Atmos. Chem. Phys., 20, 7693–7716, 2020
https://doi.org/10.5194/acp-20-7693-2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Chemistry and Physics
Open Access
EGU

Deconvolution of FIGAERO–CIMS thermal desorption profiles using positive matrix factorisation to identify chemical and physical processes during particle evaporation

Angela Buchholz[1], Arttu Ylisirniö[1], Wei Huang[2], Claudia Mohr[3], Manjula Canagaratna[4], Douglas R. Worsnop[4], Siegfried Schobesberger[1], and Annele Virtanen[1]

[1]Department of Applied Physics, University of Eastern Finland, Kuopio, Finland
[2]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Department of Environmental Science and Analytical Chemistry, Stockholm University, Stockholm, Sweden
[4]Aerodyne Research Inc., Billerica, MA 08121, USA

Correspondence: Angela Buchholz (angela.buchholz@uef.fi)

Received: 10 October 2019 – Discussion started: 20 November 2019
Revised: 15 April 2020 – Accepted: 1 May 2020 – Published: 2 July 2020

Abstract. The measurements of aerosol particles with a filter inlet for gases and aerosols (FIGAERO) together with a chemical ionisation mass spectrometer (CIMS) yield the overall chemical composition of the particle phase. In addition, the thermal desorption profiles obtained for each detected ion composition contain information about the volatility of the detected compounds, which is an important property for understanding many physical properties like gas–particle partitioning. We coupled this thermal desorption method with isothermal evaporation prior to the sample collection to investigate the chemical composition changes during isothermal particle evaporation and particulate-water-driven chemical reactions in α-pinene secondary organic aerosol (SOA) of three different oxidative states. The thermal desorption profiles of all detected elemental compositions were then analysed with positive matrix factorisation (PMF) to identify the drivers of the chemical composition changes observed during isothermal evaporation. The keys to this analysis were to use the error matrix as a tool to weight the parts of the data carrying most information (i.e. the peak area of each thermogram) and to run PMF on a combined data set of multiple thermograms from different experiments to enable a direct comparison of the individual factors between separate measurements.

The PMF was able to identify instrument background factors and separate them from the part of the data containing particle desorption information. Additionally, PMF allowed us to separate the direct desorption of compounds detected at a specific elemental composition from other signals with the same composition that stem from the thermal decomposition of thermally instable compounds with lower volatility. For each SOA type, 7–9 factors were needed to explain the observed thermogram behaviour. The contribution of the factors depended on the prior isothermal evaporation. Decreased contributions from the factors with the lowest desorption temperatures were observed with increasing isothermal evaporation time. Thus, the factors identified by PMF could be interpreted as volatility classes. The composition changes in the particles due to isothermal evaporation could be attributed to the removal of volatile factors with very little change in the desorption profiles of the individual factors (i.e. in the respective temperatures of peak desorption, T_{max}). When aqueous-phase reactions took place, PMF was able to identify a new factor that directly identified the ions affected by the chemical processes.

We conducted a PMF analysis of the FIGAERO–CIMS thermal desorption data for the first time using laboratory-generated SOA particles. But this method can be applied to, for example, ambient FIGAERO–CIMS measurements as well. There, the PMF analysis of the thermal desorption data identifies organic aerosol (OA) sources (such as biomass burning or oxidation of different precursors) and types, e.g. hydrocarbon-like (HOA) or oxygenated organic aerosol (OOA). This information could also be obtained with

the traditional approach, namely the PMF analysis of the mass spectra data integrated for each thermogram. But only our method can also obtain the volatility information for each OA source and type. Additionally, we can identify the contribution of thermal decomposition to the overall signal.

1 Introduction

To understand the impact of secondary organic aerosol (SOA) on the Earth's climate and human health, we need to know more about the chemical and physical properties of these particles and how they evolve over time in the atmosphere. The physical properties of SOA particles are controlled by the physical properties of their constituents and the interaction of the compounds in these complex mixtures. The volatility of SOA constituents is one of the defining characteristics of SOA particles, as it plays a key role in understanding (and predicting) the partitioning behaviour of a compound between the gas and particle phase (Pankow, 1994a, b; Pankow et al., 2001). Generally, the partitioning of a compound into the particle phase is controlled by the saturation vapour pressure (volatility) of the compound involved, its concentrations, and the available condensation sink. In addition, particle-phase processes also play an important role, especially when particle-phase compounds are partitioning back into the gas phase. In highly viscous or solid particles, mass transfer limitations exist that reduce the apparent particle volatility (Buchholz et al., 2019; Wilson et al., 2015; Yli-Juuti et al., 2017). The partitioning process is further complicated by particle-phase chemical reactions. Accretion reactions can convert more volatile compounds into larger and heavier compounds, thereby changing the overall properties of the SOA particles again (Herrmann, 2003; Kroll and Seinfeld, 2008). Particulate water plays a special role in these particle-phase processes. On the one hand, it acts as a plasticiser that reduces the particle viscosity (Renbaum-Wolff et al., 2013; Virtanen et al., 2010) and thus the mass transport limitation in the particles. These transport limitations are responsible for the reduced evaporation under dry conditions (Liu et al., 2016; Wilson et al., 2015; Yli-Juuti et al., 2017). On the other hand, the presence of an aqueous phase enables a wide range of chemical reactions with the potential to form low-volatility compounds via oligomerisation reactions (e.g. Surratt et al., 2007; Tolocka et al., 2004). Hydrolysis of labile bonds (e.g. peroxides or esters) is also possible, which would lead to more volatile products.

There are many challenges involved in trying to fully characterise SOA particles and their volatility. Already the sheer number of precursor compounds and their reaction products, which may contribute to the particle phase by forming new particles or condensing on existing ones, makes it almost impossible to fully characterise the chemical composition of SOA particles (Glasius and Goldstein, 2016; Goldstein and Galbally, 2007). However, the development of the filter inlet for gases and aerosols (FIGAERO, Lopez-Hilfiker et al., 2014) for the chemical ionisation mass spectrometer (CIMS) was a big step forward in the chemical characterisation of SOA particles. This is because the FIGAERO–CIMS provides more detailed information about the molecular composition and, at the same time, records the thermal desorption behaviour (thermogram) of each detected ion. Hence, in addition to composition information, FIGAERO–CIMS measurements enable the determination of the volatility of SOA constituents. In an ideal case, the peak desorption temperature (T_{max} – temperature at peak of ion thermogram) of a single ion thermogram is correlated to the ion volatility expressed by its effective saturation vapour pressure, C^*_{sat} (Lopez-Hilfiker et al., 2014; Schobesberger et al., 2018). This relationship can be calibrated for a specific FIGAERO–CIMS set-up and temperature ramp by measuring compounds with known volatilities, e.g. carboxylic acids (Lopez-Hilfiker et al., 2014) or polyethylene glycol (Bannan et al., 2019). Unfortunately, in most cases the data interpretation is more complicated as some compounds will not desorb from the FIGAERO filter at a temperature corresponding to their volatility. They will instead decompose at a lower temperature, and the decomposition products will be detected in the mass spectrometer (D'Ambro et al., 2019; Lopez-Hilfiker et al., 2015; Stark et al., 2017; Wang and Hildebrandt Ruiz, 2018). The decomposition products may have the same sum formula as other constituents of the particles. Thus, only the shape of the ion thermogram may give a hint as to whether an ion stems from the desorption (typically a sharp peak) or decomposition of one or several different larger compounds (typically a broad peak or broad tailing on a peak, Schobesberger et al., 2018). A further complication for the interpretation of the T_{max} values arises from the presence of multiple isomers with different volatilities. Depending on how close the T_{max} values of the isomers are and on the contribution of each isomer to the signal at this ion mass, the resulting ion thermogram may be multimodal, broadened, or with considerable tailing/fronting.

To overcome the issues related to thermal decomposition and further the interpretation of ion thermograms, we utilised positive matrix factorisation (PMF) in the interpretation of FIGAERO data. Traditionally, PMF has been used to analyse complex mass spectra data sets, and it is mostly used to identify the contribution of different sources to the total organic aerosol mass (Jimenez et al., 2009; Lanz et al., 2007; Ulbrich et al., 2009). But, for PMF, it does not matter if the "source" of a mass spectra signal is a real, physical source (e.g. biomass burning, or traffic emissions) or if the source is particles collected on a filter being desorbed. PMF identifies the characteristic changes in the contribution from a source to the total signal; this means in the case of the FIGAERO–CIMS data one or more compounds desorbing at a specific temperature range. In this study, we apply PMF to FIGAERO–CIMS data analysis for the first time to

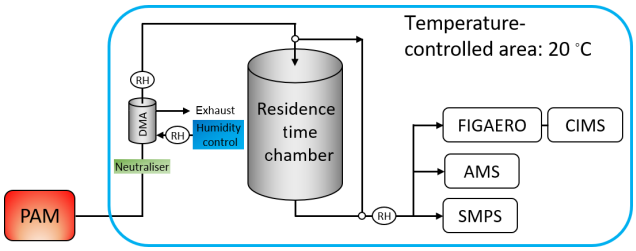

Figure 1. Schematic diagram of experimental set-up.

distinguish the direct desorption (controlled by C^*_{sat}) from the thermal decomposition of thermally labile compounds with lower volatility (i.e. controlled by the strength of the weakest bond in the molecule). Furthermore, we combine the PMF analysis of FIGAERO–CIMS data with the information gained from isothermal evaporation experiments in which the particle composition evolves during the isothermal evaporation of the particles to understand the processes controlling particle volatility.

2 Methodology

2.1 Data set

The acquisition of the data set investigated in this study was described in detail in Buchholz et al. (2019) and in the Supplement. The schematic overview of the set-up is shown in Fig. 1. Briefly, three types of SOA were formed via combined ozonolysis and photo-oxidation of α-pinene in an oxidative flow reactor (OFR). They are characterised as low, medium, and high O : C based on their elemental composition (oxygen-to-carbon (O : C) ratio of 0.53, 0.69, and 0.96, respectively, as derived from the aerosol mass spectrometer data). A nano differential mobility analyser (NanoDMA, TSI Inc.) was used to select a quasi-monodisperse particle distribution (with an electrical mobility diameter of 80 nm) and, at the same time, dilute the surrounding gas phase by orders of magnitude, which initiates isothermal evaporation at the NanoDMA outlet. The monodisperse particles were then filled into a stainless-steel residence time chamber (RTC) to study their isothermal evaporation behaviour by measuring the particle size in 1 h intervals for up to 10 h. Two sets of evaporation experiments were conducted for each SOA type, namely dry (RH < 2 %) and wet (RH 80 %). To achieve the different RH conditions and control the RH of the selected sample and in the RTC, only the RH of the sheath flow in the NanoDMA was adjusted. The conditions of α-pinene SOA formation in the OFR were not changed. The instruments, tubing, RTC, and OFR were flushed with particle-free purified air or nitrogen between experiments.

The chemical composition of the particles was investigated with a filter inlet for gases and aerosols (FIGAERO, Aerodyne Research Inc.; Lopez-Hilfiker et al., 2014) sampling unit in combination with a chemical ionisation mass spectrometer (CIMS, Aerodyne Research Inc.; Lee et al., 2014) using iodide as reagent ion. Samples were taken directly after the size selection ("fresh" particles: $t_{evap} = 0.25$ h) and after 3–4 h of isothermal evaporation in the RTC ("RTC" particles: $t_{evap} = 4$ h). Note that the evaporation time of 0.25 h for the "fresh" sample does not stem from residence in the RTC but rather from the collection time on the filter (see Sect. S1.1 in the Supplement for details). Due to this minimum evaporation time, the FIGAERO–CIMS measurements will underestimate the contribution of volatile compounds in the particles as they leave the OFR.

The combined analysis of the evaporation behaviour and FIGAERO–CIMS thermogram with the composition information in Buchholz et al. (2019) revealed increasing average desorption temperatures with increasing O : C ratio of the particles, while the overall particle volatility (measured by isothermal evaporation) decreased. The residual particles after isothermal evaporation in the RTC exhibited an increase in desorption temperature in all cases, which indicates that the more volatile species had left the particles. Under wet conditions, evaporation was enhanced due to the lowering of particle viscosity and, thus, kinetic transport limitations as described previously (D'Ambro et al., 2018; Wilson et al., 2015; Yli-Juuti et al., 2017). But in the high-O : C case, strong indications for aqueous-phase chemistry were found in the data, namely the shift of some ion thermograms to much higher desorption temperatures and a relative increase in low molecular weight (M_w) compounds. Thus, this data set is perfect for testing the performance of PMF with FIGAERO–CIMS data and determining whether PMF can capture the evaporation behaviour and separate it from aqueous-phase processes in the high-O : C case.

2.2 FIGAERO–CIMS measurements

It is necessary to understand the operation and data structure of FIGAERO–CIMS to comprehend the challenges of analysing this data with PMF. In the FIGAERO inlet, particles are collected on a polytetrafluoroethylene (PTFE) filter. A gradually heated nitrogen gas flow evaporates increasingly less-volatile compounds and transports them into the CIMS for detection. Hereafter, the resulting signal versus desorption temperature curves will be called ion thermogram for individual ions and total thermogram for the sum of all detected ions apart from the reagent ions. Each desorption cycle ("thermogram scan") consists of the following three parts: the particle collection; the linear increase of the desorption temperature (here, $\sim 25\,^\circ$C $\rightarrow \sim 190\,^\circ$C in 15 min); and a "soak" period at the highest temperature (here, $> 190\,^\circ$C for 15 min). The soak period ensures that low-volatility compounds have been removed from the FIGAERO filter before the next sample is collected. Note that only the part of the thermogram with a near-linear increase in the desorption temperature can be used to derive volatility information. The

relationship between a compound's desorption temperature, specifically T_{max}, and volatility (e.g. expressed as saturation vapour pressure) can be calibrated for a specific FIGAERO–CIMS set-up and temperature ramp by measuring, for example, polyethylene glycol aerosol with a range of molecular weights and volatilities (similar to the method described by Bannan et al., 2019).

