Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

Article

Fusing Observational, Satellite Remote Sensing and Air Quality Model Simulated Data to Estimate Spatiotemporal Variations of $PM_{2.5}$ Exposure in China

Tao Xue [1], Yixuan Zheng [1], Guannan Geng [1,2], Bo Zheng [2], Xujia Jiang [1], Qiang Zhang [1,3,*], Kebin He [2,3]

[1] Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China;

[2] State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China;

[3] The Collaborative Innovation Center for Regional Environmental Quality, Beijing 100084, China;

* Correspondence: qiangzhang@tsinghua.edu.cn Tel.: +86-10-62795090

Abstract: Estimating ground surface $PM_{2.5}$ with fine spatiotemporal resolution is a critical technique for exposure assessments in epidemiological studies of its health risks. Previous studies have utilized monitoring, satellite remote sensing or air quality modeling data to evaluate the spatiotemporal variations of $PM_{2.5}$ concentrations, but such studies rarely combined these data simultaneously. Through assembling techniques, including linear mixed effect regressions with a spatial-varying coefficient, a maximum likelihood estimator and the spatiotemporal Kriging together, we develop a three-stage model to fuse $PM_{2.5}$ monitoring data, satellite-derived aerosol optical depth (AOD) and community multi-scale air quality (CMAQ) simulations together and apply it to estimate daily $PM_{2.5}$ at a spatial resolution of 0.1˚ over China. Performance of the three-stage model is evaluated using a cross-validation (CV) method step by step. CV results show that the finally fused estimator of $PM_{2.5}$ is in good agreement with the observational data (RMSE = 23.0 μg/m^3 and R^2 = 0.72) and outperforms either AOD-derived $PM_{2.5}$ (R^2 = 0.62) or CMAQ simulations (R^2 = 0.51). According to step-specific CVs, in data fusion, AOD-derived $PM_{2.5}$ plays a key role to reduce mean bias, whereas CMAQ provides spatiotemporally complete predictions, which avoids sampling bias caused by non-random incompleteness in satellite-derived AOD. Our fused products are more capable than either CMAQ simulations or AOD-based estimates in characterizing the polluting procedure during haze episodes and thus can support both chronic and acute exposure assessments of ambient $PM_{2.5}$. Based on the products, averaged concentration of annual exposure to $PM_{2.5}$ was 55.7 μg/m^3, while averaged count of polluted days ($PM_{2.5}$ > 75 μg/m^3) was 81, across China during 2014. Fused estimates will be publicly available for future health-related studies.

Keywords: Fine particulate matter ($PM_{2.5}$); Aerosol optical depth (AOD); Community multi-scale air quality (CMAQ) model; Data fusion; Exposure assessment.

1. Introduction

Many epidemiological studies have associated particulate matter with an aerodynamic diameter of ≤ 2.5 μm ($PM_{2.5}$) with adverse health outcomes, including cardiovascular and respiratory diseases [1,2], infant birth defects [3-5], DNA damages [6,7], cancer mortality [8,9] and many others. Severe

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

$PM_{2.5}$ pollution in China has attracted considerable public attention [10-12] and inspired numerous epidemiological studies to investigate the health effects of air pollution in China since 2013 [13-18]. Accurately assessing $PM_{2.5}$ exposure is critical for estimating its health risks in such epidemiological studies. However, due to the limited number of ground monitors in China, previous studies generally ignored the spatial variation of $PM_{2.5}$ and assessed the ambient exposure uniformly using one monitor or averages of several monitors located within a city or a municipality [13-16], which causes exposure misclassification. Therefore, accurately estimating the fine-scale spatiotemporal variation of ground $PM_{2.5}$ may lay a foundation for future health-related studies of $PM_{2.5}$ in China.

Three types of numerical values have been applied in exposure assessments of ambient particles: (1) monitoring observations, (2) satellite remote sensing measurements of aerosol, and (3) air quality model simulations. Routine monitors were widely used to predict air pollution concentrations across an area using geostatistical methods such as Kriging [19,20] or land use regression (LUR) to incorporate external spatial covariates [21,22], but such monitors may be sparsely distributed in sub-urban or rural areas. The $PM_{2.5}$ monitoring network has been rapidly spreading over China. In 2013, only approximately 70 cities or municipalities were covered by official sites of the China Environmental Monitoring Center (CEMC), whereas by 2015, the number had increased to approximately 330. However, monitoring data remain inadequate for characterizing the national-scale spatial variability of $PM_{2.5}$ in China.

The satellite remote sensing technique can retrieve integrated column concentrations of gases and aerosol from the bottom to top of the atmosphere and has been applied to assess ground surface air pollution [23]. Satellite-derived aerosol optical depth (AOD) has been successfully associated with ground $PM_{2.5}$ [24] and has thus been used to generate spatiotemporal estimators of $PM_{2.5}$ by acting as a primary predictor in statistical models such as LUR [25-27] or being calibrated by ratios ($PM_{2.5}$/AOD) simulated by a chemical transport model (e.g., GEOS-Chem) [28,29]. However, due to meteorological or geographical conditions, non-randomly missing values in satellite-derived AOD caused absent estimates of $PM_{2.5}$ in specific periods (e.g., winter [26]) or areas (e.g., deserts [28]).

Air quality models, such as the community multi-scale air quality model (CMAQ), simulate pollution concentrations based on emission inventories and chemical and physical processes driven by a meteorological model, such as the weather research and forecasting model (WRF) [30], and can provide exposure estimates with spatiotemporally complete coverage [31,32]. However, the accuracy of air quality models depends on the uncertainty of emission inventories and meteorological inputs and has thus been reported to vary with seasons and locations [33].

Hybrid models have been developed to combine different numerical values of air pollutants to improve exposure estimates. Beckerman, *et al.* [34] estimated monthly $PM_{2.5}$ on an 8.9 km × 8.9 km grid over the contiguous United States (US) through combining LUR of monitors and satellite $PM_{2.5}$ derived from GEOS-Chem and AOD. Mcmillan, *et al.* [35] developed a hierarchical Bayesian spatiotemporal model to bring monitors and CMAQ together and generated daily $PM_{2.5}$ and O_3 on both 36 km × 36 km and 12 km × 12 km grids over the US. Friberg, *et al.* [36] introduced a method to fuse CMAQ simulations and monitoring observations for daily estimates of multiple air pollutants on a 12 km × 12 km grid over Georgia, US. Liu, *et al.* [18] utilized the Ensemble Kalman Filter (EnKF) to assemble in situ observations with daily stimulations of $PM_{2.5}$ from an air quality model across China, and applied the analyzed products in risk assessment of chronic exposure to ambient pollution.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

Beloconi, *et al.* [27] mixed spatiotemporal Kriging maps of monitoring data and satellite AOD to estimate fine-scale (1 km × 1 km) daily estimates of $PM_{2.5}$ and PM_{10} during 2002-2012 over London.

To further increase the accuracy of exposure assessments of $PM_{2.5}$ via making full use of available data, this study aims to develop a fused estimator to join monitoring, satellite remote sensing and air quality modeling data together. Our approach incorporated (1) monitoring records from routine sites as reference measurements of $PM_{2.5}$, (2) CMAQ simulations as prior knowledge, which -provides completely spatiotemporal coverage of $PM_{2.5}$ and (3) satellite AOD as alternative observations with wider spatial coverage than monitors and higher accuracy than CMAQ. We applied a three-stage model for the fused estimator. In step 1, we derived ground surface $PM_{2.5}$ from satellite AOD and calibrated CMAQ simulations by monitoring data using two separate regression models. In step 2, we combined AOD-derived $PM_{2.5}$ and calibrated-CMAQ $PM_{2.5}$ using a maximum likelihood method. In step 3, we incorporated the spatiotemporal autocorrelation of the monitoring data in the final estimator through interpolating the residuals in step 2. We illustrated the three-stage model by a practice to develop daily maps of $PM_{2.5}$ on a regular grid of 0.1° × 0.1 ° over China. The performance of the statistical models was assessed using cross-validation (CV) methods by steps.

2. Materials and Methods

2.1. Data description

2.1.1. $PM_{2.5}$ monitoring data

We collected hourly $PM_{2.5}$ measurements from three monitoring networks in the year 2014, including the CEMC sites (http://113.108.142.147:20035/emcpublish/), the sites of the Beijing Municipal Environmental Monitoring Center (http://zx.bjmemc.com.cn/) and the sites of the Guangdong Environmental Monitoring Center (http://113.108.142.147:20031/AQIPublish/AQI.html). Duplicate monitoring sites among these three networks were removed, leaving a total of 944 sites (Figure 1). According to the Chinese National Ambient Air Quality Standard (CNAAQS, GB3095-2012) released in 2012, the ground-based $PM_{2.5}$ data are measured using the tapered element oscillating microbalance (TEOM) technique or the beta-attenuation method. For each monitor, we averaged $PM_{2.5}$ by day and excluded the dates with less than 19 hourly measurements. Figure 1 presents locations of the monitors.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

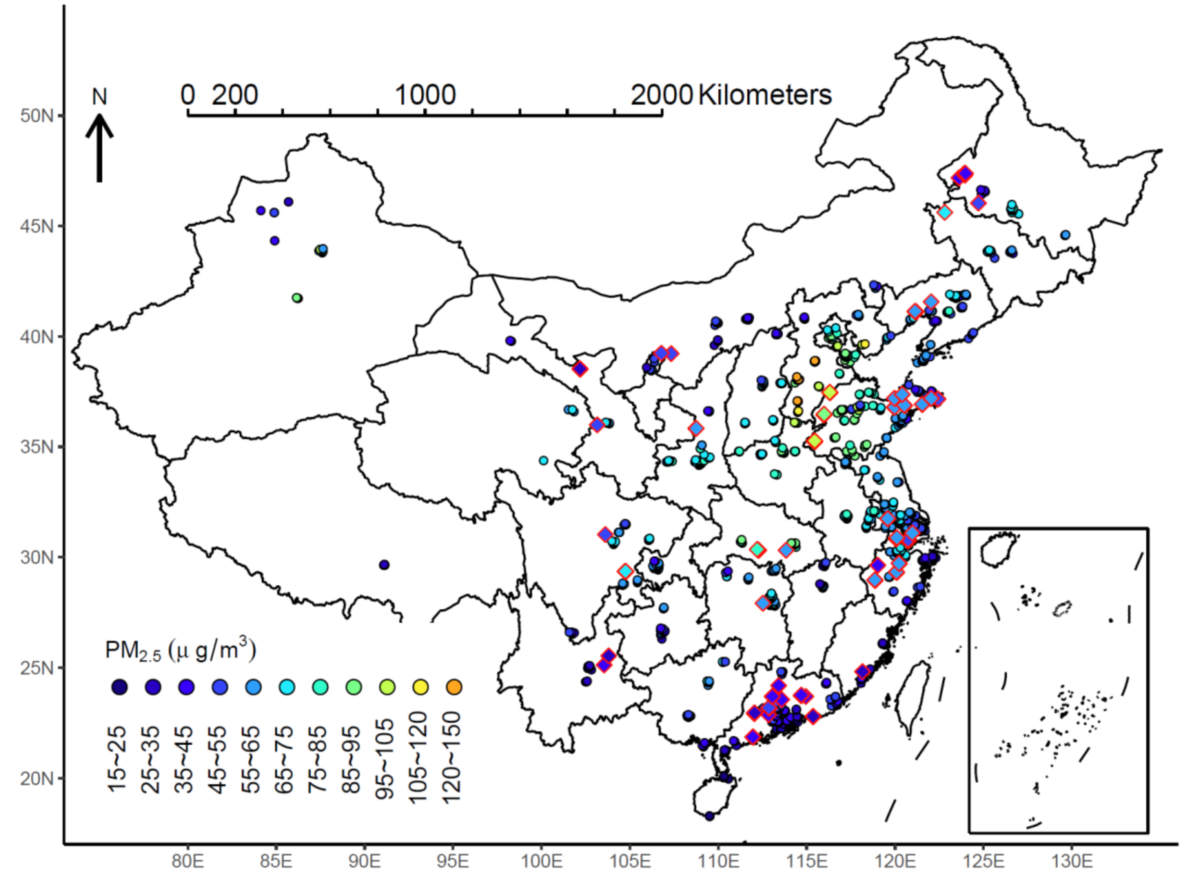


Figure 1. Locations of monitoring sites of $PM_{2.5}$. The diamond points present the sites, which are held-out in cross-validation.

