remote sensing

Technical Note

Exploring the Conversion Model from Aerosol Extinction Coefficient to PM_1, $PM_{2.5}$ and PM_{10} Concentrations

Huanhuan Shao [1], Hui Li [1], Shikuan Jin [2], Ruonan Fan [2], Weiyan Wang [2], Boming Liu [2,*], Yingying Ma [2], Ruyi Wei [1] and Wei Gong [1]

1 School of Electronic Information, Wuhan University, Wuhan 430079, China
2 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China
* Correspondence: liuboming@whu.edu.cn

Abstract: Particle matter (PM) mass concentrations have an important influence on human and environmental health. Lidar plays an important role in the monitoring of PM concentrations. However, the accuracy of PM concentrations retrieved via lidar depends on the performance of the conversion model from the aerosol extinction coefficient (EC) to PM concentration. Therefore, surface PM concentrations, aerosol EC and five meteorological factors are used to build the conversion model that can also be applicable to lidar for retrieving PM concentrations. In this study, the traditional linear model (LM), random forest (RF) and artificial neural network (ANN) algorithms are used to estimate the mass concentrations of PM with aerodynamic diameters < 1 μm (PM_1), 2.5 μm ($PM_{2.5}$) and 10 μm (PM_{10}). The influence of meteorological factors on the conversion model is analyzed. The results show that the meteorological parameters play a non-ignorable role in the model of PM retrieval based on EC, especially when retrieving PM_{10}. Moreover, the performance of three models is investigated by comparing with the surface measurements. The results indicate that the RF and ANN models are more suitable to estimate PM than the LM model. The diurnal variations in mean relative error (MRE) from the three models are then analyzed. There is a diurnal pattern in MRE values, meaning that the maximum values occur in the afternoon and the minimum values occur at night. In addition, there are subtle differences in performance between two machine learning (ML) models. After analysis, it is found that for PM_{10}, the RF method is superior to the ANN when the EC value is small, while the ANN method is superior to the RF when the EC value is relatively high, and the EC threshold is set to 0.6 km^{-1}. For PM_1 and $PM_{2.5}$ estimation, the ANN is the most appropriate model. Finally, accurate diurnal variations in PM_1 and $PM_{2.5}$ based on the ANN model and PM_{10} based on the combined model of RF and ANN (named RA) are investigated. The results exhibit that the daily maximum values of PM_1, $PM_{2.5}$ and PM_{10} in the Wuhan area all occur at approximately 08:00–10:00 local time (LT), which is mainly due to the impact of commuter vehicle emissions and the impact of secondary photochemistry response aggravated by sufficient illumination and temperature rises after sunrise. These research results provide an important basis for particulate matter monitoring.

Keywords: particle matter; aerosol extinction coefficient; random forest; artificial neural network

Citation: Shao, H.; Li, H.; Jin, S.; Fan, R.; Wang, W.; Liu, B.; Ma, Y.; Wei, R.; Gong, W. Exploring the Conversion Model from Aerosol Extinction Coefficient to PM_1, $PM_{2.5}$ and PM_{10} Concentrations. *Remote Sens.* **2023**, *15*, 2742. https://doi.org/10.3390/rs15112742

Academic Editor: Jing Wei

Received: 10 April 2023
Revised: 22 May 2023
Accepted: 23 May 2023
Published: 25 May 2023

1. Introduction

Particulate matter (PM) refers to fine solid or liquid aerosol particles suspended in the atmosphere [1,2]. PM with aerodynamic diameters < 1 μm (PM_1), 2.5 μm ($PM_{2.5}$) and 10 μm (PM_{10}) is the main atmospheric pollutant and is also important in affecting the atmospheric environment [3–6]. With the rapid development of society, the intense consumption of energy increases the emissions of PM, resulting in serious air pollution and affecting the human living environment [7–10]. A large amount of PM will cause haze and reduce atmospheric visibility [11–13]. The combination of PM and chemical substances seriously endangers human health and causes a variety of diseases [14–16]. Therefore, continuous monitoring of the variations in PM concentrations is crucial [17].

At present, there are many monitoring methods for PM concentrations. Surface in situ measurement is a common method to obtain the surface PM concentrations. Chen et al. [18] used surface PM_{10} concentrations collected from ground monitoring stations to investigate the relationship between daily mortality and PM_{10} in 16 cities in China. Pu et al. [19] used the measurement of ground $PM_{2.5}$ concentration to study the influence of long-distance transport on aerosol properties at Shangdianzi station. To obtain the large-scale surface PM concentrations, aerosol optical depth (AOD) derived from satellite measurements was applied to establish the relationship with surface PM [2,3,20,21]. Chen et al. [22] incorporated daily surface PM_1, AOD obtained via satellite observation and other predictive information to retrieve the PM_1 concentrations across China from 2005 to 2014. Based on the meteorological information and other predictors, Hu et al. [23] built the association of AOD with ground $PM_{2.5}$ using geographically weighted regression. Meng et al. [24] estimated surface PM_{10} concentrations in Shanghai by using AOD. These studies have improved the understanding of the change in surface PM concentrations.

With further research, PM concentration profiles have gradually attracted attention because it is important to understand the transport of pollutants and the formation process of regional pollution [25–27]. The original method of measuring the PM concentration profile is to use meteorological towers installed with PM monitors. Yang et al. [28] employed $PM_{2.5}$ monitors on a meteorological tower to derive $PM_{2.5}$ profiles in Beijing. Due to the limited detection height of this measurement method, some studies have retrieved PM concentration profiles from aerosol extinction coefficient (EC) profiles obtained from lidar systems. Liu et al. [29] explored the association of the surface particle number concentration, with EC obtained via lidar data, and derived the vertical distribution of the particle number concentration. Raut and Chazette [30] examined the vertical distribution of PM_{10} in the area around the Paris Peripherique by developing two approaches for converting aerosol EC retrieved from mobile lidar measurements into PM_{10}. These studies confirm that lidar has good time and height resolution and can be applied to derive the PM concentration profiles [17]. However, there are still two difficulties in obtaining the full profile of PM concentrations from lidar. One is that, due to the influence of the overlap factor, a lidar system is usually unable to obtain near-ground extinction coefficient signals [29]. On the other hand, the accuracy of PM profiles retrieved by lidar depends on the performance of the conversion model from EC to PM concentration. Therefore, here, we use the surface EC not derived from lidar data to develop the conversion model that can be applicable to lidar for retrieving PM concentration profiles. Currently, the traditional linear model (LM) and machine learning (ML) algorithms are applied to retrieve PM concentration profiles. Lv et al. [31] adopted the LM method to study the association of $PM_{2.5}$ concentration with EC and derived the $PM_{2.5}$ profile. Ma et al. [32] derived $PM_{2.5}$ profiles by using ML methods. Zhu et al. [33] applied a random forest model (RF) in ML to derive the $PM_{2.5}$ profiles in Wuhan. The aim of these studies is to develop an accurate approach to describe the nonlinear transformation from EC to PM concentrations. The artificial neural network (ANN) algorithm, as an ML algorithm, can map complex nonlinear relationships between multiple inputs and outputs [34]. However, it has rarely been applied in the estimation of PM concentrations. Therefore, it is worth attempting to utilize the ANN algorithm to construct the conversion model from EC to PM concentrations.

In this study, surface PM_1, $PM_{2.5}$ and PM_{10} concentrations, aerosol EC and five meteorological parameters from November 2014 to May 2017 in Wuhan are utilized to build the conversion model. The purpose of building the conversion model from EC to PM concentrations is to obtain an accurate model that can be applicable to lidar for retrieving PM concentrations. The comparison analyses of the LM and two ML models are then carried out. Finally, the most accurate model is built to analyze the diurnal variations in these three particle concentrations in Wuhan. The remainder of this paper is organized as follows. The observation station and data are illustrated in Section 2. In Section 3, three retrieval methods of PM concentrations are provided. In Section 4, the comparison results of model accuracy are discussed. The main conclusions are provided in Section 5.

2. Station and Data

2.1. Observation Station

Located in Central China, Wuhan is an important industrial city in the middle reaches of the Yangtze River [35,36]. With the industrial development and the continuously increasing number of people and vehicles, the environment around Wuhan has been seriously affected, and the air quality problem has aroused people's concern [37]. An atmospheric observatory is located in the State Key Laboratory of Surveying, Mapping and Remote Sensing Information Engineering (LIESMARS) of Wuhan University, China (39.98°N, 116.38°W), with an altitude of about 23 m. The observation station is located in an urban area, surrounded by buildings. It is equipped with devices, including nephelometer, aethalometer, PM detector and automatic weather station.

