



Supplement of

Comprehensive non-targeted molecular characterization of organic aerosols in the Amazon rainforest

Denis Leppla et al.

Correspondence to: Thorsten Hoffmann (t.hoffmann@uni-mainz.de)

The copyright of individual parts of the supplement might differ from the article licence.

Supporting Information

Table S1: Values of n_c^0 and b according to the compound class to calculate C_0 . The parameterization is obtained from the detailed study by Li et al. (2016).

Compound class	n_c^0	b_c	b_o	b_{co}	b_N	b_S
CHO	22.66	0.4481	1.656	-0.7790		
CHON	24.13	0.3667	0.7732	-0.0779	1.114	
CHOS	24.06	0.3637	1.327	-0.3988		0.7579
CHONS	28.50	0.3848	1.011	0.2921	1.053	1.3160

Table S2: Average values¹ of the detected background ions for the wet and dry seasons in 2018 and 2019. The listed molecules were detected in at least 75% of all corresponding samples. The signals are divided into four subgroups regarding their elemental composition. The relative contribution of the subgroups was calculated by dividing the number of compounds of the particular subgroup by the total number of compounds. Additionally, the average values¹ of molecular weight (MW), carbon oxidation state (OSc) and aromaticity index (Xc) are listed.

Season	Height m	Number of compounds detected	CHO %	CHON %	CHONS %	CHOS %	MW Da	OSc ²	Xc
WS18	42	36	89	0	0	11	170	0.031	0.060
	150	45	82	5	2	11	177	-0.039	0.180
	320	28	82	0	0	18	174	0.040	0.024
DS18	80	211	90	4	1	5	191	-0.404	0.724
	150	215	90	4	1	4	188	-0.418	0.738
	320	209	89	4	1	5	188	-0.407	0.755
WS19	80	51	94	0	2	4	197	-0.497	0.696
	150	52	98	0	2	0	195	-0.464	0.577
	320	60	97	0	2	2	196	-0.499	0.606
DS19	0	72	94	4	1	0	196	-0.367	0.424
	80	108	95	4	0	1	182	-0.424	0.480
	320	75	97	3	0	0	191	-0.398	0.279

¹ Average values are calculated based on the molecular composition of each compound.

² Only CHO compounds are considered in the calculation of OSc.

Table S 3: Number of active fires each month in the Amazon region since 1998 until 2020. The sampling periods are highlighted in yellow. Data retrieved by INPE - Instituto Nacional de Pesquisas Espaciais, 2020 (Portal do Monitoramento de Queimadas e Incêndios Florestais, <http://www.inpe.br/queimadas>).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1998	-	-	-	-	-	1549	3192	20075	19214	8777	3833	2547
1999	160	358	130	70	449	1439	3675	21525	16106	12794	4449	1703
2000	87	182	405	92	930	3211	1510	12791	10062	10226	5497	3175
2001	165	699	1134	617	916	4227	1816	17679	15528	14292	8346	4256
2002	590	667	901	405	1490	5702	7529	43484	48549	27110	23660	9174
2003	3704	1573	1997	1038	1983	6848	15918	34765	47789	25341	19631	13813
2004	2178	805	1035	1012	3131	9179	19179	43320	71522	23928	26424	16924
2005	4314	1048	758	832	1746	2954	19364	63764	68560	26624	16790	6966
2006	1973	879	903	709	843	2522	6995	34208	51028	18309	17474	8579
2007	1918	1761	1431	760	1176	3519	6196	46385	73141	28731	16025	5437
2008	938	527	860	569	383	1248	5901	21445	26469	23518	15450	6145
2009	1095	354	584	435	673	1023	2327	9732	20527	19323	19104	6505
2010	1697	1147	1176	633	1026	1911	5868	45018	43933	14798	12167	5240
2011	771	271	427	465	528	1083	2445	8002	16987	9760	9815	7632
2012	1203	438	484	473	855	1875	3095	20687	24067	14814	13259	5469
2013	1181	374	738	518	796	1450	2531	9444	16786	10242	6615	8013
2014	1573	473	1010	632	673	1628	2766	20113	20522	13222	12169	7773
2015	2042	1047	572	762	407	1287	2817	20471	29326	19469	16935	11303
2016	4657	1559	2024	1075	895	1663	6120	18340	20460	14234	11610	5124
2017	796	379	736	618	805	1759	7986	21244	36569	14457	14105	7985
2018	1444	888	1359	513	772	1980	4788	10421	24803	10654	8881	1842
2019	1419	1368	3383	1702	854	1880	5318	30900	19925	7855	11297	3275
2020	1200	1196	1641	789	829	2248	6803	29307	32017	17326	6321	2498

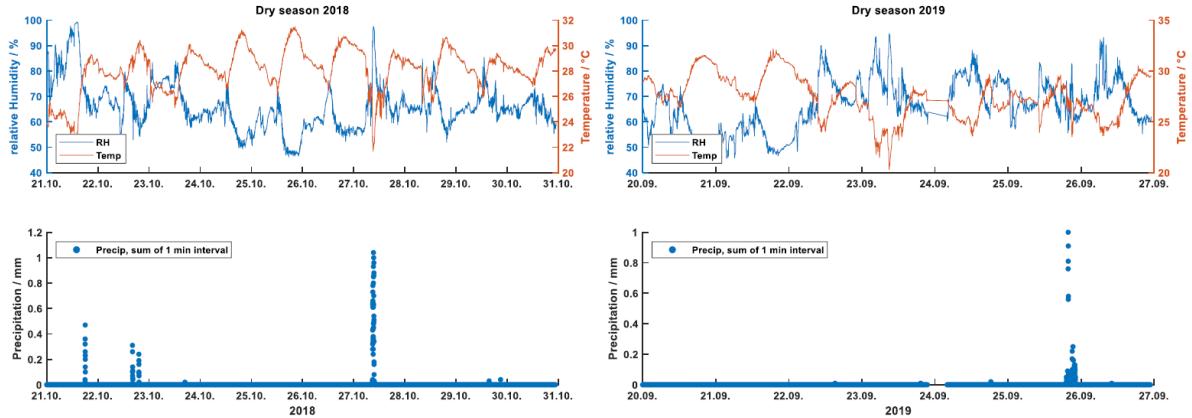


Figure S 1: The upper panel illustrates the relative humidity and temperature at ATTO for the dry season 2018 (left) and the dry season 2019 (right), respectively. The lower panel shows the precipitation as a sum of 1 min intervals. The instrument was located at 325 m.

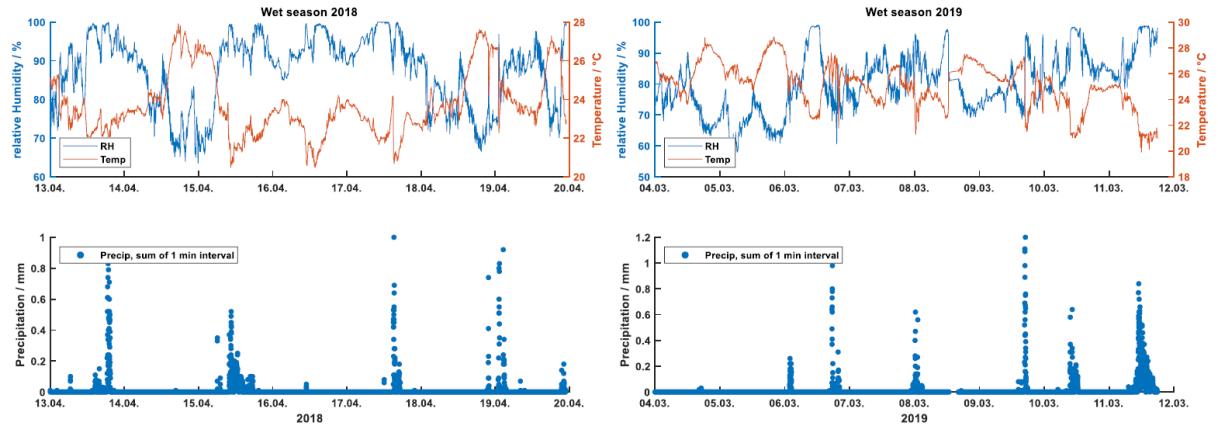


Figure S 2: The upper panel illustrates the relative humidity and temperature at ATTO for the wet season 2018 (left) and the wet season 2019 (right), respectively. The lower panel shows the precipitation as a sum of 1 min intervals. The instrument was located at 325 m.

