Pol. J. Environ. Stud. Vol. 31, No. 4 (2022), 3879-3891
DOI: *10.15244/pjoes/147084* ONLINE PUBLICATION DATE: 2022-05-30

Original Research

Spatial-Temporal Pattern and Influencing Factors of $PM_{2.5}$ Pollution in North China Plain

Xiaoyang Wang[1,2], Mingliang Ma[1]*, Lin Guo[1], Yuqiang Wang[3], Guobiao Yao[1], Fei Meng[1], Mingyang Yu[1]

[1]School of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan 250101, China
[2]School of Geographic Sciences, East China Normal University, Shanghai 200241, China
[3]Disaster Reduction Center of Shandong Province, Jinan 250102, China

Received: 2 February 2022
Accepted: 3 March 2022

Abstract

It is of great significance to reveal the spatiotemporal pattern of $PM_{2.5}$ pollution in North China Plain (NCP), and relative importance of influencing factors for controlling $PM_{2.5}$ pollution in NCP. In this work, statistical models between $PM_{2.5}$ and a set of influential factors were firstly established to pinpoint the dominant factors contributing to $PM_{2.5}$ variations in NCP and to estimate the daily $PM_{2.5}$ dataset. The results of model attribution analysis showed that emission factors such as NO_2 and SO_2 were the two dominant influencing factors of $PM_{2.5}$ pollution in NCP, while temperature, relative humidity and wind speed were three main meteorological factors that significantly contributed to $PM_{2.5}$ pollution. Subsequently, spatiotemporal analysis results showed that there was obvious spatial agglomeration of $PM_{2.5}$ pollution in NCP, and the most serious $PM_{2.5}$ pollution occurred at Beijing and the junction of Hebei, Shanxi and Henan. Meanwhile, $PM_{2.5}$ pollution has a strong seasonal pattern, and the highest pollution is in winter while the lowest is in summer. In addition, $PM_{2.5}$ pollution level in NCP showed a downward trend from 2015 to 2020. Among them, the most serious $PM_{2.5}$ pollution in 2020 decreased most significantly, which was 71.52 % lower than that of 2019.

Keywords: $PM_{2.5}$ pollution, spatial-temporal analysis, statistical modeling, air quality management

Introduction

$PM_{2.5}$ refers to atmospheric particulate matter with a dynamic diameter less than 2.5 μm [1]. With the rapid urbanization in China in recent years, the issue of $PM_{2.5}$ pollution has attracted more and more attention. $PM_{2.5}$ pollution may cause adverse effects on human health [2-7] and the environment [8-9], and even influence the climate change [10]. Due to the small particle size of $PM_{2.5}$, it is easy to be inhaled, and then it is also easy to cause respiratory diseases, affecting the cardiovascular system [11] and nervous system [3]. Besides, it can also cause to cancer [2] or premature death [4-6]. Once $PM_{2.5}$ pollution increases, atmospheric visibility decreases [8], which will affect people's travel, meanwhile, the solar radiation time on the ground has become shorter, which will affect plant photosynthesis and ecosystems. In addition, $PM_{2.5}$ particles dissolved in rainwater will produce sulfuric acid and nitric acid through a series

*e-mail: mamingliang19@sdjzu.edu.cn

of chemical reactions, which will cause harm to buildings and vegetation health [9]. Finally, the presence of $PM_{2.5}$ can even influence climate change by altering the earth's energy balance [10].

Therefore, satellite-based $PM_{2.5}$ pollution mapping and the $PM_{2.5}$ pollution mechanism revealing are of great significance for the $PM_{2.5}$ pollution forecasting, public health protecting, environmental pressure relieving, and air pollution improvement measures formulating. In addition, obtaining large-scale continuous dynamic $PM_{2.5}$ data through model estimation can provide basic data guarantee for current atmospheric environmental governance, which is a hot issue in current $PM_{2.5}$ pollution research.

$PM_{2.5}$ ground monitoring, atmospheric chemical transport modeling, and statistical satellite-based $PM_{2.5}$ estimation modeling are three mainly used methods for $PM_{2.5}$ prediction. First, $PM_{2.5}$ ground monitoring is estimated based on the values measured by ground-based monitoring stations. This method was widely used before the development of remote sensing technology, the measurement method is simple, and the measurement results are not affected by cloud coverage and surface conditions [12]; however, the ground measurement method is greatly affected by the distribution and terrain, etc. When the sites in the research area are densely distributed, the estimation results are more accurate, while the $PM_{2.5}$ estimation results are not that accurate in areas with sparse sites [13-14]. Second, the complex chemical and physical mechanism inherent between $PM_{2.5}$ and influence factors is considered in the atmospheric chemical transport model, and this kind of models is theoretically reliable [15]. However, the computational time and memory requirements of this model are high, which limits the application of it [16-17]. Finally, the $PM_{2.5}$ statistical model based on satellite remote sensing is based on the statistical relationship between AOD and $PM_{2.5}$, coupling with other factors affecting $PM_{2.5}$. The $PM_{2.5}$ estimations results based on remote sensing has high timeliness, certain periodicity and is continuous in a large spatial coverage.

With the progress of research, the spatial representation of the first method, and higher operation speed cost of the second method limited the extensive continuous dynamic estimation of $PM_{2.5}$. However, the $PM_{2.5}$ statistical model based on satellite remote sensing has become a main method to obtain a large range of continuous $PM_{2.5}$ concentration. $PM_{2.5}$ statistical model based on satellite remote sensing is constantly evolving, from the simplest linear model [18-20] to complex statistical model. In addition, many studies have shown that there is a complex non-linear relationship between $PM_{2.5}$ and many influencing factors. Machine learning model can better characterize the complex non-linear relationship between $PM_{2.5}$ and many influencing factors, and it can also effectively process big data and measure the importance of each influencing factor. Taking into account the high accuracy and the simple implementation of the machine learning methods, they are introduced into $PM_{2.5}$ estimation at this stage, such as BP's artificial neural network model, support vector machine model and random forest [21-22]. Chen et al. estimated daily $PM_{2.5}$ concentrations across China from 2005-2016, showing that daily random forest models had much higher accuracy than conventional regression models [23]. Zhao et al. established a random forest model to estimate high resolution daily $PM_{2.5}$ concentrations in Beijing-Tianjin with the R^2 value of 0.83 [24].

To obtain higher $PM_{2.5}$ estimation accuracy, it is necessary to comprehensively sort out the influencing factors of $PM_{2.5}$. Previous studies found that the main factors affecting $PM_{2.5}$ concentration include: AOD, SO_2, NO_2, meteorological factors, socio-economic factors, land use [25-27]. Meanwhile, energy consumption and NO_2 and SO_2 produced by human activities (such as automobile exhaust and factory exhaust) are the main drivers of $PM_{2.5}$ concentration [28-31]. Besides, meteorological factors can influence the distribution of $PM_{2.5}$ by affecting the formation, diffusion and settlement of particulate matter [32-35]. In addition, social factors such as regional GDP, population density, as well as vegetation cover [36-38] can also affect the distribution of $PM_{2.5}$.

This paper attempted to build a $PM_{2.5}$ estimation model for NCP and generated a spatially surface $PM_{2.5}$ dataset over NCP by using sites-based $PM_{2.5}$ observations, MODIS/AOD data, MODIS/NDVI data, satellite-observed emissions data (i.e., OMI/SO_2 and OMI/NO_2), meteorological factors, population density, gross regional product (GDP), and DEM. Subsequently, the random forest model was utilized in to build the $PM_{2.5}$ estimation model, and then daily $PM_{2.5}$ dataset with a resolution of 0.1°×0.1° in NCP was calculated and generated from this $PM_{2.5}$ estimation model. Finally, the spatial and temporal distribution of $PM_{2.5}$ concentration was revealed. The study aims to: (1) construct a high resolution daily $PM_{2.5}$ estimation model suitable for NCP, and (2) explore the spatial and temporal distribution patterns of $PM_{2.5}$ in NCP using $PM_{2.5}$ estimation results.

