Satellite-based Estimation of the Impacts of Summertime Wildfires on Particulate Matter Air Quality in 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, AL, 35806, USA

Abstract

Frequent and widespread wildfires in North Western United States and Canada has become the "*new normal*" during the northern hemisphere summer months, which significantly degrades particulate matter air quality in the United States ~~significantly~~. 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 (PM2.5) air quality in the United States. We use a Geographically Weighted Regression method to estimate surface PM2.5 in the United States between low (2011) and high (2018) fire activity years. Our results indicate that smoke aerosols caused significant pollution changes over half of the United States. We estimate that nearly 29 states have increased PM2.5 during the fire active year and 15 of these states have PM2.5 concentrations more than 2 times than that of the inactive year. Furthermore, these fires increased daily mean surface PM2.5 concentrations in Washington and Oregon by 38 to 259μgm^{-3} posing significant health risks especially to vulnerable populations. Our results also show that the GWR model can be

successfully applied to PM2.5 estimations from wildfires thereby providing useful information for various applications including 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 that 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 PM2.5 due to fires in North Western 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 (PM2.5) 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 accelerate 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 that 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 μgm^{-3} increase in PM2.5 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 PM2.5 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 PM2.5 estimate on a square kilometer basis. Since surface monitors are limited, satellite data has been used with numerous ancillary data sets to estimate surface PM2.5 at various spatial scales. ~~Given the sparsity of measurements in many parts of the world, it is not possible to use interpolation techniques between monitors to provide PM2.5 estimate on a square kilometer basis.~~ Several techniques have been developed to estimate surface PM2.5 using satellite observations from regional to global scales including simple linear regression, multiple linear regression, mixed-effect model, chemical transport model (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 550nm (mid-visible) aerosol optical depth (AOD) which is a unitless columnar measure of aerosol extinction ~~(Wang and Christopher, 2006)~~. Simple linear regression method uses satellite AOD as the only independent variable, which shows limited predictability compared to other method and correlation coefficients vary from 0.2 to 0.6 from the Western to Eastern United States (Zhang et al., 2009). Multiple linear regression method uses meteorological variables along with AOD data, and the prediction accuracy varies with different conditions including the height of boundary layer and other meteorological conditions (Goldberg et al., 2019; Gupta and Christopher, 2009b; Liu et al., 2005)~~(Liu, et al, 2005; Gupta and Christopher, 2009b)~~. For both univariate model and multi-variate models, AOD shows stronger correlation with PM2.5

during-fire episodes compared to pre-fire and post-fire periods (Mirzaei et al., 2018). Chemistry transport models (CTM) that scale the satellite AOD by the ratio of PM2.5 to AOD simulated by models can provide PM2.5 estimations without ground measurements, which are different than other statistical methods (Donkelaar et al., 2019, 2006). However, the CTM models that depend on reliable emission data usually show limited predictability at shorter time scales, and is largely useful for studies that require annual averages (Hystad et al., 2012).

The relationship among PM2.5, AOD and other meteorological variables is not spatially consistent (Hoff and Christopher, 2009; Hu, 2009). Therefore, methods that consider spatial variability can replicate surface PM2.5 with higher accuracy. One such method is the GWR, which is a non-stationary technique that models spatially varying relationships by assuming 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 PM2.5– (Hu, 2009) followed by the use of meteorological parameters and land use information (Hu et al., 2013). Several studies (Guo et al., 2021; Ma et al., 2014; You et al., 2016) successfully applied GWR model in estimating PM2.5 in China by using AOD and meteorological features as predictors. Similar to all the statistical methods, however, the GWR relies on 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 PM2.5.

In this paper, we use satellite data from the Moderate Resolution Imaging Spectroradiometer (MODIS) and surface PM2.5 data combined with meteorological and other ancillary information to develop and use the GWR method to estimate PM2.5. The use of the GWR method is not novel and we merely use a proven method to apply this towards surface PM2.5 estimations for forest fires. We calculate the change in PM2.5 between a high fire activity (2018)

with low fire activity (2011) periods during summer to assess the role of NWUSC wildfires on surface PM2.5 in the United States. The paper is organized as follows: We describe the data sets 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-day period (August 9th to August 25th) in 2018 (high fire activity) and 2011 (low fire activity) was selected based on analysis of total fires (details in methodology section) to assess surface PM2.5 (Table 1).

2.1 Ground level PM2.5 observations: Daily surface PM2.5 from the Environment Protection Agency (EPA) are 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. PM2.5 values less than 2 μgm^{-3} are discarded since they are lower than the established detection limit (Hall et al., 2013).

2.2 Satellite Data: The MODIS mid visible AOD from the Multi-Angle Implementation of Atmospheric Correction (MAIAC) product (MCD19A2 Version 6 data product) is used in this study. We used MAIAC retrieved Terra and Aqua MODIS AOD product at 1 km pixel resolution (Lyapustin et al., 2018). Different orbits are averaged to obtain mean daily values. Since thick smoke plumes generated by the wildfires can be detected as cloud by a large chance, we preserve possible cloud contaminated pixels to preserve the thick smoke pixels, and only AOD less than 0 will be discarded.Validation with AERONET studies show that 66% of the MAIAC AOD data

agree within ±0.5~±0.1 AOD (Lyapustin et al., 2018). Largely due to cloud cover, grid cells may have limited number of AOD observations within a certain period. On average, cloud free AOD data are available about 40% of the time during August 9th to August 25th in 2018 when fires were active in the region bounded by 25~50°N, 65~125°W. Smoke flag from the same product is used as a predictor in estimating surface PM2.5. The smoke detection is performed using MODIS red, blue and deep blue bands, and separate smoke pixels from dust and clouds based on absorption parameter, size parameter and thermal threshold (Lyapustin et al., 2018, 2012). Smoke flag data can provide the percentage of 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 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 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), 2m temperature (T2M), 10m wind speed (WS), surface relative humidity (RH) and surface pressure (SP) are obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5) product, with a spatial resolution of 0.25 degrees and temporal resolution of 1 hour and is matched temporally with the satellite overpass time. The BLH can provide information of aerosol layer height as aerosols are often found to be well-mixed within the boundary layer (Gupta and Christopher, 2009b). A higher RH will increase the hygroscopicity, change scattering properties of certain aerosols and can lead to a higher AOD value (Zheng et al., 2017). In addition, high surface temperatures can also accelerate the formation of secondary particles in the atmosphere.

