
Zongwei Ma, Emory University  
Xuefei Hu, Emory University  
Andrew M. Sayer, Universities Space Research Association  
Robert Levy, NASA Goddard Space Flight Center  
Qiang Zhang, Tsinghua University  
Yingang Xue, Changzhou Environmental Monitoring Center  
Shilu Tong, Queensland University of Technology  
Jun Bi, Nanjing University  
Lei Huang, Nanjing University  
Yang Liu, Emory University

Journal Title: Environmental Health Perspectives  
Volume: Volume 124, Number 2  
Publisher: National Institute of Environmental Health Sciences (NIEHS) | 2016-02, Pages 184-192  
Type of Work: Article | Final Publisher PDF  
Publisher DOI: 10.1289/ehp.1409481  
Permanent URL: https://pid.emory.edu/ark:/25593/rk1vg

Final published version: http://dx.doi.org/10.1289/ehp.1409481

Copyright information:
Publication of EHP lies in the public domain and is therefore without copyright. All text from EHP may be reprinted freely. Use of materials published in EHP should be acknowledged (for example, “Reproduced with permission from Environmental Health Perspectives”); pertinent reference information should be provided for the article from which the material was reproduced. Articles from EHP, especially the News section, may contain photographs or illustrations copyrighted by other commercial organizations or individuals that may not be used without obtaining prior approval from the holder of the copyright.

Accessed October 17, 2017 5:55 AM EDT
Satellite-Based Spatiotemporal Trends in PM$_{2.5}$ Concentrations: China, 2004–2013

Zongwei Ma,$^{1,2}$ Xuefei Hu,$^2$ Andrew M. Sayer,$^{3,4}$ Robert Levy,$^4$ Qiang Zhang,$^5$ Yingang Xue,$^6$ Shilu Tong,$^7$ Jun Bi,$^1$ Lei Huang,$^1$ and Yang Liu$^2$

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

**Background:** Three decades of rapid economic development is causing severe and widespread PM$_{2.5}$ (particulate matter ≤ 2.5 μm) pollution in China. However, research on the health impacts of PM$_{2.5}$ exposure has been hindered by limited historical PM$_{2.5}$ concentration data.

**Objectives:** We estimated ambient PM$_{2.5}$ concentrations from 2004 to 2013 in China at 0.1° resolution using the most recent 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) and assimilated meteorology, land use data, and PM$_{2.5}$ 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$_{2.5}$ concentrations from 2004 to early 2014 using ground observations.

**Results:** The overall model cross-validation $R^2$ and relative prediction error were 0.79 and 35.6%, respectively. Validation beyond the model year (2013) indicated that it accurately predicted PM$_{2.5}$ concentrations with little bias at the monthly ($R^2 = 0.73$, regression slope = 0.91) and seasonal ($R^2 = 0.79$, regression slope = 0.92) levels. Seasonal variations revealed that winter was the most polluted season and that summer was the cleanest season. Analysis of predicted PM$_{2.5}$ levels showed a mean annual increase of 1.97 μg/m$^3$ between 2004 and 2007 and a decrease of 0.46 μg/m$^3$ between 2008 and 2013.

**Conclusions:** Our satellite-driven model can provide reliable historical PM$_{2.5}$ estimates in China at a resolution comparable to those used in epidemiologic studies on the health effects of long-term PM$_{2.5}$ exposure in North America. This data source can potentially advance research on PM$_{2.5}$ health effects in China.


Introduction

Fine particulate matter (PM$_{2.5}$; particles with aerodynamic diameter ≤ 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 occurring in China, severe, widespread PM$_{2.5}$ pollution has attracted nationwide attention (Xu et al. 2013). However, research on the adverse health impacts of PM$_{2.5}$ exposure has been hindered because a nationwide regulatory PM$_{2.5}$ monitoring network did not exist until the end of 2012.