Data and Methods

Study Area

NCP was the study area in this paper (shown in Fig. 1). NCP locates at 32°N~42°N, 110°E~120°E, and there are five provinces in this region including Beijing, Tianjin, Hebei, Henan and Shandong. Besides, it is also the political, economic, and cultural center of China, and it has high temperature and rain in summer, frequent sand and dust storms in spring, cold and dry winter, increased coal heating, large population,

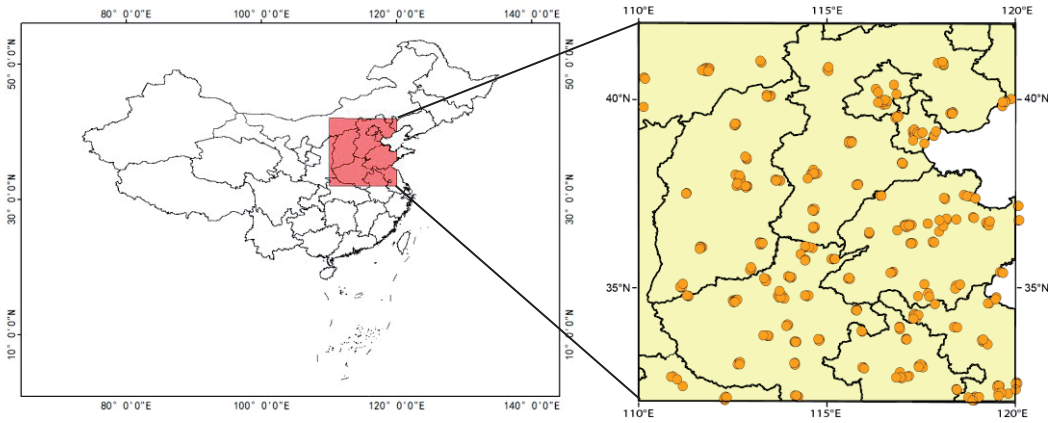

Fig. 1. Location of the study area (left), and spatial distribution of $PM_{2.5}$ monitoring sites in NCP (right).

industrial pollution emissions, automobile gas and dust from construction sites, and sharp contradictions between resource environment and urban development. The air quality in this region has always been the focus and urgent need of environmental research today of China [39]. Establishing a regional $PM_{2.5}$ estimation model and obtaining a large-scale continuous coverage of $PM_{2.5}$ estimation results are of great significance for mitigating regional $PM_{2.5}$ pollution, formulating relevant measures, and alleviating development contradictions.

Data

In this study, 16 kinds of data were utilized, and the data types, data representation, time range, spatial resolution, temporal resolution, and data sources of these datasets were shown in the Table 1.

$PM_{2.5}$ Monitoring Sites

$PM_{2.5}$ measurements from the ground-based sites in NCP were used to build the $PM_{2.5}$ simulation model.

Table 1. Overview of the relevant data in this study.

Data Types	Abbreviation	Time	Spatial Resolution	Temporal Resolution	Data Sources
AOD	MODIS/AOD	2015~2020	0.05°×0.05°	24h	NASA LAADS DAAC
NDVI	MODIS/NDVI	2015~2020	0.05°×0.05°	16d	NASA LAADS DAAC
SO_2	OMI/SO_2	2015~2020	0.25°×0.25°	24h	NASA EARTH DATA
NO_2	OMI/NO_2	2015~2020	0.25°×0.25°	24h	NASA EARTH DATA
Elevation	DEM	2015	1km×1km	-	RESDC
Economic	GDP	2015	0.1°×0.1°	-	RESDC
Population	POPU	2015	0.1°×0.1°	-	RESDC
Boundary Layer Height	BLH	2015~2020	0.1°×0.1°	1h	ECMWF
Surface Solar Radiation	SSRD	2015~2020	0.1°×0.1°	1h	ECMWF
Relative Humidity	RH	2015~2020	0.1°×0.1°	1h	ECMWF
Wind Speed	WS	2015~2020	0.1°×0.1°	1h	ECMWF
Wind Direction	WD	2015~2020	0.1°×0.1°	1h	ECMWF
Surface Pressure	SP	2015~2020	0.1°×0.1°	1h	ECMWF
Temperature	T	2015~2020	0.1°×0.1°	1h	ECMWF
Total Precipitation	TP	2015~2020	0.1°×0.1°	1h	ECMWF
Total Cloud Cover	TCC	2015~2020	0.1°×0.1°	1h	ECMWF

China established the national air pollution (including $PM_{2.5}$) monitoring network at the end of 2012, and the $PM_{2.5}$ data was released on the national air quality release platform from 2013 (http://106.37.208.233:20035/). The spatial distribution of the ground-based monitoring sites in NCP is shown in Fig. 1.

Satellite Data

AOD is a measure of the sun's beam blocked by air pollutants and is widely used as an indicator of near-Earth pollutants [40], is of great significance in the study of air pollution, MODIS AOD product is the most widely used aerosol optical thickness dataset in the $PM_{2.5}$ simulation. The MODIS AOD product is one of the atmospheric products carried by the medium-resolution imaging spectrometer on the Terra and Aqua satellites, which were successfully launched by NASA on December 1, 1999 and April 18, 2002, respectively. MODIS AOD has the characteristics of large spatial-coverage, wide spectrum range, fast data update, and MODIS data can be freely downloaded. Therefore, MODIS AOD data was selected in this study as one of the important input factors for $PM_{2.5}$ estimation. In addition, Terra and Aqua satellites have different transit times, and AOD products from these two satellites cover different spatial ranges. Therefore, in order to improve the temporal representation of AOD data and expand the spatial coverage of AOD data, MOD09CMA and MYD09CMA aerosol products from Terra and Aqua satellites from 2015 to 2020 were averaged in this study, and then the averaged AOD data was utilized to estimate surface $PM_{2.5}$ data in NCP during 2015 and 2020. The spatial resolution of these two daily MODIS AOD products is 0.05°×0.05°.

In addition to AOD, the SO_2, and NO_2 emitted by human activities also have significant influence on the production of $PM_{2.5}$. The SO_2 and NO_2 data from 2015 to 2020 used in this work were L3 daily products observed by the Ozone Measurement Instrument (OMI) loaded in the AURA satellite, and the spatial resolution of these daily datasets is 0.25°× 0.25°. The OMI is a sensor jointly developed by the Netherlands and Finland to measure the concentrations of ozone, HCHO, NO_2, and SO_2, as well as multiple data on aerosols, clouds, and surface ultraviolet radiation. Besides, the Normalized Difference Vegetation Index (NDVI) can reveal the variations of vegetation coverage, impact the environment and climate, and also influence the $PM_{2.5}$ concentration changes. MODIS 16 days NDVI product MYD13C1 from 2015 to 2020 with the spatial resolution of 0.05°× 0.05°was used in this work. Finally, MODIS AOD data [18, 41-42], OMI SO_2 and OMI NO_2 data [27, 43-45] and NDVI product involved in this work were widely used in previous studies and the quality and product accuracy of them are well illustrated in these articles.

Meteorological Reanalysis Data

Meteorological factors can affect the distribution of $PM_{2.5}$ concentration by influencing the formation, diffusion, and sedimentation of particulate matter. [33-35,46]. Through the literature review, boundary layer height, surface solar radiation, relative humidity, wind speed, wind direction, surface air pressure, temperature, total precipitation, and total cloud cover were selected as main $PM_{2.5}$ influencing meteorological factors. Hourly ERA5 meteorological reanalysis data from the European Mid-scale Weather forecast Center (ECMWF, European Center for Medium-Range Weather Forecasts) with a spatial resolution of 0.1°× 0.1° was used in this work to build the $PM_{2.5}$ estimation model, and the time range of the meteorological data was from 1 January 2015 to 31 December 2020. The quality and product accuracy of these reanalysis data were documented in previous studies [47-50].

DEM and Socioeconomic Factors Data

Digital Elevation Model (DEM) affects the transmission and diffusion of air pollutants. Besides, Gross Domestic Product (GDP) and population density (POPU) indicates the level of urban development, the human activity intensity, and they can also influence the pollutant emissions [51-53]. Therefore, the DEM, GDP, and POPU data served as the terrain and socioeconomic influencing factors of $PM_{2.5}$. All these datasets were obtained from the Resource and Environmental Science and Data Center of Chinese Academy of Sciences. The spatial resolution of the DEM, GDP and POPU data are 1km×1km, and all these data were resampled to 0.1°× 0.1°grids. The latest available data provided by the website is the GDP and POPU data of 2015, and these two data of 2015 were used in this work.

Research Technique

Space-Temporal Matching of Multi-Source Data

All the $PM_{2.5}$ influencing factors were spatiotemporal matched and resampled to establish the 0.1°×0.1°grids for all data. The spatial resolution of OMI/SO_2, OMI/NO_2 is 0.25°×0.25°and these two datasets were interpolated to the 0.1°×0.1°grids. Besides, the fused MODIS/AOD, MODIS/NDVI (0.05°×0.05°), and DEM, GDP and POPU data (1km×1km) were resampled to 0.1°×0.1°. Moreover, the temporal resolution of MODIS/NDVI data is 16 days, so we use the NDVI value of the day when the NDVI observation value is available to replace the NDVI value of the next 16 days. In addition, SSRD and TP used in this work are the daily cumulative data, and the T used is the daily highest temperature. While other meteorological factors used in this work are daily average values (local 8 am to 22 am).

$PM_{2.5}$ Estimation Models

Random Forest is first proposed as a classifier by Leo Breiman and trained on training samples using multiple trees [54-55]. It takes the decision tree as the basic unit, combined with the bootstrap resampling method (random sampling), and improves the classification accuracy through the combination of multiple decision trees. Later, the random forest model can also be used for nonlinear regression, and the algorithm is easy to implement and the accuracy of it is high [56]. Meanwhile, the Extreme Gradient Boosting (XGBoost) model is also a machine learning algorithm implemented in the Gradient Boosting framework, and it classifies and predicts the datasets based on the CART regression tree [57-58].

