Environmental Science and Pollution Research (2021) 28:66978–66989
https://doi.org/10.1007/s11356-021-14487-0

RESEARCH ARTICLE

Source apportionment of $PM_{2.5}$ during different haze episodes by PMF and random forest method based on hourly measured atmospheric pollutant

Xin Du[1] · Junwei Yang[1] · Zhimei Xiao[2] · Yingze Tian[1] · Kui Chen[2] · Yinchang Feng[1]

Received: 19 January 2021 / Accepted: 14 May 2021 / Published online: 9 July 2021

Abstract

Hourly measured $PM_{2.5}$-bound species, gases, and meteorological data were analyzed by the PMF receptor model to quantify source contributions, and by the random forest to estimate decisive factors of variations of $PM_{2.5}$, sulfur oxidation ratio (SOR), and nitrogen oxidation ratio (NOR) during different haze episodes. $PM_{2.5}$ variation was influenced by CO (17%), SO_2 (19%), NH_3 (12%), O_3 (10%), air pressure (P, 9.9%), and temperature (T, 10%) during the whole period. SOR was determined by SO_2 (15%), temperature (T, 9.8%), relative humidity (RHU, 15%), and pondus hydrogenii (pH, 35%), and NOR was influenced by NO_x (19%), O_3 (14%), NH_3 (13%), and RHU (15%). Three types of pollution episodes were captured. Process I was characterized by high CO (contributing 40% of $PM_{2.5}$ concentration variation estimated by the random forest) due to coal combustion for heating during winter in northern China. According to the PMF, coal combustion (32%) and secondary sources (38%) were both the most important contributors in the first stage, and then, when the RHU increased to above 80%, the highest contribution was from secondary sources (40%). Process II was during the Spring Festival and was characterized by 8.8 μg m^{-3} firework contribution. High SO_2 during this process, especially on the CNY's Eve, was observed due to the firework displays, and SO_2 gave a high contribution (24%) to $PM_{2.5}$ variation. Process III showed high ions and high RHU in summer with sulfate and nitrate contributing 44% and 22%, respectively. Furthermore, meteorological parameters and NH_3 play a key role on SOR and NOR.

Keywords Source appointment · Random forest · High time-resolved species · Air · Haze

Introduction

Particulate matter (PM) and gaseous pollutants have led to major environmental problems in the last few decades in southern and eastern Asia (Huang et al. 2014; Zhang et al. 2017). In the recent year, China has experienced extremely severe and persistent haze pollution, especially in the North China Plain (Wang et al. 2016; Cheng et al. 2016; Zou et al. 2017). Severe pollution is accompanied by extremely poor air quality and poor visibility, which threatens human health (Jerrett et al. 2017; Li et al. 2018; Almetwally et al. 2020). Considerable research has reported that the mechanisms are very complex and differ for different pollution episodes (Wu et al. 2019). However, pollution mechanisms and influence factors during different pollution episodes are still not well understood (Cheng et al. 2016; Xie et al. 2019). In order to investigate pollution mechanisms, severe pollution episodes were captured and studied by environmentalist.

The origins of PM, which can be emitted from primary sources and formed through chemical reactions from precursors, may rapidly change during haze episode (Notario et al. 2013; Zheng et al. 2014; Wu et al. 2020). So as to perform source apportionment of PM, receptor models based on high time-resolved species data have been used to investigate timely variations of source contributions (Eatough et al. 2008; Pancras et al. 2013; Gao et al. 2016; Han et al. 2016;), which is fundamental to understanding the influence factors of heavy pollution episodes. Meteorologists have confirmed that gaseous pollutants and meteorological parameters play important roles in haze (Kuwata

Responsible Editor: Gerhard Lammel

✉ Yingze Tian
tianyingze@hotmail.com

✉ Yinchang Feng
fengyc@nankai.edu.cn

[1] The State Environmental Protection Key Laboratory of Urban Air Particulate Matter Pollution Prevention and Control, College of Environmental Science and Engineering, Nankai University, Tianjin 300071, China

[2] Tianjin Environment Monitoring Center, Tianjin 300071, China

et al. 2017; Wang et al. 2019a). Gaseous pollutants can act as precursors of secondary aerosol (Schelden et al. 2017; Yao et al. 2018). Adverse meteorological conditions can lead to accumulation of pollutants and favor formation of secondary aerosol (Wang et al. 2019b). Thus, gaseous pollution and meteorological parameters are important for $PM_{2.5}$ heavy pollution and formation of secondary aerosol, but their relationships are very complex (Trivedi et al. 2014; Wang et al. 2014a). Nevertheless, there is nonlinear relationship between the decisive influencing factors (e.g., gaseous pollution and meteorological parameters) and $PM_{2.5}$ concentration variation. At present, machine learning algorithm would be used to better identify the decisive influencing factors on $PM_{2.5}$ levels and secondary formation. The algorithm provides novel machine learning–based framework for data analysis. In this study, random forest algorithm, a classification and regression tool (Svetnik et al. 2003), is introduced to study the formation mechanism and important influencing factors of air pollutants based on gaseous pollutants and meteorological parameters. In brief, this study has one main contribution that the research attempts to investigate the relationships between $PM_{2.5}$ concentrations and the gaseous pollution and between $PM_{2.5}$ and meteorological factors by machine learning method from the macroscopic perspective.

Comprehensive observations of chemical species, gases, PM mass concentrations, and meteorological parameters at a 1-h time resolution were conducted to explore influence factors of haze episodes. Different types of pollution episodes were observed during the sampling period. The levels of $PM_{2.5}$-bound species (such as NO_3^-, Cl^-, SO_4^{2-}, NH_4^+, Ca^{2+}, Na^+, Mg^{2+}, K^+, OC, EC), pollutants (such as PM_1, $PM_{2.5}$, PM_{10}, TSP, SO_2, O_3, NH_3, NO_x, CO), and meteorological parameters (such as wind speed, relative humidity, temperature) were investigated. Positive Matrix Factorization (PMF) would be applied for source apportionment of $PM_{2.5}$ based on chemical species. Then, random forest method was applied to estimate the impacts of gaseous pollutants and meteorological parameters on $PM_{2.5}$ concentrations, sulfur oxidation ratio (SOR), and nitrogen oxidation ratio (NOR). Our work used the machine learning algorithm to reveal the characteristics and the reasons for the formation of typical heavy pollution episodes. The severe pollution episodes provide researchers an opportunity to study different mechanisms of pollution. The methods and influence factors reported in this paper are applicable to other emerging economies or developing countries and have significance for efforts to design effective management strategies.

Methodology

Sampling site and study period

The high time-resolved measurements of chemical species, gases, and PM mass concentrations were conducted on the rooftop of a five-story building (the Tianjin Environmental Protection Bureau) in downtown Tianjin, which is a municipality directly under the Central Government of China. Tianjin, near the capital of China (Beijing), is in the Bohai Economic Circle. Tianjin is a megacity under rapid industrialization and urbanization in the last few decades and currently has a population of over 14 million and an area of 11,947 km^2. The sampling site was located in a mixed residential and commercial area surrounded by few direct industrial sources, high vehicular emissions, and construction areas. The observations were conducted in February, March, June, July, August, and September of 2015 and produced data for 4,344 h. The details of sampling have been reported in our previous publication (Tian et al. 2018a).

Sampling methods and instrumentation

To obtain high time-resolved speciate data for $PM_{2.5}$, instruments were used to detect ambient $PM_{2.5}$. Hourly concentrations of elemental carbon (EC) and organic carbon (OC) were detected through a semi-continuous EC/OC carbon aerosol analyzer (Model-4, Sunset Laboratory Inc., USA) (Dall'Osto et al. 2014) using the basic thermal/optical transmittance measurement protocol of the National Institute for Occupational Safety and Health (NIOSH). Two temperature stages were used to determine OC and EC: an aliquot of sample filter (2.1 cm^2) was heated stepwise to 820 °C in a furnace in a non-oxidizing atmosphere (100% He); the oxidizing oven was then cooled to 550 °C, and the filter was again gradually heated to 870 °C in an oxidizing atmosphere (98% He, 2% O_2). Evolved carbon was oxidized to CO_2 and detected by a non-dispersive infrared detector (NDIR) during each temperature step. The split point was quantified as the carbon evolved after the introduction of oxygen but before the point where transmittance became equal to its initial value. Calibration was performed by introducing a known amount of methane into the oven and measuring its constant response. The carbon that evolved before the split point was OC, whereas EC was measured as the carbon evolved after this point but prior to the methane calibration peak (Tiwari et al. 2013).

