Journal of Cleaner Production xxx (xxxx) 126493

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: http://ees.elsevier.com

High-resolution prediction of the spatial distribution of $PM_{2.5}$ concentrations in China using a long short-term memory model

Zhige Wang[a,b], Yue Zhou[c], Ruiying Zhao[a,b], Nan Wang[a,b], Asim Biswas[d], Zhou Shi[a,b,*]

[a] Institute of Agricultural Remote Sensing and Information Technology Application, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, 310058, China
[b] Key Laboratory of Environment Remediation and Ecological Health, Ministry of Education, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, 310058, China
[c] Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, 1348, Louvain-la-Neuve, Belgium
[d] School of Environmental Sciences, University of Guelph, Alexander Hall 135, 50 Stone Road East, Guelph, ON, N1G 2W1, Canada

ARTICLE INFO

Article history:
Received 10 November 2020
Received in revised form 7 February 2021
Accepted 22 February 2021
Available online xxx
Handling editor Bin Chen

Keywords
$PM_{2.5}$ prediction
AOD
LSTM
Time series data modeling

ABSTRACT

The concentration of fine particulate matter ($PM_{2.5}$) has a significant impact on the environment and human health. However, strong spatial heterogeneity and spatiotemporal dependence increases the difficulty of prediction. Moreover, due to the lag of the update of auxiliary variables at national scale in the prediction application, it is still difficult to achieve the timely nationwide $PM_{2.5}$ prediction at present. To better model and predict real time concentrations and spatial distributions of $PM_{2.5}$, this study developed a workflow of future $PM_{2.5}$ concentrations prediction based on long short-term memory (LSTM) model. Using ground-based station $PM_{2.5}$ data in 2014–2018, the 1 km Multi-Angle Implementation of Atmospheric Correction (MAIAC) aerosol optical depth (AOD) product and other auxiliary data to predict $PM_{2.5}$ concentrations in the next year and generate a high-resolution national $PM_{2.5}$ concentration spatial distribution map. The LSTM model outperformed random forest (RF) and Cubist approaches for prediction $PM_{2.5}$ because of its recurrent neural network structure that can capture time dependence and nonlinear relationships among $PM_{2.5}$ concentrations and other independent variables, and exhibited a stable accuracy with an R^2 of 0.83, by applying the annual time series, with an improvement of 0.04–0.09, compared to daily and monthly data. The results indicated that $PM_{2.5}$ pollution had gradually decreased in 2019 after application of pollution controls, with annual mean $PM_{2.5}$ concentrations of 27.33 ± 15.56 μg m^{-3}, although there were still some areas with severe pollution, including the North China Plain, parts of the Loess Plateau, and the Taklimakan Desert. The LSTM model makes it possible to predict fine-scale $PM_{2.5}$ spatial distributions nationwide in the future and may thus be useful for sustainable management and control of air pollution at a national scale.

1. Introduction

Fine particles with an aerodynamic diameter less than or equal to 2.5 μm or $PM_{2.5}$ are mainly caused by fuel combustion and other human activities. Because of the nature of the particles, other toxic substances are easily adsorbed by $PM_{2.5}$, which can be inhaled into the lungs through breathing and cause respiratory and cardiovascular diseases, or even death when integrated into the blood (Kampa and Castanas 2008; Raaschou-Nielsen et al., 2013; Zhang et al., 2019). Additionally, $PM_{2.5}$ has also been associated with adverse environmental issues for their strong absorption and scattering effect on visible light (Che et al., 2014). In recent years, environmental pollution and human health problems caused by high concentrations of $PM_{2.5}$ in China have attracted widespread attention (Liang et al., 2016; Peng et al., 2016; Wang et al., 2017). Measuring and monitoring concentration of $PM_{2.5}$ is the basis of pollution control, while at present, the number and distribution of air quality monitoring stations in China are

* Corresponding author. Institute of Agricultural Remote Sensing and Information Technology Application, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, 310058, China.
E-mail address: shizhou@zju.edu.cn (Z. Shi)

https://doi.org/10.1016/j.jclepro.2021.126493
0959-6526/© 2021.

insufficient to control $PM_{2.5}$ and support research. Advanced prediction of $PM_{2.5}$ distribution on fine resolution and its spatial changes can provide basis for developing pollution control strategies which is of great significance for comprehensive prevention of atmospheric environmental pollution and the control of human health and environmental risks.

Compared with traditional ground monitoring, satellite remote sensing technology is macroscopic, capable of obtaining information quickly, and has gradually become an important monitoring method. Aerosol optical depth (AOD), an integral of the extinction coefficient of the medium in vertical direction, describes the reduction effect of aerosol on light and could be obtained fast and over a large area through remote sensing. This is possible due to the correlation between AOD and the concentrations of near-surface particulate matters (Engel-Cox et al., 2004; Hu et al., 2014; Zhang and Li 2015). Currently, an increasing number of satellite remote sensing products are being used for retrieval of $PM_{2.5}$ concentrations (Alston et al., 2012; Li et al., 2015a; Meng et al., 2015; Pang et al., 2018; Philip et al., 2014; Tsai et al., 2011; Wang and Chen 2016). However, the coarse-resolution (5–40 km) AOD products limits the capacity for high-resolution prediction of $PM_{2.5}$ concentration on a national scale. The newly released MCD19A2 data is a Moderate Resolution Imaging Spectrometer (MODIS) Terra and Aqua combined Multi-angle Implementation of Atmospheric Correction (MAIAC) Land AOD gridded Level 2 product. This is produced daily and at 1 km pixel resolution (Lyapustin et al., 2018) and makes possible a high-resolution and high accuracy $PM_{2.5}$ prediction.

The AOD products to retrieve and predict ground $PM_{2.5}$ concentration were often based on empirical statistical methods. There are two major category of methods including statistical regression models, for example, simple linear regression (SLR) model (Engel-Cox et al., 2004; Koelemeijer et al., 2006), geographically weighted regression (GWR) model (Xiao et al., 2018; Yu et al., 2017), geographically and temporally weighted regression (GTWR) model (Bai et al., 2016), and machine learning-based methods, such as random forest (RF) model (Chen et al., 2018), and support vector regression (SVR) model (Liu et al., 2019). However, the $PM_{2.5}$ related physical and human determinants are diverse which may cause spatial stratified heterogeneity (SSH) (Bai et al., 2019; Wang et al., 2016), meanwhile concentrations of $PM_{2.5}$ are affected by many variables, including weather, terrain, and human activities which are often related in complex nonlinear fashion (Li et al., 2015b; Guo et al., 2014). Machine learning-based methods can solve huge data processing problems and fit nonlinear relationships well. However, these relationships ignore complex temporal correlations among $PM_{2.5}$ concentrations and their controlling variables (Zhang and Cao 2015). Moreover, delayed (or lagged) update and availability of the auxiliary variables make it difficult to timely predict nationwide $PM_{2.5}$ concentrations. Capturing temporal correlations and applying lagged data to timely predict $PM_{2.5}$ concentration remain a significant research problem that must be resolved.

Recently, deep learning-based methods, including long short-term memory (LSTM) model, has been proposed to provide a good solution in solving temporal correlation problems (Fan et al., 2017). The LSTM model builds a recurrent neural network (RNN) structure, which represents a deep network structure containing a long time-series. The LSTM model has shown promise to adapt different types and representations of data, recognize sequential patterns over a long-time span, and capture complex nonlinear relationships (Hochreiter and Schmidhuber, 1997). It has been applied to various time series prediction fields including human trajectory prediction (Alexandre et al., 2016), crop yield estimation (Jiang et al., 2020) and air pollutant prediction (Pak et al., 2020; Qi et al., 2019; Li et al., 2017a).

Characterizing and quantifying nonlinear relationships among numerous controlling variables and $PM_{2.5}$ concentrations are critical in providing a higher precision prediction of $PM_{2.5}$. Since air pollution has a strong time-dependent effect, $PM_{2.5}$ can be modeled using the deep learning methods which taking full advantage of long-term data. Li et al. (2017b) introduced the LSTM model into the field of air pollutant concentrations prediction, and proved the superiority of LSTM model compared with other statistics-based models. Pak et al. (2020) proposed a spatiotemporal convolutional neural network (CNN) and long short-term (LSTM) memory (CNN-LSTM) model to predict the next day's daily average $PM_{2.5}$ concentration. However, they didn't consider satellite-derived products such as aerosol data, boundary layer height (BLH) and normalized difference vegetation index (NDVI) which have significant correlation with $PM_{2.5}$ (Gupta et al., 2006; Saide et al., 2011; Di et al., 2016). On this basis, in order to capture the change of $PM_{2.5}$ concentration more accurately, the convolutional neural network (CNN) and recurrent neural network (RNN) model integrated multi-source data including AOD data and meteorological data was proposed by Wen et al. (2019) and Wu et al. (2020). However, most of their studies focus on the prediction of $PM_{2.5}$ concentration over ground monitoring station level, few attempts to capture the spatial continuous change and the large geographic distribution of $PM_{2.5}$ using deep learning models. Meanwhile, the current $PM_{2.5}$ prediction based on deep learning is mostly based on the daily and hourly short time scales, the application of intelligent algorithms at different time scales needs further exploration.

