

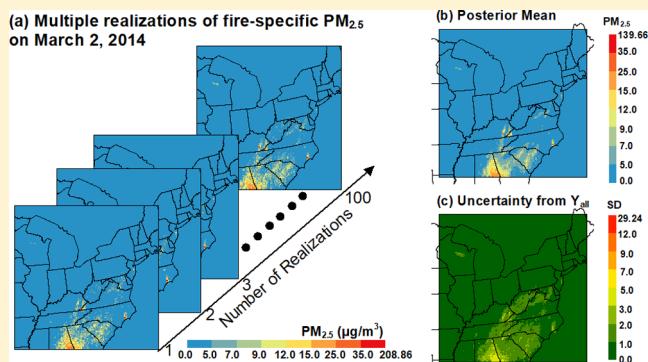
Modeling Wildland Fire-Specific PM_{2.5} Concentrations for Uncertainty-Aware Health Impact Assessments

Xiangyu Jiang*^{ID} and Eun-Hye Enki Yoo

Department of Geography, University at Buffalo—The State University of New York, Buffalo, New York 14261, United States

Supporting Information

ABSTRACT: Wildland fire is a major emission source of fine particulate matter (PM_{2.5}), which has serious adverse health effects. Most fire-related health studies have estimated human exposures to PM_{2.5} using ground observations, which have limited spatial/temporal coverage and could not separate PM_{2.5} emanating from wildland fires from other sources. The Community Multiscale Air Quality (CMAQ) model has the potential to fill the gaps left by ground observations and estimate wildland fire-specific PM_{2.5} concentrations, although the issues around systematic bias in CMAQ models remain to be resolved. To address these problems, we developed a two-step calibration strategy under the consideration of prediction uncertainties. In a case study of the eastern U.S. in 2014, we evaluated the calibration performance using three cross-validation methods, which consistently indicated that the prediction accuracy was improved with an R^2 of 0.47–0.64. In a health impact study based on the wildland fire-specific PM_{2.5} predictions, we identified regions with excess respiratory hospital admissions due to wildland fire events and quantified the estimation uncertainty propagated from multiple components in health impact function. We concluded that the proposed calibration strategy could provide reliable wildland fire-specific PM_{2.5} predictions and health burden estimates to support policy development for reducing fire-related risks.



1. INTRODUCTION

Wildland fires, including both prescribed fires and wildfires, have occurred at a higher level of intensity and frequency across the western U.S. in recent years due to arid conditions and increased temperature related to climate changes.^{1,2} About 1 400 000 wildland fires have burned more than 560 000 km² of forests, woodland, and rangeland in the past 2 decades over the entire U.S.,³ generating large amounts of fine particulate matter (PM_{2.5}; particles with an aerodynamic diameter $\leq 2.5 \mu\text{m}$).^{4,5} This massive amount of particulate matter deteriorated air quality^{6–8} and caused adverse health problems as evidenced in a growing number of epidemiological studies.^{9–11} A number of studies have shown that exposure to intense PM_{2.5} concentrations during wildfire episodes was associated with an increased risk of morbidity and mortality from respiratory and cardiovascular disease,^{12–16} although others, such as Johnston et al.¹⁷ and Yao et al.,¹⁸ concluded that the effect of wildfire smoke on cardiovascular morbidity was inconclusive. These inconsistent results might be attributed to misclassified exposures to PM_{2.5} concentrations stemming from sporadic and episodic wildland fires.^{19,20}

Most fire-related health studies have estimated exposures to PM_{2.5} using ground observations obtained from the network of state or local air monitoring stations.¹⁹ Ground PM_{2.5} observations are highly accurate, but their spatial coverage is sparse, especially over rural areas where wildfires frequently

occur.^{18,21,22} Moreover, many monitoring stations operated inconsistently, thereby resulting in gaps in the continuous monitoring of the spatial and temporal variations of PM_{2.5} concentrations.^{22,23} Satellite aerosol optical depth (AOD) data have been increasingly used as a proxy for ground PM_{2.5} observations.^{21,23–25} Many air quality studies have combined satellite-based AOD data with ground observations to improve the spatial and temporal coverages of PM_{2.5} estimates,^{26,27} although the accuracy of predicted PM_{2.5} is affected by missing or abnormal AOD values associated with the presence of bright surfaces, heavy cloud cover, and thick wildfire smoke.^{24,28,29} Moreover, estimates of exposure to wildland fire-related PM_{2.5} based on either ground observations or satellite AOD-based predictions cannot distinguish wildland fire-sourced PM_{2.5} concentrations from other emission sources, such as transportation and power plants.^{19,30,31} This inability to parse out PM_{2.5} concentrations from wildland fire emissions versus other sources is problematic for epidemiological studies and policy development, because PM_{2.5} coming from wildland fire emissions might be more harmful to public health than other sourced PM_{2.5}.³²

Received: May 2, 2019

Revised: September 18, 2019

Accepted: September 19, 2019

Published: September 19, 2019

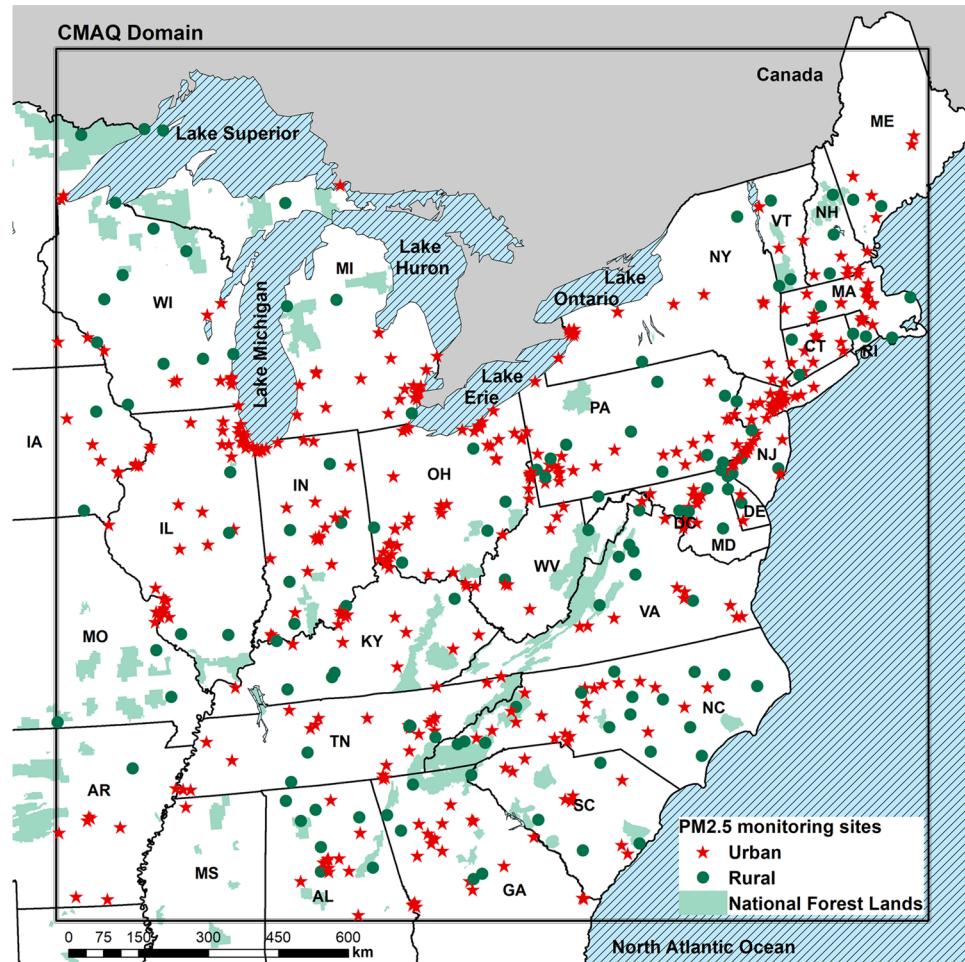


Figure 1. Study domain (thick solid lines), monitoring stations (circles and stars), lakes and ocean (shaded polygon in blue), Canada (shaded polygon in gray), and national forest lands (shaded polygon in light green).

Some studies^{23,33,34} have used numerical air quality models to estimate air pollutant concentrations specifically from wildland fires at hourly intervals with complete coverage of the study domain. The atmospheric chemistry models are designed to simulate source-specific air pollutants at grid level by integrating meteorological fields and various emission sources.³⁵ The model outputs have been used as deterministic predictions of wildland fire-related air pollution in several health impact studies.^{16,19,36} For example, Johnston et al.³⁷ characterized the distribution of global mortality due to landscape fire-related PM_{2.5} exposure using the GEOS-Chem model. Kollanus et al.³⁸ predicted daily vegetation fire-specific PM_{2.5} concentrations in Europe using a mesoscale atmospheric model and quantified fire-related annual mortality. Recently, Fann et al.³⁹ concluded that exposure to wildland fire-related PM_{2.5} resulted in approximately 11 000 annual mortality and 6400 respiratory admissions to a hospital across the U.S. in 2012, basing their PM_{2.5} estimates on Community Multiscale Air Quality (CMAQ) model simulations.

Although atmospheric chemistry models, such as the CMAQ model, have been used to estimate exposures to wildland fire-specific PM_{2.5}, little attention has been paid to systematic biases present in the CMAQ model that are associated with uncertainties in the model inputs, such as meteorological fields and emissions, and model parameterizations.^{40–42} Statistical calibration is a viable option to adjust such biases in atmospheric

model simulations as shown in both Wang et al.⁴³ and Solazzo et al.⁴⁴ in which calibration improved the prediction accuracy. Several statistical approaches are readily applicable for the atmospheric model calibration, including quantile mappings,^{45,46} inverse modeling techniques,^{44,47} and neural-network-based models,²⁵ although they are not necessarily designed for uncertainty assessments.

Alternatively, the Bayesian downscaler model^{48–50} can be used to improve the quality of CMAQ-based deterministic predictions and quantify uncertainties associated with PM_{2.5} predictions within the Bayesian hierarchical framework. Specifically, the downscaler model combines ground PM_{2.5} observations with CMAQ outputs treating them as a proxy variable within a linear regression model with spatially and temporally varying coefficients. The downscaler models have been widely used in both data fusion and health studies. For example, Chang et al.⁵¹ fused ground observations with CMAQ outputs to estimate total PM_{2.5} concentrations over North Carolina to investigate associations between ambient PM_{2.5} and preterm birth. Similarly, Geng et al.⁵² integrated accurate ground PM_{2.5} observations with collocated CMAQ outputs and other land use variables under the Bayesian hierarchical framework to predict total PM_{2.5} concentrations during fire days in Colorado. The Bayesian downscaler method, however, has neither been used for calibrating CMAQ-based wildland fire-specific PM_{2.5} concentrations nor been used for quantifying their prediction

uncertainties, possibly due to the absence of ground wildland fire-specific PM_{2.5} observations.

