Main Manuscript for

Wildfire smoke impacts on indoor air quality assessed using crowdsourced data in California

Yutong Liang[a]*, Deep Sengupta[a], Mark J. Campmier[b], Joshua S. Apte[b,c], David M. Lunderberg[d], Allen H. Goldstein[a,b]*

[a]Department of Environmental Science, Policy, and Management, University of California at Berkeley, Berkeley, CA 94720, USA.

[b]Department of Civil and Environmental Engineering, University of California at Berkeley, Berkeley, CA 94720, USA.

[c]School of Public Health, University of California at Berkeley, Berkeley, CA 94720, USA.

[d]Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720, USA.

*Yutong Liang or Allen H. Goldstein

Email: yutong.liang@berkeley.edu or ahg@berkeley.edu.

Author Contributions: Y.L., D.S., J.S.A. and A.H.G. designed research; Y.L., M.J.C. and D.M.L. performed the study; Y.L., M.J.C., D.S., J.S.A and A.H.G. jointly analyzed the data and wrote the article.

Competing Interest Statement: The authors declare no competing interests.

Classification: Physical Sciences - Environmental Sciences

Keywords: biomass burning, $PM_{2.5}$, indoor air, exposure, low-cost $PM_{2.5}$ sensors

This PDF file includes:

Main Text
Figures 1 to 5
Table 1

Abstract

Wildfires have become the dominant source of particulate matter ($PM_{2.5}$, < 2.5 μm diameter) leading to unhealthy air quality index occurrences in the western United States. Since people mainly shelter indoors during wildfire smoke events, the infiltration of wildfire $PM_{2.5}$ into indoor environments is a key determinant of human exposure, and is potentially controllable with appropriate awareness, infrastructure investment, and public education. Using time-resolved observations outside and

inside over 1400 buildings from the crowdsourced PurpleAir sensor network in California, we found that infiltration ratios (indoor $PM_{2.5}$ of outdoor origin/outdoor $PM_{2.5}$) were reduced on average from 0.4 during non-fire days to 0.2 during wildfire days. Even with reduced infiltration, mean indoor concentration of $PM_{2.5}$ nearly tripled during wildfire events, with lower infiltration in newer buildings and those utilizing air conditioning or filtration.

Significance Statement

Wildfires have become the dominant source of particulate matter in the western United States. Previous characterizations of exposure to wildfire smoke particles were based mainly on ambient concentration of $PM_{2.5}$. Since people mainly shelter indoors during smoke events, the infiltration of wildfire $PM_{2.5}$ into indoor environments determines exposure. We present analysis of infiltration of wildfire $PM_{2.5}$ into more than 1400 buildings in California using more than 2.4 million sensor-hours of data from the PurpleAir low-cost sensor network. Findings reveal that infiltration of $PM_{2.5}$ during wildfire days was substantially reduced compared with non-fire days, related to people's behavioral change. These results improve understanding of exposure to wildfire particles and facilitate informing the public about effective ways to reduce their exposure.

Main Text

Introduction

Fine particulate matter ($PM_{2.5}$) air pollution is the single-largest environmental risk factor for human health and death in the United States (US) (1). Wildfires are a major source of $PM_{2.5}$, and are documented to cause adverse respiratory health effects and increased mortality (2). Toxicological and epidemiological studies suggest that $PM_{2.5}$ from wildfires is more harmful to the respiratory system than equal doses of non-wildfire $PM_{2.5}$ (3, 4). The number and magnitude of wildfires in the western US has increased in recent decades due to climate change and land management (5–7). Although the annual mean level of $PM_{2.5}$ has substantially declined over this period following the implementation of extensive air quality policies to reduce emissions from controllable sources, the frequency and severity of smoke episodes with $PM_{2.5}$ exceedances has increased sharply due to wildfires in the Pacific Northwest and California (8, 9). The annual mean $PM_{2.5}$ in Northern California has increased since 2015 (*SI Appendix*, Fig. S1) due to massive seasonal fire events, and these events have become the dominant cause of $PM_{2.5}$ exceedances.

People in the United States spend 87% of their time indoors (10). However, the protection against air pollutants of outdoor origin provided by buildings is commonly overlooked in air quality, epidemiologic, and risk assessment studies (11). To accurately characterize and reduce population exposures to wildfire $PM_{2.5}$, it is necessary to understand then optimize how buildings are used by their occupants to mitigate exposure. Previous estimations of indoor particles of outdoor origin typically relied on measurements from a limited number of buildings, and extrapolation of these measurements to other buildings based on the empirical infiltration and removal parameters (12, 13). However, such extrapolation is not applicable to wildfire events because it does not take into account the distribution of protection provided by buildings (including natural and mechanical ventilation) due to lack of data measuring infiltration under representative conditions. The infiltration of outdoor particles is dependent on people's behavior (11, 14, 15), which changes during wildfires (and in 2020 during the COVID-19 pandemic). Pollution levels during wildfire events, and knowledge of those pollution levels through available air quality data, directly impact human

responses aimed at controlling the infiltration of outdoor $PM_{2.5}$ including reducing ventilation, using air conditioning, and using active filtration. Statistically robust observations of the variability of $PM_{2.5}$ infiltration during actual wildfire events across a broad cross-section of normally occupied residences provides the opportunity to understand the distribution of real infiltration rates affecting human exposure, and the factors controlling them, potentially informing guidance towards improvement.

Here, we exploit a recent trend in air quality sensing – public data from a network of ubiquitous crowdsourced low-cost $PM_{2.5}$ sensors – to characterize how indoor air quality during wildfire episodes is affected by buildings and their occupants. We demonstrate that buildings provide substantial protection against wildfire $PM_{2.5}$, and that behavioral responses of building occupants contribute to effective mitigation of wildfire smoke. Real-time $PM_{2.5}$ sensors based on aerosol light scattering have proliferated as easy-to-use and low-cost consumer devices in recent years, providing a novel opportunity to explore the indoor intrusion of wildfire $PM_{2.5}$. Among various devices available, the crowdsourced PurpleAir network has developed the most extensive public-facing network currently available. As of June 2, 2021, there are 15,885 publicly accessible active PurpleAir sensors reporting data from across the earth, 76% are outdoor (12,088), and 24% are indoor (3,797). Of these PurpleAir sensors, 57% are installed in California (9,072), split into 69% outdoor (6,273) and 31% indoor (2,799). As shown in Fig. 1, California accounts for 74% of all indoor PurpleAir sensors worldwide, with adoption increasing most rapidly following individual wildfire episodes, as noted by prior work (16). We focus here on analyzing the data from these sensors deployed across the metropolitan regions of San Francisco and Los Angeles, California, where the public adoption of indoor and outdoor PurpleAir sensors is especially high, at least partially in response to the high frequency of recent wildfire events. Analyses are presented for the wildfire season in the San Francisco Bay Area of Northern California (NC) during August-September 2020 (denoted NC 2020) and November 2018 (NC 2018), and for the Los Angeles area of Southern California (SC) in August-September 2020 (SC 2020). Maps of the measurement regions are provided in *SI Appendix,* Fig. S2 and S3. We analyzed the data from over 1,400 indoor