In this study, the nonlinear relationship model between $PM_{2.5}$ and various variables established with random forest model and XGBoost model can be described as follows:

$$PM_{2.5} \sim RF/XGBoost\ (Month, NO_2, SO_2, T, RH, WS, WD, SSRD, BLH, TCC, TP, GDP, POPU, DEM, NDVI, AOD) \quad (1)$$

Where NO_2 and SO_2 represent the concentration of nitrogen dioxide and sulfur dioxide, T, RH, WS, WD, SSRD, BLH, TCC, TP are meteorological data, and the relevant details of these parameters are shown in Table 1. Besides, GDP denotes gross domestic product, POPU is population density, DEM represents the elevation of the corresponding position, and NDVI means the Normalized Difference Vegetation Index.

In the modeling progress of RF and XGBoost, 10% of the sites were randomly selected as the validation sites, and the remaining 90% of the sites were used as the modeling training data and testing data. In these remaining 90% sites, 80% of the data samples are randomly selected as model training data, and the remaining 20% of the data samples are used as testing data. Finally, the accuracy of both models was evaluated by the validation sites data.

Results and Discussion

Models Verification Accuracy

The sites data were randomly divided into validation sites (10%), modeling training data (90%×80%), and modeling testing data (90%×20%). The cross-validation results of the RF $PM_{2.5}$ estimation model and XGBoost $PM_{2.5}$ estimation model are shown in Fig. 2. As shown in Fig. 2, the accuracy of the test results of these two models is relatively high, and the correlation coefficients (R value) between the model simulation results of these two models and the observation data are both higher than 0.81. Similarly, the accuracy of the validation results of these two models is also relatively high with the R values both higher than 0.80. However, when the test and validation results of the RF model and the XGBoost model are compared, the results show that the accuracy of the XGBoost model are

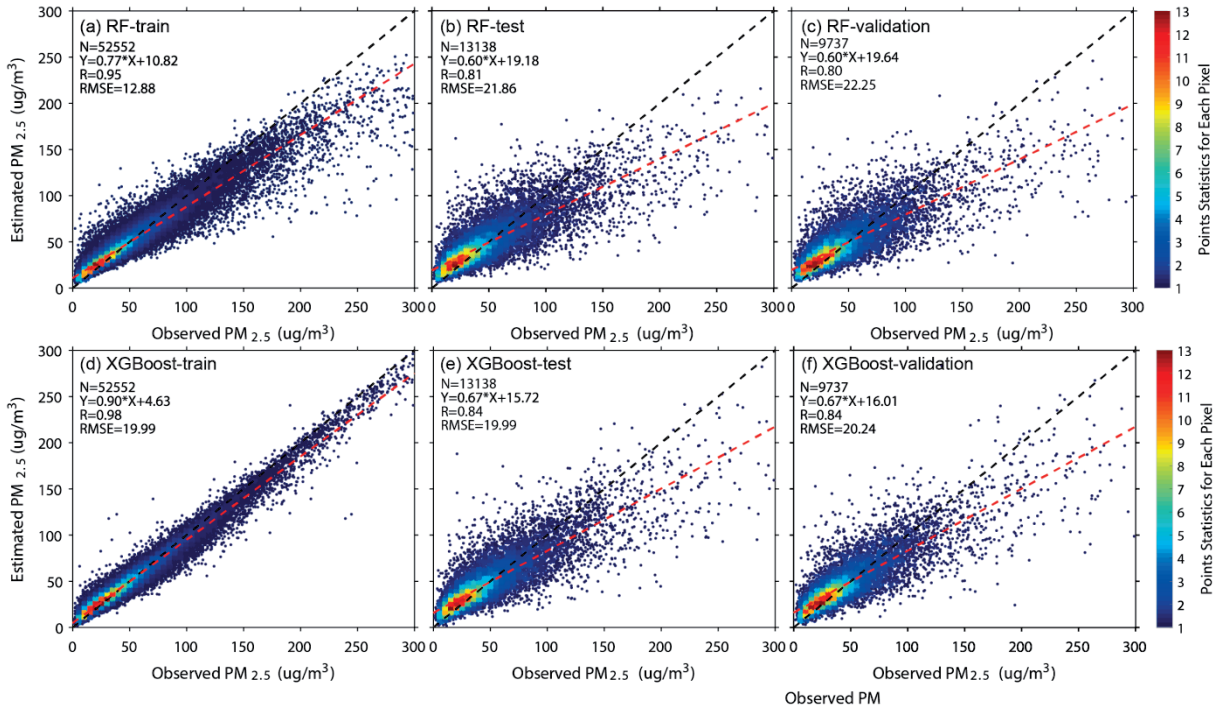

Fig. 2. Comparison results of the cross validation accuracy between RF model and XGBoost model. (a-c) present the scatter plots between observed and estimated $PM_{2.5}$ using RF model. (d-e) are similar with (a-c) but for the results of XGBoost model.

Table 2. Statistical results of RF model and XGBoost model in each season in NCP.

	Season	Sample number	R^2	RMSE ($\mu g\ m^{-3}$)	Mean error ($\mu g\ m^{-3}$)
RF	Spring	22138	0.84	12.42	7.37
	Summer	16680	0.80	8.45	5.29
	Autumn	17113	0.86	13.24	7.54
	Winter	19496	0.87	19.46	11.55
XGBoost	Spring	22138	0.85	11.63	7.60
	Summer	16680	0.79	8.56	5.90
	Autumn	17113	0.87	12.08	7.39
	Winter	19496	0.89	16.78	9.51

higher with the both R values of the test and validation results are 0.84, which are higher than the results of the RF model (R value of 0.81 and 0.80).

Table 2 shows the statistical results of the R^2, RMSE and mean error between the observed $PM_{2.5}$ and model estimated $PM_{2.5}$ using RF model and XGBoost model. The R^2 values of four seasons are all higher than 0.79, and this means that these two $PM_{2.5}$ estimation models have higher model accuracy. Besides, the prediction accuracy of both models in autumn and winter are higher than that in spring and summer, and the highest value appears in winter while the lowest is in summer. From the perspective of RMSE and the mean error, the largest RMSE and mean error value are in winter while the smallest values are in summer. This is due to the serious pollution and the large base of pollutant concentration in winter. Although the absolute error value is relatively large, it is not inconsistent with the high relative accuracy of R^2. In conclusion, the accuracy of the XGBoost model is higher than that of the RF model, so the $PM_{2.5}$ estimation results using XGBoost model were derived and selected for further analysis.

Order of Influence Factor Importance

The relative importance of the 16 variables used in the $PM_{2.5}$ estimation model is shown in Fig. 3. It reveals that NO_2 is the most important influencing factor in the $PM_{2.5}$ estimation model with the highest relative importance, accounting for about 11.4%. Previous studies have shown that $PM_{2.5}$ mainly derived from the NO_2 and SO_2 emitted from fossil fuel combustion and human activity [28-31]. Besides, Fig. 9 revealed that the severe $PM_{2.5}$ pollution generally showed a significantly downward trend with the proportion of severe $PM_{2.5}$ pollution in NCP reduced from 1.9% (2015) to 0.4% (2020). Generally, $PM_{2.5}$ pollution is mainly controlled by emissions caused by human activities, and then the generation of $PM_{2.5}$ pollution is ultimately caused by human activities, so the main reasons for the decline in $PM_{2.5}$ pollution in NCP from 2015 to 2020 was due to the reduction of NO_2 and SO_2 emissions from human activities in the same period of time. This result is consistent with previous relevant research results [59-

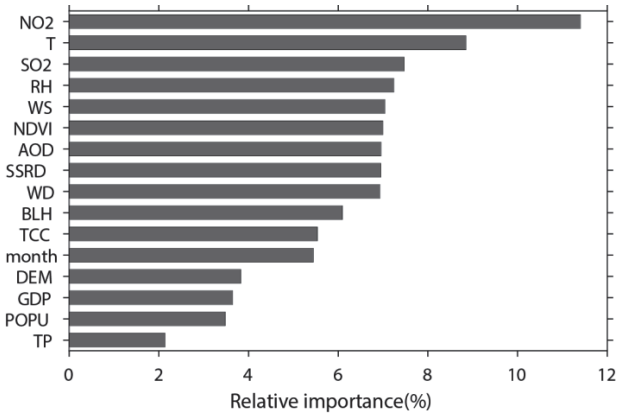

Fig. 3. The relative importance of each influencing factor in the $PM_{2.5}$ estimation model.

