



## Satellite-Based Spatiotemporal Trends in $PM_{2.5}$ Concentrations: China, 2004–2013

Zongwei Ma, Xuefei Hu, Andrew M. Sayer, Robert Levy,  
Qiang Zhang, Yingang Xue, Shilu Tong, Jun Bi, Lei Huang,  
and Yang Liu

<http://dx.doi.org/10.1289/ehp.1409481>

Received: 14 November 2014

Accepted: 17 July 2015

Advance Publication: 24 July 2015

This article will be available in a 508-conformant form upon final publication. If you require a 508-conformant version before then, please contact [ehp508@niehs.nih.gov](mailto:ehp508@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.

## Satellite-Based Spatiotemporal Trends in PM<sub>2.5</sub> Concentrations: China, 2004–2013

Zongwei Ma,<sup>1,2</sup> Xuefei Hu,<sup>2</sup> Andrew M. Sayer,<sup>3,4</sup> Robert Levy,<sup>4</sup> Qiang Zhang,<sup>5</sup> Yingang Xue,<sup>6</sup>  
Shilu Tong,<sup>7</sup> Jun Bi,<sup>1</sup> Lei Huang,<sup>1</sup> and Yang Liu<sup>2</sup>

<sup>1</sup> State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu, China; <sup>2</sup> Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; <sup>3</sup> Goddard Earth Sciences Technology and Research, Universities Space Research Association, Greenbelt, Maryland, USA; <sup>4</sup> NASA Goddard Space Flight Center, Greenbelt, Maryland, USA; <sup>5</sup> Center for Earth System Science, Tsinghua University, Beijing, China; <sup>6</sup> Changzhou Environmental Monitoring Center, Changzhou, Jiangsu, China; <sup>7</sup> Queensland University of Technology, Brisbane, Australia

**Address correspondence to** Yang Liu, Emory University, Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322 USA. Telephone: +1 404 727 2131. E-mail: [yang.liu@emory.edu](mailto:yang.liu@emory.edu). Lei Huang, School of the Environment, Nanjing University, 163 Xianlin Avenue, Nanjing 210023, China. Telephone: +86 25 89680535. E-mail: [huanglei@nju.edu.cn](mailto:huanglei@nju.edu.cn)

**Running title:** Satellite-estimated PM<sub>2.5</sub> trends in China

**Acknowledgment:** 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 Y. Liu was partially supported by NASA Applied Sciences Program (grant No. NNX11AI53G). The work of L. Huang and J. Bi was

partially supported by the Key Program of Chinese National Natural Science Foundation (71433007) and National High-tech R&D Program (863 Program) of China (2013AA06A309).

The work of Z. Ma was supported by the China Scholarship Council (CSC).

**Competing financial interests:** The authors declare that they have no actual or potential competing financial interests.

## ABSTRACT

**Background:** Three decades of rapid economic development is causing severe and widespread PM<sub>2.5</sub> pollution in China. However, research on the health impacts of PM<sub>2.5</sub> exposure has been hindered by limited historical PM<sub>2.5</sub> concentration data.

**Objectives:** We estimated ambient PM<sub>2.5</sub> concentrations from 2004 to 2013 in China at 0.1 degree resolution using the latest satellite data and evaluated model performance with available ground observations.

**Methods:** We developed a two-stage spatial statistical model using the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 aerosol optical depth (AOD), assimilated meteorology, land use data, and PM<sub>2.5</sub> concentrations from China's recently established ground monitoring network. An inverse variance weighting (IVW) approach was developed to combine MODIS Dark Target and Deep Blue AOD to optimize data coverage. We evaluated model-predicted PM<sub>2.5</sub> concentrations from 2004 to early 2014 using ground observations.

**Results:** The overall model cross-validation  $R^2$  and relative prediction error are 0.79 and 35.6%, respectively. Validation beyond the model year (2013) shows that it makes accurate predictions of PM<sub>2.5</sub> concentrations with little bias at the monthly ( $R^2 = 0.73$ , regression slope = 0.91) and seasonal levels ( $R^2 = 0.79$ , regression slope = 0.92). Seasonal variations show that winter is the most polluted season and summer is the cleanest season. Analysis of predicted PM<sub>2.5</sub> levels showed a mean annual increase of 1.97  $\mu\text{g}/\text{m}^3$  between 2004 and 2007, and a decrease of 0.46  $\mu\text{g}/\text{m}^3$  between 2008 and 2013.

**Conclusions:** Our satellite-driven model can provide reliable historical PM<sub>2.5</sub> estimates in China at a resolution comparable to those used in epidemiologic studies on the health effects of long-

term PM<sub>2.5</sub> exposure in North America. This data source can potentially advance PM<sub>2.5</sub> health effects research in China.

## Introduction

Fine particulate matter (PM<sub>2.5</sub>, particles with aerodynamic diameter less than 2.5 μm) has been strongly associated with adverse health effects (e.g., cardiovascular and respiratory morbidity and mortality) by numerous epidemiologic studies conducted primarily in developed countries (Pope and Dockery 2006). With the rapid economic development and urbanization, severe, widespread PM<sub>2.5</sub> pollution in China has attracted nationwide attention (Xu et al. 2013). However, research on the adverse health impacts of PM<sub>2.5</sub> exposure has been hindered since a nationwide regulatory PM<sub>2.5</sub> monitoring network did not exist until the end of 2012.

Estimating ground-level PM<sub>2.5</sub> from satellite-retrieved aerosol optical depth (AOD) is a promising, new method that has rapidly advanced recently (Hu et al. 2014b; Kloog et al. 2011; Lee et al. 2011; Liu et al. 2009). Satellite-driven statistical models have the potential to fill the spatiotemporal PM<sub>2.5</sub> gaps left by ground monitors with high-quality predictions. Several recent studies of the health effects due to long-term PM<sub>2.5</sub> exposure have adopted satellite-estimated PM<sub>2.5</sub> levels as their exposure estimates (Crouse et al. 2012; Madrigano et al. 2013). Since sufficient ground PM<sub>2.5</sub> measurements are needed to fit and validate statistical models, model development in China prior to 2013 has been difficult. van Donkelaar et al. (2010) estimated long-term (2001-2006) average global PM<sub>2.5</sub> concentrations at 0.1 degree resolution using the PM<sub>2.5</sub>/AOD ratios derived from a global chemical transport model (CTM). Two follow-up studies estimated the global PM<sub>2.5</sub> time series from 1998 to 2012 (Boys et al. 2014; van Donkelaar et al. 2015). Both studies only validated their seasonal average estimates with ground observations mostly from North America and the Pearson coefficients ranged from ~0.37 to ~0.68 ( $R^2 \sim 0.14-0.46$ ). Yao and Lu (2014) estimated PM<sub>2.5</sub> levels in China from 2006 to 2010 using an artificial neural network (ANN) model. However, their ANN was trained partially using

PM<sub>2.5</sub> and satellite data in the U.S., which may have introduced substantial prediction error.

Taking advantage of the newly available national PM<sub>2.5</sub> measurements, Ma et al. (2014) estimated PM<sub>2.5</sub> levels in 2013 in China using satellite AOD and a geographically weighted regression (GWR) model. Using an earlier version Dark Target (DT) algorithm (Remer et al. 2005), this study adopted a relatively coarse spatial resolution of 50 km, but did not attempt to estimate historical PM<sub>2.5</sub> levels. The coarse resolution is a result of the limited coverage of AOD values retrieved by the Moderate Resolution Imaging Spectroradiometer (MODIS, <http://modis.gsfc.nasa.gov>) instrument aboard the Terra and Aqua satellites launched by the National Aeronautics and Space Administration (NASA). In early 2014, more accurate Aqua MODIS Collection 6 (C6) AOD products retrieved by the enhanced DT (Levy et al. 2013) and Deep Blue (DB) algorithms (Hsu et al. 2013) were released. Despite better coverage over deserts and urban centers than DT AOD, DB AOD has rarely been used in PM<sub>2.5</sub> studies due to poorly characterized retrieval errors in earlier versions. As we demonstrate in the following sections, including the MODIS C6 DB AOD data substantially increases the spatiotemporal coverage of model predictions in China.

