

RESEARCH ARTICLE

Estimating PM_{2.5} Concentrations in Xi'an City Using a Generalized Additive Model with Multi-Source Monitoring Data

Yong-Ze Song¹, Hong-Lei Yang^{1*}, Jun-Huan Peng¹, Yi-Rong Song², Qian Sun³, Yuan Li⁴

1 School of Land Science and Technology, China University of Geosciences, Beijing, China, **2** Department of Geological Engineering, Qinghai University, Xining, Qinghai Province, China, **3** School of Water Resources and Environment, China University of Geosciences, Beijing, China, **4** School of Geophysics and Information Technology, China University of Geosciences, Beijing, China

* yanghonglei226@126.com



OPEN ACCESS

Citation: Song Y-Z, Yang H-L, Peng J-H, Song Y-R, Sun Q, Li Y (2015) Estimating PM_{2.5} Concentrations in Xi'an City Using a Generalized Additive Model with Multi-Source Monitoring Data. PLoS ONE 10(11): e0142149. doi:10.1371/journal.pone.0142149

Editor: Qinghua Sun, The Ohio State University, UNITED STATES

Received: February 22, 2015

Accepted: October 19, 2015

Published: November 5, 2015

Copyright: © 2015 Song et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the following research projects: the National Natural Science Foundation of China (41304012, 61427802, 41330634, 41374016); the National High-tech R&D Program of China (SS2012AA120804); the Doctoral Fund of Ministry of Education of China (20100022110008); the Open Research Fund of Key Laboratory of Digital Earth Science, Center for Earth Observation and Digital Earth, Chinese Academy of Sciences (2010LDE002); the Fundamental Research Funds for the Central Universities (2652015180). The

Abstract

Particulate matter with an aerodynamic diameter <2.5 μm (PM_{2.5}) represents a severe environmental problem and is of negative impact on human health. Xi'an City, with a population of 6.5 million, is among the highest concentrations of PM_{2.5} in China. In 2013, in total, there were 191 days in Xi'an City on which PM_{2.5} concentrations were greater than 100 μg/m³. Recently, a few studies have explored the potential causes of high PM_{2.5} concentration using remote sensing data such as the MODIS aerosol optical thickness (AOT) product. Linear regression is a commonly used method to find statistical relationships among PM_{2.5} concentrations and other pollutants, including CO, NO₂, SO₂, and O₃, which can be indicative of emission sources. The relationships of these variables, however, are usually complicated and non-linear. Therefore, a generalized additive model (GAM) is used to estimate the statistical relationships between potential variables and PM_{2.5} concentrations. This model contains linear functions of SO₂ and CO, univariate smoothing non-linear functions of NO₂, O₃, AOT and temperature, and bivariate smoothing non-linear functions of location and wind variables. The model can explain 69.50% of PM_{2.5} concentrations, with R² = 0.691, which improves the result of a stepwise linear regression (R² = 0.582) by 18.73%. The two most significant variables, CO concentration and AOT, represent 20.65% and 19.54% of the deviance, respectively, while the three other gas-phase concentrations, SO₂, NO₂, and O₃ account for 10.88% of the total deviance. These results show that in Xi'an City, the traffic and other industrial emissions are the primary source of PM_{2.5}. Temperature, location, and wind variables also non-linearly related with PM_{2.5}.

Introduction

Particulate matter with an aerodynamic diameter less than 2.5 μm, termed PM_{2.5}, is a mixture of various gaseous and particulate components, and the primary source of PM_{2.5} is believed to be anthropogenic emission such as combustion process [1]. The monitoring and study of

fundings had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

PM_{2.5} is important because PM_{2.5} can negatively impact human health [2, 3]. Because of its small size, PM_{2.5} has a long lifetime in the atmosphere and can be transmitted great distance, leading to remote deposit locations in the respiratory tract of human beings. It can readily infiltrate people's lungs, reach alveoli and hardly be eliminated from the body [4]. The primary source of toxicity is the heavy metal elements and toxic organic matters attached on PM_{2.5}, which could affect the normal functions of human body, delay the human development, and even cause heart disease and cancer after entering the body [5–7]. Individuals are more prone to respiratory disease after long-term exposure to high concentrations of PM_{2.5}; additionally, the morbidity and mortality rates increase with exposure [8–10]. To protect public health, PM_{2.5} standards were proposed by US Environmental Protection Agency (USEPA) in 1997, with a short-term (24-h) standard value of 65 µg/m³ and a long-term (annual) standard value of 15 µg/m³ for PM_{2.5} [11]. Following several modifications derived from scientific research on the influence of PM_{2.5} on human health, the current short-term and long-term standards are 35 µg/m³ and 12 µg/m³, respectively [12].

In recent years, days with continuously high concentrations of PM_{2.5} have occurred frequently in central and eastern China. Early death of 1.23 million people and 25 million DALYs (disability-adjusted life years) lost are attributed to high concentrations of outdoor PM_{2.5} (2010) [3]. In January, 2013, approximately 600 million people living in a quarter of China's land area are at risk of PM_{2.5} [13]. Therefore, monitoring, predicting, and controlling PM_{2.5} is of great importance. On February 29, 2012, the Ministry of Environmental Protection of China unveiled a new revised "Ambient air quality standard" (GB3095-2012), which set the primary and secondary standards of PM_{2.5} concentrations for nature protection areas and residential districts respectively. The primary standards set an average of 35 µg/m³ for 24-hour period and 15 µg/m³ for one year, while the secondary standards are 75 µg/m³ as a 24-hour average and 35 µg/m³ as an annual average, respectively. With an annual average PM_{2.5} concentration of 133 µg/m³ in 2013, Xi'an City features one of the highest average concentrations of PM_{2.5} in China. Among 365 days in 2013, in total, there were 352 days and 259 days with concentrations greater than 35 µg/m³ and 75 µg/m³, respectively [14]. Approximately 6.5 million people in nine municipal districts, according to the sixth National Census of China in 2010, are often exposed to high PM_{2.5} concentrations. The health impact of this exposure is a serious concern. Research has shown that deaths caused by PM_{2.5} accounted for 1.6% of all deaths in Xi'an in 2010 [15]. Therefore, it is important to investigate the spatiotemporal and source patterns of PM_{2.5} concentrations in Xi'an City.

A series of models have been proposed to discern the relationships of pollutants and meteorological variables with PM_{2.5} concentrations. PM_{2.5} source apportionment models require the physical and chemical features of the chemical components and the quantitative determination theories [16, 17]. For these models, various chemical elements in PM_{2.5} are obtained with the help of chemical experiments by using the measurement of dozens of chemical elements (e.g., carbon), certain water-soluble ions, and several types of organic matter [18, 19].

Satellite data is a source of air quality variables for estimating PM_{2.5} concentrations. The Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical thickness (AOT) product, an integral of aerosol extinction coefficients in the vertical direction from ground to top of atmosphere, is the relevant satellite remote sensing product that could be used to study variations in PM_{2.5} concentrations. MODIS product can be used to estimate and continuously monitor air quality variation and factors globally, while the ground-based measurements are only representatives at specific positions, primarily in urban regions, with a lack of spatial coverage and sparse monitoring data in rural areas [20]. The experiment conducted in Jefferson County, Alabama in 2002 demonstrated a strong relationship between AOT and PM_{2.5} with a linear correlation coefficient of 0.7 [21]. The strong correlation, especially under clear sky

conditions with less than 40–50% relative humidity, was also validated in six other locations worldwide: Sydney, Delhi, Hong Kong, New York City, and Switzerland, where the linear correlation coefficient between bin-averaged AOT and PM_{2.5} concentration is 0.96 [22], and China [23–26]. Therefore, this product is substantiated by a series of experiments to demonstrate that AOT values have strong positive correlations with PM_{2.5} concentrations [27–31]. However, the relationship between AOT values and PM_{2.5} concentrations is also complex and non-linear, as they are both mixture that the ratios of components are different [32–36], and the relationship depends on multiple factors such as aerosol concentrations, relative humidity, and cloud coverage [22]. Seasonality appeared in the relationship between AOT and PM_{2.5} concentrations in the experiments conducted in the United States, Guangdong Province, and North China [37–39]. In northern China, the strong seasonal variation of the relationship is caused by the dominant aerosol types, implying that large dust particles and soil aerosols are the primary types in spring and summer, and smoke and soot aerosols are the dominant types in autumn and winter [39, 40].

In addition, gas-phase concentrations such as SO₂ and NO₂ could be the auxiliary variables for PM_{2.5} concentrations prediction, since they are the precursor gas contents of secondary ions of water-soluble inorganic salts [41], and can share a common emission source, be processed in the atmosphere and partition to the particulate-phase. In China, the mass concentrations of water-soluble inorganic salts and carbonaceous components make up more than 50% of PM_{2.5} [42–44]. Moreover, a few studies showed that the variations of PM_{2.5} concentrations were sensitive to meteorological variables [45, 46]. The integrated action of these variables, such as temperature, humidity and wind renders PM_{2.5} concentrations highly variable between different periods and regions.

A regular diurnal cycle with two maxima and two minima per day was obvious in the hourly PM_{2.5} concentrations of Xi'an City [47]. The AOT values during particulate matter events including haze, dust storms, straw combustion, and fireworks displays, were twice as large as those of normal days [48]. Based on the emissions of organic and elemental carbon, gasoline engine exhaust, diesel exhaust and coal burning were the main contributors to PM_{2.5} in fall and winter in Xi'an City [49]. In 2010, the concentrations of organic and elemental carbon in PM_{2.5} in winter were 2.62 and 1.75 times higher than those in summer, primarily due to traffic emissions and adverse weather conditions [50]. Thus, the variation of PM_{2.5} in Xi'an city was related to the above emissions and meteorological conditions.

For these studies, linear regression analysis is utilized to determine the types of variables related to PM_{2.5}. However, non-linear relationships between many variables render linear regression model descriptions inaccurate. Moreover, due to the complex and various chemical components of PM_{2.5} and the potential variables in different places and periods, their relationships and variation patterns can be inadequate in Xi'an City. For instance, the dominant species and their ratios are different for AOT in various seasons and areas, indicating that the relationships between AOT and PM_{2.5} concentration to be seasonal and non-linear. Additionally, the relationships (e.g., the effects, degrees and patterns) between meteorological variables and air quality indexes, and PM_{2.5} concentrations are complex. We expect that effects of temperature, wind directions and other variables are non-linear. Therefore, to understand the relationships among the multiple sources and various types of variables, we will have to use non-linear models.

The objective of this work is to explore statistical relationships between the variables and PM_{2.5} concentrations in Xi'an City in 2013, identify the primary variables, and to interpret the patterns of the relationships and understand the source patterns on the basis of the spatio-temporal evolution of PM_{2.5} concentrations. Ground-based monitoring data are provided by several ministries. Satellite remote sensing products will be used as well. Different types of data

complement each other. Additionally, the comprehensive use of multi-source monitoring data is valuable to determine the values and causes of PM_{2.5} concentration levels in different periods and regions. Furthermore, a relationship between PM_{2.5} and its corresponding statistical variables such as pollutants and meteorological variables is constructed by applying a generalized additive model (GAM), which describes the non-linear relationship between variables and responses via nonparametric smoothing functions [51]. A series of experiments have demonstrated that a generalized additive model works better than a linear regression model for variables with non-linear relationships [52–54]. The generalized additive model approach is reliable and flexible for constructing relationships, exploring variables and making predictions [55–58].

Materials and Methods

Study area and sampling sites

The study area includes nine urban districts in Xi'an City, which is the capital city of Shaanxi Province in China. Xi'an City, with an area of 3,581 square kilometers and a population of 6.5 million people, is located in the middle of the Yellow River's Guanzhong Plain, with the Qinling Mountains to the south and the Weihe River to the north. The terrain features high elevations in the southeast and low elevations in the northwest. Days with continuous high concentration of PM_{2.5} occur frequently in central and eastern China in recent years. The annual average of PM_{2.5} concentration is 133 µg/m³ in Xi'an City, one of the cities with the most severe air pollution in China.

