Environmental Science & Technology

Article

pubs.acs.org/est

Daily Estimation of Ground-Level $PM_{2.5}$ Concentrations over Beijing Using 3 km Resolution MODIS AOD

Yuanyu Xie,[†] Yuxuan Wang,*[,†,‡,§] Kai Zhang,[∥] Wenhao Dong,[†] Baolei Lv,[†] and Yuqi Bai[†]

[†]Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China

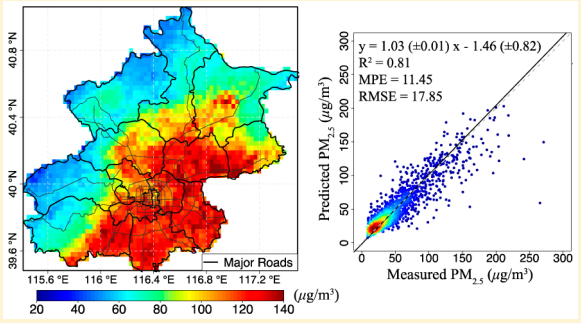
[‡]Department of Marine Science, Texas A&M University at Galveston, Galveston, Texas 77553, United States

[§]Department of Atmospheric Science, Texas A&M University, College Station, Texas 77853, United States

[∥]Department of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, Texas 77030, United States

Supporting Information

ABSTRACT: Estimating exposures to $PM_{2.5}$ within urban areas requires surface $PM_{2.5}$ concentrations at high temporal and spatial resolutions. We developed a mixed effects model to derive daily estimations of surface $PM_{2.5}$ levels in Beijing, using the 3 km resolution satellite aerosol optical depth (AOD) calibrated daily by the newly available high-density surface measurements. The mixed effects model accounts for daily variations of AOD-$PM_{2.5}$ relationships and shows good performance in model predictions (R^2 of 0.81−0.83) and cross-validations (R^2 of 0.75−0.79). Satellite derived population-weighted mean $PM_{2.5}$ for Beijing was 51.2 $\mu g/m^3$ over the study period (Mar 2013 to Apr 2014), 46% higher than China's annual-mean $PM_{2.5}$ standard of 35 $\mu g/m^3$. We estimated that more than 19.2 million people (98% of Beijing's population) are exposed to harmful level of long-term $PM_{2.5}$ pollution. During 25% of the days with model data, the population-weighted mean $PM_{2.5}$ exceeded China's daily $PM_{2.5}$ standard of 75 $\mu g/m^3$. Predicted high-resolution daily $PM_{2.5}$ maps are useful to identify pollution "hot spots" and estimate short- and long-term exposure. We further demonstrated that a good calibration of the satellite data requires a relatively large number of ground-level $PM_{2.5}$ monitoring sites and more are still needed in Beijing.



INTRODUCTION

Particulate matter (PM) air pollution in China is a major public health problem. According to the Global Burden of Disease report, ambient PM air pollution is responsible for over 1.2 million premature deaths annually in China, and 50% of its 1.3 billion population is currently exposed to ambient $PM_{2.5}$ (particulate matter with aerodynamic diameters of less than 2.5 μm) in exceedance of an annual average of 35 $\mu g/m^3$ (China's National Ambient Air Quality Standard, or NAAQS, available at http://kjs.mep.gov.cn/).[1] In recent years, extremely high $PM_{2.5}$ concentrations have been reported over many urban areas in China, notably in Beijing.[2−5] Some of the worst pollution episodes saw $PM_{2.5}$ concentrations exceeding 500 $\mu g/m^3$ and lasting multiple days.[6,7] Accurate assessment of population exposure to such elevated levels of $PM_{2.5}$ over densely populated urban areas requires surface $PM_{2.5}$ concentrations at high temporal and spatial resolutions, but ground-level monitors have been sparse in China which hinders our ability to accurately assess the health impact of $PM_{2.5}$ pollution.

Satellite retrievals of AOD provide larger spatial coverage and have been widely employed as a proxy to infer surface $PM_{2.5}$ concentrations.[8−13] Early studies obtained the AOD-$PM_{2.5}$ relationships using simple linear regression models or chemical transport models (CTM) at a global or continental scale, including the United States (US) and Europe.[10,11,14−17] Those studies reported R^2, a measure of goodness of fit of linear regression, generally between 0.3 and 0.6 and the AOD-$PM_{2.5}$ relationship was typically assumed to be constant in time and location.[18] Advanced statistical models (e.g., generalized linear model, land use regression model, geographically weighted regression, and generalized additive model) have been applied in order to account for variability in the AOD-$PM_{2.5}$ relationships associated with meteorology or land use.[13,19−22] These advanced models yielded higher prediction performance (R^2 at 0.6−0.8) for large regions such as the northeastern and southeastern US and China, but the increased complexity of these models requires additional data as inputs. Lee et al.[23] first introduced an AOD daily calibration approach using a mixed

Received: March 20, 2015
Revised: August 19, 2015
Accepted: August 27, 2015
Published: August 27, 2015

