Title
Exposure, Vulnerability and Adaptation to Heat and Wildfire in the Southwestern United States

Author
Rosenthal, Noam

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Exposure, Vulnerability and Adaptation to Heat and Wildfire in the Southwestern United States

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Environment and Sustainability

by

Noam Rosenthal

2023

ABSTRACT OF THE DISSERTATION

Exposure, Vulnerability and Adaptation to Heat and Wildfire in the Southwestern United States

by

Noam Rosenthal

Doctor of Philosophy in Environment and Sustainability

University of California, Los Angeles, 2023

Professor Alan Irwin Barreca, Co-Chair

Professor Miriam Elizabeth Marlier, Co-Chair

The dissertation presents three papers examining exposure to extreme heat and wildfire in the Western United States. In the first chapter, I develop a framework for analyzing transit passenger exposure to extreme heat in Maricopa County and then implement an optimization algorithm for minimizing wait times through the reallocation of buses across the transit network. In simulating the reconfiguration of buses, I find the potential for small adjustments to produce large reductions in wait time for vulnerable populations. This work also formulates a way to measure passenger vulnerability with an activity-based model that accounts for the distinct demographics of transit riders.

In the second chapter, I study the prevalence of ground level wildfire smoke, specifically particulate matter 2.5μm in diameter, in California during the 2020 wildfire season - the most severe wildfire season ever recorded by the state. For the first time, I study how frequently extreme smoke levels at surface level coincide with extreme heat in space and time. These

interactions can influence adaptive behaviors and studies show evidence of increased hospitalizations when these hazards co-occur. I find that a majority of Californians experienced at least one day of concurrent heat and smoke in 2020 and that these events were concentrated in more rural areas of the State. This case study motivates the integration of multi-hazard frameworks in both public and private sector risk planning.

In the final chapter, I examine wildfire risk factors for residential property in California. I leverage a dataset collected by CAL FIRE enumerators who record the features of a home and categorize the level of damage after every named incident. I enhance this dataset using remotely sensed detections of wildfire to impute the date when a home burned from which I then estimate time-varying weather risk factors like humidity, temperature and wind as well as fire intensity. I then use these features to train a predictive model to be used by homeowners or insurance carriers to better estimate the vulnerability of their property.

The dissertation of Noam Rosenthal is approved.

Mikhail Chester

Adam S. Millard-Ball

Miriam Elizabeth Marlier, Co-Chair

Alan Irwin Barreca, Co-Chair

University of California, Los Angeles

2023

DEDICATION

This dissertation is dedicated to my family and my grandfather Shulim, a Holocaust survivor, who was denied the opportunities of formal education. My accomplishments are owed to his hard work and sacrifice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ACKNOWLEDGEMENTS

For the first chapter, I must thank Prof. Mikhail Chester for helping conceive the project and supplying the Maricopa Activity Based Model data; Andrew Fraser for sharing his literature review; as well as David Hondula and David Eisenman for their feedback on vulnerability frameworks and benchmarks for heat safety. Additional thanks goes to Marty Wachs for reviewing manuscripts, his mentorship and signature kindness. For Chapter 2, I would like to thank my co-authors Miriam Marlier, Tarik Benmarhnia, Ravan Ahmadov and Eric James. Prof. Marlier provided indispensable mentorship throughout the project, Prof. Benmarhnia reviewed manuscripts and Dr. Ahmadov and Dr. James generously shared the archived 2020 HRRR-Smoke forecasts. I'd also like to thank Nick Clinton and Noel Gorelick on the Google Earth Engine team for their financial and technical support. Additional thanks to the LA Urban Center for supporting me financially and Dr. Francisco Escobedo for his mentorship. I owe a great deal to Prof. Adam Millard Ball who provided feedback on many manuscripts in Chapter 3 and who supported me at the beginning of my PhD journey as a prospective applicant. Additional thanks to my colleagues at Kettle, including Dr. Yaling Liu, Faran Sikandar, Dr. Maxime Dion and Son Le who provided engineering support. I couldn't have reached this milestone without the friendship of Will Krantz, Jonathan York, Edith de Guzman, Frida Winkelmann and Miriam Pinski. Last but not least, I'd like to thank Alan Barreca, my co-chair, for his indispensable guidance and support throughout all three chapters. Alan provided feedback on the writing and methodologies, and most importantly, taught me how to think critically and causally.

VITA

2015, B.S. Atmosphere/Energy, Stanford University

2018, M.S. Earth Systems, Stanford University

PUBLICATIONS

Rosenthal, Noam, Tarik Benmarhnia, Ravan Ahmadov, Eric James, and Miriam E Marlier. "Population Co-Exposure to Extreme Heat and Wildfire Smoke Pollution in California during 2020." Environmental Research: Climate, 2022. https://doi.org/10.1088/2752-5295/ac860e.

Rosenthal, Noam, Mikhail Chester, Andrew Fraser, David M. Hondula, and David P. Eisenman. "Adaptive Transit Scheduling to Reduce Rider Vulnerability during Heatwaves." Sustainable and Resilient Infrastructure 0, no. 0 (February 23, 2022): 1–12. https://doi.org/10.1080/23789689.2022.2029324.

CHAPTER ONE: Introduction

This dissertation presents new frameworks for climate risk management. The first chapter challenges conventional hardening of physical assets to withstand more extreme stressors through capital improvements. For example, instead of relying solely on shade structures or urban forestry to reduce heat exposure of bus passengers, I demonstrate how the revised operation of existing assets can realize similar benefits. Operational changes like the one I simulate offer flexible responses to new conditions and constraints, what is referred to as "adaptive resilience". This framework differs from otherwise fail-safe strategies in which infrastructure is designed to withstand stresses for the most extreme, plausible scenario. These designs are considered "brittle" in that there are no contingencies if they fail. Furthermore, fail-safe infrastructure often relies on physical reinforcement that risks obsolescence when the underlying dynamics of the earth system - such as temperature projections - change. In turn, large amounts of capital and time are needed to reverse course, rendering people and physical assets vulnerable. In contrast, operational changes such as the changes to bus routing and frequencies that I examine require less capital and time to implement. They are also amenable to changes in technology or human behavior - for example, shifts to remote work or budget cuts.

Chapter one also demonstrates how high-resolution traffic data enables more equitable policy and planning. Specifically, my use of an activity based model where individual trips are modeled at the person-level with demographic details for each "passenger" allowed me to consider the age, car-dependency and income of ridership when modeling vulnerability. This improves on the standard application of resident population metrics to represent commuters who travel *through* an area and whose sociodemographics deviate from the population average. The advent of activity-based models augurs significant improvements in equitable city planning.

The results from my second chapter pose a new question to risk planners: can someone's adaptive behavior to one climate stressor inadvertently increase their exposure or vulnerability to another? It is a particularly relevant question when two hazards affect the same location at the same time. To illustrate, if someone seeks shelter from extreme heat in their non-air-conditioned apartment by going to the beach, may they be exposing themselves to worse air quality? Interactions may also exist physiologically - for example, increases in blood pressure stemming from heat may cause respiratory sensitivity for someone inhaling wildfire particulate matter. Physiological interactions, while not a part of the dissertation, are examined in a follow-on study alongside researchers at University of California, San Diego.

My final chapter specifies a model for insurers to estimate the likelihood of a wildfire damaging a building. Whereas existing studies quantify damage risk associated with topography, vegetation canopy and building construction, they overlook the effects of wildfire intensity, temperature, wind and dryness. Yet weather and fire intensity influence fire ember production, firefighter behavior and the combustibility of a structure. Unlike building modifications that rely on error-prone human data collection, fire intensity can be directly observed for historical events and used to train empirical models. Fire intensity can also be integrated with forecasting models to explore warming's impact on the risk of building damage from wildfires.

Notably, this model has practical implications for insurance availability and affordability in California. Empowering insurers to quantify avoided losses may enable discounts on insurance premiums as an incentive for private risk mitigation as well as more accurate underwriting. An efficient insurance marketplace helps homeowners protect their wealth and shelter while transmitting price signals to preemptively discourage high-risk land use.

Future Adaptation

This dissertation highlights the risks of engineering infrastructure to withstand "worst-case" scenarios when the intensity, frequency and duration of climate hazards are uncertain. I quantify heat-smoke co-occurrences to illustrate one source of such uncertainty. Through my simulation of fleet reallocation across Phoenix's Valley Metro bus network, I demonstrate an adaptation strategy that can operate under uncertainty by responding to changes in conditions (e.g., temperature) and resources (e.g., buses). I also study the predictability of wildfire damage, recognizing insurance as a critical tool for "adaptive resilience" because it allows for infrastructure failure (e.g. building destruction) while enabling physical assets to bounce back to a more resilient state (e.g. fire-resistant construction in a lower-risk location).

The characteristics of exposure and vulnerability that I document point to key considerations in adaptation planning. First, where possible, policymakers should design infrastructure to render multi-functional benefits that can scale the returns of any investment. For example, cooling shelters that protect people from heat should be equipped with HEPA air filters that can remove pollutants since I find that heatwaves are often accompanied by reductions in air quality. Additionally, transit systems may provision face masks, water or hazard information as a service to passengers who are more vulnerable and exposed than "typical" residents due to their lower incomes and tendency to walk to stations.

Second, public and private stakeholders must account for interactions between simultaneous events when forecasting climate impacts. For example, evidence of increased morbidity and mortality due to heat-smoke interactions can help inform hospital staffing to ensure the availability of care. Public notification systems informing the public of health risks

due to smoke or heat should also caution against behaviors with spillover effects. To wit, turning on a window-AC unit during a heat wave will exacerbate smoke exposure.

Finally, democratizing information on climate risk can support private actions to mitigate vulnerability and promote adaptive capacity. My development of a model that predicts building damage from wildfires enables homeowners to take actions irrespective of federal investments in prescribed burning or mechanical thinning. Empowered homeowners can also change their behavior and implement new technologies more quickly and nimbly than systems that rely on institutional decision making. They may also incorporate local knowledge in ways that improve the efficiency and efficacy compared to public programs.

It's also worth noting that the above recommendations are embedded in systems of political and community governance. In the context of insurance, the application of any building vulnerability model will be subject to the state's department of insurance, which under Proposition 103 has final authority over insurance premium rates that a carrier can charge. Officials who are elected by homeowners have few incentives to rule in an insurance company's favor, even if the company proposes actuarially sound rates.

Interdisciplinary Challenges in Measuring Vulnerability and Exposure

My three chapters also highlight the importance and challenge of considering social, ecological and technological systems when quantifying exposure and vulnerability to extreme heat and wildfire. In the first chapter of this dissertation I focus on the potential for technology to reduce exposure by intelligently deploying buses in a transit network. Then, using an activity-based model, I incorporate social factors like income, car ownership and age to capture vulnerability. However, there are many other ecological and social variables that could improve

my estimates of exposure and vulnerability, respectively. For example, mean radiant temperature would better approximate thermal comfort than the regular temperature values I used. Additionally quantifying the prevalence of air-conditioning in one's home or workplace would help contextualize the contribution of transportation to one's cumulative daily exposure. Fortunately, improvements in synthetic datasets and physical modeling of weather are enabling more comprehensive exposure and vulnerability assessments in these contexts.

In the second chapter, I map the frequency, intensity and duration of extreme heat and extreme smoke individually and jointly. Notably, my measurement of exposure in this chapter accounts for ecological dynamics alone. Lacking additional data on social variables such as one's awareness, work environment or housing type as well as data on technological interventions like air filters and air quality warning systems, an individual's exposure cannot be precisely measured. Future improvements in the modeling of particulate emission and transport, including influences of the built environment and ownership of personal air filters, are critical for better measuring human exposure to extreme wildfire smoke. Additionally, efforts to quantify wildfire impacts in the future must account for hazard interdependencies; wildfire risk cannot just be measured in acres burned, researchers must also quantify the number of subsequent landslides, the amount of particulates emitted and, as I highlight, the intensity of heat that co-occurs.

My final chapter focuses on the vulnerability of buildings to wildfire. To date, researchers have focused exclusively on technological aspects like building characteristics and a subset of ecological factors, like topography and canopy cover, to identify drivers of risk. I expand on these efforts, including fire intensity and weather variables. These factors improve our understanding of the ecological drivers of risk and indirectly reveal social drivers because of the influence of wildfire spread rate on firefighter behavior. Still, information on the exact locations

of firefighting crews, hardening improvements in the form of vent meshes as well as data on homeowner occupancy, would all enhance the performance of my model. Additionally, higher precision weather models that can account for wildfire feedback on wind speed and precipitation would also improve my results.

To summarize, this dissertation reveals three challenges that should inform future research. First, is the generation of data on human behavior and the indoor/outdoor environments they navigate. These data are critical for true estimates of exposure and adaptive behaviors. Smartphone tracking with appropriate privacy considerations, an expansion of existing transit and time use surveys and higher resolution information about the built environment can help bridge this gap. Second, is the complete, accurate and precise reproduction of natural hazard events. Better resolved, spatially contiguous data on heat and smoke, reported at the same spatiotemporal resolution with which human behavior and demographics vary can unlock natural experiments to elucidate human responses to environmental stressors. Finally, there is a need for social and political scientists to help translate new insights into effective policy, urban planning and ecosystem management. Despite the potential for rerouting to reduce passenger heat exposure or defensible space to reduce a home's combustion likelihood, unions and agencies that represent bus drivers or municipalities that enforce brush clearance may have contravening interests, limited resources or lack the jurisdiction to unilaterally implement change. Without strong governance, even the best ideas will lie dormant.

Final Remarks

Accelerated warming portends rapid increases in the frequency and magnitude of wildfire smoke and extreme heat events. Meanwhile cities in the Southwestern United States are among

the fastest growing communities in the country. Their enduring appeal is a testament to the impressive technologies humans have crafted to stay cool in the desert sun, to insure one another when disaster strikes and to filter the air inside one's home as wildfires rage. As these hazards intensify and become less predictable, society will need to respond in-kind with continued technical ingenuity, deeper understanding of natural systems and social structures to ensure that these communities can continue to coexist with the singular environments that define the American Southwest.

CHAPTER TWO: Adaptive transit scheduling to reduce rider vulnerability during heatwaves

ABSTRACT

Extreme heat events induced by climate change present a growing risk to transit passenger comfort and health. To reduce exposure, agencies may consider changes to schedules that reduce headways on heavily trafficked bus routes serving vulnerable populations. This paper develops a schedule optimization model to minimize heat exposure and applies it to local bus services in Phoenix, Arizona, using agent-based simulation to inform travel demand and rider characteristics. Rerouting as little as 10% of a fleet is found to reduce network-wide exposure by as much as 35% when operating at maximum fleet capacity. Outcome improvements are notably characterized by diminishing returns, owing to skewed ridership and the inverse relationship between fleet size and passenger wait time. Access to spare vehicles can also ensure significant reductions in exposure, especially under the most extreme temperatures. Rerouting, therefore, presents a low-cost, adaptable resilience strategy to protect riders from extreme heat exposure.

INTRODUCTION

It is widely believed that the sustainable growth of cities relies on increasing public transit use (Dulal, Brodnig & Onoriose 2011; Hodges 2010). However, increasing frequency, duration, and severity of heat waves and other extreme weather events caused by climate change threaten both engineered infrastructure that supports public transit as well as passenger comfort and health. Phoenix, Arizona, with average maximum daytime temperatures of 111°F during summer months, and the fastest growing population in the United States, is particularly vulnerable to extreme heat (Chow et al. 2012; Census Bureau 2020; National Oceanic and Atmospheric Administration 2020). As climate change advances, the frequency of extreme heat

events in Phoenix is projected to increase 14-fold from 2.0 to 24.4 annual events by the year 2070 (Grossman-Clarke et al. 2010). Importantly, lower-income communities that have less residential air conditioning, fewer cooling centers, and streetscapes lacking vegetation, are more susceptible to the impacts from extreme heat (Chow et al. 2012; Voelkel et al. 2018). Households residing in these areas are also more dependent on public transportation and may be disproportionately exposed to heat during travel (Taylor & Fink 2003).

Existing studies show that seasonal changes in temperature and weather push would-be transit users to different modes of travel and may result in the delay or cancellation of non-essential trips (Liu, Susilo, & Karlström 2017). The drivers of these behavior changes are numerous and complex. For one, the discomfort of waiting for transit is exacerbated under inclement weather conditions (Guo 2007, Singhal, Kamga & Yazici 2014). Additionally, the journey to the bus, usually made by foot, can be physically taxing during high temperatures, especially for the elderly and persons with a disability or a chronic illness. These responses are furthermore shaped by the culture, climate (average weather) and built environment of an area (Böcker, Dijst & Prillwitz 2013; Dijst, Böcker & Kwan 2013; Liu, Susilo & Karlström 2017).

Transit agencies are already equipped with a range of tools and actions to shield passengers during periods of unpleasant and potentially dangerous weather. These include investments in shading structures and tree cover in the surrounding areas of a bus stop (Lanza & Durand 2021). Air conditioning on vehicles and in nearby retail can also be critical for reducing exposure. Physical infrastructure investments are not the only tool available to agencies, however.

Adjusting transit schedules to reduce waiting times can lower the duration of exposure without needing additional engineered infrastructure and can be implemented quickly. When opting to adapt transit schedules, agencies may avoid the expenses of large capital investments, help maintain ridership levels, as well as limit health risks. Understanding the potential benefits of reallocation as a means of adapting to extreme weather and how disruptive such schemes would be to current fleet assignments could help agencies make informed decisions for their customers. This study develops and implements a rerouting model for minimizing passenger heat exposure for the regional transportation agency in Phoenix, Arizona.

Heat, transit and human health

Despite the extensive efforts to understand the effects of weather on travel behavior, the negative impact of heat on ridership has not been sufficiently integrated into existing transport planning processes (Liu et al. 2017). Moreover, these studies neglect the potential effects heat may have on riders who continue to use public transit (Dzyuban et al. 2021). There are likely many reasons individuals continue to travel during dangerous heat conditions including, but not limited to, modal captivity, weather indifference, and inelastic travel demand with respect to weather. Mode captives refers to people who have no other mobility options available to them and weather indifference refers to the subjective experience of adverse weather conditions and, in this context, one's willingness to endure certain temperatures for a sustained period (Jacques, Manaugh & El-Geneidy 2013). Such willingness is likely informed by a traveler's opportunity cost from canceling or postponing work travel, namely, their elasticity of demand (Liu, Susilo & Karlström 2015).

Exposure to heat extremes accounts for more weather-related fatalities than nearly all other extreme weather events combined and is a leading cause of weather-related deaths in the U.S. most years (Hyland 2016; Berko et al. 2014). Heat stroke and the aggravation of existing medical conditions from heat exposure are common causes of emergency room visits, hospitalizations, and early mortality (Kenney, Craighead & Alexander 2014; Kovats & Hajat 2008; Michelozzi et al. 2009). Globally, exposure to extreme heat was estimated to result in 480,000 excess deaths per year (Zhao et al. 2021).

Epidemiologic studies have identified demographic, economic, and community characteristics that are associated with increased mortality and morbidity during periods of extreme heat (Reid et al., 2009). Determinants of increased exposure and sensitivity to heat include: lower income, older age, higher population density, lower tree density, outdated construction, and lack of air conditioning (Aminipouri, Knudby & Ho 2016). Research has shown that these characteristics cluster spatially in urban areas where transit ridership is concentrated and where discriminatory redlining has resulted in historical underinvestment in infrastructure (Harlan et al. 2013; Hoffman, Shandas & Pendleton 2020; Reid et al. 2009).

The potential for prolonged exposure to extreme temperatures for transit riders is far greater than for drivers. In most cases transit use requires riders to expose themselves to the environment in three phases: ingress, waiting, and egress. Ingress and egress exposure are a function of the mode used and distance traveled to access transit stops and final destinations. Seeing as more than 75% of all transit riders walk to transit, the location of the nearest stop relative to a person's origin is a critical factor in transit related exposure (Hess 2012). By contrast, waiting based exposure depends on the frequency of individual transit lines, their reliability, their capacity, and the physical characteristics of a station (Fraser & Chester 2017).

Early research around heat exposure and transit use combines simulated urban meteorology with transportation activity diaries to assess outdoor heat exposure during non-motorized travel, including access trips to transit stops. In such studies, researchers find that socially disadvantaged groups are disproportionately exposed to transport-heat (Karner, Hondula & Vanos 2015). Additional research shows that transit stop location and transit schedules contributed to variable heat exposure across transit systems and that users from areas with low density, few high capacity roadways, and irregular street networks are more likely to experience prolonged exposure via transit access and waiting at transit stops (Fraser & Chester 2017).

Adjusting bus transit schedules during heat waves

Bus transit network design and scheduling is a complex process that balances service quality, coverage, and directness. In most cases, transportation agencies are not profit driven but their resources and operations remain constrained by available budgets (Desaulniers & Hickman 2007). The primary competing alternatives are transit systems that offer a large service area with limited frequency and those with smaller service areas but high frequency. The public transit planning process is typically divided into five steps, (1) network design (route structure and stop placement), (2) route frequencies, (3) timetabling, (4) vehicle scheduling, and (5) crew scheduling and rostering (Guihaire & Hao 2008). This analysis focuses on step two, specifically increasing transit frequencies to reduce waiting times for passengers.

Vehicle arrivals per hour along a given route is the most important factor affecting overall wait time. Traditionally, average waiting times at any stop have been estimated as half the “headway”, or the time between bus arrivals; on-board surveys, however, contradict this assumed uniform arrival distribution (Fraser & Chester 2017). Along infrequent routes,

passenger wait times are significantly less than half the headway indicating rider knowledge of existing transit schedules. Increasing frequencies along such routes without providing advanced notice to riders may not significantly reduce average waiting due to the passenger arrival behavior. Conversely, average wait times along frequent routes are typically greater than the times predicted by transit schedules. Thus, adding vehicles to routes with intermediate frequencies may represent the best opportunity to significantly reduce passenger waiting times.

Transit frequencies are usually developed from demand estimates and agency standards for vehicle occupancy and minimum frequency (Ceder 2016). Such estimates are typically derived from travel demand models that draw from an area's economic activity and population to produce top-down projections of trip generation and mode choice. Yet these models are incapable of "chaining" trips together nor do they attribute trips to specific households, limiting the ability for planners to fully measure the equity-impacts of bus scheduling. To address these and other shortcomings, agencies have, in recent years, begun implementing advanced activity-based models (ABM); bottom-up estimations of travel demand that are generated from household attributes and an individual's anticipated behavior in areas beyond their home and across different times of day (Hafezi et al. 2018).

There are well-established models for determining transit frequencies to optimize economic and efficiency outcomes (Hadas & Shnaiderman 2012). Increased calls to consider equity in transit service, coupled with an ABM's provision of disaggregated sociodemographic data at fine spatial resolutions, however, warrants the addition of new optimization criteria in planning models, not least, ridership's exposure to extreme heat. This could either be accomplished by adding vehicles to service, as agencies already do for special events, or reallocating existing in-service vehicles from other routes. This paper develops an optimization

framework that reallocates existing vehicles and dispatches spare fleet capacities based on the heat vulnerability of riders and explores that framework using a case study of Valley Metro, the transit agency serving Phoenix, AZ.

