Science of the Total Environment 658 (2019) 1256–1264

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Estimation of spatiotemporal $PM_{1.0}$ distributions in China by combining $PM_{2.5}$ observations with satellite aerosol optical depth

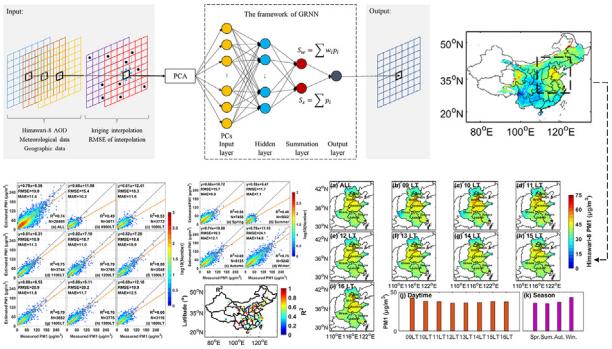
Lin Zang [a,b], Feiyue Mao [b,c,d,*], Jianping Guo [e], Wei Wang [f], Zengxin Pan [b], Huanfeng Shen [g], Bo Zhu [h], Zemin Wang [a]

[a] Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[c] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[d] Collaborative Innovation Center for Geospatial Technology, Wuhan 430079, China
[e] State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China
[f] School of Geosciences and Info-Physics, Central South University, Changsha 410083, China
[g] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[h] Hubei Environmental Monitoring Center, Wuhan 430079, China

HIGHLIGHTS

- Fusing $PM_{2.5}$ observations is significant to improve estimation accuracy of $PM_{1.0}$.
- Estimation performance shows significant daily, seasonal and spatial variations.
- Estimation accuracy largely dependences on quality of AOD and interpolated $PM_{2.5}$

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 27 September 2018
Received in revised form 19 December 2018
Accepted 20 December 2018
Available online 21 December 2018

Editor: Jianmin Chen

Keywords:
Aerosol optical depth
Air pollution
Himawari-8
Neural network
$PM_{1.0}$

ABSTRACT

Particulates smaller than 1.0 μm ($PM_{1.0}$) have strong associations with public health and environment, and considerable exposure data should be obtained to understand the actual environmental burden. This study presented a $PM_{1.0}$ estimation strategy based on the generalised regression neural network model. The proposed strategy combined ground-based observations of $PM_{2.5}$ and satellite-derived aerosol optical depth (AOD) to estimate $PM_{1.0}$ concentrations in China from July 2015 to June 2017. Results indicated that the $PM_{1.0}$ estimates agreed well with the ground-based measurements with an R^2 of 0.74, root mean square error of 19.0 μg/m^3 and mean absolute error of 11.4 μg/m^3 as calculated with the tenfold cross-validation method. The diurnal estimation performance displayed remarkable single-peak variation with the highest R^2 of 0.80 at noon, and the seasonal estimation performance showed that the proposed method could effectively capture high-pollution events of $PM_{1.0}$ in winter. Spatially, the most polluted areas were clustered in the North China Plain, where the average estimates presented a bimodal distribution during daytime. In addition, the quality of satellite-derived AOD, the robustness of the interpolation algorithm and the proportion of $PM_{1.0}$ in $PM_{2.5}$ were confirmed to affect the estimation accuracy of the proposed model.

* Corresponding author at: School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China.
E-mail address: maofeiyue@whu.edu.cn (F. Mao).

https://doi.org/10.1016/j.scitotenv.2018.12.297
0048-9697/

1. Introduction

As the major components of aerosols, particulate matter (PM) exert remarkable impacts on the climate system through their direct and indirect effects (Qin et al., 2017; Wang et al., 2015). The report of IPCC5 indicates that, the impact of aerosol–cloud interaction on climate remains the most uncertain factor in global radiation changes (IPCC, 2013). Many studies on aerosol climatic effects have been conducted (Charlson et al., 1992; Pan et al., 2017; Lu et al., 2018; Mao et al., 2018). Large amounts of particle emissions seriously reduce atmospheric visibility and affect daily traffic. Moreover, particles smaller than 10 μm are inhalable by humans and threaten human health (Bartell et al., 2013; Kulshrestha, 2018).

Studies on the regional spatiotemporal evolution of PM with different diameters acquired through station measurements have indicated that fine particles, especially $PM_{1.0}$, are the major contributor to air pollution (Shi et al., 2014; Zhuang et al., 2014). Statistically, $PM_{1.0}$ accounts for more than 80% of the average $PM_{2.5}$ and even reaches 98% in several regions of China (Li et al., 2015). $PM_{1.0}$ data is more useful than $PM_{2.5}$ data in assessing anthropogenic emissions. Morawska et al. (2008) investigated the size distribution of ambient particles and found that $PM_{1.0}$ enables improved distinction between combustion and mechanically generated aerosols. Furthermore, the health hazards of PM increase with the decrease in particulate diameter (Huang et al., 2003; Elder et al., 2015) because fine particles contain considerable harmful substances and could enter the respiratory system, especially the alveoli, deeply, thereby causing serious threats to humans (Zhang et al., 2016).

Given the strong associations of $PM_{1.0}$ with public health and the environment, large amounts of exposure data should be obtained to strengthen particle pollution control. However, $PM_{1.0}$ monitoring stations are sparse and unevenly distributed, especially in sparsely populated western parts of China; this feature is not conducive for conducting continuous spatiotemporal monitoring of ultrafine particulates. Advanced satellite remote sensing with wide observation horizons can effectively compensate for limitations in ground measurements (Guo et al., 2009). Studies have demonstrated that aerosol extinction is closely related to particles with diameters between 0.4 and 1.0 μm (Madronich and Flocke, 1999). Therefore, satellite-derived aerosol optical depth (AOD), which is the column integral of aerosol's extinction coefficient, could be a useful tool for mapping $PM_{1.0}$ distributions in theory. Current studies on $PM_{1.0}$ estimation with satellite remote sensing are only in the exploratory stage. Chen et al. (2018) investigated $PM_{1.0}$ estimation on the basis of moderate resolution imaging spectroradiometer (MODIS) AOD in China and obtained 59% variability for daily estimations. Qin et al. (2018) estimated $PM_{1.0}$ concentrations from MODIS over Yangtze River Delta, China, by using a geographically and temporally weighted regression model with R^2 of 0.74 for daily estimations. However, hourly $PM_{1.0}$ concentration with high spatial coverage is needed when evaluating a diurnal transport and control approach of air pollution. As a geostationary satellite, Himawari-8 provides an unprecedented opportunity. Our previous work (Zang et al., 2018) showed that the limited number of $PM_{1.0}$ sites and insufficient ground-based observations cannot effectively constrain the estimated results of $PM_{1.0}$ (R^2 of 0.65), and model accuracy needs further improvement.

In this study, considering the high correlation between $PM_{1.0}$ and $PM_{2.5}$ and large amount of available observation data on $PM_{2.5}$, we estimated $PM_{1.0}$ by combining ground-based $PM_{2.5}$ measurements with hourly satellite aerosol retrievals in China into an improved generalised regression neural network (GRNN) model (hybrid GRNN model with the principal component analysis [PCA] method, namely, PCA–integrated GRNN model). We designed three groups of comparative experiments, which consisted of multiple linear regression (MLR), GRNN and PCA-integrated GRNN models with or without $PM_{2.5}$ observations. We compared the tenfold cross-validation (CV) results of these models by fusing particle observations to evaluate the improvement in each model. In addition, we estimated hourly $PM_{1.0}$ concentrations by using a well-trained model and depicted the spatial distribution of annual and seasonal averages of $PM_{1.0}$. The error sources of estimates were discussed and analysed for future improvements of the proposed method.

