Changing impacts of plant emissions on urban air quality in a changing climate

Barbora Bracinik

MSc by Research

University of York

Department of Chemistry

November 2025

Abstract

Particulate matter acts as a significant pollutant affecting climate and health, particularly in urban environments such as Beijing which experiences consistent poor air quality. The most dangerous pollutant to health in China is fine particulate matter ($PM_{2.5}$), and substantial reductions are required to reduce pollution-related mortality. A large fraction of the sub-micron particulate matter (PM_1) mass consists of secondary organic aerosols (SOA), which form from the oxidation of volatile organic compounds (VOCs) emitted from biogenic or anthropogenic sources.

In this study, non-target analysis was used to develop SOA tracer libraries containing the 100 most abundant compounds for 6 monoterpenes, 2 monoterpenols and 2 sesquiterpenes via photo-oxidation chamber experiments. $PM_{2.5}$ filter samples were extracted and analysed using Ultra-High-Performance Liquid Chromatography with tandem Mass Spectrometry (UHPLC-MS2), and imported into the open source MZmine 4.3.0 software for library generation. To investigate ambient $PM_{2.5}$, filter samples were collected from 11th-25th June 2017 at the Institute of Atmospheric Physics (IAP) in Beijing, China, and consequently imported into MZmine. A data processing workflow was set up using the 'Spectral Library Search' function on MZmine, allowing matching of the libraries to Beijing samples, leading to the identification of 311 biogenic SOA tracers.

For monoterpenoids, fragmentation processes generating species with less than 10 carbons dominated oxidation products for all libraries except linalool, which produced more functionalised C_{11}-C_{20} products. Sesquiterpene libraries were also dominated by fragmentation products, yielding mostly C_1-C_{15} compounds. Tracer concentrations were generally stable apart from a pronounced peak on 20th June, driven by *p*-cymene SOA, linked to southeasterly transport of biogenic emissions. Diurnal patterns showed morning maximums and night-time minimums, while correlation analysis indicated that atmospheric processes, including oxidant levels and NO_x regime, were significant in controlling SOA formation. These findings demonstrate the influence of regional transport and atmospheric complexity in urban biogenic SOA diversity.

Table of Contents

Title Page 1
Abstract 2
Table of Contents 3
List of Tables 5
List of Figures 7
Acknowledgements 11
Author's Declaration 12
1. Introduction 13
1.1. Introduction to Aerosols 13
1.2. Types of Aerosols 14
1.3. Secondary Organic Aerosol formation 15
1.4. Global contributions of SOA 16
1.5. Contributions of SOA to total PM masses 16
1.6. Impact of atmospheric aerosols: climate 18
1.7. Impact of atmospheric aerosols: health 21
1.8. BVOCs and BSOA 23
1.9. Global BVOC emissions 24
1.10. Known BVOC oxidation mechanisms 26
1.11. SOA pollution in Beijing 29
1.12. BVOC sources in Beijing 29
1.13. Similar studies 30
1.14. Project aims 31
2. Methodology 32
2.1. Chamber experiments 32
2.2. Beijing Filter Collection 38
2.3. Filter Extraction 39
2.4. UHPLC-MS2 40
2.5. Auxiliary measurements in Beijing 40
2.6. Data analysis 41
2.7. Generation of BSOA mass spectral libraries-workflow 42
2.8. Biogenic SOA tracer compounds in Beijing 45
3. Results and Discussion 46
3.1.1. Composition of chamber SOA 46

3.1.2. Monoterpenes 47
3.1.3. Monoterpenols 56
3.1.4. Impact of structure on composition 59
3.1.5. Sesquiterpenes 59
3.2. Spectral library search results 63
3.2.1. Overview of BSOA tracers identified in Beijing 63
3.2.2. C_1-C_{10} tracers: Monoterpenoid SOA 64
3.2.3. C_{11}-C_{20} tracers: Monoterpenoid SOA 64
3.2.4. C_{20+} tracers: Monoterpenoid SOA 65
3.2.5. Sesquiterpene SOA tracers 65
3.2.6. Limitations of spectral library search 66
3.3. Tracer time series 68
3.4. Precursor contributions 69
3.4.1. Trends in BSOA tracer contributions 69
3.4.2. p-Cymene BSOA in Beijing 72
3.5. Diurnal trends 72
3.5.1. Diurnal variability of total BSOA tracers 72
3.5.2. Individual diurnal plots 75
3.6. Causes of Diurnal Variations 77
3.7. Time series trends 82
3.8. Temperature 85
3.9. Humidity 88
3.10. Spatial variability of BSOA tracers 89
3.11. Comparison of BSOA trends to other pollutants 92
3.12. Contrasting individual polar plots 99
3.13. Trends between tracers 100
3.13.1. Filtering process 100
3.13.2. Corr plot 101
3.13.3. Causes of Corr plot trends 103
3.14. Conclusion 105
References 107

List of Tables

Table 1: Names and structures of 10 biogenic SOA precursors used in the chamber experiments, including 6 monoterpenes, 2 monoterpenols and 2 sesquiterpenes 33
Table 2: Start and end times, along with the durations of each experiment. The end time for the β-caryophyllene experiment was not recorded, therefore, its total duration is unknown 36
Table 3: Initial and final filter masses (mg), as well as the mass differences, which represent the mass of SOA deposited onto the filters. Masses were weighed three times, and averages were taken. The initial filter mass for the terpinolene experiment was not recorded, therefore, the mass difference is unknown 37
Table 4: MS1 noise levels for each individual biogenic library and the corresponding number of tracers retained 44
Table 5: The 20 most abundant SOA tracers in the terpinolene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 47
Table 6: The 20 most abundant SOA tracers in the sabinene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 48
Table 7: The 20 most abundant SOA tracers in the α-pinene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 49
Table 8: The 20 most abundant SOA tracers in the p-cymene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 50
Table 9: The 20 most abundant SOA tracers in the ocimene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 51
Table 10: The 20 most abundant SOA tracers in the 3-carene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 52
Table 11. SOA tracers detected in three or more SOA tracer tables. The table lists molecular formulas, the number of tables in which each tracer appears, and the tables where the tracer is observed 55
Table 12: The 20 most abundant SOA tracers in the linalool library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 56
Table 13: The 20 most abundant SOA tracers in the α-terpineol library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 57
Table 14: The 20 most abundant SOA tracers in the β-caryophyllene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 60
Table 15: The 20 most abundant SOA tracers in the α-cedrene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided 61
Table 16: The number of SOA tracer compounds matched to each monoterpenoid precursor library, along with the number of tracers containing 1-10 C atoms, 11-20 C atoms and 20+ C atoms 63

Table 17: The number of SOA tracer compounds matched to each sesquiterpene precursor library, along with the number of tracers containing 1-15 C atoms, 16-30 C atoms and 30+ C atoms 65
Table 18: The number of SOA tracer compounds matched to the Beijing samples and the number of those appearing among the 20 most abundant compounds, along with their corresponding molecular formulas 66

List of Figures

Figure 1: Illustration of different modes in a typical atmospheric particle size distribution. The main sources and processes influencing each mode are shown, along with their classification as ultrafine particles, $PM_{2.5}$, PM_{10}, or total suspended particles (TSP).....13

Figure 2: Examples of sources of volatile organic compounds (VOCs) in the atmosphere, both anthropogenic and biogenic, and the processes they undergo throughout their life cycles in order to become aerosols......14

Figure 3. Total mass concentrations (μg m^{-3}) and mass fractions of inorganic species and organic components in submicrometer (PM_1) aerosols at multiple surface sites across the Northern Hemisphere.....17

Figure 4: Mechanism of Mie scattering on smaller and larger particles, showing the strong favouritism towards re-emission in the direction of the incoming light source.....18

Figure 5: The indirect effects of aerosols on cloud formation, illustrated by differences between polluted environments with high aerosol concentrations and clean environments with low aerosol concentrations, both having the same water vapour concentration.....21

Figure 6: Geographic distribution of cardiovascular and respiratory disease rates across the USA (left) compared with concentrations of $PM_{2.5}$ across the USA (right).....23

Figure 7: Relative contribution of monoterpene species to monoterpene annual global total averaged over the 1980-2010 period. The considered monoterpene species are listed to the right, and the 'other monoterpenes' fraction has been excluded.....24

Figure 8: Spatial distribution of monthly mean emissions of (a) isoprene, (b) monoterpenes, (mg m^{-2} day^{-1}) for January (top) and July (bottom) averaged over the modelled period 1980-2010.....25

Figure 9: Classification of precursor compounds by saturation concentration (C^0) and O:C ratios. Categories range from VOCs to ELVOCs. Arrows indicate chemical processes that contribute to SOA formation. (n) represents the precursor VOC, therefore (n+1) represents the 1st generation oxidation product......26

Figure 10: The proposed reaction scheme showing the initial O_3 addition to terpinolene. The diagram illustrates fragmentation occurring due to the fact that O_3 attacks the exocyclic double bond.....27

Figure 11: The ozonolysis of endocyclic α-pinene, showing O_3 attacking the double bond within the carbon ring. This leads to the production of Criegee intermediates, which lead to formation of products in which the carbon backbone is retained.....27

Figure 12: Simplified proposed mechanism for α-pinene oligomer formation, leading to the production of the MW 368 ester.....28

Figure 13: The 0.3 m^3 continuous-flow reactor (CFR) setup, based in York.....35

Figure 14: Detailed schematic of the continuous-flow reactor (CFR), showing connections into the chamber, the chamber itself and connections coming out of the chamber. MFC is a mass flow controller, and PTR-MS is a proton transfer reaction mass spectrometer 35
Figure 15: A map of Beijing, highlighting the location of the Institute of Atmospheric Physics (IAP), along with a zoomed-out view showing its position within East China 38
Figure 16: MZmine workflows for BSOA library generation (blue) and the spectral library search (green), with key parameters indicated for each step 42
Figure 17: A plot of peak area against the cumulative number of tracers detected in the linalool library, showing that most of the total peak area was from the 100 most abundant tracers. 44
Figure 18: Flowchart showing the workflow for LC-MS-based BSOA spectral library building and matching to Beijing filter samples 46
Figure 19: Carbon number distribution of SOA tracers formed from different VOC precursors. Each bar represents the relative fraction of the number of tracers grouped by carbon number: <C10, C10, C11-C20, and >C20. The y-axis maximum is set to 20, reflecting the number of compounds listed in the tables of the 20 most abundant tracers 59
Figure 20: Total SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017, highlighting a spike on the 20th of June 69
Figure 21: Total SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017. The y-axis is set from 0-100,000 to zoom in on baseline periods, and gaps in the line show points at which the y-axis values exceed 100,000 69
Figure 22: Average contributions of each precursor compound to SOA mass in Beijing. Percentages represent each compound's share of the total measured SOA mass 70
Figure 23: Normalised stacked bar chart showing changes in the contributions of each precursor to SOA mass across 98 samples in Beijing 71
Figure 24: Diurnal variations in total SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. Sample 279 and all zero values were removed . 73
Figure 25: Diurnal variations in total SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. Sample 279 and all 0 values were retained 74
Figure 26: Diurnal variations in p-cymene (top left), α-terpineol (top right), 3-carene (middle left), ocimene (middle right) and sabinene (bottom). SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. All zero values were removed, and for p-cymene, α-terpineol, 3-carene and sabinene, Sample 279 was excluded. For ocimene, Sample 295 was removed 75
Figure 27: Diurnal variations in linalool (top left), β-caryophyllene (top right), α-cedrene (middle left), α-pinene (middle right) and terpinolene (bottom). SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. All zero values were removed, and for linalool, β-caryophyllene, α-cedrene and α-pinene, Sample 279 was excluded 76

Figure 28: Variation of monoterpene (β-pinene) emissions as a function of temperature. The line represents the best fit to the data points and demonstrates an exponential increase in β-pinene emissions with respect to temperature 78
Figure 29: Diurnal variations in O_3 concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. 79
Figure 30: Diagram showing the typical diurnal changes in the boundary layer (BL) depth. At night, a shallow stable boundary layer (SBL) forms near the surface due to cooling, suppressing turbulence. After sunrise, surface heating creates turbulence that grows the convective boundary layer (CBL) upward throughout the day. The residual layer (RL) above the SBL consists of the previous day's CBL. After sunset, the CBL collapses and the SBL reforms 81
Figure 31: Time series graphs of p-cymene (top left), α-terpineol (top right), 3-carene (middle left), ocimene (middle right) and sabinene (bottom) SOA tracer concentrations (peak area/m^3) from the 11th-25th of June 2017, demonstrating overall variations in tracer levels during the monitoring period 83
Figure 32: Time series graphs of linalool (top left), β-caryophyllene (top right), α-cedrene (middle left), α-pinene (middle right) and terpinolene (bottom) SOA tracer concentrations (peak area/m^3) from the 11th-25th of June 2017, demonstrating overall variations in tracer levels during the monitoring period 84
Figure 33: Time series graph of temperature (°C) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period 86
Figure 34: Time series graph of temperature (red, °C), and α-pinene concentrations (purple, peak area/m^3) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period 86
Figure 35: Boxplot showing total SOA tracer concentrations across four temperature intervals: 15-25°C, 25-30°C, 30-35°C and 35-40°C. Temperature measurements were averaged to filter times, and Sample 279 and all zero values were removed 87
Figure 36: Time series graph of relative humidity (%) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period. 89
Figure 37: Polar plot showing the mean total SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed 90
Figure 38: Polar plots showing the mean p-cymene (top left), α-terpineol (top right), 3-carene (middle left), linalool (middle right) and B-caryophyllene (bottom) SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed 91
Figure 39: Time series graph of O_3 concentrations (ppbv) from the 11th-25th of June 2017, plotted using raw 15-minute data (top) and averaged to BSOA filter times (bottom) 93

Figure 40: Polar plot showing the mean O_3 concentrations (ppbv) as a function of wind direction and wind speed..... 93
Figure 41: Time series graph of NOx (NO + NO2) concentrations (ppbv) from the 11th-25th of June 2017, plotted using 15-minute data (top) and averaged to SOA filter times (bottom)……..……....95
Figure 42: Polar plot showing the mean NOx (NO + NO2) concentrations (ppbv) as a function of wind direction and wind speed using 15-minute NOx data 95
Figure 43: Time series graph of OOA (OOA1 + OOA2 + OOA3) concentrations (µg/m^3) from the 11th-25th of June 2017, plotted using raw 15-minute data (top) and averaged to SOA filter times (bottom). Data between the 14th-17th is missing. 97
Figure 44: Polar plot showing the mean total OOA (OOA1 + OOA2 + OOA3) concentrations (µg/m^3) as a function of wind direction and wind speed..... 97
Figure 45: Polar plots showing the mean α-pinene (top left), α-cedrene (top right) and sabinene (bottom) SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed. 99
Figure 46: Total SOA tracer concentrations (peak area/m^3) from the 11th-25th of June 2017 before (peach) and after (blue) filtering. 101
Figure 47: Correlation heatmap with the 59 SOA tracers on each axis. Red shows a +1 correlation coefficient, yellow indicates a 0 correlation coefficient, and blue indicates a -1 correlation coefficient. The dark line across the middle represents +1 correlations of each compound's relationship with itself. Dendrons show hierarchical clustering of SOA tracers, with branches being short within clusters to show tight relationships and long between clusters of more distinct groups..... 103

Acknowledgements

Firstly, I would like to thank my supervisor, Prof. Jacqui Hamilton, for her support throughout the project and for helping me integrate into a research environment effectively. Secondly, I'd like to thank Dr Dan Bryant for his patience and willingness to devote time to the project alongside his own commitments. Lastly, I would like to thank my family for their endless support and curiosity regarding my research.

Author's Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

This research was conducted under the supervision of Professor Jacqui Hamilton.

I also acknowledge the contributions of Dr Dan Bryant, who assisted in laboratory work and method development.

Barbora Bracinik
16/10/2025

1. Introduction

1.1. Introduction to Aerosols

Atmospheric aerosols refer to tiny solid or liquid particles suspended in the atmosphere. The particles range in size from a few nanometres (nm) to several tens of microns (µm), covering a size range of more than five orders of magnitude. Fine particle sizes range from a few nm to 2.5 µm, and sizes greater than 2.5 µm are classified as coarse particles.[1] Within the fine fraction, the nucleation mode consists of newly formed particles smaller than ~20 nm, the Aitken mode spans roughly 20-100 nm, and the accumulation mode covers sizes from ~100 nm up to 2.5 µm.

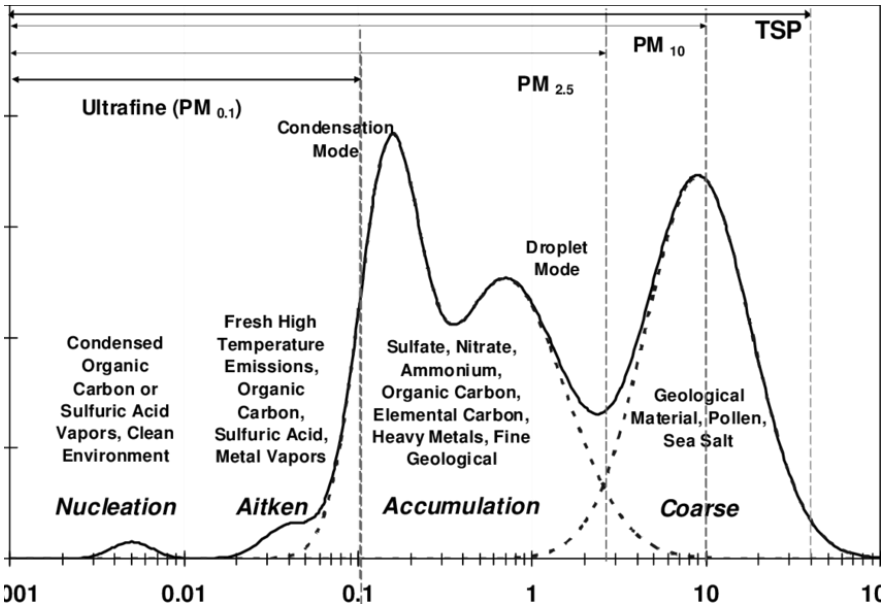


Figure 1: Illustration of different modes in a typical atmospheric particle size distribution. The main sources and processes influencing each mode are shown, along with their classification as ultrafine particles, $PM_{2.5}$, PM_{10}, or total suspended particles (TSP).[2]

The ultrafine particles (sizes 1 nm-100 nm) mainly form through condensation of hot vapours during combustion processes, leading to nucleation of gaseous species to form new particles.[1] The slightly larger accumulation particles, however, are generated through coagulation (sticking together) of small particles, and condensation of vapours onto existing particles, causing them to grow significantly. Among the coarse particles, those with sizes ranging from 2.5 µm to the upper limit of 100 µm mainly contain man-made, sea salt and

natural dust aerosols, which are subject to sufficiently high settling speeds and therefore are efficiently removed in short times.