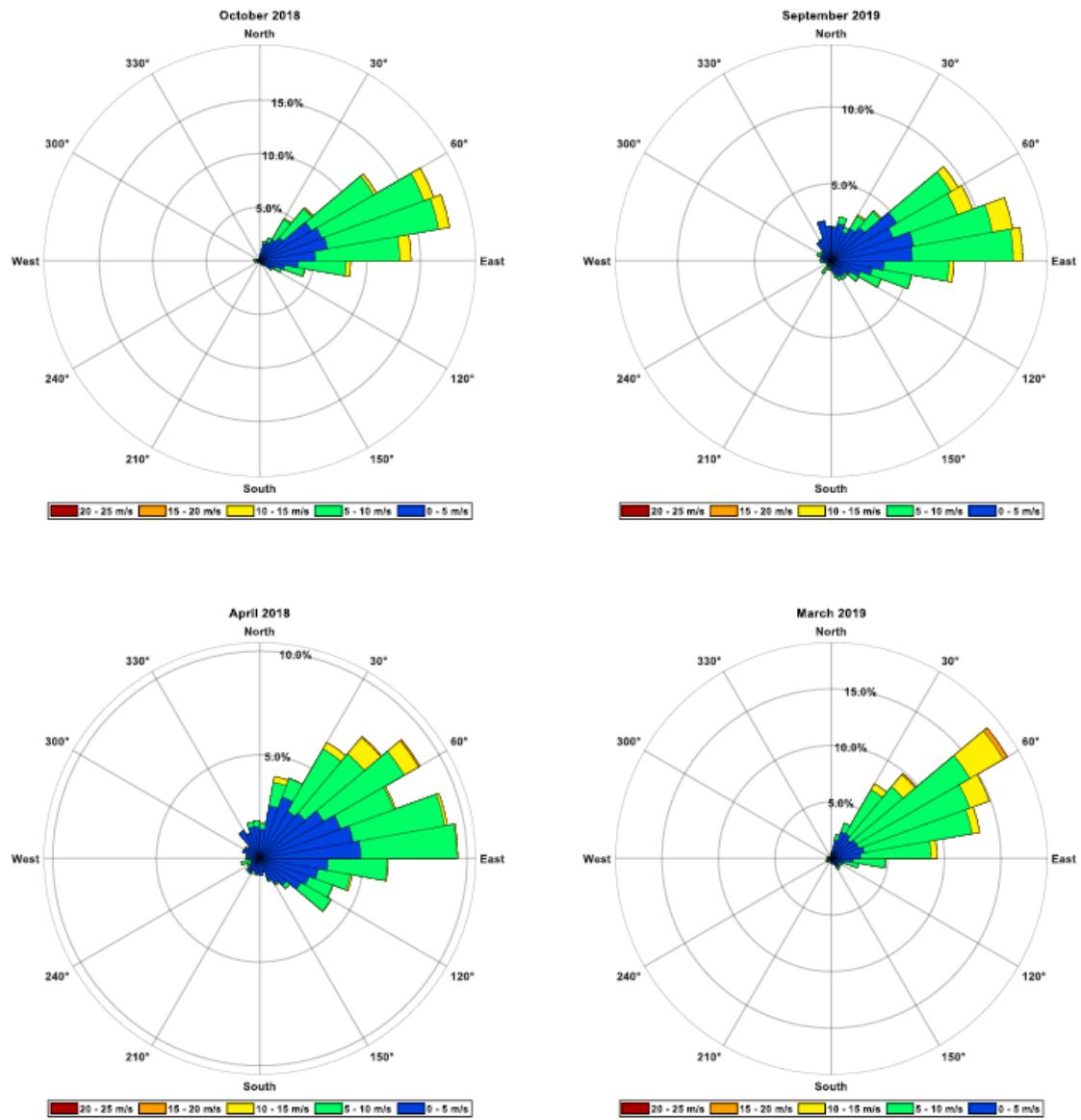


Figure S 3: Wind roses for the sampling period during the dry seasons in October 2018 and September 2019 (upper part) and during the wet seasons in April 2018 and March 2019 (lower part).

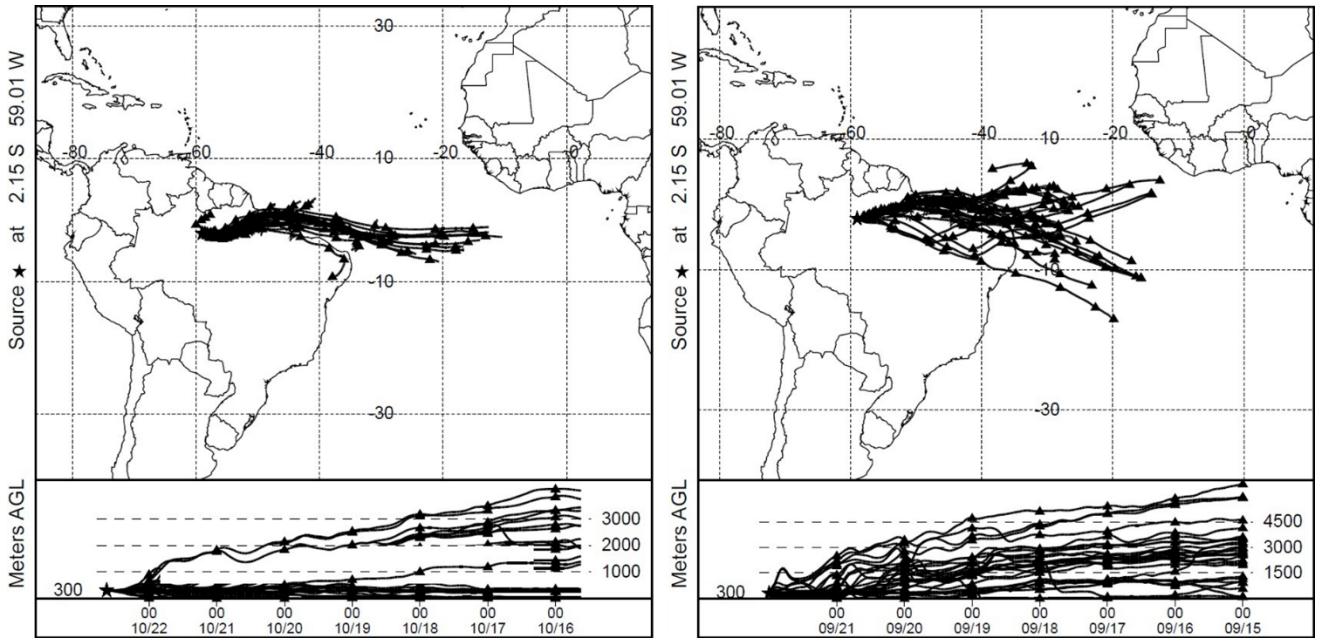


Figure S 4: The 7 d HYSPLIT backward trajectory ensembles starting at 300 m above ground level for the dry season 2018 (left) and dry season 2019 (right) (Stein et al., 2015).

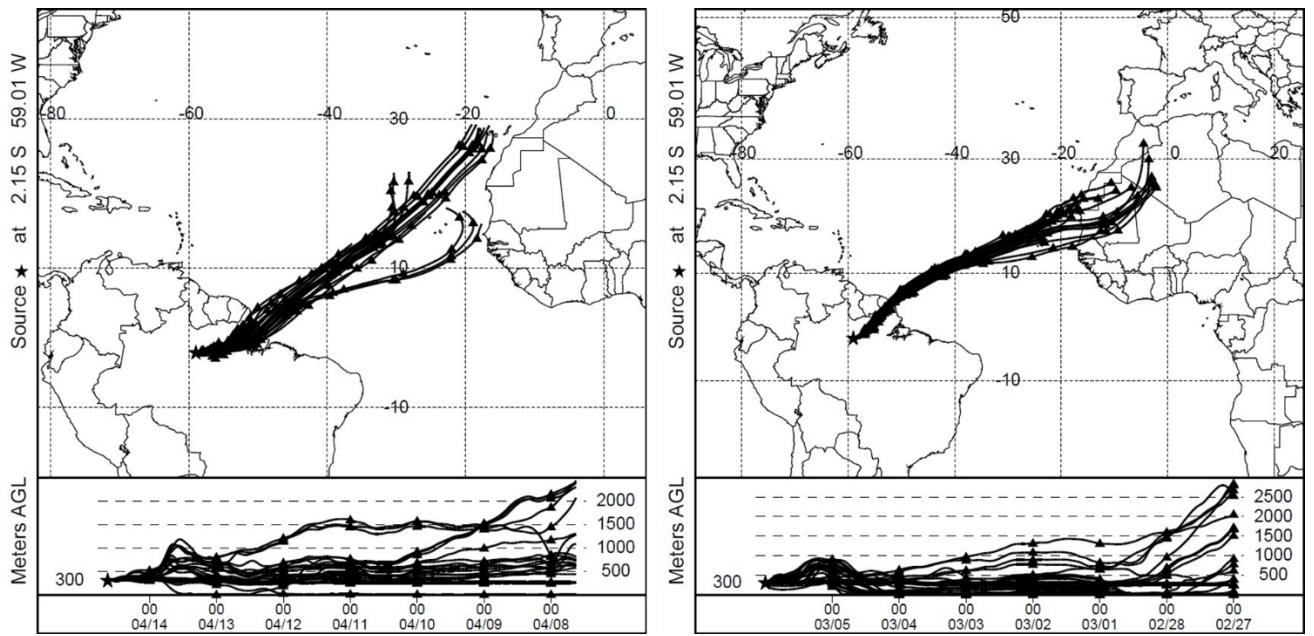


Figure S 5: The 7 d HYSPLIT backward trajectory ensembles starting at 300 m above ground level for the wet season 2018 (left) and the wet season 2019 (right) (Stein et al., 2015).

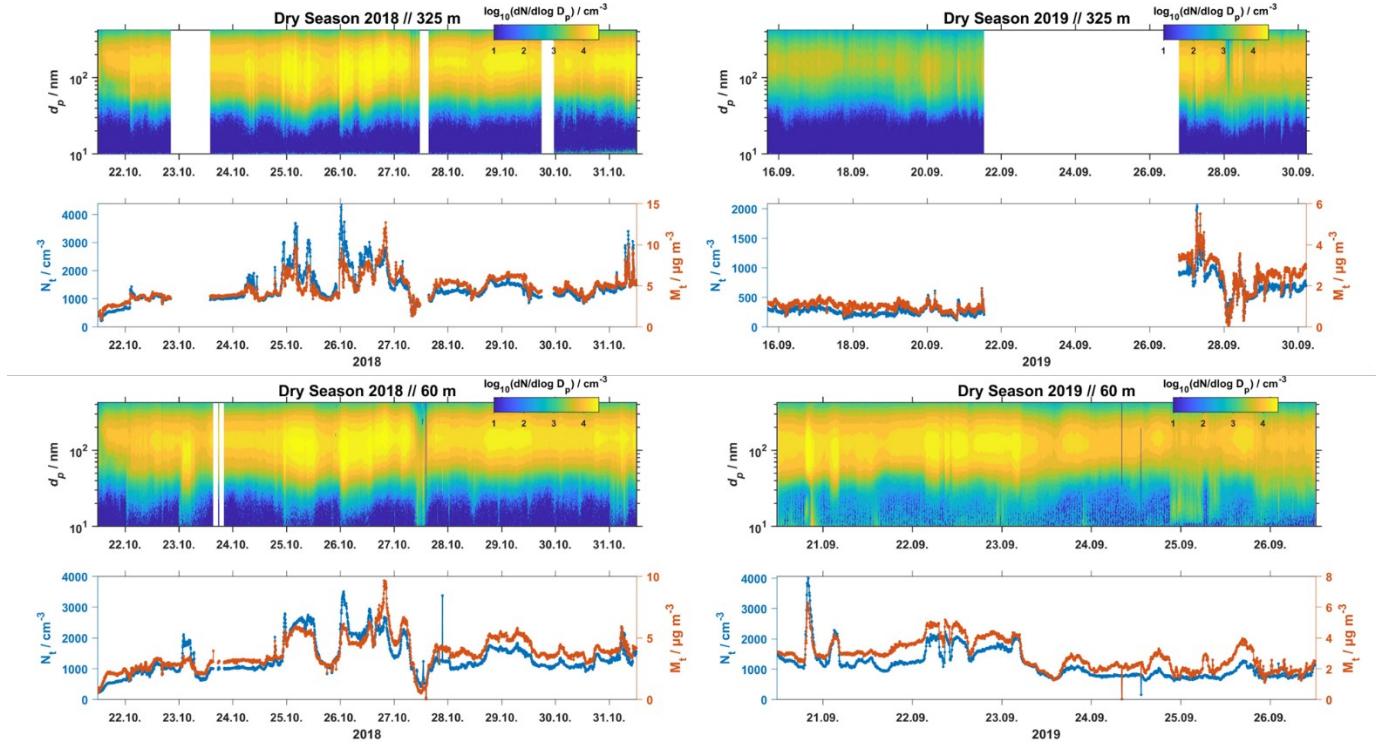


Figure S 6: SMPS data for the dry season 2018 and 2019 at 60 m and 325 m altitude. The calculated total particle numbers and total particle masses are illustrated in the lower panels, respectively. The instruments did not work during the 23., 27., and 30.10.2018 and during the 21.09.2019 and 27.09.2019. Thus, a longer period is displayed.

