A Section 508–conformant HTML version of this article
is available at https://doi.org/10.1289/EHP15672.

A State of the Science Review of Wildfire-Specific Fine Particulate Matter Data Sources, Methods, and Models

Ava Orr,[1,2] Claire E. Adam,[1] Jon Graham,[1,3] Zachary A. Holden,[4] Lu Hu,[5] Zeina Jaffar,[2] Cindy Leary,[1] Christopher T. Migliaccio,[2] Katrina Mullan,[1,6] Curtis Noonan,[1] Erin O. Semmens,[1] Shawn Urbanski,[4] Ethan Walker,[1] and Erin L. Landguth[1]

[1]Center for Population Health Research, University of Montana, Missoula, Montana, USA
[2]Center for Environmental Health Sciences, University of Montana, Missoula, Montana, USA
[3]Mathematics Department, University of Montana, Missoula, Montana, USA
[4]Rocky Mountain Research Station, USDA Forest Service, Missoula, Montana, USA
[5]Department of Chemistry and Biochemistry, University of Montana, Missoula, Montana, USA
[6]Economics Department, University of Montana, Missoula, Montana, USA

BACKGROUND: Despite progress in reducing industrial air pollution, rising wildfire frequency and intensity, driven in part by climate change, pose significant health risks. Accurate estimates of wildfire-generated fine particulate matter with an aerodynamic diameter <2.5 μm ($PM_{2.5}$) are needed for advancing health research, policymaking, and environmental protection.

OBJECTIVE: This review evaluates existing methodologies and data sources for estimating wildfire-generated $PM_{2.5}$, aiming to improving accuracy and accessibility for health research, policy development, and environmental management strategies.

METHODS: We conducted a systematic literature search across Medline, Scopus, Web of Science, Google Scholar, and Embase (January 2018 to March 2024) using keywords such as "$PM_{2.5}$ exposure," and "wildfire $PM_{2.5}$." Studies were included if they were publicly available, focused on North America (primarily the US), and provided wildfire-attributable $PM_{2.5}$ data. Of 2,757 articles identified, 418 full texts were screened, and 33 met inclusion criteria. Four studies offered wildfire-specific estimates of $PM_{2.5}$, and one dataset was excluded due to accessibility issues, leaving three for analysis. We processed data using R (version R 4.3.1; R Development Core Team) at the ZIP code level for consistency and examined total and wildfire-specific $PM_{2.5}$ estimates for California in 2010 (low fire activity) and 2018 (high fire activity), focusing on Los Angeles (densely monitored) and Modoc (no monitors) counties. Analyses included Pearson correlation, cross-correlation, and Granger causality to assess temporal relationships and consistency.

RESULTS: From the 33 studies included, three main estimation approaches emerged: chemical extraction, thresholding, and integration of satellite and fire-specific data (e.g., smoke plumes and fire perimeters). Most studies combined ground-based monitor data, satellite-derived aerosol optical depth, and explanatory data like meteorology and land use. The three public datasets indicated that in California, wildfire-specific $PM_{2.5}$ contributed 11.2%–36.9% of total $PM_{2.5}$ in 2010 and 13.7%–21.2% in 2018 with stronger agreement in 2018. Correlations were stronger in Modoc County (no monitors) (0.44–0.51 in 2010; 0.79–0.88 in 2018) than in Los Angeles County (densely populated area, 20 EPA monitors, where correlations ranged from 0.19–0.21 in 2010 and 0.54–0.79 in 2018). Overall, the datasets estimating total $PM_{2.5}$ were more consistent than wildfire-specific $PM_{2.5}$ estimates.

CONCLUSIONS: We offer a review of current data sources used for wildfire-specific $PM_{2.5}$ estimation and compare publicly available datasets. As expected, the contribution of wildfire smoke to overall $PM_{2.5}$ increased with wildfire activity. However, limited publicly available datasets hinder comprehensive comparisons and generalizations for health research and outcomes. https://doi.org/10.1289/EHP15672

Introduction

Recent decades have seen a dramatic increase in wildfires in the Western US,[1] both in frequency (+50%) and area affected (+390%), exacerbating air pollution levels despite concurrent reductions in emissions from industrial and traffic sources.[2] While climate change creates warmer, drier conditions that heighten wildfire risk, other factors also play a role. Expanded development in the wildland–urban interface introduces more ignition sources, and historical fire suppression has led to dense forests with high fuel loads, making fires harder to control.[3] Migration into these vulnerable regions further increases risks to health and property. Wind patterns carry wildfire smoke beyond the immediate fire zones, increasing exposure in distant communities.[4–8] Extensive studies have unveiled a multitude of adverse effects on human health, linking short-term exposure to wildfire smoke with a range of acute health effects, including heightened rates of cardiovascular mortality,[9–11] adverse birth outcomes,[10,12] mental health effects,[13] elevated instances of asthma-related hospitalizations,[8] and an uptick in emergency room visits.[14,15] With climate change, wildfires are expected to become more frequent, emphasizing the need to better understand the short-term and long-term health implications of wildfire-specific air pollution.[16]

Decades of extensive research have focused on the adverse health effects of air pollution, specifically on particulate matter composed of tiny solid particles and liquid droplets suspended in the air. Particulate matter remains a critical global environmental health risk, linked to ~ 7 million premature deaths annually.[17] In the US, it has been estimated that total air pollution exposure contributed to 230,000–300,000 deaths in 2012 alone.[18] Accurate, source-specific estimates of human exposure to inhaled air pollutants are essential for assessing health impacts and devising mitigation strategies.

This review specifically focuses on fine particulate matter with an aerodynamic diameter <2.5 μm ($PM_{2.5}$) due to its particularly harmful characteristics and the substantial body of evidence

Address correspondence to Ava Orr. Email: ava.orr@mso.umt.edu

Supplemental Material is available online (https://doi.org/10.1289/EHP15672).

The authors declare they have no conflicts of interest related to this work to disclose.

Conclusions and opinions are those of the individual authors and do not necessarily reflect the policies or views of EHP Publishing or the National Institute of Environmental Health Sciences.

EHP is a Diamond Open Access journal published with support from the NIEHS, NIH. All content is public domain unless otherwise noted. Contact the corresponding author for permission before any reuse of content. Full licensing information is available online.

Received 28 June 2024; Revised 2 April 2025; Accepted 28 April 2025; Published 11 June 2025.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

linking it to severe health outcomes. Although wildfire smoke contains other pollutants, such as particulate matter with aerodynamic diameter <10 μm (PM_{10}), polycyclic aromatic hydrocarbons (PAHs), and volatile organic compounds (VOCs), $PM_{2.5}$ is considered the most critical component for health risk assessment. Its small size allows it to penetrate deep into the respiratory system, reaching the lungs, and even entering the bloodstream, leading to acute cardiovascular and pulmonary issues, dementia,[19–21] and other longer-term conditions. In 1996, the US Environmental Protection Agency (EPA) recognized $PM_{2.5}$ as a major pollutant contributing to smog, acid rain, and various health issues under the US National Overall Air Quality Standards (NAAQS).[22] The EPA air monitoring network measures $PM_{2.5}$ nationwide for compliance with the recently revised 2024 primary annual mean standard set at $9\ \mu g/m^3$ (recently reduced from $12\ \mu g/m^3$) and a 35-$\mu g/m^3$ 24-h average.[23] In comparison, the World Health Organization has the annual mean standard set at $5\ \mu g/m^3$ and the 24-h average standard at $15\ \mu g/m^3$.[24]

The regulations within the US, established by the Clean Air Act Amendments,[25] do not account for chemical composition or speciation differences in $PM_{2.5}$ concentrations. Wildfire-generated $PM_{2.5}$ is often considered an "exceptional event" under EPA regulations.[26] As a result, $PM_{2.5}$ emitted from wildfires is categorized similarly to emissions from industrial and traffic sources despite potentially distinct chemical composition and impacts on health. For example, a study revealed that $PM_{2.5}$ specific to wildfires was ten times more detrimental to respiratory hospitalizations than $PM_{2.5}$ from other origins.[27] This generalization is partly due to the challenges in pinpointing the specific components of $PM_{2.5}$, especially those from wildfires, which involve complex measurement techniques, the sporadic nature of fires, and varying seasonal and atmospheric distribution patterns.[27–29] In addition, emerging studies now offer insights into the immunotoxic effects of wildfire-derived $PM_{2.5}$, suggesting that it may be significantly more harmful than its overall counterpart.[30,31] Studies have shown that the chemical composition of wildfire $PM_{2.5}$ has a higher fraction of organic components than other PM.[32,33] *In vitro* and *in vivo* studies indicate that wildfire-specific $PM_{2.5}$ can lead to disease-causing inflammatory cascades through increased inflammation, oxidative stress, and respiratory infection by impacting pulmonary macrophage functions.[30,34] Further insight comes from human-based research that found that increases in wildfire-specific $PM_{2.5}$ caused increases in respiratory hospitalizations and cardiovascular visits.[27,35,36]

Researchers and policymakers select from a variety of $PM_{2.5}$ exposure modeling strategies. Yet, there is limited clarity around the tradeoffs in accuracy and accessibility between different approaches, especially in the context of wildfire seasons. Implicitly, one of the goals of these analyses is to better characterize exposure at locations to understand the impacts of that exposure. For instance, studies have been able to highlight the importance of accurate exposure assessments in particular areas to better evaluate the health effects of $PM_{2.5}$ exposure during wildfire events.[37,38] Here, we aim to clarify the data sources, methodologies, and modeling techniques employed in estimating wildfire-specific $PM_{2.5}$ concentrations. We start with an overview of the general methods and models for compiling data on $PM_{2.5}$ from all sources. This foundation allows us to explore the specialized data and approaches for isolating wildfire-derived $PM_{2.5}$ (see Figure 1 for a roadmap). We then introduce the publicly accessible wildfire-specific $PM_{2.5}$ datasets available for research and policymaking with a comparative analysis and summary statistics of select wildfire-specific $PM_{2.5}$ data. We conclude with an update on the challenges, progress, and future directions for wildfire-specific $PM_{2.5}$ estimation.

Methods

Literature Search and Study Selection

We conducted literature searches in multiple databases, including MEDLINE, Web of Science, Google Scholar, EMBASE, and Scopus, to identify publicly available wildfire-specific $PM_{2.5}$ exposure datasets. We focused on publications published from January 2018 to March 2024, although data did not have to be within this period. We chose this timeframe to build upon prior reviews and to capture the most recent research, methods, and regulatory updates about wildfire-specific $PM_{2.5}$ exposure estimation (Table S1). Our search strategy utilized a range of keywords and phrases specifically targeting wildfire-specific $PM_{2.5}$ exposure estimates, including terms such as "$PM_{2.5}$ exposure estimates," "wildfire-specific $PM_{2.5}$," "Estimating $PM_{2.5}$ exposure," and "wildfire $PM_{2.5}$." These terms were individually searched without quotation marks using API queries, combined using the AND Boolean operator within each database. These search terms were selected to ensure the inclusion of all relevant studies addressing wildfire-specific $PM_{2.5}$. Figure 2 outlines the specific methodology and search terms used. Initially, our search yielded ∼220,000 results. Due to extensive indexing in databases such as Google Scholar and Scopus, initial queries yielded very high numbers of results, including many irrelevant articles, such as those related to indoor air pollution, general air quality studies not focused on wildfires, epidemiological studies unrelated to $PM_{2.5}$ exposure from wildfires, and broader environmental or climate studies that did not specifically address wildfire emissions. Using Python scripts, we systematically filtered out irrelevant articles as described above, along with reviews and non-English articles, resulting in 2,757 studies; after the removal of duplicates, we were left with 1,102 studies. Two authors (A.O. and E.L.L.) independently screened these results based on predefined inclusion criteria. Disagreements were resolved through discussion and consensus; if no consensus could be reached, a third reviewer (C.L.) was consulted.

Our review was conducted in two stages. First, we identified all studies conducted in North America (US and Canada) that attempted to distinguish a wildfire-specific $PM_{2.5}$ component from overall ambient $PM_{2.5}$. This initial search resulted in 33 studies that met these general criteria. Second, from among the 33 studies, we applied a more specific inclusion criteria for comparison: studies had to *a*) provide or reference publicly available (downloadable) wildfire-specific $PM_{2.5}$ datasets and *b*) include daily $PM_{2.5}$ data covering California, due to its frequent and severe wildfire activity.

Of the 33 studies, 29 studies did not offer publicly available datasets but are summarized in Table S1. These studies contributed valuable insights into different methodological approaches used in the field, highlighting the range of strategies for wildfire source attribution. By including them, we aimed to provide a broader understanding of how different research groups have attempted to separate wildfire emissions from other $PM_{2.5}$ sources, complementing the direct dataset comparisons conducted in our analysis. Four of the 33 studies provided publicly available wildfire-specific $PM_{2.5}$ datasets. However, after further evaluation, only three datasets (Aguilera et al.,[39] Childs et al.,[6] and Zhang et al.[40]) were used in the final analysis. One dataset was excluded (Reid et al.[41]) due to spatial and temporal incompatibilities. Building on the insights from the 29 studies, we focused on these three articles that provided both methodological detail and publicly available data.

