Atmos. Chem. Phys., 21, 11243–11256, 2021
https://doi.org/10.5194/acp-21-11243-2021

Atmospheric Chemistry and Physics
Open Access
EGU

Satellite-based estimation of the impacts of summertime wildfires on $PM_{2.5}$ concentration in the United States

Zhixin Xue[1], **Pawan Gupta**[2,3], **and Sundar Christopher**[1]

[1]Department of Atmospheric and Earth Science, The University of Alabama in Huntsville, Huntsville, 35806 AL, USA
[2]STI, Universities Space Research Association (USRA), Huntsville, 35806 AL, USA
[3]NASA Marshall Space Flight Center, Huntsville, 35806 AL, USA

Correspondence: Sundar Christopher (sundar@nsstc.uah.edu)

Received: 4 November 2020 – Discussion started: 9 December 2020
Revised: 7 June 2021 – Accepted: 21 June 2021 – Published: 27 July 2021

Abstract. Frequent and widespread wildfires in the northwestern United States and Canada have become the "new normal" during the Northern Hemisphere summer months, which significantly degrades particulate matter air quality in the United States. Using the mid-visible Multi Angle Implementation of Atmospheric Correction (MAIAC) satellite-derived aerosol optical depth (AOD) with meteorological information from the European Centre for Medium-Range Weather Forecasts (ECMWF) and other ancillary data, we quantify the impact of these fires on fine particulate matter concentration ($PM_{2.5}$) air quality in the United States. We use a geographically weighted regression (GWR) method to estimate surface $PM_{2.5}$ in the United States between low (2011) and high (2018) fire activity years. Our results indicate an overall leave-one-out cross-validation (LOOCV) R^2 value of 0.797 with root mean square error (RMSE) between 3 and 5 $\mu g\,m^{-3}$. Our results indicate that smoke aerosols caused significant pollution changes over half of the United States. We estimate that nearly 29 states have increased $PM_{2.5}$ during the fire-active year and that 15 of these states have $PM_{2.5}$ concentrations more than 2 times that of the inactive year. Furthermore, these fires increased the daily mean surface $PM_{2.5}$ concentrations in Washington and Oregon by 38 to 259 $\mu g\,m^{-3}$, posing significant health risks especially to vulnerable populations. Our results also show that the GWR model can be successfully applied to $PM_{2.5}$ estimations from wildfires, thereby providing useful information for various applications such as public health assessment.

1 Introduction

The United States (US) Clean Air Act (CAA) was passed in 1970 to reduce pollution levels and protect public health and has led to significant improvements in air quality (Hubbell et al., 2010; Samet, 2011). However, the northern part of the US continues to experience an increase in surface $PM_{2.5}$ due to fires in the northwestern United States and Canada (hereafter NWUSC), especially during the summer months, and these aerosols are a new source of "pollution" (Coogan et al., 2019; Dreessen et al., 2016). The smoke aerosols from these fires increase fine particulate matter ($PM_{2.5}$) concentrations and degrade air quality in the United States (Miller et al., 2011). Moreover, several studies have shown that from 2013 to 2016, over 76 % of Canadians and 69 % of Americans were at least minimally affected by wildfire smoke (Munoz-Alpizar et al., 2017). Although wildfire pre-suppression and suppression costs have increased, the number of large fires and the burnt areas in many parts of western Canada and the United States have also increased (Hanes et al., 2019; Tymstra et al., 2019). Furthermore, in a changing climate, as surface temperature increases and humidity decreases, the flammability of land cover also increases and thus accelerates the spread of wildfires (Melillo et al., 2014). The accumulation of flammable materials like leaf litter can potentially trigger severe wildfire events, even in those forests that hardly experience wildfires (Calkin et al., 2015; Hessburg et al., 2015; Stephens, 2005).

Wildfire smoke exposure can cause small particles to be lodged in lungs, which may lead to exacerbations of asthma chronic obstructive pulmonary disease (COPD), bronchitis, heart disease and pneumonia (Apte et al., 2018; Cascio,

2018). According to a recent study, a $10\,\mu g\,m^{-3}$ increase in $PM_{2.5}$ is associated with a 12.4 % increase in cardiovascular mortality (Kollanus et al., 2016). In addition, exposure to wildfire smoke is also related to massive economic costs due to premature mortality, loss of workforce productivity, impacts on the quality of life and compromised water quality (Meixner and Wohlgemuth, 2004).

Surface $PM_{2.5}$ is one of the most commonly used parameters to assess the health effects of ambient air pollution. Given the sparsity of measurements in many parts of the world, it is not possible to use interpolation techniques between monitors to provide $PM_{2.5}$ estimates on a square kilometer basis. Since surface monitors are limited, satellite data have been used with numerous ancillary datasets to estimate surface $PM_{2.5}$ at various spatial scales. Several techniques have been developed to estimate surface $PM_{2.5}$ using satellite observations from regional to global scales, including simple linear regression, multiple linear regression, mixed-effect models, chemical transport models (scaling methods), geographically weighted regression (GWR), and machine-learning methods (see Hoff and Christopher, 2009, for a review). The commonly used global satellite data product is the 550 nm (mid-visible) aerosol optical depth (AOD), which is a unitless columnar measure of aerosol extinction. Simple linear regression methods use satellite AOD as the only independent variable, which shows limited predictability compared to other methods, and correlation coefficients vary from 0.2 to 0.6 from the western to eastern United States (Zhang et al., 2009). Multiple linear regression methods use numerous variables along with AOD data, and the prediction accuracy varies with different conditions, including the height of the boundary layer and other meteorological conditions (Goldberg et al., 2019; Gupta and Christopher, 2009b; Liu et al., 2005). For both univariate and multi-variate models, the AOD shows stronger correlation with $PM_{2.5}$ during fire episodes compared to pre-fire and post-fire periods (Mirzaei et al., 2018). Chemistry transport models (CTMs) that scale the satellite AOD by the ratio of $PM_{2.5}$ to AOD simulated by models can provide $PM_{2.5}$ estimations without ground measurements, which are different from other statistical methods (Van Donkelaar et al., 2006, 2019). However, the CTM models depend on reliable emission data, show limited predictability at shorter timescales and are largely useful for studies that require annual averages (Hystad et al., 2012). Different machine-learning methods, including neuron networks, random forests, and deep belief networks, show improvements in prediction accuracy (with CV R^2 values larger than 0.8), which is hard to accomplish for other parametric regression models (Gupta and Christopher, 2009a; Hu et al., 2017; Li et al., 2017; Wei et al., 2019, 2020, 2021). However, these methods also require a large number of samples to train the model, which means it is more suitable for daily $PM_{2.5}$ estimation rather than short-term wildfire events with relatively low occurrence frequency.

The relationship among $PM_{2.5}$, AOD and other meteorological variables is not spatially consistent (Hoff and Christopher, 2009; Hu, 2009). Therefore, methods that consider spatial variability can replicate surface $PM_{2.5}$ with higher accuracy. One such method is the GWR, which is a non-stationary technique that models spatially varying relationships by assuming that the coefficients in the model are functions of locations (Brunsdon et al., 1996; Fotheringham et al., 1998, 2003). In 2009, satellite-retrieved AOD was introduced in the GWR method to predict surface $PM_{2.5}$ (Hu, 2009), followed by the use of meteorological parameters and land use information (Hu et al., 2013). Meteorological variables are crucial for simulating surface $PM_{2.5}$ since they interact with $PM_{2.5}$ through different processes (Chen et al., 2020), which will be discussed in detail in the data section. Several studies (Guo et al., 2021; Ma et al., 2014; You et al., 2016a) successfully applied the GWR model in estimating $PM_{2.5}$ in China by using AOD and meteorological features as predictors. Similarly to all the statistical methods, however, the GWR relies on an adequate number and density of surface measurements (Chu et al., 2016; Gu, 2019; Guo et al., 2021), underscoring the importance of adequate ground monitoring of surface $PM_{2.5}$.

In this paper, we use satellite data from the Moderate Resolution Imaging Spectroradiometer (MODIS) and surface $PM_{2.5}$ data combined with meteorological and other ancillary information to develop and use the GWR method to estimate $PM_{2.5}$. The use of the GWR method is not novel, and we merely use a proven method to estimate surface $PM_{2.5}$ from forest fires. We calculate the change in $PM_{2.5}$ between high fire activity (2018) and low fire activity (2011) periods during summer to assess the role of NWUSC wildfires in surface $PM_{2.5}$ in the United States. The paper is organized as follows. We describe the datasets used in this study, followed by the GWR method. We then describe the results and discussion, followed by a summary with conclusions.

2 Data

A 17 d period (9 to 25 August) in 2018 (high fire activity) and 2011 (low fire activity) was selected based on analysis of total fires (details in the methodology section) to assess surface $PM_{2.5}$ (Table 1).

2.1 Ground-level $PM_{2.5}$ observations

Daily surface $PM_{2.5}$ from the Environmental Protection Agency (EPA) is used in this study. These data are from Federal Reference Methods (FRM), Federal Equivalent Methods (FEM), or other methods that are to be used in the National Ambient Air Quality Standards (NAAQS) decisions. A total of 1003 monitoring sites in the US are included in our study, with 949 having valid observations in the study period in 2018 and a total of 873 sites with 820 having valid observations in the study period in 2011. $PM_{2.5}$ values less than

Table 1. Datasets used in the study with sources.