Estimating ground-level PM$_{2.5}$ from satellite-retrieved aerosol optical depth (AOD) data is a promising new method that has advanced rapidly in recent years (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$_{2.5}$ gaps left by ground monitors with high-quality predictions. Several recent studies of the health effects caused by long-term PM$_{2.5}$ exposure have adopted satellite-estimated PM$_{2.5}$ levels as their exposure estimates (Crouse et al. 2012; Madrigano et al. 2013). Because sufficient ground PM$_{2.5}$ measurements are needed to fit and validate statistical models, development of models in China was difficult before 2013. van Donkelaar et al. (2010) estimated long-term (2001–2006) average global PM$_{2.5}$ concentrations at 0.1° resolution using the PM$_{2.5}$/AOD ratios derived from a global chemical transport model (CTM). Two follow-up studies estimated the global PM$_{2.5}$ time series from 1998 to 2012 (Boys et al. 2014; van Donkelaar et al. 2015). Both studies validated their seasonal average estimates only with ground observations mostly obtained from North America, and the Pearson coefficients ranged from -0.37 to ~ 0.68 ($R^2 = ~0.14-0.46$). Yao and Lu (2014) used an artificial neural network (ANN) model to estimate PM$_{2.5}$ levels in China from 2006 to 2010. However, their ANN was trained partially using PM$_{2.5}$ and satellite data from the United States, which may have introduced substantial prediction error.

Taking advantage of the newly available national PM$_{2.5}$ measurements for China, Ma et al. (2014) estimated PM$_{2.5}$ levels for 2013 in China using satellite AOD and a geographically weighted regression (GWR) model. Using an early version of the 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$_{2.5}$ levels. The coarse resolution was a result of the limited coverage of AOD values retrieved by the Moderate Resolution Imaging Spectroradiometer (MODIS; http://modis.gsfc.nasa.gov) instruments 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 from the Terra MODIS collection (C6) were made available and are used in this study.

Satellite-based monitoring of PM$_{2.5}$ concentration data 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$_{2.5}$ concentrations from 2004 to early 2014 using ground observations.

**Results:** The overall model cross-validation $R^2$ and relative prediction error were 0.79 and 35.6%, respectively. Validation beyond the model year (2013) indicated that it accurately predicted PM$_{2.5}$ concentrations with little bias at the monthly ($R^2 = 0.73$, regression slope = 0.91) and seasonal ($R^2 = 0.79$, regression slope = 0.92) levels. Seasonal variations revealed that winter was the most polluted season and that summer was the cleanest season. Analysis of predicted PM$_{2.5}$ levels showed a mean annual increase of 1.97 μg/m$^3$ between 2004 and 2007 and a decrease of 0.46 μg/m$^3$ between 2008 and 2013.

**Conclusions:** Our satellite-driven model can provide reliable historical PM$_{2.5}$ estimates in China at a resolution comparable to those used in epidemiologic studies on the health effects of long-term PM$_{2.5}$ exposure in North America. This data source can potentially advance research on PM$_{2.5}$ health effects in China.


Address correspondence to Y. Liu, Emory University, Rollins School of Public Health, 1518 Clifton Rd. NE, Atlanta, GA 30322 USA. Telephone: (404) 727-2131. E-mail: yang.liu@emory.edu, or L. Huang, School of the Environment, Nanjing University, 163 Xianlin Ave., Nanjing 210023, China. Telephone: 86 25 89608535. E-mail: huangjiei@nju.edu.cn

Supplemental Material is available online (http://dx.doi.org/10.1289/ehp.1409481). This publication was made possible by U.S. Environmental Protection Agency (EPA) grant R834799. The work of Y.L. was partially supported by the NASA Applied Sciences Program (grant NNX11AP59G). The work of L.H. and J.B. was partially supported by the Key Program of the Chinese National Natural Science Foundation (71430007) and the National High-tech R&D Program (863 Program) of China (2013AA06A309). The work of Z.M. was supported by the China Scholarship Council (CSC).