In the present study, we predicted PM_{2.5} concentrations originating from wildland fires while quantifying the prediction uncertainties. Specifically, we proposed a two-step strategy based on the Bayesian downscaler model to correct biases in CMAQ-based total PM_{2.5} predictions, which were used to extract wildland fire-specific PM_{2.5} concentrations. Finally, we evaluated health impacts attributable to the exposure to wildland fire-specific PM_{2.5} concentrations, while accounting for the uncertainties in fire-related health burdens propagated from both uncertain wildland fire-specific PM_{2.5} predictions and risk coefficients.

2. MATERIALS AND METHODS

2.1. Study Region and Data Sources. The study domain includes the eastern U.S., Canadian provinces, and a subset of the North Atlantic Ocean, which covers approximately 1872 × 1872 km² areas (see Figure 1). For the purpose of prediction, we discretized the study region into grid cells with a 12 km resolution. Here, we focused on the year of 2014, because this was the latest year that U.S. Environmental Protection Agency (EPA) provides comprehensive emission inventories for the CMAQ model simulations at the time of this study, which enabled us to avoid substantial bias in CMAQ simulations from insufficient emission sources.

2.1.1. CMAQ Model Simulations. We used CMAQ model version 5.2.1⁵³ to simulate 24 h average PM_{2.5} concentrations at a 12 km resolution, from which we identified the contribution of wildland fire emissions to total PM_{2.5} concentrations. The CMAQ simulations were carried out over the eastern U.S., which was discretized into a total of 24 336 grid cells and 35 vertical layers. Two primary inputs were prepared for the CMAQ simulations in the study domain. The first input was hourly meteorological fields generated by the Weather Research and Forecasting (WRF) model version 3.8.⁵⁴ The physical options used in WRF simulations are listed in Table S1. More detailed information about the WRF configuration settings and model evaluation performance can be found in the WRF technical report.⁵⁵ Emissions, including both fire and nonfire emission sources, were taken from the 2014 emission modeling platform version 7.1.⁵⁶ This platform was developed based on the 2014 National Emission Inventory,⁴ which provided all possible emission data and ancillary files for the subsequent emission processing.⁵⁷ The information on the fire activity locations and burned area was collected from satellite images, such as the U.S. National Oceanic and Atmospheric Administration Hazard Mapping System, and ground-based wildland fire reports, such as the Incident Command System (ICS-209).^{58,59} These data were reconciled by SMARTFIRE2 and digested by BlueSky emission modeling framework to estimate day-specific wildland fire-related emissions at point level.^{4,58,59} About 70 930 and 234 227 t of PM_{2.5} emissions were generated by wildfires and prescribed burning over the study region.⁵ Additional details on the fire and nonfire emission sources can be found in the EPA emission report.⁵⁷ We applied the Sparse Matrix Operator Kernel Emission (SMOKE) model version 4.5⁶⁰ to process the above emission sources and generate emission estimates at hourly intervals across grid cells.

We used CMAQ predefined static initial conditions and boundary conditions to prepare the chemical conditions over the study region for the first hour of air quality simulations and along the study region's boundaries for the entire year,

respectively. To estimate wildland fire-specific PM_{2.5}, we ran the CMAQ model under two scenarios with a chemical module of Carbon-Bond 06 and an aerosol module of AERO6 without considering any spin-up periods. In the first scenario, we simulated hourly average PM_{2.5} levels by integrating meteorological fields provided by the WRF model and all emission sources processed by the SMOKE processor. For the second case, the same input data except wildland fire emissions were used. That is, we simulated both CMAQ-based PM_{2.5} concentrations with all emission sources Y_{all} and without wildland fire emissions Y_{noFire} , respectively. The discrepancy between these two simulation outputs ($Y_{\text{all}} - Y_{\text{noFire}}$), denoted as Y_{fire} , represents the amount of hourly average PM_{2.5} concentrations stemming from wildland fires at each grid cell. Given most health studies require daily time scale of surface air pollutant data, we aggregated the hourly CMAQ-based PM_{2.5} predictions obtained from the lowest layer (about 20 m deep) to 24 h averages based on local time.

We quantified the daily contribution $F(g_k, d_j)$ of wildland fire-specific PM_{2.5} to total PM_{2.5} concentrations at each grid cell g_k , $k = 1, \dots, K$ (24 336) on each day d_j , $j = 1, \dots, T$ (365). Given that CMAQ-based PM_{2.5} predictions were available on an hourly basis, we summarized them via a set of summary measures. For example, we quantified the daily contribution as the 24 h average of the hourly conversion ratios $F(g_k, h_l)$, $l = 1, \dots, 24$ as

$$F(g_k, d_j) = \frac{1}{24} \sum_{l=1}^{24} F(g_k, h_l) \quad (1)$$

where $F(g_k, h_l)$ is the fraction of hourly CMAQ-based wildland fire-specific to all-source PM_{2.5} concentrations, obtained as $\frac{Y_{\text{fire}}(g_k, h_l)}{Y_{\text{all}}(g_k, h_l)}$. We also considered two quantiles of the hourly conversion ratios, i.e., the 5th and 95th percentile as the lower and upper bounds of $F(g_k, d_j)$, respectively. We assumed that the contribution of wildland fire emissions to total PM_{2.5} concentrations is invariant to the statistical calibration.^{33,46}

2.1.2. Ground PM_{2.5} Observations. We collected ground PM_{2.5} measurements from the EPA air quality system (AQS).⁶¹ There were a total of 587 monitoring stations in the study domain in 2014, which were assigned to 478 collocated grid cells for a modeling purpose. If more than one monitoring stations were located within a grid cell, we paired the ground PM_{2.5} observations with the same CMAQ value. Each monitoring site was classified as either urban or rural based on their locations with respect to the urban boundary retrieved from U.S. Census Bureau.⁶² As shown in Figure 1, most monitoring sites (77.3%) were located in metropolitan areas and the remaining sites (22.7%) were in rural areas. These monitoring sites reported hourly or 24 h rolling average PM_{2.5} concentrations every 1, 3, 6, or 12 days during the study period of 01/01/2014–12/31/2014. Negative PM_{2.5} values were discarded for the subsequent analyses as they were viewed as biased measurements. Similarly, we excluded four observations (marked as “excluded”) based on their event-type label to avoid potential bias.⁶³ Further details on the AQS data processing are available in Section 3 of Supporting Information. Hourly observations at each monitoring sites were aggregated to 24 h averages, which yielded a total of 139 027 daily average PM_{2.5} measurements.

2.1.3. Ancillary Data Sets. Following previous studies,^{52,64} we initially included several covariates, such as the distance to the nearest fire, elevation, percentage of forest coverage, percentage of impervious surface, road density, and population

density, in the calibration model. However, the preliminary results revealed that only two covariates, the percentage of forest coverage at each grid cell and the distance to the nearest fire, were relatively strong indicators of PM_{2.5} concentrations during wildfires. Forest coverage was retrieved from the 2011 National Land Cover Database,^{65,66} which was aggregated from its original resolution of 0.03–12 km, at which the fraction of forest areas was calculated. We detected daily fire events at point level using the Moderate Resolution Imaging Spectroradiometer (MODIS)⁶⁷ that were taken from the Remote Sensing Application Center of the U.S. Department of Agriculture Forest Service. We calculated the distance between each monitor station and its nearest fire event on days when fire events occurred.

2.2. Modeling Wildland Fire-Specific PM_{2.5} Concentrations. Calibration may adjust biases in CMAQ-based wildland fire-specific PM_{2.5} concentrations, but it requires wildland fire-specific ground PM_{2.5} observations that were not readily available. The ground observations collected from monitoring stations were measurements of total PM_{2.5} concentrations from various emissions, including wildland fires, motor vehicles, and power plants, etc. To accommodate the absence of source-specific PM_{2.5} measurements, we developed a two-step calibration strategy. First, we calibrated the CMAQ-based all-source PM_{2.5} values $Y_{\text{all}}(\cdot, \cdot)$ to ground evidence using the Bayesian statistical downscaling approach.⁴⁹ In the second step, we extracted daily fire-contributed PM_{2.5} at each grid cell from the calibrated all-source PM_{2.5} concentrations obtained from the first step by multiplying the conversion ratio $F(\cdot, \cdot)$ estimated from eq 1, while quantifying the uncertainties in fire-specific PM_{2.5} predictions.

2.2.1. Stage 1: Statistical Calibration of All-Source PM_{2.5} Concentrations. For the statistical calibration, we combined daily all-source PM_{2.5} concentrations obtained from the CMAQ model (see Section 2.1.1) with ground observations via a linear regression model with spatially and temporally varying coefficients as⁴⁹

$$\begin{aligned} \text{PM}(s_i, d_j) = & \alpha_1(s_i) + \beta_1(d_j) + [\alpha_2(s_i) + \beta_2(d_j)] \cdot Y_{\text{all}}(g_{s_i}, d_j) \\ & + \gamma \cdot Z(g_{s_i}, d_j) + \epsilon(s_i, d_j) \end{aligned} \quad (2)$$

where $\text{PM}(s_i, d_j)$ denotes the daily PM_{2.5} ground observation obtained at i th monitoring site s_i , $i = 1, \dots, 587$ on the j th day d_j , $j = 1, \dots, T$; $Y_{\text{all}}(g_{s_i}, d_j)$ is the same day CMAQ-based all-source PM_{2.5} concentration at a grid cell g_{s_i} within which monitoring station $s_i \in g_{s_i}$ is located. The two predictors $Z(g_{s_i}, d_j) = [Z_1(g_{s_i}) Z_2(s_i, d_j)]'$ consist of the percentage of forest within a grid cell g_{s_i} containing monitoring site s_i , denoted as $Z_1(g_{s_i})$, and the distance between monitoring site s_i to its nearest fire event if fire occurred on day d_j , denoted as $Z_2(s_i, d_j)$; the coefficients for the predictors are denoted as a vector of fixed effects as $\gamma = [\gamma_1 \gamma_2]$. The spatially and temporally varying random coefficients for the intercept and slope are denoted as $\alpha_1(s_i)$, $\alpha_2(s_i)$, and $\beta_1(d_j)$, $\beta_2(d_j)$, respectively. The residual term $\epsilon(s_i, d_j)$ follows the normal distribution with mean zero and variance σ^2 , i.e., $\epsilon(s_i, d_j) \sim N(0, \sigma^2)$.

