

## A model framework to reduce bias in ground-level PM<sub>2.5</sub> concentrations inferred from satellite-retrieved AOD

Fei Yao <sup>a,\*</sup>, Paul I. Palmer <sup>a,b</sup>

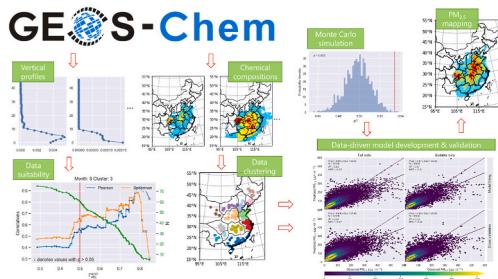
<sup>a</sup> School of GeoSciences, University of Edinburgh, UK

<sup>b</sup> National Centre for Earth Observation, University of Edinburgh, UK

### HIGHLIGHTS

- We propose a model framework to infer ground-level PM<sub>2.5</sub> from satellite AOD.
- The model framework considers the representativeness of AOD for ground-level PM<sub>2.5</sub>.
- The model framework reduces bias in ground-level PM<sub>2.5</sub> estimates by 9–15%.
- The model framework captures more variations in ground-level PM<sub>2.5</sub> by up to 8%.

### GRAPHICAL ABSTRACT



### ARTICLE INFO

#### Keywords:

PM<sub>2.5</sub>  
AOD  
Model of atmospheric chemistry and transport  
Statistical regression  
Machine learning

### ABSTRACT

We present a new method to infer ground-level fine particulate matter (PM<sub>2.5</sub>) from satellite remote sensing observations of aerosol optical depth (AOD). The conventional method generally uses a range of modelling approaches to determine PM<sub>2.5</sub>:AOD relationships that are subsequently used to infer ground-level PM<sub>2.5</sub> concentrations from satellite-retrieved AOD. Here, we use a high-resolution atmospheric chemistry simulation to explore how changes in the vertical distribution of aerosol extinction coefficients affects the PM<sub>2.5</sub>:AOD relationship and how we can use that information to improve the robustness of inferred estimates of ground-level PM<sub>2.5</sub> over eastern China. We define a metric,  $\Gamma_{PBL}^{AOD}$ , that describes the fraction of AOD that resides in the planetary boundary layer compared with the total columnar AOD. We determine physically-meaningful PM<sub>2.5</sub>:AOD relationships using data for which  $\Gamma_{PBL}^{AOD} \geq 50\%$ , a criterion based on sensitivity analyses on data clusters that we identify using a hierarchical clustering method. We use statistical and machine learning methods to develop independent models that describe these PM<sub>2.5</sub>:AOD relationships, and use a Monte Carlo approach to quantify the improvement after our selection of more physically relevant data records. Benefiting from the improved representativeness of AOD for ground-level PM<sub>2.5</sub>, our method effectively reduces bias in inferred estimates of ground-level PM<sub>2.5</sub> by 10–15% (9–12%) for space-borne sensors passing over in the morning (afternoon). It also captures more variations in ground-level PM<sub>2.5</sub> by up to 8% (5%) for space-borne sensors passing over in the morning (afternoon), particularly over areas dominated by natural aerosols such as dust. Accordingly, our method improves the seasonal ground-level PM<sub>2.5</sub> maps, e.g. the bias of the autumn (winter) mean of ground-level PM<sub>2.5</sub> estimates over Qinghai and Gansu (Shaanxi, Shanxi, and Henan) provinces reduces from –8% to –5% (11%–6%).

\* Corresponding author.

E-mail addresses: [Fei.Yao@ed.ac.uk](mailto:Fei.Yao@ed.ac.uk) (F. Yao), [Paul.Palmer@ed.ac.uk](mailto:Paul.Palmer@ed.ac.uk) (P.I. Palmer).



## 1. Introduction

Fine particulate matter, particles with an aerodynamic diameter  $\leq 2.5\mu\text{m}$  ( $\text{PM}_{2.5}$ ), have well-documented deleterious impacts on human health (Lelieveld et al., 2015; Burnett et al., 2018), and under certain environment conditions can result in cloud condensation nuclei that subsequently alter cloud optical properties that perturb Earth's radiative balance (Stephens, 2005; Seinfeld et al., 2016). Current knowledge of ground-level  $\text{PM}_{2.5}$  and its chemical constituents is limited to *in situ* instruments that are typically located in regional clusters across the globe, with generally fewer calibrated sensors available in less developed countries. To address this measurement gap, the science community has developed progressively more sophisticated methods to relate satellite remote sensing observations of aerosol optical depth (AOD), an integrated measure of aerosol extinction of light as it passes through the atmosphere, to ground-level  $\text{PM}_{2.5}$  concentrations.

Translating satellite-retrieved AOD into estimates of ground-level  $\text{PM}_{2.5}$  is non-trivial. Previous studies have used a wide range of process-driven and data-driven models to determine  $\text{PM}_{2.5}$ :AOD relationships (Fig. 1). Process-driven models, such the GEOS-Chem global 3-D model of atmospheric chemistry and transport, take into account emission distributions, atmospheric chemistry and transport to simulate self-consistent  $\text{PM}_{2.5}$ :AOD (Van Donkelaar et al., 2010, 2015; Boys et al., 2014; Hammer et al., 2020). However, sources of model error, e.g. vertical mixing of surface emissions and heterogeneous chemistry, will introduce uncertainties in the corresponding ground-level  $\text{PM}_{2.5}$  estimates. The alternative is to develop data-driven  $\text{PM}_{2.5}$ :AOD relationships using statistical methods, e.g. the linear mixed effects model (Lee et al., 2011; Ma et al., 2016b; Zhang et al., 2018), geographically weighted regression model (Hu et al., 2013; Ma et al., 2014; Guo et al., 2017; He and Huang, 2018b, a), and two-stage model (Hu et al., 2014; Ma et al., 2016a; Yao et al., 2019), and machine learning, e.g. artificial neural networks (Gupta and Christopher, 2009; Wu et al., 2012; Guo et al., 2013) and the random forest (RF) (Hu et al., 2017; Wei et al., 2019, 2020). While these methods show promise, they are sometimes difficult to interpret without a supporting in-depth investigation.

A major criticism of data-driven models is that they do not explicitly

take into consideration that AOD and ground-level  $\text{PM}_{2.5}$  are different measures of atmospheric aerosols. Ground-level  $\text{PM}_{2.5}$  measurements are mass concentrations ( $\mu\text{gm}^{-3}$ ) of fine particles measured under controlled relative humidity (RH) conditions: 35% in the US (U.S. Environmental Protection Agency, 1997), and 50% in Europe (European Committee for Standardization (CEN), 1998) and China (HJ 618-2011, 2011). In contrast, AOD (unitless) reflects a vertical integration of aerosol extinction coefficients ( $\text{m}^{-1}$ ) at a specific wavelength (typically 550 nm). There are a range of factors that can influence AOD and/or ground-level  $\text{PM}_{2.5}$  and consequently strengthen or weaken their relationship. For example, covariation of free tropospheric aerosols (e.g. desert dust) and ground-level  $\text{PM}_{2.5}$  may result in a strong relationship between the two measurements without any physical connection.

