Quantification of primary and secondary organic aerosol sources by combined factor analysis of extractive electrospray ionisation and aerosol mass spectrometer measurements (EESI-TOF and AMS)

Yandong Tong[1], Lu Qi[1], Giulia Stefenelli[1], Dongyu S. Wang[1], Francesco Canonaco[1], Urs Baltensperger[1], André S.H. Prévôt[1], Jay G. Slowik[1,*]

[1]Laboratory of Atmospheric Chemistry, Paul Scherrer Institute (PSI), 5232 Villigen, Switzerland

*Correspondence to Jay Slowik (jay.slowik@psi.ch)

Abstract:

Source apportionment studies have struggled to quantitatively link secondary organic aerosol (SOA) to its precursor sources, due largely to instrument limitations. For example, aerosol mass spectrometers (AMS) provide quantitative measurements of the total SOA fraction, but lack the chemical resolution to resolve most SOA sources. In contrast, instruments based on soft ionisation techniques, such as extractive electrospray ionisation mass spectrometry (EESI, e.g., the EESI time of flight mass spectrometer, EESI-TOF), have demonstrated the resolution to identify specific SOA sources but provide only a semi-quantitative apportionment due to uncertainties in the dependence of instrument sensitivity on molecular identity. We address this challenge by presenting a method for positive matrix factorisation (PMF) analysis on a single dataset which includes measurements from both AMS and EESI-TOF instruments, denoted "combined PMF" (cPMF). Because each factor profile includes both AMS and EESI-TOF components, the cPMF analysis maintains the source resolution capability of the EESI-TOF, while also providing quantitative factor mass concentrations. Therefore, the bulk EESI-TOF sensitivity to each factor can also be directly determined from the analysis. We present metrics for ensuring both instruments are well-represented in the solution, a method for optionally constraining factor profile contributions for one or both instruments, and a protocol for uncertainty analysis.

As a proof of concept, the cPMF analysis was applied to summer and winter measurements in Zurich, Switzerland. Factors related to biogenic and wood burning-derived SOA are quantified, as well as POA sources such as wood burning, cigarette smoke, cooking, and traffic. The retrieved EESI-TOF factor-dependent sensitivities are consistent with both laboratory measurements of SOA from model precursors and bulk sensitivity parameterisations based on ion chemical formulae. The cPMF analysis shows that with the standalone EESI-TOF PMF, in which factor-dependent sensitivities are not accounted for, some factors are significantly under/overestimated. For example, when factor-dependent sensitivities are not considered in the winter dataset, the SOA fraction is underestimated by ~25% due to the high EESI-TOF sensitivity to components of primary biomass burning such as levoglucosan. In the summer dataset, where both SOA and total OA are dominated by monoterpene oxidation products, the uncorrected EESI-TOF underestimates the fraction of daytime SOA relative to nighttime SOA (in which organonitrates and less oxygenated $C_xH_yO_z$ molecules are enhanced). Although applied here to an AMS/EESI-TOF pairing, cPMF is suitable for the general case of a multi-instrument dataset, thereby

1   **providing a framework for exploiting semi-quantitative, high-resolution instrumentation for**
2   **quantitative source apportionment.**
3
4   **1. Introduction**
5
6   Atmospheric aerosols negatively affect visibility (Chow et al., 2002), human health (Beelen et al., 2014;
7   Laden et al., 2006; Pope et al., 2002), and urban air quality (Fenger, 1999; Mayer, 1999) on local and
8   regional scales. Aerosols also provide the largest uncertainties for global radiation balance and climate
9   change (Myhre et al., 2014; Penner et al., 2011; Forster et al., 2007; Lohmann and Feichter, 2005).
10   Therefore, to develop appropriate mitigation policies, it is of vital importance to understand aerosol
11   chemical composition, sources, and evolution. Organic aerosol (OA) is a major component of
12   atmospheric aerosol and accounts for 20 to 90 % of the submicron aerosol mass (Jimenez et al., 2009).
13   OA is typically classified as either primary organic aerosol (POA), which is directly emitted to the
14   atmosphere, or secondary organic aerosol (SOA), which is produced by atmospheric reactions of
15   emitted volatile organic compounds (VOCs). Both POA and SOA can exert serious health effects,
16   including protein and DNA damage caused by reactive oxygen species (ROS), which can be either
17   contained in the particles or induced by oxidation reactions following inhalation (Fuller et al., 2014;
18   Kelly and Fussell, 2012; Reuter et al., 2010; Li et al., 2003; Halliwell and Cross, 1994). Recent studies
19   indicate that the oxidation potential of SOA is source-dependent. Therefore, different sources likely
20   carry different health risks, highlighting the importance of OA source identification and quantification
21   (Daellenbach et al., 2020; Zhou et al., 2018). Previous studies have been relatively successful in
22   quantitatively linking POA to its sources. However, quantification of SOA sources and/or formation
23   pathways is more challenging due to 1) the chemical complexity of SOA, which can consist of
24   thousands of unique oxidation products, including highly oxygenated molecules and high molecular
25   weight organic oligomers, and 2) limitations of traditional instrumentation for characterising OA
26   chemical composition, especially the SOA fraction. Therefore, the effects of individual SOA sources
27   on health and climate remain poorly constrained.
28
29   Positive matrix factorisation (PMF) is a widely used source apportionment technique. PMF is a bilinear
30   receptor model which represents the measured mass spectral time series as a linear combination of
31   factor mass spectra and their corresponding time-dependent concentrations (Paatero and Tapper, 1994).
32   These factors may then be related to emission sources, and/or atmospheric processes, depending on
33   their chemical and temporal characteristics. PMF has been implemented in extensive online and offline
34   studies worldwide to quantify OA sources. The Aerodyne aerosol mass spectrometer (AMS) is widely
35   used in OA source apportionment studies because it provides online, quantitative measurements of non-
36   refractory PM_1 or $PM_{2.5}$ (particulate matter with aerodynamic diameter smaller than 1 or 2.5 μm,
37   respectively) chemical composition with high time resolution. Source apportionment studies using PMF
38   based on AMS data have successfully separated and quantified POA sources based on different
39   chemical signatures, e.g., hydrocarbon-like OA (HOA) (Zhao et al., 2019; Xu et al., 2019; Sun et al.,
40   2016a; Elser et al., 2016; Zhang et al., 2014; Ng et al., 2011b), cooking-related OA (COA) (Xu et al.,
41   2019; Zhao et al., 2019; Hu et al., 2016; Sun et al., 2016a; Sun et al., 2016b; Crippa et al., 2013a; Mohr
42   et al., 2012), biomass burning OA (BBOA) (Alfarra et al., 2007; Lanz et al., 2007; Sun et al., 2011),
43   and coal combustion OA (CCOA) (Elser et al., 2016; Hu et al., 2016; Sun et al., 2016a; Zhang et al.,
44   2014; Zhang et al., 2008). However, SOA is typically reported as either a single SOA factor (denoted
45   oxygenated organic aerosol, OOA), or as two factors distinguished by degree of oxygenation (i.e., less
46   oxygenated OOA, LO-OOA, and more oxygenated OOA, MO-OOA) or by volatility (i.e., semi-volatile
47   OOA, SV-OOA, and low-volatility OOA, LV-OOA) (Xu et al., 2019; Elser et al., 2016; Sun et al.,
48   2016a; Sun et al., 2013; Jimenez et al., 2009; Zhang et al., 2011; Crippa et al., 2013a) rather than in

1   terms of sources and/or formation processes. This limitation is due to the vaporisation/ionisation scheme
2   in the AMS, which causes significant thermal decomposition and ionisation-induced fragmentation
3   (Decarlo et al., 2006). The corresponding decrease in chemical resolution, particularly for the
4   multifunctional and/or highly oxygenated SOA components molecules of which SOA is comprised (e.g.,
5   multifunctional acids, peroxides, organonitrates, organosulfates, oligomers), limits the resolution of
6   SOA source apportionment.

8   The development of continuous or semi-continuous instruments with softer vaporisation/ionisation
9   schemes has provided new insights into SOA composition, and is thus of considerable interest for source
10 apportionment. Recent examples include the (semi-continuous) Filter Inlet for Gases and AEROsols
11 chemical ionisation time-of-flight mass spectrometer (FIGAERO-CIMS) (Lopez-Hilfiker et al., 2014),
12 and the (continuous) extractive electrospray ionisation time-of-flight mass spectrometer (EESI-TOF)
13 (Lopez-Hilfiker et al., 2019), which implement soft ionisation schemes at lower temperatures than the
14 AMS, thereby reducing thermal decomposition and increasing chemical resolution (i.e., providing
15 chemical formulae of molecular ions). A recent source apportionment study using FIGAERO-CIMS at
16 a rural site in the southeastern USA successfully resolved three SOA factors, characterised by isoprene-
17 derived species such as carboxylic acids from aqueous phase processes, highlighting the chemistry of
18 biogenic species (Chen et al., 2020). Source apportionment studies in Zurich using an EESI-TOF
19 identified SOA factors from monoterpene oxidation in summer (Stefenelli et al., 2019) and oxidation
20 of biomass burning emissions in winter (Qi et al., 2019). EESI-TOF measurements identified SOA
21 factors related to solid fuel combustion and aqueous-phase processes in Beijing (Tong et al., 2021) and
22 SOA factors with aromatic and biogenic origins in Delhi (Kumar et al., 2021). However, to date the
23 factor concentrations returned by PMF analyses using these instruments are not quantitative.

25 Quantification of the measurements by instruments such as EESI-TOF and CIMS is challenging,
26 because the instrument sensitivity varies strongly with molecular identity. For CIMS, the sensitivity to
27 different compounds is determined by the frequency of collisions between reagent ions and analytes,
28 the ion–molecule reaction time, and the transmission efficiency of product ions to the detector, which
29 depends on ion-molecule binding energy. Lopez-Hilfiker et al. (2016) developed methods to estimate
30 the binding energy of iodide (I^-) adduct ions of multifunctional organic compounds for species whose
31 formation is collision-limited, providing a lower limit to their mass concentrations. Another method to
32 explore the sensitivity is to measure single-compound aerosols or SOA generated from different
33 precursors simultaneously by an EESI-TOF and a scanning mobility particle sizer (SMPS) to determine
34 the mass concentration (Lopez-Hilfiker et al., 2016). Lopez-Hilfiker et al. (2019) explored EESI-TOF
35 sensitivities to selected reference compounds with different functional groups (including saccharides,
36 polyols and carboxylic acids) and bulk SOA generated from oxidation of a single precursor VOC. For
37 pure compounds, relative sensitivities vary by two orders of magnitude, with some composition-
38 dependent trends evident (e.g., increasing sensitivity of saccharides with decreasing molecular weight,
39 and high sensitivities for polyols relative to other functionalities). In addition, a trend of decreasing
40 sensitivity with decreasing molecular weight of the precursors was found for bulk SOA. While
41 calibration with standard compounds is straightforward, the quantification of individual species within
42 SOA is extremely challenging, due to its complex composition, the lack of chemical standards for most
43 molecules, and the potential for structural isomers to have significantly different sensitivities. These
44 issues were investigated recently for the EESI-TOF by generating SOA in the presence of a variable
45 seed surface area, and comparing the difference in SOA ion concentrations measured by the EESI-TOF
46 and the corresponding gas-phase concentrations measured by a Vocus proton transfer reaction-mass
47 spectrometer (Vocus-PTR-MS) (Wang et al., 2021). The observed sensitivities for different SOA
48 components produced from the oxidation of limonene, *o*-cresol, or 1,3,5-trimethylbenzene ranged from

1 10^3 to 10^5 ion s^{-1} ppb^{-1}. A regression model was developed that was able to predict the ion-by-ion
2 sensitivities to within a factor of 5 of the experimental value when the precursor VOC is known *a priori*.
3 However, the study also showed significantly different sensitivities (up to a factor of 20) for structural
4 isomers derived from different VOC precursors. Similar isomer sensitivity differences for I$^-$-CIMS was
5 also reported by (Bi et al., 2021). The fact that these isomers cannot be distinguished by 1-dimensional
6 mass spectrometry, represents a fundamental limitation of calibration/parameterisation-based
7 quantification and complicates interpretation of the binding energy-based approach (Lopez-Hilfiker et
8 al., 2016), because ambient SOA may derive from unknown or complex mixtures of VOCs. Therefore,
9 for source apportionment purposes, source-based sensitivities are preferred and essential to quantify
10 SOA sources and formation processes.
11
12 Here we present a new approach for quantification of SOA sources retrieved from source apportionment.
13 This is achieved by PMF analysis of a single input matrix consisting of data from both a quantitative
14 instrument with lower chemical resolution (i.e., AMS) and an instrument with high chemical resolution
15 and a linear but molecule-dependent response (i.e., EESI-TOF). This method is based on the combined
16 PMF (cPMF) analysis previously performed on combined OA/VOC data from AMS and PTR-MS,
17 respectively (Crippa et al., 2013b; Slowik et al., 2010), but utilises a more robust metric for ensuring
18 adequate representation of both instruments in the model solution, optionally allows constraints to be
19 placed on the factor profile contributions for one or both instruments, and provides a method for
20 uncertainty analysis. The cPMF method is applied to AMS/EESI-TOF datasets collected during summer
21 and winter campaigns in Zurich, Switzerland, for which single-instrument PMF analyses were
22 previously reported (Qi et al., 2019; Stefenelli et al., 2019). The present study is the first application of
23 cPMF to a joint EESI-TOF/AMS dataset, and the first attempt at quantitative EESI-TOF-driven source
24 apportionment.
25
26 **2. Methodologies**
27   **2.1 The measurement site and field campaigns**
28 Field campaigns were conducted at the Swiss National Air Pollution Monitoring Network (NABEL)
29 station, an urban background site located in the Alte Kaserne, central Zurich (47°22' N, 8°33' E, 410 m
30 above sea level), previously described in detail (Canonaco et al., 2013; Lanz et al., 2007). The
31 measurements used in the current analysis are from 20 June to 26 June 2016 and 25 January to 4
32 February 2017. These periods are excerpted from longer campaigns, and correspond to the times during
33 which both the AMS and EESI-TOF achieved stable operation. The measurement site is located in a
34 courtyard, although influences from nearby restaurants, local minor roads, and human activities (e.g.,
35 cigarette smoking) are often observed (Lanz et al., 2007; Daellenbach et al., 2017; Stefenelli et al., 2019;
36 Qi et al., 2019; Qi et al., 2020). Gas-phase species, e.g., nitrogen dioxide (NO$_2$), nitrogen oxide (NO)
37 and sulfur dioxide (SO$_2$) and meteorological data, e.g., temperature (T), relative humidity (RH),
38 radiation, wind speed (WD) and wind direction (WD) are recorded by the monitoring station.
39
40 During the intensive campaigns, a separate trailer was deployed to house an additional suite of gas and
41 particle instrumentation. A PM$_{2.5}$ cyclone was installed ~75 cm above the trailer roof (~5 m above
42 ground) to remove coarse particles. After passing through the cyclone, the sampled air passed through
43 a stainless steel (~6 mm outer diameter, O.D.) tube to the particle instrumentation, which included a
44 high-resolution time-of-flight aerosol mass spectrometer (HR-TOF-AMS, Aerodyne Research Inc.) and
45 an extractive electrospray ionisation time-of-flight mass spectrometer (EESI-TOF) to measure the OA
46 composition, and a scanning mobility particle sizer (SMPS) to measure the particle concentration and
47 size distribution. The summer and winter campaign results, including OA source apportionment from
48 the standalone AMS and EESI-TOF datasets, were previously presented in detail (Qi et al., 2019;

Stefenelli et al., 2019). In this study, we focus on the OA source apportionment using positive matrix factorisation (PMF) on the combined dataset from AMS and EESI-TOF, collected during the two campaigns.

2.2 Instrumentation
2.2.1 High-resolution time-of-flight aerosol mass spectrometer (HR-TOF-AMS)

The AMS (Aerodyne Research, Inc.) provides fast, online, quantitative measurements of the size-resolved composition of non-refractory PM_1 ($NR-PM_1$). A detailed description of the instrument can be found elsewhere (Decarlo et al., 2006; Canagaratna et al., 2007), while operational details and data treatment are documented in Stefenelli et al. (2019) and Qi et al. (2019). Briefly, in both campaigns, the organic composition of $NR-PM_1$ was measured by AMS with a time resolution of 1 min. At the beginning and at the end of the both campaigns, the instrument was calibrated for ionisation efficiency (IE) using 400 nm NH_4NO_3 particles by the mass-based method (Canagaratna et al., 2007; Jimenez et al., 2003). The HR-TOF-AMS data was analysed using the SQUIRREL (v.1.57) and PIKA (v.1.16) software packages in IGOR Pro 6.37 (Wavemetrics, Inc., Portland, OR, USA). Before further single-instrument and cPMF analysis, a composition-dependent collection efficiency (CDCE) was implemented to correct the measured aerosol mass (Middlebrook et al., 2012). For both single-instrument PMF and cPMF analysis, the input matrices consisted of the time series of fitted OA ions from highresolution mass spectral analysis, together with their corresponding uncertainties estimated from ion counting statistics and detector variability according to Allan et al. (2003). Following Ulbrich et al. (2009), a minimum error value was applied to the error matrix.

The AMS PMF input matrices are identical to those used by Stefenelli et al. (2019) and Qi et al. (2019), with the exception that they include not only the OA ions retrieved from spectral analysis, but also NO^+ and NO_2^+. These ions are added because they represent the major products measured from organonitrate fragmentation (Farmer et al., 2010), and standalone EESI-TOF PMF suggested a significant role for organonitrates and other nitrogen-containing species during both the summer and winter campaigns (Qi et al., 2019; Stefenelli et al., 2019). Detailed descriptions of the final input matrices from AMS (e.g., number of measurements, number of ions and time resolution) in summer and in winter are presented in Table 1.

2.2.2 Extractive electrospray ionisation time-of-flight mass spectrometer (EESI-TOF)

The EESI-TOF provides online, fast, near-molecular-level measurement (i.e., chemical formulae of molecular ions) of OA composition, without thermal decomposition or ionisation-induced fragmentation. A detailed description can be found elsewhere (Lopez-Hilfiker et al., 2019) and the operational details for the summer and winter campaigns are documented in Stefenelli et al. (2019) and Qi et al. (2019), respectively. Briefly, aerosol particles were continuously sampled through a 6 mm O.D., 5 cm long multi-channel extruded carbon denuder. Particles then intersected a spray of charged droplets generated by a conventional electrospray probe and the soluble fraction was extracted into the droplets. The droplets passed through a heated stainless-steel capillary (~250 °C), wherein the electrospray solvent evaporated and ions were ejected into the mass spectrometer. Due to the short residence time (~1 ms) in the capillary, no thermal decomposition was observed. The analyte ions were detected by a high-resolution time-of-flight mass spectrometer with an atmospheric pressure interface (API-TOF) (Junninen et al., 2010). In the summer campaign, the electrospray consisted of a 1:1 water/methanol (MeOH, UHPLC-MS grade, LiChrosolv) mixture doped with 100 ppm NaI (>99 %, Sigma-Aldrich). In the winter campaign, a 1:1 water/acetonitrile mixture (> 99.9 %, Sigma-Aldrich)

1 mixture with 100 ppm NaI (99 %, Sigma-Aldrich) was utilised, which reduced background signal. In
2 both campaigns, the mass spectrometer was configured to detect positive ions. Because of NaI use,
3 analyte ions were detected almost exclusively as $[M]Na^+$ and other ionisation pathways were suppressed
4 (the only notable exception being nicotine, which was detected as $[C_{10}H_{14}N_2]H^+$). This yields a linear
5 response to mass, avoids matrix effects, and simplifies spectral interpretation (Lopez-Hilfiker et al.,
6 2019). Adducts of an analyte with acetonitrile or methanol molecule(s) may also be detected by the
7 instrument, depending on the voltage settings in the ion transfer optics (i.e., collision energy), but these
8 adducts were observed to have negligible signals with our voltage configurations in both campaigns.
9 The EESI-TOF alternates between direct sampling and sampling through a particle filter to provide a
10 measurement of instrument background (including spray).

12 Data analysis, including high-resolution peak fitting, was performed using Tofware version 2.5.7
13 (Tofwerk AG, Thun, Switzerland). Detailed data treatment processes can be found in Stefenelli et al.
14 (2019) and Qi et al. (2019). The EESI-TOF alternates between periods of direct ambient sampling (M_{amb})
15 and filter sampling (M_{bkgd}), with the filter periods interpolated to yield an estimated background
16 spectrum during ambient measurements ($M_{bkgd,est}$). The spectra corresponding to aerosol composition
17 (M_{diff}) are determined by the difference of M_{amb} and $M_{bkgd,est}$ as shown in Eq. (1a). The corresponding
18 error matrix was estimated by adding in quadrature the uncertainties of the total sampling measurement
19 $s_{amb}(i,j)$ and the filter sampling measurement $s_{bkdg,est}(i,j)$ as shown in Eq. (1b), which are in turn
20 calculated from ion counting statistics and detector variability (Allan et al., 2003):

$$M_{diff}(i.j) = M_{amb}(i,j) - M_{bkgd,est}(i,j) \tag{1a}$$

$$s_{diff}(i,j) = \sqrt{s_{amb}^2(i,j) + s_{bkgd,est}^2(i,j)} \tag{1b}$$

24 where the unit of all quantities in both equations is counts per second (cps). Ions with a mean SNR
25 smaller than 2 were removed from both matrices, because the signals of these ions were predominantly
26 caused by electrospray and/or instrumental background. Input matrix dimensions are summarised in
27 Table 1.

29 In theory, EESI-TOF signal for an ion x can be converted from ion flux (cps) to mass concentration
30 ($\mu g\ m^{-3}$), according to Eq. (2):

$$Mass_x = I_x \cdot \frac{MW_x}{EE_x + CE_x + IE_x + TE_{m/z}} \cdot \frac{1}{F} \tag{2}$$

33 where $Mass_x$ and I_x are the mass concentration (in $\mu g\ m^{-3}$), and the ion flux (cps) reaching the detector
34 for an ion x, respectively. MW_x represents the molecular weight of the measured ion (e.g., $[M]Na^+$)
35 (Lopez-Hilfiker et al., 2019; Qi et al., 2019; Stefenelli et al., 2019). EE_x, CE_x, IE_x and $TE_{m/z}$ denote
36 EESI extraction efficiency (the probability that a molecule dissolves in the spray), EESI collection
37 efficiency (the probability that the analyte-laden droplet enters the inlet capillary), ionisation efficiency
38 (the probability that an ion forms and subsequently survives declustering forces induced by evaporation
39 and electric fields), and ion transmission efficiency (the probability that a generated ion is transmitted
40 to the detector, which is independent from chemical identity but depends only on m/z), respectively. F
41 indicates the flow rate. In practice, several of these parameters are ion-dependent and remain
42 uncharacterised, and therefore conversion to mass concentration on an ion-by-ion basis cannot currently
43 be achieved (Lopez-Hilfiker et al., 2019). Instead, to facilitate comparison with bulk quantities, we
44 define an "apparent sensitivity (AS)" to describe the EESI-TOF response to a measured concentration
45 of species x, as shown in Eq. (3):

$$AS_x = \frac{MW_x}{EE_x \cdot CE_x \cdot IE_x \cdot TE_{m/z}} \cdot \frac{1}{F} = \frac{I_x}{Mass_x} \tag{3}$$

where I_x is the measured ion flux (counts per second, cps) for the ion or factor x detected by EESI-TOF, $Mass_x$ is measured mass concentration ($\mu g\ m^{-3}$) from a reference instrument for the same ion or factor x, thus the AS is in the unit of cps ($\mu g\ m^{-3}$)$^{-1}$. In this study, we calculated the apparent sensitivities for different factors from the cPMF results by utilising AMS contribution to the factor profile ($\mu g\ m^{-3}$) as $Mass_x$ and EESI-TOF contribution (cps) as I_x. Calculation of these contributions is discussed later in in Sect. 4 using these factor-dependent sensitivities.

2.2.3   Estimation of EESI-TOF sensitivities from a multi-variate model

For comparison to the factor-dependent sensitivities determined by the cPMF analysis (see Sect. 3.3), we also estimated sensitivities for SOA factors from molecular formulae of individual analyte ions using parameterisations developed from laboratory measurements of SOA generated from oxidation of limonene (LMN) by ozone and o-cresol (cresol) and 1,3,5-trimethylbenzene (TMB) by OH radicals (Wang et al., 2021). As discussed in Sect. 1, the parameterisation can predict the relative sensitivities of ions measured by the EESI-TOF to within a factor of 5, provided that the SOA is derived from a single, known VOC. However, for ambient data, SOA derives from multiple precursor VOCs, increasing uncertainties. For example, SOA isomers generated from different precursors can differ by up to a factor of 20 in relative sensitivity (Wang et al., 2021). This represents a significant source of uncertainty for calibration/parameterisation-based approaches for quantifying SOA factors from source apportionment, but is nonetheless a useful point of comparison.

