Supplementary Material for

Contributions of biomass-burning, urban, and biogenic emissions to the concentrations and light-absorbing properties of particulate matter in central Amazonia during the dry season

Suzane S. de Sá (1), Luciana V. Rizzo (2), Brett B. Palm[a] (3), Pedro Campuzano-Jost (3), Douglas A. Day (3), Lindsay D. Yee (4), Rebecca Wernis (5), Gabriel Isaacman-VanWertz[b] (4), Joel Brito[c] (6), Samara Carbone[d] (6), Yingjun J. Liu[e] (1), Arthur Sedlacek (7), Stephen Springston (7), Allen H. Goldstein (4), Henrique M. J. Barbosa (6), M. Lizabeth Alexander (8), Paulo Artaxo (6), Jose L. Jimenez (3), Scot T. Martin[*] (1,9)

(1) John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA
(2) Department of Environmental Sciences, Universidade Federal de São Paulo, Diadema, São Paulo, Brazil
(3) Department of Chemistry and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA
(4) Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, California, USA
(5) Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, California, USA
(6) Institute of Physics, University of São Paulo, São Paulo, Brazil
(7) Brookhaven National Laboratory, Upton, New York, USA
(8) Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA
(9) Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA
[a] Now at Department of Atmospheric Sciences, University of Washington, Seattle, USA
[b]Now at Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, Virginia, USA
[c] Now at IMT Lille Douai, Université Lille, SAGE, Lille, France
[d] Now at Agrarian Sciences Institute, Federal University of Uberlândia, Minas Gerais, Brazil
[e] Now at College of Environmental Science and Engineering, Peking University, Beijing, China

[*]To Whom Correspondence Should be Addressed

E-mail: scot_martin@harvard.edu

https://martin.seas.harvard.edu/

Supplementary Text

S1. Aerosol mass spectrometry and positive-matrix factorization

S1.1 AMS collection efficiency

The collection efficiency (CE) of the AMS was calculated as composition dependent (Middlebrook et al., 2012) with a default value of 0.5, which yielded a CE of 0.51 ± 0.02 throughout IOP2. A comparison of particle volume concentrations measured by the AMS with the CE applied and two co-located Scanning Mobility Particle Sizers (SMPS) is shown in Figure S1. The volume of refractory black carbon (rBC) accounted for 2 ± 1 % of the total volume measured by SMPS, for an assumed material density of 1.8 g cm^{-3} for BC. Therefore, the BC concentration was not subtracted from the abscissa of Figure S1.

S1.2 Estimates of organic and inorganic nitrates from AMS data

The total nitrate reported by the AMS includes fragments originating from both organic and inorganic nitrates. In the absence of external measurements of inorganic nitrate, the speciation of nitrate was estimated using the ratio of NO_2^+ to NO^+ signals according to the methods of Farmer et al. (2010); Fry et al. (2013). Results are shown in Figure S2. Calculations were done on a 60-min time base to increase signal over noise, and the resulting inorganic and organic nitrate time series were then interpolated into the original AMS timestamp for ambient measurements. The analysis excluded points that had total nitrate below the estimated detection limit, $DL_{Nitrate}$, which was estimated as three times the standard deviation for "closed AMS spectra", i.e., when the chopper was in closed position and particles did not reach the vaporizer. Mathematically, $DL_{Nitrate} = 3\times\sqrt{E}$, where E is the "closed" error calculated by the standard *PIKA* software (Ulbrich et al., 2009). The resulting mean fraction of organic nitrate in total nitrate was 92% for IOP2 (Figure S2b).

S1.3 Diagnostics of positive-matrix factorization

The time series of organic mass spectra measured by the AMS was analyzed by positive-matrix factorization (PMF) using a standard analysis toolkit (Ulbrich et al., 2009). While the higher-resolution "W"-mode data was used to aid the choice of ions to fit, the higher-sensitivity "V-mode" data were used for quantification of mass concentrations and for the PMF analysis. The first week of collected data (August 18 to August 24, 2014) was excluded from the PMF analysis due to the unusual and overwhelming signal intensities at m/z 44 during that period, which would bias the characterization of the PMF factors as representative of IOP2 as a whole. Technical diagnostics of the six-factor solution are presented in Figures S3 and S4. The analysis was run for a number of factors from 1 to 10, and the rotational ambiguity parameter f_{peak} was varied from -1 to 1 in intervals of 0.2. In Figure S3, panel c shows the quality of fit parameter $Q/Q_{expected}$ (Ulbrich et al., 2009) as a function of the number of factors, suggesting that the solution should have at the very least three factors. Based on panels a and b, there was a large improvement in residuals by inclusion of a fifth factor for IOP2, indicating that the best solution for this dataset should contain at least five factors.

The six-factor solution offered meaningful factors, which showed important correlations with external measurements and allowed for the study of specific sources and processes (Figure 6). Figure S4 corroborates this analysis, showing the factor profiles and loading time series of the 5- and 7-factor solutions. In the 5-factor solution, only one characteristic BBOA factor is resolved. Although this solution was also physically meaningful, the separation of BBOA factors in the 6-factor solution allowed for a more detailed scientific investigation into their sources and properties. In the 7-factor solution, the factors associated with primary sources are further split

(factors 4 through 7) and their interpretation becomes difficult. In conjunction with all the other diagnostics described, these results suggested that the 6-factors solution was the best choice.

Finally, panel d of Figure S3 shows $Q/Q_{expected}$ as a function of the rotational ambiguity parameter f_{peak} (Ulbrich et al., 2009) for the six-factor solution. A plausible range for f_{peak} was determined according to the best practice of limiting $Q/Q_{expected}$ to a value that does not exceed 0.1% of the minimum value (occurring at $f_{peak} = 0$). The default value of $f_{peak} = 0$ was chosen for the final six-factor solution. It yielded the minimum quality of fit parameter $Q/Q_{expected}$, and no significant improvements in the external validation of factors were observed by varying f_{peak}.

S2. Fuzzy c-means clustering

Fuzzy c-means (FCM) clustering was applied to datasets consisting of pollution indicators, namely concentrations of particle number, NO_y, ozone, rBC, CO, and sulfate (Bezdek et al., 1984). The use of a fuzzy clustering method stems from the understanding that any point in time may be affected by a combination of different sources and processes and could therefore be anywhere on the scale between pristine background and extreme polluted conditions, as opposed to a simpler binary classification. Given the scope of the analysis as non-overcast afternoon times, data points were restricted to (i) local 12:00-16:00 h, (ii) local solar radiation over the past 4 h not in the lowest 10 percentile, and (iii) no precipitation over the previous 10 h along backward trajectory. The data were normalized prior to the FCM analysis using the z-score method, which transforms all variables into a common scale with a mean of 0 and a standard deviation of 1.

The FCM algorithm used was the same as in de Sá et al. (2018). It minimizes the objective function represented in Eq. S2-1, which is a weighted sum of squared errors where the error is the Euclidean distance between each data point and a cluster centroid.

$$J(U,v)= \sum_{k=1}^{N}\sum_{i=1}^{c} u_{ik}{}^{m} \, ||y_k - v_i||^2 \qquad \text{(S2-1)}$$

The input data is given by the matrix $Y = [y_1, y_2, \ldots, y_N]$, where y_k is a vector of length X at the *k-th* time point. X is the number of variables (i.e., measurements) used as input in the analysis. The number of time points is represented by N, and the associated running index is k. N was 397 for this study. The number of clusters is represented by c, and the corresponding running index is i. The coordinates of the centroid of each cluster i are represented by v_i, a vector of length X. The exponent of the Fuzzy partition matrix is represented by m. The algorithm returns (1) the Fuzzy partition matrix of Y, given by $U = [u_{ik}]$ where u_{ik} is the degree of membership of time point k to cluster i, (2) the vectors of coordinates of cluster centers, given by $v = [v_i]$, as well as (3) the value J of the objective function.

The analysis was performed in MATLAB® using the "fcm" function in the Fuzzy logic toolbox™. A default value of 2 was used for the exponent m of the partition matrix (Bezdek et al., 1984; Hathaway and Bezdek, 2001; Chatzis, 2011). Further technical details have been described in de Sá et al. (2018). The analysis was run for a number of clusters ranging from two to ten, and the value of the objective function after convergence is shown in Figure S8. The choice of number of clusters hinges on a trade-off between additional information provided by each extra cluster and increased complexity. The objective function largely improved from two to four clusters, with marginal improvements beyond four clusters. The location of cluster centroids was also examined for evaluation of cluster overlap (Figure S9 for IOP2). In this study, three clusters described the system in a meaningful way. The backtrajectories and PM chemical composition typically associated with each of the clusters corroborated the physical interpretation of the 3-cluster solution.

The PM composition associated with each of the clusters was determined by calculating the corresponding coordinates of the centroids for AMS species concentrations and PMF factor loadings, which were not input to the FCM analysis (except for sulfate). The calculation followed the mathematical definition of the centroid (Eq. S2-2). The typical particle optical properties and concentrations of nitrogen-containing families for each cluster in were also determined by the same equation. The resulting characterization of clusters was shown in Figure 10 and Table S1.

$$v_i = \frac{\sum_{k=1}^{N} (u_{ik})^m y_k}{\sum_{k=1}^{N} (u_{ik})^m} \quad \text{(S2-2)}$$

In analogy to the weighted mean of Eq. S2-2, a weighted standard deviation was defined as a measure of cluster variability (Eq. S2-3). All points are considered in the calculation of the standard deviation for a variable in any given cluster. Because clusters have a fuzzy nature, large standard deviations may be expected (Table S1).

