

Supplement Information

1.1 Experimental setup of isothermal evaporation experiments

SOA was formed via combined photooxidation and ozonolysis of α -pinene in a potential aerosol mass reactor (PAM, Aerodyne Research Inc., Kang et al., 2007; Lambe et al., 2011). The α -pinene concentration (190 ppb), relative humidity (RH = 40%) and temperature (27 °C) in the PAM reactor were held constant throughout the experiments, but three different oxidative settings were chosen to create three types of α -pinene SOA characterised by their oxidative age (Table S 2). The gas phase was monitored with a high-resolution time-of-flight proton transfer reaction mass spectrometer (PTR-MS, Ionicon model 8000) while the particles were characterised with a scanning mobility particle sizer (SMPS, TSI Inc., model 3082+3775), a high resolution time-of-flight aerosol mass spectrometer (AMS, Aerodyne Research Inc., DeCarlo et al., 2006), and a filter inlet for gases and aerosols sampling unit (FIGAERO, Aerodyne Research Inc., Lopez-Hilfiker et al., 2014) in combination with a chemical ionisation mass spectrometer (CIMS, Aerodyne Research Inc., Lee et al., 2014) using iodide as reagent ion.

Two types of particles samples were collected with the FIGAERO-CIMS: “fresh” particles (labelled $t_{\text{evap}} = 0.25\text{h}$), size selected with a nano differential mobility analyser (NanoDMA, 80 nm electro mobility size) and “RTC” particles (labelled $t_{\text{evap}} = 4\text{h}$) which were left to evaporate at ~ 20 °C for 3 - 4 h in a residence time chamber (RTC) prior to collection on the FIGAERO filter. The isothermal evaporation method has been described previously in detail (Buchholz et al., 2019; Yli-Juuti et al., 2017). The NanoDMA was operated with an open loop sheath flow which together with the extremely short residence time inside the NanoDMA (≤ 0.3 s) limited the diffusion of gaseous compounds into the selected sample flow. This created a sudden shift in the gas-particle equilibrium which initiates the particle evaporation at the NanoDMA outlet. For evaporation experiments, this monodisperse particle sample was filled for 75 min into a 100 L stainless steel RTC and the RTC then closed off. After 3 – 4 h of isothermal evaporation, the RTC was reopened and the remaining particles sampled onto the FIGAERO filter. In addition, filter blank measurements were conducted on each experiment day to quantify the instrument background.

Generally, a mass loading of several 10s of ng is needed on the FIGAERO filter to achieve a sufficiently high signal to noise ratio in CIMS. In this study, 20 or 30 min of collection time were necessary to accumulate enough particular mass on the FIGAERO filter. The mass loadings in the RTC at the beginning of the evaporation experiment and on the FIGAERO filter estimated from AMS measurements are given in Table S 1. It has to be noted that in this specific setup, particles already collected on the FIGAERO filter will continue to evaporate during the remaining collection time as no new gas-particle equilibrium can be reached after the dilution in the NanoDMA. Thus, the “fresh” particles on the

FIGAERO filter actually represent particles with an average evaporation time of 15 min and may have lost some of the very volatile compounds before the thermal desorption begins. Particles sampled after hours of evaporation in the RTC exhibit the same behaviour. But as the majority of particle evaporation occurs within the first 1 h, the additional residence time on the FIGAERO filter does not contribute significantly to the overall evaporation (Buchholz et al., 2019).

5 1.2 PLError calculation

The error for mass spectra data consists of two parts: the counting statistic error and instrument noise (electronic background noise). Assuming a Poisson type distribution of the counting error the total error S_{ij} for each ion i at time j can be described as (same as Eq. 7 in main text):

$$S_{ij} = a \cdot \sqrt{\frac{X_{ij}}{t_s}} + \sigma_{noise,i} \quad (S1)$$

10 with X_{ij} signal intensity of the ion i , t_s sampling (averaging) interval in s, $\sigma_{noise,i}$ the electronic noise for ion i , and the empirically derived parameters a . Yan et al. (2016) suggested a method to derive a from a dataset of CIMS gas-phase measurements utilising that for most part of that dataset real changes were much slower than the noise in the data. The problem with FIGAERO data is that the changes in signal due to temperature are fast, and it is difficult to separate the noise/error from these changes. However, most ion traces reach a steady state in the last minutes of the measurements as
 15 most of the material has evaporated already and background signals are stable or slowly changing. Additionally, the temperature ramp reaches a plateau (“soak” period). Thus, we modified the method from Yan et al. (2016). The last 20 data points of each thermogram (400 sec, ~195 – 205 °C) are considered to be in steady state. To account for still persisting trends in the data, a linear fit (LineFit) is performed to these points for each ion trace. The residual between the data and this fit, res_{ij} , is calculated for each ion (20 values per ion):

$$20 \quad res_{ij} = data_{ij} - LineFit_{ij} \quad (S2)$$

Yan et al. (2016) showed that the analytical uncertainty in gas-phase CIMS measurements is independent of the m/z range of the instrument and the specific ion. Thus, the values of all ions can be combined for further analysis. σ_{noise} is calculated as the median of the standard deviation of res_{ij} for one ion (Figure S 8).