METHODS

Optimization framework

To minimize negative outcomes from extreme heat, an agency should maximize service to the areas that are most likely to rely on transit and most likely to suffer from prolonged exposure to high temperatures. In other words, an agency should minimize the combination of heat exposure and heat sensitivity, what the model hereinafter refers to as vulnerability. Individual exposure is assumed to be negatively correlated with income and cars per household, as wealthier households will retain the option of driving or foregoing travel completely (He & Thøgersen 2017). Regional vegetation abundance is also included as it offers cooling and shading for pedestrians (Lanza & Durand 2021). Heat sensitivity, on the other hand, is modeled by passenger age, which studies have identified as the primary risk factor for susceptibility to heat stress during physical activity (McGinn et al. 2017). The final component of passenger vulnerability in this model is the duration of exposure. Because agencies independently determine the frequency of bus arrivals, the optimization model described herein solves for the bus frequencies that will minimize wait times. Simply stated, the model aims to reduce wait times by as much as possible, for as many people as possible while accounting for each passenger's sensitivity to heat and their dependency on travel. It uses a non-linear constrained optimization solver for the entire bus service area of Phoenix's transit agency. The model is specified as:

Equation 2-1: Objective Function

$$Min \sum_{i} \sum_{j \in C_i} W_i(f_i) D_j V_j S.T. \quad 1) \sum_{i} f_i < B \quad 2)\, M \leq f_i \leq N, \ \in Z \ \forall i \, Where: i$$

$$= Transit\ route\ \&\ j = Microanalysis\ zone W_i(f_i)$$
$$= Average\ waiting\ time\ for\ route\ i \left(\frac{runtime}{vehicles}\right) C_i$$
$$= Microanalysis\ zones\ served\ by\ transit\ route\ i D_j$$
$$= Transit\ demand\ in\ microanalysis\ zone\ j\ (\#\ of\ Transit\ Riders) V_j$$
$$= Heat\ vulnerability\ of\ microanalysis\ zone\ j f_i$$
$$= Frequency\ of\ route\ i \left(\frac{vehicles}{runtime}\right) B$$
$$= The\ total\ number\ of\ buses\ currently\ operating M$$
$$= NIOSH\ Minimum\ allowable\ frequency \left(\frac{vehicles}{runtime}\right) N$$
$$= Maximum\ allowable\ frequency \left(\frac{vehicles}{runtime}\right)$$

The model assumes exposure to primarily be a function of wait time $W_i(f_i)$, demand D_j, and vulnerability V_j, subject to two main constraints. The first constraint, B, represents the total agency fleet size, which provides an upper bound on the total number of buses that can be assigned across all routes. While total fleet size is typically a constant, Valley Metro, like many agencies, has the .capacity to dispatch spare vehicles to increase capacity. To observe the sensitivity of outcomes to the fleet size, B, the model is run for five different capacity multiples ranging from 0% to 20% increases in normal fleet capacity; the latter is the official spare fleet capacity reported in Valley Metro's 2020-2024 Short Range Transit Program and is also the maximum spare fleet size allowed by the Federal Transit Administration (Valley Metro 2019).

The second constraint, N, ensures that the number of vehicles servicing a route produces wait times that are below the National Institute of Occupational Safety and Health's (NIOSH) heat exposure duration standards and above an impractical lower limit of five minutes. NIOSH's standards were developed in 2016, with the Centers for Disease Control and Prevention, to

inform employers about heat safety standards. These are based on the Wet Bulb Globe Temperature, a holistic measure of "experienced" heat, the metabolic rate of an activity, and the availability of engineering controls to reduce heat stress (e.g. air conditioning, shade). In the absence of alleviating heat stress, the institute recommends "administrative controls", or more simply stated, rest periods to allow for the body to cool. These guidelines are drafted separately for "heavy", "medium", and "light" forms of work. The administrative control guidelines for "light" work are used to parametrize maximum wait times in the above model for five different extreme temperature scenarios – 106°F, 107°F, 108°F, 109°F and 110°F. Finally, vehicle bunching, a reliability issue caused by excess vehicles, is known to occur on high frequency routes in high demand areas during peak periods (Camps & Romeu 2016). Adding vehicles to routes already experiencing bunching may exacerbate this problem. Accordingly, the minimum allowable headway in the model, or maximum allowable frequency N, is set to correspond to five minutes.

To assess the optimization's sensitivity to the two imposed constraints, the model was run for each possible combination of fleet size, B, and maximum headway, N, equaling 25 scenarios total. The optimization problem was solved in MATLAB using the 'fmincon' function which implements an interior point algorithm to find a globally optimal fleet allocation. The model produces a non-integer value for bus vehicle allocation and therefore can be interpreted as either a theoretical representation of service capacity or buses servicing only a segment of the route.

Transit schedule and data

The geography, service frequencies, and current fleet allocations of individual routes are derived from the 2020 General Transit Feed Specification data for Phoenix's regional transit

agency, Valley Metro (Valley Metro 2020). The number of vehicles needed to produce a certain headway on a given route, f_i, is estimated by dividing the time it takes to complete a route, its runtime, by the route's average headway. Conversely, the wait time for a given route can be determined by dividing the runtime by the route's fleet allocation. For example, a route that on average takes half-an-hour to complete with buses arriving every 10 minutes, would require three buses. Given that headways and runtimes are not equal across all times of day nor all days of the week but rather fluctuate in response to demand across different stops at different times, estimates are based on the modal headway of weekday service.

Transit demand

Transit demand D_j is determined using the output of Maricopa Association of Governments (MAG) ABM in 2018 (Maricopa Association of Governments 2018a). Activity-based travel demand models capture household-level and person-level travel choices including intra-household interactions between household members across a wide range of activity and travel dimensions (Parsons Brinckerhoff Inc & Arizona State University 2010). The ABM used in this simulation is informed, in large part, by a 2017 household travel survey conducted by MAG that includes GPS activity-travel data from 6,073 surveyed households, as well as data from the American Community Survey (Maricopa Association of Governments 2018b). Importantly, transit trips in the ABM do not indicate the specific mode of travel (e.g. light-rail, bus); rather they are categorized between premium and conventional transit accessed by walking, kiss-and-ride, and park-and-ride. Premium transit traditionally includes express buses, bus rapid transit, light rail transit, and commuter rail whereas conventional transit typically refers to regularly scheduled local services. This model simulates heat exposure for all premium and conventional transit trips.

The ABM output details the daily travel movements of a simulated population of 3.8 million agents and 18.4 million daily trips. Only a small fraction (~150,000) of the modeled daily trips in MAGs region occur by transit. In addition to the location of transit demand the model output also allows one to isolate the transit demand by time of day. The general movement of agents reflects a pattern of leaving residential areas during the morning peak period (6-9am) for employment locations and the opposite pattern during the afternoon/evening peak (3-7pm).

At its most resolved spatial scale, the ABM identifies the origin and destination microanalysis zones (MAZ) – the smallest transportation spatial unit used by the planning agency – for each transit trip. The open-source routing software Open Trip Planner (OTP) is used to translate origin-destination pairings –imputed to be the centroid of each respective MAZ - into a specific transit bus line. OTP is a graph-based multimodal routing system that operates on a unified graph including links representing road, pedestrian, and transit facilities and services (Hillsman & Barbeau 2011). Focusing on all MAZ with sizes below the 98th percentile, the trip-weighted median MAZ area is 0.09km^2 (μ = 0.38 km^2), such that the "true" origin point of a trip within any single MAZ does not differ much from the centroid. Variance in size across the 8662 unique MAZs that were routed by OTP, though larger, remains small in absolute terms with a weighted standard deviation of 0.9 km^2. As MAZs are population weighted, the MAZs with the largest area and greatest potential error, account for only a small share of passenger trips.

OTP routing assumes a maximum walking distance of one mile to reach a transit stop and accounts for all possible transfers. All routing requests were made at the nearest hour of departure based on weekday service for the agency in February of 2020. OTP routing produced transit routes for 99.5% of trips across 72 bus lines. Importantly, for the optimization model,

non-local and non-bus transit lines are excluded as they tend to be low-frequency, pre-scheduled commuter services, with predictable wait times.

Passenger vulnerability

Area population heat vulnerability indices V_M were developed for each MAZ in Maricopa based on the ABM-reported characteristics of each transit rider who begins a trip from that MAZ as well as the area's vegetation abundance. The formulation of vulnerability was adapted from relevant literature to develop an individual-scale, transit-specific metric. Specifically, standardized scores for income, cars per household, and age were combined with equal weight, given their correlations with transit dependency and health risk (Taylor & Fink 2003). The inclusion of transit dependency assumes that under extreme heat conditions there will be a decline in transit usage, specifically among passengers with alternative private travel options. Accordingly, this model prioritizes servicing areas with the greatest number of residents who lack such options. Finally, to estimate heat vulnerability from the physical environment, the normalized difference vegetation index (NDVI) was estimated for each MAZ by computing the median NDVI 30 meter pixel value from 2020 July and August LANDSAT 8 imagery in the one mile area surrounding the MAZ centroid. This area matches the maximum walking shed allowed by the routing algorithm for any single traveler. All social and physical variables mentioned were min-max normalized and added with equal weight to produce a vulnerability index. The product of the vulnerability index and the ridership demand for a given bus route can be interpreted as the weights that drive the prioritization of fleet allocation in the model.

Equation 2-2. Area population vulnerability index.

$$V_j = \sum_{i \in A_i} N[Z(I_i) + Z(C_i) + Z(G_i)] + N(NDVI_j)$$

$$Where: i = agent\ and\ A_i = set\ of\ agents\ with\ trips\ originating\ in\ MAZ\ j$$

$$Z(x) = \frac{x - \mu(x)}{\sigma(x)}\ where\ \mu = mean\ and\ \sigma = standard\ deviation$$

$$N(x) = \frac{x - min(x)}{max(x) - min(x)}$$

$$I_i = Individual's\ Household\ Income$$

$$C_i = Cars\ per\ household\ of\ individual \quad G_i = Age\ of\ individual$$

$NDVI_j = Normalized\ Difference\ Vegetation\ Index$ surrounding centroid of MAZ j

Our approach marks a noteworthy advancement in precision compared to current vulnerability estimations by using ABM household and person characteristics. That is, while traditional ridership characteristics are imputed from aggregated census data for *residents* in a spatial unit, our vulnerability estimates are based on *traveler* characteristics and, in turn, capture a more accurate sample that includes non-home-originating trips.

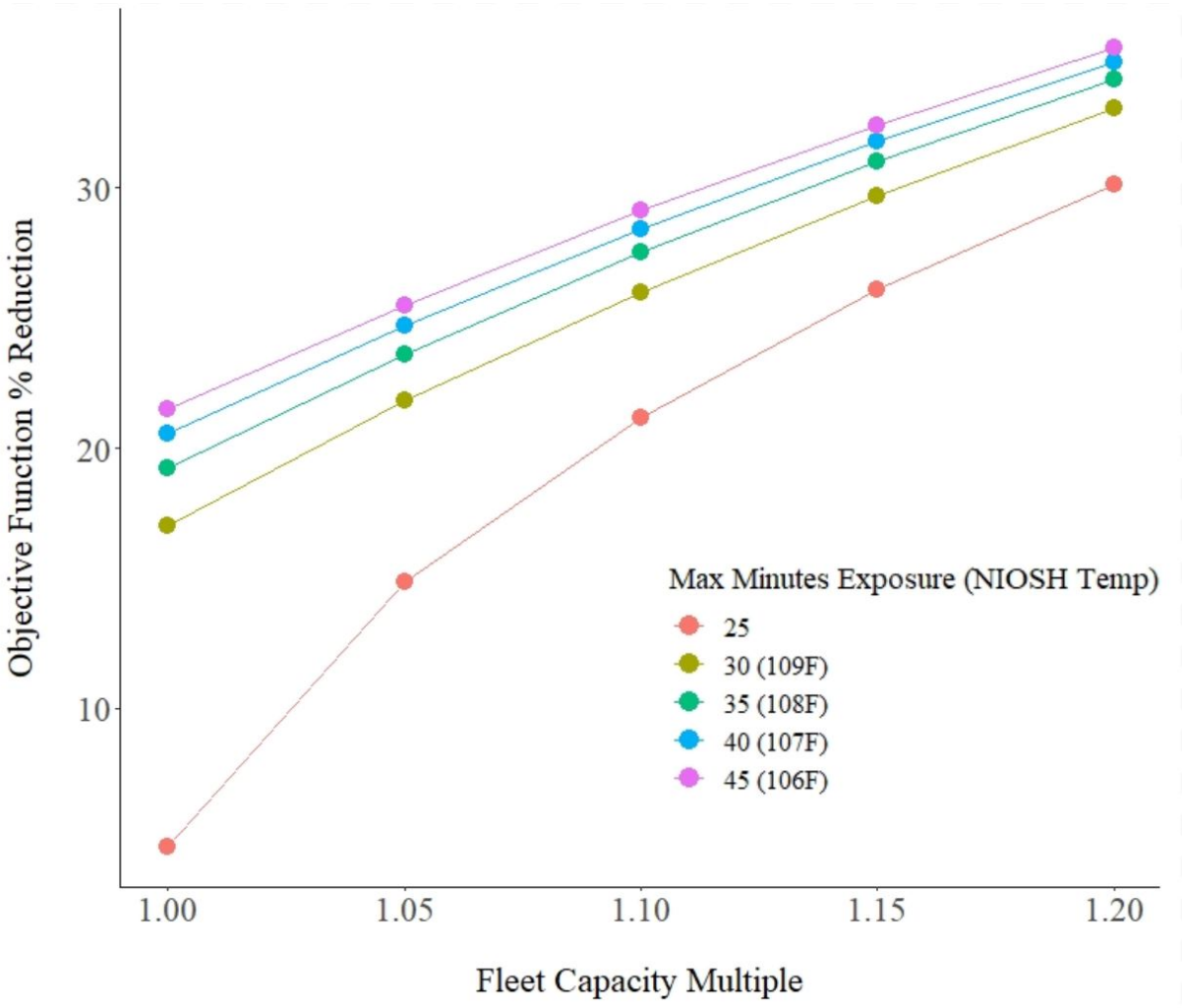
RESULTS

The ABM model counts 115,129 transit trips on an average weekday between the hours of 7 am and 6 pm, representing 0.79% of all trips completed during daytime hours, when the combination of sunlight and high air temperatures can be hazardous. According to the ABM, transit riders in Maricopa County are overall 11 years younger, with a median age of 25, and hail from households with median incomes of $56,500, which is approximately 15% less than their

non-transit counterparts. Cars per household for transit passengers is 1.77 compared with 2.14 for non-transit travelers.

Across all fleet capacity multiples, the mildest heat exposure of 106F, benefitted the most from rerouting, realizing reductions in the demand-headway weighted vulnerability of over 20% for the standard bus fleet capacity and up to 40% when the maximum available fleet is used. Improvements stemming from capacity increases tend to be linear (Figure 2-1), suggesting that routes served by additional bus capacity contribute equally to the objective outcome. By contrast, when keeping capacity fixed, there is a nonlinear decline in improvement as the temperature increases, i.e. maximum allowable wait time decreases. Notably, at the extreme temperature of 110°F, there are not enough buses in the agency fleet to meet the wait-time constraint of 15 minutes for all bus lines. Rather, the minimum headway that can be realized for the entire local bus network using its standard fleet is 25 minutes. It is at this temperature scenario that capacity increases produce the greatest benefit by reducing headways for the most vulnerable routes; specifically, a 5% increase in fleet size corresponds to a 15% improvement in the objective outcome.

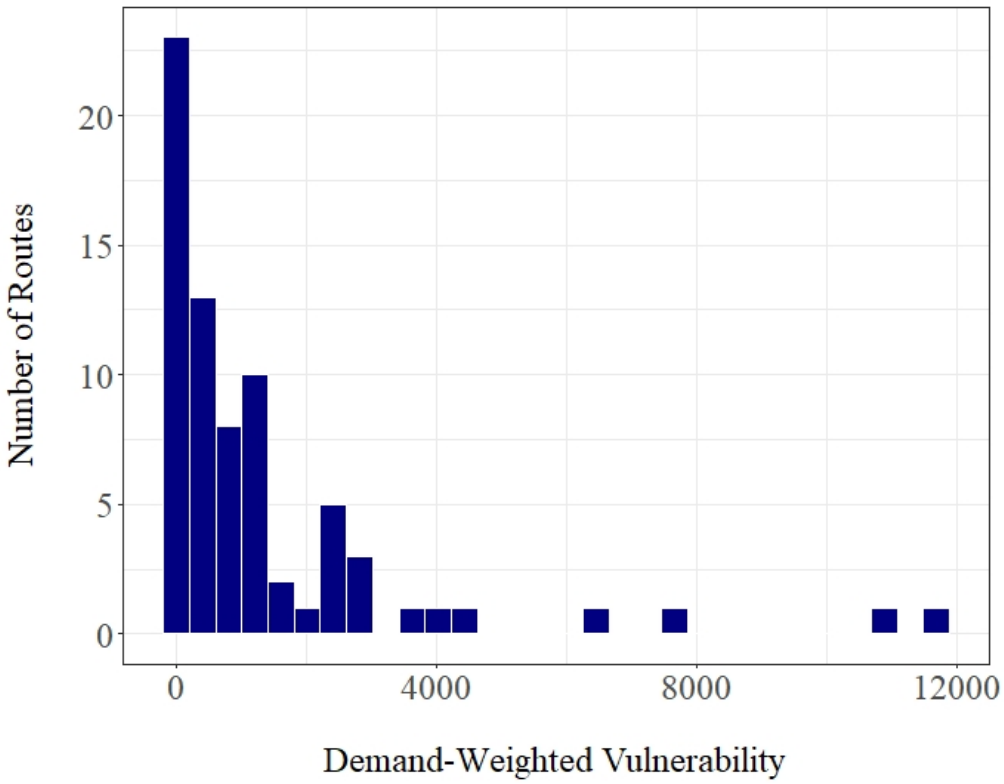
Figure 2-1. Optimization objective function improvements and fleet change magnitude. The Y-axis is the percent reduction of the optimized objective outcomes relative to the baseline schedule's unoptimized objective outcome. As the vertical spacing between points for a given capacity multiple illustrate, there is a nonlinear reduction in exposure as the maximum allowable wait time is reduced. As capacity increases, there is a linear reduction in risk – except for scenarios with most extreme heat, which exhibit non- linear improvements in the objective outcome. The minimum achievable wait time (maximum exposure duration) for all routes was 25 minutes, for which there is no official corresponding NIOSH temperature.



The observed results stem from both the uneven demand across the system and the inherent nonlinearity in headways which are inversely related to fleet size. Indeed, each incremental bus added to a route has a diminishing return on wait time equal to the inverse fleet size, squared. It follows then that a bus line with pre-existing small headways will see little gain

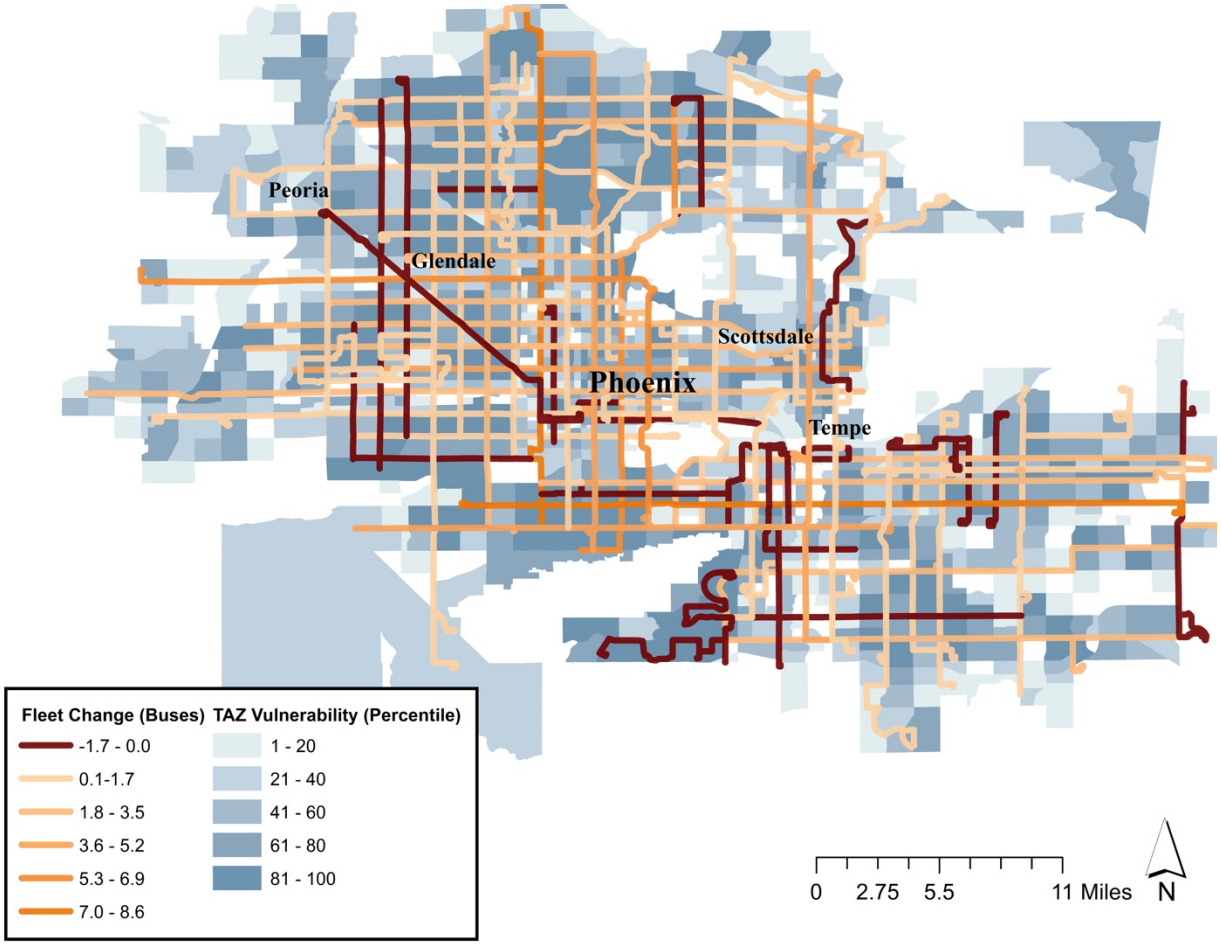
from fleet additions. By contrast, a line with large headways and moderate exposure, will see a large improvement from an increase in service. Simply stated, few lines carry most of the exposure and the model prioritizes those few lines with more buses to realize large reductions in riders' heat exposure. Any subsequent service improvements will have more moderate effects on outcomes. Spare fleet capacity makes the greatest difference under extreme temperature scenarios when the requirements of NIOSH compliance overwhelm the network.

Figure 2-2. Distribution of MAZ demand-weighted vulnerabilities. The optimization model aims to minimize the wait time for the routes with the highest demand and vulnerability. The demand- vulnerabilities for all routes are represented as weights to the variable of interest – fleet allocation – when running the optimization model. The frequency distribution of these weights represents the distribution of vulnerable demand, or the total person-minutes scaled by the vulnerability indices for all origin MAZs boarding a given bus route. The positively skewed distribution demonstrates that a small number of routes carry a disproportionate number of vulnerable passenger.



NIOSH standards interact with the skewedness of passenger demand as well. This can be seen in Figure 2-2, which shows select few outliers in the objective function coefficients, i.e. the vulnerability index multiplied by the demand. When temperatures increase, the model's requirement for meeting maximum allowable exposure time on *all* lines results in the diversion of buses from lines with high vulnerability and high ridership to lines with lower vulnerability and ridership. Furthermore, given that allowable heat exposure has nonlinear stringency for each incremental degree, there is a corresponding nonlinear reduction in objective outcomes for increasingly severe temperatures when capacity is fixed.

Figure 2-3. Map of Phoenix fleet reallocation. Background polygons correspond to transportation analysis zones (TAZ) – an MAZ's parent spatial unit. Each zone's vulnerability quintile is shaded in blue so that darker shades correspond to higher vulnerability. TAZ vulnerability is measured as the weighted average of all MAZ vulnerabilities subsumed by a TAZ and is used for visualization purposes only. Polylines correspond to the 72 local bus routes that were studied and are shaded according to the change in buses per runtime under the optimized scenario relative to baseline. Red colored lines experienced a reduction in service, with a maximum reduction of 1.7 buses per runtime. Orange shaded routes saw an increase in service with a maximum increase of 8.6 buses. For the few segments where routes overlap, the route with the highest frequency is shown.



Of practical importance is the number of buses that would need to be reallocated to achieve the optimal schedule. On average, across all scenarios, rerouting would require the reassignment of approximately 13.5% of the bus fleet (SD=2.2%). Notably, 27 of the 72 lines see on average a 0.9 *decrease* in fleet size, based on the median outcome for each bus line across all

scenarios. The remaining lines all show median increases in bus allocations compared to the existing agency schedule, with a maximum increase of 8.6 buses per hour for Line 19 which runs through Downtown Phoenix. Compared to our GTFS estimates of the agency's fleet sizes and assuming a bus capacity of 36 seats, the optimized model decreases the average load factor by 27% from 1.07 to 0.78. This is unsurprising given the model's emphasis on serving routes with large demand. Arterial lines, that traverse the downtown area as well as more densely populated neighborhoods to the east and west of the city center, are most served in the model's output (Figure 2-3). This reflects the higher population density west of Downtown Phoenix and coincident lower median family incomes. These lines also serve areas with younger residents who are less likely to own vehicles, including affiliates of Arizona State University.

