

Disruption of outdoor activities caused by wildfire smoke shapes circulation of respiratory pathogens

Beatriz Arregui García^{1†}, Claudio Ascione^{2†}, Arianna Pera^{3†}, Davide Stocco^{4†},
Boxuan Wang^{2,5†}, Eugenio Valdano², Giulia Pullano^{*6}

1 Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus UIB, 07122 Palma de Mallorca, Spain.

2 Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, F75012, Paris, France

3 Department of Computer Science, IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen, Denmark

4 Department of Mathematics, Politecnico di Milano, Via Bonardi 9, I-20133 Milan, Italy (IT)

5 EHESP French School of Public Health, F-35000 Rennes, France

6 Department of Biology, Regents Hall, Georgetown University, 37th and O Streets NW, Washington, DC, 20057-1229, USA

† These authors contributed equally to this work

*e-mail: giulia.pullano@georgetown.edu

Abstract

The consequences of wildfires on public health extend beyond injury. Smoke can traverse vast distances, compromising air quality in unaffected areas and exacerbating chronic respiratory conditions. But smoke may affect the circulation and burden of communicable diseases, too. The disruption in air quality and adherence to safety guidelines can impact the time people spend indoors, and this in turn may increase exposure to airborne pathogens like influenza, SARS-CoV-2, RSV. However, the quantification of such disruptions and their implications for the transmission of respiratory diseases remain unclear. Here we study the effects of smoke generated by severe wildfires in the U.S. states of California, Oregon, Washington in September 2020. We assess the impact on human behavior and the potential consequences for the emergence of respiratory diseases. Our findings reveal a significant shift towards indoor activities in counties within Oregon and Washington during wildfires. However, a discernible change in mobility patterns is not evident in California. This discrepancy may arise from the familiarity of Californian residents with wildfires and air quality index alerts, which have become integrated into their daily routines. Consequently, their mobility patterns may be less affected during such incidents compared to individuals in other regions. We then use a deterministic compartmental model of epidemic spread to quantify the impact of the describe behavioral changes on epidemic circulation. We found that counties with disrupted air exhibited higher cumulated and peak incidence of cases compared to unaffected counties, with the exception of California. Additionally, we found that flu-like epidemics – low reproduction ratio and short generation time – are most affected by the behavioral changes under study. Our findings may help improve public health response in a context of larger, more frequent wildfires triggered by climate change.

1 Introduction

Climate change is causing larger, longer and more devastating wildfires everywhere, and this is especially visible in the United States of America. Adopting this perspective, the New York Times defines wildfires ravaging the West Coast as climate fires [1]. The 2020 wildfire season in the Western United States occurred during the warmest period since global climate records began in 1880, and it was one of the seasons with the highest number of wildfires in the last two decades. Specifically, in August 2020, thunderstorms triggered multiple wildfires across the states of Oregon, Washington, and California which were then followed by additional fire outbreaks along the West Coast in early September [2].

Wildfires cause harm beyond the direct destruction they cause: atmospheric circulation carries the smoke to areas unaffected by the wildfires, compromising air quality over large swathes of land [3]. Smoke exacerbates pre-existing respiratory chronic diseases, increasing their morbidity and mortality [4]. For this reason, the Centers for Disease Control (CDC) provides guidelines to limit outdoor physical activity when and where wildfire smoke makes the air unsafe [5]. But wildfires may threaten human health in other ways, by changing behavioral patterns [6, 7]. Directly perceived air quality deterioration and guidelines induce changes in the ability and willingness of individuals to engage in outdoor activities, altering

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

recurrent human behavior habits, and these behavioral changes may affect the spread of communicable disease [8–14]. This is especially true in the case of airborne respiratory pathogens: influenza, SARS-CoV-2, and Respiratory Syncytial Virus (RSV). Their spread is greatly driven by the social behavior of individuals [15], and indoor activities facilitate transmission and play a significant role in shaping the variability of disease dynamics [16]. While seasonal human behavior changes and their impact on hampering or mitigating epidemic activity have been largely studied [17–19], the impact of Wildfire smokes on human behaviors and the cascade effect for respiratory disease transmission is still unclear.

To address this question, we use a metric of the relative tendency of human interactions to take place indoors at a detailed spatial and temporal scale throughout the United States, as defined in the study published in [16]. We select the U.S. counties with the highest Air Quality Index during the 2020 Western United States wildfire season and quantify the impact of deteriorated air quality on seasonal indoor and outdoor activities. Subsequently, we incorporate this information into an infectious disease transmission model and compare the results to scenarios where indoor activity remains at baseline levels. Our research provides evidence of a shift in human behavior across states in the U.S. during wildfires, enhancing our understanding of the interplay between human behavior and the risk of infection in the context of climate change.

This report describes the results of an educational project carried out in 3 days' time within the framework of the workshop Complexity72h. As such, it provides the starting point for additional research into a problem of great public health relevance, and which climate change is likely to further exacerbate.

2 Methodology

2.1 Data

To select affected counties by wildfire smoke, we analyze the daily Air Quality Index (AQI) in all US counties [20]. The AQI monitors the levels of significant air pollutants that are regulated under the Clean Air Act [21]. Data are collected by Environmental Protection Agency (EPA) and are available on their website[22].

