*Article*

# Estimation and Analysis of PM$_{2.5}$ Concentrations with NPP-VIIRS Nighttime Light Images: A Case Study in the Chang-Zhu-Tan Urban Agglomeration of China

Mengjie Wang [1,2,3], Yanjun Wang [1,2,3,*], Fei Teng [1,2,3], Shaochun Li [1,2,3], Yunhao Lin [1,2,3] and Hengfan Cai [1,2,3]

1   Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying, Mapping and Remote Sensing, Hunan University of Science and Technology, Xiangtan 411201, China; wangmengjie@mail.hnust.edu.cn (M.W.); tengfei@mail.hnust.edu.cn (F.T.); lsc_gis@mail.hnust.edu.cn (S.L.); linyunhao@mail.hnust.edu.cn (Y.L.); chf@mail.hnust.edu.cn (H.C.)
2   National-Local Joint Engineering Laboratory of Geo-Spatial Information Technology, Hunan University of Science and Technology, Xiangtan 411201, China
3   School of Earth Sciences and Spatial Information Engineering, Hunan University of Science and Technology, Xiangtan 411201, China
*   Correspondence: wangyanjun@hnust.edu.cn; Tel./Fax: +86-731-5829-0092

**Abstract:** Rapid economic and social development has caused serious atmospheric environmental problems. The temporal and spatial distribution characteristics of PM$_{2.5}$ concentrations have become an important research topic for sustainable social development monitoring. Based on NPP-VIIRS nighttime light images, meteorological data, and SRTM DEM data, this article builds a PM$_{2.5}$ concentration estimation model for the Chang-Zhu-Tan urban agglomeration. First, the partial least squares method is used to calculate the nighttime light radiance, meteorological elements (temperature, relative humidity, and wind speed), and topographic elements (elevation, slope, and topographic undulation) for correlation analysis. Second, we construct seasonal and annual PM$_{2.5}$ concentration estimation models, including multiple linear regression, support random forest, vector regression, Gaussian process regression, etc., with different factor sets. Finally, the accuracy of the PM$_{2.5}$ concentration estimation model that results in the Chang-Zhu-Tan urban agglomeration is analyzed, and the spatial distribution of the PM$_{2.5}$ concentration is inverted. The results show that the PM$_{2.5}$ concentration correlation of meteorological elements is the strongest, and the topographic elements are the weakest. In terms of seasonal estimation, the spring estimation results of multiple linear regression and machine learning estimation models are the worst, the winter estimation results of multiple linear regression estimation models are the best, and the annual estimation results of machine learning estimation models are the best. At the same time, the study found that there is a significant difference in the temporal and spatial distribution of PM$_{2.5}$ concentrations. The methods in this article overcome the high cost and spatial resolution limitations of traditional large-scale PM$_{2.5}$ concentration monitoring, to a certain extent, and can provide a reference for the study of PM$_{2.5}$ concentration estimation and prediction based on satellite remote sensing technology.

**Keywords:** multisource data; machine learning; PM$_{2.5}$ concentration estimation; partial least squares

## 1. Introduction

In recent years, with the rapid development of China's industrialization and urbanization, air quality problems have become increasingly intensified. In 2012, the Chinese government included PM$_{2.5}$ concentration as an important pollution source indicator in the national environmental air quality standards [1,2]. PM$_{2.5}$ can remain in the air for a long time, which will not only cause serious environmental problems, such as haze [3–8], but will also have a certain negative impact on meteorological changes, and it also has many health effects, such as premature mortality [9,10], hypertension [11], burden of disease [12,13], and

*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

2 of 18

health risks [14,15]. PM$_{2.5}$ concentration monitoring is the key to the scientific management of PM$_{2.5}$.

Traditional PM$_{2.5}$ concentration monitoring methods include the manual particle sampling weight method, micro-oscillation balance method, and β-ray absorption method. These three ground monitoring methods have high accuracy and strong real-time performance. They are relatively used to common PM$_{2.5}$ long-term monitoring methods, but the monitoring cost is too high, and the observation data from limited monitoring sites can only be used to characterize the PM$_{2.5}$ concentration in the entire area. It is difficult to accurately monitor a large-scale geographic scene. Remote sensing data can be used to monitor the geographic phenomena of continuous ground surfaces for a long time. It has been widely used in PM$_{2.5}$ concentration monitoring [16–24]. Kahn et al. [16] found that the particle size corresponding to the aerosol optical depth (AOD), obtained by the MISR inversion of the multiangle imaging spectrometer, was similar to the PM$_{2.5}$ particle size, which proved the feasibility of establishing the correlation model between AOD and PM$_{2.5}$. Li et al. [20] used satellite remote sensing parameters of AOD, fine mode fraction (FMF), planetary boundary layer height (PBLH), and atmospheric relative humidity (RH) to estimate PM$_{2.5}$ concentrations and obtain high estimation accuracy. At the same time, a series of satellite images, such as Landsat, have also been used in PM$_{2.5}$ concentration estimation [25,26]. In the above studies, satellite remote sensing technology is becoming more and more mature for daytime PM$_{2.5}$ concentration estimation. However, it is difficult to monitor changes in PM$_{2.5}$ concentration at night based on images obtained from visible light observations. At present, low-cost sensors are gradually being used in air quality monitoring. Relevant studies have shown that low-cost sensor sites with adequate monitoring conditions can provide high-quality PM$_{2.5}$ concentration data, and they can effectively monitor the temporal and spatial changes of regional PM$_{2.5}$ concentrations [27]. However, in some countries or regions, PM$_{2.5}$ air pollution is not taken seriously, so the deployment of low-cost sensors on a large scale is still a long time away for developing countries.

Nighttime light images can effectively reflect the intensity of human activities, provide more spatial details of human society, and realize the time-series monitoring of the temporal and spatial dynamic changes of human social activities. Today's nighttime light images have been widely used in socioeconomic and ecological environmental monitoring such as carbon emissions [28,29], GDP [30], poverty [31], city development [32,33], population density [34], and marine ships [28–35]. In addition to remote sensing images, commonly used for PM$_{2.5}$ concentration estimation, nighttime light data have also been used to estimate PM$_{2.5}$ concentrations at night. These nighttime light data are mainly from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLP) [36–38] and National Polar-orbiting Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) [39–42]. Wang et al. [40] used the day/night band (DNB) from radiation data of the Suomi National Polar-orbiting Partnership (S-NPP) satellite's visible infrared imaging radiometer suite (VIIRS) to estimate PM$_{2.5}$ concentration, and they found that nighttime light images can provide a good inversion of PM$_{2.5}$ concentrations. The correlation coefficient R, between the estimated PM$_{2.5}$ concentration and the measured PM$_{2.5}$ concentration, is 0.67. Fu et al. [41] used data from the Day/Night Band (DNB) of the Visible Infrared Imaging Radiometer Suite (VIIRS) and hourly PM$_{2.5}$ data, at 35 stations in Beijing, to develop a mixed-effects model to estimate nighttime PM$_{2.5}$ concentrations. The results of cross-validation showed that the estimation accuracy of PM$_{2.5}$ concentration in the four seasons was high, and the R$^2$ of the model was greater than 0.80. Xu et al. [37] explored the influence of meteorological and social factors on PM$_{2.5}$ concentrations, and their results showed that the nighttime light index was one of the main influencing factors of PM$_{2.5}$ concentrations. Zhang et al. [42] combined meteorological data and satellite observation data, such as Luojia (LJ) 1-01 nighttime light images, to build a PM$_{2.5}$ concentration estimation model. The LJ1-01 satellite is the first dedicated nighttime light remote sensing satellite in the world, and it launched in July 2018. The results showed that adding nighttime light image information can improve the performance of PM$_{2.5}$ prediction models. The spatiotemporal distribution of PM$_{2.5}$

*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

3 of 18

concentration is a complex geographical phenomenon affected by multiple factors. It is difficult to explore the spatiotemporal relationship between nighttime light images and $PM_{2.5}$ concentration from a smaller time scale. Long-term $PM_{2.5}$ concentration estimation is an important part of air quality monitoring, but few studies have applied nighttime light images to seasonal and annual $PM_{2.5}$ concentration estimations. The relationship between nighttime light images and long-term $PM_{2.5}$ concentration temporal and spatial changes remains to be further studied.

The temporal and spatial distribution of $PM_{2.5}$ concentration is a complex geographical phenomenon. Topographic and meteorological factors are important influencing factors for the temporal and spatial distribution of $PM_{2.5}$ concentration [43]. Meteorological factors mainly depend on meteorological conditions, such as wind, precipitation, and temperature, to affect the regional $PM_{2.5}$ concentration [44]. Wind acts on the temporal and spatial distribution of $PM_{2.5}$ concentration by affecting air diffusion. Precipitation increases humidity and causes $PM_{2.5}$ particles to clump together, unable to stay in the air, and fall to the ground. Changes in air temperature will affect the characteristics of atmospheric flow and, thus, the diffusion of $PM_{2.5}$. Although topographic factors have less influence on the temporal and spatial distribution of $PM_{2.5}$ concentration than meteorological factors [43], topographical factors such as altitude and slope affect the changes of $PM_{2.5}$ concentration by changing the flow characteristics of air [45].

This paper analyzes the ability of nighttime light images to estimate seasonal and annual $PM_{2.5}$ concentrations. This paper uses partial least squares to analyze the correlation between meteorological elements, terrain elements, nighttime light radiance, and $PM_{2.5}$ concentration. Then, a multivariate linear and machine learning regression model for $PM_{2.5}$ concentration estimation in the Chang-Zhu-Tan urban agglomeration was constructed, combined with the ground monitoring station data, to evaluate the accuracy of the model results, and finally, the spatial continuous distribution of $PM_{2.5}$ concentration in the Chang-Zhu-Tan urban agglomeration was inverted.

