

UC Merced

UC Merced Electronic Theses and Dissertations

Title

Spatiotemporal burn severity, emissions, and health impact analyses of California wildfires

Permalink

<https://escholarship.org/uc/item/1gg2j0ss>

Author

Xu, Qingqing

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

Spatiotemporal burn severity, emissions, and health impact analyses
of California wildfires

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

In

Environmental Systems

By

Qingqing Xu

Committee in charge:

Professor Anthony LeRoy Westerling, Chair
Professor Roger Bales
Professor Erin Hestir
Dr. Christine Wiedinmyer

2022

Copyright

Qingqing Xu, 2022

All rights reserved

The Dissertation of Qingqing Xu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Roger Bales

Erin Hestir

Anthony LeRoy Westerling, Chair

Christine Wiedinmyer

University of California, Merced

2022

Table of Contents

List of Tables	vi
List of Figures.....	vii
Acknowledgements	ix
Curriculum Vita.....	x
Abstract of the Dissertation	xii
1 Introduction.....	1
References.....	3
2 Wildfire Burn Severity and Emissions Inventory: An example implementation over California	6
Abstract.....	6
2.1 Introduction.....	6
2.2 Methods.....	8
2.3 Results and Discussion.....	14
2.3.1 Burn Severities	14
2.3.2 Emissions Estimates.....	16
2.3.3 Daily Emissions.....	18
2.3.4 Limitations and Uncertainties	19
2.4 Conclusions.....	20
References.....	21
3 Spatial and Temporal Patterns of Wildfire Burn Severity and Biomass Burning-induced Emissions in California	25
Abstract.....	25
3.1 Introduction.....	25
3.2 Data and Methods.....	26
3.2.1 Study Area.....	26
3.2.2 Data	27
3.2.3 Statistical Estimation of Annual Trends.....	27
3.3 Results and Discussions.....	28
3.3.1 The Spatial Distribution of Burn Severity and Emissions	28
3.3.2 Annual Severity and Emissions.....	29
3.3.3 Emission Seasonality.....	32
3.3.4 Extreme Events	36
3.3.5 Limitations	37
3.4 Conclusions.....	38
References.....	39

4 Development of Methods to Predict Respiratory Hospitalization with Wildfire Emissions Inventory Data	41
Abstract.....	41
4.1 Introduction.....	41
4.2 Data and Methods.....	43
4.2.1 PM _{2.5} Emissions	44
4.2.2 Health Data.....	46
4.2.3 Covariate Data.....	46
4.2.4 Statistical Analysis	48
4.3 Results and Discussions	49
4.3.1 Descriptive Statistics	49
4.3.2 Analysis of Logistic Generalized Additive Model for Predicting Presence of Respiratory Hospitalization	50
4.3.3 Analysis of Negative Binomial Generalized Additive Model for Predicting Numbers of Respiratory Hospitalization.....	53
4.3.4 Discussions.....	55
4.4 Conclusions.....	56
References.....	56
Appendix.....	60

List of Tables

Table 2. 1 Emission factors (g species emitted per kg biomass burned) assigned to grasslands, shrublands, and forests.	13
Table 2. 2 Pixel level burn severity classification comparison of MTBS and WBSE (unit: %)	16
Table 2. 3 Summary statistics of emissions during 1984-2020.	17
Table 4. 1 Continuous variables tested and selected in GAM models	51
Table 4. 2 Frequency of having large wildfire induced PM _{2.5} emissions in semi-months and total emissions aggregated over the years of 1984-2019	50
Table 4. 3 Parametric coefficients in GAM models	52

List of Figures

Figure 2. 1 Diagram of Wildfire Burn Severity and Emissions Inventory (WBSE) inventory steps and datasets.....	9
Figure 2. 2 Diagram of composite the differenced Normalized Burn Ratio (dNBR) calculation with Google Earth Engine.....	11
Figure 2. 3 A decision tree framework to apply models to convert dNBR/NBR values to CBI values (adapted from Picotte et al 2021).....	12
Figure 2. 4 A diagram of assigning day-of-burning with the fire progression raster tool	14
Figure 2. 5 The annual area burned in unchanged (fire refugia within perimeters), low, moderate, high severity, and grass burn categories of large wildfires(>404 ha) in the California region during 1984-2020	15
Figure 2. 6 Burn severity map of all large wildfires(>404 ha) burned in each California air basin during 1984-2020. Some burned areas overlap since the same place could be burned more than once between 1984 and 2020.....	15
Figure 2. 7 Annual fire PM _{2.5} emissions (Tg yr ⁻¹), over 1984-2020 in California, from five global fire emissions inventories (GFEDv4s, FINNv1.5, GFASv1.2, QFEDv2.5r, FEERv1.0-G1.2), two regional (CARB and WFEIS) inventories, and WBSE	18
Figure 2. 8 Average daily PM _{2.5} emissions aggregated by day of the year from 2002 to 2020. The dashed line represents the value of average daily PM _{2.5} emissions from all other sectors in California over 2002-2017; Bars colored in red and cyan indicate when average daily emissions from wildfires are above and below the average emissions from all other sectors, respectively	19
Figure 2. 9 Total emissions from 2002-2019 aggregated by day of year compared to daily emissions in 2020.....	19
Figure 3. 1 The spatial distribution of burn severity in 13 ecoregions and fire induced PM _{2.5} emissions in 15 air basins during 1984-2020 in California for all fires over 404 ha. Some burned area and emissions 30m pixels can overlap since the same place could be burned more than once between 1984 and 2020	28
Figure 3. 2 Annual area burned in unchanged, low, moderate, high severity, and grassland of large wildfire Fires (>404 ha) during 1984-2020. The blue indicates annual large wildfire numbers	29
Figure 3. 3 Annual area burned in shrub, low elevation forest (< 5500 ft), moderate elevation forest (5500-7500 ft), high elevation forest (> 7500 ft) in low (top), moderate (middle), and high (bottom) severity during 1984-2020. Grey bars indicated areas ignored by WBSE when classifying LANDFIRE Existing Vegetation Type to coarse land cover types, and not classified as grass burn as the burn severity. Those areas are mainly developed or vegetation types that were not considered in the emission model	30
Figure 3. 4 Annual PM _{2.5} emissions from low, moderate, and high severities during 1984-2020.....	31
Figure 3. 5 Annual PM _{2.5} emissions from grassland, shrubland, and forest during 1984-2020.....	32
Figure 3. 6 Annual PM _{2.5} emissions from the area burned in low, moderate, high severity in low elevation forest (< 5500 ft), mid elevation forest (5500-7500 ft), and high elevation forest (> 7500 ft) during 1984-2020.....	32

Figure 3. 7 Annual area burned in unchanged (fire refugia within perimeters), low, moderate, and high severity, and grass burn categories of large wildfires (>404 ha) in California during 1984-2020. Blue line shows the total numbers of large wildfires burned in each month during the whole period.....	33
Figure 3. 8 Cumulative total emissions by month during the periods of 1984-2001, 2002-2019, and the year of 2020.....	34
Figure 3. 9 The spatial distribution of PM _{2.5} emissions from large wildfires ignited during the summer (top) and autumn (bottom) over the 1984-1990 (1980s), 1991-200(1990s), 2001-2010 (2000s), and 2011-2020 (2010s) periods.....	35
Figure 3. 10 Monthly total emissions from grass, shrub, and forest ecosystems during 1984-2020	35
Figure 3. 11 Fire numbers (bars) and emission contribution pareto graph (line). Large wildfires are classified into six groups based on size (less than 1000 ha, 1000 - 5000 ha, 5000 -10000ha, 10000 - 20000 ha, 20000 - 40468 ha, and over 40468 ha (Megafires)) .	36
Figure 3. 12 Annual PM _{2.5} emissions from largest fire, megafires, and other large wildfires (> 404 ha) during 1984-2020. Black dots show megafire numbers	37
Figure 4. 1 Schematic of our approach to predict health outcomes with wildfire emissions inventory data	42
Figure 4. 2 ZIP Code Tabulation Areas (ZCTAs) in California, colored by 2019 population data from American Community Survey (ACS)	43
Figure 4. 3 California air basins. Dots show the air monitors, colored by operation start year for monitoring particulate matter data (EPA, 2022)	45
Figure 4. 4 Observed counts of all respiratory hospitalizations during 1984-2020.....	49
Figure 4. 5 Observed counts of monthly respiratory hospitalizations aggregated over the years 1984-2020.....	49
Figure 4. 6 Observed counts of daily respiratory hospitalizations during 1984-2020. Cyan color indicates weekdays and red color indicates days as holidays or weekends.....	49
Figure 4. 7 Partial effect plots of logistic GAM output. The output on y axis is the probability of the outcome if all other variables were at their average value. Dotted lines show the confidence intervals	53
Figure 4. 8 Partial effect plots of negative binomial GAM output. The output on y axis is the log value of hospitalization numbers. Dotted lines show the confidence intervals	54
Figure 4. 9 Predicted and observed hospitalization numbers for the testing data.....	55

Acknowledgements

I have been fortunate to have the support from my family, friends, colleagues, and the university completing my Ph.D. program and this dissertation. This journey is an exceptional experience and memorable.

First, I would like to express my deepest gratitude to my advisor, Dr. Anthony LeRoy Westerling. I was encouraged and inspired by his academia experience to pursue a doctoral degree in Environmental Systems and develop research interests in wildfires. This journey has not been easy and won't be possible without his guidance. He has been supportive to my research interests and providing insightful suggestions to the overall pictures of this thesis.

Many thanks Andrew Notohamiprodjo, Dr. Christine Wiedinmyer, Joshua J Picotte, Dr. Sean A. Parks, Dr. Matthew D. Hurteau, Dr. Miriam E Marlier, Dr. Crystal A. Kolden, Dr. Jonathan A Sam, Dr. W Jonathan Baldwin, and Dr. Christiana Ade for all their help and advice on the manuscript of chapter 2.

I'd like to acknowledge Dr. Haiganoush K. Preisler, Dr. Jonathan A Sam, Dr. Colleen E. Reid, Dr. Fan Jia, Dr. Yueqi Yan, Dr. Yang Ju, and Shuangshuang Xu for their suggestions on the statistics in Chapter 4.

My deepest gratitude goes to my dissertation committee members, Dr. Roger Bales, Dr. Erin Hestir, Dr. Christine Wiedinmyer, for all their advice, suggestions, and support throughout the dissertation process.

I am grateful to my families and friends for their support and encouragement. Life won't be joyful and wonderful without you all.

This work was funded by the California Energy Commission EPC-18-026, National Oceanic and Atmospheric Administration Climate Program – NOAA NA170AR4310284, California Department of Insurance 18028CA-AM 1, Strategic Growth Council of California CCR20021, and UC Lab Fees LFR-20-651032. MH acknowledges support from the California Department of Forestry and Fire Protection as part of the California Climate Investments Program, grant #8GG14803. This research was supported in part by the USDA Forest Service, Rocky Mountain Research Station, Aldo Leopold Wilderness Research Institute. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

Curriculum Vita

Qingqing Xu

School of Engineering

Environmental Systems Graduate Group

University of California Merced, CA 95343

Email: qxu6@ucmerced.edu, qingqingxu619@gmail.com

Education

Ph.D. in Environmental Systems	2016 - 2022
University of California, Merced, USA	
Dissertation title: Spatiotemporal burn severity, emissions, and health impact analyses of California wildfires.	
Advisor: Dr. Anthony LeRoy Westerling	
M.B.A. in International Business in Developing and Emerging Markets	2014 - 2016
Technische Universität Bergakademie Freiberg, Saxony, Germany	
M.S. in Management Science and Engineering	2013 - 2016
China University of Geosciences Beijing, Beijing, China	
B.S. in Business Administration	2009 - 2013
China University of Geosciences Beijing, Beijing, China	

Publications

Xu, Q., Westerling, A. L., Notohamiprodjo, A., Wiedinmyer, C., Picotte, J. J., Parks, S. A., ... & Ade, C. (2022). Wildfire burn severity and emissions inventory: an example implementation over California. *Environmental Research Letters*, 17(8), 085008.

Xu, Q., Westerling, A. L., & Baldwin, W. J. (2022). Spatial and temporal patterns of wildfire burn severity and biomass burning-induced emissions in California. *Environmental Research Letters*, 17(11), 115001.

Sam, J. A., Baldwin, W. J., Westerling, A. L., Preisler, H. K., **Xu, Q.**, Hurteau, M. D., ... & Thapa, S. B. (2022). Simulating burn severity maps at 30 meters in two forested regions in California. *Environmental Research Letters*, 17(10), 105004.

Westerling, L., Medellín-Azuara, J., & Viers, J. (2018) California's fourth climate change assessment San Joaquin Valley region report. (**Contributing author**)

Invited Talks and Posters

Xu, Q., Westerling, L., Notohamiprodjo, A., Picotte, J., Parks, S., & Wiedinmyer, C. (2021). An Integrated System for Estimating Burn Severity, Day-of-Burning, and Emissions with Google Earth Engine. Poster presentation at AGU Fall Meeting.

Xu, Q., Westerling, A.L., Schnier K., Wiedinmyer, C., & Hurteau, M.D. (2019). Wildfire Emissions and Public Health in California. Poster presentation at 2019 AGU Annual Meeting.

Xu, Q., Westerling, A.L., Baldwin, W.J., Sam, J., & Preisler, K. (2019). Controls on Severity of Large Wildfires in California. Oral presentation at the 8th International Fire Ecology Management Congress.

Xu, Q. Wildfires, smoke and public health in California. (2019). Invited talk at UC Merced Enviro-lunch.

Xu, Q., Westerling, A. L., Wiedinmyer, C., Hurteau, M. D., & Baldwin, W.J. (2019). Emissions from Sierra Nevada Wildfires. Poster presentation at the symposium on Climate Change and the Ecology of Sierra Nevada Forests.

Honors and Awards

Travel, Research, and Educational Experience (TREE) Grants from The Association for Fire Ecology (AFE) 2019

University of California, Merced travelling fellowship ESA meeting 2019

University of California, Merced summer fellowship 2018, 2019

Excellent Graduate Thesis, 2016

Outstanding Graduates of Beijing, 2013

Outstanding Student Leaders in College, 2013

Working experience

Air pollution specialist, California Air Resources Board, 2021 – current

Guest lecturer, University of California, Merced, 2022 Fall

Course: Remote Sensing of Environment

Graduate research assistant, University of California, Merced, 2017 – 2021

Teaching assistant, University of California, Merced, 2016 Fall

Course: Engineering Economic Analysis

Abstract of the Dissertation

Spatiotemporal burn severity, emissions, and health impact analyses of California wildfires

by

Qingqing Xu

Doctor of Philosophy, Environmental Systems
University of California, Merced, 2022
Professor Anthony LeRoy Westerling, Chair

Concerns over the potential impacts of increasing wildfires with continued climate change are growing. Comprehensive inventories of severity and emissions are essential for assessing these impacts and setting appropriate management strategies. To improve the availability of accurate fire severity and emissions estimates, we developed the Wildfire Burn Severity and Emissions Inventory (WBSE). WBSE is a retrospective spatial burn severity and emissions inventory at 30-m resolution for event-based assessment and 500-m resolution for daily emissions calculation. We applied the WBSE framework to calculate burn severity and emissions for large wildfires that burned during 1984-2020 in the state of California, U.S., a substantially more extended period than existing inventories. The framework described here can be applied to estimate severity for smaller wildfires and can also be used to estimate emissions for fires simulated in California for future climate and land-use scenarios. Based on the data generated with WBSE, we analyzed spatial and temporal patterns of burn severity and emissions from large wildfires burning in California during 1984-2020. Results show vegetation and severity play critical roles in controlling the spatial and seasonal distribution of emissions. California's annual area burned, and emissions are increasing, notably in early and late parts of what once was the typical fire season. Emissions and areas burned in moderate to high severity were particularly high and increasing in North Coast and Sierra Nevada forests. The year of 2020—with the most megafires in history—had fifteen times the annual average emissions of 1984-2015. Another key aim of this dissertation is to explore whether the existing health and wildfire emissions inventory data are sufficient to develop computationally efficient statistical methods that link wildfire pollution emissions to health impacts, to support the simulation of wildfire-related health impacts for very large numbers of future climate and land use scenarios. We presented our effort of developing a logistic Generalized Additive Model (GAM) and a negative binomial GAM model based on 36-years data in California to predict the probability of presence and numbers of respiratory hospitalizations using wildfire $PM_{2.5}$ emissions inventory data at ZIP Code Tabulation Areas (ZCTAs) spatial scale and semimonthly temporal scale.

1 Introduction

A wildfire is any non-structure fire—excluding prescribed fire—that occurs in wildlands. Though naturally ignited wildfires are important to the earth system and have some beneficial effects on ecosystems that have evolved with fire, human influences have increased wildfire risks through human-caused ignitions, human triggered climate change, and cumulative effects of land management and use (Bowman *et al* 2009, Stephens *et al* 2013, Balch *et al* 2017, Abatzoglou *et al* 2018, McLauchlan *et al* 2020). Comprehensive, accurate inventories of severity and emissions are essential for assessing these impacts and setting appropriate fire management and health care preparedness strategies, as is the ability to project emissions for future wildfires.

Burn severity, defined as the degree to which a fire-disrupted ecosystem has changed (Eidenshink *et al* 2007, Keeley 2009), is a key criteria to identify the immediate and long-term influence on fire regimes, air quality, emissions, and other impacts. Ground observations of burn severity have been conducted on hundreds of wildfires and prescribed fires in recent decades via the Composite Burn Index (CBI) protocol in the United States (Picotte *et al* 2019). However, on-the-ground burn severity data is still considered limited given that there are more than 50,000 wildfires on average annually according to National Interagency Fire Center (NIFC 2022). It is not possible to have a complete record via ground observation due to budget constraints and difficulties to reach some burn sites when the fires are on steep inclines in mountainous regions. Normalized spectral indices has been widely used to assess fire impact owing to the development of remotes sensing technologies. The Landsat derived differenced Normalized Burn Ratio (dNBR) is widely used for its advantages including temporal continuity, coverage and cost. Remote sensing-based assessment has limitations (e.g., cannot observe the surface when cloud cover is present, subjective and not field validated severity classification thresholds), that make them less than ideal for research (Kolden *et al* 2015).

Emission inventories are a critical input for a variety of modeling applications that seek to understand how emissions affect air quality, climate, and human health (Larkin *et al* 2014, Jaffe *et al* 2020, Pouliot *et al* 2020). Several global biomass burning emissions inventories have been produced and updated in recent decades. "Bottom-up" inventories such as the Global Fire Emissions Database (GFED; van der Werf *et al* 2004, 2006, 2010, 2017) and the Fire INventory from NCAR (FINN; Wiedinmyer *et al* 2011). "Top-down" inventories such as the Global Fire Assimilation System (GFAS; Kaiser *et al* 2011), Quick Fire Emissions Database (QFED; Koster *et al* 2015), and Fire Energetics and Emissions Research (FEER; Ichoku and Ellison 2014). Several emission inventories for the CONUS are currently available. The United States Environmental Protection Agency National Emissions Inventory (US EPA NEI; Raffuse *et al* 2012), the Wildland Fire Emissions Information System (WFEIS) (French *et al* 2014). Comparisons of fire emissions inventories have revealed variability in inventory results, with each having advantages and disadvantages (Larkin *et al* 2014, Faulstich *et al* 2022).

Fire has been an important ecological factor in California for millennia. Its Mediterranean climate is characterized by hot and dry summers that dry fuels produced during the winter wet season. Expanding residence in wildland-urban interfaces, fuel accumulations from fire suppression legacies over last century, and increasing temperatures and more variable precipitation, are making wildfires catastrophic in California, increasing in size and severity in recent decades (Keeley and Syphard 2019, Westerling *et al* 2006, Westerling 2016). Studies have analyzed the spatial and temporal distributions of wildfire occurrence (Li and Banerjee 2021), ignitions and controlling

factors (Chen and Jin 2022), prehistoric fire emissions (Stephens *et al* 2007), wildfire emissions calculation and contributors (Jaffe *et al* 2008, Hurteau *et al* 2014, CARB 2021), daily fire emissions and seasonal variability (Wiedinmyer and Hurteau 2010), and the projected effects of climate change and development on California wildfire emissions through 2100 (Hurteau *et al* 2014). Despite all the efforts, the spatiotemporal patterns of large wildfires' burn severity and emission regimes throughout California over the past four decades have not been investigated with high spatial and temporal resolution data.

Wildfires do not just affect the individuals who live in the fire zone. A much larger population is at risk of health impacts from exposure to wildfire emissions that can transport pollutants thousands of miles away (Reid *et al* 2016, Cascio 2018, Reid and Maestas 2019). The pollutant of most concern from wildfires is PM_{2.5} (atmospheric particulate matter that has a diameter of smaller than 2.5 micrometers). Many epidemiological studies have measured exposures using pollutant concentration data from ground-based air quality monitoring stations (Chen *et al* 2006, Morgan *et al* 2010, Aguilera *et al* 2021), and combinations of transport model and satellite observation approaches (Reid *et al* 2019 a & b, Gan *et al* 2020). Different degrees of uncertainty are associated with these methods. There are limited air quality monitoring stations, especially in rural areas where wildfires often occur, and stations might be damaged by fires which can interfere with incomplete or lost data. It's also challenging to separate the pollutions from wildfires. Atmospheric transport modeling offers the advantage to isolate the emission sources but could be biased due to input emissions and meteorology data. The expense of running transport models also severely limited the number of future scenarios that could be analyzed, limiting the value of the analyses. Satellite imagery may have low quality or missed information when clouds are present. There are also a few other exposure measurements based on the number of fires or data from questionnaires (Caamano-Isorna *et al* 2011, Marshall *et al* 2007, Künzli *et al* 2006). The number of fires may not be representative of actual threats since a large fire may emit much more emissions than a couple of small fires. Exposure measurement from questionnaires may underestimate real exposure (Youssef *et al* 2014). The current approaches are not ideal for future impact assessment.

This dissertation took the advantage of the recently available regression models developed at national and regional scales (Picotte *et al* 2021) to transform the remote sensing index the differenced Normalized Burn Ratio (dNBR)/NBR to the Composite Burn Index (CBI) to consistently map burn severity and calculating emissions in a transparent manner. We aimed to offer a valuable, open-source resource for the fire research community with a framework that allows for iterative improvements. We applied this framework and utilized this publicly available dataset to study the historical patterns of burn severity and emissions in California. We also explore the statistical approach to link this wildfire emissions inventory data and the existing health to develop computationally efficient statistical methods to support the simulation of wildfire-related health impacts for very large numbers of future climate and land use scenarios.

The objectives of this dissertation are presented in three self-contained chapters (2-4). In Chapter 2, "Wildfire Burn Severity and Emissions Inventory: An example implementation over California", my objective was to present the methods we developed to assess wildfire burn severity and emissions. Chapter 3, "Spatial and temporal patterns of wildfire burn severity and biomass burning-induced emissions in California" seeks to answer questions of (1) How have annual and seasonal patterns of burn severity and emissions changed over the period of 1984-2020? (2) What are the spatial distribution characteristics of burn severity and emissions? (3) What are the contributions of the most extreme fire events to emissions? In Chapter 4, "Development of Methods to Predict Respiratory Hospitalization with Wildfire Emissions Inventory Data", we presented the

development of methods to connect wildfire emissions to respiratory health impact without going through the computational expensive atmospheric transport modeling. All chapters were written using the pronoun ‘we’ to refer to myself and co-authors of each manuscript.

References

- Abatzoglou J T, Williams A P and Barbero R 2018 Global emergence of anthropogenic climate change in fire weather indices *Geophys. Res. Lett.* **46** 326–36
- Aguilera R, Corringham T, Gershunov A and Benmarhnia T 2021 Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California *Nat Commun* **12** 1493
- Balch J K, Bradley B A, Abatzoglou J T, Nagy R C, Fusco E J and Mahood A L 2017 Human-started wildfires expand the fire niche across the United States *Proc. Natl Acad. Sci.* **114** 2946–51
- Bowman D M J S *et al* 2009 Fire in the Earth system *Science* **324** 481–4
- California Air Resource Board 2021 Wildfire emissions & burned area estimates 2000–2020 (available at: https://ww2.arb.ca.gov/sites/default/files/2021-07/Wildfire%20Emission%20Estimates%20for%202020%20_Final.pdf) (Accessed 12 December 2021)
- Caamano-Isorna F, Figueiras A, Sastre I, Montes-Martínez A, Taracido, M and Piñeiro-Lamas M 2011 Respiratory and mental health effects of wildfires: an ecological study in Galician municipalities (north-west Spain) *Environ Health* **10** 48
- Cascio W E 2018 Wildland fire smoke and human health *Science of The Total Environment* **624** 586–95
- Chen B and Jin Y 2022 Spatial patterns and drivers for wildfire ignitions in California *Environ. Res. Lett.* **17** 055004
- Chen L, Verrall K and Tong S 2006 Air particulate pollution due to bushfires and respiratory hospital admissions in Brisbane, Australia *International Journal of Environmental Health Research* **16** 181–91
- Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B and Howard S 2007 A project for monitoring trends in burn severity *Fire Ecol.* **3** 3–21
- Faulstich S D, Schissler A G, Strickland M J and Holmes H A 2022 Statistical comparison and assessment of four fire emissions inventories for 2013 and a large wildfire in the Western United States *Fire* **5** 27
- French N H F *et al* 2014 Modeling regional-scale wildland fire emissions with the wildland fire emissions information system *Earth Interact.* **18** 1–26
- Gan R W, Liu J, Ford B, O’Dell K, Vaidyanathan A, Wilson A, Volckens J, Pfister G, Fischer E V, Pierce J R and Magzamen S 2020 The association between wildfire smoke exposure and asthma-specific medical care utilization in Oregon during the 2013 wildfire season *J Expo Sci Environ Epidemiol* **30** 618–28
- Hurteau M D, Westerling A L, Wiedinmyer C and Bryant B P 2014 Projected effects of climate and development on California wildfire emissions through 2100 *Environ. Sci. Technol.* **48** 2298–304
- Ichoku C and Ellison L 2014 Global top-down smoke-aerosol emissions estimation using satellite fire radiative power measurements *Atmos. Chem. Phys.* **14** 6643–67
- Jaffe D, Chand D, Hafner W, Westerling A and Spracklen D 2008 Influence of fires on O₃ concentrations in the Western U.S. *Environ. Sci. Technol.* **42** 5885–91

- Jaffe D A, O'Neill S M, Larkin N K, Holder A L, Peterson D L, Halofsky J E and Rappold A G 2020 Wildfire and prescribed burning impacts on air quality in the United States *J. Air Waste Manage. Assoc.* **70** 583–615
- Kaiser J W *et al* 2011 Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power *Biogeoscience* **9** 527–54
- Keeley J E and Syphard A D 2019 Twenty-first century California, USA, wildfires: fuel-dominated vs. wind-dominated fires *Fire Ecol.* **15** 1–5
- Kolden C A, Smith A M S and Abatzoglou J T 2015 Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA *Int. J. Wildland Fire* **24** 1023–8
- Koster R D, Darmenov A S and da Silva A M 2015 *The Quick Fire Emissions Dataset (QFED): Documentation of Versions 2.1, 2.2 and 2.4 (Technical Report Series on Global Modeling and Data Assimilation vol 38)* ed R D Koster (Greenbelt, MD: National Aeronautics and Space Administration) p 201
- Künzli N, Avol E, Wu J, Gauderman W J, Rappaport E, Millstein J, Bennion J, McConnell R, Gilliland F D, Berhane K, Lurmann F, Winer A and Peters J M 2006 Health Effects of the 2003 Southern California Wildfires on Children *Am J Respir Crit Care Med* **174** 1221–8
- Larkin N K, Raffuse S M and Strand T M 2014 Wildland fire emissions, carbon, and climate: U.S. emissions inventories *For. Ecol. Manage.* **317** 61–69
- Li S and Banerjee T 2021 Spatial and temporal pattern of wildfires in California from 2000 to 2019 *Sci. Rep.* **11** 8779
- Marshall G N, Schell T L, Elliott M N, Rayburn N R and Jaycox L H 2007 Psychiatric Disorders Among Adults Seeking Emergency Disaster Assistance After a Wildland-Urban Interface Fire *PSYCHIATRIC SERVICES* **58** 6
- McLauchlan K K *et al* 2020 Fire as a fundamental ecological process: research advances and frontiers ed G Durigan *J. Ecol.* **108** 2047–69
- Morgan G, Sheppard V, Khalaj B, Ayyar A, Lincoln D, Jalaludin B, Beard J, Corbett S and Lumley T 2010 Effects of Bushfire Smoke on Daily Mortality and Hospital Admissions in Sydney, Australia *Epidemiology* **21** 47–55
- National Interagency Coordination Center 2022 Wildfires and Acres (available at: <https://www.nifc.gov/fire-information/statistics/wildfires>)
- Picotte J *et al* 2019 Composite Burn Index (CBI) Data for the Conterminous US, Collected Between 1996 and 2018: U.S. Geological Survey data release, <https://doi.org/10.5066/P91BH1BZ>.
- Picotte J J, Cansler C A, Kolden C A, Lutz J A, Key C, Benson N C and Robertson K M 2021 Determination of burn severity models ranging from regional to national scales for the conterminous United States *Remote Sens. Environ.* **263** 112569
- Pouliot G, Pierce T, Benjey W, O'Neill S M and Ferguson S A 2020 Wildfire emission modeling: integrating bluesky and SMOKE *14th Annual Int. Emission Inventory Conf.* 11–14
- Raffuse S M, Larkin N K and Lahm P W 2012 Development of version 2 of the wildland fire portion of the National Emissions Inventory *20th Int. Emission Inventory Conf.*
- Reid C E, Brauer M, Johnston F H, Jerrett M, Balmes J R and Elliott C T 2016 Critical review of health impacts of wildfire smoke exposure *Environ. Health Perspect.* **124** 1334–43
- Reid C E, Considine E M, Watson G L, Telesca D, Pfister G G and Jerrett M 2019 Associations between respiratory health and ozone and fine particulate matter during a wildfire event *Environment International* **129** 291–8
- Reid C E and Maestas M M 2019 Wildfire smoke exposure under climate change: impact on respiratory health of affected communities *Current Opinion in Pulmonary Medicine* **25**

- 179–87 Stephens S L *et al* 2013 Temperate and boreal forest mega-fires: characteristics and challenges *Front. Ecol. Environ.* **12** 115–22
- Stephens S L, Martin R E and Clinton N E 2007 Prehistoric fire area and emissions from California's forests, woodlands, shrublands, and grasslands *For Ecol. Manage.* **251** 205–16
- Stephens S L *et al* 2013 Temperate and boreal forest mega-fires: characteristics and challenges *Front. Ecol. Environ.* **12** 115–22
- Stevens J T, Collins B M, Miller J D, North M P and Stephens S L 2017 Changing spatial patterns of stand-replacing fire in California conifer forests *For Ecol. Manage.* **406** 28–36
- van der Werf G R *et al* 2017 Global fire emissions estimates during 1997–2016 *Earth Syst. Sci. Data* **9** 697–720
- van der Werf G R, Randerson J T, Collatz G J, Giglio L, Kasibhatla P S, Arellano A F, Olsen S C and Kasischke E S 2004 Continental-Scale Partitioning of Fire Emissions During the 1997–2001 El Niño/La Niña Period *Science* **303** 73–76
- van der Werf G R, Randerson J T, Giglio L, Collatz G J, Kasibhatla P S and Arellano A F 2006 Interannual variability in global biomass burning emissions from 1997 to 2004 *Atmos. Chem. Phys.* **6** 3423–41
- van der Werf G R, Randerson J T, Giglio L, Collatz G J, Mu M, Kasibhatla P S, Morton D C, DeFries R S, Jin Y and van Leeuwen T T 2010 Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009) *Atmos. Chem. Phys.* **10** 11707–35
- Westerling A L, Hidalgo H G, Cayan D R and Swetnam T W 2006 Warming and earlier spring increase Western U.S. forest wildfire activity *Science* **313** 940–3
- Westerling A L 2016 Increasing Western US forest wildfire activity: sensitivity to changes in the timing of spring *Phil. Trans. R. Soc. B* **371** 20150178
- Wiedinmyer C and Hurteau M D 2010 Prescribed fire as a means of reducing forest carbon emissions in the Western United States *Environ. Sci. Technol.* **44** 1926–32
- Wiedinmyer C, Akagi S K, Emmons L K, Al-Saadi J A, Orlando J J, Soja A J and Soja A J 2011 The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning *Geosci. Model Dev.* **4** 625–41
- Youssef H, Lioussé C, Roblou L, Assamoi E M, Salonen R O, Maesano C, Banerjee S and Annesi-Maesano I 2014 Quantifying wildfires exposure for investigating health-related effects *Atmospheric Environment* **97** 239–51

2 Wildfire Burn Severity and Emissions Inventory: An example implementation over California

Abstract

Wildfire severity is a key indicator of both direct ecosystem impacts and indirect emissions impacts that affect air quality, climate, and public health far beyond the spatial footprint of the flames. Comprehensive, accurate inventories of severity and emissions are essential for assessing these impacts and setting appropriate fire management and health care preparedness strategies, as is the ability to project emissions for future wildfires. The frequency of large wildfires and the magnitude of their impacts have increased in recent decades, fueling concerns about decreased air quality. To improve the availability of accurate fire severity and emissions estimates, we developed the wildfire burn severity and emissions inventory (WBSE). WBSE is a retrospective spatial burn severity and emissions inventory at 30 m resolution for event-based assessment and 500 m resolution for daily emissions calculation. We applied the WBSE framework to calculate burn severity and emissions for historically observed large wildfires (>404 hectares (ha)) that burned during 1984–2020 in the state of California, U.S., a substantially more extended period than existing inventories. We assigned the day of burning and daily emissions for each fire during 2002–2020. The framework described here can also be applied to estimate severity for smaller wildfires and can also be used to estimate emissions for fires simulated in California for future climate and land-use scenarios. The WBSE framework implemented in R and Google Earth Engine can provide quick estimates once a desired fire perimeter is available. The framework developed here could also easily be applied to other regions with user-modified vegetation, fuel data, and emission factors.