sensors and their outdoor counterparts to characterize levels of and dynamics of indoor $PM_{2.5}$ and the fraction of outdoor $PM_{2.5}$ that entered buildings, comparing wildfire and non-fire periods. The vast majority (> 87%) of sensors in our dataset are in buildings that are unambiguously identified as residential. We mainly focus on residential buildings, which is facilitated by linking individual PurpleAir sensor locations with a dataset of detailed home property characteristics (Zillow).

Results and Discussion

$PM_{2.5}$ inside and outside an example building. Fig. 2 displays the $PM_{2.5}$ concentrations measured by an indoor sensor and its nearest outdoor counterpart on wildfire days and non-wildfire days (classified by whether the daily average $PM_{2.5}$ level measured by the nearest EPA Air Quality Measurement Station was above or below 35 μg m^{-3}). The outdoor $PM_{2.5}$ concentration was clearly affected by wildfire plumes for August 14-28, September 6-15, and September 28-30. On fire days, the 10-min average outdoor $PM_{2.5}$ exceeded 500 μg m^{-3} several times. The indoor concentration was more than doubled in these periods due to the infiltration of wildfire particles. We also observed peaks of indoor $PM_{2.5}$ exceeding the outdoor $PM_{2.5}$ even on the most polluted days. These peaks typically lasted between 1 hour and 4 hours, which match well with the characteristics of cooking/cleaning peaks, reported in studies such as Patel *et al*. and Tian *et al.* (17, 18). Fig. 2C shows the concentration profiles of indoor and outdoor $PM_{2.5}$, and 1D shows the outdoor $PM_{2.5}$ and indoor $PM_{2.5}$ with outdoor origins (after removal of identified indoor emission events). The infiltration of outdoor wildfire smoke caused the concentration of indoor $PM_{2.5}$ to exceed 150 μg m^{-3} in this building (Fig. 2D).

Differences of infiltration on fire days and non-fire days. Taking all the buildings in the NC 2020 case into consideration, we found that the mean concentration of indoor $PM_{2.5}$ more than doubled on the fire days compared to the non-fire days due to the infiltration of outdoor smoke (Table 1, *SI Appendix,* Fig. S4). On the fire days, the average outdoor concentration of $PM_{2.5}$ was more than 4 times the mean indoor $PM_{2.5}$. Fig. 3A displays the distribution of the mean indoor/outdoor $PM_{2.5}$ ratio of each building on the fire days and the non-fire days. The average indoor/outdoor $PM_{2.5}$

ratios for many buildings exceeded 1 due to indoor emission events, particularly on non-fire days. On fire days, the majority of indoor $PM_{2.5}$ infiltrated from outdoors, but the indoor/outdoor $PM_{2.5}$ ratios were much lower because people closed their buildings and many also filtered their indoor air for protection from the smoke. Figure 3B shows the ratio of indoor $PM_{2.5}$ of outdoor origin to outdoor $PM_{2.5}$ (defined as the infiltration ratio). The infiltration factor (F_{in}) is the steady-state fraction of outdoor $PM_{2.5}$ that enters the indoor environment and remains suspended there (14). It quantifies the extent that the building provides protection against outdoor particles (11). For particulate matter, F_{in} can be obtained from the ratio of indoor/outdoor concentration when there are not additional indoor sources or loss processes (19, 20). On fire days ($PM_{2.5}$ > 35 μg m^{-3}), due to the predominance of $PM_{2.5}$ of outdoor origin, the infiltration ratio approaches the infiltration factor. The infiltration factors of $PM_{2.5}$ for different buildings in NC 2020 have a geometric mean (GM) of 0.22 (0.15, 0.35 for 25th and 75th percentiles, same below). On non-fire days ($PM_{2.5}$ < 35 μg m^{-3}), the GM infiltration ratio increases to 0.4 (0.32, 0.54), while on days with unhealthy air quality ($PM_{2.5}$ > 55.4 μg m^{-3}), the GM infiltration ratio reduces to 0.19 (0.12, 0.31) (Table 1). However, around 25% of buildings had $PM_{2.5}$ infiltration factors above 0.35 on the fire days (Fig. 3B). Occupants of these exposure hotspot buildings experienced much higher levels of wildfire smoke. For context, infiltration factors of homes and commercial buildings measured in the US are usually above 0.5 (14, 21), and the infiltration factor of office buildings with 85% ASHRAE filters were predicted to be around 0.18 (22). The difference in mean infiltration ratio between fire days and non-fire days are most apparent in the daytime (*SI Appendix,* Fig. S5), consistent with more ventilation typically occurring during daytime (23). The lower infiltration factors for the buildings on fire days indicates the efficacy of reduced ventilation and enhanced removal of particles as people took measures to protect themselves from smoke exposure, and that more behavioral changes happened in daytime. Infiltration ratios of $PM_{2.5}$ were not significantly different between fire days and non-fire days in the SC 2020 case (Fig. 4), in contrast to the 2020 NC observations. This difference is probably because the hotter weather in Southern California caused more frequent use of air conditioning systems (and shutting windows), which is implied by a higher 2 pm mean indoor-outdoor temperature

difference (~4ºC) than buildings in the San Francisco Bay Area (~2ºC). Another possibility is that the $PM_{2.5}$ pollution levels in the Greater Los Angeles area were not high enough to induce people to change their behaviors (*SI Appendix,* Fig. S6-S9).