$$\sigma_{noise} = median(stdev(res_{ij})) \quad (S3)$$

25 These σ_{noise} values are used in the CNerror scheme (Eq. 8 in main manuscript). Note that the values are higher for ions which have a higher plateau value, i.e., which show higher desorption during the end of the soak period. As this is an indication for a stronger impact of background compounds from the instrument/filter, it seemed justified to apply this higher error value for the whole thermogram of these ions, decreasing their weight in the PMF optimisation algorithm.

To account for the minimum error needed in the PMF algorithm, σ_{noise} values smaller than the median of all σ_{noise} values ($\sigma_{min} = 0.013$) were set to the σ_{min} value.

Similar to the method in Yan et al. (2016), the res_{ij} values are grouped by signal strength (X_{ij}) into 12 classes. The intervals were chosen to have at least 30 values in each class and the first class starts at $\sim 1 \text{ ct s}^{-1}$ (i.e., $X > 60\%$ of max signal, 50%-60%, 40%-50%, 30%-40%, 25%-30%, 20%-25%, 15%-20%, 10%-15%, 5%-10%, 2%-5%, 1%-2%, 0-1%). Note, that this is the average signal strength for the last 400 sec of the thermogram. The peak values during desorption may be much higher. A histogram is calculated for each class and the data is fitted with a gauss function (examples in Figure S 9b and c). The half-width of that function (σ_{fit} : “width” in plots) is considered to be the overall measurement error S_{ij} .

To gain the factor a in equation 1, σ_{fit} is plotted vs the average signal intensity of each class (I_{class}) and a power law is fitted:

$$S_{ij} = \sigma_{fit} = a' \cdot (I_{class})^c + b \quad S4$$

With $a' = a/\sqrt{(t_s)}$, $b = \sigma_{noise}$, and c should be 0.5 (i.e., the square root for the case where the error is following a Poisson distribution). The 95th percentile of σ_{fit} is used as “one-sigma-standard-deviation” to weight the data points. Figure S 9 shows the combined S_{ij} dataset for all thermograms for all OC cases (4*12 points per OC case). The parameters yielded by freely fitting Eq. S4 (black line in Figure S 9) and fitting with $c = 0.5$ (grey line) are given in. Note that the value for c is > 0.5 in the unconstrained case. This may be caused by some random error on the data, the fact that the error does not follow a Poisson style distribution, or because there are very few points with good quality above signal strength 10 constraining the function. To avoid assumptions about the exact shape of the error distribution, we decided to use the parameters of the unconstrained fit to Eq. S4 to calculate the error values labelled PLError. In addition, we apply the same minimum error criterium as in the CNerror case.

As a test case not shown in the manuscript, PMF analysis was also conducted with error values calculated according to Eq. S1 with $\sigma_{noise} = 0.013$ and $a = 1$ or $a = 1.28$ (value typically used for AMS data). The results were very similar to the ones with the PLError scheme.

1.3 SI References

- 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(6), 4061–4073, doi:10.5194/acp-19-4061-2019, 2019.
- Decarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath,

- T., Docherty, K. S., Worsnop, D. R. and Jimenez, J. L.: Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, *Anal. Chem.*, 78(24), 8281–8289, doi:10.1021/ac061249n, 2006.
- 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(8), 2635–2643, doi:10.1021/es052297c, 2006.
- 5 Kang, E., Root, M. J., Toohey, D. W. and Brune, W. H.: Introducing the concept of Potential Aerosol Mass (PAM), *Atmos. Chem. Phys.*, 7(22), 5727–5744, doi:10.5194/acp-7-5727-2007, 2007.
- Lambe, A. T., Ahern, A. T., Williams, L. R., Slowik, J. G., Wong, J. P. S., Abbatt, J. P. D., Brune, W. H., Ng, N. L., Wright, J. P., Croasdale, D. R., Worsnop, D. R., Davidovits, P. and Onasch, T. B.: Characterization of aerosol photooxidation flow reactors: heterogeneous oxidation, secondary organic aerosol formation and cloud condensation nuclei activity measurements, *Atmos. Meas. Tech.*, 4, 445–461, doi:10.5194/amt-4-445-2011, 2011.
- 10 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(11), 6309–6317, doi:10.1021/es500362a, 2014.
- Lopez-Hilfiker, F. D., Mohr, C., Ehn, M., Rubach, F., Kleist, E., Wildt, J., Mentel, T. F., Lutz, A., Hallquist, M., Worsnop, D. and Thornton, J. A.: A novel method for online analysis of gas and particle composition: description and evaluation of a Filter Inlet for Gases and AEROsols (FIGAERO), *Atmos. Meas. Tech.*, 7, 983–1001, doi:10.5194/amt-7-983-2014, 2014.
- 15 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(19), 12715–12731, doi:10.5194/acp-16-12715-2016, 2016.
- 20 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(5), 2562–2570, doi:10.1002/2016GL072364, 2017.