DISCUSSION

The model results highlight the context dependency of schedule optimization and rerouting to promote climate resilience in hot climates. Whereas under less extreme conditions, capacity increases confer proportional benefits to the objective, under more severe temperatures - when all routes must be equally serviced and baseline capacity is constrained -surplus vehicles offer significant benefits to reducing vulnerable exposure. Importantly, adding bus capacity to a line is limited by both a practical "minimum" headway and the diminishing returns of additional buses. And finally, no amount of fleet reallocation would be able to satisfy NIOSH guidance on maximum allowable exposure for temperatures exceeding 110°F, without a dramatic increase in fleet size. That is because the minimum number of buses per route needed to fulfill the maximum allowable wait time are so large, that the total number of required buses exceeds the total fleet size. Therefore, under the most extreme heat conditions, agencies would need to pursue alternative interventions that directly reduce heat exposure for a passenger and/or avoids the

need to commute and wait for a bus. These include, but are not limited to, cooled shade structures, improved tree canopy, first-last-mile micromobility connections, and on-demand transit services.

It is important to note that as a demonstration, the model presented in this paper makes simplifying assumptions that would need to be addressed prior to any real implementation. For one, rerouting and the dispatching additional vehicles could increase an agencies' operating expenses beyond allocated budgets. Further, many agencies have designated vehicles serving routes that are determined based on travel distance, terrain, and powertrain; by contrast, our model assumes that all vehicles can serve all routes. It may also be simplistic to assume that drivers would be amenable to sudden adjustments in route assignments and schedules. From the passenger standpoint, it is also worth noting that any changes to schedules would need to be communicated effectively, especially for the select routes that would experience a reduction in service.

The simulated scenarios also bear simplifying assumptions. For one, they assume uniform temperature throughout the study area, when the local built environment is known to cause variability in microclimates that impact human thermal comfort and health (Park et al. 2017). Yet, even with complete temperature information an agency would unlikely be able to adapt service to account for such high spatial variability in temperatures. Using fixed-point weather station data, like those available from a regional weather station, seems more likely and practical. Additionally, NIOSH's "light work" exposure thresholds while informative, are not directly based on travel activity. For this reason, the exposure thresholds introduced in the model should be interpreted as rudimentary benchmarks.

Finally, there are several improvements that could be made to the model, pending data availability. The first would be to obtain empirical wait times for different routes and stops to accurately quantify exposure. Having additional information on the bus stop infrastructure and egress exposure would also enhance such measurements. Additionally, the estimates of fleet assignments in the model are based on expected wait times and runtimes derived from GTFS and as a result overlook day-to-day and hour-by-hour fluctuations in schedules. Lastly, cost simulations associated with each fleet reallocation could further demonstrate the feasibility of any adaptive scheduling in applied settings.

These challenges notwithstanding, the model presents an advancement in its use of an ABM model that offers unmatched precision in traveler behavior and characteristics. These data inform changes to service operation that might offer significant cost savings over traditional hard infrastructure. They can be implemented relatively quickly and they can be easily adapted to unanticipated changes in infrastructure, demand, and weather. Future areas of research that could advance these models include enabling the partitioning of routes, whereby certain segments would be served at a higher or lower frequency, as well as logistic considerations that include bus depot location and bus driver availability. It would also be worth replicating this analysis for an agency like Los Angeles Metro, which has a larger bus fleet and covers an area with more temperature variability across space and time.

CONCLUSION

Cities throughout the United States have tasked public transit systems with securing myriad social benefits that include alleviating automobile congestions, reducing carbon emissions, and providing mobility to lower-income residents. To realize these goals, cities have

invested heavily in the expansion of transit services and improving reliability. By comparison, the first-last-mile comfort of passengers prior to boarding and after alighting, especially as it relates to weather, has been overlooked. With climate change expected to increase temperatures in Phoenix and the nation, extreme heat has the potential to reverse hard earned improvements in service and safety.

This paper presents one method for protecting passenger comfort and health, leveraging the flexibility of bus systems to better serve routes with more vulnerable riders. It highlights the complex interaction between allowable heat exposure and the effect of bus capacity on wait-times – both of which follow nonlinear trends. The findings show that during milder summer heat (<110°F), agencies can achieve significant improvements with modest route adjustments and that during more severe heat events, the deployment of spare vehicles can secure large gains in passenger welfare. This is particularly true for agencies with skewed ridership, where a few lines carry most passengers. Given these findings, agencies might consider investments in building an adaptable workforce – training drivers for multiple routes and negotiating more flexible working arrangements– as well as ensuring the availability of spare vehicles before any extreme heat event. In doing so, agencies will both help protect the health and comfort of their customers as well as equitably enhance the resilience of their systems.

BIBLIOGRAPHY

Aminipouri, M., Knudby, A., Ho, H.C., 2016. Using multiple disparate data sources to map heat vulnerability: Vancouver case study. The Canadian Geographer / Le Géographe canadien 60, 356–368. https://doi.org/10.1111/cag.12282

Berko, J., Ingram, D.D., Saha, S., Parker, J.D., 2014. Deaths attributed to heat, cold, and other weather events in the United States, 2006-2010. Natl Health Stat Report 1–15.

Böcker, L., Dijst, M., Prillwitz, J., 2013. Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review. Transport Reviews 33, 71–91. https://doi.org/10.1080/01441647.2012.747114

Ceder, A., 2016. Public Transit Planning and Operation.

Chow, W., Chuang, W.-C., Gober, P., 2012. Vulnerability to Extreme Heat in Metropolitan Phoenix: Spatial, Temporal, and Demographic Dimensions. The Professional Geographer 64, 286–302. https://doi.org/10.1080/00330124.2011.600225

Desaulniers, G., Hickman, M.D., 2007. Chapter 2 Public Transit, in: Barnhart, C., Laporte, G. (Eds.), Handbooks in Operations Research and Management Science, Transportation. Elsevier, pp. 69–127. https://doi.org/10.1016/S0927-0507(06)14002-5

Dijst, M., Böcker, L., Kwan, M.-P., 2013. Exposure to weather and implications for travel behaviour: introducing empirical evidence from Europe and Canada. Journal of Transport Geography 28, 164–166. https://doi.org/10.1016/j.jtrangeo.2013.01.004

Dulal, H.B., Brodnig, G., Onoriose, C.G., 2011. Climate change mitigation in the transport sector through urban planning: A review. Habitat International 35, 494–500. https://doi.org/10.1016/j.habitatint.2011.02.001

Dzyuban, Y., Hondula, D.M., Coseo, P.J., Redman, C.L., 2021. Public transit infrastructure and heat perceptions in hot and dry climates. Int J Biometeorol. https://doi.org/10.1007/s00484-021-02074-4

Eisenman, D.P., Wilhalme, H., Tseng, C.-H., Chester, M., English, P., Pincetl, S., Fraser, A., Vangala, S., Dhaliwal, S.K., 2016. Heat Death Associations with the built environment, social vulnerability and their interactions with rising temperature. Health & Place 41, 89–99. https://doi.org/10.1016/j.healthplace.2016.08.007

Fraser, A.M., Chester, M.V., 2017. Transit system design and vulnerability of riders to heat. Journal of Transport & Health 4, 216–225. https://doi.org/10.1016/j.jth.2016.07.005

Grossman-Clarke, S., Zehnder, J.A., Loridan, T., Grimmond, C.S.B., 2010. Contribution of Land Use Changes to Near-Surface Air Temperatures during Recent Summer Extreme Heat Events in the Phoenix Metropolitan Area. Journal of Applied Meteorology and Climatology 49, 1649–1664. https://doi.org/10.1175/2010JAMC2362.1

Guo, Z., Wilson, N.H.M., Rahbee, A., 2007. Impact of Weather on Transit Ridership in Chicago, Illinois. Transportation Research Record 2034, 3–10. https://doi.org/10.3141/2034-01

Hadas, Y., 2013. Assessing public transport systems connectivity based on Google Transit data. Journal of Transport Geography 33, 105–116. https://doi.org/10.1016/j.jtrangeo.2013.09.015

Hafezi, M.H., Millward, H., Liu, L., 2018. Activity-Based Travel Demand Modeling: Progress and Possibilities, in: International Conference on Transportation and Development 2018. Presented at the International Conference on Transportation and Development 2018, American Society of Civil Engineers, Pittsburgh, Pennsylvania, pp. 138–147. https://doi.org/10.1061/9780784481561.014

Harlan, S.L., Declet-Barreto, J.H., Stefanov, W.L., Petitti, D.B., 2013. Neighborhood Effects on Heat Deaths: Social and Environmental Predictors of Vulnerability in Maricopa County, Arizona. Environ Health Perspect 121, 197–204. https://doi.org/10.1289/ehp.1104625

He, S.Y., Thøgersen, J., 2017. The impact of attitudes and perceptions on travel mode choice and car ownership in a Chinese megacity: The case of Guangzhou. Research in Transportation Economics 62, 57–67. https://doi.org/10.1016/j.retrec.2017.03.004

Hess, D.B., 2012. Walking to the bus: perceived versus actual walking distance to bus stops for older adults. Transportation 39, 247–266. https://doi.org/10.1007/s11116-011-9341-1

Hodges, T., 2010. Public Transportation's Role in Responding to Climate Change.

Hoffman, J.S., Shandas, V., Pendleton, N., 2020. The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. Climate 8, 12. https://doi.org/10.3390/cli8010012

Hyland, D., 2016. Climate Change and Extreme Heat: What You Can Do to Prepare. US EPA. https://doi.org/10.13140/RG.2.2.15996.74889

Jacklitsch, B., Musolin, K., Williams, J., Coca, A., Kim, J.-H., Turner, N., 2016. Criteria for a Recommended Standard: Occupational Exposure to Heat and Hot Environments.

Jacques, C., Manaugh, K., El-Geneidy, A.M., 2013. Rescuing the captive [mode] user: an alternative approach to transport market segmentation. Transportation 40.

Karner, A., Hondula, D.M., Vanos, J.K., 2015. Heat exposure during non-motorized travel: Implications for transportation policy under climate change. Journal of Transport & Health 2, 451–459. https://doi.org/10.1016/j.jth.2015.10.001

Kenney, W.L., Craighead, D.H., Alexander, L.M., 2014. Heat waves, aging, and human cardiovascular health. Med Sci Sports Exerc 46, 1891–1899. https://doi.org/10.1249/MSS.0000000000000325

Kovats, R.S., Hajat, S., 2008. Heat stress and public health: a critical review. Annu Rev Public Health 29, 41–55. https://doi.org/10.1146/annurev.publhealth.29.020907.090843

Lanza, K., Durand, C.P., 2021. Heat-Moderating Effects of Bus Stop Shelters and Tree Shade on Public Transport Ridership. International Journal of Environmental Research and Public Health 18, 463. https://doi.org/10.3390/ijerph18020463

Liu, C., Susilo, Y., Karlström, A., 2015. Investigating the impacts of weather variability on individual's daily activity–travel patterns: A comparison between commuters and non-commuters in Sweden. Transportation Research Part A Policy and Practice 82, 47–64. https://doi.org/10.1016/j.tra.2015.09.005

Liu, C., Susilo, Y.O., Karlström, A., 2017. Weather variability and travel behaviour – what we know and what we do not know. Transport Reviews 37, 715–741. https://doi.org/10.1080/01441647.2017.1293188

Maricopa Association of Governments, 2018a. Maricopa Association of Government Activity-Based Model 2018.

Maricopa Association of Governments, 2018b. 2017 MAG Household Travel Survey Report.

McGinn, R., Poirier, M.P., Louie, J.C., Sigal, R.J., Boulay, P., Flouris, A.D., Kenny, G.P., 2017. Increasing age is a major risk factor for susceptibility to heat stress during physical activity. Appl. Physiol. Nutr. Metab. 42, 1232–1235. https://doi.org/10.1139/apnm-2017-0322

Mension, J., Estrada Romeu, M.Á., 2016. Headway adherence: detection and reduction of the bus bunching effect. Presented at the AET papers repository, pp. 1–15.

Michelozzi, P., Accetta, G., De Sario, M., D'Ippoliti, D., Marino, C., Baccini, M., Biggeri, A., Anderson, H.R., Katsouyanni, K., Ballester, F., Bisanti, L., Cadum, E., Forsberg, B., Forastiere, F., Goodman, P.G., Hojs, A., Kirchmayer, U., Medina, S., Paldy, A., Schindler, C., Sunyer, J., Perucci, C.A., PHEWE Collaborative Group, 2009. High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. Am J Respir Crit Care Med 179, 383–389. https://doi.org/10.1164/rccm.200802-217OC

National Weather Service Corporate Image Web, n.d. NOWData - NOAA Online Weather Data [WWW Document]. NOWData - NOAA Online Weather Data. URL https://www.weather.gov/wrh/climate?wfo=psr (accessed 6.8.21).

Park, C., Ha, J., Lee, S., 2017. Association between Three-Dimensional Built Environment and Urban Air Temperature: Seasonal and Temporal Differences. Sustainability 9, 1338. https://doi.org/10.3390/su9081338

Parsons Brinckerhoff, Inc, Arizona State University, 2010. Design and Development Plan for the MAG CT-RAMP Activity-Based Model (ABM).

Reid, C.E., O'Neill, M.S., Gronlund, C.J., Brines, S.J., Brown, D.G., Diez-Roux, A.V., Schwartz, J., 2009. Mapping community determinants of heat vulnerability. Environ Health Perspect 117, 1730–1736. https://doi.org/10.1289/ehp.0900683

Singhal, A., Kamga, C., Yazici, A., 2014. Impact of weather on urban transit ridership. Transportation Research Part A: Policy and Practice 69, 379–391.

Southern and Western Regions Experienced Rapid Growth This Decade [WWW Document], n.d. URL https://www.census.gov/newsroom/press-releases/2020/south-west-fastest-growing.html (accessed 6.8.21).

Taylor, B.D., Fink, C.N.Y., 2003. The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature.

U.S. Census Bureau, 2020. Southern and Western Regions Experienced Rapid Growth This Decade. URL https://www.census.gov/newsroom/press-releases/2020/south-west-fastest-growing.html

Valley Metro, 2020. Valley Metro Bus Schedule -City of Phoenix Open Data.

Valley Metro - Regional Public Transportation Authority (RPTA), 2019. Short Range Transit Program (SRTP) FY 2020-2024.

Voelkel, J., Hellman, D., Sakuma, R., Shandas, V., 2018. Assessing Vulnerability to Urban Heat: A Study of Disproportionate Heat Exposure and Access to Refuge by Socio-Demographic Status in Portland, Oregon. International Journal of Environmental Research and Public Health 15, 640. https://doi.org/10.3390/ijerph15040640

Zhao, Q., Guo, Y., Ye, T., Gasparrini, A., Tong, S., Overcenco, A., Urban, A., Schneider, A., Entezari, A., Vicedo-Cabrera, A.M., Zanobetti, A., Analitis, A., Zeka, A., Tobias, A., Nunes, B., Alahmad, B., Armstrong, B., Forsberg, B., Pan, S.-C., Íñiguez, C., Ameling, C., De la Cruz Valencia, C., Åström, C., Houthuijs, D., Dung, D.V., Royé, D., Indermitte, E., Lavigne, E., Mayvaneh, F., Acquaotta, F., de'Donato, F., Di Ruscio, F., Sera, F., Carrasco-Escobar, G., Kan, H., Orru, H., Kim, H., Holobaca, I.-H., Kyselý, J., Madureira, J., Schwartz, J., Jaakkola, J.J.K., Katsouyanni, K., Hurtado Diaz, M., Ragettli, M.S., Hashizume, M., Pascal, M., de Sousa Zanotti Stagliorio Coélho, M., Valdés Ortega, N., Ryti, N., Scovronick, N., Michelozzi, P., Matus Correa, P., Goodman, P., Nascimento Saldiva, P.H., Abrutzky, R., Osorio, S., Rao, S., Fratianni, S., Dang, T.N., Colistro, V., Huber, V., Lee, W., Seposo, X., Honda, Y., Guo, Y.L., Bell, M.L., Li, S., 2021. Global, regional, and national burden of mortality associated with non-optimal ambient temperatures from 2000 to 2019: a three-stage modelling study. The Lancet Planetary Health 5, e415–e425. https://doi.org/10.1016/S2542-5196(21)00081-4

CHAPTER THREE: Population co-exposure to extreme heat and wildfire smoke pollution in California during 2020

ABSTRACT

Excessive warming from climate change has increased the total wildfire burned area over the past several decades in California. This has increased population exposure to both hazardous concentrations of air pollutants from fires such as fine particulate matter (smoke $PM_{2.5}$) and extreme heat events. Exposure to $PM_{2.5}$ and extreme heat are individually associated with negative health impacts and recent epidemiological evidence points to synergistic effects from concurrent exposures. This study characterizes the frequency and spatial distribution of co-occurring extreme heat and smoke $PM_{2.5}$ events in California during the record-setting wildfire season of 2020. We measure exceedances over extreme thresholds of modeled surface-level smoke $PM_{2.5}$ concentrations and heat index based on observed temperature and humidity. We estimate that, during the studied period, extreme smoke and heat co-occurred at least once within 68% of the state's area (~288,000 km^2) and an average 2.5 times across all affected areas. Additionally, 16.5 million people, mostly in lower population density areas, were impacted at least once in 2020 by such synergistic events. Our findings suggest that public health guidance and adaptation policies should account for co-exposures, not only distinct exposures, when confronting heat and smoke $PM_{2.5}$.

INTRODUCTION

An increase in the frequency and severity of climate-related hazards has renewed interest in the distribution of multi-hazard events that can produce extraordinary risks (Field et al., 2012). Hazards may coincide in space and time by random chance, shared meteorological drivers or causal interdependency (Zscheischler et al., 2020). Examples include a flood after an earthquake;

the co-occurrence of extreme wind and flooding during severe storms (Nielsen et al., 2015); or the increased likelihood of landslides in wildfire-damaged areas, respectively (Mazdiyasni and AghaKouchak, 2015; Moftakhari et al., 2017). Whereas multi-hazard systems have been examined theoretically, empirical characterizations of their dynamics and drivers, not least in the wildfire pollution context, are more limited (Gill and Malamud, 2014).

Wildfires contribute to increased trace gas and aerosol concentrations that are harmful to human health. Fine particulate matter contributed by fires ("smoke $PM_{2.5}$"; particles smaller than 2.5 microns in diameter) is particularly dangerous because it directly enters bloodstreams and alveoli, impairing cardiorespiratory functions (Brook et al., 2010; Guo et al., 2018) as well as other organs such as the brain (Weuve et al., 2021). Moreover, in comparison with $PM_{2.5}$ from other sources, researchers have identified distinct mutagenic and oxidative stresses in humans from smoke $PM_{2.5}$ (Aguilera et al., 2021; DeFlorio et al., 2019; Nakayama Wong et al., 2011). In the western United States and California, smoke $PM_{2.5}$ was found to increase respiratory hospitalizations by as much as 7% and 3.3%, respectively, over a six-year period (Heaney et al., 2022; Liu et al., 2017).

While all-source $PM_{2.5}$ concentrations declined in the Eastern United States from 2006-2016, many areas in the West experienced an increase in $PM_{2.5}$ attributable to summertime wildfires that offset declines in non-fire anthropogenic sources (O'Dell et al., 2019). Wildfire smoke accounted for as much as half of the overall $PM_{2.5}$ exposure in the western United States in recent years, compared to approximately 20% on average in the mid-2000s (Burke et al., 2021). Future climate scenarios project that, by 2100, wildfire smoke will account for more than 50% of total $PM_{2.5}$ across the entire continental United States (Ford et al., 2018).

Extreme heat often precedes fire ignition as high temperatures predispose vegetational fuels to ignite and burn (Goss et al., 2020). Heat presents a sizable health risk of its own; it elevates heart and respiratory rates as well as blood viscosity and cholesterol, which may aggravate pre-existing conditions (Cheng et al., 2019; Davies and Maconochie, 2009; Ebi et al., 2021; Keatinge et al., 1986; Sherbakov et al., 2018). In California, select heat waves have been estimated to cause as much as a 6% increase in excess deaths (Hoshiko et al., 2010) and as much as a 39% and 47% increase in the likelihood of hospitalization for dehydration and renal failure, respectively (Schwarz et al., 2020).

Rising global temperatures and more frequent extreme heat events are expected to increase wildfire size and intensity, signaling a growing public health threat from concurrent heat-smoke exposure (Abatzoglou and Williams, 2016; Perkins et al., 2012; Westerling, 2018). However, exposure inventories of heat-smoke co-occurrence (HSC) that include smoke $PM_{2.5}$, which would help elucidate the drivers of hospitalization and death, are comparatively scarce. Austin et al. (2020) examined exposure to HSC among outdoor agricultural workers at the county level in Washington and found strong spatiotemporal variability in areas exposed to high heat and high levels of $PM_{2.5}$, with the largest co-occurrences during the summer wildfire season (Austin et al., 2020). More recently, researchers examined the co-occurrence of heat, ozone and $PM_{2.5}$ in the Western United States (Kalashnikov et al., 2022). These two studies, however, used concentration measurements that include all sources of $PM_{2.5}$ (versus smoke-specific $PM_{2.5}$) and relied on either unevenly distributed air quality stations or a coarse 10 km resolution, respectively. These spatially coarse measures of compound exposure constrain policymaking because they preclude investigation into the sociodemographic correlates of exposure (Schwarz et al., 2021).

In this study, we investigate the frequency, intensity and duration of individual and combined extreme heat and smoke $PM_{2.5}$ exposures in California from June through November 2020. During this time period wildfires burned over four million acres - the largest burned area in the State's recorded history dating back to 1878 - and coincided with the fourth hottest summer since 1895 (CalFire, 2020; National Weather Service, 2022). Our analysis is done at a 3 km spatial resolution and maps single and compound hazard exposures at surface level across different population characteristics, including race, ethnicity, income and health, to identify the communities most exposed to HSC. We contribute to the literature of compound climate exposures in California by: (1) modeling smoke $PM_{2.5}$ rather than total $PM_{2.5}$ to isolate fire contributions, (2) mapping compound exposures at fine spatial scale, and (3) identifying sociodemographic correlates of exposure.

METHODS

Smoke $PM_{2.5}$ exposure

We quantify smoke pollution exposure using NOAA's High-Resolution Rapid Refresh coupled with smoke (HRRR-Smoke) atmospheric model. Based on the Weather Research and Forecasting model coupled to Chemistry (WRF-Chem) model, HRRR-Smoke provides surface-level smoke $PM_{2.5}$ estimates across the United States in near real-time. Fire emissions estimates are based on satellite observations of fire radiative power (FRP) as detected by the VIIRS and MODIS satellites (Ahmadov et al., 2017). The model is initialized every 12 hours at a 3 km horizontal grid spacing; for this study, we utilize 48-hour forecasts initialized at 00 and 12 UTC. HRRR-Smoke $PM_{2.5}$ estimates do not account for non-fire sources of pollution (from traffic, industry, etc.) and are therefore specific to fire contributions only. Previous validation of HRRR-

Smoke with all-source ground station measurements during the 2018 Camp Fire identified strong spatiotemporal agreement with observed progressions of smoke plume locations and magnitudes (Chow et al., 2021).

We use the average of 24 individual hourly HRRR-Smoke forecasts to estimate daily smoke $PM_{2.5}$. All hourly concentration forecasts are based on the most recent available 00 or 12 UTC model initialization (n=364) given the increased accuracy of meteorological conditions closer to the forecast. In the event that an initialization is skipped (n=57) predictions from the most proximate initialization timestamp are used (i.e. the previous day's forecasts). Following this correction there are 3 days, out of 182 total, that are missing data for all initializations. These gaps, resulting from computer outages during the model run, are omitted from our analysis.