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

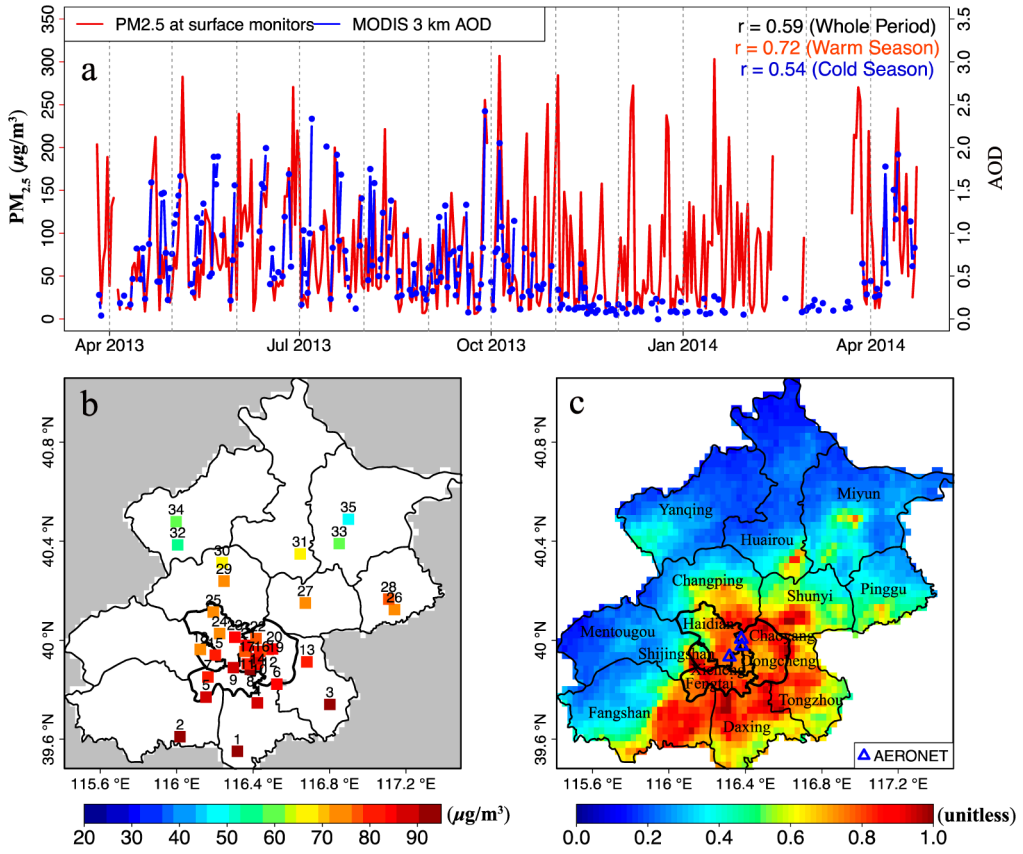


Figure 1. Temporal and spatial variations of $PM_{2.5}$ and AOD: (a) Daily time series of the mean $PM_{2.5}$ concentrations of the 35 monitoring sites and mean site-collocated AOD from the MODIS 3 km product (correlation coefficients r between different portions of the two time series are indicated); (b) Mean $PM_{2.5}$ concentrations at each site (sites numbered and marked) averaged over the study period; (c) Distribution of the 3 km resolution MODIS AOD averaged during the study period, with the boundaries of 16 districts of Beijing shown in black lines (district names in black character) and the location of the three AERONET sites shown as blue triangles.

effects model to account for day-to-day and location-specific variability in the AOD-$PM_{2.5}$ relationships over New England. They demonstrated that daily adjustment of the AOD-$PM_{2.5}$ relationships leads to a significant improvement in model performance with R^2 reaching 0.92, as compared to the linear regression ($R^2 \sim 0.51$) applied to the same region. The mixed effects model provides a valuable tool for estimating $PM_{2.5}$ concentrations in urban areas where ground monitoring sites are not of sufficient density.

Dense surface networks and high-resolution satellite AOD data are essential for improving the accuracy of exposure assessment models. Since Mar 2013, 35 $PM_{2.5}$ monitoring sites have been deployed in Beijing Municipality by Beijing Environmental Protection Bureau, providing hourly $PM_{2.5}$ concentrations at the urban and suburban areas of Beijing. In early 2014, the Moderate Resolution Imaging Spectroradiometer (MODIS) team released the new collection 6 (C006) AOD products at a 3 km resolution.[24] Prior standard AOD products from MODIS have a spatial resolution of 10 km or coarser. Most of the previous studies applying the MODIS AOD to derive surface $PM_{2.5}$ concentrations over China focused on the regional scale with resolutions ranging from 10 km to 100 km.[12,13,22] Utilizing the newly available surface monitoring data in Beijing, this study is the first attempt to develop a mixed effects model to estimate ground-level $PM_{2.5}$ concentrations using the 3 km resolution AOD data on a daily basis for a heavily polluted urban environment in China. The severe pollution events offer a valuable opportunity to assess the performance of the daily AOD calibration approach through this mixed effects model. The paper is organized as follows: surface and satellite data as well as model development are introduced in the "Materials and Methods" section; the model performance, $PM_{2.5}$ prediction results, and uncertainties are provided in the "Results and Discussion" section.

MATERIALS AND METHODS

Ground-Level $PM_{2.5}$ Data. Hourly $PM_{2.5}$ concentrations observed at 35 monitoring stations in Beijing (Figure 1b) were obtained from Beijing Municipal Environmental Monitoring Center (http://zx.bjmemc.com.cn/). Surface $PM_{2.5}$ mass concentrations are measured by the Tapered Element Oscillating Microbalance (TEOM) method, and the measurements have undergone the calibration processes and quality controls according to the environmental protection standard of China (HJ 618-2011; MEPCN).[25] Beijing Municipality consists of 16 districts with six urban districts located in the center of the city and ten suburban districts surrounding the urban regions (Figure 1c). Among the 35 $PM_{2.5}$ monitors, 17 are located in the urban districts and 18 in suburban districts. We averaged the surface concentrations measured between 13:00 pm and 14:00 pm local time at each site to derive the daily $PM_{2.5}$ concentrations that match with the overpassing time of the Aqua satellite. The study period is from March 26th 2013 to April 23th 2014, spanning a total of 394 days.

MODIS 3 km AOD Products and Calibration. *MODIS 3 km AOD Product.* MODIS on board the NASA Aqua satellite

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

Table 1. Descriptive Statistics of $PM_{2.5}$ Observations from the 35 Surface Sites and Site-Collocated AOD from the 3 km MODIS Product during the Study Period (March 26th 2013 to Apr 23th 2014)

	Site Average $PM_{2.5}$ ($\mu g/m^3$)					
averaging period	N^a	mean	SD^b	min	max	median
all	11126	81.04	73.23	3.78	385.80	57.84
warm season (Apr 15th–Oct 14th)	5929	80.61	65.53	3.94	331.47	61.68
cold season (Oct 15th–Apr 14th)	4773	81.26	80.39	4.02	377.86	51.40
	Site-Collocated Average AOD (unitless)					
averaging period	N	mean	SD	min	max	median
all	2818	0.68	0.49	0.09	2.39	0.53
warm season (Apr 15th–Oct 14th)	2219	0.75	0.51	0.10	2.39	0.61
Cold season (Oct 15th–Apr 14th)	594	0.34	0.26	0.12	0.98	0.28
	$PM_{2.5}$ (AOD) during Periods When Both Data Are Available					
averaging period	N	mean	SD	min	max	median
all	1933	49.58 (0.64)^c	41.06 (0.48)	4.80 (0.10)	186.44 (2.23)	36.10 (0.50)
warm season (Apr 15th–Oct 14th)	1482	52.59 (0.71)	42.12 (0.50)	5.26 (0.11)	184.30 (2.23)	39.56 (0.58)
cold season (Oct 15th–Apr 14th)	451	35.39 (0.33)	31.97 (0.23)	8.94 (0.13)	101.81 (0.85)	25.75 (0.28)

^{a}N denotes the number of valid observations. ^{b}SD represents the standard deviation of the data. ^{c}Numbers in parentheses are for the MODIS AOD.