A total of 139 027 daily PM_{2.5} measurements were matched with the collocated CMAQ outputs and the two predictors, from which we estimated the model in eq 2. The Markov chain Monte Carlo (MCMC) was used for the model inference within a Bayesian hierarchical framework. Specifically, we assigned a

prior distribution for each parameter and used the MCMC technique to obtain samples from posterior distributions of parameters given observations. We ran 50 000 MCMC iterations and discarded the first 25 000 parameter samples as burn-in period and thinned the remaining 25 000 samples by 250 to gain computational efficiency. We used the 100 sets of parameters obtained from MCMC procedure to generate 100 realizations of calibrated all-source daily PM_{2.5} concentrations at all grid cells, which are denoted as $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ for the 100 sets of realizations ($r = 1, \dots, 100$). The two spatial random coefficients $\hat{\alpha}_1(s_i)$ and $\hat{\alpha}_2(s_i)$ estimated at monitoring sites were interpolated at the centroid of each grid cell using the tapered conditional kriging approach. Further details on the parameter specifications and MCMC algorithm can be found in Chang et al.⁴⁹

The calibration performance was evaluated via three different cross-validation (CV) approaches, including the 10-fold, leave-one-state-out (LOSO), and leave-one-month-out (LOMO) CV. In all CVs, the data set was split into training and validation sites, in which the former was used to fit the model. The model predictions obtained at validation sites were compared to the corresponding PM_{2.5} observations. In the 10-fold CV, we randomly split the data set into 10-folds and used 9-folds to fit the model and the remaining fold for validation. To eliminate potential biases associated with spatial and temporal clusters in the training sites, we assessed the calibration performance at locations without any nearby monitors in the LOSO CV by dropping all PM_{2.5} observations from one state for model training and used the dropped values for model validation. Similarly, we excluded PM_{2.5} observations for each month and fitted the model with data collected from the remaining 11 months in LOMO CV. This process was repeated until each of 10-fold, each state, and each month was used as a validation data set. In all CVs, the gap between calibrated and observed PM_{2.5} at validation sites was summarized using statistical indicators, including mean bias (MB), coefficient of determination (R^2), fractional bias (FB), and fractional error (FE). See the definition of each statistical indicator in Table S2.

2.2.2. Stage 2: Wildland Fire-Specific PM_{2.5} Concentrations. We extracted daily fire-contributed PM_{2.5} from the set of multiple realizations of daily calibrated all-source PM_{2.5} concentration surface $\tilde{Y}_{\text{all}}^{(r)}(\cdot, \cdot)$ obtained from stage 1 by multiplying the conversion ratio as

$$\begin{aligned} \tilde{Y}_{\text{fire}}^{(q,r)}(g_k, d_j) = & F^{(q)}(g_k, d_j) \times \tilde{Y}_{\text{all}}^{(r)}(g_k, d_j), \quad q = 1, 2, 3, \\ r = 1, \dots, 100 \end{aligned} \quad (3)$$

where $F^{(q)}(g_k, d_j)$ denotes the q th set of time and location-specific conversion ratio. The three sets of daily conversion ratios consist of the 5th percentile ($q = 1$), the 24 h average ($q = 2$), and the 95th percentile ($q = 3$) of hourly conversion ratio distribution.

We obtained a total of 300 realizations of daily wildland fire-related PM_{2.5} by multiplying the 100 sets of $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ with the three sets of $F^{(q)}(g_k, d_j)$. For each q , we summarized simulated daily wildland fire-related PM_{2.5} by calculating the posterior mean as $\frac{1}{100} \sum_{r=1}^{100} \tilde{Y}_{\text{fire}}^{(q,r)}(g_k, d_j)$ at each grid cell g_k , $k = 1, \dots, K$, which was further aggregated over time for an annual average. We also explored the daily variability of wildland fire-specific PM_{2.5} concentrations by calculating the spatial average of fire-related PM_{2.5} levels for each $F^{(q)}$.

We quantified the uncertainty in the daily wildland fire-related PM_{2.5} prediction $\tilde{Y}_{\text{fire}}^{(q,r)}$ by calculating the standard deviation

(SD) of the wildland fire-specific PM_{2.5} realizations at each grid cell on a daily basis. To explore the uncertainty propagated from each component of $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ and $F^{(q)}(g_k, d_j)$, we calculated two sets of SDs. First, we calculated SD from the 100 sets of wildland fire-specific PM_{2.5} realizations based on the 24 h averaged conversion ratio $F^{(2)}$. Further, we calculated SD from the 300 sets of realizations where the uncertainty in the conversion ratio was accounted for. We summarized the spatial variability of the prediction uncertainty for the entire year by classifying daily SD value at each grid cell into 1 if the value was greater than a cutoff value ($c = 1 \mu\text{g}/\text{m}^3$), 0 otherwise. In summary, a grid cell with the uncertainty indicator of 1 indicates that there is high uncertainty based on the cutoff value, whereas an indicator of 0 suggests that the prediction was reliable. We counted the total number of days that are associated with high uncertainties per grid cell. We also assessed the sensitivity of the cutoff value by examining the variability in the number of days with high uncertainty with multiple cutoff values ($c = 2, 3, 5 \mu\text{g}/\text{m}^3$) and summarized the results in Figure S6.

2.3. Uncertainty-Aware Health Impact Assessments. Several studies have investigated the exposure to wildland fire-related PM_{2.5}^{36,38,39} in which the extent of mortality or morbidity attributed to fire-specific PM_{2.5} was estimated from a health impact function. However, the impact of uncertain exposure estimates on estimated health burdens has not been investigated in previous studies. In what follows, we estimated the excess respiratory hospital admissions (RHAs) caused by the exposure to wildland fire-specific PM_{2.5}, while accounting for multiple uncertainty components in the following health impact function.

2.3.1. Health Impact Function. The daily RHA attributable to the exposure to wildland fire-related PM_{2.5} concentrations can be estimated using the health impact function⁶⁸

$$\Delta\text{RHA}_{\text{fire}}(g_k, d_j) = (1 - e^{-\beta \cdot \tilde{Y}_{\text{fire}}(g_k, d_j)}) \cdot m_0(g_k, d_j) \cdot \text{Pop}(g_k) \quad (4)$$

where β denotes the risk coefficient following $\beta \sim N(0.00276, 0.00067^2)$. The risk coefficient was estimated based on the relative risk of RHA associated with PM_{2.5} increment during wildfires in southern California.⁶⁹ The detailed information regarding the β distribution is provided in Section 4 of the Supporting Information. The term $m_0(g_k, d_j)$ refers to the rate of daily baseline hospital visits for the exposed population at a grid cell g_k on day d_j . The annual hospitalization rates at regional, county, and state levels taken from the Healthcare Cost and Utilization Project⁷⁰ were converted to daily rates at the grid level using the Benefit Mapping and Analysis Program – Community Edition version 1.4.^{68,71} $\text{Pop}(g_k)$ represents the 2014 population data (aged from 0 to 99) within the grid cell g_k , as generated by the PopGrid program.⁷² The spatial distribution of the population is presented in Figure S2. The hospital admission rates and population data over Canada were not included, and thus their RHAs caused by fire-related PM_{2.5} were set to zeros.

2.3.2. Wildland Fire PM_{2.5}-Related RHA and Uncertainty Assessment. As a baseline for comparison, we estimated daily wildland fire-related RHA from eq 4 with the posterior mean of all-source PM_{2.5}, the 24 h averaged conversion ratio, and the mean of risk coefficient distribution. We summarized the spatial distribution of fire-related RHA by aggregating the daily RHA to annual total RHA. We also examined the daily variations of fire-related RHA over the entire study region.

To account for the uncertainties in health impact function, we developed four scenarios in which the effect of uncertainty in $\tilde{Y}_{\text{all}}^{(r)}(\cdot, \cdot)$, $F^{(q)}(\cdot, \cdot)$, and β on the fire PM_{2.5}-related RHA estimates was assessed. Table 1 summarizes the sources of uncertainty

Table 1. Sources of Uncertainty in the Fire-Related RHA Estimates

scenario	sources of uncertainties
1	$\tilde{Y}_{\text{all}}(g_k, d_j)$
2	$F(g_k, d_j)$
3	β
4	$\tilde{Y}_{\text{all}}(g_k, d_j)$, $F(g_k, d_j)$, β

considered in each scenario. In the top three scenarios, we considered one source of uncertainty at a time while controlling other sources being fixed. Specifically, we took lower and upper percentile (the 5th and 95th percentile) values of the uncertain component and estimated corresponding daily fire-related RHA. In scenario 4, however, we pooled all of the uncertainties in the three components together to assess the full uncertainties in fire-related RHA estimates. We generated a large set of fire-related RHA based on a total of eight combinations of both the 5th and 95th percentiles of $\tilde{Y}_{\text{all}}^{(r)}$, $F^{(q)}$, and β for scenario 4 and calculated the lower and upper bounds of daily fire-related RHA. The annual total fire-related RHA over the entire region was obtained by adding up the daily RHA values over T days at K grid cells for each scenario. Here, we assumed that other components of the model, such as baseline incident rate $m_0(g_k, d_j)$ and population $\text{Pop}(g_k)$, were fixed despite their potential impacts on health burden analysis,^{73,74} as we focused on the three sources in the present paper. The detailed workflow of estimating fire PM_{2.5}-related health burdens can be found in Figure S3.

3. RESULTS

We illustrated the proposed two-step calibration strategy for the prediction of annual CMAQ-based wildland fire-specific PM_{2.5} concentrations using data collected in the eastern U.S. during 2014. We also demonstrated its application to a health impact assessment, while accounting for the multiple sources of uncertainty associated with the health impact function.

3.1. CMAQ-Based PM_{2.5} Concentrations. The CMAQ-based annual PM_{2.5} estimation associated with all emission sources ranged from 1.51 to 29.12 $\mu\text{g}/\text{m}^3$. As shown in Figure 2a, the high levels of all-source PM_{2.5} ($>12 \mu\text{g}/\text{m}^3$) were concentrated in metropolitan cities, such as New York City (NYC) and Montreal, and in the national forest lands of southeastern North Carolina, eastern Tennessee, and central Alabama. The ground PM_{2.5} monitoring sites are denoted as either circles or stars in Figure 2a whose observations were related to the corresponding CMAQ-based annual PM_{2.5} with a correlation coefficient of 0.58.