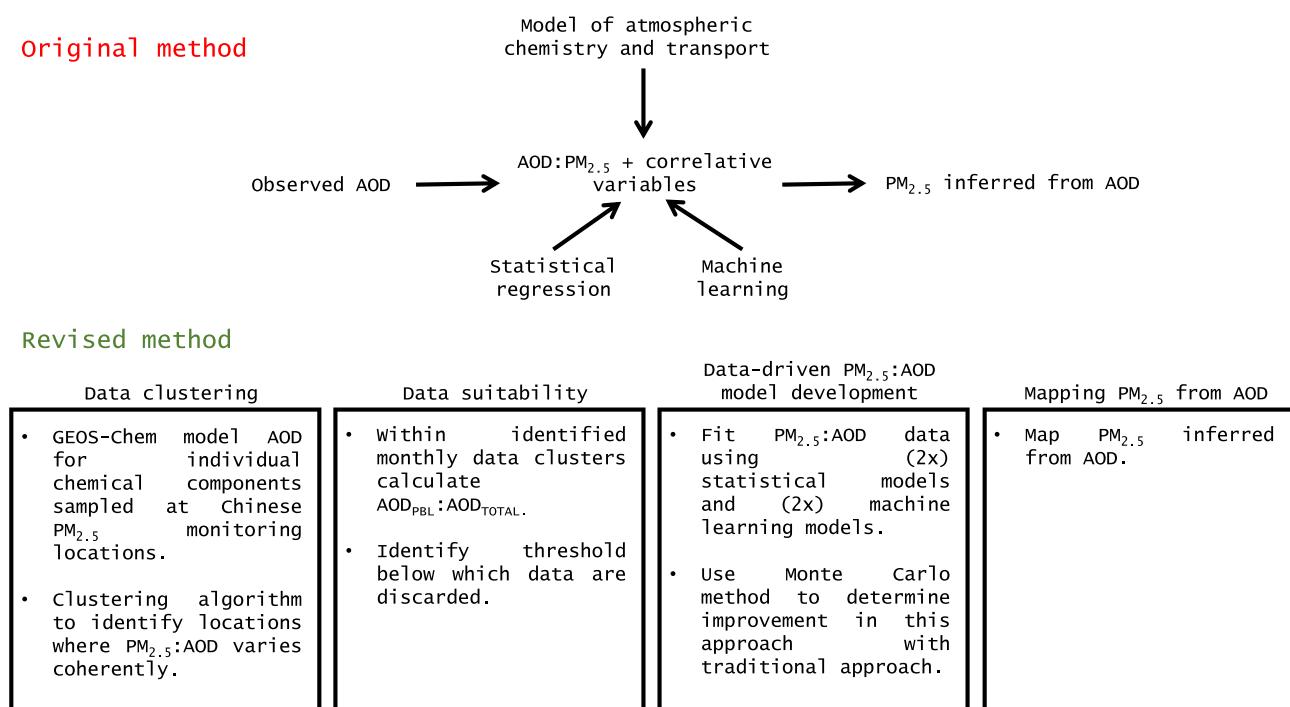
In this study, we propose a model framework to infer ground-level  $\text{PM}_{2.5}$  from satellite-retrieved AOD. Most importantly, the method identifies observed scenes for which aerosol extinction coefficients is mostly within the planetary boundary layer (PBL), where variations are mostly likely to reflect changes in ground-level  $\text{PM}_{2.5}$ . In the next section we describe the ground-level  $\text{PM}_{2.5}$  and AOD measurements, the nested version of the GEOS-Chem global 3-D model of atmospheric chemistry and transport, and the statistical regression and machine learning models that we use to develop our  $\text{PM}_{2.5}$ :AOD relationships. In section 3, we report the results of our method and compare them with conventional methods. In section 4, we conclude the paper by reflecting on our results, outlining limitations to our approach, and discussing future research directions.

## 2. Materials and methods

### 2.1. Data

#### 2.1.1. Ground-level $\text{PM}_{2.5}$ measurements

We collect hourly ground-level  $\text{PM}_{2.5}$  measurements for 2014 from the China National Environmental Monitoring Center (CNEMC, <http://www.cnemc.cn/>). The data are measured either by the tapered element oscillating microbalance method (TEOM) or the beta-attenuation method (HJ 618-2011, 2011). We remove values by



**Fig. 1.** A brief description of the original method of inferring ground-level  $\text{PM}_{2.5}$  concentrations from satellite observations of AOD, and the revised method that has developed in this study and most importantly takes into account the fraction of columnar AOD that resides in the planetary boundary layer.

following the first three steps of the *ad hoc* quality control protocol developed by Jiang et al. (2020). We limit our study area to eastern China, defined as 95–140° E and 15–55° N (Fig. 2) due to limited ground-level PM<sub>2.5</sub> monitoring sites over western China. For our analysis we consider mean ground-level PM<sub>2.5</sub> measurements taken between 1000 and 1100 and between 1300 and 1400 local times, corresponding to the 1030 and 1330 local equatorial overpass times of Terra and Aqua satellites. These mean values are described on the GEOS-Chem model 0.25° (latitude) × 0.3125° (longitude) nested grid, described below.

### 2.1.2. MODIS MAIAC and AERONET AOD retrievals

We use data retrieved from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the Terra and Aqua satellites. In particular, we use the MODIS Collection 6 Multi-Angle Implementation of Atmospheric Correction (MAIAC) AOD product available on a 1 km spatial resolution (Lyapustin et al., 2011, 2018) covering eastern China for 2014 downloaded from the Level-1 and Atmosphere Archive & Distribution System (<https://ladsweb.modaps.eosdis.nasa.gov/>). We use Terra and Aqua MAIAC AOD at 550 nm, corresponding to particle sizes of 0.1–2 μm and comparable to the size range of PM<sub>2.5</sub> (Kahn et al., 1998). To evaluate the MAIAC AOD data over China, we compare values against AOD retrievals from the NASA AErosol RObotic NETwork (AERONET, version 3: <https://aeronet.gsfc.nasa.gov/>) as these ground-based data are generally recognized as a ‘ground truth’ for satellite observations of aerosol optical properties (Holben et al., 1998). Our analysis of this comparison is presented in the Supporting Information (SI, Text S1, Figs. S1 and S2). We find that the accuracy of MAIAC AOD is insensitive to the different values of MODIS retrieval quality assurance flag, allowing us to use the complete MAIAC dataset. We regrid the 1 km Terra and Aqua MAIAC AOD data product to the coarser GEOS-Chem model 0.25° (latitude) × 0.3125° (longitude) nested grid.

## 2.2. Methods

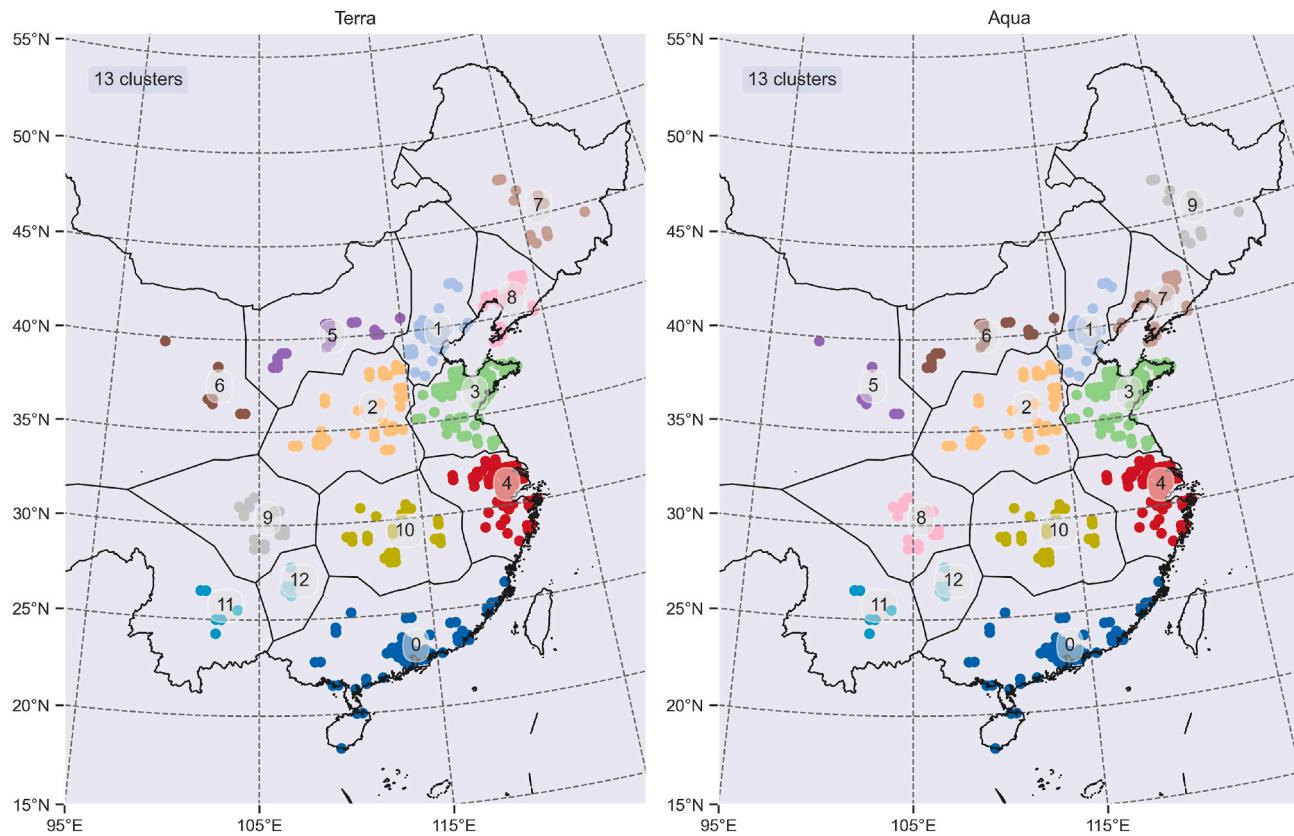
Fig. 1 describes the computational approach we propose to infer ground-level PM<sub>2.5</sub> concentrations from MODIS MAIAC AOD retrievals. The approach is described by four broad steps: 1) data clustering to identify spatially proximate areas where ground-level PM<sub>2.5</sub> and AOD behave in a coherent manner; 2) data suitability to retain data where a certain fraction of the columnar AOD is found in the PBL, which we assume is well-mixed and consequently reflects variations of ground-level PM<sub>2.5</sub>; 3) statistical and machine learning data-driven models of the relationship between ground-level PM<sub>2.5</sub> and AOD, which is subsequently used to (step 4) infer ground-level PM<sub>2.5</sub> concentrations from satellite observations of AOD. In this section, we describe each step in our approach and clearly outline any assumptions we have made.

