



Published in final edited form as:

*Remote Sens Environ.* 2021 February ; 253: . doi:10.1016/j.rse.2020.112203.

## Estimating PM<sub>2.5</sub> concentrations in Northeastern China with full spatiotemporal coverage, 2005–2016

Xia Meng<sup>a,1</sup>, Cong Liu<sup>a,1</sup>, Lina Zhang<sup>a</sup>, Weidong Wang<sup>a</sup>, Jennifer Stowell<sup>c</sup>, Haidong Kan<sup>a,b,\*\*</sup>, Yang Liu<sup>d,\*</sup>

<sup>a</sup>School of Public Health, Fudan University, Shanghai, China

<sup>b</sup>Children's Hospital of Fudan University, National Center for Children's Health, Shanghai 201102, China

<sup>c</sup>Boston University, School of Public Health, Boston, MA, USA

<sup>d</sup>Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

### Abstract

Predicting long-term spatiotemporal characteristics of fine particulate matter (PM<sub>2.5</sub>) is important in China to understand historical levels of PM<sub>2.5</sub>, to support health effects research of both long-term and short-term exposures to PM<sub>2.5</sub>, and to evaluate the efficacy of air pollution control policies. Satellite-retrieved aerosol optical depth (AOD) provides a unique opportunity to characterize the long-term trends of ground-level PM<sub>2.5</sub> at high spatial resolution. However, the missing rate of AOD in Northeastern China (NEC) is very high, especially in winter, and challenges the accuracy of long-term predictions of PM<sub>2.5</sub> if left unresolved. Using random forest algorithms, this study developed a gap-filling approach combining satellite AOD, meteorological data, land use parameters, population and visibility in the NEC during 2005–2016. The model, including all predictors, combined with a model without AOD was able to fill the gap of PM<sub>2.5</sub> predictions caused by missing AOD at 1-km resolution. The R<sup>2</sup> (RMSE) of the full-coverage predictions was 0.81 (18.5 μg/m<sup>3</sup>) at the daily level. Gap-filled PM<sub>2.5</sub> predictions on days with missing AOD reduced the relative prediction error from 28% to 2.5% in winter. The leave-one-year-out-cross-validation R<sup>2</sup> (RMSE) of the full-coverage predictions was 0.65 (16.3 μg/m<sup>3</sup>) at the monthly level, indicating relatively high accuracy of predicted historical PM<sub>2.5</sub> concentrations. Our results suggested that AOD helped increase the reliability of historical PM<sub>2.5</sub> prediction when ground PM<sub>2.5</sub> measurements were unavailable, even though predictions from the AOD model only accounted for approximate 37% of the whole dataset. Predicted PM<sub>2.5</sub> level in NEC have

\* Correspondence to: Y. Liu, Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Road, Atlanta, GA 30322, USA. yang.liu@emory.edu (Y. Liu). \*\* Correspondence to: H. Kan, Department of Environmental Health, School of Public Health, Fudan University, P.O. Box 249, 130 Dong-An Road, Shanghai 200032, China. kanh@fudan.edu.cn (H. Kan).

<sup>1</sup>Xia Meng and Cong Liu contributed equally to this work.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2020.112203>.

increased since 2005, reached its peak during 2013–2015, then saw a major decline in 2016. Our high-resolution predictions also showed a south to north gradient and many pollution hot spots in the city clusters surrounding provincial capitals, as well as within large cities. Overall, by combining predictions from the AOD model with higher accuracy and predictions from the non-AOD model to achieve full coverage, our modeling approach could produce long-term, full-coverage historical PM<sub>2.5</sub> levels in high-latitude areas in China, despite the widespread and persistent AOD missingness.

## Keywords

AOD; PM<sub>2.5</sub>; Gap-filling; Random forest; Northeastern China

---

## 1. Introduction

It has been widely acknowledged that exposure to fine particulate matter (PM<sub>2.5</sub>, particulate matter with aerodynamic diameter less than 2.5 μm) is related to adverse effects on various health outcomes (Liu et al., 2019; Nachman et al., 2016; Pope et al., 2019; Power et al., 2015). According to the estimates from the Global Burden of Diseases (GBD) 2017, ambient PM<sub>2.5</sub> ranks as the eighth leading risk factor for death worldwide, by causing 2.94 million deaths globally (Stanaway et al., 2018). The GBD study estimated the disease burden attributed to PM<sub>2.5</sub> using an integrated exposure-response (IER) model (Burnett et al., 2014). Few cohort studies in areas with high PM<sub>2.5</sub> levels, e.g. China, were included in the IER model. Hence, conducting cohort studies in China on the associations between long-term PM<sub>2.5</sub> exposure and health outcomes could provide more evidence necessary to estimate disease burden attributable to PM<sub>2.5</sub>, which has been hindered by exposure assessment of PM<sub>2.5</sub> to a large extent. Relatively few sources of China PM<sub>2.5</sub> data are available prior to 2013 when the national air quality network was established. Additionally, these monitoring stations tend to be unevenly-distributed, with most located in urban centers. Developing models capable of predicting PM<sub>2.5</sub> concentrations at high spatial resolution and estimating historical PM<sub>2.5</sub> levels with high accuracy is crucial for understanding the long-term spatiotemporal trends of PM<sub>2.5</sub> and supporting PM<sub>2.5</sub>-related cohort studies in China.

Chemical transport models (CTMs) are a classical method that can provide simulations of PM<sub>2.5</sub> concentrations without relying on PM<sub>2.5</sub> measurements. For example, the Community Multi-scale Air Quality (CMAQ) modeling system was used to predict PM<sub>2.5</sub> levels in Eastern China during 2010, and the correlation coefficients between predictions and measurements of PM<sub>2.5</sub> were 0.64 and 0.56 at two ground monitoring stations (Zhao et al., 2013). Nationwide PM<sub>2.5</sub> levels were also predicted in China in 2013 with CMAQ. However, the results tended to underestimate PM<sub>2.5</sub>—especially in spring and summer (Hu et al., 2016a). Due to the missingness of mechanisms and imperfect emission inventories, the simulations of CTMs tend to have high bias (Geng et al., 2017; Hu et al., 2016b; Zheng et al., 2015). Satellite data, such as aerosol optical depth (AOD) with wide temporal and spatial coverage, combined with advanced modeling strategies, have been increasingly used in predicting ground PM<sub>2.5</sub> levels in China, especially since 2013 when PM<sub>2.5</sub> measurements

became available nationwide(He and Huang, 2018; Li et al., 2017; Ma et al., 2014; Wei et al., 2019; Zheng et al., 2016). As a measure of aerosol light extinction and a column measurement of aerosol loading, AOD is a quantitative estimate of aerosol abundance in the atmosphere. It is widely known that the relationship between AOD and  $PM_{2.5}$  is influenced by weather and land use, and can vary in both space and time (Paciorek and Liu, 2009). Previous studies have proposed various methods to characterize this relationship, such as calibration based on simulations from chemical transport models (Donkelaar et al., 2010), identifying daily variation with linear mixed effects model (Ma et al., 2016), and assessing data-driven, non-linear  $PM_{2.5}$ -AOD associations with machine learning algorithms (Di et al., 2016; Xiao et al., 2018). Along with the demands of understanding the long-term trend of  $PM_{2.5}$  and evaluating the effectiveness of air quality control policies, a few studies have begun to predict historical  $PM_{2.5}$  levels in China (Lin et al., 2018; Ma et al., 2016; Xiao et al., 2018; Xue et al., 2019b). These studies have utilized multiple approaches, including development of a two-stage spatial statistical model to predict  $PM_{2.5}$  concentrations in China during 2004–2013 at 10-km resolution (Ma et al., 2016); evaluating long-term trends of  $PM_{2.5}$  in 19 provinces of China during 2001–2015 at 1-km resolution(Lin et al., 2018); and assessing  $PM_{2.5}$  variations between 2006 and 2017 in China based on a two-stage statistical model at 1 degree resolution (Xue et al., 2019b). However, gap-filling of daily predictions due to missing AOD was not utilized in these studies.

One limitation of employing AOD to predict ground level of  $PM_{2.5}$  is the non-random missingness of data. The absence of  $PM_{2.5}$  predictions at the daily level induced by missing AOD could bias time-aggregated  $PM_{2.5}$  estimates. For example, annual mean  $PM_{2.5}$  concentration was overestimated by approximately  $10 \mu\text{g}/\text{m}^3$  without gap-filling missing AOD in Yangtze River Delta (YRD) of China in 2013–2014 (Xiao et al., 2017). Geostatistical and regression methods have been proposed in previous studies to fill missing AOD or  $PM_{2.5}$  predictions caused by non-random missingness of AOD and ground  $PM_{2.5}$  observations were commonly used as predictors in some of these studies when estimating the spatiotemporal trends of missing AOD or  $PM_{2.5}$  predictions (de Hoogh et al., 2018; Goldberg et al., 2019; Kloog et al., 2012; Lv et al., 2016). For example, the ratio of daily and seasonal  $PM_{2.5}$  measurements joint with interpolation method was used to estimate missing AOD (Lv et al., 2016). Additionally, a spatial smoothing model based on  $PM_{2.5}$  measurements was developed in the Mid-Atlantic region of the United States to increase the coverage of predicted  $PM_{2.5}$  when using models that include AOD (Kloog et al., 2012). Models dependent on  $PM_{2.5}$  observations are unable to predict historical  $PM_{2.5}$  levels prior to the establishment of the national air quality network. Other studies have employed multiple imputation, machine learning or deep learning methods by including meteorological parameters, land cover data and other predictors to fill missing AOD (Bi et al., 2019; Li et al., 2020; Xiao et al., 2017; Zhao et al., 2019a). In the NEC, where AOD retrieval failures are extensive for a prolonged period, associations derived from data outside the region and time period might fail to capture the true spatiotemporal patterns and lead to inaccurate predictions. Models without AOD have also been proposed to fill missing PM predictions. For example, linear mixed effects models or gradient boosting algorithms were used to predict  $PM_{2.5}$  or  $PM_{10}$  on grid cell-days without valid AOD (Meng et al., 2016;

Zhan et al., 2017). However, the model performance for gap-filling PM<sub>2.5</sub> in high-latitudes and historical periods have not been fully evaluated.

Cloud cover is the main reason for AOD missingness, while high surface brightness can also result in AOD missingness over desert or snow or ice cover (Belle et al., 2017; Bi et al., 2019; Chu et al., 2002; Hsu et al., 2013). A large number of people reside in high-latitude areas of the Northern Hemisphere, including Northeastern China (NEC), the New England region of the US, Northern European countries and Russia, where the land is covered by snow or ice for an extended period in winter. So far, only a few regional studies, e.g. in northeastern U.S. and Switzerland (Bi et al., 2019; de Hoogh et al., 2018; Kloog et al., 2012), have attempted gap-filling of predictions resulting from missing AOD in these regions. Both PM<sub>2.5</sub> levels and variability are significantly higher in China than in the US (Cheng et al., 2016), hence the previous methods may not be applicable in the NEC. Specifically, approximately 50% of the NEC areas within the study domain have no valid AOD value in winter. Therefore, this study aims to estimate long-term PM<sub>2.5</sub> levels at full temporal and spatial coverages in NEC.