The research in this paper is based on daily PM_{2.5} concentrations, reported by the Xi'an environmental monitoring station [14] and obtained from 13 observation stations in Xi'an between January 1 and December 31, 2013. Fig 1 shows the location of Xi'an City in China and the distribution of the 13 observation stations on a digital elevation model (DEM) map. Stations 7 and 8 are located in mountainous regions with the elevations 500.2 m and 500.8 m respectively, and other stations are in plain area with the elevations ranging from 388.1 m (station 5) to 438.4 m (station 2). PM_{2.5} concentrations and four gas-phase concentrations (daily average concentrations of SO₂, NO₂, CO and O₃) are measured daily at the 13 observation stations.

In Table 1, the PM_{2.5} concentrations and air quality are classified according to the definitions in Technical Regulation on Ambient Air Quality Index (under review) (HJ 633–2012, China's environment protection standard). The air quality conditions can be presented with different colors. The time series of Fig 2 show the variation and seasonality of PM_{2.5} concentrations in Xi'an in 2013. We can see from the time series that excellent and good air qualities are primarily from May to August, and seriously polluted and severely polluted air qualities are from January to March and from October to December. Especially, severely polluted air qualities are in the whole winter (December, January, and February). Table 2 lists the summary statistics of the PM_{2.5} monitoring data (raw dataset is S1 Appendix). Missing PM_{2.5} concentrations data accounts for 10.7% (509/4745 data points). In other words, we use 89.3% of the raw data (4236 data points) to conduct the following experiment.

Measurement of gas-phase concentrations and meteorological data

Daily average of gas-phase pollutant concentrations in the study area are also collected from the 13 monitoring stations of the Xi'an environmental monitoring network [14] in Xi'an City. These indices include daily average concentrations of SO₂, NO₂, CO and O₃ measured at each station. Table 2 lists the summary statistics of these indices with a few missing observations. Similar to the time series of PM_{2.5} concentrations shown in Fig 2, the time series of four

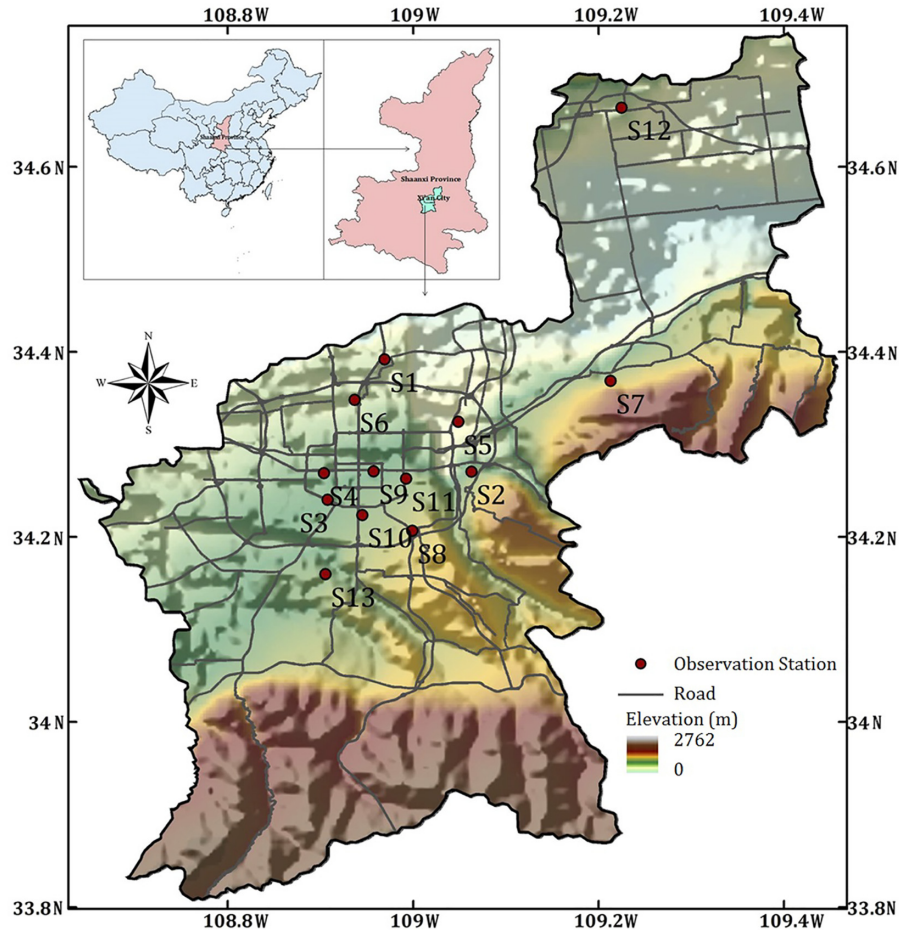


Fig 1. The location of Xi'an City in China and the distribution of 13 observation stations.

doi:10.1371/journal.pone.0142149.g001

gas-phase concentrations and temperature in Fig 3 also show seasonality. The season with relatively high PM_{2.5} concentrations is winter. The pattern is identical with that of SO₂ concentrations. While the largest values of NO₂ and CO concentrations appear in both winter and spring, and the largest O₃ concentrations occur in summer. Besides, the spatial differences of each variable change with the ranges of daily observations in 13 stations, as shown in the vertical black lines.

The meteorological data in study area consists of daily temperature (daily maximum temperature, minimum temperature and mean temperature), wind direction and wind scale in

Table 1. Categories of air quality and corresponding PM_{2.5} concentrations.

Level of air quality index	Category of air quality	PM _{2.5} concentration (µg/m ³)	Color
1	Excellent	0–35	Green
2	Good	36–75	Yellow
3	Mildly polluted	76–115	Orange
4	Moderately polluted	116–150	Red
5	Seriously polluted	151–250	Purple
6	Severely polluted	>250	Maroon

doi:10.1371/journal.pone.0142149.t001

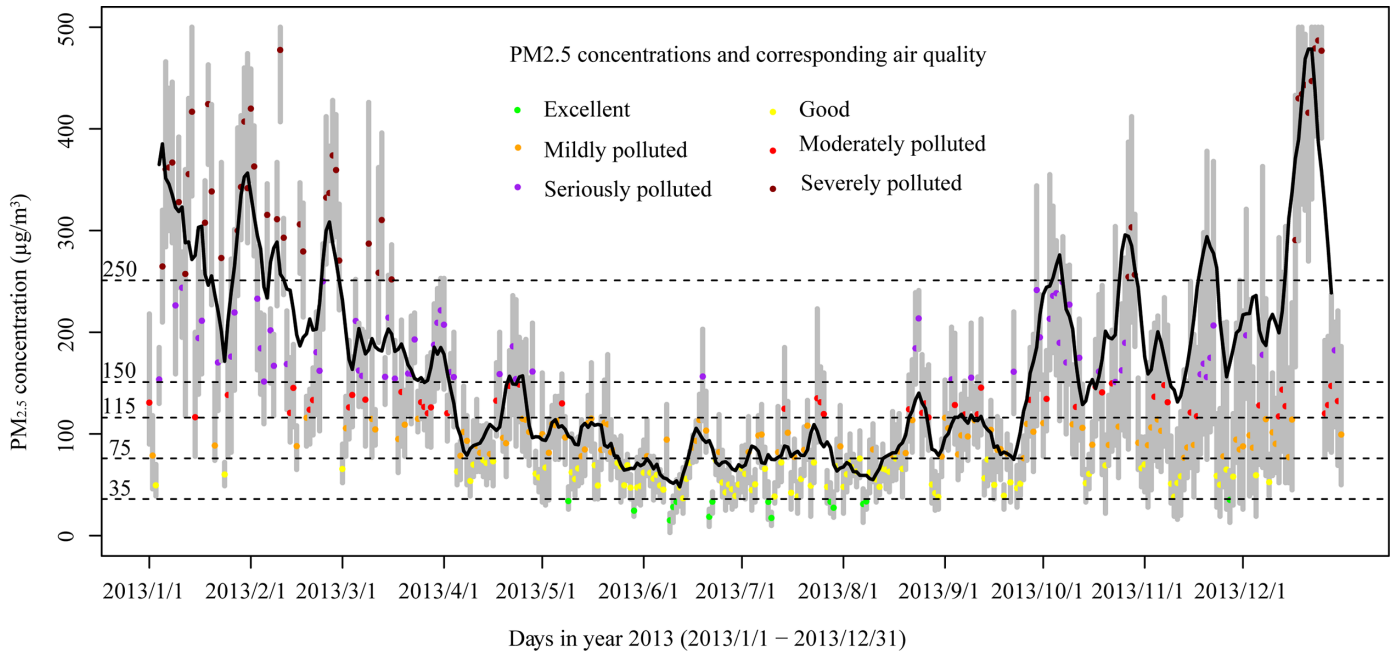


Fig 2. Time series of daily PM_{2.5} concentrations at 13 monitoring stations of Xi'an in 2013. The gray vertical lines indicate the range of daily PM_{2.5} concentrations at the 13 monitoring stations (from maximum to minimum), the colored dots are the average concentrations of each day (according to the defined colors in Table 1), and the black line illustrates the continuous 7-day average concentration.

doi:10.1371/journal.pone.0142149.g002

Xi'an City throughout 2013. The meteorological data are obtained from the China Meteorological Administration [59]. Table 2 lists the summary statistics of the daily mean temperature and wind scale, and Fig 4 shows the summaries of wind directions in each season (raw dataset is S1 Appendix). Because the actual observed meteorological data are invariant across Xi'an City for a given day [60] and the scale of the study area is not too large to discuss the variations in meteorological data, the application of a single weather observation to the whole region is reasonable.

A circular area with a radius of 14 km covers 11 observation stations (with stations S7 and S12 excluded). The distances between the circle's center and the two excluded stations are about 25 km and 50 km, respectively. The correlation between the 47-year daily temperature series of Beijing, a city with the same monsoon climate and similar terrain as Xi'an, is significant ($\alpha = 0.01$). For all the reference stations located within 340 km the correlation coefficients are larger than 0.94 [61]. Furthermore, the correlation coefficient between the daily

Table 2. Summary of data (except for wind direction) used to model the PM_{2.5} concentrations.

Variables	Units	Observation stations	No. of observations (missing percent, %)	Min	1st Qu.	Median	Mean	3rd Qu.	Max
PM _{2.5}	µg/m ³	13	4236 (10.73)	3	65	102	135.6	172	500
SO ₂	µg/m ³	13	4137 (12.81)	0	18	32	38.83	55	135
NO ₂	µg/m ³	13	3927 (17.24)	0	42	62	66.34	89	199
CO	mg/m ³	13	4181 (11.89)	0	30	43	51.36	68	135
O ₃	µg/m ³	13	4016 (15.36)	1	18	33	40.48	52	199
Temperature	°C	Xi'an City	365 (/)	-3.50	7.00	17.50	16.61	25.50	32.50
Wind scale	level	Xi'an City	365 (/)	3	3	3	3.038	3	5
AOT	/	13	4745 (/)	0.05	0.34	0.52	0.62	0.78	3.80