In the context of air pollution, aerosols are typically categorised based on the aerodynamic diameter of the particles irrespective of their origin and chemical composition, since this heavily affects how they behave in the atmosphere and how deeply they can penetrate into the respiratory system.[3] Common categories are PM_{10} (particulate matter with an aerodynamic diameter of less than 10 µm), $PM_{2.5}$ (less than 2.5µm) and PM_1 (less than 1µm). Smaller aerosol particles pose a greater risk to human health because they can penetrate deeper into the respiratory system, and in some cases, enter the bloodstream.

1.2. Types of Aerosols

Atmospheric aerosols originate from a wide variety of natural and anthropogenic sources, as illustrated in Figure 2.

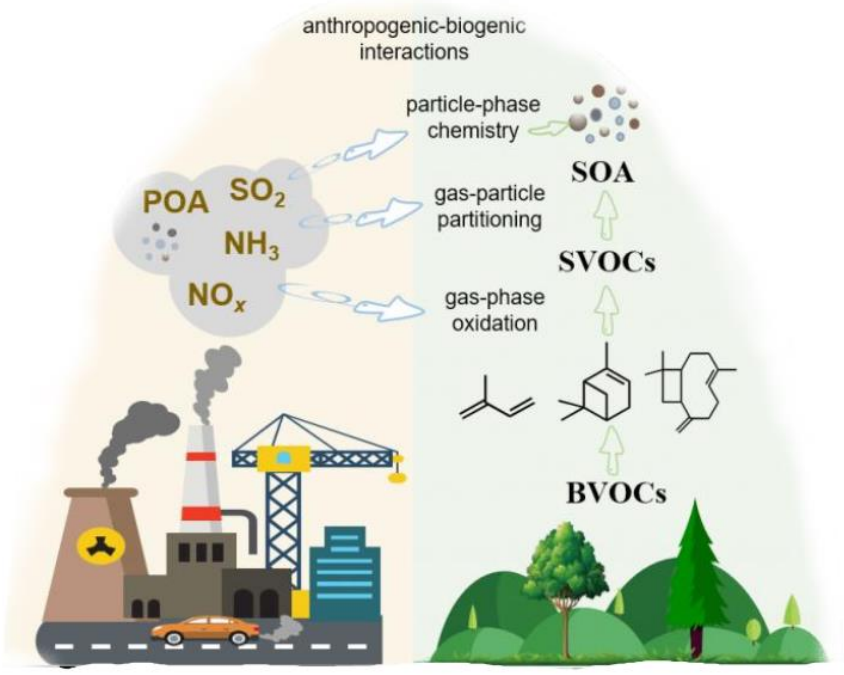


Figure 2: Examples of sources of volatile organic compounds (VOCs) in the atmosphere, both anthropogenic and biogenic, and the processes they undergo throughout their life cycles in order to become aerosols.[4]

Aerosols are broadly classified into two main categories: primary and secondary particles. Primary particles are emitted directly from their source, with typical sources

including biomass burning, combustion of fossil fuels, volcanic eruptions and wind-driven suspension of soil, mineral dust, sea salt, and biological materials.[5,6]

In contrast, secondary particles are formed in the atmosphere from emissions of volatile pollutants, which undergo chemical reactions with oxidising agents, leading to gas-particle conversion processes such as condensation of gases. This causes nucleation of new particles and heterogeneous and multiphase chemical reactions.

An alternative classification distinguishes aerosols based on their chemical composition as either inorganic or organic. Inorganic aerosols encompass both primary inorganic aerosols (PIA), such as mineral dust, volcanic ash, and sea salt, and secondary inorganic aerosols (SIA). SIA particles form in the atmosphere through in situ gas-to-particle conversion processes. They are produced either by condensation of oxidised sulphur and nitrogen-containing species onto existing particles, or by nucleation of new particles.[2]

Much like inorganic aerosols, organic aerosols can be classified as primary organic aerosol (POA) and secondary organic aerosol (SOA). POA is emitted directly in particulate form from various emission sources, including vehicular exhaust, fuel combustion, and cooking.[7] SOA, by contrast, is produced in the atmosphere from the oxidation of volatile organic compounds (VOCs), forming oxidation products which subsequently partition into the particle phase.

1.3. Secondary Organic Aerosol formation

SOA formation is a pathway that remains highly uncertain, as atmospheric oxidation of VOCs may lead to either gas-particle partitioning into low volatility species or to complete conversion into CO_2 and H_2O. These oxidation steps can occur throughout the daytime (mainly with OH radicals and ozone (O_3)) or during the night-time with O_3 and other oxidants such as nitrate radicals (NO_3).[8]

The initial oxidation step of VOCs leads to the generation of a set of organic products containing one or more polar oxygenated functional group, such as aldehyde (−C(=O)H), ketone (−C(=O)−), alcohol (−OH), nitrate (−ONO_2), peroxyacyl nitrate (−C(=O)$OONO_2$), carboxylic acid (−C(=O)OH), hydroperoxide (−OOH) or percarboxylic acid (−C(=O)OOH) groups.[3] This tends to make the products less volatile and more water soluble, with a high tendency to stick to pre-existing particles or to nucleate new particles in the presence of less volatile species such as sulfuric acid.[9] Further oxidation may introduce additional functional

groups, producing a second generation and subsequent generations of even lower volatility and higher solubility products.[6]

The complexity of the situation is compounded further as the oxidation mechanisms often lead to fragmentation of the carbon chains to form lower molecular weight oxygenates, which may be more volatile. This creates a complex mixture of oxidised VOCs in the atmosphere, along with their various possible fragments.[9] Given enough time, the gas phase oxidation would eventually convert all carbon in the emitted VOCs not sequestered into particulate matter or lost to wet or dry deposition, into CO_2 or H_2O.[6]

1.4. Global contributions of SOA

When assessing the significance of SOA, often models are used to quantify the total aerosol burden.[10] The traditional approach is a bottom-up estimate, where known biogenic (most notably isoprene and terpenes) and/or anthropogenic VOC precursor fluxes are combined in global models with laboratory data from oxidation experiments to obtain a global organic aerosol flux.

Bottom-up modelling approaches estimate total biogenic SOA (BSOA) fluxes in the range of 12-70 Tg/yr, corresponding to biogenic secondary organic carbon (BSOC) fluxes of 9-50 TgC/yr.[6] In addition to these BSOA fluxes, the bottom-up models typically account for emissions of primary organic aerosol (POA) at about 35 TgC/yr, which includes ~9 TgC/yr from anthropogenic sources and ~25 TgC/yr from open biomass burning. These models also incorporate a smaller anthropogenic SOA contribution, estimated at 2-12 Tg/yr.

When contributions from all modelled sources (biogenic SOA, anthropogenic SOA, and primary organic aerosol from both biomass burning and fossil fuels) are combined, bottom-up estimates place the global organic aerosol burden at 50-90 TgC yr^{-1}, with values generally clustering toward the lower end of this range. Other studies, however, report a much wider spread for bottom-up estimates, ranging from 16-121 TgC yr^{-1}.[11]

1.5. Contributions of SOA to total PM masses

Figure 3 below shows mass fractions of different components of PM_1 aerosols at multiple sites across the Northern Hemisphere.[12] HOA denotes hydrocarbon-like organic aerosol, typically associated with primary emissions (POA). SV-OOA and LV-OOA correspond to semi-volatile and low-volatility oxygenated organic aerosols, respectively, both

representing secondary organic aerosol (SOA) at various stages of ageing. The central box plot shows the oxygen-to-carbon (O:C) ratios of the organic aerosol components.

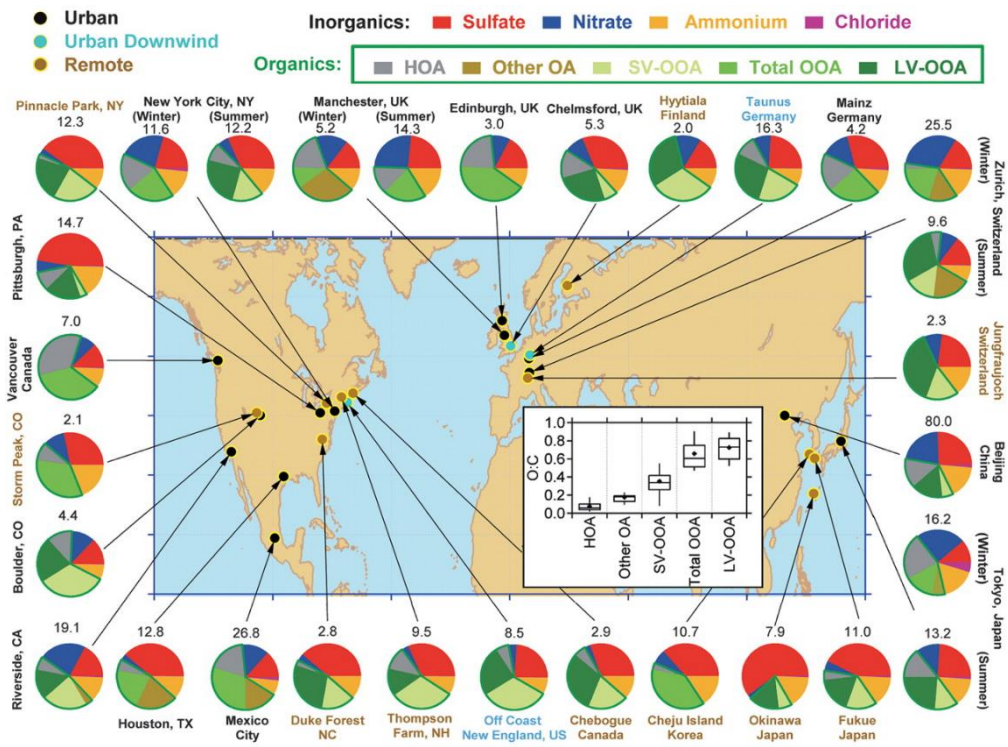


Figure 3. Total mass concentrations (μg m⁻³) and mass fractions of inorganic species and organic components in submicrometer (PM_1) aerosols at multiple surface sites across the Northern Hemisphere.[12]

Figure 3 illustrates that organic aerosol (shown in the green outlined section of the pie charts) often makes up the largest fraction of submicron tropospheric particulate mass in the Northern hemisphere (ranging from about 20 % up to around 70 %).[12] Oxygenated OA (OOA), shown in the green sections of the pie chart, generally dominates over hydrocarbon-like OA (HOA), shown in grey.

Similarly, a comprehensive analysis of 22 European cities reported that OA contributes 21-75 % of total submicron aerosol (PM_1).[13] Within this OA fraction, oxygenated organic aerosol (OOA) dominates, accounting for 71.1 % of submicron OA on average (range: 43.7-100 %). By summing the more oxidised (MO-OOA) and less oxidised (LO-OOA) components, the study demonstrated that secondary OA is the primary contributor to PM_1, representing on average 34.5 % (range: 11.7-62.4 %) of total mass and contributing to the majority of the OA burden.

Building on this, earlier work has also examined how SOA and POA contributions vary in broader fine particulate matter ($PM_{2.5}$) across different regions. SOA fluxes make up a large fraction of $PM_{2.5}$ in some places, while in others, POA dominates.[9] Studies in Monterrey, Mexico and Bengbu, China, both reported that SOA accounts for around 32-45 % of total $PM_{2.5}$ mass, with Monterrey showing a higher contribution of 59-87 % to organic carbon (OC).[14,15]

However, research in Nanchang, China, an equally urbanised yet slightly larger city than Bengbu, indicated that SOA contributes a modest 4 % of $PM_{2.5}$ mass, therefore showing POA is the primary $PM_{2.5}$ component at this location.[16]

It has also been shown that the contribution of SOA to submicron aerosol dominates even more in remote forested locations such as the Amazon, accounting for as much as 50-85 % of the total organic aerosol burden.[10]

1.6. Impact of atmospheric aerosols: climate

Aerosols impact the Earth's radiative balance through their direct effects of absorption and Mie scattering of solar radiation, and through their indirect effects of altering cloud microphysical properties, both of which affect the planetary energy budget.[10]

Incoming solar radiation intensity peaks at wavelengths between 380 nm and 750 nm.[17] When light hits a particle that is roughly the same diameter as the light's wavelength, the particle scatters the light. The result of this interaction is the re-emission of light in all directions, also known as Mie scattering. As seen in Figure 4, re-emission of light in the direction of the incoming light source is favoured.

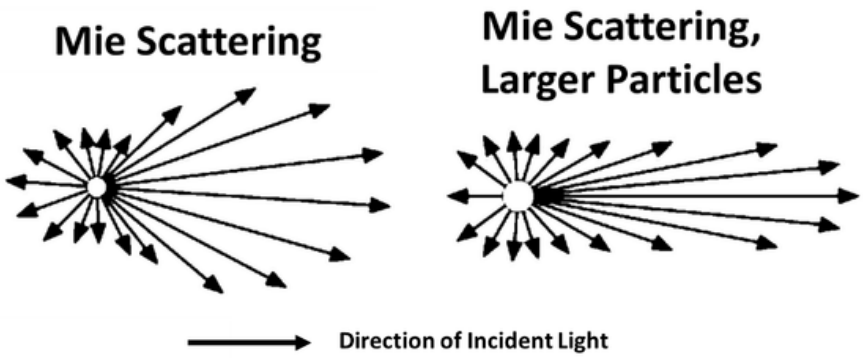


Figure 4: Mechanism of Mie scattering on smaller and larger particles, showing the strong favouritism towards re-emission in the direction of the incoming light source.[18]

Processes that aid in the growth of particles up to these sizes, such as the condensation of secondary organic material, will aid in the formation of an aerosol layer in the atmosphere capable of causing significant Mie scattering.[17]

Aerosol optical depth is a measure of how much sunlight is blocked or scattered by aerosols in a column of the atmosphere. Studies in the Beijing-Tianjin-Hebei (BTH) region showed that a 41 % decrease in monthly mean $PM_{2.5}$ concentrations led to a 20 % decrease in aerosol optical depth.[19] The increase in scattering albedo caused by this layer will have an overall cooling effect, due to less radiation reaching the Earth's surface.

A small fraction of aerosol particles, such as black carbon (BC) and brown carbon (BrC) emitted from fossil fuel combustion and biomass burning, have the ability to absorb sunlight, a characteristic determined by their optical properties.[20] These aerosols contribute to localised atmospheric warming by increasing the absorption of solar radiation, which is then released as heat, causing the particles to behave similarly to greenhouse gases in their warming effect.

The effect of climate factors on temperatures is defined using a measure known as radiative forcing. This measure is defined as the change in net radiative energy at the tropopause or top of the atmosphere, usually measured in watts per square meter (W m^{-2}), due to a change in a specific climate factor (like CO_2, methane, or aerosols) relative to a baseline period (often pre-industrial times).[21]

The direct annual radiative effect (DRE), which accounts for the absorption and scattering of sunlight by aerosols, has been estimated to be -5.36 ± 0.04 W m^{-2} at the top of the atmosphere (TOA), with the corresponding effect at the surface (SFC) measured at -9.64 ± 0.07 W m^{-2}.[22] The difference of $+4.27 \pm 0.04$ W m^{-2} between these values corresponds to energy retained within the atmosphere, demonstrating a net atmospheric warming. While scattering aerosols dominate and produce a net cooling effect at the top of the atmosphere, the absorption by certain aerosol types contributes appreciably to heating within the rest of the atmosphere.

Regional effects such as those over tropical forest regions may be much higher, however, due to the DRE being most sensitive to the yield of SOA production and strongest in regions of high biogenic VOC (BVOC) emission.[17]

A study modelling the shift from pre-industrial to present-day anthropogenic SOA levels found the SOA-related radiative forcing to be approximately -0.09 W m^{-2} (when including partitioning to organic and sulphate aerosols), and -0.06 W m^{-2} when only partitioning to organic aerosols was considered.[23]

These direct effects also lead to many secondary impacts, such as altering the photolysis rate of NO_2, a key parameter in determining O_3 formation (see equations R1 and R2). Both numerical simulations and observations showed that aerosols with a scattering effect may accelerate photochemical production of O_3 due to increased UV scattering within the atmosphere, therefore meaning aerosols with strong absorption properties have the opposite effect.[17]

$$NO_2 + h\nu\ (< 410\ \text{nm}) \rightarrow NO + O\ (^3P) \qquad \text{(R1)}$$

$$O_2 + O\ (^3P) + M \rightarrow O_3 + M \qquad \text{(R2)}$$

The indirect effects of atmospheric aerosols mainly occur through their influence on cloud microphysics. Aerosol particles can serve as cloud condensation nuclei (CCN) or ice-nucleating particles (INPs), influencing cloud formation and development. These particles provide surfaces upon which water vapour can condense, facilitating the growth of cloud droplets or ice crystals.[8,17] Aerosols become active as cloud condensation or ice-nucleating particles based on their physical and chemical properties, particularly their size, ability to absorb water, and chemical composition. When aerosol particles reach a sufficient size and possess suitable chemical properties, they can activate under supersaturated conditions and initiate cloud droplet formation. This process affects cloud albedo, lifetime, and precipitation efficiency, representing a key factor contributing to global cooling.