The contents of this publication are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. EPA. Further, the U.S. EPA does not endorse the purchase of any commercial products or services mentioned in the publication. The authors declare they have no actual or potential competing financial interests.
by the enhanced DT (Levy et al. 2013) and Deep Blue (DB) algorithms (Hsu et al. 2013) were released. Despite providing better coverage over deserts and urban centers than DT AOD, DB AOD has rarely been used in PM$_{2.5}$ studies owing 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°, which is approximately 10 km) statistical model to estimate historical ambient PM$_{2.5}$ concentrations in China from 2004 to 2013 using MODIS C6 AOD data. First, we present our approach to generating a custom “combined” AOD parameter using the operational DT and DB AOD values, and we describe our two-stage spatial statistical model for estimating daily ambient PM$_{2.5}$ levels. We then evaluate predicted PM$_{2.5}$ concentrations at seasonal, monthly, and daily levels using ground PM$_{2.5}$ measurements in China not included in the model development. Finally, we analyze the 10-year spatiotemporal trend of PM$_{2.5}$ levels.

Materials and Methods

Ground PM$_{2.5}$ measurements. The daily average PM$_{2.5}$ concentrations for China (January 2013 to June 2014) were collected primarily from the website of the China Environmental Monitoring Center (CEMC). We collected additional data that are not included in the CEMC data from the websites of local environmental monitoring centers of several provinces (e.g., Shandong, Shanxi, Zhejiang, Guangdong) and municipalities (e.g., Beijing, Tianjin). Daily PM$_{2.5}$ data for Macao (2013), Hong Kong (2005 to June 2014), and Taiwan (2004 to June 2014) were also collected from the 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 web links for the abovementioned PM$_{2.5}$ data sources are shown in Supplemental Material, Table S1. Data for Changzhou City in Jiangsu province were provided by the Changzhou Environmental Monitoring Center. Monthly and seasonal mean PM$_{2.5}$ measurements for Beijing from 2005 to 2007 were obtained from Zhao et al. (2009). All ground PM$_{2.5}$ measurements were obtained using tapered element oscillating microbalances (TEOMs) or beta attenuation monitors, both of which are subject to measurement errors due to the loss of semivolatile components (Duncan et al. 2014). However, because PM$_{2.5}$ compliance in China is based on measurements obtained from these monitors, we used these PM$_{2.5}$ measurements to develop and evaluate our model. Our study included 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), and data from other years were used to evaluate the predicted historical PM$_{2.5}$ concentrations.

Satellite data. We extracted DT and DB AOD data for the period from January 2004 to June 2014 at 550 nm from the Aqua MODIS Level 2 aerosol data product, which were downloaded from the Level 1 and Atmosphere 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 is equal to DT AOD if NDVI > 0.3 and is equal to DB AOD if NDVI < 0.2. When 0.2 ≤ NDVI ≤ 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. Detailed descriptions of the MODIS operational combined AOD algorithm and the QA flags can be found elsewhere (Levy et al. 2013).

Figure 1. Spatial distribution of ground PM$_{2.5}$ monitoring sites. Open circles denote the sites with data available only from January to June 2014. Solid circles denote the sites with data available for not only 2014 but also for 2013 or earlier years. Note that many clustered sites are overlapped because of their proximity. The spatial resolution of the background gridded population is 0.1° × 0.1°.
We did not use the operational combined AOD data set of the MODIS C6 aerosol product because it discarded all DB AOD data with NDVI values > 0.3 (see 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 the 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 (see 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:

$$\text{AOD}_c = \frac{\text{AOD}_{DT}/\text{Var}_{DT,m} + \text{AOD}_{DB}/\text{Var}_{DB,m}}{1/\text{Var}_{DT,m} + 1/\text{Var}_{DB,m}}$$