doi:10.1371/journal.pone.0142149.t002

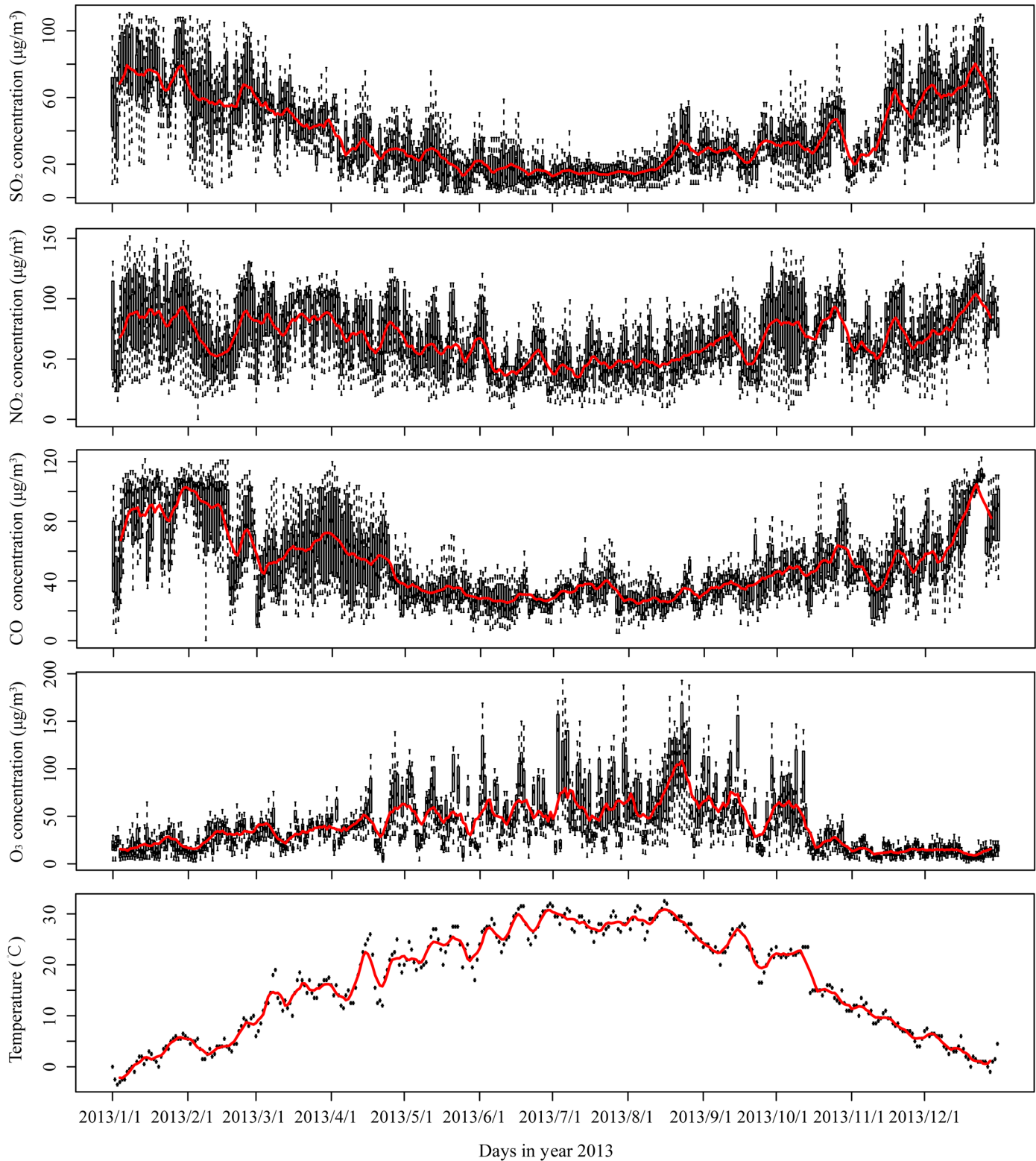


Fig 3. The time series of gas-phase concentrations (SO₂, NO₂, CO, and O₃, µg/m³) and temperature (°C). The black vertical lines are the daily ranges of concentrations at 13 observation stations, and red line illustrates the continuous 7-day average concentration.

doi:10.1371/journal.pone.0142149.g003

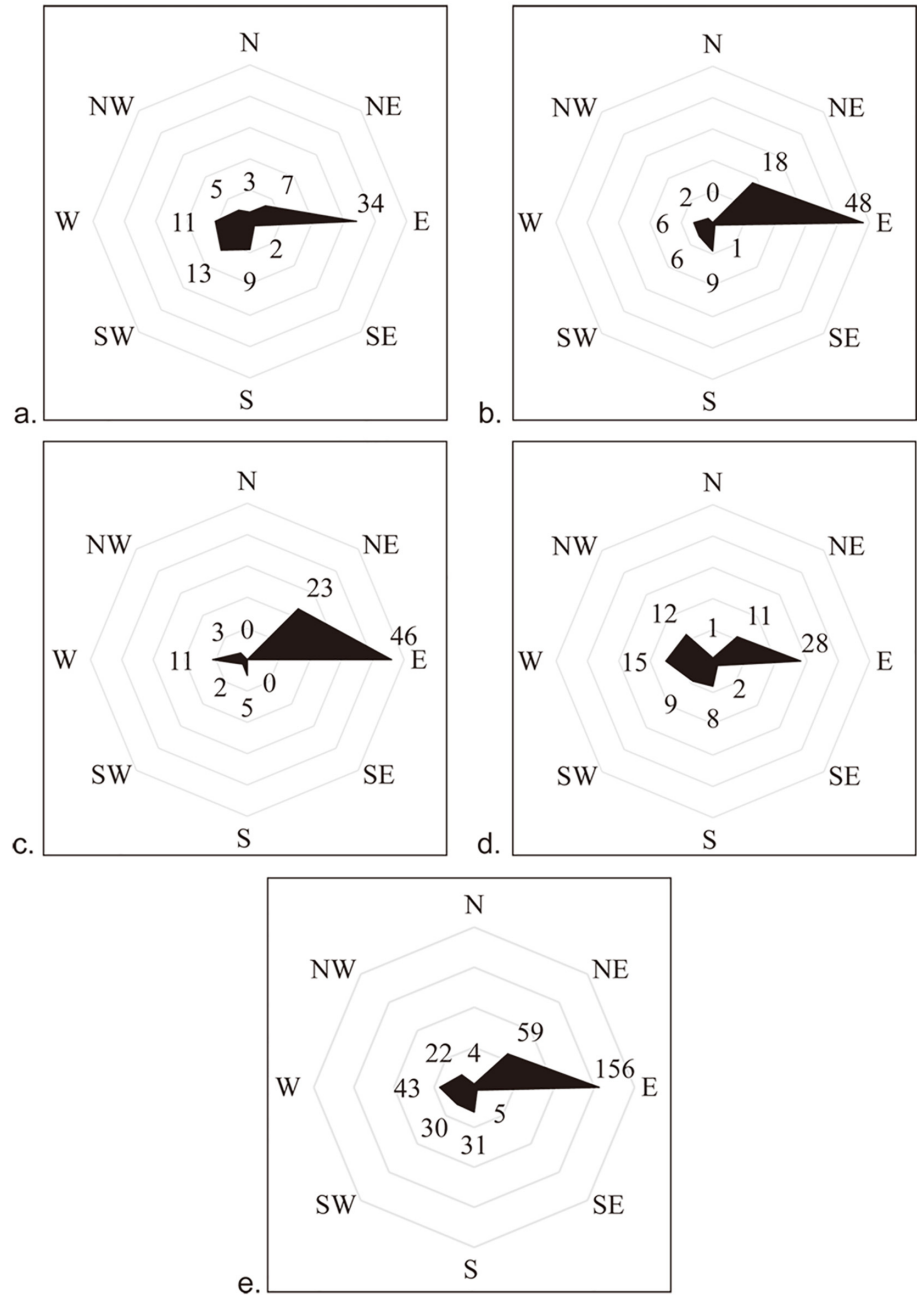


Fig 4. Number of days with the same wind directions in each season (a. spring; b. summer; c. autumn; d. winter) and the whole year (e.).

doi:10.1371/journal.pone.0142149.g004

temperatures of the two locations with a distance of 15.74 km is larger than 0.96, and the correlation between the stations with a distance of 55.51 km is larger than 0.94 [62]. Daily temperature variations in these 13 observation stations are small even constant in all seasons [63], so the use of single daily mean temperature for the whole study area is reasonable.

According to the Chinese National Standards (GB/T 19201–2006), wind speed is divided into a number of levels: level 3 (a breeze) defined at 3.4 m/s–5.4 m/s, level 4 (a soft breeze) at 5.5 m/s–7.9 m/s, and level 5 (a fresh breeze) at 8.0 m/s–10.7 m/s. Wind scale is the

substitution variable for wind speed if wind speed is not available. Therefore, the wind-related variables are wind direction and wind scale. The wind scale for 96.4% of the days in 2013 (352/365) is at level 3 and the occurrence of levels 4 or 5 is rare. In 2013, easterly wind is the primary direction (42.7%, 156 days), as shown in Fig 4, and there are only 15 days with almost no wind. Wind direction and wind speed in Xi'an are primarily affected by the Pacific Ocean and Siberian air current. North-westerly wind is formed by the Siberian high-pressure system in autumn and winter, and south-westerly or south-easterly winds are related to the increasing or decreasing Pacific Ocean pressure in spring and summer [64–66]. Because the study area is much smaller than the area affected by air pressure, wind is usually constant among the observation sites in this study area. Ratios of days with easterly wind and north-easterly wind increase from 11.2% in spring to 18.1% in summer, but decrease from 18.9% in autumn to 10.7% in winter. Ratios of days with south-westerly wind, southerly wind, and westerly wind decrease from 9.0% in spring to 4.9% in autumn, but increase to 28.5% in winter.

Aerosol optical thickness data

AOT, which is the integral of aerosol extinction coefficients in the vertical direction from the ground to the top of atmosphere, is an alternative satellite product to effectively predict PM_{2.5} concentrations. The strong correlation between AOT and PM_{2.5} concentrations has been documented by a series of studies in recent years [21, 23, 28]. The daily AOT (550 nm) data is a MODIS Terra Atmosphere level 3 product, downloaded from the Global Space Flight Center MODIS Level 1 and Atmosphere Archive and Distribution System Web [67]. The AOT values for the 13 observation stations are generated using the Ordinary Kriging method and the MODIS Terra Atmosphere level 3 product with a spatial resolution of 1°. Fig 5 shows the spatial distribution of the annual mean AOT (550 nm) obtained from the MODIS satellite data product for 2013 in Xi'an and its surrounding cities. This figure also shows that Xi'an City is one of the areas with relatively high AOT (550 nm) values.

Modeling PM_{2.5} concentrations

The flowchart of the experiment process is illustrated in Fig 6. The three major steps of the experiment are the exploratory analysis, variable selection and estimation model. The first step exploratory analysis aims at charactering PM_{2.5} concentration data, testing its statistical features, and removing the outliers. The target of the second step is to select the variables that could be statistically related to PM_{2.5} concentration for statistical models. Finally, the last step is to find potential non-linear relationships, to estimate model parameters and to quantify the variances of the potential variables.

Exploratory analysis and variable selection

Exploratory analysis, based on the summary statistics of PM_{2.5} concentrations and the independent variables listed in Table 1, is a key step for processing raw data. Both dependent and independent variables are tested with Quantile-Quantile (QQ) plots and those that don't follow a normal distribution are transformed to a normal distribution with natural logarithm. In this process, the variables PM_{2.5} concentration, O₃ concentration, and AOT (550 nm) are transformed to the data with a normal distribution. Outliers, which generate anomalous analysis results, are removed by using the criteria. More specifically, a value with more than three times of the estimated standard deviation ($> 3\sigma$) from the median is treated as an outlier from an assumed normal distribution [68]. As a result, a small number of outliers in PM_{2.5} concentrations data and certain other variables are removed.

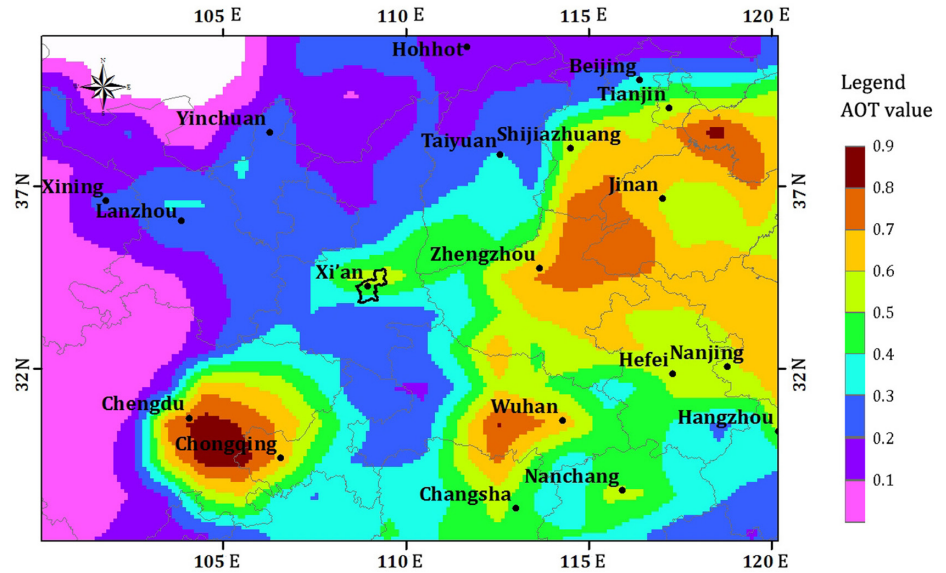


Fig 5. Spatial distribution of annual mean AOT in 2013 for Xi'an and its surrounding cities.

doi:10.1371/journal.pone.0142149.g005

The normalized variables without outliers are divided into the following groups: (a) the satellite-based data is AOT (550 nm) values; (b) the gas-phase concentrations are the daily average concentrations of SO₂, NO₂, CO and O₃; (c) the meteorological variables of daily mean temperature, wind direction and wind scale; and (d) the spatial location variables of longitude and latitude. Temporal variables, such as days or seasons, are not considered, because the meteorological variables are physical reflections of temporal variables, and they are equivalent to the physical meanings of temporal variables in this particular case.