~~**2.4 Land cover and population data:** Land cover and population density is highly related to anthropogenic aerosol emissions, which also affects surface PM2.5. Land cover information from the European Space Agency (ESA) with a spatial resolution of 300m and temporal resolution of one year (ESA, 2017) and population data from 'Gridded Population of the World', v4 (GPWv4) with a spatial resolution of 5 km are used as variables for the GWR method used in this study. The ESA land cover product uses global time series input datasets acquired by the Envisat Medium Resolution Imaging Spectrometer (MERIS) and from the SPOT-Vegetation (SPOT-VGT) sensors. The global land surface reflectance values are produced from the MERIS evel-1 dataset, which along with SPOT-VGT S1 (daily synthesis product) are used as input to the classification module and interprets into land cover classes. The population data uses results of the 2010 Population and Housing Census as input data.~~

3. Methodology

To assess the impact of NWUSC fires on PM2.5 in the United States, we first estimate the PM2.5 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 Northwestern (NW) US (35~49°N, 105~125°W). Table 2 shows the total FRP in Canada and 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 9 years, which provides the basis for the study. In order to create a 0.1° surface PM2.5, the GWR model is used to estimate the relationships of PM2.5 and AOD. Detailed processing steps for GWR model are shown in Figure 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 PM2.5 data and average the values if there is more than one value in one grid. Then the MAIAC AOD and smoke flag~~, land cover and population data~~ 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 & averaging: Next we select data where AOD and ground PM2.5 are both available (AOD > 0 and PM2.5 > 2.0 $\mu g\ m^{-3}$) and average them for the study period. This is to ensure that the AOD, PM2.5 and other variables match with each other, because PM2.5 is not a continuous measurement for some sites and AOD have 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's Information Criterion (AIC) is used for the GWR model (Loader, 1999). For locations that already have PM2.5 monitors, we calculate the mean AOD of a 0.5×0.5° box centered at the ground location and estimate the GWR coefficients (β) for AOD and meteorological~~/land cover~~ variables to estimate PM2.5. The model structure can be expressed as:

$$PM_{2.5i} = \beta_{0,i} + \beta_{1,i}AOD_i + \beta_{2,i}BLH_i + \beta_{3,i}T2M_i + \beta_{4,i}U10M_i + \beta_{5,i}RH_{sfci} + \beta_{6,i}SP_i ~~+ \beta_{7,i}LC_i~~ ~~+ \beta_{8,i}POP_i~~ + \beta_{7,i}SF_i + \varepsilon_i$$

where $PM_{2.5i}$ $(\mu g\ m^{-3})$ is the selected ground-level PM2.5 concentration at location i; $\beta_{0,i}$ is the intercept at location i; $\beta_{1,i}$~$\beta_{8,i}$ are the location-specific coefficients; AOD_i is the resampled AOD selected from MAIAC daily AOD data at location i; $BLH_i, T2M_i, U10M_i, RH_{sfci}, SP_i$ are selected meteorological parameters (BLH, T2M, WS, RH and PS) at location i; ~~LC_i is the resampled land cover data at location i; POP_i ($person/km^2$) is the~~

~~resampled population density 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 the Leave One Out Cross Validation (LOOCV) to test the model predictive performance (Kearns and Ron, 1999). Since the GWR model relies on adequate number of observations, the prediction accuracy will be lower if we preserve too much data for validation. Therefore, we choose the LOOCV method, which preserve only one data for validation at a time and repeat the process until all the data are used. In addition, R^2 and RMSE are calculated for both model fitting and model validation process to detect overfitting. Model overfitting will lead to low predictability, which means it fits too close to the limited number of data to predict for other places and will cause large bias.

3.4 Model prediction: While predicting the ground-level PM2.5 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 further than the 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 PM2.5 is not measured every day in all locations.

4. Results and Discussion

We first discuss the surface PM2.5 for a few select locations that are impacted by fires followed by the spatial distribution of MODIS AOD and the FRP for August 2018. We then assess the spatial distribution of surface PM2.5 from the GWR method. The validation of the GWR

method is then discussed. To further demonstrate the impact of the NWUSC fires on PM2.5 air quality in the United States, we show the spatial distribution of the difference between August 2018 and August 2011. We further quantify these results for ten 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 northwest US that burnt forest and grassland also affected air quality. Starting with the Cougar Creek Fire, then Crescent Mountain and Gilbert Fires, different wildfires in in NWUSC caused severe air pollution in various US cities. Figure 2a shows the rapid increase in PM2.5 of selected US cities from July 1st to August 31st, due to the transport of smoke from these wildfires. For all sites, July had low PM2.5 concentrations (<10 $\mu g\ m^{-3}$) and rapidly increases as fire activity increases. Calculating only from the EPA ground observations, the mean PM2.5 of the 17 days for the whole US is 13.7 $\mu g\ m^{-3}$ and the mean PM2.5 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 PM2.5 of all US stations (black line with no markers) and the mean PM2.5 of all WA stations (grey line with no markers). Ground-level PM2.5 reaches its peak between August 17th-21st and daily PM2.5 values during this time period far exceeds the 17-day mean PM2.5. For example, mean PM2.5 in WA on August 20th is 86.75 $\mu g\ m^{-3}$, which is more than two times the 17-day average of this region. On August 19th, Omak which is located in the foothills of the Okanogan Highlands in WA had PM2.5 values exceed 250 $\mu g\ m^{-3}$. According to a review of US wildfire caused PM2.5 exposures, 24-h mean PM2.5 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~~some~~ statistics of 15 states that have at least one daily record of non-attainment of ~~violating~~ EPA standard (>35 μg m^{-3} ~~^(-3))~~. From the frequency records of ~~violation~~non attainment in the 17-day period (last column), four states (Montana, Washington, California and Idaho) were~~as~~ consistently affected by the wildfires, and large portion of ground stations in these states were influenced by smoke aerosols~~, which represent that the pollution area is large~~. Most of the neighboring states also suffered from short-term but broad air pollution (third column). ~~One thing n~~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 total 8 EPA stations in Idaho, only two of them have consistent observations during the fire event; the other~~another~~ two stations have no valid observations, and the remaining~~other~~ four stations have only 2~6 observations during the 17-day period. Limited valid data along with unevenly distributed stations makes it hard to quantify smoke pollution in Northwestern US during the fire event period. Therefore, we utilize~~bring in~~ satellite data to enlarge the spatial coverage and estimate pollution at~~in~~ a finer spatial resolution.

Formatted: Superscript

The spatial distribution of AOD shown in Figure 2b indicates that the smoke from Canada is concentrated mostly in Northern US states such as WA, Oregon, Idaho, Montana, North Dakota and Minnesota. The black arrow shows the mean 800hPa-level mean wind for 17 days, and the length of the arrow represents the wind speed in ms^{-1}. Also shown in Figure 2b are wind speeds close to the fire sources which are about 4~5 ms^{-1}, and according to the distances and wind directions, it can take approximately 28~36 hours for the smoke to transport southeastward to Washington state. 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 fire radiative power (FRP) of every 2.3×3.5-degree box. The reason for not choosing a smaller grid for the FRP is to not clutter Figure 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 the Tweedsmuir Provincial Park of British Columbia in Canada (53.333N, 126.417W). The four separate lightning-caused wildfires burnt nearly 301,549 hectares of the boreal forest. The total FRP of August 2018 in Canada is about 5362 (*1000 MW), while the total FRP of August 2011 in Canada is 48 (* 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 PM2.5 in the United States.

4.2 Model Fitting and validation

The main goal for using GWR model is to help predict the spatial distribution of PM2.5 for places with no ground monitors while leveraging the satellite AOD and therefore it is important to ensure that the model is robust. Figure 3a and 3b 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 PM2.5 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 PM2.5 and using this estimated relationship we are able to assess surface PM2.5 using other parameters at locations where PM2.5 monitors are not available. The LOOCV process tests the model performance in predicting PM2.5. If the results of LOOCV has a large bias from the model fitting, then the predictability of the model

is low. Higher R^2 difference and RMSE difference value indicate that the model is overfitting and not suitable. The R^2 for the model fitting is 0.834, and the R^2 for the LOOCV is 0.797~~804~~; the RMSE for the GWR model fitting is 3.464 $\mu g\ m^{-3}$, and for LOOCV the RMSE is 3.84 ~~77~~$\mu g\ m^{-3}$. There are minor differences between fitting R^2 and validation R^2 (0.037~~0.036~~) and between fitting RMSE and validation RMSE (0.376~~0.37~~ $\mu g\ m^{-3}$) suggesting that the model is not over-fitting and has stable predictability further indicating that the model can predict surface PM2.5 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 R^2 of 0.745~~6~~ and RMSE of 4.3~~15~~ $\mu g\ m^{-3}$. The increase/decrease in the cross validated R^2 and RMSE indicates the importance of sufficient data 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 3~5 $\mu g\ m^{-3}$, which is a reasonable error range for 17-day average prediction of PM2.5. For data greater than the 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 ~~whole model~~ RMSE when using the entire model. Therefore, the model has a tendency for underestimating PM2.5 exceedances by around 12.07 $\mu g\ m^{-3}$. The larger the PM2.5 is, the greater the model underestimates.