Hourly concentrations of ionic species in $PM_{2.5}$, including NO_3^-, Cl^-, F^-, SO_4^{2-}, NH_4^+, Ca^{2+}, Na^+, Mg^{2+}, and K^+, were measured using a URG 9000D ambient ion monitor (AIM) (Chapel Hill, NC). The AIM separated and analyzed each anion and cation through a particle collection system and ion chromatographs (ICs). The $PM_{2.5}$ samples were collected at a flow rate of 3 L min^{-1} by using a sharp cut cyclone (Manigrasso et al. 2010). Then, a liquid diffusion parallel-plate denuder was used to separate the gases from the aerosol samples. The water-soluble compositions of aerosol and gaseous pollutants were collected through four syringes installed into pre-concentrators and then injected into the ICs. Anion detection was conducted in a gradient elution program using a KOH solution at a flow rate

of 1.0 mL min^{-1}, and the cation analyzer was run with methane sulfonic acid at a flow rate of 0.5 mL min^{-1}.

Gaseous pollutants included SO_2, O_3, NO, NO_2, and CO, and PM mass concentrations were of four sizes: PM_1, $PM_{2.5}$, PM_{10}, and TSP. Mass concentrations of PMs were continuously measured by a tapered element oscillating microbalance (TEOM) mass sensor. The SO_2 in the atmosphere was determined by pulsed fluorescence technology. The O_3 was analyzed by a dual cell photometer, a concept adopted by the US NIST as a national standard. Atmospheric NOx was measured by chemiluminescence technology; the analyzer had isolated outputs for NO and NO_2 that could be individually calibrated. The CO was quantified based on the absorption of infrared radiation at a 4.6-μm wavelength.

For quality assurance and quality control (QA/QC), the particle stream entered an aerosol super-saturation chamber to increase particle growth to obtain higher efficiencies. After collection through four syringes, the aerosols and the gaseous pollutants were injected into the ICs within 1 h. Anion/cation calibration solutions were used for the calibration of the ICs on the AIM for at least 1 month. The minimum detection limits (MDLs) were as follows: 0.2 μg m^{-3} (Cl^-), 0.2 μg m^{-3} (F^-), 0.2 μg m^{-3} (NO_3^-), 0.2 μg m^{-3} (NO_2^-), 0.3 μg m^{-3} (SO_4^{2-}), 1.8 μg m^{-3} (NH_4^+), 2.3 μg m^{-3} (Ca^{2+}), 0.8 μg m^{-3} (Mg^{2+}), 0.5 μg m^{-3} (K^+), and 0.6 μg m^{-3} (Na^+). The EC/OC carbon aerosol analyzer was calibrated each month through a blank punch of a pre-heated quartz fiber filter and standard sucrose solutions (3.2 μg C μl^{-1}). The quartz fiber filter was changed each week during the analysis process. The MDLs were 0.45 μg cm^{-2} and 0.06 μg cm^{-2} for OC and EC, respectively. The MDLs for PM, NO, NO_2, SO_2, CO, and O_3 were as high as 0.1 μg m^{-3}, 0.40 ppb, 0.40 ppb, 0.5 ppb, 0.04 ppm, and 0.50 ppb, respectively. The flow calibration, gas tightness test, blank filter test, and standard sample calibration were all conducted for QA/QC.

Meteorological parameters, including temperature (T), relative humidity (RHU), air pressure (P), wind speed (WS), and wind direction (DD), used to evaluate the impact of meteorological conditions were acquired from an online database at http://rp5.by. The meteorological parameters were available for 2:00 AM, 5:00 AM, 8:00 AM, 11:00 AM, 2:00 PM, 5:00 PM, and 8:00 PM each day. The meteorological parameters were acquired at the Beimalu Weather Station in downtown Tianjin. It is about 7 km far away from the Tianjin Environmental Protection Bureau Building.

Except for chemical species, gases, and PM mass concentrations, SOR and NOR were also used as follows:

$$SOR = SO_4^{2-}/(SO_4^{2-} + SO_2)$$

$$NOR = NO_3^-/(NO_3^- + NO_x)$$

where SO_4^{2-} is the S mass in SO_4^{2-}, SO_2 is the S mass in SO_2, NO_3^- is the N mass in NO_3^-, and NO_x is the N mass in NO_x. TS and TN were used to indicate the total emissions of S and N (Chen and Xie 2014). SOR and NOR were used to indicate the reaction of sulfur and nitrogen (Chen and Xie 2014; Wang et al. 2018).

In this research, the aerosol pH was calculated by utilizing the thermodynamic model ISORROPIA-II (Fountoukis and Nenes 2007), which could predict the physical state and composition of atmospheric inorganic aerosols. The ISORROPIA-II model can solve forward problems in which T, relative humidity, and the concentrations of gas + aerosols were known (e.g., NH_3 NH_4^+), and reverse problems in which T, relative humidity, and the concentrations of aerosol (but not gas) species were known. The pH calculation utilized measurements of NH_3, NH_4^+, Na^+, K^+, Mg^{2+}, Ca^{2+}, SO_4^{2-}, NO_3^-, and Cl^- for Tianjin from February to September of 2015. In this study, the ISORROPIA-II model was run in the forward mode and assumed that the aerosol solutions were metastable, as a high degree of accuracy determined on the basis of measurements of semivolatile partitioning of certain species (e.g., NH_3/NH_4^+) (Guo et al. 2015; Song et al. 2018)

Source apportionment

EPA PMF 5.0 model (Paatero and Tapper 1994; Paatero 1997; EPA 2014) was applied for $PM_{2.5}$ source apportionment based on hourly measured $PM_{2.5}$-bound chemical species. The PMF is used to solve the mass balance between observed species concentrations X, estimated source profiles F, and estimated contributions G (Ogulei et al. 2006; Tian et al. 2014):

$$X_{ij} = \sum_{h=1}^{p} g_{ih} f_{hj} + e_{ij}$$

where x_{ij} is the measured concentration of the jth species in the ith sample, f_{hj} (g g^{-1}) is the estimated species profile of the hth source, g_{ih} is the estimated contribution by the hth factor to each sample, e_{ij} is the residuals, and p is the number of factors. Moreover, ME-2 can incorporate prior information such as chemical properties to apportion sources through auxiliary equations (Tian et al. 2018b).

In this work, the random forest algorithm, a non-linear, data-driven model, is employed to study the influence factor of $PM_{2.5}$, NOR, and SOR. Random forest performs tree-ensemble of Classification and Regression Trees (CART). Regression tree is trained to perform differentiable mapping from the m-dimensional input x to its leaf index (Chen and Guestrin 2016):

$$f(x) = \omega_q(x) \ (q : \mathbb{R}^m \rightarrow T, \omega \in \mathbb{R}^T)$$

Here, f is parameterized by the learned tree structure q (of T leaves) and leaf weights ω. Independent regression trees f can be trained by corresponding tree structure q and leaf weights ω.

Results and discussion

Concentrations

It was calculated that the average concentrations of PM_1, $PM_{2.5}$, PM_{10}, and TSP were 57, 96, 104, and 152 μg m^{-3}, respectively, from February 1, 2015, to September 30, 2015. The $PM_1/PM_{2.5}$, $PM_{2.5}/PM_{10}$, and PM_{10}/TSP ratios were 0.59, 0.92, and 0.70, respectively. These ratios usually varied with different pollution processes and can indicate different pollution mechanisms. As reported by Alastuey et al. (2004), the fine particles can be caused by a low proportion of secondary and primary combustion emitted species, and the high coarse part can be caused by the high load of other contributions such as natural and anthropogenic mineral dust.

The important components of $PM_{2.5}$ include NO_3^- (13%), SO_4^{2-} (11%), NH_4^+ (14%), OC (9.5%), EC (2.7%), and Cl^- (2.5%). And the hourly data were collected from February 1, 2015, to September 30, 2015. The annual average concentration of Cl^-, NO_3^-, NO_2^-, SO_4^{2-}, NH_4^+, Ca^{2+}, Mg^{2+}, K^+, Na^+, OC, EC, CO, SO_2, NO_x, O_3, and NH_3 were 2.5 μg m^{-3}, 11 μg m^{-3}, 0.40 μg m^{-3}, 9.7 μg m^{-3}, 13 μg m^{-3}, 0.18 μg m^{-3}, 0.07 μg m^{-3}, 0.87 μg m^{-3}, 0.77 μg m^{-3}, 9.4 μg m^{-3}, 2.7 μg m^{-3}, 1.3 μg m^{-3}, 28 μg m^{-3}, 56 μg m^{-3}, 103 μg m^{-3}, and 29 μg m^{-3}, respectively. Moreover, RHU, pH, and WD were recorded and their annual mean were 58%, 4.2, and 2.7 m/s, respectively.

