ONE DAY AHEAD PREDICTION OF PM2.5 SPATIAL DISTRIBUTION USING MODIS 3 KM AOD AND SPATIOTEMPORAL MODEL OVER BEIJING-TIANJIN-HEBEI, CHINA

Xinpeng Wang [1,3], Wanzeng Liu [1,3], Wenbin Sun [2,3], Yunlu Peng [1,3 *], Ye Zhang [1,3], Xi Zhai [1,3], Tingting Zhao [1,3], Ran Li [1,3]

[1]National Geomatics Center of China, Beijing 100830, China – (wangxinpeng, pengyunlu, zhangye, zhaixi, zhaotingting, liran)@ngcc.cn; luwnzg@163.com.
[2]College of Geoscience and Surveying Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China – swb1996@126.com
[3]Key Laboratory of Spatio-temporal Information and Intelligent Services (LSIIS), MNR

Commission III, WG III/8

KEY WORDS: MODIS AOD, PM2.5, PM2.5 spatiotemporal distribution, Spatiotemporal autoregressive model, Machine learning model.

ABSTRACT:

Accurate prediction of PM2.5 concentration is the premise and guarantees to effectively control PM2.5 concentration and avoid the adverse effects of high PM2.5 concentration on human health. However, given the existing statistical models can only predict the pollutant concentration at the monitoring sites, the spatial distribution of PM2.5 concentration cannot be predicted, which greatly limits the application of the model in PM2.5 concentration prediction. This study combined the PM2.5 spatial distribution data predicted using the Moderate Resolution Imaging Spectroradiometer 3 km aerosol optical depth (AOD) and meteorological factors into a spatiotemporal autoregressive (STAR) model to predict the regional PM2.5 concentration and quantify the short-term spatial distribution change of PM2.5 in Beijing–Tianjin–Hebei region (JingJinJi) one day in advance. Five simulation functions were used to simulate the STAR model, and the 2014 data of JingJinJi were used to verify its accuracy. Results showed that the STAR model had the best prediction performance when gradient boosting decision tree was used as the simulation function compared with other simulation functions. The coefficient of determination (R^2), root mean square prediction error (RMSE), index of agreement (IA), and mean absolute error (MAE) of the STAR model were 0.85, 27.08 $\mu g/m^3$, 0.96, and 20 $\mu g/m^3$, respectively. The spatial distribution prediction results of PM2.5 showed that the +1-day PM2.5 spatial distribution prediction results were in good agreement with the PM2.5 spatial distribution results predicted by AOD to provide accurate spatiotemporal distribution data for reducing air pollution and air pollution early warning.

1. INTRODUCTION

Air quality refers to the chemical state of the atmosphere at a specific time and place and reflects the degree of air pollution. The concentration of pollutants in the atmosphere will have adverse effects on human health, such as damage to immune and nervous systems and premature death, when it exceeds a certain boundary (Zhang et al., 2012). Among the many atmospheric pollutants, PM2.5 is the most harmful to human health because it contains toxic and harmful substances and can directly enter the alveoli (Belleudi et al., 2010; Crouse et al., 2012; Dominici et al., 2006; Pope III et al., 2002). With the rapid economic development and intensified anthropogenic emissions annually, China has suffered from serious air pollution, especially in northern China (Wang et al., 2019; Zheng et al., 2016). To effectively control PM2.5 concentration and avoid its adverse effects on human health, the government should conduct precautionary measures in advance (such as closing the main emission sources and vehicle restrictions) to reduce air pollution and issue air pollution warning for providing guidance for public travel (Zhang et al., 2012). Therefore, real-time air quality information should be obtained and the temporal and spatial change trends of PM2.5 concentration should be predicted (Chen et al., 2013).

Prediction methods of air pollution concentration can be mainly divided into two types, namely, statistical and deterministic models. Considering that meteorological and air pollutant concentration variables are statistically related, statistical models use different functions to simulate the relationship of measured pollutant variables and various selected predictors for predicting pollutant concentration (Hx, et al., 2021; Pak, et al., 2020; Wen et al., 2019; Cobourn 2007; Elangasinghe et al., 2014; Kurt and Oktay 2010; Nieto et al., 2013; Li et al., 2016; Sánchez et al., 2013; Qi et al., 2019; Soh et al., 2018). Statistical models have better prediction accuracy. However, the existing statistical models can only predict the pollutant concentration at the monitoring site and cannot be extended to other regions with different meteorological conditions or without monitoring sites (Cortina Januchs et al., 2015; Elangasinghe et al., 2014; Hooyberghs et al., 2005), that is, the existing statistical models cannot predict the regional PM2.5 concentration.

The deterministic model simulates pollutant discharge, diffusion, and disappearance in a model-driven manner (Li et al., 2016), thus enabling the prediction of pollutant concentration in areas without monitoring sites (Cortina Januchs et al., 2015). During simulation, pollutant discharge volume, chemical composition of pollution gas, and physical processes of the atmosphere should be fully understood, however, these key knowledge are often insufficient (Hrust et al., 2009). In addition, the model is

* Corresponding author

https://doi.org/10.5194/isprs-annals-V-3-2022-303-2022 |

simplified during prediction because of its incomplete theoretical basis and calculation complexity, thereby leading to low prediction accuracy (Feng et al., 2015).