In the present study, we utilise two well-performing models from Wang et al. (2021), namely the gradient boosting regression and linear ridge regression models, denoted GBR and LRR, respectively, developed in scikit-learn packages in Spyder 4.1.4 and Python 3.8.3. The SOA parameterisation derived from LMN was used to predict the sensitivities for summer SOAs (which are predominantly terpene-derived SOAs), and SOA systems derived from cresol and TMB were used to predict the sensitivities for winter SOAs (which are characterised by aromatics from biomass burning activities). The regression models provide compound-dependent relative sensitivities (AS_x) based only on molecular formulae. Then, the EESI-TOF signals for factor are calculated as a signal-weighted average from the respective factor profiles, as shown in Eq. (4):

$$AS_{factor} = \frac{\sum_x I_x}{\sum_x (I_x/AS_x)} \tag{4}$$

Here I_x denotes the contribution to the factor profile of each ion x. Because the model parameterisations are based on laboratory SOA that contained only the CHO group, while the resolved OA sources in this study include both CHO and CHON, we approximate the total factor sensitivity by assuming the average EESI-TOF sensitivity to CHON ions is equal to the average sensitivity of CHO ions (on a factor-by-factor basis). Note that the ions from the CHO group contribute a major fraction in SOA mass for each factor, comprising 85.2 %, 78.1 %, 57.3 % and 76.3 % for DaySOA1, DaySOA2, NightSOA1 and NightSOA2 for summer and 77.9 % and 75.0 % to SOA1 and SOA2 for winter, reducing the uncertainties introduced by this assumption (these factors will be discussed in Sect. 3.2).

2.3 Source Apportionment Method

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

1   In this paper, source apportionment was performed using the positive matrix factorisation (PMF) model
2   on a single dataset containing both AMS and EESI-TOF data. We denote the overall method governing
3   analysis of such a merged dataset as "combined PMF" (cPMF), while "PMF" denotes both the general
4   PMF model and single-run executions by the Multilinear Engine solver (see Sect. 2.3.1), which are
5   identical for PMF and cPMF. This section presents an overview of the cPMF method, with detailed
6   descriptions of each step in the referenced sub-sections. Section 3.1 then presents its application to the
7   test datasets, including dataset-specific decisions (e.g., which factors to constrain, criteria for
8   accepting/rejecting solutions) required during certain steps. The overall procedure is outlined in Fig. 1,
9   with the main steps as follows:
10
11   1) Conventional PMF analyses are conducted on the standalone EESI-TOF and AMS datasets
12   with synchronised time resolution, including constraints on factor profiles as necessary.
13   Residuals from the optimised solutions are used as a reference to retrieve a balanced solution
14   (step 4).
15   2) The EESI-TOF and AMS datasets with synchronised time resolution are combined into a single
16   input matrix. This input matrix contains OA spectra from EESI-TOF and AMS, as well as the
17   NO^+ and NO_2^+ ions measured by the AMS due to the contributions of organonitrates to these
18   ions (Sect. 2.3.1).
19   3) For any factors that are to be constrained, joint AMS/EESI-TOF profiles are constructed (Sect.
20   2.3.2 and 3.1.2).
21   4) An exploratory PMF analysis is conducted on the joint AMS/EESI-TOF matrix. This consists
22   of a 2-D exploration of the solution space defined by the number of factors (p) and relative
23   instrument weight (C) (Sect. 2.3.3). The instrument weight ensures that both instruments are
24   well-represented in the solution and is assessed by comparing residuals from cPMF and
25   standalone PMF. For computational efficiency, the profiles of all constrained factors are not
26   allowed to deviate from their reference profiles. Solutions in which both instruments receive
27   approximately equal weight are evaluated for environmental interpretability, with the most
28   interpretable solution utilised as the base case for further analysis. Note that the base case is
29   fully defined by C, p, and the set of constrained factor profiles.
30   5) From the selected base case, 1000 PMF runs are conducted, which combine bootstrap analysis
31   with random selection of a-values (i.e., tightness of constraint) for the constrained factors
32   within predetermined limits that are defined on a factor-by-factor basis (Sect. 2.3.4). This
33   requires the following as prerequisites:
34   a. Definition of dataset-specific criteria for acceptance/rejection of individual runs (Sect.
35   3.1.4).
36   b. Determination of the a-value range on a factor-by-factor basis giving a reasonable
37   acceptance probability, i.e., sufficient rejection rate to ensure adequate exploration
38   while maintaining computational efficiency (Sect. 3.1.4).
39   6) The final cPMF result is taken as the mean of all accepted solutions from the bootstrap/a-value
40   analysis (step 5), with uncertainties represented by the standard deviation. From this mean
41   solution, quantitative time series and EESI-TOF factor-specific sensitivities are calculated.

Atmospheric
Measurement
Techniques
Discussions
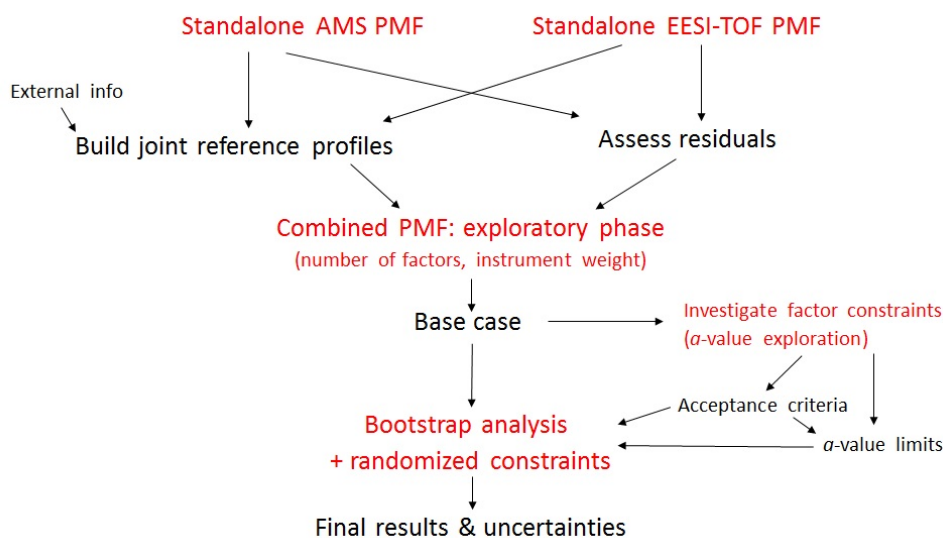
1
2 Figure 1. Flow chart summary of cPMF analysis workflow. Red text denotes PMF model operations
3 while black text denotes inputs, outputs, and/or analysis decisions.
4

### 2.3.1   Positive matrix factorisation (PMF)

6   Positive matrix factorisation (PMF) is implemented using the Multilinear Engine (ME-2) (Paatero,
7   1999), with model configuration and post-analysis performed with the Source Finder (SoFi, version 6B)
8   (Canonaco et al., 2013), programmed in Igor Pro 6.39 (Wavemetrics, Inc.). PMF is a bilinear receptor
9   model, which operates on an input data matrix $\mathbf{X}$ (here the mass spectral time series collected by EESI-
10 TOF and/or AMS) and uncertainty matrix $\mathbf{S}$, which corresponds point-by-point to $\mathbf{X}$. PMF describes $\mathbf{X}$
11 as a linear combination of static factor profiles (in this case characteristic mass spectra, representing
12 specific sources and/or atmospheric processes) and their corresponding time-dependent source
13 contributions, as described in Eq. (5):

14 $$\mathbf{X} = \mathbf{G} \times \mathbf{F} + \mathbf{E} \qquad (5)$$

15

16 Here $\mathbf{X}$ has dimensions of $m \times n$, representing m measurements of n variables (here ions), $\mathbf{G}$ and $\mathbf{F}$ are
17 respectively the factor time series with the dimension of $m \times p$, and factor profiles with the dimension of
18 $p \times n$, where p is the number of factors in the PMF solution, and is determined by the user. $\mathbf{E}$ is the
19 residual matrix and defined by Eq. (5). Figure 2 shows a conceptual representation of the combined
20 EESI-TOF and AMS input data matrix $\mathbf{X}$. The corresponding uncertainty matrix $\mathbf{S}$ and residual matrix
21 $\mathbf{E}$ are constructed in the same way (Slowik et al., 2010). Note that the AMS component of $\mathbf{X, S}$ and $\mathbf{E}$
22 is in $\mu g\ m^{-3}$, and the EESI-TOF component is in cps. Also, $\mathbf{X}$ includes not only organic ions from the
23 AMS, but also NO^+ and NO_2^+, which contain a large fraction of the AMS signal derived from
24 organonitrates (Farmer et al., 2010).

1
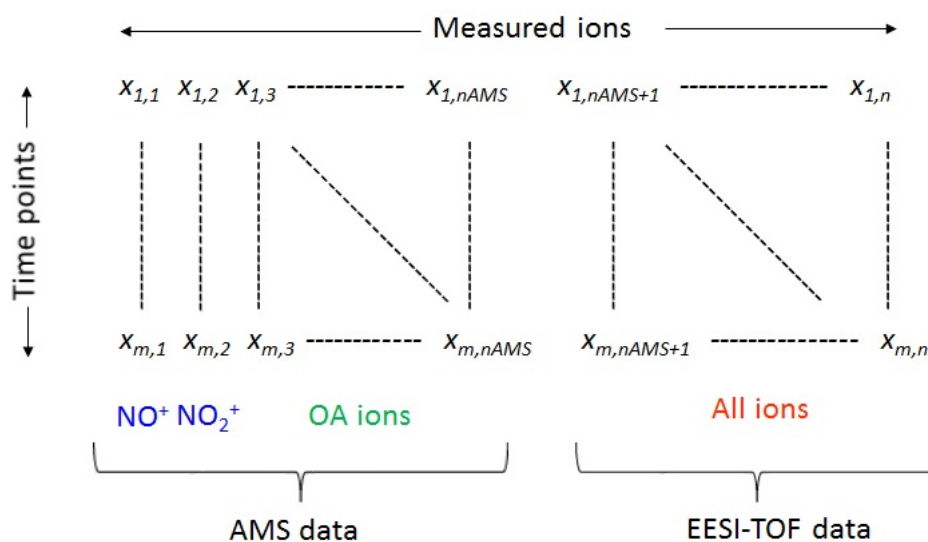
Figure 2. Schematic of the combined EESI-TOF and AMS input data matrix (**X**) for cPMF. Matrix
dimensions for the summer and winter datasets are provided in Table 1.

Equation 5 is solved by a least-squares algorithm that iteratively minimises the quantity Q, which is
defined in Eq. (6) as the sum of the squares of the uncertainty-weighted residuals:

$$Q = \sum_i \sum_j \left(\frac{e_{ij}}{s_{ij}}\right)^2 \qquad (6)$$

Here e_{ij} is an element in the residual matrix **E**, and s_{ij} is the corresponding element in the uncertainty
matrix, where i and j are the indices representing time and ion (or m/z), respectively. The theoretical Q,
denoted as Q_{expected}, is estimated by Eq. (7):

$$Q_{\text{expected}} \cong mn - p(m + n) \qquad (7)$$

where m and n denote the number of measurements (here time points) and the number of variables (ions
or m/z), respectively, and p denotes the number of factors in this PMF solution.

Due to the nature of the cPMF **X** matrix, each retrieved factor has a single time series, which can be
expressed in the concentration units of either instrument, and the factor profile contains both an AMS
and an EESI-TOF component. The factor time series for a single factor k is calculated as follows:

$$(g_{i,k})_{inst} = g_{i,k} \cdot \sum_{j=inst} f_{k,j} \qquad (8)$$

10

1   Here $\left(g_{i,k}\right)_{inst}$ refers generally to the time series in the measurement units of a given instrument, which

2   we denote $\left(g_{i,k}\right)_{AMS}$ or $\left(g_{i,k}\right)_{EESI}$, and the $j{=}inst$ formalism denotes the set of ions measured by the

3   respective instrument. For ease of interpretation, we report the instrument contribution to each factor

4   profile as the mass spectrum (in the respective instrument units) that would be obtained for a factor

5   mass concentration of 1 µg m^{-3}. This is expressed as follows, for a single factor k:

6

$$\left(f_{k,j}\right)_{inst} = \left(\frac{f_{k,j}\overline{\left(g_{i,k}\right)_{AMS}}}{g_0}\right)_{j=inst} \tag{9}$$

7

8   Here $\overline{\left(g_{i,k}\right)_{AMS}}$ denotes the mean of the factor time series in AMS units (µg m^{-3}), g_0 is a reference mass

9   concentration (chosen here as 1 µg m^{-3}), the $j{=}inst$ formulation again refers to all ions measured by a

10   given instrument. We refer to the organic fraction of AMS profile components and EESI-TOF profile

11   components as $\left(f_{k,j}\right)_{AMS}$ and $\left(f_{k,j}\right)_{EESI}$, respectively. The EESI-TOF apparent sensitivity (AS_x,

12   defined in Eq. (3)) can then be calculated for a single factor k as:

13

14

$$AS_k = \left(\frac{\overline{\left(g_{i,k}\right)_{EESI}}}{\overline{\left(g_{i,k}\right)_{AMS}}}\right)_{j=inst} \tag{10}$$

15   Evaluation of factor interpretability for PMF analysis of the data from a single instrument typically

16   includes: 1) correlation of the time series with external data; 2) comparison of factor diurnal cycles with

17   known source activity and previous measurements; 3) identification of source-specific spectral features.

18   In addition to these three points, factors from cPMF were also interpreted by considering the consistency

19   of spectral features between the AMS and EESI-TOF, e.g., factors originated from fresh biomass

20   burning activities are characterised by elevated signal from $C_2H_4O_2^+$ in the AMS spectrum and

21   levoglucosan in the EESI-TOF spectrum.

22

23   2.3.2   Constraints on factor profiles

24

25   Different combinations of the **G** and **F** matrices may result in solutions with the same or similar Q

26   (rotational ambiguity), which in practice leads to mixed or unresolvable factors. Here we explore a

27   subset of the possible PMF/cPMF solutions in which one or more factor profiles are constrained using

28   the a-value approach to direct solutions towards environmentally meaningful rotations. These factors

29   are constrained using reference profiles, with the scalar a ($0{\leq}\ a\ {\leq}1$) determining the tightness of

30   constraint as follows:

31

$$\left(f_{k,j}\right)_{sol} = \left(f_{k,j}\right)_{ref} \pm a \times \left(f_{k,j}\right)_{ref} \tag{11}$$

32   Here $\left(f_{k,j}\right)_{ref}$ represents the reference profile and $\left(f_{k,j}\right)_{sol}$ the final profile returned by the model. Due

33   to the renormalisation of matrices after PMF runs, the final values in $\left(f_{k,j}\right)_{sol}$ may slightly exceed the

34   prescribed range. This approach has been shown to significantly improve the model performance

35   relative to unconstrained PMF (Crippa et al., 2014; Canonaco et al., 2013; Daellenbach et al., 2016;

36   Stefenelli et al., 2019; Qi et al., 2019).

37

38   As shown in Eqs. (8-10), the EESI-TOF factor sensitivity is intrinsic to a given factor (via its profile).

39   However, in the cPMF, it may be desirable to constrain a factor for which a single reference profile

40   incorporating both AMS and EESI-TOF mass spectra is not available. For example, a factor may be

41   detectable by only one instrument, or reference profiles may have been retrieved independently for each

1   instrument (e.g., from different studies). In such cases, the cPMF reference profile, $\left(f_{k,j}\right)_{j=all,ref}$ is

2   constructed from merged individual profiles as follows:

3

$$\frac{\left(f_{k,j}\right)_{j=all,ref}}{1\ \mu\text{g m}^{-3}} = \begin{cases} \dfrac{\left(f_{k,j}\right)_{j}}{\sum_{j}\left(f_{k,j}\right)_{j}}, & j \in AMS, ref \\[4mm] AS_{k} \cdot \dfrac{\left(f_{k,j}\right)_{j}}{\sum_{j}\left(f_{k,j}\right)_{j}}, & j \in EESI, ref \end{cases} \tag{12}$$

5

6

7   Here $\left(f_{k,j}\right)_{j}$ denotes standalone reference profiles for the AMS and EESI-TOF, respectively. Note that

8   although Eq. (12) requires an initial value of AS_k to be assumed prior to PMF execution and utilised

9   during the exploratory phase of cPMF (Sect. 2.3, step 4), selection of a non-zero a-value during

10   bootstrap analysis (Sect. 2.3, step 5) allows the final AS_k to be determined by the algorithm within the

11   designated boundaries. Therefore, only a reasonable *a priori* estimate is required. In the case that a

12   factor is undetectable by the EESI-TOF (e.g., non-oxygenated hydrocarbons comprising traffic-related

13   factors), a value of AS_k is assumed that fixes the EESI-TOF contribution near zero. In the present study,

14   we utilised $AS_k = 0.01$ cps $(\mu\text{g m}^{-3})^{-1}$ when this situation arose. For contrast, AS_k for factors detectable

15   by both instruments ranged from approximately 100 to 1000 cps $(\mu\text{g m}^{-3})^{-1}$.

16

17

18   2.3.3   Instrument weighting

19   For both factor interpretation and quantitative analysis, it is important that both instruments be well-

20   represented in any accepted PMF solution. In principle, the extent to which PMF can explain a variable

21    $x_{i,j}$ is limited by the measurement uncertainty, $s_{i,j}$; that is, the expectation value of the scaled residual

22    $(e_{i,j}/s_{i,j})$ is 1 (i.e., $Q/Q_{expect} \sim 1$). In practice, $e_{i,j}/s_{i,j}$ may be systematically above or below 1, and differ

23   between instruments, for several reasons. First, the accuracy of the error calculation may be

24   systematically different between instruments, leading to systematic differences in the effect of residuals

25   from a given instrument on Q. Second, the extent of internal correlations in the dataset may differ

26   between instruments. For example, fragmentation/thermal decomposition in the AMS can lead to

27   sequences of correlated ions (e.g., $C_nH_{2n+1}^{+}$ for alkanes). In contrast, for the EESI-TOF measurement of

28   individual molecular ions, ion-to-ion correlations depend solely on particle composition. Finally, even

29   for a case where ion-by-ion signal-to-noise and the extent of internal correlations is equal between

30   instruments, the relative number of variables (ions) included in the dataset may affect the weight due to

31   small drifts in instrument performance, modelling errors in PMF, and the prevalence of

32   transient/variable sources not fully captured by PMF. Therefore, it is important to assess the relative

33   weight of the two instruments and rebalance if necessary. We define a balanced solution as one in which

34   there are no systematic differences between quality of fit for different instruments (Crippa et al., 2013b;

35   Slowik et al., 2010). However, note that variable-to-variable differences in the $e_{i,j}/s_{i,j}$ within the dataset

36   of a single instrument are permitted (as in standalone PMF).

37

38   The instrument weighting process follows the method previously proposed by Slowik et al., (2010), in

39   which weighting is performed by applying a weighting factor C to the uncertainties and evaluated by

40   comparison of the AMS vs. EESI-TOF residuals. Here we utilise the same weighting method, but

41   propose an improved evaluation metric. Instrument weighting is performed by applying a weighting

Atmospheric
Measurement
Techniques
Discussions
EGU
Open Access

1   factor C to the components of the uncertainty matrix $\mathbf{S}$ corresponding to one of the two instruments.

2   This increases/decreases the contribution of that instrument's residuals to Q, thereby changing its

3   weight within the PMF solver. In this paper, we applied the weighting factor, denoted C_{EESI}, to the

4   columns of $\mathbf{S}$ corresponding to ions measured by the EESI-TOF, according to Eq. (13):

5

$$\begin{cases} (s'_{i,j})_{j=EESI} = \dfrac{(s_{i,j})_{j=EESI}}{C_{\text{EESI}}} \\ (s'_{i,j})_{j=AMS} = (s_{i,j})_{j=AMS} \end{cases} \tag{13}$$

6

7   Note that $C_{\text{EESI}} = 1$ is equivalent to an unweighted solution; and $C_{\text{EESI}} > 1$ means the uncertainty matrix

8   of EESI-TOF decreases, which upweights the EESI-TOF.

9

10   As noted above, a balanced solution is defined as one in which the quality of fit to a given ion (assessed

11   via scaled residuals, e_{ij}/s_{ij}) is independent of the instrument performing the measurement. In previous

12   work (Slowik et al., 2010; Crippa et al., 2013), the metric used to assess this was the mean of the

13   absolute scaled residuals. This metric assumes that the optimised solution for each individual instrument

14   yields approximately the same Q/Q_{exp}. In practice, this may vary between instruments for the reasons

15   described above. Further, this metric can be unduly influenced by a few large outliers. Therefore, we

16   employ a new approach which references the residuals from the combined dataset to those obtained

17   from the final solutions from single-instrument PMF, which having been selected as the optimal

18   representation of environmental data are assumed to likewise provide the optimised distributions of

19   single-instrument residuals. The new method is as follows:

20   1) From the result of each single instrument PMF (here AMS PMF, EESI-TOF PMF), calculate

21   the scaled residual (e_{ij}/s_{ij}) probability distribution over the entire (single instrument) dataset. Here we

22   denote the scaled residual probability distribution function in the scaled residual (e_{ij}/s_{ij}) space for EESI-

23   TOF and AMS as $P_{\text{EESI}}(e_{ij}/s_{ij})$ and $P_{\text{AMS}}(e_{ij}/s_{ij})$, respectively.

24   2) Calculate the overlap fraction F_{overlap} between the AMS and EESI-TOF scaled residual

25   probability distributions from the single instrument solutions, according to Eq. (14):

26

27

$$F_{\text{overlap}} = \int \min\left(P_{\text{EESI}}\left(\frac{e_{ij}}{s_{ij}}\right), P_{\text{AMS}}\left(\frac{e_{ij}}{s_{ij}}\right) \right) \tag{14}$$

28   where $P_{\text{EESI}}(e_{ij}/s_{ij})$ and $P_{\text{AMS}}(e_{ij}/s_{ij})$ indicates the probability of occurrence of AMS and EESI-TOF at the

29   point e_{ij}/s_{ij} in scaled residual space, respectively. Given the previously mentioned assumption that the

30   single-instrument solutions represent the optimal representation of the data for the individual

31   instruments, the F_{overlap} calculated at this step is the value that should likewise be obtained from a

32   balanced solution to the combined dataset. Therefore, we define the quantity F^*_{overlap} as the F_{overlap} of

33   the final single-instrument PMF solutions.