The raw FIGAERO–CIMS data were processed by using tofTools, a MATLAB-based software package developed for analysing time-of-flight CIMS data (Junninen et al., 2010). The data were averaged to a 20 s time grid, and a baseline correction was applied before the high-resolution mass spectra data were fitted. The filter blank measurements were processed in the same fashion as the collected samples. For the PMF analysis, we did not subtract the filter blank measurements but instead added the corresponding filter blank thermograms to the data set to help with the identification of the background factors i.e. factors dominated by compounds from the instrument and/or filter background (more details on factor identification in Sect. 3.1).

Due to suboptimal settings in the instrument's ion guidance unit, an atypically high number of declustered ions (not containing the reagent ion iodide) were observed. This was discussed in detail in Buchholz et al. (2019). For this study, we will not make any assumptions about the declustering process and treat the iodide clusters and declustered ions as separate variables. However, this impacts neither the application of PMF to the data set nor the validity of this method for other data sets as the variables (ions) are all treated independently in the model, and the variables with the same behaviour will be grouped into the same factor.

2.3 Positive matrix factorisation (PMF)

2.3.1 Working principles of PMF

Since its introduction by Paatero and Tapper (1994), PMF has been established as a useful tool for analysing long time series of mass spectra data mostly from ambient observations.

In the PMF model, it is assumed that the measured data can be expressed by the combination of an (unknown) number p of constant source profiles with varying concentrations over time (Ulbrich et al., 2009). This can be mathematically expressed as follows:

$$\mathbf{X} = \mathbf{GF} + \mathbf{E}, \tag{1}$$

where $\mathbf{X}$ is a $m \times n$ matrix containing the measured mass spectra with m rows of mass spectra ("observations"), which are each averaged over 20 s of measurement time in the CIMS, and n columns representing the time series of one specific ion. $\mathbf{G}$ is a $m \times p$ matrix containing the factor time series as columns. The rows of the $p \times n$ matrix $\mathbf{F}$ contain the factor mass spectra, and then the $m \times n$ matrix $\mathbf{E}$ contains the residuals between the measured data and the fitted values. No a priori information about the values of $\mathbf{G}$ and $\mathbf{F}$ or the number of factors (p) is required, but the user has to decide which solution (i.e. how many factors) characterises the data best. To account for uncertainties in the measurement data, the PMF model weights the data points with their measurement error ($\mathbf{S}_{ij}$). Values for $\mathbf{G}$ and $\mathbf{F}$ are constrained to be positive and are iteratively found by minimising the quantity, Q, with a least square algorithm (Paatero and Tapper, 1994) as follows:

$$Q = \sum_{j=1}^{m} \sum_{i=1}^{n} \left(\frac{\mathbf{E}_{ij}}{\mathbf{S}_{ij}} \right)^2, \tag{2}$$

where $\mathbf{S}_{ij}$ is the error (uncertainty) of each measurement data point. In an ideal case, the Q value of the model should approach the expected Q value (Q_{exp}) that is equal to the degree of freedom of the model solution. For mass spectra data, this is approximately equal to the size of the original data matrix, $\mathbf{X}$, as follows:

$$Q_{exp} \approx n \cdot m. \tag{3}$$

Different algorithms have been developed to solve the PMF model (e.g. Hoyer, 2004; Lu and Wu, 2004; Paatero, 1999). In this study, we used the PMF2 algorithm with robust least square optimisation, which is included in the PMF Evaluation Tool (PET, Ulbrich et al., 2009) for Igor Pro 7 (WaveMetrics, Inc., Portland, Oregon). We calculated solutions with 1 to 12 factors. For each solution, 5 rotations ($f_{peak} -1.0$ to $+1.0$) were calculated, and 6 different seed values were tested for each original solution ($f_{peak} = 0$).

As an additional measure for the goodness of fit, we calculated the fraction of explained absolute variance (Ratio_{exp}) as follows:

$$\mathrm{absVar}_{total} = \sum_{ij} \left| \mathbf{X}_{ij} - \overline{X_i} \right| \tag{4}$$

$$\mathrm{absVar}_{explained} = \sum_{ij} \left| \mathbf{R}_{ij} - \overline{X_i} \right| \tag{5}$$

$$\mathrm{Ratio}_{exp} = \frac{\mathrm{absVar}_{explained}}{\mathrm{absVar}_{total}}, \tag{6}$$

where $\mathbf{R}_{ij}$ is the value in the reconstructed data matrix ($\mathbf{R} = \mathbf{GF}$) for each ion i and observation j, $\overline{X_i}$ is the average measured value of the ion i, and absVar_{total} and $\mathrm{absVar}_{explained}$ are the total and explained absolute variance. Note that we use the absolute distance between the average values and the measured or reconstructed data instead of the square of this distance.

PMF has been widely used for analysing the time series of mass spectra data in the atmospheric science community. However, the model does not utilise the information of the time axis in the optimisation process. Rather, it is a method that can be used to analyse a set of mass spectra, which were obtained at different time points during the desorption cycle of FIGAERO, and for different particle sampling conditions.

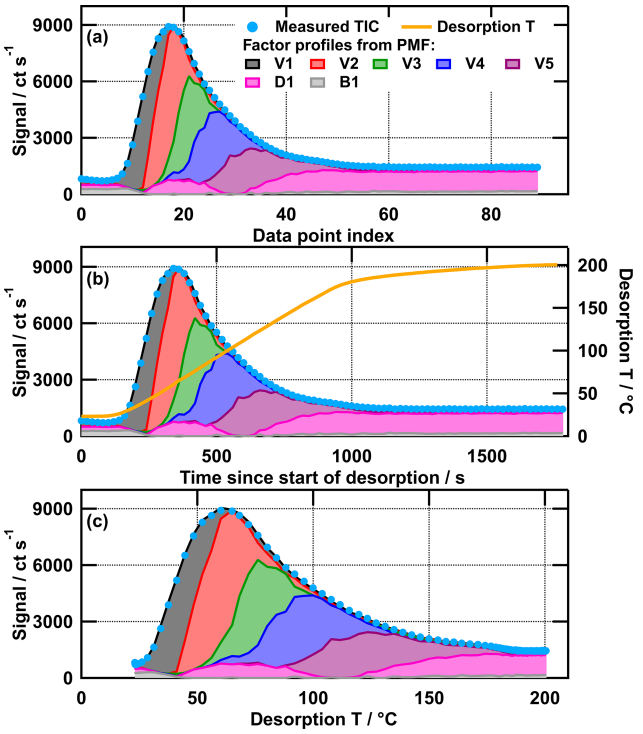

Figure 2. Measured total ion thermogram with the colour-coded contributions of PMF model output factors for the medium-O : C wet ($t_{evap} = 4$ h) case plotted against data point index **(a)**; time since start of desorption **(b)**; and desorption temperature **(c)**. Note that the desorption temperature ramp **(b)** is not increasing linearly after ∼ 1000 s. This "soak" period ensures that all organic material is removed from the filter before the next collection.

This means that PMF will create the same model output if the x values in the data set are a real time series (Fig. 2b), a temperature ramp (Fig. 2c), or simply an index with numbers (Fig. 2a). Thus, data from separate thermogram scans with FIGAERO–CIMS can be combined to larger data sets and analysed together with PMF. Analysing multiple thermogram scans together has the advantage that more data points are utilised to identify the factors (90 mass spectra for each thermogram in this case) and that factors can be directly compared between scans. The real time series/temperature ramp is only of interest when evaluating the model output for the interpretation of the identified factors and comparing their desorption temperature profiles between thermogram scans. In the graphic presentation of these combined "time series" (e.g. Fig. 4), a data index was used as the x values, which is the desorption temperature of each thermogram plus an offset (200 per thermogram). This choice of x values preserves the shape of the thermogram in the desorption temperature space. The individual thermograms are marked with roman numerals and the sampling conditions are given in the figure captions. For an easier comparison of the shape of the desorption behaviour of the factors, they are plotted individually for each SOA type (e.g. Fig. 5).

When performing PMF with the combined data set with all available thermogram scans, the large number of factors (13 or more) necessary to explain the observed variability complicates the analysis and interpretation (see the case study in Sect. S1.4). Thus, the thermogram scans were grouped by SOA type (i.e. $t_{evap} = 0.25$ h and 4 h particles, dry and wet conditions of one SOA type: four thermogram scans per group). This pre-grouping reduced the number of factors in each group enhancing their interpretability while still enabling a direct investigation of the changes due to the evaporation and/or humidification for one SOA type. But, generally, splitting the data by SOA type or even knowing about different SOA types and/or sources in the data is not a requirement for analysing a thermogram data set with PMF.

2.3.2 Error schemes for PMF

To perform the PMF analysis, a data error $\mathbf{S}_{ij}$ must be defined. As seen in Eq. (2), the $\mathbf{S}_{ij}$ values have a strong influence on the outcome of the PMF model. The measurement error can be understood as a weighting mechanism that gives more weight to data points with less uncertainty (Paatero and Hopke, 2003). Ideally, $\mathbf{S}_{ij}$ is the true measurement error of the data set. For gas-phase CIMS data, Yan et al. (2016) suggested calculating the measurement error by assuming a Poisson-type distribution of the counting error as follows:

$$\mathbf{S}_{ij} = a \cdot \sqrt{\frac{\mathbf{X}_{ij}}{t_s}} + \sigma_{\mathrm{noise},i}, \tag{7}$$

where $\mathbf{X}_{ij}$ is the signal intensity of the ion i, t_s is the sampling (averaging) interval in s, and $\sigma_{\mathrm{noise},i}$ is the electronic noise for ion i. We applied a procedure equivalent to the one introduced by Yan et al. (2016) to derive the parameter a from analysing the distribution of signal noise. The detailed calculation for this type of error is given in the Supplement. The resulting error values (Poisson-like – "PLerror") will trace the shape of the thermogram signal with higher absolute values for those parts of the thermogram that have higher intensity (i.e. the "peak") and give less weight to this region (Fig. S1 in the Supplement). This is the correct approach for the analysis of long time series data in which rapid changes are most likely caused by instrument noise or data outliers.

For FIGAERO–CIMS thermograms, the main information lies in the rapidly increasing and decreasing part of the data (i.e. the "peak"; data points 10–50 in Fig. 2a) when the compounds are desorbing from the FIGAERO filter and not in the slowly changing (or constant) part at high desorption temperatures (i.e. the "tail"; data points 50–90 in Fig. 2a). During this analysis it was found that the thermal desorption peaks could not be modelled well with error values calculated using Eq. (7) (see Sect. 2.3.3 and Appendix A). Thus, a new error scheme that allowed for an increased weighting of the thermal desorption peaks was also tested. In this scheme, a constant error value corresponding to the noise in the data at

the end of the thermogram scan is used for each thermogram scan (constant noise – "CNerror") as follows:

$$\mathbf{S}_{ij} = \sigma_{\text{noise},i}, \tag{8}$$

where $\sigma_{\text{noise},i}$ of each ion i is calculated in the same way as for the PLerror (see the Supplement for details). Note that by omitting the first term in Eq. (7), Eq. (8) does not correspond to the true measurement error of the FIGAERO–CIMS data. Rather, it is the simplest way of weighting the PMF runs to put more emphasis on each thermogram peak and less on the fronts and tails (Fig. S1 shows an example of the values for the two error schemes for one exemplary ion). The signal-to-noise values are up to 3 orders of magnitude higher in the peak region for the CNerror case, which clearly gives them a stronger weight in the optimisation. As a direct consequence of the modified error value, the value for Q/Q_{exp} is not expected to approach 1 but will instead reach a larger (used error values smaller than real measurement error) or smaller (used error values larger than real measurement error) value. Thus, most solutions from PMF with PLerror will have (much) lower Q and Q/Q_{exp} values than any solution from PMF with CNerror. This also means that comparing the absolute Q or Q/Q_{exp} values between results from the different error schemes is not meaningful as a higher absolute error value will result in a lower Q value.

2.3.3 Selection of error scheme and number of factors ("best" solution)

Before the "best" solution from PMF can be identified by investigating the factor profiles and spectra, the impact of the two different error schemes on the PMF output needs to be determined by running PMF for all combined data sets with both error schemes and comparing the output. Since the comparison of the Q/Q_{exp} values between the error schemes is not meaningful, as pointed out previously, the fraction of explained variance ($\text{Ratio}_{\text{exp}}$) and the reconstruction of the characteristic shape of the thermograms (i.e. time series of residuals) were the decisive criteria. In addition to the single Q/Q_{exp} value summed over all ions and observations (i.e. mass spectra) in each data set, we calculated the time series of the Q contributions (Q_j) summed over all ions for each observation (mass spectrum), j, to identify which periods in the data set were not captured well by the investigated PMF solution. The calculation is as follows:

$$Q_j = \sum_{i=1}^{n} \left(\frac{\mathbf{E}_{ij}}{\mathbf{S}_{ij}} \right)^2. \tag{9}$$

Similarly, we calculated Q_i as the value sum over all observations (mass spectra), j, as follows, to investigate which ion has the strongest contribution to the overall Q value:

$$Q_i = \sum_{j=1}^{m} \left(\frac{\mathbf{E}_{ij}}{\mathbf{S}_{ij}} \right)^2. \tag{10}$$

For a given number of factors, the CNerror scheme results in higher $\text{Ratio}_{\text{exp}}$ values than the PLerror (Fig. 3), i.e. a larger fraction of the observed variance is captured by the model. With the PLerror the maximum $\text{Ratio}_{\text{exp}}$ is 0.9 – even with up to 12 factors – while the values for $\text{Ratio}_{\text{exp}}$ with the CNerror are already > 0.95 with 7 factors.

