Short period PM$_{2.5}$ prediction based on multivariate linear regression model

Rui Zhao[1], Xinxin Gu[1], Bing Xue[2]*, Jianqiang Zhang[1], Wanxia Ren[3]

1 Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China, **2** Institute for Advanced Sustainability Studies e. V., Potsdam, Germany, **3** Key Lab of Pollution Ecology and Environmental Engineering, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China

* bing.xue@iass-potsdam.de

Abstract

A multivariate linear regression model was proposed to achieve short period prediction of PM$_{2.5}$ (fine particles with an aerodynamic diameter of 2.5 μm or less). The main parameters for the proposed model included data on aerosol optical depth (AOD) obtained through remote sensing, meteorological factors from ground monitoring (wind velocity, temperature, and relative humidity), and other gaseous pollutants (SO$_2$, NO$_2$, CO, and O$_3$). Beijing City was selected as a typical region for the case study. Data on the aforementioned variables for the city throughout 2015 were used to construct two regression models, which were discriminated by annual and seasonal data, respectively. The results indicated that the regression model based on annual data had ($R^2 = 0.766$) goodness-of-fit and ($R^2 = 0.875$) cross-validity. However, the regression models based on seasonal data for spring and winter were more effective, achieving 0.852 and 0.874 goodness-of-fit, respectively. Model uncertainties were also given, with the view of laying the foundation for further study.

Introduction

With the rapid economic development of China, air pollutants are also growing rapidly in recent decades, which became one of the country's most serious environmental issues, and attracted increasing public attention [1–3]. PM$_{2.5}$ refers to fine particulate matter with aerodynamic diameters equal to or smaller than 2.5 microns [4], which is recognized as a major component for air pollution, which has been shown to lead to multiple adverse health outcomes [5]. A number of epidemiological studies have indicated that long-term exposure to air containing high PM$_{2.5}$ concentrations may increase incidences of respiratory and cardiovascular diseases, and even result in death [6–8].

Generally, PM$_{2.5}$ is monitored by ground stations, and the coverage is extended from individual points to broader planes via spatial interpolation methods such as nearest-neighbour [9] and kriging [10]. However, the monitoring results may contain uncertainties due to the limited number and uneven distribution of ground monitoring stations and sampling points for spatial interpolation [11]. To compensate for this information gap, satellite remote sensing is gradually being applied to the monitoring of air quality [12]. The aim is to establish a quantitative relationship between Aerosol Optical Depth (AOD) obtained by satellite observations

and PM$_{2.5}$ concentrations [13]. This facilitates real-time and continuous monitoring of air quality for specific regions [14].

AOD is a multiphase system formed by gases, liquid, and solid particles suspended in the atmosphere, at a scale ranging from 10^{-3} to 10^2 microns [15]. AOD reflects the amount of transmittance through a unit section of an atmospheric column [13]. PM$_{2.5}$ is a major component of aerosol, and is suspended in the air under a dispersed phase [16]. Studies have shown a linear correlation between PM$_{2.5}$ and AOD [17,18].

In earlier studies, this relationship was generally established via a simplistic linear regression model between PM$_{2.5}$ and AOD, with the understanding that PM$_{2.5}$–AOD is stable within a certain spatiotemporal range, given as follows [19–24]:

$$PM_{2.5} = \alpha + \beta \times AOD \tag{1}$$

where PM$_{2.5}$ refers to its concentration near the ground (μg/m^3), which can be measured by using the taper element oscillating microbalance (TEOM); AOD is the aerosol optical thickness (dimensionless); and α and β represent the intercept and slope, respectively.

The proposal by van Donkelaar et $al.$ (2010) [17] used the ratio between PM$_{2.5}$ and AOD for PM$_{2.5}$ prediction. This method contained uncertainties because of insufficient PM$_{2.5}$ data from ground monitoring [10]. After studying the relationship between aerosol and air quality in Beijing from 2005 to 2014, Chen et $al.$ [25] found that there was no significant nor consistent correlation between the two, with disparities being especially large for winters and summers. This indirectly showed that improvements are needed for models that use AOD as the main variable for the prediction of PM$_{2.5}$ concentrations.

Based on this premise, a number of studies have delineated the PM$_{2.5}$–AOD relationship through the introduction of concomitant variables in the form of meteorological parameters, such as boundary layer height, temperature, relative humidity, and wind velocity, since the matting property of particles can drastically affect the degree of vertical mixing and increase the moisture absorption of aerosol [18, 26–28]. After modification of Eq (1), a generic model was given as follows [28]:

$$PM_{2.5} = (\alpha + \varepsilon_1) + (\beta_1 + \varepsilon_2) \times AOD + (\beta_2 + \varepsilon_3) \times TEMP + (\beta_3 + \varepsilon_4) \times RH + (\beta_4 + \varepsilon_5) \\ \times SPD \tag{2}$$

where TEMP is temperature (˚C); RH is relative humidity (%); SPD is wind velocity (m/s); α and β are fixed coefficients; and ε is a random error.

Related research has shown that organic carbides, ammonium nitrates, and sulphates are major components of PM$_{2.5}$ [29, 30]. In addition, under certain environmental conditions, the Air Quality Index (AQI) monitoring indicators such as SO$_2$, NO$_2$, and CO may convert into important precursors that form PM$_{2.5}$ [31]. Based on the data from monitoring stations in 31 Chinese cities between 2013 and 2014, Xie et $al.$ [32] found a moderate correlation between PM$_{2.5}$ and the heights at which SO$_2$, NO$_2$, and CO were present, but the correlation with the presence of O$_3$ was weak.

Our literature review revealed that when modelling the relationship between these factors and the changes in PM$_{2.5}$ mass concentrations, quite a few studies comprehensively considered the synergistic effects of AOD, meteorological parameters, and gaseous pollutants. In that context, this study aims to establish a quantitative model that would allow continuous monitoring of PM$_{2.5}$ to be conducted more effectively, and that would provide insight into the spatio-temporal distribution of PM$_{2.5}$. Herein, the multivariate linear regression method is applied, with PM$_{2.5}$ concentration as the dependent variable, and the following as variables: AOD data,

meteorological parameters (wind velocity, temperature, and relative humidity), and physical and chemical factors (SO$_2$, NO$_2$, CO, and O$_3$).

In addition, to improve the R^2, by means of advanced statistical models such as generalized additive regression, geographically weighted regression, and land use regression [33–35]. These studies aimed to improve model accuracy or land use information (such as altitude, population, and vegetation coverage). However, these methods could not reflect the constituent components of PM$_{2.5}$ [36].

This paper will elaborate on the construction of the proposed model. First of all, Beijing, the capital city of China, was used as a typical case study, and all related data needed by the model for the year 2015 were collected. And then, fitting and cross-validation results of the model were obtained for the entire year and the respective seasons, before model uncertainties were discussed. The conclusion was that the constructed model could be an effective means to supplement ground monitoring for PM$_{2.5}$ prediction. Nevertheless, there were inadequacies in the study that required further improvement.

