

# A SPATIAL CAUSAL ANALYSIS OF WILDLAND FIRE-CONTRIBUTED $PM_{2.5}$ USING NUMERICAL MODEL OUTPUT

BY ALEXANDRA LARSEN<sup>1,a</sup>, SHU YANG<sup>2,b</sup>, BRIAN J. REICH<sup>2,c</sup> AND ANA G. RAPPOLD<sup>3,d</sup>

<sup>1</sup>*Department of Biostatistics and Bioinformatics, Duke University, [alexandra.larsen@duke.edu](mailto:alexandra.larsen@duke.edu)*

<sup>2</sup>*Department of Statistics, North Carolina State University, [syang24@ncsu.edu](mailto:syang24@ncsu.edu), [bjreich@ncsu.edu](mailto:bjreich@ncsu.edu)*

<sup>3</sup>*National Health and Environmental Effects Research Laboratory—Environmental Public Health Division, US Environmental Protection Agency, [Rappold.Ana@epa.gov](mailto:Rappold.Ana@epa.gov)*

Wildland fire smoke contains hazardous levels of fine particulate matter ( $PM_{2.5}$ ), a pollutant shown to adversely effect health. Estimating fire attributable  $PM_{2.5}$  concentrations is key to quantifying the impact on air quality and subsequent health burden. This is a challenging problem since only total  $PM_{2.5}$  is measured at monitoring stations and both fire-attributable  $PM_{2.5}$  and  $PM_{2.5}$  from all other sources are correlated in space and time. We propose a framework for estimating fire-contributed  $PM_{2.5}$  and  $PM_{2.5}$  from all other sources using a novel causal inference framework and bias-adjusted chemical model representations of  $PM_{2.5}$  under counterfactual scenarios. The chemical model representation of  $PM_{2.5}$  for this analysis is simulated using Community Multiscale Air Quality Modeling System (CMAQ), run with and without fire emissions across the contiguous U.S. for the 2008–2012 wildfire seasons. The CMAQ output is calibrated with observations from monitoring sites for the same spatial domain and time period. We use a Bayesian model that accounts for spatial variation to estimate the effect of wildland fires on  $PM_{2.5}$  and state assumptions under which the estimate has a valid causal interpretation. Our results include estimates of the contributions of wildfire smoke to  $PM_{2.5}$  for the contiguous U.S. Additionally, we compute the health burden associated with the  $PM_{2.5}$  attributable to wildfire smoke.

**1. Introduction.** Wildfires are a leading contributor to unhealthy air quality in many communities across the world. Among the pollutants found in smoke, fine particulate matter (mixtures of particles smaller than  $2.5 \mu\text{m}$  in diameter or  $PM_{2.5}$ ) associated with a number of respiratory and cardiovascular outcomes, is of the largest public health concern (Dennekamp and Abramson (2011), Dennekamp et al. (2015), Haikerwal et al. (2015, 2016), Johnston et al. (2012), Rappold et al. (2011), Wettstein et al. (2018)). The objective of this study is to estimate the wildland fire-attributable fraction of ambient  $PM_{2.5}$  in order to quantify the related health burden. We introduce a potential outcomes framework to estimate the causal effect of wildland fires on ambient  $PM_{2.5}$  in the presence of spatial correlation. The framework leverages numerical model simulations of air quality serving as biased representations of the potential outcomes. A Bayesian spatial downscaling model is used to learn the relationship between the spatially and temporally resolved numerical model output and the sparsely observed  $PM_{2.5}$  from air quality monitors and to provide unbiased estimates of counterfactual outcomes, quantification of uncertainty, and predictions that are both spatially and temporally resolved.

To quantify the magnitude of the health burden attributable to the smoke from fire events, we need to distinguish the  $PM_{2.5}$  composition mixture attributable to fire from the  $PM_{2.5}$  mixture due to all other sources. Total ambient  $PM_{2.5}$  concentrations are recorded at the

---

Received August 2020; revised November 2021.

*Key words and phrases.* Bayesian analysis, downscaling, interference, spillover effect.

monitoring sites across the country; however, these observations do not provide insight into the potential composition of particles that would have formed had there been no wildfires. The mixture of particles measured on any given day depends on multiple sources of emissions and conditions of combustion by which particles were produced. Once released, particles and gases coalesce and interact with those already present in the atmosphere through nonadditive chemical and physical processes. Formation of PM<sub>2.5</sub> is additionally confounded by external factors, including fire weather conditions, vegetation, burned areas, and areas unable to burn again as well as anthropogenic and other natural emissions (McKenzie et al. (2014), Stavros, McKenzie and Larkin (2014)). Finally, in the presence of fire, nonfire emissions themselves can be altered through feedbacks. Together, these factors lead to complex dependencies of PM<sub>2.5</sub> concentrations across space and time.

To distinguish fire-contributed PM<sub>2.5</sub> from total ambient concentrations, we utilize numerical model representations of air quality. The model simulates chemical reactions and transport of particle mixtures in the atmosphere using deterministic representations of chemical processes under a set of input emissions and external forcings. By removing the forcing for wildfire emissions, these models produce air quality simulations from the counterfactual scenario, that is, PM<sub>2.5</sub> composition that would have formed had there not been wildfires. In this study we use the Community Multiscale Air Quality (CMAQ) numerical model to simulate air quality under observed and counterfactual forcings. The difference in PM<sub>2.5</sub> under CMAQ representations of air quality, with and without wildfire emissions, is considered to be a modeled representation of fire-contributed PM<sub>2.5</sub>.

Numerical models have been used to simulate counterfactual environmental conditions in other contexts, most notably to investigate future unobserved or distant past climate trajectories (Allen and Stott (2003), Hammerling, Katzfuss and Smith (2017), Hegerl and Zwiers (2011), Knutson et al. (2017), National Academies of Science (2016)). These studies, referred to as detection and attribution (D&A) studies, use global climate models to detect changes by varying an exogenous forcing while holding all else constant and to attribute the change to the specific forcings. These studies have been linked to causal counterfactual theory in Hannart et al. (2016) in which authors demonstrate the utility of deriving the probability of necessary and sufficient causality in formulating causal claims. However, when outcomes are not directly observable (e.g., future or paleo climates), causal inference is limited due to lack of accounting for error and uncertainty.

Even when the outcomes are observable, such as in the case of air quality, numerical models are subject to systematic bias arising from misspecification of inputs or processes governing model behavior. To calibrate the CMAQ PM<sub>2.5</sub> output in our study, we develop a Bayesian statistical downscaling method that relates data at a lower observed resolution to a spatially resolved, higher resolution of CMAQ model and allows for spatial prediction at all locations. The calibration model is similar to the spatiotemporal downscaling method introduced by Berrocal, Gelfand and Holland (2010) in that it uses spatially-varying coefficients to estimate the relationship between sparse observations and numeric model output where data is available (Berrocal, Gelfand and Holland (2010), Gelfand et al. (2003, 2004), Schmidt and Gelfand (2003)). The method is computationally efficient and has high predictive performance relative to other downscaling methods (Chilès and Delfiner (2012), Cressie (1993), Fuentes and Raftery (2005)).

The second challenge to estimating wildfire attributed PM<sub>2.5</sub> concentrations within a causal inference paradigm is the spatial interference between the observed PM<sub>2.5</sub> at sites according to whether or not the site (observation unit) is impacted by wildfires (treatment). Indirect or spillover effects across spatial locations violates the stable unit treatment value assumption (SUTVA) which is central to the potential outcomes framework for causal inference (Rubin (1978)). Estimating valid causal effects in the presence of interference has previously

been addressed in the context of vaccines and infectious diseases (Hong and Raudenbush (2006), Hudgens and Halloran (2008), Rosenbaum (2007), Tchetgen Tchetgen and VanderWeele (2012)). Commonly, interference has been addressed using the less stringent partial interference assumption, which assumes interference amongst units in the same group, but not between groups (Halloran and Struchiner (1991), Hudgens and Halloran (2008), Sobel (2006)).

There is a growing body of literature addressing spatial interference (Reich et al. (2021)) and which concerns air pollution in particular (Papadogeorgou, Mealli and Zigler (2019), Zigler and Papadogeorgou (2021)), owing to the fact that causal methods are particularly relevant in air pollution regulatory settings (Dominici, Greenstone and Sunstein (2014), Zigler, Choirat and Dominici (2018)). For instance, Zigler, Dominici and Wang (2012) introduced the first application of spatial models to predict unobserved potential outcomes and estimate the causal effect of air-pollution regulations via principal stratification. To address the interference among units of observation, the authors rely on an assumption of partial interference. However, the correlation between units observed under opposite treatments is unidentifiable under their framework.

The main methodological contribution of this paper is to present a counterfactual framework that utilizes bias-corrected numerical model output to produce valid causal inference and to capture spatial spillover effects. The proposed framework estimates counterfactual outcomes for each day and study site under two treatment regimes: an observed regime with wildfires and an unobservable regime without wildfires. We specify a Bayesian model to fuse the numerical model output with monitor data. We assume that, conditional on numerical model output, observations in the areas not affected by smoke are representative of the counterfactual regime without wildfires. This allows us to bias-correct the CMAQ output from the counterfactual regimes with observed data, a limitation of previous studies. Additionally, by running numerical model simulations under both regimes we are able to estimate correlation between the units observed under different regimes. Finally, we clarify the assumptions required for the estimates to have a causal interpretation and show if these assumptions hold, then the proposed method accounts for spillover effects and that all model parameters are identified.

We apply our method to estimating the effect of wildfires on ambient  $PM_{2.5}$  during 2008–2012 wildfire seasons in the contiguous U.S. We use these estimates to conduct a health-burden analysis. Our Bayesian model provides full uncertainty quantification about the causal effects and resulting health-burden assessment. While we apply the method to the example of wildfire-contributed  $PM_{2.5}$ , the approach is relevant to many applications.

**2. Description of the  $PM_{2.5}$  data.** The analysis of fire-contributed  $PM_{2.5}$  is conducted over the 2008 to 2012 wildfire seasons (May 1–October 31) in the contiguous U.S. There are two sources of  $PM_{2.5}$  data: monitor data from the Environmental Protection Agency’s (EPA) Air Quality System (AQS) and simulated  $PM_{2.5}$  from the CMAQ model. Both data sources cover the contiguous U.S., but because of the large size, we partition the data into nine regions with similar climates and conduct the analysis separately by region (USEPA (2020)). In the Supplementary Material (Section 1) (Larsen et al. (2022)), we show a sensitivity analysis to demonstrate that model results are robust to blocking by region.