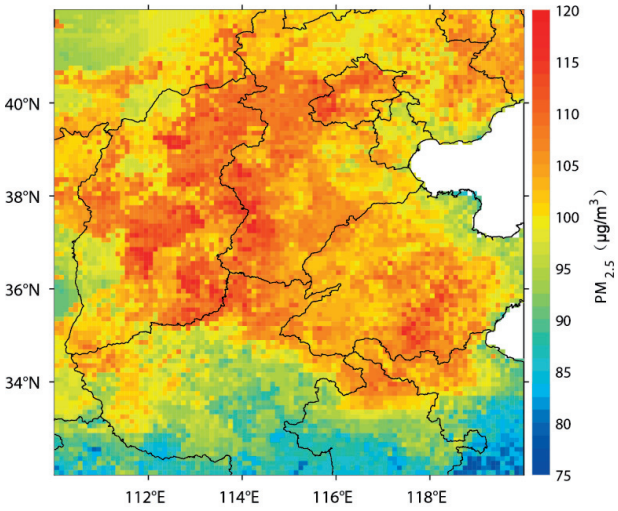

Fig. 4. Spatial distribution of daily mean estimated $PM_{2.5}$ in NCP from 2015 to 2020.

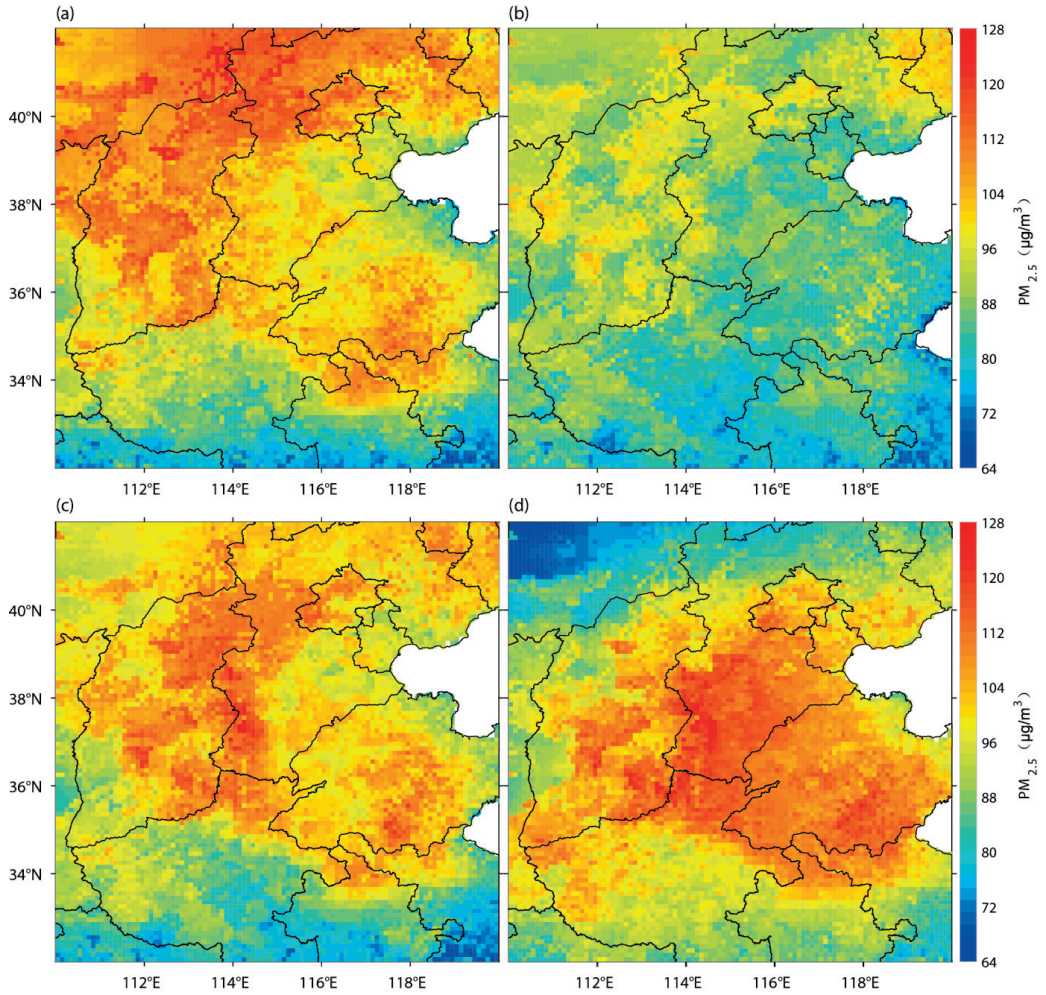

Fig. 5. Spatial distribution of mean estimated $PM_{2.5}$ in four seasons, (a-d) present the results of spring, summer, autumn and winter, respectively.

61]. And it also shows that the air cleaning plan started by the Chinese government in 2013 is more effective.

Besides, temperature also has a great influence on $PM_{2.5}$, and higher temperature will increase atmospheric turbulence, which is conducive to the diffusion of pollutants and can reduce the concentration of pollutants [62], while low temperature in the winter heating season often leads to the burning of fossil fuels and the emission of polluting gases, which has an impact on the formation of $PM_{2.5}$. Moreover, the contribution of SO_2, RH, WS, NDVI, AOD, SSRD, and WD are roughly similar with a relative importance of more than 6.9%. Among them, meteorological factors including RH, WS, SSRD, and WD affect the concentration of $PM_{2.5}$ by influencing the diffusion and deposition of particulate matter [32-35, 62-63], and these results are in consistence with previous studies [64]. Under high humidity conditions, aerosols can attach more impurities, so the $PM_{2.5}$ concentration also increases [65]. AOD reflects the concentration of near-ground pollutants [18, 25, 40], while areas with high vegetation cover have less human activity, and NDVI have a positive effect (reducing pollutant concentrations) on the deposition of $PM_{2.5}$ particles [66, 67], and then both factors contribute to $PM_{2.5}$ estimation. The relative importance of BLH, TCC and month to $PM_{2.5}$ is between 5.4% and 6.5%. In addition, the relative importance of DEM, GDP, POPU and TP for $PM_{2.5}$ is less than 4%, indicating that social factors, altitude information and precipitation are not the key factors affecting $PM_{2.5}$ concentration.

Spatial and Temporal Variation of $PM_{2.5}$

According to the $PM_{2.5}$ estimation model established above, daily gridded $PM_{2.5}$ results were calculated, and the spatial distribution of the average $PM_{2.5}$ from 2015 to 2020 is shown in Fig. 4. It is indicated that the $PM_{2.5}$ concentration in NCP is in the range of 75~120 μg m^{-3}, and most parts of this region are in heavy $PM_{2.5}$ pollution with the high $PM_{2.5}$ concentration in the range of 90~110 μg m^{-3}. The distribution of $PM_{2.5}$ concentration has obvious spatial agglomeration, and the pollution in the south of 40°N, the north of 34°N and the east side of Taihang Mountain is heavier, while the highest concentration of $PM_{2.5}$ occurs at the junction of Beijing, Hebei, Shanxi and Henan provinces. Shanxi

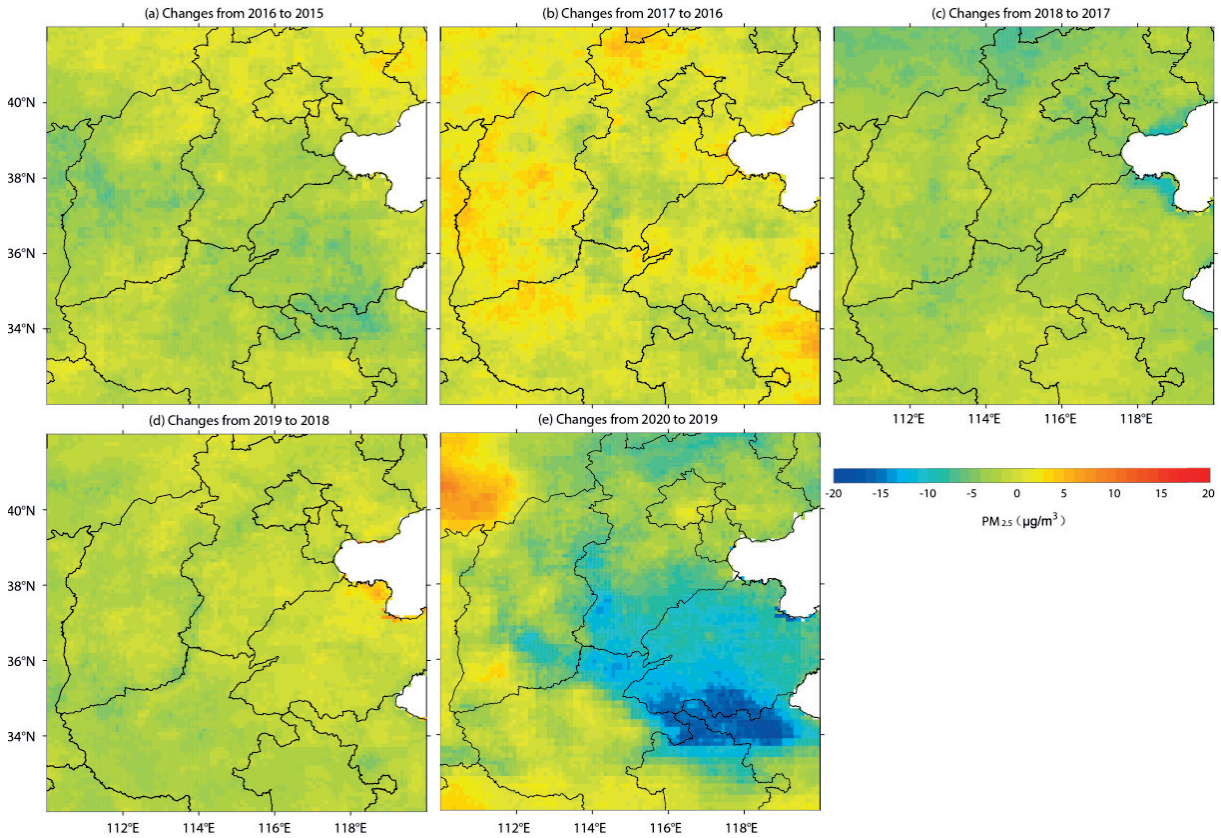

Fig. 6. Spatial distribution of $PM_{2.5}$ concentration changes in adjacent two years.

