

Supplementary Material

Effect of uncertainty on source contributions from the positive matrix factorization model for a source apportionment study

Guo-Liang Shi¹, Hao Chen¹, Ying-Ze Tian^{1,*}, Dan-Lin Song², Lai-Dong Zhou², Fei Chen³, Hao-Fei Yu⁴, Yin-Chang Feng^{1,*}

¹ State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution Prevention and Control, College of Environmental Science and Engineering, Nankai University, Tianjin, 300071, P. R. China

² Chengdu Research Academy of Environmental Sciences, China

³ Nanjing Institute of Environmental Sciences, MEP, China

⁴ Georgia Institute of Technology, Atlanta, Georgia 30332, United States

* Corresponding author. Tel.: +86 22 23503397

E-mail: tianyingze@hotmail.com (Y.Z. Tian) and fengyc@nankai.edu.cn (Y.C. Feng)

Methods and materials

PMF and the uncertainty σ_{ij} for PMF

Positive Matrix Factorization (PMF), developed by Paatero and Tapper (1994), is a powerful method for source apportionment. PMF attempts to apportion source profiles and source contributions based on observations at a receptor site:

$$x_{ij} = \sum_{h=1}^p g_{ih} f_{hj} + e_{ij} \quad (\text{S1})$$

where x_{ij} is the j^{th} elemental concentration measured in the i^{th} sample; f_{hj} ($\mu\text{g}/\mu\text{g}$) is the fraction of the j^{th} element in the h^{th} source; g_{ih} is the contribution of the h^{th} source to the i^{th} sample; e_{ij} is the residuals; and p the number of factors extracted. The g_{ih} and f_{hj} in PMF are constrained to non-negative values (Paatero, 2007).

The goal of PMF is to minimize an ‘object function’ Q which is defined as:

$$Q = \sum_{i=1}^n \sum_{j=1}^m (e_{ij} / \sigma_{ij})^2 \quad (\text{S2})$$

where the value σ_{ij} is the “uncertainty” in the j^{th} species for the i^{th} sample, which is the focus of this work. The uncertainty σ_{ij} (also noted as standard deviation s_{ij} or $X_std\text{-dev}$ in some works) is used to down weight observations that include various errors. Generally, users are expected to specify uncertainty values so that agreement are achieved between model and observed data array. Initial values of σ_{ij} are determined by users, the solution is then obtained through an iterative minimization algorithm (Paatero, 2007).

User can specify an uncertainty matrix that includes both laboratory uncertainties and model uncertainties (EPA, 2009). Uncertainty σ_{ij} is computed based on three uncertainty parameters (C1, C2, and C3) using “Errormodel” (EM) (Table S2). The uncertainty parameters are defined as follows: C1 is associated with laboratory uncertainty; C2 applies only to Poisson-distributed data and is set to zero for the

majority of environmental applications; C3 is model uncertainty coefficient, which describes expected residuals that are not caused by laboratory errors, such as variations of source profiles with time (EPA, 2009). Table S3 summarizes some of the available EM and how uncertainties are calculated using uncertainty parameters C1, C2, and C3. In most EM, C3 is multiplied by the species concentration x_{ij} , as it is assumed to vary with the magnitude of the observed value.

Several PMF models are available, including the PMF2, EPA PMF 5.0, and ME-2. PMF2 (provided by Dr. Pentti Paatero at the University of Helsinki, Finland) was selected for this study in consideration of its fast performance, which is beneficial for large number of model runs.

Synthetic Dataset Development

To evaluate the impact of uncertainties σ on the accuracy of PMF results, synthetic datasets with various errors (inherent characteristics of all environmental datasets) were developed and each dataset was analyzed using PMF with different uncertainty levels.

The method for simulating datasets can be found at (Zeng and Hopke, 1992). Two types of errors were considered: errors associated with the magnitude of the measured concentrations (denoted as e_x); and errors due to the temporal variation of source profiles (denoted as e_f). The Monte Carlo method was used to simulate datasets (Lee and Russell, 2007). The Equations for simulating datasets are as follows:

$$x_{ij} = \sum_{h=1}^p g_{ih} (f_{hj} + e_f^{hj}) + e_x^{ij} \quad (S3)$$

The magnitude of the errors was controlled by parameters H_f and H_x :

$$e_f^{ih} = H_f f_{hi}(\mathit{rand}) \quad (S4)$$

$$e_x^{ij} = H_x x_{ij}(\mathit{rand}) \quad (S5)$$

where rand is a random number generator producing normally distributed random numbers from $N(0,1)$. Profiles of secondary sources were kept constant (Lee and Russell, 2007). H_f was set to 0.01 for primary sources (Zeng and Hopke (1992)), and to 0 for secondary sources. H_x were specified as 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 0.7, 0.9, to simulate datasets with various error levels.

For practical reasons, collinearity of source profiles and source contributions were not taken into account. Three ambient particle sources, Light-duty gasoline vehicles (LDGV), soil dust (SDUST) and ammonium nitrate (AMNITR) were selected and their source profiles are described in Marmur et al. (2005) and Shi et al. (2011), and also summarized in Table S4. The contributions of each source are simulated by combining their corresponding mean contributions with randomly generated errors that are normally distributed and with variances as shown in Table 1 (Habre et al. (2011)). Each of the synthetic dataset contains 300 “daily samples”, 22 species, and 3 sources.

Data Analysis

First, the impacts of uncertainties σ on synthetic datasets with various errors ($H_x = 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 0.7, 0.9$) were investigated and the C3 value was same for all of the species in this scenario (SCE 1). For each H_x , ten simulated datasets were generated, resulting in a total of 100 datasets. Each of these datasets was analyzed by PMF with ten different levels of uncertainties σ , so at least a total of 1000 computations were calculated by PMF in SCE 1.

Typically in environmental work, C1 is associated with detection limits and is usually considered as a constant for one species in a work (Paatero, 2007). In terms of the recommendation in the user’s Guide (Paatero, 2007), C1 equals to two units of the least significant digit reported for x_{ij} for all the computations (as shown in Table S4).

Usually, appropriate values of C3 need to be explored by users, so diverse C3 were specified in this work. C3 equals to ten values (C3 = 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 0.7, 0.9), so that the uncertainties σ introduced to different computations elevated with the C3 (as shown in Table S5).

Furthermore, as mentioned above, one can subjectively decrease the uncertainties for some variables to emphasize their weights. To study the influence of this question on the PMF performance, three scenarios were investigated: (1) if all the sources can be indentified correctly, the C3 for all the tracers of three sources (EC and OC for LDGV, Si for SDUST, and NO_3^- for AMNITR) were decreased (SCE 2); (2) if only AMNITR can be determined, the C3 for the tracer of AMNITR source (NO_3^-) were decreased (SCE 3); (3) if a subjective Fe-industry source was identified mistakenly, the C3 of an additional element – Fe (defined as noise species) were decreased (SCE 4). Similar with SCE 1, in each scenario, error values $H_x = 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 0.7, 0.9$ were used to simulate datasets; ten datasets were produced for each H_x ; C3 = 0.0125/0.025, 0.025/0.05, 0.05/0.1, 0.075/0.15, 0.1/0.2, 0.125/0.25, 0.15/0.3, 0.25/0.5, 0.35/0.7, 0.45/0.9 for tracers/other elements. The C1 values were same with SCE 1 as shown in Table S4; and C3 values for SCE 2-4 were listed in Table S5. At least, a total of 3000 computations were performed by PMF for these three scenarios.

During model execution of PMF2, the FPEAK option was set to 0 for all PMF runs in consideration of feasibility and model performance. In addition, the PMFs model were executed in robust mode, with 3 factors, 5 base runs for each case and error model code EM=-14.