### 2.2.1. GEOS-chem model of atmospheric chemistry and transport

We use v12.5.0 of the GEOS-Chem global 3-D model of atmospheric chemistry and transport (Bey et al., 2001) to relate emissions and atmospheric chemistry to hourly ground-level PM<sub>2.5</sub> and 3-D fields of aerosol extinction coefficients and its individual chemical components over eastern China during 2014 on a 0.25° (latitude) × 0.3125° (longitude) nested grid. We describe in detail the emission inventory, meteorology, chemistry, and evaluation of the GEOS-Chem model in the SI (Text S2, Figs. S3–S6). We find that the GEOS-Chem generally reproduces the observed variations of AOD and ground-level PM<sub>2.5</sub>, providing confidence in use of this model in our study.

### 2.2.2. Data clustering

The relationship between ground-level PM<sub>2.5</sub> and AOD varies spatially (Xiao et al., 2018; Yao et al., 2019), so we develop spatially adaptive models to infer ground-level PM<sub>2.5</sub> concentrations from satellite observations of AOD. We assume that ground-level PM<sub>2.5</sub> and AOD



**Fig. 2.** Spatial clusters of ground-level PM<sub>2.5</sub> monitoring locations over eastern China during 2014 for (left) Terra and (right) Aqua overpass times. Clusters are determined by one year of GEOS-Chem 3-D model fields of individual chemical components of AOD using the UPGMA hierarchical clustering method.

behave in a coherent manner over spatially proximate areas where individual chemical components of AOD vary similarly. We identify these areas using the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering method. The UPGMA hierarchical clustering method includes three steps:

**2.2.2.1. Step 1.** Calculate the correlation distance between any two ground-level PM<sub>2.5</sub> monitoring locations using equation (1). The distance between two locations will be related to the extent to which we expect agreement of individual chemical components of AOD at these two locations. Repeating the calculation for all combinations of ground-level PM<sub>2.5</sub> monitoring locations returns a distance matrix:

$$D_{ij} = \frac{1}{N} \sum_{k=1}^N (1 - R_{ijk}), \quad (1)$$

where  $R_{ijk}$  denotes the Pearson correlation between ground-level PM<sub>2.5</sub> monitoring location  $i$  and  $j$  for  $k$ th chemical component of AOD using a certain period (e.g. a month and a year in this study) of time series data.  $N$  is the total number of individual chemical components of AOD, which in this study is eight. The distance  $D$  will vary between 0 and 2.

**2.2.2.2. Step 2.** Starting from the initial distance matrix formed in Step 1, in which all ground-level PM<sub>2.5</sub> monitoring locations are thought as a single cluster, we first search the pair that has the minimum distance. Supposing they are  $A$  and  $B$ , we then merge them into a new cluster called  $A \cup B$ . Next, we calculate the distances between  $A \cup B$  and the remaining clusters by equation (2). Hence we reduce the distance matrix by one row and one column. We repeat this process until the distance matrix has been downgraded to a simple scalar, so that the original ground-level PM<sub>2.5</sub> monitoring locations are all merged and we obtain a dendrogram:

$$D_{A \cup B, X} = \frac{|A| \cdot D_{A, X} + |B| \cdot D_{B, X}}{|A| + |B|}, \quad (2)$$

where  $X$  denotes a remaining cluster,  $D_{A, X}$  and  $D_{B, X}$  denotes its distance from  $A$  and  $B$ .  $|A|$  and  $|B|$  denotes the total number of original ground-level PM<sub>2.5</sub> monitoring locations that cluster  $A$  and  $B$  hold. The updated distance is then with respect to initial distances rather than mathematical procedure, i.e. why the method is called unweighted.

**2.2.2.3. Step 3.** A clear relationship exists between the distance threshold chosen for the dendrogram and the number of clusters emerged. Since no universally method can determine the best distance threshold, we choose 0.5 as our distance threshold. In this sense, the mean similarity of variations of individual chemical components of AOD within the same cluster is no less than 0.5. We finally generate Thiessen polygons (Thiessen, 1911) from the ground-level PM<sub>2.5</sub> monitoring locations and assign the GEOS-Chem model 0.25° (latitude) × 0.3125° (longitude) nested grid cells within each Thiessen polygon to the same cluster of the corresponding ground-level PM<sub>2.5</sub> monitoring location in the center.

We conduct the data clustering on a monthly and an annual scale. We identify data for which we expect the strongest physical relationship between ground-level PM<sub>2.5</sub> and AOD by combining the monthly data clusters with simple correlation techniques (section 3.2). We develop robust PM<sub>2.5</sub>:AOD models for geographical regions represented by the clustered ground-level PM<sub>2.5</sub> monitoring locations by combining the annual data clusters with sophisticated statistical and machine learning models (section 3.3).

### 2.2.3. Data suitability

Intuitively, we expect the strongest physical relationship between ground-level PM<sub>2.5</sub> and total columnar AOD when the majority of the contributing aerosol extinction coefficients resides in the PBL that is

connected to aerosols at the surface. First, we define the fraction of total columnar AOD that resides in the PBL ( $\Gamma_{PBL}^{AOD}$ ) using the 3-D fields of aerosol extinction coefficients simulated from the GEOS-Chem model. We then explore the robustness of the PM<sub>2.5</sub>:AOD relationship and how it varies with  $\Gamma_{PBL}^{AOD}$ . We quantify robustness using the Pearson and Spearman rank correlation coefficients, acknowledging that the PM<sub>2.5</sub>:AOD may vary monotonically without necessarily being linear. By examining how these correlations vary with  $\Gamma_{PBL}^{AOD}$  for each monthly data cluster we identify a threshold value for  $\Gamma_{PBL}^{AOD}$  that ensures a robust PM<sub>2.5</sub>:AOD relationship that is determined by a large number of observations. By using this approach, we identify suitable data that provide the strongest constraints on our data-driven models that we describe below.

### 2.2.4. Data-driven model development

We use two statistical methods and two machine learning methods to model the PM<sub>2.5</sub>:AOD relationship for each annual data cluster with a range of auxiliary meteorological predictors. The meteorological predictors include PBL height (PBLH), mean RH in the PBL (RH\_PBL), surface skin temperature (TS), total precipitation (PRECTOT), 10-metre eastward wind (U10M), 10-metre northward wind (V10M), and sea level pressure (SLP). Mean values of these fields from 1000 to 1100 and 1300–1400 local times, corresponding to Terra and Aqua overpass times, are taken from the Goddard Earth Observing System-Forward Processing (GEOS-FP) meteorological data (Lucchesi, 2018). We describe more details on the GEOS-FP meteorology in the SI (Text S2).

We first construct a pooled ordinary least square (PooledOLS) model (equation (3)):

$$\begin{aligned} PM_{2.5g}^d = & \beta_0 + \beta_{AOD} AOD_g^d + \beta_{PBLH} PBLH_g^d \\ & + \beta_{RH\_PBL} RH\_PBL_g^d + \beta_{TS} TS_g^d \\ & + \beta_{PRECTOT} PRECTOT_g^d + \beta_{U10M} U10M_g^d \\ & + \beta_{V10M} V10M_g^d + \beta_{SLP} SLP_g^d + \epsilon_g^d, \end{aligned} \quad (3)$$

where  $PM_{2.5g}^d$ ,  $AOD_g^d$ ,  $PBLH_g^d$ ,  $RH\_PBL_g^d$ ,  $TS_g^d$ ,  $PRECTOT_g^d$ ,  $U10M_g^d$ ,  $V10M_g^d$ ,  $SLP_g^d$  represent the ground-level PM<sub>2.5</sub>, MAIAC AOD, PBLH, RH\_PBL, TS, PRECTOT, U10M, V10M, SLP, respectively, at grid  $g$  on day  $d$  during Terra or Aqua overpass time;  $\beta_{AOD}$ ,  $\beta_{PBLH}$ ,  $\beta_{RH\_PBL}$ ,  $\beta_{TS}$ ,  $\beta_{PRECTOT}$ ,  $\beta_{U10M}$ ,  $\beta_{V10M}$ ,  $\beta_{SLP}$  denote the corresponding coefficients;  $\beta_0$  is the intercept; and  $\epsilon_g^d$  is the error term at grid  $g$  on day  $d$  during Terra or Aqua overpass time.