2.2. Ground Data

The aerosol EC is the sum of aerosol scattering coefficients (SCs) and aerosol absorption coefficients (ACs) [32]. The surface aerosol SCs are obtained using the nephelometer (model 3563, TSI, Saint Paul, MN, USA), where the error in the data is about 7%. It can provide the surface SCs at 450, 550 and 700 nm simultaneously in 5 min intervals, which can be used to derive the surface SC at 532 nm [38,39]. This device is calibrated every three months [36]. Moreover, the aethalometer (model AE31, Magee Scientific, Berkeley, CA, USA) is employed to provide the black carbon concentrations that can be applied to obtain the surface aerosol AC. This device is regularly maintained every three months [40]. The surface AC at 532 nm can be derived from black carbon concentrations at 880 nm [39,41]. Therefore, the surface EC at 532 nm is obtained from the surface SC at 532 nm and surface AC at 532 nm.

An automatic meteorological station (U3-NRC, Onset HOBO, Cape Cod, MA, USA) is employed to provide meteorological parameters. Here, five parameters, including pressure (Press), relative humidity (RH), temperature (T), wind direction (WD) and wind speed (WS), are collected. They are applied as auxiliary data for model construction. Moreover, the environmental particulate detector (Grimm EDM 180, Ainring, Bayern, Germany) is applied to provide the ground PM_1, $PM_{2.5}$ and PM_{10} concentrations, which are used as standard values for model training. The measuring principle of this instrument is to firstly obtain the size and number concentration of each particle, which are calculated by measuring the laser light scattering of each particle in the sample air chamber. Then, the obtained size and number concentration is converted into PM mass concentrations by using the protocol [42]. These observation data are processed into hourly average data. The time series of all these data are shown in Figure S1. After matching, a total of 5342 h of data are finally obtained from 16 November 2014 to 18 May 2017.

3. Methodology

In this section, the LM method and two ML methods are applied to estimate PM_1, $PM_{2.5}$ and PM_{10} concentrations, respectively. Statistical methods utilized to evaluate the predictive accuracy of models are also proposed.

3.1. Traditional Linear Model

LM method was utilized to estimate PM concentrations [31,43,44]. Lv et al. [31] interpreted the association of PM concentrations with aerosol EC and showed that EC is linear with the total particle mass concentration. Therefore, here, the LM model is built to study the association of EC with PM_1, $PM_{2.5}$ and PM_{10}. Due to the fact that the ECs of aerosols are significantly affected by RH [1], RH is a vital factor in evaluating the association of PM concentrations with aerosol ECs. The dependence of PM concentrations on aerosol ECs under different RH conditions is shown in Figure 1. The black line is the fitting result. In order to avoid unreasonable negative values, the restricted linear fitting through the origin is applied here according to the approach of Liu et al. [45], which is represented by the red line. Based on this method, the linear relationship between ECs and PM_1, $PM_{2.5}$ and PM_{10} is, respectively:

$$EC = 0.0072 \times PM_1 \tag{1}$$

$$EC = 0.0067 \times PM_{2.5} \tag{2}$$

$$EC = 0.0053 \times PM_{10} \tag{3}$$

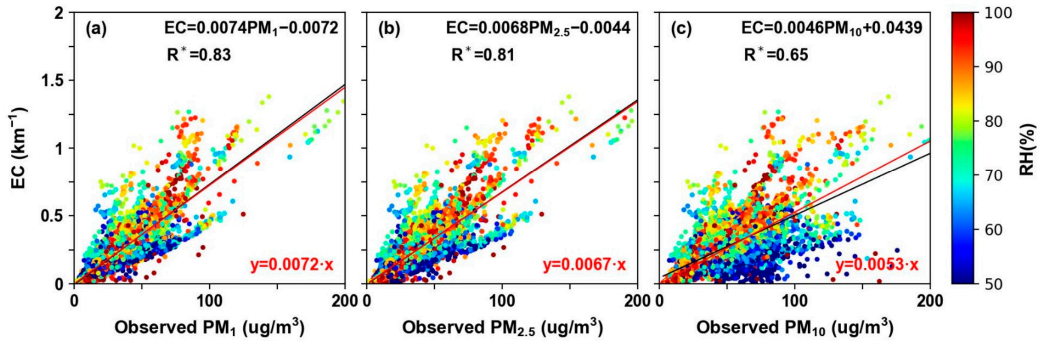

Figure 1. The linear regression relationship (**a**) between observed PM_1 and EC, (**b**) between observed $PM_{2.5}$ and EC and (**c**) between observed PM_{10} and EC. The black and red lines are the reference line and the regression line, respectively. The color bar represents the RH. The asterisk indicates that the R passed the statistical significance difference test ($p < 0.05$).

The correlation coefficient (R) between surface ECs and PM_1, $PM_{2.5}$ and PM_{10} is 0.83, 0.81 and 0.65, respectively. It shows that the correlation between PM_{10} and ECs is relatively low compared with PM_1 and $PM_{2.5}$. This is mainly related to the detection performance of the instruments and meteorological factors [1,46]. Chen et al. [1] revealed that with RH enlarging, the uncertainty of the correlation between PM and EC will exacerbate. Li et al. [46] also indicated that the correlation between PM and EC is affected by RH. Because under high RH, ECs will increase due to the enlargement of the size of aerosol particles, yet the concentrations of PM measured by the instrument are handled by drying.

3.2. Machine Learning Algorithms

Two ML algorithms, containing the RF and ANN, are utilized to retrieve PM_1, $PM_{2.5}$ and PM_{10}. Here, the input variables contain surface EC, Press, RH, T, WD and WS. We randomly assign all the ground observation data (5342 sets) to be a training set and a testing set according to 90% and 10%. Among them, there are 4808 sets of data in the training set and 534 sets of data in the testing set. The 10-fold crossover is used to train these two ML models. There are ten testing sets, nine of which contain 534 sets of data and one of which contains 536 sets of data. Therefore, all 5342 sets of data in this study participate in the results testing.

3.2.1. Random Forest Model

RF model was proposed by Breiman [47]. As shown in Figure 2a, RF model is an integrated model composed of multiple decision trees, which are irrelevant for each other [33,48]. It produces numerous independent trees and takes the average value of each tree estimation result as the final estimation result [49]. In this paper, the final estimated PM concentrations are given by the average result of all single decision trees [50]. Since the RF algorithm can deal with multiple input variables and produce the best results by considering various variables, it has been widely employed to retrieve concentrations of atmospheric pollutants [49].

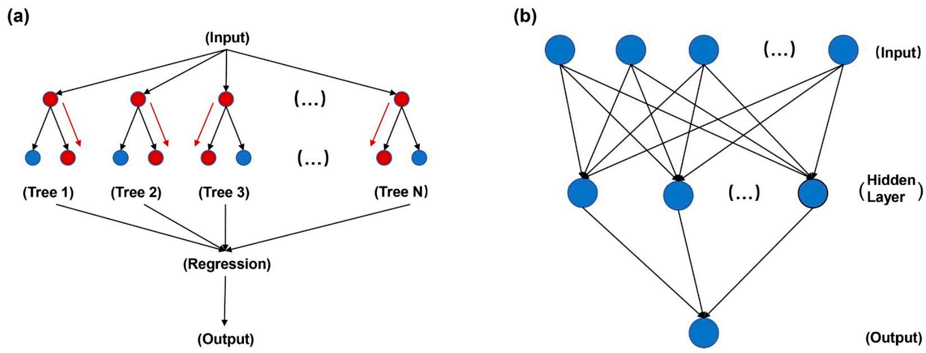

Figure 2. Schematic diagram of the (**a**) RF and (**b**) ANN algorithms used to estimate PM_1, $PM_{2.5}$, PM_{10}.

Here, surface aerosol ECs and the five meteorological parameters mentioned above are used as input variables to train the model. The number of trees (estimator num) and the maximum depth of each decision tree (max depth num) are two important parameters in RF model training. When training and optimizing the model, the best performance model can be obtained by adjusting these two parameters, as shown in Figure 3a–c. The principle of parameter adjustment is to define an appropriate value of estimator num at the minimum RMSE, which is also applied in selecting the value of max depth num. Finally, the values of estimator num and max depth num are set to 22 and 88, 14 and 84, and 53 and 67 for PM_1, $PM_{2.5}$ and PM_{10} estimations, respectively.