Table 3. $PM_{2.5}$ pollution standard values (k represents $PM_{2.5}$ concentration).

$PM_{2.5}$ concentration (μg m^{-3})	Grade
0<=k< = 35	Excellent
35<k< = 75	Good
75<k< = 115	Mild pollution
115<k< = 150	Middle level pollution
150<k< = 250	Heavy pollution
k>250	Serious pollution

Province, Hebei Province, and Beijing have large-scale areas with high concentrations of $PM_{2.5}$, which are related to the industrial structure of them. Shanxi Province has a large number of coal fields, and the high-value areas are exactly where the coal mines are located. Coal mining and transportation lead to elevated $PM_{2.5}$ pollution. Meanwhile, Hebei Province is dominated by heavy industry, and the first pillar industry is the iron and steel industry that emits a lot of nitrogen oxides. In addition, as the political and cultural center of China, it is not surprising that Beijing has a high level of pollution with a dense population, frequent human activities and heavy traffic.

To reveal the seasonal spatial patterns of $PM_{2.5}$, the seasonal average $PM_{2.5}$ results are shown in Fig. 5.

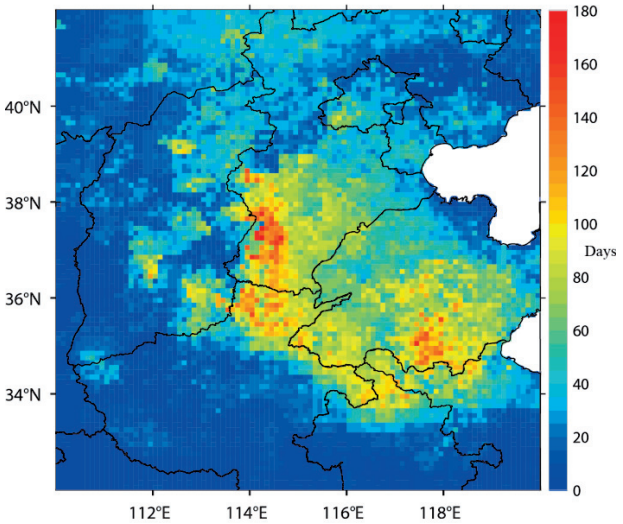

Fig. 7. Spatial distribution of $PM_{2.5}$ exceedances (daily mean $PM_{2.5}$ >150 μg m^{-3}) in each grid during 2015 and 2020 over NCP.

The results show that the $PM_{2.5}$ pollution level is the highest in winter, and the scope of pollution is also the largest. Following winter, spring and autumn are the other two seasons with the most serious $PM_{2.5}$ pollution. However, summer is the season with the lightest $PM_{2.5}$ pollution. For $PM_{2.5}$ pollution in winter, the most severely $PM_{2.5}$ pollution regions are concentrated in southern Hebei, northern Henan and southern Shandong with

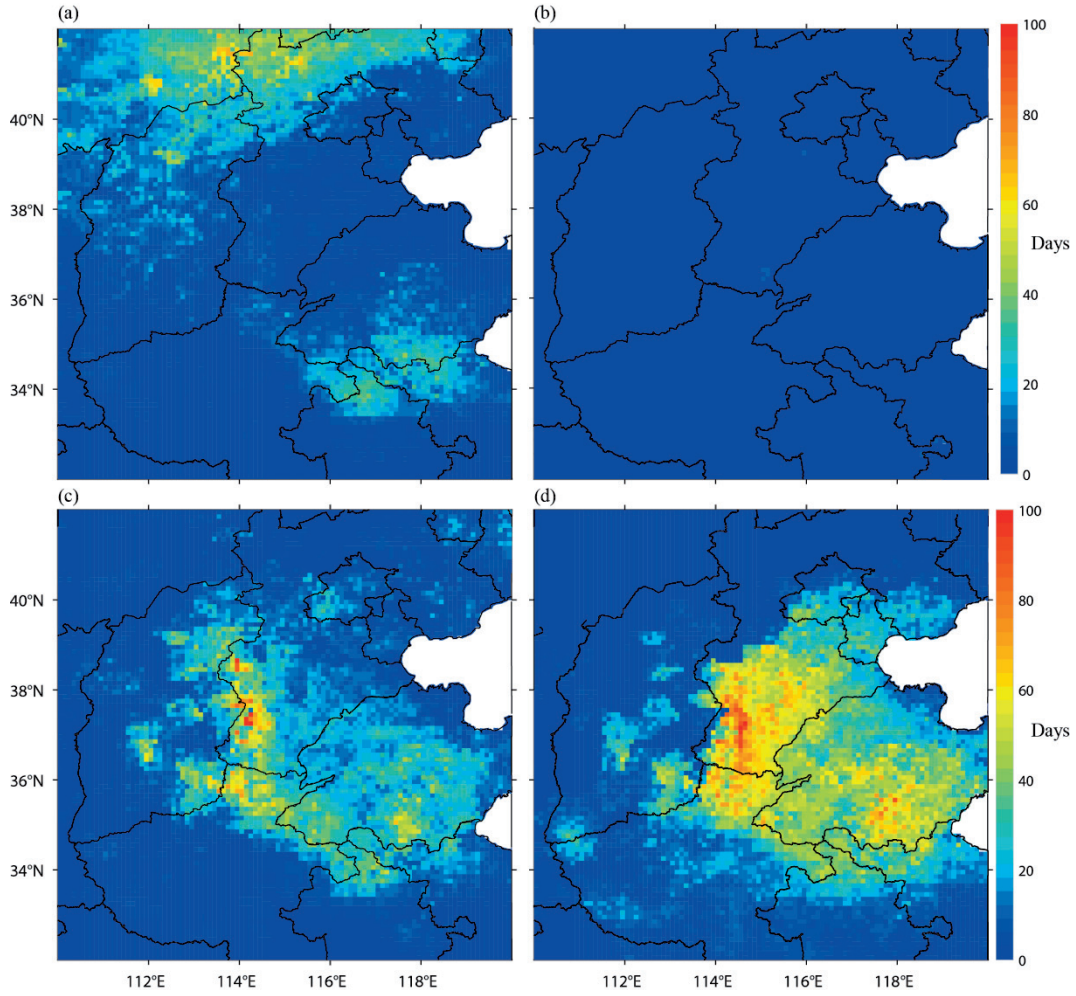

Fig. 8. $PM_{2.5}$ exceedances (daily mean $PM_{2.5}$ >150 μg m^{-3}) in each season, (a-d) present the results of spring, summer, autumn and winter, respectively.

$PM_{2.5}$ concentrations ranging in 112~128 μg m^{-3}, and it is related to winter heating in NCP, where a large amount of coal is burned, accompanied by a large amount of polluting gas emissions. Meanwhile, the temperature is low in winter and air convection slows down. These two factors make it difficult for the pollution to spread effectively, and $PM_{2.5}$ pollution increases in this region. In the terms of $PM_{2.5}$ in summer, the temperature is higher, and the air convection is accelerated, which is conducive to the diffusion of $PM_{2.5}$ pollution. Therefore, $PM_{2.5}$ pollution in summer is the lowest in four seasons with $PM_{2.5}$ concentration in most areas ranging in 64~104 μgm^{-3}. In early spring, NCP is still affected by the northwesterly wind from Mongolia-Siberia. The vegetation and loose soil in the area where the monsoon passes are sparse. Therefore, the monsoon will carry a lot of dust and be blocked by the Taihang Mountains, resulting in a high $PM_{2.5}$ concentration on the northwest side of the Taihang Mountains.

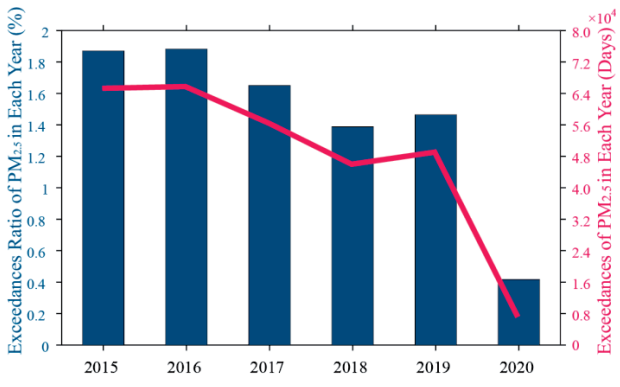

Fig. 9. Statistical results of annual cumulative exceedances ratio (left axis) and exceedances of $PM_{2.5}$ (right axis) in each year during 2015 and 2020 in NCP.