25

1.4 SI Tables

Table S 1: Mass concentration in RTC after filling and estimated collected sample mass on FIGAERO filter.

OH exposure	condition	mass conc in RTC after filling / $\mu\text{g m}^{-3}$	sampled mass / ng	
			$t_{\text{evap}}=0.25\text{h}$	$t_{\text{evap}}=4\text{ h}$
low	dry	0.67	178	33
	RH80%	0.52	186	22
medium	dry	1.0	239	72
	RH80%	1.1	258	50
high	dry	0.69	138	46
	RH80%	0.98	172	30

Table S 2: Overview on oxidation condition and oxidative state of formed SOA.

	low	medium	high
$[\text{O}_3]_{\text{inlet}} / \text{ppm}$	6.6	25	25
$[\text{O}_3]_{\text{outlet}} / \text{ppm}$	6.4	22.2	16
OH exposure / $\text{cm}^{-3} \text{ s}$	2.54e11	6.85e11	2.45e12
photochemical age / days	2.0	5.3	18.9
O:C(AMS)	0.53	0.69	0.96
O:C(FIG)	0.66	0.75	0.84

5

Table S 3: Fitting parameters for fitting Eq. S4 to the S_{ij} data presented in Figure S 9.

	a'	b	c
free fit	0.260 ± 0.003	0.056 ± 0.002	0.726 ± 0.009
square root fit	0.303 ± 0.002	0.009 ± 0.001	0.5