	Data/model	Sensor	Spatial resolution	Temporal resolution	Accuracy
1	Surface $PM_{2.5}$	TEOM	Point data	Daily	±5 %–10 %
2	Mid-visible aerosol optical depth (AOD)	MAIAC_MODIS	1 km	Daily	66 % compared to AERONET
3	Fire radiative power (FRP)	Terra/Aqua-MODIS	1 km	Daily	±7 %
4	ECMWF (meteorological variables)		0.25°	Hourly	

https://www.epa.gov/outdoor-air-quality-data, https://earthdata.nasa.gov/, https://earthdata.nasa.gov/, https://www.ecmwf.int/en/forecasts (last access: 26 July 2021).

$2\,\mu g\,m^{-3}$ are discarded since they are lower than the established lower detection limit (LDL) (EPA, 2011, 2018).

2.2 Satellite data

AOD, which represents the total column aerosol mass loading, is related to surface $PM_{2.5}$ as a function of aerosol properties (Koelemeijer et al., 2006):

$$\mathrm{AOD} = \mathrm{PM}_{2.5}\, H\, f(\mathrm{RH})\, \frac{3 Q_{\mathrm{ext,dry}}}{4\rho\, r_{\mathrm{eff}}} = \mathrm{PM}_{2.5}\, H\, S, \quad (1)$$

where H is the aerosol layer height, $f(\mathrm{RH})$ is the ratio of ambient and dry extinction coefficients, $Q_{\mathrm{ext,dry}}$ is the extinction efficiency under dry conditions, r_{eff} is the particle effective radius, ρ is the aerosol mass density and S is the specific extinction efficiency ($m^2\,g^{-1}$) of the aerosol under ambient conditions. Although AOD usually has a positive correlation with $PM_{2.5}$, this relationship depends on various meteorological parameters which will be discussed in Sect. 2.3.

The MODIS mid-visible AOD from the Multi-Angle Implementation of Atmospheric Correction (MAIAC) product (MCD19A2 version 6) is used in this study. We used the MAIAC-retrieved Terra and Aqua MODIS AOD product at 1 km spatial resolution (Lyapustin et al., 2018), where different orbits are averaged to obtain mean daily values. Since thick smoke plumes generated by wildfires can be misclassified as clouds, only AOD less than 0 will be discarded. Validation with AERONET studies shows that 66 % of the MAIAC AOD data agree within ±0.5 to ±0.1 AOD (Lyapustin et al., 2018). Largely due to cloud cover, grid cells may have a limited number of AOD observations within a certain period. On average, cloud-free AOD data are available about 40 % of the time during 9 to 25 August 2018 when fires were active in the region bounded by 25–50° N, 65–125° W. A smoke flag from the same product is used as a predictor in estimating surface $PM_{2.5}$. The smoke detection is performed using MODIS red, blue and deep blue bands, and smoke pixels are separated from dust and clouds based on absorption parameter, size parameter and thermal thresholds (see Lyapustin et al., 2012, 2018, for further discussion). The smoke flag data can provide the percentage of a smoke pixel in each grid, which is related to smoke coverage.

We also use the MODIS level-3 daily FRP (MCD14ML, fire radiative power) product which combines the Terra and Aqua fire products to assess wildfire activity. The fire radiative energy indicates the rate of combustion, and thus FRP can be used for characterizing active fires (Freeborn et al., 2014). For the purposes of the study we sum the FRP within every 2.3° × 3.5° box to represent the total fire activity in different locations.

2.3 Meteorological data

Meteorological information, including boundary layer height (BLH), 2 m temperature (T2M), 10 m wind speed (WS), surface relative humidity (RH) and surface pressure (SP), is obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5) product, with a spatial resolution of 0.25° and temporal resolution of 1 h, and is matched temporally with the satellite overpass time. The meteorological parameters provide important information on different processes affecting surface $PM_{2.5}$ concentration, which can also be seen as supplements of the AOD–$PM_{2.5}$ relationship as previously discussed.

The BLH can provide information on aerosol layer height (H in Eq. 1) as aerosols are often found to be well mixed within the boundary layer (Gupta and Christopher, 2009b). With the same amount of pollution within the boundary layer, the higher the BLH, the more $PM_{2.5}$ is distributed within that layer and vice versa (Miao et al., 2018; Zheng et al., 2017). Therefore, $PM_{2.5}$ usually has an anticorrelation with BLH. However, for wildfire events, the aerosol layer height is sometimes higher than the BLH (Haarig et al., 2018), which leads to lower correlation between AOD and $PM_{2.5}$ since we use only BLH to present the aerosol layer height. Thus BLH can provide aerosol vertical information in most cases except for suspended high-layer aerosol caused by fires, which leads to higher bias of the model for high-layer aerosols near the fire sources. Surface temperature (T2M) can affect $PM_{2.5}$ through convection, evaporation, temperature inversion and secondary pollutant generation processes (Chen et al., 2020). The first two processes are negatively related to $PM_{2.5}$ concentration: (1) higher temperature increases turbulence, which accelerates the dispersion of pollution and a decrease in $PM_{2.5}$; (2) higher temperature increases evaporation loss of $PM_{2.5}$, including ammonium nitrate and other volatile or semi-volatile components (Wang et al., 2017). The

latter two processes are positively related to $PM_{2.5}$ by limiting vertical motion and promoting photochemical reactions under high temperature (Xu et al., 2019; Zhang et al., 2015). High wind speeds are often negatively related to $PM_{2.5}$ since it increases the dispersion of pollutants. However, unique geographical conditions such as mountains with certain wind directions can cause accumulation of pollutants (Chen et al., 2017). RH may promote hygroscopic growth of particles to increase $PM_{2.5}$ (Trueblood et al., 2018; Zheng et al., 2017). Changes in surface pressure (SP) may influence the diffusion or accumulation of pollutants through formation of low-level wind convergence (You et al., 2017). Precipitation is another factor that largely influences surface $PM_{2.5}$ since it can accelerate the wash-out of suspended particles, but AOD values are not available when clouds are present.

3 Methodology

To assess the impact of NWUSC fires on $PM_{2.5}$ in the United States, we first estimate $PM_{2.5}$ over the study region during a time period with high fire activity (2018). We then use the same method during a year with low fire activity (2011) to compare the differences between the two years. The two years are selected based on the total FRP in August calculated within Canada (49–60° N, 55–135° W) and the northwestern (NW) United States (35–49° N, 105–125° W). Table 2 shows the total FRP in Canada and the northwestern US in August from 2010 to 2018. The total FRP in the two regions is lowest in 2011 and highest in 2018 during the nine years, which provides the basis for the study. In order to create a 0.1° surface $PM_{2.5}$, the GWR model is used to estimate the relationships of $PM_{2.5}$ and AOD using various parameters. Detailed processing steps for the GWR model are shown in Fig. 1.

3.1 Data preprocessing

The first step is to resample all datasets to a uniform spatial resolution by creating a 0.1° resolution grid covering the continental United States. During this process, we collocate the $PM_{2.5}$ data and average the values if there is more than one value in one grid. Then the MAIAC AOD and smoke flag are averaged into 0.1° grid cells. Meteorological datasets are also resampled to the 0.1° grid cells by applying the inverse distance method.

3.2 Time selecting and averaging

Next we select data where AOD and ground $PM_{2.5}$ are both available ($AOD > 0$ and $PM_{2.5} > 2.0\,\mu g\,m^{-3}$) and average them for the study period (since the LDL for the FRM method is $2\,\mu g\,m^{-3}$ in 2011 and $3\,\mu g\,m^{-3}$ in 2018, we decided to use the LDL for 2011) (EPA, 2011, 2018). This is to ensure that the AOD, $PM_{2.5}$ and other variables match with each other, because $PM_{2.5}$ is not a continuous measurement for some sites and AOD has missing values due to cloud cover and other reasons. Therefore, it is important to use data from days where both measurements are available to avoid sampling biases.

3.3 GWR model development and validation

The adaptive bandwidth selected by the Akaike information criterion (AIC) is used for the GWR model (Loader, 1999). For locations that already have $PM_{2.5}$ monitors, we calculate the mean AOD of a $0.5 \times 0.5°$ box centered at the ground location and estimate the GWR coefficients (β) for AOD and meteorological variables to estimate $PM_{2.5}$. The model structure can be expressed as

$$\begin{aligned} PM_{2.5i} = {} & \beta_{0,i} + \beta_{1,i}AOD_i + \beta_{2,i}BLH_i + \beta_{3,i}T2M_i \\ & + \beta_{4,i}U10M_i + \beta_{5,i}RH_{sfci} + \beta_{6,i}SP_i + \beta_{7,i}SF_i + \varepsilon_i, \end{aligned}$$

where $PM_{2.5i}$ ($\mu g\,m^{-3}$) is the selected ground-level $PM_{2.5}$ concentration at location i; $\beta_{0,i}$ is the intercept at location i; $\beta_{1,i} \sim \beta_{8,i}$ are the location-specific coefficients; AOD_i is the resampled AOD selected from MAIAC daily AOD data at location i; BLH_i, $T2M_i$, $U10M_i$, RH_{sfci}, and SP_i are selected meteorological parameters (BLH, T2M, WS, RH, and PS) at location i; SF_i (%) is the resampled smoke flag data at location i, and ε_i is the error term at location i.

We perform leave-one-out cross-validation (LOOCV) to test the model predictive performance (Kearns and Ron, 1999). Since the GWR model relies on an adequate number of observations, the prediction accuracy will be lower if we preserve too many data for validation. Therefore, we choose the LOOCV method, which preserves only one datum for validation at a time and repeats the process until all the data are used. In addition, R^2 and root mean square error (RMSE) are calculated for both model fitting and model validation processes to detect overfitting, which leads to low predictability.

