Tracking hourly PM₂.₅ using geostationary satellite sensor images and multiscale spatiotemporal deep learning

Zhige Wang [a], Ce Zhang [b,c,*], Su Ye [a], Rui Lu [a], Yulin Shangguan [a], Tingyuan Zhou [d], Peter M. Atkinson [d,e], Zhou Shi [a,f,**]

[a] College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China
[b] School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK
[c] UK Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK
[d] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[e] Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[f] Key Laboratory of Environment Remediation and Ecological Health, Ministry of Education, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

ARTICLE INFO

ABSTRACT

Spatially continuous fine particulate matter (PM₂.₅) mapping with hourly updated is essential for monitoring environmental pollution and promoting public health. The intensive observation of geostationary satellite enables accurate estimation of PM₂.₅ at a fine-scale. However, current estimation models are still limited by their weak transferability and hard to provide a robust hourly PM₂.₅ estimation. In this research, we aim to estimate the daytime PM₂.₅ concentrations at fine spatial and temporal resolution (1 km and hourly) in mainland China using an improved deep learning algorithm and the AOD products from geostationary satellite Himawari-8. An Adaptive Spatio-Temporal Multiscale Neural Network (ASTMNN) which contains three sub-networks and an adaptive weight was proposed to capture the spatiotemporal heterogeneity of hourly PM₂.₅. The three sub-networks of ASTMNN are spatial adjacency module (SaM), temporal adjacency module (TaM) and global module (GM), which used to incorporate the information from spatial neighborhood, temporal neighborhood, and global spatiotemporal range, respectively. And the weight function combines the outputs from the three subnetworks, where the weights were adaptively trained from the model optimization. The proposed model outperformed most current hourly PM₂.₅ estimation models with the sample-based, time-based, and site-based cross-validation (CV) R^2 of 0.94, 0.89 and 0.83, respectively. Besides, we used our PM₂.₅ product to track extreme dust events. Our findings provide valuable implications for tracking continuous variation in particulate pollution using geostationary satellites.

1. Introduction

Fine particulate matter with aerodynamic diameter less than or equal to 2.5 μm (PM₂.₅) could significantly impact the environment and the climate system and lead to serious health conditions (Burnett et al., 2018). According to the study of Global Burden of Disease (GBD) (Murray et al., 2020), air pollution has become the fourth risk factor for global attributable deaths in 2019 among which ambient PM₂.₅ is much more severe in Asia and Africa (Wang et al., 2022a). Great efforts have been made by the Chinese government to control PM₂.₅ pollution in China in recent years. For example, numerous ground monitoring

stations have been built in China since late 2012, which greatly improved the ability to monitor air pollution (Zhang et al., 2012). Accurate and timely PM₂.₅ information is of enormous importance for related health burden studies and air pollution management. However, problems remain about the limited spatial coverage of the current site-level dataset, particularly in rural areas and the northwest China where the stations are unevenly and sparse distributed. Therefore, it is urgent to create spatially continuous and temporally dense PM₂.₅ products to better track and study the evolution of PM₂.₅.

Compared with site-based monitoring, satellite images can be used to monitor spatially continuously over large areas and at a fine resolution.

Fig. 1. The spatial distributions of ground-level PM$_{2.5}$ sites (blue dots), Aerosol Robotic Network sites (red dots), and regions (orange borders) mentioned in this research. BTH: Beijing-Tianjin-Hebei; FWP: Fenwei Plain; YRD: Yangtze River Delta; PRD: Pearl River Delta; SCB: Sichuan Basin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The aerosol optical depth (AOD), which denotes the extinction effects of all particulates in the atmosphere, is an efficient indicator for PM$_{2.5}$ estimation (Eck et al., 1999; Engel-Cox et al., 2004; Della Ceca et al., 2018). In recent years, satellite-derived AOD products, such as the Multi-angle Imaging Spectroradiometer (MISR), and Moderate Resolution Imaging Spectroradiometer (MODIS), were adopted to estimate surface PM$_{2.5}$ (Sorek-Hamer et al., 2013; Xiao et al., 2017; van Donkelaar et al., 2006; Wei et al., 2019). While these polar orbit satellites made it possible for estimating surface PM$_{2.5}$ on a daily basis with an acceptable accuracy, the rapidly changing nature of PM$_{2.5}$ has called for mapping PM$_{2.5}$ at an hourly level (Vu et al., 2022), particularly with the emerging geostationary satellites (e.g., Himawari-8; and Geostationary Operational Environmental Satellite-16, GOES-16). Among them, Himawari-8 provides hourly/sub-hourly AOD observations for most parts of China and was utilized widely to estimate hourly surface PM$_{2.5}$ concentration in China (Mao et al., 2021; Wei et al., 2021; Xu et al., 2022). However, past research focusing on estimating hourly PM$_{2.5}$ in China remains scarce due to the limited observation range of the geostationary satellites, with most hourly PM$_{2.5}$ estimation in China covering only the central and eastern parts of China.

Numerous statistical approaches based on the relationship between PM$_{2.5}$ and AOD have been applied for mapping PM$_{2.5}$ from a large scale, which could be categorized into two broad groups: 1) global models, and 2) local models. Global models which target on establishing a global PM$_{2.5}$-AOD relationship that could be interpolated for each pixel's prediction. Earlier studies commonly used the simple linear regression to quantify the relationship (Gupta et al., 2006; Sorek-Hamer et al., 2013). Some advanced linear models and machine learning (ML) techniques were introduced to better capture the relationship between PM$_{2.5}$ and AOD by integrating an uncertainty modeling, such as generalized additive model (GAM), Random Forest (RF) and neural network models (e. g., Convolutional Neural Network, CNN) (Paciorek et al., 2008; Liu et al., 2009; Lee et al., 2011; Gupta and Christopher, 2009; Wang et al., 2021). Some advanced approaches encoded the predicted location in a spatiotemporal cube into the ML model as additional independent variables, such as using the geospatial-temporal joint code methods (Yang et al., 2022). Generally, global methods ignored local contextual information while PM$_{2.5}$ presented high heterogeneity led by various meteorological factors (e.g., precipitation and wind direction) and human activities (e.g., transportation and gas emissions) (Lei et al., 2022; Wang et al., 2022b).

Local models predict PM$_{2.5}$ by capturing the local relationship between PM$_{2.5}$ and explanatory variables at a per-pixel level from each pixel's spatial or spatiotemporal neighborhoods, relying on the techniques like geographically weighted regression (GWR) (Ma et al., 2014; Song et al., 2014), and fusing spatial adjacency information and ML methods (Di et al., 2016; Li et al., 2020; Wang et al., 2022c). Recently, with more repeated satellite image acquisition, some innovations were made by integrating temporal dependence into the spatially explicit methods. For example, the geographically and temporally weighted regression (GTWR) was applied to explore the spatiotemporal variation of the PM$_{2.5}$-AOD relationship (Bai et al., 2016; He and Huang, 2018). Local models only consider local variation in the prediction model and likely introduce a great deal of uncertainties in regions where PM$_{2.5}$ are distributed sparsely (Ma et al., 2022). Moreover, local models are less robust to outlier and are more sensitive to multicollinearity due to their smaller sample size (Harris et al., 2010; Páez et al., 2011). In summary, the current PM$_{2.5}$ estimation models considered only the spatiotemporal dependence of PM$_{2.5}$ at a single scale, which is insufficient for accommodating geographical and temporal complication of hourly PM$_{2.5}$.

Therefore, the main contribution of this research is to design an efficient hourly PM$_{2.5}$ estimation model that aggregates multiscale (i.e., local and global scale) spatiotemporal information in PM$_{2.5}$ and its auxiliary variables, and further generates the full-coverage hourly PM$_{2.5}$ products across mainland China. We firstly filled the gaps in the hourly Himawari-8 AOD product. Subsequently, an Adaptive Spatio-Temporal Multiscale Neural Network (ASTMNN) model with three-subnetworks was designed for local spatial feature learning, local temporal feature

Table 1
Summary information of the data used in this study.

Variables	Unit	Spatial Resolution	Temporal Resolution	Variables type	Source
PM$_{2.5}$	µg m^{-3}	–	hourly	site-based	CNEMC
AHI AOD	–	5 km × 5 km	hourly	satellite-	JAXA
MAIAC AOD	–	1 km × 1 km	hourly	derived	MCD19A2
GEOS FP AOD	–	0.25° × 0.3124°	3-hour	Reanalysis	NASA/GMAO GEOS-5
Total Aerosol Extinction AOT [550 nm]	–	0.5° × 0.625°	1-hour		MERRA-2
Total AOD at 550 nm	–	0.4° × 0.4°	1-hour	Forecasts	CAMS global atmospheric composition forecasts
BLH	m	0.25° × 0.25°	hourly	Reanalysis	ERA5 hourly data on single levels from 1959 to present
v10	m s^{-1}	0.1° × 0.1°	hourly		ERA5-Land hourly data from 1950 to present
u10	m s^{-1}				
SP	Pa				
TEMP	K				
PRE	m				
DTEM	kg kg^{-1}				
ET	m of water equivalent				
NDVI	–	1 km × 1 km	monthly	Satellite-	MOD13A3
LUCC	–	500 m × 500 m	annual	derived	MCD12Q1
DEM	m	90 m × 90 m	annual		RESDC

AHI AOD: Advanced Himawari Imager Aerosol Optical Depth; AOT: Aerosol Optical Thickness; BLH: Boundary Layer Height; v10: 10 m v-components of wind; u10: 10 m u-components of wind; SP: Surface pressure; TEMP: 2 m surface air temperature; PRE: Total precipitation; DTEM: 2 m dewpoint temperature; ET: Total evaporation; NDVI: Normalized difference vegetation index; LUCC: Land use/cover change; DEM: Digital elevation model.

learning, and global feature learning. Lastly, an adaptive weighting function was created to effectively combine the three learning features for final PM$_{2.5}$ estimation. The model performance was evaluated by employing multiple cross validations. Finally, we generated hourly PM$_{2.5}$ products at 1-km spatial resolution for mainland China and applied the output to track an extreme haze event.