**2.1. AQS monitor data.** We use  $PM_{2.5}$  data from Federal Reference Method (FRM) monitoring sites in the U.S. EPA’s AQS monitoring network. Monitors in the AQS network are distributed according to population density, so there are more monitors in urban areas than rural (USEPA (2015)). At each site, daily average concentrations of  $PM_{2.5}$  are measured every one, three, or six days. Figure 1 shows the monitor locations and the  $PM_{2.5}$  observed

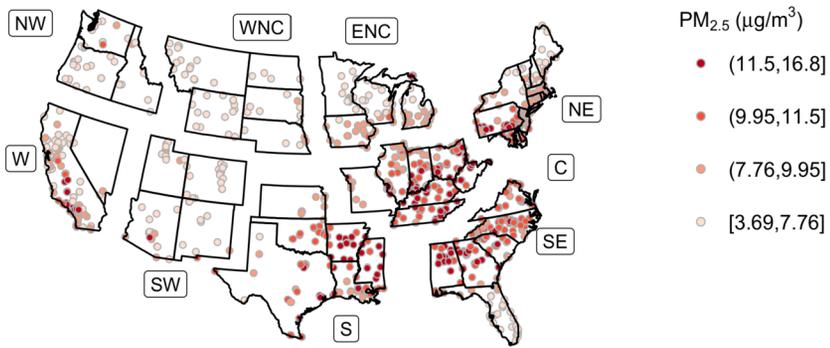


FIG. 1. EPA Air Quality System (AQS) monitor data. The average daily PM<sub>2.5</sub> concentrations ( $\mu\text{g}/\text{m}^3$ ) at EPA AQS monitoring stations over the study period: 2008 to 2012 wildfire seasons. Breakpoints correspond to the 25th, 50th, and 75th percentiles of PM<sub>2.5</sub>. Region names: West (W), Northwest (NW), West North Central (WNC), East North Central (ENC), Northeast (NE), Central (C), Southeast (SE), South (S), and Southwest (SW).

at each monitor in the network averaged over the study time period. The observed average concentrations range from 3.69 to 16.8  $\mu\text{g}/\text{m}^3$ . The highest concentrations of PM<sub>2.5</sub> are in California and the Southeast, and the lowest are in the West North Central and Southwest regions.

**2.2. CMAQ numerical model output.** CMAQ is a deterministic model of air quality which represents the most important processes related to atmospheric chemistry using cutting-edge scientific knowledge. The model utilizes emissions from a wide range of sources and transport by winds to predict concentrations of ambient composition and deposition due to precipitation. CMAQ characterizes production and loss of hundreds of particle and gas phase pollutants (USEPA (2019)). In the case of wildfire emissions, hourly information on fire location and size are determined using satellite information as well as on the ground reports. Wildland fire emissions are estimated, based on the type, load, and conditions of vegetation at the detected burning site, and uses vegetation-specific emission factors (US Forest Service (2015)).

As described below, our causal analysis assumes that CMAQ includes the variables required to explain the association between wildland fires and PM<sub>2.5</sub>. Successive iterations of the model have included the main drivers, but uncertainties remain. The largest known sources of uncertainty arise due to misspecification in characterizing variability in weather patterns and anthropogenic emissions. In the case of wildfires, additional uncertainty arises from misclassification of plume rise, fire weather conditions, and other factors.

The CMAQ-simulated PM<sub>2.5</sub> data is the average PM<sub>2.5</sub> concentrations for each day in the 2008 to 2012 wildfire seasons on a  $12 \times 12$  km grid over the contiguous U.S.; see Rappold et al. (2017) for details. The model is run with and without the forcing for wildland fire emissions. The run without fire emissions is a CMAQ estimate of PM<sub>2.5</sub> in the counterfactual scenario where no wildland fires are possible. The difference in PM<sub>2.5</sub> concentrations between the two runs is a CMAQ estimate of fire-contributed PM<sub>2.5</sub>. CMAQ captures emissions, topology, weather conditions, fate, and transport of air pollution among other factors. However, there are many possible remaining determinants or knowledge gaps that lead to either error and bias which motivates our statistical approach.

Figure 2 displays the PM<sub>2.5</sub> modeled by CMAQ averaged over the 2008 to 2012 wildfire seasons. The western half of the U.S. has predominantly low concentrations of PM<sub>2.5</sub> ( $1.16$ – $2.21$   $\mu\text{g}/\text{m}^3$ ) when fire emissions are excluded, but higher concentrations, when fire emissions are included (up to  $6.78$ – $30.4$   $\mu\text{g}/\text{m}^3$ ). This trend is particularly notable in the West and Northwest regions, where wildfire frequency is high and fire-contributed PM<sub>2.5</sub> comprises

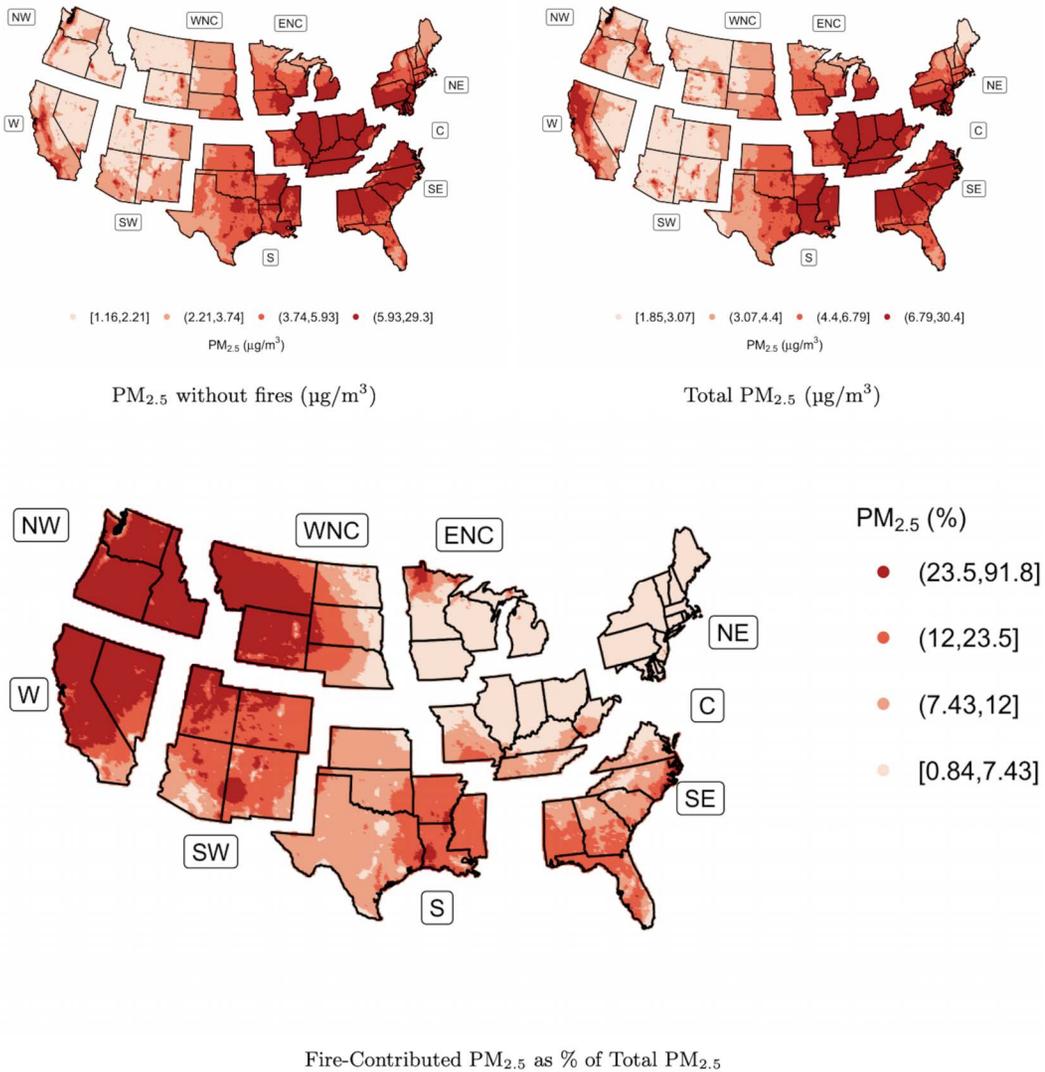


FIG. 2. CMAQ numerical model output. The average daily  $\text{PM}_{2.5}$  concentrations ( $\mu\text{g}/\text{m}^3$ ) over the 2008 to 2012 fire seasons from the CMAQ model. Clockwise from top left: Average daily  $\text{PM}_{2.5}$  from the model without fire emissions, average daily  $\text{PM}_{2.5}$  from the model with fire emissions, and the difference between the models reported as the percentage of the total. The spatial resolution is  $12 \times 12$  km. Region names: West (W), Northwest (NW), West North Central (WNC), East North Central (ENC), Northeast (NE), Central (C), Southeast (SE), South (S), and Southwest (SW).

23.5–91.8% of the total  $\text{PM}_{2.5}$  in parts of these regions (i.e., central and northern California, eastern Oregon and Washington, and central Idaho). In the South and Southeastern regions, contribution of both wildland and prescribed fires is evident.

### 3. Methods.

3.1. *Notation.* We first establish some notation for the data. Let the observed monitor data be  $Y_t(\mathbf{s})$  for day  $t$  and location  $\mathbf{s} \in \mathcal{R}^2$  in spatial domain  $\mathcal{D}$ . We denote CMAQ output from the no-fire run as  $\hat{\theta}_t(\mathbf{s})$  and the difference between the fire and no-fire runs as  $\hat{\delta}_t(\mathbf{s})$ . We denote other environmental factors that are related to both fire activity and  $\text{PM}_{2.5}$  (confounders such as nonfire natural emissions, anthropogenic sources, wind, land type, etc.) as  $X_t(\mathbf{s})$ . The smoke plumes (i.e., the set of locations that is receiving particles emitted by a

fire) associated with the fires determine which locations' air quality are affected by the fires, and so we define  $C_t(\mathbf{s}) = 1$  if site  $\mathbf{s}$  is in a plume on day  $t$  and  $C_t(\mathbf{s}) = 0$  otherwise to capture spillover effects. The collection of data across space is denoted in bold, for example,  $\mathbf{C}_t = \{C_t(\mathbf{s}) : \mathbf{s} \in \mathcal{D}\}$  for the plume indicators.

