Table S1. Conditions during oxidation flow reactor experiments.

Precursor	254 nm lamp	Carrier gas	Source temperature	Mean mass concentration ± standard deviation	
				Blank	Sample
	V	$\mathrm{ml\,min^{-1}}$	°C	$\mu\mathrm{g\,m^{-3}}$	
α-Pinene	2	37.5	26	0.34 ± 0.09	83.9 ± 3.8
α-Pinene	-	37.5	26	0.34 ± 0.09	42.5 ± 1.6
β-Pinene	2	16.6	35	0.27 ± 0.08	184.4 ± 10.4
β-Pinene	-	16.6	39	0.27 ± 0.08	61.5 ± 8.9
Limonene	2	93.6	27	0.06 ± 0.07	104.3 ± 12.6
Limonene	-	93.6	27	0.06 ± 0.07	55.1 ± 3.2
3-Carene	2	16.6	28	1.3 ± 0.4	62.5 ± 4.8
3-Carene	-	12.9	29	1.3 ± 0.4	90.1 ± 10.6
trans-Caryophyllene	2	37.5	32	0.09 ± 0.05	52.5 ± 6.7
trans-Caryophyllene	-	71.6	32	0.09 ± 0.05	47.3 ± 4.8
Toluene	2	16.6	23	0.08 ± 0.03	66.2 ± 1.8
o-Xylene	2	25.4	22	0.42 ± 0.14	66.0 ± 2.2
1,2,4-Trimethylbenzene	2	37.5	40	0.21 ± 0.08	24.2 ± 1.2
Naphthalene	2	93.6	25	2.9 ± 0.7	35.9 ± 5.7

Table S2. Composition of the standard solution with substance name, CAS number, formula, purity, retention time, concentration in the solution, and appearance in (−) or (+) HESI ionization mode.

Substance	CAS number	Formula	Purity %	Retention time min	Concentration $\mu g\,mL^{-1}$	HESI mode
Benzoic acid α-^{13}C	3880-99-7	$C_6H_5{}^{13}CO_2H$	99 atom	6.86	10	(−)
Phthalic acid	88-99-3	$C_8H_6O_4$	99.5	4.55	1	(−), (+)
Acridine	260-94-6	$C_{13}H_9N$	98	5.07	0.5	(+)
Pinonic acid	473-72-3	$C_{10}H_{16}O_3$	N/A	6.75	2	(−), (+)
Pinic acid	473-73-4	$C_9H_{14}O_4$	N/A	5.82	2	(−), (+)
Caffeine-$^{13}C_3$	78072-66-9	$^{13}C_3C_5H_{10}N_4O_2$	99 atom	4.79	1	(+)
5-Acenaphthene carboxylic acid	55720-22-4	$C_{13}H_{10}O_2$	97	10.53	10	(−), (+)
MBTCA[1]	77370-41-3	$C_8H_{12}O_6$	N/A	4.09	2	(−), (+)
2,4-Di-*tert*-butylphenole	96-76-4	$C_{14}H_{22}O$	98	12.88	20	(−)
Camphorsulfonic acid	35963-20-3	$C_{10}H_{16}O_4S$	98	5.16	2	(−), (+)
Tri-*p*-cresyl phosphate	78-32-0	$C_{21}H_{21}O_4P$	N/A	13.19	0.5	(+)
Tris(2-ethylhexyl) phosphate	78-42-2	$C_{24}H_{51}O_4P$	97	15.53	0.5	(+)
Pentaerythritol tetrahexaonate	7445-47-8	$C_{29}H_{52}O_8$	95	14.93	1	(+)

[1] 3-methyl-1,2,3-butanetricarboxylic acid

Table S3. Workflow of the non-target software Compound Discoverer 3.2. The *Search mzVault* node was only used for the representative selection of the field campaign extracts. If the node is not used, *Data Source #2* in the *Assign Compound Annotations* node has to be set to *not specified*.