Methods and data source

In addition to meteorological factors, there may be different degrees of correlation between the PM$_{2.5}$ and SO$_2$, NO$_2$, CO, and O$_3$ [32, 37–41]. In order to verify the rationality of this conclusion, this study attempted to modify Eq (2) and construct a multivariate linear regression model. The concomitant variables of the model are the meteorological parameters and four types of pollutant indices (SO$_2$, NO$_2$, CO, and O$_3$):

$$
\begin{aligned}
PM_{2.5} = {} & (\alpha + \varepsilon_1) + (\beta_1 + \varepsilon_2) \times AOD + (\beta_2 + \varepsilon_3) \times TEMP + (\beta_3 + \varepsilon_4) \times RH + (\beta_4 + \varepsilon_5) \\
& \times SPD + (\beta_5 + \varepsilon_6) \times CO + (\beta_6 + \varepsilon_7) \times NO_2 + (\beta_7 + \varepsilon_8) \times SO_2 + (\beta_8 + \varepsilon_9) \\
& \times O_3
\end{aligned}
\tag{3}
$$

where PM$_{2.5}$ is its mass concentration at ground level (μg/m^3), AOD is derived from MODIS (dimensionless), TEMP is temperature (˚C); RH is relative humidity (%), *SPD* is wind velocity (m/s); and SO$_2$, NO$_2$, CO, and O$_3$ are the mass concentrations of the four pollutants at ground level, $\beta_1, \beta_2, \ldots \beta_8$ are the slopes corresponding to the respective variables; and $(\alpha + \varepsilon_1)$ is the intercept.

Beijing City, which was taken as a typical case study for analysis, has seven national ground monitoring stations, including West Wanshou Nishinomiya, Temple of Heaven, Dongsi Subdistrict, Xicheng District, Agricultural Exhibition Center, The Institute of Atmospheric Physics, Olympic Centre. The data from these stations throughout the year 2015, specifically during the time period that satellite transits (i.e., AM 9:00, 10:00, 11:00, and 12:00) were obtained from the China National Environmental Monitoring Centre (http://www.cnemc.cn/), containing hourly mass concentrations of PM$_{2.5}$ and the four major air pollutants NO$_2$, CO, SO$_2$, and O$_3$, as well as meteorological data (including temperature, wind velocity, and relative humidity). The above data were averaged in each time node of the 4 hours respectively, to be set as the representative values for the model construction and validation. Their means, standard deviations (SD), and minimum and maximum values are shown in Table A in S1 Appendix.

In addition to ground data, this study also acquired AOD product data from the Aqua-MODIS 550 nm Collection 6. MODIS is a medium-resolution imaging spectrometer carried on the Terra and Aqua satellites of the United States' Earth Observing System, and provides daily aerosol data worldwide [42]. The standard MODIS Level-2 (L2) product provides AOD data at 10 km spatial resolution, while the resolution of the newly-released MODIS Collection 6 (C6) product (MYD04_3K) is 3 km. In addition, the MODIS C6 product has been improved

in various ways, including instrument calibration, cloud detection, the structure of the lookup table, calculation of radiation transmission, and corrections to gas absorption [43].

First, MYD04_3K data from the Aqua-MODIS 550 nm L2 Aerosol Products for the period 1 January to 31 December 2015 were downloaded through the MODIS L1 Level 1 and Atmosphere Archive and Distribution System (LAADS) (http://ladsweb.nascom.nasa.gov/). These data were verified by using the Beijing observation station's AOD data obtained from the Aerosol Robotic Network (AERONET) (http://aeronet.gsfc.nasa.gov/). To ensure comparability between the MODIS 550 nm data and the corresponding AOD data, AERONET's AOD data corresponding to the wavelengths 440 nm and 675 nm were interpolated to obtain the AOD value corresponding to 550 nm. The results showed a strong correlation between Aqua-MODIS 3 km AOD and AERONET AOD, with the Pearson correlation coefficient r between the two AODs being 0.92 (Fig A in S1 Appendix). The slope between the two was 1.23. This indicated that, overall, the values of MODIS AOD were higher than those of AERONET AOD. However, the deviations between the two could be considered as systematic because these were holistic and continuous throughout the entire data range [24]. As such, there would not be any impact on the prediction of PM$_{2.5}$ concentrations.

Due to extreme weather, e.g., impact from thick clouds, strong snowfall etc. on the regional atmospheric environment, data may have deficiencies in any of the variables related to the proposed linear regression model. For example, data missing may happen to PM2.5, AOD, or the gaseous pollutants. By the reason that satellite transits within a specific time period, i.e., AM 9:00, 10:00, 11:00, and 12:00 every day in the morning, it is impossible to obtain the AOD data during the whole day. To ensure the data availability and consistency for model construction and validation, the first step is to screen out the complete dataset without deficiency for all the variables, compliance with the data capacity of AOD. As each variable has different unit, the second step is data normalization by using SPSS software to ensure them dimensionless. After data compilation and screening, 954 sets applicable to all variables were retained. To use the available data as much as possible, their ratios for the four seasons were set as the basis for stratified sampling. Two-thirds of the data were used for regression analysis at the seasonal level, while the remaining third were used for cross-validation of the model, giving 636 and 318 sets of data for annual regression analysis and model checking, respectively. Taking into account the seasonal factor, random selections were made from the 217, 239, 228, and 270 sets of data for spring, summer, autumn, and winter, respectively.

Results and discussion

The study first gives the regression result by using the annual data of Beijing City. As the available data related to PM$_{2.5}$ have expressed seasonal differences, they have been fitted for each season to examine their respective regression performances. The coefficients corresponding to the different regression models are shown in Table 1.

Table 1. The coefficients estimated in the regression models.

	Constant	AOD	Temp	RH	SPD	CO	NO$_2$	SO$_2$	O$_3$
Annual	-104.046	12.21	-0.364	0.507	4.086	20.665	1.818	-0.173	0.611
Spring	-141.253	9.994	0.264	0.792	2.042	49.554	1.329	-0.379	1.485
Summer	-15.458	10.930	-0.900	-0.127	2.171	2.98	1.045	3.189	0.147
Autumn	-27.879	7.487	-1.258	-0.102	3.043	96.221	0.777	-3.544	0.265
Winter	-137.028	21.349	0.723	0.522	4.726	42.947	1.486	-1.998	0.994

https://doi.org/10.1371/journal.pone.0201011.t001

Regression results

The scatter distributions for the fitting and cross-validation of Beijing City PM$_{2.5}$ data for 2015 are shown in Figs 1 and 2. The fitted line is generated by Excel software packaging, which is based upon the least squares method to find out the linear trend with the best fitness among the scattered points. The R^2 and root mean square error (RMSE) for the regression model of annual data (Fig 1) were 0.766 and 30.271 μg/m^3, respectively, and for cross-validation (Fig 2), the R^2 and RMSE were 0.875 and 23.423 μg/m^3, respectively, representing an increase of 14.2% and a decrease of 22.6%, respectively, compared to the regression model of annual data.

A better R^2 results for cross-validation compared to the regression model usually might be due to the uneven distribution of sample data, which is inherent in the respective datasets for the four seasons that were used for model construction. Furthermore, the overall PM$_{2.5}$ fluctuations between the seasons were large, such that the random selection of test data could not fully guarantee that data distribution was proportional to the actual scenarios for the four seasons.