AQI values are made available daily and can be split into the following intervals:

- $0 \leq AQI \leq 50$. **Good**: air quality is satisfactory, and air pollution poses little or no risk.
- $50 < AQI \leq 100$. **Moderate**: air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.
- $100 < AQI \leq 150$. **Unhealthy for sensitive groups**: members of sensitive groups may experience health effects. People are less likely to be affected.
- $150 < AQI \leq 200$. **Unhealthy**: some persons may experience health effects; members of sensitive groups may experience more serious health effects.
- $200 < AQI \leq 300$. **Very unhealthy**: health alert, the risk of health effects is increased for everyone.
- $AQI > 300$. **Hazardous**: health warning of emergency conditions: everyone is more likely to be affected.

In our analysis, we employ the weekly average AQI for each county in the continental United States of America. We focus only on the counties of the states of Oregon (OR), Washington (WA), and California (CA) since these three states were the most affected by smoke caused by more than 100 wildfires affecting the Western US states in September 2020 [23]. As the starting date of the phenomenon, we refer to 10 September 2020 in our analyses since it is the date on which the smoke started to spread over the West Coast. For the analyses focusing on variation over a broader timeframe, we consider the period between July 1, 2020 and November 1, 2020.

We select the counties in OR, WA and CA reporting an AQI greater than 150 (AQI level corresponding at least to "unhealthy") for at least three days during the period between July 1, 2020 and November 1, 2020. We then filter such counties by selecting those that are part of the same indoor activity "Northern" cluster as defined in [16]. For each state, we consider the top 5 counties by population¹. Such a selection process results in the definition of the following affected counties:

¹All census data about the U.S. can be found in <https://data.census.gov/>

- **OR:** Multnomah, Washington, Clackamas, Lane, Marion.
- **WA:** King, Spokane, Clark, Thurston, Yakima.
- **CA:** Plumas.

We then define as *unaffected* all U.S. counties characterized by a *good* or *moderate* AQI level during the studied time window. We select unaffected counties with population size in the top 25% of the distribution (i.e. population size ≥ 67976), for a total of 50 unaffected counties. Those counties will be considered as a baseline in our analyses.

To analyze human behavior, we consider the weekly indoor seasonality index σ_{it} by county computed in Ref. [16] and available here ². The index takes values higher (lower) than 1 if, on average, indoor (outdoor) activities are preferred to outdoor (indoor) ones. The data is based on SafeGraph Weekly Patterns data [24] from 2018 to 2021, and 4.6 million POIs are sampled in all years of our study from all locations where people can spend their time, excluding home locations. The dataset is anonymized, and information on individual mobile devices is omitted. Data are representative across various demographics, including race/ethnicity, educational attainment, and income [25], and it has also been shown not to be significantly biased by age and gender [26], except for individuals younger than 16 years of age are not included [27].

2.2 Regression Discontinuity of indoor activity seasonality

To detect discontinuities in the seasonal pattern of indoor activity, we employ the Regression Discontinuity (RD) approach [28–31]. It is a statistical method used to estimate the causal effect of a treatment or intervention by leveraging a sharp discontinuity in the relationship between a continuous assignment variable and an outcome variable. In our context, the core concept of RD is based on the notion that the time-dependent indoor activity seasonality is inherently similar in the 8 weeks study period, except for the exposure to the wildfire event. By comparing the indoor seasonal activity for the affected counties at the starting date of the event, we can isolate the anomalous effect caused by wildfires from confounding factors (i.e., other potential local anomalies in the mobility).

For regression discontinuity analysis, we use a local linear regression model. It estimates the event's effect by fitting a linear regression line separately for a set of observations after and before it. The local linear regression model can be represented as follows:

$$Y_i = \tilde{\alpha} + \tilde{\beta}X_i + \tilde{\gamma}D_i + \tilde{\epsilon}_i, \quad (1)$$

where:

- Y_i is the outcome variable for observation i ;
- X_i is the assignment variable for observation i ;
- D_i is the wildfire event indicator variable, such that:

$$\begin{cases} D_i = 1 & \text{if } X_i > c \\ D_i = 0 & \text{if } X_i \leq c \end{cases} \quad (2)$$

where c is the moment of impact of the event;

- $\tilde{\alpha}$ and $\tilde{\beta}$ are the coefficients representing the intercept and slope of the regression line, respectively;
- $\tilde{\gamma}$ is the wildfire event effect, which represents the difference in outcome between the observations before and after the event;
- $\tilde{\epsilon}_i$ is the error term.

²https://github.com/bansallab/indoor_outdoor

To give a larger weight to observations closer to the moment of the event, a weighting function is often applied in regression discontinuity analysis. One commonly used weighting function is the triangular kernel. The triangular kernel weighting function can be defined as:

$$w_i = \mathbb{1}_{|X_i - c| > h} \cdot \left[1 - \left(\frac{|X_i - c|}{h} \right) \right], \quad (3)$$

where:

- w_i is the weight assigned to observation i ;
- X_i is the assignment variable for individual i ;
- c is the moment of impact of the event;
- h is the bandwidth parameter that determines the width of the window around the moment of impact.