The CMAQ-based annual PM_{2.5} concentrations without wildland fire emissions are presented in Figure 2b. Heavy nonfire PM_{2.5} levels were concentrated in major cities, with the highest concentration (29.06 $\mu\text{g}/\text{m}^3$) in Montreal. The spatial pattern of annual fire-specific PM_{2.5} concentrations is presented in Figure 2c, showing that relatively high PM_{2.5} ($>2 \mu\text{g}/\text{m}^3$) was concentrated in forest areas in Alabama, Kentucky, North Carolina, and Virginia, with a peak value of 7.31 $\mu\text{g}/\text{m}^3$ in southeastern North Carolina. The annual fire conversion ratio based on $F^{(2)}$ is presented in Figure 2d. The largest value was

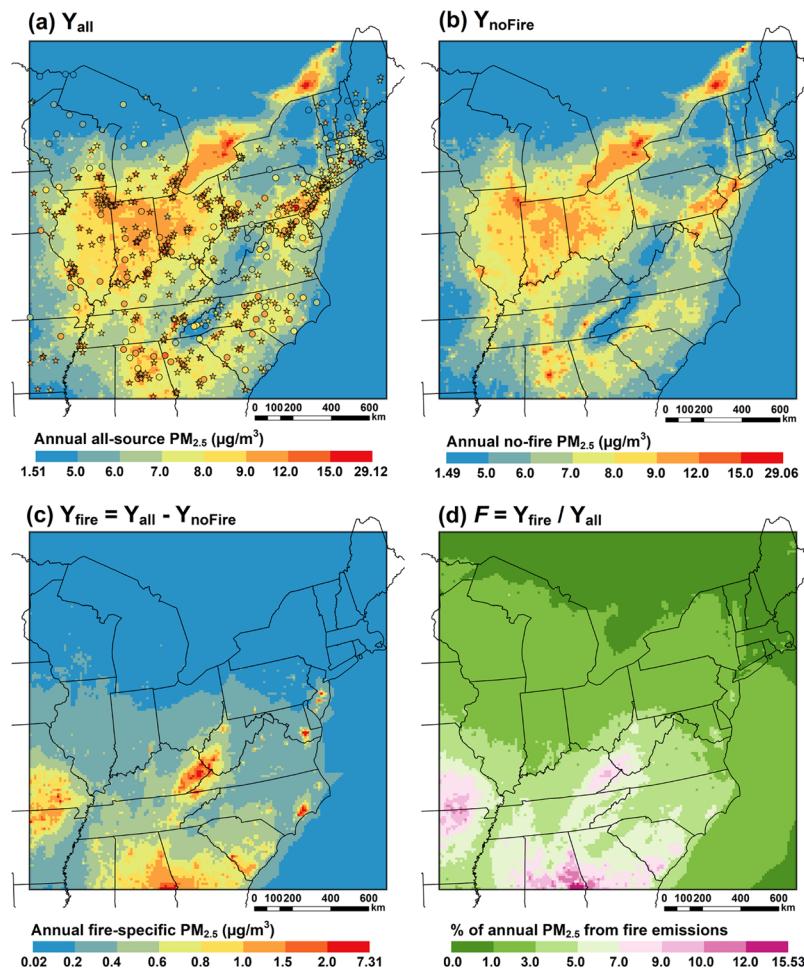


Figure 2. (a) Annual average of CMAQ-based $\text{PM}_{2.5}$ concentrations from all emission sources and $\text{PM}_{2.5}$ ground observations at monitoring sites (circles and stars); annual average of CMAQ-based $\text{PM}_{2.5}$ concentrations from (b) nonfire- and (c) fire-specific emission sources; and (d) rate of annual $\text{PM}_{2.5}$ concentrations from wildland fire emission sources to those from all sources.

found in Alabama, suggesting that 15.53% of total $\text{PM}_{2.5}$ concentrations were originated from wildland fire emissions annually. These results indicated that wildland fire was one of the essential emission sources for $\text{PM}_{2.5}$ air pollution, especially over the southern part of the study domain.

3.2. Calibrated All-Source $\text{PM}_{2.5}$ Concentrations. We fitted the Bayesian downscaler model in eq 2 with the daily ground observations at 587 monitoring sites. The calibration performance was evaluated by comparing the calibrated $\text{PM}_{2.5}$ against ground $\text{PM}_{2.5}$ observations at validation sites, which were selected from each CV criterion (see details in Section 2.2.1). To demonstrate the calibration effect, we compared the raw CMAQ-modeled outputs against ground $\text{PM}_{2.5}$ observations at the same validation sites.

As shown in Figure 3a,b, the calibrated $\text{PM}_{2.5}$ values were in the better agreement with the ground observations than the raw CMAQ outputs according to 10-fold CV. R^2 was improved from 0.40 to 0.64, and the linear regression slope was close to 1 (increasing from 0.56 to 0.99) after the biases were adjusted. The MB values for the CMAQ outputs before and after calibration were -0.36 and $0.002 \mu\text{g}/\text{m}^3$, respectively. Overall, the calibration substantially improved the accuracy of all-source $\text{PM}_{2.5}$ predictions, with the absolute FB value falling from 7.86 to 5.51% and the FE value decreasing by 12.72% points (from 39.16 to 26.44%). According to the prediction performance guideline,⁷⁵ the level of accuracy in calibrated all-source $\text{PM}_{2.5}$

concentrations was improved to “excellent” as the absolute FB and FE values were lower than the cutoff values of 15 and 35%, respectively. Gradual, but consistent patterns of the improvement was found in the two other CV methods of LOSO and LOMO. As shown in Figure 3c,d, R^2 values from the LOSO and LOMO were 0.47 and 0.51, respectively. Both were higher than the R^2 value of 0.40 obtained from the raw CMAQ outputs. Moreover, the MB values (0.01 for LOSO and 0.07 for LOMO) were smaller than the absolute MB (0.36) for the raw CMAQ outputs. As suggested from previous studies,^{76,77} the LOSO and 10-fold CV results can be treated as the lower and upper bounds of calibration performance. In summary, the overall performance of the Bayesian downscaler model fell within the R^2 range of 0.47–0.64. The realizations of all-source $\text{PM}_{2.5}$ obtained from MCMC and the posterior mean value of $\tilde{Y}_{\text{all}}^{(r)}$ can be found in Figure S4.

3.3. Wildland Fire-Specific $\text{PM}_{2.5}$ Concentrations. We mapped the annual fire-specific $\text{PM}_{2.5}$ concentrations associated with different $F^{(q)}(g_b, d_j)$, $q = 1, 2, 3$ in Figure 4a–c to demonstrate the effect of the conversion ratio on the spatial distribution of fire-related $\text{PM}_{2.5}$. Overall, heavy fire-related $\text{PM}_{2.5}$ were concentrated in the middle and southern parts of the CMAQ model domain, including the eastern Kentucky, the middle-eastern region of Alabama, southeastern North Carolina, and western Georgia, which are primarily characterized as forest lands. The highest fire-contributed annual average of $\text{PM}_{2.5}$

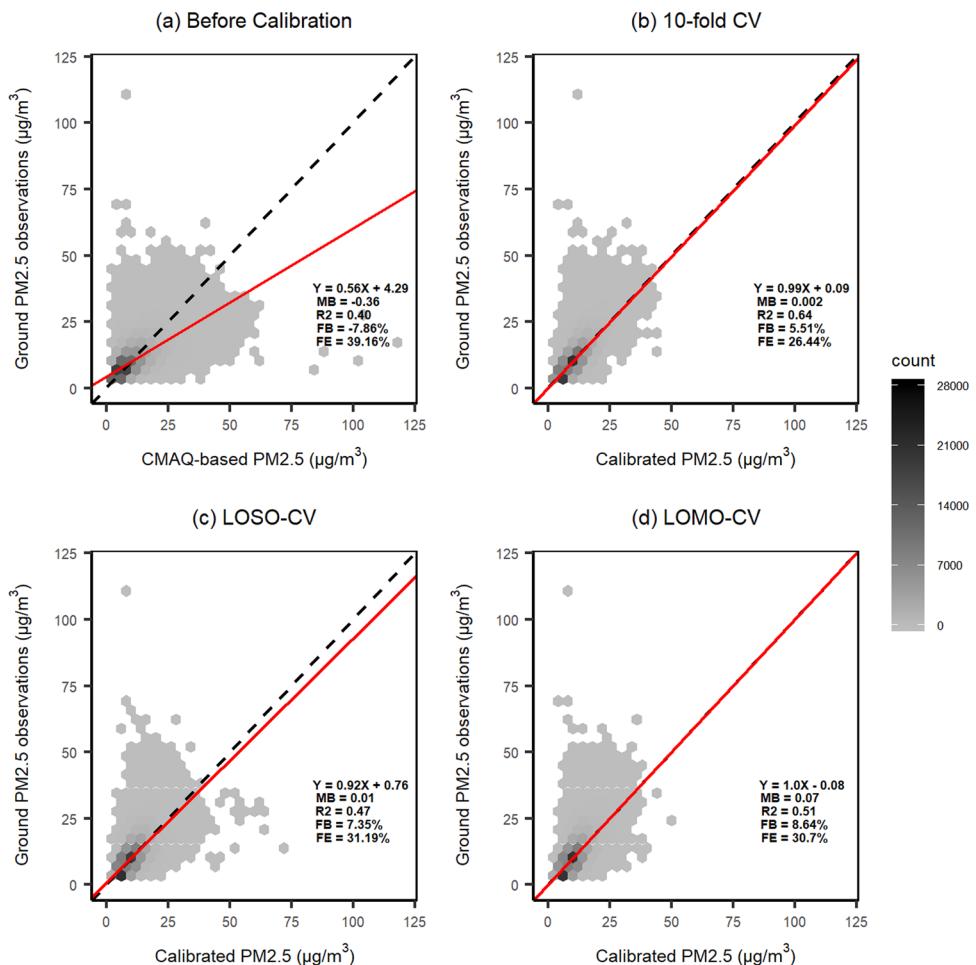


Figure 3. (a) Comparison of daily observed vs CMAQ-modeled PM_{2.5} ($\mu\text{g}/\text{m}^3$) before calibration; calibration performance evaluation using the (b) 10-fold CV method, (c) LOSO CV method, and (d) LOMO CV method. The solid red line represents the linear regression between observed and modeled/calibrated PM_{2.5} concentrations. The dashed black line represents a 1:1 relationship between observed and modeled/calibrated PM_{2.5} concentrations.

values was found in Croatan National Forest in North Carolina in all $F^{(q)}$ for $q = 1, 2, 3$ (see Figure 4a–c). This forest land also concurred with the locations at which about 78 fire events were detected by MODIS. In other areas, especially over the northern part of the study region, the annual wildland fire-specific PM_{2.5} levels were near zero. In comparison to the annual PM_{2.5} associated with $F^{(2)}$ in Figure 4b, the PM_{2.5} concentrations were substantially reduced when $F^{(1)}$ was used, especially over the southern states as shown in Figure 4a. Finally, the application of $F^{(3)}$ resulted in a substantial increase in wildland fire-related PM_{2.5} concentrations across the central and southern regions of the study area (see Figure 4c). The daily variations of the spatial averaged fire-specific PM_{2.5} concentrations associated with three conversion ratios are presented in Figure 4d, whose values varied in the range between 0 and 4.66 $\mu\text{g}/\text{m}^3$ forming a cyclical pattern. Comparatively, fire-specific PM_{2.5} concentrations were high in the spring and winter seasons, but dropped to almost zero in the summer and fall. As expected, the wildland fire-specific PM_{2.5} predictions based on $F^{(3)}$ generated the highest daily fire-related PM_{2.5} concentrations, followed by $F^{(2)}$ and $F^{(1)}$.