has been in operation since 2002, providing retrieval products of aerosol and cloud properties with nearly daily global coverage.[26] The dark target algorithms make use of surface reflectance at three wavelength channels (0.47 μm, 0.66 μm, and 2.12 μm) for AOD retrievals over land. Standard MODIS Level 2 (L2) AOD products are distributed at a 10 km resolution. To fulfill the need for higher resolution pollution detection, the most recently released MODIS Collection 6 product (MYD04_3K) provides aerosol products at a 3 km resolution in addition to the L2 10 km product. The retrieval algorithm of the higher resolution product is similar to that of the 10 km standard product, but averages 6 × 6 pixels in a single retrieval box rather than the 20 × 20 pixels after cloud screening and other surface mask processes. Pixels outside the reflectivity range of the brightest 50% and darkest 20% at 0.66 μm are discarded to reduce uncertainty. Validation against surface sun photometer shows that two-thirds of the 3 km retrievals fall within the expected error on a regional comparison but with a high bias of ~0.06 especially over urban surface.[27]

MODIS AOD Validation. The ground-based AOD measurements from three AErosol RObotic NETwork (AERONET) sites (http://aeronet.gsfc.nasa.gov/) in Beijing were used to validate the satellite-derived AOD. To be comparable with the spectrum setting of MODIS, AERONET AOD at 550 nm were calculated by interpolating AOD at 440 nm and at 675 nm using the reported angstrom exponent at the respective wavelengths. The 3 km MODIS AOD products from Aqua show high temporal correlations with the AERONET AOD (Pearson correlation coefficient r of 0.93, 0.93, and 0.94 at the three sites respectively, Text S1, Supporting Information (SI)). The mean AOD difference between MODIS and AERONET at the three sites is 0.29. The higher AOD from MODIS could be partly attributed to the reflectance bias over the brighter surface of urban areas, which was also observed over other urban regions.[24,27] Since this bias is persistent for the full range of AOD values, it can be treated as a systematic bias when deriving the $PM_{2.5}$ to AOD ratio and is not expected to influence surface $PM_{2.5}$ estimation in this study because we calibrated MODIS AOD with ground-level $PM_{2.5}$ measurements on a daily basis.

Model Development and Validation. We selected the site-collocated satellite AOD values for each surface site where it falls within a 3 km grid. If there are more than one site within a single 3 km grid, the $PM_{2.5}$ values of those sites are averaged. With this process, sites #12 and #14 are averaged, and there remain 34 pairs of AOD and $PM_{2.5}$ data for model development.

A linear regression model was first applied to the collocated AOD and $PM_{2.5}$ data sets. The linear regression model for all 34 sites follows the form of

$$PM_{2.5} = \alpha + \beta \times AOD \quad (1)$$

where α and β are the fixed intercept and slope, respectively. The intercept and slope do not have spatial and temporal variations since the linear regression model assumes that the AOD-$PM_{2.5}$ relationship is constant for all the sites throughout the study period. The log transformed AOD and $PM_{2.5}$ were also tested, and no significant improvement was found in the regression performance. Therefore, we only present the regression results with the original AOD and $PM_{2.5}$ data sets.

In reality, the AOD-$PM_{2.5}$ relationship may exhibit spatial and temporal variations due to changing meteorology and other factors. To represent the varying AOD-$PM_{2.5}$ relationship, we developed a mixed effects model following the approach proposed by Lee et al.[23] which takes into consideration daily variations of the AOD-$PM_{2.5}$ relationship and derives site-specific parameters for spatial adjustment. The mixed effects model estimates surface $PM_{2.5}$ concentrations for each site i of day j ($PM_{2.5,ij}$) from collocated MODIS AOD (AOD_{ij}) in the form of

$$PM_{2.5,ij} = (\alpha + u_j) + (\beta + v_j) \times AOD_{ij} + s_i + \varepsilon_{ij}$$
$$(u_j v_j) \sim N[(00), \sum] \quad (2)$$

where α and β are the fixed intercept and slope independent of time and location, and u_j and v_j are the random intercept and slope for all the sites at each day. When the subscript i is not present for a parameter in eq 2, it indicates that the parameter does not change by site. The random terms (u_j and v_j) reflect the day-to-day variations of the AOD-$PM_{2.5}$ relationship influenced by meteorology, satellite retrieval conditions, etc. The $s_j \sim N(0, \sigma_s^2)$ is a site term which accounts for the spatial difference of the AOD-$PM_{2.5}$ relationship due to differences in site specific characteristics (i.e., surface reflectivity, topography,

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280–12288

Table 2. Model Performance Statistics for the Linear Regression and the Mixed Effects Model over All the Surface Sites

model type	N^a	slope^b	intercept^c	R^2	MPE^d ($\mu g/m^3$)	RMSE^e($\mu g/m^3$)
MODIS 3 km AOD (Whole Period)						
linear regression	1933	58.67	10.08	0.47	21.49	32.09
mixed effects	1435	53.13	20.44	0.81	11.45	17.85
mixed effects (w/site effect)	1435	55.62	17.74	0.83	10.69	16.63
MODIS 3 km AOD (Warm Season Only)						
linear regression	1482	57.96	8.54	0.47	21.82	32.44
mixed effects	1150	46.04	19.46	0.79	11.57	18.14
mixed effects (w/site effect)	1150	49.33	17.22	0.82	10.70	16.91
MODIS 3 km AOD (Cold Season Only)						
linear regression	451	100.20	3.93	0.52	18.23	27.37
mixed effects	285	107.06	17.11	0.87	10.24	15.42
mixed effects (w/site effect)	285	108.09	14.67	0.89	10.28	14.48

[a]N denotes total available pairs of data. [b]Fixed regression slope derived from the models. [c]Fixed regression intercept derived from the models. [d]MPE is estimated as the absolute differences between predicted and measured $PM_{2.5}$ concentrations. [e]RMSE is estimated as the root mean squared differences between predicted and measured $PM_{2.5}$ concentrations.

$PM_{2.5}$ emissions, and pollution transported to the observation sites). The comparison between the mixed effects models with and without the site effect s_i provides a measure of the sensitivity in model-derived day-to-day variations of the AOD-$PM_{2.5}$ relationship to site locations. ε_{ij} represents the error term, and $\sum$ is the variance-covariance matrix for the day-specific random effects. We excluded days with less than two pairs of AOD-$PM_{2.5}$ data and performed model predictions on the remaining available days. Model performances were evaluated by comparing the predictions to ground measurements using R^2, mean prediction error (MPE), and root-mean-square error (RMSE). The correlation coefficient, r, reported in this study is the Pearson correlation coefficient unless stated otherwise.