Each of the three datasets had statewide California coverage, enabling direct comparison across spatial and temporal scales. The Aguilera et al.[39] dataset provided 1-km × 1-km $PM_{2.5}$ estimates across California, derived from satellite data and ground-based sensors, focusing on wildfire smoke contributions. The Childs et al.[6] dataset estimated $PM_{2.5}$ exposure across the US at a

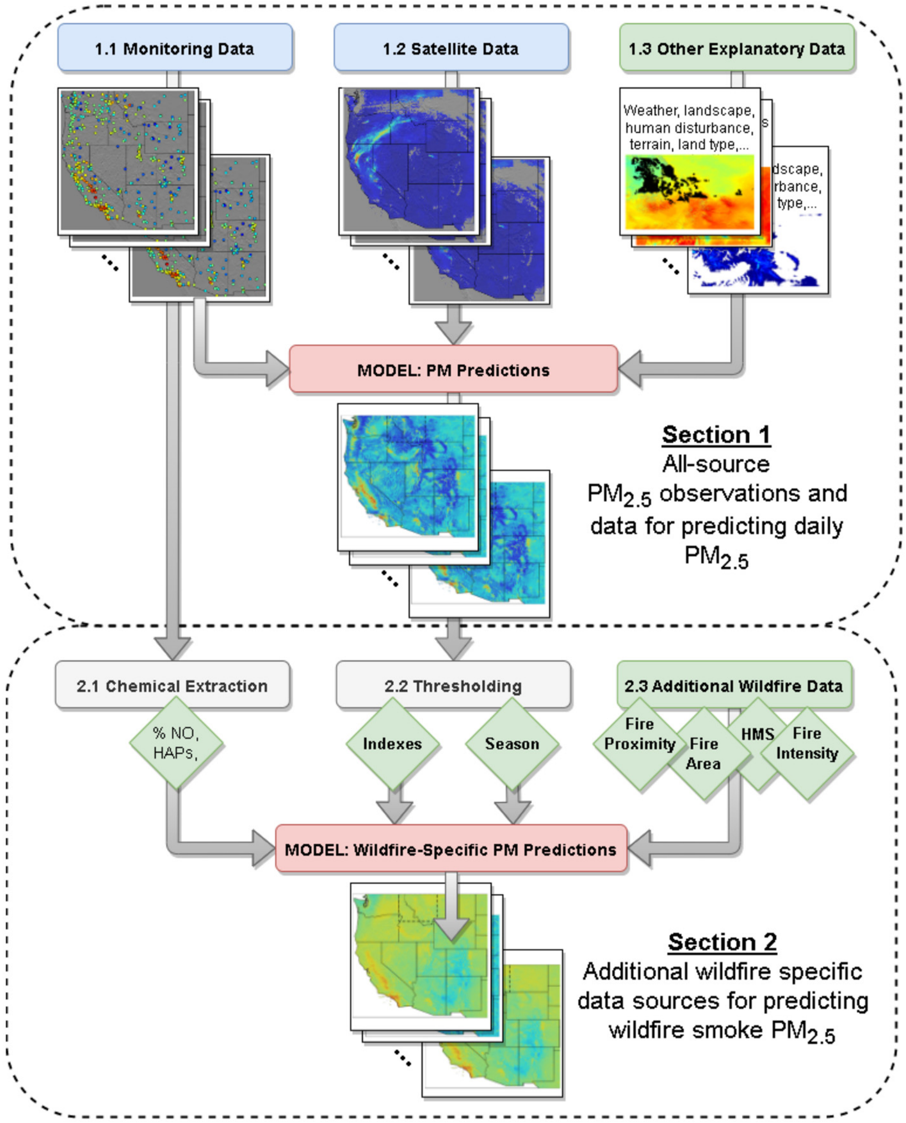

Figure 1. A conceptual diagram to explain the data sources used in estimating spatial and temporal patterns of $PM_{2.5}$ (the three basic data types of all-source $PM_{2.5}$ estimates) and additional data included for estimating spatiotemporal wildfire-specific $PM_{2.5}$ (the three basic approaches for extracting wildfire-specific $PM_{2.5}$ estimates). Blue boxes (Ground-based $PM_{2.5}$ monitoring data and Satellite data for $PM_{2.5}$ estimates) provide the observational $PM_{2.5}$ data, and gray boxes (Chemical extraction, Thresholding for wildfire-specific $PM_{2.5}$) illustrate additional techniques applied to $PM_{2.5}$ layers to extract the additional explanatory data and wildfire-specific data (green boxes and diamonds). Red boxes illustrate the PM predictions for all sources (the three basic data types of all-source $PM_{2.5}$ estimates) and wildfire-specific sources (the three basic approaches for extracting wildfire-specific $PM_{2.5}$ estimates). Note: HAPS, Health and Air Pollution Surveillance System; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm.

10-km × 10-km resolution. The Zhang et al.[40] dataset used satellite-derived smoke plumes and air quality modeling to produce 5-km wildfire-specific $PM_{2.5}$ estimates. Dataset details, including geographic scope, modeling methods, and sources, are provided in Table 1.

By comparing these three datasets, we were able to directly evaluate the consistency, strengths, and limitations of each dataset in estimating wildfire-specific $PM_{2.5}$ across different regions and fire conditions. The emphasis on US studies was practical because of the comparability of regulatory frameworks (e.g., US EPA standards), monitoring networks (e.g., AirNow and PurpleAir), and wildfire behavior patterns. Although regional differences within the US (e.g., Pacific Northwest vs. Southeast) exist, focusing within North America ensured greater consistency across datasets compared to including studies from other continents with substantially different wildfire dynamics and regulatory systems.

Data Extraction from Identified Studies

We reviewed the 29 studies that did not provide downloadable or directly usable datasets. These studies, summarized in Table S1, played a critical role in shaping our understanding of the broader methodological landscape for estimating wildfire-specific $PM_{2.5}$. From these articles, we systematically extracted key methodological information to inform the review's classification and synthesis sections. Specifically, we recorded details about each study's geographic and temporal coverage, the modeling frameworks employed (e.g., chemical transport models, machine learning algorithms, or ensemble methods), and the types of input data used, such as ground-based monitors, satellite-derived aerosol optical depth, fire emission inventories, meteorological variables, and land use data. We also noted the methods used to isolate wildfire-related $PM_{2.5}$, including plume modeling, smoke subtraction approaches, use of fire perimeter data, and thresholding

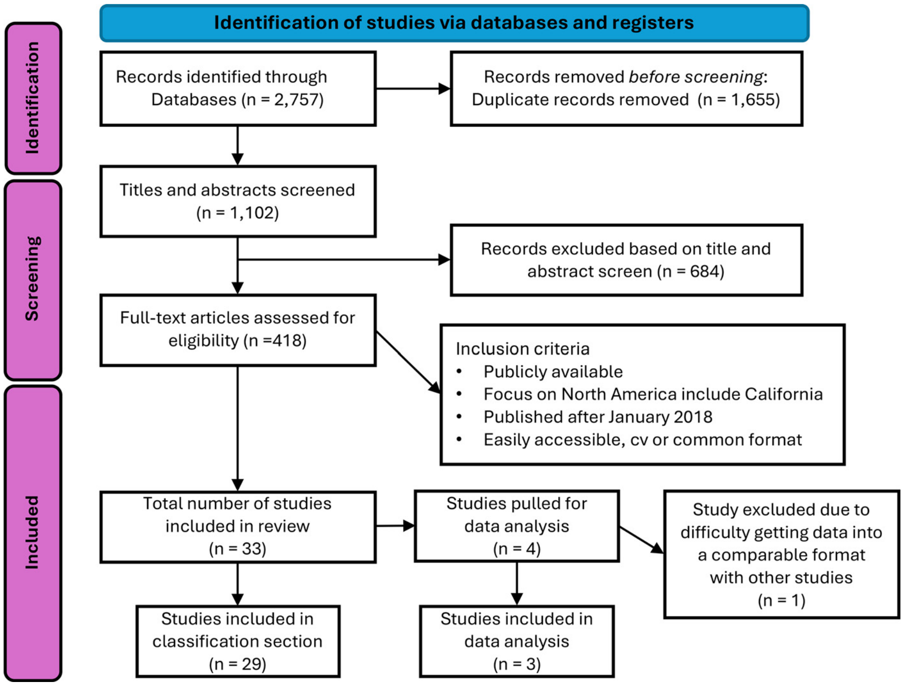

Figure 2. Flow diagram of study selection. Initially, 2,757 records were identified through our comprehensive database searches, which included MEDLINE, Web of Science, Google Scholar, EMBASE, and Scopus. After duplicates were removed, 1,102 records were screened, leading to 684 being excluded after title and abstract review based on predefined criteria. The search was conducted using an extensive list of keywords and phrases to ensure comprehensive coverage of all pertinent studies on wildfire-specific $PM_{2.5}$ exposure. These keywords included: "$PM_{2.5}$ exposure estimates," "Estimating $PM_{2.5}$ exposure," "$PM_{2.5}$ pollution," "$PM_{2.5}$ estimates," "wildfires," "wildfire-specific $PM_{2.5}$," "wildfire $PM_{2.5}$," "wildfire air quality impacts," "particulate matter from wildfires," "smoke pollution," "wildfire pollution," "airborne particulates from wildfires," "wildfire smoke assessment," "smoke exposure," "wildfire particulate analysis," "atmospheric particulates from wildfires," "fire-related air pollution," and "forest fire pollution." These terms were chosen to capture a broad spectrum of research related to $PM_{2.5}$ emissions from wildfires. Further assessment of 418 full-text articles resulted in 384 exclusions due to not being assessable (71), study covering North America (40), relevant to the study (273), leaving 33 studies included in the final review. Only 4 out of the 33 studies included in this review had publicly available data. In the final dataset comparison analysis, we only include 3 of the 4 publicly available datasets due to accessibility issues; see methods section for detailed inclusion and exclusion criteria. Note: $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm.

techniques, as well as any reported model performance metrics like R^2 and root mean square error (RMSE). The information gleaned from these 29 studies directly informed the typology described in the three basic data types of all-source $PM_{2.5}$ estimates (three basic data types used in $PM_{2.5}$ estimation: ground-based monitors, satellite data, and additional explanatory variables) and the categorization in the three basic approaches for extracting wildfire-specific $PM_{2.5}$ estimates (three major approaches to extracting wildfire-specific $PM_{2.5}$: chemical extraction, thresholding, and wildfire-specific auxiliary data). While these studies did not contribute raw datasets for side-by-side numerical comparison, their methodological contributions were foundational for building the framework used to evaluate and contextualize the three core datasets analyzed in this review.

Data Extraction from Datasets

Data extraction from the selected datasets was conducted using R (version 4.4.2; R Development Core Team). R was used to standardize the data (e.g., same time periods, same location) formats across the three datasets and perform a comparative analysis of $PM_{2.5}$ exposure estimates. This analysis focused on comparing the consistency of $PM_{2.5}$ estimates across different regions, with particular emphasis on California, which was included in all three datasets. Data cleaning and harmonization were carried out to ensure uniform temporal and spatial resolution, enabling direct comparison of $PM_{2.5}$ estimates across the studies. We processed all data at the ZIP code level. This approach was chosen because EPA monitoring data can be geolocated and assigned to specific ZIP codes, allowing for a standardized spatial unit across all datasets.

Table 1. A summary of publicly available wildfire-specific $PM_{2.5}$ datasets used in this comparative study.

Source of dataset	Region	Time period	Spatial resolution	Temporal resolution	Ground monitor	Model	Satellite	Other variables
Aguilera et al.[27]	CA	2006–2020	ZIP code	Daily	EPA	RF, ML	AOD	AB, H, P, SSR, T, TI, WD, WP, WV
Childs et al.[6]	CONUS	2006–2020	10×10 km	Daily	EPA	ML, HYSPLIT	AOD, MODIS	P, RH, SP, T, VPD, WS
Zhang et al.[40]	CONUS	2007–2018	ZIP code	Daily	EPA, PurpleAir	CMAQ, RF, SMOTE, ML	AOD, MAIAC, MODIS	LC, M, PO, RH, SP, T

Note: The URLs of the datasets are listed as follows: Aguilera et al. dataset https://github.com/benmarhnia-lab/Wildfire_PM25_California_ZIP_and_census_tract, Childs et al. dataset https://github.com/echolab-stanford/daily-10km-smokePM, and Zhang et al. dataset https://figshare.com/articles/dataset/Zip_code-level_wildfire_smoke_PM_sub_2_5_sub_data_for_CONUS_2007-2018/25016510. Table 7 shows spatially continuous wildfire-specific $PM_{2.5}$ datasets that are publicly available and free. AB, air basin; AOD, aerosol optical depth; CMAQ, community multiscale air quality; EPA, Environmental Protection Agency; H, humidity; HYSPLIT, hybrid single-particle Lagrangian integrated trajectory; LC, land cover; M, meteorological; MAIAC, multi-angle implementation of atmospheric correction; ML, machine learning; MODIS, Moderate Resolution Imaging Spectroradiometer; P, precipitation; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm; PO, population; RF, random forest; RH, relative humidity; SMOTE, synthetic minority oversampling technique; SP, smoke plumes; SSR, surface shortwave radiation; T, temperature; TI, thermal inversions; VPD, vapor pressure deficit; WD, wind direction; WP, wildfire perimeter; WV, wind velocity.

Table 2. Summary of publicly available wildfire-specific and ambient $PM_{2.5}$ datasets for California by ZIP code for 2010 and 2018.

Dataset	Type	Year	Minimum	Mean	Median	IQR	Maximum	Total	Days above WHO 15 ($\mu g/m^3$) standard	Days above EPA 35 ($\mu g/m^3$) standard
EPA[42]	Overall	2010	0.0	10.19	8.5	7.9	114.8	35,763	35.0	0.0
Aguilera et al.[27]	Overall	2010	0.0	8.81	7.7	5.82	97.15	649,335	6.0	0.0
Childs et al.[6]	Overall	2010	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2010	1.02	10.16	9.26	7.0	54.58	605,900	11.0	0.0
Aguilera et al.[27]	WF	2010	0.0	1.82	1.0	0.0	67.71	29,602	0.0	0.0
Childs et al.[6]	WF	2010	0.0	1.98	1.63	2.14	28.53	20,716	0.0	0.0
Zhang et al.[40]	WF	2010	0.0	1.68	1.4	0.0	44.61	140,323	0.0	0.0
EPA[42]	Overall	2018	0.0	11.38	8.5	7.8	417.0	51,482	54.0	8.0
Aguilera et al.[27]	Overall	2018	0.0	10.29	8.08	6.36	285.1	649,333	43.0	5.0
Childs et al.[6]	Overall	2018	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2018	1.15	12.26	10.27	7.27	279.47	605,900	60.0	7.0
Aguilera et al.[27]	WF	2018	3.96	11.15	3.33	0.0	277.86	104,574	11.0	2.0
Childs et al.[6]	WF	2018	0.0	14.16	4.68	11.87	294.48	94,215	28.0	8.0
Zhang et al.[40]	WF	2018	3.29	5.27	2.26	2.42	257.31	322,777	14.0	2.0

Note: All values reported in $\mu g/m^3$. Total is the aggregate $PM_{2.5}$ concentration over all days. Overall includes all sources, while WF is $PM_{2.5}$ specifically from wildfire smoke, based on models. Days above WHO's 15 $\mu g/m^3$ guideline and EPA's 35 $\mu g/m^3$ standard are noted. Ambient $PM_{2.5}$ data is from raw EPA monitoring, while WF $PM_{2.5}$ values are modeled. Information obtained from the Environmental Protection Agency (EPA)[6,22,39,40] and World Health Organization (WHO). IQR, interquartile range; NA, data unavailable; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm; WF, wildfire.

