Source identification of atmospheric organic vapors in two European pine forests: Results from Vocus PTR-TOF observations

Haiyan Li[1], Manjula R. Canagaratna[2], Matthieu Riva[3], Pekka Rantala[1], Yanjun Zhang[1], Steven Thomas[1], Liine Heikkinen[1], Pierre-Marie Flaud[4,5], Eric Villenave[4,5], Emilie Perraudin[4,5], Douglas Worsnop[2],
5 Markku Kulmala[1], Mikael Ehn[1], Federico Bianchi[1]

[1] Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, 00014, Finland
[2] Aerodyne Research Inc., Billerica, Massachusetts 01821, USA
[3] Univ. Lyon, Université Claude Bernard Lyon 1, CNRS, IRCELYON, 69626, Villeurbanne, France
10 [4] University of Bordeaux, EPOC, UMR 5805 CNRS, 33405 Talence Cedex, France
[5] CNRS, EPOC, UMR 5805 CNRS, 33405 Talence Cedex, France

Correspondence: Haiyan Li (haiyan.li@helsinki.fi) and Federico Bianchi (federico.bianchi@helsinki.fi)

Abstract.

15 Atmospheric organic vapors play essential roles in the formation of secondary organic aerosol. Source identification of these vapors is thus fundamental to understand their emission sources and chemical evolution in the atmosphere and their further impact on air quality and climate change. In this study, a Vocus proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF) was deployed in two forested environments, the Landes forest in southern France and the boreal forest in southern Finland, to measure atmospheric organic vapors, including both volatile organic compounds (VOCs) and their oxidation
20 products. For the first time, we performed binned positive matrix factorization (binPMF) analysis on the complex mass spectra acquired with the Vocus PTR-TOF and identified various emission sources as well as oxidation processes in the atmosphere. Based on separate analysis of low- and high-mass ranges, fifteen PMF factors in the Landes forest and nine PMF factors in the Finnish boreal forest were resolved, showing a high similarity between the two sites. Factors representing monoterpenes dominate the biogenic VOCs in both forests, with less contributions from the isoprene factors and sesquiterpene factors.
25 Particularly, various terpene reaction products were separated into individual PMF factors with varying oxidation degrees, such as lightly oxidized compounds from both monoterpene and sesquiterpene oxidations, monoterpene-derived organic nitrates, and monoterpene more oxidized compounds. These factors display similar mass profiles and diurnal variations between the two sites, revealing similar terpene reaction pathways in these forests. With the distinct characteristics of VOCs and oxygenated VOCs measured by the Vocus PTR-TOF, this study identified various primary emission sources and secondary
30 oxidation processes of atmospheric organic vapors in the European pine forests, providing a more comprehensive understanding of gas-phase atmospheric chemistry.

1 Introduction

Volatile organic compounds (VOCs) and their oxidation products are important contributors to atmospheric secondary organic aerosol (SOA) (Hallquist et al., 2009; Ehn et al., 2014) and new particle formation (Bianchi et al., 2016; Kirkby et al., 2016).

35 Therefore, the identification of these organic vapors and their sources is fundamental for understanding the effects of atmospheric aerosols on climate change and air quality (Schell et al., 2001; Maria et al., 2004). Large amounts of VOCs with varying physicochemical properties are emitted from both biogenic and anthropogenic sources (Friedrich et al., 1999; Kesselmeier et al., 1999), and their oxidation processes in the atmosphere can lead to the formation of thousands of structurally distinct products containing multiple functional groups (Hallquist et al., 2009; Wennberg et al., 2018). Due to the enormous

40 challenge in characterizing these organic vapors at molecular level, knowledge of their sources or formation pathways has remained lacking.

Globally, SOA production from biogenic sources is much larger than that from anthropogenic sources (Tsigaridis and Kanakidou, 2003). As a group of highly reactive gases, typically with one or more C = C double bounds, terpenes make up a major fraction of biogenic VOCs, including isoprene, monoterpenes, and sesquiterpenes (Guenther et al., 1995). In the

45 atmosphere, they react with various oxidants, i.e., hydroxyl radical (OH), ozone (O_3), and nitrate radical (NO_3), and produce a large variety of oxygenated molecules. Isoprene is the most emitted biogenic VOC on the global scale but has a relatively small SOA yield (Ahlberg et al., 2017; McFiggans et al., 2019). Monoterpenes are important sources of SOA (Ehn et al., 2014; Zhang et al., 2018) and their oxidation processes have been found to play important roles in new particle formation (Kirkby et al., 2016; Simon et al., 2020). High ambient concentrations of monoterpenes have been observed in numerous pine forests

50 (Hakola et al., 2012; Noe et al., 2012; Bourtsoukidis et al., 2014). While the concentrations of sesquiterpenes are generally much smaller than those of isoprene and monoterpenes (Sakulyanontvittaya et al., 2008; Sindelarova et al., 2014), sesquiterpenes could contribute significantly to SOA formation because of their reactivity and high aerosol yields (Calogirou et al., 1999; Khan et al., 2017). Previous studies found that sesquiterpenes contributed to the O_3 and OH reactivity in forest environments (Kim et al., 2011; Hellén et al., 2018). The recently developed Vocus proton-transfer-reaction time-of-flight

55 mass spectrometry (PTR-TOF) enables the real-time detection of both VOCs and their oxidation products. With a new chemical ionization source called Vocus, the instrument significantly improves its detection efficiency of product ions compared with conventional PTR instruments (Krechmer et al., 2018). Based on a laboratory comparison of different chemical ionization techniques, Riva et al. (2019) revealed that Vocus PTR-TOF is sensitive to a large range of oxygenated VOCs. With the deployment of a Vocus PTR-TOF in the French Landes forest, Li et al. (2020) observed various terpenes and terpene

60 oxidation products, including a range of organic nitrates.

With the benefit of the state-of-the-art capabilities of Vocus PTR-TOF to detect hundreds to thousands of molecules, a great challenge arises to analyze the complicated dataset where emission sources and atmospheric physical and chemical processes are mixed together. The characteristic analysis based solely on individually identified compounds cannot give the full picture of the measurements. Factor analytical techniques, e.g., positive matrix factorization (PMF), have been utilized to

65 extract information from mass spectrometer data by resolving co-varying signals with common sources or atmospheric processes into a single factor (Paatero and Tapper, 1994). For example, PMF analysis has been widely applied by the research community using aerosol mass spectrometer to identify multiple primary organic aerosol sources and SOA aging processes (Lanz et al., 2007; Ulbrich et al., 2009; Zhang et al., 2011). Yan et al. (2016) successfully applied PMF to unit-mass-resolution (UMR) nitrate ion-based chemical ionization mass spectrometer (NO_3^- CIMS) data to differentiate mainly monoterpene highly

70 oxygenated organic molecules (HOMs) formed from different formation pathways in the boreal forest. The application of PMF to high-resolution (HR) NO_3^- CIMS data by Massoli et al. (2018) identified more HOM factors at an isoprene-dominated forest site in Alabama, USA. Recently, the mass spectral binning combined with PMF (binPMF) was proposed as a novel and simple method for analyzing high-resolution mass spectra datasets (Zhang et al., 2019a). This approach divides the full mass spectra into small bins as input data to PMF, thus avoiding the time-consuming and complicated peak identification. Zhang et al.

75 (2019b) further applied binPMF to sub-ranges of ambient NO_3^- CIMS mass spectra and separated more meaningful factors related to chemical processes yielding HOMs.

In this work, we present the first factor analysis on Vocus PTR-TOF datasets to identify and apportion the contribution of different sources and formation pathways to atmospheric organic vapors. The measurements were conducted in two forest ecosystems in Europe, the French Landes forest and the boreal forest in southern Finland. Due to orders of magnitude

80 differences in the signal intensities of ions between lower mass range and higher mass range, we divided the mass spectra into two sub-ranges (50–200 Th and 201–320 Th) and performed binPMF analysis on these ranges separately. The resolved factors were linked to possible sources or chemistry processes by examining their mass profiles, time series, diurnal cycles, and correlation with molecular markers. Spatial comparisons were discussed for the common sources apportioned in both forests.

2 Materials and methods

2.1 Site description and measurement period

The measurement data were obtained during summertime in two forest environments in Europe, the Landes forest in southwestern France and the boreal forest research station SMEAR (Station for Measuring Forest Ecosystem-Atmosphere Relations) II in southern Finland. The field campaign in the Landes forest was conducted from 8 to 20 July 2018 as part of the Characterization of Emissions and Reactivity of Volatile Organic Compounds in the Landes forest (CERVOLAND) campaign.

90 An overview of the Vocus PTR-TOF measurements in the Landes forest has been presented earlier by Li et al. (2020). The ambient observations at the SMEAR II station were performed during 18 June – 18 July 2019.

The Landes forest (44°29'N, 0°57'W) is the largest man-made pine forest in Europe, mainly filled with maritime pine trees (*Pinus pinaster Aiton*). The sampling site is situated at the European Integrated Carbon Observation System (ICOS) station at Bilos. The nearest urban area of the Bordeaux metropole is around 40 km to the northwest. A more detailed description of the

95 measurement site can be found elsewhere (Kammer et al., 2018; Bsaibes et al., 2020; Li et al., 2020). Ambient meteorological

parameters, including temperature, relative humidity (RH), wind speed and direction, solar radiation, and pressure, and mixing ratios of nitrogen oxides (NO and NO_2) and O_3 were continuously monitored at the station throughout the campaign.

The SMEAR II station (61º51'N, 24º17'E) is located in a boreal mixed-coniferous forest in Hyytiälä, southern Finland (Hari and Kulmala, 2005). The site is dominated by a rather homogeneous Scots pine (*Pinus sylvestris L.*) stand and represents

100   a rural background measurement station. The nearest large city Tampere, located about 60 km to the southwest, has approximately 200 000 inhabitants. The station is equipped with extensive facilities to measure forest ecosystem-atmosphere interactions. Ambient meteorological parameters (i.e., global radiation, UVA, UVB, temperature, RH, pressure, and wind speed and direction), mixing ratios of various trace gases (i.e., carbon dioxide, carbon monoxide, sulfur dioxide, NO_x, and O_3), and particle concentration and size distribution, are continuously measured at the station.

2.2 Instrumentation

105

A Vocus PTR-TOF was deployed in both forest ecosystems to characterize atmospheric organic vapors. Equipped with a new chemical ionization source with a low-pressure reagent-ion source and focusing ion-molecule reactor (FIMR), the Vocus PTR-TOF is able to measure organic vapors with a wide range of volatilities (Krechmer et al., 2018; Riva et al., 2019; Li et al., 2020). A quadrupole radio frequency (RF) field inside the FIMR focuses ions to the central axis and improves the detection

110   efficiency of product ions. Compared with conventional PTR instruments, the sensitivity and detection efficiency of Vocus PTR-TOF are significantly improved (detection limit < 1 pptv). With a high water mixing ratio (10% v/v–20% v/v) in the FIMR, the instrument shows no humidity dependence for sensitivity. More instrumental details have been provided elsewhere (Krechmer et al., 2018; Li et al., 2020).

During both campaigns, we operated the Vocus ionization source at a pressure of 1.5 mbar. Sample air was drawn in

115   through a ~1-m-long PTFE tubing (10 mm o.d., 8 mm i.d.). A sample air flow of 4.5 L min^{-1} was used to reduce inlet wall losses and sampling delay. Around 100–150 ccm of this flow was sampled into the Vocus and the remainder was directed to the exhaust. The mass resolving power of the long TOF mass analyzer was 12 000 – 13 000 m Δm^{-1} during our measurements. Data were recorded with a time resolution of 5 s. During the campaign in the Landes forest, background checks were automatically performed every hour using ultra-high-purity nitrogen (UHP N_2). The instrument was calibrated twice a day

120   using a mixture of terpenes (α-/β-pinene+limonene; *p*-cymene). For measurements at the SMEAR II station, background measurements by injection of zero air using a built-in active carbon filter and quantitative calibrations with a multi-component standard cylinder were automatically conducted every three hours. All the *m/z* ratios mentioned in this work include the contribution from the charger ion (H$^+$, mass of 1 Th) unless stated otherwise.

2.3 binPMF data preparation and analysis

125   As described by Zhang et al. (2019a), binPMF divides the mass spectra into small bins and then takes advantage of PMF analysis to separate different sources or formation processes. The binPMF allows utilization of the high-resolution information of the complex mass spectra without the time-consuming and potentially error-prone steps of peak identification and peak

fitting before the factorization. Selected peaks of interest can be analyzed after binPMF, based on the output factors. PMF assumes that factor profiles are constant and unique, and the measured signal of a chemical component is a linear combination
130 of different factors. This approach does not require a priori information about the factors. The detailed working principle of PMF has been provided in numerous previous studies (Paatero and Tapper, 1994; Zhang et al., 2011; Yan et al., 2016).

To prepare the data and error matrices for PMF input, the Vocus PTR-TOF data were processed using the software package "Tofware" (v3.2.0; Tofwerk), which runs in the Igor Pro environment (WaveMetrics, OR, USA). The detailed data processing routines have been presented elsewhere (Stark et al., 2015). Signals were averaged over 30 min for data processing.
135 Unlike traditional UMR or HR fitting of the mass spectra, in binPMF analysis, the mass spectra were divided into small bins after mass calibration. Due to the greater mass resolving power of the TOF mass analyzer compared with former binPMF studies (Zhang et al., 2019a, 2019b), a bin width of 0.01 Th was applied in this study. At a nominal mass N, signals between N-0.15 and N+0.35 Th were included for binning. The error matrix was calculated to include uncertainty from counting statistics following Poisson distribution and instrument electronic noise, as described by Yan et al. (2016) and Zhang et al.
140 (2019a). The electric noise was estimated as the median of the standard deviation of binned noise signals between two nominal masses, with noise range between N+0.4 and N+0.6 Th.

Figure 1 shows the average mass spectra of the measurements in the Landes forest as an example. Since the signal intensity of larger molecules is generally much lower than that of low-mass molecules, we divided the mass spectra into two sub-ranges, the low mass range (51–200 Th) and the high mass range (201–320 Th). Factor analysis was separately performed on these
145 two sub-ranges using an Igor-based open-source PMF Evaluation Tool (PET; http://cires1.colorado.edu/jimenez-group/wiki/index.php/PMF-AMS_Analysis_Guide#PMF_Evaluation_Tool_Software). We ran the PMF up to ten factors for both sub-ranges. For the low mass range of 51–200 Th, the signals at m/z 81 Th ($C_6H_8H^+$, monoterpene fragment) and 137 Th ($C_{10}H_{16}H^+$, monoterpenes) were markedly higher than the others. In the initial PMF results, the mass profiles of all resolved factors were dominated by these ions. Therefore, the major mass bins of these ions were excluded for further PMF analysis,
150 but their corresponding isotopes were retained, effectively downweighting their contributions to the PMF result. However, to quantify the relative contribution of different factors, the signals of these removed mass bins were counted back into their corresponding factors. More details can be found in Sect. 4.4.

3 Dataset overview

Figure 2 shows the temporal behaviors of temperature, global radiation, concentrations of O_3 and NO_x, and concentrations of
155 isoprene and monoterpenes in the Landes forest and at the SMEAR II station. In the Landes forest, the weather was mainly sunny during the observation period (global radiation > 400 W m^{-2}), indicating strong photochemical activity. The air mass in the forest was largely influenced by local sources, with wind speeds below canopy lower than 1 m s^{-1} over the whole campaign. The O_3 concentration fluctuated dramatically between day and night, with the average daytime concentration peaking up to 50 ppb and the average nighttime level falling below 2 ppb (Li et al., 2020). The low O_3 concentration at night was probably to

160 some extent caused by its titration by monoterpenes (Fig. 2a; Kammer et al., 2018, 2020). The Landes forest is known for strong monoterpene emissions (Simon et al., 1994). During our measurements, the average mixing ratios of isoprene and monoterpenes were 0.6 ppb and 6.0 ppb, respectively. More details about this dataset can be found in Li et al. (2020). All data in the Landes forest are reported in local time and all data at the SMEAR II station in Finnish winter time (both equal UTC time + 2).

165 During the measurements at the SMEAR II station, 84% (26 out of 31) of the days had strong photochemistry (global radiation > 400 W m^{-2}), with the rest being cloudy days. The diurnal variation in O$_3$ concentration was not as dramatic as that in the Landes forest. In the daytime, the O$_3$ concentration sometimes reached up to 50 ppb. At night, the O$_3$ level still largely remained high, above 20 ppb, in contrast to the observations in the Landes forest. A possible explanation is less nighttime O$_3$ consumption by terpenes at the SMEAR II station. On average, the mixing ratios of isoprene and monoterpenes were 0.2 ppb

170 and 0.8 ppb, respectively, during the measurements, much lower than those in the Landes forest.