The weight w_i decreases as the distance between X_i and c increases, with observations exactly at the cutoff receiving the highest weight ($w_i = 1$) and observations further away receiving lower weights.

2.3 Infectious disease model

In our study, we develop a simple SIR model of disease spread to model infectious disease dynamics of the form:

$$\begin{cases} \frac{dS}{dt} = -\beta_0\beta(t)SI \\ \frac{dI}{dt} = \beta_0\beta(t)SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}, \quad (4)$$

where:

- S represents the number of susceptible individuals;
- I is the number of infectious individuals;
- R is the number of recovered individuals;
- γ is the recovery rate;
- $\beta_0\beta(t)$ is the transmissibility rate, defined as the product of homogeneous time and country-independent parts β_0 and $\beta(t)$. β_0 is a constant parameter that takes into account the overall transmissibility, while $\beta(t)$ is related to the seasonality effect and the potential disruption of human mobility σ_{it} caused by wildfires.

We define the *force-of-infection* as $\Lambda = \frac{I}{N}\beta_0\beta(t)$, and the total population N . We run all the simulations starting one week before the start of the wildfire. In our study, the SIR model is evaluated for the *affected* counties OR, WA, and CA and the *unaffected* counties described in subsection 2.1. To compare the model on the two types of counties, we introduce the *relative Attack Rate* $AR(t)$ as the relative cumulative number of infectious individuals at time t over the population size, and the *relative Peak Incidence* $PI(t)$ as the relative variation of the occurrence of new cases of disease at the incidence peak, where the *relative* quantities are obtained wrt every element of the control-group counties and then statistically analyzed.

We explore the relative attack rate and the relative peak incidence comparing *affected* and *unaffected* counties using baseline *reproduction ratios* ($R = \beta/\gamma$) values of 1.3, 1.5 and 3. We choose such values because 1.3 and 1.5 are compatible with a seasonal spread of respiratory viruses such as influenza and currently SARS-CoV-2, and 3.0 was close to the basic reproduction ratio of the wild-type of SARS-CoV-2.

3 Results

3.1 Disruptions to mobility seasonality due to wildfires response

Given the disruptive smoke spread on U.S. West Coast in September 2020 (Figure 1, left-hand side), our preliminary analysis compares the AQI for *affected* counties in Oregon (OR), Washington (WA) and California (CA) (Figure 1, right-hand side). To provide a baseline, we report the AQI of U.S. *unaffected* counties. We notice that a peak in AQI is visible for OR, WA, and CA affected counties right after 2020/09/10. Oregon (OR) shows the highest peak.

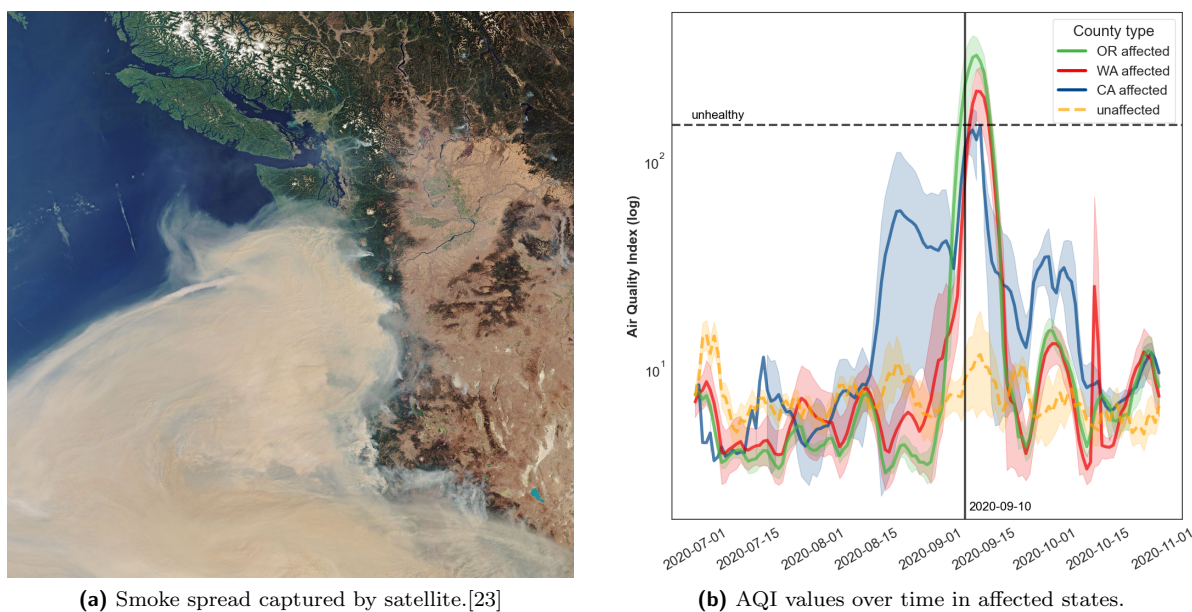


Figure 1: (a) Image captured on 2020/09/10, by Copernicus Sentinel-3, showing the extent of the smoke plume. (b) AQI for Oregon (OR), Washington (WA) and California (CA) counties affected by unhealthy air quality during the September 2020 wildfires. The reported AQI takes into account a 5-days moving average.