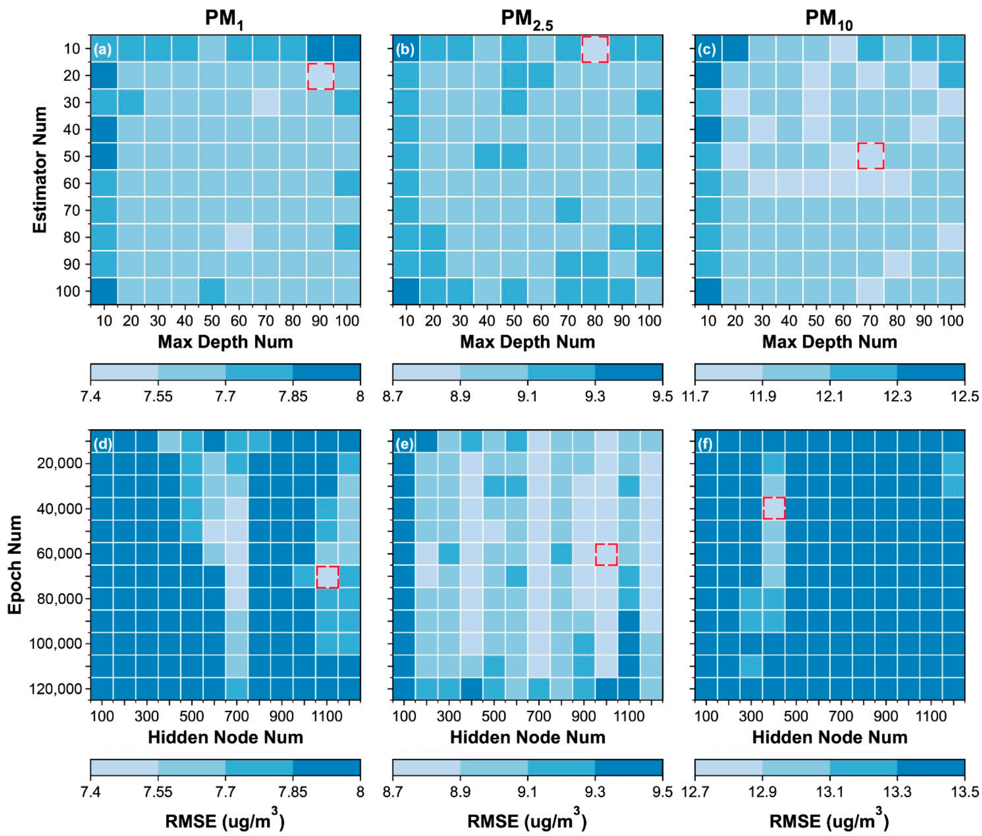

Figure 3. The value of RMSE between observed PM_1 and estimated PM_1 (**a**), observed $PM_{2.5}$ and estimated $PM_{2.5}$ (**b**), observed PM_{10} and estimated PM_{10} (**c**) based on the RF model changes with estimator num and max depth num. Similar for ANN model changes with epoch num and hidden node num (resp. **d**–**f**). The red dotted box indicates the area where the optimal parameters are located.

3.2.2. Artificial Neural Network

ANN is a computational technique that simulates the structure and function of the human brain and nervous system through mathematical modeling [51]. The ANN algorithm is a classic and common ML algorithm. Because this algorithm can build complex nonlinear relationships between input features and output data [34], it has been employed to solve the nonlinear fitting problem. As shown in Figure 2b, its network structure consists of three layers, which are input, hidden and output layers, respectively. These three parts are connected in sequence, and each layer can contain multiple neurons. When the input data enter the neural network through the input layer, the hidden layer will process the data and output them from the output layer.

Here, the ANN model adopts one hidden layer and adds the hyperbolic tangent functions, which are tanh(x) and relu(x), as activation functions. When training and optimizing the neural network model, we construct the best performance model by adjusting the number of epochs (epoch num) and the number of hidden nodes (hidden node num), as shown in Figure 3d–f. It exhibits that the performance of the ANN model built in this paper is mainly affected by the hidden node num, which indicates that we only need to select an appropriate value of hidden node num. After parameter tuning, the values of hidden node num are defined to 1050, 1038 and 410 for PM_1, $PM_{2.5}$ and PM_{10} estimations, respectively.

3.2.3. Sensibility Analysis

The PM concentrations are not only correlated with ECs but also correlated with meteorological parameters [5,52]. Therefore, the importance analysis of the input parameter for two ML models is carried out, as shown in Figure 4. For PM_1, $PM_{2.5}$ and PM_{10}, the importance values of ECs in RF (ANN) are 0.6 (0.45), 0.57 (0.7) and 0.39 (0.43), which are evidently greater than that of other input variables. The result exhibits that the concentrations of these three particles are mainly affected by ECs. Moreover, for the RF (ANN) model in predicting PM_1, the importance values of Press, RH and T are 0.17 (0.2), 0.06 (0.17) and 0.11 (0.14), which are also relatively large varying from 0.05 to 0.2. The case of $PM_{2.5}$ is similar to that of PM_1, with importance values varying from 0.05 to 0.2. It indicates that meteorological factors, such as Press, RH and T, will also be considered in the inversion process of RF and ANN models. Specifically, among these input variables, the Press is a relatively important contributor, with larger importance values of 0.15. This is because there is a positive correlation between the Press and PM concentration [53]. Li et al. [53] indicated that air Press affected convection, thereby influencing PM's transport and accumulation. However, it notes that for the RF (ANN) model in predicting PM_{10}, the importance values of Press, RH and T are 0.23 (0.27), 0.18 (0.16) and 0.12 (0.11), which are all larger than 0.1. Combined with these results, it indicates that the concentrations of PM_{10} are more affected by meteorological factors than the concentrations of PM_1 and $PM_{2.5}$. Therefore, the influence of meteorological parameters should be taken into account when estimating PM_1, $PM_{2.5}$ and PM_{10} using aerosol ECs, especially when estimating PM_{10}.

Due to the influence of ECs on PM concentration estimations, error analysis for the three models is performed. The difference between the predicted PM concentrations and the observed PM concentrations is shown in Figure 5. The results of three LM models (Figure 5a,d,g) exhibit that the predicted PM concentrations deviate from their observed values. Figure 5a and d show that the inversion results of LM are overestimated when the EC is greater than 0.6 km^{-1}. The result of LM (Figure 5g) shows an overestimation when the EC is greater than 0.3 km^{-1}. This is because the LM model does not consider the effects of RH. High RH exacerbates the complexity of predicting particle concentration due to the hygroscopic growth of aerosols [1]. By comparison, the difference in RF and ANN models is obviously smaller than that of the LM model, no matter for PM_1, $PM_{2.5}$ or PM_{10}. This is due to the fact that the meteorological parameters, such as Press, RH and T, are taken into account in the building of the two models. In addition, the deviations in RF and ANN are relatively small and steady, which do not increase with the aerosol EC. Overall, the RF and ANN models are more suitable than the LM model for PM estimation based on the EC.

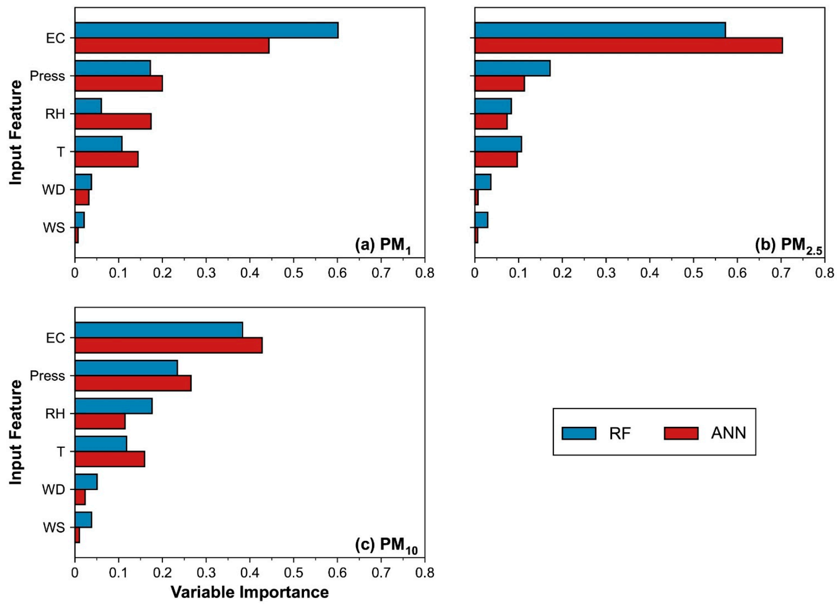

Figure 4. The input variable importance for the RF and ANN models of PM_1 (**a**), $PM_{2.5}$ (**b**) and $PM_{10,}$ (**c**) respectively.

