Supplementary Material for *Online Measurement of Gas-Phase Nitrated Phenols Utilizing CI-LToF-MS: Primary Sources and Secondary Formation*

Kai Song[1,2], Song Guo[1,2*], Haichao Wang[3], Ying Yu[1], Hui Wang[1], Rongzhi Tang[1], Shiyong Xia[4], Yuanzheng Gong[1], Zichao Wan[1], Daqi Lv[1], Rui Tan[1], Wenfei Zhu[1], Ruizhe Shen[1], Xin Li[1], Xuena Yu[1], Shiyi Chen[1], Liming Zeng[1], Xiaofeng Huang[4]

[1] State Key Joint Laboratory of Environmental Simulation and Pollution Control, International Joint Laboratory for Regional Pollution Control, Ministry of Education (IJRC), College of Environmental Sciences and Engineering, *Beijing*, 100871, China

[2] Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, *Nanjing* 210044, China P. R.

[3] School of Atmospheric Sciences, Sun Yat-sen University, *Zhuhai*, 519082, China.

[4] Key Laboratory for Urban Habitat Environmental Science and Technology, School of Environment and Energy, Peking University Shenzhen Graduate School, *Shenzhen*, 518055, China.

* Correspondence to: S. Guo: songguo@pku.edu.cn

Figures

Figure S1. Chemical structures and high-resolution peak fits of reagent ions and nitrated phenols (NPs) investigated in this study.

Figure S2. (a) Background ions and ions detected during the calibration period (local time); (b) Calibration line of ions (y) and the standard gas-phase concentration of nitrophenol (x).

Figure S3. The measured concentration of nitrated phenols and their secondary formation simulation by the box model.

Figure S4. Air quality and meteorology conditions during the sampling period in Beijing: time series of (a) wind speed, (b) RH, (c) $PM_{2.5}$, (d) NO_y and (e) CO from Dec 1 to Dec 31, 2018.

Figure S5. Consensus maps of brunet, KL, offset, lee, nsNMF and snmf/l algorithms in NMF. The consensus approach was used to estimate the proper method and cluster method of simulation. The color of the consensus map indicated the coefficient and an ideal consensus map was a color-coded heat map in which red blocks along the diagonal on a blue background (Monti et al., 2003; Simpson et al., 2010). KL approach was the optimal one.

Figure S6. NMF rank survey of factors 3 to 7. The cophenetic coefficient and RSS curves were used for the judgment of factor numbers. The first decreasing cophenetic value (Brunet et al., 2004)and an inflection point of the RSS curve (Hutchins et al., 2008)was the optimal factor number, that was, four factors in this study.

Figure S7. Diurnal profiles of coal combustion (a), biomass burning (b), industry (c) and vehicle exhaust (d) sources. Coal combustion and biomass burning displayed a nighttime peak while the source of vehicle exhaust showed peaks at rush hour which were evidence of the NMF source apportionment.

Figure S8. Source profile from the PMF model. (a) Source profile of PMF results. SO_2, chloromethane, aromatics and 1,3-butadiene as the markers of coal combustion, biomass burning, industry and vehicle exhaust sources. (b) Contribution of primary emission (in blue borderline) and second formation (in red borderline) of NPs.

