

# Predicting annual PM<sub>2.5</sub> in mainland China from 2014 to 2020 using multi-temporal satellite product: An improved deep learning approach with spatial generalization ability

Zhige Wang <sup>a</sup>, Bifeng Hu <sup>b</sup>, Bo Huang <sup>c,g</sup>, Ziqiang Ma <sup>d</sup>, Asim Biswas <sup>e</sup>, Yefeng Jiang <sup>a</sup>, Zhou Shi <sup>a,f,\*</sup>

<sup>a</sup> Institute of Agricultural Remote Sensing and Information Technology Application, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

<sup>b</sup> Department of Land Resource Management, School of Tourism and Urban Management, Jiangxi University of Finance and Economics, Nanchang 330013, China

<sup>c</sup> Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong

<sup>d</sup> Institute of Remote Sensing and Geographical Information Systems, School of Earth and Space Sciences, Peking University, Beijing 100871, China

<sup>e</sup> School of Environmental Sciences, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada

<sup>f</sup> Key Laboratory of Environment Remediation and Ecological Health, Ministry of Education, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

<sup>g</sup> Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong



## ARTICLE INFO

### Keywords:

PM<sub>2.5</sub> prediction

AOD

Attention mechanism

Bidirectional LSTM

Spatiotemporal variation

## ABSTRACT

The estimation of long-term fine particulate matter (PM<sub>2.5</sub>) concentrations and trend assessments play a critical role in preventing health risks to the human body. Deep learning methods based on time series have been highly accurate in predicting PM<sub>2.5</sub>. However, most time series models lack the ability for spatial generalization because they cannot combine the analysis at different spatial scales. In addition, temporal long-term trend analysis have not been reported in most of the studies. In this study, to reveal the spatiotemporal variability and trends of PM<sub>2.5</sub>, an improved deep learning framework named the SpatioTemporal Enhanced Neural Network (STENN) is developed for estimating PM<sub>2.5</sub> concentrations with a spatial resolution of 1 km. Based on the bidirectional long short-term memory (LSTM) structure and attention mechanisms, the model provides a geographic-data-driven approach to incorporate the impact of the spatial heterogeneity and time dependence of PM<sub>2.5</sub>, which demonstrates that it has robust spatiotemporal transferable power with an R<sup>2</sup> of 0.89 produced by cross validation (CV). High-resolution (1 km) and high-quality annual PM<sub>2.5</sub> products for mainland China from 2014 to 2020 were constructed. In comparison with the current 1-km PM<sub>2.5</sub> products, our framework demonstrates better stability in different regions, especially in terms of the high-value estimations and spatial continuity. The spatiotemporal PM<sub>2.5</sub> distributions were also analyzed based on the time-series products. After the implementation of various control policies for atmospheric pollution, a declining trend of PM<sub>2.5</sub> concentrations was observed in 88.79% of China between 2014 and 2020, with a mean decrease rate of 3.35  $\mu\text{g m}^{-3} \text{yr}^{-1}$ . This result indicates that the control policies of the Chinese government were effective in reducing PM<sub>2.5</sub> concentrations. The PM<sub>2.5</sub> concentrations in China reveal an exponential temporal trend, from a rapid decline to a gradual slowdown and a stable phase. To realize the vision of the Beautiful China Initiative, a regionally targeted policy for air pollution management is required. This study provides valuable implications for a more detailed analysis of the spatiotemporal variations in PM<sub>2.5</sub> at small and medium spatial scales by developing an improved deep learning approach with a spatial generalization ability and integration with multi-temporal satellite products.

\* Corresponding author at: 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](mailto:shizhou@zju.edu.cn) (Z. Shi).

## 1. Introduction

Ambient PM<sub>2.5</sub> (particulate matter with an aerodynamic diameter less than or equal to 2.5 μm) predominantly originates from fuel combustion and other anthropogenic emissions. The PM<sub>2.5</sub> was responsible for 2.9 million deaths worldwide in 2013 according to the Global Burden of Disease (GBD) (Brauer et al., 2016). It has been reported that ambient PM<sub>2.5</sub> was the fifth-ranking mortality risk factor in 2015 (Cohen et al., 2017). In recent years, the Chinese government has issued a series of policies (e.g., Action Plan on Air Pollution Prevention and Control, three-year Blue-Sky Defense) (Wang et al., 2020b; Zhang et al., 2019) for the prevention and control of PM<sub>2.5</sub> to limit its adverse environmental and health effects in China (Chen and Chen, 2021; Wang et al., 2020a; Zhang et al., 2021). Thus, the number of PM<sub>2.5</sub>-observation stations in China has been increasing annually since 2013. However, the geographic coverage of these monitoring networks remains limited and uneven, which leads to a lack of reliable information regarding the concentrations of PM<sub>2.5</sub>.

Recently, the development of aerosol observations from satellite remote sensing has made it possible to estimate the PM<sub>2.5</sub> distribution with broad spatial coverage and acceptable spatial resolution. The satellite aerosol optical depth (AOD) measures the light extinction of aerosols over the entire atmospheric column (Ceca et al., 2018). Because of the correlation between the aerosol optical properties and particle size distribution, composition, and shape (Engel-Cox et al., 2004; Gupta and Christopher 2008; Li et al., 2015a), a series of satellite-based AOD products (e.g., Advanced Himawari Imagery, AHI; Moderate Resolution Imaging Spectroradiometer, MODIS; Multi-angle Imaging Spectroradiometer, MISR; Visible Infrared Imaging Radiometer Suite, VIIRS) have been used to estimate PM<sub>2.5</sub> (Chen et al., 2020a; Franklin et al., 2017; Liu et al., 2019b; Meng et al., 2015; Pang et al., 2018; Wei et al., 2018; Yao et al., 2019; Zhang and Li, 2015). Recently, the emerging high-resolution (spatial resolution of 1 km and produced daily) Multi-Angle of Implementation of Atmospheric Correction (MAIAC) AOD datasets (Lyapustin et al., 2018) derived from both Terra and Aqua MODIS are promising resources for estimating PM<sub>2.5</sub> with a fine spatial resolution.

Many studies have developed various statistical methods for PM<sub>2.5</sub>, based on the relationships between PM<sub>2.5</sub> and satellite-based AOD (Engel-Cox et al., 2004; Guo et al., 2009). The methods range from simple linear regression, which only considers the AOD-PM<sub>2.5</sub> relationship (Schaap et al., 2009) to the multiple linear regression (Ma et al., 2016; Zhang et al., 2018; Xiao et al., 2017) and have been further developed for geostatistical regression that considers the spatial and temporal heterogeneity (e.g., geographically weighted regression (GWR) model and geographically and temporally weighted regression (GTWR) model) (He and Huang, 2018; Song et al., 2014; van Donkelaar et al., 2016; Xiao et al., 2018). Recently, machine learning models have been widely used to estimate PM<sub>2.5</sub> (Li et al., 2015b; Liu et al., 2019a; Liu et al., 2019c; Stafoggia et al., 2019; Yang et al., 2020). Nevertheless, the meteorological (e.g., boundary layer height, relative humidity, and wind speed) and anthropogenic determinants (e.g., industrial and vehicle emissions and agriculture fertilizer application) related to PM<sub>2.5</sub> are diverse, leading to strong spatial and temporal variations in PM<sub>2.5</sub> (Ma et al., 2019a; Zhao et al., 2019). Statistical regression methods, with weak feature learning abilities, cannot establish a stable relationship between PM<sub>2.5</sub> and AOD, especially at large spatial scales (Ma et al., 2014). Meanwhile, machine learning models can fit nonlinear relationships well (Hu et al., 2020; Fu et al., 2021; Hu et al., 2021). They are based on only features modeling with large amounts of data (Brokamp et al., 2018), whereas the geographical and temporal features hidden in time series and spatial distribution are underutilized.

In recent years, deep learning has achieved notable success in discovering the potential and intricate relationships from the temporal sequences and spatial data (Ma et al., 2019b; Mountrakis et al., 2018; Reichstein et al., 2019). In particular, the recurrent neural network (RNN) and its variant long short-term memory (LSTM) are two of the

most prominent deep learning models that specialize in linking adjacent observations, recognizing time variation patterns over a long time series, and capturing intricate nonlinear relationships (Fan et al., 2017; Hochreiter & Schmidhuber, 1997). Existing studies have confirmed that PM<sub>2.5</sub> exhibits a strong time dependency. Therefore, some researchers have applied deep learning methods to estimate and predict PM<sub>2.5</sub> using long time-series data. Qi et al. (2019) used the LSTM model to estimate hourly PM<sub>2.5</sub>, and validated the superiority of deep learning models. To further improve the spatiotemporal learning power of the neural networks, multiple neural networks (e.g., convolutional neural network (CNN), artificial neural network (ANN), LSTM) have been combined to predict the daily PM<sub>2.5</sub> concentrations (Pak et al., 2020; Soh et al., 2018). On this basis, the multiple data types (i.e., AOD and gaseous pollutant data) are integrated into combined CNN and RNN models by Wu et al. (2020). However, the weak spatial generalization ability of deep learning models limits the analysis of the spatially continuous changes and geographic distribution of PM<sub>2.5</sub> at the national scale. Furthermore, deep learning models for PM<sub>2.5</sub> predictions have mostly been utilized for short time scales (i.e., daily and hourly) (Li et al., 2017b; Wen et al., 2019). In addition, the potential of deep learning in long-term trend analysis has not yet been explored. More importantly, most RNN models are based on time series, and they ignore the spatial heterogeneity and spatial correlation (Wang et al., 2021), which have a strong influence on geographic objects (i.e., PM<sub>2.5</sub>) (Tobler, 1970). Therefore, it is critical to build a robust model that can be widely used to process remote sensing data with both spatial distribution characteristics and multi-temporal phases to estimate PM<sub>2.5</sub>.

The key aim of the present study was to develop an improved LSTM-based method with spatial generalization for annual PM<sub>2.5</sub> estimations across China. In order to overcome the drawbacks of the RNN models mentioned above, in this study, we constructed a new framework called the SpatioTemporal Enhanced Neural Network (STENN), which introduced the concept of the spatial signals. Under this framework, the model accuracy and spatial and temporal transferability were also validated. By integrating the historical multi-temporal remote sensing data (i.e., AOD, meteorological, land cover/use, and socio-economic factors for 2011–2019), we produced annual PM<sub>2.5</sub> products at a 1-km resolution for China. We investigated the long-term spatiotemporal variations of PM<sub>2.5</sub> across China since the implementation of the air pollution control policies. Overall, in this study, we aimed to address four specific research questions regarding the robust PM<sub>2.5</sub> estimations and variation analysis:

1. How well does the STENN model performance for PM<sub>2.5</sub> estimations compared with previous studies?
2. How well does the model perform for extrapolating PM<sub>2.5</sub> for the years and regions without ground-based monitoring?
3. Do the PM<sub>2.5</sub> products produced in this study have reliable accuracy and stability?
4. How did the PM<sub>2.5</sub> distribution in China change between 2014 and 2020?

## 2. Materials and methods

The workflow of this study is presented in Fig. 1. The data used in the study included AOD products, as well as ten variables related to meteorology, land cover, and socio-economy to capture their relationships with PM<sub>2.5</sub> concentrations.

### 2.1. PM<sub>2.5</sub> measurements

The 24-hour average PM<sub>2.5</sub> measurements at ground-based observed stations across China from 2013 to 2020 were obtained from the China National Environmental Monitoring Center (CNEMC, <http://www.cnemc.cn/>). The number of stations increased from 835 in 2013 to 1,700 by the end of 2020. As indicated in Fig. 2, more monitoring

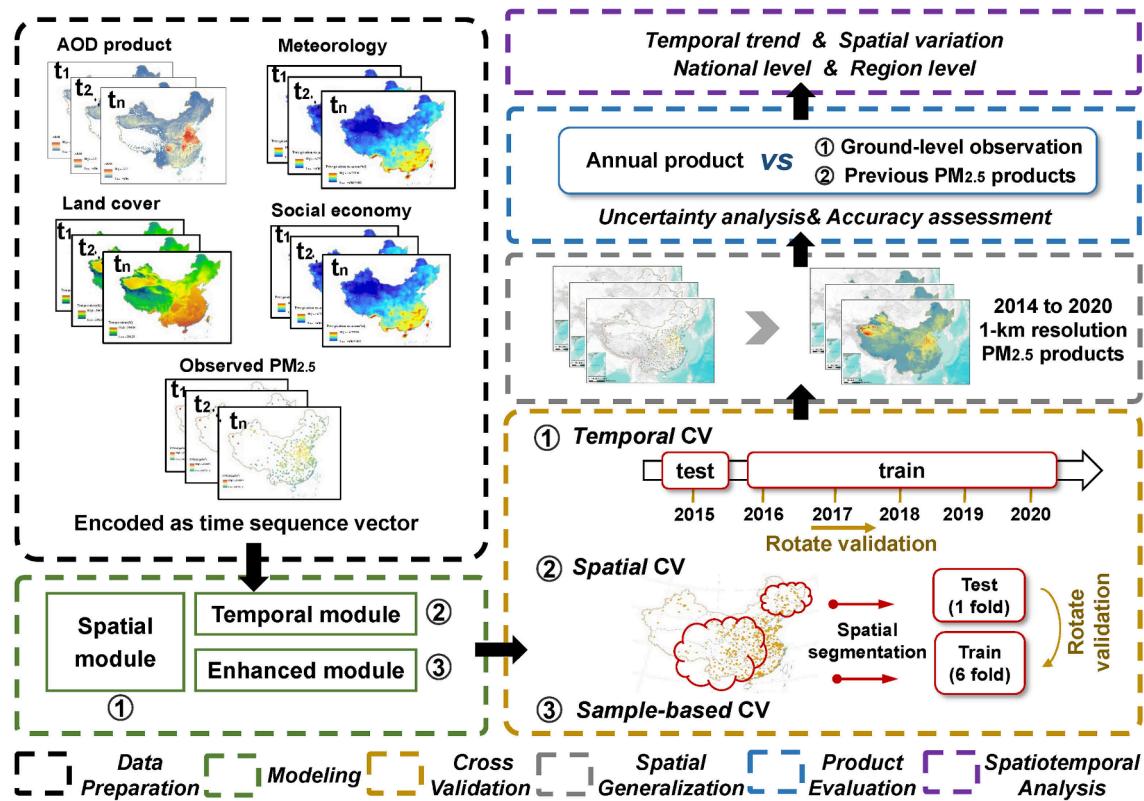


Fig. 1. Workflow of current study.

stations are distributed in northern and southeastern China, with fewer stations distributed in southwestern and northwestern China.

## 2.2. MODIS MAIAC AOD data

MODIS has collected over 20 years of global columnar aerosol properties while aboard the Terra and Aqua since 2000 and 2002, respectively. Recently, the advanced Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm has been employed in the MODIS instrument, which provides daily AOD products with high spatial resolution. The MAIAC algorithm combines the time series and spatial analysis to help improve the quality of cloud and snow detection and ensure good performance for aerosol retrievals and atmospheric corrections over both dark vegetated surfaces and bright deserts (Lyapustin et al., 2018). The MODIS Collection 6 MAIAC AOD product (MCD19A2) in China from 2011 to 2019 was obtained from the Land Processes Distributed Active Archive Center (LP DAAC, <https://ladsweb.modaps.eosdis.nasa.gov/>) in our study. The data are produced daily at a 1-km resolution and provide quality assurance (QA) to ensure retrieval quality. The AERONET AOD data were also used to validate the accuracy of the MAIAC AOD (Text S1 and Fig. S1 for details). Subsequently, the 550 nm wavelength AOD was used following projection definition and reprojection. Finally, the AOD data were mosaicked to obtain the products for all of China.