To highlight the differences in the behaviour of the two error schemes, we display the time series of the residual and Q_j values in Fig. 4 for the high-O : C case for three solutions (6, 7, and 10 factors). With the PLerror, the residuals are much larger than in the CNerror case (panels b and d). But, due to the larger values of $\mathbf{S}_{ij}$ in the PLerror case, the Q/Q_{exp} values (panels c and e) are much smaller. Thus, the optimisation algorithm sees no need to further improve the model in the PLerror case. In contrast, the smaller unscaled residual in the 6-factor solution with the CNerror leads to much higher Q/Q_{exp} values, especially in the peaks of thermograms III and IV. Here, the addition of 1 factor (from 6 to 7) improves both the residual and the Q_j/Q_{exp} values, and the new factor captures a characteristic behaviour that we discuss in Sect. 3.3.

This analysis, together with the more detailed case study in Appendix A, leads us to the conclusion that, for this study and data set, the CNerror reconstructed the measured data best and yielded the most interpretable results. Thus, hereafter we only present the results from PMF runs with the CNerror scheme.

The advantage of PMF, that no a priori information about $\mathbf{F}$, $\mathbf{G}$, and p is needed for the analysis, is also a disadvantage. There is no absolute criterion for which number of factors (p) is correct or "best", but the chosen value strongly impacts the interpretation of the factors and their profiles. In the ideal case, when the true measurement errors are used, Q/Q_{exp} approaches 1, and a solution with Q/Q_{exp} close enough to 1 may be considered as the "best" or most correct. But, as we explained previously, PMF performed much better for FIGAERO–CIMS data when the "unrealistic" CNerror scheme was used; thus Q/Q_{exp} are not necessarily meaningful. However, the shape of the Q/Q_{exp} versus the number of factors curve can be used to judge the impact of introducing another factor; i.e. a large change in Q/Q_{exp} values suggest that the new factor explains a large fraction of the variability in the data (Ulbrich et al., 2009). We investigated this for the PMF runs for each SOA type (Figs. 3 and S2). The largest changes in Q/Q_{exp} are already achieved by increasing the factors from 2 to 3. Further factor addition leads to a steady decrease of Q/Q_{exp}. In this case, the $\text{Ratio}_{\text{exp}}$ values are more helpful. Strong increases of $\text{Ratio}_{\text{exp}}$ are observed when increasing the number of factors to 6 (medium- and high-O : C cases) or 8 (low-O : C case).

As shown by Yan et al. (2016) for gas-phase CIMS data, a solution with a low overall Q/Q_{exp} value may still have large variations in the scaled residual with time or with different ions. We carefully investigated the time series (Q_j/Q_{exp}) of individual ions (e.g. $C_5H_5O_6^-$ in Fig. A1b and c) in particu-

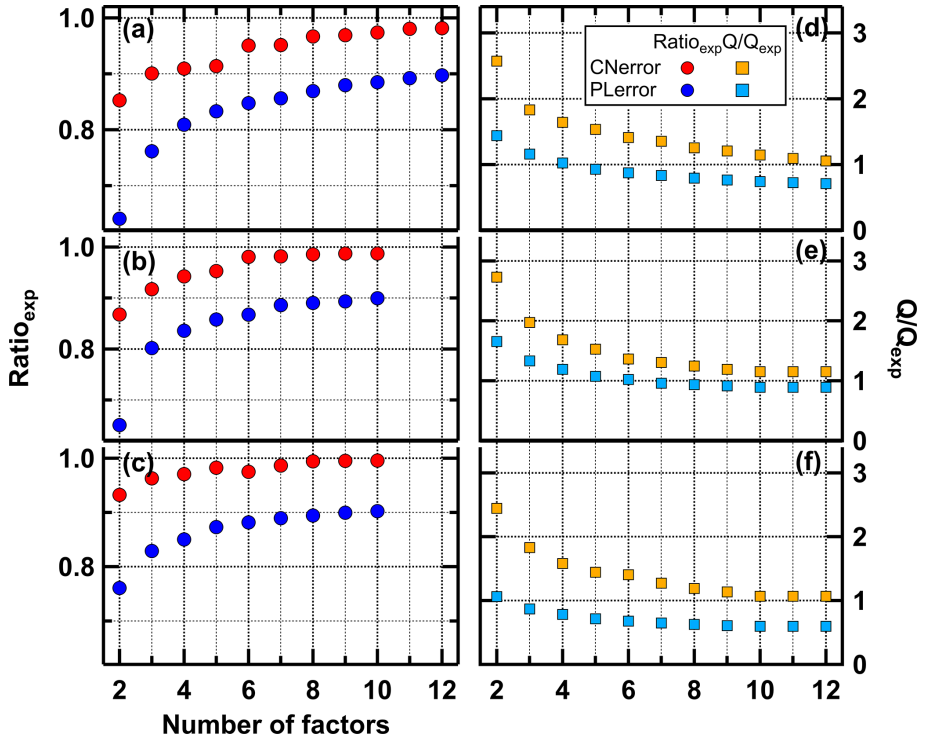

Figure 3. Fraction of explained variance (Ratio$_{exp}$, **a–c**) and Q/Q_{exp} values **(d–f)** for the low- **(a)**, medium- **(b)**, and high-O : C **(c)** data set for PLerror (blue) and CNerror (red).

lar and present the details of this case study in Appendix A. There were a few specific ions for each SOA type that were not captured well in the data set until a certain number of factors was chosen (e.g. 7 in the high-O : C case) – even if the overall fraction of explained variance for the solutions was already larger than 95 % and changed very little with further factor additions. We decided to choose the PMF solution with the smallest number of factors that still described the characteristic behaviour of most ion thermograms. These were the solutions with 9, 7, and 7 factors for the low-, medium-, and high-O : C cases, respectively.

3 Results and discussion

3.1 PMF-factor interpretation

The three evaporation data sets (one for each SOA type) were analysed with PMF by using the CNerror scheme and the results for the "best" solutions chosen are shown in Figs. 5, 6, and 7 (and with "stacked" factor contribution in Figs. S4, S5, and S6). In the following paragraphs, the first letter in the labels of factors indicates whether they are from the low- (L), medium- (M), or high (H)-O : C case, and the second letter identifies the factor type (V, B, D, and C; see below).

Generally, there were the following three main types of thermogram profiles for all factors: volatility class (type V) with a single, distinct peak (LV1 – 5, MV1 – 5, and HV1 – 5); background (type B) with mostly constant contribution over the full T_{desorp} range (LB1, MB1, and HB1); and decomposition (type D) with mostly very broad peaks at $T_{desorp} < 65\,°C$ and an increase at $T_{desorp} > 110\,°C$ (LD1, MD1, and HD1).

Factors of type V do not contribute to the filter blank thermograms (Fig. S3), which indicates that these factors are linked to compounds only present in the sampled aerosol particles. With the exception of the high-O : C wet case (which we discuss in detail in Sect. 3.3), the peak position (T_{max}) of type-V factors changes very little with aerosol age or water content (Table 2). Only the contribution of these factors to the total signal changes with isothermal evaporation or humidification. For each type-V factor, we could identify ions with thermogram shapes similar to the thermogram profile of the individual factors. This means that especially the type-V factors at high desorption temperature are not merely a better mathematical description of the tails of some ion thermograms, but they represent real compounds desorbing from the FIGAERO filter at high desorption temperatures. Thus, we interpret the type-V factors as volatility classes. Compounds with the same thermal desorption behaviour (i.e. volatility) are grouped into one type-V factor that is characterised by its T_{max} value. Note that for the three different SOA types the starting particle composition was significantly different. So, even if the T_{max} values for 2 factors of different SOA types, e.g. LV2, MV2, and HV1 (dry cases), only differ by $\sim 5\,°C$, the compounds contributing to them are not the same; i.e. the

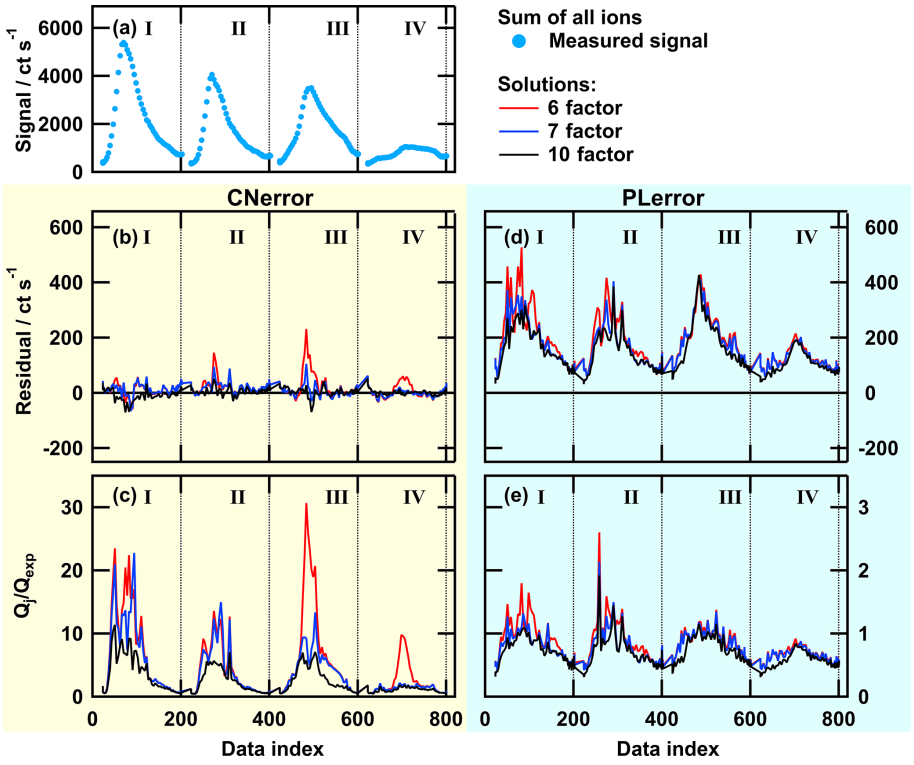

Figure 4. Total ion thermogram **(a)**, residuals **(b, d)**, and Q_j/Q_{exp} values **(c, e)** as time series for solutions with 6, 7, or 10 factors for PMF run with CNerror (**b, c** – yellow background) and PLerror (**d, e** – blue background). The data set contains thermogram scans for the high-O : C SOA particles from the following sampling conditions: dry – $t_{\mathrm{evap}} = 0.25$ h (I); dry – $t_{\mathrm{evap}} = 4$ h(II); wet – $t_{\mathrm{evap}} = 0.25$ h (III); and wet – $t_{\mathrm{evap}} = 4$ h (IV). Note that the y-axis scaling is the same in panels **(b)** and **(d)**, but it is 10 times smaller in **(e)** than in **(c)**.

factor mass spectra for LV2, MV2, and HV1 are significantly different. We elaborate on the reasons for these differences in Sect. S1.3 and S1.4.

Type-B factors show contributions to the signal of sample thermograms and filter blanks (Fig. S3). For LB1, MB1, and HB1, the very shallow thermogram profile and the similar absolute signal strength despite different mass loadings on the FIGAERO filter indicate that these are instrument background factors. For some samples, especially those with high overall filter loading, the factor thermograms of the type-B factors show a decrease to almost 0 for $T_{\mathrm{desorp}} \sim 60$–$\sim 120\,^{\circ}\mathrm{C}$. This may indicate that part of the background signal in this temperature range is assigned to other factors ("factor blending/mixing"). Since the overall background in these few cases is 200–400 ct s^{-1} while the signals of the type-V factors are more than 10 times higher, this factor blending will have a minor effect on the shape and T_{max} values of the type-V factors and thus the interpretation of the PMF solution. For the parts of the thermograms with lower signal strength and samples with low overall mass loading on the filter, the signal of the type-B factors stays constant at a value that is comparable to values observed for the filter blank samples. For all SOA types, the mass spectra of these factors are dominated by single ions typically associated with the FIGAERO–CIMS background (e.g. fluorine-containing compounds, formic acid, and lactic acid). According to the uncentred correlation method (contrast angle or dot product) MB1 and HB1 are reasonably similar. For the low-O : C case, some of the instrument background is apparently assigned to the contamination factors (LC1 and 2; see below), which decreases the degree of similarity between LB1 and the other type-B factors.