By exploring the variation of the indoor activity in the counties *affected* by the smoke of the September 2020 wildfires (Figure 2), we detect significant increases in the time spent indoors, compared to the baseline. These significant peaks are found in all the Oregon counties under study and in King, Clark, and Spokane counties in Washington. A slight signal of a peak in the indoor activity seasonality can also be noticed in the affected Plumas County, CA, even though its AQI is lower compared to that of the other *affected* counties considered. It can be noticed that, even when the peak of indoor activity is not significant with respect to the baseline in Washington counties, there is evidence of an increase in indoor activity.

We quantified discontinuities in indoor activity seasonality with respect to September 2020 wildfires. Table 1 shows the results of the analysis. We observe that all Washington *affected* counties show a significant increase in indoor activity seasonality after the starting date of the wildfires. This also holds for Multnomah and Clackamas counties in Oregon. This latter is the county showing the greatest increase in indoor activity overall.

The increase is not significant for Washington, Lane, and Marion counties in Oregon and for Plumas County in California.

3.2 Impact on respiratory disease circulation caused by disruption on indoor activity

Figure 3 shows the variation of relative attack rate and relative peak incidence between the different *affected* counties and the *unaffected*. The results show relative values above zero for both epidemic outcomes, meaning that the attack rate and the peak incidence of *affected* counties are higher than those of *unaffected* counties. Washington, OR, and Yakima, WA have the highest relative values. On the contrary, Plumas, CA, presents negative relative values, as we would have expected from the different behavior of its corresponding indoor activity seasonality. *reproduction ratio* equal to 3.0 generally results