All datasets utilized in the analysis were supplied in CSV format, containing daily mean estimates of extracted $PM_{2.5}$ data at the ZIP code level in California. By using ZIP code–level aggregation, we were able to ensure a uniform spatial resolution across the datasets for meaningful comparison. In our analysis, we aggregated EPA data from all air quality monitoring stations in California over the period 2010–2018. We extracted summary values of $PM_{2.5}$ for each ZIP code by mapping the geographic location of each ground monitor to its corresponding ZIP code. This allowed us to directly compare EPA observations[42] with estimates from the Aguilera et al.[39] and Zhang et al.[40] datasets. The comparison includes both Modoc and Los Angeles counties to explore how $PM_{2.5}$ estimates from models and monitors align in areas with starkly different monitoring availability. Modoc County, located in the northeastern corner of California, had no EPA ground monitors[42] and a relatively low population of 9,686 (2010).[43] In contrast, Los Angeles County in southern California had the highest concentration of monitors (20) and a dense population of $\sim$9.8 million (2010).[43] Including both counties allowed us to investigate how the model estimates compare in a densely monitored, high-population region vs. a sparsely monitored, low-population one. This setup helped assess the models' ability to estimate $PM_{2.5}$ across areas with varying population densities, pollution sources, and monitoring availability. It is important to note that the EPA monitors are not evenly distributed across all ZIP Codes, and they are typically placed in areas with higher population density or greater pollution levels.[42] This means that EPA data, when restricted to ZIP codes with monitors, may inherently reflect higher $PM_{2.5}$ levels compared to modeled estimates that cover broader areas, including ZIP codes without monitors. When comparing Aguilera et al. and Zhang et al.'s total $PM_{2.5}$ estimates to EPA values, we did not restrict the comparison solely to ZIP codes with monitors. This could explain why EPA data often showed higher $PM_{2.5}$ concentrations, as monitors tended to be in regions more prone to pollution. The modeled datasets (Aguilera et al.[39] and Zhang et al.[40]) provided estimates for all ZIP codes, offering a more comprehensive geographic coverage, which may have resulted in lower average $PM_{2.5}$ values when aggregated across the entire state.

Statistical analyses were conducted using R. We conducted temporal comparisons by isolating the years 2010 and 2018 from the identified datasets, representing years with low and high wildfire activity, respectively. In 2010, California experienced relatively low wildfire activity (6,554 wildfires that burned 109,529 acres),[44] whereas 2018 was the deadliest and most destructive wildfire year on record (8,527 wildfires that burned 1.8 million acres, resulting in 103 fatalities).[45] These contrasting years allowed us to explore how different datasets estimate wildfire-specific and overall $PM_{2.5}$ during periods of varying fire intensity. Tables 2–4 provide a detailed summary of all combinations of datasets (overall vs. wildfire-specific $PM_{2.5}$) across years (2010 vs. 2018) and locations (statewide and two counties). To evaluate the relationships between datasets, we performed cross-correlation and Granger causality tests. Cross-correlation was used to measure the strength and timing of associations between $PM_{2.5}$ estimates across different datasets, helping identify potential time lags in agreement between modeled and observed values. Pearson correlation coefficients assessed the linear relationships between overall $PM_{2.5}$ estimates (Aguilera et al.,[39] EPA,[42] and Zhang et al.[40]) and wildfire-specific $PM_{2.5}$ estimates (Aguilera et al.,[39] Childs et al.,[6] and Zhang et al.[40]), providing insight into the consistency of model predictions with EPA observations. Additionally, Granger causality tests were applied to examine whether past values of one dataset could predict future values of another.[46] This test helps determine whether wildfire-specific $PM_{2.5}$ estimates from different models provide predictive value over time, offering insights into how well different datasets capture wildfire smoke trends.

Results and Discussion

Our initial literature search returned 2,757 articles, with duplicates removed and titles and abstracts screened to refine the selection to 418 full-text articles for review. Application of exclusion/inclusion criteria narrowed the field to 33 articles that met our specific criteria for detailed examination. From these, we initially identified four publicly available datasets meeting our inclusion criteria. However, the Reid et al.[41] dataset was subsequently excluded due to challenges in accessibility and format standardization, leaving three datasets for detailed analysis: Aguilera et al.,[39] Childs et al.,[6] and Zhang et al.[40] These three datasets were used in our comparative analysis to assess consistency and variability in $PM_{2.5}$ estimates across different geographic and environmental contexts (publicly available wildfire-specific $PM_{2.5}$ datasets, comparisons, and discussion). The details of each dataset's geographic scope, data sources, and specific $PM_{2.5}$ estimation methods are provided in Table 1. Additionally, the other 29 articles that were not part of this comparative dataset analysis are summarized in Table S1, offering a broader overview of the research landscape and supporting data on wildfire-specific $PM_{2.5}$ exposure (the three basic data types of all-source $PM_{2.5}$ estimates and the three basic approaches for extracting wildfire-specific $PM_{2.5}$ estimates).

The Aguilera et al. dataset provided 1-km × 1-km $PM_{2.5}$ estimates across California, derived from satellite data and ground-

Table 3. Summary of publicly available wildfire-specific and ambient $PM_{2.5}$ datasets for Modoc, USA by ZIP code for 2010 and 2018.

Dataset	Type	Year	Minimum	Mean	Median	IQR	Maximum	Total	Days above WHO 15 ($\mu g/m^3$) standard	Days above EPA 35 ($\mu g/m^3$) standard
EPA[42]	Overall	2010	0.0	10.19	8.5	7.9	114.8	35,763	35.0	0.0
Aguilera et al.[27]	Overall	2010	0.0	8.81	7.7	5.82	97.15	649,335	6.0	0.0
Childs et al.[6]	Overall	2010	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2010	1.02	10.16	9.26	7.0	54.58	605,900	11.0	0.0
Aguilera et al.[27]	WF	2010	0.0	1.82	1.0	0.0	67.71	29,602	0.0	0.0
Childs et al.[6]	WF	2010	0.0	1.98	1.63	2.14	28.53	20,716	0.0	0.0
Zhang et al.[40]	WF	2010	0.0	1.68	1.4	0.0	44.61	140,323	0.0	0.0
EPA[42]	Overall	2018	0.0	11.38	8.5	7.8	417.0	51,482	54.0	8.0
Aguilera et al.[27]	Overall	2018	0.0	10.29	8.08	6.36	285.1	649,333	43.0	5.0
Childs et al.[6]	Overall	2018	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2018	1.15	12.26	10.27	7.27	279.47	605,900	60.0	7.0
Aguilera et al.[27]	WF	2018	3.96	11.15	3.33	0.0	277.86	104,574	11.0	2.0
Childs et al.[6]	WF	2018	0.0	14.16	4.68	11.87	294.48	94,215	28.0	8.0
Zhang et al.[40]	WF	2018	3.29	5.27	2.26	2.42	257.31	322,777	14.0	2.0

Note: All values reported in $\mu g/m^3$. Total is the aggregate $PM_{2.5}$ concentration over all days. Overall includes all sources, while WF is $PM_{2.5}$ specifically from wildfire smoke, based on models. Days above WHO's 15 $\mu g/m^3$ guideline and EPA's 35 $\mu g/m^3$ standard are noted. Ambient $PM_{2.5}$ data is from raw EPA monitoring, while WF $PM_{2.5}$ values are modeled. Information obtained from the Environmental Protection Agency (EPA)[6,22,39,40] and World Health Organization (WHO). IQR, interquartile range; NA, data unavailable; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm; WF, wildfire.

based sensors, with a focus on wildfire smoke contributions. The ensemble model in Aguilera's study achieved an R^2 of 0.78 for the hold-out test dataset, with a RMSE of 3.51 $\mu g/m^3$, reflecting strong predictive performance in estimating $PM_{2.5}$ concentrations at the ZIP code level across California over a 15-year period.[39] The Childs et al. dataset covered $PM_{2.5}$ exposure estimates across the US at a 10-km × 10-km resolution, incorporating wildfire-specific data, particularly in the western US. Childs et al. reported an R^2 of 0.67 and did not report an RMSE. The Zhang et al. dataset focused on $PM_{2.5}$ estimates attributed to wildfires using satellite-derived smoke plume data and air quality models, covering key wildfire regions in the US, including California at a 5-km resolution. The model's overall R^2 was 0.75 with a RMSE of 4.59 $\mu g/m^3$[40] for smoke-impacted regions.[6] The no-smoke regions yielded an overall R^2 of 0.68 and RMSE of 3.35 $\mu g/m^3$.[40] The details of each dataset, including their geographic scope, data sources, and specific $PM_{2.5}$ estimation methods, are provided in Table 1.

The Three Basic Data Types of All-Source $PM_{2.5}$ Estimates

The goal of this review was to assess methodologies for estimating wildfire-specific $PM_{2.5}$ by first establishing a foundation in all-source $PM_{2.5}$ data. Understanding how baseline $PM_{2.5}$ estimates are generated is essential for identifying the strengths and limitations of different approaches in isolating wildfire smoke contributions. Based on our synthesis of the 33 studies identified through the literature search, we categorized these data sources into three key types, ground-based monitor data, satellite data, and other explanatory data, each offering distinct advantages and constraints for air quality estimation. These 33 studies were not only used to identify candidate datasets for comparative analysis but also served as the basis for characterizing the broader methodological landscape of wildfire-specific $PM_{2.5}$ estimation. We used these studies to develop a classification framework for both all-source $PM_{2.5}$ data inputs and the approaches used to isolate wildfire-specific contributions.

Ground-based $PM_{2.5}$ monitoring data. Ground-based monitoring data play a crucial role in air quality assessment and regulatory compliance in the US, as mandated by the Clean Air Act.[25] State, local, and tribal monitoring agencies are required to assess overall $PM_{2.5}$ levels across the continental US for adherence to the $PM_{2.5}$ National Ambient Air Quality Standards (NAAQS) and support policy research to protect public health and the environment. The EPA officially recognizes these data collected using Federal Reference Methods (FRM), Federal Equivalent Methods (FEM), or non-FRM techniques (e.g., PurpleAir), ensuring that the data meet air quality standards for $PM_{2.5}$ and other pollutants.

Table 4. Summary of publicly available wildfire-specific and ambient $PM_{2.5}$ datasets for Los Angeles, California by ZIP code for 2010 and 2018.

Dataset	Type	Year	Minimum	Mean	Median	IQR	Maximum	Total	Days above WHO 15 ($\mu g/m^3$) standard	Days above EPA 35 ($\mu g/m^3$) standard
EPA[42]	Overall	2010	3.2	12.87	12.42	6.46	31.28	4,326	109.0	0.0
Aguilera et al.[27]	Overall	2010	3.54	11.97	12.13	4.13	25.31	105,120	51.0	0.0
Childs et al.[6]	Overall	2010	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2010	4.99	13.8	13.88	3.02	26.31	105,485	97.0	0.0
Aguilera et al.[27]	WF	2010	0.0	0.03	0.0	0.0	1.62	3,231	0.0	0.0
Childs et al.[6]	WF	2010	0.0	1.38	1.3	1.63	3.8	972	0.0	0.0
Zhang et al.[40]	WF	2010	0.0	0.1	0.0	0.0	2.24	8,818	0.0	0.0
EPA[42]	Overall	2018	0.0	12.24	10.8	8.2	111.0	4,646	92.0	2.0
Aguilera et al.[27]	Overall	2018	0.0	11.87	11.35	6.14	96.57	105,120	69.0	0.0
Childs et al.[6]	Overall	2018	NA	NA	NA	NA	NA	NA	NA	NA
Zhang et al.[40]	Overall	2018	1.99	13.51	13.62	5.62	80.65	105,485	106.0	1.0
Aguilera et al.[27]	WF	2018	0.0	2.72	1.21	0.0	86.42	11,094	0.0	0.0
Childs et al.[6]	WF	2018	0.0	4.74	3.2	3.16	131.68	7,605	2.0	0.0
Zhang et al.[40]	WF	2018	1.8	2.34	1.65	1.48	66.33	47,488	2.0	0.0

Note: All values reported in $\mu g/m^3$. Total is the aggregate $PM_{2.5}$ concentration over all days. Overall includes all sources, while WF is $PM_{2.5}$ specifically from wildfire smoke, based on models. Days above WHO's 15 $\mu g/m^3$ guideline and EPA's 35 $\mu g/m^3$ standard are noted. Ambient $PM_{2.5}$ data is from raw EPA monitoring, while WF $PM_{2.5}$ values are modeled. Information obtained from the Environmental Protection Agency (EPA)[6,22,39,40] and World Health Organization (WHO). IQR, interquartile range; NA, data unavailable; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 μm; WF, wildfire.

The data collected by these networks are categorized into continuous (hourly averages) and intermittent (3- or 24-h samples) measurements that provide near-surface concentration estimates of $PM_{2.5}$. For environmental health studies, researchers will use these $PM_{2.5}$ observations that are closest to or aggregated to the spatial unit of reference (residence-, census block group–, ZIP code–level, etc.) and/or through spatial weighting or interpolation (e.g., Dockery and Pope[47]).