As shown in Figure 5, in more polluted environments where aerosol numbers are higher, clouds tend to form with smaller droplets because the available water vapour is distributed amongst more particles. This results in longer-lasting clouds with reduced precipitation potential and increased reflectivity, contributing to a higher cloud albedo.

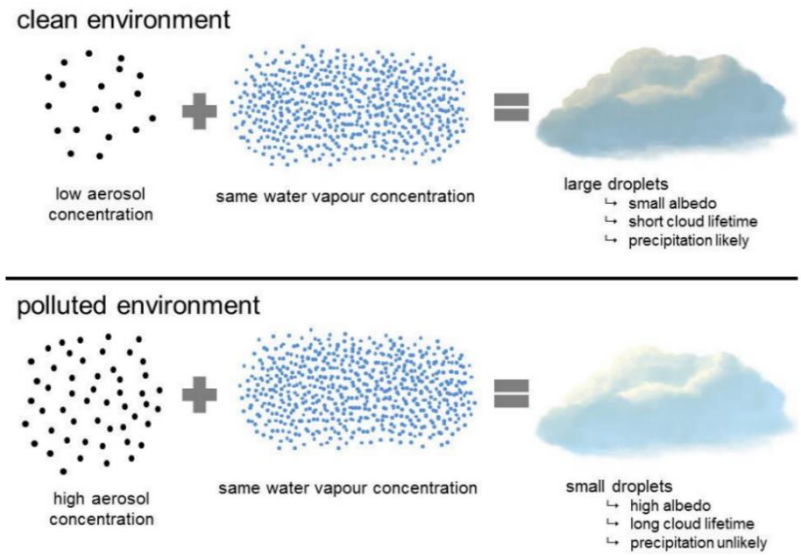


Figure 5: The indirect effects of aerosols on cloud formation, illustrated by differences between polluted environments with high aerosol concentrations and clean environments with low aerosol concentrations, both having the same water vapour concentration.[24]

A recent modelling study suggested that low-volatility SOA may be responsible for more than half of the CCN in many continental locations.[8] Biogenic SOA can increase the global annual mean concentration of CCN by 3.6-21.1 %, depending on various factors. This increase in CCN leads to a rise in global annual mean cloud droplet number concentration (CDNC) of 1.9-5.2 %, and a global mean first aerosol indirect effect (AIE) ranging between +0.01 W m^{-2} and −0.12 W m^{-2}, showing an overall cooling effect due to increased scattering of radiation.[17]

1.7. Impact of atmospheric aerosols: health

Recently, there has been growing interest into $PM_{2.5}$ and its significant health effects.[25] It is estimated to be responsible for up to 9 million deaths per year worldwide, including 15 deaths per 100,000 people in the US, with multiple studies showing a link between exposure to $PM_{2.5}$, hospital admissions and mortality.[26] Exposure of PM to the lungs, the first organs being targeted, has been shown to be associated with airspace damage and alterations of lung function, resulting in a number of chronic pulmonary diseases such as chronic obstructive pulmonary disease.[25]

In Beijing, the health burden remains substantial despite significant improvements in air quality over recent years. Recent estimates using the Global Exposure Mortality Model

(GEMM) indicate that $PM_{2.5}$ exposure in Beijing was still responsible for approximately 48,969 premature deaths in 2021 alone.[27] While mortality attributable to $PM_{2.5}$ decreased by around 34% between 2016 and 2021 the overall health burden remains high even at concentrations below China's national air quality standard (35 μg m^{-3}).

Animal studies showed chronic exposure to PM resulted in lung inflammation and tissue destruction.[26] Similarly, a series of human cell culture studies also reported the capacity of PM to induce cell toxicity, while other studies reported increased RNA oxidation in cells exposed to PM, resulting in transcript alteration and cell function disruption.[25] Long-term exposure to fine PM has also been associated with a heightened risk of pulmonary diseases, including acute respiratory distress, asthma, chronic obstructive pulmonary disease (COPD), and lung cancer.

According to the size, PM is deposited in either bronchi or alveolar spaces, where it can be incorporated into cells. As well as this, $PM_{2.5}$ has the capacity to cross the alveolar capillary barrier and reach other organs, contributing to the development of other pathologies like cardiovascular diseases, diabetes, or birth defects.

Interestingly, a study done in the U.S. showed that SOA was strongly associated with death rates independent of the total $PM_{2.5}$ mass.[28] The study provided evidence that SOA had a higher rate of mortality by comparison to general $PM_{2.5}$. Biogenic and anthropogenic carbon sources both played a role in the overall SOA association with mortality throughout the U.S., except for parts of the southeast, in which anthropogenic SOA was found to be insignificant.

Not only was SOA associated with higher cardiorespiratory death rates than other components considered, but it also showed a larger association per unit mass than total $PM_{2.5}$. Across the contiguous U.S., SOA was associated with 8.9 additional deaths per 100,000 people for a 1 μg m^{-3} increase in concentration, indicating about 6.5× higher association than total $PM_{2.5}$. In the southeastern U.S., the same increase in SOA was associated with a 3.0× greater impact on per capita cardiorespiratory mortality compared to the national average.

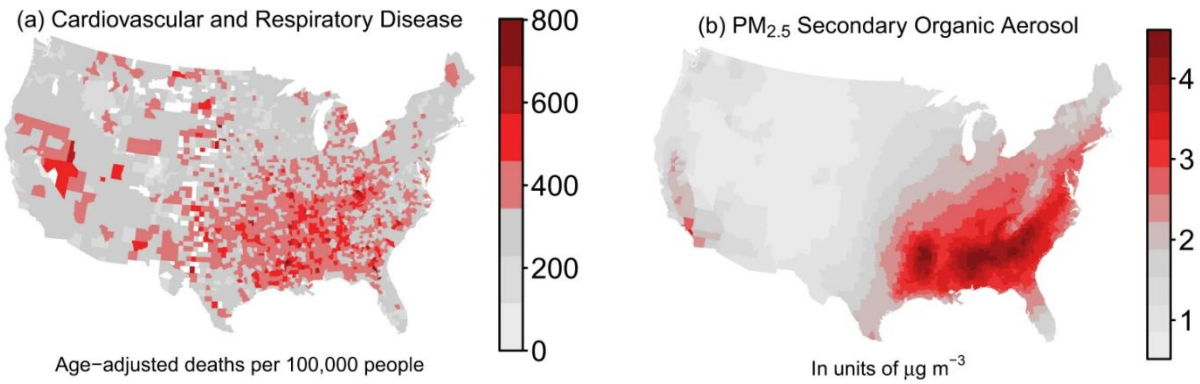


Figure 6: Geographic distribution of cardiovascular and respiratory disease rates across the USA (left) compared with concentrations of $PM_{2.5}$ across the USA (right).[28]

1.8. BVOCs and BSOA

Biogenic secondary organic aerosol (BSOA) makes up a significant fraction of global SOA formation and is formed through oxidation of biogenic VOCs (BVOCs), which are emitted naturally from plants and animals.[17] Subtraction of the estimated anthropogenic and biomass burning SOA from the estimate for total volatile SOA leads to a net BSOA production rate of around 90 TgC/yr.[6]

Globally, BVOC emissions are as high as 1000 Tg/year, accounting for 70-90 % of the total VOCs emissions, far more than the VOCs released from anthropogenic activities.[29] Uncertainties in BVOC emission estimates, however, contribute to the overall uncertainties associated with model predictions of SOA production rates.

The sources of BVOCs are mainly derived from terrestrial ecosystems, such as terpenes, which are a class of hydrocarbons that evaporate from leaves.[1] About 80 % of these emissions oxidise to organic aerosols within about 1 h. Isoprene accounts for the largest fraction of BVOCs (70 %, ~700 Tg/year), largely due to its abundance in many forested regions.[30] Once in the atmosphere, it rapidly reacts with OH, giving it a lifetime of ~1.8 h for [OH] ~1.5×10^6 cm^3, making it a key contributor to SOA formation.[31] It occupies a prominent position in tropospheric chemistry, and its oxidation mechanism is of critical importance to understanding the VOC budget of the troposphere and SOA formation.

Other BVOC sources include monoterpenes accounting for 11 % (~110 Tg/year), methanol for 6 % (~60 Tg/year) and other VOCs for 14 % (~140 Tg/year) of the BVOC total expressed as the emission of carbon.[30] Examples of monoterpenes are α-pinene, β-pinene, limonene, trans-β-ocimene, myrcene, sabinene and 3-carene, and their relative contribution to the monoterpene annual global emissions averaged over the 1980-2010 period is presented in Figure 7.

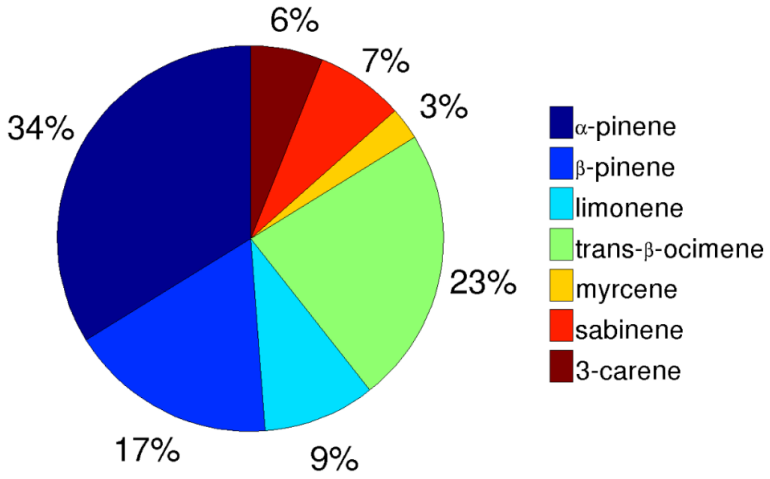


Figure 7: Relative contribution of monoterpene species to monoterpene annual global total averaged over the 1980-2010 period. The considered monoterpene species are listed to the right, and the 'other monoterpenes' fraction has been excluded.[30]

Amongst the other 14 % of BVOC emissions are sesquiterpenes, a class of BVOCs mainly emitted from coniferous trees, scots pines and deciduous trees.[32] Despite their lower atmospheric emission concentrations compared to isoprene and monoterpenes, they are important in the atmosphere because of their high reactivities with O_3 and large aerosol formation potentials. Among all the sesquiterpene species, β-caryophyllene (BCP) has been shown to be one of the most abundant.

1.9. Global BVOC emissions

BVOC emissions are highly sensitive to environmental conditions, especially temperature, solar radiation, and land cover, including vegetation type and amount.[8] As shown in Figure 8, modelled spatial distributions of isoprene and monoterpene emissions averaged over 1980-2010 demonstrate pronounced seasonal and latitudinal variability, with emissions peaking during the summer months in vegetated regions.

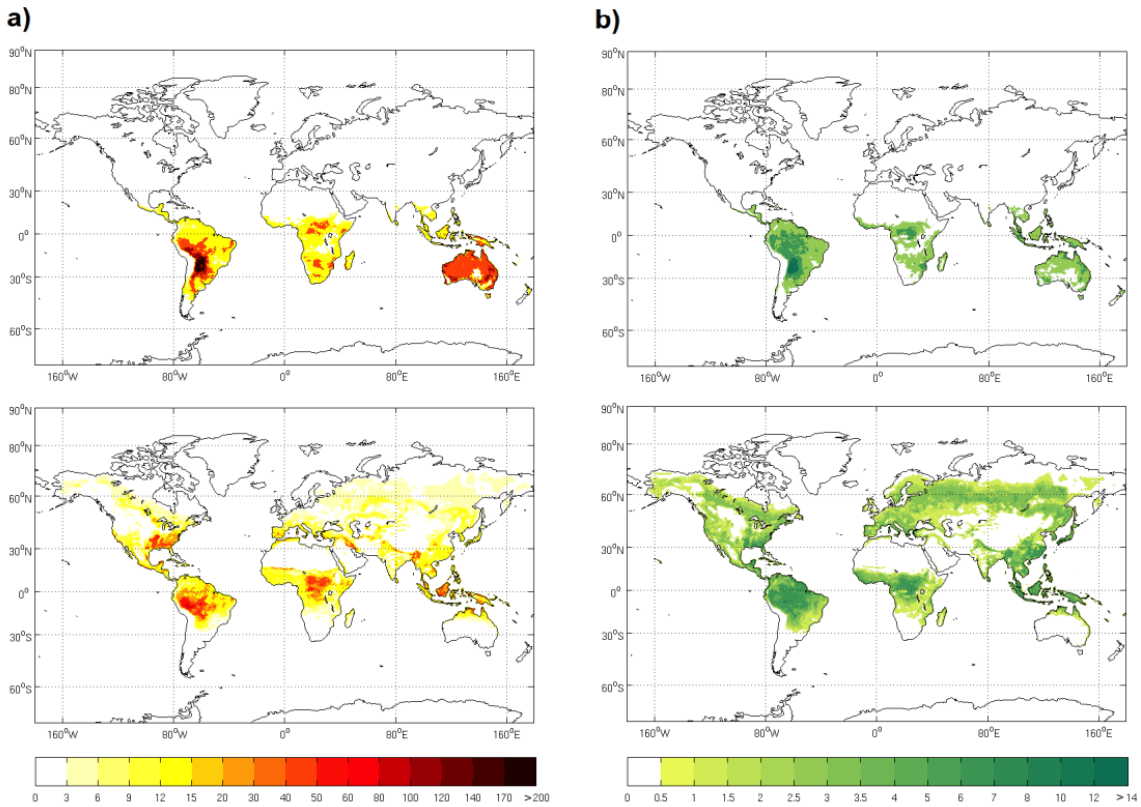


Figure 8: Spatial distribution of monthly mean emissions of (a) isoprene, (b) monoterpenes, (mg m^{-2} day^{-1}) for January (top) and July (bottom) averaged over the modelled period 1980-2010.[31]

Isoprene is mainly emitted in the south-tropical region, which contributes about 56 % to the global total isoprene emissions in January, followed by the northern tropics (32 %) and the northern and southern temperate regions contributing 7 % and 4 % respectively.[31] Less than 1 % of global isoprene is emitted in the northern Arctic. Isoprene emissions in July shift toward the Northern Hemisphere, reflecting higher temperatures and therefore increased plant activity in northern regions.

In contrast, monoterpene emissions show a more balanced hemispheric distribution, with their annual maximum occurring during the Northern Hemisphere summer. Results highlight the main emitting regions as the southern and northern tropics, contributing 48 % and 35 % to the global monoterpene total in July, respectively. There is also a significant source of monoterpenes in the northern temperate region (13 %), while the southern temperate region contributes about 3 %, and the northern Arctic less than 2 %.

1.10. Known BVOC oxidation mechanisms

Ozonolysis of BSOA has been shown to yield a wide array of products via different formation pathways. The 2D volatility basis set, as shown in Figure 9, can be used to group the organic molecules based on their saturation concentration (C^0, a measure of volatility) and their O:C ratios.[33] They are classified into extremely low volatility (ELVOC), low volatility (LVOC), semi-volatile (SVOC), intermediate volatility (IVOC), and VOC. The arrows represent typical chemical processes that transform a VOC precursor (n) into its products, contributing to secondary organic aerosol (SOA) formation.

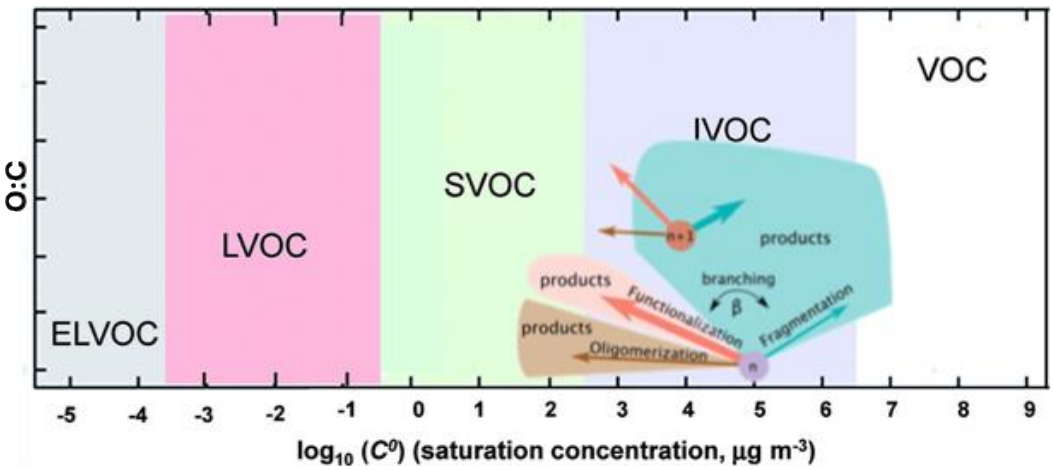


Figure 9: Classification of precursor compounds by saturation concentration (C^0) and O:C ratios. Categories range from VOCs to ELVOCs. Arrows indicate chemical processes that contribute to SOA formation. (n) represents the precursor VOC, therefore (n+1) represents the 1st generation oxidation product. [33]

For monoterpenes ($C_{10}H_{16}$), two dominant SOA formation processes are functionalisation and fragmentation. The pathway that leads to SOA formation depends largely on the location of the double bond(s) within the precursor molecule. When the double bond is exocyclic (outside the ring), or aliphatic, ozonolysis and OH reactions are more likely to cause fragmentation, breaking the carbon chain and producing smaller molecules with fewer than ten carbon atoms.[34] This is shown in Figure 10 for terpinolene.

