

Supplement of Atmos. Chem. Phys., 19, 3645–3672, 2019  
<https://doi.org/10.5194/acp-19-3645-2019-supplement>  
© Author(s) 2019. This work is distributed under  
the Creative Commons Attribution 4.0 License.



*Supplement of*

## **Constructing a data-driven receptor model for organic and inorganic aerosol – a synthesis analysis of eight mass spectrometric data sets from a boreal forest site**

**Mikko Äijälä et al.**

*Correspondence to:* Mikael Ehn ([mikael.ehn@helsinki.fi](mailto:mikael.ehn@helsinki.fi))

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

## S.1 Chemical Mass Balance model

Chemical mass balance (CMB) models are typically used as receptor models for cases where source profiles are known, and only the mass loading information needs resolving (Friedlander, 1973; Gordon, 1988; Hopke, 1991a; Miller et al., 1972). In receptor modelling based on mass conservation, the observed loading  $x$  of variable  $v$  at sampling time  $t$  can be modelled as a  
5 sum of contributions from all sources  $s = 1 \dots p$  (e.g. Hopke, 1991b):

$$x_{v,t} = \sum_{s=1}^p f_{v,s} m_{s,t} + \varepsilon_{v,t}. \quad (\text{S.1})$$

The contribution from a single source ( $f_{v,s} m_{s,t}$ ) is the total mass,  $m_{s,t}$ , emitted from that source multiplied by the fraction of  
10 variable  $v$  in the emission profile of said source,  $f_{v,s}$ . The residual  $\varepsilon_{v,t}$  encompasses the difference between observation and the model.

## S.2 Mass scaling optimisation

Beside the selected solutions described in Sections 3.1.1 to 3.1.3, there are other solutions of mathematically (near) equal quality available, resulting in slightly different classifications. When deciding on clustering solutions, we also explored the  
15 highest, alternative local maxima. While a comprehensive description of the clustering solution space is out of the scope of what we can present here, we note the main differences between the solutions relate to outlier groups and how they can be combined or separated from the main classes. In phase I, the other viable solution would have been the maxima at  $k = 14$  ( $s_m = 1.88$  gives approximately the same classification as the selected solution at  $k = 17$  ( $s_m = 2.12$ ). We selected the latter solution since it yields a higher silhouette a) overall (0.5645 vs 0.5628) and b) for the AN cluster specifically (0.48 vs 0.35).  
20 For both P-I and P-II, there exist a number of unstable (in relation to minor  $s_m$  change), high  $k$  solutions. However, upon inspection, their high silhouettes (compared to lower  $k$  solutions) seem to result from division of outlier clusters ( $n = 2 \dots 3$ , with low within-cluster silhouettes), into singleton ( $n = 1$ ) clusters. We find that including solutions with such further divisions does not incorporate new information, but rather smears the picture when aiming to interpret the outlier clusters (Section 3.4).

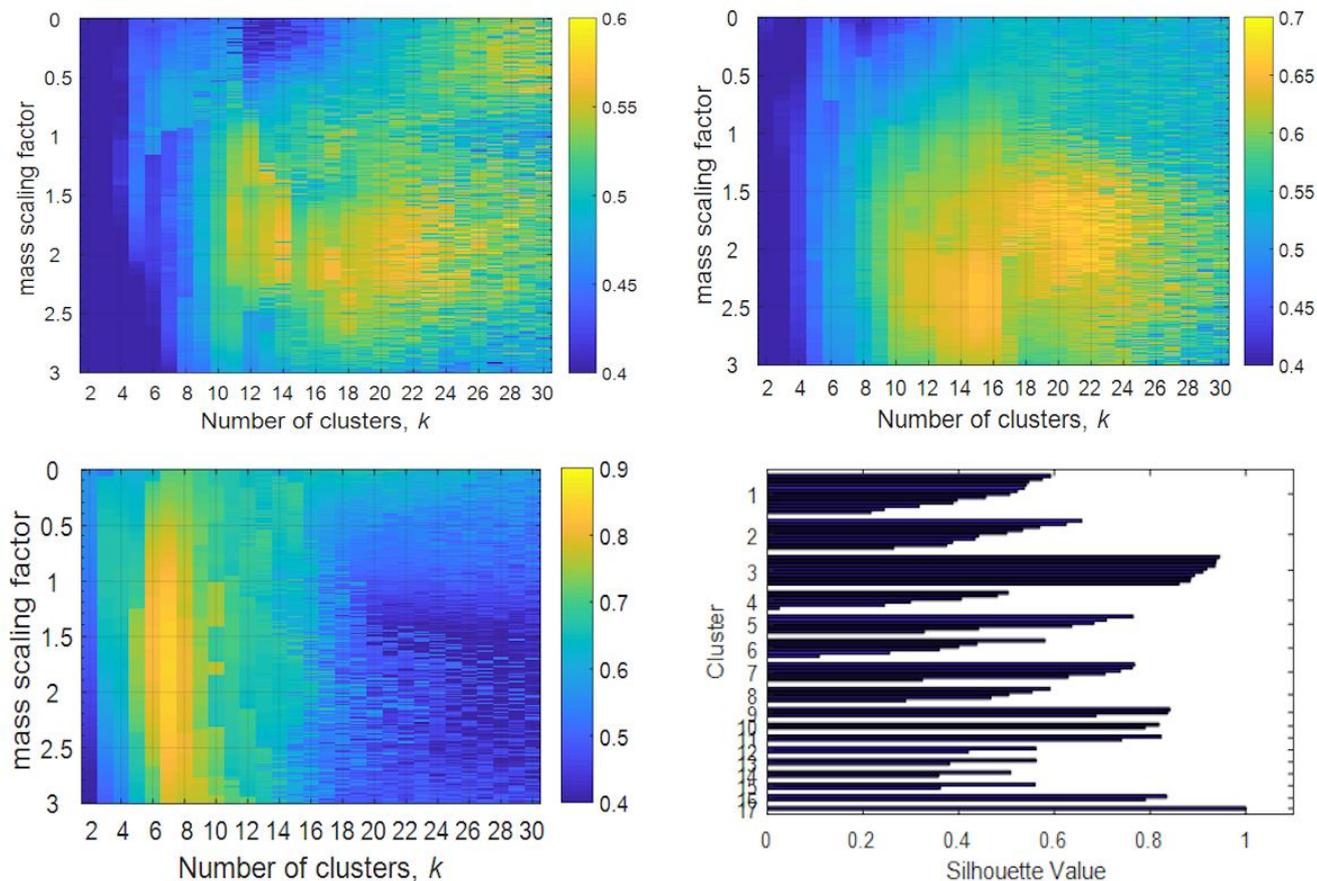


Figure S.1. Clustering silhouette scans for phases I (upper left), II (upper right) and III (lower left) panels. x-axis comprises cluster number ( $k$ ) and y-axis the mass scaling factor  $s_m$  (Equation 4). Brighter colour (higher silhouette) indicates better solution quality. Note the different colour scales in panels. Stable silhouette maxima (i.e. not sensitive to minor, 0.01 changes in  $s_m$ ) solutions were preferred. The solutions we selected were for P-I:  $k=17$ ,  $s_m=2.12$ , for P-II:  $k=15$ ,  $s_m=2.41$  and for P-III:  $k=7$ ,  $s_m=1.81$ ). In lower right panel: object specific silhouettes ( $s_i$ ) by cluster for the selected P-I solution.

### S.3 Ion pairing scheme

In the pairing schemes of Gysel et al. (2007) and Hong et al. (2017), nitrate, sulphate or ammonium are not allowed to exist separately from inorganic salts (e.g. organonitrate, organic sulphate or organic acid salt with ammonium). Since there is evidence of especially organonitrates representing a considerable fraction of total particulate nitrate (e.g. Kiendler-Scharr et al., 2016), we modify the Hong et al. scheme to allow organic nitrate and “excess” ammonium (i.e.  $\text{NH}_4$  existing with other species, such as organics). Organic sulphate was not included since it is not possible to differentiate between the two possible forms of “excess sulphur” (from the perspective of inorganic salt balance): sulphuric acid and organosulphate. To maintain compatibility with previous schemes, a sulphuric acid class was included. Any potential organosulphates would also fall into the sulphuric acid class.

Our ion balance scheme, modified from Hong et al., (2017), is divided into three cases:

**Case 1:** Aerosol is acidic, not enough  $[NH_4^+]$  to convert all sulphate to  $NH_4HSO_4$ . Sulphate exists as sulphuric acid and ammonium bisulphate. All nitrate is organic, as sulphate formation is assumed preferential. All  $[NH_4^+]$  goes to ammonium bisulphate.