4 Results and discussion

4.1 Choice of PMF solution and factor interpretation

To interpret the PMF results, the most critical decision is to choose the best number of factors. More factors introduce more degrees of freedom to explain variations in the data, but too many factors may cause splitting of real factors and lead to

175 mathematical artifacts without physical meaning (Ulbrich et al., 2009). The factor interpretation results in this work are summarized in Table 1. In the factor name, L means the Landes forest and S means the SMEAR II station.

For the low mass range of the Landes forest dataset, the Q/Q_{exp} varied from 15.5 to 6.0 for two to ten factors (Q is the total sum of the squares of the scaled residuals for PMF solutions). The larger Q/Q_{exp} indicates underestimation of the errors or high residuals for some bins that cannot be simply modeled by the solution (Ulbrich et al., 2009). After seven factors,

180 increasing the factor number does not significantly decrease the Q/Q_{exp}. The optimal solution of seven factors was chosen after a detailed evaluation following the procedures proposed by Ulbrich et al. (2009) and Zhang et al. (2011). Figure S1 shows the distribution of scaled residuals as a function of m/z. For some bins the residuals are still high. The seven factors include Factor L1 closely related to the C$_4$H$_8$H$^+$ ion, Factor L2 attributed to a plume event occurring on a single night during the campaign, Factor L3 mainly containing lightly oxidized compounds with six or seven carbon atoms ("C$_6$" or "C$_7$"), Factor L4 representing

185 monoterpenes, Factor L5 indicative of isoprene and its oxidation products, Factor L6 identified as unknown source with large contributions from unknown peaks, and Factor L7 dominated by monoterpene lightly oxidized compounds. The direct comparison of the mass spectra, time series, and diurnal cycles of six-factor and eight-factor solutions are shown in Fig. S2 and Fig. S3. In the six-factor case, the C$_4$H$_8$H$^+$ ion-related factor cannot be separated. With eight-factor results, the factor representing isoprene and its oxidation products is split into two components with similar time series. For the high mass range

190 of the Landes forest dataset, the Q/Q_{exp} decreased from 2.5 to 0.9 for two to ten factors. After evaluation, we choose the eight-factor solution to explain the data. The Q/Q_{exp} value of the eight-factor solution was 1.1 and the decreasing trend in Q/Q_{exp}

obviously slowed down after eight factors. The distribution of scaled residuals as a function of *m/z* for the eight-factor solution is shown in Fig. S4. The eight factors are interpreted as Factor L8 dominated by lightly oxygenated compounds containing 13 carbon atoms ("C_{13}"), Factor L9 attributed to a plume event occurring on a single night during the campaign, Factor L10

195   mainly related to sesquiterpene lightly oxidized compounds, Factor L11 representing more oxidized products mainly from monoterpene oxidations, Factor L12 indicating sesquiterpenes, Factor L13 largely composed of monoterpene-derived organic nitrates, Factor L14 mainly containing oxidized compounds with twelve, fourteen or sixteen carbon atoms ("C_{12}", "C_{14}" or "C_{16}")and Factor L15 as unknown source largely contributed by siloxane compounds. Figure S5 and Figure S6 display the mass spectra, time series, and daily variations of seven-factor and nine-factor solutions. In the seven-factor case, monoterpene

200   more oxidized products and monoterpene-derived organic nitrates are mixed together into a single factor. However, in the nine-factor solution, the unknown factor mainly composed of siloxane compounds is split into two factors with similar mass profiles and similar diurnal trends.

For the SMEAR II dataset, the optimal solutions of five-factor and four-factor are chosen for the low and high mass ranges, respectively. The Q/Q_{exp} varied from 7.2 to 2.5 for two to ten factors in the low mass range and from 2.0 to 1.0 for two

205   to ten factors in the high mass range. The five factors for the low mass range are identified as Factor S1 - $C_4H_8H^+$ ion-related, Factor S2 - monoterpenes, Factor S3 - lightly oxidized compounds with six to nine carbon atoms, Factor S4 - isoprene and its oxidation products, and Factor S5 - monoterpene lightly oxidized compounds. The mass spectra, time series, and diurnal profiles of the four-factor and six-factor solutions for the low mass range are presented in Fig. S7 and Fig. S8. For the four-factor solution, monoterpene lightly oxidized products are not separated as a single factor and mixed into the others. In the six-

210   factor case, the factor indicative of monoterpene lightly oxidized products is split into two factors. The four factors for the high mass range include Factor S6 - sesquiterpene lightly oxidized products, Factor S7 - sesquiterpenes, Factor S8 - more oxidized compounds, and Factor S9 - unknown source. The direct comparison of the mass spectra, time series, and diurnal variations of three-factor and five-factor solutions are shown in Fig. S9 and Fig. S10. The three-factor solution does not identify a factor representing sesquiterpenes. In the five-factor case, the factor of unknown source mainly contributed by siloxane compounds

215   is split into two factors with similar mass profiles.

4.2 Source identification in the Landes forest

4.2.1 Low mass range (51–200 Th)

Figure 3 and Figure 4 illustrate the factor profiles, time series, and diurnal variations of the seven factors resolved in the low mass range. The high-resolution peak fitting was further performed on the mass profile to identify the fingerprint peaks in the

220   factors. Fingerprint peaks are defined largely based on their distribution in the factors rather than their absolute intensity in the mass profile. The correlation map of each factor with various compounds is shown in Fig. 5.

Factor L1: $C_4H_8H^+$ ion-related

Factor L1 shows irregular diurnal variations with spiky peaks in the time series (Fig. 4b). The major bins that are largely distributed into this factor are $C_4H_8H^+$ and $C_4H_{10}O_2H^+$. Factor L1 closely correlates with these fingerprint peaks. Considering

225   the high signal intensity of $C_4H_8H^+$ ion and its large contribution to this factor, we name Factor L1 as $C_4H_8H^+$ ion-related. According to the discussions by Li et al. (2020), the observation of $C_4H_8H^+$ in the Landes forest can be attributed to several sources. For instance, the protonated butene may contribute to the $C_4H_8H^+$ signal, which is emitted by biogenic or anthropogenic sources (Hellén et al., 2006; Zhu et al., 2017). Another possible explanation is that the $C_4H_8H^+$ ion is produced during the fragmentation of many VOCs in the PTR instruments (Pagonis et al., 2019). The green leaf volatiles (GLV) have

230   been found to fragment at m/z 57 Th inside the PTR instruments, which are a group of six-carbon aldehyde, alcohols and their esters released by plants. Furthermore, butanol can easily lose an OH during the PTR source ionization and produce prominent $C_4H_8H^+$ peaks (Spanel and Smith, 1997). Therefore, the condensation particle counters (CPCs) using butanol for aerosol measurements at the site could also be an important source of $C_4H_8H^+$ ions, although the exhaust air from these instruments has been filtered using charcoal denuder.

235   *Factor L2: A plume event*

Factor L2 is identified as a plume event occurring on a single night during the campaign. As shown in Fig. 4a, the time series of this factor are characterized by much higher intensities at midnight of 9 July 2018 than over the other days. Fingerprint peaks in this factor are aromatic compounds such as $C_6H_6H^+$, $C_7H_6H^+$, and $C_6H_6OH^+$. Factor L2 is well correlated with benzene and phenol ($r^2 = 0.88$; Fig. 5), indicating the major influence of anthropogenic sources. As mentioned above, the air masses in

240   the forest were relatively stable during our observations with wind speed below canopy < 1 m s^{-1}. Therefore, the influence of long-range regional transport on the atmosphere in the forest is expected to be minor. We speculate that Factor L2 is a result of local anthropogenic disturbances favored by the lower boundary layer height at night.

Factor L3: C_6 and C_7 lightly oxidized products

The diurnal cycle of Factor L3 exhibits a small morning peak at 9:00 and significantly elevated intensities during

245   nighttime, peaking at around 22:00 (Fig. 4b). As illustrated in the mass profile of Factor L3, this factor is mainly composed of lightly oxidized compounds containing six or seven carbon atoms such as $C_6H_{10}OH^+$, $C_7H_{10}OH^+$, $C_6H_{10}O_2H^+$, and $C_7H_{12}O_2H^+$. The C_6 oxygenated compounds have been observed during the oxidation processes of benzene and C_7 oxygenated compounds from toluene oxidations (Sato et al., 2012; Zaytsev et al., 2019). These compounds can also be directly emitted from biogenic or anthropogenic sources (Conley et al., 2005; Pandya et al., 2006; Rantala et al., 2015). Factor L3 shows tight correlations

250   with $C_7H_{10}OH^+$ and $C_6H_{10}O_2H^+$ ($r^2 > 0.92$; Fig. 5).

Factor L4: monoterpenes

The mass profile of Factor L4 is dramatically characterized by a monoterpene peak ($^{13}CC_9H_{16}H^+$) and its major fragments (i.e., $^{13}CC_5H_8H^+$ and $C_7H_8H^+$) inside the instrument. As shown in Fig. 4b, the diurnal variation of this factor follows a similar pattern to that of monoterpenes (Li et al., 2020). The signal intensity of the factor starts to increase at 20:00, peaks at midnight,

255   and then drops to around the detection limit during daytime. Monoterpene emissions are mainly influenced by temperature (Hakola et al., 2006; Kaser et al., 2013). Therefore, with the continuous emissions of monoterpenes and the shallow boundary

layer at night, the signal intensities of monoterpenes are observed to be elevated. The lower signal of Factor L4 in the daytime is likely a combination of enhanced atmospheric mixing after sunrise and the rapid photochemical consumption of monoterpenes. The signal of $C_{10}H_{16}OH^+$ is also mostly resolved into this factor. $C_{10}H_{16}O$ could be primary emissions of

260   oxygenated monoterpenes or monoterpene oxidation products (Kallio et al., 2006; McKinney et al., 2011). Previous ambient observation has demonstrated that the atmospheric behavior of $C_{10}H_{16}O$ has high similarity to that of monoterpenes (Li et al., 2020).

Factor L5: isoprene and its oxidation products

The marker peaks in Factor L5 are highly dominated by isoprene and its major oxidation products in the atmosphere, i.e.,

265    $C_5H_8H^+$ and $C_4H_6OH^+$ (Wennberg et al., 2018). Isoprene emissions strongly depend on light intensity (Monson and Fall, 1989; Kaser et al., 2013) and generally show high concentrations in the day. Similarly, the daily variations of Factor L5 display maximum signal during daytime and minima at night.

Factor L6: unknown source

Factor L6 is characterized by increased signals in the afternoon. The major peaks in its factor profile are $C_6H_4O_2H^+$,

270    $C_6H_6O_3H^+$, and numerous unidentified peaks with negative mass defect. As this factor is clearly separated as a single source with high signals during our observations and the molecule markers remain unidentified, we name this factor as an unknown source.

Factor L7: monoterpene lightly oxidized products

Fingerprint peaks in this factor are monoterpene oxidation products with oxygen number from one to three, such as

275    $C_9H_{14}OH^+$, $C_{10}H_{14}OH^+$, $C_{10}H_{16}O_2H^+$, and $C_{10}H_{16}O_3H^+$. This factor displays clear morning and evening peaks, similar to the behavior of these lightly oxidized compounds (Li et al., 2020). By calculating the reaction rates of monoterpenes with OH and O_3, Li et al. (2020) demonstrated that both OH- and O_3-initiated oxidation processes contribute to the formation of these compounds in the Landes forest.

4.2.2 High mass range (201–320 Th)

280   The mass spectra of the five factors identified in the high mass range are shown in Fig. 6, and their time series and daily variations in Fig. 7. Figure 5 includes the correlations of these five factors with fingerprint molecules.

Factor L8: C_{13} lightly oxidized products

The mass profile of Factor L8 is characterized by high peaks of lightly oxidized compounds containing 13 carbon atoms, like $C_{13}H_{18}O_2H^+$ and $C_{13}H_{20}O_3H^+$. Similar to C_6 and C_7 lightly oxidized compounds, this factor shows a morning peak at 9:00

285   and an evening peak at around midnight (Fig. 7b). The time series of this factor correlate well with those of Factor L3 and Factor L7 (r^2 = 0.64 and 0.40; Fig. S11), indicating potentially similar formation mechanisms of these lightly oxygenated compounds. Therefore, the C_{13} oxidized compounds are speculated to be produced through the dimer formation mechanisms of C_6 and C_7 peroxy radicals (Valiev et al., 2019). In addition, $C_{13}H_{20}O_3$ can be direct emissions of methyl jasmonate (Meja), which is a typical green leaf volatile used in plant-plant communications for defensive purposes (Cheong and Choi, 2003). But

290   with similar temporal behaviors to Factor L3 and Factor L7, we conclude that these C_{13} lightly oxidized compounds are formed

from atmospheric oxidation processes, not direct plant emissions.

Factor L9: A plume event

Similar to Factor L2, Factor L9 is characterized with much high intensities on a single night (9 July 2018) during the

campaign (Fig. 7a). Fingerprint peaks in the mass profile of Factor L9 are numerous unidentified ions. The time series of

295   Factor L9 correlate tightly with those of Factor L2 ($r^2 = 0.93$) and aromatic compounds C_6H_6 and C_6H_6O ($r^2 = 0.75$). Therefore,

we define Factor L9 as a similar plume event to Factor L2 resolved in the high mass range.

Factor L10: sesquiterpene lightly oxidized products

The fingerprint peaks identified in this factor are $C_{15}H_{22}OH^+$, $C_{15}H_{24}OH^+$, $C_{15}H_{22}O_2H^+$, $C_{15}H_{24}O_2H^+$, and $C_{15}H_{24}O_3H^+$,

which are typical reaction products from sesquiterpene oxidations (Fu et al., 2009; Yee et al., 2018). The signal intensity of

300   this factor is generally high during nighttime, but shows another morning peak at 8:00. In addition to the production from

sesquiterpene oxidation processes, $C_{15}H_{22}O$ and $C_{15}H_{24}O$ can be oxygenated sesquiterpene alcohols and aldehydes directly

emitted from vegetation (Kännaste et al., 2014).

Factor L11: monoterpene more oxidized products

The mass spectrum of this factor is mainly characterized by more oxidized compounds from monoterpene oxidations such

305   as $C_{10}H_{16}O_4H^+$, $C_{10}H_{14}O_5H^+$, $C_{10}H_{16}O_5H^+$, and $C_{10}H_{16}O_6H^+$. As shown in Fig. 5, the time series of Factor L11 show good

correlations with these compounds. Compared with monoterpene lightly oxidized compounds, the diurnal cycle of this factor

shows a broad daytime distribution peaking between 14:00 and 20:00, caused by strong and complicated photochemical

reactions during the day.

Factor L12: sesquiterpenes

310   The mass spectra of Factor L12 are clearly dominated by a big single peak of $C_{15}H_{24}H^+$, indicating the influence of

sesquiterpenes. Sesquiterpene emissions from plants are found to exhibit a strong dependence on temperature (Duhl et al.,

2008). Therefore, similar to the diurnal cycle of Factor L4, this factor shows prominently enhanced signals during nighttime.

As shown in Fig. S11, Factor L12 and Factor L4 correlate quite well with each other ($r^2 = 0.69$).

Factor L13: monoterpene-derived organic nitrates

315   The signal intensity of this factor starts to increase in the early morning (around 7:00) and presents a distinct morning

peak at 9:00. In addition, a much smaller evening peak is observed at 21:00. The daily variations of this factor are quite similar

to those of monoterpene-derived organic nitrates measured in the Landes forest (Li et al., 2020). Consistently, the major peaks

in the factor profile are $C_{10}H_{15}NO_4H^+$, $C_{10}H_{15}NO_5H^+$, $C_9H_{13}NO_6H^+$, and $C_{10}H_{15}NO_6H^+$, indicating the dominant contribution

of organic nitrates formed from monoterpene oxidation processes. According to the calculation of the reaction rates of

320   monoterpenes with OH radical, O_3, and NO_3 radical, the big morning peak came from O_3- and OH-initiated monoterpene

oxidations in the presence of NO_x, while for the small evening peak the additional contribution of NO_3 radical-induced

monoterpene oxidations should be included (Li et al., 2020). $C_{10}H_{17}NO_4$, $C_{10}H_{15}NO_5$, $C_{10}H_{17}NO_5$, and $C_{10}H_{15}NO_6$ have been

found to be major gas-phase organic nitrates from α-pinene and/or β-pinene + NO_3 oxidation systems (Wängberg et al., 1997; Perraud et al., 2010; Boyd et al., 2015).