Processing node	Parameter	Settings
Select Spectra		
	1. Spectrum Properties Filter:	
		Lower RT Limit: 0
		Upper RT Limit: 0
		First Scan: 0
		Last Scan: 0
		Ignore Specified Scans: (not specified)
		Lowest Charge State: 0
		Highest Charge State: 0
		Min. Precursor Mass: 50 Da
		Max. Precursor Mass: 5000 Da
		Total Intensity Threshold: 0
		Minimum Peak Count: 1
	2. Scan Event Filters:	
		Mass Analyzer: Is FTMS
		MS Order: Any
		Activation Type: Is HCD
		Min. Collision Energy: 0
		Max. Collision Energy: 1000
		Scan Type: Any
		Polarity Mode: Is -
	3. Peak Filters:	
		S/N Threshold (FT-only): 5
	4. Replacements for Unrecognized Properties:	
		Unrecognized Charge Replacements: 1
		Unrecognized Mass Analyzer Replacements: FTMS
		Unrecognized MS Order Replacements: MS2
		Unrecognized Activation Type Replacements: HCD
		Unrecognized Polarity Replacements: -
		Unrecognized MS Resolution@200 Replacements: 60000
		Unrecognized MSn Resolution@200 Replacements: 30000
	5. General Settings:	
		Precursor Selection: Use MS1 Precursor
		Use Isotope Pattern in Precursor Reevaluation: True
		Provide Profile Spectra: Automatic
		Store Chromatograms: False
Align Retention Times		
	1. General Settings:	
		Alignment Model: Linear
		Alignment Fallback: Use Linear Model
		Maximum Shift [min]: 0.1

continued on next page

Processing node	Parameter	Settings
		Shift Reference File: True
		Mass Tolerance: 1 ppm
		Remove Outlier: True
Detect Compounds		
	1. General Settings:	
		Mass Tolerance [ppm]: 1 ppm
		Intensity Tolerance [%]: 30
		S/N Threshold: 5
		Min. Peak Intensity: 500000
		Ions: [2M-H]-1; [M-CO_2-H]-1; [M-H]-1; [M-H-H_2O]-1
		Base Ions: [M-H]-1
		Min. Element Counts: H
		Max. Element Counts: C_{90} H_{190} Br_4 Cl_4 N_4 O_{20} S_4
	2. Peak Detection:	
		Filter Peaks: True
		Max. Peak Width [min]: 0.1
		Remove Singlets: True
		Min. # Scans per Peak: 10
		Min. # Isotopes: 1
	3. Isotope Grouping:	
		Min. Spectral Distance Score: 0
		Remove Potentially False Positive Isotopes: True
Group Compounds		
	1. Compound Consolidation:	
		Mass Tolerance: 1 ppm
		RT Tolerance [min]: 0.2
	2. Fragment Data Selection:	
		Preferred Ions: [M-H]-1
Fill Gaps		
	1. General Settings:	
		Mass Tolerance: 1 ppm
		S/N Threshold: 5
		Use Real Peak Detection: True
Mark Background Compounds		
	1. General Settings:	
		Max. Sample/Blank: 5
		Max. Blank/Sample: 0
		Hide Background: True
Predict Compositions		
	1. Prediction Settings:	
		Mass Tolerance: 1 ppm
		Min. Element Counts: H
		Max. Element Counts: C_{90} H_{190} Br_4 Cl_4 N_4 O_{20} S_4
		Min. RDBE: 0
		Max. RDBE: 40

continued on next page

Processing node	Parameter	Settings
		Min. H/C: 0.1
		Max. H/C: 3.5
		Max. # Candidates: 10
		Max. # Internal Candidates: 200
	2. Pattern Matching:	
		Intensity Tolerance [%]: 30
		Intensity Threshold [%]: 0.1
		S/N Threshold: 5
		Min. Spectral Fit [%]: 30
		Min. Pattern Cov. [%]: 90
		Use Dynamic Recalibration: True
	3. Fragments Matching:	
		Use Fragments Matching: True
		Mass Tolerance: 1 ppm
		S/N Threshold: 5
Search mzVault		
	1. Search Settings:	
		mzVault Library: *add libraries created with mzVault*
		Max. # Results: 10
		Match Factor Threshold: 50
		Search Algorithm: HighChem HighRes
		Match Analyzer Type: True
		IT Fragment Mass Tolerance: 0.4 Da
		FT Fragment Mass Tolerance: 10 ppm
		Use Retention Time: True
		Precursor Mass Tolerance: 10 ppm
		Apply Intensity Threshold: True
		Match Ionization Method: True
		Ion Activation Energy Tolerance: 20
		Match Ion Activation Energy: Match with Tolerance
		Match Ion Activation Type: True
		Compound Classes: All
		Remove Precursor Ion: True
		RT Tolerance [min]: 0.2
Assign Compound Annotations		
	1. General Settings:	
		Mass Tolerance: 1 ppm
	2. Data Sources:	
		Data Source #1: Predicted Compositions
		Data Source #2: mzVault Search
	3. Scoring Rules:	
		Use mzLogic: True
		Use Spectral Distance: True
		SFit Threshold: 20
		SFit Range: 20
Calculate Mass Defect		