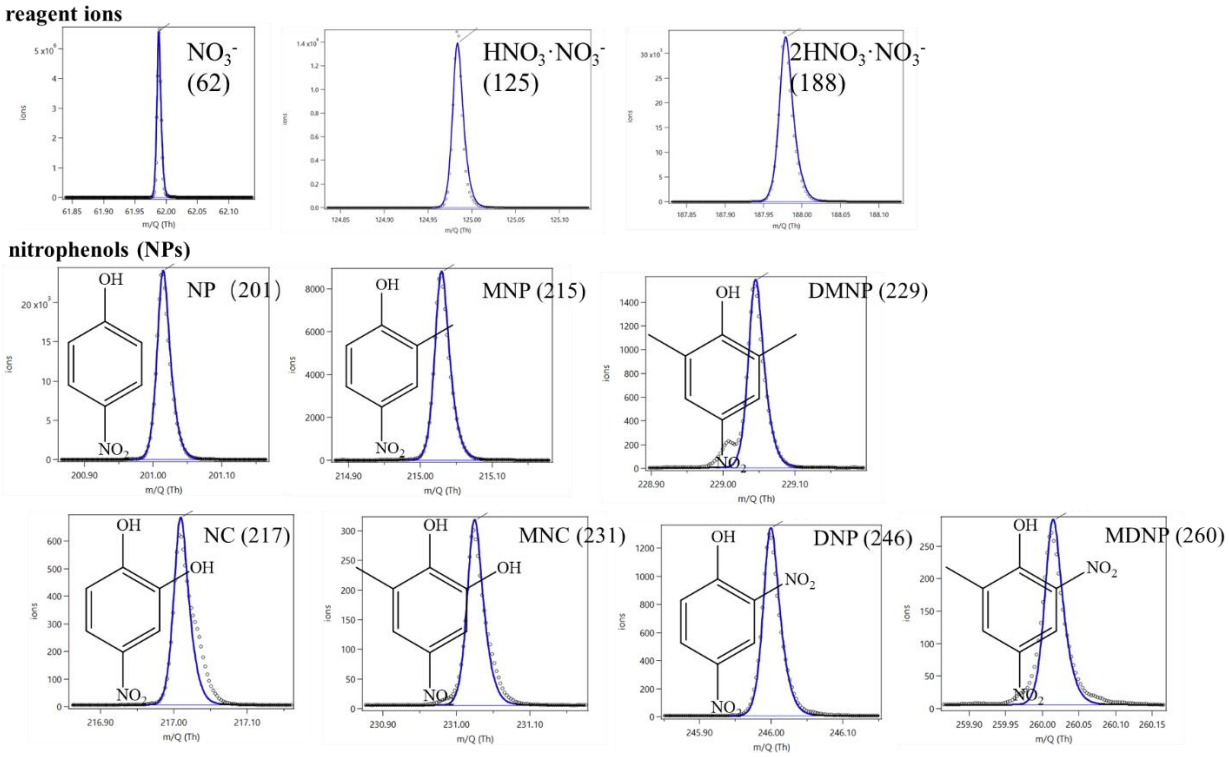

Figure S1. Chemical structures and high-resolution peak fits of reagent ions and nitrated phenols (NPs) investigated in this study.

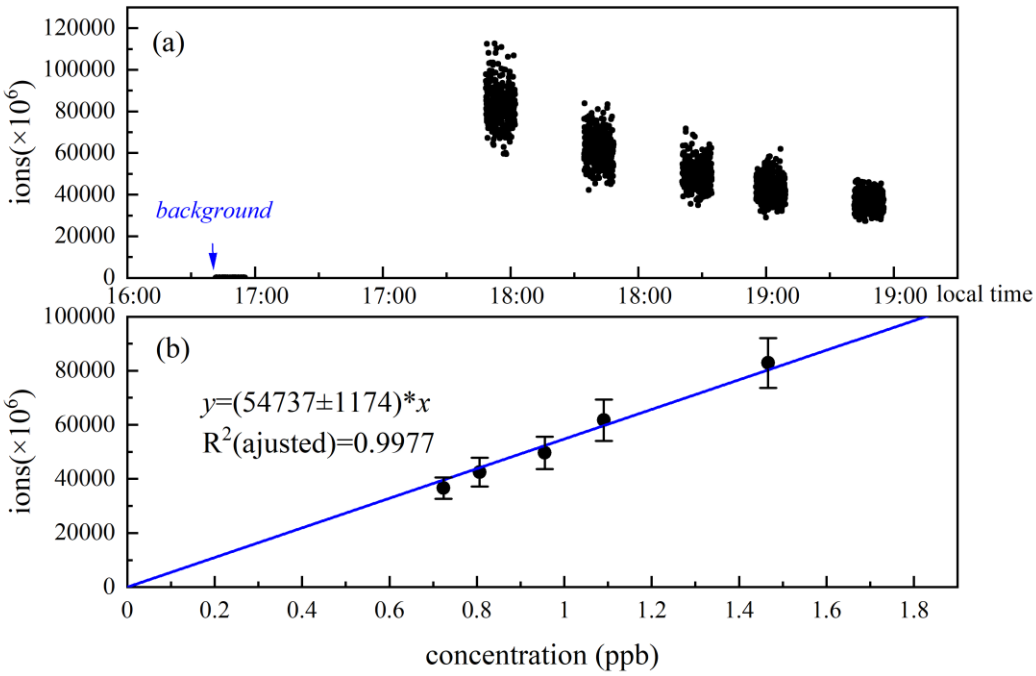

Figure S2. (a) Background ions and ions detected during the calibration period (local time); (b) Calibration line of ions (y) and the standard gas-phase concentration of nitrophenol (x).