325 *Factor L14: C_{12}, C_{14} or C_{16} lightly oxidized compounds*

The mass profile of Factor L14 is characterized with distinct peaks of C_{12}, C_{14} or C_{16} lightly oxidized compounds, i.e., $C_{12}H_{26}O_3H^+$, $C_{14}H_{26}O_2H^+$, $C_{16}H_{30}O_2H^+$, and $C_{16}H_{30}O_3H^+$. The time series of Factor L14 correlate very well with those of $C_{12}H_{26}O_3$ ($r^2 = 0.83$), characterized with enhanced signals during daytime and low intensities at night (Fig. 7b). $C_{12}H_{26}O_3$ has been found during the photooxidation of dodecane (Zhang et al., 2014).

330 *Factor L15: unknown source*The mass profile of Factor L15 is predominantly characterized by high cyclic volatile methyl siloxanes (VMSs) peaks and some unidentified peaks (Fig. 6). The major cyclic VMSs are protonated D3 siloxane, D4 siloxane, and their H_3O^+ cluster ions, which have been widely used in cosmetics and personal care products (Buser et al., 2013; Yucuis et al., 2013). The diurnal cycle of this factor shows a bit higher intensity during daytime but also big background signals at night. A similar factor has also been identified at the SMEAR II station. More detailed discussions can be found in Sect. 4.3.2.

335 **4.3 Source identification in the southern Finnish boreal forest**

4.3.1 Low mass range (51–200 Th)

The factor profiles, time series, and diurnal cycles of the five-factor solution for the low mass range are presented in Fig. 8 and Fig. 9. Figure 10 illustrates the correlation of the five factors with various molecules.

Factor S1: $C_4H_8H^+$ ion-related

340 Similar to the source identification in the Landes forest, a factor related to $C_4H_8H^+$ ion is clearly resolved at the SMEAR II station. The major peaks in this factor are $C_4H_8H^+$, $C_4H_{12}O_2H^+$, and $C_4H_{14}O_3H^+$. As discussed in Section 4.2.1, several sources could contribute to the detection of $C_4H_8H^+$ ion. However, at this site, the bivariate polar plot where the concentrations of air pollutants are shown as a function of WS and WD indicates that high signals of $C_4H_8H^+$ generally occur when the wind comes from the north (Fig. S12). Located in the north of the measurement container is a particle measurement cottage with

345 several CPCs inside using butanol. A previous study at this station also found that $C_4H_8H^+$ signals detected by PTR-TOF mainly come from butanol used by aerosol instruments (Schallhart et al., 2018). Therefore, it is expected that Factor S1 at the SMEAR II station is mainly contributed by butanol fragmentation inside the instrument where butanol comes from nearby aerosol instruments.

Factor S2: monoterpenes

350 A factor representing monoterpenes is also identified at the SMEAR II station, with fingerprint peaks of $^{13}CC_5H_8H^+$, $C_7H_{10}H^+$, and $^{13}CC_9H_{16}H^+$. Monoterpenes undergo some degree of fragmentation within PTR instruments, and $C_6H_8H^+$ and $C_7H_{10}H^+$ have been observed to be the major fragments of monoterpenes (Tani et al., 2003; Tani, 2013; Kari et al., 2018). The signal intensity of monoterpenes at the SMEAR II station is much lower than that in the Landes forest.

Factor S3: C_6-C_9 lightly oxygenated compounds

355   The mass profile of Factor S3 is characterized by lightly oxygenated compounds with carbon atoms varying from six to nine (C_6-C_9) such as $C_6H_{10}OH^+$, $C_6H_{12}OH^+$, $C_7H_{10}OH^+$, $C_8H_{14}OH^+$, and $C_9H_{18}OH^+$. The signal intensity of this factor shows high peaks at night and low appearance during daytime. As discussed in Section 4.2.1, these lightly oxygenated molecules can be directly emitted from anthropogenic and biogenic sources or come from oxidation processes of various VOC precursors (Conley et al., 2005; Pandya et al., 2006; Rantala et al., 2015; Hartikainen et al., 2018). For instance, $C_7H_{10}O$ has been found

360   from direct soil emissions (Abis et al., 2020) or oxidation processes of 1,2,4-trimethyl benzene (Mehra et al., 2020). Therefore, we expect the molecules in this factor to be either directly emitted or as oxidation products of forest emissions.

Factor S4: isoprene and its oxidation products

       At the SMEAR II station, a factor largely composed of isoprene and its oxidation products is also resolved. The outstanding peaks in the factor profile are $C_5H_8H^+$, $C_4H_6OH^+$, $C_4H_8O_2H^+$, and $C_5H_8O_2H^+$. The signal intensity of this factor is

365   around ten times lower than that of Factor L5 measured in the Landes forest. Similar to previous isoprene observations at the sampling site (Hakola et al., 2012), this factor shows a broad daytime peak and low signals at night.

Factor S5: monoterpene lightly oxidized products

       Similar to Factor L7 identified in the Landes forest, this factor is characterized by major peaks of monoterpene lightly oxidized compounds, as shown in Fig. 8. The signal intensity of this factor starts to increase at 20:00 and presents an obvious

370   morning peak at 7:00.

4.3.2 High mass range (201-320 Th)

Figure 11 and Figure 12 present the mass spectra, time series, and daily variations of the four factors identified in the higher mass range at the SMEAR II station. The correlation coefficients among each factor and various fingerprint compounds can be found in Fig. 10.

375   *Factor S6: sesquiterpene lightly oxidized products*

       This factor is identified as sesquiterpene lightly oxidized compounds with high peaks of $C_{14}H_{22}OH^+$, $C_{14}H_{24}OH^+$, $C_{15}H_{22}OH^+$, and $C_{15}H_{24}OH^+$, similar to Factor L10 in the Landes forest. The time series of this factor show strong correlations with the lightly oxidized products of sesquiterpenes (Fig. 10; $r^2 > 0.88$).

Factor S7: sesquiterpenes

380   Similar to Factor L12 in the Landes forest, this factor is characterized by the big peak of $C_{15}H_{24}H^+$, demonstrating the dominance of sesquiterpenes in the factor. Figure 10 shows that this factor closely correlates with monoterpenes and sesquiterpenes, with r^2 being 0.73 and 0.85, respectively. Compared with the identification of Factor L12, representing sesquiterpenes in the Landes forest, the signal intensity of this factor at the SMEAR II station is approximately three times lower. Including the lower signals of monoterpenes and isoprene, the results indicate weaker biogenic VOC emissions in the

385   Hyytiälä boreal forest than in the Landes forest.

Factor S8: monoterpene more oxidized products including organic nitrates

Factor S8 is mainly composed of more oxidized compounds, particularly from monoterpene oxidation processes, including monoterpene-derived organic nitrates. The major peaks are shown in Fig. 11. Mixed with monoterpene-derived organic nitrates, this factor of more oxidized compounds displays a small morning peak at 8:00 and generally high signals during daytime (Fig. 12). Utilizing non-negative matrix factorization analysis on iodide-adduct CIMS data at the SMEAR II station, Lee et al. (2018) found that the gas-phase organic species subgroup of $C_{6-10}H_yO_{\geq 7}$ showed distinct daytime diel trends. Yan et al. (2016) conducted source apportionment of HOMs at the SMEAR II station and separated various HOM formation pathways, such as monoterpene ozonolysis and monoterpene oxidation initiated by NO_3 radical. Unfortunately, due to the similar time series of monoterpene more oxidized compounds and monoterpene-derived organic nitrates, these different formation mechanisms cannot be separated in this study. For example, the time series of $C_{10}H_{15}NO_5H^+$ correlate well with those of $C_{10}H_{16}O_4H^+$ and $C_{10}H_{16}O_5H^+$ ($r^2 > 0.61$).

Factor S9: unknown source

The marker peaks of Factor S9 are mainly high cyclic volatile methyl siloxanes (VMSs) and unidentified compounds (Fig. 11), i.e., protonated D3 siloxane, D4 siloxane, and their H_3O^+ cluster ions. In addition to cosmetics and personal care products, siloxanes can also be emitted by silicone oils (Schweigkofler et al., 1999), which have been widely used in instrument pumps (Gonvers et al., 1985). In this study, the temporal behaviors of Factor S9 are contributed by high background signals and present a very regular diurnal cycle with higher signal intensities during daytime and lower ones at night, which basically follow the variations in ambient temperature. Therefore, we speculate that Factor S9 is mainly caused by emissions from silicone oil pumps used by several instruments in the container, and these emissions are influenced by daily temperature changes.

4.4 Comparison between the two forests

To give an overview of the source distributions in the two forest ecosystems, we calculated the mass fraction of each factor based on their average signal intensities. We acknowledge that it is not a perfect method to quantify the contributions of various sources and formation processes. The sensitivities of different VOCs measured by the PTR instruments may vary by a factor of 2-3 (Sekimoto et al., 2017; Yuan et al., 2017). The uncertainties can come from the challenge to convert the signal intensity to atmospheric concentrations because of problematic calibrations, especially given that many unknown molecules exist in the mass spectra. The major bins at *m/z* 81 Th and 137 Th, which were initially excluded to perform PMF analysis, were counted into their corresponding factors. For example, the signals of the discarded bins at *m/z* 81 Th and 137 Th were estimated by multiplying their isotope signals by the corresponding scale number and added to the factor representing monoterpenes. The average mass fractions of various PMF factors in total measured organic vapors are shown in Fig. 13.

While the atmospheric environment and ecosystem processes differ markedly in the Landes forest and the southern Finnish boreal forest, the results of this study reveal similar biogenic sources and oxidation processes in these forest environments. For instance, the biogenic VOCs at the two sites are both dominated by monoterpenes, with the average fractions of 29% in the Landes forest and at the SMEAR II station. These two forests are both characterized by pine trees, with dominant

420   emissions of α-pinene and β-pinene (Riba et al., 1987; Simon et al., 1994; Hellén et al., 2018). According to the PMF results, isoprene and its major oxidation products in these environments (mainly C_4H_6O) contribute 14% and 21% in the two ecosystems, respectively. Factors indicative of sesquiterpenes are identified in the high mass range at both sites. The average contribution of sesquiterpenes is much smaller than that of monoterpenes and isoprene. Figure 14 presents the comparison between the factor profiles of common sources identified at both sites, where the fractions of different bins in the mass spectra of the factors are plotted. As shown in Fig. 14c, d, and f, the mass spectra of the factors indicative of monoterpenes, isoprene

425   and its oxidation products, and sesquiterpenes, match quite well between the two forests, particularly for the dominant mass bins in the factor profile.

Terpenes undergo varying degrees of oxidations in the atmosphere and produce a large variety of organic compounds with different volatilities (Donahue et al., 2012; Ehn et al., 2014). With the sub-range PMF analysis performed in this study,

430   terpene reaction products with varying oxidation degrees are successfully separated. The sources of monoterpene lightly oxidized products, sesquiterpene lightly oxidized products, more oxidized compounds, and monoterpene-derived organic nitrates are identified in both forests with distinct characteristics. For the lightly oxygenated compounds from monoterpene or sesquiterpene reactions, they present similar temporal behaviors at the two sites, with a small morning peak and increased signal intensities at night. The mass spectra of these factors show high similarities in the two forests (Fig. 14e and g).

435   Monoterpene-derived organic nitrates are mainly characterized by a distinct morning peak at 9:00. With the active photochemical processes during daytime, more oxidized reaction products have a broad high distribution throughout the day. At the SMEAR II station, more oxidized compounds are mixed together with monoterpene organic nitrates and resolved into a single factor. Therefore, the mass spectra comparison between the two forests are more scattered for the factors of monoterpene-derived organic nitrates and more oxidized compounds. Overall, these common sources and their similar

440   characteristics indicate the similar atmospheric chemical processes in the two forest ecosystems.

5 Concluding remarks

In this study, we conducted Vocus PTR-TOF measurements in two forest environments and performed binPMF analysis on these complex mass spectra. In addition to VOC species, Vocus PTR-TOF is able to measure large amounts of oxygenated VOCs with enhanced detection efficiency. According to the results in this work, factor analysis on Vocus PTR-TOF mass

445   spectra separated VOC precursors and their reaction products with varying oxidation degrees into different factors. These factors showed distinct characteristics in the atmosphere. Comparatively, the conventional PTR instruments or gas chromatograph-mass spectrometry (GC-MS) largely detect VOC precursors of low-mass molecules (Dewulf et al., 2002; de Gouw et al., 2007). Previous source apportionment studies on these datasets mainly identified primary biogenic and anthropogenic emission sources (Vlasenko et al., 2009; Patokoski et al., 2014; Baudic et al., 2016; Debevec et al., 2017; Sarkar

450   et al., 2017; Wang et al., 2020). Recently, factorization methods have been applied on NO_3^- CIMS dataset to identify various atmospheric formation pathways of HOMs (Yan et al., 2016; Massoli et al., 2018; Zhang et al., 2019b). Here, for the first time,

source apportionment of Vocus PTR-TOF data identified various primary emission sources and secondary formation pathways of atmospheric organic vapors, highlighting the novelty of Vocus PTR-TOF in measuring both VOCs and oxygenated VOCs and providing new perspectives to understand gas-phase chemical processes.

455   Compared with VOC species, VOC reaction products are generally present in much smaller amounts in the atmosphere. Therefore, utilizing a sub-range PMF analysis, or other similarly weighting method, is particularly important for Vocus PTR-TOF observations, where several orders of magnitude differences are expected between VOC precursors and their oxidation products. Compared with the low mass range, the average contributions of the high mass range in total signals are significantly smaller, 2% and 9%, in the Landes forest and at the SMEAR II station, respectively. However, the identified sources in the

460   high mass range, such as sesquiterpenes, sesquiterpene lightly oxidized products, monoterpene-derived organic nitrates, and more oxidized compounds, provide crucial insights into atmospheric physicochemical processes.

     To summarize, this study successfully performed binPMF analysis on sub-ranges of mass spectrometry dataset acquired with a Vocus PTR-TOF in two European forest ecosystems, the Landes forest and a southern Finnish boreal forest. Similar sources and formation pathways of organic vapors were identified in the two environments, particularly for terpenes and their

465   reaction products with varying oxidation degrees (including organic nitrates). With the broad coverage of various organic vapors measured by Vocus PTR-TOF, this study provides a more comprehensive picture of gas-phase source identifications in the European forest ecosystems, covering both primary emissions and secondary oxidation processes.

Data Availability. The time series of the measured trace gases, meteorological parameters, and the concentrations of isoprene and monoterpenes in the Landes forest and at the SMEAR II station are available from https://doi.org/10.5281/zenodo.3946644

470   (Li, 2020).

Author contributions. HL and ME conceived the study. HL, MR, ST, LH, PMF, EV, and EP conducted the field measurements. HL carried out the data analysis. MC, YZ, ME, and FB participated in the discussions on data analysis. HL wrote the paper with inputs from all coauthors.

Competing interests. Manjula R. Canagaratna and Douglas Worsnop both work for Aerodyne Research Inc.

475   *Acknowledgements.* This work was supported by the H2020 European Research Council (grant nos. ATM-GP (742206), COALA (638703), and CHAPAs (850614)) and the Academy of Finland (grant nos. 317380 and 320094). We thank the SMEAR II station staff for their help during field measurements in Hyytiälä. The authors also would like to thank the PRIME-QUAL program for financial support (ADEME, convention#1662C0024) and the French National Research Agency (ANR) in the frame of the "Investments for the Future" program, within the Cluster of Excellence COTE (ANR-10-LABX-45) of the

480 University of Bordeaux for financial support. Harald Stark and Donna T. Sueper from Aerodyne Research Inc. are acknowledged for helpful discussions.

References

Abis, L., Loubet, B., Ciuraru, R., Lafouge, F., Houot, S., Nowak, V., Tripied, J., Dequiedt, S., Maron, P. A., and Sadet-Bourgeteau, S.: Reduced microbial diversity induces larger volatile organic compound emissions from soils, Scientific
485   Reports, 10, 6104, 10.1038/s41598-020-63091-8, 2020.

Ahlberg, E., Falk, J., Eriksson, A., Holst, T., Brune, W. H., Kristensson, A., Roldin, P., and Svenningsson, B.: Secondary organic aerosol from VOC mixtures in an oxidation flow reactor, Atmos Environ, 161, 210-220, https://doi.org/10.1016/j.atmosenv.2017.05.005, 2017.

Baudic, A., Gros, V., Sauvage, S., Locoge, N., Sanchez, O., Sarda-Estève, R., Kalogridis, C., Petit, J.-E., Bonnaire, N., Baisnée,
490   D., Favez, O., Albinet, A., Sciare, J., and Bonsang, B.: Seasonal variability and source apportionment of volatile organic compounds (VOCs) in the Paris megacity (France), Atmos. Chem. Phys., 16, 11961–11989, https://doi.org/10.5194/acp-16-11961-2016, 2016.