Infiltration and building characteristics. Differences in fire-day infiltration ratios may also stem from differences in building characteristics. As shown in Table S4 in *SI Appendix*, buildings with fire-day infiltration ratio < 0.15 were widely distributed in the study area. However, buildings with fire-day infiltration ratio > 0.35 were mostly located in San Francisco where the climate is cooler and air conditioning is much less common. Buildings in California Climate Zone 12 (Northern California Central Valley) had lower infiltration ratios than any other climate zones in the San Francisco Bay Area (*SI Appendix,* Fig. S10). Due to the summer hot weather, substantial cooling is required for buildings in this zone (24). Air conditioning and associated filtration systems apparently decrease the indoor $PM_{2.5}$ in those buildings. In addition, since the mid-late 1990s, most new residential buildings in the US are equipped with air conditioning systems (25). Since 2008, new buildings in California are mandated to have mechanical ventilation systems (26). Many of the newer buildings also have filtration systems (27). The changes in the building stock are apparent in the resulting data, as residences built after 2000 had significantly lower infiltration ratios on both fire days and non-fire days compared with older buildings (*SI Appendix,* Fig. S10), which is consistent with the findings of a recent wildfire smoke infiltration study in Seattle (28). We further classified the buildings in the NC 2020 case into cool buildings and hot buildings based on whether the 95th percentile indoor temperature reached 30ºC. These cool buildings were more likely to have air conditioning systems on. As shown in *SI Appendix,* Fig. S11, the cool buildings have significantly lower fire-day infiltration ratios than the hot ones ($p < 0.01$), and around 20% of cool buildings had extremely low infiltration ratios (< 0.1). In sum, these results demonstrate that (i) this sensing and analysis approach yields findings in line with mechanistic plausibility (ii) and that the diversity of building characteristics within a region leads to substantial heterogeneity in the degree to which populations are protected indoors from wildfire $PM_{2.5}$.

Decay rate constants for $PM_{2.5}$ were determined for all indoor observations using a box model (Equation 2). The difference in the decay rate constants of $PM_{2.5}$ indoors further reveals why the infiltration ratio was lower on fire days. Fig. 5 shows the distribution of mean total loss rate constant of $PM_{2.5}$ on fire days and non-fire days in the buildings. The mean and median total loss rate constants (λ_t) are 1.6 h^{-1} and 1.3 h^{-1} on fire days, and 2.0 h^{-1} and 2.2 h^{-1} on non-fire days, respectively. Comparing individual buildings on fire days and non-fire days, 70% of them have lower particle loss rate constants on fire days, indicating a high percentage of buildings whose occupants took effective action to reduce $PM_{2.5}$ infiltration. During the fire days, the decrease in air exchange rate exceeded the enhanced indoor filtration, making the loss rate smaller. Since the infiltration ratio (infiltration rate/total loss rate, $aP/(a+k_{loss})$) was also lower on fire days, it can be inferred that the infiltration rate (air exchange rate × penetration factor, aP) was lower on fire days (Equations 1 and 3). We expect both air exchange rate and penetration factor to drop on fire days. Closure of windows and doors will lead to a lower air exchange rate. The usage of filtration systems on incoming air and closure of openings will lead to a lower penetration factor (12). For the SC 2020 case, the estimated particle loss rate constants (1.48 h^{-1} for both fire days and non-fire days) are lower than in the San Francisco Bay Area (*SI Appendix,* Fig. S12), which further implies that a larger fraction of PurpleAir sensor owners in the Los Angeles area kept their windows/doors closed.

People are more likely to open the windows when the indoor temperature is higher than the outdoor temperature in summer (29, 30). In the NC 2020 and SC 2020 cases, the difference in daytime indoor/outdoor temperature alternated between positive and negative values (*SI Appendix,* Fig. S13). However, in the NC 2018 case, due to the colder outdoor temperatures in November, we infer that people probably closed their windows for a longer time, explaining the lower loss rate constants observed. This was expected to reduce the difference between the infiltration ratio on fire days and non-fire days. However, this ratio is still statistically significantly higher ($p < 0.05$) on fire days, which suggests the widespread application of filtration systems.

Our conclusions come with caveats. First, we treated each building as a well-mixed box, which assumes the indoor sensor measurement can represent the $PM_{2.5}$ levels of the entire building. Second, our algorithm to remove the indoor-source peaks could miss lower indoor emission events. In addition, we assumed a universal quasi-linear response for all the PurpleAir sensors throughout the analysis period. Such treatment could lead to biases, but our results should still reflect the average trend. Indoor environments with PurpleAir sensors may not be representative of the entire distribution of buildings (details are provided in the *SI Appendix*). Adoption of PurpleAir sensors (at least ~200 US dollars per sensor) is higher among affluent people concerned about exposure to $PM_{2.5}$. Consistent with the expectation of an affluent "early-adopter" effect, PurpleAir owners live in homes with estimated average property values 21% greater than the median property value for their cities (*SI Appendix,* Table S3 and Fig. S14). The 2015 U.S. Residential Energy Consumption Survey shows that households with less than $40,000 annual income are less likely to use air-conditioning equipment than other households (31). Low-income houses tend to be older, and they are shown to have larger leakage than other houses (32, 33). Lower-income households can therefore have disproportionately higher exposure to wildfire smoke. Finally, although we were not able to disentangle the influence of multiple regionally varying parameters (such as building type, floor area, property values) on penetration of wildfire smoke with the current distribution of indoor sensors, more extensive sensor adoption in coming years may allow future work to address this limitation.

This work demonstrates that crowdsourced environmental sensing can provide valuable information about how people are protecting themselves from the increasingly severe environmental hazard of wildfire smoke. We find that common adaptation measures, including reducing ventilation and active air filtration, effectively mitigate the average indoor exposures of all the buildings by 17% and 74% relative to baseline and outdoor conditions, respectively. This work further suggests that such protective measures could be enhanced through public education to substantially mitigate indoor exposures at the population scale in the future. Given anticipated increases in wildfire smoke in coming decades, it is critical to evaluate these findings in other

settings, including in lower-income communities and in other climate regions affected by wildfires. While our data imply that early adoption of crowdsourced indoor PurpleAir sensors seems to be propelled by wildfire events (Figure 1), gaining more broadly representative insight into the distribution of indoor PM conditions might benefit from complementary approaches to disseminating these sensors, such as targeted deployments in lower-income communities. Overall, our results suggest the increasing ubiquity of indoor and outdoor air pollution sensors can aid in understanding exposures to episodic pollution sources such as wildfires.