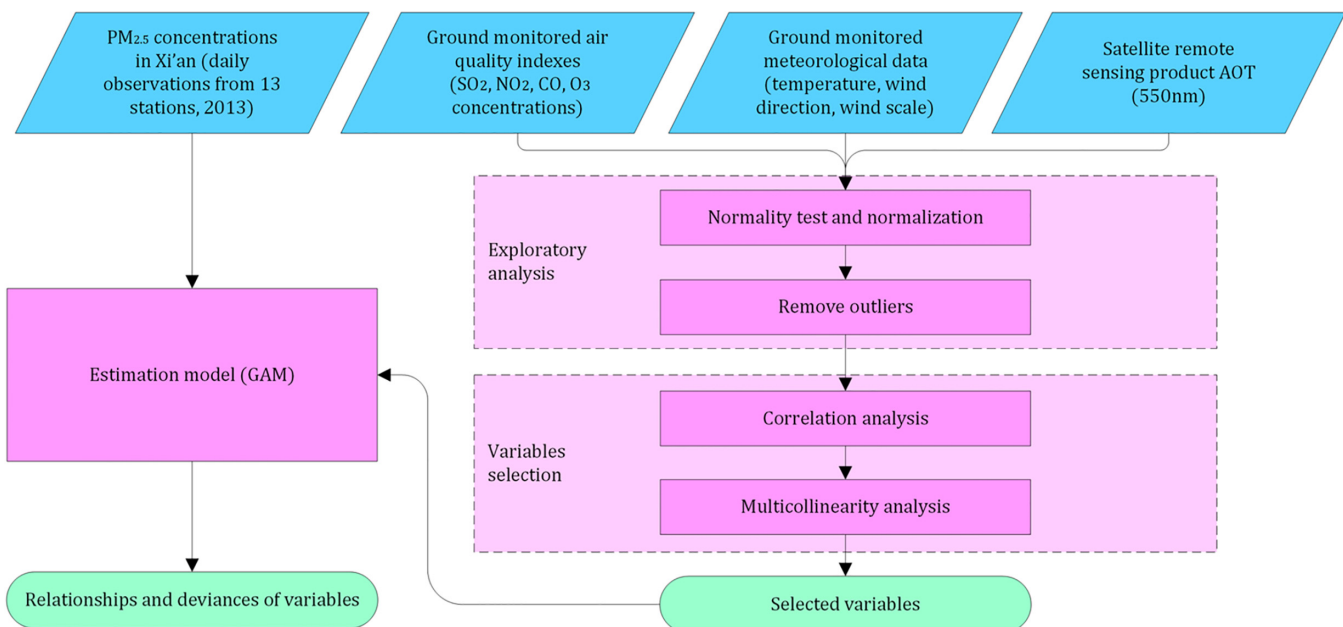


Fig 6. Schematic overview of the estimation experiment.

doi:10.1371/journal.pone.0142149.g006

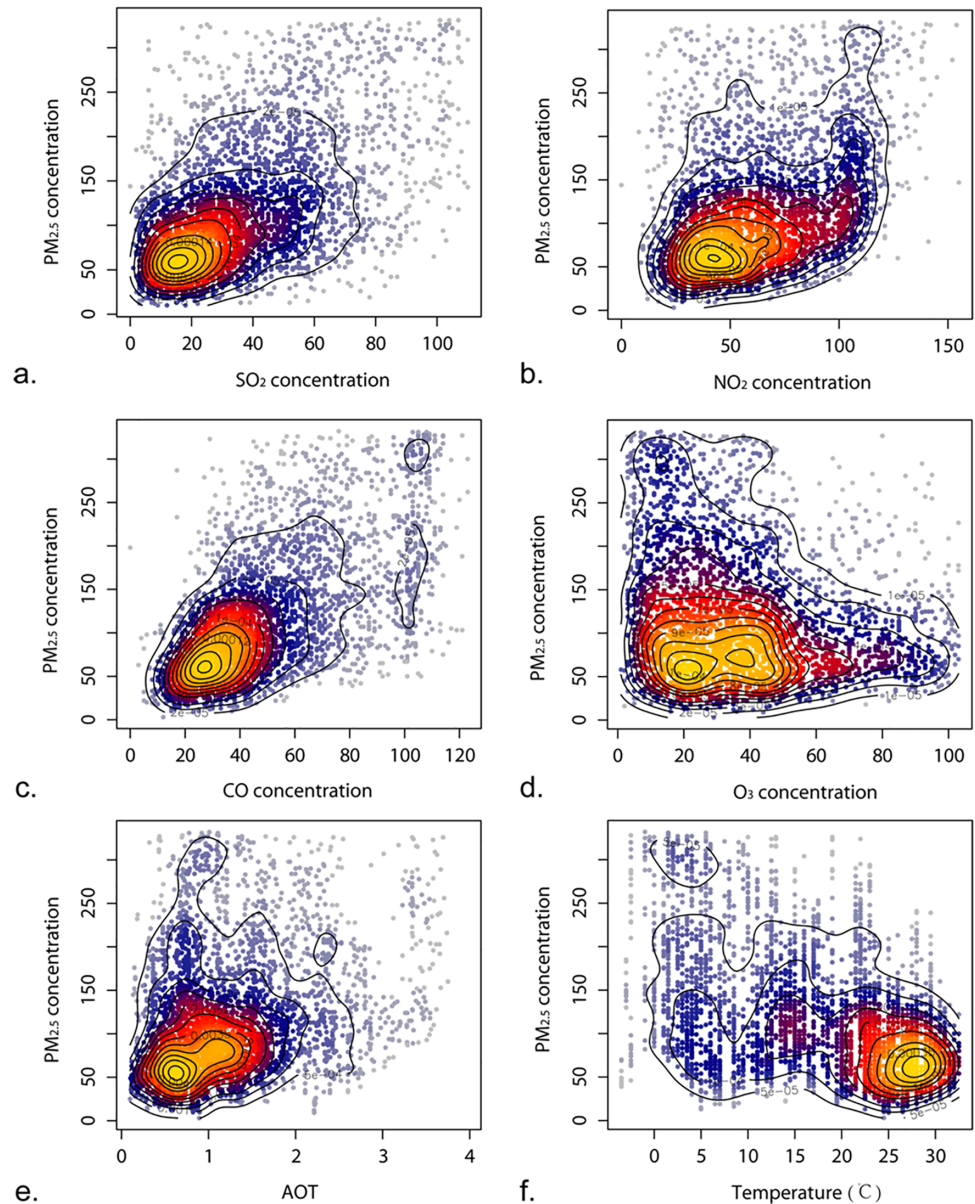


Fig 7. Scatter plots between PM_{2.5} concentration ($\mu\text{g}/\text{m}^3$) and the variables gas-phase concentrations (SO₂, NO₂, CO, and O₃, $\mu\text{g}/\text{m}^3$), AOT values, and Temperature ($^{\circ}\text{C}$).

doi:10.1371/journal.pone.0142149.g007

Correlation analysis is conducted for these groups of variables to remove those without significant correlation with PM_{2.5} concentrations. Scatter plots between PM_{2.5} concentrations and the variables in group (a), (b), and temperature are used to depict and determine the linear or non-linear relationships. Fig 7 shows the relations between PM_{2.5} concentrations and the variable four gas-phase concentrations, AOT values, and daily mean temperature. The colors and contours represent the density of the scatters. It is clear from the figure that the relationships between PM_{2.5} concentration and concentrations of SO₂ and CO are linear, but other relationships are complex and regarded as non-linear relationships. Further, in Fig 8, the seasonality between AOT (550 nm) and PM_{2.5} concentration is depicted with scatter plots and linear regression functions in the four seasons in Xi'an City. The regressions illustrate that the

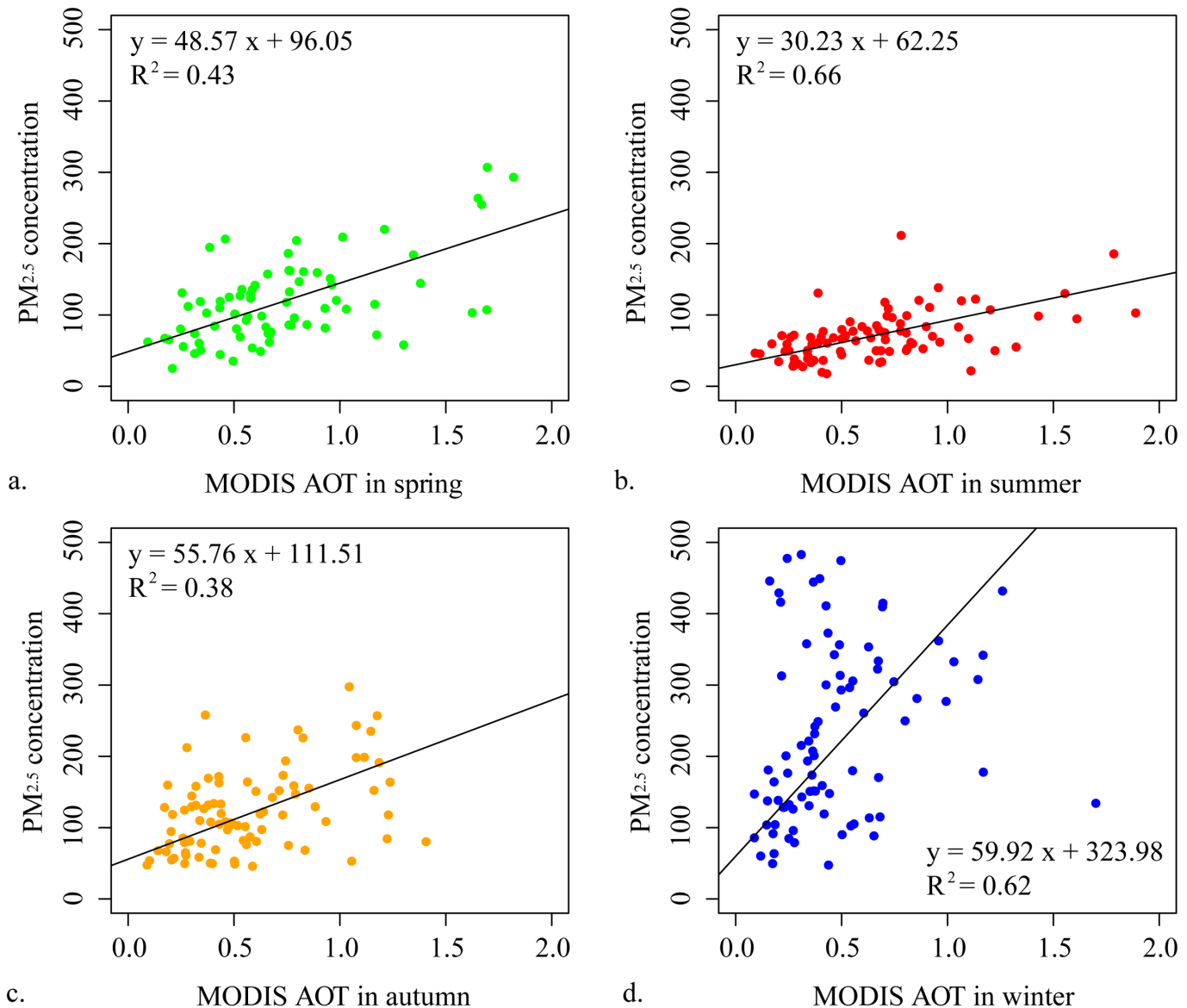


Fig 8. Linear regression functions between MODIS AOTs and PM_{2.5} concentrations (µg/m³) in different seasons.

doi:10.1371/journal.pone.0142149.g008

relationships between AOT (550 nm) and PM_{2.5} concentrations are distinct among the four seasons.