Bianchi, F., Tröstl, J., Junninen, H., Frege, C., Henne, S., Hoyle, C. R., Molteni, U., Herrmann, E., Adamov, A., Bukowiecki, N., Chen, X., Duplissy, J., Gysel, M., Hutterli, M., Kangasluoma, J., Kontkanen, J., Kürten, A., Manninen, H. E., Münch,
495   S., Peräkylä, O., Petäjä, T., Rondo, L., Williamson, C., Weingartner, E., Curtius, J., Worsnop, D. R., Kulmala, M., Dommen, J., and Baltensperger, U.: New particle formation in the free troposphere: A question of chemistry and timing, Science, 352, 1109-1112, 10.1126/science.aad5456, 2016.

Bourtsoukidis, E., Williams, J., Kesselmeier, J., Jacobi, S., and Bonn, B.: From emissions to ambient mixing ratios: online seasonal field measurements of volatile organic compounds over a Norway spruce-dominated forest in central Germany,
500   Atmos. Chem. Phys., 14, 6495–6510, https://doi.org/10.5194/acp-14-6495-2014, 2014.

Boyd, C. M., Sanchez, J., Xu, L., Eugene, A. J., Nah, T., Tuet, W. Y., Guzman, M. I., and Ng, N. L.: Secondary organic aerosol formation from the β-pinene+NO3 system: effect of humidity and peroxy radical fate, Atmos. Chem. Phys., 15, 7497–7522, https://doi.org/10.5194/acp-15-7497-2015, 2015.

Bsaibes, S., Al Ajami, M., Mermet, K., Truong, F., Batut, S., Hecquet, C., Dusanter, S., Léornadis, T., Sauvage, S., Kammer,
505   J., Flaud, P.-M., Perraudin, E., Villenave, E., Locoge, N., Gros, V., and Schoemaecker, C.: Variability of hydroxyl radical (OH) reactivity in the Landes maritime pine forest: results from the LANDEX campaign 2017, Atmos. Chem. Phys., 20, 1277–1300, https://doi.org/10.5194/acp-20-1277-2020, 2020.

Buser, A. M., Kierkegaard, A., Bogdal, C., MacLeod, M., Scheringer, M., and Hungerbühler, K.: Concentrations in Ambient Air and Emissions of Cyclic Volatile Methylsiloxanes in Zurich, Switzerland, Environmental Science & Technology, 47,
510   7045-7051, 10.1021/es3046586, 2013.

Calogirou, A., Larsen, B. R., and Kotzias, D.: Gas-phase terpene oxidation products: a review, Atmos Environ, 33, 1423-1439, https://doi.org/10.1016/S1352-2310(98)00277-5, 1999.

Cheong, J.-J. and Choi, Y. D.: Methyl jasmonate as a vital substance in plants, Trends Genet., 19, 409–413, doi:10.1016/S0168-9525(03)00138-0, 2003.

515 Conley, F. L., Thomas, R. L., and Wilson, B. L.: Measurement of Volatile Organic Compounds in the Urban Atmosphere of Harris County, Texas, Journal of Environmental Science and Health, Part A, 40, 1689-1699, 10.1081/ESE-200067996, 2005.

de Gouw, J., and Warneke, C.: Measurements of volatile organic compounds in the earth's atmosphere using proton-transfer-reaction mass spectrometry, Mass Spectrometry Reviews, 26, 223-257, 10.1002/mas.20119, 2007.

520 Debevec, C., Sauvage, S., Gros, V., Sciare, J., Pikridas, M., Stavroulas, I., Salameh, T., Leonardis, T., Gaudion, V., Depelchin, L., Fronval, I., Sarda-Esteve, R., Baisnée, D., Bonsang, B., Savvides, C., Vrekoussis, M., and Locoge, N.: Origin and variability in volatile organic compounds observed at an Eastern Mediterranean background site (Cyprus), Atmos. Chem. Phys., 17, 11355–11388, https://doi.org/10.5194/acp-17-11355-2017, 2017.

Dewulf, J., Van Langenhove, H., and Wittmann, G.: Analysis of volatile organic compounds using gas chromatography, TrAC

525 Trends in Analytical Chemistry, 21, 637-646, https://doi.org/10.1016/S0165-9936(02)00804-X, 2002.

Donahue, N. M., Kroll, J. H., Pandis, S. N., and Robinson, A. L.: A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution, Atmos. Chem. Phys., 12, 615–634, https://doi.org/10.5194/acp-12-615-2012, 2012.

Duhl, T. R., Helmig, D., and Guenther, A.: Sesquiterpene emissions from vegetation: a review, Biogeosciences, 5, 761–777, https://doi.org/10.5194/bg-5-761-2008, 2008.

530 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, 10.1038/nature13032, 2014.

535 Friedrich, R., and Obermeier, A.: Chapter 1 - Anthropogenic Emissions of Volatile Organic Compounds, in: Reactive Hydrocarbons in the Atmosphere, edited by: Hewitt, C. N., Academic Press, San Diego, 1-39, 1999.

Fu, P., Kawamura, K., Chen, J., and Barrie, L. A.: Isoprene, Monoterpene, and Sesquiterpene Oxidation Products in the High Arctic Aerosols during Late Winter to Early Summer, Environmental Science & Technology, 43, 4022-4028, 10.1021/es803669a, 2009.

540 Gonvers, M.: A New Silicone Oil Pump, American Journal of Ophthalmology, 99, 210, https://doi.org/10.1016/0002-9394(85)90236-3, 1985.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., McKay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., and Zimmerman, P.: A global model of natural volatile organic compound emissions, J. Geophys. Res.-Atmos., 100, 8873–8892, https://doi.org/10.1029/94JD02950, 1995.

545 Hakola, H., Tarvainen, V., Bäck, J., Ranta, H., Bonn, B., Rinne, J., and Kulmala, M.: Seasonal variation of mono- and sesquiterpene emission rates of Scots pine, Biogeosciences, 3, 93–101, https://doi.org/10.5194/bg-3-93-2006, 2006.

Hakola, H., Hellén, H., Hemmilä, M., Rinne, J., and Kulmala, M.: In situ measurements of volatile organic compounds in a boreal forest, Atmos. Chem. Phys., 12, 11665–11678, https://doi.org/10.5194/acp-12-11665-2012, 2012.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George,
550 C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, Th. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9, 5155–5236, https://doi.org/10.5194/acp-9-5155-2009, 2009.

Hari, P. and Kulmala, M.: Station for measuring ecosystem-atmosphere relations, Boreal Environ. Res., 10, 315–322, 2005.

555 Hartikainen, A., Yli-Pirilä, P., Tiitta, P., Leskinen, A., Kortelainen, M., Orasche, J., Schnelle-Kreis, J., Lehtinen, K. E. J., Zimmermann, R., Jokiniemi, J., and Sippula, O.: Volatile Organic Compounds from Logwood Combustion: Emissions and Transformation under Dark and Photochemical Aging Conditions in a Smog Chamber, Environmental Science & Technology, 52, 4979-4988, 10.1021/acs.est.7b06269, 2018.

Hellén, H., Hakola, H., Pystynen, K.-H., Rinne, J., and Haapanala, S.: C2-C10 hydrocarbon emissions from a boreal wetland
560 and forest floor, Biogeosciences, 3, 167–174, https://doi.org/10.5194/bg-3-167-2006, 2006.

Hellén, H., Praplan, A. P., Tykkä, T., Ylivinkka, I., Vakkari, V., Bäck, J., Petäjä, T., Kulmala, M., and Hakola, H.: Long-term measurements of volatile organic compounds highlight the importance of sesquiterpenes for the atmospheric chemistry of a boreal forest, Atmos. Chem. Phys., 18, 13839-13863, 10.5194/acp-18-13839-2018, 2018.

Kallio, M., Jussila, M., Rissanen, T., Anttila, P., Hartonen, K., Reissell, A., Vreuls, R., Adahchour, M., and Hyötyläinen, T.:
565 Comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry in the identification of organic compounds in atmospheric aerosols from coniferous forest, Journal of Chromatography A, 1125, 234-243, https://doi.org/10.1016/j.chroma.2006.05.050, 2006.

Kännaste, A., Copolovici, L., and Niinemets, Ü.: Gas Chromatography–Mass Spectrometry Method for Determination of Biogenic Volatile Organic Compounds Emitted by Plants, in: Plant Isoprenoids: Methods and Protocols, edited by:
570 Rodríguez-Concepción, M., Springer New York, New York, NY, 161-169, 2014.

Kammer, J., Perraudin, E., Flaud, P. M., Lamaud, E., Bonnefond, J. M., and Villenave, E.: Observation of nighttime new particle formation over the French Landes forest, Sci. Total Environ., 621, 1084–1092, https://doi.org/10.1016/j.scitotenv.2017.10.118, 2018.

Kammer, J., Flaud, P. M., Chazeaubeny, A., Ciuraru, R., Le Menach, K., Geneste, E., Budzinski, H., Bonnefond, J. M.,
575 Lamaud, E., Perraudin, E., and Villenave, E.: Biogenic volatile organic compounds (BVOCs) reactivity related to new particle formation (NPF) over the Landes forest, Atmospheric Research, 237, 104869, https://doi.org/10.1016/j.atmosres.2020.104869, 2020.

Kari, E., Miettinen, P., Yli-Pirilä, P., Virtanen, A., and Faiola, C. L.: PTR-ToF-MS product ion distributions and humidity-dependence of biogenic volatile organic compounds, International Journal of Mass Spectrometry, 430, 87-97, https://doi.org/10.1016/j.ijms.2018.05.003, 2018.

Kaser, L., Karl, T., Guenther, A., Graus, M., Schnitzhofer, R., Turnipseed, A., Fischer, L., Harley, P., Madronich, M., Gochis, D., Keutsch, F. N., and Hansel, A.: Undisturbed and disturbed above canopy ponderosa pine emissions: PTR-TOF-MS measurements and MEGAN 2.1 model results, Atmos. Chem. Phys., 13, 11935–11947, https://doi.org/10.5194/acp-13-11935-2013, 2013.

Kesselmeier, J., and Staudt, M.: Biogenic Volatile Organic Compounds (VOC): An Overview on Emission, Physiology and Ecology, J Atmos Chem, 33, 23-88, 10.1023/A:1006127516791, 1999.

Khan, M. A. H., Jenkin, M. E., Foulds, A., Derwent, R. G., Percival, C. J., and Shallcross, D. E.: A modeling study of secondary organic aerosol formation from sesquiterpenes using the STOCHEM global chemistry and transport model, Journal of Geophysical Research: Atmospheres, 122, 4426-4439, 10.1002/2016JD026415, 2017.

Kim, S., Guenther, A., Karl, T., and Greenberg, J.: Contributions of primary and secondary biogenic VOC tototal OH reactivity during the CABINEX (Community Atmosphere-Biosphere INteractions Experiments)-09 field campaign, Atmos. Chem. Phys., 11, 8613–8623, https://doi.org/10.5194/acp-11-8613-2011, 2011.

Kirkby, J., Duplissy, J., Sengupta, K., Frege, C., Gordon, H., Williamson, C., Heinritzi, M., Simon, M., Yan, C., Almeida, J., Tröstl, J., Nieminen, T., Ortega, I. K., Wagner, R., Adamov, A., Amorim, A., Bernhammer, A.-K., Bianchi, F., Breitenlechner, M., Brilke, S., Chen, X., Craven, J., Dias, A., Ehrhart, S., Flagan, R. C., Franchin, A., Fuchs, C., Guida, R., Hakala, J., Hoyle, C. R., Jokinen, T., Junninen, H., Kangasluoma, J., Kim, J., Krapf, M., Kürten, A., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Mathot, S., Molteni, U., Onnela, A., Peräkylä, O., Piel, F., Petäjä, T., Praplan, A. P., Pringle, K., Rap, A., Richards, N. A. D., Riipinen, I., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Scott, C. E., Seinfeld, J. H., Sipilä, M., Steiner, G., Stozhkov, Y., Stratmann, F., Tomé, A., Virtanen, A., Vogel, A. L., Wagner, A. C., Wagner, P. E., Weingartner, E., Wimmer, D., Winkler, P. M., Ye, P., Zhang, X., Hansel, A., Dommen, J., Donahue, N. M., Worsnop, D. R., Baltensperger, U., Kulmala, M., Carslaw, K. S., and Curtius, J.: Ion-induced nucleation of pure biogenic particles, Nature, 533, 521-526, 10.1038/nature17953, 2016.

Krechmer, J., Lopez-Hilfiker, F., Koss, A., Hutterli, M., Stoermer, C., Deming, B., Kimmel, J., Warneke, C., Holzinger, R., Jayne, J., Worsnop, D., Fuhrer, K., Gonin, M., and de Gouw, J.: Evaluation of a New Reagent-Ion Source and Focusing Ion–Molecule Reactor for Use in Proton-Transfer-Reaction Mass Spectrometry, Analytical Chemistry, 90, 12011-12018, 10.1021/acs.analchem.8b02641, 2018.

Kundu, S., Fisseha, R., Putman, A. L., Rahn, T. A., and Mazzoleni, L. R.: Molecular formula composition of β-caryophyllene ozonolysis SOA formed in humid and dry conditions, Atmos Environ, 154, 70-81, https://doi.org/10.1016/j.atmosenv.2016.12.031, 2017.

610 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., 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,
615   Atmos. Chem. Phys., 18, 11547–11562, https://doi.org/10.5194/acp-18-11547-2018, 2018.

Li, H., Riva, M., Rantala, P., Heikkinen, L., Daellenbach, K., Krechmer, J. E., Flaud, P.-M., Worsnop, D., Kulmala, M., Villenave, E., Perraudin, E., Ehn, M., and Bianchi, F.: Terpenes and their oxidation products in the French Landes forest: insights from Vocus PTR-TOF measurements, Atmos. Chem. Phys., 20, 1941–1959, https://doi.org/10.5194/acp-20-1941-2020, 2020.

620 Li, H.: Data for "Source identification of atmospheric organic vapors in two European pine forests: Results from Vocus PTR-TOF observations", http://doi.org/10.5281/zenodo.3946644, 2020.

Maria, S. F., Russell, L. M., Gilles, M. K., and Myneni, S. C. B.: Organic Aerosol Growth Mechanisms and Their Climate-Forcing Implications, Science, 306, 1921-1924, 10.1126/science.1103491, 2004.

Massoli, P., Stark, H., Canagaratna, M. R., Krechmer, J. E., Xu, L., Ng, N. L., Mauldin, R. L., Yan, C., Kimmel, J., Misztal,
625   P. K., Jimenez, J. L., Jayne, J. T., and Worsnop, D. R.: Ambient Measurements of Highly Oxidized Gas-Phase Molecules during the Southern Oxidant and Aerosol Study (SOAS) 2013, ACS Earth and Space Chemistry, 2, 653-672, 10.1021/acsearthspacechem.8b00028, 2018.

McFiggans, G., Mentel, T. F., Wildt, J., Pullinen, I., Kang, S., Kleist, E., Schmitt, S., Springer, M., Tillmann, R., Wu, C., Zhao, D., Hallquist, M., Faxon, C., Le Breton, M., Hallquist, Å. M., Simpson, D., Bergström, R., Jenkin, M. E., Ehn, M.,
630   Thornton, J. A., Alfarra, M. R., Bannan, T. J., Percival, C. J., Priestley, M., Topping, D., and Kiendler-Scharr, A.: Secondary organic aerosol reduced by mixture of atmospheric vapours, Nature, 565, 587-593, 10.1038/s41586-018-0871-y, 2019.

McKinney, K. A., Lee, B. H., Vasta, A., Pho, T. V., and Munger, J. W.: Emissions of isoprenoids and oxygenated biogenic volatile organic compounds from a New England mixed forest, Atmos. Chem. Phys., 11, 4807–4831,
635   https://doi.org/10.5194/acp-11-4807-2011, 2011.

Mehra, A., Wang, Y., Krechmer, J. E., Lambe, A., Majluf, F., Morris, M. A., Priestley, M., Bannan, T. J., Bryant, D. J., Pereira, K. L., Hamilton, J. F., Rickard, A. R., Newland, M. J., Stark, H., Croteau, P., Jayne, J. T., Worsnop, D. R., Canagaratna, M. R., Wang, L., and Coe, H.: Evaluation of the Chemical Composition of Gas and Particle Phase Products of Aromatic Oxidation, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-161, in review, 2020.