3.4 Model prediction

While predicting the ground-level $PM_{2.5}$ for unsampled locations, we make use of the estimated parameters for sites within a 5° radius to generate new slopes for independent variables based on the spatial weighting matrix (Brunsdon et al., 1996). The closer to the predicted location, the closer to 1 the weighting factor will be, while the weighting factor for sites farther than 5° in distance is zero. It is important to note that AOD and other independent variables used for prediction in this step are averaged values for days that have valid AOD, which is different from the data used in the fitting process since $PM_{2.5}$ is not measured every day in all locations.

4 Results and discussion

We first discuss the surface $PM_{2.5}$ for a few select locations that are impacted by fires followed by the spatial distribu-

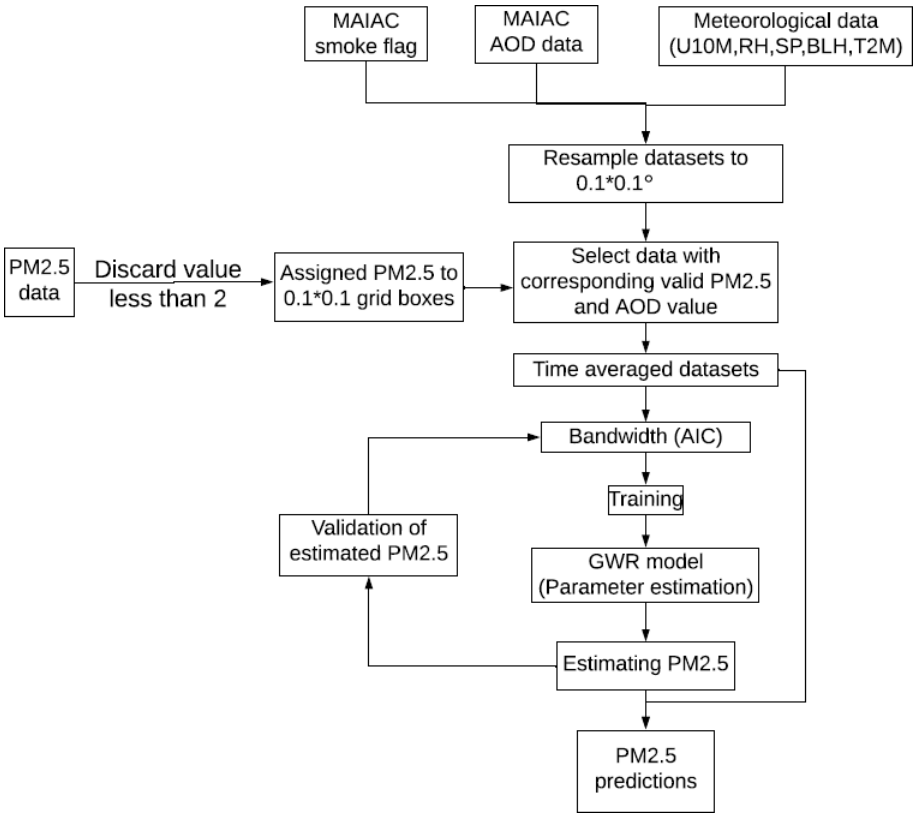

Figure 1. Flow chart for the geographically weighted regression model used. All satellite, ground, and meteorological data are gridded to 0.1 by 0.1°.

Table 2. Total FRP in Canada and the northwestern US in August of different years (unit: 10^4 MW).

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
CA	148.24	4.84	19.93	70.54	107.78	10.39	4.6	307.3	542.99
NW US	16.41	42.84	320.39	192.06	67.01	339.58	112.9	195.64	296.91

tion of MODIS AOD and the FRP for August 2018. We then assess the spatial distribution of surface $PM_{2.5}$ from the GWR method. The validation of the GWR method is then discussed. To further demonstrate the impact of the NWUSC fires on $PM_{2.5}$ air quality in the United States, we show the spatial distribution of the difference between August 2018 and August 2011 and quantify these results for 10 US EPA regions.

4.1 Descriptive statistics of satellite data and ground measurements

The 2018 summertime Canadian wildfires started around the end of July in British Columbia and continued until mid-September. The fires spread rapidly to the south of Canada during August, causing high concentrations of smoke aerosols to drift down to the US and affecting particulate matter air quality significantly. From late July to mid-September, wildfires in the northwestern US that burnt forest and grassland also affected air quality in the United States. Starting with the Cougar Creek Fire and then the Crescent Mountain and Gilbert fires, different wildfires in NWUSC caused severe air pollution in numerous US cities. Figure 2a shows the rapid increase in $PM_{2.5}$ in selected US cities from 1 July to 31 August, due to the transport of smoke from these wildfires. For all sites, July had low $PM_{2.5}$ concentrations ($< 10\,\mu g\,m^{-3}$) and rapidly increases as fire activity increases. Calculating only from the EPA ground observations, the mean $PM_{2.5}$ of the 17 d for the entire US is $13.7\,\mu g\,m^{-3}$ and the mean $PM_{2.5}$ for Washington (WA) is $40.6\,\mu g\,m^{-3}$, which indicates that the PM pollution is concentrated in the northwestern US for these days. This trend is obvious when comparing the mean $PM_{2.5}$ of all US stations (black line with no markers) and the mean $PM_{2.5}$ of all WA stations (grey line with no markers). Ground-level $PM_{2.5}$ reaches

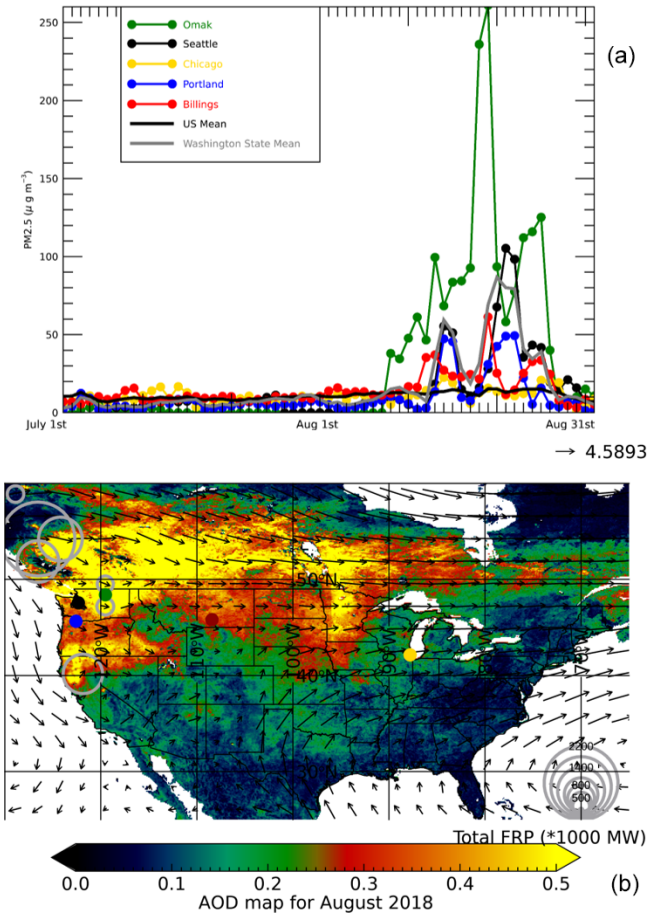

Figure 2. **(a)** Variations of EPA ground-observed $PM_{2.5}$ in different cities from July to August 2018 (Omak – Washington, Seattle – Washington, Chicago – Illinois, Portland – Oregon, Billings – Montana). Black line without markers shows the mean variation of all the US stations, and the grey line without markers shows the mean variation of stations in Washington State. **(b)** Mean MAIAC satellite AOD distribution from 9 to 25 August 2018. AOD values equal to or larger than 0.5 are shown as the same color (yellow). Also shown are circles with FRP. Black arrow shows the wind direction, and the length of it represents the wind speed. The round spots of different colors on the map show the locations of the five selected cities (green – Omak, black – Seattle, yellow – Chicago, blue – Portland, red – Billings).

peak values between 17 and 21 August, and daily $PM_{2.5}$ values during this time period far exceed the 17 d mean $PM_{2.5}$. For example, mean $PM_{2.5}$ in Washington on 20 August is 86.75 $\mu g\,m^{-3}$, which is more than 2 times the 17 d average of this region. On 19 August, Omak, which is located in the foothills of the Okanogan Highlands in WA, had $PM_{2.5}$ values that exceed 250 $\mu g\,m^{-3}$. According to a review of US wildfire-caused $PM_{2.5}$ exposures, 24 h mean $PM_{2.5}$ concentrations from wildfires ranged from 8.7 to 121 $\mu g\,m^{-3}$, with a 24 h maximum concentration of 1659 $\mu g\,m^{-3}$ (Navarro et al., 2018).

Table 3 shows relevant statistics for 15 states that have at least one daily record of non-attainment of EPA standards ($> 35\,\mu g\,m^{-3}$). From the frequency records of non-attainment in the 17 d period (last column), four states (Montana, Washington, California and Idaho) were consistently affected by the wildfires, and large portions of ground stations in these states were influenced by smoke aerosols. Most of the neighboring states also experienced significant air pollution (third column). Noticeable from these records is that the total number of ground stations in some of the highly affected states (such as Idaho) is not sufficient for capturing the smoke. Although there are a total of eight EPA stations in Idaho, only two of them have consistent observations during the fire event; the other two stations have no valid observations, and the remaining four stations have only two to six observations during the 17 d period. Limited availability of valid data along with unevenly distributed stations makes it hard to quantify smoke pollution in the northwestern US during the fire event period. Therefore, we utilize satellite data to enlarge the spatial coverage and estimate pollution at a finer spatial resolution.