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^{N} (u_{ik})^m (y_k - v_i)^2}{\sum_{k=1}^{N} (u_{ik})^m}} \quad \text{(S2-3)}$$

S3. Comparison of PM_1 between IOP1 and IOP2

PM_1 mass concentrations at the T3 site during the dry and wet seasons differed by almost an order of magnitude. Figure S10a shows the statistics of mass concentrations of the NR-PM_1 components. The organic mass concentrations had the largest increase between IOP1 (wet season) and IOP2 (dry season), corresponding to a factor of 8. Mass concentrations increased by a factor of 6 for sulfate and ammonium, of 4 for nitrate, and of 2 for chloride. Mass concentrations of PM in the basin seem to have large interannual variability especially in the dry season due to the variability in biomass burning emissions (van Marle et al., 2017). Even so, the inter-season increases found in this study for the year of 2014 are in line with values previously

reported for other years, which vary between 3 and 10 (Artaxo et al., 1994; Holben et al., 1996; Fuzzi et al., 2007).

The observed increases can be rationalized in terms of important differences between the wet and dry seasons. One relevant aspect is that meteorological factors such as less precipitation and lower relative humidity (RH) in the dry season may lead to lower wet deposition (Figure S11). In addition, higher solar irradiance may favor the photochemical processing of VOCs and thereby the production of PM, in spite of higher temperatures which may favor partitioning to the gas phase (Figure S11). As a direct result of lower wet deposition, higher particle number and mass concentrations may be maintained (Figure S10a-b). As an indirect result, particles with a longer atmospheric lifetime can continue to grow to larger sizes through condensation, especially given the increased solar irradiance, also leading to increased mass concentrations. A comparison of volume-diameter distributions for the two seasons (Figure S10c) shows that there was a shift from peak D_m of 340 nm in IOP1 to 400 nm in IOP2.

Another relevant feature of the dry season is the basin-wide increased occurrence of fires (Artaxo et al., 2013; Martin et al., 2016). Biomass burning can contribute both primary particles and gaseous emissions that may be precursors for the production of secondary material. As shown in Figure S10b, there was a significant shift to large particle number concentrations from IOP1 to IOP2, with median values of 1060 cm^{-3} and 3240 cm^{-3}, respectively. Taken together, these results suggest that the increased mass concentrations observed in the dry season compared to the wet season were due to a combination of larger number concentrations and larger particle diameters, driven both by meteorological and anthropogenic factors.

S4. Calculations of PM optical properties

In order to estimate the absorption coefficient of BrC at 370 nm, $b_{abs,BrC}$, the absorption coefficient of BC at the same wavelength, $b_{abs,BC}$, had to first be determined (Eq. 1). The calculation of $å_{abs,BC}$ and consequently of $b_{abs,BC}$ was done through four methods, of which Method 2 was used in the analysis described in the main text. Herein, details and assumptions of the four methods as well as a comparison among their results are presented.

The assumptions of each method and a description of their meaning is presented in Table S2. Methods 1 and 2 assume a constant absorption Ångstrom exponent $å_{abs,BC}$ across the spectrum. Methods 3 and 4 assume a varying $å_{abs,BC}$, and the difference between $å_{abs,BC}$ at longer wavelengths and shorter wavelengths is accounted for by δ, which is the wavelength dependence of the absorption Ångstrom exponent, also known as WDA (Wang et al., 2016). The value of δ is calculated theoretically using Mie Theory and assuming spherical particles. The calculation also assumes a range of BC size distributions and coatings unless measurements are available. Once $å_{abs,BC}$ is estimated, $b_{abs,BC}$ can be calculated through Eq. 2.

A comparison of the estimated values for $b_{abs,BrC}$ through the different methods is presented in Figure S13. On average, $b_{abs,BrC}$ values from method 1 are 45% larger than method 2, and values from methods 3 and 4 are 6 to 20% larger than method 2. Method 2 was chosen because (i) it represents an improvement over method 1, as it calculates $å_{abs,BC}$ sample by sample and does not simply assume a value of 1, and (ii) although methods 3 and 4 consider a wavelength dependence of $å_{abs,BC}$, this dependence is unknown for our study. Method 3 relies purely on Mie modeling, and assumes spherical particles and ranges of BC size distribution and coating taken from global averages that might not be representative of our site. Method 4 uses BC size distribution data from a different site, which is an improvement but might still not be representative, and the mixing state is also not known. Because these methods might bring

additional uncertainty, method 2 is chosen as the base case. Because method 2 yields the lowest values, it can also be seen as a conservative method that establishes a lower bound for the particle absorption properties.

S5. Attribution of BrC absorption for the clusters

The attribution of BrC absorption for the three clusters is shown in Figure S15. The differences were overall small. Biomass burning factors represented the dominant brown carbon components in the dry season afternoons, accounting for about 50% of $b_{abs,BrC}$ under all conditions. For the urban cluster, between the two BBOA factors, a larger proportion of absorption was attributed to the MO-BBOA factor. This result highlights that the increased concentrations of secondary products from biomass burning emissions possibly driven by the interaction with the oxidant-rich Manaus plume affected the total absorption by organic PM_1. Regarding the IEPOX-SOA factor, its relative importance was larger when the influence of urban emissions was lower, corresponding to the baseline and event clusters. The HOA factor was associated with the largest estimated E_{abs} but its loadings were usually small. As a result, its contribution to BrC absorption became comparable to the LO-BBOA and IEPOX-SOA factors only for the urban cluster, which had the highest HOA loadings due to Manaus emissions.

References

Artaxo, P., Gerab, F., Yamasoe, M. A., and Martins, J. V.: Fine mode aerosol composition at three long-term atmospheric monitoring sites in the Amazon Basin, J. Geophys. Res. Atmos., 99, D11, 22857-22868, https://doi.org/10.1029/94JD01023 1994.

Artaxo, P., Rizzo, L. V., Brito, J. F., Barbosa, H. M. J., Arana, A., Sena, E. T., Cirino, G. G., Bastos, W., Martin, S. T., and Andreae, M. O.: Atmospheric aerosols in Amazonia and land use change: from natural biogenic to biomass burning conditions, Faraday Disc., 165, 0, 203-235, https://doi.org/10.1039/C3FD00052D, 2013.

Bezdek, J. C., Ehrlich, R., and Full, W.: FCM: The fuzzy c-means clustering algorithm, Comput. Geosci., 10, 2, 191-203, https://doi.org/10.1016/0098-3004(84)90020-7, 1984.

Canagaratna, M. R., Jimenez, J. L., Kroll, J. H., Chen, Q., Kessler, S. H., Massoli, P., Hildebrandt Ruiz, L., Fortner, E., Williams, L. R., Wilson, K. R., Surratt, J. D., Donahue, N. M., Jayne, J. T., and Worsnop, D. R.: Elemental ratio measurements of organic compounds using aerosol mass spectrometry: characterization, improved calibration, and implications, Atmos. Chem. Phys., 15, 1, 253-272, https://doi.org/10.5194/acp-15-253-2015, 2015.

Chatzis, S. P.: A fuzzy c-means-type algorithm for clustering of data with mixed numeric and categorical attributes employing a probabilistic dissimilarity functional, Exp. Syst. Appl., 38, 7, 8684-8689, https://doi.org/10.1016/j.eswa.2011.01.074, 2011.

de Sá, S. S., Palm, B. B., Campuzano-Jost, P., Day, D. A., Hu, W., Isaacman-VanWertz, G., Yee, L. D., Brito, J., Carbone, S., Ribeiro, I. O., Cirino, G. G., Liu, Y. J., Thalman, R., Sedlacek, A., Funk, A., Schumacher, C., Shilling, J. E., Schneider, J., Artaxo, P., Goldstein, A. H., Souza, R. A. F., Wang, J., McKinney, K. A., Barbosa, H., Alexander, M. L., Jimenez, J. L., and Martin, S. T.: Urban influence on the concentration and composition of submicron particulate matter in central Amazonia, Atmos. Chem. Phys. Discuss., 2018, 1-56, https://doi.org/10.5194/acp-2018-172, 2018.

Farmer, D. K., Matsunaga, A., Docherty, K. S., Surratt, J. D., Seinfeld, J. H., Ziemann, P. J., and Jimenez, J. L.: Response of an aerosol mass spectrometer to organonitrates and organosulfates and implications for atmospheric chemistry, Proc. Natl. Acad. Sci. USA, 107, 15, 6670-6675, https://doi.org/10.1073/pnas.0912340107, 2010.

Fry, J. L., Draper, D. C., Zarzana, K. J., Campuzano-Jost, P., Day, D. A., Jimenez, J. L., Brown, S. S., Cohen, R. C., Kaser, L., Hansel, A., Cappellin, L., Karl, T., Hodzic Roux, A., Turnipseed, A., Cantrell, C., Lefer, B. L., and Grossberg, N.: Observations of gas- and aerosol-phase organic nitrates at BEACHON-RoMBAS 2011, Atmos. Chem. Phys., 13, 17, 8585-8605, https://doi.org/10.5194/acp-13-8585-2013, 2013.

Fuzzi, S., Decesari, S., Facchini, M. C., Cavalli, F., Emblico, L., Mircea, M., Andreae, M. O., Trebs, I., Hoffer, A. s., Guyon, P., Artaxo, P., Rizzo, L. V., Lara, L. L., Pauliquevis, T., Maenhaut, W., Raes, N., Chi, X., Mayol-Bracero, O. L., Soto-Garcı, L. L., Claeys, M., Kourtchev, I., Rissler, J., Swietlicki, E., Tagliavini, E., Schkolnik, G., Falkovich, A. H.,

Rudich, Y., Fisch, G., and Gatti, L. V.: Overview of the inorganic and organic composition of size-segregated aerosol in Rondonia, Brazil, from the biomass-burning period to the onset of the wet season, J. Geophys. Res. Atmos., 112, D01201, https://doi.org/10.1029/2005JD006741, 2007.