2.1.2 Satellite remote sensing of AOD

Satellite-derived AOD data during 2014 were obtained from moderate resolution imaging spectroradiometer (MODIS), equipped on two earth observing system satellites, Terra and Aqua, which are operated by US National Aeronautics and Space Administration (NASA). Terra (Aqua) scans the earth at 10:30 a.m. (1:30 p.m.) with a global coverage of 1~2 days. They retrieved AOD from visible and near-infrared electromagnetic signals at nadir. Level 2 MODIS AOD products (collection 6) with a spatial resolution of 10 km × 10 km were collected from Atmosphere Archive and Distribution System (LAADS, http://ladsweb.nascom.nasa.gov). L2 swath data were resampled into a fixed grid, which covers the whole investigated regions with a spatial resolution of 0.1° × 0.1° via area-weighted averaging. The AOD_550_Dark_Target_Deep_Blue_Combined dataset with QA Flag equal to 2 or 3 were utilized in this study. According to Ma, *et al.* [37], we combined Terra/Aqua MODIS AOD measurements together to increase the spatial coverage of AOD measurements.

2.1.3 Satellite remote sensing covariates for AOD-derived $PM_{2.5}$

Satellite normalized difference vegetation index (NDVI) and fire spots (FS) were obtained from combined MODIS products. We aggregated monthly products of NDVI with a spatial resolution of 1km × 1km into seasonal averages over the regular grid of 0.1° × 0.1°. Daily counts of FS within a 75 km buffer around each centroid of the grid were calculated based on the location and time of fires, collected from MODIS burned area products. Integrated column concentrations of NO_2 were obtained from Ozone Monitoring Instrument (OMI), launched on Aura. Level 2 products of column NO_2 with a spatial resolution of 13 km × 24 km were prepared into seasonal means with a spatial resolution of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

0.1° × 0.1°. The above data can be accessed from https://lpdaac.usgs.gov/, http://modis-fire.umd.edu/ and http://disc.sci.gsfc.nasa.gov/.

2.1.4 WRF-CMAQ simulation

In this study, the WRF model version v3.5.1 (http://www.wrf-model.org/) and the CMAQ model version 5.1 were used to simulate the daily variations of $PM_{2.5}$ over China. The WRF model is driven by the National Centers for Environmental Prediction Final Analysis (NCEP-FNL) reanalysis data as initial and boundary conditions (ICs and BCs). Meteorological parameters simulated by WRF model were applied to drive CMAQ. Our CMAQ simulations utilized CB05 as the gas-phase mechanism, AERO6 as the aerosol module, and Regional Acid Deposition Model (RADM) as the aqueous-phase chemistry model. Boundary conditions for our CMAQ model were provided by dynamic GEOS−Chem simulation [38]. The anthropogenic emission for mainland China during 2014 are derived from the Multi-resolution Emission Inventory of China (http://www.meicmodel.org/). Detailed model configurations for WRF-CMAQ were presented in our previous study [39]. We simulated meteorological variables including ground wind speed (WS), planetary boundary layer height (PBL), ground ambient pressure (PS), and ground relative humidity (RH) by WRF and $PM_{2.5}$ by CMAQ with a spatial resolution of 36 km × 36 km, which were further downscaled to the 0.1° × 0.1° grid using an offline ordinary Kriging method [40]. The daily means of simulations were interpolated in spatial dimensions for each variable separately. The purpose of downscaling is to spatially match WRF-CMAQ simulations with the rest data. Validations for CMAQ-simulated $PM_{2.5}$ at both spatial resolutions (0.1° and 36 km) were performed using monitoring data, which are presented in Figures S2 and S3 and briefly illustrated in discussion section. After downscaling, CMAQ-simulated $PM_{2.5}$ covered 100% of spatiotemporal coordinates (99,351 pixels × 365 days), while the in situ observations or AOD measurements only covered 0.54% or 31.56% of spatiotemporal coordinates, respectively.

2.2 Statistical analysis

The modeling framework of exposure assessment included three steps, which were presented in Figure 2. Briefly, we first developed two regression models (steps 1.1 & 1.2) to associated AOD or CMAQ with in situ observations of $PM_{2.5}$, separately; then the estimates from the two models were combined based on a maximum likelihood (step 2); finally, we incorporated spatiotemporal autocorrelations of the monitoring $PM_{2.5}$ (step 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1
Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

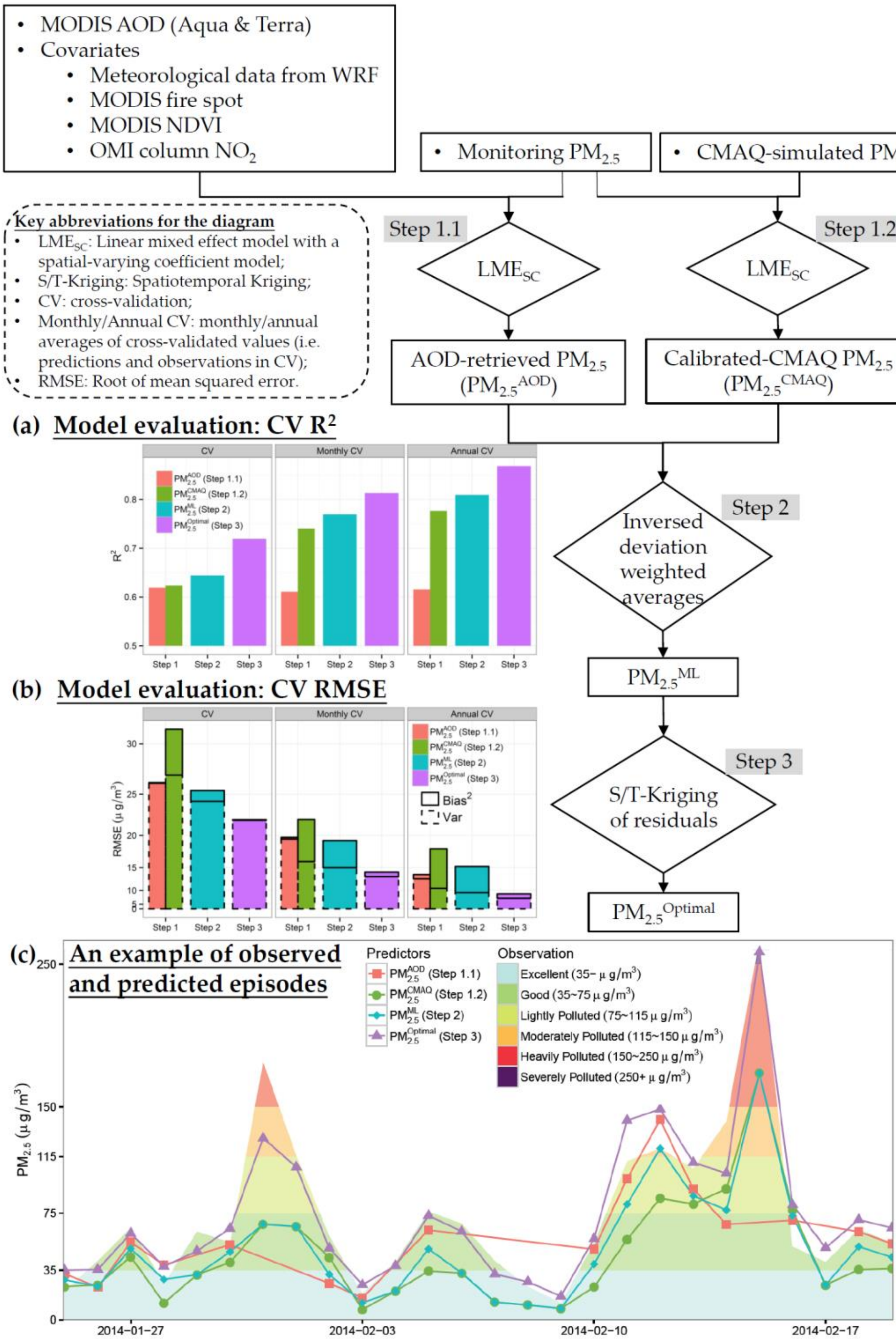


Figure 2. Framework for the three-stage model and cross-validation (CV) results by steps. (a) CV R^2s by steps. In (a), R^2s were derived based on all available CV samples. (b) CV RMSEs by steps. In (b), RMSEs were calculated using the records, where AOD-derived estimates or their averages were existing. Scale of y-axis is logarithm-transformed. Squared RMSE can be divided into two components: squared bias and variance of the estimates, which are highlighted by rectangles in (b). (c) An example of observed episodes by a CV_{IS} testing site located in (121.12° E, 41.12° N). The corresponding predictors (dots and lines) are presented with the monitoring observations (the polygons filled by colors, which reflect air pollution levels.). The location of the site is visualized by the red box in Figure 7.