2. Datasets and methods

2.1. Datasets

The datasets employed in the $PM_{1.0}$ estimation are summarised in Table 1.

2.1.1. PM observations

Hourly $PM_{2.5}$ observations were obtained from the monitoring network operated by the China National Environmental Monitoring Center (CNEMC). This network has been updated since its establishment in 2013, and the number of monitoring stations has increased from 500 to more than 1400. Observation samples from 1430 matched sites were collected in this study. Unlike $PM_{2.5}$, $PM_{1.0}$ has not been used as a routine indicator of air monitoring, and the stations of $PM_{1.0}$ is much less than those of $PM_{2.5}$. In this study, $PM_{1.0}$ concentration data were collected from 73 matched sites of the monitoring network operated by the China Meteorological Administration (CMA). The station distributions of the two monitoring networks are shown in Fig. S1.

2.1.2. AOD data

Hourly AOD data were obtained from Himawari-8, which was launched in October 2014 and is operated by the Japan Aerospace Exploration Agency (JAXA). The sensor on this satellite is called Advanced Himawari Imager, which has 16 bands from visible to infrared wavelengths with three different spatial resolutions of 0.005°, 0.01° and 0.02° (Kikuchi et al., 2018). Two types of AOD products of Himawari-8 have been released, namely, L2 products with a temporal resolution of 10 min and L3 products with temporal resolutions of 1 h, 1 day and 1 month. All aerosol products have the same spatial resolution of 0.05°. Fukuda et al. (2013) and Higurashi (1999) examined the retrieval

Table 1
Datasets employed in the $PM_{1.0}$ estimation.

Dataset	Variable	Unit	Height	Temporal resolution	Spatial resolution	Data source
$PM_{2.5}$	$PM_{2.5}$	$\mu g/m^3$	Ground-level	1 h	In-situ	CNEMC
$PM_{1.0}$	$PM_{1.0}$	$\mu g/m^3$		5 min	In-situ	CMA
AOD	AOD	–	Column- integral	1 h	0.05°	Himawari-8
Meteorological data	TEMP	K	2 m	6 h	0.125°	ECMWF
	RH	%	Surface	6 h	0.125°	
	SP	Pa	Surface	6 h	0.125°	
	WS	m/s	10 m	6 h	0.125°	
	BLH	m	–	3 h	0.125°	
Geographic data	NDVI	–	–	16 days	0.05°	MODIS C6
	DEM	m	–	–	90 m	USGS

algorithm of Himawari-8 AOD. Satellite AOD data were compared with ground-based measurements from the aerosol robotic network (AERONET) in a previous study (Zang et al., 2018), and the results indicated that the two AOD data have good consistency with a correlation of 0.74 and uncertainty of 0.24.

2.1.3. Meteorological and geographic data

Meteorological reanalysis datasets were obtained from the ERA-Interim Project, which is operated by the European Centre for Medium-Range Weather Forecasts (ECMWF). This data product has been in service since 1979. On the basis of existing publications (Li et al., 2017a; Wang et al., 2017; He and Huang, 2018), five meteorological parameters, namely, near-surface air temperature (TEMP), relative humidity (RH), wind speed (WS), surface pressure (SP) and height of the planetary boundary layer (BLH), were selected in this study. The effects of topography and vegetation coverage on the distribution of suspended particulates were considered. DEM data obtained from the United States Geological Survey (USGS) and NDVI data derived from MODIS were used to describe topographic relief and vegetation coverage, respectively.

2.2. Methods

Affected by many explanatory variables, such as RH and BLH, the relationship between PM concentration and its predictors tends to be nonlinear and is thus difficult to be expressed by a simple function; for this reason, artificial neural network (ANN) models have been widely used to estimate or predict particle concentrations (Wu et al., 2012; Guo et al., 2013; Li et al., 2017a). For example, the GRNN model is well known for its outstanding self-organising, self-learning and adaptive functions and its high-accuracy prediction of PMs (Li et al., 2017a). In this study, we used GRNN as the base model to depict the spatiotemporal distributions of $PM_{1.0}$ in China. The implementation is shown in Fig. 1. Specifically, to fuse ground-based particle measurements, we adopted the Kriging method and obtained spatial interpolations of $PM_{2.5}$. Then, all input variables, including AOD, interpolated $PM_{2.5}$, meteorological and geographic parameters, were recombined using the PCA method. The updated input variables were introduced into the GRNN model to train the network and estimate $PM_{1.0}$ concentrations with the well-trained model.

2.2.1. Interpolation of $PM_{2.5}$

As a geographical variable, PM concentration has a spatial correlation and satisfies Tobler's first law of geography, which posits that near things are more related than distant things. In this study, the Kriging interpolation algorithm was used to obtain interpolations of ground-based $PM_{2.5}$, which is generally considered to be optimal in geoscience applications due to its unbiased interpolations and small variances (Oliver and Webster, 1990). Kriging method is essentially based on local weighted averaging, of which the weights are determined by a variogram, the core of regionalised variable theory that depends on the size and distribution of samples (Matheron, 1971). In addition, confidence of this algorithm can be calculated and inputted into the estimates because the estimation variances (i.e. theoretical root mean square error [RMSE] of interpolation) can be mapped like interpolations (Oliver and Webster, 1990). This feature differentiates kriging from other interpolators and makes it more sensitive.

The interpolates with theoretical RMSE (i.e. interpolation variance) greater than the 99th percentile of the average theoretical RMSE in the study domain were rejected to control interpolation accuracy, and the interpolated value and its theoretical RMSE were used as the original input parameters of the proposed model.

2.2.2. Parameter recombination

The proposed method requires four types of explanatory variables, namely, AOD, meteorological and geographical parameters and interpolation results of ground-based particle observations. These data, especially meteorological parameters, show a degree of collinearity, which may lead to model overfitting and may reduce estimation accuracy. Therefore, the PCA method, which can solve the problem of multicollinearity by converting the original predictors into a new set of independent variables (i.e. principal components or PCs) via orthogonal transformation, was adopted in this study. Sun and Sun (2016) provided a detailed algorithm introduction. The cumulative explained variance of the selected PCs should not be less than 0.95. These independent variables obtained through PCA were transmitted to the designed neural network as input nodes.

2.2.3. $PM_{1.0}$ estimation

The GRNN model was utilised to estimate $PM_{1.0}$ concentrations. This neural network is a normalised radial basis function (RBF) network proposed and developed by Dr. Specht (Specht, 1991). It performs nonparametric estimation of the Parzen window with samples as the posterior condition on the basis of nonparametric regression theory and calculates the network output according to the principle of maximum probability.

