

UNIVERSITY OF TARTU
FACULTY OF SCIENCE AND TECHNOLOGY
Institute of Chemistry



Organic Constituents of Atmospheric Aerosols in a Hemi-boreal Forest
Master's Thesis (30 EAP)
Applied Measurement Science

Nieves Maria Flores March

Supervisor: Professor Heikki Junninen

Tartu 2022

Abbreviations

AMS Aerosol mass spectrometer

ASOA Anthropogenic Secondary Organic Aerosol

AVOC Anthropogenic Volatile Organic Compound

BC Black Carbon

BrC Brown Carbon

BSMs Biogenic small molecules

BSOA Biogenic Secondary Organic Aerosol

BVOC Biogenic Volatile Compound

CCN Cloud condensation nuclei

CI-API-TOF Chemical Ionization Atmospheric pressure interface Time-Of-Flight Mass Spectrometer

FIGAERO Filter Inlet for Gases and Aerosols

FTIR Fourier Transform infra-red spectrometry

GC-MS Gas Chromatography-Mass Spectrometry

LC-MS Liquid Chromatography-Mass Spectrometry

LV-OOA Low Volatility Oxygenated Organic Aerosol

NMR Nuclear magnetic resonance

OA Organic aerosol

OOA Oxygenated organic aerosol

PILS Particle-into-Liquid Sampler

PM Particulate matter

POA Primary Organic Aerosol

SOA Secondary Organic Aerosol

SV-OOA Semi-Volatile Oxygenated Organic Aerosol

TSP Total Suspended Particulate

VCP Volatile Chemical Products

VOC Volatile organic compounds

WISOM Water Insoluble Organic Matter

WSOC Water Soluble Organic Compounds

Abstract

Atmospheric aerosols have been demonstrated to be a highly dynamic system, playing a significant role in climate change and human health. In nature, ecosystems like boreal forests can modify the atmospheric particles producing a warming or cooling effect on climate. However, the regional and global impact of boreal forest on climate is still difficult to determine, especially due to the heterogeneous chemistry of aerosol samples, the need for multiple instruments for identification, and their limited library of compounds. To overcome these issues, in this thesis, we used a molecular networking technique based on the Global Natural Products Social web platform in combination with Nuclear Magnetic Resonance (NMR) to perform a screening of organic aerosols during the winter, spring, and summer seasons from a Hemi-boreal forest. The aerosol samples were recollected in a glass filter weekly from SMEAR Station (Estonia) and analyzed by Gas Chromatography-Mass spectrometry and NMR. A variety of chemical functional groups including carboxylic acids, phthalates, and organophosphate among the most abundant were annotated in the studied seasons. Furthermore, it was analyzed the presence of *n*-alkanol, carboxylic acid, and *n*-alkane to evaluate any hydrocarbon contamination. Phthalates-based compounds like Dibutyl phthalate (~20.59% in winter), and Bis(2-ethylhexyl) phthalate (~3.87% in summer), altogether with organophosphates like Tris(2,4-di-tert-butylphenyl) phosphate (~24.13% in spring) and tris(2,4-di-tertbutylphenyl) phosphite (~5.13% in summer) were annotated as a possible air pollutant. Besides that, the presence of conifer burning tracers such as 7-Oxodehydroabietic (~1.18% in spring) and dehydroabietic acid (~0.49% in summer) were annotated. These findings presented in this work give an insightful impact on the atmospheric aerosol composition presented in a Hemi-boreal forest using a straightforward and versatile technique such as molecular networking.

Keywords: Organic aerosols, GC/MS/EI, GNPS, NMR, Molecular networking, NMRfilter

Acknowledgment

Thanks to God, Dr. Heikki Junninen (my advisor), my Family (Carlos Flores, Hilda March, Juan Jose, and Juan Carlos), Franco Centurion, Kristo Hõrrak, Dr. Biswapriya Misra, Jing Guo, Ritudisha Biswas. Team members of Laboratory of Environmental Physics. Many thanks as well Dr. Lauri Toom, Dr. Eliise Tammekivi, and Professor Ivo Leito from the Chemistry department.

Table of content

Abbreviations	1
Abstract	2
Acknowledgment	3
Table of figures	6
Table of tables	7
Table of appendix	8
1. Introduction	9
1.1. Fundamental Concepts	11
1.1.1. Volatile Organic Compounds	11
1.1.2. Aerosols	12
1.1.2.1. Aerosol Chemical Composition	12
1.1.2.1.1. Black and Brown Carbon	12
1.1.2.1.2. Organic Aerosols	13
1.1.2.1.3. Secondary Inorganic Aerosols	13
1.1.2.1.3.1. Atmospheric Oxidants	14
1.1.2.2. Organic Aerosol Measurement Techniques	15
1.1.2.2.1. Online Methods	15
1.1.2.2.2. Off-line Methods	17
1.1.3. Molecular Networking Analysis	18
2. Materials and Equipment	20
2.1. Sampling Site	20
2.2. Sample Collection	20
2.3. α -Pinene Ozonolysis Experiment	21
2.4. GC/MS Analysis	23
2.5. NMR Analysis	24
3. Results	25
3.1. Molecular Networking Analysis of Interfering Compounds from the Background ..	25
3.2. Molecular Networking Analysis of α -Pinene Oxidation by Ozone	27
3.3. Molecular Networking of Atmospheric Aerosols in Summer, Winter, and Spring ...	30
3.4. NMR Analysis	34
3.5. Compound Identification on Two-dimensional (2D) NMR with NMRfilter	35
4. Discussion	36

4.1.	Molecular Networking Analysis of Interfering Compounds from the Background..	36
4.2.	Molecular Networking Analysis of α -Pinene oxidation by Ozone.....	37
4.3.	Molecular Networking of Atmospheric Aerosols in Summer, Winter, and Spring...	38
4.3.1.	Organophosphates.....	38
4.3.2.	Organosulfur compounds.....	38
4.3.3.	Phthalates	39
4.3.4.	Polycyclic Aromatic Hydrocarbons (PAHs).....	39
4.3.5.	Organosilicon Compounds.....	40
4.3.6.	Aliphatic compounds	40
4.3.6.1.	Alkanols	40
4.3.6.2.	Carboxylic acids.....	41
4.3.6.3.	Alkanes.....	42
4.3.7.	Organohalogen.....	42
4.3.8.	Terpenoids.....	43
4.4.	Analysis of ^1H -NMR Spectra.....	43
4.5.	2D NMR Analysis.....	44
4.6.	Compound Identification on Two-dimensional (2D) NMR with NMRfilter	46
5.	Conclusion.....	48
	References.....	49
	Appendix.....	58
	Non-exclusive licence to reproduce the thesis and make the thesis public	86
	Information Sheet.....	87

Table of figures

Figure 1. The processing pipeline and performance. a, Spectra is aligned and binned; noise is filtered and (b) baseline corrected. c, Common profile across the data set and peaks in rT dimension are aligned using FFT-accelerated correlation. d, Generation of both peak integrals for all samples and their common fragmentation patterns. e, Separation of overlapping peaks with patterns across samples using NMF. f, Peak integrals for all samples and canonical fragmentation patterns. NIST, National Institute of Standards and Technology. g, Annotation with public or private libraries. rTI, retention time index. h, Molecular networks. i, Data and results are shared between users. Adapted from ¹²	19
Figure 2. Location map of the sampling site in Järvelja, Tartu county, Estonia	20
Figure 3. Setup of aerosol sampling in SMEAR Estonia. (a) Normal setup. (b) Special setup for background sampling.	21
Figure 4. Setup of α -pinene ozonolysis. (a) For samples F3, F4. (b) For samples F2, F2*	22
Figure 5. Distribution of annotations in different chemical families found in background sample by GC/MS obtained from GNPS public library. MQScore suggests the chemical similarity and the MQScore value range from 0 to 1. These annotations have an MQscore between 0.5 to 0.99, this list of compounds exclude nodes also found in a blank sample.	25
Figure 6. Chemical component distribution in the liquid and particle-phase from the α -pinene oxidation obtained from the molecular networking analysis. MQScore suggests the chemical similarity and the MQScore value ranges from 0.5 to 1.	27
Figure 7. Molecular networking clusters from α -pinene oxidation. (a) Representative cluster #5 from the liquid phase of the α -pinene oxidation. (b) Representative cluster #7 from the particle phase of the α -pinene oxidation.	29
Figure 8. Comparison of the putative chemical families identified in the particle phase from winter, spring, and summer using the GNP public library. MQScore suggests the chemical similarity and the MQScore value ranges from 0 to 1.	30
Figure 9. <i>n</i> -Alkane distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.	32
Figure 10. <i>n</i> -Alkanol distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.	32
Figure 11. <i>n</i> -Carboxylic acid distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.	33
Figure 12. <i>n</i> -Iodoalkanes distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.	33

Table of tables

Table 1. Reaction conditions and particle size distributions.....	22
Table 2. Annotated compounds from aerosol particles in the background sample by GNPS public library with a normalized intensity >1%.....	26
Table 3. Chemical shift ranges of the regions defined from the 2D-NMR spectra according to Chalbot ⁵⁷	34
Table 4. NMRfilter analysis with HSQC and HMBC providing a matching rate and quality parameters ('distance' and 'standard deviation').....	35

Table of appendix

Appendix 1. Summary of sampling of aerosol particles and analysis conditions by GC and NMR	58
Appendix 2. Compounds annotated from aerosol particles from the background, by GC/MS obtained from GNPS public library	60
Appendix 3. Annotated compounds with a relative intensity higher than 1% identified from aerosol particles in winter, spring, and summer obtained from GNPS public library.	61
Appendix 4. The number of oxygen atoms vs the number of carbon atoms for each chemical formula in the liquid phase (Asolution2), and particle-phase (Afilter2, background and sample seasons in SMEAR Estonia). The blue and black lines in the figures correspond to O:C = 0.5 and O:C = 1 ratios, respectively, for reference.	63
Appendix 5. Analysis of the functional groups identified in the 700 MHz ¹ H NMR spectra of the CDCl ₃ extraction of blank and background samples. (a) Full ¹ H NMR spectra. (b) Zoom view from 5 to 9.3 ppm. (c) Zoom view from 1.8 to 5 ppm. (d) Zoom view from 0 to 1.8 ppm.	64
Appendix 6. Analysis of the functional groups identified in the 700 MHz ¹ H NMR spectra of the CDCl ₃ extraction of blank, background, and ambient aerosol samples.	65
Appendix 7. ¹ H- ¹ H TOCSY NMR spectra of the CDCl ₃ extraction of α-pinene experiment, background, and ambient aerosol samples.	66
Appendix 8. ¹ H- ¹³ C HSQC NMR spectra of the CDCl ₃ extraction of α-pinene experiment and ambient aerosol samples.	67
Appendix 9. ¹ H- ¹³ C HMBC NMR spectra of the CDCl ₃ extraction of α-pinene experiment and spring sample.	68
Appendix 10. HMBC and HSQC spectra with NMRfilter simulation	69
Appendix 11. Weight of filter collected in Smear Station	70
Appendix 12. Node annotations . Note: Functional group reported before derivatization.	71

1. Introduction

Atmospheric aerosols are always under scientific attention, mainly due to their heterogeneous chemistry and their impact on air quality, climate change, and human health. Aerosols, known as a particular matter, can be originated from anthropogenic and natural resources, and their main influence is based on the modification of the atmosphere's radiative balance ¹. The modification of the radiation balance can be performed directly *via* scattering or absorbing sunlight, producing a cooling/warming effect on the atmosphere, or indirectly by acting as cloud condensation nuclei (CCN); resulting in the modification of the cloud properties ². For instance, water in the gas phase does not condense readily on its own; its condensation is greatly expedited by the presence of a nucleus on which to condense ^{3,4}. Aerosol particles can also provide such surfaces in the atmosphere acting as cloud condensation nuclei. Consequently, the rise of cloud condensation nucleus generates a higher cloud droplet density which produces a higher albedo and therefore cools the atmosphere more effectively ⁵.

Forests are known to play a significant role in climate modulation, and this regulation is controlled by the surface-atmosphere exchange of energy, water, carbon dioxide, and aerosol species ⁶. Due to the great vegetation in forests, volatile organic compounds (VOC) can be oxidized and condensed, forming products with low enough vapor pressure which also condense on existing aerosol particles, creating secondary organic aerosol (SOA). Furthermore, considering the chemical composition of atmospheric aerosols, organic aerosol (OA) makes up a large fraction (20 to 90 %) of the total aerosol mass ⁷ and their exact chemical composition is remaining unknown.

Mass spectrometry (MS) has surged as a valuable tool to analyze a different variety of samples, providing identification from small molecules to complex biomolecules. However, a central problem for MS methods is the analysis of complex mixture and the full elucidation of molecular structures ⁸. To achieve a high correlation and better characterization, the combination of methods such as nuclear magnetic resonance spectroscopy (NMR), high-resolution mass spectrometry (HRMS), and chiroptical spectroscopic methods and vibrational circular dichroism (VCD) are typically used. Although all the previous methods are known to be effective, they can be slow and time-consuming ⁸.

Molecular networking (MN) analysis is an emerging tool used to cluster molecules into molecular families based on their MS spectra similarities⁹⁻¹¹. In particular, the combination of GC-MS data can be used as a valuable input for MN, providing benefits such as guiding the annotation

of the molecular family focusing more on the connected nodes than individual nodes, and cluster visualization. The use of the Global Natural Product Social Molecular Networking (GNPS) platform can facilitate the performance, sharing of MN analyses and provide freely available reference data of more than 19,808 spectra ¹².

In this context, given the importance of knowing the chemical composition of atmospheric aerosols and their impact on climate change, this thesis will first review some fundamental concepts in atmospheric chemistry. Then, the author will use the GNPS platform and NMR analysis for the annotation of organic compounds found in atmospheric aerosols from a Hemi-boreal Forest, located at the Järvelja Experimental Forestry station in Estonia. Finally, the author will present the insightful implications of this study.

1.1. Fundamental Concepts

1.1.1. Volatile Organic Compounds

Volatile Organic Compounds (VOC) are carbon-based compounds that have a boiling point less than or equal to 250 °C measured at a standard atmospheric pressure of 101.3 kpa¹³. Excluding the emission of methane, global VOC emissions are ~10% anthropogenic and ~90% biogenic¹⁴. However, anthropogenic VOCs are often dominant in urban areas due to the industrial emissions which typically include a complex mixture of solvents, waste, and other resources¹⁵.

Anthropogenic VOCs sources include traditional emissions (transportation) and volatile chemical products (VCPs). As such, VCPs are a category of VOCs that share two important features. The first one is the ability of chemical products to be evaporated from them, and the second feature is to carry or to cause a residue to stick to a surface. Examples of VCPs include personal care products, pesticides, coatings, inks, adhesives, and cleaning agents, as well as cooking emissions, asphalt emissions, and solid-fuel emissions from residential wood burning or cookstoves¹⁶. Furthermore, anthropogenic VOCs are also known to interact with other molecules such as oxides of nitrogen (NOx). Consequently, when these molecules are in the presence of sunlight, a harmful molecule like ozone is produced at the ground level. To control the resulting air pollutants, regulating agencies such as U.S. Environmental Protection Agency (EPA) have implemented policies to control man-made VOC emissions.

On the other hand, Biogenic Volatile Organic Compound (BVOC) includes a broad spectrum of atmospheric hydrocarbons (excluding methane) that are emitted from terrestrial and marine ecosystems. For instance, terpenes are the largest and most structurally diverse class of volatiles produced by plants. Therefore, these types of VOCs like terpenes are used to attract beneficial organisms, repel potential herbivores, and protect themselves against pathogens¹⁷. Among the terpenoid family, isoprene (a hemiterpene, containing five carbon atoms) and monoterpenes (10 carbon atoms) like α -pinene, β -pinene, and Δ^3 -Carene are the most abundant BVOC¹⁸. Besides that, it is also common the presence of sesquiterpenes (15 carbon atoms), and diterpenes (20 carbon atoms). In particular, the high reactivity of sesquiterpenes allows them to efficiently contribute to aerosol formation compared to other BVOCs¹⁹.

1.1.2. Aerosols

Aerosols are defined as a suspension of fine solid or liquid particles in gas and are also known as particulate matter (PM). The size of atmospheric aerosols is ranging from a few nanometers to tens of micrometers in diameter⁵. Because of this nano-micro meter size, aerosols can penetrate deep into the lungs of people and have been linked to severe short- and long-term health effects²⁰. Moreover, when aerosols interact with both solar and terrestrial radiation, aerosols have a cooling effect overall by scattering and reflecting the sunlight into space²¹. However, aerosols can also absorb part of the incoming sunlight. As an example, black carbon is a strong sunlight absorber and therefore contributes to global warming²².

Among the classification of aerosols, primary aerosols are those which are directly emitted as liquids or solids from sources such as biomass burning, incomplete combustion of fossil fuels, volcanic eruptions, soil, road dust, mineral dust, sea salt, and biological materials (plant fragments, microorganism, pollen, etc). Unlike primary aerosol, secondary aerosols are not emitted directly in the particulate phase but come from the condensation of atmospheric gas-phase species²³.

1.1.2.1. Aerosol Chemical Composition

The study of the chemical composition of aerosol particles is an essential factor to understand the aerosol interaction with climate change, atmospheric environment, and human health. In this section, the three main components of particulate matter such as black and brown carbon, organic and inorganic aerosols are presented.

1.1.2.1.1. Black and Brown Carbon

Black carbon (BC) or refractory carbon cores, refers to particles primarily composed of elemental carbon and characterized by broad absorbance across the visible spectrum from ultraviolet to infrared wavelengths²⁴. BC is typically produced as a result of incomplete combustion of fossil fuels, biofuels, and biomass. Additionally, the amount of organic material compared with BC produced seems to depend on the type of burned material and the temperature of combustion²⁵. In addition, BC is also known as the strongest absorber of solar radiation of all the constituents in the atmosphere²², playing a crucial role in global warming. Like BC, brown carbons (BrC) are also light-absorbing organic aerosols and account for ~20% of aerosol-driven atmospheric heating^{26,27}. In particular, BrC is respired by microbes and then reintroduced to the atmosphere as CO₂, contributing to global warming as well. Furthermore, atmospheric BrC can be accumulated on the surface of the snow, diminishing the snow albedo and accelerating snowmelt²⁸.

1.1.2.1.2. Organic Aerosols

The organic fraction of PM is referred to as organic aerosols (OA). Primary OA (POA) are directly emitted from sources such as fossil fuel combustion, plant debris, fungal spore, and biomass burning²⁹. Meanwhile, secondary organic aerosols (SOA) are designated as particles that have not emitted directly in the particulate phase but are produced by homogenous reactions of biogenic and anthropogenic VOCs with ozone (O_3), hydroxyl ($OH\cdot$), nitrate ($NO_3\cdot$) radicals³⁰, and by autoxidation in the atmosphere.

Biogenic secondary organic aerosol (BSOA) refers to SOA produced from the photooxidation of BVOCs and their importance lies in the function of acting as cloud condensation nuclei (CCN). Besides that, the contribution of BSOA to the urban concentration of submicron aerosols (PM_{10}) is accounted to be from 1 to 20%¹⁶. Unlike BSOA, anthropogenic secondary organic aerosol (ASOA) refers to SOA produced from the photooxidation of AVOCs and constitutes a substantial fraction of the mass of PM_{10} in cities around the world. ASOA particles are a significant indirect source of deadly fine particles¹⁶, and its contribution of ASOA to the urban concentration of non-refractory submicron aerosol (PM_{10}) is ranging between 17 to 39%¹⁶.

Overall, when organic aerosols age in the atmosphere, they become more oxygenated, less volatile, and more hygroscopic²⁹. In consequence, the formation of organic peroxide compounds is favored. Organic peroxide compounds have emerged as an important particle-phase constituent of secondary organic aerosol (SOA), and they are considered as components of a family of compounds denoted as highly oxygenated molecules (HOMs) which play a leading role in new particle formation³¹.

1.1.2.1.3. Secondary Inorganic Aerosols

As it was described earlier, PM can be classified as primary and secondary aerosols. The generation of the latter fraction (secondary aerosols) is the result of physical and chemical processes involving reactions between SO_2 , NO_x , NH_3 , and several VOCs. Therefore, both secondary organic and inorganic aerosols (SIA) are generated. In particular, sulfate (SO_4^{2-}), nitrate (NO_3^-), and ammonium (NH_4^+) are the major SIA components in fine PM fractions. For example, It was reported that SIA is the predominant component of fine particles ($PM_{2.5}$) in China, making up about 30-40% of total $PM_{2.5}$ mass annually³², and it is estimated that the composition of urban PM_{10} worldwide is being dominated by sulfate (SO_4^{2-} , 6 to 48%), nitrate (NO_3^- , 1 to 30%) and ammonium (NH_4^+ , 8 to 24 %) ¹⁶. Unlike SIA, primary inorganic aerosols are generally very

hygroscopic and have a higher backscatter solar radiation than SIA²³. For example, primary sulfate aerosols can contribute to the net cooling of the Earth's atmosphere due to their ability to backscatter light^{33,34}.

1.1.2.1.3.1. Atmospheric Oxidants

Given the wide heterogeneity of chemical compounds in the atmosphere, the primary pollutants like CO, NO_x, and VOCs are transformed by the presence of atmospheric oxidants such as hydroxyl radicals (OH·), nitrate radicals (NO₃·), and ozone (O₃). Due to the high reactivity of these oxidants, the resulting products lead to the formation of a series of secondary pollutants like SOA^{35,36}.

Hydroxyl radicals are powerful oxidants that can have the main impact on the daytime oxidation capacity of the atmosphere³⁶, initiating and participating in many oxidation reactions. For instance, hydroxyl radicals can react by adding OH groups to or abstracting protons to double bonds (C=C), depending on the structure of the organic molecule^{15,35}.

When the OH· levels are reduced because of the absence of ozone photolysis during the nighttime, the nitrate radical (NO₃·) is formed from the slow reaction between nitrogen dioxide and O₃ (NO₂ + O₃ → NO₃· + O₂). Note: This reaction occurs in the daytime as well but is photolyzed, reducing its concentration. As a result, the presence of the nitrate radical is considered the main initiator of nighttime oxidation chemistry in the troposphere³⁷. Moreover, NO₃· can react effectively with unsaturated non-methane VOCs, such as certain alkenes or aromatics *via* additions to double bonds, which can initiate the formation of peroxy radicals (HOO· and ROO·) and even OH·³⁷. Another nitrogen oxide that is relevant for air pollution is nitrogen dioxide (NO₂). This compound is typically emitted in small quantities from combustion processes along with nitric oxide (NO) and is also formed in the atmosphere by the oxidation of NO.

Ozone (O₃) is a reactive oxidant gas produced naturally in trace amounts in Earth's atmosphere⁵ and plays a major role due to the high mixing ratios in clean or contaminated atmospheres³⁸. In particular, tropospheric ozone is considered an air pollutant, causing respiratory effects in humans and affecting plant growth. In addition, some ozone is also transported down to Earth's surface from the stratosphere. Along with reactions with OH· and NO₃·, ozone is also responsible to oxidize atmospheric gases in the troposphere. Generally, O₃ reacts only with alkenes, adding to a C=C bond to produce a primary ozonide, which then decomposes to form a carbonyl plus a Criegee intermediate^{15,35}. This organic reaction is known as ozonolysis and is often the most efficient

oxidation mechanism of unsaturated BVOCs, which leads to the formation of products with lower volatility (extremely low volatile organic compounds ELVOC)³⁹. For example, α -pinene (BVOC with an endocyclic double bond) appears to produce ELVOC with a much higher molar yield from ozonolysis (~7%) than from the OH radical reaction (<1%)³¹.

In summary, the oxidation mechanism of VOCs in the atmosphere can be simplified as the following: 1) initial attack of the VOCs by oxidants (OH \cdot , O₃, NO₃ \cdot), 2) organic peroxy radicals reactions, and in some cases, 3) alkoxy radical reactions. Organic peroxy radicals can react with other species in the atmosphere (*e.g.* NO, NO₂, HO₂, etc) and undergo functionalization or form alkoxy radicals. Alkoxy radicals can be fragmented and form smaller organic compounds in the atmosphere that can be oxidized further. This fragmentation process leads to increased volatility whereas functionalization decreases volatility and increases solubility in water⁴⁰. These complex, multigenerational, gas-phase oxidation processes result in the formation of a large variety of organic compounds, which can undergo gas-particle partitioning and/or nucleation to form SOA⁴¹.

1.1.2.2. Organic Aerosol Measurement Techniques

As it was presented before, the chemistry of aerosol particles remains a complex process that involves multiple oxidation reactions, transforming a single molecule into thousands of oxidized products. Hence, the atmospheric organic species display a wide range of physical and chemical properties. Due to the chemical complexity of aerosol particles, a wide range of techniques have been developed and used showing advantages and limitations. In the next section, a brief description of current online and offline methods for the determination of atmospheric aerosol is presented. Additionally, a sub-section is dedicated to Molecular Networking for the analysis of GC/MS data.

1.1.2.2.1. Online Methods

Online (*in situ*) methodologies can analyze the sample at the same time that the particles are being collected. Typically, they are used in-field observations to provide information about the chemical and physical properties of OA on short timescales (one hour or less) and near real-time data. Based on their high sensitivity to environmental changes, online methods can be set up in a wide range of platforms like cars and airplanes. However, these methods have drawbacks like low fragmentation capabilities and complex calibration processes for organic compounds⁴².

Due to the high complexity of aerosol samples, the use of mass spectrometric techniques has been applied to detect a large range of chemical species, providing a better understanding of the atmospheric chemical composition. Among the main online techniques based on mass spectrometry, the most used techniques include Aerosol mass spectrometry (AMS), Atmospheric pressure interface Time-Of-Flight Mass Spectrometry (APITOF), The Filter Inlet for Gases and Aerosols (FIGAERO), and Proton Transfer Reaction-Time-Of-Flight mass spectrometry (PTR-TOF).

Aerosol mass spectrometry (AMS) is an online measurement technique that provides fast acquisition time. AMS is designed to obtain the size-resolved real-time composition of non-refractory submicron aerosols, which relies upon thermal vaporization at 600°C and 70 eV electron-impact ionization (EI) in a vacuum^{43,44}. This technique has a vast number of applications, providing information about the relationship between aerosols and atmospheric chemistry, emissions sources, human exposure to pollutants, radiative transfer, and cloud microphysics⁴⁵.

Mass spectrometry methods using a chemical ionization source like Nitrate Chemical Ionization – Atmospheric pressure interface – Time-Of-Flight Mass Spectrometer (Nitrate-CI-API-TOF) has been used for identification of HOMs in the gas phase^{31,46,47}. This technique provides information about the molecular mass but not fragmentation. In principle, the nitrate ions are specific to highly functionalized organic compounds (featuring at least two suitable located hydroperoxy groups) which makes them ideal reagent ions for the observation of HOM and their radical precursor RO₂. However, the detection efficiency of this method decreases dramatically for less oxidized organic compounds having higher volatility⁴⁸.

More recent techniques allow the analysis of semi-continuous gas-phase and particle-phase measurements of organic aerosol samples. Techniques like Filter Inlet for Gases and Aerosols (FIGAERO) coupled to a High-Resolution Time of Flight – Chemical Ionization Mass Spectrometer (HRTof-CIMS), can collect information from gas and particulate OA phases using programmed thermal desorption of particles collected on a Teflon filter with subsequent detection and speciation of desorbed vapors⁴⁹. As the FIGAERO particle-phase analysis is a filter-based method, the instrument can also be used in an offline mode, where filter samples are collected elsewhere and then analyzed at a larger stage in the laboratory. Over the last years techniques like proton transfer reaction-time-of-flight mass spectrometry (PTR3) have been developed to detect

aerosol formation precursors, like VOCs and their oxidation products (HOMs), the PTR3 can also help to understand the formation of SVOC, LVOC, and even ELVOC ⁴⁸.

1.1.2.2.2. Off-line Methods

Offline methods are based on the collection of aerosols on a surface, or into a liquid, and then analyzed by a specific technique. Due to the low concentration (trace levels) of organic aerosol samples, the most typical offline methods are based on liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS). Those techniques are known to provide separation resolution, chemical identification, and sensitivity compatible with aerosol particles ⁵⁰.

The most widely used technique for the characterization of SOA from complex mixture aerosols is GC-MS, which is utilized as a powerful separation and analytical technique. The GC-MS is usually used with electron ionization (EI) or chemical ionization (CI) detectors. Additionally, derivatization for GC analysis is done via three different reactions, *i.e.*, acylation, alkylation, and silylation ⁵¹. Although GC/MS is a robust technique for quantitative characterization of organic molecules in aerosol samples, it is limited to small, thermally stable molecules ⁵².

Similarly, LC-MS is used with atmospheric pressure chemical ionization (APCI) and electrospray ionization (ESI) detectors in positive and negative ion modes, which is used for the characterization of thermally unstable, high molecular weight, and highly polar compounds (41).

Other offline techniques are based on Fourier transform infrared (FTIR) spectroscopy, nuclear magnetic resonance (NMR), and Ion chromatography. Fourier transforms infrared (FTIR) spectroscopy is used for the detection of functional groups in compounds. One of the major advantages of this technique is that it can measure some functional groups, *e.g.*, amines and organic sulfates, which are difficult to measure using other techniques ⁵³.

NMR identifies functional groups and substructures present OA ⁵⁴, this technique has been used for the study of the water-soluble fraction (WSOM). The most widely used NMR method in the analysis of aerosol WSOM is the one-dimensional (1D) liquid state ¹H NMR. However, it is often characterized by the presence of strong overlapping signals ⁵⁵. Another set of NMR-based multidimensional methodologies for the structural analysis of WSOM is the 2D liquid-state NMR techniques, these solve the spectral information overlaid in 1D NMR spectra by separating the data into a second frequency dimension ^{56,57}. Two-dimensional NMR spectra are unique in that they provide direct evidence for atom connectivity and spatial arrangements.

Ion chromatography methods are most often used for the measurement of inorganic ions such as SO_4^{2-} , NO_3^- and metals. Some advantages are minimal sample preparation, inexpensive separation, and high sensitivity for multiple ion analysis. In particular, this method was used for the simultaneous determination of methylamines (methylamine, dimethylamine, trimethylamine) and trimethylamine-N-oxide in particulate matter air samples by IC with non-suppressed conductivity detection⁵⁸. Another use associated with IC includes the detection of sulfate and hydroxymethanesulfonate⁵⁹. Characterization of low molecular weight organic acid in atmospheric aerosols using IC-MS/MS⁶⁰.