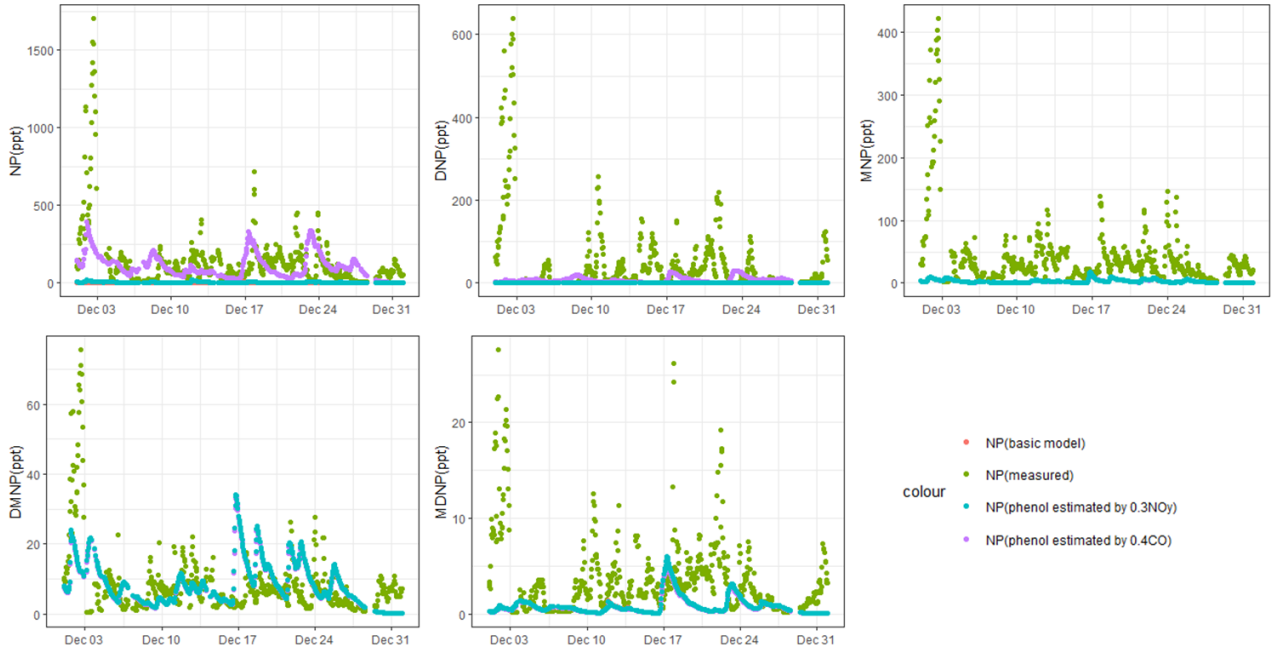

Figure S3. The measured concentration of nitrated phenols and their secondary formation simulation by the box model.