Beijing is located in a mid-latitude region and has a temperate monsoon climate characterized by cold winters, hot summers, and distinct seasonal characteristics [44]. After the PM$_{2.5}$ monitoring data were further grouped by the four seasons (Spring: January–March; Summer: April–June; Autumn: July–September; and Winter: October–December), the seasonality of the data was apparent. As can be seen in Table 2, the mean PM$_{2.5}$ concentrations for spring and winter far exceeded those of summer and autumn. This observation prompted data fitting for each season, to examine the respective performance of regression. Table 3 shows that the regression results for the PM$_{2.5}$ data in the four seasons revealed high fitness for spring and winter, both with R^2 greater than 0.85; however, that for summer was low, with R^2 of only
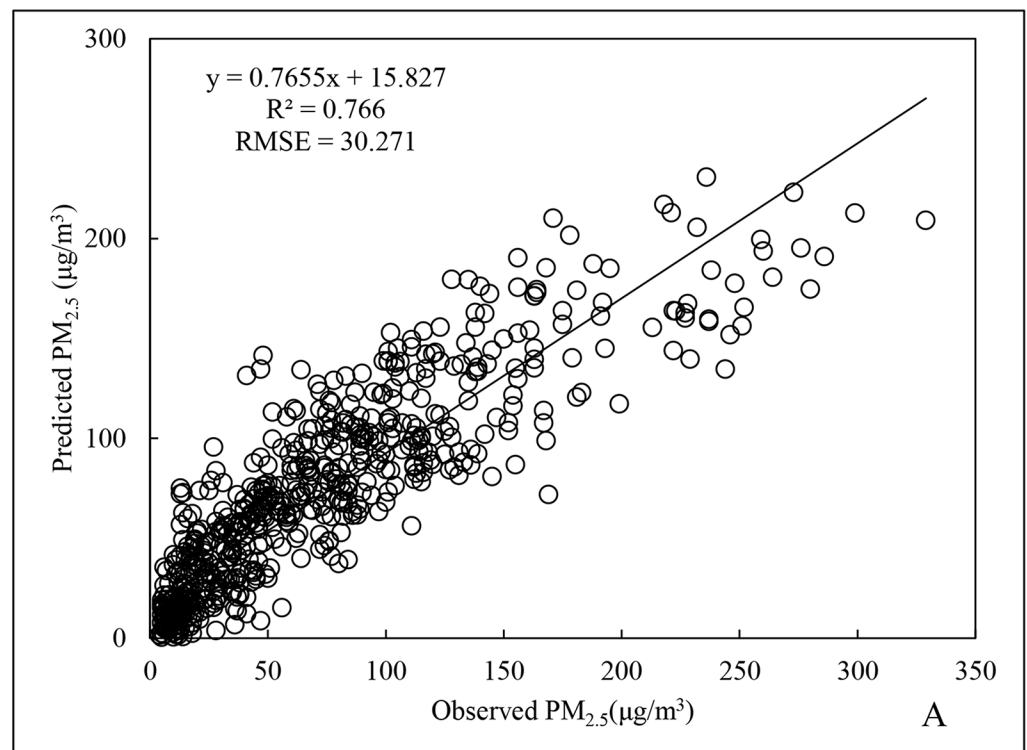

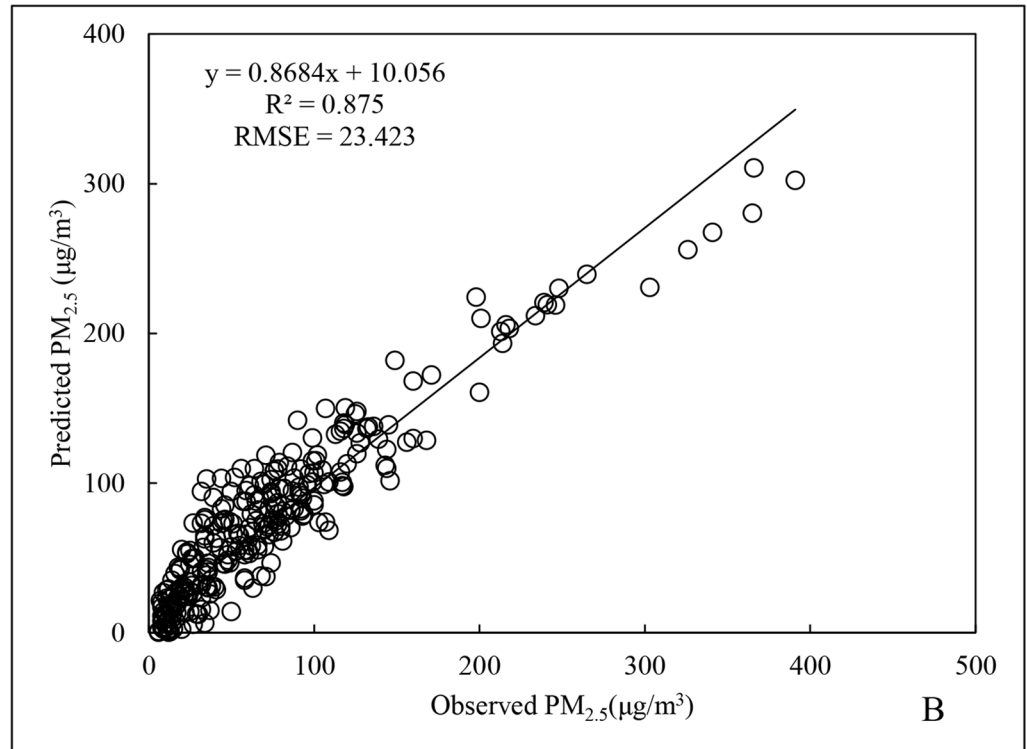**Fig 1. Fitting results of annual PM$_{2.5}$ data for 2015.**

Fig 2. Cross-validation results of annual PM$_{2.5}$ data for 2015.

0.761. Comparison indicated that, with the exception of the data for summer, the fitting results and prediction validity for the other three seasons were all better than that of the annual data. This confirmed that the main factor affecting prediction validity was seasonal variations.

For spring, the R^2 for the PM$_{2.5}$ data was 0.852 (Fig 3), representing an increase of 11.2% compared to the annual data. For RMSE, there was a decrease of 1.5% instead. The R^2 for cross-validation was 0.822 (Fig 4), 3.5% lower than that for fitting of the spring data, indicating slight overfitting. Nevertheless, the R^2 was still better than that for the annual data. This indirectly demonstrates that the regression model for the spring season data could accurately reflect changes in PM$_{2.5}$ over that time period.

For the summer PM$_{2.5}$ data, R^2 was 0.761and RMSE was 17.977 µg/m^3 (Fig 5), being 2.4% and 40.6% lower than the results for the annual data, respectively. In addition, the R^2 for cross-validation (Fig 6) of the summer data was 18.8% less than that for the overall data fitting, indicating slight overfitting. From the Table 1, it is apparent that the associated coefficient of carbon monoxide (CO) in summer model was much smaller than in other models. This indicated that CO had little influence on PM$_{2.5}$ in summer, to result in an inferior performance on the

Table 2. Descriptive statistics and summary of regression results.

Seasonal PM$_{2.5}$	N	Mean	SD	Min	Max
Annual	954	76.246	74.112	3	479
Spring	217	90.443	78.505	4	425
Summer	239	65.319	49.678	5	260
Autumn	228	49.639	38.364	4	199
Winter	270	100.688	102.127	3	479

Table 3. Summary of R² and error measure for fitting and cross-validation result.

Seasonal Regression	Fitting results		Cross-validation results	
	R²	RMSE	R²	RMSE
Annual	0.766	30.271	0.875	23.423
Spring	0.852	29.802	0.822	21.718
Summer	0.761	17.977	0.618	21.082
Autumn	0.788	18.721	0.803	16.190
Winter	0.874	25.692	0.940	30.449

summer prediction. The possible reason might be a rapid conversion of gaseous pollutants to form nitrates and sulphates, due to high summer temperatures in Beijing and longer duration of direct sunlight. Consequently, the correlation between the variable of gaseous pollutants (CO, NO$_2$, SO$_2$, and O$_3$) and PM$_{2.5}$ was weakened [41].

During the autumn, the value of R² regarding PM$_{2.5}$ data is 0.788 (Fig 7), 2.9% higher than that of annual value. In contrast, the RMSE was 18.721 μg/m³, being 38.2% lower than the annual data and similar to the fitting performance for the summer. The obvious reduction in RMSE might be because the mean and maximum PM$_{2.5}$ concentrations for autumn were far lower than those for both in spring and winter. The R² for cross-validation was 0.803 (Fig 8), which was slightly better than the regression result. This provides indirect evidence that the autumn data could better predict PM$_{2.5}$ concentrations during that period.