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

2.2.1 Step 1.1: AOD-derived $PM_{2.5}$

Based on the mature methodology developed by Ma, *et al.* [37], we first derived $PM_{2.5}$ from satellite-retrieved AOD with the auxiliary variables, which were selected according to experimental findings (e.g. RH [41,42]) or empirical results on $PM_{2.5}$-AOD associations (e.g. NO_2 [43] and FS [26]). Instead of using the linear mixed effect model (LME) [37], we developed an updated version, a linear mixed effect model with a spatial-varying coefficient (LME_{SC}) model, as shown in follows:

$$\mathrm{PM}_{2.5,\mathbf{st}} = \mu + [\beta_1 + f(s)]AOD_{st} + (\beta_2 + \beta'_{2,j})WS_{st} + (\beta_3 + \beta'_{3,j})PBL_{st} + (\beta_4 + \beta'_{4,j})PS_{st} + (\beta_5 + \beta'_{5,j})RH_{st} + (\beta_6 + \beta'_{6,j})FS_{st} + (\beta_7 + \beta'_{7,j})NDVI_{sj} + (\beta_8 + \beta'_{8,j})NO_{2,sj} + \epsilon_{st}, \quad (1)$$

where

$$f(\boldsymbol{s}) = b'_{1,t}\eta_1(\boldsymbol{s}) + b'_{2,t}\eta_2(\boldsymbol{s}) + \cdots + b'_{\mathrm{k},t}\eta_k(\boldsymbol{s}),$$

$$\epsilon_{st} \sim N(0, \sigma^2),$$

$$[b'_{i,t=1}, b'_{i,t=2,} \cdots, b'_{i,t=T,}]' \sim N(0, \Phi_i), i = 1, \cdots, \mathrm{k};$$

$$[\beta'_{i,j=1}, \beta'_{i,j=2,}\beta'_{i,j=3}, \beta'_{i,j=4}]' \sim N(0, \Phi_i),\ i = 2, \cdots, 9.$$

In the LME_{SC}, **s**, t or j denotes spatial coordinates (longitude and latitude at the centroid of each pixel), daily or seasonal index; $\mathrm{PM}_{2.5,\mathbf{st}}$ denotes in situ observations at spatial location **s** and date t; $\mu, \beta_1, \cdots, \beta_8$ denote fixed intercept and slopes for covariates including (1) daily values of AOD, WS, PBL, PS, RH and FS and (2) seasonal values of NDVI and NO_2; $\beta'_{2,j}, \ldots, \beta'_{8,j}$ denote seasonally-specific random slopes for the other covariates than AOD. $f(s)$ denotes a spatial-varying coefficient for AOD and is expanded by a given set of k-dimensional basis functions (e.g. local bisquare functions [44]) and daily-specific random slopes ($b'_{\cdot,t}$). In this study, for computing efficiency, we expanded $f(s)$ by 2-D splines provided by R package mgcv [45]. $\boldsymbol{\eta}$s became known values depended on spatial coordinates (s), once the specific form of basis functions was determined. Thus the inference of coefficients ($b'_{\cdot,t}$) in $f(\boldsymbol{s})$ was done in regression procedure, simultaneously with other parameters (e.g. $\boldsymbol{\beta}$s) in equation (1). If $f(\boldsymbol{s})$ is simplified as a one-dimensional daily-specific random slope ($\beta'_{1,t}$), the LME_{SC} will be reduced to a LME, which has been utilized in previous studies to generate AOD-derived $PM_{2.5}$ [37]. LME method has disadvantages in generating spatially smoothing predictors, especially near the provincial boundaries. Through introducing spatial-varying coefficients, LME_{SC} fixed the problem and was evidenced to outperform LME by our cross-validation results (as shown by supplemental Figure S1 and Figure 4 (a)). Detailed comparisons are presented in discussion section. Spatial and temporal patterns for $PM_{2.5}$-AOD associations ($\beta_1 + f(s)$) are presented in Figure S8. Fitted value and its standard deviation (SD) from Equation 1 are denoted by $PM_{2.5}^{AOD}$ and SD^{AOD}, respectively. We named $PM_{2.5}^{AOD}$ as "AOD-derived $PM_{2.5}$" in this study. Equation (1) was fitted based on 92,644 in situ observations collocated with AOD data, and $PM_{2.5}^{AOD}$ was estimated at all spatiotemporal coordinates, where AOD existed.

2.2.2 Step 1.2: calibrated-CMAQ $PM_{2.5}$

We calibrated CMAQ simulated $PM_{2.5}$ with the in situ observations by a similar LME_{SC} model, shown as follows:

$$\mathrm{PM}_{2.5,\mathbf{st}} = \mu^* + [\beta_1^* + f^*(s)]\mathrm{CMAQ}_{\mathbf{st}} + \epsilon_{st}, \quad (2)$$

where $\mathrm{CMAQ}_{\mathbf{st}}$ denotes downscaled CMAQ-simulated $PM_{2.5}$ with a spatial resolution of 0.1° × 0.1°. In equation (2), we utilized original scale instead of log-scale of $PM_{2.5}$ in order to guarantee comparable error terms (ϵ_{st}) to that in equation (1), although logarithm transform was usually used to reduce the bias caused by violation of normality assumption of $PM_{2.5}$ in the regression analysis. Spatial and temporal variations of estimated coefficients of CMAQ-simulated $PM_{2.5}$ ($\beta_1^* + f^*(s)$) are presented in

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

Figure S9. Fitted value and its SD from Equation 2 are denoted by $PM_{2.5}^{CMAQ}$ and SD^{CMAQ}, respectively. We named $PM_{2.5}^{CMAQ}$ as "calibrated-CMAQ $PM_{2.5}$" in this study. Equation (2) was fitted based on all 294,122 in situ observations collocated with CMAQ data, and $PM_{2.5}^{CMAQ}$ was estimated at all spatiotemporal coordinates.

2.2.3 Step 2: inversed deviation weighted averages

To minimize the uncertainty, we derived a maximum likelihood estimator ($PM_{2.5}^{ML}$) for the collocated AOD-derived $PM_{2.5}$ and calibrated-CMAQ $PM_{2.5}$. Assuming the normality for the fitted values ($PM_{2.5}^{AOD}$ and $PM_{2.5}^{CMAQ}$), the maximum likelihood estimator can be simplified as inversed deviation weighted averages, shown as follows:

$$\mathrm{PM}_{2.5,\mathbf{st}}^{\mathrm{ML}} = \frac{\mathrm{PM}_{2.5,\mathrm{st}}^{\mathrm{AOD}}/\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{AOD}}\right)^2 + \mathrm{PM}_{2.5,\mathrm{st}}^{\mathrm{C}}/\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{CMAQ}}\right)^2}{1/\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{ML}}\right)^2} \tag{3}$$

where

$$\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{ML}}\right)^2 = \frac{1}{1/\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{AOD}}\right)^2 + 1/\left(\mathrm{SD}_{\mathbf{st}}^{\mathrm{CMAQ}}\right)^2}.$$

For the places, where the AOD is missing, the $PM_{2.5}^{ML}$ is defined identically as $PM_{2.5}^{CMAQ}$.

2.2.4 Step 3: spatiotemporal Kriging of the residuals

Taking spatiotemporal autocorrelation of $PM_{2.5}$ into consideration, we interpolated the residuals ($\mathrm{e}_{\mathbf{st}} = \mathrm{PM}_{2.5,\mathbf{st}} - \mathrm{PM}_{2.5,\mathbf{st}}^{\mathrm{ML}}$) using spatiotemporal Kriging (S/T-Kriging) based on a product-sum covariance function [46]. Assuming a stationary multivariate normal distribution for the residuals ($\mathrm{e}_{\mathbf{st}}$), the variance-covariance matrix can be captured by a function (C) of the spatiotemporal coordinates, as shown in follows:

$$[\mathrm{e}_{\mathbf{st}}] \equiv \mathbf{E} \sim MVN(0, \boldsymbol{\Sigma}),$$

$$Cov\left(\mathrm{e}_{\mathbf{s}_i\mathrm{t}_i}, \mathrm{e}_{\mathbf{s}_j\mathrm{t}_j}\right) \equiv \boldsymbol{\Sigma}_{\mathrm{i,j}} = \mathrm{C}\left(\left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\|_2, \left\|t_i - t_j\right\|_1 \;|\boldsymbol{\theta}\right),$$

where $\boldsymbol{\theta}$ denotes the tuning parameters in the covariance function (C) and can be estimated using variogram approach. For a spatiotemporal point ($\mathbf{s}^*$, t*), where in situ observation of $PM_{2.5}$ does not exist, the residual can be interpolated as $\hat{\mathrm{e}}_{\mathbf{s}^*\mathrm{t}^*} = Cov(\mathrm{e}_{\mathbf{s}^*\mathrm{t}^*}, \mathbf{E})\boldsymbol{\Sigma}^{-1}\mathbf{E}$. Therefore, the optimal estimates of $PM_{2.5}$ can be derived as

$$\mathrm{PM}_{2.5,\mathbf{st}}^{\mathrm{Optimal}} = \mathrm{PM}_{2.5,\mathbf{st}}^{\mathrm{ML}} + \hat{\mathrm{e}}_{\mathbf{st}} \tag{4}$$

For more details of S/T-Kriging, please refer to chapter 6 in Cressie and Wikle [46].

2.3 Model evaluation

Previous studies usually evaluated statistical performance of $PM_{2.5}$ estimators by the 10-fold cross validation (CV_{10}), which randomly divides the monitoring data into ten folds and iteratively leaves one fold as the testing dataset to assess the predictions from a model trained by the rest data. For independent data, the root of mean squared error (RMSE) has been considered as an unbiased estimator for prediction accuracy [47]. However, for spatiotemporally auto-correlated $PM_{2.5}$ data, CV_{10} may underestimate prediction errors [48]. To fairly evaluate the models, we designed isolated-site cross-validation (CV_{IS}), in which, we held out about 10% of the monitoring sties and used all measurements from the testing sites to validate the modeling results based on the rest data. The testing sites were randomly selected with two constraints: (1) they should be separated from the training sites by more than 25 km; and (2) they should be universally spanned over the study domain,

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

especially areas with dense population. In this study, we involved 91 testing sites with a minimum distance from the remained sites of 26.2km, as shown in Figure 1. The testing set contained 27,800 samples out of 294,122 total daily values of monitoring measurements. We kept multiple testing values located within one grid at the same time point, because the discrepancy among those values represents the error caused by spatially aggregation, which should not be ignored in model evaluation. A comparison between CV_{10} and CV_{IS} was performed based on AOD-derived $PM_{2.5}$ from a LME model (Figure S1), and more detailed rationale of CV_{IS} is presented in discussion section. We also evaluated the three-stage model using the same CV_{IS} data step by step. The CV_{IS} analysis of the intermediate estimators illustrated how the errors propagate in our data fusion model.

3. Results

3.1 Descriptive statistics for inputs of data fusion

Figure 3 presents the frequency distributions and summary statistics of in situ observations, CMAQ simulations and AOD-derived estimates of $PM_{2.5}$. CMAQ simulated or AOD-derived $PM_{2.5}$ concentrations were extracted at the same spatiotemporal coordinates of monitoring data, in order to compare the three types of inputs in our model. During 2014, the overall mean of the monitoring $PM_{2.5}$ is 61.3 μg/m^3, which is slightly higher that of CMAQ-simulated $PM_{2.5}$ (57.4 μg/m^3) but lower than that of AOD-derived $PM_{2.5}$ (66.4 μg/m^3), which suggests systematic bias in the latter two datasets. However, after excluding the observational $PM_{2.5}$ at the time points, when AOD is missing, the mean of monitoring data is increased to 66.6 μg/m^3 (Figure 3 (b)), which is close to that of AOD-derived $PM_{2.5}$. A Kolmogorov–Smirnov test indicated that monitoring data presented significantly different distributions depended on the missing status of AOD. According to our findings, in China AOD incompleteness occurred non-randomly and was influenced by the ambient concentrations of $PM_{2.5}$, which leads to sampling errors in AOD-derived $PM_{2.5}$. The systemic bias between frequency distribution of AOD-derived $PM_{2.5}$ and that of overall monitoring data was partially caused by the sampling errors of satellite-derived AOD.