## 2.3. Ancillary data

### 2.3.1. Meteorological data

The distribution and chemical and optical properties of PM<sub>2.5</sub> are strongly affected by the meteorological conditions (Chen et al., 2020b; Liu et al., 2009). Consequently, we selected six variables related to meteorology for 2011 to 2019: surface air relative humidity (RHU), 2-m surface air temperature (TEMP), surface pressure (P), precipitation flux (PRE), boundary layer height (BLH), and 10-m wind speed (WS). The

data were obtained from the reanalysis products (Hersbach et al., 2020) developed at the European Center for Medium Weather Forecasting (ECMWF, <https://www.ecmwf.int/>). The dataset was tested for high reliability in China (Guo et al., 2016; Liu et al., 2021; Zhou et al., 2018).

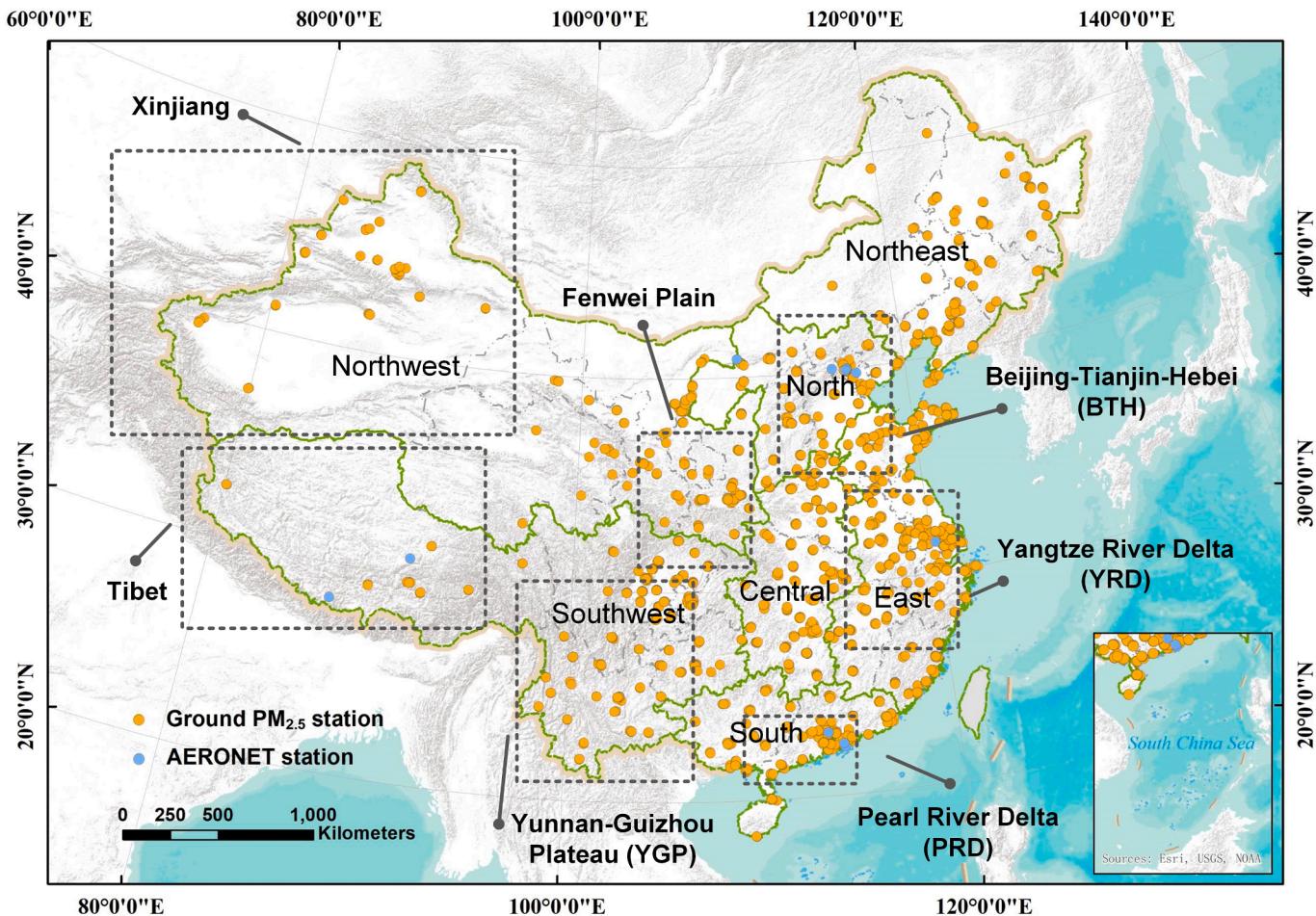
### 2.3.2. Land cover related and socio-economic data

The land cover and socio-economic variables, including the normalized difference vegetation index (NDVI), digital elevation model (DEM), land use/cover change (LUCC), and population distribution (POP) obtained from the Resource and Environment Science and Data Center (<http://www.resdc.cn/>), were used in our study to better estimate the distribution of PM<sub>2.5</sub> (Table S2 for details and sources of these data).

A correlation significance test was conducted to ensure a significant correlation between the PM<sub>2.5</sub> concentrations and the selected variables. All the explanatory variables passed the correlation significance test (significant confidence at the 0.01 level). AOD ( $r = 0.497$ ) and PRE ( $r = -0.302$ ) were significantly correlated with the PM<sub>2.5</sub> concentrations (Table S3). To test the sensitivity of the selected variables, the idea of Gini index (Nembrini et al., 2018) was adopted. The basic idea is to remove each variable in turn and observe the change (increase/decrease) in the residual sum of squares (RSS) to quantify the importance of that variable.

## 2.4. Data processing

The 24-hour PM<sub>2.5</sub> observations were averaged over a year for the annual estimation. To encode the data into a time series format for inputting into the temporal module structure, the discontinuous PM<sub>2.5</sub> station data were removed. For the model fitting and testing, the multi-source remote sensing products (i.e., AOD, meteorological, land use/cover, and socio-economic data) were first preprocessed to match the annual PM<sub>2.5</sub> ground-level observations. The final number of matched samples was 10,647 from 2013 to 2020. For the model estimation, the



**Fig. 2.** Locations of PM<sub>2.5</sub> ground monitoring stations (orange dots), Aerosol Robotic Network stations (green dots), and classification of seven geographical regions in China (green borders) and the main regions mentioned in the study. The Fenwei Plain includes parts of Shaanxi Province, Shanxi Province and Henan Province. The YRD includes Shanghai, Jiangsu Province, Zhejiang Province and Anhui Province. The PRD refers to parts of Guangdong Province. The YGP mainly covers Guizhou Province, Yunnan Province, northern Guangxi, and the border areas of Sichuan province. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bilinear interpolation method was used to resample the meteorological, LUCC, and DEM data with 1-km grid cells. The annual AOD values and meteorological data were calculated using the arithmetic average method to ensure consistency in the spatiotemporal resolution of environmental variables (the spatial and temporal resolutions are 1 km and annual, respectively). The data were processed using ArcGIS 10.3 and R.

### 2.5. Deep learning framework for PM<sub>2.5</sub> estimation

To account for the complex temporal dependency and spatial heterogeneity of PM<sub>2.5</sub>, a three-component spatiotemporal enhanced neural network (STENN) framework with the ability to predict PM<sub>2.5</sub> concentrations across China was developed (Fig. 3). This improved deep learning model utilizes the multi-temporal and multi-source remote sensing products as well as PM<sub>2.5</sub> ground observations, as inputs, acquires spatial signals according to the spatial correlation, obtains temporal features through the LSTM structure, and enhances the spatiotemporal features through the attention mechanism. Finally, the PM<sub>2.5</sub> concentrations at the next time step are predicted as outputs. Furthermore, the framework also has the capacity for spatial generalization, which can extend the site-level PM<sub>2.5</sub> predictions to the national level with 1-km-resolution. The three components of the model framework are the spatial, temporal, and enhanced modules.

#### 2.5.1. Spatial module

Although deep learning models perform well in time series predictions, they need to be improved to handle the spatial correlation and heterogeneity of spatial objects in the spatial estimation of PM<sub>2.5</sub>. As a result, we first introduced the spatial signal (SS) in the spatial module (Li et al., 2017a). The SS can be calculated as:

$$SS_i = \frac{\sum_{j=1}^n \frac{1}{d_{ji}^2} X_j}{\sum_{j=1}^n \frac{1}{d_{ji}^2}} \quad (1)$$

where  $X_j$  refers to the PM<sub>2.5</sub> concentrations of point  $j$  which is near the target data point  $i$ ,  $d_{ji}$  is the distance between points  $i$  and  $j$ , and  $n$  represents the number of the nearest points to point  $i$ . The determination of  $n$  is essential for measuring the spatial correlation; we made different attempts which can be seen in the supporting information (Table S4). The performance of the model has no significant change when  $n$  changes and the model performs best when  $n$  is 3.

The  $d_{ij}$  is defined as the Euclidean distance, calculated using the coordinates of points  $i$  ( $x_i, y_i$ ) and point  $j$  ( $x_j, y_j$ ),

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2)$$

The spatial signal considers the influence of several near observed points  $j$  ( $j = 1, 2, \dots$ ) near the target point  $i$ , and uses distance weighting to enhance or weaken the influence of the nearby observed points.

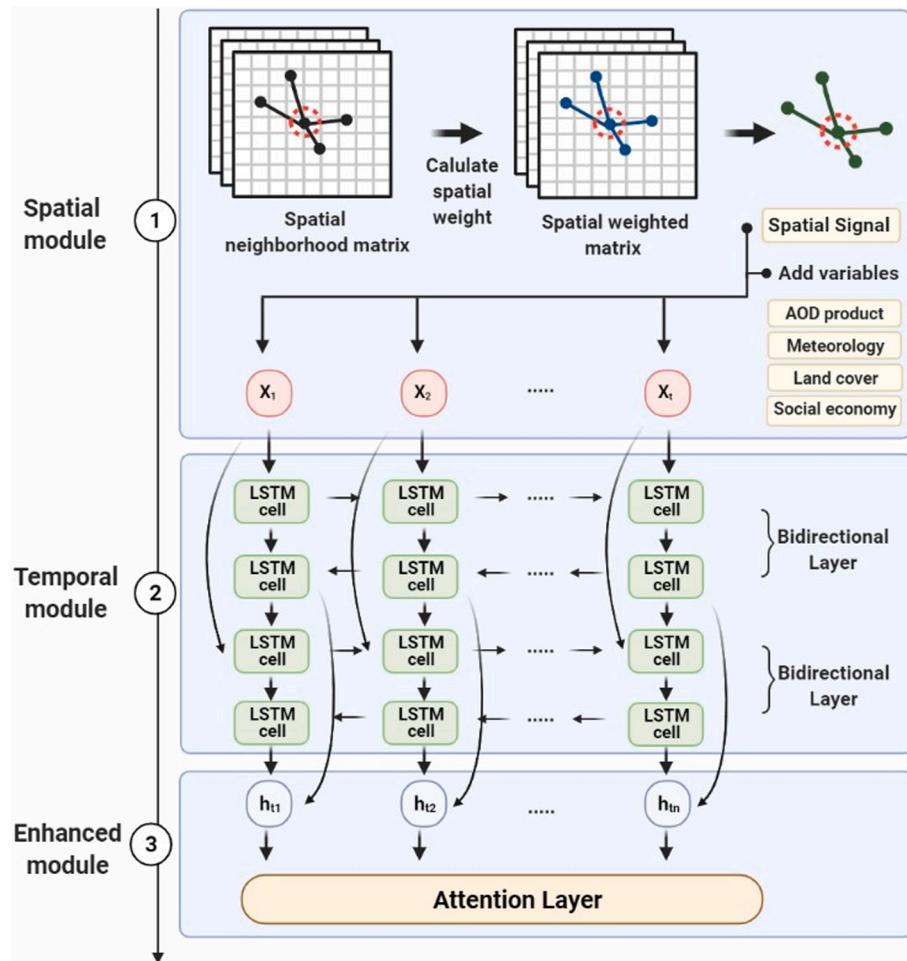


Fig. 3. The overall architecture of the STENN model for  $\text{PM}_{2.5}$  estimation.

Because observations nearby are more informative than those farther away (Fotheringham et al., 2002), the observed point  $j$  near to point  $i$  was assumed to have a greater effect on the  $\text{PM}_{2.5}$  estimation of point  $i$  than the points located farther from object point  $i$ .

The spatial signal, along with other preprocessed variables, at each time node  $t$  was then encoded as a vector  $x_t$ , that is,  $[v_1, v_2, \dots, v_n]$ , which consists of multiple environmental variables. The input of the predicted framework is the time sequence  $[x_1, x_2, \dots, x_T]$ , where  $T$  is the time step length fed into the model structure of the next component.

To validate whether the spatial signal can improve the  $\text{PM}_{2.5}$  estimation accuracy, we established a model without SS as a comparison. The results are presented in Table S5. There is a considerable decrease in the model accuracy without the spatial signal.

#### 2.5.2. Temporal module

To better capture the abstract temporal features from the input dataset, we built a Bidirectional LSTM (BiLSTM) model in the temporal module.

As an evolution of RNN, the LSTM model (Hochreiter and Schmidhuber, 1997) compensates for the exploding or vanishing gradient problem in the long sequence data processing of RNN (Zhang, 2019).

The key idea of LSTM is the adaptive gating mechanism, which determines whether and to what extent the status of the LSTM block is updated. Each LSTM block contains a forget gate  $f_t$ , which determines how much information is retained from time step  $t-1$ , an input gate  $i_t$ , which decides how much information is stored from the current time step  $t$ , a cell state  $C_t$ , which updates the current cell state, and an output gate  $o_t$ , which determines how much information in the current cell state

$C_t$  is transferred to output.

The LSTM equations at time step  $t$  can be demonstrated as:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * \tanh(C_t) \quad (7)$$

where  $f_t$ ,  $i_t$ ,  $o_t$ , and  $C_t$  denote the vectors of the forget gate, input gate, output gate, and cell state;  $W_f$ ,  $W_i$ ,  $W_C$  and  $W_o$  are the weights;  $b_f$ ,  $b_i$ ,  $b_c$ , and  $b_o$  represent the bias vector of the corresponding gate and cell state, respectively;  $h_{t-1}$  and  $h_t$  are the output vectors at different time steps  $t$  and  $t-1$ ;  $x_t$  is the current input,  $[h_{t-1}, x_t]$  combines two vectors into one longer vector;  $\sigma$  and  $\tanh$  are regarded as the activation functions.

Bidirectional LSTM (Bi-LSTM) is an extension of LSTM, which changes the defect of the time sequence only flowing forward in LSTM, allowing the time sequence to flow forward and backward (Xu et al., 2020; Zhou et al., 2016). In bidirectional LSTM, the output vector at time step  $t$  expands to two vectors:  $h_t^{\text{forward}}$  and  $h_t^{\text{backward}}$ . The two vectors from opposite time directions combine to create the final output vector  $[h_t^{\text{forward}}, h_t^{\text{backward}}]$ , as  $h_t$ .