Keywords: burn severity, emission inventory, daily progression estimates, Google Earth Engine

2.1 Introduction

Though some wildfires are essential to the Earth system and have beneficial effects on ecosystems that have evolved with fire, human influences have altered wildfire risk via increasing ignitions, climate change, and cumulative impacts of land management and use (Bowman *et al* 2009, Stephens *et al* 2013, Balch *et al* 2017, Abatzoglou *et al* 2018, McLauchlan *et al* 2020). In addition to ecosystem effects, wildfires directly impact people through property losses, injuries, mortality, and evacuations, while smoke and its health impacts are how the largest numbers of people are affected by wildfires through the regional transport of smoke in the atmosphere (Finlay *et al* 2012, Reid *et al* 2016, Bowman *et al* 2017, Thomas *et al* 2017, Burke *et al* 2021).

The two most commonly used criteria for analyzing fire effects are burned area and burn severity (Key and Benson 2006, Kolden *et al* 2012, Meng and Zhao 2017). The degree to which a fire-disrupted ecosystem has changed is referred to as burn severity (Eidenshink *et al* 2007, Keeley 2009). Because of differences in wind, topography, fuel conditions, and other factors, burned areas typically consist of complex landscape mosaics of low, moderate, and high burn severity (Turner *et al* 1994, Perry *et al* 2011, Birch *et al* 2015, Prichard *et al* 2020). Thus, landscape-scale quantification of burn severity is vital to identify the immediate and long-term influence on fire regimes, air quality, emissions, and other impacts. Passive spectral reflectance satellites and field observation have been used to determine burn severity. The change in Normalized Burn Ratio (NBR) between pre-and post-fire satellite images (dNBR) has become the primary index for remote

sensing of severity. In the United States, the Monitoring Trends in Burn Severity Project (MTBS) (Eidenshink *et al* 2007), which is derived from long-running Landsat earth observing sensors, mapped burn severity for large fires dating back to 1984. However, the MTBS products have limitations, with sources of potential errors (e.g., subjective and not field validated severity classification thresholds) and a 2-year delay in including new fires in the database, that make them less than ideal for research (Kolden *et al* 2015). The Composite Burn Index (CBI) field method tabulates severity within 30-m plots by averaging measurements of several rating factors on a continuous scale ranging from 0.0 (unburned) to 3.0 (high severity) (Key and Benson 2006). On-the-ground burn severity was assessed following the CBI protocol only for 234 fires that burned in the contiguous United States (CONUS) during 1996-2018 (Picotte *et al* 2019).

Emission inventories are a critical input for a variety of modeling applications that seek to understand how emissions affect air quality, climate, and human health (Larkin *et al* 2014, Jaffe *et al* 2020, Pouliot *et al* 2020). Several global biomass burning emissions inventories have been produced and updated in recent decades. "Bottom-up" inventories such as the Global Fire Emissions Database (GFED; van der Werf *et al* 2004, 2006, 2010, 2017) and the Fire INventory from NCAR (FINN; Wiedinmyer *et al* 2011), rely on satellite burned area and/or active fire detections, coupled with fuel loads, combustion completeness, and emissions factors to estimate fire emissions. GFED has provided monthly, daily, and 3-hourly (since 2003) emissions data at 0.25 degrees since 1997. FINN has provided daily estimates at 1 km resolution since 2002. "Top-down" inventories such as the Global Fire Assimilation System (GFAS; Kaiser *et al* 2011), Quick Fire Emissions Database (QFED; Koster and Darmenov 2015), and Fire Energetics and Emissions Research (FEER; Ichoku and Ellison 2014) rely on fire radiative power observations to compute emissions, along with different approaches to address cloud cover and to compute scaling factors for fire emissions. GFAS, QFED and FEER produce daily estimates at 0.1 degrees since 2003, 2000, and 2003, respectively. Several emission inventories for the CONUS are currently available. The United States Environmental Protection Agency National Emissions Inventory (US EPA NEI, Raffuse *et al* 2012) estimates emissions from all sources including wildfire and prescribed fire emissions every 3 years. The Wildland Fire Emissions Information System (WFEIS) (French *et al* 2014) provides tools for assessing wildland fire emissions in area-of-interest or for an identified fire event. It also provides pre-calculated emissions data on state or country level on monthly/yearly scale with the earliest records back to 1984. Comparisons of fire emissions inventories have revealed variability in inventory results, with each having advantages and disadvantages (Larkin *et al* 2014, Faulstich *et al* 2022). Despite all efforts, there is an urgent need for burn severity and emissions inventories at the scale of actionable management (e.g., fuel reduction treatments) and over longer temporal periods.

Here, we introduce the Wildfire Burn Severity and Emissions Inventory (WBSE) and present a detailed description of the inventory model, initial results using California historical large wildfires as an example, comparison with other inventory estimates, and a discussion of uncertainties. The primary objective of this manuscript is to describe the model and our comparison with other emissions inventories is limited in scope. Similarly, the historical reconstructions of wildfire severity maps and emissions are analyzed in more detail in a companion paper which is currently in review for this issue. This study benefits from the recently available regression models developed at national and regional scales transforming the dNBR/NBR to CBI to consistently map burn severity in a transparent manner (Picotte *et al* 2021). Following Wiedinmyer *et al* (2011), we use modified algorithms and updated fuel emissions and consumption factors drawing on recently published field and remote sensing data to better capture emissions from more extreme fires. Larger, severe wildfires in California negatively impact ecosystems, human health and life, communities,

and infrastructure due to expanding wildland-urban interfaces, fuel accumulation from a century of fire exclusion policies, and increasing aridity, which increases the amount of biomass available to burn. (Abatzoglou and Williams 2016, Westerling 2016, Williams *et al* 2019, Goss *et al* 2020, Goodwin *et al* 2021). The comprehensive, long-term event and daily emissions records described here could be used to study health effects of wildfire smoke, either by combining them with transport modeling to model air quality and estimate exposures, or by incorporating them into statistical models predicting health impacts as a direct function of estimated emissions. These data will also facilitate analyses of changing emissions impacts on the carbon cycle over the last three decades. High resolution severity and emissions raster maps are generated for each fire event to support further spatial analysis. While the emissions calculated for California with WBSE is not a substitute for real-time daily emissions estimates, it is designed to extend the estimated emissions record back to 1984 with a finer spatial resolution and provide more up-to-date estimates on emissions factors reflecting information from California's recent extreme fires.

2.2 Methods

The Wildfire Burn Severity and Emissions Inventory (WBSE) provides estimates of 30-m resolution burn severity, and emissions of CO₂, CO, CH₄, non-methane organic compounds (NMOC), SO₂, NH₃, NO, NO₂, nitrogen oxides (NO_x = NO + NO₂), PM_{2.5}, OC, and BC. We implemented WBSE for California large wildfires on a per-fire event scale since 1984 and also a daily scale since 2002. The inventory implementation steps, input datasets, and output data are summarized in Figure 2.1. Emissions of all species are calculated as a function of area burned, fuel loading, the fraction of vegetation burned based on burn severity, and an emissions factor specific to each vegetation type using the following equation modified from the FINN model (Wiedinmyer *et al* 2011):

$$E_i = A_{v,s} \times F_v \times cr_{v,s} \times ef_{v,i} \quad (1)$$

Where E_i is the mass of emission species i emitted, $A_{v,s}$ and $cr_{v,s}$ are the area burned and fuel consumption rate of general vegetation class v (grass, shrub, forest < 5500 ft, forest 5500-7500 ft, forest > 7500 ft) under severity class s (low, moderate, and high severity), F_v is the fuel loading of vegetation class v , and $ef_{v,i}$ is the emission factor of emission species i for vegetation v .

The FINN model was developed to estimate global emissions from open burning at 1-km resolution. It has been used in various modeling studies for modeling atmospheric chemistry and air quality (Jiang *et al* 2012, Val Martin *et al* 2013, Brey *et al* 2018). The primary differences of emission calculation methods between WBSE and FINN are that WBSE calculates emissions per fire event while FINN provides global emission estimates in real time. Other differences include (1) obtaining the area burned by fire severity class from a new CBI-based burned area product (based on fire perimeters published by MTBS and California Department of Forestry and Fire Protection (CAL FIRE)) while the MODIS Thermal Anomalies Product is used in FINN; (2) assigning vegetation type from the LANDFIRE existing vegetation type (EVT) product to general vegetation classes with an updated crosswalk from Hurteau *et al* (2014) while the MODIS land cover product is used in FINN; (3) calculating fuel loading and consumption based on burn severity for each vegetation class while fuel consumption is based on a function of tree cover and fuel loading for generic land cover classifications for the global regions in FINN; (4) Using region specific emissions factors for California and Western U.S. ecosystems to calculate emissions for the example implementation over California; and (5) assigning the day of burning based on MODIS and VIIRS data for each large wildfire burned since 2002. The inventory has a spatial resolution of 30-m for per-fire emission estimates and 500-m for daily emissions rather than the 1-km resolution in FINN. The area burned in low, moderate, and high burn severity is calculated from Landsat-derived indices

(NBR and dNBR) using a hierarchy of regression models developed by Picotte *et al* (2021). We recognize that other fire severity metrics have been developed (e.g. RdNBR and RBR; Miller and Thode 2007 & Parks *et al* 2014, respectively), but we opted to use NBR/dNBR because of the regionally specific models developed by Picotte *et al* (2021) that used NBR/dNBR. We don't have a quantitative assessment of the uncertainty; however, we assign a factor of 2 to the uncertainty due to all of the inputs to the model following Wiedinmyer *et al* (2011). A matrix table of 30-m pixel level PM_{2.5} emission uncertainty for 5 general vegetation classes due to burn severity classification error is given in Supplementary Table S2.3.

Wildfire Burn Severity and Emission Inventory (WBSE)

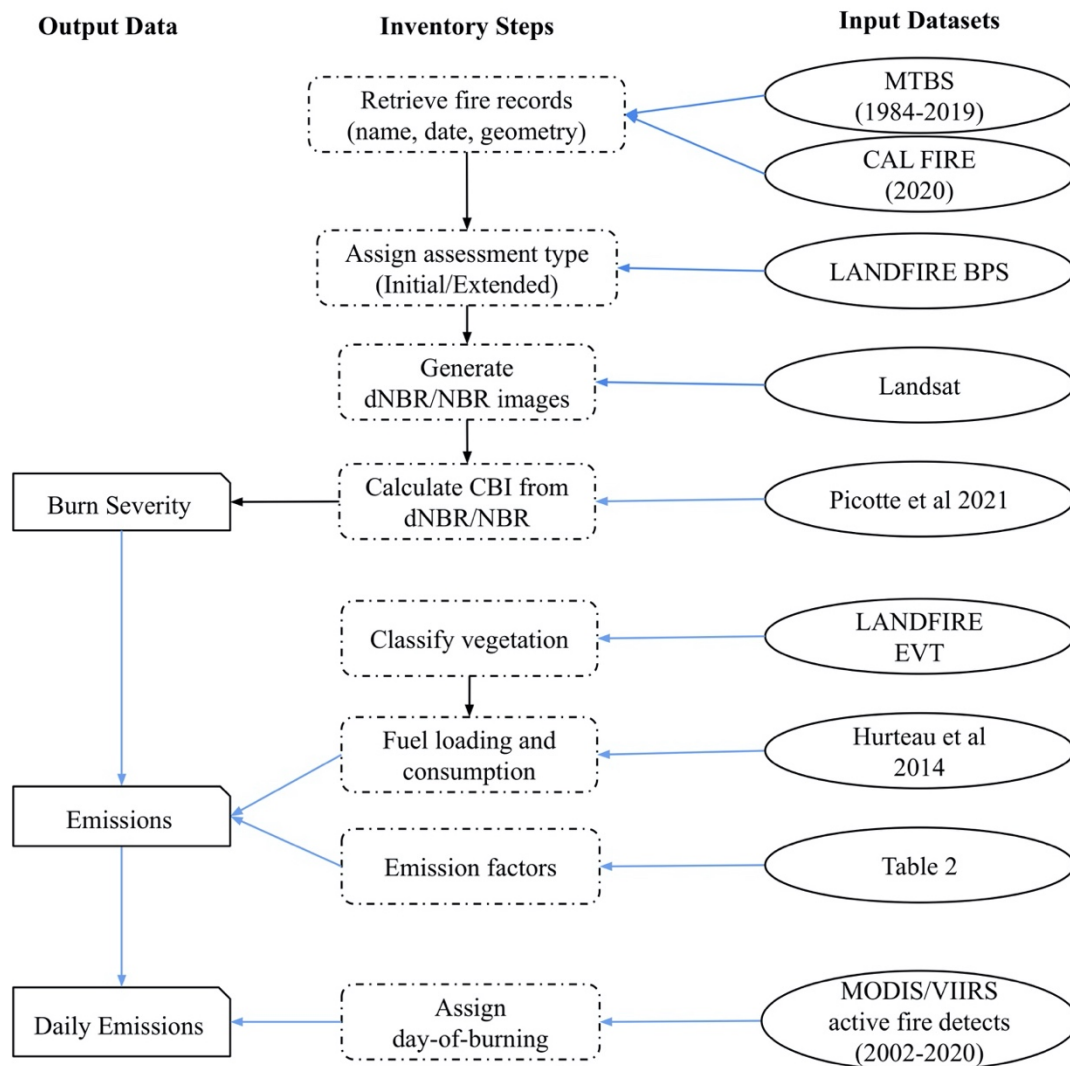


Figure 2. 1 Diagram of Wildfire Burn Severity and Emissions Inventory (WBSE) inventory steps and datasets.

Fire records for California from 1984 - 2019 were retrieved from MTBS (<https://mtbs.gov/viewer/index.html>) via interactive viewer on May 8th, 2021, resulting in a dataset with a total of 1623 wildfires. We also acquired fire perimeters for 74 large wildfires in 2020 from

CAL FIRE (<https://frap.fire.ca.gov/frap-projects/fire-perimeters/>) and calculated dNBR for each 2020 fire using the dNBR calculation tool with Google Earth Engine (Figure 2.2). This process first selects either initial assessment or extended assessment for each fire. The initial assessment utilizes Landsat images acquired immediately after a fire to capture first-order fire effects. The extended assessment uses images obtained during the growing season following the fire to identify delayed first-order effects and dominant second-order effects (Eidenshink *et al* 2007). We utilized LANDFIRE Biophysical Settings (BPS) to determine which assessment type to apply for each fire burned in 2020. Per Picotte *et al* (2021), we used extended assessment if the majority of general vegetation groups within the fire perimeter are forests, while initial assessment is used when the majority of general vegetation groups are grassland/shrubland. By contrast, MTBS uses extended assessment for forest and shrubland types. We did not delineate grasslands into burn severity categories. Instead, we classified them as burned ("grass burn") because of difficulties in assessing vegetation change. Post-fire images for extended assessment were selected during the next peak of the green season (June - September) using the mean compositing approach suggested by Parks *et al* (2018). Composite post-fire images acquired immediately within two months after the fire containment dates were used for the initial assessment. Composite pre-fire images for extended and initial assessments were acquired with the matching periods from the preceding year. The dNBR images were produced by quantifying the spectral difference between composite pre-fire and post-fire Landsat scenes.

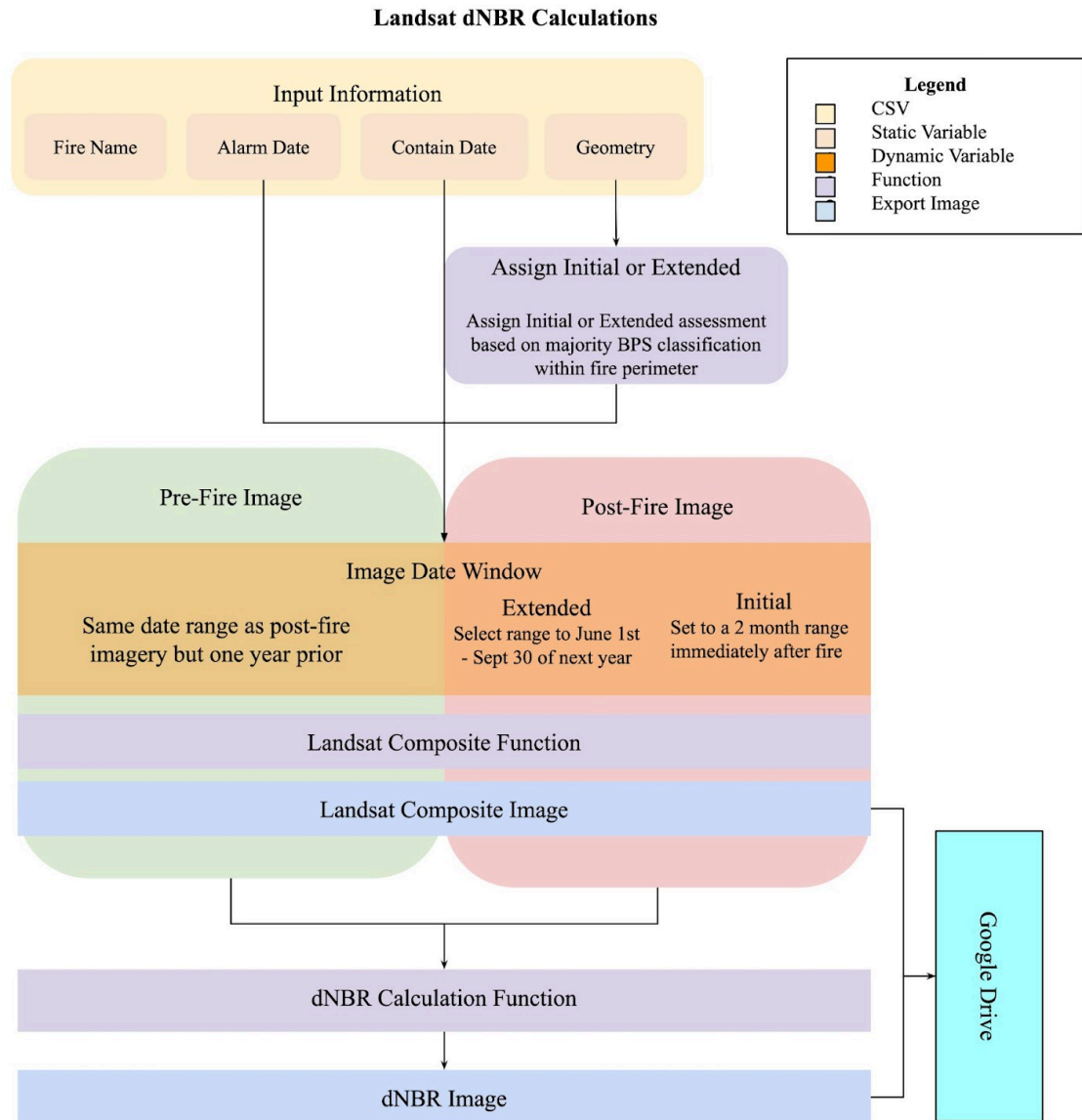


Figure 2. 2 Diagram of composite the differenced Normalized Burn Ratio (dNBR) calculation with Google Earth Engine.

Burn severity for a given fire can have either a linear or non-linear relationship to dNBR/NBR (Picotte *et al* 2021). So to make an equivalent burn severity product across all fires, we calculated the unitless, continuous CBI variable from dNBR/NBR values using the linear and Sigmoid B regression models developed for the CONUS by Picotte *et al* (2021). We first applied the MTBS or CAL FIRE perimeter shapefile for each wildfire event to crop the BPS raster file. The BPS data are categorized into one of twelve vegetation groups. For CBI calculations, we further grouped "Conifer", "Hardwood", "Hardwood-Conifer", "Shrubland", and "Riparian" into the "forest/shrub" classification; "Savanna", "Sparse", and "Grassland" into the "grass" classification; and the rest "Barren-Rock/Sand/Clay", "Open Water", and "Perennial Ice/Snow" into the "unchanged" group.

We applied the decision tree framework (Figure 2.3) for each pixel categorized as forest/shrub to convert the dNBR/NBR value into a CBI value. CBI values less than or equal to 0 were arbitrarily assigned a value of 0.001 to be classified into the "unburned" group. Modeled CBI values larger than 3 were assigned the value 3 to be classified into the "high severity" group. For the grass group, we set a CBI value of 3.5 to indicate that this pixel was a grass burn. CBI values were then classified following thresholds modified based on Crotteau *et al* (2014) into six severity classes: unburned, low severity, moderate severity, high severity, grass burn, and non-processing area.

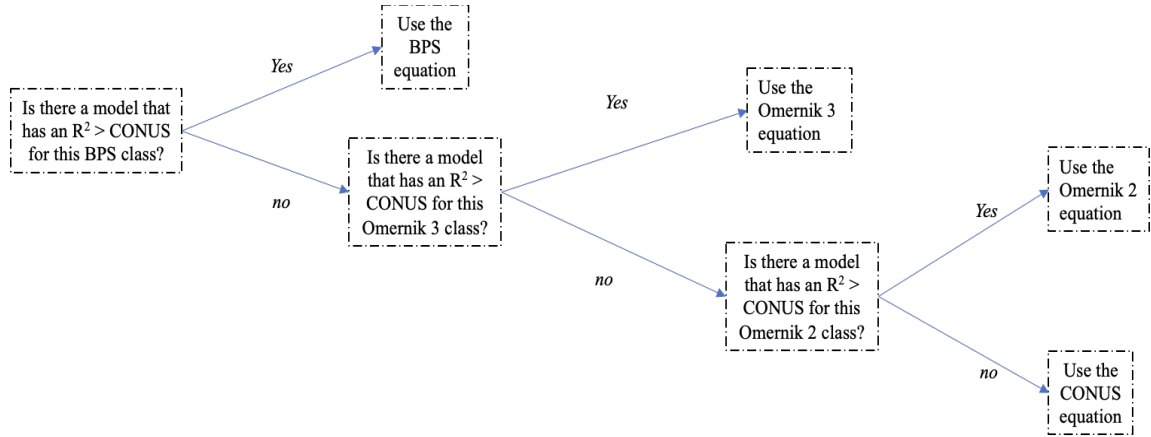


Figure 2. 3 A decision tree framework to apply models to convert dNBR/NBR values to CBI values (adapted from Picotte et al 2021)

Fuel categories were assigned from LANDFIRE EVT products. EVT represents the current distribution of the terrestrial ecological systems classification (Rollins 2009), and the vegetation layers were updated every couple of years to account for landscape disturbances. The most appropriate version was applied for each fire event depending on the burn year -- EVT data version LF 1.0.5, 1.2.0, 1.3.0, 1.4.0, 2.0.0 were applied to wildfires burned before the year of 2002, 2002-2010, 2011-2012, 2013-2014, and after 2014, respectively. For emissions calculations, EVT data were then categorized into five general vegetation categories: grass, shrub, forest under 5500 feet (1676 m), forest between 5500-7500 feet (1676-2286 m), and forest above 7500 feet (2286 m), updated for California ecosystems (Supplementary Table S2.1). Fuel consumption was determined following Hurteau *et al* (2014) assigning fuel loading and consumption values for each severity class for the five general vegetation categories based on the First Order Fire Effects Model v5 (Reinhardt *et al* 1997) (Supplementary Table S2.2).

The LF EVT classes were assigned to generic vegetation classes of FINN so that emission factors could be applied. Emission factors for greenhouse gases, particulate matter, and reactive trace gases were updated with recent data for each general vegetation class using results from recent field campaigns and studies specific for California ecosystems and Western U.S. ecosystems (see Table 2.1).

Table 2. 1 Emission factors (g species emitted per kg biomass burned) assigned to grasslands, shrublands, and forests

Emission Factors	Temperate Forest	Shrublands	Grasslands
CO ₂	1650 ^a	1588 ^b	1421 ^h
CO	80 ^a	97 ^b	56 ^h
CH ₄	3.7 ^a	2.2 ^b	2.9 ^h
NMOC	24 ^c	24 ^d	28 ^e
SO ₂	0.94 ^a	0.73 ^b	3.0 ^h
NH ₃	1.17 ^f	1.68 ^g	0.56 ^h
NO	0.90 ^f	1.3 ^g	2.9 ^h
NO ₂	0.99 ^f	0.94 ^g	3.1 ^h
NO _x (as NO)	1.2 ^f	2.1 ^g	3.9 ⁱ
PM _{2.5}	10.6 ^a	7.9 ^a	7.2 ⁱ
OC	11.6 ^j	3.7 ^k	2.6 ^k
BC	0.4 ^j	1.31 ^k	0.37 ^k

^a Average of Forest emission factors, Table 5, Prichard *et al* (2020)

^b Shrublands and Grasslands emissions factors, Table 5, Prichard *et al* (2020)

^c Average NMOC emissions for forest fires from Permar *et al* (2021)

^d Average NMOC emissions from sagebrush fires from Permar *et al* (2021)

^e Sum of NMOCs in Akagi *et al* (2011), updated Feb 2015 for Savanna

^f Average of Forest emission factors, Table 5, Prichard *et al* (2020) and forest fire E.F.s from Lindaas *et al* (2020)

^g Average of Shrubland emission factors, Table 5, Prichard *et al* (2020) and sagebrush fire E.F.s from Lindaas *et al* (2020)

^h Grassland emission factors, Table 5, Prichard *et al* (2020)

ⁱ Akagi *et al* (2011)- updated Feb. 2015 for Savanna

^j Supplemental Table 2, Permar *et al* (2021)

^k Wiedinmyer *et al* (2011)

To assign the day of burning for individual pixels, NASA FIRMS active fire products from MODIS (Collection 6) within 750 m of the fire perimeter shapefiles supplied by MTBS or CAL FIRE were selected for interpolation to account for detections that might be outside the boundary due to

detection radius. VIIRS 375 m data, when available since 2012, was added to complement MODIS data with improved performance to assign burn dates using the fire progression raster tool (Figure 2.4). We filtered the MODIS/VIIRS detection points to the date range of interest and created a 500m buffer around each point. Points were then converted to circle polygons to represent each point's detection extent properly. The average date was selected as the proper date in regions of overlapping buffers. We then calculated daily emissions and assigned them to the centroids of the aggregated daily progression polygons. Versions were also implemented using the earliest or latest date of detection points to assign the day of burning (results not reported here, but maps are available).

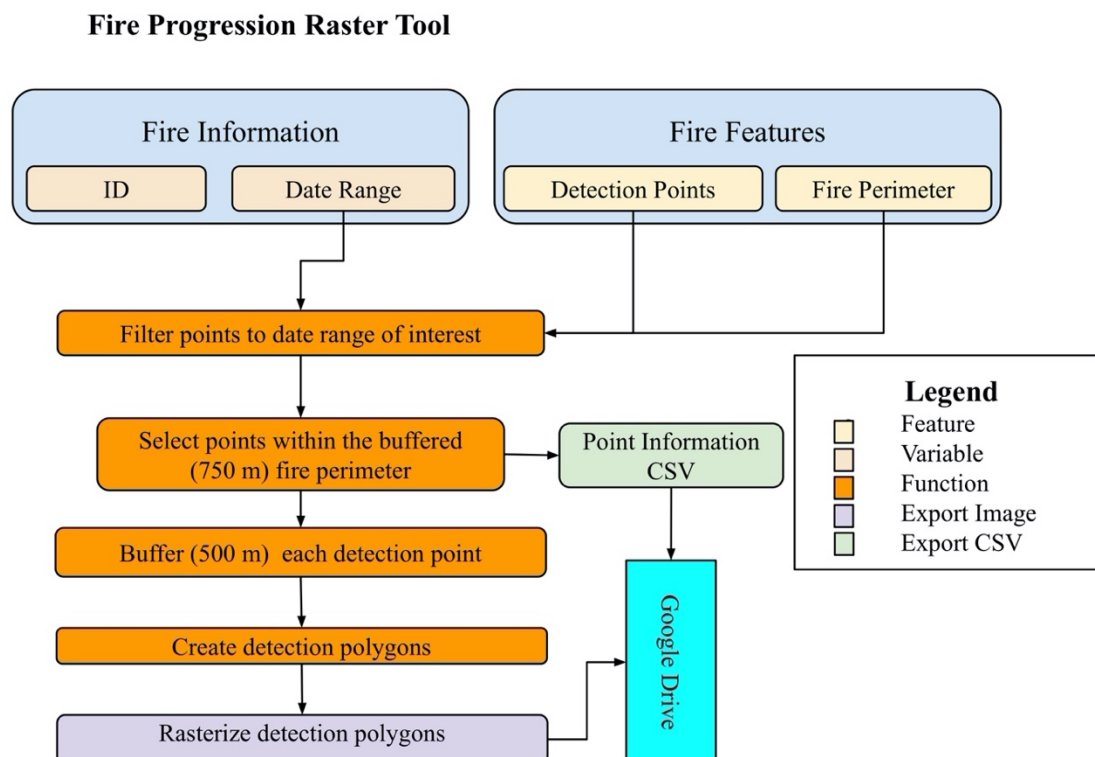


Figure 2. 4 A diagram of assigning day-of-burning with the fire progression raster tool

2.3 Results and Discussion

Burn severity and emissions were calculated for 1697 large wildfire events in California during 1984 - 2020 at a 30-m resolution scale using the WBSE framework described above. Day of burning and daily emissions were calculated for each fire during 2002-2020. Data processing and calculations were implemented using R (R Core Team 2021) and Google Earth Engine (Gorelick *et al* 2017). The datasets are available for download and use.

2.3.1 Burn Severities

Over the 1984-2020 period, the average annual area burned (excluding obscured area) was 0.23 million ha with substantial interannual variability (Figure 2.5). Obscured area includes data gaps associated with clouds, smoke, or the Landsat 7 ETM+ scan-line corrector failure (Key 2006), affecting 0.3% of the area mapped within historical burn perimeters. The maximum annual area

burned was 1.6 million ha in 2020, about 220 times the area burned in the minimum year of 1991. Of the approximately 8.5 million ha mapped by this project, about 63 percent burned at moderate or high severity (Figure 2.6).