We then consider a time-fixed-effects model (TFEM), equivalent to equation (3) but with an additional day-specific offset  $\beta_d$ , that allows us to capture temporal heterogeneities in the PM<sub>2.5</sub>:AOD relationship (Yao et al., 2018). We choose to fix time in the fixed-effects model because our data analysis suggests that aerosols are more variable in the time domain than in the space domain (SI, Fig. S7).

We use a RF model as an alternative approach to describe the relationship between ground-level PM<sub>2.5</sub> and AOD. RF is an ensemble of decision trees, with each tree constructed using bootstrap samples drawn from the original data (Breiman, 2001). For our study, the final prediction is the mean of predictions from all individual decision trees. Unlike statistical regression models, such as the PooledOLS model, RF typically involves fewer, less restrictive assumptions regarding independence of predictors and distributions of outcomes, with no *a priori* requirement of a formula between the predictors and the outcome variables.

We generate two RF models (RF1: equation (4) and RF2: equation (5)) that correspond to the PooledOLS and TFEM regression models, respectively:

$$\begin{aligned} PM_{2.5g}^d = & f(AOD_g^d, PBLH_g^d, RH\_PBL_g^d, TS_g^d, \\ & PRECTOT_g^d, U10M_g^d, V10M_g^d, SLP_g^d) \end{aligned} \quad (4)$$

$$= f(AOD_g^d, PBLH_g^d, RH\_PBL_g^d, TS_g^d, PRECTOT_g^d, U10M_g^d, V10M_g^d, SLP_g^d, DOY_g^d) \quad (5)$$

where  $f(\cdot)$  denotes a function representing the RF,  $DOY_g^d$  denotes the day of year at grid  $g$  on day  $d$  during Terra or Aqua overpass time, and the remaining variables are as previously defined. Each RF model is determined by using 1,000 fully grown and unpruned decision trees.

We use both the full data (ignoring the data suitability step described above) and the suitable data (taking into consideration  $\Gamma_{PBL}^{AOD}$ ) to train the statistical and machine learning models. We fit a linear regression between predicted and measured ground-level PM<sub>2.5</sub>. We use the coefficient of determination ( $R^2$ ) of the linear regression and the mean percentage error  $MPE = \frac{1}{N} \sum \left( \frac{PM_{2.5obs} - PM_{2.5pred}}{PM_{2.5obs}} \right)$ , where  $PM_{2.5obs}$  and  $PM_{2.5pred}$  are observed and predicted PM<sub>2.5</sub>, respectively, and  $N$  is the number of comparison points, to evaluate the goodness of model fitting (MF). Positive (negative) MPE values denote negative (positive) model bias.

### 2.2.5. Data-driven model validation and comparison

We conduct a sample-based 10-fold cross validation (CV) to evaluate model performance on yet unseen data. First, we split randomly and evenly the full/suitable data into 10 subsets. For each subset, we then calculate predictions using the model trained from the remaining nine subsets. Next, we fit a linear regression between the CV predictions and observations and report the  $R^2$  and the MPE. These CV statistics reflect the model performance on yet unseen data. The difference between MF and CV statistics can help identify whether a model under- or over-fits the data.

By comparing the MF and CV statistics of models trained by the full data and by the suitable data we can evaluate the comparative model improvement. We use a Monte Carlo approach to determine whether the potential model improvement is statistically significant. First, we randomly select a subset of the full data that matches the length of the suitable data and we repeat that process 1,000 times. Using each selected data sample, we build the statistical and machine learning models and perform the sample-based 10-fold CV. We subsequently obtain a series of MF and CV statistics, based on which we calculate the possibility (p value) of obtaining a model that performs no worse than our approach but has no skill of AOD vertical partitioning.

### 2.2.6. Ground-level PM<sub>2.5</sub> mapping

Finally, we use the statistical and machine learning models to predict ground-level PM<sub>2.5</sub> concentrations during Terra and Aqua overpass times on a daily basis across eastern China at the GEOS-Chem nested model horizontal resolution of 0.25° (latitude) × 0.3125° (longitude). We use both the models trained by the full and the suitable data for ground-level PM<sub>2.5</sub> prediction. For the latter, we only predict at locations and times where  $\Gamma_{PBL}^{AOD}$  is no less than the determined threshold. We further compile daily ground-level PM<sub>2.5</sub> estimates to seasonal values, based on which we make another comparison between models trained by the full and the suitable data.

## 3. Results

### 3.1. Results of data clustering

**Fig. 2** shows the data clustering results based on one year of individual chemical components of AOD simulated from the GEOS-Chem model using the UPGMA hierarchical clustering method, described above. We determine a total of 13 spatial clusters with similar extent across China. Generally, we find little difference between results using model values taken from the 1030 and 1330 Terra and Aqua overpass times.

We find that the majority of these spatial clusters have boundaries similar to those of urban agglomerations, e.g. Pearl River Delta urban agglomeration (cluster 0), Yangtze River Delta urban agglomeration (cluster 4), and the Yangtze River Mid-Reaches urban agglomeration (cluster 10). This, to some extent, suggests remarkable influence of anthropogenic activities on aerosols. We also find that most clusters are dominated by secondary aerosols except cluster 6 for Terra and cluster 5 for Aqua overpass time, where dust aerosol is prevalent (SI, **Table S1**). Moreover, these dust aerosols are frequently found in the free troposphere, e.g. the median ± median absolute deviation of the fraction of dust aerosols above the PBL at ground-level PM<sub>2.5</sub> monitoring locations in cluster 6 for Terra and cluster 5 for Aqua overpass time are 61 ± 21% and 44 ± 21%, respectively.

**Fig. S8** extends our data clustering analysis to monthly scales. Generally we find, irrespective of the local overpass times, more spatial clusters in warmer months (e.g. April–September) than during colder months (e.g. January–March, October–December), suggesting a wider spatial association of aerosols during colder months, e.g. from domestic heating.

### 3.2. Results of data suitability

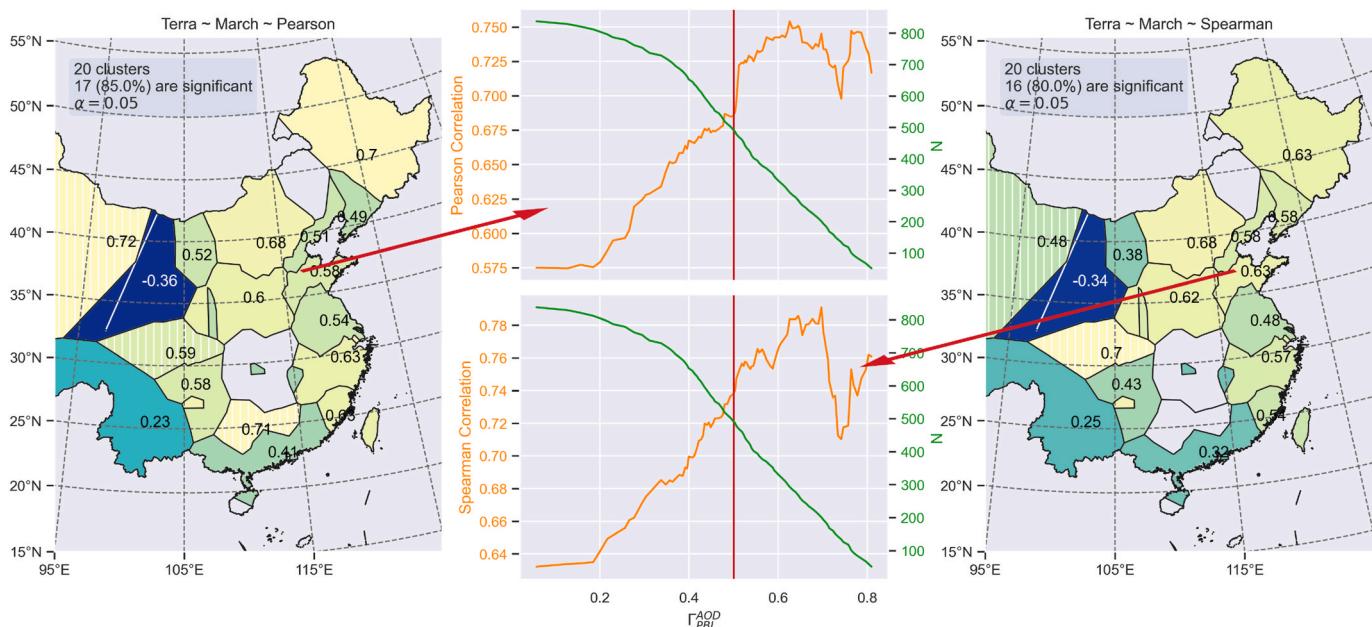
For each cluster in each calendar month shown in **Fig. S8** we analyze how the PM<sub>2.5</sub>:AOD relationship changes with  $\Gamma_{PBL}^{AOD}$ . **Fig. 3** shows an example analysis. We find that the Pearson and Spearman rank correlation coefficients between ground-level PM<sub>2.5</sub> and AOD generally increase with increasing  $\Gamma_{PBL}^{AOD}$ , as expected, despite a progressively decreasing number of observations that results in some fluctuations. There are variations in our analysis between clusters and months, but **Fig. 3** illustrates how  $\Gamma_{PBL}^{AOD}$  can be used effectively to identify a subset of data that ensures a physically-meaningful PM<sub>2.5</sub>:AOD relationship. Based on our analysis over all clusters and months, we use a  $\Gamma_{PBL}^{AOD}$  value of 0.5 as a threshold, below which we discard the corresponding data.