#### 2.5.3. Enhanced module

After extracting the spatiotemporal information of the input features,

the attention layer is built to selectively focus on more important information from the multi-source environmental variables. The attention mechanism first obtains the attentional vector  $\tilde{h}_t$  by combining the information of the input vectors  $h_t$  and target  $y_t$ . Then, the Softmax function is used to calculate the weight value  $\alpha_t$ . Finally, the adjusted vectors  $h_t^*$  obtained from the weighted sum are based on the weight value  $\alpha_t$ . The calculation is as follows:

$$\tilde{h}_t = \tanh(W_C[y_t; h_t]) \quad (8)$$

$$\alpha_t = \text{softmax}(W_s \tilde{h}_t) \quad (9)$$

$$h_t^* = \sum_i \alpha_t h_t \quad (10)$$

where the  $W_C$  and  $W_s$  are the model parameters for the calculation,  $\tanh$  represents the activation function (Luong et al., 2015).

In this study, we used Keras, a deep learning API in Python, to employ the STENN model to predict PM<sub>2.5</sub>. All data were first scaled by calculating the mean and standard deviation of the dataset being fed into the model. The temporal module contained three bidirectional LSTM layers with 512 hidden units, and there was one attention layer in the enhanced module. To prevent overfitting, we set one dropout layer in the enhanced module. More details related to the deployment of the model are provided in Table S6.

Because the environmental variables are all derived from multi-source satellite-based products and are available at a fine spatial resolution, it is possible for spatial generalization. These derived variables could be further used to feed into the model and they serve to map the PM<sub>2.5</sub> concentrations at a 1-km spatial resolution.

## 2.6. Model assessment and product evaluation

Three cross-validation scheme methods were adopted to assess the performance of the STENN model: (1) temporal cross-validation (CV), (2) spatial CV, and (3) sample-based CV. To better understand the spatiotemporal transferability of the STENN model, model validation was conducted on the temporal and spatial scales. The temporal CV was performed by omitting one year out for validation, and the others were used for model fitting. The spatial CV was performed by dividing the dataset into calibration and validation sets according to the geographical division of China (Fig. 2) to conduct cross-validation. In the sample-based CV, the dataset was randomly divided into 10 folds randomly. One fold was used for validation and the remaining nine folds were used as calibration folds, and then they were rotated until ten sets were used for the validation once again.

The prediction results of PM<sub>2.5</sub> produced by the STENN model were compared with the estimation produced by LSTM, random forest (RF), Cubist model, and results provided by other researchers to verify the performance of the STENN model. The tuning parameters of RF and Cubist are presented in Text S2.

To test the spatial generalization ability of our model and the accuracy of our products, we also evaluated the accuracy of our products for each year using the ground-based observation data and compared our products with the widely used 1-km PM<sub>2.5</sub> products from Wei et al. (2021) and van Donkelaar et al. (2016). They produced the annual PM<sub>2.5</sub> products, which can be downloaded from the ChinaHighPM<sub>2.5</sub> dataset (<https://weijing-rs.github.io/product.html>) and SEDAC (Socioeconomic Data and Applications Center, <https://sedac.ciesin.columbia.edu/>). The spatial resolution of both products is 1 km.

Some commonly used indices such as the determination coefficient ( $R^2$ ), RMSE, mean absolute error (MAE), relative prediction error (PRE, %), and predicted bias were calculated to evaluate the performance of the different models (Text S3 for details).

## 2.7. Trend analysis methods

In the current study, we integrated the Theil-Sen Median trend analysis method (Theil, 1950; Sen, 1968) and the Mann-Kendall (MK) test (Kendall, 1938) to analyze the trend of the annual PM<sub>2.5</sub> changes at the pixel scale. The Theil-Sen median slope determines the median of the slopes between all  $\frac{n(n-1)}{2}$  pairwise combinations over time and is based on nonparametric variables; therefore, it is particularly effective for estimating trends in small series (Fensholt et al., 2012; Jiang et al., 2015). The slope of the Theil-Sen Median represents the increase or decrease in the PM<sub>2.5</sub> concentrations over 7 years from 2014 to 2020 for each pixel, which can be obtained by:

$$S_{PM2.5} = \text{median}\left(\frac{PM_{2.5j} - PM_{2.5i}}{j - i}\right) \quad (11)$$

where  $PM_{2.5j}$  and  $PM_{2.5i}$  refer to the PM<sub>2.5</sub> values for years  $j$  and  $i$ ,  $S_{PM2.5}$  represents the Theil-Sen median slope, with a positive  $S_{PM2.5}$  indicating a rising trend and a negative  $S_{PM2.5}$  indicating a decreasing trend.

The MK test is a non-parametric statistical test used to measure the significance of the trend. It offers a great advantage in that outliers do not interfere with the results (de Jong et al., 2011). The Z-test statistics can be calculated as:

$$Z = \begin{cases} \frac{S - 1}{\sqrt{\text{var}(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S + 1}{\sqrt{\text{var}(S)}}, & S < 0 \end{cases} \quad (12)$$

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(PM_{2.5j} - PM_{2.5i}) \quad (13)$$

$$\text{var}(S) = \frac{n(n-1)(2n+5)}{18} \quad (14)$$

where  $n$  is the duration of time, and  $\text{sgn}$  represents a sign function. The index  $Z$  changes in the range of  $(-\infty, +\infty)$ .  $|Z| > u_{1-\alpha}$  indicates that the time series exhibits significant variations at the level of  $\alpha$ . In this study,  $\alpha = 0.05$ .

## 3. Results and discussion

### 3.1. Variable importance in modeling

The results of the variable-sensitivity tests are presented in Fig. 4. The importance mentioned here refers to the importance of each variable in the STENN model design and the reduction of the model accuracy after each variable is removed. Nevertheless, this importance cannot fully represent the physical correlation, although this technique is widely used to quantify the importance (Boulesteix et al., 2012). According to the variable selection strategy of Wei et al. (2021), variables with importance scores lower than 2% should be excluded. The importance scores of all the variables in this study were greater than 2%; consequently, all variables were used in the final model fitting.

The most important variable in the model fitting is AOD, which is known to have a high correlation with the PM<sub>2.5</sub> concentrations (Wang and Christopher, 2003). Meteorological variables (i.e., TEMP, RHU, PRE, BLH and P) also have a crucial influence on PM<sub>2.5</sub>. The thick temperature inversion layer, lower precipitation, and lower BLH increased the air stagnation, stabilizing the polluted air over the area without diffusion (Zheng et al., 2015; Wang et al., 2018). High relative humidity facilitated the formation of secondary aerosols, which are also a significant part of PM<sub>2.5</sub> (Zhang et al., 2015). These results indicate that the highly sensitive variables within STENN have physical interpretability.

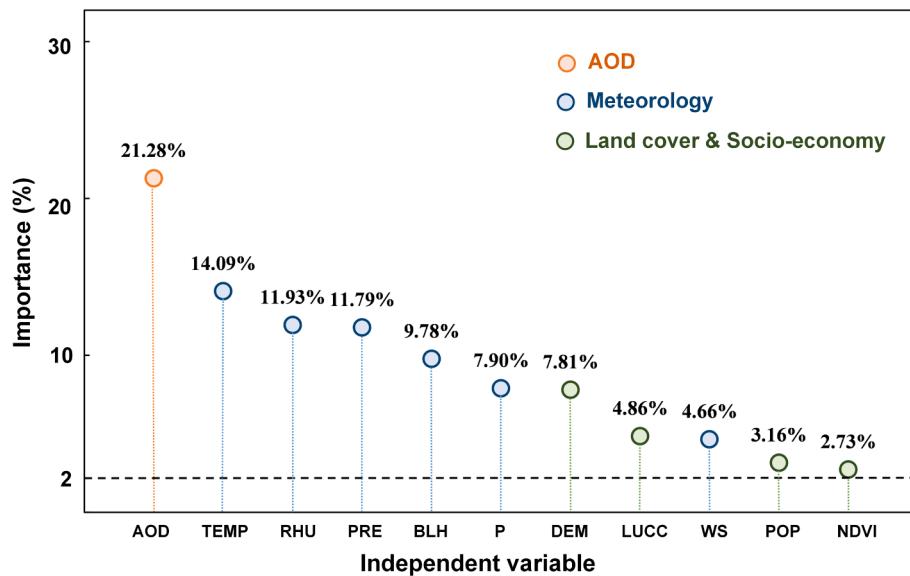


Fig. 4. The importance of independent variables selected for PM<sub>2.5</sub> estimations.

### 3.2. Model and product assessment

#### 3.2.1. Spatiotemporal transferability

Temporal and spatial CVs were also conducted to estimate the spatiotemporal transferable power of the STENN model. Fig. 5 presents

the result of temporal CVs of the annual PM<sub>2.5</sub>, estimated from ground-based observations from 2015 to 2020 in mainland China. The performance of the STENN model for each year was not consistent and was the best for 2020, with the estimation uncertainty decreased largely (i.e., RMSE = 4.25 µg m<sup>-3</sup>, MAE = 2.88 µg m<sup>-3</sup>, and RPE = 12.16%). This was

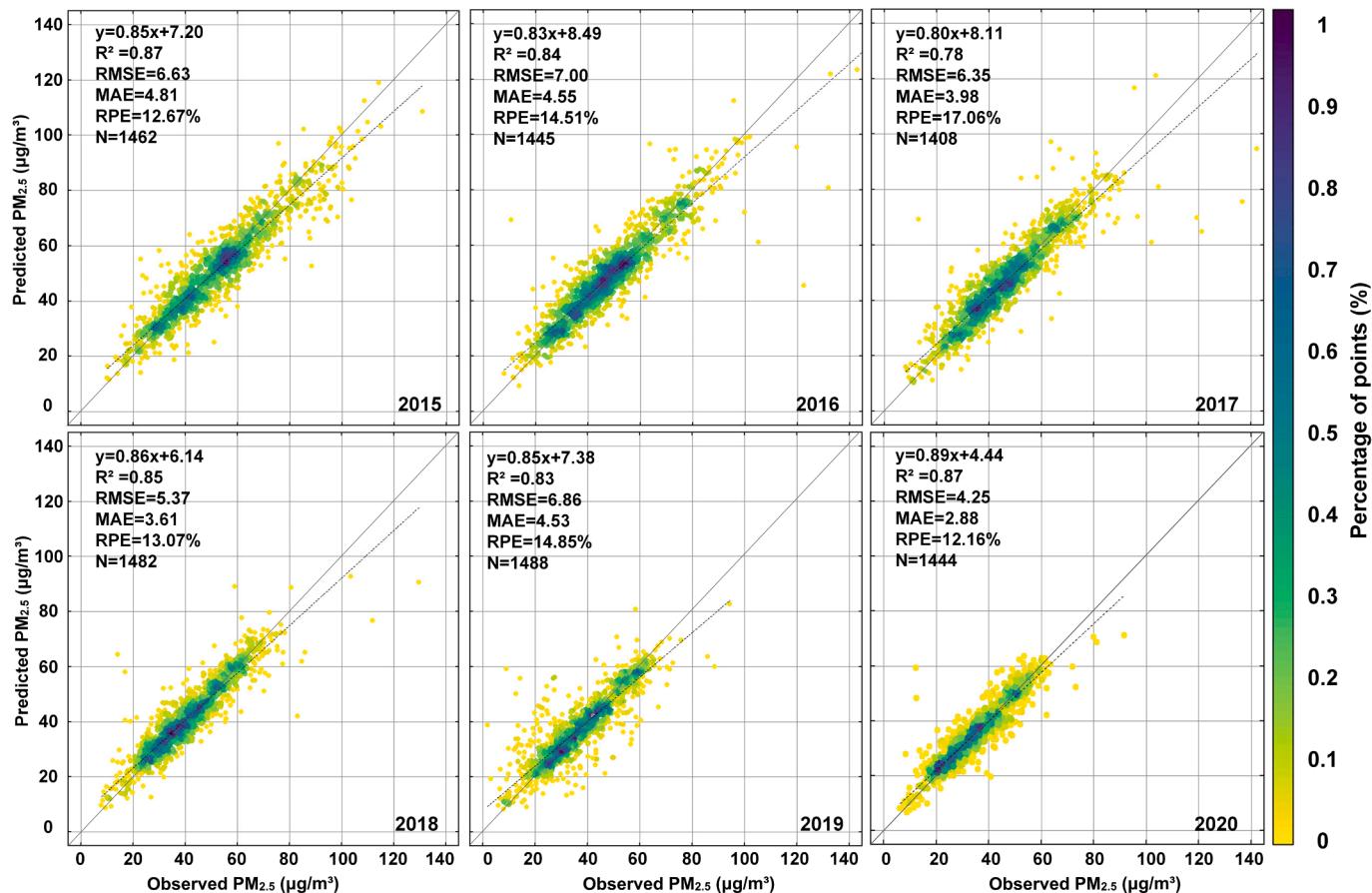
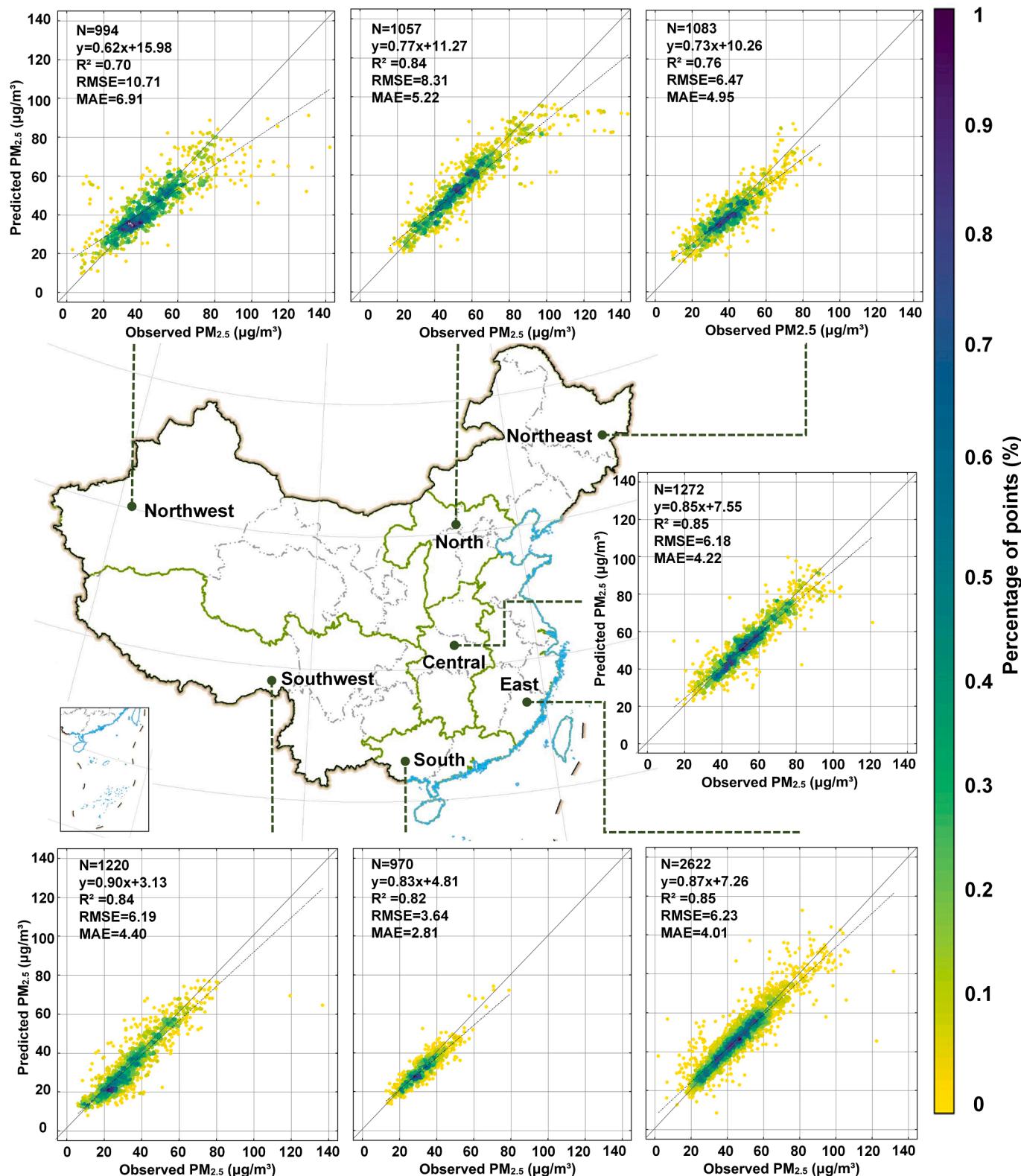


Fig. 5. Density scatter plots of the results of temporal CV using the STENN model in 2015 to 2020 across China. The color of points represents the percentage of the total number of points in this value range. A higher percentage means more data points are within this value range. Statistical metrics (i.e., the fitting lines, 1:1 lines, R<sup>2</sup>, RMSE, MAE and RPE) and the linear regression relationships are given. Units of RMSE and MAE are µg m<sup>-3</sup>.

primarily due to the improved air quality in 2020, which reduced the extremely high concentrations of PM<sub>2.5</sub>. In general, the estimation results were well consistent with the surface observations ( $R^2 = 0.78\text{--}0.87$ , slope = 0.80–0.89) over different years, with the RMSE and MAE values ranging from 4.25 to 7.00  $\mu\text{g m}^{-3}$  and 2.88 to 4.81  $\mu\text{g m}^{-3}$ , respectively.