1.1.3. Molecular Networking Analysis

Molecular networking (MN) is a computational approach that organizes MS/MS data based on spectral similarity. This approach allows the identification of similarities among MS/MS spectra within a database, establishing correlations between related molecules. Hence, the correlated structural information is represented in a network where unknown but structurally related molecules are identified and dereplicated. Common examples where MN analysis is used include metabolite identification and toxicological screening^{61,62}.

The Global Natural Product Social (GNPS) platform was used to analyze MS data obtained from GC. The visualization of molecular networks in GNPS represents each spectrum as a node, and spectrum-to-spectrum alignments as edges (connections) between nodes. Nodes can be supplemented with metadata, including dereplication matches. In addition, the user can also include information such as abundance, the origin of the product, biochemical activity, or hydrophobicity which can be reflected in a node's size or color. This map of all related molecules is visualized as a global molecular network⁶³.

The molecular networking for GC is created through spectral similarity of the deconvoluted fragmentation spectrum. Molecular networking patterns for the electron impact ionization data can further guide the annotation at the molecular family level by using information from connected nodes rather than focusing on individual annotations¹². The MS structural similarity with library spectra is typically expressed by the cosine scores⁶³, which refers to a mathematical measure of spectral similarity between two fragmentation spectra. A cosine score (MQScore) of 1 represents identical spectra while a cosine score (MQScore) of 0 denotes no similarity at all. In the present study, the aim was to construct a training network from the *in-situ* generated α -pinene oxidation

products and for the untargeted analysis of organic constituents of atmospheric aerosols in a Hemi-boreal Forest.

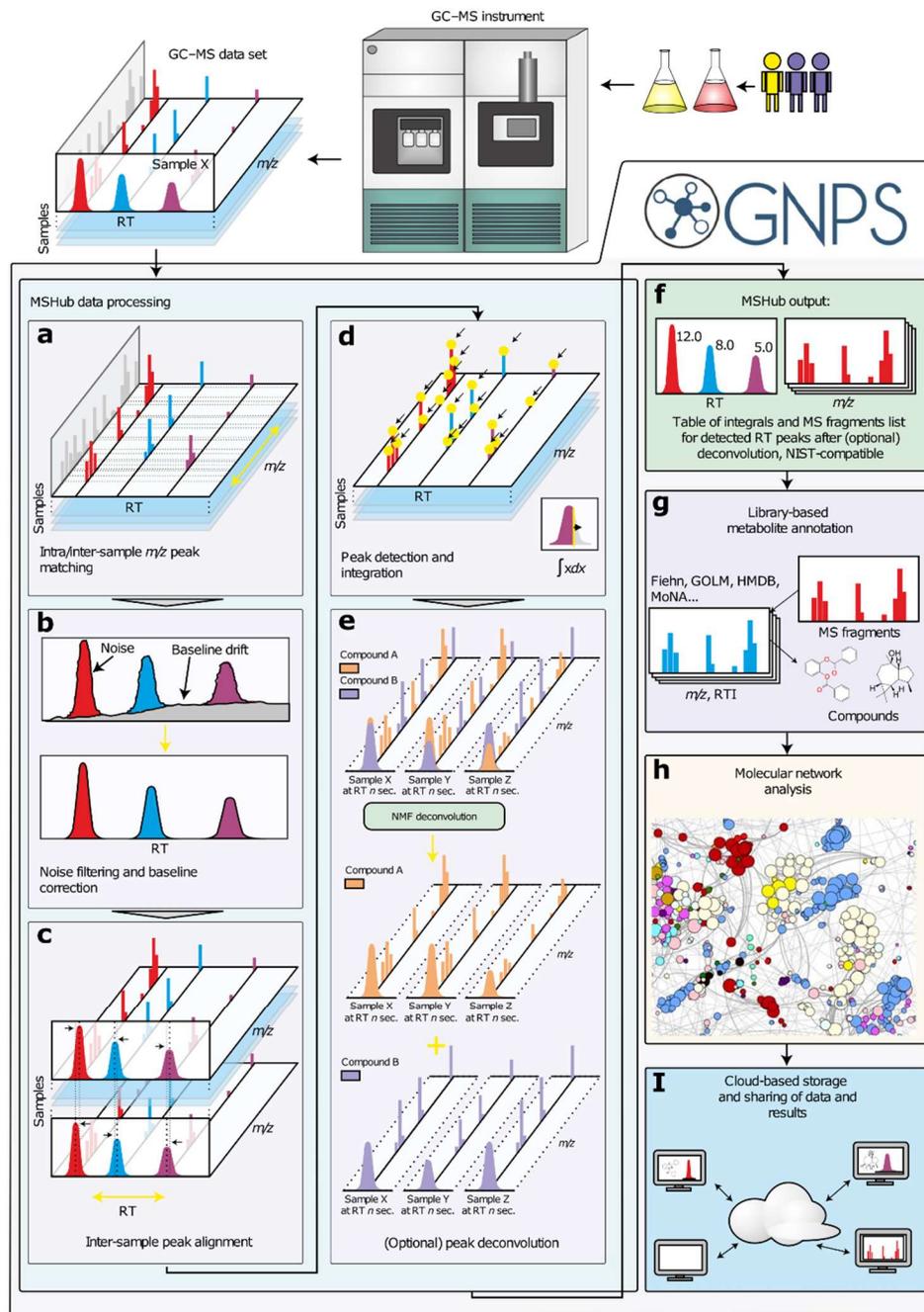


Figure 1. The processing pipeline and performance. a, Spectra is aligned and binned; noise is filtered and (b) baseline corrected. c, Common profile across the data set and peaks in RT dimension are aligned using FFT-accelerated correlation. d, Generation of both peak integrals for all samples and their common fragmentation patterns. e, Separation of overlapping peaks with patterns across samples using NMF. f, Peak integrals for all samples and canonical fragmentation patterns. NIST, National Institute of Standards and Technology. g, Annotation with public or private libraries. rTI , retention time index. h, Molecular networks. i, Data and results are shared between users. Adapted from reference¹².

2. Materials and Equipment

2.1. Sampling Site

Aerosol cottage in SMEAR Estonia station (a.s.l.) is situated in a mixed, Hemi-boreal forest stand comprising Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), silver birch (*Betula pendula*), and downy birch (*Betula pubescens*)⁶⁴.

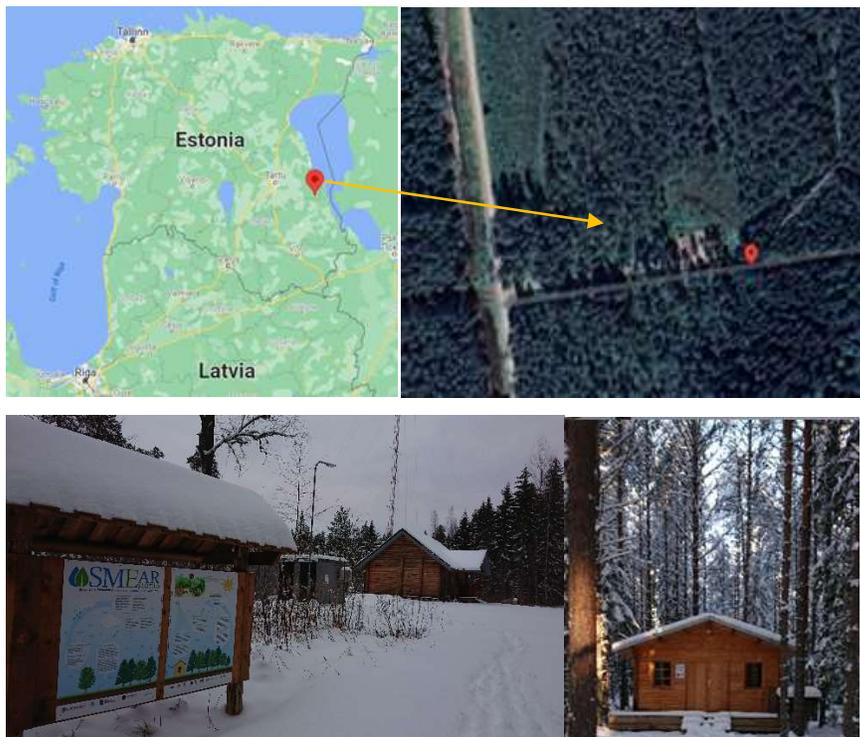


Figure 2. Location map of the sampling site in Järvelja, Tartu county, Estonia

2.2. Sample Collection

Aerosols samples were collected in pretreated (4h, 200 °C) glass microfiber (GM) filters (38 cm², Whatman No 1820-070). The GM filters were weighed before and after sampling to determine the total suspended particles (TSP) from the mass loadings. Then, the samples were packed in aluminum foils (foils were pretreated at 200 °C for 4h).

The aerosols obtained for winter, summer, spring, and the background sample were collected with an airflow rate of 50 L min⁻¹ (Note: filters were changed every week [Appendix 11]). Figure 3a shows the setup used for winter, summer, and spring samples. The sample collection for daytime and nighttime shares an inlet, with the purpose of maintain similar sampling conditions (Figure 3a). Timers were used to switch pumps on and off for daytime and nighttime sample collection, collection times were from 6-18 for daytime and 18-6 for nighttime samples. In addition, a blank

sample was also performed which consisted of the filter in contact with the filter holder for 5 min without any pumping system. This measurement was carried out to verify possible artifacts due to the adsorption of constituents into glass filters during and after sampling. It was verified that sample treatment and transportation did not cause any artifact and contamination. Also, a background sample was done sampling one week in daytime shift (Appendix 1, Figure 3b). All the constituents of the whole system include timers, pumps, needle valves, a sampling holder, a glass filter for every pump, tubing, and copper tubing for the inlet. The dates of sample collection are reported in Appendix 1.

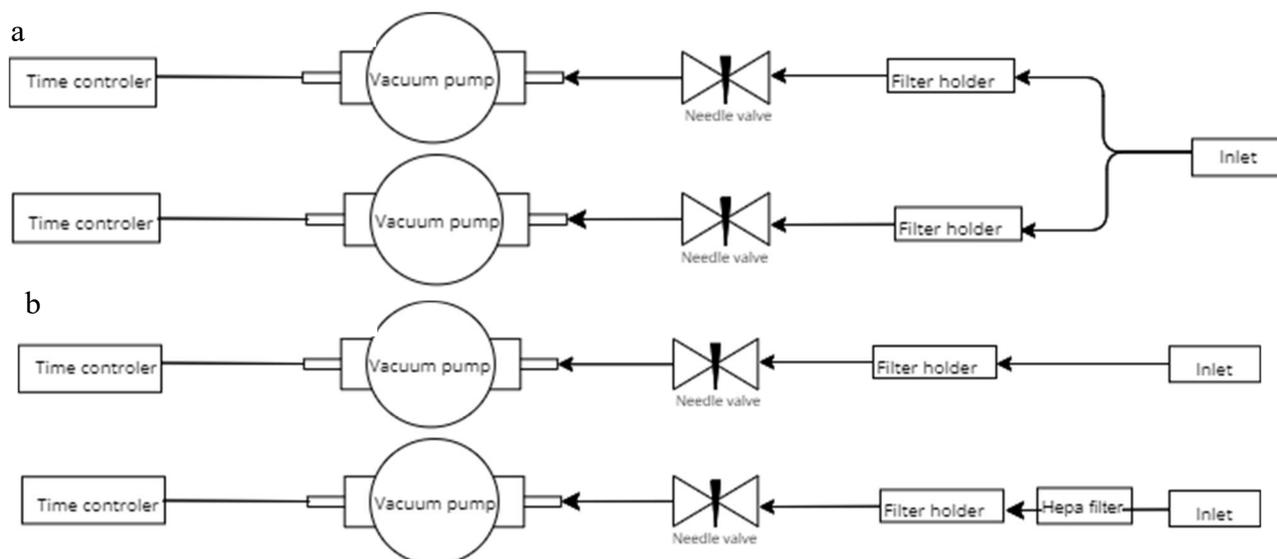


Figure 3. Setup of aerosol sampling in SMEAR Estonia. (a) Normal setup. (b) Special setup for background sampling.

2.3. α -Pinene Ozonolysis Experiment

A stainless-steel ultrahigh vacuum chamber (2 inches inner diameter and height $\sim 4/5$ inches) was used to study the *in-situ* formation of particles from α -pinene ozonation. Ozone was generated with a commercial generator (Mag-Pro), and the α -pinene vapor was collected by passing an airflow over the liquid (α -pinene). To transport all the produced aerosols, a filtered (using HEPA filter) air-dried was used. Particle size distributions were measured with a Fast mobility particle sizer spectrometer (FMPS, model 3091). Subsequently, a set of experiments were designed to examine the liquid and particle phases of the oxidation products of α -pinene, and were defined as follows:
 Filter F2: Oxidation products of α -pinene in the **particle phase**, GC analysis (Figure 4b).

Solution F2*: Oxidation products of α -pinene in the **liquid phase**, GC analysis (Figure 4b).

Filter F3: Oxidation products of α -pinene in the **particle phase**, NMR extraction with D_2O water (Figure 4a).

Filter F4: Oxidation products of α -pinene in the **particle phase**, NMR extraction with CDCl_3 (Figure 4a).

Table 1. Reaction conditions and particle size distributions

Experiments	Temperature (°C)	Mean particle diameter (nm)	α -pinene (mL)	Dry air flow (L/min) ^a	Ozone flow ^b (L/min)	Reaction time (h)	dN/d(logdp) at mean particle diameter ^c ($\times 10^6$)	Aerosol concentration (10^6 particles/cm ³)
Filter F2 Solution F2*	25.8±0.0 4	191.1	(0.6) x 3	8.05 + 8.68	7.74	4.8	3.37 ± 0.1	1.40 ± 0.07
Filter F3	26.8 ± 0.4	107.5	2.5	15.5 + 1.8	10	3.5	3.5 ± 0.6	1.50 ± 0.2
Filter F4	27.0 ± 0.4	107.5	2.5	15.5 + 1.8	7	3	2.2 ± 0.4	0.87 ± 0.1

^a Dry airflow includes two air inlets

^b Ozone Generation: 2000 mg/h

^c Particle size distributions were measured with a Fast mobility particle sizer spectrometer (FMPS, model 3091). The mass lost to the walls is not considered, mass and number concentrations in this work are not corrected.

* α -pinene in Eppendorf tubes of 0.6 mL

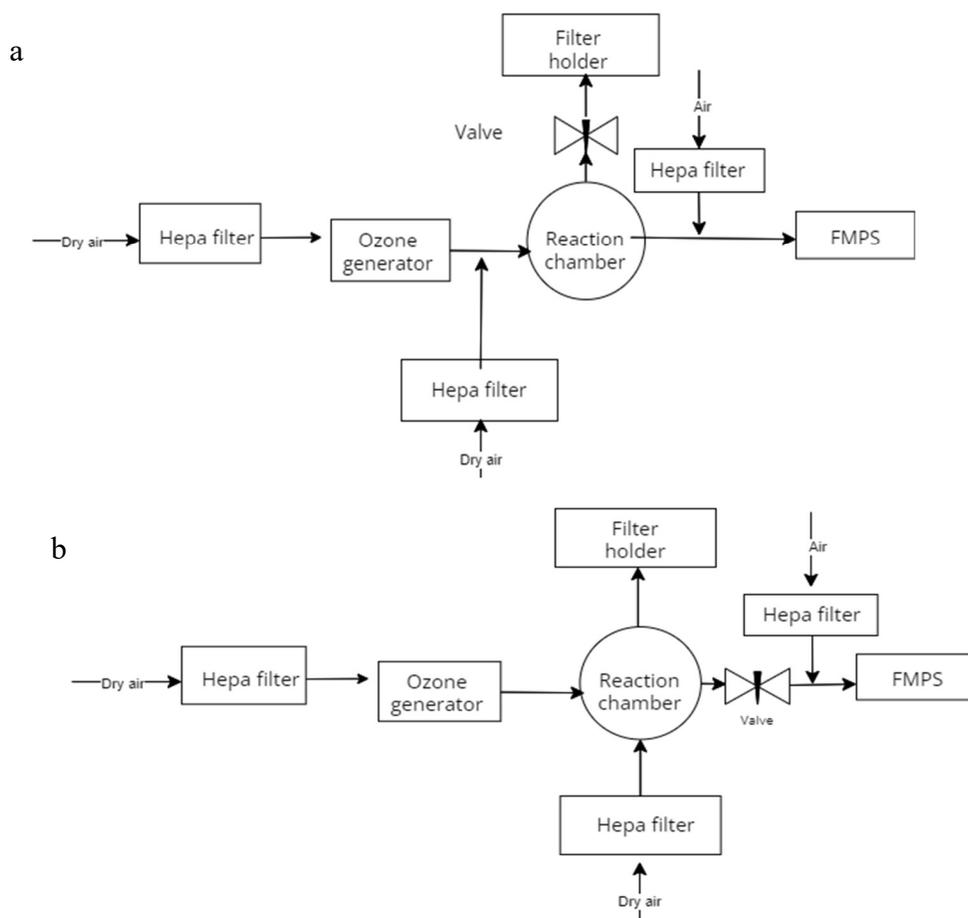


Figure 4. Setup of α -pinene ozonolysis. (a) For samples F3, F4. (b) For samples F2, F2*

2.4. GC/MS Analysis

Instrument Setup

The GC-MS analysis for the identification of compounds was carried out in a Hewlett Packard 5890 series II GC, with ionization achieved by electron impact at 70 eV. The capillary column used was a J&W Scientific DB-5 (30 m x 0.25 mm ID, 0.25 μm thick film). A 5 μL sample was injected into the GC injector split less. The GC instrument is equipped with an Automatic Sampler HP 6890 series injector, carrier gas helium with the constant flow of 1 mL/min. The oven was programmed initially from 270 $^{\circ}\text{C}$ to the final temperature of 315 $^{\circ}\text{C}$ with a 1.25 $^{\circ}\text{C}/\text{min}$ ramp. The final temperature hold time was 22 min. The inlet, MS transfer line, and ion source temperatures were kept at 250 $^{\circ}\text{C}$, 250 $^{\circ}\text{C}$, and 230 $^{\circ}\text{C}$ respectively. The mass spectra of aerosol components were identified in scan mode within an m/z range of 30 to 500, and the scan time started at 9.5 min. Then, some injected samples were firstly derivatized with *m*-(trifluoromethyl)phenyltrimethylammonium hydroxide (TMTFTH) reagent. Information about which samples were derivatized is shown in Appendix 1. This derivatization procedure increased the volatility of the compounds and produced methyl esters ⁵¹.

Data analysis: Molecular Networking analysis

Raw data obtained from GMPS of forest samples, blanks, background, and α -pinene experiments were firstly converted to CDF format and subsequently converted to mzXML format using Mass ++ 2.7.5 (build Jun 11, 2015). Later, the data was converted to mzML format using Proteowizard MSConvert. These data files were uploaded on GNPS by WinSCP (version 5.19.2).

The Molecular networking analysis was performed using the GNPS platform (<http://gnps.ucsd.edu>), and library annotations were obtained from the comparison between the MS spectra with several spectral libraries, including GNPS ¹² and NIST 98. Additionally, the IUPAC name of each compound was assigned, and each structure was classified following a standardized chemical taxonomy algorithm and analyzed in two computer programs (ClassyFire and NP Classifier) ^{65,66}. Next, all the links between nodes were created when the cosine score was higher than 0.7. Subsequently, the resulting spectral network was uploaded into Cytoscape 3.8.2⁶⁷ to obtain the Molecular network visualization, the nodes from the MN were labeled with name and the thickness of the edges is proportional to the cosine score. Nodes were colored in different colors according to the group where the precursor was detected. All the software programs used in these steps are open source and can be accessed freely online.

2.5. NMR Analysis

The NMR spectra were recorded on a Bruker spectrometer operating at 700 MHz for ^1H which used a 5 mm probe at room temperature. Tetramethylsilane (TMS) was used as internal standard 0.03% (v/v) for calibration purposes (0 ppm). The ^1H NMR spectra were signaled in 5 different ranges^{68,69}: (i) H-C (δH 0.6-1.8 ppm) that included R-CH₃, R-CH₂, and R-CH protons; (ii) H-C-C= (δH 1.8-3.2 ppm) included protons adjacent to a double bond and amines (H-C-NR₂); (iii) H-C-O (δH 3.2-4.4 ppm) contained alcohols, ethers, and esters; (iv) O-CH-O and H-C (δH 5.0-6.4 ppm) included anomeric protons of carbohydrates and olefins; and (v) H-Ar (δH 6.5-8.3 ppm) contained aromatic protons (Appendix 5 and 6). 2D-NMR spectra were obtained from the samples indicated in Appendix 1.

The TOCSY, HSQC, or HMBC spectra were collected with 90 and 170 scans and pulse sequences of mlevph, hsqcedetgpsisp2.3, and hmbcgpplndqf, respectively. Post-NMR data processing was conducted in MestReNova software.

Smiles strings of selected molecules, list of HSQC and HMBC peak of spring and α -pinene sample were used as input for an open-access program called NMRfilter⁷⁰ which was developed to quantify the similarity of the predicted spectra of related structures to the measured spectra of the mixture sample.

3. Results

3.1. Molecular Networking Analysis of Interfering Compounds from the Background

To understand the composition of aerosol samples in the studied seasons, it was necessary to determine first the composition of background samples. In this work, the dereplication of the CDCl_3 extracts from the molecular network GNPS was performed to annotate the compounds. The analysis of the annotated compounds found in the background samples are considered as interfering compounds from the system and therefore are presented.

The molecular networking analysis of GC-MS identified 162 nodes in the “Background”, representing more than ~67% of total relative intensity. Among all the nodes, 121 different compounds were annotated, and 20 unknown compounds were estimated. Based on relative intensity, the GC-MS of the “Background” showed a high concentration of carboxylic acids compounds, constituting ~20% for the particle phase as shown in Figure 5. Besides that, the second-largest group of background constituents was esters with ~17%, while organosulfur compounds, phthalates, and alcohol contributed less than ~24%.

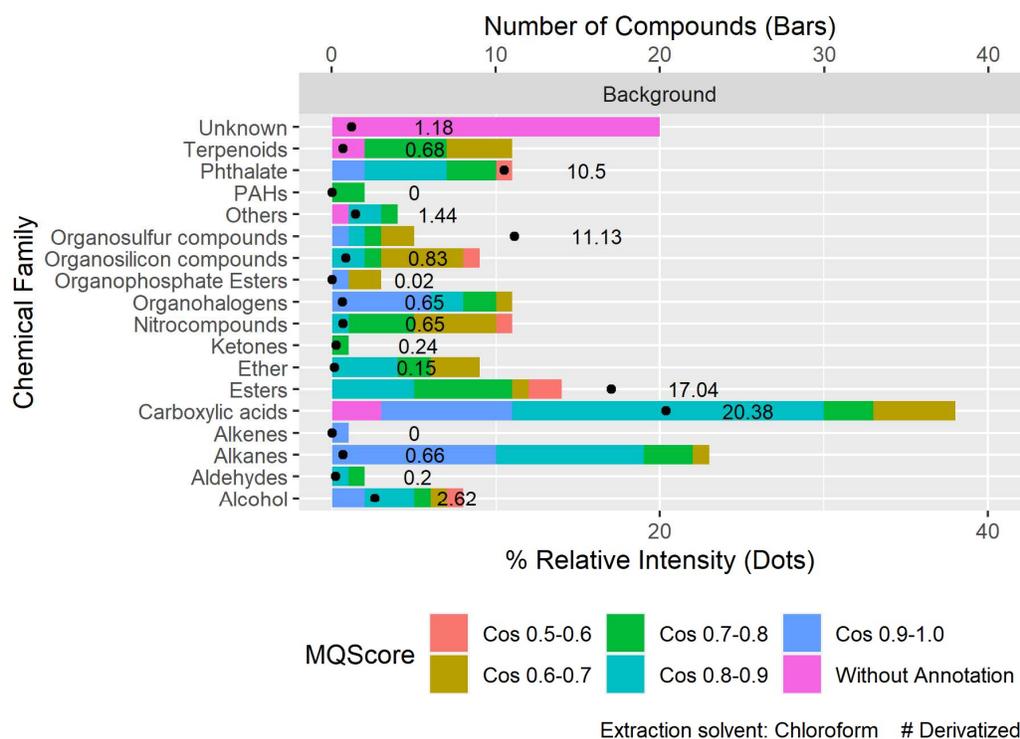


Figure 5. Distribution of annotations in different chemical families found in background sample by GC/MS obtained from GNPS public library. MQScore suggests the chemical similarity and the MQScore value range from 0 to 1. These annotations have an MQScore between 0.5 to 0.99, this list of compounds exclude nodes also found in a blank sample.

To determine the probability of the annotations, the MQScore analysis was used. As it was found in Figure 5, carboxylic acids (71% of annotations) have a high probability of being well-annotated, with the MQScore between 0.8 to 1. Interestingly, alkenes (100% of annotations) showed a high MQScore between 0.9 to 1, and alkanes (85% of annotations) showed an MQScore between 0.8 to 1. However, the relative intensity for these compounds is considered minimal (<1%).

After the chemical family identification, a set of chemical compounds were annotated using the GNPS public library as is shown in Table 2. The compounds were grouped considering the MQScore range between 0.6 to 1. It was annotated that the highest intensity (~16%) belonged to decanedioic acid, dibutyl ester (CAS# 109-43-3) followed by ~10% of Benzenesulfonyl isocyanate (CAS#2845-62-7). Among phthalate, diisobutyl phthalate (CAS# 84-69-5) just represented ~9% of relative intensity. In addition, phthalates compounds formed a distinctive cluster of 18 nodes in the molecular network. Other compounds like Decanedioic acid, dimethyl ester only accounted for ~3%, and the remained compounds (<3%) were classified as carboxylic acids.

Further analysis using GNPS public library in “Background” revealed the presence of some natural products from the forest like terpenoid compounds. Besides that, terpenoids were spread around different clusters in the molecular network, identifying monoterpenoids (CAS# 3796-70-1, 10373-81-6, 1195-92-2, 78-69-3, 106-26-3, 22422-34-0, 78-70-6), diterpenoids (CAS# 1740-19-8, 505-32-8, 110936-78-2) and sesquiterpenoids (CAS# 5172-21-4) (Annexes 2).

Table 2. Annotated compounds from aerosol particles in the background sample by GNPS public library with a normalized intensity >1%.

	Ret Time (min)	Compound Name	CAS #	Node	Functional Group before derivatization	Cluster	MQScore	Relative Intensity (%)
1	37.33	Decanedioic acid, dibutyl ester	109-43-3	743	Esters	11	0.78	16.3
2	27.58	Benzenesulfonyl isocyanate	2845-62-7	492	Organosulfur compounds	46	0.68	10.5
3	32.52	Diisobutyl phthalate	84-69-5	606	Phthalate	6	0.86	8.9
5	31.23	Decanedioic acid, dimethyl ester	106-79-6	571	Carboxylic acids	11	0.68	3.4
6	11.97	Methyl pelargonate	1731-84-6	116	Carboxylic acids	1	0.99	3.0
7	54.59	Triacontanoic acid, methyl ester	629-83-4	1180	Carboxylic acids	1	0.79	2.3
9	53.19	Octacosanoic acid, methyl ester	55682-92-3	1145	Carboxylic acids	1	*	1.7
10	52.99	Methyl triacontyl ether	629-83-4	1144	Alcohol	43	0.94	1.6

11	16.61	1-Heptyn-3-ol, 3-ethyl-5-methyl-	207974-16-1	227	Others	7	0.73	1.4
12	43.97	Methyl dodecanoate	111-82-0	922	Carboxylic acid	8	0.87	1.3

* Identified by NIST library without MQScore analysis

3.2. Molecular Networking Analysis of α -Pinene Oxidation by Ozone

Atmospheric aerosols undergo chemical and physical changes when exposed to ozone. To simulate the formation of oxygenated species from the pine forest, the oxidation of α -pinene (terpenoid) by ozone was selected as a model reaction. In this experiment, firstly, the α -pinene liquid was in contact with ozone (assigned as the liquid phase) and analyzed by GC-MS after a nonpolar extraction with CDCl_3 and without derivatization. Then, the oxidation products generated from the α -pinene oxidation were collected in a glass microfiber filter (assigned as the particle phase) and further analyzed as the liquid phase. As was expected, the liquid phase contained terpenoids (from α -pinene) as the main component, meanwhile, the particle phase showed a large quantity of carboxylic acid (~75.7%), followed by terpenoids (~10.9%), and ketones (~3.26%) as it is shown in Figure 6. Other chemical families like alkynes accounted for just ~4.5%, and unknown compounds accounted for ~1.7%.