The network architecture of GRNN contains four layers of neurons, namely, input, hidden RBF, summation and output layers (Fig. 1). Each neuron of the input layer represents an input variable, and this layer is usually connected with the hidden RBF layer by a density function. The number of hidden neurons is determined by training samples. In particular, the output of the hidden layer in GRNN is not directly

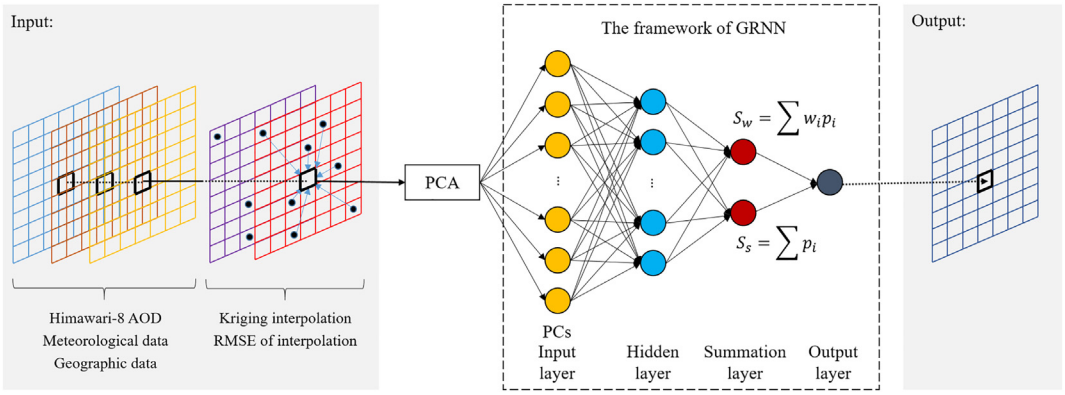


Fig. 1. Implementation of $PM_{1.0}$ estimation based on the Kriging and PCA–integrated GRNN model.

Table 2
Performance of six $PM_{1.0}$ estimation models in three groups of comparative experiments assessed through tenfold CV.

Model	Predictors	R^2	RMSE	MAE	AIC
MLR	Without $PM_{2.5}$	0.21	33.3	23.0	280,596.4
	With $PM_{2.5}$	0.38	29.3	20.0	273,421.2
GRNN	Without $PM_{2.5}$	0.62	23.6	14.1	261,103.8
	With $PM_{2.5}$	0.67	22.0	13.0	257,074.6
PCA–integrated GRNN	Without $PM_{2.5}$	0.68	21.4	12.9	253,390.1
	With $PM_{2.5}$	0.74	19.0	11.4	252,773.2

connected to the output layer by a linear function but by a summation layer; in this summation layer, two summations, i.e., S_s and S_w, are calculated, and weights w are related to hidden outputs P (Specht, 1996; Zhou et al., 2014). The final output of the GRNN model can be presented as

$$output = {}^{S_w}/_{S_s}.$$

Compared with other common neural network models, GRNN is easier to train without the requirement of pre-setting the hidden layer and demonstrates better nonlinear approximation performance. In this study, nearly 30,000 samples were used for model training and simulation.

2.2.4. Accuracy evaluation

We designed three groups of comparative experiments, namely, MLR (Pawan and Christopher, 2009), GRNN and PCA-integrated GRNN models with or without $PM_{2.5}$ observations, to verify whether fusing ground-based $PM_{2.5}$ observations enhances $PM_{1.0}$ estimation accuracy or not. Tenfold CV and four statistical indexes, namely, determination coefficient (i.e. R^2), RMSE, mean absolute error (MAE) and Akaike information criterion (AIC), were adopted in this study. The selection of evaluation indexes was performed in reference to Wang et al. (2017).

3. Results and discussion

3.1. Evaluation of model performance

Table 2 summarises the $PM_{1.0}$ estimation performance of each model from three groups of comparative experiments. Comparisons of the first and second experimental groups indicated that the nonlinear ANN model could better learn the potential relationship between $PM_{1.0}$ and its predictors, and its estimation accuracy was improved by nearly twice. Specifically, the values of R^2, RMSE and MAE of the MLR model without $PM_{2.5}$ observations were 0.21, 33.3 μg/m^3 and 23.0 μg/m^3, respectively, whereas the GRNN model without $PM_{2.5}$ observations presented a much higher estimation accuracy with R^2 of 0.62, RMSE of 23.6 μg/m^3 and MAE of 14.1 μg/m^3, respectively. When the PCA method was integrated into the GRNN model, the accuracy of $PM_{1.0}$ estimation was further improved with R^2 increasing to 0.68, RMSE decreasing to 21.4 μg/m^3 and MAE decreasing to 12.9 μg/m^3, respectively, as shown in the third experimental group.

Comparisons within each experimental group showed that the estimation accuracy of each model was remarkably improved after fusing $PM_{2.5}$ observations. Specifically, the R^2 of the MLR model increased by nearly twice, and those of GRNN and PCA–integrated GRNN models improved by 8% and 9%, respectively. The values of R^2, RMSE and MAE of the PCA–integrated GRNN model fusing $PM_{2.5}$ observations were 0.74, 19.0 μg/m^3 and 11.4 μg/m^3, respectively. In addition, the proposed method had the lowest value of AIC among all the compared models.

Fig. 2 presents the estimation performance of the proposed method at different hours during daytime (09 to 16 local times). The values of R^2 between the estimated and measured $PM_{1.0}$ ranged from 0.49 to 0.80, RMSEs varied between 15.4 and 20.9 μg/m^3 and MAEs ranged from

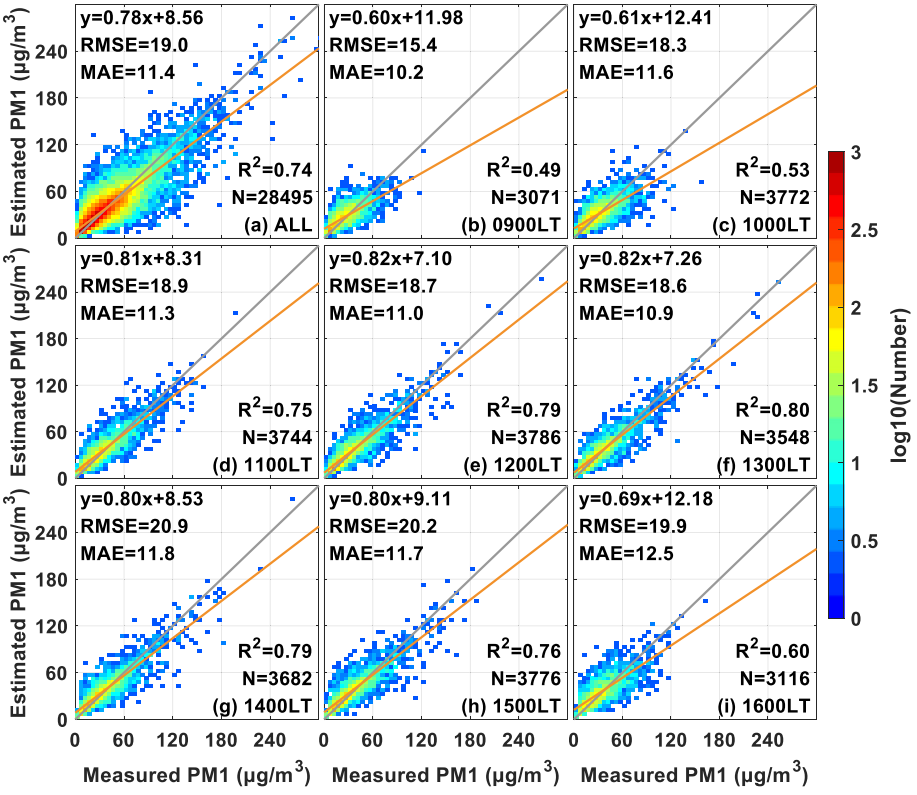


Fig. 2. Scatter plots of $PM_{1.0}$ estimations versus $PM_{1.0}$ measurements during daytime: (a) all available data and (b)–(i) different hours during daytime (09–16 LTs). 'N' represents the sample size in each subplot. Colours represent the logarithm of data size for a corresponding pixel. The grey lines represent 1:1 reference lines.