Materials and Methods

Selection of Sensor Correction Models. The performance of low-cost $PM_{2.5}$ sensors is dependent on humidity, temperature, particle size distribution and level of particulate matters (34–40). To evaluate the performance of the PurpleAir sensors against reference US EPA $PM_{2.5}$, we linked hourly average measurements from all 16 reference monitors in the study domain (for the entire study period) with surrounding (within 5 km) outdoor PurpleAir sensors, as detailed in the *SI Appendix* (section "Selection of Sensor Correction Models", Figures S6-S9, Tables S1 and S2). We then evaluated the relationship between $PM_{2.5}$ data from PurpleAir sensors and US EPA monitors for multiple calibration schemes in four categories: (i) the manufacturer's default calibration, (ii) previously reported calibration factors for wildfire smoke from the literature (35, 41), (iii) parsimonious empirical calibration relationships based on linear regression using this dataset, and (iv) a machine learning (random forest) based calibration scheme using this dataset. As described in the *SI Appendix,* for the main case (NC 2020), we found that a parsimonious ordinary least-squares model fit nearly precisely matched the manufacturer's default calibration (slope = 1.00, intercept = 0), and provided good agreement with the EPA measurements for this dataset, with R^2 = 0.79 and normalized root mean square error = 0.63. This parsimonious model performed as well or better than other published linear calibration models for wildfire smoke, a finding that underscores that the specific calibration of the Plantower aerosol photometer varies among settings and episodes. For the range of increasingly complex calibration models considering extra

parameters for the PurpleAir vs. reference monitor that we developed, we found moderate further improvement to sensor precision and accuracy, but with qualitatively unchanged results (see *SI Appendix*). Accordingly, we rely on our most parsimonious calibration equation – effectively the same as the manufacturer's default settings – for its more straightforward interpretability in our core analyses.

Decomposition of Indoor $PM_{2.5}$ In addition to infiltration of $PM_{2.5}$ from outdoors, cooking, cleaning and resuspension are the main sources of indoor $PM_{2.5}$ (17, 18, 42). Prior to assessing the amount of indoor $PM_{2.5}$ resulting from infiltration of wildfire smoke, we first identified and removed the events (peaks) caused by indoor sources based on the magnitude and duration of indoor $PM_{2.5}$ peaks. Details of the algorithm can be found in the *SI Appendix.*

Other QA and QC

As described in detailed QA/QC procedures in the *SI Appendix,* we sought to ensure appropriate sensor selection, and to exclude sensors that were likely mislabeled.

Mass Balance Model. We explored the dynamics of indoor $PM_{2.5}$ with a well-mixed box model. When the indoor and outdoor particles are in steady state, and the indoor source is small, we have:

$$\frac{dC_{in}}{dt} = 0 = aPC_{out} - (a + k_{loss})C_{in} \Rightarrow F_{in} = \frac{C_{in}}{C_{out}} = \frac{aP}{a + k_{loss}} \quad \textbf{[1]}$$

where a is the air exchange rate, P is the penetration factor of particles, k_{loss} is the loss rate constant including deposition and indoor filtration. C_{in} and C_{out} are the indoor and outdoor concentrations, respectively (14, 19). F_{in} is the infiltration factor (which is close to the infiltration ratio).

Particle Loss Rate Constant Calculation. After major indoor emission events, the indoor concentration of $PM_{2.5}$ will decay following:

$$\frac{dC_{in}}{dt} = -(a + k_{loss})C_{in} \quad \textbf{[2]}$$

Therefore, (a+k_{loss}) can be estimated by fitting the curve of $C_{in}(t)$ (43). We define the total indoor particle loss rate constant (λ_t) as:

$$\lambda_t = a + k_{loss} \quad \textbf{[3]}$$

Details of the derivation of these equations and the algorithms are provided in the *SI Appendix*.

Building information. Property data were obtained by matching coordinates associated with the PurpleAir sensors to addresses. The list of addresses was then inputted to Zillow, a publicly accessible website to find the publicly available building information such as building age and livable area. Zillow uses existing building information and a proprietary algorithm to derive an estimate of the current (as of December 2020) price of the home or apartment. More details are provided in the *SI Appendix*.

Data availability Data used in this work can be freely downloaded from the PurpleAir and EPA websites (links are provided in the *SI Appendix*).

Acknowledgments

The authors acknowledge Tongshu Zheng at Duke University for his advice on processing the data from the PurpleAir sensors. We thank the owners of PurpleAir sensors who generously shared the measurement data online. This work was supported by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office's AC4 program (Award No. NA16OAR4310107). This publication was also developed as part of the Center for Air, Climate and Energy Solutions (CACES), which was supported under Assistance Agreement No. R835873 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

References

1. GBD 2019 Diseases and Injuries Collaborators, Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).

2. C. E. Reid, *et al.*, Critical review of health impacts of wildfire smoke exposure. *Environ. Health Perspect.* **124**, 1334–1343 (2016).

3. T. C. Wegesser, K. E. Pinkerton, J. A. Last, California wildfires of 2008: Coarse and fine particulate matter toxicity. *Environ. Health Perspect.* **117**, 893–897 (2009).

4. R. Aguilera, T. Corringham, A. Gershunov, T. Benmarhnia, Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nat. Commun.* **12** (2021).

5. P. E. Dennison, S. C. Brewer, J. D. Arnold, M. A. Moritz, Large wildfire trends in the western United States, 1984-2011. *Geophys. Res. Lett.* **41**, 2928–2933 (2014).

6. J. T. Abatzoglou, A. P. Williams, Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11770–11775 (2016).

7. A. L. Westerling, Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. *Science.* **313**, 940–943 (2006).

8. K. O'Dell, B. Ford, E. V Fischer, J. R. Pierce, Contribution of Wildland-Fire Smoke to US PM 2.5 and Its Influence on Recent Trends. *Environ. Sci. Technol.* **53**, 1797–1804 (2019).

9. C. D. McClure, D. A. Jaffe, US particulate matter air quality improves except in wildfire-prone areas. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 7901–7906 (2018).

10. N. E. Klepeis, *et al.*, The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. *J. Expo. Anal. Environ. Epidemiol.* **11**, 231–252 (2001).