continued on next page

Processing node	Parameter	Settings
	1. Mass Defect:	
		Fractional Mass: False Standard Mass Defect: False Relative Mass Defect: False Kendrick Mass Defect: True Nominal Mass Rounding: Round
	2. Kendrick Formula:	
		Formula 1: C H_2
Merge Features		
	1. Peak Consolidation:	
		Mass Tolerance: 1 ppm RT Tolerance [min]: 0.1
Find Expected Compounds		
	1. General Settings:	
		Mass Tolerance: 1 ppm Intensity Tolerance [%]: 30 Intensity Threshold [%]: 0.1 S/N Threshold: 5 Min. # Isotopes: 2 Min. Peak Intensity: 10000 Average Peak Width [min]: 0
Group Expected Compounds		
	1. Compound Consolidation:	
		RT Tolerance [min]: 0.1
	2. Fragment Data Selection:	
		Preferred Ions: [M-H]-1
Generate Expected Compounds		
	1. Compound Selection:	
		Compounds: Benzoic acid ^{13}C ($C_6H_5{}^{13}CO_2H$)
	2. Dealkylation:	
		Apply Dealkylation: False Apply Dearylation: False Max. # Steps: 1 Min. Mass [Da]: 200
	3. Transformations:	
		Phase I: (not specified) Phase II: (not specified) Others: (not specified) Max. # Phase II: 1 Max. # All Steps: 3
	4. Ionization:	
		Ions: [M-CO_2-H]-1; [M-H]-1; [M-H-H_2O]-1

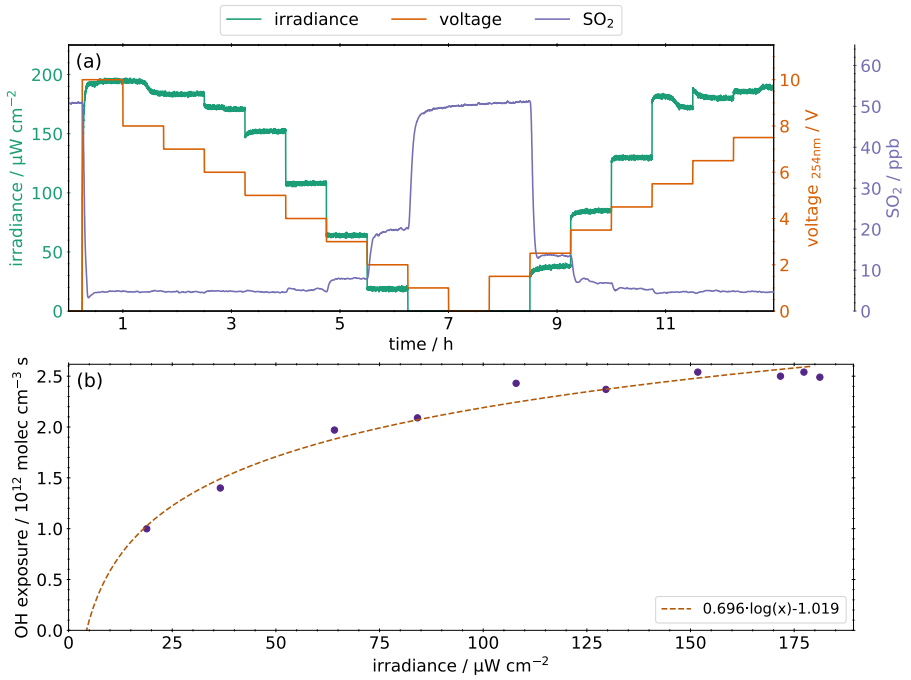

Figure S1. Results of the OH exposure approximation experiment: (a) the adjusted voltage of the 254 nm lamps as well as the measured irradiance and SO_2 concentration. (b) OH exposure, calculated with Equation (S1), vs. irradiance .