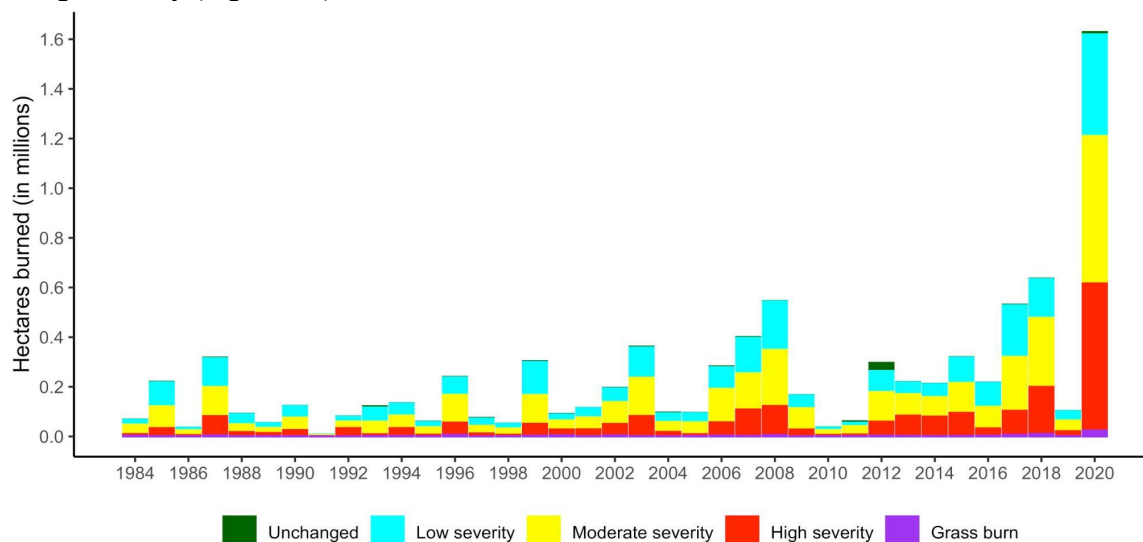


Figure 2. 5 The annual area burned in unchanged (fire refugia within perimeters), low, moderate, high severity, and grass burn categories of large wildfires(>404 ha) in the California region during 1984-2020

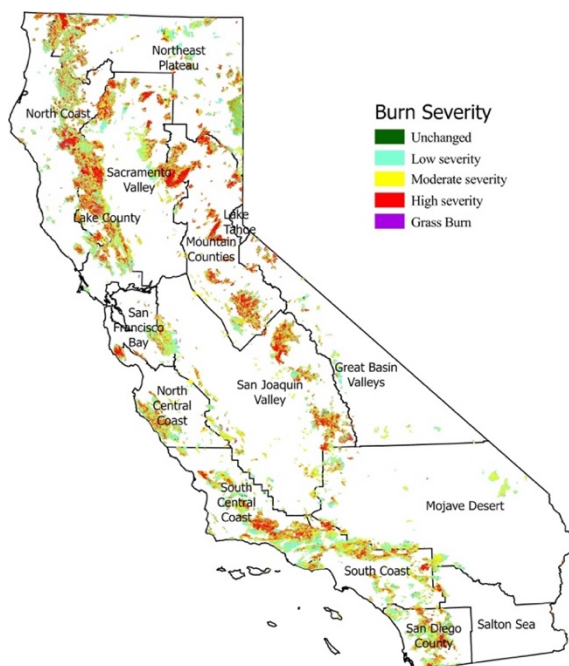


Figure 2. 6 Burn severity map of all large wildfires(>404 ha) burned in each California air basin during 1984-2020. Some burned areas overlap since the same place could be burned more than once between 1984 and 2020

As MTBS has been previously established as a rough estimate of burn severity rather than a rigorous calculation for a specified need (e.g., emissions calculations), a direct comparison of the

total area burned by severity category between MTBS and WBSE is helpful to contextualize the more rigorous approach applied here. MTBS analysts subjectively classify NBR and dNBR into six classes (unburned to low, low, moderate, high, increased greenness, and non-processing area mask) (Kolden *et al* 2015). Here, we empirically classified modeled CBI values into four classes (unchanged, low, moderate, and high) and two categories of "grass burn" and "non-processing area" using an objective, standardized procedure. A pixel-level classification comparison (excluding non-processing area) indicates that ~45% of the pixels in both datasets were classified as the same burn severity (cells highlighted in blue in Table 2.2); however, WBSE burn severity was generally higher than MTBS (cells highlighted in pink in Table 2.3). We found that 22% percent of pixels classified as low burn severity by MTBS were classified as moderate burn severity by WBSE. By contrast, less than 2% of pixels were classified as having higher burn severity by MTBS than WBSE (cells highlighted in green in Table 2.2). The remaining pixels were mainly classified as low burn severity by WBSE but unburned to low burn severity by MTBS.

Table 2. 2 Pixel level burn severity classification comparison of MTBS and WBSE (unit: %)

MTBS	WBSE	Unchanged	Low	Moderate	High	Grass burn	Total
Unburned to Low		< 1	19	< 1	< 1	< 1	21
Low		< 1	14	22	< 1	< 1	37
Moderate		< 1	< 1	17	8	< 1	25
High		< 1	< 1	2	14	< 1	16
Increased Greenness		< 1	< 1	< 1	< 1	< 1	1
Total		1	34	42	22	1	100

Blue: same burn severity classification in MTBS and WBSE.

Pink: higher severity classification in WBSE.

Green: lower severity classification in WBSE.

White: class not comparable between products.

2.3.2 Emissions Estimates

Summary statistics of trace gases and particulate species emitted by wildfires in California during 1984–2020 are shown in Table 2.3. The annual average of CO₂ and PM_{2.5} emitted are 18 Tg and 108 Gg, respectively, with a significant interannual variation. Emissions from wildfires varied significantly and depended on fire size, fire severity, and vegetation characteristics. The largest wildfire was the 2020 August Complex, which burned 417,898 ha in six counties in the Coast Range of Northern California and produced the maximum amount of all types of emissions. This single fire emitted approximately 2.5 times the annual average wildfire emissions for all of California for the past 37 years.

Table 2. 3 Summary statistics of emissions during 1984-2020.

	Min(t)*	Median(Gg)*	Mean(Gg)*	Max(Gg)*	Annual average(Gg)^
CO ₂	8.185	73.59	388.519	48154	17819
CO	0.323	3.856	19.834	2397	910
CH ₄	0.017	0.137	0.798	104	37
NMOC	0.161	1.103	5.734	704	263
SO ₂	0.017	0.043	0.222	27	10
NH ₃	0.003	0.06	0.303	36	14
NO	0.017	0.05	0.245	28	11
NO ₂	0.018	0.048	0.243	29	11
NO _x	0.022	0.072	0.346	38	16
PM _{2.5}	0.041	0.426	2.365	302	108
OC	0.015	0.355	2.302	315	106
BC	0.002	0.032	0.143	15	7

*Min, median, mean, and max is the value in the context of each fire event;

^Annual average is the amount averaged over the 1984-2020 period;

t: metric tons.

Gg: Gigagrams.

We compared our emissions estimates of PM_{2.5} to five global fire emissions inventories (GFEDv4s, FINNv1.5, GFASv1.2, QFEDv2.5r, FEERv1.0-G1.2) that are commonly used in atmospheric modeling simulations by extracting monthly and annual emissions for California from 2003-2020 using the FIRECAM online tool (Liu *et al* 2020). We note that GFEDv4s values assessed here for 2017-2020 emissions are preliminary. Annual emissions were also extracted from California Air Resource Board (CARB 2021) for 2000-2020 and WFEIS for 1984-2020. The estimates agree reasonably well with other emissions inventories, although they may vary by as much as a factor of 3 annually across inventories (Figure 2.7).

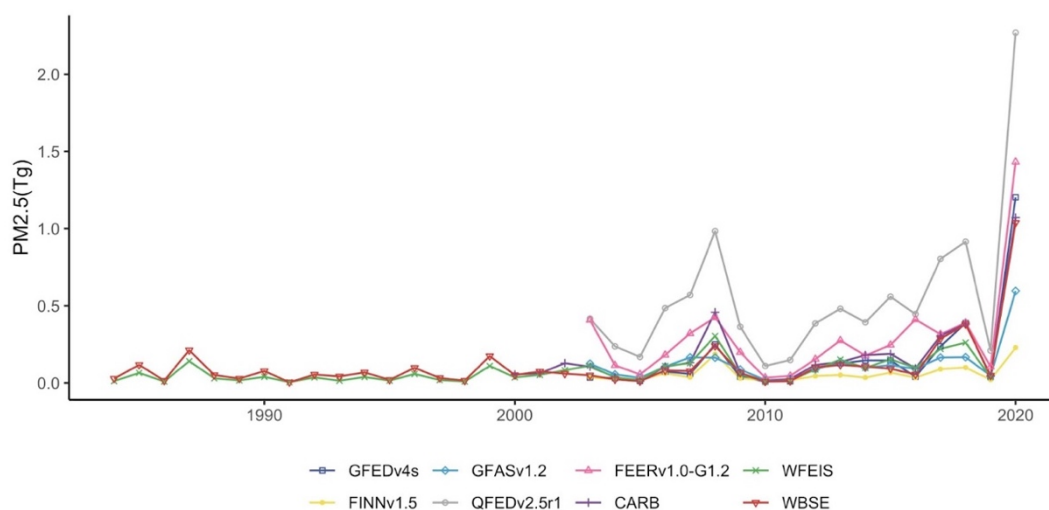


Figure 2. 7 Annual fire $PM_{2.5}$ emissions ($Tg\ yr^{-1}$), over 1984-2020 in California, from five global fire emissions inventories (GFEDv4s, FINNv1.5, GFASv1.2, QFEDv2.5r, FEERv1.0-G1.2), two regional (CARB and WFEIS) inventories, and WBSE

2.3.3 Daily Emissions

We calculated daily emissions for a total of 841 fires since 2002 when MODIS data were available. For 2002-2020, 6% percent of the original FIRMS fire points were removed annually due to low detection confidence. Seventy-four fire events were removed from the daily emissions database due to less than 10 MODIS/VIIRS fire detections falling within the WBSE fire perimeter; the daily emissions from these events are considered uncertain given the low number of fire detections (after Parks 2014). These 74 fires were also small fire events, and their total emissions only accounted for less than 1% of the emissions during 2002-2020. The average daily emissions over the summer during 2002-2020 are higher compared to the average amount ($0.47\ Gg^1$) of daily $PM_{2.5}$ emissions from all other sectors in California over 2002-2017 (Figure 2.8). The extremes that peak around the middle of August and September are primarily contributed by the 2020 wildfires (Figure 2.9).

¹calculated based on air pollutant emissions trends data from the United States Environmental Protection Agency <https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data>

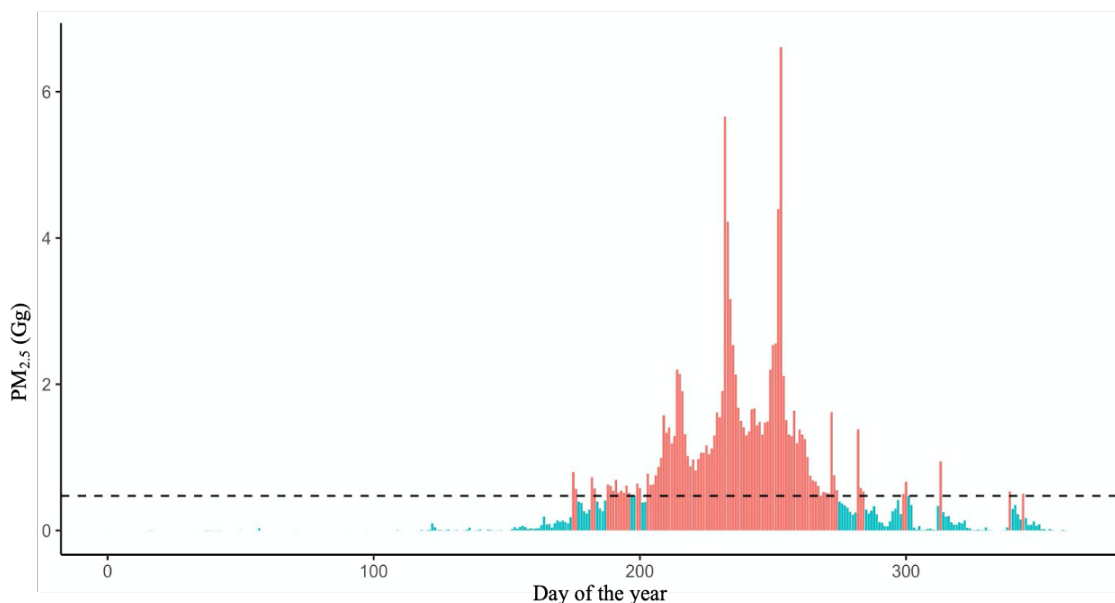


Figure 2. 8 Average daily $PM_{2.5}$ emissions aggregated by day of the year from 2002 to 2020. The dashed line represents the value of average daily $PM_{2.5}$ emissions from all other sectors in California over 2002-2017; Bars colored in red and cyan indicate when average daily emissions from wildfires are above and below the average emissions from all other sectors, respectively

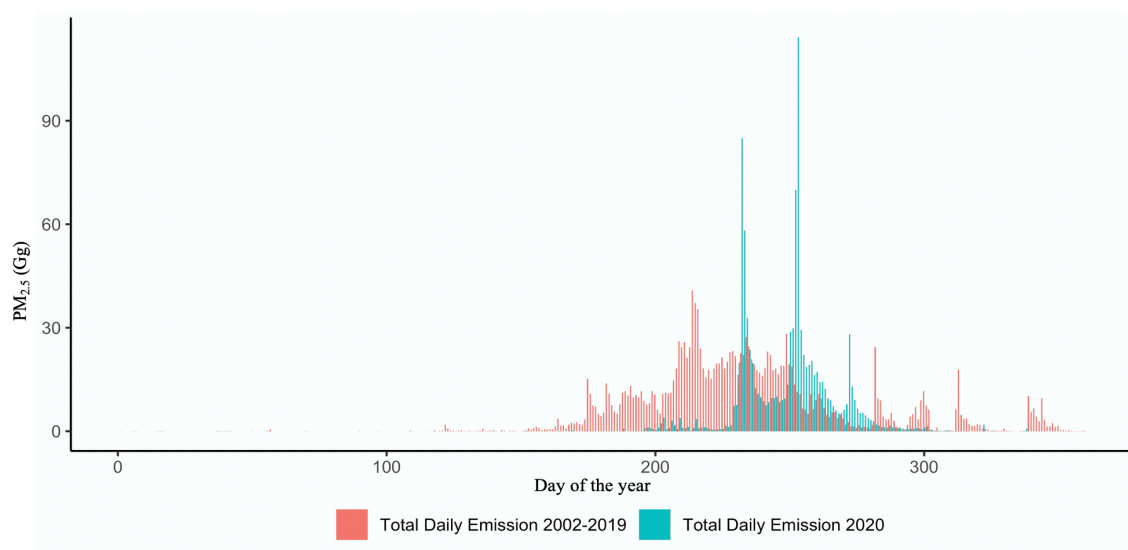


Figure 2. 9 Total emissions from 2002-2019 aggregated by day of year compared to daily emissions in 2020

2.3.4 Limitations and Uncertainties

The results of this research should be considered with some limitations in mind, as uncertainties are associated with several aspects of the calculation process. Uncertainties related to the FINN model, including fire location, timing and area burned, vegetation identification, fuel loading, and fuel consumption, and emission factors, are described in detail by Wiedinmyer *et al* (2011). The

modified FINN model used here reduces uncertainties related to fire numbers and sizes by using documented fire records rather than satellite detections. However, fire perimeters used in this calculation for wildfires during 1984-2019 were generated by on-screen interpretation and delineation of dNBR images by MTBS analysts. We used MTBS perimeters to better compare the burn severity difference between WBSE and MTBS. The results may change slightly when using other documented fire perimeter data; for example, we also calculated emissions since 1984 using perimeters from CAL FIRE (not shown here but available). Uncertainty also arises from misidentification of land cover: (1) the burn severity classification is based on the LF BPS dataset, which characterizes vegetation systems that may have been dominant on the landscape before Euro-American settlement, but may not be representative of current vegetation; (2) fuel loading is calculated based on LF EVT, which represents the current distribution of the terrestrial ecological systems classification, but this dataset is infrequently updated and may not accurately represent the conditions at the time of burning. Fuel loading, consumption, and emission rates were constant values for general vegetation categories and severity classes, which may not reproduce the full heterogeneity of land cover. The day of burning was assigned with an average date in regions of overlapping buffers in this analysis, and the results change slightly if the earliest date or latest date is used as the proper date. In addition, around 5% of emissions were not assigned to a valid date since no fire points were detected by satellites. This may lead to a slight underestimate of daily emissions and thus the derived impact on air quality and public health. The open-source, open data framework presented here can be further refined as needed to address any of these limitations. Likewise, Picotte *et al* (2021) 's severity algorithms implemented here are designed to be robust to limitations in the available plot-level CBI data and can be extended to incorporate new plot-level data as they become available.

2.4 Conclusions

The WBSE framework implemented in R and GEE produces quick estimates of burn severity and emissions from wildfires at a 30-m spatial resolution and day of burning and daily emissions at 500 m spatial resolution. The burn severity estimates have advantages over MTBS products in that: (1) burn severity is categorized consistently based on field-validated equations, rather than subjective thresholds for each fire event; (2) an initial assessment can be estimated as soon as the desired fire perimeter is available; (3) they enable consistent annual estimates for wildfires burned in California during 1984-2020, which facilitates objectively quantifying changes over time and across fires. Further, daily emissions estimates can be calculated for periods when daily fire activity data are available (beginning in 2002). The daily emission estimates are currently being incorporated into chemical transport models to evaluate the impact of wildfire smoke on downwind air quality and population exposure. This work facilitates climate vulnerability assessments that include the effects of smoke on public health and provides estimates of greenhouse gas emissions during a fire. The publicly available WBSE fire severity history and emissions inventory offers a valuable, open-source resource for the fire research community, readily updated after each fire season, with a framework that allows for iterative improvements as more data become available. Future planned improvements include automated fire perimeter mapping of observed fires and incorporating interactive web tools to enable users to estimate fire emissions for simulated fire severity maps with user-modified climate and fuels management. For California's Fifth State Climate Assessment, the Pyrengence Consortium is simulating individual wildfire events (location, date, size, and severity fraction) for future climate, land use and fuels management scenarios, mapping simulated severity at 30-m across the state, and using the framework presented here for estimating emissions from individual large fire simulations through 2100.

Code availability: The code to implement our methods is available at the following GitHub repository: <https://github.com/qxu6/WBSE.git>.

Data availability: Data used for this study could be found from our GitHub repository or downloaded following the R code. All the data generated from this study are available at the following URL: <https://doi.org/10.6071/M3QX18>

References

- Abatzoglou J T and Williams A P 2016 Impact of anthropogenic climate change on wildfire across Western US forests *Proc. Natl Acad. Sci.* **113** 11770–5
- Abatzoglou J T, Williams A P and Barbero R 2018 Global emergence of anthropogenic climate change in fire weather indices *Geophys. Res. Lett.* **46** 326–36
- Akagi S K, Yokelson R J, Wiedinmyer C, Alvarado M J, Reid J S, Karl T, Crounse J D and Wennberg P O 2011 Emission factors for open and domestic biomass burning for use in atmospheric models *Atmos. Chem. Phys.* **11** 4039–72
- Balch J K, Bradley B A, Abatzoglou J T, Nagy R C, Fusco E J and Mahood A L 2017 Human-started wildfires expand the fire niche across the United States *Proc. Natl Acad. Sci.* **114** 2946–51
- Birch D S, Morgan P, Kolden C A, Abatzoglou J T, Dillon G K, Hudak A T and Smith A M S 2015 Vegetation, topography and daily weather influenced burn severity in central Idaho and western Montana forests *Ecosphere* **6** 1–23
- Bowman D M J S *et al* 2009 Fire in the Earth system *Science* **324** 481–4
- Bowman D M J S, Williamson G J, Abatzoglou J T, Kolden C A, Cochrane M A and Smith A M S 2017 Human exposure and sensitivity to globally extreme wildfire events *Nat. Ecol. Evol.* **1** 0058
- Brey S J, Barnes E A, Pierce J R, Wiedinmyer C and Fischer E V 2018 Environmental conditions, ignition type, and air quality impacts of wildfires in the southeastern and Western United States *Earth's Future* **6** 1442–56
- Burke M, Driscoll A, Heft-Neal S, Xue J, Burney J and Wara M 2021 The changing risk and burden of wildfire in the United States *Proc. Natl Acad. Sci.* **118** e2011048118
- California Air Resource Board 2021 Wildfire emissions & burned area estimates 2000–2020 (available at: https://ww2.arb.ca.gov/sites/default/files/2021-07/Wildfire%20Emission%20Estimates%20for%202020%20_Final.pdf) (Accessed 12 December 2021)
- Crotteau J S, Ritchie M W and Varner J M 2014 A mixed-effects heterogeneous negative binomial model for postfire conifer regeneration in Northeastern California, USA *For. Sci.* **60** 275–87
- Darmenov A S and da Silva A M 2015 *The Quick Fire Emissions Dataset (QFED): Documentation of Versions 2.1, 2.2 and 2.4 (Technical Report Series on Global Modeling and Data Assimilation vol 38)* ed R D Koster (Greenbelt, MD: National Aeronautics and Space Administration) p 201
- Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B and Howard S 2007 A project for monitoring trends in burn severity *Fire Ecol.* **3** 3–21
- Faulstich S D, Schissler A G, Strickland M J and Holmes H A 2022 Statistical comparison and assessment of four fire emissions inventories for 2013 and a large wildfire in the Western United States *Fire* **5** 27

- Finlay S E, Moffat A, Gazzard R, Baker D and Murray V 2012 Health impacts of wildfires *PLoS Curr.* **4** e4f959951cce2c
- French N H F *et al* 2014 Modeling regional-scale wildland fire emissions with the wildland fire emissions information system *Earth Interact.* **18** 1–26
- Goodwin M J, Zald H S J, North M P and Hurteau M D 2021 Climate-driven tree mortality and fuel aridity increase wildfire's potential heat flux *Geophys. Res. Lett.* **48** e2021GL094954
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D and Moore R 2017 Google earth engine: planetary-scale geospatial analysis for everyone *Remote Sens. Environ.* **202** 18–27
- Goss M, Swain D L, Abatzoglou J T, Sarhadi A, Kolden C A, Williams A P and Diffenbaugh N S 2020 Climate change is increasing the likelihood of extreme autumn wildfire conditions across California *Environ. Res. Lett.* **15** 094016
- Hurteau M D, Westerling A L, Wiedinmyer C and Bryant B P 2014 Projected effects of climate and development on California wildfire emissions through 2100 *Environ. Sci. Technol.* **48** 2298–304
- Ichoku C and Ellison L 2014 Global top-down smoke-aerosol emissions estimation using satellite fire radiative power measurements *Atmos. Chem. Phys.* **14** 6643–67
- Jaffe D A, O'Neill S M, Larkin N K, Holder A L, Peterson D L, Halofsky J E and Rappold A G 2020 Wildfire and prescribed burning impacts on air quality in the United States *J. Air Waste Manage. Assoc.* **70** 583–615
- Jiang X, Wiedinmyer C and Carlton A G 2012 Aerosols from fires: an examination of the effects on ozone photochemistry in the Western United States *Environ. Sci. Technol.* **46** 11878–86
- Kaiser J W *et al* 2011 Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power *Biogeoscience* **9** 527–54
- Keeley J E and Syphard A D 2019 Twenty-first century California, USA, wildfires: fuel-dominated vs. wind-dominated fires *Fire Ecol.* **15** 1–5
- Key C H 2006 Ecological and sampling constraints on defining landscape fire severity *Fire Ecol.* **2** 34–59
- Key C H and Benson N C 2006 Landscape assessment (LA) *FIREMON: Fire Effects Monitoring and Inventory System* ed D Lutes *et al* (Fort Collins, CO: Department of Agriculture, Forest Service, Rocky Mountain Research Station) p 55 RMRS-GTR-164
- Kolden C A, Lutz J A, Key C H, Kane J T and van Wageningen J W 2012 Mapped versus actual burned area within wildfire perimeters: characterizing the unburned *For. Ecol. Manage.* **286** 38–47
- Kolden C A, Smith A M S and Abatzoglou J T 2015 Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA *Int. J. Wildland Fire* **24** 1023–8
- Larkin N K, Raffuse S M and Strand T M 2014 Wildland fire emissions, carbon, and climate: U.S. emissions inventories *For. Ecol. Manage.* **317** 61–69
- Lindaas J *et al* 2020 Emissions of reactive nitrogen from Western U.S. wildfires during summer 2018 *J. Geophys. Res. Atmos.* **126** e2020JD032657
- Liu T, Mickley L J, Marlier M E, DeFries R S, Khan M F, Latif M T and Karambelas A 2020 Diagnosing spatial biases and uncertainties in global fire emissions inventories: Indonesia as regional case study *Remote Sens. Environ.* **237** 111557
- McLauchlan K K *et al* 2020 Fire as a fundamental ecological process: research advances and frontiers ed G Durigan *J. Ecol.* **108** 2047–69

- Meng R and Zhao F 2017 Remote sensing of fire effects: a review for recent advances in burned area and burn severity mapping *Remote Sensing of Hydrometeorological Hazards* (Boca Raton, FL: CRC Press) pp 261–83
- Miller J D and Thode A E 2007 Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR) *Remote Sens. Environ.* **109** 66–80
- Parks S A 2014 Mapping day-of-burning with coarse-resolution satellite fire-detection data *Int. J. Wildland Fire* **23** 215–23
- Parks S, Holsinger L, Voss M, Loehman R and Robinson N 2018 Mean composite fire severity metrics computed with Google Earth Engine offer improved accuracy and expanded mapping potential *Remote Sens.* **10** 879
- Permar W *et al* 2021 Emissions of trace organic gases from Western U.S. wildfires based on WE-CAN aircraft measurements *J. Geophys. Res. Atmos.* **126** e2020JD033838
- Perry D A, Hessburg P F, Skinner C N, Spies T A, Stephens S L, Taylor A H, Franklin J F, McComb B and Riegel G 2011 The ecology of mixed severity fire regimes in Washington, Oregon, and Northern California *For. Ecol. Manage.* **262** 703–17
- Picotte J J 2019 Composite burn index (CBI) data for the conterminous US, collected between 1996 and 2018: U.S. Geological Survey Data Release (<https://doi.org/10.5066/P91BH1BZ>)
- Picotte J J, Cansler C A, Kolden C A, Lutz J A, Key C, Benson N C and Robertson K M 2021 Determination of burn severity models ranging from regional to national scales for the conterminous United States *Remote Sens. Environ.* **263** 112569
- Pouliot G, Pierce T, Benjey W, O'Neill S M and Ferguson S A 2020 Wildfire emission modeling: integrating bluesky and SMOKE *14th Annual Int. Emission Inventory Conf. 11-14*
- Prichard S J, O'Neill S M, Eagle P, Andreu A G, Drye B, Dubowy J, Urbanski S and Strand T M 2020 Wildland fire emission factors in North America: synthesis of existing data, measurement needs and management applications *Int. J. Wildland Fire* **29** 132
- R Core Team 2021 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing)
- Raffuse S M, Larkin N K and Lahm P W 2012 Development of version 2 of the wildland fire portion of the National Emissions Inventory *20th Int. Emission Inventory Conf.*
- Reid C E, Brauer M, Johnston F H, Jerrett M, Balmes J R and Elliott C T 2016 Critical review of health impacts of wildfire smoke exposure *Environ. Health Perspect.* **124** 1334–43
- Reinhardt E D, Keane R E and Brown J K 1997 *First order fire effects model: FOFEM 4.0, user's guide* (No. 344-345) (US Department of Agriculture Forest Service Intermountain Research Station)
- Rollins M G 2009 LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment *Int. J. Wildland Fire* **18** 235–49
- Stephens S L *et al* 2013 Temperate and boreal forest mega-fires: characteristics and challenges *Front. Ecol. Environ.* **12** 115–22
- Thomas D, Butry D, Gilbert S, Webb D and Fung J 2017 *The Costs and Losses of Wildfires* NIST Special Publication 1215 (Gaithersburg, Maryland, United States of America: National Institute of Standards and Technology) p 1215
- Turner M G, Hargrove W W, Gardner R H and Romme W H 1994 Effects of fire on landscape heterogeneity in Yellowstone National Park, Wyoming *J. Veg. Sci.* **5** 731–42

- Val Martin M, Heald C L, Ford B, Prenni A J and Wiedinmyer C 2013 A decadal satellite analysis of the origins and impacts of smoke in Colorado *Atmos. Chem. Phys.* **13** 7429–39
- van der Werf G R *et al* 2017 Global fire emissions estimates during 1997–2016 *Earth Syst. Sci. Data* **9** 697–720
- van der Werf G R, Randerson J T, Collatz G J, Giglio L, Kasibhatla P S, Arellano A F, Olsen S C and Kasischke E S 2004 Continental-Scale Partitioning of Fire Emissions During the 1997–2001 El Niño/La Niña Period *Science* **303** 73–76
- van der Werf G R, Randerson J T, Giglio L, Collatz G J, Kasibhatla P S and Arellano A F 2006 Interannual variability in global biomass burning emissions from 1997 to 2004 *Atmos. Chem. Phys.* **6** 3423–41
- van der Werf G R, Randerson J T, Giglio L, Collatz G J, Mu M, Kasibhatla P S, Morton D C, DeFries R S, Jin Y and van Leeuwen T T 2010 Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009) *Atmos. Chem. Phys.* **10** 11707–35
- Westerling A L 2016 Increasing Western US forest wildfire activity: sensitivity to changes in the timing of spring *Phil. Trans. R. Soc. B* **371** 20150178
- Wiedinmyer C, Akagi S K, Emmons L K, Al-Saadi J A, Orlando J J, Soja A J and Soja A J 2011 The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning *Geosci. Model Dev.* **4** 625–41
- Williams A P, Abatzoglou J T, Gershunov A, Guzman-Morales J, Bishop D A, Balch J K and Lettenmaier D P 2019 Observed Impacts of Anthropogenic Climate Change on Wildfire in California *Earth's Future* **7** 892–910

3 Spatial and Temporal Patterns of Wildfire Burn Severity and Biomass Burning-induced Emissions in California

Abstract

Wildfires are an important disturbance in the Earth system, and their emissions have regional and even global impacts on radiation, clouds, and climate. The increased frequency and magnitude of California wildfires in recent years is altering the terrestrial carbon cycle, undermining the state's efforts to reduce the Greenhouse Gases (GHGs) to confront climate change. Air quality and public health are also greatly affected by air pollution from wildfires. The severity of wildfire burns is a critical indicator of both their direct and indirect ecological and human impacts. To formulate targeted mitigation strategies, it is imperative to understand the evolving scale, location and timing of wildfire burn severity and emissions. This study analyzed spatial and temporal patterns of burn severity and emissions at 30 m resolution from large wildfires (>404 hectares) burning in California during 1984–2020 from the recently developed Wildfire Burn Severity and Emissions Inventory. Results show vegetation and severity play critical roles in controlling the spatial and seasonal distribution of emissions. California's annual burned area and emissions increased, notably in early and late parts of what once was the typical fire season, although peak wildfire burned area and emissions continue to occur in mid-Summer. Emissions and areas burned in moderate to high severity were particularly high and increasing in North Coast and Sierra Nevada forests. The 2020 fire year—with the most megafires in California history—had 15 times the annual average emissions that occurred during 1984–2015.

Keywords: PM_{2.5}, trend, seasonality, megafires

3.1 Introduction

Fire is a natural phenomenon, sparked by humans as well as natural ignitions such as lightning and other far less common ways, including volcanic activity and meteors. Fire serves an important ecological role in clearing low-growing underbrush, removing organic debris, opening the ground up to sunlight, and nourishing the soil (He *et al* 2019, McLauchlan *et al* 2020). Fire also facilitated human evolution over two million years ago when our ancestors began to control and make uses of fire. Fire and humans have then contributed together to landscape diversity. Not all fires are harmful and need to be extinguished. However, uncontrollable and destructive wildfires that burn with uncharacteristic frequency, extent, and severity due to climate change and legacies of past management can have negative impacts on nature and human society, and are a phenomenon we have been increasingly challenged by in recent decades and expect to continue to face in the near future (Westerling 2006, 2016, 2018, Abatzoglou and Williams 2016, Williams *et al* 2019).

Fire has been a vital force in California's diverse ecosystems for millennia, such as facilitating germination of seeds for certain tree species, shaping the composition of land cover, and altering biological diversity (Van Wagtendonk, 2018). Expanding development in wildland-urban interfaces, fuel accumulations from fire suppression legacies over the last century, increasing temperatures, and more variable precipitation are making wildfires more catastrophic in California, with increases in total and high severity burned area in recent decades (Lutz *et al* 2011, Miller and Safford 2012, Stevens *et al* 2017, Keeley and Syphard 2019, Keyser and Westerling 2019, Goss *et*

al 2020, Swain 2021). In addition to direct damage to the environment and human systems in the wildfire burned area, wildfire severity affects emissions which in turn can have effects on air quality, climate, and public health over a much larger area that can long outlast their flames (Reid *et al* 2016, O'Neill *et al* 2021, Wang *et al* 2021). Recent catastrophic wildfires driven by climate change and human activities also make achieving carbon neutrality more challenging in California. Air regulators and policy makers are increasingly interested in understanding the patterns of wildfire burn severity and emissions, in order to develop plans to reduce the risk of uncharacteristically severe wildfires and to mitigate emissions from wildfires.