### 3.3. Results of data-driven model development, validation, and comparison

#### 3.3.1. Results on the national scale

**Table 1** shows the overall MF and CV results of the PooledOLS, TFEM, RF1, and RF2 models. The difference between MF and CV statistics are smallest for the PooledOLS model and largest for the RF models. We interpret this as the PooledOLS model underfitting the data and the RF models overfitting the data. The  $R^2$  in the RF models are generally higher than those in the PooledOLS model and the TFEM model, suggesting the non-linearity in the PM<sub>2.5</sub>:AOD relationship has been captured by the non-linear machine learning models. The models that account for day-to-day variations generally have a higher  $R^2$  and a smaller absolute MPE. This is consistent with of day-specific offsets in the TFEM model and the DOY predictor in the RF2 model having skill in reproducing temporal variations in the PM<sub>2.5</sub>:AOD relationship. We generally find that models perform better for data collected during Aqua overpass time when the PBLH is close to its highest values ([Chatfield et al., 2020](#)). We also find that fewer data are discarded due to the  $\Gamma_{PBL}^{AOD} \geq 0.5$  criterion at the Aqua overpass time (≈ 16%) compared to the earlier Terra overpass time (≈ 37%).

We find that the model trained by the suitable data generally has smaller absolute MPE values than the model using the full data, e.g. the model trained by the suitable data reduces the model bias using the full data by 10–15% and 9–12% for the Terra and Aqua overpass times, respectively, during cross validation. We find that the model trained by the suitable data generally has comparable or larger  $R^2$  values than the model using the full data, e.g. the model trained by the suitable data improves the model  $R^2$  using the full data by up to 8% and 5% for the Terra and Aqua overpass times, respectively, during cross validation. Moreover, we find the p values are all equivalent to zero, suggesting that



**Fig. 3.** (Orange) Pearson and Spearman rank correlation coefficients between ground-level PM<sub>2.5</sub> and AOD, and the (green) corresponding number of observations, as a function of  $\Gamma_{PBL}^{AOD}$ , in a specific monthly cluster. We use a  $\Gamma_{PBL}^{AOD}$  value of 0.5 (red vertical lines) as a threshold, below which we discard the corresponding data. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

**Table 1**

Overall model fitting and cross validation results of the PooledOLS, TFEM, RF1, and RF2 models in 2014 over eastern China. N,  $R^2$ , and MPE represent statistics of the model trained by the full data, while  $N'$ ,  $R'^2$ , and MPE' represent those of the model trained by the suitable data.  $R_p^2$  and MPE<sub>p</sub> represent the possibility of obtaining a model performance no worse than that of our approach by randomly selecting a subset of the full data that matches the length of the suitable table to train the model.

	Model	Terra								Aqua							
		N	$N'$	$R^2$	$R'^2$	$R_p^2$	MPE	MPE'	MPE <sub>p</sub>	N	$N'$	$R^2$	$R'^2$	$R_p^2$	MPE	MPE'	MPE <sub>p</sub>
MF	PooledOLS	57819	36692	0.37	0.39	0.0	-0.48	-0.41	0.0	55939	46961	0.43	0.45	0.0	-0.45	-0.41	0.0
	PanelOLS	57819	36692	0.63	0.67	0.0	-0.25	-0.21	0.0	55939	46961	0.69	0.71	0.0	-0.24	-0.21	0.0
	RF1	57819	36692	0.96	0.96	0.0	-0.12	-0.10	0.0	55939	46961	0.96	0.97	0.0	-0.11	-0.10	0.0
	RF2	57819	36692	0.97	0.96	0.0	-0.10	-0.09	0.0	55939	46961	0.97	0.97	0.0	-0.10	-0.09	0.0
CV	PooledOLS	57819	36692	0.36	0.39	0.0	-0.48	-0.41	0.0	55939	46961	0.43	0.45	0.0	-0.45	-0.41	0.0
	PanelOLS	57819	36692	0.58	0.58	0.0	-0.27	-0.24	0.0	55939	46961	0.64	0.66	0.0	-0.26	-0.23	0.0
	RF1	57819	36692	0.63	0.63	0.0	-0.32	-0.28	0.0	55939	46961	0.67	0.69	0.0	-0.31	-0.28	0.0
	RF2	57819	36692	0.68	0.66	0.0	-0.29	-0.26	0.0	55939	46961	0.73	0.73	0.0	-0.28	-0.25	0.0

the comparative model improvement is statistically significant.

### 3.3.2. Results on the regional scale

For brevity, we focus on CV results in each cluster shown in Table 2. In terms of CV MPE, we find that most clusters obtain a statistically significant smaller absolute value. The number of these clusters are 11, 7, 10, and 10 in the PooledOLS, TFEM, RF1, and RF2 models, respectively, during the earlier Terra overpass time. Corresponding values for the Aqua overpass time are 11, 7, 11, and 9, respectively. In addition, the decrease of the absolute MPE values is spatially homogeneous. For instance, we observe the biggest decrease in both clusters 3 and 4 dominated by anthropogenic aerosols and clusters 5 and 6 dominated by natural aerosols (SI, Table S1). A few clusters show larger absolute MPE values but the increase are all limited to  $\leq 0.03$ , a very small value. Different model performances across clusters under the same criterion,  $\Gamma_{PBL}^{AOD} \geq 0.5$ , reflects the importance of individual chemical components that comprise PM<sub>2.5</sub> and AOD (and thereby the data clustering step) to the development of robust PM<sub>2.5</sub>: AOD models.

In terms of CV  $R^2$ , its behaviour is slightly different during Terra and Aqua overpass times. For the Aqua overpass time, we find that most clusters obtain a statistically significantly higher  $R^2$ . For the Terra overpass time, we admit that there is a smaller number of clusters

showing statistically significant increased CV  $R^2$ . A few clusters present declined CV  $R^2$  to a slightly large extent, suggesting that  $\Gamma_{PBL}^{AOD} \geq 0.5$  might be too strict a criterion for the Terra overpass time from the perspective of capturing ground-level PM<sub>2.5</sub> concentrations. Nonetheless, we find that cluster 6 (5), corresponding to the Terra (Aqua) overpass time, where dust aerosol is prevalent (SI, Table S1) generally shows the largest improvement in CV  $R^2$  that is distinguishable from those in other clusters. This highlights that using more AOD data blindly without considering their representativeness for ground-level PM<sub>2.5</sub> will not always lead to better estimates of ground-level PM<sub>2.5</sub> even if a sophisticated machine learning model is used. Careful filtering of these data is required to build confidence and robustness of resulting estimates, and eventually scientific exploitation of satellite-derived ground-level PM<sub>2.5</sub> data products.

### 3.4. Ground-level PM<sub>2.5</sub> maps

For illustrative purposes, we show seasonal mean values inferred from RF2 using data from the Aqua overpass time. Fig. 4 shows the ground-level PM<sub>2.5</sub> estimates inferred from the full data and from the subset of these data where  $\Gamma_{PBL}^{AOD} \geq 0.5$ . We find that the seasonal distributions of PM<sub>2.5</sub> inferred from the full and reduced AOD datasets are

**Table 2**

Cross validation results of the PooledOLS, TFEM, RF1, and RF2 models in each cluster of eastern China in 2014.

