Influence of anthropogenic emissions on the composition of highly oxygenated organic molecules in Helsinki: a street canyon and urban background station comparison

Magdalena Okuljar¹, Olga Garmash², Miska Olin², Joni Kalliokoski², Hilika Timonen¹, Jarkko V. Niemi², Pauli Paasonen¹, Jenni Kontkanen¹, Yanjun Zhang¹, Heini Kuuluvainen³, Mina Aurela¹, Hanna E. Manninen⁴, Mikko Sipili¹, Topi Rönkkö², Tuukka Päijälä¹, Markku Kulmala¹, Miikka Dal Maso² and Mikael Ehn¹

¹Institute of Atmospheric and Earth System Science, Faculty of Science / Physics, Faculty of Science, University of Helsinki, FI-0014, Helsinki, Finland
²Aerosol Physics Laboratory, Physics Unit, Tampere University, PO Box 692, FI-33101, Tampere, Finland
³Atmospheric Composition Research, Finnish Meteorological Institute, Helsinki, Finland
⁴Helsinki Region Environmental Services Authority, HSY, PO Box 100, FI-00066, Helsinki, Finland

Abstract. Condensable vapors, including highly oxygenated organic molecules (HOM), govern secondary organic aerosol formation and thereby impact the amount, composition, and properties (e.g. toxicity) of aerosol particles. These vapors are mainly formed in the atmosphere through the oxidation of volatile organic compounds (VOCs). Urban environments contain a variety of VOCs from both anthropogenic and biogenic sources, as well as other species, for instance nitrogen oxides (NOx), that can greatly influence the formation pathways of condensable vapors like HOM. During the last decade, our understanding of HOM composition and formation has increased dramatically, with most experiments performed in forests or heavily polluted urban areas. However, studies on the main sources for condensable vapors and secondary organic aerosols (SOA) in biogenically influenced urban areas, such as suburbs or small cities, has been limited. Here, we studied the HOM composition, measured with two nitrate-based chemical ionization mass spectrometers and analyzed using positive matrix factorization (PMF), during late spring at two locations in Helsinki, Finland. Comparing the measured concentrations at a street canyon site and a nearby urban background station, we found a strong influence of NOx on the HOM formation at both stations, in agreement with previous studies conducted in urban areas. Even though both stations are dominated by anthropogenic VOCs, most of the identified condensable vapors originated from biogenic precursors. This implies that in Helsinki anthropogenic activities mainly influence HOM formation by the effect of NOx on the biogenic VOC oxidation. At the urban background station, we found condensable vapors formed from two biogenic VOC groups (monoterpens and sesquiterpenes), while at the street canyon, the only identified biogenic HOM precursor was monoterpens. At the street canyon, we also observed oxidation products of aliphatic VOCs, which were not observed at the urban background station. The only factors that clearly correlate (temporally and composition-wise) between the two stations contained monoterpene-derived dimers. This suggests that HOM composition and formation mechanisms are strongly dependent on localized emissions and the oxidative environment in these biogenically influenced urban areas, and they can change considerably also within distances of one kilometer within the urban environment. This further suggests that studies should be careful when extrapolating single-point measurements in an urban setting to be representative for district or city scales.

1. Introduction
Urban environments can contain various anthropogenic and biogenic sources of volatile organic compounds (VOCs). Biogenic emissions come mostly from urban vegetation, for example, trees and bushes in parks, gardens, and may contain biogenic volatile organic compounds (BVOCs) such as isoprene, monoterpenes (MT), or sesquiterpenes. The sources of anthropogenic emissions consist of traffic, cooking, industrial processes and production of customer goods, and volatile chemical products (VCP) (Li et al., 2022; Koppnann, 2007; Watson et al., 2001). Gas-phase compounds emitted from anthropogenic sources contain trace gases, including nitrogen oxides (NO\(_x\)), as well as anthropogenic volatile organic compounds (AVOCs), for example aromatic compounds or aliphatic hydrocarbons (Timonen et al., 2017; McDonald et al., 2018). In densely populated areas, VCPs can dominate AVOCs concentrations and compounds typically known as BVOC (e.g., monoterpenes) are also emitted from anthropogenic sources, such as personal care products and cleaning agents (Gkatzelis et al., 2021; Li et al., 2022).

Under atmospheric conditions, VOCs can undergo oxidation to form condensable vapors (Pandis et al., 1992; Ehn et al., 2014). The most common ambient oxidants are ozone (O\(_3\)), hydroxyl radical (OH), and nitrate radical (NO\(_3\)) (Wayne, 2000). O\(_3\) is a trace gas produced in the troposphere mostly by photolysis of NO\(_x\) (Liu et al., 1980), and present in the ambient air during the entire day. O\(_3\) can oxidize only VOC containing at least one double or triple bond, or, with a slower reaction rate, carbonyls (Bianchi et al., 2019). OH is a short-lived, highly-reactive compound produced mostly by the photolysis of O\(_3\) (Crutzen et al., 1999), thus OH is present in the atmosphere mainly during the daytime. NO\(_3\) is a product of the reaction between O\(_3\) and NO\(_x\), which gets rapidly destroyed by photolysis and reactions with NO during the daytime (Wayne et al., 1991). Both radicals can react with most closed-shell VOCs (Seinfeld and Pandis, 2016), but in the atmosphere, NO\(_3\) reacts mostly with alkenes while OH reacts with almost all compounds, including aromatic hydrocarbons (Seinfeld and Pandis, 2016). Oxidation of VOCs almost always leads to peroxy radical (RO\(_2\)) intermediates, typically with long enough lifetimes to participate in bimolecular reactions, primarily with NO, HO\(_2\), or other RO\(_2\). The RO\(_2\) may also undergo various unimolecular isomerizations, and both these and the bimolecular reactions can lead to either propagation or termination of the organic radical (Bianchi et al., 2019). The structure of the final product depends on multiple factors, including the structure of the initial VOC and the “oxidative conditions”, meaning available oxidants and the bimolecular reaction partners. The latter can be referred to as “terminators” when they terminate the oxidation process, and in some cases the product composition can tell a lot about the oxidative conditions. Additionally, NO\(_x\) can terminate oxidation chain in reaction leading in most cases, which decompose back to substrates (Atkinson and Arey, 2003). For example, RO:\_ oxidation by NO and oxidation by NO\(_x\) can produce organic nitrogen compounds (ONCs), organonitrites (Atkinson and Arey, 2003; Bianchi et al., 2019), while RO\(_2\) termination by NO\(_x\) can form relatively unstable peroxy nitrates. RO\(_2\) cross reactions are the only reactions that can form accretion products, ROOR, referred to here as “dimers” (Valiev et al., 2019). RO\(_2\) intermediates can also undergo autoxidation, where the RO\(_2\) isomerizes through a hydrogen shift (H-shift) creating an alkyl radical to which molecular oxygen can attach (Bianchi et al., 2019; Ehn et al., 2014; Crounse et al., 2013). In the end, a new, more oxidized RO\(_2\) is formed, which can either undergo additional H-shifts or bimolecular reactions, with both potentially terminating or propagating the oxidation (Bianchi et al., 2019) chain. In cases where the radical can undergo multiple autoxidation H-shifts, the end product can reach high enough oxidation levels to be classified as HOM (Bianchi et al., 2019). The structure of a VOC strongly influences its propensity to undergo autoxidation and, consequently, the molar yield of HOM. This results in the very variable HOM yields, which can reach high values for different anthropogenic and biogenic compounds (Molteni et al., 2018; Bianchi et al., 2019; Garmash et al., 2020).
Differences in the structural composition affect both the physical and chemical properties of HOM, with more oxidized products typically being less volatile (Kroll and Seinfeld, 2008). However, the exact functionalities are important, and e.g. oxygen atoms in nitrate groups lower the volatility much less than if the oxygen was found in some other functional group (Kroll and Seinfeld, 2008). In general, the high oxygen content of HOM makes them an important contributor to secondary organic aerosol (SOA) formation, influencing e.g. air quality.