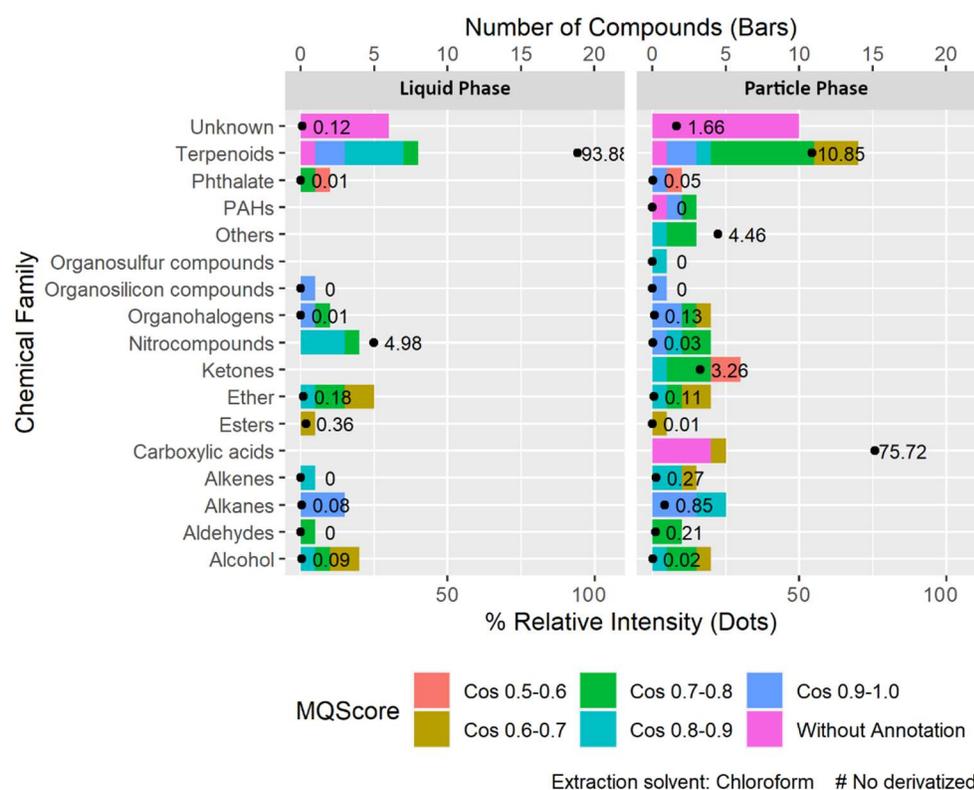
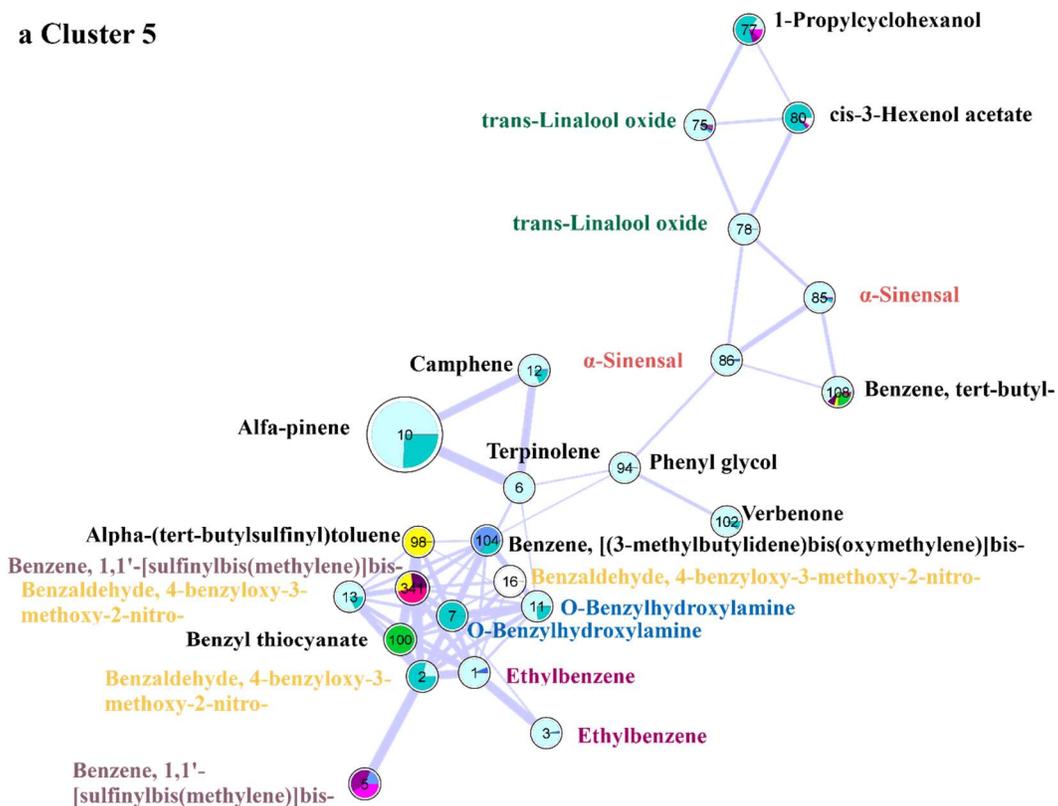


Figure 6. Chemical component distribution in the liquid and particle-phase from the α -pinene oxidation obtained from the molecular networking analysis. MQScore suggests the chemical similarity and the MQScore value ranges from 0.5 to 1.

The GC-MS analysis of the liquid phase of the α -pinene led to the identification of 38 nodes, 28 different constituents, and 6 unknown compounds (>99.5 % of the total). Additionally, 24 nodes from the liquid phase of the α -pinene experiment constituted a differentiated cluster in the molecular network (Figure 7a). Besides that, the identification of the oxidation products in the liquid phase showed the presence of camphene (~0.96%) (CAS# 79-92-5) as one oxidation product of the ozonolysis of α -pinene. It was observed a considerable fraction of Benzaldehyde, 4-benzyloxy-3-methoxy-2-nitro- (~4%) (CAS# 2450-27-3). Other compounds that were also presented in cluster 5 are O-benzylhydroxylamine (CAS# 2450-27-3), α -Sinensal (CAS# 17909-77-2), ethylbenzene (CAS# 100-41-4), and terpinolene (CAS# 586-62-9).

The GC-MS analysis of the particle phase aerosols led to the annotation of 72 nodes, 50 different constituents, and 10 unknown compounds (>97.6 % of the total). It was identified that pinonic acid (CAS# 473-72-3) is the most abundant oxidation product from the ozonolysis of α -pinene, together with (-)-2,3-pinandiol (CAS# 22422-34-0) making up ~81 % of the total. The α -pinene oxidation products also contain traces of L-menthyl 3-oxohexanoate (CAS# 92811-10-4), geranyl acetone (CAS# 3796-70-1), acetylcyclohexane (CAS# 823-76-7), 2,5,5-trimethyl-3-hexyn-2-ol (CAS#1522-16-3), as it is observed in cluster 7 of the molecular network (Figure 7b).

a Cluster 5



b Cluster 7

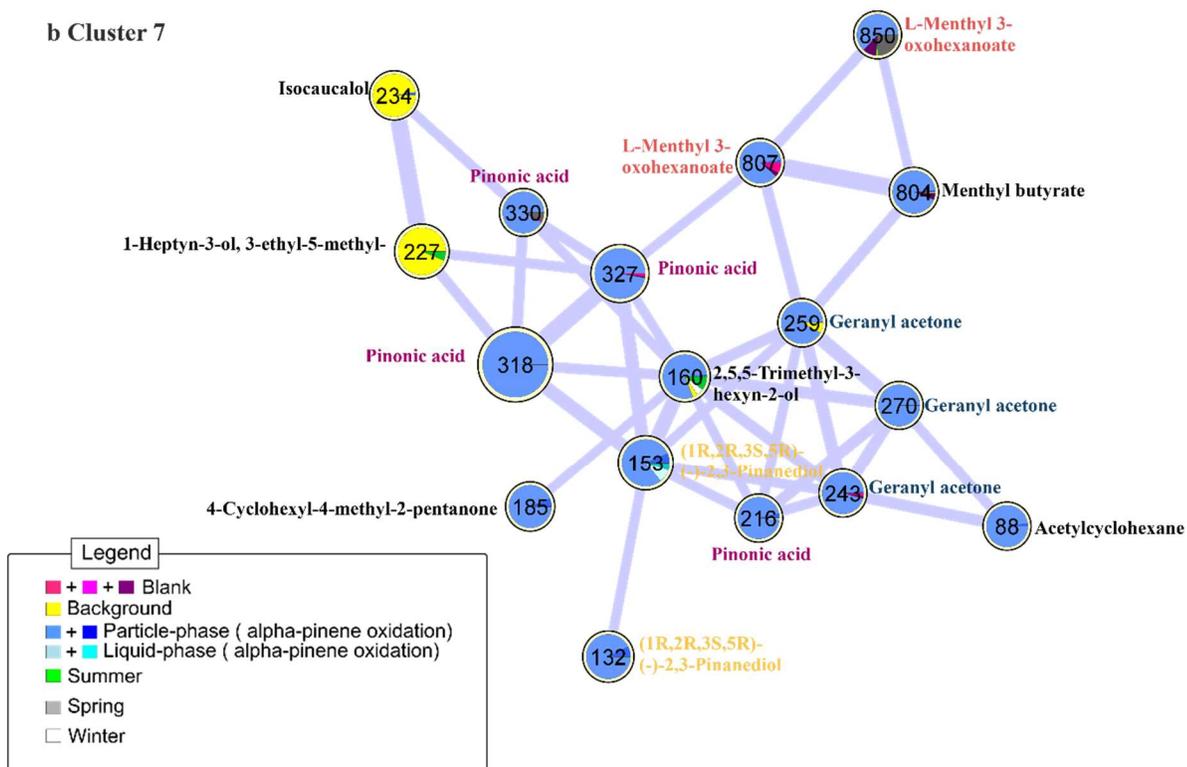


Figure 7. Molecular networking clusters from α -pinene oxidation. (a) Representative cluster #5 from the liquid phase of the α -pinene oxidation. (b) Representative cluster #7 from the particle phase of the α -pinene oxidation.

3.3. Molecular Networking of Atmospheric Aerosols in Summer, Winter, and Spring

The atmospheric aerosols collected during winter, spring, and summer were analyzed by GC/MS and molecular networking. As is shown in Figure 8, a set of diverse chemical families were identified and reported according to the MQScore values. It was found that the major chemical constituents in the winter were phthalates (~34.4%), followed by organosulfur compounds (~19.3%), organohalogen compounds (~17.9%), carboxylic acids (~9.3), ether (~4.5), and esters (~4.2%). Meanwhile, the spring season showed a high number of organophosphate esters (tris(2,4-di-tert-butylphenyl) phosphate), accounting for ~24.2 % of the normalized intensity. Furthermore, it was observed the presence of terpenoids (~2.7%), carboxylic acids (~1.7%), organosulfur compounds (~1.6%), alkanes (~1.4%), and alcohols (~0.5%). Over the summer, it was noticed that the main functional group was carboxylic acids (~30.9%), finding the presence of stearic acid as the main representative, and fatty acids and pinonic acid at low concentrations. Besides that, it was also found organophosphate esters (~8.4%), phthalates (~7.4%), organosulfur compounds (~7.0%), alkanes (~6.6%), alkenes (~3.8%), alcohols (~3.6%), esters (~2.4%), organohalogen (~2.4%), and ether (~1.49%).

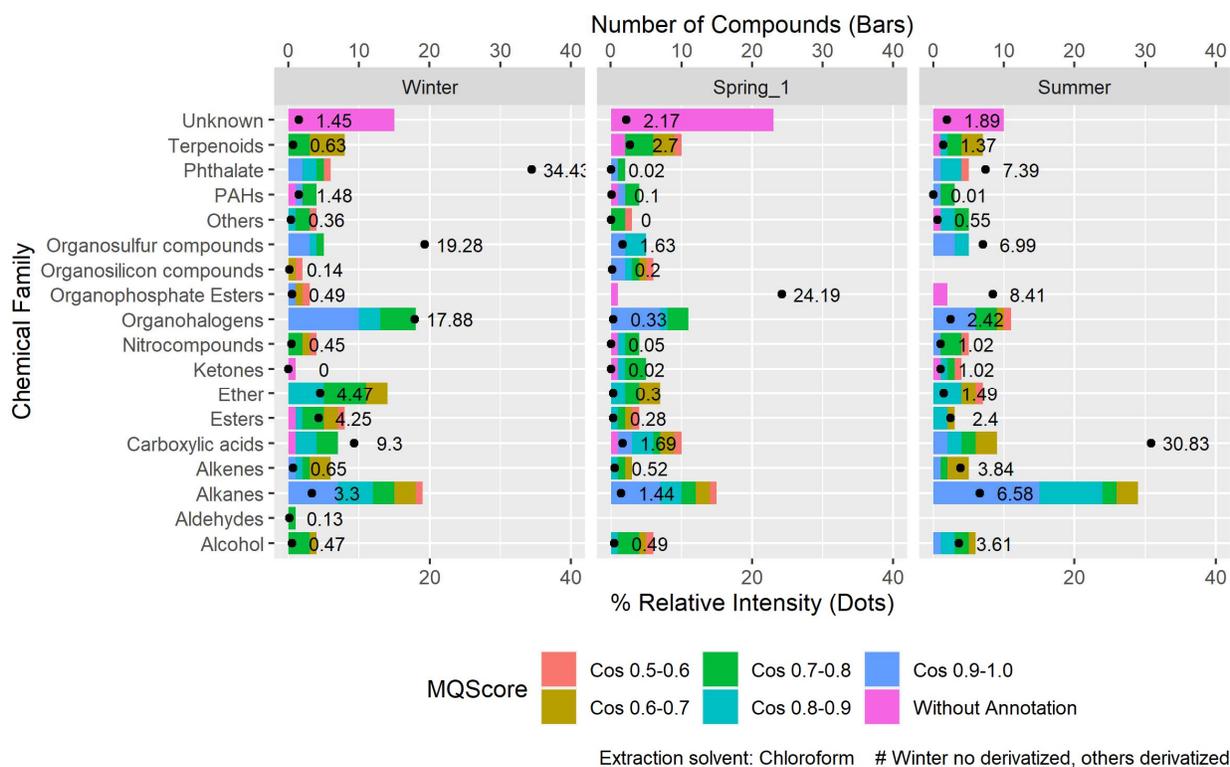


Figure 8. Comparison of the putative chemical families identified in the particle phase from winter, spring, and summer using the GNP public library. MQScore suggests the chemical similarity and the MQScore value ranges from 0 to 1.

Further analysis using the GNPS public library was performed for the studied seasons as is shown in Annexes 3. It was identified that a substantial number of phthalate compounds were present in the winter, accounting for a third part of the relative intensity (~34.4%), followed by organohalogen compounds like iodoalkane (~14.29%), hexadecenoic, and stearic acid (~9.3%), and alkanes (~2.0%). In the case of the summer season, it was found a certain concentration of alcohols like Methyl triacontyl ether (3.3%, CAS # 237742-64-2) Note: the compound is presented as a methyl ether due to the derivatization process. Furthermore, during the background analysis of the spring season, a significant concentration of diisobutyl phthalate (~45%) was identified. Other findings include the presence of Bis(2-ethylhexyl) phthalate during the winter (~10.9%) and summer (~3.87%) seasons, and also the presence of benzenesulfonamide, N-butyl (~18.35% in winter, ~6.47% in summer, and 1.44% in spring) in all the studied seasons.

Finally, it was observed that benzenesulfonamide, n-butyl, Tetracosane, cis-pinonic acid, pyrene (node 733), and 4-tert-Butyl-2,6-diisopropylphenyl acetate are present in all the forest samples and not in blank or background samples. Besides that, all the forest samples and background have the following common compounds: N-butyl, Tris(2,4-di-tert-butylphenyl) phosphate, octadecane, pentadecane, pentacosane, cyclopentadecane, (Z)-9-heptacosene, hexadecanoic acid, fluoranthene, 7H-Benz[de]anthracen7one (node 936) and octacosane,1-iodo.

To identify the presence of possible hydrocarbon molecules, the total *n*-alkane distribution is presented. As is shown in Figure 9, it was observed that the straight-chain *n*-alkanes ranged from C11 to C29 during the studied seasons. In addition, the molecular distribution of *n*-alkanes was characterized by an odd-carbon-number predominant, showing high abundance (~2.06%) for hexacosane (C26) during the summer season. On the other hand, tetracosane (C24) was dominant (~2.02 %) in winter, and octadecane (C18) was present (~1.02%) in spring with an even-carbon prevalence.

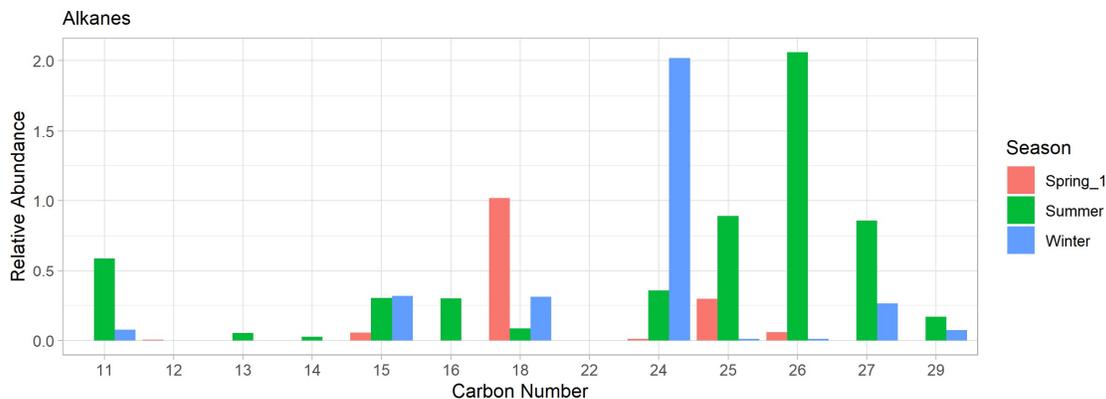


Figure 9. *n*-Alkane distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.

Further analysis of the aerosol composition was performed on the alkanol distribution. As is shown in Figure 10, it was identified that *n*-alkanols ranged from C8 to C30 with a high concentration (~3.32%) of C30 during the summer, ~ 0.01 % for C13 in the spring, and C13 during the winter (~0.37%). Besides that, a high number of alkanols with an odd number were found in winter (~0.37%) and spring (~0.01%). Meanwhile, summer presents less concentration of even alkanol (~3.32%).

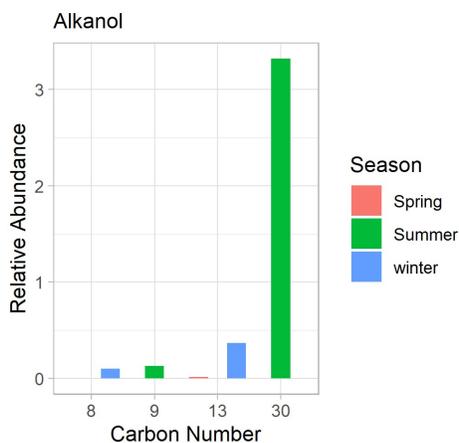


Figure 10. *n*-Alkanol distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.

Another important molecular marker for organic matter in the atmosphere is carboxylic acids. As shown in Figure 11, carboxylic acids ranged from C9 to C23, finding a dense concentration of C16 during the winter (~6.60%) and spring season (~0.50%), and high levels of C18 during summer. In particular, a significant concentration of stearic acid (C18) was found in summer (~29.35%), and lower concentrations (~2.65%) in winter.

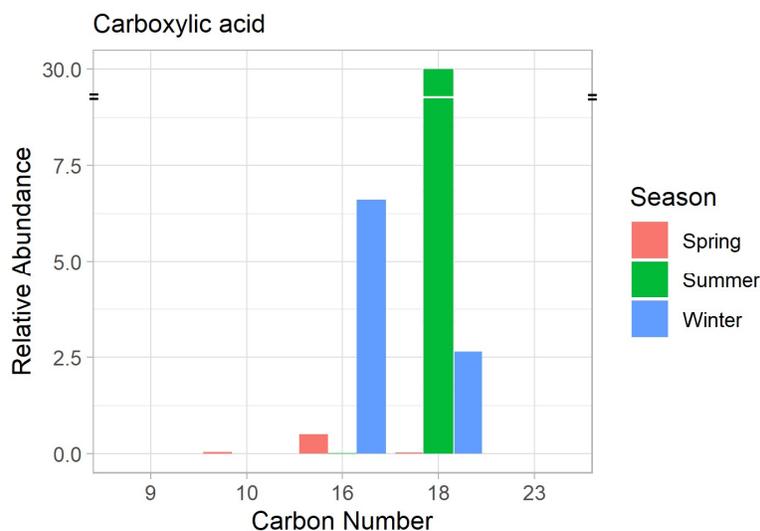


Figure 11. *n*-Carboxylic acid distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.

To identify the presence of possible *n*-alkyl iodide compounds, the total *n*-alkyl iodide distribution is presented. As is shown in Figure 12, it was observed that the straight-chain *n*-alkyl iodide ranged from C10 to C36 during the studied seasons. In addition, the molecular distribution of *n*-iodoalkanes was characterized by an even-carbon-number predominant, showing high abundance (~5.27%) for octacosane, 1-iodo- (C28) during the winter season. On the other hand, octacosane, 1-iodo- (C28) was dominant (~0.17 %) in spring, and dotriacontane, 1-iodo- (C32) was present (~0.038%) in summer with minor prevalence.

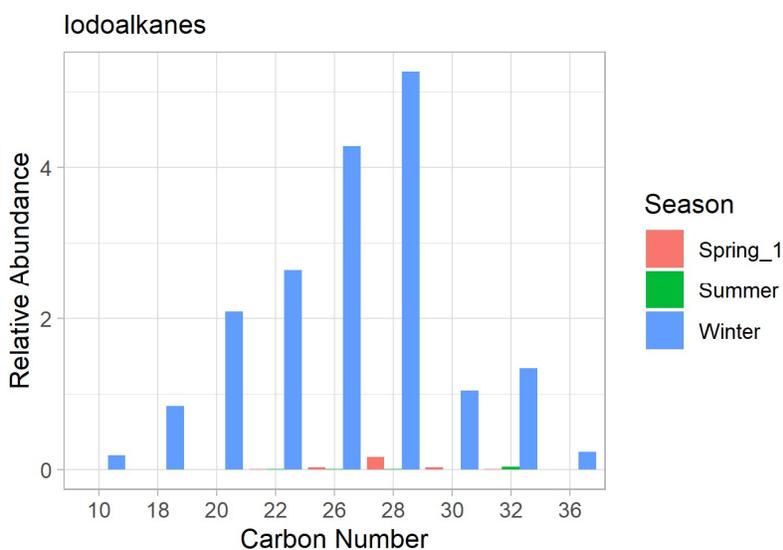


Figure 12. *n*-Iodoalkanes distribution patterns in total suspended particulate aerosols collected during winter, summer, and spring in SMEAR Estonia.

3.4. NMR Analysis

In a similar way that was described previously, the background (sample collection with the HEPA filter before filter holder for 1 week with a pumping system) and blank (sample collection with the filter holder for 5 min without any pumping system) samples were firstly analyzed by NMR (Appendix 5). In addition, it was identified a variety of functional groups in the studied seasons including aromatic compounds, carbohydrates, olefins, alcohols, ether, esters, amines, and hydrocarbon chains (Appendix 6). A detailed analysis is described in the discussion section.

To investigate deeper the functional group characterization, 2D NMR analyses were performed in selected samples including ^1H - ^1H TOCSY (Appendix 7), ^1H - ^{13}C HSQC (Appendix 8), and ^1H - ^{13}C HMBC (Appendix 9). Then, specific regions which belong to functional groups were identified similarly that it was early defined by Chalbot in 2021⁵⁷ (Table 3).

Table 3. Chemical shift ranges of the regions defined from the 2D-NMR spectra according to Chalbot⁵⁷.

Regions ^1H - ^1H NMR	F2/F1 Ranges (ppm)	Regions ^1H - ^{13}C -HSQC	F2/F1 Ranges (ppm)	Regions ^1H - ^{13}C -HMBC	F2/F1 Ranges (ppm)	Compounds
A	0.5-2.5 0.5-2.5	I	0.5-1.6 10-40	I _a I _b I _c	0.7-1.8 10-40 0.7-1.8 40-60 0.7-1.8 60-100	-Intra-aliphatic chain couplings in aliphatic compounds, monocarboxylic acids, long-chain carboxylic acids, amino acids -Methylene adjacent to amines (R-CH ₂ -N) or chlorine (R-CH ₂ -Cl) -Methylene adjacent to hydroxyl (R-CH ₂ -O)
B	2.4-3.2 0.8-1.8	II	1.6-3.2 16-56	II _d	2.4-3.2 30-56	-Amine
C	1.8-3.2 1.8-3.2	II	1.6-3.2 16-56	II _a II _b II _c	1.8-2.5 20-50 1.8-2.5 160-190 1.8-2.5 200-230	-Oxo-acids -hydroxyacids -compounds with carboxylic and ester -compounds with ketones
D	3.0-4.6 0.8-2.5	III, IV	3.0-4.4 44-60	III, IV	3.0-4.8 50-115	-Methyl-polyols, Secondary organic Aerosols, Hydroxy acids, amino acids
E	3.2-4.6 3.2-4.6	III, IV	3.1-4.4 44-60	III, IV	3.0-4.8 50-115	-Glucose, fructose Sucrose, Levoglucosan Ethanolamine Choline HMSA
F	4.8-5.25 3.4-4.4	V	4.4-5.6 84-115	V	4.9-5.6 60-110	-Anomeric carbons of carbohydrate
G	5.25-5.8 3.4-4.6	V	4.4-5.6 84-115	V	4.9-5.6 60-110	-Anomeric carbons of anhydrohexose
H	6.6-8.2 5.5-6.8	VI	6.6-9 115-140	n/a		-Alkene
I	6.4-8.8 6.4-8.8	VI	6.6-9 115-140	n/a		-Aromatic
J	4.6-6.6 0.8-2.6	V	4.4-5.6 84-115	V	4.9-5.6 60-110	-Olefinic compounds

3.5. Compound Identification on Two-dimensional (2D) NMR with NMRfilter

To add a layer of confidence for the identity of the annotated compounds, the 2D NMR data were imported to the NMRfilter software⁷¹ for evaluation of the coupling constant networks and matching profiles. Peak lists from HSQC (heteronuclear single quantum correlation), and heteronuclear multiple-bond correlation (HMBC) were used as input, and compounds suggested by the GC–MS identification routine were listed as SMILES strings and used as candidates for NMRfilter⁷⁰. The shift prediction is done using data from nmrshiftdb2 and an extended HOSE code algorithm⁷². The output is shown in Table 4, and Appendix 10.

‘Distance’ and ‘standard deviation’ in this scheme are specific parameters from NMRfilter that represent how close the experimental data are to the simulations and how sparse the peaks are in correlated networks, respectively. (Note: ‘Distance’ values close to 0 indicate spectral similarities and ‘standard deviation’ close to 1 indicate that peaks are clustered together)

Table 4. NMRfilter analysis with HSQC and HMBC providing a matching rate and quality parameters (‘distance’ and ‘standard deviation’).

	Compound name	Group	Formula	Node	Distance	Std deviation	Match HMB C	Matching rate	Match HSQ C	Matching rate
Spring	Isopropanol	Alcohol	C ₃ H ₈ O	-	0	0.01	1/1	1.00	2/2	1.0
	tert-Butanol	Alcohol	C ₄ H ₁₀ O	-	0	0.13	1/1	1.00	1/1	1.0
	Cyclopentadecane	Alkanes	C ₁₅ H ₃₀	678,737,974,996	0	0	1/1	1.00	1/1	1.0
	Dodecane	Alkanes	C ₁₂ H ₂₆	65,170,1348,1354	0	0.48	10/10	1.00	6/6	1.0
	Dotriacontane	Alkanes	C ₃₂ H ₆₆	1079	0	0.69	14/14	1.00	8/8	1.0
	n-Octadecane	Alkanes	C ₁₈ H ₃₈	33,35,484,601,851,997,1002	0	0.69	14/14	1.00	8/8	1.0
	Octacosane	Alkanes	C ₂₈ H ₅₈	1077	0	0.69	14/14	1.00	8/8	1.0
	Octadecane	Alkanes	C ₁₈ H ₃₈	33,35,484,601,851,997,1002	0	0.69	14/14	1.00	8/8	1.0
	Pentacosane	Alkanes	C ₂₅ H ₅₂	788,920,931,977,1015	0	0.69	14/14	1.00	8/8	1.0
	Pentadecane	Alkanes	C ₁₅ H ₃₂	186,454,604,620	0	0.68	13/13	1.00	8/8	1.0
	Pentatriacontane	Alkanes	C ₃₅ H ₇₂	1118	0	0.69	14/14	1.00	8/8	1.0
	Tetracosane	Alkanes	C ₂₄ H ₅₀	474, 883	0	0.69	14/14	1.00	8/8	1.0
	Tetradecane	Alkanes	C ₁₄ H ₃₀	950,1308,1335	0	0.48	12/12	1.00	7/7	1.0
	Acetoacetate	Carboxylic acids	C ₄ H ₅ O ₃	-	0	0.01	2/2	1.00	2/2	1.0
	Adipic acid	Carboxylic acids	C ₆ H ₁₀ O ₄	349	0.01	0.12	2/3	0.67	2/2	1.0
	Decanoic acid	Carboxylic acids	C ₁₀ H ₂₀ O ₂	455	0	0.52	14/15	0.93	9/9	1.0
	Heptadecanoic acid	Carboxylic acids	C ₁₇ H ₃₄ O ₂	627	0	0.64	25/26	0.96	14/14	1.0
	Hexadecanoic acid	Carboxylic acids	C ₁₆ H ₃₂ O ₂	613, 644, 794	0.01	0.87	24/25	0.96	14/14	1.0
	n-Tricosanoic acid	Carboxylic acids	C ₂₃ H ₄₆ O ₂	926	0	0.64	25/26	0.96	14/14	1.0
	Stearic acid	Carboxylic acids	C ₁₈ H ₃₆ O ₂	709,760	0.01	0.87	25/26	0.96	14/14	1.0
	Tricarballic acid	Carboxylic acids	C ₆ H ₈ O ₆	-	0	0.34	3/3	1.00	2/2	1.0
	Decanamide-	Nitrocompounds	C ₁₀ H ₂₁ NO	749	0.01	0.69	14/15	0.93	9/9	1.0
	Docosane, 1-iodo-	Organohalogens	C ₂₂ H ₄₅ I	770	0.01	0.75	24/25	0.96	13/14	0.9
	Octacosane, 1-iodo-	Organohalogens	C ₂₈ H ₅₇ I	978, 1139	0.01	0.75	24/25	0.96	13/14	0.9
Tetracosane, 1-iodo-	Organohalogens	C ₂₄ H ₄₉ I	825	0.01	0.75	24/25	0.96	13/14	0.9	

	Cyclopentasiloxane, decamethyl-	Organosilicon compounds	C ₁₀ H ₃₀ O ₅ Si ₅	1378,1405	0	0.16	1/1	1.00	1/1	1.0
	Cyclotetrasiloxane, octamethyl-	Organosilicon compounds	C ₈ H ₂₄ O ₄ Si ₄	17,20	0	0.04	1/1	1.00	1/1	1.0
	Tetrasiloxane, decamethyl-	Organosilicon compounds	C ₁₀ H ₃₀ O ₃ Si ₄	1364	0	0.17	2/2	1.00	2/2	1.0
	Disulfide, di-tert-dodecyl	Organosulfur compounds	C ₂₄ H ₅₀ S ₂	1169,1260	0	0.69	18/18	1.00	10/10	1.0
	tert-Hexadecanethiol	Organosulfur compounds	C ₁₆ H ₃₄ S	1320	0	0.87	25/24	1.04	14/14	1.0
α-pinene experiment	Cyclopentadecane	Alkanes	C ₁₅ H ₃₀	678,737,974,996	0	0	1/1	1.00	1/1	1.0
	Adipic acid	Carboxylic acids	C ₆ H ₁₀ O ₄	-	0	0.12	3/3	1.00	2/2	1.0
	Acetoacetate	Carboxylic acids	C ₄ H ₅ O ₃	-	0	0	2/2	1.00	2/2	1.0
	3-Formyl-2,2-dimethylcyclobutane-carboxylic acid	Carboxylic acids	C ₈ H ₁₂ O ₃	-	0.37	0.56	10/12	0.83	3/5	0.6
	Norpinic acid	Carboxylic acids	C ₈ H ₁₂ O ₄	-	0.01	0.01	5/6	0.83	3/3	1.0
	Pinic acid	Carboxylic acids	C ₉ H ₁₄ O ₄	-	0.01	0.7	13/13	1.00	3/5	0.6
	2-Propanone	Ketones	C ₃ H ₆ O	-	0	0	1/1	1.00	1/1	1.0
	2,5,5-Trimethyl-3-hexyn-2-ol	Others	C ₉ H ₁₆ O	160	0	0	3/4	0.75	2/2	1.0
P-Menthan-4-ol	Terpenoids	C ₁₀ H ₂₀ O	260	0	0.01	11/11	1.00	4/6	0.7	

Note: 'Distance' values close to 0 indicate spectral similarities and 'standard deviation' close to 1 indicates that peaks are clustered together.