To fill the research gap, we compared the performance of LSTM models in three different timescales and extended the deep learning workflow which taking full advantage of long-term data and heterogeneous geospatial data, integrating AOD products, meteorological variables, and land use factors for predicting and mapping annual $PM_{2.5}$ at a spatial resolution of 1 km. In this context, we addressed the following specific research questions, across mainland China, regarding robust $PM_{2.5}$ prediction:

1. Does the LSTM model improve performance of $PM_{2.5}$ prediction over other machine learning approaches?
2. How does the LSTM model perform under different timescales?
3. Does the LSTM model show robust performance in terms of national scale $PM_{2.5}$ mapping?

2. Materials and methods

The prediction and mapping workflow (Fig. 1) for $PM_{2.5}$ concentrations were developed by integrating auxiliary data, AOD products, and surface-based observations of $PM_{2.5}$ concentrations. Eight variables related to meteorology and land use were selected to establish relationships with $PM_{2.5}$ concentrations.

2.1. Ground-level $PM_{2.5}$

The daily average $PM_{2.5}$ concentrations recorded at 1467 monitoring sites across China from January 2014 to December

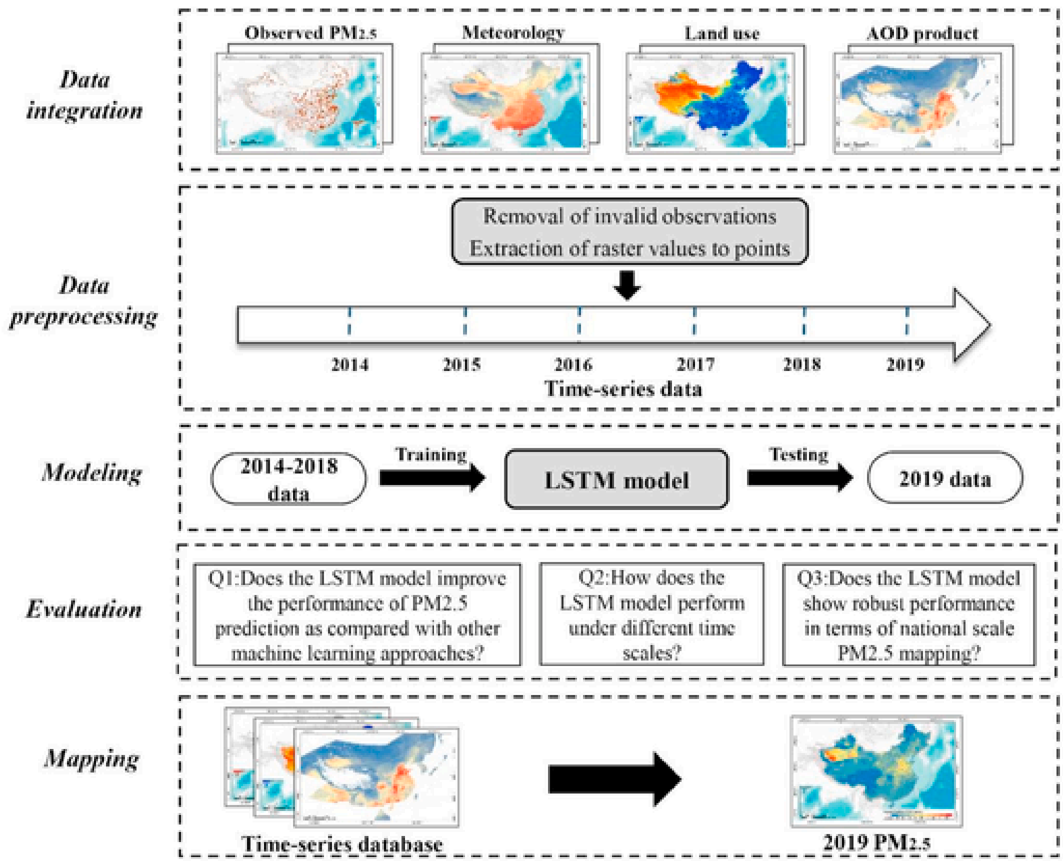

Fig. 1. A workflow of national $PM_{2.5}$ concentration prediction and mapping in the study.

2019, were obtained from the official website of the China Environmental Monitoring Center (CEMC, http://106.37.208.233: 20035/). The distribution of monitoring sites is quite heterogeneous with higher number of stations in southeastern China and scattered stations across northwestern China (Fig. 2).

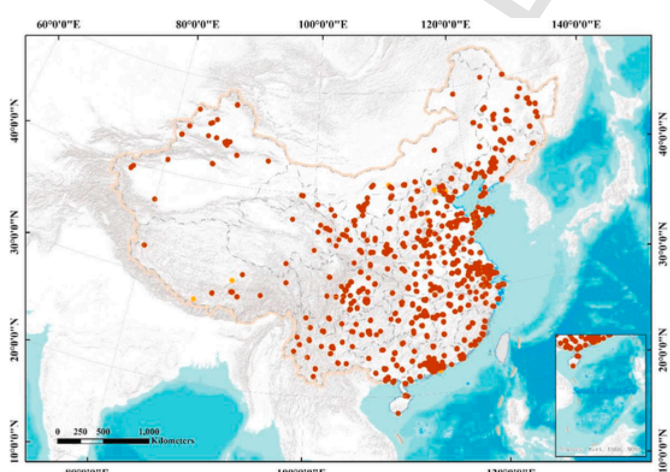

Fig. 2. Geographical distribution of $PM_{2.5}$ (red dots) and Aerosol Robotic Network (orange dots) monitoring stations included in the study.

2.2. MODIS AOD products

MODIS has been on board the NASA Terra since 1999 and Aqua since 2002, providing two daily observations of columnar aerosol properties at approximately 10:30 a.m. (Terra) and 1:30 p.m. (Aqua) local time.

2.2.1. MODIS 1 km MAIAC AOD products

In our study, the newly released MAIAC AOD products from 2011 to 2018 were downloaded from NASA (https://ladsweb.modaps.eosdis.nasa.gov/). They were subject to quality assurance (QA), cloud screening ($QA_{CloudMask}$ = Clear), and adjacency ($QA_{AdjacencyMask}$ = Clear) tests. The data are a MODIS Terra and Aqua combined AOD gridded Level 2 product produced daily at 1 km pixel resolution, based on the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm. The MAIAC is a new advanced algorithm that uses time-series analysis and a combination of pixel- and image-based processing to improve the accuracy of cloud detection, aerosol retrievals, and atmospheric correction. To adapt to global processing and improve cloud/snow detection, aerosol retrievals, and atmospheric correction of MODIS data, the MAIAC MODIS, which added polarization correction of MODIS Terra, removed residual trends of both Terra and Aqua, and cross-calibrated Terra to Aqua (Lyapustin et al., 2018).

The data processing steps were as follows: (a) the 550 nm wavelength aerosol optical depth data were selected following a geometric correction; and (b) after geometric correction, the 550 nm wavelength aerosol optical depth data were reprojected and mosaicked to obtain the daily product. Average

AOD value of the overlapping areas in the mosaic image was then selected for further analysis.

2.2.2. Validation against AERONET measurements

Aerosol Robotic Network (AERONET) data have been widely used for validating satellite derived AOD because of their high accuracy (Kahn et al., 2010; van Donkelaar et al., 2010). The operational MODIS AOD products (as described previously) were then validated against AERONET Level 1.5 data (quality-assured) at 19 sites across China, from 2011 to 2018. Table S1 provides detailed information about the selected AERONET sites in China.

To compare with the MODIS AOD values, AERONET AOD at 550 nm was interpolated from AERONET AOD at 440 and 870 nm using the Angstrom Exponent provided by AERONET (Li et al., 2009; Ma et al., 2014). When compared to the AERONET AOD observations (Fig. S1), the MAIAC AOD (retrievals falling within the expected error bounds, EE = 77.44% and RMSE = 0.1341) exhibited a high accuracy, indicating that the MAIAC product has improved spatial coverage and higher spatial resolution in China, and was sufficient to support our study.

2.3. Auxiliary data

A number of meteorological and land-cover-related variables were used in this study. According to previous studies (He and Huang 2018; Wei et al., 2019), several variables have significant positive or negative effects on $PM_{2.5}$ concentrations ($p < 0.01$), and consequently we selected the following variables (Table S3): specific humidity (SHUM), temperature (TEMP), wind speed (WS), precipitation rate (PREC), pressure (P), and boundary layer height (BLH). The land-cover-related variables were digital elevation (DEM) and normalized difference vegetation index (NDVI). Table S2 summarizes the details of the data used in this study. It is worth noting that the national 1 km resolution meteorological data from National Tibetan Plateau Data Center is only up to 2018. Thus, the prediction of $PM_{2.5}$ in our study were conducted in the next year 2019. All independent variables passed the significance test of correlation, especially AOD ($r = 0.327$, $p < 0.01$) and BLH ($r = 0.287$, $P < 0.01$), which showed high correlation with $PM_{2.5}$ concentrations (Table S3). However, the collinearity of independent variables must be taken into account to avoid duplication of information conveyed in seemingly independent variables. For this, the variance inflation factor (VIF) was used to diagnose collinearity among the selected variables. The results in Table S4 indicate a strong collinearity (DEM ~6.34 and P ~6.30) when both P and DEM were included in the model. The P data with time variable characteristics was included in the model to help present detailed spatiotemporal variation of $PM_{2.5}$ along with the other seven independent variables.