5

$$\frac{[NH_4^+]}{[SO_4^{2-}]} < 1$$

sulphuric acid:  $[H_2SO_4] = [SO_4^{2-}] - [NH_4^+]$  (S.2)

ammonium bisulphate:  $[NH_4HSO_4] = [NH_4^+]$

ammonium sulphate:  $[(NH_2)SO_4] = 0$

10 ammonium nitrate:  $[NH_4NO_3] = 0$

organonitrate:  $[org\ NO_3] = [NO_3^-]$

excess ammonium:  $[\text{excess } NH_4] = 0$

**Case 2:** Aerosol is partly neutralised, sulphate exists as ammonium bisulphate and ammonium sulphate. Any nitrate is organic. All of  $[NH_4^+]$

15 associates with sulphate.

$$1 \leq \frac{[NH_4^+]}{[SO_4^{2-}]} < 2$$

sulphuric acid:  $[H_2SO_4] = 0$  (S.3)

ammonium bisulphate:  $[NH_4HSO_4] = \left(2 - \frac{[NH_4^+]}{[SO_4^{2-}]}\right) * [SO_4^{2-}]$

20 ammonium sulphate:  $[(NH_4)SO_4] = \left(\frac{[NH_4^+]}{[SO_4^{2-}]} - 1\right) * [SO_4^{2-}]$

ammonium nitrate:  $[NH_4NO_3] = 0$

organonitrate:  $[org\ NO_3] = [NO_3^-]$

excess ammonium:  $[\text{excess } NH_4] = 0$

25 **Case 3:** Aerosol is fully neutralised, sulphate exists as ammonium sulphate. Leftover  $[NH_4^+]$  from sulphate neutralisation goes with  $[NO_3^-]$ .

In case all of  $[NO_3^-]$  can be neutralised by  $[NH_4^+]$ , leftover  $[NH_4^+]$  is labelled “excess  $NH_4^+$ ”, and assumed to bind with other species (e.g. organic acids).

$$2 < \frac{[NH_4^+]}{[SO_4^{2-}]}$$

sulphuric acid:  $[H_2SO_4] = 0$  (S.4)

30 ammonium bisulphate:  $[NH_4HSO_4] = 0$

ammonium sulphate:  $[(NH_2)SO_4] = [SO_4^{2-}]$

ammonium nitrate:  $[NH_4NO_3] = \max([NO_3^-], [NH_4^+] - 2 * [SO_4^{2-}])$

organonitrate:  $[org\ NO_3] = [NO_3^-] - [NH_4NO_3]$

excess ammonium:  $[\text{excess } NH_4] = \min(0, [NH_4^+] - 2 * [SO_4^{2-}] - [NO_3^-])$

We echo Hong et al. (2017) in emphasising that this approximation is only a first estimate of the dynamics of inorganic salt formation; the assumption of perfect internal mixing is likely unrealistic for atmospheric aerosol particles, due to the nature of atmospheric processing often affecting the particle outer surface, resulting in layered structures (e.g. Buajarnern et al., 2007). The model also does not adequately describe the thermodynamic equilibriums of salt formation or model the competition between  $[\text{SO}_4^{2-}]$  and  $[\text{NO}_3^-]$  for  $[\text{NH}_4^+]$  realistically.

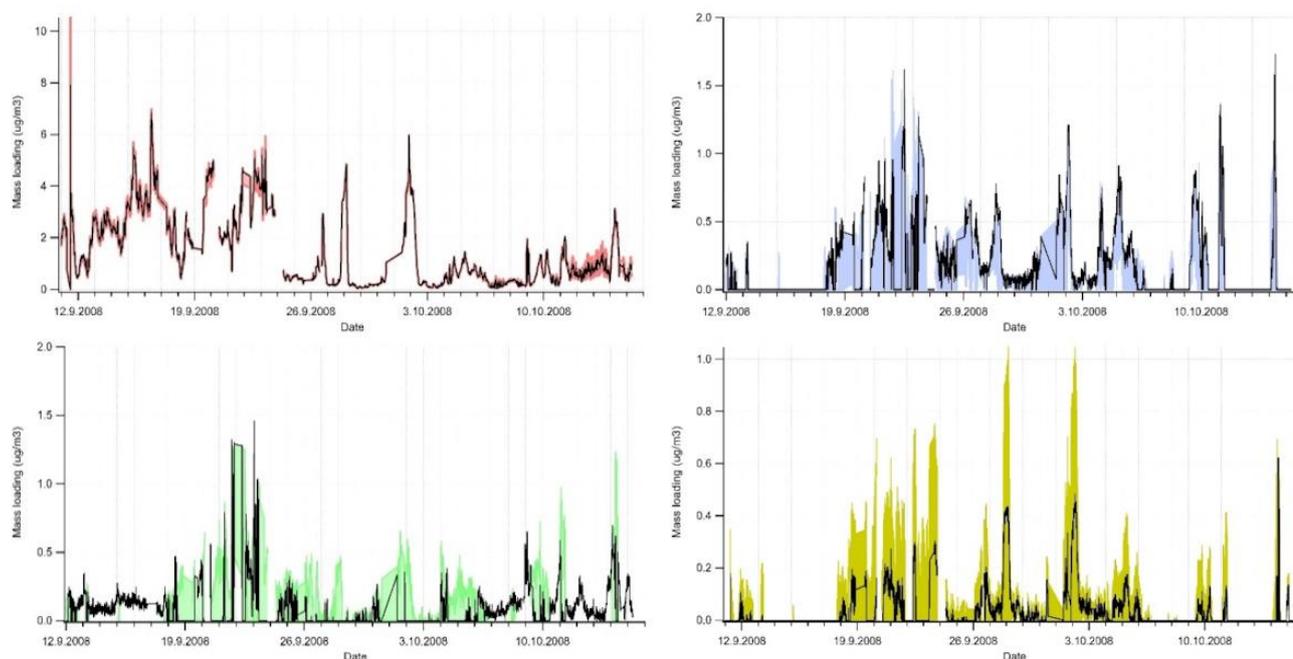
#### S.4 Relative ionisation efficiencies (RIEs) and ion pairing scheme sensitivity to $\text{RIE}_{\text{NH}_4}$ changes

To convert from nitrate equivalent mass units (used in factorisation steps until P-III), the factorisation output data needs to be corrected for species-specific Relative Ionisation Efficiency (RIE; Allan et al., 2003; Jimenez et al., 2003). We thus summed the signals from individual ion species, and corrected them for RIE. Due to high uncertainties (standard deviation 51% of mean RIE) and for some data sets unavailability of relative ionisation efficiency calibrations for  $\text{NH}_4$ , we generally did not apply RIE correction for data sets individually, but instead used the default value of  $\text{RIE}_{\text{NH}_4}$  of 4.0 for AMS instruments. Ammonium sulphate calibrations were not routinely performed, so  $\text{RIE}_{\text{SO}_4}$  data was not available – we thus used the default  $\text{RIE}_{\text{SO}_4}$  of 1.2. For September 2008 (data set II) we had IE and  $\text{RIE}_{\text{NH}_4}$  data available, but it similarly showed a very high variability between subsequent estimates for RIE. For the 12 brute-force-single-particle (BFSP; Drewnick et al., 2015) type of calibration results available, the average  $\pm$  standard deviation of RIE observed was  $3.39 \pm 1.20$  ( $\pm 35\%$ ; min 1.03, max 5.83). Especially for BFSP-type  $\text{NH}_4\text{NO}_3$  calibration, uncertainty and/or bias for  $\text{RIE}_{\text{NH}_4}$  is high, due to low ions-per-particle (IPP) rate for  $\text{NH}_4$  fragment ions, and sensitivity to threshold set for the detector (AMS best practices, Jayne, 2018). To evaluate results for our inorganic r-CMB components, we calculated ion balance ratios for AN and AS, and compared them with theoretical values derived from compound molecular weights. The theoretical mass ratio of  $\text{NH}_4$  to  $\text{NO}_3$  for  $\text{NH}_4\text{NO}_3$  is 0.29. With default  $\text{RIE}_{\text{NH}_4}$  of 4.0 the r-CMB AN factor yields a ratio of 0.46, and with  $\text{RIE}_{\text{NH}_4}$  from calibrations ( $3.39 \pm 1.20$ ), an average ratio of 0.54 (min 0.40, max 0.84). Results by Sun et al. (2012) produce a value of 0.36, so overall there remains discrepancy between the theoretical and observed ratios. The reason for it remains unclear, although one should note that the uncertainties involved in such estimates are generally very large if RIE uncertainty is high. The respective theoretical mass ratio for ammonium sulphate,  $\text{NH}_4 : \text{SO}_4$ , is between 0.18 ( $\text{NH}_4\text{HSO}_4$ ) and 0.36 ( $(\text{NH}_4)_2\text{SO}_4$ ). Our AS r-CMB component features a ratio of 0.20 for  $\text{RIE}_{\text{NH}_4} = 4.0$ , and 0.24 for  $\text{RIE}_{\text{NH}_4} = 3.39$ . The min/max estimates due to uncertainty of  $\text{RIE}_{\text{NH}_4}$  ( $\pm 1.20$ ) are (0.17...0.39). The ratio reported by Sun et al (2012) is 0.34. As RIE is not incorporated in the factorisation phase, uncertainty in RIE does not play a role in PMF/ME-2 outcome. We did briefly test the sensitivity of the inorganic apportionment scheme to changes in RIE of  $\text{NH}_4$  ion species, by running the model with  $\text{RIE}_{\text{NH}_4} \pm 33\%$  for the September 2008 data set (data set III; also presented in Section 3.2.2). The resulting model output changes are shown below, in Figure S.2 and Table S.1.