Regarding the winter season, the R² of the winter PM$_{2.5}$ data was 0.874 (Fig 9), 14.1% higher than for the annual data. Furthermore, the RMSE was 25.692 μg/m³, with a decrease of 15.1%. The R² for cross-validation of the winter PM$_{2.5}$ data reached 0.940 (Fig 10), indicating that
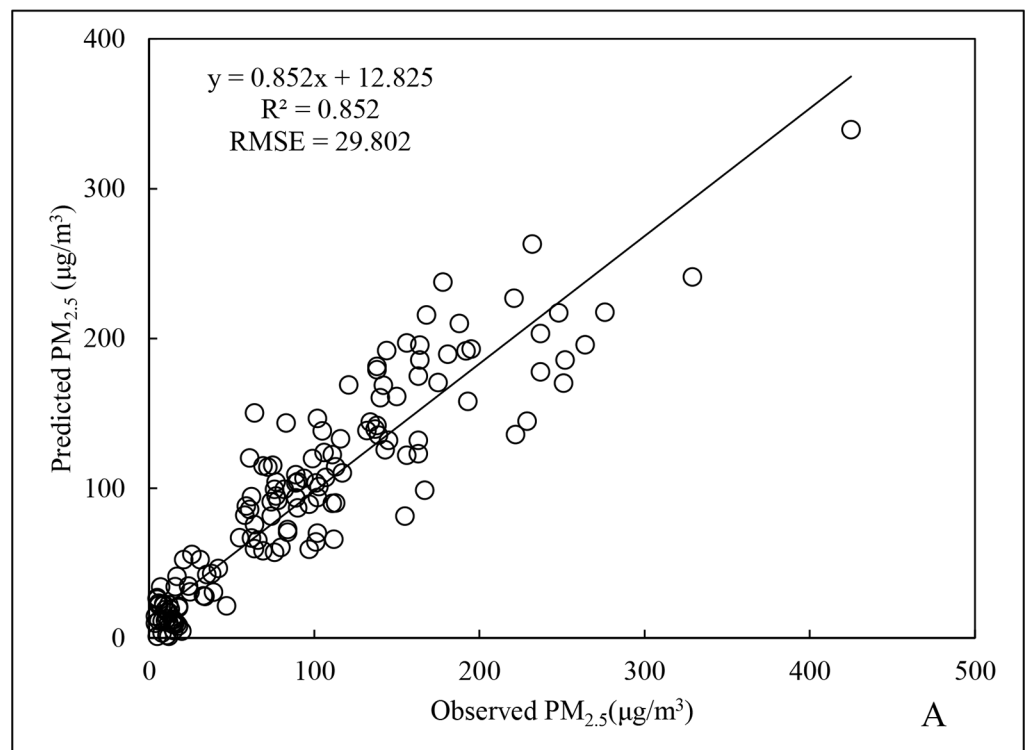

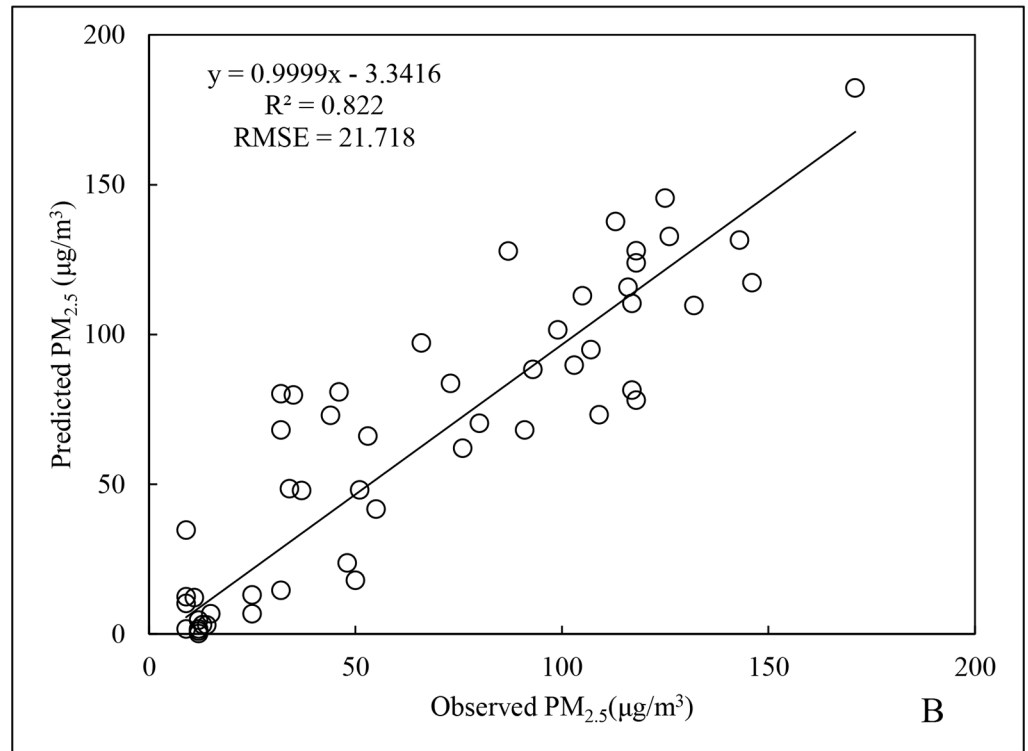**Fig 3. Fitting results of PM$_{2.5}$ data for spring 2015.**

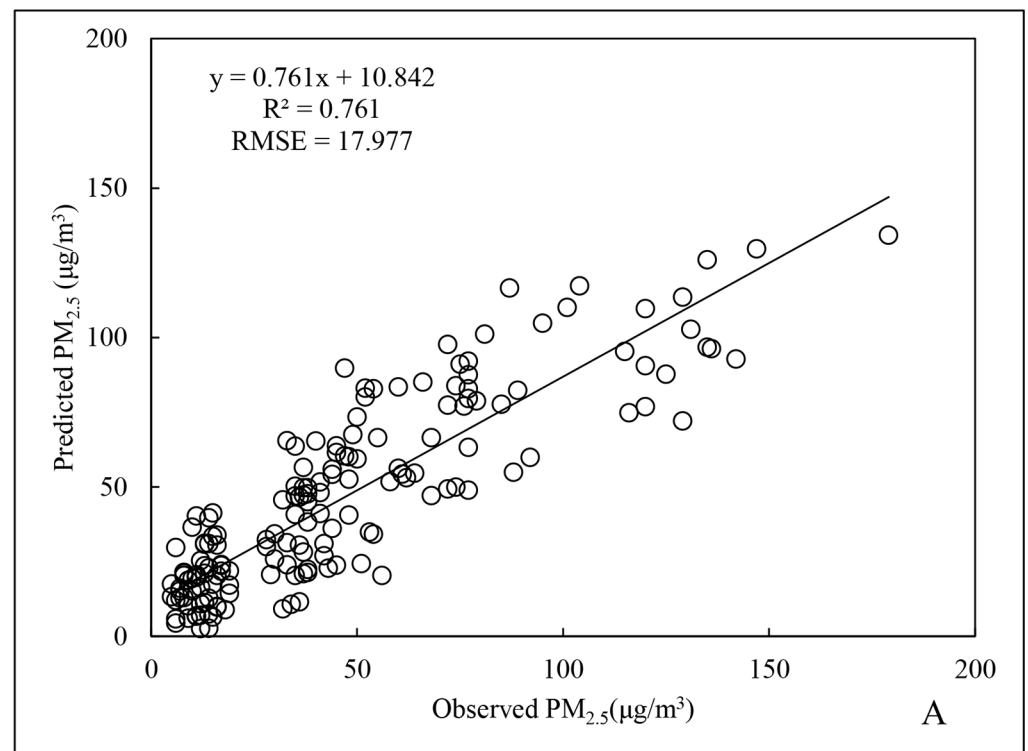**Fig 4. Cross-validation results of PM$_{2.5}$ data for spring 2015.**