During the last decade, HOM formation from biogenic emissions has been extensively studied in forests (Ehn et al., 2014; Yan et al., 2016; Bianchi et al., 2017; Massoli et al., 2018), and in agricultural environments (Kürtén et al., 2016). Recently, research showed that also the oxidation of AVOCs can noticeably contribute to the HOM population (Molteni et al., 2018; Garmash et al., 2020; Wang et al., 2021) and SOA formation (Timonen et al., 2017). Additionally, NOx can alter the HOM formation mechanism and influence SOA formation (Fry et al., 2014; Ng et al., 2017; Pullinen et al., 2020; Mutzel et al., 2021). Due to these findings, the research on condensable vapors and their origin focused stronger on urban environments, especially very polluted ones, heavily influenced by anthropogenic emissions (Brean et al., 2019; Liu et al., 2021, Guo et al., 2022b; Nie et al., 2022; Yan et al., 2022). In very polluted environments, formation of condensable vapors is greatly impacted by NOx (Brean et al., 2019; Liu et al., 2021; Guo et al., 2022b; Nie et al., 2022; Yan et al., 2022) and HOM composition is often dominated by AVOC precursors (Nie et al., 2022).

While the composition and formation of condensable vapors have been studied in the above-mentioned forests and highly polluted locations, environments with considerable influence from both anthropogenic and biogenic emission sources have received much less attention. Such areas include urban environments with lots of green areas, for example suburbs, or cities surrounded by large forests. A better understanding of such locations may also help to assess the impact on air quality from adding vegetation such as green roofs to already built-up areas. Helsinki is an example of a city with forests in close proximity, and Saarikoski et al. (2023) estimated that there, even at a street canyon site strongly affected by traffic emissions, BVOCs are the main contributor to oxidation products. While Saarikoski et al. (2023) measured only the composition of VOCs, and not their oxidation products, this finding makes us expect that the relative role of BVOCs is even higher for HOM, as BVOCs typically have higher propensity for autoxidation than AVOC (Bianchi et al., 2019). Another important aspect to consider is the spatial representativeness of typical urban measurements. As cities are very inhomogeneous in terms of local emissions and the oxidative environment, and HOM are short-lived compounds, HOM studies in urban environments that were performed at one specific location may not be comparable to other nearby locations with different urban sub-environments.

Here we investigate the composition of condensable vapors at two nearby stations in Helsinki, which are differently influenced by anthropogenic emissions. The first station is located in a busy street canyon while the second is in an urban background area, at less densely built part of Helsinki, 150 meters from the nearest busy road. We studied the composition of condensable vapors, mostly HOM, at these sites using two nitrate-based chemical ionization mass spectrometers. To identify different HOM types from the mass spectra and connect them to different formation pathways, we applied Positive Matrix Factorization (PMF) to separate co-variying species. We compared the drivers of HOM formation between the two urban sub-environments and explored the roles of biogenic and anthropogenic emissions on HOM composition, in order to understand how these can affect the air quality in urban environments with a strong biogenic influence.

2. Methods
We measured the composition of condensable vapors at two stations in Helsinki situated in contrasting environments: the Helsinki Region Environmental Services Authority (HSY) air quality station (60°11′47.0″ N, 24°57′07.7″ E) and the Station for Measuring Ecosystem-Atmosphere Relations (SMEAR III, 60°12′10.4″ N, 24°57′40.2″ E) (Fig. 1). The HSY supersite is located at a street canyon, less than a meter from Mäkelänkatu street (around 28,000 vehicles/weekday) (Kuuluvainen et al., 2018). The street canyon is 42 m wide and the height of buildings on the both sides of the supersite is 19 and 16 m, leading to the average height-to-width ratio of 0.45 (Jarvi et al., 2023). It contains a pavement and three lines of road for both directions separated by the two tram lines and trees. SMEAR III is 900 m north-east of the HSY station and with 150 m distance from the closest busy road (Hämeentie street). SMEAR III is classified as an urban background station (Jarvi et al., 2009). The neighborhood of these stations was previously described in Okuljar et al. (2021). Here we refer to them as “street canyon” (later also “SC”) and “urban background station” (later also abbreviated “UB”), respectively.

The measurement campaign was conducted during 11 May 2018 – 03 June 2018 at the urban background station and 27 April 2018 – 24 May 2018 at the street canyon. The measurement period was during change of season and by 14 May 2018 the deciduous trees in the surrounding area had fully developed their leaves. To study the influence of traffic emissions, we analyzed separately the data measured during workdays as well as only during weekends and public holidays (1 May 2018 and 10 May 2018). We refer to them as ‘workdays’ and ‘weekends’, respectively. As the nighttime concentrations are often influenced by the emission from the previous day, we separate these categories in 24 h periods starting at 4 a.m. The corresponding analysis of size distribution of 1-1000 nm particles measured during the same time at both stations in Helsinki is presented by Okuljar et. al. (2021).

The detailed meteorological description of the transport of pollutants at both stations with the emphasis on the mechanism affecting this transport at the street canyon is presented by Jarvi et al. (2023). In summer the atmosphere is very stable during nighttime and very unstable during daytime at the urban background station. The meteorological conditions at the urban background station (Fig. 2) resemble the one described by Jarvi et al. (2023) for summer, which suggests limited vertical mixing of the atmosphere during nighttime and very well-mixed lower atmosphere during daytime during our measurement. At both stations, during the warm period the mechanical and thermal mixing is stronger than during cold periods resulting in conditions more favorable for pollution dispersion (Jarvi et al., 2023). At the street canyon, not only mean wind but also the turbulent mixing is important for the transport of pollutants. This may suggest that even though the air is not stagnant, the pollutants are not efficiently transported from the street canyon.
The composition of condensable vapors was measured simultaneously at both stations by two nitrate-ion based chemical ionization atmospheric pressure interface (CI-APIs-TOF) mass spectrometers (MS) (Jokinen et al., 2012). Nitrate ions (NO$_3^-$), produced by interactions between soft X-ray and sheath air containing nitric acid (HNO$_3$), binds to the analyzed compound through hydrogen bonds or charges the analyte via proton transfer reactions. NO$_3^-$ is primarily selective towards organic molecules containing at least two suitably positioned hydroxyl (-OH) or hydroperoxyl groups (-OOH) (Hyttinen et al., 2015), or compounds with higher gas-phase acidity than HNO$_3$. After the sample gets ionized, the ions are focused in the APi module and ultimately separated in the time-of-flight (TOF) analyzer based on their mass-to-charge ratio (m/Q, reported in units of Th). The CI-APIs-TOF and its working principle was described in detail by Jokinen et al. (2012). The resolving power of the MS at both stations was approximately 3000-4000 Th/Q for signals with m/Q higher than 200 Th. The mass spectra were analyzed using the software package tofTools (Juminnen et al., 2010).

In measured mass spectra, we observed multiple peaks at every m/Q. To perform high-resolution (HR) analysis requires us to fit closely set signals and could increase uncertainties of results. Therefore, most of our analysis is based on unit mass resolution (UMR) data and we apply HR analysis follow only when we can narrow it down at times where PMF indicated low interference to this factor as described in section 2.4. Additionally, we noted that condensable vapor measurement at street canyon had lower transmission for higher m/Q than at urban background station. Here, we discuss quantitative changes in condensable vapors based on their measured signal in counts per second (cps) normalized by the cps of the reagent ions, (NO$_3^-$, HNO$_3$ NO$_2^-$, and (HNO$_3$)NO$_2^-$) using the unit ncps (normalized cps). The ambient concentrations can be estimated by using previously determined instrument-specific calibration coefficients for sulfuric acid (Okuljar et al., 2021) equal to 4·10$^9$ cm$^{-3}$ for the street canyon station and 7·10$^9$ cm$^{-3}$ for the urban background station. However, usage of these calibration coefficient determined for sulfuric acid to calculate HOM concentration comes with very large uncertainties, and we therefore concentrate on comparison of ion signal strength.