34   3) For the combined dataset, calculate F_{overlap} as a function of a two-dimensional exploration

35   of the space defined by weighing factor (C_{EESI}) and the number of factors (p). This exploration is

36   necessary because the scaled residuals have been empirically observed to depend not only on C but also

37   p (Crippa et al., 2013b; Slowik et al., 2010), likely because p affects the degrees of freedom in the

38   solution. We select for further analysis the set of solutions in which F_{overlap} does not greatly differ from

39   F^*_{overlap}, as given by Eq. (15):

40

41

$$\left| F_{\text{overlap}}(C,p) - F^*_{\text{overlap}} \right| < \beta \tag{15}$$

42   where the threshold of absolute difference is defined as β. Here β is a subjective parameter chosen to

43   allow a manageable number of solutions to be selected for detailed inspection. For computational

1   efficiency, if one or more factors are constrained, we choose $a = 0$ for all constrained factors at this
2   preliminary exploration stage and will explore the a-value range(s) for constraint(s) for further
3   bootstrapping analysis once the C and p are determined.
4
5   The balanced solutions satisfying Eq. (15) are then evaluated using the same metrics as in standard
6   PMF analysis to select the solution with the greatest explanatory power. This solution is used as the
7   base case for bootstrap analysis and, if one or more factors are constrained, simultaneous randomised
8   a-value trials.
9
10   2.3.4   Bootstrap/constraint sensitivity analysis on the combined dataset
11   Bootstrap analysis (Davison and Hinkley, 1997) is frequently used to characterise solution stability,
12   reproducibility and estimate uncertainties. In typical bootstrap analysis, a set of new input and error
13   matrices are created by random resampling of rows from the original input data and error matrices. The
14   resulting resampled matrices preserve the original dimensions of the input data matrix, but randomly
15   duplicate some time points while excluding others (Paatero et al., 2014). In the present analysis, we
16   combined bootstrap analysis with randomised selection of a-values for all constrained factors within
17   predetermined limits defined on a factor-by-factor basis. Since the constrained factors use reference
18   profiles constructed with an estimated AS_k (see Eq. (12)), this combined bootstrap/constraint analysis
19   allows recalculation of AS_k within PMF. As a result, the final reported solution is the average of all
20   accepted bootstrap runs, with uncertainties in factor profiles and time series taken as the standard
21   deviation.
22
23   Within this analysis, the range of a-values explored for a given factor may have a significant effect on
24   the acceptance probability. A very low acceptance probability is undesirable because it is
25   computationally inefficient, while a very high acceptance probability is also undesirable because it
26   implies the solution space is inadequately explored due to excessively restrictive a-values (Canonaco
27   et al., 2021). Therefore, we conduct pre-tests to estimate the a-value range leading to a reasonable
28   acceptance probability. This is done by a set of 2-dimensional a-value ("multi-2D") scans in which the
29   a-values of two constrained factors are varied stepwise from 0 to 1 with a step size of 0.1 (i.e., 121 runs),
30   while the a-values of other constrained factors are held at 0. The results of all multi-2D runs for a given
31   factor are combined to determine the acceptance probability as a function of a-value, and upper and
32   lower a-value boundaries are assessed. The acceptance criteria are dataset-specific and discussed later
33   (Sect. 3.1.4). When the number of constrained factors (p_{ref}) = 2, the multi-2D algorithm is equivalent to
34   an explicit exploration of all possible a-value combinations. However, for $p_{ref} > 2$, multi-2D is much
35   more computationally efficient, because it increases as $p_{ref} (p_{ref} -1)/2$, whereas the explicit method
36   increases as the factorial of p_{ref}. For the datasets used here, in which p_{ref} is 3 (summer) and 4 (winter),
37   the multi-2D approach decreases the number of runs required for a-value pre-scans by factors of ~4 and
38   ~20, respectively.
39
40   Acceptance criteria consist of both the assessment of specific features of selected factor profiles/time
41   series (see Sect. 3.1.4), as well as a general evaluation of whether the solution is qualitatively similar to
42   the base case. That is, we require that the time series of each factor from a PMF run to be unambiguously
43   related to the corresponding base case factor (Stefenelli et al., 2019; Vlachou et al., 2019; Tong et al.,
44   2021). The key steps of this method are summarised below: 1) identify a base case, which as discussed
45   above is defined by a weighting factor C, number of factors p, and set of constrained factors with the a-
46   value set to 0; 2) calculate the Spearman correlation between the time series of base case and the multi-
47   2D scans, which yields a correlation matrix with the highest correlation values on the diagonal; 3) each
48   correlation coefficient on the matrix diagonal must be by a statistically significant margin (using

1   different confidence levels from a t test) than any value on the intersecting row or column. In the current
2   study, we selected a confidence level of 0 for this base case/bootstrap correlation test, representing the
3   most permissive application of this criterion. That is, we require only that the diagonal matrix mentioned
4   above can be constructed, i.e., that there is a unique 1:1 correspondence between base case factors and
5   factors from the bootstrap/a-value analysis.

7   The final set of PMF runs consisted of 1000 bootstrap runs, conducted at a single combination of C_{EESI}
8   and p, with a-values randomly selected with a step size of 0.05 for summer and 0.1 for winter within
9   the factor-specific limits determined via the multi-2D pre-scans. The same acceptance criteria utilised
10   for the multi-2D pre-scans were also used for the bootstrap runs. As a final solution, we report the mean
11   factor profiles and time series determined from all accepted bootstrap runs, with the standard deviation
12   taken to represent the uncertainty of the analysis procedure. Although not currently implemented within
13   the analysis software used, we note that in theory it would be possible to additionally include random
14   C_{EESI} selection (within a predefined range corresponding to balanced solutions) and randomised AS_k for
15   constrained profiles (within a user-defined range) in this stage of the analysis and in calculation of the
16   final model outputs.

3. Results

22   Due to the complexity of the analysed datasets (2 seasons × 3 PMF methods), we use the following
23   convention for identifying factors: factorName$_{season,method}$, where "factorName" is the name of the factor
24   (e.g., COA for cooking-related organic aerosol), "season" denotes either the summer ("S") or winter
25   ("W") dataset, and "method" refers to PMF on standalone AMS dataset ("A"), standalone EESI-TOF
26   dataset ("E"), or combined dataset ("C"). For example, COA$_{S,C}$ stands for the cooking-related factor
27   retrieved from cPMF applied to the summer dataset.

3.1 Method validation and solution selection

3.1.1   PMF analysis of single-instrument datasets

32   Single-instrument AMS and EESI-TOF PMF analysis was previously conducted and validated for both
33   the summer and winter datasets (Qi et al., 2019; Stefenelli et al., 2019). To determine the $F^*_{overlap}$, the
34   EESI-TOF-only PMF was re-run on only the period when both AMS and EESI-TOF were operating.
35   In addition, the AMS PMF analysis was re-run on the same period, but with the NO^+ and NO_2^+ ions
36   included. As discussed above, these ions contain a large fraction of the AMS signal deriving from
37   organonitrates. For EESI-TOF-only PMF analysis in both datasets, we used the same constraints as in
38   the referenced studies, that is, cooking-influenced OA (COA$_{S,E}$) was constrained for the summer dataset
39   and cigarette-smoking OA (CSOA$_{W,E}$) was constrained for the winter dataset. For AMS-only PMF
40   analysis, the only constrained factor in the original studies was hydrocarbon-like OA during winter
41   (HOA$_{W,A}$). We additionally constrained inorganic nitrate (InorgNit) in both the summer and winter
42   datasets, by including 1) the $CO_2^+/(NO^+ + NO_2^+)$ ratio, where the CO_2^+ signal was produced by reaction
43   of nitrate on the vaporiser (Pieber et al., 2016), as well as minor organic contaminants, and 2) NO^+/NO_2^+
44   ratio. In summer, we took the mass spectrum acquired from the NH_4NO_3 calibration period during the

1   campaign to calculate the ratios in 1) and 2), whereas in winter, we constructed the reference using the

2   two ratios from the ambient measurements (2.54) during periods of high nitrate to organic ratios.

3   Fig. S1 and Fig. S2 show the results from these single-instrument AMS and EESI-TOF PMF analyses

4   for summer and winter, respectively, as well as a comparison with the factor time series from the

5   original studies. Because the results are very similar to the single-instrument studies, they are discussed

6   only briefly here. The AMS-only PMF yielded five OA factors consistent with those of Stefenelli et al.

7   (2019), namely hydrocarbon-like OA (HOA$_{S,A}$), cooking-influenced OA (COA$_{S,A}$), cigarette-smoking

8   OA (CSOA$_{S,A}$), more oxygenated OA, MO-OOA$_{S,A}$, and less oxygenated OA (LO-OOA$_{S,A}$), and

9   additionally a factor dominated by NO^+ and NO_2^+ in a ratio consistent with that of ammonium nitrate,

10   denoted InorgNit$_{S,A}$. The main difference between these results and those reported by Stefenelli et al.

11   (2019) is some exchange of signal between MO-OOA$_{S,A}$ and LO-OOA$_{S,A}$. In addition, the contribution

12   from NO^+ and NO_2^+ is not solely apportioned to InorgNit$_{S,A}$ but also to factors such as LO-OOA$_{S,A}$;

13   however, this does not affect the identity and interpretation of these factors.

14   Similarly, for the winter dataset, seven factors were resolved consistent with the OA factors determined

15   by Qi et al. (2019), namely HOA$_{W,A}$, COA$_{W,A}$, LO-OOA$_{W,A}$, MO-OOA$_{W,A}$, biomass burning OA

16   (BBOA$_{W,A}$), event-specific OA (EVENT$_{W,A}$) and nitrogen-rich OA (NitrogenOA$_{W,A}$), as well as a new

17   factor consistent with InorgNit$_{W,A}$. Apart from being apportioned to InorgNit, NO^+ and NO_2^+ were also

18   apportioned to non-InorgNit factors, indicating organonitrate content and/or imperfect attribution of

19   inorganic NO^+ and NO_2^+ to these factors. Although the NO^+ and NO_2^+ contributions in some non-

20   InorgNit factors are significant, causing some changes in the factor time series compared to those in Qi

21   et al. (2019), the main features of the spectra from other OA components (i.e., ions other than NO^+ and

22   NO_2^+) in these factors are retained.

23   As discussed in Sect. 2.3.3, scaled residual probability distributions, i.e., $P(e_{ij}/s_{ij})$, for the selected

24   single-instrument solutions were calculated and are shown in Fig. 3. As discussed in Eq. (14), this yields

25   values for $F_{overlap}^*$, which are calculated to be 0.769 in summer and 0.899 in winter.

a)

b)

1

2
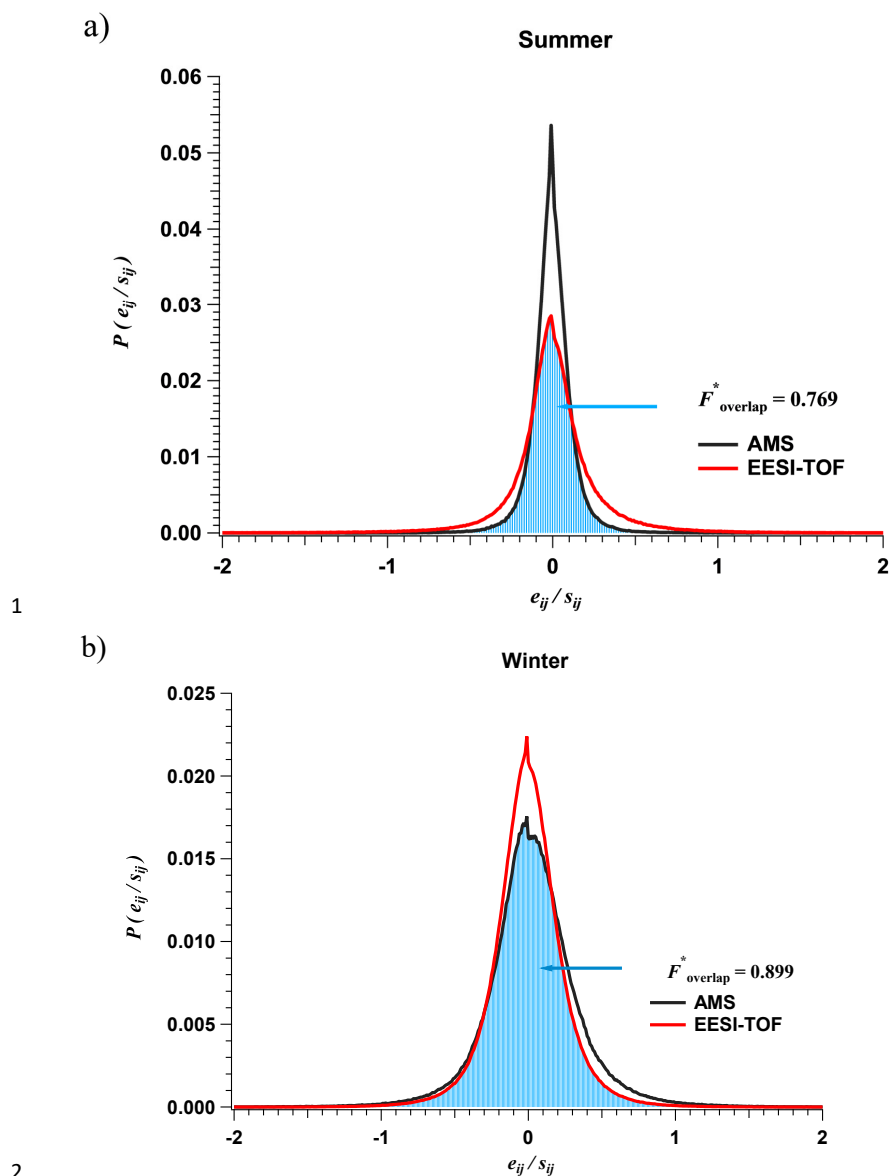
3   Figure 3. Scaled residual probability distributions and region of overlap from individual AMS PMF
4   solution and EESI-TOF PMF solutions for the summer (a) and winter (b) datasets. Red and black lines
5   show the residual distributions for the EESI-TOF and AMS, respectively; shading denotes the region
6   of overlap.

7

8

9

1   **Table. 1** Summary of parameters for the PMF analysis of re-analysed summer and winter datasets, and
2   the combined dataset. All datasets include AMS measurements of NO^+ and NO_2^+.

		EESI-TOF	AMS	Combined
Summer	Matrix dimensions (time points × m/z)	1779×507	1779×287	1779×794
	Time period	20 to 26 June 2016	20 to 26 June 2016	20 to 26 June 2016
	Time resolution (min)	5	5	5
	Range of p analysed	6	6	5-10
Winter	Matrix dimensions (time points × m/z)	6142×892	6142×258	6142×1150
	Time period	25 Jan to 4 Feb 2017	25 Jan to 4 Feb 2017	25 Jan to 4 Feb 2017
	Time resolution (min)	1	1	1
	Range of p analysed	12	8	7-14

3

4   ### 3.1.2   Construction of reference profiles
5

6   In the cPMF analysis, the factor profiles for HOA, COA, and InorgNit were constrained in both the
7   summer and winter datasets, while CSOA was constrained in winter only. All reference profiles were
8   constructed according to Eq. (12). Here we discuss the methods used to determine $\left(f_{k,j}\right)_{j=AMS,ref}$,
9    $\left(f_{k,j}\right)_{j=EESI,ref}$, and the estimated AS_k used to synthesise the reference profile. Note that COA and
10   CSOA are retrieved by both AMS and EESI-TOF, while HOA and InorgNit are not retrieved by the
11   EESI-TOF in the configuration used for these campaigns. Specifically, no HOA-sensitive EESI-TOF
12   extraction/ionisation scheme has yet been developed, while the measurable ion corresponding to
13   inorganic nitrate, $[NaNO_3]Na^+$, has been detected in other studies (Tong et al., 2021) but falls below the
14   m/z transmission window used here.

15   For summer $COA_{S,C}$, $\left(f_{k,j}\right)_{j=AMS,ref}$ and $\left(f_{k,j}\right)_{j=EESI,ref}$ were taken from the factor profiles for
16   $COA_{S,A}$ and $COA_{S,E}$, respectively. AS_{COA_S} was calculated as the ratio of the mean signals of $COA_{S,E}$
17   (cps) to $COA_{S,A}$ (µg m^{-3}). For $HOA_{S,C}$, $\left(f_{k,j}\right)_{j=AMS,ref}$ the HOA profile of Crippa et al. (2013b) was
18   used, and for $InorgNit_{S,C}$, it was taken to be the mass spectrum acquired from the NH_4NO_3 calibration
19   period during the campaign. The latter included the CO_2^+ signal produced by reaction of nitrate on the
20   vaporiser (Pieber et al., 2016), here observed with a $CO_2^+/(NO^+ + NO_2^+)$ ratio of 0.0345, as well as
21   minor organic contaminants. For both $HOA_{S,C}$ and $InorgNit_{S,C}$ all ions in $\left(f_{k,j}\right)_{j=EESI,ref}$ were set at
22   the same intensity, and AS_k was selected to be 0.01 cps (µg m^{-3})$^{-1}$.

23   The $COA_{W,C}$ reference profile was constructed using the identical method as for $COA_{S,C}$, with $COA_{W,A}$
24   and $COA_{W,E}$ as references. For $CSOA_{W,C}$, $\left(f_{k,j}\right)_{j=EESI,ref}$ was taken to be the $CSOA_{W,E}$ profile.
25   However, because the AMS did not resolve CSOA in the winter, we used the $CSOA_{S,A}$ profile for
26   $\left(f_{k,j}\right)_{j=AMS,ref}$ and estimated $AS_{CSOA,w}$ as follows:

27
$$AS_{CSOA,w} = \frac{AS_{COA,w}}{AS_{COA,s}} \cdot AS_{CSOA,s} \tag{16}$$

Atmospheric
Measurement
Techniques

Open Access

EGU

Discussions

1   where $AS_{\text{COA,s}}$, $AS_{\text{CSOA,s}}$, and $AS_{\text{COA,w}}$ are the EESI-TOF apparent sensitivities of the corresponding

2   factors, calculated assuming direct correspondence between the AMS and EESI-TOF factors sharing

3   the same name (Stefenelli et al., 2019; Qi et al., 2019).

4   The reference profile for $\text{HOA}_{\text{W,C}}$ is identical to $\text{HOA}_{\text{S,C}}$, and constructed in the same way using the

5   same profile as in the summer dataset. Unlike summer, the calibration mass spectrum of NH_4NO_3 was

6   not used as the reference profile for $\text{InorgNit}_{\text{W,C}}$, because the NO^+/NO_2^+ in the NH_4NO_3 calibration

7   period (1.58) was not consistent with that observed from ambient measurements (2.54) during periods

8   of high nitrate to organic ratios, possibly indicating contributions from non-NH_4^+ cations. Instead, the

9   InorgNit reference profile of AMS ions was constructed based on these features: 1) the NO^+/NO_2^+ ratio

10   (2.54) from 26 Jan 2016 to 31 Jan 2016, when the instrument remained stable and the ratio of nitrate to

11   OA was high, suggesting the contribution from organonitrates to NO^+ and NO_2^+ was low, 2) the

12    $CO_2^+/(NO^+ + NO_2^+)$ ratio (0.00026) was assumed to be the same as during the calibration period in the

13   Zurich winter campaign and 3) the ratio of intensity of each organic ion to CO_2^+ was kept the same as

14   during the calibration period in the Zurich winter campaign. Then $\left(f_{k,j}\right)_{j=EESI,ref}$ and $AS_{\text{InorgNit,w}}$

15   were determined using the same method as in summer.

16

17   **3.1.3   Determination of C_{EESI} and number of solutions**

18   Because F_{overlap} depends on both the weighting factor C_{EESI} and the number of factors p, an exploration

19   of this two-dimensional space is required. As discussed earlier, for computational efficiency the a-

20   values of all constrained factor profiles were set to zero during this initial exploration. For the summer

21   dataset, in which both the AMS-only and EESI-TOF-only PMF analyses yielded 6 factors, the cPMF

22   was explored from 5 to 12 factors with $\text{HOA}_{\text{S,C}}$, $\text{COA}_{\text{S,C}}$ and $\text{InorgNit}_{\text{S,C}}$ constrained. For the winter

23   dataset, in which the AMS-only and EESI-TOF-only PMF analyses yielded 8 and 11 factors,

24   respectively, the cPMF was explored from 7 to 15 factors with $\text{HOA}_{\text{W,C}}$, $\text{COA}_{\text{W,C}}$, $\text{CSOA}_{\text{W,C}}$ and

25    $\text{InorgNit}_{\text{W,C}}$ constrained. For the summer dataset, C_{EESI} was explored from 0.1 to 100, and in winter

26   from 0.001 to 50. The results of this exploration are shown in Fig. 4a and Fig. 4b, which present

27    $|F_{\text{overlap}} - F_{\text{overlap}}^*|$ as a function of C_{EESI} and p for the summer and winter datasets, respectively.

28

29   The Zurich summer dataset displays the expected trend of $|F_{\text{overlap}} - F_{\text{overlap}}^*|$ with respect to C_{EESI}.

30   Balanced solutions are found at intermediate values of C_{EESI}, with lower and higher values yielding

31   solutions in which the AMS and EESI-TOF, respectively, are overweighted. Examples of scaled

32   residual distributions for these three cases (AMS overweighted, balanced, and EESI-TOF overweighted)

33   are shown in Fig. S3. The black box in Fig. 4a denotes a set of solutions satisfying the criterion in Eq.

34   (15), which are selected for further inspection. The value of β is selected empirically to yield a practical

35   number of solutions for manual inspection, with 0.02 chosen for summer and 0.005 for winter. Factor

36   profiles and time series for solutions satisfying the β criterion, comprising solutions with 6 to 9 factors

37   (black box in the figure) are shown in Figs S4 to Fig. S13. An 8-factor solution was chosen as the best

38   representation of the data, and included $\text{HOA}_{\text{S,C}}$, $\text{COA}_{\text{S,C}}$, $\text{CSOA}_{\text{S,C}}$, $\text{InorgNit}_{\text{S,C}}$, two daytime SOAs

39   ($\text{DaySOA1}_{\text{S,C}}$ and $\text{DaySOA2}_{\text{S,C}}$) and two nighttime SOAs ($\text{NightSOA1}_{\text{S,C}}$ and $\text{NightSOA2}_{\text{S,C}}$), discussed

40   in detail in Sect. 3.2.1. Solutions with higher numbers of factors yielded uninterpretable splits in the

41   SOA or CSOA factors. Among the balanced 8-factor solutions, we selected the solution with $C_{\text{EESI}} = 2$,

42   which has the minimum value of $|F_{\text{overlap}} - F_{\text{overlap}}^*|$. This solution serves as the base case for further

43   analysis. The other 8-factor solutions exhibit time series and profiles that are similar to the selected

44   solutions. Therefore, we simply select the 8-factor solution with minimum $|F_{\text{overlap}} - F_{\text{overlap}}^*|$.

45

46   For the winter dataset, solutions with 12 or more factors are similar to the summer in which balanced

47   solutions (i.e., $\beta < 0.005$) are clustered narrowly around a single value of C_{EESI} (in this case 0.05), as

48   shown in the right black box in Fig. 4b. However, in addition, solutions with 10 to 11 factors show

1   balanced solutions over a relatively broad range, C_{EESI} = 0.001 to 0.01, as shown in the left black box
2   in Fig. 4b. This complex behaviour highlights the importance of fully exploring the two-dimensional
3   space. Solutions from the left black box (e.g., a 10-factor solution with C_{EESI} = 0.01, and 11-factor
4   solutions with C_{EESI} = 0.001, 0.005, and 0.001 which are shown in Fig. S14 to Fig S17) exhibited mixed
5   factors, in which biomass burning was not clearly separable from other sources. In contrast, the 12-
6   factor solution (see Fig. S18) and 13-factor solution (see Fig. S19) in the narrow band successfully
7   resolves these factors. The 12-factor and 13-factor solutions differ in that the 13-factor solution includes
8   uninterpretable splitting of biomass-burning-related factors. Similarly, higher-order solutions also result
9   in uninterpretable factor splitting. Therefore, the 12-factor solution with C_{EESI} of 0.05 is selected as the
10 best representation of the combined dataset.

11

12
13
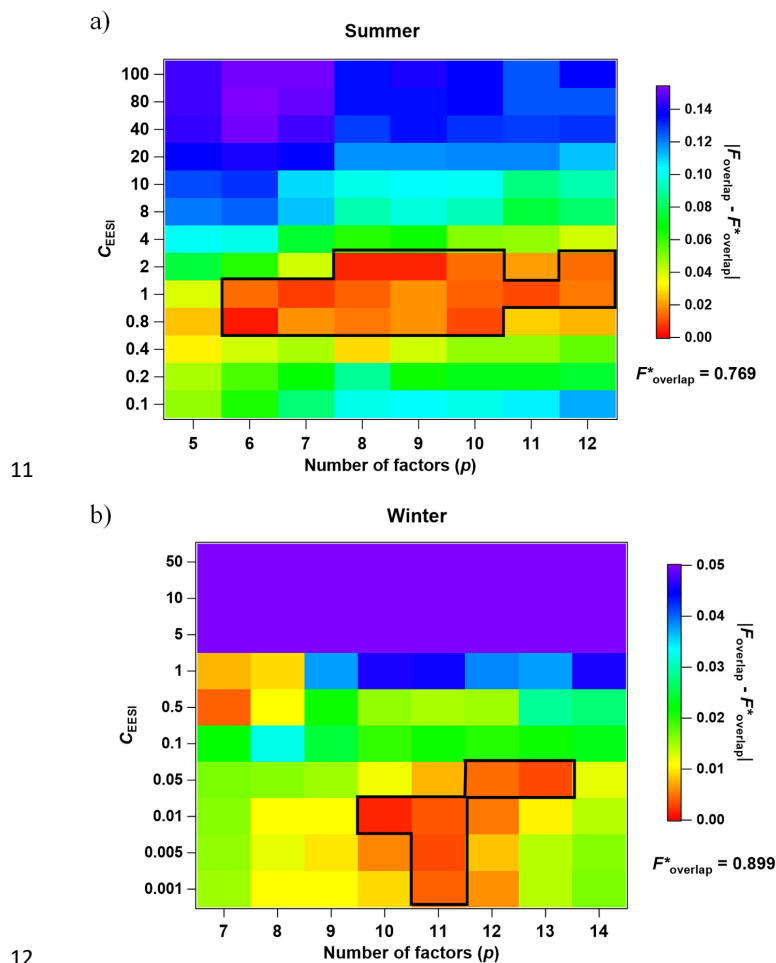
14 Figure 4. Identification of balanced solutions in the combined dataset (i.e., $| F_{overlap} - F^*_{overlap} |$ as a
15 function of C_{EESI} and p) for summer (a) and winter (b) datasets. Note that $| F_{overlap} - F^*_{overlap} |$ = 0 defines
16 a balanced solution. Solutions within the black box satisfied the $| F_{overlap} - F^*_{overlap} | < \beta$ criterion
17 defined in Eq. (15) (β is set to be 0.02 and 0.005 for summer and winter, respectively) and were selected
18 as base case candidates, from which the base case that can best represent the combined data was selected
19 by manual inspection.