A cross-validation (CV) method was implemented to test the performance of the linear regression model and mixed effects model. We isolated one site at a time, performed model fitting with the remaining 33 sites, and validated model performances on the isolated site. The process was repeated for each of the 34 sites. The CV statistics are represented with R^2, MPE, and RMSE. We also varied the number of isolated sites during cross-validation to 5, 10, and 20 as a way of testing the sensitivity of the mixed effects model to the density of surface monitors needed.

RESULTS AND DISCUSSION

Descriptive Statistics. Table 1 presents the descriptive statistics of measured $PM_{2.5}$ concentrations at the 35 surface sites and site-collocated MODIS AOD during the study period. Daily time series of the site-average $PM_{2.5}$ concentrations and site-collocated MODIS AOD (Figure 1a) indicate an overall good correlation between them (r = 0.6). Average $PM_{2.5}$ from the ground-based monitors is 81.04 $\mu g/m^3$ during the study period, more than a factor of 2 higher than China's NAAQS of 35 $\mu g/m^3$ for annual-mean $PM_{2.5}$, and the standard deviation (SD) is 73.23 $\mu g/m^3$. The average AOD is 0.68 with a SD of 0.49. The maximum site-average daily mean $PM_{2.5}$ is 385.80 $\mu g/m^3$, and the maximum single-site daily concentration reaches 500 $\mu g/m^3$. Many sites have high AOD values larger than 2.0. The data are further divided into the warm season (Apr 15th to Oct 14th) and the cold season (Oct 15th to Apr 14th). The average $PM_{2.5}$ concentrations does not differ much between the two seasons, but the AOD values are much lower in the cold season than in the warm season. Besides the impact of lower planetary boundary layer (PBL) depths in the cold season as well as other meteorological conditions (e.g., lower relative humidity), a lower percent of available satellite AOD retrievals over snow also explains the lower mean and SD of AOD in the cold season. The detailed statistics for each site and description for seasonal comparisons are shown in Text S2 (SI).

Spatially surface $PM_{2.5}$ from the ground-level monitors presents a decreasing gradient from south to north due to topography and land use difference in Beijing (Figure 1b). The south to north $PM_{2.5}$ gradient generally demonstrates the emission difference between suburban and urban regions. In addition, the southern sites are more influenced by pollution transported from the cities south of Beijing, while north of Beijing is surrounded by mountains. Satellite AOD averaged for the entire study period exhibits a similarly strong spatial gradient (Figure 1c), with the highest values over the southeast urban districts. The correlation between the period-mean $PM_{2.5}$ and site-collocated AOD is 0.64 across the 35 sites, indicating a relatively tight spatial consistency between the two data sets.

Linear Regression Model. The linear regression model gives a regression slope of 58.67 and an intercept of 10.08 between measured $PM_{2.5}$ and site-collocated MODIS AOD, with an overall R^2 of 0.47 ($p < 0.0001$) (Table 2). MPE and RMSE are 21.49 $\mu g/m^3$ and 32.09 $\mu g/m^3$, respectively. The regression performance at each of the 34 sites (sites #12 and #14 were averaged) is shown in Text S2 (SI). The correlation is slightly higher in the cold season (R^2 = 0.52) than that in the warm season (R^2 = 0.47). The linear regression slope for the warm season is 42.2% lower than that for the cold season due mainly to lower PBL and lower RH in the cold season. The simple linear regression model will be used as a benchmark to assess the improvement in predictability by the mixed effects model.

Mixed Effects Model Fitting and Validation. After the data selection process described above, there are a total of 120 valid days including 1435 pairs of AOD-$PM_{2.5}$ data available for model fitting. The parameters and performances of the mixed effects model are compared with those from the linear regression model in Table 2. The overall R^2 between the predicted and measured $PM_{2.5}$ is 0.81 from the mixed effects model without a site term, which is a significant improvement compared to the R^2 of 0.47 from the linear regression model. The prediction performance at each site is displayed in Text S2 (SI). The site-specific R^2 ranges from 0.54 to 0.96 with the

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

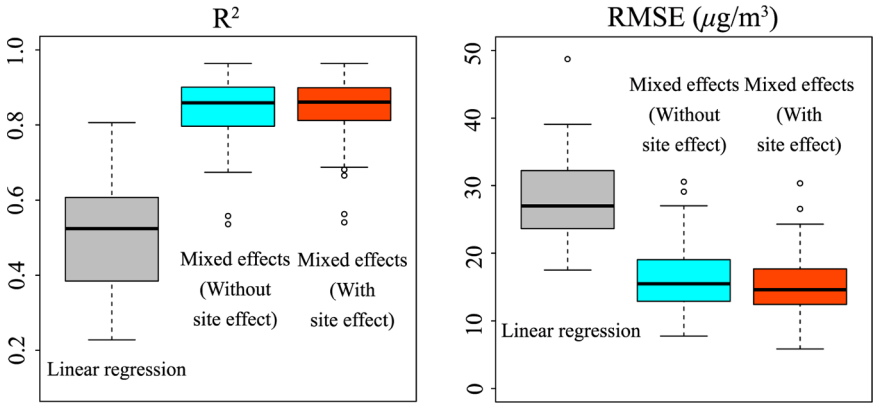


Figure 2. Box plots of prediction performance of the linear regression model and mixed effects model (with and without site term) at each site for (left) R^2 and (right) RMSE ($\mu g/m^3$).

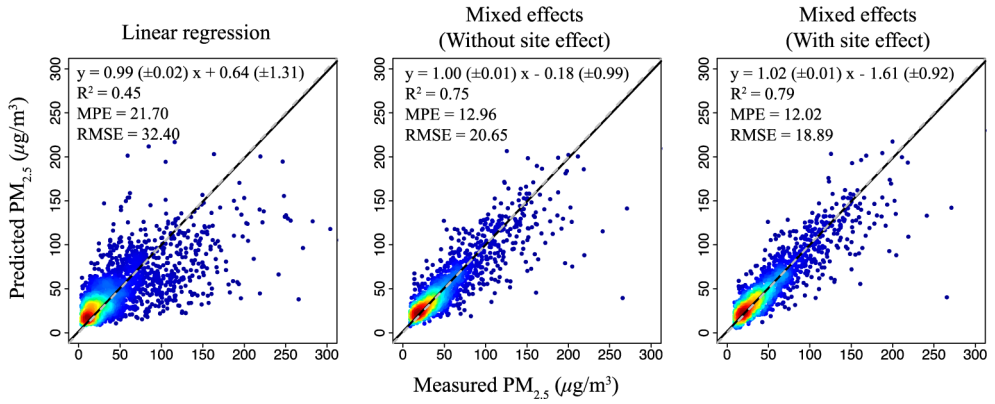


Figure 3. Scatter plots for cross-validation results between measured and predicted $PM_{2.5}$ concentrations from the linear regression model (left), the mixed effects model without site effect (center), and the mixed effects model with site effect (right).