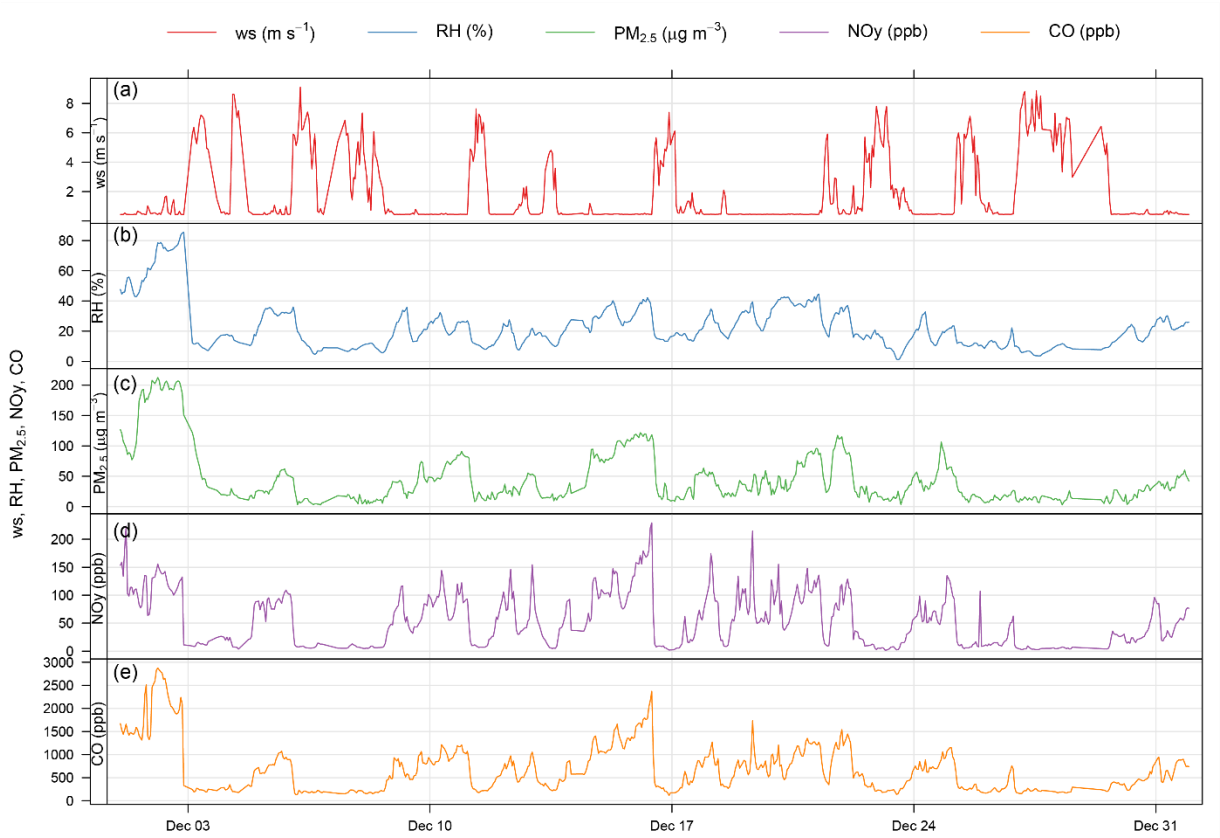

Figure S4. Air quality and meteorology conditions during the sampling period in Beijing: time series of (a) wind speed, (b) RH, (c) $PM_{2.5}$, (d) NO_y and (e) CO from Dec 1 to Dec 31, 2018.