2.2. VOC measurements

VOC concentrations were measured at the street canyon with an offline method in which ambient samples were first collected on a Tenax TA-Carpack B sorbent tube and later analyzed by thermal desorption gas chromatography coupled with mass spectrometry (TD-GC-MS). We measured VOC concentrations during the period 15 – 25 May 2018 with 4 h time resolution. 13 analytes were classified as AVOCs: benzene, toluene, ethylbenzene, p/m-xylene, styrene, o-xylene, propylbenzene, 3-ethyltoluene, 4-ethyltoluene, 1,3,5-trimethylbenzene, 2-ethyltoluene, 1,2,4-trimethylbenzene, and 1,2,3-trimethylbenzene, and 15 as BVOCs: monoterpenoids (α-pinene, camphene, β-pinene, A3-carene, p-cymene, linalool, limonene, terpinolene), terpene alcohol (linalool), an oxidation product of β-pinene (nopinone), bornyl acetate, and sesquiterpenes (longicyclene, iso-longifolene, β-caryophyllene and α-farnesene). More detailed description of the method can be found e.g. in Helin et al. (2020).

2.3. Other instrumentation

CO$_2$, NO, NO$_2$, SO$_2$ as well as meteorological variables were measured at both stations. Table S1 contains information about measurements of additional variables used in this paper.
2.4. Positive Matrix Factorization

Collected datasets from the measurement of condensable vapors at both stations consist of an enormous amount of information and it is challenging to filter data that contain relevant information for analysis of HOM formation. As both stations are located in a city, the composition of condensable vapors is dependent on different types of VOC sources as well as chemical and metrological conditions. To extract relevant information, separate different pathways of HOM formation, and find processes affecting condensable vapor composition at both stations, we applied Positive Matrix Factorization (PMF) (Paatero, 1997; Paatero and Tapper, 1994; Paatero and Hopke, 2003). PMF is a multivariate factor analysis model which has been widely used on aerosol mass spectrometry data (Ng et al., 2011; Zhang et al., 2011; Chen et al., 2022) and more recently on ambient gas-phase chemical ionization mass spectrometry data (Yan et al., 2016; Massoli et al., 2018; Zhang et al., 2019; Liu et al., 2021; Nie et al., 2022).

We performed PMF analysis on three different m/Q ranges from UMR data at both stations: 200-350 Th, 350-500 Th, and 500-650 Th. In this paper, we will refer to these ranges as ranges 1, 2, and 3, respectively. The loss rate of HOM due to condensation is roughly a function of their mass (Peräkylä et al., 2020), thus, analyzing mass spectra in ranges allows us to group HOM with similar loss rates and focus specifically on separating the HOM sources (Zhang et al., 2020).

Additionally, when a m/Q range has lower signal than other ranges, it will only have a minor weight on the PMF solution and relevant information may be lost (Zhang et al., 2020). Using m/Q ranges for PMF analysis is important especially at the street canyon as it may partly counteract the loss of information due to lower transmission for higher m/Q. The focus of our analysis is on compounds in a range of 200 to 650 Th as in this reach we can find majority of the condensable vapors containing C_{2-35}. Smaller m/Q are unlikely to condense, while larger m/Q had very low, or even negligible, signals.

We prepared data and error matrices with 30 min time resolution, separately for each range at each station according to the methods described by Yan et al. (2016). To conduct PMF analysis we used the Igor-based interface Source Finder (SoFi, version 6.6) (Canonica et al., 2013) and ME-2 solver (Paatero, 1999). Detailed information about data preparation and validation of PMF solutions can be found in S1.

To describe the chemical composition of ions in obtained factors, we determined the times for each factor when that factor had the highest relative contribution to the total signal and then fit peaks to the HR data to identify the key compounds. Choosing times when the analyte is dominant across all factors in the same m/Q range and at the same site is necessary to ensure that the identified compound is correctly assigned to the factor. In this paper, we performed a more detailed interpretation only of chosen factors from each station, which we refer to further on as “selected factors”. A factor was chosen for further interpretation only when we could reasonably identify ions in it, and relate it to a real atmospheric source, i.e. not impurities. We refer to other factors as “not selected factors”. Examples of each type will be given later to better clarify this selection process.

2.5. Limitation of data for interpretation

There are several limitations for interpreting the data. At the street canyon, a low signal is observed for higher m/Q. That leads to a low signal-to-noise ratio (S/N) for HOM measured in range 2 and 3, and in some cases makes it impossible to identify the compounds. As a result of fast decrease of measured signal with an increase of m/Q, at the street canyon, over 90% of the signal of 200-650 Th is located in range 1. This could be caused by the lower transmission rate of CI-APi-TOF m/Q of the CI-APi-TOF measurement at that station. Transmission is a result of voltage settings in CI-APi-TOF, which are optimized for each instrument separately. We could not correct our data for transmission, Zha et al. (2018)
showed that the ratio of the signal for the same sampled air measured by two CI-APi-TOF can change drastically with an increase of m/Q due to the difference in the transmission between instruments. Thus, the highest uncertainty caused by inconsistent transmission between two instruments is observed in range 3. Nevertheless, this uncertainty does not influence identification of peaks that have sufficient S/N.

Due to the chemical complexity of the samples, we cannot achieve high accuracy of mass calibration on some of the measured days. This is the reason why we have performed PMF analysis on UMR data. Limitations of peak identification due to the MS resolution and the presence of multiple overlapping peaks also hinder the identification of some ions, and hence we are confident to report only the dominant ions in each factor. We are not able to report key compounds for factors that have minor contribution to their m/Q range or have too many similar peaks with other factors, as we cannot unambiguously assign identified compounds to these specific factors.

Lastly, we need to keep in mind that chemical ionization with NO3- is very selective, mostly towards highly functionalized compounds. Overall, this ionization method is optimal for detection of HOM, however, it limits observations of other oxidation products.

3. Results and discussion

We start this section by providing a short overview of the meteorological conditions during the campaign. In section 3.2, we present our main findings, starting from the PMF results and the subsequent interpretations of important formation pathways of condensable vapors at the two measurement sites. In the last part of this section, we discuss the potential implications of our findings on the air quality in Helsinki.

Concerning notations, we focus our study on HOM, but we also detect abundant organic compounds which contain less than six oxygen atoms, which do not classify as HOM. Thus, we often use a broader term ‘condensable vapors’ when discussing observed products more broadly. In addition, we observe monomeric (mostly C9-C10) and dimeric (mostly C19-C20) oxidation products of MT, which we refer to as ‘MT-derived monomers’ and ‘MT-derived dimers’, respectively. For simplicity, we call factors containing monomeric oxidation products of MT ‘MT monomers’ while ‘MT dimers’ factors contain dimeric oxidation products of MT.