Many studies have analyzed patterns of wildfires and fire emissions at various temporal and spatial scales. Li and Banerjee (2021) studied the spatial and temporal distributions of wildfire occurrence in California and found their frequency and total burned area increased significantly during 2000–2019 compared to 1920–1999. They also found temperature, vapor pressure deficit, grass cover, and distance to roads highly affected risks of wildfire occurrence. Chen and Jin (2022) analyzed spatial patterns of wildfire ignitions during 1992–2015 and found precipitation, slope, human settlement, and road network controlled human-caused ignitions while snow water equivalent, lightning density, and fuel amount shaped lightning-caused ignition in California. Stephens *et al* (2007) estimated annual average prehistoric fire emissions in California using a First Order Fire Effects Model by each vegetation type. They found a high amount of wildfire emissions was produced annually before the influences of Euro-American settlement on fuels and ignitions led to a period of relative fire exclusion. Estimates of modern California wildfire emissions found that they exhibited large inter-annual variability, and forests and woodlands represented the largest contributors (Jaffe *et al* 2008, Hurteau *et al* 2014, CARB 2021). Estimated daily fire emissions of CO₂ for 11 states in the western U.S.—including California—during 2001–2008 by Wiedinmyer and Hurteau (2010) indicated that emissions were elevated from July through October with the peak in August. For a medium-high temperature scenario, projected effects of climate change and development on California wildfire emissions through 2100 were projected to increase by 19%–101% over 1961–1990, with the highest increases centered in northern California forests (Hurteau *et al* 2014). Despite all the efforts, the spatiotemporal patterns of large wildfires' burn severity and emission regimes throughout California over the past four decades have not been investigated with high spatial and temporal resolution data.

Here, we take advantage of the recently available Wildfire Burn Severity and Emission Inventory (WBSE), a retrospective inventory that was recently developed to calculate burn severity and emissions at 30 m resolution for each large wildfire observed since 1984 (Xu *et al* 2022). This study attempts to utilize the most up to date data to investigate the spatiotemporal patterns of burn severity and emissions from large wildfires burned in California over the period of 1984–2020 to answer the following questions: (1) How have annual and seasonal patterns of burn severity and emissions changed over the period of 1984–2020? (2) What are the spatial distribution characteristics of burn severity and emissions? (3) What are the contributions of the most extreme fire events to emissions?

3.2 Data and Methods

3.2.1 Study Area

California is the third-largest state in the United States in area, bordered by the Pacific Ocean to the west, and geographically diverse. In the middle of the state lie the fertile farmlands of the Central Valley, bounded by the Sierra Nevada mountains in the east, the Coastal Mountain Ranges in the west, the Klamath Mountains in the northwest, and the Mojave Desert in the southeast. The

climate of most areas in California is described as a Mediterranean climate, characterized by winter precipitation and long dry, hot summers. The natural environmental conditions make California a high fire risk area. There are currently 13 level III ecoregions (Griffith *et al* 2016) and 15 Air Basins in California that we use here for context (Figure 3.1).

3.2.2 Data

The burn severity and emissions data used for this analysis are from the WBSE, a retrospective inventory that was recently developed by Xu *et al* (2022). Historical (1984–2019) remote sensing data previously used by the Monitoring Trends in Burn Severity project (MTBS) to estimate burn severity were calibrated by WBSE based on regression relationships developed by Picotte *et al* (2021) between the field-based burn severity index-the Composite Burn Index (CBI), and Landsat imagery-derived differenced normalized burn ratio (dNBR)/normalized burn ratio to produce objective, validated severity classifications at 30 m resolution. For large wildfires in 2020 recorded by the California Department of Forestry and Fire Protection (CAL FIRE), burn severity maps were also generated with a Landsat dNBR calculator developed for WBSE in order to extend the record through the very active 2020 fire season (MTBS data for comparison were not yet available for 2020). WBSE burn severity products differ from MTBS in that the spectral indices are first converted to CBI and then classified with standard thresholds other than the subjective interpretation per fire based on the experience of MTBS analysts. The WBSE burn severity dataset was generated following the regression models developed by Picotte *et al* (2021), where the details of accuracy and evaluation of the regression models for each landcover classification were presented. There were only 12 wildfires in California during 1984–2020 that have CBI field data (collected between 1996 and 2018). Thus, the companion paper by Xu *et al* (2022) compared the WBSE dataset with the burn severity dataset from the MTBS project. Their pixel-level classification comparison indicated that around 45% of the pixels in both datasets were classified as the same burn severity. The rest generally have a higher burn severity in WBSE than do MTBS products.

WBSE calculated the amount of particulate matter (PM_{2.5}), greenhouse gases (CO₂, CH₄), and other trace gases emitted from each large wildfire (>400 ha) as a function of the area burned by severity classification, fuel consumption under low, moderate, and high severity, and emission factors for coarse vegetation types, using a modification of the Fire Inventory from NCAR (FINN) model (Wiedinmyer *et al* 2011). Fuel type was assigned to five general vegetation categories (grass, shrub, forest under 5500 feet (1676 m), forest between 5500 and 7500 feet (1676–2286 m), and forest above 7500 feet (2286 m)) using the LANDFIRE existing vegetation type (EVT) product (Rollins 2009). Fuel loadings and consumption rates for the five vegetation classes were assigned under low, moderate, and high severity. Normalized emission factors were compiled and updated for coarse vegetation types in California eco- systems. Fire event-based emissions were estimated at 30 m resolution. The estimates agree reasonably well with global and regional emissions inventories (Xu *et al* 2022).

3.2.3 Statistical Estimation of Annual Trends

The non-parametric Mann–Kendall test (Mann 1945, Kendall 1975) and Sen’s slope estimator (Sen 1968) were used to determine if a trend exists and to estimate the magnitude of the changes. Non-parametric methods were considered to be effective here since they are less influenced by extremes, and do not have to meet the assumption of normality. Both methods were applied together to detect consistent increasing or decreasing trends on a variety of metrics of burn severity and emissions data (annual total burned area, annual area burned and fractions in low, moderate, and high severity,

annual total emissions, and emissions from different fire severities and land covers). Results were considered significant at $p < 0.05$.

3.3 Results and Discussions

3.3.1 The Spatial Distribution of Burn Severity and Emissions

WBSE provides 30 m resolution burn severity and 12 types of biomass burning-induced emissions. Annual emissions of CO_2 , CO, CH_4 , non-methane organic compounds, SO_2 , NH_3 , NO, NO_2 , nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$), $\text{PM}_{2.5}$, OC and BC averaged about 17.819, 0.910, 0.037, 0.263, 0.010, 0.014, 0.011, 0.011, 0.016, 0.108, 0.106, and 0.007 million tons, respectively. Annual emissions of all types are summarized in supplementary Table 3.1. Spatial and temporal patterns of the 12 types of emissions were the same except for the scale, due to the fact that the modified FINN model calculates emissions as a function of burned area, vegetation consumption, and emission factors. We use $\text{PM}_{2.5}$ as an example to illustrate the patterns of emissions since it is a major concern due to its potential for public health impacts. The spatial distribution of burn severity and fire-induced emissions was highly heterogeneous in different ecoregions and air basins in California (Figure 3.1). The North Coast Ranges, Southern California Mountains, and Sierra Nevada are regions of large burned areas. The spatial pattern of $\text{PM}_{2.5}$ emissions is similar to that of burned area. However, the magnitude of emissions did not resemble that of burn severity in the Southern Coast due to the fact that the dominant shrub eco- system there emits fewer emissions compared to the forest ecosystems that are characteristic of the north- ern mountain regions.

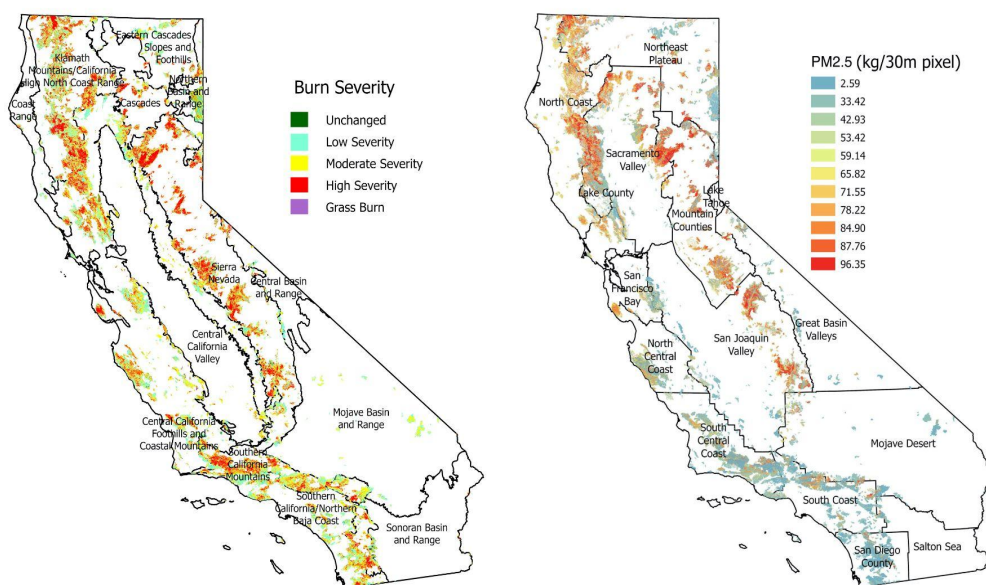


Figure 3. 1 The spatial distribution of burn severity in 13 ecoregions and fire induced $\text{PM}_{2.5}$ emissions in 15 air basins during 1984-2020 in California for all fires over 404 ha. Some burned area and emissions 30m pixels can overlap since the same place could be burned more than once between 1984 and 2020

3.3.2 Annual Severity and Emissions

Annual area burned from large wildfires (>404 ha) averages about 0.23 million ha, with a range of 0.008–1.63 million ha during 1984–2020. Area burned significantly increased during the past four decades at a mean annual rate of 0.07 million ha (Figure 3.2; supplementary Table S3.1). On average, 32%, 41% and 25% burned in low, moderate, and high severity, respectively. Area burned in grassland only accounted for 1% of total area burned. The remaining 1% of total area burned was in barren-rock/sand/clay, open water, or perennial ice/snow, which were categorized as unchanged. Annual area burned in low, moderate, and high severity also significantly increased, but there was no significant trend in the fractions of each severity category over the available record. The year 1991 had the smallest number (8) of large wildfires while the year 2008 had the largest number (102) of large wildfires (Figure 3.2). On average, 46 large wildfires burned each year in California. However, there is no significant trend of annual large fire numbers (supplementary Table S3.2).

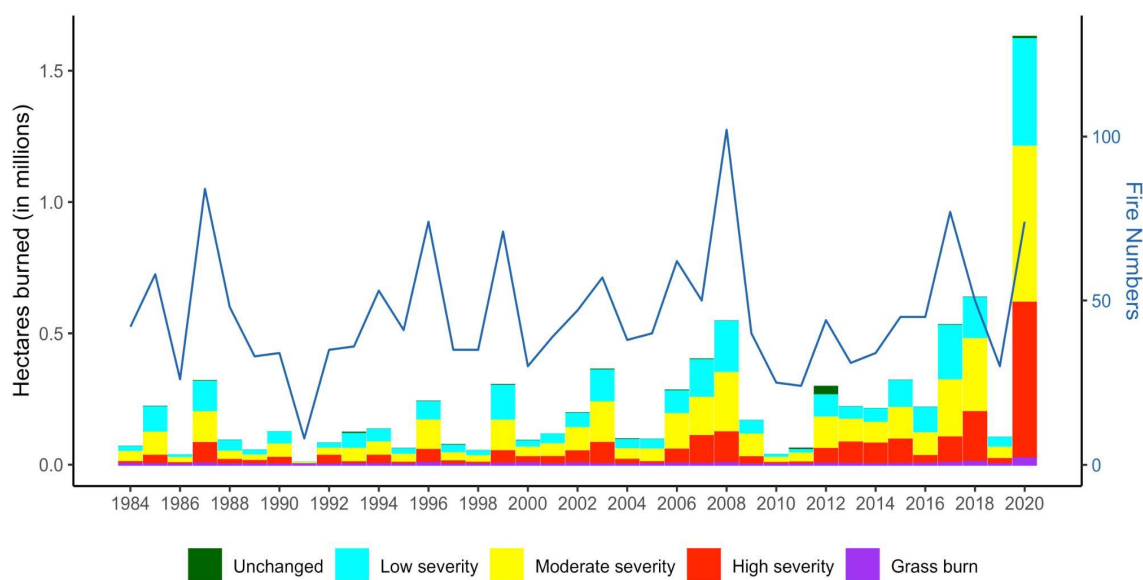


Figure 3. 2 Annual area burned in unchanged, low, moderate, high severity, and grassland of large wildfire Fires (>404 ha) during 1984-2020. The blue indicates annual large wildfire numbers

Nearly half of the area burned in low, moderate, and high severities was in shrubland before 2002 but decreased to 13% during the period of 2002–2015 (Figure 3.3). The fraction of area burned in low elevation forests (<5500 ft) and mid elevation forests (5500–7500 ft) increased, especially in low and moderate severity, from 45% before the year of 2002 to 78% during the period of 2002–2015. During the last five years (2016–2020), shrubs and forests (at low and moderate elevation) burned an average of 26% and 54% of total area, similar to the average percentage burned during the past four decades. The high elevation forests burned only about 2% of total area during 1984–2020. However, the annual average area burned in those high elevation forests has increased about four times since 2008, greatly contributed to by megafires that burned in the year of 2020. This increase is comparable in scale to what has been previously projected for these forests in some scenarios for end-of-century climate change (Westerling 2018).

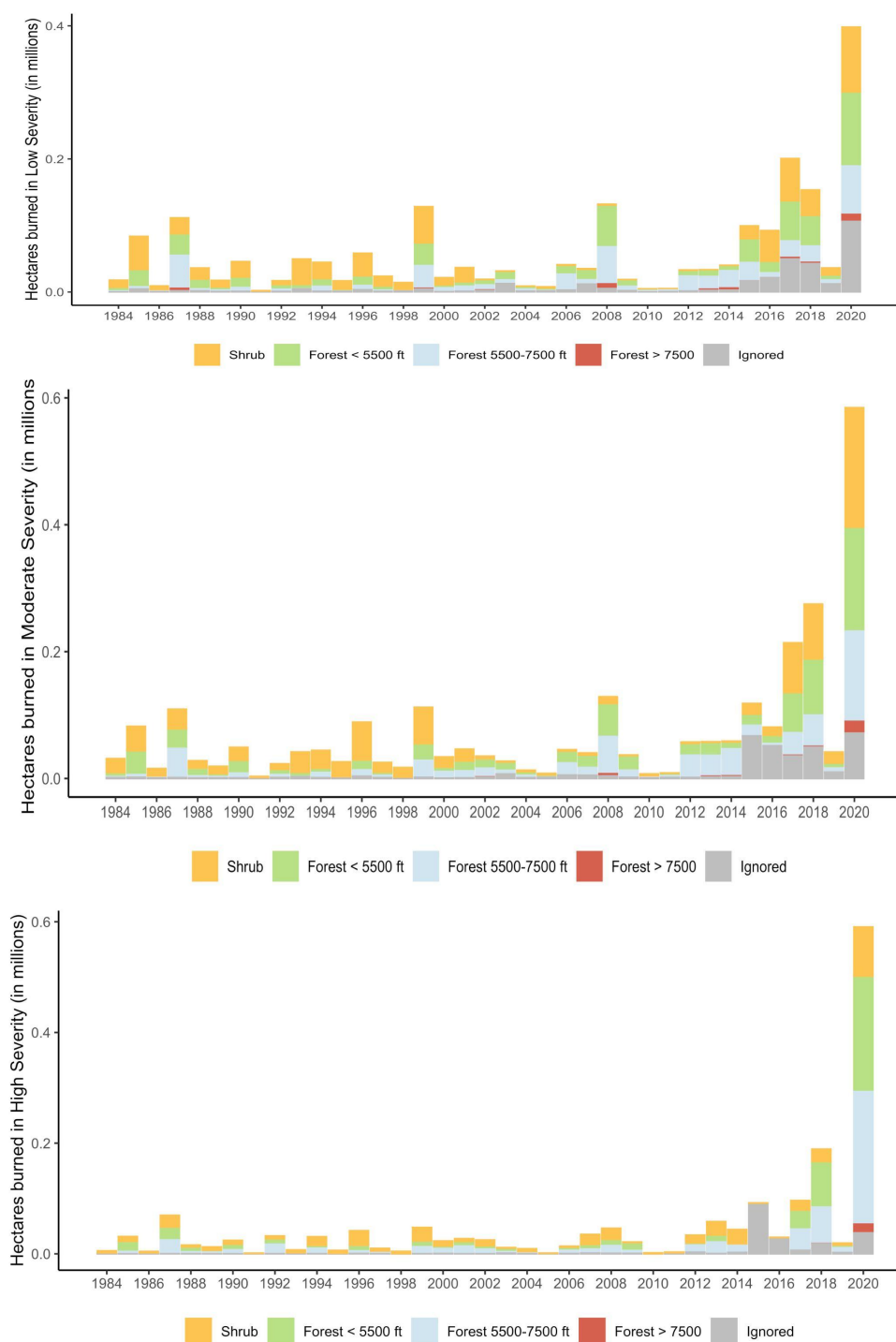


Figure 3. 3 Annual area burned in shrub, low elevation forest (< 5500 ft), moderate elevation forest (5500-7500 ft), high elevation forest (> 7500 ft) in low (top), moderate (middle), and high (bottom) severity during 1984-2020. Grey bars indicated areas ignored by WBSE when classifying LANDFIRE Existing Vegetation Type to coarse land cover types, and not classified as grass burn as the burn severity. Those areas are mainly developed or vegetation types that were not considered in the emission model

Emissions from low, moderate, and high severity burns are coupled with area burned in the corresponding severity category (Figure 3.2 & Figure 3.4). PM_{2.5} emissions showed considerable inter-annual variations, with a significant increasing trend at a mean rate of about 1.98 Gg yr⁻¹ over 1984-2020 (supplementary Table S3.2). California's annual PM_{2.5} emissions from large wildfires ranged between 4 and 1038 Gg yr⁻¹ during 1984-2020, with an average of 108 Gg yr⁻¹. Emissions from the last five years (2016-2020) are more than five times the average annual emissions during 1984-2015. California experienced a record-breaking total area burned in large wildfires in the year of 2020, which led to fifteen times the amount of annual average emissions compared to 1984-2015.

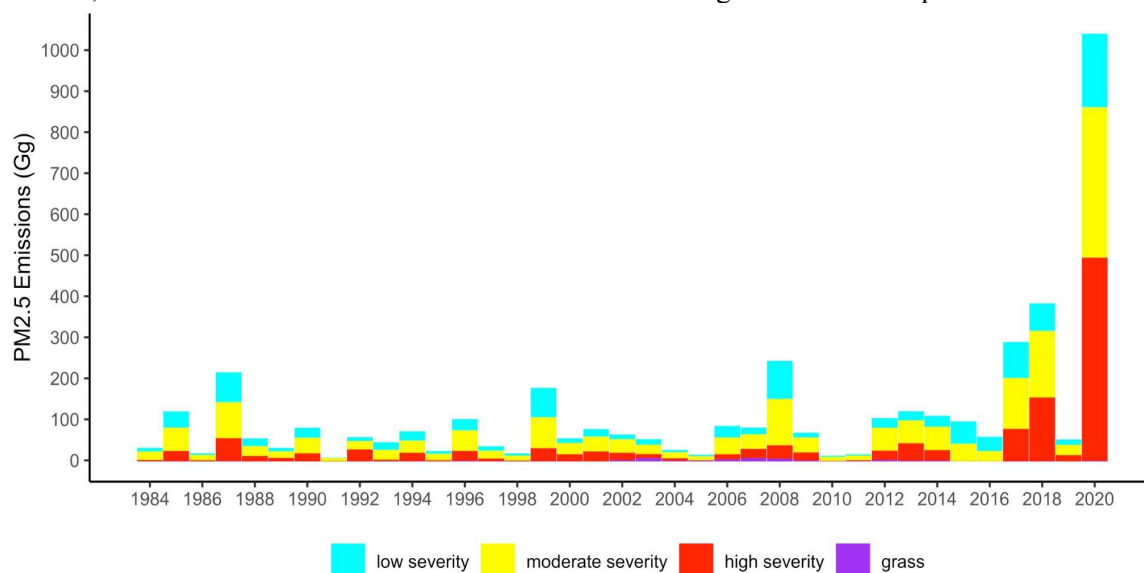


Figure 3. 4 Annual PM_{2.5} emissions from low, moderate, and high severities during 1984-2020

In total, 80.9%, 17.7%, and 1.4% of emissions were emitted from forests, shrublands, and grasslands, respectively. Annual area burned in forests was ranged from 0.004 to 0.99 million ha, accounting for 7 - 67% of total area burned, and 22-95% of annual emissions (Figure 3.5). There is a significant trend of 2 Gg yr⁻¹ increase from forest over 1984-2020 (supplementary Table S3.2). Forests, specifically low and mid elevation forests burned in moderate and high severity, are the major sources of emissions (Figure 3.6). Emissions from moderate severity burned areas in mid elevation forests significantly increased at 0.3 % per year and accounted for 16% of total emissions during 1984-2020.

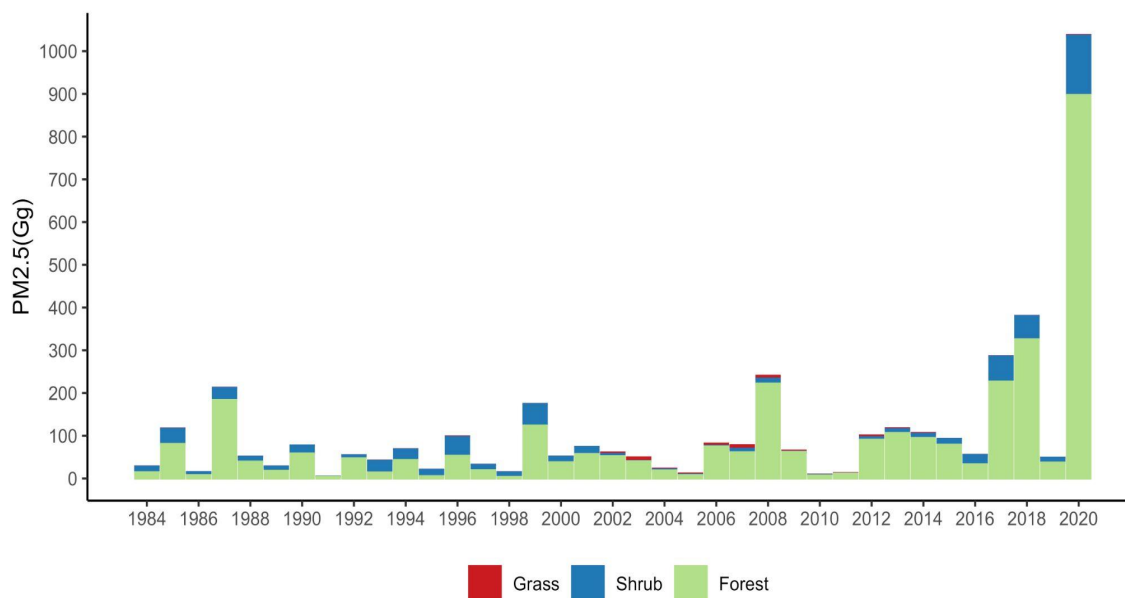


Figure 3. 5 Annual PM_{2.5} emissions from grassland, shrubland, and forest during 1984-2020

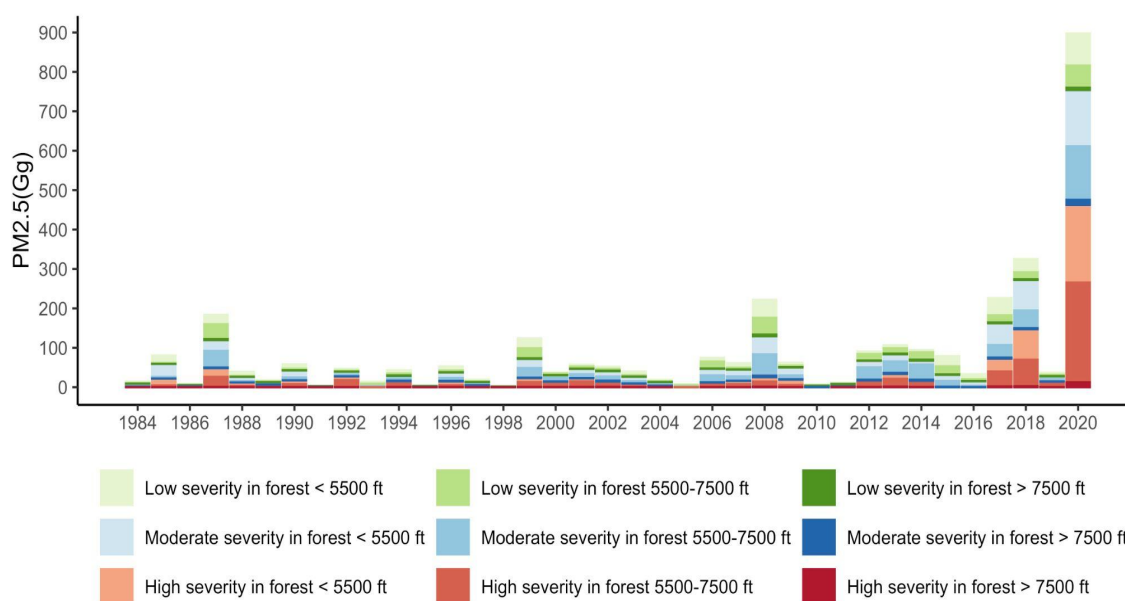


Figure 3. 6 Annual PM_{2.5} emissions from the area burned in low, moderate, high severity in low elevation forest (< 5500 ft), mid elevation forest (5500-7500 ft), and high elevation forest (> 7500 ft) during 1984-2020

3.3.3 Emission Seasonality

Some years wildfires burn all year round. Large wildfires burn frequently starting early summer through the end of the autumn in most parts of California (Figure 3.7). The peak emission month

was in August during 1984-2001, and then shifted to July during 2002-2019 (Figure 3.8). The peak month was back to August in the year of 2020 with multiple megafires including the August Complex fires, North Complex fires, SCU Complex fires, and LNU Lightning Complex fires that all ignited in Northern California. The peak month emissions from this single year are at a similar magnitude to the peak month emissions cumulated over 18 years during the 1984-2001 and the 2002-2019 periods (Figure 3.8). While peak emissions continue to coincide with fires beginning in July or August, the scale of emissions from fires starting earlier and later in the season has increased substantially over the past two decades, with very substantial increases in June (~400%) and in October - December (~125%), comparing 2002-2020 to 1984-2001. Note, however, that fires beginning in any particular month may have some portion of their emissions occurring over subsequent months.

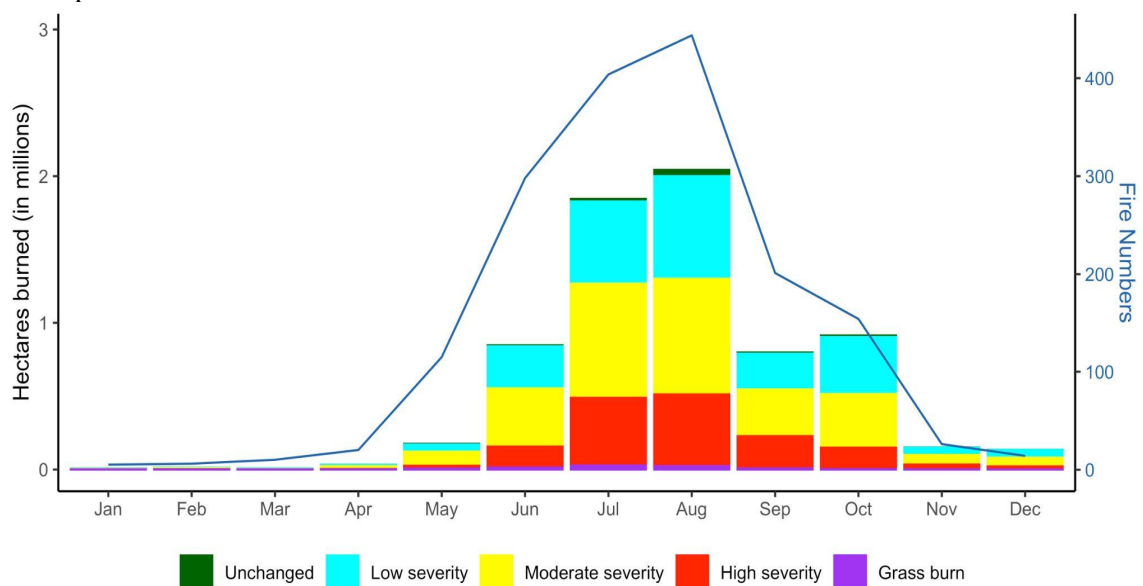


Figure 3. 7 Annual area burned in unchanged (fire refugia within perimeters), low, moderate, and high severity, and grass burn categories of large wildfires (>404 ha) in California during 1984-2020. Blue line shows the total numbers of large wildfires burned in each month during the whole period

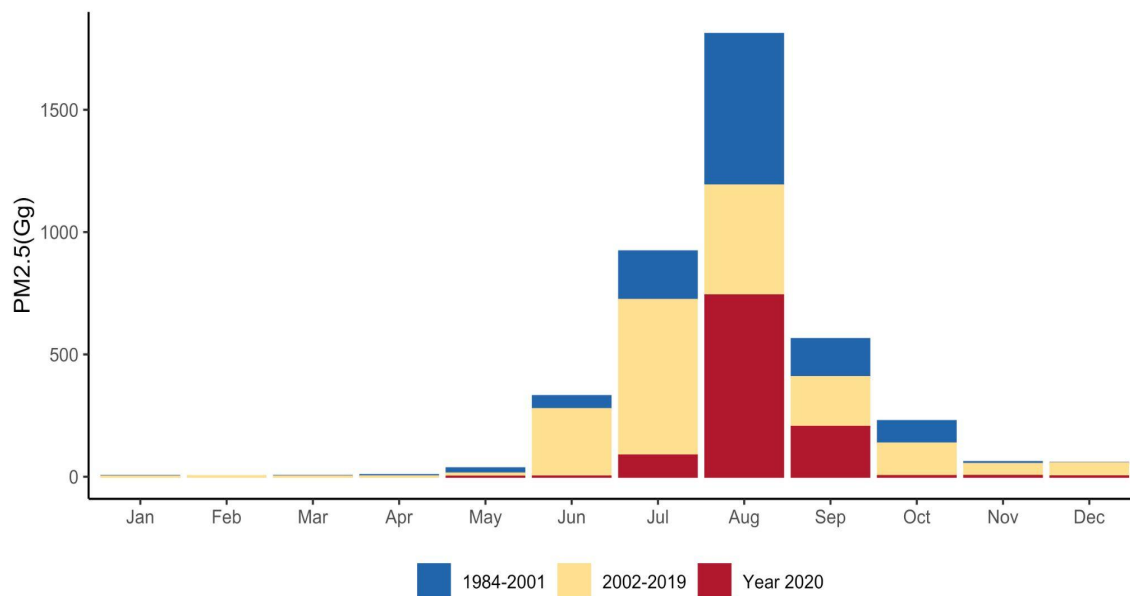


Figure 3. 8 Cumulative total emissions by month during the periods of 1984-2001, 2002-2019, and the year of 2020

The seasonal distribution of emissions is also uneven across California. Figure 3.9 shows the spatial distribution of $PM_{2.5}$ emissions from large wildfires ignited in summer and autumn over the 1984-1990 (1980s), 1991-2000 (1990s), 2001-2010 (2000s), and 2011-2020 (2010s) periods. The peak emissions start month usually occurred in summer in Northern California, and later in Fall in the Southern California region. There was a large increase of areas burned with low emissions in the South Coast region in the summers during the 2001-2010 period. The greatest increase occurred in the summers and autumns during the 2011-2020 period in the North Coast region, areas largely in forest ecosystems under federal management. Changes were relatively small in the Great Basin Valleys, Mojave Desert, and Salton Sea regions, which are also sometimes classified as the Southeastern Desert bioregion, which usually has low fire risks due to the natural limitation of fuel availability.