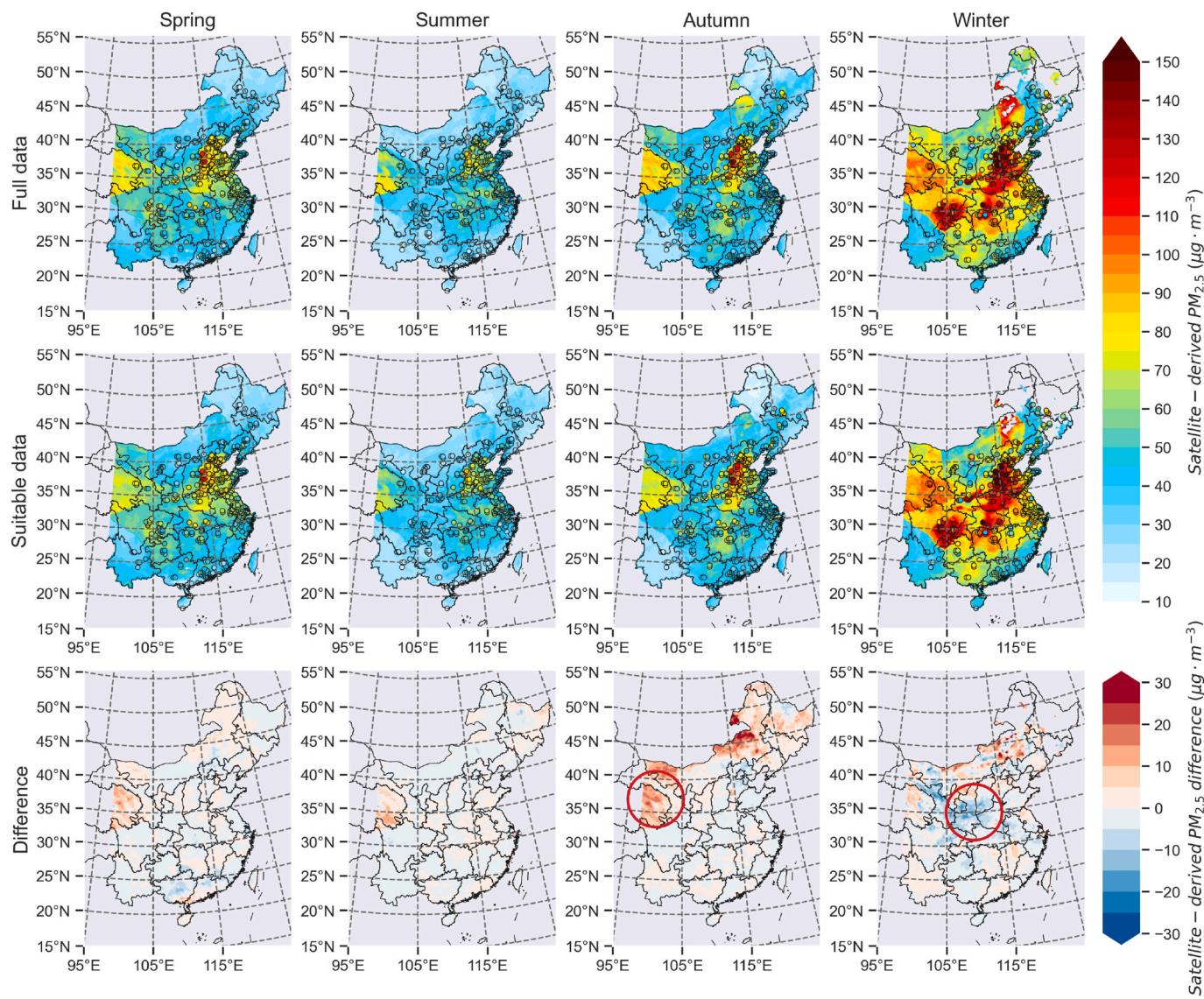
	Cluster	Terra							Aqua								
		N	N'	R <sup>2</sup>	R <sup>2'</sup>	R <sup>2</sup> <sub>p</sub>	MPE	MPE'	MPE <sub>p</sub>	N	N'	R <sup>2</sup>	R <sup>2'</sup>	R <sup>2</sup> <sub>p</sub>	MPE	MPE'	MPE <sub>p</sub>
PooledOLS	0	7857	5410	0.39	0.42	0.00	-0.37	-0.34	0.00	7776	6776	0.46	0.51	0.00	-0.30	-0.29	0.00
	1	6172	3458	0.34	0.33	0.73	-0.84	-0.77	0.00	5650	4716	0.46	0.47	0.07	-0.80	-0.77	0.02
	2	6312	4246	0.34	0.36	0.05	-0.46	-0.44	0.25	6039	5475	0.37	0.38	0.13	-0.50	-0.49	0.03
	3	10975	7202	0.30	0.34	0.00	-0.54	-0.42	0.00	10277	8388	0.35	0.38	0.00	-0.48	-0.40	0.00
	4	7672	5301	0.31	0.29	0.99	-0.31	-0.25	0.00	7243	6045	0.35	0.37	0.00	-0.29	-0.25	0.00
	5	4105	2020	0.24	0.24	0.42	-0.52	-0.40	0.00	1612	1313	0.31	0.35	0.01	-0.41	-0.35	0.00
	6	1603	756	0.16	0.22	0.00	-0.44	-0.37	0.00	3974	3145	0.30	0.35	0.00	-0.64	-0.54	0.00
	7	2278	1380	0.42	0.47	0.00	-0.44	-0.41	0.09	5376	3953	0.33	0.35	0.02	-0.49	-0.46	0.00
	8	5206	3006	0.24	0.29	0.00	-0.54	-0.50	0.00	1203	1097	0.47	0.51	0.00	-0.33	-0.28	0.00
	9	1189	667	0.40	0.38	0.76	-0.32	-0.27	0.02	2222	1757	0.58	0.65	0.00	-0.39	-0.36	0.02
	10	2920	1808	0.19	0.26	0.49	-0.36	-0.31	0.00	2861	2611	0.29	0.27	0.84	-0.32	-0.31	0.06
TFEM	11	1266	1233	0.28	0.30	0.00	-0.26	-0.25	0.00	1341	1331	0.28	0.28	0.10	-0.24	-0.24	0.76
	12	264	205	0.39	0.45	0.02	-0.37	-0.31	0.02	365	354	0.42	0.44	0.03	-0.36	-0.39	0.36
	0	7857	5410	0.63	0.63	0.36	-0.22	-0.22	0.02	7776	6776	0.63	0.65	0.00	-0.22	-0.20	0.00
	1	6172	3458	0.54	0.53	0.14	-0.48	-0.42	0.00	5650	4716	0.65	0.67	0.01	-0.40	-0.39	0.07
	2	6312	4246	0.51	0.53	0.00	-0.27	-0.27	0.33	6039	5475	0.56	0.59	0.00	-0.28	-0.26	0.00
	3	10975	7202	0.59	0.59	0.14	-0.26	-0.23	0.00	10277	8388	0.63	0.64	0.01	-0.24	-0.21	0.00
	4	7672	5301	0.60	0.59	0.44	-0.17	-0.14	0.00	7243	6045	0.67	0.69	0.00	-0.15	-0.13	0.00
	5	4105	2020	0.40	0.36	0.37	-0.38	-0.31	0.00	1612	1313	0.32	0.38	0.00	-0.32	-0.26	0.00
	6	1603	756	0.19	0.27	0.00	-0.33	-0.25	0.02	3974	3145	0.49	0.53	0.00	-0.45	-0.38	0.00
	7	2278	1380	0.63	0.49	1.00	-0.25	-0.23	0.11	5376	3953	0.59	0.68	0.00	-0.24	-0.22	0.00
	8	5206	3006	0.57	0.60	0.00	-0.26	-0.23	0.00	1203	1097	0.70	0.72	0.08	-0.15	-0.14	0.12
	9	1189	667	0.62	0.44	0.98	-0.17	-0.18	0.51	2222	1757	0.69	0.74	0.00	-0.25	-0.23	0.08
	10	2920	1808	0.57	0.48	0.99	-0.21	-0.18	0.00	2861	2611	0.60	0.60	0.06	-0.16	-0.17	0.43
RF1	11	1266	1233	0.50	0.50	0.94	-0.13	-0.13	0.87	1341	1331	0.39	0.40	0.71	-0.14	-0.15	0.68
	12	264	205	0.44	0.34	0.47	-0.26	-0.27	0.50	365	354	0.56	0.58	0.13	-0.29	-0.25	0.24
	0	7857	5410	0.67	0.68	0.00	-0.24	-0.22	0.00	7776	6776	0.65	0.67	0.01	-0.24	-0.22	0.00
	1	6172	3458	0.62	0.60	0.00	-0.53	-0.45	0.00	5650	4716	0.70	0.71	0.04	-0.49	-0.43	0.00
	2	6312	4246	0.55	0.57	0.00	-0.36	-0.35	0.33	6039	5475	0.62	0.62	0.01	-0.34	-0.31	0.00
	3	10975	7202	0.64	0.64	0.00	-0.30	-0.26	0.00	10277	8388	0.68	0.70	0.00	-0.26	-0.22	0.00
	4	7672	5301	0.64	0.63	0.09	-0.19	-0.16	0.00	7243	6045	0.68	0.71	0.00	-0.19	-0.17	0.00
	5	4105	2020	0.40	0.37	0.12	-0.44	-0.37	0.00	1612	1313	0.37	0.39	0.06	-0.39	-0.37	0.01
	6	1603	756	0.24	0.38	0.00	-0.40	-0.32	0.00	3974	3145	0.50	0.52	0.00	-0.56	-0.49	0.00
	7	2278	1380	0.68	0.63	0.54	-0.37	-0.35	0.00	5376	3953	0.59	0.62	0.00	-0.30	-0.26	0.00
	8	5206	3006	0.59	0.62	0.00	-0.32	-0.29	0.00	1203	1097	0.70	0.73	0.01	-0.24	-0.22	0.01
	9	1189	667	0.64	0.55	0.85	-0.23	-0.23	0.10	2222	1757	0.79	0.80	0.08	-0.35	-0.32	0.00
	10	2920	1808	0.56	0.43	1.00	-0.27	-0.25	0.01	2861	2611	0.61	0.59	0.87	-0.23	-0.22	0.04
RF2	11	1266	1233	0.50	0.50	0.20	-0.21	-0.21	0.35	1341	1331	0.41	0.41	0.18	-0.22	-0.22	0.41
	12	264	205	0.52	0.52	0.20	-0.30	-0.27	0.03	365	354	0.61	0.63	0.13	-0.36	-0.37	0.25
	0	7857	5410	0.72	0.72	0.12	-0.22	-0.21	0.00	7776	6776	0.70	0.70	0.32	-0.21	-0.21	0.00
	1	6172	3458	0.68	0.64	0.14	-0.47	-0.41	0.00	5650	4716	0.75	0.75	0.05	-0.44	-0.40	0.00
	2	6312	4246	0.58	0.59	0.00	-0.34	-0.33	0.33	6039	5475	0.68	0.67	0.13	-0.30	-0.28	0.00
	3	10975	7202	0.69	0.67	0.15	-0.26	-0.24	0.00	10277	8388	0.73	0.73	0.00	-0.23	-0.20	0.00
	4	7672	5301	0.73	0.69	0.91	-0.17	-0.15	0.00	7243	6045	0.77	0.77	0.04	-0.16	-0.14	0.00
	5	4105	2020	0.44	0.40	0.11	-0.42	-0.36	0.00	1612	1313	0.40	0.42	0.09	-0.38	-0.36	0.00
	6	1603	756	0.27	0.40	0.00	-0.39	-0.32	0.00	3974	3145	0.55	0.58	0.00	-0.51	-0.45	0.00
	7	2278	1380	0.70	0.65	0.56	-0.33	-0.31	0.00	5376	3953	0.69	0.68	0.03	-0.27	-0.25	0.00
	8	5206	3006	0.64	0.64	0.00	-0.29	-0.28	0.00	1203	1097	0.73	0.75	0.06	-0.21	-0.20	0.03
	9	1189	667	0.66	0.59	0.93	-0.22	-0.22	0.32	2222	1757	0.80	0.80	0.11	-0.32	-0.31	0.01
	10	2920	1808	0.64	0.51	1.00	-0.22	-0.21	0.03	2861	2611	0.70	0.68	0.66	-0.18	-0.18	0.07
	11	1266	1233	0.60	0.60	0.62	-0.16	-0.16	0.82	1341	1331	0.51	0.51	0.04	-0.17	-0.17	0.13
	12	264	205	0.57	0.55	0.35	-0.28	-0.26	0.03	365	354	0.65	0.68	0.05	-0.34	-0.34	0.30