There are approximately 100 million people living in NEC, but little is known about the long-term trend of air pollution there. The NEC, home to the Northeast Old Industrial Base (NOIB), used to be China's industrial heartland and has suffered a chronic economic recession since the 1990s. To stimulate regional development, the Chinese central government implemented the Northeast Old Industrial Base Revitalization Strategy in 2003. These measures have increased energy consumption rapidly in this region (Li et al., 2016), with potentially significant impacts on air quality. Additionally, the NEC has severe PM<sub>2.5</sub> air pollution especially in winter, when PM<sub>2.5</sub> concentrations can be more than twice the summer level during 2013–2016 (Table 1). For example, the heating systems ran for 146 days in 2015 in Shenyang City with population of more than 8 million in NEC. The daily mean concentrations of PM<sub>2.5</sub> were higher than 75 µg/m<sup>3</sup> on 95 days out of these 146 days, with the highest levels measured at over 600 µg/m<sup>3</sup> in Shenyang in 2015. On the other hand, the AOD missing rate in NEC is high, caused by both cloud cover and snow cover. Therefore, a gap-filling approach is needed to correct the prediction errors resulting from AOD sampling bias to produce long-term estimates of PM<sub>2.5</sub> with high accuracy in NEC. These methods could also provide references of modeling strategies for areas similar to NEC. Finally, though PM<sub>2.5</sub> levels are high in NEC, fewer studies on associations of PM<sub>2.5</sub> and various health outcomes have been done in the region compared with other polluted areas of China. Since the majority of previous studies have been based on measurements from a limited number of monitoring stations, this could introduce exposure misclassification that are necessary for epidemiological studies. Therefore, high quality exposure models are crucial to better understand the spatiotemporal characteristics of PM<sub>2.5</sub> and support further epidemiological studies in NEC.

In this study, a gap-filling approach with random forest models at 1-km resolution was developed during 2013–2016 in NEC, where AOD missing rate was extremely high due to cloud cover, haze as well as lands covered by snow or ice, but few previous studies have done there. This study also analyzed the representativeness of predictions with respect to absence of PM<sub>2.5</sub> predictions on days without available AOD and the role of AOD

in predicting historical PM<sub>2.5</sub> concentrations. Long-term trends in 2005–2016 and spatial characteristics of PM<sub>2.5</sub> concentrations in NEC were also characterized based on the full-coverage predictions of PM<sub>2.5</sub> produced by the gap-filling approach.

## 2. Data and methods

Our study domain covered three provinces located in NEC with an area of more than 780,000 km<sup>2</sup> and approximately 100 million people, including Heilongjiang, Jilin and Liaoning Provinces from north to south. NEC is dominated by a temperate monsoon climate, characterized by cold and dry winters. There are up to 173 monitoring stations of PM<sub>2.5</sub> in NEC (Fig. 1). Most of the regional population resides along the Northeast China Plain (approximate 350,000 km<sup>2</sup>), which is a main agricultural and heavy industrial center in China (Zhao et al., 2019b).

A significant challenge for developing models of PM<sub>2.5</sub> in NEC is the high rate of missing AOD. Throughout the entire study domain, the annual mean AOD coverage was 32% during 2005–2016, with decreased coverage in the east and south (Fig. 2). Most of NEC had extremely low AOD coverage in winter. For example, approximately 50% of areas within the study domain have no valid AOD value during the 2015 winter (December 2015 – February 2016) (Fig. 2). In addition, the monthly mean AOD coverage in Harbin, the capital city of Heilongjiang Province, is highest in autumn (47%) and as low as 11% in winter. A 50-km buffer was added around the provincial boundary within China to make sure the accuracy of predictions near the boundary of the study domain.

### 2.1. Ground measurements of PM<sub>2.5</sub> concentrations

Ground measured hourly PM<sub>2.5</sub> concentrations within the study domain from the National Air Quality Monitoring Network are collected from the China National Environmental Monitoring Center (CNEMC) at <http://www.cnemc.cn/>. Average daily PM<sub>2.5</sub> concentrations of up to 173 monitoring stations were calculated from at least 18 valid hourly values per day during 2013–2016 for each station for model development.

### 2.2. Remote sensing AOD data

High-resolution AOD products have been increasingly applied in capturing local gradients of PM<sub>2.5</sub> spatial distribution (Bi et al., 2019; Di et al., 2016; Xiao et al., 2017). In this study, the MAIAC (Multi-Angle Implementation of Atmospheric Correction) AOD data with a spatial resolution of 1-km was used. High resolution AOD was retrieved based on the MAIAC algorithm of the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua satellites (Lyapustin et al., 2011a; Lyapustin et al., 2011b). Gridded MAIAC AOD data were downloaded from the Earth Data portal (<https://search.earthdata.nasa.gov>) and values with quality assurance flags indicating cloud contaminated or covered by snow/water were cleaned. AOD data from 2013 to 2016 were used for model development and data from 2005 to 2012 were used for prediction. The 1-km resolution modeling grid based on MAIAC AOD was used to integrate all parameters. PM<sub>2.5</sub> monitoring sites were spatially matched to the 1-km grid cells. Model development and PM<sub>2.5</sub> predictions were done at the grid level.

### 2.3. Other predictors

Additional predictors related to  $PM_{2.5}$  concentrations were used to improve the accuracy of predicting  $PM_{2.5}$ , especially when predicting historical  $PM_{2.5}$  levels when  $PM_{2.5}$  measurements were unavailable. Overall, simulated  $PM_{2.5}$  values, visibility, meteorological parameters, land use data as well as population density were included for modeling and predicting  $PM_{2.5}$  concentrations in NEC.

Daily simulated  $PM_{2.5}$  concentrations were calculated based on simulations from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Randles et al., 2017). MERRA-2 simulations have full spatial and temporal coverage and provide additional information regarding  $PM_{2.5}$  trends. The MERRA-2  $PM_{2.5}$  at  $0.5^\circ \times 0.625^\circ$  resolution was interpolated into the 1-km grid cells by inverse distance weighted (IDW) method at daily level.

Previous studies have found visibility closely associated with ground  $PM_{2.5}$  levels and a promising predictor to estimate historical  $PM_{2.5}$  levels (Liu et al., 2017). Therefore, hourly visibility data from ground meteorological monitoring stations were downloaded from National Oceanic and Atmospheric Administration (NOAA) of United States. The hourly visibility data was aggregated to the daily level and interpolated to the 1-km grid cells by IDW method.

Meteorological data have been proven to play crucial roles in predicting  $PM_{2.5}$  concentrations; therefore, we included relative humidity and air temperature at 2 m, east-west and north-south components of the wind vector plus wind speed at 10 m, total precipitation, as well as planetary boundary layer height (PBLH) in this study. All meteorological variables were obtained from Goddard Earth Observing System Data Assimilation System GEOS-5 Forward Processing (GEOS5-FP) at  $0.25^\circ \times 0.3125^\circ$  for 2013–2016 and from Goddard Earth Observing System Model, version 5 (GEOS5) at  $0.5^\circ \times 0.666^\circ$  for 2005–2012. All meteorological data were interpolated to the 1-km grid cells by IDW method. Since model-simulated aerosol vertical profiles are prone to large errors and lidar-measured aerosol vertical profiles are too sparse for daily, full-coverage  $PM_{2.5}$  modeling, we included weather parameters such as boundary layer height (indicator of vertical mixing) and wind speed (indicator of horizontal mixing) as predictors in our model. Vertical profile of aerosol with high spatiotempory resolution and coverage matched with AOD could be considered in future studies.

Land use, cloud, distance and population variables were provided by several sources. Cloud fraction (CF) data at 5-km resolution from Aqua and Terra (MYD06\_L2 and MOD06\_L2), were downloaded from the Earth Data portal (<https://search.earthdata.nasa.gov>). The daily mean CF values of Aqua and Terra were interpolated to MAIAC grids by IDW method. Normalized Difference Vegetation Index (NDVI) at 1-km resolution from MODIS 16-day NDVI product (MOD13A2), elevation data at 30-m resolution from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Map (GDEM), road length within the 1-km grid, population data at 1-km resolution from the LandScan Global Population Database were matched or integrated with the 1-km grid cells.

## 2.4. Model development

Machine learning algorithms have been increasingly used in predicting concentrations of particulate matter in recent years for its flexibility in building associations between dependent and independent variables and high capacity of handling large amount of predictors (Di et al., 2016; Hu et al., 2017; Meng et al., 2018a; Stafoggia et al., 2019; Wei et al., 2019). Random forest was employed in this study to develop models during 2013–2016 and predict  $PM_{2.5}$  levels during 2005–2016, by incorporating MAIAC AOD, MERRA-2 simulated  $PM_{2.5}$  concentrations, visibility, meteorological variables, land use data and population data. Random forest models with AOD on days and at grid cells of available AOD (AOD model) and without AOD (non-AOD model) on all days and at all grid cells were developed separately. The non-AOD model included the same predictors as the AOD model except MAIAC AOD. Ground  $PM_{2.5}$  measurements were used as the dependent variable and all mentioned predictors were used as independent variables to train the random forest. Random forest, as one of machine learning algorithms, is based on decision trees. It has been introduced in various previous studies and were briefly described here (Breiman, 2001; Hu et al., 2017; Liaw, 2015; Meng et al., 2018b). To train each tree in the forest, a training dataset is drawn from the original dataset using a bootstrap sampling method with replacement, then predictions are made based on data outside of the training dataset. Predictions from all trees are averaged and compared with their corresponding measurements to evaluate the model performance. To develop each tree in the forest, a random subset of the predictors is selected at each candidate split in the learning process, with the assumption that, if a few variables are very predictive for the dependent variable, they will be selected in many of these trees. The AOD-model and non-AOD model could be considered as we split the overall training dataset manually according to the availability of AOD data. In this manner,  $PM_{2.5}$  concentrations were predicted using the AOD model in grid cells and on days with available AOD, or using the non-AOD model where and when AOD was missing. The gap-filling approach combined results from two separate random forest models to produce predictions at a 100% spatiotemporal coverage. One strength of random forest is to provide importance ranking of predictors by comparing the mean squared error (MSE, defined as the average of the squares of the differences between the predictions and measurements) between predictions before and after that variable is permuted while all others are left unchanged.

Model overfitting was tested by 10-fold cross validation (CV) for predictions from AOD model, non-AOD model and combined predictions, respectively. The spatial interpolation accuracy was evaluated by a spatial 10-fold CV, i.e., randomly holding out 10% of monitoring stations, training the model with the remaining 90% of the stations, making predictions on the held-out stations, and repeating this process 10 times. Previous studies generally implemented regular temporal 10-fold CV by randomly choosing 10% days of data each time. However, a much stricter CV procedure was selected in this study by dropping an entire year of data each time (leave-one-year-out-cross-validation) to test the capacity of predicting  $PM_{2.5}$  beyond modeling period. The regression  $R^2$  and root mean square error (RMSE) were calculated between predictions and measurements to indicate the accuracy of combined predictions for spatial and temporal CV.