However, these stations can be unevenly distributed, especially in remote, rural, or economically disadvantaged areas, leading to significant spatial gaps in data accuracy.[48] Efforts to expand monitoring, such as temporary monitoring, auxiliary networks, and grassroots initiatives (e.g., PurpleAir Network), have begun to address these gaps over time (see Table S3). The availability of low-cost sensors that can be deployed and operated with little technical expertise has fostered expanded $PM_{2.5}$ monitoring in rural states in recent years (see the AirNow Fire and Smoke Map online for a map of low-cost sensor location). While independent monitoring networks are growing, EPA FRM-approved monitors have also expanded (as illustrated in Figure S1), indicating a growing capability to monitor air quality across diverse geographic areas. Additionally, to address the impact of wildfire smoke, federal, state, and tribal agencies have started programs to deploy temporary $PM_{2.5}$ monitors as part of wildfire incident response.[49] While these auxiliary networks utilize non-FRM/non-FEM methods and do not conform to EPA-approved monitoring standards, they can provide valuable insights for scientific studies and air quality assessments when properly calibrated.[50–52]

Satellite data for $PM_{2.5}$ estimates. Despite adding these auxiliary ground-based monitors described previously, gaps persist. Researchers and environmental agencies utilize satellite data as a supplementary $PM_{2.5}$ source to achieve more comprehensive spatial coverage. Satellite remote sensing offers a broad view of the Earth's surface and atmosphere, capturing data from the "top-down" on particulate matter that can be used to estimate $PM_{2.5}$ levels within the atmosphere (see Table S4 for examples of satellite data sources). Satellite-derived aerosol optical depth (AOD) observations measure how much light is absorbed or scattered by aerosol particles in the atmosphere.[53] Higher AOD values indicate more particles in the air, which can be due to various sources, such as pollution, dust storms, volcanic dust, and wildfires.[54] Integrating ground-based $PM_{2.5}$ measurements with satellite-derived AOD observations has enhanced air quality monitoring for several reasons.[53] Firstly, in the absence of interfering clouds or snow features, AOD provides a broad, spatially continuous view of aerosol distribution over the Earth's surface, such as in remote areas with sparse ground-based air quality monitoring stations. Secondly, when AOD data are calibrated and correlated with ground-based $PM_{2.5}$ measurements, it can help estimate $PM_{2.5}$ concentrations with higher spatial resolution than either data source used alone (Figure S2).[55] This is particularly valuable for tracking the spatial extent of wildfire smoke plumes, which surface monitoring stations may not capture.[56]

Satellite observations are not without their limitations. As illustrated in Figure S2, not all ground-based $PM_{2.5}$ monitoring stations correlate well with AOD. This is particularly true in regions with frequent cloud cover, extensive and/or persistent snow cover, bright land surfaces, or less than optimal viewing angles, which can result in data gaps[57] (see Figure S3 for an example of missingness). In addressing the challenge of missing data and ensuring the reliability of satellite-derived information, researchers have devised correction techniques to refine the precision of their analyses. This includes strategies for gap-filling missing pixels, such as when AOD values are obscured by heavy clouds or snow cover and for filtering out erroneous pixels where the provided AOD values may be unreliable. For example, Swanson et al.[58] employed a combination of methods to tackle both issues, including quality screening processes, a high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) normalized difference vegetation index (NDVI) filter, and applying a static mask. The careful selection of these correction methods significantly influences the accuracy of the results, especially when confronted with contamination from factors like snow, clouds, or other impurities.[58] Another constraint to consider is that AOD integrates aerosol concentrations throughout the entire atmospheric column,[53] not just at the ground level where human exposure occurs. This becomes particularly relevant for wildfire smoke, as $PM_{2.5}$ from wildfires often remains suspended at high altitudes, especially during transport over long distances. In such cases, the relationship between AOD and surface-level $PM_{2.5}$ can be weaker, as AOD may overestimate surface concentrations when most of the wildfire smoke is aloft. Understanding the vertical distribution of PM often relies on data from chemical transport models, as PM composition can significantly influence AOD. Since PM composition can vary both spatially and temporally, relating AOD to surface PM concentrations frequently depends on insights gleaned from these models.

Using other explanatory data for predicting $PM_{2.5}$. Ground-based monitor data and satellite observations form the backbone of most $PM_{2.5}$ air quality observations and predictions. To effectively estimate $PM_{2.5}$ levels, researchers typically combine these data sources along with additional explanatory data. For example, the complexity of atmospheric conditions and the multitude of factors affecting air pollution levels necessitate incorporating additional explanatory data to enhance the accuracy and reliability of $PM_{2.5}$ estimates. Meteorological variables, such as temperature, humidity,[59] and wind speed, significantly influence the atmosphere's dispersion, formation, and PM lifetime.[60] For instance, high temperatures can increase the rate of chemical reactions in the atmosphere, leading to secondary PM—particulate matter formed through atmospheric chemical reactions from precursor gases emitted by human activities or natural sources.[61] Similarly, wind speed and direction can disperse or accumulate pollutants in specific areas. At the same time, humidity levels affect the hygroscopic growth of PM, altering its concentration and impact on human health.[59,60] Beyond these, variables like human population density, industrial activity, traffic volume,[62] land use patterns, and vegetation[63] cover also play critical roles in determining the spatial and temporal distribution of $PM_{2.5}$ concentrations. For examples of explanatory variables that have been shown to improve $PM_{2.5}$ model fitting, see Table S5.

The Three Basic Approaches for Extracting Wildfire-Specific $PM_{2.5}$ Estimates

Extracting wildfire-specific $PM_{2.5}$ datasets from the overall $PM_{2.5}$ presents unique challenges. The unique and variable chemical composition of wildfire smoke, which is subject to changes based on the fuel source, geographic location, distance traveled, and combustion conditions, significantly differs from other $PM_{2.5}$ sources, such as vehicle emissions or industrial activities.[29] Atmospheric mixing and the chemical transformation of smoke particles further obscure the origins of $PM_{2.5}$, as aging processes alter the smoke's properties, making it chemically resemble other particulates.[64] The spatial and temporal variability of wildfire events, coupled with the often-limited coverage of air quality monitoring networks, especially in rural or remote areas, adds another layer of difficulty. Using the data sources and modeling techniques described in the three basic data types of all-source $PM_{2.5}$ estimates, researchers have turned to many different methodologies to address the complexities of extracting wildfire-specific $PM_{2.5}$ and isolating the impact of wildfire smoke on air quality and health effects. Building on our past research and patterns that we have observed in the literature search, we discuss three different approaches for isolating wildfire-specific $PM_{2.5}$: chemical extraction, thresholding techniques, and additional wildfire-specific data.

Chemical extraction. In the three basic data types of all-source $PM_{2.5}$ estimates, it is noted that ground-based monitors, like the EPA stations in the US, are commonly relied upon for measuring $PM_{2.5}$ levels during wildfire occurrences. However, a significant limitation of these monitors is their inability to inherently distinguish between $PM_{2.5}$ originating from wildfires and that from other sources, as they only measure the total $PM_{2.5}$ mass concentrations. The chemical composition of fresh wildfire $PM_{2.5}$ is mostly organic (organic aerosol; $>85\%$) but also includes soot (black carbon) and other inorganic particles (sulfate, nitrate, ammonium, and chloride).[65,66] The variability in emission composition and concentration is influenced by factors like vegetation type, combustion stage, distance traveled and age, and atmospheric conditions. To tackle the challenge of discerning wildfire contributions, scientists typically resort to a method called chemical speciation[67] to distinguish wildfire-specific $PM_{2.5}$ from other particulate matter. This is done by deploying specialized air sampling equipment, such as high-volume or low-volume samplers, or passive samplers, strategically positioned to collect air samples. These samplers draw air through filters designed to capture particulate matter without altering its chemical makeup. Subsequently, the collected $PM_{2.5}$ particles undergo extraction using techniques like solvent extraction, ultrasonic extraction, or digestion with strong acids or bases, depending on the analytical approach employed. Following extraction, various analytical methods, including gravimetric analysis, chromatography, mass spectrometry, and microscopic analysis, are used to characterize the composition and morphology of the collected $PM_{2.5}$ samples. By utilizing chemical speciation techniques, scientists can pinpoint and quantify specific chemical species within $PM_{2.5}$ emissions, distinguishing them from those originating from alternate combustion sources, such as vehicular emissions or industrial activities. Chemical speciation allows for the distinction between organic compounds originating from plant material combustion and inorganic compounds associated with soil and ash particles. For example, a study identified levoglucosan as a promising biomarker of biomass smoke exposure.[68] One can measure hundreds of organic compounds in smoke plumes, but the significant variability in fuel composition, combustion conditions, and atmospheric aging makes it challenging to isolate a single or small set of wildfire smoke markers. Isotopic analysis and molecular markers provide further insights into the unique chemical profiles of wildfire smoke. In Italy, researchers successfully utilized the isotopic carbon analysis to detect "wildfire smoke signatures" within the broader $PM_{2.5}$ mixture, enabling a more accurate assessment of wildfire-induced air quality impacts.[69] While this approach is considered by the EPA[42] to be one of the best ways to measure the "true" concentration of wildfire smoke in $PM_{2.5}$, its accuracy depends on characterizing each fire individually due to the substantial variation between fires in terms of fuel composition and other factors. Moreover, this approach is resource intensive, and its applicability to large spatiotemporal datasets remains limited. Additionally, as discussed in ground-based $PM_{2.5}$ monitoring data on ground-based monitoring, sparse data issues pose challenges to this method.

Thresholding for wildfire-specific $PM_{2.5}$. Thresholding techniques rely upon the episodic and acute nature of wildfire events to estimate wildfire-specific $PM_{2.5}$.[70] Thresholding involves setting a certain concentration level of $PM_{2.5}$, above which the $PM_{2.5}$ is deemed to be predominantly from wildfires. By establishing this threshold, researchers can differentiate between baseline $PM_{2.5}$ levels and those attributed to wildfire events. This method enables the identification and quantification of the specific impact of wildfires on air quality, aiding in developing strategies for mitigation and adaptation. We note co-measured $PM_{2.5}$ and CO, and their ratios are often used to indicate wildfire smoke in cities (i.e., Jaffe et al.[71]). However, we do not include that approach in the review due to, in part, the required high data quality and the scarcity of CO monitoring sites compared to $PM_{2.5}$. Below are some of the most common methods used for thresholding.

Air quality index values identifying wildfire periods. The air quality index (AQI) is a widely used simplified method for communicating air quality conditions to the public, transforming complex air quality data into a color-coded scale that indicates whether air quality is considered "healthy" or "unhealthy," based on concentrations of regulated air pollutants, including $PM_{2.5}$ (Table S6). However, it's important to note that the AQI is not uniform across countries and does not always exhibit a linear association with $PM_{2.5}$ levels because AQI also comprises PM_{10}, ozone, sulfur dioxide, nitrogen dioxide, and carbon monoxide, each with its own weighting and thresholds.[72] This inclusion of other pollutants can sometimes obscure the specific contribution of wildfire-generated $PM_{2.5}$ to AQI values, especially when carbon monoxide and other components are elevated due to nearby fires. This blending effect reduces AQI's precision as a tool for assessing wildfire-specific smoke impacts and makes it less suitable for research purposes. For public health communications during wildfire events, using AQI values above a threshold is common because it provides a straightforward way to inform the public about overall air quality conditions and associated health risks. For example, AQI levels may rise dramatically in regions affected by or downwind of wildfires, triggering public health alerts. For instance, in Seeley Lake, Montana, researchers utilized AQI values alongside EPA ground monitor data to establish a correlation between poor air quality and decreased lung function in adults.[73] While the AQI is a valuable tool and effectively conveys general air quality conditions, it has limitations in accurately determining wildfire-specific $PM_{2.5}$ levels due to its lack of specificity, inability to capture spatial and temporal variability, influence of background pollution, and difficulty in interpretation. In contrast, $PM_{2.5}$ concentrations over a threshold offer a more precise method for measuring the impact of wildfire smoke on air quality, especially in research and regulatory contexts. $PM_{2.5}$ measurements focus directly on the fine particulate matter from smoke, providing clearer insights into exposure and health risks. Therefore, while the AQI is useful for public health advisories, $PM_{2.5}$ concentration thresholds are preferred for detailed assessments of wildfire smoke impacts.[72]

Wildfire qualifier codes within monitoring networks. The EPA assigns wildfire-specific qualifier codes to air quality data in its Air Quality System (AQS) based on a combination of multiple data sources and analytical methods.[74] The codes are used to identify when elevated $PM_{2.5}$ levels are influenced by wildfire smoke, distinguishing these events from other pollution sources. The process begins with real-time data from ground-based air quality monitoring stations, where sudden spikes in $PM_{2.5}$ concentrations can indicate a potential wildfire-related event. However, the EPA does not rely solely on this data; it also integrates meteorological information, including wind patterns, temperature, and humidity, which help determine how wildfire smoke is transported and dispersed across regions. In addition to ground-level data, the EPA uses satellite imagery, such as AOD measurements, to track smoke plumes from a broader perspective. These remote sensing data allow for the detection of smoke even in areas with few or no ground monitors. Fire emission inventories further assist by providing detailed records of known wildfire events, including the location, intensity, and duration of fires. This information helps confirm whether elevated $PM_{2.5}$ levels correspond with specific fire events. Air quality models are also utilized to simulate the transport and dispersion of smoke particles. These models integrate data on fire emissions and meteorological conditions to predict where wildfire smoke will likely impact air quality, further informing the decision to assign wildfire-related qualifier codes. The EPA uses these combined data sources to apply specific qualifier codes to indicate when and where air quality data might have been affected by smoke, enabling researchers and policymakers to distinguish between $PM_{2.5}$ pollution caused by wildfires and other sources.[74] Some examples include the following: "W" - Wildfires: elevated PM levels are primarily attributed to wildfires in the area. "FW" - Fire/Wildfires: specifies the presence of PM originating from both structural fires and wildfires. "F" - Fire: while not exclusively for wildfires, denotes the presence of PM from fires, including controlled burns and wildfires. "WF" - Windblown Fugitive Dust/Fire: indicates PM originating from windblown fugitive dust and fires, including wildfires. By leveraging these qualifier codes, analysts can more accurately filter and analyze air quality data to isolate the impacts of wildfire smoke. For example, air quality monitors may detect elevated $PM_{2.5}$ levels during a wildfire event. Qualifier codes help confirm that these spikes are related to wildfire smoke rather than other pollution sources, such as industrial emissions. However, limitations exist within the AQS, data tagging, and qualifier codes. These limitations, as demonstrated in previous work, encompass a potential lack of specificity in accurately distinguishing $PM_{2.5}$ sources and the ineffective capture of spatial and temporal variability, given that wildfire impacts can fluctuate rapidly and locally. Additionally, interpretation challenges may arise due to the complexity of tagging and coding systems, potentially leading to data misinterpretation. Notably, the absence of a wildfire/fire-related qualifier code does not confirm the absence of smoke impact at the site. Marginal impacts may not be reflected by a qualifier code, mainly if the total $PM_{2.5}$ level was not sufficiently high to trigger a NAAQS exceedance and/or if the potential smoke contribution was relatively minor. Table S7 shows a case use example where AQS qualifier codes were used to determine wildfire windows for six Western US states.[70]

Smoke wave days. "Smoke wave days," as first introduced by Liu et al.,[75] is a method that uses time series analysis of air quality monitoring data to spot spikes or unusual patterns in $PM_{2.5}$ levels that align with known wildfire incidents, utilizing daily or hourly data resolutions. This approach differentiates wildfire smoke from other $PM_{2.5}$ sources by identifying patterns and seasonal variations that are characteristic of wildfire emissions, often aligning with wildfire incident records such as location, duration, and intensity. For instance, "smoke wave days" are recognized when $PM_{2.5}$ levels exceed a baseline by a predetermined threshold, indicating a significant rise likely due to wildfire smoke. A common threshold example is setting smoke wave days at levels where $PM_{2.5}$ concentrations surpass the baseline by 10 $\mu g/m^3$ for at least two consecutive days, highlighting extended periods of elevated air pollution due to wildfire smoke.[75] These thresholds allow researchers to focus on periods with substantial deviations from typical air quality. Statistical models and time-series analysis techniques assess air quality data over time, identifying specific periods that meet the criteria for smoke wave days. However, while valuable, the reliance on fixed thresholds and durations may not capture the nuances of each wildfire event, especially given the potential for smoke to impact regions far from the original fire source. Additionally, short-term fluctuations in $PM_{2.5}$ during dynamic fire conditions might be overlooked, emphasizing the need for flexible thresholding in future assessments.