average R^2 of 0.83 and SD of 0.10 (Figure 2). By adding the daily calibration of the AOD-$PM_{2.5}$ relationship through the mixed effects model, the 3 km AOD product from MODIS explains on average 83% of the observed surface $PM_{2.5}$ variability. The overall MPE and RMSE are 11.45 $\mu g/m^3$ and 17.85 $\mu g/m^3$, respectively, 47.3% and 44.4% lower than the corresponding values from the simple linear regression. The fixed term of the intercept and slope is 20.44 ($p < 0.001$) and 53.13 ($p < 0.001$), with the standard errors being 2.30 and 5.00, respectively. The daily specific intercept and slope have a standard deviation of 13.46 and 32.98, respectively. The mixed effects model has a higher R^2 (0.87) and smaller MPE and RMSE for predicting $PM_{2.5}$ in the cold season than for the warm season. The mean daily specific slope for the warm and cold season is −6.6 (SD = 28.2) and 12.7 (SD = 37.8), respectively. The larger slope for the cold season reflects a higher portion of $PM_{2.5}$ concentrated near the surface due to lower PBL, consistent with the linear regression model, while the higher SD of the cold season slope indicates the larger day to day variation. The intercepts have less variation between the two seasons.

The site effect term is incorporated in the mixed effects model to examine if the AOD-$PM_{2.5}$ relationship can be further improved with spatial variation. The site effect term at each monitoring location varies from −10.55 to 18.71 (SD = 6.03). Larger values are found at sites located in southern (sites #1, 2, 3) and eastern (site #26) Beijing. These southern sites are subject to more frequent regional transport of aerosols from the heavily industrialized cities southward and southwestward of Beijing. The other two sites near traffic (sites #8, 12) also show a higher site effect term, indicating the influence of local pollution. The mixed effects model with a site effect term has an overall R^2 of 0.83, with MPE of 10.69 $\mu g/m^3$, and RMSE of 16.63 $\mu g/m^3$ (Table 2). Detailed statistics are shown in Text S2 (SI). Compared with the mixed effects model without the site effect, adding the site effect shows only a slight improvement in the model performance with a 6% decrease of MPE and RMSE (Table 2). This comparison further confirms that day-to-day variation of the AOD-$PM_{2.5}$ relationship is a more important factor than the spatial variation of this relationship. In support of this, we found that exclusion of the two southern sites (#1, 2) or the two traffic sites (#8, 12), which are not represented by the 3 km AOD product, does not affect the performance of the mixed effects model (Text S3, SI).

Given the difference in data statistics between the cold and the warm season, we tested if the linear regression model and the mixed effects model would give better results when fitted separately for each season (Table 2). Based on detailed comparison (Text S2), we found that it is unlikely that the model itself can be improved through separate fitting by season. Therefore, we chose to use the mixed effects model fitted with data from the entire period for $PM_{2.5}$ predictions over Beijing.

Cross-validation results for different models are shown in Figure 3. The R^2 values for the linear regression model and mixed effects models with and without site effects are 0.45, 0.75, and 0.79, respectively. The CV MPE and RMSE of the

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

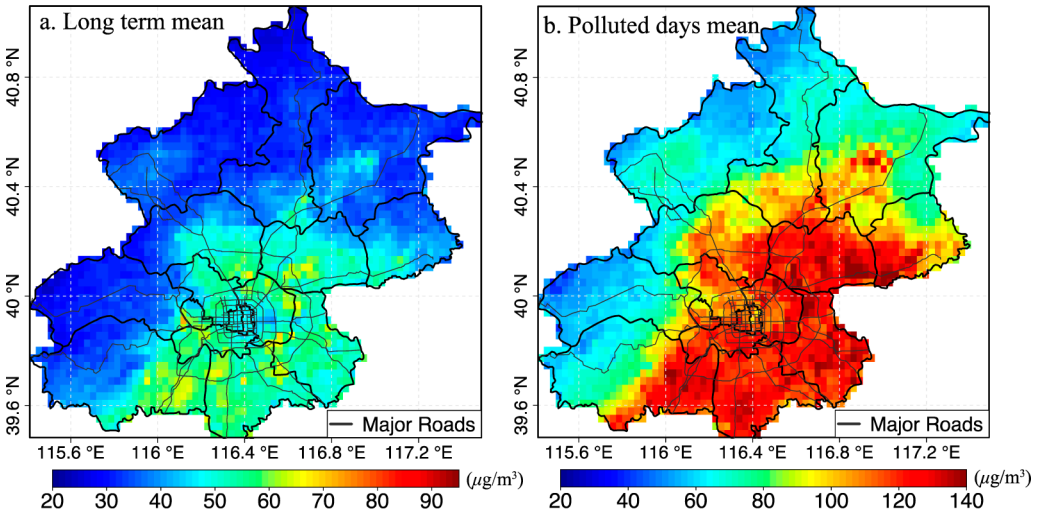


Figure 4. (a) Mean $PM_{2.5}$ concentrations ($\mu g/m^3$) derived from the 3 km resolution MODIS AOD product for Beijing averaged over all the days with valid data; (b) same as left but averaged over the heavily polluted days (daily population-weighted mean $PM_{2.5}$ larger than 75 $\mu g/m^3$). The thin black lines indicate major roads in Beijing.

mixed effects models are reduced by ~50% compared to the linear regression model. To assess the model dependence on the density of surface monitors needed for calibration, we increased the number of isolated sites during cross-validation of the mixed effects model (without site effect) to 5, 10, and 20 and compared the resulting effects on model agreement. When the number of isolated sites increases from 1 to 10, the CV R^2 decreases from 0.75 to 0.71, while MPE and RMSE increase by ~11%. The CV statistics show little change when the number of isolated sites is less than three. When 20 sites are isolated, the R^2 decreases to 0.64 with MPE and RMSE increased by up to 25%. This sensitivity analysis demonstrates the advantage of using more surface monitors for model fitting and calibration. It also suggests that at least 31 sites are needed for a successful calibration of the mixed effects model over Beijing.