Hathaway, R. J. and Bezdek, J. C.: Fuzzy c-means clustering of incomplete data, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 31, 5, 735-744, https://doi.org/10.1109/3477.956035, 2001.

Holben, B. N., Setzer, A., Eck, T. F., Pereira, A., and Slutsker, I.: Effect of dry-season biomass burning on Amazon basin aerosol concentrations and optical properties, 1992–1994, J. Geophys. Res. Atmos., 101, D14, 19465-19481, https://doi.org/10.1029/96JD01114 1996.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: observations and modeling of the green ocean Amazon (GoAmazon2014/5), Atmos. Chem. Phys., 16, 8, 4785-4797, https://doi.org/10.5194/acp-16-4785-2016, 2016.

Middlebrook, A. M., Bahreini, R., Jimenez, J. L., and Canagaratna, M. R.: Evaluation of composition-dependent collection efficiencies for the aerodyne aerosol mass spectrometer using field data, Aerosol Sci. Technol., 46, 3, 258-271, https://doi.org/10.1080/02786826.2011.620041, 2012.

Saturno, J., Pöhlker, C., Massabò, D., Brito, J., Carbone, S., Cheng, Y., Chi, X., Ditas, F., de Angelis, I. H., Morán-Zuloaga, D., Pöhlker, M., Rizzo, L. V., Walter, D., Wang, Qiaoqiao, Artaxo, P., Prati, P., and Andreae, M. O.: Comparison of different Aethalometer correction schemes and a reference multi-wavelength absorption technique for ambient aerosol data, Atmos. Meas. Tech., 10, 8, 2837, https://doi.org/10.5194/amt-10-2837-2017, 2017.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 9, 2891-2918, https://doi.org/10.5194/acp-9-2891-2009, 2009.

van Marle, M. J. E., Field, R. D., Werf, G. R., Estrada de Wagt, I. A., Houghton, R. A., Rizzo, L. V., Artaxo, P., and Tsigaridis, K.: Fire and deforestation dynamics in Amazonia (1973–2014), Global Biogeochem. Cy., 31, 1, 24-38, https://doi.org/10.1002/2016GB005445, 2017.

Wang, X., Heald, C. L., Sedlacek, A. J., de Sá, S. S., Martin, S. T., Alexander, M. L., Watson, T. B., Aiken, A. C., Springston, S. R., and Artaxo, P.: Deriving brown carbon from multiwavelength absorption measurements: method and application to AERONET and Aethalometer observations, Atmos. Chem. Phys., 16, 19, 12733-12752, https://doi.org/10.5194/acp-16-12733-2016, 2016.

List of Supplementary Figures

Figure S1. Scatter plot of AMS PM volume concentrations and SMPS PM volume concentrations for IOP2. SMPS1 measured particles having mobility diameters of 10 to 461 nm, and SMPS 2, 10 to 510 nm. SMPS1 measurements were available from August 16 to October 10, and SMPS2 measurements were available from August 16 to October 15. Material densities used in the calculation of AMS volume from AMS mass were based on a mixing rule for the five AMS-measured species. The material density of the organic component was calculated following the method of Kuwata et al. (2011) based on O:C and H:C values, which in turn were calculated following the method of Canagaratna et al. (2015).

Figure S2. Summary of the analysis for estimating organic and inorganic nitrates from AMS bulk measurements for IOP2. (a) Resulting time series of organic and inorganic nitrates are shown together with the original nitrate AMS times series. (b) Time series of the fraction of organic nitrate in total nitrate. (c) Time series of the measured NO_2^+/NO^+ ratio is shown in red and values of NO_2^+/NO^+ from ammonium nitrate calibrations are shown in gray triangles. The reference ratio for inorganic nitrate over time is represented by the dashed dark blue line, which was the mean of the calibration values (grey triangles). The reference ratio for organic nitrates over time is represented by the dashed light blue line, and it was assumed to be a factor of 2.25 lower than that or inorganic nitrate based on previous field studies (Farmer et al., 2010; Fry et al., 2013). Calculations were done for data binned to one hour (as plotted), and the resulting time series were interpolated to the native time stamp for employment in data analyses.

Figure S3. Diagnostics of the PMF analysis for IOP2. (a) Time series of total ion residuals of PMF solutions from one to six factors, (b) Statistics of the sum of residuals presented in panel a. Box plots show the interquartile ranges, including the medians as a horizontal line. Red markers show the means. Whiskers show the 5 and 95 percentiles. (c) Dependence of the quality-of-fit parameter $Q/Q_{expected}$ on the number of factors for $f_{peak} = 0$, (d) Dependence of the quality-of-fit parameter $Q/Q_{expected}$ on f_{peak} for number of factors = 6. The red line represents $Q/Q_{expected}$ that exceeds in 0.1% the minimum value at $f_{peak} = 0$. This limit determines the range of plausible f_{peak} values as indicated by the dashed black lines.

Figure S4. Results of the PMF analysis for 5 factors (a and b) and 7 factors (c and d) for IOP2. Panels on the left (a and c) show the time series of factor loadings and panels on the right (b and d) show the profiles of factors. The signals shown in panels b and d were summed to unit mass resolution.

Figure S5. Average total organic mass spectra (top row) and PMF factor profiles. The signals are classified by their ion families and colored accordingly. For m/z > 60, signals are multiplied by three for clearer visualization.

Figure S6. Average mass spectra for nitrogen-containing organic material (top row) and PMF factor profiles showing only the nitrogen-containing ion families.

Figure S7. Peak fitting for important nitrogen-containing ions measured: (a) CHN (m/z = 27.011), (b) CH_4N (30.034) (c) C_2HN (39.011), (d) C_2H_3N (41.027), (e) CHNO (43.042), (f) CH_2NO_2 (60.009). For each panel, the rows show (a) spectra with chopper open, (b) spectra with chopper closed, and (c) the difference spectra between the cases of open and closed chopper. The spectra shown correspond to V-

mode data averaged over IOP2. The black circles are data, and the purple lines are the overall fits. The fits to each individual ion are shown in different colors. All ions that were fitted are labeled by vertical arrows and tags. In each panel, the column on the right shows the fitting that was used in the analysis, with the nitrogen-containing ion indicated in dark red, and the column on the left offers a comparison for the case when the ion is not included.

Figure S8. Value of the objective function of the FCM analysis in the last iteration plotted against the number of clusters.

Figure S9. Locations of cluster centroids from the FCM analysis for IOP2 as visualized by a 2-D projection on the plane defined by each pair of input variables. Results for two to five clusters are shown in panels a to d. Red circles are observational data and black squares are cluster centroids.

Figure S10. Statistical comparison of NR-PM_1 at T3 between the wet and dry seasons (IOP1 and IOP2, respectively). (a) Mass concentrations of components measured by the AMS. The left axis scale refers to organic species, and the right axis refers to the inorganic species, as indicated by arrows at the top of plot. Solid markers represent means, whiskers show 5 percentiles and 95 percentiles, boxes span interquartile range, and horizontal line inside boxes indicates medians. Concentrations were adjusted to standard temperature (273.15 K) and pressure (10^5 Pa). (b) Probability density function for particle numbers concentrations. Vertical dashed lines indicate the medians of distributions. (c) Volume-diameter distributions measured by a Scanning Mobility Particle Sizer (SMPS). D_m represents the mobility diameter.

Figure S11. Statistical comparison of meteorological variables at T3 between the wet and dry seasons (IOP1 and IOP2, respectively). (a) Relative humidity, (b) Temperature, and (c) Solar irradiance. Boxes represent interquartile ranges, markers inside boxes represent medians, and lines represent means.

Figure S12. Diel trends of families of nitrogen-containing ions. Organic and inorganic nitrates were estimated based on the ratio of NO^+ and NO_2^+ ions measured by the AMS (Section S1.2). $C_xH_yO_zN_p^+$ refers to the sum of ions containing at least one carbon atom and one nitrogen atom as measured by the AMS.

Figure S13. Comparison of $b_{abs,Brc}$ values calculated through method 2 and through the three other methods. Scatter plots depict results from methods (a) 1, (b) 3, and (c) 4 on the ordinate against method 2 on the abscissa.

Figure S14. Relationships between the brown-carbon absorption coefficient and the fractional contributions to organic PM_1 of (a) the $C_xH_yO_z^+$ family and (b) the $C_xH_yO_zN_p^+$ family. Boxes indicate interquartile ranges, and horizontal lines within the boxes indicate medians.

Figure S15. Attribution of BrC absorption, as represented by $b_{abs,BrC}$, to the components of organic PM_1, as represented by the PMF factors. Calculations were made based on the typical cluster composition as described by centroid values (Table 1) and on the estimated E_{abs} values for each of the PMF factors (Table 3). Absorption by "other" corresponds to the contribution of the model constant B.

List of Supplementary Tables

Table S1. Coordinates of cluster centroids for input variables, AMS species concentrations, PMF factor loadings and other PM characteristics and properties. Table entries for AMS species and PMF factors are plotted in Figure 10. Concentrations of NO_y, O_3, BC, CO, sulfate, and particle number were used as input in the clustering analysis.