Type-D factors are the most difficult to interpret as they contribute to the signal for both the filter blank and sample thermograms, but the contribution can vary with the collected mass loading on the filter for sample thermograms. The factor mass spectra (LD1, MD1, and HD1) mostly show contributions from ions with $M_{\mathrm{w}} < 200$ Da, but the thermogram profiles exhibit a strong increase at $T_{\mathrm{desorp}} > 110\,^{\circ}\mathrm{C}$, especially in filter blank thermograms. This suggests that the detected low M_{w} compounds in these factors are thermal decomposition products of larger, low-volatility compounds that are thermally unstable. But in some cases (e.g. medium-O : C dry – $t_{\mathrm{evap}} = 0.25$ h and 4 h; Fig. 6a and b) a second peak occurs at much lower T_{desorp} ($< 65\,^{\circ}\mathrm{C}$), which is in the range where compounds of the detected composition are expected to desorb. This suggests that the ions grouped into the type-D factors can stem from two "sources", namely

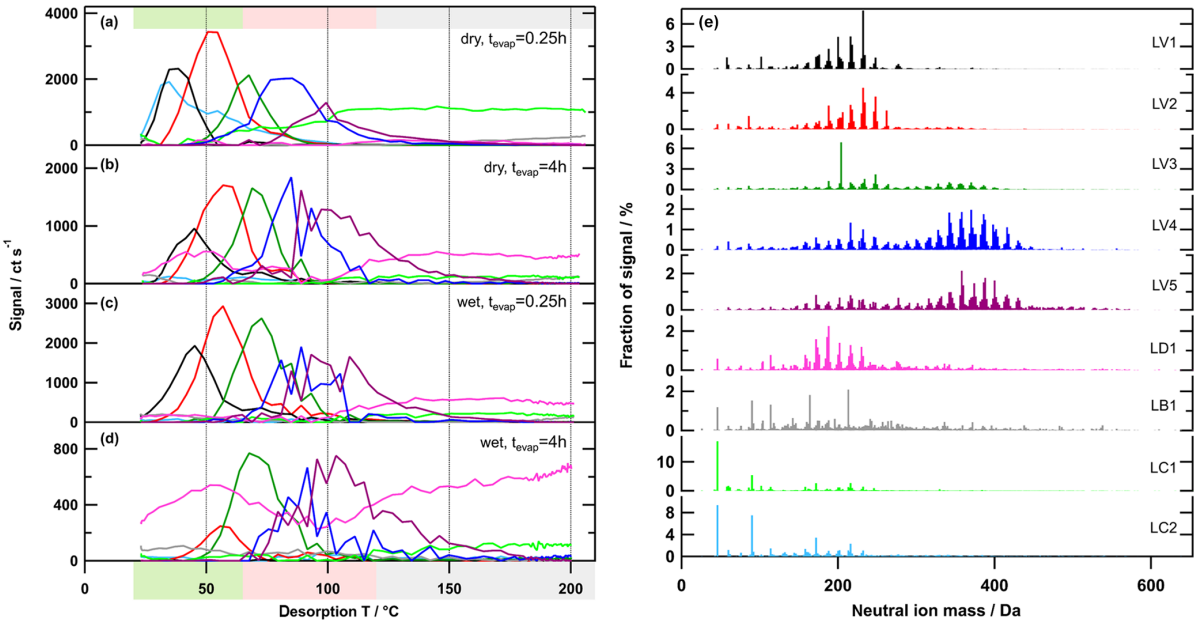

Figure 5. Factor thermograms **(a–d)** and factor mass spectra **(e)** for the 9-factor solution for the low-O : C case. Each factor mass spectrum is normalised. The colour code is the same for both panels. Background colours in the panel on the left indicate volatility classifications according to Donahue et al. (2006) and are derived from T_{max}–C^*_{sat} calibrations (green: semi-volatile organic compounds (SVOCs), red: low-volatility organic compounds (LVOCs), and grey: extremely low volatility organic compounds (ELVOCs)). Note the different scaling for the y axes in **(a)**–**(d)**.

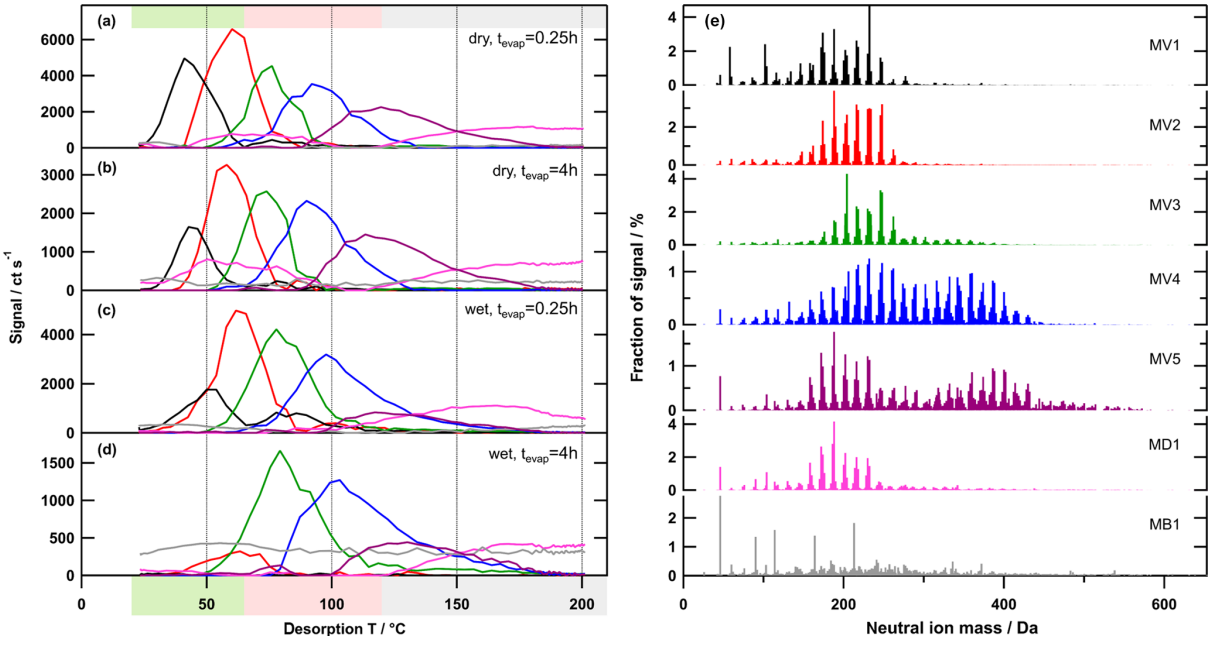

Figure 6. Factor thermograms **(a–d)** and factor mass spectra **(e)** for the 7-factor solution for the medium-O : C case. Each factor mass spectrum is normalised. The colour code is the same for both panels. Background colours in the panel on the left indicate the volatility classification derived from T_{max}–C^*_{sat} calibrations (green – SVOC; red – LVOC; and grey – ELVOC). Note the different scaling for the y axes in **(a)**–**(d)**.

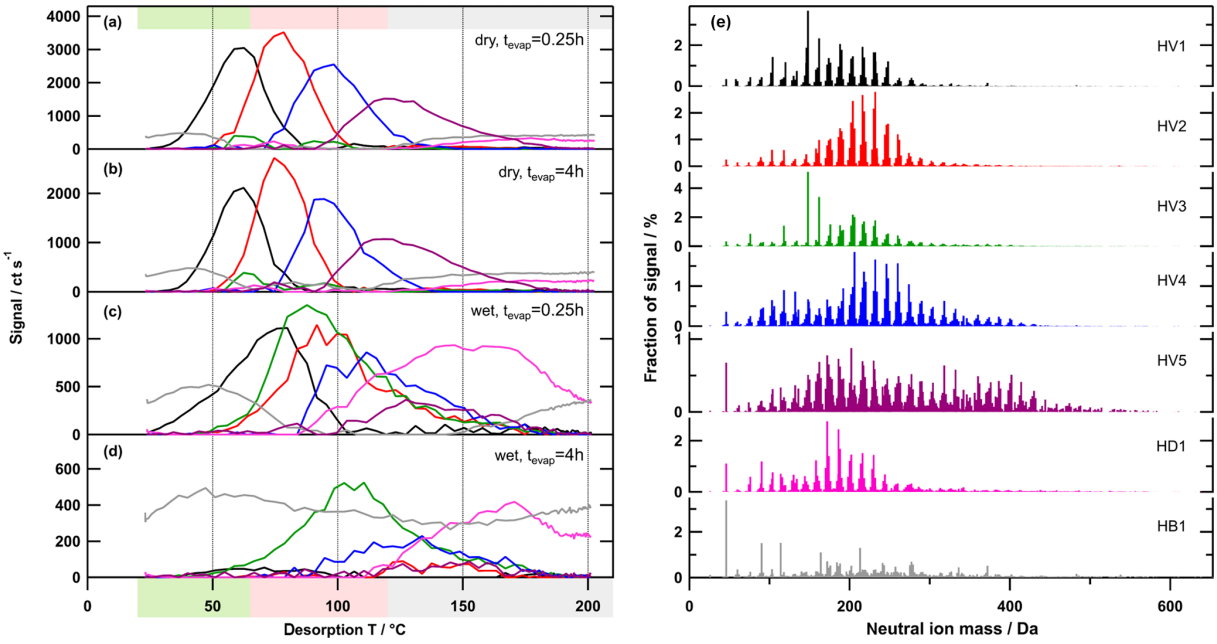

Figure 7. Factor thermograms **(a–d)** and factor mass spectra **(e)** for the 7-factor solution of the high-O : C case. Each factor mass spectrum is normalised. The colour code is the same for both panels. Background colours in the panel on the left indicate the volatility classification derived from T_{max}–C^*_{sat} calibrations (green – SVOC; red – LVOC; and grey – ELVOC). Note the different scaling for the y axes in **(a)**–**(d)**.

direct desorption ($T_{desorp} < \sim 100\,°C$) and thermal decomposition ($T_{desorp} > \sim 100\,°C$), and PMF is not able to separate them because either their composition or their desorption behaviour is too similar. Consequently, type-D factors have to be carefully analysed and interpreted as desorption at low T_{desorp} and decomposition at high T_{desorp}. Also, the instrument background contribution needs to be estimated from the filter blank thermograms. For the low-O : C case, LD1 is dominated by compounds coming from the filter/instrument background as the factor thermogram does not change with the collected sample mass, and there is still a contribution from the factor below $T_{desorp} < 100\,°C$ after 4 h of isothermal evaporation (Fig. S3a). In the medium-O : C case, the direct desorption part ($T_{desorp} < 100\,°C$) of MD1 is removed with isothermal evaporation, which suggests that at least this part of the factor stems from the collected sample and not just the instrument/filter background. The high-O : C case is discussed in Sect. 3.3 below.

For the low-O : C dry ($t_{evap} = 0.25$ h) sample, 2 additional factors (type C) were found. The factor mass spectra of LC1 and 2 are dominated by extremely high signals for formic and lactic acid, which are typically strong indications of a contamination on the FIGAERO filter due to handling. Retrospectively, we could not determine what happened to this specific sample collection to cause this obvious contamination, but the FIGAERO filter was replaced between this and the next sample collection, and several heating cycles were performed to ensure that no other sample was affected. However, since PMF has identified the ions affected by this contamination and grouped them into LC1 and 2, these 2 factors can be omitted from further analysis and remove the bias caused by the contamination.

Note that almost the same factors are produced by PMF regardless of whether the filter blank measurements are added to the data sets or not. This shows that PMF can be a very helpful tool for data interpretation when no reliable instrument background measurements are available or if the background varies strongly between samples. Then the identification of the type-B, type-D, and type-C factors has to rely on the factor thermograms and factor mass spectra.

3.2 Composition changes due to evaporation

One set of type-V factors (i.e. volatility classes) was identified and separated from the instrument background contributions for each data set consisting of one SOA type sampled after different time intervals of isothermal evaporation under dry and wet conditions. The contribution of a single factor to the total signal is calculated as the ratio of the integral of the thermogram of this factor to the total signal. The relative contribution of factors V1–V5 for each sampling condition is shown in Fig. 8 and is plotted against the volume fraction remaining (VFR) that was measured in separate isothermal evaporation measurements (VFR values from Buchholz et al., 2019). The corresponding figure with absolute signal contributions is shown in the Supplement (Fig. S7). Note that the residual particles after isothermal evaporation or humidification were collected on the FIGAERO filter. This means

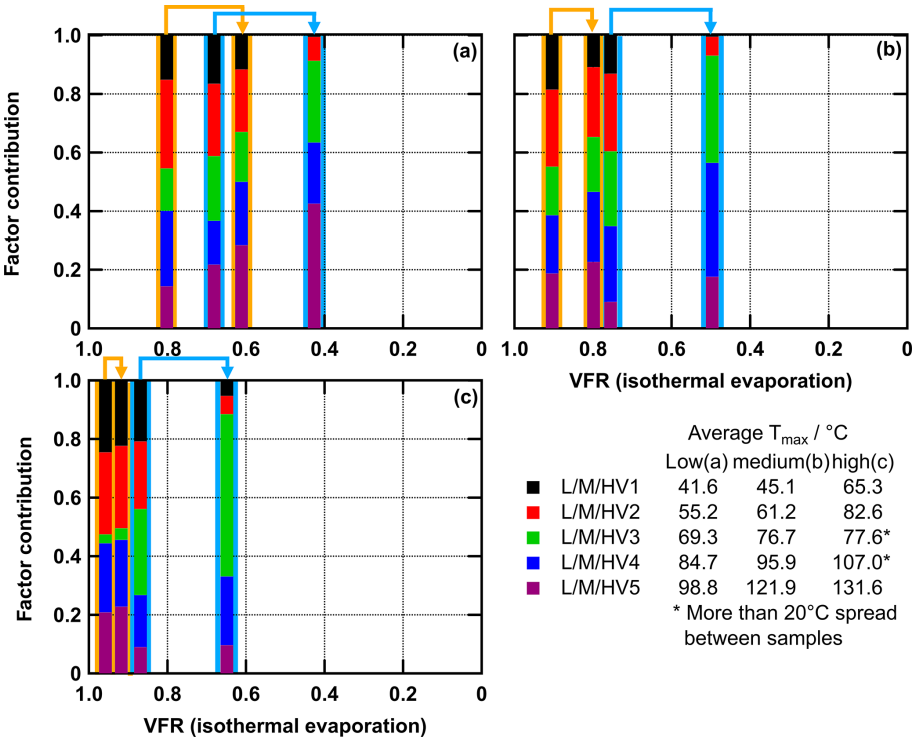

Figure 8. Contribution of type-V factors to total signal for low- **(a)**, medium- **(b)**, and high-O : C cases **(c)**. The x axis is the average volume fraction remaining (VFR) after comparable time intervals of isothermal evaporation observed in separate RTC experiments. Orange and blue arrows indicate the change from $t_{evap} = 0.25$ h to $t_{evap} = 4$ h particles for dry and wet conditions, respectively. Note that the colour code is the same in all panels, but LV1 is not equal to MV1, etc. VFR values are from isothermal evaporation measurements described in Buchholz et al. (2019). Average T_{max} values are for the comparison of the volatility of the factors. Detailed values are given in Table 2.

that, with decreasing VFR, a larger fraction of the particle mass evaporated prior to the FIGAERO–CIMS measurements. In the low- and medium-O : C cases (Fig. 8a and b), the relative contributions of MV1 and 2, and LV1 and 2 (T_{max} in the range of semi-volatile organic compounds, SVOCs), decreased with decreasing VFR while those of LV3–5 and MV3–5 (T_{max} in the range of low- and extremely low volatility organic compounds, LVOCs and ELVOCs) increased. During 4 h of dry isothermal evaporation, a similar volume fraction was removed as in 0.25 h of isothermal evaporation under wet conditions. The very similar relative contribution of the type-V factors in these two samples suggests that the observed changes in chemical composition in the particles are indeed connected to the change in VFR (i.e. how much of the volatile material was removed before sampling) and are not directly driven by other water-induced processes. For these SOA types, the main process during physical ageing in the RTC (i.e. long residence time in clean air) under dry and wet conditions was isothermal particle evaporation. In this case, the particulate water mostly decreased the viscosity in the particles, thus decreasing the kinetic transport limitations in the particle phase and increasing evaporation. This observation is in agreement with previous interpretations of this and other comparable data sets (Buchholz et al., 2019; Yli-Juuti et al., 2017). The high-O : C case (Fig. 8c) will be discussed in Sect. 3.3.