**Table S.1.  $RIE_{NH_4}$  sensitivity test results for the inorganic apportionment scheme.  $\pm$  estimate is given as  $((\max-\min) / 2) / \text{default result}$ .**

$RIE_{NH_4}$	AS ( $\mu\text{g}/\text{m}^3$ )	AN ( $\mu\text{g}/\text{m}^3$ )	orgNO <sub>3</sub> ( $\mu\text{g}/\text{m}^3$ )	excess_NH <sub>4</sub> ( $\mu\text{g}/\text{m}^3$ )
4.00 (default)	1.562	0.139	0.105	0.032
2.68 (- 33 %)	1.473	0.031	0.202	0.003
5.32 (+ 33 %)	1.620	0.186	0.063	0.108
+/-	4.7 %	55.8 %	66.2 %	164.1 %

It thus seems the inorganic model AS component is rather insensitive to 33 % change in  $RIE_{NH_4}$ . However, the nitrate components are much more sensitive (AN  $\pm$  56 % and orgNO<sub>3</sub>  $\pm$  66 %), and the sensitivity of “excess\_NH<sub>4</sub>” class to a 33 % change in  $RIE_{NH_4}$  is extremely high ( $\pm$  164 % in mass concentrations). The changes, however, seem connected mostly to differences in scaling rather than different temporal variability, so the  $RIE_{NH_4}$  sensitivity does not seem to explain the difference between the inorganic ion balance scheme and the r-CMB model result. While the relative uncertainty becomes large for AN, orgNO<sub>3</sub> and excess\_NH<sub>4</sub>, we note this may be partly due to the generally low nitrate concentrations at SMEAR II, and the overall large uncertainty (arising from other error sources, e.g. fragmentation table calculations) in measuring low concentrations of NH<sub>4</sub> typical of the site.



**Figure S.2.  $RIE_{NH_4}$  sensitivity test results for the inorganic apportionment model (Sect 2.4.1), tested for data set III. Black line indicates the result with default  $RIE_{NH_4}$  of 4.0. Coloured areas corresponds to the min/max results for when  $RIE_{NH_4}$  was changed by  $\pm 33\%$ . Components are colour-coded similar to main text Figure 10. (red: AS, blue: AN, green: orgNO<sub>3</sub>, yellow: excess NH<sub>4</sub>).**

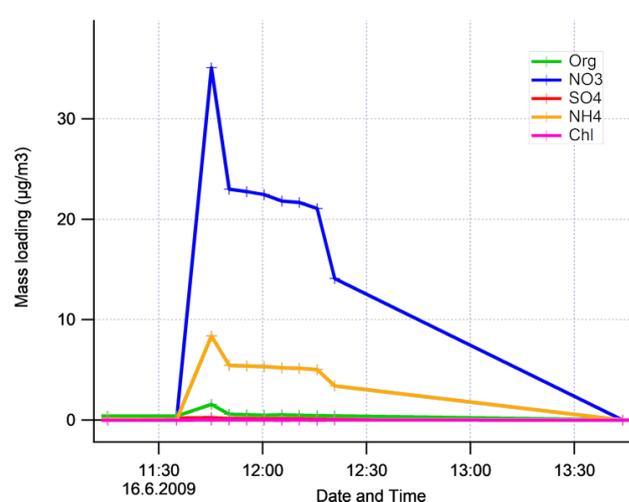
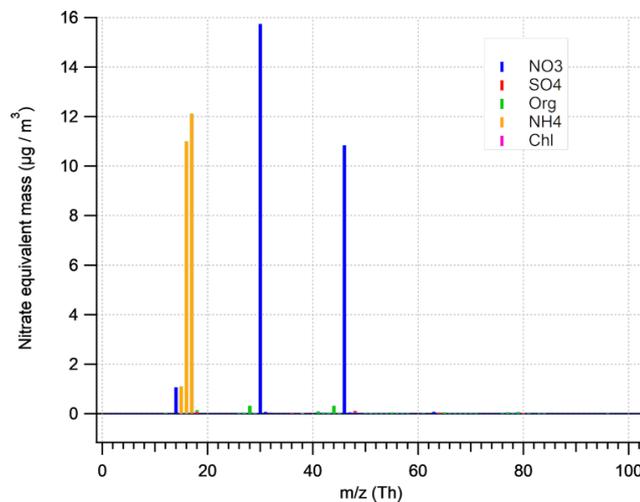
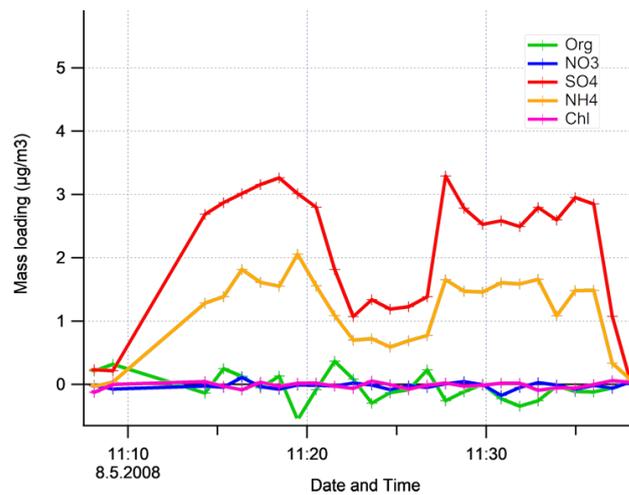
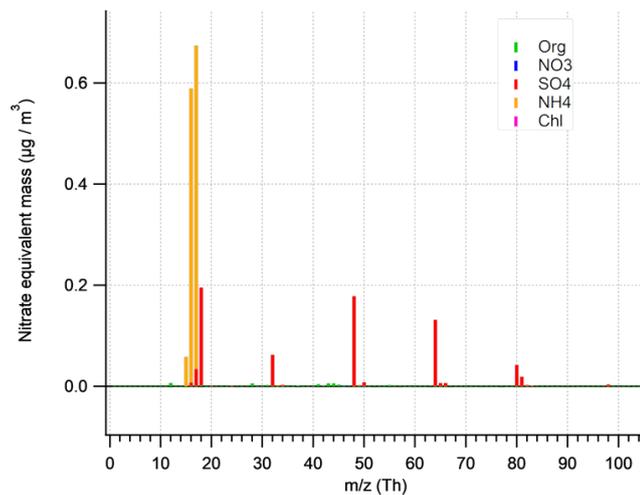
## S.5 Approximation for error propagation for signal ratios

Propagation of uncertainty can be approximated by the variance formula (Ku, 1966), which yields for a ratio of two signals ( $s_1$  and  $s_2$ ) and their uncertainties ( $\sigma_1$  and  $\sigma_2$ ) an uncertainty

$$\sigma(s_1:s_2) = \sqrt{\frac{1}{s_2^2} * \sigma_1^2 + \frac{s_1^2}{s_2^4} * \sigma_2^2}. \quad (\text{S.5})$$

- 5 In this work the above is applied in estimating the uncertainties of various signal ratios, such as  $\sigma(\text{NO}_2^+:\text{NO}^+)$ .