Therefore, regional PM2.5 spatiotemporal distribution trend prediction with high precision should be explored to reduce the impacts of PM2.5 concentration on public health and provide guidance for public travel and government pollution prevention. The formation of PM2.5 is a dynamic system under the influence of meteorological and natural factors. The PM2.5 concentration at a certain location is related to its own and nearby PM2.5 concentration and meteorological factors. Therefore, PM2.5 has a strong spatiotemporal autocorrelation. A spatiotemporal autoregressive (STAR) model considers the temporal evolution of adjacent spatial grid points. In the STAR model, each target pixel value is simulated as a weighted sum of neighboring available raster pixel values (Zhang et al., 2009), thus, the model is frequently used for remote sensing images prediction (Cheng et al., 2017; Crespo et al., 2007; Das and Ghosh 2016). This study introduced the daily complete PM2.5 spatial distribution (raster image) data obtained by Moderate Resolution Imaging Spectroradiometer (MODIS) AOD and meteorological factors into the STAR model to predict the one day ahead spatiotemporal distribution trend of PM2.5 and provide guidance for public travel and government pollution prevention. The rest of this paper is organized as follows. Section 2 introduces the study area, input data, and model structure. Section 3 presents the model's validation results, PM2.5 spatiotemporal distribution, and feasibility analysis. Section 4 provides the conclusions.

2. MATERIALS AND METHOD

2.1 Study area

The study area includes Beijing, Tianjin, and Hebei (JingJinJi) (Figure 1(a)). Excessive emissions, unfavorable terrain and meteorological conditions make JingJinJi a typical heavily polluted area in China. Therefore, this area is selected as the research area of this paper.

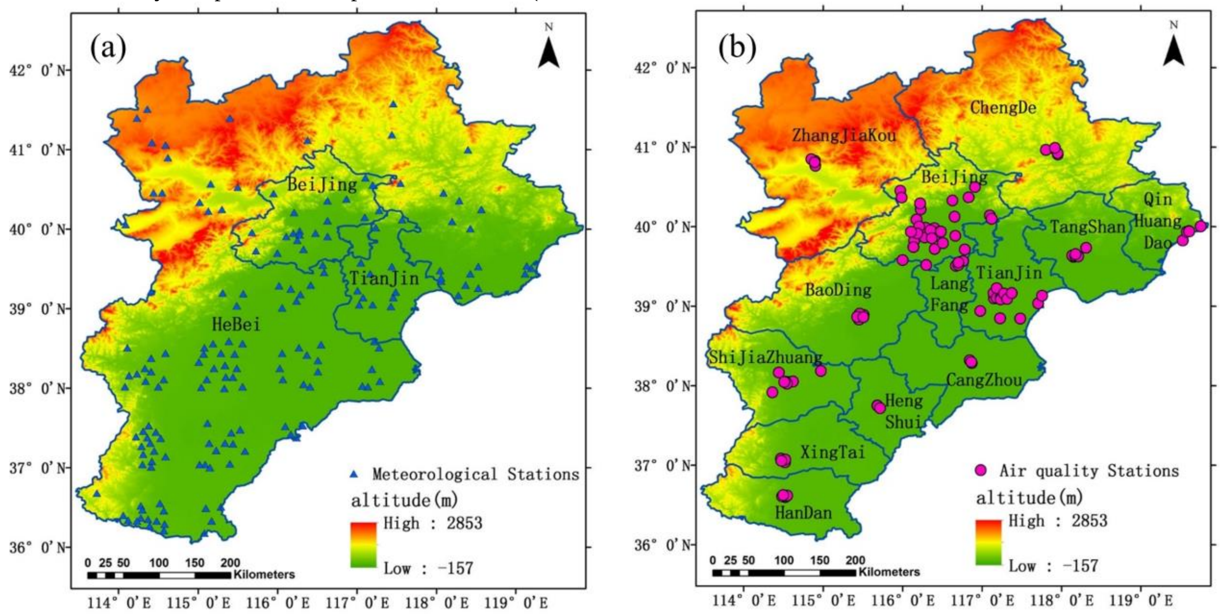

Figure 1. Schematic diagram of study area.

2.2 MODIS AOD Data

The Dark Target algorithm Collection 6 MODIS Aqua AOD data of 2014 were downloaded from the NASA official website (Yang et al., 2019), and the AOD resolution used in this study was 3 km.

2.3 Grounding Monitoring Data

The hourly meteorological data (including wind speed, wind direction, pressure, sea level pressure, water vapor pressure, temperature, and humidity) and pollutant monitoring data of 2014 were collected from the related official websites (http://113.108.142.147:20035/emcpublish/, http://zx.bjmemc.com.cn/).

2.4 Data Processing and Integration

The research results of Wang et al (2020) has showed that full-coverage and high-precision PM2.5 spatial distribution data in JingJinJi can be generated based on AOD, gaseous pollutants and meteorological factors. Details can be found in related article (Wang et al., 2020)). Therefore, the data processing results were the 3 km resolution grid PM2.5 concentration values covering the entire JingJinJi obtained using MODIS 3 km Aqua AOD and the 3 km resolution grid meteorological data generated through Kriging interpolation.

During data integration, the grid PM2.5 concentration on day t was matched to the t-day grid meteorological data and t-1, t-2, t-3…day grid PM2.5 concentrations. Day of year (DOY, range 1–365) and grid position (GP, row number m and column number n of the 3 km resolution grid) were used as predictors to reflect the spatiotemporal heterogeneity.

