Environmental Science & Technology

Article

pubs.acs.org/est

Estimating Ground-Level $PM_{2.5}$ in China Using Satellite Remote Sensing

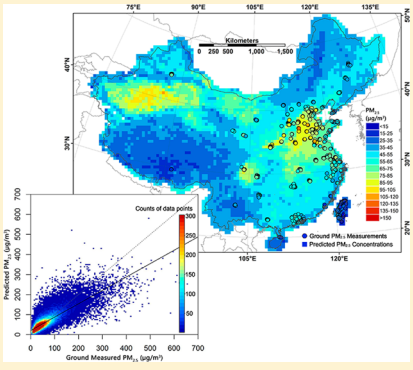
Zongwei Ma,[†,‡] Xuefei Hu,[‡] Lei Huang,[†] Jun Bi,*[,†] and Yang Liu*[,‡]

[†]State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Xianlin Campus, Box 624, 163 Xianlin Avenue, Nanjing 210023, China

[‡]Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, Georgia 30322, United States

Supporting Information

ABSTRACT: Estimating ground-level $PM_{2.5}$ from satellite-derived aerosol optical depth (AOD) using a spatial statistical model is a promising new method to evaluate the spatial and temporal characteristics of $PM_{2.5}$ exposure in a large geographic region. However, studies outside North America have been limited due to the lack of ground $PM_{2.5}$ measurements to calibrate the model. Taking advantage of the newly established national monitoring network, we developed a national-scale geographically weighted regression (GWR) model to estimate daily $PM_{2.5}$ concentrations in China with fused satellite AOD as the primary predictor. The results showed that the meteorological and land use information can greatly improve model performance. The overall cross-validation (CV) R^2 is 0.64 and root mean squared prediction error (RMSE) is 32.98 $\mu g/m^3$. The mean prediction error (MPE) of the predicted annual $PM_{2.5}$ is 8.28 $\mu g/m^3$. Our predicted annual $PM_{2.5}$ concentrations indicated that over 96% of the Chinese population lives in areas that exceed the Chinese National Ambient Air Quality Standard (CNAAQS) Level 2 standard. Our results also confirmed satellite-derived AOD in conjunction with meteorological fields and land use information can be successfully applied to extend the ground $PM_{2.5}$ monitoring network in China.



INTRODUCTION

Epidemiologic studies have shown that exposure to fine particulate matter ($PM_{2.5}$, particles with aerodynamic diameter less than 2.5 μm) is associated with increased cardiovascular and respiratory morbidity and mortality.[1,2] With the rapid economic development, $PM_{2.5}$ pollution is becoming a severe environmental problem in China and has been attracting great attention.[3,4] Limited studies show that $PM_{2.5}$ pollution has had a severe health impact in China. For example, one study shows that heavy smog is associated with an increased risk of hospital visits and hospital admissions in Beijing, especially for cardiovascular and respiratory disease.[5] Another study shows that all-cause and cardiovascular mortalities are significantly associated with $PM_{2.5}$ exposure in Xi'an City.[6] Epidemiologic research on the health impacts of $PM_{2.5}$ exposure continues to be a research hotspot in China.

Epidemiologic studies of $PM_{2.5}$ require long-term, accurate $PM_{2.5}$ exposure data, which historically have been provided by ground monitoring networks. Before 2013, however, there were only a few cities in China (e.g., Nanjing and Guangzhou City) where had research monitoring sites.[7,8] The lack of spatially and temporally continuous ground $PM_{2.5}$ measurements before 2013 makes it difficult to assess the historical spatial and temporal variability of $PM_{2.5}$ and has substantially restrained epidemiological studies of $PM_{2.5}$ in China.

The application of satellite remote sensing in air quality studies has evolved greatly in recent years.[9] Because satellite remote sensing provides comprehensive spatial and temporal coverage, estimating $PM_{2.5}$ using satellite remote sensing can supplement the ground $PM_{2.5}$ monitoring network. The National Aeronautics and Space Administration (NASA) launched two Earth Observing System (EOS) satellites, the Terra and Aqua, in 1999 and 2002, respectively.[10] The Moderate Resolution Imaging Spectroradiometer (MODIS) and Multiangle Imaging SpectroRadiometer (MISR) aboard the Terra satellite and a second MODIS aboard the Aqua satellite measure aerosol optical depth (AOD), which is retrieved at visible and near-infrared wavelengths. The corresponding particle sizes of AOD range from 0.1 to 2 μm, which is similar to the size range of $PM_{2.5}$.[11]

Early studies examined the relationship between $PM_{2.5}$ and satellite-derived AOD using simple linear regression model[12] or applying local scaling factors from a global atmospheric chemistry model,[13] and found that AOD derived from satellite remote sensing can be an effective tool for $PM_{2.5}$ pollution monitoring. One study estimated global $PM_{2.5}$ concentrations

Received: February 23, 2014
Revised: May 23, 2014
Accepted: June 5, 2014
Published: June 5, 2014

dx.doi.org/10.1021/es5009399 | Environ. Sci. Technol. 2014, 48, 7436−7444

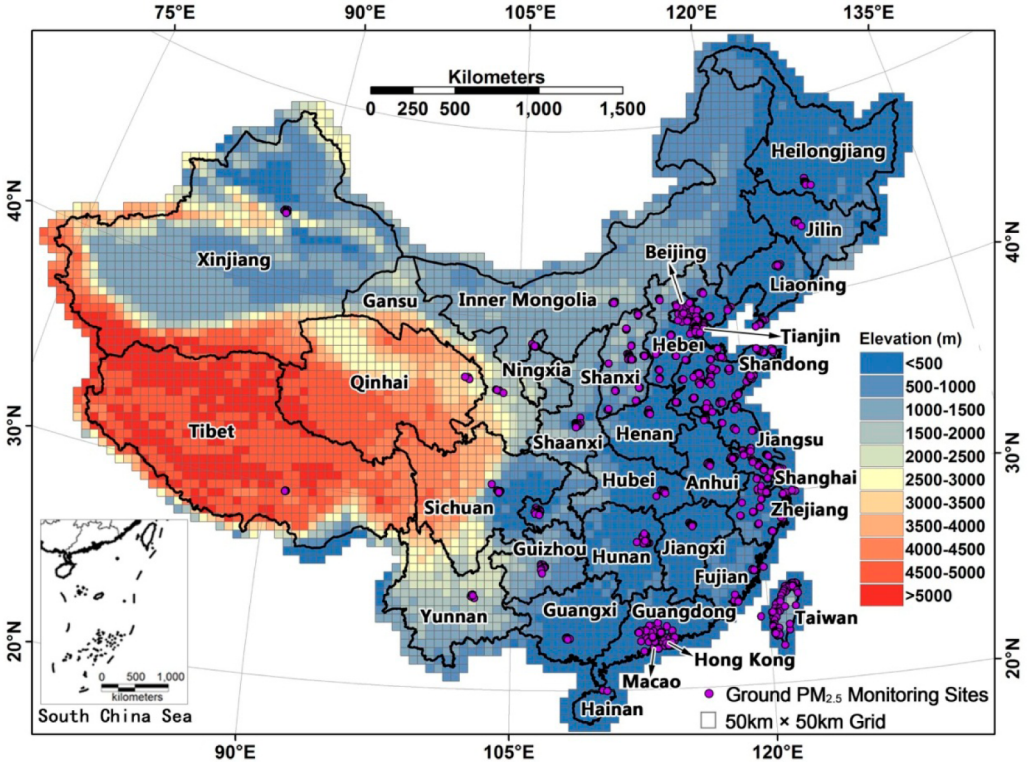


Figure 1. Spatial distribution of $PM_{2.5}$ monitoring sites involved in this study. There are 835 monitoring sites in 113 cities. It should be noted that the monitoring sites are clustering in the urban areas of major cities whereas rural areas have little coverage. Many monitoring sites are overlaid in this map.

from satellite-derived AOD[14] using the $PM_{2.5}$/AOD ratios derived from a global chemical transport model (CTM). This method does not require ground $PM_{2.5}$ data to develop the model. However, it is difficult to validate these results since sufficient ground measurements do not exist for most of the world, including China.