Influence factors on $PM_{2.5}$, SOR, and NOR

To comprehensively analyze the role of meteorological factors and gaseous pollutant for affecting local $PM_{2.5}$ concentration, NOR and SOR, we employed random forest model. A number of factors were selected as follows: T, RHU, P, WS, DD, pH, CO, SO_2, NO_x, O_3, and NH_3. By analyzing variables through multiple nonlinear regression, the model quantitatively calculated the relative importance of different factors about $PM_{2.5}$, NOR, and SOR. As shown in Table 1.

Influence factors on $PM_{2.5}$

Based on the random forest, the results displayed that CO had the most important influence on $PM_{2.5}$ concentration and contributed 17% to $PM_{2.5}$ concentration variations. SO_2 and NO_x (totally 24%) also strongly linked with $PM_{2.5}$ concentration. It has been reported that in the condition of stable meteorological and enough residence time, SO_2 and NO_x can be transformed into sulfate and nitrate. P (9.9%) made the great contribution to $PM_{2.5}$ concentration among meteorological parameters. It may happen that when there was high pressure, down draft impeded going up of $PM_{2.5}$ and led to accumulation of particles. T (10%) was also an important factor for $PM_{2.5}$ concentration variation. Luo et al. (2017) noted that air convection relied on temperature, high temperature could lead to dilution of and $PM_{2.5}$, and it was on the contrary when temperature reduction, at the same time, high temperature could elevate formation rates of secondary aerosols. RHU (6.7%) and pH (4.8%) could accelerate formation of secondary aerosol and increased the chance of collision and adsorption between particles, which resulted in decreasing $PM_{2.5}$ concentration. DD (4.7%) and WS (1.1%) can also influence on $PM_{2.5}$ concentration variations. Wind speed could favor particles spread and diffusion of $PM_{2.5}$ and low $PM_{2.5}$ concentration (Wang et al. 2019b; Karimian et al. 2019). And different speed directions had a significant effect on the spatial distribution of $PM_{2.5}$ and could change transport of atmosphere pollutants.

Table 1 Different influence factors on $PM_{2.5}$, SOR and NOR during the whole period

Influence factors	$PM_{2.5}$	SOR	NOR
CO	0.17	0.030	0.080
SO_2	0.19	0.15	0.070
NOx	0.050	0.020	0.19
O_3	0.10	0.090	0.14
NH_3	0.12	0.050	0.13
Temperature (T)	0.10	0.098	0.050
air pressure (P)	0.099	0.030	0.080
Relative humidity (RHU)	0.067	0.15	0.15
Wind direction (DD)	0.047	0.020	0.040
Pondus hydrogenii (pH)	0.048	0.35	0.060
Wind speed (WS)	0.011	0.010	0.020

Influence factors on SOR and NOR

The importance analysis of gaseous pollutants (CO, SO_2, NO_x, O_3, NH_3) and meteorological parameters (T, RHU, P, DD, WS, and pH) on SOR and NOR were also constructed based on random forest method. It was found that pH (35%), SO_2 (15%), RHU (15%), T (9.8%), and O_3 (9%) played important role on SOR. As reported in this work and related work, SOR in the atmosphere was closely related to pH (Fuzzi 1978; Brimblecombe and Spedding 1972). Sakamoto et al. (2004) found that O_3 had noticeable effect on the amount of SO_2 oxidation. And RHU was significantly positively correlated with SOR and NOR (Yao et al. 2020). In addition, temperature could promote the rate of sulfur oxidation. Meanwhile, NO_x (19%), O_3 (14%), NH_3 (13%), and RHU (15%) had a greater influence on NOR. Seinfeld and Pandis (1998) have found that the formation of nitrogen oxidation was dominated by the following reaction:

$$NO_2 + OH + M \rightarrow HNO_3 + M$$
$$NO_2 + O_3 \rightarrow NO_3 + O_2$$
$$NO_2 + O_3 + M \rightarrow N_2O_5 + M$$
$$N_2O_5 + H_2O \rightarrow 2HNO_3$$
$$NH_3 + H_2O \rightarrow NH_4^+ + OH$$

Characteristics of typical pollution episodes

During the sampling period, three typical pollution episodes were captured and selected for further analysis. The concentrations of species in $PM_{2.5}$ during each episode were introduced into the USEPA PMF 5.0 model for source apportionment, and the gaseous pollutants and meteorological parameters were introduced into the random forest model to explore the influence factors. The factor profiles of three episodes estimated USEPA PMF 5.0 is exhibited in Fig. 1. Coal combustion, traffic emission, and resuspended dust were consistently identified for three pollution episodes. Tian et al. (2020) has explored how to better conduct PMF during haze episodes, showing that the PMF performance was poor for some episodes through whole-based mode (using all data of whole sampling period as input). Thus, the episode-based mode (using data of each episode) was used in this work to conduct PMF. In addition, the consistent results of PMF and random forest based on different datasets demonstrate the reliability of two methods. We add the bootstrapping (BS) to estimate the uncertainty of the PMF solution. BS involves resampling the input dataset, fitting PMF model parameters for this resampled dataset, and then using the variations among these resampled or "bootstrapped" fitted profiles to estimate the uncertainty of the initial PMF solution (Norris et al. 2014). The BS results of each episode are listed in Table 2. Thus, 4, 6, and 5 factors were selected for episodes I, II and III, respectively, because the BS values were all higher than 75%, indicating the BS uncertainties can be interpreted. The coal combustion was