Lambe et al. (2011) define the OH exposure as

$$OH_{\mathrm{exp}} = -\frac{1}{k_{\mathrm{OH,SO_2}}} \cdot \ln\left(\frac{SO_{2,f}}{SO_{2,i}}\right), \tag{S1}$$

with the initial SO_2 concentration ($SO_{2,i}$), the final SO_2 concentration ($SO_{2,f}$) and the rate constant between SO_2 and OH ($k_{OH,SO_2} = 9 \times 10^{-13}$ cm^3 molecule^{-1}).

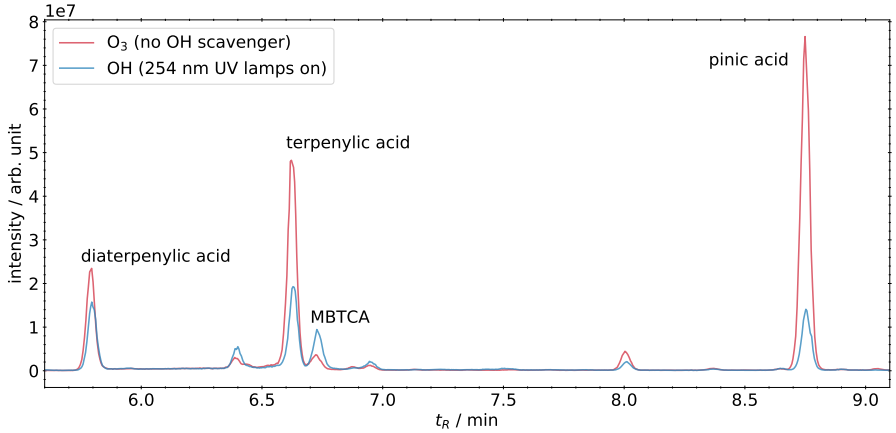

Figure S2. Chromatograms of diaterpenylic acid, terpenylic acid, MBTCA, and pinic acid from α-pinene oxidation experiments. Diaterpenylic-, terpenylic-, and pinic acid show lower absolute intensities under OH conditions. The absolute signal intensity of MBTCA increases under OH conditions.

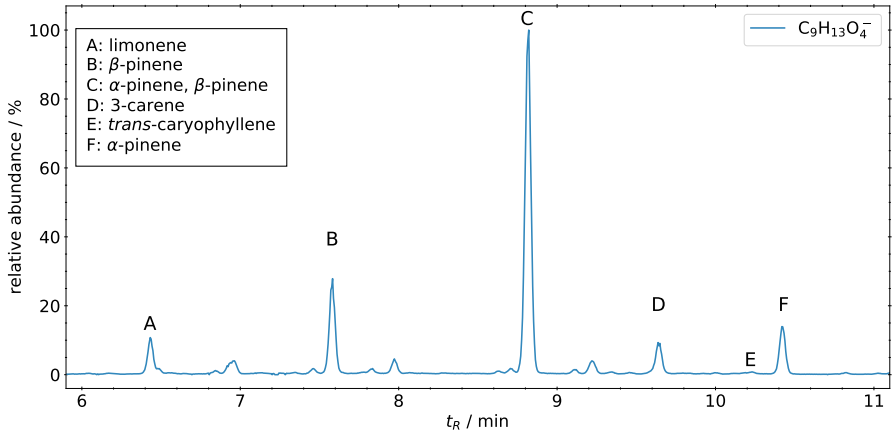

Figure S3. Ambient air chromatogram of *m/z* 185.0819 and the assigned isomers based on the PAM-OFR experiments. With exception of the isomer C appearing at 8.79 minutes, the other five isomers A, B, D, E, and F can be used as specific tracer isomers for the respective precursor.