2. Data and methods

2.1. PM$_{2.5}$ measurements

Hourly PM$_{2.5}$ ground observations across China from January 1st to December 31th in 2021 were acquired from the China National Environmental Monitoring Center (CNEMC, https://www.cnemc.cn/). There are 1,748 *in situ* stations in total. The stations are mostly located in the north and southeast of China, while fewer are situated in the southwest and northwest of China (Fig. 1). The PM$_{2.5}$ observation data from the CNEMC is limited to the mainland China. We excluded the null PM$_{2.5}$ values and only analyzed the data collected from 8:00 to 17:00 (UTC + 8).

2.2. AOD data

2.2.1. Himawari AHI AOD

Himawari-8 is a geostationary meteorological satellite launched by Japan in 2014. The multispectral imager (i.e., Advanced Himawari Imager, AHI) with 16 channels is aboard the satellites. They provide visible light and infrared images of the Western Pacific Ocean, East and Southeast Asia and Oceania (80°E to 160°W and 60°N to 60°S) (Fig. S1) and sub-hourly aerosol observations (Kikuchi et al., 2018; Letu et al., 2020). We used the level-3 hourly (500 nm) AOD product (version 3.1) with 5 km spatial resolution download from the Japan Aerospace eXploration Agency (JAXA, https://www.eorc.jaxa.jp/ptree/). The band "AOT_Merged" which provides hourly AOD at 500 nm were selected. The 500 nm AOD were transformed to 550 nm to maintain the same wavelength with other multi-source AOD products, following the method of Jiang et al. (2021) and Zhang et al. (2019).

2.2.2. MODIS MAIAC AOD

MODIS onboard the Terra and Aqua satellites provides fine

resolution daily aerosol observations. We used MODIS AOD products retrieved using the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm (Lyapustin et al., 2011). The MODIS Collection 6 MAIAC AOD data (MCD19A2) produce 550 nm AOD with spatial resolution of 1 km (Lyapustin et al., 2018) and are available from https://ladsweb.modaps.eosdis.nasa.gov/. The band "Optcial_Depth_055" with QA$_{CloudMask}$ = Clear and QA$_{AdjacencyMask}$ = Clear were chosen. In this research, we extracted the data at the different transit times of the two satellites according to the "Orbit_time_stamp" in the files to obtain hourly AOD during daytime (i.e., UTC + 8 8:00–17:00). The coverage of hourly MAIAC AOD is shown in Fig. S2.

The accuracy of two satellite-based AOD products was evaluated using Aerosol Robotic Network (AERONET) AOD data (Giles et al., 2019) across China (Fig. 1 and Table S1). The AERONET AOD at 440 nm and 870 nm were used to obtain the AOD at 550 nm by calculating the Angstrom Exponent (referring to Text S1 in the supplementary for detail). The data were download from https://aeronet.gsfc.nasa.gov/. Fig. S1 shows that the two satellite AOD products are consistent with *in situ* AOD observation.

2.2.3. Reanalysis AOD products

Since the AHI AOD does not cover all China, we adopted some other hourly AOD products from reanalysis datasets, including the 3-hourly Goddard Earth Observing System (GEOS) Forward Processing (FP) AOD data, hourly Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) AOD, and hourly Copernicus Atmosphere Monitoring Service (CAMS) AOD (Molod et al., 2015; Gelaro et al., 2017; Inness et al., 2019). The detail information of these AOD data are provided in the Text S2 and Table 1.

2.3. Ancillary data

Eight meteorological variables and three surface cover related variables were selected in this research due to their strong affects on the spatiotemporal variation of PM$_{2.5}$ documented in the existing study (Chen et al., 2018). The variables related to meteorology included boundary layer height (BLH), relative humidity (RH), total evaporation (ET), 2 m surface air temperature (TEMP), precipitation flux (PRE), surface pressure (SP), the 10 m u-components (u10) of wind, and the 10 m v-components (v10) of wind. The data were obtained from the fifth-
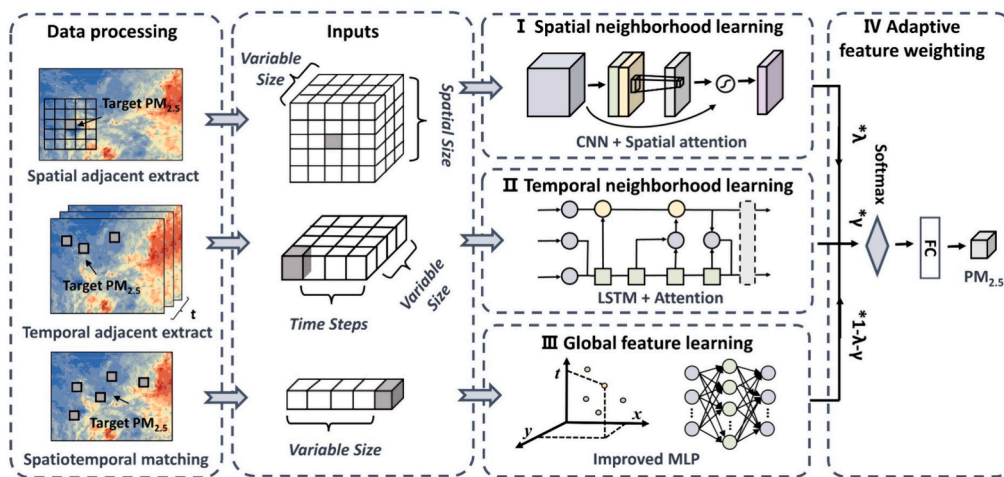
Fig. 2. The structure of the ASTMNN model.

generation ECMWF global atmospheric reanalysis products (ERA5) (Muñoz-Sabater et al., 2021). The RH was calculated using the 2 m dewpoint temperature and TEMP. The surface cover related variables, including normalized difference vegetation index (NDVI) from MODIS/Terra vegetation indices (MOD13A3), land use/cover change (LUCC) from MODIS/Terra and Aqua Land Cover Dynamics (MCD12Q1), and digital elevation model (DEM) data, were adopted for PM$_{2.5}$ estimation. The details of these dataset are listed in Table 1. All variables passed the correlation test and significant at the 0.01 confidence level (Table S2).

2.4. Data preprocessing

We firstly reconstructed the hourly AOD data using multi-source AOD and other related covariates, then corrected the hourly AOD based on the hourly MAIAC AOD. To ensure the accuracy and consistency of our hourly AOD products, we validated the reconstructed AOD against AERONET AOD measurements (Fig. S3) and assessed the ability of the reconstructed data to preserve the spatial patterns of AOD (Fig. S4 and S5). Detailed information on AOD reconstruction and validation are shown in Text S3.

All variables were resampled to a 1 km grid using bilinear interpolation in order to estimate PM$_{2.5}$ concentration of 1 km spatial resolution. We also created a gridded dataset with 1 km spatial resolution and extracted the PM$_{2.5}$ value at each site for the corresponding pixel. If one pixel included multiple sites, the PM$_{2.5}$ values were averaged.

For global feature learning, the multiple remote sensing variables were extracted for the corresponding PM$_{2.5}$ pixel according to spatial and temporal location. For local spatial feature learning, we moved a sliding window of size $w \times w$ across every pixel of target PM$_{2.5}$ such as to extract the adjacent remote sensing variables around the target PM$_{2.5}$ value. For local temporal learning, temporally discontinuous PM$_{2.5}$ data were removed. Then, all independent variables of the previous t hours were extracted for the corresponding PM$_{2.5}$ pixel for temporal feature learning.

The final valid training dataset included 6,049,940 samples during 2021. It is time consuming for model training using such a large amount of data. Thus, we sampled randomly 8-day of sample data from each month dataset. This process was repeated 3 times until all samples were processed. And the set of data samples with the best performance was used for final model calibration and validation (Table S3 displays the performances in different datasets). Finally, a set of 1,604,994 samples were selected for model calibration. All data were processed using Python, R and ArcGIS 10.3.

2.5. Methodology

In this research, the ASTMNN framework was designed for PM$_{2.5}$ estimation. It contains three subnetworks (Fig. 2):

(1) a spatial adjacency module (SaM) for spatial neighborhood feature learning,
(2) a temporal adjacency module (TaM) for temporal neighborhood feature learning,
(3) a global module (GM) for spatiotemporal feature learning at the global scale.

The first two modules were designed for fixed local scale feature learning and the third module was designed for global scale feature learning. In the ASTMNN, the learning features from the three modules are combined adaptively by sharing a common loss function and the weights of each module are adjusted automatically during network optimization.

2.5.1. Spatial adjacency module

The SaM was designed with an input layer, three Conv-ReLU-BN layers, one spatial attention composite layer, one dropout layer and a fully connected layer. A Conv-ReLU-BN layer consists of a convolutional (Conv) layer, ReLU and a batch normalization (BN) layer. The CNN was chosen as the basic network of SaM to capture the spatial features of PM$_{2.5}$ since it operates by exploiting local spatial information (Albawi et al., 2017).