Predicted Surface $PM_{2.5}$. Given the high predicting power of the mixed effects model, it is used to derive surface $PM_{2.5}$ maps for Beijing at 3 km resolution. The site effect is not used when mapping surface $PM_{2.5}$ for the following reasons: first, the site effect is a site specific term and such information is not available for every 3 km grid; second, while the site effect term can be interpolated spatially onto each grid, this process brings in additional uncertainty in predicted $PM_{2.5}$ which cannot be well characterized given the sparseness of the existing monitors; third, the mixed effects model without the site effect has similar performance statistics compared to the one with the site effect. The predicted long-term mean surface $PM_{2.5}$ map is shown in Figure 4a. Here the long-term mean refers to the average from all valid model-days during the entire study period. There is a southeast-to-northwest gradient in the predicted $PM_{2.5}$ distribution, with higher concentrations over downtown urban areas and southern regions and lower values over north and western suburban districts. The predicted long-term mean $PM_{2.5}$ for the whole Beijing city is 40.97 $\mu g/m^3$. The district-mean $PM_{2.5}$ concentrations, calculated by averaging the predicted $PM_{2.5}$ from all the satellite grids that fall within the boundaries of individual districts, range from 32.2 $\mu g/m^3$ for the Mentougou district to 55.5 $\mu g/m^3$ for the Daxing district. For comparison, we calculated the monitor-derived district-mean $PM_{2.5}$ by averaging the measured $PM_{2.5}$ from all the sites that fall within each district (Text S4, SI). The number of available sites for individual districts ranges from one to a maximum of four. For the densely populated Xicheng and Chaoyang districts, the satellite-derived mean $PM_{2.5}$ is ~20% and 12% higher than the monitor-derived mean, respectively. This suggests that those urban regions are particularly undersampled by the existing surface monitors.

The population-weighted mean $PM_{2.5}$ is derived by weighting the satellite-derived $PM_{2.5}$ by population in each grid. The gridded population data was adopted from the 1 km population data set by Fu et al.[28] The population data were regridded to the same 3 km resolution as the satellite-derived $PM_{2.5}$. The resident population of Beijing was 19.61 million in 2010, consisting of 11.72 million (59.8%) in the six urban districts and 7.89 million (40.2%) in the ten suburban districts according to the sixth national population census. The population-weighted mean $PM_{2.5}$ for Beijing as a whole is 51.2 $\mu g/m^3$, which is 25.0% higher than the area mean of 40.97 $\mu g/m^3$ and 46.3% higher than China's annual mean standard of 35 $\mu g/m^3$. The predicted $PM_{2.5}$ levels indicate that approximately 58.7% of the area of Beijing Municipality, corresponding to a total of 19.2 million or 98% of the total population, is exposed to risky $PM_{2.5}$ pollution levels, which is a serious health concern for Beijing. Higher $PM_{2.5}$ concentrations are predicted near main traffic routes, and several $PM_{2.5}$ hotspots are located in areas of higher road density, for example over the urban center, near the main expressways extending southward to Fangshan and Daxing districts, and at road conjunctions in Pinggu and Shunyi districts (Figure 4a),[29] suggesting emissions from road transportation as an important contributor to $PM_{2.5}$ pollution in Beijing. Both AOD-derived $PM_{2.5}$ and surface measurements indicate lower $PM_{2.5}$ levels within the second ring road at the urban center where there is lower industry activity and higher percentage of parks and common spaces.

Heavily polluted days were selected as the days with city-average daily population-weighted mean $PM_{2.5}$ concentrations greater than 75 $\mu g/m^3$ (China NAAQS level 2 daily standard). Among all the 120 days with valid data, a total of 30 days (25%) are heavily polluted days. As shown in Figure 4, the mean $PM_{2.5}$ of those heavily polluted days displays a similar spatial pattern with the mean $PM_{2.5}$ of all the days but with enhanced $PM_{2.5}$ levels over the whole city of Beijing. The population-weighted

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

mean (area-averaged) $PM_{2.5}$ during the heavily polluted days is 109.5 $\mu g/m^3$ (90.02 $\mu g/m^3$) over the whole Beijing, 46% (20%) higher than the daily standard. During the heavily polluted days, the southern and southeastern edge of Beijing had averaged $PM_{2.5}$ concentration almost twice the mean concentration of all the days, suggesting that regional pollution transported from south of Beijing contributes to high pollution levels in Beijing during those polluted days. Given the regional influence, those polluted days also saw 123.9% and 85.1% higher $PM_{2.5}$ levels than the mean conditions in the relatively cleaner northern and northwestern districts, respectively. Under the polluted conditions, more than 58.67% of the city, which includes all the urban districts and 55.01% of the suburban districts, is exposed to $PM_{2.5}$ levels higher than 75 $\mu g/m^3$. These areas correspond to a total population of 19.07 million, including 11.72 million in urban districts and 7.35 million in suburban districts. The high $PM_{2.5}$ concentrations and high population exposures presented here provide clear evidence for the serious health risk of PM pollution for Beijing residents, calling for aggressive emission control measures both within the city and in the surrounding regions.

Model Performance Comparison. To demonstrate the benefit of the fine-resolution AOD product, we averaged the 3 km AOD product onto 6 km and 9 km resolution and used those regridded coarser-resolution AOD to develop the mixed effects model. The details on data processing, model developments, and model performances comparison are provided in the Text S5 (SI). The mixed effects model developed from the 3 km AOD product has better performances than that developed from the 6 km and 9 km products, with R^2 higher by up to 0.02 and MPE and RMSE lower by ~10%.

To compare with previously published studies which used the standard MODIS L2 10 km AOD products as predictors of surface $PM_{2.5}$, we also tested the change in model performances when using the standard 10 km product (Text S5, SI).[30] The total available pairs of AOD-$PM_{2.5}$ data decrease by 32.9% when using the 10 km AOD product. Spatially, the 10 km product shows similar patterns with the 3 km products but provides much less details in terms of spatial variability. The model R^2 based on the 10 km AOD product is ~0.02 lower than that based on the 3 km product, with mean MPE and RMSE higher by ~13% and 17%, respectively. The regression performance based on the 10 km product is also lower than that using the regridded 6 km and 9 km products, further demonstrating the benefit of using finer-resolution AOD. The AOD variation of a single 10 km pixel can be as high as 0.2. Therefore, the 3 km resolution product provides better spatial representation and allows capturing intraurban variations of the $PM_{2.5}$ concentrations, offering a direct benefit for air quality and pollution exposure assessment.