1.5 SI Figures

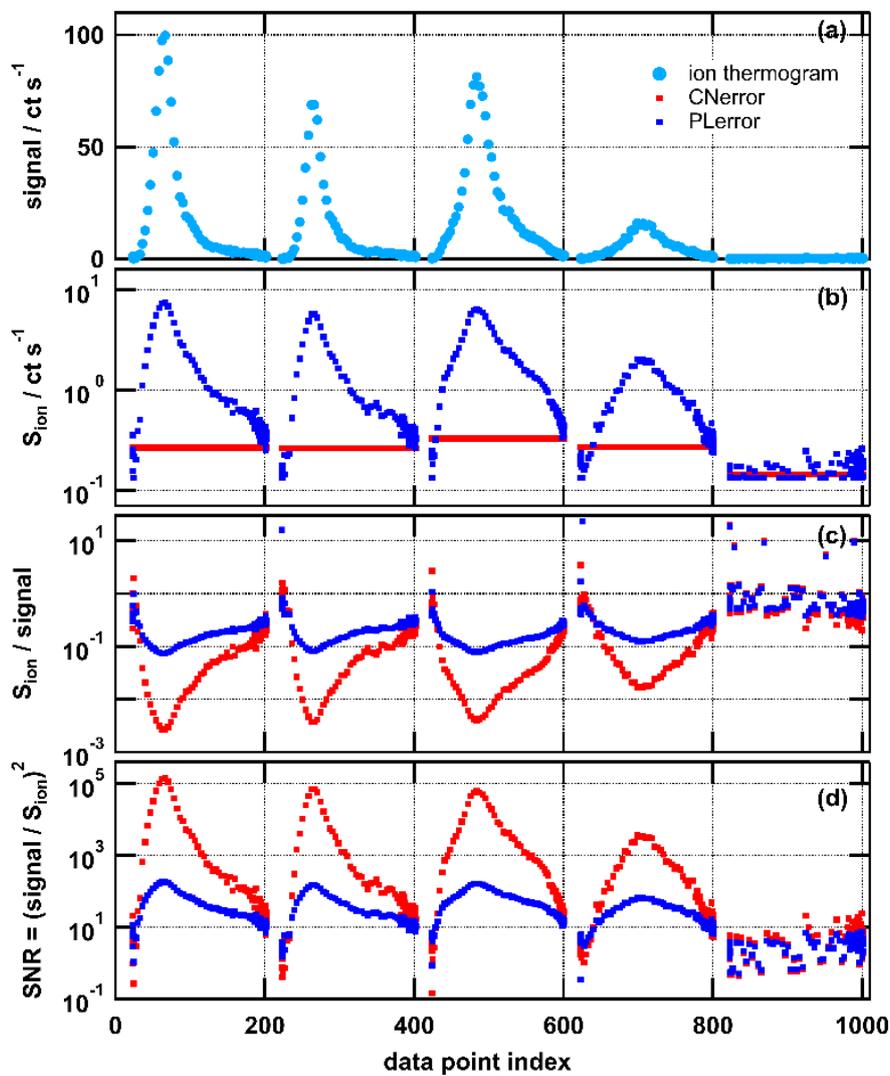


Figure S 1: Single ion thermogram (a) and time series of errors (S_{ion} , b), relative error (S_{ion} / signal , c), and signal-to-noise-ratio (SNR, d) for $\text{C}_5\text{H}_5\text{O}_6^-$. Values calculated for CError are marked in red and those for PError in dark blue. Note that panels b, c, and d have logarithmic y axis scaling.

5

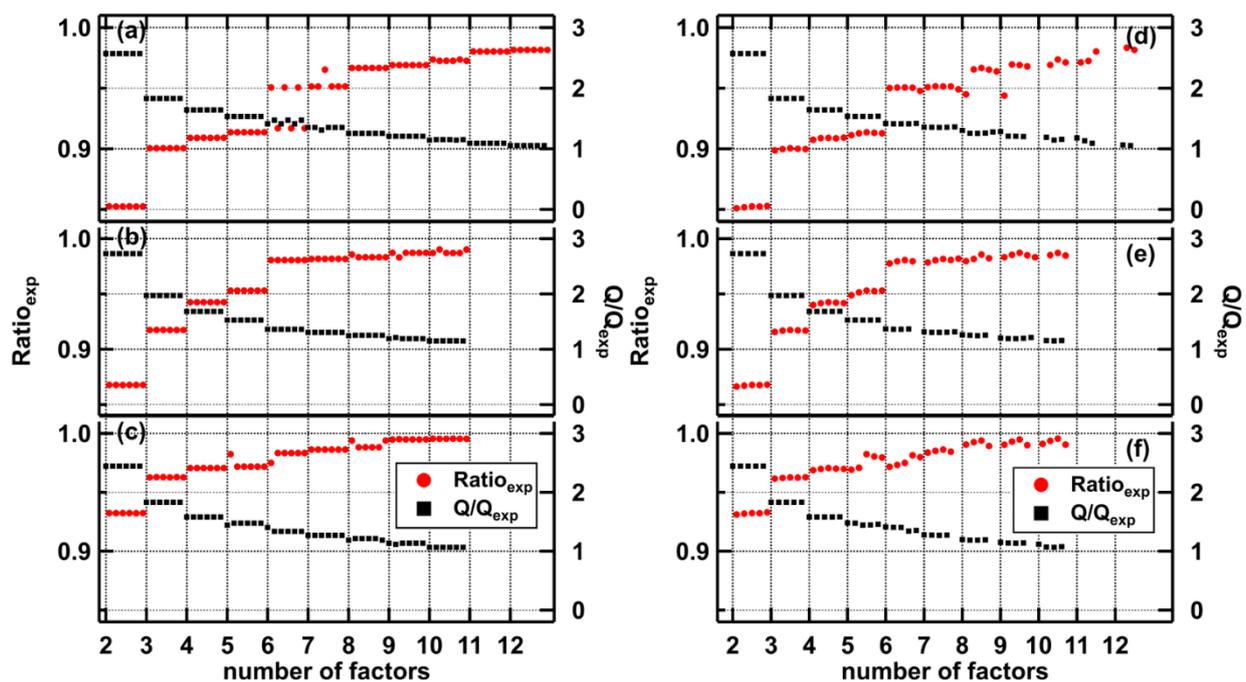


Figure S 2: Fraction of explained variance ($Ratio_{exp}$, red, left axis) and Q/Q_{exp} values (black, right axis) for all PMF solutions with seed values varying from 1 to 6 (a, b, c) and fpeak values varying from -1.0 to +1.0 (d, e, f) for all SOA types (low- (a&d), medium-(b&e), and highOC (c&f)).