## S.6 Ammonium nitrate and ammonium sulphate calibrations and the $\text{CO}_2^+$ artefact



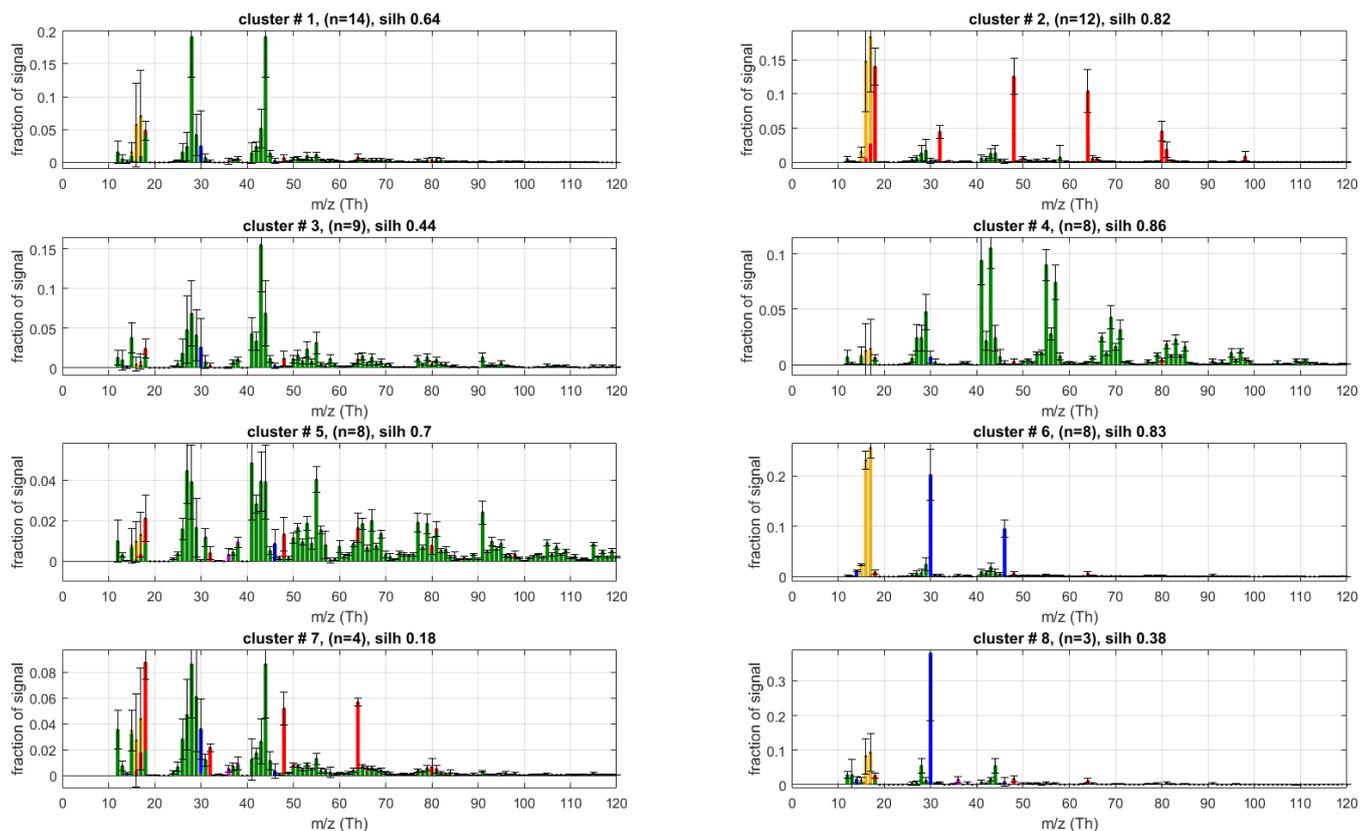
10 **Figure S.3.** AS (upper panels) and AN (lower panels) calibrations in ‘MS’ mode. Left: mass spectra, right: time series of a calibration. The data is not quality assured and was not used for (R)IE analysis.

Pieber et al (2016) discuss the  $\text{CO}_2^+$  artefact induced by thermal decomposition of inorganic salts. They find a  $m/z$  44 Th artefact with a magnitude of 3.4 % of total nitrate signal. For the phase III solution we obtain a corresponding value (for signal ratio of  $m/z$  44 Th to  $\text{NO}_3$ ) of 2.9 %, suggesting most of the  $m/z$  44 Th signal may indeed be explained by this artefact. When accounting for fragmentation-table-calculated  $m/z$  44 Th derivatives, the same applies most of total organic signal for AN. For AS (P-III), we observe a similar ratio ( $m/z$  44 Th to  $\text{SO}_4$ ) of 2.8 %, when Pieber (2016) estimate a contribution between 0.1 to 0.3 % for AS. Similar  $m/z$  44 Th artefact is also seen for the AN calibration data (Figure S.3) above. For ON, the observed signal ratio of  $m/z$  44 Th to  $\text{NO}_3$  (P-III) is 22 %. Notably, the ON solution for P-I differs from the P-III solution here, with a ratio of 8.8 %. For  $\text{KNO}_3$  Pieber et al (2016) report a ratio of 4.5 %. This also gives rise to the speculation discussed in Section 3.2.2, that the ON factor might arise from thermal decomposition of  $\text{KNO}_3$ . Furthermore, Pieber et al (2016) importantly note that thermal decomposition of  $\text{KNO}_3$  is temperature dependent, so vaporiser temperatures above 600 degrees Celsius would increase the observed ratio. While the standard operating temperature of the vaporiser (as per instrument readout) was used by default, oven temperature calibrations were not performed. Thus we do not have vaporiser power data available for the various campaigns to further examine the possibility of a bias in oven temperature readout, and if such a bias could connect to our observations of ON.

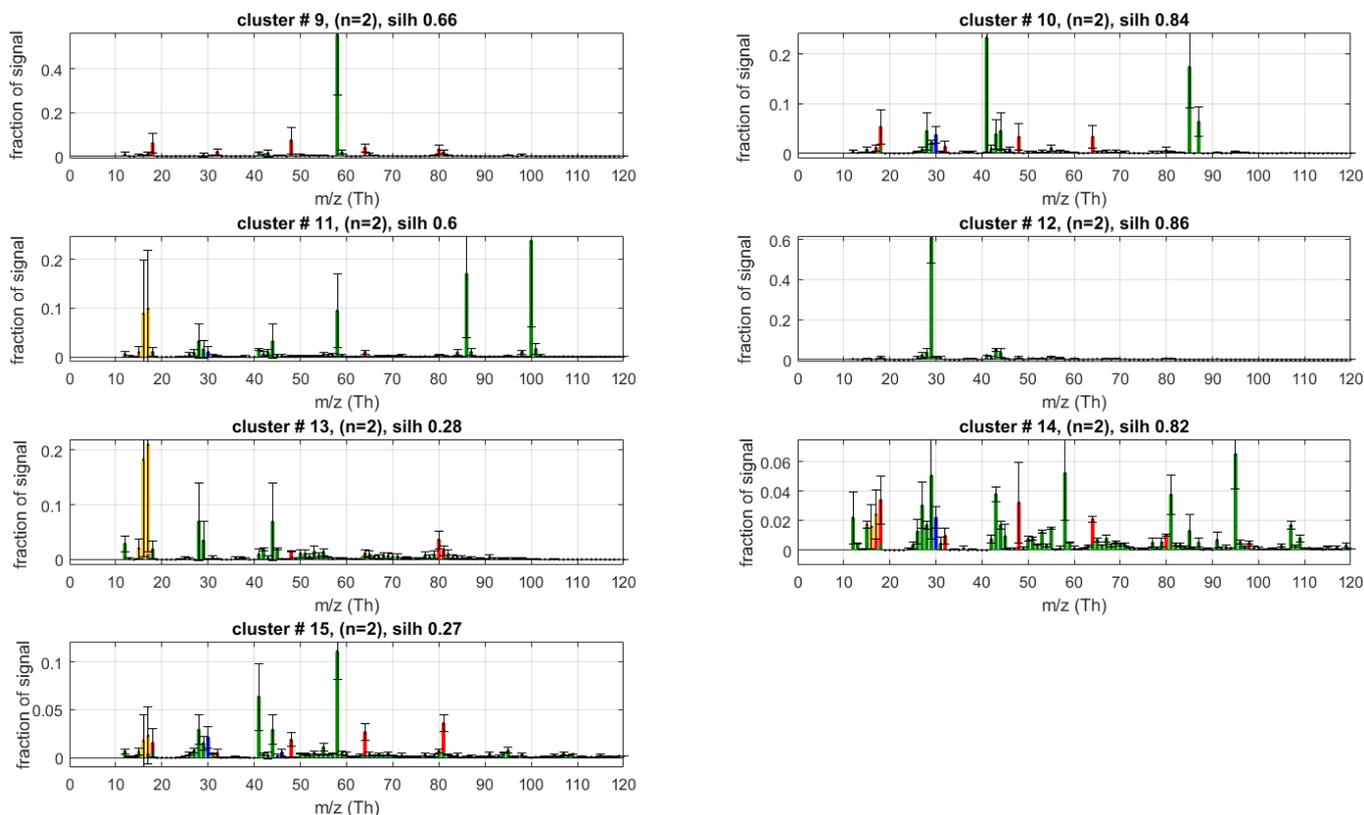
15

## S.7 Phase II clustering solution

Phase II solution was obtained for  $k = 15$ ,  $s_m = 2.41$  (silh 0.6557), and is presented below.