4.3 Predictors' influence during wildfires

Table 4 shows the mean and different region coefficients from the GWR model. Boxes are shown in figure 4c in different colors: box1 (red) located in Washington state is nearby the fire sources; box2 (gold) located in Montana state is influenced from both neighboring states and ~~remote~~ smoke from Canada; box3 (green) in Minnesota which is~~state~~ located further from the fires and has minor increase in PM2.5 due to remote smoke; box4 (black) in Pennsylvania state is the furthest from fires and has no obvious pollution

increase. By comparing the coefficients in these boxes, predictors have different influence in different locations. AOD has stronger influence on predicting PM2.5 closer to fire sources, but local emissions become more dominant if the distances is large enough. The smoke flag is overall positive related to surface PM2.5, while it could slightly negatively relate to PM2.5 around fire sources and northeastern coasts. PBL is~~are~~ negatively related to PM2.5 when the pollution is concentrated near the~~around~~ surface (fires or human-made emissions), while it appears to be positively related to PM2.5 at locations where the main pollution source comes from remote wildfire smoke. Surface temperature have a relative stable positive correlation with surface PM2.5, however, surface pressure and wind speeds are~~have~~ negatively correlated~~ion~~ with PM2.5. Relative humidity, on the other hand, shows large variations on PM2.5 influence across the nation. Around the wildfires where the RH is relative low, RH has a positive correlation with PM2.5 since hygroscopicity would increase and leads to accumulation of PM2.5, but increasing RH can also decrease PM2.5 concentration by overgrowing the PM2.5 particles to deposition at high RH environment (Chen et al., 2018).

Formatted: Font: 11 pt

4.43 Predicted PM2.5 Distribution

The mean PM2.5 distributions over the United States shown in Figure 4a is calculated by averaging the surface PM2.5 data from ground monitors for the 17 days, which matches well with the GWR model-predicted PM2.5 distributions shown in Figure 4b. The model estimation extends the ground measurements and provide pollution assessments across the entire nation. Comparing the AOD map (Figure 2b) with the PM2.5 estimations (Figure 4b), demonstrates the differences between columnar and surface-level pollution. Differences between the AOD and PM2.5 distributions are due to various reasons including 1) Areas with high PM2.5 concentrations in figure 4b correspond to low AOD values in figure 2b (Southern California, Utah, and southern US); 2) and high AOD regions in figure 2b correspond to low PM2.5 concentrations in figure 4b (Minnesota). The first situation usually occurs at the edge of polluted areas that are relative far

from the fire source, which is consistent with previous studies that reported smaller particles (<10 μg) are able to travel longer distances compared to large particles (>10 μg) (Gillies et al., 1996), and that lager particles tend to settle closer to their source (Sapkota et al., 2005; Zhu et al., 2002).

We use the same method for August 9th to August 25th in 2011 that had low fire activity, ensuring consistency for estimating coefficients for different variables for 2011. Figure 4c shows the difference in spatial distribution of mean ground PM2.5 of the 17 days between 2018 and 2011. High values of PM2.5 differences 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 the 48 states within the study region, there are 29 states that have a higher PM2.5 value in 2018 than 2011, and 15 states have 2018 PM2.5 value more than two times their 2011 value (shown in ~~table 3~~figure 5). The mean PM2.5 for WA increases from 5.87 in 2011 to ~~47.1~~46.47 $\mu g\ m^{-3}$ in 2018, which is about 8 times more than 2011 values. The PM2.5 values in Oregon increases from 4.97 (in 2011) to 33.3~~1~~ $\mu g\ m^{-3}$ in 2018, which is nearly seven times more than in 2011. For states from Montana to Minnesota, the mean PM2.5 decreases from east to west, which reveals the path of smoke transport. As shown in Figure 4c, there is a clear transport path of smoke from North Dakota all the way to Texas. Along the path, smoke increases PM2.5 concentrations by ~~173~~168% in North Dakota and 2~~76.2~~% in Texas. Smoke aerosols transported over long distances contains fine fraction PM which significantly affect the health of children, adults, and vulnerable groups.

Figure 5 shows the mean PM2.5 predicted from the GWR model of different EPA regions for the 17 days in 2011 and 2018 (Hawaii and Alaska are not included). The most influenced region is region 10, which has a 2018 mean PM2.5 value of 34.2~~7~~ $\mu g\ m^{-3}$ that is 6 times larger than the values in 2011 (5.8 $\mu g\ m^{-3}$) values. The PM2.5 of region 8 and 9 have 2.~~7 4~~ and 2.6~~5~~ times increase in 2018 compared to 2011. Region 1~4 have lower PM2.5 in 2018 than 2011 possibly

due to Clean Air Act initiatives, absence of any major fire activites and further away for transported aerosols. The emission reduction improves the US air quality and lower the PM2.5 every year, but 6 out of 10 EPA regions show significant increases in PM2.5 during the study period, which indicates that the long-range transported wildfire smoke has become the new major pollutant in the US.

4.54 Estimation of Canadian fire pollution

To evaluate the pollution caused only from Canadian fires, we did a rough assessment according to the total FRP and PM2.5 values. There are three states in the US have wildfires during the study period: California, Washington and Oregon, and they have total FRP of 1186, 518 and 439 (*1000 MW) respectively. Assuming that California was only influenced by the local fires, then fires of 1186 (*1000 MW) cause 13 $\mu g\ m^{-3}$ increase in PM2.5. Accordingly, wildfires in Washington and Oregon State will cause 6 and 5 $\mu g\ m^{-3}$ increase in state mean PM2.5. Therefore, Canadian fires caused PM2.5 increase in Washington and Oregon is about 35 and 23 $\mu g\ m^{-3}$. Since the FRP of Canadian wildfires are approximately 5 times larger than that of the California fires, which is the strongest fire in US, we assume the pollution affecting the states located in the downwind directions other than the three states are mainly coming from Canadian wildfires. States with no local fires such as Montana, North Dakota, South Dakota and Minnesota have PM2.5 increase of 18.31, 12.8, 10.4 and 10.13 $\mu g\ m^{-3}$. The decrease of these numbers reveal that the smoke is transport in a SE direction. This influence of Canadian wildfires on US air quality is only a rough quantity estimation, thus additional work is needed for understand long-range transport smoke pollution and its impact on public health. One way to do this would be assessing the difference of pollution by turning on and off US fires in chemistry models.