3.2. *Fire regimes and potential outcomes framework.* To estimate fire-attributable PM<sub>2.5</sub>, that is, the amount of PM<sub>2.5</sub> that would not have occurred were it not for wildland fires, we use a potential outcomes framework (Rubin (1978)). Our objective is to estimate the average difference in PM<sub>2.5</sub> between the status quo where there are fires, and thus  $C_t(\mathbf{s})$  takes on values of 0 and 1, following the fires that occur, with the counterfactual case of no fires and  $C_t(\mathbf{s}) = 0$  for all  $\mathbf{s}$  and  $t$ . Because our interest is in comparing these two cases rather than all possible fire/plume combinations, we define regimes  $R = 1$  and  $R = 0$  to summarize the plume distribution over the entire spatial domain  $\mathcal{D}$ : the fire regime ( $R = 1$ ), under which wildfires occur in  $\mathcal{D}$ , and the no-fire regime ( $R = 0$ ) under which fires do not occur anywhere in  $\mathcal{D}$ . We also define the potential PM<sub>2.5</sub> under regimes  $R = 0$  and  $R = 1$  as  $Y_t(\mathbf{s}, 0)$  and  $Y_t(\mathbf{s}, 1)$ , respectively, and model each as

$$(1) \quad \begin{aligned} Y_t(\mathbf{s}, 0) &= \theta_t(\mathbf{s}), \\ Y_t(\mathbf{s}, 1) &= \theta_t(\mathbf{s}) + \delta_t(\mathbf{s}), \end{aligned}$$

where  $\theta_t(\mathbf{s})$  and  $\delta_t(\mathbf{s})$  are stochastic processes representing nonfire and fire-attributed PM<sub>2.5</sub>, respectively.

Under regime  $R = 1$ , the data-generating process for the potential outcomes is to sample  $\mathbf{X}_t, \mathbf{C}_t|\mathbf{X}_t$ , and, finally,  $\mathbf{Y}_t(1)|\mathbf{X}_t, \mathbf{C}_t$  in sequence. Similarly, under regime  $R = 0$ , the potential outcome are generated by sampling  $\mathbf{X}_t$  and then  $\mathbf{Y}_t(0)|\mathbf{X}_t$  setting  $\mathbf{C}_t = 0$ . Therefore, the amount of PM<sub>2.5</sub> caused by wildland fires is quantified by the average (over  $\mathbf{X}_t$  and  $\mathbf{C}_t$ ) difference in the potential outcomes (Hernán et al. (2008), Holland (1986), Rubin (1978)),

$$\Delta(\mathbf{s}) = E[Y_t(\mathbf{s}, 1) - Y_t(\mathbf{s}, 0)] = E[\delta_t(\mathbf{s})].$$

In our analysis we average over time throughout the entire fire season and years of the study, although this framework could be applied to estimate the causal effect annually, seasonally, or even daily.

3.3. *Assumptions.* The fundamental problem in causal inference is that not all potential outcomes are observed for each  $\mathbf{s}$  and  $t$  (Holland (1986)). Therefore, the potential outcome models and the causal effect,  $\Delta(\mathbf{s})$ , are not identifiable without satisfying SUTVA, the components of which we discuss in the following.

We assume there exist bias-correction functions,  $B_0$  and  $B_1$ , and we observe binary indicator  $C_t(\mathbf{s}) \in \{0, 1\}$  where  $\mathbf{s}$  is affected by wildfire smoke if and only if  $C_t(\mathbf{s}) = 1$  so that the following hold.

ASSUMPTION 1 (Potential Outcomes Model). The counterfactual processes can be decomposed as

$$\theta_t(\mathbf{s}) = B_0(\hat{\theta}_t(\mathbf{s}), \mathbf{s}) + e_{0t}(\mathbf{s}) \quad \text{and} \quad \delta_t(\mathbf{s}) = B_1(\hat{\delta}_t(\mathbf{s}), \mathbf{s}) + e_{1t}(\mathbf{s}),$$

where  $[e_{0t}(\mathbf{s}), e_{1t}(\mathbf{s})]$  is a bivariate spatial process independent of  $\mathbf{X}_t$  and  $\mathbf{C}_t$ , given  $\hat{\theta}_t$  and  $\hat{\delta}_t$ .

This assumption relates the CMAQ estimates  $\hat{\theta}_t(\mathbf{s})$  and  $\hat{\delta}_t(\mathbf{s})$  with the true processes  $\theta_t(\mathbf{s})$  and  $\delta_t(\mathbf{s})$ . The bias correction functions  $B_0$  and  $B_1$  can be flexible nonlinear functions (e.g., splines) and vary by spatial location, and the discrepancy terms  $e_{0t}$  and  $e_{1t}$  account for model

misspecification and are modeled as spatial processes (Kennedy and O’Hagan (2001)). To allow for learning where the model is underperforming, relative to the truth, we model CMAQ bias function  $B_j$  to be a flexible spatially varying surface. As such, this bias function enables us to gain insights into possible spatially-varying confounding by examining the residual variation. If it is suspected that model error depends on  $\mathbf{X}_t$  (or a subset of  $\mathbf{X}_t$ ), then this assumption could be relaxed by including  $\mathbf{X}_t$  in the model as, for example,  $\delta_t(\mathbf{s}) = B_1(\hat{\delta}_t(\mathbf{s}), \mathbf{X}_t, \mathbf{s}) + e_{1t}(\mathbf{s})$ .

Equation (1) and Assumption 1 specify the full joint model between counterfactual outcomes  $Y_t(\mathbf{s}, 0)$  and  $Y_t(\mathbf{s}, 1)$ , given  $[\hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})]$ . Under Assumption 1,  $[Y_t(\mathbf{s}, 0), Y_t(\mathbf{s}, 1)]$  are independent of  $\mathbf{X}_t$ , given  $[\hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})]$ . Therefore,  $[\hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})]$  can be called the prognostic score of Hansen (2008) which is the prognostic analogue of the propensity score. Also, Assumption 1 implies that all confounders with the regime realizations and potential outcomes are captured through CMAQ output, since Assumption 1 implies  $[\theta_t(\mathbf{s}), \delta_t(\mathbf{s})]$ , and thus  $[Y_t(\mathbf{s}, 0), Y_t(\mathbf{s}, 1)]$  are independent of  $\mathbf{C}_t$ , given  $[\hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})]$ . This is similar to the unconfounded network influence assumption of Kao (2017) under the social network framework. Although Assumption 1 is key for identification and dramatically simplifies the analysis, compared to modeling the effect of  $\mathbf{X}_t$ , it cannot be verified from the observed data. Because CMAQ uses most important meteorological and environmental factors for fire activity, smoke transportation, and  $\text{PM}_{2.5}$  as well as state-of-the-art computer simulations (see Section 2.2), this assumption is plausible.

**ASSUMPTION 2 (Consistency).** Ignoring measurement errors, the observation at  $\mathbf{s}$  equals the potential outcome at  $\mathbf{s}$  under regime given by  $C_t(\mathbf{s})$ ,

$$Y_t(\mathbf{s}) = \begin{cases} Y_t(\mathbf{s}, 0) & \text{if } C_t(\mathbf{s}) = 0, \\ Y_t(\mathbf{s}, 1) & \text{if } C_t(\mathbf{s}) = 1. \end{cases}$$

Assumption 2 links the potential outcomes with the observed outcomes. In particular, it allows for partial realizations of  $Y_t(\mathbf{s}, 0)$ , removing the need to actualize a situation under the counterfactual no-fire regime. For example, this assumption implies that a set of monitors far removed from fires and plumes on day  $t$  can be assumed to follow the potential outcomes distribution under the no fires regime and thus be used to identify parameters in this distribution, such as those that determine  $B_0$  and the spatial covariance of  $e_{0t}(\mathbf{s})$ . As long as an appropriate variable  $C_t(\mathbf{s})$  can be identified from the observed data, this assumption is plausible. In our analysis we use CMAQ output to estimate  $C_t(\mathbf{s})$ . Namely, we set  $C = 1(\hat{\delta} > \tau)$ , where  $\tau$  is a fixed threshold chosen through cross-validation and sensitivity analysis.

Theorem 1 gives the main identification result with the proof deferred to the Appendix.

**THEOREM 1.** *Under Assumptions 1 and 2 and further assuming that  $\mathbf{C}_t$  is not degenerate, the parameters in the potential outcome models are identifiable via the distribution of the observed data, that is, the distribution of  $\mathbf{Y}_t$ , given  $(\hat{\theta}_t, \hat{\delta}_t, \mathbf{C}_t)$ .*

While we never observe complete  $\mathbf{Y}_t$  (i.e., for all  $\mathbf{s}$ ) under the no fires regime, Assumptions 1 and 2, along with the nondegeneracy of  $\mathbf{C}_t$ , are sufficient for identification. By Theorem 1, causal parameter estimation only requires inspecting the implied model for  $Y_t(\mathbf{s})$  and confirming parameter identification. In Section 3.4 we specify parametric models for the bias correction functions  $B_0$  and  $B_1$  and the spatial process  $\mathbf{e}_t(\mathbf{s}) = [e_{0t}(\mathbf{s}), e_{1t}(\mathbf{s})]^T$ . We then argue in Section 3.5 that all parameters, including the correlation between counterfactuals, are identifiable in our spatial setting. This setup serves as a basis for using a Bayesian approach to estimating the causal effect,  $\Delta(\mathbf{s})$ .

Defining the intervention as the fire regime instead of individual fires ( $C_t(\mathbf{s})$ ) is key for two reasons. First, this is parallel to how the numerical model simulates fire and no-fire PM<sub>2.5</sub>. Second, the amount of fire-contributed PM<sub>2.5</sub> at any site in the spatial domain depends on the fire status at other sites, because the smoke from fires at neighboring sites is transported. This is called interference or spillover, and it is problematic because we could not reasonably claim that changes in PM<sub>2.5</sub> at site  $\mathbf{s}$  were only due to fire presence or absence at site  $\mathbf{s}$ , that is,  $Y_t(\mathbf{s}, C_t(\mathbf{s}))$  is not well defined. There would be a different potential outcome for every possible  $C_t$ , resulting in  $2^n$  potential outcomes per site for a spatial domain containing  $n$  sites.