Additional wildfire data. **Satellite observations specific for wildfire screening.** As described in the three basic types of all-source $PM_{2.5}$ estimates, satellite observations significantly contribute to the enhancement of $PM_{2.5}$ estimation efforts. In addition, satellite observations are used specifically to map on-the-ground wildfire detection and monitoring [i.e., active fire detection (AFD)]. Notable examples of these satellite systems (sees Table S4 and S5 for more examples) include Moderate Resolution Imaging Spectroradiometer (MODIS)[76] and Visible Infrared Imaging Radiometer Suite (VIIRS).[77] MODIS excels in identifying active fires and tracking smoke plumes. Similarly, VIIRS detects wildfires and monitors their evolution by observing thermal anomalies and reflected radiation from smoke plumes. Sensors like the Sea and Land Surface Temperature Radiometer (SLSTR)[78] and the Ocean and Land Color Instrument (OLCI)[79] contribute to wildfire detection by identifying thermal anomalies and assessing smoke plume spatial extents. Landsat satellites, primarily dedicated to land imaging, also play a vital role in wildfire monitoring, offering detailed analysis of burn scars and fire-affected areas. On the positive side, these observations excel in identifying active fire locations in real-time, aiding in early detection and rapid response efforts.[80] However, interpreting satellite data for $PM_{2.5}$ estimation poses significant challenges with limitations in spatial resolution and geostatistical considerations, including cloud cover and sensor characteristics, impacting data quality and reliability.[81]

Hazard mapping system. Another satellite-based tool is the Hazard Mapping System (HMS)[82] Smoke Plumes product developed by the National Oceanic and Atmospheric Administration (NOAA), which employs satellite imagery to monitor and trace the pathways of smoke plumes. By aggregating satellite images, HMS Smoke Plumes provide an expansive overview of how smoke disperses over large areas. Although HMS Smoke Plumes do not directly quantify ground-level $PM_{2.5}$ levels, they offer insight into the trajectory and extent of wildfire smoke, information that is important for evaluating the potential impact of smoke on communities located downwind. When integrated with other satellite-derived data, such as AOD values and trace gas detections, HMS Smoke Plumes' insights substantially improve the assessment of wildfires' impact on air quality.[83] However, limitations exist, including the potential for inaccuracies in capturing ground-level $PM_{2.5}$ concentrations due to variations in smoke composition, atmospheric conditions, and other aerosols.[83] Furthermore, cloud cover can obstruct visibility and hinder the accurate detection of smoke plumes.[83] At the same time, coarse spatial resolutions may limit the ability to capture localized variations in $PM_{2.5}$ concentrations near wildfire-affected areas. One example of a methodology paper employing HMS comes from Vargo et al.[84] They employed HMS smoke plume data in combination with air quality and meteorological datasets to estimate the contribution of wildfire smoke to $PM_{2.5}$ concentrations. Vargo et al. focused on understanding the temporal and spatial patterns of wildfire smoke exposure and its effects on $PM_{2.5}$ levels across various regions. Their methodology highlighted how integrating satellite-derived smoke data could improve spatial coverage and provide more accurate estimations of $PM_{2.5}$ exposure, especially in areas lacking ground-based monitoring stations. Another example of a study that used HMS smoke plumes with health results is Aguilera et al.[27] This study utilized HMS smoke plumes to evaluate the impact of wildfire smoke on respiratory health in Southern California, focusing on the years 1999–2012. This research innovatively combined HMS data with other environmental and health datasets to pinpoint how wildfire-related $PM_{2.5}$ affects respiratory hospital admissions. By identifying smoke plume days within a 160 km radius of wildfires, the study distinguished between wildfire-specific and other $PM_{2.5}$ sources, ultimately demonstrating the heightened health risks posed by wildfire smoke.[27]

Wildfire permitter mapping and biomass burning emission inventories. Geospatial techniques, particularly wildfire perimeter mapping, provide additional data sources specific to wildfires, such as proximity measures, intensity of fires, area, and type of vegetation burned. These data can then be used as additional explanatory variables in estimating the contribution of wildfire to overall $PM_{2.5}$. By understanding the extent and location of the fire, researchers and air quality managers can more accurately estimate the amount of particulate matter generated and its potential impact on air quality in surrounding areas. As with all techniques, limitations arise due to variations in fire behavior and environmental conditions, which can affect smoke emissions, complicating $PM_{2.5}$ estimation. Additionally, perimeter mapping alone struggles to differentiate between $PM_{2.5}$ from wildfires and other sources, thus requiring other data to be used alongside.[58] The resolution of mapping data may also limit capturing fine-scale $PM_{2.5}$ variations. Finally, wildfire mapping has led to wildfire smoke emission inventories (Table S8) that have been used in health studies.[85] It should be noted that most of the inventories listed in Table S8 include, or are based on, satellite-mapped burned areas and/or active fire detections. However, the selection of emission inventories could lead to drastically different outcomes for estimated wildfire-induced $PM_{2.5}$ concentrations, with disparities of up to 20-fold for specific fires/cases.[86] Another study found that different fire emissions inventories can yield significantly varying outcomes, largely due to differences in methods for calculating factors like burned area, emissions, and fuel loading.[87] In addition, recent aircraft observations provided top-down constraints, suggesting these inventories could underestimate regional fire emissions by a factor of 2 or more.[88,89]

Publicly Available Wildfire-Specific $PM_{2.5}$ Datasets, Comparisons, and Discussion

Publicly and easily accessible $PM_{2.5}$ datasets offer invaluable resources for researchers, policymakers, and the public to enhance our understanding of air quality, evaluate health risks, and support informed decision-making. Next, we used the three papers mentioned above, Aguilera et al.,[39] Childs et al.,[6] and Zhang et al.,[40] in our comparative analysis to assess consistency and variability in $PM_{2.5}$ estimates across different geographic and environmental contexts. We used the smallest common spatial scale—referred to as the least common spatial denominator—for California and in 2010 and 2018, years with low and high fire activity, respectively. Our goal was to explore how different datasets vary in estimating wildfire-specific $PM_{2.5}$ exposure.

Description of the three wildlife-specific $PM_{2.5}$ datasets from the literature. The Aguilera et al.[39] model used an ensemble-based statistical approach that combined multiple machine learning algorithms to estimate daily wildfire-specific $PM_{2.5}$ in California. The model integrated a comprehensive set of explanatory variables, including satellite-derived AOD, meteorological data like temperature, and land-use variables. It also utilized geographical information, such as basin delineations, to refine its predictions based on local topography and meteorological conditions. The approach of Childs et al.[6] focused on a broader geographic scale, providing 10-km resolution estimates across the contiguous United States. This model used satellite data to track smoke plumes and employed hybrid single-particle Lagrangian integrated trajectory (HYSPLIT) simulations from identified fire locations, coupled with machine learning techniques to predict daily $PM_{2.5}$ concentrations. While this approach also incorporates meteorological data and AOD from satellites, it does not utilize local geographic features such as basins, focusing instead on a broader application suitable for a national scale. The Zhang et al.[40] method incorporated both remote sensing and ground-based monitoring data to estimate $PM_{2.5}$ concentrations from wildfires and other sources. This model also differentiated between smoke-derived $PM_{2.5}$ and other types of particulate matter. Each of these models illustrates the critical role of explanatory variables in shaping the estimates of wildfire-specific $PM_{2.5}$. Aguilera et al.'s model was particularly detailed in its use of local data, such as temperature and basin information, to enhance accuracy in a specific region. In contrast, Childs et al. and Zhang et al. expand the geographical focus but vary in their use of data, with Childs et al. leveraging HYSPLIT trajectory data for a nationwide application and Zhang et al. emphasizing the differentiation of smoke impacts from other pollutants for comprehensive air quality modeling.

Processing comparative datasets. We processed all data at the ZIP code level to compare EPA observations[42] with estimates from the Aguilera et al.,[39] Childs et al.,[6] and Zhang et al.[40] datasets. The comparison focused on Modoc and Los Angeles counties representing areas with stark differences in monitoring availability but both experiencing wildfire activity during the study years. Modoc County, located in the northeastern corner of California, has no EPA ground monitors[42] and a low population of 9,686 (2010).[43] In contrast, Los Angeles County in southern California has the highest concentration of monitors (20) and a dense population of $\sim$9.8 million (2010).[43] Since Los Angeles had multiple monitors, we averaged $PM_{2.5}$ values across the 20 EPA monitors to create a single county-level estimate for comparison with modeled datasets. This aggregation allowed us to focus on broader patterns rather than site-specific variability within the county. This setup allowed us to assess how model estimates perform in regions with no ground-based $PM_{2.5}$ monitoring compared to regions with abundant monitoring, an important question for assessing model generalizability to data-sparse areas. We chose this comparison deliberately to explore how models estimate wildfire-specific $PM_{2.5}$ under contrasting conditions of data availability. Although wildfire behavior, urbanization, and background PM sources may differ between the two counties, both experienced wildfire smoke exposure, and this comparison highlights the potential strengths and weaknesses of each model across varying infrastructure and population density settings. EPA monitors are typically placed in higher-population or high-pollution areas,[42] meaning their data may inherently reflect higher $PM_{2.5}$ levels. We acknowledge that differences in local wildfire characteristics and background pollution may introduce some confounding, which represents a limitation of this analysis. To examine differences across wildfire activity levels, we selected two contrasting years. We analyzed California in 2010 (low wildfire activity: 6,554 wildfires that burned 109,529 acres)[44] and 2018 (high wildfire activity: 8,527 wildfires that burned 1.8 million acres, resulting in 103 fatalities).[45] These contrasting years allowed us to explore how different datasets estimate wildfire-specific and overall $PM_{2.5}$ during periods of varying fire intensity. Tables 2–4 provide a detailed summary of all combinations of datasets (overall vs. wildfire-specific $PM_{2.5}$) across years (2010 vs. 2018) and locations (statewide and two counties).

Summary statistics of wildfire-specific $PM_{2.5}$ data for the state of California, USA. Figure 3 shows the mapped annual mean $PM_{2.5}$ values across California for the three modeled wildfire-specific $PM_{2.5}$ datasets for years 2010 and 2018, including the observed overall $PM_{2.5}$ values from the EPA. Tables 2–4 describe the summary statistics for all combinations of year, region, and overall vs. wildfire-specific $PM_{2.5}$. In general, the statistics for 2010 were lower than 2018, with the modeled results generally underestimating EPA's measured values. When comparing the wildfire-specific

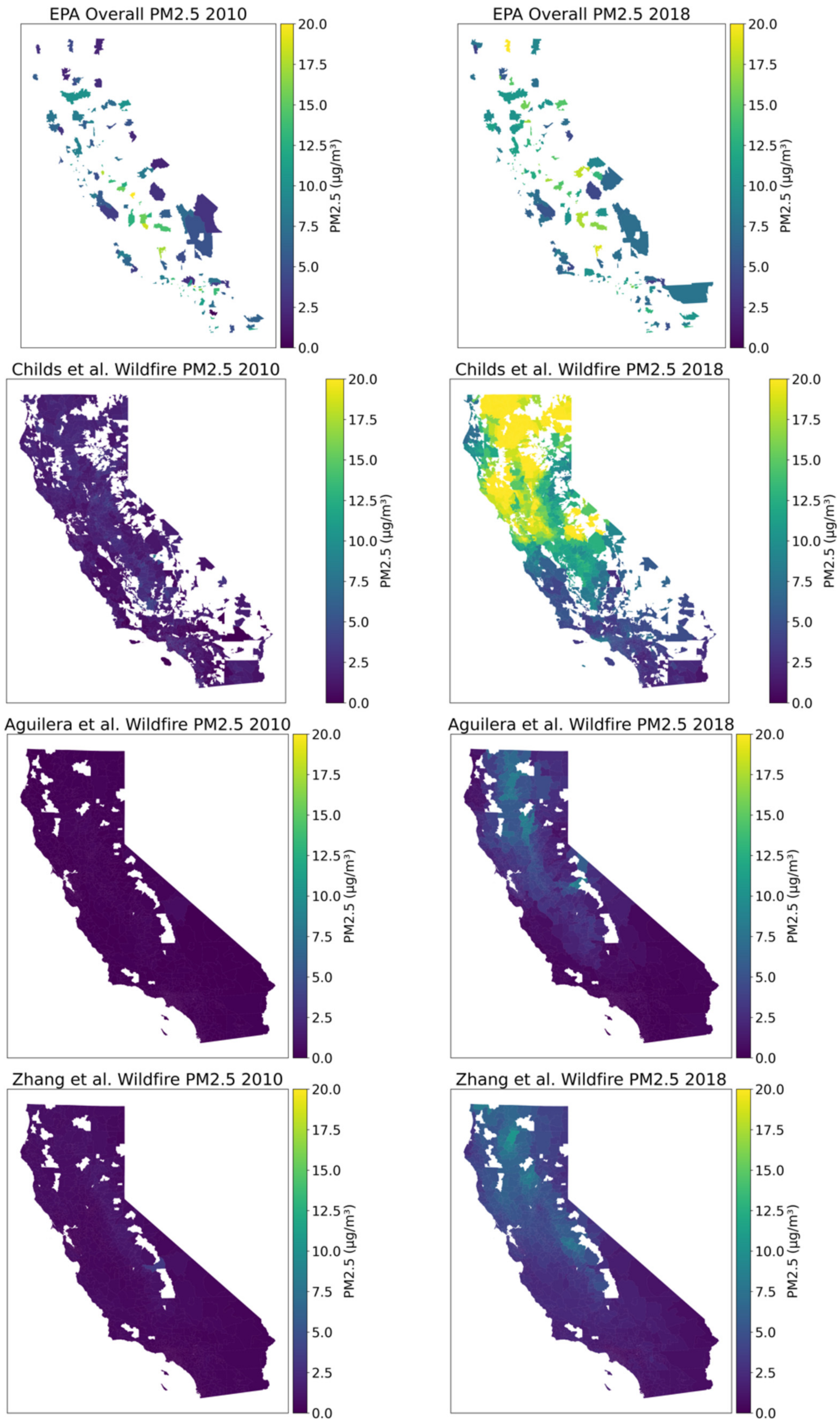

Figure 3. Predicted wildfire-specific $PM_{2.5}$ datasets for California during the years of 2010 (low wildfire year, left column) and 2018 (one of the highest wildfire years, right column). The top row shows the EPA monitoring stations aggregated to the mean for each ZIP code and reflects all-source $PM_{2.5}$ levels as a reference dataset. All wildfire-specific $PM_{2.5}$ estimates are summarized to the mean values for each ZIP code. Note: Each dataset, EPA,[42] Aguilera et al.,[27] Childs et al.,[6] and Zhang et al.,[40] is presented at its obtained scales with white values for missing ZIP code data. Graphs created using python.[90] $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 µm.