### 3. Results and discussion

#### 3.1. Description of PM<sub>2.5</sub> measurements

Fig. 1 showed the study domain with PM<sub>2.5</sub> monitoring stations during 2013–2016 and elevation in NEC. The monitoring network of air quality has been expanding since it was established in 2013 with 35 stations in NEC, and the number increased up to 173 in 2016. The mean PM<sub>2.5</sub> measured concentrations of all stations during 2013–2016 was 48.6 µg/m<sup>3</sup> with standard deviation of 41.9 µg/m<sup>3</sup>. The annual mean values of PM<sub>2.5</sub> concentrations were 58.9, 61.9, 60.3 and 46.3 µg/m<sup>3</sup> for each year from 2013 to 2016, respectively, at 35 sites with measurements for all four years. This time trend based on measurements suggests that the PM<sub>2.5</sub> level in NEC was much higher than the level 2 annual limit of PM<sub>2.5</sub> (35 µg/m<sup>3</sup>) in the Chinese Ambient Air Quality Standard (GB 3095–2012) in 2013–2015, and dropped substantially by ~23% in 2016 compared to levels in 2013. A total of 145,026 PM<sub>2.5</sub> ground observations were included in the non-AOD model; while only 37% of them were matched with valid AOD values in the AOD model. The correlation coefficient was 0.40 for all the pairs of PM<sub>2.5</sub> measurements and AOD, which was comparable with previous study (Yang et al., 2019).

#### 3.2. Model performance

The 10-fold CV R<sup>2</sup> (RMSE) values were 0.81 (16.3 µg/m<sup>3</sup>), 0.81 (18.6 µg/m<sup>3</sup>), and 0.81 (18.5 µg/m<sup>3</sup>) for AOD model, non-AOD model and combined predictions, respectively. The importance ranking of predictors produced by random forest models were shown in Fig. S1. AOD was the second most important predictor in the AOD-model. MERRA-2 PM<sub>2.5</sub> simulations, humidity, temperature, PBLH and wind were important predictors in both the AOD-model and non-AOD model. Fig. 3 illustrated the density plots of linear regressions between measurements (y-axis) and 10-fold CV predictions of PM<sub>2.5</sub> (x-axis) for AOD model (a), non-AOD model (b) and combined predictions (c). The results showed that the AOD-model ( $N = 54,328$ ) and non-AOD model ( $N = 145,026$ ) performed similarly well. However, the AOD model had smaller prediction errors with lower RMSE. With the constraint of PM<sub>2.5</sub> measurements and the advantage of larger data size, the predictions from the non-AOD model had relatively high accuracy during our modeling period. Therefore, predictions from the non-AOD model could be used to gap-fill PM<sub>2.5</sub> values with relatively high accuracy. Two-stage statistical models have been developed in seven regions of China to predict PM<sub>2.5</sub> concentrations at 1-degree resolution based on one-year data of 2016, and the model 10-fold CV R<sup>2</sup> and RMSE in NEC region were 0.71 and 13.5 µg/m<sup>3</sup>, respectively (Xue et al., 2019b). As an improvement over previous studies, the gap-filling random forest model in this study has significantly finer spatial resolution (1-km resolution) and higher CV R<sup>2</sup> (0.81) than the two-stage model; though the RMSE (18.5 µg/m<sup>3</sup>) was slightly higher. This increase may be due to higher PM<sub>2.5</sub> concentrations included in the modeling dataset (up to ~600 µg/m<sup>3</sup>) compared with levels in the previous study (up to ~300 µg/m<sup>3</sup>).

Since the rate of AOD missingness varied significantly by season, the representativeness of predictions was evaluated (Table 1). Monthly mean PM<sub>2.5</sub> concentrations were calculated with all available daily PM<sub>2.5</sub> measurements, combined predictions and predictions from the AOD model only. Seasonal mean values were calculated based on monthly mean values.



The missing rate for each season were given as a reference in Table 1. Measurements and combined predictions of PM<sub>2.5</sub> were of 100% coverage, and the coverage for predictions from the AOD-model only was calculated as (1 - missing rate). Overall, the prediction errors at seasonal level were within 10% in spring, summer and fall, when the AOD missing rates and the PM<sub>2.5</sub> levels were both relatively low. The prediction error in winter was 28% when using predictions only from AOD-model with 80% missing rate of AOD, but only 2.5% when using the combined predictions with full coverage. The predictions without gap-filling dramatically underestimated PM<sub>2.5</sub> concentrations in winter, when the AOD missing rate was highest. A comparison of the annual mean and seasonal mean of PM<sub>2.5</sub> measurements, full-coverage predictions, and predictions from the AOD-model only in NEC during 2013–2016 was shown in Fig. S2. Nearly half of the grid cells lack predictions in winter without gap filling, which would lead to a significant underestimation of PM<sub>2.5</sub> in winter and at the annual level in NEC. Previous research reported that non-gap-filled predictions tended to overestimate the PM<sub>2.5</sub> level in YRD of China, since missing AOD normally occurs on cloudy days when PM<sub>2.5</sub> tends to be lower than on clear days (Xiao et al., 2017). Unlike YRD, mean measured concentrations of PM<sub>2.5</sub> in winter during 2013 and 2016 in NEC was 73.5 µg/m<sup>3</sup> on days without valid AOD, versus 57.2 µg/m<sup>3</sup> on days with valid AOD. Consequently, missing AOD results in underestimation of PM<sub>2.5</sub> in winter in NEC, demonstrating the necessity to predict PM<sub>2.5</sub> with full coverage.

The slight differences between measurements and our full-coverage predictions might because measurements are PM<sub>2.5</sub> levels at monitoring points, while predictions are average PM<sub>2.5</sub> levels within the 1 km by 1 km grids. The other possible explanation is that model training might be insufficient when PM<sub>2.5</sub> levels are significant high (mostly in fall and winter) or low (mostly in spring and summer); therefore, the models tend to underestimate PM<sub>2.5</sub> levels in fall and winter and overestimate in spring and summer.

The performance of our gap-filling approach is comparable with previous gap-filling models. The two-stage models from the study in YRD of China utilized gap-filled AOD by multiple imputation (MI) and had cross validation R<sup>2</sup> values of 0.81 and 0.73 for predicting PM<sub>2.5</sub> in 2013 and 2014, respectively (Xiao et al., 2017). A national study employing a machine learning algorithm based on satellite AOD and CMAQ AOD also achieved 100% spatiotemporal coverage in China with an R<sup>2</sup> of 0.61 (Xue et al., 2019a). Yet another study in New York State developed random forest models to predict missing AOD using meteorological parameters in the first step, and prediction of PM<sub>2.5</sub> with gap-filled AOD. The New York State example results showed a cross validation R<sup>2</sup> of 0.82 (Bi et al., 2019). To compare with previous research, we also explored multiple imputation methods (MI) and random forest to predict missing AOD, but the widespread and persistent AOD missingness in NEC prevented these methods from being effective. Previously reported MI or random forest models firstly built associations between AOD and meteorological parameters at grids and on days with valid AOD, then predicted missing AOD based on these associations. In NEC, where AOD retrieval failures happened in extensive areas for a prolonged period, the models built with data beyond these regions and periods fail to capture the true patterns of AOD and lead to inaccurate predictions of missing AOD values. The combined results in this study achieved high accuracy because many meteorological parameters, which were used to predict AOD in previous studies, were also included in our models to provide similar

information for predicting  $PM_{2.5}$ . Additionally, both the AOD model and non-AOD model were constrained by ground  $PM_{2.5}$  measurements.

The  $R^2$  (RMSE) of the 10-fold spatial CV was 0.78 ( $20.2 \mu\text{g}/\text{m}^3$ ) at the daily level, or 0.03 lower than the overall  $R^2$  of the combined predictions. The density plot of the spatial CV was shown in Fig. 4. The slope and intercept of the linear regression of  $PM_{2.5}$  measurements by  $PM_{2.5}$  predictions of spatial CV were 1.11 and  $-6.0$ , respectively, which were similar to the overall model performance. The spatial CV results indicated that the capacity of spatial interpolation of the combined predictions was stable, though moderate unobserved spatial trend was not captured by our model.

### 3.3. Leave-one-year-out-cross-validation

To further explore the role of AOD in predicting historical  $PM_{2.5}$  concentrations, leave-one-year-out-cross-validation was conducted separately for the AOD-model and non-AOD model. Here we assessed whether missing AOD would influence the capacity of models in predicting historical  $PM_{2.5}$  concentrations, which limited studies have done so far. We also compared  $PM_{2.5}$  measurements with  $PM_{2.5}$  predictions from the AOD model and non-AOD model only on grid cell-days with valid AOD values ( $N = 54,328$ ) to avoid the influence of the different sizes of the training datasets. In this instance, the  $R^2$  (RMSE) values were 0.50 ( $26.6 \mu\text{g}/\text{m}^3$ ) and 0.41 ( $28.9 \mu\text{g}/\text{m}^3$ ) at daily level for AOD-model and non-AOD model, respectively; and 0.62 ( $15.5 \mu\text{g}/\text{m}^3$ ) and 0.52 ( $17.3 \mu\text{g}/\text{m}^3$ ) at monthly level for the two models, respectively. These results suggested that AOD could help enhance the predicting accuracy significantly when predicting historical  $PM_{2.5}$  concentrations.

The overall  $R^2$  (RMSE) values of leave-one-year-out-cross-validation for the combined predictions were 0.43 ( $31.7 \mu\text{g}/\text{m}^3$ ) at the daily level and 0.65 ( $16.3 \mu\text{g}/\text{m}^3$ ) at the monthly level ( $N = 145,026$ ), indicating that the gap-filling approach had relatively high accuracy in predicting historical  $PM_{2.5}$  concentrations, especially at the monthly level (Fig. 5). The linear regression plot between measurements and by year temporal CV predictions of  $PM_{2.5}$  at the monthly level were highly consistent for most pairs of measurements and predictions. Fig. 5 illustrated that the combined predictions tended to underestimate  $PM_{2.5}$  concentrations when monthly mean measurements were higher than  $150 \mu\text{g}/\text{m}^3$ , and most of these occurrences were found in 2013. Additionally, leave-one-year-out-cross-validation  $R^2$  was 0.63 at the monthly level for the non-AOD model ( $N = 145,026$ ), which was 0.02 lower than  $R^2$  of combined predictions and suggested that integrating predictions from the AOD-model ( $N = 54,328$ ), while only accounts for 37% of the whole dataset, can help to stabilize the accuracy of historical  $PM_{2.5}$  concentration predictions. Since AOD is directly related to  $PM_{2.5}$  by containing information of aerosol loading, the high spatial resolution of AOD may help the model capture fine gradient of  $PM_{2.5}$  variability. Thus, we combined predictions from the AOD-model to help capture fine-scale  $PM_{2.5}$  variability and predictions from the non-AOD model to achieve full coverage.

Because few historical  $PM_{2.5}$  measurements were available, the results from the leave-one-year-out cross-validation could indicate the capacity of predicting historical  $PM_{2.5}$  levels. To better illustrate the validation results of leave-one-year-out cross-validation, we showed the out-of-sample results in 2013 in Fig. S3. As a scenario for predicting historical  $PM_{2.5}$

before the period of model training, random forest models were developed based on data from 2014 to 2016, and then were used to predict  $PM_{2.5}$  concentrations in 2013. The  $R^2$  (RMSE) values were 0.46 (42.3  $\mu\text{g}/\text{m}^3$ ) and 0.76 (23.9  $\mu\text{g}/\text{m}^3$ ) at the daily and monthly levels, respectively; with the estimation error mainly caused by the underestimation at higher  $PM_{2.5}$  levels. A previous study reported that by-year CV  $R^2$  of a random forest model at 10-km resolution in Northeastern China was 0.72 at monthly level in 2013–2017 with gap-filled predictions (Xiao et al., 2018). This difference may be attributed to differences in spatial resolution, study domain and study period. Another study developed a spatiotemporal random forest model in 2015, which was then used to predict  $PM_{2.5}$  concentrations in mainland China in 2016 (Wei et al., 2019). The  $R^2$  (RMSE) in 2016 were 0.55 (27.38  $\mu\text{g}/\text{m}^3$ ) and 0.73 (14.88  $\mu\text{g}/\text{m}^3$ ) at the daily and monthly level, respectively (Wei et al., 2019). Using similar methods, our model using data from 2013 to 2015 to predict  $PM_{2.5}$  in 2016; the leave-one-year-out cross-validation  $R^2$  (RMSE) in 2016 is 0.71 (13.3  $\mu\text{g}/\text{m}^3$ ) at the monthly level and was comparable with previous results. Unfortunately, the absence of  $PM_{2.5}$  measurements before 2013 prevents an independent assessment of the accuracy of predicted historical  $PM_{2.5}$  levels.