## 2. Study Areas and Data Sources

### 2.1. Study Areas

The Chang-Zhu-Tan urban agglomeration is located in the middle-eastern part of Hunan Province (Figure 1). It has a mid-subtropical monsoon climate with four distinct seasons, short winters, long summers, and abundant rainfall. As the core growth pole of economic development in Hunan Province, the Chang-Zhu-Tan urban agglomeration industry has achieved rapid development in recent years [46]. At the same time, the problem of air pollution has become increasingly prominent, and the concentration of various air pollutants in the urban agglomeration remains high [47,48]. The air quality level ranks last in the province year round. Regional air pollution seriously affects public health and ecological safety, and the serious haze problem has also attracted great attention from all walks of life [49]. In recent years, the relevant air pollution control measures of the Chinese government have resulted in a significant decrease in the $PM_{2.5}$ concentration in the Chang-Zhu-Tan urban agglomeration, effectively improving the air quality of the urban agglomeration [50].
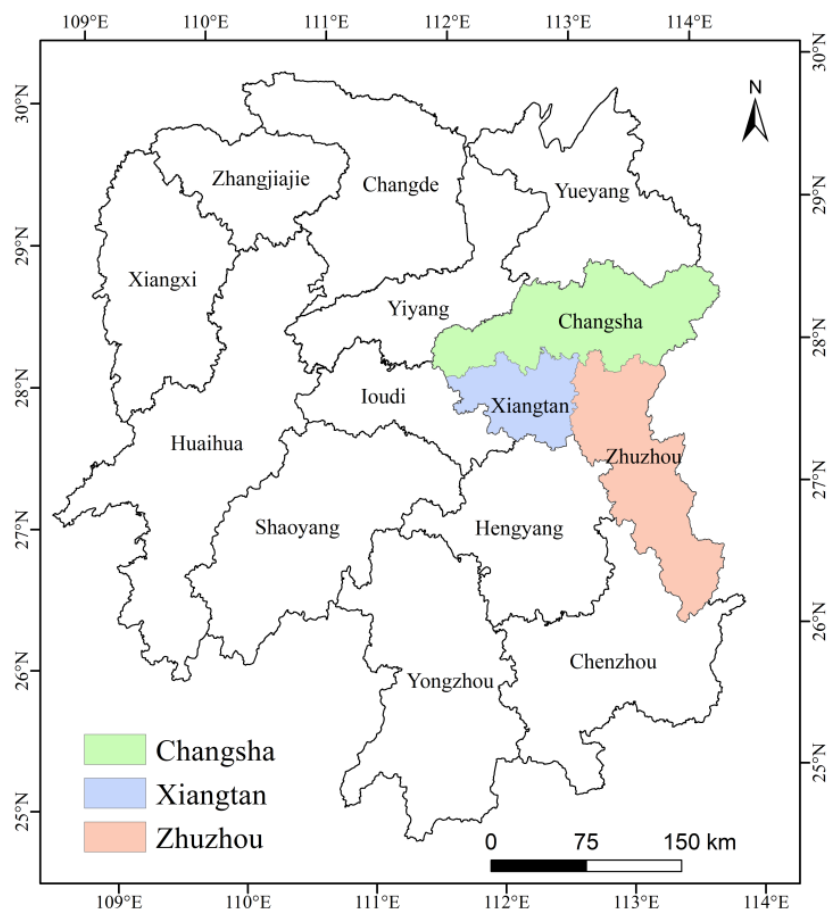
**Figure 1.** Location of the Chang-Zhu-Tan urban agglomeration.

### 2.2. Data Sources

The data for this research include PM$_{2.5}$ concentration data, meteorological data, NPP-VIIRS nighttime light images, and Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) data in the Chang-Zhu-Tan urban agglomeration in 2015 and 2018.

PM$_{2.5}$ concentration data: The PM$_{2.5}$ concentration data used in this article came from the national urban air quality real-time release platform of the China Environmental Monitoring Station (CEMS. http://106.37.208.233:20035/ (accessed on 15 October 2019)). The quarterly and annual average PM$_{2.5}$ concentrations were derived from the hourly monitoring data of 24 ambient air quality assessment monitoring points in the Chang-Zhu-Tan urban agglomeration (Figure 2a). In order to ensure the accuracy, continuity, and integrity of PM$_{2.5}$ concentration measurement data, the Chinese government stipulates that, when automatic monitoring equipment is used for monitoring, the monitoring equipment needs to run continuously, 365 days a year. The daily average of PM$_{2.5}$ concentration measurements requires at least 20 h of average concentration values or adoption time. The PM$_{2.5}$ concentration measurement data in this paper are obtained by the continuous automatic monitoring method. The Chinese government stipulates that the PM$_{2.5}$ automatic monitoring method with different principles can only be used to measure PM$_{2.5}$ if it is consistent with the monitoring results of the manual gravimetric method. Therefore, the PM$_{2.5}$ concentration measurement values used in this paper are subject to strict quality control and are effective.
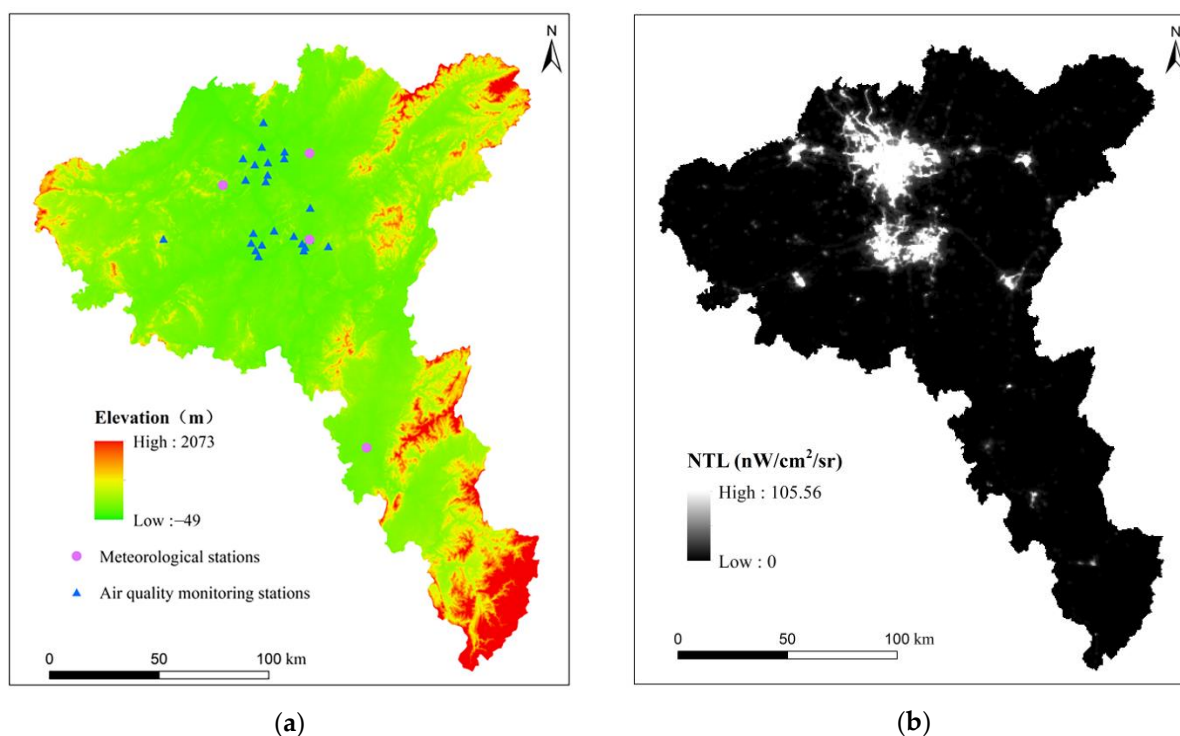
*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

5 of 18



**Figure 2.** Datasets used in this study. (**a**) SRTM DEM data and spatial distribution of the monitoring stations; (**b**) NPP-VIIRS nighttime light (NTL) image.

Meteorological data: The meteorological data came from the National Meteorological Science Data Sharing Service Platform (NMSDSSP. http://data.cma.cn. (accessed on 15 October 2019)) and mainly include precipitation, temperature, relative humidity, and wind speed. The quarterly and annual average weather data came from the daily average values of meteorological stations in the Chang-Zhu-Tan urban agglomeration (Figure 2a). Meteorological factors have a great impact on the spatial distribution of PM$_{2.5}$ in the Chang-Zhu-Tan urban agglomeration [47]. The meteorological information of the air quality monitoring stations comes from four ground meteorological stations. Since the air quality monitoring stations are distributed in the plain area and are concentrated near the four ground meteorological stations, the uniformity of meteorological factors in a small range is considered [51]. Therefore, it is feasible that the meteorological information of the air quality monitoring station comes from four ground meteorological stations in this study.

NPP-VIIRS nighttime light images were obtained from the Earth Observation Group (EOG). This article used the monthly data from NPP-VIIRS nighttime light images in 2015 and 2018, with a resolution of 500 m (Figure 2b). The monthly nighttime light image was composed of the cloudless nighttime light image of the month, which was the average radiation image. The monthly nighttime light images were also processed with stray light correction. The processed monthly NPP-VIIRS nighttime light images can effectively monitor the status quo of regional socioeconomic development [52–54]. Nighttime light images can effectively reflect the development status of human society and provide more spatial details of human activities [55,56].

SRTM DEM data: The DEM data of the experimental area came from the SRTM data of the U.S. Space Shuttle Endeavour. This dataset was based on the latest SRTM V4.1 data, through collation and splicing, to generate 90 m resolution DEM data (Figure 2a). Topography not only affects the spatial distribution of pollutant emissions by affecting the intensity of human activities but also has a profound impact on the diffusion of PM$_{2.5}$, which is an important factor affecting the spatial distribution of PM$_{2.5}$ [43,57].

### 3. Methods

*3.1. Correlation Analysis between Remote Sensing Data and PM$_{2.5}$ Concentration*

Based on the theory of radiative transmission, the relationship model between nighttime light radiance and PM$_{2.5}$ concentration in the near-surface layer can be constructed [40]. First, it is assumed that there is no change in the distribution of surface features (especially buildings and city lights) around the ground air quality monitoring site. Then, there is the nighttime light radiance, after reflection/scattering by various physical media from lights emitting upwards, from what is considered a Lambertian body, which is a constant with spatial differences [40]. Assuming negligible multiple scattering from aerosols, the nighttime light radiance reaching the sensor follows Beer's law. Assuming that there is a good and stable aerosol extinction coefficient profile structure in the boundary layer at night, and PM$_{2.5}$ is uniformly mixed at the effective height, the relationship between PM$_{2.5}$ and nighttime light radiance can be established [40]. In this paper, the average value of nighttime light, 2 km around the environmental detection site, was extracted as its nighttime light radiance value.