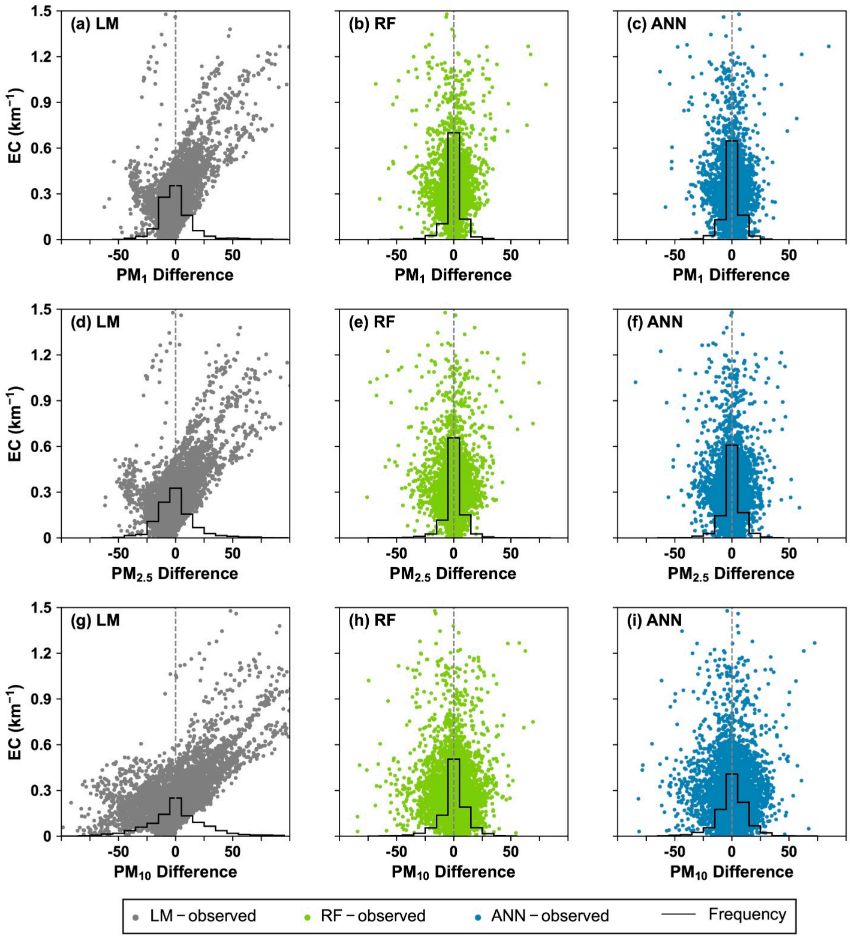

Figure 5. Difference in observed PM_1 and predicted PM_1, observed $PM_{2.5}$ and predicted $PM_{2.5}$ and observed PM_{10} and predicted PM_{10} as a function of EC based on the (**a**,**d**,**g**) LM, (**b**,**e**,**h**) RF and (**c**,**f**,**i**) ANN models. The gray, green and blue dots represent the difference for LM-observed, RF-observed and ANN-observed PM, respectively. The black line represents the frequency of difference.

3.3. Statistical Methods

In this study, we select three statistical metrics to assess the predictive abilities of each model: R, root mean square error (RMSE) and mean relative error (MRE). R indicates the correlation between estimated and observed values. RMSE and MRE are indicators used to quantify the difference between estimated and observed values. In addition, RMSE is an important indicator for adjusting model parameters. The calculation formulas used to calculate R, RMSE and MRE are as follows:

$$R = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \tag{5}$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - x_i|}{x_i} \tag{6}$$

where n represents the total number of samples, x_i and y_i are the ith sample point of the observed and estimated values and $\overline{x}$ and $\overline{y}$ are the mean value of observed and estimated value.

4. Results and Discussion

4.1. Intercomparison of Prediction Results

Figure 6 presents the comparisons between the predicted PM concentrations and the actual concentrations based on three methods. Consistent with the result in Figure 5, the concentrations of these three particles estimated via the LM model deviate from their observed values. The results of three LM models (Figure 6a,d,g) indicate that PM_{10} has the largest number of deviation samples compared with PM_1 and $PM_{2.5}$, which is because the correlation between PM_{10} and EC is relatively low, as described in Section 3.1. The simple linear relationship between EC and PM is unable to obtain the surface PM concentrations with high accuracy. By contrast, the PM concentrations estimated using the two ML models are closer to their observed values. For PM_1, the R of the LM, RF and ANN models is 0.83, 0.93 and 0.93, respectively. The RMSE of these three models is 15.41, 7.95 and 7.90 $\mu g/m^3$, respectively. The R and RMSE of the three models for $PM_{2.5}$ and PM_{10} can be seen in Figure 6. These results show that compared with the LM, the accuracy of the RF and ANN models is significantly improved, and the RMSE is about half of the LM model. This is because the RF and ANN models take into account the influence of meteorological factors, thus improving the performance of the models. Furthermore, it notes that for PM_1 ($PM_{2.5}$), the R of both ML models is improved by 0.1 (0.12) compared with the LM model. While for PM_{10}, the R of the RF (ANN) model is improved by 0.24 (0.22) compared with the LM, which is nearly double that of PM_1 and $PM_{2.5}$. This is because the PM_{10} concentrations are more affected by meteorological factors than PM_1 and $PM_{2.5}$, which is also confirmed by the results in Figure 4. Overall, for these three particles, the accuracy of the RF and ANN models is relatively similar, and both are evidently superior to that of the LM model.

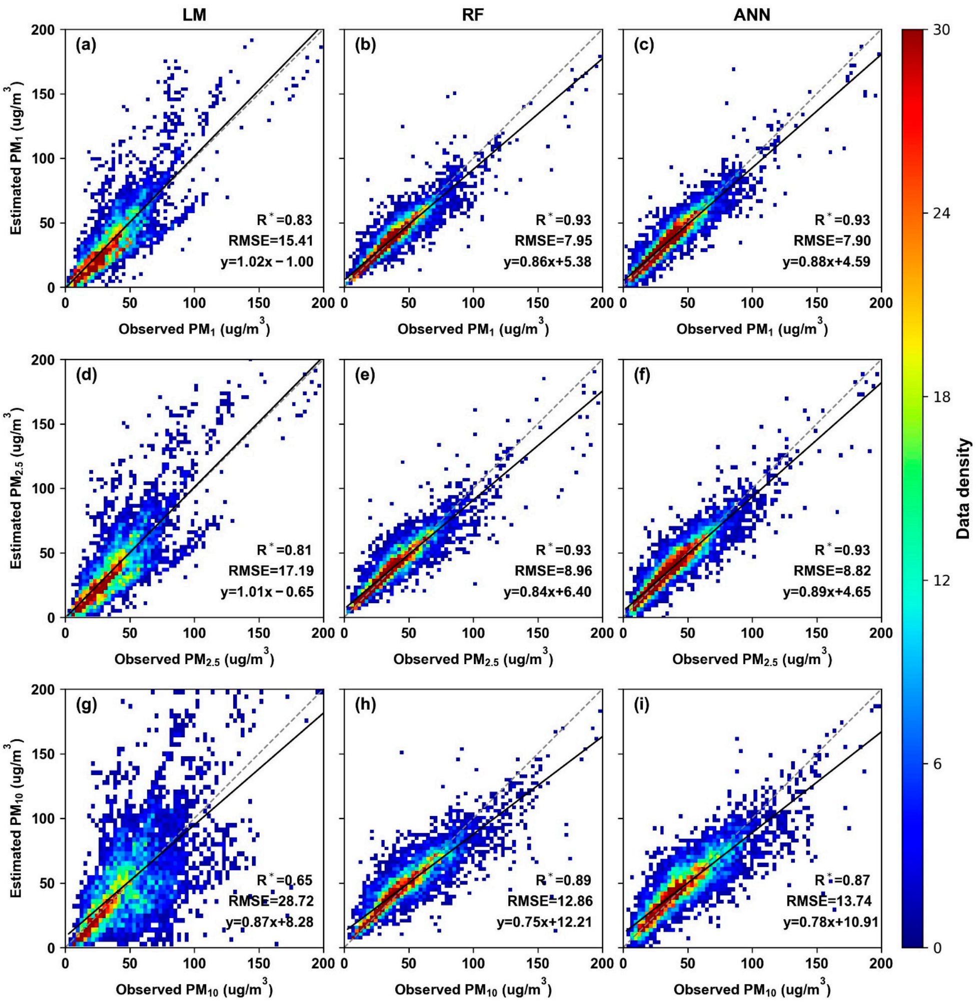

Figure 6. Comparison of predicted PM_1 and observed PM_1, predicted $PM_{2.5}$ and observed $PM_{2.5}$ and predicted PM_{10} and observed PM_{10} based on the (**a**,**d**,**g**) LM, (**b**,**e**,**h**) RF and (**c**,**f**,**i**) ANN models. The gray and black lines are the reference and regression line, respectively. The asterisk indicates that the R passed the statistical significance difference test ($p < 0.05$).