640 Monson, R. K. and Fall, R.: Isoprene Emission from Aspen Leaves, Influence of Environment and Relation to Photosynthesis and Photorespiration, J. Plant. Physiol., 90, 267–274, https://doi.org/10.1104/pp.90.1.267, 1989.

Noe, S. M., Hüve, K., Niinemets, Ü., and Copolovici, L.: Seasonal variation in vertical volatile compounds air concentrations within a remote hemiboreal mixed forest, Atmos. Chem. Phys., 12, 3909–3926, https://doi.org/10.5194/acp-12-3909-2012, 2012.

645 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, 10.1002/env.3170050203, 1994.

Pagonis, D., Sekimoto, K., and de Gouw, J.: A Library of Proton-Transfer Reactions of H3O+ Ions Used for Trace Gas Detection, Journal of The American Society for Mass Spectrometry, 30, 1330-1335, 10.1007/s13361-019-02209-3, 2019.

Pandya, G. H., Gavane, A. G., Bhanarkar, A. D., and Kondawar, V. K.: Concentrations of volatile organic compounds (VOCs)
650 at an oil refinery, International Journal of Environmental Studies, 63, 337-351, 10.1080/00207230500241918, 2006.

Patokoski, J., Ruuskanen, T. M., Hellén, H., Taipale, R., Grönholm, T., Kajos, M. K., Petäjä, T., Hakola, H., Kulmala, M., and Rinne, J.: Winter to spring transition and diurnal variation of VOCs in Finland at an urban background site and a rural site, Boreal Environ. Res., 19, 79–103, 2014.

Perraud, V., Bruns, E. A., Ezell, M. J., Johnson, S. N., Greaves, J., and Finlayson-Pitts, B. J.: Identification of Organic Nitrates
655 in the NO3 Radical Initiated Oxidation of α-Pinene by Atmospheric Pressure Chemical Ionization Mass Spectrometry, Environmental Science & Technology, 44, 5887-5893, 10.1021/es1005658, 2010.

Rantala, P., Aalto, J., Taipale, R., Ruuskanen, T. M., and Rinne, J.: Annual cycle of volatile organic compound exchange between a boreal pine forest and the atmosphere, Biogeosciences, 12, 5753–5770, https://doi.org/10.5194/bg-12-5753-2015, 2015.

660 Riba, M. L., Tathy, J. P., Tsiropoulos, N., Monsarrat, B., and Torres, L.: Diurnal variation in the concentration of α- and β-pinene in the landes forest (France), Atmospheric Environment (1967), 21, 191-193, https://doi.org/10.1016/0004-6981(87)90285-X, 1987.

Riva, M., Rantala, P., Krechmer, J. E., Peräkylä, O., Zhang, Y., Heikkinen, L., Garmash, O., Yan, C., Kulmala, M., Worsnop, D., and Ehn, M.: Evaluating the performance of five different chemical ionization techniques for detecting gaseous
665 oxygenated organic species, Atmos. Meas. Tech., 12, 2403–2421, https://doi.org/10.5194/amt-12-2403-2019, 2019.

Sakulyanontvittaya, T., Duhl, T., Wiedinmyer, C., Helmig, D., Matsunaga, S., Potosnak, M., Milford, J., and Guenther, A.: Monoterpene and Sesquiterpene Emission Estimates for the United States, Environmental Science & Technology, 42, 1623-1629, 10.1021/es702274e, 2008.

Sarkar, C., Sinha, V., Sinha, B., Panday, A. K., Rupakheti, M., and Lawrence, M. G.: Source apportionment of NMVOCs in
670 the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization, Atmos. Chem. Phys., 17, 8129–8156, https://doi.org/10.5194/acp-17-8129-2017, 2017.

Sato, K., Takami, A., Kato, Y., Seta, T., Fujitani, Y., Hikida, T., Shimono, A., and Imamura, T.: AMS and LC/MS analyses of SOA from the photooxidation of benzene and 1,3,5-trimethylbenzene in the presence of NOx: effects of chemical structure on SOA aging, Atmos. Chem. Phys., 12, 4667–4682, https://doi.org/10.5194/acp-12-4667-2012, 2012.

675 Schallhart, S., Rantala, P., Kajos, M. K., Aalto, J., Mammarella, I., Ruuskanen, T. M., and Kulmala, M.: Temporal variation of VOC fluxes measured with PTR-TOF above a boreal forest, Atmos Chem Phys, 18, 815-832, 2018.

Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., and Ebel, A.: Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, Journal of Geophysical Research: Atmospheres, 106, 28275-28293, 10.1029/2001JD000384, 2001.

680 Schweigkofler, M., and Niessner, R.: Determination of Siloxanes and VOC in Landfill Gas and Sewage Gas by Canister Sampling and GC-MS/AES Analysis, Environmental Science & Technology, 33, 3680-3685, 10.1021/es9902569, 1999.

Sekimoto, K., Li, S.-M., Yuan, B., Koss, A., Coggon, M., Warneke, C., and de Gouw, J.: Calculation of the sensitivity of proton-transfer-reaction mass spectrometry (PTR-MS) for organic trace gases using molecular properties, International Journal of Mass Spectrometry, 421, 71-94, https://doi.org/10.1016/j.ijms.2017.04.006, 2017.

685 Simon, V., Clement, B., Riba, M.-L., and Torres, L.: The Landes experiment: Monoterpenes emitted from the maritime pine, J. Geophys. Res., 99, 16501–16510, https://doi.org/10.1029/94JD00785, 1994.

Simon, M., Dada, L., Heinritzi, M., Scholz, W., Stolzenburg, D., Fischer, L., Wagner, A. C., Kürten, A., Rörup, B., He, X.-C., Almeida, J., Baalbaki, R., Baccarini, A., Bauer, P. S., Beck, L., Bergen, A., Bianchi, F., Bräkling, S., Brilke, S., Caudillo, L., Chen, D., Chu, B., Dias, A., Draper, D. C., Duplissy, J., El Haddad, I., Finkenzeller, H., Frege, C., Gonzalez-
690 Carracedo, L., Gordon, H., Granzin, M., Hakala, J., Hofbauer, V., Hoyle, C. R., Kim, C., Kong, W., Lamkaddam, H., Lee, C. P., Lehtipalo, K., Leiminger, M., Mai, H., Manninen, H. E., Marie, G., Marten, R., Mentler, B., Molteni, U., Nichman, L., Nie, W., Ojdanic, A., Onnela, A., Partoll, E., Petäjä, T., Pfeifer, J., Philippov, M., Quéléver, L. L. J., Ranjithkumar, A., Rissanen, M., Schallhart, S., Schobesberger, S., Schuchmann, S., Shen, J., Sipilä, M., Steiner, G., Stozhkov, Y., Tauber, C., Tham, Y. J., Tomé, A. R., Vazquez-Pufleau, M., Vogel, A., Wagner, R., Wang, M., Wang, D.
695 S., Wang, Y., Weber, S. K., Wu, Y., Xiao, M., Yan, C., Ye, P., Ye, Q., Zauner-Wieczorek, M., Zhou, X., Baltensperger, U., Dommen, J., Flagan, R. C., Hansel, A., Kulmala, M., Volkamer, R., Winkler, P. M., Worsnop, D. R., Donahue, N. M., Kirkby, J., and Curtius, J.: Molecular understanding of new-particle formation from alpha-pinene between −50 °C and 25 °C, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-1058, in review, 2020.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and
700 Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317–9341, https://doi.org/10.5194/acp-14-9317-2014, 2014.

Spanel, P., and Smith, D.: SIFT studies of the reactions of H3O+, NO+ and O2+ with a series of alcohols, International Journal of Mass Spectrometry and Ion Processes, 167-168, 375-388, https://doi.org/10.1016/S0168-1176(97)00085-2, 1997.

Stark, H., Yatavelli, R. L. N., Thompson, S. L., Kimmel, J. R., Cubison, M. J., Chhabra, P. S., Canagaratna, M. R., Jayne, J.
705 T., Worsnop, D. R., and Jimenez, J. L.: Methods to extract molecular and bulk chemical information from series of complex mass spectra with limited mass resolution, Int. J. Mass Spectrom., 389, 26–38, https://doi.org/10.1016/j.ijms.2015.08.011, 2015.

Tani, A., Hayward, S., and Hewitt, C. N.: Measurement of monoterpenes and related compounds by proton transfer reaction-mass spectrometry (PTR-MS), International Journal of Mass Spectrometry, 223-224, 561-578, https://doi.org/10.1016/S1387-3806(02)00880-1, 2003.

Tani, A.: Fragmentation and Reaction Rate Constants of Terpenoids Determined by Proton Transfer Reaction-mass Spectrometry, Environmental Control in Biology, 51, 23-29, 10.2525/ecb.51.23, 2013.

Tsigaridis, K. and Kanakidou, M.: Global modelling of secondary organic aerosol in the troposphere: a sensitivity analysis, Atmos. Chem. Phys., 3, 1849–1869, https://doi.org/10.5194/acp-3-1849-2003, 2003.

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.

Valiev, R. R., Hasan, G., Salo, V.-T., Kubečka, J., and Kurten, T.: Intersystem Crossings Drive Atmospheric Gas-Phase Dimer Formation, The Journal of Physical Chemistry A, 123, 6596-6604, 10.1021/acs.jpca.9b02559, 2019.

Vlasenko, A., Slowik, J. G., Bottenheim, J. W., Brickell, P. C., Chang, R. Y. W., Macdonald, A. M., Shantz, N. C., Sjostedt, S. J., Wiebe, H. A., Leaitch, W. R., and Abbatt, J. P. D.: Measurements of VOCs by proton transfer reaction mass spectrometry at a rural Ontario site: Sources and correlation to aerosol composition, Journal of Geophysical Research: Atmospheres, 114, 10.1029/2009JD012025, 2009.

Wang, L., Slowik, J. G., Tripathi, N., Bhattu, D., Rai, P., Kumar, V., Vats, P., Satish, R., Baltensperger, U., Ganguly, D., Rastogi, N., Sahu, L. K., Tripathi, S. N., and Prévôt, A. S. H.: Source characterization of volatile organic compounds measured by PTR-ToF-MS in Delhi, India, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-11, in review, 2020.

Wängberg, I., Barnes, I., and Becker, K. H.: Product and Mechanistic Study of the Reaction of NO3 Radicals with α-Pinene, Environmental Science & Technology, 31, 2130-2135, 10.1021/es960958n, 1997.

Wennberg, P. O., Bates, K. H., Crounse, J. D., Dodson, L. G., McVay, R. C., Mertens, L. A., Nguyen, T. B., Praske, E., Schwantes, R. H., Smarte, M. D., St Clair, J. M., Teng, A. P., Zhang, X., and Seinfeld, J. H.: Gas-Phase Reactions of Isoprene and Its Major Oxidation Products, Chem Rev, 118, 3337-3390, 2018.

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, 10.5194/acp-16-12715-2016, 2016.

Yee, L. D., Isaacman-VanWertz, G., Wernis, R. A., Meng, M., Rivera, V., Kreisberg, N. M., Hering, S. V., Bering, M. S., Glasius, M., Upshur, M. A., Gray Bé, A., Thomson, R. J., Geiger, F. M., Offenberg, J. H., Lewandowski, M., Kourtchev, I, Kalberer, M., de Sá, S., Martin, S. T., Alexander, M. L., Palm, B. B., Hu, W., Campuzano-Jost, P., Day, D. A., Jimenez, J. L., Liu, Y., McKinney, K. A., Artaxo, P., Viegas, J., Manzi, A., Oliveira, M. B., de Souza, R., Machado, L. A. T.,

Longo, K., and Goldstein, A. H.: Observations of sesquiterpenes and their oxidation products in central Amazonia during the wet and dry seasons, Atmos. Chem. Phys., 18, 10433–10457, https://doi.org/10.5194/acp-18-10433-2018, 2018.

Yuan, B., Koss, A. R., Warneke, C., Coggon, M., Sekimoto, K., and de Gouw, J. A.: Proton-Transfer-Reaction Mass Spectrometry: Applications in Atmospheric Sciences, Chem Rev, 117, 13187-13229, 2017.

745 Yucuis, R. A., Stanier, C. O., and Hornbuckle, K. C.: Cyclic siloxanes in air, including identification of high levels in Chicago and distinct diurnal variation, Chemosphere, 92, 905-910, https://doi.org/10.1016/j.chemosphere.2013.02.051, 2013.

Zaytsev, A., Koss, A. R., Breitenlechner, M., Krechmer, J. E., Nihill, K. J., Lim, C. Y., Rowe, J. C., Cox, J. L., Moss, J., Roscioli, J. R., Canagaratna, M. R., Worsnop, D. R., Kroll, J. H., and Keutsch, F. N.: Mechanistic study of the formation of ring-retaining and ring-opening products from the oxidation of aromatic compounds under urban atmospheric 750 conditions, Atmos. Chem. Phys., 19, 15117–15129, https://doi.org/10.5194/acp-19-15117-2019, 2019.

Zhang, H., Yee, L. D., Lee, B. H., Curtis, M. P., Worton, D. R., Isaacman-VanWertz, G., Offenberg, J. H., Lewandowski, M., Kleindienst, T. E., Beaver, M. R., Holder, A. L., Lonneman, W. A., Docherty, K. S., Jaoui, M., Pye, H. O. T., Hu, W., Day, D. A., Campuzano-Jost, P., Jimenez, J. L., Guo, H., Weber, R. J., de Gouw, J., Koss, A. R., Edgerton, E. S., Brune, W., Mohr, C., Lopez-Hilfiker, F. D., Lutz, A., Kreisberg, N. M., Spielman, S. R., Hering, S. V., Wilson, K. R., Thornton, 755 J. A., and Goldstein, A. H.: Monoterpenes are the largest source of summertime organic aerosol in the southeastern United States, Proceedings of the National Academy of Sciences, 115, 2038-2043, 10.1073/pnas.1717513115, 2018.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, Anal Bioanal Chem, 401, 3045-3067, 10.1007/s00216-011-5355-y, 2011.

760 Zhang, X., Schwantes, R. H., Coggon, M. M., Loza, C. L., Schilling, K. A., Flagan, R. C., and Seinfeld, J. H.: Role of ozone in SOA formation from alkane photooxidation, Atmos. Chem. Phys., 14, 1733–1753, https://doi.org/10.5194/acp-14-1733-2014, 2014.

Zhang, Y., Peräkylä, O., Yan, C., Heikkinen, L., Äijälä, M., Daellenbach, K. R., Zha, Q., Riva, M., Garmash, O., Junninen, H., Paatero, P., Worsnop, D., and Ehn, M.: A novel approach for simple statistical analysis of high-resolution mass 765 spectra, Atmos. Meas. Tech., 12, 3761-3776, 10.5194/amt-12-3761-2019, 2019a.

Zhang, Y., Peräkylä, O., Yan, C., Heikkinen, L., Äijälä, M., Daellenbach, K. R., Zha, Q., Riva, M., Garmash, O., Junninen, H., Paatero, P., Worsnop, D., and Ehn, M.: Insights on Atmospheric Oxidation Processes by Performing Factor Analyses on Sub-ranges of Mass Spectra, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-838, in review, 2019b.

Zhu, Y., Yang, L., Kawamura, K., Chen, J., Ono, K., Wang, X., Xue, L., and Wang, W.: Contributions and source identification 770 of biogenic and anthropogenic hydrocarbons to secondary organic aerosols at Mt. Tai in 2014, Environ Pollut, 220, 863-872, https://doi.org/10.1016/j.envpol.2016.10.070, 2017.

Table 1. Summary of source identification results for the two forest sites (L, Landes; S, SMEAR II).