In this study, we developed a high-resolution (0.1 degree, which is approximately 10 km) statistical model to estimate historical ambient PM<sub>2.5</sub> concentrations from 2004 to 2013 in China using MODIS C6 AOD data. We first present our approach to generate a custom “combined” AOD parameter using the operational DT and DB AOD values and describe our two-stage spatial statistical model to estimate daily ambient PM<sub>2.5</sub> levels. We then evaluate predicted PM<sub>2.5</sub> concentrations at seasonal, monthly, and daily level using ground PM<sub>2.5</sub> measurements in China not included in model development. Finally, we analyze the 10-year spatiotemporal trend of PM<sub>2.5</sub> levels.

## Data and Methods

### *Ground PM<sub>2.5</sub> measurements.*

The daily average PM<sub>2.5</sub> concentrations of China (01/2013-06/2014) were primarily collected from the website of China Environmental Monitoring Center (CEMC). We collected additional data that are not included in the CEMC from the websites of local environmental monitoring centers of several provinces (e.g., Shandong, Shanxi, Zhejiang, and Guangdong) and municipalities (e.g., Beijing and Tianjin). Daily PM<sub>2.5</sub> data of Macao (2013), Hong Kong (2005-06/2014), and Taiwan (2004-06/2014) were also collected from websites of local environmental protection agencies. Data from the U.S. consulate sites in Beijing (2008-2013), Shanghai (2011-2013), Guangzhou (2011-2013), Shenyang (2013), and Chengdu (2012-2013) were also included. The website links of the above PM<sub>2.5</sub> data sources are shown in Table S1 (Supplemental Material). Data of Changzhou City in Jiangsu province were provided by Changzhou Environmental Monitoring Center. Monthly and seasonal mean PM<sub>2.5</sub> measurements of Beijing from 2005 to 2007 were obtained from Zhao et al. (2009). All ground PM<sub>2.5</sub> measurements were made with Tapered Element Oscillating Microbalances (TEOM) or beta-attenuation method, both of which are subject to measurement errors due to loss of semivolatile components (Duncan et al. 2014). However, since PM<sub>2.5</sub> compliance in China is based on these monitors, we used their PM<sub>2.5</sub> measurements to develop and evaluate our model. Our study includes a total of 1,185 monitoring sites in 205 cities or regions (Figure 1). The 2013 data were used for model fitting and cross-validation (CV), while data from other years were used to evaluate the predicted historical PM<sub>2.5</sub> concentrations.

*Satellite data.*

We extracted DT and DB AOD data for the period 01/2004 to 06/2014 at 550 nm from the Aqua MODIS Level 2 aerosol data product, which were downloaded from the Level 1 and Atmospheric Archive and Distribution System (<http://ladsweb.nascom.nasa.gov/>). Aqua MODIS C6 includes an operational combined AOD product calculated from DB and DT AOD in three normalized difference vegetation index (NDVI) categories (Levy et al. 2013). This combined AOD equals DT AOD if  $NDVI > 0.3$ , and DB AOD if  $NDVI < 0.2$ . When  $0.2 \leq NDVI \leq 0.3$ , the combined AOD equals the mean of DT and DB AOD if both values have high quality assurance (QA) flags. If one of the algorithms reports a higher QA than the other, then that AOD value is used. A detailed description of MODIS operational combined AOD algorithm as well as the QA flags can be found elsewhere (Levy et al. 2013). We did not use the operational combined AOD dataset of MODIS C6 aerosol product because it discards all DB AOD data with NDVI values  $> 0.3$  (Supplemental Material, Validation of Aqua MODIS C6 AOD products). We developed a three-step customized approach to combine DT and DB AOD. First, we performed a regression analysis between daily collocated DT and DB AOD. The resulting regression coefficients were then used to predict the missing DB AOD in those pixels with only DT AOD and vice versa (Puttaswamy et al. 2014). Second, level 2 validated AOD observations from 33 Aerosol Robotic Network (AERONET) sites (Supplemental Material, Figure S1) in China were matched with the gap-filled MODIS DT and DB AOD retrievals. The variance of the differences between gap-filled DT (or DB) AOD and AERONET AOD values for each season was calculated. Finally, we combined the gap-filled DT and DB AOD data using the inverse variance weighting (IVW) approach as follows:

$$AOD_c = \frac{AOD_{DT}/Var_{DTm} + AOD_{DB}/Var_{DBm}}{1/Var_{DTm} + 1/Var_{DBm}} \quad [1]$$

where  $AOD_c$  is IVW combined AOD;  $AOD_{DT}$  and  $AOD_{DB}$  are gap-filled DT and DB AOD, respectively; and  $VAR_{DTm}$  and  $VAR_{DBm}$  are the variances of the differences between gap-filled DT and DB AOD and AERONET AOD of season  $m$ , respectively. When compared with the AERONET observations, our combined AOD performs similarly ( $R^2 = 0.80$ , mean bias = 0.07) to MODIS's operational combined AOD ( $R^2 = 0.81$ , mean bias = 0.07), but has 90% greater coverage. Spatially, the improvement in temporal coverage varies by land use type (Figure 2). Coverage for densely populated southern and eastern China improve by 50-100%. The Tibetan plateau has the most improvement (~200%), while the Gobi and Taklamakan Deserts have the least (20-30%).

To account for the impact of fire smoke on  $PM_{2.5}$  levels (Hu et al. 2014c), we downloaded Aqua and Terra MODIS active fire spots from 2004 to 2014 from the NASA Fire Information for Resource Management System (<https://earthdata.nasa.gov/data/near-real-time-data/firms>).

#### *Meteorological and land use data.*

The Goddard Earth Observing System Data Assimilation System GEOS-5 Forward Processing (GEOS 5-FP) (Lucchesi 2013) and GEOS-5.2.0 meteorological data were used in this study. GEOS-5 FP is the latest version of GEOS-5 meteorological data with a spatial resolution of  $0.25^\circ$  latitude  $\times$   $0.3125^\circ$  longitude in a nested grid covering China and has been available since 04/2012. GEOS-5.2.0 is the previous version of GEOS-5 FP and has a resolution of  $0.50^\circ \times 0.666^\circ$ . GEOS-5.2.0 data are available from 01/2004 to 05/2013. We averaged GEOS-5 FP data to GEOS-5.2.0 grid to maintain a consistent spatial resolution across all model years. We used GEOS-5 FP data in 2013 for model development, and GEOS-5.2.0 data from 2004 to 2012 for estimating historical  $PM_{2.5}$  levels. The overlaid period (04/2012-05/2013) of GEOS-5 FP and GEOS-5.2.0 data were used to evaluate the influence of the change in meteorological data source

(Supplemental Material, Comparison of model performance using GEOS-5 FP and GEOS-5.2.0 meteorological data). We extracted planetary boundary layer height (PBLH, 100m), wind speed (WS, m/s) at 10 m above ground, mean relative humidity in PBL (RH\_PBLH, %), and surface pressure (PS, hPa) between 1 pm and 2 pm local time (Aqua satellite overpass time corresponds to 1:30pm local time), as well as cumulative precipitation of the previous day (Precip\_Lag1, mm). Land use variables at 300 m resolution were obtained from the European Space Agency (ESA) Global Land Cover Product (GlobCover, [http://due.esrin.esa.int/page\\_globcover.php](http://due.esrin.esa.int/page_globcover.php)) (Bontemps et al. 2011). We extracted urban and forest cover data from GlobCover 2005-2006 to represent study years 2004 to 2008 and GlobCover 2009 to represent the years of 2009 for 2014.

#### *Data integration.*

We created a 0.1 degree grid (100,699 grid cells in total) for data integration and model development. Ground PM<sub>2.5</sub> data from multiple monitors in each grid cell were averaged. Since the sizes and geographical locations of MODIS AOD pixels vary in space and time, a 0.1 degree grid cell may have multiple AOD pixels (e.g., near the center of each satellite swath) or an AOD pixel may cover multiple 0.1 degree grid cells (e.g., near the edge of each swath). Therefore, Thiessen polygons representing individual MODIS AOD pixels were created and then mapped to the 0.1 degree grid to spatially assign combined AOD values to the grid cells. We interpolated the GEOS-5 FP and GEOS-5.2.0 data to 0.1 degree grid using the inverse distance weighting (IDW) method. We calculated the percentage forest cover and urban area in each grid cell and daily total counts of MODIS fire spots for each grid cell using a 75-km radius buffer. Finally, all the variables in 2013 were matched by grid cell and Day-of-Year (DOY) for model fitting. The model prediction dataset is composed of all spatiotemporally matched variables, except PM<sub>2.5</sub> concentrations from 01/2004 to 06/2014. Before model development, the independent variables

in the fitting and prediction datasets were centered by subtracting their respective mean value computed from the fitting dataset.