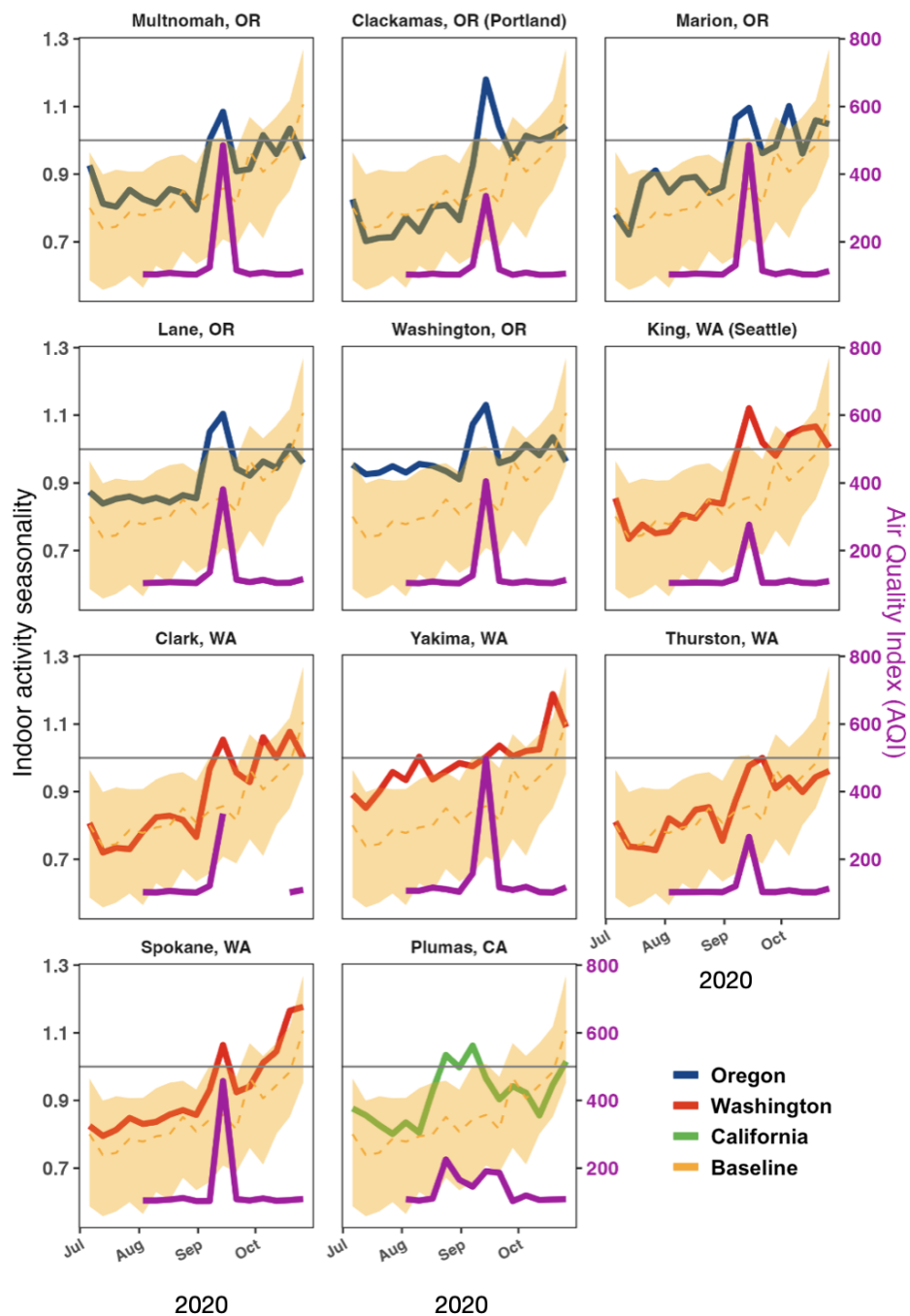


Figure 2: Indoor activity seasonality index between 2020/07/01 and 2020/11/01 in the 11 selected *affected* counties. In each subplot, we also take into account the median and 95% CI of indoor activity seasonality index for the baseline *unaffected* counties (yellow curves). Violet curves represent the AQI of the *affected* counties in the studied time period.

| Country | County | $\Delta\tilde{\gamma}$ | $\Delta\tilde{\gamma}_{5\%}$ | $\Delta\tilde{\gamma}_{95\%}$ |
|---------|------------|------------------------|------------------------------|-------------------------------|
| OR | Multnomah | 0.09 | 0.00 | 0.19 |
| OR | Washington | 0.07 | -0.01 | 0.14 |
| OR | Clackamas | 0.30 | 0.23 | 0.36 |
| OR | Lane | 0.07 | -0.01 | 0.14 |
| OR | Marion | 0.02 | -0.06 | 0.10 |
| WA | King | 0.15 | 0.10 | 0.20 |
| WA | Spokane | 0.13 | 0.08 | 0.18 |
| WA | Clark | 0.10 | 0.02 | 0.18 |
| WA | Thurston | 0.15 | 0.09 | 0.21 |
| WA | Yakima | 0.02 | 0.01 | 0.04 |
| CA | Plumas | 0.02 | -0.01 | 0.04 |

Table 1: Difference in $\tilde{\gamma}$ before and after 10 September 2020 and 90% CI.

in the highest relative attack rate, while the *reproduction ratio* equal to 1.3 leads overall to the highest relative peak incidence values.

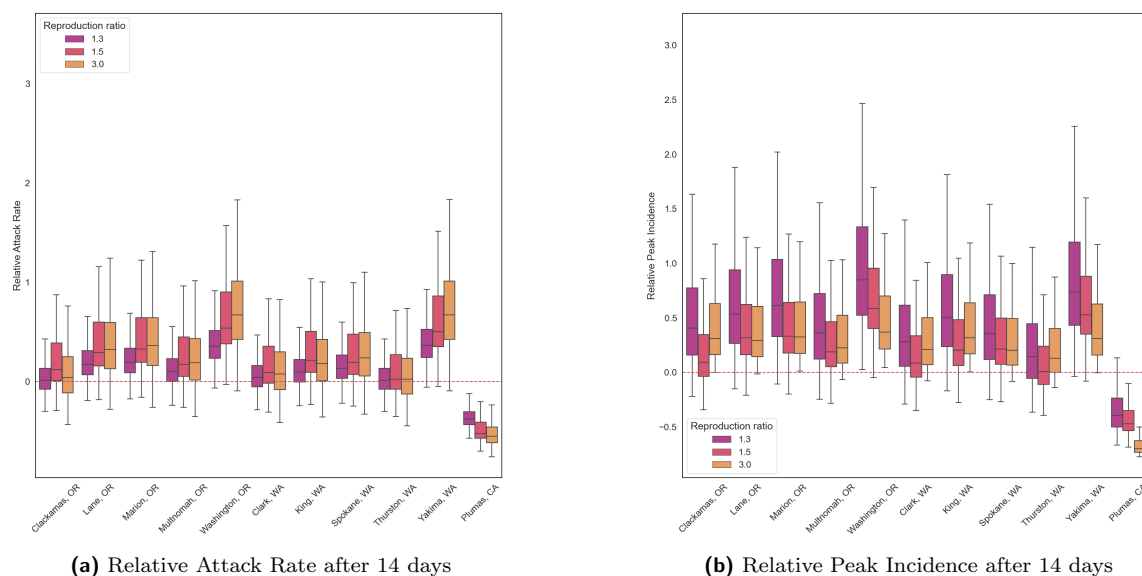


Figure 3: Relative attack rate (left) and relative peak incidence (right) for *affected* and *unaffected* counties after 14 days from the air quality alert due to wildfires (2020/09/10). We set the generation time equal to 14 days.

To characterize the impact of wildfires on different airborne respiratory diseases, we study model epidemic outcomes varying the generation time. We explore generation time ranging from 2 days (Flu-like) to 25 days (Pertussis-like) [32]. For this analysis, we focus on Yakima, WA (Figure 4). We study the variation in the relative values of attack rate and peak incidence compared again with the *unaffected* counties, while exploring different lengths of the generation time and three different *reproductive ratio* values. On the top row of the figure (*reproductive ratio* = 1.3), we observe a decreasing trend of the relative values of the epidemic's outcomes. For short generation times (less than one week), we observe higher differences and also higher variability. We observe the same behavior, but with lower values of the relative measures, when we increase the *reproductive ratio* up to 1.5. This suggests that within generation times shorter than the period where the mobility of human activity occurs, the outbreak of the epidemic in an *affected* county is faster than in *unaffected* counties. The fact that within short ranges of generation time, the infections occur faster explains why we observe a higher variability of the results. When we look at higher values of the *reproductive ratio* (bottom row, Figure 4) we no longer see the same pattern. Since the epidemic grows faster, the disruption of human activity's effect on the outbreak becomes very smooth.