The absolute biases of the temporal CVs are provided in Fig. S3. STENN showed the best performance for 2020, with almost 80% of the absolute bias of less than 4  $\mu\text{g m}^{-3}$ . In contrast, the worst performance was observed for 2017, with higher absolute biases distributed in central and northeastern China.



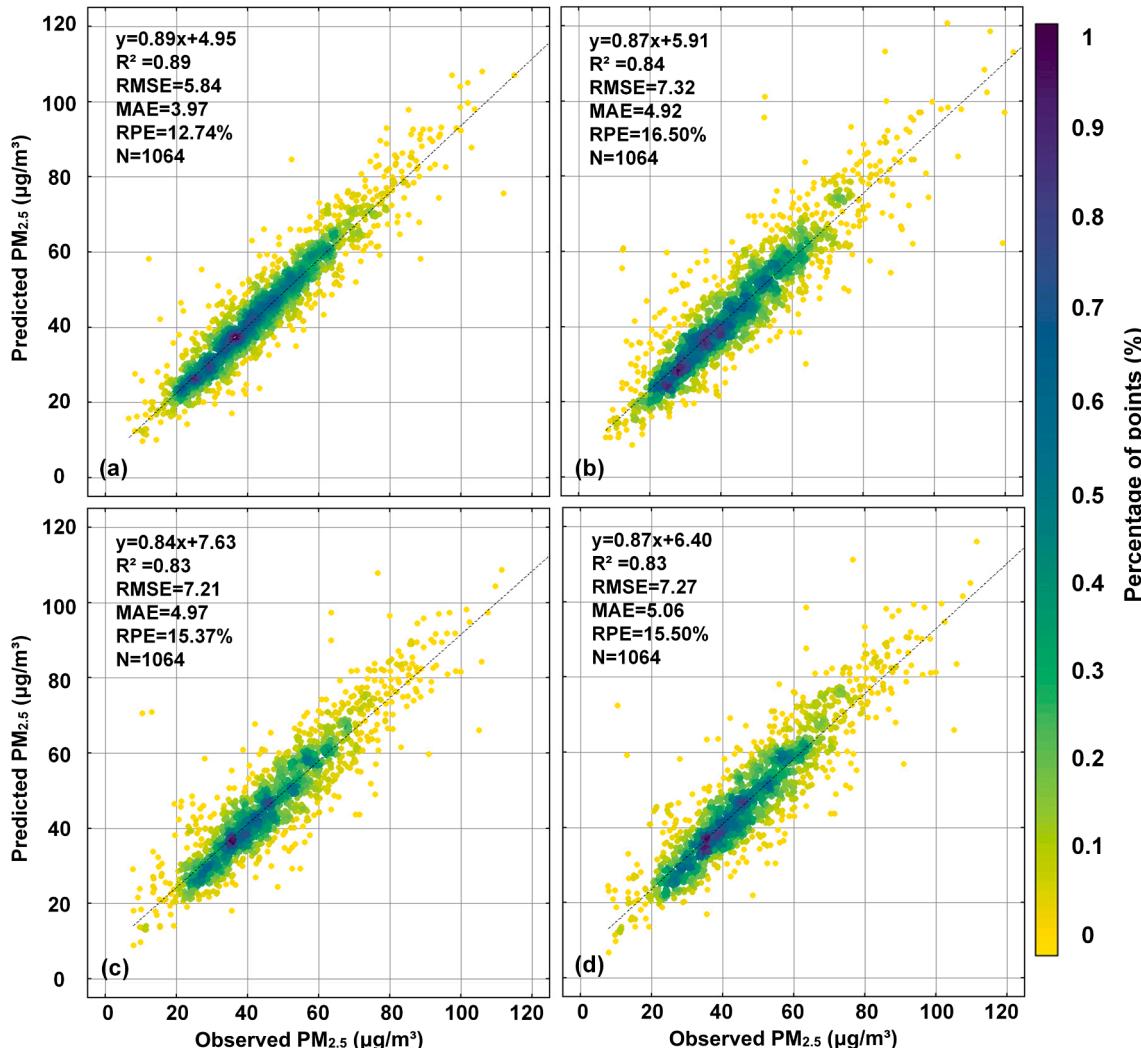
**Fig. 6.** Density scatter plots of the results of spatial CV using the STENN model in seven geographic regions (i.e., northwest, northeast, north, central, east, south, and southwest) across China. The linear regression relation fitting lines, 1:1 lines, sample size (N),  $R^2$ , RMSE, and MAE are given. Units of RMSE and MAE are  $\mu\text{g m}^{-3}$ .

The spatial CV results of the STENN model for annual PM<sub>2.5</sub> predictions over seven geographic regions in China are presented in Fig. 6. The model exhibited different performances for each region. The best prediction performance was observed in the eastern and central China, which had numerous evenly distributed ground sites, with the high R<sup>2</sup> values of 0.85 and a low estimation bias (i.e., RMSE = 6.18 µg m<sup>-3</sup>, MAE = 4.22 µg m<sup>-3</sup> in central China). By contrast, the model performed worst in northwestern China with fewer sites and sparse distribution, with the R<sup>2</sup>, RMSE, and MAE of 0.70, 10.71 µg m<sup>-3</sup>, and 6.91 µg m<sup>-3</sup>, respectively. The annual spatial CV results are provided in Fig S4. The validation results in northern China fluctuated the most over the years due to the significant variations in PM<sub>2.5</sub> concentrations during the 7 years.

Several factors may have contributed to the model performance. In regions with sparsely distributed observations, the effect of spatial signals in response to the spatial correlation of PM<sub>2.5</sub> would decrease because of the greater distances. In contrast, in the regions with numerous and densely distributed observations, the spatial signal could better reflect the variations in PM<sub>2.5</sub> caused by spatial correlation and heterogeneity. Second, the meteorological conditions and surface coverage in different regions were diverse, especially in Xinjiang, in northwestern China, which has a typical continental warm temperate arid climate (Peng et al., 2019) and is varies considerably from the other

regions of China. This also led to poor spatial CV results in northwestern China. In addition, the coverage of the AOD data also affected the results. The MAIAC AOD has a relatively high quality control and removes the AOD pixels in the ice/snow cover areas (Tao et al., 2019). In northeast and southwest China, owing to the snow cover, the AOD data in winter had numerous missing values (Fig. S5 (a)). AOD values were also missing in southern China because of the cloud contamination. When the arithmetic average method was used for the annual AOD synthesis, errors were inevitably introduced, which also caused a bias in spatial CV for the PM<sub>2.5</sub> estimation.

Similar studies have also examined the spatiotemporal transferability of their models. For example, He et al. (2021) used a temporal CV to test their model's predictive power. For estimations at the annual scale, the temporal CV R<sup>2</sup> reached 0.75. The same validation was also conducted by Wei et al. (2021). They compared the temporal transferability of four traditional models (i.e., MLR, LME, GWR, and two-stage model) with STRF (Wei et al., 2019a) and the STET model developed in their study. Their model had the best performance, with an R<sup>2</sup> value of 0.82 at the annual scale. Though these models also demonstrated temporal transferability, the STENN model performed the best with an average R<sup>2</sup> value of 0.84 in the temporal CV. These results indicate that our model accurately predicted the PM<sub>2.5</sub> in the year without the observation data.



**Fig. 7.** Density scatter plots of the results of sample-based CV using the STENN (a), LSTM (b), RF (c), and Cubist (d). The dashed and solid lines denote best-fit lines from linear regression and 1:1 lines, respectively. The color of points represents the percentage of the total number of points in this value range. A higher percentage means more data points are within this value range. The R<sup>2</sup>, RMSE, MAE and RPE are also given. Units of RMSE and MAE are µg m<sup>-3</sup>.

Regarding the spatial transferability, the validation methods vary in different studies (Li et al., 2020b). Li et al. (2017a) used leave-one-province-out cross-validation (LOPOCV), and the  $R^2$  value was 0.54. He and Huang (2018) adopted the space CV (i.e., randomly omitted 10% of the monitoring stations) to test the spatial predictive power of their model, with an  $R^2$  value of 0.75. Wei et al. (2019a) adopted the same method as He and Huang (2018), and the  $R^2$  value was 0.63. The  $R^2$  of the spatial CV of our STENN model reached 0.81, indicating a strong spatial transferability. The difference in accuracy may also be partly due to the different standards of the spatial CV. Therefore, it is necessary to establish reasonable and general validation standards.

### 3.2.2. Overall accuracy

The overall accuracy was evaluated using a sample-based CV. The density scatter plots of the results are presented in Fig. 7. By combining the spatiotemporal variations, the STENN model performed the best, with the lowest RMSE, MAE, and RPE values of  $5.84 \mu\text{g m}^{-3}$ ,  $3.97 \mu\text{g m}^{-3}$ , and 12.74%, respectively. The STENN model yielded a higher average  $R^2$  of 0.89, which was superior to that of LSTM (0.84), RF (0.83), and Cubist (0.83). This conclusion can also be drawn from the distribution of the scattered points. The fitted regression line of the STENN model demonstrated a close tendency to the 1:1 line. The performances of the other three models were generally similar. Although they all have considerable accuracy, without spatial signals introduced, the shortcomings of a poor predictive ability for abnormally high and low values are also evident (Fig. S2).

Other similar studies that have predicted  $\text{PM}_{2.5}$  concentrations in China are listed in Table 1. The  $R^2$  values of these studies varied between 0.64 and 0.89, and the RMSE ranged between  $5.4 \mu\text{g m}^{-3}$  and  $32.98 \mu\text{g m}^{-3}$  in the validation dataset. The GWR model exhibited the worst performance. The GTWR model and timely structure adaptive modeling (TSAM) model, which are two improved forms of the GWR model, improved significantly compared with the GWR model. XGBoost ( $R^2 = 0.86$ ) and RF ( $R^2 = 0.81$ ) were proven to have a better ability to estimate  $\text{PM}_{2.5}$ . In comparison, the improved versions based on geographical

factors, including GTW-GRNN (Geographically and temporally weighted generalized regression neural network), Geo-DBN (Deep belief network), and STRF, achieved higher accuracy with  $R^2$  values of 0.80, 0.88, and 0.85, respectively, indicating the value of improving the model based on the characteristics of the research objects.

Overall, the STENN model demonstrated the best performance for estimating  $\text{PM}_{2.5}$ , with the highest value of  $R^2$  (0.89) and a low value of RMSE ( $5.84 \mu\text{g m}^{-3}$ ). These results indicated that the model had the capacity to better estimate the  $\text{PM}_{2.5}$  concentrations in China in different years. Compared with the traditional geostatistical models (i.e., GWR, stage-1, and stage-2), the STENN model built a data-driven neural network structure and had the ability to learn the data features. Furthermore, because the recurrent neural network could powerfully process the time series, the ability of the model to extract the time dependence of  $\text{PM}_{2.5}$  was enhanced, and the model outperformed the machine learning models (e.g., RF and Cubist) (Yuan et al., 2020). Meanwhile, the spatial signal introduced by the STENN model accounts for the influence of the spatial heterogeneity and spatial correlation, which makes the prediction results more stable; therefore, the prediction accuracy surpassed the general neural network.

### 3.2.3. Product assessment

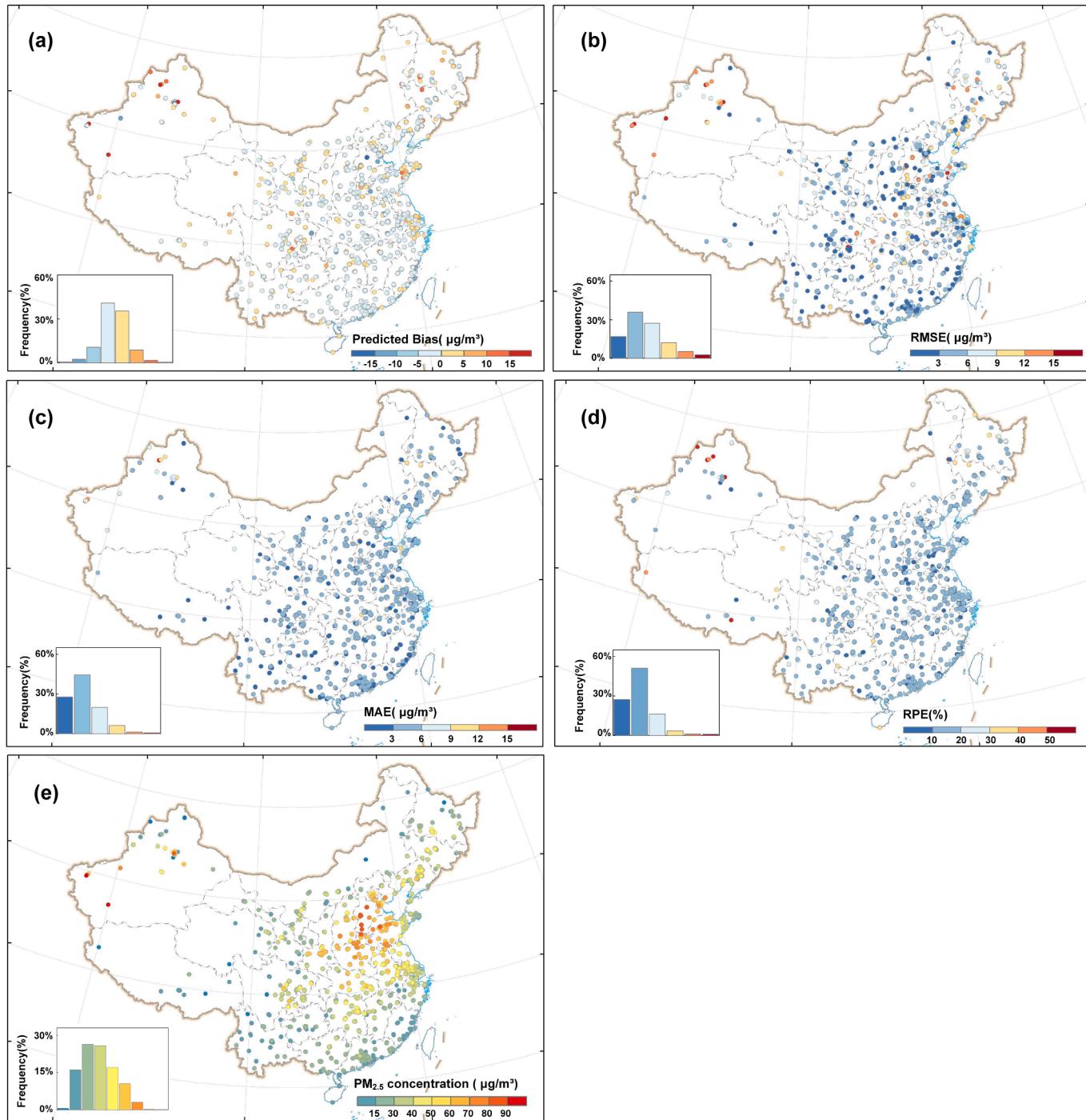
Fig. 8 presents the assessment results of our 1-km resolution annual products against the average annual observations from the ground stations. A series of statistical metrics were calculated by matching the spatial-temporal  $\text{PM}_{2.5}$  values of our products with the observation values of the ground monitoring stations. The products produced high accuracy when compared with the ground-observed  $\text{PM}_{2.5}$ . Overestimations and underestimations accounted for 45.88% and 54.12%, respectively. Most of the predicted biases (76.1%) were distributed within  $\pm 5 \mu\text{g m}^{-3}$ . More than 90% (98%) of the stations had low estimation errors, with the RMSE (MAE) values of less than  $12 \mu\text{g m}^{-3}$ , indicating the spatial generalization ability of the STENN model. The MAE and RPE values were low in most areas of mainland China, with a few exceptions in Xinjiang and Tibet provinces. The spatial distribution of the predicted bias and RMSE also demonstrated variability. Larger estimation errors (i.e.,  $\text{RMSE} > 15 \mu\text{g m}^{-3}$ ) were observed with our products for the northwest, northeast, and northern parts of China. The estimation errors in the northwest and northern parts of China may be attributed to the extremely high  $\text{PM}_{2.5}$ . The sparse distribution of  $\text{PM}_{2.5}$  observation stations in northwestern China also increased the error in the model estimation. In addition, the quality and coverage of AOD also have a significant impact on estimation of  $\text{PM}_{2.5}$ . The two high RPE values in Tibet in Fig. 8(d), which are also reflected in two outliers of over  $100 \mu\text{g m}^{-3}$  in the spatial CV of the southwest (Fig. 6), are primarily caused by the missing AOD data in this area (Fig. S5 (b)). However, due to the low  $\text{PM}_{2.5}$  concentrations in Tibet, the abnormal results are not reflected in the predicted bias (i.e., Fig. 8 (a)) and RMSE (i.e., Fig. 8(b)). The large estimation errors for northeast China may be contributed from the uncertainties in the AOD products due to the bright surface covered by snow in winter.