11.6 μg/m^3 to 12.5 μg/m^3. The estimation accuracy displayed a single-peak variation that peaked between 12 and 14 local times (LTs).

Comparisons between Himawari-8 and AERONET-observed AODs showed that the aerosol detection accuracy of Himawari-8 presented a remarkable diurnal variation with a single peak between 12 and 14 LTs (Zang et al., 2018). This finding proves that the estimation accuracy of $PM_{1.0}$ is affected by the quality of satellite AOD data, and this finding is consistent with the conclusion of Qin et al. (2018). High temperatures at noon cause the planetary boundary layer to develop upward and promote the mixing of particulates in the vertical direction (Guo, et al., 2016; Li, et al., 2017b), which is helpful to the retrieval of fine particulate concentrations. Moreover, the average proportion of $PM_{1.0}$ in $PM_{2.5}$ was slightly higher between 12 and 15 LTs (Table S1), which is conducive for enhancing the estimation accuracy of $PM_{1.0}$ over the same period based on the proposed model.

The estimation performance in different seasons was evaluated to assess the capability of the proposed method to capture heavy-pollution incidents (Fig. 3). The statistical results indicated that the highest seasonal estimation accuracy of $PM_{1.0}$ occurred in winter with R^2 of 0.75, whereas the estimation accuracy in summer was the lowest with R^2 of 0.48 only. The R^2 values in spring and autumn were 0.58 and 0.69, respectively. RMSE and MAE were higher because the absolute concentrations of particles are higher in autumn and winter than in spring and summer.

The step-by-step $PM_{1.0}$ estimation performance for model validation with the addition of each type of variables (AOD, meteorological parameters, geological parameters and interpolated $PM_{2.5}$) is shown in Table S3. The results showed that estimation accuracy of the model increased with the introduction of variables (R^2 from 0.44 to 0.74), and the meteorological parameters also exerted a remarkable impact on model performance. Fig. 4 presents the correlation between each parameter and $PM_{1.0}$ in different seasons. Specially, $PM_{2.5}$ and AOD were the two variables with the highest correlation with $PM_{1.0}$, and meteorological parameters showed different impacts in different season, among which, wind speed was the most effective meteorological parameter that influenced estimation accuracy of the model, and the correlation between BLH and $PM_{1.0}$ in autumn and winter was much higher than that in spring and summer. Overall, comparisons among four seasons showed that the correlation between input parameters and $PM_{1.0}$ was the lowest in summer and the highest in winter.

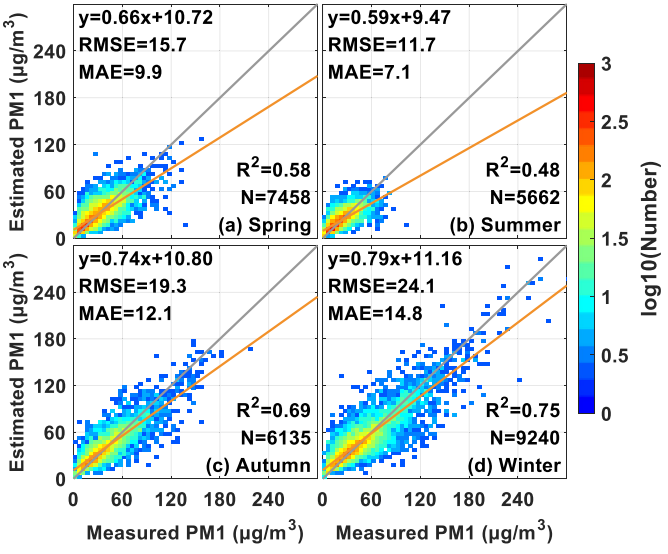


Fig. 3. Scatter plots of $PM_{1.0}$ estimations versus $PM_{1.0}$ measurements for different seasons: (a) spring, (b) summer, (c) autumn and (d) winter. The grey lines represent 1:1 reference lines.

To investigate the cause of the different correlations between parameters and $PM_{1.0}$ in summer and winter, we further analysed the correlations between $PM_{2.5}$ and gas pollutants (SO_2, NO_2 and O_3) to explore the different sources of fine particles in the two seasons because no synchronous observations of $PM_{1.0}$ and gas pollutant are available. Fig. 5 indicates that the concentration of $PM_{2.5}$ was closely related to that of gas pollutants. The high correlation coefficient between $PM_{2.5}$ and O_3 in summer indicated that the active photochemical reaction remarkably accelerated the formation of secondary fine particles in this period, as reported by Wen et al. (2015). This finding is also consistent with the conclusions of Guo et al. (2011). The correlation coefficient between $PM_{2.5}$ and NO_2 and SO_2 was high in winter, indicating that $PM_{2.5}$ in winter is closely related to anthropogenic combustion emissions. Furthermore, different sources of aerosols could lead to different proportions of $PM_{1.0}$ in $PM_{2.5}$. In this study, the average proportion of $PM_{1.0}$ in $PM_{2.5}$ in summer (winter) was 83% (86%), and the Pearson correlation coefficient between them is 0.88 (0.94). In addition, there were considerable anomalies of PM observations in summer. An example is that there were more samples with $PM_{1.0}$ concentration greater than $PM_{2.5}$ in summer, as shown in Fig. 5-(s–d). All of these factors caused the different correlations between variables and $PM_{1.0}$ in summer and winter, as well as the different estimation performance based on the proposed model.

3.2. Mapping of $PM_{1.0}$ estimates

Fig. 6 displays the spatial distributions of the estimated $PM_{1.0}$ in China. Spatially, the average $PM_{1.0}$ concentration in the east is higher than that in the west, and the $PM_{1.0}$ pollution level is higher in inland regions than in south-eastern coasts. In particular, the highly polluted zones are clustered in the North China Plain. The total average of $PM_{1.0}$ estimates over the mainland of China was 32.5 μg/m^3, which is slightly smaller than the average of ground-based observations (40.2 μg/m^3, as shown in Table S2). This result was obtained because the former represents the average pollution level of $PM_{1.0}$ in the entire study area, but the mean value of ground-based observations from uneven stations mainly represents the average pollution level of urban areas, especially in central and eastern China, which are the most polluted areas (He and Huang, 2018). The estimation error also affected this result. Temporally, $PM_{1.0}$ concentrations showed a remarkable seasonal variation, with the highest pollution level in winter (>40 μg/m^3) and the lowest pollution level in summer (<30 μg/m^3). The levels of the estimated $PM_{1.0}$ were similar in spring and autumn, with mean values of 30.5 and 31.0 μg/m^3, respectively.