Figure 3. 9 The spatial distribution of PM_{2.5} emissions from large wildfires ignited during the summer (top) and autumn (bottom) over the 1984-1990 (1980s), 1991-200(1990s), 2001-2010 (2000s), and 2011-2020 (2010s) periods

The seasonal distribution of emissions from grass, shrub, and forest ecosystems is consistent with the burn patterns. The majority of emissions were from forests burning throughout the year (Figure 3.10). The fraction of emissions from shrubland in December is relatively higher compared to the other months due to the 2017 massive Thomas wildfire that mostly burned in shrub vegetation types.

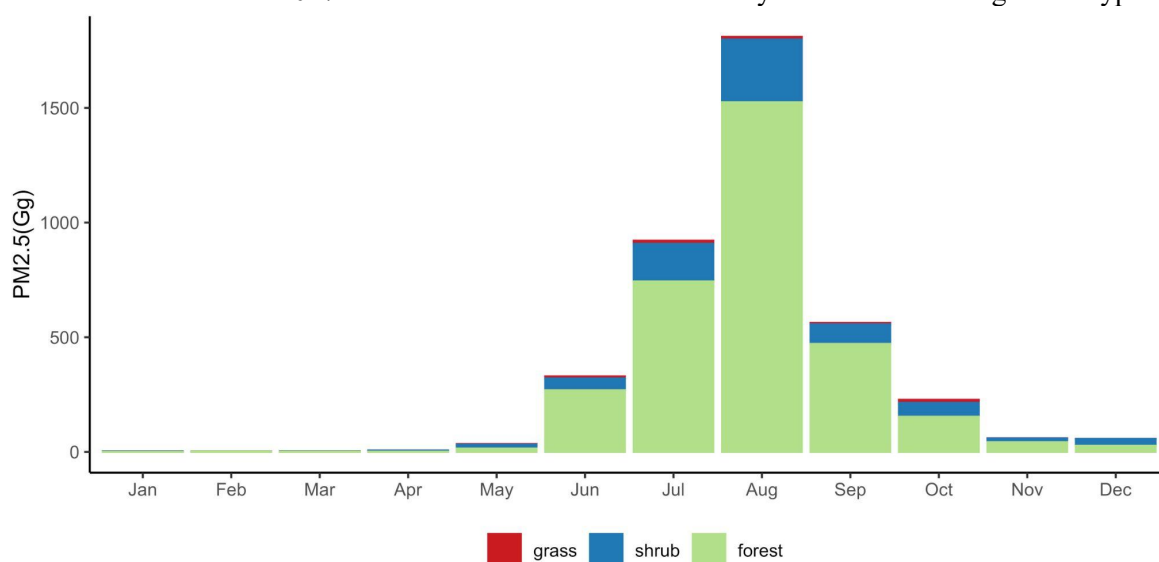


Figure 3. 10 Monthly total emissions from grass, shrub, and forest ecosystems during 1984-2020

3.3.4 Extreme Events

Extremely large wildfires, or *megafires*, are the major contributors to fire emissions. The concept of megafires has not been formally defined but most frequently characterized by size. They have been classified by fire size larger than 10 000 ha (24 710 acres) (Jones *et al* 2016, Stevens *et al* 2017) and over 40 468 ha (100 000 acres) in more recent years (Weber and Yadav 2020, Jones and McDermott 2021, Calhoun *et al* 2022). We defined megafires as fire size over 40 468 ha here to capture the effects of the most extreme fires. We further grouped the 1697 large wildfires included in the analysis into six groups based on size: less than 1000 ha, 1000–5000 ha, 5000– 10 000 ha, 10 000–20 000 ha, 20 000–40 468 ha, and over 40 468 ha (with fires over 40 468 ha defined here as megafires). Figure 3.11 shows the numbers of fires in each size group and the cumulative percent- ages of emissions. The less than 1000 ha group burned with the largest number (739) of large wildfires but only accounted for 4% of total emissions. The numbers of fires decrease rapidly with increasing size class, while the emissions contribution increases with fire size. There was not much difference in the fractions of vegetation burned among size groups (supplementary Figure S3.1). Increasing wildfire size has been the primary driver of increased emissions.

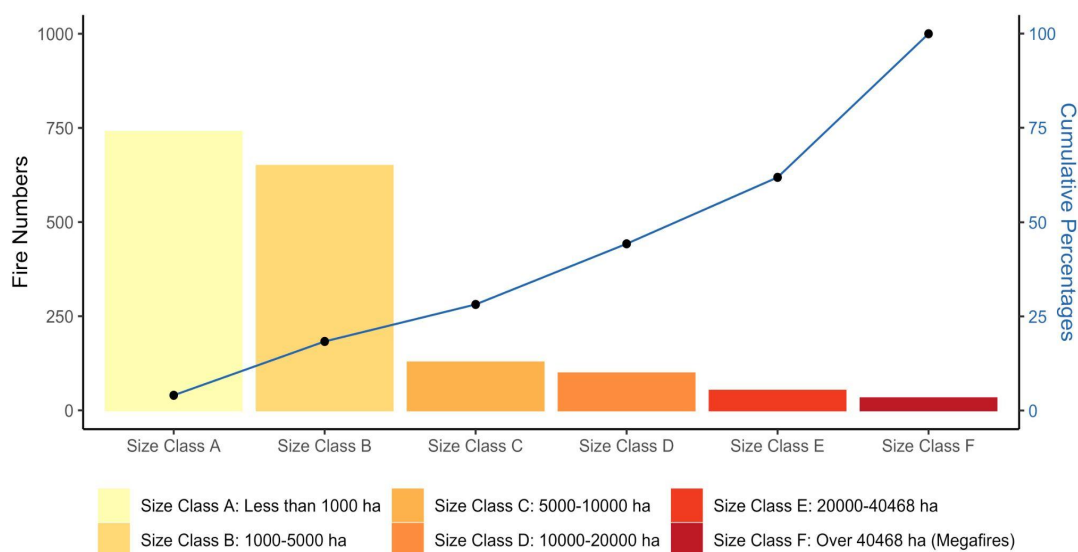


Figure 3. 11 Fire numbers (bars) and emission contribution pareto graph (line). Large wildfires are classified into six groups based on size (less than 1000 ha, 1000 - 5000 ha, 5000 -10000ha, 10000 - 20000 ha, 20000 - 40468 ha, and over 40468 ha (Megafires))

The size of the annual largest fire ranges from 0.002 million ha to 0.42 million ha with a significant increasing trend (supplementary Figure 3.2). Those fires accounted for about 25% of the total PM_{2.5} emissions (Figure 3.12). Some years there are multiple megafires, defined here as a fire larger than 40 468 ha (100 000 acres). Those megafires (not including megafires that were also classified as the annual largest fire) contributed about 17% of the total emissions. Overall, megafires accounted for about 40% of total emissions, and 10 of the 32 total megafires burned in the year of 2020, accounting for 21% of total emissions from all large wildfires during 1984–2020.

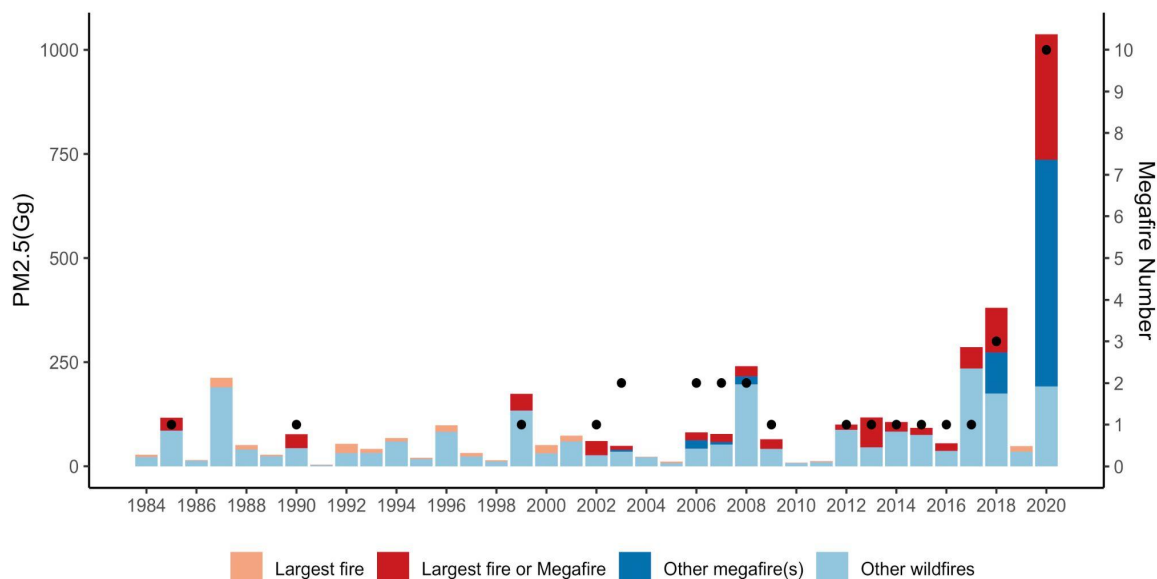


Figure 3.12 Annual PM_{2.5} emissions from largest fire, megafires, and other large wildfires (> 404 ha) during 1984–2020. Black dots show megafire numbers

3.3.5 Limitations

The models used to estimate emissions and severity were developed using different input data: fuel loadings were assigned based on the most appropriate version of LANDFIRE EVT products while burn severities were classified based on LANDFIRE Bio-physical Settings (BPS) products. This resulted in differences in land cover classification of some pixels, as discussed in Xu *et al* this issue. Emissions were calculated assuming grass burn when either EVT or BPS identified the pixel as grassland, except when a pixel is classified as barren-rock/sand/clay, open water, or perennial ice/snow in BPS. Emissions from 22% total area burned in low, moderate, or high severity were calculated as grass burn, a much larger fraction than the fraction of grass burn classified in burn severity calculations (Figure 3.1 and supplementary Figure S3.2). This may lead to an underestimate of the emissions since fuel consumption rates of shrubland and forest are about 12–25 times the number from grassland. The magnitude of changes in emissions varies among different types of emissions depending on emission factors. The current inventory also included 0.2% of total area burned as grass burned in emissions calculations, accounting for 0.1% of emissions. Also, about 11% of burned areas were ignored in emission calculations due to the fact that urban and some other vegetation types were not considered in the current methodology (see supplementary Figure S3.3 for a map of burned area in different vegetation classes and ignored area). A great fraction of total area burned (on average 33% each year) were ignored in emission calculation over the 2016–2020 period, while this number only averages 5% each year during 1984–2015. The increase of emissions is still large even though we underestimate the total emission during the recent five years. This also implies that more urban ecosystems were damaged by recent wildfires. Future versions of these open-source models will seek to reconcile the vegetation types used to calculate fire severity and emissions, and to assign reasonable emissions factors to areas burned in developed areas.

3.4 Conclusions

Burn severity and emissions have significant implications for assessing fire effects on ecosystems. Using the data from a new wildfire burn severity and emissions inventory, we systematically assessed the temporal and spatial distribution of California's large wildfires over the last four decades. The main conclusions from this study can be summarized as follows:

1. The spatial distribution of burn severity and fire-induced emissions was highly heterogeneous, with the North Coast Ranges, Southern California Mountains, and Sierra Nevada being the hot spots and main sources. The spatial pattern of PM_{2.5} emission is similar to that of burned area, albeit the magnitude of emissions did not resemble that of burn severity in the Southern Coast.
2. On average, 32%, 41%, and 25%, respectively, burned in low, moderate, and high severity. Nearly half of the areas burned in low, moderate, and high severities were shrubland before 2002. The fraction of area burned in low elevation forests (< 5500 ft) and moderate elevation forests (5500-7500 ft) increased, especially in low and moderate severity, during the period of 2002-2015. During the last five year (2016-2020), shrub, and low and mid elevation forests burned at the average rate of the past four decades.
3. Emissions from low, moderate, and high severity burns are coupled with area burned in the corresponding severity category. PM_{2.5} emissions averaged 108 Gg yr⁻¹ with considerable inter-annual variations, and a significant increasing trend. Emissions from the last five years (2016-2020) are more than five times the average annual emissions during 1984-2015. The year of 2020 had fifteen times the amount of annual average emissions during 1984-2015.
4. Most emissions were emitted from forests. There is a significant increasing trend from forest emissions over 1984-2020 (Supplementary Table S3.2). Forests, specifically low and mid elevation forests burned in moderate severity, are the major sources of emissions. However, emissions from high elevation forests appear to be increasing rapidly over the last decade of the record, and this potential emerging trend merits close attention in coming years.
5. Large wildfires burned frequently, starting in early summer through the end of the autumn in most parts of California. The peak emissions month was in August during 1984-2001, and then shifted to July during 2002-2019. The peak month was back to August in the year of 2020 with multiple megafires that all ignited in Northern California. The seasonal distribution of emissions is also uneven across California. The peak emissions start month usually occurred in summer in Northern California, and later in fall in the Southern California region. The seasonal distribution of emissions from grass, shrub, and forest ecosystems is consistent with the burn patterns.
6. Extreme fires such as megafires, have serious consequences for ecosystems and human society. Their impacts far exceed those of more numerous smaller fires. Less than 2% of those extreme events accounted for about 40% of the total emissions. Megafires are also burning more high elevation forests than in previous decades.

Wildfires are a natural and necessary component of California's ecosystems, and play complex roles with biological, physical, climatic, and social elements. As a warming climate, increasing

variability in precipitation, and changing land use patterns continue to alter the ecosystems that sustain fire, California will also continue to see changing patterns of wildfire severity, size, timing, location and emissions. However, understanding the changing spatio-temporal patterns of burn severity and emissions can help to facilitate management actions to mitigate the risks of uncharacteristically severe large wildfires and their impacts. The fire and emissions inventory data analyzed here are being used to calibrate models for simulating extreme fire events over a range of future climate, development footprint, and fuels management scenarios for California's Fifth State Climate Assessment, with the goal of providing a guide to public and private resource managers, policymakers and communities undertaking these mitigation efforts.

References

- Abatzoglou J T and Williams A P 2016 Impact of anthropogenic climate change on wildfire across western US forests *Proc. Natl Acad. Sci.* **113** 11770–5
- Calhoun K L, Chapman M, Tubbesing C, McInturff A, Gaynor K M, Van Scoyoc A, Wilkinson C E, Parker-Shames P, Kurz D and Brashares J 2022 Spatial overlap of wildfire and biodiversity in California highlights gap in non-conifer fire research and management *Divers. Distrib.* **28** 529–41
- California Air Resource Board 2021 Wildfire emissions & burned area estimates 2000–2020 (available at: https://ww2.arb.ca.gov/sites/default/files/2021-07/Wildfire%20Emission%20Estimates%20for%202020%20_Final.pdf) (Accessed 12 December 2021)
- Chen B and Jin Y 2022 Spatial patterns and drivers for wildfire ignitions in California *Environ. Res. Lett.* **17** 055004
- Goss M, Swain D L, Abatzoglou J T, Sarhadi A, Kolden C A, Williams A P and Diffenbaugh N S 2020 Climate change is increasing the likelihood of extreme autumn wildfire conditions across California *Environ. Res. Lett.* **15** 094016
- Griffith G E, Omernik J M, Smith D W, Cook T D, Tallyn E, Moseley K and Johnson C B 2016 Ecoregions of California (poster): U.S. geological survey *Open-File Report 2016–1021, with map, scale 1:1,100,000* (<https://doi.org/10.3133/ofr20161021>)
- He T, Lamont B B and Pausas J G 2019 Fire as a key driver of Earth's biodiversity *Biol. Rev.* **94** 1983–2010
- Hurteau M D, Westerling A L, Wiedinmyer C and Bryant B P 2014 Projected effects of climate and development on California wildfire emissions through 2100 *Environ. Sci. Technol.* **48** 2298–304
- Jaffe D, Chand D, Hafner W, Westerling A and Spracklen D 2008 Influence of fires on O₃ concentrations in the Western U.S. *Environ. Sci. Technol.* **42** 5885–91
- Jones B A and McDermott S 2021 The local labor market impacts of US megafires *Sustainability* **13** 9078
- Jones G M, Gutiérrez R, Tempel D J, Whitmore S A, Berigan W J and Peery M Z 2016 Megafires: an emerging threat to old-forest species *Front. Ecol. Environ.* **14** 300–6
- Keeley J E and Syphard A D 2019 Twenty-first century California, USA, wildfires: fuel-dominated vs. wind-dominated fires *Fire Ecol.* **15**
- Kendall M G 1975 *Rank Correlation Methods* 4th edn (London: Griffin) p 202
- Keyser A R and Westerling A L 2019 Predicting increasing high severity area burned for three forested regions in the western United States using extreme value theory *For. Ecol. Manage.* **432** 694–706
- Li S and Banerjee T 2021 Spatial and temporal pattern of wildfires in California from 2000 to 2019 *Sci. Rep.* **11** 8779

- Lutz J A, Key C H, Kolden C A, Kane J T and van Wagtendonk J W 2011 Fire frequency, area burned, and severity: a quantitative approach to defining a normal fire year *Fire Ecol.* **7** 51–65
- M O S *et al* 2021 A multi-analysis approach for estimating regional health impacts from the 2017 Northern California wildfires *J. Air Waste Manage. Assoc.* **71** 791–814
- Mann H B 1945 Nonparametric tests against trend *Econometrica* **13** 245–59
- McLauchlan K K *et al* 2020 Fire as a fundamental ecological process: research advances and frontiers *J. Ecol.* **108** 2047–69
- Miller J D and Safford H 2012 Trends in wildfire severity: 1984–2010 in the Sierra Nevada, Modoc Plateau, and southern Cascades, California, USA *Fire Ecol.* **8** 41–57
- P K S 1968 Estimates of the regression coefficient based on Kendall's tau *J. Am. Stat. Assoc.* **63** 1379–89
- Picotte J J, Cansler C A, Kolden C A, Lutz J A, Key C, Benson N C and Robertson K M 2021 Determination of burn severity models ranging from regional to national scales for the conterminous United States *Remote Sens. Environ.* **263** 112569
- Reid C E, Brauer M, Johnston F H, Jerrett M, Balmes J R and Elliott C T 2016 Critical review of health impacts of wildfire smoke exposure *Environ. Health Perspect.* **124** 1334–43
- Rollins M G 2009 LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment *Int. J. Wildland Fire* **18** 235
- Stephens S L, Martin R E and Clinton N E 2007 Prehistoric fire area and emissions from California's forests, woodlands, shrublands, and grasslands *For Ecol. Manage.* **251** 205–16
- Stevens J T, Collins B M, Miller J D, North M P and Stephens S L 2017 Changing spatial patterns of stand-replacing fire in California conifer forests *For Ecol. Manage.* **406** 28–36
- Swain D L 2021 A shorter, sharper rainy season amplifies California wildfire risk *Geophys. Res. Lett.* **48**
- Van Wagtendonk J W, Sugihara N G, Stephens S L, Thode A E, Shaffer K E and Fites-Kaufman J A 2018 *Fire in California's Ecosystems* (Oakland, CA: University of California Press)
- Wang D *et al* 2021 Economic footprint of California wildfires in 2018 *Nat. Sustain.* **4** 252–60
- Weber K T and Yadav R 2020 Spatiotemporal trends in wildfires across the Western United States (1950–2019) *Remote Sens.* **12** 2959
- 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 CCA4-CEC-2018014* (Merced: University of California)
- Westerling A L 2016 Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring *Phil. Trans. R. Soc. B* **371** 20150178
- Westerling A L, Hidalgo H G, Cayan D R and Swetnam T W 2006 Warming and earlier spring increase Western U.S. forest wildfire activity *Science* **313** 940–3
- Wiedinmyer C, Akagi S K, Emmons L K, Al-Saadi J A, Orlando J J, Soja A J and Soja A J 2011 The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning *Geosci. Model Dev.* **4** 625–41
- Wiedinmyer C and Hurteau M D 2010 Prescribed fire as a means of reducing forest carbon emissions in the Western United States *Environ. Sci. Technol.* **44** 1926–32
- Williams A P, Abatzoglou J T, Gershunov A, Guzman-Morales J, Bishop D A, Balch J K and Lettenmaier D P 2019 Observed Impacts of anthropogenic climate change on wildfire in California *Earth's Future* **7** 892–910
- Xu Q *et al* 2022 Wildfire burn severity and emissions inventory: an example implementation over California *Environ. Res. Lett.* **17** 085008

4 Development of Methods to Predict Respiratory Hospitalization with Wildfire Emissions Inventory Data

Abstract

Wildfire risks are projected to increase with climate change, raising both public and political concerns of health impacts from wildfire emissions. Numerous studies have documented health impact from wildfire PM_{2.5} emissions. However, there are few studies on developing methods that would be practical for predicting health impact from wildfire emissions under climate change scenarios. This study aims to develop methods to connect wildfire emissions to health outcomes that could be applied in future climate change assessment. PM_{2.5} emissions calculated from every large wildfire (> 404 hectares (ha)) were aggregated to air basin level semimonthly and linked respiratory hospitalization at Zip Code Tabulation Areas (ZCTAs) level in California during 1984-2019 using the Generalized Additive Modeling (GAM) statistic approach. The estimated relationship based on 1719 California ZCTAs with 36-years data showed that the presence of at least one hospitalization visit in ZCTAs is not significantly affected by PM_{2.5} from wildfires while there is a significant impact on the numbers of visit. While the wildfire emissions impact is relatively low compared to demographic and socioeconomic factors, this finding is still notable and may be useful in informing resilience actions.

Keywords: PM_{2.5}, GAM, California, ZCTAs, respiratory hospitalization

4.1 Introduction

Wildfire activity has been increasing in the Western US due to human influences through ignitions, climate change, land use, and fire management in the past decades (Abatzoglou and Williams 2016, Bowman *et al* 2011). Wildfires do not just affect the individuals who live in the fire zone. A much larger population is at risk of health impacts from exposure to wildfire emissions that can transport pollutants thousands of miles away (Cascio 2018, Reid *et al* 2016, Reid and Maestas 2019, Matz *et al* 2020). When it comes to health impacts, the pollutant of most concern from wildfires is PM_{2.5} (atmospheric particulate matter that has a diameter of smaller than 2.5 micrometers). PM_{2.5} can remain in the atmosphere for days to weeks and travel far from sources. PM_{2.5} exposure can cause many negative health outcomes. Numerous studies have documented that PM_{2.5} is a major threat to the respiratory and cardiovascular system because it can be inhaled into the deepest recesses of the lung causing respiratory symptoms and can be toxic to the cardiovascular system (Finlay *et al* 2012, Alman *et al* 2016, Heaney *et al* 2022).

Most fire-related health impact studies have quantified wildfire-attributable PM_{2.5} using pollutant concentration data from ground-based air quality monitoring stations (Chen *et al* 2006, Morgan *et al* 2010, Aguilera *et al* 2021), combinations of monitor observations and satellite derived products (Sorensen *et al* 2021), and a combinations of monitor observations, transport model and satellite observation approaches (Johnston *et al* 2012, Reid *et al* 2019 a & b, Cleland *et al* 2020). Different degrees of uncertainty are associated with these methods. There are limited air quality monitoring stations, especially in rural areas where wildfires often occur, and stations might be damaged by fires which can interfere with incomplete or lost data. It's also challenging to separate other pollution sources from wildfires. Atmospheric transport modeling offers the advantage to

isolate the emission sources but could be biased due to input emissions and meteorology data. Satellite imagery may have low quality or missed information when clouds are present. The adoption of a combination of satellite data, ground measurement, and chemical transport models may provide improved exposure estimates, its expense greatly limits the applicability of this approach to future impact assessments. There are also a few other exposure measurements based on the number of fires or data from questionnaires (Künzli *et al* 2006, Marshall *et al* 2007, Caamano-Isorna *et al* 2011). The number of fires may not be representative of actual threats since a large fire may emit much more emissions than a couple of small fires. Exposure measurement from questionnaires may underestimate real exposure (Youssef *et al* 2014).

Smoke transport and dispersion is affected by a complicated interaction of wind, relative humidity, temperature, smoke injection height, and topography (Goodrick *et al* 2013), thus making the projection difficult. Spracklen *et al* (2007 & 2009) and Jaffe *et al* (2008 a & b) attempted to quantify wildfire influences on current concentrations of ozone, PM_{2.5} and organic carbon aerosols, as well as future carbonaceous aerosols in the western United States. Their work applied both an *in-situ* record and transport models to characterize exposures with a focus on California. However, the sparse instrumental record combined with the complexity of California's topography and wind regime made it difficult to characterize exposures with precision, while the expense of running transport models severely limited the number of future scenarios that could be analyzed, limiting the value of the analyses.

A key aim of this research is to explore whether the existing health and wildfire data are sufficient to develop computationally efficient statistical methods that link wildfire emissions to health impacts, to support the simulation of wildfire-related health impacts for very large numbers of future climate and land use scenarios. In this paper we presented our effects for developing statistical models using the Generalized Additive Model (GAM) approach (Hastie and Tibshirani 1990, Wood 2006 & 2017). In our framework, there are three major components: (1) backward forecasting the historical demographic and socioeconomic data at annual time scale for each ZCTA in California as part of the covariate variables for modeling; (2) establishing an algorithm to spatially and temporally estimate the wildfire PM_{2.5} emissions experienced by downwind ZCTAs; (3) developing a two-stage model for estimating the probability of hospitalization and hospitalization numbers. An overview of our framework is presented in Figure 4.1.

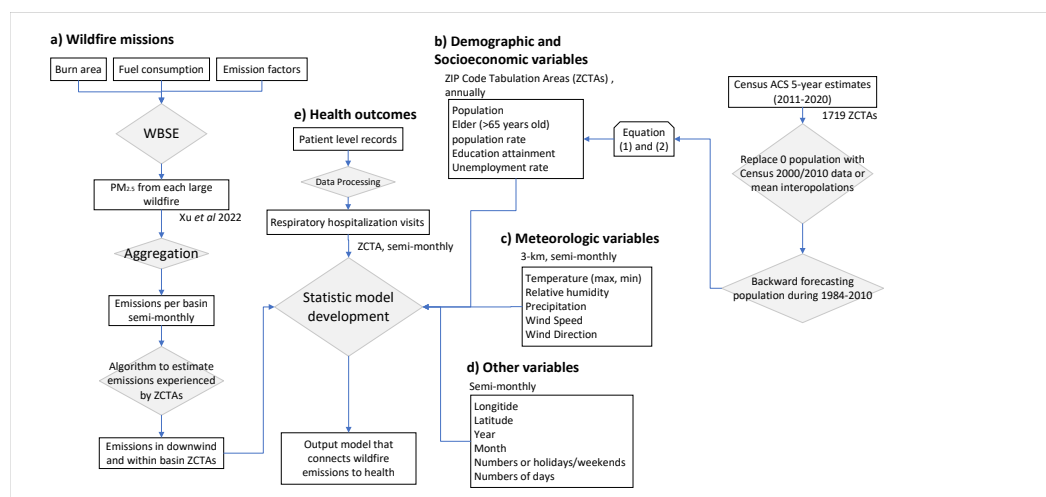


Figure 4. 1 Schematic of our approach to predict health outcomes with wildfire emissions inventory data

4.2 Data and Methods

4.2.1 Spatial and Temporal Domain

We conducted the modeling analysis at ZCTA spatial level on a semimonthly time scale. The patient level health records from California Department of Health Care Access and Information (HCAI) used in this study comes with zip code of residency. ZIP Code Tabulation Areas (ZCTAs), first introduced in the Decennial Census 2000, is a generalized area representations of United States Postal Service ZIP Code service areas. We linked patients' zip codes to ZCTAs to get a summary linked with demographic and socioeconomic status (see Figure 4.2 for an example).

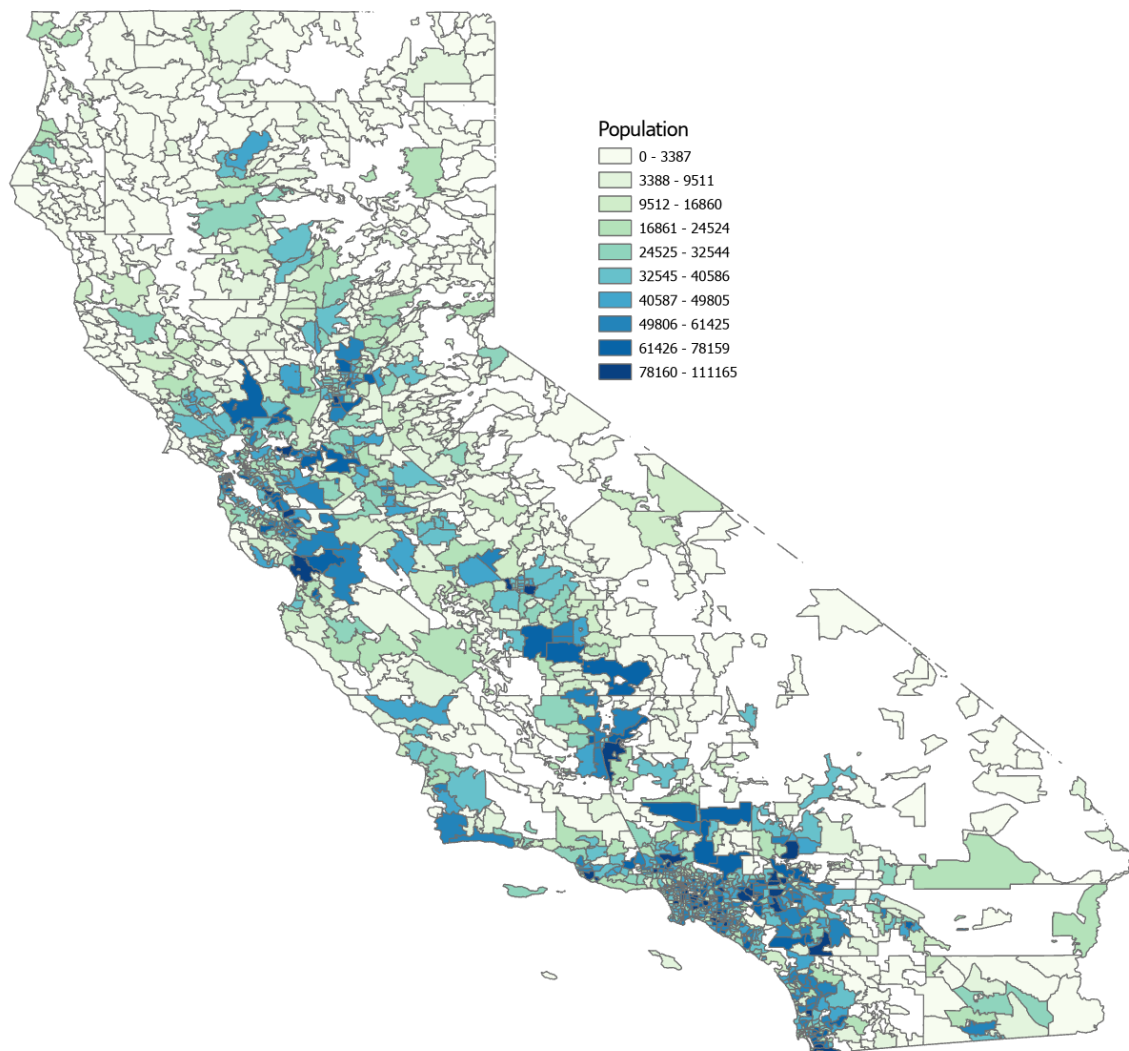


Figure 4. 2 ZIP Code Tabulation Areas (ZCTAs) in California, colored by 2019 population data from American Community Survey (ACS)

We set the time scale to be semimonthly (twice a month, from the first day to the 15th day, and the rest of days). Fire records retrieved from the Monitoring Trends and Burn Severity project (MTBS)

(Eidenshink *et al* 2007) only come with the dates of fire ignition. Large wildfires usually burn from days to months. We set the time scale to semimonthly to ease the problem of assigning emissions.

4.2.1 PM_{2.5} Emissions

PM_{2.5} emissions of each large wildfire (>404 ha) in California from 1984 to 2020 were acquired from Wildfire Burn Severity and Emissions Inventory (WBSE) (Xu *et al* 2022). WBSE estimates emissions as a function of the area burned in different vegetation types and severities, fuel loading, consumption rates under different severities, and emission factor of that vegetation species. WBSE acquired fire records during 1984-2019 from MTBS and fire records for the year 2020 from the California Department of Forestry and Fire Protection's Fire and Resource Assessment Program (CAL FIRE 2022). California is divided geographically into 15 air basins for air pollution management purpose by California Air Resource Board (Figure 4.3). Air basins were determined for areas with similar meteorological and geographic conditions. Air pollutions generally move freely in an air basin. We merged the adjacent Mountain Counties air basin and Lake Tahoe air basin to be one air basin since there were only 1 fire record during the study period in Lake Tahoe air basin. We aggregated PM_{2.5} emissions per air basin by overlying the burn scar map from MTBS to the California air basin GIS layer. For burn scars that cover more than one air basin, we assigned the emissions proportional to the burned area in each basin.