Species	Clusters (centroid ± standard deviation)		
	Baseline	Event	Urban
Input variables			
Particle number (cm^{-3})	2007 ± 752	3179 ± 938	4638 ± 1040
NO_y (ppb)	1.27 ± 0.36	1.70 ± 0.44	2.64 ± 0.56
O_3 (ppb)	32.5 ± 6.4	42.8 ± 7.2	56.4 ± 8.9
rBC ($\mu g\ m^{-3}$)	0.18 ± 0.10	0.28 ± 0.12	0.33 ± 0.14
CO (ppb)	159 ± 30	179 ± 35	178 ± 33
Sulfate ($\mu g\ m^{-3}$)	1.41 ± 0.51	2.32 ± 0.87	1.85 ± 0.50
AMS species concentrations ($\mu g\ m^{-3}$)			
Organic	7.28 ± 2.54	9.19 ± 2.60	9.67 ± 2.49
Ammonium	0.38 ± 0.12	0.59 ± 0.19	0.52 ± 0.12
Nitrate	0.11 ± 0.04	0.15 ± 0.04	0.22 ± 0.05
Chloride	0.012 ± 0.004	0.014 ± 0.004	0.014 ± 0.004
PMF factor loadings ($\mu g\ m^{-3}$)			
MO-OOA	1.64 ± 0.98	2.12 ± 1.16	2.48 ± 0.87
LO-OOA	2.84 ± 0.83	2.88 ± 0.97	3.52 ± 0.99
IEPOX-SOA	1.39 ± 0.82	2.30 ± 1.06	1.41 ± 0.58
MO-BBOA	0.70 ± 0.70	0.95 ± 0.85	1.30 ± 0.92
LO-BBOA	0.34 ± 0.37	0.49 ± 0.53	0.39 ± 0.51
HOA	0.12 ± 0.20	0.22 ± 0.27	0.31 ± 0.22
Optical properties			
$b_{abs,BrC}$ (Mm^{-1})	1.5 ± 1.6	2.6 ± 1.9	2.4 ± 1.9
$å_{abs}$	1.5 ± 0.4	1.7 ± 0.4	1.7 ± 0.4
Nitrogen-containing families			
$C_xH_yO_zN_p^+$ family ($\mu g\ m^{-3}$)	0.25 ± 0.09	0.34 ± 0.10	0.33 ± 0.09
Organic nitrates ($\mu g\ m^{-3}$)	0.12 ± 0.04	0.15 ± 0.04	0.21 ± 0.04
Inorganic nitrates ($\mu g\ m^{-3}$)	0.005 ± 0.014	0.007 ± 0.020	0.005 ± 0.016

Table S2. Assumptions and description of the four methods used to calculate the absorption Ångstrom exponent of BC, $å_{abs,BC}$.

Method	Assumptions	Description
1	$å_{abs,BC} = 1$	• $å_{abs,BC}$ is not wavelength dependent • $å_{abs,BC}$ is equal to 1 at any point in time (absorption is constant over the spectrum)
2	$å_{abs,BC} = å_{abs}(700,880)$	• $å_{abs,BC}$ is not wavelength dependent • $å_{abs,BC}$ is calculated for each point in time from aethalometer measurements at the two largest wavelengths
3	$å_{abs,BC} = å_{abs}(700,880) + \delta$ $\delta = -0.1$	• $å_{abs,BC}$ is wavelength dependent • δ value was based on the theoretical calculations of Wang et al. (2016)
4	$å_{abs,BC} = å_{abs}(700,880) + \delta$ $\delta = -0.3$	• $å_{abs,BC}$ is wavelength dependent • δ value was based on Saturno et al. (2017), which relied on BC size distribution measurements at the T0a site

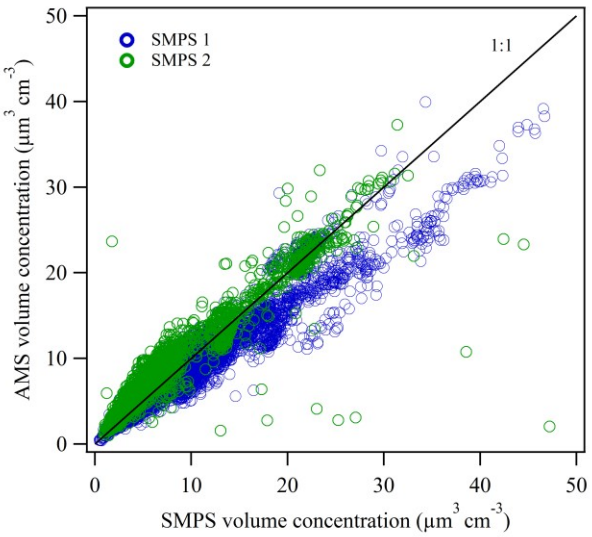
SMPS 1
SMPS 2
1:1
AMS volume concentration (µm^3 cm^{-3})
SMPS volume concentration (µm^3 cm^{-3})
0
10
20
30
40
50

Figure S1

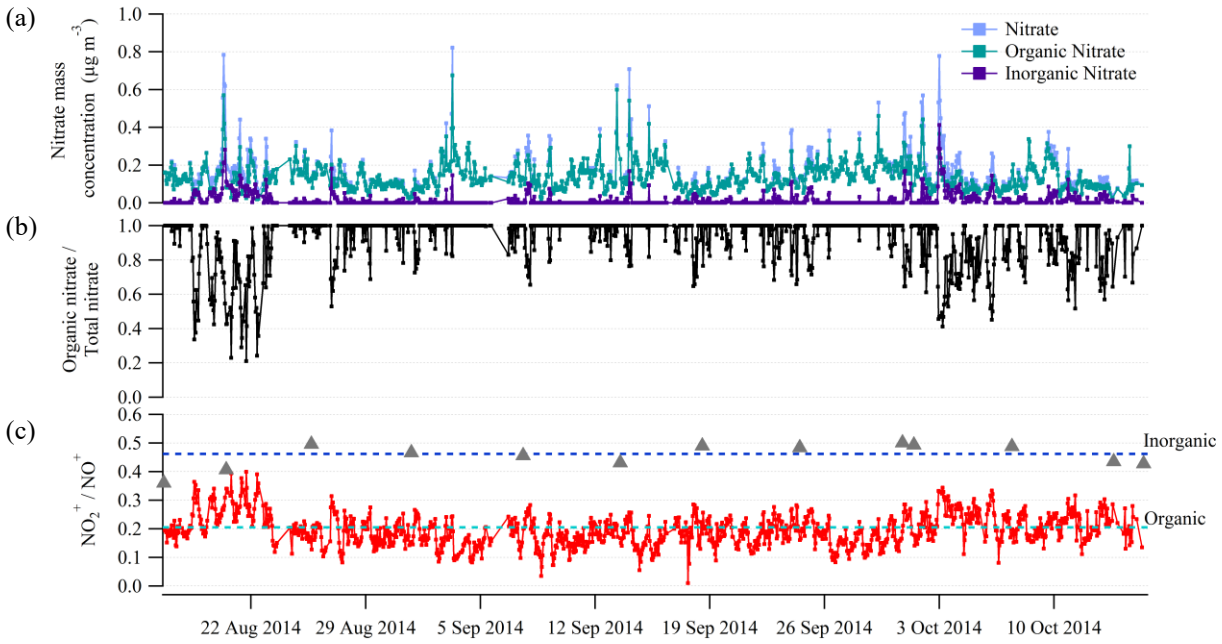
(a)
Nitrate mass concentration (μg m-3)
Nitrate
Organic Nitrate
Inorganic Nitrate
(b)
Organic nitrate / Total nitrate
(c)
NO2+ / NO+
Inorganic
Organic
22 Aug 2014
29 Aug 2014
5 Sep 2014
12 Sep 2014
19 Sep 2014
26 Sep 2014
3 Oct 2014
10 Oct 2014

Figure S2

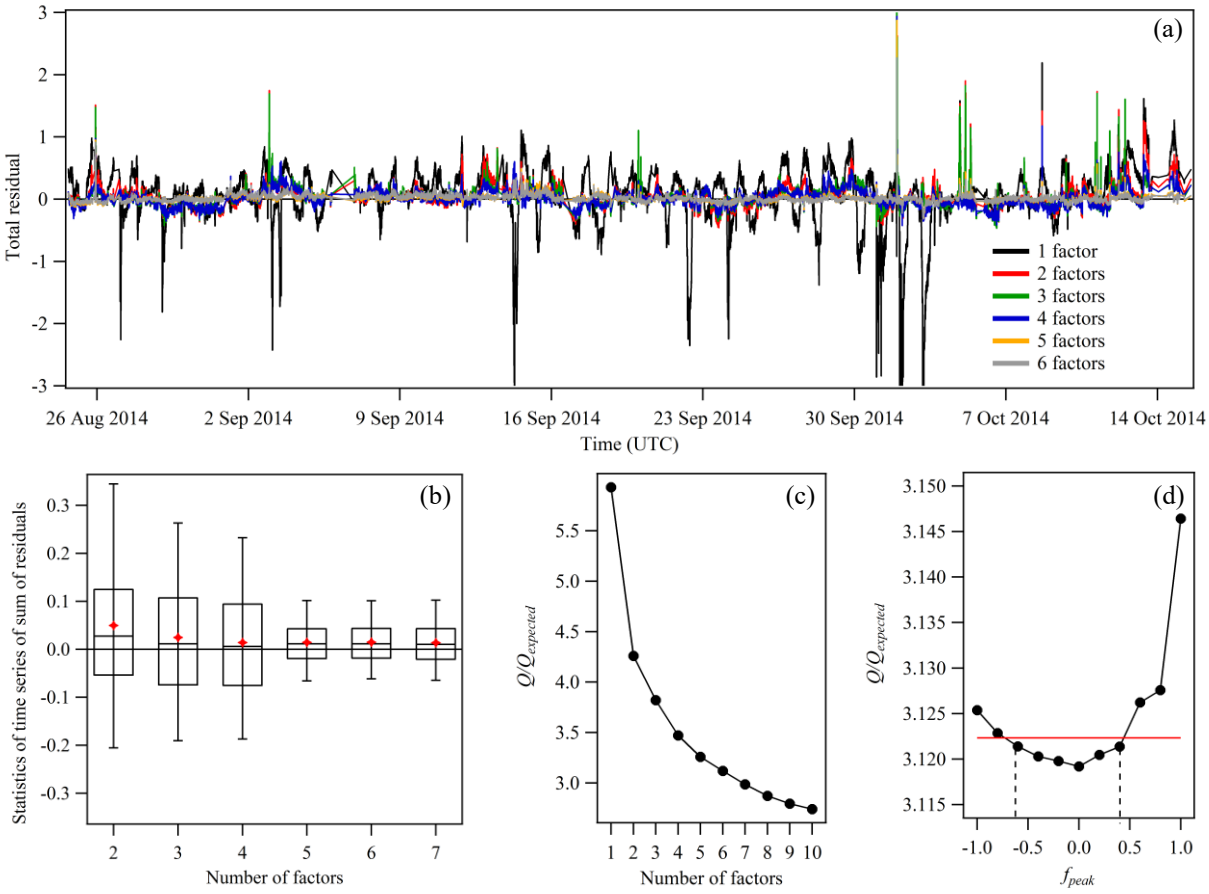
(a)
Total residual
Time (UTC)
26 Aug 2014
2 Sep 2014
9 Sep 2014
16 Sep 2014
23 Sep 2014
30 Sep 2014
7 Oct 2014
14 Oct 2014
1 factor
2 factors
3 factors
4 factors
5 factors
6 factors
(b)
Statistics of time series of sum of residuals
Number of factors
(c)
Q/Q_expected
Number of factors
(d)
Q/Q_expected
f_peak