4.64 Model uncertainties and limitations

There are various sources of uncertainties and limitations for studies that use satellite data to estimate surface PM2.5 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 data sets. Although satellite observations have excellent spatial coverage, missing data due to cloud cover is a limitation. As discussed in the paper, the prediction error (RMSE) of the model is between 3~5 $\mu g\ m^{-3}$. The GWR model is largely influenced by the distribution of ground stations, and the prediction error will be different in different places due to unevenly distributed PM2.5 stations. For locations that have a dense ground-monitoring distribution, the prediction error will be low, while the prediction error will be relative larger at 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 PM2.5 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 (Donkelaar et al., 2015; Hu et al., 2013; You et al., 2015; Zou et al., 2016). ~~While Land cover and land use information can improve PM2.5 estimation predictability, redundant information such as NDVI can cause overfitting of models. Therefore, in order to control the number of predictors used in the GWR model, we use only one piece of land cover information. However, which of the land cover and land use information performs better in predicting surface PM2.5 is still to be assessed in the future.~~ 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.

One limitation of this study is that analysis based on 17-day mean values cannot capture daily pollution variations, which is also very important for pollution estimation during rapid-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 PM2.5 for 17 days in August for a high fire active year (2018) and compare that with a low fire activity year using the Geographically Weighted Regression (GWR) method to assess the increase in PM2.5 in the United States due to smoke transported from fires. The difference in PM2.5 between the two years indicates that more than half of the US states (29 states) are influenced by the NWUSC wildfires, and half of the affected states have 17-day mean PM2.5 increases larger than 100% of the baseline value. The peak PM2.5 during the wildfires can be much larger than the 17-day average and can affect vulnerable populations susceptible to air pollution. Some of the most affected states are in Washington, California, Wisconsin, Colorado and Oregon, all of which have populations greater than 4 million. According to CDC (Centers for Disease Control and Prevention), 8% of the population have asthma (CDC, 2011). Therefore, for asthma alone, there are about 3 million people facing significant health issue due to the long-range transport smoke in these states.

For states that show decrease in PM2.5 due to the Clean Air Act, the mean decrease is about 16% of the baseline after 7 years. This is consistent with EPA's report that there is a 23% decrease of PM2.5 in national average from 2010 to 2019(U.S. Environmental Protection Agency, 2019). Comparing with the dramatic increase (132%) caused by wildfires, pollution from the fires is counteracting our effort on emission controls. Although wildfires are often episodic and short-

term, high frequency of fire occurrence and increasing longer durations of summertime wildfires in recent years has made them now a long-term influence on public lives. Our results show a significant increase of 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.

~~This study is novel in: 1) applying PM2.5 estimation methods on wildfire events and calculate the prediction error at high pollution concentration condition; 2) analyzing predictors' different influences in estimating PM2.5 under various conditions; 3) quantify the air pollution from fires by states and EPA regions.~~

Formatted: Indent: First line: 0.5"

Acknowledgements.

Pawan Gupta was supported by a NASA Grant. MODIS data were acquired from the Goddard DAAC. We thank all the data providers for making this research possible.

References

Formatted: Justified

Apte, J.S., Brauer, M., Cohen, A.J., Ezzati, M., Pope, C.A., 2018. Ambient PM2.5 Reduces Global and Regional Life Expectancy. Environ. Sci. Technol. Lett. 5, 546–551. https://doi.org/10.1021/acs.estlett.8b00360

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

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

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

438 CDC, 2011. Asthma in the US. CDC Vital Signs 1–4.

439 Chen, Z., Xie, X., Cai, J., Chen, D., Gao, B., He, B., Cheng, N., Xu, B., 2018. Understanding
440 meteorological influences on PM2.5 concentrations across China: A temporal and spatial
441 perspective. Atmos. Chem. Phys. 18, 5343–5358. https://doi.org/10.5194/acp-18-5343-2018

442 Chu, Y., Liu, Y., Li, X., Liu, Z., Lu, H., Lu, Y., Mao, Z., Chen, X., Li, N., Ren, M., Liu, F., Tian,
443 L., Zhu, Z., Xiang, H., 2016. A review on predicting ground PM2.5 concentration using
444 satellite aerosol optical depth. Atmosphere (Basel). 7, 129.
445 https://doi.org/10.3390/atmos7100129

446 Coogan, S.C.P., Robinne, F.N., Jain, P., Flannigan, M.D., 2019. Scientists' warning on wildfire
447 — a canadian perspective. Can. J. For. Res. 49, 1015–1023. https://doi.org/10.1139/cjfr-
448 2019-0094

449 Donkelaar, A. Van, Martin, R. V., Li, C., Burnett, R.T., 2019. Regional Estimates of Chemical
450 Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method
451 with Information from Satellites, Models, and Monitors. Environ. Sci. Technol. 53, 2595–
452 2611. https://doi.org/10.1021/acs.est.8b06392

453 Donkelaar, A. Van, Martin, R. V, Park, R.J., 2006. Estimating ground-level PM 2 . 5 using

454 aerosol optical depth determined from satellite remote sensing. J. Geophys. Res. Atmos.
455 111. https://doi.org/10.1029/2005JD006996

456 Donkelaar, A. Van, Martin, R. V, Spurr, R.J.D., Burnett, R.T., 2015. High-Resolution Satellite-
457 Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over
458 North America. Environ. Sci. Technol. 49, 10482–10491.
459 https://doi.org/10.1021/acs.est.5b02076

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

463 Fotheringham, A.S., Charlton, M.E., Brunsdon, C., 1998. Geographically weighted regression: a
464 natural evolution of the expansion method for spatial data analysis. Environ. Plan. A 30,
465 1905–1927.

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

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

472 Goldberg, D.L., Gupta, P., Wang, K., Jena, C., Zhang, Y., Lu, Z., Streets, D.G., 2019. Using gap-
473 filled MAIAC AOD and WRF-Chem to estimate daily PM2.5 concentrations at 1 km
474 resolution in the Eastern United States. Atmos. Environ. 199, 443–452.
475 https://doi.org/10.1016/j.atmosenv.2018.11.049

476 Gu, Y., 2019. Estimating PM2 . 5 Concentrations Using 3 km MODIS AOD Products : A Case
477 Study in British Columbia , Canada. University of Waterloo.

478 Guo, B., Wang, X., Pei, L., Su, Y., Zhang, D., Wang, Y., 2021. Identifying the spatiotemporal
479 dynamic of PM2.5 concentrations at multiple scales using geographically and temporally
480 weighted regression model across China during 2015–2018. Sci. Total Environ. 751.
481 https://doi.org/10.1016/j.scitotenv.2020.141765

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

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

488 Hall, E.S., Kaushik, S.M., Vanderpool, R.W., Duvall, R.M., Beaver, M.R., Long, R.W.,
489 Solomon, P.A., 2013. Intergrating Sensor Monitoring Technology into Current Air
490 Pollution Regulatory Support Paradign: Practical Considerations. Am. J. Environ. Eng 4,
491 147–154. https://doi.org/10.5923/j.ajee.20140406.02

492 Hessburg, P.F., Churchill, D.J., Larson, A.J., Haugo, R.D., Miller, C., Spies, T.A., North, M.P.,
493 Povak, N.A., Belote, R.T., Singleton, P.H., Gaines, W.L., Keane, R.E., Aplet, G.H.,
494 Stephens, S.L., Morgan, P., Bisson, P.A., Rieman, B.E., Salter, R.B., Reeves, G.H., 2015.
495 Restoring fire-prone Inland Pacific landscapes: seven core principles. Landsc. Ecol. 30,
496 1805–1835. https://doi.org/10.1007/s10980-015-0218-0

497 Hoff, R.M., Christopher, S.A., 2009. Remote Sensing of Particulate Pollution from Space : Have

498 We Reached the Promised Land ? J. Air Waste Manage. Assoc. 59, 645–675.
499 https://doi.org/10.3155/1047-3289.59.6.645

500 Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M., Quattrochi,
501 D.A., Sarnat, J.A., Liu, Y., 2013. Estimating ground-level PM2.5 concentrations in the
502 southeastern U.S. using geographically weighted regression. Environ. Res. 121, 1–10.
503 https://doi.org/10.1016/j.envres.2012.11.003

504 Hu, Z., 2009. Spatial analysis of MODIS aerosol optical depth, PM2.5, and chronic coronary
505 heart disease. Int. J. Health Geogr. 8, 1–10. https://doi.org/10.1186/1476-072X-8-27

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

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

513 J.A.Gillies, W.G.Nickling, G.H.Mctainsh, 1996. Dust concentration s and particle-size
514 characteristics of an intense dust haze event: inland delta region. Atmos. Environ. 30, 1081–
515 1090.