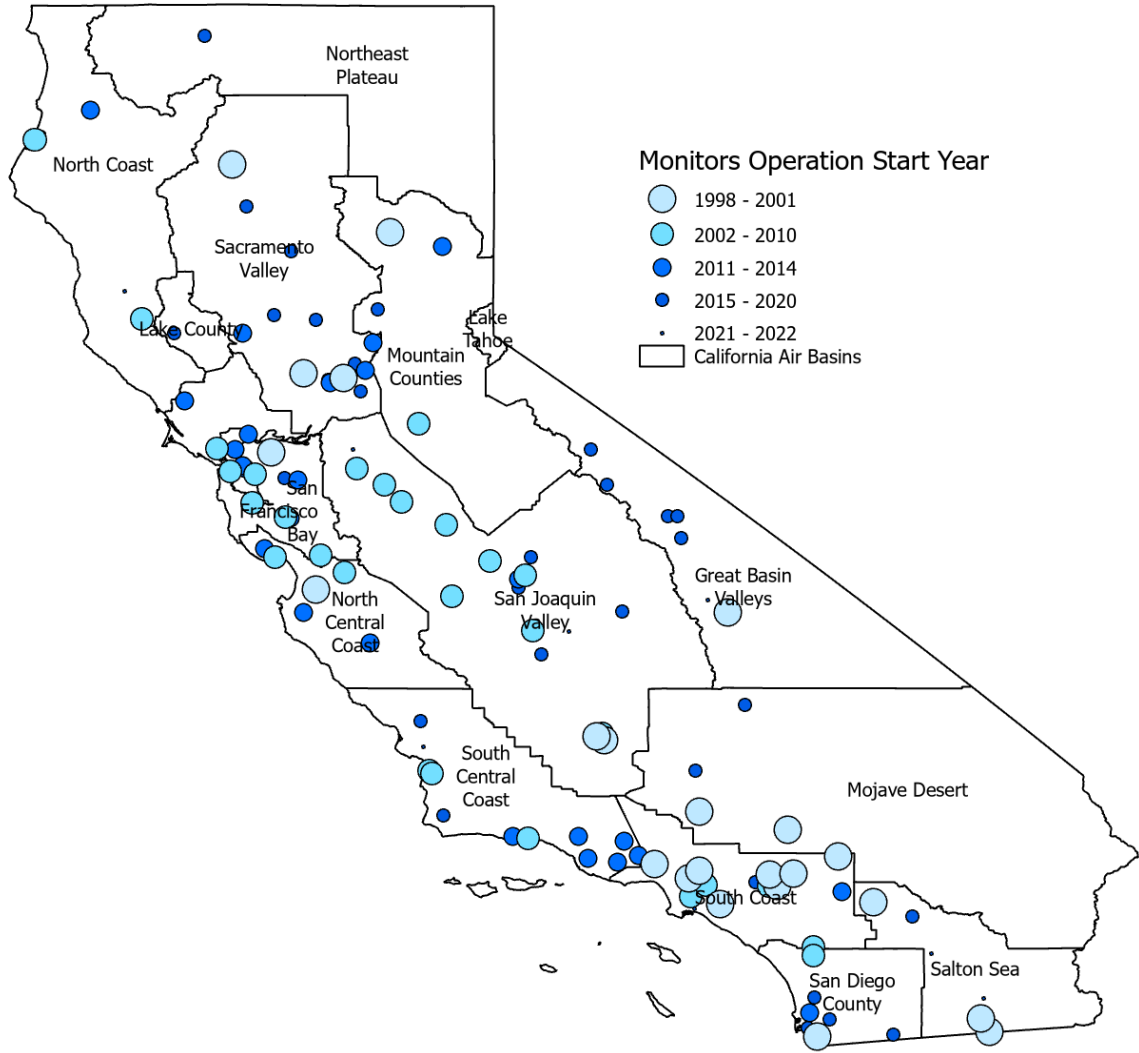


Figure 4. 3 California air basins. Dots show the air monitors, colored by operation start year for monitoring particulate matter data (EPA 2022)

We calculated the distance and angle from the centroid of air basins to the centroid of each ZCTA. We estimated the $PM_{2.5}$ emissions that each ZCTA exposed based on the assumption that prominent wind in the air basin would carry the pollutant in the direction of the maximum speed wind during the semi-month:

$$PM_{j,t} = \sum_{j \text{ not in } i} E_{i,t} \cos(\theta_{ij}) / d_{ij} * s_{i,t} + \sum_{j \text{ in } i} E_{i,t} \quad (1)$$

Where

$PM_{j,t}$ is the $PM_{2.5}$ emissions from basins to ZCTA j ;

$E_{i,t}$ is the total emissions aggregated from wildfires started to burn in basin i during the semi-month t ;

d_{ij} is the distance between basin i to ZCTA j ;

θ_{ij} is the angle between direction from basin i to ZCTA j , and wind direction of the maximum wind speed s_i in basin i ;

$s_{i,t}$ is the maximum wind speed in basin i during semi-month t .

We also considered the lag effects of emissions emitted before the current period, up to 3 months (k up to 6):

$$L_{j,t,k} = \frac{PM_{j,t-k}}{1+k} \quad (2)$$

Where

$L_{j,t,k}$ is the $PM_{2.5}$ emissions from basins to ZCTA j at time t from previous k semi-months.

4.2.2 Health Data

Daily records of the major diagnosis related to the diseases of the respiratory and circulatory system for each inpatient discharged from a California-licensed hospital (hereafter “PDD”) during 1984-2020 and encounters from emergency departments, general acute care hospitals and licensed freestanding ambulatory surgery centers (hereafter “EDAS”) during 2005-2020 were obtained from the California Department of Health Care Access and Information (HCAI). Institutional Review Board approvals were obtained from the Committee for the Protection of Human Subjects and from the Office of Research Compliance and Integrity at University of California, Merced. Principal causes of admission were coded by version 9 of the International Classification of Diseases (ICD-9) for the records during 1984-September 2015, and version 10 of the International Classification of Diseases (ICD-10) since October 2015. Details of ICD categories and codes used in this paper were summarized in Supplementary Table S4.1. We kept records under the following the criteria:

1. unique patient records;
2. the major diagnosis falls into the diseases categories selected for this study;
3. the zip code given by the patient is a California zip code;
4. there were census data for the ZIP Code Tabulation Areas (ZCTAs) which matches with the patient’s zip code;
5. the admission date is during the year of 1984-2020

We further aggregated the daily records to semi-monthly time scale at ZCTAs level for statistical analysis using the ZIP Code to ZCTA crosswalk table based on Census 2010 developed by John Snow, Inc, retrieved from the “zcta” R package. A total of 1719 ZCTAs were included in this study.

4.2.3 Covariate Data

Meteorology data, including semi-monthly 2-m temperature (maximum and mean), relative humidity, mean 10-m wind speed, total precipitation at 3-km resolution of historical simulations, were acquired from ERA5 Weather Research and Forecasting Model (WRF) downscaled re-analysis climate data (Rahimi *et al* 2022), bias corrected following the methodology of Pierce *et al* (2022). The gridded climate data were assigned to each ZCTA based on the latitude and longitude. The predominant wind was determined for each air basin with the maximum speed wind.

Demographic and socioeconomic data at the ZCTAs geographical level were acquired from the data dissemination platform to access data from the U.S. Census Bureau (Data.census.gov). The Census Bureau surveys the entire population every ten years starting 1790. Decennial Census 2000 and 2010 provided data at ZCTAs levels. American Community Survey (ACS) is another demographics survey program conducted by the U.S. Census Bureau that annually publishes results based on a subset of the population. Multiple years of data are pooled together to provide estimates for small population areas. ACS supported Geography ZCTAs in 2007-2011 ACS 5-year estimates and later. We used population by age, education attainment, and employment status data from 2011-2020 ACS 5-Year Estimates Detailed Tables for consistency. There were a total of 1769 ZCTAs fully/partially in California based on ACS 2011-2020 data. There were 48 ZCTAs with a total population estimated as 0 for some or all years. We then acquired Census 2000 and 2010 population data for California from the State of California Department of Finance to replace the ACS data for those ZCTAs when available. For the 1716 ZCTAs with a complete record from ACS, we extrapolated population data during 1984-2010 using backward forecasting constrained to an interval (a, b) to ensure forecasts to be within specified range. We transformed the data using a scaled logit transform:

$$y = \log \left(\frac{x-a}{b-x} \right) \quad (3)$$

Where x is the original scale and y is the transformed data, (a, b) is the constrained prediction internals. The population has been growing at the state level though showing interannual variability within each ZCTA. We therefore set a value to be the 50% of the minimum population number during 2011-2020. We set the b value to be 105% of the maximum historical population.

To reverse the transformation, we used:

$$x = \frac{(b-a)e^y}{1+e^y} + a \quad (4)$$

We removed 11 ZCTAs that have no population data in ACS or the Decennial Census. We used the mean value of each ZCTA for the years with 0 population since we cannot take the logarithm of zero using the equations above. We conducted the same process to calculate the percentage of population over 65 years old, and percent of population 25 years and over with a high school graduate or higher degree.

Unemployment rate, which is measured as the percentage of unemployed civilians in the labor force, remained high around 11% during 2011-2015 and decreased since 2016 to 6.7% in 2020, according to ACS estimates (Supplementary Figure S4.4). These data are not optimistic for backward forecasts. We acquired the monthly unemployment rate of each county during 1984-2020 from the State of California Employment Development Department (EDD, 2022). We assigned the county level unemployment rate to ZCTAs based on the crosswalk from R package “zcta”. The unemployment rate of the county which the major population lived in was selected when a ZCTA overlays more than 1 county. Unemployment rate data during 1984-1989 was only available for Los Angeles County. Thus, we compared the monthly mean unemployment rate of each ZCTA with the state mean rate over the period of 1990-2020 and calculated the monthly unemployment rate based on the ratio during 1984-1989.

Our backward estimate for the years 1984-2019 agreed reasonably well with records from other sources at the state level. We aggregated total population and the percentage of population over 65 years old data interpolated in this study to the state level and compared to estimation by State of California Department of Finance (DOF, 2022) (Supplementary Figure S4.1 & S4.2). We also compared the percent of population 25 years and over with a high school graduate or higher degree

to Decennial Census 1990 and 2000 state level data (Supplementary Figure S 4.3). We compared unemployment rate to ACS estimates (Supplementary Figure S4.4).

Though wildfire emissions play an important role affecting public health, the general air quality has been greatly affected by the total emissions from all the sectors in the state. Thus, we acquired the California annual total GHG emission inventory data (excluded emissions related to "Imports", "Fire", "Out of State") from California Air Resources Board (<https://ww2.arb.ca.gov/ghg-inventory-data>) as a proxy for other air pollutant emissions.

Since data were aggregated at a semi-monthly time scale and span around four decades, we added a time variable with the numbers of days in each semi-month period. Hospital admissions were generally lower during weekends and holidays. We added another time variable with the number of days that was identified as weekends or holidays (Supplementary Table S4.2).

4.2.4 Statistical Analysis

Given the complexity of human behavior of seeking medical care and some people may be more sensitive to weather extremes, we expected some covariates to exhibit nonlinear characteristics. Thus, we chose the Generalized Additive Model (GAM) (Hastie and Tibshirani 1990, Wood 2006 & 2017) for relating the covariates to the health outcomes in two stages.

We first used a logistic GAM for calculating the probability of the presence of respiratory hospitalizations; more specifically, we modeled the presence of respiratory hospitalization from any semi-month t and ZCTA j , as being binomial with probability P_{jt} related to the covariates via the logistic link function:

$$\text{logit}(P_{jt}) = \log\left(\frac{P_{jt}}{1 - P_{jt}}\right) = \Sigma_1^m f_m x_{jt} + \Sigma_1^n \beta_n x_{jt} + \text{offset} \quad (5)$$

Where the first term on the righthand side of the equation represents the nonlinear functions, and the second term is the linear portion of the model. We included the log of the total population in each ZCTA as an offset term to adjust for population size.

Hospitalization numbers are counts making a Gaussian distribution unlikely. Negative binomial and negative binomial distribution are usually selected for modeling count data. We started with Poisson GAM models, but they were over dispersed. We then used a negative binomial generalized additive model, which could be applied with overdispersion data for estimating numbers of visits conditional on the presence of admissions:

$$\log(R_{jt}|R_{jt} > 0) = \Sigma_1^m f_m x_{jt} + \Sigma_1^n \beta_n x_{jt} + \text{offset} \quad (6)$$

Where R_{jt} is the number of hospitalizations in ZCTA j and semi-month t on the condition that there is at least 1 visit; The rest of the terms are the same as the logistic GAM described above.

We applied the backward elimination method for variable selection. This method simply starts with a full model with all variables and then deleted one by one until all the remaining variables have a significant impact on the outcome. Variables were also selected by testing for concurvity (collinearity concept in linear models) and visualization (Wood, 2006). Model selection was based on the Akaike's Information Criterion (Akaike 1974 & 1981) for parsimony. Models were developed on random selected 80% of the data and the remaining 20% were reserved for cross-validation. GAMs were performed using the R package mgcv (Wood 2017). All statistical analyses were performed in R (R Core Team 2021).

4.3 Results and Discussions

4.3.1 Descriptive Statistics

Respiratory hospitalization admissions show an increasing trend during before 2000 and then a decreasing trend (Figure 4.4). Hospitalization numbers vary annually with a strong seasonal cycle which peaks in the winter months (Figure 4.5). Hospitalization admissions were generally lower on weekends or California state holidays (Figure 4.6). The numbers drop sharply in the year 2020 most likely due to people delaying or forgo care when the COVID-19 pandemic declared in March 2020 (Figure 4.4). We thus excluded data from the year 2020 from the modeling analysis.

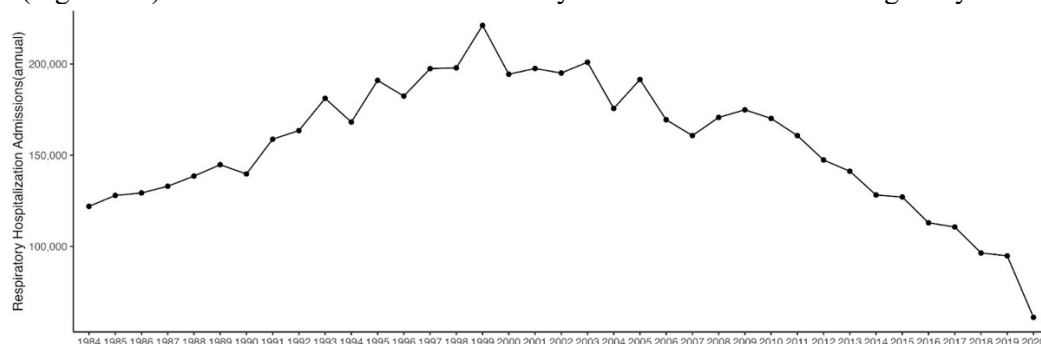


Figure 4. 4 Observed counts of all respiratory hospitalizations during 1984-2020

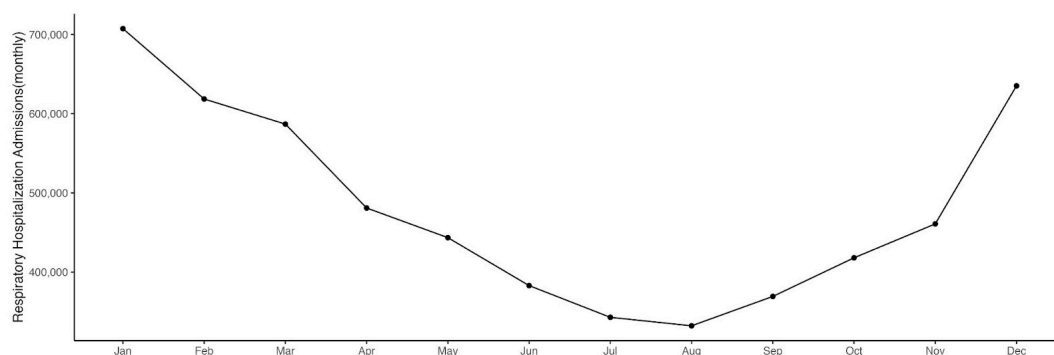


Figure 4. 5 Observed counts of monthly respiratory hospitalizations aggregated over the years 1984-2020

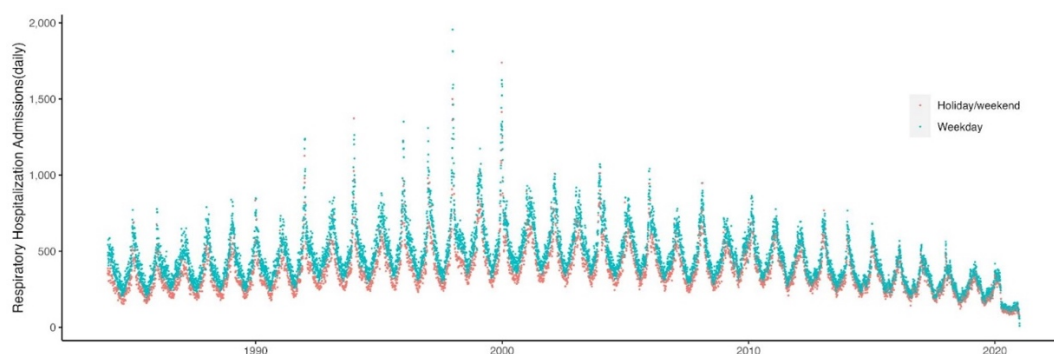


Figure 4. 6 Observed counts of daily respiratory hospitalizations during 1984-2020. Cyan color indicates weekdays and red color indicates days as holidays or weekends

The spatial distribution of fire induced PM_{2.5} emissions were highly heterogeneous in different air basins across California. The Mountain Counties, Sacramento Valley, Northeast Plateau, North Coast air basins had the majority emissions from large wildfires in terms of both frequency and total emissions (Table 4.1). We merged the emissions data from Lake Tahoe air basin to Mountain counties air basin since there were emissions in only one semi-month period during the study period and these two counties were geospatially bounded. When aggregated to the state level by month, August has the highest average emissions.

Table 4. 1 Frequency of having large wildfire induced PM_{2.5} emissions in semi-months and total emissions aggregated over the years of 1984-2019

Air Basin	Emission Frequency	PM _{2.5} Emissions (Metric Tons)
Great Basin Valleys	51	40.84
Lake County	29	116.94
Lake Tahoe	1	0.62
Mojave Desert	85	95.98
Mountain Counties	104	504.04
North Central Coast	63	133.61
North Coast	63	344.46
Northeast Plateau	95	442.67
Sacramento Valley	101	479.12
Salton Sea	24	11.09
San Diego County	87	75.84
San Francisco Bay	35	57.52
San Joaquin Valley	166	233.02
South Central Coast	113	290.73
South Coast	147	149.25

4.3.2 Analysis of Logistic Generalized Additive Model for Predicting Presence of Respiratory Hospitalization

We applied the logistic GAM model using semi-monthly wildfire PM_{2.5} emissions, meteorology, demographic, socioeconomic, location and time variables to model the presence/absence of respiratory hospitalization in ZCTAs in which the wildfire emissions in the current semi-month or the leading periods were assigned. All the 1719 ZCTAs experienced wildfire emissions on average around 50% of the time during the study period of 1984-2019. We first run a model with all the variables including an interaction term with longitude and latitude to account for spatial autocorrelation. We tested nonlinearity on the meteorological and emissions variables. The concavity check indicated that there were large collinearities among the meteorological variables. We inspected the model and removed some variables. We also adjusted the model by removing the interaction term of longitude and latitude since the basic dimension was too low for this smooth

term and increasing the basic dimension didn't improve the model performance. Instead, we used a tensor spline interaction term with to create a proxy for the metrology observed in the spatial domain. All continuous variables and the final selected variables after the concurvity and visualization check could be found in Table 4.2.

Table 4. 2 Continuous variables tested and selected in GAM models

Continuous Variables	presence model (logistic GAM)	number model (negative binomial GAM)
lon,lat		
Elder	TRUE	TRUE
Education	TRUE	TRUE
Unemployment	TRUE	TRUE
GHG	TRUE	
Temperature_Max	TRUE	
Temperature_Mean		
Temperature_Min		TRUE
Precipitation_Sum		
RelativeHumidity_Mean		TRUE
RelativeHumidity_M_max	TRUE	
PM _{2.5} _Current		TRUE
PM _{2.5} _Lag		TRUE

Included terms in the final model formulations are indicated by "True".

The final logistic GAM model (deviance explained = 44.1%) included the interaction of maximum temperature and relative humidity, demographic, socio-economic variables, and time related factor variables. Wildfire emissions variables were not significant nonlinear or linear terms while some meteorologic, demographic and socioeconomic variables were significant (Table 4.1). Descriptive statistics for non-parametric terms were reported in Table 4.3. The intercept of log-odds value -10 in the presence model means that the model predicts a close to 0.004 % baseline chance of a positive outcome when expecting the other variables at their average values. There is in general not a significant impact of numbers of days that are holidays/weekends. There is a large increase in the probability of at least 1 hospitalization visit when there are more days in a semi-month. The relationship between timing of the year (month variable) and hospitalization is reasonably well with the patterns we observed in Figure 4.5 that the probability of hospitalization decreases in summer and increases in winter.

Table 4. 1 Parametric coefficients in GAM models

parametric terms	presence model (logistic GAM)	number model (negative binomial GAM)
Intercept	-10***	-7.15***
GHG	0.01***	
holweekend5	-	0.02***
holweekend6	-	-0.02***
holweekend7	-	0.06**
days14	0.73***	0.18***
days15	0.20**	0.11***
days16	0.29***	0.18***
month2	-0.12**	-0.18***
month3	-0.02	-0.22***
month4	-0.25***	-0.31***
month5	-0.54***	-0.42***
month6	-0.73***	-0.46***
month7	-0.85***	-0.55***
month8	-0.84***	-0.55***
month9	-0.63***	-0.43***
month10	-0.57***	-0.41***
month11	-0.42***	-0.34***
month12	-0.10***	-0.07***

Significant codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The partial effects of the nonlinear variables in the final model were presented in Figure 4.7. They show the effect of smooth terms in the model, which add up to the overall model prediction. The contour plot represented the interactions of maximum temperature and relative humidity. This shows moderate temperature with high relative humidity, and high temperature with low relative humidity led to a higher probability of respiratory hospitalizations.

Model generalization performance on the remaining 20% of data was accessed by accuracy, precision and recall (sensitivity). We set the threshold of 0.5 to determine the model output to binary outcomes (presence/absence). An accuracy of 0.84 indicated that overall, around 84% is classified correctly. The precision 0.85 and recall 0.91 also indicated that the model had a good performance to correctly classify the presence of hospitalization.

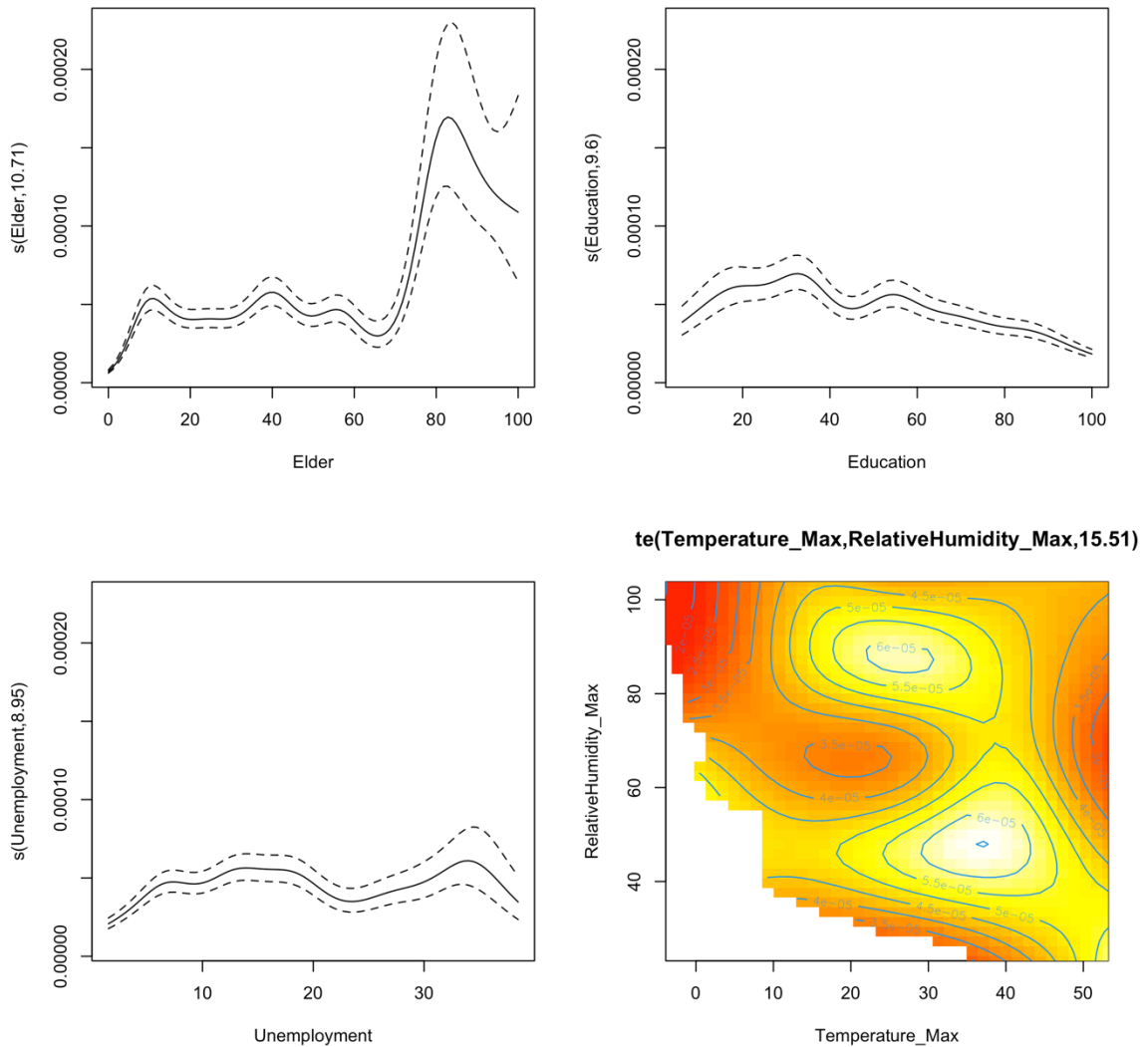


Figure 4. 7 Partial effect plots of logistic GAM output. The output on y axis is the probability of the outcome if all other variables were at their average value. Dotted lines show the confidence intervals

4.3.3 Analysis of Negative Binomial Generalized Additive Model for Predicting Numbers of Respiratory Hospitalization

We conducted the same procedure for the negative binomial generalized additive modeling. The best model (deviance explained = 19.8%) we got included minimum temperature, mean relative humidity, percentage of elder population, education attainment, employment rate, current $\text{PM}_{2.5}$ emissions, lag $\text{PM}_{2.5}$ emissions, and all the time factor variables (days of holidays/weekends, total days in the semi-month, month). The partial effects of the nonlinear variables in negative binomial GAM model were presented in Figure 4.8 and the parametric coefficients were in Table 4.3. Wildfire emissions variables were significant nonlinear terms though the partial effects were much

smaller when compared to metrological, demographic and socioeconomic variables. It is surprising that the impact from current wildfire emissions first increases but decrease with the increase of emissions. This can still be possibly true since people may choose to stay home to avoid the pollutions from outdoor. The lag effect of emissions shows a nonlinear relationship which needs more investigation. Cooler temperature and extreme relative humidity lead to higher numbers of hospitalization. The impact of demographic and socioeconomic factors is in general reasonable. We validated the model on the testing data. The predicted hospitalization numbers showed similarity to the observations and coincident with the one-to-one line (Figure 4.9). The total visits from predicted and observed were also close (results not reported).

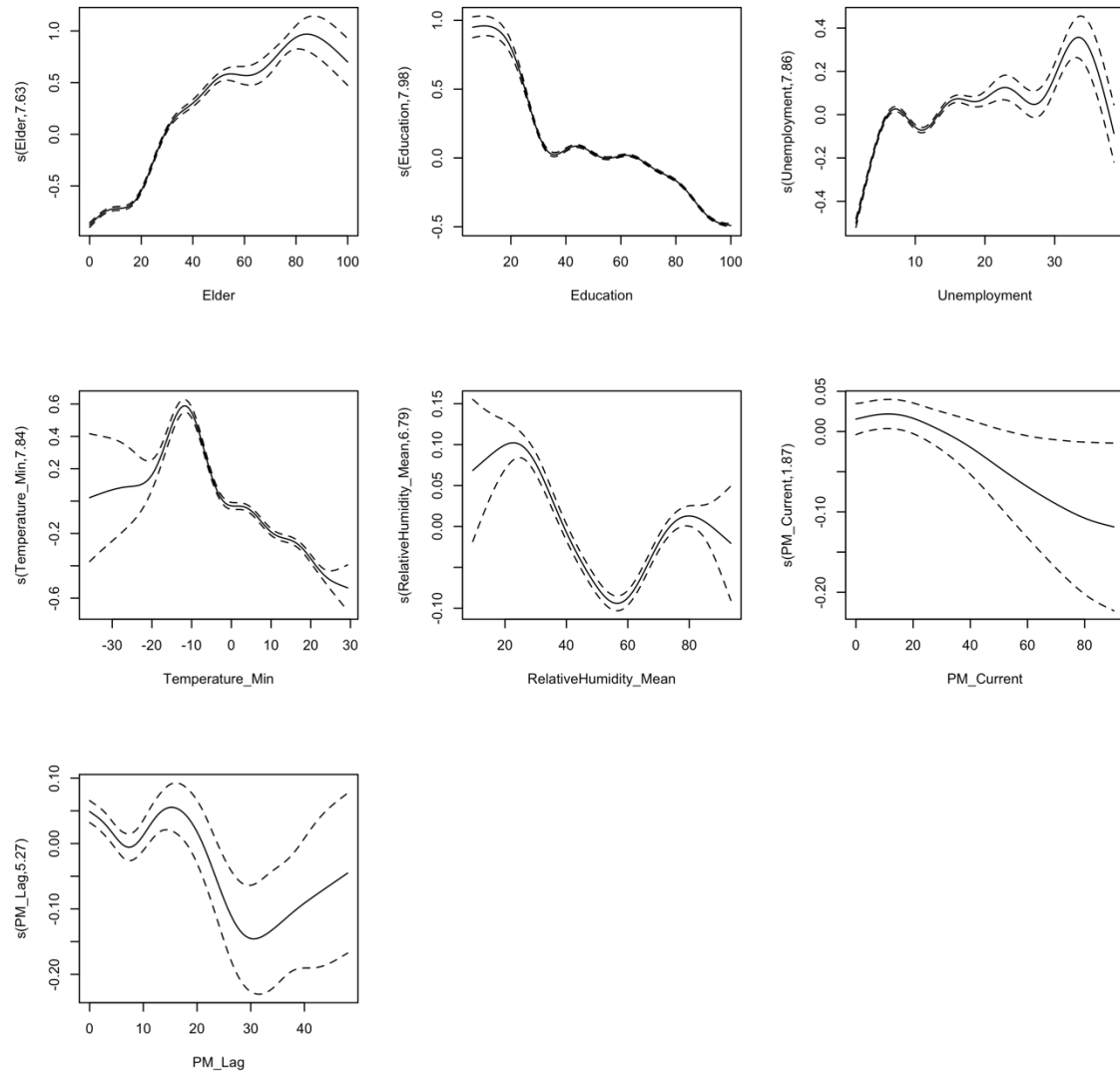


Figure 4. 8 Partial effect plots of negative binomial GAM output. The output on y axis is the log value of hospitalization numbers. Dotted lines show the confidence intervals

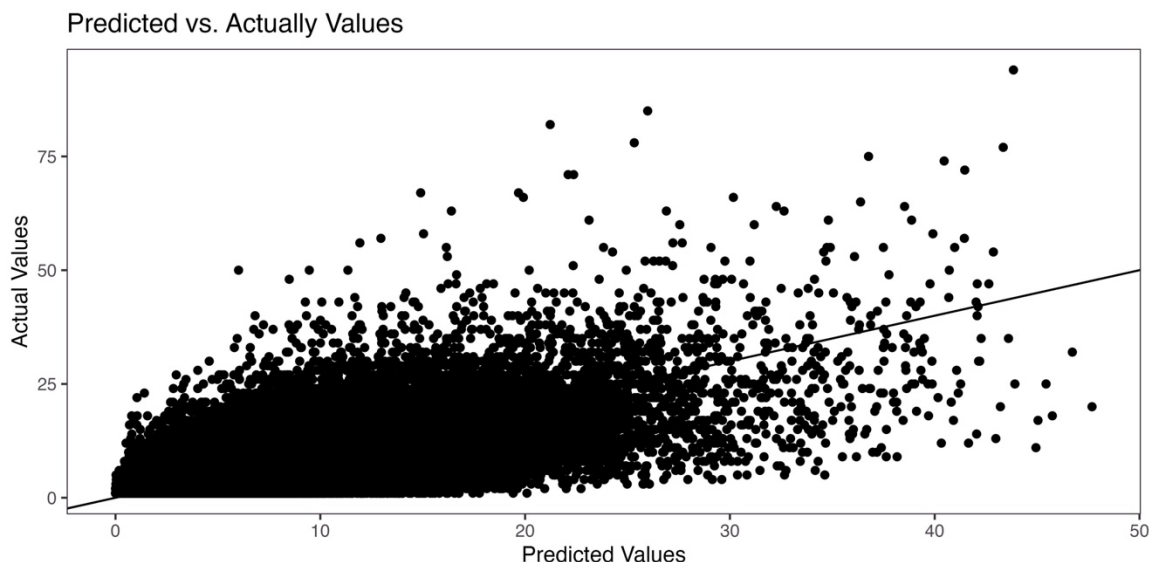


Figure 4. 9 Predicted and observed hospitalization numbers for the testing data

4.3.4 Discussions

We observed that demographic and socioeconomic variables, specifically, the percentage of elder (over 65 years old) population, education attainment and unemployment rates, were the main predictors of respiratory hospitalization in California Zip Code Tabulation Areas. Meteorological factors, including temperature and relative humidity has a smaller impact. The annual state level greenhouse gases variable, which was expected to serve as an indicator of general anthropological air pollutants, has an impact on the presence of hospitalization. The interpolated wildfire $PM_{2.5}$ emissions data has a small influence on predicting hospitalization numbers. There is a drop in hospitalization at the highest levels of $PM_{2.5}$ likely due to the fact that most probably people will stay indoors and protect themselves from the smoke.