In recent years, many studies have established the relationship between $PM_{2.5}$ and satellite AOD using advanced statistical models (e.g., generalized linear regression, generalized additive, geographically weighted regression, and land use regression models).[15−18] These studies included meteorological parameters (e.g., boundary level height, temperature, relative humidity, and wind speed) and land use information (e.g., elevation, population, and vegetation cover) as covariates to improve model performance. Their results showed that meteorological fields and land use information can significantly improve the model performance.

Empirical statistical methods require sufficient ground $PM_{2.5}$ measurements to fit and validate. The development of this research area has been very limited in China, mainly due to the lack of extensive ground $PM_{2.5}$ measurements. To date, there were only a handful of studies conducted in some Chinese cities, including Beijing, Xi'an, and Zhengzhou.[19−22] Because these studies were conducted only in specific sites or cities, their models could not be easily generalized to estimate $PM_{2.5}$ elsewhere. Furthermore, few of these studies considered the impact of land use information. To our knowledge, no national-scale studies on such empirical modeling of $PM_{2.5}$ and satellite-derived AOD in conjunction with meteorological and land use information in China have been reported. China began the development of a nationwide $PM_{2.5}$ monitoring network in 2012. The capital cities of all provinces and the major cities in Hebei, Shandong, Shanxi, Jiangsu, Zhejiang, and Guangdong provinces finished the $PM_{2.5}$ monitoring sites construction and began to release $PM_{2.5}$ concentrations information in December 2012, which provides data for AOD and $PM_{2.5}$ modeling.

To date, advanced spatial statistical $PM_{2.5}$ exposure models have not been reported in China. As the first attempt, we present a national-scale geographically weighted regression (GWR) model for $PM_{2.5}$ estimation in China using satellite-derived AOD data, meteorological and land use information, and the newly released ground $PM_{2.5}$ monitoring data. The rest of the paper is organized such that various input data sets and the model structure are described in the "Materials and Methods" section. In the "Results and Discussion" section, we provide the model fitting and cross-validation (CV) results, discuss the model-predicted spatial and temporal variability of $PM_{2.5}$ in China, and we also compare our results with two previous studies.

MATERIALS AND METHODS

Ground $PM_{2.5}$ Measurements. Daily average $PM_{2.5}$ measurements from December 22, 2012 to November 30, 2013 were collected primarily from the official Web site of the China Environmental Monitoring Center (CEMC) (http://113.108.142.147:20035/emcpublish/). Some provinces (such as Shandong, Shanxi, Zhejiang, and Guangdong province) and municipalities (such as Beijing and Tianjin City) have established additional monitoring sites that are not included in the CEMC's Web site. Data from those additional sites were also collected. According to the Chinese National Ambient Air Quality Standard (CNAAQS, GB3095-2012, available on the Chinese Ministry of Environmental Protection (MEP) Web site http://kjs.mep.gov.cn/), the ground $PM_{2.5}$ data of China's

dx.doi.org/10.1021/es5009399 | Environ. Sci. Technol. 2014, 48, 7436−7444

mainland are measured by the tapered element oscillating microbalance method (TEOM) or the beta-attenuation method. In addition, ground $PM_{2.5}$ monitoring data of Macao, Hong Kong, and Taiwan were also collected from the Web sites of the Macao Meteorological and Geophysical Bureau (MCMGB) (http://www.smg.gov.mo), the Hong Kong Environmental Protection Department (HKEPD) (http://www.epd-asg.gov.hk), and the Taiwan Environmental Protection Administration (TWEPA) (http://taqm.epa.gov.tw), respectively. $PM_{2.5}$ data for Taiwan, Hong Kong, and Macao are also measured by TEOM or beta-attenuation. A total of 835 monitoring sites in 113 cities are included in the present work (Figure 1).

Satellite AOD Retrievals and Calibration. *MODIS Level 2 Aerosol Products.* The MODIS instruments scan a swath of 2330 km and have a global coverage of 1−2 days. Terra MODIS and Aqua MODIS cross the equator at approximately 10:30 a.m. and 1:30 p.m. local time, respectively. MODIS retrieves AOD at 10 km spatial resolution at nadir. Level 2 MODIS aerosol product data (Collection 5.1) from 2000 to November 30, 2013 (Terra) and 2002 to November 30, 2013 (Aqua) were downloaded from the Level 1 and Atmospheric Archive and Distribution System (LAADS Web; http://ladsweb.nascom.nasa.gov/). MODIS AOD data before 2013 were used for AOD calibration ("Satellite AOD Calibration" section). Data after December 22, 2012 were used for modeling. The AOD data (MODIS parameter name: Image_Optical_Depth_Land_And_Ocean) at 550 nm with Quality Assurance Confidence Flag = 2 and 3 were used.

MISR Level 2 Aerosol Products. MISR was launched aboard the Terra satellite in 1999. The spatial resolution of MISR AOD data is 17.6 km. MISR has a global coverage between 2 and 9 days depending on location and crosses the equator at approximately 10:30 a.m. local time. The MISR Level 2 aerosol data from 2000 to November 30, 2013 were downloaded from the NASA Langley Research Center Atmospheric Science Data Center (http://l0dup05.larc.nasa.gov/MISR/cgi-bin/MISR/main.cgi). Like the MODIS data, we used MISR AOD data before 2013 for AOD calibration and data after December 22, 2012 for modeling. The best-fit AOD data (MISR parameter name: RegBestFitSpectralOptDepth) at 558 nm, with the smallest chi-square fitting parameter from all aerosol mixtures, were extracted from the MISR data files.

Satellite AOD Calibration. To improve data coverage, Aqua MODIS, Terra MODIS, and MISR AOD data were fused together for modeling. MODIS and MISR have different value ranges and accuracies; MODIS allows small negative values, but MISR does not. Therefore, Terra MODIS, Aqua MODIS, and MISR AOD data need to be calibrated separately before they can be fused together. The AOD data from the Aerosol Robotic Network (AERONET) were used to calibrate the MODIS and MISR AOD data.[23] AERONET data have been widely used for validating the satellite-derived AOD due to their high accuracy.[24,25] AERONET Level 2 data (quality-assured) from 15 sites in China from 2000 to 2012 were downloaded from the AERONET Web site (http://aeronet.gsfc.nasa.gov/). Each site has operated for at least three seasons in the period of 2000 through 2012. To compare to the MODIS and MISR AOD values, AERONET AOD at 550 nm was interpolated from AERONET AOD at 440 and 675 nm using the Angstrom Exponent provided by AERONET.[26] A previous validation study showed that 50 km is the maximum appropriate radius window for satellite-/ground-based comparisons over land.[27] In this study, Terra MODIS, Aqua MODIS, and MISR AOD values within 25 km of the AERONET sites were averaged and compared with AERONET AOD values measured within the hour of the satellite overpass time, respectively. A simple linear regression between the AERONET and each satellite AOD data set was established each season (Supporting Information (SI), Text S1). Then, Terra MODIS, Aqua MODIS, and MISR data were calibrated using the established relationships, respectively.