Fig. 7 shows a zoomed-in map of the $PM_{1.0}$ distribution in the North China Plain. The total average of $PM_{1.0}$ estimates in this region was 36.4 μg/m^3, and the averages ranged from 23 μg/m^3 to 53 μg/m^3 for different subareas, showing remarkable regional variation. In addition, the average $PM_{1.0}$ concentrations during daytime (09–16 LTs) exhibited a bimodal distribution with peaks of 42.5 and 38 μg/m^3 in the morning

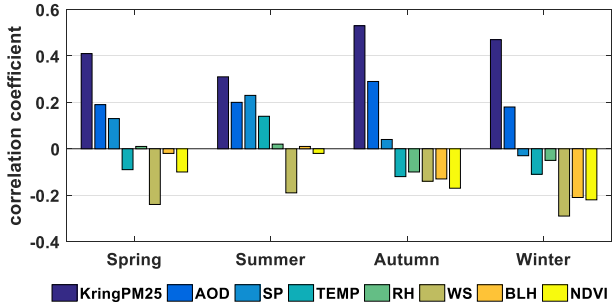


Fig. 4. Correlation between $PM_{1.0}$ and each parameter in different seasons.

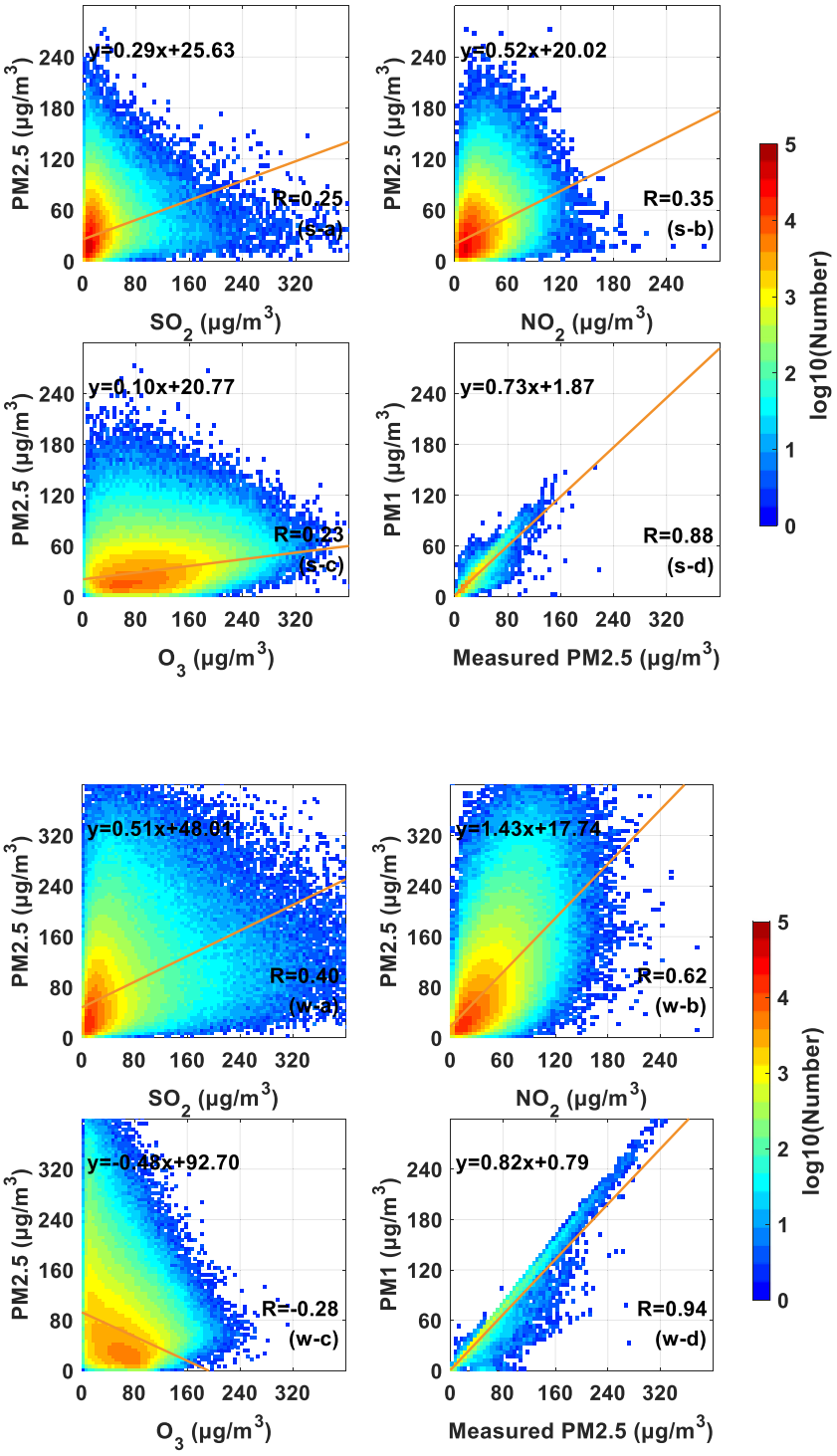


Fig. 5. Correlation between $PM_{2.5}$ and gas pollutants and proportion of $PM_{1.0}$ in summer ((s–a)–(s–d)) and winter ((w–a)–(w–d)).

and afternoon, respectively, which fits well with rush hours. The $PM_{1.0}$ pollution in winter was the heaviest among all seasons in this region (mean concentration of 43.8 μg/m^3). According to existing studies, the heavy pollution in the North China Plain is mainly due to human activities related to rapid economic development and accelerated urbanisation. Moreover, heavy pollution can be aggravated by the adverse topography of Taihang Mountains located in the west of Hebei Province; these mountains confine the pollutants in local areas (Tao et al., 2012; Fu et al., 2014; He and Huang, 2018).

3.3. Analysis of error sources of $PM_{1.0}$ estimates

Comparisons of the three groups of comparative experiments showed that the proposed method outperformed other methods in investigating the relationship between $PM_{1.0}$ and its predictors. The spatial accuracy of the estimated $PM_{1.0}$ is shown in Fig. 8. Specifically, the average R^2 and RMSE for all sites were 0.6 and 17.4 ± 6.6 μg/m^3, respectively. About half of the sites (47%) had a mean MAE of <10 μg/m^3, and the proportion of sites with a mean relative error (RE, i.e. ratio of

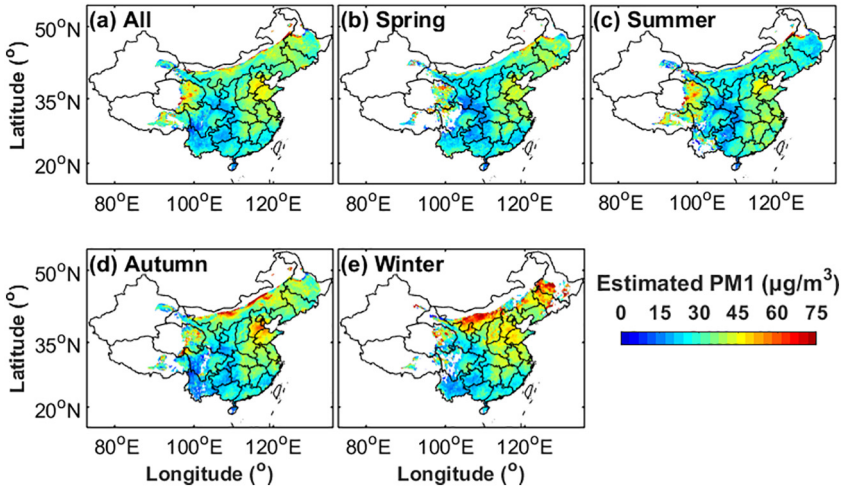


Fig. 6. Spatial distributions of the average $PM_{1.0}$ estimates in China: (a) all available data and (b)–(i) different seasons.