In the absence of smoke-specific observed $PM_{2.5}$ mass concentrations, we compare HRRR-Smoke forecasts with a network of 166 ground station all-source $PM_{2.5}$ measurements managed by the EPA's Air Quality System (AQS) (Figure S4). This dataset is an imperfect validation dataset since it includes anthropogenic sources of $PM_{2.5}$ in addition to wildfire smoke. However, during extreme smoke events when wildfire contributions dominate, AQS measurements may converge towards HRRR-Smoke estimates. AQS measurements are sourced from national, state and local air stations associated with parameter codes 88502 and 88101, providing daily average $PM_{2.5}$ concentrations from all sources (US EPA, 2014). The point geometries of AQS stations are coupled with gridded HRRR-Smoke estimates based on their intersection.

Temperature exposure

Extreme heat exposures are calculated from the Gridded Surface Meteorological dataset (GRIDMET) that includes daily surface measurements of maximum and minimum temperature, humidity and other meteorological variables across the contiguous United States (Abatzoglou, 2013). We resampled the data from its original 4 km resolution to 3 km to align with HRRR-Smoke output. Humidity and temperature are combined to estimate apparent temperature or heat index, a stronger correlate of biological heat stress that is referenced by the federal Occupational Safety and Health Administration in its exposure guidelines (Jacklitsch et al., 2016). We follow the National Weather Service's Weather Prediction Center's method, which adapts the Rothfusz regression model (Equation 3-1) to account for more extreme conditions and is reported in Fahrenheit (National Weather Service, 2014; Rothfusz, 1990).

Equation 3-1. Rothfusz Heat Index.

$$\textit{Daily Maximum Heat Index} = -42.379 + 2.04901523T + 10.14333127R - 0.22475541TR - 6.83783\times10^{-3}T^2 - 5.481717\times10\text{-}2R^2 + 1.22874\times10^{-3}T^2R + 8.5282\times10^{-4}TR^2 - 1.99\times10^{-6}T^2R^2$$

Where: T = Daily Maximum Temperature (°F)

R =Daily Minimum Relative Humidity (%)

These heat index measurements are then used to identify "exceedances" for our population exposure analysis (see Section 2.3). Previous validation work found the median correlations between GRIDMET and a national sample of over 1500 remote automated weather stations to be 0.94–0.95 and 0.87–0.90 for maximum and minimum temperature, respectively, with median mean absolute error (MAE) between 1.7 and 2.3 °C. Daily maximum and minimum

RH featured median correlation values between 0.77 and 0.81 and median MAE between 6 and 12% (Abatzoglou, 2013).

Exceedance thresholds

We define an extreme smoke exposure event as any day with 24-hour smoke $PM_{2.5}$ exceeding 20 $\mu g/m^3$. Our threshold of 20 $\mu g/m^3$ corresponds to the 98th percentile of smoke $PM_{2.5}$ measured by a global atmospheric chemistry model across the western United States between 2004-2009 (Liu et al., 2017). Finally, we define a *smoke wave*, designed to be analogous to a heatwave, as two or more consecutive exceedances. This accounts for the potential risks of sporadic, yet persistent, exposure to high concentrations of smoke $PM_{2.5}$ that are common during long-lasting conflagrations (Liu et al., 2017).

For heat, we define an extreme threshold as the greater of two prespecified intensity values. First, we calculate the 85th percentile historical heat index for the months of July and August, within a grid cell, for 1970-2010 (US EPA, 2021). This location-specific threshold accounts for any local behavioral or physical adaptations to extreme heat. Second, we apply an absolute minimum heat index cutoff of 80°F for daily maximum temperatures, which corresponds to the National Institute for Occupational Safety and Health's lowest "caution" heat-index for worker safety (Jacklitsch et al., 2016). Accordingly, colder areas where the summertime 85th percentile corresponds to a mild temperature will instead be compared to an absolute cutoff. Conversely, to avoid "false negatives", we also apply an absolute maximum heat index cutoff of 105°F for areas with extreme summertime 85th percentile historical heat indices; this threshold corresponds to the National Weather Service's excessive heat warning trigger.

Finally, in addition to single event exceedances, we adopt a definition of *heat wave* as two or more consecutive exceedances in a given location.

Using Google Earth Engine, a cloud-based geocomputation engine, we intersect both hazard exceedances to identify HSC at a daily interval at a resampled scale of 3 km (Gorelick et al., 2017). Since there are no standardized definitions for extreme heat or extreme smoke, we further analyze the sensitivity of our results to different definitions: for temperature we increase the threshold to the 95th percentile and for smoke we test an absolute cutoff of 35 $\mu g/m^3$, which corresponds to the EPA's 24-hour national ambient air quality standard for all source $PM_{2.5}$ (US EPA, 2014). Finally, we test the differences between using daily maximum and minimum heat indices for temperature percentile thresholds as well as the impact of focusing only on persistent exceedances (smoke waves or heat waves). These alterations are informed by evidence of hotter evening temperatures and persistent heat exposure increasing morbidity and mortality (Rey et al., 2007; Zhang et al., 2012).

Population characteristics

To quantify total human exposure and the density of affected areas, population data are taken from Gridded Population of the World Version 4.11, which is an interpolation of decennial census population counts at a 1 km resolution (CIESIN, 2018; OEHHA, 2021). Additionally, we use California's Office of Environmental Health Hazard Assessment CalEnviroScreen 4.0 and the 2019 5-year American Community Survey to measure existing pollution burdens and socioeconomic characteristics for all census tracts across the State, respectively (U.S. Census Bureau, 2019) (Figure S5). We resolve spatial mismatch between the 3 km hazard layer and

variably sized census tracts by computing the aggregated hazard metrics within a 9 km radius (3x3 pixel window) of each census tract's centroid.

We also measure the correlation between each hazard's magnitude and the prevalence of different sociodemographic variables. Each variable is first ranked and binned into deciles and then the mean magnitude exceedance is computed for each hazard in a given census tract's vicinity for each bin. For analyses related to population exposure (incidence), where we compute the relative ratio between a group's share of the HSC impacted population and their share of the total State population, we sum population counts for each census tract in proportion to the area of the tract affected at least once by HSC.

RESULTS

HRRR-Smoke model comparison

In comparison with the EPA's AQS network of ground stations, smoke $PM_{2.5}$ concentrations simulated by the HRRR-Smoke model were typically lower than all-source $PM_{2.5}$ station measurements; the median HRRR-Smoke grid cell estimate corresponding to each site was 0.4 $\mu g/m^3$ versus 10.0 $\mu g/m^3$ for AQS. This difference is expected since AQS accounts for all sources of $PM_{2.5}$, most days do not feature any smoke pollution and a majority of the AQS stations are located in urban areas with heavier anthropogenic contributions. Indeed, biases between AQS and HRRR-Smoke were highest in less populated areas and lowest in urban areas where car exhaust and residential gas appliances, for example, dominate $PM_{2.5}$ emissions (Figure S2). The correlation coefficient between both datasets, which was greatest for same day comparisons without lag, equaled 0.62, indicating sufficient temporal coherence. When comparing all AQS measurements with the nearest HRRR-Smoke forecast, the normalized mean

absolute error (NMAE) equaled 77%; after filtering for "extreme" HRRR-Smoke forecasts (>20 $\mu g/m^3$) NMAE drops to 60%. These statistics match reported metrics for previous comparisons with smoke-enhanced aerosol optical depth measurements during the Williams Flats Fire in 2019 (Ye et al., 2021).

Spatiotemporal Trends

Both hazards co-occurred at least once across 288,505 km^2 (~68%) of California during the 2020 study period (Figure 3-1). The statewide average of the season's maximum observed exceedance over the baseline thresholds was 3.0°F for the daily maximum heat index and 51.1 $\mu g/m^3$ for smoke $PM_{2.5}$. The latter is more than twice the smoke event threshold and almost 150% above the EPA 24-hour standard of 35 $\mu g/m^3$. Notably, exceedance magnitudes during compounding events were lower for temperature, averaging 2.6°F. Smoke exceedances were also lower during HSC events, averaging 40.9 $\mu g/m^3$.

HSC occurred a maximum of seventeen times across four different grid cells in California's Carmel Valley and for as long as nine consecutive days along the western edge of the Mojave Desert, southeast of Sequoia National Park. The Monterey Bay area was also one of the most frequently affected areas in the State. Most co-occurrences coincided spatiotemporally with season's most severe conflagrations (Figure 3-2) including the August Complex, North Complex and Creek Fires that affected the Northern California coastline, Upper Sierras and Central Valley, respectively.

We also identified trends across the four Level 1 North American ecoregions in California which are: Mediterranean California (Mediterranean), North American Deserts (Desert), Marine West Coast Forests (Marine) and Northwestern Forested Mountains (Forested).

These ecoregions delineate distinct ecologies and climates that can affect smoke production and temperature (Appendix 3A) (Omernik, 1987). California's Marine and Forested ecoregions were the two most affected by HSC, with nearly equal frequency-weighted impacted areas of 1.2%. For extreme heat alone, using the baseline cutoff defined in Section 2.3, we find that Mediterranean coastal regions as well as Desert areas in the Mojave were most frequently affected (Appendix 3F). These areas, in addition to the Central Valley, a primary agricultural region, also experienced the most intense heat events. For smoke $PM_{2.5}$, the season's average exceedance over the baseline threshold was greatest in Northwestern Forested Mountains, at 84.1 $\mu g/m^3$, and these events also lasted longest, for an average of 4.5 days. The exceedance magnitude in Forested areas was significantly larger than the next-most impacted Marine ecoregion (μ =50.2 $\mu g/m^3$).

Time-series analysis shows peak heat and smoke PM2.5 during the months of August and September. The Pearson correlation coefficient between daily HSC and smoke areas (ρ=0.45) was smaller than for HSC and heat (ρ=0.63). There is also a discernible lagged trend between hazards. Time-lagged cross correlation, which identifies the offset (number of days) at which cross-correlation is maximized between variables, peaks at 6 days for maximum heat index and smoke (ρ=0.57). This suggests that an increase in the frequency of heat waves longer than six days may increase co-occurrences, assuming stationarity in heat-wildfire dynamics.

Figure 3-1. Spatial distribution of individual and co-occurring climate hazards. (a) The HSC map shows the frequency of HSC events in a given 3 km grid cell. Areas in white did not experience any event throughout the studied period. (b) The smoke magnitude displays the season's mean exceedance above 20 μg m−3 for each grid cell. (c) The smoke frequency displays the total number of days with smoke magnitudes exceeding the 'extreme' threshold of 20 μg m−3. (d), (e) The same definitions for magnitude and frequency apply relative to each grid cell's 85th percentile historical heat index.

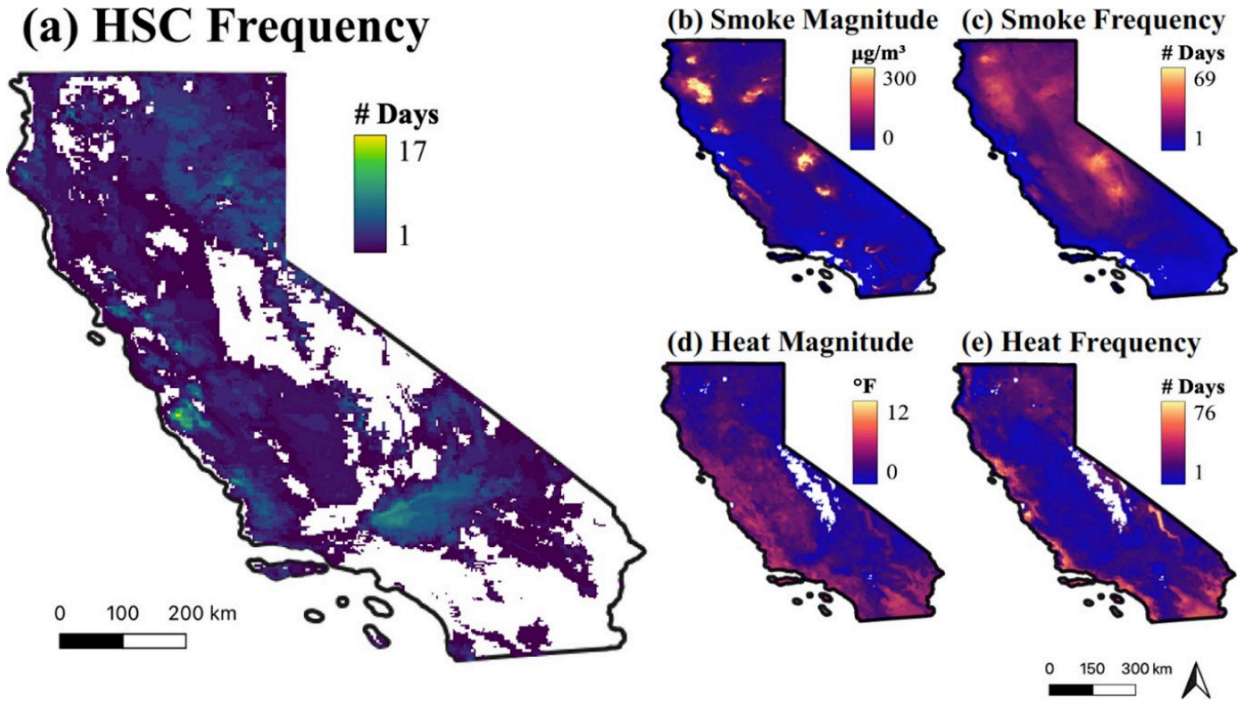
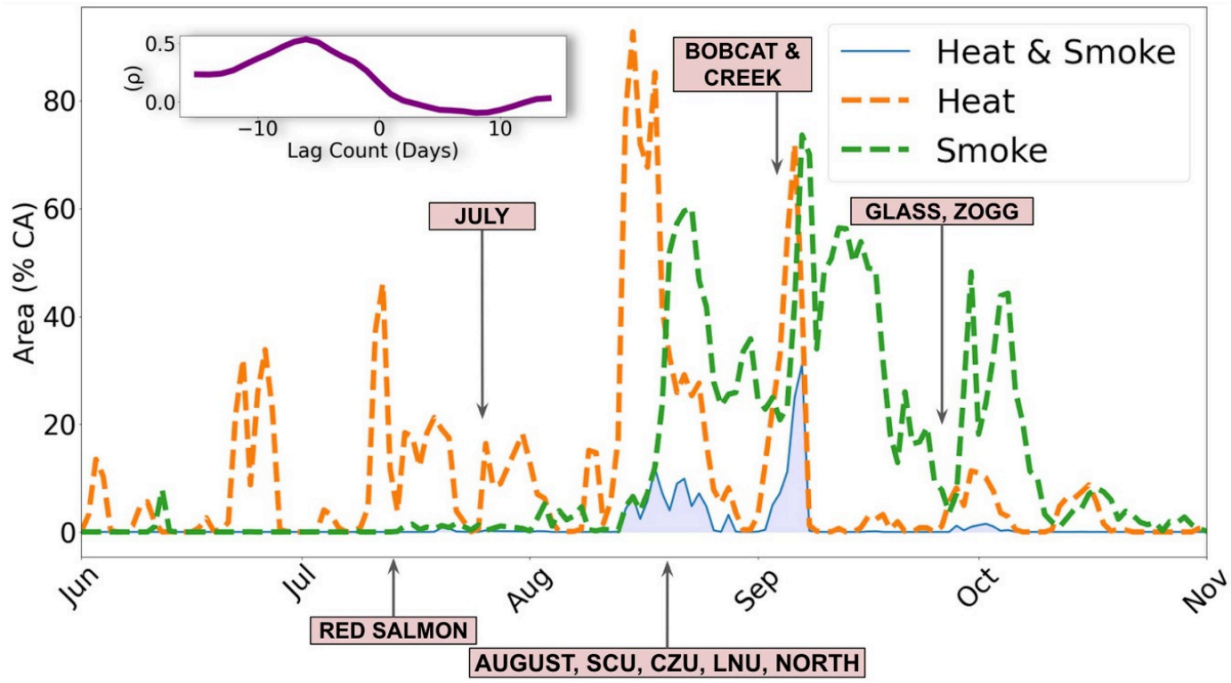


Figure 3-2. Area of California in extreme heat, smoke, and combined hazard categories for summer 2020. Dashed lines show the percent of California affected by individual hazard exceedances for the studied period. The solid blue line corresponds to the area of concurrent heat and smoke. Boxes with arrows indicate name and ignition dates of wildfires (±2 d). The inset line chart on the top left shows the cross-correlation between the daily area of heat and smoke exceedance at different time lags.



Sensitivity

There is a nonlinear decline in the frequency of HSC when we examine persistent smoke waves and/or heat waves. As Table 3-1 illustrates, applying a persistence threshold of two or more consecutive exceedances for both smoke and heat diminishes the total HSC-affected area by 27% from ~288,000km^2 to approximately ~211,000km^2 of the State. This decline in area corresponds to a slightly larger 33% decline in the total affected population. Our results are also sensitive to the extreme heat and smoke thresholds - an increase from the 85th to 95th percentile decreases frequency-weighted affected area by nearly 60% and an increase in the smoke threshold from 20 µg/m^3 to 35 µg/m^3 results in a drop of 47%. For reference, the average 85th and 95th percentiles for maximum heat index in California correspond to 92°F and 95°F, respectively.

Finally, persistent events, namely heat waves and smoke waves, are associated with more intense magnitudes, exceeding non-persistent scenarios by an average 9.5% and 10.4%, respectively.

Table 3-1. Sensitivity Analysis. (a) Frequency weighted area is the occurrence frequency of HSC in a given pixel, multiplied by the area of that pixel; (b) minimum extent is the percent of CA that experiences at least one HSC event in the studied timespan; (c) maximum pixel frequency is the maximum number of HSC events observed in a pixel, statewide; (d) HSC average frequency is the statewide average of all HSC counts among pixels with HSC counts greater than zero (i.e. unaffected areas are masked); (e) mean heat magnitude is the statewide mean of each pixel's seasonal mean exceedance above the heat threshold for days and locations with HSC; (f) mean smoke magnitude is the same as e but for smoke; (g) HSC maximum persistence is the largest observed consecutive run length of HSC events in a pixel throughout the state; (h) HSC average persistence is the same as g but the average; (i) population affected is the total number of people residing in locations that are within the minimum extent (see b).

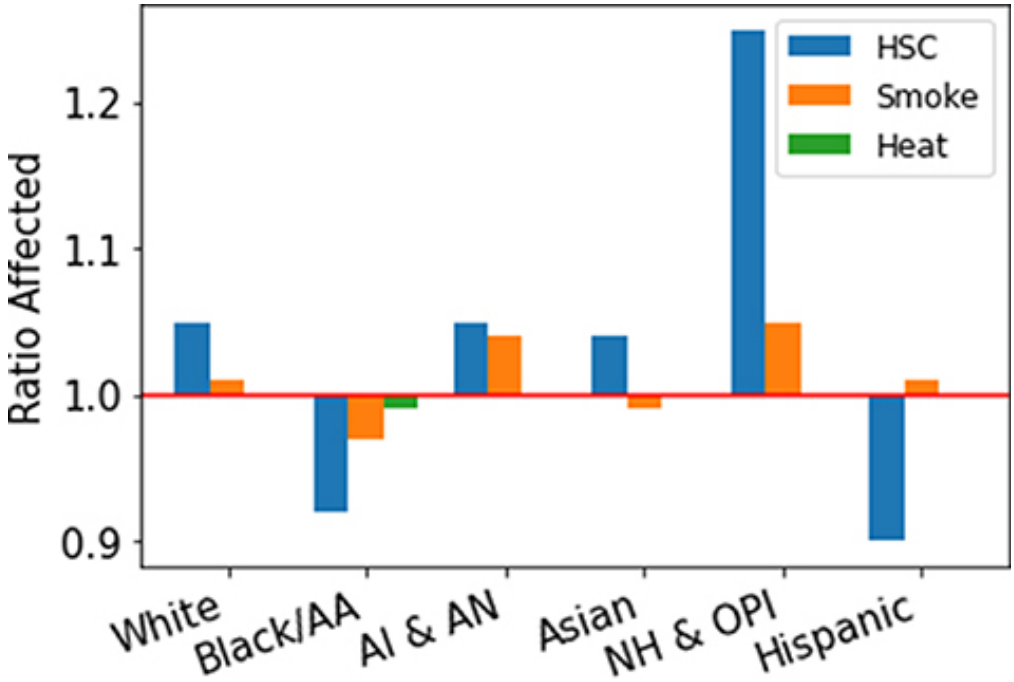
MAGNITUDE THRESHOLDS			PERSISTENCE THRESHOLD	HSC OUTCOMES (STATEWIDE)								
HEAT (%ile)	SMOKE ($\mu g/m^3$)	HEAT METRIC	HEAT & SMOKE (days)	FREQUENCY WEIGHTED AREA (km^2-day)	MINIMUM EXTENT (%)	HSC MAXIMUM FREQUENCY	HSC AVERAGE FREQUENCY	MEAN HEAT MAGNITUDE EXCEEDANCE (°F)	MEAN SMOKE MAGNITUDE EXCEEDANCE ($\mu g/m^3$)	HSC MAXIMUM PERSISTENCE (days)	HSC AVERAGE PERSISTENCE (days)	POPULATION AFFECTED (people)
85	20	max	0	7.2E+05	68.1	17	2.5	2.6	40.9	9	1.6	1.7E+07
85	20	min	0	1.7E+06	79.7	26	5.1	4.0	47.5	22	1.9	1.8E+07
95	20	max	0	2.8E+05	35.8	13	1.8	2.2	29.9	8	1.4	8.4E+06
95	20	min	0	8.2E+05	63.4	21	3.0	3.2	50.3	11	1.6	1.2E+07
85	35	max	0	3.8E+05	48.6	12	1.9	2.3	54.9	8	1.4	8.9E+06
85	35	min	0	1.1E+06	65.7	22	3.9	3.9	58.6	19	1.8	1.1E+07
95	35	max	0	1.3E+05	20.0	8	1.5	1.8	44.5	6	1.3	3.4E+06
95	35	min	0	4.8E+05	48.2	14	2.4	3.1	64.0	10	1.4	7.0E+06
85	20	max	2	5.2E+05	49.6	12	2.5	2.8	46.8	9	2.0	1.1E+07
85	20	min	2	1.4E+06	68.9	26	4.9	4.1	53.2	22	2.3	1.4E+07
95	20	max	2	1.9E+05	21.8	9	2.0	2.4	32.5	8	1.7	5.7E+06
95	20	min	2	6.1E+05	47.4	20	3.0	3.5	47.5	11	1.9	7.7E+06
85	35	max	2	2.5E+05	31.0	8	1.9	2.6	63.4	8	1.8	5.4E+06
85	35	min	2	8.5E+05	53.8	22	3.7	3.9	68.0	19	2.2	8.6E+06
95	35	max	2	7.2E+04	10.1	6	1.7	2.3	54.2	6	1.6	1.4E+06
95	35	min	2	3.3E+05	31.6	13	2.5	3.4	64.8	10	1.8	3.6E+06

Population Exposure

Out of approximately 40 million residents, we estimate that 35.3, 30.2 and 16.5 million residents were affected by at least one occurrence of extreme heat, smoke and HSC, respectively

(Figure S3). Relative to the statewide average population density of 98.9 persons/km^2, heat-affected areas were denser (μ=102.43) and smoke-affected areas less dense (μ=86.5). Gridded areas with at least one HSC event had, on average, lower population densities (μ=67.5, σ = 392.8, n=40,425) than unaffected areas (μ=171.3, σ = 771.7, n=17,735).

Figure 3-3. Proportionality. Bar plots show the ratio above or below one at which different ethnicities and races in California were affected by individual hazard exceedances and HSC, relative to their share of the general population. Abbreviations: Black & AA—Black and African American; AI & AN—American Indian and Alaska Native; NH & OPI—Native Hawaiian & Other Pacific Islander; Latino—Latino or Hispanic.



Proportionality tests show that certain populations are overrepresented in areas with HSC compared to their total statewide representation (Figure 3-3). Based on census demographic data from 2019, “White Alone” respondents were 1.05 times as likely to be exposed to HSC than would be expected based on their overall share of the population. Conversely, “Black or African American Alone” individuals, who represent 5.7% of the total population in California,

represented 5.2% of the population exposed to HSC. Hispanic and Latino populations were even less likely to reside in HSC affected areas - comprising 39.2% of the State's population but only 34.0% of the affected population (Appendix 3H).