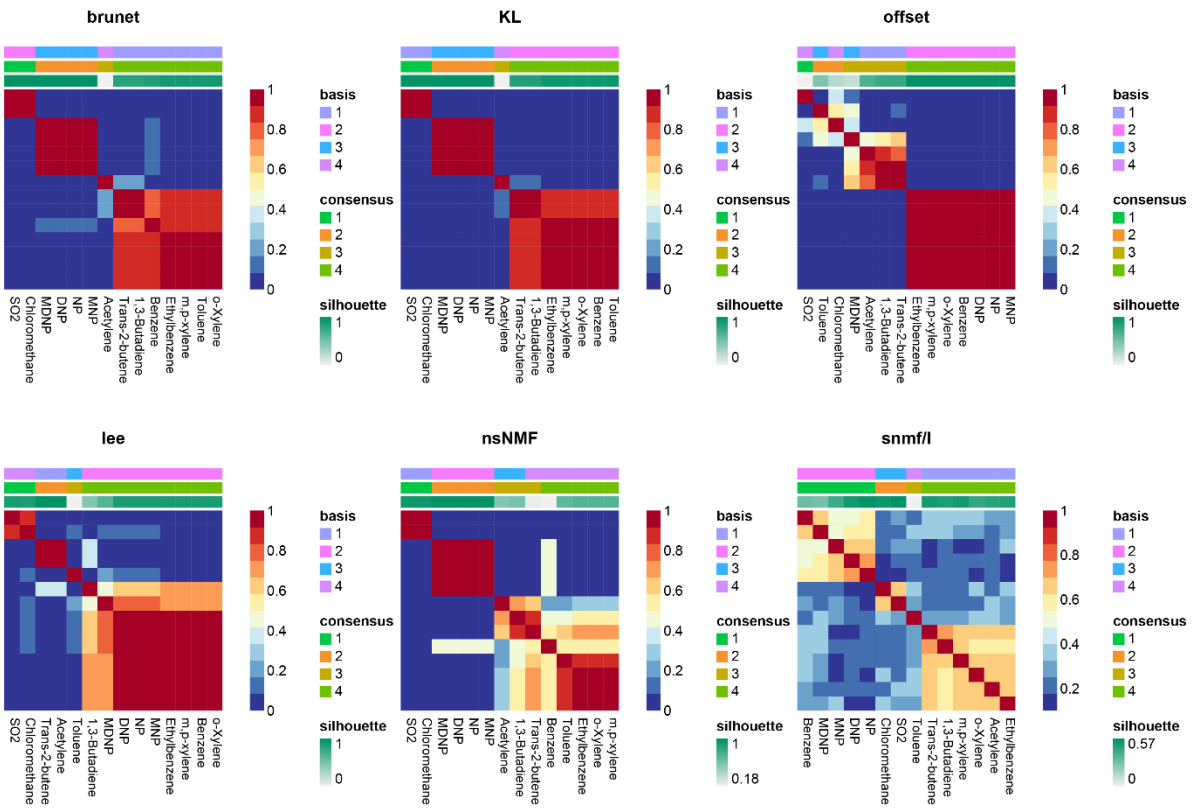

Figure S5. Consensus maps of brunet, KL, offset, lee, nsNMF and snmf/l algorithms in NMF. The consensus approach was used to estimate the proper method and cluster method of simulation. The color of the consensus map indicated the coefficient and an ideal consensus map was a color-coded heat map in which red blocks along the diagonal on a blue background (Monti et al., 2003; Simpson et al., 2010). KL approach was the optimal one.