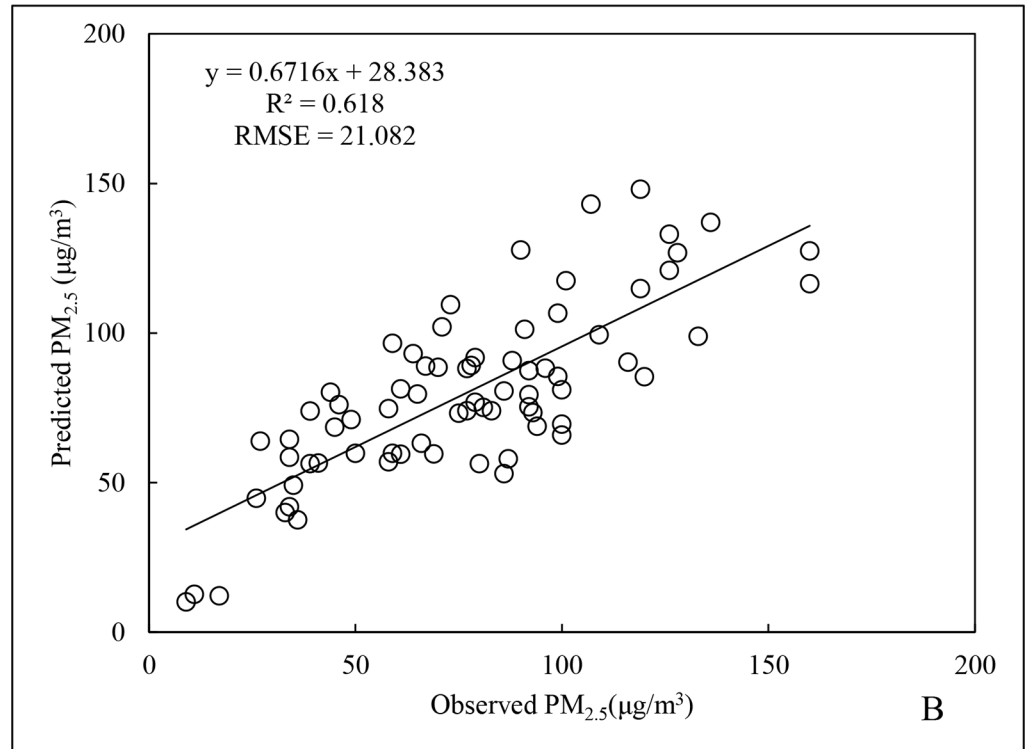**Fig 5. Fitting results of PM$_{2.5}$ data for summer 2015.**

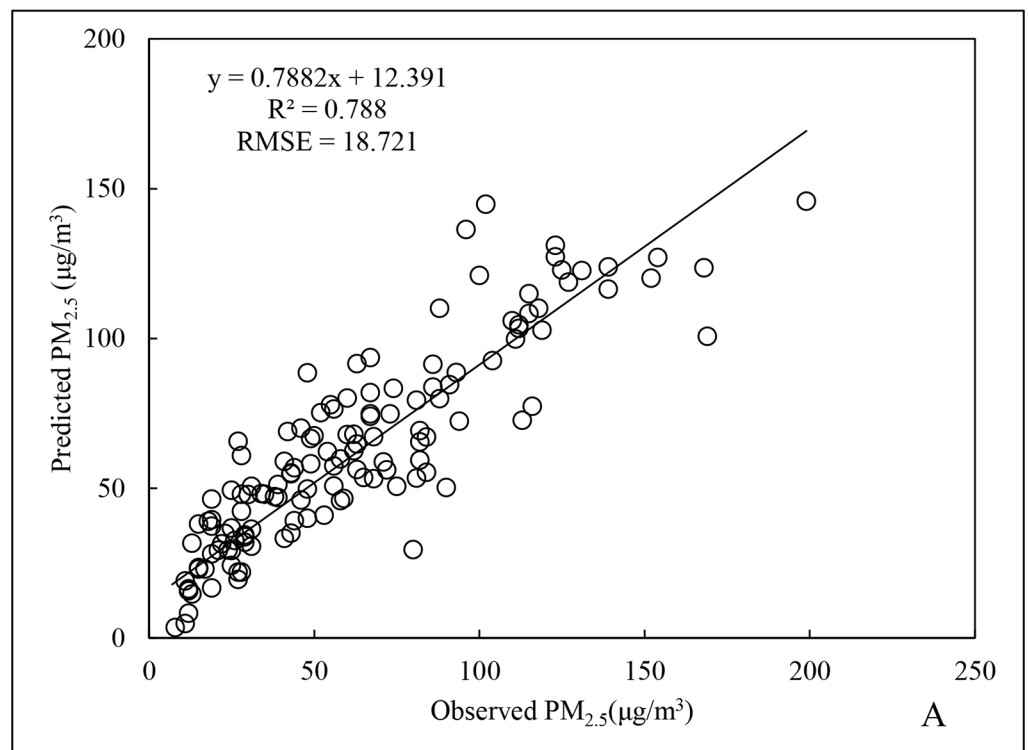**Fig 6. Cross-validation results of PM2.5 data for summer 2015.**

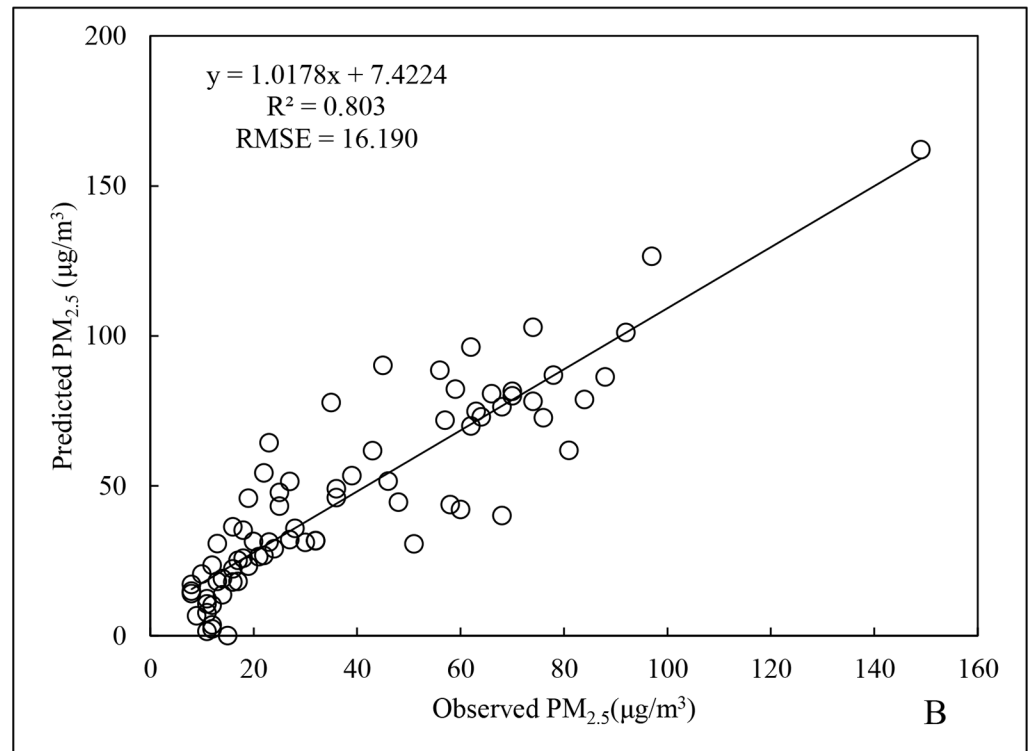**Fig 7. Fitting results of PM2.5 data for autumn 2015.**