Meteorological elements are important factors influencing the changes in PM$_{2.5}$ concentration [44,58–60]. Wang et al. [58] discussed whether meteorological elements can affect PM$_{2.5}$ concentrations and found that meteorological elements, such as humidity and air temperature, can affect the temporal and spatial distributions of PM$_{2.5}$ concentrations. In addition, topographic elements affect the change in regional PM$_{2.5}$ concentration to a certain extent [43,45,57]. He et al. [45] added the information extracted from DEM data to the PM$_{2.5}$ estimation model, and the results showed that the model with topography, meteorology, and other elements can better estimate PM$_{2.5}$ concentrations. Therefore, the PM$_{2.5}$ concentration estimation model that takes into account the influence of multiple factors, such as weather and topography, at the same time can obtain higher-precision PM$_{2.5}$ concentration simulation results. Therefore, the characteristic factors determined in this paper include nighttime light radiance I, elevation E, slope S, precipitation R, temperature T, relative humidity RHU, and wind speed W.

*3.2. Selection of Characteristic Factors for the PM$_{2.5}$ Concentration Estimation Model*

The correlation analysis was carried out by constructing a partial least squares model of Factor Set A and PM$_{2.5}$ concentration. The partial least squares method uses the algorithm of decomposing and screening the data information in the model, extracts the comprehensive variable with the strongest explanatory power for the dependent variable, and can calculate the importance of each factor. The partial least squares method can better solve the factor collinearity problem and obtain more objective and accurate factor importance results [61]. The variable importance in projection (VIP) value of partial least squares is used as the factor importance result [62], and the VIP value calculation formula is as follows:

$$VIP_j = \sqrt{\frac{p \sum_{k=1}^{h} \left( c_k^2 t_k' t_k \right) w_{jk}^2}{\sum_{k=1}^{h} c_k^2 t_k' t_k}} \tag{1}$$

where: $VIP_j$ is the VIP value of the j-th variable; $p$ is the number of variables participating in the analysis; $h$ is the number of iteration calculations; $c_k^2 t_k' t_k$ is the interpretation of the dependent variable from the *k*-th independent variable mapping result interpretation degree; $w_{jk}^2$ is the weight of variable *j* in the *k*-th iteration.

*3.3. Construction of the PM$_{2.5}$ Concentration Estimation Model*

Simple models have limitations in simulating complex geographic phenomena, with multiple factors, at high precision [63]. Zhang et al. [63] found that simple models cannot effectively estimate the spatial distribution of PM$_{2.5}$ concentrations affected by multiple factors. In this paper, referring to the research results of Wang et al. [40], a multiple linear regression model was selected to construct the PM$_{2.5}$ concentration estimation Model I of

*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

7 of 18

the Chang-Zhu-Tan urban agglomeration. There are 24 air quality monitoring stations in the Chang-Zhu-Tan urban agglomeration.

$$PM_{2.5} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \qquad (2)$$

where: $PM_{2.5}$ is the estimated $PM_{2.5}$ concentration of the air quality monitoring site; $X_1$, $X_2$, and $X_n$ are the 1st, 2nd$\cdots$nth estimated model factors, respectively; $\beta_1$, $\beta_2$, and $\beta_n$ are the regression coefficients of each model, respectively.

When there is no definite estimation method of $PM_{2.5}$ concentration, the application of machine learning can extract key feature information to find the relationship between known datasets, and the machine model trained with a large amount of data can be used for accurate prediction. Machine learning methods have been increasingly used in socioeconomic parameter estimation and geographic phenomenon inversion, and there have also been related studies using machine learning methods for $PM_{2.5}$ concentration estimation. Among them, there are many studies on the use of random forest models for $PM_{2.5}$ concentration estimation [64–66], and other machine learning models are gradually applied to $PM_{2.5}$ concentration estimation [67,68]. Based on the $PM_{2.5}$ concentration data from ground stations and the known data of nighttime light radiance I, elevation E, slope S, precipitation R, temperature T, relative humidity RHU, and wind speed W, three machine learning $PM_{2.5}$ concentration estimation models were constructed in this paper: random forest Model II, support vector machine Model III, and Gaussian process regression Model IV. These three models are more commonly used and more mature machine learning regression models. Each of them has some advantages. For example, support vector machines can solve machine learning problems with small samples and can find the nonlinear relationship between variables well. For unbalanced data sets, ensemble trees can balance errors to a certain extent. Gaussian process regression can quantify the prediction uncertainty in a principled way.

In this paper, the three machine learning estimation models were trained with multiple samples, and the fivefold cross-validation method was used to test model accuracy. Finally, the model parameters, when the goodness of fit ($R^2$) of the model is the highest, are determined. According to the training results, the important parameters of the machine learning model with the highest $R^2$ are selected (see Table 1). Among them, the parameter of random forest Model II is the minimum leaf size, and the parameters of Model III support vector machine and Model IV Gaussian process regression are the kernel function.

**Table 1.** Important parameters of the various $PM_{2.5}$ concentration estimation models based on machine learning.

| Model Parameters | Spring | Summer | Autumn | Winter | Annual |
|---|---|---|---|---|---|
| Model II smallest leaf | 12 | 4 | 12 | 12 | 12 |
| Model III kernel function | Linear | Linear | Linear | Linear | Quadratic |
| Model IV kernel function | Exponential | Exponential | Matern 5/2 | Exponential | Matern 5/2 |
| Model parameters | Spring | Summer | Autumn | Winter | Annual |

## 4. Results

### 4.1. Importance Analysis of PM2.5 Concentration Estimation Model Factors

To explore the influence of characteristic factors on the model estimation results, nighttime light radiance, elevation, slope, precipitation, air temperature, relative humidity, and wind speed were selected as Factor Set A. In addition, the more relevant feature factors from the Factor Set A were selected as Factor Set B. Finally, the precipitation, temperature, relative humidity, and wind speed of commonly used meteorological elements were selected from Factor Set A as Factor Set C.

In this paper, the partial least squares method was used to analyze the importance of the model factors. The VIP score of each factor obtained by the formula (1) determines the correlation between the factor and the $PM_{2.5}$ concentration. The results showed that (Table 2)

four meteorological factors (air temperature T, relative humidity RHU, precipitation R, and wind speed W) had high VIP scores. The mean VIP scores of quarterly and annual were 1.552, 0.795, 0.835, and 1.100, respectively. The air temperature T factor is the most important factor affecting the temporal and spatial distribution of $PM_{2.5}$ concentration. There was a high correlation between nighttime light radiance I and $PM_{2.5}$ concentration, with an average VIP score of 0.504. The topographic factors (elevation E and slope S) had a low correlation with the $PM_{2.5}$ concentration, with average VIP scores of 0.320 and 0.304, respectively. Therefore, this paper selected temperature T, relative humidity RHU, precipitation R, wind speed W, and nighttime light radiance I as Factor Set B.

**Table 2.** VIP scores of different factors for the $PM_{2.5}$ concentration estimation.

| Factor | Spring | Summer | Autumn | Winter | Annual |
|:------:|:------:|:------:|:------:|:------:|:------:|
| *I* | 1.138 | 0.464 | 0.366 | 0.302 | 0.249 |
| *T* | 1.381 | 1.507 | 1.658 | 1.465 | 1.748 |
| *RHU* | 1.157 | 1.449 | 0.530 | 0.662 | 0.178 |
| *W* | 0.508 | 0.985 | 0.979 | 0.986 | 0.719 |
| *R* | 0.943 | 0.526 | 0.742 | 1.384 | 1.907 |
| *E* | 0.414 | 0.257 | 0.322 | 0.442 | 0.164 |
| *S* | 0.723 | 0.249 | 0.175 | 0.283 | 0.091 |

*4.2. The Results and Accuracy Evaluation of the $PM_{2.5}$ Concentration Estimation Model for the Chang-Zhu-Tan Urban Agglomeration*

Based on the multiple linear regression model and three machine learning regression models, combined with the environmental monitoring site data of the Chang-Zhu-Tan urban agglomeration, model verification was carried out for the four seasons as well as annually (see Tables 3 and 4).

Since the temporal and spatial distribution of $PM_{2.5}$ concentration is a complex geographical phenomenon, the variation law of $PM_{2.5}$ concentration, under the action of multiple factors, may be different in different time periods. Therefore, this paper considers selecting a variety of models to analyze the relationship between $PM_{2.5}$ concentration and factors, in order to improve the estimation accuracy of $PM_{2.5}$ concentration. The results showed that there were obvious differences in the estimation results of $PM_{2.5}$ concentration models in different seasons, among which the $PM_{2.5}$ concentration estimation model in spring had the worst results, and the $R^2$ value was significantly lower than those from the other three seasonal and annual estimation models. The annual estimation model had the best effect, followed by the winter, summer, and autumn estimation models, which had similar effects.

There were also obvious differences in the estimation effects of different models. The multiple linear regression models had better estimation results for the seasonal $PM_{2.5}$ concentration, while the machine learning model had better estimation results for the annual $PM_{2.5}$ concentration. The number of sample points for the construction of seasonal and annual $PM_{2.5}$ concentration estimation models was different. The number of sample points for seasonal $PM_{2.5}$ concentration was small, only one-fourth of the number of annual $PM_{2.5}$ concentration sample points, resulting in opposite results in the season and year for $PM_{2.5}$ concentration estimation accuracy based on multivariate linear and machine learning models.

The effect of the estimation model of Factor Set B was obviously better than that of Factor Set C, indicating that adding nighttime light image information can effectively improve the performance of the estimation model. In addition, the estimation model effect of Factor Set A was better than that of Factor Set B, which also shows that adding topographic information can also effectively improve the model estimation ability.