Figure 7 presents the hourly MRE between the PM concentrations retrieved using the LM, RF and ANN models and their observed values. The MRE values of LM are higher than that of RF and ANN in every hour for these three particles. These results exhibit that there are smaller errors in estimating PM concentrations using the RF and ANN methods compared with LM, which is consistent with the results in Figure 6. Moreover, the hourly MRE shows a diurnal pattern that the maximum values occur at 13:00–17:00 local time (LT) and the minimum values occur at 00:00–06:00 LT. Taking PM_{10} as an example, the MRE values of the LM, RF and ANN models reach the maximum in the afternoon, which are, respectively, 46%, 28% and 31% and remain at a low level at night, where the lowest MRE values are, respectively, 40%, 14% and 18%. This may be due to the fact that PM concentrations are also affected by the planetary boundary layer height and atmospheric turbulence [44,54,55]. Li et al. [54] revealed that with the enhancement in solar radiation in the afternoon, the turbulent diffusion rate increased, and the planetary boundary layer height also increased. These two parameters peaked at approximately 15:00 LT, and both then declined with solar radiation and kept steady after sunset. For these reasons, the error in estimating PM concentrations is higher in the afternoon than at night. Li et al. [55] also indicated that the higher RMSE and MAE in estimating PM_1 concentration over Zhejiang province was concentrated during 12:00–15:00 local solar time (LST).

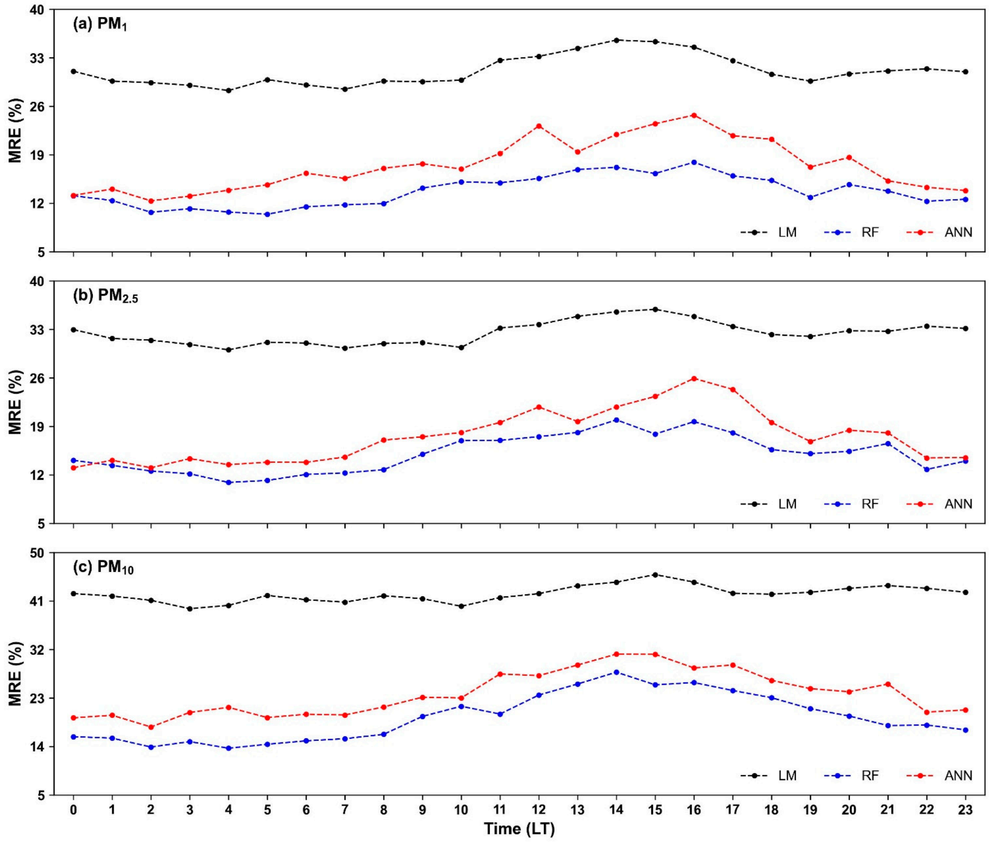

Figure 7. Hourly MRE between (**a**) observed PM_1 and estimated PM_1, (**b**) observed $PM_{2.5}$ and estimated $PM_{2.5}$, and (**c**) observed PM_{10} and estimated PM_{10} based on three models, respectively. The black, blue and red lines represent the MRE between observed PM and estimated PM using LM, RF and ANN models, respectively.

4.2. Performance Difference between RF and ANN with Different EC Thresholds

Combined with the results in Section 4.1, we find that although both ML models are superior to the LM model, there are still subtle differences in performance between these two models. Since the PM concentrations are mainly affected by aerosol EC, we explore which of the two algorithms is better under different EC thresholds, as shown in Figure 8. The solid dots indicate that the corresponding case passed the statistical significance difference test ($p < 0.05$). The number of samples for each bin is listed in Table S1. It can be seen that for PM_1 and $PM_{2.5}$, the RMSE (R) of RF is greater than (less than) that of the ANN model when the EC value is higher than these thresholds. These results exhibit that for PM_1 and $PM_{2.5}$, the ANN model is more appropriate for their estimation based on the EC. For PM_{10}, the RMSE (R) of the ANN model is greater than (less than) that of the RF model when the EC value is greater than a low threshold, while the RMSE (R) of RF is greater than (less than) that of the ANN model when the EC value is greater than a high threshold. These results exhibit that for PM_{10}, the RF model is more appropriate for estimating low values and the ANN model is more suitable for estimating high values. This is because there are fewer training samples at high values, which leads to reduced generalization ability of RF model. The generalization ability refers to the adaptability of the ML algorithm to fresh samples. The RF model is an integrated method based on the decision trees and the final estimation result depends on the output of multiple decision trees, which means that its generalization ability is affected by the number of training samples. By contrast, ANN is a computational method to deal with complex nonlinear fitting problems through mathematical modeling, where the generalization ability is relatively less affected by the number of training samples. These results exhibit that the combined use of RF and ANN is likely to produce more accurate results for PM_{10}. Therefore, it is necessary to choose an appropriate threshold to reach the best result. Figure 8c,f show that for PM_{10}, the ANN model outperforms the RF model when the threshold is > 0.6 km^{-1}. On the contrary, when the threshold is less than 0.6 km^{-1}, the RF model outperforms the ANN model. Therefore, the suitable threshold is set to 0.6 km^{-1} for PM_{10}.

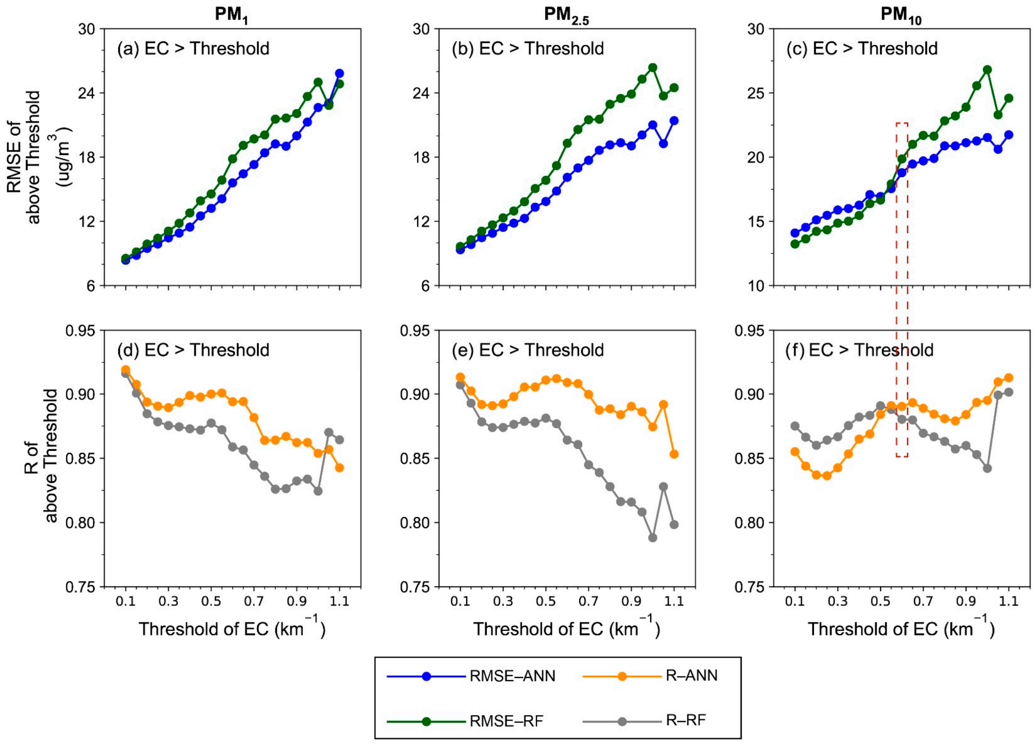

Figure 8. Change trends of RMSE between estimated PM_1 (**a**), $PM_{2.5}$ (**b**) and PM_{10} (**c**) by two ML models and their actual values when EC is higher than the threshold. Similar for the case of the change trends of R (resp. **d**–**f**). The red dotted box indicates the selected EC threshold.