Pearson correlation coefficients between PM_{2.5} concentrations and the four air quality indices observed at the 13 stations each day were calculated. At this step, the non-linear relationships between the spatial locations of the observation stations and wind direction with PM_{2.5} concentrations must be considered. Additionally, the effect of AOT values on PM_{2.5} concentrations is seasonal and non-linear. Then, multicollinearity analysis is performed using variance inflation factors (VIFs) for the variables. The variables are divided into two groups, weakly correlated variables (VIF < 10) and highly correlated variables (VIF ≥ 10). One highly correlated variable is selected for combination with the weakly correlated variables, and VIFs are

recalculated. This procedure is repeated, and the optimal combination of experimental variables is eventually chosen. In this experiment, the maximum values of VIFs are less than 4, so all VIFs are less than 10.

Estimation model

A generalized additive model is used to construct the unknown statistical relationships, linear or non-linear, between PM_{2.5} concentration and the potential variables. In this work, the PM_{2.5} data form a continuous time series, but thirteen observation stations are not enough to perform spatial interpolation using geostatistics methods. Additionally, linear regression modeling is inappropriate when the dependent and independent variables are non-linearly related. The relationships between dependent variable, the pre-processed PM_{2.5} concentration, and independent variables are constructed via a generalized additive model under the assumption of normal distributions as follows:

$$y = \beta_0 + \sum_{i=1} \beta_i x_i + \sum_{j=1} f_j(x_j) + \sum_{k=1} f_k(x_{1,k}, x_{2,k}) + \varepsilon \quad (1)$$

where y is the dependent variable; x_i is a continuous linear variable; x_j is a continuous non-linear variable; $x_{1,k}, x_{2,k}$ are continuous non-linear variable-pairs; β_0 and β_i are unknown coefficients; f_j and f_k are nonparametric smoothing functions between y and non-linear variables; and ε is the normal random error term ($\varepsilon \sim N(0, \sigma^2)$).

In the model, wind variables are calculated as w_c ([wind scale] × cosine([wind direction])) and w_s ([wind scale] × sine([wind direction])). Additionally, the coordinates of the thirteen observation stations are used as variables. Because a few raw data points of certain variables were lost, the data used in the subsequent calculation are less than 4,745 lines (365 days × 13 observation stations). Therefore, data in 349 days (3067 lines) is available which is less 365 days of data because of missing data, and it becomes 333 days of data (2880 lines) when outliers are removed.

From the spatial perspective of geostatistics, generalized additive models are more flexible and reliable than spatial interpolation methods. To apply this method, a tensor product of longitude and latitude is constructed to describe the spatial distribution of PM_{2.5} concentrations, because the number of stations to monitor PM_{2.5} concentrations is only 13 and is not sufficient to estimate the PM_{2.5} concentrations of unknown coordinates via kriging interpolation. The reason is that too few observations will cause bias estimation [69]. However, estimates can be produced by taking longitude and latitude as variables and applying a bivariate smoothing function in a generalized additive model.

In this research, the *mgcv* package in the program R is used. The parameters of the smoothing functions are selected automatically by the generalized cross-validation (GCV) criterion due to its iterative approach, which improves calculation efficiency, which can help determine whether GCV score decrease when one variable is removed [70]. Three aspects are considered for model assessment: residual analysis for the model, comparison between the fitted values and observations, and fitness evaluation with the result comparison of generalized additive model to stepwise linear regression.

Results and Discussion

In this research, satellite-based data AOT (550 nm), together with four gas-phase concentrations (SO₂, NO₂, CO, and O₃), meteorological data (temperature, wind scale, and wind direction), and locations of observation stations (longitude and latitude), are treated as independent variables for constructing the relationships with dependent variable, namely, PM_{2.5}

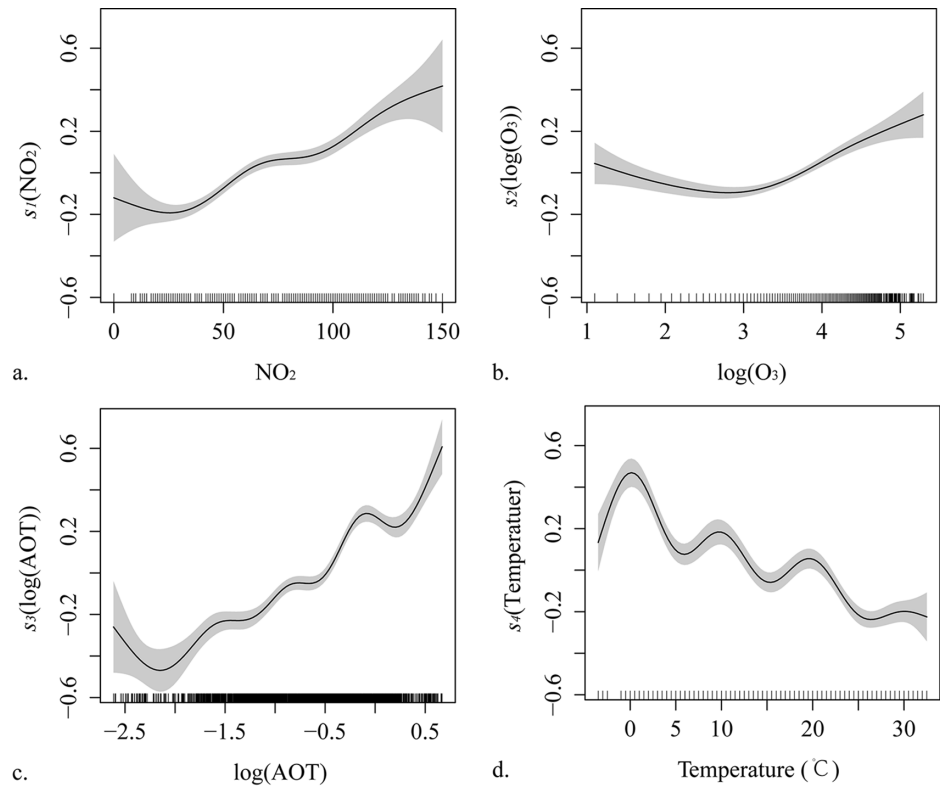


Fig 9. The effects of univariate smoothing functions for log(PM_{2.5}) in the generalized additive model: (a) $s_1(\text{NO}_2, \mu\text{g}/\text{m}^3)$, (b) $s_2(\log(\text{O}_3), \mu\text{g}/\text{m}^3)$, (c) $s_3(\log(\text{AOT}))$, and (d) $s_4(\text{Temperature}, ^\circ\text{C})$. The gray shaded areas are the estimated 95% confidence intervals. The vertical lines adjacent to the x-axis indicate the presence of data.

doi:10.1371/journal.pone.0142149.g009

concentration, in Xi'an City. These variables reflect both air quality variations and meteorological conditions [71–73]; thus, the necessity of using both kinds of variables is affirmed. With the pre-process and consideration of each variable, the constructed generalized additive model is given as follows:

$$\log(\text{PM}_{2.5}) = 3.97 + 4.17 \times 10^{-3} \times \text{SO}_2 + 1.07 \times 10^{-2} \times \text{CO} + s_1(\text{NO}_2) + s_2(\log(\text{O}_3)) + s_3(\log(\text{AOT})) + s_4(\text{Temperature}) + t_1(w_c, w_s) + t_2(Lo, La) + \varepsilon \quad (2)$$

where s_i ($i = 1, 2, 3, 4$) is a univariate smoothing function, the spline based smooth function; t_j ($j = 1, 2$) is a bivariate smoothing function, the tensor product smooth function; w_c ([wind scale]×cosine([wind direction])) and w_s ([wind scale]×sine([wind direction])) denote the products of wind direction and wind scale, i.e., the substitution variable of wind speed; Lo and La are longitude and latitude, respectively.

The formula shows the concentrations of SO₂ and CO are positively and linearly related with PM_{2.5} concentration, while other statistical relationships are non-linear. Different types of non-linear relationships can be expressed with these smoothing functions in the generalized additive model. Fig 9 shows the effects of the univariate smoothing functions on log(PM_{2.5} concentrations) in the model. In general, the smoothing function of NO₂ concentration is an increasing non-linear function, but other three functions contain both increasing and reduction parts. When log(O₃) is smaller than 3, the values are sparse and part of function is decreasing. On the contrary, when it is larger than 3, the values of log(O₃) are dense and part of

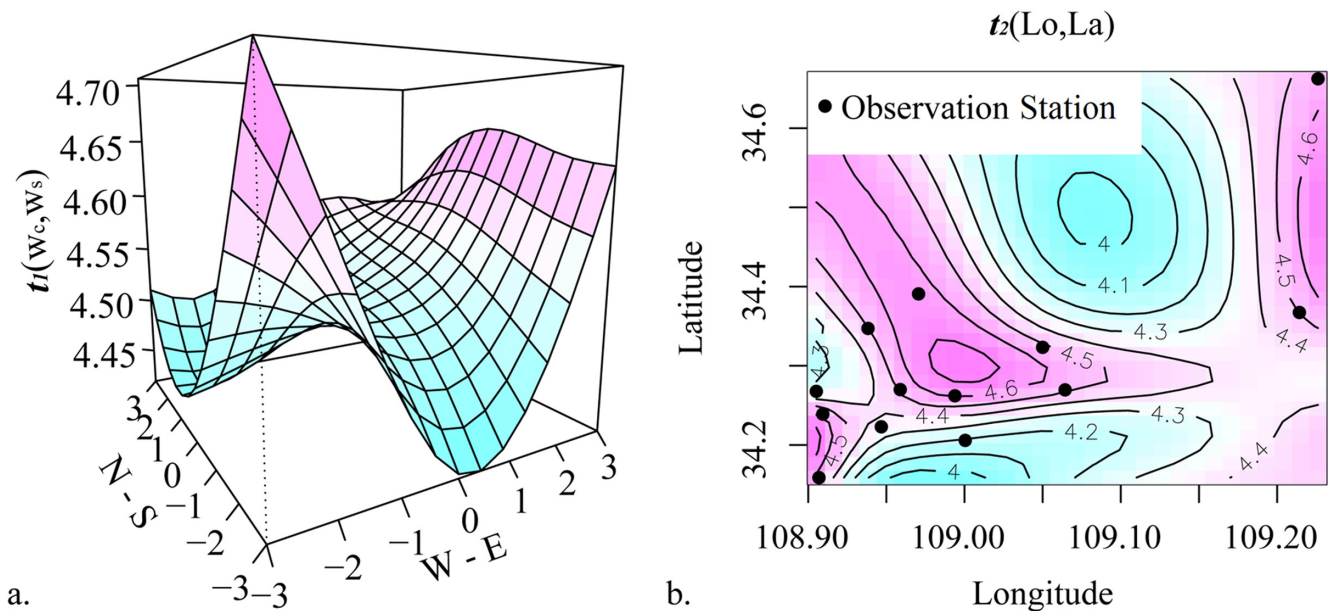


Fig 10. Bivariate smoothing functions of fitted components for log(PM_{2.5}) in the generalized additive model. (a) Bivariate smoothing function for the wind components and the corresponding fitted values. (b) Bivariate smoothing function for longitude and latitude and the corresponding fitted values. The black dots represent the 13 observation stations that monitor PM_{2.5} concentrations.

doi:10.1371/journal.pone.0142149.g010

function is increasing. The function of log(AOT) rises with fluctuations though it declines at the small value part. Based on the smoothing function of daily mean temperature with fluctuation changes, the fitted PM_{2.5} values are highest when the daily mean temperature is 0°C. This result demonstrates that 0°C represents the suitable weather conditions for PM_{2.5} accumulation. Bivariate smoothing functions of fitted components for log(PM_{2.5} concentrations) in the generalized additive model are shown in Fig 10. In Fig 10A, vertical axis shows the bivariate smoothing function, $t_1(w_c, w_s)$, for the wind components and its corresponding fitted values.