Since the invalid neighboring information may introduce bias into the final results, we adopted the spatial attention mechanism to emphasize valid information in SaM. The spatial attention feature matrix was generated using the method from Woo et al. (2018). The max-pooling operation and average-pooling operation were first adopted across the channel dimension, which has been demonstrated useful for improving target detection (Zagoruyko and Komodakis, 2016). Then, the two outputs were concatenated as a new feature descriptor. The final spatial attention feature map was output by adding a convolution layer to the concatenated feature as followed:

$$SA(F) = \sigma(\text{Conv}_{9 \times 9}([\text{AvgPool}(Fc); \text{MaxPool}(Fc)])) \tag{1}$$

where SA $\in R^{w \times w \times 1}$ and $Fc \in R^{w \times w \times c}$ represent the spatial attention matrix and 3-D input feature matrix, respectively. The $w \times w$ denotes the spatial size of the input matrix and c denotes the features amount. The σ and Conv$_{9 \times 9}$ denote the sigmoid function and the 9×9 convolutional layer. To set the spatial size (w), referring to previous studies (Lei et al., 2022), we experimented with different values of w (Table S4) and the

model was found to perform the best when *w* was 9. Therefore, the input data size of the spatial adjacency module is $9 \times 9 \times 12$. Here, 9×9 refers to the range of pixels around the target $PM_{2.5}$, and 12 represents the 12 independent variables.

2.5.2. Temporal adjacency module

In the TaM, we built an attention-based Bidirectional LSTM (BiLSTM) model with a goal of uncovering features hidden in the time-series data.

The LSTM model was proposed by Hochreiter and Schmidhuber (1997) to process long sequence data. The adaptive gating mechanism (including input gate i_t, forget gate f_t, and output gate o_t), which decides whether and how much the state of the LSTM block will be updated, is the essential idea of LSTM. Bidirectional LSTM (BiLSTM) was designed to overcome the flaw of single flowing direction in LSTM, which enabled the sequence to flow in two directions (Graves et al., 2005, 2013). We set different time steps and the model performances with different time steps are listed in Table S5. The results revealed that the model performed the best when *t* was set to 3. Consequently, the input data size for the temporal adjacency module is 3×12. Since AOD products are only available from 8:00 to 17:00, we used the AOD value at 8:00 as a substitute for the AOD values at previous time steps when estimating the target $PM_{2.5}$ at 8:00.

The attention layer was adopted behind the Bi-LSTM layer to capture the more crucial features from the multi-timesteps. The main idea of the attention layer proposed by Luong et al. (2015) is to calculate the weight vector a_t using the similarity between the input vectors and target state. Text S4 provided the detailed formulation of attention layer used in this research.

2.5.3. Global module

Different from the preceding modules which focus on limited local scale features, the GM focuses more on the spatiotemporal characteristics of $PM_{2.5}$ throughout the entire dataset. Therefore, in this module, we encoded the geospatial and temporal location of target $PM_{2.5}$ and fed them into a multilayer perceptron (MLP) neural network (Khotanzad and Chung, 1998) to learn the global information. Unlike ordinary local regressions that use the spatial distance of neighboring sites as inverse distance weights, the geospatial and temporal encoding determines the proximity of locations based on the similarity of spatial and temporal location (Yang et al., 2022).

The geospatial location is based on the entire dataset to denote the global spatial characteristics of $PM_{2.5}$. For the spatial location, the longitude ε_i and latitude θ_i of sample *i* are first converted from degrees to radians (φ_i, λ_i):

$$(\varphi_i, \lambda_i) = \left(\varepsilon_i \frac{\pi}{180}, \theta_i \frac{\pi}{180}\right) \quad (2)$$

where φ_i and λ_i denote the longitude and latitude in radians of sample *i*, respectively. Then, the spatial location $S_{i(x,y)}$ is calculated by transferring the polar coordinate into Cartesian coordinate:

$$S_{i(x,y)} = [S_{ix}, S_{iy}] \quad (3)$$

$$= [r\cos(\varphi_i)\cos(\lambda_i), r\cos(\varphi_i)\sin(\lambda_i)]$$

where *r* denotes the radius of the Earth. Since we have introduced the DEM as an ancillary variable, the Cartesian coordinates here only include *x* and *y*, and do not consider the *z* axis.

The temporal location is based on the whole time series to denote the global temporal characteristics of $PM_{2.5}$. The first hour (i.e., 1:00) and last hour (i.e., 24:00) in one day are both midnight and share the same meteorological conditions, the temporal location T_i is calculated using the time radian difference for each hour in the whole day:

$$T_i = \sin\left(2\pi \frac{t_i}{T}\right) \quad (4)$$

Table 2
Setting of the ASTMNN.

Layer Name	Layer	Kernel Size, stride	Output W×H×N
SaM			
Conv	Input	–	$9 \times 9 \times 12$
	Conv2D + ReLU + BN	3×3, 1	$7 \times 7 \times 128$
Conv1	Conv2D + ReLU + BN	3×3, 1	$5 \times 5 \times 256$
Conv2	Conv2D + ReLU + BN	3×3, 1	$3 \times 3 \times 256$
	Dropout	–	$3 \times 3 \times 256$
Conv3	Conv2D + ReLU + BN	2×2, 3×1	$1 \times 3 \times 512$
Spatial attention (SA)	Max pool	1×3, –	$1 \times 3 \times 1$
	Average pool	1×3, –	$1 \times 3 \times 1$
	Concat [Max, Ave]	–	$1 \times 3 \times 2$
	Conv2D + ReLU	$7 \times 7, 3 \times 1$	$1 \times 3 \times 1$
	Multiply [Conv3, SA]	–	$1 \times 3 \times 512$
TaM			
Bi-LSTM1	Input	3	3×12
	BiLSTM1	512	3×1024
Bi-LSTM2	BiLSTM2	512	3×1024
Bi-LSTM3	BiLSTM3	256	3×512
Attention	Attention Layer	–	256
GM			
FC1	Input	–	16
	FC + ReLU	–	128
FC2	FC + ReLU	–	256
	Dropout	–	256
FC3	FC + ReLU	–	512

where t_i represents the time position of sample *i* (i.e., hour of the day), and *T* means the total number of times in question (i.e., 24-hour). Since we only estimated the daytime (i.e., 8:00 to 17:00) $PM_{2.5}$, we employed the sine function to encode the temporal location instead of cosine function used in previous studies (Wei et al., 2020) to avoid the same temporal location in the morning and afternoon.

2.5.4. Adaptive weight function

After extracting the multiscale learning features from the three subnetworks, we built an adaptive feature weighting function to better combine the features from three components. The outputs of the three subnetworks were fed into a fully-connected (FC) layer to ensure the same dimension of three outputs. Then the score weighted method was adopted to merge the three outputs and the final output is obtained by a weighted sum of the three modules. We set weight parameters which can be trained adaptively during model optimization. Specifically, the parameters with the minimal loss are selected as the final weights of each output of the three modules. The adaptive feature weighting function could be represented as:

$$F = \sigma(\lambda_1 F_s + \lambda_2 F_t + (1 - \lambda_1 - \lambda_2)F_g) \quad (5)$$

where *F* is the final output, σ is the activation function, and λ_1 and λ_2 are weighting parameters in the range of [0,1], which will be adaptively adjusted during the model-optimization stage.

In this research, we used the package Keras in Python to perform the ASTMNN model. Table 2 provides the detailed deployment and output of each module in the model.

The setting of Attention Layer according to Luong et al. (2015) in TaM refer to https://github.com/philipperemy/keras-attention

2.6. Model evaluation

The performance of the proposed model was assessed through adopting sample-based CV, time-based CV, and site-based CV. In sample-based CV, the dataset was partitioned into 10 equal folds
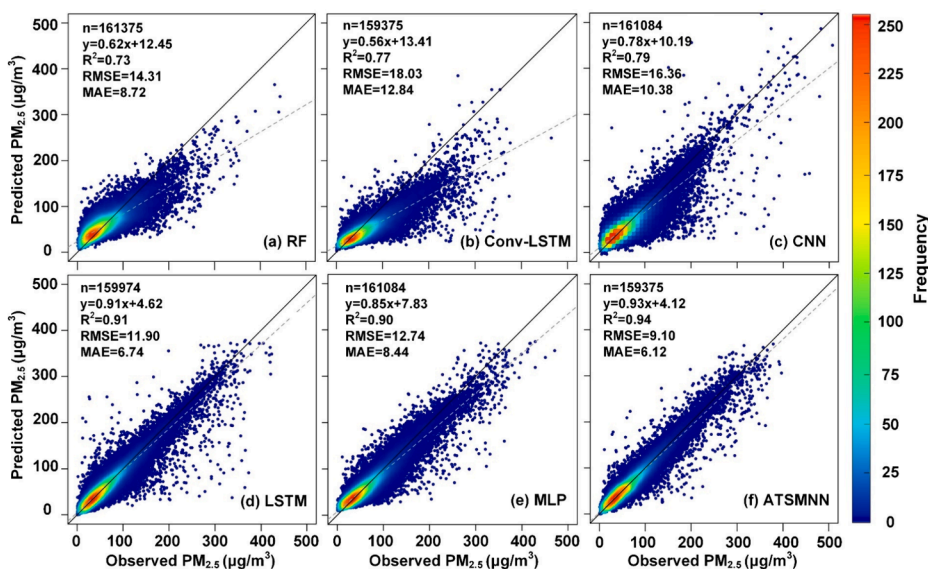
Fig. 3. Density scatterplots of sample-based CV results. The linear regression fitting line and the 1:1 line are shown by the dashed and solid lines, respectively. The proportion of the number of points is represented as the color of the points. The number of samples n, coefficient of determination R^2, RMSE ($\mu g \ m^{-3}$) and MAE ($\mu g \ m^{-3}$) are provided.