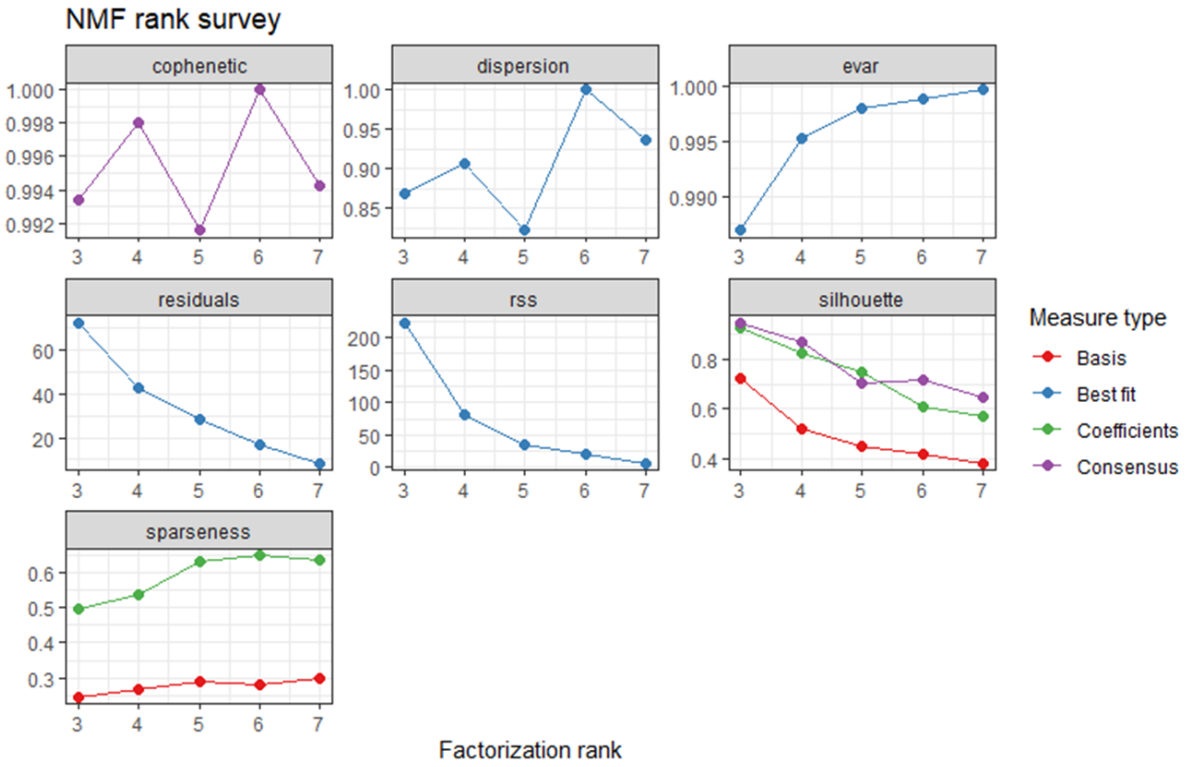

Figure S6. NMF rank survey of factors 3 to 7. The cophenetic coefficient and RSS curves were used for the judgment of factor numbers. The first decreasing cophenetic value (Brunet et al., 2004) and an inflection point of the RSS curve (Hutchins et al., 2008) was the optimal factor number, that was, four factors in this study.

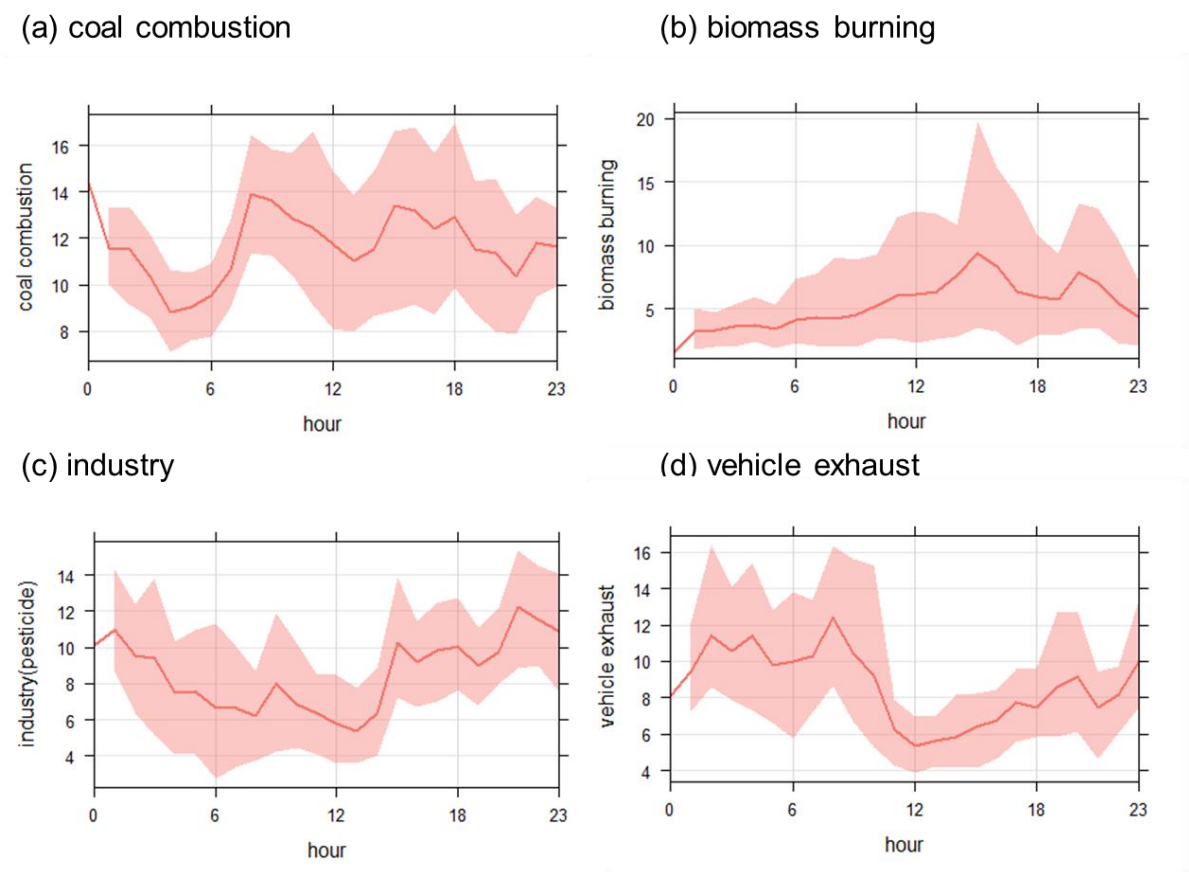

Figure S7. Diurnal profiles of coal combustion (a), biomass burning (b), industry (c) and vehicle exhaust (d) sources. Coal combustion and biomass burning displayed a nighttime peak while the source of vehicle exhaust showed peaks at rush hour which were evidence of the NMF source apportionment.