3.4. *Bayesian hierarchical model.* Assumption 1 and the addition of measurement error give the following model for the observed PM<sub>2.5</sub>:

$$(2) \quad Y_t(\mathbf{s}) = \theta_t(\mathbf{s}) + C_t(\mathbf{s})\delta_t(\mathbf{s}) + \epsilon_t(\mathbf{s}),$$

where  $\epsilon_t(\mathbf{s}) \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$  are measurement errors. To separate background PM<sub>2.5</sub> from fire-contributed PM<sub>2.5</sub>, we assign priors to  $\theta_t(\mathbf{s})$  and  $\delta_t(\mathbf{s})$  based on bias-adjusted CMAQ runs, as per Assumption 1. We model the means of both processes as

$$B_j(z, \mathbf{s}) = \alpha_j(\mathbf{s}) + \beta_j(\mathbf{s})z$$

for  $j = 0, 1$ , where  $\alpha_j(\mathbf{s})$  is the additive bias and  $\beta_j(\mathbf{s})$  is the multiplicative bias. The bias terms have Gaussian process priors with means  $E[\alpha_j(\mathbf{s})] = \mu_{\alpha_j}$  and  $E[\beta_j(\mathbf{s})] = \mu_{\beta_j}$  and covariances  $\text{Cov}[\alpha_j(\mathbf{s}), \alpha_j(\mathbf{s}')] = \sigma_{\alpha_j}^2 \exp(-\|\mathbf{s} - \mathbf{s}'\|/\phi_2)$  and  $\text{Cov}[\beta_j(\mathbf{s}), \beta_j(\mathbf{s}')] = \sigma_{\beta_j}^2 \exp(-\|\mathbf{s} - \mathbf{s}'\|/\phi_2)$ .

The background and fire-contributed PM<sub>2.5</sub> then have the following form:

$$(3) \quad \begin{aligned} \theta_t(\mathbf{s}) &= \alpha_0(\mathbf{s}) + \beta_0(\mathbf{s})\hat{\theta}_t(\mathbf{s}) + e_{0t}(\mathbf{s}), \\ \delta_t(\mathbf{s}) &= \alpha_1(\mathbf{s}) + \beta_1(\mathbf{s})\hat{\delta}_t(\mathbf{s}) + e_{1t}(\mathbf{s}), \end{aligned}$$

where  $\mathbf{e}_t(\mathbf{s}) = [e_{0t}(\mathbf{s}), e_{1t}(\mathbf{s})]^T$  is a bivariate spatial process with mean  $E[e_{jt}(\mathbf{s})] = 0$  and separable exponential covariance  $\text{Cov}[\mathbf{e}_t(\mathbf{s}), \mathbf{e}_t(\mathbf{s}')] = \Sigma \exp(-\|\mathbf{s} - \mathbf{s}'\|/\phi_1)$ . The  $2 \times 2$  cross-covariance matrix  $\Sigma$  has diagonal elements  $\sigma_1^2$  and  $\sigma_2^2$ , and off-diagonal element  $\sigma_{12} = \sigma_1\sigma_2\gamma$ , so  $\gamma$  gives the correlation between counterfactual outcomes.

Our Bayesian analysis requires prior distributions for the model parameters which are given in the Supplementary Material (Section 2) along with a sensitivity analysis that suggests the results are not sensitive to prior specification. Exploration of the spatial model goodness of fit and assumption of temporal independence are provided in the Supplementary Material (Section 3 and 4).

3.5. *Estimability.* In this section we argue that all parameters in the joint model specified above are estimable.

First, consider the parameters in the mean,

$$(4) \quad E[Y_t(\mathbf{s})] = \mu_t(\mathbf{s}) = \alpha_0(\mathbf{s}) + \beta_0(\mathbf{s})\hat{\theta}_t(\mathbf{s}) + \alpha_1(\mathbf{s})C_t(\mathbf{s}) + \beta_1(\mathbf{s})[C_t(\mathbf{s})\hat{\delta}_t(\mathbf{s})].$$

Assuming the four covariates  $(1, \hat{\theta}_t(\mathbf{s}), C_t(\mathbf{s}), C_t(\mathbf{s})\hat{\delta}_t(\mathbf{s}))$  are not linearly dependent at  $\mathbf{s}$ , then the four parameters  $\alpha_0(\mathbf{s}), \alpha_1(\mathbf{s}), \beta_0(\mathbf{s}),$  and  $\beta_1(\mathbf{s})$  are estimable. For example, ordinary least squares would provide an unbiased and consistent estimator (as the number of days increases). This result clearly relies on Assumption 1, or it would not be possible to identify both  $\alpha_0(\mathbf{s})$  and  $\alpha_1(\mathbf{s})$ .

Under the model  $Y_t(\mathbf{s}) = \mu_t(\mathbf{s}) + e_{0t}(\mathbf{s}) + C_t(\mathbf{s})e_{1t}(\mathbf{s}) + \epsilon_t(\mathbf{s})$ , the covariance is

$$(5) \quad \text{Cov}[Y_t(\mathbf{s}), Y_t(\mathbf{s}') | \hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})] = \begin{cases} \sigma_1^2 \exp(-h/\phi_1) & C_t(\mathbf{s}) = C_t(\mathbf{s}') = 0, \\ \sigma_1^2 \left(1 + \frac{\sigma_2^2}{\sigma_1} \gamma\right) \exp(-h/\phi_1) & C_t(\mathbf{s}) \neq C_t(\mathbf{s}'), \\ (\sigma_1^2 + 2\sigma_1\sigma_2\gamma + \sigma_2^2) \exp(-h/\phi_1) & C_t(\mathbf{s}) = C_t(\mathbf{s}') = 1, \end{cases}$$

where  $h = \|\mathbf{s} - \mathbf{s}'\|$ . By examining the spatial correlation between pairs of points separately, according to their values of  $C_t(\mathbf{s})$ , the parameters  $\sigma_1^2$ ,  $\sigma_2^2$ ,  $\gamma$ , and  $\phi_1$  are estimable. For example, simple variogram-based methods could be used to estimate these parameters. More importantly, under the full Bayesian model,  $\text{Cov}[Y_t(\mathbf{s}, 0), Y_t(\mathbf{s}, 1) | Y_t(\mathbf{s})] = \sigma_1^2(1 + \frac{\sigma_2^2}{\sigma_1} \gamma)$  is estimable, although  $Y_t(\mathbf{s}, 0)$  and  $Y_t(\mathbf{s}, 1)$  are never jointly observed at one location. In our analysis we use Bayesian modeling to jointly estimate the mean and covariance parameters.

3.6. *Posterior inference.* The causal effect at  $\mathbf{s}$  is approximated as

$$(6) \quad \Delta(\mathbf{s}) \approx \frac{1}{T} \sum_{t=1}^T C_t(\mathbf{s}) \delta_t(\mathbf{s}).$$

Our fully Bayesian analysis produces the entire posterior distribution of the causal effect, including the posterior mean  $\bar{\Delta}(\mathbf{s}) = \frac{1}{T} \sum_{t=1}^T C_t(\mathbf{s}) \bar{\delta}_t(\mathbf{s})$ , where  $\bar{\delta}_t(\mathbf{s})$  is the posterior mean of  $\delta_t(\mathbf{s})$ . The estimate,  $\bar{\delta}_t(\mathbf{s})$ , includes both bias-corrected CMAQ,  $B_0(\hat{\theta}_t(\mathbf{s}), \mathbf{s})$ , and  $B_1(\hat{\delta}_t(\mathbf{s}), \mathbf{s})$ , and observed data,  $Y_t(\mathbf{s})$ , allowing us to account for any daily variation in fire-attributable  $\text{PM}_{2.5}$  not captured by CMAQ. Finally,  $\bar{\delta}_t(\mathbf{s})$  is based on estimable parameters, defined in the previous section, making  $\bar{\Delta}(\mathbf{s})$  an estimable quantity as well.

We multiply  $\bar{\delta}_t(\mathbf{s})$  by  $C_t(\mathbf{s})$  because, given Assumption 2, this relates the observations to the potential outcomes, thereby imparting the causal interpretation on  $\bar{\Delta}(\mathbf{s})$ . Multiplying by  $C_t(\mathbf{s})$  also allows the model to only identify  $\delta_t(\mathbf{s})$  as a causal quantity if  $C_t(\mathbf{s}) = 1$ , which is important, as we are not interested in the  $\text{PM}_{2.5}$  if there were fires affecting  $\mathbf{s}$  every day, but the causal estimate if the fires we observed were removed.

Assuming conditional independence of  $C_t(\mathbf{s})$  and  $\delta_t(\mathbf{s})$ , given  $Y_t(\mathbf{s}), \hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})$  over time,  $\bar{\Delta}(\mathbf{s})$  satisfies

$$\begin{aligned} E[\bar{\Delta}(\mathbf{s})] &= E[C_t(\mathbf{s}) \bar{\delta}_t(\mathbf{s})] \\ &= E[C_t(\mathbf{s}) E[\delta_t(\mathbf{s}) | Y_t(\mathbf{s}), \hat{\theta}_t(\mathbf{s}), \hat{\delta}_t(\mathbf{s})]] \\ &= E[C_t(\mathbf{s}) \delta_t(\mathbf{s})]. \end{aligned}$$

Hence, it is reasonable to use  $\bar{\Delta}(\mathbf{s})$  to approximate the causal effect.

3.7. *Computation.* To approximate the posterior of the causal effect  $\Delta(\mathbf{s})$ , we implement the spatial Bayesian analysis using a Markov chain Monte Carlo sampling. The missing values of observed  $\text{PM}_{2.5}$  are imputed, and every model parameter is iteratively updated by the algorithm, conditional on all other parameters. The spatial range parameters,  $\phi_1$  and  $\phi_2$ , are estimated empirically using variograms. All other model parameters have conditionally-conjugate priors and are accordingly updated with Gibbs steps where each step samples from their respective full conditional distributions (see Supplementary Material Section 5 for derivations of the full conditional distributions). We use Gaussian Kriging to estimate smooth spatial surfaces across each study region for both the posterior means and standard deviations of each model parameter. We Kriged each estimate to the centroids of the  $12 \times 12$  km CMAQ

grid. Our MCMC has a burn-in period of length 5000, after which we collect samples every 100 iterations until a total of 30,000 iterations have been completed. To verify that the MCMC algorithm converged, we computed the effective sample size of the causal effect estimate,  $\Delta(\mathbf{s})$ , for each  $\mathbf{s}$ . We also monitored convergence using visual inspection of trace plots for several representative parameters. Summary statistics and figures of the effective sample sizes and trace plots are included in the Supplementary Material (Section 6).