4. Discussion

4.1. Molecular Networking Analysis of Interfering Compounds from the Background

As it was presented earlier, the identification of compounds in the background sample is crucial to get a clear understanding of the real composition of the organic aerosols in the studied seasons. During the MN analysis, all the putative annotations are considered as level 3 which means that the annotated compounds are characterized by spectral similarity in a chemical family of compounds according to the Metabolomic Standard Initiative⁷³. Based on this, it was found that the background compositions were mainly influenced by carboxylic acids (~20.38%, Figure 5). In addition, malic acid (~0.053%), a photo-oxidation product of unsaturated fatty acids⁷⁴ was also identified. Even when the presence of carboxylic acids like fatty acids is known compounds to be typically found in the forest⁷⁵, the presence of carboxylic acids in the background sample could indicate a possible accumulation of atmospheric compounds in the aerosol collection system. Nonanedioic acid is formed from the ozonolysis of oleic and linoleic acid⁷⁶ or the oxidation of particulate unsaturated fatty acids with the double bond at the C-9 position⁷⁶. This dicarboxylic acid was found in the background sample (~0.03%) with Decanedioic acid (3.4%). In another example, we found a significant concentration (~45.7%) of DiBP (CAS # 84-69-5) during the spring season. This compound is classified as a plasticizer and can be used for coatings, paints, pipes, and tanks. In addition, this compound may also be coming from plastic burning materials⁷⁷,

as it was also reported to be widespread worldwide ⁷⁸. Besides that, it was also noted the presence of other plasticizers like DEHP (CAS # 117-81-7), DEP (CAS # 84-66-2) as is shown in Annexes 2.

Additionally, two resin acids were identified in the Background sample. The first one was Dehydroabietic acid (~0.0023%) which is known to be a biomarker from the burning of softwood or conifer trees ⁷⁹. The second one was the 7-Oxodehydroabietic acid which is a derived compound from the dehydroabietic acid ⁸⁰. These findings could make us hypothesize that the filter holder setup may be retaining those compounds. Furthermore, dehydroabietic acid was also detected to be present in winter (~0.27%), and its influence is discussed in the next sections.

Furthermore, (Z)-9-Heptacosene is a characteristic compound of epicuticular waxes of male houseflies ⁸¹, and female long-horned beetle ⁸². This compound was found in all the forest and background samples. It had a relative intensity of 3.4% during the summer season (Annexes 3).

4.2. Molecular Networking Analysis of α -Pinene oxidation by Ozone

α -Pinene is the second foremost released non-methane VOC from biological resources and a major contributor to aerosol formation. In this experiment, the α -pinene oxidation was chosen as a model reaction to generate aerosols (a size of 100 μ m, Table1) that can mimic the natural oxidation of VOCs in the forest. It was found that (-)-2,3-Pinenediol (node 153, cluster 7, CAS# 22422-34-0) was the main oxidation product (~1.9%) in the liquid phase and the particle phase (~5.5%) after the oxidation process. This compound is reported to be an early-stage oxidation product formed *via* the photooxidation or ozonolysis of α -pinene ⁸³. Similarly, (-)-2,3-Pinenediol is also reported to be an oxidation product of β -pinene *via* the metabolism of a fungus species like *Botrytis cinerea*⁸⁴ in a culture flask. Consequently, (-)-2,3-Pinenediol was also found in spring (~0.006%), summer (~0.345%), and in the Background (~0.0007%) at very low concentrations. This fact may be possible due to the continuous oxidation process in the forest which leads to an attenuation of the precursor aerosols. Furthermore, other related molecules to α -pinene were also identified as camphene and terpinolene from the liquid phase (cluster 5) with an MQScore greater than 0.88.

In the particle-phase analysis of α -pinene oxidation, it was also found that (-)-2,3-Pinenediol was present, and this observation was in agreement with other reports in the particulate phase ⁸⁵. Moreover, the main oxidation product detected in this phase was pinonic acid (cluster 7, nodes 216, 318, 327, 330), representing ~75% of the relative intensity (Figure 6). This compound was reported to be a secondary organic aerosol tracer worldwide ⁸⁶ due to the lower volatility

properties, and their release from the particle phase into the gas phase may depend on the acidity of the environment ⁸⁷.

4.3. Molecular Networking of Atmospheric Aerosols in Summer, Winter, and Spring

The application of a non-targeted analysis based on GC-MS to study aerosols let the identification of different groups of compounds: organophosphate esters, organosulfur compounds, phthalates, polycyclic aromatic hydrocarbons, aliphatic compounds (*n*-alkanes, *n*-alkanols, *n*-alkanoic acids), organohalogen, and terpenoids as it is shown in Figure 8. In this section, the annotated compounds are being discussed based on their environmental relevance and their possible origin.

4.3.1. Organophosphates

In terms of risk potential, organophosphate flame retardants and plasticizers represent chemical emerging concerns because their transformation products can be more toxic and persistent than the parent chemicals in the environment ⁸⁸. In this work, organophosphate compounds were annotated in all the seasons (Annexes 3). The tris(2-ethylhexyl) phosphate (TEHP), a non-chlorinated organophosphate, was found in winter (~0.34%). This compound is reported to be also present in PM_{2.5} ⁸⁹, TSP ⁹⁰⁻⁹², and remote areas such as the arctic ^{93,94}. In addition, this compound can be bound to glass materials ⁹⁵ such as the GM filter used in this study. Another organophosphate triester is the Tris(2,4-di-tert-butylphenyl) phosphate (T24DtBPP) which is commonly reported to be an indicator of the open-burning of plastic⁷⁷. This compound represented ~24% of the total components in spring (node 1243), ~3.27 % in summer (node 1274), and ~0.01% in the background sample (node 1342). Besides, it is worth mentioning that high concentrations of T24DtBPP in the aerosol samples may also indirectly come from the oxidation of tris(2,4-di-tertbutylphenyl) phosphite (CAS# 31570-04-4), which is a common compound used as an organophosphite antioxidant for plastic production worldwide. Also, the latter compound is found to be a toxic pollutant in urban PM_{2.5} ⁹⁶. In this study, it was identified that tris(2,4-di-tertbutylphenyl) phosphite (CAS# 31570-04-4) represented ~5.13% during summer (node 1207), and the intensity ratio between tris(2,4-di-tert-butylphenyl) phosphate (T24DtBPP) and tris(2,4-di-tertbutylphenyl) phosphite was accounted for 0.64, evidencing a high proportion of this precursor compound.

4.3.2. Organosulfur compounds

Organosulfur compounds like N-butyl-benzenesulfonamide are used commercially as a plasticizer in the polymerization of polyamide compounds. In this study, N-butyl-benzenesulfonamide (CAS#

3622-84-2) was found in all the seasons and was absent in blank and background samples. This finding is in agreement with other reports where this compound was found with high intensity in urban aerosol samples TSP ⁹⁷, PM_{2.5} ⁹⁸, and in the boreal Forest Hyytiälä ⁹⁹. In our case, this compound was distributed in winter (~18.61%), spring (~1.44%), and summer (~6.66%).

4.3.3. Phthalates

Phthalates are ubiquitous in the atmosphere, and they are also reported in air indoor environments ¹⁰⁰. Among their distribution, lower molecular weight phthalates are used as solvents in personal care products, while higher molecular weight phthalates are used as plasticizers in various polymers, especially in the manufacture of polyvinylchloride (PVC) ^{101,102}. As an example, Bis(2-ethylhexyl) phthalate (DEHP) is typically added to rigid PVC to impart flexibility ¹⁰³. While diethyl-phthalate (DEP), and Dibutyl phthalate (DnBP) are widely used as solvents to hold color and scent in various consumer and personal care products ¹⁰¹.

It is known that Glass Microfiber filters can retain some Phthalates like DEHP and DBP ¹⁰⁴, suggesting the possibility to be attached to our filters during the aerosol collection. For instance, it is reported that DEHP, a phthalate ester, was prevalent in the particles phase ¹⁰⁵. As a result, the partitioning of high molecular phthalates to particles may reduce their photolysis rates, and thus increase their persistence in the atmosphere, as atmospheric reactions often deplete compounds in the gas phase. In our samples, DEHP presents ~10.9% in winter, 3.9% in summer, and it also was found in spring with ~0.02%. Like DEHP, DEP was also found in summer and winter but with lower concentrations (<~0.11%). In addition to these phthalates, DnBP and Diisobutyl phthalate were identified in winter with a significant relative intensity (~20.6%) for the former, and just ~1.9% for the latter. Another phthalate is di(2-ethyl hexyl)isophthalate (DEHIP) which was found in summer (~3.37%). Interestingly, this compound has been found in human milk samples in China, indicating its possible presence in the environment ¹⁰⁶. The European Chemicals Agency reported DEHIP as very toxic to aquatic life and may damage the fertility of an unborn child ¹⁰⁷.

4.3.4. Polycyclic Aromatic Hydrocarbons (PAHs)

The MN analysis showed the presence of two PAHs nodes during the winter season. Those nodes were identified as fluoranthene and pyrene. Both compounds were found in all the studied seasons. In addition, Fluoranthene (CAS# 206-44-0, node 696) presented ~0.63% in winter, ~0.087% in spring, and ~0.008 % in summer. Meanwhile, Pyrene (CAS# 129-00-0, node 733) was distributed ~0.64% in winter, ~0.002% in spring, and ~0.0008% in summer. Altogether those PAHs are

typically originating from incomplete combustion of organic material (both fossil and modern carbon)¹⁰⁸ also in freshly emitted smoke from green vegetation¹⁰⁹. An oxy-PAH that was found in the winter season was 7H-Benz[de]anthracen7one (CAS# 82-05-3, node 866 and node 936, with ~0.21% of relative intensity in winter). This compound was also reported in particulate PM₁₀ in ambient air of Grenoble (France)¹¹⁰.

4.3.5. Organosilicon Compounds

Volatile methylsiloxanes can be classified as volatile linear (IVMS) or cyclic methylsiloxanes (cVMS). IVMS and cVMS are widely used in cleaning agents, lubricants, and personal care products, such as cosmetics, antiperspirants, and skin and hair care products. Some representatives of IVMS are Octamethyltrisiloxane (L3); Decamethyltetrasiloxane (L4). Several cVMS including octamethylcyclotetrasiloxane (D4), decamethylcyclopentasiloxane (D5), and dodecamethylcyclohexasiloxane (D6) have been prioritized in several regulatory jurisdictions due to their persistence and bioaccumulation potential and environmental toxicity^{111,112}.

Global average half-lives for this reaction have been determined to be approximately 30 days for D3, 15 days for D4, and 10 days for D5, while IVMS are estimated to have half-lives around 9 days¹¹³. Modeling studies have predicted that these half-lives are sufficient for cVMS to undergo long-range transport to remote regions, including the Arctic¹¹⁴. In the forest sample from SMEAR Estonia, the IVMS have L4 in spring (~0.0007 %). On the other hand, spring has a cVMS-predominance, with D4 (~0.18%) and following D5 (~0.021%), meanwhile in winter D5 (~0.00013 %).

4.3.6. Aliphatic compounds

Series of lipid class compounds, including *n*-alkanes (C11 to C35), *n*-alkanols (C8 to C30), and *n*-alkanoic acids (C5 to C23), were detected in the forest. *n*-Alkanoic acids are the major lipid class compounds (31.0%) in Järvelja during summer, but alkanes are the dominant lipid class compound in winter (11.7%) and spring (2.6%) (Figure 9).

4.3.6.1. Alkanols

Fatty acids (*n*-alkanoic acid) from the flora and fauna usually range from C12 to C32 with even-numbered carbon chain length homologs. Because fatty alcohols (*n*-alkanols) in plants are biosynthesized from fatty acids by enzymatic reductions, they have similar carbon chain lengths with predominantly even number chains¹¹⁵. As such, fatty alcohols homologs with less than C20 are not a major constituent of plant waxes but are derived from microbiological sources⁸⁵.

Meanwhile, the homologs higher than C20 may derive from epicuticular vegetation ¹¹⁶. Also, a high concentration of alcohols containing C13 (~0.36%) and C30 (~3.32%) was identified.

The natural vegetation waxes consist of the longer chain of *n*-alkanes (C23-C35), *n*-alkanoic acids (C22-C32), and *n*-alkanols (C22-C32). Typically, a layer of wax exists on the surface of plant leaves that acts as a protective barrier between the cuticula and the environment ¹¹⁷. Moreover, the wax can become airborne due to wind-induced mechanical shear and the rubbing of leaves against each other ¹¹⁸. Plant waxes are mostly non-volatile and are expected to maintain their original (as emitted) distribution during transport ¹¹⁹. In our work, *n*-alkanols ranged from C8 to C30 with a high concentration of C30 during the summer, C13 during the spring, and the winter. This finding could indicate the microbiological sources of alkanols in spring and winter, the possible origin from plant waxes in summer. Other alkanols with even numbers are also found in winter, spring, and summer.

4.3.6.2. Carboxylic acids

Alkenoic acids (carboxylic acids) located in the atmosphere are unstable and are often used as an indication of the age of aerosols ¹²⁰. Carboxylic acid ranged from C9 to C23, with a high distribution of C16 during the winter and spring season, and C18 during the summer. In particular, ~29 % of stearic acid (C18) before derivatization was found in the summer, meanwhile, its concentration was reduced (~2.65%.) during the winter. In addition, it was found the presence of other unsaturated acids (alkenoic acid) like 11-*trans*-Eicosenoic acid in winter (0.0003%), spring (0.001%), and summer (0.29%) season.

n-Alkanoic and *n*-alkenoic acids are known to be synthesized by biological systems and are frequently constituted for even carbon numbered homologs between C16 and C18 acids as major compounds. Consistent with the expectation, both the natural Oak and pine smoke particulate matter exhibits the typical even to odd carbon numbered *n*-alkanoic acid distribution that is characteristic for biosynthetic organic matter ¹²¹. In our study, even carbon numbered *n*-alkanoic acid was predominant in summer and winter but in spring was not the case. The *n*-alkanoic acids C23 were interpreted to derive from higher plant waxes, meanwhile, the *n*-alkanoic acid distributions (<C20) are probably derived from microorganisms hosted by plant foliage, which are ubiquitous in all biota.

Another type of carboxylic acid is hexadecenoic acid, which was observed to be ~6.6% during the winter, and just ~0.5% during spring. In contrast, the summer season just showed a marginal

concentration (~0.01%), in which the occurrence of metabolites was less visible in the molecular network of the annotated molecules.

Regarding dicarboxylic acids in the forest, they ranged from C5 to C8 in summer. The following dicarboxylic acids were identified: pentanedioic acid and hexanedioic acid. In general, dicarboxylic acids are detectable in wood smoke but not in synthetic log smoke. In wood smoke, only short-chain homologs of dicarboxylic acid from C3 to C6 are observed¹²¹. A series of dicarboxylic acids, ranging from C3 to C16, were also reported in the smoke samples from a wildfire in Portugal. It seems that higher emissions of dicarboxylic acids take place during the flaming phase¹²². Pentanedioic acid and Hexanedioic acid are molecular tracers of ozonolysis of cyclic olefins (cyclohexene)^{123,124}.

Among branched dicarboxylic acids, 3-Methylheptanedioic acid was found in summer (0.38 %). This compound also was reported in PM1 summer samples collected in a forest clearing in the middle area of a confined mixed forested area in Hungary¹²⁵. Other types of carboxylic acids like pinonic acid were found in spring (~0.02%), summer (~0.00019%), and winter (~0.00011%). Typically, pinonic, norpinic, and pinic acids are produced by the photooxidation of α/β -pinene *via* reactions with O₃ and OH· radicals¹²⁶.

4.3.6.3. Alkanes

Alkanes are the primary fuel type in combustion engines, and they are considered an important urban trace gas¹²⁷. The saturated hydrocarbons are mainly emitted by incomplete fuel combustion, the alkanes in the dust could be attributed to the primary emissions of the traffic wastes and some other combustion processes¹²⁸. Most petroleum contains *n*-alkanes ranging from C10 to C35 with no carbon number predominance. In the forest sample from SMEAR Estonia, the straight-chain *n*-alkanes range from C11 to C29, and their molecular distribution is characterized by an odd-carbon-number predominance with maxima at hexacosane (C26) in summer (~2.06 %). On the other hand, spring and winter have an even-carbon-predominance, with octadecane (C18) dominating in spring (~1.02 %) and tetracosane (C24) in winter (~2.02%).

4.3.7. Organohalogen

The presence of organohalogen compounds was identified with a higher concentration during the winter (~17.7%) in comparison with just traces in summer (<0.1%) and spring (~0.24%). Some examples include decane, 1-iodo-(C₁₀H₂₁I); eicosane, 1-iodo (C₂₀H₄₁I); docosane, 1-iodo (C₂₂H₄₅I); and octacosane, 1-iodo(C₂₈H₅₇I) (Figure 12). In general, *n*-alkyl iodide compounds are

photolabile entities, and they are typically used in photolysis reactions. For instance, the gas phase of alkyl iodides like 1-octyl iodides (C8), 1-decal iodides (C10), and 1 dodecyl iodides (C12) can act as precursors to produce organic peroxy radicals that act as SOA ¹²⁹. The high abundance of these organohalogen compounds may be explained by the reduced light intensity during the winter ¹²⁹ which contributes to a reduced oxidation rate of these compounds. Another possible explanation may be attributed to certain fungi species which can proliferate in the trees with more incidence during the winter. For instance, Octacosane, 1-iodo was reported in the headspace of rice plants infected with *Scirpophaga* ¹³⁰. In addition, decane, 1-iodo was present in a dichloromethane extraction of *Sordariomycetes* sp. endophytic fungi isolated from *Strobilanthes crispus* ¹³¹.

4.3.8. Terpenoids

According to the molecular networking analysis of the studied seasons, the forest samples displayed the presence of terpenoid nodes which were also included in the two clusters of the previous α -pinene laboratory experiment. For example, verbenone, and benzenemethanol, 4-ethyl were identified during the summer and spring season respectively, and these compounds were also present in cluster 5 (Figure 7a) from the α -pinene oxidation experiment. Similarly, other compounds from cluster 7 (Figure 7b) were also observed in traces during the studied seasons. For instance, 2,3-Pinenediol was annotated in summer (~0.35%) and at very low quantities in spring (<0.01%). In addition, dehydroabietic acid was found in winter (<~0.27%), and this diterpenoid is considered a major marker compound emitted from conifer burning ¹³². Also, 7-Oxodehydroabietic acid was present in the spring sample (~1.18%), it is considered a thermal degradation product of resin acids ¹³³ it is characteristic of ponderosa pine duff, needles, and sticks burning ¹³⁴. Overall, these findings suggest that those compounds are structurally relevant to the oxidation of terpenoids such as α -pinene, and the presence in trace levels may correspond to the continuous oxidation process of these compounds.

4.4. Analysis of ¹H-NMR Spectra

As was mentioned previously, the analysis of the control samples from the background and blank were necessary to determine the real composition of the studied seasons. As is shown in Appendix 5a, the ¹H NMR spectrum from the background showed a higher signal than our blank, and also displayed other relevant signals. This observation can be explained due to the long collection time (~69 h) of the background sample in comparison to the blank sample. The resulting additional signals are accumulated on the filter from the gas phase, during the sample collection

all particles were filtered out. As it was observed during the GC analysis, a set of compounds was identified in the background sample like Phthalates. In line with this observation, additional peaks were identified with a chemical shift of δ H7.71, δ H7.53, δ H4.3 in the NMR spectrum of the background, suggesting the possible presence of aromatic species and esters (Appendix 5b,c). In addition, the NMR spectrum of the background and blank sample allowed the identification of a hydrocarbon signal (Appendix 5d) which was used later for the study of the seasons.

Next, the analysis of the functional groups in the winter, spring, and summer seasons was performed by NMR using a non-polar solvent (CDCl_3). During the analysis, it was annotated the presence of aromatic species during the summer daytime (09-16/06/21) and summer nighttime (05-12/08/20) as is shown in Appendix 6 (red labels). Also, it was identified amides with a chemical shift of δ H8.75 in summer nighttime (05-12/08/20) and aldehydes with a chemical shift of δ H9.77 in background daytime (09-16/06/21). In addition, the presence of carbohydrates (pink labels) was noted in summer daytime (09-16/06/21) and winter daytime (13-21/01/21). Interestingly, species containing double bonds and amines (blue labels) were identified during winter daytime (13-21/01/21). Finally, summer seasons including daytime and nighttime displayed the possible presence of alcohols, ether, and esters (light brown labels). Overall, it was observed that the background sample was overlapping the real composition of the aerosols from the seasons. That can be explained due to the small concentration of the aerosol molecules in the filters. This observation leads us to be cautious in the identification of aerosol molecules.

4.5. 2D NMR Analysis

Typically, one dimensional NMR is based under the influence of a magnetic field where all non-exchangeable hydrogen atoms (i.e., bonded to a carbon atom) in a molecule present unique resonances due to electron density as determined by its molecular environment. As a result, a single molecule displays several unique peaks in the spectrum that increases the complexity of NMR spectra for mixtures of compounds. To get a further understanding of the interaction of the molecules and their environment, 2D NMR was performed due to the ability to distinguish which proton nuclei are coupled to each other *via* the chemical bonds. Examples of this include ^1H - ^1H total correlation spectroscopy (TOCSY) which is used to obtain information on protons connected by a chain of couplings, HSQC, and HMBC to determine the coupling between carbon and proton atoms separated by one bond and 2–4 bonds, respectively.

In this thesis, TOCSY, HSQC, and HMBC were needed due to the high complexity of the ^1H NMR spectra from the aerosol particles. Even, when 2D NMR spectroscopy is rarely used within aerosol science due to the mass required for analysis ($1 > \text{mg}$)¹³⁵, the obtained spectrum was studied to annotate possible atmospheric aerosols. However, one of the main inconveniences was the sample concentration. As a result, TOCSY spectra for Summer (09-16/06/21) and Winter (13-21/01/21) (Appendix 7) was limited due to the low obtained concentration. Similarly, the HSQC spectra from Winter (13-20/01/21) samples didn't show good resolution (Appendix 8). In contrast, the HMBC spectra from spring and α -pinene show many cross-peaks in region II_b (Appendix 9), being optimal for analysis.

The following analysis is based on Chalbot⁵⁷ regions between ^1H - ^1H correlations (A through J) and ^1H - ^{13}C correlations (I through VI) which correspond to functional group regions' chemical shift ranges (in ppm) as was described in Table 3.

Region A contains cross peaks corresponding to terminal CH_3 , mid-chain CH_2 groups, CH in aliphatic structures, long-chain α , ω -dicarboxylic acids, and hydrophobic amino acids ($\text{H}_2\text{N}-\text{CH}(\text{R})-\text{COOH}$)¹³⁶. In our analysis, TOCSY experiments demonstrated similar signals for all our samples (Appendix 7).

Region B includes intra-aliphatic couplings associated with alkylated amines, cross-peaks of long-chain alkyls, and highly oxidized carboxylic acid (keto-carboxylic acids or hydroxy-carboxylic acids). In our analysis α -pinene sample showed unique signals in regions B which were distinct from the background sample. In addition, it was identified the presence of amine compounds by cross-peaks in the regions II of the HSQC spectra (Appendix 8), and the portion I_b and II_d of HMBC from α -pinene and spring sample (Appendix 9). Note: a background sample analysis from HSQC and HMBC is required to further establish the identification of the signal.

Region C includes cross-peaks between aliphatic chain in the vicinity of oxo-acids and hydroxy-acids of general formula $\text{R}-\text{C}(=\text{O})-\text{COOH}$ and $\text{R}-\text{C}(\text{-OH})-\text{COOH}$, or $\text{R}-\text{CH}_2-\text{C}(=\text{O})-(\text{CH}_2)_n-\text{COOH}$. Signals of oxo-acids were exclusively observed in the spring and the α -pinene experiment (Appendix 7). The corresponding ^1H - ^{13}C cross-peaks were in region II in the HSQC spectra (Appendix 8) and the regions II_{a-d} of the HMBC spectra (Appendix 9), where the cross-peaks in the regions II_b and II_c indicated the presence of molecules containing carbonyl and carboxyl groups

in the α -pinene experiment. Interestingly, the signals in II_c (carboxyl) were not present in the spring sample.

Region D includes resonances of hydrogen in methyl (-CH₃), methylene (-CH₂-), or methyne (=CH-) groups and hydrogen adjacent to a carbon linked to a heteroatom. These compounds can be attributed to CH₂-OH groups (C₆ carbon) from carbohydrates, oxymethylene (-OCH₂-O-), aromatic methyl esters (R_{ar}-COOCH₃), aliphatic methyl esters (R_{al}-COOCH₃), and aromatic methyl ethers (R_{ar}-OCH₃), and polyols (-C-CH-OH)¹³⁷. In our samples, the spring and α -pinene experiments showed unique signals in this region evidencing the presence of these mentioned chemical groups.

Region E shows resonances associated with polyols (-C-CH-OH) including cyclic and linear carbohydrates, alcohol-sugars, and anhydro-sugars. It was found the presence of this compound in the TOCSY experiment (spring, Appendix 7) and further observed as well in the ¹H-¹³C HSQC (region III, IV, Appendix 8) and ¹H-¹³C HMBC (region III-IV, Appendix 9) spectra of spring.

Region G shows correlations with anomeric carbons of anhydro-sugars. In the TOCSY experiments, it was evidenced the presence of this region during the spring and winter seasons (δ 5.26, δ 4.21). In addition, the corresponding carbon-hydrogen correlations were found in the region V of the HSQC spectra (Appendix 8) from the spring season.

Region J shows the cross-peaks associated with long-range coupling between -CH₃ or (-CH₂-) groups with olefinic protons (C=CHR). It was observed a high number of signals (region J) during the spring season (05-12/08/20), and in the α -pinene experiment (22/10/20). Besides that, it was observed unique signals (δ 4.87, δ 1.27) share in summer nighttime (05-12/08/20), and spring season (05-12/08/20).

Region I includes cross peaks corresponding to aromatic functional groups. It was observed the absence of aromatic signals in the α -pinene experiment.

4.6. Compound Identification on Two-dimensional (2D) NMR with NMRfilter

The use of 2D-NMR spectroscopy allowed an enhanced identification of individual compounds and therefore a cleared picture of the remaining resonances associated with macromolecules. The results provided by NMRfilter software indicated the presence of alkanes, carboxylic acids,

nitrocompounds, organosulfur compounds with an HSQC matching rate of 100% and the parameter standard deviation over 0.5 (Table 4).

Using this methodology, some compounds annotated by GC-MS were confirmed by NMR analysis. For example, it was confirmed the following compounds in the spring season: alkanes, carboxylic acids, decanamide; disulfide, di-tert-dodecyl; tert-hexadecanethiol. In another analysis from the α -pinene experiment, it was found pinic acid and 3-formyl-2,2-dimethylcyclobutane-carboxylic acid.

Finally, possible reasons for a lack of detection during the NMR analysis of some compounds could be attributed to low concentration, symmetry of the molecule, the overlap of signals. However, the use NMRfilter for improvement of dereplication studies was demonstrated to help in annotations of compounds.

5. Conclusion

In the present study, we explored the possible organic composition of atmospheric aerosols in a Hemi-boreal forest during the winter, spring, and summer seasons. We used GC/MS, molecular networking, and NMR to characterize the atmospheric aerosols. In particular, we identified a variety of chemical families during the studied seasons where phthalates were predominant during winter, carboxylic acid in the summer, and organophosphate esters in the spring season. Furthermore, we presented distribution patterns of *n*-alkanol, carboxylic acid, and *n*-alkane in the suspended particulate aerosols as indicators of soil organic matter and hydrocarbon pollution. Finally, the presence of some air contaminants was annotated during the seasons including phthalates like Dibutyl phthalate (~20.59% in winter), Bis(2-ethylhexyl) phthalate (~3.87% in summer), and organophosphate compounds like Tris(2,4-di-tert-butylphenyl) phosphate (24.13% in spring) and tris(2,4-di-tertbutylphenyl) phosphite (5.13% in summer). Interestingly, compounds such as 7-Oxodehydroabietic (1.18% in spring) and dehydroabietic acid (0.27% in winter) were noted as combustion products from conifer trees. Overall, the use of molecular networking for the identification of organic aerosols was a straightforward and versatile tool to characterize a variety of possible chemical compounds, and the use of software such as NMRfilter can aid in further corroboration of the annotated compounds. We envisioned that such GC-MS workflow, combined with other metabolomics platforms like molecular networking and NMR would enable much wider applicability to aerosol research and be of great potential to obtain a faster screening of the chemical composition of atmospheric aerosols.