11. A. H. Goldstein, W. W. Nazaroff, C. J. Weschler, J. Williams, How Do Indoor Environments Affect Air Pollution Exposure? *Environ. Sci. Technol.* **55**, 100–108 (2021).

12. E. Diapouli, A. Chaloulakou, P. Koutrakis, Estimating the concentration of indoor particles of outdoor origin: A review. *J. Air Waste Manag. Assoc.* **63**, 1113–1129 (2013).

13. K. K. Barkjohn, *et al.*, Real-time measurements of PM2.5 and ozone to assess the effectiveness of residential indoor air filtration in Shanghai homes. *Indoor Air* **31**, 74–87 (2021).

14. C. Chen, B. Zhao, Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmos. Environ.* **45**, 275–288 (2011).

15. L. K. Baxter, C. Stallings, L. Smith, J. Burke, Probabilistic estimation of residential air exchange rates for population-based human exposure modeling. *J. Expo. Sci. Environ. Epidemiol.* **27**, 227–234 (2017).

16. B. Krebs, J. Burney, J. G. Zivin, M. Neidell, Using Crowd-Sourced Data to Assess the Temporal and Spatial Relationship between Indoor and Outdoor Particulate Matter. *Environ. Sci. Technol.* **55**, 6107–6115 (2021).

17. S. Patel, *et al.*, Indoor Particulate Matter during HOMEChem: Concentrations, Size Distributions, and Exposures. *Environ. Sci. Technol.* **54**, 7107–7116 (2020).

18. Y. Tian, *et al.*, Indoor emissions of total and fluorescent supermicron particles during HOMEChem. *Indoor Air* **31**, 88–98 (2021).

19. L. Wallace, R. Williams, Use of personal-indoor-outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons. *Environ. Sci. Technol.* **39**, 1707–1714 (2005).

20. S. Bhangar, N. A. Mullen, S. V Hering, N. M. Kreisberg, W. W. Nazaroff, Ultrafine particle concentrations and exposures in seven residences in northern California. *Indoor Air* **21**, 132–144 (2011).

21. X. Wu, M. G. Apte, D. H. Bennett, Indoor particle levels in small- and medium-sized commercial buildings in california. *Environ. Sci. Technol.* **46**, 12355–12363 (2012).

22. W. J. Riley, T. E. McKone, A. C. K. Lai, W. W. Nazaroff, Indoor particulate matter of outdoor origin: Importance of size-dependent removal mechanisms. *Environ. Sci. Technol.* **36**, 200–207 (2002).

23. H. Erhorn, Influence of meteorological conditions on inhabitants' behaviour in dwellings with mechanical ventilation. *Energy Build.* **11**, 267–275 (1988).

24. Pacific Energy Center, "Pacific Energy Center's Guide to California Climate Zones and Bioclimatic Design" (2006) (January 26, 2021).

25. U.S. Census Bureau, "Characteristics of New Housing" (2019) (February 8, 2021).

26. California Energy Commission, 2008 Building Energy Efficiency Standards (2008).

27. B. C. Singer, W. R. Chan, Y. S. Kim, F. J. Offermann, I. S. Walker, Indoor air quality in California homes with code-required mechanical ventilation. *Indoor Air* **30**, 885–899 (2020).

28. J. Xiang, *et al.*, Field measurements of PM2.5 infiltration factor and portable air cleaner effectiveness during wildfire episodes in US residences. *Sci. Total Environ.* **773**, 145642 (2021).

29. D. Yan, *et al.*, Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **107**, 264–278 (2015).

30. R. Andersen, V. Fabi, J. Toftum, S. P. Corgnati, B. W. Olesen, Window opening behaviour modelled from measurements in Danish dwellings. *Build. Environ.* **69**, 101–113 (2013).

31. EIA, Residential Energy Consumption Survey (RECS). 2 (2013).

32. W. R. Chan, W. W. Nazaroff, P. N. Price, M. D. Sohn, A. J. Gadgil, Analyzing a database of residential air leakage in the United States. *Atmos. Environ.* **39**, 3445–3455 (2005).

33. G. Adamkiewicz, *et al.*, Moving environmental justice indoors: Understanding structural influences on residential exposure patterns in low-income communities. *Am. J. Public Health* **101** (2011).

34. W. W. Delp, B. C. Singer, Wildfire smoke adjustment factors for low-cost and professional PM2.5 monitors with optical sensors. *Sensors* **20**, 1–21 (2020).

35. A. L. Holder, *et al.*, Field evaluation of low-cost particulate matter sensors for measuring wildfire smoke. *Sensors* **20**, 1–17 (2020).

36. K. Ardon-Dryer, Y. Dryer, J. N. Williams, N. Moghimi, Measurements of PM2.5 with PurpleAir under atmospheric conditions. *Atmos. Meas. Tech.* **13**, 5441–5458 (2020).

37. T. Zheng, *et al.*, Field evaluation of low-cost particulate matter sensors in high-and low-concentration environments. *Atmos. Meas. Tech.* **11**, 4823–4846 (2018).

38. K. K. Barkjohn, B. Gantt, A. L. Clements, Development and Application of a United States wide correction for PM2.5 data collected with the PurpleAir sensor. *Atmos. Meas. Tech. Discuss.* (2020) https:/doi.org/10.5194/amt-2020-413 (January 5, 2021).

39. J. Bi, A. Wildani, H. H. Chang, Y. Liu, Incorporating Low-Cost Sensor Measurements into High-Resolution PM2.5 Modeling at a Large Spatial Scale. *Environ. Sci. Technol.* **54**, 2152–2162 (2020).

40. J. Kuula, *et al.*, Laboratory evaluation of particle-size selectivity of optical low-cost particulate matter sensors. *Atmos. Meas. Tech* **13**, 2413–2423 (2020).

41. K. K. Barkjohn, A. Holder, S. Frederick, G. Hagler, A. Clements, PurpleAir PM2.5 U.S. Correction and Performance During Smoke Events in *International Smoke Symposium*, (2020).

42. A. R. Ferro, R. J. Kopperud, L. M. Hildemann, Source Strengths for Indoor Human Activities that Resuspend Particulate Matter. *Environ. Sci. Technol.* **38**, 1759–1764 (2004).

43. B. Stephens, J. A. Siegel, Penetration of ambient submicron particles into single-family residences and associations with building characteristics. *Indoor Air* **22**, 501–513 (2012).