The spatial distribution of AOD shown in Fig. 2b indicates that the smoke from Canada is concentrated mostly in northern US states such as Washington, Oregon, Idaho, Montana, North Dakota and Minnesota. The black arrow shows the mean 800 hPa-level mean wind for 17 d, and the length of the arrow represents the wind speed in $m\,s^{-1}$. Also shown in Fig. 2b are wind speeds close to the fire sources which are about 4–5 $m\,s^{-1}$, and according to the distances and wind directions, it can take approximately 28–36 h for the smoke to be transported southeastward to Washington. Then the smoke continues to move east to other northern states such as Montana and North Dakota. In addition, the grey circle represents the total FRP of every $2.3 \times 3.5°$ box. The reason for not choosing a smaller grid for the FRP is to not clutter Fig. 2b with information from small fires. The bigger the circle is, the stronger the fire is in that grid, and different sizes and its corresponding FRP values are shown in the lower-right corner. It is clear that the strongest fires in 2018 are located in Tweedsmuir Provincial Park of British Columbia in Canada (53.333° N, 126.417° W). The four separate lightning-caused wildfires burnt nearly 301 549 ha of the boreal forest. The total FRP of August 2018 in Canada is about 5362 ($\cdot$ 1000 MW), while the total FRP of August 2011 in Canada is 48 ($\cdot$ 1000 MW). The 2011 fire was relatively weak compared to the 2018 Tweedsmuir Complex fire, and we therefore use the 2011 air quality data as a baseline to quantify the 2018 fire influence on $PM_{2.5}$ in the United States.

4.2 Model fitting and validation

The main goal for using the GWR model is to help predict the spatial distribution of $PM_{2.5}$ for places with no ground monitors while leveraging the increased spatial resolution of

Table 3. Statistics of 15 states that violate EPA standards ($35\,\mu g\,m^{-3}$) during the 17 d wildfire period.

State	Number of sites violating standard	Number of sites in the state	Percentage of sites violating standard (%)	Number of days violate standard
Montana	14	15	93.34	16
Washington	18	20	90	16
Oregon	12	14	85.71	5
North Dakota	7	11	63.63	4
Idaho	5	8	62.5	8
Colorado	11	21	52.38	2
South Dakota	5	10	50	1
California	57	119	47.9	14
Utah	7	15	46.67	4
Nevada	4	13	30.77	1
Wyoming	7	24	29.2	2
Minnesota	4	26	15.4	2
Texas	3	37	8.1	1
Louisiana	1	14	7.1	1
Arizona	1	20	5	1

satellite AOD, and therefore it is important to ensure that the model is robust. Figure 3a and b show the results for 2018 for GWR model fitting for the entire US and the LOOCV models, respectively. The color of the scatter plots represents the probability density function (PDF) which calculates the relative likelihood that the observed ground-level $PM_{2.5}$ would equal the predicted value. The lighter the color is, the more points are present, with a higher correlation. The model fitting process estimates the slope for each variable, and therefore the model can be fitted close to the observed $PM_{2.5}$, and using this estimated relationship, we are able to assess surface $PM_{2.5}$ using other parameters at locations where $PM_{2.5}$ monitors are not available. The LOOCV process tests the model performance for predicting $PM_{2.5}$. If the results of LOOCV have a large bias from the model fitting, then the predictability of the model is low. Higher R^2 and RMSE differences indicate that the model is overfitting and therefore not suitable. The R^2 for the model fitting is 0.834, and the R^2 for the LOOCV is 0.797, while the RMSE for the GWR model fitting is $3.46\,\mu g\,m^{-3}$, and for LOOCV the RMSE is $3.84\,\mu g\,m^{-3}$. There are minor differences between fitting R^2 and validation R^2 (0.037) and between fitting RMSE and validation RMSE ($0.376\,\mu g\,m^{-3}$), suggesting that the model is not overfitting and has stable predictability, further indicating that the model can predict surface $PM_{2.5}$ reliably. In addition, we also performed a 20-fold cross-validation by splitting the dataset into 20 consecutive folds, and each fold is used for validation, while the 19 remaining folds form the training set. The 20-fold cross-validation has an R^2 of 0.745 and a RMSE of $4.3\,\mu g\,m^{-3}$. The increase/decrease in the cross-validated R^2 and RMSE indicates that sufficient data are used for fitting since a small decrease in the number of fitting data can reduce the model prediction accuracy. Overall, the prediction error of the model is between 3 and $5\,\mu g\,m^{-3}$, which is a reasonable error range for 17 d average prediction of $PM_{2.5}$. For data greater than the daily mean EPA standard ($35\,\mu g\,m^{-3}$), the model has a RMSE of $12.07\,\mu g\,m^{-3}$, which is a lot larger than the RMSE when using the entire model. Therefore, the model has a tendency to underestimate $PM_{2.5}$ exceedances by around $12.07\,\mu g\,m^{-3}$. The larger the $PM_{2.5}$ is, the more the model underestimates. To examine the model performance for high- and low-pollution areas, the results are divided into two parts (larger than $35\,\mu g\,m^{-3}$ and less than $35\,\mu g\,m^{-3}$). Areas with high pollution have an R^2 of 0.64 and areas with low pollution have an R^2 of 0.67; therefore, the model performance is relatively stable for both large and small $PM_{2.5}$ values. Also, the inclusion of low aerosol concentration areas does not influence the model performance for high values (seen in the Supplement in Figs. S1 and S2), which means that the high R^2 is not a reason for the large number of low values. The GWR model fitting and validation results for the 17 d in 2011 are shown in Fig. S3.

4.3 Predictors' influence during wildfires

Table 4 shows the GWR model mean coefficients for the whole US region and for different selected regions. The selected boxes are shown in Fig. 4c in different colors: box1 (red) located in the NW US includes major fire sources in the US; box2 (gold) located in Montana is influenced by both neighboring states and smoke from Canada; box3 (green) in Minnesota is located further from the fires and has a minor increase in $PM_{2.5}$ due to remote smoke; box4 (black) in the NE (northeastern) US is furthest from the fires and has no obvious pollution increase. The second column of the tables shows the conditions for sample selection, and the third

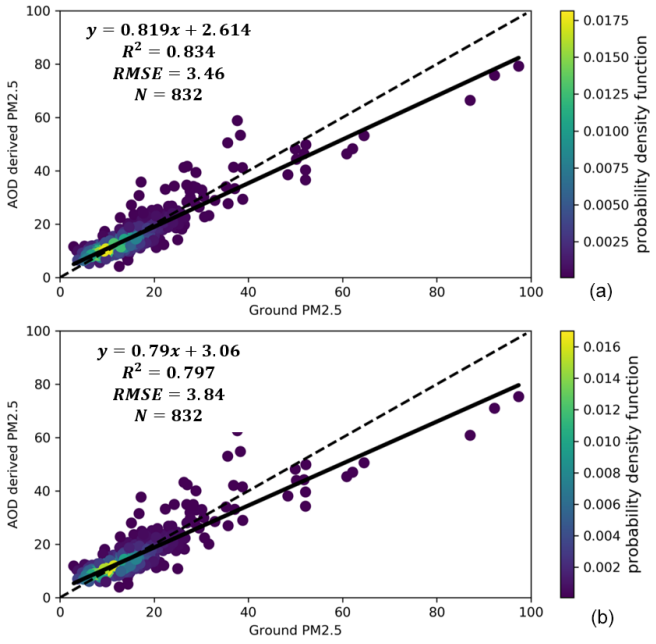

Figure 3. Results of model fitting and cross-validation for the GWR model for the entire US region averaged from 9 to 25 August 2018. **(a)** GWR model fitting results; **(b)** GWR model LOOCV results. The dashed line is the 1 : 1 line as a reference and the black line shows the regression line. The color of the scatter plots represents the probability density function which provides a relative likelihood that the value of the random variable would equal a certain sample.

column shows the number of pixels selected for each box. By comparing the coefficients of samples selected in these boxes, predictors have different influences in different locations. AOD has a stronger influence on predicting $PM_{2.5}$ closer to fire sources, but local emissions become more dominant if the distances are large enough. The smoke flag is overall positively related to surface $PM_{2.5}$, while it could slightly negatively relate to $PM_{2.5}$ around fire sources and the northeastern coasts. PBL is negatively related to $PM_{2.5}$ when the pollution is concentrated near the surface (fires or human-made emissions), while it appears to be positively related to $PM_{2.5}$ in locations where the main pollution source comes from remote wildfire smoke. Surface temperature has a relatively stable positive correlation with surface $PM_{2.5}$; however, surface pressure and wind speeds are negatively correlated with $PM_{2.5}$. Relative humidity, on the other hand, shows large variations in $PM_{2.5}$ influence across the nation. Around the wildfires where the RH is relatively low, RH has a positive correlation with $PM_{2.5}$ since hygroscopicity would increase and leads to accumulation of $PM_{2.5}$, but increasing RH can also decrease $PM_{2.5}$ concentration by overgrowing the $PM_{2.5}$ particles to deposition in a high-RH environment (Chen et al., 2018).