Two metrics were used for model evaluation: the percentage of reasonable results (DP) and the average absolute error (AAE). In many cases, especially when errors

added to the synthetic datasets are large, unreasonable or undivided sources (two or more sources were extracted into one factor) might be obtained, but undesired. To avoid such unwanted results, the divided ratio, defined as the number of divided tests to that of total tests for each case (equal to 10 in this work), was calculated first. Cases with low DP values were excluded for further evaluations.

The AAE (the percent absolute errors of average mass contributions) between “true” and estimated contributions is defined as: (Javitz et al., 1988):

$$AAE(\%) = 100 \times \frac{1}{n} \sum_{i=1}^n \frac{|E_i - T_i|}{T_i} \quad (S6)$$

where n is the number of samples (300 in this work); E_i is the estimated contributions of a certain source; T_i is the true (simulated) contributions corresponding to E_i . The AAEs for undivided cases were specified as 100.

Sampling and Chemical Analysis of the Ambient PM_{2.5} Dataset

To investigate the impact of uncertainties on the source appointment of PM, an ambient PM_{2.5} dataset was also analyzed using PMF and with diverse uncertainty inputs. The PM_{2.5} samples were collected in 2012 in the megacity of Chengdu in China. Chengdu, also known as the home of giant pandas, is one of the most important economic, transportation and communication centers in southwestern China, and is the capital of Sichuan Province. Population of this city is more than 10 million.

Ambient PM_{2.5} samples were collected using filter-based samplers (TH-150 Medium Volume Sampler, Wuhan Tianhong Intelligence Instrumentation Facility, China). Two parallel medium-volume air samplers were used: one with polypropylene membrane filters (90mm diameter, Beijing Synthetic Fiber Research Institute, China), and the other with quartz-fiber filters (90mm diameter, type 2500QAT-UP, Pall Life Sciences, USA). Blank quartz-fiber filters and polypropylene membrane filters were

baked in an oven at 400 and 60°C for 2 h, respectively, to remove any volatile matters. After that, each blank filter was placed in a silica gel desiccator for more than three days. The samplers were then pre-calibrated, and the sampling volume of the pumps were set at 100 L/min and run continuously for 24 h. A total of 133 PM_{2.5} samples were obtained in the sampling campaign.

Then, the concentrations of 17 elements (i.e., Na, Mg, Al, Si, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, and Pb), total carbon (TC) and 2 water soluble ions (i.e., NO₃⁻ and SO₄²⁻) were measured. Elements in the ambient samples collected on the polypropylene membrane filters were analyzed via inductively coupled plasma (IRIS Intrepid II, Thermo Electron). PM_{2.5} samples collected on the quartz fiber filters were used to analyze the water soluble ions and carbon species present. Water soluble ions (i.e., NO₃⁻ and SO₄²⁻) were extracted using an ultrasonic extraction system (AS3120, AutoScience) and analyzed via ion chromatography (DX-120, DIONEX). A clip from each quartz-fiber filter was taken for the TC measurement, which was determined by DRI/OGC carbon analyzers.

Detailed information on the sampling area, sampling method, chemical analysis and quality assurance/quality control were reported in previous studies (Shi et al., 2009; Tian et al., 2013b).

References

- Habre, R.; Coull, B.; Koutrakis, P. Impact of source collinearity in simulated PM_{2.5} data on the PMF receptor model solution. *Atmos. Environ.* **2011**, *45*, 6938-6946.
- Javitz, H. S.; Watson, J. G.; Robinson, N. Performance of the chemical mass balance model with simulated local-scale aerosols. *Atmos. Environ.* **1988**, *22*, 2309-2322.

- Marmur, A.; Unal, A.; Mulholland, J. A.; Russell, A. G. Optimization-based source apportionment of PM_{2.5} incorporating gas-to-particle ratios. *Environ. Sci. Technol.* **2005**, *39*, 3245-3254.
- Paatero, P.; Tapper, U. Positive Matrix Factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **1994**, *5*, 111-126.
- Paatero, P. User's guide for Positive Matrix Factorization Programs PMF2 and PMF3, Part 1-2: tutorial. *University of Helsinki, Helsinki, Finland.* **2004**.
- Shi, G. L.; Tian, Y. Z.; Zhang, Y. F.; Ye, W. Y.; Li, X.; Tie, X. X.; Feng, Y. C.; Zhu, T. Estimation of the concentrations of primary and secondary organic carbon in ambient particulate matter: Application of the CMB-Iteration method. *Atmos. Environ.* **2011**, *45*, 5692-5698.
- U.S. Environmental Protection Agency. Guidance Document for PMF Applications with the Multilinear Engine. **2009**.
- Zeng, Y. S.; Hopke, P. K. A new receptor model: A direct trilinear decomposition followed by a matrix reconstruction. *J. Chemometr.* **1992**, *6*, 65-83.

Table S1. Two examples for introduction. In each example, one dataset was analyzed with different uncertainties.

(a)

		Factor 1	Factor 2	Factor 3
1	Tracers	NO ₃ ⁻	OC and EC	Si
	Source category	AMNITR	LDGV	SDUST
	Source contribution	5.12	9.20	12.55
2	Tracers	NO ₃ ⁻ , OC and EC	Si	Fe
	Source category	AMNITR and LDGV	SDUST	Fe-industrial source
	Source contribution	11.52	12.09	3.23

(b)

		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	Tracers	Ca	AL, Si, Fe, TC	NO ₃ ⁻	SO ₄ ²⁻	Ze
	Source category	cement dust	soil dust and vehicular exhaust	secondary nitrate	secondary sulphate	Zn-related source
	Source contribution	0.67	108.79	0.10	9.13	8.40
2	Tracers	AL, Si, TC	Ca and SO ₄ ²⁻	Fe	NO ₃ ⁻ and SO ₄ ²⁻	Ze
	Source category	Soil dust	Cement dust and secondary sulphate	Fe-industrial source	secondary nitrate and secondary sulphate	Zn-related source
	Source contribution	53.58	6.20	59.42	3.73	3.05

Two examples (the data were from Table 1 and Figure 3 in the following work) listed in Table S1 indicate that widely different results might be obtained by PMF with different uncertainties, in spite of one dataset being analyzed.

Table S2. The definition of uncertainty parameters (Paatero, 2007).

parameter	names	definition
Uncertainty parameters	C1	laboratory uncertainty, dependent on data set, typically = the detection limit in environmental work
	C2	0 except for poisson-distributed data
	C3	Model uncertainty coefficient. It is assumed to vary with the magnitude of the observed value.

Table S3. Some of the available Errormodel codes (EM) and how each is calculated using the uncertainty parameters C1, C2, and C3 (Paatero, 2004).

EM	Equation	Notes
-5	$\sigma_{ij} = C1_{ij} + C2_{ij}\sqrt{sumabs} + C3_{ij}sumabs$	Sumabs is the sum of absolute values of contributions to the fitted value. Used in equations where the difference (or sum) of two values is pulled toward zero. $s = x_{ij} $, where x_{ij} is the measured data.
-12	$\sigma_{ij} = C1_{ij} + C2_{ij}\sqrt{ s } + C3_{ij} s $	Commonly used in environmental applications.
-13	$\sigma_{ij} = C1_{ij} + C2_{ij}\sqrt{ s } + C3_{ij} s $	$s = y_{ij} $, where y_{ij} is the model-fitted value
-14	$\sigma_{ij} = C1_{ij} + C2_{ij}\sqrt{ s } + C3_{ij} s $	$s = \max(x_{ij} , y_{ij})$ Commonly used in environmental applications.