Table 3
Station-based and time-based CV results of different models.

	Site-based CV					
	RF	LSTM	CNN	MLP	Conv-LSTM	ASTMNN
R^2	0.720	0.817	0.763	0.814	0.753	**0.832**
MAE	8.932	**6.694**	7.770	7.010	14.100	6.720
RMSE	14.785	11.229	12.798	12.307	22.361	**10.930**
RPE	45.67 %	35.57 %	43.31 %	39.42 %	45.48 %	**34.88 %**
Slope	0.606	0.830	0.792	0.780	0.579	**0.861**
Bias	12.686	5.025	**4.572**	5.625	16.651	4.765

	Time-based CV					
	RF	LSTM	CNN	MLP	Conv-LSTM	ASTMNN
R^2	0.751	0.869	0.813	0.804	0.773	**0.890**
MAE	14.036	7.939	9.802	10.590	13.845	**7.739**
RMSE	21.559	15.053	17.707	17.806	21.541	**12.110**
RPE	37.65 %	27.32 %	32.01 %	32.73 %	40.31 %	**25.27 %**
Slope	0.605	0.851	0.868	0.809	0.570	**0.831**
Bias	23.204	6.641	6.173	8.811	20.922	**5.483**

randomly, with one fold allocated for validation and the remaining nine folds utilized for model training, iterated ten times. The time-based CV and sited-based CV were designed to test the spatiotemporal transferability of our model. Time-based CV involved a similar process but focused on temporal splits of the dataset into 10 folds. Specifically, it was performed by using one hour of data for validation, with the others used for model fitting. And this process was repeated, until each of the ten sets (data for each hour from 8:00–17:00) had been used for validation once. The site-based CV segmented the dataset into 10 folds based on site locations randomly to perform the similar process. The CV results of our model were compared to other typical machine or deep learning models (i.e., RF, MLP, CNN, LSTM and Conv-LSTM). We also compared the results of sample-based CV to the results obtained by other researchers. The determination coefficient (R^2), root mean squared error (RMSE), mean absolute error (MAE), and relative prediction error (RPE, %) were used to assess the model performance (the formulas are presented in Text S5).

3. Results

3.1. Model performance

3.1.1. Overall accuracy

The model performance of STAMNN was compared to five typical machine or deep learning models (Fig. 3). The R^2 values of the six models ranged from 0.73 to 0.94. The RF model produced the lowest accuracy, suggesting that the ML methods applied without considering the spatiotemporal variation in $PM_{2.5}$ are insufficient for predicting hourly $PM_{2.5}$ with severe fluctuations. The Conv-LSTM was less accurate than our model with an R^2 of 0.77. Although Conv-LSTM incorporated spatiotemporal adjacency information, it failed to explore the effectiveness or relevance of this information to the target features. This highlights the significance of implementing an adaptive weighting algorithm. Our model is the most accurate amongst all models with an R^2 of 0.94 and RMSE of 9.10 $\mu g \ m^{-3}$ and low estimated bias (MAE of 6.12 $\mu g \ m^{-3}$). This demonstrates the effectiveness of ASTMNN model through an integration of improved local and global features via a weighting function.

The site-based and time-based CV results are shown in Table 3. Due to the strong spatiotemporal variation in hourly $PM_{2.5}$, it is challenging to transfer $PM_{2.5}$ models at different times or sites. The prediction accuracy of all models decreased in terms of the time-based and site-based CV. The time-based CV results of all models outperformed the site-based CV results, indicating that the spatial variation in $PM_{2.5}$ is more difficult to capture than the temporal variation. The accuracy of the ASTMNN model in terms of the site-based CV was higher than for the other models with an R^2 of 0.832 and RMSE of 10.930 $\mu g \ m^{-3}$. In general, by extracting the local and global features, the ASTMNN model produced a robust performance under different validation methods and is sufficient to support $PM_{2.5}$ estimation in space and time without observation data.

3.1.2. Model accuracy across space and through time

The model accuracy was also evaluated across space and through time. According to Fig. 4, an R^2 greater than 0.85 was present in more than 72 % of the stations, and an R^2 greater than 0.8 was present in more than 81 % of the stations. Meanwhile, most of the stations exhibited low estimation biases with the MAE of 89 % stations being less than 9 μg m^{-3}. Only ~ 2.9 % of the stations, located in in southwest China, had a
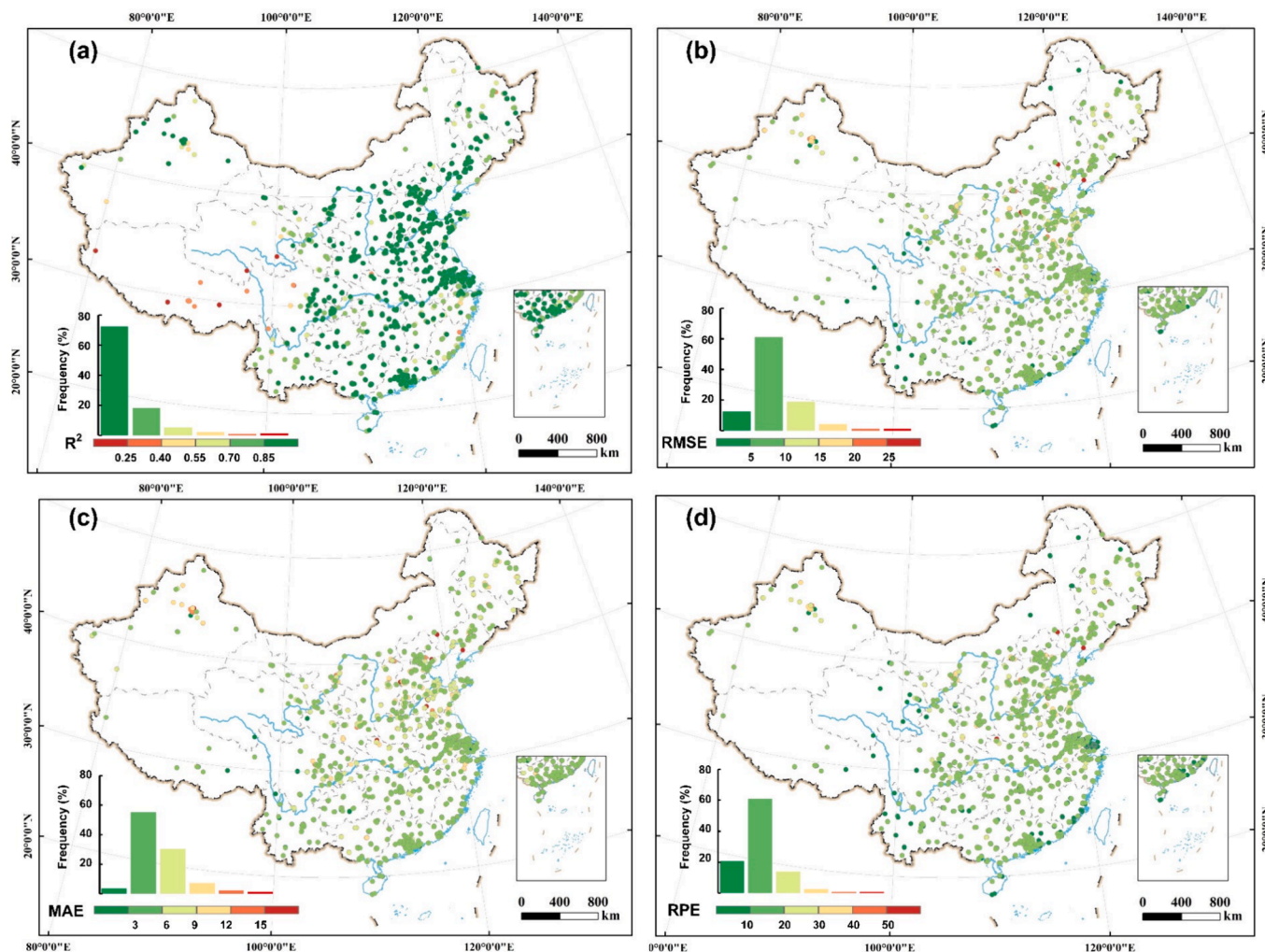
Fig. 4. Spatial distribution of the sample-based CV results of ATSMNN model. (a) The correlation of determination (R^2); (b) RMSE (µg m^{-3}), (c) MAE (µg m^{-3}) and RPE (%) in 2021 across mainland China.
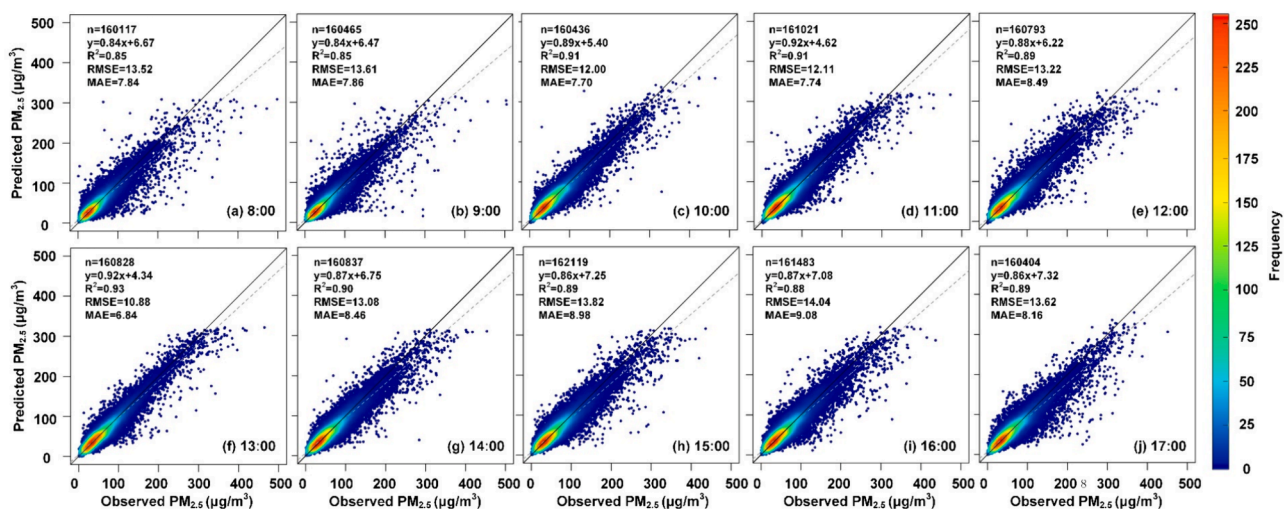