Meteorological Data. Meteorological data were obtained from the Goddard Earth Observing System-Forward Processing (GEOS-FP).[28] GEOS-FP is the latest GEOS-5 meteorological data produced with version 5.11.0 of the GEOS Data Assimilation System (DAS).[29] GEOS-FP has a native horizontal resolution of 0.25° latitude × 0.3125° longitude and hourly or 3-hourly temporal resolution in a nested grid covering China. Corresponding to the satellite overpass time, the mean values of planetary boundary layer height above surface (PBLH, m), temperature at 2 m above displacement height (T2M, K), wind speed (WS, m/s) at 10 m above displacement height, mean relative humidity in PBLH (RH_PBLH, %), and surface pressure (PS, hPa) between 10 and 11 a.m. local time were extracted.

Land Cover and Population Data. The impacts of land cover and population on the relationship between the $PM_{2.5}$ and satellite AOD were also examined in the present study. The MODIS Level 3 monthly mean normalized difference vegetation index (NDVI)[30] products with a spatial resolution of 0.25° × 0.25° were downloaded from NASA Earth Observations (NEO) (http://neo.sci.gsfc.nasa.gov/).[30] The population data were obtained from the latest version LandScan, which provides 30-arc-second global population density.[31] (http://www.ornl.gov/sci/landscan/). The population and monthly mean NDVI were used as covariates in the $PM_{2.5}$ model.

Data Integration. For data integration, a 50 km-resolution grid (4553 grid cells in total) covering entire China was created with each grid cell roughly corresponding to the buffer size (25 km) when comparing satellite AOD with AERONET observations. Additional considerations when selecting this grid size are computational efficiency and improved spatial coverage from averaging over multiple satellite pixels. $PM_{2.5}$; calibrated Terra MODIS, Aqua MODIS, and MISR AOD; and meteorological, NDVI, and population data were averaged over the 50 km grid cells, respectively. The gridded satellite AOD values were fused into one AOD parameter in each grid cell as follows. Based on the satellite overpass times, the gridded Terra MODIS AOD can be merged with MISR AOD, but not with Aqua MODIS AOD. To address this issue, for each day, simple linear regression analysis was conducted between Terra MODIS and Aqua MODIS AOD data in those 50 km grid cells where both MODIS products are present (SI, Text S2). Then, the regression coefficients were applied to each grid cell to predict the missing Terra MODIS AOD using the Aqua MODIS AOD that is present.[32] For those grid cells where both MODIS AOD values are missing, we used MISR AOD data instead. Combining MODIS and MISR AOD has been shown to be a useful way to improve the AOD coverage.[14] Unlike van Donkelaar et al.,[14] we used both Terra and Aqua MODIS AOD. Our results show that the data coverage of AOD was substantially improved after the fusion of MODIS and MISR AOD, especially in Western and Northwestern China (SI, Text S2). Finally, $PM_{2.5}$, fused AOD, meteorological parameters,

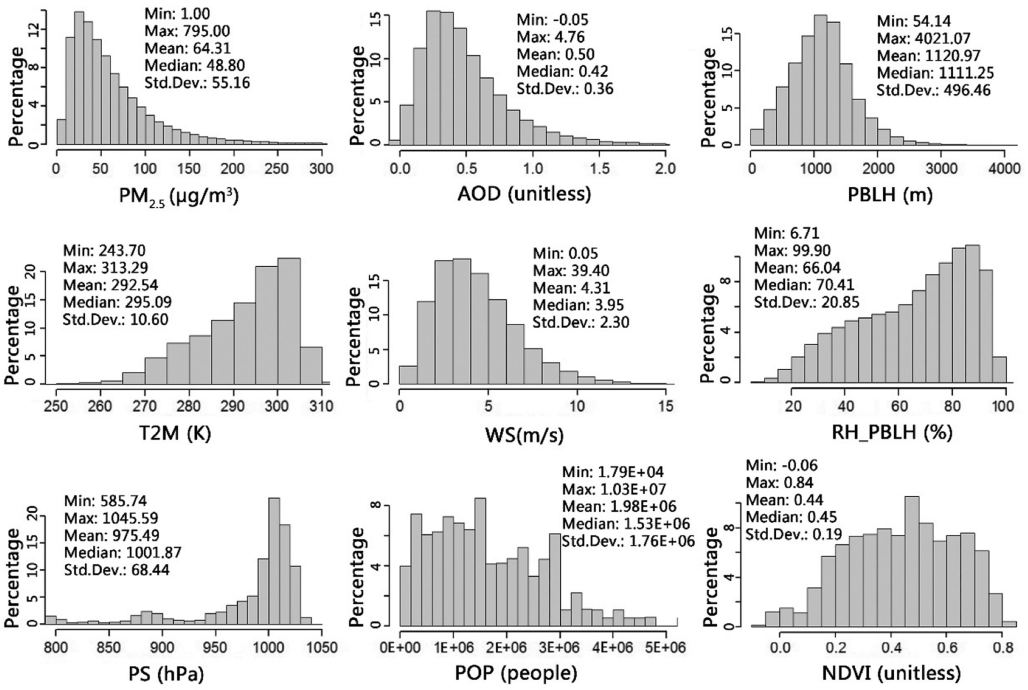


Figure 2. Histograms and descriptive statistics of the variables in the whole model fitting data set. (N = 58 164).

NDVI, and population data for all days and all grid cells were matched by grid cell ID and Day of Year (DOY) for model development.

Model Development and Validation. Since our study area is large, the relationship between $PM_{2.5}$ and AOD will likely vary in space. To address the spatial heterogeneity of the $PM_{2.5}$-AOD relationship, we developed a GWR model.[33] Unlike other statistical models that estimate global parameters, GWR can generate a continuous surface of parameter values by taking measurements of the parameters at each local observation to denote the spatial variations of the surface. GWR has been successfully applied in modeling the $PM_{2.5}$-AOD association in the Southeastern U.S.[17] In this study, a separate GWR model was fitted for each day, with the general structure as follows:

$$\begin{aligned} PM_{2.5st} = {} & \beta_{0,st} + \beta_{1,st}AOD_{st} + \beta_{2,st}PBLH_{st} + \beta_{3,st}T2M_{st} \\ & + \beta_{4,st}WS_{st} + \beta_{5,st}RH\_PBLH_{st} + \beta_{6,st}PS_{st} \\ & + \beta_{7,st}POP_{s} + \beta_{8,st}NDVI_{sm} + \varepsilon_{st} \end{aligned} \quad (1)$$

where $PM_{2.5\ st}$ ($\mu g/m^3$) is the daily ground-level $PM_{2.5}$ concentration at grid cell s on day t; $\beta_{0,st}$ denotes the location-specific intercept on day t; $\beta_{1,st}$–$\beta_{8,st}$ are the location-specific slopes on day t; AOD_{st} (unitless) is the AOD fused from Terra MODIS, Aqua MODIS, and MISR AOD products at grid cell s on day t; $PBLH_{st}$, $T2M_{st}$, WS_{st}, RH_PBLH_{st}, PS_{st} are meteorological parameters at grid cell s on day t (definitions in "Meteorological Data" section); POP_s is the total population at grid cell s; $NDVI_{sm}$ (unitless) is the MODIS NDVI value at grid cell s in month m; and ε_{st} is the error term at grid cell s on day t. The selection of the meteorological and land use variables in the final model is based on findings of previous studies, statistical significance of the variables and the GWR model performance measured by model R^2. Detailed comparisons of the full model with various reduced models are provided in the SI (Text S3).