Last, we compare smoke and heat exceedance magnitudes with a series of sociodemographic and risk variables at the census tract level (Appendix 3G), namely the shares of the population that are White, Hispanic, African American and Native American; the percent of the population that suffers from cardiovascular illness; the level of poverty and linguistic isolation; summertime average 8-hour maximums of ozone concentrations from 2017-2019; and the OEHAA's CalEnviroScreen score, a holistic score that combines an area's pollution burden with its population vulnerability. Among variables that show non-zero linear trends ($p<0.01$), the largest observed magnitude was a decrease in the exceedance magnitude of temperature -0.1°F for each decile increase in the summertime maximum ozone and -0.06°F for each decile increase in the percent of the population that is White-identifying. Nearly all other variables' associations with smoke were less than +/- 0.02 $\mu g/m^3$ per decile increase or statistically insignificant.

DISCUSSION

Our study shows that 68% of California's land area and 42% of the population simultaneously experienced hazardous smoke $PM_{2.5}$ and extreme temperatures at least once in 2020. While these results may represent an "upper bound" considering 2020's wildfire season was the largest in California's modern history, four million burned acres was in fact typical for the State prior to European settlements and concomitant fire suppression (Safford et al., 2022). These compound events peaked in August and lasted through October. The Forested ecoregion in northern California and neighboring Marine ecoregion, where large wildfires were observed in

2020, were most affected by HSC. These spatiotemporal patterns match findings from another smoke $PM_{2.5}$ exposure assessment in earlier years which, using a coarser smoke-specific model, found peak concentrations in August concentrated in northern California (Koman et al., 2019). These patterns may be attributable to available fuel loads in these ecoregions, as well as differences in plant sensitivity to dryness that influence wildfire risk (McKinnon et al., 2021; Rao et al., 2022).

Overall, persistent HSC events (at least 2 days) were found to be more intense than single day hazard extremes. Hazard magnitudes were also found to be lower during HSC events than for individual hazards. This is likely attributable to the relatively temperate climates in which smoke events were concentrated. Notwithstanding, previous studies have found that milder climates with less adaptive capacity are similarly vulnerable to heightened morbidity and mortality from persistent heat exposure (Knowlton et al., 2009). Finally, the average extreme smoke $PM_{2.5}$ concentration during HSC events was more than double the EPA's 24-hour standard (35 $\mu g/m^3$) and presents a significant health risk to affected communities.

We do not find many associations between the amount of threshold exceedance for location-specific "extremes" for heat and smoke and the selected sociodemographic covariates at the census-tract level. That is to say that no population group or social pattern was associated with a disproportionate increase or decrease in the intensity of extreme events. The lack of meaningful trends, especially in the smoke context, are likely because wildfire ignition and the subsequent meteorological conditions that influence smoke transport are stochastic and affect broad regions. Yet, in line with previous work that finds minority groups to be less exposed to wildfire (Burke et al., 2021; Masri et al., 2021), we similarly find evidence that white individuals

are more likely to be exposed to concurrent heat and smoke. These predispositions are likely a result of certain groups' propensity to live in rural areas nearer smoke emitting fires.

There are several limitations to this analysis. Because there are no empirical datasets of surface-level smoke $PM_{2.5}$ across the State, we relied on short-term forecasts that inherently feature some imprecision. However, unlike observed measurements from AQS that cover a small fraction of the State in mostly urban areas, HRRR-Smoke is spatially continuous and isolates the fire contribution to pollution. Although our case study of California in 2020 is instructive due to the State's ecological diversity and size, our results cannot be directly translated to other geographies and years. Nonetheless, climate driven fire risk in California will only worsen through the middle of the century; seasons of similar intensity for both heat and smoke are therefore plausible, if not likely, to reoccur. Additionally, our heat index measurements, while based on observed temperatures, do not account for microclimate impacts or wind which are both known to shape heat stress during outdoor activities (Thorsson et al., 2014). Finally, because we lack a definitive understanding of people's physiological and adaptive responses to varying intensities of heat and smoke, our analysis makes population-level assumptions about the magnitudes that may be considered hazardous.

CONCLUSION

We report population-level exposure for individual and co-occurring climate-related hazards and find that a majority of California's land area, especially ecoregions in northern California with dense fuel loads, were affected by HSC. The location of these events are unassociated with social indicators of vulnerability, however, they tend to cluster in rural areas near observed fire perimeters (Appendix A). To our knowledge, this is the first study to describe

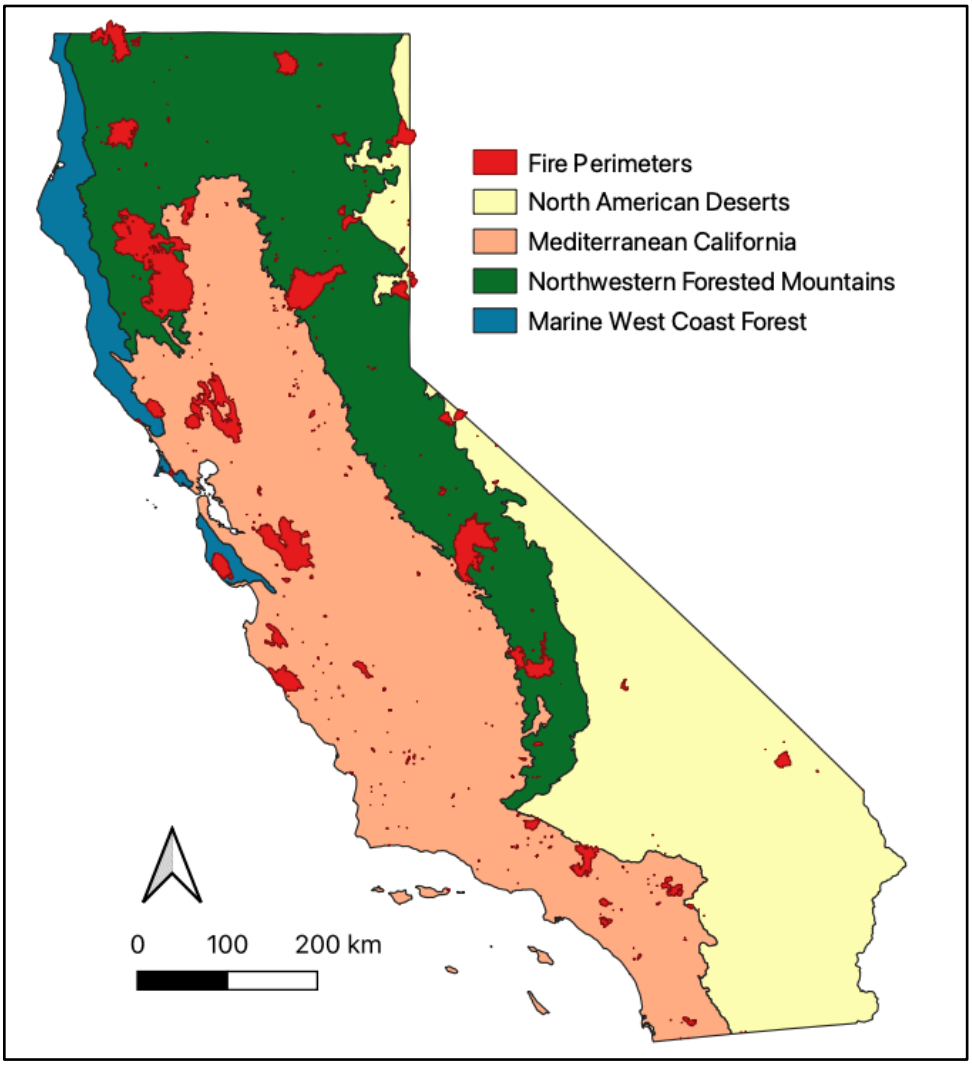
the spatiotemporal dynamics of HSC throughout California and to examine their disproportionate impacts on certain communities in the State. We advance previous work by examining exposures at a relatively high spatial resolution (3 km) with smoke-specific $PM_{2.5}$ estimates to isolate wildfire contributions from other sources of air pollution.

This study suggests several promising areas for future research. First, researchers can leverage this analysis to estimate the excess morbidity and mortality resulting from the *interactions* between heat and smoke $PM_{2.5}$ while accounting for mediating sociodemographic factors that may otherwise be obfuscated at coarser spatial resolutions. Second, longer-term studies can examine the meteorological and geophysical drivers of heat and smoke to identify multi-year HSC exposures as well as the causal mechanisms behind HSC, and in turn, enable better prediction. This would require an expanded time series of smoke $PM_{2.5}$ concentrations that is not currently available from the HRRR-Smoke model that became operational in 2020; therefore, other observational or modeling datasets are needed to backfill historical smoke patterns. Third, future climate change scenarios may alter HSC frequency and duration (Kalashnikov et al., 2022), but this was beyond the scope of this analysis. Finally, we can further investigate the social drivers of differential exposure between racial and ethnic groups including housing stock, urban tree canopy cover and occupation.

With the intensity and duration of extreme heat events and wildfires projected to increase over the coming decades (Westerling, 2018, 2016), HSC is likely to become increasingly frequent. Accordingly, public health officials must account for hazard interactions in their planning efforts and their potential to incur harms on human health that exceed the sum of their parts.

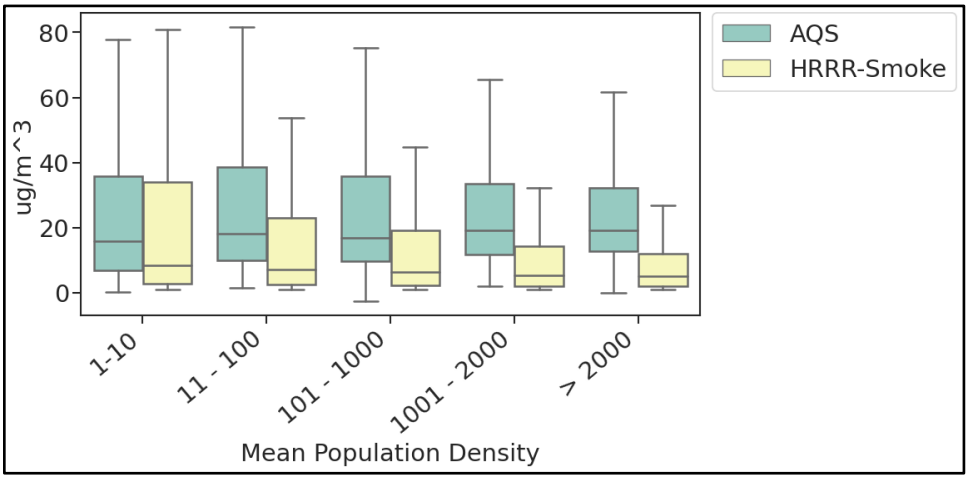
APPENDIX 3

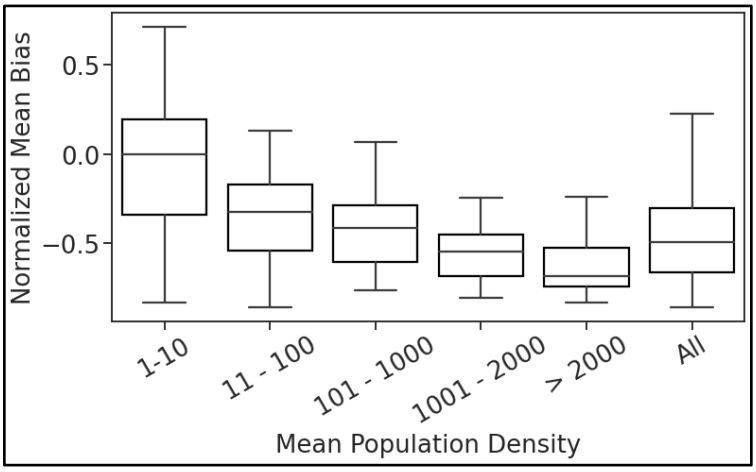
Appendix 3A. Ecoregions of California. Each polygon corresponds to a level one ecoregion according to the EPA's Ecoregions of North America. The North American Desert region contains the warm Sonoran and Mojave deserts as well as the colder Central Basin and Range. The Mediterranean California region spans regions with coastal sage, chaparral and oak woodlands, the State's central valley as well as pine-oak mountain ranges. The Marine West Coast Forest is also known as the Coast Range, featuring highly productive, rain-drenched coniferous forests, including redwoods. Finally, the Forested Mountains contain the Sierra Nevadas, Cascades, Klamath Mountains and Eastern Cascade Slopes and Foothills. Overlaid are the footprints of the 20 largest fires in 2020 from CalFire's FRAP perimeter database.



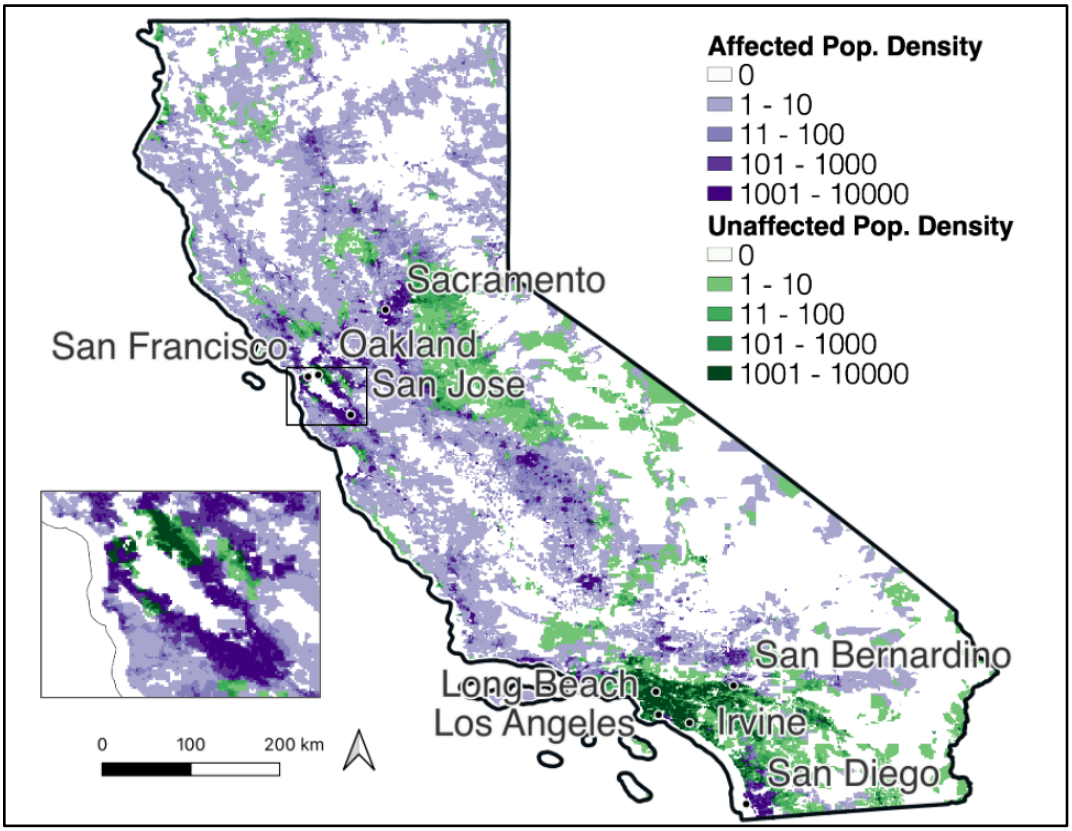
Appendix 3B. HRRR Validation.

(Top) Interquartile ranges between AQS station daily mean $PM_{2.5}$ measurement and the corresponding HRRR-Smoke forecasted smoke $PM_{2.5}$ mass concentrations across bins of increasing population density (average within a 3 km grid cell). Population densities (persons/km^2) are 1-10 (n=19), 11-100 (n=24), 101-1000 (n=52), 1001-2000 (n=28), >2000 (n=43). Discrepancies between median and interquartile ranges increase with population density. ***(Bottom)*** Normalized mean biases between HRRR-Smoke and all daily AQS measurements per station; stations are binned according to the surrounding population density. Normalized mean bias is the total difference between HRRR-Smoke estimates and AQS measurements for a single station, normalized by the station's total observed mass concentrations. Larger negative biases (model underestimation) are observed as population densities (i.e. anthropogenic aerosols) increase. Both plots are filtered for dates where HRRR-Smoke forecasts are non-zero (>1 $\mu g/m^3$).

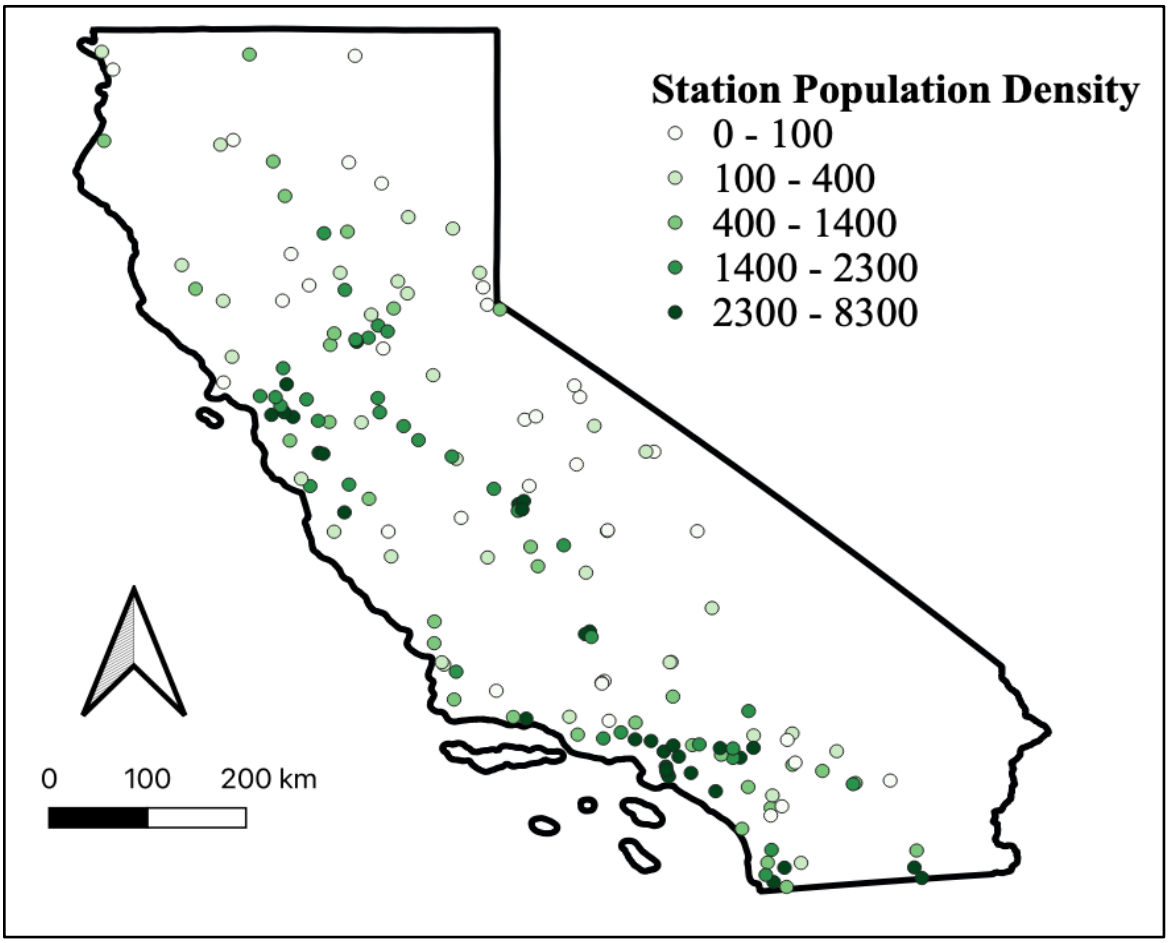




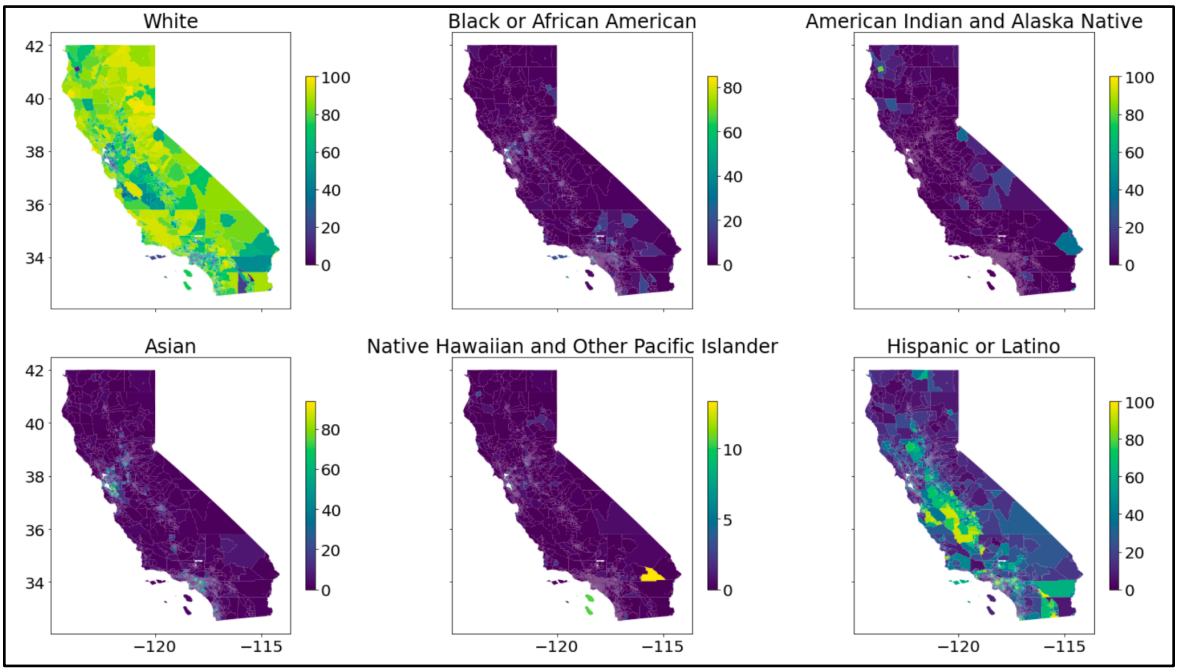
Appendix 3C. Heat-Smoke Co-Occurrence Affected Populations. Purple and green areas correspond to areas that were affected and unaffected by heat-smoke co-occurrence, respectively. Color shading is the log scaled population density. White areas are not populated. The zoomed inset map of the Bay Area highlights the spatial variability in HSC across relatively small distances as well as the concentration of population in urban areas.



Appendix 3D. AQS Stations. Below is the spatial distribution of EPA's ground air quality monitoring stations across California that were compared to HRRR-Smoke estimates. Coloring corresponds to the population density (persons/km^2) within a 4 km buffer around the station.



Appendix 3E. California Population Maps. Each map displays the percent of each census tract's population represented by Census respondents of each race or ethnicity "alone". Estimates are based on the 5 year American Community Survey from 2019. Shared X and Y axes correspond to longitude and latitude, respectively.



Appendix 3F. Ecoregion Summary Statistics. **"**Ecoregion Area" is the total area of each national ecoregion within California. The "Frequency Weighted Area Affected Percentage" is the sum of all affected grid cell areas multiplied by their frequencies in each ecoregion, divided by the product of the total area of the ecoregion and the total number of days in the summer season (n=179). The "Average Maximum Exceedance" is the statewide average of all maximum exceedances for all affected 3 km grid cells in an ecoregion.

Ecoregion	**Ecoregion Area (km^2)**	**Frequency Weighted Area Affected (%)**			**Average Maximum Exceedance (ug/m^3) | °F**	
		HSC	**Smoke**	**Heat**	**Smoke**	**Heat**
Northwestern Forested Mountains	1.16e+5	1.17	17.54	7.56	84.06	2.03
Marine West Coast Forest	1.33e+4	1.17	5.00	12.10	50.19	3.62
North American Deserts	1.20e+5	0.87	6.34	10.53	31.69	2.78
Mediterranean California	1.60e+5	0.88	9.73	8.04	42.33	3.72

Appendix 3G. Population Exposure. Slope values, Pearson correlation coefficients, p-values and standard errors for all social indicator deciles and smoke or heat magnitude exceedances. Smoke and heat outcomes are the statewide average of each non-zero pixel's seasonal average exceedance relative to the baseline threshold (85th percentile HI, 20 μg/m^3).