We assessed the uncertainties in fire-specific PM_{2.5} predictions $\tilde{Y}_{\text{fire}}^{(q,r)}(g_k, d_j)$ with respect to the two components, all-source PM_{2.5} $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ and the day and grid cell-specific conversion ratio $F^{(q)}(g_k, d_j)$. The uncertainties in $\tilde{Y}_{\text{fire}}^{(q,r)}(g_k, d_j)$ due to the error-contaminated $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ are summarized in Figure

4e (see Section 2.2.2 for details). The total number of days with high uncertainty varied from 0 to 69 days, where zero indicated that small uncertainty existed at the location throughout the year as the fire-specific PM_{2.5} concentrations remained comparatively stable among the realizations. The relatively high uncertainties in fire-specific PM_{2.5} predictions were found in the southeastern North Carolina and eastern Kentucky, as well as the southern border of the study domain, which experienced heavily fire-related pollution. Contrarily, the prediction uncertainties in the northern part of the CMAQ model domain and over the North Atlantic Ocean were low.

When both $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ and $F^{(q)}(g_k, d_j)$ are considered, the levels of uncertainty ranged from 1 to 142 days, as shown in Figure 4f. The relatively high uncertainties were found in the central and southern parts of the study area. The maximum uncertainty level (142 days) associated with both sources of $F^{(q)}$ and $\tilde{Y}_{\text{all}}^{(r)}$ was almost twice larger than that (69 days) from a single source $\tilde{Y}_{\text{all}}^{(r)}$. Moreover, the fraction of grid cells with high uncertainty (>65 days) increased from 0.02 to 8.59% as both factors were taken into account.

3.4. Fire-Related RHA Estimations and Uncertainty Assessment. As a baseline, we estimated a total of 1397 annual RHA associated with the exposure to wildland fire-specific PM_{2.5} concentrations. The spatial distribution of the RHA estimates in Figure 5a indicated that RHA cases were concentrated in

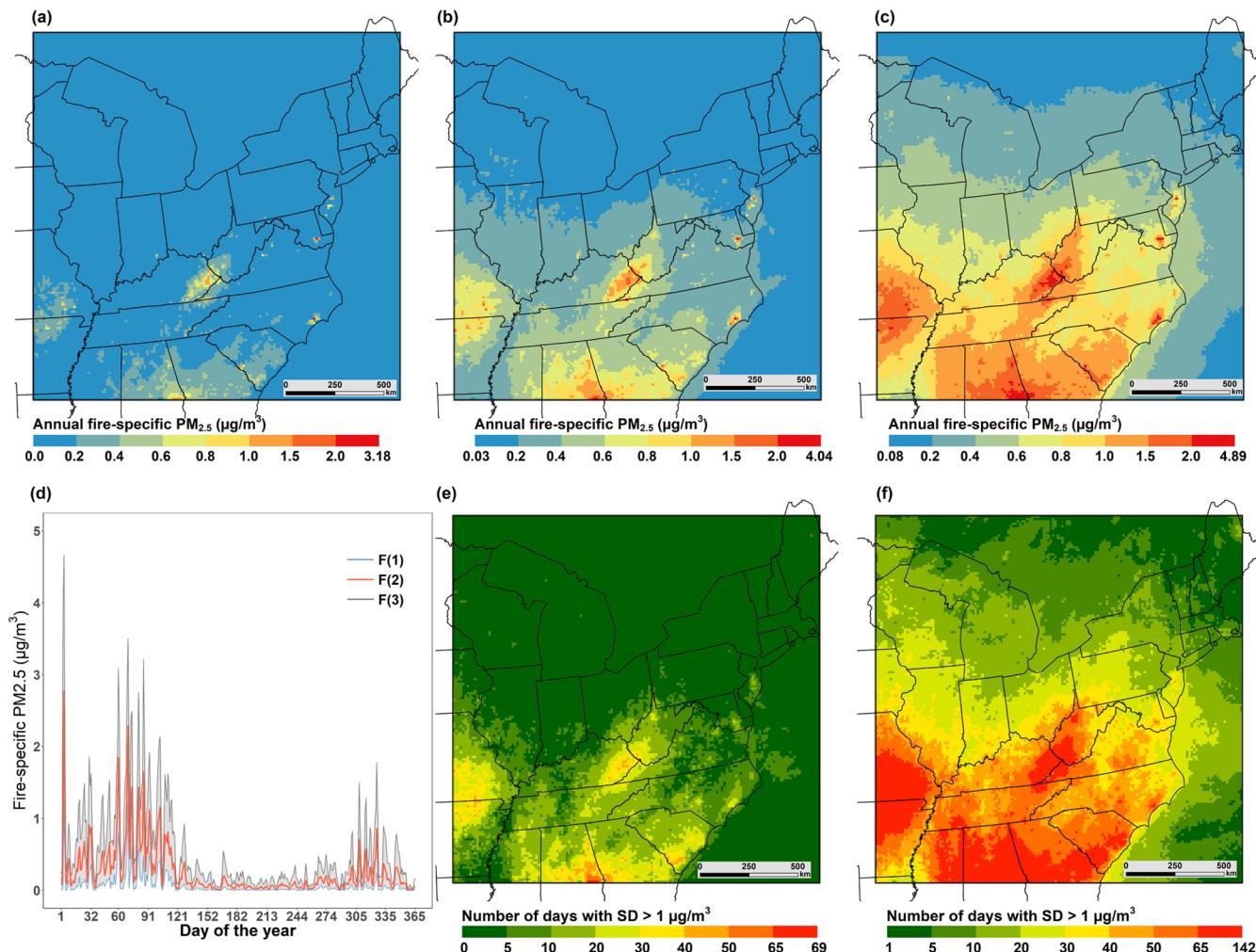


Figure 4. Posterior mean of annual wildland fire-specific PM_{2.5} concentrations based on $\tilde{Y}_{all}^{(r)}$ and (a) $F^{(1)}$; (b) $F^{(2)}$; and (c) $F^{(3)}$; (d) daily variations of the spatial average of wildland fire-specific PM_{2.5} concentrations with $F^{(1)}$ (blue), $F^{(2)}$ (red), and $F^{(3)}$ (gray) between January 1 and December 31, 2014. Shaded area (gray) denotes the range between daily fire-specific PM_{2.5} concentrations with $F^{(1)}$ and $F^{(3)}$; number of days with high uncertainty in wildland fire-specific PM_{2.5} predictions due to (e) uncertain $\tilde{Y}_{all}^{(r)}$; (f) both uncertain $\tilde{Y}_{all}^{(r)}$ and $F^{(q)}$, throughout 365 days.

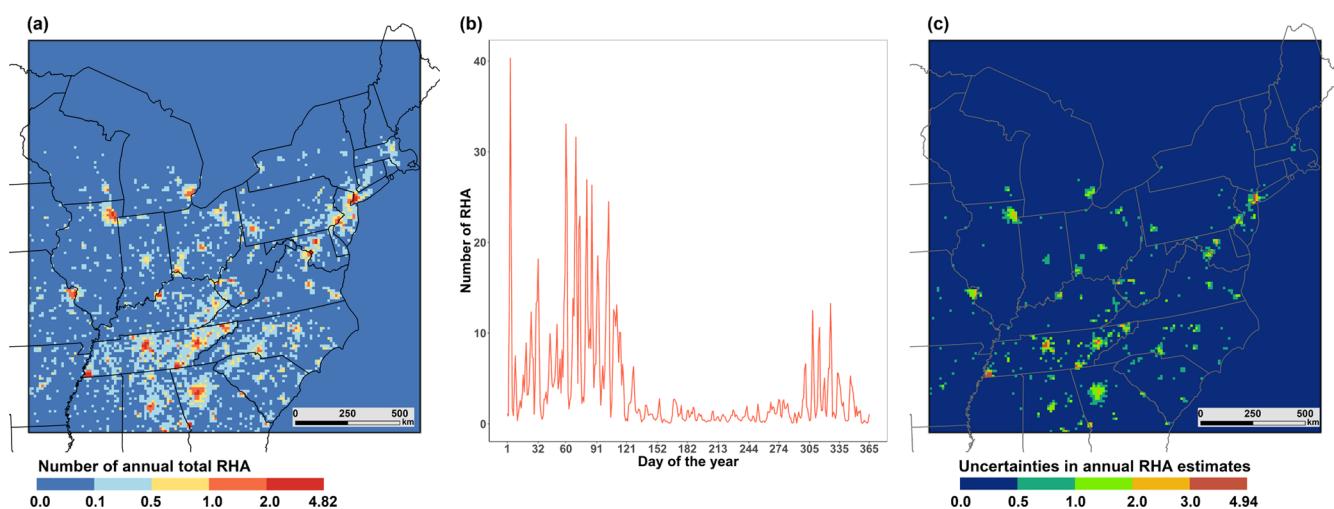


Figure 5. (a) Annual RHA attributable to wildland fire-related PM_{2.5} concentrations; (b) daily variations of wildland fire-related RHA; and (c) uncertainties in wildland fire-related RHA estimations propagated from uncertain all-source PM_{2.5} predictions.

densely populated urban centers, such as Atlanta, Nashville, Memphis, and NYC. The area with the highest RHA estimate was located in the city of Knoxville, Tennessee, although its

annual wildland fire-related PM_{2.5} value was less than 1 $\mu\text{g}/\text{m}^3$. Comparatively small numbers of RHA estimates were found in forest lands, such as the Croatan National Forest, despite the

high PM_{2.5} levels attributed to wildland fire. We found that there were large discrepancies between the annual fire-specific PM_{2.5} (Figure 4b) and the fire-related RHA (Figure 5a), which was likely due to the uneven population density as shown in Figure S2. This difference indicates that health burdens of PM_{2.5} from wildland fires were strongly associated with the population density. In general, the estimated annual total RHAs were higher in urban areas in the southern states, such as Alabama, Tennessee, Georgia, compared to those in the north. The fraction of RHA to the total population at each grid cell is presented in Figure S7. We also examined the daily variability of total RHA attributable to wildland fire-contributed PM_{2.5} over the study region (see Figure 5b). The fire PM_{2.5}-related RHAs were relatively high in the spring and winter seasons, and the estimated number of RHA caused by fire-specific PM_{2.5} was near zero during summer and fall.