As the AOD product, meteorological, and land-cover-related data were all raster data, they were first extracted at the $PM_{2.5}$ ground-based observations to prepare calibration datasets using the Spatial Analyst Tools in ArcGIS 10.3.

2.4. LSTM model

Long short-term memory (LSTM) models (Hochreiter and Schmidhuber, 1997), an improved approach of recurrent neural networks (RNN) (Fan et al., 2017), are designed to deal with time-series data. This approach is proposed to compensate for the weaknesses of RNNs including the vanishing or exploding gradient problem in the prediction of long time-series data, and performed well in addressing long-range dependency in previous research (Fan et al., 2017; Li et al., 2017b). The model used in the study consisted of one input layer, three LSTM layers, and one output layer (Fig. 3a).

The structure of the LSTM block is presented in Fig. 3b. Each LSTM block contained a cell state $C_t$and three gates, namely, forget gate f_t, input gate, $i_t$and output gate o_t. The three gates, with a sigmoid activation function control changes in the cell state.

The forget gate can be computed as:

$$f_t = \sigma \left(W_f \cdot \left[h_{t-1}, x_t \right] + b_f \right) \tag{1}$$

where f_t is the vector of the input gate, W_f and b_f are the weight and bias vector of the forget gate, respectively, $h_{t-1}$is the output result at the last (*t*-1) moment, x_t is the current (*t*) input, $\left[h_{t-1}, x_t \right]$ means connecting two vectors into a longer vector, σ, and is used as the activation function.

The input gate and output gate can be computed as:

$$i_t = \sigma \left(W_i \cdot \left[h_{t-1}, x_t \right] + b_i \right) \tag{2}$$

$$C_t = f_t * C_{t-1} + i_t * tanh \left(W_C \cdot \left[h_{t-1}, x_t \right] + b_C \right) \tag{3}$$

$$\begin{aligned} o_t &= \sigma \left(W_o \cdot \left[h_{t-1}, x_t \right] + b_o \right) \\ h_t &= o_t * tanh \left(C_t \right) \end{aligned} \tag{4a}$$

where, i_t, o_t and C_t are the vectors for the input gate, output gate, and cell state, respectively; $h_t$is the output vector; W_i, W_C and W_o are the weights of the corresponding gate; b_i, b_c and b_o are the bias vectors of the corresponding gate, *tanh*, and is used as the activation function.

The LSTM model for $PM_{2.5}$ concentrations prediction was implemented on Keras, which is an advanced neural network API in Python. All data were first normalized using the min-max method before feeding into the model. In order to achieve the best performance of the model, the LSTM model was set with a fixed batch size of 72, and three unidirectional LSTM layers with 300 hidden units. For network parameter optimization, we set the number of epochs to 300 and employed a gradient descending-based optimizer Adam and a learning rate scheduler in which the learning rate was scheduled to reduce by 1/10th of the original for every 50 epochs. The timestep is set as 3, which means predict the $PM_{2.5}$ concentrations in the fourth year used the independent variables in the first three years (Table S5 for details).

2.5. Model fitting and validation

Two evaluation cases were designed to evaluate the LSTM model, corresponding to questions 1 and 2 posed in the Introduction. First, two other widely used machine learning approaches, random forest (RF) and Cubist as benchmarking models were simultaneously applied to evaluate the performance of the LSTM.

The RF created multiple randomly generated decision trees based on the bagging algorithm, then each node of the decision tree was split and selected resulting in the maximum variance between sub-groups. An average was then used for prediction (Zhou et al., 2019). Cubist is a piecewise linear tree model that uses a recursive partitioning of the predictor variable space. It takes a divide-and-conquer strategy and seeks to minimize the intra-subset variation at each node (Rossel and

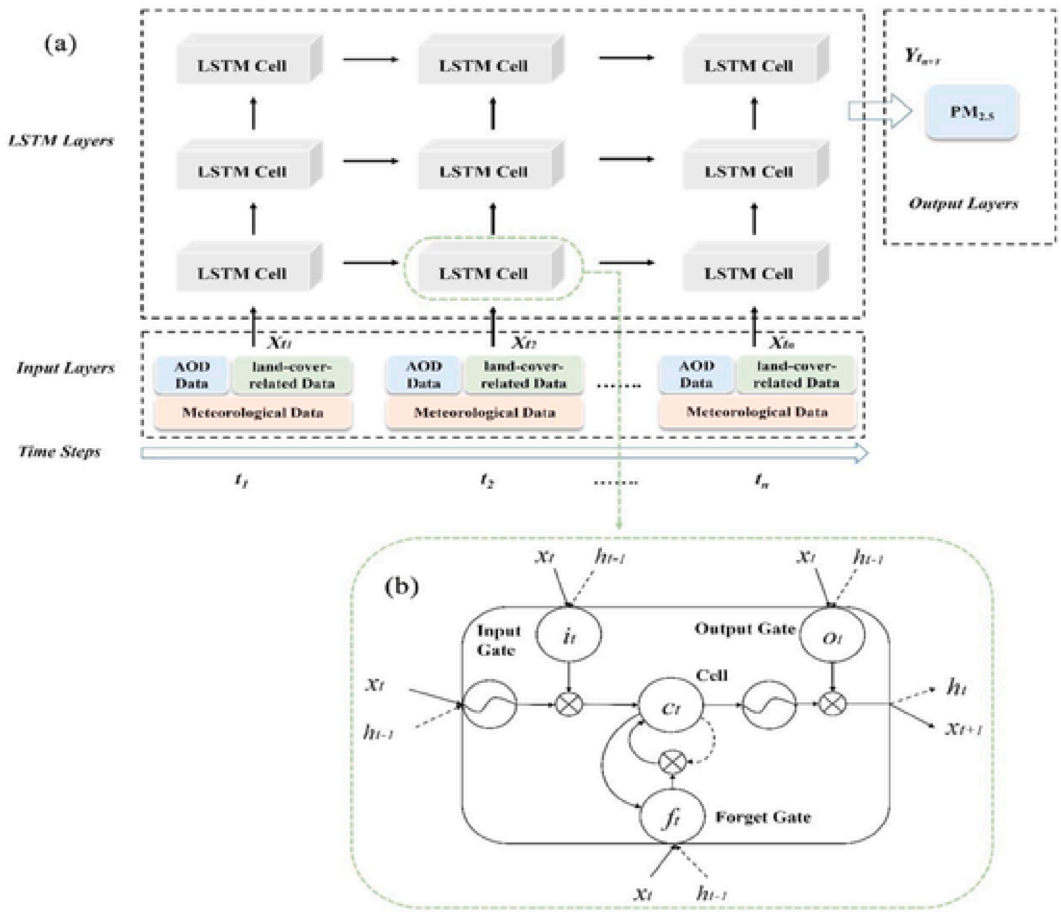

Fig. 3. (a)The structure of the LSTM model; (b) The structure and components of each cell in the LSTM.

Chen 2011; Ma et al., 2017; Peng et al., 2019), and can better solve the SSH shown by $PM_{2.5}$.

LSTM, RF and Cubist three models were trained using data from 2014 to 2017 and tested in 2018 because the RF and Cubist models could only predict the $PM_{2.5}$ concentrations in the current year with auxiliary data, the validation could not conduct in 2019. Previous studies of $PM_{2.5}$ prediction in China using different methods were also compared in this study.

We further compared the performance of LSTM models from three different timescales including daily, monthly, and annual. At each scale, the model was trained using the same amount of data from 2014 to 2018 and tested in 2019 to ensure comparability.

To evaluate the predictive ability of the models, four indices, including the determination coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE), and estimation bias, were calculated:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4b}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - f_i)^2}{n}} \tag{5}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|(f_i - y_i)\right|}{n} \tag{6}$$

$$Bias = f_i - y_i \tag{7}$$

3. Results and discussion

3.1. Descriptive statistics

Boxplot of ground-based $PM_{2.5}$ concentrations is shown in Fig. 4. The concentrations of $PM_{2.5}$ exhibited a steady annual decline, and reached its lowest value over the study period in 2018, with an annual mean concentration of 39.24 μg m^{-3} (Table S6). The most volatile year of $PM_{2.5}$ was 2016 and 2017 with some very high concentrations, and the most stable year of $PM_{2.5}$ was 2018.