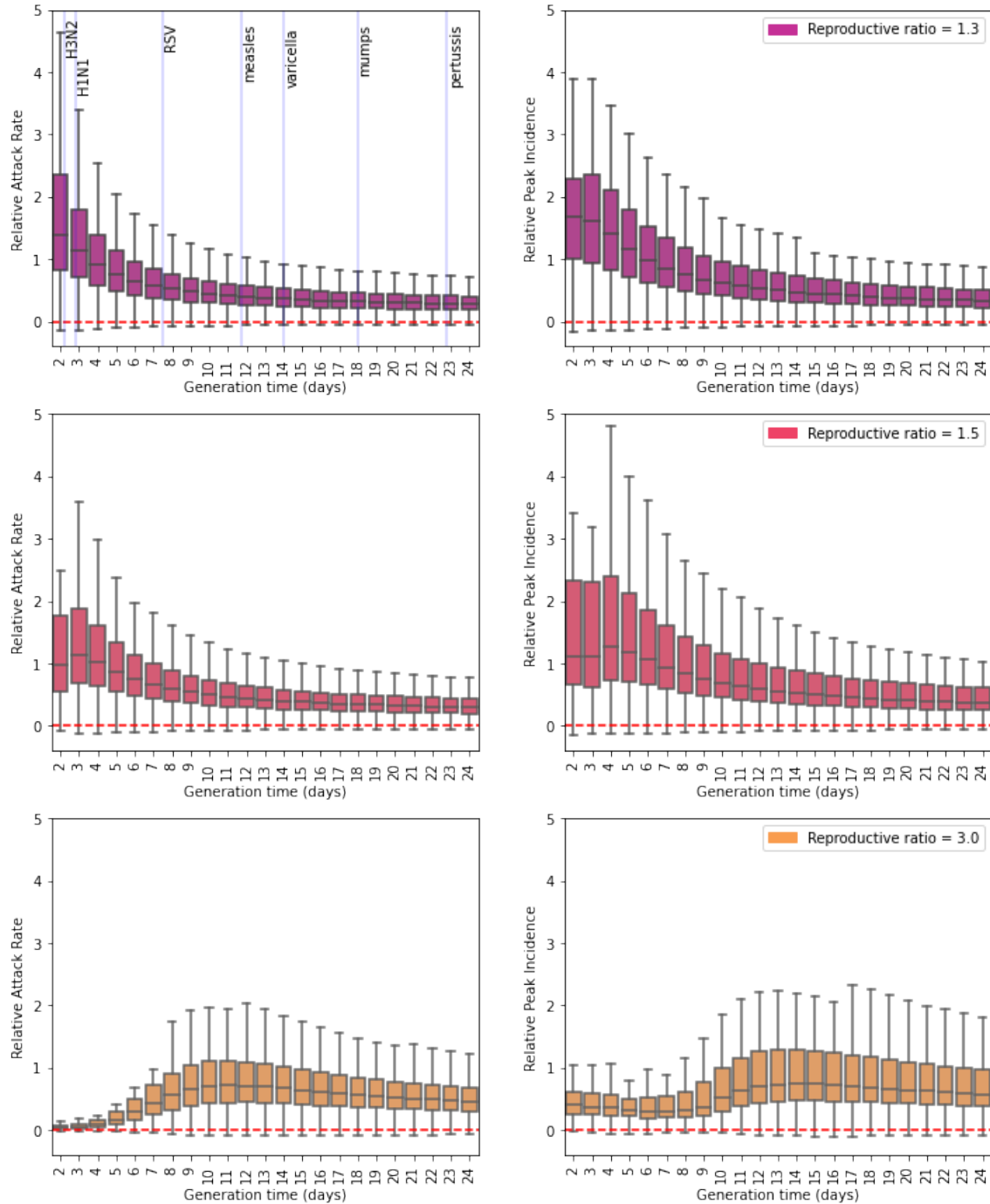


Figure 4: Relative attack rate (left) and relative peak incidence (right) after 14 days at one of the *affected* counties (Yakima, WA). On the top left panel, we show as a reference the mean serial interval for different airborne diseases [33].

4 Discussion

The impact of wildfires on public health extends beyond injuries. Smoke from wildfires can be transported by atmospheric circulation to areas unaffected by the fires, compromising air quality over vast stretches of land [3] and aggravating chronic respiratory conditions, like chronic obstructive pulmonary disease, emphysema, asthma [4]. But wildfire smoke may aggravate the burden of communicable respiratory disease, too. The perceived decline in air quality and adherence to safety guidelines can alter individuals' ability and willingness to participate in outdoor activities, disrupting established behavioral patterns [18]. The disruption of human activity and the resulting influence on the transmission of communicable diseases is still unclear.