2.5 Method

During prediction, the input variables of STAR model are the pixel values of the same and adjacent pixel positions, and the prediction result is the pixel values of each pixel (or window), as shown in Eq. (1). That is, the pixel value $p(x, y, t)$ of at GP (x, y)time t is the spatiotemporal function of the grid pixel value of the adjacent image.

$$p(x, y, t) = \varphi(p(x \pm \Delta x_i, y \pm \Delta y_i, t - \Delta t_i)), \qquad (1)$$

https://doi.org/10.5194/isprs-annals-V-3-2022-303-2022

where (x, y, t) is the location of raster cell at a given time, (Δx_i, Δy_i, Δt_i) denotes the spatiotemporal structure of the adjacent raster cells, and φ denotes an simulation function that can be linear or nonlinear.

As shown in Figure 2, considering the PM2.5 concentration of the two previous days (T=2) adjacent to the ($\Delta x_i = \Delta y_i = 1$) grid to predict PM2.5 concentration $p(m, n, t)$ of m-row and n-column grid at time t, Eq. (1) can be rewritten as follows:

$$p(m,n,t) = \sum_{i=1}^{2}\left[\sum_{y=n-1}^{n+1}\sum_{x=m-1}^{m+1}(W_{xy}^{i}\, p(x,y,t-i)\right] + W_{mn}^{t} q(m,n,t) + \varepsilon(m,n,t), \tag{2}$$

where $p(x, y, t - \Delta t_i)$ indicates the inversion results of Aqua AOD at position (x, y) and time $t - \Delta t_i$, $q(m, n, t)$ is the meteorological variable at position (m, n) and time t, W_{xy}^{i} and W_{mn}^{t} are the corresponding weight coefficients, and $\varepsilon(m, n, t)$ is the error term.

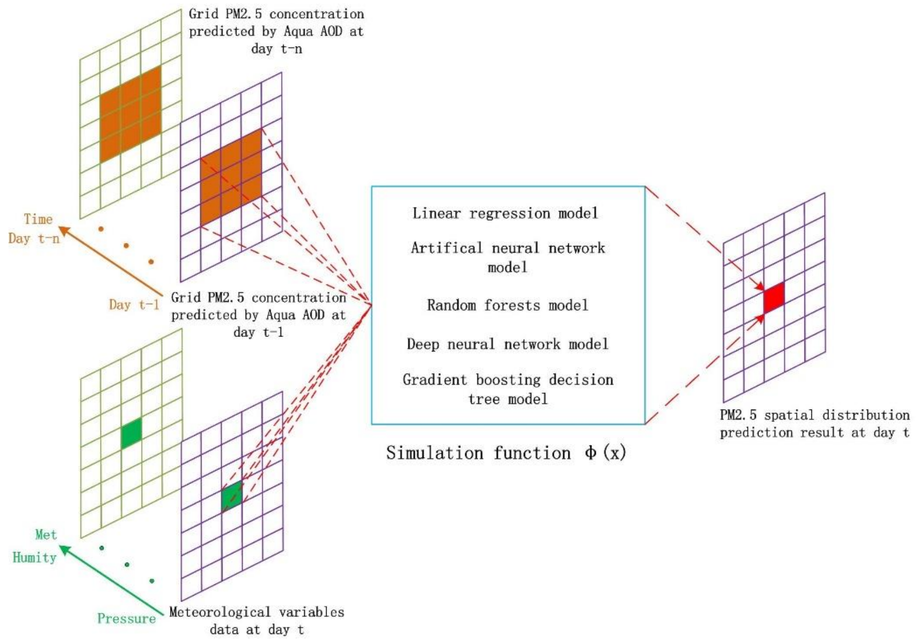

Figure 2. Schematic of the STAR model considering the influence of meteorological factors.

This study explored the effects of nonlinear functions, such as ANN, random forest (RF), deep NN (DNN), and gradient boosting decision tree (GBDT), on model accuracy.

2.6 Model Validation

During model validation, all the data were randomly divided into two groups, where 90% of the matching data were used for model fitting, and the remaining 10% were used for model validation. The coefficient of determination (R^2), root mean square prediction error (RMSE), index of agreement (IA), and mean absolute error (MAE) were used to estimate model performance and were defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \tag{3}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i| \tag{4}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2} \tag{5}$$

$$\text{IA} = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|O_i - \bar{O}| + |P_i - \bar{O}|)^2} \tag{6}$$

where N is the number of samples, O_i and P_i are the observation and prediction results, respectively, and $\bar{O}$ is the average of the observations.

2.7 Parameters Settings

The machine learning models used in this study were all built using the "Keras" model in Python 3.6.0. The main parameter settings of each model are shown in Table 1.

Model	Parameters	Value
ANN	Activation function	'relu'
	Hidden layer size	12
	Learning rate	'constant'
RF	Max depth	8
	Random state	0
DNN	Activation function	'relu'
	Loss function	'mse'
	Hidden layers	6
	Hidden layers nodes	200
GBDT	Loss function	'ls'
	Learning rate	0.1
	Boosting stages numbers	3000
	Max depth	4

Table 1. Parameter settings of machine learning models.