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

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

523 Liu, Y., Sarnat, J.A., Kilaru, V., Jacob, D.J., Koutrakis, P., 2005. Estimating ground-level PM2.5
524 in the eastern United States using satellite remote sensing. Environ. Sci. Technol. 39, 3269–
525 3278. https://doi.org/10.1021/es049352m

526 Loader, C.R., 1999. BANDWIDTH SELECTION: CLASSICAL OR PLUG-IN? Ann. Stat. 27,
527 415–438.

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

531 Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS Collection 6 MAIAC Algorithm.
532 Atmos. Meas. Tech. 11, 5741–5765. https://doi.org/10.5194/amt-2018-141

533 Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level PM2.5 in china using
534 satellite remote sensing. Environ. Sci. Technol. 48, 7436–7444.
535 https://doi.org/10.1021/es5009399

536 Meixner, T., Wohlgemuth, P., 2004. Wildfire Impacts on Water Quality. J. Wildl. Fire 13, 27–
537 35.

538 Melillo, J.M., Richmond, T., Yohe, G.W., 2014. Climate Change Impacts in the United States.
539 Third Natl. Clim. Assess. 52. https://doi.org/10.7930/J0Z31WJ2.

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

Mirzaei, M., Bertazzon, S., Couloigner, I., 2018. 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

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., Bouchet, V., 2017. Multi-Year (2013–2016) PM2.5 Wildfire Pollution Exposure over North America as Determined from Operational Air Quality Forecasts. Atmosphere (Basel). 8, 179. https://doi.org/10.3390/atmos8090179

Navarro, K.M., Schweizer, D., Balmes, J.R., Cisneros, R., 2018. 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

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

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., 2005. 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

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

564 U.S. Environmental Protection Agency, 2019. Particulate Matter (PM2.5) Trends.

565 You, W., Zang, Z., Pan, X., Zhang, L., Chen, D., 2015. Estimating PM2.5 in Xi'an, China using
566 aerosol optical depth: A comparison between the MODIS and MISR retrieval models. Sci.
567 Total Environ. 505, 1156–1165. https://doi.org/10.1016/j.scitotenv.2014.11.024

568 You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., Wang, W., 2016. National-scale estimates of
569 ground-level PM2.5 concentration in China using geographically weighted regression based
570 on 3 km resolution MODIS AOD. Remote Sens. 8. https://doi.org/10.3390/rs8030184

571 Zhang, H., Hoff, R.M., Engel-Cox, J.A., 2009. The relation between moderate resolution
572 imaging spectroradiometer (MODIS) aerosol optical depth and PM2.5 over the United
573 States: A geographical comparison by U.S. Environmental Protection Agency regions. J.
574 Air Waste Manag. Assoc. 59, 1358–1369. https://doi.org/10.3155/1047-3289.59.11.1358

575 Zheng, C., Zhao, C., Zhu, Y., Wang, Y., Shi, X., Wu, X., Chen, T., Wu, F., Qiu, Y., 2017.
576 Analysis of influential factors for the relationship between PM2.5 and AOD in Beijing.
577 Atmos. Chem. Phys. 17, 13473–13489. https://doi.org/10.5194/acp-17-13473-2017

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

581 Zou, B., Pu, Q., Bilal, M., Weng, Q., Zhai, L., Nichol, J.E., 2016. High–resolution Satellite
582 Mapping of Fine Particulates Based on Geographically Weighted Regression. Ieee Geosci.

583 Remote Sens. Lett. 13, 495–499.

584

585

586

Table 1. Datasets used in the study with sources.

	Data /Model	Sensor	Spatial Resolution	Temporal Resolution	Accuracy
1	Surface PM2.5	TEOM	Point data	daily	±5~10%
2	Mid visible aerosol optical depth (AOD)	MAIAC_ MODIS	1km	daily	66% compared to AERONET
3	Fire Radiative Power (FRP)	Terra/Aqua-MODIS	1km	daily	± 7%
4	ECMWF (Meteorological variables)		0.25 degree	hourly	
~~5~~	~~Land cover~~	~~MERIS SR~~	~~300m~~	~~Annual~~	

Formatted Table

1) https://www.epa.gov/outdoor-air-quality-data
2) https://earthdata.nasa.gov/
3) https://earthdata.nasa.gov/
4) https://www.ecmwf.int/en/forecasts
~~5) https://www.esa-landcover-cci.org~~
~~6) https://sedac.ciesin.columbia.edu/data/collection/gpw-v4~~

Table 2. Total FRP in Canada and 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

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

State	number of site violate standard	number of site in the state	Percentage of site violate 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

Formatted Table

Table 4. Coefficients of different predictors

	AOD	Smoke flag	PBL	T2M	RH	U	SP
box1(red)	92.01	-0.13	-1.5	0.2	-0.01	-1.6	-0.037
box2(gold)	63.97	0.002	-2.86	0.09	-0.11	-1.5	-0.02
box3(green)	5.9	0.044	0.3	0.16	0.017	-0.2	-0.015
box4(black)	6.72	-0.02	-1.3	0.28	-0.03	0.13	-0.007
mean	28.1	0.02	-0.89	0.06	-0.19	-0.67	-0.002

Formatted Table

Formatted: Line spacing: single

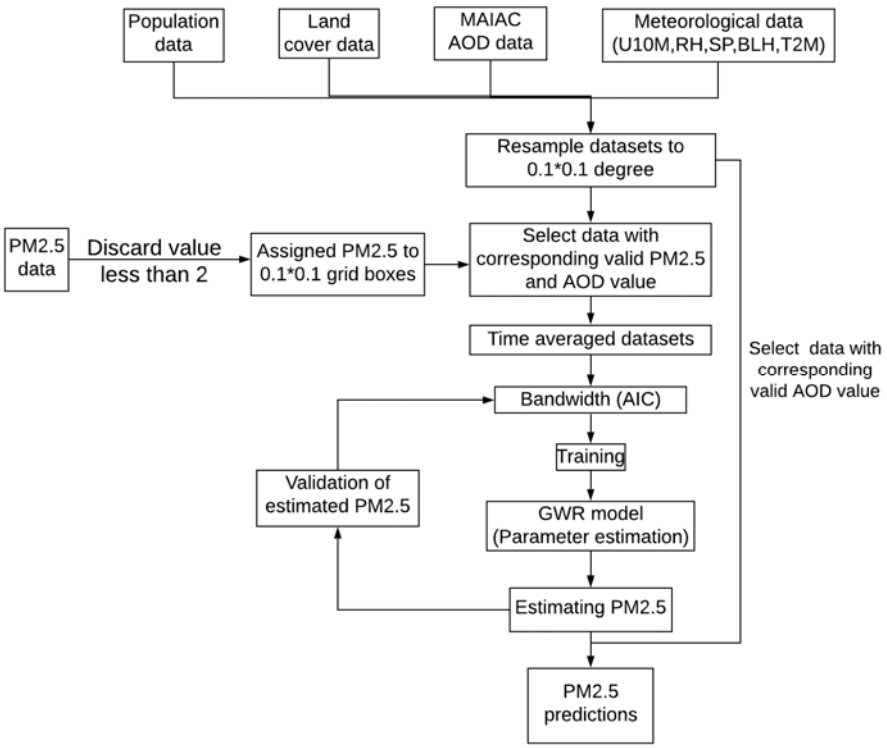
606 Table 3. Mean PM2.5 from August 9th to August 25th in 2018 and 2011 of different states