References

- 1 France OB, France CG, Germany CH, Uk AJ. Clouds and aerosols. Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA., 2013.
- 2 Yli-juuti T, Mielonen T, Heikkinen L, Arola A, Ehn M, Isokääntä S *et al.* Significance of the organic aerosol driven climate feedback in the boreal area. *Nature Communications* 2021; **12**. doi:10.1038/s41467-021-25850-7.
- 3 Kulmala M, Vehkamäki H, Petäjä T, Dal Maso M, Lauri A, Kerminen VM *et al.* Formation and growth rates of ultrafine atmospheric particles: A review of observations. *Journal of Aerosol Science* 2004; **35**: 143–176.
- 4 McFiggans G, Artaxo P, Baltensperger U, Coe H, Facchini MC, Feingold G *et al.* The effect of physical and chemical aerosol properties on warm cloud droplet activation. *Atmospheric Chemistry and Physics* 2006; **6**: 2593–2649.
- 5 Seinfeld JH, Pandis SN. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. 3rd ed. John Wiley & Sons: Hoboken, New Jersey, 2016.
- 6 Spracklen D v., Bonn B, Carslaw KS. Boreal forests, aerosols and the impacts on clouds and climate. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2008; **366**: 4613–4626.
- 7 Zhang Q, Jimenez JL, Canagaratna MR, Allan JD, Coe H, Ulbrich I *et al.* Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophysical Research Letters* 2007; **34**: 1–6.
- 8 Wolfender JL, Nuzillard JM, van der Hooft JJJ, Renault JH, Bertrand S. Accelerating Metabolite Identification in Natural Product Research: Toward an Ideal Combination of Liquid Chromatography-High-Resolution Tandem Mass Spectrometry and NMR Profiling, in Silico Databases, and Chemometrics. *Analytical Chemistry* 2019; **91**: 704–742.
- 9 Allard PM, Péresse T, Bisson J, Gindro K, Marcourt L, Pham VC *et al.* Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Analytical Chemistry* 2016; **88**: 3317–3323.
- 10 Watrous J, Roach P, Alexandrov T, Heath BS, Yang JY, Kersten RD *et al.* Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences of the United States of America* 2012; **109**: 1743–1752.
- 11 van der Hooft JJJ, Padmanabhan S, Burgess KEV, Barrett MP. Urinary antihypertensive drug metabolite screening using molecular networking coupled to high-resolution mass spectrometry fragmentation. *Metabolomics* 2016; **12**: 1–15.
- 12 Aksenov AA, Laponogov I, Zhang Z, Doran SLF, Belluomo I, Veselkov D *et al.* Auto-deconvolution and molecular networking of gas chromatography–mass spectrometry data. *Nature Biotechnology* 2021; **39**: 169–173.
- 13 EUR-Lex. Directive 2004/42/CE of the European Parliament and the Council of 21 April 2004. .
- 14 Goldstein AH, Galbally IE. Known and unexplored organic constituents in the earth’s atmosphere. *Environmental Science and Technology* 2007; **41**: 1514–1521.
- 15 Atkinson R, Arey J. Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. *Atmospheric Environment* 2003; **37**: 197–219.
- 16 Nault BA, Jo DS, McDonald BC, Campuzano-jost P, Day DA. Secondary organic aerosols from anthropogenic volatile organic compounds contribute substantially to air pollution mortality. *Atmospheric Chemistry and Physics* 2021; **21**: 11201–11224.

- 17 Cofer TM, Seidl-Adams I, Tumlinson JH. From Acetoin to (Z)-3-Hexen-1-ol: The Diversity of Volatile Organic Compounds that Induce Plant Responses. *Journal of Agricultural and Food Chemistry* 2018; **66**: 11197–11208.
- 18 Lathièrè J, Hauglustaine DA, Friend AD, de Noblet-Ducoudré N, Viovy N, Folberth GA. Impact of climate variability and land use changes on global biogenic volatile organic compound emissions. *Atmospheric Chemistry and Physics* 2006; **6**: 2129–2146.
- 19 Matsunaga SN, Mochizuki T, Ohno T, Endo Y, Kusumoto D, Tani A. Monoterpene and sesquiterpene emissions from Sugi (*Cryptomeria japonica*) based on a branch enclosure measurements. *Atmospheric Pollution Research* 2011; **2**: 16–23.
- 20 de Gouw J, Jimenez JL. Organic aerosols in the earth's atmosphere. *Environmental Science and Technology* 2009; **43**: 7614–7618.
- 21 Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zh[Masson-Delmotte, V., P. Zh RY, Zhouou B. Climate change 2021: The physical science basis Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. 2021 doi:10.1260/095830507781076194.
- 22 Bond TC, Doherty SJ, Fahey DW, Forster PM, Berntsen T, Deangelo BJ *et al.* Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research Atmospheres* 2013; **118**: 5380–5552.
- 23 Boucher O. *Atmospheric Aerosols. Propertis and Climate Impacts*. Springer, 2015 doi:10.1002/cpt1969103287.
- 24 Hems RF, Schnitzler EG, Liu-Kang C, Cappa CD, Abbatt JPD. Aging of Atmospheric Brown Carbon Aerosol. *ACS Earth and Space Chemistry* 2021; **5**: 722–748.
- 25 IPCC. Aerosol, their Direct and Indirect Effects. 2001 doi:10.1029/JD091iD01p01089.
- 26 Nguyen TB, Laskin A, Laskin J, Nizkorodov SA. Brown carbon formation from ketoaldehydes of biogenic monoterpenes. *Faraday Discussions* 2013; **165**: 473–494.
- 27 Feng Y, Ramanathan V, Kotamarthi VR. Brown carbon: A significant atmospheric absorber of solar radiation. *Atmospheric Chemistry and Physics* 2013; **13**: 8607–8621.
- 28 Yan J, Wang X, Gong P, Wang C, Cong Z. Review of brown carbon aerosols: Recent progress and perspectives. *Science of the Total Environment* 2018; **634**: 1475–1485.
- 29 Jimenez JL, Canagaratna MR, Donahue NM, Prevot ASH, Zhang Q, Kroll JH *et al.* Evolution of organic aerosols in the atmosphere. *Science* 2009; **326**: 1525–1529.
- 30 Atkinson R, Arey J. Atmospheric degradation of Volatile Organic Compounds. *Chemical Reviews* 2003; **103**: 4605–4638.
- 31 Ehn M, Thornton JA, Kleist E, Sipilä M, Junninen H, Pullinen I *et al.* A large source of low-volatility secondary organic aerosol. *Nature* 2014; **506**: 476–479.
- 32 Fu X, Wang S, Chang X, Cai S, Xing J, Hao J. Modeling analysis of secondary inorganic aerosols over China: Pollution characteristics, and meteorological and dust impacts. *Scientific Reports* 2016; **6**: 4–10.
- 33 CHARLSON RJ, LANGNER J, RODHE H, LEOVY CB, WARREN SG. Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic sulfate aerosols. *Tellus B* 1991; **43**: 152–163.
- 34 Dubois C, Cholleton D, Gemayel R, Chen Y, Surratt JD, George C *et al.* Decrease in sulfate aerosol light backscattering by reactive uptake of isoprene epoxydiols. *Physical Chemistry Chemical Physics* 2021; **23**: 5927–5935.

- 35 Palm BB, Campuzano-Jost P, Day DA, Ortega AM, Fry JL, Brown SS *et al.* Secondary organic aerosol formation from in situ OH, O₃, and NO₃ oxidation of ambient forest air in an oxidation flow reactor. *Atmospheric Chemistry and Physics* 2017; **17**: 5331–5354.
- 36 Heard DE, Pilling MJ. Measurement of OH and HO₂ in the Troposphere. *Chemical Reviews* 2003; **103**: 5163–5198.
- 37 Geyer A, Alicke B, Konrad S, Schmitz T, Stutz J, Platt U. Chemistry and oxidation capacity of the nitrate radical in the continental boundary layer near Berlin. *Journal of Geophysical Research Atmospheres* 2001; **106**: 8013–8025.
- 38 Yang Y, Wang Y, Zhou P, Yao D, Ji D, Sun J *et al.* Atmospheric reactivity and oxidation capacity during summer at a suburban site between Beijing and Tianjin. *Atmospheric Chemistry and Physics* 2020; **20**: 8181–8200.
- 39 Thomsen D, Elm J, Rosati B, Skønager JT, Bilde M, Glasius M. Large Discrepancy in the Formation of Secondary Organic Aerosols from Structurally Similar Monoterpenes. *ACS Earth and Space Chemistry* 2021. doi:10.1021/acsearthspacechem.0c00332.
- 40 Kroll JH, Seinfeld JH. Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere. *Atmospheric Environment* 2008; **42**: 3593–3624.
- 41 Joo T, Rivera-Rios JC, Alvarado-Velez D, Westgate S, Ng NL. Formation of Oxidized Gases and Secondary Organic Aerosol from a Commercial Oxidant-Generating Electronic Air Cleaner. *Environmental Science & Technology Letters* 2021; **8**: 691–698.
- 42 Junninen H, Ehn M, Pet T, Kulmala M, Worsnop DR. Techniques A high-resolution mass spectrometer to measure atmospheric ion composition. *Atmospheric Measurement Techniques* 2010; **3**: 1039–1053.
- 43 Zhou W, Xu W, Kim H, Zhang Q, Fu P, Worsnop DR *et al.* A review of aerosol chemistry in Asia: Insights from aerosol mass spectrometer measurements. *Environmental Science: Processes and Impacts* 2020; **22**: 1616–1653.
- 44 Zhu C, Kawamura K, Fu P. Seasonal variations of biogenic secondary organic aerosol tracers in Cape Hedo, Okinawa. *Atmospheric Environment* 2016; **130**: 113–119.
- 45 Pratt KA, Prather KA. Mass spectrometry of atmospheric aerosols—Recent developments and applications. Part II: On-line mass spectrometry techniques. *Mass Spectrometry Reviews* 2012; **31**: 17–48.
- 46 Bianchi F, Kurte T, Riva M, Mohr C, Rissanen MP, Roldin P *et al.* Highly Oxygenated Organic Molecules (HOM) from Gas-Phase Autoxidation Involving Peroxy Radicals: A Key Contributor to Atmospheric Aerosol. *Chemical Reviews* 2019; **119**: 3472–3509.
- 47 Zha Q, Yan C, Junninen H, Riva M, Sarnela N, Aalto J *et al.* Vertical characterization of highly oxygenated molecules (HOMs) below and above a boreal forest canopy. *Atmospheric Chemistry and Physics* 2018; **18**: 17437–17450.
- 48 Breitenlechner M, Fischer L, Hainer M, Heinritzi M, Curtius J, Hansel A. PTR3: An Instrument for Studying the Lifecycle of Reactive Organic Carbon in the Atmosphere. *Analytical Chemistry* 2017; **89**: 5824–5831.
- 49 Thornton JA, Mohr C, Schobesberger S, D’Ambro EL, Lee BH, Lopez-Hilfiker FD. Evaluating Organic Aerosol Sources and Evolution with a Combined Molecular Composition and Volatility Framework Using the Filter Inlet for Gases and Aerosols (FIGAERO). *Accounts of Chemical Research* 2020; **53**: 1415–1426.
- 50 Klyta J, Czaplicka M. Determination of secondary organic aerosol in particulate matter – Short review. *Microchemical Journal* 2020; **157**: 104997.

- 51 Tammekivi E, Vahur S, Kekišev O, van der Werf ID, Toom L, Herodes K *et al.* Comparison of derivatization methods for the quantitative gas chromatographic analysis of oils. *Analytical Methods* 2019; **11**: 3514–3522.
- 52 Laskin A, Laskin J, Nizkorodov SA. Mass spectrometric approaches for chemical characterisation of atmospheric aerosols: critical review of the most recent advances. *Environmental Chemistry* 2012; **9**: 163–189.
- 53 Mahilang M, Deb MK, Pervez S. Biogenic secondary organic aerosols: A review on formation mechanism, analytical challenges and environmental impacts. *Chemosphere* 2021; **262**: 127771.
- 54 Chen Q, Ikemori F, Nakamura Y, Vodicka P, Kawamura K, Mochida M. Structural and Light-Absorption Characteristics of Complex Water-Insoluble Organic Mixtures in Urban Submicrometer Aerosols. *Environmental Science and Technology* 2017; **51**: 8293–8303.
- 55 Duarte RMBO, Piñeiro-Iglesias M, López-Mahía P, Muniategui-Lorenzo S, Moreda-Piñeiro J, Silva AMS *et al.* Comparative study of atmospheric water-soluble organic aerosols composition in contrasting suburban environments in the Iberian Peninsula Coast. *Science of the Total Environment* 2019; **648**: 430–441.
- 56 Matos JTV, Duarte RMBO, Lopes SP, Silva AMS, Duarte AC. Persistence of urban organic aerosols composition: Decoding their structural complexity and seasonal variability. *Environmental Pollution* 2017; **231**: 281–290.
- 57 Chalbot MC, Siddiqui S, Kavouras IG. Molecular speciation of size fractionated particulate water-soluble organic carbon by two-dimensional nuclear magnetic resonance (NMR) spectroscopy. *International Journal of Environmental Research and Public Health* 2021; **18**: 1–19.
- 58 Erupe ME, Liberman-Martin A, Silva PJ, Malloy QGJ, Yonis N, Cocker DR *et al.* Determination of methylamines and trimethylamine-N-oxide in particulate matter by non-suppressed ion chromatography. *Journal of Chromatography A* 2010; **1217**: 2070–2073.
- 59 Dovrou E, Lim CY, Canagaratna MR, Kroll JH, Worsnop DR, Keutsch FN. Measurement techniques for identifying and quantifying hydroxymethanesulfonate (HMS) in an aqueous matrix and particulate matter using aerosol mass spectrometry and ion chromatography. *Atmospheric Measurement Techniques* 2019; **12**: 5303–5315.
- 60 Brent LC, Reiner JL, Dickerson RR, Sander LC. Method for characterization of low molecular weight organic acids in atmospheric aerosols using ion chromatography mass spectrometry. *Analytical Chemistry* 2014; **86**: 7328–7336.
- 61 Wang M, Carver JJ, Phelan V v., Sanchez LM, Garg N, Peng Y *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* 2016; **34**: 828–837.
- 62 Vincenti F, Montesano C, di Ottavio F, Gregori A, Compagnone D, Sergi M *et al.* Molecular Networking: A Useful Tool for the Identification of New Psychoactive Substances in Seizures by LC–HRMS. *Frontiers in Chemistry* 2020; **8**: 1–9.
- 63 Bandeira N. Protein identification by spectral networks analysis. *Methods in molecular biology (Clifton, NJ)* 2011; **694**: 151–168.
- 64 Noe SM, Niinemets Ü, Krasnova A, Krasnov D, Motallebi A, Kängsepp V *et al.* SMEAR Estonia: Perspectives of a large-scale forest ecosystem – atmosphere research infrastructure. *Forestry Studies* 2016; **63**: 56–84.

- 65 Djoumbou Feunang Y, Eisner R, Knox C, Chepelev L, Hastings J, Owen G *et al.* ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics* 2016; **8**: 1–20.
- 66 Kim HW, Wang M, Leber CA, Nothias L-F, Reher R, Kang K bin *et al.* NPCClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *Journal of Natural Products* 2021; **84**: 2795–2807.
- 67 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 2003; **13**: 2498–2504.
- 68 Decesari S, Facchini MC, Fuzzi S, Tagliavini E. Characterization of water-soluble organic compounds in atmospheric aerosol: A new approach. *Journal of Geophysical Research Atmospheres* 2000; **105**: 1481–1489.
- 69 Decesari S, Mircea M, Cavalli F, Fuzzi S, Moretti F, Tagliavini E *et al.* Source attribution of water-soluble organic aerosol by nuclear magnetic resonance spectroscopy. *Environmental Science and Technology* 2007; **41**: 2479–2484.
- 70 Kuhn S, Colreavy-Donnelly S, Santana de Souza J, Borges RM. An integrated approach for mixture analysis using MS and NMR techniques. *Faraday Discussions* 2019; **218**: 339–353.
- 71 Kuhn S, Colreavy-Donnelly S, de Andrade Silva Quaresma LE, de Andrade Silva Quaresma E, Borges RM. Applying NMR compound identification using NMRfilter to match predicted to experimental data. *Metabolomics* 2020; **16**: 1–5.
- 72 Kuhn S, Johnson SR. Stereo-Aware Extension of HOSE Codes. *ACS Omega* 2019; **4**: 7323–7329.
- 73 Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* 2007; **3**: 211–221.
- 74 Kawamura K, Seméré R, Imai Y, Fujii Y, Hayashi M. Water soluble dicarboxylic acids and related compounds in Antarctic aerosols. *Journal of Geophysical Research Atmospheres* 1996; **101**: 18721–18728.
- 75 Anttila P, Rissanen T, Shimmo M, Kallio M, Hyötyläinen T, Kulmala M *et al.* Organic compounds in atmospheric aerosols from a Finnish coniferous forest. *Boreal Environment Research* 2005; **10**: 371–384.
- 76 Yokouchi Y, Ambe Y. Characterization of polar organics in airborne particulate matter. *Atmospheric Environment* 1986; **20**: 1727–1734.
- 77 Simoneit BRT, Medeiros PM, Didyk BM. Combustion products of plastics as indicators for refuse burning in the atmosphere. *Environmental Science and Technology* 2005; **39**: 6961–6970.
- 78 Fu P, Kawamura K, Barrie LA. Photochemical and other sources of organic compounds in the Canadian high Arctic aerosol pollution during winter-spring. *Environmental Science and Technology* 2009; **43**: 286–292.
- 79 Fine PM, Cass GR, Simoneit BRT. Chemical characterization of fine particle emissions from fireplace combustion of woods grown in the northeastern United States. *Environmental Science and Technology* 2001; **35**: 2665–2675.
- 80 Oros DR, Simoneit BRT. Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 1. Temperate climate conifers. *Applied Geochemistry* 2001; **16**: 1513–1544.

- 81 Mpuru S, Blomquist GJ, Schal C, Roux M, Kuenzli M, Dusticier G *et al.* Effect of age and sex on the production of internal and external hydrocarbons and pheromones in the housefly, *Musca domestica*. *Insect Biochemistry and Molecular Biology* 2001; **31**: 935.
- 82 Zhang A, Oliver JE, Chauhan K, Zhao B, Xia L, Xu Z. Evidence for contact sex recognition pheromone of the Asian longhorned beetle, *Anoplophora glabripennis* (Coleoptera: Cerambycidae). *Naturwissenschaften* 2003; **90**: 410–413.
- 83 Ye P, Zhao Y, Chuang WK, Robinson AL, Donahue NM. Secondary organic aerosol production from pinanediol, a semi-volatile surrogate for first-generation oxidation products of monoterpenes. *Atmospheric Chemistry and Physics* 2018; **18**: 6171–6186.
- 84 Farooq A, Choudhary MI, Atta-ur-Rahman, Tahara S, Baser KH, Demirci F. The Microbial Oxidation of (-)- β -Pinene by *Botrytis cinerea*. *Zeitschrift fur Naturforschung - Section C Journal of Biosciences* 2002; **57**: 686–690.
- 85 Alves C, Pio C, Duarte A. Composition of extractable organic matter of air particles from rural and urban Portuguese areas. *Atmospheric Environment* 2001; **35**: 5485–5496.
- 86 Yu J, Griffin RJ, Cocker DR, Flagan RC, Seinfeld JH, Blanchard P. Observation of gaseous and particulate products of monoterpene oxidation in forest atmospheres. *Geophysical Research Letters* 1999; **26**: 1145–1148.
- 87 Ding X, Wang XM, Zheng M. The influence of temperature and aerosol acidity on biogenic secondary organic aerosol tracers: Observations at a rural site in the central Pearl River Delta region, South China. *Atmospheric Environment* 2011; **45**: 1303–1311.
- 88 Liu Q, Li L, Zhang X, Saini A, Li W, Hung H *et al.* Uncovering global-scale risks from commercial chemicals in air. *Nature* 2021; **600**: 456–461.
- 89 Zhang W, Wang P, Zhu Y, Wang D, Yang R, Li Y *et al.* Occurrence and human exposure assessment of organophosphate esters in atmospheric PM_{2.5} in the Beijing-Tianjin-Hebei region, China. *Ecotoxicology and Environmental Safety* 2020; **206**: 111399.
- 90 Wu Y, Venier M, Salamova A. Spatioseasonal Variations and Partitioning Behavior of Organophosphate Esters in the Great Lakes Atmosphere. *Environmental Science and Technology* 2020; **54**: 5400–5408.
- 91 Clark AE, Yoon S, Sheesley RJ, Usenko S. Spatial and Temporal Distributions of Organophosphate Ester Concentrations from Atmospheric Particulate Matter Samples Collected across Houston, TX. *Environmental Science and Technology* 2017; **51**: 4239–4247.
- 92 Castro-Jiménez J, González-Gaya B, Pizarro M, Casal P, Pizarro-Álvarez C, Dachs J. Organophosphate ester flame retardants and plasticizers in the global oceanic atmosphere. *Environmental Science and Technology* 2016; **50**: 12831–12839.
- 93 Li J, Xie Z, Mi W, Lai S, Tian C, Emeis KC *et al.* Organophosphate Esters in Air, Snow, and Seawater in the North Atlantic and the Arctic. *Environmental Science and Technology* 2017; **51**: 6887–6896.
- 94 Sührling R, Diamond ML, Scheringer M, Wong F, Pućko M, Stern G *et al.* Organophosphate esters in Canadian Arctic air: Occurrence, levels and trends. *Environmental Science and Technology* 2016; **50**: 7409–7415.
- 95 Liang K, Niu Y, Yin Y, Liu J. Evaluating the blank contamination and recovery of sample pretreatment procedures for analyzing organophosphorus flame retardants in waters. *Journal of Environmental Sciences (China)* 2015; **34**: 57–62.

- 96 Shi J, Xu C, Xiang L, Chen J, Cai Z. Tris(2,4-di- tert-butylphenyl)phosphate: An Unexpected Abundant Toxic Pollutant Found in PM2.5. *Environmental Science and Technology* 2020; **54**: 10570–10576.
- 97 Rincón AG, Calvo AI, Dietzel M, Kalberer M. Seasonal differences of urban organic aerosol composition an ultra-high resolution mass spectrometry study. *Environmental Chemistry* 2012; **9**: 298–319.
- 98 Özel MZ, Hamilton JF, Lewis AC. New sensitive and quantitative analysis method for organic nitrogen compounds in urban aerosol samples. *Environmental Science and Technology* 2011; **45**: 1497–1505.
- 99 Kourtchev I, Fuller S, Aalto J, Ruuskanen TM, McLeod MW, Maenhaut W *et al.* Molecular composition of boreal forest aerosol from Hyytiälä, Finland, using ultrahigh resolution mass spectrometry. *Environmental Science and Technology* 2013; **47**: 4069–4079.
- 100 Hwang HM, Park EK, Young TM, Hammock BD. Occurrence of endocrine-disrupting chemicals in indoor dust. *Science of the Total Environment* 2008; **404**: 26–35.
- 101 Cao XL. Phthalate Esters in Foods: Sources, Occurrence, and Analytical Methods. *Comprehensive Reviews in Food Science and Food Safety* 2010; **9**: 21–43.
- 102 Weschler CJ, Salthammer T, Fromme H. Partitioning of phthalates among the gas phase, airborne particles and settled dust in indoor environments. *Atmospheric Environment* 2008; **42**: 1449–1460.
- 103 Erythropel HC, Maric M, Nicell JA, Leask RL, Yargeau V. Leaching of the plasticizer di(2-ethylhexyl)phthalate (DEHP) from plastic containers and the question of human exposure. *Applied Microbiology and Biotechnology* 2014; **98**: 9967–9981.
- 104 Giam CS, Atlas E, Chan HS, Neff GS. Phthalate esters, PCB and DDT residues in the gulf of mexico atmosphere. *Atmospheric Environment (1967)* 1980; **14**: 65–69.
- 105 Xie Z, Ebinghaus R, Temme C, Lohmann R, Caba A, Ruck W. Occurrence and air-sea exchange of phthalates in the arctic. *Environmental Science and Technology* 2007; **41**: 4555–4560.
- 106 Deng M, Liang X, Du B, Luo D, Chen H, Zhu C *et al.* Beyond Classic Phthalates: Occurrence of Multiple Emerging Phthalate Alternatives and Their Metabolites in Human Milk and Implications for Combined Exposure in Infants. *Environmental Science and Technology Letters* 2021; **8**: 705–712.
- 107 Substance information Bis(2-ethylhexyl) isophthalate. European Chemicals Agency. 2021. <https://echa.europa.eu/substance-information/-/substanceinfo/100.004.826> (accessed 30 Dec2021).
- 108 Abdel-Shafy HI, Mansour MSM. A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation. *Egyptian Journal of Petroleum* 2016; **25**: 107–123.
- 109 Medeiros PM, Simoneit BRT. Source profiles of organic compounds emitted upon combustion of green vegetation from temperate climate forests. *Environmental Science and Technology* 2008; **42**: 8310–8316.
- 110 Tomaz S, Shahpoury P, Jaffrezo J-L, Lammel G, Perraudin E, Villenave E *et al.* One-year study of polycyclic aromatic compounds at an urban site in Grenoble (France): Seasonal variations, gas/particle partitioning and cancer risk estimation. *Science of The Total Environment* 2016; **565**: 1071–1083.

- 111 Wang DG, Norwood W, Alaei M, Byer JD, Brimble S. Review of recent advances in research on the toxicity, detection, occurrence and fate of cyclic volatile methyl siloxanes in the environment. *Chemosphere* 2013; **93**: 711–725.
- 112 Xu J, Harrison RM, Song C, Hou S, Wei L, Fu P *et al.* PM_{2.5}-bound silicon-containing secondary organic aerosols (Si-SOA) in Beijing ambient air. *Chemosphere* 2022; **288**: 132377.
- 113 Whelan MJ, Estrada E, van Egmond R. A modelling assessment of the atmospheric fate of volatile methyl siloxanes and their reaction products. *Chemosphere* 2004; **57**: 1427–1437.
- 114 McLachlan MS, Kierkegaard A, Hansen KM, van Egmond R, Christensen JH, Skjøth CA. Concentrations and fate of decamethylcyclopentasiloxane (D5) in the atmosphere. *Environmental Science and Technology* 2010; **44**: 5365–5370.
- 115 Rushdi AI, El-Mubarak AH, Lijotra L, Al-Otaibi MT, Qurban MA, Al-Mutlaq KF *et al.* Characteristics of organic compounds in aerosol particulate matter from Dhahran city, Saudi Arabia. *Arabian Journal of Chemistry* 2017; **10**: S3532–S3547.
- 116 Oros DR, Simoneit BRT. *Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 1. Temperate climate conifers*. 2001 doi:10.1016/S0883-2927(01)00021-X.
- 117 Kolattukudy PE. Biopolyester membranes of plants: Cutin and suberin. *Science* 1980; **208**: 990–1000.
- 118 Abas M, Simoneit B. Wax Lipids from Leaf Surfaces of Some Common plants of Malaysia. *Pertanika Journal of Science & Technology* 1998; **6**: 171–182.
- 119 Fang M, Zheng M, Wang F, Chim KS, Kot SC. The long-range transport of aerosols from northern China to Hong Kong - A multi-technique study. *Atmospheric Environment* 1999; **33**: 1803–1817.
- 120 Kawamura K, Gagosian RB. Implications of omega-oxocarboxylic acids in the remote marine atmosphere for photo-oxidation of unsaturated fatty acids. *Nature* 1987; **325**: 330–332.
- 121 Rogge WF, Hildemann LM, Mazurek MA, Cass GR, Simoneit BRT. Sources of fine organic aerosol. 9. Pine, oak, and synthetic log combustion in residential fireplaces. *Environmental Science and Technology* 1998; **32**: 13–22.
- 122 Vicente A, Alves C, Monteiro C, Nunes T, Mirante F, Evtyugina M *et al.* Measurement of trace gases and organic compounds in the smoke plume from a wildfire in Penedono (central Portugal). *Atmospheric Environment* 2011; **45**: 5172–5182.
- 123 Tanonaka T, Weng JH, Hatakeyama S, Bandow H, Takagi H, Akimoto H. Ozone-Cyclohexene Reaction in Air: Quantitative Analysis of Particulate Products and the Reaction Mechanism. *Environmental Science and Technology* 1985; **19**: 935–942.
- 124 Ziemann PJ. Evidence for low-volatility diacyl peroxides as a nucleating agent and major component of aerosol formed from reactions of O₃ with cyclohexene and homologous compounds. *Journal of Physical Chemistry A* 2002; **106**: 4390–4402.
- 125 Gelencsér A, Mészáros T, Blazsó M, Kiss G, Krivácsy Z, Molnár A *et al.* Structural characterisation of organic matter in fine tropospheric aerosol by pyrolysis-gas chromatography-mass spectrometry. *Journal of Atmospheric Chemistry* 2000; **37**: 173–183.
- 126 Kołodziejczyk A, Pyrcz P, Pobudkowska A, Błaziak K, Szmigielski R. Physicochemical Properties of Pinic, Pinonic, Norpinic, and Norpinonic Acids as Relevant α -Pinene Oxidation Products. *Journal of Physical Chemistry B* 2019; **123**: 8261–8267.

- 127 Wang Z, Ehn M, Rissanen MP, Garmash O, Quéléver L, Xing L *et al.* Efficient alkane oxidation under combustion engine and atmospheric conditions. *Communications Chemistry* 2021; **4**: 1–8.
- 128 Zhang J-G, Liang Y-N, Wan J-W, Sun B-S. Analysis of compounds in airborne dust collected in Beijing. In: *Electrical Contacts - 1998. Proceedings of the Forty-Fourth IEEE Holm Conference on Electrical Contacts (Cat. No.98CB36238)*. 1998, pp 166–171.
- 129 Kessler SH, Nah T, Carrasquillo AJ, Jayne JT, Worsnop DR, Wilson KR *et al.* Formation of secondary organic aerosol from the direct photolytic generation of organic radicals. *Journal of Physical Chemistry Letters* 2011; **2**: 1295–1300.
- 130 Nishintha N, Premalatha K, Chinniah C, Vellaikumar S, Nalini R, Shanthi M. Volatile profile of yellow stem borer, *Scirpophaga incertulas* (Walker) damaged rice plants. *Ijcs* 2019; **7**: 2636–2638.
- 131 Jinfeng EC, Mohamad Rafi MI, Chai Hoon K, Kok Lian H, Yoke Kqueen C. Analysis of chemical constituents, antimicrobial and anticancer activities of dichloromethane extracts of *Sordariomycetes* sp. endophytic fungi isolated from *Strobilanthes crispus*. *World Journal of Microbiology and Biotechnology* 2017; **33**: 1–19.
- 132 Leithead A, Li SM, Hoff R, Cheng Y, Brook J. Levoglucosan and dehydroabietic acid: Evidence of biomass burning impact on aerosols in the Lower Fraser Valley. *Atmospheric Environment* 2006; **40**: 2721–2734.
- 133 Oliveira TS, Pio CA, Alves CA, Silvestre AJD, Evtugina M, Afonso J v. *et al.* Seasonal variation of particulate lipophilic organic compounds at nonurban sites in Europe. *Journal of Geophysical Research Atmospheres* 2007; **112**. doi:10.1029/2007JD008504.
- 134 Smith JS, Laskin A, Laskin J. Molecular Characterization of Biomass Burning Aerosols Using High-Resolution Mass Spectrometry. *Analytical Chemistry* 2009; **81**: 1512–1521.
- 135 Duarte RMBO, Duarte AC. Unraveling the structural features of organic aerosols by NMR spectroscopy: A review. *Magnetic Resonance in Chemistry* 2015; **53**: 658–666.
- 136 Chalbot MCG, Chitranshi P, Gamboa da Costa G, Pollock E, Kavouras IG. Characterization of water-soluble organic matter in urban aerosol by ¹H-NMR spectroscopy. *Atmospheric Environment* 2016; **128**: 235–245.
- 137 Duarte RMBO, Matos JTV, Paula AS, Lopes SP, Pereira G, Vasconcellos P *et al.* Structural signatures of water-soluble organic aerosols in contrasting environments in South America and Western Europe. *Environmental Pollution* 2017; **227**: 513–525.