3. RESULTS AND DISCUSSIONS

3.1 Model Validation Results

Figure 3 shows the scatter plots of the STAR model with five simulation functions, namely, LR, ANN, DF, DNN, and GBDT.

https://doi.org/10.5194/isprs-annals-V-3-2022-303-2022

(a1)–(e1) show the scatter plots comparing the predicted and AOD-based $PM_{2.5}$ inversion results. (a2)–(e2) show the scatter plots comparing the predicted results and monitoring station PM2.5 concentration. As shown in (a1)–(e1), the STAR model had the best predictive performance when GBDT was used as the simulation function. The values of R^2, RMSE, IA, and MAE were 0.85, 27.08 μg/m^3, 0.96, and 20 μg/m^3, respectively. The R^2 using DNN reduced from 0.78 to 0.07 compared with GBDT. The model performance was the worst when the simulation function was LR. The R^2 and IA values were 0.68 and 0.88, respectively, and the RMSE and MAE values were 39.57 and 29.56 μg/m^3, respectively, which may be because of the complex nonlinear relationship of PM2.5 and meteorological factors (Wang and Sun 2019; Elangasinghe et al., 2014; Kukkonen et al., 2003). As shown in (a2)–(e2), the model performance in terms of the prediction results of monitoring site PM2.5 was inferior to that compared with the AOD-based PM2.5 inversion results. In this case, the simulation function of the STAR model with the best performance was GBDT, and its R^2, RMSE, IA, and MAE values were 0.75, 40.30 μg/m^3, 0.92, and 30.04 μg/m^3, respectively. Compared with the predicted performance of the model shown in Figure 1(e1), the R^2 and IA values decreased by 0.1 and 0.04, respectively, whereas the RMSE and MAE values increased by 13.22 and 10.04 μg/m^3, respectively. The model performance decreased the most when DNN was used as the simulation function. The R^2 and IA values decreased by 0.15 and 0.06, respectively, whereas the RMSE and MAE values increased by 15.89 and 12.05 μg/m^3, respectively. The performance degradation of the model was mainly because our data for +1-day PM2.5 prediction were the AOD-based PM2.5 inversion results. The inversion results had certain errors compared with the station monitoring results. Therefore, error propagation occurred during the +1-day PM2.5 spatial distribution prediction, resulting in degraded model performance.

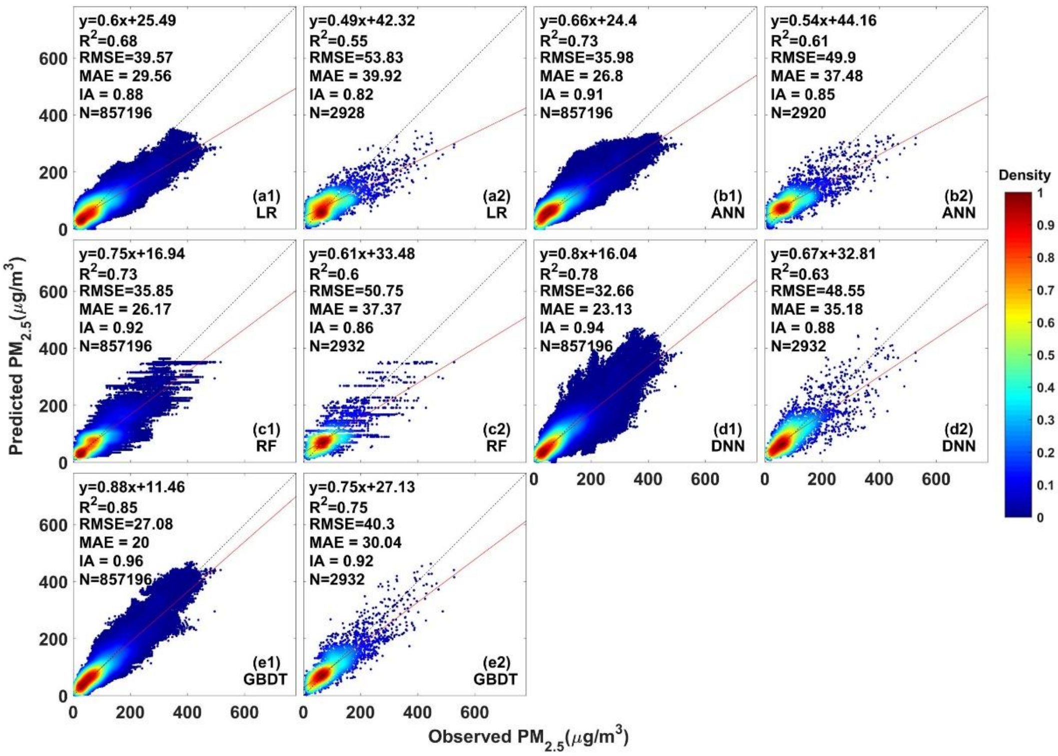

Figure 3. Scatter plots of the STAR model with five different functions. (a1)–(e1) show the scatter plots comparing the predicted and AOD-based PM2.5 inversion results. (a2)–(e2) show the scatter plots comparing the predicted results and monitoring station PM2.5 concentration.

Table 2 shows the STAR model performance statistics using different simulation functions and predictors. The STAR model under five simulation functions had similar prediction performance when the PM2.5 spatial distribution data of the previous day were used as predictors. This condition may be because the relationship of PM2.5 presented a linear relationship rather than a complex nonlinear relationship (Wang et al., 2019). The comparison experiments demonstrated that the introduction of DOY and GP immensely improve the performance of the STAR model using a nonlinear simulation function. The model performance immensely improved with the introduction of DOY and GP as predictors and GBDT as the simulation function. The R^2 and IA values decreased by 0.14 and 0.05, respectively, whereas the RMSE and MAE values increased by 10.46 and 7.79 μg/m^3, respectively, indicating that +1-day PM2.5 spatial distribution prediction results were immensely affected by time and location factors.