Another source of uncertainty regarding the validity of predicting historical  $PM_{2.5}$  is the potential differences resulting from the use of different meteorological data sources, either from GEOS-FP or GEOS-5. To address this issue, we compared daily meteorological parameters from GEOS-5 with those from GEOS-FP during the overlap period of Jan – May 2013 at grid level, with  $R^2$  values were 0.93, 0.83, 0.89, 0.90, 0.87 and 0.66 for temperature, humidity, east-west and north-south components of the wind vector, wind speed and PBLH, respectively. This comparison showed a high level of consistency between parameters from the two products. In a previous study, model performance and predictions produced by GEOS-5 and GEOS-FP during the overlap period were compared (Ma et al., 2016). The comparison showed that CV-estimated  $PM_{2.5}$  concentrations using GEOS-5 data were almost the same as the CV estimations using the GEOS-FP ( $R^2 = 0.99$ ) based on a modeling dataset; and the mean  $R^2$  value, slope, and intercept between final predictions from the two meteorological products were 0.92, 0.98, and  $-2.12$ , respectively, indicating consistency between historical estimations using GEOS-5 and GEOS-5 FP data. However, since the spatial resolutions and detailed algorithms are not fully consistent for GEOS-5 and GEOS-FP, further assessment of the performance of long-term meteorological products in predicting  $PM_{2.5}$  is warranted.

#### 3.4. Temporal and spatial patterns of predictions of $PM_{2.5}$

Predictions of  $PM_{2.5}$  daily concentrations were produced based on the gap-filling approach. Time series trends of predicted  $PM_{2.5}$  at the annual level for Liaoning (LN), Jilin (JL), Heilongjiang (HLJ) and all 3 provinces were summarized in Fig. 6; while the spatial distribution of annual mean  $PM_{2.5}$  concentrations in 2013 was shown in Fig. 7 as an example. Fig. 6 showed that, during our study period,  $PM_{2.5}$  concentrations began to increase from 2005 and reached an initial peak in 2008 or 2009.  $PM_{2.5}$  concentrations then fluctuated before arriving at the highest levels around 2013–2015, with dramatic drops seen in 2016. Even so, the annual mean in the whole NEC in 2016 was still higher than the annual Ambient Air Quality Standard limit of  $PM_{2.5}$  (35  $\mu\text{g}/\text{m}^3$ ) in China (GB 3095–2012).

The temporal pattern was similar to previous estimations of PM<sub>2.5</sub> in the Changchun city cluster, the capital city of Jilin Province (Xue et al., 2019b). The long-term trends of PM<sub>2.5</sub> indicated a lag effect in air pollution control policies in NEC. For example, the Chinese government launched the strictest air quality control policy in 2013 called the Air Pollution Prevention and Control Action Plan. With this plan, PM<sub>2.5</sub> levels in key regions such as Beijing-Tijianjin-Hebei (BTH) region, Yangtze River Delta (YRD) and Pearl River Delta (PRD) started to decline in 2014 (Geng et al., 2019; Ma et al., 2019). However, based on our predictions, the reduction of PM<sub>2.5</sub> in NEC started in 2016. Previous studies have shown that the investment in air pollution control per unit of SO<sub>2</sub> and smoke-dust emissions (used as indicators of the intensity of environmental regulation), climbed steeply in 2013 and remained high in 2014. However, subsequent drops in PM<sub>2.5</sub> concentrations were not evident until 2016 (Xue et al., 2019b), which was consistent with the trend in this study.

Considering the long heating season from mid-October to early April in NEC, we further evaluated the long-term trends during the non-heating season (April to September) and heating season (October to March). Results from this evaluation were shown in Fig. S4. While PM<sub>2.5</sub> level in the heating season have increased since 2007 and reached peak levels in 2014, it has remained relatively stable during the non-heating season from 2009 to 2014, with decreases evident in 2015. A comparison of Fig. 6 with Fig. S3 suggested that the upward trend of annual mean PM<sub>2.5</sub> since 2010 is likely dominated by rising PM<sub>2.5</sub> levels during the heating season.

Overall, PM<sub>2.5</sub> levels exhibit a decreasing trend from south to north in the NEC with high PM<sub>2.5</sub> levels found in the city clusters surrounding provincial capitals. The NEC is subject to emissions from various industrial, farming and human sources. The average annual mean concentration of PM<sub>2.5</sub> during 2005–2016 was 48.8 µg/m<sup>3</sup> in Liaoning Province, which was 13% and 26% higher than those in Jilin and Heilongjiang Provinces, respectively. The high PM<sub>2.5</sub> levels in Liaoning Province could be explained by high emissions from these sources. For example, the emissions of SO<sub>2</sub>, NO<sub>x</sub> and PM<sub>2.5</sub> were all much higher in Liaoning than Jilin and Heilongjiang Provinces (Zhao et al., 2013). Besides illustrating the spatial variability of PM<sub>2.5</sub> between cities, our predictions also captured finer gradients within the cities. The two figures in the left panel of Fig. 7 showed comparisons between predicted PM<sub>2.5</sub> at 1 km spatial resolution (upper) and a satellite image in the corresponding area of Shenyang City center (lower), indicating that the predictions could capture the lower PM<sub>2.5</sub> concentrations in grid cells associated with parks (red boxes in left lower map) at 1 km resolution.

#### 4. Conclusions

As far as we know, this is the first regional model to estimate the long-term PM<sub>2.5</sub> concentrations at a high spatial resolution and full coverage in high-latitude areas of China. The gap-filling approach reduces the missingness of PM<sub>2.5</sub> predictions and helps to correct the predictive bias caused by missing AOD that can happen in both extensive areas and prolonged time periods. These combined predictions may produce more accurate historical PM<sub>2.5</sub> concentrations with fine-scale inter-city and intra-city spatial variability. Finally, this study demonstrated that including AOD may help increase the accuracy of predicted

historical PM<sub>2.5</sub>, which few studies have reported. These full-coverage and high-accuracy predictions of PM<sub>2.5</sub> can support future evaluations of the effects of air quality control policies and facilitate epidemiological studies on the adverse health effects related to short-term and long-term exposures of PM<sub>2.5</sub> in NEC.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgment

This work was funded by the National Natural Science Foundation of China (91843302 and 91643205). The work of Y. LIU was supported by the MISR Science Team (contract No. 1363692), and the MAIA science team at the JPL, California Institute of Technology (contract No. 1588347), as well as the National Institute of Environmental Health Sciences (grant # 1R01ES032140). The authors thank Dr. Qingyang Xiao at Tsinghua University for the technical support.

## References

- Belle JH, Chang HH, Wang Y, Hu X, Lyapustin A, Liu Y, 2017. The potential impact of satellite-retrieved cloud parameters on ground-level PM<sub>2.5</sub> mass and composition. *Int. J. Environ. Res. Public Health*14, 1244.
- Bi J, Belle JH, Wang Y, Lyapustin AI, Wildani A, Liu Y, 2019. Impacts of snow and cloud covers on satellite-derived PM<sub>2.5</sub> levels. *Remote Sens. Environ*221, 665–674. [PubMed: 31359889]
- Breiman L, 2001. Random forests. *Mach. Learn*45, 5–32.
- Burnett RT, Pope CA, Ezzati M, Olives C, Lim SS, Mehta S, Shin HH, Singh G, Hubbell B, Brauer M, Anderson HR, Smith KR, Balmes JR, Bruce NG, Kan HD, Laden F, Pruss-Ustun A, Michelle CT, Gapstur SM, Diver WR, Cohen A, 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ. Health Perspect*122, 397–403. [PubMed: 24518036]
- Cheng Z, Luo L, Wang S, Wang Y, Sharma S, Shimadera H, Wang X, Bressi M, de Miranda RM, Jiang J, Zhou W, Fajardo O, Yan N, Hao J, 2016. Status and characteristics of ambient PM<sub>2.5</sub> pollution in global megacities. *Environ. Int*89–90, 212–221.
- Chu D, Kaufman Y, Ichoku C, Remer L, Tanré D, Holben B, 2002. Validation of MODIS aerosol optical depth retrieval over land. *Geophys. Res. Lett*29 (MOD2–1–MOD2–4).
- de Hoogh K, Héritier H, Stafoggia M, Künzli N, Kloog I, 2018. Modelling daily PM<sub>2.5</sub> concentrations at high spatio-temporal resolution across Switzerland. *Environ. Pollut*233, 1147–1154. [PubMed: 29037492]
- Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang YJ, Schwartz J, 2016. Assessing PM<sub>2.5</sub> exposures with high spatiotemporal resolution across the continental United States. *Environ. Sci. Technol*50, 4712–4721. [PubMed: 27023334]
- Donkelaar A.v., Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, Villeneuve PJ, 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect*118, 847–855. [PubMed: 20519161]
- Geng GN, Zhang Q, Martin RV, Lin JT, Huo H, Zheng B, Wang SW, He KB, 2017. Impact of spatial proxies on the representation of bottom-up emission inventories: a satellite-based analysis. *Atmos. Chem. Phys*17, 4131–4145.
- Geng G, Xiao Q, Zheng Y, Tong D, Zhang Y, Zhang X, Zhang Q, He K, Liu Y, 2019. Impact of China's air pollution prevention and control action plan on PM<sub>2.5</sub> chemical composition over eastern China. *Sci. China Earth Sci*1–13.
- Goldberg DL, Gupta P, Wang K, Jena C, Zhang Y, Lu Z, Streets DG, 2019. Using gap-filled MAIAC AOD and WRF-Chem to estimate daily PM<sub>2.5</sub> concentrations at 1 km resolution in the Eastern United States. *Atmos. Environ*199, 443–452.