*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

9 of 18

**Table 3.** $R^2$ values of the $PM_{2.5}$ concentration estimation model in the Chang-Zhu-Tan urban agglomeration.

| Model | Factor Set | Spring | Summer | Autumn | Winter | Annual |
|---|---|---|---|---|---|---|
| Model I | Factor set A | 0.36 | 0.81 | 0.76 | 0.89 | 0.82 |
| | Factor set B | 0.31 | 0.79 | 0.75 | 0.88 | 0.82 |
| | Factor set C | 0.25 | 0.78 | 0.75 | 0.85 | 0.82 |
| Model II | Factor set A | 0.17 | 0.65 | 0.72 | 0.79 | 0.90 |
| | Factor set B | 0.16 | 0.66 | 0.72 | 0.80 | 0.92 |
| | Factor set C | 0.07 | 0.71 | 0.67 | 0.80 | 0.91 |
| Model III | Factor set A | 0.23 | 0.69 | 0.69 | 0.77 | 0.88 |
| | Factor set B | 0.20 | 0.55 | 0.66 | 0.75 | 0.90 |
| | Factor set C | 0.13 | 0.67 | 0.69 | 0.73 | 0.90 |
| Model IV | Factor set A | 0.08 | 0.64 | 0.54 | 0.73 | 0.89 |
| | Factor set B | 0.07 | 0.63 | 0.64 | 0.72 | 0.90 |
| | Factor set C | 0.06 | 0.67 | 0.63 | 0.72 | 0.92 |

**Table 4.** Root mean square errors of the $PM_{2.5}$ concentration estimation model in the Chang-Zhu-Tan urban agglomeration.

| Model | Factor Set | Spring | Summer | Autumn | Winter | Annual |
|---|---|---|---|---|---|---|
| Model I | Factor set A | 4.48 | 3.74 | 6.06 | 7.75 | 11.80 |
| | Factor set B | 4.64 | 3.88 | 6.11 | 8.11 | 11.85 |
| | Factor set C | 4.85 | 3.94 | 6.15 | 8.91 | 11.90 |
| Model II | Factor set A | 5.14 | 5.12 | 6.79 | 11.06 | 8.65 |
| | Factor set B | 5.19 | 5.49 | 6.72 | 10.40 | 7.73 |
| | Factor set C | 5.50 | 4.64 | 7.10 | 10.69 | 8.25 |
| Model III | Factor set A | 4.94 | 4.76 | 7.19 | 11.58 | 9.85 |
| | Factor set B | 5.05 | 6.30 | 7.30 | 11.77 | 8.73 |
| | Factor set C | 5.34 | 4.98 | 6.90 | 12.45 | 8.75 |
| Model IV | Factor set A | 5.40 | 5.14 | 8.71 | 12.68 | 9.22 |
| | Factor set B | 5.44 | 5.69 | 7.57 | 12.30 | 8.67 |
| | Factor set C | 5.54 | 4.92 | 7.54 | 12.67 | 8.14 |

At the same time, this paper established a scatter diagram between the annual estimated and actual $PM_{2.5}$ concentrations (Figure 3). The results showed that there was a high correlation between the two, in which the $R^2$ values in 2015 and 2018 were 0.87 and 0.92, respectively, indicating that there were good estimation results for the $PM_{2.5}$ concentration.
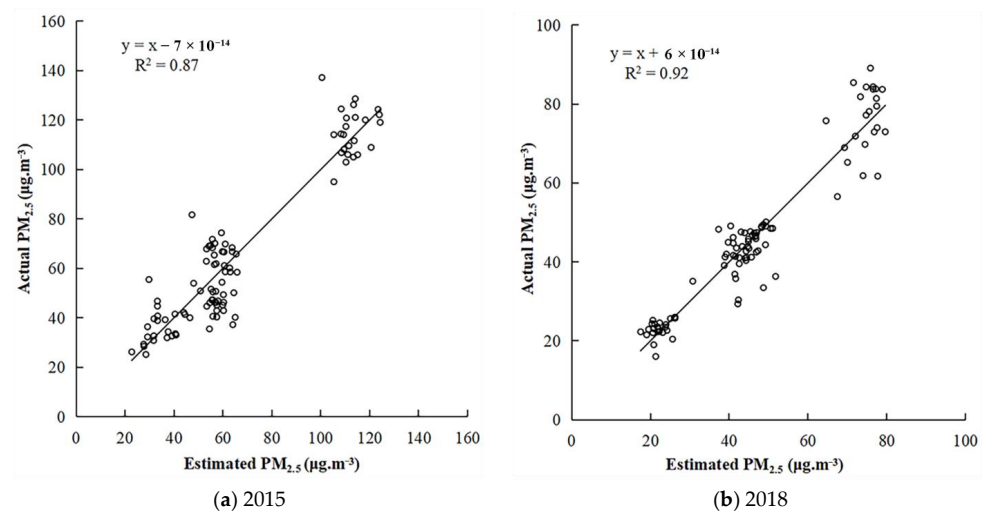


**Figure 3.** Scatter plots of estimated and actual $PM_{2.5}$ concentrations.

Int. J. Environ. Res. Public Health **2022**, 19, 4306

10 of 18

*4.3. Spatial Analysis of the PM$_{2.5}$ Concentration in the Chang-Zhu-Tan Urban Agglomeration*

In this paper, kriging interpolation analysis was performed on the seasonal PM$_{2.5}$ concentration of the Chang-Zhu-Tan urban agglomeration in 2018, and the continuous spatial interpolation of PM$_{2.5}$ concentration was realized. The results are shown in Figure 4. According to the inversion results, the temporal and spatial distributions of seasonal PM$_{2.5}$ concentrations in the Chang-Zhu-Tan urban agglomeration were analyzed. The results showed that the PM$_{2.5}$ concentration of the Chang-Zhu-Tan urban agglomeration in winter was significantly higher than that in the other three seasons, with the lowest PM$_{2.5}$ concentration in summer and similar PM$_{2.5}$ concentrations in spring and autumn.
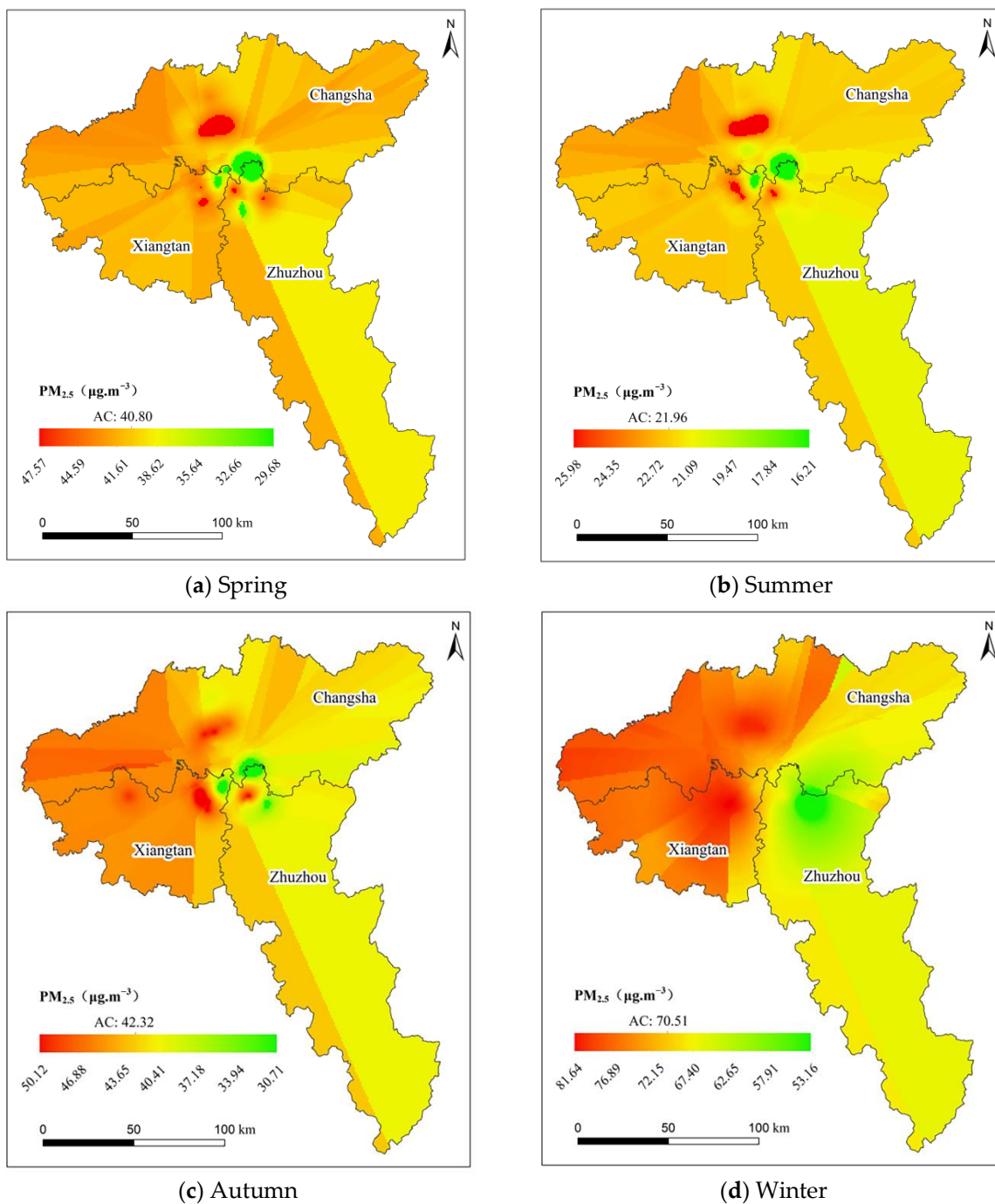


(**a**) Spring

(**b**) Summer

(**c**) Autumn

(**d**) Winter

**Figure 4.** Inversion of seasonal PM$_{2.5}$ concentration in 2018 in the Chang Zhu Tan urban agglomeration. AC means average PM$_{2.5}$ concentration.

The study area is located in the subtropical monsoon region. The northerly wind prevails in the Chang-Zhu-Tan urban agglomeration in winter, the atmospheric structure is stable, and the meteorological conditions are not conducive to the diffusion of $PM_{2.5}$ and other particles. The study area is prone to temperature inversion in winter, which makes $PM_{2.5}$ particles gradually accumulate on the surface. In addition, the burning of a large amount of coal for heating in winter increases the $PM_{2.5}$ concentration.