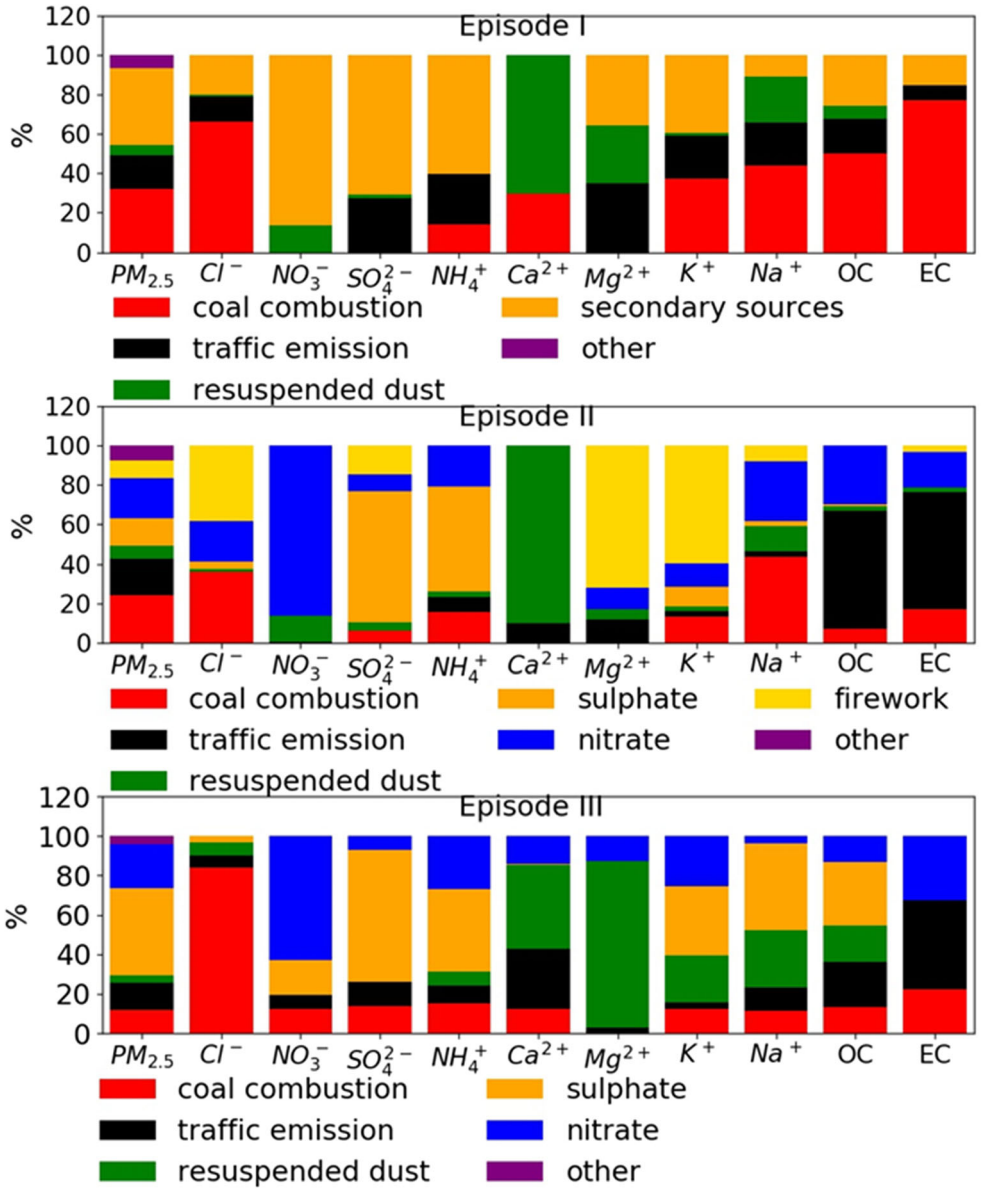

Fig. 1 The factor profiles of three episodes estimated USEPA PMF 5.0. Episode I is characterized by heavy coal combustion; the five factors, coal combustion, traffic emission, resuspended dust, secondary sources, and other, make different contributions to various components in episode I; episode II is characterized by the Spring Festival; the seven factors, coal combustion, traffic emission, resuspended dust, sulfate, nitrate, firework, and other, make different contributions to various components in episode II; episode III is characterized by high inorganic ions and high RH; the six factors, coal combustion, traffic emission, resuspended dust, sulfate, nitrate, and other, make different contributions to various components in episode III

Table 2 The bootstrapping (BS) results of three typical pollution episodes

	Episode I		Episode II		Episode III	
	4 factors	5 factors	6 factors	7 factors	5 factors	6 factors
BS	80–90%	50–90%	75–100%	30–100%	100%	50–100%

identified by high loadings of Cl^-, OC, and EC (Shen et al. 2010). The traffic emission was characterized by OC and EC but low Cl^- (Liu et al. 2017). The resuspended dust was identified by relatively high weights of Ca^{2+} and Mg^{2+} (Shen et al. 2011). Factors associated with NO_3^-, SO_4^{2-}, and NH_4^+ were distinguished as secondary sources or nitrate and sulfate (Guan et al. 2019). The OC and EC in the secondary sources may be linked with secondary organic carbon and transport of combustion emissions (Guan et al. 2019).

Process I

The first interesting episode was found from February 14, 2015, 0:00 AM to February 15, 2015, 4:00 AM; $PM_{2.5}$ concentrations ranged from 157 to 313 μg m^{-3}; then, after a short decrease, $PM_{2.5}$ returned to relatively higher levels (February 15, 2015, 4:00 PM–February 16, 2015, 12:00 PM). As shown in Fig. 2(b), the abundances of OC and EC were higher than those for other sampling period. In general, OC and EC were markers of combustion (Chow et al. 2004). According to the PMF results in Fig. 1 (episode I), for $PM_{2.5}$ concentration, the average contributions of coal combustion, traffic emission, resuspended dust, and secondary sources were 32%, 17%, 5.5%, and 38%, respectively. As demonstrated in Fig. 2(a), in the first stage, coal combustion (88 μg m^{-3}, 40%) and secondary sources (98 μg m^{-3}, 41%) were both the most important contributors, and in the second stage, the highest contribution was from secondary sources (63 μg m^{-3}, 38%). In this stage, coal combustion had emitted high SO_2, which formed secondary SO_4^{2-} under the high RH in the second stage. And the secondary sources were primary pollution source.

Based on the result of random forest in Fig. 2(c), CO (40%) remained at a high contribution for $PM_{2.5}$ concentration variation. Previous study noted that CO mainly resulted from the coal combustion for heating during the winter in northern China (Du et al. 2016; Wang et al. 2010). The results indicated that P (20%) and T (16%) also had a great influence on $PM_{2.5}$ contribution variation in the heavy pollution process. Therefore, PMF and the random forest results were consistent and both of them demonstrated that this heavy pollution was characterized by coal combustion.

In addition, SOR was analyzed by in putting gaseous pollutants and meteorological parameter in random forest model; the results are shown in Fig. 2(d); it was obvious that, among all the factors, SO_2 (43%) was the most important factor for SOR. In this pollution process, DD (22%) greatly affected the SOR transformation and it was the second largest factor. A suggested explanation is that wind direction determined gaseous pollutant pathway. Definitely, pH (6%) and RHU (7%) also served as pivotal role. When it came to NOR in Fig. 2(e), NO_x (24%) and O_3 (17%) occupied primary important factor for NOR formation variation. Moreover, in this process, CO (11%), DD (10%), and pH (12%) had more impact on NOR.

Process II

The second interesting episode was observed from February 18, 2015, 12:00 AM to February 21, 2019, 11:00 PM, which covered the Spring Festival. During the period of sampling, the concentration of $PM_{2.5}$ ranged from 31 to 407 μg m^{-3}. The chemical compositions of $PM_{2.5}$ during this process were characterized by high K^+ (3.9%) and Cl^- (4.0%), which were much higher than the average of the entire period (1.0% for K^+ and 2.7% for Cl^-). Compared with the episode one, NH_4^+ and SO_4^{2-} were much less; however, both of them were slightly higher than the annual average, respectively (as shown in Fig. 3 (b)). As reported in our previous work and the literature (Tian et al. 2014; Sarkar et al. 2010), Cl^- and K^+ were considered as firework-related species and could be markers for fireworks. K^+ and Cl^-, commonly in the form of perchlorate or chlorate, are major components in black powder and act as the main oxidizers during burning. The corresponding chemical equations are $2KClO_3 = 2KCl + 3O_2$ and $KClO_4 = KCl + 2O_2$.

Based on the PMF results in Fig. 1 (episode II), in process II, for $PM_{2.5}$ concentration, the average contributions of traffic emission, coal combustion, nitrate, resuspended dust, sulfate, firework, and unknown sources were 18%, 24%, 20%, 6.6%, 14%, 8.8%, and 7.5%, respectively. As demonstrated in Fig. 3(a), in this process, coal combustion (24 μg m^{-3}) and nitrate (20 μg m^{-3}) were important contributors, followed by traffic emission (18 μg m^{-3}), sulfate (13 μg m^{-3}), firework (8.6 μg m^{-3}), and resuspended dust (6.4 μg m^{-3}). It was found that firework was an important pollution source. A further discussion that gaseous pollutants and meteorological parameter had impact on $PM_{2.5}$ concentration variation by the random forest, as visualized in Fig. 3(c), the result showed that SO_2 (24%) kept at a high contribution for heavy pollution. Previous study noted that the number of firework was lighted, and it became a major pollution source in the Spring Festival (Wang et al. 2007; Li et al. 2017). CO (12%) and NOx (13%) were also considered major contribution for $PM_{2.5}$ contribution, and both were closely to related to coal combustion emissions and vehicle exhaust (Kota et al. 2014). Besides, P (25%) and pH (11%) also had a great influence on the heavy pollution. Therefore, combining PMF and the random results showed