Variable	Slope	Correlation Coeff.	P Value	Std. Error
Smoke Magnitude				
African American	-0.02	-0.95	0.00	0.00
Ozone	-0.00	-0.06	0.87	0.01
Linguistic Isolation	-0.01	-0.89	0.00	0.00
Cardiovascular	-0.01	-0.48	0.16	0.00
Hispanic	-0.00	-0.42	0.23	0.00
White	0.01	0.74	0.01	0.00
CalEnviroScreen	-0.01	-0.74	0.02	0.00
Poverty	-0.00	-0.28	0.43	0.00
Native American	-0.00	-0.27	0.45	0.00
Heat Magnitude				
African American	0.01	0.39	0.26	0.01
Ozone	-0.10	-0.76	0.01	0.03
Linguistic Isolation	0.04	0.94	0.00	0.01
Cardiovascular	-0.05	-0.94	0.00	0.01
Hispanic	0.00	0.03	0.94	0.01
White	-0.06	-0.91	0.00	0.01
CalEnviroScreen	-0.02	-0.55	0.10	0.01
Poverty	-0.05	-0.92	0.00	0.01
Native American	0.01	0.16	0.65	0.02

Appendix 3H. Population HSC Exposure Proportionality. **"**Total Population" is the statewide share of the population represented by each race/ethnicity. "Affected Population" corresponds to the total population of each race/ethnicity in census tracts with non-zero HSC areas, proportional to the ratio of area affected and the total area of a tract. The "Affected Ratio" is the fraction of the affected population share relative to the total population share.

		HSC		**Smoke**	**Heat**
Race/Ethnicity	**Total Population (%)**	**Affected Population (%)**	**Affected Ratio**	**Affected Ratio**	**Affected Ratio**
Total Population: White Alone	59.70	59.70	1.05	1.00	1.00
Total Population: Black or African American Alone	5.79	5.79	0.92	0.97	1.00
Total Population: American Indian and Alaska Native Alone	0.77	0.77	1.05	1.04	0.99
Total Population: Asian Alone	14.50	14.49	1.04	0.99	1.00
Total Population: Native Hawaiian and Other Pacific Islander Alone	0.40	0.40	1.25	1.05	1.00
Total Population: Hispanic or Latino	39.02	39.02	0.90	1.01	1.00

BIBLIOGRAPHY

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol. 33, 121–131. https://doi.org/10.1002/joc.3413

Abatzoglou, J.T., Williams, A.P., 2016. Impact of anthropogenic climate change on wildfire across western US forests. PNAS 113, 11770–11775. https://doi.org/10.1073/pnas.1607171113

Aguilera, R., Corringham, T., Gershunov, A., Benmarhnia, T., 2021. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. Nat Commun 12, 1493. https://doi.org/10.1038/s41467-021-21708-0

Ahmadov, R., Grell, G., James, E., Csiszar, I., Tsidulko, M., Pierce, B., McKeen, S., Benjamin, S., Alexander, C., Pereira, G., Freitas, S., Goldberg, M., 2017. Using VIIRS fire radiative power data to simulate biomass burning emissions, plume rise and smoke transport in a real-time air quality modeling system, in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2806–2808. https://doi.org/10.1109/IGARSS.2017.8127581

Austin, E., Kasner, E., Seto, E., Spector, J., 2020. Combined Burden of Heat and Particulate Matter Air Quality in WA Agriculture. Journal of Agromedicine 0, 1–10. https://doi.org/10.1080/1059924X.2020.1795032

Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate Matter Air Pollution and Cardiovascular Disease. Circulation 121, 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., Wara, M., 2021. The changing risk and burden of wildfire in the United States. Proc Natl Acad Sci USA 118, e2011048118. https://doi.org/10.1073/pnas.2011048118

Cheng, J., Xu, Z., Bambrick, H., Prescott, V., Wang, N., Zhang, Y., Su, H., Tong, S., Hu, W., 2019. Cardiorespiratory effects of heatwaves: A systematic review and meta-analysis of global epidemiological evidence. Environmental Research 177, 108610. https://doi.org/10.1016/j.envres.2019.108610

Chow, F.K., Yu, K.A., Young, A., James, E., Grell, G., Csiszar, I., Tsidulko, M., Freitas, S., Pereira, G., Giglio, L., Friberg, M.D., Ahmadov, R., 2021. High-resolution smoke forecasting for the 2018 Camp Fire in California. Bulletin of the American Meteorological Society 1. https://doi.org/10.1175/BAMS-D-20-0329.1

CIESIN, 2018. Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11.

Davies, P., Maconochie, I., 2009. The relationship between body temperature, heart rate and respiratory rate in children. Emergency Medicine Journal 26, 641–643. https://doi.org/10.1136/emj.2008.061598

DeFlorio, -Barker Stephanie, Crooks, J., Reyes, J., Rappold, A.G., 2019. Cardiopulmonary Effects of Fine Particulate Matter Exposure among Older Adults, during Wildfire and Non-Wildfire Periods, in the United States 2008–2010. Environmental Health Perspectives 127, 037006. https://doi.org/10.1289/EHP3860

Ebi, K.L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R.S., Ma, W., Malik, A., Morris, N.B., Nybo, L., Seneviratne, S.I., Vanos, J., Jay, O., 2021. Hot weather and heat extremes: health risks. The Lancet 398, 698–708. https://doi.org/10.1016/S0140-6736(21)01208-3

Field, C.B., Barros, V., Stocker, T.F., Dahe, Q. (Eds.), 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139177245

Ford, B., Val Martin, M., Zelasky, S.E., Fischer, E.V., Anenberg, S.C., Heald, C.L., Pierce, J.R., 2018. Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. GeoHealth 2, 229–247. https://doi.org/10.1029/2018GH000144

Gill, J.C., Malamud, B.D., 2014. Reviewing and visualizing the interactions of natural hazards. Reviews of Geophysics 52, 680–722. https://doi.org/10.1002/2013RG000445

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, Big Remotely Sensed Data: tools, applications and experiences 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Goss, M., Swain, D.L., Abatzoglou, J.T., Sarhadi, A., Kolden, C.A., Williams, A.P., Diffenbaugh, N.S., 2020. Climate change is increasing the likelihood of extreme autumn wildfire conditions across California. Environ. Res. Lett. 15, 094016. https://doi.org/10.1088/1748-9326/ab83a7

Guo, C., Zhang, Z., Lau, A.K.H., Lin, C.Q., Chuang, Y.C., Chan, J., Jiang, W.K., Tam, T., Yeoh, E.-K., Chan, T.-C., Chang, L.-Y., Lao, X.Q., 2018. Effect of long-term exposure to fine particulate matter on lung function decline and risk of chronic obstructive pulmonary disease in Taiwan: a longitudinal, cohort study. The Lancet Planetary Health 2, e114–e125. https://doi.org/10.1016/S2542-5196(18)30028-7

Heaney, A., Stowell, J.D., Liu, J.C., Basu, R., Marlier, M., Kinney, P., 2022. Impacts of Fine Particulate Matter From Wildfire Smoke on Respiratory and Cardiovascular Health in California. GeoHealth 6. https://doi.org/10.1029/2021GH000578

Jacklitsch, B., Musolin, K., Williams, J., Coca, A., Kim, J.-H., Turner, N., 2016. Criteria for a Recommended Standard: Occupational Exposure to Heat and Hot Environments.

Kalashnikov, D.A., Schnell, J.L., Abatzoglou, J.T., Swain, D.L., Singh, D., 2022. Increasing co-occurrence of fine particulate matter and ground-level ozone extremes in the western United States. Science Advances. https://doi.org/10.1126/sciadv.abi9386

Keatinge, W.R., Coleshaw, S.R., Easton, J.C., Cotter, F., Mattock, M.B., Chelliah, R., 1986. Increased platelet and red cell counts, blood viscosity, and plasma cholesterol levels during heat stress, and mortality from coronary and cerebral thrombosis. Am J Med 81, 795–800. https://doi.org/10.1016/0002-9343(86)90348-7

Knowlton, K., Rotkin-Ellman, M., King, G., Margolis, H.G., Smith, D., Solomon, G., Trent, R., English, P., 2009. The 2006 California heat wave: impacts on hospitalizations and emergency department visits. Environ Health Perspect 117, 61–67. https://doi.org/10.1289/ehp.11594

Koman, P.D., Billmire, M., Baker, K.R., de Majo, R., Anderson, F.J., Hoshiko, S., Thelen, B.J., French, N.H.F., 2019. Mapping Modeled Exposure of Wildland Fire Smoke for Human Health Studies in California. Atmosphere 10, 308. https://doi.org/10.3390/atmos10060308

Liu, J.C., Wilson, A., Mickley, L.J., Dominici, F., Ebisu, K., Wang, Y., Sulprizio, M.P., Peng, R.D., Yue, X., Son, J.-Y., Anderson, G.B., Bell, M.L., 2017. Wildfire-specific Fine Particulate Matter and Risk of Hospital Admissions in Urban and Rural Counties. Epidemiology 28, 77–85. https://doi.org/10.1097/EDE.0000000000000556

Masri, S., Scaduto, E., Jin, Y., Wu, J., 2021. Disproportionate Impacts of Wildfires among Elderly and Low-Income Communities in California from 2000–2020. Int J Environ Res Public Health 18, 3921. https://doi.org/10.3390/ijerph18083921

Mazdiyasni, O., AghaKouchak, A., 2015. Substantial increase in concurrent droughts and heatwaves in the United States. Proc Natl Acad Sci U S A 112, 11484–11489. https://doi.org/10.1073/pnas.1422945112

McKinnon, K.A., Poppick, A., Simpson, I.R., 2021. Hot extremes have become drier in the United States Southwest. Nature Climate Change 11, 598–604.

Moftakhari, H.R., Salvadori, G., AghaKouchak, A., Sanders, B.F., Matthew, R.A., 2017. Compounding effects of sea level rise and fluvial flooding. Proc Natl Acad Sci USA 114, 9785–9790. https://doi.org/10.1073/pnas.1620325114

Nakayama Wong, L.S., Aung, H.H., Lamé, M.W., Wegesser, T.C., Wilson, D.W., 2011. Fine particulate matter from urban ambient and wildfire sources from California's San Joaquin

Valley initiate differential inflammatory, oxidative stress, and xenobiotic responses in human bronchial epithelial cells. Toxicology in Vitro 25, 1895–1905. https://doi.org/10.1016/j.tiv.2011.06.001

National Weather Service, N.W.S., 2014. Heat Index Equation [WWW Document]. URL https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml (accessed 2.21.22).

Nielsen, E.R., Herman, G.R., Tournay, R.C., Peters, J.M., Schumacher, R.S., 2015. Double Impact: When Both Tornadoes and Flash Floods Threaten the Same Place at the Same Time. Weather and Forecasting 30, 1673–1693. https://doi.org/10.1175/WAF-D-15-0084.1

O'Dell, K., Ford, B., Fischer, E.V., Pierce, J.R., 2019. Contribution of Wildland-Fire Smoke to US PM2.5 and Its Influence on Recent Trends. Environ. Sci. Technol. 53, 1797–1804. https://doi.org/10.1021/acs.est.8b05430

OEHHA, 2021. CalEnviroScreen 4.0.

Omernik, J.M., 1987. Ecoregions of the Conterminous United States. Annals of the Association of American Geographers 77, 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x

Perkins, S.E., Alexander, L.V., Nairn, J.R., 2012. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. Geophysical Research Letters 39. https://doi.org/10.1029/2012GL053361

Rao, K., Williams, A.P., Diffenbaugh, N.S., Yebra, M., Konings, A.G., 2022. Plant-water sensitivity regulates wildfire vulnerability. Nat Ecol Evol 6, 332–339. https://doi.org/10.1038/s41559-021-01654-2

Rey, G., Jougla, E., Fouillet, A., Pavillon, G., Bessemoulin, P., Frayssinet, P., Clavel, J., Hémon, D., 2007. The impact of major heat waves on all-cause and cause-specific mortality in France from 1971 to 2003. Int Arch Occup Environ Health 80, 615–626. https://doi.org/10.1007/s00420-007-0173-4

Rothfusz, L.P., 1990. The heat index "equation" (or, more than you ever wanted to know about heat index). NWS Tech. Attachment SR 90-23 2.

Safford, H.D., Paulson, A.K., Steel, Z.L., Young, D.J.N., Wayman, R.B., 2022. The 2020 California fire season: A year like no other, a return to the past or a harbinger of the future? Global Ecology and Biogeography 0, 1–21. https://doi.org/10.1111/geb.13498

Schwarz, L., Hansen, K., Alari, A., Ilango, S.D., Bernal, N., Basu, R., Gershunov, A., Benmarhnia, T., 2021. Spatial variation in the joint effect of extreme heat events and ozone on respiratory hospitalizations in California. PNAS 118. https://doi.org/10.1073/pnas.2023078118

Sherbakov, T., Malig, B., Guirguis, K., Gershunov, A., Basu, R., 2018. Ambient temperature and added heat wave effects on hospitalizations in California from 1999 to 2009. Environmental Research 160, 83–90. https://doi.org/10.1016/j.envres.2017.08.052

Thorsson, S., Rocklöv, J., Konarska, J., Lindberg, F., Holmer, B., Dousset, B., Rayner, D., 2014. Mean radiant temperature – A predictor of heat related mortality. Urban Climate, ICUC8: The 8th International Conference on Urban Climate and the 10th Symposium on the Urban Environment 10, 332–345. https://doi.org/10.1016/j.uclim.2014.01.004

U.S. Census Bureau, 2019. 2014-2018 American Community Survey 5-year estimates.

US EPA, O., 2021. Climate Change Indicators: Heat Waves [WWW Document]. URL https://www.epa.gov/climate-indicators/climate-change-indicators-heat-waves (accessed 12.29.21).

US EPA, O., 2014. NAAQS Table [WWW Document]. URL https://www.epa.gov/criteria-air-pollutants/naaqs-table (accessed 12.29.21).

Westerling, A.L., 2018. Wildfire Simulations for California's Fourth Climate Change Assessment: Projecting Changes in Extreme Wildfire Events with a Warming Climate. California's Fourth Climate Change Assessment, California Energy Commission.

Westerling, A.L., 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, 20150178. https://doi.org/10.1098/rstb.2015.0178

Weuve, J., Bennett, E.E., Ranker, L., Gianattasio, K.Z., Pedde, M., Adar, S.D., Yanosky, J.D., Power, M.C., 2021. Exposure to Air Pollution in Relation to Risk of Dementia and Related Outcomes: An Updated Systematic Review of the Epidemiological Literature. Environmental Health Perspectives 129, 096001. https://doi.org/10.1289/EHP8716

Ye, X., Arab, P., Ahmadov, R., James, E., Grell, G.A., Pierce, B., Kumar, A., Makar, P., Chen, J., Davignon, D., Carmichael, G.R., Ferrada, G., McQueen, J., Huang, J., Kumar, R., Emmons, L., Herron-Thorpe, F.L., Parrington, M., Engelen, R., Peuch, V.-H., da Silva, A., Soja, A., Gargulinski, E., Wiggins, E., Hair, J.W., Fenn, M., Shingler, T., Kondragunta, S., Lyapustin, A., Wang, Y., Holben, B., Giles, D.M., Saide, P.E., 2021. Evaluation and intercomparison of wildfire smoke forecasts from multiple modeling systems for the 2019 Williams Flats fire. Atmospheric Chemistry and Physics 21, 14427–14469. https://doi.org/10.5194/acp-21-14427-2021

Zhang, K., Rood, R.B., Michailidis, G., Oswald, E.M., Schwartz, J.D., Zanobetti, A., Ebi, K.L., O'Neill, M.S., 2012. Comparing exposure metrics for classifying 'dangerous heat' in heat wave and health warning systems. Environment International 46, 23–29. https://doi.org/10.1016/j.envint.2012.05.001

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R.M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M.D., Maraun, D., Ramos, A.M., Ridder, N.N., Thiery, W., Vignotto, E., 2020. A typology of compound weather and

climate events. Nat Rev Earth Environ 1, 333–347. https://doi.org/10.1038/s43017-020-0060-z

CHAPTER FOUR: Weather and Fire Intensity Predicts Building Damage from Wildfire in California

ABSTRACT

Increasingly large and intense wildfires have caused record-setting property loss in California over the last decade. Existing building vulnerability models study the effects of structural hardening on damage likelihood, but ignore the effects of time-varying weather and fire intensity. Yet, these variables influence the combustibility of structures and fire suppression tactics. Using a public database of home damages collected over a five year period by the California Department of Forestry and Fire Protection, we train a gradient boosted decision tree model to infer a home's combustion risk in the event of a wildfire. Notably, we impute time-varying risk factors using climate reanalysis data and remotely sensed detections of active wildfire to supplement more commonly studied building features. Our model is able to identify nearly 91% of homes that burned in a test sample of more than six thousand homes. Importantly, meteorological conditions, fire intensity and the recency of the wildfire ignition, had the largest influence on damage predictions. These findings suggest meteorology and wildfire behavior are more predictive than structural characteristics such as roof material and shape. Catastrophe models, like those used by insurers, should include dynamic, time-varying factors in wildfire risk modeling to improve property loss estimates. Our predictive model and performance metrics can support more accurate underwriting and empower individual homeowners to reduce their risk.

INTRODUCTION

Warming and expanded residential development in wilderness areas have resulted in unprecedented economic damages from wildfire in the United States (Westerling, 2016). The National Center for Environmental Information estimates $49 billion direct losses from

catastrophic wildfires in 2017 and 2018 - more than the preceding twenty-seven years' losses, combined (Buechi et al., 2021; National Centers for Environmental Information, 2022). California is particularly affected by catastrophic wildfire: in 2018 the state's insurance department reported a record-setting $12 billion in insured losses. These losses, stemming from larger, more-intense wildfires, have prompted increases in the premiums insurance companies charge and, in turn, have constrained access to critical protection from the damages of natural disasters (Nguyen & Noy, 2020).

Short of transferring risk to an insurer or relocating, homeowners may invest in construction "hardening" and maintenance that reduce the likelihood of their home combusting in a wildfire (Hedayati et al., 2023). Structure and site-level risk factors include roof material and siding material; the topographic and vegetational features that influence spread; as well as the home's location relative to firefighting crews and infrastructure (A. Syphard & Keeley, 2019). Studies have measured the marginal risk stemming from these different variables, however, many lack empirical data over a multi-year period. Additionally, these studies focus on small geographies that cannot be extrapolated to the diverse ecologies and building types in California, limiting their policy applications, for example the drafting of building codes and zoning maps (Hakes et al., 2017).

Notably, there is no study that empirically examines the effect of fire radiative power (intensity) and weather (temperature, air dryness and wind) on an individual building's vulnerability to combustion. The intensity of a fire is known to affect ember production, ember transport, the flame length and height as well as the likelihood of a fire's containment – all of which have implications for property damage. Weather conditions, including high-speed winds and low air moisture content, also influence the likelihood of wildland fuel or building timber

combustion and subsequent ember production (Seager et al., 2015). Beyond its physical effects, extreme fire weather threatens firefighter safety which reduces suppression activity and, in turn, increases the risk of damage to a home (Butler, 2014; Penney et al., 2019).

Since no empirical study measures risk response to changes in wildfire intensity or weather, current vulnerability assessments are, as we identify, omitting key interactions with weather and fire intensity that may cause them to misjudge risk. Moreover, these assessments premised on stationary variables related to building construction cannot capture changes in wildfire risk to building damage due to climate change. This study helps answer whether an increase in the frequency and intensity of hot and dry days and, accordingly, more intense wildfires will result in more property damage (McKinnon et al., 2021).

We accomplish this by overcoming key data limitations. First, we impute the estimated burn date of a home using satellite thermal anomaly detections to obtain the prevailing weather when a structure was threatened by wildfire. Additionally, we enhance a dataset of building characteristics with a property's historical canopy cover to train a prediction model that insurers and other public stakeholders can use to better understand home vulnerability throughout California. Finally, we describe how these insights can be integrated into current risk modeling frameworks for improved underwriting.

Hardening and Building Risk

To protect built structures from wildfire, mitigation experts recommend the use of fire resistant roofing and siding that "'harden" homes against induced heat or direct flame impingement, as well as meshes and screens that prevent embers from entering vents and eaves (CalFire, n.d.). Building codes now mandate many of these practices and researchers have

attributed code compliance with a 40% reduction in home ignition likelihood (Baylis & Boomhower, 2021). Beyond modifications to the structure itself, fire mitigation specialists target fuels within 100 feet of a building envelope that can ignite homes from conducted, radiated or convected heat. This "defensible space" not only reduces the likelihood of a home igniting from direct flame contact but also allows firefighters to set up safer tactical positions (A. D. Syphard et al., 2014).

Understanding home risk has implications not only for safety but also property affordability and value. Recently, insurance carriers have increased premiums and policy non-renewals in response to elevated fire risks and financial losses (California Department of Insurance, 2019; Dixon et al., 2018). These increased homeownership costs and risks have pushed some homeowners to relocate and reduce property values for both high and low income households (Chase & Hansen, 2021). To that end, the California's Department of Insurance recently established rules requiring insurers to reduce premiums in exchange for hardening measures homeowners undertake.

Quantifying risk reduction, no less translating these reductions monetarily, is difficult in practice, however. For one, myriad variables influence risk and researchers have yet to disentangle their interactions and heterogeneous effects. Furthermore, public data on property damage does not mention the burn date nor the prevailing conditions when a home was threatened by fire (CAL FIRE, 2023). Modeling to-date has leveraged historical claims data against different independent variables that influence the likelihood of a home igniting, albeit for small geographies over a short time window (Arrowsmith et al., 2021; Bhandary & Muller, 2009; A. D. Syphard et al., 2014) or at the building cluster level (Alexandre et al., 2016). These studies confirm increases in risk when canopies overhang roofs or when there is abundant

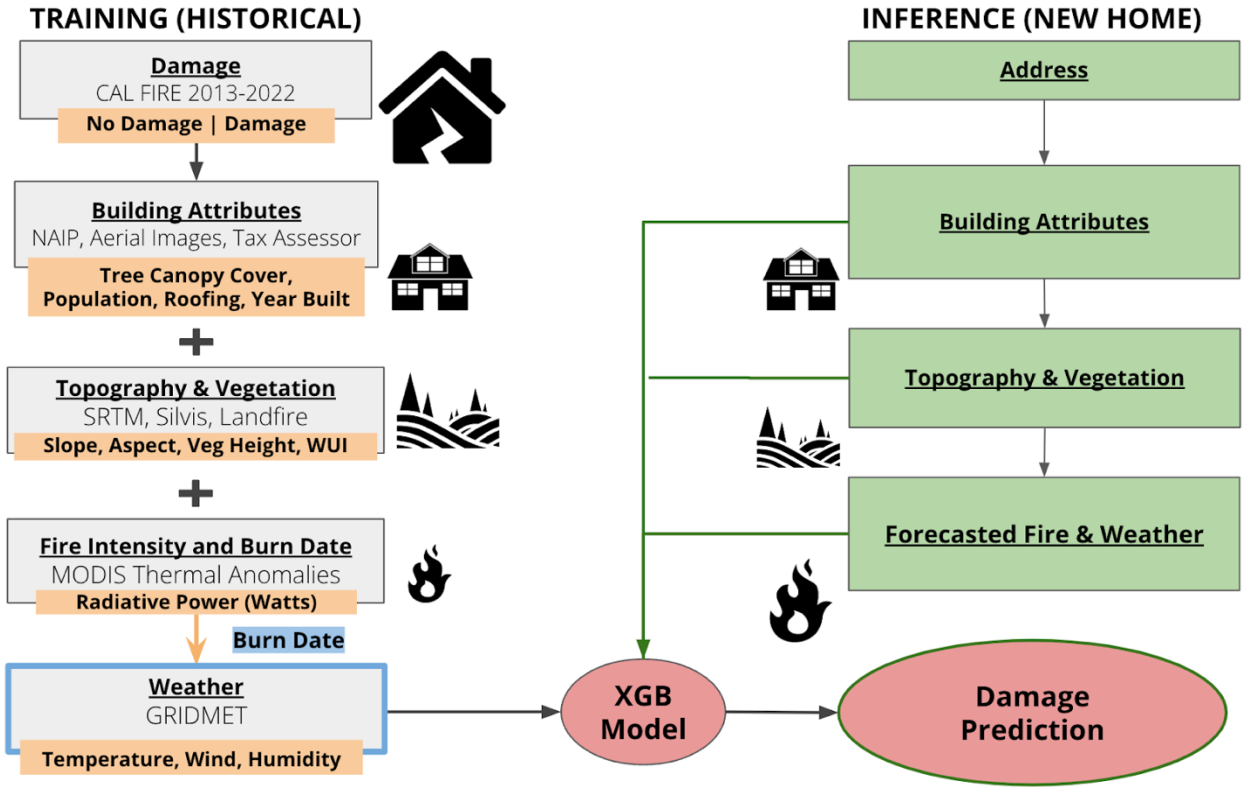
vegetation within 100ft - the defensible space - of a home. Still, none of these studies account for the impact of weather and fire intensity in shaping combustion risk. Last, no study has leveraged decision tree machine learning models that, unlike linear regression models, can account for data missingness, nonlinear interactions and multicollinearity across predictors.