- He QQ, Huang B, 2018. Satellite-based mapping of daily high-resolution ground PM<sub>2.5</sub> in China via space-time regression modeling. *Remote Sens. Environ*206, 72–83.
- Hsu N, Jeong MJ, Bettenhausen C, Sayer A, Hansell R, Seftor C, Huang J, Tsay SC, 2013. Enhanced deep blue aerosol retrieval algorithm: the second generation. *J. Geophys. Res*118, 9296–9315.
- Hu J, Chen J, Ying Q, Zhang HJAC, Physics, 2016a. One-year Simulation of Ozone and Particulate Matter in China Using WRF/CMAQ Modeling System, 16 (10333–10333).
- Hu JL, Chen JJ, Ying Q, Zhang HL, 2016b. One-year simulation of ozone and particulate matter in China using WRF/CMAQ modeling system. *Atmos. Chem. Phys*16, 10333–10350.
- Hu XF, Belle JH, Meng X, Wildani A, Waller LA, Strickland MJ, Liu Y, 2017. Estimating PM<sub>2.5</sub> concentrations in the conterminous United States using the random forest approach. *Environ. Sci. Technol*51, 6936–6944. [PubMed: 28534414]
- Kloog I, Nordio F, Coull BA, Schwartz J, 2012. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM<sub>2.5</sub> exposures in the mid-Atlantic States. *Environ. Sci. Technol*46, 11913–11921. [PubMed: 23013112]
- Li H, Lo K, Wang M, Zhang PY, Xue LY, 2016. Industrial energy consumption in Northeast China under the revitalisation strategy: a decomposition and policy analysis. *Energies*9.
- Li TW, Shen HF, Zeng C, Yuan QQ, Zhang LP, 2017. Point-surface fusion of station measurements and satellite observations for mapping PM<sub>2.5</sub> distribution in China: methods and assessment. *Atmos. Environ*152, 477–489.
- Li L, Franklin M, Girguis M, Lurmann F, Wu J, Pavlovic N, Breton C, Gilliland F, Habre R, 2020. Spatiotemporal imputation of MAIAC AOD using deep learning with downscaling. *Remote Sens. Environ*237, 111584. [PubMed: 32158056]
- Liaw A, 2015. Package ‘randomForest’.
- Lin CQ, Liu G, Lau AKH, Li Y, Li CC, Fung JCH, Lao XQ, 2018. High-resolution satellite remote sensing of provincial PM<sub>2.5</sub> trends in China from 2001 to 2015. *Atmos. Environ*180, 110–116.
- Liu MM, Bi J, Ma ZW, 2017. Visibility-based PM<sub>2.5</sub> concentrations in China: 1957–1964 and 1973–2014. *Environ. Sci. Technol*51, 13161–13169. [PubMed: 29063753]
- Liu C, Chen R, Sera F, Vicedo-Cabrera AM, Guo YM, Tong SL, Coelho MSZS, Saldiva PHN, Lavigne E, Matus P, Ortega NV, Garcia SO, Pascal M, Stafoggia M, Scortichini M, Hashizume M, Honda Y, Hurtado-Diaz M, Cruz J, Nunes B, Teixeira JP, Kim H, Tobias A, Iniguez C, Forsberg B, Astrom C, Ragettli MS, Guo YL, Chen BY, Bell ML, Wright CY, Scovronick N, Garland RM, Milojevic A, Kysely J, Urban A, Orru H, Indermitte E, Jaakkola JJK, Rytty NRI, Katsouyanni K, Analitis A, Zanobetti A, Schwartz J, Chen JM, Wu TC, Cohen A, Gasparini A, Kan HD, 2019. Ambient particulate air pollution and daily mortality in 652 cities. *N. Engl. J. Med*381, 705–715. [PubMed: 31433918]
- Lv B, Hu Y, Chang HH, Russell AG, Bai Y, 2016. Improving the accuracy of daily PM<sub>2.5</sub> distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China. *Environ. Sci. Technol*50, 4752–4759. [PubMed: 27043852]
- Lyapustin A, Martonchik J, Wang Y, Laszlo I, Korkin S, 2011a. Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *J. Geophys. Res*116.
- Lyapustin A, Wang Y, Laszlo I, Kahn R, Korkin S, Remer L, Levy R, Reid JS, 2011b. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *J. Geophys. Res*116.
- Ma ZW, Hu XF, Huang L, Bi J, Liu Y, 2014. Estimating ground-level PM<sub>2.5</sub> in China using satellite remote sensing. *Environ. Sci. Technol*48, 7436–7444. [PubMed: 24901806]
- Ma ZW, Hu XF, Sayer AM, Levy R, Zhang Q, Xue YG, Tong SL, Bi J, Huang L, Liu Y, 2016. Satellite-based spatiotemporal trends in PM<sub>2.5</sub> concentrations: China, 2004–2013. *Environ. Health Perspect*124, 184–192. [PubMed: 26220256]
- Ma Z, Liu R, Liu Y, Bi J, 2019. Effects of air pollution control policies on PM<sub>2.5</sub> pollution improvement in China from 2005 to 2017: a satellite-based perspective. *Atmos. Chem. Phys*19, 6861–6877.
- Meng X, Fu Q, Ma Z, Chen L, Zou B, Zhang Y, Xue W, Wang J, Wang D, Kan H, Liu Y, 2016. Estimating ground-level PM<sub>10</sub> in a Chinese city by combining satellite data, meteorological information and a land use regression model. *Environ. Pollut*208, 177–184. [PubMed: 26499934]

- Meng X, Hand JL, Schichtel BA, Liu Y, 2018a. Space-time trends of PM<sub>2.5</sub> constituents in the conterminous United States estimated by a machine learning approach, 2005–2015. *Environ. Int*121, 1137–1147. [PubMed: 30413295]
- Meng X, Hand JL, Schichtel BA, Liu Y, 2018b. Space-time trends of PM<sub>2.5</sub> constituents in the conterminous United States estimated by a machine learning approach, 2005–2015. *Environ. Int*121, 1137–1147. [PubMed: 30413295]
- Nachman RM, Mao GY, Zhang XY, Hong XM, Chen Z, Soria CS, He H, Wang GY, Caruso D, Pearson C, Biswal S, Zuckerman B, Wills-Karp M, Wang XB, 2016. Intrauterine inflammation and maternal exposure to ambient PM<sub>2.5</sub> during preconception and specific periods of pregnancy: the Boston birth cohort. *Environ. Health Perspect.* 124, 1608–1615. [PubMed: 27120296]
- Paciorek CJ, Liu Y, 2009. Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. *Environ. Health Perspect*117, 904–909. [PubMed: 19590681]
- Pope CA III, Coleman N, Pond ZA, Burnett RT, 2019. Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ. Res*183, 108924. [PubMed: 31831155]
- Power MC, Kioumourtoglou MA, Hart JE, Okereke OI, Laden F, Weisskopf MG, 2015. The relation between past exposure to fine particulate air pollution and prevalent anxiety: observational cohort study. *Br. Med. J*350.
- Randles CA, da Silva AM, Buchard V, Colarco PR, Darmenov A, Govindaraju R, Smirnov A, Holben B, Ferrare R, Hair J, Shinozuka Y, Flynn CJ, 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: system description and data assimilation evaluation. *J. Clim*30, 6823–6850. [PubMed: 29576684]
- Stafoggia M, Bellander T, Bucci S, Davoli M, de Hoogh K, De' Donato F, Gariazzo C, Lyapustin A, Michelozzi P, Renzi M, Scortichini M, Shtein A, Viegi G, Kloog I, Schwartz J, 2019. Estimation of daily PM<sub>10</sub> and PM<sub>2.5</sub> concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ. Int*124, 170–179. [PubMed: 30654325]
- Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, Abdollahpour I, Abdulkader RS, Abebe M, Abebe Z, Abera SF, Abil OZ, Abraha HN, Abrham AR, Abu-Raddad LJ, Abu-Rmeileh NME, Accrombessi MMK, Acharya D, Acharya P, Adamu AA, Adane AA, Adebayo OM, Adedoyin RA, Adekanmbi V, Ademi Z, Adetokunboh O, Adib MG, Admasie A, Adsuar JC, Afanvi KA, Afarideh M, Agarwal G, Aggarwal A, Aghayan SA, Agrawal A, Agrawal S, Ahmadi A, Ahmadi M, Ahmadi H, Ahmed MB, Aichour AN, Aichour I, Aichour MTE, Akbari ME, Akinemiju T, Akseer N, Al-Aly Z, Al-Eyadhy A, Al-Mekhlafi HM, Alandab F, Alam K, Alam S, Alam T, Alashi A, Alavian SM, Alene KA, Ali K, Ali SM, Alijanzadeh M, Alizadeh-Navaei R, Aljunid SM, Alkerwi A, Alla F, Alsharif U, Altirkawi K, Alvis-Guzman N, Amare AT, Ammar W, Anber NH, Anderson JA, Andrei CL, Androudi S, Anmut MD, Anjomshoa M, Ansha MG, Anto JM, Antonio CAT, Anwari P, Appiah LT, Appiah SCY, Arabloo J, Aremu O, Amlöv J, Artaman A, Aryal KK, Asayesh H, Ataro Z, Ausloos M, Avokpaho EFGA, Awasthi A, Quintanilla BPA, Ayer R, Ayuk TB, Azzopardi PS, Babazadeff A, Badali H, Badawi A, Balakrishnan K, Bali AG, Ball K, Bellew SH, Banach M, Banoub JAM, Barac A, Barker-Collo SL, Bamighausen TW, Barrero LH, Basu S, Baune BT, Bazargan-Hejazi S, Bedi N, Beghi E, Behzadifar M, Behzadifar M, Bejoy Y, Bekele BB, Bekru FT, Belay E, Belay YA, Bell ML, Bello AK, Bennett DA, Bensenor IM, Bergeron G, Berhane A, Bemabe E, Bemstein RS, Beuran M, Beyranvand T, Bhala N, Bhalla A, Bhattarai S, Bhutta ZA, Biadgo B, Bijani A, Bikbov B, Bilano V, Bililign N, Bin Sayeed MS, Bisanzio D, Biswas T, Bjorge T, Blacker BF, Bleyer A, Borschmann R, Bou-Orm IR, Boufous S, Bourne R, Brady OJ, Brauer M, Brazinova A, Breitborde NJK, Brenner H, Briko AN, Britton G, Brugha T, Buchbinder R, Burnett RT, Busse R, Butt ZA, Cahill LE, Cahuana-Hurtado L, Campos-Nonato IR, Cardenas R, Carreras G, Carrero JJ, Carvalho F, Castaneda-Orjuela CA, Rivas JC, Castro F, Catala-Lopez F, Causey K, Cercy KM, Cerin E, Chaiah Y, Chang HT, Chang JC, Chang KL, Charlson FJ, Chattopadhyay A, Chattu VK, Chee ML, Cheng CY, Chew A, Chiang PPC, Chimed-Ochir O, Chin KL, Chitheer A, Choi JYJ, Chowdhury R, Christensen H, Christopher DJ, Chung SC, Cicuttini FM, Grillo M, Cohen AJ, Collado-Mateo D, Cooper C, Cooper OR, Coresh J, Cornaby L, Cortesi PA, Cortinovis M, Costa M, Cousin E, Criqui MH, Cromwell EA, Cundiff DK, Daba AK, Dachew BA, Dadi AF, Damasceno AAM, Dandona L, Dandona R, Darby SC, Dargan PI, Daryani A, Das Gupta R, Das Neves J, Dasa TT, Dash AP, Davitoiu DV, Davletov K, De la Cruz-Gongora V, De la Hoz FP, De Leo D, De Neve JW, Degenhardt L, Deiparine