From Table 4, we know that the weighting for AOD is much larger than other predictors, but predictors other than AOD are important for the prediction. We tested our model with AOD as the only predictor to conduct a comparison to the original model, and the R^2 decreases from 0.83 to 0.79 and RMSE increases from 3.46 to 3.8. This is consistent with a previous study (Jiang et al., 2017) which shows improvements of R^2 from 0.69 to 0.78 and RMSE from 7.25 to 6.18 by adding four meteorological parameters in summer in eastern China. Other predictors have higher weighting in the fire source region (box1), where BLH cannot provide the aerosol vertical distribution information since smoke tends to be injected to higher levels. For high AOD regions where aerosol tends to be suspended at high levels, adding predictors other than AOD tends to have lower improvement of the model compared to low AOD values, because adding BLH can significantly improve the prediction for low-level aerosols. For regions with AOD less than 35, R^2 increases by 0.09 from the AOD-only model (0.6 to 0.69), while R^2 increases by 0.05 for areas with AOD larger than 35. RMSE decreases by 12 % and 7 % for AOD less and larger than 35 conditions, respectively. Overall, the meteorological factors have larger improvements for low-pollution areas (low-level aerosol in this case).

4.4 Predicted $PM_{2.5}$ distribution

The mean $PM_{2.5}$ distributions over the United States shown in Fig. 4a are calculated by averaging the surface $PM_{2.5}$ data from ground monitors for the 17 d, which matches well with the GWR model-predicted $PM_{2.5}$ distributions shown in Fig. 4b. The model estimation extends the ground measurements and provides pollution assessments across the entire nation. Comparing the AOD map (Fig. 2b) with the $PM_{2.5}$ estimations (Fig. 4b) demonstrates the differences between columnar and surface-level pollution. Differences between the AOD and $PM_{2.5}$ distributions are for various reasons, including (1) areas with high $PM_{2.5}$ concentrations in Fig. 4b corresponding to low AOD values in Fig. 2b (southern California, Utah, and the southern US) and (2) and high AOD regions in Fig. 2b corresponding to low $PM_{2.5}$ concentrations in Fig. 4b (Minnesota). The first situation usually occurs at the edge of polluted areas that are relatively far from the fire source, which is consistent with previous studies that reported smaller particles ($< 10\,\mu g$) being able to travel longer distances compared to large particles ($> 10\,\mu g$) (Gillies et al., 1996) and larger particles tending to settle closer to their source (Sapkota et al., 2005; Zhu et al., 2002).

We use the same method for 9 to 25 August in 2011 that had low fire activity, ensuring consistency in estimating coefficients for different variables for 2011. Figure 4c shows the difference in spatial distribution of mean ground $PM_{2.5}$ of the 17 d between 2018 and 2011. Larger differences in $PM_{2.5}$ are in the northwestern and central parts of the United States, with the southern states having very little impact due to the fires. Of all 48 states within the study region, there are 29 states that have a higher $PM_{2.5}$ value in 2018 than

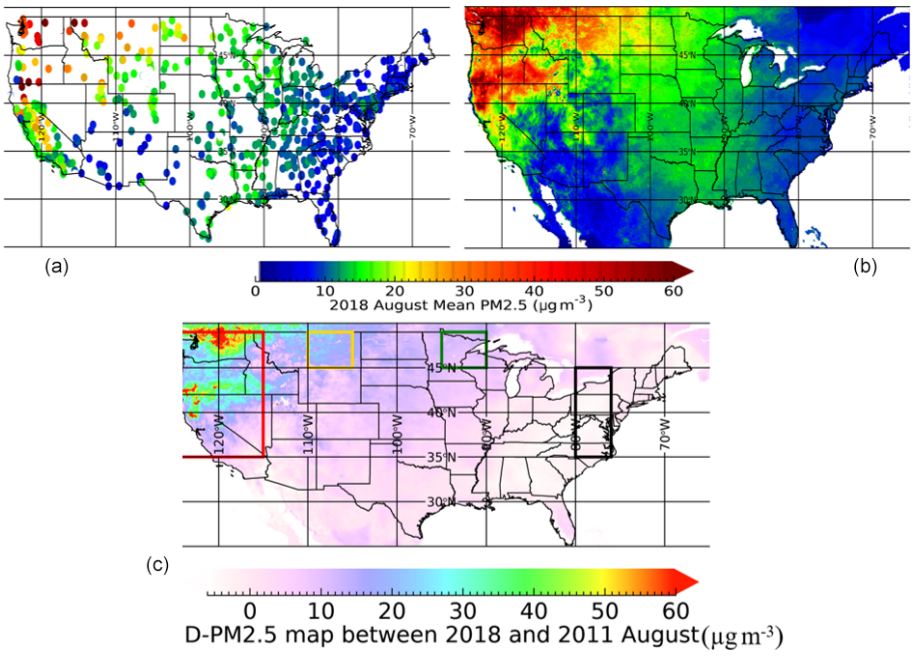

Figure 4. **(a)** EPA ground-observed $PM_{2.5}$ distribution over the US averaged from 9 to 25 August 2018. **(b)** GWR-predicted 17 d mean $PM_{2.5}$ distribution. **(c)** Difference map of predicted ground $PM_{2.5}$ of the 17 d mean values between 2018 and 2011. $PM_{2.5}$ values equal to or larger than 60 $\mu g\,m^{-3}$ are shown as the same color (red). Note that the D-$PM_{2.5}$ has a different color scale to make the negative values more apparent (blue).

Table 4. Coefficients of different predictors.

Mean coefficients	Sample selection	*N*	AOD	Smoke flag	PBL	T2M	RH	*U*	SP
box1 (red)	FRP > 1000	213	91.94	−0.14	−2.25	0.33	0.08	−2	−0.06
box2 (gold)	$PM_{2.5}$ > 30	362	60.1	0.013	−2.9	0.23	−0.08	−1.6	−0.03
box3 (green)	$PM_{2.5}$ > 17	278	6.2	0.05	0.2	0.2	0.014	−0.3	−0.02
box4 (black)	17 > $PM_{2.5}$ > 10	938	7.1	−0.02	−1.2	0.22	−0.035	0.06	−0.005
Whole US region	∼	106 352	28.1	0.024	−0.9	0.06	−0.04	−0.7	−0.002

2011, and 15 states have a 2018 $PM_{2.5}$ value of more than 2 times their 2011 value (shown in Fig. 5). The mean $PM_{2.5}$ for WA increases from 5.87 in 2011 to 46.47 $\mu g\,m^{-3}$ in 2018, which is about 8 times more than 2011 values. The $PM_{2.5}$ values in Oregon increase from 4.97 in 2011 to 33.3 $\mu g\,m^{-3}$ in 2018, which is nearly a 7-fold increase. For states from Montana to Minnesota, the mean $PM_{2.5}$ decreases from east to west, which reveals the path of smoke transport. As shown in Fig. 4c, there is a clear transport path of smoke from North Dakota all the way to Texas. Along the path, smoke increases $PM_{2.5}$ concentrations by 168 % in North Dakota and 27 % in Texas. Smoke aerosols transported over long distances typically contain fine-fraction PM, which significantly affects the health of children, adults, and vulnerable groups.

Figure 6 shows the mean $PM_{2.5}$ predicted from the GWR model of different EPA regions for the 17 d in 2011 and 2018 (Hawaii and Alaska are not included). The most influenced region is region 10, which has a 2018 mean $PM_{2.5}$ value of 34.2 $\mu g\,m^{-3}$ that is 6 times larger than the values in 2011 (5.8 $\mu g\,m^{-3}$). The $PM_{2.5}$ of regions 8 and 9 has 2.4 and 2.6 times increases in 2018 compared to 2011. Regions 1–4 have lower $PM_{2.5}$ in 2018 than 2011, possibly due to Clean Air Act initiatives, absence of any major fire activities and being further away for transported aerosols. The emission reduction improves the US air quality and lowers the $PM_{2.5}$ every year, but 6 out of the 10 EPA regions show significant increases in $PM_{2.5}$ during the study period, which indicates that the long-range transported wildfire smoke has become the new major pollutant in the US.

4.5 Estimation of Canadian fire pollution

To evaluate the pollution caused only by Canadian fires, we did a rough assessment according to the total FRP and $PM_{2.5}$ values. There are three states in the US that have wildfires during the study period, California, Washington and Oregon, and they have total FRPs of 1186, 518 and

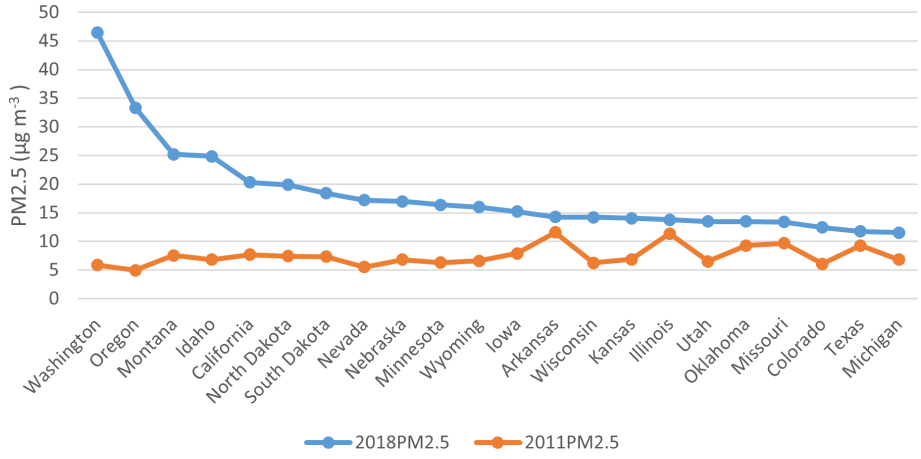

Figure 5. Mean $PM_{2.5}$ from 9 to 25 August 2018 and 2011 of most affected states.