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

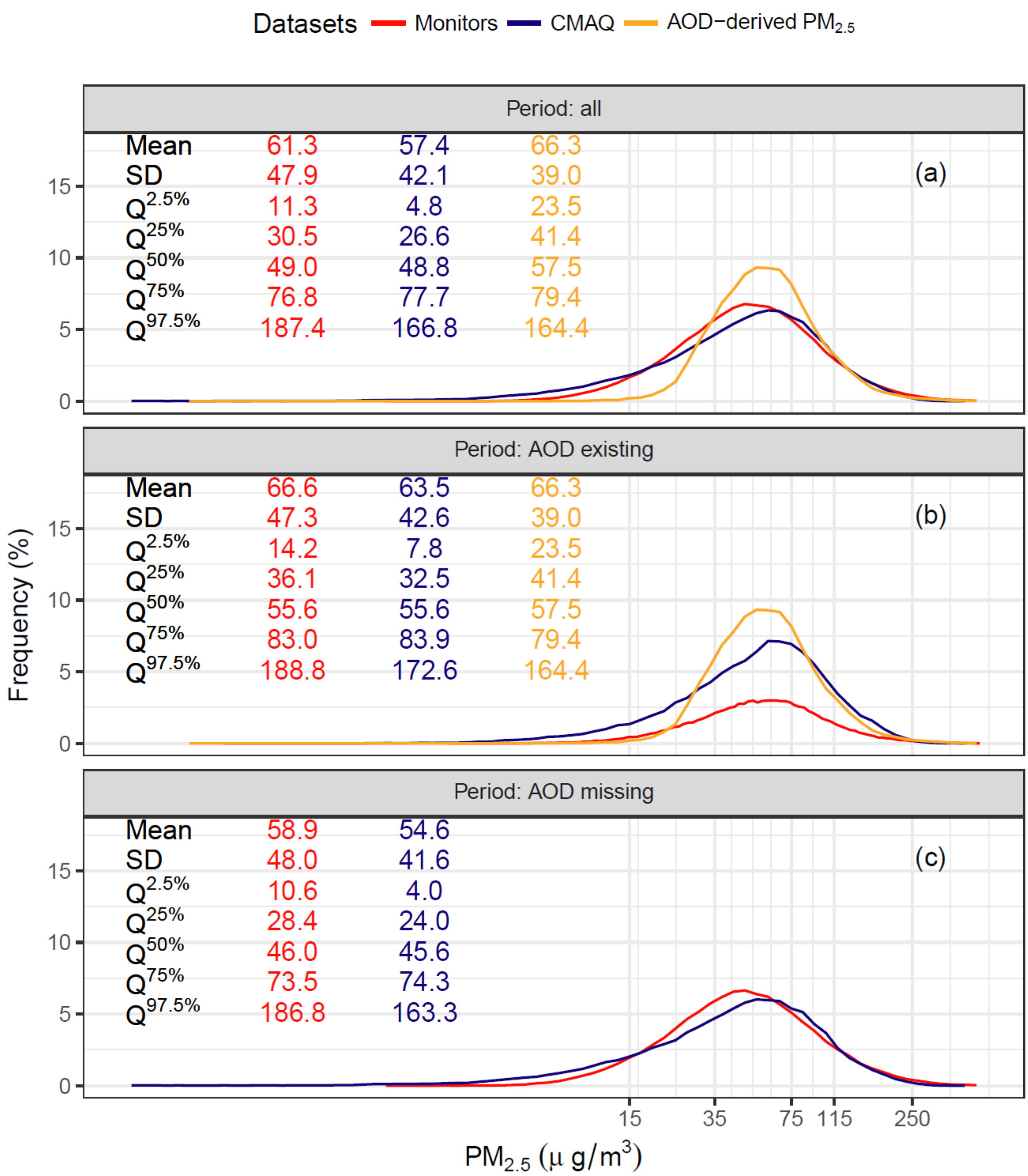


Figure 3. (a) Distributions of in situ observations, CMAQ simulations (0.1° × 0.1°) and AOD-derived estimates of $PM_{2.5}$ at the same spatiotemporal coordinates of monitoring data. (b) Distributions of the subsets, conditioned that AOD data are existing. (c) Distributions of the subsets, conditioned that AOD data are missing. $Q^{X\%}$ denotes the X^{th} percentile of a distribution.

3.2 Cross-validation results for the estimates of the three-stage model

Figure 4 (a) presents the CV_{IS} results of the three-stage model. The final estimator of the model ($PM_{2.5}^{Optimal}$) is in good agreement with the observational data (R^2=0.72). The root of mean squared error (RMSE) is 23.0 μg/m³, which accounts for 55% of the SD of observational $PM_{2.5}$ (defined as normalized root of mean squared error, NRMSE) and 41% of the mean of observational $PM_{2.5}$ (defined as relative prediction error, RPE). The mean bias is 4.9 μg/m³, which suggests that $PM_{2.5}^{Optimal}$

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

underestimate the true values. The slope of a linear regression of the predictors against the observations is 0.76, lower than 1, which indicates that $PM_{2.5}^{Optimal}$ may be over-smoothed. Among 27,800 testing observations in CV_{IS}, collocated satellite-derived AOD data are available for 9,530 of them. In another word, at one third of the CV points, $PM_{2.5}^{Optimal}$ were estimated based on AOD, CMAQ and ground monitors; while at the rest CV points, $PM_{2.5}^{Optimal}$ were estimated based on the latter two. To evaluate the capacity of $PM_{2.5}^{Optimal}$ to assess the long-term exposure to ambient $PM_{2.5}$, we averaged both the predicted and the observed values in CV_{IS} by month or by year (Figure 4 (a)). Because averaging can lower the variance of predictors, the CV_{IS} R^2 respectively increases to 0.81 or 0.87 for monthly or annually averages, which indicates that $PM_{2.5}^{Optimal}$ may be more appropriate to study chronic exposure than acute exposure to $PM_{2.5}$.

CV_{IS} results for the intermediate estimators (i.e. $PM_{2.5}^{AOD}$, $PM_{2.5}^{CMAQ}$ and $PM_{2.5}^{ML}$) of the three-stage model are shown in Figure 2 and Figure 4 (b)-(c). Generally speaking, the predicting errors were decreased step by step in our modeling process. For example, in daily scale, CV_{IS} R^2 increases from 0.62 for either $PM_{2.5}^{AOD}$ or $PM_{2.5}^{CMAQ}$ in step 1, to 0.64 for $PM_{2.5}^{ML}$ in step 2 and further to 0.72 for $PM_{2.5}^{Optimal}$ in step 3. The decreasing trend in CV_{IS} RSME is mostly dominated by the shrinkage in variations of predicting errors due to aggregations of multiple predictors at each testing site. As the more biased estimator, $PM_{2.5}^{CMAQ}$ is mixed with the less biased estimator, $PM_{2.5}^{AOD}$ in data fusion, biasness of the combined estimators ($PM_{2.5}^{ML}$ and $PM_{2.5}^{Optimal}$) lays between the former two. Although $PM_{2.5}^{AOD}$ is less biased than the others; it may fail to capture some $PM_{2.5}$ episodes due to incompleteness of satellite data (Figure 2 (c)). Such weaknesses are partially overcome by data fusion (Figure 2 (c)). The detailed CV_{IS} scatterplots for the intermediate estimators are presented in the supplemental Figure 4 (b)-(c).

We also explored temporal (Figure S6) and spatial (Figure S7) variations of CV_{IS} results. To evaluate the temporal variation of CV_{IS} errors, we calculated the statistics, including R^2, RMSE and NRMSE by dates. The daily CV_{IS} results reflected the final estimator's capacity to capture spatial variations of $PM_{2.5}$. The CV_{IS} RMSE for $PM_{2.5}^{Optimal}$ is proportional to the observed value and thus was varied seasonally (higher in colder season, but lower in warmer season). However, we found significantly trend neither in daily NRMSEs nor in daily R^2s (Figure S6), which indicates that the accuracy of $PM_{2.5}^{Optimal}$ is temporally constant. Analogously, we also calculated CV_{IS} statistics by sites to evaluate the final estimator's capacity to capture temporal variations of $PM_{2.5}$. CV_{IS} results by sites displayed significantly spatial patterns, which indicates that $PM_{2.5}^{Optimal}$ is more accurate in eastern China, but less in western China (Figure S7). Partial reason is that the accuracy of $PM_{2.5}^{Optimal}$ tends to increase with the density of training sites (Figure 8), which are more clustered in eastern China, especially the urban areas (e.g. Yangtze River Delta or Pearl River Delta metropolitan region).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