Fig. 9 presents a detailed comparison among the different  $\text{PM}_{2.5}$  products (Wei et al., 2021; van Donkelaar et al., 2016; van Donkelaar et al., 2018), the ground-measured  $\text{PM}_{2.5}$  concentrations, and ours for 2016. Considering the availability of  $\text{PM}_{2.5}$  products and the typicality of  $\text{PM}_{2.5}$  pollution, the central China and the east coast of China were selected as examples. For the east coast of China, three products accurately estimated the  $\text{PM}_{2.5}$ , especially for the estimation of high values. The three products exhibited close agreement over most of the areas and were highly consistent with the ground-measured  $\text{PM}_{2.5}$ . As presented in Fig. 9(a-c), the high  $\text{PM}_{2.5}$  concentrations were predominantly distributed in the upper left corner (southwestern part of Hebei Province), which is consistent with the spatial pattern of  $\text{PM}_{2.5}$  observed at the monitoring stations (Fig. 9(d)). As shown in Fig. 9(a), the Yangtze River section in the Anhui and Jiangsu provinces had distinctly high values with long and narrow strip distributions, which did not match the

**Table 1**  
Comparisons of model performance in  $\text{PM}_{2.5}$  estimation for China at national scale.

Reference	Temporal period	Spatial resolution (km)	Model	Model validation	
				$R^2$	RMSE
[1]	2004–2013	10	Stage-1:LME	0.78	27.99
			Stage-2:GAM	0.79	27.42
[2]	2012–2013	10	GWR	0.64	32.98
[3]	2013	10	Gaussian	0.81	21.87
[4]	2014–2015	10	TSAM	0.80	22.75
[5]	2014–2015	3	XGBoost	0.86	14.98
[6]	2015	3	Geo-GRNN	0.82	16.93
			Geo-BPNN	0.84	15.23
			Geo-DBN	0.88	13.03
			GW-GRNN	0.78	18.19
[7]	2015	10	GTW-GRNN	0.80	17.38
[8]	2015	3	GTWR	0.80	18.00
[9]	2015–2016	3	Bayesian	0.88	11.39
[10]	2016	1	Maximum Entropy-GWR		
			RF	0.81	17.91
			STRF	0.85	15.57
[11]	2000–2016	10	ML-GAM	0.77	10.1
[12]	2000–2018	1	ASTR	0.77	8.55
[13]	2013–2019	1	Three-stage RF	0.88	15.73
[14]	2019	1	LSTM	0.84	5.40
This study	2014–2020	1	STENN	0.89	5.84

[1] Ma et al., 2016; [2] Ma et al., 2014; [3] Yu et al., 2017; [4] Fang et al., 2016; [5] Chen et al., 2019b; [6] Li et al., 2017a; [7] Li et al., 2020a; [8] He and Huang, 2018; [9] Xiao et al., 2018; [10] Wei et al., 2019a; [11] Xue et al., 2019; [12] He et al., 2021; [13] Huang et al., 2021; [14] Wang et al., 2021.

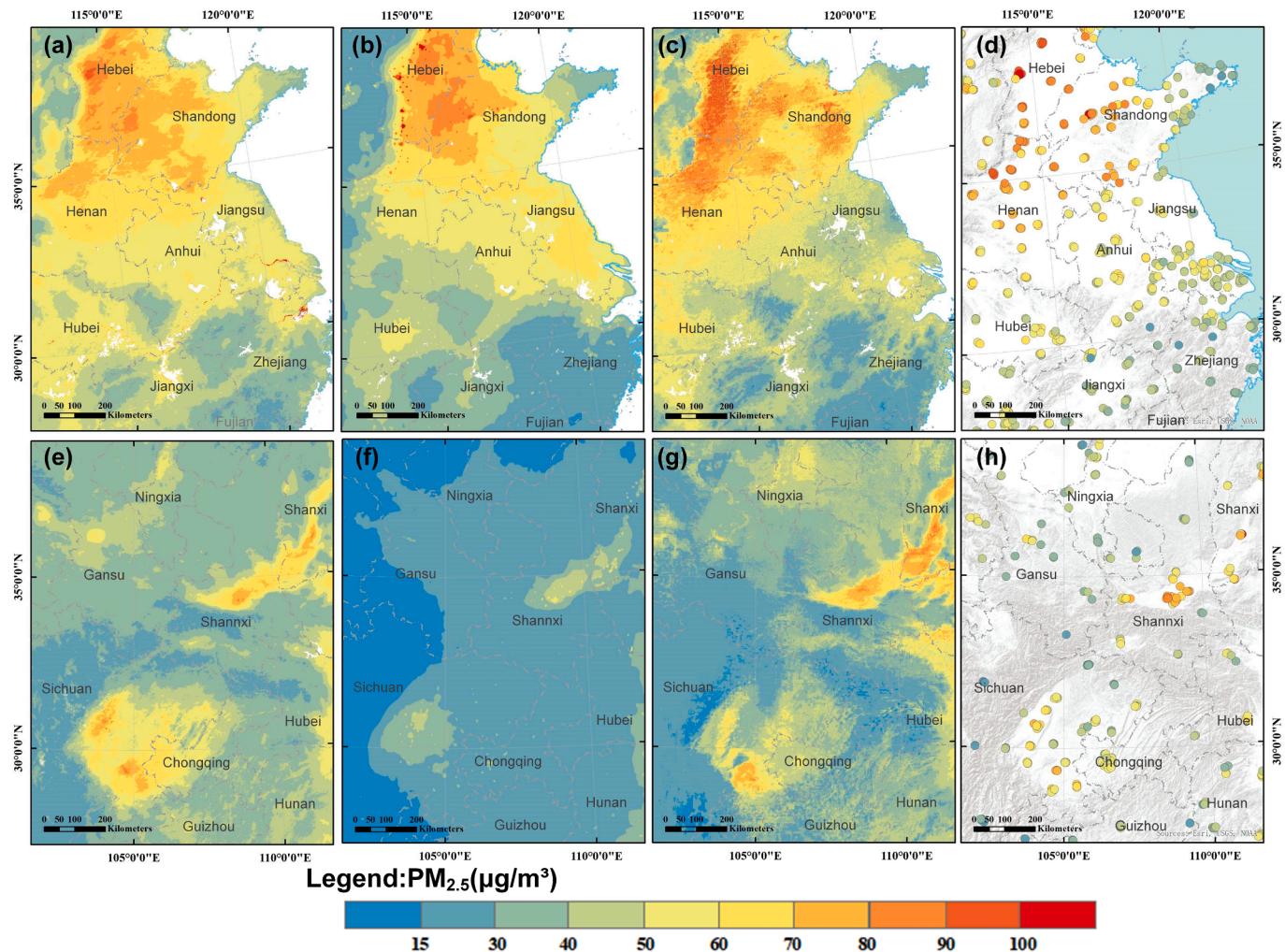


**Fig. 8.** The spatial distribution of the mean (a) predicted bias, (b) RMSE, (c) MAE, and (d) RPE of the 2014–2020 multi-year products compared to the (e) mean PM<sub>2.5</sub> concentrations observed by ground-based stations.

observations at the monitoring stations. The bias is consistent with the bias of AOD in 2016 (Fig. S6). In central China, Fig. 9 (e) and (g) are both in agreement with the ground measurements. However, a significant underestimation can be seen in the products produced by van Donkelaar et al. (2016, 2018) (Fig. 9(f)), and the spatial distribution of PM<sub>2.5</sub> concentrations is relatively smooth, which is dissimilar to the estimation results of the other two products. In contrast, our products provided a better fit with the ground measurements of PM<sub>2.5</sub>, and the distribution was continuous with no abnormal changes and was able to present a more detailed information.

The bias of the two other products can be attributed to several

factors. First, the Space-Time Extra-Tree (STET) method used by Wei et al. (2021) did not consider the time series. It only used data from a designated year to predict PM<sub>2.5</sub> in the same year, with environmental variables from the same year as ancillary. Among these environmental variables, AOD was the most important covariate for estimating PM<sub>2.5</sub>. However, the MODIS AOD products yielded a high bias for deserts and water covers due to the surface brightness and sub-pixel water pollution (Abdou et al., 2005) (Fig. S6). This bias may be further amplified through the continuous splitting of the tree when training the model and causing a further estimation bias. Furthermore, van Donkelaar et al. (2016) used multi-satellite-derived AOD products (i.e., MODIS, MISR,



**Fig. 9.** Detail comparison between PM<sub>2.5</sub> products produced by (a, e) (Wei et al., 2021), (b, f) (van Donkelaar et al., 2016; van Donkelaar et al., 2018), (c, g) our study, and (d, h) ground-measured PM<sub>2.5</sub> concentrations in 2016 (The blank areas represent the rivers and water cover).

SeaWiFS, and CALIPSO) but did not introduce any other auxiliary variables that have a higher spatial resolution and more information (e.g., meteorology and land cover data). This limited their products from providing more detailed information on PM<sub>2.5</sub>.

### 3.3. Spatial pattern of PM<sub>2.5</sub> in China

Based on multi-temporal satellite products, applying the STENN model, products regarding the PM<sub>2.5</sub> distributions, with a 1-km resolution were produced over whole mainland China from 2014 to 2020. The estimated spatial distribution of the annual mean PM<sub>2.5</sub> concentration distribution for each year from 2014 to 2020 are presented in Fig. 10 (a-g), and the 7-year mean PM<sub>2.5</sub> concentration distribution is shown in Fig. 10(h).

The average PM<sub>2.5</sub> concentrations between 2014 and 2020 throughout China was  $32.95 \pm 13.88 \mu\text{g m}^{-3}$ . In terms of regions, the multi-year averaged PM<sub>2.5</sub> concentrations in the Beijing-Tianjin-Hebei (BTH) region, the Yangtze River Delta (YRD) region, and the Pearl River Delta (PRD) region were  $49.26 \pm 16.13 \mu\text{g m}^{-3}$ ,  $41.56 \pm 11.19 \mu\text{g m}^{-3}$ , and  $28.35 \pm 4.62 \mu\text{g m}^{-3}$ , respectively. Air quality exhibited a trend of improvement over this period. The highest annual PM<sub>2.5</sub> concentration across mainland China was  $39.46 \pm 16.02 \mu\text{g m}^{-3}$  in 2014 and the lowest was  $26.18 \pm 14.15 \mu\text{g m}^{-3}$  in 2020.

The spatial distribution patterns of the PM<sub>2.5</sub> concentrations in China for multiple years were similar, where the most polluted areas were

indicated in northern China, northwest China, and a small part of central China, where the main provinces and cities include Beijing, Tianjin, Hebei, Henan, Xinjiang, and Sichuan. The primary cause of severe air pollution in these areas is anthropogenic emissions (e.g., fossil fuel combustion) (Lv et al., 2021; Ye et al., 2018). In 2019, the production of crude steel in the high pollution areas (i.e., the BTH region, and Fenwei plains) still exceeded 50% of the national production level, of which the rank of production in Hebei remained the highest in China. Furthermore, secondary aerosol formation (e.g., sulfate, which has been suggested as a significant component of PM<sub>2.5</sub>) produced by the interactions between SO<sub>2</sub> and NO<sub>2</sub> (Wang et al., 2020c) and NH<sub>3</sub> (Gu et al., 2021) also become an substantial source of PM<sub>2.5</sub> pollution. In terms of natural conditions, unfavorable meteorology (i.e., low wind speed) and topography, such as basins that do not favor the dispersion of fine particulate matter, were also key contributors to PM<sub>2.5</sub> pollution (Bao et al., 2019; Sulaymon et al., 2021). The lowest average PM<sub>2.5</sub> values were located in southwestern China (i.e., Yunnan-Guizhou Plateau) and the southeast coastal region of China (i.e., PRD). The fine air quality could be attributed to both the natural conditions (e.g., flat terrain, adequate precipitation, and the monsoon climate) and policy management. (The detailed locations of these regions are indicated in Fig. 2)