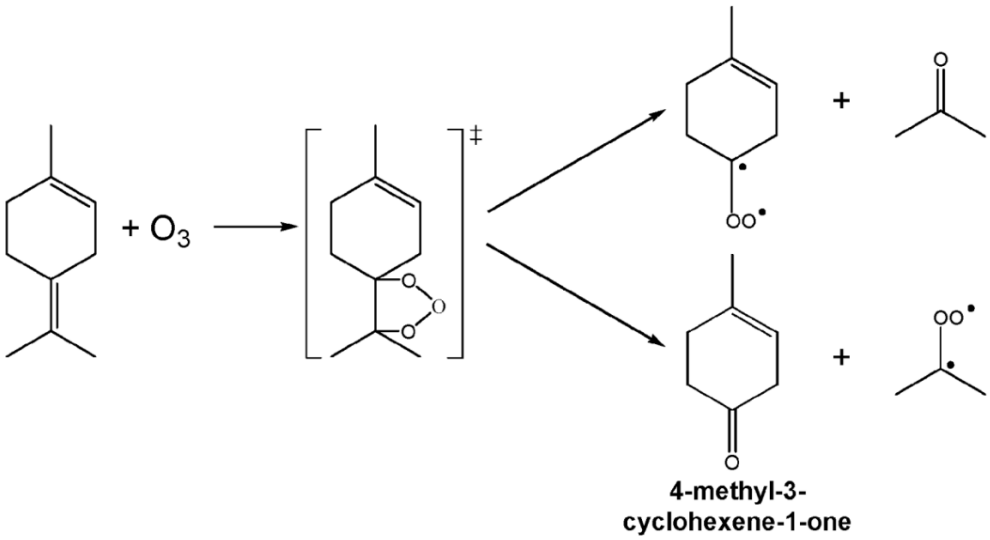


Figure 10: The proposed reaction scheme showing the initial O_3 addition to terpinolene. The diagram illustrates fragmentation occurring since O_3 attacks the exocyclic double bond.[34]

When the double bond is endocyclic (within a ring), ozonolysis typically leads to production of Criegee intermediates that rearrange to form functionalised products that retain the original carbon skeleton, as shown for α-pinene in Figure 11.[34] OH chemistry can lead to both ring-retained and ring-opened first-generation products.

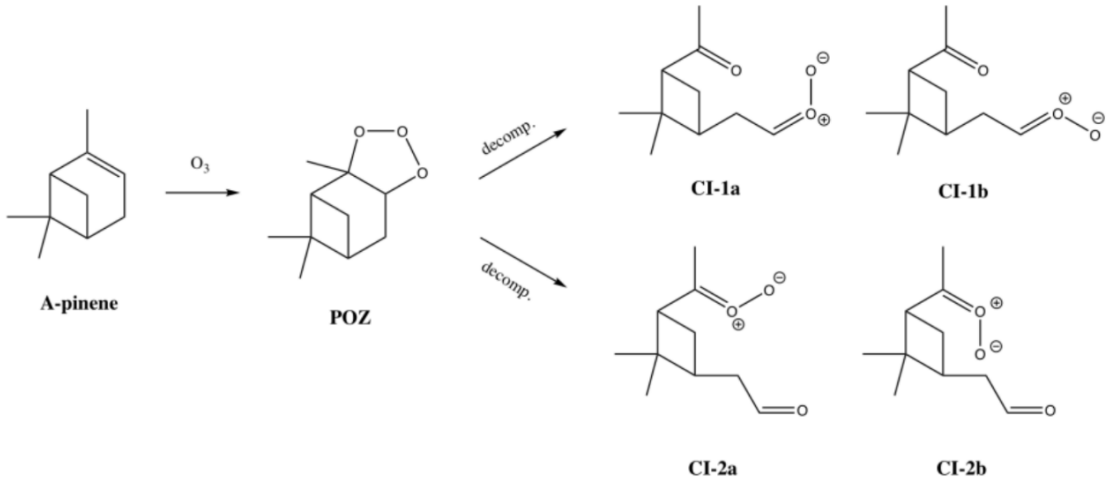


Figure 11: The ozonolysis of endocyclic α-pinene, showing O_3 attacking the double bond within the carbon ring. This leads to the production of Criegee intermediates, which lead to formation of products in which the carbon backbone is retained.[35]

Another formation pathway is oligomerisation or accretion reactions. For α-pinene and other BVOCs, products generated through reactions like ozonolysis or hydroxyl radical (OH) oxidation can undergo condensation reactions, such as esterification, hemiacetal formation, or peroxy radical cross-reactions.[36] These reactions link smaller molecules together without completely breaking down their carbon backbones, producing oligomers with higher molecular weights.

Oligomer formation increases the size and lowers the volatility of the organic compounds, enhancing their ability to remain in the particle phase and contribute to SOA mass and stability.[6] These oligomers can also influence the physical properties of aerosols, like viscosity and hygroscopicity, affecting cloud formation and climate impacts.

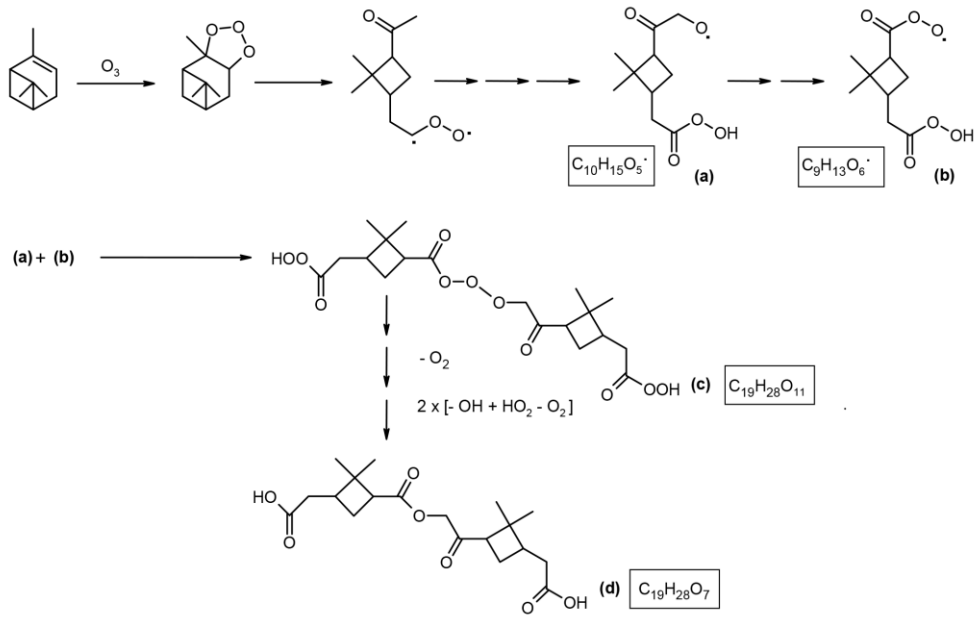


Figure 12: Simplified proposed mechanism for α-pinene oligomer formation, leading to the production of the MW 368 ester.[36]

In addition to ozonolysis mechanisms, the OH radical plays a significant role in the atmospheric oxidation of VOCs. As a highly reactive species, OH can initiate reactions through hydrogen abstraction or addition to double bonds, producing a wide range of oxygenated products.[37] These initial oxidation steps can lead to functionalisation, where oxygen-containing groups are introduced into the molecule. Under high oxidant conditions or through subsequent reactions, OH-driven oxidation can also result in fragmentation, breaking the carbon skeleton into smaller, more volatile products.[38] The products formed by OH oxidation can also subsequently undergo particle-phase reactions, including esterification and

peroxy radical interactions, contributing to the formation of high molecular weight oligomers.[39] In some cases, OH-initiated oxidation can also trigger auto-oxidation of VOC-derived peroxy radicals, a process leading to formation of highly oxygenated organic molecules (HOMs).[40]

1.11. SOA pollution in Beijing

Due to its rapidly developing nature, Beijing has been shown to have poor air quality and suffers from regular extreme air pollution events.[41] It is estimated that 1.36 million premature deaths in China in 2010 were a result of exposure to outdoor air pollution. The most dangerous pollutant to health in China is $PM_{2.5}$, with a recent study suggesting that a 50 % reduction in excess mortality requires a 62 % reduction in $PM_{2.5}$ in the Beijing-Tianjin-Hebei (BTH) region. The BTH region has, however, already seen a substantial decline in $PM_{2.5}$-attributable mortality between 2013 and 2020, with the attributable fraction dropping from ~28.8 to ~13.5 % in Beijing City.[42]

Measurements using Aerosol Mass Spectrometry (AMS), a technique commonly used for organic aerosol characterisation and source apportionment, indicated that SOA makes up a substantial fraction of the PM_1 mass, even in urban areas, but the sources are still poorly understood.[43] ^{14}C measurements, a technique used for distinguishing between fossil and non-fossil sources of carbonaceous aerosol, determined that non-fossil fuel emissions are generally a dominant contributor of secondary organic carbon (SOC) in Beijing, with a larger contribution in summer as a result of increased BVOC emissions.[44]

Studies at 14 Chinese sites found that isoprene SOA (iSOA) dominated the apportioned SOA mass (46 ± 14 %), contributing between 0.4 and 2.17 μg m^{-3} with an average of 1.59 μg m^{-3} in Beijing. This means between 40 and 68 % of SOA in Beijing originates from sources other than isoprene. Given the significant contribution of non-fossil fuel emissions to SOC, it is likely that a substantial portion comes from BSOA formed from non-isoprene BVOC emissions.

1.12. BVOC sources in Beijing

A study measuring emission factors across six dominant tree species in Beijing estimated total biogenic volatile organic compound (BVOC) emissions at approximately 38.2 Gg C/year, with monoterpenes contributing around 49 % and isoprene about 11 %.[45] The

dominance of monoterpenes is likely due to many of Beijing's dominant tree species, particularly coniferous trees, emitting large amounts of monoterpenes. A study which considered the mass of the whole compound rather than just the carbon content reported the BVOC emissions as 75.9 Gg/year for the Beijing area using model-based BVOC emission inventories created from known environmental and vegetation data.[46] Isoprene, monoterpenes, sesquiterpenes and other VOCs account for 37.6, 14.6, 1.8 and 46 % of the total, respectively.

Additionally, a field campaign conducted from June to September 2018 across 24 dominant tree species showed clear species-specific emission patterns.[47] Broad-leaved species exhibited high isoprene emission rates. In contrast, coniferous species were dominant emitters of monoterpenes. Emissions of sesquiterpenes were found to be typically 10 to 100 times lower than isoprene and monoterpenes.

Regionally, within the Jing-Jin-Ji area, Beijing accounts for 17.6 Gg C/year (24.9 %) of total BVOC emissions, second only to Chengde (28.8 %).[48] A 2018 study focusing on summer conditions further revealed that nearly 75 % of annual BVOC emissions occur during this season, with forested areas to the west and northeast of urban Beijing acting as key emission hotspots that contribute to the city's overall atmospheric reactivity.[49] These regions are a source of isoprene oxidation products, such as formaldehyde, which can be transported into urban areas and affect urban air chemistry.

1.13. Similar studies

Several previous studies have investigated the sources and formation pathways of SOA in Beijing, with a primary focus on bulk aerosol composition. For instance, Zhang et al. (2021) used ^{14}C isotope analysis to distinguish between fossil and non-fossil contributions to carbonaceous aerosols.[50] Similarly, Zhou et al. (2021) combined isotope measurements with organic tracers to assign primary and secondary organic aerosol at both urban and rural sites in the Beijing region.[51] These studies consistently reported that SOA constitutes a substantial fraction of total particulate matter, highlighting the importance of secondary formation processes in the region's aerosol pollution.

Despite their contributions, these studies are primarily limited to the identification of broad source categories or a small set of known chemical tracers. Methodologically, most rely on AMS, isotope analysis, or targeted gas and liquid chromatography, which often fail to

capture the full molecular diversity of SOA components. As a result, our understanding of the specific precursors, oxidation products, and formation mechanisms is incomplete.

With regards to research done on BSOA, a limited number of modelling studies have employed tools such as the Model of Emissions of Gases and Aerosols from Nature (MEGAN) and the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) to simulate BVOC emissions and assess their contributions to SOA and O_3 formation in the BTH region.[46,52] While these studies provide valuable regional insights, they are typically restricted by uncertainties in BVOC emission factors, which limits their ability to accurately represent the spatial and chemical complexity of urban BSOA formation.

1.14 Project aims

This project aims to provide an improved understanding of the role of BVOCs as key SOA precursors in Beijing during the summer. While previous research has highlighted uncertainties in SOA formation pathways, there remains a need for detailed measurements and identification of BSOA tracer compounds to allow for accurate identification of BVOC contributions in real-world environments.

The main question guiding this work is: how do specific BVOC precursors contribute to the formation and composition of SOA in an urban environment under photochemically active conditions?

To address this question, the study has the following objectives:

- To conduct controlled atmospheric simulation chamber experiments using the York Aerosol Chamber to produce BSOA from the reactions of O_3 and OH radicals with ten key BVOC precursors, including monoterpenoids and sesquiterpenes.
- To identify and characterise BSOA products using ultra-high-performance liquid chromatography coupled with tandem mass spectrometry (UHPLC-MS2) and non-target analysis.
- To compile these products into high-resolution mass spectral libraries for source identification.
- To apply these libraries to $PM_{2.5}$ filter samples collected in Beijing in 2017 as part of the Air Pollution and Human Health project, enabling identification of BSOA tracer compounds in ambient aerosol.[53]

- To investigate the influence of temporal dynamics and meteorological factors, including photochemistry and regional transport, on urban BSOA composition.

Controlled atmospheric simulation chamber experiments were conducted using the York Aerosol Chamber, and the resulting BSOA formed was characterised using UHPLC-MS2. The developed spectral libraries were used to identify BSOA tracer compounds in the ambient $PM_{2.5}$ filter samples, allowing for identification of tracer compounds in a real-world urban environment. By developing these detailed chemical fingerprints of BSOA composition, this project aims to enhance emission source identification of ambient aerosols and improve the accuracy of air quality assessments in polluted areas.

In contrast to previous work, this study offers a higher level of molecular resolution by combining chamber-generated spectral libraries with high-resolution mass spectrometry. This approach enables a more detailed characterisation of relationships between precursors and products, oxidation pathways, and the chemical complexity of SOA than is possible with bulk aerosol techniques. It provides insights into biogenic contributions to urban aerosol composition, particularly under photochemically active summer conditions.

Together, these findings emphasise the value of detailed molecular characterisation in improving our understanding of urban air quality. This work builds upon and extends previous research by integrating high-resolution analytical techniques with atmospheric influences, offering a greater understanding of BSOA formation in the context of a rapidly developing city like Beijing.

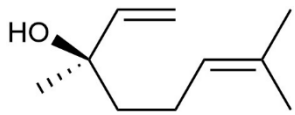
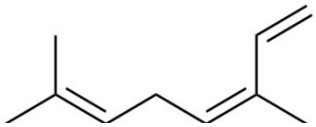
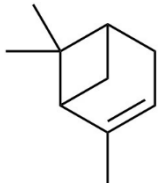
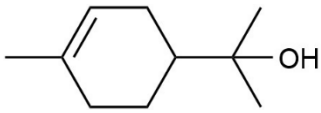
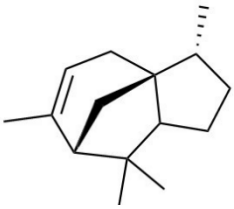
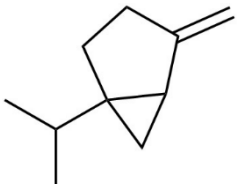
2. Methodology

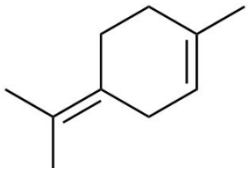
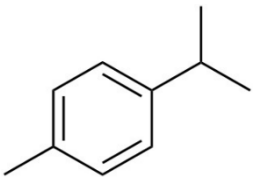
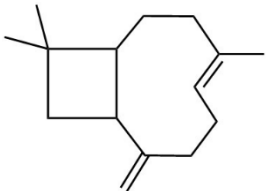
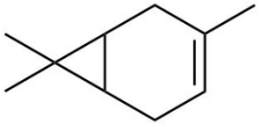
2.1. Chamber experiments

Atmospheric formation of SOA occurs through complex interactions between VOCs, oxidants, and environmental conditions, making it difficult to attribute specific oxidation products to individual precursors in ambient samples. Chamber experiments provide a controlled environment in which individual VOC precursors can be reacted with specific oxidants. This approach reduces atmospheric complexity and enables the identification of precursor-specific oxidation products.

In this study, ten key biogenic VOCs were selected for chamber experiments to generate BSOA for compositional analysis. These compounds are shown in Table 1 and include six monoterpenes with the highest emissions globally, two monoterpenols and two sesquiterpenes. *p*-Cymene has a molecular formula of $C_{10}H_{14}$ but is often classified as a mono-aromatic monoterpene, and so for this analysis, it has been included in the monoterpene group.

Table 1: Names and structures of 10 biogenic SOA precursors used in the chamber experiments, including 6 monoterpenes, 2 monoterpenols and 2 sesquiterpenes.

SOA precursor	**Structure**	**Molecular formula**
Linalool		$C_{10}H_{18}O$
Ocimene		$C_{10}H_{16}$
α-pinene		$C_{10}H_{16}$
α-terpineol		$C_{10}H_{18}O$
α-cedrene		$C_{15}H_{24}$
Sabinene		$C_{10}H_{16}$

Terpinolene		$C_{10}H_{16}$
p-Cymene		$C_{10}H_{14}$
β-caryophyllene		$C_{15}H_{24}$
3-Carene		$C_{10}H_{16}$

Biogenic SOA from the photo-oxidation of each BVOC precursor was formed during a series of atmospheric simulations in the York Aerosol Chamber, which is a continuous flow reactor (CFR).[54]

Figure 13: The 0.3 m^3 continuous-flow reactor (CFR) setup, based in York.