*Model development and validation.*

We developed a two-stage statistical model to calibrate the spatiotemporal relationships between  $PM_{2.5}$  and AOD. The first-stage linear mixed effect (LME) model includes day-specific random intercepts and slopes for AOD and season-specific random slopes for meteorological variables:

$$\begin{aligned} PM_{2.5,st} = & (\mu + \mu') + (\beta_1 + \beta_1')AOD_{st} + (\beta_2 + \beta_2')WS_{st} + (\beta_3 + \beta_3')PBLH_{st} + (\beta_4 + \beta_4')PS_{st} + \\ & (\beta_5 + \beta_5')RH\_PBLH_{st} + \beta_6Precip\_Lag1_{st} + \beta_7Fire\_spots_{st} + \varepsilon_{1,st}(\mu', \beta_1') \sim N[(0,0), \\ & \Psi_1] + \varepsilon_{2,sj}(\beta_2', \beta_3', \beta_4', \beta_5') \sim N[(0,0,0,0), \Psi_2] \end{aligned} \quad [2]$$

where  $PM_{2.5,st}$  is the average observed  $PM_{2.5}$  concentration at grid cell  $s$  on DOY  $t$ ;  $AOD_{st}$  is IWV combined AOD;  $WS_{st}$ ,  $PBLH_{st}$ ,  $PS_{st}$ ,  $RH\_PBLH_{st}$ ,  $Precip\_Lag1_{st}$  are meteorological variables;  $Fire\_spots_{st}$  is the fire count;  $\mu$  and  $\mu'$  are the fixed and day-specific random intercepts, respectively;  $\beta_1$ - $\beta_7$  are fixed slopes for independent variables;  $\beta_1'$  is the day-specific random slope for AOD;  $\beta_2'$ - $\beta_5'$  are the season-specific random slopes for meteorological variables;  $\varepsilon_{1,st}$  is the error term at grid cell  $s$  on day  $t$ ;  $\varepsilon_{2,sj}$  is the error term at grid cell  $s$  in season  $j$ ; and  $\Psi_1$  and  $\Psi_2$  are the variance-covariance matrices for the day- and season-specific random effects, respectively. In addition to modeling season-specific meteorological random effects, we tested alternative models with day- and month-specific random effects for meteorological variables and found that this may cause over-fitting (data not shown).

We fitted the first-stage model for each province separately. Because the provinces in western China (e.g., Tibet, Xinjiang, and Qinghai) do not have enough  $PM_{2.5}$  monitoring sites (Figure 1) to produce a robust model-fitting dataset, we created a buffer zone for each province to include

at least 3,000 data records and at least 300 days in 2013. We averaged overlapped predictions from neighboring provinces to generate a smooth national  $PM_{2.5}$  concentration surface.

The second-stage generalized additive model (GAM) is expressed as follows:

$$PM_{2.5\_resid_{st}} = \mu_0 + s(X, Y)_s + s(ForestCover)_s + s(UrbanCover)_s + \varepsilon_{st} \quad [3]$$

where  $PM_{2.5\_resid_{st}}$  is the residual from the first-stage model at grid cell  $s$  on day  $t$ ;  $\mu_0$  is the intercept term;  $s(X, Y)_s$  is the smooth term of the coordinates of the centroid of grid cell  $s$ ;  $s(ForestCover)_s$  and  $s(UrbanCover)_s$  are the smooth functions of percent forest cover and urban area for grid cell  $s$ ; and  $\varepsilon_{st}$  is the error term.

Statistical indicators, such as the coefficient of determination ( $R^2$ ), mean prediction error (MPE), root mean squared prediction error (RMSE), and relative prediction error (RPE, defined as RMSE divided by the mean ground  $PM_{2.5}$ ), were calculated and compared between model fitting and cross-validation to assess model performance and test the potential model over-fitting.

#### *Prediction, evaluation, and time series analysis of historical $PM_{2.5}$ .*

The historical daily  $PM_{2.5}$  concentrations (2004-2012) were estimated using the model developed in the current study based on 2013 data, assuming that the daily relationship between  $PM_{2.5}$  and AOD is constant for the same DOY in each year. Since there were few ground  $PM_{2.5}$  measurements for Mainland China before 2013, we estimated daily  $PM_{2.5}$  concentrations in the first half of 2014 using the model established in 2013 and compared them with the ground measurements to validate the accuracy of the historical  $PM_{2.5}$  estimations. We evaluated historical  $PM_{2.5}$  predictions (including 2014) at the daily, monthly, and seasonal scales. Since AOD-derived  $PM_{2.5}$  estimates are always missing due to the cloud and snow surface, we conducted a sensitivity analysis to test how many AOD-derived  $PM_{2.5}$  estimations can represent

the true monthly and seasonal mean PM<sub>2.5</sub> concentrations. We required each evaluation grid cell to have at least 25 PM<sub>2.5</sub> ground measurements in a given month to calculate the monthly mean PM<sub>2.5</sub> concentrations and at least 25 measurements in each month of a season to calculate the seasonal mean PM<sub>2.5</sub> concentrations.

We calculated the monthly mean PM<sub>2.5</sub> anomaly time series by subtracting the 10-year average PM<sub>2.5</sub> concentration of the corresponding month for each grid cell and analyzed the PM<sub>2.5</sub> trend for each grid cell using least squares regression (Weatherhead et al. 1998), which has been applied to global analyses of monthly mean AOD anomaly time-series data (Hsu et al. 2012). For each grid cell, we required at least six daily PM<sub>2.5</sub> predictions in each month to calculate the monthly mean PM<sub>2.5</sub> and at least six months of anomaly data per year in order to be included in the time-series analysis.

The work flow of estimating the spatiotemporal PM<sub>2.5</sub> concentrations in this study is shown in Figure 3.

## **Results**

### *Descriptive statistics of the model fitting dataset.*

There are 63,031 data records included in the final 2013 model-fitting dataset. The overall mean PM<sub>2.5</sub> concentration is 77.05 µg/m<sup>3</sup>, and the mean value of our combined AOD is 0.69 (Supplemental Material, Table S3). These results are approximately five times higher than those found in the eastern and southeastern U.S. (Hu et al. 2013; Liu et al. 2005).

### *Results of model fitting and cross-validation.*

We summarized the fixed effect estimates, model fitting, and CV results of the first-stage LME model for each province in Table S4 (Supplemental Material). AOD is the only variable that is

statistically significant in all provincial models ( $p < 0.05$ ). Wind speed, relative humidity, and precipitation are significant in most provincial models. Fire spots are not significant in some provinces, probably because these regions have infrequent fire activity. The CV  $R^2$  values of the first-stage LME model range from 0.64 in Ningxia to 0.82 in Zhejiang. The spatial distribution of first-stage LME CV  $R^2$  (not shown here) indicates that our LME model generally performs better in South, East, North, and Northeast China than in West and Northwest China, which have fewer  $PM_{2.5}$  monitoring networks (Figure 1).

Figure 4 shows the model-fitting and CV results for the first-stage and full models. The full model fitting and CV  $R^2$  values are 0.82 (Figure 4B) and 0.79 (Figure 4D), respectively, indicating that this model is not substantially over-fitted. Comparing the first-stage (Figure 4A and C) with the full model (Figure 4B and D), it is clear that the second-stage GAM model marginally increases  $R^2$  values. However, the GAM model does increase the slope (from 0.77, Figure 4C to 0.79, Figure 4D) and reduce the intercept (from 18.38, Figure 4C to 16.57  $\mu\text{g}/\text{m}^3$ , Figure 4D) of the linear regression between model-estimated and observed  $PM_{2.5}$  concentrations in 2013. More importantly,  $PM_{2.5}$  levels in the Hebei Province predicted by the full model are approximately 20  $\mu\text{g}/\text{m}^3$  higher than those predicted by the first-stage model; the predicted  $PM_{2.5}$  levels in Tibet are about 15  $\mu\text{g}/\text{m}^3$  lower in the full model (Supplemental Material, Figure S4), showing that the spatial pattern of the predicted  $PM_{2.5}$  levels by the full model is more consistent with that of the ground observations.