Figure S3

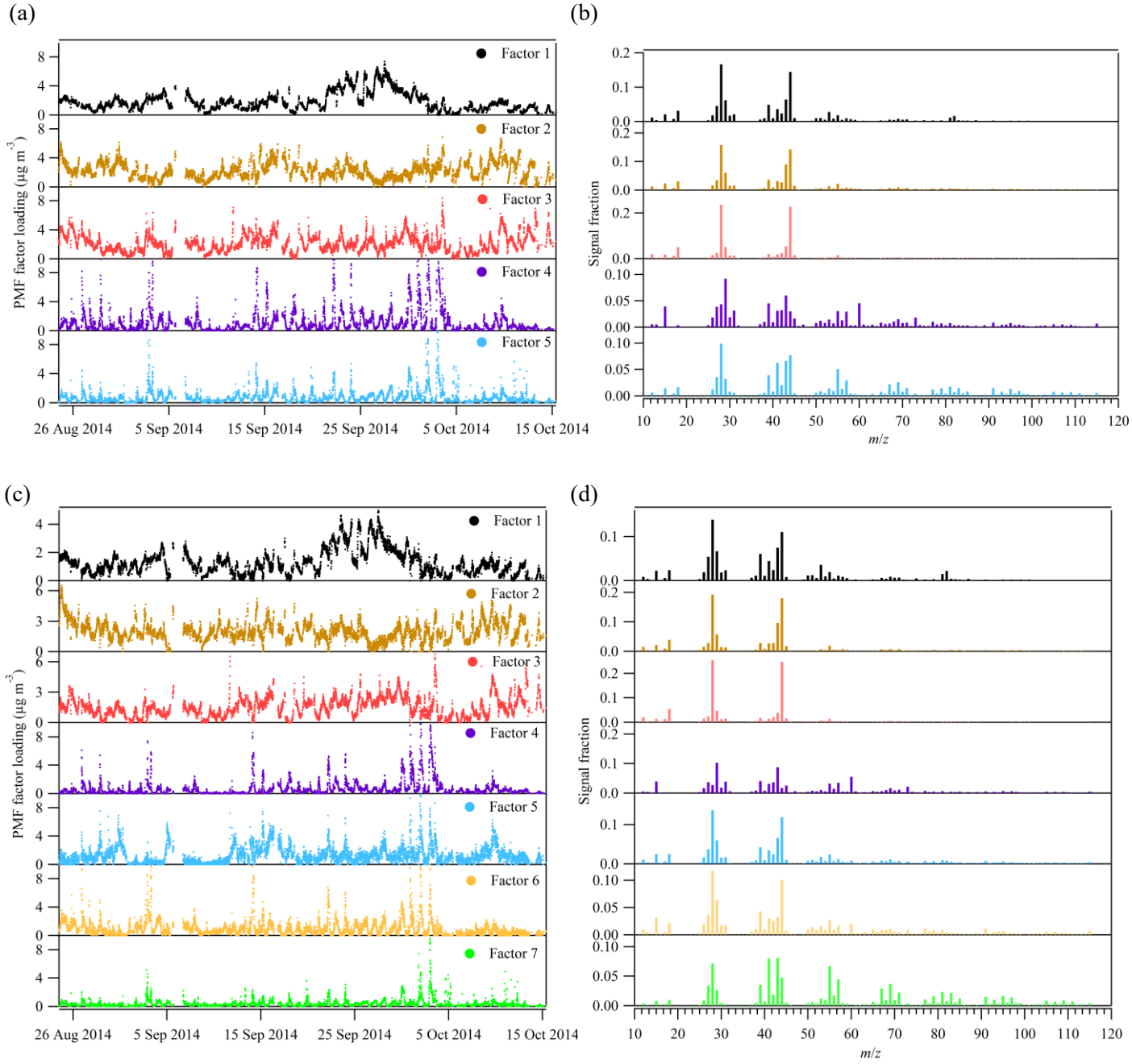
(a)
(b)
(c)
(d)
Factor 1
Factor 2
Factor 3
Factor 4
Factor 5
Factor 6
Factor 7
PMF factor loading (μg m-3)
Signal fraction
m/z
26 Aug 2014
5 Sep 2014
15 Sep 2014
25 Sep 2014
5 Oct 2014
15 Oct 2014

Figure S4

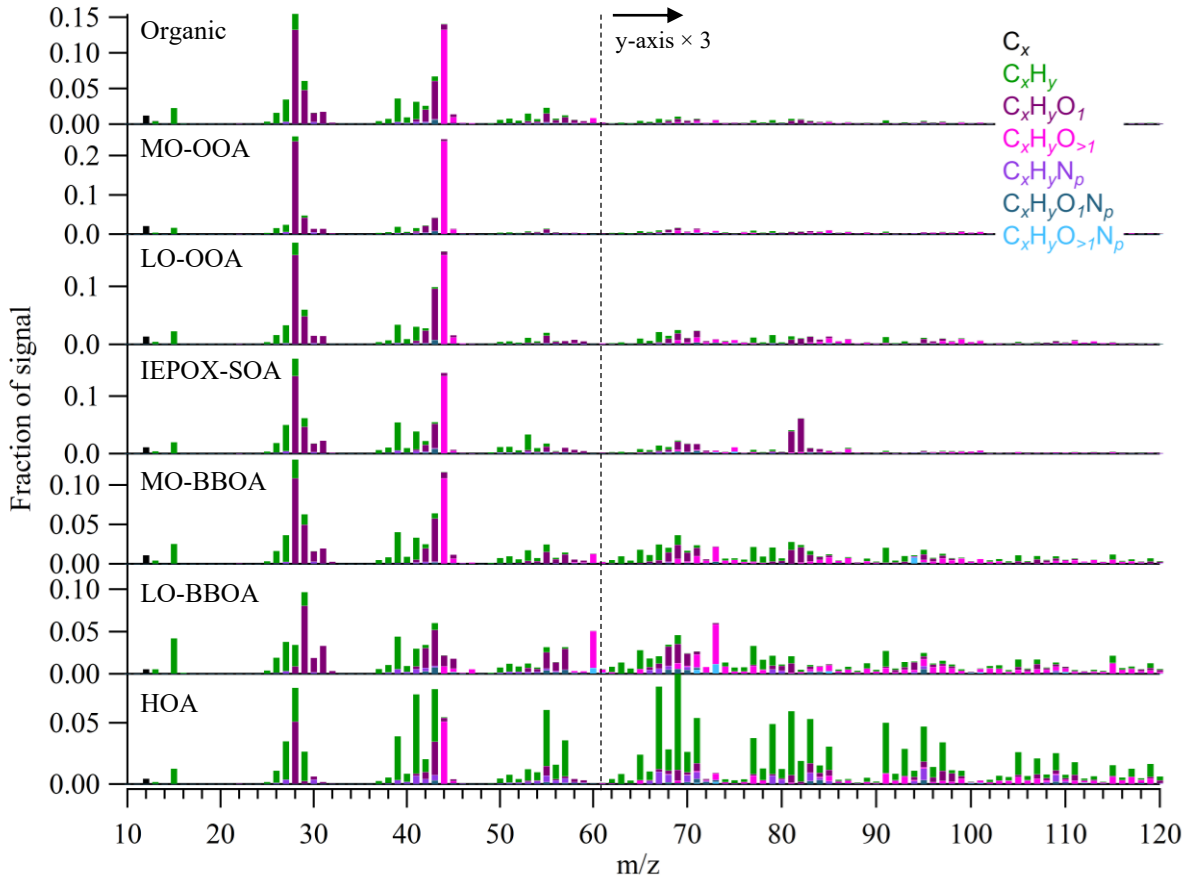
Organic
MO-OOA
LO-OOA
IEPOX-SOA
MO-BBOA
LO-BBOA
HOA
y-axis × 3
C_x
C_xH_y
$C_xH_yO_1$
$C_xH_yO_{>1}$
$C_xH_yN_p$
$C_xH_yO_1N_p$
$C_xH_yO_{>1}N_p$
Fraction of signal
m/z