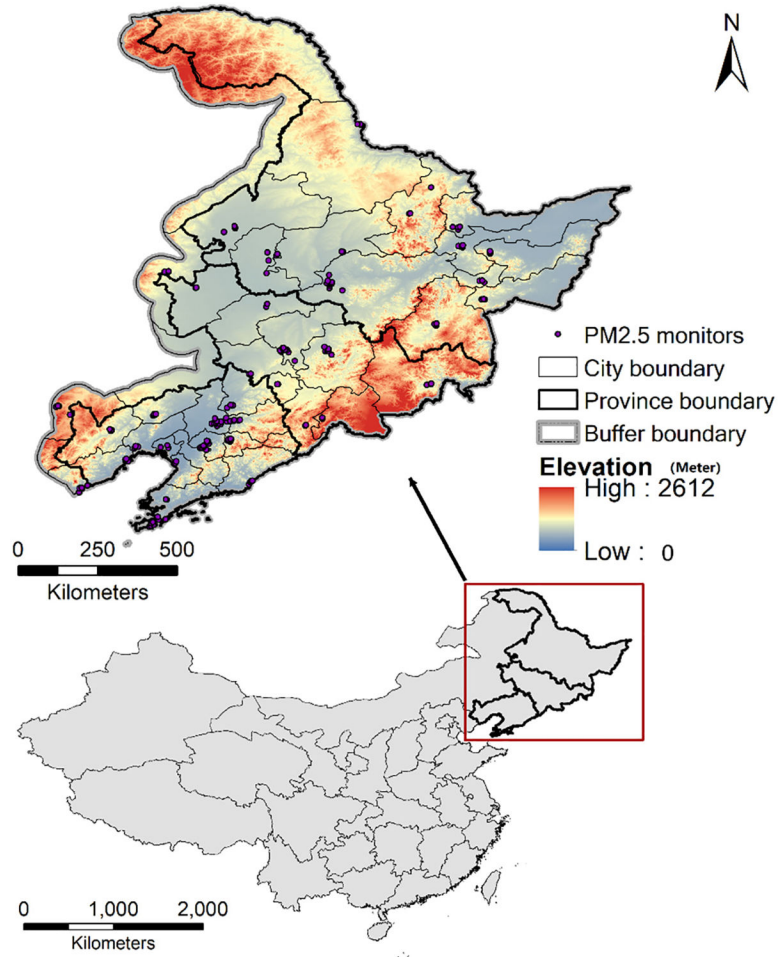
S, Dellavalle RP, Demoz GT, Denova-Gutierrez E, Deribe K, Dervenis N, Deshpande A, Jarlais DCD, Dessie GA, Deveber GA, Dey S, Dharmaratne SD, Dhimal M, Dinberu MT, Ding EL, Diro HD, Djalalinia S, Do HP, Dokova K, Doku DT, La Doyle K, Driscoll TR, Dubey M, Dubljanin E, Duken EE, Duncan BB, Duraes AR, Ebert N, Ebrahimi H, Ebrahimpour S, Edvardsson D, Elfiog A, Eggen AE, El Bcheraoui C, El-Khatib I, Elyazar IR, Enayati A, Endries AY, Er B, Erskine HE, Eskandarieh S, Esteghamati A, Estep K, Fakhim H, Faramarzi M, Fareed M, Farid TA, Farinha CSES, Farioli A, Faro A, Farvid MS, Farzaei MH, Fatima B, Fay KA, Fazaeli AA, Feigin VL, Feigl AB, Fereshtehnejad SM, Fernandes E, Fernandes JC, Ferrara G, Ferrari AJ, Ferreira ML, Filip I, Finger JD, Fischer F, Foigt NA, Foreman KJ, Fukumoto T, Fullman N, Furst T, Furtado JM, Futran ND, Gall S, Gallus S, Gamkrelidze A, Ganji M, Garcia-Basteiro AL, Gardner WN, Gebre AK, Gebremedhin AT, Gebremichael TG, Gelano TF, Geleijnse JM, Geramo YCD, Gething PW, Geese KE, Ghadimi R, Ghadiri K, Falavarjani KG, Ghasemi-Kasman M, Ghimire M, Ghosh R, Ghoshal AG, Giampaoli S, Gill PS, Gill TK, Gillum RF, Ginawi IA, Giussani G, Gnedovskaya EV, Godwin WW, Goli S, Gomez-Dantes H, Gona PN, Gopalani SV, Goulart AC, Grada A, Grams ME, Grosso G, Gugnani HC, Guo Y, Gupta R, Gupta R, Gupta T, Gutierrez RA, Gutierrez-Torres DS, Haagsma JA, Habtewold TD, Hachinski V, Hafezi-Nejad N, Hagos TB, Hailegiyorgis TT, Hailu GB, Haj-Mirzaian A, Haj-Mirzaian A, Hamadeh RR, Hamidi S, Handal AJ, Hankey GJ, Liao Y, Harb HL, Harikrishnan S, Haro JM, Hassankhani H, Hassen HY, Havmoeller R, Hawley CN, Hay SI, Hedayatizadeh-Omran A, Heibati B, Heidari B, Heidari M, Hendrie D, Henok A, Heredia-Pi I, Herteliu C, Heydarpour F, Heydarpour S, Hibstu DT, Higazi TB, Hilawe EH, Hoek HW, Hoffiman HJ, Hole MK, Rad EH, Hoogar P, Hosgood HD, Hosseini SM, Hosseinzadeh M, Hostiuc M, Hostiuc S, Hoy DG, Hsairi M, Hsiao T, Hu G, Ha H, Huang JJ, Hussen MA, Huynh CK, Iburg KM, Ikeda N, Ilesanmi OS, Iqbal U, Irvani SSN, Irvine CMS, Islam SMS, Islami F, Jackson MD, Jacobsen KH, Jahangiry L, Jahanmehr N, Jain SK, Jakovljevic M, James SL, Jassal SK, Jayatilleke AU, Jeemon P, Jha RP, Jha V, Ji JS, Jonas JB, Jonnagaddala J, Shushtari ZJ, Joshi A, Jozwiak JJ, Jurisson M, Kabir Z, Kahsay A, Kalani R, Kanchan T, Kant S, Kar C, Karami M, Matin BK, Karch A, Karema C, Karimi N, Karimi SM, Kasaeian A, Kassa DH, Kassa GM, Kassa TD, Kassebaum NJ, Katikireddi SV, Kaul A, Kawakami N, Kazemi Z, Karyana AK, Kefale AT, Keiyoro PN, Kemp GR, Kengne AP, Keren A, Kesavachandran CN, Khader YS, Khafaei B, Khafaie MA, Khajavi A, Khalid N, Khalil IA, Khan G, Khan MS, Khan MA, Khang YH, Khater MM, Khazaei M, Khazaie H, Khoja AT, Khosravi A, Khosravi MH, Kiadaliri AA, Kiirithio DN, Kim CI, Kim D, Kim YE, Kim YJ, Kimokoti RW, Kinfh Y, Kisa A, Kissimova-Skarbek K, Kivimaki M, Knibbs LD, Knudsen AKS, Kochhar S, Kokubo Y, Kolola T, Kopec JA, Kosen S, Koul PA, Koyanagi A, Kravchenko MA, Krishan K, Krohn KJ, Kromhout H, Defo BK, Bicer BK, Kumar GA, Kumar M, Kuzin I, Kyu HH, Lachat C, Lad DP, Lad SD, Lafranconi A, Laloo R, Lallukka T, Lami FH, Lang JJ, Lansingh VC, Larson SL, Latifi A, Lazarus JV, Lee PH, Leigh J, Leili M, Leshargie CT, Leung J, Levi M, Lewycka S, Li S, Li Y, Liang J, Liang X, Lian Y, Liben ML, Lim LL, Linn S, Liu S, Lodha R, Logroscino G, Lopez AD, Lorkowski S, Lotufo PA, Lozano R, Lucas TCD, Lunevicius R, Ma S, Macarayan ERK, Machado IE, Madotto F, Mai HT, Majdan M, Majdzadeh R, Majeed A, Malekzadeh R, Malta DC, Mamum AA, Manda AL, Manguerra H, Mansournia MA, Mantovani LG, Maravilla JC, Marcenes W, Marks A, Martin RV, Martins SC, Martins-Melo FR, Marz W, Marzan MB, Massenburg BB, Mathur MR, Mathur P, Matsushita K, Maulik PK, Mazidi M, McAlinden C, McGrath JJ, Mckee M, Mehrotra R, Mehta KM, Mehta V, Meier T, Mekonnen FA, Melaku YA, Melese A, Melku M, Memiah PTN, Memish ZA, Mendoza W, Mengistu DT, Mensah GA, Mensink GBM, Mereta ST, Meretola A, Meretoja TJ, Mestrovic T, Mezgebe HB, Miazgowski B, Miazgowski T, Milllear AI, Miller TR, Miller-Petrie MK, Mini GK, Mirarefin M, Mirica A, Mirrahimov EM, Misganaw AT, Mitiku H, Moazen B, Mohajer B, Mohammad KA, Mohammadi M, Mohammadifard N, Mohammadnia-Afrouzi M, Mohammed S, Mohebi F, Mokdad AH, Molokhia M, Momeniha F, Monasta L, Moodley Y, Moradi G, Moradi-Lakeh M, Moradinazar M, Moraga P, Morawska L, Morgado-Da-Costa J, Morrison SD, Moschos MM, Mouodi S, Mousavi SM, Mozaffarian D, Mruts KB, Muche AA, Muchie KF, Mueller UO, Muhammed OS, Mukhopadhyay S, Muller K, Musa KI, Mustafa G, Nabhan AF, Naghavi M, Naheed A, Nahvijou A, Naik G, Naik N, Najafi F, Nangia V, Nansseu JR, Nascimento BR, Neal B, Neamati N, Negoit IT, Negoit RI, Neupane S, Newton CRJ, Ngunjiri JV, Nguyen AQ, Nguyen G, Nguyen HT, Nguyen HLT, Nguyen HT, Nguyen M, Nguyen NB, Nichols E, Nie J, Ningrum DNA, Nirayo YL, Nishi N, Nixon MR, Nojomi M, Nomura S, Norheim OF, Noroozi M, Norrving B, Noubiap JJ, Nouri HR, Shiadeh MN,