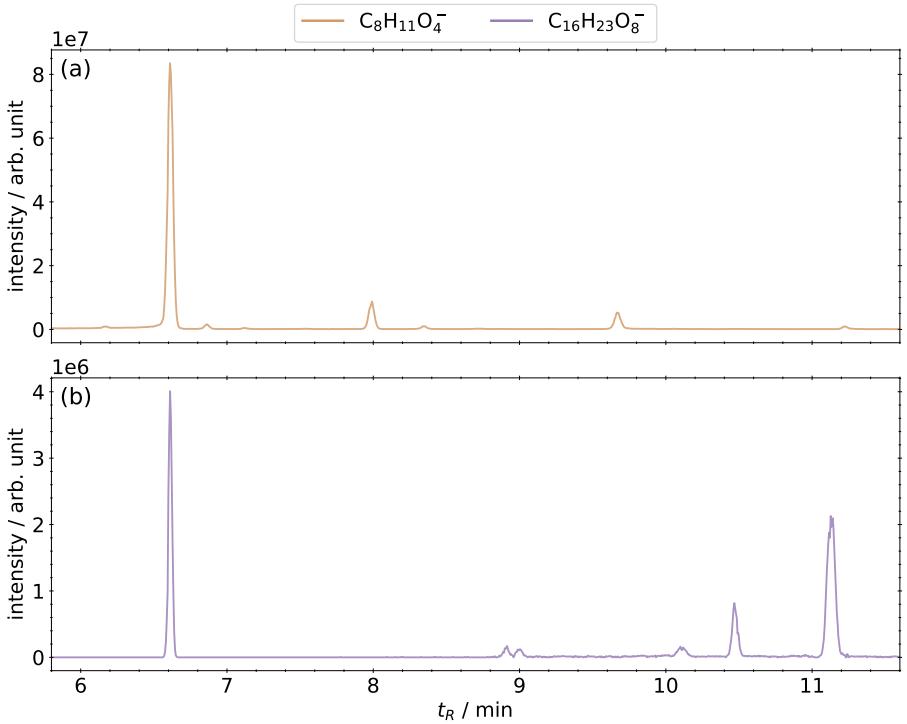

Figure S4. (a) Mass traces of *m/z* 171.0662 and (b) *m/z* 343.1397. The high concentrated monomer at 6.67 minutes can generate ionsource dimers. Real atmospheric generated dimers appear at higher retention times above 10 minutes.

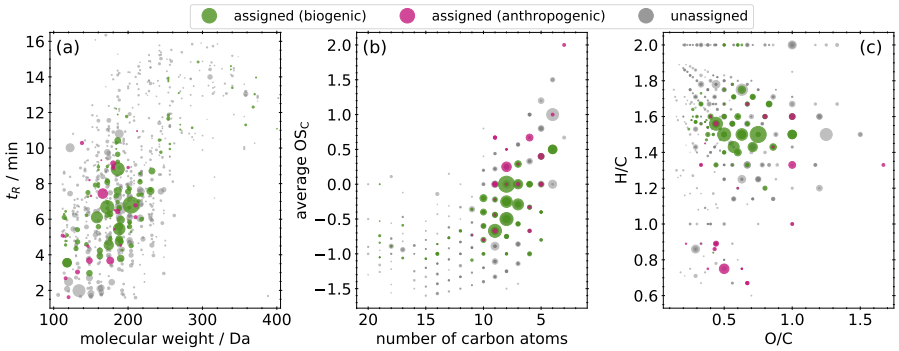

Figure S5. Assigned and unassigned CHO compounds from the representative selection of the field campaign samples. (a) Retention time vs. molecular weight. (b) Kroll diagram. (c) Van Krevelen diagram.