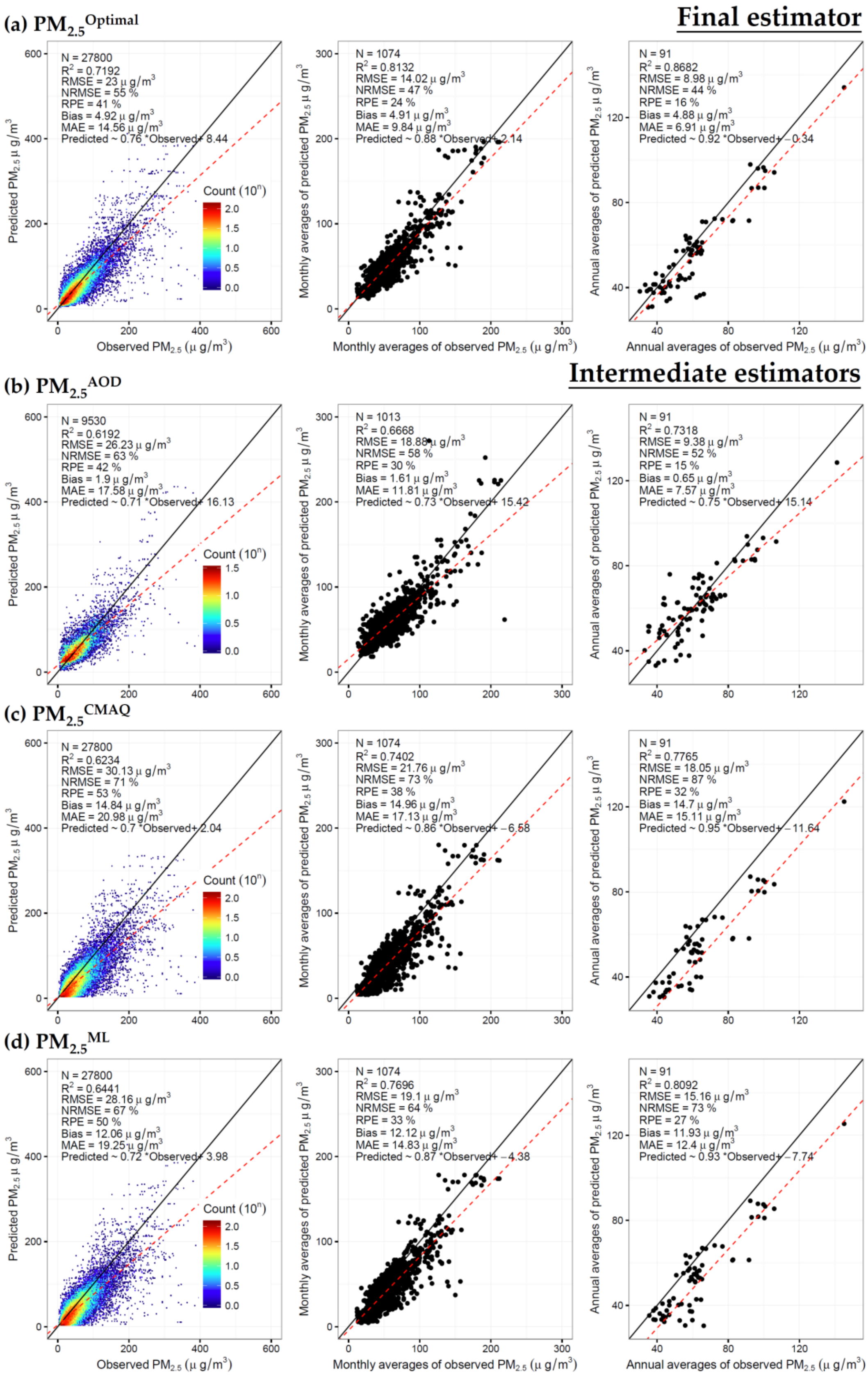


Figure 4. Scatterplots of cross-validated values and their monthly or annual averages for final estimator ($PM_{2.5}^{Optimal}$) and intermediate estimators of the three-stage model ($PM_{2.5}^{AOD}$, $PM_{2.5}^{CMAQ}$ and $PM_{2.5}^{ML}$).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

3.3 The fitted spatial and seasonal patterns of $PM_{2.5}$ in China

Figure 5 presents the annual maps of $PM_{2.5}$ fitted by the three-stage model and its intermediate steps. Different methods displayed consistent patterns in spatial variation of $PM_{2.5}$, particularly across eastern China, where $PM_{2.5}$ pollutants were dominated by anthropogenic sources. During 2014, the hot-spots of $PM_{2.5}$ ($PM_{2.5}^{Optimal}$ = 85~120 μg/m^3) spanned over North China Plain (the municipalities of Beijing and Tianjin, and the provinces of Hebei, Henan and Shandong). The moderately polluted areas ($PM_{2.5}^{Optimal}$ = 45~85 μg/m^3) occupied Sichuan Basin (Sichuan Province and Chongqing Municipality), Loess Plateau (Shanxi Province and middle of Shaanxi Province), Yangtze Plain (Shanghai Municipality, the provinces of Anhui, Jiangsu, Hunan and Hubei) and Northeast China Plain (the provinces of Heilongjiang, Liaoning and Jilin). The major divergence among these maps exists in the deserted areas of northwestern China. CMAQ-based estimators (i.e. CMAQ-simulated $PM_{2.5}$ and calibrated-CMAQ $PM_{2.5}$) failed to capture $PM_{2.5}$ from natural sources and underestimated the concentrations across the Taklamakan desert. In the fused estimators (i.e. $PM_{2.5}^{ML}$ and $PM_{2.5}^{Optimal}$), the problem was fixed by introducing AOD data. Figure 6 presents seasonal maps of $PM_{2.5}$ fitted by $PM_{2.5}^{Optimal}$, which confirms that $PM_{2.5}$ concentrations are higher during winter (DJF) and autumn (SON), but lower in summer (JJA) and spring (MAM). The sever pollution of $PM_{2.5}$ in colder seasons might be attributed by fossil fuel combustions, especially across northern China. Seasonal maps for the other estimators are presented in supplemental Figure S4.

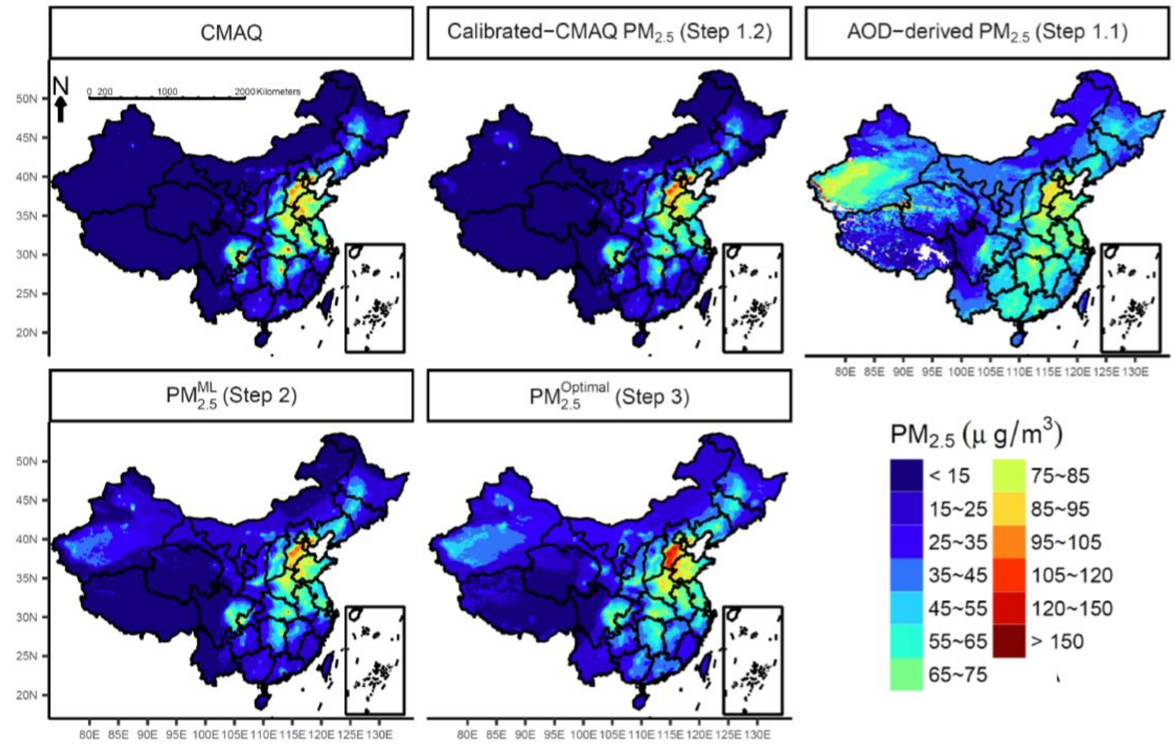


Figure 5. Annual maps (0.1° × 0.1°) of $PM_{2.5}$ during 2014 over China, produced by CMAQ, intermediate and final estimators of the three-stage model.

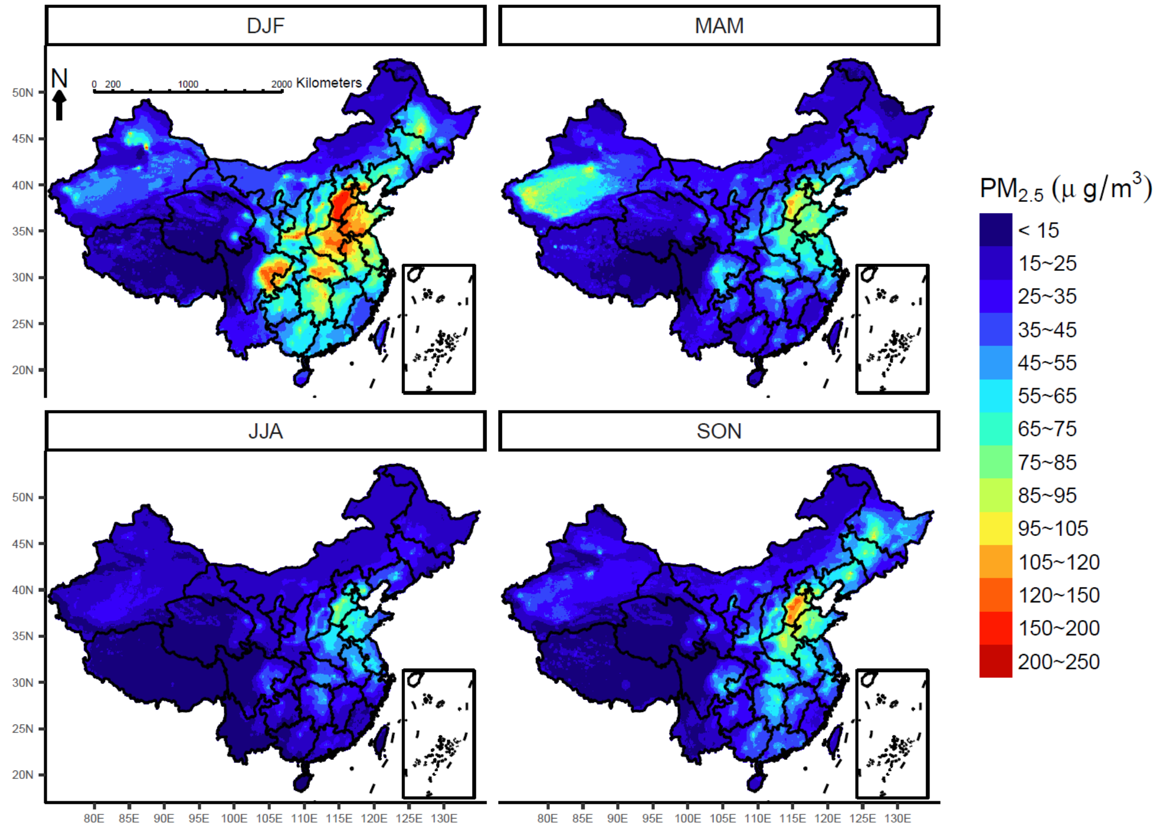
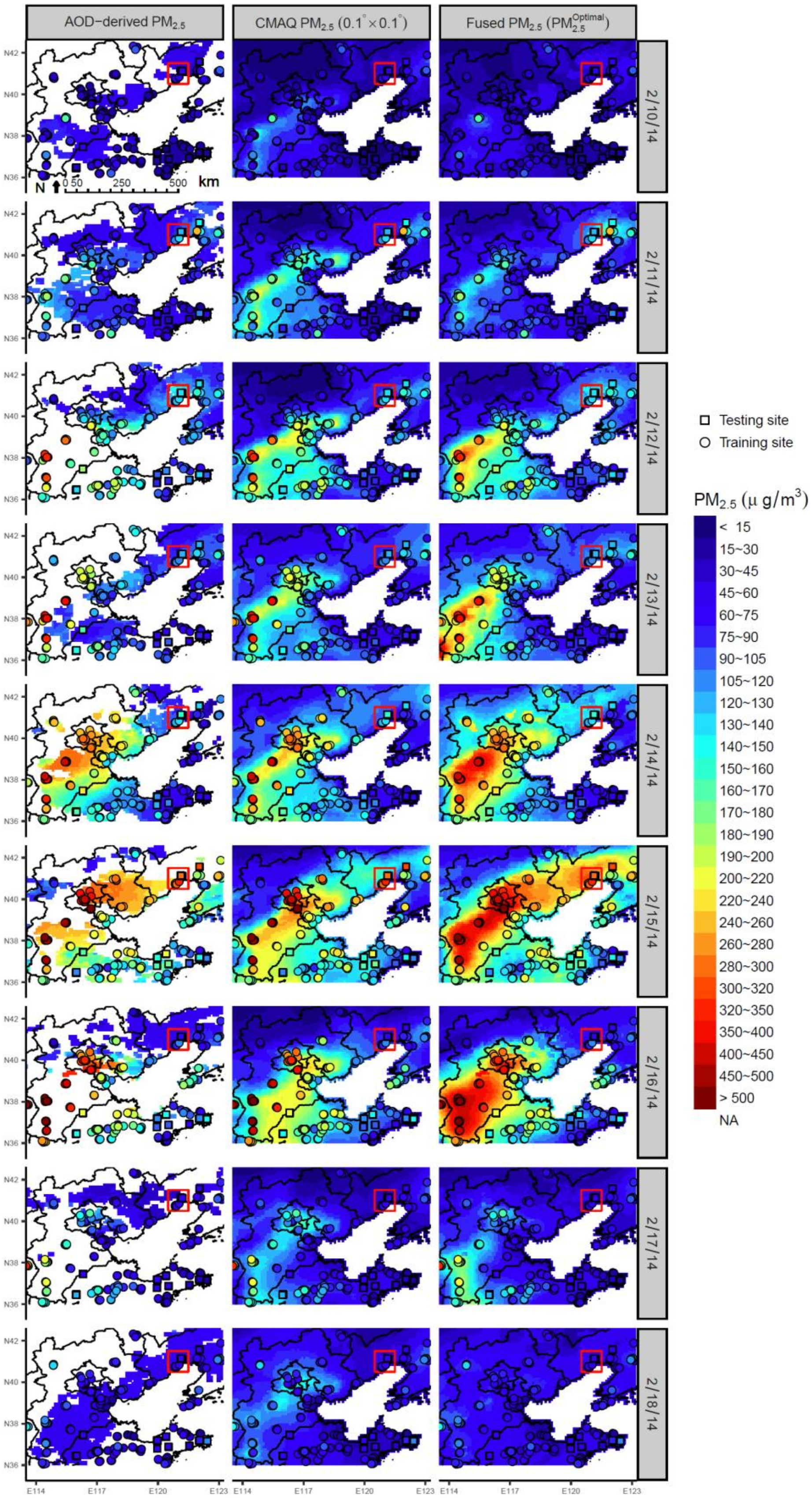