5 **Figure S.4 (upper panel).** P-II cluster centroids and variabilities are silhouette-weighted averages and standard deviations for the clusters 1 through 8. The main SOA types were identified as cluster #1 ('low-volatile oxidised organic aerosol', LV-OOA) clusters #3 ('semi-volatile oxidised organic aerosol', SV-OOA) and cluster #7 ('organic nitrogen', ON).



**Figure S.4. (lower panel) P-II cluster centroids and variabilities are silhouette-weighted averages and standard deviations for the clusters 9 through 15.**

## 5 S.8 CMB component inter-correlations

We calculated the mass-scaled  $r$ -CMB component profile inter-correlations to evaluate  $r$ -CMB component similarities. The result is shown in Tables S.2 (full spectrum) and S.3 (organics only). Since the scaled correlation similarity ( $r_s^2$ ) also corresponds to the quantity that the clustering algorithm aims to minimise, high values for  $r_s$  are not to be expected of a robust (hard) classification solution. From the resulting similarities, it seems the distinction between AN and ON seems to be the hardest for the algorithm (with the selected metric). Also, SV-OOA shares many similar features with the LV-OOA, BBOA and HOA.

Table S.2. Squared correlation coefficients ( $r_s^2$ ;  $s_m = 1.81$ ) for CMB component spectra inter-correlations, with both organics and inorganics included. Only correlations with p-values < 0.05 are shown.

	AS	LV-OOA	SV-OOA	BBOA	ON	HOA	AN
AS	-	-	0.01	-	0.07	-	-
LV-OOA	-	-	0.43	0.07	0.23	0.07	0.01
SV-OOA	0.01	0.43	-	0.41	0.16	0.38	0.03
BBOA	-	0.07	0.41	-	0.02	0.27	-
ON	0.07	0.23	0.16	0.02	-	0.02	0.52
HOA	-	0.07	0.38	0.27	0.02	-	-
AN	-	0.01	0.03	-	0.52	-	-

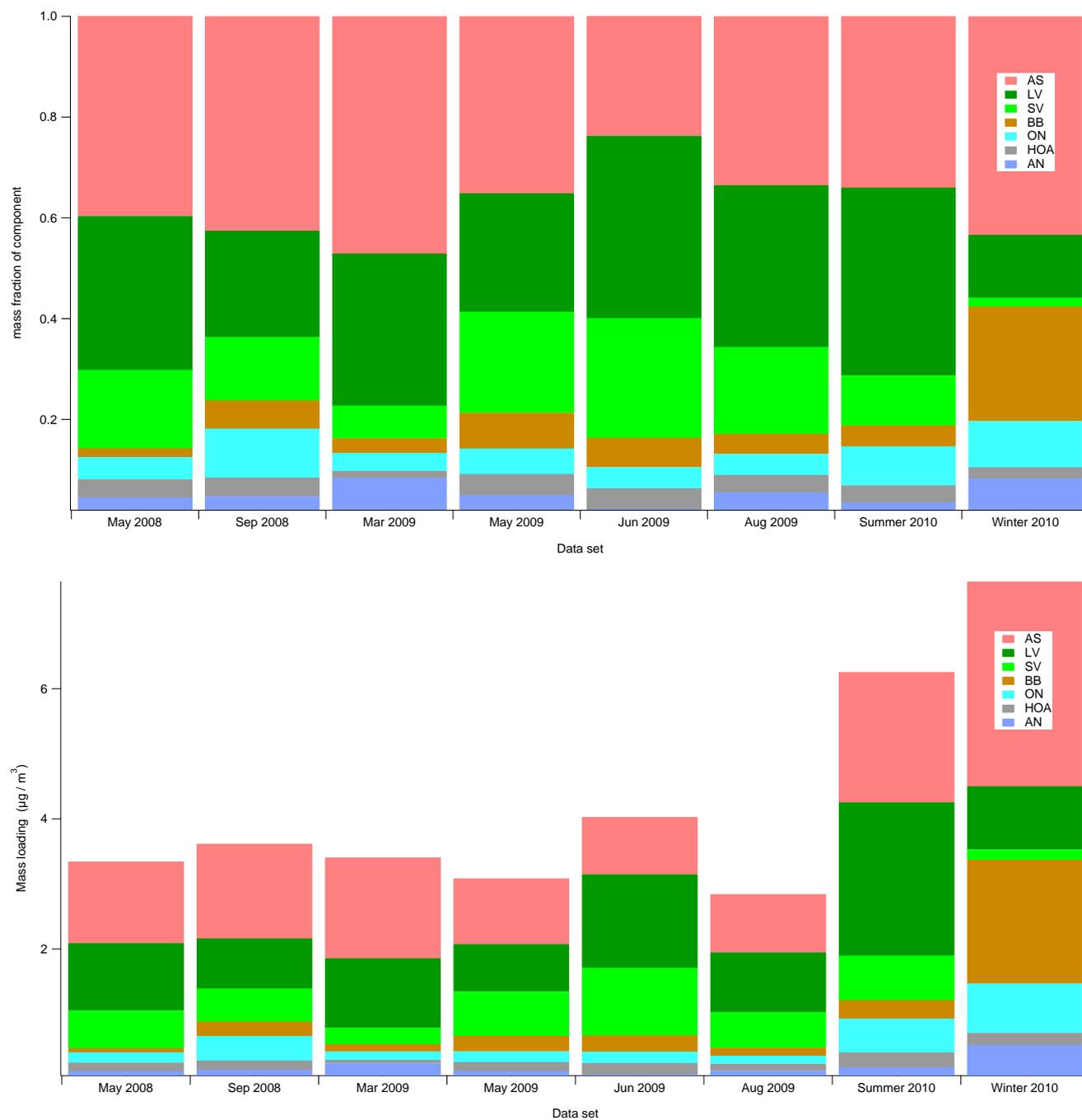
5 Table S.3. Squared correlation coefficients ( $r_s^2$ ;  $s_m = 1.81$ ) for CMB component spectra inter-correlations, for organics only. Only correlations with p-values < 0.05 are shown.

	AS	LV-OOA	SV-OOA	BBOA	ON	HOA	AN
AS	-	0.57	0.69	0.17	0.61	0.22	0.62
LV-OOA	0.57	-	0.39	-	0.90	0.04	0.30
SV-OOA	0.69	0.39	-	0.34	0.51	0.31	0.77
BBOA	0.17	-	0.34	-	0.09	0.18	0.26
ON	0.61	0.90	0.51	0.09	-	0.09	0.43
HOA	0.22	0.04	0.31	0.18	0.09	-	0.29
AN	0.62	0.30	0.77	0.26	0.43	0.29	-

Table S.3. Pearson correlation coefficients for CMB component time series inter-correlations. Only correlations with p-values < 0.05 are shown.

	AS	LV-OOA	SV-OOA	BBOA	ON	AN	HOA
AS	-	0.21	0.00	0.25	0.21	0.16	0.07
LV-OOA	0.21	-	0.24	0.02	0.09	0.10	0.22
SV-OOA	0.00	0.24	-	0.00	0.00	0.01	0.11
BBOA	0.25	0.02	0.00	-	0.48	0.25	0.09
ON	0.21	0.09	0.00	0.48	-	0.28	0.10
AN	0.16	0.10	0.01	0.25	0.28	-	0.08
HOA	0.07	0.22	0.11	0.09	0.10	0.08	-