The detailed changes in particle composition, due to isothermal evaporation, can be derived from the factor contribution by analysing the trends in the factor mass spectra. With increasing T_{max} of the factors (i.e. decreasing volatility), the average M_w as well as the carbon-chain length and the number of oxygen increased continuously from V1 to V5 (Table 1). The contribution of compounds with more than 10 carbon atoms (C > 10) also increased, which suggests an increasing contribution of dimers/oligomers. This may explain why no clear trend could be observed for the type-V factors in the O : C (or OS_c) values. While the lower volatility compounds did indeed contain more oxygen, the simultaneous increase of the carbon-chain length seems to compensate for this, which results in no obvious systematic increase in the O : C ratios. Thus, we observe a correlation of the volatility with average M_w but not with the average O : C ratio of the factors.

As the more volatile factors (LV1 and 2; MV1 and 2) were systematically removed with isothermal evaporation, the composition of the residual particles was dominated more and more by the less volatile factors (LV3–5 and MV3–5), i.e. by larger, higher M_w compounds with many of them be-

Table 1. Signal-weighted average values of elemental composition, O : C, OSc, and the contribution of C > 10 compounds for all factors.

ID	Composition	M_w/g mol^{-1}	O : C	OSc	C > 10/%
LV1	$C_{8.6}H_{13.3}O_{5.2}$	213.4	0.66	−0.27	9.3
LV2	$C_{9.0}H_{14.2}O_{5.6}$	200.3	0.64	−0.30	12.7
LV3	$C_{10.3}H_{16.7}O_{6.8}$	249.7	0.70	−0.21	36.5
LV4	$C_{12.6}H_{21.8}O_{7.9}$	300.5	0.66	−0.39	65.2
LV5	$C_{12.5}H_{21.3}O_{8.5}$	308.2	0.72	−0.24	65.1
LD1	$C_{9.8}H_{15.4}O_{6.3}$	235.7	0.74	−0.05	33.8
LB1	$C_{9.7}H_{15.5}O_{6.2}$	234.8	0.78	0.00	38.3
LC1	$C_{8.6}H_{14.3}O_{5.2}$	201.9	0.81	-0.05	28.2
LC2	$C_{6.2}H_{9.5}O_{3.9}$	148.0	0.93	0.21	7.6
MV1	$C_{7.8}H_{11.6}O_{5.0}$	185.8	0.70	−0.12	7.6
MV2	$C_{8.1}H_{11.6}O_{5.8}$	202.3	0.76	0.07	4.1
MV3	$C_{9.0}H_{13.4}O_{6.5}$	226.6	0.76	0.01	15.4
MV4	$C_{10.1}H_{16.0}O_{7.4}$	257.1	0.80	0.02	38.3
MV5	$C_{11.3}H_{18.7}O_{7.6}$	276.5	0.73	−0.17	51.4
MD1	$C_{8.8}H_{13.5}O_{5.9}$	214.8	0.75	−0.01	7.6
MB1	$C_{9.8}H_{15.6}O_{6.1}$	235.5	0.79	0.02	38.9
HV1	$C_{7.0}H_{9.5}O_{5.6}$	184.7	0.90	0.43	6.2
HV2	$C_{7.9}H_{11.1}O_{6.3}$	208.6	0.86	0.29	8.5
HV3	$C_{7.7}H_{10.6}O_{6.3}$	204.3	0.92	0.44	12.9
HV4	$C_{8.4}H_{12.4}O_{6.6}$	219.4	0.87	0.23	18.6
HV5	$C_{10.0}H_{16.2}O_{6.8}$	247.4	0.76	−0.09	39.8
HD1	$C_{8.3}H_{12.3}O_{5.9}$	207.1	0.82	0.18	19.7
HB1	$C_{9.7}H_{15.5}O_{6.1}$	232.8	0.77	−0.02	38.4

Table 2. T_{max} values for all type-V factors. The "–" symbol indicates that there was not enough signal to determine the T_{max} value.

ID	dry ($t_{evap} = 0.25$ h)	dry ($t_{evap} = 4$ h)	wet ($t_{evap} = 0.25$ h)	wet ($t_{evap} = 4$ h)
LV1	37.4	42.5	44.7	–
LV2	51.7	56.5	56.0	56.8
LV3	66.5	70.3	71.2	69.2
LV4	82.0	83.6	86.5	86.6
LV5	95.8	97.6	99.3	102.7
MV1	42.9	44.1	48.1	–
MV2	59.7	58.2	63.7	63.2
MV3	74.9	73.5	78.7	79.6
MV4	93.6	91.6	97.3	101.1
MV5	118.8	116.5	122.5	129.9
HV1	60.7	61.0	75.3	–
HV2	77.2	76.7	93.7	136.5
HV3	58.1	60.0	87.8	104.3
HV4	95.8	94.7	109.0	128.5
HV5	121.6	120.3	136.5	148.0

ing dimers/oligomers. However, the V4 and 5 factors still had a significant contribution from low M_w compounds as well (Figs. 5 and 6). The ion and factor thermograms of $[C_8H_{12}O_5 + I]^-$ are shown as an example of a relatively small, low M_w ion in Fig. 9a and b. This ion had contributions from all 5 factors. In principle, it is possible that there are several isomers of this composition with significantly different volatilities being grouped into V1–5, spreading ~ 4 orders of magnitude in C^*_{sat}. But it seems more likely that the compounds of this composition contributing to V4 and 5 were products of thermal decomposition. If this was indeed the case, it means that there were compounds in the particles that have a volatility which corresponds to even higher T_{max} than that of factors V4 and 5, but they are grouped into these

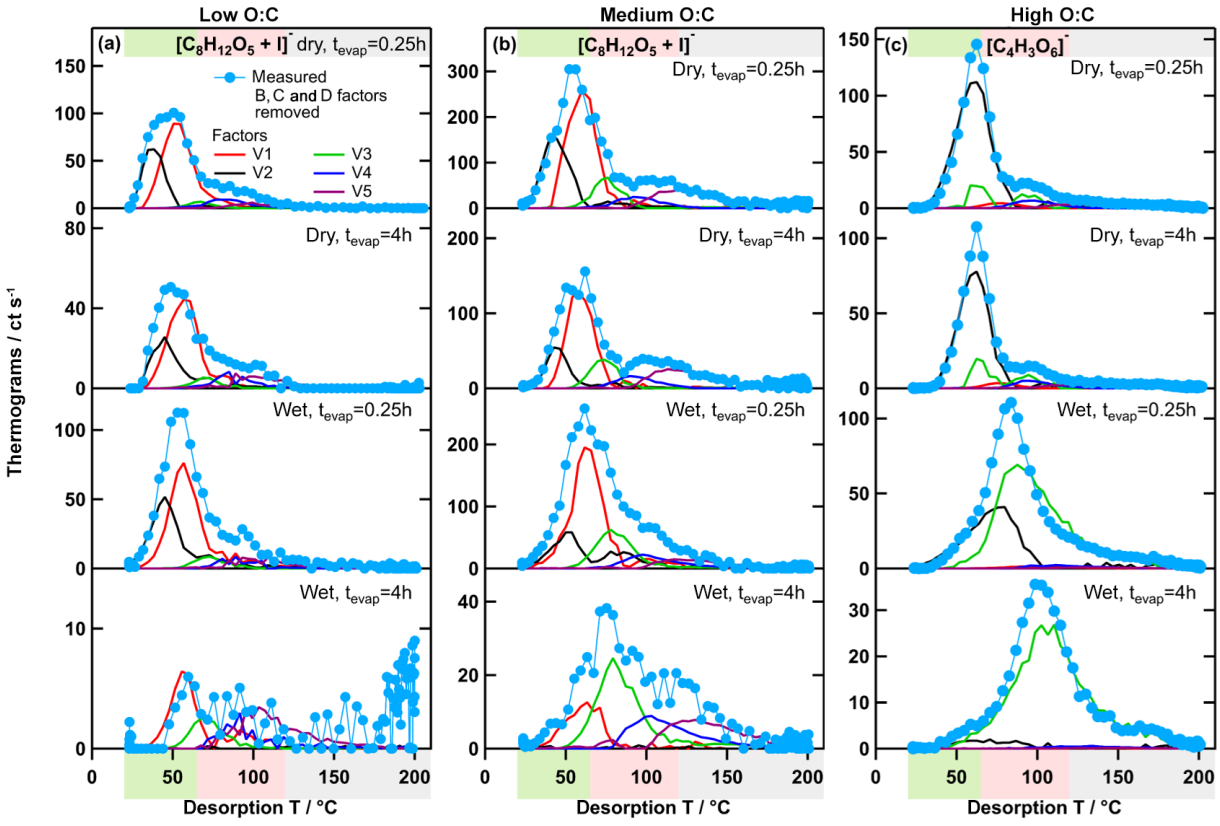

Figure 9. Measured ion thermograms and factor thermograms for ion $[C_8H_{12}O_5 + I]^-$ in the low- **(a)** and medium-O : C cases **(b)** and $C_4H_3O_6^-$ in the high-O : C case **(c)**. Note that the contributions of the type-B, type-C, and type-D factors were subtracted from the measured ion thermograms to enhance comparability between the measured data and the PMF results.

factors/volatility classes because they decompose at desorption temperatures > 100 °C. This is an indication that, as suggested previously, the FIGAERO–CIMS data overestimate the volatility (Lopez-Hilfiker et al., 2015; Schobesberger et al., 2018; Stark et al., 2017), and care has to be taken when using these volatility values for modelling purposes.

3.3 Composition changes due to aqueous-phase chemistry

Similar to the low- and medium-O : C cases, high-O : C SOA particles showed enhanced evaporation under wet conditions (Buchholz et al., 2019). But, in addition, strong signs for aqueous-phase chemistry in the high-O : C wet case were already visible when comparing the mass spectra integrated over the whole thermogram scan. Several very small compounds ($M_w < 200$ Da and C_4–C_7) increased their contribution under wet conditions. Also, the thermograms of these ions showed distinct shifts to higher T_{max} values in the wet cases (by up to 20 °C) and even the formation of new low-volatility material under wet conditions. As discussed by Buchholz et al. (2019), the different behaviour of the high-O : C SOA is most likely due to higher fractions of (hydro-)peroxides in the particles that are caused by the much higher HO_2 concentrations in the OFR at the high-O : C oxidation conditions. Most peroxides are sensitive to hydrolysis which will initiate a range of reactions in the aqueous phase. The low-volatility products of these reactions thermally decompose into similar fragments as the peroxide precursor. Thus, the same groups of ions are detected but at a higher T_{desorp}.

In the PMF analysis results, the different behaviour in the high-O : C case is also directly visible when comparing the dry ($t_{evap} = 0.25$ h) and wet ($t_{evap} = 0.25$ h) cases (Fig. 7a and c). The contribution of the (semi-)volatile factor (HV1) was reduced, but the factor thermograms and T_{max} also changed. HV2 and 4 shifted to higher T_{max} values and a new factor, HV3, was introduced, which contained mostly low M_w compounds. The least volatile factor, HV5, which contains mostly high M_w compounds, had much less of a contribution. It is also noteworthy that HD1 showed a strong increase in the wet case, not only in relative contribution but also in absolute strength. Also, the shape of the factor thermogram (strong increase at $T_{desorp} > 100$ °C) indicates that HD1 was dominated by thermal decomposition products in this case. With further isothermal evaporation under wet conditions, HV3 increased its contribution while HV1 and 2

were almost completely removed (Figs. 7 and 8). Note that HV3 also exhibited an increase in absolute contribution to the signal; i.e. compounds contributing to this factor were being produced (Fig. S7c).

The removal of HV1 can still be explained by particulate water acting as a plasticiser, enhancing the isothermal evaporation comparable to the low- and medium-O : C cases. But HV2 has a T_{max} value already in the LVOC range, like LV3 or MV3, which do not show a similar decrease with isothermal evaporation under wet conditions. Thus, the observed changes can only be explained by chemical processes that are induced by the presence of water in the particles. These processes consume compounds that were mostly grouped into the factors HV2 and HV5. The T_{max} shift of HV1 and HV4 indicates that some of the compounds grouped into these factors might have been affected as well. The reaction products are mostly detected as low M_w compounds in HV3 and HD1. While the compounds grouped into HV3 might still be desorbing from the filter as such, this seems extremely unlikely for the compounds in HD1 as they only start to appear at desorption temperatures > 100 °C. Thus, many of the formed low-volatility compounds must be thermally unstable.