Figure 6. Seasonal maps (0.1° × 0.1°) of $PM_{2.5}$ during 2014 over China, produced by the three-stage model ($PM_{2.5}^{Optimal}$).

3.4 Exposure assessments based on the fused estimates

The fused estimator of $PM_{2.5}$ ($PM_{2.5}^{Optimal}$) will support exposure assessments in future health-related studies. AOD or CMAQ based estimator of $PM_{2.5}$ has been utilized to study long-term rather than short-term exposure to ambient pollution [18,49], because of data availability or data accuracy on daily scale. For example, we visualized spatiotemporal distributions of CMAQ simulations and AOD-derived $PM_{2.5}$ with the corresponding monitoring data during an episode of haze around Beijing-Tianjin-Hebei region in Figure 7. According to the maps, AOD-based method overlooked some hotspots due to incompleteness and could not capture the whole polluting procedure; whereas CMAQ simulations underestimated the severity of haze due to systematic errors. Unlike them, the fused estimates accurately characterized the growth, expansion and elimination of the haze. Therefore, $PM_{2.5}^{Optimal}$ can serves as exposure estimates to study either acute or chronic effects of $PM_{2.5}$. For example, combining $PM_{2.5}^{Optimal}$ with county-level data of China's sixth census, we assessed both annual and daily exposures to $PM_{2.5}$ across China in 2014. Accordingly, population-weighted concentration of annual exposure to ambient $PM_{2.5}$ was 55.7 μg/m³ and 82% of total population inhabited in the places exceeding WHO Air Quality Interim Target-1, 35 μg/m³; whereas population-weighted count of polluted or heavily-polluted days (defined as daily mean of $PM_{2.5}$ > 75 μg/m³ or 150 μg/m³ by CNAAQS) was 81 or 14 days, respectively.

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221



378

379 Figure 7 Spatiotemporal distributions of AOD-derived $PM_{2.5}$, CMAQ simulations and finally fused
380 $PM_{2.5}$ during an episode of haze around Beijing-Tianjin-Hebei region. In situ observational values are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at Remote Sens. 2017, 9, 221; doi:10.3390/rs9030221

visualized by the dots. The time-series of cross-validated values for the testing site located within the red rectangle are presented in Figure 2 (c).

4. Discussion

In this paper, we developed a three-stage model to estimate spatiotemporal variations of $PM_{2.5}$ through fusing CMAQ simulations, satellite remote sensing measurements and ground monitoring data together. We illustrated the method by a practice to generate daily $PM_{2.5}$ maps with a spatial resolution of 0.1° × 0.1° across China during 2014. The CV results evidenced that the fused estimator ($PM_{2.5}^{Optimal}$) was in good agreement with the observational $PM_{2.5}$, and outperformed the estimators based on either AOD or CMAQ data alone.

AOD-based methods have been widely utilized to estimate $PM_{2.5}$ concentrations on regional [43], national [37] or global scale [29]. Among them, LME or its extension has been more widely used because of computing efficiency. For example, in China, Ma, *et al.* [37] developed high-quality estimates of $PM_{2.5}$ covering a long-period from 2004 to 2013, through joining a LME and a generalized additive model (GAM) to retrieve $PM_{2.5}$ from MODIS AOD at 0.1° resolution; and tested their estimates by a CV_{10} of monitoring data in 2013 (RMSE = 27.99 μg/m³, R^2 = 0.78 and RPE = 36.3% for the first-stage estimator from LME; RMSE = 27.42 μg/m³, R^2 = 0.79 and RPE = 35.6% for the final estimator from LME+GAM). In this paper, we fitted a similar LME model through reducing the spatial-varying coefficient ($f(s)$) in Equation (1) to a one-dimensional random slope; and evaluated it by CV_{10} (RMSE = 23.3 μg/m³, R^2 = 0.69, RPE = 36.7%), which suggests that the LME method performs equally well on both our datasets and Ma, *et al.'s* [37]. However, LME has one disadvantage in modeling AOD in large scale. To incorporate geographical variations of the fitted parameters, LME were usually fitted separately by sub-regions (e.g. provinces), which resulted in spatially non-smoothing predictions near the boundaries of two sub-regions. Ma, *et al.* [37] addressed this issue by creating buffer zones around each province and averaging the overlapped predictions from neighboring provinces. The buffer-zone-averaging method introduced a side effect that the uncertainty (standard errors) of predictions averaged from two different LME models could not be quantified directly. Our LME_{SC} approach incorporated spatial variations of the modeling parameters by a nonlinear regression coefficient ($f(s)$), rather than fitting separate models, so that it produced more spatially smoothed estimates than LME. Our CV_{IS} analysis also confirmed that LME_{SC} slightly outperformed LME in developing AOD-derived $PM_{2.5}$ (RMSE = 26.2 μg/m³ and R^2 = 0.62 for LME_{SC}; RMSE = 26.8 μg/m³ and R^2 = 0.61 for LME).

Another weakness of AOD-based estimators was caused by non-random incompleteness in satellite measurements. In another word, AOD-derived values are more likely to be absent, when estimating $PM_{2.5}$ concentrations within a specific range. At the testing sites of CV_{IS}, AOD-derived $PM_{2.5}$ approximately covered 32%, 43%, 44% or 36% of unpolluted ($PM_{2.5}$ < 75 μg/m³), lightly-polluted (75 μg/m³ ≤ $PM_{2.5}$ < 115 μg/m³), moderately-polluted (115 μg/m³ ≤ $PM_{2.5}$ < 150 μg/m³) or heavily-polluted ($PM_{2.5}$ ≥ 150 μg/m³) days, respectively (as shown in Figure S5). In China, rainfalls AOD data tend to be missing during rainfalls, when particle concentrations are usually lower due to wet deposition. Such effect partially explains the lower sampling rate of AOD-derived $PM_{2.5}$ at unpolluted days (Figure S5). Whereas hazed episodes, especially in northern China, may be falsely classified as clouds by satellite and be neglected in current AOD algorithm, so that the sampling rate of AOD was also lower at heavily or severely polluted days. Long-termed averages of AOD-derived $PM_{2.5}$ can be biased from the truth due to the unevenly missing rates at different concentrations, which is known as sampling bias. Because the extreme values are less captured in their estimates, AOD-based methods may over−smooth the variability of $PM_{2.5}$. Previous studies showed sampling bias of AOD may lead to ±20% error in chronic exposure assessment of $PM_{2.5}$ [49]. Combining AOD-derived $PM_{2.5}$ with an spatiotemporally complete estimator, such as CMAQ simulations, can reduce the bias. Our step-specific CV_{IS} results showed that comparing model performance before and after fusing with $PM_{2.5}^{CMAQ}$, accuracy of intermediate estimator of $PM_{2.5}$ was considerably improved it in monthly (R^2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

= 0.67 for $PM_{2.5}^{AOD}$ vs. R^2 = 0.77 for $PM_{2.5}^{ML}$) or yearly scale (R^2 = 0.73 for $PM_{2.5}^{AOD}$ vs. R^2 = 0.81 for $PM_{2.5}^{ML}$), which may be explained by the reduction of sampling bias.

Air quality modeling results have been utilized in risk assessment of ambient pollutants [32] but rarely in epidemiological studies because of their low accuracy and potential bias. Data assimilation methods have been applied to improve predictability of air quality models. In China, Tang, *et al.* [50] first developed an EnKF to combine numerical outputs from the Nested Air Quality Prediction Modeling System (NAQPMS) [51] and in situ observations of Ozone; then Liu, *et al.* [18] applied a similar method to estimate daily $PM_{2.5}$ across China during 2013 and reported a RMSE of 30.2 μg/m³ by a five-fold CV, which is as accurate as our intermediate estimator, $PM_{2.5}^{CMAQ}$ (CV_{IS} RSME = 30.1 μg/m³). In this study, we improved raw CMAQ estimates (CV_{IS} RSME = 33.4 μg/m³, shown in Figure S3) is three aspects: (1) downscaling spatial resolution of CMAQ simulations to 0.1° (CV_{IS} RSME = 33.0 μg/m³, shown in Figure S2), (2) calibrating them with in situ observations (CV_{IS} CV_{IS} RSME = 30.1 μg/m³ for $PM_{2.5}^{CMAQ}$, shown in Figure 4 (c)) and (3) fusing them with AOD-derived $PM_{2.5}$ (CV_{IS} RSME = 28.2 μg/m³ for $PM_{2.5}^{ML}$, shown in Figure 4(d)). Although the data fusion step increased little on CV_{IS} RMSE, but significantly decreased the bias of CMAQ-based estimator (Bias = 14.8 μg/m³ for $PM_{2.5}^{CMAQ}$ vs. Bias = 7.7 μg/m³ for the $PM_{2.5}^{ML}$, fused by both $PM_{2.5}^{CMAQ}$ and $PM_{2.5}^{AOD}$), which reflected that AOD played a key role to control systemic error in data fusion.