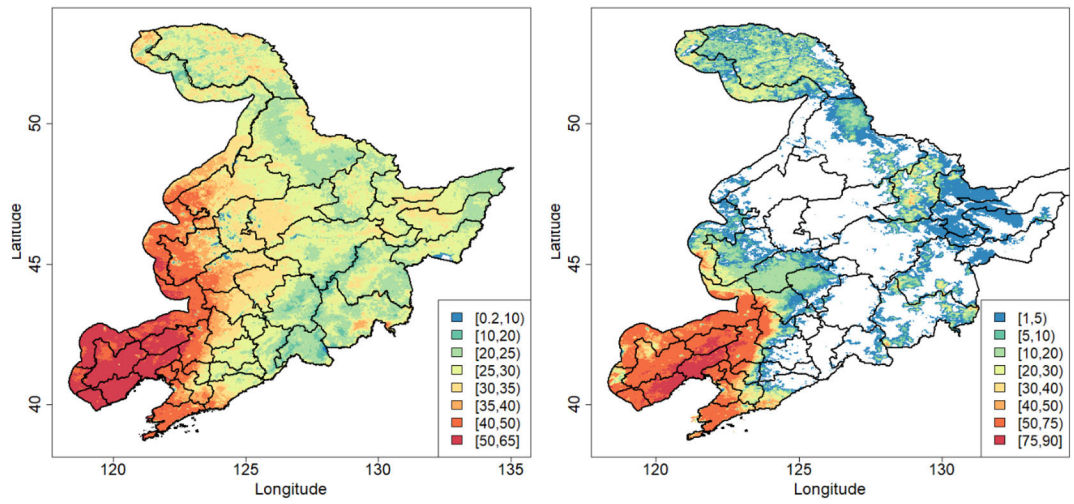
Nowroozi MR, Nsoesie EO, Nyasulu PS, Obermeyer CM, Odell CM, Ofori-Asenso R, Ogbo FA, Oh IH, Oladimeji O, Olagunju AT, Olagunju TO, Olivares PR, Olsen HE, Olusanya BO, Olusanya JO, Ong KL, Ong SK, Oren E, Orpana HM, Ortiz A, Ota E, Otstavnov SS, Overland S, Owolabi MO, Pacella MPA, Pakhare AP, Pakpour AH, Pana A, Panda-Jonas S, Park EK, Parry CDH, Parisian H, Patel S, Pati S, Patil ST, Patle A, Patton GC, Paudel D, Paulson KR, Ballesteros WCP, Pearce N, Pereira A, Pereira DM, Perico N, Pesudovs K, Petzold M, Pham HQ, Phillips MR, Pillay JD, Piradov MA, Pirsaeheb M, Pischon T, Pishgar F, Plana-Ripoll O, Plass D, Polinder S, Polkinghorne KR, Postma MJ, Poulton R, Pourshams A, Poustchi H, Prabhakaran D, Prakash S, Prasad N, Purcell CA, Purwar MB, Qorbani M, Radfar A, Rafay A, Rafiei A, Rafiei A, Rahimi Z, Rahimi-Movaghar A, Rahimi-Movaghar V, Rahman M, Rahman MHU, Rahman MA, Rai RK, Rajati F, Rajsic S, Raju SB, Ram U, Ranabhat CL, Ranjan P, Rath GK, Rawaf DL, Rawaf S, Reddy KS, Rehm CD, Rehm J, Reiner RC, Reitsma MB, Remuzzi G, Renzaho AMN, Resnikoff S, Reynales-Shigematsu LM, Rezaei S, Ribeiro ALP, Rivera JA, Roba KT, Rodrigues-Ramirez S, Roever L, Roman Y, Ronfani L, Roshandel G, Rostami A, Roth GA, Rothenbacher D, Roy A, Rubagotti E, Rushton L, Sabanayagam C, Sachdev PS, Saddik B, Sadeghi E, Moghaddam SS, Safari H, Safari Y, Safari-Faramani R, Safdarian M, Safi S, Safiri S, Sagar R, Sahebkar A, Sahraian MA, Sajadi HS, Salam N, Salamati P, Saleem Z, Salimi Y, Salimeadeh H, Salomon JA, Salvi DD, Satz I, Samy AM, Sanabria J, Sanchez-Nino MD, Sanchez-Pimienta TG, Sanders T, Sang Y, Santomauro DF, Santos IS, Santos JV, Milicevic MMS, Jose BPS, Sardana M, Sacker AR, Sarmiento-Suarez R, Sarrafzadegan N, Sartorius B, Sarvi S, Sathian B, Satpathy M, Sawant AR, Sawhney M, Saylan M, Sayyah M, Schaeffner E, Schmidt MI, Schneider IJC, Schottker B, Schutte AE, Schwehel DC, Schwendicke F, Scott JG, Seedat S, Sekerija M, Sepanlou SG, Serre ML, Servan-Mori E, Seyedmousavi S, Shabaninejad H, Shaddick G, Shafieesabet A, Shahbazi M, Shaheen AA, Shaikh MA, Levy TS, Shams-Beyranvand M, Shamsi M, Sharafi H, Sharafi K, Sharif M, Sharif-Alhoseini M, Sharifi H, Sharma J, Sharma M, Sharma R, She J, Sheikh A, Shi P, Shibuya K, Shiferaw MS, Shigematsu M, Shin MJ, Shirt R, Shirkoobi R, Shiue I, Shokraneh F, Shoman H, Shrimel MG, Shupler MS, Si S, Siabani S, Sibai AM, Siddiqi TJ, Sigfusdottir ID, Sigurvinsdottir R, Silva DAS, Silva JP, Silveira DGA, Singh JA, Singh NP, Singh V, Sinha DN, Skiadaresi E, Skirbekk V, Smith DL, Smith M, Sobaih BH, Sobhani S, Somayaji R, Soofi M, Sorensen RJD, Soriano JB, Soyiri IN, Spinelli A, Sposato LA, Sreeramareddy CT, Srinivasan V, Starodubow VI, Steckling N, Stein DJ, Stein MB, Stevanovic G, Stockfelt L, Stokes MA, Sturua L, Subart ML, Sudaryanto A, Sufiyanto MB, Sulo G, Sunguya BF, Sur PJ, Sykes BI, Szoeki CEI, Tabares-Seisdedos R, Tabuchi T, Tadakamadla SK, Takahashi K, Tandon N, Tassew SG, Tavakkoli M, Taveira N, Tehrani-Banihashemi A, Tekalign TG, Tekelemedhin SW, Tekle MG, Temesgen H, Temsah MR, Temsah O, Terkawi AS, Tessema B, Teweldemedhin M, Thankappan KR, Theis A, Thirunavukkarasu S, Thomas HJ, Thomas ML, Thomas N, Thurston GD, Tilahun B, Tillmana T, To QG, Tobollik M, Tonelli M, Topor-Madry R, Torre AE, Tortajada-Girbes M, Touvier M, Tovani-Palone MR, Towbin JA, Tran BX, Tran KB, Truelsen TC, Truong NT, Tsadik AG, Car LT, Tuzcu EM, Tymeson HD, Tyrovolas S, Ukwaja KN, Ullah I, Updike RL, Usman MS, Uthman OA, Vaduganathan M, Vaezi A, Valdez PR, Van Donkelaar A, Varavikova E, Varughese S, Vasankari TJ, Venkateswaran V, Venketasubramanian N, Villafaina S, Violante FS, Vladimirov SK, Vlassov V, Vollset SE, Vos T, Vsboughi K, Vu GT, Vujcic IS, Wagnew FS, Waheed Y, Waller SG, Walson JL, Wang Y, Wang Y, Wang YP, Weiderpass E, Weintraub RG, Weldegebreal F, Werdecker A, Weiknah AA, West JJ, Westerman R, Whiteford HA, Widecka J, Wijeratne T, Winkler AS, Wiyeh AB, Wiysonge CS, Wolfe CDA, Wong TY, Wu S, Xavier D, Xu G, Yadgir S, Yadollahpour A, Jabbari SHY, Yamada T, Yan LL, Yano Y, Yaseri M, Yasin YJ, Yeshanely A, Yimer EM, Yip P, Yisma E, Yonemoto N, Yoon SJ, Yotebieng M, Younis MZ, Yousefifard M, Yu C, Zaidi Z, Bin Zaman S, Zamani M, Zavala-Arciniega L, Zhang AL, Zhang H, Zhang K, Thou M, Zimsen SM, Zodpey S, Murray CJL, 2018. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*392, 1923–1994. [PubMed: 30496105]

Wei J, Huang W, Li Z, Xue W, Peng Y, Sun L, Cribb M, 2019. Estimating 1-km-resolution PM<sub>2.5</sub> concentrations across China using the space-time random forest approach. *Remote Sens. Environ*231, 111221.

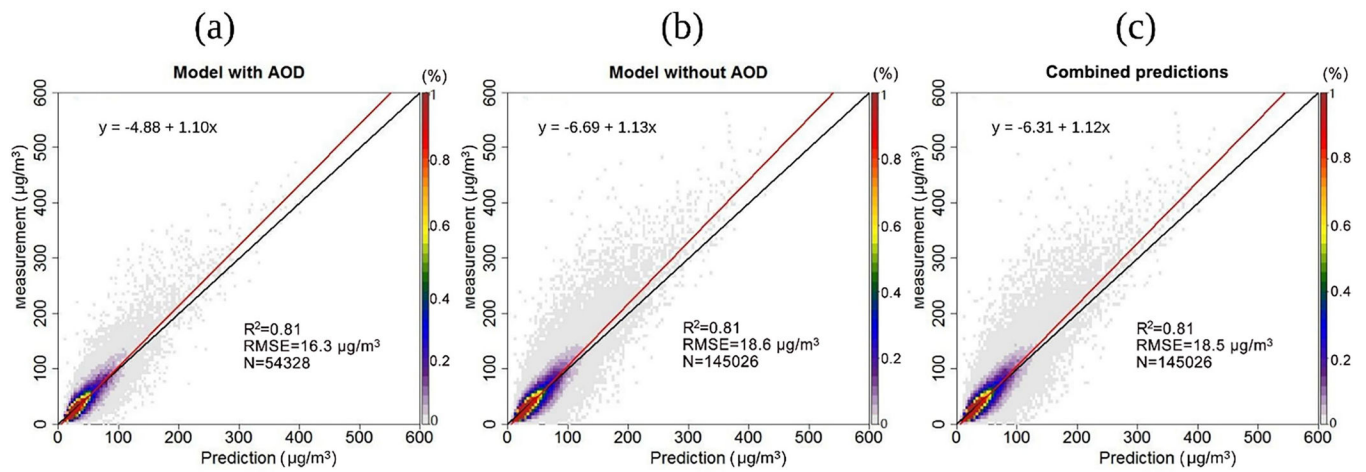
- Xiao QY, Wang YJ, Chang HH, Meng X, Geng GN, Lyapustin A, Liu Y, 2017. Full-coverage high-resolution daily PM<sub>2.5</sub> estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sens. Environ*199, 437–446.
- Xiao Q, Chang HH, Geng G, Liu Y, 2018. An ensemble machine-learning model to predict historical PM<sub>2.5</sub> concentrations in China from satellite data. *Environ. Sci. Technol*52, 13260–13269. [PubMed: 30354085]
- Xue T, Zheng Y, Tong D, Zheng B, Li X, Zhu T, Zhang Q, 2019a. Spatiotemporal continuous estimates of PM<sub>2.5</sub> concentrations in China, 2000–2016: a machine learning method with inputs from satellites, chemical transport model, and ground observations. *Environ. Int*123, 345–357. [PubMed: 30562706]
- Xue W, Zhang J, Zhong C, Ji D, Huang W, 2019b. Satellite-derived spatiotemporal PM<sub>2.5</sub> concentrations and variations from 2006 to 2017 in China. *Sci. Total Environ*134577. [PubMed: 31812394]
- Yang QQ, Yuan QQ, Yue LW, Li TW, Shen HF, Zhang LP, 2019. The relationships between PM<sub>2.5</sub> and aerosol optical depth (AOD) in mainland China: about and behind the spatio-temporal variations. *Environ. Pollut*248, 526–535. [PubMed: 30831349]
- Zhan Y, Luo Y, Deng X, Chen H, Grieneisen ML, Shen X, Zhu L, Zhang M, 2017. Spatiotemporal prediction of continuous daily PM<sub>2.5</sub> concentrations across China using a spatially explicit machine learning algorithm. *Atmos. Environ*155, 129–139.
- Zhao B, Wang S, Wang J, Fu JS, Liu T, Xu J, Fu X, Hao JJA, 2013. Impact of National NO<sub>x</sub> and SO<sub>2</sub> Control Policies on Particulate Matter Pollution in China, 77, pp. 453–463.
- Zhao C, Liu Z, Wang Q, Ban J, Chen NX, Li T, 2019a. High-resolution daily AOD estimated to full coverage using the random forest model approach in the Beijing-Tianjin-Hebei region. *Atmos. Environ*203, 70–78.
- Zhao Y, Feng D, Yu L, Cheng Y, Zhang M, Liu X, Xu Y, Fang L, Zhu Z, Gong PJRS, 2019b. Long-term Land Cover Dynamics (1986–2016) of Northeast China Derived From a Multi-temporal Landsat Archive, 11, p. 599.
- Zheng B, Zhang Q, Zhang Y, He KB, Wang K, Zheng GJ, Duan FK, Ma YL, Kimoto T, 2015. Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China. *Atmos. Chem. Phys*15, 2031–2049.
- Zheng YX, Zhang Q, Liu Y, Geng GN, He KB, 2016. Estimating ground-level PM<sub>2.5</sub> concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. *Atmos. Environ*124, 232–242.