3.1.4   Acceptance criteria and factor-specific a-value boundaries

As discussed in Sect. 2.3.4, the combined bootstrap/a-value randomisation analysis requires (1) a set of criteria for solution acceptance/rejection and (2) factor-specific boundaries for randomised a-value selection to maintain computational efficiency. The final set of acceptance criteria and a-value boundaries are presented in Table 2. Here we discuss their selection, which is determined synergistically by consideration of 1) unique correlations of factor time series with the base case (see Sect. 2.3.4), 2) factor-based acceptance criteria, which are here based on selected key mass spectral features (see Sect. 3.2.1 and Sect. 3.2.2 for a complete discussion of factor characteristics). Both (1) and (2) are evaluated as a function of changing a-values within the multi-2D scanning algorithm (see Sect. 2.3.4). For assessing the solution/base case correlations, we utilise a confidence level of 0, meaning that the only requirement is the ability to construct a correlation matrix with the values on the diagonal being higher than any vertical or horizontal transect. This accepts the largest possible number of solutions while requiring an unambiguous relationship between base case and bootstrapped factors. Recall that the multi-2D algorithm consists of two-dimensional a-value scans in which the a-values of constrained factors are scanned from 0 to 1 with a step size of 0.1, the a-values of other constrained factors are set to zero, and the remaining factors are left free.

Here we describe the general steps to determine acceptance criteria and a-value boundaries. A factor-based acceptance criterion is defined by the combination of a diagnostic quantity relating to one or more factors and a corresponding acceptance/rejection threshold (θ). Solutions that fulfil all criteria simultaneously are classified as accepted solutions. We calculate the acceptance probability as a function of a-value for a given factor (this is calculated independently for each factor). For a given factor, the acceptance probability is defined as the ratio of the number of accepted solutions to the total number of solutions, for which the factor has the selected a-value and the a-value of at most one other constrained factor is non-zero (that is, we consider only multi-2D runs where the factor in question is being scanned against a single other factor, while discarding runs for which the factor in question is fixed at a=0 while two other factors are scanned; this is relevant only for analyses with at least 3 constrained factors). The acceptance probability is not only a function of the a-value of the target constraint but also a function of the threshold θ. When an appropriate value of θ cannot be defined *a priori*, it is selected via sensitivity tests. The final selection of the threshold θ and a-value ranges is a compromise between (1) maintaining a reasonably high acceptance probability, thereby providing sufficient statistics without an excessive number of bootstrap runs; and (2) ensuring a sufficiently broad exploration of the solution space to encompass most environmentally reasonable solutions and thus accurately assess errors. Therefore, we determine the threshold θ and a-value upper limit for each constrained factor at which a steep drop-off from high to low probability of acceptance occurs.

For the summer dataset, three factors are constrained: $HOA_{S,C}$, $COA_{S,C}$, and $InorgNit_{S,C}$, yielding three pairs ($C(3,2) = 3$) of two-dimensional a-value scans. Two factor-based diagnostic quantities with acceptance/rejection thresholds (θ) were selected: 1) the ratio of $C_3H_3O^+$ to $C_3H_5O^+$ for $COA_{S,C}$ should be higher than the threshold $\theta_{COA_{S,C}}$ (Mohr et al., 2012), and 2) the ratio of $CO_2^+/(NO^+ + NO_2^+)$ for $InorgNit_{S,C}$ should not be higher than $\theta_{InorgNit_{S,C}}$, because the CO_2^+ signal in $InorgNit_{S,C}$ should not greatly exceed the CO_2^+ signal produced by reaction of nitrate on the vaporiser (Pieber et al., 2016); excessively high values would indicate mixing with OA. To explore the sensitivity of the acceptance probability to the threshold θ, we varied $\theta_{COA_{S,C}}$ from 4.5 to 5.1 with a step of 0.1 (note that 5.0 is the ratio of $C_3H_3O^+ / C_3H_5O^+$ in the reference profile) and $\theta_{InorgNit_{S,C}}$ from 0.034 to 0.040 with a step of 0.01, (note that 0.0345 is the ratio of $CO_2^+/(NO^+ + NO_2^+)$ in the reference profile).

1 The acceptance probability as a function of a-value and the various thresholds (θ's) for $COA_{S,C}$,
2 $InorgNit_{S,C}$, and $HOA_{S,C}$ are shown in Fig. S22. Vertical dashed lines denote the final selected a-values,
3 while the thicker traces denote the selected θ values (both of which are also given in Table 2). For
4 $\theta_{COA_{S,C}} > 5.0$, very few runs are accepted. Within the range $4.5 \leq \theta_{COA_{S,C}} \leq 5.0$, $\theta_{COA_{S,C}}$ does not affect
5 the relationship between acceptance probability and a-value for $InorgNit_{S,C}$ (Fig. S22b), but has a
6 considerable effect for $COA_{S,C}$ and $HOA_{S,C}$, with a decreasing $\theta_{COA_{S,C}}$ leading to the acceptance
7 probability remaining high at larger a-values. Visual inspection of the solutions suggests that this is due
8 to increased mixing, mostly between $COA_{S,C}$ and $HOA_{S,C}$. Therefore, we select a value of $\theta_{COA_{S,C}} = 5.0$,
9 corresponding to the $C_3H_3O_3^+/C_3H_5O_3^+$ in the factor profile. For $\theta_{InorgNit_{S,C}}$, values smaller than 0.0345
10 (i.e., reference profile) result in a very low acceptance probability, whereas choice of $\theta_{InorgNit_{S,C}}$ results
11 in similar acceptance probabilities as a function of a-value. Therefore, we select 0.0345, as the
12 acceptance probability for $\theta_{InorgNit_{S,C}}$ of 0.035 is not substantially different from 0.0345.Having
13 selected these θ values, we set a-value limits at the point where an incremental increase/decrease in a
14 yields a large change in acceptance probability (i.e. transition from high probability to low probability).
15 For the current dataset, constrained factors, and selected θ's, there is no such transition at low a-values,
16 and we therefore select only an upper limit for the a-values. For $COA_{S,C}$, there is a clear decrease for
17 both criteria between $a_{COA_{S,C}} = 0.1$ and $a_{COA_{S,C}} = 0.2$, and we therefore set the a-value boundaries as 0
18 $\leq a_{COA_{S,C}} \leq 0.2$. $InorgNit_{S,C}$ maintains an acceptance probability of ~50 % for $a_{InorgNit_{S,C}} \leq 0.4$, before
19 decreasing to <20 % at $a_{InorgNit_{S,C}} = 0.5$ and ~0 for $a_{InorgNit_{S,C}} > 0.5$; therefore the range $0 \leq$
20 $a_{InorgNit_{S,C}} \leq 0.5$ is chosen. Finally, for $HOA_{S,C}$, the acceptance probability decreases from ~55 % at
21 $a_{HOA_{S,C}} \leq 0.1$ to ~35 % at $a_{HOA_{S,C}} \leq 0.2$, so the a-value range for $HOA_{S,C}$ is selected as $0 \leq a_{HOA_{S,C}} \leq$
22 0.2. The a-values selected for constraints for the further summer bootstrap analysis are summarised in
23 Table 2. However, we also see that for $HOA_{S,C}$ the acceptance probability increases and stays high again
24 for the a-value of 0.4 to 0.8. Therefore, we made an additional bootstrap analysis to explore the result
25 when the a-value of $HOA_{S,C}$ randomises from 0 to 0.8, as discussed in the last paragraph in this section.

26 In the winter dataset, four factors ($HOA_{W,C}$, $COA_{W,C}$, $CSOA_{W,C}$, and $InorgNit_{W,C}$) are constrained,
27 yielding six pairs ($C(4,2) = 6$) of two-dimensional a-value scans. Compared to the summer dataset, the
28 unique base case/bootstrap correlation requirement yields a much smaller number of accepted solutions,
29 probably due to the more complicated aerosol sources and/or evolution conditions in winter (e.g.,
30 multiple biomass burning-related factors). Three factor-based diagnostic quantities were selected: 1)
31 the fraction of the nicotine signal ($[C_{10}H_{14}N_2]H^+$) apportioned to $CSOA_{W,C}$, 2) the relative intensity of
32 the AMS primary biomass burning tracer $C_2H_4O_2^+$ (Alfarra et al., 2007; Cubison and Jimenez, 2015) in
33 the factor profiles (AMS part) of less-aged biomass burning ($LABB_{W,C}$) vs. more-aged biomass burning
34 ($MABB_{W,C}$), and 3) the relative intensity of the EESI-TOF primary biomass burning tracer levoglucosan
35 ($[C_6H_{10}O_5]Na^+$) (Qi et al., 2019; Stefenelli et al., 2019; Lopez-Hilfiker et al., 2019) in the factor profiles
36 of $LABB_{W,C}$ vs. $MABB_{W,C}$. For 2) and 3), we require that the contribution of the primary tracer is higher
37 for the profile of $LABB_{W,C}$ than $MABB_{W,C}$ as follows:

38 $$\frac{LABB_{W,C,ion} - MABB_{W,C,ion}}{(LABB_{W,C,ion} + MABB_{W,C,ion})/2} > \theta_{ion} \qquad (17)$$

39 where $LABB_{W,C,ion}$ and $MABB_{W,C,ion}$ are the "ion" intensity in the $LABB_{W,C}$ and $MABB_{W,C}$ factor
40 profiles, and "ion" in Eq. (17) denotes either AMS $C_2H_4O_2^+$ (criterion 2) or EESI-TOF levoglucosan
41 ($[C_6H_{10}O_5]Na^+$) (criterion 3), and θ_{ion} denotes the acceptance threshold.

1   For criterion 1), we select the threshold $\theta_{CSOA_{W,C}}$ from investigation of Fig. S23, which shows the

2   frequency distribution of the fraction of total nicotine signal apportioned to $CSOA_{W,C}$, derived from the

3   multi-2D scans used to assess criteria 2 and 3 (see below). The figure shows that for nearly all runs, the

4   fraction of total nicotine mass apportioned to this factor is higher than 0.96. The exceptions are clear

5   outliers, and we therefore select $\theta_{CSOA_{W,C}} = 0.96$ which was therefore chosen as the criterion threshold.

6   The acceptance probability as a function of a-value is shown in Fig. S24 for $HOA_{W,C}$, $COA_{W,C}$,

7     $InorgNit_{W,C}$, and $CSOA_{W,C}$. For criteria 2 and 3, sensitivity tests are conducted using $\theta_{C_2H_4O_2^+}$ and

8     $\theta_{levoglucosan}$, which were varied from 0 to 1 with a step of 0.1, and the final selected values are shown

9   as a thicker line. The acceptance probability decreases to near-zero for $\theta_{C_2H_4O_2^+} \geq 0.1$ and $\theta_{levoglucosan}$

10   ≥ 0.2. We select 0 for both thresholds, which is the most permissive value, requiring only that $MABB_{W,C}$

11   appear more aged than $LABB_{W,C}$ (i.e., reduced contribution from POA tracers). Similar to the summer

12   dataset, there is no major decrease in acceptance probability at low a-values, and we therefore impose

13   only an upper limit. For $HOA_{W,C}$, we set the upper a-value boundary at 0.9, due to the large decrease in

14   acceptance probability at $a_{HOA_{W,C}} = 1.0$. However, for the other constrained factors, the acceptance

15   probability decreases steadily without a steep drop-off. We target an acceptance probability of ~0.4 (by

16   considering the unmixing status) as a subjective compromise between exploration and computational

17   efficiency, and select as an upper boundary the largest a-value that achieves this. This results in upper

18   a-value limits of 0.3 for $COA_{W,C}$, and 0.5 for $InorgNit_{W,C}$. For $CSOA_{W,C}$. the high acceptance probability

19   is kept high from the a-value of 0 to 0.6. Therefore, we chose the a-value range of $CSOA_{W,C}$ to be 0 to

20   0.6. However, it is also observed that the acceptance probability for this factor dips at 0.7 and stays high

21   again at a-values of 0.8 and 0.9, so we made an addition bootstrap analysis with the a-value range for

22    $CSOA_{W,C}$ of 0 to 0.9 to explore the influence of the a-value of this factor on overall result, as discussed

23   in the following paragraph. The a-values selected for the four constraints for the further winter bootstrap

24   analysis are summarised in Table 2.

25   After a-value selection, 1000 bootstrap runs were performed for summer and winter, respectively, and

26   in each bootstrap run, an a-value was randomly selected for each constrained factor, with a step size of

27   0.05 for summer and 0.1 for winter within the corresponding range. The criteria for accepted solutions

28   in the bootstrap analysis are exactly the same as the criteria and θ in Sect. 3.1.4, and are given in Table

29   2. As noted above, accepted solutions must simultaneously satisfy all criteria including the time-series-

30   based mixing status exploration and mass-spectral-based criteria. Note that we also did an additional

31   bootstrap analysis for summer and winter, respectively, as mentioned in previous paragraphs, to explore

32   the bootstrap result with larger a-value range of $HOA_{S,C}$ and $CSOA_{W,C}$. In the additional bootstrap

33   analysis for summer, a-value range for $HOA_{S,C}$ was set to be $0 \leq a_{HOA_{S,C}} \leq 0.8$, while the a-value ranges

34   of the other two constraints were kept the same as indicated in Table 2. Likewise, we only changed the

35   a-value range of $CSOA_{W,C}$ to be $0 \leq a_{CSOA_{W,C}} \leq 0.9$, while keeping the a-value ranges of the other

36   three constraints the same as in Table 2. Since the results of these additional bootstrap analysis are not

37   qualitatively different from the bootstrap analysis with a-value ranges in Table 2, we only present the

38   bootstrap results with a-value ranges in Table 2.

39

40   **Table. 2** Summary of a-value range for constrained factors, criteria for a-value range and accepted

41   bootstrap run selection and the number of accepted runs from the final combined bootstrap/a-value

42   analysis for the summer and winter datasets.

43

Dataset	Constrained factor	a-value range	Criteria	Accepted runs
Zurich summer	HOA$_{S,C}$	$0 \leq a \leq 0.2$	1). COA$_{S,C}$: $\frac{C_3H_3O^+}{C_3H_5O^+} \geq 5$ 2). InorgNit$_{S,C}$: $\frac{CO_2^+}{NO^+ + NO_2^+} \leq 0.035$ 3). Base case vs. Bootstrap correlation test at confidence level = 0	764 (76.4 %)
	COA$_{S,C}$	$0 \leq a \leq 0.2$		
	Inorganic nitrate (InorgNit$_{S,C}$)	$0 \leq a \leq 0.5$		
Zurich winter	HOA$_{W,C}$	$0 \leq a \leq 0.9$	1). CSOA$_{W,C}$: f$_{mass}$(nicotine) ≥ 0.96 2). $C_2H_4O_2^+$ intensity: LABB$_{W,C}$ − MABB$_{W,C}$ > 0 3). $C_6H_{10}O_5$ intensity: LABB$_{W,C}$ − MABB$_{W,C}$ > 0 4). Base case vs. Bootstrap correlation test at confidence level = 0	308 (30.8 %)
	COA$_{W,C}$	$0 \leq a \leq 0.3$		
	Inorganic nitrate (InorgNit$_{W,C}$)	$0 \leq a \leq 0.5$		
	CSOA$_{W,C}$	$0 \leq a \leq 0.6$		

1

3.2 cPMF results

Here we present final results from the cPMF analysis of the summer and winter campaigns. The final solutions are reported as the average of all accepted bootstrap/a-value randomisation runs (764 for summer, 308 for winter), with uncertainties corresponding to the standard deviation. We compare the cPMF factors to their counterparts from the standalone AMS and EESI-TOF solutions, for cases where a clear factor-to-factor correspondence exists.

A complication in this analysis is that the NO^+ and NO_2^+ signal can result from either organic or inorganic nitrate. Ideally, all inorganic NO^+ and NO_2^+ would apportion to the InorgNit$_{S,C}$ and InorgNit$_{W,C}$ factors, however inspection of the solutions reveals that this is not the case, as discussed in the factor presentations (Sect. 3.2.1 and Sect. 3.2.2). Therefore, we estimate the organic and inorganic contributions to these ions by the method of Kiendler-Scharr et al. (2016), as follows:

$$frac_{ON,k} = \frac{(1 + R_{ON})(R_k - R_{cal})}{(1 + R_k)(R_{ON} - R_{cal})} \tag{18}$$

Here we apply this analysis on a factor-by-factor basis, where $frac_{ON,k}$, defined in Eq. (19a), represents the fraction of ON apportioned to the kth factor, and R_k denotes the intensity ratio of NO_2^+ to NO^+ in the factor profile. R_{cal} is the reference NO_2^+/NO^+ ratio for inorganic nitrate, taken as that of the InorgNit$_{W,C}$ and InorgNit$_{S,C}$ reference profiles for their respective datasets. R_{ON}, defined in Eq. (19b), is the intensity ratio of NO_2^+ to NO^+ for organonitrate, which ranges from 0.08 to 0.20 (Boyd et al., 2015; Bruns et al., 2010; Fry et al., 2011; Fry et al., 2009; Rollins et al., 2009).

$$frac_{ON,k} = \frac{(f_{ON,k,NO^+} + f_{ON,k,NO_2^+})}{(f_{k,NO^+} + f_{k,NO_2^+})} \tag{19a}$$

$$R_{ON} = \frac{f_{ON,k,NO_2^+}}{f_{ON,k,NO^+}} \tag{19b}$$

Here f_{k,NO^+} and f_{k,NO_2^+} denote the total NO^+ and NO_2^+ signal, respectively in the kth factor profile, while f_{ON,k,NO^+} and f_{ON,k,NO_2^+} denote the organonitrate contribution to these ions. Because f_{k,NO^+} and f_{k,NO_2^+} are directly available from the factor profile, $frac_{ON,k}$ is independently calculated via Eq. (18),

1 and R_{ON} is assumed, Eqs. (19a) and (19b) constitute a system of 2 equations with 2 unknowns, which
2 can be solved algebraically for f_{ON,k,NO^+} and f_{ON,k,NO_2^+}, yielding:

$$f_{ON,k,NO^+} = \frac{(R_k - R_{cal})(f_{k,NO^+} + f_{k,NO_2^+})}{(1 + R_k)(R_{ON} - R_{cal})} \tag{20a}$$

$$f_{ON,k,NO_2^+} = \frac{(R_k - R_{cal})(f_{k,NO^+} + f_{k,NO_2^+})}{(1 + R_k)(R_{ON} - R_{cal})} \cdot R_{ON} \tag{20b}$$

6 These calculations are important not only for profile interpretation, but also for quantitative
7 apportionment of OA. Specifically, as noted earlier, calculations of the OA contribution to the factor
8 time series, $(g_{i,k})_{AMS}$, and the EESI-TOF sensitivity to a given factor, AS_k, should consider only the
9 organic contribution to NO^+ and NO_2^+. In this study, we estimated the contribution from organonitrates
10 for all factors in summer and winter assuming the midpoint of the R_{ON} range (R_{ON} = 0.14).
11 Organonitrate contributions ($frac_{ON,k}$) to the total nitrate signal for each factor and the corresponding
12 OA fraction $\sum_j (f_{k,j})_{AMS}$ are shown in Table S1. We also include the same calculations performed
13 assuming an R_{ON} of 0.08 or 0.20, which as discussed above consitute the lower and upper estimates
14 from previous studies. For R_{ON} = 0.14, the $frac_{ON,k}$ for all SOAs in summer are higher than 75 %, and
15 for winter, this fraction $frac_{ON,k}$ varies by factor from 0 to 100 %, with four factors having $frac_{ON,k}$
16 = 100 % (SOA1$_{W,C}$, MABB$_{W,C}$, LABB$_{W,C}$ and NitOA1$_{W,C}$), suggesting the NO^+ and NO_2^+ signals are
17 strongly influenced by ON. If R_{ON} = 0.08 is assumed, the estimated $frac_{ON,k}$ decreases by ~12 % for
18 the summer SOA factors and by 10 % to 20 % for the winter SOA factors, whereas assuming R_{ON} =
19 0.20 increases $frac_{ON,k}$ by ~15 % in the summer and 16% in the winter OA factors. The effect of this
20 assumption on the factor OA concentration and thus AS_k is much smaller, with all factors below ±2 %
21 except for one wintertime SOA factor (SOA1$_{W,C}$, ±6 %).
22 .
23

24 **3.2.1   cPMF analysis: Zurich summer**
25 Eight factors were resolved from the Zurich summer campaign: HOA$_{S,C}$, COA$_{S,C}$, CSOA$_{S,C}$, InorgNit$_{S,C}$,
26 two daytime SOA factors (DaySOA1$_{S,C}$ and DaySOA2$_{S,C}$), and two nighttime SOA factors
27 (NightSOA1$_{S,C}$ and NightSOA2$_{S,C}$). The mean time series, diurnal cycle and the mass spectra of these
28 factors over 764 accepted runs are shown in Fig. 5, together with the time series from AMS-only PMF
29 and/or EESI-TOF-only PMF when the corresponding standalone factor(s) exist. An estimate of
30 campaign-average percent uncertainty in the mass concentration of each factor, calculated as the median
31 of the standard deviation across all accepted runs, is given in Table S2. Many factor characteristics from
32 cPMF resemble those previously discussed in detail for single-instrument AMS PMF and/or EESI-TOF
33 PMF (Stefenelli et al., 2019). Therefore, only a summary discussion of these characteristics are
34 presented here, and we focus on new information and/or differences obtained by the cPMF analysis.
35 Recall that factor profiles for HOA$_{S,C}$, COA$_{S,C}$, and InorgNit$_{S,C}$ are constrained as discussed above.
36

37 **HOA$_{S,C}$** --- The AMS mass spectrum is dominated by the $C_nH_{2n+1}^+$, and $C_nH_{2n-1}^+$ series, consistent with
38 n-alkanes and branched alkanes (Ng et al., 2011a; Ulbrich et al., 2009; Lanz et al., 2007; Zhang et al.,
39 2005; Qi et al., 2019; Stefenelli et al., 2019). The diurnal cycle of HOA$_{S,C}$ has three clear peaks (see
40 Fig. 5b), however, compared to HOA$_{S,A}$ from Stefenelli et al. (2019), their intensities are weaker.
41 Specifically, the morning peak intensity ratio to the evening peak intensity is almost 1 in the HOA$_{S,A}$
42 factor, whereas in HOA$_{S,C}$, the morning peak is ~1/3 of the evening peak. In terms of contribution to
43 total OA, the HOA$_{S,A}$ factor contributes 5.8 % (0.177 µg m^{-3}) of the total OA, whereas in the cPMF
44 analysis, this factor only contributes 3.1 % (0.092 µg m^{-3}) of the total OA.

1

2 **COA$_{S,C}$** --- This factor is characterised by long-chain fatty acids and alcohols, e.g., coronaric acid and/or

3 its isomers at m/z 319.2 ([C$_{18}$H$_{32}$O$_3$]Na$^+$), oleic acid and/or its isomers at m/z 305.2 ([C$_{18}$H$_{34}$O$_2$]Na$^+$),

4 and 2-oxo-tetredecanoic acid and/or its isomers at m/z 293.2 ([C$_{16}$H$_{30}$O$_3$]Na$^+$). Similar to previous work,

5 the AMS profile shows both alkyl fragments and slightly oxygenated ions, consistent with aliphatic

6 acids from cooking oils (Hu et al., 2016). The AMS profile is characterised by a high ratio of C$_3$H$_3$O$^+$

7 to C$_3$H$_5$O$^+$ (~5 here), slightly higher than in other studies (Xu et al., 2019; Zhao et al., 2019; Sun et al.,

8 2016a; Sun et al., 2016b), as well as high contributions from C$_5$H$_8$O$^+$, C$_6$H$_{10}$O$^+$ and C$_7$H$_{12}$O$^+$. Both

9 cPMF and single instrument PMF analyses yield peaks during lunch and dinner. The time series of

10 COA$_{S,C}$ is strongly correlated with those of the single instrument solutions, with Pearson's r^2 of 0.846

11 and 0.634 against COA$_{S,A}$ and COA$_{S,E}$, respectively.

12

13 **CSOA$_{S,C}$** --- The EESI-TOF factor profile is dominated by nicotine (detected as [C$_{10}$H$_{13}$N$_2$]H$^+$) at m/z

14 163.12 and levoglucosan at m/z 185.042 ([C$_6$H$_{10}$O$_5$]Na$^+$), which derives from pyrolysis of the cellulose

15 present in tobacco (Talhout et al., 2006). In the AMS profile, this factor accounts for 79.3 % of the

16 signal from C$_5$H$_{10}$N$^+$ at m/z 84.081, which is attributed to a fragment of n-methyl pyrrolidine and

17 previously identified as a tracer for cigarette smoke (Struckmeier et al., 2016). The time series of

18 CSOA$_{S,C}$ correlates with that of the AMS-only and EESI-TOF solutions, with r^2 of 0.922 and 0.965,

19 respectively. The diurnal cycles from the combined and single-instrument solutions are likewise

20 correlated, showing high concentrations at night and low concentration during daytime.