METHODS

Training data

To build a model that can predict a home's likelihood of combustion we require data on observed historical damages (and non-damages) for homes that were inside a wildfire perimeter. Our model can then map associations between these observed damages and a series of predictor variables, or features, that previous studies have determined to influence risk (Table 4-1). Since the record of observed damages does not report the date when a home burns and, in turn, the weather and wildfire conditions at that time, we impute the burn date from satellite detections to then derive weather conditions. The final result is a model that can intake any set of predictors and produce either a likelihood (probability) or prediction (classification) of damage (Figure 4-1).

Figure 4-1. Modeling framework. Model training is performed on loss data enhanced with structure information documented by CAL FIRE, canopy cover from aerial imagery, "static" topography information as well as observed wildfire conditions. For inference, given any new home, similar datasets can be queried to obtain model inputs, including updated canopy conditions and, if applicable, forecasted fire dynamics.



We obtain historical damage data from CalFire's Damage Inspection (DINS) program spanning from wildfires spanning 2013 until 2022. This program, which began in the 1980s, dispatches trained inspection specialists to survey wildfire affected locations within the CAL FIRE reported perimeter on a daily basis according to a standardized protocol that includes post-collection quality control (Hawks, 2020). Damage outcomes are categorized according to a qualitative criteria (Appendix A) set by enumerators as 'No Damage', 'Affected (1-9%)', 'Minor (10-25%)', 'Major (26-50%)' and 'Destroyed (>50%)' (Wallingford, 2018). Notably there are few 'No Damage' homes in years 2013-2017 which CalFire confirmed as a data collection error; we therefore restrict our analysis to post-2017 wildfire events. Additionally, we filter the fire-

affected sites to structures with residential uses and exclude commercial facilities such as hospitals, churches, and schools given their unique structural characteristics and limited observational record. Finally, only four percent of our filtered data features intermediate levels of damage, we therefore group all non-zero levels of damage into a single "Damage" outcome and, in doing so, binarize the target variable.

Additional Features

Each outcome is trained against a series of covariates that previous studies and CalFire identify as correlating with home ignition. These include whether a building was wildfire code-compliant (all new construction in state managed areas starting in 2017) (Miller et al., 2020); the aspect, slope and elevation of the home which correspond to fire spread rates and fuel dryness and were derived from the 30-meter SRTM digital elevation model; the proximity to a CalFire facility, a proxy for response times; and the population density in a one kilometer-squared area surrounding the site to reflect building density, road network density and firefighter resources. Population density is based on the 2020 Gridded Population of World Version 4 (GPWv4) which interpolates census population data at a one kilometer-squared resolution every 5 years.

CalFire also records a series of structural characteristics based on forensic assessment, which include eaves treatments, fencing, roof material and siding material. Unfortunately, these fields have imbalanced coverage with respect to the outcome variable, and others, such as where and how a home ignited, have less than one thousand observations - we therefore omit these variables from the analysis. Instead, we use a dataset from Vexcel Imagery that extracts roof material and roof shape for homes within 10 meters of CAL FIRE's reported coordinates using

7.5cm imagery from plane-mounted cameras. This dataset also reports the number of observable roof vents, chimneys and building footprints within 30 and 100 feet of a structure.

We also calculate the defensible space around the home using the Vexcel Imagery where available and National Agricultural Imaging Program (NAIP) 60cm-1m resolution imagery everywhere else. Vexcel imagery is captured annually and NAIP imagery is captured every 3-years throughout the continental United States. The near-infrared and red imaging bands from the most recent image that precedes the ignition date of the fire are used to derive the Normalized Difference Vegetation Index (NDVI) which is then classified as vegetation based on absolute threshold of 0.25 - this cutoff value has strong recall (0.89) against tree detections from a deep learning segmentation model for Riverside, Los Angeles, San Diego and the San Francisco Bay Area - regions that capture considerable variability in California land cover (Zhang et al., 2022). This cutoff showed lower precision (0.43) due to the misclassification of grasses and shrubs as tree canopy. Accordingly, our defensible space estimates reflect both tree canopy and surface fuels. For measures of defensible space, circular distance buffers (10m, 30m, 100m) are drawn around the property's coordinates, as reported by CAL FIRE. Regional vegetation dynamics are captured by the University of Wisconsin's SILVIS Wildland-Urban Interface map, which delineates areas where urban and wildland land cover are either intermixed or interfacing (Radeloff et al., 2018).

The radiative power (megawatts) and location of a wildfire incident are based on NASA's MODIS Active Fire Product (MYD14A1.061) and the date of a wildfire event is imputed based on the nearest fire pixel from the MODIS Burned Area (MCD64A1 Version 6) monthly data product at a 500 meter resolution (Giglio, Louis et al., 2015). The difference between the estimated burn date and the CalFire reported ignition date is the presumptive day of

the incident in which a home burned. Two percent of the total dataset have estimated burn dates preceding the CalFire ignition date - in those cases, which average -1.6 days, the time elapsed was set to zero. There are also instances when the satellite pass-over occurs too late to detect a fire - in these cases fire intensity and weather variables are left null.

Weather characteristics during a wildfire event are derived from GRIDMET's daily 4 km-resolution data (Abatzoglou, 2013). We include daily maximum temperature, daily specific humidity, vapor pressure deficit and 10-meter wind speed; increased intensities of these three variables have physical associations with fire ignition and fire spread. Notably this dataset does not account for fire-generated extreme weather conditions and may be biased in areas with steep spatial gradients (Abatzoglou & Brown, 2012). All weather variables were binned into quintiles to simplify interpretation and to capture possible non-linear effects. Weather and fire intensity data were retrieved and processed in Google Earth Engine.

Gradient Boosted Model

We employ an extreme gradient boosting model for predicting the likelihood of combustion. The XGBoost algorithm uses decision trees or flowchart-like algorithms that develop predictive skill based on observed patterns and interactions across historical observations. These algorithms have demonstrated success in spatial prediction problems and offer better interpretability than deep-learning methods (Li, 2022). Additional advantages include automated parallelization, handling of missing features and multicollinear variables to best leverage available data. We use the XGBoost library wrapper from *scikit* to test a binary hinge objective function to predict a binary target variable representing building damage.

Next, we iteratively test different model training parameter combinations to identify that which produces the best model performance according to the objective function. Example parameters include the maximum tree depth and learning rate which regulate the number of separating branches in a tree and the contribution of any single tree to the final model prediction. These hyperparameters were "tuned" using a grid search that evaluated 192 unique combinations.

The model performance was trained using fire-incident-grouped four-fold cross validation and evaluated with an out of sample test set. That is, to assess the accuracy of the model's predictions we present it with new samples that were not involved in the model training and we compare its predictions for these never-before-seen samples to the observed outcomes. Our group-based sample splitting ensures that any single wildfire event's data is represented in either the training or testing datasets, but never both. This separation prevents overfitting where, for example, the algorithm could base predictions of new events based on their spatial correlation with outcomes it has already observed, not the underlying risk. An overfit model would overestimate the model's accuracy.

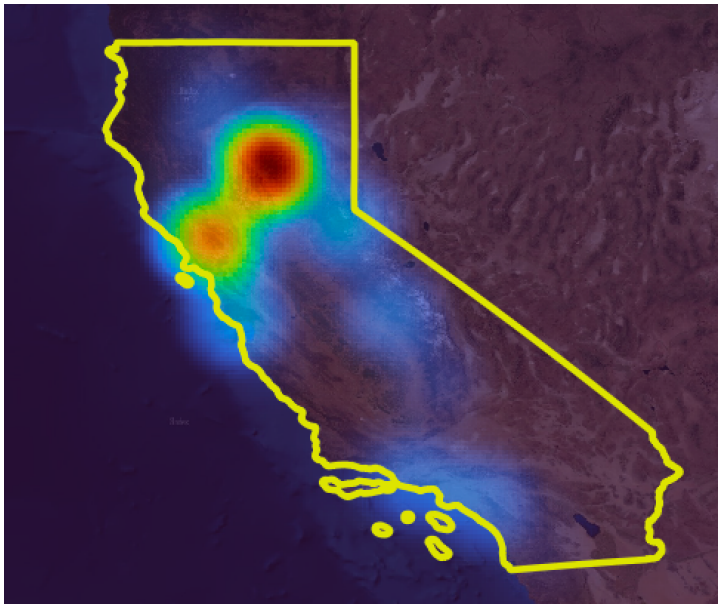
Table 4-1. Model Features with Descriptions

Continuous	Description	Source
Aspect	Degrees	SRTM 30 meter
Elevation	Meters	SRTM 30 meter
Slope	Meters / Kilometer	SRTM 30 meter
Population Density	Persons per km^2	GPW (Census)
Tree Canopy Area (10m buffer)	m^2	NAIP
Tree Canopy Area (30m buffer)	m^2	NAIP
Minimum Distance to Fire Station	m	CAL FIRE
Mean Fire Intensity	Watts/km^2	MODIS
Day of Fire	Count	MODIS
Year Built	Year	CAL FIRE
DSB 30/100	Percent of the region within (30/100) ft of the structure that is other buildings	VEXCEL
DST 30/100	Percentage of region within (30/100) ft of the structure that is tree coverage	VEXCEL
Area	Building area (geographic coordinates)	VEXCEL
Structure Tree Cover	Percentage of structure covered by trees	VEXCEL
Daily Maximum Temperature	Celsius	GRIDMET
Specific Humidity	Mass fraction	GRIDMET
Wind Speed (10 meter)	Meters per second	GRIDMET
Categorical		
WUI Classification	Intermix/Interface	SILVIS
Year Built >2008 Indicator	Yes/No	–
Burn Month	e.g. January	MODIS
Roof Shape	Hip: all sides slope down Gable: opposing sides slope down Flat: less than 2/12 slope	VEXCEL
Roof Material	Shingle, Tile, Metal, Shake, Gravel, Membrane	VEXCEL

RESULTS

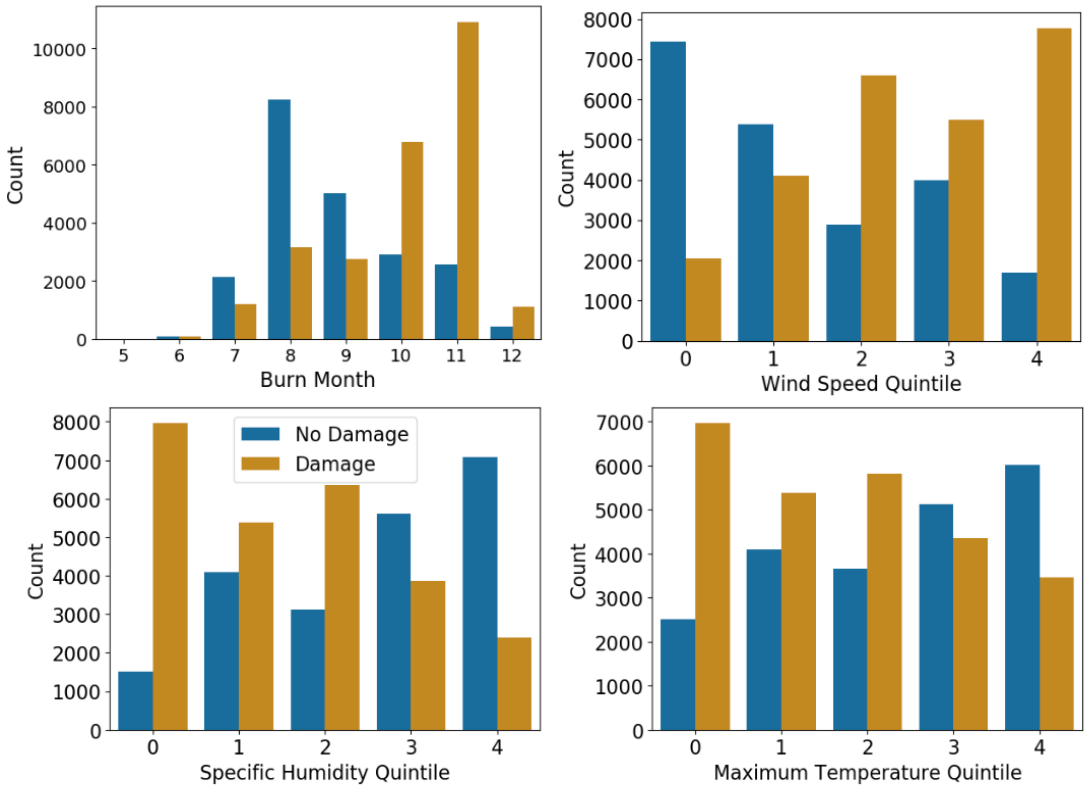
CalFire recorded 91,884 structures as being within or near 132 distinct wildfire perimeters between 2013-2022 - 60,665 of these structures were single family homes. After applying our filter criteria for residential structures that burned after 2017, a total of 49,810 homes, or 82% of all residential structures remain, highlighting the increased frequency and size of wildfires in recent years (Figure 4-2). Almost half of the wildfire events in the CalFire dataset count zero *un*burned homes and only six events count zero burned homes. While imbalances are plausible for a majority of small fires (the median recorded wildfire contains 144 affected homes), the Woolsey and Carr wildfire, which both burned over 1000 homes, omit any unburned records and are therefore included in our descriptive analysis but excluded from the model training and validation.

Figure 4-2. Damage Heat Map. Concentration of homes affected by wildfire (burned and unburned). Red regions show the highest density. As shown, the Camp Fire in Paradise (Northern California) is disproportionately represented in the dataset as the most destructive wildfire in California history.



Our imputation of a home's actual burn date suggests that 72% of affected homes burned within the first day of a fire's spread and approximately 94% within the first week. We also cross-tabulate damage outcomes against ordinal bins of weather intensity (Figure 4-3). The lowest quintile of specific humidity shows the largest ratio of "Damage" to "No Damage" homes; this ratio declines and then reverses as specific humidity increases. Wind speed also exhibits a clear positive trend between damage-to-undamaged ratio and intensity. Contrary to expectations, however, daily maximum temperature shows a general decline in the ratio of damaged homes, as do summer months.

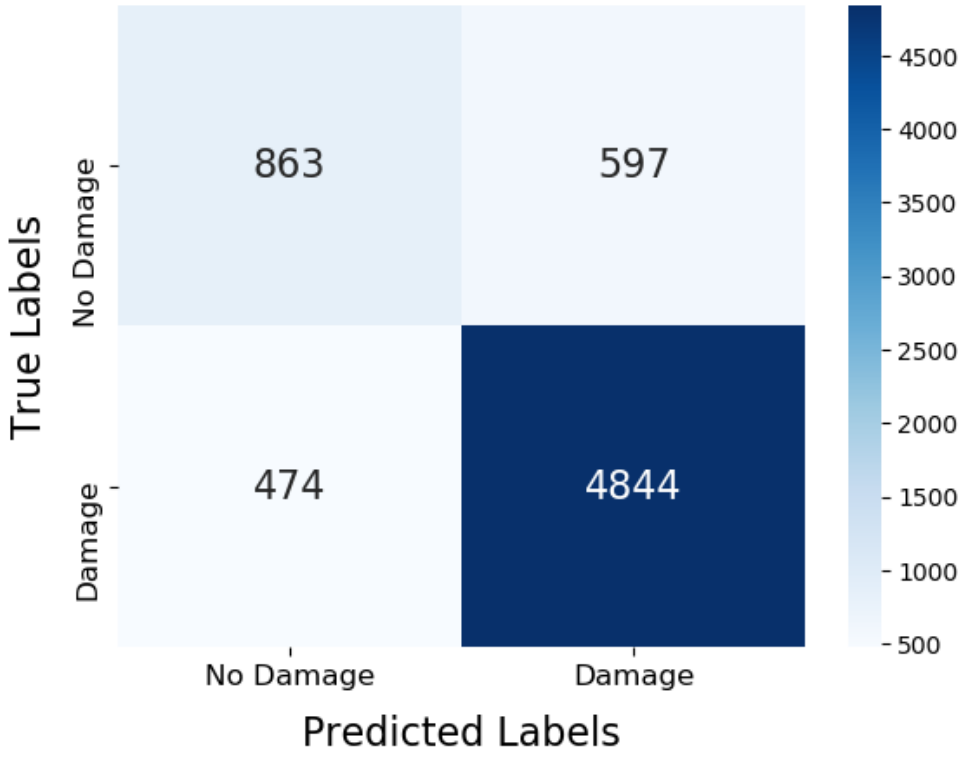
Figure 4-3. Cross Tabulations. Damage versus no-damage outcomes as a function of quintile-binned intensities for wind and specific humidity. Box and whiskers plot for the distribution of fire intensities across both damage groups.



Almost half of the DINS homes (n= 22,906) were captured by Vexcel's aerial imaging *before* the wildfire ignition date and featured a building within 10 meters of the CAL FIRE reported coordinates. Of these buildings, shingle roofs were most common (n=13,342), followed by metal (n=4430), tile (n=1596) and membrane (n=844). Metal roofs had a three-to-one ratio of undamaged to damaged homes, more than any other roof type. Other image-derived features, like the percent area occupied by buildings, show a strong positive correlation with population density (0.61) and negative correlation with vegetation abundance within the 100-foot defensible space (-0.4).

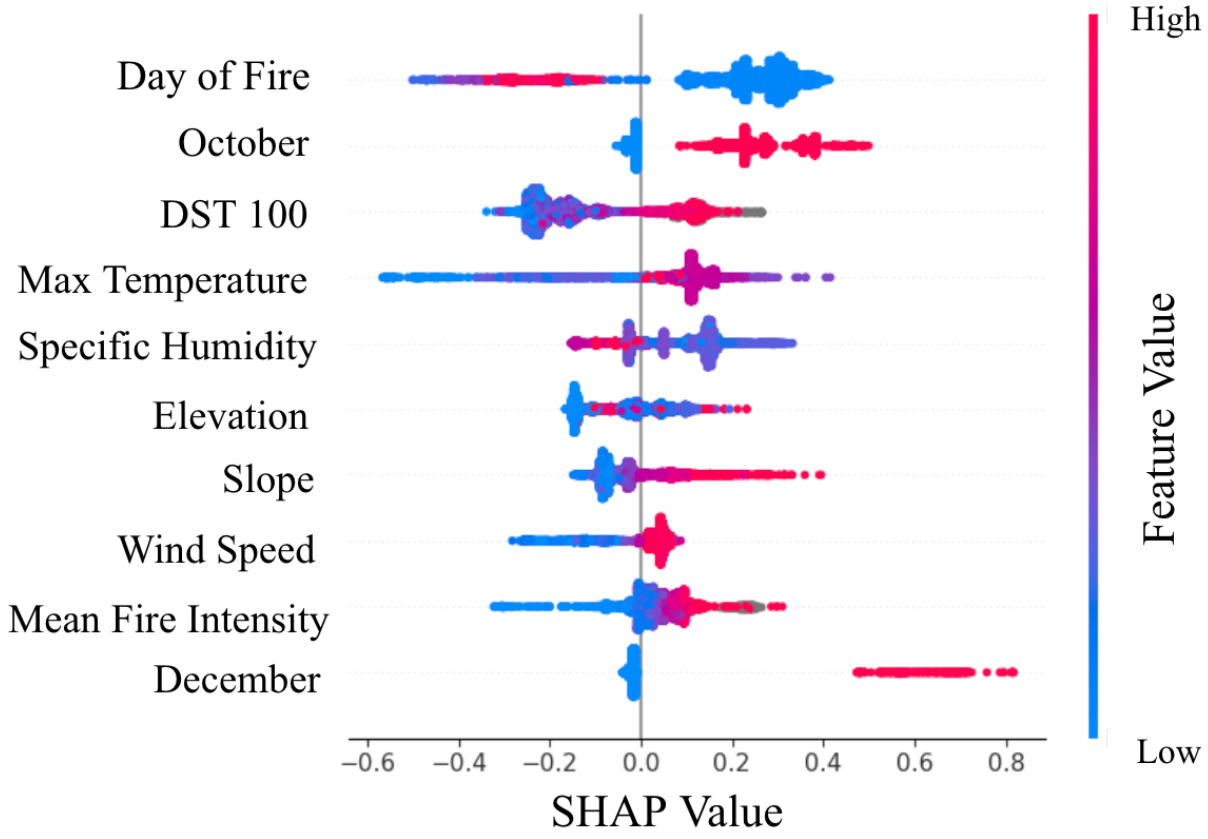
Our validation of the test set, comprising one-sixth of the data, for the XGBoost classifier using a binary hinge loss function produced a recall of damaged homes equal to 0.91 and precision of 0.89. That is, for our out-of-sample observations, the model correctly identified 91% of homes that actually burned and, of all its burn predictions, 89% of these were indeed correct (Figure 4-4). Conversely, 11% of the model's "burn" predictions were falsely positive. The F1 score, which measures the balance of precision and recall, was 0.90. This balance is important for ensuring that the model is neither too conservative nor too liberal in its prediction of damage. The model shows comparatively weaker recall and precision of homes that survive wildfires, equaling 0.60 and 0.64, respectively. The weaker performance is likely attributable to fewer no-damage observations in our test sample. However, future model iterations can apply weights to prioritize performance for either outcome.

Figure 4-4. Confusion Matrix. Predicted labels axis corresponds to the model's predicted class for each test set observation and true labels corresponds to the actual observed class in the historical data. Clockwise from the top left, quadrants show counts of true negatives, false positives, true positives and false negatives.



To identify which features most influence predictions we compute Shapley values (SHAP) that measure the average marginal contribution of each feature to the prediction across many different ordered combinations of features (Figure 4-5). Global SHAP values can be interpreted as the average sensitivity of the prediction to a feature being added or removed from the feature set, relative to the expected target value (Appendix B) (Messalas et al., 2019). In this case, the global expected target value represents the ratio of damaged to undamaged homes (p=0.5).

Figure 4-5. SHAP Feature Importance. The relationship between features and SHAP values, in order of importance. Positive SHAP-values correspond to increased odds of home damage. Higher magnitude features are redder and lower magnitude features are bluer, while categorical features are either blue (zero) or red (one). Each point represents the value of the specified feature for a single observation in the test data and the marginal SHAP value for that observation.



Out of 60 total feature values, our imputed, time-varying features of temperature, specific humidity, wind speed and fire intensity were all ranked among the ten most influential. The number of days that elapse post-ignition and the month of October, which are both correlated with extreme wind events, were the top two most influential features. The directional impacts of the top ten features match our original hypotheses (Figure 5) - dryer air, faster wind speeds and warmer temperatures increase the predicted likelihood of damage. Topographic variables of slope and elevation, which are correlated with weather, vegetation density, and fire spread

dynamics also influenced outcomes. Finally, as other studies report, more tree canopy within 100 feet of a building's envelope increases the predicted likelihood of damage. Notably these SHAP values do not imply causation nor do they imply prediction accuracy but instead measure the magnitude and direction of the feature's addition on the observed prediction.