In our previous work (Buchholz et al., 2019), we used the unexpectedly large shift of T_{max} of specific ions, together with the formation of low-volatility material at wet conditions, as evidence for aqueous-phase chemistry in the high-O : C case. With the results from PMF, we can now show how this T_{max} shift in the high-O : C case is indeed different from those smaller ones observed for the other SOA types. The single ion thermograms for $[C_8H_{12}O_5 + I]^-$ (strong ion in low- and medium-O : C samples) and for $C_4H_3O_6^-$ (strong ion in high-O : C samples which was identified to be affected by aqueous-phase chemistry) are shown in Fig. 9. In the low- and medium-O : C cases (Fig. 9a and b), T_{max} changed by ∼ 10 °C between the samples with the least (dry – $t_{evap} = 0.25$ h) and the most isothermal evaporation (wet – $t_{evap} = 4$ h). This shift is solely caused by the removal of LV1 and MV1, and partly LV2 and MV2, i.e. by the isothermal evaporation of the volatile fraction with this composition. In the high-O : C case (Fig. 9c), HV1 is also removed with isothermal evaporation, but the new factor, HV3, dominates under wet conditions. The change in T_{max} by 40 °C between the dry ($t_{evap} = 0.25$ h) case, when HV1 dominates, and the wet ($t_{evap} = 4$ h) case, when HV3 is the only contribution, is then simply the difference in the volatility of the original compounds detected with this composition and the ones formed by aqueous-phase chemistry.

In the dry case, there is a small contribution of HV3 around 100 °C. This is most likely due to the described aqueous-phase processes already happening inside the OFR that was operated at ∼ 40 % RH. The drying during the size selection stopped these processes, which leads to a very minor contribution from the reaction products to the particle phase. If the particle stayed at wet conditions, then the reactions continued and created the compounds that were grouped into HV3. But, apart from this, there has to be another source for the compounds in HV3 in the dry case as there is a small peak at 63 °C. However, this peak is a very minor contribution to the overall signal in the dry case, while HV3 at 100 °C dominates the thermograms in the wet case.

The 7-factor solution presented here clearly identifies the main features in the thermal desorption data that are caused by aqueous-phase processes in the high-O : C case. But this solution still has "mixed" factors containing some compounds that are affected by chemical processes and others that are not, which leads to changes in T_{max} as described previously. Increasing the number of factors in the solution to 10 (results not shown) splits HB1 into 3 type-B and/or type-D factors as well as HV5 into HV5'a, which contributed mostly to the dry cases, and HV5'b, which only occurred in the wet cases. This shows that a better separation of the compounds affected by aqueous-phase processes may be achieved with more factors, but many more factors than 10 may be needed and a strong factor splitting (i.e. artificial separation) may occur.

4 Conclusions

To our knowledge, this is the first study applying a PMF analysis to high-resolution FIGAERO–CIMS thermal desorption data and interpreting the PMF factors as volatility classes characterised by their T_{max} values. Although we used a very specific data set from a focussed laboratory study, the introduced method can be applied to other FIGAERO–CIMS data sets. The nature of PMF allows us to combine multiple separate FIGAERO–CIMS thermograms and investigate them together.

We found that it is very important to study the impact of the chosen "measurement error" on the PMF solutions before interpreting the results of the PMF analysis. Instead of the most realistic measurement error, an error scheme best suited for focussing on the part of the data relevant to the research question should be chosen. In our case, the most interpretable results were achieved by applying a CNerror based on the noise of each ion.

The PMF was able to separate the measured signal of each ion into instrument background, contamination, and collected aerosol mass. This separation worked even if no filter blank data were added to the data sets. However, adding filter blank measurements to the data set simplified the identification of background factors. Identifying background factors in this way, instead of simply subtracting filter blank measurements taken periodically, is especially helpful if an insufficient number of filter blank measurements were collected or if the background changed between filter blank samples. While there was some evidence that a small portion of the background signal was misassigned to other factors (factor mixing/blending) for higher mass loading samples, this did not occur in low-concentration measurements

for which being able to determine the actual contribution of background compounds is more important. At low concentrations, the shape of the combined thermogram of the background may significantly alter the overall shape of the thermogram (e.g. shift the T_{max} value) and thus change the interpretation of the volatility of the collected aerosol.

The collected aerosol mass signal part was separated into (mostly) direct desorption factors (i.e. volatility classes) and thermal decomposition factors. Thermal decomposition became the dominant process for many low M_w ions observed at temperatures above 120 °C. Then the observed "desorption" temperatures are actually the decomposition temperatures and thus give an upper limit for the true volatility of the parent compounds. This shows again that FIGAERO–CIMS measurements may overestimate the volatility of aerosol particles based on the parameterisation of the overall composition but also on desorption temperatures as described in some previous studies (Lopez-Hilfiker et al., 2016; Schobesberger et al., 2018; Stark et al., 2017). The information about the contributions of thermal decomposition to a thermogram measurement obtained with the PMF method that is presented here can be used, for example, to improve the input into process models. An example of such an application is presented in Tikkanen et al. (2019).

For each SOA type (i.e. α-pinene SOA of different oxidative age) five main volatility classes were identified in the chosen PMF solution. Isothermal evaporation prior to sampling with FIGAERO–CIMS systematically removed the more volatile factors with T_{max} values corresponding to SVOCs. Low M_w compounds remaining in the particles after evaporation were attributed to low-volatility factors, which indicates that they were most likely products of thermal decomposition above $\sim$ 100 °C. However, between $\sim$ 100 and 120 °C thermal decomposition was still a minor process. In the high-O : C case, the aqueous-phase chemistry occurring under wet conditions was captured by introducing a new factor and shifts in T_{max} for other factors. Both the educts and products (or their thermal decomposition products) could be identified. This highlights how PMF analysis can help with identifying processes in the particle phase.

The high-O : C SOA in our study may not be representative of ambient SOA with the same O : C ratio as it was formed under extremely strong oxidation conditions in an OFR. But the type of compounds affected by aqueous-phase chemistry (i.e. organic compounds containing (hydro-)peroxides or other functional groups which easily hydrolyse and then continue to react) is not unique to OFR reactors. One formation path of compounds containing several hydroperoxyl or peroxy acid groups is the auto-oxidation of terpenes in the gas phase that leads to highly oxygenated material (HOM; Bianchi et al., 2019; Ehn et al., 2014). These compounds play an important role in particle growth and are detected more and more in ambient measurements (Lee et al., 2018; Mohr et al., 2017). Another compound class that is possibly susceptible to hydrolysis is organo-nitrates (which did not occur in our study due to the experiment design). Thus, ambient aerosol will probably not show as clear signs of aqueous-phase chemistry as our high-O : C case, but it is very likely that such processes occur to some degree and may be detected by the PMF analysis of the FIGAERO thermogram data. We would like to point out that picking the "best" solution of PMF may have a subjective bias, and there is no guarantee that we selected the truly optimal solution. But even if a higher number of factors were chosen, the overall interpretation of the factors would be the same because the additional factors were added in all thermograms in the data set, and they typically split one of the previously identified factors. The influence of the background and thermal decomposition was still separated from the type-V factors, and there was very little variation in T_{max} values in one set of type-V factors for one SOA type. Different degrees of isothermal evaporation of the particles prior to FIGAERO sampling were still reconstructed by decreasing the contribution of the most volatile factors. Only in the high-O : C case, in which the chemical processes altered the particle composition enough, did the factor interpretability improve by increasing the number of factors because the "wet chemistry" factors will likely be completely separated from the other factors at some point. This means there will be multiple factors only occurring in the wet samples, and the T_{max} values of the other factors will be constant. But, as long as the chosen solution can capture the main features of the chemical process, mixed/blended factors are no hindrance for identifying the compounds affected by the chemical processes. Thus, even without a hard criterion to determine the "correct" number of factors, the PMF analysis of FIGAERO–CIMS data gives valuable insight into processes in the particle phase.

The example ions shown in Fig. 9 highlight how important it is to allow a single ion to contribute to more than one class/factor when analysing FIGAERO–CIMS data. Clustering techniques that assign each detected ion/composition to a single cluster, as described by Koss et al. (2020) or Li et al. (2020) for example, are incapable of capturing this behaviour, i.e. the shift of T_{max} between 2 measured thermograms due to the selective removal of some of the isomers/thermal decomposition products. For the data set investigated, we artificially removed the volatile fraction at a set ion composition with the prior isothermal evaporation. However, as the composition of ambient aerosol changes with time, e.g. by changes in the gas-particle partitioning or due to ageing processes, the ratio between different isomers or the educts for thermal decomposition will change and cause similar features in single ion thermograms of FIGAERO–CIMS data.

The next step for this method (PMF analysis of thermal desorption data) is its application to ambient measurements. Typically, the thermal desorption data of each sample in a time series of FIGAERO–CIMS measurements are integrated over the desorption cycle to create a time series of chemical composition that can be analysed with PMF to

identify sources and organic aerosol (OA) types. By including the "extra dimension" of the thermal desorption in the analysis, this time series is changed to a sequence of thermograms. Real time-series-specific (i.e. source-specific) grouping in PMF will then only occur for compounds that also have similar thermograms (i.e. similar volatility). A simplified demonstration of this is found in Sect. S1.3. However, the additional volatility information will help with factor interpretation. Preliminary tests with a data set of ambient FIGAERO–CIMS measurements show that PMF analysis of thermal desorption data creates factors that can be associated with ambient sources (i.e. which precursors and/or processes created the aerosol) and/or OA type (e.g. fresh and aged OA). In addition, the analysis provides detailed information about the volatility of each of these sources, or OA types, while also showing how much of the signal is affected by thermal decomposition. This information about the contribution of thermal decomposition is crucial when the FIGAERO–CIMS data are used to identify the detailed composition or volatility of SOA particles. It will be very interesting to compare the factors identified by PMF of thermal desorption data with the "traditional" PMF analysis of the mass spectra data integrated over each thermogram scan. The details of this investigation will be the content of a future paper.

Appendix A: Case study on the impact of different error schemes

As briefly described in Sect. 2.3.2 and 2.3.3, we investigated the impact of two different error schemes (CNerror and PLerror) on the results of PMF. The high-O : C data set was selected for this case study as the ions affected by aqueous-phase chemistry proved to be the most difficult to capture.

In the case of the PLerror, the residual time series for the total ion signal (Fig. 4d) was positive at all times (i.e. the total reconstructed signal was lower than the measured data) and decreased very little when increasing the factor number from 6 to 10. While the residual time series of individual ions did exhibit negative values (Figs. A1d and A2d), their distribution was still biased towards positive values (i.e. underpredicting the measured data overall). In the case of the CNerror (Fig. 4b) in particular, the residual time series is spread more symmetrically around 0. Additionally, it exhibits much lower values than in the comparable case of the PLerror – particularly for thermograms III and IV (particles under wet conditions).

To illustrate why there is no further improvement in the PMF results with the PLerror scheme and to show at which part of the data set the error schemes create different results, we investigated the behaviour of the PMF solutions for individual ions. We selected two ions with similar signal strength and different error scheme responses. Both error schemes provided comparable PMF results for the ion $[C_7H_8O_6+I]^-$ (Fig. A2) while the thermogram behaviour of the ion $C_5H_5O_6^-$ (Fig. A1) was not captured well with the PLerror scheme. Note that the latter represents the group that mostly contained ions that were affected by aqueous-phase chemistry. For the 6-factor solution (red line in Fig. A1b and d), the residual time series for this ion has similar values for thermogram scans III and IV in both error schemes, but increasing the number of factors by 1 seems to only have a noticeable effect in the case of the CNerror. This is because, in this case, the Q_{ion} values $\left(Q_{\mathrm{ion}} = \left(\frac{\mathbf{E}_{\mathrm{ion}}}{\mathbf{S}_{\mathrm{ion}}}\right)^2\right)$ are extremely high for that part of the data set (red line in panel c). Investigating the Q_i values summed over all observations (mass spectra) shows that this ion ($C_5H_5O_6^-$) has the fifth-highest contribution to overall Q/Q_{exp}. The other ions with such a high single contribution to Q/Q_{exp} exhibit very similar behaviour of their residuals and Q_{ion} values. Together, they account for 15 % of the overall Q/Q_{exp} value in the 6-factor case. So, adding an additional factor to describe that portion of the data set will strongly decrease Q_{ion} and also Q/Q_{exp}, which indicates a better fit. In the case of the PLerror, the Q_{ion} values exhibit very similar profiles for all four thermogram scans (Fig. A1d and e). Thus, changing any parameter for $C_5H_5O_6^-$ will have little effect on the Q_{ion} values and, therefore, on the overall Q/Q_{exp}. This example clearly shows how the selection of the error values guides the focus of PMF, i.e. which part of the data set still needs improvement when the number of factors is increased. In Fig. A3, the contribution of each factor to the signal of $C_5H_5O_6^-$ is shown with coloured areas for the 6- (top) and 7-factor (bottom) solutions for CNerror (a and c) and PLerror (b and d) to highlight the changes caused by the increase of the number of factors for this ion. In addition to reducing the residual for the peaks in thermograms III and IV by using the CNerror, the additional factor substantially alters the factor time series for this ion and is therefore likely affecting our interpretation of these factors – presumably towards improved accuracy. Indeed, the "new" factor F3 was identified in Sect. 3.3 as HV3, which contains the products of the chemical reactions in the aqueous phase.