#### *Evaluation of historical $PM_{2.5}$ predictions*

Although our model predictions for daily level observations were poor compared to the historical observations ( $R^2=0.41$ ,  $N=79,989$ ) (Figure 5A), it performed much better at the monthly and seasonal levels (Figure 5B and C, respectively). The sensitivity analysis showed that more daily

predictions yield more accurate monthly or seasonal estimations (Supplemental Material, Figure S5). Figure 5B shows that the monthly mean satellite  $PM_{2.5}$  calculated from more than five predicted daily  $PM_{2.5}$  concentrations can be a fairly accurate ( $R^2 = 0.73$ ) representation of monthly  $PM_{2.5}$  level measured from ground observations with only a slight bias (regression slope = 0.91). This threshold of six days per month is consistent with the method of a previous global AOD trend study (Hsu et al. 2012). At the seasonal level (Figure 5C), satellite  $PM_{2.5}$  calculated from more than 10 predicted daily  $PM_{2.5}$  concentrations can be a more accurate ( $R^2 = 0.79$ ) representation of seasonal  $PM_{2.5}$  levels with little bias (regression slope = 0.92).

*Spatial and temporal  $PM_{2.5}$  concentration trends.*

Figure 6 shows the spatial patterns of 10-year mean  $PM_{2.5}$  estimations (2004-2013) of China and four sub-regions (including the Beijing-Tianjin Metropolitan Region, Yangtze River Delta, Pearl River Delta, and Sichuan Basin). The highest  $PM_{2.5}$  estimations appear in the Beijing-Tianjin Metropolitan Region (including Beijing, Tianjin, and Hebei), followed by the Sichuan Basin, Yangtze River Delta (including Jiangsu, Shanghai, and Anhui), and Pearl River Delta. The 10-year mean  $PM_{2.5}$  estimations in the Beijing-Tianjin Metropolitan Region are generally higher than  $100 \mu\text{g}/\text{m}^3$ , with the highest concentrations greater than  $120 \mu\text{g}/\text{m}^3$ . Similarly, the 10-year mean  $PM_{2.5}$  concentrations are generally higher than  $85 \mu\text{g}/\text{m}^3$  in the Sichuan Basin, Yangtze River Delta. The mean  $PM_{2.5}$  concentrations are generally higher than  $55 \mu\text{g}/\text{m}^3$  in the Pearl River Delta. High  $PM_{2.5}$  levels also occur Taklamakan Desert in Xinjiang, an area that is a major dust source (Figure 6A). Figure S6 (Supplemental Material) illustrates the seasonal patterns of the 10-year mean  $PM_{2.5}$  concentrations in China. Winter is the most polluted season (mean  $PM_{2.5}$ :  $72.24 \mu\text{g}/\text{m}^3$ ), while summer is the cleanest season ( $32.90 \mu\text{g}/\text{m}^3$ ).

Figure 7 shows that while China has an overall positive 10-year average  $PM_{2.5}$  trend of 0.22

$\mu\text{g}/\text{m}^3$  per year (Figure 7A), there is significant regional variability. For example, the Beijing-Tianjin Metropolitan Region has had more rapid increases ( $0.75 \mu\text{g}/\text{m}^3$  per year) than the rest of the nation (Figure 7B), whereas the Pearl River Delta has experienced a rapid decrease ( $0.96 \mu\text{g}/\text{m}^3$  per year) (Figure 7D). The  $\text{PM}_{2.5}$  level in the Yangtze River Delta region has remained steady (Figure 7C). In addition,  $\text{PM}_{2.5}$  levels in most of China has been increasing by  $1.97 \mu\text{g}/\text{m}^3$  per year before 2008, but decreasing by  $0.46 \mu\text{g}/\text{m}^3$  per year afterwards (Figures 7A, 7E and 7F). Similar trends were observed in the Beijing-Tianjin Metropolitan Region (Figure 7B). The  $\text{PM}_{2.5}$  level has remained relatively constant in the Pearl River Delta for 2004-2007, followed by a negative trend of  $1.53 \mu\text{g}/\text{m}^3$  per year after 2008 (Figure 7D).

## Discussion

Compared with our previous GWR model ( $\text{CV } R^2 = 0.64$ ) (Ma et al. 2014), the current two-stage model demonstrated superior performance ( $\text{CV } R^2 = 0.79$ ). The CV RPE decreased from 51.3% (Ma et al. 2014) to 35.6% (this study), approaching results seen in regional-scale studies conducted in the U.S. (Hu et al. 2014a; Lee et al. 2011). This improvement is particularly encouraging for our national model because, unlike regional-scale models, the  $\text{PM}_{2.5}$ -AOD relationship will inevitably vary in space (e.g., variable  $\text{PM}_{2.5}$  composition and vertical distribution caused by different emission sources; synoptic weather patterns vary by province). The first-stage  $\text{CV } R^2$  drops to 0.63 if a single LME model is fitted for the whole domain, further illustrating that a constant daily  $\text{PM}_{2.5}$ -AOD relationship is a valid assumption only for relatively small geographic regions. Using both MODIS C6 DT and DB AOD products to obtain a custom combined AOD yielded a 25-fold increase in spatial resolution (from 50 km to 10 km) and greatly improved the AOD data coverage. There are 120% more matched DB AOD values than DT AOD values when comparing to AERONET observations (Supplemental Material, Figure

S2). Furthermore, our analysis indicated that DB AOD has a smaller mean bias overall than DT AOD in China (0.01~0.05 vs. 0.13~0.18) (Supplemental Material, Figure S2), enabling us to estimate lower PM<sub>2.5</sub> levels.

To our best knowledge, this is the first national-scale study in China to estimate and evaluate historical PM<sub>2.5</sub> levels in the years beyond the modeling year using advanced statistical models. The poor agreement between daily historical PM<sub>2.5</sub> predictions and ground measurements is caused by the strong model assumption that the daily PM<sub>2.5</sub>-AOD relationship derived from 2013 data remained constant for the same DOY in each year. This limitation of our model cannot be resolved without sufficient historical PM<sub>2.5</sub> data which would allow annual model adjustments before 2013. Nonetheless, our monthly ( $R^2 = 0.73$ , slope = 0.91) and seasonal ( $R^2 = 0.79$ , slope = 0.92) mean PM<sub>2.5</sub> predictions are accurate representations of the ground measurements with relatively low biases and can serve as exposure estimates to study the health impacts of long-term PM<sub>2.5</sub> exposure in China. The seasonal patterns show that the most polluted season is winter and the cleanest one is summer, which is consistent with the results of our previous study (Ma et al. 2014). Looking forward, this model can be fitted every year after 2013 to provide accurate daily PM<sub>2.5</sub> concentrations and fill the spatial gaps left by the monitoring network.

Two approaches (including statistical and scaling models) can be applied to retrieve ground PM<sub>2.5</sub> levels from satellite remotely sensed AOD data (Liu 2014). For statistical models to function properly, substantial ground data support is necessary. With the recently established ground monitoring network, we are able to develop this high-performance spatial model in China. The same model cannot be applied in regions with sparse or no ground observations. In this case the scaling approach described in Brauer et al. (2012) is the only applicable method.

We compared the 9-year (2005-2013) AOD-derived and ground-measured PM<sub>2.5</sub> trends in Hong

Kong (no PM<sub>2.5</sub> monitoring sites in 2004) and Taiwan (few sites in 2004). The results show that the AOD-derived PM<sub>2.5</sub> trend in Hong Kong (-1.28 μg/m<sup>3</sup> per year) is similar to ground measurements (-1.35 μg/m<sup>3</sup> per year). However, the trend of AOD-derived PM<sub>2.5</sub> in Taiwan is -0.17 μg/m<sup>3</sup> per year, which is much higher than that of ground measurements (-0.72 μg/m<sup>3</sup> per year). This inconsistency is probably due to missing satellite AOD retrievals. For example, only 34.5% of the grid cells in Taiwan have more than 50% of months which have AOD-derived PM<sub>2.5</sub> data. Missing AOD values are a major limitation and challenge in PM<sub>2.5</sub>-AOD modeling (Liu 2014) and the methods to account for the missing AOD data in China will be a focus of our future research.