By considering daily calibration of the AOD-$PM_{2.5}$ relationship, the mixed effects model gives higher modeled R^2 (0.83) and CV R^2 (0.79) than previously published studies, e.g., the geographically weighted regression model over the entire China region (R^2 = 0.64) and for Pearl River Delta (R^2 = 0.74), the empirical nonlinear model for Xi'an (R^2 = 0.67).[13,22,31] A recent work estimated surface $PM_{2.5}$ for China using MODIS AOD at a 1 km resolution, but the R^2 between the observed and predicted $PM_{2.5}$ was only 0.49 for Beijing, possibly due to less surface observations used for daily model calibration.[32] The performance statistics of the mixed effects model developed here for Beijing are also comparable with those from the U.S. studies using the 1 km resolution MODIS AOD derived from the Multi-Angle Implementation of Atmospheric Correction (MAIAC) program (R^2 = 0.67−0.84).[33,34]

Prediction Uncertainties. While the satellite-predicted $PM_{2.5}$ provides larger spatial coverage than the unevenly distributed ground-based monitors, the satellite has less temporal coverage due to its sampling limitation associated with surface conditions, clouds, and other factors especially in the cold season. With only 120 days of model-predicted $PM_{2.5}$ during the whole study period, the satellite-predicted mean $PM_{2.5}$ for those model-days is on average 32.0% lower than the monitor-derived mean $PM_{2.5}$ for the whole study period at the site locations. The difference at individual sites ranges from 11.7% to 41.8%. To account for different data availability during the warm and cold season, we calculated the long-term mean of predicted $PM_{2.5}$ by weighting the seasonal means by the number of available predictions in each season. The weighted long-term mean of predicted $PM_{2.5}$ at site locations is 18.9% lower than the monitor-derived long-term mean. This suggests that the sampling bias by the satellite may affect the long-term mean $PM_{2.5}$ derived from the AOD-based daily calibration model, and the estimation of long-term mean population exposure from this study is probably too low.

Other possible errors of the predicted surface $PM_{2.5}$ concentrations include those from both satellite AOD products and surface $PM_{2.5}$ measurements. The 3 km AOD products from MODIS have shown comparable quality with the widely validated 10 km products with an expected error of 0.05 ± 0.25 AOD.[24] However, improper characterization of surface reflectance will adversely impact retrieval accuracy of the higher-resolution products.[27] While this study validated the 3 km MODIS AOD products with ground-based AOD at three urban AERONET sites in Beijing, there are no ground-based AOD measurements to characterize the bias of MODIS AOD over other urban surfaces or the suburban regions in the city. Measurement bias in surface $PM_{2.5}$ concentrations from the TEOM instrument also contributes to the prediction uncertainty.[35,36] In addition, different spatial scales could bring uncertainties in the AOD-$PM_{2.5}$ relationship because the MODIS AOD reflect the average condition of a 3 km grid, while the site $PM_{2.5}$ is based on point measurement. This partly explains higher $PM_{2.5}$ at site #1 which is not observed by satellite. The mixed effects model also introduces uncertainties with the assumption of linearity and AOD as the single predictor.

Despite the uncertainties, this study is the first application of the most recent higher-resolution MODIS AOD products to predict surface $PM_{2.5}$ concentrations over Beijing at 3 km resolution. The advantage of the mixed effects model developed here is that it calibrates AOD-derived $PM_{2.5}$ models on a daily basis to account for the time varying AOD-$PM_{2.5}$ relationship. By allowing the AOD-$PM_{2.5}$ relationship to change temporally, the mixed effects model considers, at least partially, the impact of other factors correlated with time, such as emissions from anthropogenic activities and vegetation as well as weather conditions, because those factors are expected to affect the daily variability of the AOD-$PM_{2.5}$ relationship. The high prediction performance (R^2 = 0.83) of the mixed effects model has demonstrated the value of using the higher-resolution AOD products at the urban scale for Beijing, where population density varies from 23407/km^2 in urban center to 958/km^2 in suburban developing districts. Nevertheless, the model cannot include the full extent of other explanatory factors. Given the R^2 of 0.81−0.83 resulting from the mixed effects model calibrated

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288

in this study, the impact of other explanatory factors is unlikely to play a dominant role, but it is worth future investigation to further improve the model. In addition, surface observations from more sites and longer periods are useful to calibrate the AOD-$PM_{2.5}$ relationships.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.5b01413.

Texts S1_S5, Figures S1−S4, and Tables S1−S5 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 86-13701032461. E-mail: yxw@tsinghua.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the National Key Basic Research Program of China (2014CB441302), the CAS Strategic Priority Research Program (Grant No. XDA05100403), and the Beijing Nova Program (Z121109002512052). The authors thank Mr. Yi Zou for providing the photographs shown on the journal cover. The photo series "Beijing: Be Clear at a Glance" features daily photos of Beijing taken from the same location since 2013.

■ REFERENCES

(1) Lim, S. S.; Vos, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; Amann, M.; Anderson, H. R.; Andrews, K. G.; Aryee, M.; et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990−2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **2012**, *380*, 2224−2260.

(2) Guo, S.; Hu, M.; Zamora, M. L.; Peng, J. F.; Shang, D. J.; Zheng, J.; Du, Z. F.; Wu, Z.; Shao, M.; Zeng, L. M.; et al. Elucidating severe urban haze formation in China. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (49), 17373−17378.

(3) Che, H.; Xia, X.; Zhu, J.; Li, Z.; Dubovik, O.; Holben, B.; Goloub, P.; Chen, H.; Estelles, V.; Cuevas-Agulló, E.; et al. Column aerosol optical properties and aerosol radiative forcing during a serious haze-fog month over North China Plain in 2013 based on ground-based sunphotometer measurements. *Atmos. Chem. Phys.* **2014**, *14* (4), 2125−2138.

(4) Andersson, A.; Deng, J.; Du, K.; Zheng, M.; Yan, C.; Skold, M.; Gustafsson, O. Regionally-Varying Combustion Sources of the January 2013 Severe Haze Events over Eastern China. *Environ. Sci. Technol.* **2015**, *49* (4), 2038−43.

(5) Tao, M.; Chen, L.; Wang, Z.; Tao, J.; Su, L. Satellite observation of abnormal yellow haze clouds over East China during summer agricultural burning season. *Atmos. Environ.* **2013**, *79*, 632−640.

(6) Bi, J. R.; Huang, J. P.; Hu, Z. Y.; Holben, B. N.; Guo, Z. Q. Investigating the aerosol optical and radiative characteristics of heavy haze episodes in Beijing during January of 2013. *J. Geophys. Res. Atmos.* **2014**, *119* (16), 9884−9900.

(7) Wang, Y.; Zhang, Q.; Jiang, J.; Zhou, W.; Wang, B.; He, K.; Duan, F.; Zhang, Q.; Philip, S.; Xie, Y. Enhanced sulfate formation during China's severe winter haze episode in January 2013 missing from current models. *J. Geophys. Res. Atmos.* **2014**, *119* (17), 10425−10440.

