Supporting Information

Source apportionment of ambient fine particle from combined size distribution and chemical composition data during summertime in Beijing

Z. R. Liu, Y. S. Wang, Q. Liu, B. Hu, and Y. Sun

State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry (LAPC), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

Correspondence to: Y.S. Wang (wys@dq.cern.ac.cn)

Meteorological Data

The temperature ranged from 19.8 to 38.5 ℃ with a median value of 27.2 ℃. The interquartile range of barometric pressure was from 998 to 1005 mb, and that of wind speed was from 0.3 to 1.4m/s, indicating a static stability and quite variety of dry and damp weather as reflected by the relative humidity which ranged from 53 to 77% across the interquartile range.

Details of PMF Method

PMF (positive matrix factorization) was developed by Paatero (Paatero and Tapper 1994; Paatero, 1997) due to problems inherent to eigenvector based methods. PMF solves the general receptor modeling problem using constrained, weighted, least-squares (Reff et al., 2007) for a review of PMF methods. Receptor models determine the factors that are responsible for the data measured at the sampling site (in this case the IAP monitoring site). The general model assumes there are p sources, source types or source regions (termed factors) impacting a receptor (in this case the IAP monitoring site), and linear combinations of the impacts from the p factors to the observed concentrations of the various species or in this case size bins for the SMPS+APS spectra plus auxiliary measurements (Harrison et al., 2011). The model equation is:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \qquad (1)$$

Where x_{ij} is the measured concentration of the *j*th species in the ith sample, f_{kj} is the concentration of the *j*th species in material emitted by source k, g_{ik} is the contribution of the *k*th source to the *i*th sample, and e_{ij} is portion of the measurements that cannot be fitted by the model (residuals). For this analysis, “species” refers to either the chemical species or the size bin. In PMF, only x_{ij} are known and the goal is to estimate the contributions (g_{ik}) and the fractions (f_{kj}). It is assumed that the contributions and number fractions are all non-negative, hence the “constrained” part of the least-squares. Furthermore, PMF uses uncertainties measured for each of the x_{ij}. The task of PMF is then to minimize the residual sum of squares (Q) defined by equation (2):

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{e_{ij}}{s_{ij}}\right)^2 \qquad (2)$$

The uncertainties were computed from the measurement errors by equation (3):

$$s_{ij} = \sigma_{ij} + C_3 \max(|x_{ij}|, |y_{ij}|) \qquad (3)$$

Where y_{ij} is the calculated value for x_{ij}, σ_{ij} is the measurement or estimated error, and C_3 is a dimensionless constant value, 0.08 in this study. The estimation of the measurement errors of size distribution data were based on the combination of size bins and the detailed procedure was provided in the work of Zhou et al. (2004b). C_3 is used as the estimation of the relative uncertainties of large values (Norris, et al., 2008).The choice of the number of factors is a compromise according to Lee et al. (1999). Using too few factors will combine sources of different nature together and using too many factors will make a real factor further dissociate into two or more non-existing factors. F_{peak} is a parameter in PMF for controlling rotations (Paatero et al., 2002). When the F_{peak} value is positive, the following additional term is included in the object function Q:

$$Q^P = \beta^2 (\sum_{K=1}^{P} \sum_{J=1}^{N} f_{kj})^2 \qquad (4)$$

where β^2 corresponds to the F_{peak} value. The term defined above attempts to pull the sum of all the elements of F toward zero and makes the program do elementary transformations for F and G by subtracting the F vectors from each other and adding corresponding G vectors to obtain a more physically realistic solution. Different numbers of factors and F_{peak} values have been explored to obtain the most meaningful results which have been described by Zhou et al (2004a). For this study, eight factors were selected, and the F_{peak} value was set to -0.4 for both the number size distribution and volume size distribution.

References

Harrison, R. M., Beddows, D. C. S., and Dall'Osto, M.: PMF analysis of wide-range particle size spectra collected on a major highway. Environ. Sci. Technol., 45, 5522-5528, 2011.

Lee, E.; Chan, C. K.; and Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmos. Environ., 33, 3201-3212, 1999.

Norris G., Vedantham R., Wade K., Brown S., Prouty J., and Foley C.: EPA Positive Matrix Factorization (PMF)3.0. Fundamentals User Guide US Environmental Protection Agency, Office of Research and Development Washington DC, 20460, USA, 2008.

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

Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, Chemom. Intell. Lab. Syst., 60, 253-264, 2002.

Paatero, P.: Least squares formulation of robust nonnegative factor analysis. Chemom. Intell. Lab.

Syst., 37, 23-35, 1997.

Reff, A., Eberly, S. I., and Bhave, P. V.: Receptor modelling of ambient particle matter data using positive matrix factorization: review of existing methods, *JAWMA.*, 57, 146-154, 2007.