State	2018	2011
WA	47.0	5.874
OR	33.10	4.97
ID	26.26	6.79
MT	25.86	7.55
CA	21.22	7.66
ND	20.20	7.41
NV	17.92	5.51
SD	17.72	7.31
MN	16.41	6.27
WY	15.71	6.59
NE	15.69	6.81
UT	14.87	6.51
IA	14.73	7.87
KS	14.13	6.84
AR	13.92	11.59
WI	13.70	6.26
OK	13.53	9.26
MO	13.25	9.64
LA	13.24	13.07
IL	12.98	11.36
MS	12.86	13.67
CO	12.31	6.07
MI	11.97	6.82
TX	11.70	9.27
TN	11.66	14.39
AL	11.65	14.97
IN	11.50	12.51
KY	11.02	13.19
DC	10.65	13.16
NJ	10.56	9.76
DE	10.37	11.16
GA	10.22	14.02
CT	10.20	9.74
OH	10.14	11.84
FL	10.13	10.68
MD	10.07	12.59
NM	9.842	6.03
SC	9.829	12.66
PA	9.75	12.64
NC	9.67	12.44

~~RI~~	~~9.633~~	~~8.602~~
~~MA~~	~~9.56~~	~~9.413~~
~~VA~~	~~9.38~~	~~13.74~~
~~NY~~	~~9.33~~	~~9.731~~
~~WV~~	~~9.28~~	~~13.58~~
~~NH~~	~~9.11~~	~~9.33~~
~~AZ~~	~~9.08~~	~~7.00~~
~~VT~~	~~8.96~~	~~9.34~~
~~ME~~	~~7.972~~	~~10.52~~

607
608
609
610
611
612
613

Formatted: Left


Population data
Land cover data
MAIAC AOD data
Meteorological data (U10M,RH,SP,BLH,T2M)
Resample datasets to 0.1*0.1 degree
PM2.5 data
Discard value less than 2
Assigned PM2.5 to 0.1*0.1 grid boxes
Select data with corresponding valid PM2.5 and AOD value
Time averaged datasets
Select data with corresponding valid AOD value
Bandwidth (AIC)
Training
Validation of estimated PM2.5
GWR model (Parameter estimation)
Estimating PM2.5
PM2.5 predictions

614

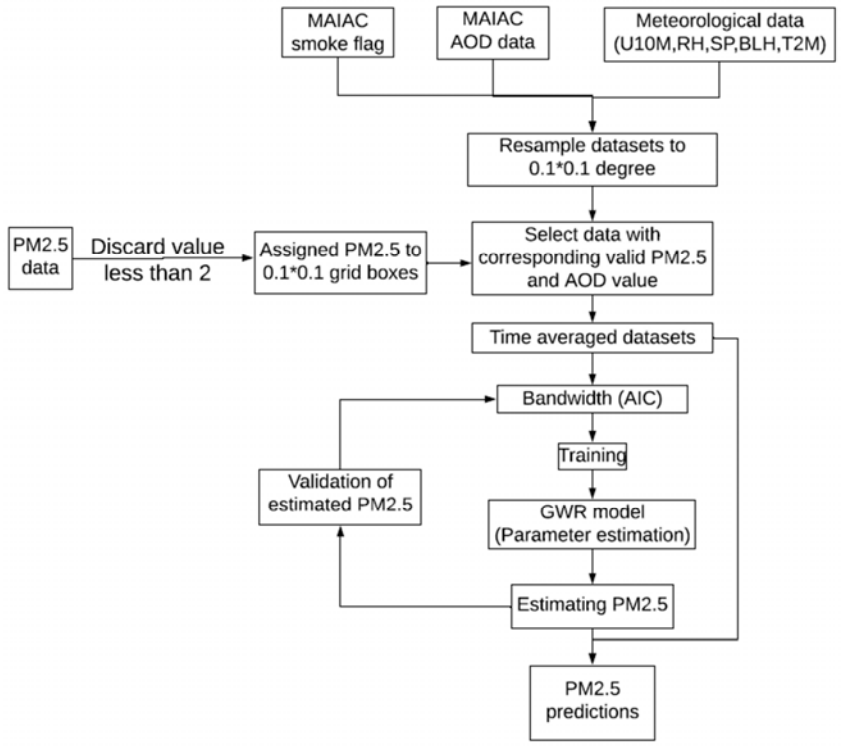


Figure 1. Flow chart for the Geographically Weighted Regression model used. All satellite, ground, meteorological data are gridded to 0.1 by 0.1 degrees.