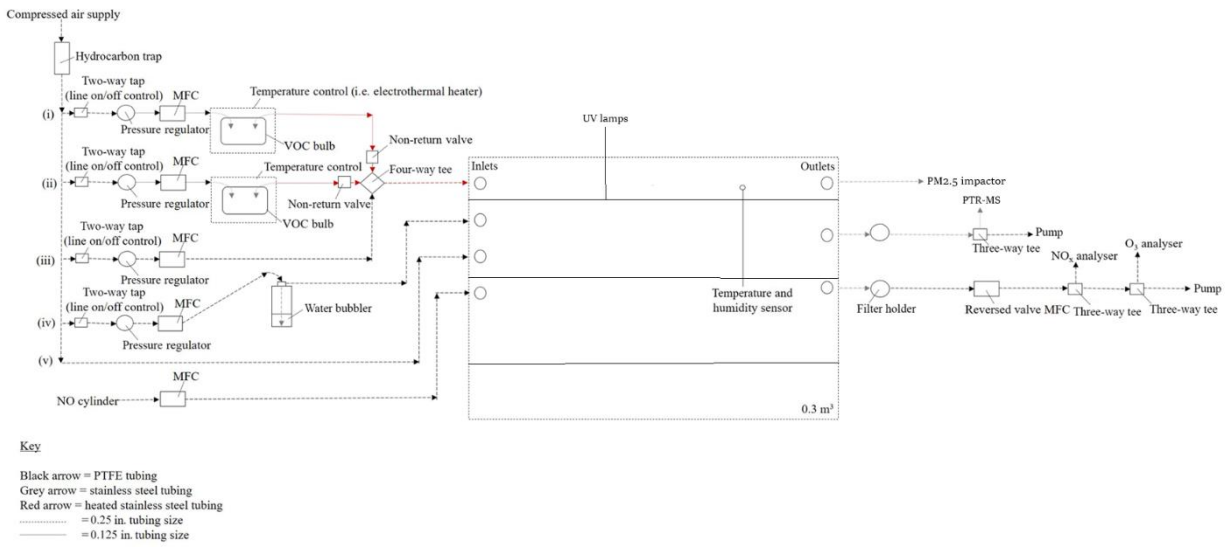


Figure 14: Detailed schematic of the continuous-flow reactor (CFR), showing connections into the chamber, the chamber itself and connections coming out of the chamber. MFC is a mass flow controller, and PTR-MS is a proton transfer reaction mass spectrometer.[54]

The CFR consists of a 0.3 m^3 custom-made polyvinyl fluoride (PVF) sampling bag with the following dimensions: 1.5 m (L) ×1.0 m (W) ×0.2 m (H) (Adtech, Gloucester, UK). The sampling bag is enclosed within a rectangular aluminium housing and supported by a 0.75 in stainless steel hanging rail mounted inside the enclosure.[54]

The CFR has four inlets, three outlets and two septum seals. The two septum seals allow a coupled temperature and humidity sensor to be mounted inside the sampling bag. UV radiation is generated using UVC Germicidal 17W (T8) lamps (UV Light Technology, Suffolk, UK), which emit at 254 nm (primary energy) and 185 nm, forming O_3 and hydroxyl radicals inside the reactor. Approximately 60 cm of 0.0625 in. stainless steel tubing was wrapped around the UV pen ray in a cone shape, preventing the UV lamp from touching the sampling bag without blocking the light source.

Humid air, O_3 and the SOA precursors were continuously injected for 4-7 hours. Table 2 below shows the exact lengths of time for each experiment, starting from when the VOC was injected.

Table 2: Start and end times, along with the durations of each experiment. The end time for the β-caryophyllene experiment was not recorded, therefore, its total duration is unknown.

SOA precursor	**Start time**	**End time**	**Duration of experiment (hours)**
Linalool	10:16	15:45	5.29
Ocimene	9:25	16:00	6.75
α-pinene	11:19	15:45	4.26
α-terpineol	9:18	15:52	6.34
α-cedrene	9:20	16:15	6.95
Sabinene	9:40	15:30	5.90
Terpinolene	10:08	16:00	5.92
p-Cymene	9:40	16:15	6.75
β-caryophyllene	10:04	-	-
3-Carene	9:58	15:00	5.42

The overall flow through the chamber was ~10-12 L/min with mass flow controllers used to balance the inflow of the reactor, ensuring the reactor volume remains constant using an in-house-built program in DAQFactory linked to a TE HTM2500LF sensor (Future Electronics, Surrey, UK) operated via a U3 LabJack (LabJack, Colorado, USA). O_3 was generated by passing air through an ozone generator (UVP, Fisher Scientific) at a flow rate of 2 L/min, and the remaining flow consisted of humid air, generated by flowing air through a water bubbler. VOCs were introduced to the chamber by injection into a heated bulb at a constant 10 μL/hr rate, with a 0.1 L/min flow of humid air.

The experimental protocol started by turning the UV lamps on and injecting humid air and O_3 into the CFR until the O_3 that was being measured using an ozone analyser (49i, Thermo Scientific) stabilised around 400-600 ppb, after which the VOC was continuously injected. Generated SOA was collected on PTFE filters at a flow rate of 10 L/min for analysis. Reactor temperatures and humidity were measured using a TE HTM2500LF sensor (Future Electronics, Surrey, UK) operated via a U3 LabJack (LabJack, Colorado, USA). Table 3 below shows the masses of SOA deposited onto the filters for each precursor.

Table 3: Initial and final filter masses (mg), as well as the mass differences, which represent the mass of SOA deposited onto the filters. Masses were weighed three times, and averages were taken. The initial filter mass for the terpinolene experiment was not recorded, therefore, the mass difference is unknown.

SOA precursor	**Initial filter mass (mg)**	**Final filter mass (mg)**	**Mass difference (mg)**
Linalool	116.71	116.83	0.12
Ocimene	137.71	138.43	0.72
α-pinene	121.70	122.53	0.83
α-terpineol	122.15	123.15	1.00
α-cedrene	137.61	140.33	2.72
Sabinene	149.93	150.40	0.47
Terpinolene	-	132.68	-
p-Cymene	140.90	141.36	0.46
β-caryophyllene	131.80	132.91	1.11
3-Carene	124.12	125.08	0.96

Due to the use of a CFR and the high concentrations used, SOA yields were unlikely to be representative of the real atmosphere and therefore were not calculated.

2.2. Beijing Filter Collection

$PM_{2.5}$ samples were collected between 18/05/2017-24/06/2017 at the Institute of Atmospheric Physics (IAP) in Beijing, China, situated between the 3rd and 4th ring roads in the northern part of the city.[55]

Figure 15: A map of Beijing, highlighting the location of the Institute of Atmospheric Physics (IAP), along with a zoomed-out view showing its position within East China.[56]

This sampling was part of the Sources and Emissions of Air Pollutants in Beijing (AIRPOLL-Beijing) project, as part of the wider Atmospheric Pollution and Human Health in a Chinese Megacity (APHH-Beijing) programme. $PM_{2.5}$ filter samples (Whatman PTFE TE35 filter discs, 0.2 µm pore size) were collected using an ECOTECH HiVol 3000 (Ecotech, Australia) high-volume air sampler with a selective $PM_{2.5}$ inlet, with a flow rate of 1.33 m^3 min^{-1}. Filters were pre-baked at 500 °C for five hours before use. Samples were collected at a height of 8 m on top of a building in the IAP complex, and were collected roughly every 3 hours during the day, at approximately between 08:30 and 17:30, with one sample being collected overnight between 17:30 and 08:30. On two highly polluted days

($PM_{2.5}$ > 100 µg m^{-3}) 20/06/2017 and 21/06/2017, hourly daytime samples were collected.[57] After collection, samples were wrapped in foil and then stored at -20 °C and shipped to York for offline analysis.[55]

2.3. Filter Extraction

98 filters were extracted, covering the sampling period of 11/06/2017-25/06/2017. Using a 6.2 cm by 6.2 cm square (38.44 cm^2) standard cutter, an aliquot of the filter was taken and cut into roughly 1 cm^2 pieces and stored in vials at -20 °C before extraction. The cutting was carried out on a piece of clean foil using pre-washed scissors and tweezers to reduce contamination. A volume of 8 mL of LC-MS grade methanol (Optima, Fisher Scientific, UK) was then added to the samples and sonicated for 45 minutes. Ice packs were used to keep the bath temperature below room temperature, with the water swapped mid-way through.

Using a 2 mL syringe, the extract was pushed through 0.22 µm filters (Millipore, Germany) into another sample vial. An additional 2 mL of methanol was added to further wash the filter, syringe filtered and added to the initial extraction to give a combined extract. These extracts were then evaporated to dryness using a vacuum solvent evaporator (Biotage, Sweden). The dry samples were reconstituted in 1 mL of 1:1 MeOH:H_2O to yield clear solutions for offline chemical analysis.

For the 10 biogenic SOA chamber-generated filter samples, the whole filter was stored in vials at -20 °C before extraction. 8 mL of LC-MS grade MeOH was then added to the samples, followed by the samples being sonicated for 45 minutes using cooling as described previously. Using a 5 mL syringe, the solutions were then pushed through 0.22 µm filters (Millipore, Germany) into clean sample vials. An additional 2 mL of methanol was added to further wash the filter, syringe filtered and added to the initial extraction to give a combined extract. These extracts were then reduced to dryness using a vacuum solvent evaporator (Biotage, Sweden). The dry samples were then reconstituted in 0.5 mL 9:1 H_2O:MeOH to yield clear solutions for offline chemical analysis.

The final solution of α-cedrene SOA was cloudy, suggesting the components had not fully dissolved. An additional 0.5 mL of 9:1 H_2O:MeOH was added, followed by 10 minutes of sonication. This was repeated once more with another 0.5 mL of solvent and a further 10 minutes of sonication, resulting in a clear solution.

A filter blank was also collected and extracted using the same method as the 10 BSOA samples. This was done to remove impurities introduced during the extraction process, such as contaminants from glassware.

2.4. UHPLC-MS2

The filter extracts were analysed using UHPLC-MS2, using an Ultimate 3000 UHPLC (Thermo Scientific, USA) coupled to a Q-Exactive Orbitrap MS (Thermo Fisher Scientific, USA) with a heated electrospray ionisation (HESI). The UHPLC method uses a polar end-capped reverse phase 5 µm, 4.6 x 100 mm, Accucore column (Thermo Scientific, UK) held at 40 °C. The mobile phase consists of LC-MS grade water (Optima, Fisher Scientific, UK) and 100 % MeOH (Fisher Chemical, USA). The water was acidified using 0.1 % formic acid to improve peak resolution, and the injection volume was 4 µL. Gradient elution was used, starting at 90 % (A) with a 1-minute post-injection hold, decreasing to 10 % (A) at 26 minutes, returning to the starting mobile phase conditions at 28 minutes, followed by a 2-minute hold allowing the re-equilibration of the column. The mass spectrometer was operated in negative mode using full scan data-dependent (dd) MS2. The scan range was set between *m/z* 50 - 750, with a mass resolution of 70,000. The ESI voltage was 4 kV, with capillary and auxiliary gas temperatures of 320 °C. The number of most abundant precursors for MS2 fragmentation per scan was set to 10. The MS was frequently mass calibrated using ESI negative ion calibration solution (Pierce, Thermo Scientific).[57]

2.5. Auxiliary measurements in Beijing

Additional gas-phase measurements were collected at the Beijing site from an elevated inlet at 8 m by the University of York.[55] Data included Nitrogen oxide, NO, measured by chemiluminescence with a Thermo Scientific Model 42i NOx analyser and Nitrogen dioxide, NO_2, was measured using a Teledyne Model T500U Cavity Attenuated Phase Shift (CAPS) spectrometer. O_3 was measured using a Thermo Scientific Model 49i UV photometric analyser. All instruments were calibrated throughout the measurement period, with a 'zero' or 'background' calibration using a Sofnofil/charcoal trap. Span (high concentration) calibrations were carried out using gas standards. Both the Thermo Scientific 42i and 42C instrument calibrations are traceable to the National Physical Laboratories (NPL)

NO scale. The meteorological variables of wind speed, wind direction, relative humidity (RH), and temperature were measured at 102 m on the IAP 325 m meteorological tower.

Aerosol mass spectrometry (AMS) measurements were undertaken by the University of Manchester.[58] The size-resolved non-refractory submicron aerosol species at the same site in Beijing were measured by an Aerodyne high-resolution time-of-flight aerosol mass spectrometer (HR-ToFAMS) at a time resolution of 5 minutes. The elemental ratios of hydrogen-to-carbon (H:C) and oxygen-to-carbon (O:C) of OA were determined, and the sources of OA were analysed with positive matrix factorisation. Three oxidised OA factors were identified with increasing degrees of oxidation: OOA1 (O:C = 0.53), OOA2 (O:C = 0.74), OOA3 (O:C = 1.18).

2.6 Data analysis

A non-target analysis workflow was developed in MZmine for the generation of the BSOA libraries. A separate workflow was used for spectral library processing. The flow diagram seen in Figure 16 outlines both workflows, along with key parameters applied at each step.

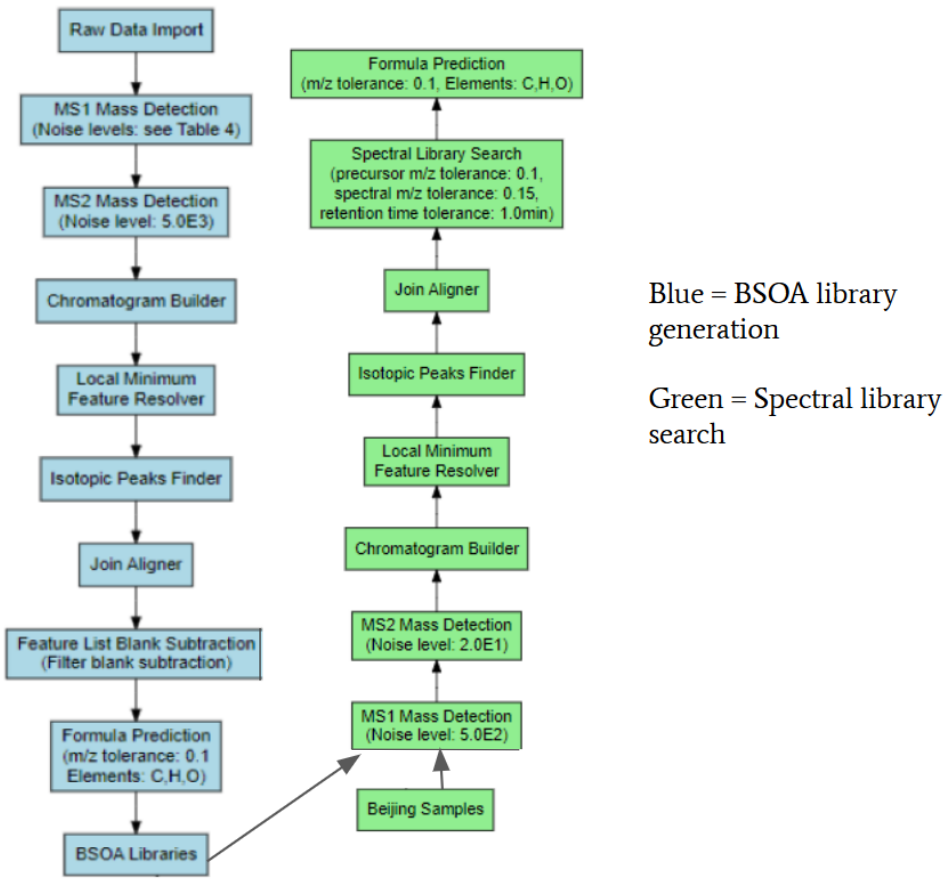


Figure 16: MZmine workflows for BSOA library generation (blue) and the spectral library search (green), with key parameters indicated for each step.

Following the spectral library search, the data was imported into RStudio for further analysis. This step was carried out to transform the dataset into a more interpretable format, enabling the separation of isomers based on differences in retention time. This allowed for accurate quantification of distinct compounds and identification of their presence across specific samples.

RStudio was also used to generate the visual figures presented throughout this paper. All graphs, box plots, and pie charts were created using the ggplot function. Polar plots were produced using the polarPlot function, and the correlation matrix was generated using the corrplot function.

2.7. Generation of BSOA mass spectral libraries-workflow

The Orbitrap MS data from the 10 biogenic filter samples and the filter blank were converted from .RAW files to .mzML files using MSconvert (ProteoWizard), to ensure

compatibility with the software used. Feature detection and library generation was carried out using MZmine 4.3.0. The .mzML data files were loaded into the software as MS data files. A workflow was implemented with MS1 and MS2 noise levels of 5.0E4 and 5.0E3, respectively, where MS1 refers to the initial mass spectrum of the precursor ions and MS2 refers to the fragmentation spectra used for structural identification, formula prediction using an *m/z* tolerance of 0.001, and the composition limited to C, H and O. Based on the fact that the precursors are all compounds consisting solely of C, H and O, it is likely that products formed via homogenous nucleation will contain only these compounds, with any exceptions occurring due to impurities.

Initially, a subtraction method was applied to each biogenic SOA filter sample to remove any compounds also detected in any of the 9 other biogenic filter samples, as well as a blank subtraction removing compounds detected in the blank filter. This removed potential background contaminants introduced by carrying out consecutive chamber experiments and filter extractions. However, due to the structural similarity of the 8 monoterpeneoids examined, it is likely that many undergo similar oxidation pathways, resulting in overlapping oxidation and fragmentation products. The applied subtraction is likely to have removed genuine SOA-derived signals, therefore acting as a limitation when analysing structurally similar products.

To address this limitation, the subtraction method was removed, and only a blank subtraction of the compounds found in the filter blank was carried out. This resulted in a drastic increase in compounds detected, many of which were low-intensity tracers which had no predicted molecular formula match and were removed. However, the number of compounds was still large, and generating libraries with such a high number of compounds would be too time-consuming. To combat this and focus on relevant BSOA tracers, the MS1 noise thresholds were adjusted in the workflow of each biogenic filter sample to retain approximately the top 100 most abundant tracers.