We quantified the uncertainties in annual RHA estimates with respect to multiple sources of uncertainties, including all-source PM_{2.5} predictions $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$, conversion ratio $F^{(q)}(g_k, d_j)$, and risk coefficients β . We assessed the effect of each source of uncertainty, as well as the impact of a pooled case under the four scenarios discussed in Section 2.3.2. The uncertain all-source PM_{2.5} predictions contributed RHA estimates with a 90% interval of 623–2174. Figure 5c shows the spatial distribution of uncertainties associated with fire-related RHA (quantified by the difference between 5th and 95th percentiles of RHA) due to uncertain $\tilde{Y}_{\text{all}}^{(r)}(\cdot, \cdot)$. The relatively high uncertainties (>3.0) occurred primarily in populous urban centers, especially over the south, whereas the uncertainties in the northern states of the U.S., including Maine, New Hampshire, Vermont, and Wisconsin, were relatively low. The uncertainty in RHA estimates associated with the uncertainty of risk coefficient was slightly smaller than the uncertainty from all-source PM_{2.5} predictions (90% interval: 623–2174) as the 90% interval was narrow (841–1952). Compared to uncertain $\tilde{Y}_{\text{all}}^{(r)}(g_k, d_j)$ and β , conversion ratio $F^{(q)}(g_k, d_j)$ contributed to much larger uncertainties in annual RHA estimates as the upper bound of RHA, 3215, was about 8.71 times larger than the lower bound of RHA estimates, 369. Overall, the uncertainty range was extended to 109–7058 once all possible uncertain factors were taken into account. The spatial distribution of uncertainties in RHA estimates for scenarios 2–4 is presented in Figure S8.

4. DISCUSSION

We proposed a two-step calibration strategy to predict wildland fire-specific PM_{2.5} concentrations using the CMAQ model, while accounting for the uncertainty associated with the predictions. The proposed calibration method based on the Bayesian downscaler model improved the prediction accuracy, and its hierarchical framework enabled us to quantify uncertainty in the wildland fire-specific PM_{2.5} concentrations. The latter constitutes a unique contribution of the present study in comparison to similar works by Fann et al.³³ and Liu et al.,⁴⁶ in which fire-specific PM_{2.5} predictions obtained from atmospheric models were calibrated but the uncertainties in the predictions were ignored.

The results based on the comparison between bias-adjusted all-source PM_{2.5} predictions and ground observations were consistent with the previous findings,^{50,52} in that the calibrated all-source PM_{2.5} predictions achieved better agreement with ground PM_{2.5} observations than raw CMAQ-based PM_{2.5} concentrations, with an improved 10-fold CV R^2 (0.64) and a small MB value (near zero). Although the R^2 values obtained by

two stringent CV methods were reduced to 0.47 for LOSO and 0.51 for LOMO, the accuracy of calibrated all-source PM_{2.5} predictions was still improved in comparison to the raw CMAQ-based outputs with the R^2 of 0.4. The results indicated that the Bayesian downscaler model enabled us to calibrate the CMAQ-based all-source PM_{2.5} predictions over the regions far from monitoring sites and on days without any PM_{2.5} measurements. We were not able to assess the accuracy of the wildland fire-specific PM_{2.5} obtained from stage 2 due to the absence of fire-specific PM_{2.5} measurements, but we expected that the quality improvement in the calibrated all-source PM_{2.5} predictions was likely to reduce the bias in the wildland fire-specific PM_{2.5} predictions.

The analyses indicated high concentrations of the wildland fire-originated PM_{2.5} values over the forest lands in North Carolina, Georgia, Alabama, and Kentucky (see Figure 4a–c), which corresponded to the report published by the National Interagency Fire Center,⁷⁸ in which these four states were noted due to their experiences of the highest number of wildland fire in the eastern U.S., 2014. This correspondence suggests that the proposed two-step calibration strategy enabled us to capture the elevated PM_{2.5} concentrations over fire-impacted areas. Moreover, the temporal analysis showed that the fire-originated PM_{2.5} values were relatively high in spring and winter, which was also supported by fire data obtained from MODIS.⁶⁷

In the health impact assessments, we used the bias-adjusted instead of the raw CMAQ-based outputs to improve the accuracy of final health burden estimates. We assessed the uncertainties in the estimated health impacts with respect to the uncertain risk coefficient and PM_{2.5} predictions. Similarly, Kollanus et al.³⁸ and Fann et al.³⁹ characterized the extent of fire PM_{2.5}-contributed mortality and morbidity that accounted for the uncertainties in β , but they considered PM_{2.5} predictions obtained from atmospheric model simulations being deterministic, and thus uncertainty deriving from modeled pollutant concentrations was ignored. The present study is among the first effort to account for the effect of uncertainties in both risk coefficient and fire PM_{2.5} estimates on the wildland fire-related RHA. We explored the uncertainties in RHA estimates contributed by each uncertain component and found that conversion ratio F had the largest influence in health impact estimates. These results highlighted the significance of using accurate fire emission data in air quality simulations for subsequent fire PM_{2.5}-related health impact estimates. Our baseline estimate of RHA (1397) was lower than the annual average of RHA (6520) during the 5 year period between 2008 and 2012 in the U.S.,³⁹ but this difference was expected because the population in the study domain was about 52.87% the size of the population in the U.S. and the contribution of wildland fire to ambient PM_{2.5} concentration was less obvious in the eastern U.S. than in western and southern U.S.^{8,34}

Our study is not short of limitations, some of which may have considerable impacts on the conclusions. First, we applied the Bayesian downscaler method to calibrate CMAQ outputs by relating ground observations with modeled PM_{2.5} using spatially and temporally varying coefficients. On the other hand, we assumed that the spatial and temporal structures among measured PM_{2.5} and two additional covariates were static over the study domain during the study period, which may not be realistic in certain circumstances.⁷⁹ This issue might be resolved by dividing the study area into subregions and/or over a shorter time period and applying the proposed calibration method separately, if there is sufficient evidence for the nonstationarity.

Second, wildfire aerosols possibly can be transported up to thousands of kilometers, thereby degrading air quality over regions far from fire centers.^{38,80,81} The CMAQ model was unable to capture fire emissions outside the study domain, such as emissions from fire-prone areas in the western U.S., and this might have resulted in underestimations of fire contribution ratio F . Third, the uncertainty estimations in wildland fire-related PM_{2.5} propagated from F were conservative because the uncertainties in daily conversion ratio were characterized by summary measures of hourly F . One possible solution is to apply various fire emission products, such as Fire INventory from National Center for Atmospheric Research⁸² and Global Fire Emission Database,⁸³ to estimate a plausible range of F . Similarly, the uncertainties in both population and hospitalization rates were not taken into account in health impact assessments, which might have led to the underestimation of overall uncertainty range of health impact estimates. A further investigation is warranted to examine the pooled effect of all possible uncertain components to fire-related RHA estimates. Finally, we did not directly quantify the accuracy of fire-specific PM_{2.5} estimations due to the absence of source-specific PM_{2.5} observations, although source apportionment techniques, such as positive matrix factorization⁸⁴ and receptor-oriented models,^{30,31,85} can be considered to isolate ground PM_{2.5} observations from each emission source. We expect that this approach will allow us to use wildland fire-specific PM_{2.5} estimates from monitor stations as reference data to validate the reliability of calibrated fire-specific PM_{2.5} concentrations.

Notwithstanding these limitations, this study has several strengths and implications. First, we estimated PM_{2.5} concentrations specific to wildland fires using the CMAQ model by separating PM_{2.5} concentrations of wildland fires from other emission sources. We further improved the prediction accuracy using a two-step calibration strategy to adjust the biases in the raw CMAQ-based outputs. The proposed method is easily applicable to other studies of source-specific PM_{2.5} predictions over different regions and periods. Second, we assessed the uncertainties in wildland fire-specific PM_{2.5} predictions. Future work includes the identification of areas with high uncertainty to further improve the accuracy of PM_{2.5} exposure estimates and increase reliability. Finally, we provided a range of fire-related health impacts by taking into account the uncertainties introduced by all-source PM_{2.5} predictions, conversion ratio, and risk coefficient. This information has the potential to assist state and local regulatory agencies to develop a cost-effective and targeted strategy for reducing risks of fire-related adverse health outcomes.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.est.9b02660](https://doi.org/10.1021/acs.est.9b02660).

WRF model configuration settings, the definition of statistical indicators, the description of event types labeled by AQS monitoring sites, the method of β estimations, the spatial distribution of population data, a workflow of performing uncertainty-aware wildland fire-related health impact assessments, the spatial distribution of calibrated all-source PM_{2.5} concentrations, the sensitivity analysis of cutoff values, the spatial distribution of the fraction of RHAs to total population, and the spatial distribution of

uncertainties in fire-related RHA estimates with respect to multiple uncertain components ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

*E-mail: xiangyuj@buffalo.edu.

ORCID

Xiangyu Jiang: [0000-0002-6620-4984](https://orcid.org/0000-0002-6620-4984)

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors appreciate the support provided by the Center for Computational Research and the Research and Education in Energy, Environment & Water (RENEW) seed grant at the University at Buffalo. The authors thank EPA for sharing their WRF model outputs on the manuscript, and Dr. Howard Chang for sharing the R codes of Bayesian downscaler model.