This error-scheme-dependent performance of PMF is not controlled by the signal strength of the ion or the ratio between signals of combined thermograms. The two example ions were explicitly chosen because of their similar signal strength in all thermograms (compare Figs. A1a and A2a). Instead, it seems that the PLerror does not assign enough weight to the peak region of the ion thermograms. Thus, it cannot resolve the changes in peak shape (i.e. the large shift towards higher desorption temperatures). As the shift is caused by specific processes in the particle phase, PMF with the PLerror will not identify these processes.

These two observations, the CNerror explaining more of the observed variance in general and capturing the complex chemical processes in the particles, lead us to the conclusion that the CNerror yields the more interpretable results for this study and data set and should be used – even though it is not the "true" measurement error of the data.

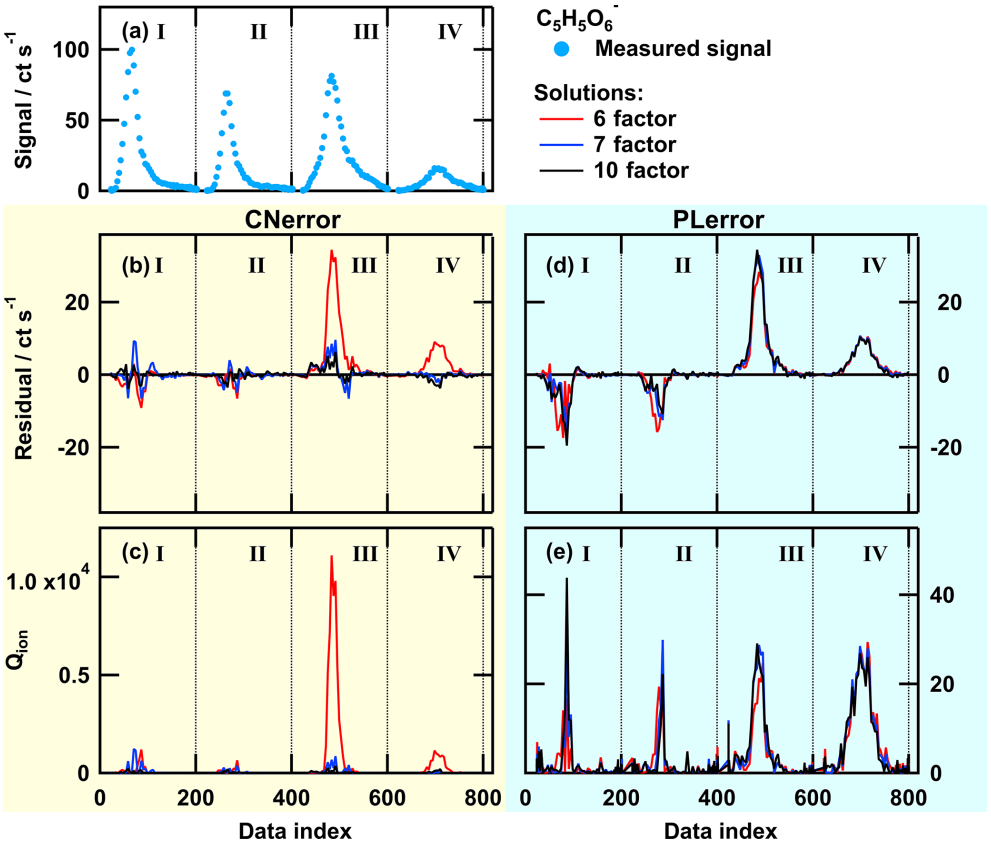

Figure A1. Single ion thermogram **(a)**, residual **(b, d)**, and Q_{ion} values **(c, e)** as time series for solutions with 6, 7, or 10 factors for PMF run with CNerror **(b, c)** and PLerror **(d, e)** for the ion $C_5H_5O_6^-$. The data set contains thermogram scans for high-O : C SOA particles for the following sampling conditions: dry – $t_{\mathrm{evap}} = 0.25$ h (I); dry – $t_{\mathrm{evap}} = 4$ (II); wet – $t_{\mathrm{evap}} = 0.25$ h (III); and wet – $t_{\mathrm{evap}} = 4$ h (IV). Note that the y scaling is the same in panels **(b)** and **(d)**, but it is much smaller in **(e)** than in **(c)**.

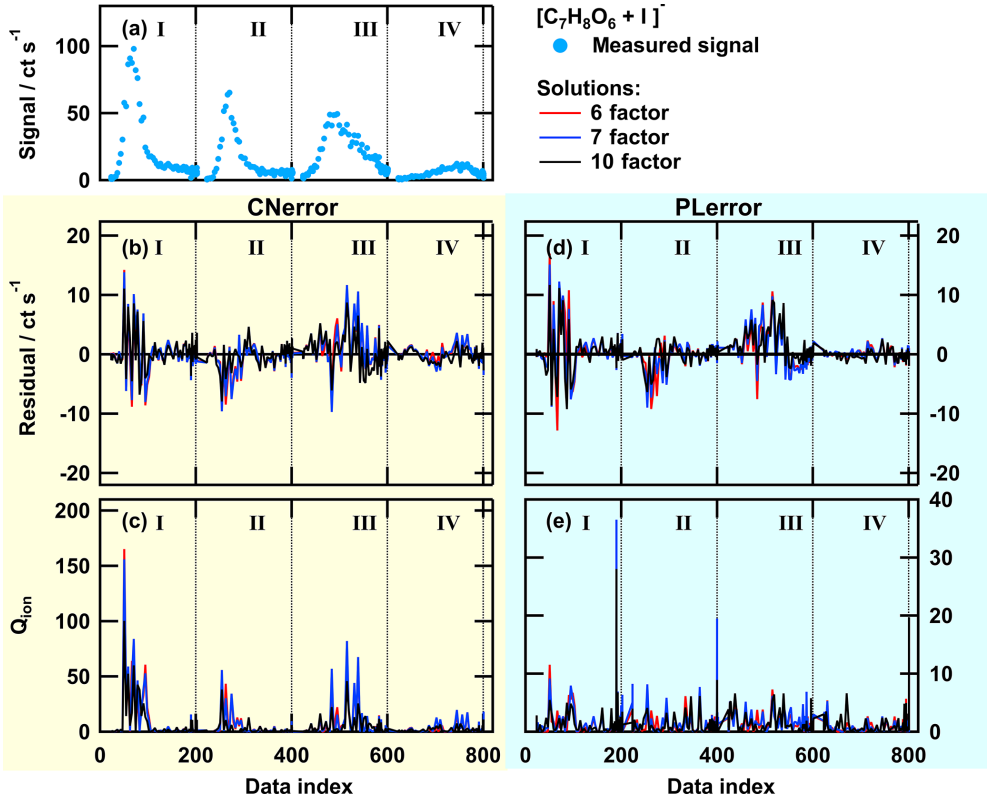

Figure A2. Single ion thermogram **(a)**, residuals **(b, d)**, and Q_{ion} values **(c, e)** as time series for solutions with 6, 7, or 10 factors for PMF run with CNerror **(b, c)** and PLerror **(d, e)** for the ion $[C_7H_8O_6 + I]^-$. The data set contains thermogram scans for high-O : C SOA particles for the following sampling conditions: dry – $t_{\mathrm{evap}} = 0.25$ h (I); dry – $t_{\mathrm{evap}} = 4$ h (II); wet – $t_{\mathrm{evap}} = 0.25$ h (III); and wet – $t_{\mathrm{evap}} = 4$ h (IV). Note that the y scaling in **(e)** is much smaller than in **(c)**.

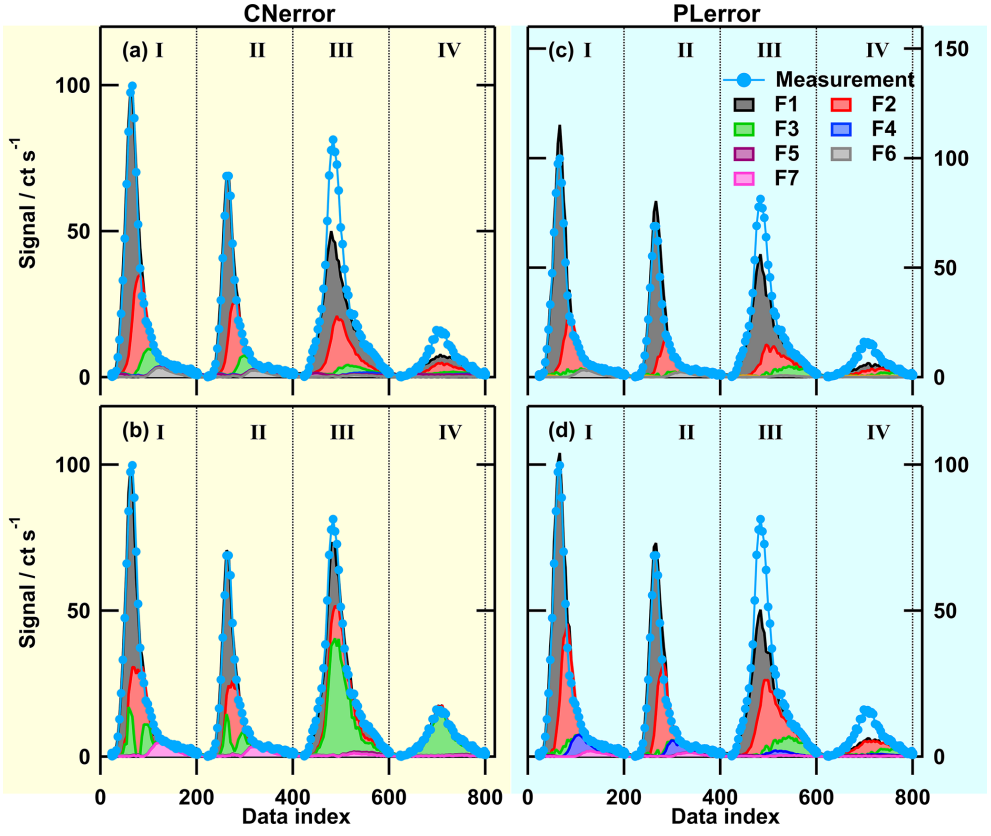

Figure A3. Combined single ion thermograms of the ion $C_5H_5O_6^-$ and PMF factor thermograms for 6- **(a, c)** and 7-factor **(b, d)** solutions. Left column values **(a, b)** are calculated with CNerror and right column values **(c, d)** are calculated with PLerror. The data set contains thermogram scans for high-O : C SOA particles for the following sampling conditions: dry – $t_{evap} = 0.25$ h (I); dry – $t_{evap} = 4$ h (II); wet – $t_{evap} = 0.25$ h (III); and wet – $t_{evap} = 4$ h (IV). Note that generally the factors are not the same between the two error schemes or the two solutions (i.e. F1 in the 6-factor solution with CNerror is different to F1 in the 7-factor solution with CNerror, etc.).

Appendix B: Mathematical symbols and notations used in the equations throughout the paper

Symbol	Explanation
$\mathbf{X}$, $\mathbf{X}_{ij}$	Data matrix ($n \times m$) and data matrix element
p	Number of factors
m	Number of observations (mass spectra) in the data set
n	Number of ions in the data set
$\mathbf{G}$	Factorisation matrix containing the factor thermograms as columns ($n \times p$)
$\mathbf{F}$	Factorisation matrix containing the factor mass spectra as rows ($p \times m$)
$\mathbf{E}$, $\mathbf{E}_{ij}$	Residual matrix and residual matrix element
$\mathbf{R}$, $\mathbf{R}_{ij}$	Reconstructed data matrix ($\mathbf{R} = \mathbf{GF}$) and reconstructed data matrix element
$\mathbf{S}$, $\mathbf{S}_{ij}$	Measurement error matrix and error matrix element
$\mathrm{absVar}_{\mathrm{total}}$	Total absolute variance
$\mathrm{absVar}_{\mathrm{exp}}$	Explained absolute variance
$\mathrm{Ratio}_{\mathrm{exp}}$	Ratio of explained to total absolute variance
Q	Square of the residual scaled with the error summed over all ions and observations (mass spectra)
Q_{exp}	Expected Q value, in the ideal case, with the "true" measurement error equal to $n \times m$
Q_j	Square of the residual scaled with the error summed over all observations (mass spectra)
Q_i	Square of the residual scaled with the error summed over all ions
Q_{ion}	Square of the residual scaled with the error for a single ion as time series
Q/Q_{exp}	Optimisation parameter in PMF

Data availability. The data set used in this study is available upon request from the corresponding author.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/acp-20-7693-2020-supplement.

Author contributions. AB, AY, and CM conducted the FIGAERO–CIMS measurements. The data were processed by AY, WH, and CM. AB conducted the PMF analysis. All authors participated in the interpretation of the data. AB wrote the manuscript with contributions from all the co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to acknowledge the support of Celia Faiola, Eetu Kari, Liqing Hao, and Sergey Nizkorodov during the collection of this data set. Furthermore, we wish to thank Andrew Lambe for his help with installing PAM-OFR.

Financial support. This research has been supported by the European Research Council 2019 Starting Grant (ERC-2019-StG (grant no. QAPPA 335478)), the Centre of Excellence programme of the Academy of Finland (grant no. 307331), and the Academy of Finland (grant nos. 299544, 317373, and 310682).

Review statement. This paper was edited by Joachim Curtius and reviewed by three anonymous referees.

References

Bannan, T. J., Le Breton, M., Priestley, M., Worrall, S. D., Bacak, A., Marsden, N. A., Mehra, A., Hammes, J., Hallquist, M., Alfarra, M. R., Krieger, U. K., Reid, J. P., Jayne, J., Robinson, W., McFiggans, G., Coe, H., Percival, C. J., and Topping, D.: A method for extracting calibrated volatility information from the FIGAERO-HR-ToF-CIMS and its experimental application, Atmos. Meas. Tech., 12, 1429–1439, https://doi.org/10.5194/amt-12-1429-2019, 2019.