**4. Fire-contributed PM<sub>2.5</sub> estimates.** We let  $C_t(\mathbf{s}) = I[\hat{\delta}_t(\mathbf{s}) > \tau]$ , where  $\hat{\delta}_t(\mathbf{s})$  is the CMAQ estimate of fire-attributed PM<sub>2.5</sub> and  $\tau$  is a fixed threshold. To select the threshold, we ran several models for a range of values of  $\tau$  and used five-fold cross-validation to evaluate each model's ability to predict total PM<sub>2.5</sub>. We found little variation between the prediction metrics between each model. For example, mean-squared error (MSE) ranged from 12.58  $\mu\text{g}/\text{m}^3$  ( $\tau = 1 \mu\text{g}/\text{m}^3$ ) to 12.71  $\mu\text{g}/\text{m}^3$  ( $\tau = 5 \mu\text{g}/\text{m}^3$ ) (Supplementary Material, Section 7). We also examined variation in the causal effect when estimated with different values of  $\tau$  and found the differences to be negligible except if  $\tau$  is selected to be extreme (e.g., 0 or 10  $\mu\text{g}/\text{m}^3$ ) (Supplementary Material, Section 7). Based on these findings, we concluded that the model is robust to moderate choices for the threshold, and we let  $\tau = 1 \mu\text{g}/\text{m}^3$  for the remaining analysis.

We display the posterior mean and standard deviation of the bias parameter estimates for each region in the analysis in the Supplementary Material (Section 8). For the multiplicative bias parameter in the fire-contribution process,  $\beta_1(\mathbf{s})$ , the highest *second* percent of  $\beta_1(\mathbf{s})$  values reached (0.991, 2.01  $\mu\text{g}/\text{m}^3$ ), meaning that the strongest estimated association between CMAQ estimates and the monitor data occurs in the Northwest and West North Central (WNC) regions, along with parts of the East North Central region, the Southwest, and parts of the Southeast region. The lowest values (−0.39, −0.018  $\mu\text{g}/\text{m}^3$ ) in the northern part of the East North Central region, the South, and parts of the Northeast region. These have fewer wildfires (Figure 2), and thus it is more difficult for CMAQ to estimate the relationship between model-estimated contribution and observed PM<sub>2.5</sub>. This is neither surprising nor problematic because these regions rarely experience fire smoke.

The posterior mean of the correlation between the counterfactual processes is summarized in the Supplementary Material (Section 9). Observing a positive correlation in a given region is indicative of fire smoke occurring in areas where nonfire PM<sub>2.5</sub> emissions are present. A negative correlation indicates the converse. The highest estimated correlation is in the West region ( $0.44 \pm 0.05$ ), followed by the WNC region ( $0.31 \pm 0.15$ ), and then the South, Central and Southeast regions ( $0.26 \pm 0.08$ ,  $0.26 \pm 0.06$ ,  $0.25 \pm 0.02$ ). The Northeast region exhibits low correlation ( $0.16 \pm 0.04$ ). The correlation estimate for the Southwest region ( $-0.21 \pm 0.06$ ) is negative, and the only areas for which the correlation was plausibly zero were the ENC and the Northwest regions. To further illustrate the spatial correlation between observations, we also provide plots of equation (5) evaluated at the posterior mean of the model parameters for each region and combination of  $C_t(\mathbf{s})$  in the Supplementary Material (Section 9).

Figure 3 displays the causal effect estimates (Panel a), posterior standard deviation (Panel b), and the causal effect as percent of total estimated PM<sub>2.5</sub> (Panel c). The largest estimates occur in the West, Northwest, and WNC regions, where wildfires are most prevalent. In these areas, between 29.5% and 72.9% of PM<sub>2.5</sub> is attributable to wildfire smoke (Figure 3c). Moderate effects are estimated in areas of the South and Southeast, where prescribed burning is prevalent. The causal estimates in the East North Central region are in both the top and bottom two percent of fire-contributed PM<sub>2.5</sub>. This area is typically only effected by long-range smoke transport from the western U.S. or from Canadian wildland fires further north. Large areas of the Northeast region have estimates near zero (some locations have very small negative values, likely due to statistical uncertainty).

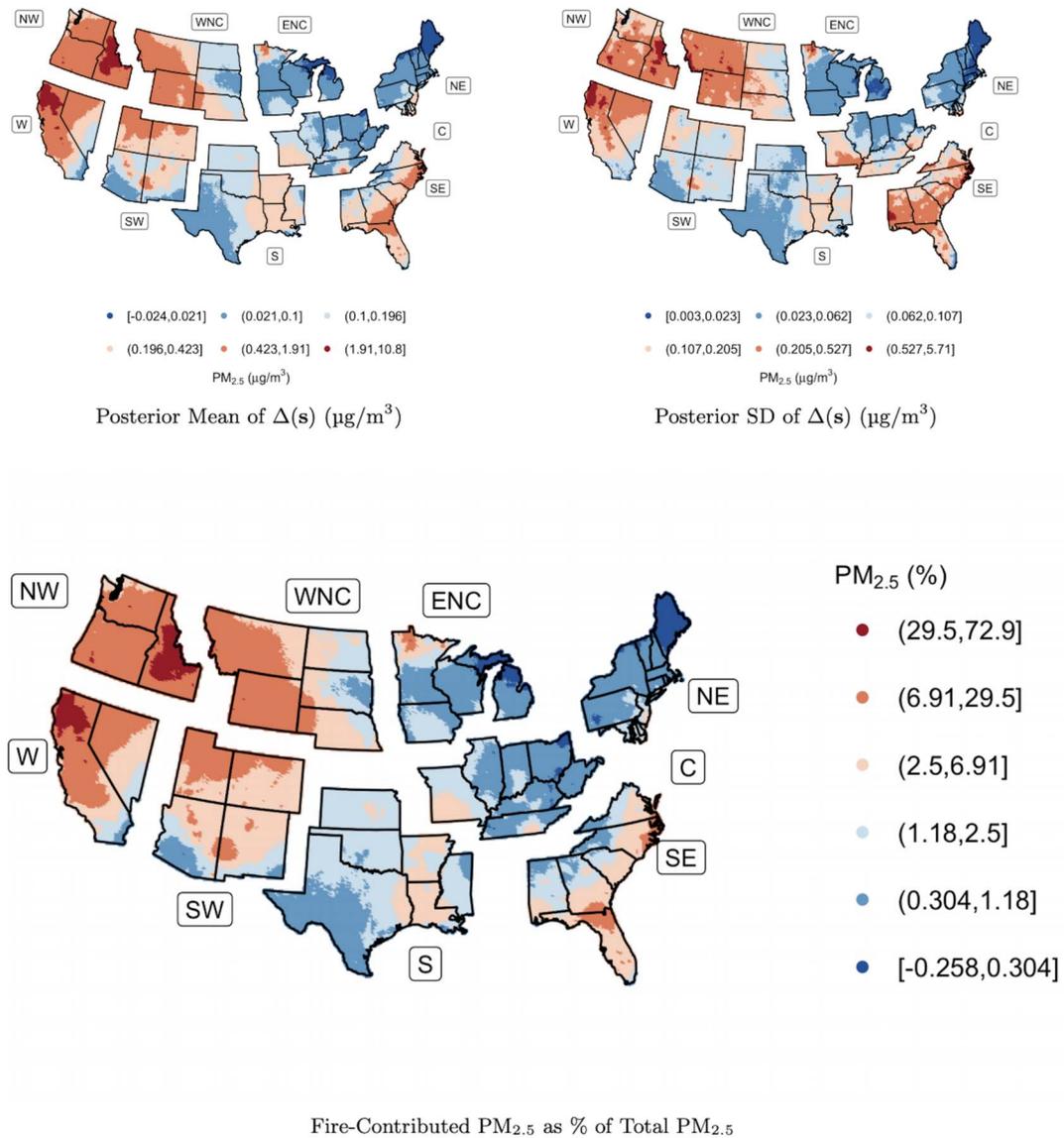


FIG. 3. Estimates of the causal effect,  $\Delta(s)$ . Clockwise from top left: Posterior mean, posterior standard deviation, fire-contributed  $\text{PM}_{2.5}$  as a percent of total  $\text{PM}_{2.5}$ . Estimates are average daily concentrations over the 2008 to 2012 wildfire seasons at a  $12 \times 12$  km spatial resolution. Region names: West (W), Northwest (NW), West North Central (WNC), East North Central (ENC), Northeast (NE), Central (C), Southeast (SE), South (S), and Southwest (SW).

Figure 3 shows Bayesian model estimates of background and total  $\text{PM}_{2.5}$ , CMAQ-simulated background, and total  $\text{PM}_{2.5}$  as well as observed  $\text{PM}_{2.5}$  during the 2008 wildfire season. Although the spatial pattern of the causal estimates resembles the CMAQ estimates, there are notable differences in the range of the estimates. Figure 4 illustrates these differences at one site on Northern California. The estimates from the Bayesian causal model tend to fit closely to the observed values of  $\text{PM}_{2.5}$  from the monitor rather than to the CMAQ-simulated total  $\text{PM}_{2.5}$  and that the CMAQ model estimates are, on average, much higher.

We also compare the estimates from the Bayesian causal model to those from CMAQ for all monitoring sites (Figure 5) and the prediction sites (Supplementary Material, Section 10). As in Figure 4, the Bayesian model generally produces lower estimates of fire-contributed  $\text{PM}_{2.5}$  than CMAQ at all regions, both at monitoring sites (Figure 5) and prediction sites.