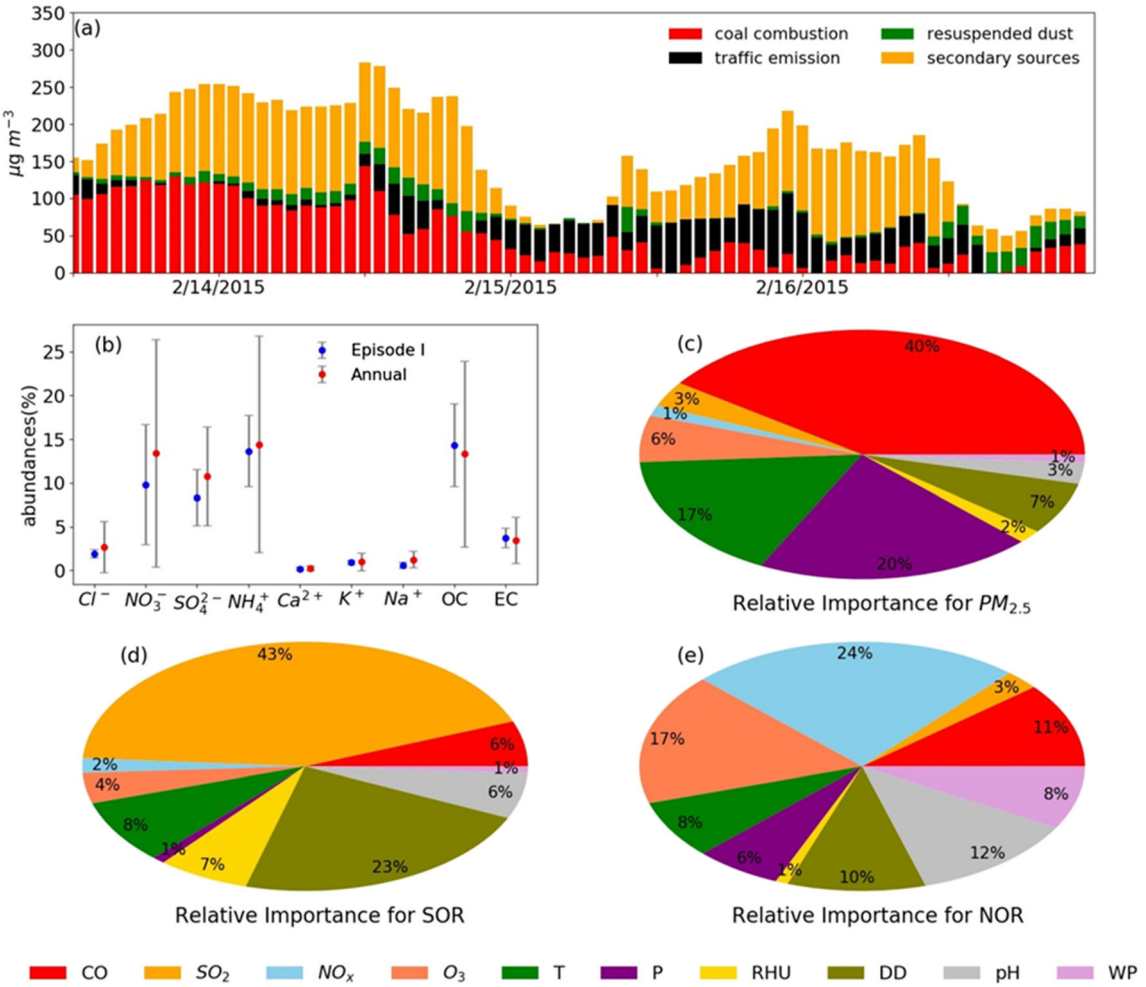

Fig. 2 The abundance and concentration of chemical species, PMF result, and random forest result during process I (episode characterized by heavy coal combustion). **a** the concentration contribution from coal combustion, traffic emission, resuspended dust, and secondary sources change over time in episode I; **b** the difference between the average mass percent of $PM_{2.5}$ components in episode I and all year; **c** gases and meteorological parameters have different contributions to $PM_{2.5}$ variation in episode I; **d** gases and meteorological parameters have different contributions to SOR in episode I; **e** gases and meteorological parameters have different contributions to NOR in episode I

that this heavy pollution was characterized by firework combustion.

In terms of the random forest results for SOR in Fig. 3(d), it was obvious that, among all the gaseous pollutants and meteorological parameters, precursor SO_2 (27%) was the most important factor for SOR; in this pollution process, P (22%) greatly affected the SOR transformation. Definitely, pH (7%) and RHU (13%) also served as pivotal role. When it came to NOR, as visualized in Fig. 3(e), NO_x (35%) occupied primary important factor, followed by P (20%), RHU (10%), and O_3 (8%).

Process III

The third pollution episode was characterized by high inorganic ions and high RHU level (51–95%) in summer. The $PM_{2.5}$ mass concentrations were at relatively higher levels, ranging from 81 to 291 μg m^{-3}. The value of SOR and NOR were 0.55 and 0.27, respectively. As visualized in Fig. 4(b), high fractions of NO_3^- (15%) and SO_4^{2-} (16%) during the pollution episode can be found and their mass concentrations were much higher than that during other sampling period.

Based on the PMF results, as demonstrated in Fig. 1 (episode III), the average contributions of traffic emission, coal combustion, resuspended dust, nitrate, sulfate, and other sources for $PM_{2.5}$ mass concentration were 13%, 12%, 4.1%, 22%, 44%, and 3.9%, respectively. As shown in Fig. 4(a), sulfate (82 μg m^{-3}) and nitrate (42 μg m^{-3}) were the most important contributors. Other source contributors were traffic emission (25 μg m^{-3}), coal burning (22 μg m^{-3}), and resuspended dust (7.7 μg m^{-3}).

According to the results of the random forest in Fig. 4(c), in process III, it was estimated that the relative importance of CO, SO_2, NH_3, and RHU to $PM_{2.5}$ concentration variation

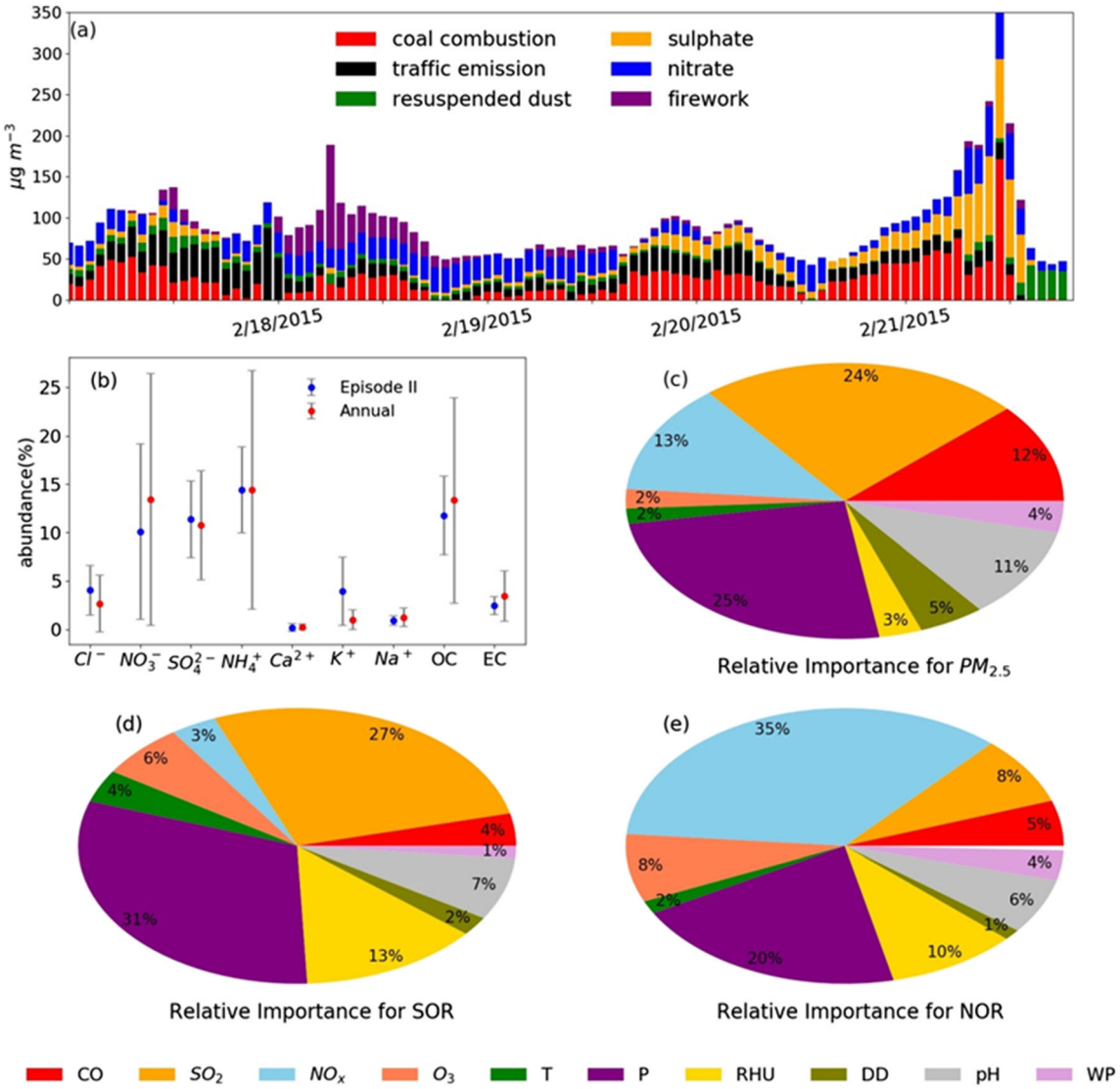

Fig. 3 The abundance and concentration of chemical species, PMF result, and random forest result during process II (episode characterized by the Spring Festival). **a** The concentration contribution from coal combustion, traffic emission, resuspended dust, sulfate, nitrate, and firework change over time in episode II; **b** the difference between the average mass percent of $PM_{2.5}$ components in episode II and all year; **c** gases and meteorological parameters have different contributions to $PM_{2.5}$ variation in episode II; **d** gases and meteorological parameters have different contributions to SOR in episode II; **e** gases and meteorological parameters have different contributions to NOR in episode II