The MS1 noise levels were determined by plotting the peak area against the cumulative number of tracers detected, as shown in Figure 17 for linalool SOA. Tracers were first ranked from most to least abundant by their MS1 signal intensity, therefore, each point on the curve corresponds to a total number of tracers with an area above a given threshold. In this analysis, a cutoff was applied at the point where 100 tracers are retained, giving the MS1 noise level. MS1 noise levels were determined for all 10 biogenic libraries using plots like Figure 17, the resulting values are presented in Table 4 below.

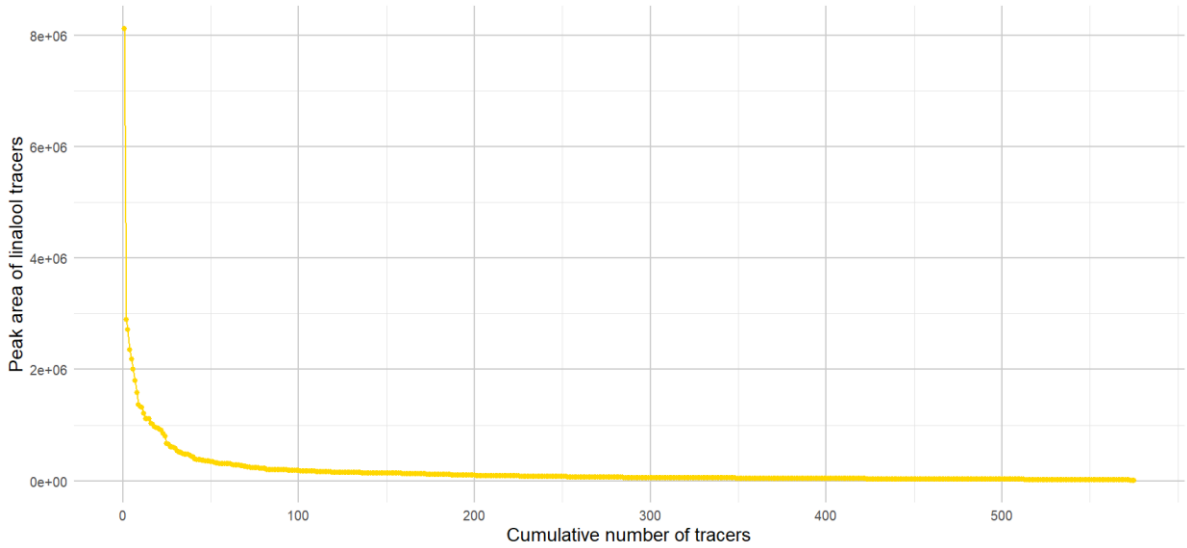


Figure 17: A plot of peak area against the cumulative number of tracers detected in the linalool library, showing that most of the total peak area was from the 100 most abundant tracers.

Table 4: MS1 noise levels for each individual biogenic library and the corresponding number of tracers retained.

SOA precursor	**MS1 noise level**	**Number of tracers in library**
Linalool	2.7E5	98
Ocimene	1.9E5	99
α-pinene	4.5E5	100
α-terpineol	8.8E5	99
α-cedrene	5.9E5	102
Sabinene	4.6E5	104

Terpinolene	4.0E5	101
p-Cymene	5.0E5	100
β-caryophyllene	8.6E5	99
3-Carene	4.0E5	103

To use the libraries for spectral library searching, conversion to a .mgf format, which retains both MS2 peak data and associated molecular formula information, is required. This was achieved using the 'Annotate Manually' function in MZmine, which enabled annotation of key data, including adduct type, compound name, molecular formula, precursor *m/z*, and retention time (RT). The 'Spectral Library Batch Generation' feature was then used to export .mgf files containing only annotated compounds, allowing them to be used as custom spectral libraries within MZmine.

2.8. Biogenic SOA tracer compounds in Beijing

Raw data from the 98 Beijing filter samples were converted from a.RAW to a .mzML format using MSConvert (ProteoWizard) to ensure software compatibility, and they were subsequently loaded into MZmine as MS data files. A data processing workflow was set up with an MS1 noise level of 5.0E2 and an MS2 noise level of 2.0E1. The 'Spectral Library Search' function was implemented with a precursor *m/z* tolerance of 0.1 Da, a spectral *m/z* tolerance of 0.15 Da, and a retention time tolerance of 1.0 min. This workflow generated an aligned feature list, which was exported as a .CSV file for further analysis.

The .CSV file was imported into RStudio for processing and analysis. Due to variability between sample durations and procedures, a two-step normalisation process was implemented. Samples were collected over differing sampling durations, with most being collected every three hours during the day and a single overnight sample, and hourly sampling being conducted on 20/06/2017 and 21/06/2017. To account for this, all samples were normalised to the volume of air sampled, providing concentrations in peak area per cubic metre of air (peak area/m^3). This allowed for comparison of approximate concentrations across all timepoints. It should be noted that ionisation efficiency is highly structure-

dependent, and calibration was not performed due to the lack of commercially available standards.

Secondly, during filter extraction, the majority of samples were reconstituted in 1 mL of a 1:1 MeOH:H_2O solution, however, some samples were diluted to 2 mL or 3 mL due to varying solubilities. To correct for this, an additional normalisation step was carried out to adjust concentration values according to the final volume, ensuring consistency across all samples.

After normalisation, the 98 Beijing samples, originally labelled by Sample ID, were matched to their corresponding dates and times. A 'midpoint time' column was then created to provide the most representative values for plotting. Figure 18 shows a schematic representation of the overall methodology used.

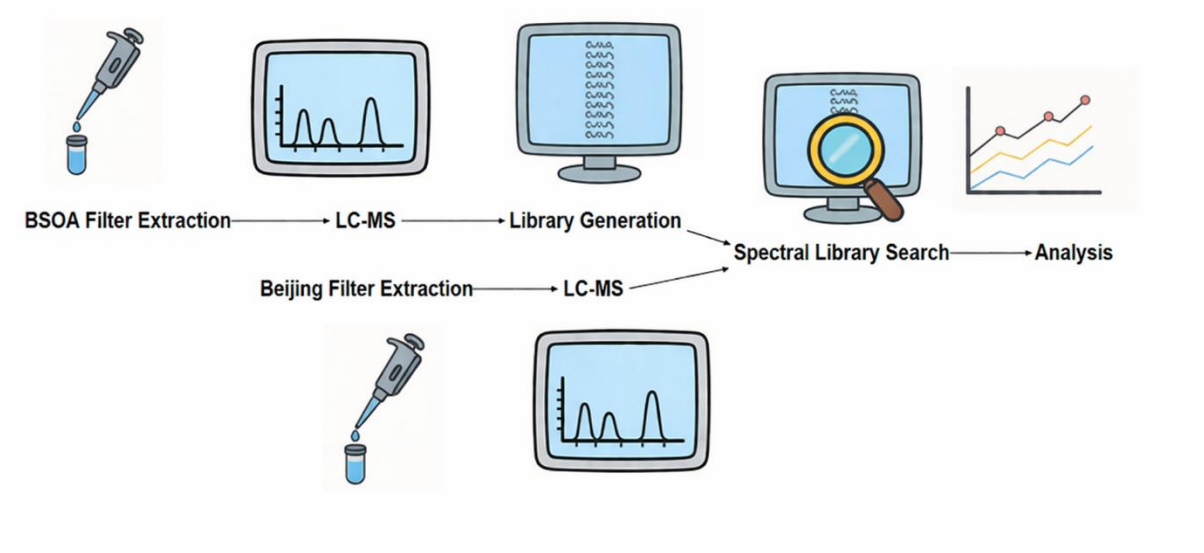


Figure 18: Flowchart showing the workflow for LC-MS-based BSOA spectral library building and matching to Beijing filter samples.[65]

3. Results and Discussion

3.1.1. Composition of chamber SOA

To address the question of how individual BVOC precursors influence SOA composition and dominant oxidation products under controlled conditions, the chemical composition of the SOA produced in each chamber experiment was examined. Both O_3 and OH radicals are present in the chamber, so oxidation products are likely to be formed from a mix of both mechanisms. NO_x was not added to the chamber to simplify the chemistry, and so only a subset of potential atmospheric reaction products will be formed. However, during the

sampling period in Beijing, NO levels were suppressed as a result of O_3 titration and other unknown pathways, resulting in the formation of biogenic oxidation products that are normally associated with low-NO pathways.[59] For each BSOA library, the 20 most abundant tracers, ranked by peak area, were identified to highlight dominant compounds as well as potential marker species in order to start to deduce potential SOA formation pathways. A full analysis of formation pathways is beyond the scope of this work, but here we summarise some general trends observed in the BSOA tracers, as shown in Tables 5-15.

3.1.2. Monoterpenes

Tables 5-10 below show the top 20 most abundant tracers found in each of the 6 monoterpene tracer tables.

Table 5: The 20 most abundant SOA tracers in the terpinolene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

Terpinolene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	8.16	185.0823	1.2E7
$C_8H_{12}O_5$	0.63	2.96	187.0616	6.8E6
$C_8H_{12}O_4$	0.50	4.79	171.0668	3.5E6
$C_7H_{10}O_2$	0.29	2.96	125.0612	3.3E6
$C_7H_{10}O_5$	0.71	1.69	173.0459	3.1E6
$C_{21}H_{36}O_{10}$	0.48	20.70	447.2240	2.8E6
$C_{16}H_{26}O_6$	0.38	22.47	313.1647	2.7E6
$C_{17}H_{28}O_6$	0.35	23.31	327.1803	2.6E6
$C_8H_{12}O_4$	0.50	6.69	171.0668	2.5E6
$C_{21}H_{36}O_{10}$	0.48	19.51	447.2239	2.4E6

$C_{16}H_{26}O_{14}$	0.88	22.47	441.1267	2.1E6
$C_{35}H_{20}O$	0.03	23.31	455.1425	2.0E6
$C_{10}H_{14}O_6$	0.60	4.51	229.0720	1.9E6
$C_{17}H_{26}O_4$	0.24	19.47	293.1762	1.8E6
$C_{10}H_{16}O_4$	0.40	6.62	199.0980	1.7E6
$C_{10}H_{16}O_5$	0.50	10.16	215.0928	1.7E6

Table 6: The 20 most abundant SOA tracers in the sabinene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

Sabinene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	11.36	185.0823	7.0E6
$C_8H_{14}O_5$	0.63	6.20	189.0773	4.7E6
$C_8H_{12}O_4$	0.50	4.83	171.0668	4.4E6
$C_{16}H_{26}O_6$	0.38	22.51	313.1647	3.1E6
$C_{10}H_{18}O_5$	0.50	7.54	217.1085	3.1E6
$C_{32}H_{22}O_{21}$	0.66	22.13	741.0575	2.4E6
$C_{17}H_{28}O_6$	0.35	23.31	327.1803	2.2E6
$C_{35}H_{20}O$	0.03	23.31	455.1425	2.2E6
$C_{34}H_{18}O$	0.03	22.51	441.1268	2.2E6
$C_9H_{16}O_4$	0.44	8.38	187.0980	2.1E6
$C_{18}H_{30}O_6$	0.33	24.07	341.1959	2.1E6
$C_{19}H_{30}O_5$	0.26	21.58	337.2025	1.8E6

$C_8H_{14}O_6$	0.75	2.86	205.0721	1.7E6
$C_{15}H_{24}O_3$	0.20	18.68	251.1656	1.7E6
$C_8H_{12}O_5$	0.63	3.00	187.0616	1.6E6
$C_9H_{14}O_3$	0.33	10.49	169.0875	1.5E6
$C_{36}H_{22}O$	0.03	24.07	469.1580	1.4E6
$C_8H_{12}O_4$	0.50	6.24	171.0668	1.3E6
$C_8H_{14}O_2$	0.25	11.36	141.0926	1.2E6
$C_{15}H_{24}O_6$	0.40	21.61	299.1490	1.1E6

Table 7: The 20 most abundant SOA tracers in the α-pinene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

α-pinene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	8.31	185.0822	3.5E7
$C_8H_{12}O_4$	0.50	4.94	171.0667	1.3E7
$C_{10}H_{16}O_4$	0.40	6.45	199.0980	8.8E6
$C_9H_{14}O_3$	0.33	8.62	169.0875	5.3E6
$C_{17}H_{26}O_8$	0.47	14.27	357.1560	5.1E6
$C_{10}H_{16}O_5$	0.50	7.36	215.0928	3.3E6
$C_{16}H_{26}O_6$	0.38	22.57	313.1647	3.1E6
$C_{17}H_{28}O_6$	0.35	23.37	327.1803	3.0E6
$C_{10}H_{16}O_5$	0.50	9.25	215.0928	2.7E6
$C_{32}H_{22}O_{21}$	0.66	22.19	41.0577	2.6E6

$C_{16}H_{26}O_{14}$	0.88	22.57	441.1267	2.3E6
$C_8H_{14}O_2$	0.25	8.31	141.0926	2.3E6
$C_{18}H_{30}O_8$	0.44	14.30	373.1871	2.2E6
$C_{35}H_{20}O$	0.03	23.37	455.1425	2.2E6
$C_{10}H_{16}O_4$	0.40	6.73	199.0981	2.1E6
$C_9H_{14}O_3$	0.33	16.82	169.0876	2.0E6
$C_{10}H_{16}O_4$	0.40	7.08	199.0981	2.0E6
$C_8H_{12}O_4$	0.50	6.83	171.0668	1.9E6
$C_{19}H_{28}O_7$	0.37	16.22	367.1766	1.8E6
$C_9H_{14}O_5$	0.56	6.62	201.0774	1.8E6

Table 8: The 20 most abundant SOA tracers in the p-cymene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

***p*-Cymene SOA tracer**	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	6.51	185.0823	1.8E7
$C_8H_{12}O_4$	0.50	4.89	171.0668	7.3E6
$C_9H_{14}O_4$	0.44	5.49	185.0823	5.1E6
$C_7H_{12}O_2$	0.29	4.72	127.0769	3.5E6
$C_7H_{10}O_4$	0.57	6.68	157.0511	3.0E6
$C_{15}H_{24}O_3$	0.20	18.76	251.1656	3.0E6
$C_8H_{12}O_5$	0.63	3.03	187.0617	2.6E6
$C_{16}H_{26}O_6$	0.38	22.52	313.1648	2.5E6

$C_{17}H_{28}O_6$	0.35	23.36	327.1803	2.4E6
$C_9H_{14}O_4$	0.44	10.02	185.0824	2.3E6
$C_{32}H_{22}O_{21}$	0.66	22.14	741.0577	2.3E6
$C_8H_{12}O_5$	0.63	3.28	187.0616	2.1E6
$C_7H_{10}O_5$	0.71	1.59	173.0459	2.1E6
$C_{34}H_{18}O$	0.03	22.52	441.1269	1.9E6
$C_{35}H_{20}O$	0.03	23.36	455.1426	1.8E6
$C_9H_{14}O_5$	0.55	3.63	201.0773	1.5E6
$C_8H_{14}O_2$	0.25	6.51	141.0926	1.4E6
$C_{10}H_{16}O_4$	0.40	5.98	199.0980	1.3E6
$C_{10}H_{18}O_4$	0.40	6.47	201.1136	1.3E6
$C_{14}H_{22}O_4$	0.29	14.67	253.1449	1.3E6

Table 9: The 20 most abundant SOA tracers in the ocimene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

Ocimene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_{32}H_{22}O_{21}$	0.66	22.09	741.0575	2.5E6
$C_{17}H_{28}O_6$	0.35	23.34	327.1803	2.2E6
$C_{16}H_{26}O_6$	0.38	22.50	313.1647	2.2E6
$C_8H_{12}O_4$	0.5	4.77	171.0668	1.8E6
$C_{34}H_{18}O$	0.03	22.50	441.1268	1.8E6
$C_{35}H_{20}O$	0.03	23.30	455.1425	1.7E6

$C_9H_{14}O_4$	0.44	6.39	185.0823	1.6E6
$C_{18}H_{30}O_6$	0.33	24.07	341.1959	1.5E6
$C_9H_{14}O_3$	0.33	15.52	169.0875	1.4E6
$C_{15}H_{12}O_4$	0.27	18.04	255.0667	1.2E6
$C_{20}H_{22}O_7$	0.35	10.29	373.1296	1.1E6
$C_{18}H_{30}O_{14}$	0.78	24.07	469.1580	1.1E6
$C_9H_{14}O_4$	0.44	11.31	185.0823	1.1E6
$C_{18}H_{34}O_2$	0.11	25.05	281.2489	8.9E5
$C_{15}H_{24}O_6$	0.40	21.60	299.1490	8.4E6
$C_{23}H_{24}O_5$	0.22	22.50	379.1563	8.4E5
$C_{20}H_{22}O_7$	0.35	10.57	373.1296	8.3E5
$C_{15}H_{24}O_{14}$	0.93	21.63	427.1111	8.2E5
$C_{24}H_{26}O_5$	0.21	23.30	393.1721	8.0E5
$C_{29}H_{30}O_8$	0.28	18.49	505.1874	7.1E5

Table 10: The 20 most abundant SOA tracers in the 3-carene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

3-Carene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	9.99	185.0822	7.0E7
$C_{10}H_{16}O_4$	0.40	8.38	199.0979	1.6E7
$C_{19}H_{30}O_5$	0.26	17.07	337.2025	6.8E6
$C_{10}H_{16}O_3$	0.30	11.26	183.1030	6.2E6

$C_8H_{14}O_2$	0.25	9.99	141.0925	4.8E6
$C_9H_{14}O_3$	0.33	10.31	169.0875	3.8E6
$C_{12}H_{26}O_{15}$	1.25	9.99	409.1182	3.8E6
$C_{16}H_{26}O_6$	0.38	22.55	313.1647	3.4E6
$C_{17}H_{28}O_6$	0.35	23.35	327.1803	3.2E6
$C_9H_{14}O_4$	0.44	10.80	185.0823	3.1E6
$C_{34}H_{18}O$	0.03	22.52	441.1268	2.5E6
$C_{35}H_{20}O$	0.03	23.35	455.1426	2.4E6
$C_{32}H_{22}O_{21}$	0.66	22.13	741.0576	2.1E6
$C_{18}H_{30}O_6$	0.33	24.11	341.1959	2.1E6
$C_{19}H_{28}O_7$	0.37	17.25	367.1764	2.1E6
$C_{19}H_{30}O_5$	0.26	21.06	337.2025	1.9E6
$C_7H_{10}O_4$	0.57	3.78	157.0510	1.6E6
$C_{19}H_{30}O_5$	0.26	20.11	337.2025	1.5E6
$C_{18}H_{30}O_{14}$	0.78	24.11	469.1580	1.5E6
$C_5H_8O_4$	0.8	4.69	131.0354	1.4E6

The most abundant BSOA tracers for cyclic monoterpene VOCs in Tables 5-8 (terpinolene, sabinene, α-pinene, *p*-cymene) are all dominated by C_7-C_{10} compounds. Given that the parent VOCs are all C_{10}, the presence of these products indicates fragmentation and functionalisation during oxidation.