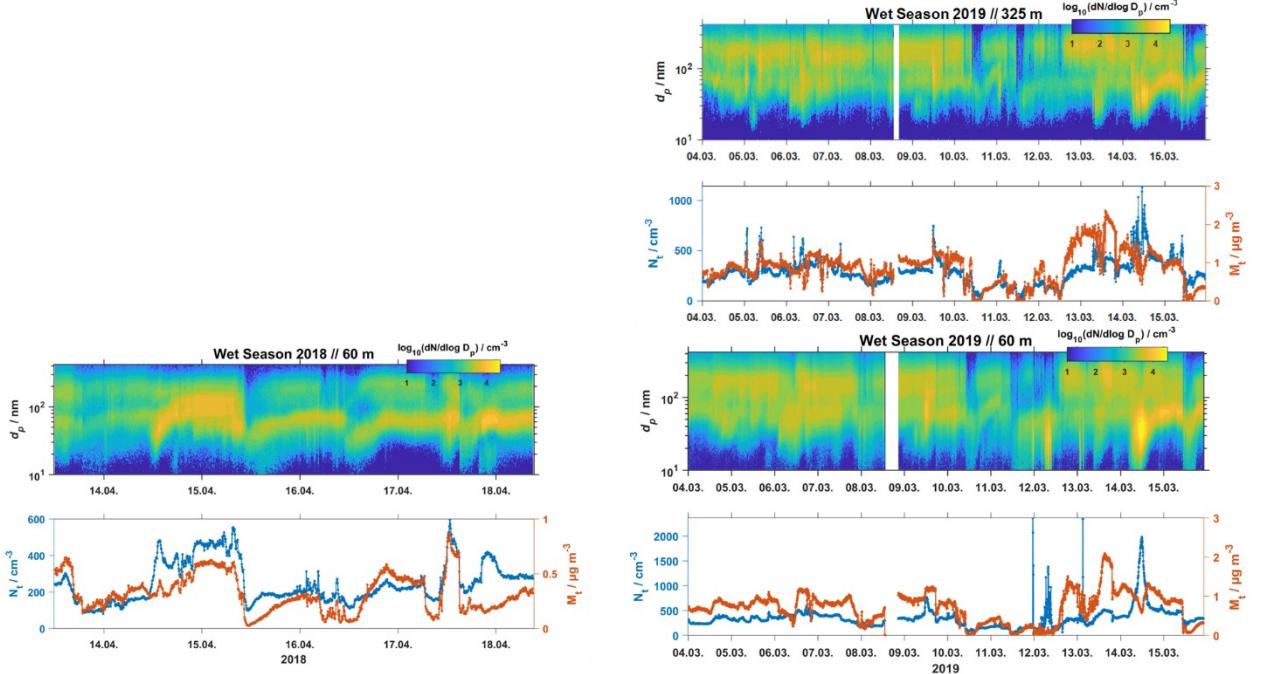


Figure S 7: SMPS data for the wet season 2018 and 2019 at 60 m and 325 m altitude. The calculated total particle numbers and total particle masses are illustrated in the lower panels, respectively. No data were collected at 325 m during the wet season 2018.

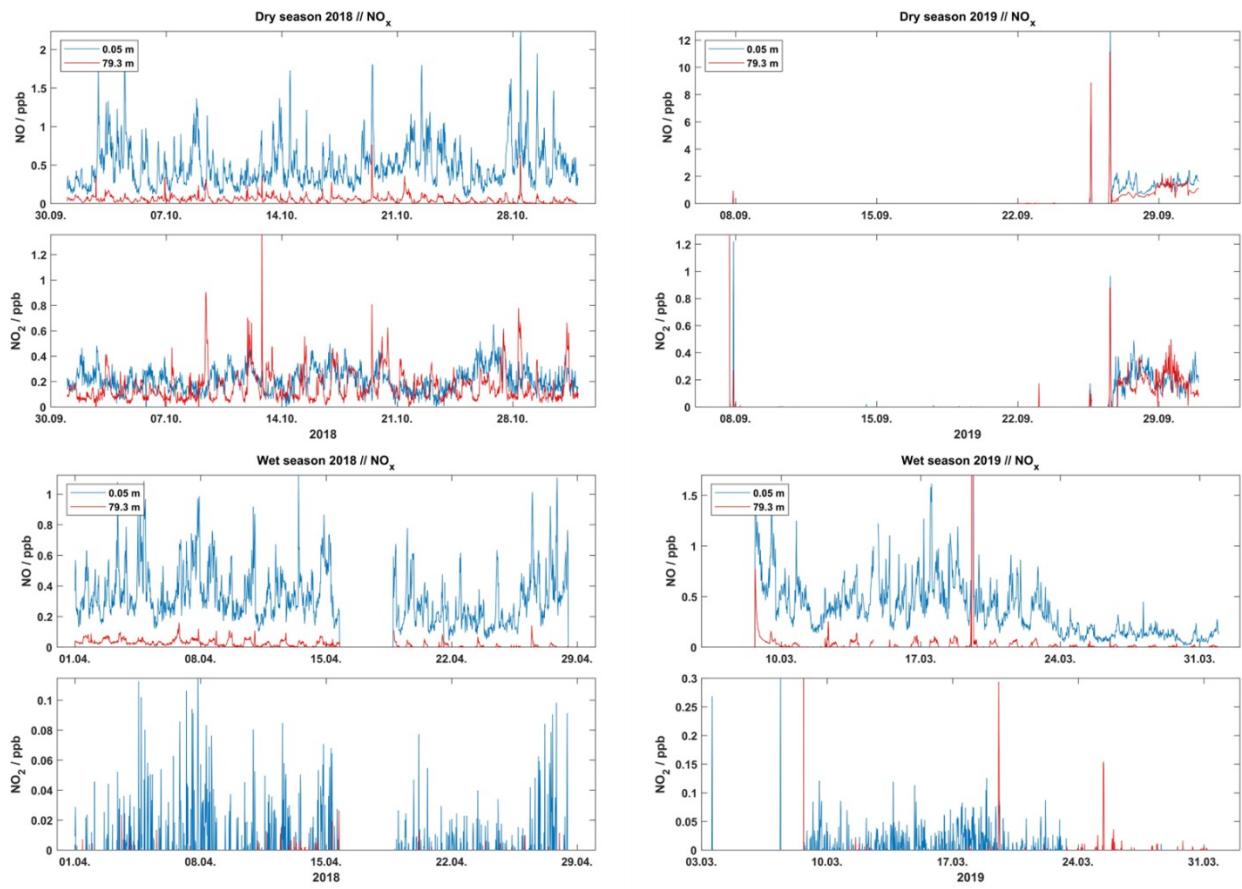


Figure S 8: NO and NO₂ concentrations at 0.05 m and 79.3 m height during the dry seasons 2018 and 2019 (upper part). The instrument was malfunctioning during the dry season 2019 resulting in few datapoints. The NO_x concentration for the wet seasons 2018 and 2019 are illustrated in the lower part.