In our study, we aimed to bridge this gap by examining the effects of the severe wildfires that struck the West Coast of the United States of America in September 2020, and specifically the states of Oregon, Washington and California. Our focus was on assessing the impact of compromised air quality on human behavior and the potential consequences of the circulation of respiratory pathogens. We observed a significant increase in the time spent indoors in the affected areas of Oregon and Washington. This was instead not evident in California: we propose that this discrepancy comes from the fact that California residents are accustomed to regular exposure to wildfire smoke and Air Quality Index alerts, causing fatigue in their behavioral response. Then we studied the impact on epidemic circulation using a Susceptible-Infected-Recovered compartmental model to simulate the circulation of a respiratory pathogen right before, during and in the aftermath of the Air Quality Index alert, signaling compromised air quality. Specifically, we simulated the impact on the cumulative incidence of cases at four weeks after the event – called *relative attack rate* –, and on the peak daily incidence of cases over the same period – called *relative peak incidence*. We find that both the cumulative incidence and the peak daily incidence in *affected* counties are overall higher than those of *unaffected* counties in Oregon and Washington, not in California, which is consistent with the fact that no significant disruption in the time spent outdoors was visible there. In terms of *reproduction ratio*, a value of 3.0 (compatible with the SARS-CoV-2 onset) generally results in the highest relative cumulative incidence while a value of 1.3 (compatible with seasonal spreads of respiratory diseases) leads to the highest relative peak daily incidence, given a generation time of 14 days. Then we focused on a single *affected* county – Yakima Co., WA – and tested the impact of compromised air quality on disease circulation at varying average generation times. While with lower values of *reproductive ratio* we observed that, within short generation times, there is a higher difference of cumulative incidence and daily peak incidence between *affected* and *unaffected* counties, with higher values of the ratio the disruption of human activity's effect on the outbreak becomes very smooth.

This study is preliminary and has limitations. We used a deterministic model of disease spread: this overlooks the inherent stochasticity of both natural hazard events and epidemic dynamics. Additionally, we study only one major wildfire event, and our findings may not be generalizable to similar events. Moreover, the study relies solely on the ratio of indoor-outdoor activities and does not incorporate information about disruptions in the time spent at home, which could affect epidemic circulation, too.

Our study is preliminary and educational in purpose. Notwithstanding, it highlights the need for additional modeling research on the impact that short periods of compromised air quality due to wildfire smoke may have on the circulation of respiratory pathogens. This is crucial to inform targeted interventions and awareness campaigns during wildfire seasons, emphasizing the importance of protecting vulnerable populations and promoting preventive measures. In particular, the proposed analysis underscores the importance of enhancing emergency preparedness protocols in regions prone to wildfires along with policy guidelines preventing epidemics spread.

Acknowledgments

This work is the output of the Complexity72h workshop³, held at IFISC in Palma, Spain, 26-30 June 2023. We want to acknowledge Shweta Bansal for insightful discussions on the design of the project and for data provision.