Fig. 5. Density scatterplots of the results of sample-based CV results at different hours (8:00–17:00 UTC + 8) across China.

R^2 less than 0.5. The limited numbers and sparse distributions of sites in southwest China result in low R^2 performance in this region (Fig. 4a). Despite this, the model demonstrated relatively low RMSE and MAE with values under 6 µg m^{-3} (Fig. 4b-c). This indicates that due to the

relatively low PM$_{2.5}$ concentrations in southwest China, the estimation biases had a minimal impact on the final estimation results.

Fig. 5 shows the sample-based CV results of our model across various times of the day to validate the model robustness throughout different
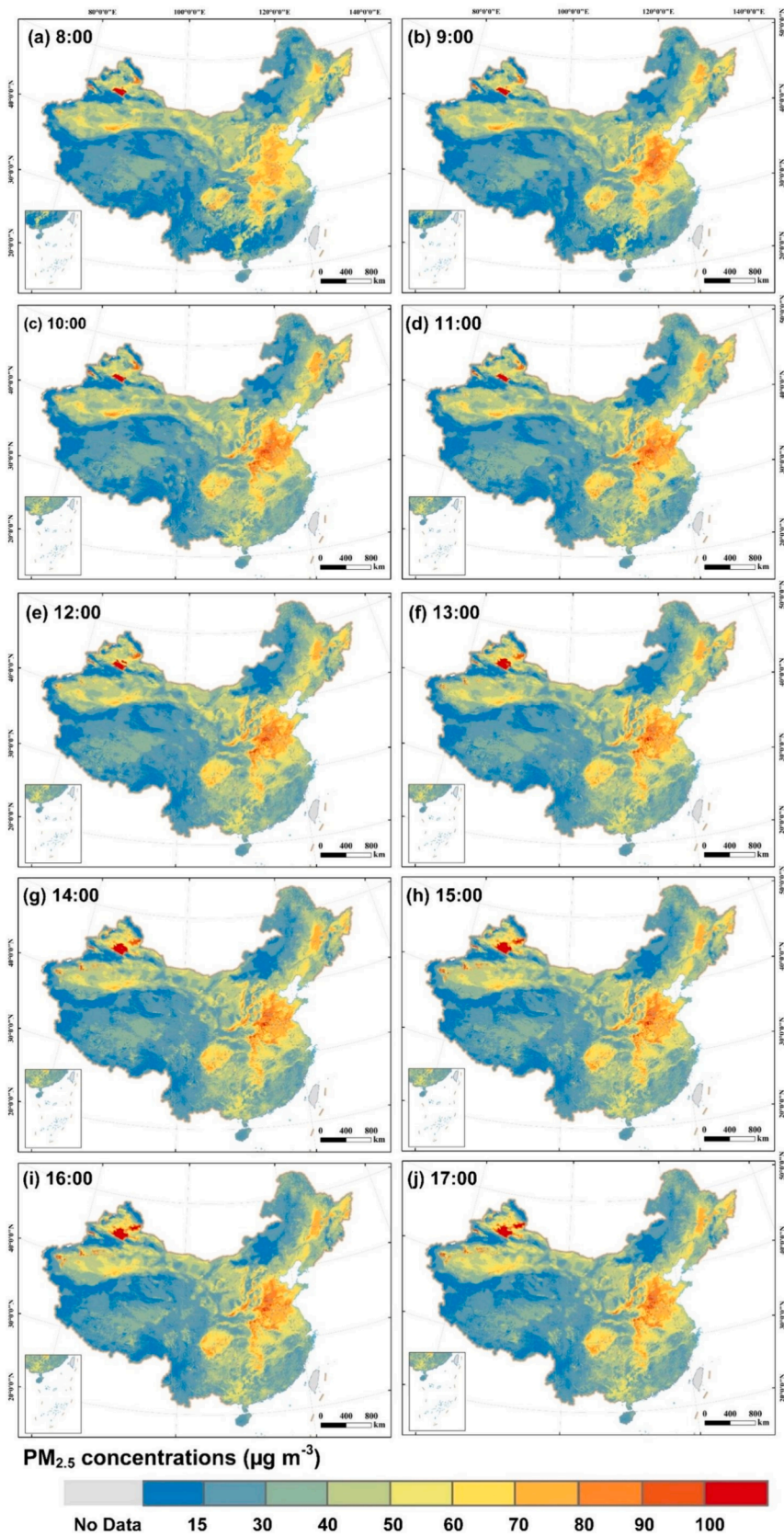
Fig. 6. Spatial distribution of hourly PM$_{2.5}$ concentrations (1 km) on January 16th, 2021 across China.
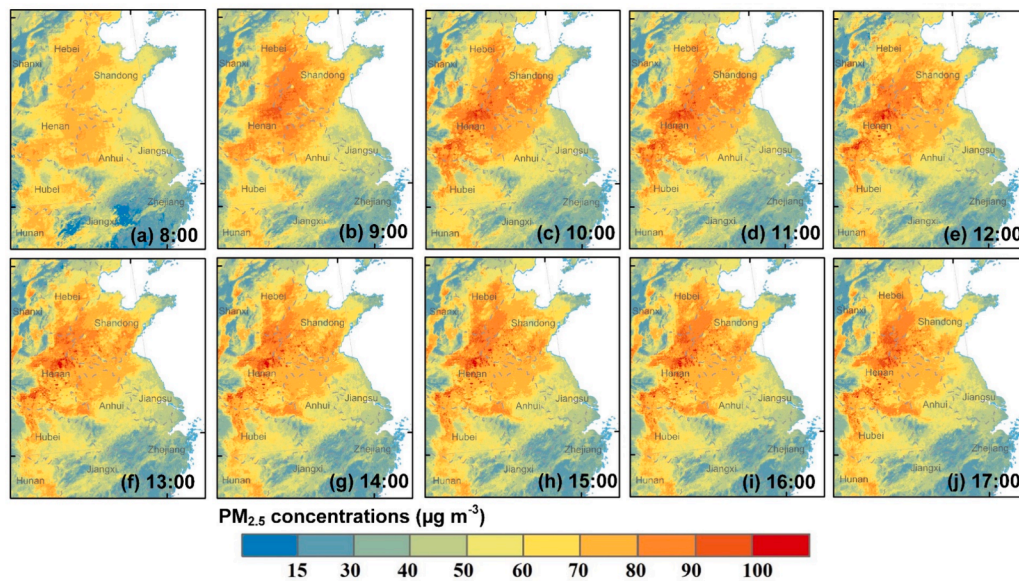
Fig. 7. Spatial distribution of hourly PM$_{2.5}$ concentrations on January 16th, 2021 across east costal China.

hours. The model accuracy was relatively stable at different hours, with the R^2 values and MAE ranging from 0.85 to 0.93, and 6.84 to 9.08 µg m^{-3}, respectively. The estimation results at adjacent hours were more similar because of the similar spatiotemporal variation in PM$_{2.5}$. The highest accuracy occurred at 3:00 UTC + 8 (Fig. 5 (f)) with an R^2 of 0.93 and the fitted line closest to 1:1 (slope of 0.92). The least accuracy occurred at 17:00 UTC + 8 and 09:00 UTC + 8 (Fig. 5 (j) and (b)) mainly caused by the poor AOD quality (Wang et al., 2022b). Overall, the model provides stable and robust performance over geographical locations and different times and is suitable for estimating hourly PM$_{2.5}$ concentrations across large regions like mainland China.

3.2. Spatiotemporal pattern of hourly PM$_{2.5}$

By adopting the ASTNN model, hourly PM$_{2.5}$ were estimated in mainland China at 1 km resolution. Fig. 6 presents an example of hourly PM$_{2.5}$ concentrations observed on January 16th, 2021. The results recorded the highest and lowest PM$_{2.5}$ concentrations at 16:00 and 8:00 (UTC + 8), respectively, with a difference of 3.70 µg m^{-3}. Fig. 7 presents a detailed spatial distribution of elevated pollution levels in the east coast of China across various hours. The zoomed-in results exhibit a prominent trend that PM$_{2.5}$ concentrations tend to rise from 8:00 to noontime and remain relatively stable in the afternoon.