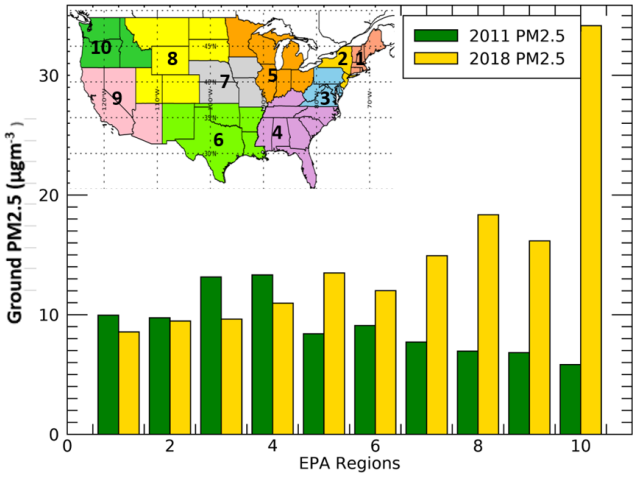

Figure 6. Mean $PM_{2.5}$ of EPA regions from 9 to 25 August 2011 and 2018. Inset shows the map of 10 EPA regions in different colors. Yellow column represents the 2018 mean $PM_{2.5}$, and green column represents for 2011 mean $PM_{2.5}$.

439 (· 1000 MW), respectively. Assuming that California was only influenced by the local fires, then fires of 1186 (· 1000 MW) cause a 13 μg m^{-3} increase in $PM_{2.5}$. Accordingly, wildfires in Washington State and Oregon State will cause 6 and 5 μg m^{-3} increases in state mean $PM_{2.5}$. Therefore, Canadian fires caused $PM_{2.5}$ increases in Washington and Oregon of about 35 and 23 μg m^{-3}, respectively. Since the FRP of Canadian wildfires is approximately 5 times larger than that of the California fires, which are the strongest fires in the US, we assume that the pollution affecting the states located in the downwind directions other than the three states is mainly coming from Canadian wildfires. States with no local fires such as Montana, North Dakota, South Dakota and Minnesota have $PM_{2.5}$ increases of 18.31, 12.8, 10.4 and 10.13 μg m^{-3}. The decrease in these numbers reveals that the smoke is transported in a southeasterly direction. This influence of Canadian wildfires on US air quality is only a rough quantity estimation, and thus additional work is needed to understand long-range transport smoke pollution and its impact on public health.

4.6 Comparison to previous studies

Compared to the Bayesian ensemble model developed by Geng et al. (2018) using MAIAC AOD and CMAQ (Community Multiscale Air Quality) model and ground $PM_{2.5}$ measurements, our GWR model has larger R^2, but with the CTM, their method can provide more vertical distribution information, which is important for wildfire smoke. GWR usually has better accuracy than the CTM since there are large uncertainties related to different CTM inputs such as emission and meteorological and land cover data, but for regions with fewer or no ground measurements, the CTM provides a good approach for estimating surface $PM_{2.5}$. Other studies which used machine-learning methods to predict surface $PM_{2.5}$ have better performance for long-term prediction rather than monthly estimations (Liang et al., 2020; Xiao et al., 2018) but can better resolve complex relationships between different predictors than statistical models (Geng et al., 2020). For wildfire events, the available data are much fewer than the long-term aerosol analysis, so the performance of a machine-learning method could be less accurate compared to long-term prediction. Our study also shows slightly larger R^2 compared to other GWR studies (Hu et al., 2013; Ma et al., 2014; You et al., 2016b) due to the inclusion of more meteorological and other related predictors.

4.7 Model uncertainties and limitations

There are various sources of uncertainties and limitations for studies that use satellite data to estimate surface $PM_{2.5}$ concentrations. Since wildfires develop quickly, it is important to have continuous observations to capture the rapid changes. This study uses polar-orbiting high-quality satellite aerosol

products, but the temporal evolution can only be estimated by geostationary datasets. Although satellite observations have excellent spatial coverage, missing data due to cloud cover are a limitation. As discussed in the paper, the prediction error (RMSE) of the model is between 3 and 5 $\mu g\,m^{-3}$, while the RMSE increased for locations with high aerosol concentration. This is partly due to lack of accurate vertical distribution information, which is very important for wildfire smoke. The GWR model is largely influenced by the distribution of ground stations, and the prediction error will be different in different locations due to unevenly distributed $PM_{2.5}$ stations. For locations that have a dense ground-monitoring distribution, the prediction error will be low, while the prediction error will be relatively larger in other places with sparse surface stations. Although there are obvious limitations, complementing surface data with satellite products and meteorological and other ancillary information in a statistical model like the GWR has provided robust results for estimating surface $PM_{2.5}$ from wildfires. We also note that we did not consider some variables used in other studies, such as NDVI, forest cover, vegetation type, industrial density, visibility and chemical constituents of smoke particles (Van Donkelaar et al., 2015; Hu et al., 2013; You et al., 2015; Zou et al., 2016). Visibility mentioned in some studies may improve the model performance, but unlike AOD, it has limited measurement across the nation, which will restrict the applicability of training data. Another uncertainty comes from the 2011 wildfires, which we assumed to be zero fire events, but there are actually few fire events in EPA regions 6, 8, 9 and 10, and this will lead to underestimation of $PM_{2.5}$ increase due to the 2018 fires in these regions.

One limitation of this study is that analysis based on 17 d mean values cannot capture daily pollution variations, which is also very important for pollution estimation during rapidly changing wildfire events. To extend this analysis to daily estimation, the cloud contaminations of satellite observations become a major problem. Therefore, future work is needed using chemistry transport models and other data to fill in the gaps on missing AOD data due to cloud coverage.

5 Summary and conclusions

We estimate the surface mean $PM_{2.5}$ for 17 d in August for a high fire activity year (2018) and compare that to a low fire activity year using the geographically weighted regression (GWR) method to assess the increase in $PM_{2.5}$ in the United States due to smoke transported from fires. The difference in $PM_{2.5}$ between the two years indicates that more than half of the United States (29 states) is influenced by the NWUSC wildfires, and half of the affected states have 17 d mean $PM_{2.5}$ increases larger than 100 % of the baseline value. The peak $PM_{2.5}$ during the wildfires can be much larger than the 17 d average and can affect vulnerable populations susceptible to air pollution. Some of the most affected states are Washington, California, Wisconsin, Colorado and Oregon, all of which have populations greater than 4 million. According to the Centers for Disease Control and Prevention (CDC), 8 % of the population has asthma (CDC, 2011). Therefore, for asthma alone, there are about 3 million people facing significant health issues due to the long-range transport of smoke in these states.

For states that show a decrease in $PM_{2.5}$ due to the Clean Air Act, the mean decrease is about 16 % of the baseline after 7 years. This is consistent with the EPA's report that there is a 23 % decrease in $PM_{2.5}$ on national average from 2010 to 2019 (U.S. Environmental Protection Agency, 2019). Compared to the dramatic increase (132 %) caused by wildfires, pollution from the fires counteracts our effort on emission controls. Although wildfires are often episodic and short term, high frequency of fire occurrence and increasingly longer durations of summertime wildfires in recent years have made them now a long-term influence on public lives. Our results show a significant increase in pollution in a short time period in most of the US states due to the NWUSC wildfires, which affects millions of people. With wildfires becoming more frequent during recent years, more effort is needed to predict and warn the public about the long-range transported smoke from wildfires.

Code availability. The official Python code for the GWR model is available at https://github.com/pysal/mgwr (last access: 26 July 2021).

Data availability. The MAIAC AOD and FRP data are accessible at https://earthdata.nasa.gov/ (last access: 26 July 2021); the ECMWF meteorological datasets are accessible at https://www.ecmwf.int/en/forecasts (last access: 26 July 2021); the surface PM2.5 observations are accessible at https://www.epa.gov/outdoor-air-quality-data (last access: 26 July 2021).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/acp-21-11243-2021-supplement.

Author contributions. The authors confirm contribution to the paper as follows: study conception and design were by SC, ZX, and PG; data collection was by ZX and PG; analysis and interpretation of results were by ZX, SC, and PG; draft manuscript preparation was done by ZX, SC, and PG. All the authors reviewed the results and approved the final version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Pawan Gupta and Sundar Christopher were supported by a NASA grant. MODIS data were acquired from the Goddard DAAC. Sincerest thanks to the MAIAC, MODIS, EPA and ECMWF teams for their datasets that made this research possible.

Financial support. This research has been supported by NASA (grant no. NNM11AA01A).

Review statement. This paper was edited by Zhanqing Li and reviewed by four anonymous referees.

References

Apte, J. S., Brauer, M., Cohen, A. J., Ezzati, M., and Pope, C. A.: Ambient $PM_{2.5}$ Reduces Global and Regional Life Expectancy, Environ. Sci. Technol. Lett., 5, 546–551, https://doi.org/10.1021/acs.estlett.8b00360, 2018.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E.: Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity, Geogr. Anal., 28, 281–298, https://doi.org/10.1111/j.1538-4632.1996.tb00936.x, 1996.

Calkin, D. E., Thompson, M. P., and Finney, M. A.: Negative consequences of positive feedbacks in us wildfire management, For. Ecosyst., 2, 1–10, https://doi.org/10.1186/s40663-015-0033-8, 2015.

Cascio, W. E.: Wildland Fire Smoke and Human Health, Sci. Total Environ., 624, 586–595, https://doi.org/10.1016/j.scitotenv.2017.12.086, 2018.