where AODc is the IVW–combined AOD; AOD_{DT} and AOD_{DB} are the gap-filled DT and DB AOD, respectively; and Var_{DT,m} and Var_{DB,m} are the variances of the differences between the gap-filled DT and DB AOD and the AERONET AOD of season m, respectively. When compared with the AERONET observations, our combined AOD performed similarly ($R^2 = 0.80$, mean bias = 0.07) to MODIS’s operational combined AOD ($R^2 = 0.81$, mean bias = 0.07) but had 90% greater coverage. Spatially, the improvement in temporal coverage varied by land use type (Figure 2). Coverage for densely populated southern and eastern China improved by 50–100%. The Tibetan plateau showed the most improvement (~ 200%), whereas the Gobi and Taklamakan Deserts showed 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.** 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° latitude × 0.3125° longitude in a nested grid covering China and has been available since April 2012. GEOS-5.2.0 is the previous version of GEOS-5 FP and has a resolution of 0.50° × 0.666°. GEOS-5.2.0 data are available from January 2004 to May 2013. We averaged GEOS-5 FP data to the GEOS-5.2.0 grid to maintain a consistent spatial resolution across all model years. We used GEOS-5 FP data from 2013 for model development and GEOS-5.2.0 data from 2004 to 2012 for estimating historical PM$_{2.5}$ levels. The period of overlap (April 2012 to May 2013) for GEOS-5 FP and GEOS-5.2.0 data was used to evaluate the influence of the change in meteorological data source (see Supplemental Material, Figure S3, Table S2, “Comparison of model performance using GEOS-5 FP and GEOS-5.2.0 meteorological data”). We extracted planetary boundary layer height (PBLH, 100 m), wind speed (WS, meters per second) at 10 m above the ground, mean relative humidity in PBL (RH$_{PBLH}$, percent), and surface pressure (PS, hectopascals) between 1300 and 1400 hours local time (Aqua satellite overpass time corresponds to 1330 hours local time), as well as

![Figure 2](https://example.com/figure2.png)

**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).
as cumulative precipitation from the previous day (Precip_Lag1; millimeters). Land use variables at 300-m resolution were obtained from the European Space Agency (ESA) Global LandCover data portal (GlobCover; http://due.esrin.esa.int/page_glbcover.php) (Bontemps et al. 2011). We extracted urban and forest cover data from GlobCover 2005–2006 to represent study years 2004–2008 and from GlobCover 2009 to represent the years 2009–2014.

Data integration. We created a 0.1° grid (100,699 grid cells in total) for data integration and model development. Ground PM$_{2.5}$ data from multiple monitors in each grid cell were averaged. Because the sizes and geographical locations of MODIS AOD pixels vary in space and time, a 0.1° 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° 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° grid to spatially assign combined AOD values to the grid cells. We interpolated the GEOS-5 FP and GEOS-5.2.0 data to the 0.1° grid using the inverse distance weighting (IDW) method. We calculated the percentage forest cover and urban area in each grid cell and the daily total counts of MODIS fire spots for each grid cell using a 75-km radius buffer. Finally, all of the variables in 2013 were matched by grid cell and day-of-year (DOY) for model fitting. The model prediction data set was composed of all spatiotemporally matched variables except PM$_{2.5}$ concentrations from January 2004 to June 2014. Before model development, the independent variables in the fitting and prediction data sets were centered by subtracting their respective mean values computed from the fitting data set.

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-effects (LME) model included day-specific random intercepts and slopes for AOD and season-specific random slopes for meteorological variables:

$$PM_{2.5_{\text{resid}}} = \beta_0 + s(X, Y) + s(\text{ForestCover}) + s(\text{UrbanCover}) + \varepsilon_{\text{resid}} \tag{3}$$

where $PM_{2.5_{\text{resid}}}$ is the residual from the first-stage model at grid cell $s$ on day $t$; $\mu_0$ is the intercept term; $s(X, Y)$ is the smooth term of the coordinates of the centroid of grid cell $s$; $s(\text{ForestCover})$, and $s(\text{UrbanCover})$ are the smooth functions of percent forest cover and urban area for grid cell $s$, respectively; and $\varepsilon_{\text{resid}}$ 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 to test for potential model over-fitting.