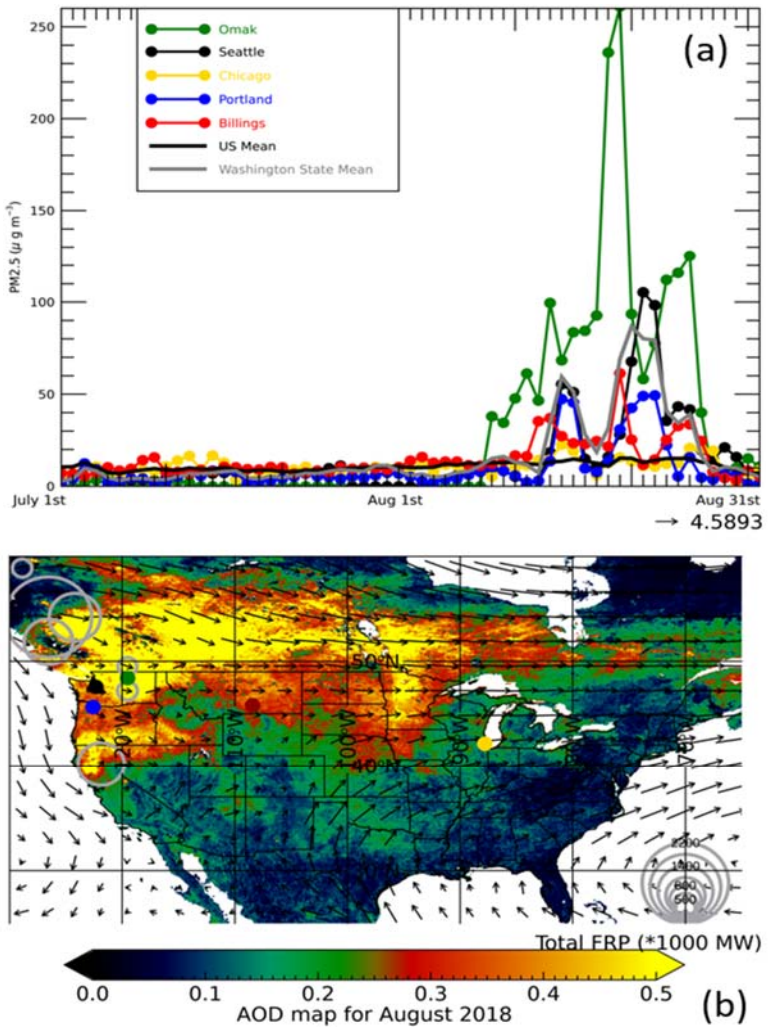
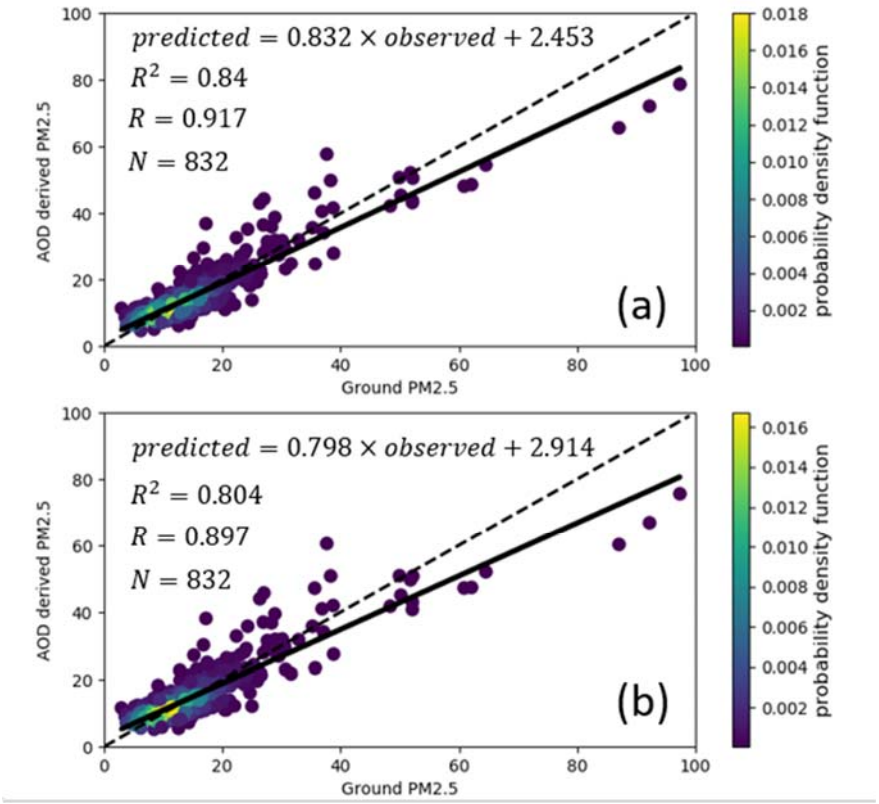


Figure 2. (a) Variations of EPA ground observed PM2.5 in different cities from July to August 2018 (Omak-Washington, Seattle-Washington, Chicago-Illinois, Portland-Oregon, Billings-Montana). Black line without markers shows the mean variation of the whole US stations and the grey line without markers shows the mean variation of stations in Washington state. (b) Mean MAIAC satellite AOD distribution from August 9th to August 25th, 2018. AOD values equal or larger than 0.5 are shown as the same color (yellow). Also shown are circles with Fire Radiative Power (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).


predicted = 0.832 × observed + 2.453
R^2 = 0.84
R = 0.917
N = 832
(a)
AOD derived PM2.5
Ground PM2.5
probability density function
predicted = 0.798 × observed + 2.914
R^2 = 0.804
R = 0.897
N = 832
(b)
AOD derived PM2.5
Ground PM2.5
probability density function

631

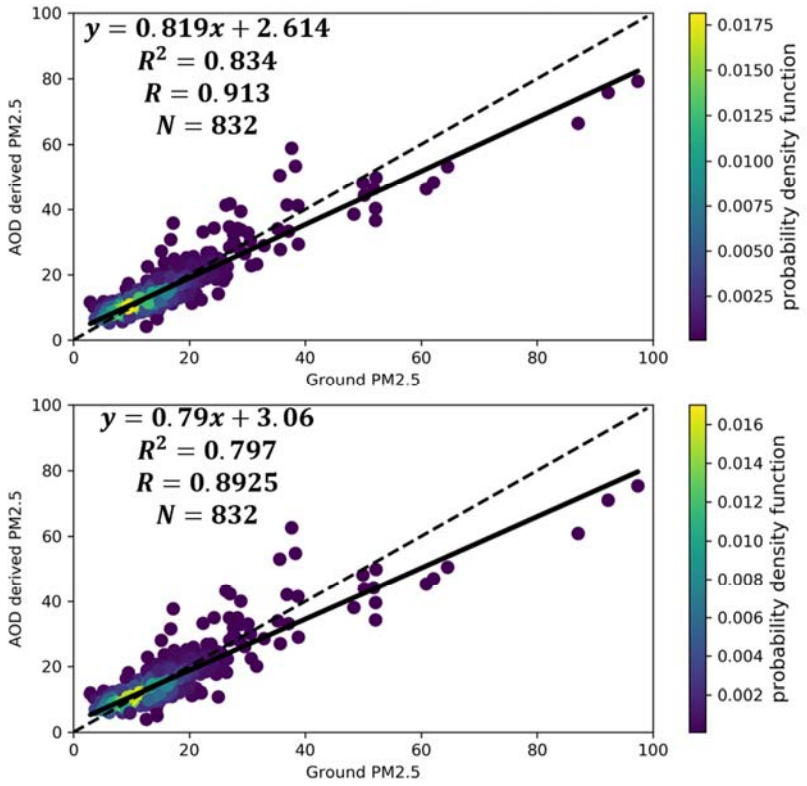


Figure 3. Results of model fitting and cross validation for GWR model for the entire US region averaged from August 9th to August 25th, 2018. (a) GWR model fitting results (b) GWR model LOOCV results. The dash line is the 1:1 line as reference and the black line shows the regression line. The color of 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.