The spatial distributions of mean $PM_{2.5}$ concentrations, from 2014 to 2018, and the relevant independent variables selected in this study, are presented in Fig. 5. The mean concentrations of $PM_{2.5}$ over the study period ranged from 14.78 μg m^{-3} to 128.92 μg m^{-3} (Fig. 5a). The distribution of $PM_{2.5}$ exhibited clear spatial heterogeneity. High concentrations were observed across Northwest China, North China, and parts of the southwestern hinterland in China, and showed high similarity with the distribution of AOD. Due to cloud contamination, the AOD images had insufficient retrievals for some areas (Fig. 5b). Spatial pattern of precipitation was consistent with that of the specific humidity: high values were mostly distributed in the southeast and low values in the west (Fig. 5d and e). Temperature and pressure also exhibited a very similar spatial distribution (Fig. 5c and g). Strong spatial variability in these variables resulted in strong spatial variation in $PM_{2.5}$.

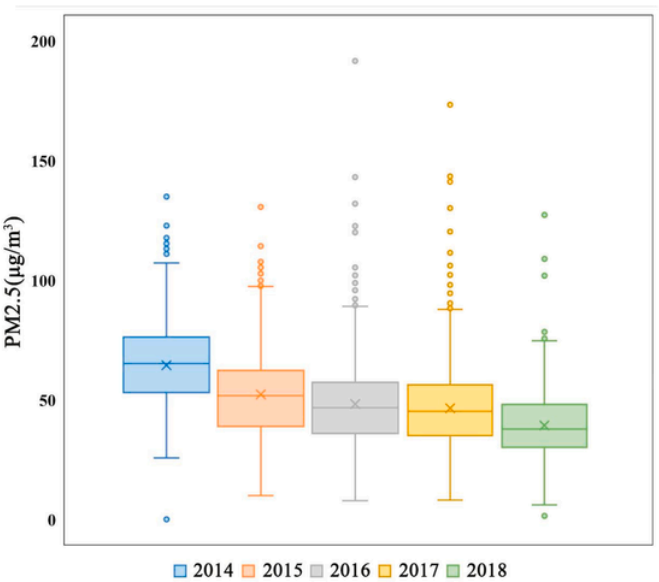

Fig. 4. Boxplots of site based $PM_{2.5}$ concentrations from 2014 to 2018.

3.2. Evaluation of the methods

3.2.1. Predictive power of different models

The $PM_{2.5}$ predicted values against observed values are shown in Fig. 6. Overall, the LSTM model outperformed RF and Cubist in the prediction based on inter-annual data. The LSTM model was robust in estimating and predicting $PM_{2.5}$ concentrations. The fitted regression lines of the RF model had the shallowest slopes and the largest intercepts, indicating the underestimation of the model. Though the fitted regression line of the Cubist model was the closest to the 1:1 line, the predicted results were relatively scattered with large deviations and reduced accuracy with the lowest R^2 of 0.73 and highest RMSE and MAE values of 8.65 μg m^{-3} and 6.49 μg m^{-3}, respectively. In contrast, the scatter plot of the LSTM model was the most concentrated and the prediction errors were low. By combining time and space information, the LSTM model performed the best, according to all evaluation metrics, with the highest R^2 value of 0.84, and the lowest RMSE and MAE values of 5.40 μg m^{-3} and 3.68 μg m^{-3}, respectively.

The studies on satellite-aerosol optical depth-based $PM_{2.5}$ prediction over whole China were select for comparison (Table 1). In model validation, R^2 values ranged from 0.64 to 0.85, and the RMSE from 4.1 μg m^{-3} to 30 μg m^{-3}. The traditional spatial statistical model GWR performed the worst, with an R^2 value of 0.64 and the largest RMSE of 32.98 μg m^{-3}. There was an obvious improvement (R^2 improved by 0.16 compared with the GWR model) in timely structure adaptive modeling (TSAM) model and GTWR model, which are two improved form of GWR model. The two-stage model, combined by stage-1 linear mixed-effects (LME) model and stage-2 GWR model, outperformed the GWR model by an increased R^2 values of 0.78 and 0.79, respectively. As two intelligent algorithms, the generalized regression neural network (GRNN) model (R^2 = 0.82) and RF model (R^2 = 0.81) could better capture the relationship between $PM_{2.5}$ and AOD and other auxiliary variables. Comparing the results, LSTM model performed the best with R^2 of 0.84 and the RMSE of 5.40 μg m^{-3}

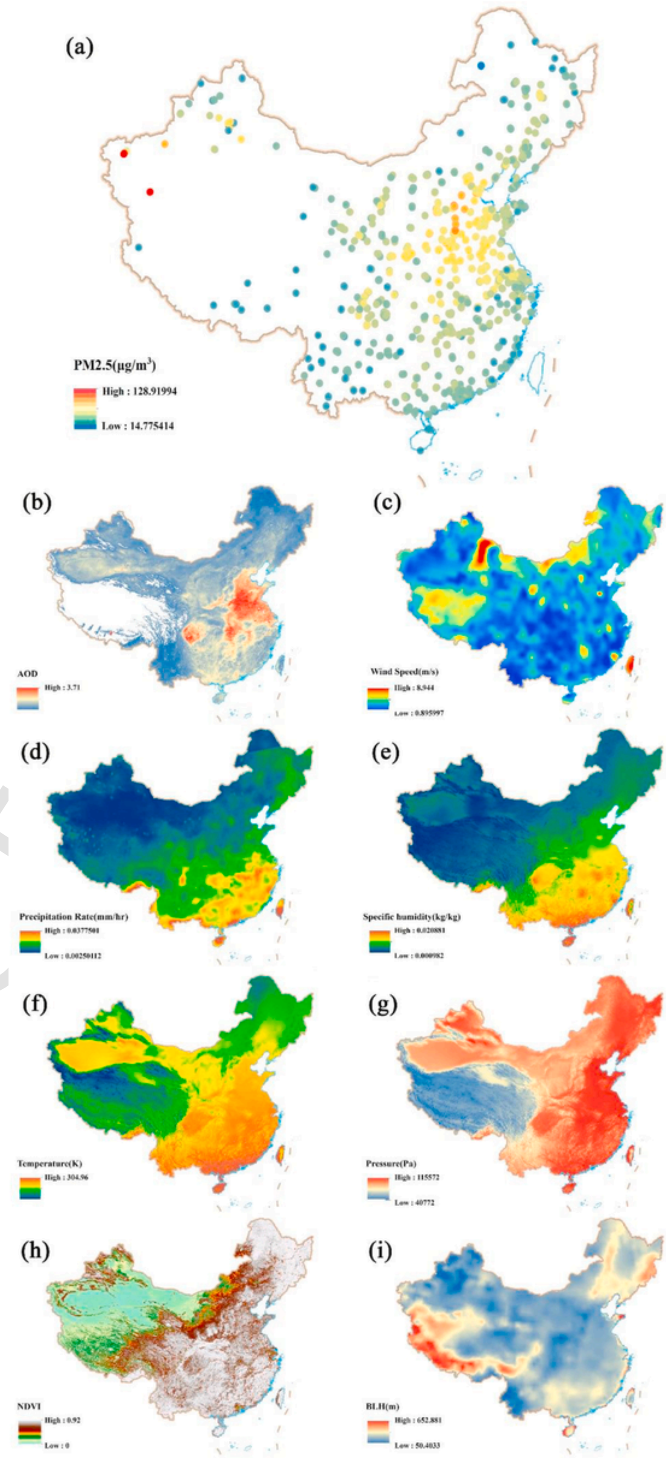

Fig. 5. (a) The mean ground-measured $PM_{2.5}$ concentrations in 2014–2018, and variables used in the prediction of $PM_{2.5}$: (b) AOD, (c) Pressure, (d) Precipitation, (e) Wind speed, (f) NDVI, (g) Temperature, (h) Specific humidity, and (i) BLH.

and thus has the capacity to better predict from the time series data .

3.2.2. Predictive power at different timescales

The LSTM model achieved the highest and most robust prediction accuracy at the annual timescale (Table 2). The annual prediction resulted in the highest R^2 value (0.83) and the low-