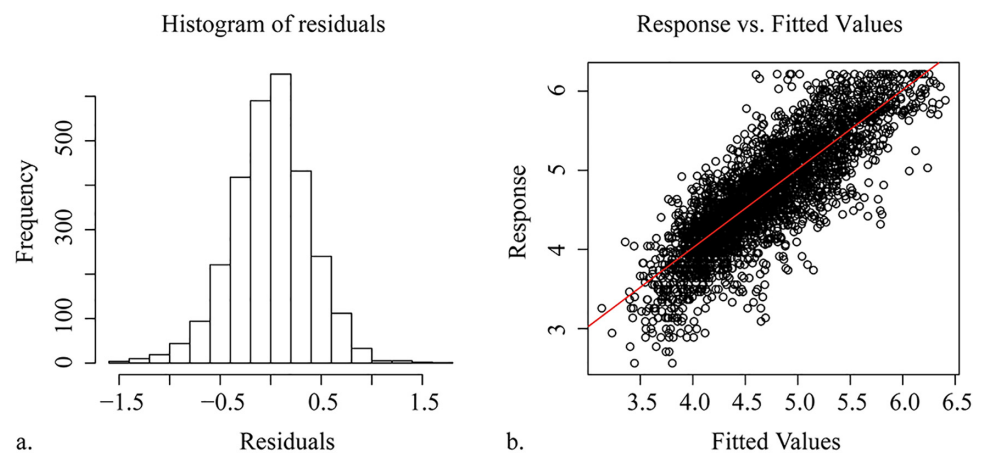


Fig 11. Evaluations of the model: (a) histogram of residuals and (b) fitted PM_{2.5} concentrations vs. observations.

doi:10.1371/journal.pone.0142149.g011

Table 3. Deviance explained by variables related to PM_{2.5} concentrations in Xi'an.

Variables	Deviance contribution
SO ₂ concentration	2.23%
NO ₂ concentration	5.86%
CO concentration	20.65%
O ₃ concentration	2.79%
Location (longitude, latitude)	6.70%
Wind direction (W-E, N-S) and scale	3.91%
Daily mean temperature	7.82%
Aerosol optical thickness	19.54%
Total deviance explained	69.50%

doi:10.1371/journal.pone.0142149.t003

Southwest and east are the major wind directions affecting the variation of PM_{2.5} concentrations. In Fig 10B, the contour lines show the bivariate smoothing function ($t_2(Lo,La)$) of location variable, i.e. longitude and latitude. In this figure, the black dots are the 13 observation stations monitoring PM_{2.5} concentrations. The location function suggests that PM_{2.5} concentrations in urban areas of Xi'an are higher than in suburbs.

The R² value of the generalized additive model is 0.691, and the fitted residuals are normally distributed and not skewed (Fig 11). The model explains 69.50% of the total deviance in the PM_{2.5} concentrations data. To evaluate the model, a comparison between the estimated results and the results of a stepwise linear regression is considered, and evaluations of the fitted residuals and fitted values in this model are calculated. The fitting R² value is 0.582 for the stepwise linear regression with the same variables and processed data after the exploratory analysis as those used in the generalized additive model. Therefore, the model in this paper is better than the stepwise linear model with the fitness improvement of 18.73%.

Since understanding the contributions of each variable to the total deviance is important, we calculate the contributions of the eight groups of variables and list the results in Table 3. Among these groups of variables, the linear variable CO concentration and log-transformed AOT (550 nm) values explain 20.65% and 19.54% of the deviance in PM_{2.5} concentrations, respectively. The total deviance that can be explained by gas-phase concentrations of SO₂, NO₂, and O₃ is 10.88%. In this case, even half of this value is accounted for NO₂ concentration. The other variables, with the decreasing contributions to the deviance, are the temperature, explaining 7.82% of deviance, the location, explaining 6.70% of deviance, and finally the wind, explaining 3.91% of deviance.

These results suggest that the PM_{2.5} concentrations in Xi'an City are most statistically related to carbon monoxide and aerosols. Carbon aerosols, a vital chemical component of PM_{2.5}, and carbon monoxide share the same origin [74]. The common sources of carbon aerosols and carbon monoxide are likely due to traffic and other industrial emissions. This relationship supports the finding that CO concentrations and aerosols contribute most to the statistical relationships of PM_{2.5} concentrations in Xi'an. The two largest anthropogenic sources of carbon aerosols at the global scale are biomass combustion and fossil fuel combustion, while the major natural sources are plant emissions and natural fires [75]. The population density in study area is more than 1,800 people per square kilometer. In such a large city, human activities, such as fuel combustion, emission-generating vehicle operation and biomass burning, play an important role in forming large quantities of carbon aerosols. In Xi'an City, gasoline engine exhaust, diesel exhaust and coal burning accounted for the PM_{2.5} mass of 48.8%±10% in fall and 45.9%±7.5% in winter, respectively [49].

Apart from CO, the other three gas-phase concentrations, SO₂, NO₂, and O₃, also contribute to PM_{2.5} concentration, though they appear to be less important than the above two variables, CO and aerosols. Sulfur dioxide and nitrogen dioxide are mainly released by traffic and industrial emissions in urban areas [76, 77]. Ozone is formed when oxygen molecules are combined with atoms released from nitrogen dioxide molecules in hot days [78–80]. It is formed primarily in urban areas from the end spring to early autumn, as shown in Fig 3. Therefore, the relationships between PM_{2.5} concentration and gas-phase concentrations could indicate that one of the important sources of PM_{2.5} is the traffic and industrial emissions.

The impact of daily mean temperature, a concrete expression of time, ranks after aerosols and gas-phase concentrations. Seasonality, with the lowest temperatures in winter, is the feature of temperature variation in Xi'an City. The period of highest PM_{2.5} concentrations occurs in winter because the weather conditions at 0°C are stable and favorable for PM_{2.5} accumulation.

In addition, 6.70% of the deviance is explained by the coordinates of the observation stations. The result of location non-linear function in Fig 10 shows that the PM_{2.5} concentrations in the urban areas are higher than those of suburbs in Xi'an spatially. This result confirms that AOT product is more significant for the global estimation of PM_{2.5} concentration, especially in the rural areas, because PM_{2.5} concentrations are different for urban and rural areas and because most of the ground-based measurements only provide the monitoring concentrations in urban areas.

Wind direction and wind scale have limited but still important influences on the concentration of PM_{2.5}. Southwesterly wind and easterly wind contribute to higher concentrations relative to the other wind directions. According to this dataset and previous records of the climate in Xi'an, southwesterly wind and easterly wind primarily occur in the spring and summer respectively as a result of increasing and decreasing Pacific Ocean air pressure systems. Given the topographic conditions of Xi'an City and the distribution of its surrounding AOT (550 nm) values, easterly wind enables PM_{2.5} in the eastern areas with relatively high AOT (550 nm) values to spread to Xi'an City because the eastern area is connected with other AOT (550 nm) concentrated areas through the channel shown in Fig 5. Thus, wind has a weak influence on the variation in PM_{2.5} concentrations. The reasons are that southwesterly and easterly winds play the most important role in increasing concentrations. Nevertheless, the concentrations are relatively low in the seasons when they frequently occur. However, during winter and autumn, northwesterly winds have a more limited effect than southwesterly and easterly winds. The resuspension of dust is enhanced in spring and summer due to higher wind speeds and is reduced in winter when the soil is damp, which limits the release of soil-derived particles [81–83]. Thus, the relatively high concentrations are mainly produced locally in winter and autumn instead of coming from other regions through the channel mentioned above. From the above discussion, we may conclude that the high concentrations in winter primarily come from local emissions in Xi'an, such as home heating and the imperfect combustion of fuel. Whereas in spring and summer, the higher concentrations are possibly related to the concentrations of neighboring areas.

Still, there are certain limitations in the experiment. To ensure the strictness and accuracy of the model and its calculating process, a normal distribution is assumed in the model, which leads to normalizing and outlier removal for large amounts of data. In addition, wind scale is used in the model as an approximation of wind speed because of the lack of wind speed data. Using a single meteorological monitoring site for the whole study area is limited, for wind direction might vary at different local places, though the trend in large spatial range is consistent. Because the focus of this article is on modeling the non-linear regression, a variety of variables are not introduced into the model as independent variables. Certain natural attributes,

such as highly spatially resolved elevation, land use and even social attributes (e.g., population density) [84], could be treated as variables related to the variation in PM_{2.5} concentrations on account of previous authoritative studies and strict experiments.

Conclusion

A generalized additive model was constructed and used to estimate PM_{2.5} concentrations in Xi'an City. Various variables from multiple monitoring data sources have been used, including satellite remote sensing product AOT, ground-based gas-phase concentrations, and meteorological monitoring data. The observed data are significant for constructing the statistical model. The model can explain 69.50% ($R^2 = 0.691$) of the total deviance in the PM_{2.5} concentrations. The most significant variables, CO and AOT (550 nm), account for 20.65% and 19.54% of the deviance, respectively, indicating that the variation in PM_{2.5} concentrations is strongly correlated with CO and AOT. The total deviance that can be explained by three other gas-phase concentrations, SO₂, NO₂, and O₃ is 10.88%. These results demonstrate biomass and fossil fuel combustion, likely produced by traffic and other industrial emissions, are the primary source of PM_{2.5} in Xi'an City. Additionally, the results have shown that temperature influences the concentrations, especially in winter when source emissions increase due to home heating and favorable temperature conditions of around 0°C for PM_{2.5} accumulation. PM_{2.5} is more likely to be accumulated in urban areas than in suburbs, which also affirms that satellite-based data is essential for predicting PM_{2.5}, since most of the ground observation stations are located in urban areas. Wind has a weaker influence in winter and autumn than in spring and summer; thus, PM_{2.5} is primarily produced locally in the cold seasons from heating and fuel combustion but is spread from other nearby areas in the warm seasons.

Supporting Information

S1 Appendix. Raw dataset of PM_{2.5} concentration, gas-phase concentrations (SO₂, NO₂, CO, and O₃), AOT, temperature, wind direction, wind scale and locations of 13 observation stations (longitude and latitude).
(XLSX)

Author Contributions

Conceived and designed the experiments: YZS HLY QS YL. Performed the experiments: YZS QS YL. Analyzed the data: YZS HLY JHP YRS QS YL. Wrote the paper: YZS HLY QS YL.

References

1. Seinfeld JH (1989) Urban Air-Pollution—State of the Science. *Science* 243(4892): 745–752 doi: [10.1126/science.243.4892.745](https://doi.org/10.1126/science.243.4892.745) PMID: [ISI:A1989T16040027](https://pubmed.ncbi.nlm.nih.gov/16040027/).
2. Guo YM, Li SS, Tian ZX, Pan XC, Zhang JL, Williams G (2013) The burden of air pollution on years of life lost in Beijing, China, 2004–08: retrospective regression analysis of daily deaths. *Bmj-British Medical Journal* 347 doi: [10.1136/Bmj.F7139](https://doi.org/10.1136/Bmj.F7139) PMID: [ISI:000328574400006](https://pubmed.ncbi.nlm.nih.gov/22857440/).
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. (2012) A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380 (9859): 2224–2260. PMID: [ISI:000312387000017](https://pubmed.ncbi.nlm.nih.gov/22242260/). doi: [10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)
4. Sun Z, Cui Y (2013) An overview of PM_{2.5} impacts on human health (in Chinese). *Environmental Science & Technology* 26(4): 75–78.
5. Meng Z, Zhang Q (2007) Damage effects of dust storm PM_{2.5} on DNA in alveolar macrophages and lung cells of rats. *Food and Chemical Toxicology* 45(8): 1368–1374 doi: [10.1016/j.fet.2007.01.014](https://doi.org/10.1016/j.fet.2007.01.014) PMID: [ISI:000248970300007](https://pubmed.ncbi.nlm.nih.gov/17030007/).