In summer, the southerly wind prevails, and the meteorological conditions are conducive to the diffusion of $PM_{2.5}$ and other particles. In summer, strong winds are more likely to lead to the diffusion of $PM_{2.5}$. In addition, it is rainy and humid in summer, and it is difficult for $PM_{2.5}$ particles to stay in the air. The high temperature in summer makes it less likely for temperature inversion to occur, and the atmosphere is prone to convection, which is conducive to the diffusion of $PM_{2.5}$ particles. Therefore, the concentration of $PM_{2.5}$ is relatively high in winter and low in summer. At the same time, there are differences in the spatial distribution of $PM_{2.5}$ concentrations. The $PM_{2.5}$ concentration in the northwestern part of the Chang-Zhu-Tan urban agglomeration is relatively high, and the $PM_{2.5}$ concentration in some central areas is low, which is significantly different from the adjacent areas.

## 5. Discussion

With the rapid development of industry and the increasing number of vehicles, the problem of air pollution is becoming increasingly serious [69]. Monitoring the spatial and temporal distribution of polluted gases is the key to solving the problem of air pollution. Among them, $PM_{2.5}$ has always been one of the main air pollutants monitored by humans. At present, the model used by daytime remote sensing satellite technology for $PM_{2.5}$ concentration estimation is relatively mature, and it can better perform spatial processing of large-scale $PM_{2.5}$ concentrations. Human production and living activities greatly affect the temporal and spatial distributions of $PM_{2.5}$ concentrations. Human social activities at night can reflect the intensity of human activities and reflect the state of human production, and living, to a certain extent. Therefore, this paper added nighttime light image information to $PM_{2.5}$ concentrations. In the concentration estimation model, the results showed that the accuracy of the $PM_{2.5}$ concentration estimation results has been somewhat improved, indicating that nighttime light images are of practical significance for $PM_{2.5}$ concentration estimation.

In this paper, the partial least squares method was used to calculate the factor importance of the $PM_{2.5}$ concentration. The partial least squares method can better solve the multicollinearity problem on the basis of retaining all factors, and the partial least squares method extracts, as much as possible, real $PM_{2.5}$ concentration-related factor information to obtain a more objective and reliable correlation between factors and $PM_{2.5}$ concentration. Compared with other factor analysis methods, the partial least squares method can calculate factor VIP scores on the basis of more effectively solving the multicollinearity problem.

In this paper, the multivariate linear model was used to obtain the estimated value of the seasonal $PM_{2.5}$ concentration, and scatter plots (Figure 5) of the estimated value and the actual value of the $PM_{2.5}$ concentration in the four seasons were constructed. The results showed that the estimated value and the actual value of the $PM_{2.5}$ concentration in the four seasons was very close to y = x, indicating that the error distribution of the model, underestimating and overestimating $PM_{2.5}$ concentration, was relatively balanced. The estimated $R^2$ value of the $PM_{2.5}$ concentration model in spring was significantly lower than that in the other three seasons, while the estimated $R^2$ value of the $PM_{2.5}$ concentration model in winter was significantly higher than that in the other three seasons, indicating that the model estimation accuracy had seasonality.
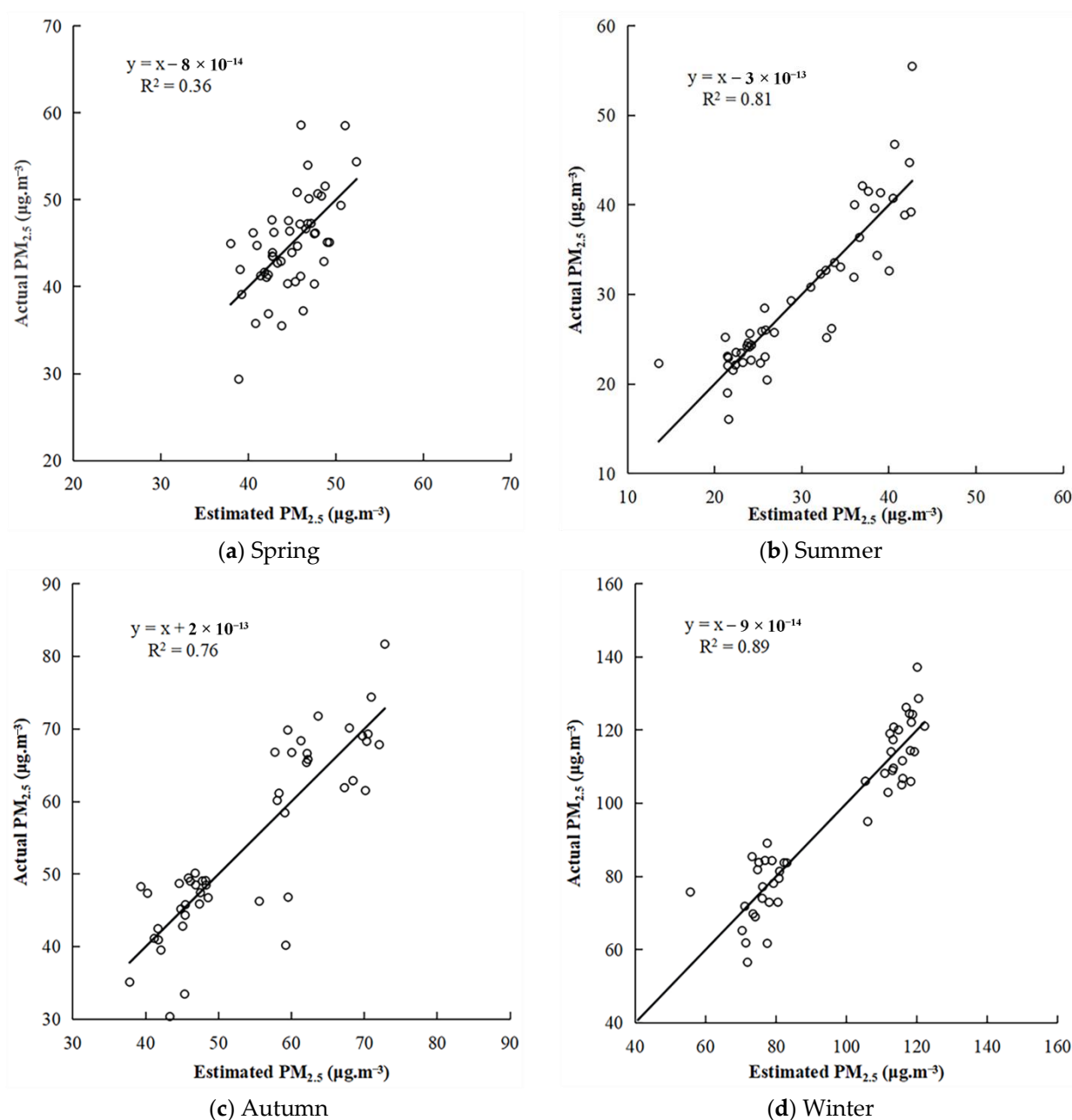
**(a)** Spring



**(b)** Summer



**(c)** Autumn



**(d)** Winter

**Figure 5.** Scatter plots of estimated and actual $PM_{2.5}$ concentrations in the four seasons.

In addition, the spatial distribution of $PM_{2.5}$ concentration is a complex geographic phenomenon, and the spatial characteristics of different air quality monitoring stations are different, resulting in obvious spatial differences in the accuracy of $PM_{2.5}$ concentration estimation models. In this paper, the multivariate linear estimation model, with high estimation accuracy of seasonal $PM_{2.5}$ concentration, was used to obtain the estimated $PM_{2.5}$ concentration in the four seasons, and the estimated $PM_{2.5}$ concentration in the four seasons was compared with the actual value (Figure 6). The results showed that the estimated and actual $PM_{2.5}$ concentrations in the four seasons had similar trends, indicating that the overall effect of the model estimation was good, but there were still obvious local differences. The estimated value of the $PM_{2.5}$ concentration, at some stations, was quite different from the actual value. To further analyze the spatial difference in model estimation accuracy, this paper also analyzed the actual error of $PM_{2.5}$ concentration estimation at the stations. At the same time, it can be seen from the figure that the spring $PM_{2.5}$ concentration of most air quality monitoring stations in the Chang-Zhu-Tan urban agglomeration was

higher than the Level 1 standard but lower than the Level 2 standard. The summer $PM_{2.5}$ concentration of most air quality monitoring stations was lower than the Level 1 standard, and the autumn $PM_{2.5}$ concentration of air quality monitoring stations was similar to spring but significantly higher than the spring $PM_{2.5}$ concentration. The $PM_{2.5}$ concentration of air quality monitoring stations in winter was significantly higher than that of the other three seasons, and the winter $PM_{2.5}$ concentration of most air quality monitoring stations was higher than the Level 2 standard.