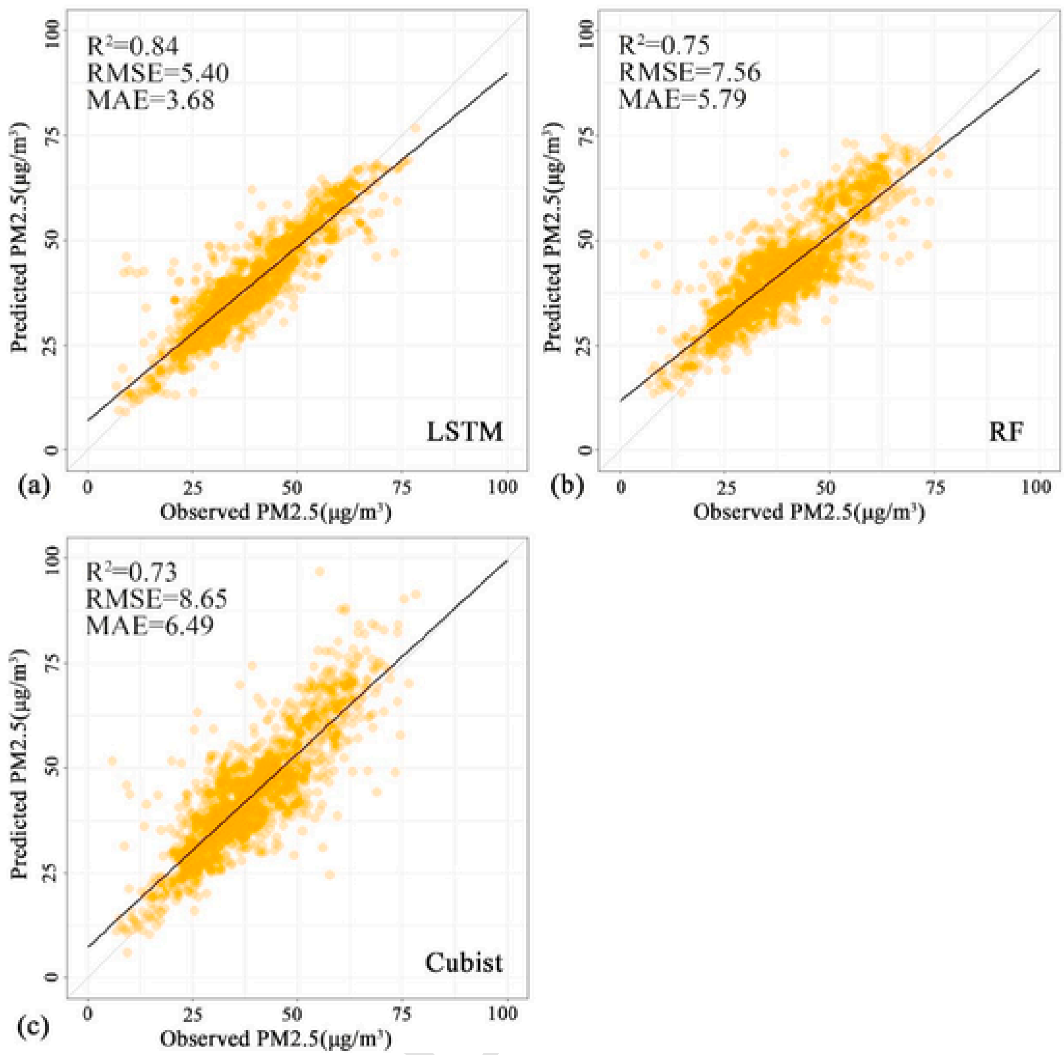

Fig. 6. Scatter plots of three models in validation. (a)–(c) are the results for LSTM, RF, and Cubist, respectively.

Table 1
Summary of different models performance in $PM_{2.5}$ prediction in China.

Related study	Study period	Spatial resolution	Model	Model validation	
				R^2	RMSE
Ma et al. (2014)	2012–2013	10 km	GWR	0.64	32.98
Ma et al. (2016)	2004–2013	10 km	Stage-1	0.78	27.99
			Stage-2	0.79	27.42
Fang et al. (2016)	2014–2015	10 km	TSAM	0.80	22.75
(Li et al., 2017)	2013–2014	3 km	GRNN	0.82	20.93
He and Huang (2018)	2015	3 km	GTWR	0.80	18.00
Wei et al. (2019)	2016	1 km	RF	0.81	17.91
Our study	2018–2019	1 km	LSTM	0.84	5.40

Table 2
Comparisons of the model performance across different time-scales.

Timescale	R^2	RMSE ($\mu g\ m^{-3}$)	MAE ($\mu g\ m^{-3}$)
Daily	0.79	22.23	14.78
Monthly	0.72	17.97	8.23
Annual	0.83	5.54	3.44

est RMSE (5.54 $\mu g\ m^{-3}$) and MAE (3.44 $\mu g\ m^{-3}$). The daily and monthly predictions showed relatively less accuracy with R^2 of 0.79 for the daily data. However, RMSE was 22.23 $\mu g\ m^{-3}$ and MAE was 14.78 $\mu g\ m^{-3}$, indicating that the daily values of $PM_{2.5}$ were the hardest to accurately model.

The boxplot of prediction bias demonstrates that the daily and monthly predictions were mostly distributed on the negative side, indicating an underestimation of $PM_{2.5}$ concentrations. The strong volatility and irregularities of daily and monthly $PM_{2.5}$ concentrations might result in errors in the prediction of $PM_{2.5}$. The annual prediction bias was mainly distributed around zero, with a few distributed on the positive side (Fig. 7). These results indicated that the LSTM model had a better ability to capture changes in mean annual $PM_{2.5}$ con-

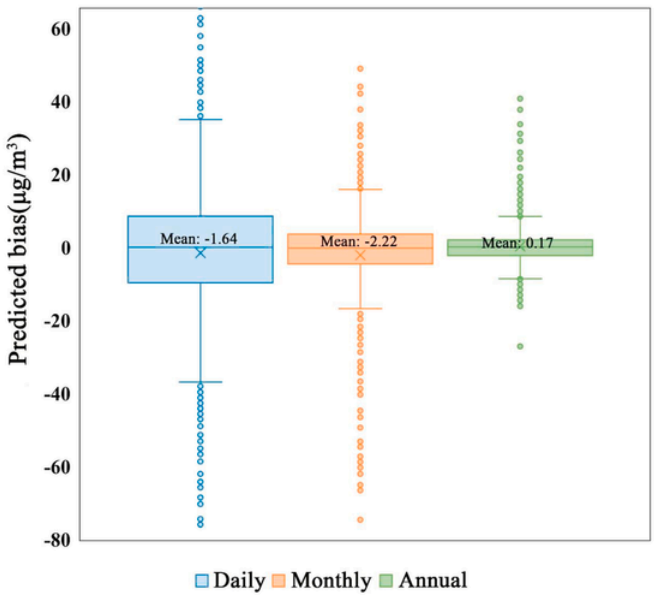

Fig. 7. Boxplot of the predicted bias of the LSTM model at daily, monthly, and annual timescales.

centrations, on the premise of the same amount of data, compared to that of the daily and monthly changes.

3.2.3. LSTM model performance of $PM_{2.5}$ prediction in 2019

The spatial distributions of ground-based measurements of $PM_{2.5}$ concentrations in 2019, and the predicted bias of the LSTM model, are shown in Fig. 8. The LSTM model exhibited high prediction accuracy against the ground-based $PM_{2.5}$ concentrations in 2019, with an R^2 of 0.83, and RMSE, and MAE of 5.54 μg m^{-3} and 3.44 μg m^{-3}, respectively. The bias in prediction ranged from −27.20 μg m^{-3} to 40.61 μg m^{-3}, and most of the biases (>78.5%) were distributed between −5 μg m^{-3} and 5 μg m^{-3}. Few biases (<5%) were exceeding 10 μg m^{-3} or less than −10 μg m^{-3}, indicating a robust performance of the LSTM model.

Spatial distribution of predicted bias showed variations, and may be attributed to the spatial heterogeneity of the response of $PM_{2.5}$ concentration to environmental variability. The biases of the LSTM model were mostly positive, and the overestimation was mainly distributed in the northwest China, Shandong province, and parts of the Sichuan Basin (Fig. 8b). The $PM_{2.5}$ pollution in these areas was generally high due to topography, weather, and anthropogenic emissions, which revealed that the capability of LSTM to predict extreme values or outliers was lower than its ability to predict average values. In the southern coastal region, Tibet, and the Northeast, areas with low $PM_{2.5}$ concentrations, the predicted bias were relatively small, and the LSTM showed good prediction performance.

The LSTM model achieved prediction of $PM_{2.5}$ concentration with a relatively high accuracy. The complex non-linear relationship between inputs of independent variables and $PM_{2.5}$ concentration was captured in neurons. However, the learning and modeling process of each neuron was not examined in depth, and thus the traceability and interpretability of the LSTM model was relatively low. Therefore, advanced analysis tools, such as attention networks, can be further studied to illustrate the process of machine learning in the future.

3.3. Predicted $PM_{2.5}$ concentrations over China

3.3.1. Spatial variation of $PM_{2.5}$ in China

The predicted spatial distribution of surface $PM_{2.5}$ concentrations in 2019 for the mainland China, with a 1-km grid derived using the LSTM model, are shown in Fig. 9. The spatial forecasts agreed well with the surface observations. The overall annual mean of predicted $PM_{2.5}$ concentrations for the whole of China was 27.33 ± 15.56 μg m^{-3}, lower than the annual mean of observed $PM_{2.5}$ (37.02 μg m^{-3}). This is because the annual mean of estimated $PM_{2.5}$ was calculated for the entirety of China, but the observed $PM_{2.5}$ concentrations only represented areas with monitoring sites.

High values of $PM_{2.5}$ were mainly distributed cross the North China Plain, the Loess Plateau, parts of central China, and decreased at the periphery. The highest concentrations of $PM_{2.5}$ were concentrated in the Taklimakan Desert in the southern Xinjiang Autonomous Region southwest of China, different from other severely polluted areas. The $PM_{2.5}$ concentrations in central of this region were lower than those in the surrounding areas. The heavy pollution in these regions were mainly caused by increasing anthropogenic emissions (i.e., fossil fuel combustion), as well as unfavorable meteorological conditions and topography (i.e., basin) that are not conducive to the dispersion of pollution. In contrast, the lowest $PM_{2.5}$ concentrations in the southwest and northeast regions of China are the lowest, followed by the southeast coastal areas. These areas are in Yunnan-Guizhou Plateau, Northeast Plain and Pearl River Delta. The low $PM_{2.5}$ concentrations can be attributed to the flat terrain and favorable climatic conditions including sufficient precipitation and the monsoon climate.