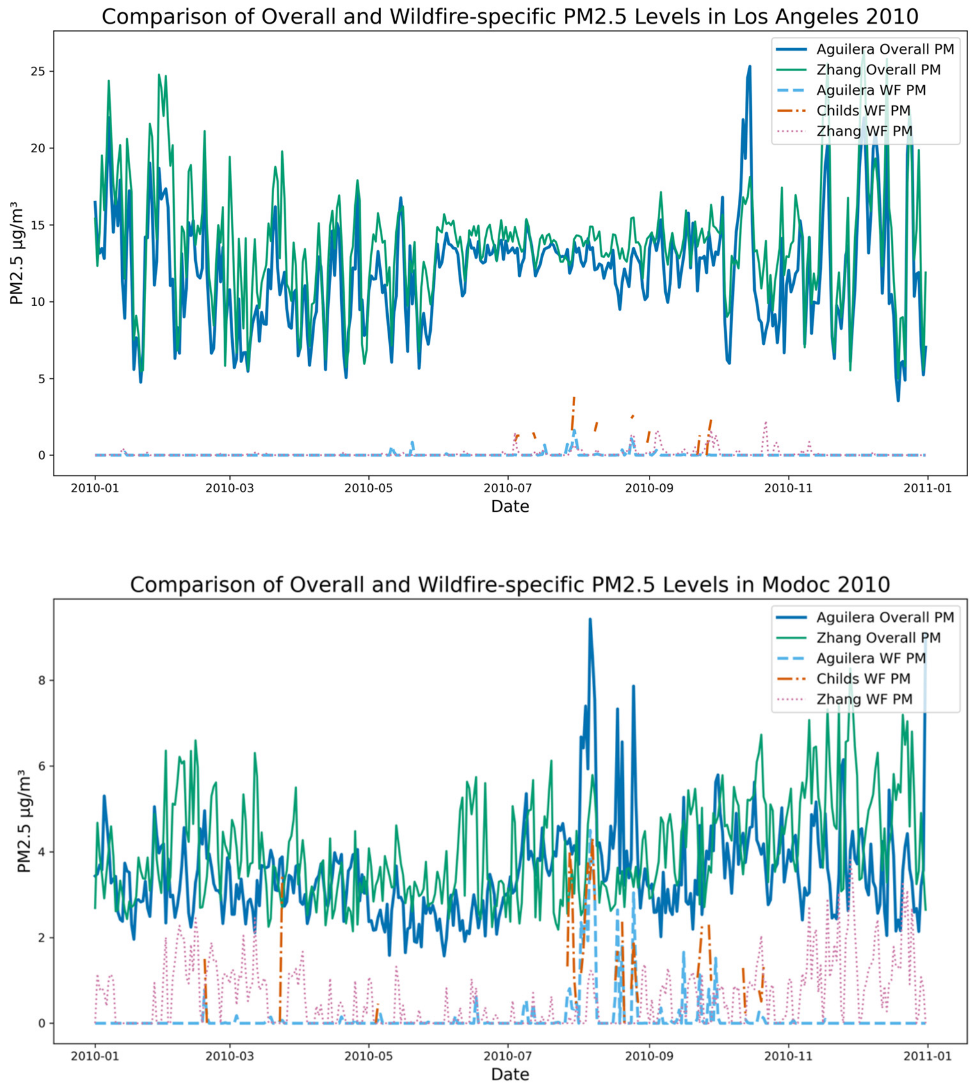

Figure 4. Graphical representation of ambient and wildfire-specific $PM_{2.5}$ estimates for the year 2010 comparing an area with (Los Angeles) and without a ground monitor (Modoc). Solid lines represent the ambient estimates, and dashed lines show the wildfire-specific estimates. Los Angeles County also shows the EPA monitoring sites average within that county. No EPA data exist for Modoc County. Data sources: EPA,[42] Aguilera et al.,[27] Childs et al.,[6] and Zhang et al.[40] Graph generated using python matplotlib.[90] Note: $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 µm.

$PM_{2.5}$ contribution to overall $PM_{2.5}$ in 2010 and 2018, the median values showed that wildfire-specific $PM_{2.5}$ comprised between 11.2% and 36.9% per Aguilera et al. and between 13.7% and 21.2% per Zhang et al. of the total annual $PM_{2.5}$. As expected, $PM_{2.5}$ estimates for all datasets were higher in 2018, a year marked by significant wildfire activity in northern California (Ranch Fire, 410,203 acres; Carr Fire, 229,651 acres; Camp Fire, 153,336 acres), aligning with the highest concentrations shown in northern California[91] (Figure 3). Additionally, more days exceeding recommended $PM_{2.5}$ standards were recorded in 2018 across all datasets, especially in Los Angeles County (see Table 4).

Comparative analysis for wildfire-specific $PM_{2.5}$ data for two counties in California, USA. The next comparison focused on two counties: Modoc, situated in the northeastern corner of California, devoid of EPA ground monitors,[42] with a population of 9,686 in 2010[43] and 538 miles away (measured from the edge of counties) from southern Los Angeles County, which boasts the highest concentration of monitors (20)[42] as well as a population of 9.8 million in 2010.[43] Tables 2–4 show the summary statistics for each dataset scenario, and Figures 4 and 5 plot their time series. In general, the trends observed for the entire state of California, such as differences in overall $PM_{2.5}$ and wildfire-specific $PM_{2.5}$ estimates, are also reflected in these two more local regions; however, some spatial differences were seen between these counties. First and not surprisingly, Los Angeles County showed the highest levels of overall $PM_{2.5}$ as compared to Modoc County, most likely a reflection of the higher concentration of people and mobile and industrial sources. Second, Modoc County observed higher ratios of wildfire-specific to overall $PM_{2.5}$, further pointing to the influence of the human population in $PM_{2.5}$ makeup. Next, we discuss the results from the cross-correlation and Granger causality tests that compared the relationships between datasets, including overall $PM_{2.5}$ levels (Aguilera et al., EPA, and Zhang et al.) and wildfire-specific $PM_{2.5}$ levels (Aguilera et al., Zhang et al., and Childs et al.) (see Table 5; Table S2).

Cross-correlation tests. As shown in Table 5, the correlation between the datasets improved markedly from 2010 to 2018, reflecting an increase in alignment between model estimates and EPA observations over time. In 2018, Los Angeles County, with its dense population and numerous EPA ground monitors, exhibited stronger correlations between the EPA $PM_{2.5}$ data and

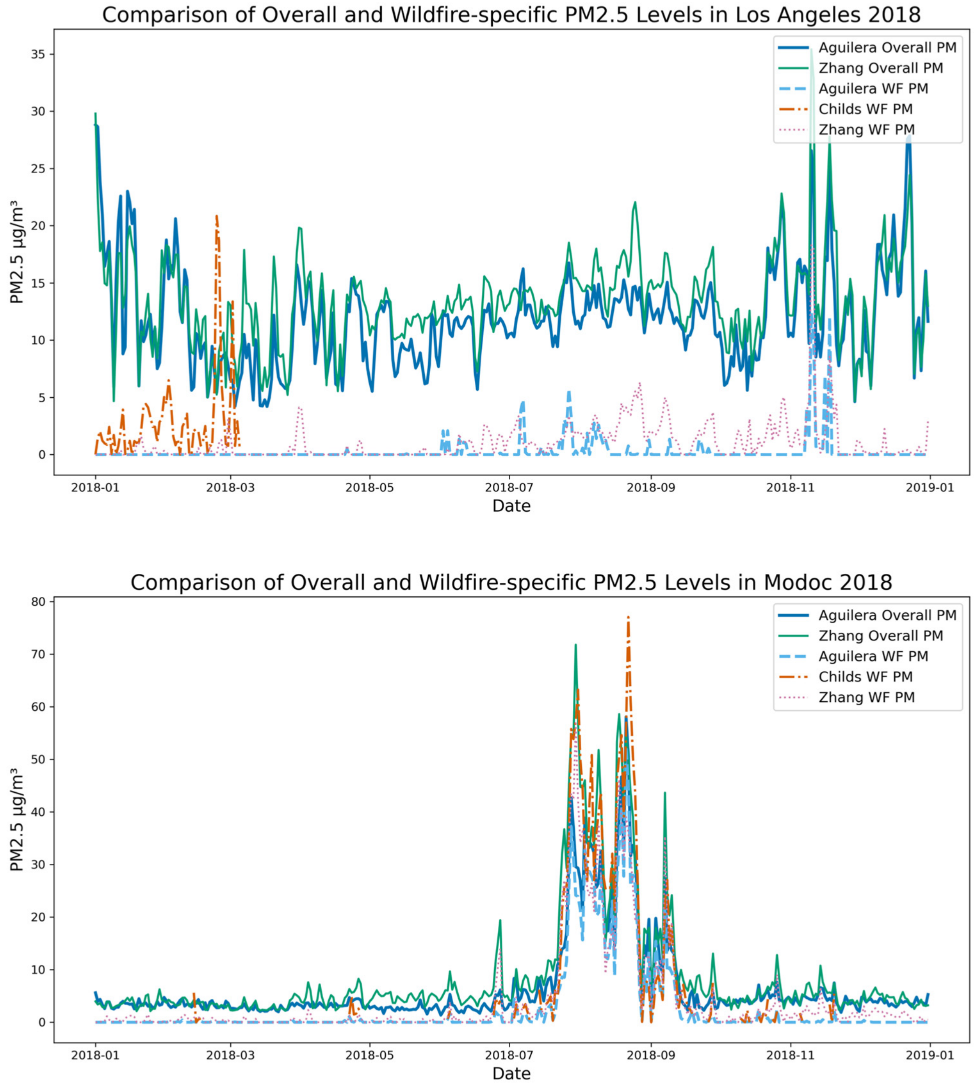

Figure 5. Graphical representation of ambient and wildfire-specific $PM_{2.5}$ estimates for the year 2018 comparing an area with (Los Angeles) and without a ground monitor (Modoc). Solid lines represent the ambient estimates, and dashed lines show the wildfire-specific estimates. Los Angeles County also shows the EPA monitoring sites average within that county. No EPA data exist for Modoc County. Data sources: EPA,[42] Aguilera et al.,[27] Childs et al.,[6] and Zhang et al.[40] Graph generated using python matplotlib.[90] Note: $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 µm.

wildfire-specific $PM_{2.5}$ datasets (Aguilera et al. and Zhang et al.) compared to 2010. The highest correlations were seen in comparisons between EPA and Zhang et al. $PM_{2.5}$ data (0.80), as well as between Aguilera et al. wildfire-specific $PM_{2.5}$ and Childs et al. wildfire-specific $PM_{2.5}$ (0.69). This improvement in correlations suggests that the models became more accurate in predicting both overall and wildfire-specific $PM_{2.5}$ levels as data quality and modeling techniques improved. Conversely, Modoc County, a rural and sparsely monitored region, showed different trends. While there was no EPA data available for direct comparison, the correlations between modeled datasets were notably high in 2018. For example, the correlation between Aguilera et al. and Zhang et al. $PM_{2.5}$ datasets in Modoc increased from 0.35 in 2010 to 0.87 in 2018. Similarly, the correlation between Aguilera et al. wildfire-specific $PM_{2.5}$ and Childs et al. wildfire-specific $PM_{2.5}$ rose to 0.79. This suggests that despite the absence of

Table 5. Pearson correlation tests between datasets compared in this study for years 2010 and 2018.

	EPA $PM_{2.5}$[42] vs. Aguilera et al. $PM_{2.5}$[27]	EPA $PM_{2.5}$[42] vs. Zhang et al. $PM_{2.5}$[40]	Aguilera et al. $PM_{2.5}$[27] vs. Zhang et al. $PM_{2.5}$[40]	Aguilera et al. WF $PM_{2.5}$[27] vs. Zhang et al. WF $PM_{2.5}$[40]	Aguilera et al. WF[27] $PM_{2.5}$[27] vs. Childs et al. WF $PM_{2.5}$[6]	Childs et al. WF $PM_{2.5}$[6] vs. Zhang et al. WF $PM_{2.5}$[40]
California 2010	0.76	0.68	0.70	0.29	0.46	0.51
California 2018	0.89	0.87	0.88	0.89	0.90	0.91
Los Angeles 2010	0.74	0.64	0.74	0.19	0.31	0.28
Los Angeles 2018	0.79	0.80	0.79	0.54	0.69	0.79
Modoc 2010	NA	NA	0.35	0.45	0.44	0.51
Modoc 2018	NA	NA	0.87	0.86	0.79	0.88

Note: Results displayed statewide for California and for the two counties of Los Angeles and Modoc. Gray boxes assess association between the ambient $PM_{2.5}$ datasets, including the EPA ground monitors, and white boxes assess the wildfire-specific (WF) $PM_{2.5}$ datasets. Information obtained from the Environmental Protection Agency (EPA)[6,22,39,40] and World Health Organization (WHO). IQR, interquartile range; NA, data unavailable; $PM_{2.5}$, fine particulate matter with an aerodynamic diameter <2.5 µm; WF, wildfire.

ground-based monitoring, the models in Modoc were more consistent with each other in predicting wildfire-specific $PM_{2.5}$ in 2018. Overall, the data show that correlations between $PM_{2.5}$ datasets, especially wildfire-specific estimates, were stronger in 2018 than in 2010 across both urban (Los Angeles) and rural (Modoc) areas. This suggests that model improvements and increased wildfire activity in 2018 led to better agreement across datasets.