DISCUSSION

Based on our imputation of burn date, the greatest wildfire risk to a home is within the first day of its ignition - over eighty percent of homes burn within this window. This is significant given that the average California wildfire in 2020 lasted over thirty days. Accordingly, the gradient boosting model leveraged the day of fire variable more than any other in its predictions of damage. Heightened risk in the first day of a fire is likely attributable to the destructiveness of wind driven events that can both cause ignition (e.g. conductor line ignitions) and drive rapid spread that forces firefighters to prioritize the evacuation of local residents as well as their personal safety (Keeley et al., 2021). Potentially for this reason, October and December, fall months when Santa Ana and Diablo winds commonly occur, were influential in the model's predictions.

As in previous studies, we find that the amount of vegetation within 100-feet of a building increases damage likelihood. This signal was most salient in the Vexcel imagery dataset that measures pre-burn canopy cover at a 7.5cm resolution around the detected building envelope. Conversely, the greater the total area of neighboring buildings within 100 feet of a home, the lower the estimated likelihood of damage, likely due to increased impervious surface. Interestingly, this outcome contrasts a relatively strong positive correlation between census measured population density and predicted damage. One takeaway is the importance of spatial

scale in analyzing risk - at the 100 foot resolution density reduces risk, however, at a 1 km resolution, density may drive contagion effects. Correlates of population density that may mitigate risk, such as firefighting resources and road connectivity, may therefore need to be included explicitly.

For the first time we include weather and wildfire intensity metrics in modeling home damage. Notably, our model assigned higher burn probabilities to homes near wildfire detections with strong radiative power and on days with lower specific humidity. This could be attributed to less effective firefighter suppression and/or increased risk to a home due to ember production and larger flame fronts. Importantly, none of the building structure characteristics that we included in our analysis - including roof vents, roof material and roof shape - were ranked in the top 10 most important features. Only the home's construction year showed a meaningful negative correlation with damage likelihood, but even then, its contribution is one-tenth the marginal contribution of the most important feature, day of fire. Our argument is not that structural characteristics are unimportant - this study does not consider other meaningful variables like siding material and treatments to vents, eaves and gutters that limit ember intrusion. However, our results show a predominance of intense weather and fire dynamics over roofing in predicting damage. This should motivate further research on the interactions between hardening measures, fire intensity and weather, and possible heterogeneity in their efficacy.

There are several limitations to our approach. For one, we assume accuracy in the damage, building attributes and location that CalFire reports. Additionally, homes that are more accessible to enumerators may have certain characteristics that are correlated with susceptibility to damage. While our study is the first in the literature to leverage a statewide database of aerial imagery over a five year time period to improve the availability and accuracy of pre-burn

characteristics, there are still many homes for which imagery was unavailable or homes were undetectable; this may depress the feature importances of these variables.

Our estimates of fire intensity, location and burn date are also prone to error in satellite measurement and imprecision stemming from the semi-daily overpass that may result in missed detections or deviations from the actual conditions that threatened a home, especially in areas with "light" fuels that extinguish quickly. Future studies can use burn scar or geostationary satellites to fill these gaps. Finally, our study does not explicitly consider regional wildfire risk mitigation like fuel breaks and forest thinning. However, these interventions would likely mitigate fire intensity and are therefore implicitly reflected in our model.

CONCLUSION

Here we present an extreme gradient boosting model that can predict home ignition with significant accuracy. Notably, weather and fire intensity are large contributors to the likelihood of home damage. We find suggestive evidence of firefighting resources and extreme wind events influencing outcomes through indirect measures of the time elapsed since ignition and the month of year in which a home burned, respectively. These features' importances outweighed building variables related to roof material, shape, and year of construction, suggesting a dominating influence of fire and weather dynamics on the predictiveness of damage.

Our results can be readily integrated into current Monte Carlo catastrophe modeling frameworks that rely on simulated wildfire footprints for estimating risk. When available, modeled weather and fuel conditions and the fire dynamics they influence should be inputted into models of building vulnerability to capture variation in home-level risk due to an area's wildfire ecology. In doing so, models can begin to identify areas that are susceptible to intense or

rapidly spreading flame fronts and, furthermore, can predict how climate change will affect the vulnerability of buildings over the next several decades.

By illustrating the possibility for damage prediction as well as the quantifiability of marginal risk reductions, our study can support market efforts to reward homeowners for hardening their home with policy renewals and possible premium discounts. As wildfire risk increases across California, precise underwriting will be critical for incentivizing private actions to reduce risk and for insurance carriers to continue operating and providing risk transfer to those in harm's way.

APPENDIX 4

Appendix 4A. Damage Criteria Published by CAL FIRE.

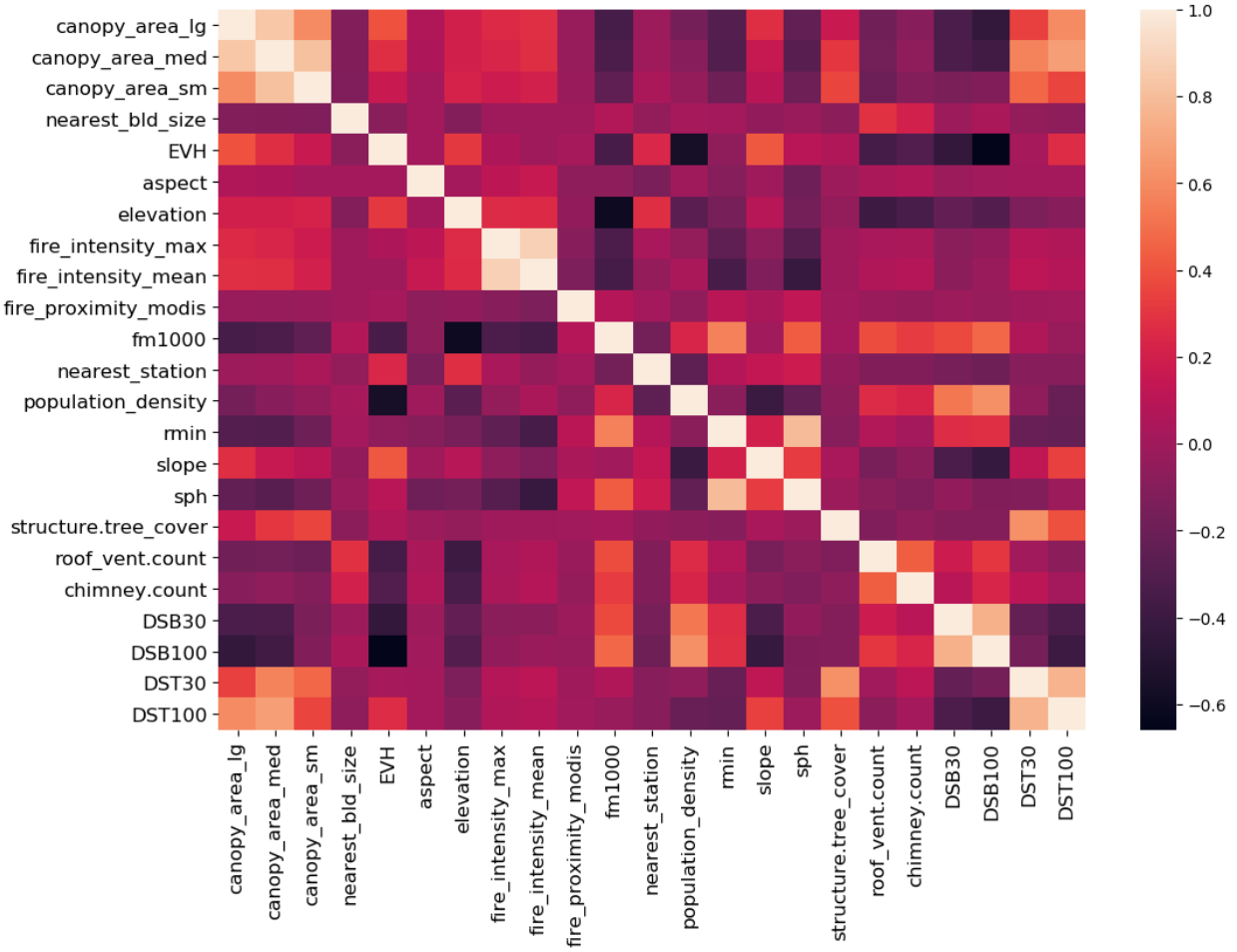
'Affected (1-9%)'	Minimal damage to the exterior and/or contents of the building. Building is habitable/usable and requires mostly cosmetic repairs.	Partially damaged shingles or siding, but roof structure is intact. Cosmetic damages such as paint discoloration, blistering or melted siding. Broken windows. Gutter damage. Damage to an attached structure like a deck, porch, carport, or patio cover.
'Minor (10-25%)'	Encompasses a wide range of damage that does not affect the structural integrity of the building. Building is not habitable/usable.	Nonstructural damage to roof components (e.g. roof covering, fascia board, soffit, flashing, and skylight). Nonstructural damage to the interior wall components (e.g. drywall and insulation). Nonstructural damage to exterior components (e.g. door and windows. Substantial damage to exterior covering (e.g. siding, vinyl or stucco). Damage to mechanical components (e.g. furnace, boiler, water heater, HVAC, etc.).
'Major (26-50%)'	A building that has sustained significant structural damage and requires extensive repairs. Building is not habitable/usable.	Failure or partial failure of structural elements to include rafters, ceiling joists, ridge boards, etc. Failure or partial failure to structural elements of the walls to include framing, sheathing, etc.
'Destroyed (>50%)'	The building is a total loss, or damaged to such an extent that repair is not feasible.	Complete failure to major components (foundation, walls, roof, etc.). Two or more walls destroyed and roof substantially damaged. Only the foundation remains. The building will have to be torn down and rebuilt as it is unsafe.

Appendix 4B. Global SHAP Value for All Features

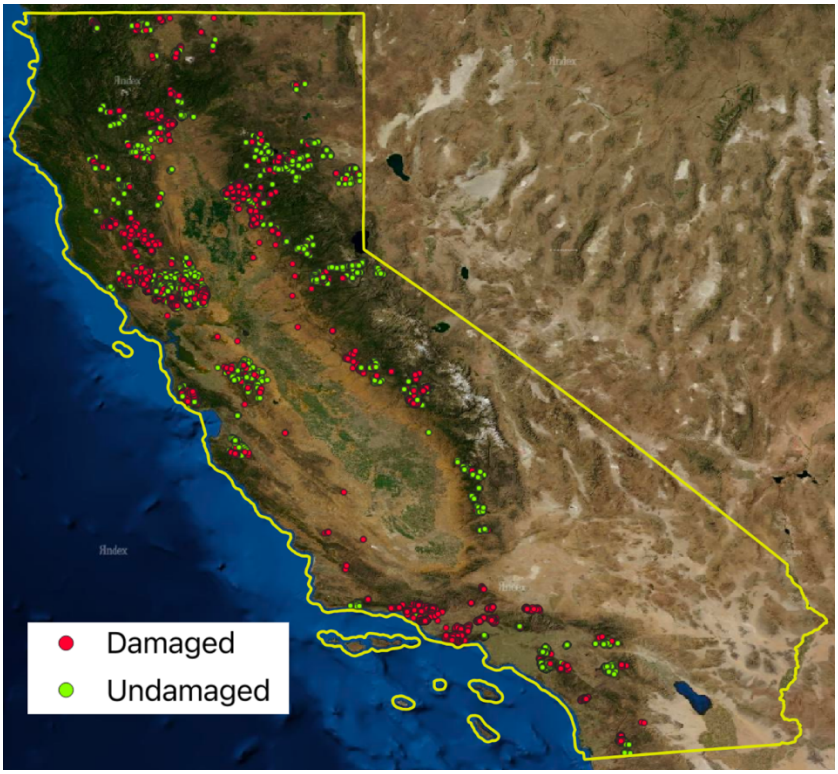
Absolute Rank	Feature	SHAP Absolute
1	Day of Fire	0.264
2	October	0.206
3	DST100	0.169
4	Daily Maximum Temperature	0.125
5	Specific Humidity	0.111
6	Elevation	0.076
7	Slope	0.066
8	Wind Speed	0.060
9	Mean Fire Intensity	0.057
10	December	0.052
11	Vapor Pressure Deficit	0.048
12	Population Density	0.043
13	November	0.037
14	Construction Year	0.032
15	NAIP Canopy 100m	0.029
16	CAL FIRE Facility Distance	0.028
17	DSB100	0.022
18	NAIP Canopy 30m	0.010
19	DST30	0.010
20	Building Area	0.009
21	Vegetation Height	0.007
22	Vegetation Height NaN	0.005
23	August	0.005
24	July	0.003
25	Roof Vent Count	0.002
26	Pre-Code Construction	0.002
27	Chimney Count	0.002
28	Structure Tree Cover	0.002
29	Aspect	0.001
30	Low_Dens_Intermix	0.001
31	WUI NaN	0.001
32	Very_Low_Dens_Veg	0.001
33	Med_Dens_Interface	0.001
34	NAIP Canopy 5m	0.001
35	DSB30	0.001
36	September	0.001
37	Tile Roof	0.001
38	Med_Dens_Intermix	0.001
39	High_Dens_Interface	0.000
40	Low_Dens_Interface	0.000
41	Gable Roof	0.000

42	Med_Dens_NoVeg	0.000
43	Very_Low_Dens_NoVeg	0.000
44	Uninhabited_Veg	0.000
45	NaN Roof Material	0.000
46	Hip Roof	0.000
47	June	0.000
48	Metal Roof	0.000
49	Uninhabited_NoVeg	0.000
50	Membrane Roof	0.000
51	Shingle Roof	0.000
52	Up-to-Code Construction	0.000
53	Flat Roof	0.000
54	Low_Dens_NoVeg	0.000
55	Shake Roof	0.000
56	Water	0.000
57	High_Dens_NoVeg	0.000
58	High_Dens_Intermix	0.000
59	May	0.000
60	Gravel Roof	0.000

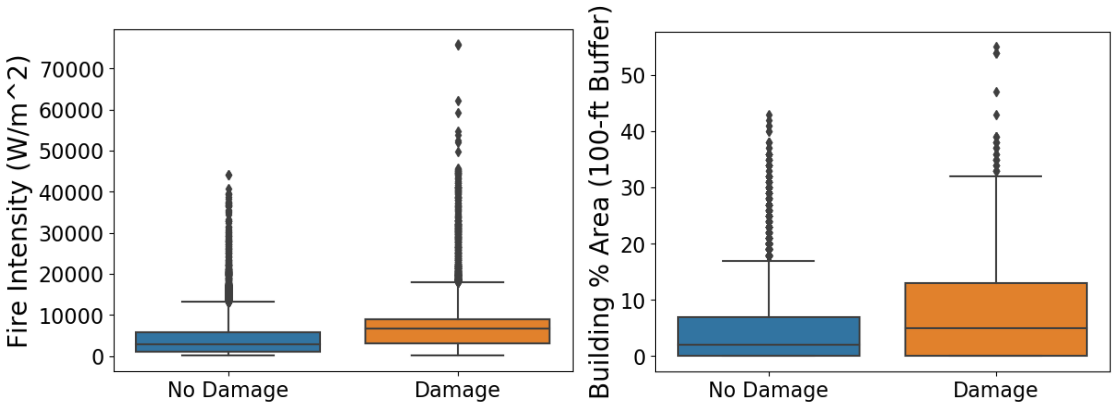
Appendix 4C. Correlation matrix.



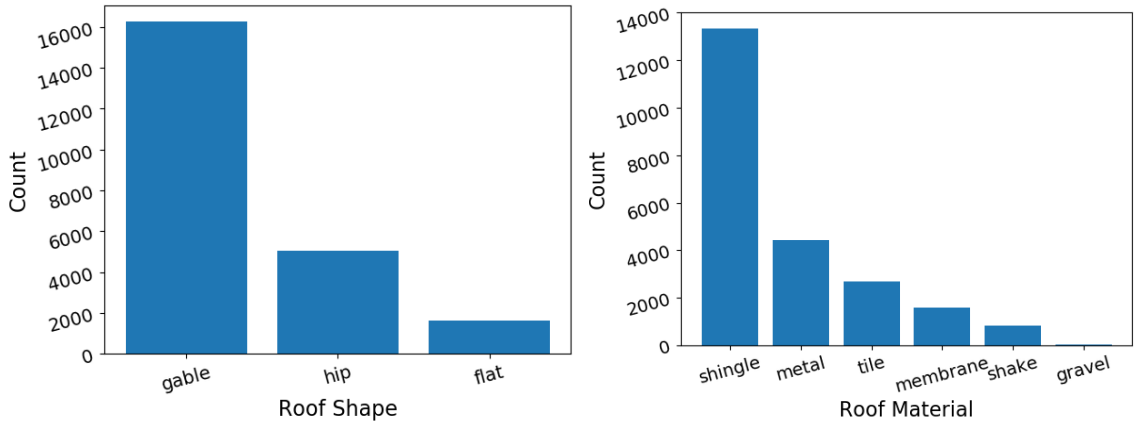
Appendix 4D. Spatial distribution of homes recorded by CAL FIRE. While many dots are overlapping, this points to the relatively equal spatial distribution of burned and unburned homes.



Appendix 4E. Bivariate Plots of Neighboring Building Area and Fire Intensity Versus Damage



Appendix 4F. Roof Shape and Material Frequency Distribution



BIBLIOGRAPHY

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131. https://doi.org/10.1002/joc.3413

Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, *32*(5), 772–780. https://doi.org/10.1002/joc.2312

Alexandre, P. M., Stewart, S. I., Keuler, N. S., Clayton, M. K., Mockrin, M. H., Bar-Massada, A., Syphard, A. D., & Radeloff, V. C. (2016). Factors related to building loss due to wildfires in the conterminous United States. *Ecological Applications*, *26*(7), 2323–2338. https://doi.org/10.1002/eap.1376

Arrowsmith, E., Fortier, F. D., & Cope, A. (2021). *Wildfire Fuel Management and Risk Mitigation*.

Baylis, P., & Boomhower, J. (2021). *Building codes and community resilience to natural disasters*. 31.

Bhandary, U., & Muller, B. (2009). Land use planning and wildfire risk mitigation: An analysis of wildfire-burned subdivisions using high-resolution remote sensing imagery and GIS data. *Journal of Environmental Planning and Management*, *52*(7), 939–955. https://doi.org/10.1080/09640560903181147

Buechi, H., Weber, P., Heard, S., Cameron, D., & Plantinga, A. J. (2021). Long-term trends in wildfire damages in California. *International Journal of Wildland Fire*, *30*, 757–762.

Butler, B. W. (2014). Wildland firefighter safety zones: A review of past science and summary of future needs. *International Journal of Wildland Fire*, *23*(3), 295–308. https://doi.org/10.1071/WF13021

CAL FIRE. (2023). *CAL FIRE Wildfire Damage Inspection* (Version 2023) [Data set].

CalFire. (n.d.). *Hardening Your Home*. Ready for Wildfire. Retrieved April 30, 2023, from https://www.readyforwildfire.org/prepare-for-wildfire/get-ready/hardening-your-home/

California Department of Insurance. (2019, August 20). *New Data Shows Insurance Is Becoming Harder to Find as a Result of Wildfires*. http://www.insurance.ca.gov/0400-news/0100-press-releases/2019/release063-2019.cfm

Chase, J., & Hansen, P. (2021). Displacement after the Camp Fire: Where are the Most Vulnerable? *Society & Natural Resources*, *34*(12), 1566–1583. https://doi.org/10.1080/08941920.2021.1977879

Dixon, L., Tsang, F., & Fitts, G. (2018). *The Impact of Changing Wildfire Risks on California's Residential Insurance Market:* [Product Page].

https://www.rand.org/pubs/external_publications/EP67670.html

Giglio, Louis, Justice, Christopher, Boschetti, Luigi, & Roy, David. (2015). *MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006* [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD64A1.006

Hakes, R. S. P., Caton, S. E., Gorham, D. J., & Gollner, M. J. (2017). A Review of Pathways for Building Fire Spread in the Wildland Urban Interface Part II: Response of Components and Systems and Mitigation Strategies in the United States. *Fire Technology*, *53*(2), 475–515. https://doi.org/10.1007/s10694-016-0601-7

Hawks, S. (Director). (2020). *Damage inspection and research implications on the California structure ignition problem*. https://www.youtube.com/watch?v=XMl7ylRDrIE

Hedayati, F., Quarles, S. L., & Hawks, S. (2023). *Wildland Fire Embers and Flames: Home Mitigations That Matter*. Insurance Institute for Business & Home Safety.

Keeley, J. E., Guzman-Morales, J., Gershunov, A., Syphard, A. D., Cayan, D., Pierce, D. W., Flannigan, M., & Brown, T. J. (2021). Ignitions explain more than temperature or precipitation in driving Santa Ana wind fires. *Science Advances*. https://doi.org/10.1126/sciadv.abh2262

Li, Z. (2022). Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*, *96*, 101845. https://doi.org/10.1016/j.compenvurbsys.2022.101845

McKinnon, K. A., Poppick, A., & Simpson, I. R. (2021). Hot extremes have become drier in the United States Southwest. *Nature Climate Change*, *11*(7), 598–604.

Messalas, A., Kanellopoulos, Y., & Makris, C. (2019). Model-Agnostic Interpretability with Shapley Values. *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1–7. https://doi.org/10.1109/IISA.2019.8900669

Miller, R. K., Field, C. B., & Mach, K. J. (2020). Factors influencing adoption and rejection of fire hazard severity zone maps in California. *International Journal of Disaster Risk Reduction*, *50*, 101686. https://doi.org/10.1016/j.ijdrr.2020.101686

National Centers for Environmental Information. (2022). *U.S. Billion-dollar Weather and Climate Disasters, 1980—Present (NCEI Accession 0209268)*. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:0209268

Nguyen, C. N., & Noy, I. (2020). Measuring the impact of insurance on urban earthquake recovery using nightlights. *Journal of Economic Geography*, *20*(3), 857–877. https://doi.org/10.1093/jeg/lbz033

Penney, G., Habibi, D., & Cattani, M. (2019). Firefighter tenability and its influence on wildfire suppression. *Fire Safety Journal*, *106*, 38–51.

https://doi.org/10.1016/j.firesaf.2019.03.012

Radeloff, V. C., Helmers, D. P., Kramer, H. A., Mockrin, M. H., Alexandre, P. M., Bar-Massada, A., Butsic, V., Hawbaker, T. J., Martinuzzi, S., Syphard, A. D., & Stewart, S. I. (2018). Rapid growth of the US wildland-urban interface raises wildfire risk. *Proceedings of the National Academy of Sciences*, *115*(13), 3314–3319. https://doi.org/10.1073/pnas.1718850115

Seager, R., Hooks, A., Williams, A. P., Cook, B., Nakamura, J., & Henderson, N. (2015). Climatology, Variability, and Trends in the U.S. Vapor Pressure Deficit, an Important Fire-Related Meteorological Quantity. *Journal of Applied Meteorology and Climatology*, *54*(6), 1121–1141. https://doi.org/10.1175/JAMC-D-14-0321.1

Syphard, A. D., Brennan, T. J., & Keeley, J. E. (2014). The role of defensible space for residential structure protection during wildfires. *International Journal of Wildland Fire*, *23*(8), 1165. https://doi.org/10.1071/WF13158

Syphard, A., & Keeley, J. (2019). Factors Associated with Structure Loss in the 2013–2018 California Wildfires. *Fire*, *2*(3), 49. https://doi.org/10.3390/fire2030049

Wallingford, N. (2018). *Camp Incident Damage Inspection Report*. CalFire. https://www.nist.gov/system/files/documents/2020/11/16/2018%20Camp%20Incident%20DINS%20Final%20Report.pdf

Westerling, A. L. (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. https://doi.org/10.1098/rstb.2015.0178

Zhang, Y., Chen, G., Myint, S. W., Zhou, Y., Hay, G. J., Vukomanovic, J., & Meentemeyer, R. K. (2022). UrbanWatch: A 1-meter resolution land cover and land use database for 22 major cities in the United States. *Remote Sensing of Environment*, *278*, 113106. https://doi.org/10.1016/j.rse.2022.113106