Figure S5

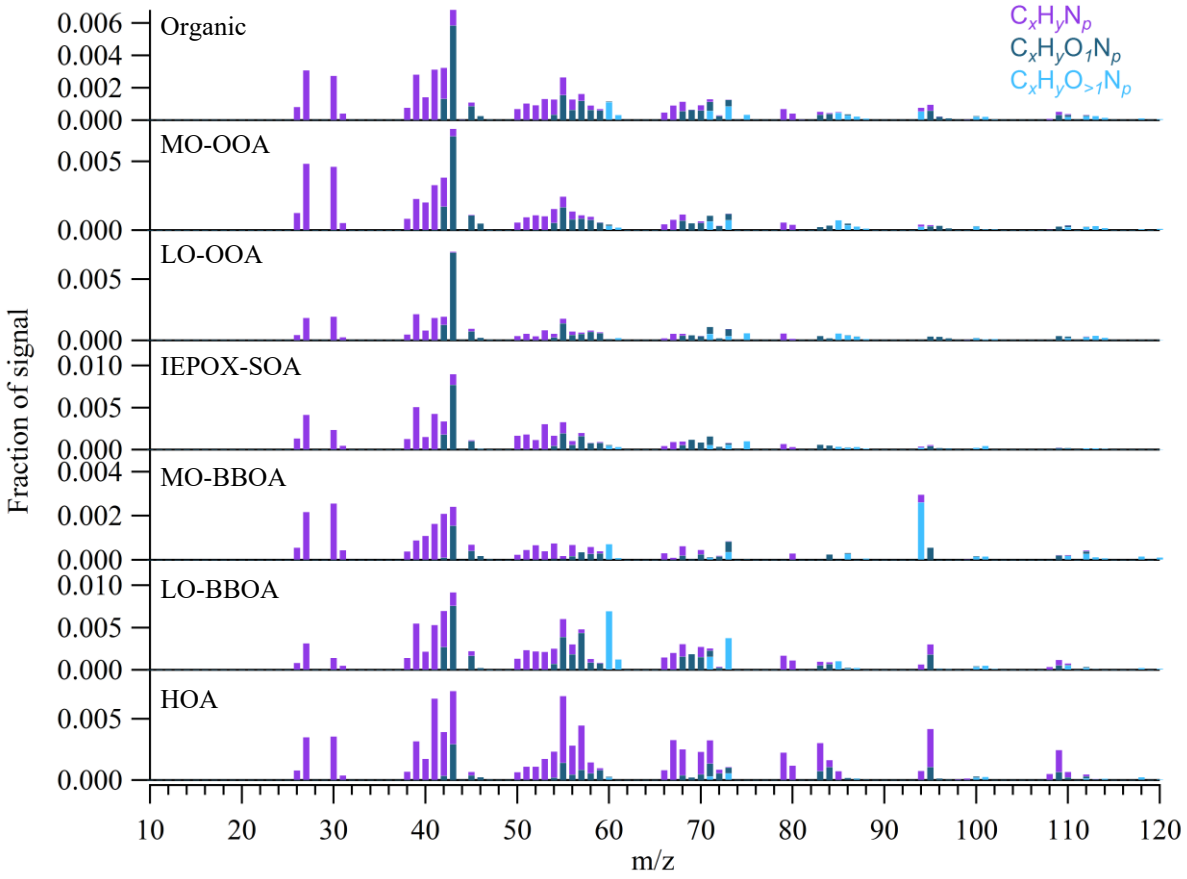
Organic
MO-OOA
LO-OOA
IEPOX-SOA
MO-BBOA
LO-BBOA
HOA
$C_xH_yN_p$
$C_xH_yO_1N_p$
$C_xH_yO_{>1}N_p$
Fraction of signal
m/z
0.006
0.004
0.002
0.000
0.005
0.010
10
20
30
40
50
60
70
80
90
100
110
120

Figure S6

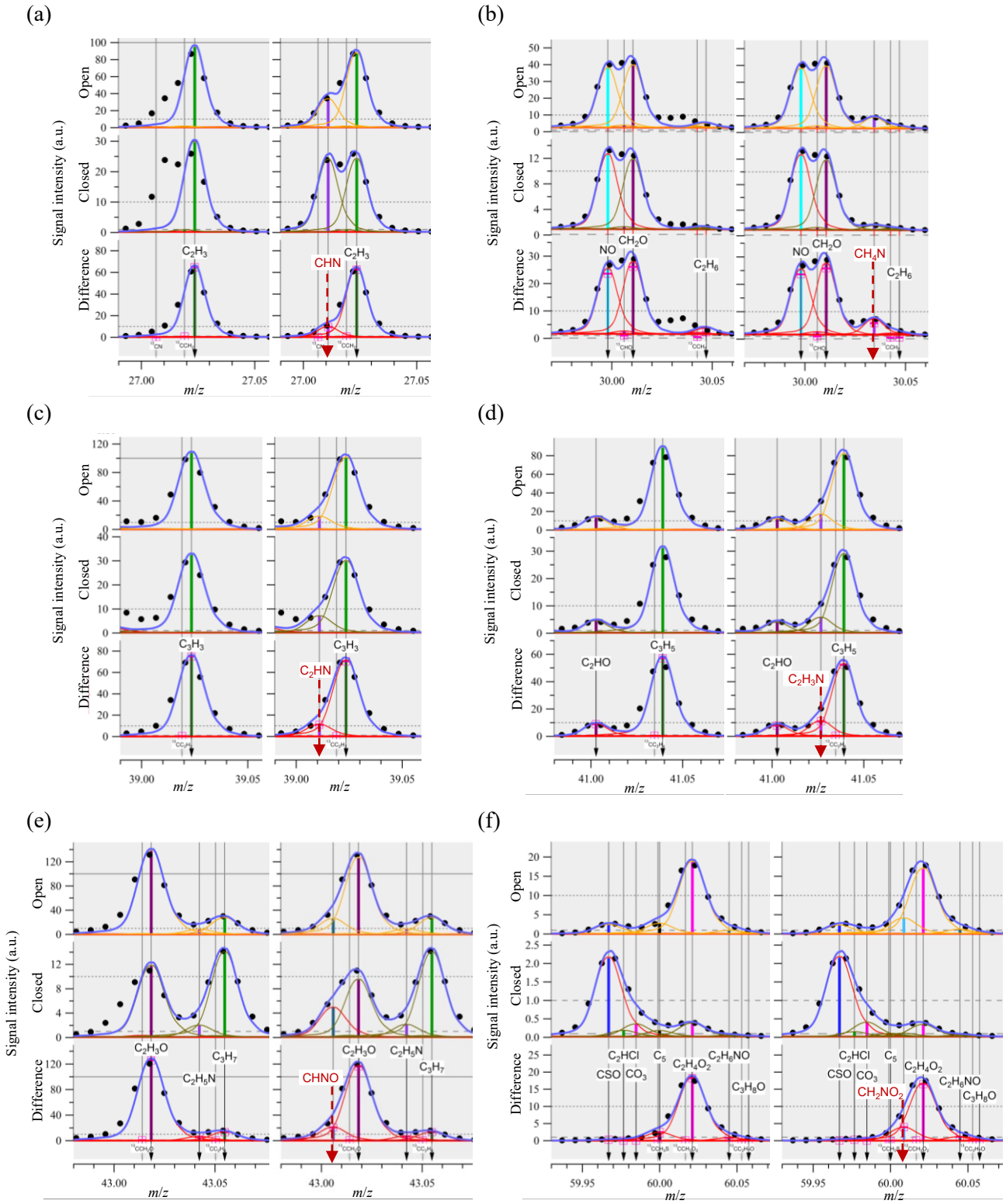
(a)
(b)
(c)
(d)
(e)
(f)
Signal intensity (a.u.)
Open
Closed
Difference
m/z
CHN
C2HN
CHNO
CH4N
C2H3N
C2H5N
C2HO
C2HCl
CSO
CO3
CH2NO2
C2H3
C3H3
C3H5
C3H7
C2H3O
C2H5N
C2H6
C3H8O
C2H4O2
C2H6NO
C5
CH2O
NO

Figure S7

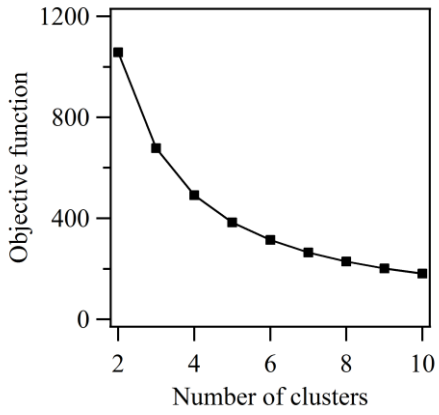
1200
800
400
0
Objective function
2
4
6
8
10
Number of clusters