Granger causality tests (Table S2). The Granger (causality) test is an alternative time-series correlation tool that determines whether one time series can predict future values of another time series rather than simply assessing their correlation at the same point in time as shown in Table 5. Unlike simple correlation, which only measures the strength of the relationship between two variables, Granger causality tells us if one dataset (the predictor) contains information that helps forecast another dataset (the response variable). This is particularly useful in the context of wildfire-specific $PM_{2.5}$ and general $PM_{2.5}$ estimates, where time-lagged effects might exist due to the delayed transport and dispersion of wildfire smoke. Table S2 shows Granger causality results for California and two counties (Los Angeles and Modoc) in 2010 and 2018. The tests reveal stronger predictive power in 2018 compared to 2010. For instance, Aguilera et al. wildfire-specific $PM_{2.5}$ data significantly predicted Zhang et al. wildfire-specific $PM_{2.5}$ in California, with F-statistics increasing from 6,999.67 ($p < 0.0001$) in 2010 to 36,771.30 ($p < 0.0001$) in 2018, reflecting improvements in data accuracy during a more active wildfire year. In Los Angeles, Aguilera et al. $PM_{2.5}$ data showed strong predictive capability for EPA $PM_{2.5}$, with F-statistics ranging from 46.56 ($p < 0.0001$) in 2010 to 17.63 ($p < 0.0001$) in 2018. In Modoc County, where no EPA monitors exist, the predictive power between modeled datasets also strengthened, with Aguilera et al. $PM_{2.5}$ predicting Zhang et al. $PM_{2.5}$ in 2018 ($F = 560.42$; $p < 0.0001$) compared to 2010 ($F = 226.03$; $p < 0.0001$). Though most pairings showed significant Granger causality, some exceptions occurred, such as Aguilera et al. not consistently predicting Childs et al. wildfire-specific $PM_{2.5}$ in Modoc in 2010. Overall, Granger causality tests highlight improved predictive power in 2018, especially for wildfire-specific datasets, reflecting how increased wildfire activity and better modeling enhanced the dataset's forecasting ability. This is crucial for understanding the evolving impact of wildfire smoke on air quality.

Conclusions

As the prevalence of wildfires continues to rise, impacting broader regions across the US and other parts of the world, the development of wildfire-specific $PM_{2.5}$ modeling has become an area of keen research interest. Our comprehensive literature review reveals that, although the availability of publicly accessible wildfire-specific $PM_{2.5}$ datasets are limited, significant strides have been made in the field. Despite the abundance of wildfire-specific $PM_{2.5}$ exposure modeling studies, very few of these data are made publicly accessible. Only 4 out of the 33 studies included in this review had publicly available data. We found this to be surprising given the large body of literature on the principles of data sharing and reproducibility. For example, in 2016, "The FAIR guiding principles for scientific data management and stewardship" were published in *Scientific Data*.[92] The authors intended to provide guidelines to improve the findability, accessibility, interoperability, and reuse of digital assets. Indeed, open-sourced data are vital for advancing research, driving innovation and enhancing transparency across scientific fields. They allow for the replication of studies, validation of findings, and exploration of new hypotheses without the limitations of proprietary datasets. Open-sourced datasets are indispensable, particularly in environmental research, such as studies on air quality and wildfire impacts on $PM_{2.5}$ levels. In our study, we prioritized datasets that were readily accessible (CSV format) and easily standardized. However, due to inaccessibility or incompatibility, we could only include three datasets. Proprietary or restricted-access data, though potentially valuable, were not utilized due to the complexities of extracting and harmonizing them with other sources. Relying on open-sourced data is crucial for ensuring equity and sustainability in research practices. Moving forward, cultivating an environment of freely shared data is essential for effectively addressing global environmental issues. Additionally, it's important to note that while MODIS data has been a significant source for wildfire-related studies, its discontinuation by NASA in February 2023 poses a new challenge, as this data is no longer available, emphasizing the urgency of transitioning to alternative data sources and promoting open data initiatives to sustain research efforts in this critical area.

Overall, in the comparison analyses, wildfire-specific datasets were most similar when more and large wildfires were present, suggesting better model agreement and fitting with more wildfire data or possibly improvement in monitoring over time. We also observed lower dataset agreement in a highly urban area with many monitoring sites. This might suggest that urban complexity, including diverse pollution sources and potential $PM_{2.5}$ variability, may introduce challenges in achieving dataset concordance at high resolution. Because essential model features and simplifying assumptions will vary from study to study, it would not be surprising that the available wildfire-specific $PM_{2.5}$ exposure models focus on different processes and represent environmental heterogeneity in different ways. However, as wildfire-specific $PM_{2.5}$ models and datasets begin to proliferate, health researchers have a unique opportunity to explore intermodal comparisons, which allows further insights and possibly more robust conclusions. It is quite possible that two different wildfire-specific $PM_{2.5}$ datasets will lead to complementary answers about certain research questions (e.g., health outcome risk estimates[93]), thus allowing insights about the influences of model structure, parameters, or other assumptions. We, therefore, emphasize the importance of continuous methodological advancements and the adoption of a multifaceted approach to air quality assessment, as well as intermodal comparisons for environmental health assessments. This is crucial for improving our understanding of environmental health impacts and developing robust public health strategies to mitigate the effects of wildfire smoke.

Acknowledgments

Conceptualization: A.O., E.L.L., and C.N. Data curation: A.O. Methodology: A.O., E.L.L., and J.G. Formal analysis: A.O., C.L., and S.U. Funding acquisition: C.N., E.L.L., E.O.S., and J.G. Project Administration: E.L.L. Writing–original draft: A.O. Writing–review & editing: All.

The authors would like to thank the editors and reviewers.

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health, United States (award numbers P20GM130418).

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

All code to replicate these analyses provided upon reasonable request.

References

1. Burke M, Childs ML, de la Cuesta B, Qiu M, Li J, Gould CF, et al. 2023. The contribution of wildfire to $PM_{2.5}$ trends in the USA. Nature 622(7984):761–766, PMID: 37730996, https://doi.org/10.1038/s41586-023-06522-6.

2. Westerling AL. 2016. Increasing Western US Forest wildfire activity: sensitivity to changes in the timing of spring. Philos Trans R Soc Lond B Biol Sci 371(1696):20150178, PMID: 27216510, https://doi.org/10.1098/rstb.2015.0178.
3. Ascoli D, Moris J, Marchetti M, Sallustio L. 2021. Land use change towards forests and wooded land correlates with large and frequent wildfires in Italy. Ann Silvicultural Res 46:177–188, https://doi.org/10.12899/asr-2264.
4. Borchering RK, Gunning CE, Gokhale DV, Weedop KB, Saeidpour A, Brett TS, et al. 2021. Anomalous influenza seasonality in the United States and the emergence of novel influenza B viruses. Proc Natl Acad Sci USA 118(5):e2012327118, PMID: 33495348, https://doi.org/10.1073/pnas.2012327118.
5. Ford B, Val Martin M, Zelasky SE, Fischer EV, Anenberg SC, Heald CL, et al. 2018. Future fire impacts on smoke concentrations, visibility, and health in the contiguous United States. GeoHealth 2(8):229–247, PMID: 32159016, https://doi.org/10.1029/2018GH000144.
6. Childs ML, Li J, Wen J, Heft-Neal S, Driscoll A, Wang S, et al. 2022. Daily local-level estimates of ambient wildfire smoke $PM_{2.5}$ for the contiguous US. Environ Sci Technol 56(19):13607–13621, PMID: 36134580, https://doi.org/10.1021/acs.est.2c02934.
7. Vaidyanathan A, Saha S, Vicedo-Cabrera AM, Gasparrini A, Abdurehman N, Jordan R, et al. 2019. Assessment of extreme heat and hospitalizations to inform early warning systems. Proc Natl Acad Sci USA 116(12):5420–5427, PMID: 30833395, https://doi.org/10.1073/pnas.1806393116.
8. O'Dell K, Bilsback K, Ford B, Martenies SE, Magzamen S, Fischer EV, et al. 2021. Estimated mortality and morbidity attributable to smoke plumes in the United States: not just a Western US problem. GeoHealth 5(9):e2021GH000457, PMID: 34504989, https://doi.org/10.1029/2021GH000457.
9. Chen H, Samet JM, Bromberg PA, Tong H. 2021. Cardiovascular health impacts of wildfire smoke exposure. Part Fibre Toxicol 18(1):2, PMID: 33413506, https://doi.org/10.1186/s12989-020-00394-8.
10. Cromar K, Gladson L, Gohlke J, Li Y, Tong D, Ewart G. 2024. Adverse health impacts of outdoor air pollution, including from wildland fires, in the United States: "health of the air," 2018–2020. Ann Am Thorac Soc 21(1):76–87, PMID: 37906164, https://doi.org/10.1513/AnnalsATS.202305-455OC.
11. Williams VA, Perreault LR, Yazbeck CT, Micovic NA, Oakes JM, Bellini C. 2024. Impact of wildfires on cardiovascular health. Circ Res 134(9):1061–1082, PMID: 38662865, https://doi.org/10.1161/CIRCRESAHA.124.323614.
12. Amjad S, Chojecki D, Osornio-Vargas A, Ospina MB. 2021. Wildfire exposure during pregnancy and the risk of adverse birth outcomes: a systematic review. Environ Int 156:106644, PMID: 34030071, https://doi.org/10.1016/j.envint.2021.106644.
13. Eisenman DP, Galway LP. 2022. The mental health and well-being effects of wildfire smoke: a scoping review. BMC Public Health 22(1):2274, PMID: 36471306, https://doi.org/10.1186/s12889-022-14662-z.
14. Black C, Tesfaigzi Y, Bassein JA, Miller LA. 2017. Wildfire smoke exposure and human health: significant gaps in research for a growing public health issue. Environ Toxicol Pharmacol 55:186–195, PMID: 28892756, https://doi.org/10.1016/j.etap.2017.08.022.
15. Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT. 2016. Critical review of health impacts of wildfire smoke exposure. Environ Health Perspect 124(9):1334–1343, PMID: 27082891, https://doi.org/10.1289/ehp.1409277.
16. Sun Q, Miao C, Hanel M, Borthwick AGL, Duan Q, Ji D, et al. 2019. Global heat stress on health, wildfires, and agricultural crops under different levels of climate warming. Environ Int 128:125–136, PMID: 31048130, https://doi.org/10.1016/j.envint.2019.04.025.
17. Kuehn BM. 2014. WHO: more than 7 million air pollution deaths each year. JAMA 311(15):1486, PMID: 24737355, https://doi.org/10.1001/jama.2014.4031.
18. Azimi P, Stephens B. 2020. A framework for estimating the US mortality burden of fine particulate matter exposure attributable to indoor and outdoor microenvironments. J Expo Sci Environ Epidemiol 30(2):271–284, PMID: 30518794, https://doi.org/10.1038/s41370-018-0103-4.
19. Schuller A, Montrose L. 2020. Influence of woodsmoke exposure on molecular mechanisms underlying Alzheimer's disease: existing literature and gaps in our understanding. Epigenet Insights 13:2516865720954873, PMID: 32974607, https://doi.org/10.1177/2516865720954873.
20. Weuve J, Bennett EE, Ranker L, Gianattasio KZ, Pedde M, Adar SD, et al. 2021. Exposure to air pollution in relation to risk of dementia and related outcomes: an updated systematic review of the epidemiological literature. Environ Health Perspect 129(9):096001, PMID: 34558969, https://doi.org/10.1289/EHP8716.
21. Wilker EH, Osman M, Weisskopf MG. 2023. Ambient air pollution and clinical dementia: systematic review and meta-analysis. BMJ 381:e071620, PMID: 37019461, https://doi.org/10.1136/bmj-2022-071620.
22. EPA (US Environmental Protection Agency). 2022. *US EPA Air Quality System*. https://www.epa.gov/aqs [accessed 12 December 2024].
23. EPA. 2024. *National Ambient Air Quality Standards (NAAQS) for PM*. https://www.epa.gov/pm-pollution/national-ambient-air-quality-standards-naaqs-pm [accessed 12 December 2024].
24. WHO (World Health Organization). 2021. *Air Pollution*. https://www.who.int/health-topics/air-pollution [accessed 12 December 2024].
25. U.S. Congress. 1965. Text of public law 89–272: the clean air act amendment. J Air Pollut Contr Assoc 15(12):597–600, https://doi.org/10.1080/00022470.1965.10468431.
26. EPA. 2024. *$PM_{2.5}$ Wildland Fire Exceptional Events Tiering Document*. EPA-457/R-24-001. Research Triangle Park, NC: EPA.
27. Aguilera R, Corringham T, Gershunov A, Benmarhnia T. 2021. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. Nat Commun 12(1):1493, PMID: 33674571, https://doi.org/10.1038/s41467-021-21708-0.
28. Peng RD, Bell ML, Geyh AS, McDermott A, Zeger SL, Samet JM, et al. 2009. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ Health Perspect 117(6):957–963, PMID: 19590690, https://doi.org/10.1289/ehp.0800185.
29. Liu JC, Peng RD. 2019. The impact of wildfire smoke on compositions of fine particulate matter by ecoregion in the Western US. J Expo Sci Environ Epidemiol 29(6):765–776, PMID: 30185941, https://doi.org/10.1038/s41370-018-0064-7.
30. Kim YH, Warren SH, Krantz QT, King C, Jaskot R, Preston WT, et al. 2018. Mutagenicity and lung toxicity of smoldering vs. flaming emissions from various biomass fuels: implications for health effects from wildland fires. Environ Health Perspect 126(1):017011, PMID: 29373863, https://doi.org/10.1289/EHP2200.
31. Wegesser TC, Pinkerton KE, Last JA. 2009. California wildfires of 2008: coarse and fine particulate matter toxicity. Environ Health Perspect 117(6):893–897, PMID: 19590679, https://doi.org/10.1289/ehp.0800166.
32. Liu X, Huey LG, Yokelson RJ, Selimovic V, Simpson IJ, Müller M, et al. 2017. Airborne measurements of Western US wildfire emissions: comparison with prescribed burning and air quality implications. JGR Atmos 122(11):6108–6129, https://doi.org/10.1002/2016JD026315.
33. Na K, Cocker DR. 2008. Fine organic particle, formaldehyde, acetaldehyde concentrations under and after the influence of fire activity in the atmosphere of riverside, California. Environ Res 108(1):7–14, PMID: 18538316, https://doi.org/10.1016/j.envres.2008.04.004.
34. Migliaccio CT, Kobos E, King QO, Porter V, Jessop F, Ward T. 2013. Adverse effects of wood smoke PM(2.5) exposure on macrophage functions. Inhal Toxicol 25(2):67–76, PMID: 23363038, https://doi.org/10.3109/08958378.2012.756086.
35. Liu JC, Wilson A, Mickley LJ, Dominici F, Ebisu K, Wang Y, et al. 2017. Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. Epidemiology 28(1):77–85, PMID: 27648592, https://doi.org/10.1097/EDE.0000000000000556.
36. Heaney A, Stowell JD, Liu JC, Basu R, Marlier M, Kinney P. 2022. Impacts of fine particulate matter from wildfire smoke on respiratory and cardiovascular health in California. GeoHealth 6(6):e2021GH000578, PMID: 35795228, https://doi.org/10.1029/2021GH000578.
37. Moeltner K, Kim MK, Zhu E, Yang W. 2013. Wildfire smoke and health impacts: a closer look at fire attributes and their marginal effects. J Environ Econ Manage 66(3):476–496, https://doi.org/10.1016/j.jeem.2013.09.004.
38. Stowell JD, Geng G, Saikawa E, Chang HH, Fu J, Yang C-E, et al. 2019. Associations of wildfire smoke $PM_{2.5}$ exposure with cardiorespiratory events in Colorado 2011–2014. Environ Int 133(pt A):105151, PMID: 31520956, https://doi.org/10.1016/j.envint.2019.105151.
39. Aguilera R, Luo N, Basu R, Wu J, Clemesha R, Gershunov A, et al. 2023. A novel ensemble-based statistical approach to estimate daily wildfire-specific $PM_{2.5}$ in California (2006–2020). Environ Int 171:107719, PMID: 36592523, https://doi.org/10.1016/j.envint.2022.107719.
40. Zhang D, Wang W, Xi Y, Bi J, Hang Y, Zhu Q, et al. 2023. Wildland fires worsened population exposure to $PM_{2.5}$ pollution in the contiguous United States. Environ Sci Technol 57(48):19990–19998, PMID: 37943716, https://doi.org/10.1021/acs.est.3c05143.
41. Reid CE, Considine EM, Maestas MM, Li G. 2021. Daily $PM_{2.5}$ concentration estimates by county, ZIP code, and census tract in 11 Western states 2008–2018. Sci Data 8(1):112, PMID: 33875665, https://doi.org/10.1038/s41597-021-00891-1.
42. EPA. 2025. Outdoor Air Quality Data. https://www.epa.gov/outdoor-air-quality-data [accessed 23 March 2025].
43. US Census Bureau. 2021. County Population Totals: 2010–2020. https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-counties-total.html [accessed 12 December 2024].
44. National Interagency Fire Center. 2015. National Report of Wildland Fires and Acres Burned by State 2010. https://www.nifc.gov/sites/default/files/NICC/2-Predictive%20Services/Intelligence/Annual%20Reports/2010/annual_report_2010_508.pdf [accessed 12 December 2024].
45. National Interagency Fire Center 2024. National Year-to-Date Report on Fires and Acres Burned. https://www.nifc.gov/sites/default/files/NICC/2-Predictive%20Services/Intelligence/Annual%20Reports/2024/annual_report_2024.pdf [accessed 12 December 2024].
46. Matteo F, Angela M. 2022. A bootstrap method to detect test Granger-causalities in the frequency domain. Computational Economics 59(3): 935–966, https://doi.org/10.1007/s10614-021-10112-x.