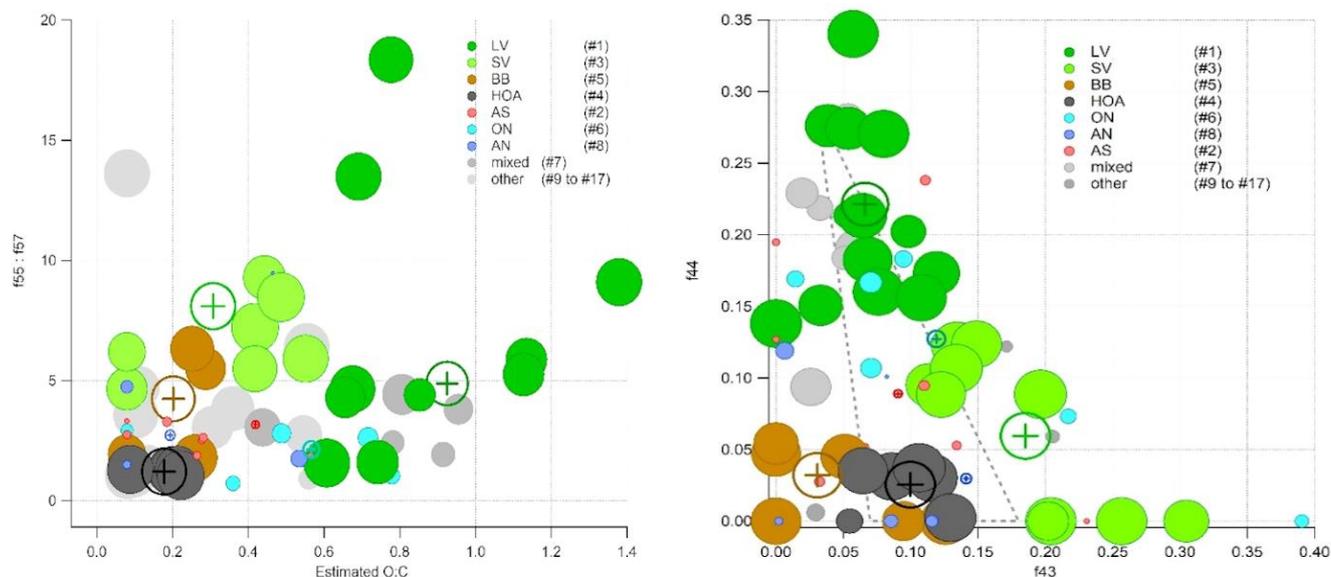
## S.9 Mass attribution by campaign



5 **Figure S.5. Mass fractions (upper panel) and absolute mass concentrations (lower panel) of r-CMB component by data set. Data set VIII (Winter 2010) results are less reliable due to pronounced surface ionisation effects.**

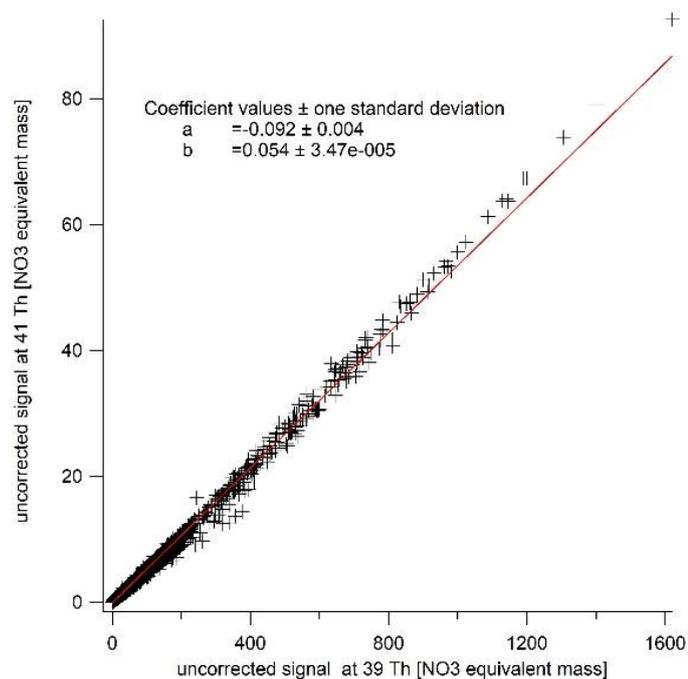
As discussed in 3.2.1, we consider results for data set VIII less reliable due to effects very likely stemming from pronounced surface ionisation processes. The high uncertainty is also reflected in phase I (PMF) high object function values ( $Q/Q_{\text{expected}}$ ; see Paatero et al., 2014). The (P-I)  $Q/Q_{\text{exp}}$  diagnostic value, optimally approaching unity, was 15.29 for data set VIII, while for the other 7 data sets the mean  $\pm$  standard deviation of  $Q/Q_{\text{exp}}$  was  $1.32 \pm 0.53$ , indicating the PMF model performance was good for data sets I to VII, but poor for set VIII. While a high object function value alone does not disqualify a result, it does indicate the uncertainties in data were likely underestimated (e.g. Norris et al., 2008), and the model does not capture the uncertainty-weighted data variation properly, even with 10 unconstrained factors.

### S.10 Cluster projections for Phase I (unconstrained PMF) result

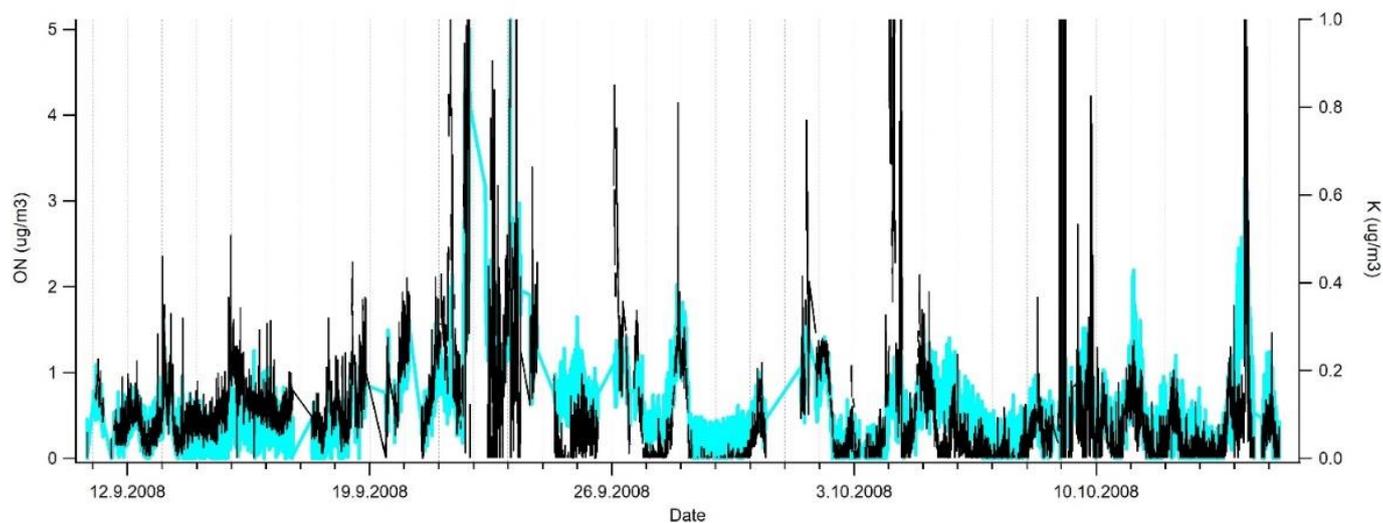


10 **Figure S.6.** (left panel) P-I solution - cluster projections onto a  $f55/f57$  (Mohr et al., 2012),  $O:C$  (estimated, Aiken et al., 2008) plane. Circles correspond to the members of the cluster and the cross markers to cluster centroids. Marker size indicates organic mass fraction in spectra. Axes are truncated. (right panel) P-I solution, projected onto  $f44, f43$  plane (i.e. the ‘Sally’s triangle’ plot; Ng et al, 2010). Circles correspond to objects in cluster and the cross markers to cluster centroids. Marker size indicates organic mass fraction in spectra. Dotted line marks the area where most laboratory data for organic aerosol falls (Ng et al., 2010).

## S.11 Potassium signal and surface ionisation effects



5 **Figure S.7.** X-axis: measured signal at  $m/z$  39 Th (mainly K-39). Y-axis: measured signal at  $m/z$  41 Th. Data is for Winter 2010 campaign (data set VIII), and contains the highest K-39 signals observed. Although corrected for in fragmentation table, the K-41 isotope correction does not fully negate the effect of potassium signal on  $m/z$  41, which may partly explain the high  $m/z$  41 Th signal observed affecting e.g. the BBOA component in data set VIII.



10 **Figure S.8.** On left axis: orgNO<sub>3</sub> mass estimate from the Kiendler-Scharr parameterisation (Eq. 7) in turquoise. On the right axis: total Potassium signal in black. Although the similar temporal behaviour is at least to some extent explained by total aerosol loading, it does raise the question if surface ionisation effects could contribute to the ON component, e.g. in form of KNO<sub>3</sub> signal.