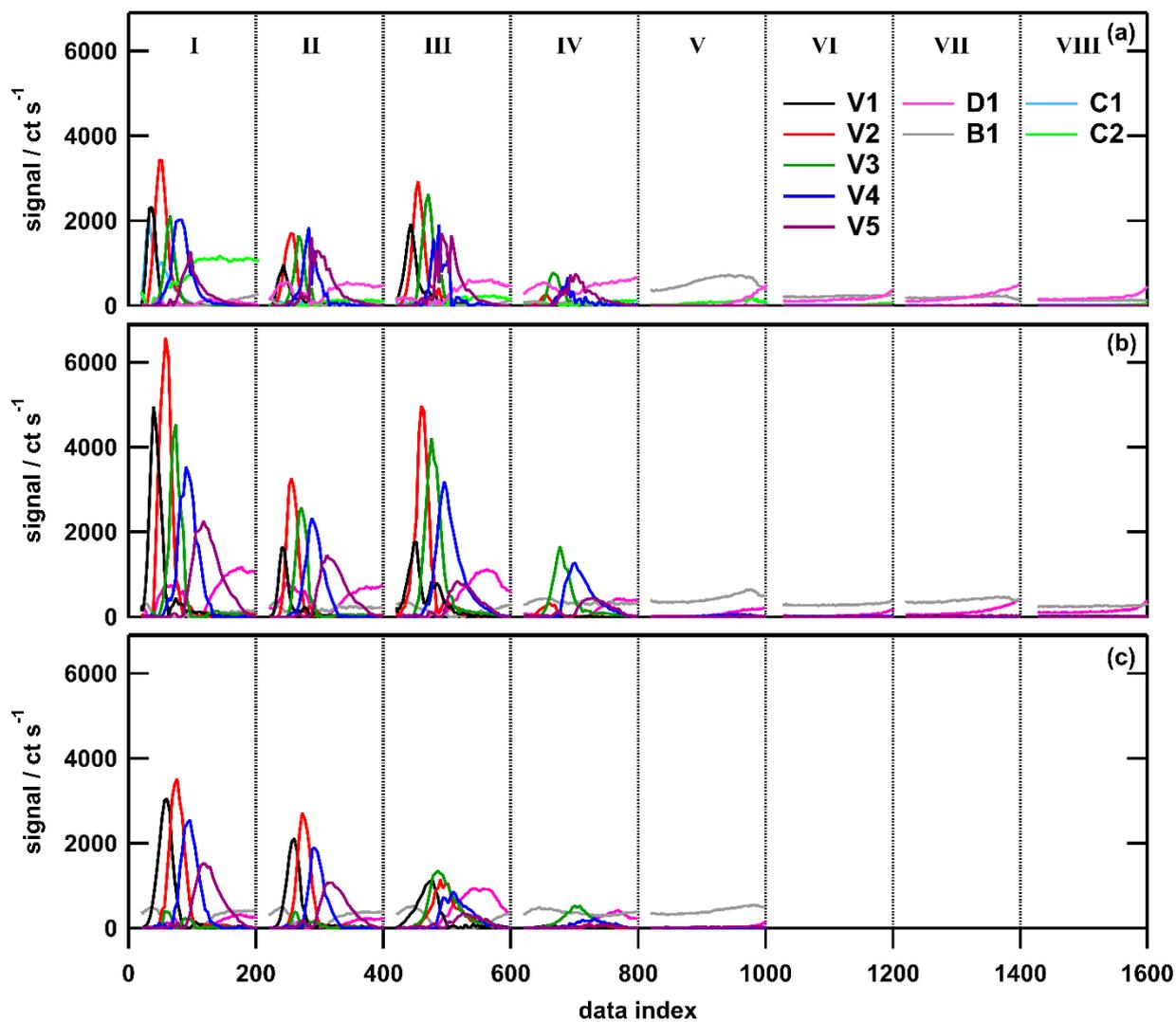


Figure S 3: Full dataset of factor profiles for samples (I - IV) and 1 or 4 filter blank measurements (V - VIII) for low- (a), medium- (b) and highOC (c) cases plotted vs the data index. Note that factors are different between SOA types (i.e., factor V1 in panel a and b have different factor mass spectra, see Figure 5 and Figure 6 in main text). Colour code for the factors is the same as in main text figures.