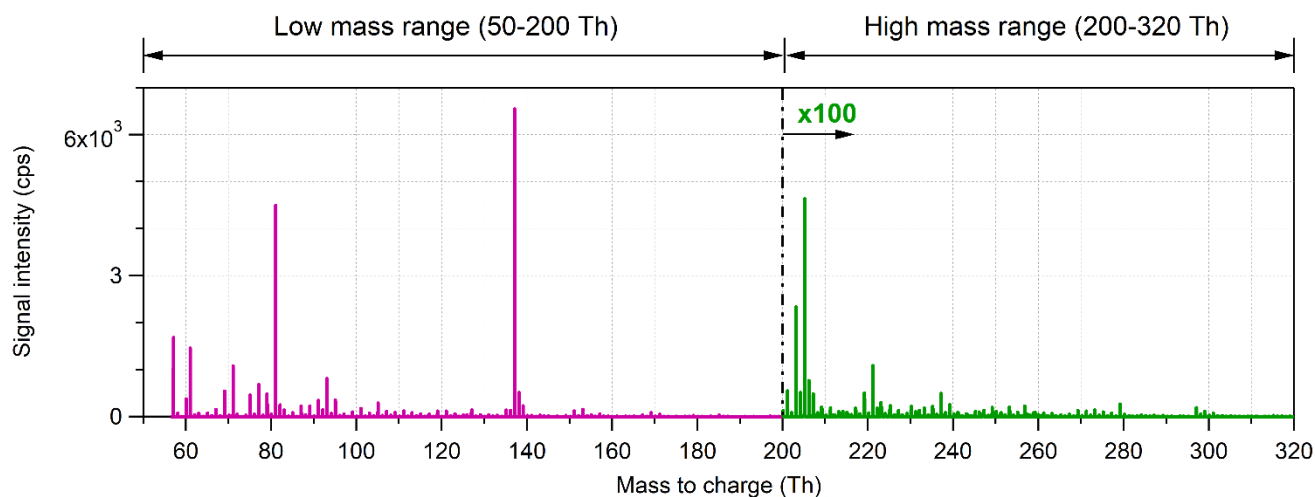
	Factor name	Possible source/chemistry	Fingerprint molecules
Landes forest			
	Factor L1	$C_4H_8H^+$ ion-related	$C_4H_8H^+$, $C_4H_{10}O_2H^+$
	Factor L2	A plume event	$C_6H_6H^+$, $C_7H_6H^+$, $C_6H_6OH^+$, unidentified peaks
	Factor L3	C_6 and C_7 lightly oxidized products	$C_6H_{10}OH^+$, $C_7H_{10}OH^+$, $C_6H_{10}O_2H^+$, $C_7H_{12}O_2H^+$,
	Factor L4	Monoterpenes	$C_6H_8H^+$, $C_7H_{10}H^+$, $C_{10}H_{16}H^+$
	Factor L5	Isoprene and its oxidation products	$C_5H_8H^+$, $C_4H_6OH^+$, $C_4H_6O_3H^+$
	Factor L6	Unknown source	$C_6H_4O_2H^+$, $C_6H_6O_3H^+$, unidentified peaks with negative mass defect
	Factor L7	Monoterpene lightly oxidized products	$C_9H_{14}OH^+$, $C_{10}H_{14}OH^+$, $C_{10}H_{16}O_2H^+$, $C_{10}H_{16}O_3H^+$
	Factor L8	C_{13} lightly oxidized products	$C_{13}H_{18}O_2H^+$, $C_{13}H_{20}O_3H^+$
	Factor L9	A plume event	Unidentified peaks
	Factor L10	Sesquiterpene lightly oxidized products	$C_{15}H_{22}OH^+$, $C_{15}H_{24}OH^+$, $C_{15}H_{22}O_2H^+$, $C_{15}H_{24}O_2H^+$, $C_{15}H_{24}O_3H^+$
	Factor L11	Monoterpene more oxidized products	$C_{10}H_{16}O_4H^+$, $C_{10}H_{14}O_5H^+$, $C_{10}H_{16}O_5H^+$, $C_{10}H_{16}O_6H^+$
	Factor L12	Sesquiterpenes	$C_{15}H_{24}H^+$
	Factor L13	Monoterpene-derived organic nitrates	$C_{10}H_{15}NO_4H^+$, $C_{10}H_{15}NO_5H^+$, $C_9H_{13}NO_6H^+$, $C_{10}H_{15}NO_6H^+$
	Factor L14	C_{12}, C_{14} or C_{16} lightly oxidized products	$C_{12}H_{26}O_3H^+$, $C_{14}H_{26}O_2H^+$, $C_{16}H_{30}O_2H^+$
	Factor L15	Unknown source	D3 siloxane, D4 siloxane, unidentified peaks
SMEAR II			
	Factor S1	$C_4H_8H^+$ ion-related	$C_4H_8H^+$, $C_4H_{12}O_2H^+$, $C_4H_{14}O_3H^+$
	Factor S2	Monoterpenes	$C_6H_8H^+$, $C_7H_{10}H^+$, $C_{10}H_{16}H^+$
	Factor S3	C_6-C_9 lightly oxygenated compounds	$C_6H_{10}OH^+$, $C_6H_{12}OH^+$, $C_7H_{10}OH^+$, $C_8H_{14}OH^+$, $C_9H_{12}O_2H^+$
	Factor S4	Isoprene and its oxidation products	$C_5H_8H^+$, $C_4H_6OH^+$
	Factor S5	Monoterpene lightly oxidized products	$C_9H_{14}OH^+$, $C_{10}H_{14}OH^+$, $C_{10}H_{16}OH^+$, $C_{10}H_{16}O_2H^+$, $C_{10}H_{16}O_3H^+$
	Factor S6	Sesquiterpene lightly oxidized products	$C_{14}H_{22}OH^+$, $C_{14}H_{24}OH^+$, $C_{15}H_{22}OH^+$, $C_{15}H_{24}OH^+$
	Factor S7	Sesquiterpenes	$C_{15}H_{24}H^+$
	Factor S8	Monoterpene more oxidized products including organic nitrates	$C_{10}H_{16}O_4H^+$, $C_{14}H_{22}O_3H^+$, $C_{15}H_{24}O_3H^+$, $C_{10}H_{17}NO_3H^+$, $C_9H_{13}NO_6H^+$
	Factor S9	Unknown source	D3 siloxane, D4 siloxane, unidentified peaks

775

Figure 1. Average mass spectrum measured in the Landes forest. The mass spectrum is divided into two sub-ranges for further source identification analysis. The intensity scale is shown 100-fold for the high mass range (201-320 Th).

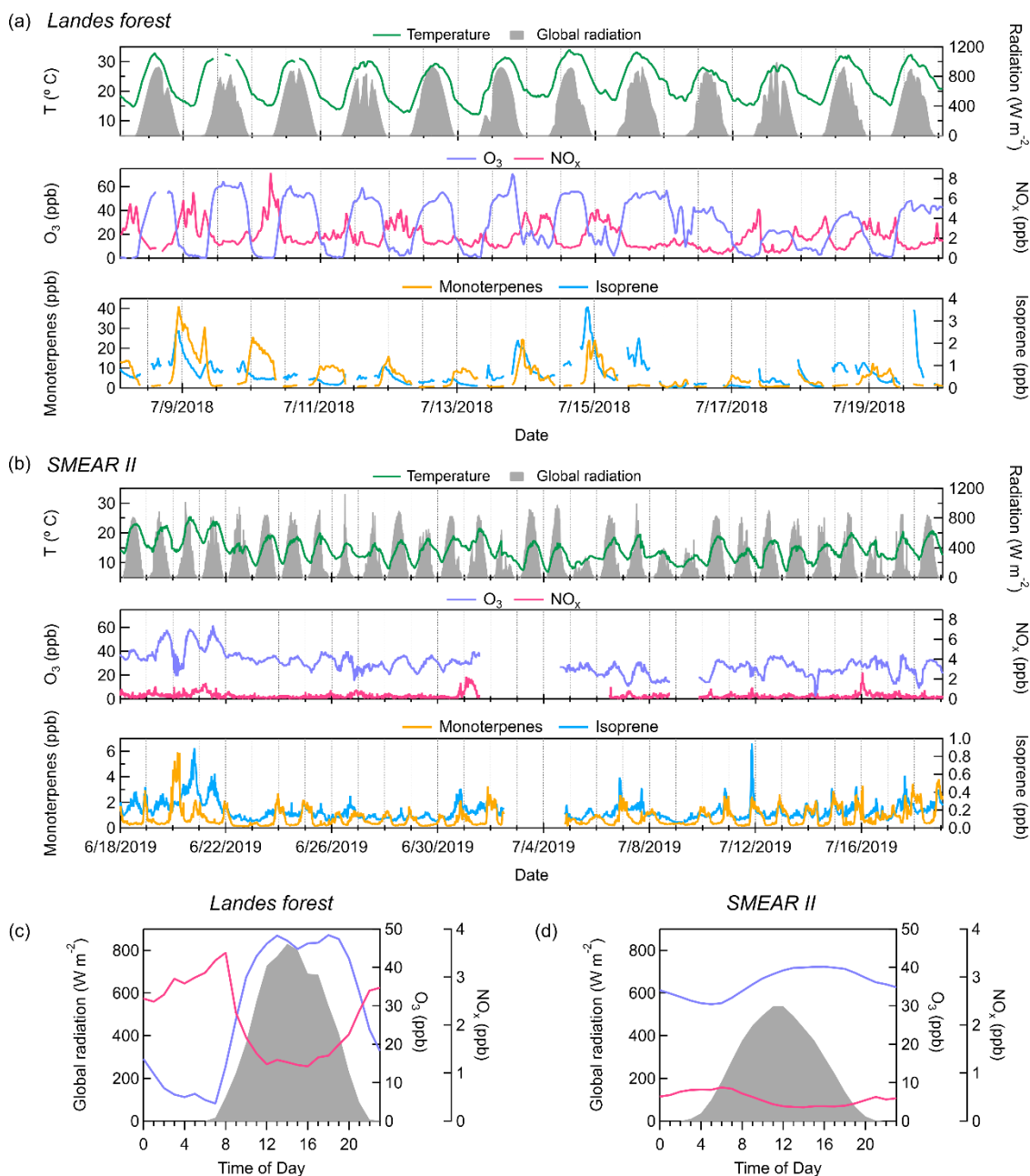



Figure 2. Time series of temperature, global radiation, concentrations of O₃ and NOₓ, and mixing ratios of monoterpenes and isoprene throughout the measurements (a) in the Landes forest and (b) at the SMEAR II station. Average diurnal cycles of global radiation, O₃ concentration (in blue), and NOₓ concentration (in pink) (c) in the Landes forest and (d) at the SMEAR II station. All parameters, except monoterpenes and isoprene, are shown in the same y-axis scale for the two sites. Monoterpene and isoprene concentrations are much lower at the SMEAR II station than in the Landes forest.

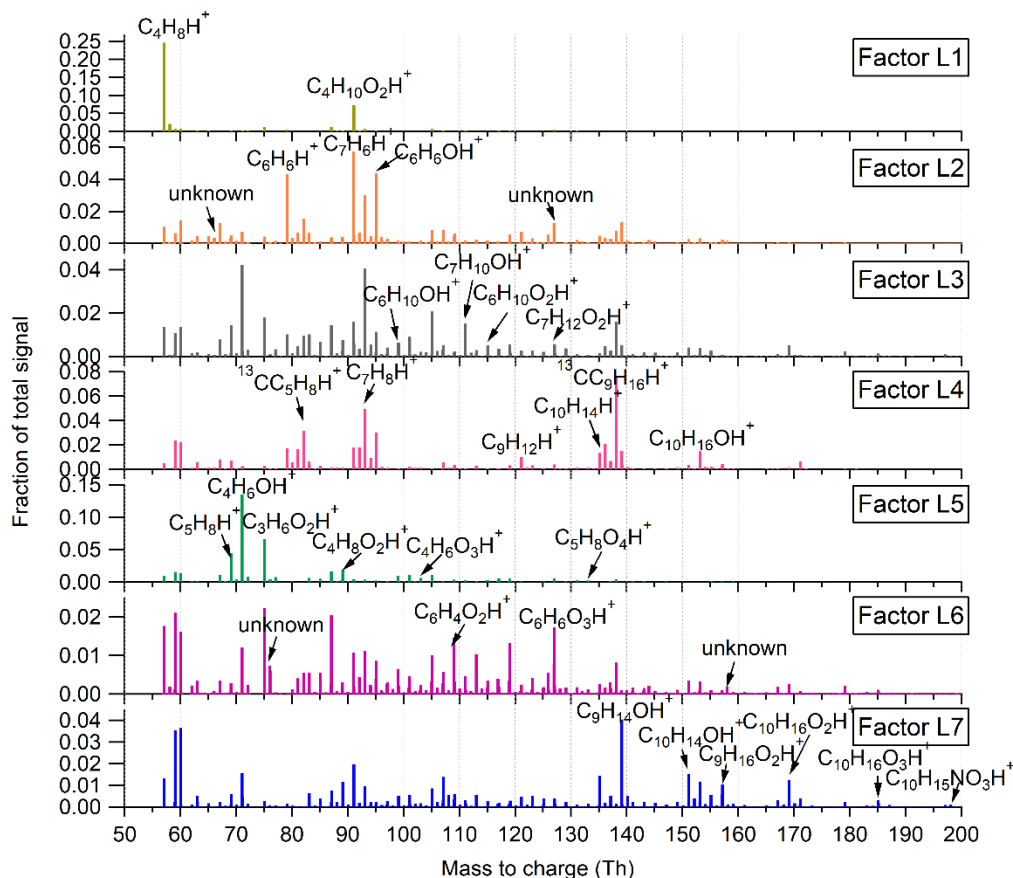



Figure 3. Mass profiles of the seven factors resolved in the low mass range in the Landes forest. Fingerprint peaks identified by high-resolution peak fitting are shown in the mass spectra.

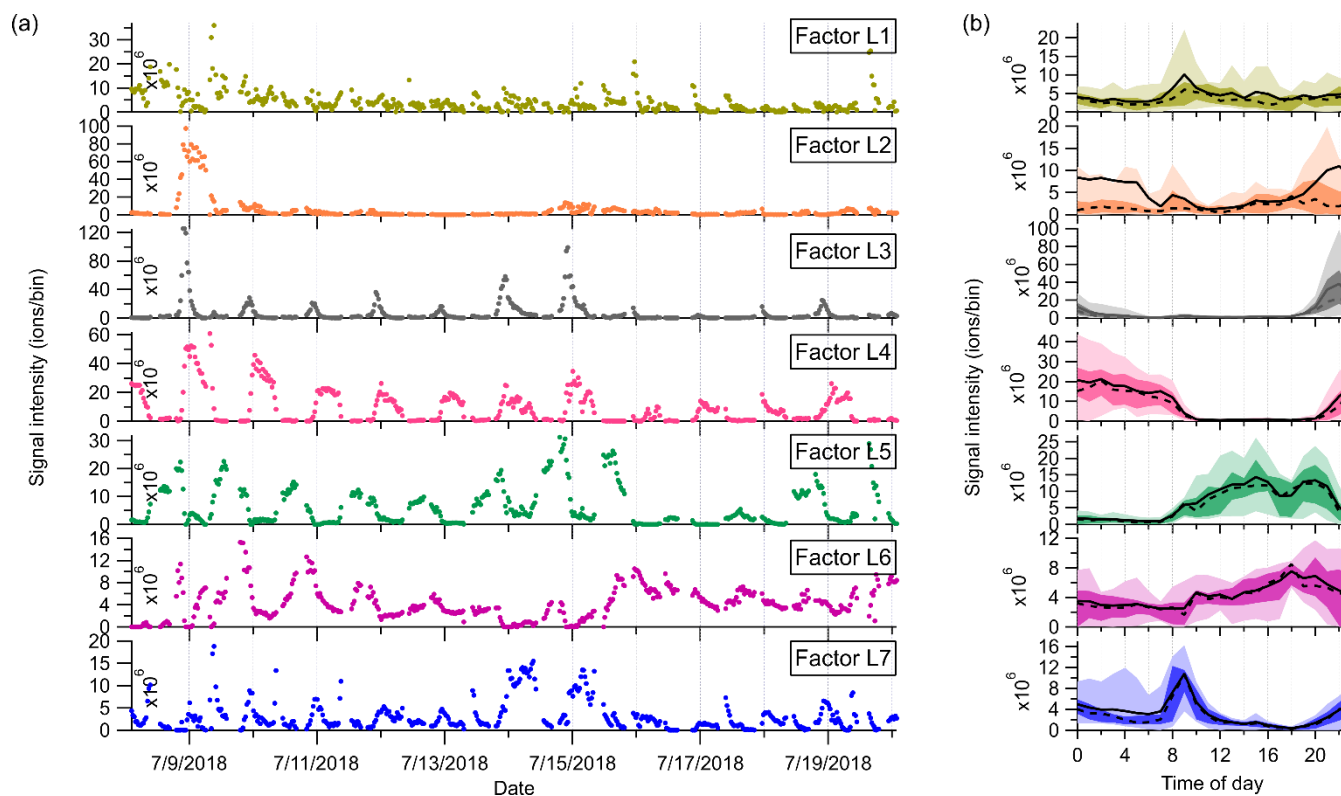



Figure 4. (a) Time series and **(b)** diurnal variations of the seven factors identified in the low mass range in the Landes forest. The solid and dashed lines in the diurnal plots show the mean and median values, respectively, and the shaded area shows 10th, 25th, 75th, and 90th percentiles.