Figures and Tables

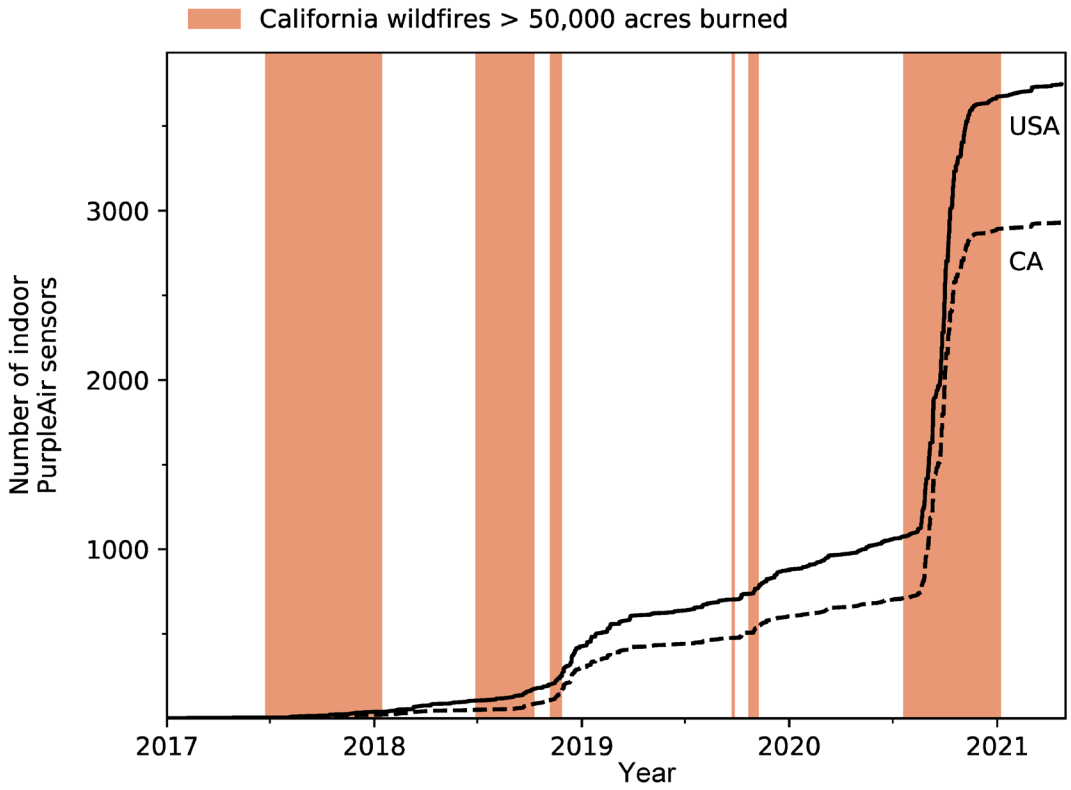

Figure 1. Number of publicly accessible indoor PurpleAir sensors in the United States and California. Shadings show major wildfire periods (start date to containment date of fires with > 50,000 total acres burned) in California. Wildfire periods are from Cal Fire website (https://www.fire.ca.gov/incidents/).

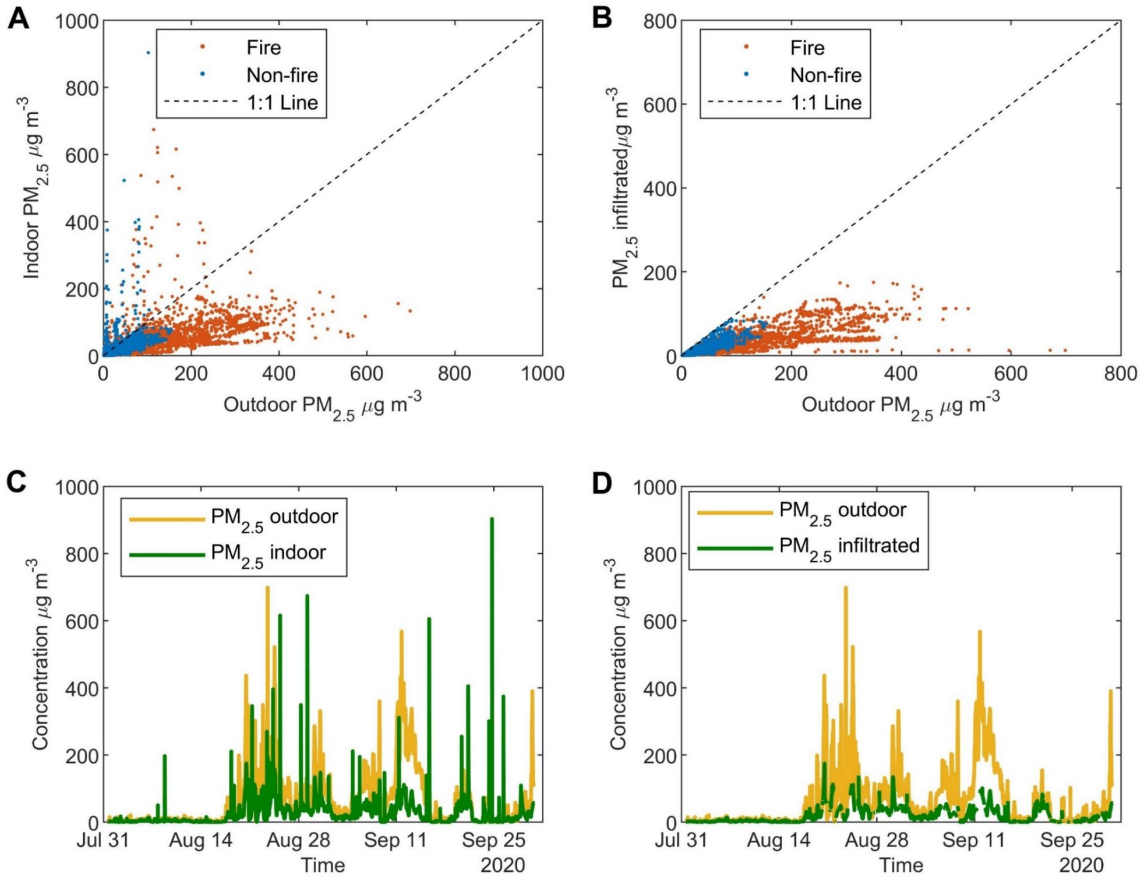

Figure 2. Relationship of indoor and outdoor $PM_{2.5}$ for an example house **(A)** Scatterplots of $PM_{2.5}$ measured at 10-min resolution by an indoor PurpleAir sensor against the nearest outdoor PurpleAir measurement, differentiating fire days (red) and non-fire days (blue), illustrative of the levels of $PM_{2.5}$ pollution of buildings in the NC 2020 case. **(B)** Scatterplots of indoor $PM_{2.5}$ of outdoor origin against outdoor $PM_{2.5}$. **(C)** Concentration time profile of indoor and outdoor $PM_{2.5}$ measured by the two sensors. **(D)** Concentration time profile of infiltrated $PM_{2.5}$ and outdoor $PM_{2.5}$. The figures demonstrate the indoor $PM_{2.5}$ were clearly affected by the outdoor smoke, and our algorithm can effectively remove the indoor peaks due to indoor emissions.