Model	Variables	R^2	RMSE(μg/m^3)	IA	MAE(μg/m^3)
LR	PM	0.62	42.91	0.85	31.02
	PM+MET	0.68	39.28	0.89	29.48
	PM+MET+GP	0.68	39.57	0.88	29.56
ANN	PM	0.62	42.9	0.85	31.02
	PM+MET	0.69	38.76	0.89	28.96

	PM+MET+GP	0.73	35.98	0.91	26.80
RF	PM	0.63	42.30	0.86	30.17
	PM+MET	0.68	39.18	0.9	28.32
	PM+MET+GP	0.74	35.85	0.92	26.16
DNN	PM	0.61	43.31	0.85	30.57
	PM+MET	0.60	43.95	0.89	31.62
	PM+MET+GP	0.78	32.66	0.94	23.13
GBDT	PM	0.63	42.32	0.86	30.13
	PM+MET	0.71	37.54	0.91	27.79
	PM+MET+GP	**0.85**	**27.08**	**0.96**	**20.00**

Table 2. Performance statistics of the STAR model under different simulation functions and predictors.

Figure 4 shows the statistics of IA and MAE values of the STAR model under five simulation functions at different monitoring sites. As shown in Figure 4, the spatial distribution of MAE at each monitoring was regional. The areas with small MAE were mainly distributed in the northern areas, such as Chengde, Qinhuangdao, and Zhangjiakou, whereas the areas with large MAE were mainly concentrated in southern Beijing, Baoding, Shijiazhuang, Handan, and Tangshan. The two main reasons for this phenomenon were provided as follows: First, the PM2.5 concentration in different regions was different. MAE was high in areas with high PM2.5 concentration, whereas MAE was relatively low in areas with low PM2.5 concentration. Figure 5(a) shows the PM2.5 annual average concentration of each monitoring station in JingJinJi in 2014. The areas with high annual average PM2.5 concentration were concentrated in the south of Beijing, Baoding, Shijiazhuang, Handan, and Tangshan, which was consistent with high MAE areas of the prediction results. As shown in the scatter plot in Figure 3, underestimation was serious when the PM2.5 concentration was high, resulting in high MAE. Second, the regional MAE spatial distribution may be related to the AOD-based PM2.5 inversion performance. Figure 5(b) shows the AOD-based PM2.5 inversion performance statistics for each monitoring site. The areas with large RMSE were consistent with high MAE areas of the prediction results. As shown in Figure 4, the MAE of each monitoring site was significantly lower than those of the four other simulation functions when GBDT was used as the simulation function. From a regional perspective, the areas with large MAE improvement of GBDT were mainly the south of Beijing, Baoding, and Shijiazhuang. In the southern part of Beijing, the MAE of GBDT reduced by approximately 30 μg/m^3 compared with the other simulation functions. The above comparison experiments showed that the STAR model had the best predictive performance when GBDT was used to simulate the STAR model during the +1-day PM2.5 spatial distribution prediction. The experimental results based on GBDT are shown as follows.

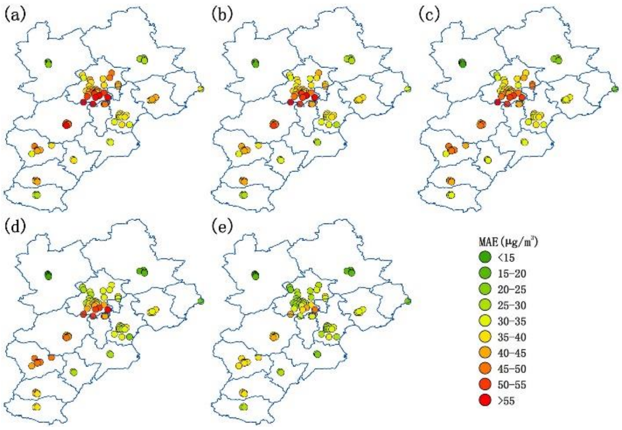

Figure 4. Spatial distribution of MAE of the STAR model at different monitoring sites using different simulation functions. (a)–(e) is LR, ANN, RF, DNN, and GBDT, respectively.

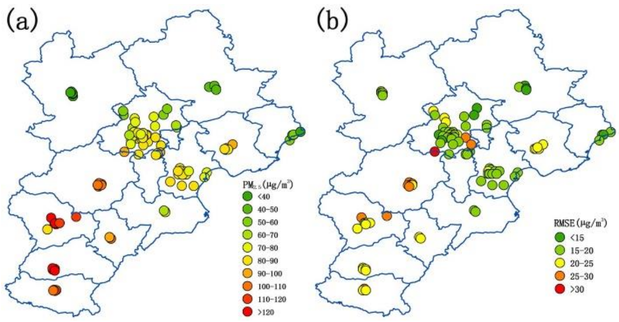

Figure 5. (a) PM2.5 annual average concentration of the monitoring sites in JingJinJi in 2014; (b) PM2.5 inversion performance statistics of each monitoring site.

3.2 Model Parameters Determination

Table 3 shows the performance statistics of the STAR model under different parameters using GBDT as the simulation function. As shown in Table 3, the model performance slightly changed with the increase in Δx_i and T. This condition may be because many prediction variables were introduced into the model with the increase in Δx_i or T. These variables were far from the destination grid and did not contribute to the model performance. Variables Δx_i and T were set to one for reducing the model complexity.