Nonetheless, the overall regional trends are consistent with the environmental policy and regulation change in China. We found an inflection point around 2008 for the monthly mean PM<sub>2.5</sub> time series. The PM<sub>2.5</sub> level steadily increased between 2004 and 2007, but the trend reversed or became non-significant after 2008, especially in the Beijing-Tianjin Metropolitan Region. A recent study (Boys et al. 2014) also found that PM<sub>2.5</sub> levels steadily rose until 2007 and then became stable in East Asia. China experienced a rapid growth of energy consumption before 2005 (Yuan et al. 2011), resulting in missed environmental quality targets between 2001-2005 (Xue et al. 2014). The growth in energy demand led to a stricter energy conservation and emissions reduction (ECER) policy, which required 20% reduction in energy usage intensity by the end of 2010, compared with the level in 2005 (Lo and Wang 2013). The ECER policy was implemented in late 2006, and the overall achievement obtained by 2010 was 19.06% (Lo and Wang 2013). A recent study also showed that the production related PM<sub>2.5</sub> emissions of China peaked at around 2007 and quickly dropped afterwards (Guan et al. 2014). A sharp reduction of PM<sub>2.5</sub> levels induced by this ECER policy may explain the inflection point.

## **Conclusions**

The two-stage satellite AOD model developed in this study can generate reliable historical monthly and seasonal PM<sub>2.5</sub> predictions at 10 km resolution in China with little bias, including data from the past decade, when the regulatory PM<sub>2.5</sub> monitoring network did not exist. Since several long-term PM<sub>2.5</sub> health effects studies in North America and the Global Burden of Disease project are driven by satellite exposure estimates at this resolution (Brauer et al. 2012; Crouse et al. 2012; Madrigano et al. 2013), our model predictions can greatly enhance the research of long-term PM<sub>2.5</sub> health effects in China. With the release of Terra MODIS C6 product in early 2015, the predicted historical PM<sub>2.5</sub> time series can now be extended to early 2000, if consistent meteorological and land use parameters are found to cover 2000 to 2003. From 2013 onward, our model can provide daily PM<sub>2.5</sub> exposure estimates to fill the gaps left by the PM<sub>2.5</sub> monitoring network in China. Finally, given the wider dynamic range of PM<sub>2.5</sub> concentrations in China compared to North America, likely due to intensive local sources, it is possible to further improve the performance of our model with detailed land use (e.g., road network) and emissions information (e.g., major point sources), which were not available during this study (Kloog et al. 2014).

## References

- Bontemps S, Defourny P, Bogaert EV, Arino O, Kalogirou V, Perez JR. 2011. GlobCover 2009-Products description and validation report. Available from [http://due.esrin.esa.int/page\\_globcover.php](http://due.esrin.esa.int/page_globcover.php) [accessed 9 July 2015].
- Boys B, Martin R, van Donkelaar A, MacDonell R, Hsu C, Cooper M, et al. 2014. Fifteen-year global time series of satellite-derived fine particulate matter. *Environ Sci Technol* 48:11109-11118.
- Brauer M, Amann M, Burnett RT, Cohen A, Dentener F, Ezzati M, et al. 2012. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. *Environ Sci Technol* 46:652-660.
- Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, et al. 2012. Risk of non accidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A Canadian national-level cohort study. *Environ Health Perspect* 120:708-714.
- Duncan BN, Prados AI, Lamsal LN, Liu Y, Streets DG, Gupta P, et al. 2014. Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmos Environ* 94:647-662.
- Guan D, Su X, Zhang Q, Peters GP, Liu Z, Lei Y, et al. 2014. The socioeconomic drivers of China's primary PM<sub>2.5</sub> emissions. *Environ Res Lett* 9:024010.
- Hsu N, Jeong MJ, Bettenhausen C, Sayer A, Hansell R, Seftor C, et al. 2013. Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *J Geophys Res Atmos* 118:9296-9315.
- Hsu NC, Gautam R, Sayer AM, Bettenhausen C, Li C, Jeong MJ, et al. 2012. Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010. *Atmos Chem Phys* 12:8037-8053.
- Hu X, Waller LA, Al-Hamdan MZ, Crosson WL, Estes Jr MG, Estes SM, et al. 2013. Estimating ground-level PM<sub>2.5</sub> concentrations in the southeastern US using geographically weighted regression. *Environ Res* 121:1-10.
- Hu X, Waller LA, Lyapustin A, Wang Y, Al-Hamdan MZ, Crosson WL, et al. 2014a. Estimating Ground-Level PM<sub>2.5</sub> Concentrations in the Southeastern United States Using MAIAC AOD Retrievals and a Two-Stage Model. *Remote Sens Environ* 140:220-232.
- Hu X, Waller LA, Lyapustin A, Wang Y, Liu Y. 2014b. 10-year spatial and temporal trends of PM<sub>2.5</sub> concentrations in the southeastern US estimated using high-resolution satellite data. *Atmos Chem Phys* 14:6301-6314.

- Hu X, Waller LA, Lyapustin A, Wang Y, Liu Y. 2014c. Improving Satellite-Driven PM<sub>2.5</sub> Models with MODIS Fire Counts in the Southeastern U.S. *J Geophys Res Atmos*:2014JD021920.
- Kloog I, Koutrakis P, Coull BA, Lee HJ, Schwartz J. 2011. Assessing temporally and spatially resolved PM<sub>2.5</sub> exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos Environ* 45:6267-6275.
- Kloog I, Chudnovsky AA, Just AC, Nordio F, Koutrakis P, Coull BA, et al. 2014. A new hybrid spatio-temporal model for estimating daily multi-year PM<sub>2.5</sub> concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmos Environ* 95:581-590.
- Lee HJ, Liu Y, Coull BA, Schwartz J, Koutrakis P. 2011. A novel calibration approach of MODIS AOD data to predict PM<sub>2.5</sub> concentrations. *Atmos Chem Phys* 11:7991-8002.
- Levy R, Mattoo S, Munchak L, Remer L, Sayer A, Patadia F, et al. 2013. The Collection 6 MODIS aerosol products over land and ocean. *Atmos Meas Tech* 6:2989-3034.
- Liu Y, Sarnat JA, Kilaru V, Jacob DJ, Koutrakis P. 2005. Estimating ground-level PM<sub>2.5</sub> in the eastern United States using satellite remote sensing. *Environ Sci Technol* 39:3269-3278.
- Liu Y, Paciorek CJ, Koutrakis P. 2009. Estimating Regional Spatial and Temporal Variability of PM<sub>2.5</sub> Concentrations Using Satellite Data, Meteorology, and Land Use Information. *Environ Health Perspect* 117:886-892.
- Liu Y. 2014. Monitoring PM<sub>2.5</sub> from space for health: past, present, and future directions. *Environ Manager* 6:6-10.
- Lo K, Wang MY. 2013. Energy conservation in China's Twelfth Five-Year Plan period: Continuation or paradigm shift? *Renew Sust Energ Rev* 18:499-507.
- Lucchesi R. 2013. File Specification for GEOS-5 FP. GMAO Office Note No. 4 (Version 1.0). Available from [http://gmao.gsfc.nasa.gov/pubs/office\\_notes](http://gmao.gsfc.nasa.gov/pubs/office_notes) [accessed 9 July 2015].
- Ma Z, Hu X, Huang L, Bi J, Liu Y. 2014. Estimating ground-level PM<sub>2.5</sub> in China using satellite remote sensing. *Environ Sci Technol* 48:7436-7444.
- Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J. 2013. Long-term exposure to PM<sub>2.5</sub> and incidence of acute myocardial infarction. *Environ Health Perspect* 121:192-196.
- Pope, Dockery D. 2006. Health effects of fine particulate air pollution: Lines that connect. *J Air Waste Manage Assoc* 56:709-742.
- Puttaswamy SJ, Nguyen HM, Braverman A, Hu X, Liu Y. 2014. Statistical data fusion of multi-sensor AOD over the Continental United States. *Geocarto Int* 29:48-64.