### 3.4. Spatiotemporal variations and trends

At the end of 2013, the Chinese government implemented a five-year

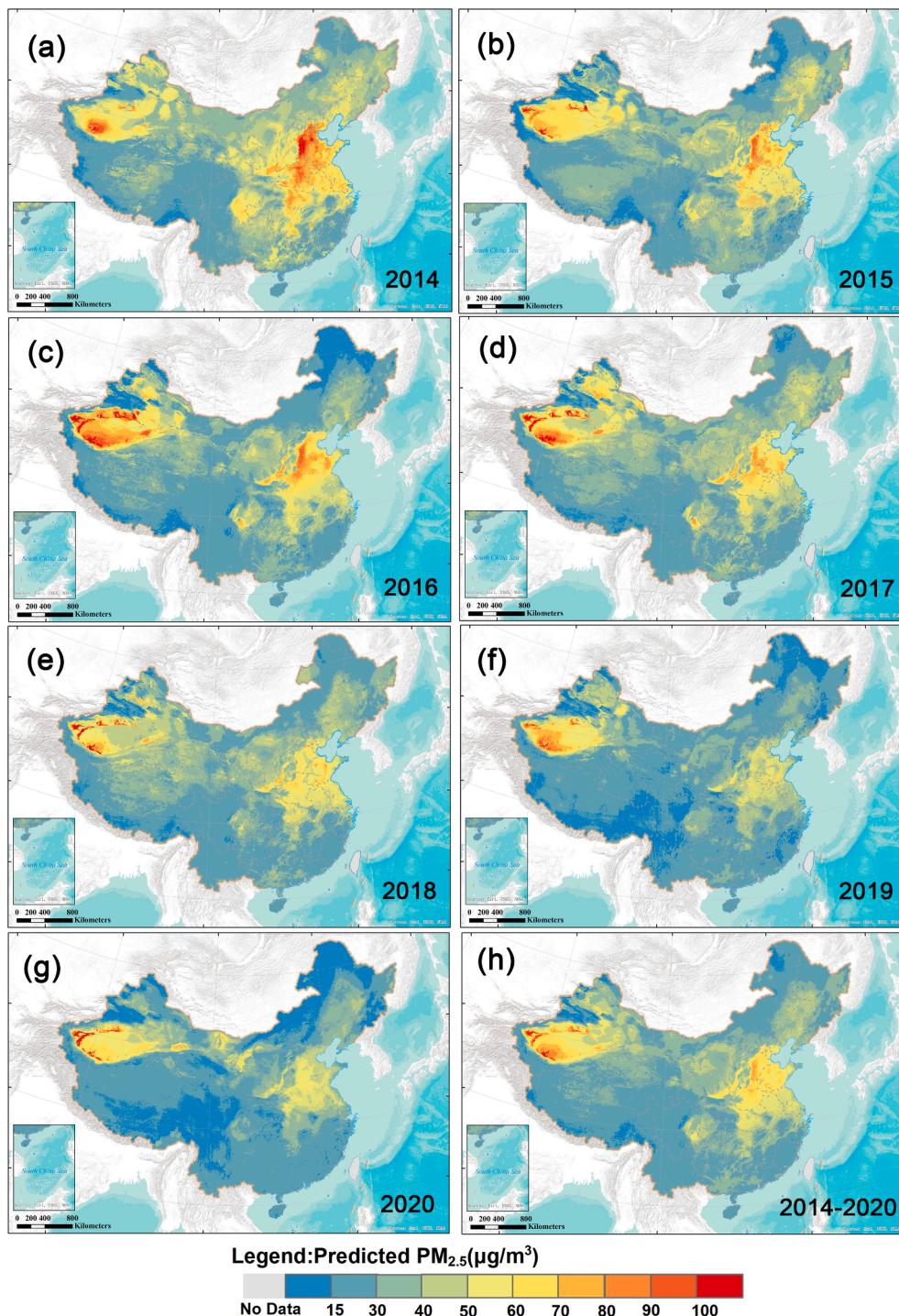


Fig. 10. Spatial distribution of annual mean (a-g), multi-year (2014–2020) mean (h) satellite-derived PM<sub>2.5</sub> concentrations across China.

Action Plan on Air Pollution Prevention (2013–2017). In 2017, the government issued a three-year Action Plan to Win the Blue-Sky Defense (2018–2020). Meanwhile, the 13th Five-Year Plan (2016–2020) and other action plans also put forward clear control requirements for air pollution. In order to explore the effectiveness of the different policies based on the multi-year PM<sub>2.5</sub> distribution products, the analysis was carried out on the spatial variability and temporal trends of the PM<sub>2.5</sub> for different periods in this study based on the multi-year PM<sub>2.5</sub> distribution products. Fig. 11 presents the spatiotemporal variability of PM<sub>2.5</sub> during the periods 2014–2020, 2014–2017, and 2017–2020, respectively (Table S8 for the detailed statistics information). Compared with 2014,

the PM<sub>2.5</sub> concentrations had decreased in 88.79% of the total area of China by 2020, with the mean decreased concentrations of 13.28 µg m<sup>-3</sup> and the decreased percentage reaching 33.65%. The regions that decreased most significantly were northern China, followed by central China, and the east coast. Only northwest China exhibited a slightly increasing trend. The spatial distribution of the PM<sub>2.5</sub> variations differed among the periods. During 2014–2017, the most significant decrease was concentrated in northern and central China, while other areas also experienced a slight decrease, with the decrease ranging from 0 to 10 µg m<sup>-3</sup>, which was predominantly due to the efforts of the government by implementing the Action Plan on Air Pollution Prevention. In contrast,

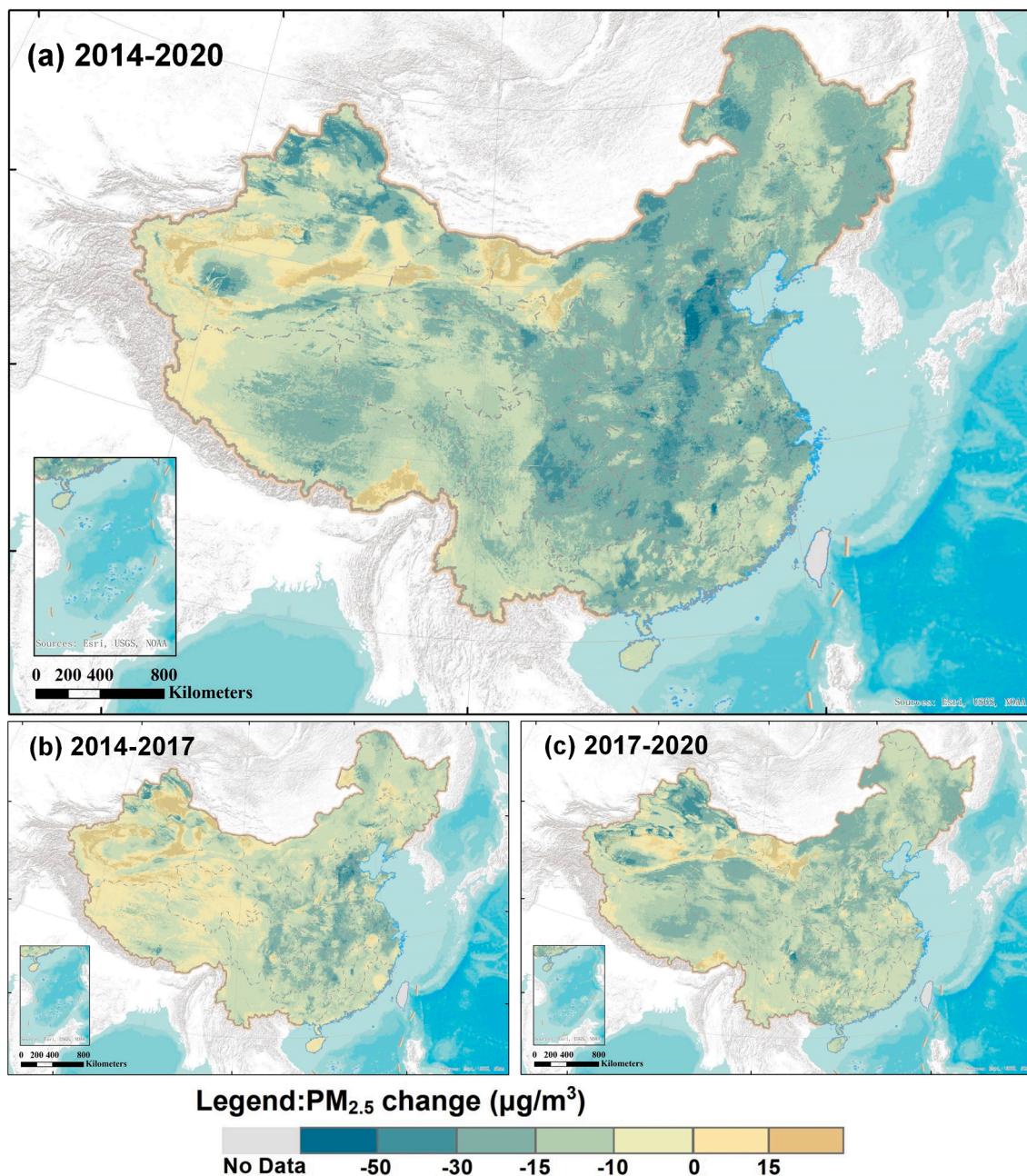


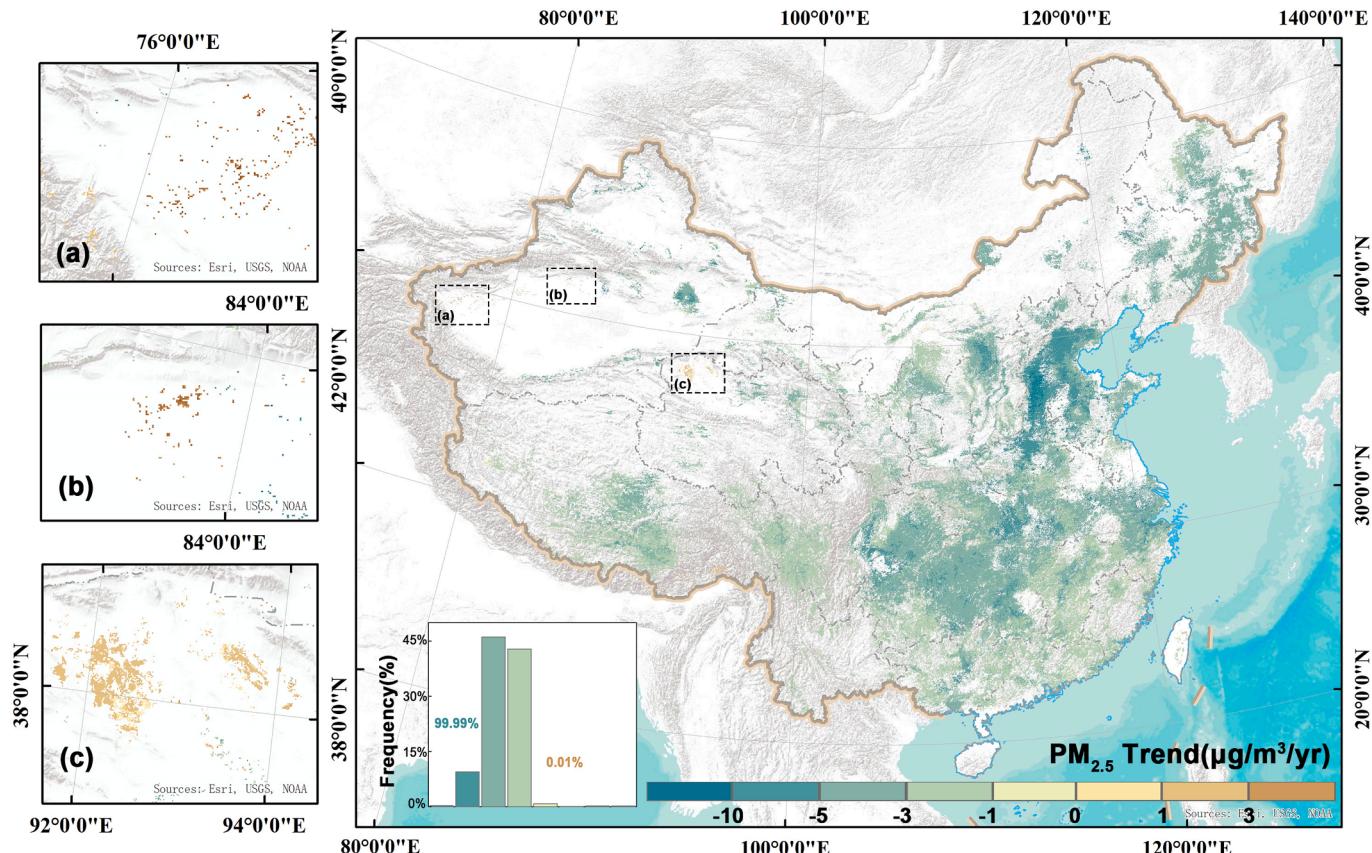
Fig. 11. Spatial variation of PM<sub>2.5</sub> concentrations during the periods (a) 2014–2020, (b) 2014–2017, (c) 2017–2020 across China.

the pollution situation was still severe in northwest China, with the PM<sub>2.5</sub> concentrations increasing from 2014 to 2017 (Fig. 11b).

During 2017–2020, the areas with significant decreases were extended to almost all of mainland China, including southwest and northeast China, with decreased concentrations of 0 to 30  $\mu\text{g m}^{-3}$ , and the regions with the largest decline were the BTH region, PRD region, Fenwei Plain, and Sichuan Province (Fig. 11c). The 13th Five-Year Plan (2016–2020) and the Blue Sky Defense (2018–2020) played critical roles in the improvement of air quality during this period. The former plan proposed zoning policies to control air pollution. It considered regional differences and emphasized the reduction in the concentrations of particulate matter in the BTH. The latter plan listed the BTH region and Fenwei Plain as prior regions for the prevention and control of air pollution. Under the proactive PM<sub>2.5</sub> control action, China has taken great efforts to reduce emissions from industries and vehicles, phasing out outdated industrial capacity and promoting clean fuels (Zhang et al.,

2019). These policies led to a prominent decrease in PM<sub>2.5</sub> after 2013. In addition, the meteorological conditions also contributed to the regional PM<sub>2.5</sub> variations. The high precipitation accelerated the PM<sub>2.5</sub> decline in southeast China, and the strong northwest winds in the BTH, to some extent, reduced PM<sub>2.5</sub> in 2017 (Ding et al., 2019; Chen et al., 2019a). In contrast, the presence of air stagnation (i.e., low wind speed, little or no precipitation, and shallow boundary layer heights) (Wang et al., 2018) events in the Sichuan province could not reduce PM<sub>2.5</sub>. Therefore, the impact of the different meteorological conditions and topographical effects should not be neglected when formulating clean air clean action plans.

We also investigated the PM<sub>2.5</sub> trend to clarify the improvement in air pollution in recent years. Fig. 12 presents the spatial distribution and frequency of the 7-year PM<sub>2.5</sub> trend of the annual mean PM<sub>2.5</sub>. The PM<sub>2.5</sub> trend exhibited great variability, ranging from  $-18.65$  to  $9.13 \mu\text{g m}^{-3} \text{ yr}^{-1}$ . The PM<sub>2.5</sub> concentrations tended to decrease in 99.99% of the



**Fig. 12.** Spatial distribution and frequency of PM<sub>2.5</sub> trends for 2014–2020 across China (The areas without color indicated where are not significant at 95%). (a, b, c) showed the detailed information.

statistically significant study area ( $p < 0.05$ , Fig. 12), with a mean decrease rate of  $3.35 \mu\text{g m}^{-3} \text{ yr}^{-1}$  at the pixel level from 2014 to 2020. The PM<sub>2.5</sub> decreased most rapidly in northern China, and the decreasing trend was also significant in central China, the YRD, and the PRD region. In contrast, the PM<sub>2.5</sub> trend in western China demonstrated a slight increase with less significance (Fig. 12a, b, c).