Table S4. C1 for each species.

	C1
SO ₄ ²⁻	2.00E-02
NO ₃ ⁻	2.00E-02
Cl ⁻	2.00E-02
NH ₄ ⁺	2.00E-02
EC	2.00E-02
OC	2.00E-02
Al	2.00E-05
As	2.00E-06
Ba	2.00E-04
Br	2.00E-06
Ca	2.00E-05
Cu	2.00E-06
Fe	2.00E-05
K	2.00E-05
Mn	2.00E-06
Pb	2.00E-05
Db	2.00E-05
Se	2.00E-06
Si	2.00E-04
Sn	2.00E-05
Ti	2.00E-05
Zn	2.00E-06

Table S5. C3 introduced to different computations for four scenarios.

		Errors									
		0.025	0.05	0.1	0.15	0.2	0.25	0.3	0.5	0.7	0.9
C3 in SCE 1		0.025 for all elements									
		0.05 for all elements									
		0.1 for all elements									
		0.15 for all elements									
		0.2 for all elements									
		0.25 for all elements									
		0.3 for all elements									
		0.5 for all elements									
		0.7 for all elements									
		0.9 for all elements									
C3 in SCE 2		0.0125 for EC, OC, Si and NO ₃ ⁻ and 0.025 for others									
		0.025 for EC, OC, Si and NO ₃ ⁻ and 0.05 for others									
		0.05 for EC, OC, Si and NO ₃ ⁻ and 0.1 for others									
		0.075 for EC, OC, Si and NO ₃ ⁻ and 0.15 for others									
		0.1 for EC, OC, Si and NO ₃ ⁻ and 0.2 for others									
		0.125 for EC, OC, Si and NO ₃ ⁻ and 0.25 for others									
		0.15 for EC, OC, Si and NO ₃ ⁻ and 0.3 for others									
		0.25 for EC, OC, Si and NO ₃ ⁻ and 0.5 for others									
C3 in SCE 3		0.0125 for NO ₃ ⁻ and 0.025 for others									
		0.025 for NO ₃ ⁻ and 0.05 for others									
		0.05 for NO ₃ ⁻ and 0.1 for others									
		0.075 for NO ₃ ⁻ and 0.15 for others									
		0.1 for NO ₃ ⁻ and 0.2 for others									
		0.125 for NO ₃ ⁻ and 0.25 for others									
		0.15 for NO ₃ ⁻ and 0.3 for others									
		0.25 for NO ₃ ⁻ and 0.5 for others									
C3 in SCE 4		0.0125 for Fe and 0.025 for others									
		0.025 for Fe and 0.05 for others									
		0.05 for Fe and 0.1 for others									
		0.075 for Fe and 0.15 for others									
		0.1 for Fe and 0.2 for others									

0.125 for Fe and 0.25 for others

0.15 for Fe and 0.3 for others

0.25 for Fe and 0.5 for others

0.35 for Fe and 0.7 for others

0.45 for Fe and 0.9 for others

Table S6. The DP values (a percentage defined as the number of divided tests to that of total tests for each case) in SCE 1 and SCE 2, to investigate the satiability of PMF results.

		C3									
		0.025	0.05	0.1	0.15	0.2	0.25	0.3	0.5	0.7	0.9
Hx for SCE 1	0.025	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	0.05	100%	100%	100%	100%	100%	100%	100%	100%	90%	90%
	0.1	100%	100%	100%	100%	100%	100%	100%	100%	90%	100%
	0.15	80%	100%	100%	100%	100%	100%	100%	100%	80%	80%
	0.2	50%	70%	50%	70%	60%	60%	50%	70%	50%	50%
	0.25	20%	50%	50%	50%	50%	40%	60%	20%	30%	20%
	0.3	30%	30%	50%	50%	40%	30%	60%	40%	10%	20%
	0.5	40%	20%	40%	30%	10%	20%	10%	20%	10%	30%
	0.7	40%	40%	0%	0%	0%	0%	0%	40%	20%	30%
	0.9	50%	20%	20%	20%	40%	40%	50%	60%	60%	40%
Hx for SCE 2	0.025	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	0.05	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	0.1	100%	100%	100%	100%	100%	100%	100%	100%	100%	90%
	0.15	90%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	0.2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	0.25	100%	100%	100%	90%	80%	80%	90%	90%	90%	90%
	0.3	80%	90%	90%	70%	70%	60%	60%	40%	70%	80%
	0.5	30%	40%	40%	70%	40%	40%	40%	20%	30%	40%
	0.7	70%	50%	60%	40%	30%	20%	30%	50%	40%	30%
	0.9	90%	40%	60%	60%	50%	60%	70%	50%	30%	40%

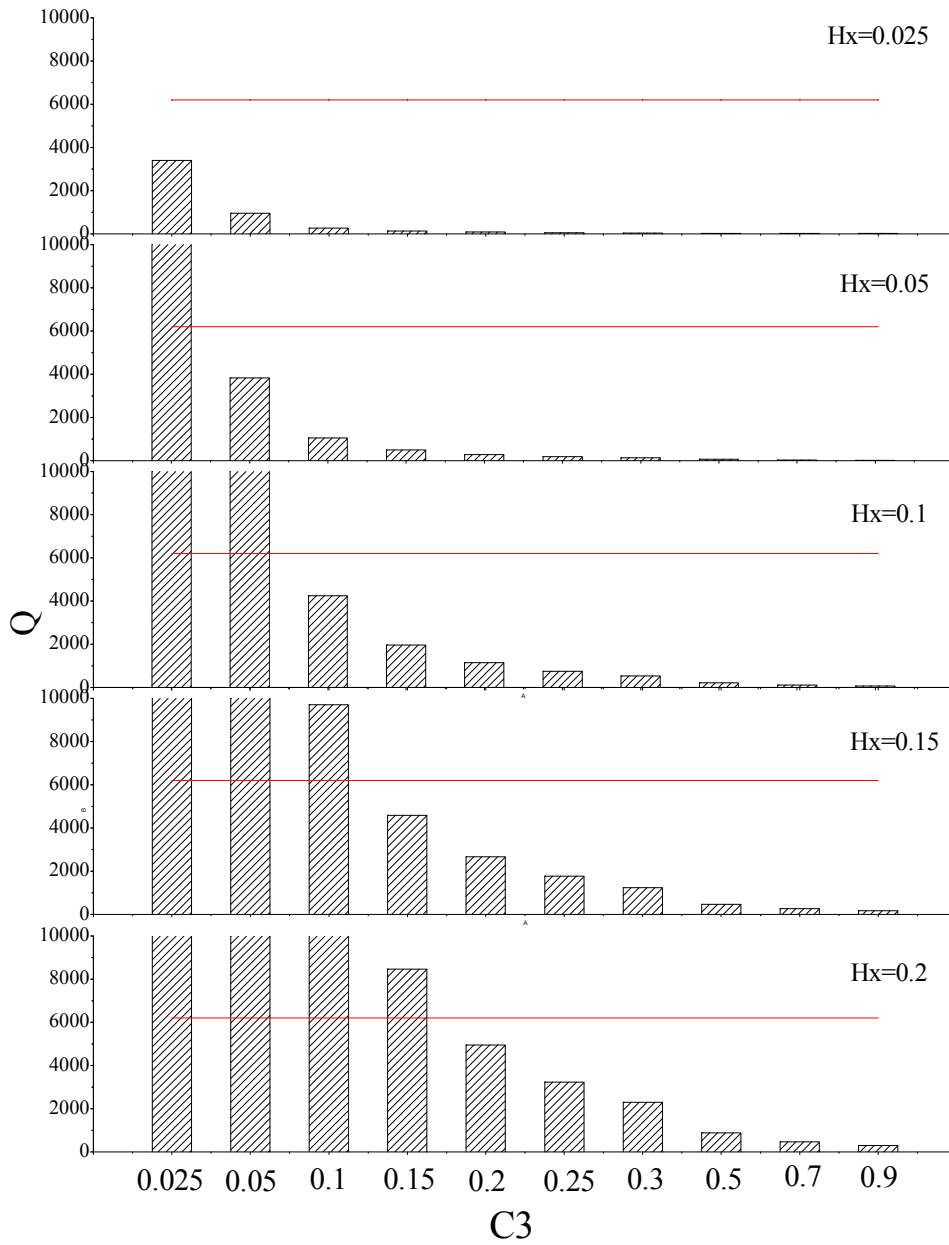


Fig.S1. The averaged Q for each Hx and each C3 as well as the theoretical Q (equal to 6204) in SCE1.