21

22 **InorgNit$_{S,C}$** --- Among the accepted bootstrap runs, the mean CO$_2^+$/(NO$^+$+NO$_2^+$) ratio is 0.0346, slightly

23 higher than the ratio of 0.0345 observed during the NH$_4$NO$_3$ calibration period, probably due to 1)

24 uncertainties in the constrained profile, and/or 2) a small amount of OA apportioned to this factor. The

25 time series of this factor correlates with AMS nitrate (NO$_3^-$), NO$^+$ and NO$_2^+$ time series, with r^2 of 0.654,

26 0.645 and 0.956, respectively. Regarding the mass fraction, approximately 48.5 % of the NO$^+$ signal

27 and 78.0 % of the NO$_2^+$ signal are apportioned to this factor, followed by the two NightSOA$_{S,C}$ factors.

28 This is consistent with the overall NO$^+$ and NO$_2^+$ signals deriving not only from inorganic nitrate, but

29 also from organonitrates (in other factors).

30

31 **DaySOA1$_{S,C}$ and DaySOA2$_{S,C}$** --- The cPMF analysis yields two SOA factors elevated during daytime,

32 denoted DaySOA1$_{S,C}$ and DaySOA2$_{S,C}$. The EESI-TOF spectra are similar to two factors retrieved from

33 EESI-TOF-only PMF analysis by Stefenelli et al. (2019), but were not resolved in AMS-only PMF,

34 where only more- and less-oxygenated SOA factors (MO-OOA$_{S,A}$ and LO-OOA$_{S,A}$) were obtained.

35 These factors contain strong signatures from terpene oxidation products, e.g., monoterpene-derived ions

36 (C$_{10}$H$_{16}$O$_x$, x=5, 6, 7) and sesquiterpene oxidation products (C$_{15}$H$_{24}$O$_x$, x=3, 4, 5). A detailed comparison

37 of the two DaySOA factors from the cPMF analysis to the LO-OOA$_{S,A}$ and MO-OOA$_{S,A}$ factors from

38 AMS-only PMF is shown in Fig. S28, and a comparison between the two DaySOA$_{S,C}$ factors and

39 DaySOA$_{S,E}$ factors are shown in Figs. S29 a) and b), respectively. The AMS ions in these two factors

40 are characterised by a strong CO$_2^+$ signal, similar to the LO-OOA$_{S,A}$ and MO-OOA$_{S,A}$ factors, indicating

41 they largely consist of oxygenated OA, consistent with the EESI-TOF spectra. We calculate $frac_{ON}$

42 for DaySOA1$_{S,C}$ and DaySOA2$_{S,C}$ to be 0.869 and 1.000, respectively, demonstrating that the NO$^+$ and

43 NO$_2^+$ signal in these factors is dominated by organonitrates. Regarding the time series, DaySOA1$_{S,C}$ and

44 DaySOA2$_{S,C}$ correlate strongly with DaySOA1$_{S,E}$ and DaySOA2$_{S,E}$, with r^2 of 0.883 and 0.977,

45 respectively. The diurnal patterns of DaySOA1$_{S,C}$ and DaySOA2$_{S,C}$ are consistent with the diurnal

46 patterns of DaySOA1$_{S,E}$ and DaySOA2$_{S,E}$. The diurnal patterns of both factors show an enhancement in

47 the afternoon and the evening, which distinguish these SOAs from other SOAs: DaySOA1$_{S,C}$ exhibits

Atmospheric
Measurement
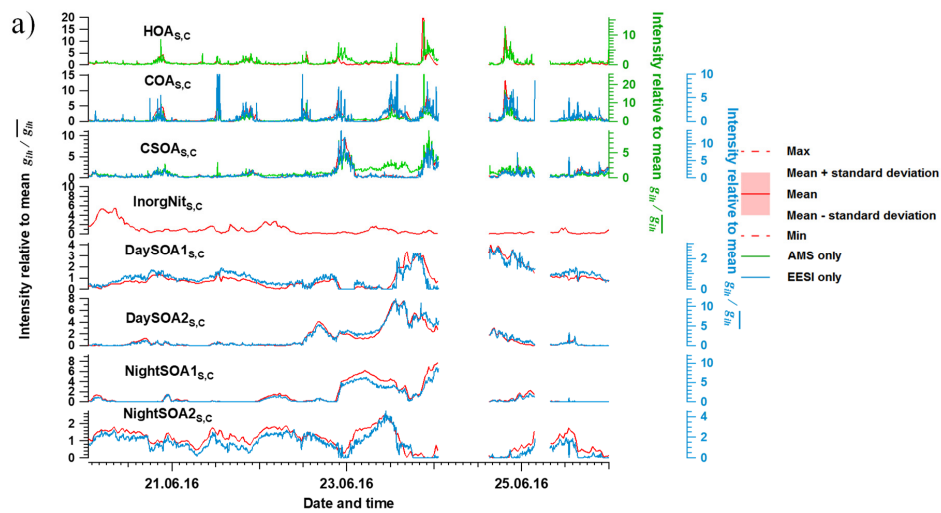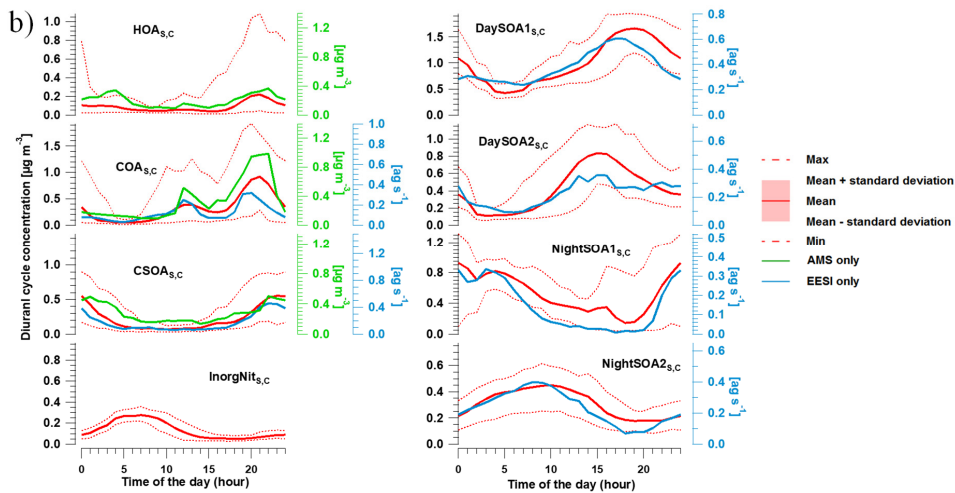Techniques

Open Access

EGU

Discussions

1   almost a factor of 2 enhancement in signal between 15:00 and 21:00 compared to the morning, whereas

2   the $DaySOA2_{S,C}$ exhibits the same magnitude of enhancement in signal around 12:00 to 17:00.

3

4   **$NightSOA1_{S,C}$ and $NightSOA2_{S,C}$** --- We retrieve two SOA factors that are enhanced overnight and in

5   the early morning, denoted $NightSOA1_{S,C}$ and $NightSOA2_{S,C}$. Their factor profiles and time

6   series/diurnals closely resemble those of $NightSOA1_{S,E}$ and $NightSOA2_{S,E}$ (see Figs. S29c and S29d).

7   Similar to the $DaySOA_{S,C}$ factors, terpene oxidation products are evident. However, the composition is

8   weighted towards less oxygenated and more volatile terpene oxidation products, e.g., $C_{10}H_{16}O_2$ and

9    $C_{10}H_{16}O_3$, which likely partition to the particle phase at night when temperature decreases. In addition,

10   signals consistent with monoterpene-derived organonitrates are also evident, e.g., the $C_{10}H_{17}O_{6-8}N$ and

11   $C_{10}H_{15}O_{6-9}N$ series, which are consistent with night time oxidation of monoterpenes by NO_3 radicals

12   (Faxon et al., 2018; Zhang et al., 2018; Xu et al., 2015). The AMS ions in these two factors are

13   characterised by a strong CO_2^+ signal and also a relatively high NO^+ signal compared to

14   $sum\_DaySOAs_{S,C}$. The ratio of NO^+/NO_2^+ ratio is 4.55 and 8.24 for $NightSOA1_{S,C}$ and $NightSOA2_{S,C}$,

15   respectively, yielding $frac_{ON}$ for $NightSOA1_{S,C}$ and $NightSOA2_{S,C}$ of 0.798 and 1, indicating high

16   organonitrate content. These two factors correlate well with $sum\_NightSOAs_{S,E}$, reaching r^2 of 0.975

17   and 0.897, following in general the same diurnal patterns, with $NightSOA1_{S,C}$ peaking from 22:00 to

18   05:00 and $NightSOA1_{S,C}$ peaking from 04:00 to 12:00.

Atmospheric
Measurement
Techniques

Open Access

EGU

Discussions

1

2

3

Atmospheric
Measurement
Techniques
Discussions
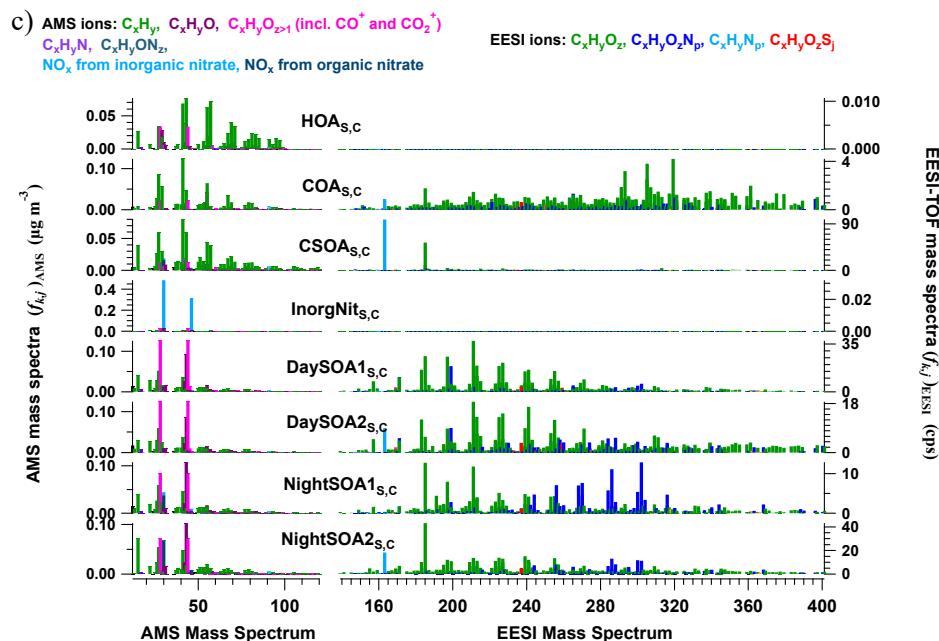
Figure 5. Mean factor time series (a), diurnal cycles (b) and factor profiles (c) from the 764 accepted bootstrap runs from cPMF analysis. In a), the average factor time series are shown in red, and corresponding AMS and/or EESI-TOF factors from standalone PMF are shown in green and blue, respectively. Shaded areas represent the standard deviation across all accepted runs and are summarised in Table S2. In b), the average diurnal cycles are displayed as red solid lines. Shaded areas denote the standard deviation over the average diurnal from individual solutions over all 764 accepted runs. Dashed lines denote the maximum and minimum mean diurnal observed within these 764 runs. For comparison, the AMS and EESI-TOF PMF factor time series and diurnal cycles from the individual dataset in Stefenelli et al. (2019) are shown in green and blue respectively for related factors. In c), the average factor profiles are coloured by different ion families. Here, the AMS factor profiles are in the unit of $\mu g \ m^{-3}$ (each factor sums to 1 $\mu g \ m^{-3}$), whereas the EESI-TOF spectra are in the unit of cps (each factor sums to the total signal derived from 1 $\mu g \ m^{-3}$ of the factor). Note that the NO^+ and NO_2^+ signal is divided into inorganic and organic contributions.

3.2.2 cPMF analysis: Zurich winter

Twelve factors were resolved from cPMF analysis of the Zurich winter campaign: $HOA_{W,C}$, $COA_{W,C}$, $InorgNit_{W,C}$, $CSOA_{W,C}$, $SOA1_{W,C}$, $SOA2_{W,C}$, a more-aged biomass burning OA ($MABB_{W,C}$), two less-aged biomass burning OAs ($LABB1_{W,C}$ and $LABB2_{W,C}$), two nitrogen-containing OA factors ($NitOA1_{W,C}$ and $NitOA2_{W,C}$), and a factor related to a specific local event ($EVENT_{W,C}$). Because no significant chemical differences are apparent between $LABB1_{W,C}$ and $LABB2_{W,C}$ (see Figs. S30 and S31), they are aggregated to a single $LABB_{W,C}$ factor for presentation. Therefore, there are 11 factors presented below. The average time series and mass spectra of these factors among 308 accepted runs are shown in Fig. 6. The factor profiles for $HOA_{W,C}$, $COA_{W,C}$, $InorgNit_{W,C}$, and $CSOA_{W,C}$ are constrained as described previously. Similar to the summer dataset, uncertainties in the factor mass concentrations are summarised in Table S2.

Atmospheric
Measurement
Techniques
Discussions

1   **HOA$_{W,C}$** --- This factor has a qualitatively a profile similar to the summer campaign, and the discussion
2   of the HOA$_{S,C}$ profile applies here as well. The HOA$_{W,C}$ time series correlates strongly with HOA$_{W,A}$ (r^2
3   of 0.913).

4   **COA$_{W,C}$** --- The COA$_{W,C}$ profile is characterised by long-chain fatty acids and alcohols e.g., coronaric
5   acid and/or its isomers at m/z 319.2 ($[C_{18}H_{32}O_3]Na^+$), oleic acid and/or its isomers at m/z 305.2
6   ($[C_{18}H_{34}O_2]Na^+$), and 2-oxo-tetredecanoic acid and/or its isomers at m/z 293.2 ($[C_{16}H_{30}O_3]Na^+$), and in
7   the AMS, a combination of alkyl fragments and slightly oxygenated ions from aliphatic acids from
8   cooking oils, including $C_5H_8O^+$, $C_6H_{10}O^+$ and $C_7H_{12}O^+$. These features are consistent with features
9   found by other studies (Qi et al., 2019; Stefenelli et al., 2019; Tong et al., 2021). The COA$_{W,C}$ time
10 series correlates with the corresponding single instrument analyses, exhibiting r^2 of 0.894, and 0.798,
11 with COA$_{W,A}$ and COA$_{W,E}$, respectively.

12 **InorgNit$_{W,C}$** --- As noted in Sect. 3.1.2, the NO^+/NO_2^+ ratio of this factor (2.42) is higher than that of
13 pure NH_4NO_3 measured onsite (1.58), consistent with the presence of other inorganic nitrate sources
14 such as KNO_3. Also, the mean $CO_2^+/(NO^++NO_2^+)$ ratio is 0.0371, higher than the ratio of 0.0261 from
15 the constructed InorgNit$_{W,C}$ profile, probably due to 1) uncertainties in the constrained profile, and/or
16 2) a small amount of OA apportioned to this factor. The time series of this factor shows high correlations
17 with the AMS nitrate (NO_3^-), NO^+ and NO_2^+ time series, with r^2 of 0.739, 0.792 and 0.754, respectively.
18 Regarding the mass fraction, only 13.7% of the NO^+ signal and 13.2 % of the NO_2^+ signal are
19 apportioned to this factor. The considerable fractions of the NO^+ and NO_2^+ signal from inorganic nitrate
20 and organonitrates in other factors are estimated as discussed above (Kiendler-Scharr et al., 2016) and
21 will be interpreted later for the relevant factors (as summarised in Table S1).

22

23 **CSOA$_{W,C}$** --- Similar to CSOA$_{S,C}$, nicotine at m/z 163.12 and levoglucosan at m/z 185.042 were found
24 to be the two highest peaks in the EESI-TOF mass spectra, contributing 8.75 % and 4.56 % of the EESI-
25 TOF signal. The time series of this factor resolved from cPMF analysis correlates with CSOA$_{W,E}$ (r^2 =
26 0.662). Similar to CSOA$_{W,C}$, the fragment of cigarette smoke tracer n-methyl pyrrolidine $C_5H_{10}N^+$ at
27 m/z 84.081 is also found here. This is a minor factor, comprising 2.4 % of OA.

28

29 **SOA1$_{W,C}$ and SOA2$_{W,C}$** --- these two factors have qualitatively similar spectra but different temporal
30 patterns. SOA1$_{W,C}$ decreased gradually from 26 to 30 January, whereas SOA2$_{W,C}$ increased from 26
31 January and fluctuated at high level from 28 to 31 January and then decreased from 1 February on.
32 From the AMS perspective, both factors are characterised by high NO^+, NO_2^+ and CO_2^+ signal compared
33 to other organic ions. Organonitrates account for all NO^+ and NO_2^+ signals in SOA1$_{W,C}$, but contribute
34 nothing in SOA2$_{W,C}$. Aside from the NO^+ and NO_2^+ ions, these AMS spectra are similar to the profiles
35 of MO-OOA $_{W,A}$ and LO-OOA$_{W,A}$ which are characterised by high CO_2^+ signal. Major ions in the EESI-
36 TOF profile include $C_{10}H_{16}O_x$ (x = 3, 4, 5), $C_9H_{14}O_x$ (x = 3, 4), $C_8H_{12}O_x$ (x = 4, 5), $C_{10}H_{18}O_4$, and
37 $C_{10}H_{14}O_5$, which are also found in secondary biomass burning (three MABB$_{W,E}$ factors) and/or terpene
38 oxidation factors (SOA1$_{W,E}$ and SOA2$_{W,E}$) from Qi et al. (2019). However, the H:C ratio of these two
39 factors from the EESI-TOF component (1.578 and 1.588 for SOA1$_{W,C}$ and SOA2$_{W,C}$, respectively) is
40 less than that of DaySOA1$_{S,C}$ (1.650) and DaySOA2$_{S,C}$ (1.672), suggesting an increased contribution
41 from aromatic precursors.

42

43 **Biomass burning factors (LABB$_{W,C}$ and MABB$_{W,C}$)** --- We resolve a less-aged biomass burning
44 factor (LABB$_{W,C}$, which, as mentioned above, is the aggregate of two similar LABB factors), and a
45 more-aged biomass burning factor (MABB$_{W,C}$). Consistent with Qi et al. (2019), the EESI-TOF
46 component of LABB$_{W,C}$ is characterised by a large signal from $[C_6H_{10}O_5]Na^+$ (mainly levoglucosan)

1   (20.4 %), and MABB$_{W,C}$ by a smaller but notably non-zero one (6.21 %). In addition, 76.7 % and 11.9 %
2   of the total levoglucosan signal is apportioned to LABB$_{W,C}$, and MABB$_{W,C}$, respectively. The difference
3   in the fraction of total levoglucosan apportioned to these two factors suggests different degrees of ageing
4   of biomass burning-emitted OA. The AMS spectrum of the BBOA$_{W,A}$ factor is characterised by
5    $C_2H_4O_2^+$ and $C_3H_5O_2^+$, which are typical fragments of anhydrosugars, such as levoglucosan (Alfarra et
6   al., 2007; Lanz et al., 2007; Sun et al., 2011). These ions are also present in LABB$_{W,C}$ and MABB$_{W,C}$
7   and are higher in LABB$_{W,C}$ (1.91 % vs 0.879 % for $C_2H_4O_2^+$ and 0.978 % vs 0.323 % for $C_3H_5O_2^+$). In
8   addition, the ratio of $C_2H_4O_2^+$ to CO_2^+ is 0.396 and 0.092 for LABB$_{W,C}$ and MABB$_{W,C}$, respectively,
9   supporting the separation of these factors based on different degrees of ageing.

10

11   **EVENT$_{W,C}$** --- This factor is low throughout the campaign except for the nights of 28 and 29 January
12   from 00.00 to 07.00 UTC+2, where large peaks are observed. Therefore, it likely corresponds to a
13   specific event near the sampling location. The mass spectrum features ions at *m/z* 174.08, 185.04 and
14   195.06, tentatively assigned to $[C_8H_{11}N_2O]Na^+$, $[C_6H_{10}O_5]Na^+$ and $[C_8H_{12}O_4]Na^+$ from the EESI-TOF
15   part and at *m/z* 15.024 (CH_3^+), 27.027 ($C_2H_3^+$), 31.018 (CH_3O^+), and 43.018 ($C_2H_3O^+$) from the AMS
16   part. Qi et al. (2019) observed a very similar factor in standalone EESI-TOF PMF, which was tentatively
17   attributed to the Zurich gaming festival and/or plastic burning in a nearby restaurant. The factor includes
18   large contributions from $C_8H_{12}O_4$, which likely represents 1,2-cyclohexane dicarboxylic acid diisononyl
19   ester, a plasticiser for the manufacture of food packaging. In the AMS spectrum, large signals from NO^+
20   (7.36%) and NO_2^+ (2.03 %) are also observed, with 46.6 % of the NO^+ signal and 23.6% of the NO_2^+
21   signal assigned to organonitrates. Similar to Qi et al. (2019), the AMS spectrum is also dominated by
22   the ions in the $C_xH_yO_z^+$ group.

23

24   **NitOA1$_{W,C}$**--- this factor is characterised by a high signal of $C_5H_{10}N^+$ at *m/z* 84.081, contributing 4.02 %
25   to the AMS intensity in this factor (no other factor exceeds 0.16 %) while 97.0 % of the $C_5H_{10}N^+$ mass
26   is apportioned to this factor. This ion is considered to be a tracer of cigarette smoking (Struckmeier et
27   al., 2016), however, different from typical CSOA mass spectra, this factor also has high signal from
28   CO_2^+, suggesting a contribution from secondary formation processes. Similar to other OA factors, this
29   factor also has a considerable fraction of NO^+ and NO_2^+ signal, attributed entirely to organonitrates. For
30   the EESI-TOF component, this factor is characterised by $[C_8H_{11}N_2O]Na^+$, levoglucosan and
31    $[C_8H_{11}N_2O]Na^+$, $[C_6H_{10}O_5]Na^+$ and $[C_9H_{12}O_4]Na^+$ and $[C_{11}H_{14}O_4]Na^+$, suggesting this factor may also
32   be influenced by fresh biomass burning.