Results of model fitting and cross-validation. We summarize the fixed effects estimates, model fitting, and CV results of the first-stage LME model for each province in Supplemental Material, Table S3. AOD is the only variable that was statistically significant in all provincial models ($p < 0.05$). Wind speed, relative humidity, and precipitation were significant in most provincial models. Fire spots were not significant in some provinces, most likely because these regions have infrequent fire activity. The CV $R^2$ values of the first-stage LME model ranged 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 performed 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 was not substantially over-fitted. Comparing the first-stage model (Figure 4A,C) with the full model (Figure 4B,D), it is clear that the second-stage GAM model marginally increased the $R^2$ values. However, the GAM model did 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 μg/m$^3$, Figure 4D) of the linear regression between the model-estimated and observed PM$_{2.5}$ concentrations for 2013. More importantly, the PM$_{2.5}$ levels in Hebei Province predicted by the full model were approximately 20 μg/m$^3$ higher than those predicted by the first-stage model; the predicted PM$_{2.5}$ levels in Tibet were approximately 15 μg/m$^3$ lower in the full model (see Supplemental Material, Figure S4) than in the first-stage model, showing that the spatial pattern of the PM$_{2.5}$ levels predicted by the full model was more consistent with that of the ground observations than PM$_{2.5}$ levels predicted by the first-stage model.

**Evaluation of historical PM$_{2.5}$ predictions.** Although our model’s predictions for daily level observations were poor compared with 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 yielded more accurate monthly or seasonal estimations (see 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 could be a fairly accurate ($R^2 = 0.73$) representation of monthly PM$_{2.5}$ levels measured from ground observations with only a slight bias (regression slope = 0.91). This threshold of 6 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 could be an 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) for China and four subregions (including the Beijing-Tianjin metropolitan region, the Yangtze River delta, the Pearl River delta, and the Sichuan Basin). The highest PM$_{2.5}$ estimations were for the Beijing-Tianjin metropolitan region (including Beijing, Tianjin, and Hebei), followed by those for the Sichuan Basin, the Yangtze River delta (including Jiangsu, Shanghai, and Anhui), and the Pearl River delta. The 10-year mean PM$_{2.5}$ estimations for the Beijing-Tianjin metropolitan region were generally ≥ 100 μg/m$^3$, and the highest concentrations were ≥ 120 μg/m$^3$. Similarly, the 10-year mean PM$_{2.5}$ concentrations were generally ≥ 85 μg/m$^3$ in the Sichuan Basin and the Yangtze River delta. The mean PM$_{2.5}$ concentrations were generally ≥ 55 μg/m$^3$ in the Pearl River delta. High PM$_{2.5}$ levels also occurred in the Taklamakan Desert in Xinjiang, an area that is a major dust source.

**Figure 3. Workflow for estimating spatiotemporal PM$_{2.5}$ concentrations.**
Satellite-estimated PM$_{2.5}$ trends in China

Figure 4. Density scatterplots of model fitting and cross-validation (CV) at the daily level ($n=63,031$). (A) and (B) are model-fitting results for the first-stage linear mixed effects (LME) model and the full LME + generalized additive model (GAM) model, respectively. (C) and (D) are model CV results for the first-stage LME model and the full LME + GAM model, respectively. Abbreviations: MPE, mean prediction error ($\mu$g/m$^3$); RMSE, root mean squared prediction error ($\mu$g/m$^3$); RPE, relative prediction error (%). The dashed line is the 1:1 line.

Figure 5. Evaluation of historical PM$_{2.5}$ estimations (2004–2012 and January–June 2014) at daily (A), monthly (B), and seasonal (C) levels. Because there were few ground PM$_{2.5}$ data for mainland China before 2013, we also estimated PM$_{2.5}$ for the first half of 2014 using the 2013 model and compared the results with the ground measurements to validate the accuracy of the historical estimations.
historical PM$_{2.5}$ levels in the years beyond the modeling year. The lack of concordance between daily historical PM$_{2.5}$ predictions and ground measurements was caused by the strong model assumption that the daily PM$_{2.5}$–AOD relationship derived from the 2013 data remained constant for the same DOY in each year. This limitation of our model cannot be resolved without sufficient historical PM$_{2.5}$ data to allow annual model adjustments before 2013. Nonetheless, our monthly ($R^2 = 0.73$, slope = 0.91) and

Figure 6. Spatial distributions of 10-year (2004–2013) mean PM$_{2.5}$ estimations for all of China (A), the Beijing-Tianjin metropolitan region (B), the Yangtze River delta (C), the Pearl River delta (D), and the Sichuan Basin (E).
seasonal ($R^2 = 0.79$, slope = 0.92) mean PM$_{2.5}$ 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$_{2.5}$ exposure in China. The seasonal patterns showed that the most polluted season was winter and the cleanest summer, 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$_{2.5}$ 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$_{2.5}$ 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 were able to develop this high-performance spatial model for China. The same model cannot be applied in regions with sparse ground observations. In this case, the scaling approach described by Brauer et al. (2012) is the only applicable method.