	Δx_i=1				Δx_i=2				Δx_i=3			
	R^2	RMSE	IA	MAE	R^2	RMSE	IA	MAE	R^2	RMSE	IA	MAE
T=1	0.85	27.08	0.96	20.00	0.84	27.88	0.96	20.51	0.85	27.29	0.96	20.22
T=2	0.85	27.24	0.96	20.25	0.85	27.09	0.96	19.95	0.85	27.13	0.96	20.08

https://doi.org/10.5194/isprs-annals-V-3-2022-303-2022 |

T=3	0.85	27.18	0.96	20.22	0.85	27.10	0.96	19.98	0.85	27.01	0.96	19.90

Table 3. Performance statistics of the STAR model under different parameters

3.3 Prediction Maps of PM2.5 Spatial Distribution

Figure 6 shows the prediction results of +1-day PM2.5 spatial distribution during heavy pollution from October 6, 2014 to October 12, 2014. The prediction results of the STAR model were consistent with the PM2.5 spatial distribution inversion and site monitoring results, thereby accurately reflecting the emergence, diffusion, and disappearance of PM2.5 during heavy pollution and providing spatiotemporal distribution data for reducing air pollution and air pollution early warning. The PM2.5 spatial distribution prediction results in this study produced high-value underestimation (Figure 6(b1) and (b2)) and low-value overestimation (Figure 6(g1) and (g2)) with the change of PM2.5. Underestimation occurred when the pollution was serious (Figure 6(d1) and (d2)).

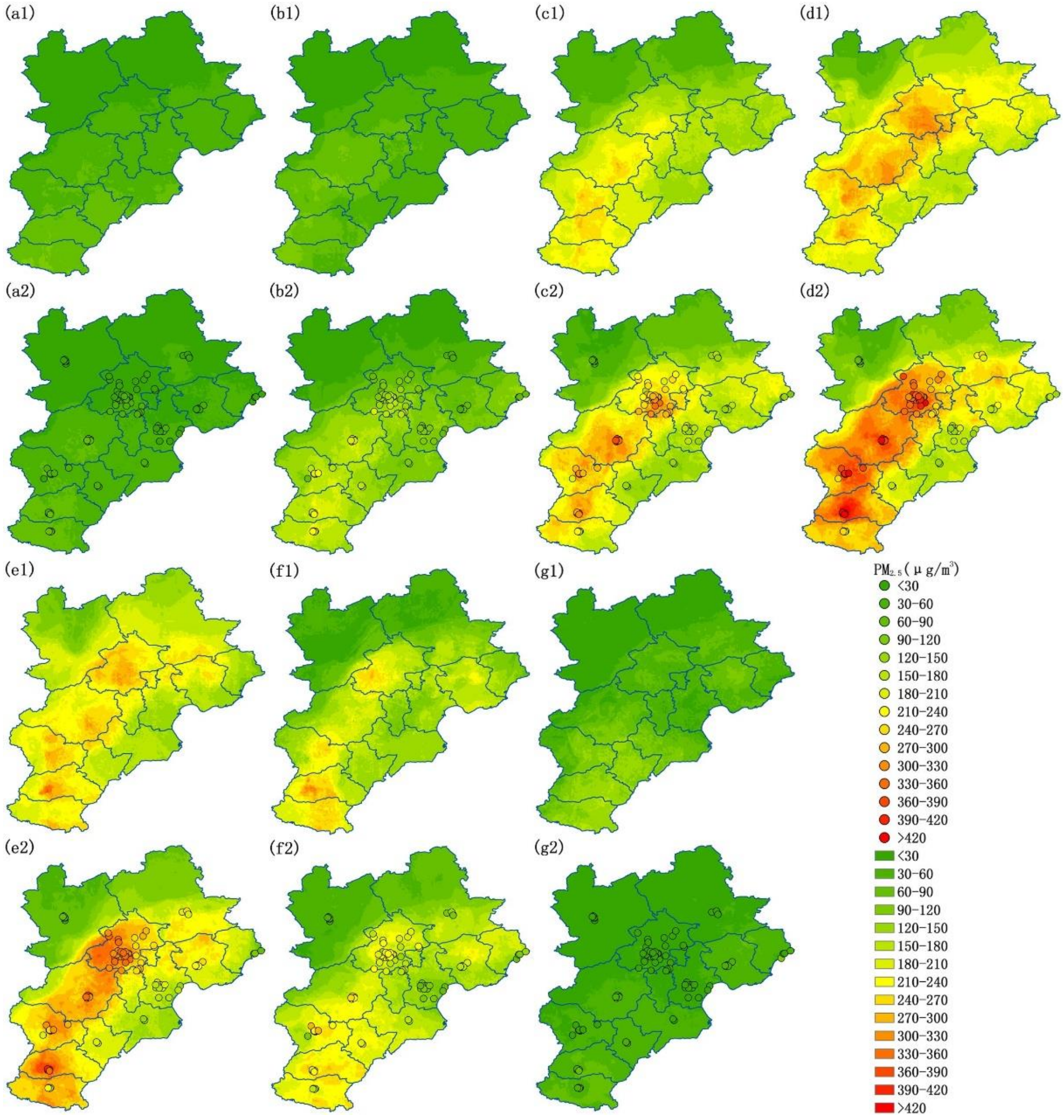

Figure 6. +1-day PM2.5 prediction results from October 6, 2014 to October 12, 2014. (a1)–(g1) show the +1-day PM2.5 prediction results; (a2)–(g2) show the PM2.5 inversion results and the site monitoring PM2.5 concentration.