The distributions of satellite-derived annual mean hourly PM$_{2.5}$ concentrations are presented in Fig. S4. The annual average PM$_{2.5}$ concentration across China in 2021 was 24.69 ± 10.43 µg m^{-3}. Spatially, northwest China (i.e., Taklamakan Desert), northern China, Fenwei plain and Sichuan basin were the most polluted regions. Temporally, the lowest and highest PM$_{2.5}$ concentrations among the daytime occurred at 10:00 and 16:00 (UTC + 8).

3.3. Hourly PM$_{2.5}$ monitoring during dust events

Satellite products with high temporal revisit frequency are essential for tracking the emission and transfer of particulate matter, especially during extreme atmospheric pollution events (Vu et al., 2022). In this research, we estimated the hourly variation in PM$_{2.5}$ during the dust events that occurred frequently in northern China in March 2021. Fig. 8 presents the spatiotemporal variation of the PM$_{2.5}$ concentrations on March 28, 2021 during a dust event which occurred in the BTH regions. Influenced by the directions of winds under the Siberian high-pressure system (Fig. 8d), the PM$_{2.5}$ showed a north-west to south-east transfer

path through time. We further tracked the hourly variation in PM$_{2.5}$ in the BTH region at a fine resolution. The high PM$_{2.5}$ levels with concentrations over 150 µg m^{-3} first occurred in central Beijing, southwestern Hebei and the border of Inner Mongolia and Liaoning province at 8:00 (UTC + 8). Compared with the true-color image from MODIS at noontime on March 28, our estimate results accurately captured the distribution of the dust plume. Moreover, the higher PM$_{2.5}$ pollution mainly affected urban areas such as Beijing and south Hebei province according to the urban built-up boundaries (Fig. 8l), and showed a transport pathway from central metropolitan areas radiating outward. These results demonstrated the impacts of land cover on the deposition of particulates. Specifically, the particulates were more likely to deposit in urban areas without vegetation and spread around.

4. Discussion

4.1. Comparison with previous studies

To further examine the effectiveness of our model, it was compared with previous studies which estimated hourly PM$_{2.5}$ concentrations across China (Table 4). Comparison within the same year would be more practical significance, however, no similar studies have been conducted on hourly scale PM$_{2.5}$ estimation in 2021 as far as we are concerned. The statistical model (i.e., IGTWR) was the least accurate (R^2 and RMSE of 0.78 and 21.10, respectively). The accuracy of PM$_{2.5}$ estimation was enhanced by the introduction of ML (e.g., RF and XGBoost), with RMSEs ranging from 11.02 µg m^{-3} to 19.60 µg m^{-3}. The accuracy of two improved ML models (i.e., STWC-DNN and STLG) adopting spatiotemporal weighting surpassed the general ML models with RMSE smaller than 14 µg m^{-3}. Our model simultaneously integrated the local spatiotemporal information and global features of PM$_{2.5}$ into deep learning models, and achieved promising accuracy.

4.2. Effectiveness of the reconstructed AOD

In this research, to achieve full-coverage hourly PM$_{2.5}$ estimation we reconstructed the hourly AOD at a spatial resolution of 1 km. To validate the effectiveness of using the reconstructed AOD for PM$_{2.5}$ estimation, we conducted a comparative experiment using both the reconstructed AOD data and without it. As shown in Fig. 9, the exclusion of AOD data led to a notable increase in both the RMSE and MAE, although the overall $R2$ remained at 0.88. Moreover, the scatterplot in Fig. 9 indicates
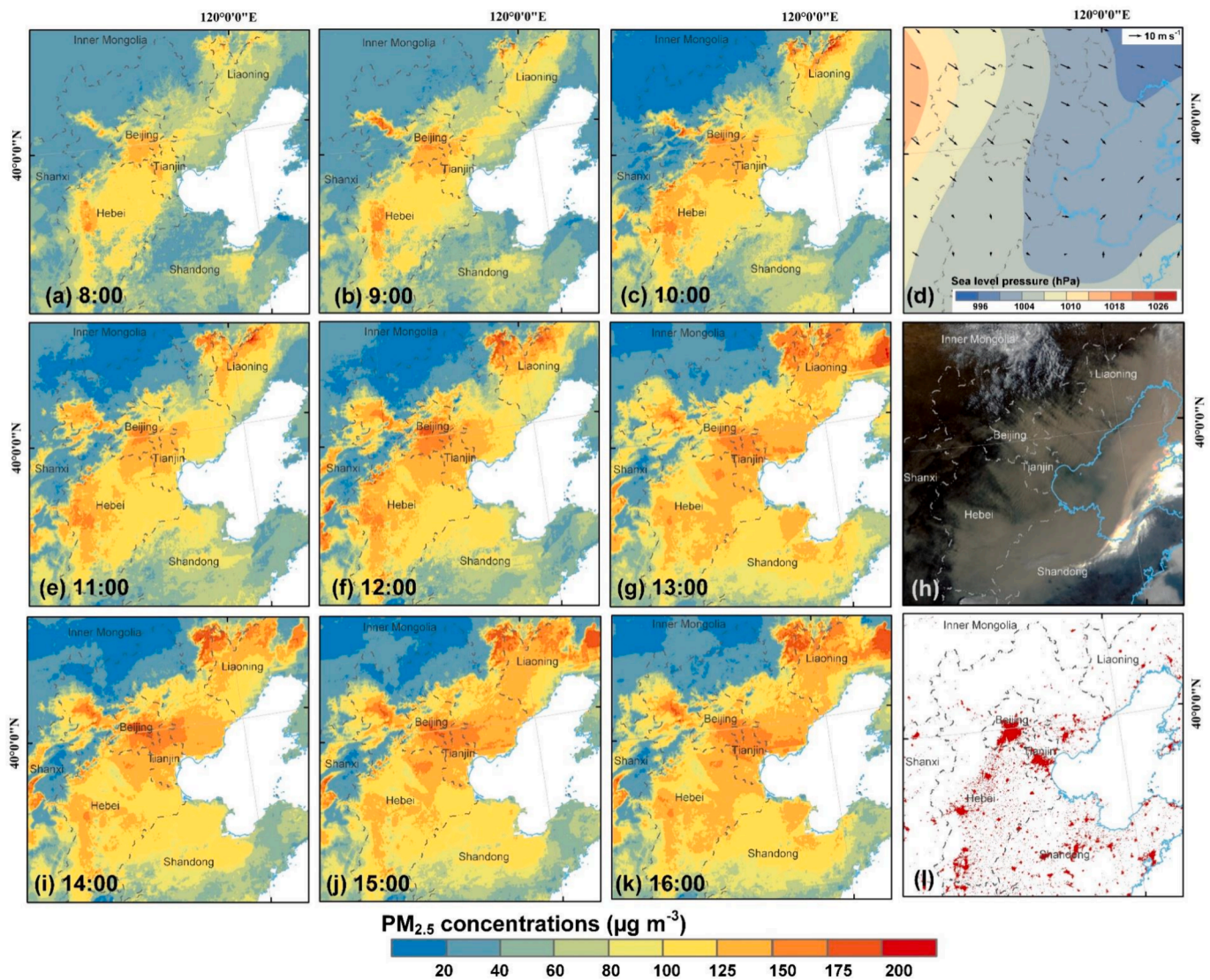
Fig. 8. Spatial distribution of hourly estimates of PM$_{2.5}$ concentrations on March 28, 2021, during a haze event in the BTH region (a-c, e-g, and i-k). (d) Sea level pressures (colored fill) and vector winds at 10 m above ground level (black arrows) on March 28, 2021, calculated based on the ERA5 hourly data. (h) True-color image from MODIS during the satellite transit on March 28. (l) The corresponding urban area distribution.

Table 4
Models' performance of previous studies for hourly PM$_{2.5}$ estimation across China.

Model	Time range	Spatial resolution	Sample-based CV			Reference
			R^2	RMSE	MAE	
ASTMNN	2021	1 km	0.94	9.82	6.11	our study
Wavelet-CatBoost	2020	0.01°	0.92	9.62	–	Ding et al. (2024)
STWC-DNN	2017	5 km	0.92	12.70	8.36	Wang et al. (2022b)
STLG	2018	5 km	0.85	13.62	8.49	Wei et al. (2021)
Two-stage RF (gap-filling RF)	2018.03–2019.02	1 km	0.85	11.02	6.73	Jiang et al. (2021)
IGTWR	2017	5 km	0.78	21.10	–	Xue et al. (2020)
Stacked model	2016	0.05°	0.85	17.30	10.50	Chen et al. (2019)
RF			0.82	19.60	12.20	
XGBoost			0.84	18.10	11.4	
AdaBoost			0.84	18.30	10.70	

that the removal of AOD data results in a significantly 'looser' and more biased distribution of points compared to Fig. 9a. This effect impacts the prediction of higher PM$_{2.5}$ values, where a significant underestimation is observed. These results indicate the necessity of incorporating reconstructed AOD for accurate PM$_{2.5}$ modelling and estimation.