REFERENCES

- (1) Abatzoglou, J. T.; Williams, A. P. Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 11770–11775.
- (2) Westerling, A. L. Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Philos. Trans. R. Soc., B* **2016**, *371*, No. 20150178.
- (3) NIFC (US National Interagency Fire Center). Total Wildland Fires and Acres (1926–2018). https://www.nifc.gov/fireInfo/fireInfo_stats_totalFires.html (accessed July 17, 2019).
- (4) 2014 National Emissions Inventory, version 2 Technical Support Document; EPA: Research Triangle Park, NC, 2018. https://www.epa.gov/sites/production/files/2018-07/documents/nei2014v2_tsd_05jul2018.pdf.
- (5) 2014 National Emissions Inventory (NEI) Data. <https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data> (accessed July 4, 2019).
- (6) Bytnerowicz, A.; Hsu, Y.-M.; Percy, K.; Legge, A.; Fenn, M. E.; Schilling, S.; Frączek, W.; Alexander, D. Ground-level air pollution changes during a boreal wildland mega-fire. *Sci. Total Environ.* **2016**, *572*, 755–769.
- (7) Jaffe, D. A.; Wigder, N.; Downey, N.; Pfister, G.; Boynard, A.; Reid, S. B. Impact of wildfires on ozone exceptional events in the western US. *Environ. Sci. Technol.* **2013**, *47*, 11065–11072.
- (8) O'Dell, K.; Ford, B.; Fischer, E. V.; Pierce, J. R. Contribution of wildland-fire smoke to US PM_{2.5} and its influence on recent trends. *Environ. Sci. Technol.* **2019**, *53*, 1797–1804.
- (9) Le, G.; Breysse, P.; McDermott, A.; Eftim, S.; Geyh, A.; Berman, J.; Curriero, F. Canadian forest fires and the effects of long-range transboundary air pollution on hospitalizations among the elderly. *ISPRS Int. J. Geo-Inf.* **2014**, *3*, 713–731.
- (10) Tinling, M. A.; West, J. J.; Cascio, W. E.; Kilaru, V.; Rappold, A. G. Repeating cardiopulmonary health effects in rural North Carolina population during a second large peat wildfire. *Environ. Health* **2016**, *15*, No. 12.
- (11) Wettstein, Z. S.; Hoshiko, S.; Fahimi, J.; Harrison, R. J.; Cascio, W. E.; Rappold, A. G. Cardiovascular and cerebrovascular emergency department visits associated with wildfire smoke exposure in California in 2015. *J. Am. Heart Assoc.* **2018**, *7*, No. e007492.
- (12) Naehler, L. P.; Brauer, M.; Lipsett, M.; Zelikoff, J. T.; Simpson, C. D.; Koenig, J. Q.; Smith, K. R. Woodsmoke health effects: A review. *Inhalation Toxicol.* **2007**, *19*, 67–106.
- (13) Elliott, C. T.; Henderson, S. B.; Wan, V. Time series analysis of fine particulate matter and asthma reliever dispensations in populations affected by forest fires. *Environ. Health* **2013**, *12*, No. 11.
- (14) Alman, B. L.; Pfister, G.; Hao, H.; Stowell, J.; Hu, X.; Liu, Y.; Strickland, M. J. The association of wildfire smoke with respiratory and

- cardiovascular emergency department visits in Colorado in 2012: A case crossover study. *Environ. Health* **2016**, *15*, No. 64.
- (15) Haikerwal, A.; Akram, M.; Sim, M. R.; Meyer, M.; Abramson, M. J.; Dennekamp, M. Fine particulate matter ($PM_{2.5}$) exposure during a prolonged wildfire period and emergency department visits for asthma. *Respirology* **2016**, *21*, 88–94.
- (16) Reid, C. E.; Brauer, M.; Johnston, F. H.; Jerrett, M.; Balmes, J. R.; Elliott, C. T. Critical review of health impacts of wildfire smoke exposure. *Environ. Health Perspect.* **2016**, *124*, 1334–1343.
- (17) Johnston, F. H.; Purdie, S.; Jalaludin, B.; Martin, K. L.; Henderson, S. B.; Morgan, G. G. Air pollution events from forest fires and emergency department attendances in Sydney, Australia 1996–2007: A case-crossover analysis. *Environ. Health* **2014**, *13*, No. 105.
- (18) Yao, J.; Eyamie, J.; Henderson, S. B. Evaluation of a spatially resolved forest fire smoke model for population-based epidemiologic exposure assessment. *J. Exposure Sci. Environ. Epidemiol.* **2016**, *26*, 233.
- (19) Liu, J. C.; Pereira, G.; Uhl, S. A.; Bravo, M. A.; Bell, M. L. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. *Environ. Res.* **2015**, *136*, 120–132.
- (20) Gan, R. W.; Ford, B.; Lassman, W.; Pfister, G.; Vaidyanathan, A.; Fischer, E.; Volckens, J.; Pierce, J. R.; Magzamen, S. Comparison of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. *GeoHealth* **2017**, *1*, 122–136.
- (21) 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.; Meredith, J. T.; Vaughan-Batten, H.; Deyneka, L.; Devlin, R. B. Peat bog wildfire smoke exposure in rural North Carolina is associated with cardiopulmonary emergency department visits assessed through syndromic surveillance. *Environ. Health Perspect.* **2011**, *119*, 1415.
- (22) About AQs Data. https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (accessed July 4, 2019).
- (23) Lassman, W.; Ford, B.; Gan, R. W.; Pfister, G.; Magzamen, S.; Fischer, E. V.; Pierce, J. R. Spatial and temporal estimates of population exposure to wildfire smoke during the Washington state 2012 wildfire season using blended model, satellite, and in situ data. *GeoHealth* **2017**, *1*, 106–121.
- (24) Van Donkelaar, A.; Martin, R. V.; Levy, R. C.; da Silva, A. M.; Krzyzanowski, M.; Chubarova, N. E.; Semutnikova, E.; Cohen, A. J. Satellite-based estimates of ground-level fine particulate matter during extreme events: A case study of the Moscow fires in 2010. *Atmos. Environ.* **2011**, *45*, 6225–6232.
- (25) Di, Q.; Kloog, I.; Kourakis, P.; Lyapustin, A.; Wang, Y.; Schwartz, J. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environ. Sci. Technol.* **2016**, *50*, 4712–4721.
- (26) Kloog, I.; Nordio, F.; Coull, B. A.; Schwartz, J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the Mid-Atlantic states. *Environ. Sci. Technol.* **2012**, *46*, 11913–11921.
- (27) Chu, Y.; Liu, Y.; Li, X.; Liu, Z.; Lu, H.; Lu, Y.; Mao, Z.; Chen, X.; Li, N.; Ren, M.; Liu, F.; Tian, L.; Zhu, Z.; Xiang, H. A review on predicting ground $PM_{2.5}$ concentration using satellite aerosol optical depth. *Atmosphere* **2016**, *7*, No. 129.
- (28) Engel-Cox, J. A.; Holloman, C. H.; Coutant, B. W.; Hoff, R. M. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmos. Environ.* **2004**, *38*, 2495–2509.
- (29) Xiao, Q.; Wang, Y.; Chang, H. H.; Meng, X.; Geng, G.; Lyapustin, A.; Liu, Y. Full-coverage high-resolution daily $PM_{2.5}$ estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sens. Environ.* **2017**, *199*, 437–446.
- (30) Nikolov, M. C.; Coull, B. A.; Catalano, P. J.; Godleski, J. J. An informative Bayesian structural equation model to assess source-specific health effects of air pollution. *Biostatistics* **2007**, *8*, 609–624.
- (31) Park, E. S.; Oh, M.-S. Accounting for uncertainty in source-specific exposures in the evaluation of health effects of pollution sources on daily cause-specific mortality. *Environmetrics* **2018**, *29*, No. e2484.
- (32) Wegesser, T. C.; Pinkerton, K. E.; Last, J. A. California wildfires of 2008: Coarse and fine particulate matter toxicity. *Environ. Health Perspect.* **2009**, *117*, 893–897.
- (33) Fann, N.; Fulcher, C. M.; Baker, K. The recent and future health burden of air pollution apportioned across US sectors. *Environ. Sci. Technol.* **2013**, *47*, 3580–3589.
- (34) Wilkins, J. L.; Pouliot, G.; Foley, K.; Appel, W.; Pierce, T. The impact of US wildland fires on ozone and particulate matter: A comparison of measurements and CMAQ model predictions from 2008 to 2012. *Int. J. Wildland Fire* **2018**, *27*, 684–698.
- (35) Byun, D.; Schere, K. L. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Appl. Mech. Rev.* **2006**, *59*, 51–77.
- (36) Lelieveld, J.; Evans, J. S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367.
- (37) Johnston, F. H.; Henderson, S. B.; Chen, Y.; Randerson, J. T.; Marlier, M.; DeFries, R. S.; Kinney, P.; Bowman, D. M.; Brauer, M. Estimated global mortality attributable to smoke from landscape fires. *Environ. Health Perspect.* **2012**, *120*, 695.
- (38) Kollanus, V.; Prank, M.; Gens, A.; Soares, J.; Vira, J.; Kukkonen, J.; Sofiev, M.; Salonen, R. O.; Lanki, T. Mortality due to vegetation fire-originated $PM_{2.5}$ exposure in Europe—assessment for the years 2005 and 2008. *Environ. Health Perspect.* **2017**, *125*, 30–37.
- (39) Fann, N.; Alman, B.; Broome, R. A.; Morgan, G. G.; Johnston, F. H.; Pouliot, G.; Rappold, A. G. The health impacts and economic value of wildland fire episodes in the US: 2008–2012. *Sci. Total Environ.* **2018**, *610–611*, 802–809.
- (40) Crooks, J. L.; Özkan, H. Simultaneous statistical bias correction of multiple $PM_{2.5}$ species from a regional photochemical grid model. *Atmos. Environ.* **2014**, *95*, 126–141.
- (41) Gilliam, R. C.; Hogrefe, C.; Godowitch, J. M.; Napelenok, S.; Mathur, R.; Rao, S. T. Impact of inherent meteorology uncertainty on air quality model predictions. *J. Geophys. Res.: Atmos.* **2015**, *120*, 12259.
- (42) Jiang, X.; Yoo, E.-H. The importance of spatial resolutions of Community Multiscale Air Quality (CMAQ) models on health impact assessment. *Sci. Total Environ.* **2018**, *627*, 1528–1543.
- (43) Wang, M.; Sampson, P. D.; Hu, J.; Kleeman, M.; Keller, J. P.; Olives, C.; Szapiro, A. A.; Vedal, S.; Kaufman, J. D. Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and $PM_{2.5}$. *Environ. Sci. Technol.* **2016**, *50*, 5111–5118.
- (44) Solazzo, E.; Riccio, A.; Van Dingenen, R.; Valentini, L.; Galmarini, S. Evaluation and uncertainty estimation of the impact of air quality modelling on crop yields and premature deaths using a multi-model ensemble. *Sci. Total Environ.* **2018**, *633*, 1437–1452.
- (45) Maraun, D. Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *J. Clim.* **2013**, *26*, 2137–2143.
- (46) Liu, J. C.; Wilson, A.; Mickley, L. J.; Dominici, F.; Ebisu, K.; Wang, Y.; Sulprizio, M. P.; Peng, R. D.; Yue, X.; Son, J.-Y.; Anderson, G. B.; Bell, M. L. Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. *Epidemiology* **2017**, *28*, 77–85.
- (47) Chen, G.; Li, J.; Ying, Q.; Sherman, S.; Perkins, N.; Sundaram, R.; Mendola, P. Evaluation of observation-fused regional air quality model results for population air pollution exposure estimation. *Sci. Total Environ.* **2014**, *485–486*, 563–574.
- (48) Berrocal, V. J.; Gelfand, A. E.; Holland, D. M. A spatio-temporal downscaler for output from numerical models. *J. Agric. Biol. Environ. Stat.* **2010**, *15*, 176–197.
- (49) Chang, H. H.; Hu, X.; Liu, Y. Calibrating MODIS aerosol optical depth for predicting daily $PM_{2.5}$ concentrations via statistical downscaling. *J. Exposure Sci. Environ. Epidemiol.* **2014**, *24*, 398.
- (50) Murray, N.; Chang, H. H.; Holmes, H.; Liu, Y. Combining Satellite Imagery and Numerical Model Simulation to Estimate Ambient Air Pollution: An Ensemble Averaging Approach. 2018, arXiv:1802.03077. arXiv.org e-Print archive. <https://arxiv.org/abs/1802.03077>.