3.4 Feasibility and Uncertainty Analysis

Previous studies have improved PM2.5 prediction accuracy by introducing the PM2.5 concentration from adjacent monitoring sites (Zheng et al., 2015; Kukkonen et al., 2003; Li et al., 2015; Wen et al., 2019), indicating that PM2.5 has a strong spatial–temporal autocorrelation. Based on this, we established the STAR model to realize the +1-day region PM2.5 spatial distribution prediction. The pollutant concentrations were cyclical because of the influence of time factors (Zhang et al., 2012). Therefore, day of week and DOY were frequently used as predictors to improve the pollutant prediction performance ADDIN(Kurt and Oktay 2010; Qi et al., 2019; Feng et al., 2015). The PM2.5 spatial distribution maps showed that the PM2.5 spatial distribution was regional. Therefore, the introduction of DOY and GP improved the prediction accuracy of the model.

Scholars have conducted numerous studies on the prediction of PM2.5 in JingJinJi using statistical models (Soh et al., 2018; Feng et al., 2015; Li et al., 2016; Qi et al., 2019). Although these studies have achieved high prediction accuracy, some limitations are found because they only predicted the PM2.5 concentration of monitoring sites. First, air quality is affected by complex factors, such as meteorological factors, transportation, and land use types, and immensely varies with time and location (Zheng et al., 2013). Therefore, single-site pollutant concentration prediction cannot effectively help people to make decisions. Different models are required in predicting the pollutant concentration at different stations (Zheng et al., 2015). Second, these studies can only predict the PM2.5 concentration in areas with monitoring stations. Taking the study area as an example, JingJinJi has few monitoring stations that are mainly concentrated in urban areas. Therefore, statistical models in previous studies cannot be used to predict the pollutant concentration in vast areas without monitoring sites. Other scholars used the deterministic models to simulate the PM2.5 concentration in eastern China (Zhou et al., 2017; Zheng et al., 2015). Although the regional PM2.5 concentration prediction can be achieved, the model accuracy is low with R^2 are 0.45 and 0.64, respectively. The model used in this study fully considered the spatial–temporal autocorrelation of PM2.5, not only extended the PM2.5 concentration prediction to other areas without monitoring sites, but also achieve regional PM2.5 concentration prediction with high accuracy, thereby overcoming the limitations of statistical models and deterministic models to some extent.

The model prediction results in this study had underestimation problems. This condition was because the AOD-based PM2.5 inversion results had certain errors, thereby resulting in error propagation. Therefore, the accuracy of the STAR model can be enhanced to some extent by improving the AOD-based inversion accuracy. At the same time, some advanced statistical models, such as long-short memory DNN (Li et al., 2017), can be used to solve the underestimation of the model.

4. CONCLUSION

Statistical models can only predict the pollutant concentration at the monitoring sites and cannot be extended to other regions with different meteorological conditions and without monitoring sites. This study establish a STAR model based on the spatial distribution of PM2.5 predicted using MODIS AOD for predicting the one day ahead PM2.5 spatial distribution in JingJinJi. The results showed that the performance of the STAR model was relatively different compared with different simulation functions. The model performance was the best when GBDT was used as the simulation function. The R^2, RMSE, IA, and MAE values were 0.85, 27.08 μg/m^3, 0.96, and 20 μg/m^3, respectively. The introduction of DOY and GP immensely improved the model performance, and R^2 and IA decreased by 0.14 and 0.05, respectively, whereas RMSE and MAE increased by 10.46 and 7.79 μg/m^3, respectively, indicating that the PM2.5 spatial distribution prediction results were immensely affected by time and location factors. The performance statistics of each monitoring station showed that the model performance distribution was regional. The regions with low PM2.5 concentration had better performance, whereas the regions with high PM2.5 concentration had poor performance. The PM2.5 prediction results of heavily polluted weather indicated that the model can accurately reflect the emergence, diffusion, and disappearance of PM2.5 during heavy pollution and provide spatiotemporal distribution data for reducing air pollution and air pollution early warning.

ACKNOWLEDGEMENTS

This work was supported by the Special Project of Science and Technology Basic Resources Survey, China Ministry of Science and Technology, under Grant 2019FY202503.

REFERENCES

Belleudi, V., Faustini, A., Stafoggia, M., Cattani, G., Marconi, A., Perucci, C. A. and Forastiere, F. (2010). Impact of fine and ultrafine particles on emergency hospital admissions for cardiac and respiratory diseases. *Epidemiology*:414-423. doi: 10.1097/EDE.0b013e3181d5c021.

Chen, Y., Shi, R., Shu, S. and Gao, W. (2013). Ensemble and enhanced PM10 concentration forecast model based on stepwise regression and wavelet analysis. *Atmos. Environ.* 74:346-359. doi: 10.1016/j.atmosenv.2013.04.002.

Cheng, Q., Liu, H., Shen, H., Wu, P. and Zhang, L. (2017). A spatial and temporal nonlocal filter-based data fusion method. *Ieee T. Geosci. Remote* 55:4476-4488. doi: 10.1109/TGRS.2017.2692802.

Cobourn, W. G. (2007). Accuracy and reliability of an automated air quality forecast system for ozone in seven Kentucky metropolitan areas. *Atmos. Environ.* 41:5863-5875. doi: 10.1016/j.atmosenv.2007.03.024.

Cortina Januchs, M. G., Quintanilla Dominguez, J., Vega Corona, A. and Andina, D. (2015). Development of a model for forecasting of PM10 concentrations in Salamanca, Mexico. *Atmos Pollut Res* 6:626-634. doi: 10.1007/s00371-007-0114-y.