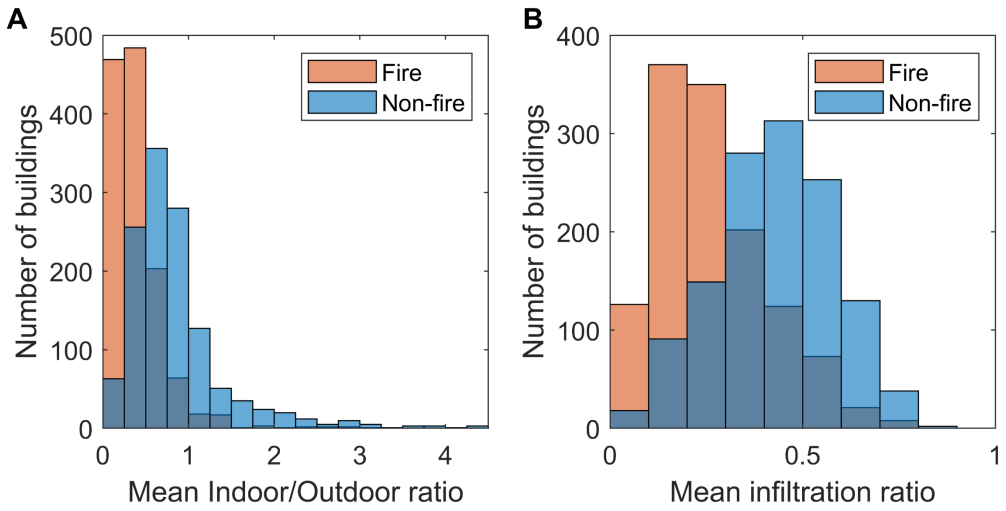

Figure 3. Distribution of the indoor/outdoor ratio and the infiltration ratio in the San Francisco Bay Area in August and September 2020. **(A)** Mean Indoor/Outdoor $PM_{2.5}$ ratio of buildings during fire days and non-fire days and (**B)** mean infiltrated $PM_{2.5}$/Outdoor $PM_{2.5}$ ratio of buildings during fire days and non-fire days. Buildings have lower indoor/outdoor $PM_{2.5}$ ratio and infiltration ratio on fire-days.

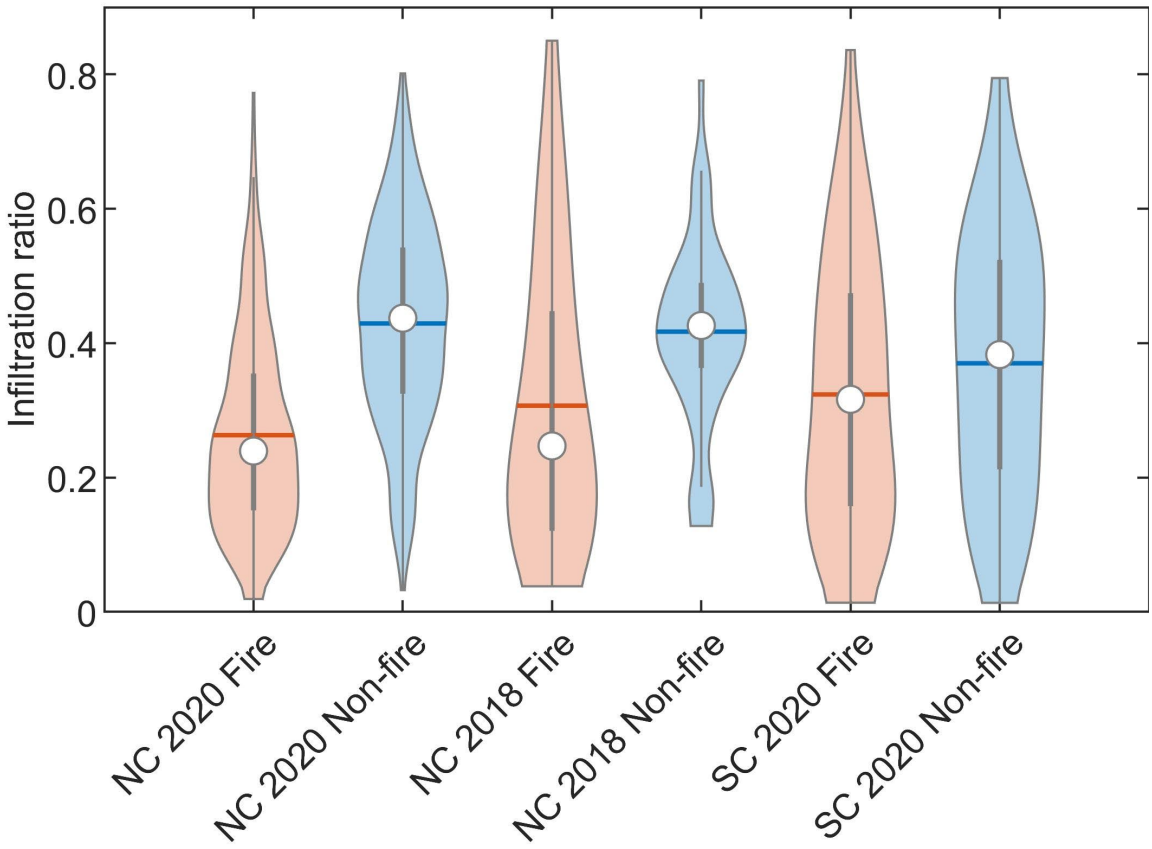

Figure 4. Violin plots of particle infiltration ratios during fire and non-fire periods. N = 1274 buildings, 2.1×10^6 sensor-hours for NC 2020, N = 115 buildings, 2.8×10^5 sensor-hours for SC 2020 and N = 52 buildings, 4.4×10^4 sensor-hours for NC 2018. Each violin plot shows the probability density of the infiltration ratio and a boxplot of interquartile range with whiskers extended to 1.5 times the interquartile range. Circles indicate the median, and horizontal lines indicate the mean.