In the final step of the three-stage model, we incorporated the spatiotemporal variations unexplained by $PM_{2.5}^{ML}$ through modeling the residuals by S/T-Kriging, which is analogous to the GAM stage in Ma, *et al.* [37]. Kriging has been proved to be mathematically equivalent to thin-plate regression splines, a specific type of GAM [40]. According to CV_{IS}, S/T-Kriging further decreased modeling error by 18% (RMSE = 28.2 μg/m³ for $PM_{2.5}^{ML}$ vs. RMSE = 23.0 μg/m³ for $PM_{2.5}^{Optimal}$), which indicated that the spatiotemporal autocorrelations should not be ignored in $PM_{2.5}$ modeling. Additionally, we also found that CV_{IS} errors of $PM_{2.5}^{Optimal}$ tended to be lower at the testing sites, which were surrounded by more training sites (Figure 8). Similar findings have been reported in previous studies, which introduced spatial or spatiotemporal autocorrelations into $PM_{2.5}$ modeling [52].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

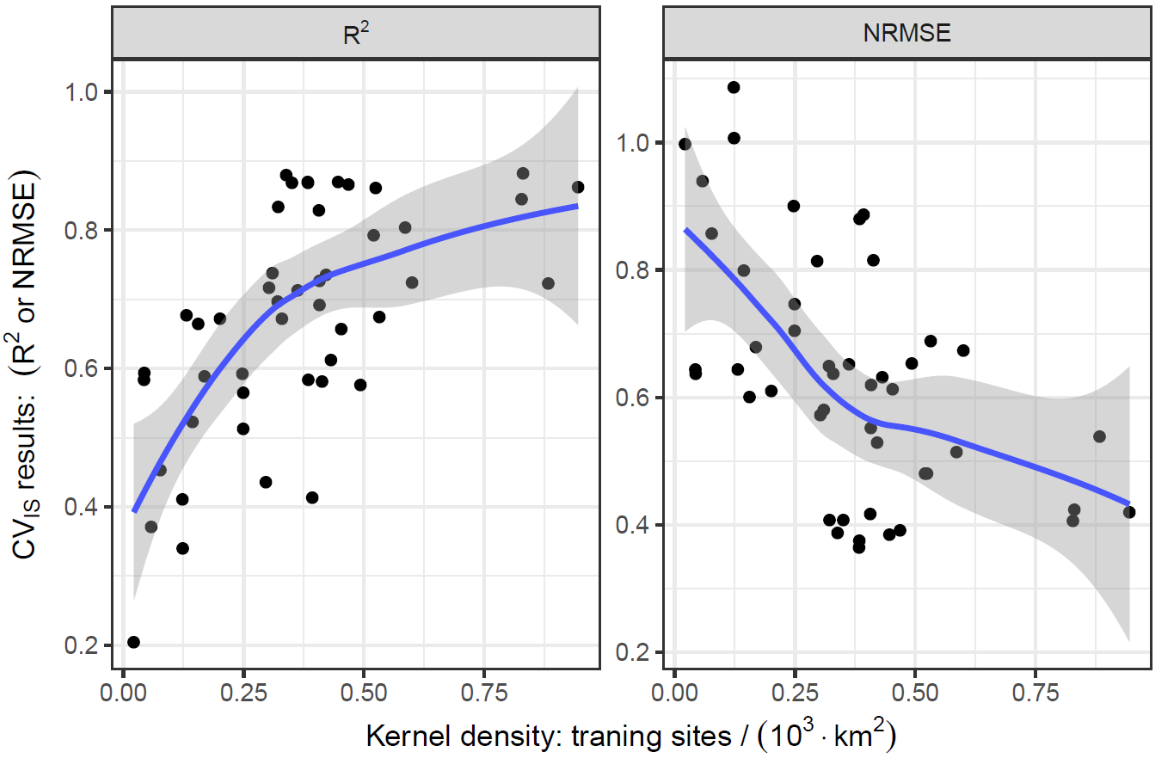


Figure 8 Different performances of CV_{IS} for the final estimator ($PM_{2.5}^{Optimal}$) by density of training sites, estimated by a 2-dimensional Kernel with a bandwidth of 50 kilometer. The blue curves and grey ribbons present the LOESS smoothing trends with corresponding confidence intervals.

Spatiotemporally autocorrelation-ship benefits prediction of $PM_{2.5}$, especially at the unmeasured locations but makes troubles for model evaluation. In CVs of auto-correlated variables, randomly selected testing data (e.g. CV_{10}) may not be independent of the training data, so that the predicting accuracy can be overestimated [48]. Through choosing the isolated monitoring sites in CV_{IS} approach, we attempted to use the testing records, which were less correlated with training data. We compared performance of CV_{IS} to that of CV_{10} in evaluating the AOD-derived $PM_{2.5}$ from the LME model, as shown in Figure S1. In comparison, we used a subset of CV_{10} to make sure that the two CVs were conducted on the same testing records. We found that CV_{10} error of the LME was consistent with the previous studies [37], but considerably lower than CV_{IS} error (CV_{10} RMSE = 23.3 μg/m^3 vs. CV_{IS} RMSE = 26.8 μg/m^3). The results suggested that CV_{10} might overestimate the predicting accuracy. Lv, *et al.* [48] addressed this issue through leaving out records from all monitors within a city simultaneously in CV, which is analogous to our approach, considering that monitors are usually clustered within cities but separated between different cities. Even though the models were evaluated by CV_{IS} in this paper, the influence of spatiotemporal autocorrelations on CVs cannot be avoided completely. In another word, the true predicting error of the three-stage model may be still underestimated in this paper.

The uncertainty of our study sources from three aspects. First, during our study period, the routine monitoring networks for ambient particles were too sparsely distributed to characterize some polluted sub-urban areas, such as undeveloped cities in the provinces of Henan and Shannxi. Second, satellite-derived AOD measurements played a key role to control bias in our approach but were only available at approximately one third of the predicting points. Increasing the spatiotemporal coverage of AOD (e.g., combing AOD from multiple satellites) will be considered in our future studies to reduce modeling uncertainty. Finally, CMAQ-WRF simulating procedures and inputted emission inventories may also contribute to the uncertainty of the three-stage model.

5. Conclusions

We developed a three-stage statistical model to estimate $PM_{2.5}$ concentrations through fusing in situ observations, satellite-derived AOD measurements and CMAQ simulations. We applied the method to produce daily maps of $PM_{2.5}$ over China at a spatial resolution of 0.1°. The final estimator of the three-stage model is shown to highly correlated with daily monitoring data (CV_{IS} R^2=0.72) and to outperform CMAQ-simulated $PM_{2.5}$ (CV_{IS} R^2=0.51) or AOD-derived $PM_{2.5}$ (CV_{IS} R^2=0.62). Our estimates will support future health-related studies on either acute or chronic exposure to ambient $PM_{2.5}$.

Supplementary Materials: The following are available online at www.mdpi.com/link, Figure S1: Scatterplots to compare CV_{10} and CV_{IS} using AOD-derived $PM_{2.5}$ from a LME model, Figure S2: Scatterplots of cross-validated values and their monthly or annual averages for downscaled CMAQ $PM_{2.5}$ (0.1° × 0.1°), Figure S3: Scatterplots of cross-validated values and their monthly or annual averages for raw CMAQ $PM_{2.5}$ (36 km × 36 km), Figure S4: Seasonal maps of $PM_{2.5}$ in 2014 over China, produced by CMAQ, intermediate and final estimators of the three-stage model, Figure S5 Comparisons of coverage rate (CR) of AOD-derived $PM_{2.5}$ by groups of observational $PM_{2.5}$ at the CV_{IS} testing sites, Figure S6 Temporal variations of CV results for the final estimator ($PM_{2.5}^{Optimal}$), Figure S7 Spatial distributions of CV results for the final estimator ($PM_{2.5}^{Optimal}$), Figure S8 Distributions of coefficients for AOD by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (1), Figure S9 Distributions of coefficients for CMAQ-simulated $PM_{2.5}$ by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (2).

Acknowledgments: This work was funded by the National Natural Science Foundation of China (41625020, 41571130032, 41222036) and the Public Welfare Program of China's Ministry of Environmental Protection (201509004 and 201309072) for Drs. Zhang, Q. and He, K..

Author Contributions: Dr. Xue, T. designed the statistical models and wrote the paper; Mr. Zheng, Y. simulated CMAQ results and analyzed the satellite data; Dr. Geng, G. analyzed monitoring and satellite data; Dr. Zheng, B. provided emission inventories; Dr. Jiang, X. reviewed literatures; Drs. Zhang, Q. and He, K. designed the whole study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dominici, F.; Peng, R.D.; Bell, M.L.; Pham, L.; Mcdermott, A.; Zeger, S.L.; Samet, J.M. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA* **2006**, *295*, 1127-1134.
2. Peng, R.D.; Bell, M.L.; Geyh, A.S.; Mcdermott, A.; Zeger, S.L.; Samet, J.M.; Dominici, F. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. *Environmental Health Perspectives* **2009**, *117*, 957-963.
3. Ritz, B.; Yu, F.; Fruin, S.; Chapa, G.; Shaw, G.M.; Harris, J.A. Ambient air pollution and risk of birth defects in southern california. *American Journal of Epidemiology* **2002**, *155*, 17-25.
4. Salam, M.T.; Millstein, J.; Li, Y.; Lurmann, F.; Margolis, H.G.; Gilliland, F.D. Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: Results from the children's health study. *Environmental Health Perspectives* **2005**, *113*, 1638-1644.
5. Sapkota, A.; Chelikowsky, A.P.; Nachman, K.E.; Cohen, A.; Ritz, B. Exposure to particulate matter and adverse birth outcomes: A comprehensive review and meta-analysis. *Air Quality, Atmosphere & Health* **2010**, *5*, 369-381.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1
Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