Due to the lack of AHI AOD coverage in parts of Xinjiang region, the impact of the reconstructed AOD on PM$_{2.5}$ estimation in this region requires further validation. Therefore, we separately extracted PM$_{2.5}$ monitoring sites in Xinjiang for validation based on the reconstructed AOD (Fig. S7). As shown in Fig. S7, the model performs well both at the
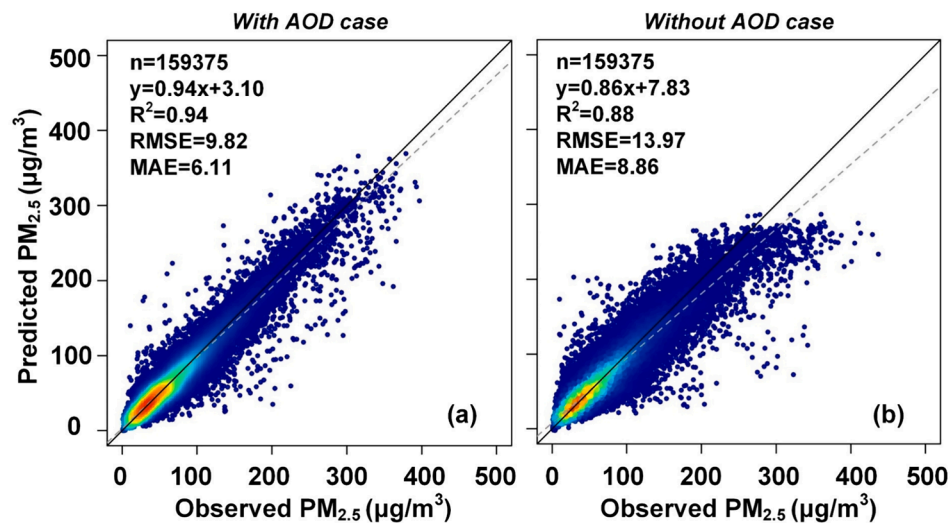
Fig. 9. Scatterplots of sample-based CV results for the cases (left) with and (right) without reconstructed AOD data. The linear regression fitting line and the 1:1 line are shown by the dashed and solid lines, respectively. The density of points is represented by the color of the points, where blue is low and red is high denisty. The number of samples n, coefficient of determination R^2, RMSE (μg m^{-3}) and MAE (μg m^{-3}) are provided. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

site scale and in the overall validation results, implying that our reconstructed AOD can characterize the relationship between PM$_{2.5}$ and AOD reliably.

4.3. Limitations and future improvements

Two main limitations of this study can be identified. Firstly, we utilized two satellite AOD products along with multiple related covariates to obtain full-coverage 1 km hourly AOD. Despite using 1 km MODIS AOD and some fine-resolution variables (e.g., NDVI, LUCC and DEM) to correct for these inconsistencies (Fig S3 displays the validation results), the coarse spatial resolution and potential errors of AHI AOD and several covariates may still degrade the spatial quality of the reconstructed AOD. Additionally, the reconstructed model could potentially propagate these errors, thereby, affecting the overall accuracy of the AOD estimates. Consequently, employing more suitable gap-filling methods to enhance the data quality and spatial resolution of the AOD products would be beneficial for fine-resolution PM$_{2.5}$ estimation. Secondly, due to the large sample size of the hourly data, our experiments were limited to the year 2021. This constraint hinders the validation of model performance across different years and reduces comparability with previous studies. In future research, we intend to extend our time-series and utilize the ASTMNN model to explore long-term temporal variation in PM$_{2.5}$.

5. Conclusions

An estimation of PM$_{2.5}$ with fine temporal and spatial resolution is required to help reduce environmental pollution and mitigate climate change effects. The key to effective PM$_{2.5}$ estimation lies in exploring the spatiotemporal characteristics of PM$_{2.5}$ and its associated physical variables. In this research, we proposed a novel framework based on deep learning for hourly PM$_{2.5}$ estimation. Combining deep learning and multiscale spatiotemporal characteristics in the proposed ASTMNN model was found to be effective for hourly PM$_{2.5}$ estimation. Based on ground-based PM$_{2.5}$ observations, gap-filled hourly AOD products and other multi-temporal remote sensing products, we generated seamless PM$_{2.5}$ concentration products with fine spatial resolution (1 km) and temporal resolution (hourly). Moreover, this research draw attention to the hourly PM$_{2.5}$ change during dust events, revealing a close association between the dust events and meteorology conditions and land cover. Our findings provide valuable insights for tracking continuous variation in particulate pollution using geostationary satellites and have the potential of supporting epidemiologic studies related to acute short-term exposure.

CRediT authorship contribution statement

Zhige Wang: Writing – original draft, Methodology, Conceptualization. **Ce Zhang:** Writing – review & editing, Investigation. **Su Ye:** Writing – review & editing, Supervision. **Rui Lu:** Methodology. **Yulin Shangguan:** Data curation. **Tingyuan Zhou:** Visualization. **Peter M. Atkinson:** Writing – review & editing. **Zhou Shi:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.104145.

References

Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, in: 2017 International Conference on Engineering and Technology (ICET). Presented at the 2017 International Conference on Engineering and Technology (ICET), pp. 1–6. 10.1109/ICEngTechnol.2017.8308186.

Bai, Y., Wu, L., Qin, K., Zhang, Y., Shen, Y., Zhou, Y., 2016. A Geographically and Temporally Weighted Regression Model for Ground-Level $PM_{2.5}$ Estimation from Satellite-Derived 500 m Resolution AOD. Remote Sens. 8. 10.3390/rs8030262.

Burnett, R., et al., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc. Natl Acad. Sci. USA 115, 9592–9597. https://doi.org/10.1073/pnas.1803222115.

Chen, Z., Xie, X., Cai, J., Chen, D., Gao, B., He, B., Cheng, N., Xu, B., 2018. Understanding meteorological influences on $PM_{2.5}$ concentrations across China: a temporal and spatial perspective. Atmos. Chem. Phys. 18, 5343–5358. https://doi.org/10.5194/acp-18-5343-2018.

Chen, J., Yin, J., Zang, L., Zhang, T., Zhao, M., 2019. Stacking machine learning model for estimating hourly $PM_{2.5}$ in China based on Himawari-8 aerosol optical depth data. Sci. Total Environ. 697, 134021 https://doi.org/10.1016/j.scitotenv.2019.134021.

Della Ceca, L.S., 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. Photogram. Rem. Sens. 145, 250–267. https://doi.org/10.1016/j.isprsjprs.2018.08.016.

Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., Schwartz, J., 2016. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. Environ. Sci. Technol. 50, 4712–4721. https://doi.org/10.1021/acs.est.5b06121.

Ding, Y., Li, S., Xing, J., Li, X., Ma, X., Song, G., Teng, M., Yang, J., Dong, J., Meng, S., 2024. Retrieving hourly seamless $PM_{2.5}$ concentration across China with physically informed spatiotemporal connection. Remote Sens. Environ. 301, 113901 https://doi.org/10.1016/j.rse.2023.113901.

Eck, T.F., Holben, B.N., Reid, J.S., Dubovik, O., Smirnov, A., O'Neill, N.T., Slutsker, I., Kinne, S., 1999. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. J. Geophys. Res.-Atmos. 104, 31333–31349. https://doi.org/10.1029/1999JD900923.

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. https://doi.org/10.1016/j.atmosenv.2004.01.039.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

Giles, D.M., Sinyuk, A., Sorokin, M.G., Schafer, J.S., Smirnov, A., Slutsker, I., Eck, T.F., Holben, B.N., Lewis, J.R., Campbell, J.R., Welton, E.J., Korkin, S.V., Lyapustin, A.I., 2019. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmos. Measure. Tech. 12, 169–209. https://doi.org/10.5194/amt-12-169-2019.

Graves, A., Fernández, S., & Schmidhuber, J. 2005. Bidirectional LSTM networks for improved phoneme classification and recognition. In Artificial Neural Networks: Formal Models and Their Applications–ICANN 2005: 15th International Conference, Warsaw, Poland, September 11-15, 2005. Proceedings, Part II 15 (pp. 799-804). Springer Berlin Heidelberg.

Graves, A., Jaitly, N., Mohamed, A.R., 2013. Hybrid speech recognition with deep bidirectional LSTM. In: 2013 IEEE Workshop on Automatic Speech Recognition and Understanding. IEEE, pp. 273–278. https://doi.org/10.1109/ASRU.2013.6707742.

Gupta, P., Christopher, S.A., 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. A neural network approach. J. Geophys. Res.-Atmos. 114 https://doi.org/10.1029/2008JD011497.

Gupta, P., Christopher, S.A., Wang, J., Gehrig, R., Lee, Y., Kumar, N., 2006. Satellite remote sensing of particulate matter and air quality assessment over global cities. Atmos. Environ. 40, 5880–5892. https://doi.org/10.1016/j.atmosenv.2006.03.016.

Harris, P., Fotheringham, A.S., Juggins, S., 2010. Robust geographically weighted regression: a technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes. Ann. Assoc. Am. Geogr. 100, 286–306. https://doi.org/10.1080/00045600903550378.

He, Q., Huang, B., 2018. Satellite-based mapping of daily high-resolution ground $PM_{2.5}$ in China via space-time regression modeling. Remote Sens. Environ. 206, 72–83. https://doi.org/10.1016/j.rse.2017.12.018.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J.J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., Suttie, M., 2019. The CAMS reanalysis of atmospheric composition. Atmos. Chem. Phys. 19, 3515–3556. https://doi.org/10.5194/acp-19-3515-2019.

Jiang, T., Chen, B., Nie, Z., Ren, Z., Xu, B., Tang, S., 2021. Estimation of hourly full-coverage $PM_{2.5}$ concentrations at 1-km resolution in China using a two-stage random forest model. Atmos. Res. 248, 105146 https://doi.org/10.1016/j.atmosres.2020.105146.

Khotanzad, A., Chung, C., 1998. Application of multi-layer perceptron neural networks to vision problems. Neural Comput. Appl. 7, 249–259. https://doi.org/10.1007/BF01414886.