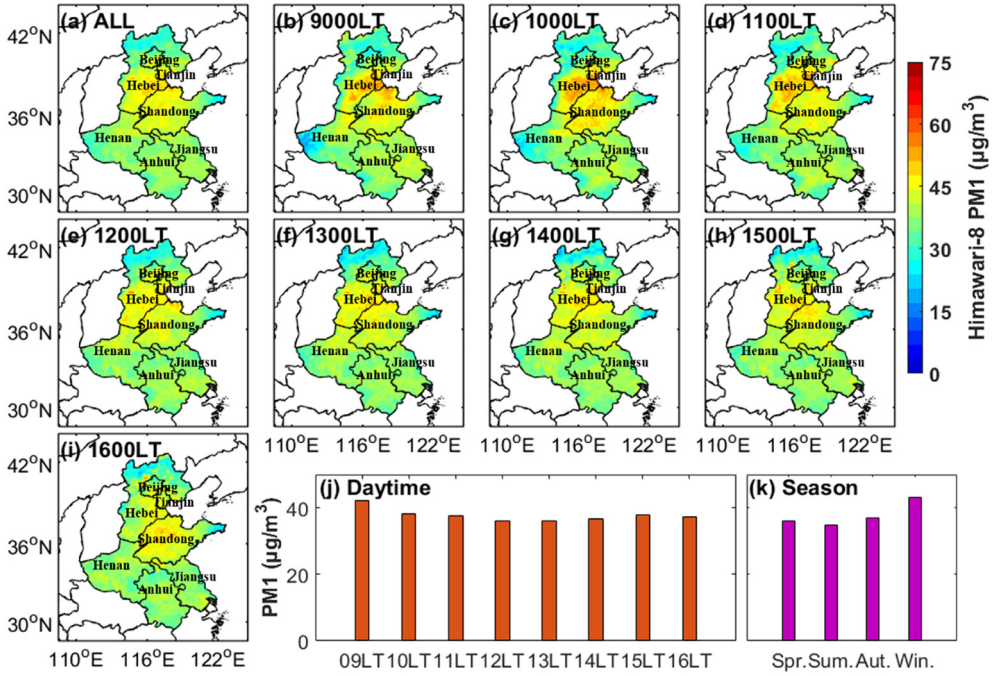


Fig. 7. Zoomed-in map and variations of the estimated $PM_{1.0}$ during daytime (09–16 LTs) and different seasons in the North China Plain.

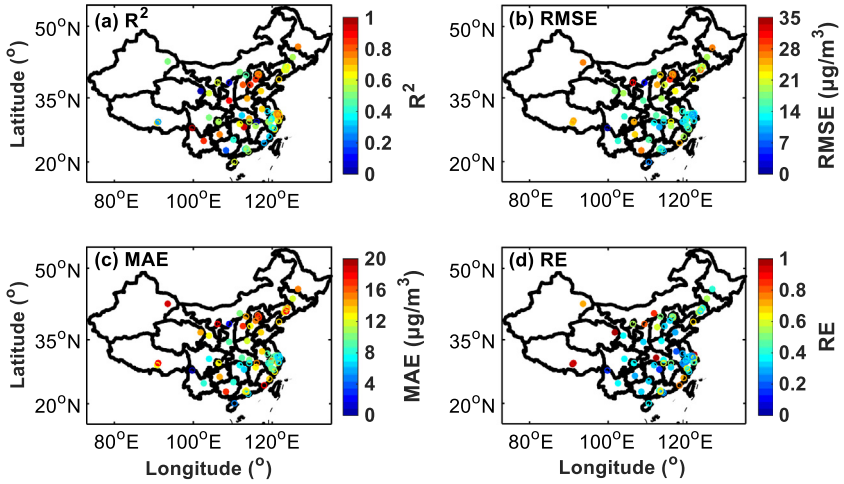


Fig. 8. Summary of site-oriented estimation performance: (a) average of R^2, (b) average of RMSE, (c) average of MAE and (d) average of RE.

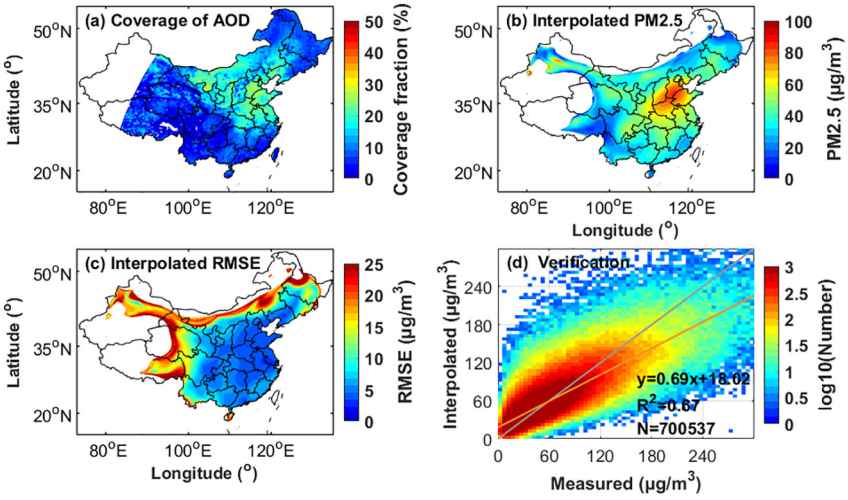


Fig. 9. Error source analysis for $PM_{1.0}$ estimations: (a) coverage of AOD; (b) interpolation of $PM_{2.5}$; (c) interpolated RMSE of $PM_{2.5}$; (d) interpolated $PM_{2.5}$ versus observed $PM_{2.5}$. The grey lines represent 1:1 reference lines.

absolute error to true value) less than 0.4 was 60%. Although the average spatial accuracy was relatively high, the estimation performance presented a large regional variability. Specifically, the values of R^2 ranged from 0.16 to 0.90, RMSEs ranged from 0.4 μg/m^3 to 31.6 μg/m^3, MAEs ranged from 0.3 μg/m^3 to 19.4 μg/m^3 and REs ranged from 0.14 to 3.3. As shown in Fig. 7(a) and (d), the estimation accuracy in central–eastern China was higher than that in the northwest, west and southeast coasts of China due to three main reasons.

The first reason is related to the retrieval of satellite AOD. According to Guo et al. (2017), most parts of western and northwestern China are arid, and the surface albedos are high. This scenario leads to high uncertainty of AOD retrievals and large biases in the estimation of $PM_{1.0}$ in these regions. The poor performance in coastal areas could be partially attributed to the difficulties in addressing aerosol types and surface albedos in the AOD inversion algorithm (Donkelaar et al., 2006; Anderson et al., 2012).