were 40%, 13%, 8.3%, and 15%, respectively. Simultaneously, it was calculated that SO_2 (24%), NH_3 (28%), and T (19%) were the main influence factors for SOR. And the results showed that NOR was greatly affected by NO_x and NH_3, and the importance of both were 12% and 45%, respectively. NOx, NH_3, and SO_2 mainly came from vehicle exhaust and coal burning (Xie et al. 2005; Meng et al. 2011). Under the higher RHU, NH_3 was an important alkaline gas in the atmosphere, and it can react with acid gases (SO_2, NO_x, etc.) to form secondary aerosols (Zhang et al. 2012). This episode can be identified as reaction caused by high RH. The strong influence of RH on the reaction of inorganic ions was worthy of notice. High RH was favorable for the formation of sulfate and nitrate in the aqueous phase and may also improve sulfate and nitrate partitioning into the liquid phase formed by a gas-phase homogeneous reaction of precursors. Alkaline aerosol particles in northern China have been reported to lead to high PH values (Kulshrestha et al. 1998), which could also enhance the equilibria reactions. Therefore, relative humidity provided a key catalyst to chemically react to secondary aerosol formation. It was apparent that this episode was a common type of haze in northern China with characteristic features of high inorganic aerosols, high RH, stagnant meteorological conditions with low mixing heights, and large emissions of primary air pollutants causing secondary pollution.

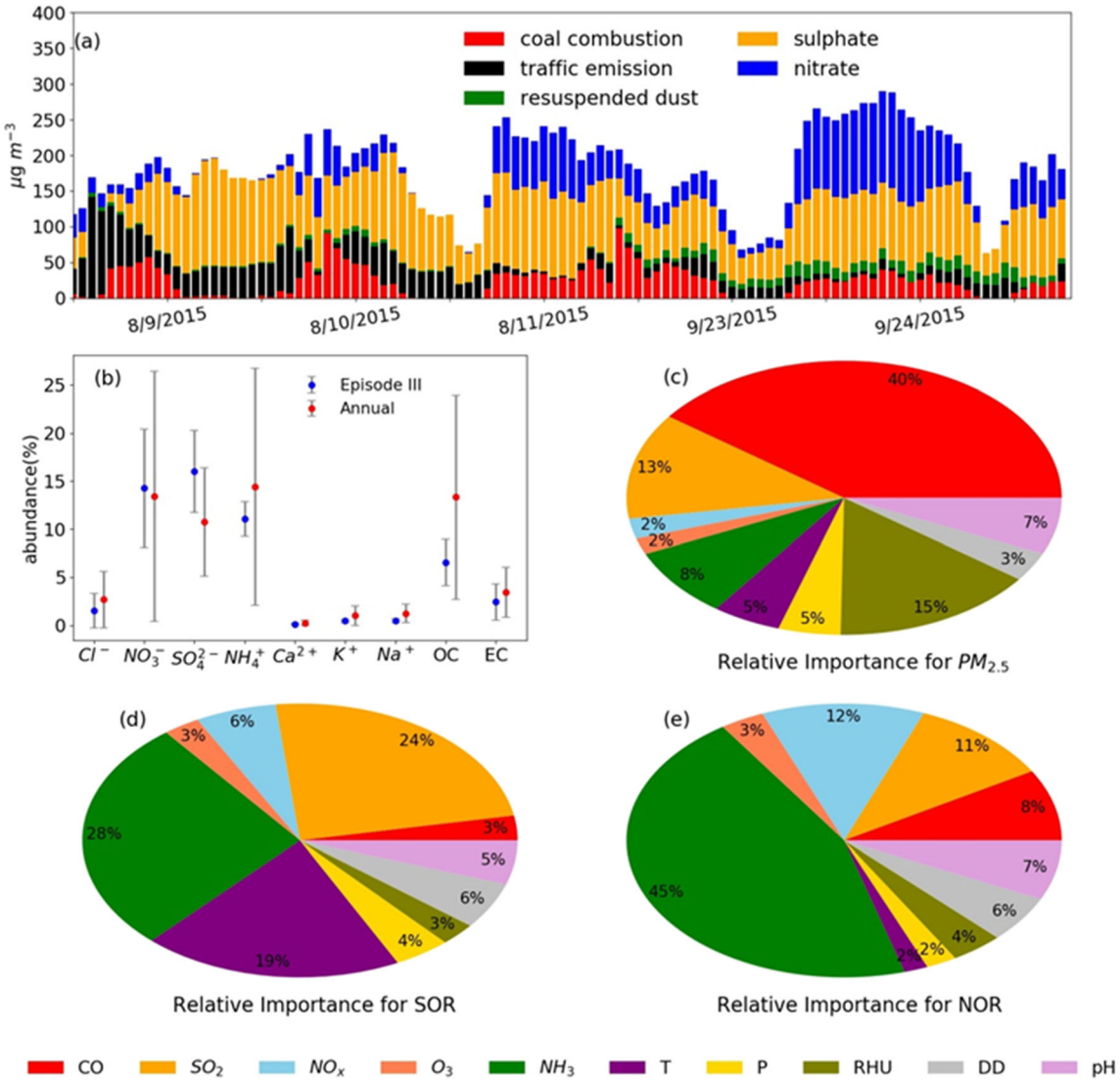

Fig. 4 The abundance and concentration of chemical species, PMF result and random forest result during process III (episode characterized by high inorganic ions and high RH). **a** The concentration contribution from coal combustion, traffic emission, resuspended dust, sulfate, and nitrate change over time in episode III; **b** the difference between the average mass percent of $PM_{2.5}$ components in episode III and all year; **c** gases and meteorological parameters have different contributions to $PM_{2.5}$ variation in episode III; **d** gases and meteorological parameters have different contributions to SOR in episode III; **e** gases and meteorological parameters have different contributions to NOR in episode III

Conclusions

Hourly measured $PM_{2.5}$-bound species, gases, and meteorological data were analyzed by the PMF receptor model to quantify source contributions and by the random forest regression method to identify the contribution of gaseous pollution and meteorological parameters on the changes of $PM_{2.5}$ concentrations, SOR, and NOR. It was found that CO, SO_2, NOx, P, and T played an important role in $PM_{2.5}$ variation; pH, SO_2, RHU, and T were the dominating influence factors for SOR; and NO_x, O_3, NH_3, and RHU determined the NOR for the whole sampling period. Proposed PMF-random forest method was used to quantify the source contributions to $PM_{2.5}$.

Three types of pollution episodes were captured. The first pollution episode (process I) was characterized by heavy coal combustion due to heating in the winter in northern China. CO remained at a high level and contributed 40% of $PM_{2.5}$ concentration variation. According to the PMF results, coal combustion and secondary sources were the most important contributors in the first stage, and then, the highest contribution was from secondary sources. The second pollution episode (process II) happened during the Spring Festival and was characterized by high K^+ and Cl^-. Firework combustion contributed 8.8 μg m^{-3} estimated by PMF. High SO_2 during process II, especially on the CNY's Eve, was observed due to the firework displays, and SO_2 gave a high contribution (24%) to

$PM_{2.5}$ concentration variation. The second pollution episode (process III) was characterized by high inorganic ions and high RH in summer. Sulfate (44 μg m^{-3}) and nitrate (22 μg m^{-3}) were the most important contributors. It was estimated that the relative importance of CO, SO_2, NH_3, and RHU to $PM_{2.5}$ concentration variation were 40%, 13%, 8.3%, and 15%, respectively. Through comparing the influence factors for SOR and NOR during episode I in winter and episode III in summer, DD was more important in episode I indicating that this episode was strongly influenced by the transportation; while NH_3 was more important in episode III indicating alkaline precursor strongly influence the episode in summer.

Authors' contributions Junwei Yang, Zhimei Xiao, and Kui Chen performed the experiments.

Xin Du made the figures. Xin Du and Yingze Tian wrote the paper. All the authors contributed to the submitted version of the manuscript. Yingze Tian and Yinchang Feng acquired the funding.

Funding This study is supported by the National Natural Science Foundation of China (91544226 and 41977181) and Tianjin Science and Technology Program (18ZXSZSF00160).

Data availability All relevant data are within the manuscript and available from the corresponding author upon request.