- Remer LA, Kaufman Y, Tanré D, Mattoo S, Chu D, Martins J, et al. 2005. The MODIS aerosol algorithm, products, and validation. *J Atmos Sci* 62:947-973.
- van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, et al. 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ Health Perspect* 118:847.
- van Donkelaar A, Martin RV, Brauer M, Boys BL. 2015. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environ Health Perspect* 132:135-143.
- Weatherhead EC, Reinsel GC, Tiao GC, Meng XL, Choi D, Cheang WK, et al. 1998. Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J Geophys Res Atmos* 103:17149-17161.
- Xu P, Chen Y, Ye X. 2013. Haze, air pollution, and health in China. *The Lancet* 382:2067.
- Xue B, Mitchell B, Geng Y, Ren W, Müller K, Ma Z, et al. 2014. A review on China's pollutant emissions reduction assessment. *Ecol Indic* 38:272-278.
- Yao L, Lu N. 2014. Spatiotemporal distribution and short-term trends of particulate matter concentration over China, 2006–2010. *Environ Sci Pollut Res*:1-11.
- Yuan J, Kang J, Yu C, Hu Z. 2011. Energy conservation and emissions reduction in China—Progress and prospective. *Renew Sust Energ Rev* 15:4334-4347.
- Zhao X, Zhang X, Xu X, Xu J, Meng W, Pu W. 2009. Seasonal and diurnal variations of ambient PM<sub>2.5</sub> concentration in urban and rural environments in Beijing. *Atmos Environ* 43:2893-2900.

## Figure Legends

**Figure 1.** Spatial distribution of ground PM<sub>2.5</sub> monitoring sites. Hollow circles denote the sites with data only available from 01/2014 to 06/2014. Solid circles denote the sites with data available not only in 2014 but also 2013 or earlier years. Note that many clustered sites are overlapped due to their proximity. The spatial resolution of the background gridded population is 0.1° × 0.1°.

**Figure 2.** Spatial distribution of annual mean available days for MODIS's operational combined AOD (A), our IVW combined AOD (B), and percentage improvement of data coverage (C).

**Figure 3.** The work flow of estimating the spatiotemporal PM<sub>2.5</sub> concentrations in this study.

**Figure 4.** Density scatter plots of model-fitting and cross-validation at daily level (N=63,031). (A) and (B) are model fitting results for first-stage LME model and full LME+GAM model, respectively. (C) and (D) are model CV results for first-stage LME model and full LME+GAM model, respectively. MPE: mean prediction error (μg/m<sup>3</sup>). RMSE: root mean squared prediction error (μg/m<sup>3</sup>). RPE: relative prediction error (%). The dashed line is the 1:1 line.

**Figure 5.** Evaluation of historical PM<sub>2.5</sub> estimations (2004-2012 and 01/2014-06/2014) at daily (A), monthly (B), and seasonal (C) levels. Since there were few ground PM<sub>2.5</sub> data for Mainland China before 2013, we also estimated PM<sub>2.5</sub> in the first half of 2014 using the 2013 model and compared them with the ground measurements to validate the accuracy of the historical estimations.

**Figure 6.** Spatial distributions of 10-year (2004-2013) mean PM<sub>2.5</sub> estimations for entire China (A), Beijing-Tianjin Metropolitan Region (B), Yangtze River Delta (C), Pearl River Delta (D), and Sichuan Basin (E).

**Figure 7.** Time series of monthly, satellite-derived PM<sub>2.5</sub> anomaly (μg/m<sup>3</sup>) for entire China (A), Beijing-Tianjin Metropolitan region (B), Yangtze Delta (C), and Pearl River Delta (D); and

spatial distribution of PM<sub>2.5</sub> trends for 2004-2007 (*E*) and 2008-2013 (*F*). White areas in (*E*) and (*F*) indicate missing data. The black lines in (*A*)-(*D*) denote the PM<sub>2.5</sub> trends for 2004-2013, while red lines represent the trends of 2004-2007 and blue lines represent the trends of 2008-2013. The PM<sub>2.5</sub> trends ( $\mu\text{g}/\text{m}^3$  per year), 95% confidence interval (CI) in brackets ( $\mu\text{g}/\text{m}^3$  per year), and significance levels (\*:  $p<0.05$ ; \*\*:  $p<0.01$ ; \*\*\*:  $p<0.005$ ) are also shown in (*A*)-(*D*).

Figure 1.

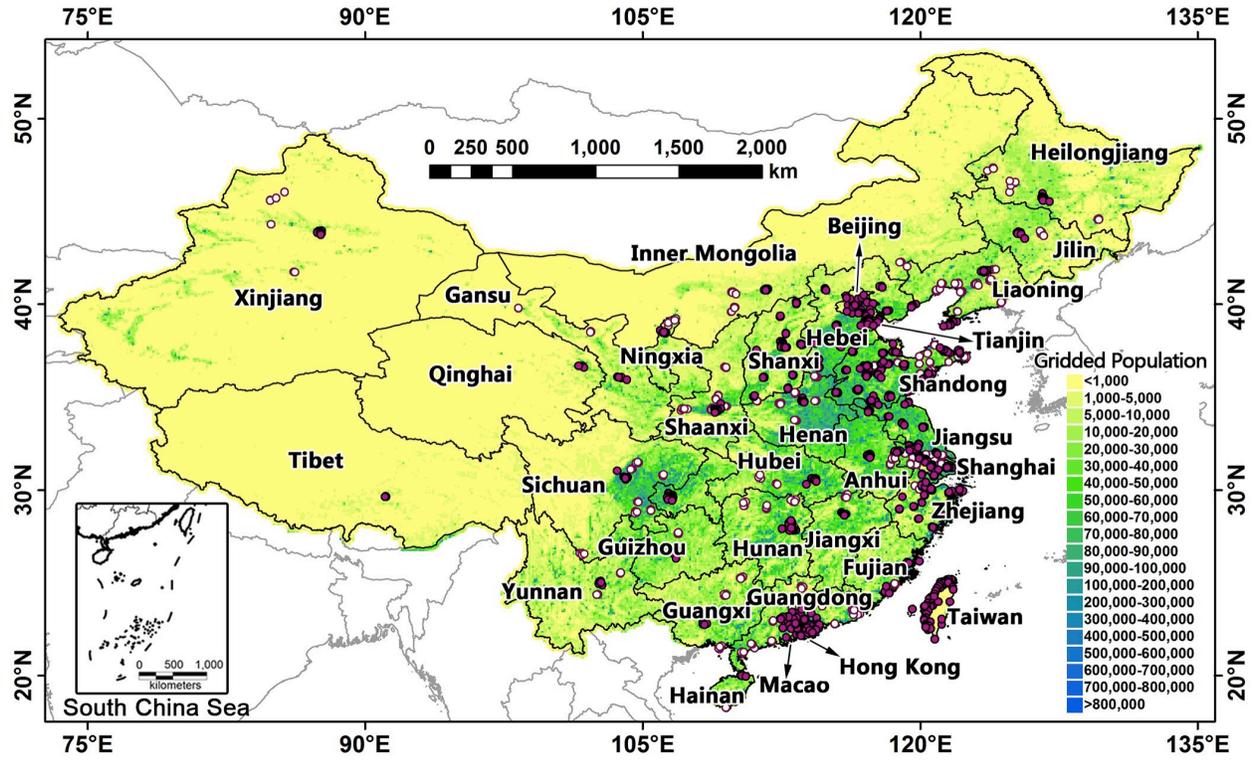


Figure 2.