CDC: Asthma in the US, CDC Vital Signs, May 2011, Center for Disease Control and Prevention, available at: https://www.cdc.gov/vitalsigns/asthma/index.html (last access: 26 July 2021), 1–4, 2011.

Chen, D., Xie, X., Zhou, Y., Lang, J., Xu, T., Yang, N., Zhao, Y., and Liu, X.: Performance evaluation of the WRF-chem model with different physical parameterization schemes during an extremely high $PM_{2.5}$ pollution episode in Beijing, Aerosol Air Qual. Res., 17, 262–277, https://doi.org/10.4209/aaqr.2015.10.0610, 2017.

Chen, Z., Xie, X., Cai, J., Chen, D., Gao, B., He, B., Cheng, N., and Xu, B.: Understanding meteorological influences on $PM_{2.5}$ concentrations across China: a temporal and spatial perspective, Atmos. Chem. Phys., 18, 5343–5358, https://doi.org/10.5194/acp-18-5343-2018, 2018.

Chen, Z., Chen, D., Zhao, C., Kwan, M. po, Cai, J., Zhuang, Y., Zhao, B., Wang, X., Chen, B., Yang, J., Li, R., He, B., Gao, B., Wang, K., and Xu, B.: Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: A review of methodology and mechanism, Environ. Int., 139, 105558, https://doi.org/10.1016/j.envint.2020.105558, 2020.

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., and Xiang, H.: A review on predicting ground $PM_{2.5}$ concentration using satellite aerosol optical depth, Atmosphere (Basel), 7, 129, https://doi.org/10.3390/atmos7100129, 2016.

Coogan, S. C. P., Robinne, F. N., Jain, P., and Flannigan, M. D.: Scientists' warning on wildfire – a canadian perspective, Can. J. Forest Res., 49, 1015–1023, https://doi.org/10.1139/cjfr-2019-0094, 2019.

Dreessen, J., Sullivan, J., and Delgado, R.: Observations and impacts of transported Canadian wildfire smoke on ozone and aerosol air quality in the Maryland region on June 9–12, 2015, J. Air Waste Manage. Assoc., 66, 842–862, https://doi.org/10.1080/10962247.2016.1161674, 2016.

EPA: Code of Federal Regulations Title 40: Protection of Environment, 694, available at: https://www.govinfo.gov/app/collection/cfr/2011/ (last access: 26 July 2021) 2011.

EPA: Code of Federal Regulations Title 40: Protection of Environment, 694, available at: https://www.govinfo.gov/app/collection/cfr/2018/ (last access: 26 July 2021) 2018.

Fotheringham, A. S., Charlton, M. E., and Brunsdon, C.: Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, Environ. Plan. A, 30, 1905–1927, 1998.

Fotheringham, S. A., Brunsdon, C., and Charlton, M.: Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, 2003.

Freeborn, P. H., Wooster, M. J., Roy, D. P., and Cochrane, M. A.: Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite-based active fire characterization and biomass burning estimation, Geophys. Res. Lett., 41, 1988–1994, https://doi.org/10.1002/2013GL059086, 2014.

Geng, G., Murray, N. L., Tong, D., Meng, X., Chang, H. H., Liu, Y., Hu, X., and Lee, P.: Satellite-Based Daily $PM_{2.5}$ Estimates During Fire Seasons in Colorado, 123, 8159–8171, https://doi.org/10.1029/2018JD028573, 2018.

Geng, G., Meng, X., He, K., and Liu, Y.: Random forest models for $PM_{2.5}$ speciation concentrations using MISR fractional AODs Random forest models for $PM_{2.5}$ speciation concentrations using MISR fractional AODs, Environ. Res. Lett., 15, 034056, https://doi.org/10.1088/1748-9326/ab76df, 2020.

Goldberg, D. L., Gupta, P., Wang, K., Jena, C., Zhang, Y., Lu, Z., and Streets, D. G.: Using gap-filled MAIAC AOD and WRF-Chem to estimate daily $PM_{2.5}$ concentrations at 1 km resolution in the Eastern United States, Atmos. Environ., 199, 443–452, https://doi.org/10.1016/j.atmosenv.2018.11.049, 2019.

Gu, Y.: Estimating $PM_{2.5}$ Concentrations Using 3 km MODIS AOD Products: A Case Study in British Columbia, Canada, University of Waterloo, 2019.

Guo, B., Wang, X., Pei, L., Su, Y., Zhang, D., and Wang, Y.: Identifying the spatiotemporal dynamic of $PM_{2.5}$ concentrations at multiple scales using geographically and temporally weighted regression model across China during 2015–2018, Sci. Total Environ., 751, 141765, https://doi.org/10.1016/j.scitotenv.2020.141765, 2021.

Gupta, P. and Christopher, S. A.: Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. A neural network approach, J. Geophys. Res.-Atmos., 114, 1–14, https://doi.org/10.1029/2008JD011497, 2009a.

Gupta, P. and Christopher, S. A.: Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach, J. Geophys. Res.-Atmos., 114, 1–13, https://doi.org/10.1029/2008JD011496, 2009b.

Haarig, M., Ansmann, A., Baars, H., Jimenez, C., Veselovskii, I., Engelmann, R., and Althausen, D.: Depolarization and lidar ratios at 355, 532, and 1064 nm and microphysical properties of aged tropospheric and stratospheric Canadian wildfire smoke, Atmos. Chem. Phys., 18, 11847–11861, https://doi.org/10.5194/acp-18-11847-2018, 2018.

Hessburg, P. F., Churchill, D. J., Larson, A. J., Haugo, R. D., Miller, C., Spies, T. A., North, M. P., Povak, N. A., Belote, R. T., Singleton, P. H., Gaines, W. L., Keane, R. E., Aplet, G. H., Stephens, S. L., Morgan, P., Bisson, P. A., Rieman, B. E., Salter, R. B., and Reeves, G. H.: Restoring fire-prone Inland Pacific landscapes: seven core principles, Landscape Ecol., 30, 1805–1835, https://doi.org/10.1007/s10980-015-0218-0, 2015.

Hoff, R. M. and Christopher, S. A.: Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land?, J. Air Waste Manage., 59, 645–675, https://doi.org/10.3155/1047-3289.59.6.645, 2009.

Hu, X., Waller, L. A., Al-Hamdan, M. Z., Crosson, W. L., Estes, M. G., Estes, S. M., Quattrochi, D. A., Sarnat, J. A., and Liu, Y.: Estimating ground-level $PM_{2.5}$ concentrations in the southeastern U.S. using geographically weighted regression, Environ. Res., 121, 1–10, https://doi.org/10.1016/j.envres.2012.11.003, 2013.

Hu, X., Belle, J. H., Meng, X., Wildani, A., Waller, L. A., Strickland, M. J., and Liu, Y.: Estimating $PM_{2.5}$ Concentrations in the Conterminous United States Using the Random Forest Approach, Environ. Sci. Technol., 51, 6936–6944, https://doi.org/10.1021/acs.est.7b01210, 2017.

Hu, Z.: Spatial analysis of MODIS aerosol optical depth, $PM_{2.5}$, and chronic coronary heart disease, Int. J. Health Geogr., 8, 1–10, https://doi.org/10.1186/1476-072X-8-27, 2009.

Hubbell, B. J., Crume, R. V., Evarts, D. M., and Cohen, J. M.: Policy Monitor: Regulation and progress under the 1990 Clean Air Act Amendments, Rev. Environ. Econ. Pol., 4, 122–138, https://doi.org/10.1093/reep/rep019, 2010.

Hystad, P., Demers, P. A., Johnson, K. C., Brook, J., Van Donkelaar, A., Lamsal, L., Martin, R., and Brauer, M.: Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study, Environ. Heal. A Glob. Access Sci. Source, 11, 1–22, https://doi.org/10.1186/1476-069X-11-22, 2012.

Gillies, J. A., Nickling, W. G., and Mctainsh, G. H.: Dust concentration s and particle-size characteristics of an intense dust haze event: inland delta region, Atmos. Environ., 30, 1081–1090, https://doi.org/10.1016/1352-2310(95)00432-7, 1996.

Jiang, M., Sun, W., Yang, G., and Zhang, D.: Modelling seasonal GWR of daily $PM_{2.5}$ with proper auxiliary variables for the Yangtze River Delta, Remote Sens., 9, 1–20, https://doi.org/10.3390/rs9040346, 2017.

Kearns, M. and Ron, D.: Algorithmic stability and sanity-check bounds for leave-one-out cross-validation, Neural Comput., 11, 1427–1453, https://doi.org/10.1162/089976699300016304, 1999.

Koelemeijer, R. B. A., Homan, C. D., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmos. Environ., 40, 5304–5315, https://doi.org/10.1016/j.atmosenv.2006.04.044, 2006.

Kollanus, V., Tiittanen, P., Niemi, J. V., and Lanki, T.: Effects of long-range transported air pollution from vegetation fires on daily mortality and hospital admissions in the Helsinki metropolitan area, Finland, Environ. Res., 151, 351–358, https://doi.org/10.1016/j.envres.2016.08.003, 2016.

Li, T., Shen, H., Yuan, Q., Zhang, X., and Zhang, L.: Estimating Ground-Level $PM_{2.5}$ by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach, Geophys. Res. Lett., 44, 11985–11993, https://doi.org/10.1002/2017GL075710, 2017.

Liang, F., Xiao, Q., Huang, K., Yang, X., Liu, F., Li, J., and Lu, X.: The 17-y spatiotemporal trend of $PM_{2.5}$ and its mortality burden in China, P. Natl. Acad. Sci. USA, 117, 25601–25608, https://doi.org/10.1073/pnas.1919641117, 2020.

Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., and Koutrakis, P.: Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing, Environ. Sci. Technol., 39, 3269–3278, https://doi.org/10.1021/es049352m, 2005.

Loader, C. R.: Bandwith selection:Classical or plug in?, Ann. Stat., 27, 415–438, 1999.

Lyapustin, A., Korkin, S., Wang, Y., Quayle, B., and Laszlo, I.: Discrimination of biomass burning smoke and clouds in MAIAC algorithm, Atmos. Chem. Phys., 12, 9679–9686, https://doi.org/10.5194/acp-12-9679-2012, 2012.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC algorithm, Atmos. Meas. Tech., 11, 5741–5765, https://doi.org/10.5194/amt-11-5741-2018, 2018.

Ma, Z., Hu, X., Huang, L., Bi, J., and Liu, Y.: Estimating ground-level $PM_{2.5}$ in china using satellite remote sensing, Environ. Sci. Technol., 48, 7436–7444, https://doi.org/10.1021/es5009399, 2014.

Meixner, T. and Wohlgemuth, P.: Wildfire Impacts on Water Quality, J. Wildl. Fire, 13, 27–35, 2004.

Melillo, J. M., Richmond, T., and Yohe, G. W.: Climate Change Impacts in the United States: The third national climate assessment, U.S. Global Change Research Program, https://doi.org/10.7930/J0Z31WJ2, 2014.

Miao, Y., Liu, S., Guo, J., Huang, S., Yan, Y., and Lou, M.: Unraveling the relationships between boundary layer height and $PM_{2.5}$ pollution in China based on four-year radiosonde measurements, Environ. Pollut., 243, 1186–1195, https://doi.org/10.1016/j.envpol.2018.09.070, 2018.

Miller, D. J., Sun, K., Zondlo, M. A., Kanter, D., Dubovik, O., Welton, E. J., Winker, D. M., and Ginoux, P.: Assessing boreal forest fire smoke aerosol impacts on U.S. air quality: A case study using multiple data sets, J. Geophys. Res.-Atmos., 116, D22209, https://doi.org/10.1029/2011JD016170, 2011.

Mirzaei, M., Bertazzon, S., and Couloigner, I.: Modeling Wildfire Smoke Pollution by Integrating Land Use Regression and Remote Sensing Data: Regional Multi-Temporal Estimates for Public Health and Exposure Models, Atmosphere (Basel), 9, 335, https://doi.org/10.3390/atmos9090335, 2018.

Munoz-Alpizar, R., Pavlovic, R., Moran, M. D., Chen, J., Gravel, S., Henderson, S. B., Sylvain, M., Racine, J., Duhamel, A., Gilbert, S., Beaulieu, P., Landry, H., Davignon, D., Cousineau, S., and Bouchet, V.: Multi-Year (2013–2016) $PM_{2.5}$ Wildfire Pollution Exposure over North America as Determined from

Operational Air Quality Forecasts, Atmosphere (Basel), 8, 179, https://doi.org/10.3390/atmos8090179, 2017.

Navarro, K. M., Schweizer, D., Balmes, J. R., and Cisneros, R.: A review of community smoke exposure from wildfire compared to prescribed fire in the United States, Atmosphere (Basel), 9, 1–11, https://doi.org/10.3390/atmos9050185, 2018.

Samet, J. M.: The clean air act and health – A clearer view from 2011, N. Engl. J. Med., 365, 198–201, https://doi.org/10.1056/NEJMp1103332, 2011.

Sapkota, A., Symons, J. M., Kleissl, J., Wang, L., Parlange, M. B., Ondov, J., Breysse, P. N., Diette, G. B., Eggleston, P. A., and Buckley, T. J.: Impact of the 2002 Canadian forest fires on particulate matter air quality in Baltimore City, Environ. Sci. Technol., 39, 24–32, https://doi.org/10.1021/es035311z, 2005.

Stephens, S. L.: Forest fire causes and extent on United States Forest Service lands, Int. J. Wildl. Fire, 14, 213–222, https://doi.org/10.1071/WF04006, 2005.

Trueblood, M. B., Lobo, P., Hagen, D. E., Achterberg, S. C., Liu, W., and Whitefield, P. D.: Application of a hygroscopicity tandem differential mobility analyzer for characterizing PM emissions in exhaust plumes from an aircraft engine burning conventional and alternative fuels, Atmos. Chem. Phys., 18, 17029–17045, https://doi.org/10.5194/acp-18-17029-2018, 2018.

U.S. Environmental Protection Agency: Particulate Matter ($PM_{2.5}$) Trends, available at: https://www.epa.gov/air-trends/particulate-matter-pm25-trends (last access: 26 July 2021), 2019.

Van Donkelaar, A., Martin, R. V., and Park, R. J.: Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing, J. Geophys. Res.-Atmos., 111, D21201, https://doi.org/10.1029/2005JD006996, 2006.

Van Donkelaar, A., Martin, R. V, Spurr, R. J. D., and Burnett, R. T.: High-Resolution Satellite-Derived $PM_{2.5}$ from Optimal Estimation and Geographically Weighted Regression over North America, Environ. Sci. Technol., 49, 10482–10491, https://doi.org/10.1021/acs.est.5b02076, 2015.

Van Donkelaar, A., Martin, R. V., Li, C., and Burnett, R. T.: Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors, Environ. Sci. Technol., 53, 2595–2611, https://doi.org/10.1021/acs.est.8b06392, 2019.

Wang, H., Shi, G., Tian, M., Zhang, L., Chen, Y., Yang, F., and Cao, X.: Aerosol optical properties and chemical composition apportionment in Sichuan Basin, China, Sci. Total Environ., 577, 245–257, https://doi.org/10.1016/j.scitotenv.2016.10.173, 2017.

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., and Cribb, M.: Estimating 1-km-resolution $PM_{2.5}$ concentrations across China using the space-time random forest approach, Remote Sens. Environ., 231, 111221, https://doi.org/10.1016/j.rse.2019.111221, 2019.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., and Song, Y.: Improved 1 km resolution $PM_{2.5}$ estimates across China using enhanced space–time extremely randomized trees, Atmos. Chem. Phys., 20, 3273–3289, https://doi.org/10.5194/acp-20-3273-2020, 2020.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1-km-resolution high-quality $PM_{2.5}$ data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sens. Environ., 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021.

Xiao, Q., Chang, H. H., Geng, G., and Liu, Y.: An Ensemble Machine-Learning Model To Predict Historical $PM_{2.5}$ Concentrations in China from Satellite Data, Environ. Sci. Technol., 52, 13260–13269,https://doi.org/10.1021/acs.est.8b02917, 2018.

Xu, T., Song, Y., Liu, M., Cai, X., Zhang, H., Guo, J., and Zhu, T.: Temperature inversions in severe polluted days derived from radiosonde data in North China from 2011 to 2016, Sci. Total Environ., 647, 1011–1020, https://doi.org/10.1016/j.scitotenv.2018.08.088, 2019.

You, W., Zang, Z., Pan, X., Zhang, L., and Chen, D.: Estimating $PM_{2.5}$ in Xi'an, China using aerosol optical depth: A comparison between the MODIS and MISR retrieval models, Sci. Total Environ., 505, 1156–1165, https://doi.org/10.1016/j.scitotenv.2014.11.024, 2015.

You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., and Wang, W.: National-scale estimates of ground-level $PM_{2.5}$ concentration in China using geographically weighted regression based on 3 km resolution MODIS AOD, Remote Sens., 8, 184, https://doi.org/10.3390/rs8030184, 2016a.

You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., and Wang, W.: National-scale estimates of ground-level $PM_{2.5}$ concentration in China using geographically weighted regression based on 3 km resolution MODIS AOD, Remote Sens., 8, 184, https://doi.org/10.3390/rs8030184, 2016b.

You, T., Wu, R., Huang, G., and Fan, G.: Regional meteorological patterns for heavy pollution events in Beijing, J. Meteorol. Res., 31, 597–611, https://doi.org/10.1007/s13351-017-6143-1, 2017.

Zhang, H., Hoff, R. M., and Engel-Cox, J. A.: The relation between moderate resolution imaging spectroradiometer (MODIS) aerosol optical depth and $PM_{2.5}$ over the United States: A geographical comparison by U.S. Environmental Protection Agency regions, J. Air Waste Manag., 59, 1358–1369, https://doi.org/10.3155/1047-3289.59.11.1358, 2009.

Zhang, H., Wang, Y., Hu, J., Ying, Q., and Hu, X. M.: Relationships between meteorological parameters and criteria air pollutants in three megacities in China, Environ. Res., 140, 242–254, https://doi.org/10.1016/j.envres.2015.04.004, 2015.

Zheng, C., Zhao, C., Zhu, Y., Wang, Y., Shi, X., Wu, X., Chen, T., Wu, F., and Qiu, Y.: Analysis of influential factors for the relationship between $PM_{2.5}$ and AOD in Beijing, Atmos. Chem. Phys., 17, 13473–13489, https://doi.org/10.5194/acp-17-13473-2017, 2017.

Zhu, Y., Hinds, W. C., Kim, S., and Sioutas, C.: Concentration and size distribution of ultrafine particles near a major highway, J. Air Waste Managas., 52, 1032–1042, https://doi.org/10.1080/10473289.2002.10470842, 2002.

Zou, B., Pu, Q., Bilal, M., Weng, Q., Zhai, L., and Nichol, J. E.: High–resolution Satellite Mapping of Fine Particulates Based on Geographically Weighted Regression, IEEE T. Geosci. Remote Sens., 13, 495–499, 2016.