Kikuchi, M., Murakami, H., Suzuki, K., Nagao, T.M., Higurashi, A., 2018. Improved hourly estimates of aerosol optical thickness using spatiotemporal variability derived from himawari-8 geostationary satellite. IEEE Trans. Geosci. Remote Sens. 56, 3442–3455. https://doi.org/10.1109/TGRS.2018.2800060.

Lee, H.J., Liu, Y., Coull, B.A., Schwartz, J., Koutrakis, P., 2011. A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations. Atmos. Chem. Phys. 11, 7991–8002. https://doi.org/10.5194/acp-11-7991-2011.

Lei, C., Xu, X., Ma, Y., Jin, S., Liu, B., Gong, W., 2022. Full coverage estimation of the PM concentration across China based on an adaptive spatiotemporal approach. IEEE Trans. Geosci. Remote Sens. 60, 1–14. https://doi.org/10.1109/TGRS.2022.3213797.

Letu, H., Yang, K., Nakajima, T.Y., Ishimoto, H., Nagao, T.M., Riedi, J., Baran, A.J., Ma, R., Wang, T., Shang, H., Khatri, P., Chen, L., Shi, C., Shi, J., 2020. High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite. Remote Sens. Environ. 239, 111583 https://doi.org/10.1016/j.rse.2019.111583.

Li, T., Shen, H., Yuan, Q., Zhang, L., 2020. Geographically and temporally weighted neural networks for satellite-based mapping of ground-level $PM_{2.5}$. ISPRS J. Photogram. Rem. Sens. 167, 178–188. https://doi.org/10.1016/j.isprsjprs.2020.06.019.

Liu, Y., Paciorek, C.J., Koutrakis, P., 2009. Estimating regional spatial and temporal variability of $PM_{2.5}$ concentrations using satellite data, meteorology, and land use information. Environ. Health Perspect. 117, 886–892. https://doi.org/10.1289/ehp.0800123.

Luong, M.-T., Pham, H., Manning, C.D., 2015. Effective approaches to attention-based neural machine translation. CoRR abs/1508.04025.

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R., Reid, J.S., 2011. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. J. Geophys. Res.-Atmos. 116. 10.1029/2010JD014986.

Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS collection 6 MAIAC algorithm. Atmos. Measure. Tech. 11, 5741–5765. https://doi.org/10.5194/amt-11-5741-2018.

Ma, Z., Dey, S., Christopher, S., Liu, R., Bi, J., Balyan, P., Liu, Y., 2022. A review of statistical methods used for developing large-scale and long-term $PM_{2.5}$ models from satellite data. Remote Sens. Environ. 269, 112827. 10.1016/j.rse.2021.112827.

Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating Ground-Level $PM_{2.5}$ in China using satellite remote sensing. Environ. Sci. Technol. 48, 7436–7444. https://doi.org/10.1021/es5009399.

Mao, F., Hong, J., Min, Q., Gong, W., Zang, L., Yin, J., 2021. Estimating hourly full-coverage $PM_{2.5}$ over China based on TOA reflectance data from the Fengyun-4A satellite. Environ. Pollut. 270, 116119. 10.1016/j.envpol.2020.116119.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., 2015. Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. Geosci. Model Develp. 8, 1339–1356. https://doi.org/10.5194/gmd-8-1339-2015.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Murray, C.J.L., et al., 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396, 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2.

Paciorek, C.J., Liu, Y., Moreno-Macias, H., Kondragunta, S., 2008. Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level $PM_{2.5}$. Environ. Sci. Technol. 42, 5800–5806. https://doi.org/10.1021/es703181j.

Páez, A., Farber, S., Wheeler, D., 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. Environ. Plan. A 43 (12), 2992–3010. https://doi.org/10.1068/a44111.

Song, W., Jia, H., Huang, J., Zhang, Y., 2014. A satellite-based geographically weighted regression model for regional $PM_{2.5}$ estimation over the Pearl River Delta region in China. Remote Sens. Environ. 154, 1–7. https://doi.org/10.1016/j.rse.2014.08.008.

Sorek-Hamer, M., Strawa, A.W., Chatfield, R.B., Esswein, R., Cohen, A., Broday, D.M., 2013. Improved retrieval of $PM_{2.5}$ from satellite data products using non-linear methods. Environ. Pollut. 182, 417–423. https://doi.org/10.1016/j.envpol.2013.08.002.

van Donkelaar, A., Martin, R.V., Park, R.J., 2006. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. J. Geophys. Res.-Atmos. 111. 10.1029/2005JD006996.

Vu, B., Bi, J., Wang, W., Huff, A., Kondragunta, S., Liu, Y., 2022. Application of geostationary satellite and high-resolution meteorology data in estimating hourly $PM_{2.5}$ levels during the Camp Fire episode in California. Remote Sens. Environ. 271, 112890. 10.1016/j.rse.2022.112890.

Wang, Z., Zhou, Y., Zhao, R., Wang, N., Biswas, A., Shi, Z., 2021. High-resolution prediction of the spatial distribution of $PM_{2.5}$ concentrations in China using a long short-term memory model. J. Clean. Prod. 297, 126493. 10.1016/j.jclepro.2021.126493.

Wang, Z., Hu, B., Zhang, C., Atkinson, P.M., Wang, Z., Xu, K., Chang, J., Fang, X., Jiang, Y., Shi, Z., 2022a. How the Air Clean Plan and carbon mitigation measures co-benefited China in $PM_{2.5}$ reduction and health from 2014 to 2020. Environ. Int. 169, 107510. 10.1016/j.envint.2022.107510.

Wang, Z., Li, R., Chen, Z., Yao, Q., Gao, B., Xu, M., Yang, L., Li, M., Zhou, C., 2022b. The estimation of hourly $PM_{2.5}$ concentrations across China based on a Spatial and Temporal Weighted Continuous Deep Neural Network (STWC-DNN). ISPRS J. Photogram. Rem. Sens. 190, 38–55. https://doi.org/10.1016/j.isprsjprs.2022.05.011.

Wang, Z., Hu, B., Huang, B., Ma, Z., Biswas, A., Jiang, Y., Shi, Z., 2022c. Predicting annual $PM_{2.5}$ in mainland China from 2014 to 2020 using multi temporal satellite product: An improved deep learning approach with spatial generalization ability.

ISPRS J. Photogram. Rem. Sens. 187, 141–158. https://doi.org/10.1016/j.isprsjprs.2022.03.002.

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019. Estimating 1-km-resolution PM$_{2.5}$ concentrations across China using the space-time random forest approach. Remote Sens. Environ. 231, 111221. 10.1016/j.rse.2019.111221.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., Song, Y., 2020. Improved 1km resolution PM$_{2.5}$ estimates across China using enhanced space-time extremely randomized trees. Atmos. Chem. Phys. 20, 3273–3289. https://doi.org/10.5194/acp-20-3273-2020.

Wei, J., Li, Z., Pinker, R.T., Wang, J., Sun, L., Xue, W., Li, R., Cribb, M., 2021. Himawari-8-derived diurnal variations in ground-level PM$_{2.5}$ pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM). Atmos. Chem. Phys. 21, 7863–7880. https://doi.org/10.5194/acp-21-7863-2021.

Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: Convolutional Block Attention Module. CoRR abs/1807.06521.

Xiao, Q., Wang, Y., Chang, H.H., Meng, X., Geng, G., Lyapustin, A., Liu, Y., 2017. Full-coverage high-resolution daily PM$_{2.5}$ estimation using MAIAC AOD in the Yangtze River Delta of China. Remote Sens. Environ. 199, 437–446. https://doi.org/10.1016/j.rse.2017.07.023.

Xu, W., Wang, W., Wang, N., Chen, B., 2022. A new algorithm for Himawari-8 aerosol optical depth retrieval by integrating regional PM$_{2.5}$ concentrations. IEEE Trans. Geosci. Remote Sens. 60, 1–11. https://doi.org/10.1109/TGRS.2022.3155503.

Xue, Y., Li, Y., Guang, J., Tugui, A., She, L., Qin, K., Fan, C., Che, Y., Xie, Y., Wen, Y., Wang, Z., 2020. Hourly PM$_{2.5}$ Estimation over Central and Eastern China Based on Himawari-8 Data. Remote Sens. 12. 10.3390/rs12050855.

Yang, N., Shi, H., Tang, H., Yang, X., 2022. Geographical and temporal encoding for improving the estimation of PM$_{2.5}$ concentrations in China using end-to-end gradient boosting. Remote Sens. Environ. 269, 112828. 10.1016/j.rse.2021.112828.

Zagoruyko, S., Komodakis, N., 2016. Paying more attention to attention: improving the performance of convolutional neural networks via attention transfer. CoRR abs/1612.03928.

Zhang, Q., He, K., Huo, H., 2012. Cleaning China's air. Nature 484, 161–162. https://doi.org/10.1038/484161a.

Zhang, Z., Wu, W., Fan, M., Tao, M., Wei, J., Jin, J., Tan, Y., Wang, Q., 2019b. Validation of Himawari-8 aerosol optical depth retrievals over China. Atmos. Environ. 199, 32–44. https://doi.org/10.1016/j.atmosenv.2018.11.024.

Zhang, T., Zang, L., Wan, Y., Wang, W., Zhang, Y., 2019a. Ground-level PM$_{2.5}$ estimation over urban agglomerations in China with high spatiotemporal resolution based on Himawari-8. Sci. Total Environ. 676, 535–544. https://doi.org/10.1016/j.scitotenv.2019.04.299.