Bianchi, F., Kurtén, T., Riva, M., Mohr, C., Rissanen, M. P., Roldin, P., Berndt, T., Crounse, J. D., Wennberg, P. O., Mentel, T. F., Wildt, J., Junninen, H., Jokinen, T., Kulmala, M., Worsnop, D. R., Thornton, J. A., Donahue, N., Kjaergaard, H. G., and Ehn, M.: Highly Oxygenated Organic Molecules (HOM) from Gas-Phase Autoxidation Involving Peroxy Radicals: A Key Contributor to Atmospheric Aerosol, Chem. Rev., 119, 3472–3509, https://doi.org/10.1021/acs.chemrev.8b00395, 2019.

Buchholz, A., Lambe, A. T., Ylisirniö, A., Li, Z., Tikkanen, O.-P., Faiola, C., Kari, E., Hao, L., Luoma, O., Huang, W., Mohr, C., Worsnop, D. R., Nizkorodov, S. A., Yli-Juuti, T., Schobesberger, S., and Virtanen, A.: Insights into the O : C-dependent mechanisms controlling the evaporation of α-pinene secondary organic aerosol particles, Atmos. Chem. Phys., 19, 4061–4073, https://doi.org/10.5194/acp-19-4061-2019, 2019.

D'Ambro, E. L., Schobesberger, S., Zaveri, R. A., Shilling, J. E., Lee, B. H., Lopez-Hilfiker, F. D., Mohr, C., and Thornton, J. A.: Isothermal Evaporation of α-Pinene Ozonolysis SOA: Volatility, Phase State, and Oligomeric Composition, ACS Earth Space Chem., 2, 1058–1067, https://doi.org/10.1021/acsearthspacechem.8b00084, 2018.

D'Ambro, E. L., Schobesberger, S., Gaston, C. J., Lopez-Hilfiker, F. D., Lee, B. H., Liu, J., Zelenyuk, A., Bell, D., Cappa, C. D., Helgestad, T., Li, Z., Guenther, A., Wang, J., Wise, M., Caylor, R., Surratt, J. D., Riedel, T., Hyttinen, N., Salo, V.-T., Hasan, G., Kurtén, T., Shilling, J. E., and Thornton, J. A.: Chamber-based insights into the factors controlling epoxydiol (IEPOX) secondary organic aerosol (SOA) yield, composition, and volatility, Atmos. Chem. Phys., 19, 11253–11265, https://doi.org/10.5194/acp-19-11253-2019, 2019.

Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N.: Coupled partitioning, dilution, and chemical aging of semivolatile organics, Environ. Sci. Technol., 40, 2635–2643, https://doi.org/10.1021/es052297c, 2006.

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann, R., Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I. H., Rissanen, M., Jokinen, T., Schobesberger, S., Kangasluoma, J., Kontkanen, J., Nieminen, T., Kurtén, T., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., Canagaratna, M., Maso, M. D., Berndt, T., Petäjä, T., Wahner, A., Kerminen, V. M., Kulmala, M., Worsnop, D. R., Wildt, J., and Mentel, T. F.: A large source of low-volatility secondary organic aerosol, Nature, 506, 476–479, https://doi.org/10.1038/nature13032, 2014.

Glasius, M. and Goldstein, A. H.: Recent Discoveries and Future Challenges in Atmospheric Organic Chemistry, Environ. Sci. Technol., 50, 2754–2764, https://doi.org/10.1021/acs.est.5b05105, 2016.

Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, Environ. Sci. Technol., 41, 1514–1521, https://doi.org/10.1021/es072476p, 2007.

Herrmann, H.: Kinetics of Aqueous Phase Reactions Relevant for Atmospheric Chemistry, Chem. Rev., 103, 4691–4716, https://doi.org/10.1021/cr020658q, 2003.

Hoyer, P. O.: Non-negative matrix factorization with sparseness constraints, J. Mach. Learn. Res., 5, 1457–1469, 2004.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb,

C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols in the 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.

Koss, A. R., Canagaratna, M. R., Zaytsev, A., Krechmer, J. E., Breitenlechner, M., Nihill, K. J., Lim, C. Y., Rowe, J. C., Roscioli, J. R., Keutsch, F. N., and Kroll, J. H.: Dimensionality-reduction techniques for complex mass spectrometric datasets: application to laboratory atmospheric organic oxidation experiments, Atmos. Chem. Phys., 20, 1021–1041, https://doi.org/10.5194/acp-20-1021-2020, 2020.

Kroll, J. H. and Seinfeld, J. H.: Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, Atmos. Environ., 42, 3593–3624, https://doi.org/10.1016/j.atmosenv.2008.01.003, 2008.

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.

Lee, B. H., Lopez-Hilfiker, F. D., Mohr, C., Kurtén, T., Worsnop, D. R., and Thornton, J. A.: An iodide-adduct high-resolution time-of-flight chemical-ionization mass spectrometer: Application to atmospheric inorganic and organic compounds, Environ. Sci. Technol., 48, 6309–6317, https://doi.org/10.1021/es500362a, 2014.

Lee, B. H., Lopez-Hilfiker, F. D., D'Ambro, E. L., Zhou, P., Boy, M., Petäjä, T., Hao, L., Virtanen, A., and Thornton, J. A.: Semi-volatile and highly oxygenated gaseous and particulate organic compounds observed above a boreal forest canopy, Atmos. Chem. Phys., 18, 11547–11562, https://doi.org/10.5194/acp-18-11547-2018, 2018.

Li, Z., D'Ambro, E. L., Schobesberger, S., Gaston, C. J., Lopez-Hilfiker, F. D., Liu, J., Shilling, J. E., Thornton, J. A., and Cappa, C. D.: A robust clustering algorithm for analysis of composition-dependent organic aerosol thermal desorption measurements, Atmos. Chem. Phys., 20, 2489–2512, https://doi.org/10.5194/acp-20-2489-2020, 2020.

Liu, P., Li, Y. J., Wang, Y., Gilles, M. K., Zaveri, R. A., Bertram, A. K., and Martin, S. T.: Lability of secondary organic particulate matter, P. Natl. Acad. Sci. USA, 113, 12643–12648, https://doi.org/10.1073/pnas.1603138113, 2016.

Lopez-Hilfiker, F. D., Mohr, C., Ehn, M., Rubach, F., Kleist, E., Wildt, J., Mentel, Th. 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.

Lopez-Hilfiker, F. D., Mohr, C., Ehn, M., Rubach, F., Kleist, E., Wildt, J., Mentel, Th. F., Carrasquillo, A. J., Daumit, K. E., Hunter, J. F., Kroll, J. H., Worsnop, D. R., and Thornton, J. A.: Phase partitioning and volatility of secondary organic aerosol components formed from α-pinene ozonolysis and OH oxidation: the importance of accretion products and other low volatility compounds, Atmos. Chem. Phys., 15, 7765–7776, https://doi.org/10.5194/acp-15-7765-2015, 2015.

Lopez-Hilfiker, F. D., Mohr, C., D'Ambro, E. L., Lutz, A., Riedel, T. P., Gaston, C. J., Iyer, S., Zhang, Z., Gold, A., Surratt, J. D., Lee, B. H., Kurten, T., Hu, W. W., Jimenez, J., Hallquist, M., and Thornton, J. A.: Molecular Composition and Volatility of Organic Aerosol in the Southeastern U.S.: Implications for IEPOX Derived SOA, Environ. Sci. Technol., 50, 2200–2209, https://doi.org/10.1021/acs.est.5b04769, 2016.

Lu, J. and Wu, L.: Technical details and programming guide for a general two-way positive matrix factorization algorithm, J. Chemometr., 18, 519–525, https://doi.org/10.1002/cem.894, 2004.

Mohr, C., Lopez-Hilfiker, F. D., Yli-Juuti, T., Heitto, A., Lutz, A., Hallquist, M., D'Ambro, E. L., Rissanen, M. P., Hao, L., Schobesberger, S., Kulmala, M., Mauldin, R. L., Makkonen, U., Sipilä, M., Petäjä, T., and Thornton, J. A.: Ambient observations of dimers from terpene oxidation in the gas phase: Implications for new particle formation and growth, Geophys. Res. Lett., 44, 2958–2966, https://doi.org/10.1002/2017GL072718, 2017.

Paatero, P.: The Multilinear Engine–A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n-Way Parallel Factor Analysis Model, J. Comput. Graph. Stat., 8, 854–888, https://doi.org/10.1080/10618600.1999.10474853, 1999.

Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, in: Analytica Chimica Acta, 490, 277–289, Elsevier, 2003.

Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111–126, https://doi.org/10.1002/env.3170050203, 1994.

Pankow, J. F.: An absorption model of gas/particle partitioning of organic compounds in the atmosphere, Atmos. Environ., 28, 185–188, https://doi.org/10.1016/1352-2310(94)90093-0, 1994a.

Pankow, J. F.: An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic aerosol, Atmos. Environ., 28, 189–193, https://doi.org/10.1016/1352-2310(94)90094-9, 1994b.

Pankow, J. F., Seinfeld, J. H., Asher, W. E., and Erdakos, G. B.: Modeling the formation of secondary organic aerosol. 1. Application of theoretical principles to measurements obtained in the α-pinene/, β-pinene/, sabinene/, Δ3-carene/, and cyclohexene/ozone systems, Environ. Sci. Technol., 35, 1164–1172, https://doi.org/10.1021/es001321d, 2001.

Renbaum-Wolff, L., Grayson, J. W., Bateman, A. P., Kuwata, M., Sellier, M., Murray, B. J., Shilling, J. E., Martin, S. T., and Bertram, A. K.: Viscosity of α-pinene secondary organic material and implications for particle growth and reactivity, P. Natl. Acad. Sci. USA, 110, 8014–8019, https://doi.org/10.1073/pnas.1219548110, 2013.

Schobesberger, S., D'Ambro, E. L., Lopez-Hilfiker, F. D., Mohr, C., and Thornton, J. A.: A model framework to retrieve thermodynamic and kinetic properties of organic aerosol from composition-resolved thermal desorption measurements, Atmos. Chem. Phys., 18, 14757–14785, https://doi.org/10.5194/acp-18-14757-2018, 2018.

Stark, H., Yatavelli, R. L. N., Thompson, S. L., Kang, H., Krechmer, J. E., Kimmel, J. R., Palm, B. B., Hu, W., Hayes, P. L.,

Day, D. A., Campuzano-Jost, P., Canagaratna, M. R., Jayne, J. T., Worsnop, D. R., and Jimenez, J. L.: Impact of Thermal Decomposition on Thermal Desorption Instruments: Advantage of Thermogram Analysis for Quantifying Volatility Distributions of Organic Species, Environ. Sci. Technol., 51, 8491–8500, https://doi.org/10.1021/acs.est.7b00160, 2017.

Surratt, J. D., Lewandowski, M., Offenberg, J. H., Jaoui, M., Kleindienst, T. E., Edney, E. O., and Seinfeld, J. H.: Effect of acidity on secondary organic aerosol formation from isoprene, Environ. Sci. Technol., 41, 5363–5369, https://doi.org/10.1021/es0704176, 2007.

Tikkanen, O.-P., Buchholz, A., Ylisirniö, A., Schobesberger, S., Virtanen, A., and Yli-Juuti, T.: Comparing SOA volatility distributions derived from isothermal SOA particle evaporation data and FIGAERO-CIMS measurements, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-927, in review, 2019.

Tolocka, M. P., Jang, M., Ginter, J. M., Cox, F. J., Kamens, R. M., and Johnston, M. V.: Formation of Oligomers in Secondary Organic Aerosol, Environ. Sci. Technol., 38, 1428–1434, https://doi.org/10.1021/es035030r, 2004.

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

Virtanen, A., Joutsensaari, J., Koop, T., Kannosto, J., Yli-Pirilä, P., Leskinen, J., Mäkelä, J. M., Holopainen, J. K., Pöschl, U., Kulmala, M., Worsnop, D. R., and Laaksonen, A.: An amorphous solid state of biogenic secondary organic aerosol particles, Nature, 467, 824–827, https://doi.org/10.1038/nature09455, 2010.

Wang, D. S. and Hildebrandt Ruiz, L.: Chlorine-initiated oxidation of n-alkanes under high-NO_x conditions: insights into secondary organic aerosol composition and volatility using a FIGAERO–CIMS, Atmos. Chem. Phys., 18, 15535–15553, https://doi.org/10.5194/acp-18-15535-2018, 2018.

Wilson, J., Imre, D., Beránek, J., Shrivastava, M., and Zelenyuk, A.: Evaporation kinetics of laboratory-generated secondary organic aerosols at elevated relative humidity, Environ. Sci. Technol., 49, 243–249, https://doi.org/10.1021/es505331d, 2015.

Yan, C., Nie, W., Äijälä, M., Rissanen, M. P., Canagaratna, M. R., Massoli, P., Junninen, H., Jokinen, T., Sarnela, N., Häme, S. A. K., Schobesberger, S., Canonaco, F., Yao, L., Prévôt, A. S. H., Petäjä, T., Kulmala, M., Sipilä, M., Worsnop, D. R., and Ehn, M.: Source characterization of highly oxidized multifunctional compounds in a boreal forest environment using positive matrix factorization, Atmos. Chem. Phys., 16, 12715–12731, https://doi.org/10.5194/acp-16-12715-2016, 2016.

Yli-Juuti, T., Pajunoja, A., Tikkanen, O. P., Buchholz, A., Faiola, C., Väisänen, O., Hao, L., Kari, E., Peräkylä, O., Garmash, O., Shiraiwa, M., Ehn, M., Lehtinen, K., and Virtanen, A.: Factors controlling the evaporation of secondary organic aerosol from a-pinene ozonolysis, Geophys. Res. Lett., 44, 2562–2570, https://doi.org/10.1002/2016GL072364, 2017.