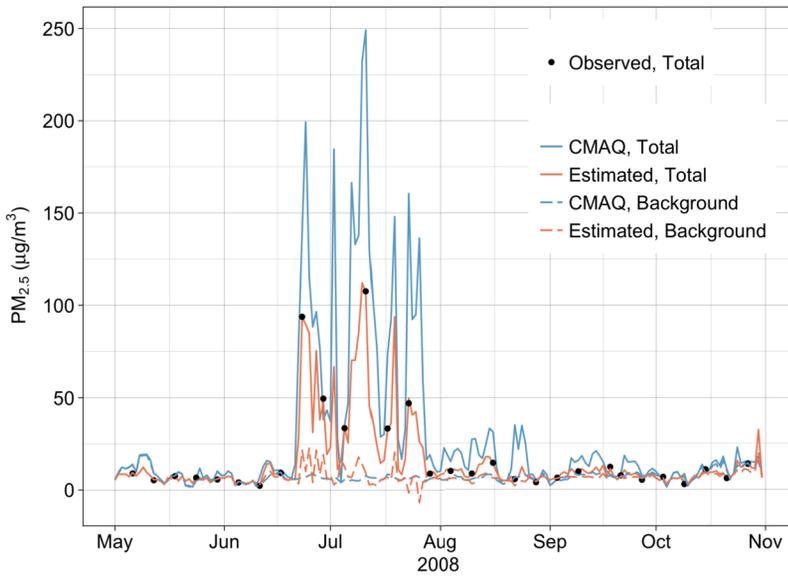


FIG. 4. PM<sub>2.5</sub> data and estimates at a monitoring site in Northern California during the 2008 wildfire season. The site in Northern California ( $-121.8^\circ$ ,  $39.8^\circ$ ) was impacted by a wildfire episode in the summer of 2008. Data and estimates include observed station measurements, total and background estimates from CMAQ ( $\hat{\theta}_T(s) + \hat{\delta}_T(s)$ ,  $\hat{\theta}_T(s)$ ), total and background estimates from the Bayesian model (posterior means of  $\theta_T(s) + C_T(s)\Delta_T(s)$ , and  $\theta_T(s)$ , respectively).

The 95% credible intervals are longer at the prediction locations than at the monitoring sites which is to be expected in an interpolation analysis. Additionally, only in regions where fires are prevalent (e.g., West, Northwest, WNC) do we see causal-effect estimates significantly different from zero.

**5. Health burden analysis.** We use a log-linear concentration-response function to describe the relationship between PM<sub>2.5</sub> and the number of hospitalizations due to respiratory illness. This analysis is conducted at the county level as well as by age group,  $a$ . We define  $\Delta_c$  as the integrated causal effect  $\Delta(s)$  for  $s$  in county  $c$ . The health impact function relating fire-contributed PM<sub>2.5</sub> to changes in the incidence rate of hospitalizations due to respiratory illness is

$$R_{ac} = r_a^0 n_c (e^{r_a \Delta_c} - 1),$$

where  $n_c$  is the population of county  $c$  based on the July 2010 U.S. Census and  $r_a^0$  is the incidence rate of hospitalizations for respiratory illness by county and age group (BenMAP (2017)). Using  $r_a$ , we calculate cumulative daily burden over all days in the study (May–October, 2008–2012) by county and age group (Supplementary Material, Section 11) (Delfino et al. (2009)). Cumulative  $R_{ac}$  over all counties in each region is summarized in Table 1, based on both the Bayesian and the CMAQ estimate of fire-contributed PM<sub>2.5</sub>. We compute 95% Bayesian credible intervals for each cumulative  $R_{ac}$  using Monte Carlo sampling over the uncertainty distribution of  $r_a$  implied by the standard errors of the relative risk estimates in Delfino et al. (2009) and the posterior distribution of the PM<sub>2.5</sub> effects  $\Delta_c$ . We note that these estimates have a causal interpretation only if the estimates in Delfino et al. (2009) have a causal interpretation. While Delfino et al. (2009) account for many known confounders for fire-contributed PM<sub>2.5</sub> and respiratory illness and the U.S. EPA (2010) declares that the adverse effects of short-term PM<sub>2.5</sub> exposure on respiratory outcomes is likely to be causal (using the Hill criteria), this remains an important caveat.

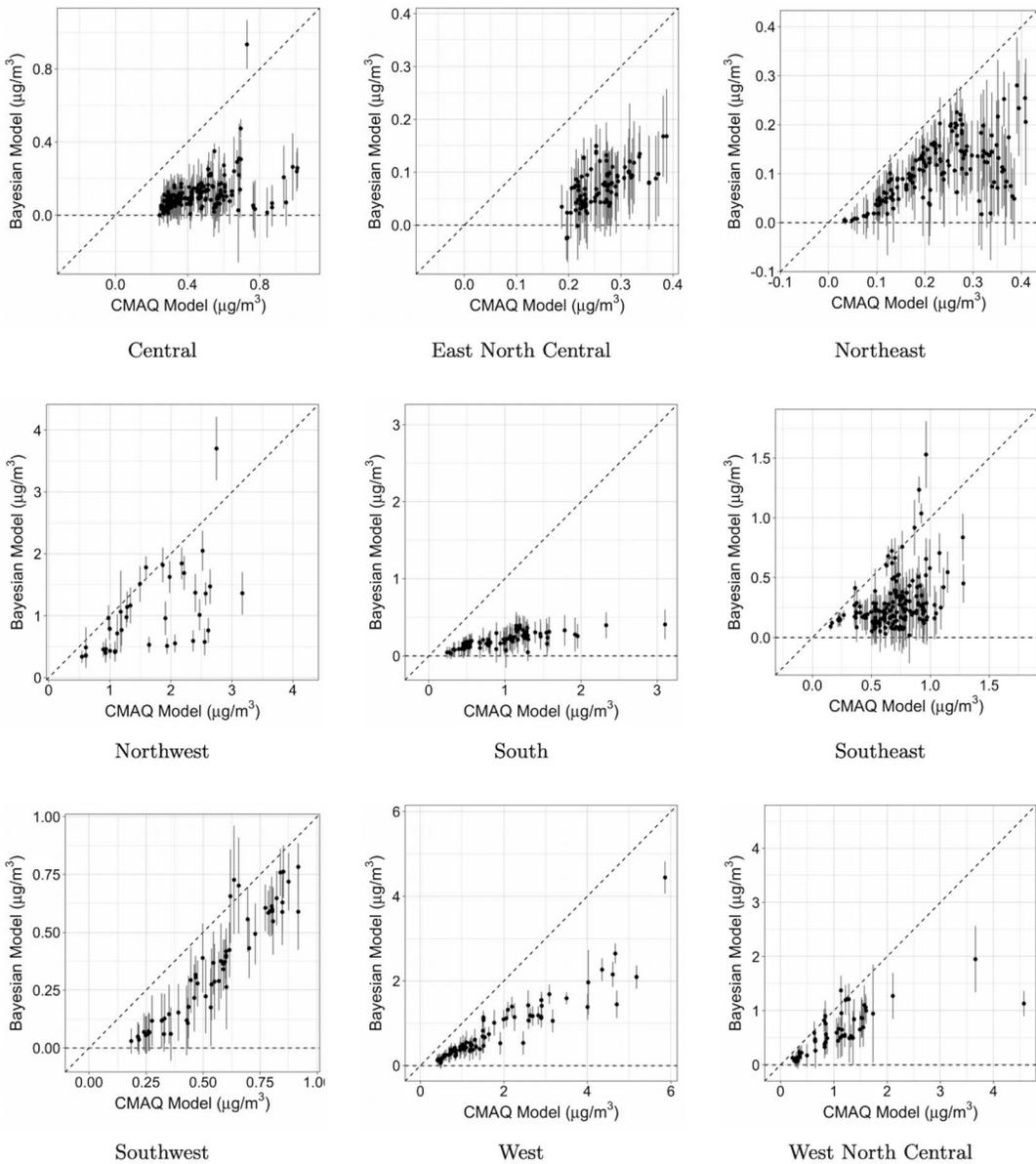


FIG. 5. Causal estimates and credible intervals at monitoring stations. Fire-contributed  $PM_{2.5}$  from the Bayesian model ( $\Delta(s)$ ) vs. the CMAQ model ( $\hat{\delta}(s)$ ) at the AQS monitoring sites. Vertical error bars denote 95% credible intervals. The dashed lines represent  $x = y$  and  $y = 0$ .

The Bayesian estimate yields more conservative estimates of the impact of fire-contributed  $PM_{2.5}$  on hospital admission rates for respiratory illness than the CMAQ-only analysis. The highest estimated burden is observed in the West region, notably in Southern California with upwards of 300 hospitalizations estimated cumulatively over the 2008 to 2012 fire seasons in some counties (Supplementary Material, Section 11). In Table 1 the highest estimated burden for any region is in the West with 1513.9 hospitalizations over the 2008–2012 fire seasons, using the Bayesian estimate of the causal effect. If the CMAQ estimate for the causal effect is used, the cumulative burden in the West is estimated to be 3500.4 hospitalizations per day. Most counties in the rest of the country exhibit lower burden with less than five hospitalizations per county over the 2008–2012 fire seasons.