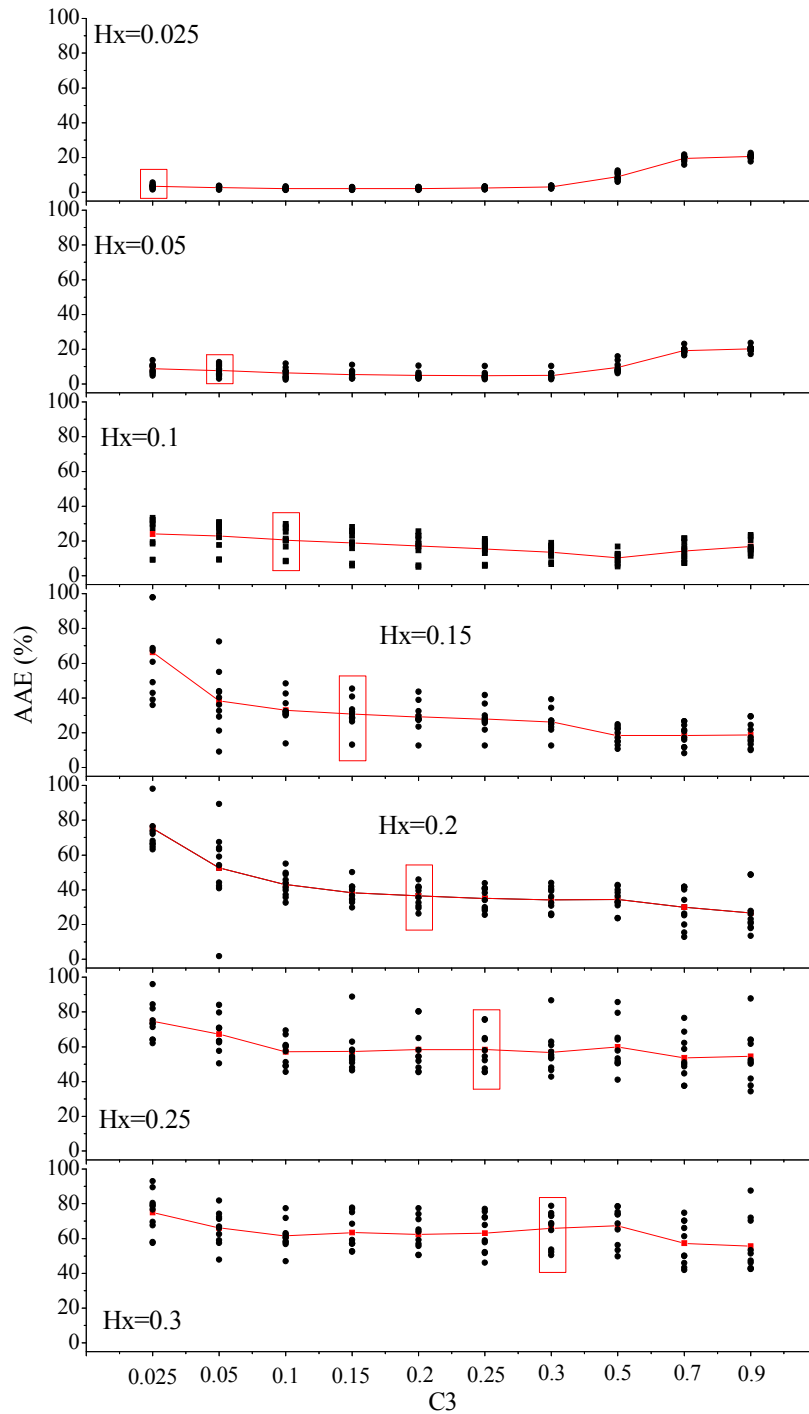


Fig. S2. AAEs of ten tests and their average values corresponding to diverse uncertainty levels (C_3) for various errors ($H_x=0.025, 0.05, 0.1, 0.15, 0.2, 0.25$ and 0.3) in SCE 2 (where C_3 for all the correct tracers were decreased). The cases with the best Q (Q were the most approximate to theoretical Q) were circumscribed for each H_x .

According to Figure S2, similar conclusion as described in SCE 1 can be observed: both of the errors of the datasets and the uncertainties inputted to PMF can influence the performance of PMF; and when errors were small, too high uncertainties would result in contributions deviating from true values, while when errors were high but in a certain degree, the PMF might obtain relatively more satisfactory results with higher uncertainties.