4.3. Diurnal Variations

The results in Figure 8 indicate that for PM_{10}, the RF model is more suitable to estimate low values, and the ANN model is more suitable to estimate high values. The combined use of RF and ANN (named RA) is developed for PM_{10} to produce the most accurate results. The RA model is that when the EC is less than the threshold selected (see Figure 8c,f), the RF model is utilized to estimate PM_{10} concentrations, and when the EC is larger than the threshold, the ANN model is utilized to estimate PM_{10} concentrations. For PM_1 and $PM_{2.5}$, the ANN model is the most appropriate model for estimation. Figure 9 presents the diurnal variations in PM_1 and $PM_{2.5}$ based on the ANN model and PM_{10} based on the RA model in Wuhan. The diurnal variations using LM, ANN and RA models are compared. The shaded area represents the range of the average value ± standard deviation. From the RMSE or standard deviation results, it exhibits that the ANN model is significantly superior to the LM model (see Figure 9a,b) and the RA model is significantly superior to the LM model (see Figure 9c). Moreover, the RMSE of the RA model is 12.77 $\mu g/m^3$, which is lower than that of the RF and ANN models. This indicates that the RA model outperforms both ML models and is the most appropriate model to predict PM_{10} concentrations.

Furthermore, the daily maximum values of the concentrations of these three particles all occur at approximately 08:00–10:00 LT, which shows a unimodal pattern of morning peak. This is because Wuhan is a metropolis with a dense population. In the urban area, the morning peaks are mainly contributed by enhanced anthropogenic activity during rush hour. Huang et al. [56] confirmed the five main contributors of $PM_{2.5}$ in Wuhan by continuously measuring $PM_{2.5}$ and its chemical composition. Among the five contributors, the most important contributors of $PM_{2.5}$ are secondary photochemistry and traffic-related emissions in Wuhan, which could explain why the daily maximum values of the three particles all occur at approximately 08:00–10:00 LT. This time period is the peak time for $PM_{2.5}$ emissions from commuting vehicles, and it is also the time when there is sufficient light and the temperature rise after sunrise, which intensifies the secondary photochemistry response. Moreover, Wang et al. [42] showed that at most stations in the middle and lower reaches of the Yangtze River, the diurnal variation in PM_1 and PM_{10} is similar to that of $PM_{2.5}$; thus, there is also a morning peak at approximately 08:00–10:00 LT.

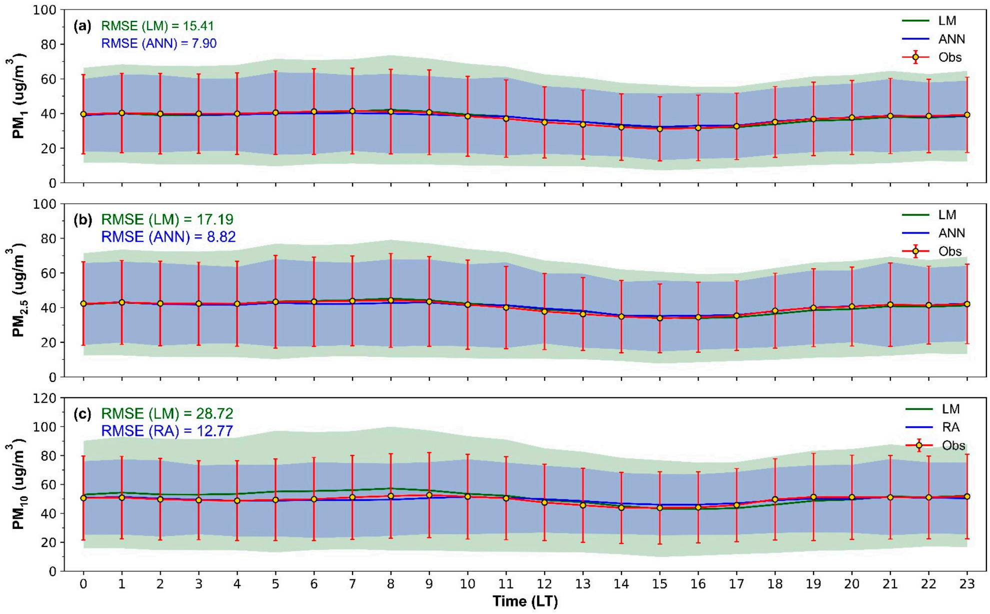

Figure 9. Diurnal variations of estimated values of PM_1 (**a**), $PM_{2.5}$ (**b**) and PM_{10} (**c**) by the different models and their observed values in Wuhan from November 2014 to May 2017. The green and red lines are the hourly average values of PM predicted by LM model and their observed values, respectively. The blue lines in (**a**–**c**) are the hourly average values predicted by ANN and RA models.

5. Conclusions

This study uses the LM, RF and ANN algorithms to estimate PM_1, $PM_{2.5}$ and PM_{10} concentrations, respectively. The performance of three models is compared based on the difference and correlation between the predicted PM concentrations and the observed values. Finally, we analyze the diurnal variations in PM_1 and $PM_{2.5}$ concentrations based on the ANN model and propose a combined model of RF and ANN (named RA) to analyze the diurnal variation in PM_{10} concentrations.

The PM concentrations retrieved using the RF and ANN models are close to their observed values, and the accuracy of both ML models is superior to the LM model. This is because these two models take into account the influence of meteorological parameters, thus improving the prediction accuracy, while the LM model estimates PM only depending on EC, which is greatly affected by RH. Moreover, the accuracy in predicting PM_{10} using the LM model is much worse than that in predicting PM_1 and $PM_{2.5}$. This is because the correlation between PM_{10} and EC is the lowest compared with PM_1 and $PM_{2.5}$. The diurnal variations in MRE from the three models are then analyzed. There exists a diurnal pattern in MRE values that the maximum values occur in the afternoon and the minimum values occur at night. In addition, the two ML models show subtle differences in estimating the PM concentrations. For PM_1 and $PM_{2.5}$, ANN is the most appropriate model for their estimation. For PM_{10}, the RF model is more suitable to estimate low values, and the ANN model is more suitable to estimate high values. Therefore, the RA model is presented for PM_{10} to produce the smallest RMSE, which indicates that RA is the most accurate model to estimate PM_{10} concentrations. Finally, diurnal variations in the concentrations of PM_1, $PM_{2.5}$ and PM_{10} are investigated, and their daily maximum values all appear at approximately 08:00–10:00 LT, which indicates that the particulate pollution in Wuhan is mainly affected by secondary photochemistry and commuter vehicles.

The above research results are helpful to provide selections of input parameters and model types for long-term monitoring and modeling of PM concentrations and have great reference value for further particulate matter concentration monitoring and modeling, characteristic research, correlation analysis and other work, with great significance for air quality and atmospheric environment assessment. Furthermore, these conversion models developed in this study can be applicable to PM inversion by lidar in future studies. However, there are some deficiencies in our work. Our research data are limited, as they are only from the Wuhan area. Therefore, these conversion models could also be applicable to regions other than Wuhan after adjusting hyperparameters in future studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs15112742/s1, Figure S1: The time series of all these data from 16 November 2014 to 18 May 2017. The red, green, blue, and black lines are the dephelometer, aethalometer, meteorological and PM data, respectively; Table S1: The number of samples above each threshold and P of the corresponding case.