A fixed bandwidth of 800 km was used for fitting the GWR. We also examined the adaptive bandwidth option, and our results show that the adaptive bandwidth can cause model overfitting problems (SI, Text S4). For some days, the number of the matched data points is limited, which might impact the model significantly. To address this issue, we interpolated AOD values with an Ordinary Kriging[34] approach to increase the sample size of the daily model fitting data set. There are 58 164 matched grid-cell days when using interpolated AOD values, compared to 24 521 matched grid-cell days without interpolated AOD values. GWR model fitting used the interpolated data set because it helps to stabilize model performance in areas with few $PM_{2.5}$ sites and the days with low AOD coverage (e.g., Western China and the winter season) (SI, Text S5). AOD interpolation was not used in model prediction. To evaluate how much the meteorological and land use parameters used in our final model could improve the model performance, we also fitted a GWR model using AOD as the only independent variable (the AOD-only model). We also fitted a model only using meteorological and land used data without AOD (the non-AOD model) to see how AOD can benefit the model performance.

Cross validation (CV) technique was applied to test for potential model overfitting, that is, the model could have better predictive performance in the data set used in model fitting than the data from the rest of the study area. We used a 10-fold CV method[35] in the present study. Furthermore, mean prediction error (MPE) and root mean squared prediction error (RMSE) were adopted to evaluate the model prediction accuracy for model fitting and CV results. We calculated Pearson's correlation coefficients among all the independent variables. The results show that the absolute values of correlation coefficients between any two of the independent variables are relatively low (r < 0.5), indicating relatively weak collinearity problems in the current model.

dx.doi.org/10.1021/es5009399 | Environ. Sci. Technol. 2014, 48, 7436–7444

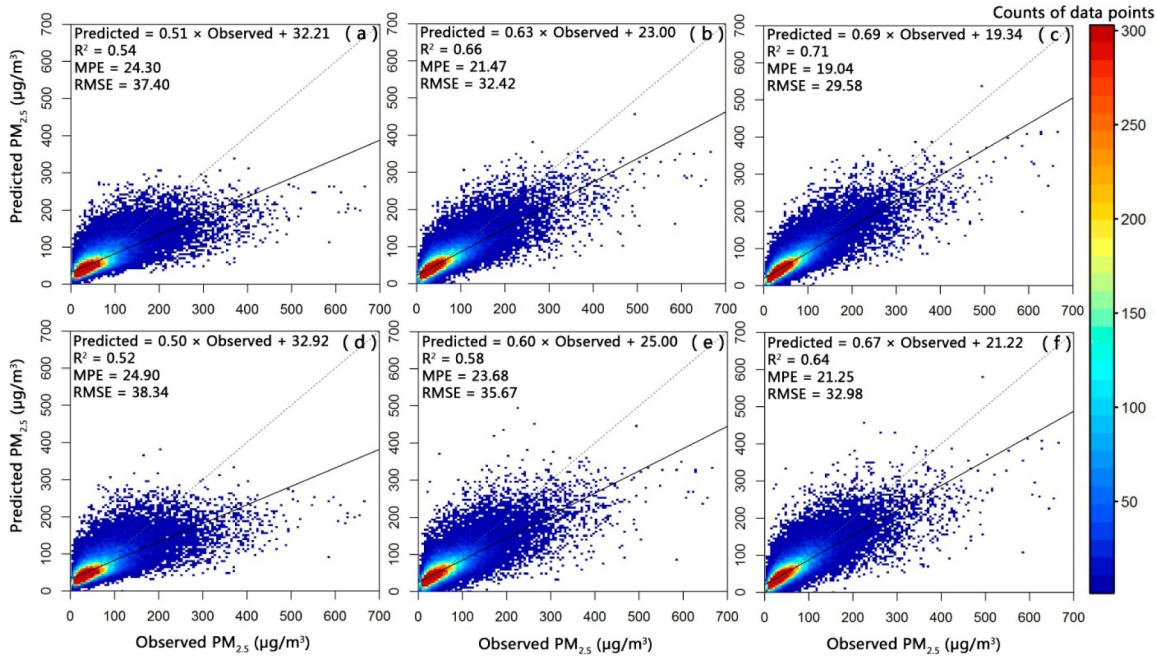


Figure 3. Results of model fitting and cross validation. MPE: mean prediction error ($\mu g/m^3$). RMSE: root mean squared prediction error ($\mu g/m^3$). The dash line is the 1:1 line as a reference. (a)−(c) are model fitting results of the AOD-only, non-AOD, and full model, respectively. (d)−(f) are CV results of the AOD-only, non-AOD, and full model, respectively.

RESULTS AND DISCUSSION

Descriptive Statistics. Figure 2 shows the histograms and descriptive statistics of the model variables in the whole model fitting data set. Overall, the mean $PM_{2.5}$ concentration for the entire domain is 64.31 $\mu g/m^3$ and the mean fused AOD is 0.50.

The values of the dependent and independent variables exhibit strong seasonality (SI, Table S3). The mean AOD value is highest in spring (0.59) and lowest in winter and summer (both 0.45).The highest mean $PM_{2.5}$ concentration is in winter (101.65 $\mu g/m^3$) and the lowest is in summer (44.28 $\mu g/m^3$). The different seasonal patterns of $PM_{2.5}$ and AOD are mainly due to their complex relationship, which can be strongly influenced by meteorological, geographical, and seasonal conditions.[36] For example, the AOD-$PM_{2.5}$ relationship can be significantly improved after calibrating AOD with relative humidity (RH) and PBLH.[37] Since RH and PBLH are temporally different, the AOD-$PM_{2.5}$ relationship also varies.

The spatial distribution of seasonal and annual mean fused AOD values and ground measured $PM_{2.5}$ concentrations is show in Figure S6 (SI). High AOD and $PM_{2.5}$ values generally occur in the North China Plain (including Beijing, Tianjin Metropolitan Area, and Hebei, Shandong, Henan, Anhui, and Jiangsu Provinces), the Tarim Basin, and the Sichuan Basin.