Zhou, L. M.; Eugene, Kim.; Hopke, P. K.; Stanier, C. O.; Pandis, S. Advanced Factor Analysis on Pittsburgh Particle Size-Distribution Data Special Issue of Aerosol Science and Technology on Findings from the Fine Particulate Matter Supersites Program. *Aerosol Sci. Technol.* 38, 118-132, 2004a.

Zhou, L., E. Kim, P. K. Hopke, C. Stanier, and S. N. Pandis, Advanced factor analysis on Pittsburgh particle size distribution data, *Aerosol Sci. Technol.* 38, 118-132, 2004b.

Table Legends

TableS1 Description of the hourly average concentrations of particle size channels, particle composition and gaseous pollutants during the summer intensive observation period in Beijing.

TableS2 Description of the hourly average concentrations of Meteorological data measured during the summer intensive observation period in Beijing.

Figure Legends

Fig. S1 Wind profile during the study period.

Fig. S2 Diurnal pattern of particle chemical composition during the study period

Fig. S3 Diurnal pattern of gaseous pollutants during the study period

Fig. S4 Diurnal pattern of meteorological parameters during the study period

TableS1 Description of the hourly average concentrations of particle size channels, particle composition and gaseous pollutants during the summer intensive observation period in Beijing.

	N	Mean	SD	Min	Max	Median	Lower25%	Upper25%
NC(cm^{-3})(1-10channels)(0.01-0.03 µm)	744	93753	65283	23243	671864	75577	54303	110380
NC(cm^{-3})(11-20channels)(0.03-0.08 µm)	744	110879	56834	18369	357335	95971	72824	136996
NC(cm^{-3})(21-30channels)(0.09-0.24 µm)	744	68708	26815	14712	159020	65834	49097	85390
NC(cm^{-3})(31-40channels)(0.27-0.60 µm)	744	7849	5090	263	32229	6829	4411	10701
NC(cm^{-3})(41-50channels)(0.64-1.22 µm)	744	305	268	8	1548	209	123	438
NC(cm^{-3})(51-60channels)(1.32-2.51 µm)	744	29	16	4	110	27	19	37
Organics(µg/m^3)	744	11.9	5.8	1.2	54.1	11.2	8.3	14.5
Nitrate(µg/m^3)	744	5.7	4.3	0.1	26.9	4.8	2.4	8.3
Sulfate(µg/m^3)	744	11.4	6.6	0.6	40.2	11.3	6.2	15.2
Ammonium(µg/m^3)	744	6.3	3.6	0.2	20.3	6.2	3.5	8.3
Chlorine(µg/m^3)	744	0.5	0.4	0.1	4.2	0.4	0.2	0.7
NO(µg/m^3)	744	5.8	8.6	0.3	79.1	2.4	1.3	6.5
NO_2(µg/m^3)	744	44.4	13.5	10.9	97.7	43.1	34.5	52.1
O_3(µg/m^3)	744	59.7	51.1	0.4	249.3	43.2	17.3	92.9
SO_2(µg/m^3)	744	13.0	7.0	0.1	41.2	11.9	8.4	16.1
CO(mg/m^3)	744	1.3	0.6	0.2	3.7	1.2	0.8	1.6

TableS2 Description of the hourly average concentrations of meteorological data measured during the summer intensive observation period in Beijing.

	N	Mean	SD	Min	Max	Median	Lower25%	Upper25%
Temperature (℃)	744	27.2	3.6	19.8	38.5	26.9	24.4	29.8
Relative Humidity (%)	744	65.3	16.8	16.0	100.0	68.0	53.0	77.0
Pressure (mBar)	744	1001.0	4.2	992.2	1011.1	1000.8	997.5	1004.6
Wind Speed (m/s)	744	0.9	0.7	0.1	4.4	0.8	0.3	1.4

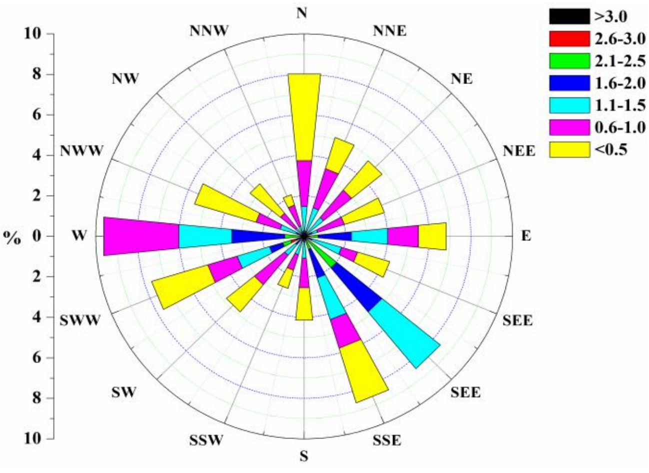

Fig. S1 Wind profile during the study period.