The second reason is related to the valid cover scope of satellite AOD. Fig. 9(a) presents the coverage of Himawari-8 AOD data used here. The available AOD data were mainly located over the central and eastern regions of China with a maximum coverage of less than 45%. The low coverage of AOD data indicates that $PM_{1.0}$ estimation cannot be performed at all times and for the entire China. In addition, AOD, as the constraint of spatiotemporal matching of all data, limits the sample size of model training. Having only a few matchups causes model overfitting and reduces estimation accuracy, and overfitting could be alleviated by increasing the number of matchups (He and Huang, 2018; Zhang et al., 2018). In this study, the training samples were mainly concentrated in the central and eastern regions, which led to a low estimation accuracy in the other regions.

The third reason is related to the interpolation deviations of the Kriging method, as shown in Fig. 9(b) to (d). Despite the high consistency between the interpolated and measured $PM_{2.5}$ with R^2 of 0.67, large interpolation deviations were presented on the interpolated edge, with a theoretical RMSE of >15 μg/m^3, due to the extremely sparse observation stations. This condition also affected the $PM_{1.0}$ estimation, especially in West and North China.

4. Conclusions

Fine PM poses a greater health risk than coarse particulates (PM_{10}) (Oberdörster et al., 2005), and the extinction coefficient increases with the increase in $PM_{1.0}$ fraction (Sabbagh-Kupelwieser, 2010). Despite the abundance of studies on $PM_{2.5}$ and PM_{10} estimations from satellite remote sensing, only a few studies have been conducted on $PM_{1.0}$ by using satellite observations. In this study, we estimated hourly $PM_{1.0}$ concentrations in China by using a PCA–integrated GRNN model that combines ground-based observations of $PM_{2.5}$ with Himawari-8 AOD data. The main findings are as follows:

(1) Fusing $PM_{2.5}$ observations was beneficial for the continuous spatiotemporal estimation of $PM_{1.0}$, and the estimation accuracy of each model was remarkably improved. Specifically, the R^2 of MLR increased from 0.21 to 0.38, and those of GRNN and PCA-integrated GRNN models improved by 8% and 6%, respectively. Comparisons of linear regression and GRNN (including PCA–integrated GRNN) models showed that the nonlinear model could determine the potential relationship between PM and its predictors.

(2) The estimation accuracies presented remarkable daily and seasonal variations. During daytime, the estimation performance displayed a single-peak variation, with the highest R^2 of 0.80 at 13 LT. This result is largely due to the fact that the aerosol observation capability of Himawari-8 presents a remarkable diurnal variation with a single peak between 12 and 14 LTs and more favourable weather conditions. For the seasonal estimation performance, the average accuracy in summer was the lowest with an R^2 of 0.48 due to the low correlation between $PM_{1.0}$ and each input parameter, which can lead to large biases in $PM_{1.0}$ estimation.

(3) The estimation performance showed large regional variability with high accuracy in central and eastern China and low accuracy in the northwest and southeast coasts of China. After analysis, the estimation errors were determined to be caused by three factors, namely, uncertainty of AOD retrievals, extremely low coverage of AOD and instability of Kriging interpolation due to sparse available stations.

To sum up, estimation of ground-level $PM_{1.0}$ concentrations from satellite AOD largely depends on the quality of satellite data. Therefore, the quality of Himawari-8 AOD should be improved to enhance the estimation accuracy. Advanced nonlinear learning algorithms are also vital in fusing many particle measurements. Moreover, the different sources of $PM_{1.0}$, their proportion in $PM_{2.5}$, and the influence of meteorological parameters on particle concentrations in different seasons, which are not fused into the current model, could be non-negligible uncertainty sources in $PM_{1.0}$ estimation with satellite remote sensing. The model should be improved further in future studies.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

This study was supported by the National Key Research and Development Program of China (grant numbers 2017YFC0212600, 2016YFC0200900 and 2017YFC1501401) and the National Natural Science Foundation of China (grant numbers 41701381 and 41627804). The authors are also grateful to CMA, Japan Aerospace Exploration Agency, ECMWF, Data Center of the US NASA, and USGS for providing the datasets used in this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2018.12.297.

References

Anderson, J.C., Wang, J., Zeng, J., Petrenko, M., 2012. Accuracy assessment of Aqua-MODIS aerosol optical depth over coastal regions: importance of quality flag and sea surface wind speed. Atmos. Meas. Tech. Discuss. 5 (4), 5205–5243.

Bartell, S.M., Longhurst, J., Tjoa, T., Sioutas, C., Delfino, R.J., 2013. Particulate air pollution, ambulatory heart rate variability, and cardiac arrhythmia in retirement community residents with coronary artery disease. Environ. Health Perspect. 121 (10), 1135.

Charlson, R.J., Schwartz, S.E., Hales, J.M., Cess, R.D., Coakley, J.A., Hansen, J.E., Hofmann, D.J., 1992. ClimateForcing by antlropogenic aerosols. Science 255, 423–430.

Chen, G., Knibbs, L.D., Zhang, W., Li, S., Cao, W., Guo, J., Ren, H., Wang, B., Wang, H., Williams, G., 2018. Estimating spatiotemporal distribution of PM1 concentrations in China with satellite remote sensing, meteorology, and land use information. Environ. Pollut. 233, 1086–1094.

Donkelaar, A.V., Martin, R.V., Park, P.J., 2006. Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing. J. Geophys. Res. Atmos. 111 (D21).

Elder, A., Schwartz, J., Oberdörster, G., 2015. Particulate Air Pollution and CNS Health. Springer, London.

Fu, G.Q., Xu, W.Y., Rong, R.F., Li, J.B., Zhao, C.S., 2014. The distribution and trends of fog and haze in the North China plain over the past 30 years. Atmos. Chem. Phys. Discuss. 14 (11), 11949–11958.

Fukuda, S., Nakajima, T., Takenaka, H., Higurashi, A., Kikuchi, N., Nakajima, T.Y., Ishida, H., 2013. New approaches to removing cloud shadows and evaluating the 380 nm surface reflectance for improved aerosol optical thickness retrievals from the GOSAT/TANSO-Cloud and Aerosol Imager. J. Geophys. Res. Atmos. 118 (24), 13–13,531.

Guo, J.P., Zhang, X.Y., Che, H.Z., Gong, S.L., An, X., Cao, C.X., Jie, G., Zhang, H., Wang, Y.Q., Zhang, X.C., 2009. Correlation between PM concentrations and aerosol optical depth in eastern China. Atmos. Environ. 43 (37), 5876–5886.

Guo, J.P., Zhang, X.Y., Wu, Y.R., Zhaxi, Y., Che, H.Z., Ba, L., Wang, W., Li, X.W., 2011. Spatio-temporal variation trends of satellite-based aerosol optical depth in China during 1980–2008. Atmos. Environ. 45 (37), 6802–6811.

Guo, J.P., Wu, Y.R., Zhang, X.Y., Li, X.W., 2013. Estimation of PM2.5 over eastern China from MODIS aerosol optical depth using the back propagation neural network. Environ. Sci. 34 (3), 817–825.

Guo, J., Miao, Y., Zhang, Y., Liu, H., Li, Z., Zhang, W., He, J., Lou, M., Yan, Y., Bian, L., Zhai, P., 2016. The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data. Atmos. Chem. Phys. 16, 13309–13319.

Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., 2017. Impact of diurnal variability and meteorological factors on the PM2.5 - AOD relationship: implications for PM2.5 remote sensing. Environ. Pollut. 221 (94), 94.