6. Bell ML, Dominici F, Ebisu K, Zeger SL, Samet JM (2007) Spatial and temporal variation in PM_{2.5} chemical composition in the United States for health effects studies. *Environmental Health Perspectives* 115 (7): 989–995 doi: [10.1289/Ehp.9621](https://doi.org/10.1289/Ehp.9621) PMID: [ISI:000247716100021](https://pubmed.ncbi.nlm.nih.gov/15100021/).
7. Schlesinger RB (2007) The health impact of common inorganic components of fine particulate matter (PM_{2.5}) in ambient air: A critical review. *Inhalation Toxicology* 19(10): 811–832 doi: [10.1080/08958370701402382](https://doi.org/10.1080/08958370701402382) PMID: [ISI:000248680900002](https://pubmed.ncbi.nlm.nih.gov/18000002/).
8. Abbey DE, Ostro BE, Petersen F, Burchette RJ (1995) Chronic Respiratory Symptoms Associated with Estimated Long-Term Ambient Concentrations of Fine Particulates Less-Than 2.5 Microns in Aerodynamic Diameter (PM_{2.5}) and Other Air-Pollutants. *Journal of Exposure Analysis and Environmental Epidemiology* 5(2): 137–159. PMID: [ISI:A1995RR23400004](https://pubmed.ncbi.nlm.nih.gov/119950004/).
9. Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, et al. (1995) Particulate Air-Pollution as a Predictor of Mortality in a Prospective-Study of Us Adults. *American Journal of Respiratory and Critical Care Medicine* 151(3): 669–674. PMID: [ISI:A1995QL39100016](https://pubmed.ncbi.nlm.nih.gov/119950016/).
10. Schwartz J, Dockery DW, Neas LM (1996) Is daily mortality associated specifically with fine particles? *Journal of the Air & Waste Management Association* 46(10): 927–939. PMID: [ISI:A1996VL82800002](https://pubmed.ncbi.nlm.nih.gov/119960002/).
11. Federal R (1987) National ambient air quality standards for particulate matter. *Federal Register* 52 (126).
12. Federal R (2013) National ambient air quality standards for particulate matter. *Federal Register* 78(10).
13. Wang G, Wang P (2014) PM_{2.5} Pollution in China and Its Harmfulness to Human Health (in Chinese). *Science A Technology Review* 32(26): 72–78.
14. Xi'an Environmental Monitoring Station. 2014. Available: <http://www.xianemc.gov.cn>.
15. Li GX, Pan XC (2013) Evaluation of excess mortality caused by PM_{2.5} in four typical Chinese big cities. *Zhonghua Yi Xue Za Zhi* 93(34): 2703–6. PMID: [24360101](https://pubmed.ncbi.nlm.nih.gov/24360101/).
16. Watson JG, Chen LWA, Chow JC, Doraiswamy P, Lowenthal DH (2008) Source apportionment: Findings from the US Supersites program. *Journal of the Air & Waste Management Association* 58(2): 265–288 doi: [10.3155/1047-3289.58.2.265](https://doi.org/10.3155/1047-3289.58.2.265) PMID: [ISI:000253017700008](https://pubmed.ncbi.nlm.nih.gov/1500253017700008/).
17. Pui DYH, Chen SC, Zuo ZL (2014) PM_{2.5} in China: Measurements, sources, visibility and health effects, and mitigation. *Particuology* 13: 1–26 doi: [10.1016/j.partic.2013.11.001](https://doi.org/10.1016/j.partic.2013.11.001) PMID: [ISI:000336114500001](https://pubmed.ncbi.nlm.nih.gov/25000336114500001/).
18. Gao B, Guo H, Wang X, Zhao X, Ling Z, Zhang Z (2013) Tracerbased source apportionment of polycyclic aromatic hydrocarbons in PM_{2.5} in Guangzhou, southern China, using positive matrix factorization (PMF). *Environmental Science and Pollution Research International* 20: 2398–2409. doi: [10.1007/s11356-012-1129-0](https://doi.org/10.1007/s11356-012-1129-0) PMID: [22926284](https://pubmed.ncbi.nlm.nih.gov/22926284/)
19. Wang Y, Zhuang G, Tang A (2005) The ion chemistry and the source of PM_{2.5} aerosol in Beijing. *Atmospheric Environment* 39(21): 3771–3784.
20. Hu Z (2009) Spatial analysis of MODIS aerosol optical depth, PM_{2.5}, and chronic coronary heart disease. *International Journal of Health Geographics* 8: 27 doi: [10.1186/1476-072X-8-27](https://doi.org/10.1186/1476-072X-8-27) PMID: [19435514](https://pubmed.ncbi.nlm.nih.gov/19435514/).
21. Wang J, Christopher SA (2003) Intercomparison between satellite-derived aerosol optical thickness and PM_{2.5} mass: Implications for air quality studies. *Geophysical Research Letters* 30(21) doi: [10.1029/2003gl018174](https://doi.org/10.1029/2003gl018174) PMID: [ISI:000186507300002](https://pubmed.ncbi.nlm.nih.gov/1500186507300002/).
22. Gupta P, Christopher SA, Wang J, Gehrig R, Lee Y, Kumar N (2006) Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment* 40(30): 5880–5892 doi: [10.1016/j.atmosenv.2006.03.016](https://doi.org/10.1016/j.atmosenv.2006.03.016) PMID: [ISI:000241217500015](https://pubmed.ncbi.nlm.nih.gov/15000241217500015/).
23. Kim HS, Chung YS, Kim JT (2014) Spatio-temporal variations of optical properties of aerosols in East Asia measured by MODIS and relation to the ground-based mass concentrations observed in central Korea during 2001 similar to 2010. *Asia-Pacific Journal of Atmospheric Sciences* 50(2): 191–200. PMID: [ISI:000332488500006](https://pubmed.ncbi.nlm.nih.gov/25000332488500006/).
24. Ma Z, Hu X, Huang L, Bi J, Liu Y (2014) Estimating Ground-Level PM_{2.5} in China Using Satellite Remote Sensing. *Environmental Science & Technology* 48: 7436–7444.
25. Wang Z, Chen L, Tao J, Zhang Y, Su L (2010) Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sensing of Environment* 114: 50–63.
26. Lin CQ LY, Yuan ZB, Alexis KH, Li CC, Jimmy CH (2015) Using satellite remote sensing data to estimate the high-resolution distribution of ground-level PM_{2.5}. *Remote Sensing of Environment* 156: 117–128.
27. Chu DA, Kaufman YJ, Ichoku C, Remer LA, Tanre D, Holben BN (2002) Validation of MODIS aerosol optical depth retrieval over land. *Geophysical Research Letters* 29(12) doi: [10.1029/2001gl013205](https://doi.org/10.1029/2001gl013205) PMID: [ISI:000178964100006](https://pubmed.ncbi.nlm.nih.gov/15000178964100006/).

28. Schaap M, Apituley A, Timmermans RMA, Koelemeijer RBA, de Leeuw G (2009) Exploring the relation between aerosol optical depth and PM_{2.5} at Cabauw, the Netherlands. *Atmospheric Chemistry and Physics* 9(3): 909–925. PMID: [ISI:000263325900011](#).
29. Hoff RM, Christopher SA (2009) Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land? *Journal of the Air & Waste Management Association* 59(6): 645–675 doi: [10.3155/1047-3289.59.6.645](#) PMID: [ISI:000266797400002](#).
30. Boys BL MR, Donkelaar AV, MacDonell RJ, Hsu NC, Cooper MJ, Yantosca RM (2014) Fifteen-Year Global Time Series of Satellite-Derived Fine Particulate Matter. *Environmental Science & Technology* 48: 11109–11118.
31. Hu XF, Waller LA, Lyapustin A, Wang YJ, Al-Hamdan MZ, William LC, et al. (2014) Estimating ground-level PM_{2.5} concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sensing of Environment* 140: 220–232.
32. Liu Y, Franklin M, Kahn R, Koutrakis P (2007) Using aerosol optical thickness to predict ground-level PM_{2.5} concentrations in the St. Louis area: A comparison between MISR and MODIS. *Remote Sensing of Environment* 107(1–2): 33–44 doi: [10.1016/j.rse.2006.05.022](#) PMID: [ISI:000244829800005](#).
33. Tian J, Chen DM (2010) A semi-empirical model for predicting hourly ground-level fine particulate matter (PM_{2.5}) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. *Remote Sensing of Environment* 114(2): 221–229. PMID: [ISI:000274370400001](#).
34. Song W, Jia H, Huang J, Zhang Y (2014) A satellite-based geographically weighted regression model for regional PM_{2.5} estimation over the Pearl River Delta region in China. *Remote Sensing of Environment* 154: 1–7.
35. Kumar N, Chu A, Foster A (2007) An empirical relationship between PM_{2.5} and aerosol optical depth in Delhi Metropolitan. *Atmospheric Environment* 41(21): 4492–4503 doi: [10.1016/j.atmosenv.2007.01.046](#) PMID: [ISI:000247810500012](#).
36. Koelemeijer RBA, Homan CD, Matthijsen J (2006) Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmospheric Environment* 40(27): 5304–5315 doi: [10.1016/j.atmosenv.2006.04.044](#) PMID: [ISI:000239864200016](#).
37. Liu Y, Park RJ, Jacob DJ, Li QB, Kilaru V, Samat JA (2004) Mapping annual mean ground-level PM_{2.5} concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *Journal of Geophysical Research-Atmospheres* 109(D22) doi: [10.1029/2004jd005025](#) PMID: [ISI:000225430000005](#).
38. Tao J, Ho KF, Chen LG, Zhu LH, Han JL, Xu ZC (2009) Effect of chemical composition of PM_{2.5} on visibility in Guangzhou, China, 2007 spring. *Particuology* 7(1): 68–75 doi: [10.1016/j.partic.2008.11.002](#) PMID: [ISI:000264010400010](#).
39. Xin JY, Zhang Q, Wang LL, Gong CS, Wang YS, Liu ZR, et al. (2014) The empirical relationship between the PM_{2.5} concentration and aerosol optical depth over the background of North China from 2009 to 2011. *Atmospheric Research* 138: 179–188 doi: [10.1016/j.atmosres.2013.11.001](#) PMID: [ISI:000333504300014](#).
40. Wang YS, Xin JY, Li ZQ, Wang SG, Wang PC, Hao WM, et al. (2011) Seasonal variations in aerosol optical properties over China. *Journal of Geophysical Research-Atmospheres* 116 doi: [10.1029/2010jd015376](#) PMID: [ISI:000295536700001](#).
41. Yang C, Kan H, Chen R (2011) Research on level, composition, source and pollution characteristics of ambient fine particles in China (in Chinese). *Journal of Environmental Health* 28(8): 735–738.
42. Sun YL, Zhuang GS, Ying W, Han LH, Guo JH, Mo D, et al. (2004) The air-borne particulate pollution in Beijing—concentration, composition, distribution and sources. *Atmospheric Environment* 38(35): 5991–6004 doi: [10.1016/j.atmosenv.2004.07.009](#) PMID: [ISI:000224654100011](#).
43. Yang F, He K, Ma Y, Zhang Q, Yang X, Chan CK, et al. (2002) Chemical characteristics of PM_{2.5} species in Beijing ambient air (in Chinese). *J Tsinghua Univ (Sci & Tech)* 42(12): 1605–1608.
44. Zhang K, Wang Y, Wen T, Liu G, Xu H (2007) On-line analysis the water-soluble chemical of PM_{2.5} in late summer and early autumn in Beijing (in Chinese). *Acta Scientiae Circumstantiae* 27(3): 459–465.
45. Tai APK, Mickley LJ, Jacob DJ (2010) Correlations between fine particulate matter (PM_{2.5}) and meteorological variables in the United States: Implications for the sensitivity of PM_{2.5} to climate change. *Atmospheric Environment* 44(32): 3976–3984 doi: [10.1016/j.atmosenv.2010.06.060](#) PMID: [ISI:000281928900014](#).
46. Liu Y, Paciorek CJ, Koutrakis P (2009) Estimating Regional Spatial and Temporal Variability of PM_{2.5} Concentrations Using Satellite Data, Meteorology, and Land Use Information. *Environmental Health Perspectives* 117(6): 886–892 doi: [10.1289/Ehp.0800123](#) PMID: [ISI:000266462600022](#).