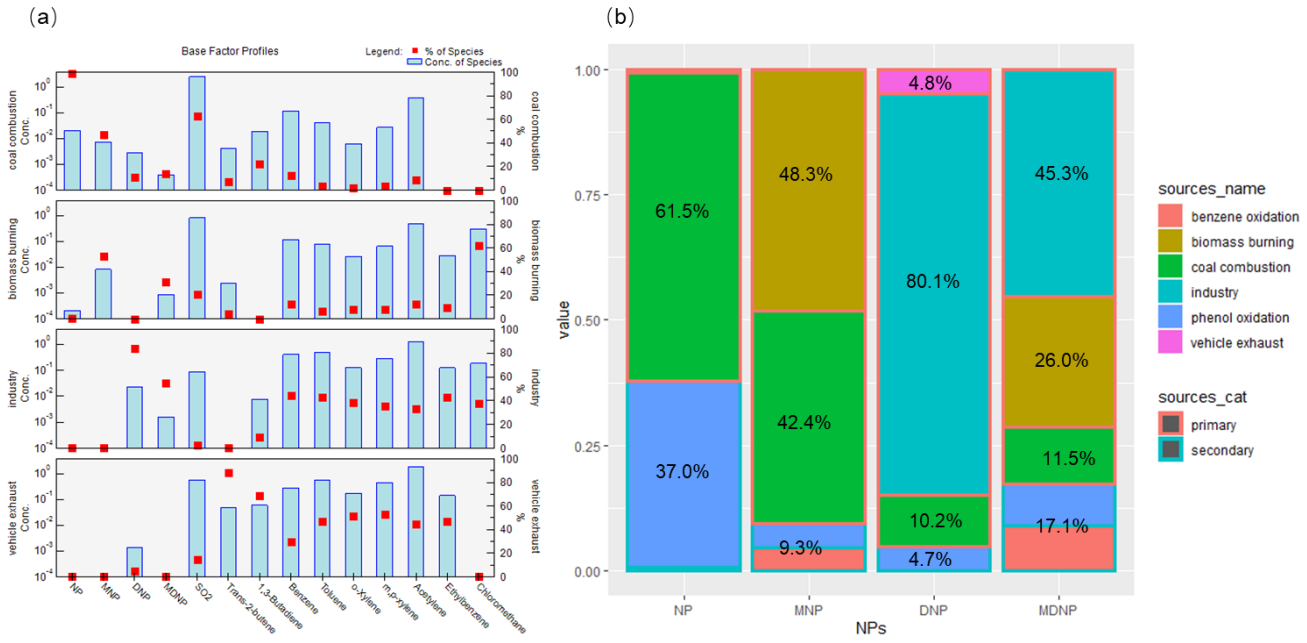

Figure S8. Source profile from the PMF model. (a) Source profile of PMF results. SO_2, chloromethane, aromatics and 1,3-butadiene as the markers of coal combustion, biomass burning, industry and vehicle exhaust sources. (b) Contribution of primary emission (in blue borderline) and second formation (in red borderline) of NPs.

References

Brunet, J. P., Tamayo, P., Golub, T. R. and Mesirov, J. P.: Metagenes and molecular pattern discovery using matrix factorization, Proc. Natl. Acad. Sci. U. S. A., doi:10.1073/pnas.0308531101, 2004.

Hutchins, L. N., Murphy, S. M., Singh, P. and Graber, J. H.: Position-dependent motif characterization using non-negative matrix factorization, Bioinformatics, doi:10.1093/bioinformatics/btn526, 2008.

Monti, S., Tamayo, P., Mesirov, J. and Golub, T.: Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data, Mach. Learn., doi:10.1023/A:1023949509487, 2003.

Simpson, T. I., Armstrong, J. D. and Jarman, A. P.: Merged consensus clustering to assess and improve class discovery with microarray data, BMC Bioinformatics, doi:10.1186/1471-2105-11-590, 2010.