## S.12 WD and mass spectral analysis for Rb-containing outlier factor

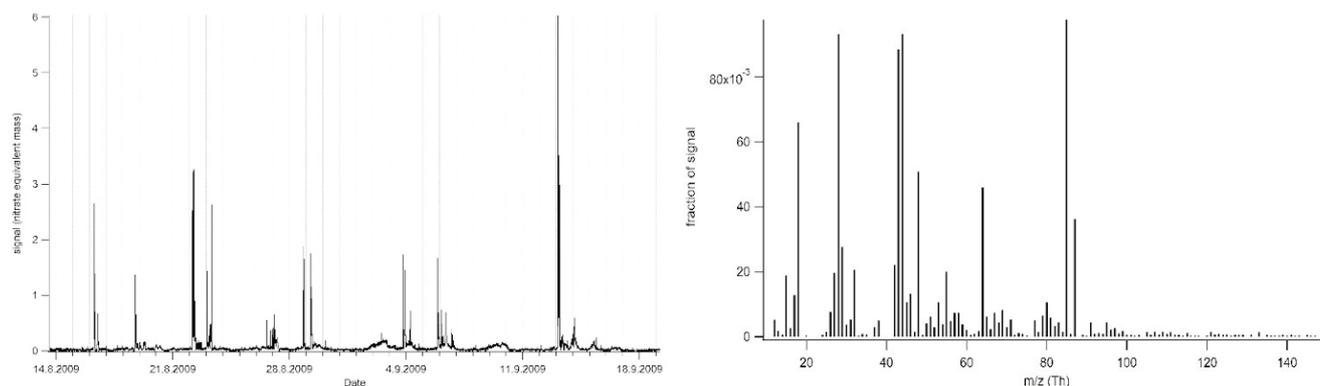
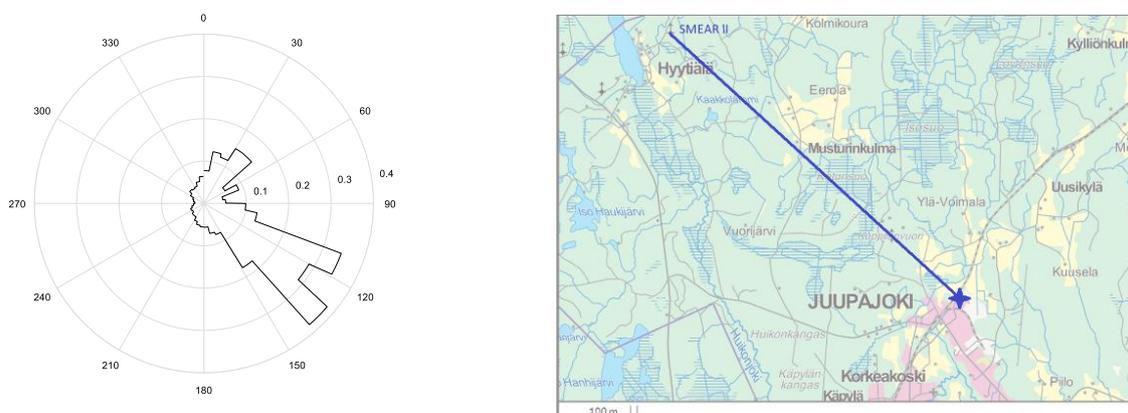
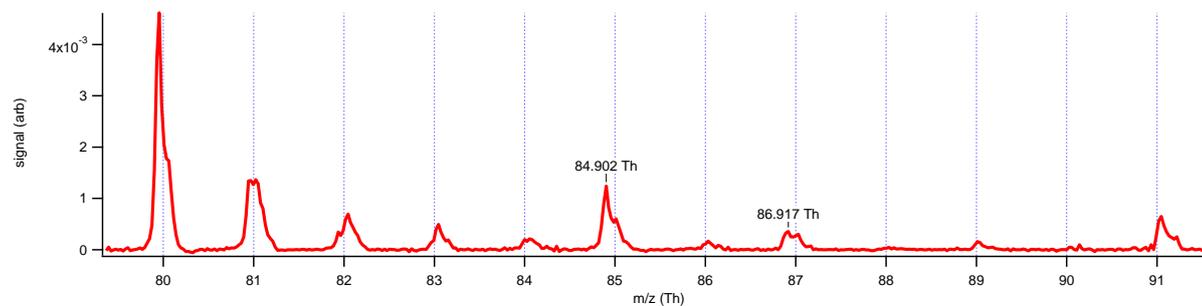


Figure S.9. Time series (left; P-I) and mass spectrum (right, P-I) of a Rb-containing factor from data set VI.



5

Figure S.10. Left: wind rose for the Rb containing factor from data set VI (P-I) average signal for the Rb-containing factor for wind sectors. Right: map showing the locations for SMEAR II and the Juupajoki district heating plant, co-located with a sawmill.



10 Figure S.11. Mass calibrated raw mass spectral signal for a plume event in Aug 14<sup>th</sup> 2010 (data set VII). While the mass resolution of the instrument ( $\sim 500$   $m/dm$ ) is not enough for proper high-resolution analysis, we can confirm existence of ions with clearly negative mass defect at  $m/z$  85 and 87 Th. The highest  $m/z$  bins (tags in figure) agree with exact masses for  $^{85}\text{Rb}^+$  (84.912 a.m.u.) and  $^{87}\text{Rb}^+$  (86.909 a.m.u.) ions.

### S.13 Diurnal cycles of r-CMB components

Daily cycles of r-CMB components are presented in Figure S.12. Due to aerosol generally being transported over long distances to the site, clear source-related peaks of e.g. rush hours for HOA are not resolved. Besides temporal behaviour of source emissions, the observed daily cycles can be connected to aerosol temperature response or boundary layer dynamics.

5

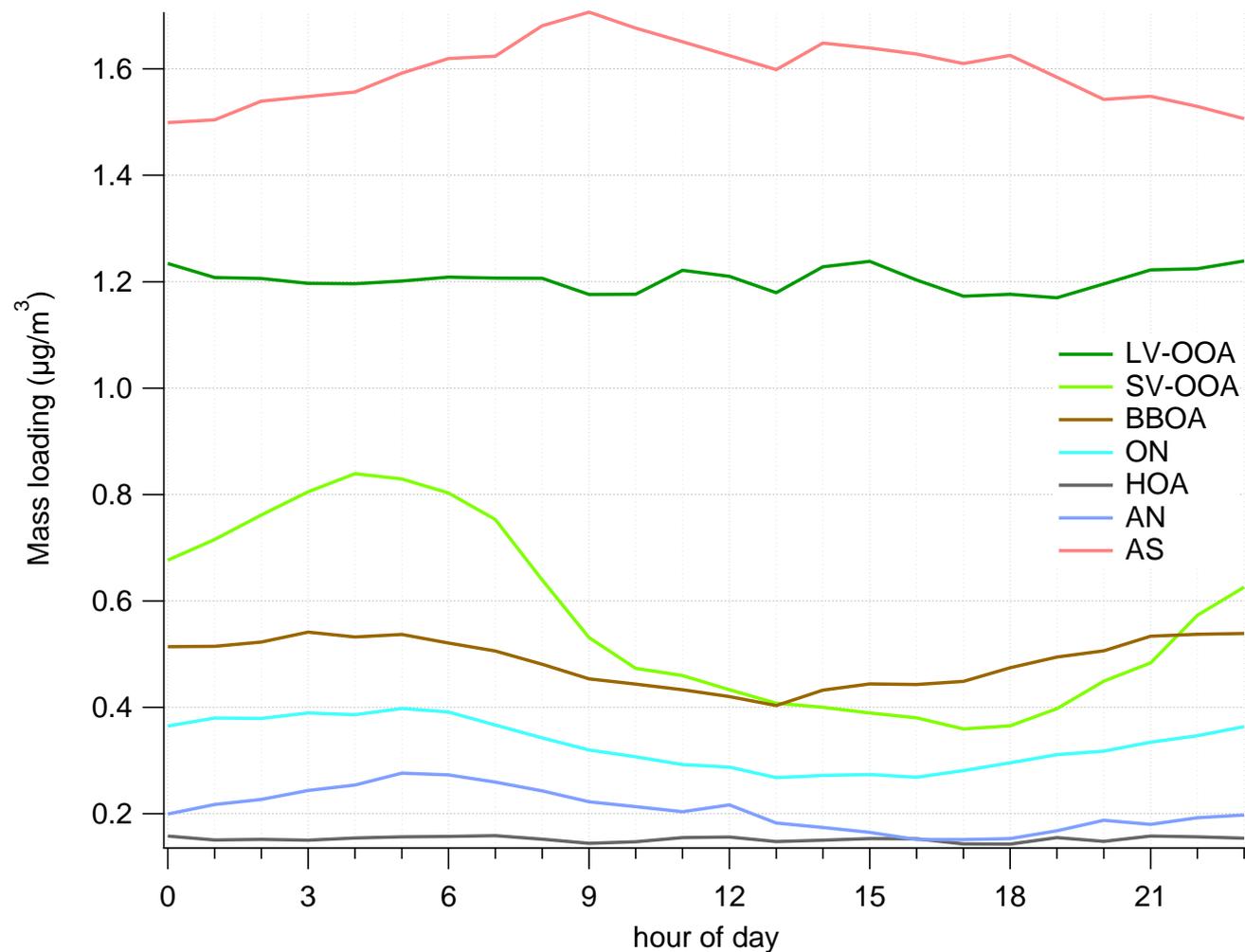
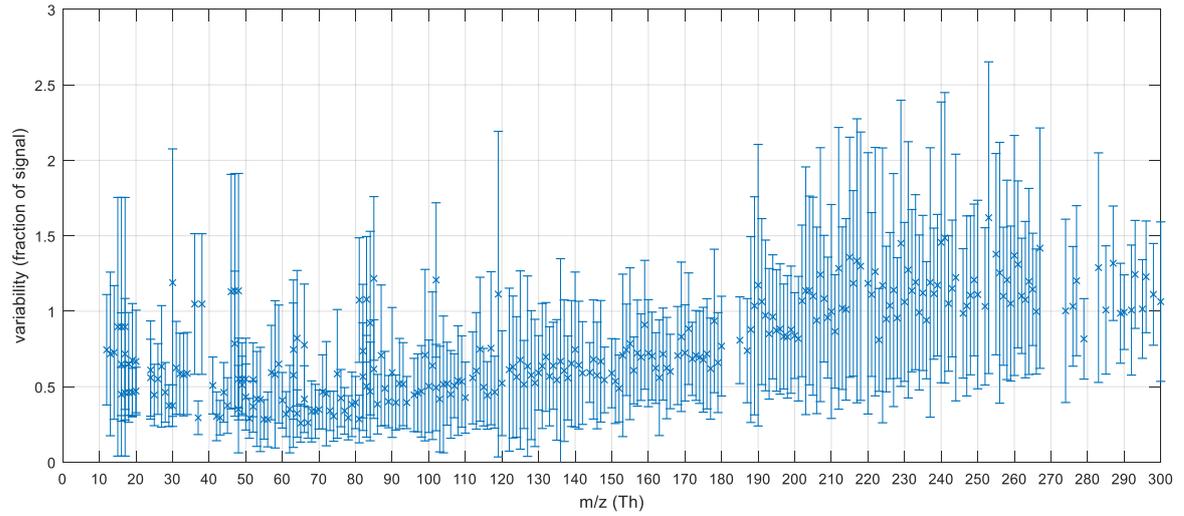