33

34   **NitOA2$_{W,C}$** --- this factor is characterised by a high fraction of total signal from the CHON group in the
35   EESI-TOF analysis (38.5 %). Among these ions, $[C_7H_{11}O_6N]Na^+$ at *m/z* 228.048, $[C_{10}H_{15}O_6N]Na^+$ at
36   *m/z* 268.079, and $[C_{10}H_{17}O_7N]Na^+$ at *m/z* 286.090 are the three highest ions, contributing 1.65 %, 1,99 %,
37   and 1.98 %, respectively. There are also some typical ions with high intensity from biomass burning
38   ageing (Qi et al., 2019; Stefenelli et al., 2019), e.g., $[C_9H_{14}O_4]Na^+$ at *m/z* 209.078, $[C_{10}H_{14}O_6]Na^+$ at *m/z*
39   253.068, and $[C_{10}H_{16}O_6]Na^+$ at *m/z* 255.084, contributing 6.47 %, 2.85 %, and 4.39 %, respectively.
40   This may suggest a contribution from biomass burning activities. From the AMS perspective, this factor
41   is characterised by high NO^+ and NO_2^+ signal, in which all of the NO^+ and NO_2^+ signals are produced
42   from inorganic nitrates (see Table S1), with the other ions being qualitatively similar to OOA-type
43   spectra.
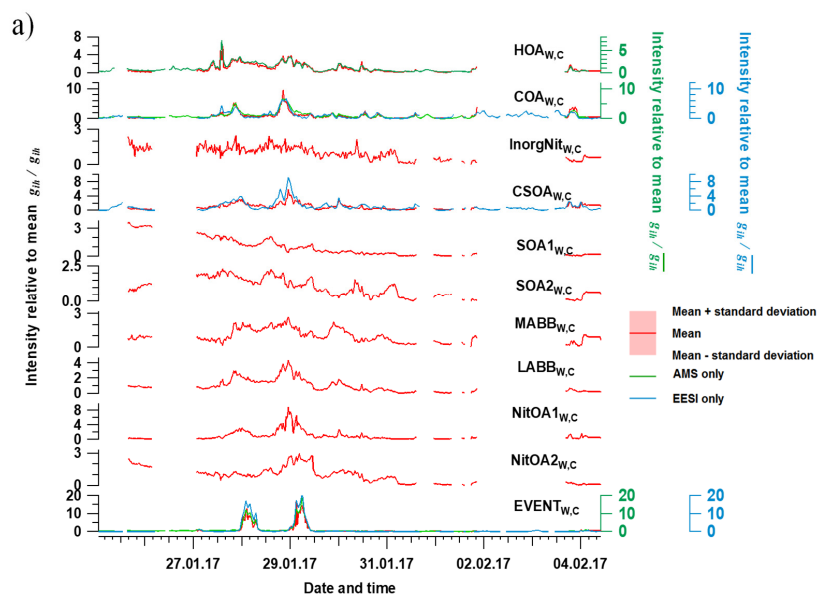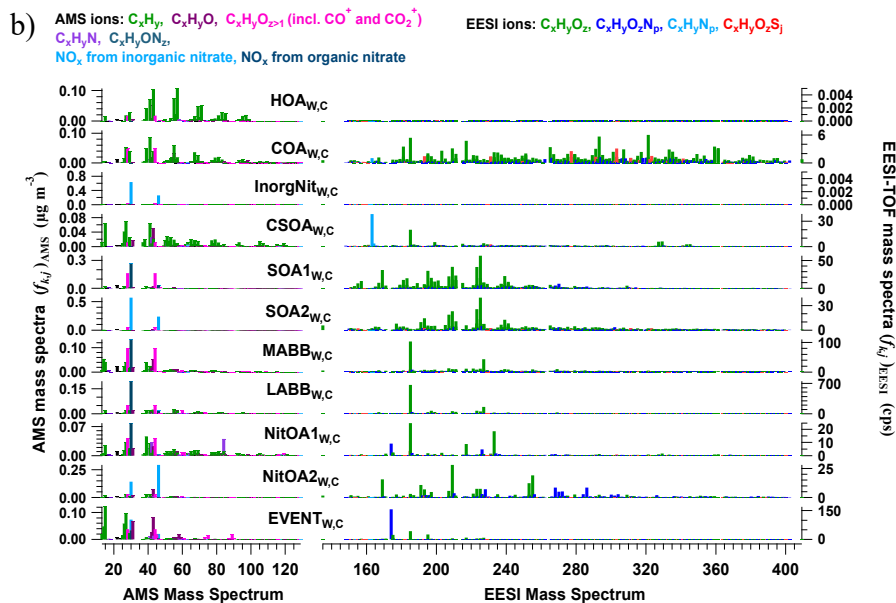
Figure 6. Average factor time series (a) and factor profiles (b), which are calculated as the mean of all accepted bootstrap runs (308 runs in total). In a), the average factor time series are shown in red, and corresponding AMS and/or EESI-TOF factors from standalone PMF are shown in green and blue, respectively. Shaded areas represent the standard deviation across all accepted runs, and are summarised in Table S2. In b), the average factor profiles are coloured by different ion families. Here, the AMS factor profiles are in the unit of $\mu g\ m^{-3}$ (each factor sums to 1 $\mu g\ m^{-3}$), whereas the EESI-TOF spectra are in the unit of cps (each factor sums to total signal derived from 1 $\mu g\ m^{-3}$ of the factor). Note that the NO^+ and NO_2^+ signal is divided into inorganic and organic contributions.

1

2 **3.3 EESI-TOF sensitivity to resolved factors**

3 AMS and EESI-TOF contributions to the factor profiles are intrinsically linked by cPMF. That is, for
4 each individual factor the two instrument profiles by definition describe the same OA fraction.
5 Therefore, the EESI-TOF sensitivity to a factor AS_k can be calculated according to Eq. (10). Note that
6 this calculation depends on the assumptions that (1) both instruments are well-represented in the
7 solution; (2) the PMF solution is of high quality (i.e., factors are all meaningful and well-separated,
8 without significant mixing or splitting); (3) solution uncertainties are not so high as to preclude
9 quantitative interpretation of the results. Assumption (1) was discussed earlier in the context of
10 instrument weighting, and assumption (2) is supported by the interpretability of the factors as presented
11 in the previous section. By performing the cPMF analysis on a large number of runs combining
12 bootstrap analysis and a-value exploration, we can estimate uncertainties in the calculated sensitivities
13 imposed by the analysis model, as presented below, thereby addressing assumptions (2) and (3).

14

15 The datasets analysed here were taken from the first field deployments of the EESI-TOF. As a result,
16 operational protocols were not yet fully standardised across campaigns. Specifically, we lack reliable
17 on-site calibration with a chemical standard common to the two campaigns (this was attempted but the
18 measurements were evaluated to be unreliable during post-analysis due to operational problems).
19 Therefore, to enable comparison of relative factor sensitivities between the summer and winter
20 campaigns, we select COA as a reference. That is, we assume $AS_{COA} = AS_{COA_{S,C}} = AS_{COA_{W,C}}$ .We
21 choose COA because it is the only factor that both (1) appears in all four single-instrument datasets (i.e.,
22 summer and winter, AMS and EESI-TOF) and (2) compared to other factors, is less likely to
23 significantly change in composition between the campaigns (in contrast to, e.g., SOA in Zurich, which
24 is known to have significantly different precursors in summer and winter). Therefore, all sensitivities
25 below are reported as $(AS_k/\overline{AS_{COA}})$, in which AS_k is calculated in every bootstrap run, and then
26 referenced to $\overline{AS_{COA}}$ (the mean AS_{COA} calculated over all bootstrap runs). Here k denotes a given factor
27 from the (summer or winter) cPMF solutions. Note that EESI-TOF sensitivities to HOA and InorgNit
28 are not discussed here, since they are undetectable by EESI-TOF (as configured for these campaigns;
29 see Sect. 2.2.2) and therefore constrained to be ~0.01 cps / (ug m^{-3}). The mean and standard deviation
30 of factor-dependent $AS_k/\overline{AS_{COA}}$ for the summer and winter datasets are shown in Fig. 7, with
31 histograms summarising all accepted runs shown in Fig. S32 and Fig. S33.

32

33 For ease of viewing, the factors in Fig. 7 are collected into related groups. We also calculate the AS_k's
34 for several factor aggregations. First, five factors that are likely related to biomass burning (LABB$_{W,C}$,
35 MABB$_{W,C}$, NitOA1$_{W,C}$, NitOA2$_{W,C}$ and EVENT$_{W,C}$), are denoted as the "Sum_BB" factor. Additionally,
36 we separately aggregate the two DaySOA$_{S,C}$ and two NightSOA$_{S,C}$ factors, denoted "sum_DaySOAs$_{S,C}$"
37 and "sum_NightSOAs$_{S,C}$", respectively. As seen in Fig. 7 (as well as Fig. S34 and Table S3), the relative
38 uncertainty from the summer factors is systematically lower than for the winter factors within the
39 accepted solutions. This may indicate higher source apportionment quality and solution stability for the
40 former, but is also related to the sub-division of factors related to primary biomass burning-related
41 factors, as discussed later.

42 For COA$_{S,C}$ and COA$_{W,C}$, the mean relative sensitivities are 1 by definition, though uncertainties are
43 still calculated due to non-zero a-values, while the reference profile utilised for CSOA$_{W,C}$, ensures that
44 CSOA$_{W,C}$ CSOA$_{S,C}$ will have similar sensitivities. Interestingly, the distribution of the sensitivities of
45 COA$_{S,C}$, COA$_{W,C}$ and CSOA$_{W,C}$ in Fig. S32 and Fig. S33 is clearly multi-modal despite a-value
46 constraints, but the reason for this remains to be explored.

1   The next group of factors (LABB$_{W,C}$, MABB$_{W,C}$, NitOA1$_{W,C}$, NitOA2$_{W,C}$ and EVENT$_{W,C}$) includes non-
2   negligible contributions from levoglucosan (C$_6$H$_{10}$O$_5$), produced typically from biomass-burning(BB)-
3   related activities. Previous work has demonstrated that the EESI-TOF sensitivity to levoglucosan is
4   higher than that of many other compounds and bulk SOA from representative precursors (Lopez-
5   Hilfiker et al., 2019; Brown et al., 2021). Indeed, although the set of studied compounds is far from
6   comprehensive, the relative sensitivity of the EESI-TOF to levoglucosan is among the highest yet
7   recorded. Therefore, although the composition of the POA-influenced factors varies considerably, it is
8   possible that the levoglucosan content may have significant predictive value with respect to the overall
9   factor sensitivity. Figure 8 shows AS_k as a function of the C$_6$H$_{10}$O$_5$ fraction for all factors for which the
10   C$_6$H$_{10}$O$_5$ signal is believed to result largely from levoglucosan. This analysis accounts for all factors
11   resolved from the cPMF of the winter dataset except CSOA$_{W,C}$, because CSOA$_{W,C}$ is dominated by the
12   signal from the protonated nicotine ([C$_{10}$H$_{14}$N$_2$]H$^+$) ion, which is both chemically different (reduced
13   nitrogen) and has a different ionisation pathway than other measured ions. The four summer SOA
14   factors are excluded as well, because the contribution from C$_6$H$_{10}$O$_5$ in these factors was previously
15   attributed to terpene and/or aromatic oxidation products (Stefenelli et al., 2019). An obvious qualitative
16   trend of increasing sensitivity with increasing levoglucosan fraction is evident with Pearson r^2 of 0.676,
17   indicating the overwhelming influence of the high sensitivity species levoglucosan on the factor
18   apparent sensitivity.

19   For the primary BB-related factors, the uncertainties are generally higher than for the other factors (see
20   Fig. S33 and Fig. S34b). In contrast, the aggregated BB factor (Sum_BB$_{W,C}$, and Sum_BB$_{W,C}$ =
21   MABB$_{W,C}$ + LABB$_{W,C}$ + NitOA1$_{W,C}$ + NitOA2$_{W,C}$ + EVENT$_{W,C}$) is less uncertain and has a narrower
22   sensitivity distribution. This suggests that the overall classification of signal as biomass burning-related
23   is robust, but the subdivision into more specific BB-related sources carries higher uncertainties.
24   Likewise, the relative sensitivities of sum_DaySOAs$_{S,C}$ and sum_NightSOAs$_{S,C}$ are less uncertain
25   compared to individual corresponding SOA factors in summer (as shown in Fig. S32 and Fig. S34a).
26   This contrast suggests that coarse classifications of factors may have higher precision, but provide less
27   information, whereas fine classifications of factors may have higher uncertainties, but potentially
28   provide more information from each factor. It also suggests that, at least for these datasets, factor mixing
29   occurs primarily between factors with closely related sources. Despite their higher uncertainties, the
30   finest classification levels explored here still appear to be meaningful. We also note that both datasets
31   investigated here are of relatively short duration, and factor separation may improve in longer datasets.

32   The final group of factors in Fig. 7 corresponds to SOA. The relative sensitivities of the SOA factors in
33   winter are shown to be lower than any of the SOA factors resolved during summer. This is consistent
34   with expectations regarding the seasonal differences in the dominant SOA precursors and the expected
35   AS_k of the resulting SOA. At this site, SOA precursors are expected to be dominated by monoterpenes
36   in summer, and biomass burning (increasing the contribution of phenols, naphthalenes, and other
37   aromatics) in winter, with traffic making a lesser contribution in both seasons (Daellenbach et al., 2016;
38   Qi et al., 2020). This is supported by analysis of the characteristics of the retrieved factors as discussed
39   above (Qi et al., 2019; Stefenelli et al., 2019). Previous studies have shown differences in the EESI-
40   TOF bulk sensitivity to SOA from different precursors, with terpene-derived SOA generally exhibiting
41   higher sensitivity than SOA from light aromatics (Lopez-Hilfiker et al., 2019; Wang et al., 2021). Figure
42   9 shows the $AS/\overline{AS_{COA}}$ for two DaySOA$_{S,C}$ and NightSOA$_{S,C}$ factors in summer, as well as the
43   sum_DaySOAs$_{S,C}$ and sum_NightSOAs$_{S,C}$, which are the aggregates of the individual DaySOA$_{S,C}$ and
44   NightSOA$_{S,C}$ factors (sum_DaySOAs$_{S,C}$ = DaySOA1$_{S,C}$ + DaySOA2$_{S\_C}$; and sum_NightSOAs$_{S,C}$ =
45   NightSOA1$_{S,C}$ + NightSOA2$_{S,C}$), respectively, and two SOA$_{W,C}$ factors in winter as a function of their

1   H:C ratio calculated from the EESI-TOF component. A trend of increasing sensitivity with increasing
2   H:C ratio is observed for the summer SOAs, as well as the winter SOAs ($SOA1_{w,C}$ and $SOA2_{w,C}$).

3   For the SOA factors, we compare AS_k retrieved to AS_k predicted using a molecular formula-based
4   parameterisation trained with laboratory SOA measurements, as described in Sect. 2.2.3 (Wang et al.,
5   2021). No parameterisations presently exist for POA factors, so these are excluded from the comparison,
6   although to allow comparison between campaigns the model is used to calculate a reference value for
7   AS_{COA}. Figure 10 compares the AS_k values based on model predictions against values determined from
8   cPMF. For summer SOAs, the LMN (limonene)-based parameterisation is applied as a surrogate for
9   terpene oxidation products. Regarding the winter SOAs, three scenarios (cresol, LMN and TMB) are
10   applied, as the winter SOAs in Zurich are mainly related to oxidation of biomass burning emissions,
11   which include monoterpenes, phenols, naphthalenes, and other aromatics (Bruns et al., 2016; Kelly et
12   al., 2018; Rouvière et al., 2006). In Fig. 10, 1:1, 1:2, 1:4, and 1:8 lines are provided to guide the eye,
13   although a 1:1 correspondence is not expected because the models are not trained on primary COA. The
14   figure shows a monotonic increase in model sensitivity predictions with increasing cPMF-derived
15   sensitivities, with the sole exception of $SOA2_{w,C}$. Specifically, the summer-derived points fall mainly
16   between the 1:1 and 1:2 lines, while for $SOA1_{w,C}$, the model predictions are roughly a factor of 2 lower
17   relative to the cPMF results. This offset may reflect differences in the appropriateness of the selected
18   precursor surrogate. The $SOA2_{w,C}$ factor is a slight outlier, probably because the AS_k for this factor is
19   more uncertain than the others (and not fully captured by the error bars in Fig. 7) due to the high
20   contribution from inorganic nitrate (~80 % of mass) in its factor profile. Given the limitations of the
21   multi-variate parameterisation (see Sect. 2.2.3) and the several orders of magnitude variation in EESI-
22   TOF sensitivities to individual compounds, the qualitative agreement between AS_k values independently
23   retrieved from multivariate parameterisation and cPMF provide support for both methods.


24
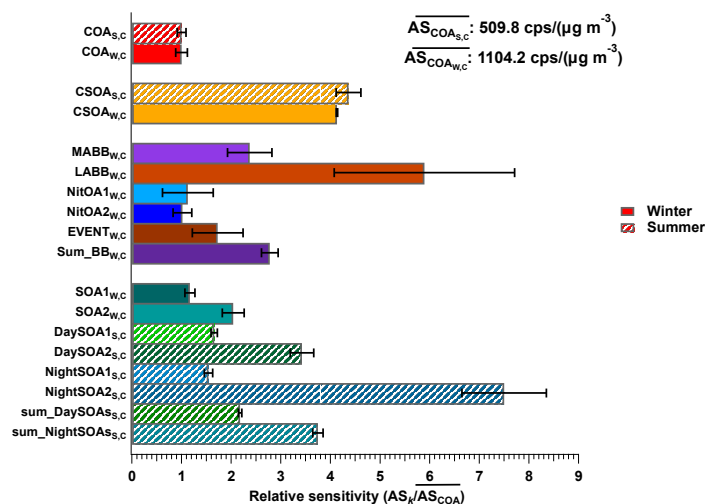
Figure 7. Comparison of $AS_k/\overline{AS_{COA_C}}$ of different factors resolved from the cPMF on the summer and
winter datasets. Mean values are shown as bars, and error bars indicate the standard deviation over all
accepted bootstrap runs. The following factor aggregations are also shown: $Sum\_BB_{w,C} = MABB_{w,C}$
$+ LABB_{w,C} + NitOA1_{w,C} + NitOA2_{w,C} + EVENT_{w,C}$; $sum\_DaySOAs_{s,C} = DaySOA1_{s,C} +$
$DaySOA2_{s\_C}$; and $sum\_NightSOAs_{s,C} = NightSOA1_{s,C} + NightSOA2_{s,C}$.
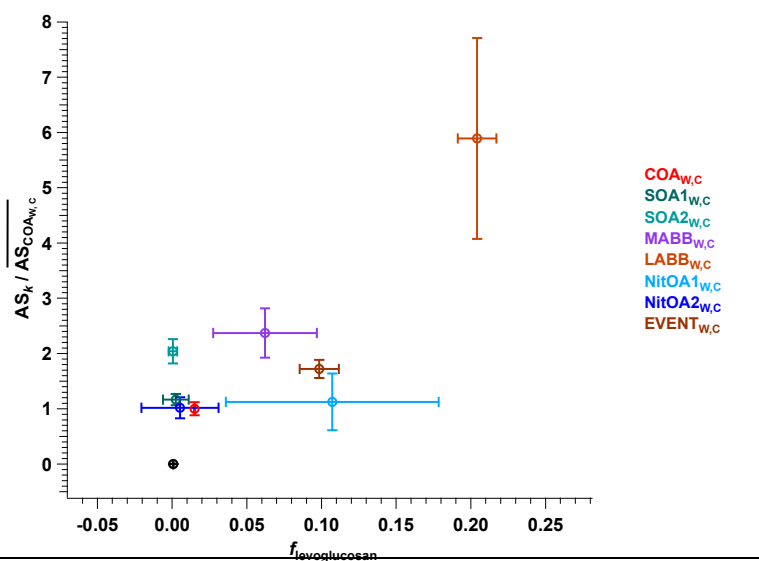
1

2   Figure 8. Relative apparent sensitivity $AS_k/\overline{AS_{COA_{W,C}}}$ as a function of levoglucosan fraction for all

3   factors resolved from the cPMF of the winter dataset except $CSOA_{W,C}$. Error bars denote standard

4   deviation.
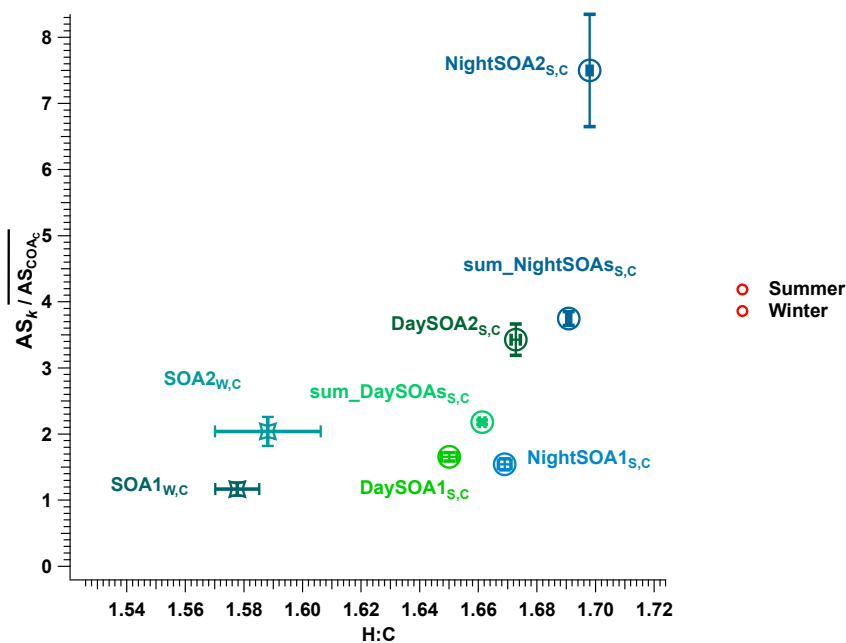
5

6   Figure 9. $AS_k/\overline{AS_{COA_C}}$ of SOA factors retrieved from the summer and winter datasets as a function of

7   the H:C ratio. Error bars denote standard deviation across all accepted runs.
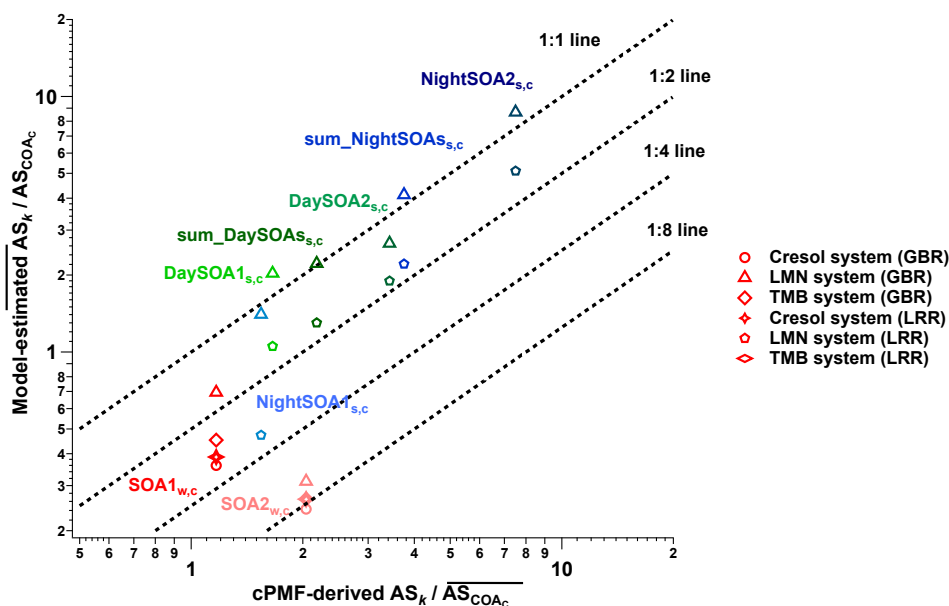
Figure 10. The estimated relative apparent sensitivity to COA ($AS_k/\overline{AS_{COA_C}}$) from the gradient boosting regression (GBR) and linear ridge regression (LRR) models as a function of cPMF-derived $AS_k/\overline{AS_{COA_C}}$). The symbols indicate the different oxidation-precursor system (LMN for SOA produced from oxidation of limonene by ozone, cresol and TMB for SOA produced from oxidation of *o*-cresol and 1,3,5-trimethylbenzene by OH radicals, respectively).

4. Atmospheric implications

The application of factor-dependent sensitivities can qualitatively and quantitatively affect the source apportionment results. Figures 11a and 11b compare the source apportionment results from cPMF on the summer and winter datasets using the calculated factor sensitivities (AS_k) (i.e., direct outputs of the cPMF analysis) vs. using a single bulk sensitivity (AS_{bulk}) for all factors, where the latter is calculated as the ratio of the total OA measured by the EESI-TOF (cps) to that measured by the AMS ($\mu g \ m^{-3}$). Figures 12a and 12b compare the total OA concentrations returned from the cPMF using AS_k and AS_{bulk} to the total OA measured by the AMS. Table S3 summarises the retrieved AS_k values for each factor (note that although the relative AS_k are believed to be intrinsic properties of the factors, the absolute sensitivities are instrument- and tuning-dependent, and will vary between campaigns).