TABLE 1

Number of hospitalizations in each region. Cumulative number of hospitalizations for respiratory illness due to wildland fires over the 2008–2012 fire seasons in each region, calculated using the Bayesian and the CMAQ model estimates of the causal effect,  $\Delta_t(s)$ , by region. 95% confidence intervals are provided. Model 1 is the Bayesian model; model 2 is the CMAQ model. C = Central, ENC = East North Central, S = South, SE = Southeast, SW = Southwest, NE = Northeast, NW = Northwest, W = West, WNC = West North Central

Reg.	Mod.	Age group (years)				
		0–1	2–34	35–64	65–99	0–99
C	1	150.5 (33.9, 270.6)	60.1 (–19.5, 146.6)	159.8 (33.5, 290.7)	283.7 (104.8, 460.1)	654.1 (152.7, 1168.0)
	2	612.8 (137.7, 1103.6)	242.3 (–77.9, 592.1)	663.5 (139.1, 1208.7)	1161.1 (428.4, 1884.9)	2679.8 (627.4, 4789.3)
ENC	1	26.7 (6.0, 48.2)	10.4 (–3.4, 25.4)	28.5 (6.0, 52.0)	56.3 (20.8, 91.4)	121.9 (29.3, 217.1)
	2	95.4 (21.4, 172.3)	37.1 (–12.0, 91.1)	103.7 (21.7, 189.1)	207.7 (76.6, 337.5)	443.9 (107.7, 790.0)
S	1	134.9 (30.4, 242.1)	47.8 (–15.4, 116.2)	127.3 (26.7, 231.5)	248.9 (92.0, 403.4)	558.9 (133.8, 993.1)
	2	604.0 (135.1, 1093.7)	211.0 (–67.2, 518.7)	562.1 (117.5, 1027.1)	1102.0 (405.4, 1794.2)	2479.1 (590.8, 4433.7)
SE	1	279.3 (62.4, 506.0)	118.7 (–37.1, 290.9)	324.8 (67.9, 593.3)	565.6 (208.2, 920.5)	1288.3 (301.4, 2310.6)
	2	642.5 (144.1, 1160.6)	272.7 (–86.8, 667.4)	746.6 (156.4, 1361.3)	1284.4 (473.4, 2087.2)	2946.1 (687.1, 5276.5)
SW	1	113.6 (25.5, 205.1)	33.8 (–13.1, 85.4)	49.1 (10.3, 89.4)	85.6 (31.6, 139.1)	282.1 (54.2, 519.2)
	2	183.6 (41.2, 330.8)	57.3 (–21.7, 144.0)	89.3 (18.7, 162.7)	157.0 (58.0, 254.9)	487.3 (96.2, 892.4)
NE	1	116.5 (26.1, 210.1)	51.1 (–18.0, 127.0)	118.4 (24.8, 215.7)	231.6 (85.4, 376.2)	517.6 (118.4, 929.1)
	2	209.9 (47.3, 377.1)	93.2 (–32.2, 229.6)	231.0 (48.5, 420.2)	456.2 (168.5, 739.6)	990.3 (232.1, 1766.5)
NW	1	101.8 (22.6, 185.2)	40.5 (–11.6, 98.6)	116.0 (24.2, 212.5)	217.0 (79.7, 354.2)	475.3 (114.9, 850.6)
	2	184.1 (40.4, 340.9)	72.6 (–20.0, 179.5)	213.3 (44.1, 394.3)	401.5 (146.1, 661.3)	871.5 (210.6, 1575.9)
W	1	391.9 (86.3, 722.1)	142.5 (–51.0, 365.2)	312.8 (64.7, 578.4)	666.7 (242.7, 1097.7)	1513.9 (342.7, 2763.3)
	2	906.3 (195.9, 1712.1)	330.5 (–116.0, 873.2)	714.8 (145.9, 1342.5)	1548.9 (556.3, 2589.3)	3500.4 (782.1, 6517.0)
WNC	1	50.4 (11.3, 91.2)	33.5 (–8.9, 80.2)	24.8 (5.2, 45.3)	30.6 (11.3, 49.7)	139.3 (18.9, 266.4)
	2	93.8 (20.7, 171.8)	61.0 (–15.9, 148.4)	45.3 (9.4, 83.2)	57.9 (21.3, 94.3)	258.0 (35.5, 497.7)

**6. Discussion.** We present a novel potential outcomes framework that leverages numerical model output to estimate fire-contributed  $PM_{2.5}$  while taking spatial correlation into account and modeling interference between sites. Using a Bayesian spatial downscaling model and monitoring data, we bias-correct CMAQ-estimated counterfactual outcomes for  $PM_{2.5}$  under fire and no-fire regimes and model correlation between potential outcomes. Assuming consistency between the potential outcomes and the observations based on a CMAQ-derived treatment indicator and that confounding is accounted for conditional on CMAQ data, we show that the resulting estimate of fire-contributed  $PM_{2.5}$  has a valid causal interpretation.

We provide a spatially-resolved estimate for fire-contributed  $PM_{2.5}$  and uncertainty across the contiguous U.S. We found that the causal estimate of wildland fires on  $PM_{2.5}$  reached the highest levels in the West, Northwest, and Southeast regions. The western parts of the U.S. are impacted by large wildfires, and frequent prescribed and agricultural burns are observed in the Southeast. The number of estimated hospitalizations, due to exposure to fire-contributed  $PM_{2.5}$ , also reached a maximum in these regions, particularly in central California. Our estimates are lower than those produced by CMAQ. This particular application can be used by health professionals and environmental managers to better understand the health burden associated with fire events in their communities. Equipped with health burden estimates and uncertainty, they would be able to better anticipate the number of patients to expect and to plan accordingly. In this analysis we estimated the number of cumulative respiratory hospitalizations per county; it is possible to compute other outcomes related to  $PM_{2.5}$  exposure, including all-cause mortality, cardiovascular outcomes, etc.

The study has limitations and strengths. The CMAQ model appears to be biased. While we account for this bias in the estimation procedure, understanding this bias could improve accuracy. Future extensions of the proposed causal framework method for deterministic model data fusion could consider characterizing uncertainties associated with CMAQ inputs. Such research efforts would be costly, as they would require multiple CMAQ runs under multiple conditions but could provide valuable insights for CMAQ performance in wildfire scenarios. We take a model-based approach that relies on relatively simple separable stationary Gaussian processes. Given the size of the data for this particular study, this approach is warranted but could be revisited if used for smaller spatial regions. The approach is, however, generalizable to related research questions concerning how fire-contributed  $PM_{2.5}$  depends on the specific features of wildland fires, such as their location, strength, etc. or attribution of  $PM_{2.5}$  to a single fire in which case CMAQ model would be run with corresponding forcings. These questions are critical in the environmental management context when it has to be shown that a specific fire caused exceedances of regulatory air quality standards. The proposed causal inference framework can also be generalized to wider range of attribution studies where potential outcomes can be represented using numerical modeling approaches, for example, in climate science, forestry, materials science, etc. In each case the potential outcomes would differ by the factor of attribution whose impact is the objective of inference. Under the given assumptions and with bias correction, we show that the resulting inference has a valid causal interpretation.

## APPENDIX

**PROOF OF THEOREM 1.** To relate the potential outcome processes to the induced model of the observed outcome process, consider  $Y_t^{\text{miss}}(\mathbf{s})$  as the observation of the potential outcome that is missing under each regime, that is,

$$Y_t(\mathbf{s}, 0) = \begin{cases} Y_t(\mathbf{s}) & \text{if } C_t(\mathbf{s}) = 0, \\ Y_t^{\text{miss}}(\mathbf{s}) & \text{if } C_t(\mathbf{s}) = 1, \end{cases} \quad \text{and} \quad Y_t(\mathbf{s}, 1) = \begin{cases} Y_t^{\text{miss}}(\mathbf{s}) & \text{if } C_t(\mathbf{s}) = 0, \\ Y_t(\mathbf{s}) & \text{if } C_t(\mathbf{s}) = 1. \end{cases}$$

Hence, the joint distribution of the potential outcomes,  $Y_t(\mathbf{s}, 0)$  and  $Y_t(\mathbf{s}, 1)$ , is the joint distribution of the observed and missing observations  $Y_t(\mathbf{s})$  and  $Y_t^{\text{miss}}(\mathbf{s})$ .

Denoting  $\Theta$  as all parameters in the potential outcomes model, the likelihood function of  $\Theta$  is

$$\begin{aligned}
 & \prod_{t=1}^T \int f(\mathbf{Y}_t, \mathbf{Y}_t^{\text{miss}} | \hat{\boldsymbol{\theta}}_t, \hat{\boldsymbol{\delta}}_t, \Theta) d\mathbf{Y}_t^{\text{miss}} \\
 &= \prod_{t=1}^T \int f(\mathbf{Y}_t, \mathbf{Y}_t^{\text{miss}} | \hat{\boldsymbol{\theta}}_t, \hat{\boldsymbol{\delta}}_t, \mathbf{C}_t, \Theta) d\mathbf{Y}_t^{\text{miss}} \\
 (7) \quad &= \prod_{t=1}^T \left[ \int f(\mathbf{Y}_t^{\text{miss}} | \mathbf{Y}_t, \hat{\boldsymbol{\theta}}_t, \hat{\boldsymbol{\delta}}_t, \mathbf{C}_t, \Theta) d\mathbf{Y}_t^{\text{miss}} \right] f(\mathbf{Y}_t | \hat{\boldsymbol{\theta}}_t, \hat{\boldsymbol{\delta}}_t, \mathbf{C}_t, \Theta) \\
 &= \prod_{t=1}^T f(\mathbf{Y}_t | \hat{\boldsymbol{\theta}}_t, \hat{\boldsymbol{\delta}}_t, \mathbf{C}_t, \Theta),
 \end{aligned}$$

where the second line follows by Assumption 1. By equation (7),  $\Theta$  depends only on the observed processes which completes the proof.  $\square$

**Acknowledgments.** *Disclaimer:* This work does not necessarily represent U.S. EPA views or policy.

**Funding.** This work was partially supported by grants from the National Institutes of Health (R01ES027892, R01DE024984-01A1, R01ES031651-01), the National Science Foundation (DMS-1513579 and DMS-1811245), the National Cancer Institute (P01CA142538), the Department of the Interior (14-1-04-9), and Oak Ridge Associated Universities.

## SUPPLEMENTARY MATERIAL

**Supplement to “A spatial causal analysis of wildland fire-contributed PM<sub>2.5</sub> using numerical model output—REVISED”** (DOI: [10.1214/22-AOAS1610SUPP](https://doi.org/10.1214/22-AOAS1610SUPP); .pdf). The supplemental materials available online include technical details and derivations pertaining to the Bayesian model, sensitivity tests, MCMC convergence diagnostics and additional results.