Crespo, J. L., Zorrilla, M., Bernardos, P. and Mora, E. (2007). A new image prediction model based on spatio-temporal techniques. *The Visual Computer* 23:419-431. doi: 10.1007/s00371-007-0114-y.

Crouse, D. L., Peters, P. A., Van, D. A., Goldberg, M. S., Villeneuve, P. J., Brion, O., Khan, S., Atari, D. O., Jerrett, M. and Pope, C. A. (2012). Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. *Environ Health Perspect* 120:708-714. doi: 10.1289/ehp.1104049.

Das, M. and Ghosh, S. K. (2016). Deep-STEP: A deep learning approach for spatiotemporal prediction of remote sensing data. *Ieee Geosci Remote S.* 13:1984-1988. doi: 10.1109/LGRS.2016.2619984.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L. and Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama* 295:1127-1134. doi: 10.1001/jama.295.10.1127.

Hx, A., Gw, B., Cl, C. and Ms, D., (2021). PM2.5 concentration modeling and prediction by using temperature-based deep belief network. *Neural Networks*, 133:157-165.

Hooyberghs, J., Mensink, C., Dumont, G., Fierens, F. and Brasseur, O. (2005). A neural network forecast for daily average PM10 concentrations in Belgium. *Atmos. Environ.* 39:3279-3289. doi: 10.1016/j.atmosenv.2005.01.050.

Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T. and Foxall, R. (2003). Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmos. Environ.* 37:4539-4550. doi: 10.1016/S1352-2310(03)00583-1.

Li, R., Gong, J., Chen, L. and Wang, Z. (2015). Estimating ground-level PM2.5 using fine-resolution satellite data in the megacity of Beijing, China. *Aerosol Air Qual. Res* 15:1347-1356. doi: 10.4209/aaqr.2015.01.0009.

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C. and Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environ. Pollut.* 231:997-1004. doi: 10.1016/j.envpol.2017.08.114.

Nieto, P. G., Combarro, E. F., Del Coz Díaz, J. J. and Montañés, E. (2013). A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study. *Appl. Math. Comput.* 219:8923-8937. doi: 10.1016/j.amc.2013.03.018.

Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K. and Pak, C., (2020). Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China. *The Science of the Total Environment*, 699 (Jan.10):133561.

Pope III, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K. and Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama* 287:1132-1141.

Sánchez, A. S., Nieto, P. G., Iglesias-Rodríguez, F. J. and Vilán, J. V. (2013). Nonlinear air quality modeling using support vector machines in Gijón urban area (Northern Spain) at local scale. *Int. J. Nonlin. Sci. Num.* 14:291-305. doi: 10.1515/ijnsns-2012-0119.

Wang, X., Sun, W., Wang, Z., Wang, Y. and Ren, H. (2019). Meteorological Parameters and Gaseous Pollutant Concentrations as Predictors of Ground-level PM2. 5 Concentrations in the Beijing-Tianjin-Hebei Region, China. *Aerosol Air Qual Res* 19:1844-1855. doi: 10.4209/aaqr.2018.12.0449.

Wang, X., Sun, W., Zheng, K., Ren, X. and Han, P. (2019). Estimating hourly PM2.5 concentrations using MODIS 3 km AOD and an improved spatiotemporal model over Beijing-Tianjin-Hebei, China. *Atmos. Environ.* doi: 10.1016/j.atmosenv.2019.117089.

Yang, Q., Yuan, Q., Yue, L., Li, T., Shen, H. and Zhang, L. (2019). The relationships between PM2. 5 and aerosol optical depth (AOD) in mainland China: About and behind the spatio-temporal variations. *Environ. Pollut.* 248:526-535. doi: 10.1016/j.envpol.2019.02.071.

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C. and Baklanov, A. (2012). Real-time air quality forecasting, part I: History, techniques, and current status. *Atmos. Environ.* 60:632-655. doi: 10.1016/j.atmosenv.2012.06.031.

Zhang, Y., Zhao, D., Ji, X., Wang, R. and Gao, W. (2009). A spatio-temporal auto regressive model for frame rate upconversion. *Ieee T. Circ. Syst. Vid.* 19:1289-1301. doi: 10.1109/TCSVT.2009.2022798.

Zheng, B., Zhang, Q., Zhang, Y., He, K. B., Wang, K., Zheng, G. J., Duan, F. K., Ma, Y. L. and 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. *Atmospheric Chemistry & Physics* 15:2031-2049. doi: 10.5194/acp-15-2031-2015.

Zheng, Y., Liu, F. and Hsieh, H. P. (2013). U-Air: When urban air quality inference meets big data, in *Acm Sigkdd International Conference on Knowledge Discovery & Data Mining*. doi: 10.1145/2487575.2488188.

Zheng, Y., Yi, X., Li, M., Li, R., Shan, Z., Chang, E. and Li, T. (2015). Forecasting fine-grained air quality based on big data, in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2267-2276. doi: 10.1145/2783258.2788573.

Zheng, Y., Zhang, Q., Liu, Y., Geng, G. and He, K. (2016). Estimating ground-level PM2.5 concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. *Atmos. Environ.* 124:232-242. doi: 10.1016/j.atmosenv.2015.06.046.

Zhou, G., Xu, J., Xie, Y., Chang, L., Gao, W., Gu, Y. and Zhou, J. (2017). Numerical air quality forecasting over eastern China: An operational application of WRF-Chem. *Atmos. Environ.* 153:94-108. doi: 10.1016/j.atmosenv.2017.01.020.