Figure S8

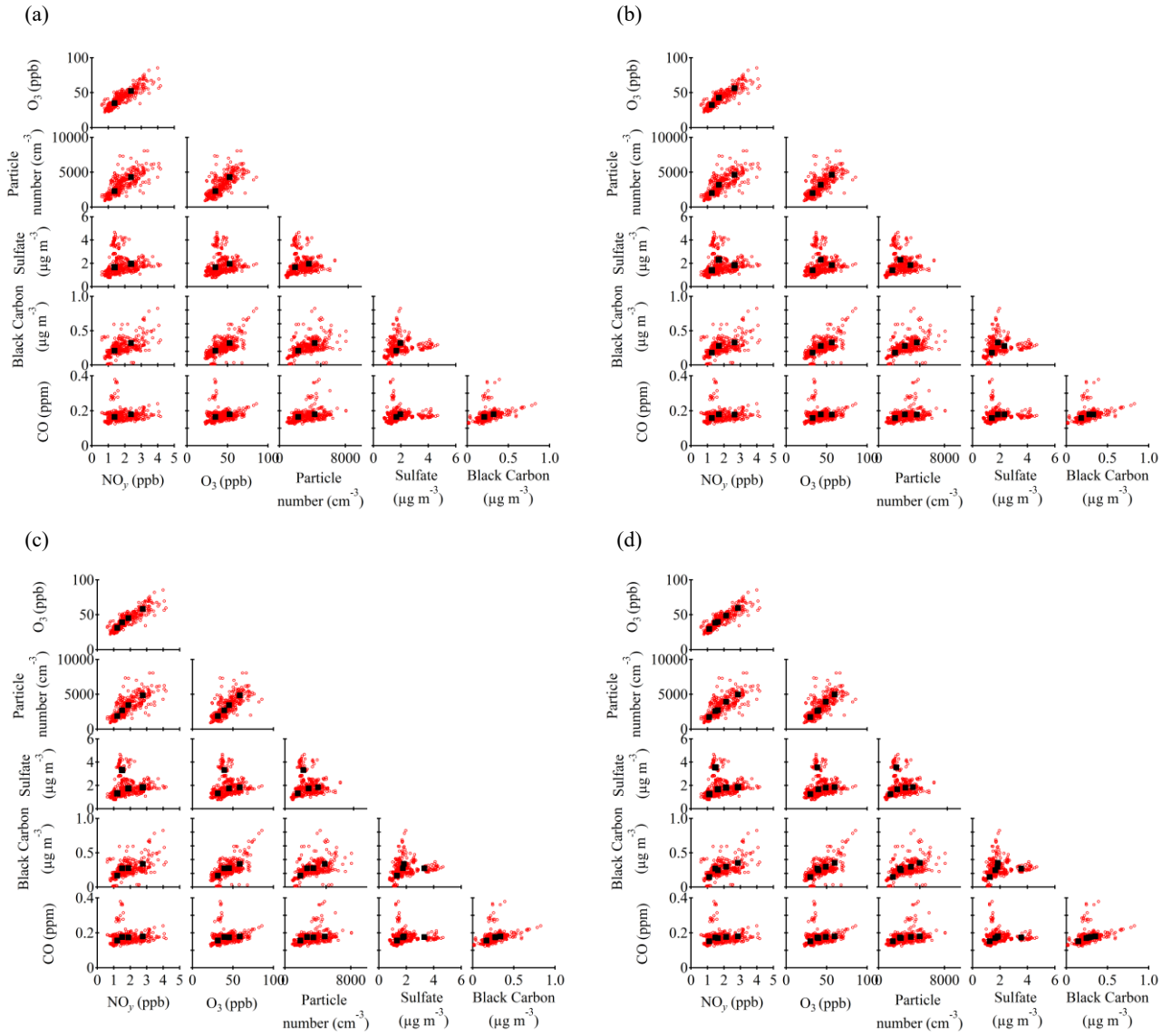
(a)
(b)
(c)
(d)
O3 (ppb)
Particle number (cm-3)
Sulfate (μg m-3)
Black Carbon (μg m-3)
CO (ppm)
NOy (ppb)
O3 (ppb)
Particle number (cm-3)
Sulfate (μg m-3)
Black Carbon (μg m-3)

Figure S9

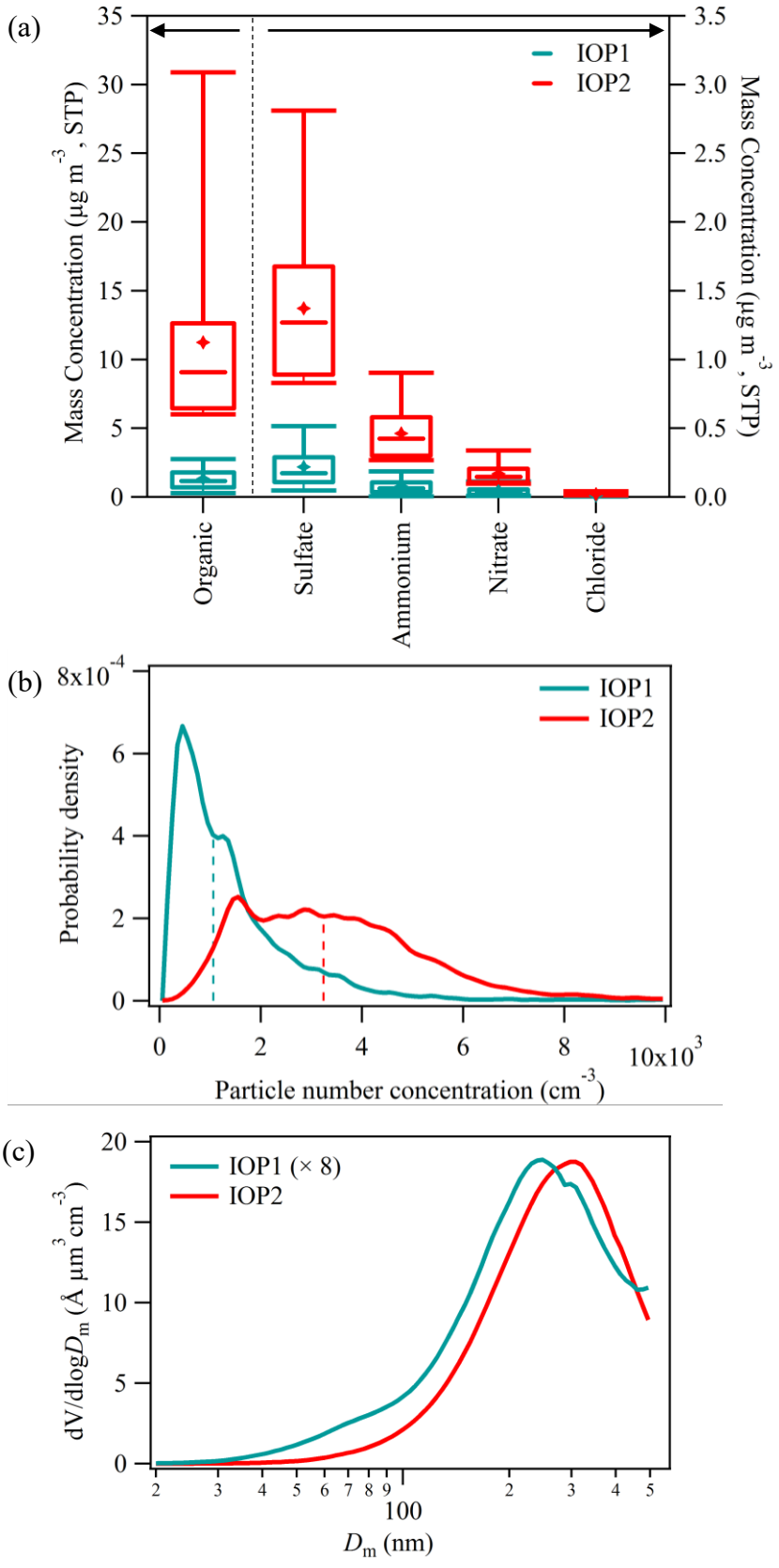
(a)
IOP1
IOP2
Mass Concentration (μg m-3, STP)
Mass Concentration (μg m-3, STP)
Organic
Sulfate
Ammonium
Nitrate
Chloride
(b)
IOP1
IOP2
Probability density
Particle number concentration (cm-3)
(c)
IOP1 (× 8)
IOP2
dV/dlogDm (Å μm3 cm-3)
Dm (nm)

Figure S10

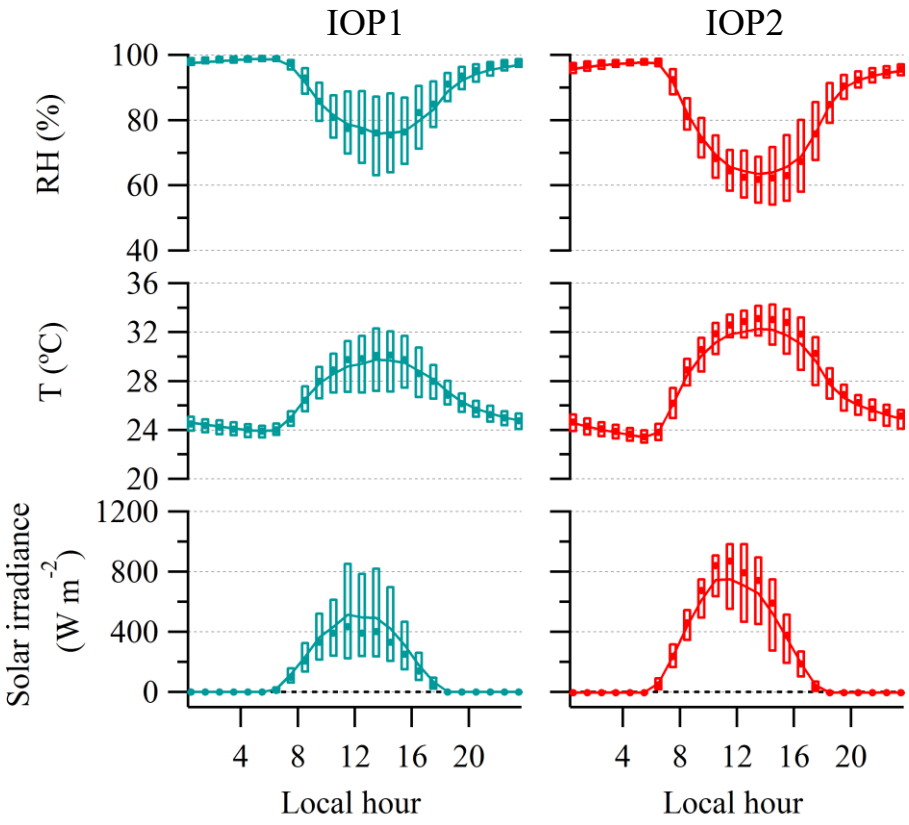
IOP1
IOP2
RH (%)
100
80
60
40
T (°C)
36
32
28
24
20
Solar irradiance
(W m-2)
1200
800
400
0
4
8
12
16
20
Local hour
4
8
12
16
20
Local hour