There are several limitations of this study:

- (1) The algorithm we proposed to assign wildfire emissions inventory data to Zip Code Tabulation Areas has not been fully field validated. We did a rough comparison of the ZCTA level wildfire emissions data and the $PM_{2.5}$ concentration data from United States Environmental Protection Agency (U.S. EPA) air quality monitors and find weak correlation (Pearson $r = 0.2$, details not reported here). This may indicate that our algorithm didn't represent the wildfire impact on air quality. However, the wildfire- $PM_{2.5}$ contribution was not separated from the monitor data making this validation questionable. Further work is required to justify or improve this algorithm to properly identify the wildfire impacted area and the magnitude of impact on air quality.
- (2) The emissions inventory data by its nature may not be the ideal data for health studies at fine spatial and temporal scale like the ones here in this study. The current emissions inventory data was calculated per fire and only associated with the fire started date. Large wildfires could burn days to months making it hard to assign the emissions to appropriate time. We have taken into the consideration of lag effect. However, this is again hard to validate. Further work could also take the advantage of satellite fire detection data (e.g. VIIRS Active Fire Data) for more precise timing allocation. Those data were only available

for fires in recent years. Daily emissions for fires since 2002 were calculated with the framework developed in Chapter 2 and available for use in the WBSE inventory. We didn't make use this data in the current study due to the consideration of the limited years availability. We also didn't take into consideration of the small size fires which may still have an impact on local communities.

- (3) Uncertainties may also arise from the other input data. For example, the demographic and socioeconomic data, especially the historical period (1984-2010) was backward forecasted by this study. The result seems reasonable while it is also difficult to find out if this represented the real-world situation. We also didn't take into the consideration the constantly change of zip codes and ZCTAs. There was little information publicly available for tracking the change for all the communities in California over three decades. The health data aggregated by HCAI annually based on the discharge year. There is large possibility that some records of patients stayed over years will be released in the future years. So, we may have slightly underestimated health impact based on current data. Uncertainties may also associate with ICD code change from the 9th version to 10th version during this study period.

4.4 Conclusions

The health impact of wildfire PM_{2.5} emissions, as well as environmental conditions and socioeconomic factors, are studied using logistic and negative binomial GAM models. We may conclude that the admission patterns of respiratory diseases in a community are affected by socioeconomic factors, specifically population size and the percentage of elder people. Education attainment and unemployment status could also affect the health care needs. The change of weather conditions including temperature, relative humidity plays a smaller role. Emissions from wildfires started burning in the current semi-month period missions from wildfires had no significant impact of the presence of hospitalization, but they both show a slight significant impact on the numbers of hospitalization given that there is at least one hospitalization. This is something to note and could be valuable for forecasting health impact under future climate change simulations. Next step we will work on fine tuning the models and applying the GAM models on cardiovascular hospitalization and emerging visits. Potential heterogeneity can be analyzed by gender, age, race, pre-existing illnesses, income level, and the distance from the hospital/clinic by stratifying health data by different socio-economic groups. The impact on a couple of non-respiratory and non-cardiovascular ailments can be assessed with a falsification test as these health problems will not be expected to be significantly influenced by wildfire emissions.

References

- Abatzoglou J T and Williams A P 2016 Impact of anthropogenic climate change on wildfire across Western US forests *Proc. Natl Acad. Sci.* **113** 11770–5
- Aguilera R, Corringham T, Gershunov A and Benmarhnia T 2021 Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California *Nat Commun* **12** 1493
- Akaike H 1974 A new look at the statistical model identification *IEEE Trans. Automat. Contr.* **19** 716–23
- Akaike H 1981 Likelihood of a model and information criteria *Journal of Econometrics* **16** 3–14
- Alman B L, Pfister G, Hao H, Stowell J, Hu X, Liu Y and Strickland M J 2016 The association of wildfire smoke with respiratory and cardiovascular emergency department visits in

- Colorado in 2012: a case crossover study *Environmental Health* **15** Online: <http://ehjournal.biomedcentral.com/articles/10.1186/s12940-016-0146-8>
- Bowman D M J S, Balch J, Artaxo P, Bond W J, Cochrane M A, D'Antonio C M, DeFries R, Johnston F H, Keeley J E, Krawchuk M A, Kull C A, Mack M, Moritz M A, Pyne S, Roos C I, Scott A C, Sodhi N S and Swetnam T W 2011 The human dimension of fire regimes on Earth: The human dimension of fire regimes on Earth *Journal of Biogeography* **38** 2223–36
- Caamano-Isorna F, Figueiras A, Sastre I, Montes-Martínez A, Taracido, M and Piñeiro-Lamas M 2011 Respiratory and mental health effects of wildfires: an ecological study in Galician municipalities (north-west Spain) *Environ Health* **10** 48
- California Department of Forestry and Fire Protection Fire Perimeters through 2021 (available at: https://frap.fire.ca.gov/media/ly2jyr4j/fire21_2.zip) (Accessed 14 January 2022)
- Cascio W E 2018 Wildland fire smoke and human health *Science of The Total Environment* **624** 586–95
- Cleland S E, West J J, Jia Y, Reid S, Raffuse S, O'Neill S and Serre M L 2020 Estimating Wildfire Smoke Concentrations during the October 2017 California Fires through BME Space/Time Data Fusion of Observed, Modeled, and Satellite-Derived PM_{2.5} *Environ. Sci. Technol.* **54** 13439–47
- Chen L, Verrall K and Tong S 2006 Air particulate pollution due to bushfires and respiratory hospital admissions in Brisbane, Australia *International Journal of Environmental Health Research* **16** 181–91
- Department of Finance (DOF) 2022 Estimates – Official population estimates of the state, counties and cities produced by the Demographic Research Unit for state planning and budgeting (available at <https://dof.ca.gov/forecasting/demographics/estimates/>) (Accessed 16 June 2022)
- Employment Development Department (EDD) 2022 Local Area Unemployment Statistics (LAUS) (available at <https://data.edd.ca.gov/Labor-Force-and-Unemployment-Rates/Local-Area-Unemployment-Statistics-LAUS-/e6gw-gvii/data>) (Accessed 5 July 2022)
- Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B and Howard S 2007 A project for monitoring trends in burn severity *Fire Ecol.* **3** 3–21
- Finlay S E, Moffat A, Gazzard R, Baker D and Murray V 2012 Health Impacts of Wildfires *PLoS Curr* Online: <https://currents.plos.org/disasters/index.html%3Fp=1881.html>
- Goodrick S L, Achtemeier G L, Larkin N K, Liu Y and Strand T M 2013 Modelling smoke transport from wildland fires: a review *International Journal of Wildland Fire* **22** 83
- Hastie T J and Tibshirani R J 1990 Generalized Additive Models Wiley Online Library.
- Heaney A, Stowell J D, Liu J C, Basu R, Marlier M and Kinney P 2022 Impacts of Fine Particulate Matter From Wildfire Smoke on Respiratory and Cardiovascular Health in California *GeoHealth* **6** Online: <https://onlinelibrary.wiley.com/doi/10.1029/2021GH000578>
- Jaffe D, Chand D, Hafner W, Westerling A and Spracklen D 2008a Influence of Fires on O₃ Concentrations in the Western U.S. *Environmental Science & Technology* **42** 5885–91
- Jaffe D, Hafner W, Chand D, Westerling A and Spracklen D 2008b Interannual Variations in PM_{2.5} due to Wildfires in the Western United States *Environmental Science & Technology* **42** 2812–8
- Johnston F H, Henderson S B, Chen Y, Randerson J T, Marlier M, DeFries R S, Kinney P, Bowman D M J S and Brauer M 2012 Estimated Global Mortality Attributable to Smoke from Landscape Fires *Environmental Health Perspectives* **120** 695–701

- Künzli N, Avol E, Wu J, Gauderman W J, Rappaport E, Millstein J, Bennion J, McConnell R, Gilliland F D, Berhane K, Lurmann F, Winer A and Peters J M 2006 Health Effects of the 2003 Southern California Wildfires on Children *Am J Respir Crit Care Med* **174** 1221–8
- Marshall G N, Schell T L, Elliott M N, Rayburn N R and Jaycox L H 2007 Psychiatric Disorders Among Adults Seeking Emergency Disaster Assistance After a Wildland-Urban Interface Fire *PSYCHIATRIC SERVICES* **58** 6
- Matz C J, Egyed M, Xi G, Racine J, Pavlovic R, Rittmaster R, Henderson S B and Stieb D M 2020 Health impact analysis of PM_{2.5} from wildfire smoke in Canada (2013–2015, 2017–2018) *Science of The Total Environment* **725** 138506
- Morgan G, Sheppard V, Khalaj B, Ayyar A, Lincoln D, Jalaludin B, Beard J, Corbett S and Lumley T 2010 Effects of Bushfire Smoke on Daily Mortality and Hospital Admissions in Sydney, Australia *Epidemiology* **21** 47–55
- Pierce D W, Cayan D R, Goodrich J, Das T and Munévar A 2022 Evaluating Global Climate Models for Hydrological Studies of the Upper Colorado River Basin *J American Water Resour Assoc* **58** 709–34
- Rahimi S, Krantz W, Lin Y, Bass B, Goldenson N, Hall A, Lebo Z J and Norris J 2022 Evaluation of a Reanalysis-Driven Configuration of WRF4 Over the Western United States From 1980 to 2020 *JGR Atmospheres* **127** Online: <https://onlinelibrary.wiley.com/doi/10.1029/2021JD035699>
- R Core Team 2021 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing)
- Reid C E, Brauer M, Johnston F H, Jerrett M, Balmes J R and Elliott C T 2016 Critical Review of Health Impacts of Wildfire Smoke Exposure *Environmental Health Perspectives* **124** 1334–43
- Reid C E, Considine E M, Watson G L, Telesca D, Pfister G G and Jerrett M 2019 Associations between respiratory health and ozone and fine particulate matter during a wildfire event *Environment International* **129** 291–8
- Reid C E and Maestas M M 2019 Wildfire smoke exposure under climate change: impact on respiratory health of affected communities *Current Opinion in Pulmonary Medicine* **25** 179–87
- Sorensen C, House J A, O'Dell K, Brey S J, Ford B, Pierce J R, Fischer E V, Lemery J and Crooks J L 2021 Associations Between Wildfire-Related PM_{2.5} and Intensive Care Unit Admissions in the United States, 2006–2015 *GeoHealth* **5** Online: <https://onlinelibrary.wiley.com/doi/10.1029/2021GH000385>
- Spracklen D V, Logan J A, Mickley L J, Park R J, Yevich R, Westerling A L and Jaffe D A 2007 Wildfires drive interannual variability of organic carbon aerosol in the western U.S. in summer: INTERANNUAL VARIABILITY OF OC AEROSOL *Geophysical Research Letters* **34** Online: <http://doi.wiley.com/10.1029/2007GL030037>
- Spracklen D V, Mickley L J, Logan J A, Hudman R C, Yevich R, Flannigan M D and Westerling A L 2009 Impacts of climate change from 2000 to 2050 on wildfire activity and carbonaceous aerosol concentrations in the western United States *Journal of Geophysical Research* **114** D20301
- United States Environmental Protection Agency 2022 Air Data: Air Quality Data Collected at Outdoor Monitors Across the US (available at <https://www.epa.gov/outdoor-air-quality-data>) (Accessed 11 September 2022)
- Wood S N 2006 Generalized additive models: an introduction with R (New York: Chapman and Hall/CRC)

- Wood S N 2017 *Generalized additive models: an introduction with R* (Boca Raton: CRC Press/Taylor & Francis Group)
- Xu Q *et al* 2022 Wildfire burn severity and emissions inventory: an example implementation over California *Environ. Res. Lett.* **17** 085008
- Youssof H, Liousse C, Roblou L, Assamoi E M, Salonen R O, Maesano C, Banerjee S and Annesi-Maesano I 2014 Quantifying wildfires exposure for investigating health-related effects *Atmospheric Environment* **97** 239–51

Appendix

Table S2.1 Crosswalk of LANDFIRE Existing Vegetation Type (EVT) to five fuel categories: grass = grassland, shrub = shrubland, <5500 = forest below 5500 ft elevation, 5500- 7500 = forest between 5500-7500 ft elevation, >7500 = forest above 7500 ft elevation

Code	Description	Fuel Category
-9999	No Data	ignore
-9999	Nodata	ignore
11	Open Water	ignore
12	North American Glacier and Ice Field	ignore
12	Snow-Ice	ignore
13	Developed-Upland Deciduous Forest	ignore
14	Developed-Upland Evergreen Forest	ignore
15	Developed-Upland Mixed Forest	ignore
16	Developed-Upland Herbaceous	ignore
17	Developed-Upland Shrubland	ignore
21	Developed-Open Space	ignore
22	Developed-Low Intensity	ignore
23	Developed-Medium Intensity	ignore
24	Developed-High Intensity	ignore
25	Developed-Roads	ignore
31	Barren	ignore
32	Quarries-Strip Mines-Gravel Pits-Well and Wind Pads	ignore
32	Quarries-Strip Mines-Gravel Pits	ignore
60	NASS-Orchard	ignore
61	NASS-Vineyard	ignore
63	NASS-Row Crop-Close Grown Crop	ignore
64	NASS-Row Crop	ignore
65	NASS-Close Grown Crop	ignore
66	NASS-Fallow/Idle Cropland	ignore
67	NASS-Pasture and Hayland	ignore
75	Herbaceous Semi-dry	ignore
76	Herbaceous Semi-wet	ignore
78	Recently Disturbed Forest	ignore
81	Agriculture-Pasture and Hay	ignore
82	Agriculture-Cultivated Crops and Irrigated Agriculture	ignore
95	Herbaceous Wetlands	ignore
2001	Inter-Mountain Basins Sparsely Vegetated Systems	grass
2001	Inter-Mountain Basins Active and Stabilized Dune	grass
2001	Inter-Mountain Basins Cliff and Canyon	grass

2001	Inter-Mountain Basins Playa	grass
2001	Inter-Mountain Basins Wash	grass
2001	Columbia Plateau Ash and Tuff Badland	grass
2002	Mediterranean California Sparsely Vegetated Systems	grass
2002	Central California Coast Ranges Cliff and Canyon	grass
2002	Klamath-Siskiyou Cliff and Outcrop	grass
2002	Mediterranean California Alpine Bedrock and Scree	grass
2002	Sierra Nevada Cliff and Canyon	grass
2002	Southern California Coast Ranges Cliff and Canyon	grass
2003	North Pacific Montane Massive Bedrock-Cliff and Talus	ignore
2003	North Pacific Active Volcanic Rock and Cinder Land	ignore
2003	North Pacific Sparsely Vegetated Systems	ignore
2004	North American Warm Desert Sparsely Vegetated Systems	grass
2004	North American Warm Desert Active and Stabilized Dune	grass
2004	North American Warm Desert Badland	grass
2004	North American Warm Desert Bedrock Cliff and Outcrop	grass
2004	North American Warm Desert Pavement	grass
2004	North American Warm Desert Playa	grass
2004	North American Warm Desert Volcanic Rockland	grass
2005	Mediterranean California Northern Coastal Dune	ignore
2005	Mediterranean California Southern Coastal Dune	ignore
2006	Rocky Mountain Alpine/Montane Sparsely Vegetated Systems	grass
2006	Rocky Mountain Cliff Canyon and Massive Bedrock	grass
2008	North Pacific Oak Woodland	<5500 5500- 7500
2011	Rocky Mountain Aspen Forest and Woodland	<5500
2014	Central and Southern California Mixed Evergreen Woodland	<5500
2015	California Coastal Redwood Forest	<5500
2016	Colorado Plateau Pinyon-Juniper Woodland	<5500
2017	Columbia Plateau Western Juniper Woodland and Savanna	<5500
2018	East Cascades Mesic Montane Mixed-Conifer Forest and Woodland	>7500
2019	Great Basin Pinyon-Juniper Woodland	<5500
2020	Inter-Mountain Basins Subalpine Limber-Bristlecone Pine Woodland	>7500 5500- 7500
2021	Klamath-Siskiyou Lower Montane Serpentine Mixed Conifer Woodland	>7500 5500- 7500
2022	Klamath-Siskiyou Upper Montane Serpentine Mixed Conifer Woodland	>7500 5500- 7500
2027	Mediterranean California Dry-Mesic Mixed Conifer Forest and Woodland	>7500 5500- 7500
2028	Mediterranean California Mesic Mixed Conifer Forest and Woodland	>7500 5500- 7500
2029	Mediterranean California Mixed Oak Woodland	<5500

2030	Mediterranean California Lower Montane Black Oak-Conifer Forest and Woodland	<5500
2030	Mediterranean California Lower Montane Conifer Forest and Woodland	<5500
2031	California Montane Jeffrey Pine(-Ponderosa Pine) Woodland	5500-7500
2031	California Montane Jeffrey Pine(-Ponderosa Pine) Woodland	5500-7500
2032	Mediterranean California Red Fir Forest	>7500
2033	Mediterranean California Subalpine Woodland	>7500
2034	Mediterranean California Mesic Serpentine Woodland and Chaparral	shrub
2034	Mediterranean California Mesic Serpentine Woodland	shrub
2034	Mediterranean California Mesic Serpentine Chaparral	shrub
2035	North Pacific Dry Douglas-fir(-Madrone) Forest and Woodland	<5500
2036	North Pacific Hypermaritime Seasonal Sitka Spruce Forest	<5500
2037	North Pacific Maritime Dry-Mesic Douglas-fir-Western Hemlock Forest	<5500
2039	North Pacific Maritime Mesic-Wet Douglas-fir-Western Hemlock Forest	<5500
2041	North Pacific Mountain Hemlock Forest	<5500
2042	North Pacific Mesic Western Hemlock-Silver Fir Forest	<5500
2043	Mediterranean California Mixed Evergreen Forest	<5500
2044	Northern California Mesic Subalpine Woodland	>7500
2045	Northern Rocky Mountain Dry-Mesic Montane Mixed Conifer Forest	5500-7500
2051	Southern Rocky Mountain Dry-Mesic Montane Mixed Conifer Forest and Woodland	5500-7500
2052	Southern Rocky Mountain Mesic Montane Mixed Conifer Forest and Woodland	5500-7500
2053	Northern Rocky Mountain Ponderosa Pine Woodland and Savanna	5500-7500
2054	Southern Rocky Mountain Ponderosa Pine Woodland	7500
2055	Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland	>7500
2056	Rocky Mountain Subalpine Mesic-Wet Spruce-Fir Forest and Woodland	>7500
2057	Rocky Mountain Subalpine-Montane Limber-Bristlecone Pine Woodland	>7500
2058	Sierra Nevada Subalpine Lodgepole Pine Forest and Woodland	>7500
2060	East Cascades Oak-Ponderosa Pine Forest and Woodland	5500-7500
2060	East Cascades Ponderosa Pine Forest and Woodland	7500
2061	Inter-Mountain Basins Aspen-Mixed Conifer Forest and Woodland	ignore
2062	Inter-Mountain Basins Curl-leaf Mountain Mahogany Woodland and Shrubland	shrub
2062	Inter-Mountain Basins Curl-leaf Mountain Mahogany Woodland	shrub
2063	North Pacific Broadleaf Landslide Forest and Shrubland	ignore
2065	Columbia Plateau Scabland Shrubland	shrub
2067	Mediterranean California Alpine Fell-Field	ignore
2068	North Pacific Dry and Mesic Alpine Dwarf-Shrubland or Fell-field or Meadow	ignore

2071	Sierra Nevada Alpine Dwarf-Shrubland	shrub
2073	Baja Semi-Desert Coastal Succulent Scrub	ignore
2079	Great Basin Xeric Mixed Sagebrush Shrubland	shrub
2080	Inter-Mountain Basins Big Sagebrush Shrubland	shrub
2081	Inter-Mountain Basins Mixed Salt Desert Scrub	shrub
2082	Mojave Mid-Elevation Mixed Desert Scrub	shrub
2083	North Pacific Avalanche Chute Shrubland	ignore
2084	North Pacific Montane Shrubland	shrub
2086	Rocky Mountain Lower Montane-Foothill Shrubland	ignore
2087	Sonora-Mojave Creosotebush-White Bursage Desert Scrub	shrub
2088	Sonora-Mojave Mixed Salt Desert Scrub	shrub
2090	Sonoran Granite Outcrop Desert Scrub	ignore
2091	Sonoran Mid-Elevation Desert Scrub	shrub
2092	Southern California Coastal Scrub	shrub
2095	Apacherian-Chihuahuan Mesquite Upland Scrub	ignore
2096	California Maritime Chaparral	shrub
2097	California Mesic Chaparral	shrub
2098	California Montane Woodland and Chaparral	shrub
2099	California Xeric Serpentine Chaparral	shrub
2103	Great Basin Semi-Desert Chaparral	shrub
2104	Mogollon Chaparral	shrub
2105	Northern and Central California Dry-Mesic Chaparral	shrub
2106	Northern Rocky Mountain Montane-Foothill Deciduous Shrubland	shrub
2107	Rocky Mountain Gambel Oak-Mixed Montane Shrubland	shrub
2108	Sonora-Mojave Semi-Desert Chaparral	shrub
2109	Sonoran Paloverde-Mixed Cacti Desert Scrub	shrub
2110	Southern California Dry-Mesic Chaparral	shrub
2112	California Central Valley Mixed Oak Savanna	<5500
2113	California Coastal Live Oak Woodland and Savanna	<5500
2114	California Lower Montane Blue Oak-Foothill Pine Woodland and Savanna	<5500
2114	California Lower Montane Foothill Pine Woodland and Savanna	<5500
2117	Southern Rocky Mountain Ponderosa Pine Savanna	5500-7500
2118	Southern California Oak Woodland and Savanna	<5500
2123	Columbia Plateau Steppe and Grassland	grass
2124	Columbia Plateau Low Sagebrush Steppe	shrub
2125	Inter-Mountain Basins Big Sagebrush Steppe	shrub
2126	Inter-Mountain Basins Montane Sagebrush Steppe	shrub
2127	Inter-Mountain Basins Semi-Desert Shrub-Steppe	shrub

2128	Northern California Coastal Scrub	shrub
2129	California Central Valley and Southern Coastal Grassland	grass
2130	California Mesic Serpentine Grassland	grass
2131	California Northern Coastal Grassland	grass
2135	Inter-Mountain Basins Semi-Desert Grassland	grass
2136	Mediterranean California Alpine Dry Tundra	shrub
2137	Mediterranean California Subalpine Meadow	grass
2138	North Pacific Montane Grassland	grass
2139	Northern Rocky Mountain Lower Montane-Foothill-Valley Grassland	grass
2142	Columbia Basin Palouse Prairie	grass
2144	Rocky Mountain Alpine Turf	ignore
2145	Rocky Mountain Subalpine-Montane Mesic Meadow	grass
2146	Southern Rocky Mountain Montane-Subalpine Grassland	grass
2151	California Central Valley Riparian Woodland and Shrubland	shrub
2151	California Central Valley Riparian Woodland	shrub
2151	California Central Valley Riparian Shrubland	shrub
2151	California Central Valley Riparian Forest and Woodland	shrub
2151	California Central Valley Riparian Herbaceous	shrub
2152	California Montane Riparian Systems	<5500
2152	Mediterranean California Foothill and Lower Montane Riparian Woodland	<5500
2152	Mediterranean California Serpentine Foothill and Lower Montane Riparian Woodland and Seep	<5500
2152	Mediterranean California Foothill and Lower Montane Riparian Shrubland	<5500
2152	Mediterranean California Serpentine Foothill and Lower Montane Riparian Shrubland and Seep	<5500
2153	Inter-Mountain Basins Greasewood Flat	shrub 5500-
2154	Inter-Mountain Basins Montane Riparian Systems	7500 5500-
2154	Great Basin Foothill and Lower Montane Riparian Woodland	7500 5500-
2154	Columbia Plateau Silver Sagebrush Seasonally Flooded Shrub-Steppe	7500 5500-
2154	Great Basin Foothill and Lower Montane Riparian Shrubland	7500 5500-
2154	Great Basin Foothill and Lower Montane Riparian Herbaceous	7500 5500-
2154	Inter-Mountain Basins Montane Riparian Forest and Woodland	7500 5500-
2154	Inter-Mountain Basins Montane Riparian Shrubland	7500
2155	North American Warm Desert Riparian Systems	shrub
2155	North American Warm Desert Riparian Woodland	shrub
2155	North American Warm Desert Lower Montane Riparian Woodland	shrub
2155	North American Warm Desert Cienega	shrub

2155	North American Warm Desert Riparian Mesquite Bosque Woodland	shrub
2155	North American Warm Desert Wash Woodland	shrub
2155	North American Warm Desert Riparian Herbaceous	shrub
2155	North American Warm Desert Riparian Shrubland	shrub
2155	North American Warm Desert Lower Montane Riparian Shrubland	shrub
2155	North American Warm Desert Riparian Mesquite Bosque Shrubland	shrub
2155	North American Warm Desert Wash Shrubland	shrub
2155	North American Warm Desert Riparian Forest and Woodland	shrub
2156	North Pacific Lowland Riparian Forest	<5500
2156	North Pacific Lowland Riparian Shrubland	shrub
2156	North Pacific Lowland Riparian Forest and Shrubland	shrub
2157	North Pacific Swamp Systems	ignore
2158	North Pacific Montane Riparian Woodland	shrub
2158	North Pacific Montane Riparian Shrubland	shrub
2158	North Pacific Montane Riparian Woodland and Shrubland	shrub
2159	Rocky Mountain Montane Riparian Systems	>7500
2159	Rocky Mountain Montane Riparian Forest and Woodland	5500-7500
2159	Rocky Mountain Montane Riparian Shrubland	shrub
2160	Rocky Mountain Subalpine-Montane Riparian Shrubland	shrub
2160	Rocky Mountain Subalpine-Montane Riparian Woodland	shrub
2160	Rocky Mountain Subalpine/Upper Montane Riparian Systems	>7500
2160	Rocky Mountain Subalpine/Upper Montane Riparian Forest and Woodland	shrub
2160	Rocky Mountain Subalpine/Upper Montane Riparian Shrubland	shrub
2163	Pacific Coastal Marsh Systems	ignore
2164	Rocky Mountain Wetland-Herbaceous	ignore
2165	Northern Rocky Mountain Foothill Conifer Wooded Steppe	ignore
2167	Rocky Mountain Poor-Site Lodgepole Pine Forest	>7500
2170	Klamath-Siskiyou Xeromorphic Serpentine Savanna and Chaparral	shrub
2170	Klamath-Siskiyou Xeromorphic Serpentine Savanna	shrub
2170	Klamath-Siskiyou Xeromorphic Serpentine Chaparral	shrub
2171	North Pacific Alpine and Subalpine Dry Grassland	grass 5500-7500
2172	Sierran-Intermontane Desert Western White Pine-White Fir Woodland	7500
2173	North Pacific Wooded Volcanic Flowage	ignore
2174	North Pacific Dry-Mesic Silver Fir-Western Hemlock-Douglas-fir Forest	<5500
2176	Rocky Mountain Alpine-Montane Wet Meadow	ignore
2177	California Coastal Closed-Cone Conifer Forest and Woodland	<5500
2180	Introduced Riparian Vegetation	shrub
2180	Introduced Riparian Shrubland	shrub

2181	Introduced Upland Vegetation-Annual Grassland	grass
2182	Introduced Upland Vegetation-Perennial Grassland and Forbland	grass
2183	Introduced Upland Vegetation-Annual and Biennial Forbland	grass
2184	California Annual Grassland	grass
2190	Recently Burned-Shrub Cover	ignore
2191	Recently Logged-Herb and Grass Cover	ignore
2192	Recently Logged-Shrub Cover	ignore
2193	Recently Logged-Tree Cover	ignore
2195	Recently Burned-Herb and Grass Cover	ignore
2197	Recently Burned-Tree Cover	ignore
2198	Recently Disturbed Other-Herb and Grass Cover	ignore
2199	Recently Disturbed Other-Shrub Cover	ignore
2200	<i>Pseudotsuga menziesii</i> - <i>Quercus garryana</i> Woodland Alliance	<5500
2200	Coastal Douglas-fir Woodland	<5500
2201	<i>Quercus garryana</i> Woodland Alliance	<5500
2202	<i>Juniperus occidentalis</i> Wooded Herbaceous Alliance	5500-7500
2203	<i>Juniperus occidentalis</i> Woodland Alliance	5500-7500
2204	Recently Disturbed Other-Tree Cover	ignore
2206	<i>Pseudotsuga menziesii</i> Giant Forest Alliance	<5500
2208	<i>Abies concolor</i> Forest Alliance	5500-7500
2210	<i>Coleogyne ramosissima</i> Shrubland Alliance	shrub
2211	<i>Grayia spinosa</i> Shrubland Alliance	shrub
2218	North American Warm Desert Sparsely Vegetated Systems II	ignore
2219	Inter-Mountain Basins Sparsely Vegetated Systems II	ignore
2220	<i>Artemisia tridentata</i> ssp. <i>vaseyana</i> Shrubland Alliance	shrub
2221	Mediterranean California Sparsely Vegetated Systems II	ignore
2222	Rocky Mountain Alpine/Montane Sparsely Vegetated Systems II	ignore
2223	Sonoran Desert Sparsely Vegetated	ignore
2224	Inter-Mountain Basins Alkaline Closed Depression	ignore
2225	North American Arid West Emergent Marsh	ignore
2227	<i>Pseudotsuga menziesii</i> Forest Alliance	5500-7500
2227	Dry-mesic Montane Douglas-fir Forest	5500-7500
2229	<i>Pinus albicaulis</i> Woodland Alliance	>7500
2230	<i>Pinus sabiniana</i> Woodland Alliance	<5500
2231	<i>Sequoiadendron giganteum</i> Forest Alliance	5500-7500
2236	Temperate Pacific Subalpine-Montane Wet Meadow	ignore