In the Zurich summer campaign, the bulk OA sensitivity $AS_{bulk_{S,C}}$ (1254.0 cps /($\mu g \ m^{-3}$)) is higher than that of $AS_{COA_{S,C}}$ (509.8 cps /($\mu g \ m^{-3}$)). Four factors (HOA$_{S,C}$, COA$_{S,C}$, DaySOA1$_{S,C}$ and NightSOA1$_{S,C}$) are underestimated, whereas three factors (CSOA$_{S,C}$, DaySOA2$_{S,C}$ and NightSOA2$_{S,C}$) are overestimated when $AS_{bulk_{S,C}}$ is used. Using the calculated AS_k, the contribution of COA$_{S,C}$ to total OA more than doubles, from 4.5 % to 11.7 % as shown in Fig. 11a). Similarly, the application of AS_k increases the contributions of DaySOA1$_{S,C}$ and NightSOA1$_{S,C}$ from 22.7 % to 35.2 %, and from 10.3 % to 17.1 %, respectively. Among the overestimated factors, the largest decrease post-correction is found for NightSOA2$_{S,C}$, the contribution of which decreases by approximately a factor of three (from 29.7 %
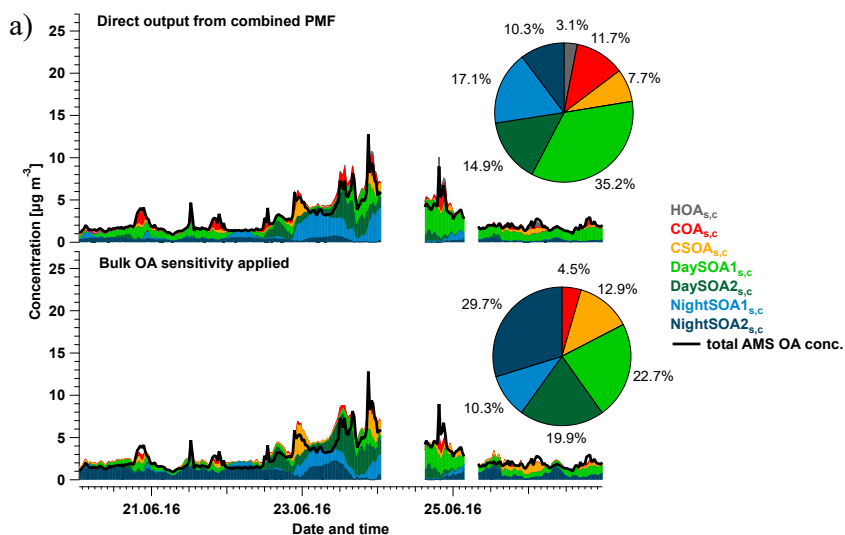
Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

1   to 10.3%). Smaller post-correction decreases are observed for the contributions of $CSOA_{S,C}$ (12.9 % to

2   7.7 %) and $DaySOA2_{S,C}$ (19.9 % to 14.9 %). If factor-dependent sensitivities were ignored,

3    $NightSOA2_{S,C}$ would be the largest contributor to total OA, followed by $DaySOA1_{S,C}$ whereas the full

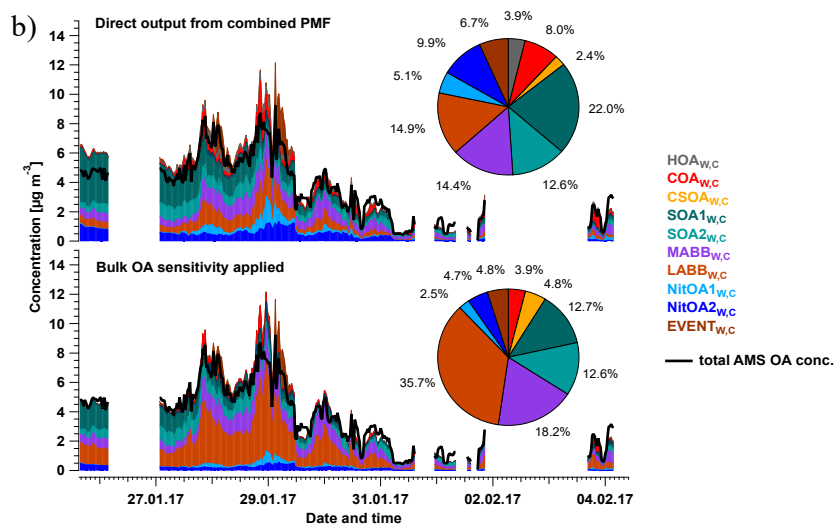4   analysis indicates that $DaySOA1_{S,C}$ is the largest contributor.

5

6   Similar to the summer campaign, application of AS_k significantly affects the source apportionment

7   results in winter. $CSOA_{W,C}$, $MABB_{W,C}$, and $LABB_{W,C}$ are shown to be overestimated, while $HOA_{W,C}$,

8    $COA_{W,C}$, $SOA1_{W,C}$, $NitOA1_{W,C}$, $NitOA2_{W,C}$ and $EVENT_{W,C}$ are underestimated. If factor-dependent

9   sensitivities were not considered, $LABB_{W,C}$ and $MABB_{W,C}$ would appear to be the dominant

10   contributors to total OA (35.7 % and 18.2 % respectively) due to their high levoglucosan content.

11   However, the full cPMF analysis indicates the $LABB_{W,C}$ and $MABB_{W,C}$ contributions to be 14.9 % and

12   14.4 %, respectively, whereas accounting for AS_k increases the contribution of $SOA1_{W,C}$ from 12.7 %

13   to 22.0 %, making it the largest contributor.

14

15   For both the summer and winter datasets, calculation of total OA from cPMF results using factor-

16   specific AS_k significantly outperforms that using a single AS_{bulk}. This is evident from an increased r^2

17   (0.966 vs 0.821) for summer. However, the r^2 is similar between the two approaches in winter (0.947

18   vs 0.943). The difference after applying AS_k and AS_{bulk} in r^2 might be related to the extent which the

19   contribution from factors with high AS_k and low AS_k to total OA changes over the time during the

20   campaign, which can vary in different datasets.

21

22   Box-and-whisker diagrams of factor contributions to total OA with/without applying AS_k values for

23   summer and winter are presented in Fig. 13. In the Zurich summer campaign, the box plots of the

24   corrected contributions of all six factors fall completely outside of the interquartile range (IQR) of the

25   uncorrected results, suggesting that the use of a single AS_{bulk} would lead to significant biases. In contrast,

26   the winter campaign exhibits a lack of overlap between the AS_k and AS_{bulk}-derived results for eight

27   factors ($HOA_{W,C}$, $COA_{W,C}$, $CSOA_{W,C}$, $SOA1_{W,C}$, $SOA2_{W,C}$, $NitOA1_{W,C}$, $NitOA2_{W,C}$ and $EVENT_{W,C}$) ,

28   whereas two factors overlap ($SOA2_{W,C}$ and $MABB_{W,C}$,). This may result from statistical uncertainties

29   in bootstrap analysis coupled with a less robust division between certain factors, yielding a wide

30   distribution, e.g., $MABB_{W,C}$, and/or AS_k values that are similar to AS_{bulk} (2271.1 cps /(μg m^{-3})), e.g.,

31    $SOA2_{W,C}$ (2253.2 cps /(μg m^{-3})), and $MABB_{W,C}$ (2619.0 cps /(μg m^{-3})).

Figure 11. Comparison of source apportionment results between direct output from cPMF (i.e., accounting for factor-dependent sensitivities) and application of a single bulk OA sensitivity, applied to the Zurich summer (a) and winter (b) datasets. Stack plots of factor time series directly from combined PMF and factor time series calculated from bulk OA sensitivity compared with total AMS OA concentration are shown in the upper and lower panel, respectively in each subfigure, together with the corresponding factor contribution shown in the pie chart. Note that here the contribution of the InorgNit factor and the contributions of NO^+ and NO_2^+ from inorganic nitrate in each factor are excluded to account only for the total OA.

a)

b)
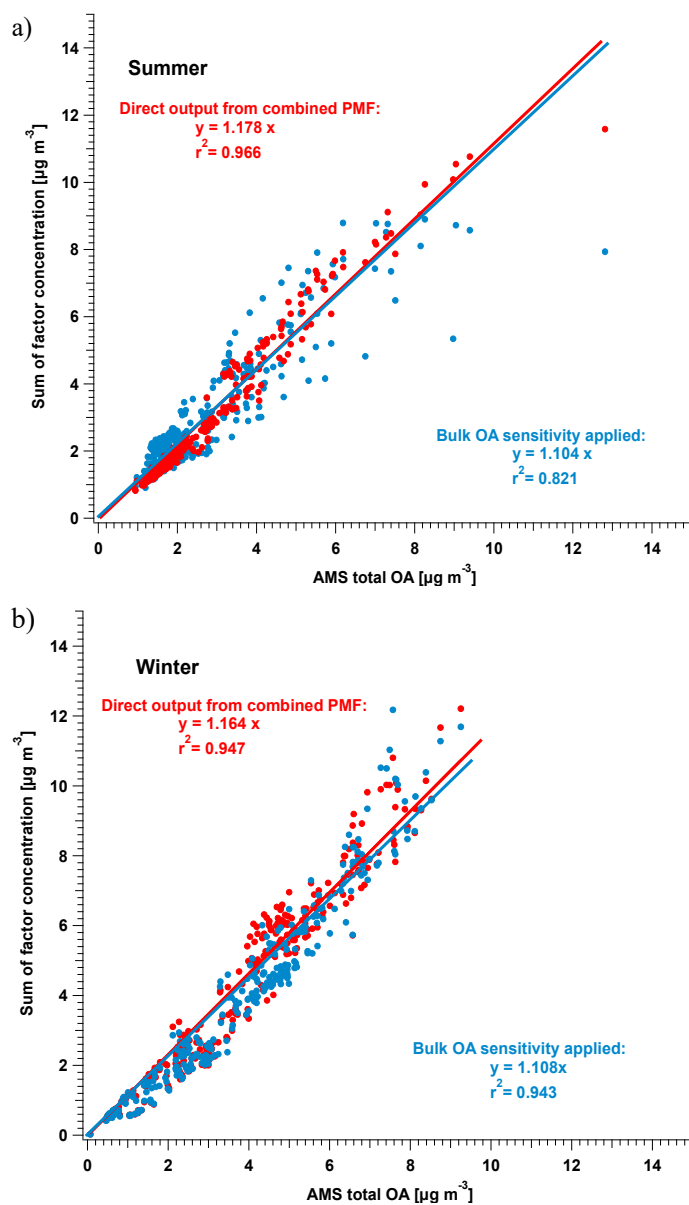
Figure 12. Comparison between the sum of factor concentrations in each time point with (in red) and without (in blue) taking the factor-dependent sensitivity into account and total OA measured by AMS for summer in a) and winter in b). A linear fit is conducted based on the Levenberg-Marquardt least orthogonal distance method. Note that here the contribution of the InorgNit factor and the contributions of NO^+ and NO_2^+ from inorganic nitrate in each factor are excluded.
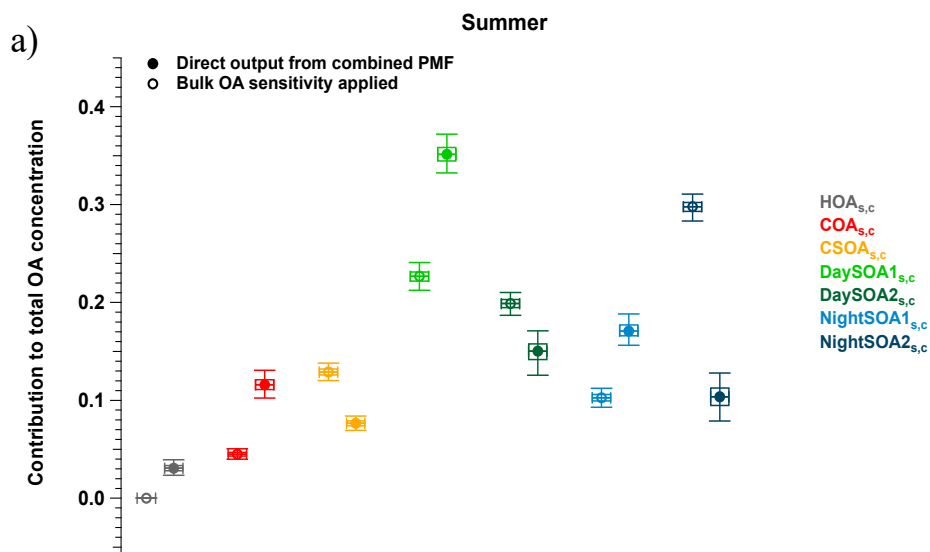
1
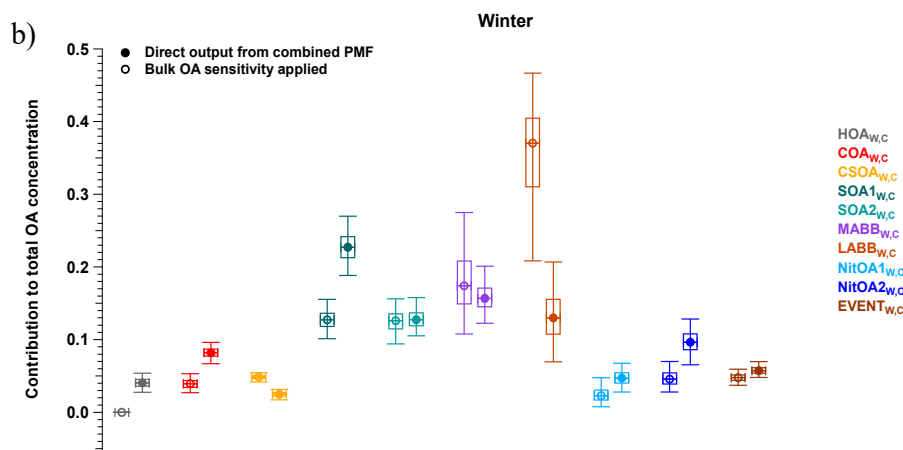
2
3
4 Figure 13. Box-and-whisker diagrams of factor contribution to total OA with/without applying the
5 factor dependent sensitivities, for summer in a) and winter in b) within accepted solutions. For each pair
6 of factors, the contribution without factor-dependent sensitivity applied is shown in the left box (open
7 symbols), whereas the contribution corrected by factor-dependent sensitivity is shown in the right box
8 (full symbols). The plot shows the mean (open/filled circle), median (horizontal bar), interquartile range
9 (rectangle), and minimum/maximum values (whiskers). Note that here the contribution of InorgNit
10 factor and contribution of NO^+ and NO_2^+ from inorganic nitrate in each factor are excluded.
11

1 **5. Conclusion**
2

3 We address the longstanding challenges in achieving quantitative source apportionment of SOA sources
4 by conducting a positive matrix factorisation (PMF) analysis of a dataset combining measurements from
5 an aerosol mass spectrometer (AMS) and an extractive electrospray ionisation time-of-flight mass
6 spectrometer (EESI-TOF). This approach combines the strengths of the two instruments, namely the
7 quantification ability of the AMS and the chemical resolution of the EESI-TOF. We demonstrate the
8 utility of this approach by PMF analysis of combined EESI-TOF/AMS datasets collected during
9 summer and winter in Zurich, Switzerland. The results retain the chemical resolution of the standalone
10 EESI-TOF PMF, while additionally providing quantitative factor time series and the EESI-TOF bulk
11 sensitivity to different OA factors.

12 We present a general procedure to conduct source apportionment on a combined dataset, including a
13 new metric for ensuring both instruments are well-represented in the solution, a method for optionally
14 constraining factor profiles for one or both instruments, and a protocol for uncertainty analysis. The
15 balancing metric references residual distributions obtained from cPMF to those of optimised single-
16 instrument PMF solutions to avoid bias due to differing instrument characteristics or error models.
17 Factor profile constraints require the construction of a reference profile, which may be challenging in a
18 multi-instrument dataset. We therefore provide methods for reference profile construction for cases
19 when (1) a single reference profile exists combining data from both instruments; (2) reference profile
20 exist independently for each instrument; and (3) a factor is detectable by one instrument but not the
21 other. To explore the solution stability and the uncertainties, a protocol for combined bootstrap
22 analysis/constraint exploration is developed.

23 The cPMF method intrinsically provides factor-dependent sensitivities (cps / (ug/m^3)) for the EESI-TOF.
24 To account for organonitrate content, the AMS ions NO^+ and NO_2^+ are included in the cPMF analysis.
25 Organic and inorganic contributions to these ions are estimated on a factor-by-factor basis using the
26 method of Kiendler-Scharr et al. (2016).

27 For practical reasons, sensitivities between winter and summer campaigns are compared using cooking-
28 related OA (COA) as a common reference. The retrieved factor sensitivities range from approximately
29 1.3 to 7.5 times the sensitivity of COA. The relative sensitivities of SOA factors are precursor-
30 dependent, and qualitatively consistent with trends observed in lab measurements of SOA from single
31 precursors (Lopez-Hilfiker et al., 2019). The SOA sensitivities estimated using our cPMF approach also
32 agree with the sensitivities predicted by multi-variate regression models (Wang et al. 2021), which
33 further demonstrates that SOA sensitivities are precursor- and/or source-dependent. Comparison of
34 source apportionment results using factor-dependent sensitivities to uncorrected results show
35 substantial differences, highlighting the importance of quantitative analysis. For example, before
36 applying factor-dependent sensitivities, the contribution of a daytime SOA factor is underestimated by
37 about 30 % (22.7 % before vs 35.2 % after), whereas the contribution of a nighttime SOA factor is
38 almost overestimated by a factor of 3 in the summer campaign (29.7 % before vs 10.3 % after). As for
39 the winter campaign, the contribution of less-aged biomass burning factor to total OA in Zurich winter
40 dataset is 35.7 %, making it a major factor in winter without considering its factor-dependent sensitivity.
41 However, this factor is significantly overestimated by more than a factor of 2 (35.7 %, before vs 14.9 %
42 after). In contrast, the SOA1 factor in winter is underestimated, with its contribution increasing from
43 12.7% to 22.0 %.

1   The cPMF method presented herein is can be utilised as-is not only for the AMS/EESI-TOF
2   combination, but to any dataset comprising data from multiple instruments. As such, it provides a
3   promising strategy for utilising instruments with high chemical resolution but semi-quantitative
4   performance (i.e., a linear but hard-to-calibrate response to mass) within the framework of a quantitative
5   source apportionment.

6

7

21

22

23

1

2 Reference

3 Alfarra, M. R., Prevot, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D.,
4 Mohr, M., and Baltensperger, U.: Identification of the mass spectral signature of organic aerosols
5 from wood burning emissions, Environ. Sci. Technol., 41, 5770-5777,
6 https://doi.org/10.1021/es062289b, 2007.
7 Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and
8 Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer: 1. Techniques
9 of data interpretation and error analysis (vol 108, art no 4090, 2003), J. Geophys. Res. Atmos., 108,
10 4090, https://doi.org/10.1029/2002JD002358, 2003.
11 Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., Wolf,
12 K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W. W., Katsouyanni, K.,
13 Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A. S., Lanki, T., Turunen, A.,
14 Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N. L., Ostenson, C. G., Fratiglioni, L.,
15 Penell, J., Korek, M., Pershagen, G., Eriksen, K. T., Overvad, K., Ellermann, T., Eeftens, M., Peeters,
16 P. H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Kramer, U., Heinrich, J., de
17 Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-
18 Hensch, N., Kunzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-
19 Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M. Y., Ricceri, F., Sacerdote, C., Galassi, C.,
20 Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P.,
21 Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., and Hoek, G.: Effects of long-term
22 exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the
23 multicentre ESCAPE project, Lancet, 383, 785-795, https://doi.org/10.1016/S0140-6736(13)62158-3,
24 2014.
25 Bi, C., Krechmer, J. E., Frazier, G. O., Xu, W., Lambe, A. T., Claflin, M. S., Lerner, B. M., Jayne, J.
26 T., Worsnop, D. R., Canagaratna, M. R., and Isaacman-VanWertz, G.: Quantification of isomer-
27 resolved iodide chemical ionization mass spectrometry sensitivity and uncertainty using a voltage-
28 scanning approach, Atmos. Meas. Tech., 14, 6835-6850, https://doi.org/10.5194/amt-14-6835-2021,
29 2021.
30 Boyd, C. M., Sanchez, J., Xu, L., Eugene, A. J., Nah, T., Tuet, W. Y., Guzman, M. I., and Ng, N. L.:
31 Secondary organic aerosol formation from the β-pinene+NO$_3$ system: effect of humidity
32 and peroxy radical fate, Atmos. Chem. Phys., 15, 7497-7522, https://doi.org/10.5194/acp-15-7497-
33 2015, 2015.
34 Brown, W. L., Day, D. A., Stark, H., Pagonis, D., Krechmer, J. E., Liu, X., Price, D. J., Katz, E. F.,
35 DeCarlo, P. F., Masoud, C. G., Wang, D. S., Hildebrandt Ruiz, L., Arata, C., Lunderberg, D. M.,
36 Goldstein, A. H., Farmer, D. K., Vance, M. E., and Jimenez, J. L.: Real-time organic aerosol chemical
37 speciation in the indoor environment using extractive electrospray ionization mass spectrometry,
38 Indoor Air, 31, 141-155, https://doi.org/10.1111/ina.12721, 2021.
39 Bruns, E. A., El Haddad, I., Slowik, J. G., Kilic, D., Klein, F., Baltensperger, U., and Prévôt, A. S. H.:
40 Identification of significant precursor gases of secondary organic aerosols from residential wood
41 combustion, Scientific Reports, 6, 27881, https://doi.org/10.1038/srep27881, 2016.
42 Bruns, E. A., Perraud, V., Zelenyuk, A., Ezell, M. J., Johnson, S. N., Yu, Y., Imre, D., Finlayson-
43 Pitts, B. J., and Alexander, M. L.: Comparison of FTIR and Particle Mass Spectrometry for the
44 Measurement of Particulate Organic Nitrates, Environ. Sci. Technol., 44, 1056-1061,
45 https://doi.org/10.1021/es9029864, 2010.
46 Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B.,
47 Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M.
48 J., DeCarlo, P. F., Kolb, C. E., Davidovits, P., and Worsnop, D. R.: Chemical and microphysical
49 characterization of ambient aerosols with the aerodyne aerosol mass spectrometer, Mass Spectrom.
50 Rev., 26, 185-222, https://doi.org/10.1002/mas.20115, 2007.
51 Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prevot, A. S. H.: SoFi, an IGOR-
52 based interface for the efficient use of the generalized multilinear engine (ME-2) for the source

apportionment: ME-2 application to aerosol mass spectrometer data, Atmos. Meas. Tech., 6, 3649-3661, https://doi.org/10.5194/amt-6-3649-2013, 2013.

Canonaco, F., Tobler, A., Chen, G., Sosedova, Y., Slowik, J. G., Bozzetti, C., Daellenbach, K. R., El Haddad, I., Crippa, M., Huang, R. J., Furger, M., Baltensperger, U., and Prévôt, A. S. H.: A new method for long-term source apportionment with time-dependent factor profiles and uncertainty assessment using SoFi Pro: application to 1 year of organic aerosol data, Atmos. Meas. Tech., 14, 923-943, https://doi.org/10.5194/amt-14-923-2021, 2021.

Chen, Y., Takeuchi, M., Nah, T., Xu, L., Canagaratna, M. R., Stark, H., Baumann, K., Canonaco, F., Prévôt, A. S. H., Huey, L. G., Weber, R. J., and Ng, N. L.: Chemical characterization of secondary organic aerosol at a rural site in the southeastern US: insights from simultaneous high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) and FIGAERO chemical ionization mass spectrometer (CIMS) measurements, Atmos. Chem. Phys., 20, 8421-8440, https://doi.org/10.5194/acp-20-8421-2020, 2020.

Chow, J. C., Bachmann, J. D., Wierman, S. S. G., Mathai, C. V., Malm, W. C., White, W. H., Mueller, P. K., Kumar, N., and Watson, J. G.: Visibility: Science and regulation - Discussion, J. Air Waste Manag. Assoc., 52, 973-999, https://doi.org/10.1016/j.atmosenv.2010.09.060, 2002.

Crippa, M., El Haddad, I., Slowik, J. G., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Sciare, J., Baltensperger, U., and Prévôt, A. S. H.: Identification of marine and continental aerosol sources in Paris using high resolution aerosol mass spectrometry, J. Geophys. Res. Atmos., 118, 1950-1963, https://doi.org/10.1002/jgrd.50151, 2013a.

Crippa, M., Canonaco, F., Slowik, J. G., El Haddad, I., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Temime-Roussel, B., Abidi, E., Poulain, L., Wiedensohler, A., Baltensperger, U., and Prévôt, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, Atmos. Chem. Phys., 13, 8411-8426, https://doi.org/10.5194/acp-13-8411-2013, 2013b.

Crippa, M., Canonaco, F., Lanz, V. A., Aijala, M., Allan, J. D., Carbone, S., Capes, G., Ceburnis, D., Dall'Osto, M., Day, D. A., DeCarlo, P. F., Ehn, M., Eriksson, A., Freney, E., Hildebrandt Ruiz, L., Hillamo, R., Jimenez, J. L., Junninen, H., Kiendler-Scharr, A., Kortelainen, A. M., Kulmala, M., Laaksonen, A., Mensah, A., Mohr, C., Nemitz, E., O'Dowd, C., Ovadnevaite, J., Pandis, S. N., Petaja, T., Poulain, L., Saarikoski, S., Sellegri, K., Swietlicki, E., Tiitta, P., Worsnop, D. R., Baltensperger, U., and Prevot, A. S. H.: Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach, Atmos. Chem. Phys., 14, 6159-6176, https://doi.org/10.5194/acp-14-6159-2014, 2014.

Cubison, M. J. and Jimenez, J. L.: Statistical precision of the intensities retrieved from constrained fitting of overlapping peaks in high-resolution mass spectra, Atmos. Meas. Tech., 8, 2333-2345, https://doi.org/10.5194/amt-8-2333-2015, 2015.

Daellenbach, K. R., Bozzetti, C., Krepelova, A. K., Canonaco, F., Wolf, R., Zotter, P., Fermo, P., Crippa, M., Slowik, J. G., Sosedova, Y., Zhang, Y., Huang, R. J., Poulain, L., Szidat, S., Baltensperger, U., El Haddad, I., and Prevot, A. S. H.: Characterization and source apportionment of organic aerosol using offline aerosol mass spectrometry, Atmos. Meas. Tech., 9, 23-39, https://doi.org/10.5194/amt-9-23-2016, 2016.

Daellenbach, K. R., Stefenelli, G., Bozzetti, C., Vlachou, A., Fermo, P., Gonzalez, R., Piazzalunga, A., Colombi, C., Canonaco, F., Hueglin, C., Kasper-Giebl, A., Jaffrezo, J. L., Bianchi, F., Slowik, J. G., Baltensperger, U., El-Haddad, I., and Prevot, A. S. H.: Long-term chemical analysis and organic aerosol source apportionment at nine sites in central Europe: source identification and uncertainty assessment, Atmos.Chem.Phys.,17, 13265-13282, https://doi.org/10.5194/acp-17-13265-2017, 2017.

Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L.-E., Leni, Z., Vlachou, A., Stefenelli, G., Canonaco, F., Weber, S., Segers, A., Kuenen, J. J. P., Schaap, M., Favez, O., Albinet, A., Aksoyoglu, S., Dommen, J., Baltensperger, U., Geiser, M., El Haddad, I., Jaffrezo, J.-L., and Prévôt, A. S. H.: Sources of particulate-matter air pollution and its oxidative potential in Europe, Nature, 587, 414-419, https://doi.org/10.1038/s41586-020-2902-8, 2020.

Davison, A. C. and Hinkley, D. V.: Bootstrap methods and their application, Cambridge University Press, Cambridge, New York, NY, USA, https://doi.org/10.1017/CBO9780511802843, 1997.

DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R., and Jimenez, J. L.: Field-deployable, high-

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

1 resolution, time-of-flight aerosol mass spectrometer, Anal. Chem., 78, 8281-8289,
2 https://doi.org/10.1021/ac061249n, 2006.
3 Elser, M., Huang, R. J., Wolf, R., Slowik, J. G., Wang, Q. Y., Canonaco, F., Li, G. H., Bozzetti, C.,
4 Daellenbach, K. R., Huang, Y., Zhang, R. J., Li, Z. Q., Cao, J. J., Baltensperger, U., El-Haddad, I.,
5 and Prevot, A. S. H.: New insights into PM2.5 chemical composition and sources in two major cities
6 in China during extreme haze events using aerosol mass spectrometry, Atmos. Chem. Phys., 16, 3207-
7 3225, https://doi.org/10.5194/acp-16-3207-2016, 2016.
8 Farmer, D. K., Matsunaga, A., Docherty, K. S., Surratt, J. D., Seinfeld, J. H., Ziemann, P. J., and
9 Jimenez, J. L.: Response of an aerosol mass spectrometer to organonitrates and organosulfates and
10 implications for atmospheric chemistry, Proc. Natl. Acad. Sci. U S A, 107, 6670,
11 https://doi.org/10.1073/pnas.0912340107, 2010.
12 Faxon, C., Hammes, J., Le Breton, M., Pathak, R. K., and Hallquist, M.: Characterization of organic
13 nitrate constituents of secondary organic aerosol (SOA) from nitrate-radical-initiated oxidation of
14 limonene using high-resolution chemical ionization mass spectrometry, Atmos. Chem. Phys., 18,
15 5467-5481, https://doi.org/10.5194/acp-18-5467-2018, 2018.
16 Fenger, J.: Urban air quality, Atmos. Environ., 33, 4877-4900, https://doi.org/10.1016/S1352-
17 2310(99)00290-3, 1999.
18 Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J.,
19 Lowe, D. C., and Myhre, G.: Changes in atmospheric constituents and in radiative forcing. Chapter 2,
20 in: Climate Change 2007. The Physical Science Basis, 2007.
21 Fry, J. L., Kiendler-Scharr, A., Rollins, A. W., Wooldridge, P. J., Brown, S. S., Fuchs, H., Dubé, W.,
22 Mensah, A., dal Maso, M., Tillmann, R., Dorn, H. P., Brauers, T., and Cohen, R. C.: Organic nitrate
23 and secondary organic aerosol yield from NO_3 oxidation of β-pinene evaluated using a
24 gas-phase kinetics/aerosol partitioning model, Atmos. Chem. Phys., 9, 1431-1449,
25 https://doi.org/10.5194/acp-9-1431-2009, 2009.
26 Fry, J. L., Kiendler-Scharr, A., Rollins, A. W., Brauers, T., Brown, S. S., Dorn, H. P., Dubé, W. P.,
27 Fuchs, H., Mensah, A., Rohrer, F., Tillmann, R., Wahner, A., Wooldridge, P. J., and Cohen, R. C.:
28 SOA from limonene: role of NO_3 in its generation and degradation, Atmos. Chem.
29 Phys., 11, 3879-3894, https://doi.org/10.5194/acp-11-3879-2011, 2011.
30 Fuller, S. J., Wragg, F. P. H., Nutter, J., and Kalberer, M.: Comparison of on-line and off-line
31 methods to quantify reactive oxygen species (ROS) in atmospheric aerosols, Atmos. Environ., 92, 97-
32 103, https://doi.org/10.1016/j.atmosenv.2014.04.006, 2014.
33 Halliwell, B. and Cross, C. E.: Oxygen-derived species: their relation to human disease and
34 environmental stress, Environ. Health Persp., 102 Suppl 10, 5-12,
35 https://doi.org/10.1289/ehp.94102s105, 1994.
36 Hu, W. W., Hu, M., Hu, W., Jimenez, J. L., Yuan, B., Chen, W. T., Wang, M., Wu, Y. S., Chen, C.,
37 Wang, Z. B., Peng, J. F., Zeng, L. M., and Shao, M.: Chemical composition, sources, and aging
38 process of submicron aerosols in Beijing: Contrast between summer and winter, J. Geophys. Res.
39 Atmos., 121, 1955-1977, https://doi.org/10.1002/2015JD024020, 2016.
40 Jimenez, J. L., Jayne, J. T., Shi, Q., Kolb, C. E., Worsnop, D. R., Yourshaw, I., Seinfeld, J. H.,
41 Flagan, R. C., Zhang, X., Smith, K. A., Morris, J. W., and Davidovits, P.: Ambient aerosol sampling
42 using the Aerodyne Aerosol Mass Spectrometer, J. Geophys. Res. Atmos., 108,
43 https://doi.org/10.1029/2001JD001213, 2003.
44 Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H.,
45 DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M.,
46 Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C.,
47 Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M.,
48 Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch,
49 T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann,
50 S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T.,
51 Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D.,
52 Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M.,
53 Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols in the
54 Atmosphere, Science, 326, 1525-1529, https://doi.org/10.1126/science.1180353, 2009.

Junninen, H., Ehn, M., Petäjä, T., Luosujärvi, L., Kotiaho, T., Kostiainen, R., Rohner, U., Gonin, M., Fuhrer, K., Kulmala, M., and Worsnop, D. R.: A high-resolution mass spectrometer to measure atmospheric ion composition, Atmos. Meas. Tech., 3, 1039-1053, https://doi.org/10.5194/amt-3-1039-2010, 2010.

Kelly, F. J. and Fussell, J. C.: Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter, Atmos. Environ., 60, 504-526, https://doi.org/10.1016/j.atmosenv.2012.06.039, 2012.

Kelly, J. M., Doherty, R. M., O'Connor, F. M., and Mann, G. W.: The impact of biogenic, anthropogenic, and biomass burning volatile organic compound emissions on regional and seasonal variations in secondary organic aerosol, Atmos. Chem. Phys., 18, 7393-7422, https://doi.org/10.5194/acp-18-7393-2018, 2018.

Kiendler-Scharr, A., Mensah, A. A., Friese, E., Topping, D., Nemitz, E., Prevot, A. S. H., Aijala, M., Allan, J., Canonaco, F., Canagaratna, M., Carbone, S., Crippa, M., Dall Osto, M., Day, D. A., De Carlo, P., Di Marco, C. F., Elbern, H., Eriksson, A., Freney, E., Hao, L., Herrmann, H., Hildebrandt, L., Hillamo, R., Jimenez, J. L., Laaksonen, A., McFiggans, G., Mohr, C., O'Dowd, C., Otjes, R., Ovadnevaite, J., Pandis, S. N., Poulain, L., Schlag, P., Sellegri, K., Swietlicki, E., Tiitta, P., Vermeulen, A., Wahner, A., Worsnop, D., and Wu, H. C.: Ubiquity of organic nitrates from nighttime chemistry in the European submicron aerosol, Geophys. Res. Lett., 43, 7735-7744, https://doi.org/10.1002/2016gl069239, 2016.

Kumar, V., Giannoukos, S., Haslett, S. L., Tong, Y., Singh, A., Bertrand, A., Lee, C. P., Wang, D. S., Bhattu, D., Stefenelli, G., Dave, J. S., Puthussery, J. V., Qi, L., Vats, P., Rai, P., Casotto, R., Satish, R., Mishra, S., Pospisilova, V., Mohr, C., Bell, D. M., Ganguly, D., Verma, V., Rastogi, N., Baltensperger, U., Tripathi, S. N., Prévôt, A. S. H., and Slowik, J. G.: Real-time chemical speciation and source apportionment of organic aerosol components in Delhi, India, using extractive electrospray ionization mass spectrometry, Atmos. Chem. Phys. Discuss., 2021, 1-35, https://doi.org/10.5194/acp-2021-1033, 2021.

Laden, F., Schwartz, J., Speizer, F. E., and Dockery, D. W.: Reduction in fine particulate air pollution and mortality - Extended follow-up of the Harvard six cities study, Am. J. Resp. Crit. Care, 173, 667-672, https://doi.org/10.1164/rccm.200503-443OC, 2006.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prévôt, A. S. H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, Atmos. Chem. Phys., 7, 1503-1522, https://doi.org/10.5194/acp-7-1503-2007, 2007.

Li, N., Sioutas, C., Cho, A., Schmitz, D., Misra, C., Sempf, J., Wang, M. Y., Oberley, T., Froines, J., and Nel, A.: Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage, Environ. Health Persp., 111, 455-460, https://doi.org/10.1289/ehp.6000, 2003.

Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, Atmos. Chem. Phys., 5, 715-737, https://doi.org/10.5194/acp-5-715-2005, 2005.

Lopez-Hilfiker, F. D., Iyer, S., Mohr, C., Lee, B. H., D'Ambro, E. L., Kurtén, T., and Thornton, J. A.: Constraining the sensitivity of iodide adduct chemical ionization mass spectrometry to multifunctional organic molecules using the collision limit and thermodynamic stability of iodide ion adducts, Atmos. Meas. Tech., 9, 1505-1512, https://doi.org/10.5194/amt-9-1505-2016, 2016.

Lopez-Hilfiker, F. D., Pospisilova, V., Huang, W., Kalberer, M., Mohr, C., Stefenelli, G., Thornton, J. A., Baltensperger, U., Prevot, A. S. H., and Slowik, J. G.: An extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF) for online measurement of atmospheric aerosol particles, Atmos. Meas.Tech.,12,4867-4886, https://doi.org/10.5194/amt-12-4867-2019, 2019.

Lopez-Hilfiker, F. D., Mohr, C., Ehn, M., Rubach, F., Kleist, E., Wildt, J., Mentel, T. F., Lutz, A., Hallquist, M., Worsnop, D., and Thornton, J. A.: A novel method for online analysis of gas and particle composition: description and evaluation of a Filter Inlet for Gases and AEROsols (FIGAERO), Atmos. Meas. Tech., 7, 983-1001, https://doi.org/10.5194/amt-7-983-2014, 2014.

Mayer, H.: Air pollution in cities, Atmos. Environ., 33, 4029-4037, https://doi.org/10.1016/S1352-2310(99)00144-2, 1999.

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, 258-271, https://doi.org/10.1080/02786826.2011.620041, 2012.

1 Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey,
2 A., Querol, X., Seco, R., Penuelas, J., Jimenez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U.,
3 and Prevot, A. S. H.: Identification and quantification of organic aerosol from cooking and other
4 sources in Barcelona using aerosol mass spectrometer data, Atmos. Chem. Phys., 12, 1649-1665,
5 https://doi.org/10.5194/acp-12-1649-2012, 2012.
6 Myhre, G., Shindell, D., and Pongratz, J.: Anthropogenic and natural radiative forcing, 2014.
7 Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Chhabra, P. S., Seinfeld, J. H., and Worsnop, D. R.:
8 Changes in organic aerosol composition with aging inferred from aerosol mass spectra, Atmos. Chem.
9 Phys.,11, 6465-6474, https://doi.org/10.5194/acp-11-6465-2011, 2011a.
10 Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-
11 Time Methods for Estimating Organic Component Mass Concentrations from Aerosol Mass
12 Spectrometer Data, Environ. Sci. Technol., 45, 910-916, https://doi.org/10.1021/es102951k, 2011b.
13 Paatero, P.: The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear
14 Problems, Including the n-Way Parallel Factor Analysis Model, J. Comput. Graph. Stat., 8, 854-888,
15 https://doi.org/10.1080/10618600.1999.10474853, 1999.
16 Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal
17 utilization of error estimates of data values, Environmetrics, 5, 111-126,
18 https://doi.org/10.1002/env.3170050203, 1994.
19 Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor
20 analytic solutions, Atmos. Meas. Tech., 7, 781-797, https://doi.org/10.5194/amt-7-781-2014, 2014.
21 Penner, J. E., Xu, L., and Wang, M. H.: Satellite methods underestimate indirect climate forcing by
22 aerosols, Proc. Natl. Acad. Sci. U S A, 108, 13404-13408, https://doi.org/10.1073/pnas.1018526108,
23 2011.
24 Pieber, S. M., El Haddad, I., Slowik, J. G., Canagaratna, M. R., Jayne, J. T., Platt, S. M., Bozzetti, C.,
25 Daellenbach, K. R., Frohlich, R., Vlachou, A., Klein, F., Dommen, J., Miljevic, B., Jimenez, J. L.,
26 Worsnop, D. R., Baltensperger, U., and Prevot, A. S. H.: Inorganic Salt Interference on CO2+ in
27 Aerodyne AMS and ACSM Organic Aerosol Composition Studies, Environ. Sci. Technol., 50, 10494-
28 10503, https://doi.org/10.1021/acs.est.6b01035, 2016.
29 Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D.: Lung
30 cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, Jama-J.
31 Am.Med.Assoc.,287, 1132-1141, https://doi.org/10.1001/jama.287.9.1132, 2002.
32 Qi, L., Chen, M. D., Stefenelli, G., Pospisilova, V., Tong, Y. D., Bertrand, A., Hueglin, C., Ge, X. L.,
33 Baltensperger, U., Prevot, A. S. H., and Slowik, J. G.: Organic aerosol source apportionment in
34 Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) -
35 Part 2: Biomass burning influences in winter, Atmos. Chem. Phys., 19, 8037-8062,
36 https://doi.org/10.5194/acp-19-8037-2019, 2019.
37 Qi, L., Vogel, A. L., Esmaeilirad, S., Cao, L., Zheng, J., Jaffrezo, J. L., Fermo, P., Kasper-Giebl, A.,
38 Daellenbach, K. R., Chen, M., Ge, X., Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G.: A 1-year
39 characterization of organic aerosol composition and sources using an extractive electrospray
40 ionization time-of-flight mass spectrometer (EESI-TOF), Atmos. Chem. Phys., 20, 7875-7893,
41 https://doi.org/10.5194/acp-20-7875-2020, 2020.
42 Reuter, S., Gupta, S. C., Chaturvedi, M. M., and Aggarwal, B. B.: Oxidative stress, inflammation, and
43 cancer How are they linked?, Free Radical Bio. Med., 49, 1603-1616,
44 https://doi.org/10.1016/j.freeradbiomed.2010.09.006, 2010.
45 Rollins, A. W., Kiendler-Scharr, A., Fry, J. L., Brauers, T., Brown, S. S., Dorn, H. P., Dubé, W. P.,
46 Fuchs, H., Mensah, A., Mentel, T. F., Rohrer, F., Tillmann, R., Wegener, R., Wooldridge, P. J., and
47 Cohen, R. C.: Isoprene oxidation by nitrate radical: alkyl nitrate and secondary organic aerosol yields,
48 Atmos. Chem. Phys., 9, 6685-6703, https://doi.org/10.5194/acp-9-6685-2009, 2009.
49 Rouvière, A., Brulfert, G., Baussand, P., and Chollet, J.-P.: Monoterpene source emissions from
50 Chamonix in the Alpine Valleys, Atmospheric Environment, 40, 3613-3620,
51 https://doi.org/10.1016/j.atmosenv.2005.09.058, 2006.
52 Slowik, J. G., Vlasenko, A., McGuire, M., Evans, G. J., and Abbatt, J. P. D.: Simultaneous factor
53 analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site,
54 Atmos. Chem. Phys., 10, 1969-1988, https://doi.org/10.5194/acp-10-1969-2010, 2010.

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

1 Stefenelli, G., Pospisilova, V., Lopez-Hilfiker, F. D., Daellenbach, K. R., Hüglin, C., Tong, Y.,
2 Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G.: Organic aerosol source apportionment in
3 Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) –
4 Part 1: Biogenic influences and day–night chemistry in summer, Atmos. Chem. Phys., 19, 14825-
5 14848, https://doi.org/10.5194/acp-19-14825-2019, 2019.
6 Struckmeier, C., Drewnick, F., Fachinger, F., Gobbi, G. P., and Borrmann, S.: Atmospheric aerosols
7 in Rome, Italy: sources, dynamics and spatial variations during two seasons, Atmos. Chem. Phys., 16,
8 15277-15299, https://doi.org/10.5194/acp-16-15277-2016, 2016.
9 Sun, Y. L., Wang, Z. F., Fu, P. Q., Yang, T., Jiang, Q., Dong, H. B., Li, J., and Jia, J. J.: Aerosol
10 composition, sources and processes during wintertime in Beijing, China, Atmos. Chem. Phys., 13,
11 4577-4592, https://doi.org/10.5194/acp-13-4577-2013, 2013.
12 Sun, Y. L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W. N., Bae, M. S., Hung, H. M.,
13 Hogrefe, O., Frank, B., Rattigan, O. V., and Lin, Y. C.: Characterization of the sources and processes
14 of organic and inorganic aerosols in New York city with a high-resolution time-of-flight aerosol mass
15 apectrometer, Atmos. Chem. Phys., 11, 1581-1602, https://doi.org/10.5194/acp-11-1581-2011, 2011.
16 Sun, Y. L., Du, W., Fu, P. Q., Wang, Q. Q., Li, J., Ge, X. L., Zhang, Q., Zhu, C. M., Ren, L. J., Xu,
17 W. Q., Zhao, J., Han, T. T., Worsnop, D. R., and Wang, Z. F.: Primary and secondary aerosols in
18 Beijing in winter: sources, variations and processes, Atmos. Chem. Phys., 16, 8309-8329,
19 https://doi.org/10.5194/acp-16-8309-2016, 2016a.
20 Sun, Y. L., Wang, Z. F., Wild, O., Xu, W. Q., Chen, C., Fu, P. Q., Du, W., Zhou, L. B., Zhang, Q.,
21 Han, T. T., Wang, Q. Q., Pan, X. L., Zheng, H. T., Li, J., Guo, X. F., Liu, J. G., and Worsnop, D. R.:
22 "APEC Blue": Secondary Aerosol Reductions from Emission Controls in Beijing, Sci. Rep-Uk., 6,
23 20668, https://doi.org/10.1038/srep20668 (2016), 2016b.
24 Talhout, R., Opperhuizen, A., and van Amsterdam, J. G. C.: Sugars as tobacco ingredient: Effects on
25 mainstream smoke composition, Food Chem. Toxicol., 44, 1789-1798,
26 https://doi.org/10.1016/j.fct.2006.06.016, 2006.
27 Tong, Y., Pospisilova, V., Qi, L., Duan, J., Gu, Y., Kumar, V., Rai, P., Stefenelli, G., Wang, L.,
28 Wang, Y., Zhong, H., Baltensperger, U., Cao, J., Huang, R. J., Prévôt, A. S. H., and Slowik, J. G.:
29 Quantification of solid fuel combustion and aqueous chemistry contributions to secondary organic
30 aerosol during wintertime haze events in Beijing, Atmos. Chem. Phys., 21, 9859-9886,
31 https://doi.org/10.5194/acp-21-9859-2021, 2021.
32 Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of
33 organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos.
34 Chem.Phys.,9, 2891-2918, https://doi.org/10.5194/acp-9-2891-2009, 2009.
35 Vlachou, A., Tobler, A., Lamkaddam, H., Canonaco, F., Daellenbach, K. R., Jaffrezo, J. L.,
36 Minguillon, M. C., Maasikmets, M., Teinemaa, E., Baltensperger, U., El Haddad, I., and Prevot, A. S.
37 H.: Development of a versatile source apportionment analysis based on positive matrix factorization:
38 a case study of the seasonal variation of organic aerosol sources in Estonia, Atmos. Chem. Phys., 19,
39 7279-7295, https://doi.org/10.5194/acp-19-7279-2019, 2019.
40 Xu, L., Suresh, S., Guo, H., Weber, R. J., and Ng, N. L.: Aerosol characterization over the
41 southeastern United States using high-resolution aerosol mass spectrometry: spatial and seasonal
42 variation of aerosol composition and sources with a focus on organic nitrates, Atmos. Chem. Phys.,
43 15, 7307-7336, https://doi.org/10.5194/acp-15-7307-2015, 2015.
44 Xu, W. Q., Sun, Y. L., Wang, Q. Q., Zhao, J., Wang, J. F., Ge, X. L., Xie, C. H., Zhou, W., Du, W.,
45 Li, J., Fu, P. Q., Wang, Z. F., Worsnop, D. R., and Coe, H.: Changes in Aerosol Chemistry From 2014
46 to 2016 in Winter in Beijing: Insights From High-Resolution Aerosol Mass Spectrometry, J. Geophys.
47 Res. Atmos., 124, 1132-1147, https://doi.org/10.1029/2018JD029245, 2019.
48 Zhang, H., Yee, L. D., Lee, B. H., Curtis, M. P., Worton, D. R., Isaacman-VanWertz, G., Offenberg,
49 J. H., Lewandowski, M., Kleindienst, T. E., Beaver, M. R., Holder, A. L., Lonneman, W. A.,
50 Docherty, K. S., Jaoui, M., Pye, H. O. T., Hu, W., Day, D. A., Campuzano-Jost, P., Jimenez, J. L.,
51 Guo, H., Weber, R. J., de Gouw, J., Koss, A. R., Edgerton, E. S., Brune, W., Mohr, C., Lopez-
52 Hilfiker, F. D., Lutz, A., Kreisberg, N. M., Spielman, S. R., Hering, S. V., Wilson, K. R., Thornton, J.
53 A., and Goldstein, A. H.: Monoterpenes are the largest source of summertime organic aerosol in the
54 southeastern United States, Proc. Natl. Acad. Sci. U S A, 115, 2038,
55 https://doi.org/10.1073/pnas.1717513115, 2018.

1 Zhang, J. K., Sun, Y., Liu, Z. R., Ji, D. S., Hu, B., Liu, Q., and Wang, Y. S.: Characterization of
2 submicron aerosols during a month of serious pollution in Beijing, 2013, Atmos. Chem. Phys., 14,
3 2887-2903, https://doi.org/10.5194/acp-14-2887-2014, 2014.
4 Zhang, Q., Worsnop, D. R., Canagaratna, M. R., and Jimenez, J. L.: Hydrocarbon-like and
5 oxygenated organic aerosols in Pittsburgh: insights into sources and processes of organic aerosols,
6 Atmos.Chem.Phys.,5, 3289-3311, https://doi.org/10.5194/acp-5-3289-2005, 2005.
7 Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.
8 L.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a
9 review, Anal.Bioanal.Chem.,401, 3045-3067, https://doi.org/10.1007/s00216-011-5355-y, 2011.
10 Zhang, Y. X., Schauer, J. J., Zhang, Y. H., Zeng, L. M., Wei, Y. J., Liu, Y., and Shao, M.:
11 Characteristics of particulate carbon emissions from real-world Chinese coal combustion, Environ.
12 Sci. Technol., 42, 5068-5073, https://doi.org/10.1021/es7022576, 2008.
13 Zhao, J., Qiu, Y. M., Zhou, W., Xu, W. Q., Wang, J. F., Zhang, Y. J., Li, L. J., Xie, C. H., Wang, Q.
14 Q., Du, W., Worsnop, D. R., Canagaratna, M. R., Zhou, L. B., Ge, X. L., Fu, P. Q., Li, J., Wang, Z.
15 F., Donahue, N. M., and Sun, Y. L.: Organic Aerosol Processing During Winter Severe Haze
16 Episodes in Beijing, J. Geophys. Res. Atmos., 124, 10248-10263,
17 https://doi.org/10.1029/2019JD030832, 2019.
18 Zhou, J., Zotter, P., Bruns, E. A., Stefenelli, G., Bhattu, D., Brown, S., Bertrand, A., Marchand, N.,
19 Lamkaddam, H., Slowik, J. G., Prévôt, A. S. H., Baltensperger, U., Nussbaumer, T., El-Haddad, I.,
20 and Dommen, J.: Particle-bound reactive oxygen species (PB-ROS) emissions and formation
21 pathways in residential wood smoke under different combustion and aging conditions, Atmos. Chem.
22 Phys., 18, 6985-7000, https://doi.org/10.5194/acp-18-6985-2018, 2018.

23