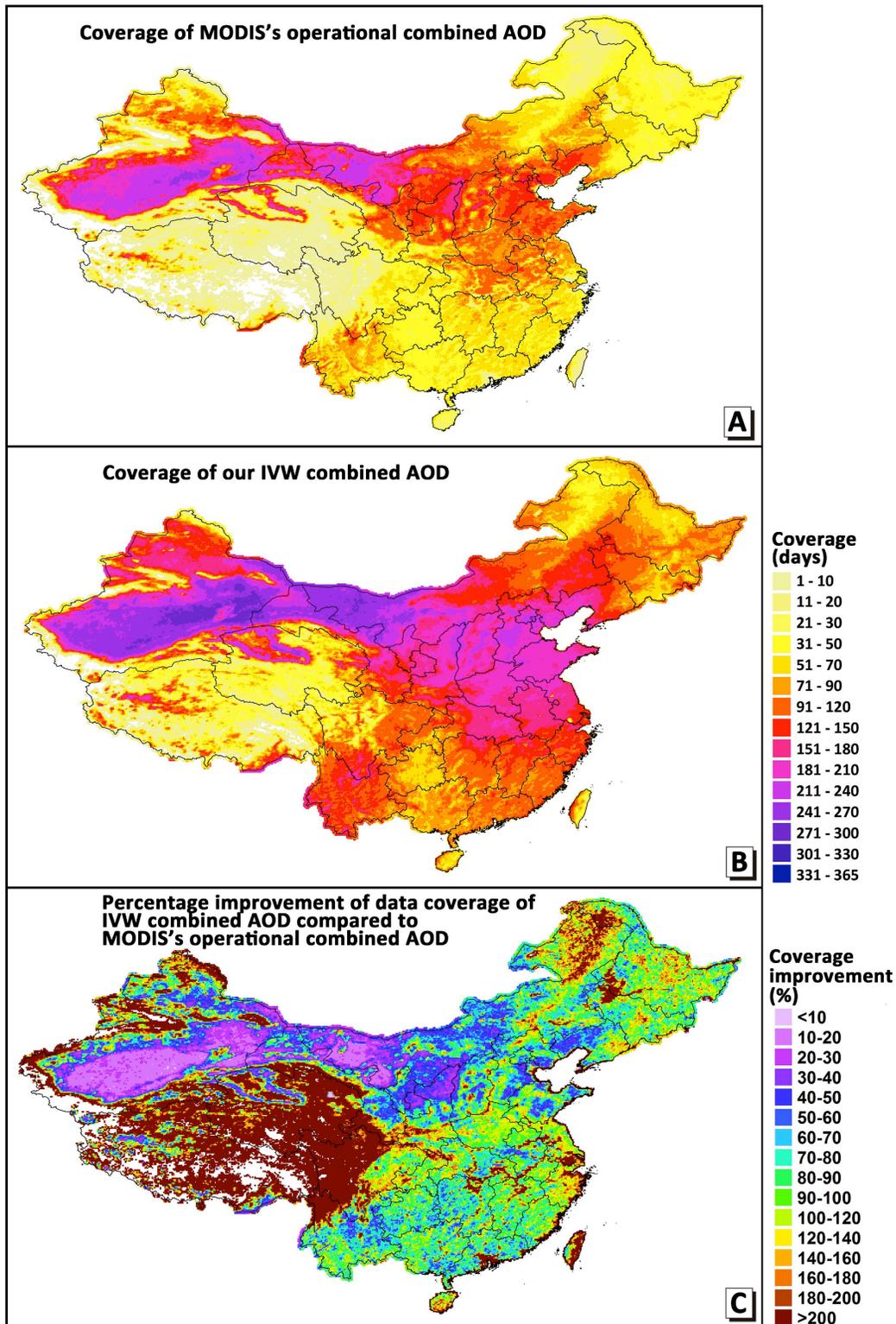


Figure 3.

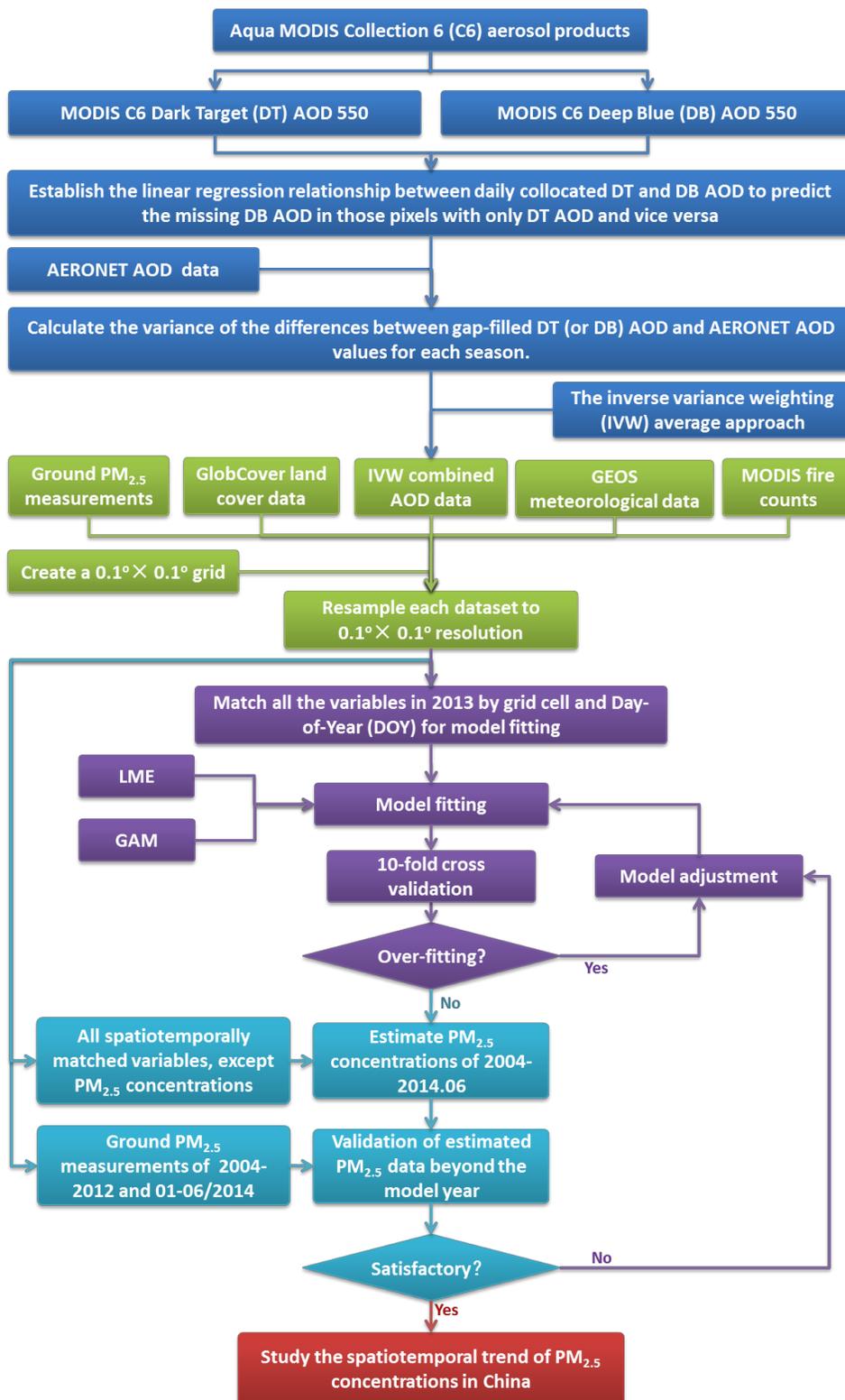


Figure 4.