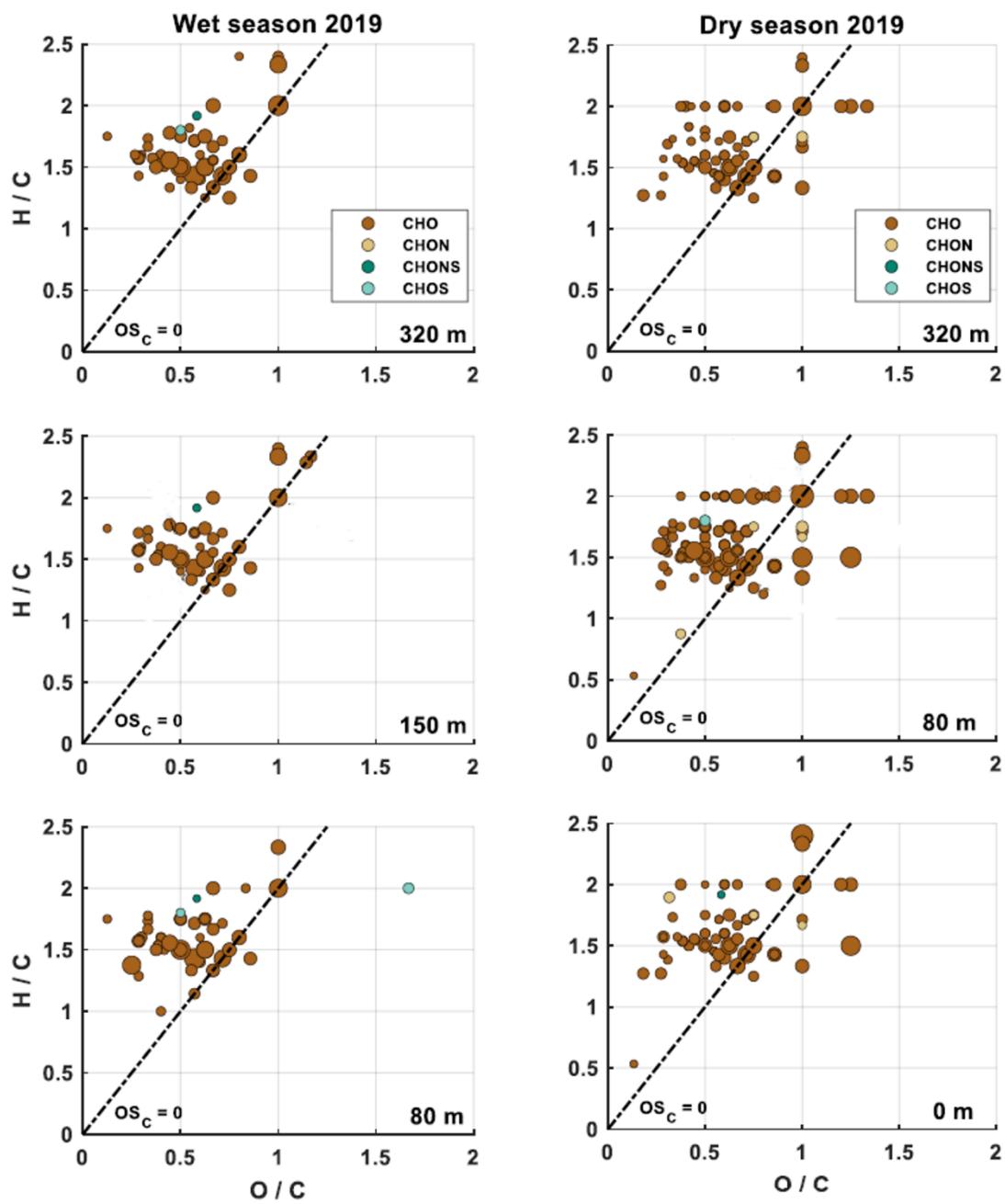


Figure S 9: Van Krevelen plots from the wet season 2019 (left panel; sampling heights: 320 m, 150 m, 80 m) and the dry season 2019 (right panel; sampling heights: 320 m, 80 m, 0 m). Included are only molecular formulae that were present in more than 75% of the samples, respectively. The size of the data points represents the signal intensity of the corresponding peak. The four subgroups are distinguished with different colors. Compounds located on the black dashed line have an average carbon oxidation state of 0 ($OSC = 0$).

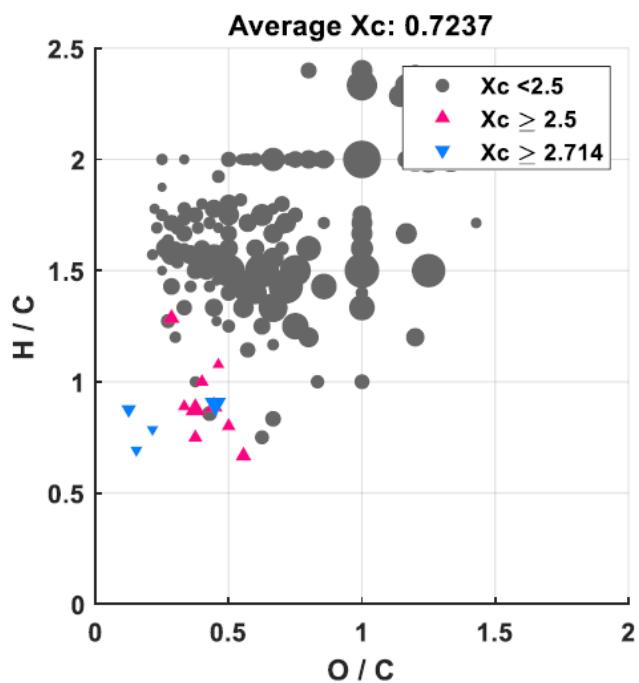


Figure S 10: Van Krevelen diagram illustrating the chemical background ions during the dry season 2018 at 80 m height. According to Yassine et al. (2014), all compounds were classified by the aromaticity equivalent Xc, highlighting unsaturated mono- and polycyclic structures.

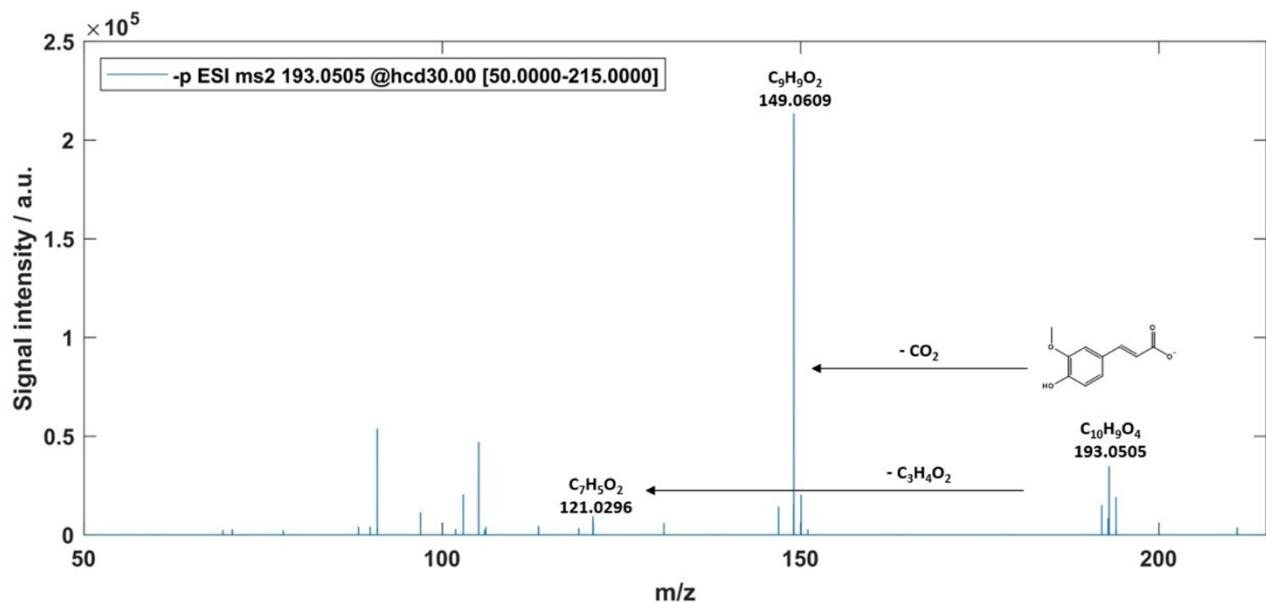


Figure S 11: MS^2 of the background aromatic species with the molecular formula $\text{C}_{10}\text{H}_{10}\text{O}_4$ (m/z 193.0505) during the dry season 2018. Presumably, the detected ion might be ferulic acid according to the characteristic losses of carbon dioxide and acrylic acid.

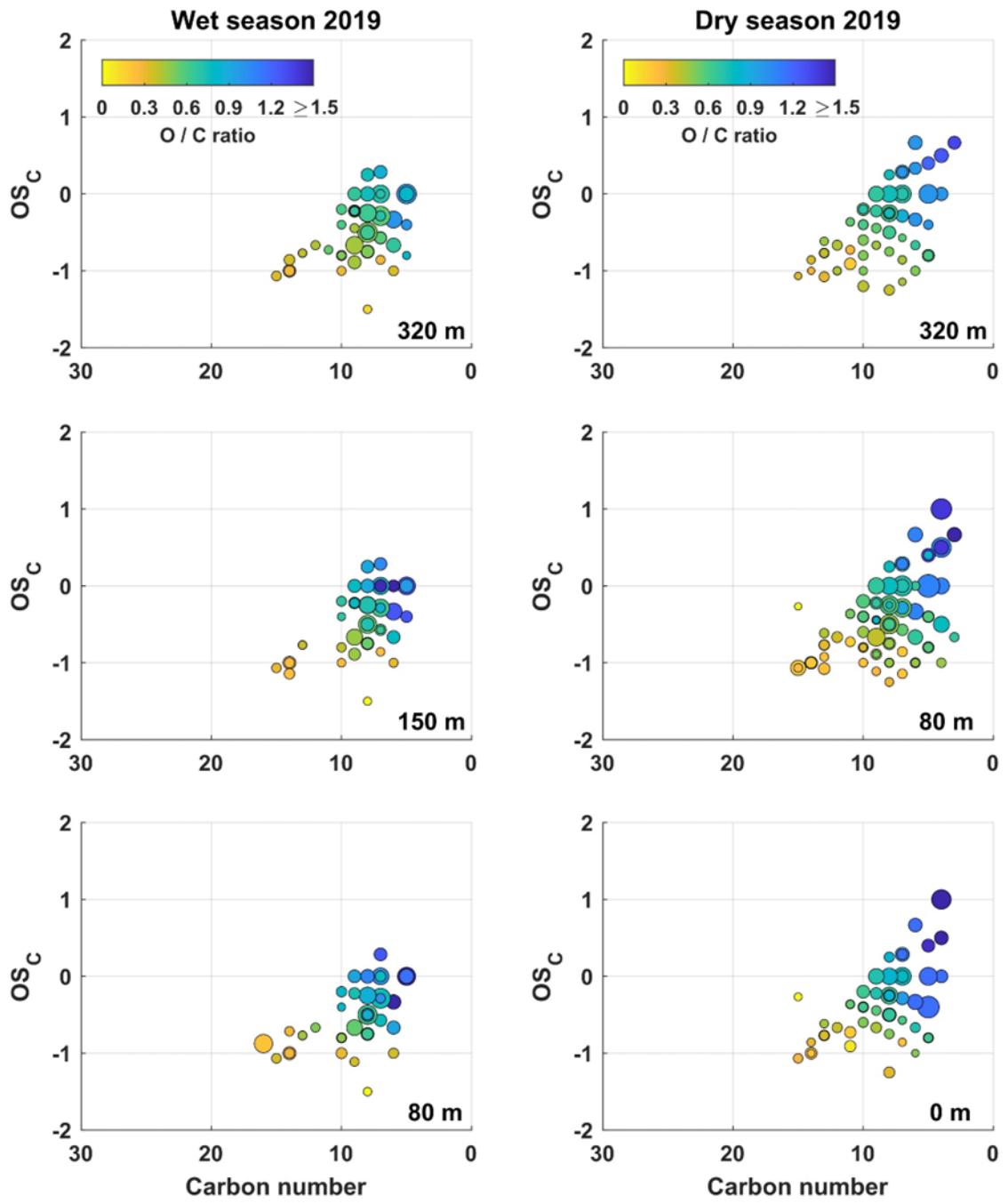


Figure S 12: Carbon oxidation state (OSC) plots for all detected CHO species during the wet season 2019 (left) and dry season 2019 (right) at different sampling heights. Only background ions are included. The size of the data points represents the signal intensity of the corresponding peak. The color code illustrates the degree of oxygenation.