Figure S11

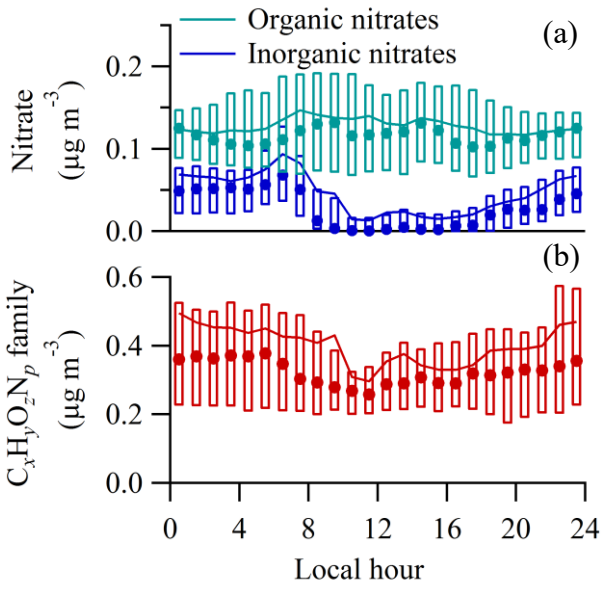
Organic nitrates
Inorganic nitrates
(a)
Nitrate
(µg m-3)
0.2
0.1
0.0
(b)
CxHyOzNp family
(µg m-3)
0.6
0.4
0.2
0.0
0
4
8
12
16
20
24
Local hour

Figure S12

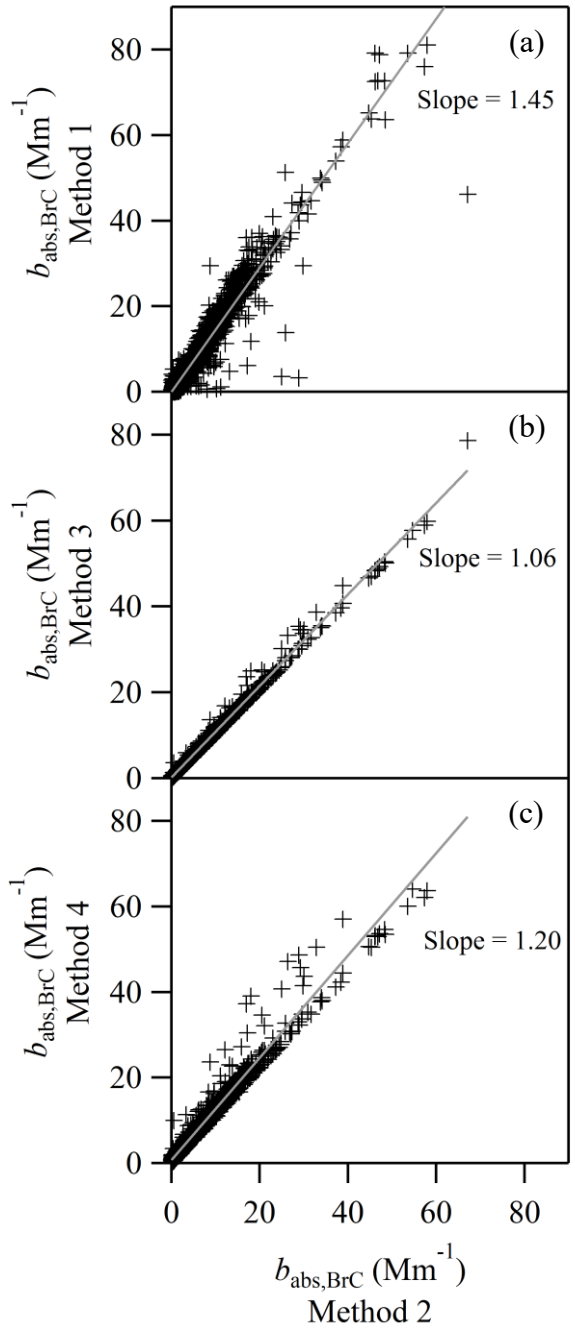
(a)
Slope = 1.45
(b)
Slope = 1.06
(c)
Slope = 1.20
$b_{abs,BrC}$ (Mm^{-1})
Method 1
$b_{abs,BrC}$ (Mm^{-1})
Method 3
$b_{abs,BrC}$ (Mm^{-1})
Method 4
$b_{abs,BrC}$ (Mm^{-1})
Method 2
0
20
40
60
80

Figure S13

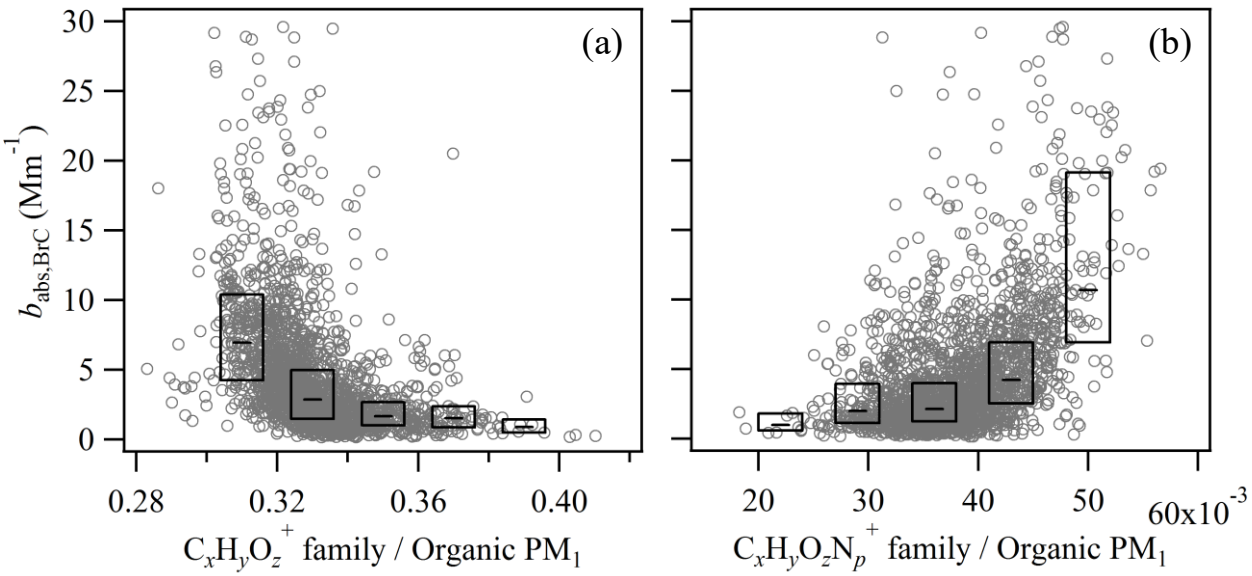
(a)
(b)
$b_{abs,BrC}$ (Mm^{-1})
30
25
20
15
10
5
0
0.28
0.32
0.36
0.40
20
30
40
50
60x10^{-3}
$C_xH_yO_z^+$ family / Organic PM_1
$C_xH_yO_zN_p^+$ family / Organic PM_1

Figure S14

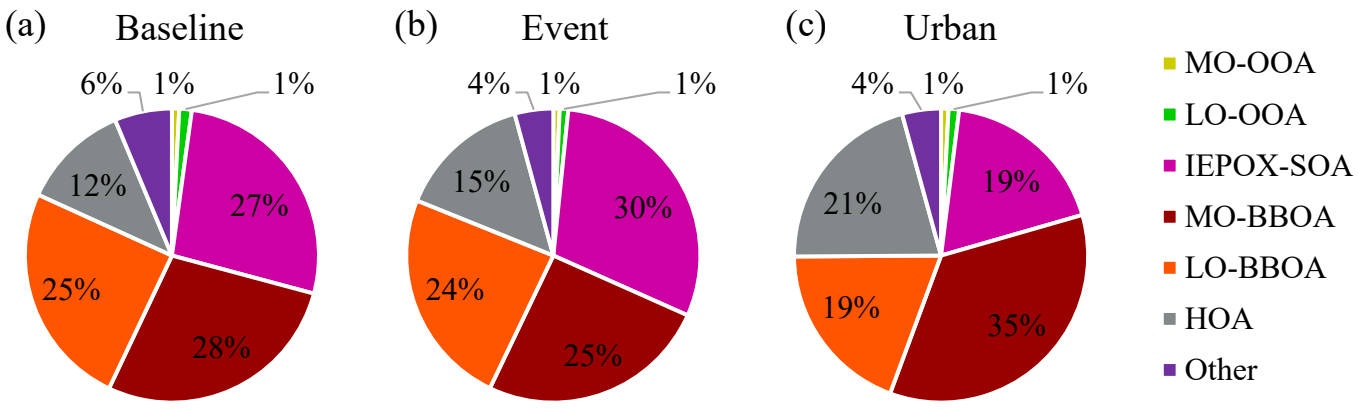
(a) Baseline
6%
1%
1%
27%
28%
25%
12%
(b) Event
4%
1%
1%
30%
25%
24%
15%
(c) Urban
4%
1%
1%
19%
35%
19%
21%
MO-OOA
LO-OOA
IEPOX-SOA
MO-BBOA
LO-BBOA
HOA
Other

Figure S15