5

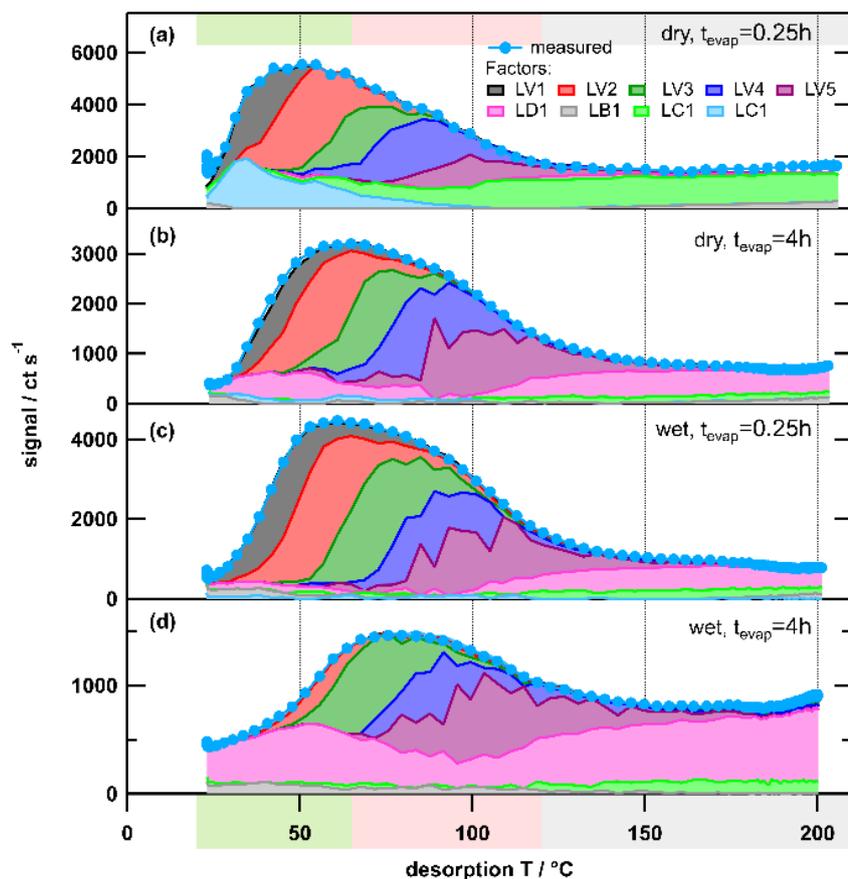


Figure S 4: Temperature profiles 9 factor solution for the lowOC case same as in main manuscript Figure 5 but stacked to highlight reconstruction of measured total thermogram signal. The colour code is the same for both panels. Background colours in the left panel indicate volatility classifications according to Donahue et al. (2006) derived from $T_{\text{max}}\text{-C}^*$ calibrations (green: SVOC, red: LVOC, grey: ELVOC).

5

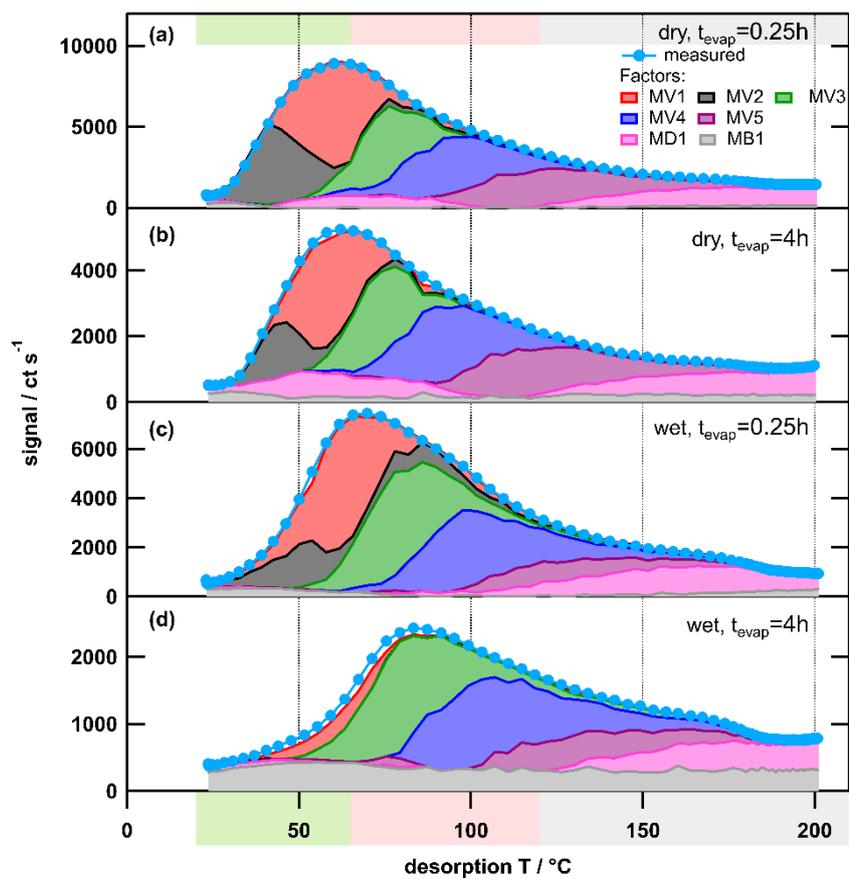


Figure S 5: Temperature profiles 7 factor solution for the mediumOC case same as in main manuscript Figure 6 but stacked to highlight reconstruction of measured total thermogram signal. The colour code is the same for both panels. Background colours in the left panel indicate volatility classifications according to Donahue et al. (2006) derived from T_{\max} - C^* calibrations (green: SVOC, red: LVOC, grey: ELVOC).