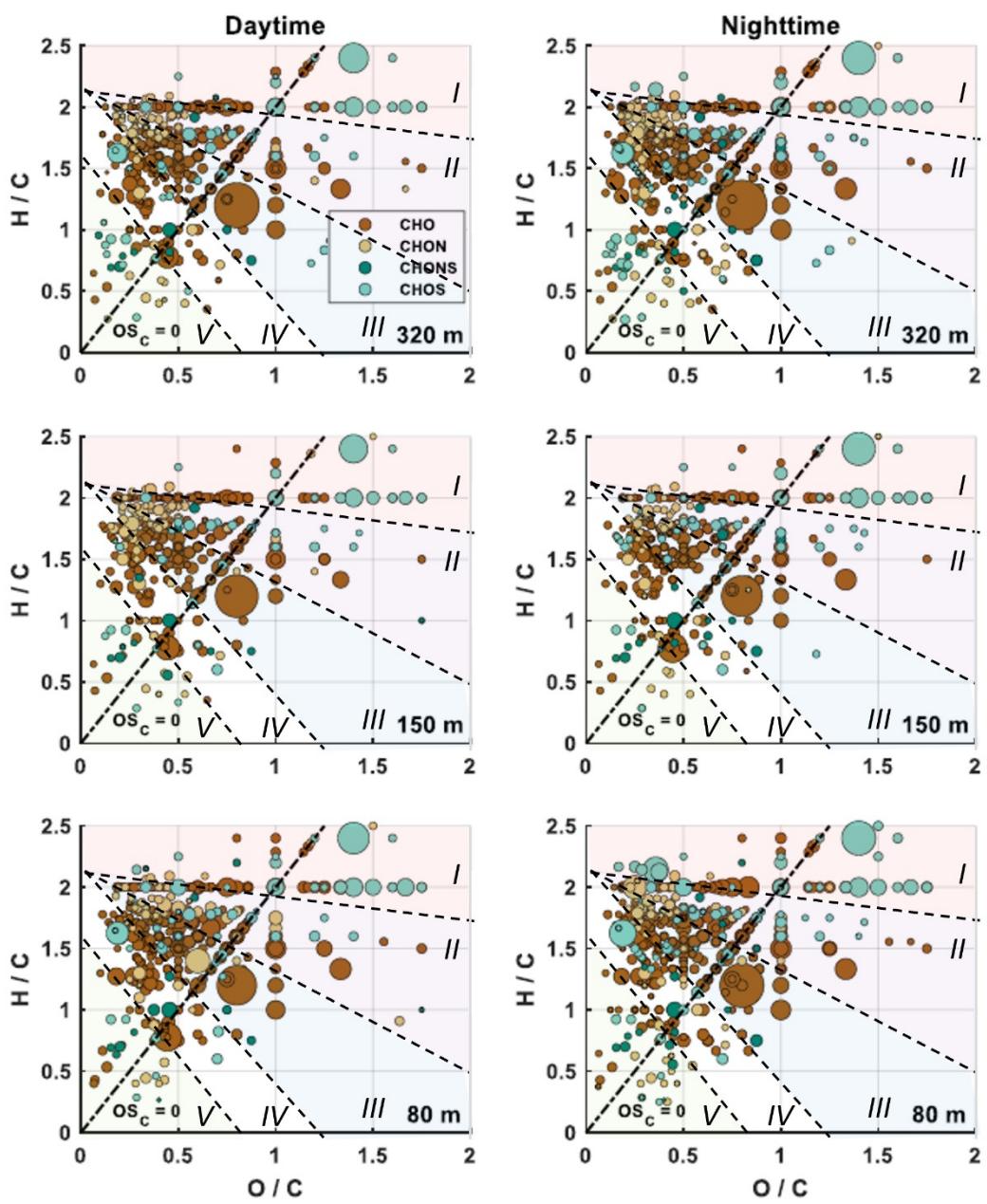


Figure S 13: Van Krevelen plots from the wet season 2019 (upper panel) during daytime (left panel) and nighttime (right panel) at 80 m, 150 m and 320 m height. Included are only molecular formulae that were present in more than one of the samples, respectively. The background signals are subtracted. The size of the data points represents the signal intensity of the corresponding peak. The four subgroups CHO, CHON, CHONS and CHOS are distinguished with different colors. Compounds located on the black dashed line have an average carbon oxidation state of 0 ($OSC = 0$).

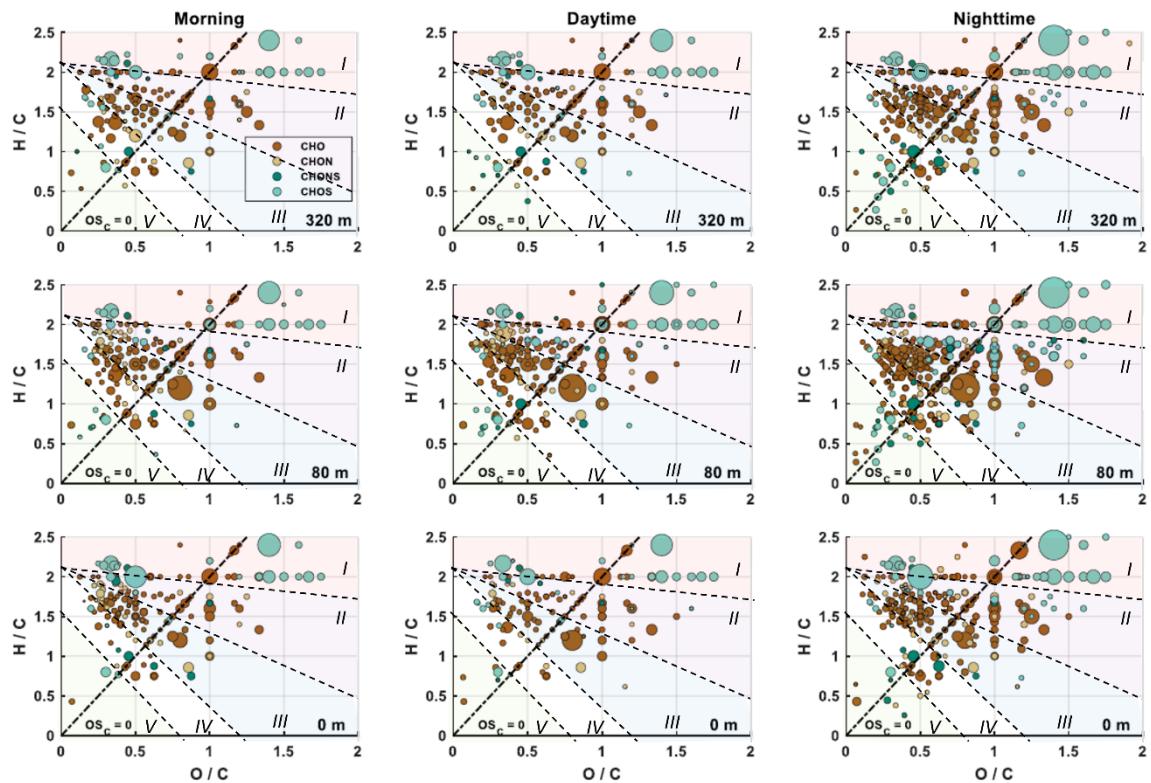


Figure S 14: Van Krevelen plots from the dry season 2019 during the morning (left panel), daytime (middle panel), and nighttime (right panel). Included are only molecular formulae that were present in more than one of the samples, respectively. The background signals are subtracted. The size of the data points represents the signal intensity of the corresponding peak. The four subgroups are distinguished with different colors. Compounds located on the black dashed line have an average carbon oxidation state of 0 ($OSC = 0$).

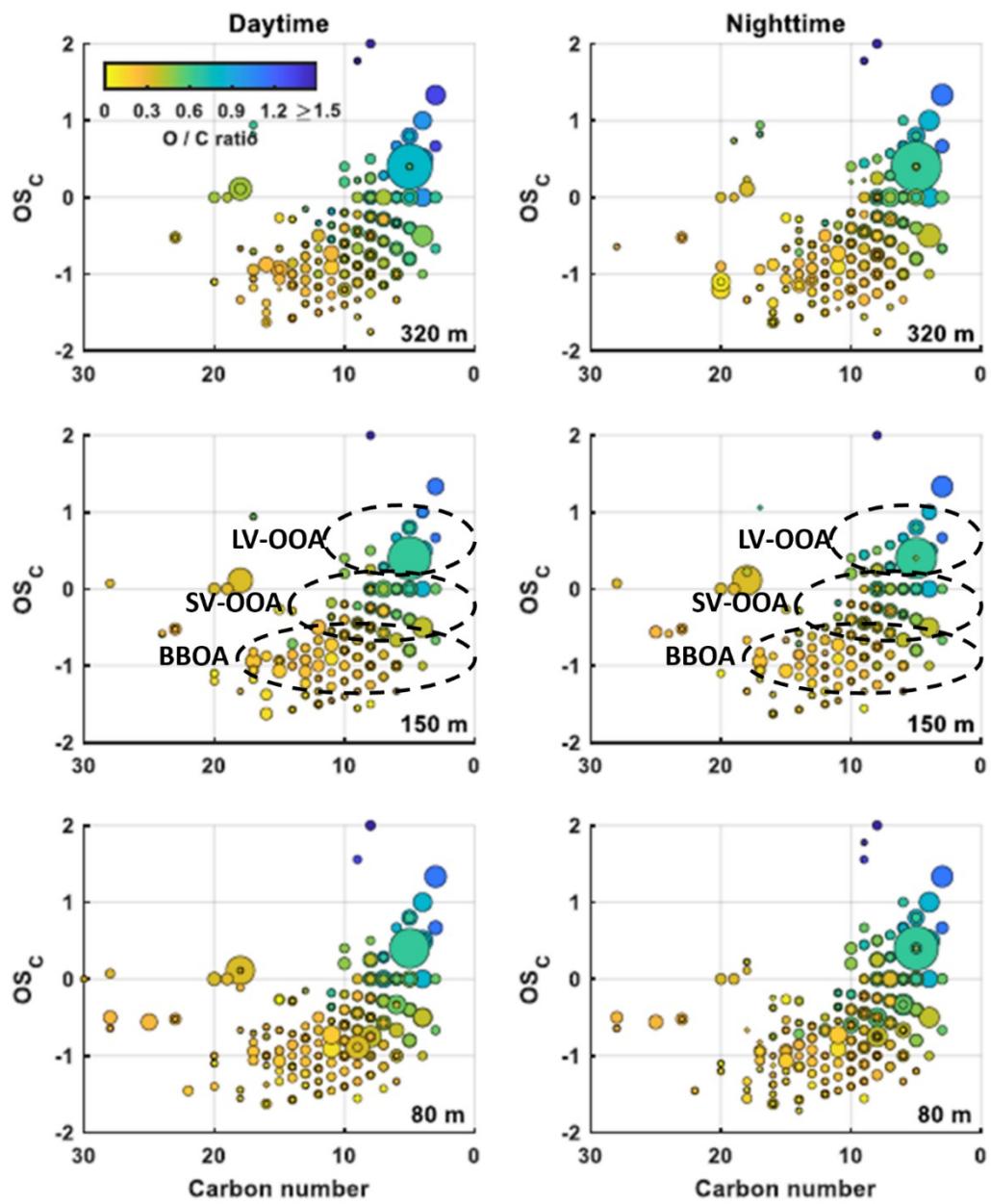


Figure S 15: Carbon oxidation state (OS_C) plots for all detected CHO species during the wet season 2019 at daytime (left panel) and nighttime (right panel) at 80 m, 150 m and 320 m sampling height. Included are only molecular formulae that were present in more than one of the samples, respectively. The background signals are subtracted. The size of the data points represents the signal intensity of the corresponding peak. The color code illustrates the degree of oxygenation. The black dashed areas are related to low- volatility (LV-OOA), semivolatile (SV-OOA) oxygenated organic aerosol and biomass burning organic aerosol (BBOA).