3.1. Overview of meteorological and trace gas conditions in Helsinki

Atmospheric conditions, for example local emissions and oxidative environment, influence HOM formation pathways. To understand HOM formation mechanisms and their differences between studied sites, we first investigated meteorological and chemical conditions at both stations. Figure 2 presents diurnal variations of measured variables that can influence HOM formation pathways: global radiation, ambient temperature (T), and concentrations of O3, NO, and NO2, as well as the wind direction measured during this campaign. As mentioned earlier, the measurement periods overlapped but were not identical between the two stations. Therefore, differences in campaign averages between sites are partly driven by differences in location and partly by differences in time. The difference in meteorological parameters (ambient temperature, global radiation, wind direction) are driven mostly by changes in the measurement period (Fig. 2, S2), while the trace gases (NO, NO2, O3) vary largely also during overlapping times of measurements (11 May 2018 – 24 May 2018) (Fig. S2).
Figure 2. Diurnal variations of (a,b) global radiation and ambient temperature, and (c-d) NO, NO₂, and O₃ concentrations at the street canyon (left) and urban background station (right). Presented data contain both workdays and weekends. The median diurnal variations are shown as solid lines with markers; 25th to 75th percentile ranges are presented as shaded areas. Time is local. The right panel (c) shows a wind rose for the overlapping time at both stations as well as outline of wind rose for the period when data was collected from the street canyon (SC, dark red) and the urban background station (UB, dark blue). Arrow in panel (c) shows the direction of the urban background station (UB) in relation to the street canyon (SC). Presented data contain both workdays and weekends.
Diurnal variation of global radiation is similar between the two stations (Figure 2a and b), though with slightly more cloudy periods at the street canyon. Global radiation initiates photolysis reactions and, as a result, enhances the formation of OH and O₃ as well as the decomposition of NOₓ. During the overlapping time of measurements, the urban background site was more likely to be the downwind station, however the wind was mostly coming from North (Fig. 2c), meaning that neither station would receive emissions from the other. Median temperatures varied between 12.0°C and 21.6°C at the urban background site and between 8.9°C and 17.3°C in the street canyon. Higher temperature at the urban background station can be explained by the difference in measurement periods as the measurements started two weeks later than in street canyon. During the period when measurements overlapped, the median temperature is very similar between stations reaching almost 22°C during daytime and dropping to 12-13.4°C during nighttime (figure not shown Fig. S2). The increase in temperature typically accelerates molecular reaction rates as well as enhances BVOCs emissions and evaporation rates. It can also affect HOM yields (Queuev et al., 2019).

At the urban background station, NO has a maximum between 8:00 and 9:00 (2.5 ppb) and it is negligible during nighttime. In contrast, at the street canyon, the median NO concentration was below the detection limit between 1:00 and 3:00, after which it rapidly increased, levelling off at 7:00 and staying elevated (ca. 9 ppb) throughout the day until 17:00. That means NO can affect oxidation reactions more at the street canyon site, even during much of the night, when it stays at 4 ppb until early morning. In the context of VOC oxidation, the presence of NO likely causes the termination of the oxidation. In the absence of NO, termination reactions with RO₂ become more favorable. NO₂ and NO (Figure 2a,b,c,d) concentrations are up to 5 and 23 times higher at the street canyon than at the urban background station, respectively. At urban background site, O₃ reaches minimum median concentration at 7:00 (30.3 ppb) and maximum at 19:00 (49.1 ppb). At the street canyon, the corresponding values are 18.5 ppb and 36.5 ppb at 13:00. During overlapping times between the sites, median O₃ concentration stays 5-25 ppb lower at street canyon than at urban background station (figure not shown Fig. S2) It could be partly associated with higher NO concentration in street canyon as NO reacts with O₃. O₃ remains relevant for VOCs oxidation throughout the day. O₃ and NO₂ concentrations affect production of NOₓ and thus its concentration.

3.2. Characterization of PMF factors

In this subsection, we examine the HOM composition and formation at both stations by investigating PMF factors in all three mQ ranges (200-350 Th, 350-500 Th, 500-650 Th); we focus our analysis on selected factors, their time-series (Fig. S3-4), and diurnal variations (Fig. 3-4) as well as mass spectra (Fig. S2-5-6). We refer to PMF factors as SCX-Y or UBX-Y where SC is the street canyon, UB is the urban background station, X is the analyzed mQ range (either 1, 2, or 3), and Y is the identifying number of the factor in that range. The factors also appear together with a descriptive name. As an example, “UB3-2: MT dimers” refers to the second PMF factor identified in mass range 3 at the urban background site and was found to mainly contain ions related to monoterpene-derived dimers. To understand the chemical composition of factors, we identify their key compounds with HR data (Table S2) as described in Sect. 2.5. All key compounds are detected as clusters with NOₓ or HNO₃(NO₃) and this is how we report them in Table S2 and on the mass spectra (Fig. S2-5-6 and S6-79-10); however, for clarity of the interpretation in this subsection, we write their chemical structures without the nitrate adducts. The PMF analysis involves several lengthy steps, including determining an optimal number of factors in the solution, as well as interpreting sources for each factor based on the supporting data available. In this study, we had six data sets to...
analyze (two sites and three mass ranges). As a result, we summarize the key characteristics of each factor and give an interpretation in the main text and present of factors. The more detailed description of the PMF analysis can be found in the supplementary information (SI) (Sect. S2 motivates the choice of factor numbers, Sect. S3 describes the main features of each factor, which lead to the interpretations given in the main text). In the following sections, we first briefly describe the overall characteristics of factors observed at the street canyon (Sect. 3.2.1) and at the urban background station (Sect. 3.2.2) after which we compare HOM formation and composition between these sites (Sect. 3.2.3).

PMF inputs and validation, S4. Factor interpretation). Only “selected” factors are described here, while characteristics of “not selected” factors are presented and described in the SI (Sect. S4, Fig. S6-S7, Table S3). Several reasons motivated us not to select factors for detailed discussion in the main text. For example, a factor was not selected if it was a contamination or an artefact (e.g., containing mainly water clusters isotopes) or if we were not confident in the meaningful separation of this factor by PMF method (this was the case for the entire range 1 at the street canyon, as described below). Overall, we selected 5 out of 13 factors from the street canyon and 10 out of 14 factors from the urban background station. These selected factors explain 34%, 100%, and 100% of the observed signal in ranges 1-3 at the urban background station and 0%, 64%, and 61% of the observed signal in ranges 1-3 at the street canyon, respectively (Table 1, Fig. S2-4, S7-8).
Figure 3. Time series of selected PMF factors (a-b), global radiation and ambient temperature (c), and The diurnal variation of PMF factors fractions (d-e) at the street canyon (SC). PMF factors are labeled as SCX-Y where X stands for the analyzed m/Q range (2 or 3), and Y is the identifying number of the factor in that range. The median diurnal variation is shown as a solid line with markers; 25th to 75th percentile ranges are presented as shaded areas. Y-axis in ncp s indicates the measured signal in normalized counts per second.
330 Figure 4. Time series of selected PMF factors (a–d), global radiation and ambient temperature (e), and the diurnal variation of PMF factors (f–i) at the urban background station (UB). PMF factors are labeled as UBX-Y where X stands for the analyzed m/Q range (1, 2, or 3), and Y is the identifying number of the factor in that range. The median diurnal variation is shown as a solid line with markers; the 25th and 75th percentile ranges are presented as shaded areas. Y-axis in ncps indicates the measured signal in normalized counts per second.

336 3.2.1. Street canyon

Here, we very briefly describe factors observed at the street canyon site in each m/Q range. The main examination of the factors is given in Sect. 3.2.3 where we also discuss them in relation to factors observed at the urban background station. As indicated already above, PMF solutions for range 1 at the street canyon were inconclusive, and therefore all factors from this range are classified as ‘not selected’. The main reason is that all factors had very similar temporal trends, mainly correlating with temperature. This may be a result of most observed molecules being semi-volatile, and increased temperatures lead to increased evaporation of these molecules. In any case, as PMF relies on temporal variability to separate factors, too much co-variance makes PMF less reliable. Nevertheless, we believe there was some useful information also in this range and will briefly discuss SC1-1: nitrophenol, SC1-2: MT monomers, and SC1-5: nitrophenol & aliphatic in this section.