5

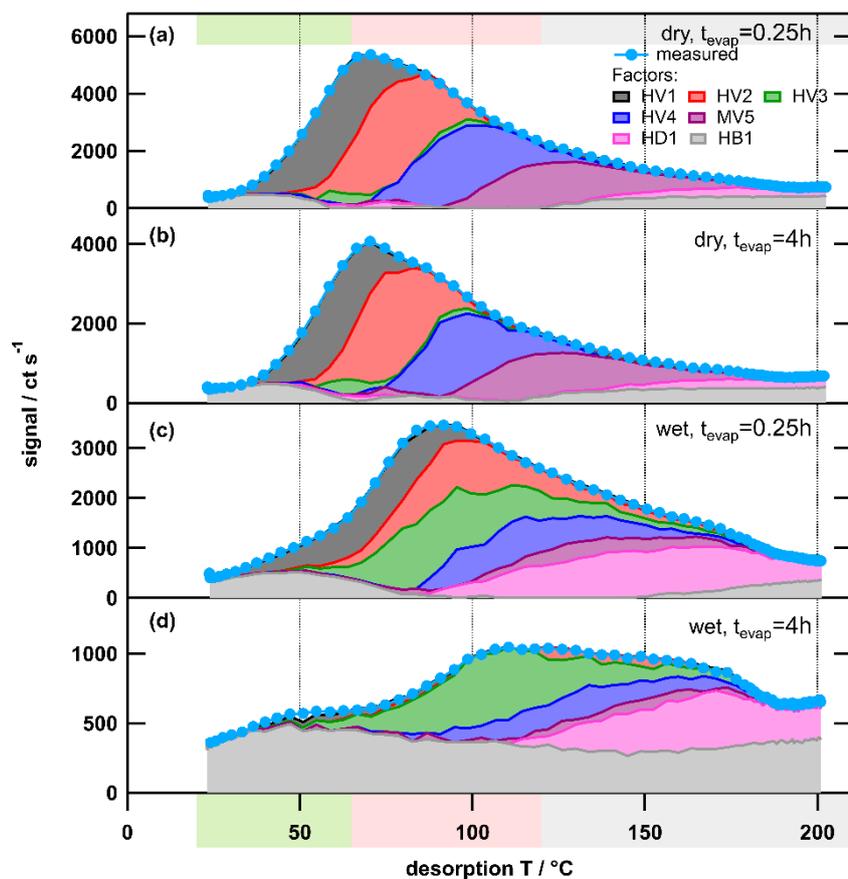


Figure S 6 Temperature profiles 7 factor solution for the highOC case same as in main manuscript Figure 7 but stacked to highlight reconstruction of measured total thermogram signal. The colour code is the same for both panels. Background colours in the left panel indicate volatility classifications according to Donahue et al. (2006) derived from $T_{\max}\text{-}C^*$ calibrations (green: SVOC, red: LVOC, grey: ELVOC).