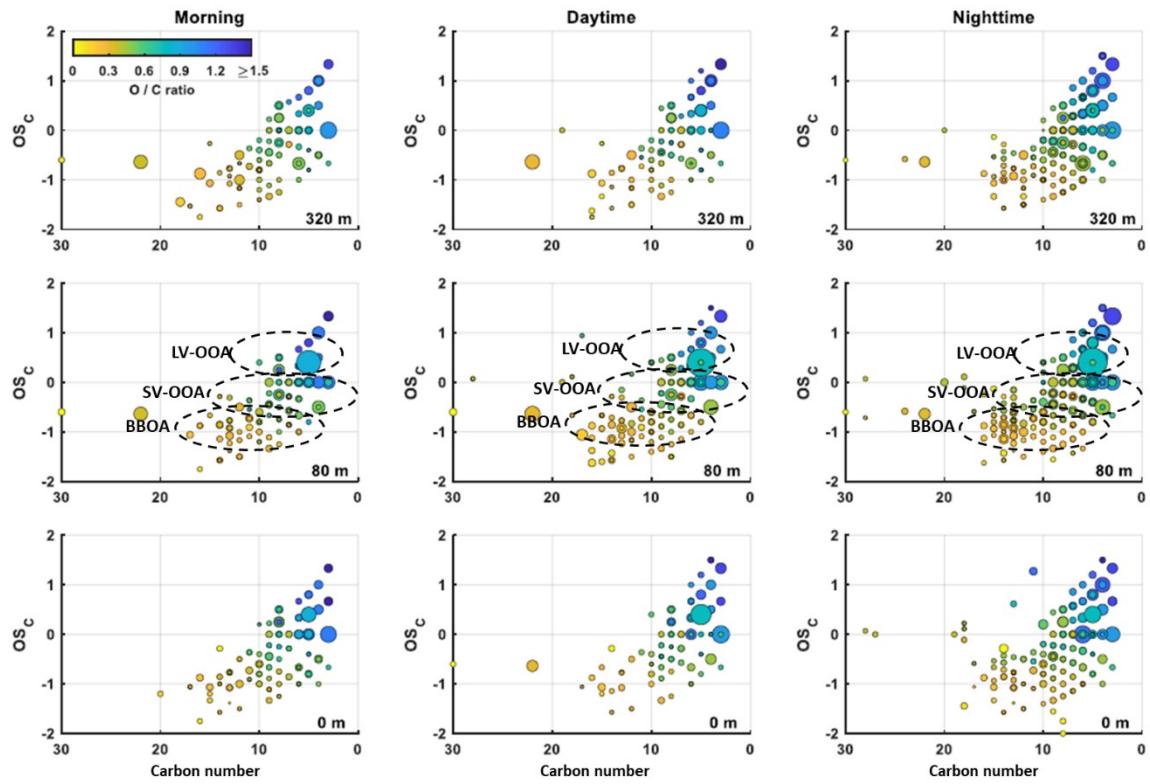


Figure S 16: Carbon oxidation state (OSC) plots for all detected CHO species during the dry season 2019 at morning (left), daytime (middle), and nighttime (right). Included are only molecular formulae that were present in more than one of the samples, respectively. The background signals are subtracted. The size of the data points represents the signal intensity of the corresponding peak. The color code illustrates the degree of oxygenation. The black dashed areas are related to low- volatility (LV-OOA), semivolatile (SV-OOA) oxygenated organic aerosol and biomass burning organic aerosol (BBOA).

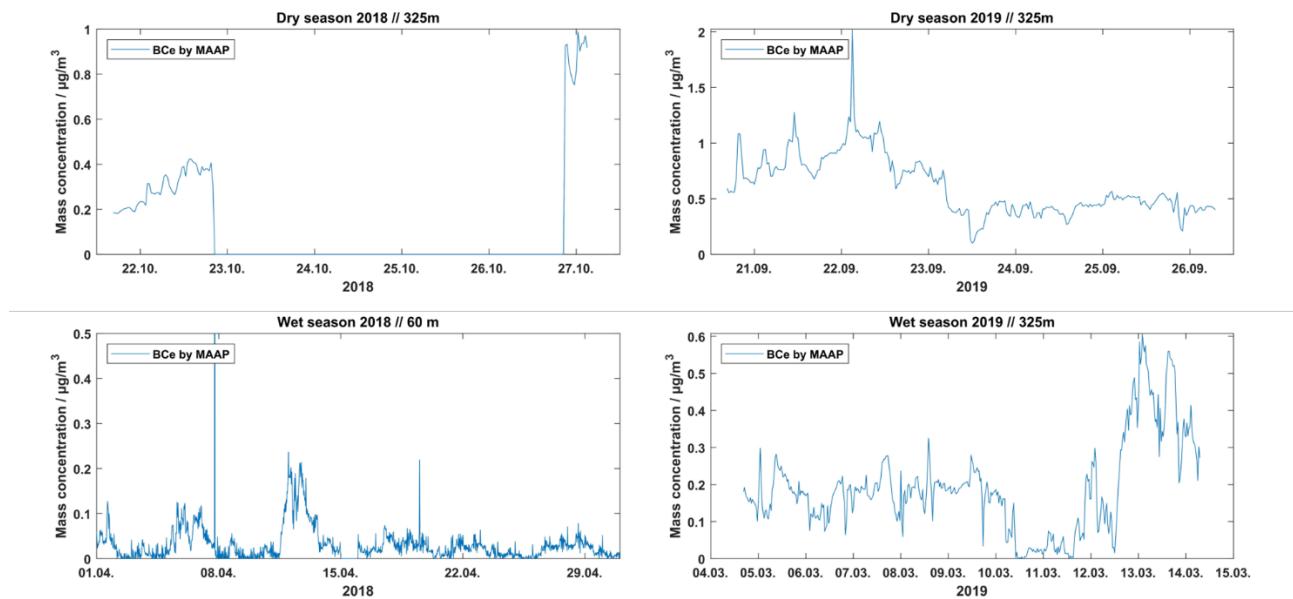


Figure S 17: Black carbon equivalent (BCe) mass concentrations during dry season 2018 and 2019 (upper part) and during the wet seasons 2018 and 2019 (lower part), respectively. Higher incidents of forest fires during the wet season 2019 are presumably responsible for increased BCe mass concentrations.

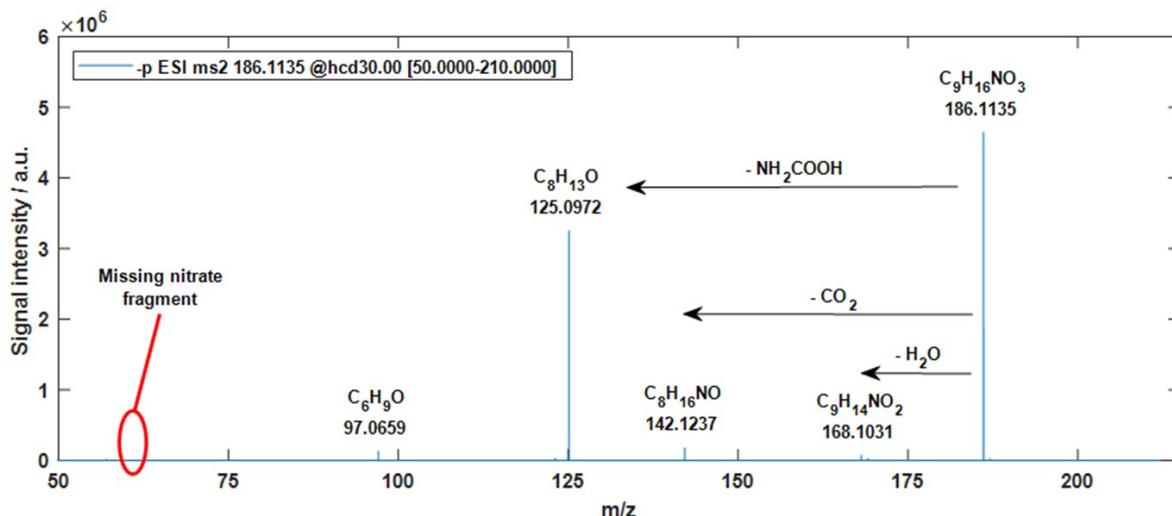


Figure S 18: MS^2 of m/z 186.1135 ($\text{C}_9\text{H}_{17}\text{NO}_3$). A typical nitrate fragment was not observed. Consequently, a nitrooxy-species can be excluded. In contrast, the observed fragments indicate an amide and a hydroxy-functional group within the molecular structure.

Method Evaluation

This section covers important aspects of the LC method development. This study focused on a single and sophisticated approach for the trace analysis of specific marker species on aerosol filter samples. Characteristic marker species can be used to identify emission sources for biogenic SOA or biomass burning events (Robinson et al., 2006a, 2006b). However, marker compounds must fulfill certain criteria, such as chemical stability against atmospheric oxidation and low volatility. These properties provide increased molecular lifetimes so that the marker species can also be detected after long-range transportation. For instance, levoglucosan (1,6-anhydro- β -D-glucopyranose) has been established as marker species for biomass burning activities. It is emitted into the atmosphere by the combustion of cellulose-containing wood (Simoneit et al., 1999). Biogenic SOA markers include oxidation products of isoprene, monoterpenes, and sesquiterpenes. Appropriate tracers for isoprene-derived SOA are 2-methyltetrols (Claeys et al., 2004a; Claeys et al., 2004b; Kourtchev et al., 2005). However, these compounds are highly polar due to the four hydroxyl groups, which complicates the analysis by RP-HPLC. The most important marker compound for aged biogenic SOA from monoterpenes is 3-methyl-1,2,3-butanetricarboxylic acid (MBTCA) formed by OH radical-initiated oxidation of α - and β -pinene (Szmigielski et al., 2007; Müller et al., 2012). Earlier-generation oxidation products, such as pinonic acid and pinic acid, can then be used to distinguish between freshly formed and processed SOA. Additionally, sesquiterpene oxidation products, such as β -caryophyllinic acid and β -nocaryophyllonic acid, are suitable marker species because of their low volatility (van Eijck et al., 2013).