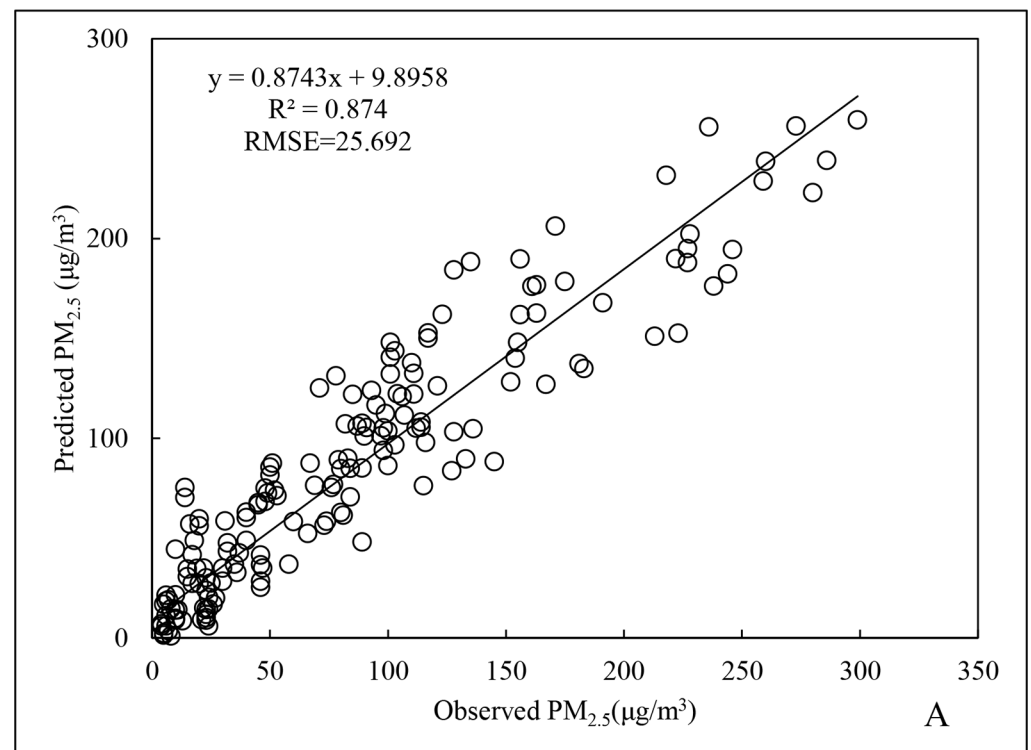**Fig 8. Cross-validation results of PM$_{2.5}$ data for autumn 2015.**

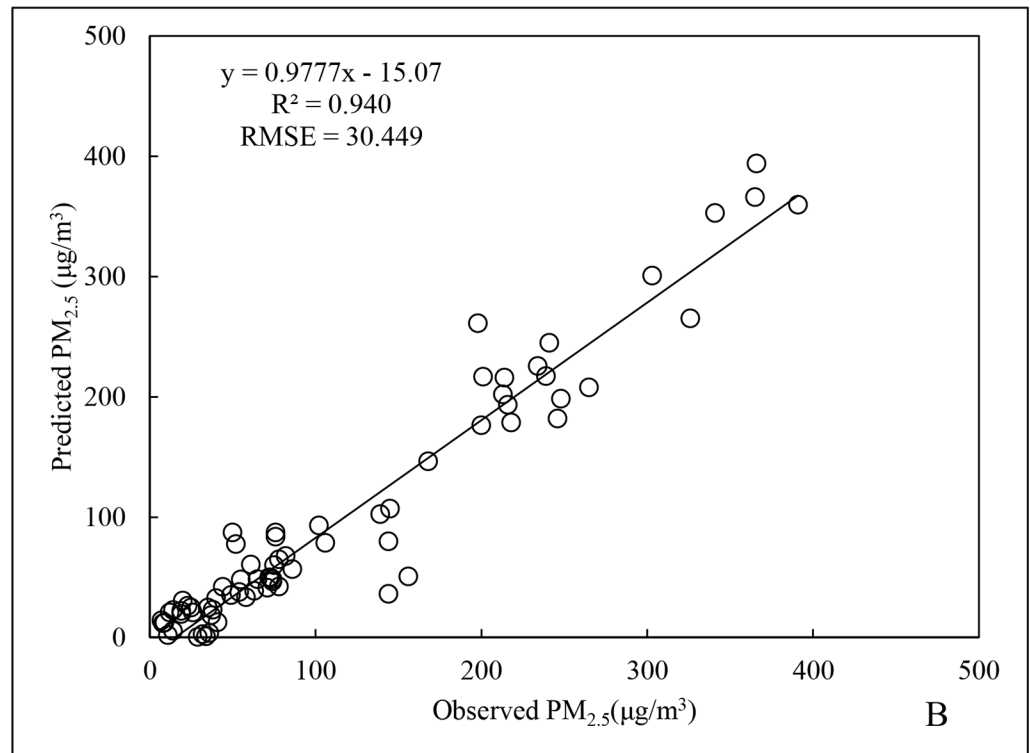**Fig 9. Fitting results of PM$_{2.5}$ data for winter 2015.**

Fig 10. Cross-validation results of PM$_{2.5}$ data for winter 2015.

https://doi.org/10.1371/journal.pone.0201011.g010

prediction validity was good. The reason would be that the mean and maximum PM$_{2.5}$ concentrations in winter were the highest for the entire year. Meteorological conditions were stable during this period, without conducive to diffusion, resulting in good spatiotemporal stability in the data for the various parameters and variables [24].

Discussions

Effective PM$_{2.5}$ prediction is a complex issue that is easily affected by numerous factors, including weather and climatic conditions, and environmental seasonality. Different factors also have different degrees of impact on the regression results used for PM$_{2.5}$ prediction. The correlation between PM$_{2.5}$ and the other parameters and variables is shown in Table 4, and the pairwise Pearson correlations among the parameters are given in Table B in S1 Appendix. During the study period, there was an extremely strong correlation between the mean concentration of PM$_{2.5}$ and those of CO, NO$_2$, and SO$_2$, and a strong correlation between the former and AOD (highlighted in blue colour). The results reflected, to a certain extent, the rationality of the variables selected for constructing the regression model. This was especially so for regression analysis between PM$_{2.5}$ and AOD: the model's predictive validity improved substantially after inclusion of the four synergistic variables CO, SO$_2$, NO$_2$, and O$_3$ to capture the contribution of gaseous pollutants to PM$_{2.5}$ formation.

Table 4. Pearson correlation between PM$_{2.5}$ and the various parameters.