2260	Mediterranean California Lower Montane Black Oak Forest and Woodland	<5500
2261	Mediterranean California Lower Montane Black Oak-Conifer Forest and Woodland	ignore
2261	Mediterranean California Lower Montane Black Oak - Conifer Forest and Woodland	<5500
2262	East Cascades Oak Forest and Woodland	<5500
2263	East Cascades Oak - Ponderosa Pine Forest and Woodland	<5500
2264	California Lower Montane Blue Oak Woodland and Savanna	<5500
2264	California Lower Montane Blue Oak Forest and Woodland	<5500
2265	California Lower Montane Blue Oak-Foothill Pine Woodland and Savanna	<5500
2265	California Lower Montane Blue Oak-Foothill Pine Forest and Woodland	<5500
2266	Oregon White Oak Woodland	<5500
2267	Douglas-fir - Oregon White Oak Woodland	<5500
2270	California Ruderal Grassland and Meadow	ignore
2270	California Ruderal Scrub	ignore
2271	Californian Ruderal Forest	ignore
2272	Great Basin & Intermountain Introduced Annual and Biennial Forbland	ignore
2273	Great Basin & Intermountain Introduced Annual Grassland	ignore
2274	Great Basin & Intermountain Introduced Perennial Grassland and Forbland	ignore
2275	North American Warm Desert Ruderal & Planted Scrub	ignore
2275	North American Warm Desert Ruderal & Planted Grassland	ignore
2278	Interior West Ruderal Riparian Forest	shrub
2278	Interior West Ruderal Riparian Scrub	shrub
2279	Interior Western North American Temperate Ruderal Shrubland	ignore
2279	Interior Western North American Temperate Ruderal Grassland	ignore
2280	Western North American Ruderal Wet Shrubland	ignore
2280	Western North American Ruderal Wet Meadow & Marsh	ignore
2282	Great Basin & Intermountain Ruderal Shrubland	ignore
2662	Temperate Pacific Freshwater Emergent Marsh	ignore
2668	Temperate Pacific Tidal Salt and Brackish Marsh	ignore
2902	Developed-Low Intensity	ignore
2903	Developed-Medium Intensity	ignore
2904	Developed-High Intensity	ignore
2905	Developed-Roads	ignore
2913	Western Cool Temperate Urban Deciduous Forest	ignore
2913	Western Warm Temperate Urban Deciduous Forest	ignore
2914	Western Cool Temperate Urban Evergreen Forest	ignore
2914	Western Warm Temperate Urban Evergreen Forest	ignore
2915	Western Cool Temperate Urban Mixed Forest	ignore
2915	Western Warm Temperate Urban Mixed Forest	ignore

2916	Western Cool Temperate Urban Herbaceous	ignore
2916	Western Warm Temperate Urban Herbaceous	ignore
2917	Western Cool Temperate Urban Shrubland	ignore
2917	Western Warm Temperate Urban Shrubland	ignore
2923	Western Cool Temperate Developed Ruderal Deciduous Forest	ignore
2923	Western Warm Temperate Developed Ruderal Deciduous Forest	ignore
2924	Western Cool Temperate Developed Ruderal Evergreen Forest	ignore
2924	Western Warm Temperate Developed Ruderal Evergreen Forest	ignore
2925	Western Cool Temperate Developed Ruderal Mixed Forest	ignore
2925	Western Warm Temperate Developed Ruderal Mixed Forest	ignore
2926	Western Cool Temperate Developed Ruderal Shrubland	ignore
2926	Western Warm Temperate Developed Ruderal Shrubland	ignore
2927	Western Cool Temperate Developed Ruderal Grassland	ignore
2927	Western Warm Temperate Developed Ruderal Grassland	ignore
2933	Western Cool Temperate Undeveloped Ruderal Deciduous Forest	ignore
2933	Western Warm Temperate Undeveloped Ruderal Deciduous Forest	ignore
2934	Western Cool Temperate Undeveloped Ruderal Evergreen Forest	ignore
2934	Western Warm Temperate Undeveloped Ruderal Evergreen Forest	ignore
2935	Western Cool Temperate Undeveloped Ruderal Mixed Forest	ignore
2935	Western Warm Temperate Undeveloped Ruderal Mixed Forest	ignore
2936	Western Cool Temperate Undeveloped Ruderal Shrubland	ignore
2936	Western Warm Temperate Undeveloped Ruderal Shrubland	ignore
2937	Western Cool Temperate Undeveloped Ruderal Grassland	ignore
2937	Western Warm Temperate Undeveloped Ruderal Grassland	ignore
2943	Western Cool Temperate Developed Ruderal Deciduous Forested Wetland	ignore
2943	Western Warm Temperate Developed Ruderal Deciduous Forested Wetland	ignore
2945	Western Cool Temperate Developed Ruderal Mixed Forested Wetland	ignore
2945	Western Warm Temperate Developed Ruderal Mixed Forested Wetland	ignore
2946	Western Cool Temperate Developed Ruderal Shrub Wetland	ignore
2946	Western Warm Temperate Developed Ruderal Shrub Wetland	ignore
2947	Western Cool Temperate Developed Ruderal Herbaceous Wetland	ignore
2947	Western Warm Temperate Developed Ruderal Herbaceous Wetland	ignore
2960	Western Cool Temperate Orchard	ignore
2960	Western Warm Temperate Orchard	ignore
2961	Western Cool Temperate Vineyard	ignore
2961	Western Warm Temperate Vineyard	ignore
2962	Western Cool Temperate Bush fruit and berries	ignore
2962	Western Warm Temperate Bush fruit and berries	ignore
2963	Western Cool Temperate Row Crop-Close Grown Crop	ignore

2963	Western Warm Temperate Row Crop-Close Grown Crop	ignore
2963	Western Cool Temperate Row Crop - Close Grown Crop	ignore
2963	Western Warm Temperate Row Crop - Close Grown Crop	ignore
2964	Western Cool Temperate Row Crop	ignore
2964	Western Warm Temperate Row Crop	ignore
2965	Western Cool Temperate Close Grown Crop	ignore
2965	Western Warm Temperate Close Grown Crop	ignore
2966	Western Cool Temperate Fallow/Idle Cropland	ignore
2966	Western Warm Temperate Fallow/Idle Cropland	ignore
2967	Western Cool Temperate Pasture and Hayland	ignore
2967	Western Warm Temperate Pasture and Hayland	ignore
2968	Western Cool Temperate Wheat	ignore
2968	Western Warm Temperate Wheat	ignore
2969	Western Warm Temperate Aquaculture	ignore
2969	Western Cool Temperate Aquaculture	ignore
9999	Nodata	ignore

Table S2.2 Fuel loadings and consumption for the five vegetation classes under low, moderate, and high burn severity

		Fire Severity		
		Low	Moderate	High
Grassland	average of chamise chaparral and ceanothus chaparral			
	pre-burn fuel load (t/ha)	4.0	4.0	4.0
	fuel consumed (t/ha)	4.0	4.0	4.0
	% reduced	100	100	100
Shrub	pre-burn fuel load (t/ha)	59	59	59
	fuel consumed (t/ha)	47	47	47
	% reduced	80	80	80
Forest<5500'	weighted average of redwood(1), DougFir-Tanoak-Madrone(2), BlueOak-GrayPine(2)			
	pre-burn fuel load (t/ha)	116	116	116
	fuel consumed (t/ha)	69	82	89
	% reduced	60	71	77
Forest 5500'-7500'	average of two sierran mixed-con models, ponderosa pine, and jeffrey pine			
	pre-burn fuel load (t/ha)	137	137	137
	fuel consumed (t/ha)	75	92	101
	% reduced	55	67	74
Forest>7500'	average of red fir, lodgepole pine, CA mixed subalpine			
	pre-burn fuel load (t/ha)	96	96	96
	fuel consumed (t/ha)	45	56	62
	% reduced	47	59	65

Table S2.3 Pixel level PM_{2.5} emission uncertainty for 5 general vegetation classes due to burn severity classification error (unit: g/30-m pixel). Negative values indicate underestimation and positive values indicate overestimation

(1) Grass					
Actual Predicted	Unchanged	Low	Moderate	High	Grass burn
Unchanged	0	-2.592			
Low	2.592	0	0	0	0
Moderate					
High					
Grass burn					
(2) Shrub					
Actual Predicted	Unchanged	Low	Moderate	High	Grass burn
Unchanged	0	-33.417			-2.592
Low	33.417	0	0	0	30.825
Moderate					
High					
Grass burn					
(3) Forest < 5500'					
Actual Predicted	Unchanged	Low	Moderate	High	Grass burn
Unchanged	0	-65.826	-78.228	-84.906	-2.592
Low	65.826	0	-12.402	-19.08	63.234
Moderate	78.228	12.402	0	-6.678	75.636
High	84.906	19.08	6.678	0	82.314
Grass burn	2.592	-63.234	-75.636	-82.314	0
(4) Forest 5500-7500'					
Actual Predicted	Unchanged	Low	Moderate	High	Grass burn
Unchanged	0	-71.55	-87.768	-96.354	-2.592
Low	71.55	0	-16.218	-24.804	68.958
Moderate	87.768	16.218	0	-8.586	85.176
High	96.354	24.804	8.586	0	93.762
Grass burn	2.592	-68.958	-13.626	-93.762	0
(5) Forest > 7500'					
Actual Predicted	Unchanged	Low	Moderate	High	Grass burn
Unchanged	0	-42.93	-53.424	-59.148	-2.592
Low	42.93	0	-10.494	-16.218	40.338
Moderate	53.424	10.494	0	-5.724	50.832
High	59.148	16.218	5.724	0	56.556
Grass burn	2.592	-40.338	-50.832	-56.556	0

Table S3.1 Annual Emissions (metric tons yr-1) from Large Wildfires in California during 1984-2020

Year	CO ₂	CO	CH ₄	NMOC	SO ₂	NH ₃	NO	NO ₂	NO _x	PM _{2.5}	OC	BC
1984	4.876	0.27	0.009	0.073	0.003	0.004	0.004	0.003	0.005	0.027	0.022	0.003
1985	19.82	1.051	0.038	0.293	0.011	0.017	0.013	0.012	0.019	0.117	0.105	0.009
1986	2.625	0.145	0.005	0.039	0.001	0.002	0.002	0.002	0.003	0.015	0.012	0.001
1987	33.886	1.713	0.071	0.496	0.019	0.026	0.02	0.02	0.028	0.209	0.214	0.012
1988	8.339	0.433	0.017	0.123	0.005	0.007	0.005	0.005	0.008	0.05	0.048	0.003
1989	4.757	0.256	0.009	0.07	0.003	0.004	0.003	0.003	0.005	0.028	0.024	0.002
1990	12.734	0.664	0.025	0.187	0.007	0.01	0.008	0.008	0.012	0.076	0.072	0.005
1991	0.666	0.034	0.001	0.01	0	0.001	0	0	0.001	0.004	0.004	0
1992	8.336	0.422	0.017	0.122	0.005	0.006	0.005	0.005	0.007	0.051	0.055	0.003
1993	7.787	0.447	0.013	0.117	0.004	0.007	0.006	0.005	0.009	0.042	0.028	0.005
1994	11.74	0.632	0.022	0.174	0.006	0.01	0.008	0.007	0.012	0.068	0.059	0.006
1995	3.838	0.224	0.006	0.058	0.002	0.004	0.003	0.002	0.005	0.02	0.013	0.003
1996	17.197	0.944	0.031	0.257	0.009	0.015	0.012	0.011	0.018	0.097	0.079	0.009
1997	5.592	0.303	0.01	0.083	0.003	0.005	0.004	0.003	0.006	0.032	0.027	0.003
1998	2.776	0.161	0.004	0.042	0.001	0.003	0.002	0.002	0.003	0.015	0.009	0.002
1999	29.036	1.535	0.056	0.428	0.016	0.024	0.019	0.017	0.027	0.172	0.159	0.013
2000	8.547	0.447	0.017	0.126	0.005	0.007	0.005	0.005	0.008	0.051	0.047	0.004
2001	12.02	0.625	0.024	0.177	0.007	0.01	0.008	0.007	0.011	0.072	0.07	0.005
2002	9.635	0.476	0.021	0.144	0.006	0.007	0.006	0.007	0.009	0.06	0.06	0.003
2003	7.848	0.368	0.017	0.122	0.007	0.005	0.007	0.007	0.009	0.048	0.047	0.002
2004	3.358	0.167	0.007	0.051	0.002	0.002	0.002	0.002	0.003	0.02	0.022	0.001
2005	1.858	0.09	0.004	0.029	0.002	0.001	0.002	0.002	0.002	0.011	0.01	0.001
2006	12.444	0.599	0.027	0.186	0.008	0.009	0.008	0.009	0.011	0.078	0.084	0.003
2007	12.422	0.609	0.026	0.19	0.009	0.009	0.01	0.01	0.013	0.075	0.074	0.004
2008	37.245	1.827	0.081	0.548	0.022	0.027	0.022	0.024	0.03	0.235	0.251	0.011
2009	9.905	0.478	0.022	0.147	0.006	0.007	0.006	0.007	0.008	0.063	0.069	0.003
2010	1.449	0.073	0.003	0.022	0.001	0.001	0.001	0.001	0.001	0.009	0.008	0.001
2011	1.842	0.089	0.004	0.028	0.001	0.001	0.001	0.001	0.002	0.011	0.012	0.001
2012	15.033	0.739	0.033	0.223	0.009	0.011	0.009	0.01	0.013	0.094	0.103	0.005
2013	18.596	0.922	0.04	0.273	0.011	0.014	0.011	0.012	0.015	0.116	0.121	0.006

2014	16.411	0.82	0.035	0.241	0.009	0.012	0.01	0.01	0.013	0.102	0.109	0.005
2015	14.893	0.756	0.031	0.218	0.008	0.011	0.009	0.009	0.013	0.092	0.093	0.005
2016	9.497	0.515	0.018	0.141	0.005	0.008	0.006	0.006	0.01	0.055	0.046	0.005
2017	46.118	2.383	0.093	0.678	0.025	0.037	0.029	0.028	0.041	0.279	0.276	0.018
2018	59.769	3.035	0.125	0.876	0.033	0.046	0.036	0.036	0.05	0.368	0.382	0.021
2019	8.011	0.417	0.016	0.118	0.004	0.006	0.005	0.005	0.007	0.048	0.046	0.003
2020	163.435	8.273	0.343	2.394	0.090	0.126	0.097	0.098	0.136	1.009	1.048	0.057

Table S3.2 Summary of Mann–Kendall test for annual wildfires and emissions in California over 1984–2020 at 5% significance level

Annual Trend analysis	Fire numbers	Area burned	Size of annual large fire	PM _{2.5} emissions	PM _{2.5} emissions from forest
<i>S</i>	0.229	0.007	0.002	1.982	2.008
<i>Z</i>	0.654	3.126	3.178	2.053	2.498
Kendall's tau	0.077	0.360	0.366	0.237	0.228
<i>p</i>	0.513	0.002	0.001	0.040	0.012
Alpha α	0.05	0.05	0.05	0.05	0.05
Significance	Insignificant increasing trend	Significant increasing trend	Significant increasing trend	Significant increasing trend	Significant increasing trend

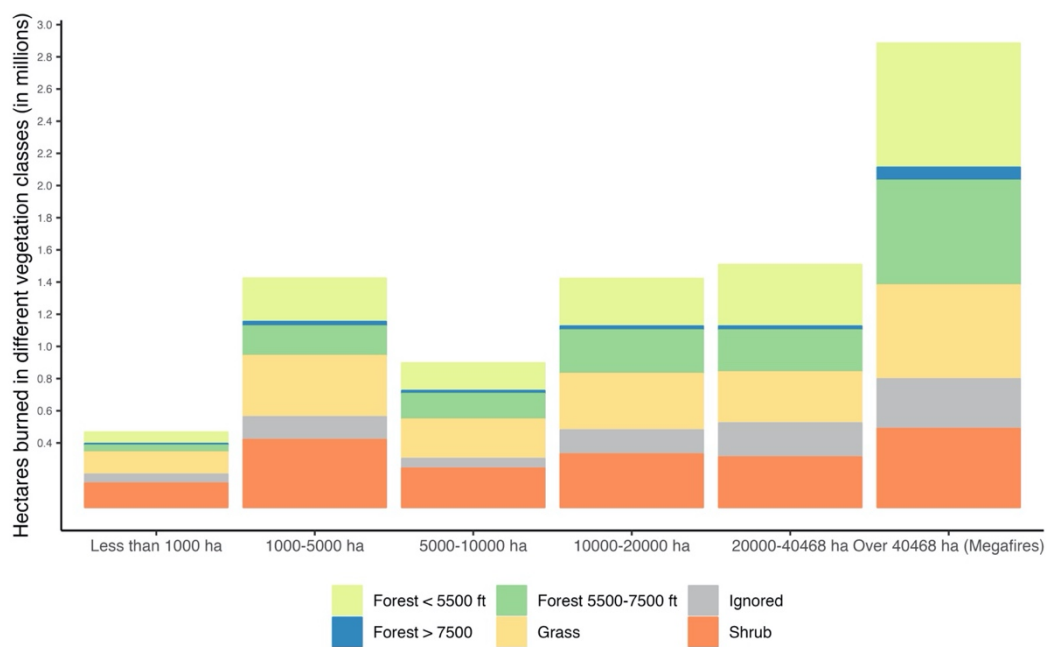


Figure S3.1 Hectares burned in different vegetation classes (in millions) by size group

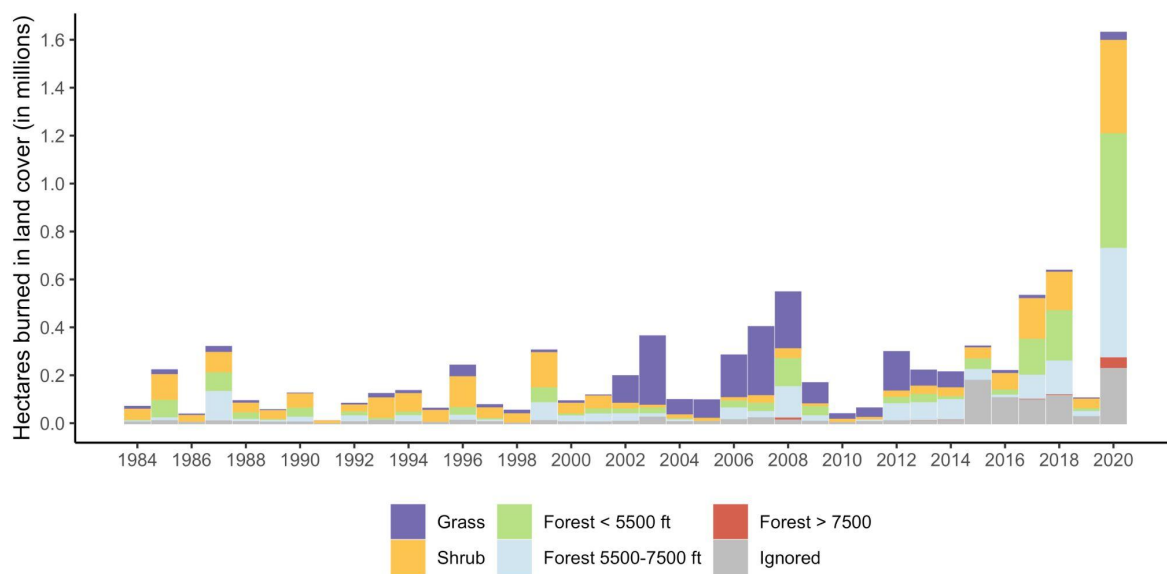


Figure S3.2 Annual area burned in different land cover

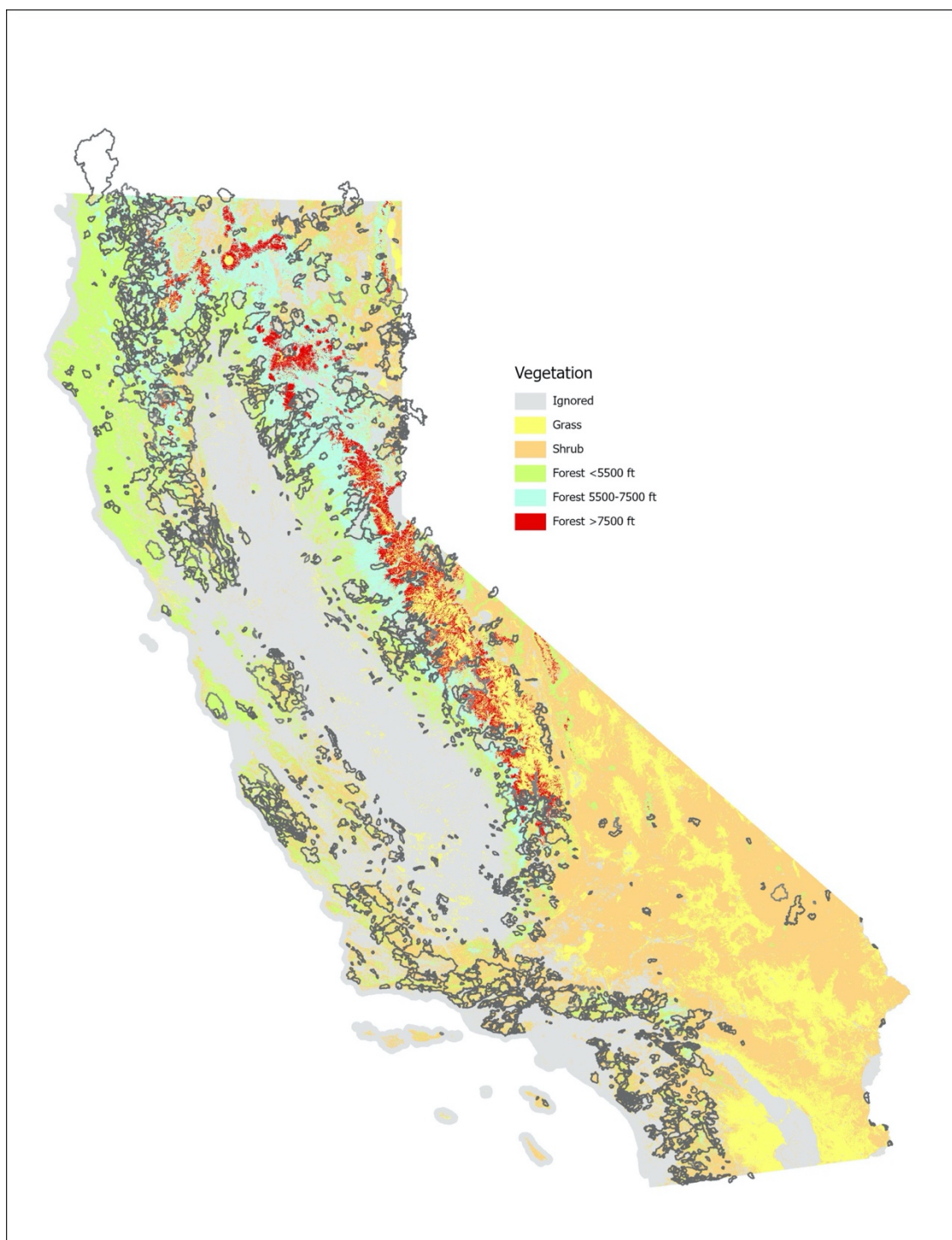


Figure S3.3 Area burned in different land cover during 1984-2020

Table S4.1 ICD-9 and ICD-10 codes for each outcome

Outcome	ICD-9 code	ICD-10 code
Asthma	493	J45
Chronic Obstructive Pulmonary Disorder (COPD)	491,492,496	J40-J44
Pneumonia	480-486	J12,J18
Ischemic Heart Disease (IHD)	410-414	I20-I25
Cardiac Dysrhythmias and Conduction Disorders	426, 427	I44, I45, I47,I49
Heart Failure	428	I50
Cerebrovascular	430-438	I60-I69
Hypertension	401-405	I10-I16

Table S4.2 ICD-9 and ICD-10 codes for each outcome

Holidays	Dates
New Year's Day	January 1
Martin Luther King Jr. Day	the third Monday of January
Presidents' Day	third Monday in February
Cesar Chavez Day	March 31, starting 2014
Memorial Day	the last Monday of May
Independence Day	July 4
Labor Day	the first Monday in September
Veterans Day	November 11
Thanksgiving Day	fourth Thursday of November
Day after Thanksgiving	fourth Thursday of November
Christmas Day	December 26

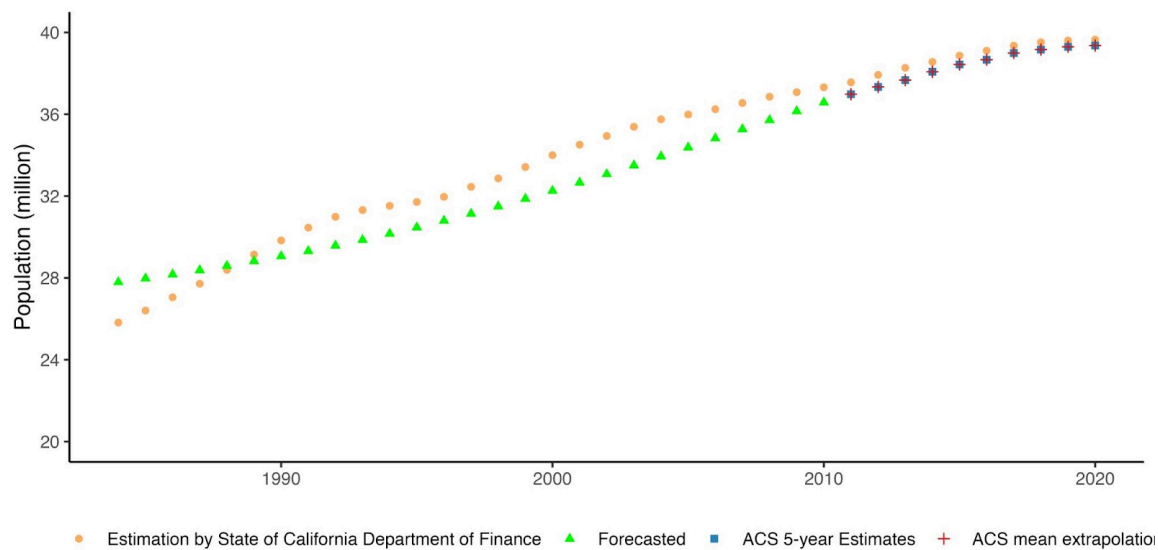


Figure S4.1 Compare backward forecasted zip code level population during 1984-2020 to American Community Survey 5-year estimates (2011-2020) and estimation by State of California Department of Finance

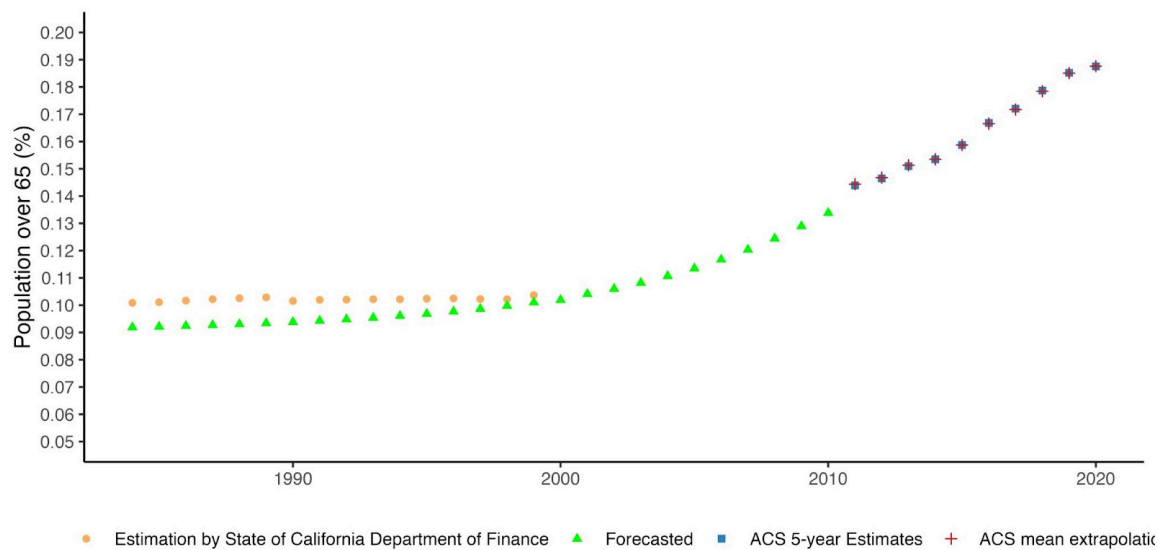


Figure S4.2 Compare backward forecasted zip code level percentage of population over 65 during 1984-2020 to American Community Survey 5-year estimates (2011-2020) and estimation by State of California Department of Finance

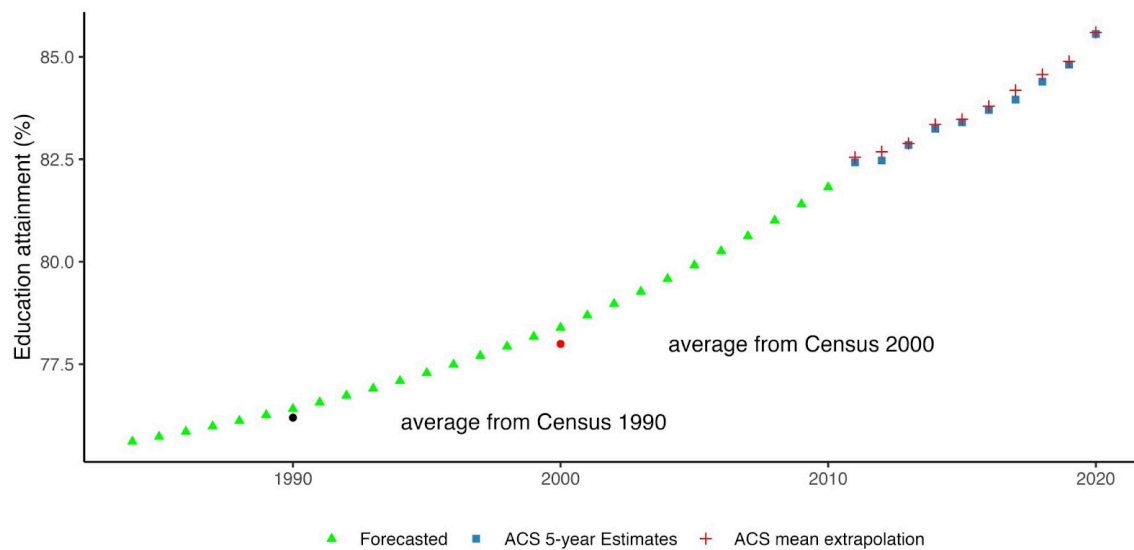


Figure S4.3 Compare backward forecasted zip code level percentage of percent high school graduate or higher in population 25 years and over during 1984-2020 to American Community Survey 5-year estimates (2011-2020) , Decennial census 1990 and 2000 state level

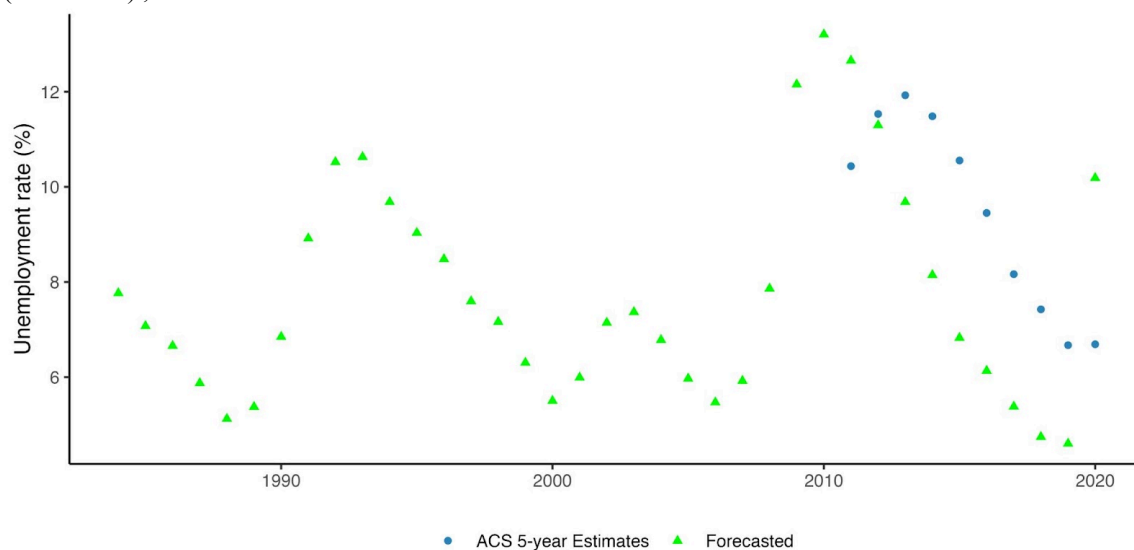


Figure S4.4 Compare annual unemployment rate aggregated to state level from monthly zip code rates derived as an approximate of county rate over during 1984-2020 to American Community Survey 5-year estimates (2011-2020)