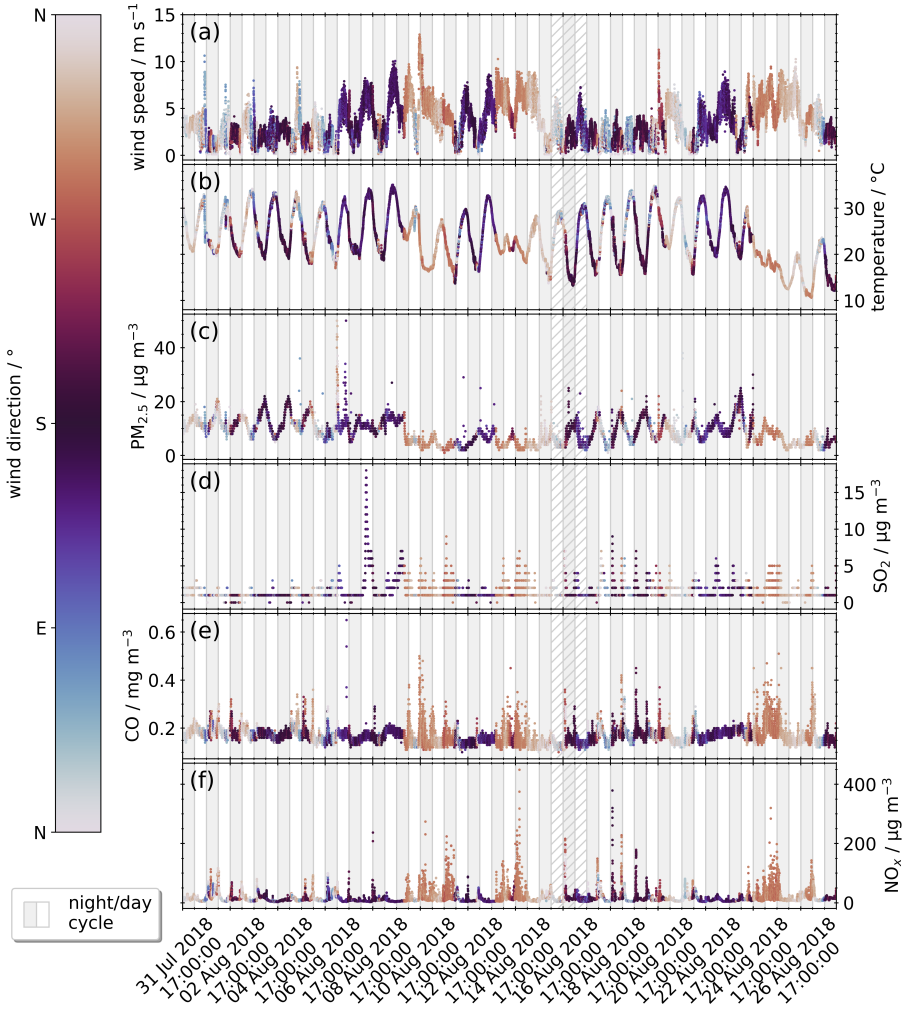

Figure S6. Meteorological data (a) wind speed and (b) temperature, (c) the $PM_{2.5}$ concentration as well as the trace gas concentrations of (d) SO_2, (e) CO, and (f) NO_x during the field campaign in August 2018 near Vienna (time is UTC). The color code indicates the wind direction, the background color the night and day cycle of the sampling interval. Between 16 August 5:00 and 17 August 17:00 (hatched background) no filter were collected.