The marker species used in this study for method development are listed in Table S 4. Salicylic acid (99%), 3-oxoglutaric acid (96%), terebic acid (98%), pimelic acid (96%), levoglucosan (99%), *cis*- pinonic acid (98%), methyltricarballylic acid ($\geq 88\%$), and camphoric acid (99%) were purchased from Sigma-Aldrich (Steinheim, Germany). D-malic acid ($\geq 99.5\%$) was obtained from Merck (Darmstadt, Germany). Vanillic acid was purchased from Alfa Aesar (Heysham, UK). Pinic acid was synthesized according to Moglioni et al. (2000). MBTCA was synthesized following the study by Dette et al. (2014). The sesquiterpene oxidation products β -caryophyllinic acid, β -caryophyllinic acid, and β - nocaryophyllonic acid were synthesized by van Eijck et al. (2013). Calibration

solutions were prepared in a mixture of water and acetonitrile (9:1) (Fisher Scientific, Optima™ grade) with concentrations ranging from 0.5 ng mL⁻¹ to 800 ng mL⁻¹. Blank solutions were prepared in the same manner.

Table S 4: Analytes used for method development. Also listed are the molecular formulae, *m/z* ratios and the calculated quantification limits (LOD).

Analyte	Molecular formula	<i>m/z</i> [M-H] ⁻	LOD ¹ / ng mL ⁻¹
D-malic acid	C ₄ H ₆ O ₅	133.0142	9.98
2-methyltetros	C ₅ H ₁₂ O ₄	135.0663	5.74
salicylic acid	C ₇ H ₆ O ₃	137.0244	2.11
3-oxoglutaric acid	C ₅ H ₆ O ₅	145.0142	2.71
terebic acid	C ₇ H ₁₀ O ₄	157.0506	3.09
pimelic acid	C ₇ H ₁₂ O ₄	159.0663	4.52
levoglucosan	C ₆ H ₁₀ O ₅	161.0455	29.68
vanillic acid	C ₈ H ₈ O ₄	167.0349	12.66
<i>cis</i> -pinonic acid	C ₁₀ H ₁₆ O ₃	183.1027	2.38
pinic acid	C ₉ H ₁₄ O ₄	185.0819	1.74
methyltricarballylic acid	C ₇ H ₁₀ O ₆	189.0405	2.17
camphoric acid	C ₁₀ H ₁₆ O ₄	199.0976	3.32
MBTCA	C ₈ H ₁₂ O ₆	203.0561	3.25
β-caryophyllonic acid	C ₁₅ H ₂₄ O ₃	251.1653	7.87
β-caryophyllinic acid	C ₁₄ H ₂₂ O ₄	253.1445	2.67
β-nocaryophyllonic acid	C ₁₄ H ₂₂ O ₄	253.1445	3.06

¹Instrumental detection limits calculated according to DIN 32645.

Each measurement was calibrated externally by six- to eight-point linear regression, which provided calibration functions for all analytes. The unknown concentrations of the filter samples were quantified after the determination of the signal areas and background subtraction. The limits of detection (LOD) were calculated according to DIN 32645.

Furthermore, the filter extraction method was evaluated to assess the extraction efficiency. Therefore, blank filters were spiked with 50 ng mL⁻¹ and 500 ng mL⁻¹ of α -pinene oxidation products, i.e., pinic acid and MBTCA, respectively. The experiments were performed in triplicates. The calibration functions are illustrated in Figure S 19. Based on the linear regression, extraction efficiencies of (93.6 \pm 3.9)% and (92.5 \pm 5.5)% were calculated for pinic acid and MBTCA, respectively. Thus, a quantitative extraction of SOA constituents from filter samples was assumed for all measurements.

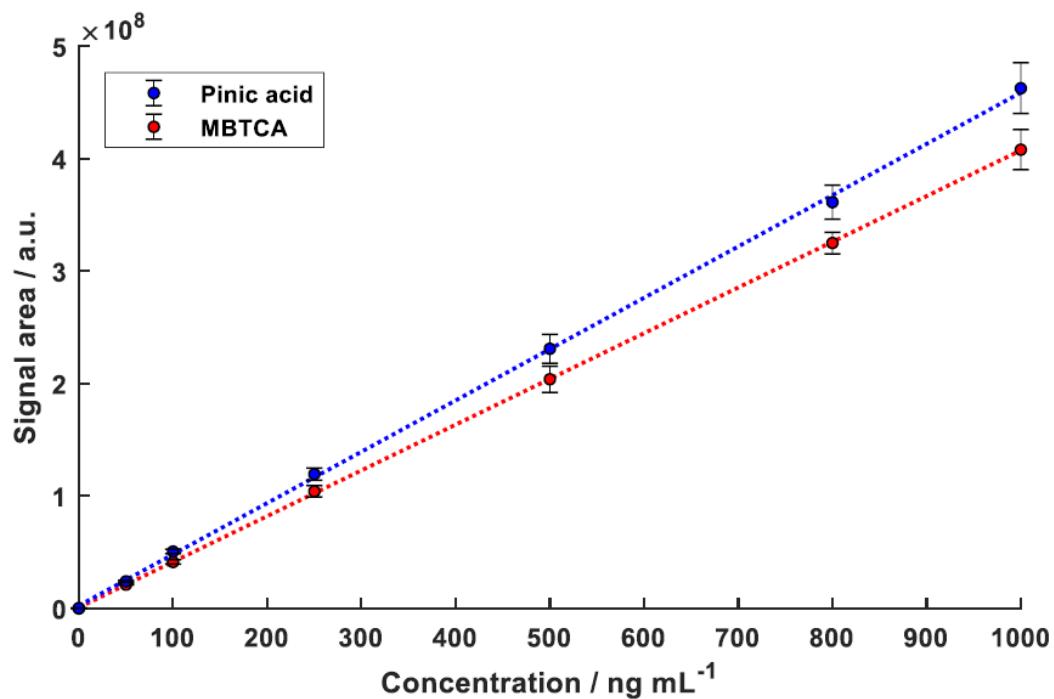


Figure S 19: External calibration functions of pinic acid (blue) and MBTCA (red) used for the validation of filter extraction efficiencies.

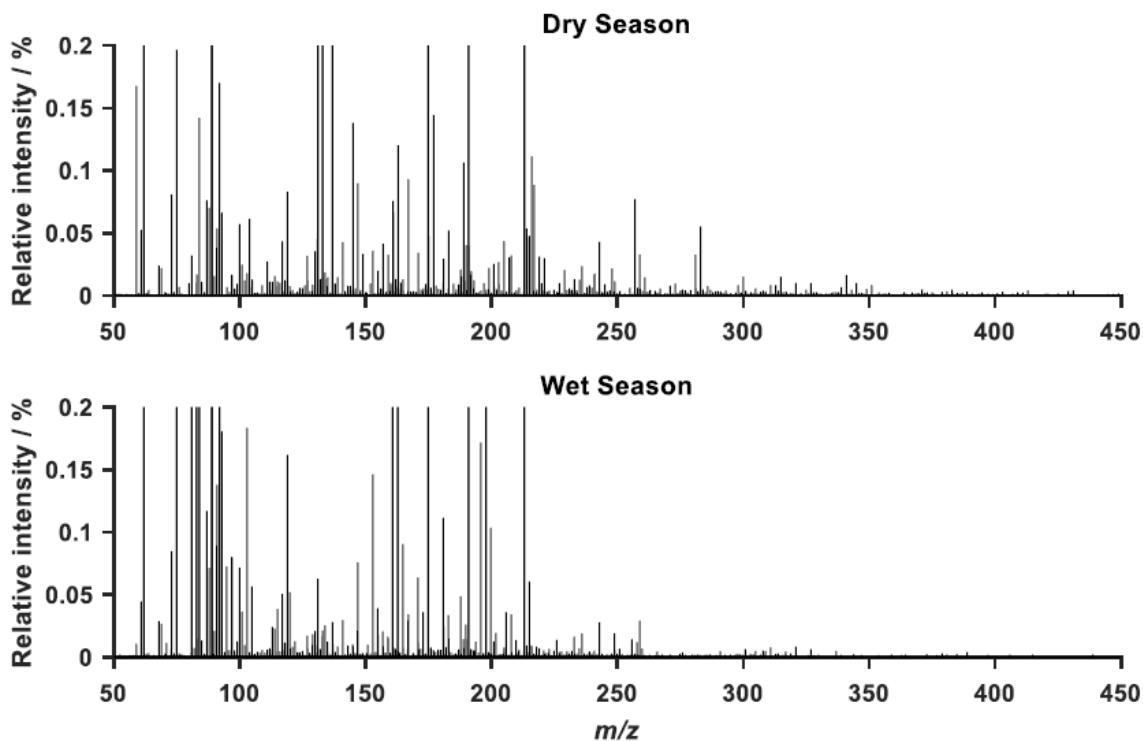


Figure S 20: Typical high-resolution mass spectra obtained in the negative ESI mode from filter samples of the dry season 2018 and wet season 2019, respectively. A UHPLC system separated the compounds before the injection into the mass spectrometer. The most intense ion signal at m/z 91.0036 was omitted for clarity reasons. It was attributed to the deprotonated formic acid dimer ($C_2H_3O_4$), which was added to the aqueous LC solvent.