To reveal the temporal variations of $PM_{2.5}$ concentration, the spatial distribution of $PM_{2.5}$ concentration variation in adjacent two years was obtained by calculating the difference of the average $PM_{2.5}$ concentration in these two years and the result is shown in Fig. 6. From 2015 to 2020, the $PM_{2.5}$ pollution concentration generally shows a decreasing trend. Most notably, compared with 2019, the $PM_{2.5}$ pollution in NCP has decreased significantly in 2020, and the

main reason for this phenomenon may be the impact of the COVID- 19 epidemic.

Spatial and Temporal Patterns of $PM_{2.5}$ Exceedances

To better reveal the temporal and spatial variation of $PM_{2.5}$ pollution, the temporal and spatial variation characteristics of $PM_{2.5}$ pollution exceedances according to the national ambient air quality standards (GB 3095-2012) [68] were calculated and analyzed. Furthermore, the 24-hour average standard values of $PM_{2.5}$ pollution is shown in the following Table 3.

In order to focus on the situation of heavy $PM_{2.5}$ pollution, this study uses k>150 μg m^{-3} as the $PM_{2.5}$ pollution standard according to the air pollution standards in the Table 3, and the counts and frequency of $PM_{2.5}$ pollution exceedances (daily mean $PM_{2.5}$ >150 μg m^{-3}) in the study area was calculated and showed in the following parts of the work.

The number of exceedances in each grid during 2015 and 2020 over NCP was calculated and shown in Fig. 7. The results reveal the spatial distribution of the cumulative counts of $PM_{2.5}$ pollution exceedances during 2015 and 2020. As shown in the figure, the distribution of exceedances is generally in line with the daily average distribution patterns revealed in Figure 5, and the severe $PM_{2.5}$ pollution with higher exceedances mainly locate in the border of Hebei and Shanxi, the border of Hebei and Henan and southern Shandong province. Compared with the higher average concentration of $PM_{2.5}$ in eastern Shanxi, the days exceeding severe $PM_{2.5}$ pollution standards in east Shanxi is lower, which shows that most region in eastern Shanxi are moderately polluted, and $PM_{2.5}$ concentration does not exceed the standard of serious pollution (<150 μg m^{-3}). At the junction of Henan, Hebei and Shanxi provinces, $PM_{2.5}$ exceedance occurs more frequently. In terms of topography, this region with severe $PM_{2.5}$ pollution is low in terrain, surrounded by mountains in the west and north, which is not conducive to the diffusion of pollutants. From the perspective of industrial structure, Hebei Province is a major province of heavy industry, Shanxi province is a major province of coal resources, and Henan Province is a major agricultural province with emissions from the straw burning and the using of nitrogen fertilizer. In addition, the number of exceedances in southern Shandong Province is also higher, which is a well-known industrial area in China.

Fig. 8 reveals the spatial distribution of the statistical results of the number of exceedances of severe $PM_{2.5}$ pollution in each season. Summer is the season with the least exceedances of $PM_{2.5}$, and there is almost no exceedance of $PM_{2.5}$ in most regions of NCP, indicating that there is rarely heavy $PM_{2.5}$ pollution in summer. On the contrary, the number of exceedances of $PM_{2.5}$ in winter is the largest in four seasons, and the spatial coverage of the exceedances of $PM_{2.5}$ is also the largest. The highly polluted regions in winter mainly located in south Hebei, the junction of Shandong and Henan, and the exceedances in most parts of these regions are between 60-80 days. It is related to the large amount of coal burning in the cold season. In addition, the numbers of exceedances in spring and autumn are less than that in winter, and the number of exceedances is less than 60 days in most regions of NCP. Meanwhile, the number of the exceedances in spring is slightly less than that in autumn. In addition, the spatial distribution of the exceedances in the two seasons is different, and the $PM_{2.5}$ pollution in spring mainly distributes in the northern region while in autumn it mainly distributes in the southeast of NCP, and it is related to wind direction, temperature in this region. Finally, spatial distribution pattern of $PM_{2.5}$ exceedances in different seasons (Fig. 8) is roughly similar with that of average $PM_{2.5}$ concentration shown in Fig. 5.

To further reveal the levels of exceedances in different years, annual cumulative exceedances ratio and exceedances of $PM_{2.5}$ in each year during 2015 and 2020 was showed in Fig. 9. From 2015 to 2020, the severe $PM_{2.5}$ pollution generally showed a significantly downward trend with the proportion of severe $PM_{2.5}$ pollution in NCP reduced from 1.9% to 0.4%. However, the severe $PM_{2.5}$ pollution levels also experienced some repetitions between 2015 and 2020, in which the severe $PM_{2.5}$ pollution in 2019 increased compared with 2018. In addition, the proportion of severe $PM_{2.5}$ pollution in 2020 has dropped significantly compared to 2019, which may be due to the impact of the COVID-19 epidemic in NCP in the first half of 2020. The "Air Pollution Prevention and Control Action Plan" was promulgated during 2013-2017 and mainly aimed to reduce ambient $PM_{2.5}$ pollution. Previous studies reported that the Action Plan led to 59% and 21% decreases in anthropogenic SO_2 and NO_X emissions [59-61], respectively.

Conclusions

The MODIS AOD, NDVI, DEM, population density, OMI/SO_2 and OMI/NO_2 data and meteorological reanalysis data from 2015 to 2020 were used to construct the $PM_{2.5}$ estimation models with RF model and XGBoost model. After two $PM_{2.5}$ estimation models were built, the accuracy of these two models was validated and compared to reveal which model is more accurate and more suitable for establishing $PM_{2.5}$ in NCP. The accuracy of both $PM_{2.5}$ estimation models is relative high, but the accuracy of XGBoost model is higher and the verification R^2 value of randomly selected site is 0.71. Meanwhile, the relative importance of each influencing factor in $PM_{2.5}$ estimation model was revealed, and NO_2, T, SO_2, RH, WS, NDVI, AOD, SSRD and WD play an important role in dominating the $PM_{2.5}$ concentration. Among these factors, NO_2 is the primary influencing factor, which is inseparable from

the formation mechanism of $PM_{2.5}$. Subsequently, the XGBoost model was chosen to estimate daily average $PM_{2.5}$ results in NCP during 2015 and 2020. In terms of spatial-temporal distribution, $PM_{2.5}$ concentration in NCP has obvious spatial agglomeration. The areas with heavier $PM_{2.5}$ pollution are concentrated in the areas south of 40°N latitude, north of 34°N latitude and east of Taihang Mountains. The highest concentration of $PM_{2.5}$ locates in Beijing and the junction of Hebei, Shanxi and Henan. $PM_{2.5}$ concentration in NCP has a strong seasonal pattern with the highest $PM_{2.5}$ pollution level appears in winter, while the lowest $PM_{2.5}$ pollution level is in summer. The $PM_{2.5}$ pollution in NCP generally showed a downward trend, and the $PM_{2.5}$ concentration in 2020 decreased significantly compared with 2019. Overall, this work provides a high precision $PM_{2.5}$ estimation model in NCP, and daily gridded $PM_{2.5}$ results with spatial resolution of 0.1°×0.1°from 2015 to 2020 in NCP were estimated by the estimation model. Furthermore, spatiotemporal variation characteristics of $PM_{2.5}$ were also revealed in this work. In addition, the relative importance of each influencing factor to $PM_{2.5}$ pollution in NCP was quantitatively revealed. The severe $PM_{2.5}$ pollution generally showed a significantly downward trend in NCP during 2015 and 2020. And the leading roles of NO_2 and SO_2 in $PM_{2.5}$ pollution in NCP were revealed in this work. Generally, the main reasons for the decline in $PM_{2.5}$ pollution in NCP from 2015 to 2020 were due to the reduction of NO_2 and SO_2 emissions from human activities. This study reveals that the Chinese government's air pollution prevention and control plan is effective, which also provides treatment ideas and technical support for further reducing $PM_{2.5}$ pollution in China, especially in NCP.

Acknowledgments

This work was supported by the project ZR2021QD034 supported by Shandong Provincial Natural Science Foundation. The authors are grateful to China National Environmental Monitoring Center (CNEMC) for making observations of surface ozone and surface $PM_{2.5}$ (http://www.cnemc.cn), the OMI group for providing NO_2 and SO_2 data (https://disc.gsfc.nasa.gov/datasets?keywords=OMI), the MODIS group for providing AOD and NDVI data (https://ladsweb.modaps.eosdis.nasa.gov/search), the China Resource and Environmental Science Data Center for providing DEM, GDP and population density data (http://www.resdc.cn/Default.aspx), and ECMWF for providing ERA5 reanalysis (https://cds. climate. copernicus.eu/#!/search?text=ERA5).