- (51) Chang, H. H.; Reich, B. J.; Miranda, M. L. Time-to-event analysis of fine particle air pollution and preterm birth: Results from North Carolina, 2001–2005. *Am. J. Epidemiol.* **2012**, *175*, 91–98.
- (52) Geng, G.; Murray, N. L.; Tong, D.; Fu, J. S.; Hu, X.; Lee, P.; Meng, X.; Chang, H. H.; Liu, Y. Satellite-based daily PM_{2.5} estimates during fire seasons in Colorado. *J. Geophys. Res.: Atmos.* **2018**, *123*, 8159–8171.
- (53) CMAQ, version 5.2.1. <https://www.cmascenter.org/>.
- (54) Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D. M.; Wang, W.; Powers, J. G. *A Description of the Advanced Research WRF*, NCAR Technical Notes. Note NCAR/TN-475+STR; NCAR, 2008.
- (55) *Weather Research Forecast, Meteorological Model Evaluation Annual 2014 12-km CONUS*, version 3.8; UNC: Chapel Hill, 2016.
- (56) *Air Emissions Modeling Platform*, 2014 version 7.1. <ftp://newftp.epa.gov/air/emismod/2014/v2/> (accessed Oct 31, 2018).
- (57) *Technical Support Document (TSD). Preparation of Emissions Inventories for the Version 7.1, 2014 Emissions Modeling Platform for the National Air Toxics Assessment*; EPA: Research Triangle Park, NC, 2018.
- (58) Larkin, N. K.; O'Neill, S. M.; Solomon, R.; Raffuse, S.; Strand, T.; Sullivan, D. C.; Krull, C.; Rorig, M.; Peterson, J.; Ferguson, S. A. The BlueSky smoke modeling framework. *Int. J. Wildland Fire* **2009**, *18*, 906–920.
- (59) Raffuse, S. M.; Larkin, N. K.; Lahm, P. W.; Du, Y. Development of Version 2 of the Wildland Fire Portion of the National Emissions Inventory. In *20th International Emission Inventory Conference*; Tampa, Florida, 2012.
- (60) Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System. <http://www.smoke-model.org/>.
- (61) U.S. Environmental Protection Agency. *Air Quality System Data Mart*. https://aqs.epa.gov/aqswb/airdata/download_files.html (accessed Dec 4, 2018).
- (62) 2014 TIGER/Line Shapefiles: Urban Areas. <https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2014&layergroup=Urban+Areas> (accessed June 25, 2019).
- (63) AirData Download Files Documentation. <https://aqs.epa.gov/aqswb/airdata/FileFormats.html> (accessed Dec 31, 2018).
- (64) Reid, C. E.; Jerrett, M.; Petersen, M. L.; Pfister, G. G.; Morefield, P. E.; Tager, I. B.; Raffuse, S. M.; Balmes, J. R. Spatiotemporal prediction of fine particulate matter during the 2008 Northern California wildfires using machine learning. *Environ. Sci. Technol.* **2015**, *49*, 3887–3896.
- (65) U.S. Geological Survey. *NLCD 2011 Land Cover*, 2011 ed. <https://www.mrlc.gov/data> (accessed June 25, 2019).
- (66) Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a Decade of Land Cover Change Information. *Photogrammetric Engineering and Remote Sensing*; American Society for Photogrammetry and Remote Sensing: Bethesda, MD, 2015; Vol. 81, pp 345–354.
- (67) Geospatial Technology and Applications Center. *MODIS Fire Detection GIS Data*. <https://fsapps.nwcg.gov/afm/gisdata.php> (accessed June 25, 2019).
- (68) Sacks, J. D.; Lloyd, J. M.; Zhu, Y.; Anderton, J.; Jang, C. J.; Hubbell, B.; Fann, N. The Environmental Benefits Mapping and Analysis Program-Community Edition (BenMAP-CE): A tool to estimate the health and economic benefits of reducing air pollution. *Environ. Modell. Software* **2018**, *104*, 118–129.
- (69) Delfino, R. J.; Brummel, S.; Wu, J.; Stern, H.; Ostro, B.; Lipsett, M.; Winer, A.; Street, D. H.; Zhang, L.; Tjoa, T.; Gillen, D. L. The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. *Occup. Environ. Med.* **2009**, *66*, 189–197.
- (70) Healthcare Cost and Utilization Project. <https://hcupnet.ahrq.gov/#setup> (accessed June 25, 2019).
- (71) U.S. Environmental Protection Agency (USEPA). *Environmental Benefits Mapping and Analysis Program – Community Edition (BenMAP-CE): User's Manual*, Research Triangle Park, NC (2017). https://www.epa.gov/sites/production/files/2015-04/documents/benmap-ce_user_manual_march_2015.pdf (accessed September 26, 2019).
- (72) PopGrid. <https://www.epa.gov/benmap/benmap-community-edition> (accessed March 9, 2019).
- (73) Chart-Asa, C.; Gibson, J. M. Health impact assessment of traffic-related air pollution at the urban project scale: Influence of variability and uncertainty. *Sci. Total Environ.* **2015**, *506–507*, 409–421.
- (74) Kodros, J.; Carter, E.; Brauer, M.; Volckens, J.; Bilsback, K.; L'Orange, C.; Johnson, M.; Pierce, J. Quantifying the contribution to uncertainty in mortality attributed to household, ambient, and joint exposure to PM_{2.5} from residential solid fuel use. *GeoHealth* **2018**, *2*, 25–39.
- (75) Morris, R. E.; McNally, D. E.; Tesche, T. W.; Tonnesen, G.; Boylan, J. W.; Brewer, P. Preliminary evaluation of the Community Multiscale Air Quality model for 2002 over the southeastern United States. *J. Air Waste Manage. Assoc.* **2005**, *55*, 1694–1708.
- (76) Lv, B.; Hu, Y.; Chang, H. H.; Russell, A. G.; Cai, J.; Xu, B.; Bai, Y. Daily estimation of ground-level PM_{2.5} concentrations at 4 km resolution over Beijing-Tianjin-Hebei by fusing MODIS AOD and ground observations. *Sci. Total Environ.* **2017**, *580*, 235–244.
- (77) Xu, H.; Bechle, M. J.; Wang, M.; Szpiro, A. A.; Vedral, S.; Bai, Y.; Marshall, J. D. National PM_{2.5} and NO₂ exposure models for China based on land use regression, satellite measurements, and universal kriging. *Sci. Total Environ.* **2019**, *655*, 423–433.
- (78) National Report of Wildland Fires and Acres Burned by State. https://www.predictiveservices.nifc.gov/intelligence/2014_Statsumm/fires_acres14.pdf (accessed Dec 31, 2018).
- (79) Wang, Y.; Hu, X.; Chang, H.; Waller, L.; Belle, J.; Liu, Y. A Bayesian downscaler model to estimate daily PM_{2.5} levels in the conterminous US. *Int. J. Environ. Res. Public Health* **2018**, *15*, No. 1999.
- (80) Sapkota, A.; Symons, J. M.; Kleissl, J.; Wang, L.; Parlange, M. B.; Ondov, J.; Breysse, P. N.; Diette, G. B.; Eggleston, P. A.; Buckley, T. J. Impact of the 2002 Canadian forest fires on particulate matter air quality in Baltimore city. *Environ. Sci. Technol.* **2005**, *39*, 24–32.
- (81) Wang, Y.; Huang, J.; Zananski, T. J.; Hopke, P. K.; Holsen, T. M. Impacts of the Canadian forest fires on atmospheric mercury and carbonaceous particles in northern New York. *Environ. Sci. Technol.* **2010**, *44*, 8435–8440.
- (82) Fire Inventory from NCAR. <https://www2.acom.ucar.edu/modeling/finn-fire-inventory-ncar> (accessed July 4, 2019).
- (83) Global Fire Emission Database. <https://www.globalfiredata.org/> (accessed July 4, 2019).
- (84) Dutton, S. J.; Vedral, S.; Piedrahita, R.; Milford, J. B.; Miller, S. L.; Hannigan, M. P. Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated PM_{2.5}. *Atmos. Environ.* **2010**, *44*, 2731–2741.
- (85) Ivey, C.; Holmes, H.; Hu, Y.; Mulholland, J.; Russell, A. Development of PM_{2.5} source impact spatial fields using a hybrid source apportionment air quality model. *Geosci. Model Dev.* **2015**, *8*, 2153–2165.