Appendix

Appendix 1. Summary of sampling of aerosol particles and analysis conditions by GC and NMR

Samples	Sample 1: Summer- daytime	Sample 2: Summer- daytime	Sample 3: Summer- nighttime	Sample 4: Summer - Blank	Sample 5: Winter- daytime	Sample 6: Spring_1- day and night	Sample 7: Background (gas phase)	Sample 8: Summer: daytime	Sample 9: α -pinene Oxidation (F4)	Sample 10: α -pinene Oxidation (F3)	Sample 11: α -pinene Oxidation (F2) AFilter1 Asolution1
Sampling date	05/08/20-12/08/20				13/01/21- 20/01/21	05/05/21- 12/05/21	09/06/21- 16/06/21	09/06/21- 16/06/21	22-10-20	22-09-20	23-02-21
Filter exposure time to forest aerosols	72 h (Day shift from 6 am to 6 pm, 1 week). Pump conditions: On		71.46 h (Night shift from 6 pm to 6 am, 1 week). Pump conditions: On	5 min Pump conditions: Off Hepa filter: Not present	71.15 h (Day shift from 6 am to 6 pm, 1 week). Pump conditions: On	143.78 h Day and night shift. Pump conditions: On	68.85 h (Day shift from 6 am to 6 pm, 1 week). Pump conditions: On Hepa filter: present between the inlet and filter holder	68.85 h (Day shift from 6 am to 6 pm, 1 week). Pump conditions: On	Reaction time =3 hours + 4 min	Reaction time =3 hours +36 min	Reaction time =4 hours +50 min
Extraction solvent	CDCl ₃ + 0.03 % (v/v) TMS	CD ₃ OD + 0.03 % (v/v) TMS	CDCl ₃ + 0.03 % (v/v) TMS	CDCl ₃ + 0.03 % (v/v) TMS	CDCl ₃ + 0.03 % (v/v) TMS	CDCl ₃ + 0.03 % (v/v) TMS	CDCl ₃ + 0.03 % (v/v) TMS	CDCl ₃ + TMS +Ag	CDCl ₃ + 0.03 % (v/v) TMS	D ₂ O	NA
Extraction conditions	Vortex agitation: 10 min Filtration: cotton filter	Vortex agitation: 10 min Filtration: cotton filter	Vortex agitation: 10 min Filtration: cotton filter	Vortex agitation: 10 min Filtration: cotton filter	Sonication time: 30 min Filtration: cotton filter	Sonication time: 99 min (11-20 °C) Filtration: cotton filter Evaporation: To 0.6 mL with N ₂ flow	Sonication time: 99 min, Filtration: cotton filter Evaporation: with N ₂ flow	Sonication time: 60 min, Filtration: cotton filter Evaporation: with N ₂ flow	Sonication time: 120 min Centrifugation: 30 min Filtration: cotton filter Evaporation: with N ₂ flow	Sonication time: 120 min Centrifugation: 30 min Filtration: cotton filter Evaporation: with N ₂ flow	NA

Collected aerosol mass (mg)	2	NA: a 2 nd extraction was performed	1.6	0.3	4	7.6	0.8	2.9	78.1	92.3	NA
Solvent mass (g)	3.12	1.33	4.26	1.50	6.03	9.09	5.82	3.65	3.76	2.17	NA
Solvent density (g/mL)	1.50	0.89	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.11	-
Volume (mL)	2.083066667	1.499324	2.841733	1	4.0184	6.0622	3.881466667	2.4362667	2.508933	1.957995	-
Volume TMS (mL)	0.00062492	0.00045	0.000853	0.0003	0.001206	0.001819	0.00116444	-	0.000753	-	-
Flow (L/min)	50		50	-	47.26	49.6/50.67	50.05	50.62	24.3	27.3	24.47-
Volume (m ³)	504	504	504	1	476.3	1010.7	504.5	255.13	4471.2	5896.8	7096.3
Particle mass concentration (mg/m ³)	0.003968		0.003175	NA	0.008398	0.00752	0.00158572	0.011367	0.017467	0.015653	
NMR analysis	¹ H NMR	¹ H NMR	¹ H NMR TOCSY	¹ H NMR	¹ H NMR, ¹³ C NMR, TOCSY, HSQC,	¹ H NMR, ¹³ C NMR TOCSY, HSQC, HMBC	¹ H NMR, TOCSY	¹ H NMR, TOCSY	¹ H NMR, ¹³ C NMR TOCSY, HSQC, HMBC	¹ H NMR, ¹³ C NMR, TOCSY, HSQC, HMBC	Not performed
Derivatization with TMTFTH	Not performed				no	yes	yes	yes	Not performed	Not performed	no
GC Analysis Dates	Not performed				14/05/21	28/06/21	17/08/21	17/09/21	Not performed	Not performed	27/03/21 29/03/21

All the filters were preconditioned by baking at 200°C for 4 hours and weighted right after preconditioning. Storage of samples at -4 °C

Appendix 2. Compounds annotated from aerosol particles from the background, by GC/MS obtained from GNPS public library

	Ret Time (min)	Compound Name	CAS #	Node	Functional Group before derivatization	Cluster	MQScore	Normalized Intensity (%)
1	30.06	Diisobutyl phthalate (DiBP)	84-69-5	543	Phthalate	6	0.92	0.5
2	48.36	Bis(2-ethylhexyl) phthalate (DEHP)	117-81-7	1016	Phthalate	6	0.73	0.4
3	32.16	Di- <i>n</i> -octyl phthalate	117-84-0	599	Phthalate	6	0.82	0.3
4	48.44	Di(2-ethyl hexyl)isophthalate	137-89-3	1017	Phthalate	6	0.59	0.2
5	48.44	Bis(2-ethylhexyl) phthalate	117-81-7	1018	Phthalate	6	0.80	0.1
6	22.79	Diethyl Phthalate (DEP)	84-66-2	381	Phthalate	6	0.81	0.04
7	12.13	Geranyl acetone	3796-70-1	121	Terpenoids	75	0.71	0.0540
8	13.72	(1R,2R,3S,5R)-(-)-2,3-Pinandediol	22422-34-0	153	Terpenoids	7	*	0.0007
9	17.85	Geranyl acetone	3796-70-1	259	Terpenoids	7	0.78	0.0109
10	16.87	Isocaulolol	5172-21-4	234	Terpenoids	7	0.77	0.3
11	27.52	D-Limonene 1,2-epoxide	1195-92-2	489	Terpenoids	27	0.66	<0.00001
12	32.98	3-Octanol, 3,7-dimethyl-	78-69-3	624	Terpenoids	4	0.71	0.2
13	39.87	3,7,11,15-Tetramethyl-1-hexadecen-3-ol	505-32-8	803	Terpenoids	21	0.79	0.0008
14	45.29	7-Oxodehydroabiatic acid, methyl ester	110936-78-2	937	Terpenoids	62	*	0.0783
15	57.95	Isophytol	505-32-8	1301	Terpenoids	53	0.83	0.0032
16	24.19	Linalool	78-70-6	406	Terpenoids	39	0.7	0.0033
17	41.58	Dehydroabiatic acid	1740-19-8	853	Terpenoids	52	0.65	0.0023
18	40.71	9,10-Anthracenedione, 2-amino-3-hydroxy	1740-19-8	829	Ketones	52	0.73	0.2

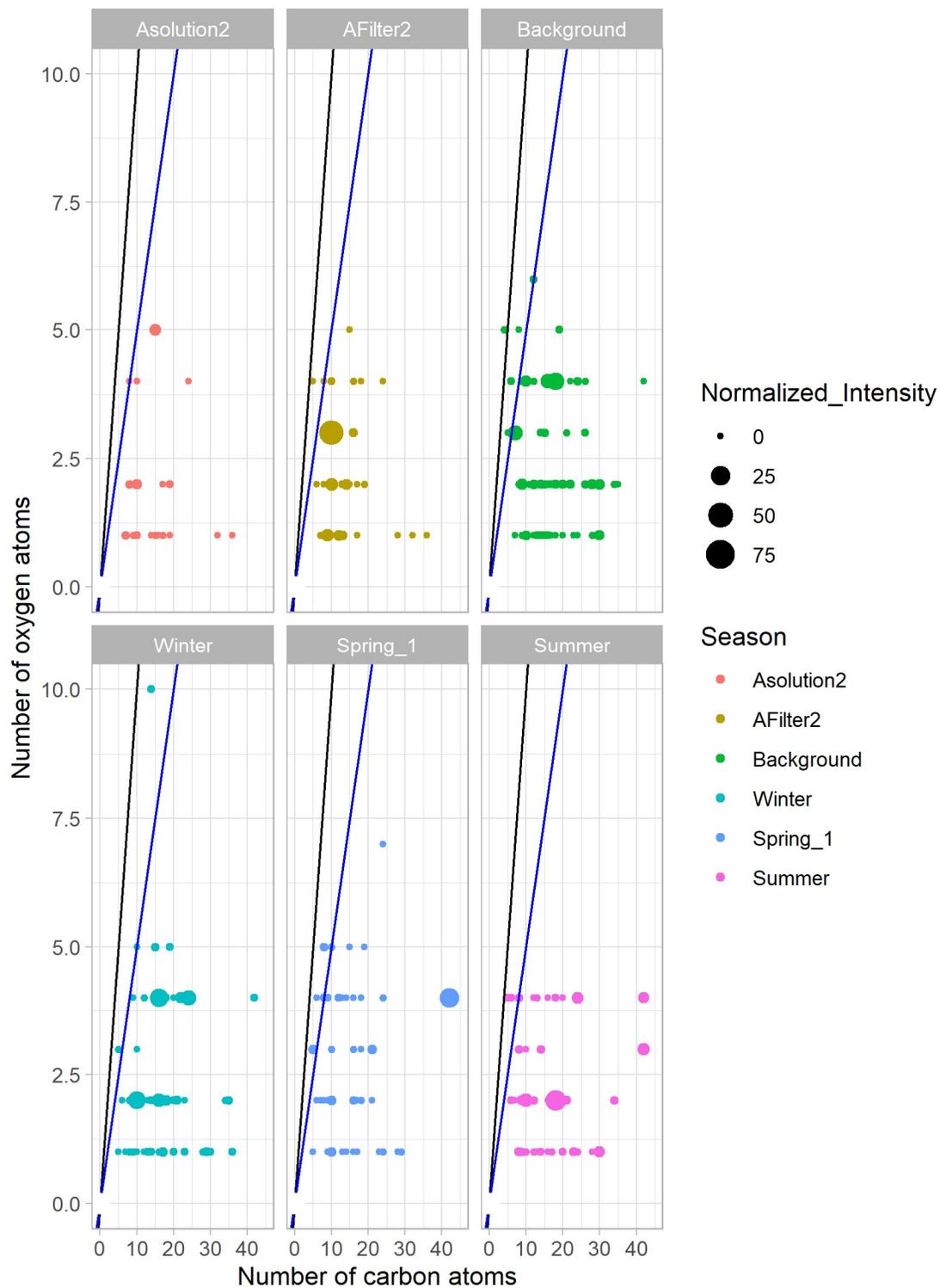
* Identified by NIST library without MQScore analysis

Appendix 3. Annotated compounds with a relative intensity higher than 1% identified from aerosol particles in winter, spring, and summer obtained from GNPS public library.

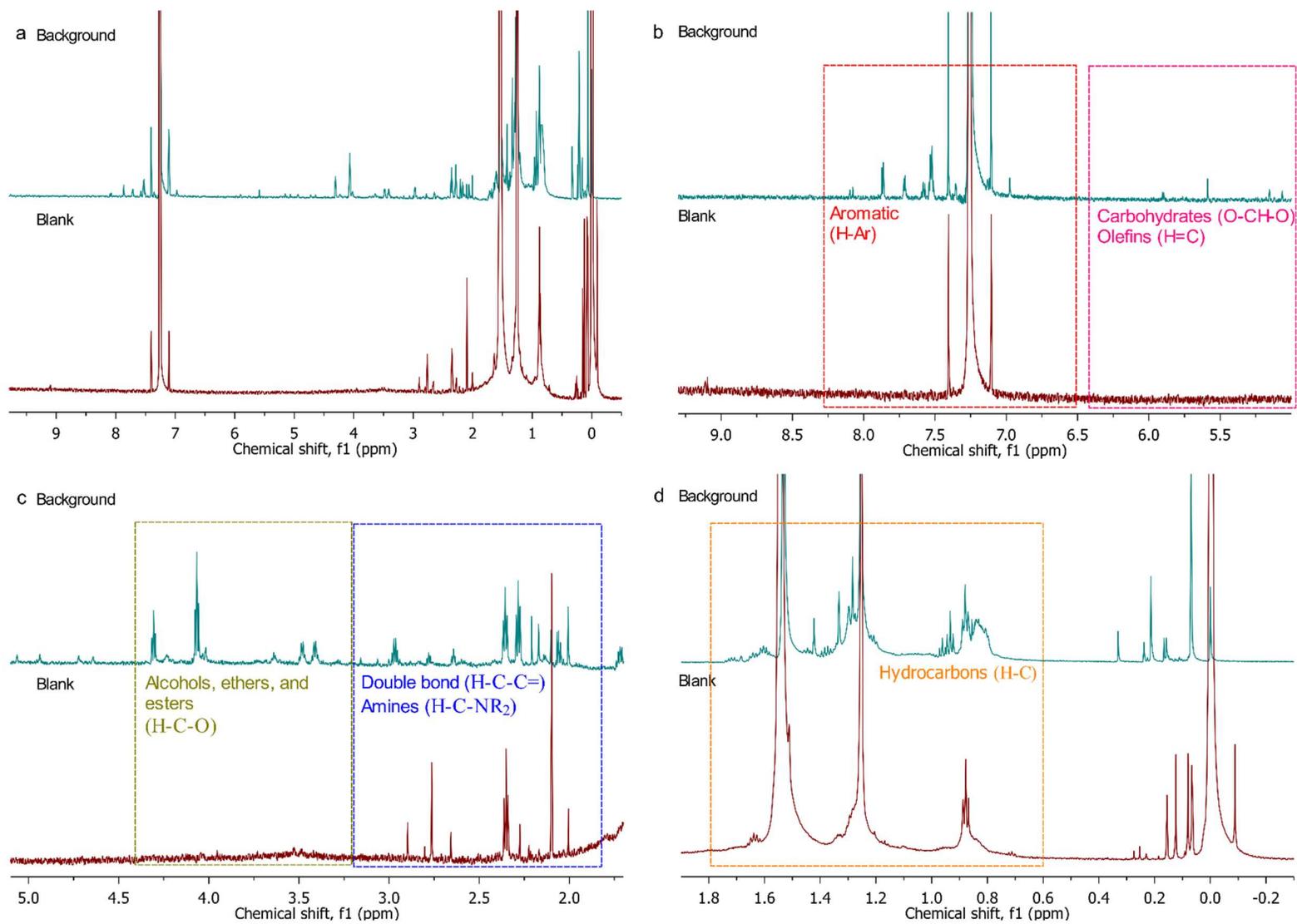
Seasons	Ret. Time (min)	Compound Name	CAS #	Node	Functional Group before derivatization	Cluster #	MQ Score	Normalized Intensity (%)	Total concentration (%)
Winter	29.44	Benzenesulfonamide, N-butyl-	3622-84-2	532	Organosulfur compounds	12	0.922	18.35	81.17
	33.84	Hexadecanoic acid	57-10-3	644	Carboxylic acid	49	0.864	6.60	
	42.57	Dioctyl adipate	123-79-5	879	Esters	34	0.885	2.75	
	38.55	Docosane, 1-iodo-	62127-53-1	770	Organohalogen	2	0.989	2.65	
	34.03	Eicosane, 1-iodo-	34994-81-5	648	Organohalogen	2	0.962	2.09	
	42.68	Tetracosane	646-31-1	883	Alkanes	2	0.936	2.02	
	46.5	Hexacosane, 1-iodo-	52644-81-2	977	Organohalogen	2	0.92	1.91	
	48.3	Octacosane, 1-iodo	62154-80-7	1015	Organohalogen	2	0.92	1.59	
	50.04	Octacosane, 1-iodo	62154-80-7	1077	Organohalogen	2	0.92	1.52	
	44.62	Hexacosane, 1-iodo-	52644-81-2	931	Organohalogen	2	0.97	1.33	
	46.82	Dotriacontane, 1-iodo-	62154-83-0	982	Organohalogen	14	0.749	1.12	
	50.3	Isopropyl tetracosyl ether		1085	Ether	14	0.77	1.07	
	51.7	1-Iodotriacontane	630-07-9	1118	Organohalogen	2	0.92	1.04	
	40.63	Hexacosane, 1-iodo-	52644-81-2	825	Organohalogen	2	0.91	1.04	
		33.29	Dibutyl phthalate	84-74-2	629	Phthalate	6	0.905	
45.43		Bis(2-ethylhexyl) phthalate	117-81-7	940	Phthalate	6	0.887	10.93	
30.83		Diisobutyl phthalate	84-69-5	560	Phthalate	6	0.938	1.92	
38.3		Stearic acid	57-11-4	760	Carboxylic acid	49	0.844	2.65	
Spring	56.41	Tris(2,4-di-tert-butylphenyl) phosphate	95906-11-9	1243	Organophosphate Esters	25	*	24.19	29.28
	29.44	Benzenesulfonamide, N-butyl-	3622-84-2	532	Organosulfur compounds	12	0.922	1.44	
	56.68	cis-Geraniol	106-26-3	1253	Terpenoids	17	0.65	1.44	
	45.29	7-Oxodehydroab	110936-78-2	937	Terpenoids	62	*	1.18	

		ietic acid, methyl ester							
	43.71	Levulinic Acid	123-76-2	912	Carboxylic acids	3	0.74	1.03	
Spring also in Background	32.52	Diisobutyl phthalate	84-69-5	606	Phthalate	6	0.86	45.72	61.2
	36.39	Methyl stearate	112-61-8	709	Carboxylic acids	1	0.935	3.06	
	40.71	9,10-Anthracenedione, 2-amino-3-hydroxy	117-77-1	829	Ketones	52	0.704	3.05	
	37.33	Decanedioic acid, dibutyl ester	109-43-3	743	Esters	11	0.78	1.35	
	31.63	Methyl myristate	124-10-7	586	Carboxylic acids	1	0.89	8.02	
Summer	36.39	Methyl stearate	112-61-8	709	Carboxylic acid	1	0.935	29.36	61.99
	28.53	Benzenesulfonamide, N-butyl-	3622-84-2	509	Organosulfur compounds	12	0.91	6.47	
	55.42	Tris(2,4-di-tert-butylphenyl) phosphite	31570-04-4	1207	Organophosphate Esters	84	*	5.13	
	48.44	Bis(2-ethylhexyl) phthalate	117-81-7	1018	Phthalate	6	0.805	3.87	
	51.04	(Z)-9-Heptacosene	36258-12-5	1103	Alkenes	43	0.961	3.44	
	48.44	Di(2-ethylhexyl)isophthalate	137-89-3	1017	Phthalate	6	0.59	3.37	
	52.99	Methyl triacontyl ether	237742-64-2	1144	Alcohol	43	0.94	3.32	
	57.19	Tris(2,4-di-tert-butylphenyl) phosphate	95906-11-9	1274	Organophosphate Esters	25	*	3.28	
	47.57	Hexacosane	593-45-3	997	Alkanes	2	0.97	2.02	
	6.81	Benzyl acetate	140-11-4	51	Esters	82	0.801	1.73	
Summer also in Background	32.52	Diisobutyl phthalate	84-69-5	606	Phthalate	6	0.86	9.41	12.03
	11.97	Methyl pelargonate	1731-84-6	116	Carboxylic acid	1	1	1.36	
	40.47	Methyl Nonadecanoate	1731-94-8	820	Carboxylic acid	1	0.94	1.26	

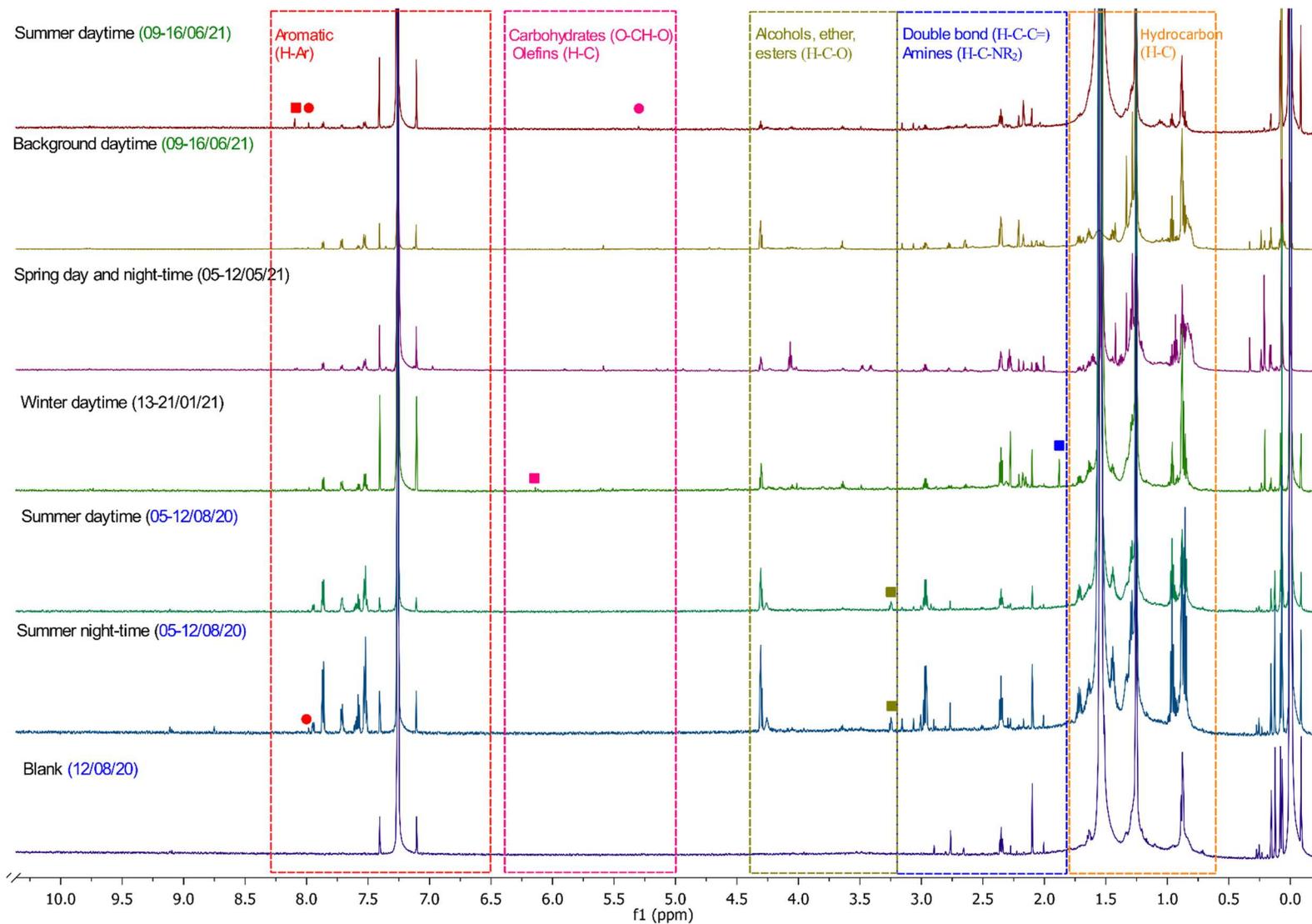
* Identified by NIST library without MQScore analysis



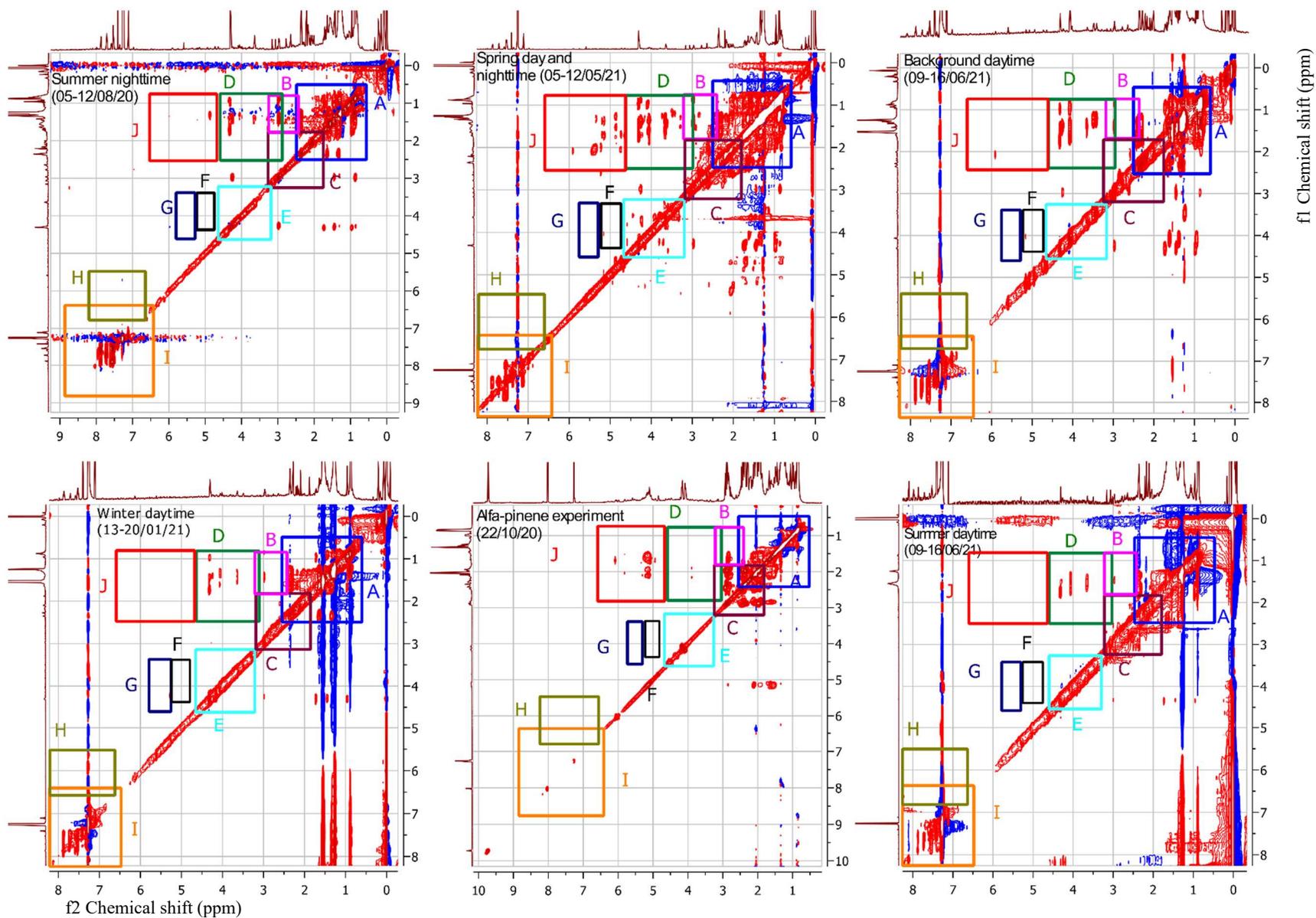
Appendix 4. The number of oxygen atoms vs the number of carbon atoms for each chemical formula in the liquid phase (Asolution2), and particle-phase (Afilter2, background and sample seasons in SMEAR Estonia). The blue and black lines in the figures correspond to O:C = 0.5 and O:C = 1 ratios, respectively, for reference.