## REFERENCES

- ALLEN, M. R. and STOTT, P. A. (2003). Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Clim. Dyn.* **21** 477–491. <https://doi.org/10.1007/s00382-003-0313-9>
- BENMAP (2017). Environmental Benefits Mapping and Analysis Program—Community Edition User’s Manual Appendices Technical Report.
- BERROCAL, V. J., GELFAND, A. E. and HOLLAND, D. M. (2010). A spatio-temporal downscaler for output from numerical models. *J. Agric. Biol. Environ. Stat.* **15** 176–197. MR2787270 <https://doi.org/10.1007/s13253-009-0004-z>
- CHILÈS, J.-P. and DELFINER, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2850475 <https://doi.org/10.1002/9781118136188>
- CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. *Wiley Series in Probability and Statistics*. John Wiley & Sons, Inc., Hoboken, NJ, USA. <https://doi.org/10.1002/9781119115151>
- DELFINO, R. J., BRUMMEL, S., WU, J., STERN, H., OSTRO, B., LIPSETT, M., WINER, A., STREET, D. H., ZHANG, L. et al. (2009). The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. *Occup. Environ. Med.* **66** 189–197. <https://doi.org/10.1136/oem.2008.041376>
- DENNEKAMP, M. and ABRAMSON, M. J. (2011). The effects of bushfire smoke on respiratory health. *Respirology* **16** 198–209. <https://doi.org/10.1111/j.1440-1843.2010.01868.x>

- DENNEKAMP, M., STRANEY, L. D., ERBAS, B., ABRAMSON, M. J., KEYWOOD, M., SMITH, K., SIM, M. R., GLASS, D. C., DEL MONACO, A. et al. (2015). Forest fire smoke exposures and out-of-hospital cardiac arrests in Melbourne, Australia: A case-crossover study. *Environ. Health Perspect.* **123** 959–964. <https://doi.org/10.1289/ehp.1408436>
- DOMINICI, F., GREENSTONE, M. and SUNSTEIN, C. R. (2014). Science and regulation. Particulate matter matters. *Science* **344** 257–259. <https://doi.org/10.1126/science.1247348>
- FUENTES, M. and RAFTERY, A. E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics* **61** 36–45. MR2129199 <https://doi.org/10.1111/j.0006-341X.2005.030821.x>
- GELFAND, A. E., KIM, H.-J., SIRMANS, C. F. and BANERJEE, S. (2003). Spatial modeling with spatially varying coefficient processes. *J. Amer. Statist. Assoc.* **98** 387–396. MR1995715 <https://doi.org/10.1198/016214503000170>
- GELFAND, A. E., SCHMIDT, A. M., BANERJEE, S. and SIRMANS, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *TEST* **13** 263–312. MR2154003 <https://doi.org/10.1007/BF02595775>
- HAIKERWAL, A., AKRAM, M., DEL MONACO, A., SMITH, K., SIM, M. R., MEYER, M., TONKIN, A. M., ABRAMSON, M. J. and DENNEKAMP, M. (2015). Impact of fine particulate matter (PM<sub>2.5</sub>) exposure during wildfires on cardiovascular health outcomes. *J. Am. Heart Assoc.* **4** e001653. <https://doi.org/10.1161/JAHA.114.001653>
- HAIKERWAL, A., AKRAM, M., SIM, M. R., MEYER, M., ABRAMSON, M. J. and DENNEKAMP, M. (2016). Fine particulate matter (PM<sub>2.5</sub>) exposure during a prolonged wildfire period and emergency department visits for asthma. *Respirology* **21** 88–94. <https://doi.org/10.1111/resp.12613>
- HALLORAN, M. E. and STRUCHINER, C. J. (1991). Study designs for dependent happenings. *Epidemiology* **2** 331–8.
- KATZFUSS, M., HAMMERLING, D. and SMITH, R. L. (2017). A Bayesian hierarchical model for climate change detection and attribution. *Geophys. Res. Lett.* **44** 5720–5728. <https://doi.org/10.1002/2017GL073688>
- HANNART, A., PEARL, J., OTTO, F. E. L., NAVEAU, P. and GHIL, M. (2016). Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Am. Meteorol. Soc.* **97** 99–110. <https://doi.org/10.1175/BAMS-D-14-00034.1>
- HANSEN, B. B. (2008). The prognostic analogue of the propensity score. *Biometrika* **95** 481–488. MR2521594 <https://doi.org/10.1093/biomet/asn004>
- HEGERL, G. and ZWIERS, F. (2011). Use of models in detection and attribution of climate change. *Wiley Interdiscip. Rev.: Clim. Change* **2** 570–591. <https://doi.org/10.1002/wcc.121>
- HERNÁN, M. A., ALONSO, A., LOGAN, R., GRODSTEIN, F., MICHELS, K. B., WILLETT, W. C., MANSON, J. E. and ROBINS, J. M. (2008). Observational studies analyzed like randomized experiments. *Epidemiology* **19** 766–779. <https://doi.org/10.1097/EDE.0b013e3181875e61>
- HOLLAND, P. W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc.* **81** 945–970. MR0867618
- HONG, G. and RAUDENBUSH, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *J. Amer. Statist. Assoc.* **101** 901–910. MR2324091 <https://doi.org/10.1198/01621450600000447>
- HUDGENS, M. G. and HALLORAN, M. E. (2008). Toward causal inference with interference. *J. Amer. Statist. Assoc.* **103** 832–842. MR2435472 <https://doi.org/10.1198/01621450800000292>
- JOHNSTON, F. H., HENDERSON, S. B., CHEN, Y., RANDERSON, J. T., MARLIER, M., DEFRIES, R. S., KINNEY, P., BOWMAN, D. M. J. S. and BRAUER, M. (2012). Estimated global mortality attributable to smoke from landscape fires. *Environ. Health Perspect.* **120** 695–701. <https://doi.org/10.1289/ehp.1104422>
- KAO, E. K. (2017). *Causal Inference Under Network Interference: A Framework for Experiments on Social Networks*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)—Harvard University. MR4172385
- KENNEDY, M. C. and O’HAGAN, A. (2001). Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 425–464. MR1858398 <https://doi.org/10.1111/1467-9868.00294>
- KNUTSON, T., KOSSIN, J. P., MEARS, C., PERLWITZ, J. and WEHNER, M. F. (2017). Detection and attribution of climate change. In *Climate Science Special Report: Fourth National Climate Assessment, Volume I* 114–132. U.S. Global Change Research Program, Washington, DC, USA. <https://doi.org/10.7930/J01834ND>
- LARSEN, A., YANG, S., REICH, B. J. and RAPPOLD, A. G. (2022). Supplement to “A spatial causal analysis of wildland fire-contributed PM<sub>2.5</sub> using numerical model output—REVISED.” <https://doi.org/10.1214/22-AOAS1610SUPP>
- MCKENZIE, D., SHANKAR, U., KEANE, R. E., STAVROS, E. N., HEILMAN, W. E., FOX, D. G. and RIEBAU, A. C. (2014). Earth’s future smoke consequences of new wildfire regimes driven by climate change. *Earth’s Future* **2** 35–59. <https://doi.org/10.1002/2013EF000180>

- NATIONAL ACADEMIES OF SCIENCE (2016). Attribution of Extreme Weather Events in the Context of Climate Change Technical Report, National Academies of Sciences, Engineering, and Medicine, Washington, D.C. <https://doi.org/10.17226/21852>
- PAPADOGEORGOU, G., MEALLI, F. and ZIGLER, C. M. (2019). Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics* **75** 778–787. MR4012083 <https://doi.org/10.1111/biom.13049>
- RAPPOLD, A. G., STONE, S. L., CASCIO, W. E., NEAS, L. M., KILARU, V. J., CARRAWAY, M. S., SZYKMAN, J. J., ISING, A., CLEVE, W. E. et al. (2011). Peat bog wildfire smoke exposure in rural North Carolina is associated with cardiopulmonary emergency department visits assessed through syndromic surveillance. *Environ. Health Perspect.* **119** 1415–1420. <https://doi.org/10.1289/ehp.1003206>
- RAPPOLD, A. G., REYES, J., POULIOT, G., CASCIO, W. E. and DIAZ-SANCHEZ, D. (2017). Community vulnerability to health impacts of wildland fire smoke exposure. *Environ. Sci. Technol.* **51** 6674–6682. <https://doi.org/10.1021/acs.est.6b06200>
- REICH, B. J., YANG, S., GUAN, Y., GIFFIN, A. B., MILLER, M. J. and RAPPOLD, A. (2021). A review of spatial causal inference methods for environmental and epidemiological applications. *Int. Stat. Rev.* **89** 605–634. MR4411920 <https://doi.org/10.1111/insr.12452>
- ROSENBAUM, P. R. (2007). Interference between units in randomized experiments. *J. Amer. Statist. Assoc.* **102** 191–200. MR2345537 <https://doi.org/10.1198/016214506000001112>
- RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. MR0472152
- SCHMIDT, A. M. and GELFAND, A. E. (2003). A Bayesian coregionalization approach for multivariate pollutant data. *J. Geophys. Res., Atmos.* **108**. <https://doi.org/10.1029/2002JD002905>
- US FOREST SERVICE (2015). PNW—Fire and Environmental Research Applications Team (FERA) Research/Studies.
- SOBEL, M. E. (2006). What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *J. Amer. Statist. Assoc.* **101** 1398–1407. MR2307573 <https://doi.org/10.1198/016214506000000636>
- STAVROS, E. N., MCKENZIE, D. and LARKIN, N. (2014). The climate-wildfire-air quality system: Interactions and feedbacks across spatial and temporal scales. *Wiley Interdiscip. Rev.: Clim. Change* **5** 719–733. <https://doi.org/10.1002/wcc.303>
- TCHETGEN TCHETGEN, E. J. and VANDERWEELE, T. J. (2012). On causal inference in the presence of interference. *Stat. Methods Med. Res.* **21** 55–75. MR2867538 <https://doi.org/10.1177/0962280210386779>
- USEPA (2010). Quantitative Health Risk Assessment for Particulate Matter Technical Report US Environmental Protection Agency.
- USEPA (2015). AirData Download Files Documentation.
- USEPA (2019). Overview of science processes in CMAQ.
- USEPA (2020). Particulate Matter (PM<sub>2.5</sub>) Trends.
- WETTSTEIN, Z. S., HOSHIKO, S., FAHIMI, J., HARRISON, R. J., CASCIO, W. E. and RAPPOLD, A. G. (2018). Cardiovascular and cerebrovascular emergency department visits associated with wildfire smoke exposure in California in 2015. *J. Am. Heart Assoc.* **7** e007492. <https://doi.org/10.1161/JAHA.117.007492>
- ZIGLER, C. M., CHOIRAT, C. and DOMINICI, F. (2018). Impact of national ambient air quality standards nonattainment designations on particulate pollution and health. *Epidemiology* **29** 165–174. <https://doi.org/10.1097/EDE.0000000000000777>
- ZIGLER, C. M., DOMINICI, F. and WANG, Y. (2012). Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics* **13** 289–302. <https://doi.org/10.1093/biostatistics/kxr052>
- ZIGLER, C. M. and PAPADOGEORGOU, G. (2021). Bipartite causal inference with interference. *Statist. Sci.* **36** 109–123. MR4194206 <https://doi.org/10.1214/19-STS749>