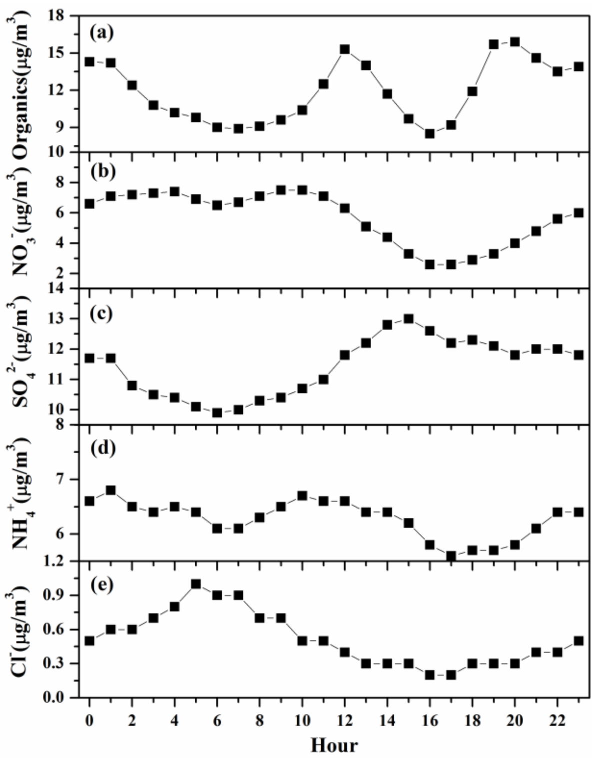

Fig. S2 Diurnal pattern of particle chemical composition during the study period

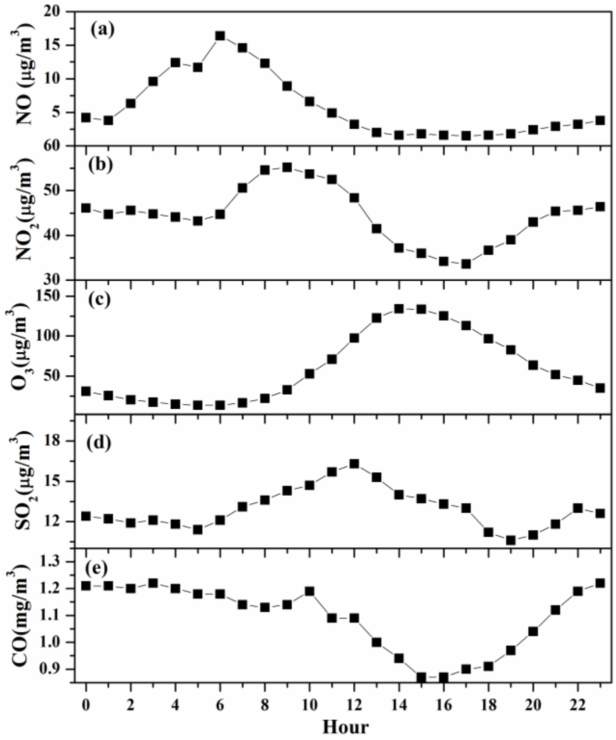

Fig. S3 Diurnal pattern of gaseous pollutants during the study period

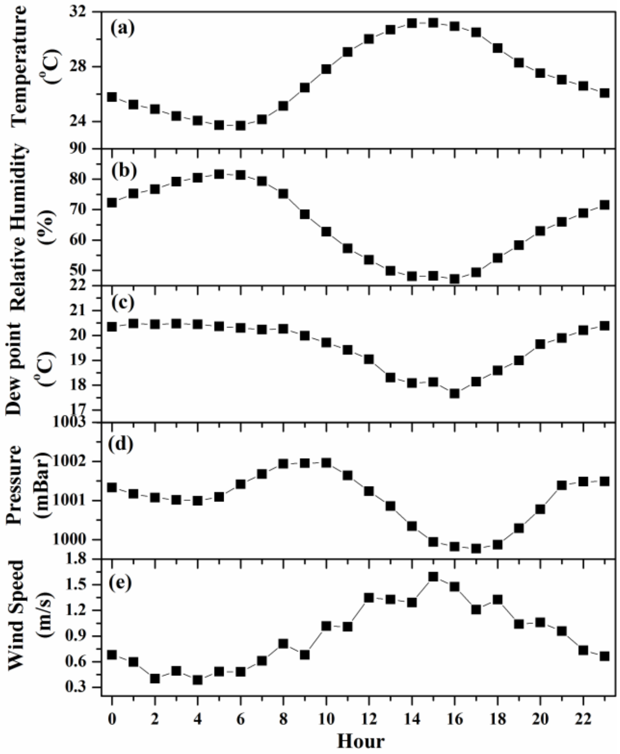

Fig. S4 Diurnal pattern of meteorological parameters during the study period