With continuous improvement of national pollution control capabilities, 80% of regions in China met the CNAAQS Level 2 standard and have annual mean of $PM_{2.5}$ concentrations less than 35 μg m^{-3}. However, pollution by $PM_{2.5}$ is still not opti-

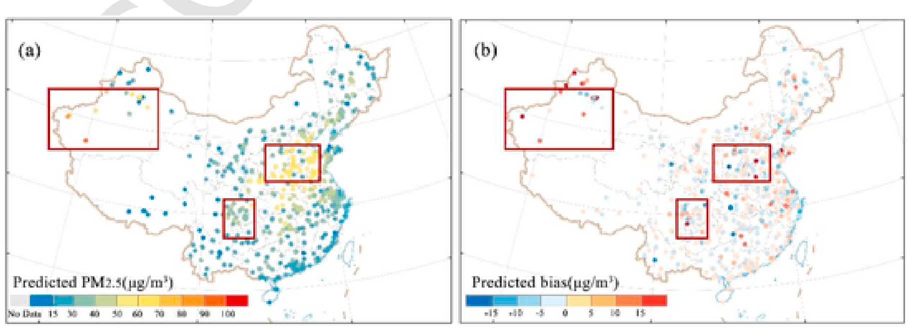

Fig. 8. The spatial distribution of (a) the ground-based measurements of $PM_{2.5}$ concentrations, and (b) the predicted bias of LSTM model in the validation dataset (year of 2019).

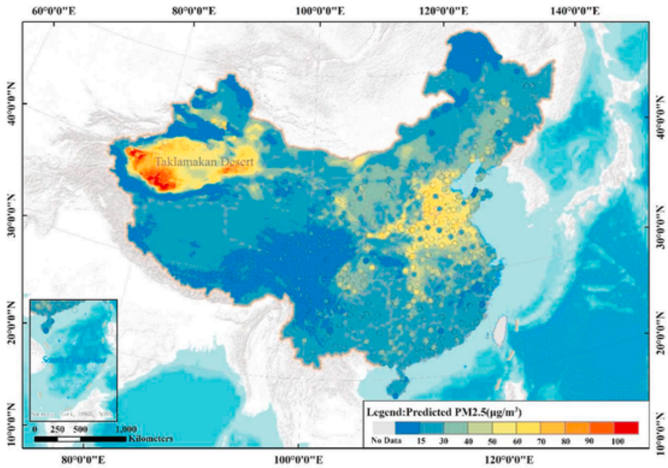

Fig. 9. Spatial distribution of 2019 satellite-derived (1 km × 1 km, background shading), and ground-measured (colored dots) $PM_{2.5}$ concentrations across China.

mistic, with over 83% of regions suffering from particulate pollution exceeding the CNAAQS Level 1 standard of 15 μg m^{-3}.

3.3.2. Heavy $PM_{2.5}$ pollution areas analysis

The distribution of high concentrations of $PM_{2.5}$ is generally aggregated and not uniform (Fig. 10), with clusters in the North China Plain, the Loess Plateau, and the Sichuan Basin. It is worth noting that while the Taklimakan Desert is also one of the most heavily polluted areas, no further analysis will be conducted because of no human settlement and few satellite and ground monitoring stations.

The North China Plain was the most heavily polluted area, and pollution was mainly distributed in Hebei, Beijing, Tianjin, northeast Henan, and Shandong (Fig. 10a). The average annual $PM_{2.5}$ concentration reached 48.53 μg m^{-3} ± 12.93 μg m^{-3}, ranging from 17.13 μg m^{-3} to 75.10 μg m^{-3}. These regions are densely populated, with rapid urban growth. In Hebei Province, heavy industries, including steel and coal, have been developed recently. Volatile organic compounds and nitrogen oxides generated by transportation among these regions have also increased $PM_{2.5}$ pollution (Fu et al., 2014; Guo et al., 2014; Zhao et al., 2013). In the Loess Plateau, pollution is mainly concentrated in the southern Shanxi and central Shaanxi (Fig. 10b). The maximum annual average $PM_{2.5}$ concentration was 39.42 ± 12.53 μg m^{-3}. Local industrial cities and mining areas in the Loess Plateau experience severe coal smoke pollution. As a result, $PM_{2.5}$ pollution mainly comes from the secondary particulate matter (Huang et al., 2014). Due to serious soil loss and the presence of vertical and horizontal gullies, industrial cities in the Loess Plateau are mostly located in the basin region. The high altitude surrounding the basin hinders surface winds, minimizing air circulation in the basin area, resulting in the accumulation of air pollutants under high emission conditions (Wang et al., 2018). The Sichuan Basin had the lowest pollution concentrations among the three pollution hotspots (Fig. 10c). The main polluted cities in the Sichuan Basin were eastern Sichuan and Chongqing, with an average annual $PM_{2.5}$ concentration of 30.77 ± 6.03 μg m^{-3}. Topographical factors in the Sichuan Basin have also contributed to $PM_{2.5}$ pollution. Chongqing, as a major industrial town in southwest China, also has serious industrial pollution, which has exacerbated $PM_{2.5}$ pollution (Li et al., 2015b; Zhang et al., 2012).

4. Conclusions

Based on the annual $PM_{2.5}$ time-series data, the newly released 1 km resolution MODIS AOD data, and 7 auxiliary variables, our study first attempted to use LSTM model in the an-

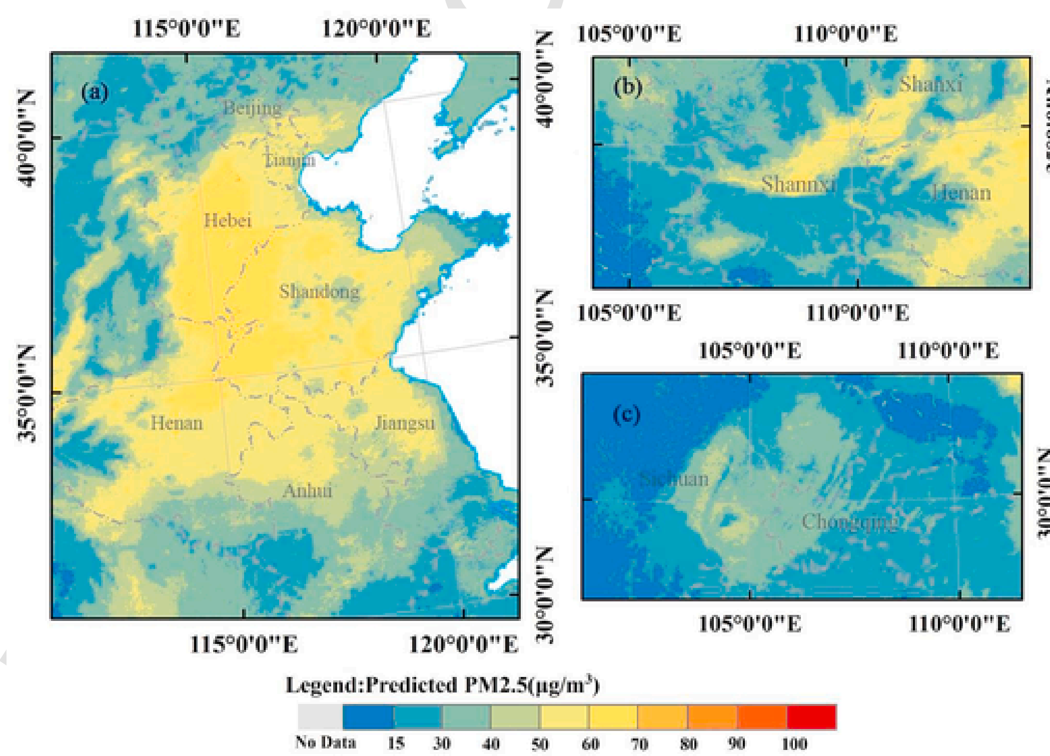

Fig. 10. Detailed map of the mean predicted $PM_{2.5}$ concentrations (1 km × 1 km) for areas experiencing high pollution: (a) the North China Plain, (b) the Loess Plateau, and (c) the Sichuan Basin.

nual scale for predicting $PM_{2.5}$ concentration of next year, meanwhile the prediction was extended from ground site point level to national surface level by means of digital mapping at spatial resolution of 1 km. The results of the study showed that: 1) The LSTM model captured the time dependence of $PM_{2.5}$ by constructing a recurrent neural network and memory gate structure framework, and achieved high performance with the lowest RMSE (5.40 μg m^{-3}) and the highest R^2 (0.84) in validation set (the year of 2018) compared with two recent machine learning approaches, the RF model and the Cubist model. 2) The LSTM model had been extended to annual scale prediction to prove the spatiotemporal scalability and could be used for predicting at different timescales. With the same amount of data, it was more stable for annual forecasts than monthly, or daily forecasts, the R^2 and RMSE reached 0.83 and 5.54 μg m^{-3}, respectively, in validation set (year of 2019). 3) Levels of $PM_{2.5}$ pollution gradually improved in 2019 after application of pollution control measures. Annual mean $PM_{2.5}$ concentrations were 27.33 ± 15.56 μg m^{-3}, and more than 80% of regions met the CNAAQS Level 2 standard, although there were still some areas with significant pollution, including the North China Plain, parts of the Loess Plateau, and the Taklimakan Desert. 4) The high levels of $PM_{2.5}$ pollution can be related to increasingly frequent anthropogenic emissions (i.e., industrial pollution, fuel combustion, and secondary production), unfavorable terrain for air circulation, and meteorological factors (including precipitation and wind speed).