(**a**) Spring             (**b**) Summer

(**c**) Autumn             (**d**) Winter
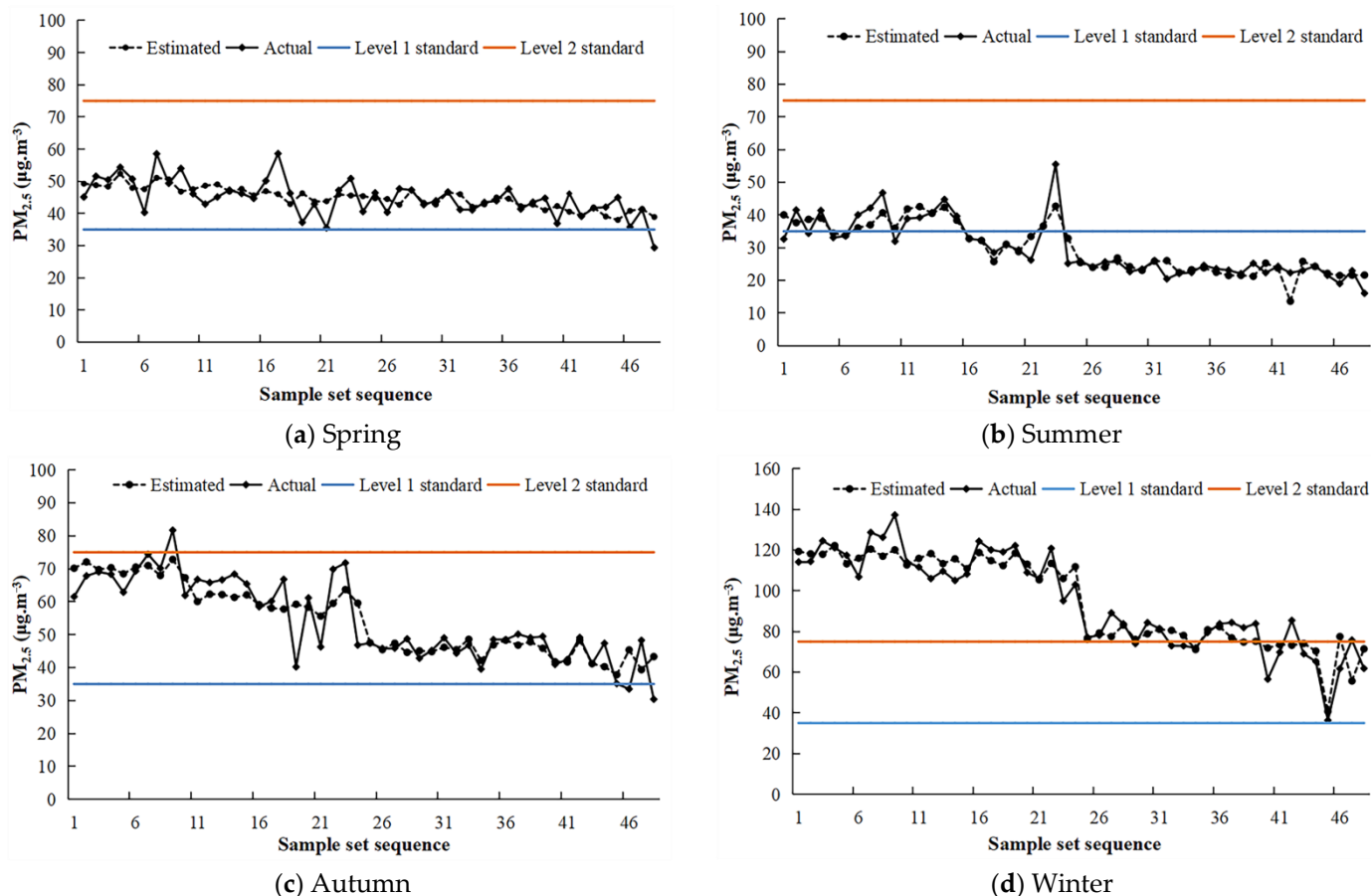
**Figure 6.** Comparison of the estimated and measured $PM_{2.5}$ concentrations, given the sample set sequence. The blue line represents the Level 1 standard, and the orange line represents the Level 2 standard. The Level 1 standard refers to the 24-h average $PM_{2.5}$ concentration lower than 35 $\mu g \cdot m^{-3}$. The Level 2 standard refers to the 24-h average $PM_{2.5}$ concentration lower than 75 $\mu g \cdot m^{-3}$.

In this paper, a total of 48 air quality monitoring stations, in 2015 and 2018, were analyzed for the real error of $PM_{2.5}$ concentration, and the average estimation errors of 48 stations in the four seasons were calculated (Figure 7). The results showed that the estimation error fluctuated greatly between stations, and there was an obvious uneven spatial distribution of model estimation errors. The total average error of 48 stations in the four seasons was 4.22 $\mu g \cdot m^{-3}$, and the estimation error of 23 stations was greater than the total average error. The spatial distribution of these 23 stations was further analyzed. Among them, 15 and 8 stations in 2015 and 2018, respectively, had estimation errors greater than the total average error, indicating that the estimation errors of $PM_{2.5}$ concentrations, at stations in 2015, were relatively large.
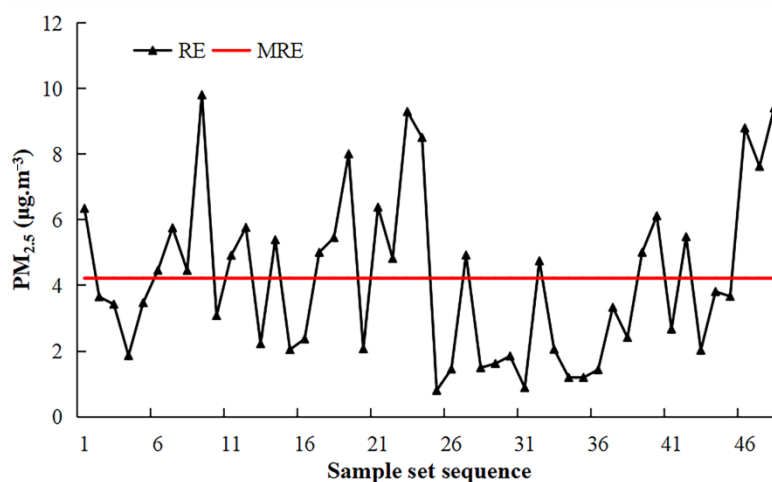
**Figure 7.** Error distribution of the estimated PM$_{2.5}$ concentration, given the sample set sequence. RE refers to the real error of each station in the four seasons. MRE refers to the mean real error of 47 stations in the four seasons.

Generally, an error higher than 4.22 µg·m$^{-3}$ is a high error site, and an error lower than 4.22 µg·m$^{-3}$ is a low error site. By analyzing the spatial locations of the 23 stations with large estimation errors, it can be found that the stations with high errors in 2015 and 2018 were mostly distributed in Xiangtan and Zhuzhou, and the economic development of these two cities was much slower than that of Changsha (Figure 8). The GDP of Changsha is 2.30 times that of the sum of the GDPs of Xiangtan and Zhuzhou, and the nighttime light area of Changsha is also much larger than that of Xiangtan and Zhuzhou. In addition, most stations distributed in dark areas at night had larger estimation errors, which was similar to the conclusion of Wang et al. [40]. The estimated models tended to underestimate PM$_{2.5}$ concentrations in darker nighttime areas.
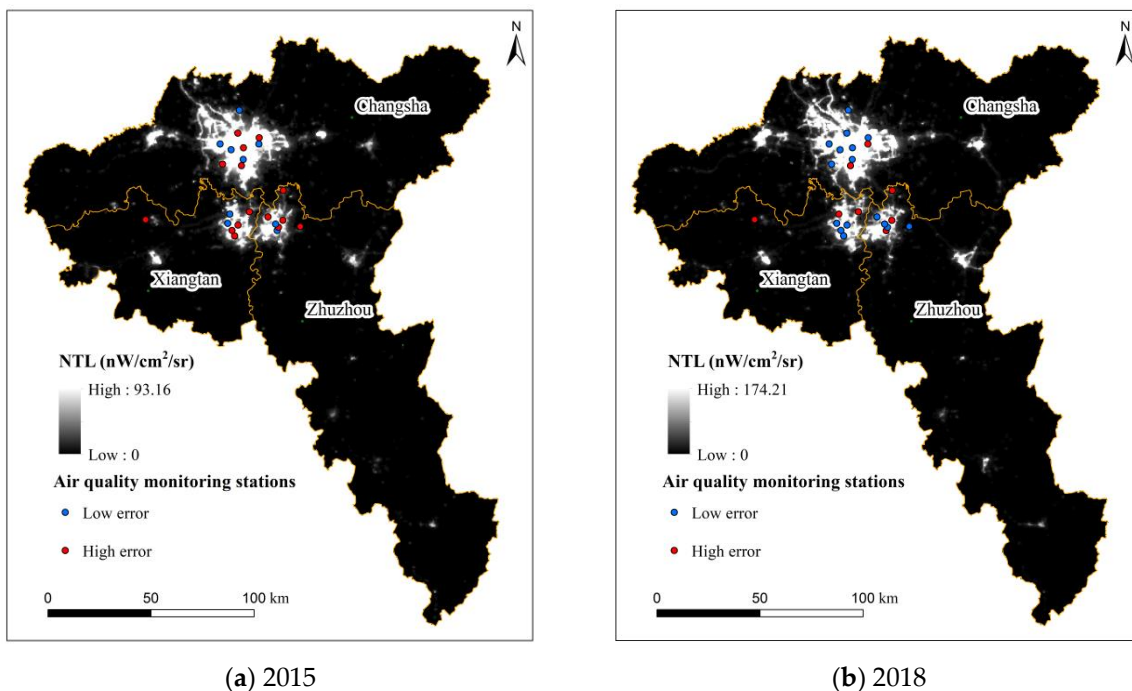


**(a)** 2015　　　　　　　　　　　　　　　　**(b)** 2018

**Figure 8.** Spatial distribution of low and high-error stations.

In this paper, a variety of estimation models for seasonal and annual PM$_{2.5}$ concentrations were constructed based on nighttime light images, meteorological data, and

topographic data. Except for spring, the models achieved high estimation accuracy, but further research is needed in terms of temporal and spatial resolution. In terms of temporal resolution, follow-up research should be more refined to the hourly scale. Nighttime light images, meteorological data, and topographic data can meet the requirements of this scale. However, in terms of spatial resolution, due to too few meteorological stations, the spatial resolution of meteorological conditions is limited. It is difficult to meet the high-precision inversion of $PM_{2.5}$ concentrations. At the same time, the spatial resolution of the nighttime light images used in this paper is low, at only 500 m, and subsequent research should attempt to select higher spatial resolution images.

## 6. Conclusions

Based on multisource data and monitoring station $PM_{2.5}$ concentration data, this paper constructed a variety of $PM_{2.5}$ concentration estimation models for the Chang-Zhu-Tan urban agglomeration. The seasonal and annual $PM_{2.5}$ concentrations of the Chang-Zhu-Tan urban agglomeration, in 2015 and 2018, were estimated, respectively, and the correlation between characteristic factors and $PM_{2.5}$ concentrations was analyzed. The results showed that, in terms of the estimation results of the seasonal $PM_{2.5}$ concentration model, the spring estimation results were the worst, and the winter estimation results were the best. Due to the increase in the number of samples in the annual $PM_{2.5}$ concentration model, the estimation results of the machine learning model were better than the seasonal estimation results. In terms of the correlation of $PM_{2.5}$ concentration, meteorological elements had a greater correlation with $PM_{2.5}$ concentration, followed by nighttime light radiance, and terrain elements and $PM_{2.5}$ concentration were the smallest.