³<https://www.complexity72h.com>

References

- [1] Stuart A. Thompson and Yaryna Serkez. “Opinion — Every Place Has Its Own Climate Risk. What Is It Where You Live?” en-US. In: *The New York Times* (Sept. 2020). ISSN: 0362-4331.
- [2] ESA. *West coast wildfires*. en.
- [3] Erum Salam. *US air quality as far south as Virginia affected by Nova Scotia wildfires — US news — The Guardian*.
- [4] Marina Romanello et al. “The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels”. English. In: *The Lancet* 400.10363 (Nov. 2022). Publisher: Elsevier, pp. 1619–1654. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(22)01540-9.
- [5] CDC. *Wildfire Smoke Guide Publications — AirNow.gov*. en. Publisher: AirNow.gov, U.S. EPA.
- [6] Lisa Thalheimer. “Compound impacts of extreme weather events and COVID-19 on climate mobilities”. en. In: *Area* 55.1 (2023). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/area.12821>, pp. 134–141. ISSN: 1475-4762. DOI: 10.1111/area.12821.
- [7] Zhenhua Chen et al. “Impact of extreme weather events on urban human flow: A perspective from location-based service data”. en. In: *Computers, Environment and Urban Systems* 83 (Sept. 2020), p. 101520. ISSN: 0198-9715. DOI: 10.1016/j.compenvurbsys.2020.101520.
- [8] Shweta Bansal, Bryan T Grenfell, and Lauren Ancel Meyers. “When individual behaviour matters: homogeneous and network models in epidemiology”. In: *Journal of The Royal Society Interface* 4.16 (July 2007). Publisher: Royal Society, pp. 879–891. DOI: 10.1098/rsif.2007.1100.
- [9] Pratha Sah et al. “Revealing mechanisms of infectious disease spread through empirical contact networks”. en. In: *PLOS Computational Biology* 17.12 (2021). Publisher: Public Library of Science, e1009604. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1009604.
- [10] Giulia Pullano et al. “Underdetection of cases of COVID-19 in France threatens epidemic control”. en. In: *Nature* 590.7844 (Feb. 2021). Number: 7844 Publisher: Nature Publishing Group, pp. 134–139. ISSN: 1476-4687. DOI: 10.1038/s41586-020-03095-6.
- [11] Laura Di Domenico et al. “Modelling safe protocols for reopening schools during the COVID-19 pandemic in France”. en. In: *Nature Communications* 12.1 (Feb. 2021). Number: 1 Publisher: Nature Publishing Group, p. 1073. ISSN: 2041-1723. DOI: 10.1038/s41467-021-21249-6.
- [12] *Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies — BMC Medicine — Full Text*.
- [13] Chenfeng Xiong et al. “Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections”. In: *Proceedings of the National Academy of Sciences* 117.44 (Nov. 2020). Publisher: Proceedings of the National Academy of Sciences, pp. 27087–27089. DOI: 10.1073/pnas.2010836117.
- [14] Ying Zhou et al. “Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data”. en. In: *The Lancet Digital Health* 2.8 (Aug. 2020), e417–e424. ISSN: 2589-7500. DOI: 10.1016/S2589-7500(20)30165-5.
- [15] Chenfeng Xiong et al. *Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections*. en. DOI: 10.1073/pnas.2010836117.
- [16] Zachary Susswein, Eva C Rest, and Shweta Bansal. “Disentangling the rhythms of human activity in the built environment for airborne transmission risk: An analysis of large-scale mobility data”. In: *eLife* 12 (Apr. 2023). Ed. by Niel Hens, Diane M Harper, and Guillaume Béraud. Publisher: eLife Sciences Publications, Ltd, e80466. ISSN: 2050-084X. DOI: 10.7554/eLife.80466.
- [17] Giulia Pullano et al. “Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study”. English. In: *The Lancet Digital Health* 2.12 (Dec. 2020). Publisher: Elsevier, e638–e649. ISSN: 2589-7500. DOI: 10.1016/S2589-7500(20)30243-0.
- [18] Amy Maxmen. “Can tracking people through phone-call data improve lives?” en. In: *Nature* 569.7758 (May 2019). Bandiera_abtest: a Cg.type: News Feature Number: 7758 Publisher: Nature Publishing Group Subject.term: Epidemiology, Malaria, Research data, Ebola virus, pp. 614–617. DOI: 10.1038/d41586-019-01679-5.
- [19] *Non-pharmaceutical interventions during the COVID-19 pandemic: A review - ScienceDirect*.

- [20] AirNow. *AQI Basics — AirNow.gov*. en. Publisher: AirNow.gov, U.S. EPA.
- [21] OAR US EPA. *Clean Air Act Text*. en. Collections and Lists. May 2015.
- [22] OAR US EPA. *Air Data: Air Quality Data Collected at Outdoor Monitors Across the US*. en. Collections and Lists. July 2014.
- [23] *West coast wildfires*. en.
- [24] Safegraph. *Places Data Curated for Accurate Geospatial Analytics — SafeGraph*.
- [25] Fox Squire. *Google Colab*. URL: <https://colab.research.google.com/drive/1u15afRytJMsizySFqA2EPLXSh3KTmNTQ#sandboxMode=true&scrollTo=xsNNli6GTN6s>.
- [26] Kathryn Zickuhr and Smith. *28% of American adults use mobile and social location-based services*. en-US. Sept. 2011.
- [27] Safegraph. *Privacy Policy*.
- [28] Guido W. Imbens and Thomas Lemieux. “Regression discontinuity designs: A guide to practice”. en. In: *Journal of Econometrics*. The regression discontinuity design: Theory and applications 142.2 (Feb. 2008), pp. 615–635. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2007.05.001.
- [29] *Causal Inference for The Brave and True*.
- [30] Catherine Hausman and David S. Rapson. “Regression Discontinuity in Time: Considerations for Empirical Applications”. en. In: *Annual Review of Resource Economics* 10.1 (Oct. 2018), pp. 533–552. ISSN: 1941-1340, 1941-1359. DOI: 10.1146/annurev-resource-121517-033306.
- [31] Wilbert Van Der Klaauw. “Regression–Discontinuity Analysis: A Survey of Recent Developments in Economics”. en. In: *LABOUR* 22.2 (2008), pp. 219–245. ISSN: 1467-9914. DOI: 10.1111/j.1467-9914.2008.00419.x.
- [32] Margaretha Annelie Vink, Martinus Christoffel Jozef Bootsma, and Jacco Wallinga. “Serial Intervals of Respiratory Infectious Diseases: A Systematic Review and Analysis”. In: *American Journal of Epidemiology* 180.9 (Nov. 2014), pp. 865–875. ISSN: 0002-9262. DOI: 10.1093/aje/kwu209.
- [33] Margaretha Annelie Vink, Martinus Christoffel Jozef Bootsma, and Jacco Wallinga. “Serial Intervals of Respiratory Infectious Diseases: A Systematic Review and Analysis”. In: *American Journal of Epidemiology* 180.9 (Nov. 2014), pp. 865–875. ISSN: 0002-9262. DOI: 10.1093/aje/kwu209.