Based on the LSTM model, conflating high-resolution AOD products, meteorological and land-cover-related indices could achieve higher precision of predictions of future $PM_{2.5}$, and generate high-resolution (1 km) $PM_{2.5}$ spatial distribution maps. This can provide more detailed national, small and medium-scale $PM_{2.5}$ information, and support and improve air quality management and control in China.

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

This work was supported by the Science and Technology Innovation Leading Talent Program of Zhejiang Province (2019R52004), and the National Key Research and Development Program (2017YFD0700501).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2021.126493.

Uncited references

; van Donkelaar et al.(2010).

References

[Alston et al., 2012] E.J. Alston, I.N. Sokolik, O.V. Kalashnikova, Characterization of atmospheric aerosol in the US Southeast from ground- and space-based measurements over the past decade, Atmospheric Measurement Techniques 5 (2012) 1667–1682.

[Bai et al., 2019] L. Bai, L. Jiang, D.Y. Yang, Y.B. Liu, Quantifying the spatial heterogeneity influences of natural and socioeconomic factors and their interactions on air pollution using the geographical detector method: a case study of the Yangtze River Economic Belt, China, J. Clean. Prod. 232 (2019) 692–704.

[Bai et al., 2016] Y. Bai, L. Wu, K. Qin, Y. Zhang, Y. Shen, Y. Zhou, A geographically and temporally weighted regression model for ground-level $PM_{2.5}$ estimation from satellite-derived 500 m resolution AOD, Rem. Sens. 8 (2016).

[Che et al., 2014] H. Che, X. Xia, J. Zhu, Z. Li, O. Dubovik, B. Holben, P. Goloub, H. Chen, V. Estelles, E. Cuevas-Agulló, L. Blarel, H. Wang, H. Zhao, X. Zhang, Y. Wang, J. Sun, R. Tao, X. Zhang, G. Shi, Column aerosol optical properties and aerosol radiative forcing during a serious haze-fog month over North China Plain in 2013 based on ground-based sunphotometer measurements, Atmos. Chem. Phys. 14 (2014) 2125–2138.

[Chen et al., 2018] G. Chen, S. Li, L.D. Knibbs, N.A.S. Hamm, W. Cao, T. Li, J. Guo, H. Ren, M.J. Abramson, Y. Guo, A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological and land use information, Sci. Total Environ. 636 (2018) 52–60.

[Di et al., 2016] Q. Di, I. Kloog, P. Koutrakis, A. Lyapustin, Y.J. Wang, J. Schwartz, Assessing PM2.5 exposures with high spatiotemporal resolution across the continental United States, Environ. Sci. Technol. 50 (2016) 4712–4721.

[Engel-Cox et al., 2004] J.A. Engel-Cox, C.H. Holloman, B.W. Coutant, R.M. Hoff, Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality, Atmos. Environ. 38 (2004) 2495–2509.

[Fan et al., 2017] J. Fan, Q. Li, J. Hou, X. Feng, H. Karimian, S. Lin, A spatiotemporal prediction framework for air pollution based on deep RNN. ISPRS annals of photogrammetry, Remote Sensing and Spatial Information Sciences IV-4/W2 (2017) 15–22.

[Fang et al., 2016] X. Fang, B. Zou, X. Liu, T. Sternberg, L. Zhai, Satellite-based ground $PM_{2.5}$ estimation using timely structure adaptive modeling, Remote Sens. Environ. 186 (2016) 152–163.

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

[Guo et al., 2014] S. Guo, M. Hu, M.L. Zamora, J. Peng, D. Shang, J. Zheng, Z. Du, Z. Wu, M. Shao, L. Zeng, M.J. Molina, R. Zhang, Elucidating severe urban haze formation in China, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 17373–17378.

[Gupta et al., 2006] P. Gupta, S.A. Christopher, J. Wang, R. Gehrig, Y. Lee, N. Kumar, Satellite remote sensing of particulate matter and air quality assessment over global cities, Atmos. Environ. 40 (2006) 5880–5892.

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

[Hochreiter and Schmidhuber, 1997] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.

[Hu et al., 2014] X. Hu, L.A. Waller, A. Lyapustin, Y. Wang, Y. Liu, 10-year spatial and temporal trends of $PM_{2.5}$ concentrations in the southeastern US estimated using high-resolution satellite data, Atmos. Chem. Phys. 14 (2014) 6301–6314.

[Huang et al., 2014] R.J. Huang, Y. Zhang, C. Bozzetti, K.F. Ho, J.J. Cao, Y. Han, K.R. Daellenbach, J.G. Slowik, S.M. Platt, F. Canonaco, P. Zotter, R. Wolf, S.M. Pieber, E.A. Bruns, M. Crippa, G. Ciarelli, A. Piazzalunga, M. Schwikowski, G. Abbaszade, J. Schnelle-

Kreis, R. Zimmermann, Z. An, S. Szidat, U. Baltensperger, I. El Haddad, A.S. Prevot, High secondary aerosol contribution to particulate pollution during haze events in China, Nature 514 (2014) 218–222.

[Jiang et al., 2020] H. Jiang, H. Hu, R. Zhong, J. Xu, J. Xu, J. Huang, S. Wang, Y. Ying, T. Lin, A deep learning approach to conflating heterogeneous geospatial data for corn yield estimation: a case study of the US Corn Belt at the county level, Global Change Biol. 26 (2020) 1754–1766.

[Kahn et al., 2010] R.A. Kahn, B.J. Gaitley, M.J. Garay, D.J. Diner, T.F. Eck, A. Smirnov, B.N. Holben, Multiangle imaging SpectroRadiometer global aerosol product assessment by comparison with the aerosol robotic network, J. Geophys. Res. 115 (2010).

[Kampa and Castanas, 2008] M. Kampa, E. Castanas, Human health effects of air pollution, Environ. Pollut. 151 (2008) 362–367.

[Koelemeijer et al., 2006] R.B.A. Koelemeijer, C.D. Homan, J. Matthijsen, Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmos. Environ. 40 (2006) 5304–5315.

[Li et al., 2017] T. Li, H. Shen, C Zeng, et al., Point-surface fusion of station measurements and satellite observations for mapping $PM_{2.5}$ distribution in China: Methods and assessment, Atmospheric Environment 152 (2017).

[Li et al., 2009] B. Li, H. Yuan, N. Feng, S. Tao, Comparing MODIS and AERONET aerosol optical depth over China, Int. J. Rem. Sens. 30 (2009) 6519–6529.

[Li et al., 2015a] J. Li, B.E. Carlson, A.A. Lacis, How well do satellite AOD observations represent the spatial and temporal variability of $PM_{2.5}$ concentration for the United States?, Atmos. Environ. 102 (2015a) 260–273.

[Li et al., 2017a] T. Li, H. Shen, Q. Yuan, X. Zhang, L. Zhang, Estimating ground-level $PM_{2.5}$ by fusing satellite and station observations: a geo-intelligent deep learning approach, Geophys. Res. Lett. 44 (11) (2017a) 985-911,993.

[Li et al., 2017b] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, T. Chi, Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation, Environ. Pollut. 231 (2017b) 997–1004.

[Li et al., 2015b] Y. Li, Q. Chen, H. Zhao, L. Wang, R. Tao, Variations in PM_{10}, $PM_{2.5}$ and $PM_{1.0}$ in an urban area of the Sichuan Basin and their relation to meteorological factors, Atmosphere 6 (2015b) 150–163.

[Liang et al., 2016] C.S. Liang, F.K. Duan, K.B. He, Y.L. Ma, Review on recent progress in observations, source identifications and countermeasures of $PM_{2.5}$, Environ. Int. 86 (2016) 150–170.

[Liu et al., 2019] W. Liu, G. Guo, F. Chen, Y. Chen, Meteorological pattern analysis assisted daily $PM_{2.5}$ grades prediction using SVM optimized by PSO algorithm, Atmospheric Pollution Research 10 (2019) 1482–1491.

[Lyapustin et al., 2018] A. Lyapustin, Y. Wang, S. Korkin, D. Huang, MODIS collection 6 MAIAC algorithm, Atmospheric Measurement Techniques 11 (2018) 5741–5765.