Author Contributions: Conceptualization, H.S. and B.L.; methodology, B.L.; software, H.S.; validation, S.J., H.L., W.W. and R.F.; formal analysis, R.W.; investigation, H.S., H.L., S.J., W.W. and R.F.; resources, W.G. and B.L.; data curation, B.L. and Y.M.; writing—original draft preparation, H.S.; writing—review and editing, B.L.; visualization, H.S. and B.L.; supervision, W.G.; project administration, Y.M.; funding acquisition, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (under grants 42001291), the Fundamental Research Funds for the Central Universities (under grants 2042022kf1003) and the Natural Science Fund of Hubei Province (under grants 2022CFB044).

Data Availability Statement: The data and codes used in this paper can be provided for non-commercial research purposes upon request (Boming Liu, Email: liuboming@whu.edu.cn).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, J.; Xin, J.; An, J.; Wang, Y.; Liu, Z.; Chao, N.; Meng, Z. Observation of aerosol optical properties and particulate pollution at background station in the Pearl River Delta region. *Atmos. Res.* **2014**, *143*, 216–227. [CrossRef]
2. Chen, G.; Li, S.; Knibbs, L.D.; Hamm, N.; 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. *Sci. Total Environ.* **2018**, *636*, 52–60. [CrossRef] [PubMed]
3. Gupta, P.; Christopher, S.A.; Wang, J.; Gehrig, R.; Lee, Y.; Kumar, N. Satellite Remote Sensing of Particulate Matter and Air Quality Assessment over Global Cities. *Atmos. Environ.* **2006**, *40*, 5880–5892. [CrossRef]
4. Wei, J.; Huang, W.; Li, Z.; Xue, W.; Peng, Y.; Sun, L.; Cribb, M. Estimating 1 km-resolution $PM_{2.5}$ concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* **2019**, *231*, 111221. [CrossRef]
5. Zhang, L.; An, J.; Liu, M.; Li, Z.; Liu, Y.; Tao, L.; Liu, X.; Zhang, F.; Zheng, D.; Gao, Q.; et al. Spatiotemporal variations and influencing factors of $PM_{2.5}$ concentrations in Beijing, China. *Environ. Pollut.* **2020**, *262*, 114276. [CrossRef]
6. Wei, J.; Li, Z.; Xue, W.; Sun, L.; Fan, T.; Liu, L.; Su, T.; Cribb, M. The China High PM_{10} dataset: Generation, validation, and spatiotemporal variations from 2015 to 2019 across China. *Environ. Int.* **2021**, *146*, 106290. [CrossRef]
7. Fu, X.; Cheng, Z.; Wang, S.; Hua, Y.; Xing, J.; Hao, J. Local and Regional Contributions to Fine Particle Pollution in Winter of the Yangtze River Delta, China. *Aerosol Air Qual. Res.* **2016**, *16*, 1067–1080. [CrossRef]
8. Fontes, T.; Li, P.; Barros, N.; Zhao, P. Trends of $PM_{2.5}$ concentrations in China: A long term approach. *J. Environ. Manag.* **2017**, *196*, 719–732. [CrossRef]
9. Zhang, M.; Ma, Y.; Gong, W.; Liu, B.; Shi, Y.; Chen, Z. Aerosol optical properties and radiative effects: Assessment of urban aerosols in central China using 10-year observations. *Atmos. Environ.* **2018**, *182*, 275–285. [CrossRef]
10. Zhang, X.; Ji, G.; Peng, X.; Kong, L.; Zhao, X.; Ying, R.; Yin, W.; Xu, T.; Cheng, J.; Wang, L. Characteristics of the chemical composition and source apportionment of $PM_{2.5}$ for a one-year period in Wuhan, China. *J. Atmos. Chem.* **2022**, *79*, 101–115. [CrossRef]
11. Zhao, P.; Zhang, X.; Xu, X.; Zhao, X. Long-term visibility trends and characteristics in the region of Beijing, Tianjin, and Hebei, China. *Atmos. Res.* **2011**, *101*, 711–718. [CrossRef]

12. Pui, D.Y.; Chen, S.C.; Zuo, Z. $PM_{2.5}$ in China: Measurements, sources, visibility and health effects, and mitigation. *Particuology* **2014**, *13*, 1–26. [CrossRef]
13. Zhang, M.; Jin, S.; Ma, Y.; Fan, R.; Wang, L.; Gong, W.; Liu, B. Haze events at different levels in winters: A comprehensive study of meteorological factors, Aerosol characteristics and direct radiative forcing in megacities of north and central China. *Atmos. Environ.* **2021**, *245*, 118056. [CrossRef]
14. Davidson, C.I.; Phalen, R.F.; Solomon, P.A. Airborne particulate matter and human health: A review. *Aerosol Sci. Technol.* **2005**, *39*, 737–749. [CrossRef]
15. Wei, J.; Li, Z.; Lyapustin, A.; Sun, L.; Peng, Y.; Xue, W.; Su, T.; Cribb, M. Reconstructing 1-km-resolution high-quality $PM_{2.5}$ data records from 2000 to 2018 in China: Spatiotemporal variations and policy implications. *Remote Sens. Environ.* **2021**, *252*, 112136. [CrossRef]
16. Lu, J.; Wu, K.; Ma, X.; Wei, J.; Yuan, Z.; Huang, Z.; Fan, W.; Zhong, Q.; Huang, Y.; Wu, X. Short-term effects of ambient particulate matter (PM_1, $PM_{2.5}$ and PM_{10}) on influenza-like illness in Guangzhou, China. *Int. J. Hyg. Environ. Health* **2023**, *247*, 114074. [CrossRef]
17. Wei, J.; Li, Z.; Sun, L.; Xue, W.; Ma, Z.; Liu, L.; Fan, T.; Cribb, M. Extending the EOS Long-Term $PM_{2.5}$ Data Records since 2013 in China: Application to the VIIRS Deep Blue Aerosol Products. *IEEE* **2022**, *60*, 4100412. [CrossRef]
18. Chen, R.; Kan, H.; Chen, B.; Huang, W.; Bai, Z.; Song, G.; Pan, G. Association of particulate air pollution with daily mortality: The China Air Pollution and Health Effects Study. *Am. J. Epidemiol.* **2012**, *175*, 1173–1181. [CrossRef]
19. Pu, W.; Zhao, X.; Shi, X.; Ma, Z.; Zhang, X.; Yu, B. Impact of long-range transport on aerosol properties at a regional background station in Northern China. *Atmos. Res.* **2015**, *153*, 489–499. [CrossRef]
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. *Atmos. Chem. Phys.* **2011**, *11*, 7991–8002. [CrossRef]
21. Ma, Z.; Hu, X.; Sayer, A.M.; Levy, R.; Zhang, Q.; Xue, Y.; Tong, S.; Bi, J.; Huang, L.; Liu, Y. Satellite-Based Spatiotemporal Trends in $PM_{2.5}$ Concentrations: China, 2004–2013. *Environ. Health Perspect.* **2016**, *124*, 184–192. [CrossRef]
22. Chen, G.; Knibbs, L.D.; Zhang, W.; Li, S.; Cao, W.; Guo, J.; Ren, H.; Wang, B.; Wang, H.; Williams, G.; et al. Estimating spatiotemporal distribution of PM_1 concentrations in China with satellite remote sensing, meteorology, and land use information. *Environ. Pollut.* **2018**, *233*, 1086–1094. [CrossRef]
23. Hu, X.; Waller, L.A.; Al-Hamdan, M.Z.; Crosson, W.L.; Estes, M.G., Jr.; Estes, S.M.; Quattrochi, D.A.; Sarnat, J.A.; Liu, Y. Estimating ground-level $PM_{2.5}$ concentrations in the southeastern U.S. using geographically weighted regression. *Environ. Res.* **2013**, *121*, 1–10. [CrossRef]
24. Meng, X.; Fu, Q.; Ma, Z.; Chen, L.; Zou, B.; Zhang, Y.; Xue, W.; Wang, J.; Wang, D.; Kan, H.; et al. Estimating ground-level PM_{10} in a Chinese city by combining satellite data, meteorological information and a land use regression model. *Environ. Pollut.* **2016**, *208*, 177–184. [CrossRef]
25. Sun, Y.; Song, T.; Tang, G.; Wang, Y. The vertical distribution of $PM_{2.5}$ and boundary-layer structure during summer haze in Beijing. *Atmos. Environ.* **2013**, *74*, 413–421. [CrossRef]
26. Liu, Y.; Tang, G.; Zhou, L.; Hu, B.; Liu, B.; Li, Y.; Liu, S.; Wang, Y. Mixing layer transport flux of particulate matter in Beijing, China. *Atmos. Chem. Phys.* **2019**, *19*, 9531–9540. [CrossRef]
27. Liu, C.; Huang, J.; Wang, Y.; Tao, X.; Hu, C.; Deng, L.; Xu, J.; Xiao, H.-W.; Luo, L.; Xiao, H.-Y.; et al. Vertical distribution of $PM_{2.5}$ and interactions with the atmospheric boundary layer during the development stage of a heavy haze pollution event. *Sci. Total Environ.* **2020**, *704*, 135329. [CrossRef]
28. Yang, L.; He, K.; Zhang, Q.; Wang, Q. Vertical distributive characters of $PM_{2.5}$ at the ground layer in autumn and winter in Beijing. *Res. Environ. Sci.* **2005**, *18*, 23–28.
29. Liu, Z.; Liu, J.; Wang, B.; Lu, F.; Huang, S.; Wu, D.; Han, D. Aerosol observation in Fengtai area, Beijing. *Particuology* **2008**, *6*, 214–217. [CrossRef]
30. Raut, J.-C.; Chazette, P. Assessment of vertically-resolved PM_{10} from mobile lidar observations. *Atmos. Chem. Phys.* **2009**, *9*, 8617–8638. [CrossRef]
31. Lv, L.; Liu, W.; Zhang, T.; Chen, Z.; Dong, Y.; Fan, G.; Xiang, Y.; Yao, Y.; Yang, N.; Chu, B.; et al. Observations of particle extinction, $PM_{2.5}$ mass concentration profile and flux in north China based on mobile lidar technique. *Atmos. Environ.* **2017**, *164*, 360–369. [CrossRef]
32. Ma, Y.; Zhu, Y.; Liu, B.; Li, H.; Jin, S.; Zhang, Y.; Fan, R.; Gong, W. Estimation of the vertical distribution of particle matter ($PM_{2.5}$) concentration and its transport flux from lidar measurements based on machine learning algorithms. *Atmos. Chem. Phys.* **2021**, *21*, 17003–17016. [CrossRef]
33. Zhu, Y.; Ma, Y.; Liu, B.; Xu, X.; Jin, S.; Gong, W. Retrieving the Vertical Distribution of $PM_{2.5}$ Mass Concentration from Lidar via a Random Forest Model. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5701209. [CrossRef]
34. Yao, L.; Lu, N. Spatiotemporal distribution and short-term trends of particulate matter concentration over China, 2006–2010. *Environ. Sci. Pollut. Res.* **2014**, *21*, 9665–9675. [CrossRef] [PubMed]
35. Zhang, Y.; Wang, W.; He, J.; Jin, Z.; Wang, N. Spatially continuous mapping of hourly ground ozone levels assisted by Himawari-8 short wave radiation products. *GIScience Remote Sens.* **2023**, *60*, 2174280. [CrossRef]
36. Liu, B.; Ma, Y.; Shi, Y.; Jin, S.; Jin, Y.; Gong, W. The characteristics and sources of the aerosols within the nocturnal residual layer over Wuhan, China. *Atmos. Res.* **2020**, *241*, 104959. [CrossRef]