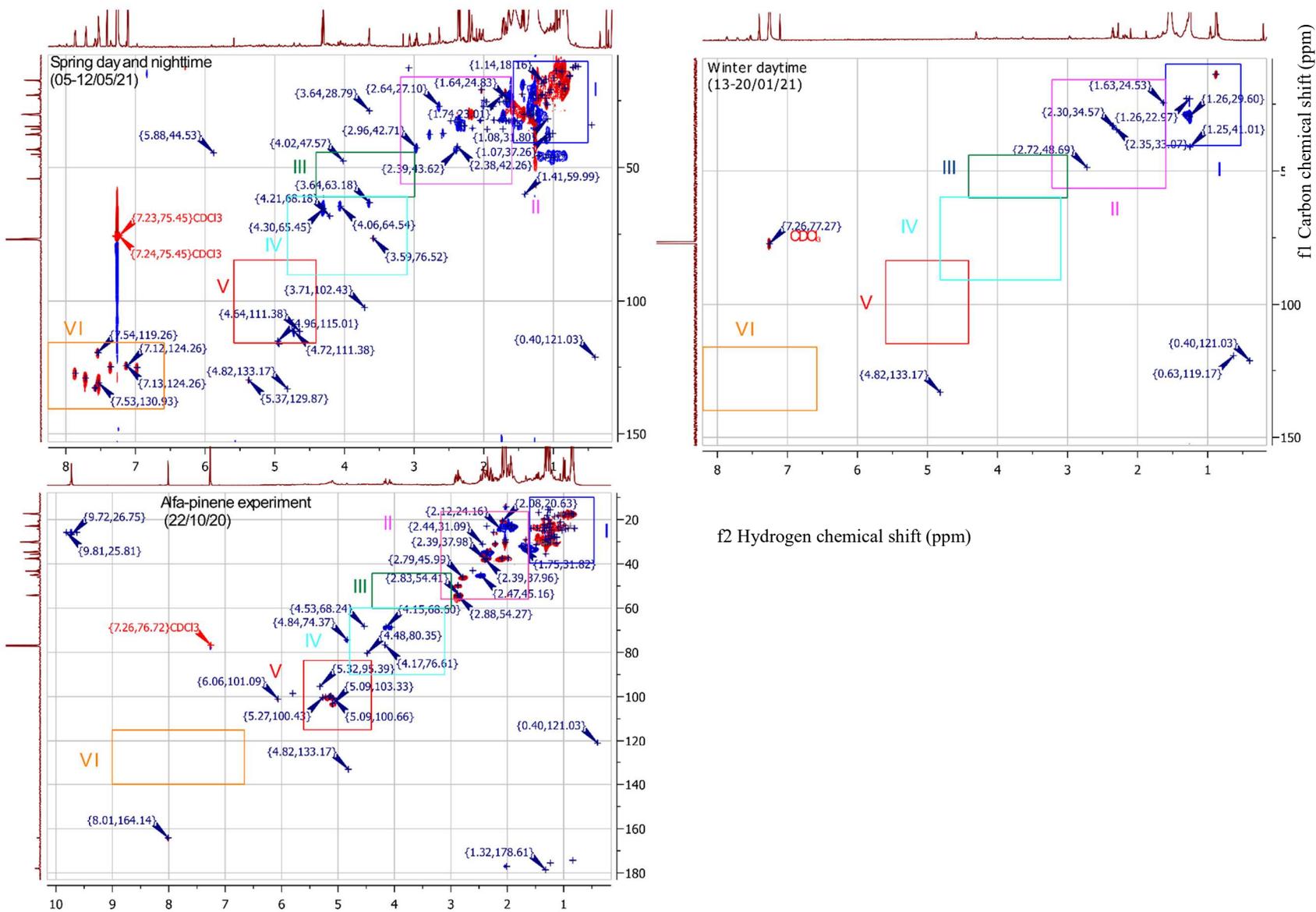
Appendix 5. Analysis of the functional groups identified in the 700 MHz ¹H NMR spectra of the CDCl₃ extraction of blank and background samples. (a) Full ¹H NMR spectra. (b) Zoom view from 5 to 9.3 ppm. (c) Zoom view from 1.8 to 5 ppm. (d) Zoom view from 0 to 1.8 ppm



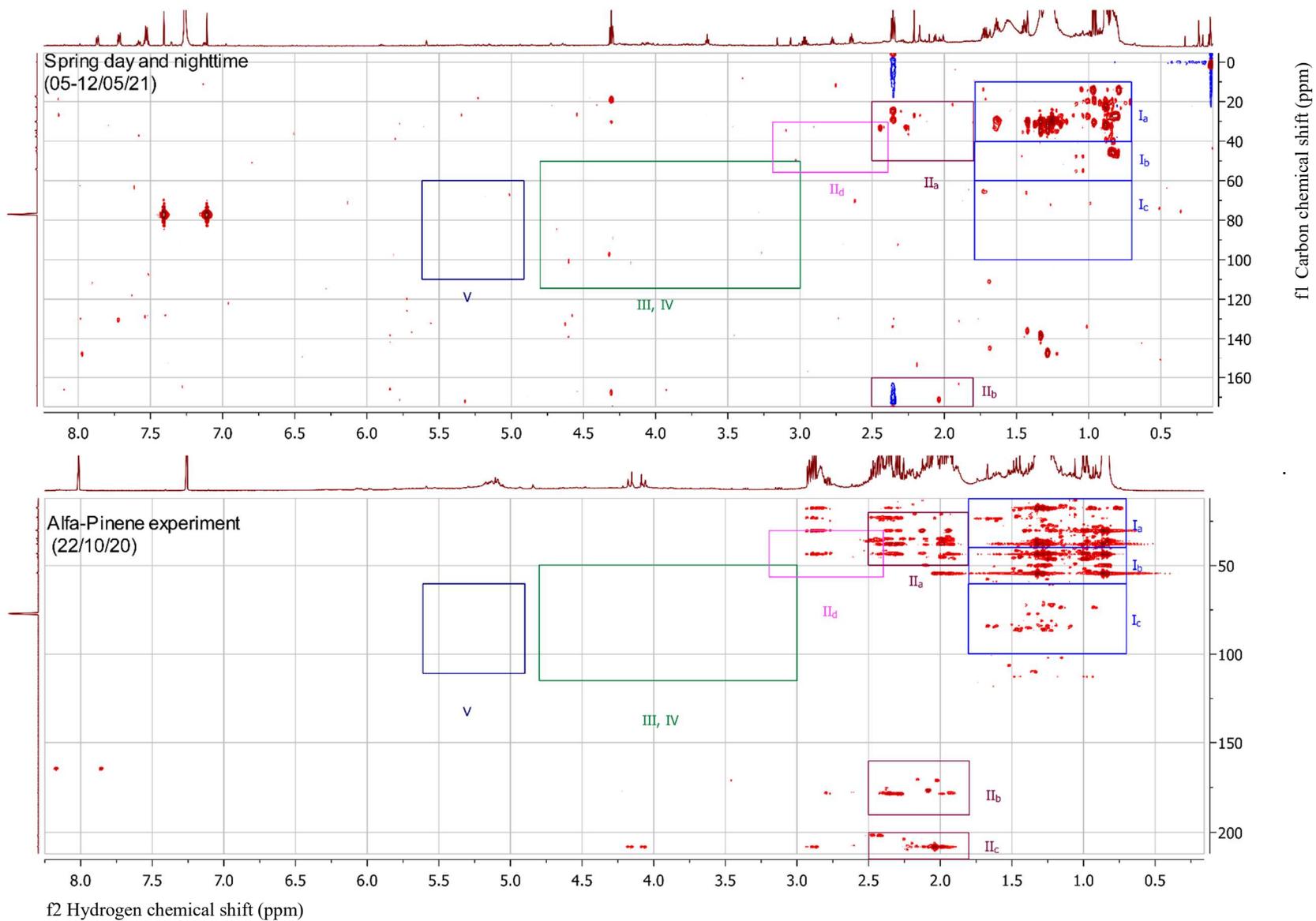
Appendix 6. Analysis of the functional groups identified in the 700 MHz ¹H NMR spectra of the CDCl₃ extraction of blank, background, and ambient aerosol samples.



Appendix 7. ^1H - ^1H TOCSY NMR spectra of the CDCl_3 extraction of α -pinene experiment, background, and ambient aerosol samples.



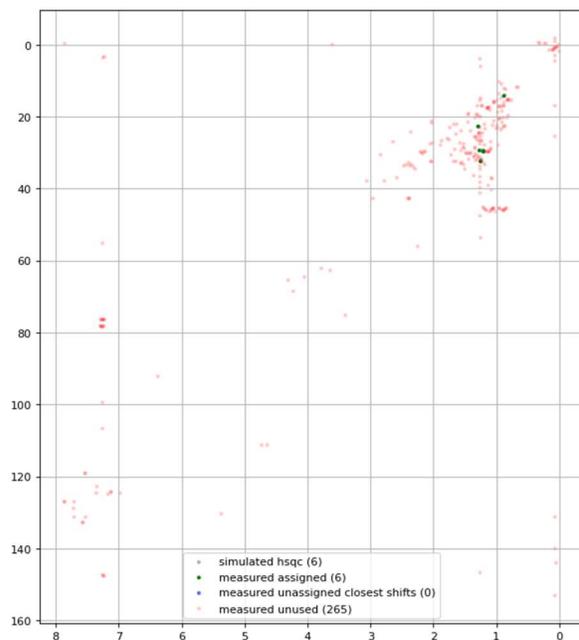
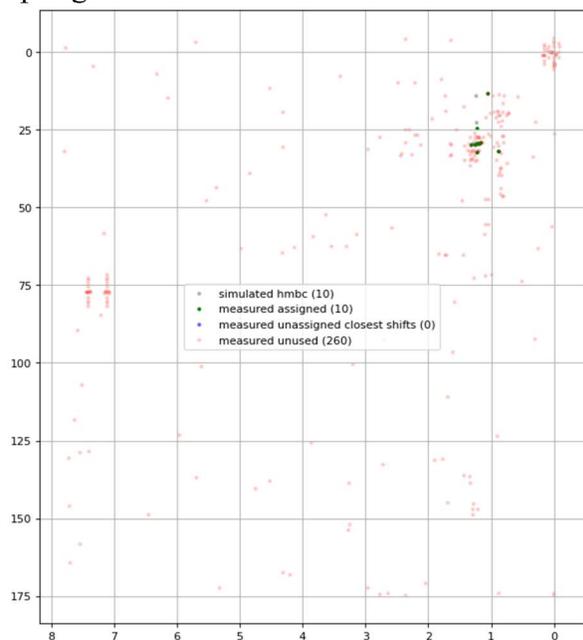
Appendix 8. ^1H - ^{13}C HSQC NMR spectra of the CDCl_3 extraction of α -pinene experiment and ambient aerosol samples.



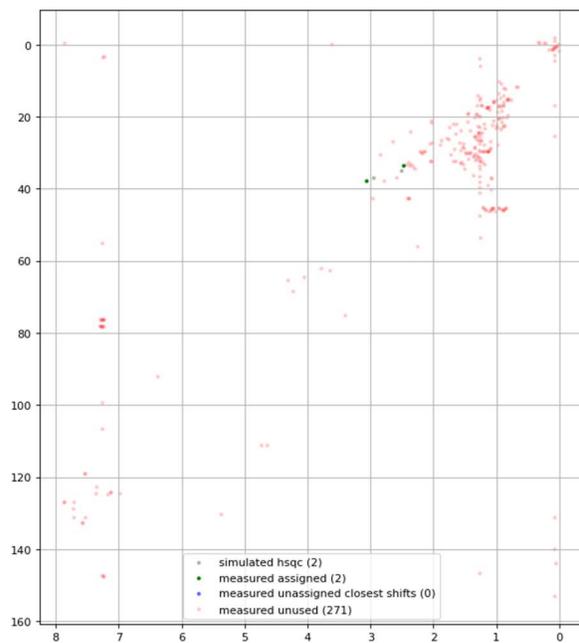
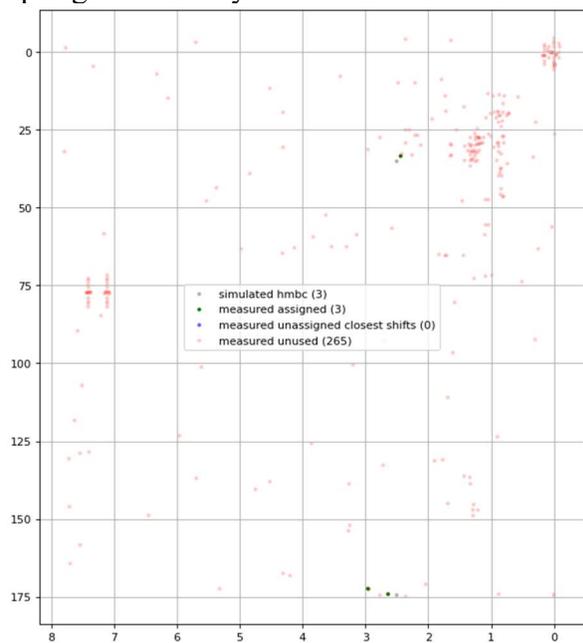
Appendix 9. ^1H - ^{13}C HMBC NMR spectra of the CDCl_3 extraction of α -pinene experiment and spring sample.

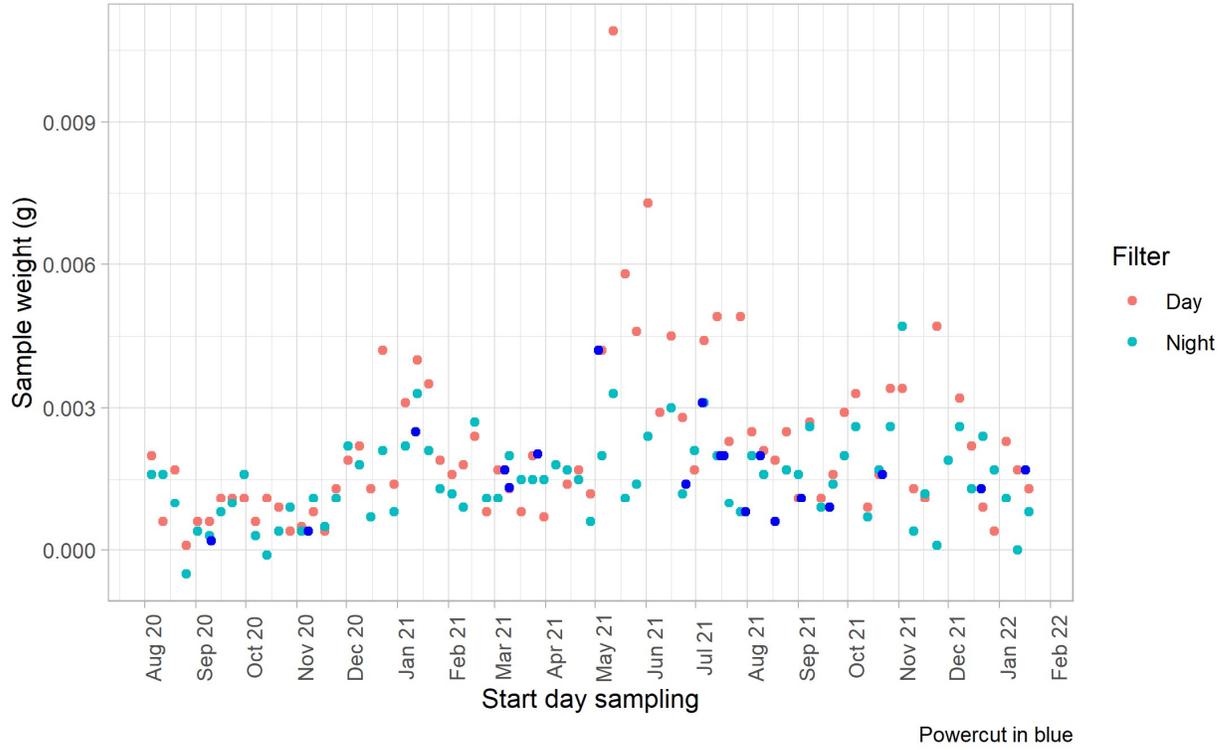
Appendix 10. HMBC and HSQC spectra with NMRfilter simulation

Spring-Dodecane



Spring- Tricarballic acid





Appendix 11. Weight of filter collected in Smear Station

Appendix 12. Node annotations . Note: Functional group reported before derivatization.

Node	Cluster	Name	Functional group	RT (min)	MQ Score	Balance. Score	Normalized Intensity (%)				
							AFilter2	Spring_1	Winter	Background	Summer
36	15	Dotriacontyl methyl ether	Alcohol	6.32	0.67	22	0.011234557				
77	5	1-Propylcyclohexanol	Alcohol	8.24	0.65	34					
94	5	Phenyl glycol	Alcohol	11.14	0.89	16	0.000580891	0.000841074		1.33E-06	
184	3	2,6-Heptadien-1-ol, 2,4-dimethyl-	Alcohol	15.02	0.82	35				0.047161737	0.154184232
370	63	Phenol, 2,6-bis(1,1-dimethylethyl)-4-ethyl-	Alcohol	22.44	0.79	81		0.001735753	0.100831653		0.00041249
433	10	4-Nonanol	Alcohol	25.5	0.81	64				0.035689598	0.126509192
573	40	1-Octyn-3-ol	Alcohol	31.27	0.79	32	0.00202646		0.10123436		
650	4	DL-3,4-Dihydroxymandelic acid	Alcohol	34.19	0.57	38		0.471446129		3.83E-06	
704	40	10-Methyl-1-dodecanol	Alcohol	36.04	0.72	20		4.66E-07	0.147834647	2.83E-08	
706	15	2,4-Di-tert-butylphenol	Alcohol	36.17	0.60	100		0.006216918	0.005783106		
876	40	10-Methyl-1-dodecanol	Alcohol	42.45	0.71	49	0.001981849	0.013827229	0.220169818		3.79E-05
960	78	3,5,5-Trimethylhexanol	Alcohol	46.04	0.69	36				0.002380547	0.00992744
1059	16	4-Hydroxy-5,8-dimethyl-5-propyl-1,7-nonadiene	Alcohol	49.53	0.82	39					
1144	43	Methyl triacontyl ether	Alcohol	52.99	0.94	100				1.61535328	3.322167345
1176	43	Methyl triacontyl ether	Alcohol	54.37	0.96	100				0.918080326	
14	24	Ethanone, 1,2-di-2-furanyl-2-hydroxy-	Aldehydes	4.8	0.75	34	0.206963878				
248	67	2,6-Heptadienal, 2,4-dimethyl-	Aldehydes	17.51	0.77	100				0.065969973	
504	63	3,5-di-tert-Butyl-4-hydroxybenzaldehyde	Aldehydes	28.16	0.72	100	0.003227204		0.131619003		
902	8	(Z)-11-Hexadecenal	Aldehydes	43.31	0.85	100				0.133894219	
33	2	Hexadecane	Alkanes	6.22	0.88	73				0.003646458	0.065376876
35	2	Hexadecane	Alkanes	6.32	0.89	78					0.232099789
63	2	Undecane	Alkanes	7.25	0.93	100				0.019042115	0.585650352
64	2	Undecane	Alkanes	7.35	0.93	77				0.13090986	
65	2	Dodecane	Alkanes	7.35	0.90	23				0.158600141	

146	2	Undecane	Alkanes	13.44	0.93	100					
147	2	Tridecane	Alkanes	13.44	0.93	100					0.052371997
170	2	Dodecane	Alkanes	14.68	0.94	27	3.34E-05	0.004850962			
186	2	Pentadecane	Alkanes	15.06	0.84	100		0.000808332			0.198479542
355	2	Undecane	Alkanes	21.74	0.91	74					0.00232951
356	2	Docosane	Alkanes	21.74	0.96	26				0.043374211	
448	2	Undecane	Alkanes	25.9	0.94	20				0.017360861	
454	2	Pentadecane	Alkanes	26.45	0.98	73	0.010100068	0.044451052	0.314411582		0.000258789
461	44	Hexacosane	Alkanes	26.59	0.85	74				0.002445413	0.02069524
474	2	Tetracosane	Alkanes	27.09	0.92	64		0.005885787	0.000986286		0.355015797
475	2	Undecane	Alkanes	27.09	0.93	36			0.000697049		
484	2	Octadecane	Alkanes	27.31	0.90	49	0.023834784				
485	2	Tridecane	Alkanes	27.37	0.89	56				0.015713573	
507	2	Nonacosane	Alkanes	28.31	0.91	100			0.073943535		0.167225759
601	83	Octadecane	Alkanes	32.2	0.71	41		0.122896384		0.006642406	0.04114613
604	2	Pentadecane	Alkanes	32.41	0.90	100	0.00862473	0.009114011			
611	37	2,2,4,4,6,8,8-Heptamethylnonane	Alkanes	32.63	0.79	22				0.034947805	
620	2	Pentadecane	Alkanes	32.84	0.94	75				0.087620683	
678	42	Cyclopentadecane	Alkanes	35.21	0.86	77				0.000101027	0.056676973
744	53	Undecane	Alkanes	37.4	0.89	43			0.074618813	0.046724748	
788	33	Pentacosane	Alkanes	39.44	0.65	30		0.29590897	0.010353825		
841	83	Hexacosane	Alkanes	41.02	0.66	22					0.01538283
851	33	Octadecane	Alkanes	41.5	0.58	51		0.891967001	0.001746314		
883	2	Tetracosane	Alkanes	42.68	0.94	100		0.004601624	2.016736595		
920	2	Pentacosane	Alkanes	43.85	0.97	82				0.060229866	0.888845316
947	2	Heptacosane	Alkanes	45.7	0.89	100			0.262797365		0.285414943
950	2	Tetradecane	Alkanes	45.75	0.93	30				0.000113091	
974	42	Cyclopentadecane	Alkanes	46.45	0.92	35				0.009713957	
996	42	Cyclopentadecane	Alkanes	47.47	0.77	100		0.011081511	0.000445868		0.288017254

997	2	Hexacosane	Alkanes	47.57	0.97	100		0.034076659	0.011473186		2.022441659
1002	14	Octadecane	Alkanes	47.76	0.77	100			0.309656645		
1045	2	2-methyl-nonadecane	Alkanes	49.29	0.93	100					0.385190838
1057	2	Docosane	Alkanes	49.43	0.93	11					0.000113978
1077	2	Octacosane, 1-iodo-	Alkanes	50.04	0.92	100	0.804249049	0.040626903	1.52400077		0.003484163
1156	2	Heptacosane	Alkanes	53.6	0.97	72					0.389243412
1192	2	Heptacosane	Alkanes	55.02	0.98	64					0.1828353
1212	23	2,2,4,4,6,8,8-Heptamethylnonane	Alkanes	55.58	0.62	78		0.000107694	0.138698069	2.30E-07	4.28E-05
1225	2	2-Methyl-nonadecane	Alkanes	55.87	0.98	100				0.007039668	0.126868829
1260	2	Pentadecane	Alkanes	56.85	0.95	100		0.001007391			0.074698883
1301	53	3-Methyloctane	Alkanes	57.95	0.81	100				0.0032	0.047764465
1308	48	n-Octadecane	Alkanes	58.27	0.84	100		0.003415664	0.000167311	8.39E-09	0.045055967
1335	2	Tetradecane	Alkanes	59.24	0.93	58				0.007434479	0.027733837
1336	23	2,2,4,4,6,8,8-Heptamethylnonane	Alkanes	59.24	0.65	42			0.033309753		0.000572923
1348	2	3-Methyloctane	Alkanes	59.83	0.96	100		0.005642777	0.018239802		
1353	21	Nonane, 2,2,4,4,6,8,8-heptamethyl-	Alkanes	59.98	0.75	100			0.019383425	0.000159366	
1354	2	3-Methyloctane	Alkanes	60.12	0.84	37			0.004997238	0.000267412	
1392	2	Pentadecane	Alkanes	61.2	0.87	78			0.003276979		0.026628974
1	5	Ethylbenzene	Alkenes	3.51	0.90	100	0.004225189				
3	5	Ethylbenzene	Alkenes	3.86	0.86	72	1.81E-12				
108	5	Benzene, tert-butyl-	Alkenes	11.79	0.77	54					0.002858374
240	56	1,19-Eicosadiene	Alkenes	17.21	0.62	100	0.265671032	0.003522003			0.392072969
764	73	1-Nonadecene	Alkenes	38.41	0.89	100		0.001688988	0.161462049		
942	20	(Z)-9-Heptacosene	Alkenes	45.6	0.77	100		0.510441276	0.216420312		
1093	24	1-Octadecene	Alkenes	50.77	0.67	63			0.097125899		3.93E-05
1103	43	(Z)-9-Heptacosene	Alkenes	51.04	0.96	100			0.026170302	0.004826685	3.444343065
1194	57	17-Pentatriacontene	Alkenes	55.06	0.65	68			0.133143653		
1363	44	cis-2-Heptene	Alkenes	60.25	0.69	39			0.017690563		0.000666693
116	1	Methyl pelargonate	Carboxylic acids	11.97	0.92	54				3.004640076	

117	1	Methyl pelargonate	Carboxylic acids	12.09	0.97	50		0.002932451		7.52E-05	
120	1	Methyl pelargonate	Carboxylic acids	12.13	0.98	50					0.000160956
127	35	Hexanedioic acid dimethyl ester	Carboxylic acids	12.45	0.97	41				0.222276434	
128	35	Hexanedioic acid dimethyl ester	Carboxylic acids	12.48	0.79	41					0.176497979
142	36	Hexanoic acid, 2-ethyl-, methyl ester	Carboxylic acids	13.18	0.61	69				0.051381872	
181	17	Dimethyl ethyl(2-oxopropyl)malonate	Carboxylic acids	14.96	0.59	46		0.081050652			
208	38	Dimethyl dl-malate	Carboxylic acids	15.71	0.83	100				0.05349848	
216	7	Pinonic acid	Carboxylic acids	16.01	0.86	85	0.40788357				0.000186943
244	29	Dimethyl glutarate	Carboxylic acids	17.4	0.66	100	0.014383938				0.198359258
251	67	3-(2-(methoxy(aryl)ethyl)-2-cyclopenten-1-one	Carboxylic acids	17.58	0.61	100				0.000650877	0.417583836
273	29	Methyl 3-methylheptane-1,7-dioate	Carboxylic acids	18.47	0.67	38					0.375861142
318	7	Pinonic acid	Carboxylic acids	20.5	0.65	100	47.68314591	0.007889845			
321	20	Dimethyl azelate	Carboxylic acids	20.6	0.70	57				0.029262052	
326	1	Methyl dodecanoate	Carboxylic acids	20.74	0.96	100				0.406358174	
327	7	Pinonic acid	Carboxylic acids	20.82		100	27.29647329		0.000109005		
330	7	Pinonic acid	Carboxylic acids	20.93		80	0.324841802	0.019694055			
349	20	Hexanedioic acid dimethyl ester	Carboxylic acids	21.57	0.61	34		0.000474802			
394	1	Hexadecanoic acid, 2-methyl-	Carboxylic acids	23.61	0.92	62				0.127715684	
409	11	Decanedioic acid, dimethyl ester	Carboxylic acids	24.32	0.82	100				0.284931702	
432	1	Methyl tetradecanoate	Carboxylic acids	25.36	0.90	50				0.023910289	
453	1	Methyl tetradecanoate	Carboxylic acids	26.37	0.94	100				0.908360835	
455	1	Methyl decanoate	Carboxylic acids	26.45	0.94	27		0.010610257			
457	1	Methyl heptadecanoate	Carboxylic acids	26.5	0.92	32					
502	1	10-Methylundecanoic acid methyl ester	Carboxylic acids	28.03	0.83	42				0.03696984	
505	1	Methyl tetradecanoate	Carboxylic acids	28.23	0.84	55				0.056739343	
520	1	Pentadecanoic acid, methyl ester	Carboxylic acids	29	0.97	100				0.558348956	

552	1	Methyl Nonadecanoate	Carboxylic acids	30.58	0.95	56				0.094033139	
571	11	Decanedioic acid, dimethyl ester	Carboxylic acids	31.23	0.69	100				3.421967148	
583	65	Hexadecanoic acid, 2-methyl-	Carboxylic acids	31.57	0.76	22			0.001123264		
612	8	11-Eicosenoic acid, methyl ester	Carboxylic acids	32.67	0.76	71	0.001329046	0.000322489		3.21E-05	0.291134683
613	1	Methyl palmitate	Carboxylic acids	32.67	0.84	29				0.12006646	
627	1	Methyl heptadecanoate	Carboxylic acids	33.2	0.90	71				0.184800914	
644	49	Hexadecanoic acid	Carboxylic acids	33.84	0.86	100	0.502475552	6.601657187			0.009989493
671	8	(Z,Z,Z)-6,12,15-Octadecatrienoic acid methyl ester	Carboxylic acids	34.97	0.78	33				0.041298083	
690	8	(Z,Z,Z)-6,9,15-Octadecatrienoic acid methyl ester	Carboxylic acids	35.42	0.90	100				0.685071108	
695	8	(z)-6-Octadecenoic acid,methyl ester	Carboxylic acids	35.58	0.84	100				0.92812541	
702	1	Methyl laurate	Carboxylic acids	36.04	0.85	35				0.201848516	
709	1	Methyl stearate	Carboxylic acids	36.39	0.94	100					29.35720846
760	49	Stearic acid	Carboxylic acids	38.3	0.84	100	0.028202619	2.649063426		6.46E-06	
761	1	Nonadecanoic acid methyl ester	Carboxylic acids	38.35	0.88	63				0.087182128	
794	8	Methyl palmitate	Carboxylic acids	39.71	0.65	43				0.13191193	
820	1	Methyl Nonadecanoate	Carboxylic acids	40.47	0.88	100				0.640242802	
897	8	(Z,Z,Z)-6,9,15-Octadecatrienoic acid methyl ester	Carboxylic acids	43.15	0.85	59				0.173382452	
912	3	Levulinic Acid	Carboxylic acids	43.71	0.88	100	1.02830253	0.04336917		0.004445959	
922	8	Methyl dodecanoate	Carboxylic acids	43.97	0.87	100				1.347303291	
926	65	Methyl tricosanoate	Carboxylic acids	44.52	0.76	48			0.000324838		
1011	1	Docosanoic acid, methyl ester	Carboxylic acids	48.14	0.86	100				0.708777677	
1115	1	Hexacosanoic acid, methyl ester	Carboxylic acids	51.56		100				0.452198808	
1145	1	Octacosanoic acid, methyl ester	Carboxylic acids	53.19		69				1.711889571	
1161	1	Docosanoic acid, methyl ester	Carboxylic acids	53.86	0.80	92				0.331434768	

1180	1	Triacontanoic acid, methyl ester	Carboxylic acids	54.59		100				2.321194269	
1240	1	Methyl Nonadecanoate	Carboxylic acids	56.25	0.82	100				0.778158586	
1313	1	Methyl Nonadecanoate	Carboxylic acids	58.44	0.83	100				0.249525771	
51	82	Benzyl acetate	Esters	6.81	0.80	81		4.59E-07			1.730418271
80	5	cis-3-Hexenol acetate	Esters	8.6	0.64	50			0.00225865		
98	5	Diphenylmethyl acetate	Esters	11.3	0.81	53				0.097686365	
237	3	Allyl decanoate	Esters	17.03	0.77	100				0.000729988	
535	41	3-Octyl hexanoate	Esters	29.63	0.50	34			0.126185491	0.002094377	
712	3	2,6,8-Trimethyl-4-nonyl acetate	Esters	36.46	0.89	13				1.26E-09	
734	11	Decanedioic acid, dibutyl ester	Esters	37.01	0.73	44				0.363602382	
737	11	Decanedioic acid, dibutyl ester	Esters	37.2	0.69	92	0.011441337	0.006417283	0.000715741		0.287961236
743	11	Decanedioic acid, dibutyl ester	Esters	37.33	0.79	100				16.33441674	
755	11	Decanedioic acid, dibutyl ester	Esters	38	0.79	100		0.006954259	0.246723001		
879	34	Diocetyl adipate	Esters	42.57	0.89	100			2.749917692	0.005542673	
893	14	Octadecanoic acid, 2-propenyl ester	Esters	43.05		88			0.937089884		
970	1	n-Dodecyl methacrylate	Esters	46.33	0.84	25				0.157925578	
1032	6	1,2-Benzenedicarboxylic anhydride	Esters	48.95	0.64	49				0.010575661	
1120	4	3,4-Dihydroxymandelic acid	Esters	51.76	0.56	25		0.270632847		0.000516078	
1249	4	Octadecyl hexadecanoate	Esters	56.55	0.76	35				0.022843787	
1267	4	Octadecyl hexadecanoate	Esters	57.01	0.72	86			0.106061311	0.000597676	
1277	27	Benzeneacetic acid, 4-chloro-, oct-3-en-2-yl ester	Esters	57.37	0.75	57			0.078776054		
1279	15	Glycerol tripropionate	Esters	57.37	0.74	10				0.010340483	
1417	4	Octadecyl hexadecanoate	Esters	62	0.82	100				0.033213045	0.381549083
40	86	Butyl decyl ether	Ether	6.59	0.88	53					0.330673777
104	5	Isovaleraldehyde dibenzyl acetal	Ether	11.65	0.83	36	0.097710814				
122	75	Ethyl dodecyl ether	Ether	12.17	0.56	100					0.301874132
472	87	Ditetradecyl ether	Ether	26.99	0.70	60	0.01102906	0.008575461	0.050935505		0.000342742