(8) Chu, D. A.; Kaufman, Y. J.; Zibordi, G.; Chern, J. D.; Mao, J.; Li, C. C.; Holben, B. N. Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). *J. Geophys. Res.* **2003**, *108* (D21), 4661.

(9) Koelemeijer, R. B. A.; Homan, C. D.; Matthijsen, J. Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmos. Environ.* **2006**, *40* (27), 5304−5315.

(10) Zhang, H.; Hoff, R. M.; Engel-Cox, J. A. The Relation between Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical Depth and $PM_{2.5}$ over the United States: A Geographical Comparison by U.S. Environmental Protection Agency Regions. *J. Air Waste Manage. Assoc.* **2009**, *59* (11), 1358−1369.

(11) Schaap, M.; Apituley, A.; Timmermans, R. M. A.; Koelemeijer, R. B. A.; de Leeuw, G. Exploring the relation between aerosol optical depth and $PM_{2.5}$ at Cabauw, the Netherlands. *Atmos. Chem. Phys.* **2009**, *9* (3), 909−925.

(12) van Donkelaar, A.; Martin, R.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.* **2010**, *118* (6), 847−855.

(13) Ma, Z.; Hu, X.; Huang, L.; Bi, J.; Liu, Y. Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing. *Environ. Sci. Technol.* **2014**, *48* (13), 7436−44.

(14) 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.

(15) Engel-Cox, J. A.; Holloman, C. H.; Coutant, B. W.; Hoff, R. M. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmos. Environ.* **2004**, *38* (16), 2495−2509.

(16) Liu, Y.; Park, R. J.; Jacob, D. J.; Li, Q. B.; 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.* **2004**, *109* (D22), D22206.

(17) Boys, B. L.; Martin, R. V.; van Donkelaar, A.; MacDonell, R. J.; Hsu, N. C.; Cooper, M. J.; Yantosca, R. M.; Lu, Z.; Streets, D. G.; Zhang, Q.; Wang, S. W. Fifteen-Year Global Time Series of Satellite-Derived Fine Particulate Matter. *Environ. Sci. Technol.* **2014**, *48* (19), 11109−11118.

(18) 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.

(19) 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.

(20) 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.

(21) 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.

(22) Song, W.; Jia, H.; Huang, J.; Zhang, Y. A satellite-based geographically weighted regression model for regional $PM_{2.5}$ estimation over the Pearl River Delta region in China. *Remote. Sens. Environ.* **2014**, *154*, 1−7.

(23) Lee, H. J.; Liu, Y.; Coull, B. A.; Schwartz, J.; Koutrakis, P. A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations. *Atmos. Chem. Phys.* **2011**, *11* (15), 7991−8002.

(24) Remer, L. A.; Mattoo, S.; Levy, R. C.; Munchak, L. MODIS 3 km aerosol product: algorithm and global perspective. *Atmos. Meas. Tech.* **2013**, *6* (7), 1829−1844.

(25) MEPCN. Determination of atmospheric articles PM_{10} and $PM_{2.5}$ in ambient air by gravimetric method. Available at http://english.mep.gov.cn/standards_reports/ (accessed Dec 10, 2014).

(26) Remer, L. A.; Kaufman, Y. J.; Tanre, D.; Mattoo, S.; Chu, D. A.; Martins, J. V.; Li, R. R.; Ichoku, C.; Levy, R. C.; Kleidman, R. G.; et al.

The MODIS aerosol algorithm, products, and validation. *J. Atmos. Sci.* **2005**, *62* (4), 947−973.
(27) Munchak, L. A.; Levy, R. C.; Mattoo, S.; Remer, L. A.; Holben, B. N.; Schafer, J. S.; Hostetler, C. A.; Ferrare, R. A. MODIS 3 km aerosol product: applications over land in an urban/suburban region. *Atmos. Meas. Tech.* **2013**, *6* (7), 1747−1759.
(28) Fu, J. Y.; Jiang, D.; Huang, Y. H. 1 KM Grid Population Dataset of China (PopulationGrid_China). *Global Change Research Data Publishing & Repository*. 2014. http://www.geodoi.ac.cn/ (accessed Aug 1, 2015). DOI:10.3974/geodb.2014.01.06.V1.
(29) Ji, W.; Wang, Y.; Zhuang, D.; Song, D.; Shen, X.; Wang, W.; Li, G. Spatial and temporal distribution of expressway and its relationships to land cover and population: A case study of Beijing, China. *Transport. Res. Part D: Trans. Environ.* **2014**, *32*, 86−96.
(30) Levy, R. C.; Mattoo, S.; Munchak, L. A.; Remer, L. A.; Sayer, A. M.; Patadia, F.; Hsu, N. C. The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* **2013**, *6* (11), 2989−3034.
(31) You, W.; Zang, Z.; Pan, X.; Zhang, L.; Chen, D. Estimating $PM_{2.5}$ in Xi'an, China using aerosol optical depth: a comparison between the MODIS and MISR retrieval models. *Sci. Total Environ.* **2015**, *505*, 1156−65.
(32) Lin, C.; Li, Y.; Yuan, Z.; Lau, A. K. H.; Li, C.; Fung, J. C. H. Using satellite remote sensing data to estimate the high-resolution distribution of ground-level $PM_{2.5}$. *Remote. Sens. Environ.* **2015**, *156*, 117−128.
(33) Hu, X. F.; Waller, L. A.; Lyapustin, A.; Wang, Y. J.; Al-Hamdan, M. Z.; Crosson, W. L.; Estes, M. G.; Estes, S. M.; Quattrochi, D. A.; Puttaswamy, S. J.; Liu, Y. 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.
(34) Chudnovsky, A. A.; Koutrakis, P.; Kloog, I.; Melly, S.; Nordio, F.; Lyapustin, A.; Wang, Y.; Schwartz, J. Fine particulate matter predictions using high resolution Aerosol Optical Depth (AOD) retrievals. *Atmos. Environ.* **2014**, *89*, 189−198.
(35) Schwab, J. J.; Felton, H. D.; Rattigan, O. V.; Demerjian, K. L. New York state urban and rural measurements of continuous $PM_{2.5}$ mass by FDMS, TEOM, and BAM. *J. Air Waste Manage. Assoc.* **2006**, *56* (4), 372−383.
(36) Engel-Cox, J.; Nguyen Thi Kim, O.; van Donkelaar, A.; Martin, R. V.; Zell, E. Toward the next generation of air quality monitoring: Particulate Matter. *Atmos. Environ.* **2013**, *80*, 584−590.

DOI: 10.1021/acs.est.5b01413
Environ. Sci. Technol. 2015, 49, 12280−12288