The PM<sub>2.5</sub> slope trend and time series of the BTH, YRD, PRD regions, and mainland China are illustrated in Fig. 13. The 1st and 99th percentiles of the PM<sub>2.5</sub> trends are identified by the dotted line of the box (Fig. 13a), which indicates that the three main regions and mainland China had a decreasing trend of PM<sub>2.5</sub>. The decreasing trend was most clearly seen in the BTH region, and the least significant was in the PRD region because of the relatively low PM<sub>2.5</sub> concentrations in the PRD since 2014. In addition, the trend in the BTH region exhibited the greatest variability, with the amplitude of change exceeding the national level. For the time series, the PM<sub>2.5</sub> concentrations in the BTH region decreased most significantly, with decreased concentrations of  $30.24 \pm 14.47 \mu\text{g m}^{-3}$  (~45.03%) and a slope of  $-0.094$ . According to this trend, it is expected that PM<sub>2.5</sub> concentrations in the BTH region will stabilize under the Chinese National Ambient Air Quality Standard (CNAQ) level 2 within five years. Over the years, the PM<sub>2.5</sub> trend in the PRD region was consistent with that throughout China, with a slope of  $-0.068$ . The slope trend in the three regions and throughout China showed an exponential change, from a rapid decline to a gradual steady change. Although the PM<sub>2.5</sub> concentrations in China continued to decrease from 2013 to 2020, the change rate also decreased. Even with the continued improvement of air quality in China, greater efforts are necessary for further improvement. The Beautiful China Outlook in the 14th Five-Year Plan (2021–2025) mentioned that by 2025, the annual average concentrations of PM<sub>2.5</sub> would drop below  $30 \mu\text{g m}^{-3}$ . According to the trend analysis over the study period, it is not difficult to

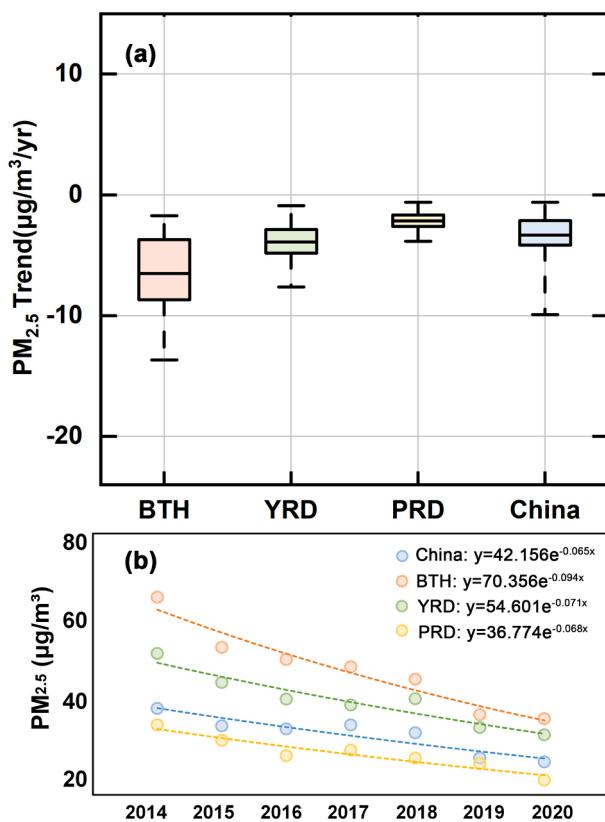
achieve this goal nationwide, but for the BTH region, due to the variability in PM<sub>2.5</sub>, it is still necessary to increase control. As a result, it is particularly vital to carry out sub-regional management and control according to the pollution levels of different regions.

#### 4. Conclusions

In this study, a spatially generalizable deep learning model, the Spatiotemporal Enhanced Neural Network (STENN), was developed to determine the relationship between PM<sub>2.5</sub>-AOD for mainland China by conflating temporal and spatial remote sensing, meteorological indices, and human activities. By integrating the spatial signals into the recurrent neural network, which is capable of processing time series, the model provided a geographic-data-driven approach to incorporate the impact of the spatial heterogeneity and time dependence of PM<sub>2.5</sub> and significantly outperformed most previous models for predicting PM<sub>2.5</sub> concentrations (e.g.,  $R^2 = 0.89$ ). Furthermore, the model also has robust spatiotemporal transferable power for extrapolating to the years and regions without ground-based monitoring.

Using our developed framework, the high-quality and high-resolution (1 km) annual PM<sub>2.5</sub> products covered the entire mainland of China from 2014 to 2020 were produced. The products exhibit high consistency with the surface PM<sub>2.5</sub>, and more than 90% (98%) of the estimations in the stations demonstrated few estimation errors, with the RMSE (MAE) values less than  $12 \mu\text{g m}^{-3}$ . In addition, compared with the widely used 1 km PM<sub>2.5</sub> products, our products provided better stability in different regions, especially in terms of high-value estimations and spatial continuity, demonstrating the strong spatial generalization ability of the model.

Based on the multi-year high-resolution products, an analysis of the spatiotemporal PM<sub>2.5</sub> variations and trends was conducted. The 7-year



**Fig. 13.** (a) The boxplot of the PM<sub>2.5</sub> trend for each region; (b) Time series of annual mean PM<sub>2.5</sub> concentrations and regression lines (dotted line) in each region from 2014 to 2020.

mean PM<sub>2.5</sub> concentrations for mainland China were  $32.95 \pm 13.88 \mu\text{g m}^{-3}$ . The most polluted areas were observed in northern China, northwest China, and a small portion of central China. After the implementation of different control policies of atmospheric pollution, compared with 2014, a decline in the PM<sub>2.5</sub> concentrations was observed in 88.79% of China by 2020, with a mean decrease rate of  $3.35 \mu\text{g m}^{-3} \text{ yr}^{-1}$ , indicating the effective governance capacity of the Chinese government. The slope trends in the BTH, YRD, and PRD regions and throughout China demonstrated an exponential change trend, from a rapid decline to a gradual slowdown and a stable phase; therefore, in order to realize the Beautiful China Initiative by 2025, a more regionally targeted policy for air pollution management is required.

Overall, the STENN model proposed in this study fully utilizes the spatiotemporal variations in PM<sub>2.5</sub> and remote sensing data, and is practical for generating a long-term, more detailed and reliable PM<sub>2.5</sub> product with high resolution, which can greatly enhance research on long-term PM<sub>2.5</sub> variations and provide valuable implications for improving the air quality 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 Key Research and Development Program of Zhejiang Province, China (2022C03078), the Ten-thousand Talents Plan of Zhejiang Province (2019R52004), and the National Key Research and Development Program of China (2018YFC1800105). We would like to thank Dr. Jing Wei and SEDAC (Socioeconomic Data and

Applications Center) for the public PM<sub>2.5</sub> grid products.

#### Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.isprsjprs.2022.03.002>.

#### References

- Abdou, W.A., Diner, D.J., Martonchik, J.V., Bruegge, C.J., Kahn, R.A., Gaitley, B.J., Crean, K.A., 2005. Comparison of coincident Multiangle Imaging Spectroradiometer and Moderate Resolution Imaging Spectroradiometer aerosol optical depths over land and ocean scenes containing aerosol robotic network sites. *J. Geophys. Res.* 110 (D10).
- Bao, Z., Chen, L., Li, K., Han, L., Wu, X., Gao, X., Azzi, M., Cen, K., 2019. Meteorological and chemical impacts on PM<sub>2.5</sub> during a haze episode in a heavily polluted basin city of eastern China. *Environ. Pollut.* 250, 520–529.
- Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., van Dingenen, R., Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* 50, 79–88.
- Brokamp, C., Jandarov, R., Hossain, M., Ryan, P., 2018. Predicting Daily Urban Fine Particulate Matter Concentrations Using a Random Forest Model. *Environ. Sci. Technol.* 52, 4173–4179.
- Boulesteix, A.L., Bender, A., Bermejo, J.L., Strobl, C., 2012. Random forest Gini importance favours SNPs with large minor allele frequency: impact, sources and recommendations. *Briefings Bioinformatics* 13, 292–304.
- Chen, F., Chen, Z., 2021. Cost of economic growth: Air pollution and health expenditure. *Sci. Total Environ.* 755, 142543.
- Chen, Z., Chen, D., Kwan, M.P., Chen, B., Gao, B., Zhuang, Y., Li, R., Xu, B., 2019a. The control of anthropogenic emissions contributed to 80% of the decrease in PM<sub>2.5</sub> concentrations in Beijing from 2013 to 2017. *Atmos. Chem. Phys.* 19, 13519–13533.
- Chen, X., de Leeuw, G., Arola, A., Liu, S., Liu, Y., Li, Z., Zhang, K., 2020a. Joint retrieval of the aerosol fine mode fraction and optical depth using MODIS spectral reflectance over northern and eastern China: Artificial neural network method. *Remote Sens. Environ.* 249.
- Chen, Z.Y., Zhang, T.H., Zhang, R., Zhu, Z.M., Yang, J., Chen, P.Y., Ou, C.Q., Guo, Y.M., 2019b. Extreme gradient boosting model to estimate PM<sub>2.5</sub> concentrations with missing-filled satellite data in China. *Atmos. Environ.* 202, 180–189.
- Chen, Z.Y., Chen, D.L., Zhao, C.F., Kwan, M.P., Cai, J., Zhuang, Y., Zhao, B., Wang, X.Y., Chen, B., Yang, J., Li, R.Y., He, B., Gao, B.B., Wang, K.C., Xu, B., 2020b. Influence of meteorological conditions on PM<sub>2.5</sub> concentrations across China: A review of methodology and mechanism. *Environ. Int.* 139, 105558.
- Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandonia, L., Dandonia, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 389, 1907–1918.
- Ceca, L.S.D., García Ferreyra, M.F., Lyapustin, A., Chudnovsky, A., Otero, L., Carreras, H., Barnaba, F., 2018. Satellite-based view of the aerosol spatial and temporal variability in the Córdoba region (Argentina) using over ten years of high-resolution data. *ISPRS J. Photogramm. Remote Sens.* 145, 250–267.
- de Jong, R., de Bruin, S., de Wit, A., Schaeppman, M.E., Dent, D.L., 2011. Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sens. Environ.* 115, 692–702.
- Ding, D., Xing, J., Wang, S.X., Liu, K.Y., Hao, J.M., 2019. Estimated Contributions of Emissions Controls, Meteorological Factors, Population Growth, and Changes in Baseline Mortality to Reductions in Ambient PM<sub>2.5</sub> and PM<sub>2.5</sub>-Related Mortality in China, 2013–2017. *Environ. Health Perspect.* 127, 12.
- Engel-Cox, J.A., Holloman, C.H., Coutant, B.W., Hoff, R.M., 2004. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmos. Environ.* 38, 2495–2509.
- Fan, J., Li, Q., Hou, J., Feng, X., Karimian, H., Lin, S., 2017. A Spatiotemporal Prediction Framework for Air Pollution Based on Deep RNN. *ISPRS J. Photogramm. Remote Sens.* IV-4/W2, 15–22.
- Fang, X., Zou, B., Liu, X., Sternberg, T., Zhai, L., 2016. Satellite-based ground PM<sub>2.5</sub> estimation using timely structure adaptive modeling. *Remote Sens. Environ.* 186, 152–163.
- Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S.D., Tucker, C., Scholes, R.J., Le, Q.B., Bondeau, A., Eastman, R., Epstein, H., Gaughan, A.E., Hellden, U., Mbow, C., Olsson, L., Paruelo, J., Schweitzer, C., Sequist, J., Wessels, K., 2012. Greenness in semi-arid areas across the globe 1981–2007 — an Earth Observing Satellite based analysis of trends and drivers. *Remote Sens. Environ.* 121, 144–158.
- Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley, New York.
- Franklin, M., Kalashnikova, O.V., Garay, M.J., 2017. Size-resolved particulate matter concentrations derived from 4.4 km-resolution size-fractionated Multi-angle Imaging