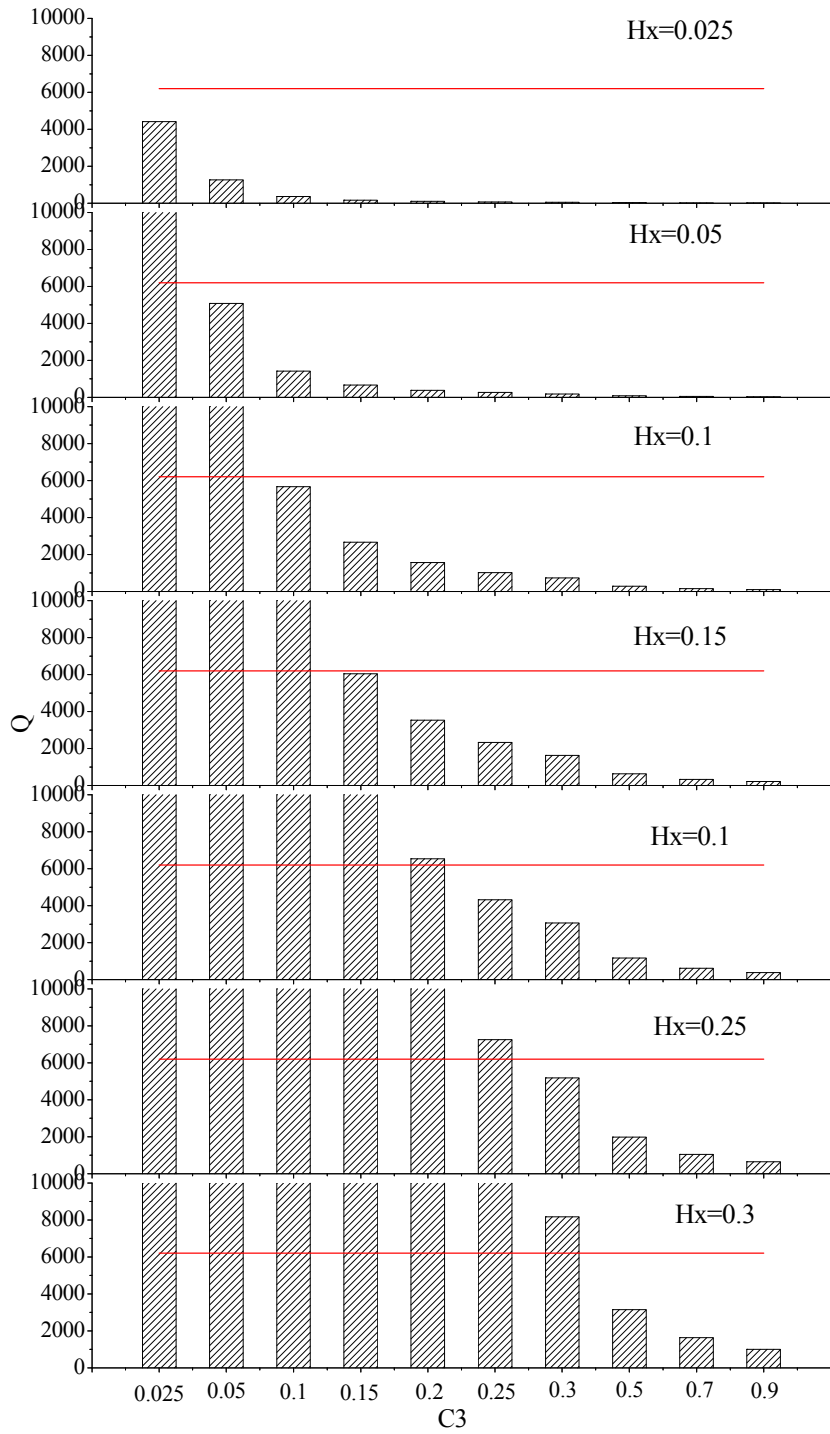


Fig. S3. The averaged Q for each Hx and each C3 as well as the theoretical Q (equal to 6204) in SCE2.

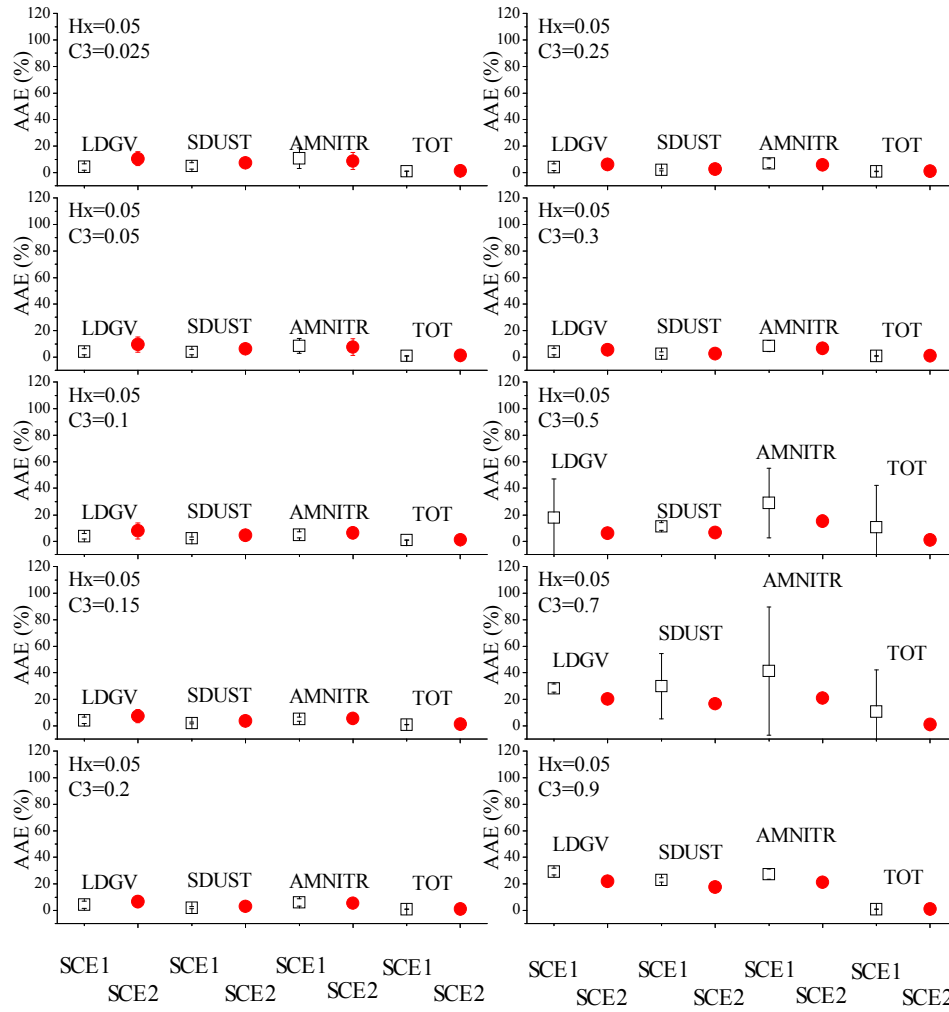


Fig. S4(a). Comparison between SCE 1 (all the species had the same C3) and SCE 2 (C3 of all the correct tracers were decreased). Averaged AAEs along with standard deviation under the same errors ($H_x=0.05$) and C3 were compared. The C3 of species except for the tracers were considered as the C3 level in SCE2.

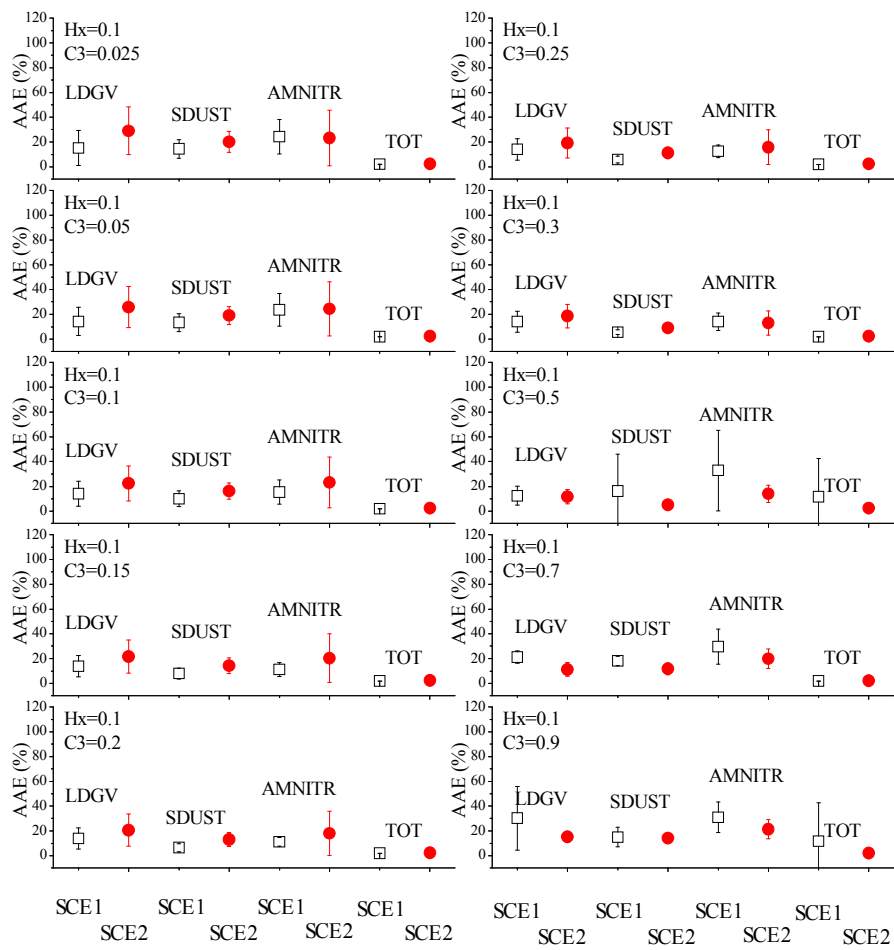


Fig. S4(b). Comparison between SCE 1 (all the species had the same C3) and SCE 2 (C3 of all the correct tracers were decreased). Averaged AAEs along with standard deviation under the same errors (Hx=0.1) and C3 were compared. The C3 of species except for the tracers were considered as the C3 level in SCE2.