Range 1, 200-350 Th (Factors selected: 0/5)

In range 1, all factors are affected by changes in ambient temperature (Fig. S6 and S7). Factors in range 1 have a daytime peak and nearly all of them could have been oxidized by OH or O3 (Fig. S6, Table S3). Most of these factors are likely formed from AVOCs and contain nitrophenol (C6H4O3N) as well as other N-containing aromatics, such as nitrocresol. Nitrophenol can be directly emitted from combustion or formed from benzene and phenol oxidation. The presence of nitrophenol in many factors can be explained by an abundance of benzene at the street canyon as it is the third most abundant VOC measured at the street canyon site (Fig. S5).
Range 2, 350-500 Th (Factors selected: 3/5)

In range 2, all selected factors respond to the changes in the ambient temperature (Fig. 4S1 and 4S13), especially factor SC2-1, which contains monomeric oxidation products of MT (MT-derived monomers) with nitrate functionalities: \( (\text{C}_9\text{H}_{21}\text{O}_3\text{N}_2) \) and \( (\text{C}_9\text{H}_{18}\text{O}_2\text{N}_2) \). Factors SC2-2 and SC2-3 are highest during the night, but they also have local maxima during the day, which suggests that competing processes influence the formation of these factors and thus their diurnal pattern. SC2-3 may be inhibited by NO, as it decreases when NO reaches its daily maximum (Figure 3d). SC2-3 consists of MT-derived monomers: \( (\text{C}_9\text{H}_{20}\text{O}_2\text{N}_2) \) and \( (\text{C}_9\text{H}_{19}\text{O}_2\text{N}_2) \), while SC2-2 is dominated by one single compound: \( \text{C}_9\text{H}_{18}\text{O}_2\text{N}_2 \).

Range 3, 500-650 Th (Factors selected: 2/3)

Range 3 contains one daytime and one nighttime factor. SC3-1 is a MT-derived dimer factor produced via oxidation by NO\(_x\) and present during the night, when NO concentrations are low enough to allow RO\(_x\) termination via RO\(_x\) cross reactions. The key compounds detected in SC3-1 are coming from one oxidation chain: \( \text{C}_9\text{H}_{19}\text{O}_2\text{N}_2 \). SC3-2 is a daytime factor containing HOM oxidized by OH. \( (\text{C}_9\text{H}_{22}\text{O}_3\text{N}_2) \). SC3-2 also has some signal from instrumental impurities containing fluorine (F-impurities), and undefined noise peaks.

3.2.2. Urban background station

Similar to the previous subsubsection, we describe briefly factors observed in all ranges at the urban background station. The discussion about these factors follows in Sect. 3.2.3, in which we compare factors found at both sites in Helsinki.

Range 1, 200-350 Th (Factors selected: 2/6)

Selected factors in range 1 contain daytime factors, from which UB1-6 is a factor correlating the best with the ambient temperature (Fig. 4S4 and 4S14). Time-series of UB1-6 correlates with O\(_3\) (Fig 4a, Fig. 4S14) and it contains key compounds with C\(_{18}\) atoms. These formulas have been detected earlier as products of MT oxidation in chamber studies (Yan et al., 2020; Saarikoski et al., 2023) and in ambient measurement (Liu et al., 2021), however, they have also been identified as oxidation products of aromatic VOCs (Guo et al., 2022b). Since CI-API-TOF does not provide information about molecular structure, we cannot unambiguously determine the origin of this factor. In contrast to UB1-6, UB1-2 factor contains nitrophilen and likely originates from AVOCs. The diurnal variation of UB1-2 resembles the one expected for OH (Saarikoski et al., 2023). Both UB1-2 and UB1-6 contain ONCs and their oxidation was likely terminated by NO or NO\(_2\).

Range 2, 350-500 Th (Factors selected: 5/5)

Range 2 contains various daytime and nighttime factors (Fig. 4). Factor UB2-1 reaches the highest concentrations at 5 am. and 10 pm., which corresponds to the time of sunrise and sunset during our measurement period. As this factor consists of MT-derived HOM with two N-atoms, we can speculate that they are formed from NO\(_x\) oxidation of MT and terminated by NO. It is typically assumed that NO\(_2\) and NO would not co-exist. However, simultaneous presence of NO\(_2\) and NO when photolysis is just high enough to form NO but not to fully deplete NO\(_2\) is a plausible explanation for the diurnal pattern of UB2-1. As this factor consists of MT-derived ONCs, possibly with at least two N-atoms \( (\text{C}_9\text{H}_{19}\text{O}_2\text{N}_2) \) or \( (\text{C}_9\text{H}_{18}\text{O}_2\text{N}_2) \), we can speculate that they are formed from NO\(_x\) oxidation of MT and terminated by NO. Chamber experiments have shown that formation of condensable vapors through NO\(_x\) oxidation and NO termination is a possible formation mechanism during NO\(_x\)-influenced oxidation of monoterpenes (Yan et al., 2020; Nie et al., 2023).
It is typically assumed that NO$_3$ and NO would not co-exist in the atmosphere. However, in urban areas NO concentration is not negligible during the night and this pathway was previously suggested as possible mechanism of ONC formation in Beijing (Guo et al., 2022h). Simultaneous presence of NO$_3$ and NO when photolysis is just high enough to form NO but not to fully deplete NO$_3$ is a plausible explanation for the diurnal pattern of UB2-1.

UB2-2 and UB2-3 are both nighttime factors oxidized mainly by NO$_3$ and inhibited by NO during daytime. UB2-2 contains MT-derived monomers and correlates with the MT-derived dimer factor (UB3-2). Following the diurnal cycle in Fig. 4a,b, it can be observed that when the concentration of UB2-2 decreases, the concentration of daytime MT-derived monomer factor, UB2-4, increases. Even though UB2-2 and UB2-4 both contain key compounds with C$_{9}$H$_{16}$, the molecular formulas are slightly different. Specifically, in UB2-2 key compounds contain mainly one or three N-atoms while in UB2-4 they have mostly zero or two N-atoms, (Table S2). UB2-2 and UB2-4 could thus be formed from competing HOM formation pathways from the same VOCs.

In contrast to other factors, UB2-3 consists of HOM with composition of C$_{10}$H$_{2}$O$_{5,m,n}$N$_{12,15,16}$, based on which we conclude that this factor is formed from sesquiterpenes (C$_{15}$H$_{24}$, Richters et al., 2016). UB2-3 correlates very well with a corresponding sesquiterpene factor from range 3, UB3-1 (R=0.93) (Fig. S5S12). The last factor UB2-5 is a daytime factor which during noon corresponds to more than 50% of the measured signal (Fig. S8). It contains C$_{10}$O$_{8}$ compounds (C$_{10}$H$_{12}$O$_{8}$N and C$_{10}$H$_{12}$O$_{8}$N$_{4}$) most likely that UB2-5 is formed in OH oxidation.

**Range 3, 500-650 Th (Factors selected: 3/3)**

Range 3 at UB site contains two nighttime factors: sesquiterpene-derived UB3-1 factor, (C$_{10}$H$_{12}$O$_{8,2}$N and C$_{10}$H$_{12}$O$_{8,2}$N$_{4}$ (or C$_{10}$H$_{12}$O$_{8,2}$N$_{4}$)), and MT-derived dimer UB3-2 factor, (mostly containing C$_{16}$O$_{23}$ compounds, for example C$_{16}$H$_{24}$O$_{28,2}$N$_{2}$). Both factors consist of ONCs, products of NO$_3$ oxidation of BVOCs, and are inhibited by NO, being absent during the day as a result. UB3-3 is the only daytime factor (Fig. 4c,d) in range 3 and it consists of OH-oxidized HOM, F-impurities, and noise.