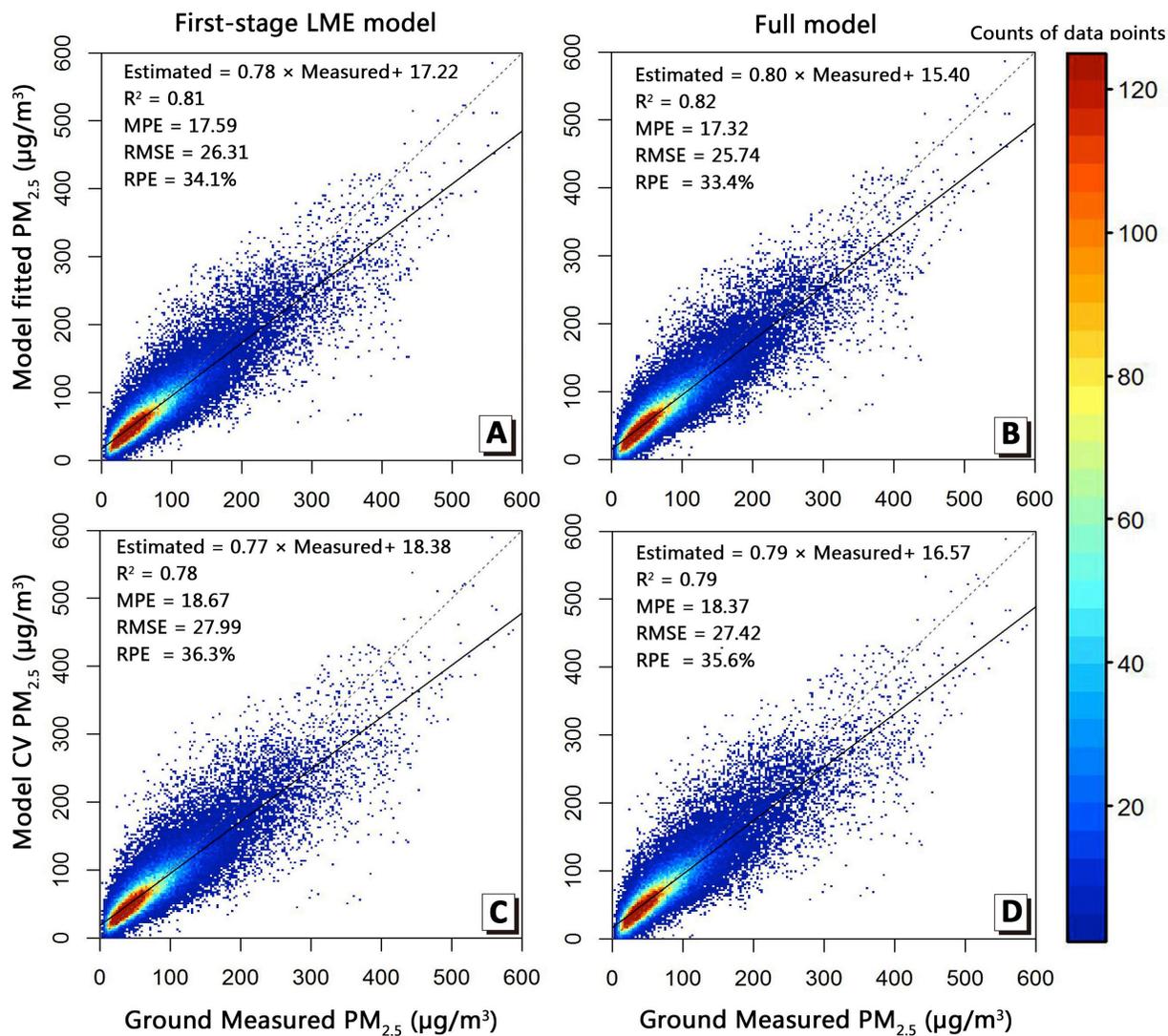


Figure 5.

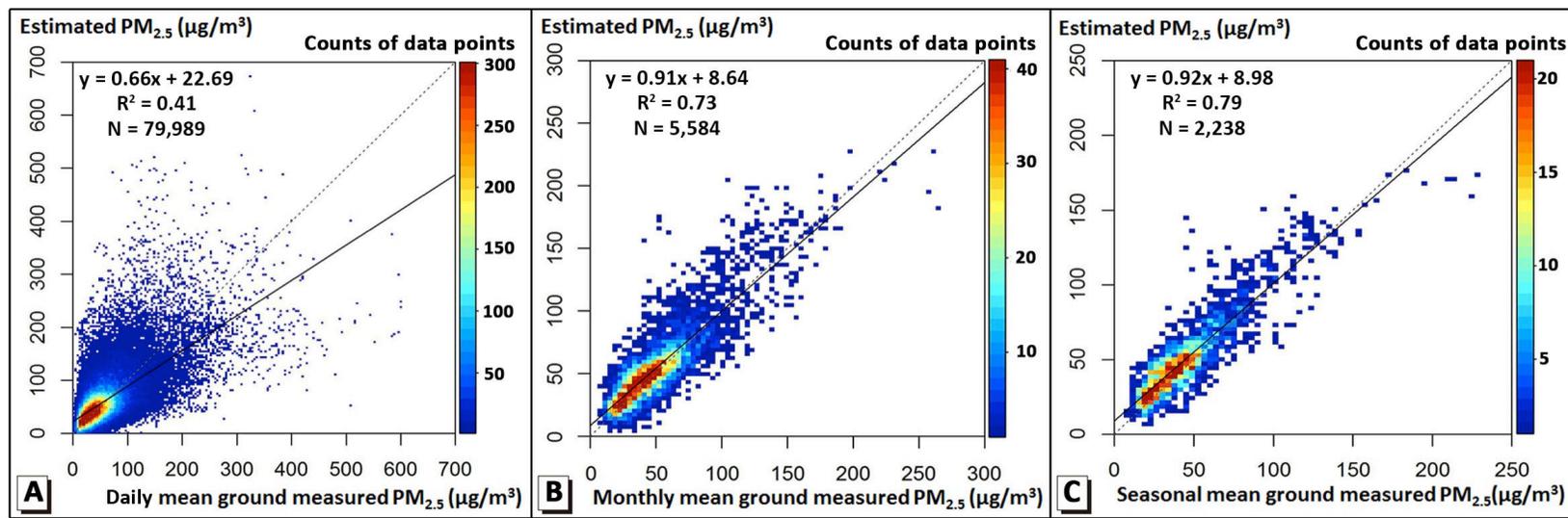


Figure 6.

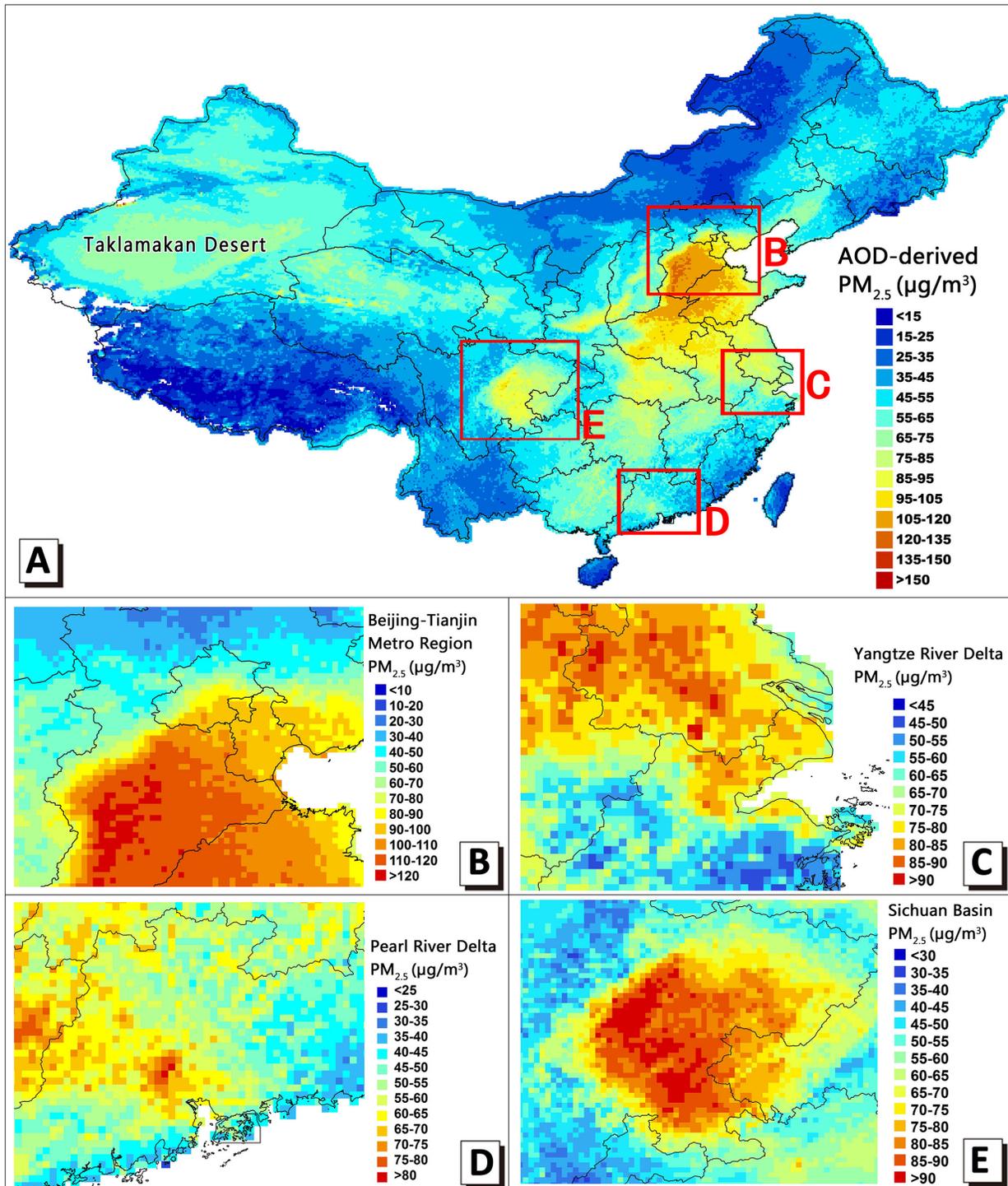


Figure 7.