795

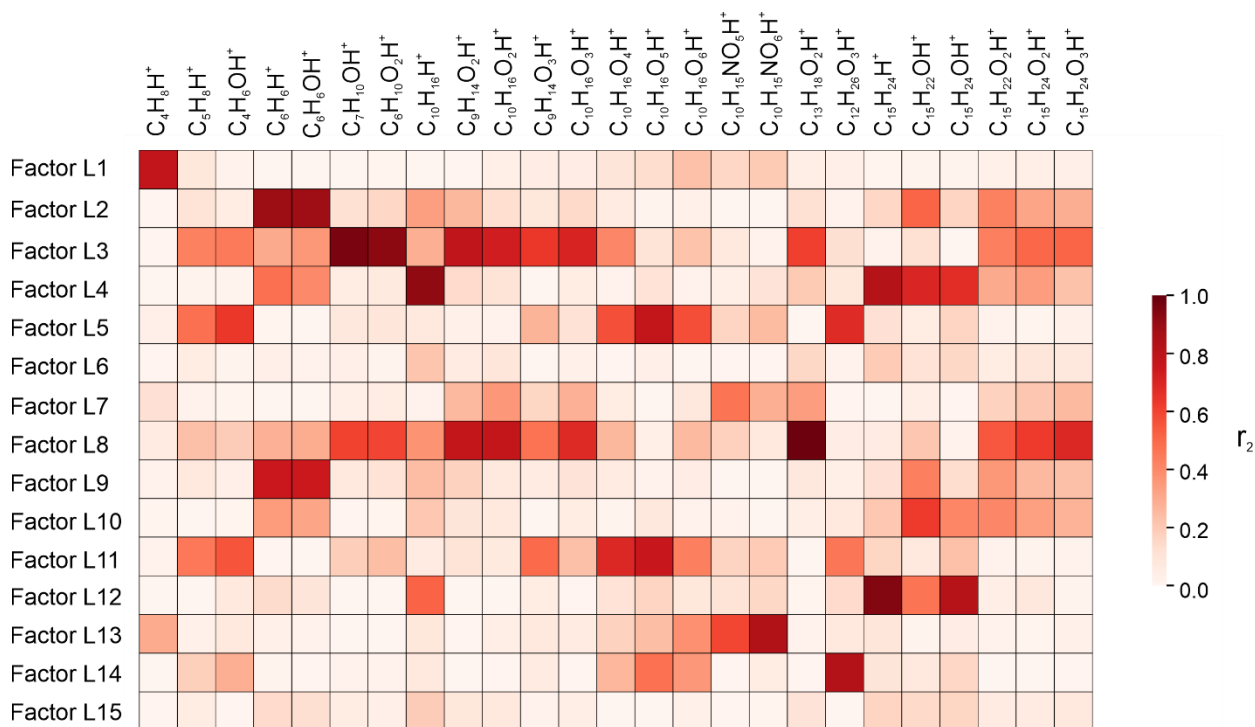



Figure 5. Correlations between PMF factors and marker molecules in the Landes forest, with the color representing the correlation coefficients (r^2).

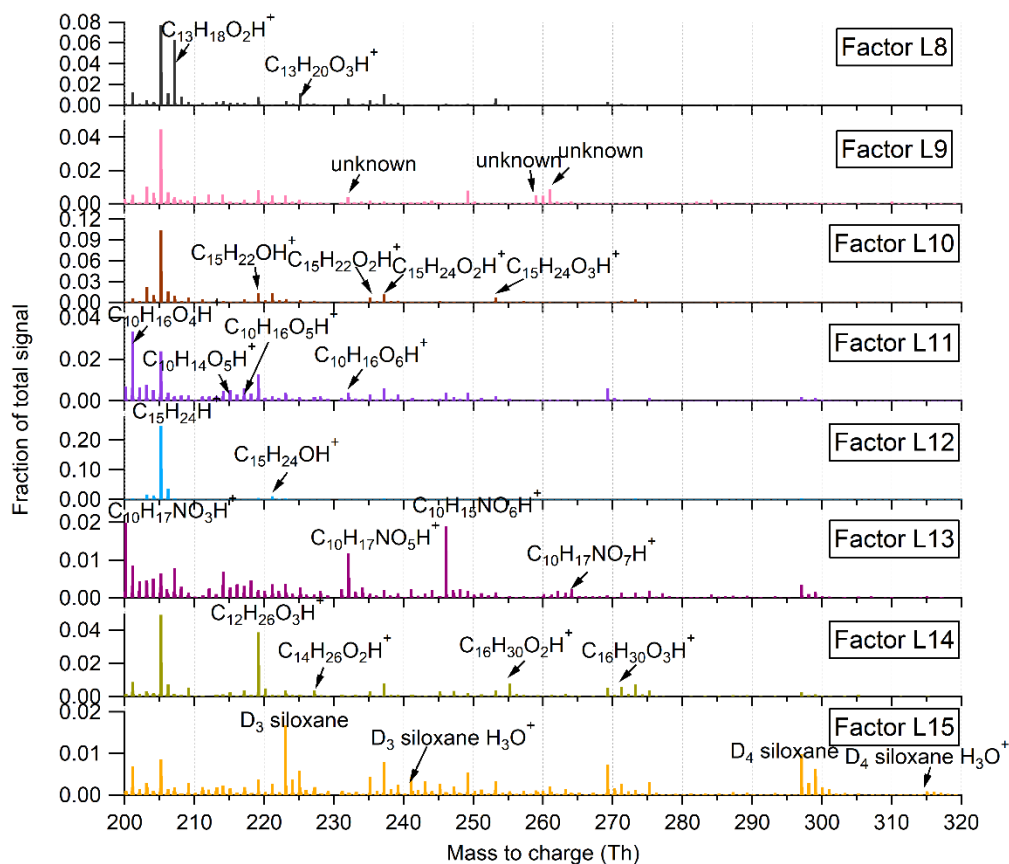



Figure 6. Mass profiles of the eight factors resolved in the high mass range in the Landes forest, with major fingerprint peaks labeled in the mass spectra.

800

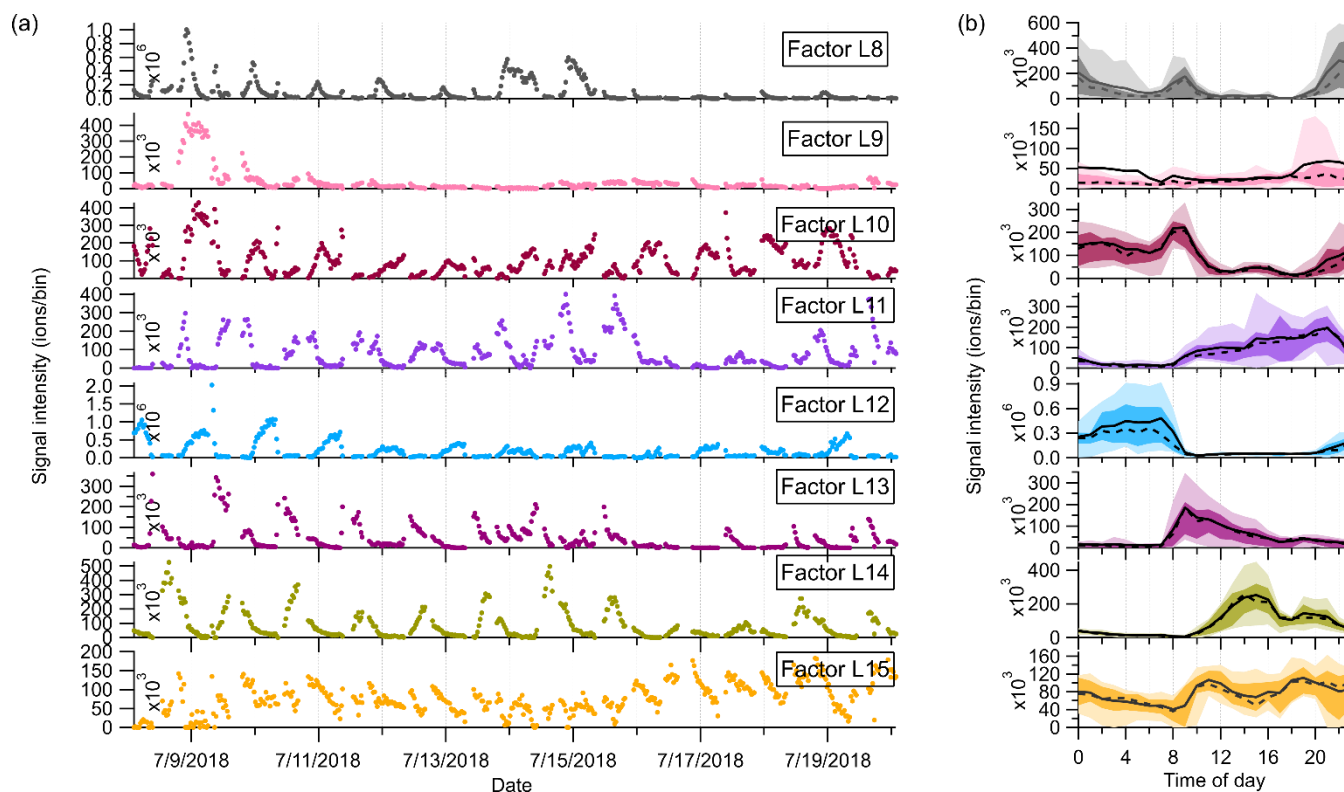



Figure 7. (a) Time series and (b) diurnal trends of the eight factors resolved in the high mass range in the Landes forest. The solid and dashed lines in the diurnal plots show the mean and median values, respectively, and the shaded area shows 10th, 25th, 75th, and 90th percentiles.

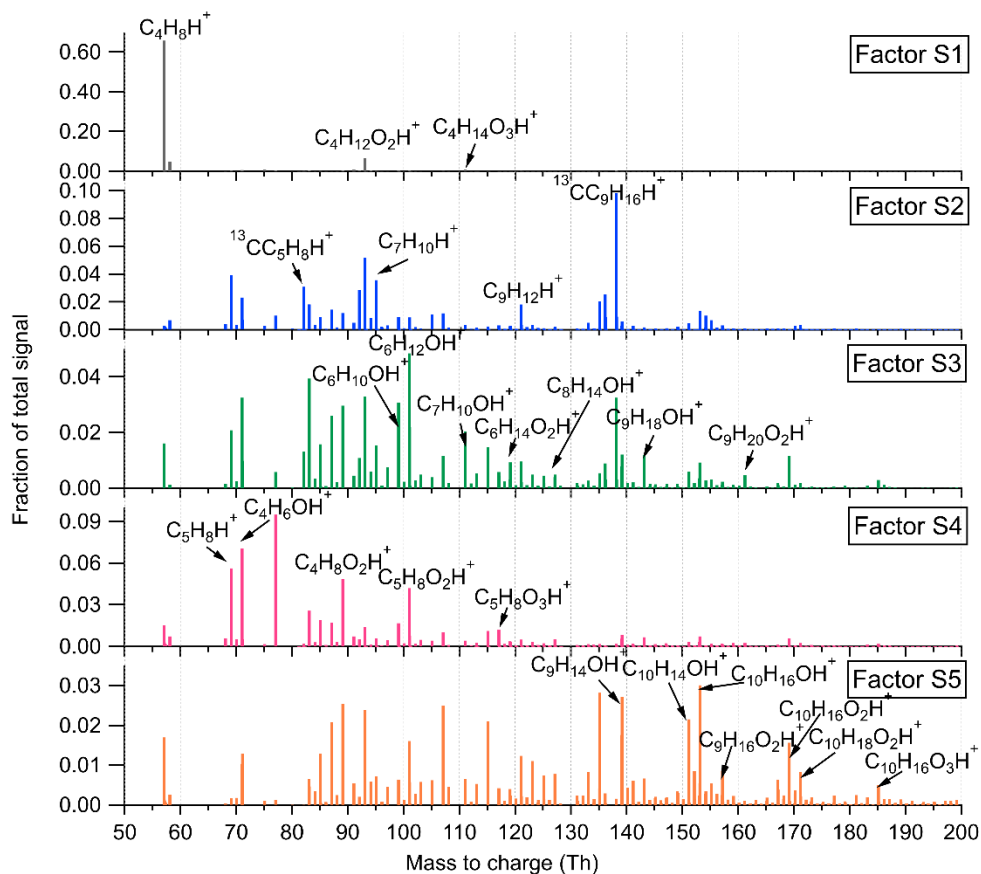



Figure 8. Mass profiles of the five factors identified in the low mass range at the SMEAR II station, with fingerprint peaks shown in the mass spectra.

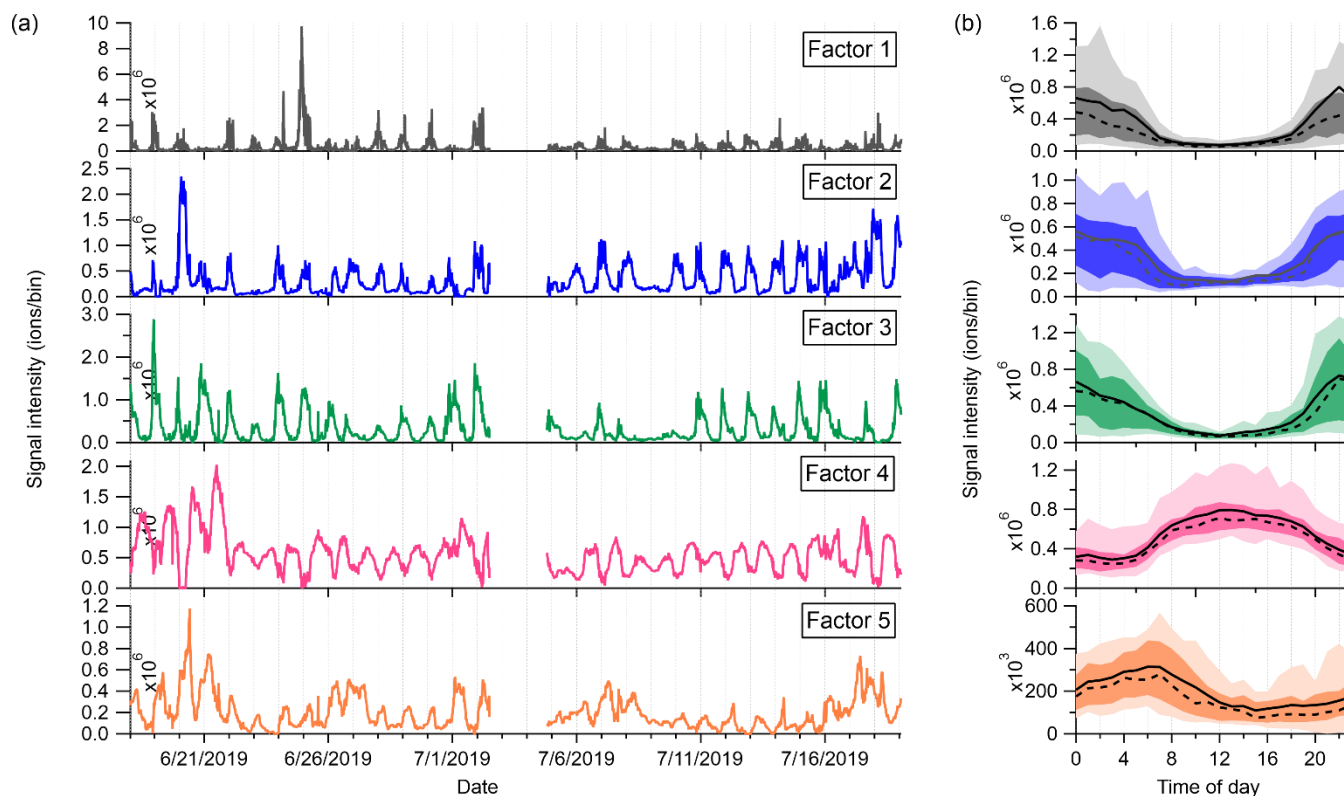



Figure 9. (a) Time series and (b) diurnal cycles of the five factors in the low mass range at the SMEAR II station. The solid and dashed lines in the diurnal plots show the mean and median values, respectively, and the shaded area shows 10th, 25th, 75th, and 90th percentiles.