**Fig. 1.** Locations of PM<sub>2.5</sub> measurement stations during 2013–2016 and elevation in the Northeast China (NEC).

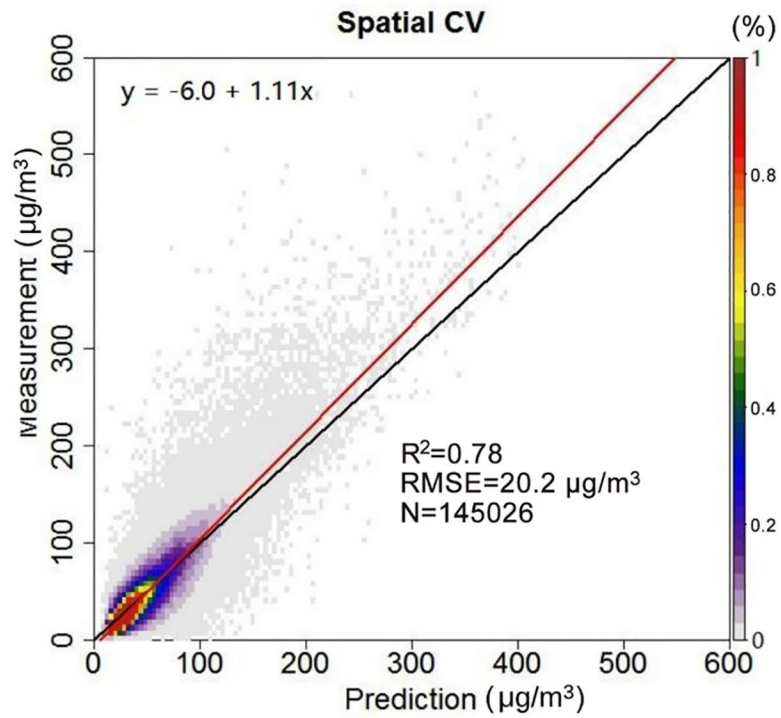


**Fig. 2.** The mean percentage of days with valid AOD value for grid cells at annual level in 2015 (left) and for grid cells in 2015 winter (right) in the Northeast China (NEC).

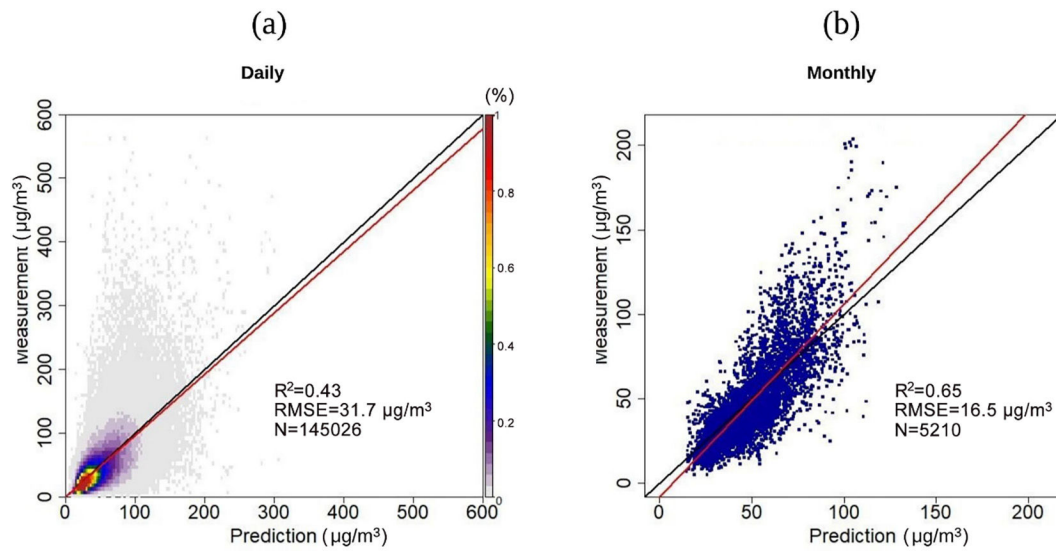


**Fig. 3.**

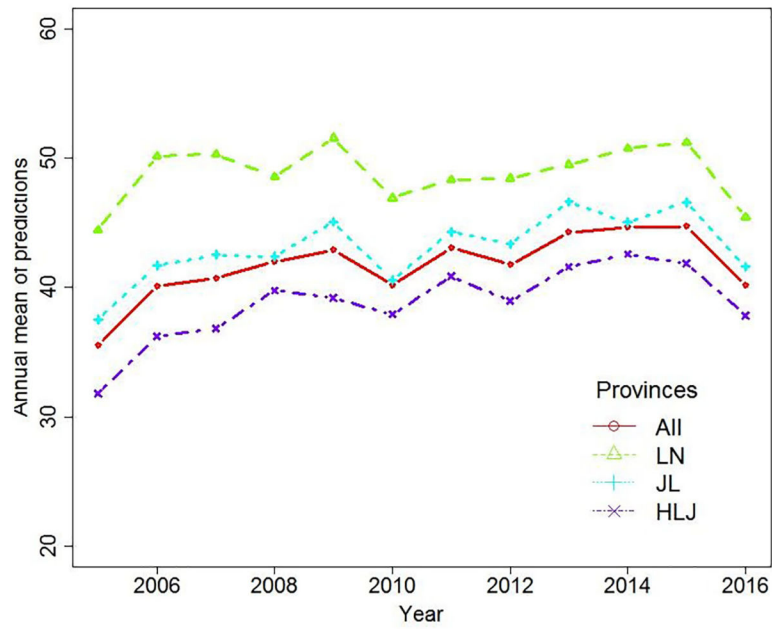
The density plots of 10-fold CV between measured and predicted  $PM_{2.5}$  concentrations for the AOD model (a), the non-AOD model (b) and combined predictions from both models (c) in 2013–2016. The red lines are the regression lines between  $PM_{2.5}$  measurements and  $PM_{2.5}$  predictions. The black lines are the 1:1 lines. The vertical colour bar represents the percentage of the total number of data points in each pixel. Higher percentage means more data points falling into the pixel of the density plot.



**Fig. 4.** The density plot of 10-fold spatial cross validation between measured and predicted PM<sub>2.5</sub> concentrations for the combined predictions. The red line is the regression line between PM<sub>2.5</sub> measurements and PM<sub>2.5</sub> predictions. The black line is the 1:1 line. The vertical colour bar represents the percentage of the total number of data points in each pixel. Higher percentage means more data points falling into the pixel of the density plot.

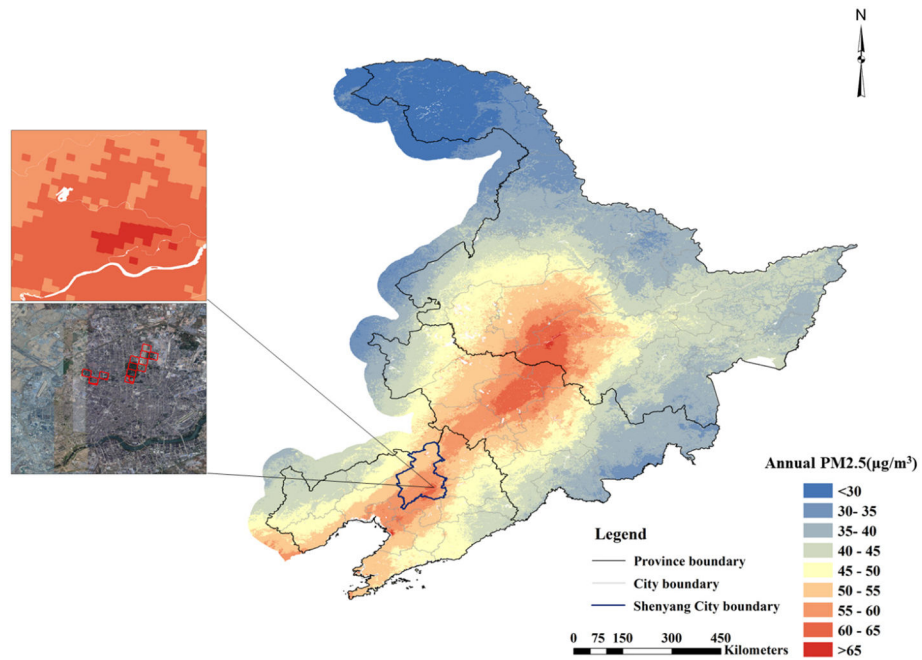


**Fig. 5.** Leave-one-year-out-cross-validation results at daily level (a) and at monthly level (b) for the combined predictions on all days during 2013–2016. The red lines are the regression lines between PM<sub>2.5</sub> measurements and PM<sub>2.5</sub> predictions. The black lines are the 1:1 lines. The vertical colour bar in (a) represents the percentage of the total number of data points in each pixel. Higher percentage means more data points falling into the pixel of the density plot.



**Fig. 6.** The long-term trends of predicted PM<sub>2.5</sub> concentrations for the Northeast China (All), Liaoning Province (LN), Jilin Province (JL) and Heilongjiang Province (HLJ) during 2005–2016. The unit is µg/m<sup>3</sup>.





**Fig. 7.** Annual mean predictions of PM<sub>2.5</sub> in Northeast China (NEC) in 2013. Left upper map shows predictions of PM<sub>2.5</sub> at 1 km spatial resolution and left lower map shows satellite image in the same area. The red boxes in left lower map were 1 km × 1 km grids corresponding to the grids with lower levels of PM<sub>2.5</sub> in the left upper map.

**Table 1**

The seasonal mean concentrations of PM<sub>2.5</sub> measurements, predictions with full coverage (combined) and predictions with missingness (AOD-model only), as well as the AOD missing rates by seasons in the modeling dataset during 2013–2016.

Seasons	No. of measurements	Missing rate*	Mean of measurements	Mean of predictions (combined)	Mean of predictions (AOD-model only)
Spring	36,145	56%	44.6 µg/m <sup>3</sup>	47.3 µg/m <sup>3</sup>	46.2 µg/m <sup>3</sup>
Summer	37,771	64%	32.5 µg/m <sup>3</sup>	32.8 µg/m <sup>3</sup>	32.0 µg/m <sup>3</sup>
Fall	37,004	51%	57.4 µg/m <sup>3</sup>	54.8 µg/m <sup>3</sup>	53.3 µg/m <sup>3</sup>
Winter	34,106	80%	80.8 µg/m <sup>3</sup>	78.7 µg/m <sup>3</sup>	57.9 µg/m <sup>3</sup>

\* Missing rate = (No. of all PM<sub>2.5</sub> observations - No. of PM<sub>2.5</sub> observations matched with valid AOD) / No. of all PM<sub>2.5</sub> observations.