Data Processing by MZmine 2.30

A non-targeted approach was developed and applied to the high-resolution mass spectra obtained from the analysis of each filter sample. Therefore, MZmine2.30 was used for data processing including toolboxes for peak detection,

filtering for shoulder peaks and duplicates, chromatogram builder, deconvolution and smoothing, alignment, adduct and complex search, and formula prediction (Pluskal et al., 2010; Wang et al., 2017). The following optimized parameterization was used in this study:

1. For each scan, only masses with signal intensities above 5000 were included for further processing. Signal intensities below 5000 were treated as noise and consequently not considered. Afterward, a shoulder peak filtering was applied with a mass resolution of 140 000 at m/z 200.
2. Chromatograms were built for each mass by using the ADAP Chromatogram Builder Module (Myers et al., 2017). The minimum intensity required to start a new chromatogram was set to 15 000 with a mass tolerance of 2 ppm.
3. The chromatograms were smoothed (filter width of 5) and deconvoluted by the ADAP Wavelet algorithm (Myers et al., 2017) (S/N threshold of 9; coefficient/area threshold of 110; peak duration of 0.03 – 1.00; retention time wavelet range of 0.01 – 0.17).
4. Isotopic peaks were removed (mass tolerance of 2 ppm; 0.1 min retention time tolerance; maximum charge of 1).
5. Adducts were removed (retention time tolerance of 0.1 min; mass tolerance of 2 ppm; maximum relative peak height of 50%; mass difference: m/z 41.0266 [M–H+ACN] $^-$, m/z 46.0055 [M–H+CH₂O₂] $^-$, m/z 67.9874 [M–H+CH₂ONa] $^-$). Ion complexes were also removed with the parameters mentioned above.
6. Peaks detected in different samples were aligned, based on the mass and retention time tolerances (mass tolerance of 2 ppm; weight for m/z of 5; retention time tolerance of 0.1 min; weight for retention time of 4). A minimum score of 85% for isotope patterns was required.
7. The formula prediction was based on [M–H] $^-$ ions with a mass tolerance of 2 ppm. Tentative molecular formulae with an elemental composition of CHONS were determined according to the following constraints: C₂–40H₂–100O₀–40N₀–4S₀–2; 0.1 \leq H/C \leq 6; N/C \leq 4; O/C \leq 3; S/C \leq 3; multiple- element counts according to Kind and Fiehn (2007); double bond equivalents (DBE) integer and between 0 – 30. Finally, all peaks were filtered for isotope pattern with a mass tolerance of 5 ppm. Therefore, the software was used in a batch mode with decreasing minimum scores for the isotope pattern match starting at 99% to 85%.
8. Peak duplicates were removed (mass tolerance of 2 ppm; retention time tolerance of 0.1 min)

The resulting peak lists contained m/z ratios, molecular formulae, retention times, and signal areas of each detected organic species. The mass spectra of blank samples were processed accordingly. All peak lists were further computed by Matlab R2017b (The MathWorks Inc., Natick, USA). Background subtraction was performed to only retain signals with sample-to-blank ratios \geq 3. Afterward, further elemental constraints were applied (H/C ratio of 0.3 – 3; N/C ratios of 0 – 0.5; S/C ratios of 0 – 0.2) to remove molecular formulae unlikely to be observed in nature (Wozniak et al., 2008; Fuller et al., 2012; Wang et al., 2017).

Table S 5: Confirmed species due to comparison with authentic standard substances. Also listed are the molecular formulae, *m/z* ratios and the retention time. The standard substances used are described in the section “method evaluation” in the supporting information. The shown species could be unambiguously assigned by comparing the exact masses and retention times.

Confirmed species	Molecular formula	<i>m/z</i> [M-H] ⁻	Retention time /min
D-malic acid	C ₄ H ₆ O ₅	133.0142	1.07
2-methyltetros	C ₅ H ₁₂ O ₄	135.0663	0.99
terebic acid	C ₇ H ₁₀ O ₄	157.0506	5.81
levoglucosan	C ₆ H ₁₀ O ₅	161.0455	1.13
pinic acid	C ₉ H ₁₄ O ₄	185.0819	7.63
MBTCA	C ₈ H ₁₂ O ₆	203.0561	6.01
β-caryophyllonic acid	C ₁₅ H ₂₄ O ₃	251.1653	12.94
β-caryophyllinic acid	C ₁₄ H ₂₂ O ₄	253.1445	12.59
β-nocaryophyllonic acid	C ₁₄ H ₂₂ O ₄	253.1445	11.43

References

Claeys, M., Wang, W., Ion, A. C., Kourtchev, I., Gelencsér, A., and Maenhaut, W.: Formation of secondary organic aerosols from isoprene and its gas-phase oxidation products through reaction with hydrogen peroxide, *Atmospheric Environment*, 38, 4093–4098, <https://doi.org/10.1016/j.atmosenv.2004.06.001>, 2004a.

Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., Cafmeyer, J., Guyon, P., Andreae, M. O., Artaxo, P., and Maenhaut, W.: Formation of secondary organic aerosols through photooxidation of isoprene, *Science (New York, N.Y.)*, 303, 1173–1176, <https://doi.org/10.1126/science.1092805>, 2004b.

Dette, H. P., Qi, M., Schröder, D. C., Godt, A., and Koop, T.: Glass-forming properties of 3-methylbutane-1,2,3-tricarboxylic acid and its mixtures with water and pinonic acid, *The journal of physical chemistry. A*, 118, 7024–7033, <https://doi.org/10.1021/jp505910w>, 2014.

Fuller, S. J., Zhao, Y., Cliff, S. S., Wexler, A. S., and Kalberer, M.: Direct surface analysis of time-resolved aerosol impactor samples with ultrahigh-resolution mass spectrometry, *Analytical chemistry*, 84, 9858–9864, <https://doi.org/10.1021/ac3020615>, 2012.

Kourtchev, I., Ruuskanen, T., Maenhaut, W., Kulmala, M., and Claeys, M.: Observation of 2-methyltetros and related photo-oxidation products of isoprene in boreal forest aerosols from Hyttiälä, Finland, *Atmos. Chem. Phys.*, 5, 2761–2770, <https://doi.org/10.5194/acp-5-2761-2005>, 2005.

Li, Y., Pöschl, U., and Shiraiwa, M.: Molecular corridors and parameterizations of volatility in the chemical evolution of organic aerosols, *Atmos. Chem. Phys.*, 16, 3327–3344, <https://doi.org/10.5194/acp-16-3327-2016>, 2016.

Moglioni, A. G., García-Expósito, E., Aguado, G. P., Parella, T., Branchadell, V., Moltrasio, G. Y., and Ortúñoz, R. M.: Divergent routes to chiral cyclobutane synthons from (-)-alpha-pinene and their use in the stereoselective synthesis of dehydro amino acids, *The Journal of organic chemistry*, 65, 3934–3940, <https://doi.org/10.1021/jo991773c>, 2000.

Müller, L., Reinnig, M.-C., Naumann, K. H., Saathoff, H., Mentel, T. F., Donahue, N. M., and Hoffmann, T.: Formation of 3-methyl-1,2,3-butanetricarboxylic acid via gas phase oxidation of pinonic acid – a mass spectrometric study of SOA aging, *Atmos. Chem. Phys.*, 12, 1483–1496, <https://doi.org/10.5194/acp-12-1483-2012>, 2012.

Myers, O. D., Sumner, S. J., Li, S., Barnes, S., and Du, X.: One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms

for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks, *Analytical chemistry*, 89, 8696–8703, <https://doi.org/10.1021/acs.analchem.7b00947>, 2017.

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M.: MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC bioinformatics*, 11, 395, <https://doi.org/10.1186/1471-2105-11-395>, 2010.

Robinson, A. L., Subramanian, R., Donahue, N. M., Bernardo-Bricker, A., and Rogge, W. F.: Source apportionment of molecular markers and organic aerosol. 2. Biomass smoke, *Environ. Sci. Technol.*, 40, 7811–7819, <https://doi.org/10.1021/es060782h>, 2006a.

Robinson, A. L., Subramanian, R., Donahue, N. M., Bernardo-Bricker, A., and Rogge, W. F.: Source apportionment of molecular markers and organic aerosol-1. Polycyclic aromatic hydrocarbons and methodology for data visualization, *Environ. Sci. Technol.*, 40, 7803–7810, <https://doi.org/10.1021/es0510414>, 2006b.

Simoneit, B.R.T., Schauer, J. J., Nolte, C. G., Oros, D. R., Elias, V. O., Fraser, M. P., Rogge, W. F., and Cass, G. R.: Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles, *Atmospheric Environment*, 33, 173–182, [https://doi.org/10.1016/S1352-2310\(98\)00145-9](https://doi.org/10.1016/S1352-2310(98)00145-9), 1999.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., and Ngan, F.: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, *Bulletin of the American Meteorological Society* [code], 96, 2059–2077, <https://doi.org/10.1175/BAMS-D-14-00110.1>, 2015.

Szmigielski, R., Surratt, J. D., Gómez-González, Y., van der Veken, P., Kourtchev, I., Vermeylen, R., Blockhuys, F., Jaoui, M., Kleindienst, T. E., Lewandowski, M., Offenberg, J. H., Edney, E. O., Seinfeld, J. H., Maenhaut, W., and Claeys, M.: 3-methyl-1,2,3-butanetricarboxylic acid: An atmospheric tracer for terpene secondary organic aerosol, *Geophys. Res. Lett.*, 34, D16312, <https://doi.org/10.1029/2007GL031338>, 2007.

van Eijck, A., Opatz, T., Taraborrelli, D., Sander, R., and Hoffmann, T.: New tracer compounds for secondary organic aerosol formation from β -caryophyllene oxidation, *Atmospheric Environment*, 80, 122–130, <https://doi.org/10.1016/j.atmosenv.2013.07.060>, 2013.

Wang, X., Hayeck, N., Brüggemann, M., Yao, L., Chen, H., Zhang, C., Emmelin, C., Chen, J., George, C., and Wang, L.: Chemical Characteristics of Organic Aerosols in Shanghai: A Study by Ultrahigh-Performance Liquid Chromatography Coupled With Orbitrap Mass Spectrometry, *J. Geophys. Res.*, 122, 11,703–11,722, <https://doi.org/10.1002/2017JD026930>, 2017.

Wozniak, A. S., Bauer, J. E., Sleighter, R. L., Dickhut, R. M., and Hatcher, P. G.: Technical Note: Molecular characterization of aerosol-derived water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Atmos. Chem. Phys.*, 8, 5099–5111, <https://doi.org/10.5194/acp-8-5099-2008>, 2008.

Yassine, M. M., Harir, M., Dabek-Zlotorzynska, E., and Schmitt-Kopplin, P.: Structural characterization of organic aerosol using Fourier transform ion cyclotron resonance mass spectrometry: aromaticity equivalent approach, *Rapid communications in mass spectrometry RCM*, 28, 2445–2454, <https://doi.org/10.1002/rcm.7038>, 2014.