5

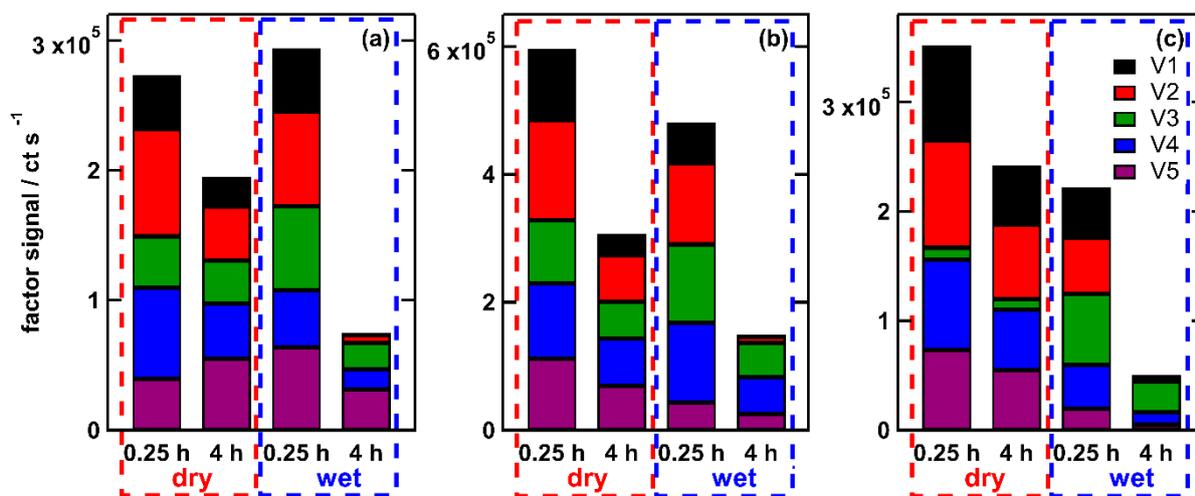
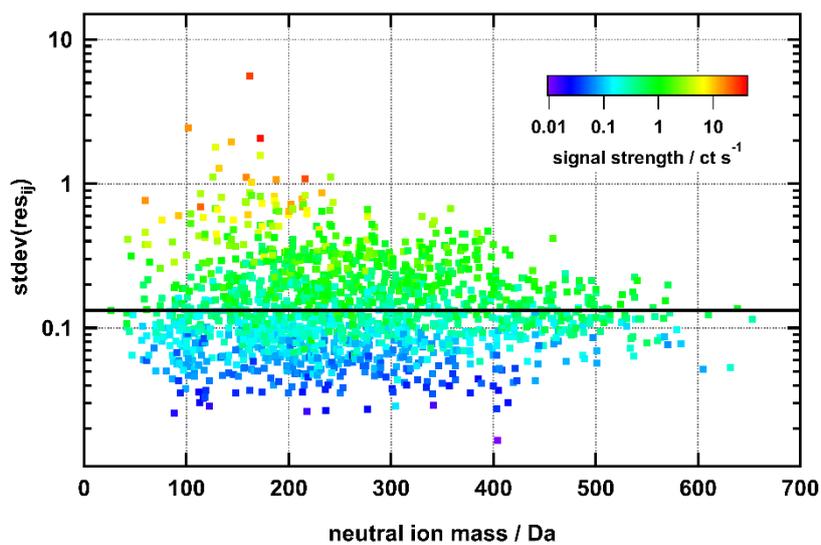


Figure S 7: Absolute contribution of V-Type factors to the measured signal for low- (a), medium- (b), and highOC cases (c). Factors are sorted by their T_{\max} values from V1 to V5. Colour code of the factors is the same as in main text.



5 **Figure S 8:** Standard deviation of residual for one thermogram (one point for each ion). Colour code is the average signal strength in the last 400 sec. Horizontal line indicates the median of the values which is defined as σ_{noise} .

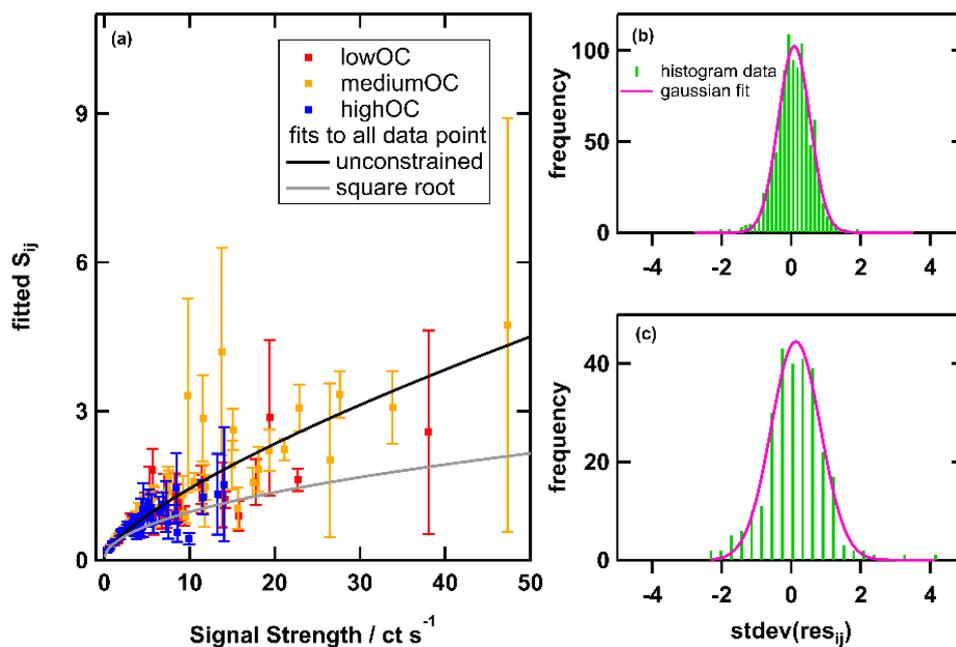


Figure S 9: (a) S_{ij} from gaussian fits to histograms vs average signal strength for all thermograms for all OC cases. Symbol colour stands for low-, medium-, and highOC. (b) and (c) histograms and gaussian fit for 2 of the 12 classes defined for calculation of S_{ij} . The fitted S_{ij} values in panel (a) are the halfwidth of these gaussian fits and the error bars are the 95th percentile ranges.