6. Wei, Y.; Han, I.; Shao, M.; Hu, M.; Zhang, J.; Tang, X. $PM_{2.5}$ constituents and oxidative DNA damage in humans. *Environmental Science & Technology* **2009**, *43*, 4757-4762.
7. Ren, C.; Park, S.K.; Vokonas, P.S.; Sparrow, D.; Wilker, E.H.; Baccarelli, A.; Suh, H.H.; Tucker, K.L.; Wright, R.O.; Schwartz, J. Air pollution and homocysteine: More evidence that oxidative stress-related genes modify effects of particulate air pollution. *Epidemiology* **2010**, *21*, 198-206.
8. Laden, F.; Schwartz, J.; Speizer, F.E.; Dockery, D.W. Reduction in fine particulate air pollution and mortality: Extended follow-up of the harvard six cities study. *American Journal of Respiratory and Critical Care Medicine* **2006**, *173*, 667-672.
9. Pope, C.A.; Burnett, R.T.; Turner, M.C.; Cohen, A.; Krewski, D.; Jerrett, M.; Gapstur, S.M.; Thun, M.J. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: Shape of the exposure–response relationships. *Environmental Health Perspectives* **2011**, *119*, 1616-1621.
10. Zhang, J.; Mauzerall, D.L.; Zhu, T.; Liang, S.; Ezzati, M.; Remais, J.V. Environmental health in china: Progress towards clean air and safe water. *The Lancet* **2010**, *375*, 1110-1119.
11. Zhang, Q.; He, K.; Huo, H. Policy: Cleaning china's air. *Nature* **2012**, *484*, 161-162.
12. Wang, Y.; Sun, M.; Yang, X.; Yuan, X. Public awareness and willingness to pay for tackling smog pollution in china: A case study. *Journal of Cleaner Production* **2016**, *112*, 1627-1634.
13. Li, G.X.; Zhou, M.G.; Zhang, Y.J.; Cai, Y.; Pan, X.C. Seasonal effects of PM_{10} concentrations on mortality in tianjin, china: A time-series analysis. *Journal of Public Health* **2012**, *21*, 135-144.
14. Chen, R.; Zhang, Y.; Yang, C.; Zhao, Z.; Xu, X.; Kan, H. Acute effect of ambient air pollution on stroke mortality in the china air pollution and health effects study. *Stroke* **2013**, *44*, 954-960.
15. Guo, Y.; Li, S.; Tian, Z.; Pan, X.; Zhang, J.; Williams, G.M. The burden of air pollution on years of life lost in beijing, china, 2004-08: Retrospective regression analysis of daily deaths. *BMJ* **2013**, *347*, 1-10.
16. Yang, Y.; Li, R.; Li, W.; Wang, M.; Cao, Y.; Wu, Z.; Xu, Q. The association between ambient air pollution and daily mortality in beijing after the 2008 olympics: A time series study. *PLOS ONE* **2013**, *8*.
17. Wu, S.; Deng, F.; Huang, J.; Wang, H.; Shima, M.; Wang, X.; Qin, Y.; Zheng, C.; Wei, H.; Yu, H. Blood pressure changes and chemical constituents of particulate air pollution: Results from the healthy volunteer natural relocation (hvnr) study. *Environmental Health Perspectives* **2013**, *121*, 66-72.
18. Liu, J.; Han, Y.; Tang, X.; Zhu, J.; Zhu, T. Estimating adult mortality attributable to $PM_{2.5}$ exposure in china with assimilated $PM_{2.5}$ concentrations based on a ground monitoring network. *Science of The Total Environment* **2016**, *568*, 1253-1262.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

19. Jerrett, M.; Burnett, R.T.; Ma, R.; Pope, C.A.; Krewski, D.; Newbold, K.B.; Thurston, G.D.; Shi, Y.; Finkelstein, N.; Calle, E.E. Spatial analysis of air pollution and mortality in los angeles. *Epidemiology* **2005**, *16*, 727-736.
20. Mercer, L.D.; Szpiro, A.A.; Sheppard, L.; Lindstrom, J.; Adar, S.D.; Allen, R.; Avol, E.L.; Oron, A.P.; Larson, T.V.; Liu, L.J.S. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (nox) for the multi-ethnic study of atherosclerosis and air pollution (mesa air). *Atmospheric Environment* **2011**, *45*, 4412-4420.
21. Henderson, S.B.; Beckerman, B.; Jerrett, M.; Brauer, M. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology* **2007**, *41*, 2422-2428.
22. Eeftens, M.; Beelen, R.; K, D.H.; Bellander, T.; Cesaroni, G.; Cirach, M.; Declercq, C.; Dėdelė, A.; Dons, E.; A, D.N. Development of land use regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PM_{coarse} in 20 european study areas; results of the escape project. *Environmental Science & Technology* **2012**, *46*, 11195-11205.
23. Martin, R.V. Satellite remote sensing of surface air quality. *Atmospheric Environment* **2008**, *42*, 7823-7843.
24. Paciorek, C.J.; Liu, Y.; Morenomacias, H.; Kondragunta, S. Spatiotemporal associations between goes aerosol optical depth retrievals and ground-level $PM_{2.5}$. *Environmental Science & Technology* **2008**, *42*, 5800-5806.
25. 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. *Environmental Science & Technology* **2012**, *46*, 11913-11921.
26. Ma, Z.; Hu, X.; Huang, L.; Bi, J.; Liu, Y. Estimating ground-level $PM_{2.5}$ in china using satellite remote sensing. *Environmental Science & Technology* **2014**, *48*, 7436-7444.
27. Beloconi, A.; Kamarianakis, Y.; Chrysoulakis, N. Estimating urban PM_{10} and $PM_{2.5}$ concentrations, based on synergistic meris/aatsr aerosol observations, land cover and morphology data ☆. *Remote Sensing of Environment* **2016**, *172*, 148-164.
28. Van Donkelaar, A.; Martin, R.V.; Park, R.J. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. *Journal of Geophysical Research* **2006**, *111*.
29. Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Kahn, R.A.; Levy, R.C.; Verduzco, C.; Villeneuve, P.J. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environmental Health Perspectives* **2010**, *118*, 847-855.
30. Byun, D.W.; Schere, K.L. Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system. *Applied Mechanics Reviews* **2006**, *59*, 51-77.

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

31. Nawahda, A.; Yamashita, K.; Ohara, T.; Kurokawa, J.; Yamaji, K. Evaluation of premature mortality caused by exposure to $PM_{2.5}$ and ozone in east asia: 2000, 2005, 2020. *Water Air and Soil Pollution* **2012**, *223*, 3445-3459.
32. Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367-371.
33. Bravo, M.A.; Fuentes, M.; Zhang, Y.; Burr, M.J.; Bell, M.L. Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation. *Environmental Research* **2012**, *116*, 1-10.
34. Beckerman, B.S.; Jerrett, M.; Serre, M.L.; Martin, R.V.; Lee, S.; Van Donkelaar, A.; Ross, Z.; Su, J.; Burnett, R.T. A hybrid approach to estimating national scale spatiotemporal variability of $PM_{2.5}$ in the contiguous united states. *Environmental Science & Technology* **2013**, *47*, 7233-7241.
35. Mcmillan, N.J.; Holland, D.M.; Morara, M.; Feng, J. Combining numerical model output and particulate data using bayesian space–time modeling. *Environmetrics* **2010**, *21*, 48-65.
36. Friberg, M.; Zhai, X.; Holmes, H.A.; Chang, H.H.; Strickland, M.J.; Sarnat, S.E.; Tolbert, P.E.; Russell, A.G.; Mulholland, J.A. Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution. *Environmental Science & Technology* **2016**, *50*, 3695-3705.
37. Ma, Z.; Hu, X.; Sayer, A.M.; Levy, R.C.; Zhang, Q.; Xue, Y.; Tong, S.; Bi, J.; Huang, L.; Liu, Y. Satellite-based spatiotemporal trends in $PM_{2.5}$ concentrations: China 2004-2013. *Environmental Health Perspectives* **2016**, *124*, 184-192.
38. Bey, I.; Jacob, D.J.; Yantosca, R.M.; Logan, J.A.; Field, B.D.; Fiore, A.M.; Li, Q.; Liu, H.Y.; Mickley, L.J.; Schultz, M.G. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research* **2001**, *106*, 23073-23095.
39. Zheng, B.; Zhang, Q.; Zhang, Y.; He, K.B.; Wang, K.; Zheng, G.T.; Duan, F.K.; Ma, Y.; Kimoto, T. Heterogeneous chemistry: A mechanism missing in current models to explain secondary inorganic aerosol formation during the january 2013 haze episode in north china. *Atmospheric Chemistry and Physics* **2015**, *15*, 2031-2049.
40. Cressie, N. Statistics for spatial data. *Terra Nova* **1993**, *4*, 613-617.
41. Randriamiarisoa, H.; Chazette, P.; Couvert, P.; Sanak, J.; Megie, G. Relative humidity impact on aerosol parameters in a paris suburban area. *Atmospheric Chemistry and Physics* **2006**, *6*, 1389-1407.
42. Wang, Z.; Chen, L.; Tao, J.; Zhang, Y.; Su, L. Satellite-based estimation of regional particulate matter (PM) in beijing using vertical-and-rh correcting method. *Remote Sensing of Environment* **2010**, *114*, 50-63.
43. Zheng, Y.; Zhang, Q.; Liu, Y.; Geng, G.; He, K. Estimating ground-level $PM_{2.5}$ concentrations over three megalopolises in china using satellite-derived aerosol optical depth measurements. *Atmospheric Environment* **2016**, *124*, 232-242.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 February 2017 doi:10.20944/preprints201702.0059.v1

Peer-reviewed version available at *Remote Sens.* **2017**, *9*, 221; doi:10.3390/rs9030221

44. Cressie, N.; Johannesson, G. Fixed rank kriging for very large spatial data sets. *Journal of The Royal Statistical Society Series B-statistical Methodology* **2008**, *70*, 209-226.
45. Wood, S. *Generalized additive models: An introduction with r*. CRC press: 2006.
46. Cressie, N.; Wikle, C.K. *Statistics for spatio-temporal data*. John Wiley & Sons: 2015.
47. Bunke, O.; Droge, B. Bootstrap and cross-validation estimates of the prediction error for linear regression models. *Annals of Statistics* **1984**, *12*, 1400-1424.
48. Lv, B.; Hu, Y.; Chang, H.H.; Russell, A.G.; Bai, Y. Improving the accuracy of daily pm2. 5 distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in north china. *Environmental science & technology* **2016**, *50*, 4752-4759.
49. Geng, G.; Zhang, Q.; Martin, R.V.; van Donkelaar, A.; Huo, H.; Che, H.; Lin, J.; He, K. Estimating long-term $PM_{2.5}$ concentrations in china using satellite-based aerosol optical depth and a chemical transport model. *Remote Sensing of Environment* **2015**, *166*, 262-270.
50. Tang, X.; Zhu, J.; Wang, Z.; Gbaguidi, A. Improvement of ozone forecast over beijing based on ensemble kalman filter with simultaneous adjustment of initial conditions and emissions. *Atmospheric Chemistry and Physics* **2011**, *11*, 12901-12916.
51. Wang, Z.; Maeda, T.; Hayashi, M.; Hsiao, L.-F.; Liu, K.-Y. A nested air quality prediction modeling system for urban and regional scales: Application for high-ozone episode in taiwan. *Water, Air, and Soil Pollution* **2001**, *130*, 391-396.
52. Lee, S.-J.; Serre, M.L.; van Donkelaar, A.; Martin, R.V.; Burnett, R.T.; Jerrett, M. Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient $PM_{2.5}$ concentrations across the continental united states. *Environmental health perspectives* **2012**, *120*, 1727.