Declarations

Ethical approval Not applicable.

Consent to participate All authors participated in this work.

Consent to publish All authors agree to publish.

Conflict of interest The authors declare that they have no conflict of interest.

References

Alastuey A, Querol X, Rodríguez S, Plana F, Lopez-Soler A, Ruiz C, Mantilla E (2004) Monitoring of atmospheric particulate matter around sources of secondary inorganic aerosol. Atmos Environ 38: 4979–4992

Almetwally AA, Bin-Jumah M, Allam AA (2020) Ambient air pollution and its influence on human health and welfare: an overview. Environ Sci Pollut Res 27:24815–24830

Brimblecombe P, Spedding DJ (1972) Rate of solution of gaseous sulphur Dioxide at atmospheric concentrations. Nature 236:225

Chen Y, Xie S (2014) Characteristics and formation mechanism of a heavy air pollution episode caused by biomass burning in Chengdu, Southwest China. Sci Total Environ 473–474:507–517

Cheng YF, Zheng GJ, Wei C, Mu Q, Zheng B, Wang ZB, Gao M, Zhang Q, He KB, Carmichael G, Pöschl U, Su H (2016) Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China. Sci Adv 2:e1601530

Chow JC, Watson JG, Kuhns H, Etyemezian V, Lowenthal DH, Crow D, Kohl SD, Engelbrecht JP, Green MC (2004) Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. Chemosphere 54:185–208

Chen TQ, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA. 785-794. https://doi.org/10.1145/2939672.2939785

Dall'Osto M, Hellebust S, Healy RM, O'Connor IP, Kourtchev I, Sodeau JR, Ovadnevaite J, Ceburnis D, O'Dowd CD (2014) Apportionment of urban aerosol sources in Cork (Ireland) by synergistic measurement techniques. Sci Total Environ 493:197–208

Du QQ, Zhang CL, Mu YJ, Cheng Y, Zhang YY, Liu CT (2016) An important missing source of atmospheric carbonyl sulfide: domestic coal combustion. Geophys Res Lett 43:8720–8727

Eatough DJ, Grover BD, Woolwine WR, Eatough NL, Long R, Farber R (2008) Source apportionment of 1 h semi-continuous data during the 2005 Study of Organic Aerosols in Riverside (SOAR) using positive matrix factorization. Atmos Environ 42:2706–2719

Fuzzi S (1978) Study of iron(III) catalysed sulphur dioxide oxidation in aqueous solution over a wide range of pH. Atmos Environ 12:1439–1442

Fountoukis C, Nenes A (2007) ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+-Ca^{2+}-Mg^{2+}-NH_4^+-Na^+-SO_4^{2-}-NO_3^--Cl^--H_2O aerosols. Atmos Chem Phys 7:4639–4659

Guo H, Xu L, Bougiatioti A, Cerully KM, Capps SL, Hite JR Jr, Carlton AG, Lee SH, Bergin MH, Ng NL, Nenes A, Weber RJ (2015) Fine-particle water and pH in the southeastern United States. Atmos Chem Phys 15:5211–5228

Gao J, Peng X, Chen G, Xu J, Shi GL, Zhang YC, Feng YC (2016) Insights into the chemical characterization and sources of $PM_{2.5}$ in Beijing at a 1-h time resolution. Sci. Total Environ 542:162–171

Guan L, Liang YL, Tian YZ, Yang ZR, Sun YM, Feng YC (2019) Quantitatively analyzing effects of meteorology and PM 2.5 sources on low visual distance. Sci. Total Environ 659:764–772

Han B, Zhang R, Yang W, Bai ZP, Ma ZQ, Zhang WJ (2016) Heavy haze episodes in Beijing during January 2013: inorganic ion chemistry and source analysis using highly time-resolved measurements from an urban site. Sci. Total Environ 544:319–329

Huang RJ, Zhang Y, Bozzetti C, Ho KF, Cao JJ, Han Y, Daellenbach KR, Slowik JG, Platt SM, Canonaco F, Zotter P, Wolf R, Pieber SM, Bruns EA, Crippa M, Ciarelli G, Piazzalunga A, Schwikowski M, Abbaszade G, Schnelle-Kreis J, Zimmermann R, An Z, Szidat S, Baltensperger U, Haddad I, Prévôt ASH (2014) High secondary aerosol contribution to particulate pollution during haze events in China. Nature 514:218–222

Jerrett M, Turner MC, Beckerman BS, Pope CA, van Donkelaar A, Martin RV, Serre M, Crouse D, Gapstur SM, Krewski D (2017) Comparing the health effects of ambient particulate matter estimated using ground-based versus remote sensing exposure estimates. Environ. Health Perspect 125:552–559

Kulshrestha UC, Saxena A, Kumar N, Kumari KM, Srivastava SS (1998) Chemical composition and association of size-differentiated aerosols at a suburban site in a semi-arid tract of India. J Atmos Chem 29:109–118

Kuwata M, Kai FM, Yang LDQ, Itoh M, Gunawan H, Harvey CF (2017) Temperature and burning history affect emissions of greenhouse gases and aerosol particles from tropical peatland fire. J Geophys Res-Atmos 122:1281–1292

Kota SH, Zhang HL, Chen G, Schade GW, Ying Q (2014) Evaluation of on-road vehicle CO and NO_x national emission inventories using an urban-scale source-oriented air quality model. Atmos Environ 85:99–108

Karimian H, Li Q, Li CC, Chen G, Mo YQ, Wu CL, Fan JX (2019) Spatio-temporal variation of wind influence on distribution of fine particulate matter and its precursor gases. Atmos Pollut Res 10:53–64

Li YJ, Wang JH, Ren BN, Wang HL, Qiao LP, Zhu JP, Li L (2018) The characteristics of atmospheric phthalates in Shanghai: a haze case study and human exposure assessment. Atmos Environ 178:80–86

Luo J, Du P, Samat A, Xia J, Che M, Xue Z (2017) Spatiotemporal pattern of $PM_{2.5}$ concentrations in Mainland China and analysis of its influencing factors using geographically weighted regression. Sci Rep 7:40607

Li JY, Xu TT, Liu XH, Chen H, Nizkorodov SA, Chen JM, Yang X, Mo ZY, Chen ZM, Liu HL, Mao JY, Liang GY (2017) Online single particle measurement of fireworks pollution during Chinese New Year in Nanning. J Environ Sci 53:184–195

Liu B, Li TK, Yang JM, Wu JH, Wang J, Gao JX, Bi XH, Feng YC, Zhang YF, Yang HH (2017) Source apportionment and a novel approach of estimating regional contributions to ambient PM2.5 in Haikou, China. Environ Pollut 223:334–345

Manigrasso M, Abballe F, Jack RF, Avino P (2010) Time-resolved measurement of the ionic fraction of atmospheric fine particulate matter. J Chromatogr Sci 48:549–552

Meng ZY, Lin WL, Jiang XM, Yan P, Wang Y, Zhang YM, Jia XF, Yu XL (2011) Characteristics of atmospheric ammonia over Beijing, China. Atmos Chem Phys 11:6139–6151

Notario A, Bravo I, Adame JA, Díaz-de-Mera Y, Aranda A, Rodríguez A, Rodríguez D (2013) Behaviour and variability of local and regional oxidant levels ($OX = O_3 + NO_2$) measured in a polluted area in central-southern of Iberian Peninsula. Environ Sci Pollut Res 20: 188–200

Norris GA, Duvall R, Brown SG, Bai S (2014) EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide. Prepared for the, U.S. Environmental Protection Agency Office of Research and Development, Washington, DC (EPA/600/R-14/108; STI910511-5594-UG, April).