Conflict of Interest

The authors declare no conflict of interest.

References

1. BROOK R.D., RAJAGOPALAN S., POPE C.A., BROOK J.R., BHATNAGAR A., DIEZ-ROUX A.V., HOLGUIN F., HONG Y., LUEPKER R.V., MITTLEMAN M.A., PETERS A., SISCOVICK D., SMITH S.C., WHITSEL L., KAUFMAN J.D. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the american heart association. Circulation, **121** (21), 2331, **2010**.
2. HAMRA G.B., GUHA N., COHEN A., LADEN F., RAASCHOU-NIELSEN O., SAMET J.M., VINEIS P., FORASTIERE F., SALDIVA P., YORIFUJI T., LOOMIS D. Outdoor particulate matter exposure and lung cancer: A systematic review and meta-analysis. Environmental Health Perspectives, **122** (9), 906, **2014**.
3. KIOUMOURTZOGLOU M.A., SCHWARTZ J.D., WEISSKOPF M.G., MELLY S.J., WANG Y., DOMINICI F., ZANOBETTI A. Long-term $PM_{2.5}$ exposure and neurological hospital admissions in the northeastern United States. Environmental Health Perspectives, **124** (1), 23, **2016**.
4. SON J.Y., LEE H.J., KOUTRAKIS P., BELL M.L. Pregnancy and Lifetime Exposure to Fine Particulate Matter and Infant Mortality in Massachusetts, 2001-2007. American Journal of Epidemiology, **186** (11), 1268, **2017**.
5. MAJI K.J., ARORA M., DIKSHIT A K. Premature mortality attributable to $PM_{2.5}$ exposure and future policy roadmap for 'airpocalypse' affected Asian megacities. Process Safety and Environmental Protection, **118**, 371, **2018**.
6. QI Z., ZHANG Y., CHEN Z.F., YANG C., SONG Y., LIAO X., LI W., TSANG S.Y., LIU G., CAI Z. Chemical identity and cardiovascular toxicity of hydrophobic organic components in $PM_{2.5}$. Ecotoxicology and Environmental Safety, **201** (March), 110827, **2020**.
7. LEE H.J. Advancing Exposure Assessment of $PM_{2.5}$ Using Satellite Remote Sensing: A Review. Asian Journal of Atmospheric Environment, **14** (4), 319, **2020**.
8. WATSON J.G. Visibility: Science and regulation. Journal of the Air and Waste Management Association, **52** (6), 628, **2002**.
9. LIU W., SONG X., HUANG F., HU L. Experimental study on the disintegration of granite residual soil under the combined influence of wetting-drying cycles and acid rain. Geomatics, Natural Hazards and Risk, **10** (1), 1912, **2019**.
10. TIE X., MADRONICH S., WALTERS S., EDWARDS D.P., GINOUX P., MAHOWALD N., ZHANG R., LOU C., BRASSEUR G. Assessment of the global impact of aerosols on tropospheric oxidants. Journal of Geophysical Research D: Atmospheres, **110** (3), 1, **2005**.
11. PUETT R.C., HART J.E., YANOSKY J.D., PACIOREK C., SCHWARTZ J., SUH H., SPEIZER F.E., LADEN F. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. Environmental Health Perspectives, **117** (11), 169, **2009**.
12. GUPTA P., CHRISTOPHER S.A. An evaluation of Terra-MODIS sampling for monthly and annual particulate matter air quality assessment over the Southeastern United States. Atmospheric Environment, **42** (26), 6465, **2008**.
13. LEE M., KLOOG I., CHUDNOVSKY A., LYAPUSTIN A., WANG Y., MELLY S., COULL B., KOUTRAKIS P., SCHWARTZ J. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003-2011. Journal of Exposure

Science and Environmental Epidemiology, **26** (4), 377, **2016**.
14. CHU J., DONG Y., HAN X., XIE J., XU X., XIE G. Short-term prediction of urban $PM_{2.5}$ based on a hybrid modified variational mode decomposition and support vector regression model. Environmental Science and Pollution Research, **28** (1), 56, **2021**.
15. LIN C., LI Y., YUAN Z., LAU A.K.H., LI C., FUNG J.C.H. Using satellite remote sensing data to estimate the high-resolution distribution of ground-level $PM_{2.5}$. Remote Sensing of Environment, **156**, 117, **2015**.
16. ZOU B., BENJAMIN ZHAN F., GAINES WILSON J., ZENG Y. Performance of AERMOD at different time scales. Simulation Modelling Practice and Theory, **18** (5), 612, **2010**.
17. ZOU B., CHEN J., ZHAI L., FANG X., ZHENG Z. Satellite based mapping of ground $PM_{2.5}$ concentration using generalized additive modeling. Remote Sensing, **9** (1), 1, **2017**.
18. WANG J., CHRISTOPHER S.A. Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies. Geophysical Research Letters, **30** (21), 2, **2003**.
19. GUPTA P., CHRISTOPHER S.A. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. Journal of Geophysical Research Atmospheres, **114** (14), 1, **2009**.
20. LEE H.J., LIU Y., COULL B.A., SCHWARTZ J., KOUTRAKIS P. A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations. Atmospheric Chemistry and Physics, **11** (15), 7991, **2011**.
21. ZHANG G., RUI X., FAN Y. Critical review of methods to estimate $PM_{2.5}$ concentrations within specified research region. Canadian Historical Review, **7** (9), **2018**.
22. RYBARCZYK Y., ZALAKEVICIUTE R. Machine learning approaches for outdoor air quality modelling: A systematic review. Applied Sciences (Switzerland), **8** (12), **2018**.
23. CHEN G., LI S., KNIBBS L.D., HAMM N.A.S., CAO W., LI T., GUO J., REN H., ABRAMSON M.J., GUO Y. A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological and land use information. Science of the Total Environment, **636**, 52, **2018**.
24. ZHAO C., WANG Q., BAN J., LIU Z., ZHANG Y., MA R., LI S., LI T. Estimating the daily $PM_{2.5}$ concentration in the Beijing-Tianjin-Hebei region using a random forest model with a $0.01° \times 0.01°$ spatial resolution. Environment International, **134** (May 2019), 105297, **2020**.
25. CHU D.A., KAUFMAN Y.J., ZIBORDI G., CHERN J.D., MAO J., LI C., HOLBEN B.N. Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). Journal of Geophysical Research: Atmospheres, **108** (21), 1, **2003**.
26. WANG G., SU J., CHU P.C. Mesoscale eddies in the South China Sea observed with altimeter data. Geophysical Research Letters, **30** (21), 2, **2003**.
27. LI C., KROTKOV N.A., LEONARD P.J.T., CARN S., JOINER J., SPURR R.J.D., VASILKOV A. Version 2 Ozone Monitoring Instrument SO_2 product ($OMSO_2$ V2): New anthropogenic SO_2 vertical column density dataset. Atmospheric Measurement Techniques, **13** (11), 6175, **2020**.
28. TAO J., GAO J., ZHANG L., ZHANG R., CHE H., ZHANG Z., LIN Z., JING J., CAO J., HSU S.C. $PM_{2.5}$ pollution in a megacity of Southwest China: Source apportionment and implication. Atmospheric Chemistry and Physics, **14** (16), 8679, **2014**.
29. WANG H.J., CHEN H.P. Understanding the recent trend of haze pollution in eastern China: Roles of climate change. Atmospheric Chemistry and Physics, **16** (6), 4205, **2016**.
30. GUO L., CHEN B., ZHANG H., XU G., LU L., LIN X., KONG Y., WANG F., LI Y. Improving $PM_{2.5}$ forecasting and emission estimation based on the bayesian optimization method and the coupled FLEXPART-WRF model. Atmosphere, **9** (11), 1, **2018**.
31. WU X., CHEN W., ZHANG S., LI R., ZHANG M., LIU J., JIANG Y., LIU Y. Temporal variation and chemical components of rural ambient $PM_{2.5}$ during main agricultural activity periods in the black soil region of Northeast China. Atmosphere, **10** (9), **2019**.
32. REQUIA W.J., JHUN I., COULL B.A., KOUTRAKIS P. Climate impact on ambient $PM_{2.5}$ elemental concentration in the United States: A trend analysis over the last 30 years. Environment International, **131** (February), 104888, **2019**.
33. CHEN Z., CHEN D., ZHAO C., KWAN M.P.O., CAI J., ZHUANG Y., ZHAO B., WANG X., CHEN B., YANG J., LI R., HE B., GAO B.,WANG K., XU B. Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: A review of methodology and mechanism. Environment International, **139** (February), 105558, **2020**.
34. LI M., WANG L., LIU J., GAO W., SONG T., SUN Y., LI L., LI X., WANG Y., LIU L., DAELLENBACH K.R., PAASONEN P.J., KERMINEN V.M., KULMALA M., WANG Y. Exploring the regional pollution characteristics and meteorological formation mechanism of $PM_{2.5}$ in North China during 2013-2017. Environment International, **134** (July 2019), 105283, **2020**.
35. LU J., ZHANG Y., CHEN M., WANG L., ZHAO S., PU X., CHEN X. Estimation of monthly 1 km resolution $PM_{2.5}$ concentrations using a random forest model over "2 + 26" cities, China. Urban Climate, **35** (October 2020), 100734, **2021**.
36. HAO Y., LIU Y. M. The influential factors of urban $PM_{2.5}$ concentrations in China: A spatial econometric analysis. Journal of Cleaner Production, **112**, 1443, **2016**.
37. LI C., ZHANG K., DAI Z., MA Z., LIU X. Investigation of the impact of land-use distribution on $PM_{2.5}$ in weifang: Seasonal variations. International Journal of Environmental Research and Public Health, **17** (14), 1, **2020**.
38. LIU X., ZOU B., FENG H., LIU N., ZHANG H. Anthropogenic factors of $PM_{2.5}$ distributions in China's major urban agglomerations: A spatial-temporal analysis. Journal of Cleaner Production, **264**, 121709, **2020**.
39. ZHANG W., ZHENG F., ZHANG W., YANG X. Estimating Ground-Level Hourly $PM_{2.5}$ Concentrations Over North China Plain with Deep Neural Networks. Journal of the Indian Society of Remote Sensing, **49** (8), 1839, **2021**.
40. LI L. A robust deep learning approach for spatiotemporal estimation of Satellite AOD and $PM_{2.5}$. Remote Sensing, **12** (2), 1, **2020**.
41. FU D., XIA X., WANG J., ZHANG X., LI X., LIU J. Synergy of AERONET and MODIS AOD products in the estimation of $PM_{2.5}$ concentrations in Beijing. Scientific Reports, **8** (1), 2, **2018**.
42. XU X., ZHANG C. Estimation of ground-level $PM_{2.5}$ concentration using MODIS AOD and corrected regression