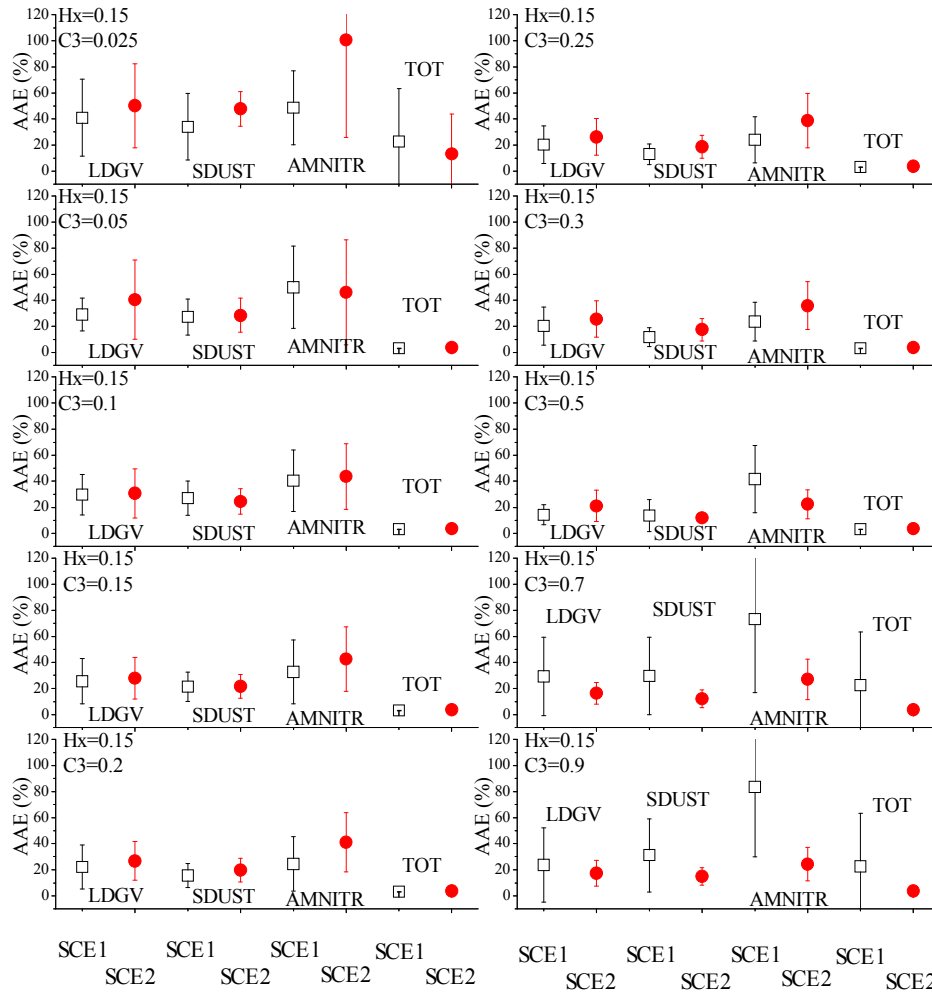


Fig. S4(c). Comparison between SCE 1 (all the species had the same C3) and SCE 2 (C3 of all the correct tracers were decreased). Averaged AAEs along with standard deviation under the same errors ($H_x=0.15$) and C3 were compared. The C3 of species except for the tracers were considered as the C3 level in SCE2.

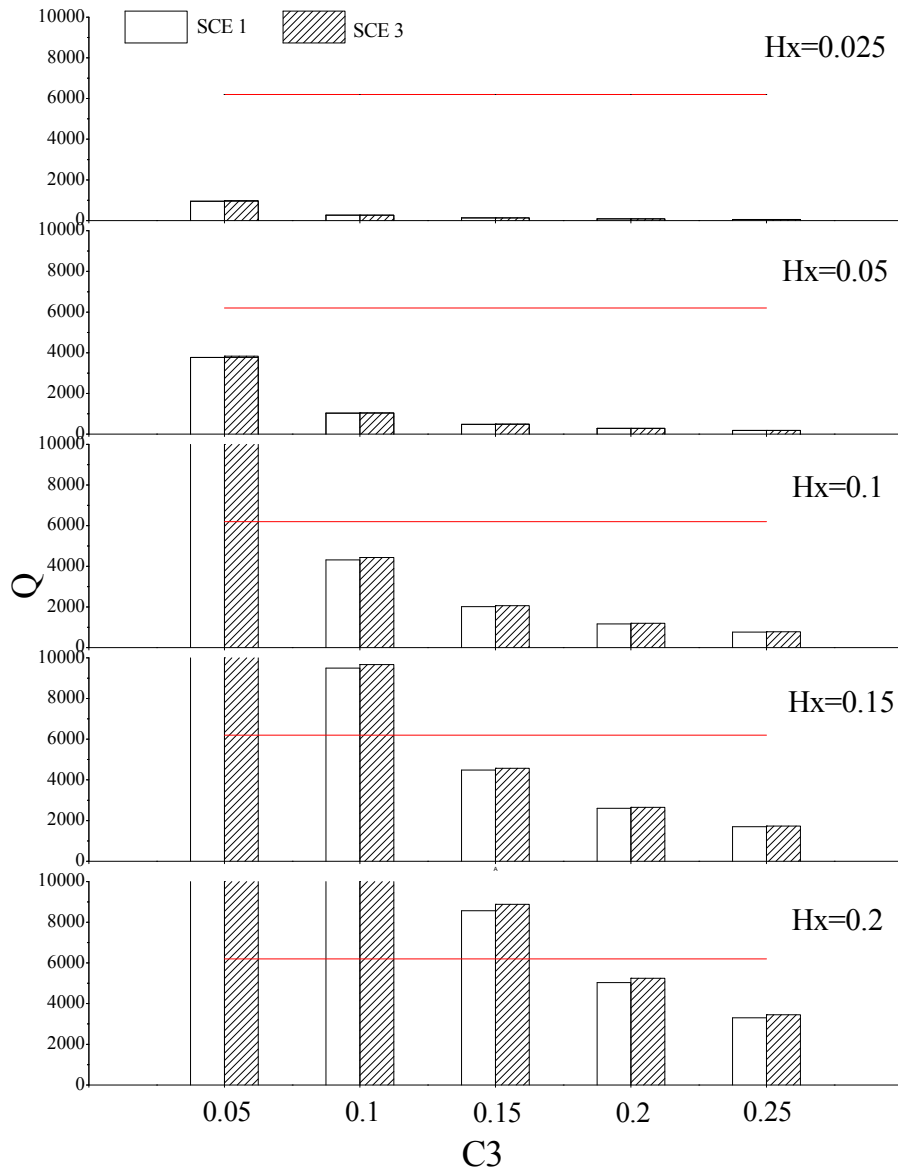


Fig. S6. Comparison between the averaged Q in SCE 1 and in SCE 3.