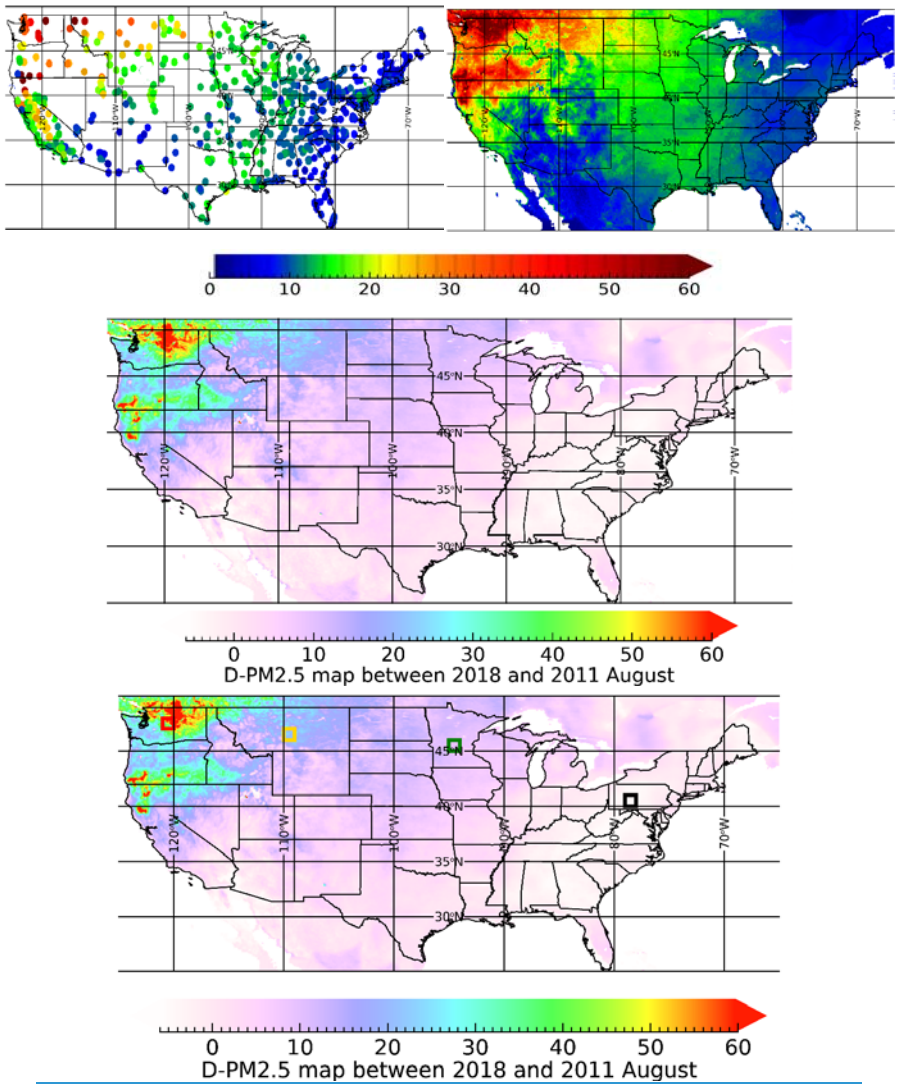


Figure 4. (a) EPA ground observed PM2.5 distribution over the US averaged from August 9th to August 25th, 2018. (b) GWR predicted 17-day mean PM2.5 distribution. (c) Difference map of predicted ground PM2.5 of the 17-day mean values between 2018 and 2011. PM2.5 values equal or larger than 30 $\mu g\ m^{-3}$ are shown as the same color (red). Note that the D-PM2.5 has a different color scale to make the negative values more apparent (blue).

655

656

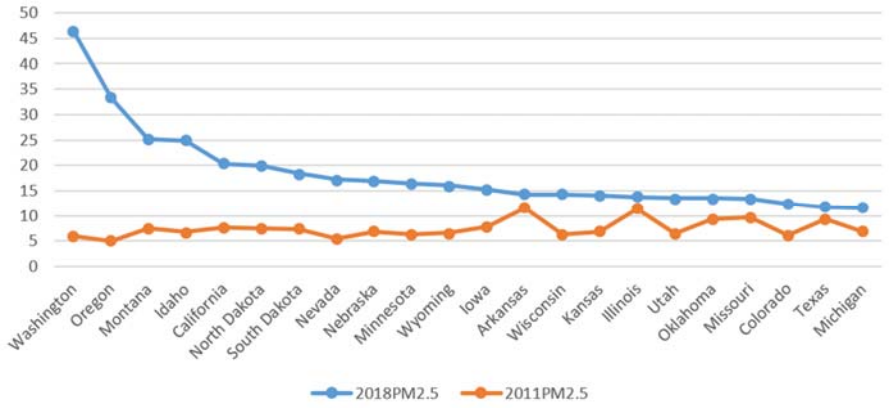
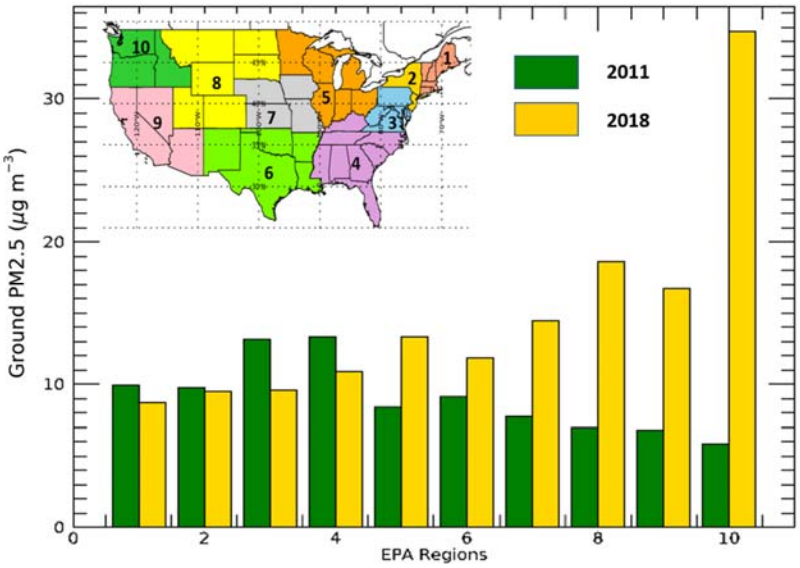


Figure 5. Mean PM2.5 from August 9th to August 25th in 2018 and 2011 of most affected states

Formatted: Centered


2011
2018
Ground PM2.5 (μg m^{-3})
EPA Regions
0
10
20
30
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10

660

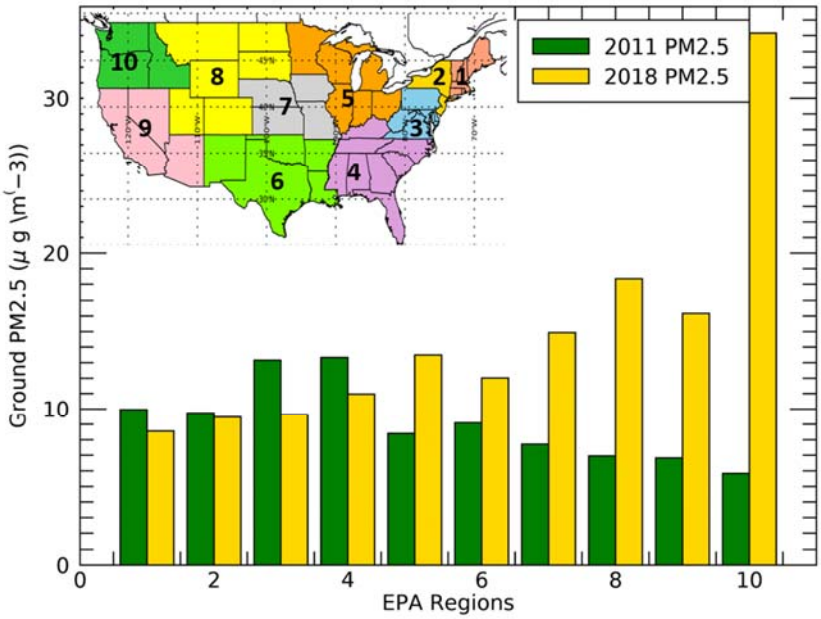


Figure 6~~5~~. Mean PM2.5 of EPA regions from August 9th to August 25th in 2011 and 2018. Inset shows the map of 10 EPA regions in different colors. Yellow column represents the 2018 mean PM2.5 and green column represents for 2011 mean PM2.5.