3.2.3. Factor interpretation and comparison between urban background and street canyon sites

Tables 1 and S3 present the most plausible interpretation of selected and not selected factors, respectively. For each factor, we propose VOC precursors, oxidants, and terminators, which were most likely to influence the formation of species in this factor. We also specify an hour of the day when factor’s signal reached its maximum as well as the contribution of this factor to the total signal both within its own m/Q sub-range and within the full analyzed range (200-650 Th). See Table 1 caption for a more detailed description of how to read the table. The findings and implications are discussed below. While discussing time series correlations between factors from both sites, it is important to keep in mind that they cannot be ideal. We estimated that the highest correlation between stations for condensable vapers is approx. 0.88 which is a correlation between concentrations of a compound that is mostly produced in the same pathway on both sites; sulfuric acid (SA) (Fig. S15).
Table 1. Suggested characterization of selected factors at both stations. Detailed factor interpretation is described in Sect. 4.3. The importance of the various species described in this table was assessed based on either factor time series (TS), factor mass spectra (MS), or both (B), as indicated by the superscript in the “Factor” column. The “Precursor” column describes which type of molecules we expect to act as precursors to the observed signals, separating (when possible) between AVOC and BVOC. The “Oxidant” and “Terminator” columns depict our estimates for the most likely species involved in the oxidation process (“M”, as in “maybe”, is used if we were unable to exclude or confirm the participation of the species). If the “yes” or “no” is marked in bold font, it means that we found a particularly clear influence of that species for that factor. The “Diurnal peak time” shows the hour when the factor had its highest concentration, and “Fraction” depicts the percentage of signal (of the given sub-range or the total analyzed m/Q range) that the factor contributed to.

<table>
<thead>
<tr>
<th>Range</th>
<th>Factor</th>
<th>Precursor</th>
<th>Oxidant</th>
<th>Terminator</th>
<th>Diurnal peak time</th>
<th>Fraction [%] within 200-650 m/Q range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>OH</td>
<td>NO₂</td>
<td>O₃</td>
<td>NO</td>
</tr>
<tr>
<td>Street canyon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>350-500</td>
<td>SC2-1TS</td>
<td>BVOCs</td>
<td>M</td>
<td>no</td>
<td>M</td>
<td>yes</td>
</tr>
<tr>
<td>350-500</td>
<td>SC2-2TS</td>
<td>VOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>350-500</td>
<td>SC2-3TS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>500-650</td>
<td>SC3-1S</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>500-650</td>
<td>SC3-2IS</td>
<td>VOCs, noise, F-impurities</td>
<td>yes</td>
<td>no</td>
<td>M</td>
<td>yes</td>
</tr>
<tr>
<td>Urban background station</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200-350</td>
<td>UB1-2RS</td>
<td>AVOCs</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>200-350</td>
<td>UB1-6TS</td>
<td>VOCs</td>
<td>M</td>
<td>no</td>
<td>M</td>
<td>yes</td>
</tr>
<tr>
<td>200-350</td>
<td>UB2-1TS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>350-500</td>
<td>UB2-2TS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>no</td>
</tr>
<tr>
<td>350-500</td>
<td>UB2-3RS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>no</td>
</tr>
<tr>
<td>350-500</td>
<td>UB2-4RS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>no</td>
</tr>
<tr>
<td>350-500</td>
<td>UB2-5RS</td>
<td>VOCs, noise</td>
<td>yes</td>
<td>no</td>
<td>M</td>
<td>yes</td>
</tr>
<tr>
<td>500-650</td>
<td>UB3-1RS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>no</td>
</tr>
<tr>
<td>500-650</td>
<td>UB3-2RS</td>
<td>BVOCs</td>
<td>no</td>
<td>yes</td>
<td>M</td>
<td>no</td>
</tr>
<tr>
<td>500-650</td>
<td>UB3-3RS</td>
<td>VOCs, noise</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
Figure 5. Time series of PMF factors with Pearson correlation coefficient higher than 0.7 (Figure S9-S12) between the street canyon (SC) and the urban background station (UB) for common measurement time. PMF factors are labeled as SCX-Y or UBX-Y where X stands for the analyzed m/Q range (1, 2, or 3), and Y is the identifying number of the factor in that range.

The total concentration of AVOCs is much higher than BVOCs at both stations (shown for street canyon in Fig. S8-S11), however, most of the selected factors are more likely to have biogenic origin, primarily based on the identified peaks in the mass spectra (Table 1, Sect. S3-S4). Our result is in agreement with the earlier study by Saarikoski et al. (2023) which concluded that, despite dominant AVOCs concentrations at street canyon site, BVOCs are estimated to be the main source
of oxidized products, due to their higher reactivities. MT are the only type of biogenic precursors identified at the street canyon while at the urban background station we find oxidation products of both MT and sesquiterpenes. This is likely due to the difference in proximity of trees and vegetation from the stations as none of measured sesquiterpenes exceed 0.2% of measured total BVOC concentrations at the street canyon. All key compounds detected in MT-derived factors were previously reported in studies investigating the influence of NOx on HOM formation from MT precursors (Pullinen et al., 2020; Yan et al., 2020; Shen et al., 2021; Dam et al., 2022; Guo et al., 2022a). Even though MT can be emitted from anthropogenic sources, for instance in form of VCPs (Gkatzelis et al., 2021; Li et al., 2022), the population density of Helsinki is low enough that that signature is likely lost among the abundant biogenic MT signals. This is in agreement with a-pinene being the most abundant BVOC and a common VCP, limonene (Coggan et al., 2021), being very low (<5% of BVOC concentration).

We observed only few PMF factors that we expect to originate from AVOC oxidation, and for both stations AVOC factors are detected only for the smallest m/Q sub-range, 200-350 Th (Table 1, Table S3). None of the factors had diurnal variations resembling traffic emissions at these sites (Olin et al., 2020; Okuljar et al., 2021). Even though concentrations of some factors differed between weekends and weekdays, the diurnal behavior of factors was very similar (Fig. S12). This suggests that the local emissions from traffic at these sites did not oxidize to form HOM at adequate fast enough, and with high enough yields or in a large of detectable compounds, to considerably contribute to our measured signals. This is in line with conclusions presented by Brean et al. (2019) and Saarikoski et al. (2023). As precursor VOC concentrations are also affected by the mixing layer height (MLH), this effect may also impact HOM formation. However, VOCs are only one of the components that affect HOM formation, the effect of the MLH on HOM observed in this study is expected to be small.

Between the two stations, there are only a few factors with similar key compounds (Table S2). SC3-1: MT dimers has partially corresponding key compounds with UB3-2: MT dimers. They both contain C_{18}H_{32}O,N_{2} compounds, where x is 11-15 for SC and 13-16 for UB. UB3-2: MT dimers contains also other types of dimers which are not usually present in SC3-1: MT dimers. These slight differences between stations may be caused by the small difference in concentration of oxidant, or the concentration or type of MT. The non-negligible concentration of NO during nighttime at SC may also impact the dimer formation there. Nevertheless, these factors are very similar and form through similar pathways (Table 1) – oxidation of MT mostly by NOx, and termination through ROx cross reactions, leading also to correlating time series (R=0.77, Fig. 5 and S812). The ROx + ROx reactions forming the MT-derived dimers will inevitably also form monomers as the dimer yield is never 100 %. However, monomers can also form through all other ROx termination channels, making them much more heterogeneous than the dimers. The time evolution of some MT-derived monomer factor time series (SC2-3: MT monomers 2 and UB2-2: MT monomers 2) correlate with the corresponding dimer factors (R= 0.67, and R=0.67 respectively) as well as with each other (R=0.5657). While both factors are dominated by C_{10} compounds, their detailed mass spectra have significant differences (Fig. S255 and S256). UB2-2: MT monomers 2 contains mainly ONCs with one N-atom while SC2-3: MT monomers 2 has more compounds with two N-atoms (Table S2). This may indicate that there is enough NO available to terminate some fraction of the ROx yet without totally shutting down the ROx + ROx channel.