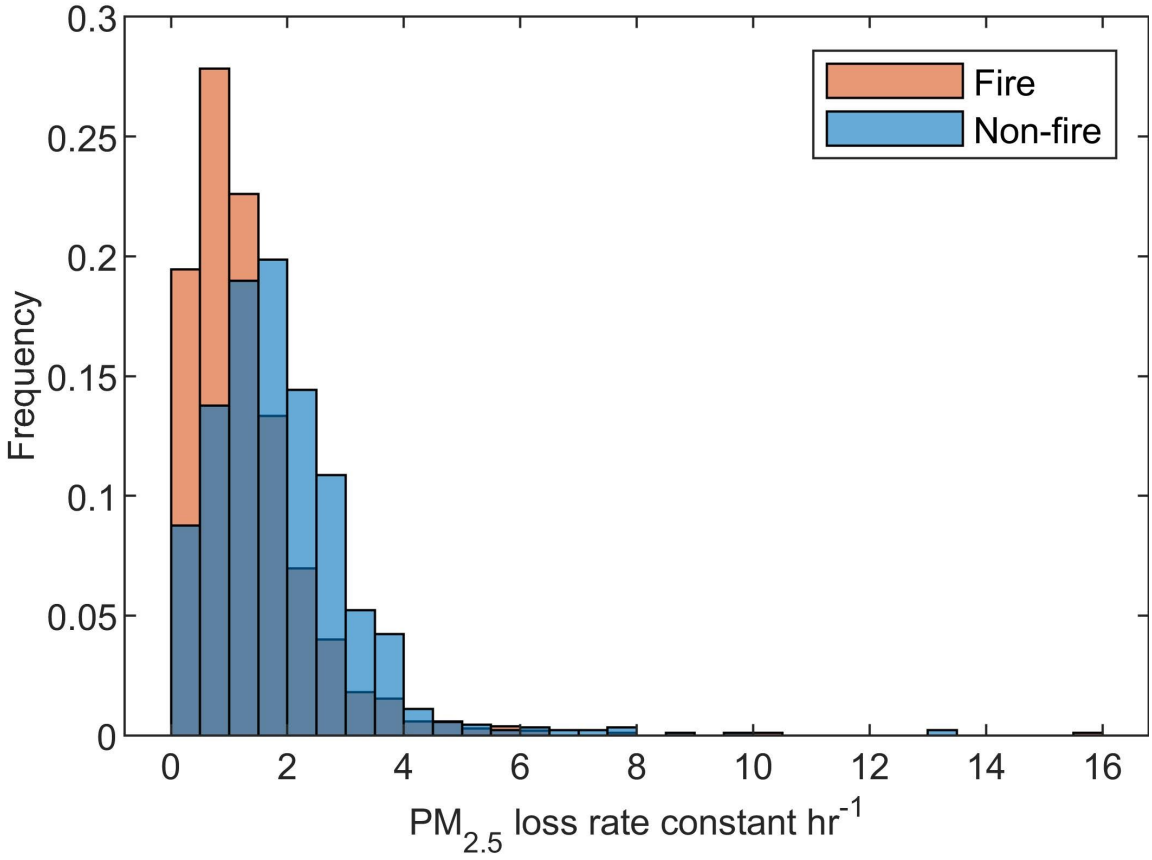

Figure 5. Frequency distribution of indoor $PM_{2.5}$ total loss rate constants (λ_t) in buildings in the San Francisco Bay Area on the fire days and non-fire days in August-September 2020 (decay peaks were found in N = 1195 buildings). A reduced total $PM_{2.5}$ loss rate constant on the fire days indicates a reduction in ventilation.

Table 1. Statistical parameters of the concentration indoor/outdoor ratios for buildings with PurpleAir sensors in August-September 2020 in the San Francisco Bay Area (35 μg m^{-3} daily average $PM_{2.5}$ concentration measured at the nearest EPA measurement site was used as the threshold for fire days and non-fire days). *N* = 1274. Unhealthy days are defined as days with daily average $PM_{2.5}$ concentration above 55.4 μg/m^3. GM = Geometric Mean, GSD = Geometric Standard Deviation.

	Mean outdoor conc μg m^{-3}	**Mean indoor conc μg m^{-3}**				**Indoor/outdoor ratio**				**Infiltration ratio**			
	Mean ± s.d.	Mean ± s.d.	GM, GSD	γ	*β*	Mean ± s.d.	GM, GSD	γ	*β*	Mean ± s.d.	GM, GSD	γ	*β*
Non-fire days	16.0 ± 6.6	7.1±4.5	6.2, 1.7	8.1	1.75	0.94±1.05	0.72, 2.0	1.03	1.24	0.43±0.15	0.40, 1.6	1.03	1.24
Fire days	85.4 ± 32.0	20.5±15.6	16.1, 2.0	22.8	1.46	0.42±0.49	0.31, 2.1	0.46	1.20	0.26±0.14	0.22, 1.8	0.30	1.94
Unhealthy days	115.3 ± 38.8	25.1±20.2	18.8, 2.2	27.6	1.37	0.32±0.42	0.23, 2.2	0.34	1.17	0.23±0.14	0.19, 2.0	0.26	1.70

Quantile-quantile plots (*SI Appendix,* Fig. S4) show the mean concentration of indoor $PM_{2.5}$ in all the buildings can be satisfactorily described by the Weibull distribution. The scale parameter and shape parameter of the Weibull fit are γ and *β,* respectively. The probability distribution function for *x* is $f(x) = \frac{\beta}{\gamma}(\frac{x}{\gamma})^{\beta-1} e^{-(x/\gamma)^{\beta}}$, where *x* > 0. Parameters of the SC 2020 and NC 2018 cases are not shown here due to the small sample sizes, which are less representative of all the buildings in these areas at that time.