[Ma et al., 2014] Z. Ma, X. Hu, L. Huang, J. Bi, Y. Liu, Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing, Environ. Sci. Technol. 48 (2014) 7436–7444.

[Ma et al., 2016] Z. Ma, X. Hu, A. Sayer, R. Levy, Q. Zhang, Y. Liu, Satellite-based spatiotemporal trends in $PM_{2.5}$ concentrations: China, 2004–2013, Environ. Health Perspect. 124 (2) (2016) 184–192.

[Ma et al., 2017] Z. Ma, Z. Shi, Y. Zhou, J. Xu, W. Yu, Y. Yang, A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed, Rem. Sens. Environ. 200 (2017) 378–395.

[Meng et al., 2015] R. Meng, F. Zhao, K. Sun, R. Zhang, C. Huang, J. Yang, Analysis of the 2014 “APEC blue” in beijing using more than one decade of satellite observations: lessons learned from radical emission control measures, Rem. Sens. 7 (2015) 15224–15243.

[Pak et al., 2020] U. Pak, J. Ma, U. Ryu, K. Ryom, U. Juhyok, K. Pak, C. Pak, Deep learning-based $PM_{2.5}$ prediction considering the spatiotemporal correlations: a case study of Beijing, China, Sci. Total Environ. 699 (2020) 133561.

[Pang et al., 2018] J. Pang, Z. Liu, X. Wang, J. Bresch, J. Ban, D. Chen, J. Kim, Assimilating AOD retrievals from GOCI and VIIRS to forecast surface $PM_{2.5}$ episodes over Eastern China, Atmos. Environ. 179 (2018) 288–304.

[Peng et al., 2019] J. Peng, A. Biswas, Q. Jiang, R. Zhao, J. Hu, B. Hu, Z. Shi, Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China, Geoderma 337 (2019) 1309–1319.

[Peng et al., 2016] J. Peng, S. Chen, H. Lü, Y. Liu, J. Wu, Spatiotemporal patterns of remotely sensed $PM_{2.5}$ concentration in China from 1999 to 2011, Rem. Sens. Environ. 174 (2016) 109–121.

[Philip et al., 2014] S. Philip, R.V. Martin, A. van Donkelaar, J.W. Lo, Y. Wang, D. Chen, L. Zhang, P.S. Kasibhatla, S. Wang, Q. Zhang, Z. Lu, D.G. Streets, S. Bittman, D.J. Macdonald, Global chemical composition of ambient fine particulate matter for exposure assessment, Environ. Sci. Technol. 48 (2014) 13060–13068.

[Qi et al., 2019] Y. Qi, Q. Li, H. Karimian, D. Liu, A hybrid model for spatiotemporal forecasting of $PM_{2.5}$ based on graph convolutional neural network and long short-term memory, Sci. Total Environ. 664 (2019) 1–10.

[Raaschou-Nielsen et al., 2013] O. Raaschou-Nielsen, Z.J. Andersen, R. Beelen, E. Samoli, M. Stafoggia, G. Weinmayr, B. Hoffmann, P. Fischer, M.J. Nieuwenhuijsen, B. Brunekreef, W.W. Xun, K. Katsouyanni, K. Dimakopoulou, J. Sommar, B. Forsberg, L. Modig, A. Oudin, B. Oftedal, P.E. Schwarze, P. Nafstad, U. De Faire, N.L. Pedersen, C.-G. Östenson, L. Fratiglioni, J. Penell, M. Korek, G. Pershagen, K.T. Eriksen, M. Sørensen, A. Tjønneland, T. Ellermann, M. Eeftens, P.H. Peeters, K. Meliefste, M. Wang, B. Bueno-de-Mesquita, T.J. Key, K. de Hoogh, H. Concin, G. Nagel, A. Vilier, S. Grioni, V. Krogh, M.-Y. Tsai, F. Ricceri, C. Sacerdote, C. Galassi, E. Migliore, A. Ranzi, G. Cesaroni, C. Badaloni, F. Forastiere, I. Tamayo, P. Amiano, M. Dorronsoro, A. Trichopoulou, C. Bamia, P. Vineis, G. Hoek, Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE), Lancet Oncol. 14 (2013) 813–822.

[Rossel and Chen, 2011] R.A.V. Rossel, C. Chen, Digitally mapping the information content of visible–near infrared spectra of surficial Australian soils, Rem. Sens. Environ. 115 (2011) 1443–1455.

[Saide et al., 2011] P.E. Saide, G.R. Carmichael, S.N. Spak, L. Gallardo, A.E. Osses, M.A. Mena-Carrasco, M. Pagowski, Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model, Atmos. Environ. 45 (2011) 2769–2780.

[Tsai et al., 2011] T.-C. Tsai, Y.-J. Jeng, D.A. Chu, J.-P. Chen, S.-C. Chang, Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008, Atmos. Environ. 45 (2011) 4777–4788.

[van Donkelaar et al., 2010] A. van Donkelaar, R.V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, P.J. Villeneuve, Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application, Environ. Health Perspect. 118 (2010) 847–855.

[Wang and Chen, 2016] B. Wang, Z. Chen, High-resolution satellite-based analysis of ground-level $PM_{2.5}$ for the city of Montreal, Sci. Total Environ. 541 (2016) 1059–1069.

[Wang et al., 2016] J.F. Wang, T.L. Zhang, B.J. Fu, A measure of spatial stratified heterogeneity, Ecol. Indicat. 67 (2016) 250–256.

[Wang et al., 2017] S. Wang, C. Zhou, Z. Wang, K. Feng, K. Hubacek, The characteristics and drivers of fine particulate matter ($PM_{2.5}$) distribution in China, J. Clean. Prod. 142 (2017) 1800–1809.

[Wang et al., 2018] X. Wang, R.E. Dickinson, L. Su, C. Zhou, K. Wang, $PM_{2.5}$ pollution in China and how it has been exacerbated by terrain and meteorological conditions, Bull. Am. Meteorol. Soc. 99 (2018) 105–119.

[Wei et al., 2019] J. Wei, W. Huang, Z. Li, W. Xue, Y. Peng, L. Sun, M. Cribb, Estimating 1-km-resolution $PM_{2.5}$ concentrations across China using the space-time random forest approach, Rem. Sens. Environ. (2019) 231.

[Wen et al., 2019] C. Wen, S. Liu, X. Yao, L. Peng, X. Li, Y. Hu, T. Chi, A novel spatiotemporal convolutional long short-term neural network for air pollution prediction, Sci. Total Environ. 654 (2019) 1091–1099.

[Wu et al., 2020] X.L. Wu, Y. Wang, S. He, Z. Wu, $PM_{2.5}/PM_{10}$ ratio prediction based on a long short-term memory neural network in Wuhan, China, Geosci. Model Dev. (GMD) 13 (2020) 1499–1511.

[Xiao et al., 2018] L. Xiao, Y. Lang, G. Christakos, High-resolution spatiotemporal mapping of $PM_{2.5}$ concentrations at Mainland China using a combined BME-GWR technique, Atmos. Environ. 173 (2018) 295–305.

[Yu et al., 2017] W. Yu, Y. Liu, Z. Ma, J. Bi, Improving satellite-based $PM_{2.5}$ estimates in China using Gaussian processes modeling in a Bayesian hierarchical setting, Sci. Rep. 7 (2017) 7048.

[Zhang et al., 2012] H. Zhang, J. Li, Q. Ying, J.Z. Yu, D. Wu, Y. Cheng, K. He, J. Jiang, Source apportionment of $PM_{2.5}$ nitrate and sulfate in China using a source-oriented chemical transport model, Atmos. Environ. 62 (2012) 228–242.

[Zhang et al., 2019] X. Zhang, Q. Wang, T. Qiu, S. Tang, J. Li, J.P. Giesy, Y. Zhu, X. Hu, D. Xu, $PM_{2.5}$ bound phthalates in four metropolitan cities of China: concentration, seasonal pattern and health risk via inhalation, Sci. Total Environ. 696 (2019) 133982.

[Zhang and Li, 2015] Y. Zhang, Z. Li, Remote sensing of atmospheric fine particulate matter ($PM_{2.5}$) mass concentration near the ground from satellite observation, Rem. Sens. Environ. 160 (2015) 252–262.

[Zhang and Cao, 2015] Y.L. Zhang, F. Cao, Fine particulate matter ($PM_{2.5}$) in China at a city level, Sci. Rep. 5 (2015) 14884.

[Zhao et al., 2013] P.S. Zhao, F. Dong, D. He, X.J. Zhao, X.L. Zhang, W.Z. Zhang, Q. Yao, H.Y. Liu, Characteristics of concentrations and chemical compositions for $PM_{2.5}$ in the region of Beijing, Tianjin, and Hebei, China, Atmos. Chem. Phys. 13 (2013) 4631–4644.

[Zhou et al., 2019] Y. Zhou, J. Xue, S. Chen, Y. Zhou, Z. Liang, N. Wang, Z. Shi, Fine-resolution mapping of soil total nitrogen across China based on weighted model averaging, Rem. Sens. 12 (2019).