model over Beijing, China. PLoS ONE, **15** (10 October), 1, **2020**.
43. BOGUMIL K., ORPHAL J., HOMANN T., VOIGT S., SPIETZ P., FLEISCHMANN O.C., VOGEL A., HARTMANN M., KROMMINGA H., BOVENSMANN H., FRERICK J., BURROWS J.P. Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: Instrument characterization and reference data for atmospheric remote-sensing in the 230-2380 nm region. Journal of Photochemistry and Photobiology A: Chemistry, **157** (2-3), 167, **2003**.
44. PIETERNEL F., LEVELT E.H., GILBERT W., LEPPELMEIER, MEMBER, IEEE, GIJSBERTUS H.J., VAN D.O., PAWAN K., BHARTIA J.T., JOHAN F., DE HAAN J.P.V. Science objectives of the ozone monitoring instrument, **44** (5), 1199, **2006**.
45. N. LAMSAL, L., A. KROTKOV N., VASILKOV A., MARCHENKO S., QIN W., FASNACHT Z., JOINER J., CHOI S., HAFFNER D., H.S.W., FISHER B., BUCSELA E. Ozone Monitoring Instrument (OMI) Aura nitrogen dioxide standard product version 4.0 with improved surface and cloud treatments. Atmospheric Measurement Techniques, **14** (1), 455, **2021**.
46. REQUIA W.J., JHUN I., COULL B.A., KOUTRAKIS P. Climate impact on ambient $PM_{2.5}$ elemental concentration in the United States: A trend analysis over the last 30 years. Environment International, **131** (February), 104888, **2019**.
47. ZHONG J., ZHANG X., WANG Y., SUN J., ZHANG Y., WANG J., TAN K., SHEN X., CHE H., ZHANG L., ZHANG Z., QI X., ZHAO H., REN S., LI Y. Relative contributions of boundary-layer meteorological factors to the explosive growth of $PM_{2.5}$ during the red-alert heavy pollution episodes in Beijing in December 2016. Journal of Meteorological Research, **31** (5), 809, **2017**.
48. MA M., BAI K., QIAO F., SHI R., GAO W. Quantifying impacts of crop residue burning in the North China Plain on summertime tropospheric ozone over East Asia. Atmospheric Environment, **194**, **2018**.
49. VU B.N., SÁNCHEZ O., BI J., XIAO Q., HANSEL N.N., CHECKLEY W., GONZALES G.F., STEENLAND K., LIU Y. Developing an advanced $PM_{2.5}$ exposure model in Lima, Peru. Remote Sensing, **11** (6), 1, **2019**.
50. MA M., YAO G., GUO J., BAI K. Distinct spatiotemporal variation patterns of surface ozone in China due to diverse influential factors. Journal of Environmental Management, **288** (September 2020), 112368, **2021**.
51. MA L., GAO Y., FU T., CHENG L., CHEN Z., LI M. Estimation of Ground $PM_{2.5}$ *Concentrations* using a DEM-assisted Information Diffusion Algorithm: A Case Study in China. Scientific Reports, **7** (1), 1, **2017**.
52. WANG H., LI J., GAO Z., YIM S.H.L., SHEN H., HO H.C., LI Z., ZENG Z., LIU C., LI Y., NING G., YANG Y. High-spatial-resolution population exposure to $PM_{2.5}$ pollution based on multi-satellite retrievals: A case study of seasonal variation in the Yangtze River Delta, China in 2013. Remote Sensing, **11** (23), **2019**.
53. WANG S., XU L., GE S., JIAO J., PAN B., SHU Y. Driving force heterogeneity of urban $PM_{2.5}$ pollution: Evidence from the Yangtze River Delta, China. Ecological Indicators, **113** (February), 106210, **2020**.
54. LEO B. Random Forests. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **45** (1), 5, **2001**.
55. SPEISER J.L., MILLER M.E., TOOZE J., IP E. A comparison of random forest variable selection methods for classification prediction modeling. Expert Systems with Applications, **134**, 93, **2019**.
56. FERNÁNDEZ-DELGADO M., CERNADAS E., BARRO S., AMORIM D. Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research, **15**, 3133, **2014**.
57. FAN J., WU L., ZHANG F., CAI H., ZENG W., WANG X. Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China. Renewable and Sustainable Energy Reviews, **100** (November 2018), 186, **2019**.
58. ZHANG W., ZHANG R., WU C., TECK A., GOH C., LACASSE S., LIU Z., LIU H. Geoscience Frontiers State-of-the-art review of soft computing applications in underground excavations. Geoscience Frontiers, **11** (4), 1095, **2020**.
59. GENG G., XIAO Q., ZHENG Y., TONG D., ZHANG Y., ZHANG X., ZHANG Q., HE K., LIU Y. Impact of China's Air Pollution Prevention and Control Action Plan on $PM_{2.5}$ chemical composition over eastern China. Science China Earth Sciences, **62** (12), 1872, **2019**.
60. ZHAI S., JACOB D.J., WANG X., SHEN L., LI K., ZHANG Y., GUI K., ZHAO T., LIAO H. Fine particulate matter ($PM_{2.5}$) trends in China, 2013-2018. separating contributions from anthropogenic emissions and meteorology. Atmospheric Chemistry and Physics, **19** (16), 11031, **2019**.
61. XIAO Q., GENG G., XUE T., LIU S., CAI C., HE K., ZHANG Q. Tracking $PM_{2.5}$ and O_3 Pollution and the Related Health Burden in China 2013-2020. Environmental Science and Technology. **2021**.
62. YANG J., JI Z., KANG S., ZHANG Q., CHEN X., LEE S.Y. Spatiotemporal variations of air pollutants in western China and their relationship to meteorological factors and emission sources. Environmental Pollution, **254**. **2019**.
63. CHEN Z., CAI J., GAO B., XU B., DAI S., HE B., XIE X. Detecting the causality influence of individual meteorological factors on local $PM_{2.5}$ concentration in the Jing-Jin-Ji region. Scientific Reports, **7** (January), 1, **2017**.
64. ZHANG H., WANG Y., HU J., YING Q., HU X. M. Relationships between meteorological parameters and criteria air pollutants in three megacities in China. Environmental Research, **140**, 242, **2015**.
65. DUO B., CUI L., WANG Z., LI R., ZHANG L., FU H., CHEN J., ZHANG H., QIONG A. Observations of atmospheric pollutants at Lhasa during 2014-2015: Pollution status and the influence of meteorological factors. Journal of environmental sciences (China), **63**, 28, **2018**.
66. BECKETT K.P., FREER-SMITH P., TAYLOR G. Effective Tree Species For Local Air-Quality Management. Journal of Arboriculture, **26** (1), 12, **2000**.
67. HWANG H.J., YOOK S.J., AHN K.H. Experimental investigation of submicron and ultrafine soot particle removal by tree leaves. Atmospheric Environment, **45** (38), 6987, **2011**.
68. MINISTRY OF ENVIRONMENTAL PROTECTION, PRC., GENERAL ADMINISTRATION OF QUALITY SUPERVISION, INSPECTION AND QUARANTINE. National Ambient Air Quality Standards. GB 3095-2012. China Environmental Science Press. 3. **2012**.