	PM$_{2.5}$	AOD	Temp	RH	SPD	CO	NO$_2$	SO$_2$	O$_3$
PM$_{2.5}$	1.000	0.416	-0.176	0.276	-0.134	0.784	0.772	0.500	-0.329

https://doi.org/10.1371/journal.pone.0201011.t004

Separately, a comparative analysis was made between the proposed regression model and those used by other scholars. One example was the geographical weighted regression model used by Ma *et al.* [10] to predict PM$_{2.5}$ in China, which had an R^2 of 0.71. The prediction accuracy of the proposed model was higher (R^2 = 0.76) than other models because the data resolution of 3 km retained much more information than studies that used data at 50 km resolution. However, this study only predicted PM$_{2.5}$ for Beijing City, but did not take into consideration any regional differences in the atmospheric environment. To this end, it is critical to consider the spatial heterogeneity presented by PM$_{2.5}$, as well as the impact of this heterogeneity when enhancing the model's accuracy. The linear mixed-effect model proposed by Zheng *et al.* [45] used Beijing–Tianjin–Hebei as the study area and had an accuracy and cross-validation of R^2 = 0.77 and 0.84, respectively. Although the fitting results were good, the model did not consider seasonal differences. Lv *et al.* [46] predicted the surface concentrations of PM$_{2.5}$ in northern China (including Beijing, Tianjin, Hebei, and Shandong) using a Bayesian hierarchical model. The R^2 of 0.78 was slightly better than our proposed model. A possible reason was that more AOD data were available for their use, which expanded the training set of their model, thus improving its prediction accuracy. In addition, their study took into consideration the effects of seasonal differences on PM$_{2.5}$. This indirectly verified the rationale of our study, in which the individual seasons were used for model construction.

Although the model proposed here had higher predictive accuracy (to a certain extent) than earlier models constructed by other scholars, few uncertainties are remained. First, uncertainties in the PM$_{2.5}$ data sources: on the one hand, this was a reflection of the uneven spatial distribution of PM$_{2.5}$ ground monitoring stations, which were mainly concentrated in cities and urban areas, but lacking in the suburbs; and on the other hand, it revealed the systemic deviations inherent in the TEOM method of measuring PM$_{2.5}$ concentrations [47]. Second, uncertainties in the AOD data: Although the quality of AOD products at 3 km resolution was relatively high, these were more prone to generating random noise, which affected prediction accuracies [43]. Third, uncertainties in the proposed model itself: This study only assumed a possible linear relationship between PM$_{2.5}$ and the numerous factors but did not consider the formation mechanism of PM$_{2.5}$, which would have an impact on the effectiveness of the model. And the fourth is the uncertainty caused by spatiotemporal heterogeneities. Depending on the region and time period, significant differences exist in PM$_{2.5}$. The good predictions achieved by the model proposed here were limited to a region and over short duration. Further examination would be needed to determine whether the model could be applied to PM$_{2.5}$ predictions at larger geographical scales and temporal dimensions.

Conclusions

A multivariate linear regression equation was developed between PM$_{2.5}$ and AOD, meteorological parameters, and various gaseous pollutants. The aim was to overcome the inadequacies in spatiotemporal observations by ground monitoring stations. The results showed that the regression model using annual data for Beijing City in 2015 could explain 76.6% of the city's PM$_{2.5}$ concentrations. Apparent seasonal differences in PM$_{2.5}$ concentrations were found, with R^2 values of 0.852, 0.761, 0.788, and 0.874 for models utilizing the data for spring, summer, autumn, and winter seasons, respectively. The results of the regression models that used spring and winter data were superior to those that used summer and autumn data. Further studies will investigate the effectiveness of the proposed model. These will include the model's sensitivity to changes in time and region, and the effects of different resolutions of satellite-acquired AODs on prediction accuracy. The constituent components and formation of PM$_{2.5}$ will also be analysed, so that the model can be further improved to validate its prediction.

Supporting information

S1 Appendix.
(DOCX)

Author Contributions

Conceptualization: Bing Xue.

Data curation: Rui Zhao, Jianqiang Zhang.

Formal analysis: Rui Zhao, Xinxin Gu.

Funding acquisition: Rui Zhao, Bing Xue.

Methodology: Xinxin Gu, Jianqiang Zhang.

Project administration: Rui Zhao.

Software: Xinxin Gu, Jianqiang Zhang.

Validation: Rui Zhao, Bing Xue, Wanxia Ren.

Visualization: Xinxin Gu, Jianqiang Zhang.

Writing – original draft: Rui Zhao, Bing Xue, Jianqiang Zhang, Wanxia Ren.

Writing – review & editing: Bing Xue.

References

1. Chang Y. China needs a tighter PM$_{2.5}$ limit and a change in priorities. Environ Sci Technol. 2012; 46 (13): 7069–7070. https://doi.org/10.1021/es3022705 PMID: 22716938

2. Yuan Y, Liu S, Castro R and Pan X. PM$_{2.5}$ monitoring and mitigation in the cities of China. Environ Sci Technol. 2012, 46(7): 3627–3628. https://doi.org/10.1021/es300984j PMID: 22448594

3. Guan D, Su X, Zhang Q, Peters G P, Liu Z, Lei Y, et al. The socioeconomic drivers of China's primary PM$_{2.5}$ emissions. Environ Res Lett. 2014; 9(2): 024010.

4. Walsh M P. PM$_{2.5}$: global progress in controlling the motor vehicle contribution front. Environ Sci Engineer. 2014; 8(1): 1–17.

5. Xue B, Butler T, Ren W, Zhang ZL, Wang Y, Mu ZL. Reviewing Air Pollution and Public Health in China. Proceedings of the Institution of Civil Engineers—Engineering Sustainability. 2017; 0 0: 0, 1–10.

6. Zhou Y, Fu J S, Zhuang G and Levy J I. Risk-based prioritization among air pollution control strategies in the Yangtze River Delta, China. Environ Health Persp. 2010; 118: 1204–1210.

7. Leitte A M, Schlink U, Herbarth O, Wiedensohler A, Pan X C, Hu M, et al. Size-segregated particle number concentrations and respiratory emergency room visits in Beijing, China. Environ Health Persp. 2011; 119(4): 508–513.

8. Langrish J P, Li X, Wang S F, Lee M M Y, Barnes G D, Miller M R, et al. Reducing personal exposure to particulate air pollution improves cardiovascular health in patients with coronary heart disease. Environ Health Persp. 2012; 120(2): 367–372.

9. Just A C, Wright R O, Schwartz J, Coull B A, Baccarelli A A, Tellez-Rojo, et al. Using high-resolution satellite aerosol optical depth to estimate daily PM$_{2.5}$ geographical distribution in Mexico City. Environ Sci Technol. 2015; 49(14): 8576–8584. https://doi.org/10.1021/acs.est.5b00859 PMID: 26061488

10. Ma Z, Hu X, Huang L, Bi J and Liu Y. Estimating ground-level PM$_{2.5}$ in China using satellite remote sensing. Environ Sci Technol. 2014; 48(13): 7436–7444. https://doi.org/10.1021/es5009399 PMID: 24901806

11. Qu L, Xiao H, Zheng N, Zhang Z and Xu Y. Comparison of four methods for spatial interpolation of estimated atmospheric nitrogen deposition in South China. Environ Sci Pollut Res. 2017; 24(3): 2578–2588.

12. Hoff R M, Christopher S A. Remote sensing of particulate pollution from space: Have we reached the promised land? J Air Waste Manage Assoc. 2009; 59(6): 645–675.

13. Hu X, Waller L A, Lyapustin A, Wang Y, Al-Hamdan M Z, Crosson W L, et al. Estimating ground-level PM$_{2.5}$ concentrations in the southeastern United States using MAIAC AOD retrievals and a two-stage model Remote Sens Environ. 2014; 140: 220–232.

14. Li J, Carlson B E and Lacis AA. How well do satellite AOD observations represent the spatial and temporal variability of PM$_{2.5}$ concentration for the United States? Atmos Environ. 2015; 102: 260–273.

15. Grigas T. Remote Sensing and In-Situ Characterisation of Atmospheric Aerosol Pollution. Ph.D. Thesis. School of Physics, Centre for Climate & Air Pollution Studies; The National University of Ireland Galway. 2017.