37. Liu, B.; Ma, Y.; Gong, W.; Zhang, M.; Yang, J. Study of continuous air pollution in winter over Wuhan based on ground-based and satellite observations. *Atmos. Pollut. Res.* **2018**, *9*, 156–165. [CrossRef]
38. Yan, W.; Yang, L.; Chen, J.; Wang, X.; Wen, L.; Zhao, T.; Wang, W. Aerosol optical properties at urban and coastal sites in Shandong Province, Northern China. *Atmos. Res.* **2017**, *188*, 39–47. [CrossRef]
39. Liu, B.; Gong, W.; Ma, Y.; Zhang, M.; Yang, J.; Zhang, M. Surface Aerosol Optical Properties during High and Low Pollution Periods at an Urban Site in Central China. *Aerosol Air Qual. Res.* **2018**, *18*, 3035–3046. [CrossRef]
40. Gong, W.; Zhang, M.; Han, G.; Ma, X.; Zhu, Z. An investigation of aerosol scattering and absorption properties in Wuhan, Central China. *Atmosphere* **2015**, *6*, 503–520. [CrossRef]
41. Xu, J.; Tao, J.; Zhang, R.; Cheng, T.; Leng, C.; Chen, J.; Huang, G.; Li, X.; Zhu, Z. Measurements of surface aerosol optical properties in winter of Shanghai. *Atmos. Res.* **2012**, *109*, 25–35. [CrossRef]
42. Wang, Y.; Zhang, X.; Sun, J.; Zhang, X.; Che, H.; Li, Y. Spatial and temporal variations of the concentrations of PM_{10}, $PM_{2.5}$ and PM_1 in China. *Atmos. Chem. Phys.* **2015**, *15*, 13585–13598. [CrossRef]
43. Zhao, Q.; Su, H.; Yi, M.; Yu, D.; Xu, C. Aerosol Horizontal Distribution Detected by Lidar in Excavation Stage of Construction Site Foundation Pit. *Chin. J. Lasers* **2021**, *48*, 2010001.
44. Tao, Z.; Wang, Z.; Yang, S.; Shan, H.; Ma, X.; Zhang, H.; Zhao, S.; Liu, D.; Xie, C.; Wang, Y. Profiling the $PM_{2.5}$ mass concentration vertical distribution in the boundary layer. *Atmos. Meas. Tech.* **2016**, *9*, 1369–1376. [CrossRef]
45. Liu, B.; Ma, Y.; Gong, W.; Zhang, M.; Shi, Y. The relationship between black carbon and atmospheric boundary layer height. *Atmos. Pollut. Res.* **2019**, *10*, 65–72. [CrossRef]
46. Li, L.; Yang, J.; Wang, Y. Retrieval of High-Resolution Atmospheric Particulate Matter Concentrations from Satellite-Based Aerosol Optical Thickness over the Pearl River Delta Area, China. *Remote Sens.* **2015**, *7*, 7914–7937. [CrossRef]
47. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
48. Liu, B.; Ma, X.; Guo, J.; Li, H.; Jin, S.; Ma, Y.; Gong, W. Estimating hub-height wind speed based on a machine learning algorithm: Implications for wind energy assessment. *Atmos. Chem. Phys.* **2023**, *23*, 3181–3193. [CrossRef]
49. Yu, S.; Vautard, R. A transfer method to estimate hub-height wind speed from 10 meters wind speed based on machine learning. *Renew. Sustain. Energy Rev.* **2022**, *169*, 112897. [CrossRef]
50. Shi, T.; Han, G.; Ma, X.; Mao, H.; Chen, C.; Han, Z.; Gong, W. Quantifying factory-scale CO_2/CH_4 emission based on mobile measurements and EMISSION-PARTITION model: Cases in China. *Environ. Res. Lett.* **2023**, *18*, 034028. [CrossRef]
51. Gao, S.; Zhao, H.; Bai, Z.; Han, B.; Xu, J.; Zhao, R.; Zhang, N.; Chen, L.; Lei, X.; Shi, W.; et al. Combined use of principal component analysis and artificial neural network approach to improve estimates of $PM_{2.5}$ personal exposure: A case study on older adults. *Sci. Total Environ.* **2020**, *726*, 138533. [CrossRef] [PubMed]
52. Zha, Y.; Gao, J.; Jiang, J.; Lu, H.; Huang, J. Monitoring of urban air pollution from MODIS aerosol data: Effect of meteorological parameters. *Tellus* **2010**, *62*, 109–116. [CrossRef]
53. Li, Y.; Chen, Q.; Zhao, H.; Wang, L.; Tao, R. Variations in PM_{10}, $PM_{2.5}$ and $PM_{1.0}$ in an Urban Area of the Sichuan Basin and Their Relation to Meteorological Factors. *Atmosphere* **2015**, *6*, 150–163. [CrossRef]
54. Li, T.; Wang, H.; Zhao, T.; Xue, M.; Wang, Y.; Che, H.; Jiang, C. The Impacts of Different PBL Schemes on the Simulation of $PM_{2.5}$ during Severe Haze Episodes in the Jing-Jin-Ji Region and Its Surroundings in China. *Adv. Meteorol.* **2016**, *2016*, 6295878. [CrossRef]
55. Li, R.; Cui, L.; Fu, H.; Meng, Y.; Li, J.; Guo, J. Estimating high-resolution PM1 concentration from Himawari-8 combining extreme gradient boosting-geographically and temporally weighted regression (XGBoost-GTWR). *Atmos. Environ.* **2020**, *229*, 117434. [CrossRef]
56. Huang, F.; Zhou, J.; Chen, N.; Li, Y.; Li, K.; Wu, S. Chemical characteristics and source apportionment of $PM_{2.5}$ in Wuhan, China. *J. Atmos. Chem.* **2019**, *76*, 245–262. [CrossRef]