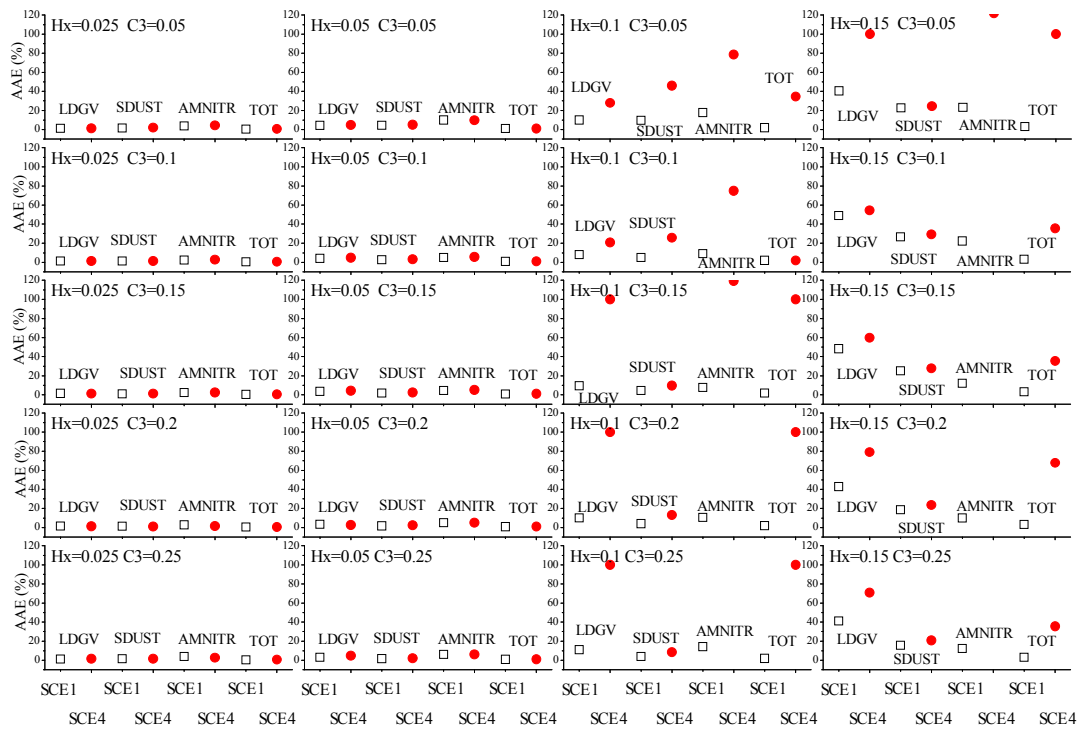


Fig. S7. Comparison between SCE 1 (all the species had the same C3) and SCE 4 (C3 of Fe were decreased). Averaged AAEs along with standard deviation under the same errors ($H_x=0.025-0.2$) and uncertainties ($C_3=0.05-0.25$) were compared. The C3 of species except for Fe were considered as the C3 level in SCE 4.

As shown in Figure S7, although unobvious difference can be observed in SCE 4 and SCE1 when errors were very low ($H_x=0.025$ and 0.05), the estimated contributions of three sources divaricated from the true values even when H_x were still low (like $H_x=0.1$)

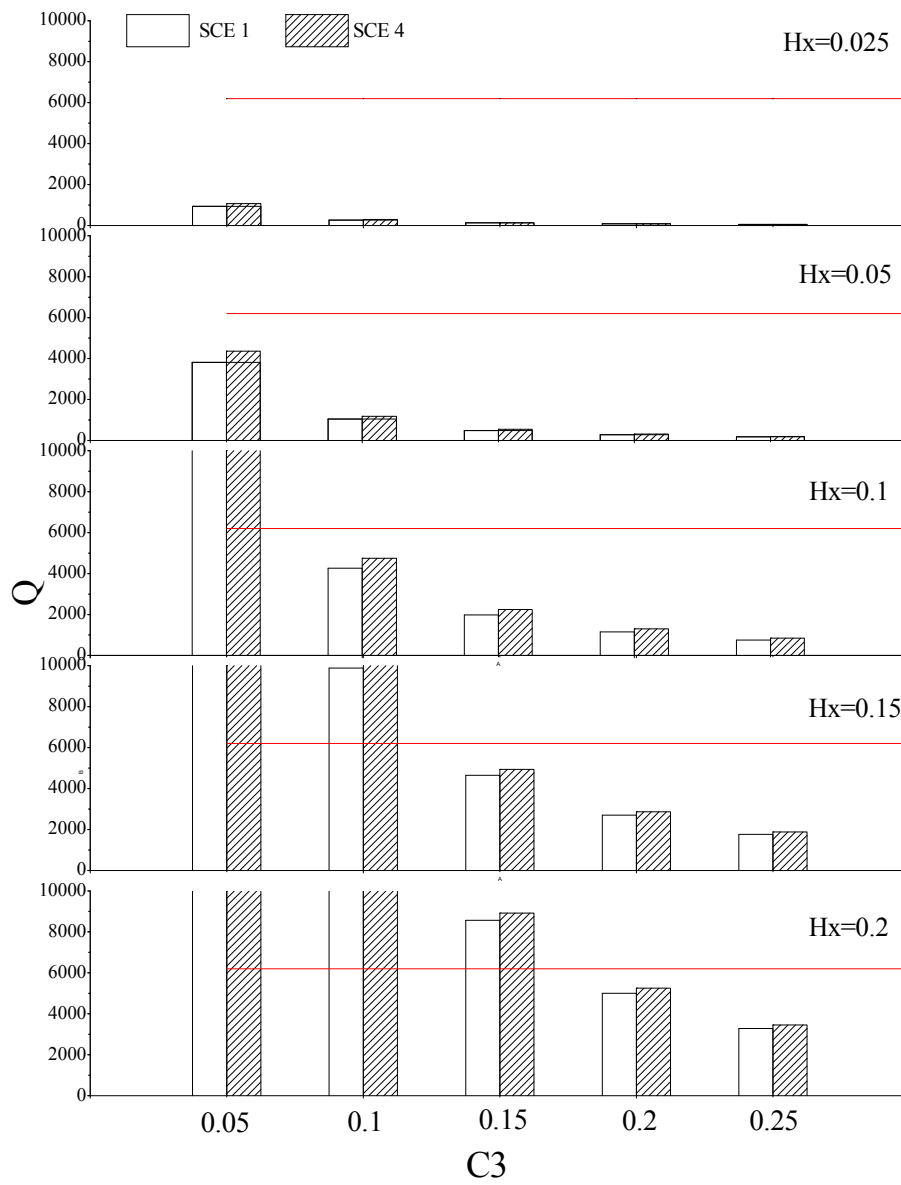


Fig. S8. Comparison between the averaged Q in SCE 1 and in SCE 4.