This paper proposes a $PM_{2.5}$ concentration estimation method based on multisource data. At the same time, there are some limitations in multisource data fusion and continuous surface $PM_{2.5}$ concentration inversion, so further exploration is needed in subsequent research.

## References

1. You, M. Addition of $PM_{2.5}$ into the national ambient air quality standards of China and the contribution to air pollution control: The case study of Wuhan, China. *Sci. World J.* **2014**, *2014*, 768405. [CrossRef] [PubMed]
2. Guan, T.; Xue, T.; Guo, J.; Wang, X.; Zheng, Y.; Chao, B.; Kang, Y.; Chen, Z.; Zhang, L.; Zheng, C.; et al. How protective is China's National Ambient Air Quality Standards on short-term $PM_{2.5}$? Findings from blood pressure measurements of 1 million adults. *Environ. Res. Lett.* **2020**, *15*, 125014. [CrossRef]
3. Song, Y.; Li, Z.; Yang, T.; Xia, Q. Does the expansion of the joint prevention and control area improve the air quality?-Evidence from China's Jing-Jin-Ji region and surrounding areas. *Sci. Total Environ.* **2020**, *706*, 136034. [CrossRef] [PubMed]
4. Wu, W.; Zhang, M.; Ding, Y. Exploring the effect of economic and environment factors on $PM_{2.5}$ concentration: A case study of the Beijing-Tianjin-Hebei region. *J. Environ. Manag.* **2020**, *268*, 110703. [CrossRef]

5.      Li, Q.; Wu, B.; Liu, J.; Zhang, H.; Cai, X.; Song, Y. Characteristics of the atmospheric boundary layer and its relation with PM$_{2.5}$ during haze episodes in winter in the North China Plain. *Atmos. Environ.* **2020**, *223*, 117265. [CrossRef]

6.      Chen, Q.; Sheng, L.; Gao, Y.; Miao, Y.; Hai, S.; Gao, S.; Gao, Y. The Effects of the Trans-Regional Transport of PM$_{2.5}$ on a Heavy Haze Event in the Pearl River Delta in January 2015. *Atmosphere* **2019**, *10*, 237. [CrossRef]

7.      Hong, Y.; Li, C.; Li, X.; Ma, Y.; Zhang, Y.; Zhou, D.; Wang, Y.; Liu, N.; Chang, X. Analysis of Compositional Variation and Source Characteristics of Water-Soluble Ions in PM$_{2.5}$ during Several Winter-Haze Pollution Episodes in Shenyang, China. *Atmosphere* **2018**, *9*, 280. [CrossRef]

8.      Wang, S.; Wang, L.; Wang, N.; Ma, S.; Su, F.; Zhang, R. Formation of droplet-mode secondary inorganic aerosol dominated the increased PM$_{2.5}$ during both local and transport haze episodes in Zhengzhou, China. *Chemosphere* **2021**, *269*, 128744. [CrossRef]

9.      Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367–371. [CrossRef]

10.     Apte, J.S.; Brauer, M.; Cohen, A.J.; Ezzati, M.; Pope, C.A. Ambient PM$_{2.5}$ Reduces Global and Regional Life Expectancy. *Environ. Sci. Technol. Lett.* **2018**, *5*, 546–551. [CrossRef]

11.     Lin, H.; Guo, Y.; Zheng, Y.; Di, Q.; Liu, T.; Xiao, J.; Li, X.; Zeng, W.; Cummings-Vaughn, L.A.; Howard, S.W.; et al. Long-Term Effects of Ambient PM$_{2.5}$ on Hypertension and Blood Pressure and Attributable Risk Among Older Chinese Adults. *Hypertension* **2017**, *69*, 806–812. [CrossRef] [PubMed]

12.     Cohen, A.J.; Brauer, M.; Burnett, R.; Anderson, H.R.; Frostad, J.; Estep, K.; Balakrishnan, K.; Brunekreef, B.; Dandona, L.; Dandona, R.; et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. *Lancet* **2017**, *389*, 1907–1918. [CrossRef]

13.     Bu, X.; Xie, Z.; Liu, J.; Wei, L.; Wang, X.; Chen, M.; Ren, H. Global PM$_{2.5}$-attributable health burden from 1990 to 2017: Estimates from the Global Burden of disease study 2017. *Environ. Res.* **2021**, *197*, 111123. [CrossRef] [PubMed]

14.     Zhang, Z.; Shao, C.; Guan, Y.; Xue, C. Socioeconomic factors and regional differences of PM$_{2.5}$ health risks in China. *J. Environ. Manag.* **2019**, *251*, 109564. [CrossRef] [PubMed]

15.     Xie, Y.; Dai, H.; Zhang, Y.; Wu, Y.; Hanaoka, T.; Masui, T. Comparison of health and economic impacts of PM$_{2.5}$ and ozone pollution in China. *Environ. Int.* **2019**, *130*, 104881. [CrossRef]

16.     Kahn, R.; Banerjee, P.; McDonald, D.; Diner, D.J. Sensitivity of multiangle imaging to aerosol optical depth and to pure-particle size distribution and composition over ocean. *J. Geophys. Res. Atmos.* **1998**, *103*, 32195–32213. [CrossRef]

17.     Liu, Y. Estimating Ground-Level PM$_{2.5}$ in the Eastern United States Using Satellite Remote Sensing. *Environ. Sci. Technol.* **2005**, *39*, 3269–3278. [CrossRef] [PubMed]

18.     Tian, J.; Chen, D. A semi-empirical model for predicting hourly ground-level fine particulate matter (PM$_{2.5}$) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. *Remote Sens. Environ.* **2010**, *114*, 221–229. [CrossRef]

19.     Lee, H.J.; Coull, B.A.; Bell, M.L.; Koutrakis, P. Use of satellite-based aerosol optical depth and spatial clustering to predict ambient PM$_{2.5}$ concentrations. *Environ. Res.* **2012**, *118*, 8–15. [CrossRef] [PubMed]

20.     Li, Z.; Zhang, Y.; Shao, J.; Li, B.; Hong, J.; Liu, D.; Li, D.; Wei, P.; Li, W.; Li, L.; et al. Remote sensing of atmospheric particulate mass of dry PM$_{2.5}$ near the ground: Method validation using ground-based measurements. *Remote Sens. Environ.* **2016**, *173*, 59–68. [CrossRef]

21.     Lin, C.; Labzovskii, L.D.; Leung Mak, H.W.; Fung, J.C.H.; Lau, A.K.H.; Kenea, S.T.; Bilal, M.; Vande Hey, J.D.; Lu, X.; Ma, J. Observation of PM$_{2.5}$ using a combination of satellite remote sensing and low-cost sensor network in Siberian urban areas with limited reference monitoring. *Atmos. Environ.* **2020**, *227*, 117410. [CrossRef]

22.     Zhang, H.; Kondragunta, S. Daily and Hourly Surface PM$_{2.5}$ Estimation From Satellite AOD. *Earth Space Sci.* **2021**, *8*, 8. [CrossRef]

23.     Song, Z.; Fu, D.; Zhang, X.; Han, X.; Song, J.; Zhang, J.; Wang, J.; Xia, X. MODIS AOD sampling rate and its effect on PM$_{2.5}$ estimation in North China. *Atmos. Environ.* **2019**, *209*, 14–22. [CrossRef]

24.     Zhang, T.; Zhu, Z.; Gong, W.; Zhu, Z.; Sun, K.; Wang, L.; Huang, Y.; Mao, F.; Shen, H.; Li, Z.; et al. Estimation of ultrahigh resolution PM$_{2.5}$ concentrations in urban areas using 160 m Gaofen-1 AOD retrievals. *Remote Sens. Environ.* **2018**, *216*, 91–104. [CrossRef]

25.     Im, E.; Yang, S.; Zhang, P.; Chen, Y.; Han, W.; Chen, S.; Tong, L. Estimating ground-level PM$_{2.5}$ concentration using Landsat 8 in Chengdu, China. In Proceedings of the Remote Sensing of the Atmosphere, Clouds, and Precipitation V, Beijing, China, 1 December 2014.

26.     Yang, L. A novel approach of Landsat 8 imagery to predict PM$_{2.5}$ concentrations in a south-eastern coastal city of China. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *619*, 619. [CrossRef]

27.     Danek, T.; Zareba, M. The Use of Public Data from Low-Cost Sensors for the Geospatial Analysis of Air Pollution from Solid Fuel Heating during the COVID-19 Pandemic Spring Period in Krakow, Poland. *Sensors* **2021**, *21*, 5208. [CrossRef] [PubMed]

28.     Lu, H.; Liu, G. Spatial effects of carbon dioxide emissions from residential energy consumption: A county-level study using enhanced nocturnal lighting. *Appl. Energy* **2014**, *131*, 297–306. [CrossRef]

29.     Levin, N.; Kyba, C.C.M.; Zhang, Q.; Sánchez de Miguel, A.; Román, M.O.; Li, X.; Portnov, B.A.; Molthan, A.L.; Jechow, A.; Miller, S.D.; et al. Remote sensing of night lights: A review and an outlook for the future. *Remote Sens. Environ.* **2020**, *237*, 111443. [CrossRef]

*Int. J. Environ. Res. Public Health* **2022**, *19*, 4306

17 of 18

30. Liang, H.; Guo, Z.; Wu, J.; Chen, Z. GDP spatialization in Ningbo City based on NPP/VIIRS night-time light and auxiliary data using random forest regression. *Adv. Space Res.* **2020**, *65*, 481–493. [CrossRef]

31. Yu, B.; Shi, K.; Hu, Y.; Huang, C.; Chen, Z.; Wu, J. Poverty Evaluation Using NPP-VIIRS Nighttime Light Composite Data at the County Level in China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1217–1229. [CrossRef]