Model Fitting and Validation. Figure 3 shows the scatter plots for the model fitting and cross validation of the AOD-only, non-AOD, and full models. For the model fittings, the R^2 values are 0.54 and 0.71 for the AOD-only and full models, respectively. The RMSE is 37.40 $\mu g/m^3$ for the AOD-only model and 29.58 $\mu g/m^3$ for the full model. Compared to the model fitting, the CV R^2 only decreases by 0.02 and CV RMSE only increases by 0.94 $\mu g/m^3$ for the AOD-only model. However, the CV R^2 decreases by 0.07 and CV RMSE increases by 3.40 $\mu g/m^3$ for the full model, which are greater than the model fitting and suggest that our full model is slightly overfitted. However, the CV R^2 of the full model (0.64) is much greater than that of the AOD-only model (0.52), and the CV RMSE of the full model (32.98 $\mu g/m^3$) is significantly lower than that of the AOD-only model (38.34 $\mu g/m^3$). Our findings suggest that despite a slight overfitting in the full model, the overall prediction accuracy is significantly improved when considering the meteorological data and land use information in the GWR model. For the non-AOD model, the model fitting R^2 and RMSE are 0.66 and 32.42 $\mu g/m^3$, respectively; and CV R^2 and RMSE are 0.58 and 35.67 $\mu g/m^3$, respectively. Comparison of the full model and non-AOD model shows that AOD can improve the model performance without causing more overfitting.

Our results show that the CV RMSE of the full model is 32.98 $\mu g/m^3$, which is much higher than results of studies in U.S. (~5 $\mu g/m^3$).[36,38] However, it should be noted that the $PM_{2.5}$ concentrations in these studies are relatively low (mainly <40 $\mu g/m^3$ and rarely >60 $\mu g/m^3$). In this study, $PM_{2.5}$ concentrations range from 1 $\mu g/m^3$ to 795 $\mu g/m^3$. The CV relative prediction error (RPE, defined as RMSE divided by the mean ground $PM_{2.5}$) of the full model is 51.3%, which is approximately 15% greater than the previous study in the U.S.[36] One possible reason is that our models tend to underestimate when ground $PM_{2.5}$ concentrations are greater than ~60 $\mu g/m^3$ (Figure 3). Similar results were also found in the Eastern U.S.,[15] where $PM_{2.5}$ was substantially underestimated at higher concentrations (>40 $\mu g/m^3$), and a sublinear relationship between $PM_{2.5}$ and AOD at high particle loadings was posited.[9] Another possible reason is that these high $PM_{2.5}$ levels might be very local, therefore cannot be well represented in a large grid cell. Given our coarse spatial resolution, this is very likely the case. When we excluded the data records with ground $PM_{2.5}$ values higher than 100 $\mu g/m^3$, and the RPE dropped to 37.3%. Our results also showed that the CV slope of the full model (0.67) is greater than that of the AOD-only model (0.50); that is, the extent of the underestimation is lower in the full model compared to the AOD-only model.

dx.doi.org/10.1021/es5009399 | Environ. Sci. Technol. 2014, 48, 7436−7444

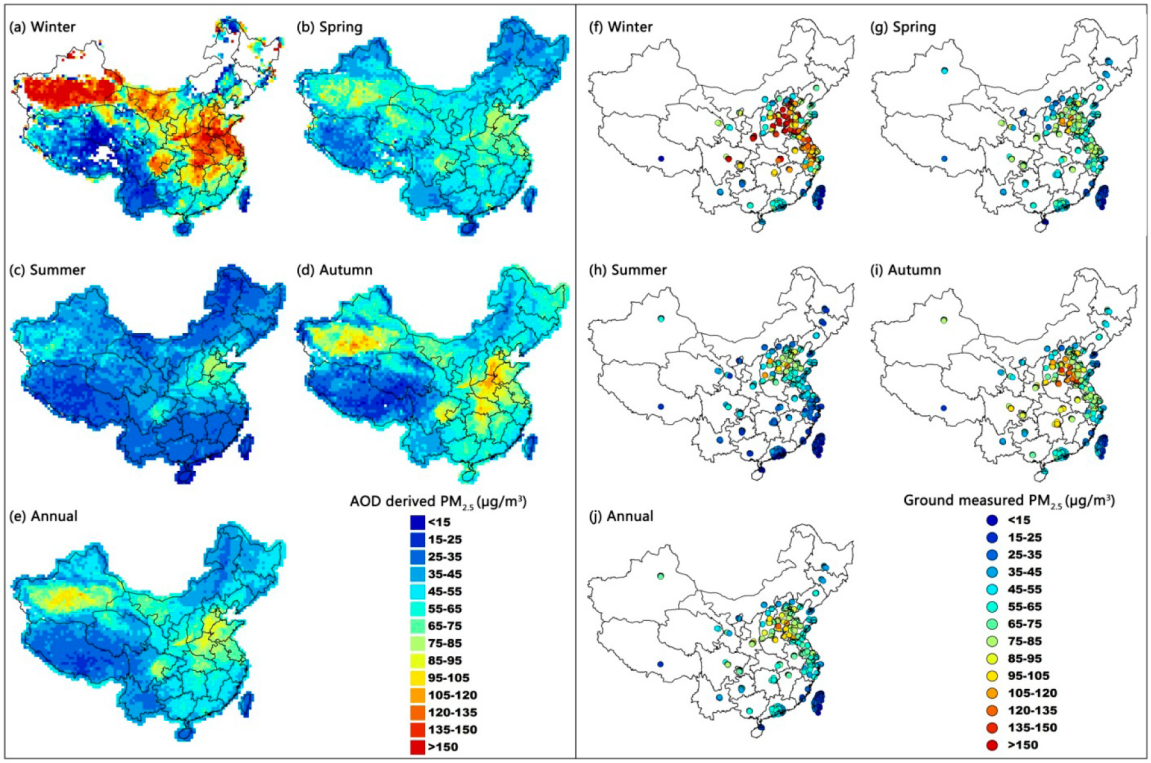


Figure 4. Comparisons of seasonal and annual mean AOD-derived $PM_{2.5}$ and ground-measured $PM_{2.5}$ concentrations from those days when corresponding gridded AOD values are available.

The local R^2 is an indicator that reflects the model performance and the spatial heterogeneity of the GWR. For the full model, the mean local R^2 values of the daily GWR models range from 0.29 to 0.90, with an overall mean of 0.56, indicating the overall performance is relatively high. For the AOD-only model, the daily mean local R^2 values range from 0.06 to 0.72, with an overall mean of 0.34. This is a substantial decrease. These findings indicate that meteorological data and land use information can significantly improve model performance and agree with previous studies.[16,17] Examining the spatial variability of mean local R^2 of the full model indicates that the local R^2 values in South China with lower $PM_{2.5}$ concentrations (<75 $\mu g/m^3$) are higher (i.e., >0.56) than those in North China with higher $PM_{2.5}$ concentrations (>75 $\mu g/m^3$) (SI, Figure S7). This finding suggests that a more flexible nonlinear model structure might better represent the sublinearity between $PM_{2.5}$ and AOD in a wide dynamic range.

Prediction Maps of $PM_{2.5}$ Concentrations. The predicted seasonal mean and annual mean $PM_{2.5}$ concentrations are shown in Figure 4. To examine the spatially and seasonal prediction accuracy of the GWR model, the ground $PM_{2.5}$ measurements which are covered with gridded AOD values are also shown in Figure 4.