Most O:C ratios fall between 0.30 and 0.70, suggesting moderate functionalisation and possibly multiple oxidation steps. Sabinene is bicyclic with a single, strained exocyclic double bond, while terpinolene is monocyclic but contains multiple double bonds (one exocyclic and one endocyclic), likely contributing to the formation of more highly oxidised

products by providing additional reactive sites for attack. α-pinene has a single bond but also has a strained bicyclic structure, also facilitating attack, whereas *p*-cymene has its double bond within an aromatic ring, therefore making it a less reactive double bond, producing fractionally lower O:C products.

The variation in retention times (RTs) for several repeated molecular formulas strongly suggests the presence of isomeric species. $C_8H_{12}O_4$ exhibits RT values which cluster around 4.7-5.0 min and 6.2-6.9 min, suggesting isomerism due to functionality differences. $C_8H_{12}O_5$ typically appears near 2.9-3.3 min, except for sabinene, where it elutes at 6.20 min. The α-pinene table shows several well-established marker compounds for α-pinene such as pinonic acid ($C_{10}H_{16}O_3$) and norpinonic acid ($C_{10}H_{16}O_4$).[60,61] The molecular formula of norpinonic acid ($C_{10}H_{16}O_4$) appears 3 times in the α-pinene table, suggesting formation of 3 isomers.

Unlike Tables 5-8, the most abundant species in Tables 9 and 10 are not C_7-C_{10} compounds, with the exception of a few observed in Table 10, and formation of larger compounds is observed. Both ocimene and 3-carene also exhibit a pathway that produces extremely oxidised products, with ocimene producing $C_{15}H_{24}O_{14}$ and 3-carene producing $C_{12}H_{26}O_{15}$. Their high O:C ratios indicate auto-oxidation and accretion reactions, resulting in highly polar, low-volatility products with increased partitioning into the particle phase.

Ocimene is an acyclic triene, therefore, it is highly reactive due to its 3 C=C bonds, meaning it is able to form these highly oxidised oligomers, and there is a lot of potential for fragmentation. Low peak areas were observed and are likely a result of the numerous possible sites for oxidant attack, leading to a wide array of products and no single dominant pathway. 3-Carene, on the other hand, is bicyclic with a single double bond and has a strained ring system, which are features it shares with multiple compounds in Tables 5-8, mainly α-pinene.

There is a significant number of overlapping SOA tracers across the six precursor tables. Table 11 lists the compounds that appear in four or more of the systems.

Table 11. SOA tracers detected in three or more SOA tracer tables. The table lists molecular formulas, the number of tables in which each tracer appears, and the tables where the tracer is observed.

Tracer Formula	**Number of BVOCs that produce it**	**Tables it is present in**
$C_9H_{14}O_4$	6	terpinolene, sabinene, α-pinene, *p*-cymene, ocimene, 3-carene
$C_{16}H_{26}O_6$	6	terpinolene, sabinene, α-pinene, *p*-cymene, ocimene, 3-carene
$C_{17}H_{28}O_6$	6	terpinolene, sabinene, α-pinene, *p*-cymene, ocimene, 3-carene
$C_{35}H_{20}O$	6	terpinolene, sabinene, α-pinene, *p*-cymene, ocimene, 3-carene
$C_{32}H_{22}O_{21}$	5	sabinene, α-pinene, *p*-cymene, ocimene, 3-carene
$C_8H_{12}O_4$	5	terpinolene, sabinene, α-pinene, *p*-cymene, ocimene
$C_{34}H_{18}O$	4	sabinene, *p*-cymene, ocimene, 3-carene
$C_9H_{14}O_3$	4	sabinene, α-pinene, ocimene, 3-carene
$C_{10}H_{16}O_4$	4	terpinolene, α-pinene, *p*-cymene, 3-carene
$C_8H_{14}O_2$	4	sabinene, α-pinene, *p*-cymene, 3-carene

Notably, tracers such as $C_9H_{14}O_4$, $C_{16}H_{26}O_6$ and $C_{17}H_{28}O_6$ appear in the SOA of all six BVOCs, suggesting they are formed through common oxidation pathways. As well as this, the repeated presence of oligomeric species (e.g. $C_{32}H_{22}O_{21}$) indicates that accretion reactions also play a major role in SOA composition, regardless of precursor structure.

Compounds $C_{35}H_{20}O$ and $C_{34}H_{18}O$ have very low O numbers and do not conform to known oxidation pathways. It is possible that these species are contaminants that are not removed by the subtraction step.

3.1.3. Monoterpenols

Tables 12-13 below show the top 20 most abundant tracers found in the 2 monoterpenol tracer tables.

Table 12: The 20 most abundant SOA tracers in the linalool library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

Linalool SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_{15}H_{24}O_3$	0.20	18.67	251.1655	8.1E6
$C_{16}H_{26}O_6$	0.38	22.51	313.1646	3.5E6
$C_{17}H_{28}O_6$	0.35	23.34	327.1803	3.3E6
$C_{14}H_{22}O_4$	0.29	14.57	253.1449	2.9E6
$C_{16}H_{26}O_{14}$	0.88	22.51	441.1267	2.7E6
$C_{35}H_{20}O$	0.03	23.34	455.1425	2.4E6
$C_{18}H_{30}O_6$	0.33	24.07	341.1959	2.3E6
$C_{14}H_{22}O_3$	0.21	17.31	237.1499	2.2E6
$C_{32}H_{22}O_{21}$	0.66	22.09	741.0575	2.1E6
$C_{13}H_{20}O_3$	0.23	16.71	223.1344	2.0E6
$C_{10}H_{20}O_3$	0.30	18.08	187.1344	1.8E6
$C_{18}H_{30}O_{14}$	0.78	24.11	469.1580	1.5E6
$C_{15}H_{24}O_6$	0.40	21.57	299.1489	1.3E6

$C_{15}H_{24}O_{14}$	0.93	21.57	427.1110	1.3E6
$C_{23}H_{24}O_5$	0.22	22.51	379.1562	1.3E6
$C_{24}H_{26}O_5$	0.21	23.31	393.1721	1.2E6
$C_{10}H_{20}O_3$	0.30	17.83	187.1344	1.1E6
$C_{11}H_{26}O_{14}$	1.28	18.67	381.1235	1.1E6
$C_{16}H_{26}O_6$	0.38	22.82	313.1647	1.1E6
$C_{14}H_{22}O_4$	0.29	18.04	253.1448	1.0E6

Table 13: The 20 most abundant SOA tracers in the α-terpineol library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

α-terpineol SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_9H_{14}O_4$	0.44	6.37	185.0822	8.5E7
$C_9H_{14}O_4$	0.44	5.35	185.0822	3.7E7
$C_8H_{12}O_4$	0.50	4.75	171.0667	1.6E7
$C_8H_{12}O_5$	0.63	2.93	187.0616	1.1E7
$C_8H_{14}O_4$	0.50	4.26	173.0823	6.6E6
$C_8H_{14}O_2$	0.25	6.37	141.0925	5.8E6
$C_7H_{10}O_2$	0.29	2.93	125.0612	4.8E6
$C_{20}H_{34}O_8$	0.4	14.00	401.2183	4.7E6
$C_{15}H_{24}O_3$	0.2	18.68	251.1656	4.3E6
$C_{10}H_{16}O_6$	0.6	8.51	231.0877	4.1E6
$C_9H_{14}O_5$	0.56	3.49	201.0773	3.9E6

$C_{10}H_{18}O_4$	0.4	6.33	201.1136	3.5E6
$C_9H_{16}O_4$	0.44	6.16	187.0980	3.1E6
$C_{19}H_{28}O_7$	0.37	12.88	367.1764	2.7E6
$C_{18}H_{28}O_7$	0.39	12.92	355.1766	2.6E6
$C_{16}H_{26}O_6$	0.38	22.49	313.1647	2.4E6
$C_{21}H_{36}O_8$	0.38	15.74	415.2341	2.4E6
$C_{18}H_{28}O_8$	0.44	11.21	371.1714	2.4E6
$C_{17}H_{28}O_6$	0.35	23.32	327.1803	2.2E6
$C_{14}H_{22}O_4$	0.29	14.59	253.1449	2.2E6

The most abundant tracers identified in the linalool ($C_{10}H_{18}O$) tracer table were compounds with carbon chain lengths greater than that of the parent molecule. The product distribution ranged from mid-sized molecules (C_{14}-C_{20}) to high-mass ones up to C_{35}, suggesting the involvement of multiple formation pathways. The other monoterpenol, α-terpineol, produced smaller, often monomeric compounds such as $C_9H_{14}O_4$ and $C_8H_{12}O_4$, pointing towards dominant fragmentation processes.

A broad range of O:C ratios was detected amongst the linalool products, demonstrating varying degrees of oxidation. The number of oxygen atoms ranges from 1 to 21, indicating the formation of oligomer species from the combination of multiple highly oxidised organic molecules (HOMs). In previous chamber studies, HOMs formation from linalool oxidation has been predicted, with autoxidation leading to the formation of first-generation oxidation products with up to 8 O atoms.[62] α-terpineol tracers are mostly clustered within a moderate O:C range of 0.2 to 0.6, suggesting that the SOA composition is dominated by monomeric oxidation products with less extensive oligomer formation. Several SOA tracers overlap between linalool and α-terpineol, such as $C_{16}H_{26}O_6$ and $C_{17}H_{28}O_6$, indicating common oxidation pathways despite their structural differences.

These trends in tracer production are linked to precursor structure, particularly the number and position of unsaturated bonds. Linalool's flexible, acyclic structure and multiple double bonds facilitate extensive oxidation and the formation of highly oxygenated products.

In contrast, α-terpineol's monocyclic structure and single endocyclic double bond limit the extent of oxidation compared to linalool.

3.1.4. Impact of structure on composition

The differences in O:C ratios and carbon numbers within the tracers highlight how precursor structure influences the extent of oxidation and oligomerisation. Figure 19 below presents a bar chart comparing the number of SOA tracers in each carbon number category ($<C_{10}$, C_{10}, C_{11}-C_{20}, and $>C_{20}$) across eight monoterpenoid precursors, allowing for easier comparison of their tracer distributions.

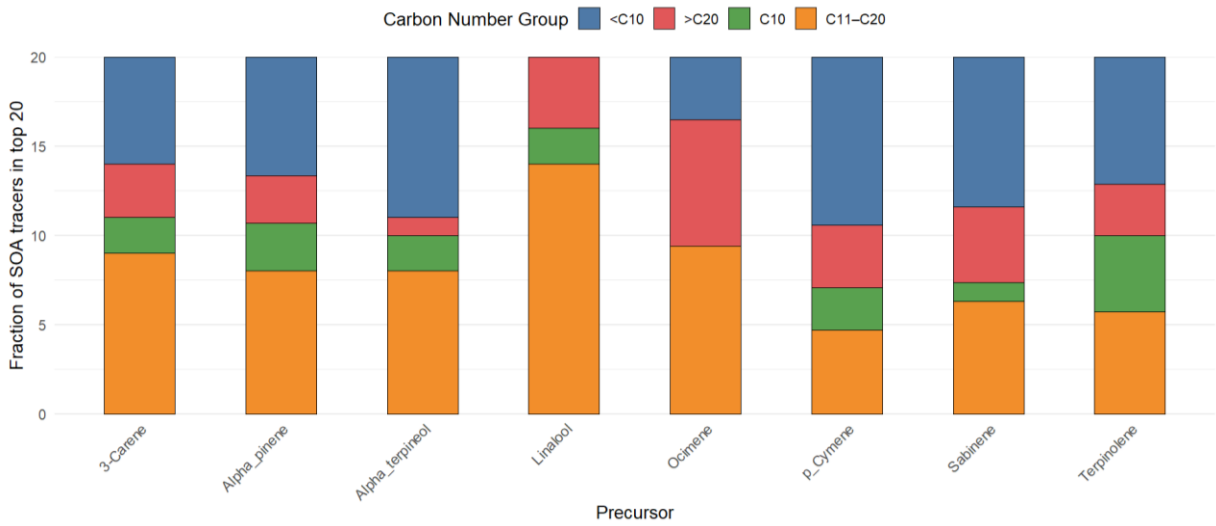


Figure 19: Carbon number distribution of SOA tracers formed from different VOC precursors. Each bar represents the relative fraction of the number of tracers grouped by carbon number: <C10, C10, C11-C20, and >C20. The y-axis maximum is set to 20, reflecting the number of compounds listed in the tables of the 20 most abundant tracers.

Figure 19 highlights how SOA tracer carbon number distributions vary by precursor. While most compounds produce a mix of all carbon number groups, linalool and ocimene show distinct trends, with linalool forming mainly C_{11}–C_{20} tracers with no $<C_{10}$ compounds, and ocimene lacking C_{10} products entirely. These differences reflect the influence of structural features such as ring strain and unsaturation on SOA formation pathways.

3.1.5. Sesquiterpenes

Tables 14-15 below show the top 20 most abundant tracers found in the 2 sesquiterpene tracer tables.

Table 14: The 20 most abundant SOA tracers in the β-caryophyllene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

β-caryophyllene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_{14}H_{22}O_4$	0.29	14.64	253.1448	4.0E7
$C_{15}H_{24}O_3$	0.20	18.72	251.1655	3.7E7
$C_{14}H_{22}O_3$	0.21	17.36	237.1499	1.7E7
$C_{16}H_{28}O_6$	0.38	16.63	315.1815	1.3E7
$C_{13}H_{20}O_3$	0.23	16.80	223.1343	9.6E6
$C_{16}H_{30}O_7$	0.44	12.93	333.1923	5.1E6
$C_{15}H_{24}O_4$	0.27	16.17	267.1604	5.0E6
$C_{16}H_{28}O_6$	0.38	16.38	315.1815	4.3E6
$C_{14}H_{22}O_4$	0.29	13.17	253.1558	4.3E6
$C_{15}H_{26}O_6$	0.40	15.93	301.1660	4.0E6
$C_{14}H_{22}O_5$	0.36	13.87	269.1397	3.2E6
$C_{15}H_{26}O_5$	0.33	13.35	285.1710	3.0E6
$C_{12}H_{18}O_5$	0.45	11.88	241.1085	2.9E6
$C_{16}H_{26}O_6$	0.38	22.52	313.1647	2.7E6
$C_{14}H_{22}O_2$	0.14	16.42	221.1550	2.6E6
$C_9H_{14}O_4$	0.44	10.00	185.0823	2.6E6
$C_{15}H_{24}O_4$	0.27	16.73	267.1604	2.4E6
$C_{14}H_{22}O_5$	0.36	14.26	269.1398	2.4E6

$C_{14}H_{24}O_5$	0.36	14.26	271.1554	2.3E6
$C_{17}H_{28}O_6$	0.35	23.35	327.1803	2.3E6

Table 15: The 20 most abundant SOA tracers in the α-cedrene library ranked by peak area. Their O:C ratios, retention times (RT) and m/z values are also provided.

α-cedrene SOA tracer	**O:C ratio**	**RT (min)**	***m/z***	**Peak area**
$C_{15}H_{24}O_3$	0.20	18.73	251.1655	1.3E8
$C_{14}H_{22}O_4$	0.29	18.14	253.1448	9.9E7
$C_{15}H_{24}O_4$	0.27	16.19	267.1604	8.3E7
$C_{14}H_{22}O_3$	0.21	18.25	237.1499	2.5E7
$C_{13}H_{20}O_3$	0.23	16.82	223.1343	2.3E7
$C_{14}H_{22}O_4$	0.28	13.22	253.1449	1.8E7
$C_{14}H_{22}O_3$	0.21	17.41	237.1499	1.2E7
$C_{30}H_{48}O_8$	0.27	16.19	535.3280	1.2E7
$C_{30}H_{48}O_6$	0.20	18.73	503.3382	1.0E7
$C_9H_{14}O_4$	0.44	10.04	185.0823	7.7E6
$C_{13}H_{20}O_4$	0.31	16.75	239.1292	6.8E6
$C_{15}H_{24}O_4$	0.27	14.65	267.1605	6.7E6
$C_{13}H_{22}O_3$	0.23	16.82	225.1498	6.2E6
$C_{14}H_{22}O_4$	0.29	12.21	253.1449	5.4E6
$C_{13}H_{20}O_3$	0.23	17.45	223.1344	4.7E6
$C_{15}H_{22}O_5$	0.33	14.76	281.1398	4.3E6

$C_{13}H_{22}O_4$	0.31	17.48	241.1449	3.7E6
$C_{14}H_{22}O_4$	0.29	14.30	253.1448	3.7E6
$C_{15}H_{24}O_4$	0.27	16.47	267.1605	3.4E6
$C_{14}H_{20}O_4$	0.29	14.34	251.1293	3.3E6

Likely due to them being sesquiterpenes, β-caryophyllene and α-cedrene (both $C_{15}H_{24}$) also show similarities in their SOA tracer production. Most detected compounds have carbon numbers similar to, or slightly lower than those of the parent molecule, suggesting minor fragmentation likely through cleavage of terminal or branched carbon atoms during oxidation. β-caryophyllene oxidation products have previously been shown to be dominated by compounds which retain similar carbon numbers as the parent molecule, such as β-caryophyllene oxide.[63] Most O:C ratios fall between 0.20 and 0.40, suggesting moderate functionalisation and few oxidation steps rather than extensive oxidation. Several tracers, including $C_{15}H_{24}O_3$, $C_{14}H_{22}O_4$, $C_{14}H_{22}O_3$, and $C_{13}H_{20}O_3$, appear in both β-caryophyllene and α-cedrene SOA, indicating similar oxidation pathways.