broadly similar. However, we find two areas characterized by large differences between these seasonal estimates.

The first is located around Qinghai and Gansu provinces (spatially corresponding to cluster 5) that is highly influenced by dust aerosols (SI, Table S1). The larger values over these regions compared to the surrounding regions particularly in the autumn, inferred from the full data, are likely erroneous. This is because the spatial distribution of seasonal ground-level PM<sub>2.5</sub> should be relatively smooth without too many sudden changes. For the same regions, using the reduced data we find there is an obvious reduction of up to 20  $\mu\text{gm}^{-3}$ , alleviating the extent to which the large anomaly occurs. Further, we calculate the MPE between these seasonal estimates and the seasonal observations compiled from daily ground-level PM<sub>2.5</sub> measurements. We find that the corresponding MPE over cluster 5 in the autumn goes from -8% to -5% when the seasonal estimates are inferred from the full to the reduced data. This indicates that the positive model bias occurs frequently over cluster 5 in autumn

months and the seasonal estimates inferred from the reduced data are much closer to observations.

The second is located around Shaanxi, Shanxi, and Henan provinces (spatially corresponding to cluster 2). The ground-level PM<sub>2.5</sub> pollution is severe in winter months, causing negative model bias in most previous studies. We find that the winter mean of ground-level PM<sub>2.5</sub> estimates inferred from the reduced data is higher than those inferred from the full data over cluster 2. Further, we find that the MPE of the winter mean of ground-level PM<sub>2.5</sub> estimates inferred from the full and the reduced data against the observations are 11% and 6%, respectively. This suggests that the negative model bias occurs frequently over cluster 2 in winter months and the seasonal estimates inferred from the reduced data are much closer to observations.



**Fig. 4.** Seasonal ground-level PM<sub>2.5</sub> maps in 2014 over eastern China inferred from the RF2 using the data from the Aqua overpass time at 0.25° (latitude) × 0.3125° (longitude) coincide with observed ground-level PM<sub>2.5</sub> measurements. The first row presents results using the RF2 trained by the full data. The second row presents results using the RF2 trained by the reduced data where  $\Gamma_{PBL}^{AOD} \geq 0.5$ . The third row presents their differences. Two red circles in the last two subplots highlight areas where large differences are located. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

#### 4. Discussion and concluding remarks

We have proposed a model framework to infer ground-level PM<sub>2.5</sub> from satellite observations of AOD. Most importantly, the method highlights the need to take into account the fraction of AOD in the PBL, where variations are most likely to reflect changes in ground-level PM<sub>2.5</sub>. The experiment over eastern China during 2014 shows that the method can effectively reduce bias in inferred estimates of ground-level PM<sub>2.5</sub> and captures more variations in ground-level PM<sub>2.5</sub> particularly over areas dominated by natural aerosols. We attribute the comparative model improvement to the two benefits due to restricting the data analysis to the suitable data based on the  $\Gamma_{PBL}^{AOD} \geq 0.5$  criterion. First, we improve the extent to which AOD is connected to ground-level PM<sub>2.5</sub>. Second, for areas where natural aerosols like dust dominate and are frequently found in the free troposphere, we reduce the disparities in aerosol compositions between AOD and ground-level PM<sub>2.5</sub>. These two aspects jointly improve the representativeness of using AOD as the proxy for ground-level PM<sub>2.5</sub>.

As a result of our criterion we do not have a temporally and spatially

continuous ground-level PM<sub>2.5</sub> field, with gaps left where AOD is not sufficiently representative of aerosol extinction in the PBL. A looser or stricter criterion would retain more or less data, respectively. Our ability to infer ground-level PM<sub>2.5</sub> depends on this criterion retaining sufficient data. It then becomes a trade-off between the quality and quantity of the AOD data being used to infer ground-level PM<sub>2.5</sub> concentrations. Alternatively, the resulting ground-level PM<sub>2.5</sub> estimates inferred from satellite observations of AOD could be incorporated into a data assimilation framework where gaps are filled by a dynamic model of aerosols. In work (not shown) we also explore whether we can improve the model performance by using  $\sum_{PBL}^{AOD} = AOD_{MAIAC} \times \Gamma_{PBL}^{AOD}$  as the main predictor to avoid data loss. However, we have not found any model improvement even on a national scale. Hence, by discarding the data for which  $\Gamma_{PBL}^{AOD} < 0.5$  we likely remove the circumstance that the AOD is not sufficiently mixed in the PBL to reflect changes in ground-level PM<sub>2.5</sub>.

A large quantity of research has developed statistical and machine learning models to infer ground-level PM<sub>2.5</sub> concentrations from satellite observations of AOD and more studies are increasingly emerging. To date, however, limited have investigated the suitability of the

methodology. Through this study we demonstrate that caution should be exercised when relating AOD to ground-level PM<sub>2.5</sub> via data-driven models. Our model framework is sufficiently generic in that it can be applied to other periods and other geographical regions, after taking into account differences in the composition of AOD and ground-level PM<sub>2.5</sub> that will help define the process-driven and data-driven models.