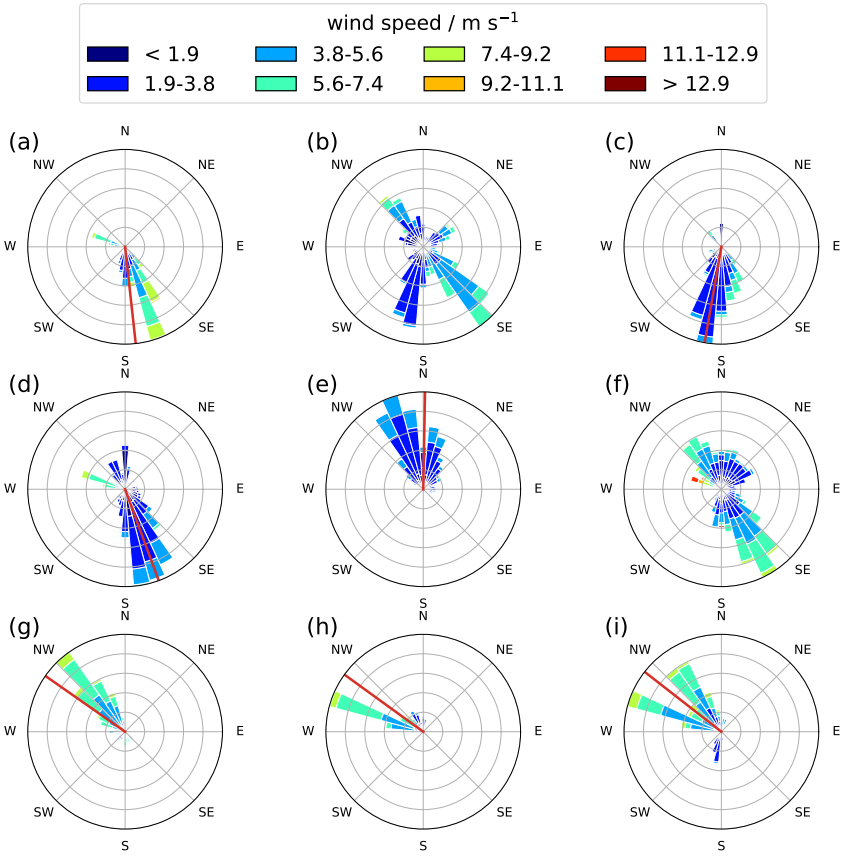

Figure S7. Wind direction and speed distribution for sample subsclusters (a) to (i). The mean wind directions are calculated with Equation (S2). Subclusters (b) and (f) do not show a predominant wind direction for which reason, no line is drawn.

From a list of wind directions (θ_i), given in radians, Yamartino (1984) defined the mean wind direction as

$$\theta_a = arctan\left(\frac{s_a}{c_a}\right) \tag{S2}$$

with

$$s_a = n^{-1}\sum_i sin\theta_i \tag{S3}$$

and

$$c_a = n^{-1}\sum_i cos\theta_i. \tag{S4}$$

In (−)HESI nearly all fragments and adducts can be explained by the four ions presented in Sect. 2.6. With the NTA those ions are identified by the software and assigned to the respective molecule. In (+)HESI this is by far not so simple. Ions can be detected as known adducts like $[M - H_2O + H]^+$, $[M + H]^+$, $[M + NH_4]^+$, $[M + Na]^+$, and $[M + K]^+$ but can also form several unknown fragments and adducts. The NTA of the standard solution measurement, in which 11 compounds can be detected in (+)HESI (Table S2), results in 53 identified compounds shown in Fig. S8. The color of the scatters indicates whether the compound was spiked into the solution (red), fragmented or formed adducts during measurement (orange), could not be identified (yellow), or is a pseudo real fragment of a spiked substance (light blue). The horizontal lines should simplify the assignment of fragments and adducts of same retention time.

Spiked compounds show mostly more then two adducts with the exception of nitrogen containing compounds caffeine (at 4.8 min) and acridine (at 5.1 min) with only one and 5-acenaphthene carboxylic acid (at 10.5 min) with two adducts. In contrast fragments and adducts formed during measurement always have a maximum number of two adducts. Consider only compounds with more then two adducts detected from the software can reduce false identification noticeably. The loss of real compounds, however, cannot be ruled out.

The signals between *m/z* 413 and 469, at a retention time of 13.7 minutes, probably refer to a fragment of pentaerythritol tetrahexaonate (528 Da at 14.9 min) formed before sampling due to the fact that five adducts were detected from the software and the retention time is lower then from pentaerythritol tetrahexaonate. The resulting chemical formula corresponds to the loss of a $C_6H_{10}O$ side chain. The five unidentified signals between *m/z* 232 and 264 at a retention time of 6.8 minutes could be caused by interference with benzoic acid, measured at this time in (−)HESI. The two unidentified signals at 12.4 minutes with *m/z* 267 and 289 could may be the hydrogen and sodium adduct of a fragment of tris(2-ethylhexyl) phosphate formed in solution. Due to the high uncertainty it is labeled as unidentified. In principal the chemical formula $C_{12}H_{27}O_4P$ could match with the losses of an entire side chain (C_8H_{16}) and a fragment (C_4H_8) of another side chain.