47. Tao L, Zhang C, Zhou B, Zhang B, Liu L (2011) Analysis of PM_{2.5} mass concentrations variation with high resolution in Xi'an City. *Environmental Science and Technology* 24(A01): 63–64.
48. Shen ZX, Cao JJ, Arimoto R, Han ZW, Zhang RJ, Han YM, et al. (2009) Ionic composition of TSP and PM_{2.5} during dust storms and air pollution episodes at Xi'an, China. *Atmospheric Environment* 43(18): 2911–2918 doi: [10.1016/j.atmosenv.2009.03.005](https://doi.org/10.1016/j.atmosenv.2009.03.005) PMID: [ISI:000266515900009](https://pubmed.ncbi.nlm.nih.gov/15900009/).
49. Cao JJ, Wu F, Chow JC, Lee SC, Li Y, Chen SW, et al. (2005) Characterization and source apportionment of atmospheric organic and elemental carbon during fall and winter of 2003 in Xi'an, China. *Atmospheric Chemistry and Physics* 5: 3127–3137. PMID: [ISI:000233609800001](https://pubmed.ncbi.nlm.nih.gov/15900001/).
50. Zhang C, Ding C, Zhou B (2013) Pollution characteristics of carbon aerosol in PM_{2.5} of Xi'an during winter and summer. *Chinese Journal of Environmental Engineering* 7(4): 1477–1481.
51. Hastie TJ, Tibshirani RJ (1990) *Generalized Additive Models* (CRC Press).
52. Li LF, Wu J, Hudda N, Sioutas C, Fruin SA, Delfino RJ (2013) Modeling the Concentrations of On-Road Air Pollutants in Southern California. *Environmental Science & Technology* 47(16): 9291–9299 doi: [10.1021/Es401281r](https://doi.org/10.1021/Es401281r) PMID: [ISI:000323471700037](https://pubmed.ncbi.nlm.nih.gov/23471700037/).
53. Sorek-Hamer M, Strawa AW, Chatfield RB, Esswein R, Cohen A, Broday DM (2013) Improved retrieval of PM_{2.5} from satellite data products using non-linear methods. *Environmental Pollution* 182: 417–423 doi: [10.1016/j.envpol.2013.08.002](https://doi.org/10.1016/j.envpol.2013.08.002) PMID: [ISI:000326661700051](https://pubmed.ncbi.nlm.nih.gov/26661700051/).
54. Zhang K, Batterman S (2010) Near-road air pollutant concentrations of CO and PM_{2.5}: A comparison of MOBILE6.2/CALINE4 and generalized additive models. *Atmospheric Environment* 44(14): 1740–1748 doi: [10.1016/j.atmosenv.2010.02.008](https://doi.org/10.1016/j.atmosenv.2010.02.008) PMID: [ISI:000277476200004](https://pubmed.ncbi.nlm.nih.gov/277476200004/).
55. Paciorek CJ, Liu Y, Moreno-Macias H, Kondragunta S (2008) Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level PM_{2.5}. *Environmental Science & Technology* 42(15): 5800–5806. PMID: [ISI:000258075100067](https://pubmed.ncbi.nlm.nih.gov/258075100067/).
56. Carslaw DC, Beevers SD, Tate JE (2007) Modelling and assessing trends in traffic-related emissions using a generalised additive modelling approach. *Atmospheric Environment* 41(26): 5289–5299 doi: [10.1016/j.atmosenv.2007.02.032](https://doi.org/10.1016/j.atmosenv.2007.02.032) PMID: [ISI:000248872900001](https://pubmed.ncbi.nlm.nih.gov/248872900001/).
57. Strawa AW, Chatfield RB, Legg M, Scarnato B, Esswein R (2011) Improving PM_{2.5} retrievals in the San Joaquin Valley using A-Train Multi-Satellite Observations. *Atmos. Chem. Phys. Discuss.* 11(11): 30563–30598 doi: [10.5194/acpd-11-30563-2011](https://doi.org/10.5194/acpd-11-30563-2011)
58. Glantz P, Kokhanovsky A, von Hoyningen-Huene W, Johansson C (2009) Estimating PM_{2.5} over southern Sweden using space-borne optical measurements. *Atmospheric Environment* 43(36): 5838–5846.
59. China Meteorological Administration. 2014. Available: <http://www.cma.gov.cn>.
60. Ho KF, Ho SS, Dai WT, Cao JJ, Huang RJ, Tian L, et al. (2014) Seasonal variations of monocarbonyl and dicarbonyl in urban and sub-urban sites of Xi'an, China. *Environ Monit Assess* 186(5): 2835–49 doi: [10.1007/s10661-013-3584-6](https://doi.org/10.1007/s10661-013-3584-6) PMID: [24420739](https://pubmed.ncbi.nlm.nih.gov/24420739/).
61. Li Z, Yan ZW (2010) Application of Multiple Analysis of Series for Homogenization to Beijing daily temperature series (1960–2006). *Advances in Atmospheric Sciences* 27(4): 777–787. PMID: [ISI:000278929300006](https://pubmed.ncbi.nlm.nih.gov/278929300006/).
62. Yan ZW, Li Z, Li QX, Jones P (2010) Effects of site change and urbanisation in the Beijing temperature series 1977–2006. *International Journal of Climatology* 30(8): 1226–1234. PMID: [ISI:000280023100011](https://pubmed.ncbi.nlm.nih.gov/280023100011/).
63. Wang J, Sun J, Wang S, Zhang H, Xue R, Hui Y (2011) Impact of Basin Terrain Urban Heat Island Effect on Precipitation in Surrounding Region of Xi'an. *Journal of Arid Meteorology* 29(2): 174–181.
64. (2013) Xi'an Municipal People's Government, Local Records Office Xi'an City. *Xi'an Year Book*: Chinese Publishing Group, World Book Inc.
65. City LRCCoXa (1996) *Local Records of Xi'an City*: Xi'an Press.
66. Lu YP DJ, Hou JZ, Li YL (2006) Spatial-temporal variation characteristics of wind speed and wind directions in Shaanxi Province. *Journal of Shaanxi Meteorology* 01: 1–4.
67. Global Space Flight Center MODIS Level 1 and Atmosphere Archive and Distribution System Web (LAADS Web). 2014. Available: <http://ladsweb.nascom.nasa.gov>.
68. Glasser L (2007) Dealing with outliers: Robust, resistant regression. *Journal of Chemical Education* 84(3): 533–534. PMID: [ISI:000244486700032](https://pubmed.ncbi.nlm.nih.gov/244486700032/).
69. Moeed RA, Papritz A (2002) An empirical comparison of kriging methods for nonlinear spatial point prediction. *Mathematical Geology* 34(4): 365–386 doi: [10.1023/A:1015085810154](https://doi.org/10.1023/A:1015085810154) PMID: [ISI:000175571800001](https://pubmed.ncbi.nlm.nih.gov/175571800001/).
70. Wood S (2006) *Generalized additive models: an introduction with R*: CRC press.
71. Edwards DP, Emmons LK, Hauglustaine DA, Chu DA, Gille JC, Kaufman YJ, et al. (2004) Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability. *Journal of*

- Geophysical Research-Atmospheres 109(D24) doi: [10.1029/2004jd004727](https://doi.org/10.1029/2004jd004727) PMID: [ISI:000225882400001](https://pubmed.ncbi.nlm.nih.gov/16111111/).
72. Bian HS, Chin MA, Kawa SR, Yu HB, Diehl T, Kucsera T (2010) Multiscale carbon monoxide and aerosol correlations from satellite measurements and the GOCART model: Implication for emissions and atmospheric evolution. *Journal of Geophysical Research-Atmospheres* 115 doi: [10.1029/2009jd012781](https://doi.org/10.1029/2009jd012781) PMID: [ISI:000285585200005](https://pubmed.ncbi.nlm.nih.gov/2000285585200005/).
 73. Chi X, Winderlich J, Mayer JC, Panov AV, Heimann M, Birmili W, et al. (2013) Long-term measurements of aerosol and carbon monoxide at the ZOTTO tall tower to characterize polluted and pristine air in the Siberian taiga. *Atmospheric Chemistry and Physics* 13(24): 12271–12298 doi: [10.5194/acp-13-12271-2013](https://doi.org/10.5194/acp-13-12271-2013) PMID: [ISI:000329930000008](https://pubmed.ncbi.nlm.nih.gov/24933333/).
 74. Chen LWA, Doddridge BG, Dickerson RR, Chow JC, Mueller PK, Quinn J, et al. (2001) Seasonal variations in elemental carbon aerosol, carbon monoxide and sulfur dioxide: Implications for sources. *Geophysical Research Letters* 28(9): 1711–1714 doi: [10.1029/2000gl012354](https://doi.org/10.1029/2000gl012354) PMID: [ISI:000168243100011](https://pubmed.ncbi.nlm.nih.gov/11511111/).
 75. Lioussé C, Penner JE, Chuang C, Walton JJ, Eddleman H, Cachier H (1996) A global three-dimensional model study of carbonaceous aerosols. *Journal of Geophysical Research-Atmospheres* 101(D14): 19411–19432 doi: [10.1029/95jd03426](https://doi.org/10.1029/95jd03426) PMID: [ISI:A1996VE25800042](https://pubmed.ncbi.nlm.nih.gov/14111111/).
 76. Huang Y, Shen HZ, Chen H, Wang R, Zhang YY, Su S, et al. (2014) Quantification of Global Primary Emissions of PM_{2.5}, PM₁₀, and TSP from Combustion and Industrial Process Sources. *Environmental Science & Technology* 48(23): 13834–13843 doi: [10.1021/Es503696k](https://doi.org/10.1021/Es503696k) PMID: [ISI:000345803800034](https://pubmed.ncbi.nlm.nih.gov/24933333/).
 77. Costabile F, Desantis F, Hong WM, Liu FL, Salvatori R, Wang FJ, et al. (2006) Representativeness of urban highest polluted zones for sitting traffic-oriented air monitoring stations in a Chinese city. *Journal of Environmental Science and Technology* 40(1): 35–41 doi: [10.1029/Jsmeb.49.35](https://doi.org/10.1029/Jsmeb.49.35) PMID: [ISI:000236432800006](https://pubmed.ncbi.nlm.nih.gov/16111111/).
 78. Saiz-Lopez A, Adame JA, Notario A, Poblete J, Bolivar JP, Albaladejo J (2009) Year-Round Observations of NO, NO₂, O₃, SO₂, and Toluene Measured with a DOAS System in the Industrial Area of Puertollano, Spain. *Water Air and Soil Pollution* 200(1–4): 277–288 doi: [10.1007/s11270-008-9912-8](https://doi.org/10.1007/s11270-008-9912-8) PMID: [ISI:000266487700025](https://pubmed.ncbi.nlm.nih.gov/16111111/).
 79. Dogruparmak SC, Ozbay B (2011) Investigating Correlations and Variations of Air Pollutant Concentrations under Conditions of Rapid Industrialization—Kocaeli (1987–2009). *Clean-Soil Air Water* 39(7): 597–604 doi: [10.1002/clen.201000478](https://doi.org/10.1002/clen.201000478) PMID: [ISI:000293696300001](https://pubmed.ncbi.nlm.nih.gov/2000293696300001/).
 80. Notario A, Adame JA, Bravo I, Cuevas CA, Aranda A, Diaz-De-Mera Y, et al. (2014) Air pollution in the plateau of the Iberian Peninsula. *Atmospheric Research* 145: 92–104 doi: [10.1016/j.atmosres.2014.03.021](https://doi.org/10.1016/j.atmosres.2014.03.021) PMID: [ISI:000337983100008](https://pubmed.ncbi.nlm.nih.gov/24933333/).
 81. Deshmukh DK, Deb MK, Tsai YI, Mkoma SL (2011) Water Soluble Ions in PM_{2.5} and PM₁₀ Aerosols in Durg City, Chhattisgarh, India. *Aerosol and Air Quality Research* 11(6): 696–708 doi: [10.4209/aaqr.2010.04.0025](https://doi.org/10.4209/aaqr.2010.04.0025) PMID: [ISI:000298003500008](https://pubmed.ncbi.nlm.nih.gov/2000298003500008/).
 82. Bhaskar BV, Mehta VM (2010) Atmospheric Particulate Pollutants and their Relationship with Meteorology in Ahmedabad. *Aerosol and Air Quality Research* 10(4): 301–315 doi: [10.4209/aaqr.2009.10.0069](https://doi.org/10.4209/aaqr.2009.10.0069) PMID: [ISI:000278764500001](https://pubmed.ncbi.nlm.nih.gov/16111111/).
 83. Khodeir M, Shamy M, Alghamdi M, Zhong MH, Sun H, Costa M, et al. (2012) Source apportionment and elemental composition of PM_{2.5} and PM₁₀ in Jeddah City, Saudi Arabia. *Atmospheric Pollution Research* 3(3): 331–340 doi: [10.5094/Apr.2012.037](https://doi.org/10.5094/Apr.2012.037) PMID: [ISI:000310515500009](https://pubmed.ncbi.nlm.nih.gov/24933333/).
 84. Lin G, Fu J, Jiang D, Hu W, Dong D, Huang Y, et al. (2014) Spatio-Temporal Variation of PM_{2.5} Concentrations and Their Relationship with Geographic and Socioeconomic Factors in China. *International Journal of Environmental Research & Public Health* 11(1): 173–186.