47. Dockery DW, Pope CA. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15(1):107–132, PMID: 8054077, https://doi.org/10.1146/annurev.pu.15.050194.000543.
48. Kilpatrick DJ, Hung P, Crouch E, Self S, Cothran J, Porter DE, et al. 2024. Geographic variations in urban-rural particulate matter ($PM_{2.5}$) Concentrations in the United States, 2010–2019. GeoHealth 8(9):e2023GH000920, PMID: 39234600, https://doi.org/10.1029/2023GH000920.
49. EPA. 2021. *Comparative Assessment of the Impacts of Prescribed Fire Versus Wildfire (CAIF): A Case Study in the Western U.S.* EPA/600/R-21/197. Washington, DC: EPA.
50. Chow JC, Doraiswamy P, Watson JG, Chen LW, Ho SS, Sodeman DA. 2008. Advances in integrated and continuous measurements for particle mass and chemical composition. J Air Waste Manag Assoc 58(2):141–163, PMID: 18318335, https://doi.org/10.3155/1047-3289.58.2.141.
51. Allen G, Sioutas C, Koutrakis P, Reiss R, Lurmann FW, Roberts PT. 1997. Evaluation of the TEOM method for measurement of ambient particulate mass in urban areas. J Air Waste Manag Assoc 47(6):682–689, PMID: 9180064, https://doi.org/10.1080/10473289.1997.10463923.
52. deSouza P, Kahn R, Stockman T, Obermann W, Crawford B, Wang A, et al. 2022. Calibrating networks of low-cost air quality sensors. Atmos Meas Tech 15(21):6309–6328, https://doi.org/10.5194/amt-15-6309-2022.
53. NOAA (National Oceanic and Atmospheric Administration) Global Monitoring Laboratory. 2025. SURFRAD Aerosol Optical Depth. https://gml.noaa.gov/grad/surfrad/aod/ [accessed 25 February 2025].
54. NASA Earth Observatory. Aerosol Optical Depth. https://earthobservatory.nasa.gov/global-maps/MODAL2_M_AER_OD [accessed 25 February 2025].
55. CDC (Centers for Disease Control and Prevention). 2001. Considerations for distinguishing influenza-like illness from inhalational anthrax. JAMA 286(20):2537–2539, PMID: 11763847, https://doi.org/10.1001/jama.286.20.2537-JWR1128-2-1.
56. Sullivan DM, Krupnick A. 2018. *Using Satellite Data to Fill the Gaps in the US Air Pollution Monitoring Network.* Washington, DC: Resources for the Future.
57. Lyapustin A, Wang Y, Korkin S, Huang D. 2018. MODIS collection 6 MAIAC algorithm. Atmos Meas Tech 11(10):5741–5765, https://doi.org/10.5194/amt-11-5741-2018.
58. Swanson A, Holden ZA, Graham J, Warren DA, Noonan C, Landguth E. 2022. Daily 1 km terrain resolving maps of surface fine particulate matter for the Western United States 2003–2021. Sci Data 9(1):466, PMID: 35918383, https://doi.org/10.1038/s41597-022-01488-y.
59. Won W-S, Oh R, Lee W, Ku S, Su P-C, Yoon Y-J. 2021. Hygroscopic properties of particulate matter and effects of their interactions with weather on visibility. Sci Rep 11:16401.
60. Chen Z, Chen D, Zhao C, Kwan M-P, Cai J, Zhuang Y, et al. 2020. Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: a review of methodology and mechanism. Environ Int 139:105558, PMID: 32278201, https://doi.org/10.1016/j.envint.2020.105558.
61. Orru H, Ebi KL, Forsberg B. 2017. The interplay of climate change and air pollution on health. Curr Environ Health Rep 4(4):504–513., https://doi.org/10.1007/s40572-017-0168-6.
62. Wu J, Xie W, Li W, Li J. 2015. Effects of urban landscape pattern on $PM_{2.5}$ pollution—a Beijing case study. PLoS One 10(11):e0142449, PMID: 26565799, https://doi.org/10.1371/journal.pone.0142449.
63. Yang W, Jiang X. Evaluating the influence of land use and land cover change on fine particulate matter. Sci Rep 11:17612.
64. Palm BB, Peng Q, Fredrickson CD, Lee BH, Garofalo LA, Pothier MA, et al. 2020. Quantification of organic aerosol and brown carbon evolution in fresh wildfire plumes. Proc Natl Acad Sci USA 117(47):29469–29477, PMID: 33148807, https://doi.org/10.1073/pnas.2012218117.
65. Urbanski SP, Hao WM, Baker S. 2008. Chapter 4. Chemical composition of wildland fire emissions. In: *Developments in Environmental Science*. Bytnerowicz A, Arbaugh MJ, Riebau AR, Andersen C, eds. Amsterdam, Netherlands: Elsevier, 79–107.
66. Garofalo LA, Pothier MA, Levin EJT, Campos T, Kreidenweis SM, Farmer DK. 2019. Emission and evolution of submicron organic aerosol in smoke from wildfires in the Western United States. ACS Earth Space Chem 3(7):1237–1247, https://doi.org/10.1021/acsearthspacechem.9b00125.
67. Boaggio K, LeDuc SD, Rice RB, Duffney PF, Foley KM, Holder AL, et al. 2022. Beyond particulate matter mass: heightened levels of lead and other pollutants associated with destructive fire events in California. Environ Sci Technol 56(20):14272–14283, PMID: 36191257, https://doi.org/10.1021/acs.est.2c02099.
68. Migliaccio CT, Bergauff MA, Palmer CP, Jessop F, Noonan CW, Ward TJ. 2009. Urinary levoglucosan as a biomarker of wood smoke exposure: observations in a mouse model and in children. Environ Health Perspect 117(1):74–79, PMID: 19165390, https://doi.org/10.1289/ehp.11378.
69. Giannoni M, Martellini T, Del Bubba M, Gambaro A, Zangrando R, Chiari M, et al. 2012. The use of levoglucosan for tracing biomass burning in $PM_{2.5}$ samples in Tuscany (Italy). Environ Pollut 167:7–15, PMID: 22522313, https://doi.org/10.1016/j.envpol.2012.03.016.
70. Landguth EL, Holden ZA, Graham J, Stark B, Mokhtari EB, Kaleczyc E, et al. 2020. The delayed effect of wildfire season particulate matter on subsequent influenza season in a Mountain West region of the USA. Environ Int 139:105668, PMID: 32244099, https://doi.org/10.1016/j.envint.2020.105668.
71. Jaffe DA, Schnieder B, Inouye D. 2022. Technical note: use of $PM_{2.5}$ to CO ratio as an indicator of wildfire smoke in urban areas. Atmos Chem Phys 22(18):12695–12704, https://doi.org/10.5194/acp-22-12695-2022.
72. AirNow. 2025. Air Quality Index (AQI) Basics. https://www.airnow.gov/aqi/aqi-basics/ [accessed 25 February 2025].
73. Orr A, Migliaccio CAL, Buford M, Ballou S, Migliaccio CT. 2020. Sustained effects on lung function in community members following exposure to hazardous $PM_{2.5}$ levels from wildfire smoke. Toxics 8(3):53, PMID: 32764367, https://doi.org/10.3390/toxics8030053.
74. EPA. 2025. Air Quality System (AQS). https://www.epa.gov/aqs/aqs-manuals-and-guides [accessed 25 February 2025].
75. Liu JC, Mickley LJ, Sulprizio MP, Dominici F, Yue X, Ebisu K, et al. 2016. Particulate air pollution from wildfires in the Western US under climate change. Clim Change 138(3):655–666, PMID: 28642628, https://doi.org/10.1007/s10584-016-1762-6.
76. NASA (National Aeronautics and Space Administration). 2025. MODIS Moderate Resolution Imaging Spectroradiometer. https://modis.gsfc.nasa.gov [accessed 16 March 2025].
77. NASA. 2025. VIIRS Visible Infrared Imaging Radiometer Suite. https://www.earthdata.nasa.gov/data/instruments/viirs#:∼:text=The%20Visible%20Infrared%20Imaging%20Radiometer,atmosphere%2C%20cryosphere%2C%20and%20ocean [accessed 16 March 2025].
78. NASA. 2025. SLSTR Sea and Land Surface Temperature Radiometer. https://www.earthdata.nasa.gov/data/instruments/slstr [accessed 16 March 2025].
79. NASA. 2025. OLCI Ocean and Land Colour Instrument. https://www.earthdata.nasa.gov/data/instruments/olci [accessed 16 March 2025].
80. NASA. 2025. Landsat's Role in Managing Wildland Fires. https://landsat.gsfc.nasa.gov/benefits/wildland-fires/ [accessed 16 March 2025].
81. Zhu S, Tang J, Zhou X, Li P, Liu Z, Zhang C, et al. 2023. Research progress, challenges, and prospects of $PM_{2.5}$ concentration estimation using satellite data. Environ Rev 31(4):605–631, https://doi.org/10.1139/er-2022-0125.
82. NOAA. 2021. Hazard Mapping System Fire and Smoke Product. https://www.ospo.noaa.gov/products/land/hms.html [accessed 16 March 2025].
83. Ruminski M, Kondragunta S, Draxler R, Rolph G. 2007. *Use of* Environmental Satellite Imagery *for* Smoke Depiction *and* Transport Model Initialization. Washington, DC: EPA.
84. Vargo JA. 2020. Time series of potential US wildland fire smoke exposures. Front Public Health 8:126, PMID: 32426315, https://doi.org/10.3389/fpubh.2020.00126.
85. Michael R, Mirabelli MC, Vaidyanathan A. 2023. Public health applications of historical smoke forecasts: an evaluation of archived BlueSky data for the coterminous United States, 2015–2018. Comput Geosci 171:105267, PMID: 39100411, https://doi.org/10.1016/j.cageo.2022.105267.
86. Koplitz SN, Nolte CG, Pouliot GA, Vukovich JM, Beidler J. 2018. Influence of uncertainties in burned area estimates on modeled wildland fire $PM_{2.5}$ and ozone pollution in the contiguous U.S. Atmos Environ (1994) 191:328–339, PMID: 31019376, https://doi.org/10.1016/j.atmosenv.2018.08.020.
87. Faulstich SD, Schissler AG, Strickland MJ, Holmes HA. 2022. Statistical comparison and assessment of four fire emissions inventories for 2013 and a large wildfire in the Western United States. Fire 5(1):27, PMID: 35295881, https://doi.org/10.3390/fire5010027.
88. Jin L, Permar W, Selimovic V, Ketcherside D, Yokelson RJ, Hornbrook RS, et al. 2023. Constraining emissions of volatile organic compounds from Western US wildfires with WE-CAN and FIREX-AQ airborne observations. Atmos Chem Phys 23(10):5969–5991, https://doi.org/10.5194/acp-23-5969-2023.
89. Tang W, Emmons LK, Buchholz RR, Wiedinmyer C, Schwantes RH, He C, et al. 2022. Effects of fire diurnal variation and plume rise on U.S. Air quality during FIREX-AQ and WE-CAN based on the multi-scale infrastructure for chemistry and aerosols (MUSICAv0). JGR Atmospheres 127(16):e2022JD036650, https://doi.org/10.1029/2022JD036650.
90. Hunter DJ. 2007. Matplotlib: a 2D graphics environment. Comput Sci Eng 9(3):90–95, https://doi.org/10.1109/MCSE.2007.55.
91. Cal Fire. 2018. 2018 Fire Season Incident Archive. https://www.fire.ca.gov/incidents/2018 [accessed 12 December 2024].
92. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. 2016. The FAIR guiding principles for scientific data management and stewardship. Sci Data 3(1):160018, PMID: 26978244, https://doi.org/10.1038/sdata.2016.18.
93. Gan RW, Ford B, Lassman W, Pfister G, Vaidyanathan A, Fischer E, et al. 2017. Comparison of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. GeoHealth 1(3):122–136, PMID: 28868515, https://doi.org/10.1002/2017GH000073.