Populate a database in (+)HESI is a further challenge. While in (−)HESI the ion with the highest signal intensity is mostly $[M - H]^-$, in (+)HESI no preferred ion can be named. Consequently from all or most measured ions a MS^2 spectra should be recorded and added to each library entry by hand. The necessity of this work can be illustrated with the MS^2 spectra of the ions $[M + H]^+$ and $[M + Na]^+$ from tri-*p*-cresyl phosphate. Despite the same molecule the spectra from both ions differ completely. The MS^2 spectra of the sodium adduct consists mainly of one signal. In contrast the hydrogen adduct produces a plausible variety of different fragments regarding the structure of tri-*p*-cresyl phosphate (Fig. S9). All of these difficulties highlight the need for further work on the (+)HESI.

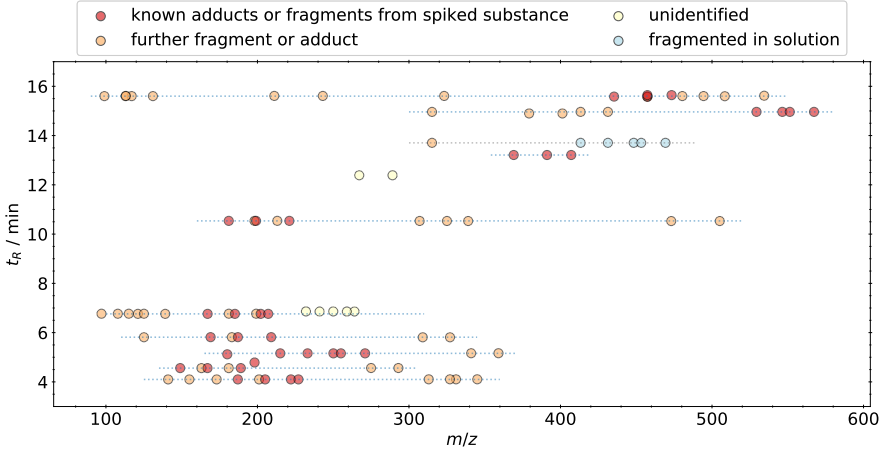

Figure S8. Retention time vs. *m/z* of all identified ions, which were not marked as background, from the measurement of the standard solution in (+)HESI. Known adducts include $[M - H_2O + H]^+$, $[M + H]^+$, $[M + NH_4]^+$, $[M + Na]^+$, and $[M + K]^+$.

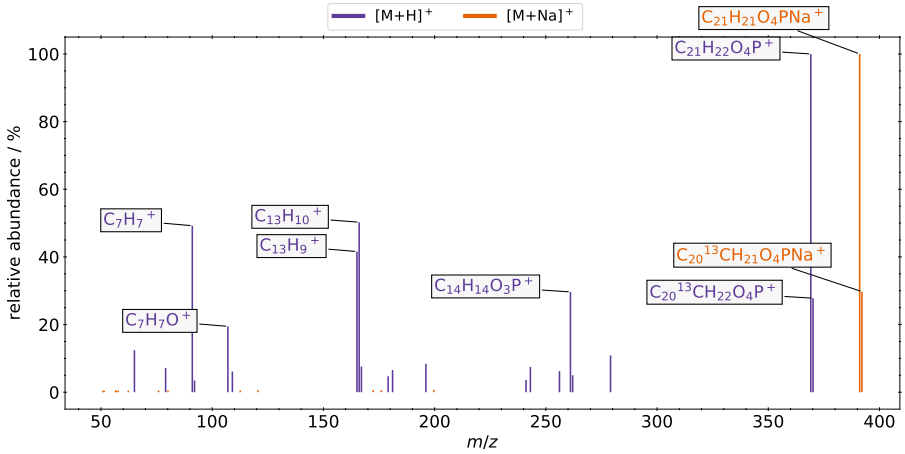

Figure S9. MS^2 spectra of the tri-*p*-cresyl phosphate addcuts $[M+H]^+$ and $[M+Na]^+$.

References

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, https://doi.org/10.5194/amt-4-445-2011, 2011.

Yamartino, R. J.: A Comparison of Several "Single-Pass" Estimators of the Standard Deviation of Wind Direction, J. App. Meteorol. Clim., 23, 1362–1366, https://doi.org/10.1175/1520-0450(1984)023<1362:ACOSPE>2.0.CO;2, 1984.