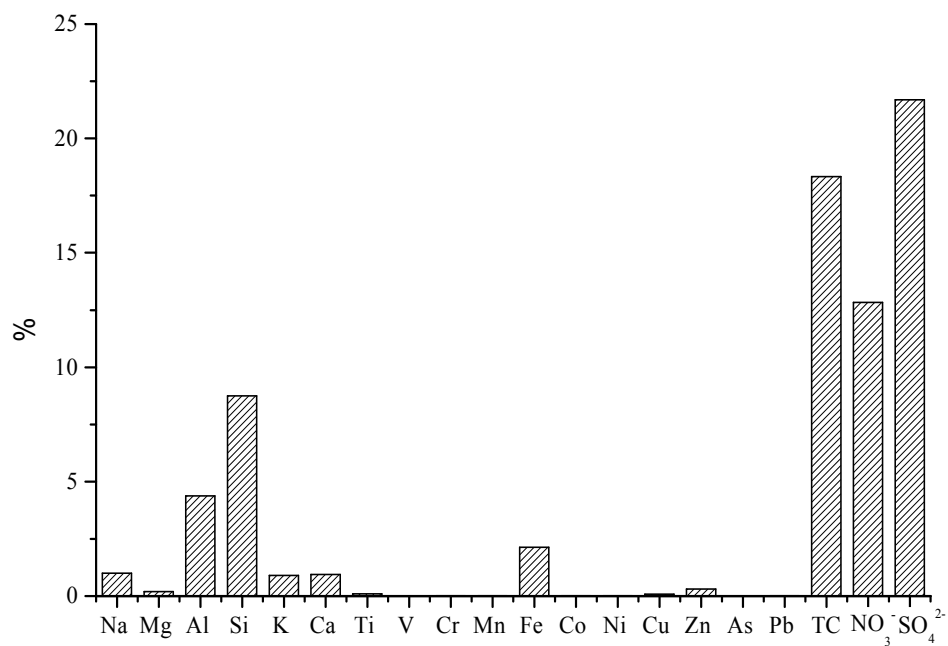


Fig. S9. Chemical compositions of PM_{2.5} in Chengdu.

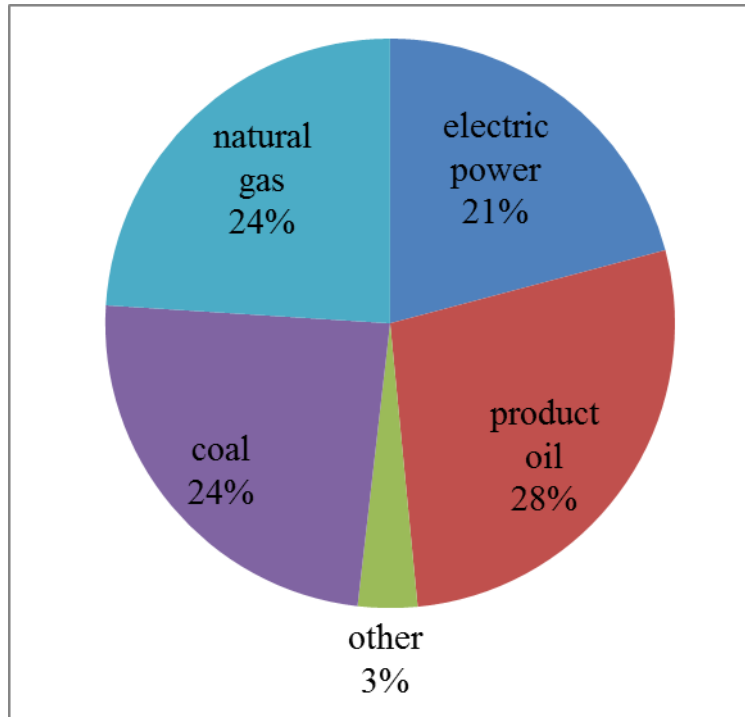


Fig. S10 Energy structure of Chengdu in 2013.

Table S7. C1 of each species for ambient PM_{2.5} dataset from Chengdu.

Species	C1
Na	0.00002
Mg	0.000002
Al	0.00002
Si	0.00002
K	0.00002
Ca	0.00002
Ti	0.000002
V	0.000002
Cr	0.000002
Mn	0.00002
Fe	0.00002
Co	0.000002
Ni	0.000002
Cu	0.000002
Zn	0.00002
As	0.000002
Pb	0.000002
TC	0.002
NO ₃ ⁻	0.002
SO ₄ ²⁻	0.002

Table S8. Estimated source profiles ($\mu\text{g}/\text{m}^3$), percentage source contributions (%) and Q values by PMF in three Tests for the ambient $\text{PM}_{2.5}$ dataset from the megacity in China.

(A)	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Na	0.00	1.13	0.00	0.00	0.00
Mg	0.19	0.00	0.00	0.00	0.00
Al	0.00	5.43	0.00	0.00	0.00
Si	0.00	10.50	0.00	0.00	0.00
K	0.00	0.26	0.65	0.00	0.00
Ca	0.76	0.00	0.00	0.00	0.00
Ti	0.00	0.11	0.00	0.00	0.00
V	0.00	0.00	0.00	0.00	0.00
Cr	0.00	0.00	0.00	0.00	0.00
Mn	0.00	0.00	0.00	0.00	0.00
Fe	0.00	2.40	0.00	0.00	0.00
Co	0.00	0.00	0.00	0.00	0.00
Ni	0.00	0.00	0.00	0.00	0.00
Cu	0.00	0.00	0.04	0.00	0.00
Zn	0.00	0.00	0.00	0.00	0.30
As	0.00	0.00	0.00	0.00	0.00
Pb	0.00	0.00	0.01	0.00	0.00
TC	0.00	22.94	0.00	0.00	0.00
NO_3^-	0.00	0.00	13.18	0.00	0.00
SO_4^{2-}	0.00	0.00	0.00	29.94	0.00
Percentage contributions (%)	0.51%	82.43%	0.07%	6.92%	6.36%
Q=	1724.73				
(B)	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Na	0.99	0.00	0.00	0.14	0.00
Mg	0.00	0.11	0.02	0.00	0.00
Al	3.76	0.13	0.61	1.06	0.00
Si	8.13	0.68	0.00	2.06	0.00
K	0.18	0.07	0.63	0.00	0.00
Ca	0.05	0.43	0.00	0.02	0.00
Ti	0.07	0.00	0.03	0.01	0.00
V	0.00	0.00	0.00	0.00	0.00
Cr	0.00	0.00	0.00	0.00	0.00
Mn	0.00	0.00	0.00	0.00	0.00
Fe	1.97	0.33	0.37	0.00	0.00
Co	0.00	0.00	0.00	0.00	0.00
Ni	0.00	0.00	0.00	0.00	0.00
Cu	0.00	0.00	0.02	0.00	0.00
Zn	0.02	0.25	0.00	0.00	0.00
As	0.00	0.00	0.00	0.00	0.00
Pb	0.00	0.02	0.00	0.00	0.00
TC	14.69	0.00	3.40	0.00	5.91
NO_3^-	0.15	0.00	15.94	0.12	0.00
SO_4^{2-}	0.01	0.00	0.01	33.41	0.00
Percentage contributions (%)	54.21%	2.35%	11.64%	18.29%	11.48%
Q=	2170.96				

(C)	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Na	1.12	0.00	0.00	0.00	0.00
Mg	0.00	0.16	0.00	0.01	0.00
Al	5.21	0.00	0.00	0.07	0.00
Si	10.21	0.00	0.02	0.00	0.02
K	0.21	0.00	0.02	0.71	0.02
Ca	0.00	0.73	0.00	0.00	0.00
Ti	0.10	0.00	0.00	0.02	0.00
V	0.00	0.00	0.00	0.00	0.00
Cr	0.00	0.00	0.00	0.00	0.00
Mn	0.00	0.00	0.00	0.00	0.00
Fe	0.00	0.00	2.81	0.00	0.00
Co	0.00	0.00	0.00	0.00	0.00
Ni	0.00	0.00	0.00	0.00	0.00
Cu	0.00	0.00	0.00	0.04	0.00
Zn	0.00	0.00	0.00	0.11	0.17
As	0.00	0.00	0.00	0.00	0.00
Pb	0.00	0.00	0.00	0.00	0.02
TC	20.69	0.02	0.02	1.66	0.01
NO ₃ ⁻	0.16	0.00	0.02	14.09	0.00
SO ₄ ²⁻	0.00	13.71	0.00	6.76	0.01
Percentage contributions (%)	40.60%	4.70%	45.02%	2.83%	2.31%
Q=			1861.96		