Figure S.12. Average loading of r-CMB model components by hour of day, over all of the data sets. Since the data spans all seasons, and the length of day varies considerably, the variability of average loading is generally high.

## S.14. Estimated variability for the final r-CMB components



**Figure S.13. Variability estimate as a function of  $m/z$  ratio for the final r-CMB model components. Mean (x) and standard deviation (error bar) of the variability-to-signal ratio, calculated over all the 7 P-III clusters.**

## References

- Jayne J.: AMS best practices, AMS user's meeting presentation, 2018. Available online at [http://cires1.colorado.edu/jimenez-group/UsrMtgs/UsersMtg19/Jayne\\_CalibrationOverview.pdf](http://cires1.colorado.edu/jimenez-group/UsrMtgs/UsersMtg19/Jayne_CalibrationOverview.pdf), viewed 19.Sep.2018.
- 5 Aiken, A. C., Decarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P. J., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., and Jimenez, J. L.: O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry, *Environ Sci Technol*, 42, 4478-4485, 2008.
- 10 Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *Journal of Geophysical Research: Atmospheres* (1984–2012), 108, 2003.
- Buajareern, J., Mitchem, L., and Reid, J. P.: Characterizing the formation of organic layers on the surface of inorganic/aqueous aerosols by Raman spectroscopy, *The Journal of Physical Chemistry A*, 111, 11852-11859, 2007.
- Drewnick, F., Diesch, J.-M., Faber, P., and Borrmann, S.: Aerosol mass spectrometry: particle–vaporizer interactions and their consequences for the measurements, *Atmospheric Measurement Techniques*, 8, 3811-3830, 2015.
- 15 Friedlander, S. K.: Chemical element balances and identification of air pollution sources, *Environmental Science & Technology*, 7, 235-240, 1973.
- Gordon, G. E.: Receptor models, *Environmental Science & Technology*, 22, 1132-1142, 1988.
- Gysel, M., Crosier, J., Topping, D., Whitehead, J., Bower, K., Cubison, M., Williams, P., Flynn, M., McFiggans, G., and Coe, H.: Closure study between chemical composition and hygroscopic growth of aerosol particles during TORCH2, *Atmospheric Chemistry and Physics*, 7, 6131-6144, 2007.
- 20 Hong, J., Äijälä, M., Häme, S. A., Hao, L., Duplissy, J., Heikkinen, L. M., Nie, W., Mikkilä, J., Kulmala, M., and Prisle, N. L.: Estimates of the organic aerosol volatility in a boreal forest using two independent methods, *Atmospheric Chemistry and Physics*, 17, 4387-4399, 2017.
- Hopke, P. K.: Receptor modeling for air quality management, Elsevier, 1991a.
- 25 Hopke, P. K.: An introduction to receptor modeling, *Chemometrics and Intelligent Laboratory Systems*, 10, 21-43, 1991b.
- Jimenez, J. L., Jayne, J. T., Shi, Q., Kolb, C. E., Worsnop, D. R., Yourshaw, I., Seinfeld, J. H., Flagan, R. C., Zhang, X. F., Smith, K. A., Morris, J. W., and Davidovits, P.: Ambient aerosol sampling using the Aerodyne Aerosol Mass Spectrometer, *Journal of Geophysical Research-Atmospheres*, 108, Artn 8425
- 10.1029/2001jd001213, 2003.
- 30 Kiendler-Scharr, A., Mensah, A. A., Friese, E., Topping, D., Nemitz, E., Prevot, A., Äijälä, M., Allan, J., Canonaco, F., and Canagaratna, M.: Ubiquity of organic nitrates from nighttime chemistry in the European submicron aerosol, *Geophysical Research Letters*, 43, 7735-7744, 2016.
- Ku, H. H.: Notes on the use of propagation of error formulas, *Journal of Research of the National Bureau of Standards*, 70, 1966.
- 35 Miller, M., Friedlander, S., and Hidy, G.: A chemical element balance for the Pasadena aerosol, in: *Aerosols and atmospheric chemistry*, Elsevier, 301-312, 1972.
- Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Penuelas, J., Jimenez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prevot, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, *Atmospheric Chemistry and Physics*, 12, 1649-1665, 10.5194/acp-12-1649-2012, 2012.
- 40 Ng, N. L., Canagaratna, M. R., Zhang, Q., Jimenez, J. L., Tian, J., Ulbrich, I. M., Kroll, J. H., Docherty, K. S., Chhabra, P. S., Bahreini, R., Murphy, S. M., Seinfeld, J. H., Hildebrandt, L., Donahue, N. M., DeCarlo, P. F., Lanz, V. A., Prevot, A. S. H., Dinar, E., Rudich, Y., and Worsnop, D. R.: Organic aerosol components observed in Northern Hemispheric datasets from Aerosol Mass Spectrometry, *Atmospheric Chemistry and Physics*, 10, 4625-4641, 10.5194/acp-10-4625-2010, 2010.
- 45 Norris, G., Vedantham, R., Wade, K., Brown, S., Prouty, J., and Foley, C.: EPA positive matrix factorization (PMF) 3.0 fundamentals & user guide, Prepared for the US Environmental Protection Agency, Washington, DC, by the National Exposure Research Laboratory, Research Triangle Park, 2008.
- Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, *Atmospheric Measurement Techniques*, 7, 781-797, 10.5194/amt-7-781-2014, 2014.
- 50 Pieber, S. M., El Haddad, I., Slowik, J. G., Canagaratna, M. R., Jayne, J. T., Platt, S. M., Bozzetti, C., Daellenbach, K. R., Fröhlich, R., and Vlachou, A.: Inorganic salt interference on CO<sub>2</sub>+ in aerodyne AMS and ACSM organic aerosol composition studies, *Environmental Science & Technology*, 50, 10494-10503, 2016.
- Sun, Y., Zhang, Q., Schwab, J., Yang, T., Ng, N., and Demerjian, K.: Factor analysis of combined organic and inorganic aerosol mass spectra from high resolution aerosol mass spectrometer measurements, *Atmospheric Chemistry and Physics*, 12, 8537-8551, 2012.