As shown in Figure 4, the temporal and spatial patterns of AOD-derived $PM_{2.5}$ are very similar to ground measurements. The MPE of the annual mean $PM_{2.5}$ concentration is 8.28 $\mu g/m^3$. Winter is the most polluted season, with the highest MPE of 25.33 $\mu g/m^3$; summer is the cleanest season, with the lowest MPE of 7.25 $\mu g/m^3$; and the MPE values of spring and autumn are 9.04 $\mu g/m^3$ and 11.18 $\mu g/m^3$, respectively. Winter heating in North China, Northeast China, and Northwest China is considered a major source of $PM_{2.5}$.[39−41] When comparing the AOD-derived $PM_{2.5}$ to the ground $PM_{2.5}$ concentrations from all available days (SI, Figure S6), the MPE of annual, winter, spring, summer, and autumn mean $PM_{2.5}$ are 13.64 $\mu g/m^3$, 32.28 $\mu g/m^3$, 12.99 $\mu g/m^3$, 8.05 $\mu g/m^3$, and 13.29 $\mu g/m^3$, respectively.

Spatially, the North China Plain, where the annual mean $PM_{2.5}$ estimates generally range from 85 to 95 $\mu g/m^3$, is one of the most industrialized and populated regions in China. Rapid economic development and urbanization have led to severe $PM_{2.5}$ pollution.[42,43] The Tarim Basin is another area of high $PM_{2.5}$ and is located in the southern part of the Xinjiang Autonomous Region. The annual mean $PM_{2.5}$ in the Tarim Basin are on average above 85 $\mu g/m^3$ and above 95 $\mu g/m^3$ at its center. The Taklamakan Desert, the largest desert in China, occupies over 60% of the Basin. In addition, a band of high pollution levels spans the Tarim Basin, Gansu, Ningxia, Shaanxi, Shanxi Provinces, and western Inner Mongolia. The Gobi Desert covers part of this region. The Taklamakan and Gobi Deserts have been considered the two major sources of dust storms in North China.[44,45] Dust storms usually occur in winter and spring, and contribute significantly to regional $PM_{2.5}$ levels in Northern China.[46−48] The Sichuan Basin, which is located in the eastern part of Sichuan Province and includes Chongqing Metropolitan Area, is another area of high pollution, with the annual $PM_{2.5}$ of 75−95 $\mu g/m^3$. Surrounded by high mountain ranges, this region suffers from persistent temperature inversion and stagnant air circulation,[49] which often lead to severe air pollution episodes. The cleanest areas are in Taiwan, Hainan, and Tibet, where the annual mean values are generally lower than 35 $\mu g/m^3$. North Inner Mongolia, West Heilongjiang, and Yunnan Province have annual mean values of 25−45 $\mu g/m^3$.

A potential source of prediction error of the GWR model is the uneven spatial distribution of ground $PM_{2.5}$ monitoring

dx.doi.org/10.1021/es5009399 | Environ. Sci. Technol. 2014, 48, 7436−7444

sites. Similar to the regulatory monitoring networks elsewhere, the Chinese network primarily covers large urban centers with very limited coverage in rural areas, especially in western part of the country, including Tibet, Qinghai, and Xinjiang. Although the GWR model developed in this work has similar model performance compared with models developed in the U.S.,[17] the absolute prediction error is much higher. Consequently, the estimated daily $PM_{2.5}$ levels in the cleaner areas of China, such as western Tibet, may contain substantial errors, especially in winter. This problem is expected to be alleviated in the near future, as the national $PM_{2.5}$ monitoring network will cover all county-level cities by 2015, according to the MEP of China [3]. This research will be helpful for selecting $PM_{2.5}$ monitoring site locations in the future, which should consider placing more sites in rural areas, especially in Western China.

According to CNAAQS, the Level 1 annual $PM_{2.5}$ standard of 15 $\mu g/m^3$ is assigned for protecting the air quality of natural reserves and scenic areas. The Level 2 annual $PM_{2.5}$ standard of 35 $\mu g/m^3$ is designated for residential, cultural, industrial, and commercial areas. The Level 1 and Level 2 concentrations are equivalent to the World Health Organization (WHO) Air Quality Interim Target-3 (IT-3) and Target-1 (IT-1)[50] levels, respectively. Our predicted annual $PM_{2.5}$ concentrations indicated that 100% and 81.8% of China exceed the CNAAQS Level 1 (or WHO IT-3) and Level 2 (or WHO IT-1) annual standards, respectively. When considering the population distribution, over 96% of the Chinese population lives in areas that exceed the CNAAQS Level 2 standard.

Comparison with Previous Studies. We compared our results with two publically available data sets of satellite-derived $PM_{2.5}$ concentrations. Based on the color scale in Figure 4, we replotted the mean estimated $PM_{2.5}$ concentration of 2001−2006 in China from the study of van Donkelaar et al.[14] at 0.1 degree spatial resolution and the mean estimated $PM_{2.5}$ concentration of 2010 from the Center for International Earth Science Information Network (CIESIN) at Columbia University[51] at 0.5 degree resolution (SI, Figure S8). Note that the CIESIN data set was generated by simply applying the AOD/$PM_{2.5}$ surface level conversion factors for 2001−2006 calculated by van Donkelaar et al.[14] to monthly AOD data from MODIS and MISR.[51]

While the GWR-estimated $PM_{2.5}$ concentrations in 2013 show similar spatial patterns to previous results, there are several important differences. The GWR-estimated $PM_{2.5}$ concentrations in 2013 are substantially higher than the CIESIN estimates for 2010. For example, our estimated mean concentrations in much of the North China Plain are around 85−95 $\mu g/m^3$ compared to 55−65 $\mu g/m^3$ in the CIESIN data set. In addition, the high $PM_{2.5}$ region in the Tarim basin is not apparent in the CIESIN map. Our results are more like those of van Donkelaar et al.,[14] but have higher $PM_{2.5}$ levels in rural regions and slightly lower levels in the Sichuan Basin. For example, van Donkelaar, et al.[14] estimated that $PM_{2.5}$ concentrations in Northeastern China are generally less than 25 $\mu g/m^3$, except along the Bo Hai Coast. Our GWR model, however, shows large regions with $PM_{2.5}$ concentrations around 35−55 $\mu g/m^3$ in Northeastern China. Our model also predicts higher $PM_{2.5}$ concentrations in the dust source region, including the Tarim Basin and the Gobi Desert.

A statistically significant increasing trend of regional mean AOD has been reported for Southern and Eastern China between 1998 and 2010, primarily due to rapid economic development.[52] This could partially explain the discrepancies among the three data sets. Another important difference between the GWR model and the other two methods is that the GWR model is calibrated against the ground observations in its development process. The development of such a model would have not been possible without the newly released ground $PM_{2.5}$ observations in China. However, the flexible model structure ensures higher prediction accuracy when compared with the other two methods. Comparing to all available ground observations (SI, Figure S6-J), the MPE of annual $PM_{2.5}$ are 13.64 $\mu g/m^3$, 17.78 $\mu g/m^3$, and 32.08 $\mu g/m^3$ for our GWR model, van Donkelaar's data set, and CIESIN data set, respectively.