We compared the 9-year (2005–2013) AOD-derived and ground-measured PM$_{2.5}$ trends for Hong Kong (no PM$_{2.5}$ monitoring sites in 2004) and Taiwan (few sites in 2004). The results revealed that the AOD-derived PM$_{2.5}$ trend for Hong Kong (–1.28 μg/m$^3$ per year) was similar to the trend for the ground measurements (–1.35 μg/m$^3$ per year). However, the AOD-derived PM$_{2.5}$ trend for Taiwan was –0.17 μg/m$^3$ per year, which was much higher than that for the ground measurements (–0.72 μg/m$^3$ per year). This inconsistency is most likely due to missing satellite AOD retrievals. For example, only 34.5% of the grid cells in Taiwan had > 50% months with AOD-derived PM$_{2.5}$ data. Missing AOD values are a major limitation of and challenge for PM$_{2.5}$-AOD modeling (Liu et al. 2014). 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 for the monthly mean PM$_{2.5}$ time series around 2008. The PM$_{2.5}$ level increased steadily 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$_{2.5}$ levels rose steadily 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 and 2005 (Xue et al. 2014). The growth in energy demand led to a stricter energy conservation and emissions reduction (ECER) policy, which required a 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 reduction achieved by 2010 was 19.06% (Lo and Wang 2013). A recent study also showed that the production-related PM$_{2.5}$ emissions in China peaked in approximately 2007 and dropped quickly afterwards (Guan et al. 2014). A sharp reduction of PM$_{2.5}$ levels induced by this ECER policy may explain the inflection point.

**Conclusions**

The two-stage satellite AOD model developed in the present study generated reliable historical monthly and seasonal PM$_{2.5}$ predictions for China at 10-km resolution and with little bias, including data from the past decade, when the regulatory PM$_{2.5}$ monitoring network did not exist. Because several long-term PM$_{2.5}$ health effects studies in North America and the Global Burden of Disease project are driven by satellite exposure estimates obtained at this resolution

![Figure 7](image)

**Figure 7.** Time series of monthly, satellite-derived PM$_{2.5}$ anomaly (μg/m$^3$) for all of China (A), the Beijing-Tianjin metropolitan region (B), the Yangtze River delta (C), and the Pearl River delta (D), and spatial distribution of PM$_{2.5}$ trends for 2004–2007 (E) and 2008–2013 (F). The white areas in (E) and (F) indicate missing data. The black lines in A–D denote the PM$_{2.5}$ trends for 2004–2013, the red lines represent the trends for 2004–2007, and the blue lines represent the trends for 2008–2013. The PM$_{2.5}$ trends (μg/m$^3$ per year), 95% confidence intervals (CIs) in parentheses (μg/m$^3$ per year), and significance levels (*p < 0.05; **p < 0.01; ***p < 0.005) are also shown in A–D.
Ma et al. (Brauer et al. 2012; Crouse et al. 2012; Madrigano et al. 2013), our model predictions could greatly enhance research on long-term PM$_{2.5}$ health effects in China. With the release of the Terra MODIS C6 product in early 2015, the predicted historical PM$_{2.5}$ time series can now be extended to early 2000, if consistent meteorological and land use parameters are found to cover 2000–2003. From 2013 onward, our model can provide daily PM$_{2.5}$ exposure estimates to fill the gaps left by the PM$_{2.5}$ monitoring network in China. Finally, given the wider dynamic range of PM$_{2.5}$ concentrations in China compared with that in 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 (e.g., major point sources) information, which was not available when this study was performed (Kloog et al. 2014).

**References**