- SpectroRadiometer (MISR) aerosol optical depth over Southern California. *Remote Sens. Environ.* 196, 312–323.
- Fu, T., Zhao, R., Hu, B., Jia, X., Wang, Z., Zhou, L., Huang, M., Li, Y., Shi, Z., 2021. Novel framework for modelling the cadmium balance and accumulation in farmland soil in Zhejiang Province, East China: Sensitivity analysis, parameter optimisation, and forecast for 2050. *J. Clean. Pro.* 279, 123674.
- Guo, B., Zhang, L., Van Dingenen, R., Vieno, M., Van Grinsven, H.J., Zhang, X., Zhang, S., Chen, Y., Wang, S., Ren, C., Rao, S., Holland, M., Winiwarter, W., Chen, D., Xu, J., Sutton, M.A., 2021. Abating ammonia is more cost-effective than nitrogen oxides for mitigating PM<sub>2.5</sub> air pollution. *Science (New York, N.Y.)* 374, 758–762.
- 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.P., Zhang, X.Y., Che, H.Z., Gong, S.L., An, X., Cao, C.X., Li, X.-W., 2009. Correlation between PM concentrations and aerosol optical depth in eastern China. *Atmos. Environ.* 43 (37), 5876–5886.
- Gupta, P., Christopher, S.A., 2008. Seven year particulate matter air quality assessment from surface and satellite measurements. *Atmos. Chem. Phys.* 8, 3311–3324.
- He, Q., Huang, B., 2018. Satellite-based mapping of daily high-resolution ground PM<sub>2.5</sub> in China via space-time regression modeling. *Remote Sens. Environ.* 206, 72–83.
- He, Q., Gao, K., Zhang, L., Song, Y., Zhang, M., 2021. Satellite-derived 1-km estimates and long-term trends of PM<sub>2.5</sub> concentrations in China from 2000 to 2018. *Environ. Int.* 156, 106726.
- Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9, 1735–1780.
- Huang, C., Hu, J., Xue, T., Xu, H., Wang, M., 2021. High-Resolution Spatiotemporal Modeling for Ambient PM<sub>2.5</sub> Exposure Assessment in China from 2013 to 2019. *Environ. Sci. Technol.* 55, 2152–2162.
- Hu, B., Xue, J., Zhou, Y., Shao, S., Fu, Z., Li, Y., Chen, S., Qi, L., Shi, Z., 2020. Modelling bioaccumulation of heavy metals in soil-crop ecosystems and identifying its controlling factors using machine learning. *Environ. Pollut.* 262, 114308.
- Hu, B., Zhou, Q., He, C., Duan, L., Li, W., Zhang, G., Ji, W., Peng, J., Xie, H., 2021. Spatial variability and potential controls of soil organic matter in the Eastern Dongting Lake Plain in southern China. *J. Soils Sediments* 21 (8), 2791–2804.
- Jiang, W., Yuan, L., Wang, W., Cao, R., Zhang, Y., Shen, W., 2015. Spatio-temporal analysis of vegetation variation in the Yellow River Basin. *Ecol. Indic.* 51, 117–126.
- Kendall, M., 1938. A new measure of rank correlation. *Biometrika* 30 (1–2), 81–89.
- Li, J., Carlson, B.E., Lacis, A.A., 2015a. How well do satellite AOD observations represent the spatial and temporal variability of PM<sub>2.5</sub> concentration for the United States? *Atmos. Environ.* 102, 260–273.
- Li, T., Shen, H., Yuan, Q., Zhang, X., Zhang, L., 2017a. Estimating Ground-Level PM<sub>2.5</sub> by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. *Geophys. Res. Lett.* 44.
- Li, T., Shen, H., Yuan, Q., Zhang, L., 2020a. Geographically and temporally weighted neural networks for satellite-based mapping of ground-level PM<sub>2.5</sub>. *ISPRS J. Photogramm. Remote Sens.* 167, 178–188.
- Li, T., Shen, H., Zeng, C., Yuan, Q., 2020b. A Validation Approach Considering the Uneven Distribution of Ground Stations for Satellite-Based PM<sub>2.5</sub> Estimation. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 13, 1312–1321.
- Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., Chi, T., 2017b. Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environ. Pollut.* 231, 997–1004.
- Li, Y., Chen, Q., Zhao, H., Wang, L., Tao, R., 2015b. Variations in PM<sub>10</sub>, PM<sub>2.5</sub> and PM<sub>1.0</sub> in an Urban Area of the Sichuan Basin and Their Relation to Meteorological Factors. *Atmos* 6, 150–163.
- Liu, J., Weng, F., Li, Z., 2019a. Satellite-based PM<sub>2.5</sub> estimation directly from reflectance at the top of the atmosphere using a machine learning algorithm. *Atmos. Environ.* 208, 113–122.
- Liu, H., Dong, L., Yan, R., Zhang, X., Guo, C., Liang, S., Tu, J., Feng, X., Wang, X., 2021. Evaluation of Near-Surface Wind Speed Climatology and Long-Term Trend over China's Mainland Region Based on ERA5 Reanalysis. *Climatic Environ. Res. (in Chinese)* 26 (3), 299–311.
- Liu, M., Zhou, G., Saari, R.K., Li, S., Liu, X., Li, J., 2019b. Quantifying PM<sub>2.5</sub> mass concentration and particle radius using satellite data and an optical-mass conversion algorithm. *ISPRS J. Photogramm. Remote Sens.* 158, 90–98.
- Liu, W., Guo, G., Chen, F., Chen, Y., 2019c. Meteorological pattern analysis assisted daily PM<sub>2.5</sub> grades prediction using SVM optimized by PSO algorithm. *Atmos. Pollut. Res.* 10, 1482–1491.
- Liu, Y., Paciorek, C.J., Koutrakis, P., 2009. Estimating regional spatial and temporal variability of PM<sub>2.5</sub> concentrations using satellite data, meteorology, and land use information. *Environ. Health Perspect.* 117, 886–892.
- Luong, M.T., Pham, H., Manning, C.D., 2015. Effective approaches to attention-based neural machine translation. arXiv preprint arXiv 1508.
- Lv, L., Chen, Y., Han, Y., Cui, M., Wei, P., Zheng, M., Hu, J., 2021. High-time-resolution PM<sub>2.5</sub> source apportionment based on multi-model with organic tracers in Beijing during haze episodes. *Sci. Total Environ.* 772, 144766.
- Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* 11, 5741–5765.
- Ma, J., Ding, Y., Cheng, J.C.P., Jiang, F., Wan, Z., 2019a. A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM<sub>2.5</sub>. *J. Clean. Prod.* 237.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019b. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 152, 166–177.
- Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level PM<sub>2.5</sub> in China using satellite remote sensing. *Environ. Sci. Technol.* 48, 7436–7444.
- Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2016. Satellite-Based Spatiotemporal Trends in PM<sub>2.5</sub> Concentrations: China, 2004–2013. *Environ. Health Perspect.* 124, 184–192.
- Meng, R., Zhao, F., Sun, K., Zhang, R., Huang, C., Yang, J., 2015. Analysis of the 2014 “APEC Blue” in Beijing Using More than One Decade of Satellite Observations: Lessons Learned from Radical Emission Control Measures. *Remote Sens.* 7, 15224–15243.
- Mountrakis, G., Li, J., Lu, X.Q., Hellwich, O., 2018. Deep learning for remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* 145, 1–2.
- Nembrini, S., Koenig, I.R., Wright, M.N., 2018. The revival of the Gini importance? *Bioinformatics* 34, 3711–3718.
- Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., Pak, C., 2020. Deep learning-based PM<sub>2.5</sub> prediction considering the spatiotemporal correlations: A case study of Beijing, China. *Sci. Total Environ.* 699, 133561.
- Pang, J., Liu, Z., Wang, X., Bresch, J., Ban, J., Chen, D., Kim, J., 2018. Assimilating AOD retrievals from GOCI and VIIRS to forecast surface PM<sub>2.5</sub> episodes over Eastern China. *Atmos. Environ.* 179, 288–304.
- Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., Shi, Z., 2019. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* 337, 1309–1319.
- Qi, Y., Li, Q., Karimian, H., Liu, D., 2019. A hybrid model for spatiotemporal forecasting of PM<sub>2.5</sub> based on graph convolutional neural network and long short-term memory. *Sci. Total Environ.* 664, 1–10.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. *Nature* 566, 195–204.
- Schaap, M., Apituley, A., Timmermans, R.M.A., Koelmeijer, R.B.A., de Leeuw, G., 2009. Exploring the relation between aerosol optical depth and PM<sub>2.5</sub> at Cabauw, the Netherlands. *Atmos. Chem. Phys.* 9, 909–925.
- Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* 63, 1379–1389.
- Song, W., Jia, H., Huang, J., Zhang, Y., 2014. A satellite-based geographically weighted regression model for regional PM<sub>2.5</sub> estimation over the Pearl River Delta region in China. *Remote Sens. Environ.* 154, 1–7.
- Stafoggia, M., Bellander, T., Bucci, S., Davoli, M., de Hoogh, K., de' Donato, F., Gariazzo, C., Lyapustin, A., Michelozzi, P., Renzi, M., Scorticini, M., Shtain, A., Viegi, G., Kloog, I., Schwartz, J., 2019. Estimation of daily PM10 and PM2.5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ. Int.* 124, 170–179.
- Soh, P.-W., Chang, J.-W., Huang, J.-W., 2018. Adaptive Deep Learning-Based Air Quality Prediction Model Using the Most Relevant Spatial-Temporal Relations. *IEEE Access* 6, 38186–38199.
- Sulaymon, I.D., Zhang, Y., Hopke, P.K., Hu, J., Zhang, Y., Li, L., Mei, X., Gong, K., Shi, Z., Zhao, B., Zhao, F., 2021. Persistent high PM<sub>2.5</sub> pollution driven by unfavorable meteorological conditions during the COVID-19 lockdown period in the Beijing-Tianjin-Hebei region, China. *Environ. Res.* 198, 111186.
- Tao, M.H., Wang, J., Li, R., Wang, L.L., Wang, L.C., Wang, Z.F., Tao, J.H., Che, H.Z., Chen, L.F., 2019. Performance of MODIS high-resolution MAIAC aerosol algorithm in China: Characterization and limitation. *Atmos. Environ.* 213, 159–169.
- Theil, H., 1950. A rank-invariant method of linear and polynomial regression analysis. I, II and III. *Proc. K. Ned. Akad. Wet.*, 53.(pp. 386–392) (521–525, 1397–1412).
- Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* 46 (Supp 1), 234–240.
- van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M., 2016. Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environ. Sci. Technol.* 50, 3762–3772.
- van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., & winker, D.M. 2018. Global Annual PM<sub>2.5</sub> Grids from MODIS, MISR and SeaWiFS Aerosol Optical Depth (AOD) with GWR, 1998–2016. In: Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
- Wang, J., Christopher, S.A., 2003. Intercomparison between satellite-derived aerosol optical thickness and PM<sub>2.5</sub> mass: implications for air quality studies. *Geophys. Res. Lett.* 30.
- Wang, J., Li, J., Ye, J., Zhao, J., Wu, Y., Hu, J., Liu, D., Nie, D., Shen, F., Huang, X., Huang, D.D., Ji, D., Sun, X., Xu, W., Guo, J., Song, S., Qin, Y., Liu, P., Turner, J.R., Lee, H.C., Hwang, S., Liao, H., Martin, S.T., Zhang, Q., Chen, M., Sun, Y., Ge, X., Jacob, D.J., 2020a. Fast sulfate formation from oxidation of SO<sub>2</sub> by NO<sub>2</sub> and HONO observed in Beijing haze. *Nat. Commun.* 11, 2844.
- Wang, L.Q., Yu, S.C., Li, P.F., Chen, X., Li, Z., Zhang, Y.B., Li, M.Y., Mahmood, K., Liu, W.P., Chai, T.F., Zhu, Y.N., Rosenfeld, D., Seinfeld, J.H., 2020b. Significant wintertime PM<sub>2.5</sub> mitigation in the Yangtze River Delta, China, from 2016 to 2019: observational constraints on anthropogenic emission controls. *Atmos. Chem. Phys.* 20, 14787–14800.
- Wang, X., Dickinson, R.E., Su, L., Zhou, C., Wang, K., 2018. PM<sub>2.5</sub> Pollution in China and How It Has Been Exacerbated by Terrain and Meteorological Conditions. *Bull. Amer. Meteorol. Soc.* 99, 105–119.
- Wang, Y., Gao, W., Wang, S., Song, T., Gong, Z., Ji, D., Wang, L., Liu, Z., Tang, G., Huo, Y., Tian, S., Li, J., Li, M., Yang, Y., Chu, B., Petäjä, T., Kermani, V.-M., He, H., Hao, J., Kulmala, M., Wang, Y., Zhang, Y., 2020c. Contrasting trends of PM<sub>2.5</sub> and surface-ozone concentrations in China from 2013 to 2017. *Natl. Sci. Rev.* 7, 1331–1339.

- Wang, Z., Zhou, Y., Zhao, R., Wang, N., Biswas, A., Shi, Z., 2021. High-resolution prediction of the spatial distribution of PM<sub>2.5</sub> concentrations in China using a long short-term memory model. *J. Clean. Prod.* 297.
- Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019a. Estimating 1-km-resolution PM<sub>2.5</sub> concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* 231.
- Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M., 2021. Reconstructing 1-km-resolution high-quality PM<sub>2.5</sub> data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. *Remote Sens. Environ.* 252.
- Wei, J., Sun, L., Huang, B., Bilal, M., Zhang, Z., Wang, L., 2018. Verification, improvement and application of aerosol optical depths in China Part 1: Inter-comparison of NPP-VIIRS and Aqua-MODIS. *Atmos. Environ.* 175, 221–233.
- Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., Chi, T., 2019. A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Sci. Total Environ.* 654, 1091–1099.
- Wu, X.L., Wang, Y., He, S., Wu, Z., 2020. PM<sub>2.5</sub> /PM<sub>10</sub> ratio prediction based on a long short-term memory neural network in Wuhan, China. *Geosci. Model Dev.* 13, 1499–1511.
- Xiao, L., Lang, Y., Christakos, G., 2018. High-resolution spatiotemporal mapping of PM<sub>2.5</sub> concentrations at Mainland China using a combined BME-GWR technique. *Atmos. Environ.* 173, 295–305.
- Xiao, Q., Wang, Y., Chang, H.H., Meng, X., Geng, G., Lyapustin, A., Liu, Y., 2017. Full-coverage high-resolution daily PM<sub>2.5</sub> estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sens. Environ.* 199, 437–446.
- Xu, J.F., Zhu, Y., Zhong, R.H., Lin, Z.X., Xu, J.L., Jiang, H., Huang, J.F., Li, H.F., Lin, T., 2020. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sens. Environ.* 247, 18.
- Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T., Zhang, Q., 2019. Spatiotemporal continuous estimates of PM<sub>2.5</sub> concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations. *Environ. Int.* 123, 345–357.
- Yao, F., Wu, J., Li, W., Peng, J., 2019. A spatially structured adaptive two-stage model for retrieving ground-level PM<sub>2.5</sub> concentrations from VIIRS AOD in China. *ISPRS J. Photogramm. Remote Sens.* 151, 263–276.
- Yang, Q., Yuan, Q., Yue, L., Li, T., Shen, H., Zhang, L., 2020. Mapping PM<sub>2.5</sub> concentration at a sub-km level resolution: A dual-scale retrieval approach. *ISPRS J. Photogramm. Remote Sens.* 165, 140–151.
- Ye, W.-F., Ma, Z.-Y., Ha, X.-Z., 2018. Spatial-temporal patterns of PM(2.5) concentrations for 338 Chinese cities. *Sci. Total Environ.* 631–632, 524–533.
- Yu, W., Liu, Y., Ma, Z., Bi, J., 2017. Improving satellite-based PM<sub>2.5</sub> estimates in China using Gaussian processes modeling in a Bayesian hierarchical setting. *Sci. Rep.* 7, 7048.
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* 241.
- Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., Xu, X., Wang, J., He, H., Liu, W., Ding, Y., Lei, Y., Li, J., Wang, Z., Zhang, X., Wang, Y., Cheng, J., Liu, Y., Shi, Q., Yan, L., Geng, G., Hong, C., Li, M., Liu, F., Zheng, B., Cao, J., Ding, A.,
- Gao, J., Fu, Q., Huo, J., Liu, B., Liu, Z., Yang, F., He, K., Hao, J., 2019. Drivers of improved PM<sub>2.5</sub> air quality in China from 2013 to 2017. *Proc. Natl. Acad. Sci. USA* 116, 24463–24469.
- Zhang, T., Zhu, Z., Gong, W., Zhu, Z., Sun, K., Wang, L., Huang, Y., Mao, F., Shen, H., Li, Z., Xu, K., 2018. Estimation of ultrahigh resolution PM<sub>2.5</sub> concentrations in urban areas using 160 m Gaofen-1 AOD retrievals. *Remote Sens. Environ.* 216, 91–104.
- Zhang, Y., Li, Z., 2015. Remote sensing of atmospheric fine particulate matter (PM<sub>2.5</sub>) mass concentration near the ground from satellite observation. *Remote Sens. Environ.* 160, 252–262.
- Zhang, Y.B., Chen, X., Yu, S.C., Wang, L.Q., Li, Z., Li, M.Y., Liu, W.P., Li, P.F., Rosenfeld, D., Seinfeld, J.H., 2021. City-level air quality improvement in the Beijing-Tianjin-Hebei region from 2016/17 to 2017/18 heating seasons: Attributions and process analysis. *Environ. Pollut.* 274, 12.
- Zhang, H., Wang, Y., Hu, J., Ying, Q., Hu, X.-M., 2015. Relationships between meteorological parameters and criteria air pollutants in three megacities in China. *Environ. Res.* 140, 242–254.
- Zhao, J., Deng, F., Cai, Y., Chen, J., 2019. Long short-term memory - Fully connected (LSTM-FC) neural network for PM<sub>2.5</sub> concentration prediction. *Chemosphere* 220, 486–492.
- Zheng, B., Zhang, Q., Zhang, Y., He, K., Wang, K., Zheng, G., Duan, F., Ma, Y., Kimoto, T., 2015. Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China. *Atmos. Chem. Phys.* 15, 2031–2049.
- Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., Xu, B., 2016. Attention-based bidirectional long short-term memory networks for relation classification. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 207–212.
- Zhou, C., He, Y., Wang, K., 2018. On the suitability of current atmospheric reanalyses for regional warming studies over China. *Atmos. Chem. Phys.* 18, 8113–8136.

## Further reading

- Chen, G., Li, S., Knibbs, L.D., Hamm, N.A.S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y., 2018. A machine learning method to estimate PM<sub>2.5</sub> concentrations across China with remote sensing, meteorological and land use information. *Sci. Total Environ.* 636, 52–60.
- Jiang, T., Chen, B., Nie, Z., Ren, Z., Xu, B., Tang, S., 2021. Estimation of hourly full-coverage PM<sub>2.5</sub> concentrations at 1-km resolution in China using a two-stage random forest model. *Atmos. Res.* 248.
- Li, B., Yuan, H., Feng, N., Tao, S., 2009. Comparing MODIS and AERONET aerosol optical depth over China. *Int. J. Remote Sens.* 30, 6519–6529.
- van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verdúzco, C., Villeneuve, P.J., 2010. Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. *Environ. Health Perspect.* 118, 847–855.
- Wei, J., Peng, Y., Mahmood, R., Sun, L., Guo, J., 2019b. Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products. *Atmos. Chem. Phys.* 19, 7183–7207.