In summary, despite the increasingly serious particle pollution in China, detailed characterization of the spatial pattern of $PM_{2.5}$ has been lacking. Satellite-derived $PM_{2.5}$ concentrations have been made possible recently but such estimates have not been evaluated against ground truth. The GWR model described in this work is the first attempt to develop advanced spatial models for $PM_{2.5}$ exposure assessment in China. This model exhibits satisfactory performance with a CV R^2, RMSE, and RPE are 0.64, 32.98 $\mu g/m^3$, and 51.3%, respectively. When extremely high $PM_{2.5}$ levels (>100 $\mu g/m^3$) were excluded, model performance was further improved to be comparable with previous studies conducted in the U.S. The inclusion of meteorological and land use information can significantly improve the model performance. It is certainly feasible to develop models with spatial resolutions compatible with the raw satellite AOD pixel size, which is a direction of our future research.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Texts S1−S5, Table S1−S3, and Figures S1−S8. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Authors

*(Y.L.) Phone: +1-404-727-2131; fax: +1-404-727-8744; e-mail: yang.liu@emory.edu.

*(J.B.) Phone: +86-25-89680566; fax: +86-25-89680586; e-mail: jbi@nju.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was partially supported by NASA Applied Sciences Program (grant no. NNX11AI53G, PI: Yang Liu). In addition, this publication was made possible by USEPA grant R834799. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication. The work of Zongwei Ma was supported by the China Scholarship Council (CSC) under the State Scholarship Fund.

■ REFERENCES

(1) Pope, C. A.; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; Thurston, G. D. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA, J. Am. Med. Assoc.* **2002**, *287* (9), 1132−1141.

(2) Dominici, F.; Peng, R. D.; Bell, M. L.; Pham, L.; McDermott, A.; Zeger, S. L.; Samet, J. M. Fine particulate air pollution and hospital

admission for cardiovascular and respiratory diseases. *JAMA, J. Am. Med. Assoc.* **2006**, *295* (10), 1127−1134.
(3) Yuan, Y.; Liu, S.; Castro, R.; Pan, X. $PM_{2.5}$ monitoring and mitigation in the cities of China. *Environ. Sci. Technol.* **2012**, *46* (7), 3627−3628.
(4) Chang, Y. China needs a tighter $PM_{2.5}$ limit and a change in priorities. *Environ. Sci. Technol.* **2012**, *46* (13), 7069−7070.
(5) Chen, R.; Zhao, Z.; Kan, H. Heavy smog and hospital visits in Beijing, China. *Am. J. Respir. Crit. Care Med.* **2013**, *188* (9), 1170−1171.
(6) Huang, W.; Cao, J.; Tao, Y.; Dai, L.; Lu, S.-E.; Hou, B.; Wang, Z.; Zhu, T. Seasonal variation of chemical species associated with short-term mortality effects of $PM_{2.5}$ in Xi'an, a Central City in China. *Am. J. Epidemiol.* **2012**, *175* (6), 556−566.
(7) Chen, Y. Pollution characteristic of $PM_{2.5}$ in urban Guangzhou. *Environ. Protect. Sci.* **2010**, *36* (3), 7−11 (in Chinese with English abstract).
(8) Wei, Y.; Yin, Y.; Yang, W.; Rui, D.; Hang, W. Analysis of the polution characteristics & influence factors of $PM_{2.5}$ in Nanjing Area. *Environ. Sci. Manage.* **2009**, *34* (9), 29−34 (in Chinese with English abstract).
(9) Hoff, R. M.; Christopher, S. A. Remote sensing of particulate pollution from space: Have we reached the promised land? *J. Air Waste Manage. Assoc.* **2009**, *59* (6), 645−675.
(10) Remer, L. A.; Kaufman, Y.; Tanré, D.; Mattoo, S.; Chu, D.; Martins, J.; Li, R.-R.; Ichoku, C.; Levy, R.; Kleidman, R. The MODIS aerosol algorithm, products, and validation. *J. Atmos. Sci.* **2005**, *62* (4), 947−973.
(11) Kahn, R.; Banerjee, P.; McDonald, D.; Diner, D. J. Sensitivity of multiangle imaging to aerosol optical depth, and to pure-particle size distribution and composition over ocean. *J. Geophys. Res. Atmos.* **1998**, *103*, 32195−32213.
(12) Wang, J.; Christopher, S. A. Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies. *Geophys. Res. Lett.* **2003**, *30* (21), 2095.
(13) Liu, Y.; Park, R. J.; Jacob, D. J.; Li, Q.; Kilaru, V.; Sarnat, J. A. Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *J. Geophys. Res., Atmos.* **2004**, *109* (D22).
(14) van Donkelaar, A.; Martin, R. V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P. J. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* **2010**, *118* (6), 847.
(15) Liu, Y.; Sarnat, J. A.; Kilaru, V.; Jacob, D. J.; Koutrakis, P. Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing. *Environ. Sci. Technol.* **2005**, *39* (9), 3269−3278.
(16) Liu, Y.; Paciorek, C. J.; Koutrakis, P. Estimating regional spatial and temporal variability of $PM_{2.5}$ concentrations using satellite data, meteorology, and land use information. *Environ. Health Perspect.* **2009**, *117* (6), 886−892.
(17) Hu, X.; Waller, L. A.; Al-Hamdan, M. Z.; Crosson, W. L.; Estes, M. G., Jr; Estes, S. M.; Quattrochi, D. A.; Sarnat, J. A.; Liu, Y. Estimating ground-level $PM_{2.5}$ concentrations in the southeastern US using geographically weighted regression. *Environ. Res.* **2013**, *121*, 1−10.
(18) Kloog, I.; Nordio, F.; Coull, B. A.; Schwartz, J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the Mid-Atlantic states. *Environ. Sci. Technol.* **2012**, *46* (21), 11913−11921.
(19) Guo, J.; Zhang, X.; Che, H.; Gong, S.; An, X.; Cao, C.; Guang, J.; Zhang, H.; Wang, Y.; Zhang, X. Correlation between PM concentrations and aerosol optical depth in eastern China. *Atmos. Environ.* **2009**, *43* (37), 5876−5886.
(20) Liu, Y.; He, K.; Li, S.; Wang, Z.; Christiani, D. C.; Koutrakis, P. A statistical model to evaluate the effectiveness of $PM_{2.5}$ emissions control during the Beijing 2008 Olympic Games. *Environ. Int.* **2012**, *44* (0), 100−105.
(21) Wang, J.; Yang, F.; Wang, D.; He, K. Characteristics and relationship of aerosol optical thickness and $PM_{2.5}$ concentration over Beijing. *J. Grad. Sch. Chin. Acad. Sci.* **2010**, *27* (1), 10−16 (in Chinese with English abstract).
(22) Guo, J.; Wu, Y.; Zhang, X.; Li, X. Estimation of $PM_{2.5}$ over Eastern China from MODIS aerosol optical depth using the back propagation neural network. *Environ. Sci.* **2013**, *34* (3), 817−825 (in Chinese with English abstract).
(23) Holben, B.; Eck, T.; Slutsker, I.; Tanre, D.; Buis, J.; Setzer, A.; Vermote, E.; Reagan, J.; Kaufman, Y.; Nakajima, T. AERONET-A federated instrument network and data archive for aerosol characterization. *Remote. Sens. Environ.* **1998**, *66* (1), 1−16.
(24) Jiang, X.; Liu, Y.; Yu, B.; Jiang, M. Comparison of MISR aerosol optical thickness with AERONET measurements in Beijing metropolitan area. *Remote. Sens. Environ.* **2007**, *107* (1), 45−53.
(25) Kahn, R.; Gaitley, B.; Martonchik, J.; Diner, D.; Crean, K.; Holben, B. Multiangle imaging spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations. *J. Geophys. Res.* **2005**, *110*, D10S04.
(26) Liu, Y.; Sarnat, J. A.; Coull, B. A.; Koutrakis, P.; Jacob, D. J. Validation of multiangle imaging spectroradiometer (MISR) aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the contiguous United States. *J. Geophys. Res., Atmos.* **2004**, *109* (D6), D06205.
(27) Breon, F. M.; Vermeulen, A.; Descloitres, J. An evaluation of satellite aerosol products against sunphotometer measurements. *Remote. Sens. Environ.* **2011**, *115* (12), 3102−3111.
(28) Lucchesi, R. File Specification for GEOS-5 FP. GMAO Office Note No. 4 (Version 1.0), 2013. http://gmao.gsfc.nasa.gov/pubs/office_notes.
(29) Rienecker, M. M.; Suarez, M. J.; Todling, R.; Bacmeister, J.; Takacs, L.; Liu, H.; Gu, W.; Sienkiewicz, M.; Koster, R.; Gelaro, R.; Nielsen, E. The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. In *Technical Report Series on Global Modeling and Data Assimilation 104606*, 2008; Vol. *27*
(30) Huete, A.; Justice, C.; Van Leeuwen, W. *MODIS vegetation index (MOD 13). Version 3. Algorithm theoretical basis document*, 1999. http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf.
(31) Bright, E. A.; Coleman, P. R.; Rose, A. N.; Urban, M. L. *LandScan 2011*; Oak Ridge National Laboratory: Oak Ridge, TN, 2012.
(32) Puttaswamy, S. J.; Nguyen, H. M.; Braverman, A.; Hu, X.; Liu, Y. Statistical data fusion of multi-sensor AOD over the Continental United States. *Geocarto Int.* **2013**, 1−17.
(33) Brunsdon, C.; Fotheringham, A. S.; Charlton, M. E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28* (4), 281−298.
(34) Cressie, N. Spatial prediction and ordinary kriging. *Math. Geol.* **1988**, *20* (4), 405−421.
(35) Rodriguez, J. D.; Perez, A.; Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32* (3), 569−575.
(36) Liu, Y.; Franklin, M.; Kahn, R.; Koutrakis, P. Using aerosol optical thickness to predict ground-level $PM_{2.5}$ concentrations in the St. Louis area: A comparison between MISR and MODIS. *Remote. Sens. Environ.* **2007**, *107* (1), 33−44.
(37) Paciorek, C. J.; Liu, Y.; Moreno-Macias, H.; Kondragunta, S. Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level $PM_{2.5}$. *Environ. Sci. Technol.* **2008**, *42*, 5800−5806.
(38) Hu, X.; Waller, L. A.; Lyapustin, A.; Wang, Y.; Al-Hamdan, M. Z.; Crosson, W. L.; Estes, M. G., Jr; Estes, S. M.; Quattrochi, D. A.; Puttaswamy, S. J. Estimating ground-level $PM_{2.5}$ concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote. Sens. Environ.* **2014**, *140*, 220−232.
(39) Zhu, S. J.; Dong, W.; Xu, J. Y. Study on pollution characteristics of $PM_{2.5}$: Tracing and tracking atmospheric particulates in Urumqi