Ogulei D, Hopke PK, Zhou L, Pancras JP, Nair N, Ondov JM (2006) Source apportionment of Baltimore aerosol from combined size distribution and chemical composition data. Atmos Environ 40:396–410

Paatero P (1997) Least squares formulation of robust non-negative factor analysis. Chemom Intell Lab Syst 37:23–35

Paatero P, Tapper U (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5:111–126

Pancras JP, Landis MS, Norris GA, Vedantham R, Dvonch JT (2013) Source apportionment of ambient fine particulate matter in Dearborn, Michigan, using hourly resolved PM chemical composition data. Sci Total Environ 448:2–13

Sarkar S, Khillare PS, Jyethi DS, Hasan A, Parween M (2010) Chemical speciation of respirable suspended particulate matter during a major firework festival in India. J Hazard Mater 184:321–330

Schelden VG, de Foy B, Herring C, Kaspari S, VanReken T, Jobson B (2017) Contributions of wood smoke and vehicle emissions to ambient concentrations of volatile organic compounds and particulate matter during the Yakima wintertime nitrate study. J Geophys Res Atmos 122:1871–1883

Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP (2003) Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. J Chem Inf Comput Sci 43:1947–1958

Song S, Gao M, Xu W, Shao J, Shi G, Wang S, Wang Y, Sun Y, McElroy MB (2018) Fine-particle pH for Beijing winter haze as inferred from different thermodynamic equilibrium models. Atmos Chem Phys 18:7423–7438

Sakamoto K, Takada H, Sekiguchi K (2004) Influence of ozone, relative humidity, and flow rate on the deposition and oxidation of sulfur dioxide on yellow sand. Atmos Environ 38:6961–6967

Seinfeld JH, Pandis SN (1998) From Air Pollution to Climate Change. Wiley, New York

Shen ZH, Cao JJ, Arimoto R, Han YM, Zhu CS, Tian J, Liu SX (2010) Chemical Characteristics of Fine Particles (PM1) from Xi'an, China. Aerosol Sci Technol 44(6):461–472

Shen ZH, Cao JJ, Liu SX, Zhu CS, Wang X, Zhang T, Xu HM, Hu TF (2011) Chemical composition of PM_{10} and $PM_{2.5}$ collected at ground level and 100 meters during a strong winter-time pollution episode in Xi'an, China. J Air Waste Manage Assoc 61(11):1150–1159

Tian YZ, Liu JY, Han SQ, Shi XR, Shi GL, Xu H, Yu HF, Zhang YF, Feng YC, Russell AG (2018a) Spatial, seasonal and diurnal patterns in physicochemical characteristics and sources of $PM_{2.5}$ in both inland and coastal regions within a megacity in China. J Hazard Mater 342:139–149

Tiwari S, Srivastava AK, Bisht DS, Safai PD, Parmita P (2013) Assessment of carbonaceous aerosol over Delhi in the Indo-Gangetic Basin: characterization, sources and temporal variability. Nat Hazards 65:1745–1764

Tian YZ, Wang J, Peng X, Shi GL, Feng YC (2014) Estimation of the direct and indirect impacts of fireworks on the physicochemical characteristics of atmospheric PM_{10} and $PM_{2.5}$. Atmos Chem Phys 14:9469–9479

Tian YZ, Xue QQ, Xiao ZM, Chen K, Feng YC (2018b) PMF-GAS Methods to estimate contributions of sources and oxygen for $PM_{2.5}$, based on highly time-resolved $PM_{2.5}$ species and gas data. Aerosol Air Qual Res 18:2956–2966

Trivedi DK, Ali K, Beig G (2014) Impact of meteorological parameters on the development of fine and coarse particles over Delhi. Sci Total Environ 478:175–183

Tian YZ, Zhang YF, Liang YL, Niu ZB, Xue QQ, Feng YC (2020) PM2.5 source apportionment during severe haze episodes in a Chinese megacity based on a 5-month period by using hourly species measurements: explore how to better conduct PMF during haze episodes. Atmos Environ 224:117364

U.S. Environmental Protection Agency (EPA), EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide, (2014) www.epa.gov.

Wang XQ, Wei W, Cheng SY, Zhang C, Duan WJ (2018) A monitoring-modeling approach to SO_4^{2-} and NO_3^- secondary conversion ratio estimation during haze periods in Beijing, China. J Environ Sci 78: 293–302

Wang S, Liao TT, Wang LL, Sun Y (2016) Process analysis of characteristics of the boundary layer during a heavy haze pollution episode in an inland megacity, China. J Environ Sci 40:138–144

Wang GH, Cheng CL, Huang Y, Tao J, Ren YQ, Wu F, Meng JJ, Li JJ, Cheng YT, Cao JJ, Liu SX, Zhang T, Zhang R, Chen YB (2014a) Evolution of aerosol chemistry in Xi'an, inland China, during the dust storm period of 2013 – Part 1: Sources, chemical forms and formation mechanisms of nitrate and sulfate. Atmos Chem Phys 14: 11571–11585

Wang SB, Yin SS, Zhang RQ, Yang LM, Zhao QY, Zhang LS, Yan QS, Jiang N, Tang XY (2019a) Insight into the formation of secondary inorganic aerosol based on high-time-resolution data during haze episodes and snowfall periods in Zhengzhou, China. Sci Total Environ 660:47–56

Wang ZF, Li J, Wang Z, Yang WY, Tang X, Ge BZ, Yan PZ, Zhu LL, Chen XS, Chen HS, Wand W, Li JJ, Liu B, Wang XY, Zhao YL, Lu N, Su DB (2014b) Modeling study of regional severe hazes over mid-eastern China in January 2013 and its implications on pollution prevention and control. Sci China Earth Sci 57:3–13

Wang SX, Wei W, Li D, Aunan K, Hao JM (2010) Air pollutants in rural homes in Guizhou, China – Concentrations, speciation, and size distribution. Atmos Environ 44:4575–4581

Wang Y, Zhuang GS, Xu C, An ZS (2007) The air pollution caused by the burning of fireworks during the lantern festival in Beijing. Atmos Environ 41:417–431

Wang XQ, Wei W, Cheng SY, Yao S, Zhang HY, Zhang C (2019b) Characteristics of PM2.5 and SNA components and meteorological factors impact on air pollution through 2013–2017 in Beijing, China. Atmos. Pollut. Res 10:1976–1984

Wu JM, Zhang YJ, Wang T, Qian YL (2020) Rapid improvement in air quality due to aerosol-pollution control during 2012–2018: an evidence observed in Kunshan in the Yangtze River Delta, China. Atmos Pollut Res 11:693–701

Wu P, Huang XJ, Zhang JK, Luo B, Luo JQ, Song HY, Zhang W, Rao ZH, Feng YP, Zhang JQ (2019) Characteristics and formation mechanisms of autumn haze pollution in Chengdu based on high time-resolved water-soluble ion analysis. Environ Sci Pollut Res 26: 2649–2661

Xie YZ, Liu ZR, Wen TX, Huang XJ, Liu JY, Tang GQ, Yang Y, Li XG, Shen RR, Hu B, Wang YS (2019) Characteristics of chemical composition and seasonal variations of $PM_{2.5}$ in Shijiazhuang, China: impact of primary emissions and secondary formation. Sci. Total Environ 677:215–229

Xie SD, Yu T, Zhang YH, Zeng LM, Qi L, Tang XY (2005) Characteristics of PM_{10}, SO_2, NO_x and O_3 in ambient air during the dust storm period in Beijing. Sci Total Environ 345:153–164

Yao L, Garmash O, Bianchi F, Zheng J, Yan C, Kontkanen J, Junninen H, Mazon SB, Ehn M, Paasonen P, Sipilä M, Wang MY, Wang XK, Xiao S, Chen HF, Lu YQ, Zhang BW, Wang DF, Fu QY, Geng FH, Li L, Wang HL, Qiao LP, Yang X, Chen JM, Kerminen VM, Petäjä T, Worsnop DW, Kulmala M, Wang L (2018) Atmospheric new particle formation from sulfuric acid and amines in a Chinese megacity. Science 361:278–281

Yao Q, Liu ZR, Han SQ, Cai ZY, Liu JL, Hao TY, Liu JY, Huang XJ, Wang YS (2020) Seasonal variation and secondary formation of size-segregated aerosol water-soluble inorganic ions in a coast megacity of North China Plain. Environ Sci Pollut Res 27:26750–26762

Zhang Q, Jiang X, Tong D, Davis SJ, Zhao H, Geng G, Feng T, Zheng B, Lu Z, Streets DG, Ni R, Brauer M, van Donkelaar A, Martin RV, Huo H, Liu Z, Pan D, Kan H, Yan Y, Lin J, He KB, Guan DB (2017) Transboundary health impacts of transported global air pollution and international trade. Nature 543:705–709

Zheng M, Zhao X, Cheng Y, Yan C, Shi W, Zhang X, Weber RJ, Schauer JJ, Wang X, Edgerton ES (2014) Sources of primary and secondary organic aerosol and their diurnal variations. J Hazard Mater 264: 536–544

Zou YF, Wang YH, Zhang YZ, Koo JH (2017) Arctic sea ice, Eurasia snow, and extreme winter haze in China. Sci Adv 3:e1602751

Zhang H, Shen Z, Wei X, Zhang M, Li Z (2012) Comparison of optical properties of nitrate and sulfate aerosol and the direct radiative forcing due to nitrate in China. Atmos Res 113:113–125