Another pair of factors showing similarities between the stations is SC2-2: 370+433 and UB2-4: MT monomers 3 (Fig. 5). Both factors are driven mostly by one compound (C_{18}H_{32}O,N_{2}), which has been detected as two clusters C_{18}H_{32}O,N_{2}+NO_{3} (370 Th) and C_{18}H_{32}O,N_{2}+HNO_{3}+NO_{x} (433 Th) in our instrument (determined by correlation analysis).
The high time series correlation (R=0.75) suggests that molecules in these factors are formed via very similar pathways between the sites. Potentially, the formation pathways are identical, but importance of some competing pathways differ between the sites. Overall, the lack of stronger resemblance between these nearby sites suggests that even if HOM have the same VOC precursors, the environmental conditions regulate the relative importance between different oxidation pathways.

While differences in emissions and oxidation reactions will lead to diverse mass spectra, also the time series are expected to vary between the sites as the wind direction changes. For example, the street canyon site will likely be impacted by the street in different ways if the wind direction is from the street or towards the street. A clearly longer campaign than ours would be needed to identify the detailed impacts from different wind directions. However, analysis of the average diurnal variation can help us understand the roles of different oxidation conditions if the impact of varying wind directions diminishes in a longer average. Most factors at both stations can be characterized by one of a few types of diurnal patterns.

Factors with a daytime diurnal variation reaching maximum concentration during noon or afternoon resemble diurnal variation of OH or O₃, respectively. However, temperature also peaks in the afternoon, and can lead to both higher BVOC emissions as well as evaporation of semi-volatile species from aerosols or surfaces, convoluting the effect of the oxidants on the observed HOM. Factors with noon or afternoon maxima are mostly found in range 1 at both sites, and are mostly sesquiterpene 1, namely minima during morning and evening (Fig. 4), around sunrise and sunset. We can speculate that these are the periods when sunlight was still available, but at limited amounts. This effect may cause an optimal situation for having both NOₓ and NO participating in the oxidation process. This is supported by the high N-atom content of the main species in this factor. Meanwhile, some other factors showed an opposite trend to UVB-2, namely minima during morning and evening, often with a strong nighttime peak and a smaller daytime increase. Some of the most prominent factors with such behavior were SC2-3: MT monomers 2, UB2-2: MT monomers 2, UB2-3: sesquiterpene 1, and UB3-1: sesquiterpene 2. NOₓ was identified as the main oxidant for these factors based on the mass spectra and the high nighttime signals, but the local maxima around noon is surprising. Saarikoski et al. (2023) did estimate that NOₓ would have a small daytime maximum, likely due to the sinks not being fast enough to fully overwhelm the very high formation rates from high O₃ and NOₓ during this time. We cannot determine to which extent the diurnal variation of NOₓ influences these diurnal patterns. As was the case also in many situations discussed above, we are often unable to separate if an increase is due to an enhanced source strength or a decrease in competing reaction pathways.

Comparison to previous research
HOM data from Helsinki show similarities with previous studies done on ambient HOM data in urban as well as rural environments. Yan et al. (2016) investigated HOM formation pathways at a boreal forest site (SMEAR II station, Hyytiälä, Finland) located approximately 190 km from Helsinki and 50 km from the closest city – Tampere (with population approximately 250,000). A factor “Nighttype type-2”, obtained from PMF analysis by Yan et al. (2016) contained MT-derived dimers formed by NO\textsubscript{3} and O\textsubscript{3} oxidation and RO\textsubscript{2} termination. That factor mostly consisted of C\textsubscript{32}H\textsubscript{56}O\textsubscript{34}N\textsubscript{4} (40%) and C\textsubscript{31}H\textsubscript{55}O\textsubscript{33}N\textsubscript{2} (20%) suggesting that the dimers detected in Hyytiälä and in Helsinki (both stations) have the same formation pathways, even though these measurement sites represent different rural and urban environment. Despite relatively similar precursors and formation pathways, far fewer similarities are found between the mass spectra of MT-derived monomer factors at these three sites. This suggests, as also mentioned above, that monomer formation pathways are much more diverse compared to dimer formation. Still, a comparison of our results with other studies done in rural environments (Massoli et al., 2018; Kürten et al., 2016) showed clearly lower resemblance between MT-derived dimers than one described in preceding sentences, which is likely a result of different biome types in their studies (isoprene-dominated south east US and rural agricultural site in Germany respectively) compared the ones conducted in Finland. In recent years, more research on condensable vapor formation has been conducted in urban environments heavily influenced by NO\textsubscript{3} (Yan et al., 2022; Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022; Zhang et al., 2022). Unlike forest environments, where the fraction of nitrogen-containing HOM is similar to the fraction of HOM without nitrogen atoms (Yan et al., 2016; Massoli et al., 2018), condensable vapor composition in Chinese megacities is dominated by nitrogen-containing compounds, which represent approximately 60-85% of all measured condensable vapors (Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022; Zhang et al., 2022). This strong influence by NO\textsubscript{3} was also observed in the present study at both stations in Helsinki. In addition, the majority of key compounds in SCI-5: nitrophenol & aliphatic are also listed as main compounds in factors originating from aliphatic AVOCs detected in Nanjing (Aliph-OOM) (Liu et al., 2021) and in Beijing (aliphatic OOMs) (Guo et al., 2022b). However, depending on the time of the year, the main precursors for condensable vapors in cities in China are either AVOCs or a mix of AVOCs and BVOCs (Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022). This is clearly during spring, 77% of all measured condensable vapors originate from AVOC in Beijing (Guo et al., 2022b). This is different compared to Helsinki, where BVOC-derived vapors were more abundant. This dissimilarity is likely due to the AVOC-BVOC ratio being much larger in Chinese cities due to closer proximity of much larger areas with anthropogenic emissions. In contrast, Helsinki AVOC:BVOC is much smaller due to larger BVOC emissions from abundant vegetation in the close surroundings. It is also important to notice that most studies of condensable vapors in Chinese cities (Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022) analyzed much smaller mass range (200-400 Th or 250-400 Th), which corresponds to range 1-2 here. In our study, range 1 is the only mass range in which we find the dominant influence of anthropogenic precursors. Brean et al. (2019) also showed that MT-derived dimer concentrations were approximately 50 times lower than MT-derived monomers in Beijing, likely due to both small MT emissions and suppression of dimer formation by NO.

3.3. Implications for air quality

Ambient air pollution was recognized as the largest environmental health risk and one of the top risk factors for the loss of healthy years (Lim et al., 2012; Anderson et al., 2012; Cohen et al., 2017). Premature deaths caused by ambient air pollution are linked to particulate matter (PM) (Cohen et al., 2017; WHO, 2021), both due to short-term (Pope and Dockery, 2012) and long-term exposure (Burnett et al., 2014). In many environments, including different urban areas, PM is dominated by secondary aerosol formed from condensable vapors, including HOM. HOM and other condensable
organic vapors impact not only PM concentration but also the chemical composition of SOA and, consequently, aerosol properties like toxicity. For example, in a recent study with human alveolar epithelial cells and human monocyte cells, the organic compounds and the aging of the aerosol were major drivers of the cell level toxicity of aerosol (Hakkarainen et al., 2022).