City. *Environ. Protect. Xinjiang* **2012**, *34* (3), 6−11 (in Chinese with English Abstract).
(40) Han, B.; Kong, S.; Bai, Z.; Du, G.; Bi, T.; Li, X.; Shi, G.; Hu, Y. Characterization of elemental species in $PM_{2.5}$ samples collected in four cities of Northeast China. *Water, Air, Soil Pollut.* **2010**, *209* (1−4), 15−28.
(41) Yu, Y.; Schleicher, N.; Norra, S.; Fricker, M.; Dietze, V.; Kaminski, U.; Cen, K.; Stuben, D. Dynamics and origin of $PM_{2.5}$ during a three-year sampling period in Beijing, China. *J. Environ. Monit.* **2011**, *13* (2), 334−346.
(42) Quan, J.; Zhang, Q.; He, H.; Liu, J.; Huang, M.; Jin, H. Analysis of the formation of fog and haze in North China Plain (NCP). *Atmos. Chem. Phys.* **2011**, *11* (15), 8205−8214.
(43) Tao, M.; Chen, L.; Su, L.; Tao, J., Satellite observation of regional haze pollution over the North China Plain. *J. Geophys. Res., Atmos.* **2012**, *117* (D12).
(44) Sun, J.; Zhang, M.; Liu, T. Spatial and temporal characteristics of dust storms in China and its surrounding regions, 1960−1999: Relations to source area and climate. *J. Geophys. Res.* **2001**, *106* (D10), 10325−10333.
(45) Wang, X.; Dong, Z.; Zhang, J.; Liu, L. Modern dust storms in China: An overview. *J. Arid Environ.* **2004**, *58* (4), 559−574.
(46) Sun, Y.; Zhuang, G.; Huang, K.; Li, J.; Wang, Q.; Wang, Y.; Lin, Y.; Fu, J. S.; Zhang, W.; Tang, A.; Zhao, X. Asian dust over northern China and its impact on the downstream aerosol chemistry in 2004. *J. Geophys. Res., Atmos.* **2010**, *115* (D7), D00K09.
(47) Song, Y.; Zhang, Y.; Xie, S.; Zeng, L.; Zheng, M.; Salmon, L. G.; Shao, M.; Slanina, S. Source apportionment of $PM_{2.5}$ in Beijing by positive matrix factorization. *Atmos. Environ.* **2006**, *40* (8), 1526−1537.
(48) Zheng, M.; Salmon, L. G.; Schauer, J. J.; Zeng, L.; Kiang, C. S.; Zhang, Y.; Cass, G. R. Seasonal trends in $PM_{2.5}$ source contributions in Beijing, China. *Atmos. Environ.* **2005**, *39* (22), 3967−3976.
(49) Qian, Y.; Giorgi, F. Regional climatic effects of anthropogenic aerosols? The case of Southwestern China. *Geophys. Res. Lett.* **2000**, *27* (21), 3521−3524.
(50) WHO, World Health Organization. *Air Quality Guidelines Global Update 2005*; WHO/Europe.: Copenhagen, Denmark, 2005.
(51) BMI&CIESIN (Battelle Memorial Institute & Center for International Earth Science Information Network—Columbia University). *Global Annual Average $PM_{2.5}$ Grids from MODIS and MISR Aerosol Optical Depth (AOD)*; NASA Socioeconomic Data and Applications Center (SEDAC), 2013.
(52) Hsu, N. C.; Gautam, R.; Sayer, A. M.; Bettenhausen, C.; Li, C.; Jeong, M. J.; Tsay, S. C.; Holben, B. N. Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010. *Atmos. Chem. Phys.* **2012**, *12* (17), 8037−8053.