32. Wu, K.; Wang, X. Aligning Pixel Values of DMSP and VIIRS Nighttime Light Images to Evaluate Urban Dynamics. *Remote Sens.* **2019**, *11*, 1463. [CrossRef]

33. Yang, Y.; Ma, M.; Zhu, X.; Ge, W. Research on spatial characteristics of metropolis development using nighttime light data: NTL based spatial characteristics of Beijing. *PLoS ONE* **2020**, *15*, e0242663. [CrossRef] [PubMed]

34. Tan, M.; Li, X.; Li, S.; Xin, L.; Wang, X.; Li, Q.; Li, W.; Li, Y.; Xiang, W. Modeling population density based on nighttime light images and land use data in China. *Appl. Geogr.* **2018**, *90*, 239–247. [CrossRef]

35. Zhong, L.; Liu, X.; Yang, P.; Lin, R. Explore the application of high-resolution nighttime light remote sensing images in nighttime marine ship detection: A case study of LJ1-01 data. *Open Geosci.* **2020**, *12*, 1169–1184. [CrossRef]

36. Johnson, R.S.; Zhang, J.; Hyer, E.J.; Miller, S.D.; Reid, J.S. Preliminary investigations toward nighttime aerosol optical depth retrievals from the VIIRS Day/Night Band. *Atmos. Meas. Tech.* **2013**, *6*, 1245–1255. [CrossRef]

37. Zhang, J.; Reid, J.S.; Miller, S.D.; Turk, F.J. Strategy for studying nocturnal aerosol optical depth using artificial lights. *Int. J. Remote Sens.* **2008**, *29*, 4599–4613. [CrossRef]

38. Wang, J.; Aegerter, C.; Xu, X.; Szykman, J.J. Potential application of VIIRS Day/Night Band for monitoring nighttime surface PM$_{2.5}$ air quality from space. *Atmos. Environ.* **2016**, *124*, 55–63. [CrossRef]

39. Fu, D.; Xia, X.; Duan, M.; Zhang, X.; Li, X.; Wang, J.; Liu, J. Mapping nighttime PM$_{2.5}$ from VIIRS DNB using a linear mixed-effect model. *Atmos. Environ.* **2018**, *178*, 214–222. [CrossRef]

40. Xu, G.; Ren, X.; Xiong, K.; Li, L.; Bi, X.; Wu, Q. Analysis of the driving factors of PM$_{2.5}$ concentration in the air: A case study of the Yangtze River Delta, China. *Ecol. Indic.* **2020**, *110*, 105889. [CrossRef]

41. Zhang, G.; Shi, Y.; Xu, M. Evaluation of LJ1-01 Nighttime Light Imagery for Estimating Monthly PM$_{2.5}$ Concentration: A Comparison With NPP-VIIRS Nighttime Light Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3618–3632. [CrossRef]

42. Yuan, M.; Huang, Y.; Shen, H.; Li, T. Effects of urban form on haze pollution in China: Spatial regression analysis based on PM$_{2.5}$ remote sensing data. *Appl. Geogr.* **2018**, *98*, 215–223. [CrossRef]

43. Yang, Q.; Yuan, Q.; Yue, L.; Li, T. Investigation of the spatially varying relationships of PM$_{2.5}$ with meteorology, topography, and emissions over China in 2015 by using modified geographically weighted regression. *Environ. Pollut.* **2020**, *262*, 114257. [CrossRef]

44. Li, X.; Feng, Y.J.; Liang, H.Y. The Impact of Meteorological Factors on PM$_{2.5}$ Variations in Hong Kong. *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *78*, 012003. [CrossRef]

45. He, Q.; Huang, B. Satellite-based high-resolution PM$_{2.5}$ estimation over the Beijing-Tianjin-Hebei region of China using an improved geographically and temporally weighted regression model. *Environ. Pollut.* **2018**, *236*, 1027–1037. [CrossRef] [PubMed]

46. He, Y.; Lin, Y.; Zhou, G.; Zhu, Y.; Tang, K. Spatial pattern and drivers of urbanization in China's mid-level developing urban agglomeration: A case study of Chang-Zhu-Tan. *Reg. Sustain.* **2021**, *2*, 83–97. [CrossRef]

47. Zhang, Y.; Jiang, W. Pollution characteristics and influencing factors of atmospheric particulate matter (PM$_{2.5}$) in Chang-Zhu-Tan area. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *108*, 42047. [CrossRef]

48. Zhu, Y.F. Characteristics of a Continuous Haze Process in the Chang-Zhu-Tan Region of China and Its Influencing Factors. *Appl. Ecol. Environ. Res.* **2019**, *17*, 17. [CrossRef]

49. Yao, L.; Huang, C.; Jing, W.; Yue, X.; Xu, Y. Quantitative Assessment of Relationship between Population Exposure to PM$_{2.5}$ and Socio-Economic Factors at Multiple Spatial Scales over Mainland China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2058. [CrossRef] [PubMed]

50. Ma, X.-Y.; Xiao, Z.-H.; He, L.-Z.; Cao, Y.-J.; Liu, J.-S. Comparison of chemical characteristics of PM$_{2.5}$ during two winters in Xiangtan City in south central China. *J. Atmos. Chem.* **2020**, *77*, 169–183. [CrossRef]

51. Murano, G.; Corrado, V.; Dirutigliano, D. The new Italian Climatic Data and their Effect in the Calculation of the Energy Performance of Buildings. *Energy Procedia* **2016**, *101*, 153–160. [CrossRef]

52. Lin, J.; Shi, W. Statistical Correlation between Monthly Electric Power Consumption and VIIRS Nighttime Light. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 32. [CrossRef]

53. Ivan, K.; Holobâcă, I.-H.; Benedek, J.; Török, I. VIIRS Nighttime Light Data for Income Estimation at Local Level. *Remote Sens.* **2020**, *12*, 2950. [CrossRef]

54. Li, Y.; Song, Z.; Wu, B.; Yu, B.; Wu, Q.; Hong, Y.; Liu, S.; Wu, J. Evaluating the ability of NOAA-20 monthly composite data for socioeconomic indicators estimation and urban area extraction. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1837–1845. [CrossRef]

55. Lopez-Ruiz, H.G.; Blazquez, J.; Hasanov, F. Estimating the Saudi Arabian Regional GDP Using Satellite Nighttime Light Images. *SSRN Electron. J.* **2018**. [CrossRef]

56. Lin, Z.; Xu, H. Anthropogenic Heat Flux Estimation Based on Luojia 1-01 New Nighttime Light Data: A Case Study of Jiangsu Province, China. *Remote Sens.* **2020**, *12*, 3707. [CrossRef]

57. Hu, W.; Zhao, T.; Bai, Y.; Shen, L.; Sun, X.; Gu, Y. Contribution of Regional PM$_{2.5}$ Transport to Air Pollution Enhanced by Sub-Basin Topography: A Modeling Case over Central China. *Atmosphere* **2020**, *11*, 1258. [CrossRef]

58. Wang, J.; Wang, Y.; Liu, H.; Yang, Y.; Zhang, X.; Li, Y.; Zhang, Y.; Deng, G. Diagnostic identification of the impact of meteorological conditions on PM$_{2.5}$ concentrations in Beijing. *Atmos. Environ.* **2013**, *81*, 158–165. [CrossRef]

59. Chen, T.; He, J.; Lu, X.; She, J.; Guan, Z. Spatial and Temporal Variations of PM$_{2.5}$ and Its Relation to Meteorological Factors in the Urban Area of Nanjing, China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 921. [CrossRef]

60. Liu, Q.; Wu, R.; Zhang, W.; Li, W.; Wang, S. The varying driving forces of PM$_{2.5}$ concentrations in Chinese cities: Insights from a geographically and temporally weighted regression model. *Environ. Int.* **2020**, *145*, 106168. [CrossRef]

61. Afanador, N.L.; Tran, T.N.; Buydens, L.M. Use of the bootstrap and permutation methods for a more robust variable importance in the projection metric for partial least squares regression. *Anal. Chim. Acta* **2013**, *768*, 49–56. [CrossRef]

62. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [CrossRef]

63. Zhang, Z.; Wu, L.; Chen, Y. Forecasting PM$_{2.5}$ and PM10 concentrations using GMCN(1,N) model with the similar meteorological condition: Case of Shijiazhuang in China. *Ecol. Indic.* **2020**, *119*, 106871. [CrossRef]

64. Hu, X.; Belle, J.H.; Meng, X.; Wildani, A.; Waller, L.A.; Strickland, M.J.; Liu, Y. Estimating PM$_{2.5}$ Concentrations in the Conterminous United States Using the Random Forest Approach. *Environ. Sci. Technol.* **2017**, *51*, 6936–6944. [CrossRef] [PubMed]

65. Chen, Y. Prediction algorithm of PM$_{2.5}$ mass concentration based on adaptive BP neural network. *Computing* **2018**, *100*, 825–838. [CrossRef]

66. Xue, T.; Zheng, Y.; Tong, D.; Zheng, B.; Li, X.; Zhu, T.; Zhang, Q. Spatiotemporal continuous estimates of PM$_{2.5}$ concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations. *Environ. Int.* **2019**, *123*, 345–357. [CrossRef] [PubMed]

67. Danesh Yazdi, M.; Kuang, Z.; Dimakopoulou, K.; Barratt, B.; Suel, E.; Amini, H.; Lyapustin, A.; Katsouyanni, K.; Schwartz, J. Predicting Fine Particulate Matter (PM$_{2.5}$) in the Greater London Area: An Ensemble Approach using Machine Learning Methods. *Remote Sens.* **2020**, *12*, 914. [CrossRef]

68. Just, A.C.; De Carli, M.M.; Shtein, A.; Dorman, M.; Lyapustin, A.; Kloog, I. Correcting Measurement Error in Satellite Aerosol Optical Depth with Machine Learning for Modeling PM$_{2.5}$ in the Northeastern USA. *Remote Sens.* **2018**, *10*, 803. [CrossRef]

69. Xu, W.; Sun, J.; Liu, Y.; Xiao, Y.; Tian, Y.; Zhao, B.; Zhang, X. Spatiotemporal variation and socioeconomic drivers of air pollution in China during 2005–2016. *J. Environ. Manag.* **2019**, *245*, 66–75. [CrossRef]