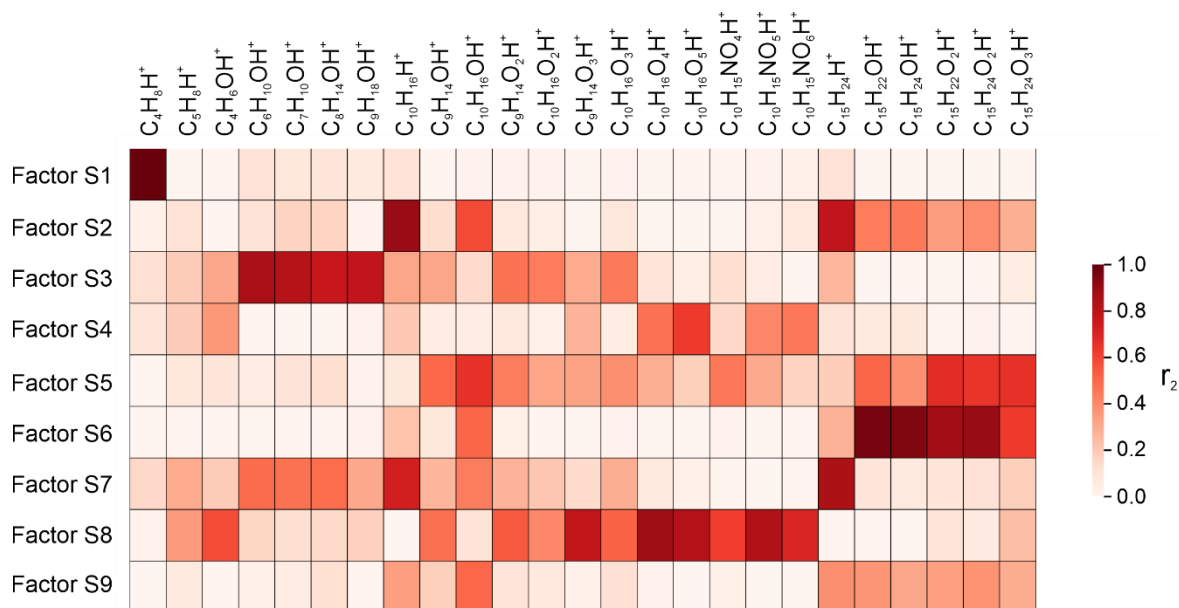



Figure 10. Correlations between PMF factors and marker molecules at the SMEAR II station, with the color indicating the correlation coefficients (r^2).

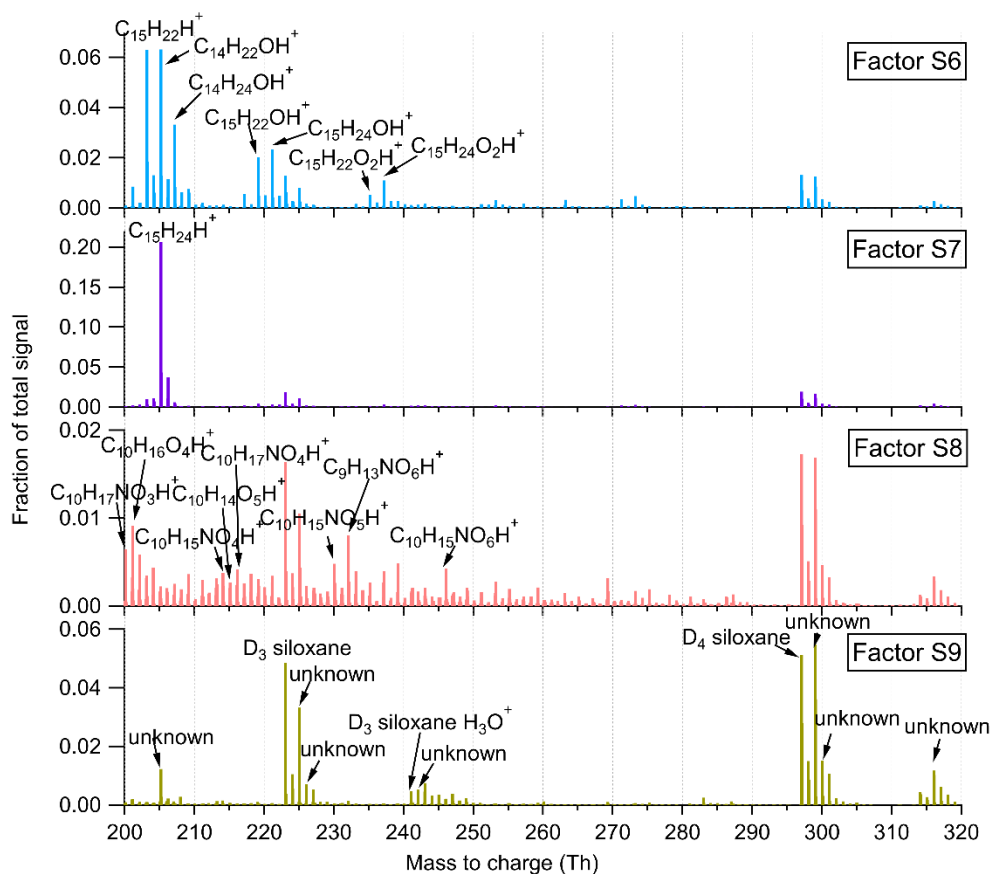



Figure 11. Mass profiles of the four factors resolved in the high mass range at the SMEAR II station. The fingerprint peaks are labeled in the mass spectra.

825

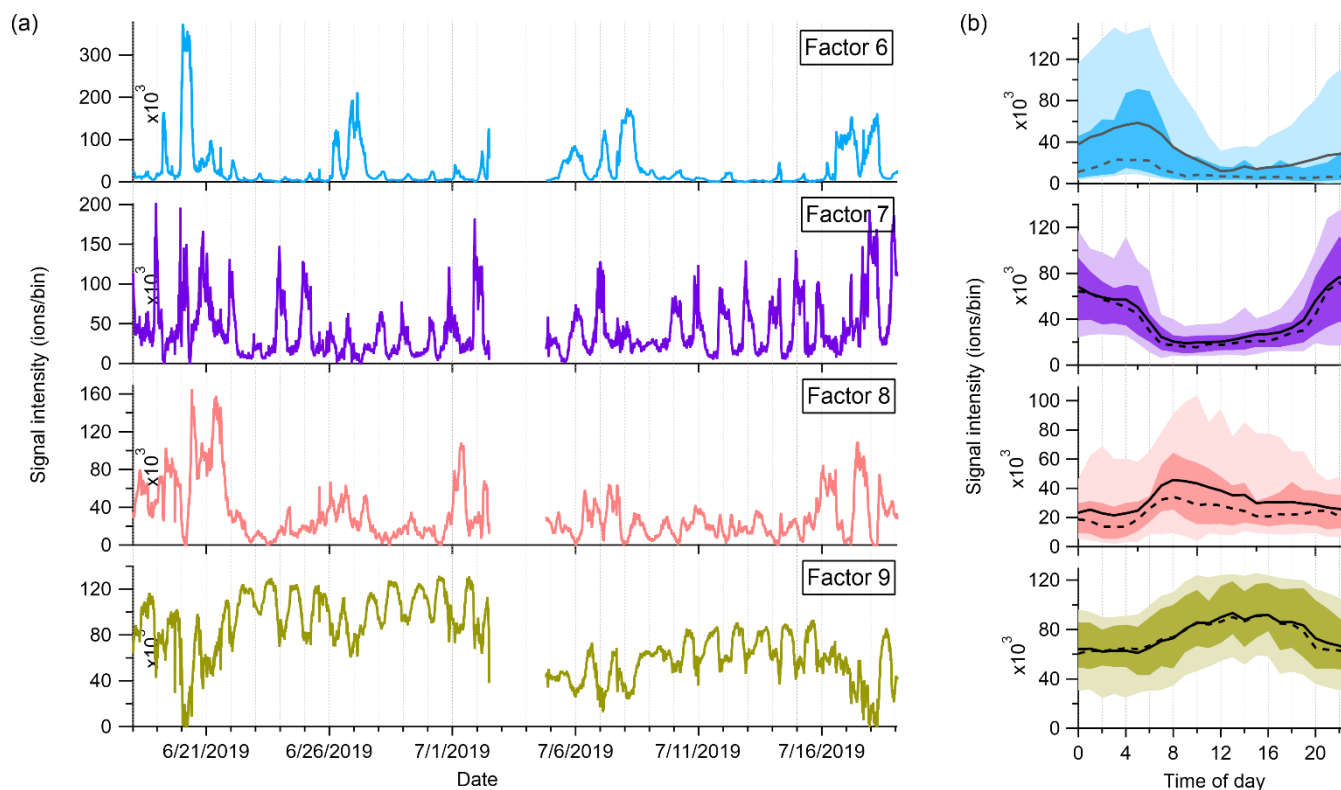



Figure 12. (a) Time series and (b) daily trends of the four factors in the high mass range at the SMEAR II station. The solid and dashed lines in the diurnal plots show the mean and median values, respectively, and the shaded area shows 10th, 25th, 75th, and 90th percentiles.

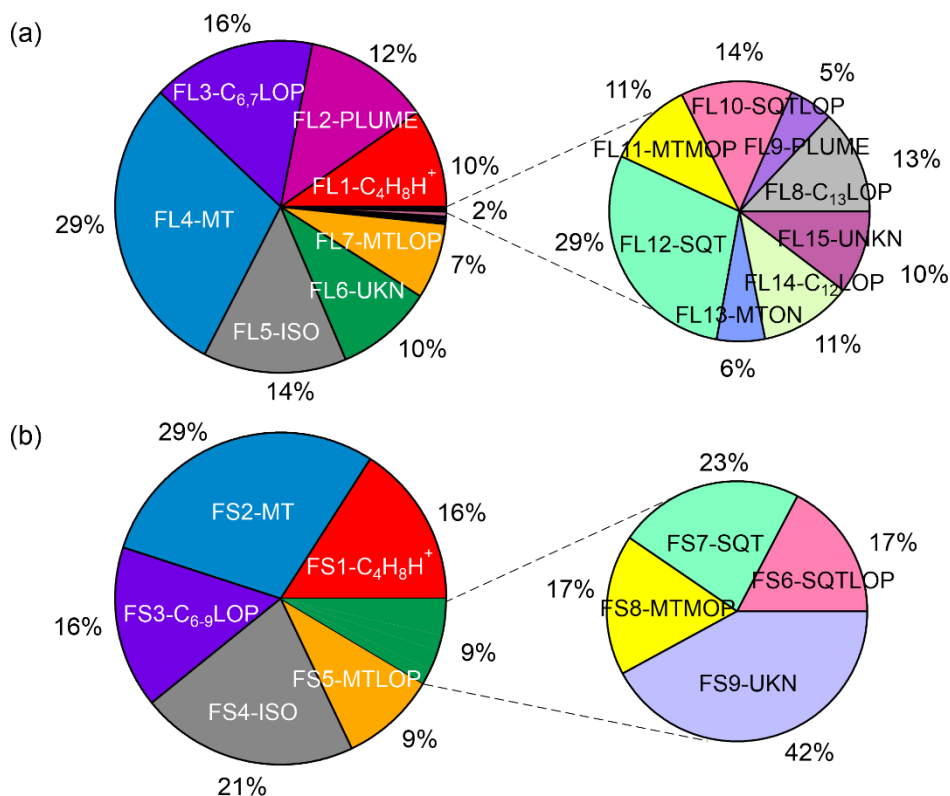



Figure 13. Average mass contributions of various identified factors in total measured organic vapors (a) in the Landes forest and (b) at the SMEAR II station. The common sources apportioned at both sites are presented in the same color in (a) and (b). FL: factors in the Landes forest; FS: factors at the SMEAR II station; $C_4H_8H^+$, $C_4H_8H^+$ ion-related; PLUME: a plume event; $C_{6,7}$LOP: C_6 and C_7 lightly oxidized products; C_{6-9}LOP: C_6-C_9 lightly oxygenated compounds; MT: monoterpenes; ISO: isoprene and its oxidation products; UKN: unknown source; MTLOP: monoterpene lightly oxidized products; C_{12}LOP: C_{12}, C_{14}, or C_{16} lightly oxidized products; C_{13}LOP: C_{13} lightly oxidized products; SQT: sesquiterpenes; MTON: monoterpene-derived organic nitrates; SQTLOP: sesquiterpene lightly oxidized products; MTMOP: monoterpene more oxidized products.

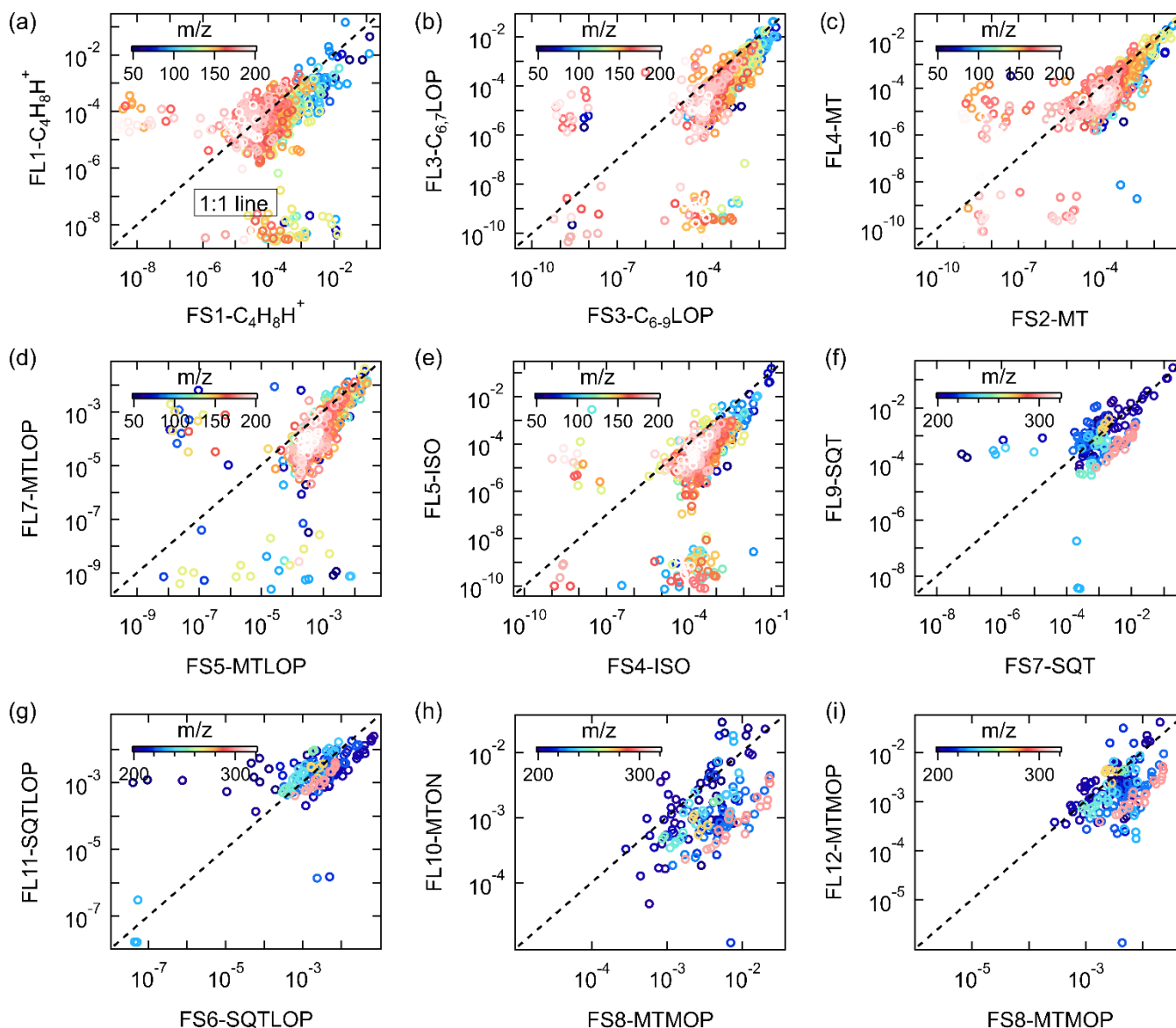



Figure 14. Comparison between factor profiles of the common sources apportioned in the Landes forest and at the SMEAR II station. The x- and y-axis show the fraction of each bin in the mass spectra of the factors. FL: factors in the Landes forest; FS: factors at the SMEAR II station; $C_4H_8H^+$, $C_4H_8H^+$ ion-related; $C_{6,7}LOP$: C_6 and C_7 lightly oxidized products; $C_{6-9}LOP$: C_6-C_9 lightly oxygenated compounds; MT: monoterpenes; ISO: isoprene and its oxidation products; MTLOP: monoterpene lightly oxidized products; $C_{13}LOP$: SQT: sesquiterpenes; MTON: monoterpene-derived organic nitrates; SQTLOP: sesquiterpene lightly oxidized products; MTMOP: monoterpene more oxidized products.