493	60	Hexadecyl octyl ether	Ether	27.68	0.80	49		0.021157829			0.09713556
503	34	Hexanal trans-2-hexenyl pentyl acetal	Ether	28.11	0.78	100	0.00107366	0.009275413	0.084806119		
724	10	1,3-Dioxolane, 2,2?-(1,3-propanediyl)bis-	Ether	36.88	0.78	32		0.149761498	1.12E-08		
771	73	Eicosyl propyl ether	Ether	38.8	0.83	100			0.174379803	0.001551244	0.010099147
870	13	1,3-Propanediol, dodecyl ethyl ether	Ether	42.31	0.79	100			0.510150461	7.47E-08	
887	60	Isobutyl tetradecyl ether	Ether	42.9	0.69	33				0.011944169	
967	13	1,3-Propanediol, ethyl triacontyl ether	Ether	46.16	0.79	100			0.516822596	0.019496555	
969	62	Octadecane, 1,1'-oxybis-	Ether	46.33	0.65	27	0.001195414		0.231340742		
1023	13	Ethyl octacosyl ether	Ether	48.62	0.76	52			0.143829799		
1051	10	Hexadecyl isopropyl ether	Ether	49.39	0.77	28					
1069	13	Ethyl octadecyl ether	Ether	49.71	0.84	52			0.255569361		
1081	28	Isobutyl tetradecyl ether	Ether	50.23	0.87	41				0.016740859	
1085	14	Isopropyl tetracosyl ether	Ether	50.3	0.77	100		0.004532847	1.070056519		
1136	4	Eicosyl isopropyl ether	Ether	52.27	0.86	100					0.615285807
1137	66	Isobutyl tetracosyl ether	Ether	52.4	0.82	100			0.733965735		
1155	13	1,3-Propanediol, ethyl tetracosyl ether	Ether	53.49	0.81	100			0.293979792		
1158	66	Isobutyl tetracosyl ether	Ether	53.77	0.86	100		0.098219808	0.384158899		
1178	59	Eicosyl octyl ether	Ether	54.48	0.80	80				0.067733604	
1223	69	Batilol	Ether	55.79	0.82	49				0.028975974	
1292	20	1-Methyl-1,2-epoxycyclohexane	Ether	57.69	0.66	51			0.018308538	0.000136531	
1328	4	Eicosyl isopropyl ether	Ether	58.83	0.64	73		0.004708935		0.000144222	0.133869645
30	24	3-Hepten-2-one	Ketones	5.82	0.80	43	0.05875781				
62	3	6-Methyl-5-octen-2-one	Ketones	7.19	0.82	100		0.00233632			0.484345723
88	7	Acetylcyclohexane	Ketones	10.4	0.89	100	0.383149568				
185	7	4-Cyclohexyl-4-methyl-2-pentanone	Ketones	15.06	0.53	35	2.375187078				
213	79	6-Methyl-3-(trimethylsilyl)methyl-1-hepten-4-one	Ketones	15.88	0.54	21	0.000917311				0.028814677
243	7	Geranyl acetone	Ketones	17.34	0.77	100	0.441945182	0.000673545			

449	3	Geranylacetone	Ketones	26.04	0.77	79	2.22E-05	0.010124282			
476	3	Ethanone, 1-(2,4,6-trihydroxyphenyl)-	Ketones	27.22	0.80	67		0.005948889			
829	52	9,10-Anthracenedione, 2-amino-3-hydroxy	Ketones	40.71	0.70	100				0.2369	0.487162018
915	19	3-(1-Acetyl-2,2-dimethyl-5-oxocyclopentyl)-propionic acid, methyl ester	Ketones	43.75		21		3.02E-13	6.07E-07		0.023613698
2	5	Benzaldehyde, 4-benzyloxy-3-methoxy-2-nitro-	Nitrocompounds	3.62	0.89	100	0.002121825				
7	5	O-Benzylhydroxylamine	Nitrocompounds	4.26	0.84	33					
11	5	O-Benzylhydroxylamine	Nitrocompounds	4.52	0.78	100					
13	5	Benzaldehyde, 4-benzyloxy-3-methoxy-2-nitro-	Nitrocompounds	4.8	0.83	66					
16	5	Benzaldehyde, 4-benzyloxy-3-methoxy-2-nitro-	Nitrocompounds	5.12	0.75	100	0.004332226	0.005781873	0.387739293		
49	86	Oct-3-enoyl amide, N-allyl-N-butyl-	Nitrocompounds	6.75	0.61	56				0.134064582	
100	5	Benzyl thiocyanate	Nitrocompounds	11.4	0.93	100	0.016708957				0.576017164
209	38	Valeramide, 2-methyl-N-hexyl-	Nitrocompounds	15.77	0.52	21			0.001522257	0.000231824	0.042423722
218	36	Oct-3-enoylamide, N-3-methylbutyl-	Nitrocompounds	16.05	0.69	87				0.165444821	
219	36	Oct-3-enoylamide, N-nonyl-	Nitrocompounds	16.1	0.75	81					0.060347628
311	77	Propanediamide, 2-ethyl-2-phenyl-	Nitrocompounds	20.3	0.60	40				0.058641782	
529	9	N-Ethyl-dodecanamide	Nitrocompounds	29.37	0.67	57				0.042185851	
656	9	N-Ethyl-dodecanamide	Nitrocompounds	34.33	0.76	15				0.080250111	
749	51	Decanamide-	Nitrocompounds	37.52	0.90	33		0.034799156		1.69E-05	
772	9	N-Ethyl-dodecanamide	Nitrocompounds	38.88	0.75	100				0.16693879	
773	9	N-Ethyl-dodecanamide	Nitrocompounds	38.93	0.75	53				0.001130719	0.081909699
856	51	9-Octadecenamamide, (Z)-	Nitrocompounds	41.74	0.72	100				0.001617151	
862	59	Tridecanenitrile	Nitrocompounds	42.07	0.79	86	0.008220835	0.012534207	0.000711583		0.255149335
992	10	Ethanediamide, N-(2-ethoxyphenyl)-N'-(2-ethylphenyl)-	Nitrocompounds	47.3		8		0.001212424			
1201	23	2-Pyrazoline, 1-isobutyl-3-methyl-	Nitrocompounds	55.18	0.70	36			0.06145995	0.000863145	
83	85	Benzene, 1-iodo-2-(trifluoromethyl)-	Organohalogens	9.15	0.63	83	0.132290543				

151	85	Benzaldehyde, 4-(trifluoromethyl)-	Organohalogens	13.57	0.60	69					0.646699189
162	16	1-Chloroeicosane	Organohalogens	14.19	0.95	74					0.432420003
165	16	1-Chloroeicosane	Organohalogens	14.39	0.92	42				0.170300839	
166	16	1-Chloroeicosane	Organohalogens	14.43	0.91	42					0.326352688
174	2	2-Bromo dodecane	Organohalogens	14.76	0.96	27				0.439314644	0.87633613
308	3	1-Chloroeicosane	Organohalogens	19.97	0.90	29			0.00173358		
319	24	3-Methyladipoyl chloride	Organohalogens	20.57	0.75	86	0.000778197				0.017377414
521	2	Octadecane, 1-iodo	Organohalogens	29.09	0.92	100			0.83601545		
648	2	Eicosane, 1-iodo-	Organohalogens	34.03	0.96	100			2.091191243		
687	3	1-Chloroeicosane	Organohalogens	35.32	0.89	15	0.090418652			3.46E-07	
770	2	Docosane, 1-iodo-	Organohalogens	38.55	0.99	100	0.003807539		2.64896		0.009665105
825	2	Hexacosane, 1-iodo-	Organohalogens	40.63	0.91	100	0.000212408	0.009079923	1.040563329		
848	21	Decane, 1-iodo-	Organohalogens	41.35	0.76	100			0.184795418	0.005366421	
931	2	Hexacosane, 1-iodo-	Organohalogens	44.62	0.97	100	0.000912881	0.023849964	1.330653752	3.04E-06	0.003180359
977	2	Hexacosane, 1-iodo-	Organohalogens	46.5	0.92	100			1.90890905	0.007743947	
978	70	Octacosane, 1-iodo-	Organohalogens	46.63	0.82	100			0.4342799	0.013158927	
982	14	Dotriacontane, 1-iodo-	Organohalogens	46.82	0.75	100			1.122141121		0.038449832
1015	2	Octacosane, 1-iodo	Organohalogens	48.3	0.92	100		0.067000803	1.587968238	0.00182849	
1079	70	Dotriacontane, 1-iodo-	Organohalogens	50.19	0.80	65	0.001163185	0.003482027	0.212577175		
1118	2	Triacontane, 1-iodo-	Organohalogens	51.7	0.92	100		0.032391217	1.042267695	1.07E-08	
1139	2	Octacosane, 1-iodo-	Organohalogens	52.63	0.98	100		0.058131002	0.724240183		
1147	2	Octacosane, 1-iodo-	Organohalogens	53.3	0.90	100			0.994995955		
1200	45	Hexadecane, 1-chloro-	Organohalogens	55.18	0.70	64				0.00807484	0.065615805
1224	72	Heptane, 3-(bromomethyl)-	Organohalogens	55.82	0.77	100		0.003242075	0.163813467	2.66E-08	
1311	57	1-Chloroeicosane	Organohalogens	58.35	0.77	85			0.028606094		0.00014968
832	76	Phosphoric acid, tris(2-ethylhexyl) ester	Organophosphate Esters	40.83	0.66	71			0.076416265	2.57E-05	
921	76	Phosphoric acid, tris(2-ethylhexyl) ester	Organophosphate Esters	43.85	0.98	18			0.26389397	0.003289571	
1207	84	Tris(2,4-di-tert-butylphenyl) phosphite	Organophosphate Esters	55.42		100					5.134191411

1243	25	Tris(2,4-di-tert-butylphenyl) phosphate	Organophosphate Esters	56.41		84		24.18768027		
1274	25	Tris(2,4-di-tert-butylphenyl) phosphate	Organophosphate Esters	57.19		100			0.148652117	3.277356756
1342	25	Tris(2,4-di-tert-butylphenyl) phosphate	Organophosphate Esters	59.5	0.66	65				0.011726014
17	32	Cyclotetrasiloxane, octamethyl-	Organosilicon compounds	5.18	0.94	78	2.40E-13	0.17579382		
20	32	Cyclotetrasiloxane, octamethyl-	Organosilicon compounds	5.33	0.92	32		0.00038271		
523	4	Cyclohexasiloxane, dodecamethyl-	Organosilicon compounds	29.2	0.90	79				0.055690944
602	4	2,4,6-Tris(trimethylsiloxy)benzoic acid trimethylsilyl ester	Organosilicon compounds	32.25	0.61	100		0.00505193	0.137781119	
665	4	2,4,6-Tris(trimethylsiloxy)benzoic acid trimethylsilyl ester	Organosilicon compounds	34.81	0.67	60				0.025611831
711	4	Cyclohexasiloxane, dodecamethyl-	Organosilicon compounds	36.46	0.82	42				0.084971669
1006	4	Cyclohexasiloxane, dodecamethyl-	Organosilicon compounds	47.94	0.76	100				0.039885562
1086	4	Cyclohexasiloxane, dodecamethyl-	Organosilicon compounds	50.35	0.67	100				0.081957956
1167	4	Cyclohexasiloxane, dodecamethyl-	Organosilicon compounds	54.02	0.65	64				0.143277261
1190	4	2,4,6-Tris(trimethylsiloxy)benzoic acid trimethylsilyl ester	Organosilicon compounds	54.96	0.62	61				0.145434625
1233	4	2-(Trimethylsilyl)oxy-eicosanoic acid trimethylsilyl ester	Organosilicon compounds	56.03	0.63	62				0.201820993
1329	4	2-(Trimethylsilyl)oxy-eicosanoic acid trimethylsilyl ester	Organosilicon compounds	58.83	0.56	27				0.050692292
1364	31	Tetrasiloxane, decamethyl-	Organosilicon compounds	60.25	0.82	11		0.000720148		
1378	31	Cyclopentasiloxane, decamethyl-	Organosilicon compounds	60.83	0.77	25		0.020917041		
1405	31	Cyclopentasiloxane, decamethyl-	Organosilicon compounds	61.69	0.57	26		0.000135505	0.000136094	
238	56	Decyl sulfide	Organosulfur compounds	17.15	0.61	100				0.572536116
488	12	N,N-Dichlorobenzenesulfonamide	Organosulfur compounds	27.44	0.83	100	0.00091542	0.000421908	0.000241193	0.3125482

492	46	Benzenesulfonyl isocyanate	Organosulfur compounds	27.58	0.64	100				10.53706004	
500	12	Benzenesulfonamide, N-butyl-	Organosulfur compounds	27.97	0.91	100					0.193421847
509	12	Benzenesulfonamide, N-butyl-	Organosulfur compounds	28.53	0.92	100			0.000160196		6.467960207
532	12	Benzenesulfonamide, N-butyl-	Organosulfur compounds	29.44	0.92	100	1.443265382		18.35276452		0.000130218
533	12	Benzenesulfonamide, N-butyl-	Organosulfur compounds	29.52	0.91	100	0.000891013		0.264536918		
828	12	N-Fluorobenzenesulfonimide	Organosulfur compounds	40.67	0.82	10	0.186875748			0.000805773	
1038	28	tert-Nonyl mercaptan	Organosulfur compounds	49.06	0.96	41				0.016405095	
1169	14	Disulfide, di-tert-dodecyl	Organosulfur compounds	54.06	0.75	100			0.66152752	0.003854912	
1320	69	tert-Hexadecanethiol	Organosulfur compounds	58.56	0.88	75	7.72E-05				0.017562102
160	7	2,5,5-Trimethyl-3-hexyn-2-ol	Others	14.06	0.83	68	4.46100049		0.046439614	0.025864031	0.069775581
227	7	1-Heptyn-3-ol, 3-ethyl-5-methyl-	Others	16.61	0.73	100				1.401003916	
260	39	2,5,5-Trimethyl-3-hexyn-2-ol	Others	17.89	0.75	85	0.000186047	0.004211384			0.09719248
268	17	7-Tetradecyne	Others	18.28	0.75	100	4.95E-05		0.00405814		
385	30	1,2,3,4-Tetra-O-acetyl-beta-d-glucopyranuronamide	Others	22.94	0.51	86			0.206340188		
534	27	1-Nonyne	Others	29.63	0.81	66				0.003038748	0.183192977
930	55	Hexaethylene glycol monododecyl ether	Others	44.56	0.58	26	0.000712188				
1299	26	2,5,5-Trimethyl-3-hexyn-2-ol	Others	57.84		13				0.011426646	0.12013899
1302	26	2,5,5-Trimethyl-3-hexyn-2-ol	Others	58.07	0.79	100		6.26E-05			0.080510828
696	74	Fluoranthene	PAHs	35.78	0.79	100		0.087841996	0.63017832	0.001191119	0.008244043
733	74	Pyrene	PAHs	37.01	0.95	56	0.000235767	0.002206903	0.644323534		0.00089774
866	71	7H-Benz[de]anthracen-7-one	PAHs	42.13		53	0.000204927	6.41E-05	0.061144141		
936	71	7H-Benz[de]anthracen-7-one	PAHs	45.07	0.72	100	4.98E-05	0.009718393	0.147934338	2.87E-07	0.000240275
381	6	Diethyl Phthalate	Phthalate	22.79	0.81	36				0.042331574	
400	6	Diethyl phthalate	Phthalate	23.74	0.85	100			0.113562016		2.38E-05
543	6	Diisobutyl phthalate	Phthalate	30.06	0.92	100				0.51918562	
560	6	Diisobutyl phthalate	Phthalate	30.83	0.94	100			1.920371589	9.69E-06	0.009700932

564	6	Di-n-octyl phthalate	Phthalate	30.93	0.73	24				8.99E-05	
599	6	Di-n-octyl phthalate	Phthalate	32.16	0.82	100				0.349339126	
606	6	Diisobutyl phthalate	Phthalate	32.52	0.86	100				8.903890659	
629	6	Dibutyl phthalate	Phthalate	33.29	0.91	100	0.049921398	1.70E-06	20.59344941		
776	6	Octyl decyl phthalate	Phthalate	39.07	0.84	42				0.024296188	
924	6	Bis(2-ethylhexyl) phthalate	Phthalate	44.35	0.60	81	0.000493538		6.40E-05		
940	6	Bis(2-ethylhexyl) phthalate	Phthalate	45.43	0.89	100			10.92950562		0.133121849
1016	6	Bis(2-ethylhexyl) phthalate	Phthalate	48.36	0.73	100				0.369942156	
1017	6	Di(2-ethyl hexyl)isophthalate	Phthalate	48.44	0.59	78				0.205973756	3.371944078
1018	6	Bis(2-ethylhexyl) phthalate	Phthalate	48.44	0.80	22				0.087936307	3.874212391
1044	6	Bis(2-ethylhexyl) phthalate	Phthalate	49.14	0.75	100		0.020732608	0.872891088	1.84E-05	
6	5	Terpinolene	Terpenoids	4.26	0.88	67	0.000255053				
10	5	Alfapinene	Terpenoids	4.43	0.94	100	0.013791008				
12	5	Camphene	Terpenoids	4.73	0.94	100	0.032065909				
69	3	11-Epi-sinulariolide	Terpenoids	7.61	0.72	100		0.000689112			
75	5	trans-Linalool oxide	Terpenoids	8.19	0.72	50					
78	5	trans-Linalool oxide	Terpenoids	8.52	0.81	100					
85	5	α -Sinensal	Terpenoids	9.47	0.75	100					
86	5	α -Sinensal	Terpenoids	9.62	0.82	100					
102	5	Verbenone	Terpenoids	11.56	0.82	18					0.00011909
121	75	Geranyl acetone	Terpenoids	12.13	0.73	100				0.054	
132	7	(1R,2R,3S,5R)-(-)-2,3-Pinane-1,2-diol	Terpenoids	12.67		62	5.539589621	0.005820521			0.344757945
153	7	(1R,2R,3S,5R)-(-)-2,3-Pinane-1,2-diol	Terpenoids	13.72		20				7.00E-04	
191	3	(5E,9E,13E)-6-Acetyl-9,13-dimethyl-3-methylene-3,3a,4,7,8,11,14a-octahydro-2H-cyclotrideca(b)furan-2-one	Terpenoids	15.16	0.73	44		0.006306258			
197	3	Ethyl geranyl acetone	Terpenoids	15.44	0.85	59					
234	7	Isocaulalol	Terpenoids	16.87	0.77	100				0.335342587	
259	7	Geranyl acetone	Terpenoids	17.85	0.78	100	1.018965319			0.0109	

270	7	Geranyl acetone	Terpenoids	18.38	0.78	73	0.395492397				
284	3	Citronellal	Terpenoids	19.13	0.78	42	3.23E-07	0.006491176			
332	50	2,4,6-Triisopropylphenetole	Terpenoids	20.98	0.63	67			0.171342056		
357	39	α -Ionol	Terpenoids	21.9	0.80	100					0.693415841
406	39	Linalool	Terpenoids	24.19	0.70	100				0.003334833	0.110951549
427	3	Widdrol hydroxyether	Terpenoids	25.21	0.71	75	0.005149477	3.86E-06	2.25E-06		
489	27	D-Limonene 1,2-epoxide	Terpenoids	27.52	0.66	65		0.020977799		0	0.06840185
624	4	3-Octanol, 3,7-dimethyl-	Terpenoids	32.98	0.71	100				0.1877	
723	26	3,3,5-Trimethylcyclohexyl isobutyrate	Terpenoids	36.88	0.71	32	0.31129257		7.69E-06		
783	26	5-Hydroxy-1-methylbicyclo(4.3.0)nonane-2,7-dione	Terpenoids	39.22	0.67	25	0.119318881				0.000285741
791	9	1,2-Epoxy-3,7,11-trimethyldodecane-3-ol	Terpenoids	39.54	0.69	13	0.008336097				
803	21	3,7,11,15-Tetramethyl-1-hexadecen-3-ol	Terpenoids	39.87	0.79	26			0.002818017	8.00E-04	0.155765749
804	7	Menthyl butyrate	Terpenoids	39.95	0.75	100	2.65408613				
807	7	L-Menthyl 3-oxohexanoate	Terpenoids	40.07	0.75	51	0.628954627				
850	7	L-Menthyl 3-oxohexanoate	Terpenoids	41.4	0.69	36	0.123082255	0.037934359			
853	52	Dehydroabiatic acid	Terpenoids	41.58	0.63	100			0.269041505	0.002313715	
937	62	7-Oxodehydroabiatic acid, methyl ester	Terpenoids	45.29		54		1.184162266		0.0783	
1210	27	6,10-Dimethyldodeca-5,9-diene-2-one	Terpenoids	55.53	0.70	52			0.167887256		
1211	17	Citronellal	Terpenoids	55.53	0.64	48			0.000255711		
1253	17	cis-Geraniol	Terpenoids	56.68	0.70	100		1.43944668	0.014787359	0.004843222	
1276	17	cis-Geraniol	Terpenoids	57.28	0.56	34		4.26E-05			
27	61		Unknown	5.51		100	8.70E-08	0.068435879	0.739740733		
28	61		Unknown	5.7		100			0.356847098		
39	3		Unknown	6.55		100					
44	19		Unknown	6.68		40				0.054293535	
45	37		Unknown	6.68		37		0.002217221			
47	81		Unknown	6.72		56					0.28302191

54	81		Unknown	6.87		37				0.085580637	
55	80		Unknown	6.87		21					0.046019949
57	80		Unknown	7.02		53		8.12E-05	2.89E-05		
135	38		Unknown	12.81		100				0.112557729	
140	3		Unknown	13.12		39	0.000111484	0.003084355			
157	44		Unknown	13.93		68	0.043130913				
159	3		Unknown	13.97		68					
198	46		Unknown	15.44		100	0.143144648				
206	3		Unknown	15.65		100	0.001003837				
232	35		Unknown	16.8		76	0.004576997				0.060509496
274	30		Unknown	18.66		40			0.109044187	7.85E-05	
283	30		Unknown	19.02		40			0.109771222		
340	3		Unknown	21.3		100	2.99E-05	6.97E-05	3.80E-09		
345	19		Unknown	21.54		29		0.000377849		5.73E-08	
364	3		Unknown	22.16		49		0.000983238			
387	29		Unknown	22.99		100				0.000435475	0.19200224
477	87		Unknown	27.22		33	0.003278347		0.058793861		
508	41		Unknown	28.42		100				0.289852838	
513	41		Unknown	28.76		41				0.064537588	
550	55		Unknown	30.43		49		0.000927173			
572	6		Unknown	31.27		68				0.120869706	
581	15		Unknown	31.52		100		0.046211743			
623	68		Unknown	32.9		16		0.252521293			
646	1		Unknown	33.93		29			1.95E-06		
655	54		Unknown	34.33		85		0.002163014	0.041918566		0.501215811
673	33		Unknown	35.05		16		0.042924243	0.000617838		
707	64		Unknown	36.22		100				0.034221001	0.073646972
747	54		Unknown	37.43		100		0.247705723		0.014285332	0.629795394
809	4		Unknown	40.07		16		0.13681961			
833	68		Unknown	40.83		29		0.156460663		2.79E-05	

894	9		Unknown	43.05		12				0.052431571	
1000	4		Unknown	47.66		25				0.001744629	
1027	64		Unknown	48.67		19	8.67E-06	0.003220873	0.001053947		
1028	6		Unknown	48.73		100				0.002085296	
1053	6		Unknown	49.39		16		0.074321952	0.005963619		
1070	25		Unknown	49.71		48	1.462768196	0.020113083			
1112	1		Unknown	51.45		43		0.240916977	2.21E-07		
1170	45		Unknown	54.1		42					0.076190147
1184	84		Unknown	54.8		100				0.000227736	
1197	58		Unknown	55.12		64		0.512814942		8.72E-05	
1209	72		Unknown	55.49		45				0.332554506	
1232	4		Unknown	56		20		0.153043817	0.000218192		1.40E-05
1278	58		Unknown	57.37		33		0.200967154			
1322	23		Unknown	58.64		100		2.17E-06	0.023788824		
1374	4		Unknown	60.78		56				0.000286878	0.024117339
1375	4		Unknown	60.78		44				0.018227144	

Non-exclusive licence to reproduce the thesis and make the thesis public

1. I, Nieves Maria Flores March, grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis

Organic Constituents of Atmospheric Aerosols in a Hemi-boreal Forest

Supervised by Professor Heikki Junninen.

2. I grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.
3. I am aware of the fact that the author retains the rights specified in points 1 and 2.
4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Nieves Maria Flores March
22/03/2022

Information Sheet

Organic Constituents of Atmospheric Aerosols in a Hemi-boreal Forest

Atmospheric aerosols have been demonstrated to be a highly dynamic system, playing a significant role in climate change and human health. In nature, ecosystems like boreal forests can modify the atmospheric particles producing a warming or cooling effect on climate. However, the regional and global impact of boreal forest on climate is still difficult to determine, especially due to the heterogeneous chemistry of aerosol samples, the need for multiple instruments for identification, and their limited library of compounds. In this thesis, to overcome these issues, we used a molecular networking technique based on the Global Natural Products Social web platform in combination with Nuclear Magnetic Resonance (NMR) to perform a screening of organic aerosols during the winter spring, and summer seasons from a Hemi-boreal forest. The aerosol samples were recollected in a glass filter weekly from SMEAR Station (Estonia) and analyzed by Gas Chromatography Mass spectrometry and NMR. A variety of chemical functional groups including carboxylic acids, phthalates, and organophosphate among the most abundant were annotated in the studied seasons. Furthermore, it was analyzed the presence of *n*-alkanol, carboxylic acid, and *n*-alkane to evaluate any hydrocarbon contamination. Phthalates-based compounds like Dibutyl phthalate (~20.59% in winter), and Bis(2-ethylhexyl) phthalate (~3.87% in summer), altogether with organophosphates like Tris(2,4-di-tert-butylphenyl) phosphate (~24.13% in spring) and tris(2,4-di-tertbutylphenyl) phosphite (~5.13% in summer) were annotated as a possible air pollutant. Besides that, conifer burning tracers such as 7-Oxodehydroabietic (~1.18% in spring) and dehydroabietic acid (~0.49% in summer) were annotated. These finding presented in this work gives an insightful impact on the atmospheric aerosol composition presented in a Hemi-boreal forest using a straightforward and versatile technique such as molecular networking.

Keywords: Organic aerosols, GC/MS/EI, GNPS, NMR, Molecular networking, NMRfilter

CERCS: P305 Environmental chemistry

INFOLEHT

Hemiboreaalsete metsade atmosfääri aerosooli orgaaniline koostis

Atmosfääri aerosoolid on väga mitmetahuline süsteem, millel on oluline roll kliimamuutustes ja inimeste tervises. Atmosfääris on mitmeid protsesse, mis mõjutavad uute osakeste teket ja osakesed on omakorda pilvetekkeks vajalikud kondensatsioonituumad. Muutused atmosfääri keemilises koostises, mis mõjutab uute osakeste teket, on põhjustatud atmosfääri heidetest, looduslikest või inimtekkelistest. Boreaalsete metsade piirkondlikku ja globaalset mõju kliimale on aga endiselt raske hinnata, eriti aerosoolide heterogeense keemia, analüütiliste meetodite keerukuse ja piiratud ühendite andmebaaside tõttu. Nende probleemide lahendamiseks kasutasime selles lõputöös aerosooliproovide analüüsimiseks gaasikromatograafiat koos massispektromeetriga ja tuumamagnetresonants spektromeetrit. Saadud andmed analüüsi Global Natural Products Social Network veebiplatvormil põhinevat multidimensionaalset statistilist analüüsi meetodit kasutades. Aerosooliproove koguti talvel, kevadel ja suvel kord nädalas SMEAR Estonia mõõtejaamas (Järvselja, Eesti) klaasfiiberfiltrile. Uuritud proovidest tuvastati mitmesuguseid keemilisi funktsionaalseid rühmi, sealhulgas karboksüülhappeid, ftalaadid ja orgaanilised fosfaadid. Lisaks analüüsi süsivesinike saastumise hindamiseks *n*-alkanooli, karboksüülhappe ja *n*-alkaani olemasolu. Ftalaatidel põhinevad ühendid nagu dibutüülftalaat (talvel ~20,59%) ja bis(2-etiülheksüül)ftalaat (suvel ~3,87%), koos organofosfaatidega nagu Tris(2,4-di-tert-butüülfenüül)fosfaat (~24.13% kevadel) ja tris(2,4-di-tertbutüülfenüül)fosfit (suvel ~5,13%) märgiti võimalikuks õhusaasteaineks. Lisaks sellele tuvastati biomassi põlemis

markerite, nagu 7-oksodehüdrobieet (~1,18% kevadel) ja dehüdrobieethape (suvel ~0,49%) olemasolu. Selles töös esitatud tulemused annavad põhjaliku ülevaate orgaanilise aerosooli koostisest hamiboreaalse metsa atmosfääris, samas ka tõestavad ka kasutatud meetodika võimekust keemiliselt heterogeense keskkonna proovi analüüsimisel.

Märksõnad: orgaanilised aerosoolid, GC/MS/EI, GNPS, NMR, molekulaarvõrgud, NMR-filter
CERCS: P305 Keskkonnakeemia