16. Pavese G, Lettino A, Calvello M, Esposito F and Fiore S. Aerosol composition and properties variation at the ground and over the column under different air masses advection in South Italy. Environ Sci Pollut Res. 2016; 23(7): 6546–6562.

17. van Donkelaar A, Martin R V, Brauer M, Kahn R, Levy R, Verduzco C, et al. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. Environ Health Persp.2010; 118(6): 847.

18. Lee H J, Liu Y, Coull B A, Schwartz J and Koutrakis P. A novel calibration approach of MODIS AOD data to predict PM$_{2.5}$ concentrations. Atmos Chem Phys. 2011; 11: 7991–8002.

19. Engel-Cox J A, Holloman C H, Coutant B W and Hoff R M. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. Atmos Environ. 2004; 38(16): 2495–2509.

20. Liu Y, Park R J, Jacob D J, Li Q B, Kilaru V and Sarnat J A. Mapping annual mean ground-level PM$_{2.5}$ concentrations using multiangle imaging spectroradiometer aerosol optical thickness over the contiguous United States. J Geophys Res. 2004; 109: D22206.

21. Zhang H, Hoff R M and Engel-Cox J A. The relation between moderate resolution imaging spectroradiometer (MODIS) aerosol optical depth and PM$_{2.5}$ over the United States: A geographical comparison by U.S. Environmental Protection Agency regions. J Air Waste Manage Assoc. 2009; 59(11): 1358–1369.

22. Schaap M, Apituley A, Timmermans R M A, Koelemeijer R B A and de Leeuw G. Exploring the relation between aerosol optical depth and PM$_{2.5}$ at Cabauw, the Netherlands. Atmos Chem Phys. 2009; 9(3): 909–925.

23. Boys B L, Martin R V, van Donkelaar A, MacDonell R J, Hsu N C, Cooper M J, et al. Fifteen-year global time series of satellite-derived fine particulate matter. Environ Sci Technol. 2014; 48(19): 11109–11118. https://doi.org/10.1021/es502113p PMID: 25184953

24. Xie Y, Wang Y, Zhang K, Dong W, Lv B and Bai Y. Daily estimation of ground-level PM$_{2.5}$ concentrations over Beijing using 3 km resolution MODIS AOD. Environ Sci Technol. 2015b; 49(20): 12280–12288. https://doi.org/10.1021/acs.est.5b01413 PMID: 26310776

25. Chen W, Tang H, Zhao H and Yan L. Analysis of aerosol properties in Beijing based on ground-based sun photometer and air quality monitoring observations from 2005 to 2014. Remote Sens. 2016; 8(2): 110.

26. Kumar N, Chu A and Foster A. An empirical relationship between PM$_{2.5}$ and aerosol optical depth in Delhi Metropolitan. Atmos Environ. 2007; 41(21): 4492–4503.

27. Guo J P, Zhang X Y, Che H Z, Gong S L, An X Q, Cao C X, et al. Correlation between PM concentrations and aerosol optical depth in eastern China. Atmos Environ. 2009; 43(37): 5876–5886.

28. Song W, Jia H, Huang J and Zhang Y. A satellite-based geographically weighted regression model for regional PM$_{2.5}$ estimation over the Pearl River Delta region in China. Remote Sens Environ. 2014; 154: 1–7.

29. Zhuang X, Wang Y, He H, Liu J, Wang X, Zhu T, et al. Haze insights and mitigation in China: An overview. J Environ Sci. 2014; 26: 2–12.

30. Wang J, Song Y, Zuo J and Wu H. Compositions and pollutant sources of haze in Beijing urban sites. Environ Sci Pollut Res. 2016; 23(9): 8827–8836.

31. Rogula-Kozłowska W, Klejnowski K, Rogula-Kopiec P, Ośródka L, Krajny E, Błaszczak B, et al. Spatial and seasonal variability of the mass concentration and chemical composition of PM$_{2.5}$ in Poland. Air Qual Atmos Health. 2014, 7(1): 41–58.

32. Xie Y, Zhao B, Zhang L and Luo R. Spatiotemporal variations of PM$_{2.5}$ and PM$_{10}$ concentrations between 31 Chinese cities and their relationships with SO$_2$, NO$_2$, CO and O$_3$. Particuology. 2015a; 20: 141–149.

33. Liu Y, Paciorek C J, Koutrakis P. Estimating regional spatial and temporal variability of PM$_{2.5}$ concentrations using satellite data, meteorology, and land use information. Environ Health Persp. 2009; 117(6): 886–892.

34. Kloog I, Nordio F, Coull B A and Schwartz J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM$_{2.5}$ exposures in the Mid-Atlantic states. Environ Sci Technol. 2012; 46(21): 11913–11921. https://doi.org/10.1021/es302673e PMID: 23013112

35. Hu X, Waller L A, Al-Hamdan M Z, Crosson W L, Estes M G, Estes S M, et al. Estimating ground-level PM$_{2.5}$ concentrations in the southeastern US using geographically weighted regression. Environ Res. 2013; 121: 1–10.

36. Di Q, Koutrakis P and Schwartz J. A hybrid prediction model for PM$_{2.5}$ mass and components using a chemical transport model and land use regression. Atmos Environ. 2016; 131: 390–399.

37. Unger N, Shindell D, Koch D and Streets D. Cross influences of ozone and sulfate precursor emissions changes on air quality and climate. Proc Natl Acad Sci. 2006; 103(12): 4377–4380.

38. Rae J G L, Johnson C E, Bellouin N, Boucher O, Haywood J M and Jones A. Sensitivity of global sulphate aerosol production to changes in oxidant concentrations and climate. J Geophys Res. 2007; 112: D10312.

39. Kloster S, Dentener F, Feichter J, Raes F, van Aardenne J, Roeckner E, et al. Influence of future air pollution mitigation strategies on total aerosol radiative forcing. Atmos Chem Phys. 2008; 8: 6405–6437.

40. Shindell D, Lamarque J F, Unger N, Koch D, Faluvegi G, Bauer S, et al. Climate forcing and air quality change due to regional emissions reductions by economic sector. Atmos Chem Phys. 2008; 8: 7101–7113.

41. Leibensperger E M, Mickley L J, Jacob D J and Barrett S R. Intercontinental influence of NOx and CO emissions on particulate matter air quality. Atmos Environ. 2011; 45(19): 3318–3324.

42. Levy R C, Remer L A, Kleldman R G, Mattoo S, Ichoku C, Kahn R, et al. Global evaluation of the collection 5 MODIS dark-target aerosol products over land. Atmos Chem Phys. 2010; 10: 14815–14873.

43. Munchak L A, Levy R C, Mattoo S, Remer L A, Holben B N, Schafer J S, et al. MODIS 3 km aerosol product: applications over land in an urban/suburban region. Atmos Meas Tech. 2013; 6(7): 1747–1759.

44. He X, Shen S, Miao S, Dou J and Zhang Y. Quantitative detection of urban climate resources and the establishment of an urban climate map (UCMap) system in Beijing. Build Environ. 2015; 92: 668–678.

45. Zheng Y, Zhang Q, Liu Y, Geng G and He K. Estimating ground-level PM$_{2.5}$ concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. Atmos Environ. 2016; 24: 232–242.

46. Lv B, Hu Y, Chang H H, Russell A G and Bai Y. Improving the accuracy of daily PM$_{2.5}$ distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China. Environ Sci Technol. 2016; 50 (9): 4752–4759.

47. Engel-Cox J, Oanh N T K, van Donkelaar A, Martin R V and Zell E. Toward the next generation of air quality monitoring: Particulate matter. Atmos Environ.2013; 80: 584–590.