He, Q., Huang, B., 2018. Satellite-based mapping of daily high-resolution ground PM2.5 in China via space-time regression modeling. Remote Sens. Environ. 206, 72–83.

Higurashi, A., 1999. Development of a two-channel aerosol retrieval algorithm on a global scale using NOAA AVHRR. J. Atmos. Sci. 56 (56), 924–941.

Huang, S.L., Miao-Kan, H., Chan, C.C., 2003. Effects of submicrometer particle compositions on cytokine production and lipid peroxidation of human bronchial epithelial cells. Environ. Health Perspect. 111 (4), 478–482.

IPCC, 2013. Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York, USA.

Kikuchi, M., Murakami, H., Suzuki, K., Nagao, T.M., Higurashi, A., 2018. Improved hourly estimates of aerosol optical thickness using spatiotemporal variability derived from Himawari-8 geostationary satellite. IEEE Trans. Geosci. Remote Sens. PP(99), 1–14.

Kulshrestha, U.C., 2018. PM1 is more important than PM2.5 for human health protection. Curr. World Environ. 13 (1), 01–02.

Li, Y., Chen, Q., Zhao, H., Wang, L., Tao, R., 2015. Variations in PM10, PM2.5 and PM1.0 in an urban area of the Sichuan Basin and their relation to meteorological factors. Atmosphere 6 (1), 150–163.

Li, T., Shen, H., Zeng, C., Yuan, Q., Zhang, L., 2017a. Point-surface fusion of station measurements and satellite observations for mapping PM 2.5 distribution in China: methods and assessment. Atmos. Environ. 152, 477–489.

Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., Zhu, B., 2017b. Aerosol and boundary-layer interactions and impact on air quality. Natl. Sci. Rev. 4 (6), 810–833. https://doi.org/10.1093/nsr/nwx117.

Lu, X., Mao, F., Pan, Z., Gong, W., Wang, W., Tian, L., Fang, S., 2018. Three-dimensional physical and optical characteristics of aerosols over Central China from long-term CALIPSO and HYSPLIT data. Remote Sens. 10 (2), 314.

Madronich, S., Flocke, S., 1999. The Role of Solar Radiation in Atmospheric Chemistry. Springer, Berlin Heidelberg.

Mao, F., Pan, Z., Henderson, D.S., Wang, W., Gong, W., 2018. Vertically resolved physical and radiative response of ice clouds to aerosols during the Indian summer monsoon season. Remote Sens. Environ. 216, 171–182.

Matheron, G., 1971. The Theory of Regionalized Variables and Its Application. 5.

Morawska, L., Keogh, D.U., Thomas, S.B., Mengersen, K., 2008. Modality in ambient particle size distributions and its potential as a basis for developing air quality regulation. Atmos. Environ. 42 (7), 1617–1628.

Oberdörster, G., Oberdörster, E., Oberdörster, J., 2005. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ. Health Perspect. 7 (113), 823–839.

Oliver, M.A., Webster, R., 1990. Kriging: a method of interpolation for geographical information systems. Int. J. Geogr. Inf. Syst. 4 (3), 313–332.

Pan, Z., Mao, F., Gong, W., Min, Q., Wang, W., 2017. The warming of Tibetan Plateau enhanced by 3D variation of low-level clouds during daytime. Remote Sens. Environ. 198, 363–368.

Pawan, G., Christopher, S.A., 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: multiple regression approach. J. Geophys. Res. Atmos. 114 (D14).

Qin, K., Wang, L., Wu, L., Xu, J., Rao, L., Letu, H., Shi, T., Wang, R., 2017. A campaign for investigating aerosol optical properties during winter hazes over Shijiazhuang, China. Atmos. Res. 198, 113–122.

Qin, K., Zou, J., Guo, J., Lu, M., Bilal, M., Zhang, K., Ma, F., Zhang, Y., 2018. Estimating PM1 concentrations from MODIS over Yangtze River Delta of China during 2014–2017. Atmos. Environ. 195, 149–158.

Sabbagh-Kupelwieser, N., 2010. Urban aerosol studies of PM1 size fraction with reference to ambient conditions and visibility. Aerosol Air Qual. Res. 10 (5), 425–432.

Shi, Y., Chen, J., Hu, D., Wang, L., Yang, X., Wang, X., 2014. Airborne submicron particulate (PM1) pollution in Shanghai, China: chemical variability, formation/dissociation of associated semi-volatile components and the impacts on visibility. Sci. Total Environ. 473, 199–206.

Specht, D.F., 1991. A general regression neural network. IEEE Trans. Neural Netw. 2 (6), 568.

Specht, D.F., 1996. Probabilistic Neural Networks and General Regression Neural Networks. McGraw-Hill, Inc.

Sun, W., Sun, J., 2016. Daily PM2.5 concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. J. Environ. Manag. 188, 144.

Tao, M., Chen, L., Su, L., Tao, J., 2012. Satellite observation of regional haze pollution over the North China Plain. J. Geophys. Res. Atmos. 117 (D12).

Wang, Y.Q., Zhang, X.Y., Sun, J.Y., Zhang, X.C., Che, H.Z., Li, Y., 2015. Spatial and temporal variations of the concentrations of PM10, PM2.5 and PM1 in China. Atmos. Chem. Phys. Discuss. 15 (23), 3585–13598.

Wang, W., Mao, F., Du, L., Pan, Z., Gong, W., Fang, S., 2017. Deriving hourly PM2.5 concentrations from Himawari-8 AODs over Beijing–Tianjin–Hebei in China. Remote Sens. 9 (8), 858.

Wen, L., Chen, J., Yang, L., Wang, X., Xu, C., Sui, X., Yao, L., Zhu, Y., Zhang, J., Zhu, T., 2015. Enhanced formation of fine particulate nitrate at a rural site on the North China Plain in summer: the important roles of ammonia and ozone. Atmos. Environ. 101, 294–302.

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 (7), 20–30.

Zang, L., Mao, F., Guo, J., Gong, W., Wang, W., Pan, Z., 2018. Estimating hourly PM1 concentrations from Himawari-8 aerosol optical depth in China. Environ. Pollut. 241, 654–663.

Zhang, J., Yang, L., Mellouki, A., Wen, L., Yang, Y., Gao, Y., Jiang, P., Li, Y., Wang, W., 2016. Chemical characteristics and influence of continental outflow on PM1.0, PM2.5 and PM10 measured at Tuoji island in the Bohai Sea. Sci. Total Environ. 573, 699–706.

Zhang, T., Zhu, Z., Gong, W., Zhu, Z., Sun, K., Wang, L., Huang, Y., Mao, F., Shen, H., Li, Z., 2018. Estimation of ultrahigh resolution PM2.5 concentrations in urban areas using 160 m Gaofen-1 AOD retrievals. Remote Sens. Environ. 216, 91–104.

Zhou, Q., Jiang, H., Wang, J., Zhou, J., 2014. A hybrid model for $PM_{2.5}$ forecasting based on ensemble empirical mode decomposition and a general regression neural network. Sci. Total Environ. 496 (2), 264–274.

Zhuang, X., Wang, Y., He, H., Liu, J., Wang, X., Zhu, T., Ge, M., Zhou, J., Tang, G., Ma, J., 2014. Haze insights and mitigation in China: An overview. J. Environ. Sci. 26 (1), 2–12.