### CRediT authorship contribution statement

**Fei Yao:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition. **Paul I. Palmer:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

### 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.

### Acknowledgements

The work of Fei Yao was supported by the China Scholarships Council/University of Edinburgh Scholarships. This study was funded as part of NERC's support of the National Centre for Earth Observation: PIP was supported by grant number #NE/R016518/1.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atmosenv.2021.118217>.

### References

- 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., 2001. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation. *J. Geophys. Res.: Atmospheres* 106, 23073–23095.
- Boys, B., Martin, R., Van Donkelaar, A., MacDonell, R., Hsu, N., Cooper, M., Yantosca, R., Lu, Z., Streets, D., Zhang, Q., et al., 2014. Fifteen-year global time series of satellite-derived fine particulate matter. *Environ. Sci. Technol.* 48, 11109–11118.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., et al., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl. Acad. Sci. Unit. States Am.* 115, 9592–9597.
- Chatfield, R.B., Sorek-Hamer, M., Esswein, R.F., Lyapustin, A., 2020. Satellite mapping of pm2.5 episodes in the wintertime san joaquin valley: a 'static' model using column water vapor. *Atmos. Chem. Phys.* 20, 4379–4397. <https://doi.org/10.5194/acp-20-4379-2020>. URL: <https://www.atmos-chem-phys.net/20/4379/2020/>.
- European Committee for Standardization (CEN), 1998. Air Quality Determination of the PM10 Fraction of Suspended Particulate Matter Reference Method and Field Test Procedure to Demonstrate Reference Equivalence of Measurement Methods (European Standard EN 12341).
- Guo, J., Wu, Y., Zhang, X., Li, X., 2013. Estimation of pm2. 5 over eastern China from modis aerosol optical depth using the back propagation neural network. *Huan jing ke xue= Huanjing xueke* 34, 817–825.
- Guo, Y., Tang, Q., Gong, D.Y., Zhang, Z., 2017. Estimating ground-level pm2. 5 concentrations in beijing using a satellite-based geographically and temporally weighted regression model. *Remote Sens. Environ.* 198, 140–149.
- Gupta, P., Christopher, S.A., 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. a neural network approach. *J. Geophys. Res.: Atmospheres* 114.
- Hammer, M.S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A.M., Hsu, N.C., Levy, R.C., Garay, M., Kalashnikova, O., Kahn, R.A., et al., 2020. Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018). *Environmental Science & Technology*.
- He, Q., Huang, B., 2018a. Satellite-based high-resolution pm2. 5 estimation over the beijing-tianjin-hebei region of China using an improved geographically and temporally weighted regression model. *Environ. Pollut.* 236, 1027–1037.
- He, Q., Huang, B., 2018b. Satellite-based mapping of daily high-resolution ground pm2. 5 in China via space-time regression modeling. *Remote Sens. Environ.* 206, 72–83.
- HJ 618-2011, 2011. Determination of atmospheric articles PM10 and PM2.5 in ambient air by gravimetric method. URL: [http://english.mee.gov.cn/Resources/standards/Air\\_Environment/air\\_method/20111/W020110914570884643311.pdf](http://english.mee.gov.cn/Resources/standards/Air_Environment/air_method/20111/W020110914570884643311.pdf).
- Holben, B.N., Eck, T.F., Slutsker, I.a., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y., Nakajima, T., et al., 1998. Aeronet—a federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* 66, 1–16.
- Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017. Estimating pm2. 5 concentrations in the conterminous United States using the random forest approach. *Environ. Sci. Technol.* 51, 6936–6944.
- Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes Jr., M.G., Estes, S.M., Quattrochi, D.A., Sarnat, J.A., Liu, Y., 2013. Estimating ground-level pm2. 5 concentrations in the southeastern us using geographically weighted regression. *Environ. Res.* 121, 1–10.
- Hu, X., Waller, L.A., Lyapustin, A., Wang, Y., Al-Hamdan, M.Z., Crosson, W.L., Estes Jr., M.G., Estes, S.M., Quattrochi, D.A., Puttaswamy, S.J., et al., 2014. Estimating ground-level pm2. 5 concentrations in the southeastern United States using maiac aod retrievals and a two-stage model. *Remote Sens. Environ.* 140, 220–232.
- Jiang, Z., Jolleys, M.D., Fu, T.M., Palmer, P.I., Ma, Y., Tian, H., Li, J., Yang, X., 2020. Spatiotemporal and probability variations of surface pm2. 5 over China between 2013 and 2019 and the associated changes in health risks: an integrative observation and model analysis. *Sci. Total Environ.* 137896.
- Kahn, R., Banerjee, P., McDonald, D., Diner, D.J., 1998. Sensitivity of multiangle imaging to aerosol optical depth and to pure-particle size distribution and composition over ocean. *J. Geophys. Res.: Atmospheres* 103, 32195–32213.
- Lee, H., Liu, Y., Coull, B., Schwartz, J., Koutrakis, P., 2011. A novel calibration approach of modis aod data to predict pm 2.5 concentrations. *Atmos. Chem. Phys. Discuss.* 11.
- Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* 525, 367–371.
- Lucchesi, R., 2018. File specification for geos fp. GAMO office note No.4 (version 1.2). URL: [https://gmao.gsfc.nasa.gov/pubs/office\\_notes](https://gmao.gsfc.nasa.gov/pubs/office_notes).
- Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. Modis collection 6 maiac algorithm. *Atmos. Meas. Tech.* 11.
- Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R., Reid, J., 2011. Multiangle implementation of atmospheric correction (maiac): 2. aerosol algorithm. *J. Geophys. Res.: Atmospheres* 116.
- Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level pm2.5 in China using satellite remote sensing. *Environ. Sci. Technol.* 48, 7436–7444.
- Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2016a. Satellite-based spatiotemporal trends in pm2. 5 concentrations: China, 2004–2013. *Environ. Health Perspect.* 124, 184–192.
- Ma, Z., Liu, Y., Zhao, Q., Liu, M., Zhou, Y., Bi, J., 2016b. Satellite-derived high resolution pm2. 5 concentrations in yangtze river delta region of China using improved linear mixed effects model. *Atmos. Environ.* 133, 156–164.
- Seinfeld, J.H., Bretherton, C., Carslaw, K.S., Coe, H., DeMott, P.J., Dunlea, E.J., Feingold, G., Ghan, S., Guenther, A.B., Kahn, R., et al., 2016. Improving our fundamental understanding of the role of aerosol–cloud interactions in the climate system. *Proc. Natl. Acad. Sci. Unit. States Am.* 113, 5781–5790.
- Stephens, G.L., 2005. Cloud feedbacks in the climate system: a critical review. *J. Clim.* 18, 237–273.
- Thiessen, A.H., 1911. Precipitation averages for large areas. *Mon. Weather Rev.* 39, 1082–1089.
- U.S. Environmental Protection Agency, 1997. Reference Method for the Determination of Fine Particulate Matter as PM2.5 in the Atmosphere, 40CFR50, Appendix L.
- Van Donkelaar, A., Martin, R.V., Brauer, M., Boys, B.L., 2015. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environ. Health Perspect.* 123, 135–143.
- Van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verdúzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.* 118, 847–855.
- Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019. Estimating 1-km-resolution pm2. 5 concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* 231, 111221.
- Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., et al., 2020. Improved 1 km resolution pm2. 5 estimates across China using enhanced space-time extremely randomized trees. *Atmos. Chem. Phys.* 20,

- Wu, Y., Guo, J., Zhang, X., Tian, X., Zhang, J., Wang, Y., Duan, J., Li, X., 2012. Synergy of satellite and ground based observations in estimation of particulate matter in eastern China. *Sci. Total Environ.* 433, 20–30.
- Xiao, Q., Chang, H.H., 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.
- Yao, F., Si, M., Li, W., Wu, J., 2018. A multidimensional comparison between modis and viirs aod in estimating ground-level pm<sub>2.5</sub> concentrations over a heavily polluted region in China. *Sci. Total Environ.* 618, 819–828.
- Yao, F., Wu, J., Li, W., Peng, J., 2019. A spatially structured adaptive two-stage model for retrieving ground-level pm<sub>2.5</sub> concentrations from viirs aod in China. *ISPRS J. Photogrammetry Remote Sens.* 151, 263–276.
- Zhang, T., Zhu, Z., Gong, W., Zhu, Z., Sun, K., Wang, L., Huang, Y., Mao, F., Shen, H., Li, Z., et al., 2018. Estimation of ultrahigh resolution pm<sub>2.5</sub> concentrations in urban areas using 160 m gaofen-1 aod retrievals. *Remote Sens. Environ.* 216, 91–104.