In this work, we found that the majority of low-volatility condensable vapors in Helsinki were impacted by both biogenic and anthropogenic precursors, despite high local anthropogenic emissions. The VOC precursors themselves were mostly of biogenic origin, i.e. BVOC, but the oxidation process was strongly perturbed by anthropogenic activity, particularly by NOx. While detailed similarities in mass spectra of factors were often small between the close-by sites studied here, most observed compounds at both stations were ONCs. Previous studies have shown that NOx can change the yield of SOA formation during VOC oxidation (Mutzel et al., 2021; Jaoui et al., 2013; Ng et al., 2017), though this effect may be not as clear to observe in ambient measurements (Yan et al., 2022). In the smaller m/Q ranges studied in this work, the influence from AVOC was larger, but we cannot deduce the impact of these factors on SOA formation due to their semi-volatile nature. Nevertheless, our results indicate that in Helsinki, and likely in other biologically influenced urban areas, anthropogenic emissions affect HOM formation and composition most strongly by the participation of NOx in the (BV)OC oxidation. That influence will be propagated to the SOA, both concerning the composition as well as the effective yield of SOA from the BVOC oxidation, but quantifying the ultimate impact on either of these will require further studies.

4. Conclusions

We measured the composition of condensable vapors, HOM, during late spring at two stations separated by 900 m in different sub-environments in Helsinki, a city with considerable biogenic influence from trees. We compared HOM composition and formation pathways at the two sites, an urban background station and a street canyon, using PMF analysis to separate the complex data into covarying compound groups. We found that the majority of the HOM originated from BVOCs at both locations, despite them being dominated by AVOC emissions (Rantala et al., 2016; Saarikoski et al., 2023). However, we did observe a strong anthropogenic influence on the HOM formation, due to the elevated NOx concentrations at both stations, which is consistent with previous studies conducted in urban environments (Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022). The PMF factors, and their temporal behavior, were surprisingly different between the two sites, considering their relatively close proximity. Monoterpene-derived dimers were the compound groups that correlated best between the sites. On the contrary, at the street canyon site we observed a factor corresponding partly to AVOC-derived factors found in Chinese megacities (Guo et al., 2022b; Liu et al., 2021; Nie et al., 2022). The lack of a similar factor in the PMF solution from the urban background station highlights that HOM composition at two nearby sites in an urban environment can differ noticeably depending on the local anthropogenic influences. To a large extent, we expect this difference to be driven by differences in the environmental conditions, leading to distinct oxidation products even when the same VOC molecule becomes oxidized, due to competition between both oxidants and RO2 terminators.

Our work indicates that when analyzing and discussing the impact of HOM on SOA and air quality in urban environments, we need to keep in mind the spatial inhomogeneity of urban areas in the HOM composition and formation mechanisms. Thus, a more detailed investigation of the formation and composition of HOM in a variety of different urban sub-environments would be beneficial. Additionally, our findings are restricted to a short and biologically active period, hence follow-up research on seasonal changes is needed. Finally, we recommend that future mass spectrometric studies in urban
area employ devices with resolving power above 5000 Th/Th, as the mass spectra are extremely complex and thus even peak identification can be a major challenge.

Data availability

All data presented in this manuscript will be available in open repository before the final version of manuscript is completed. PMF factors and VOC data used for this study are available [Zenodo link will be added before printing]. SA, sub-3 nm particle concentration as well as meteorological and trace gas data from street canyon are available at https://doi.org/10.5281/zenodo.4884875 (Okuljar, 2021). Trace gas and meteorological data measured at the background station are available at the SmartSMEAR data repository (https://avaa.tdata.fi/web/smart, last access: 22 August 2023; Junninen et al., 2009).

Author contribution

The main ideas were formulated by OG, HT, JKo, JVM, MS, MDM TR, TP, MK and the results were interpreted by MOK, OG, PP, and ME. TR, HK, OG, HT prepared measurement methodology and OG, MOI, JKa, HH, and HK contributed to data collection. MOK performed the data analysis and YZ supported it. OG and EMM supervised the project. HT, MDM, and TP made a funding acquisition. MOK visualized data and prepared the manuscript with contributions from OG, ME. All the authors reviewed and commented the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This research was supported by the Regional innovations and experimentations funds AIKO (project HAQT, AIKO014), Business Finland (CITYZER project, Tekes nro: 3021/31/2015 and 2883/31/2015), Pegasor Oy and HSY, Academy of Finland (grant nos 273010, 307331, 310626, 311932, 318940, 1325656, 326437, and ACCC flagship grant no. 337549, 337552, 337551), Healthy Outdoor Premises for Everyone (HOPE), Urban Innovation Actions, Regional development funds, The Technology Industries of Finland Centennial Foundation via project Urbaani ilmanlaatu 2.0, European Commission via Horizon Europe project “Non-CO2 Forcers and their Climate, Weather, Air Quality and Health Impacts, FOCl” (101056783), Faculty of Science 3-year grant (75284132), Tampere University of Technology graduate school, European Research Council (ERC) project ADAPT (grant no. 101002728), European Union Horizon 2020 research and innovation programme (grant no 101036245, RI-URBANS; grant no. 821205, FORCeS, and ERA-PLANET project SMURBS, 689443).

We would like to thank the people who took care of instruments and helped with measurements at the SMEAR III (Pekka Rantalaa, Erkki Siivola, Pasi Aalto, Frans Kortkomen, Tsia Laurila, Lauri Quéléver, Tuuli Lehmusjärvi, Deniz Kemppainen) and the HSY Mäkieläinkatu site (Anssi Julkunen, Anders Svens, Harri Portin, Taneli Mäkelä, Tommi Wallenius, Anu Kousa).

References


perturbed by NOx in eastern China, Atmos. Chem. Phys., 21, 14789–14814, https://doi.org/10.5194/ACP-21-14789-

Massoli, P., Stark, H., Canagaratna, M. R., Krehmjer, J. E., Xu, L., Ng, N. L., Mauldin, R. L., Yan, C., Kimmel, J.,
Misztal, 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 Sp. Chem., 2, 653–672,
https://doi.org/10.1021/acsearthspacechem.8b00028, 2018.

J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S. W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H.,
largest petrochemical source of urban organic emissions, Science (80-. )., 359, 760–764,

of highly oxygenated organic molecules from aromatic compounds, Atmos. Chem. Phys., 18, 1909–1921,

Mutzel, A., Zhang, Y., Böge, O., Rodigast, M., Kolodziejczyk, A., Wang, X., and Herrmann, H.: Importance of
secondary organic aerosol formation of is/i'-pinene, limonene, and im/i'-cresol comparing day- And nighttime radical

R., Zhang, Q., Sun, Y. L., and Jayne, J. T.: An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring
of the Composition and Mass Concentrations of Ambient Aerosol, 45, 780–794,

L., Fuchs, H., Griffin, R. J., Guzman, M. I., Herrmann, H., Hodzic, A., Iinuma, Y., Jimenez, J. L., Kiendler-Scharr, A.,
A.: Nitrate radicals and biogenic volatile organic compounds: oxidation, mechanisms, and organic aerosol, Atmos.

X., Kerminen, V.-M., Ehn, M., Donahue, N. M., Wang, T., Huang, C., Kulmala, M., Worsnop, D., Jiang, J., and Ding,
A.: Secondary organic aerosol formed by condensing anthropogenic vapours over China’s megacities, Nat. Geosci.

Nie, W., Yan, C., Yang, L., Roldin, P., Liu, Y., Vogel, A. L., Molteni, U., Stolzenburg, D., Finkenzeller, H., Amorim,
A., Bianchi, F., Curtis, J., Dada, L., Draper, D. C., Duplissy, J., Hansel, A., He, X. C., Hofbauer, V., Jokinen, T., Kim,
C., Lehtipalo, K., Nichman, L., Mauldin, R. L., Makhmutov, V., Mentler, B., Mizzieli-Ojani, A., Petaja, T., Quéléver,