Several differences between Tables 4 and 5 are also evident, likely arising from differing structures of the precursors. In the case of β-caryophyllene, no compounds exceed 20 carbon atoms, implying limited oligomerisation. The presence of $C_9H_{14}O_4$, a tracer with six fewer carbon atoms than the parent, indicates significant fragmentation can occur, as seen in previous studies.[64] In contrast, α-cedrene forms higher-mass compounds such as $C_{30}H_{48}O_6$ and $C_{30}H_{48}O_8$, indicating some degree of oligomer formation.

These trends in tracer production are linked to the precursor structures, as β-caryophyllene is a bicyclic sesquiterpene with two double bonds, while α-cedrene is tricyclic, more constrained and only has a single double bond. α-cedrene shows a higher proportion of high mass tracers, suggesting that its rigid structure may favour oligomerisation or limit fragmentation.

3.2. Spectral library search results

3.2.1. Overview of BSOA tracers identified in Beijing

To address the question of how different BVOC precursors contribute to the molecular composition and size distribution of SOA in ambient Beijing samples, a total of 311 SOA tracers from the ten spectral libraries, each containing the 100 most abundant compounds, were matched to peaks in the Beijing samples during the spectral library search. In order to get a better idea of potential formation pathways of these precursors in Beijing, they were split into 3 groups. Those with 1-10 carbon atoms, those with 11-20 and those with 20+. Table 13 summarises the number of tracers matched to each SOA precursor library, and the number of tracers that fit into each size group.

Table 16: The number of SOA tracer compounds matched to each monoterpenoid precursor library, along with the number of tracers containing 1-10 C atoms, 11-20 C atoms and 20+ C atoms.

Monoterpenoid precursor	**Total no. of SOA tracers matched to Beijing samples**	**No. of C_1-C_{10} tracers**	**No. of C_{11}-C_{20} tracers**	**No. of C_{20}+ tracers**
Linalool	20	3	12	5
Ocimene	43	20	18	5
α-pinene	28	24	3	1
α-terpineol	28	19	8	1
Sabinene	17	10	3	4
Terpinolene	27	17	5	5
p-Cymene	31	28	3	0
3-Carene	25	16	8	1

The ocimene library was matched to the highest number of tracers, suggesting it may have a strong contribution to SOA formation. It was followed by *p*-cymene, with other notable contributors being α-pinene, α-terpineol, and terpinolene. In contrast, linalool, 3-carene, and sabinene were associated with fewer tracers, implying a more limited impact on SOA composition in this environment.

3.2.2. C_1-C_{10} tracers: Monoterpenoid SOA

Within the 8 C_{10} precursor libraries, a total of 137 C_1-C_{10} SOA tracers were detected in the Beijing samples. The most dominant contributions to this low carbon number fraction came from *p*-cymene (28), α-pinene (24) and ocimene (20). For ocimene, however, the products are evenly distributed across the size ranges, indicating a more balanced distribution between formation pathways.

In contrast, *p*-cymene and α-pinene exhibit a noticeable dominance towards production of these low-carbon-number compounds, with over 90 % of their total matched tracers falling into this category. Significantly, the α-pinene matches were dominated by C_{10} compounds (15), including several well-established marker compounds such as pinonic acid ($C_{10}H_{16}O_3$) and norpinonic acid ($C_{10}H_{16}O_4$).[6,61]

Although 3-carene is not among the three most dominant contributors in terms of tracer count, it exhibits a strong bias toward low carbon number compounds, with 16 out of 25 matched tracers falling within the C_1-C_{10} range. This distribution suggests a significant contribution from fragmentation pathways, with 3-carene marker compounds cis-3-caric acid ($C_9H_{14}O_4$) and cis-3-caronic acid ($C_7H_{10}O_4$) being detected.[66]

3.2.3. C_{11}-C_{20} tracers: Monoterpenoid SOA

A total of 147 C_{11}-C_{20} tracers were matched across the eight C_{10} precursor libraries. Dominant contributions came from ocimene (18), linalool (12), α-terpineol (8), and 3-carene (8). With the exception of 3-carene, the BVOC precursors will likely form compounds in the $<C_{10}$ category during the initial oxidation step and therefore, the accretion reactions may be more important in forming lower volatility products.

3.2.4. C_{20+} tracers: Monoterpenoid SOA

A total of 22 C_{20}+ SOA tracers were matched across the eight C_{10} precursor libraries. The largest contributions to this high carbon number fraction came from ocimene, terpinolene, and linalool, each with five matched tracers. However, no single precursor exhibited a strong dominance in the formation of these larger compounds, and the overall number of C_{20}^{+} tracers was relatively low.

This suggests that oligomerisation pathways were not favoured under the atmospheric conditions in Beijing during sampling. This may be explained by a combination of factors, in particular, the much lower levels of BVOC in Beijing compared to the chamber experiments, which reduces the potential for accretion reactions.[67] Additionally, there may be cross-VOC reactions taking place which were not considered in this study and could influence product distribution.

3.2.5. Sesquiterpene SOA tracers

Table 17: The number of SOA tracer compounds matched to each sesquiterpene precursor library, along with the number of tracers containing 1-15 C atoms, 16-30 C atoms and 30+ C atoms.

Sesquiterpene precursor	**Total no. of SOA tracers matched to Beijing samples**	**No. of C_1-C_{15} tracers**	**No. of C_{16}-C_{30} tracers**	**No. of C_{30}+ tracers**
β-caryophyllene	48	45	3	0
α-cedrene	44	34	5	5

Both of the C_{15} sesquiterpenes in the table above, β-caryophyllene and α-cedrene, were matched to more tracers than any of the monoterpenoids presented in Table 16. The majority of matched SOA tracers (79 in total) fall within the C_1-C_{15} range, indicating that fragmentation processes dominate. A smaller number of tracers (8) were observed in the C_{16}-C_{30} range, suggesting limited functionalisation. Only five tracers were identified in the C_{30^+} range, implying that the formation of high molecular weight oligomers from these precursors was minimal under the atmospheric conditions present in Beijing.

3.2.6. Limitations of spectral library search

If the chamber method accurately reflects real-world conditions, substantial overlap between the 20 most abundant species in the chamber experiments and those most abundant in the Beijing samples should be observed. To assess this, the number of compounds from each precursor SOA tracer table (Tables 5-10, 12-15) that were also detected in the Beijing samples was examined. The results are summarised in the table below.

Table 18: The number of SOA tracer compounds matched to the Beijing samples and the number of those appearing among the 20 most abundant compounds, along with their corresponding molecular formulas.

SOA precursor (abbreviation, full)	**Total no. of SOA tracers matched to Beijing samples**	**Number of spectral matches to top 20 compounds**	**Formulas of matched compounds**
Linalool	20	7	$C_{14}H_{22}O_4$, $C_{14}H_{22}O_4$, $C_{16}H_{26}O_{14}$, $C_{10}H_{20}O_3$, $C_{10}H_{20}O_3$, $C_{23}H_{24}O_5$, $C_{24}H_{26}O_5$
Ocimene	43	4	$C_9H_{14}O_4$, $C_9H_{14}O_4$, $C_{18}H_{34}O_2$, $C_{24}H_{26}O_5$
α-pinene	28	5	$C_8H_{12}O_4$, $C_{10}H_{16}O_4$, $C_{10}H_{16}O_5$, $C_{10}H_{16}O_5$, $C_{32}H_{22}O_{21}$
α-terpineol	28	7	$C_9H_{14}O_4$, $C_9H_{14}O_4$, $C_8H_{14}O_4$, $C_{10}H_{16}O_6$, $C_9H_{14}O_5$, $C_9H_{16}O_4$, $C_{16}H_{26}O_6$
α-cedrene	44	6	$C_{14}H_{22}O_4$, $C_{13}H_{20}O_3$, $C_{13}H_{20}O_3$, $C_{13}H_{20}O_4$, $C_{13}H_{22}O_3$, $C_{13}H_{22}O_4$
Sabinene	17	2	$C_8H_{14}O_6$, $C_8H_{12}O_5$
Terpinolene	27	3	$C_8H_{12}O_5$, $C_{17}H_{26}O_4$, $C_6H_{10}O_3$,

p-Cymene	31	5	$C_9H_{14}O_4$, $C_8H_{12}O_4$, $C_8H_{12}O_5$, $C_8H_{14}O_2$, $C_{10}H_{16}O_4$
β-caryophyllene	48	6	$C_{15}H_{24}O_3$, $C_{14}H_{22}O_3$, $C_{14}H_{22}O_5$, $C_{14}H_{22}O_5$, $C_{15}H_{26}O_5$, $C_{14}H_{24}O_5$
3-Carene	25	4	$C_9H_{14}O_4$, $C_{16}H_{26}O_6$, $C_{17}H_{28}O_6$, $C_7H_{10}O_4$

All the precursors tested produced some overlap with the Beijing samples, indicating that the chamber experiments do capture part of the real-world SOA composition. The matched formulas in the top 20 tables range from relatively small oxygenated species (e.g. C_6-C_{10}) to large, highly oxygenated compounds (e.g. $C_{23}H_{24}O_5$, $C_{32}H_{22}O_{21}$). Linalool and α-terpineol showed the strongest overlap, with seven matches each to compounds in their top 20 table, while α-cedrene and β-caryophyllene followed closely with six. Other precursors, such as sabinene and terpinolene, showed only two or three matches to compounds in their tables.

Low numbers of matches to compounds in the top 20 tables presented in Table 18 above suggests that, despite there being some overlap, the chamber experiments do not fully reflect the range of precursor sources and atmospheric processes shaping real-world aerosol compositions. This can be due to many reasons, the main reasons being the complexity of the real atmosphere being difficult to represent. Real-world factors such as sunlight intensity, humidity, temperature, lower concentrations and interactions with anthropogenic emissions all affect SOA chemistry and are very hard to replicate. Abundance of pre-existing seed particles also plays a role, as the SOA generated in the chamber was a result of homogeneous nucleation only, whereas the abundance of pre-existing particles in the atmosphere means gas-particle partitioning will be much more important.

As well as this, chambers usually test one precursor at a time, while the atmosphere contains a vast mixture of VOCs interacting simultaneously. This means cross-reactions, competition, and multiple oxidation pathways in the real world are not fully represented. In this case, only O_3 and OH were used, while in reality, NO_3 radicals also act as important oxidants, often competing with O_3 and OH to drive different oxidation pathways. As a result, combined radical pathways may play a significant role, with multiple sequential oxidation reactions taking place.

VOC concentrations in chambers are also often much higher than ambient levels in order to generate sufficient SOA mass for detection. Analytical constraints, such as the sensitivity of mass spectrometry, require higher concentrations to ensure that oxidation products can be reliably detected. At high concentrations, reaction pathways and partitioning can differ significantly from ambient conditions.

The large number of tracers identified in Beijing, however, does provide a unique dataset for understanding the sources and chemistry of BSOA formation in a polluted megacity. While many of the matches do not appear among the top 20 most abundant compounds, the method still captures a wide range of less abundant components, suggesting that the method is more informative than Table 18 indicates.

3.3. Tracer time series

To address the question of how SOA tracer concentrations vary over time, total tracer concentrations were assessed for each of the 98 Beijing samples collected. The tracers detected in each sample were summed, producing a time series graph of the total SOA tracer concentration expressed as peak area/m^3 covering the sampling period from 11/06/2017-25/06/2017. To get a better idea of whether the baseline concentrations remained stable, a plot with a y-axis of 0-100,000 peak area/m^3 was also generated. For both figures, the points on the line represent midpoint times of each sample and are irregular due to there being different periods of time between collection of each sample.

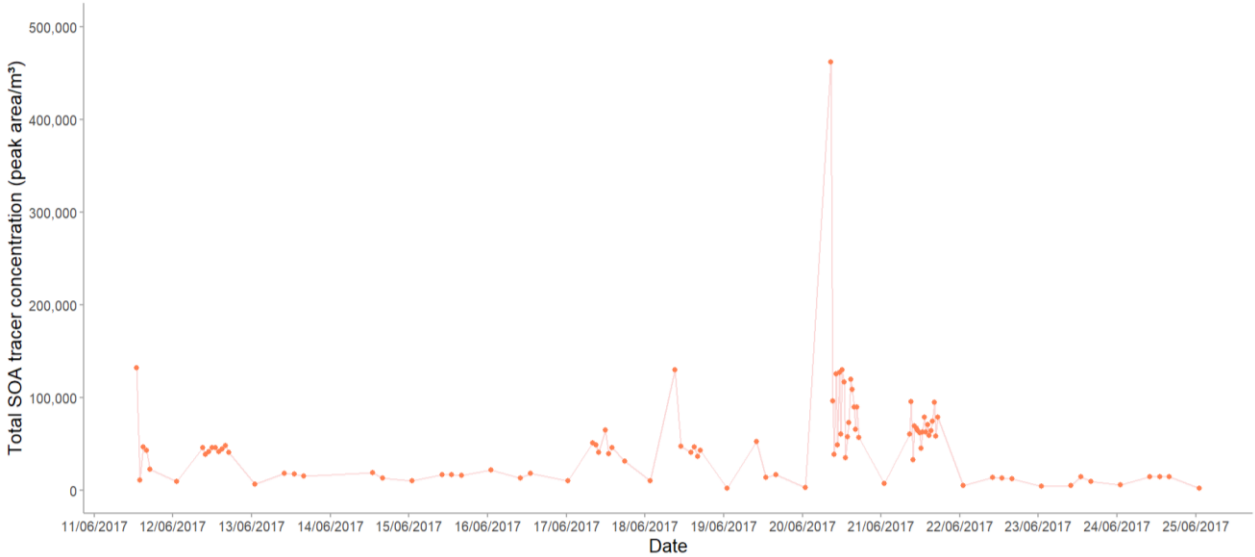


Figure 20: Total SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017, highlighting a spike on the 20th of June.

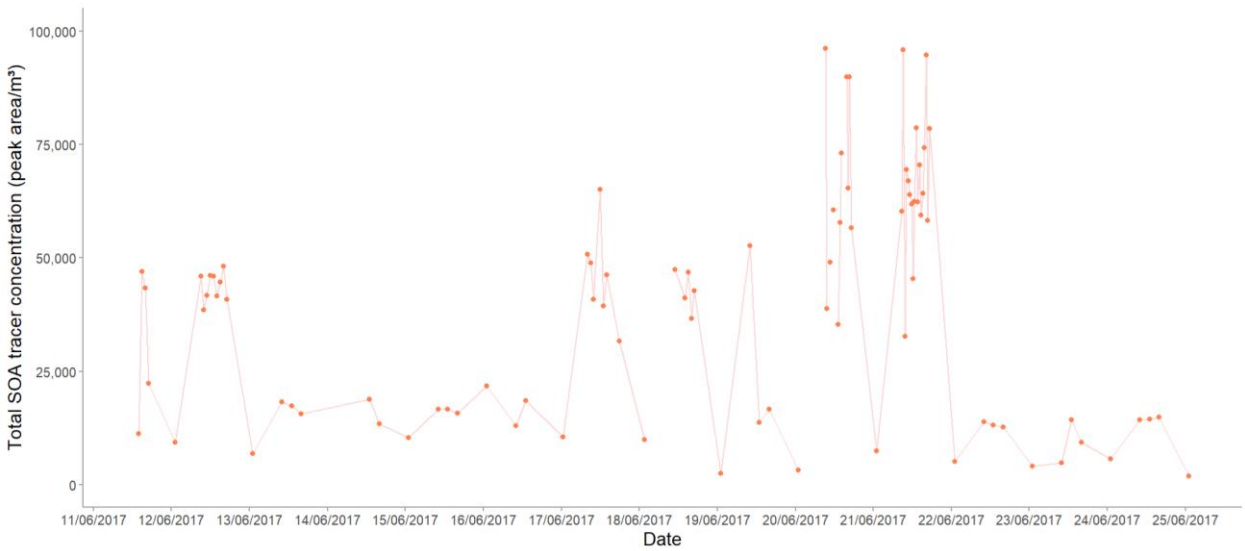


Figure 21: Total SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017. The y-axis is set from 0-100,000 to zoom in on baseline periods, and gaps in the line show points at which the y-axis values exceed 100,000.

Figure 20 gives the impression that tracer concentrations remained largely stable over the period, with only minor fluctuations and small peaks. However, Figure 21 does reveal slight diurnal trends, with higher concentrations during the day and lower concentrations at night, as well as elevated concentrations on the 12th and between the 17th-21st.

Most notably, Figure 20 shows a pronounced spike on the 20th of June, where concentrations rose to approximately five times the estimated average across the campaign. Following this event, concentrations rapidly returned to baseline.

3.4. Precursor contributions

3.4.1. Trends in BSOA tracer contributions

To address the question of how the relative contributions of different BVOC precursors to SOA formation compare, total SOA output concentrations were evaluated for each precursor library. Although Section *3.2.1. Overview of BSOA tracers identified in Beijing* showed that the β-caryophyllene library produced the highest number of matches, suggesting a potentially significant contribution to SOA formation, and sabinene was associated with the fewest matches, these findings require further examination. To resolve this, total SOA output concentrations were calculated for each precursor library by adding the concentrations of all matched tracers for each sample. This was then used to generate mean

values for each precursor library. The resulting percentages representing these values are presented in Figure 22 below.

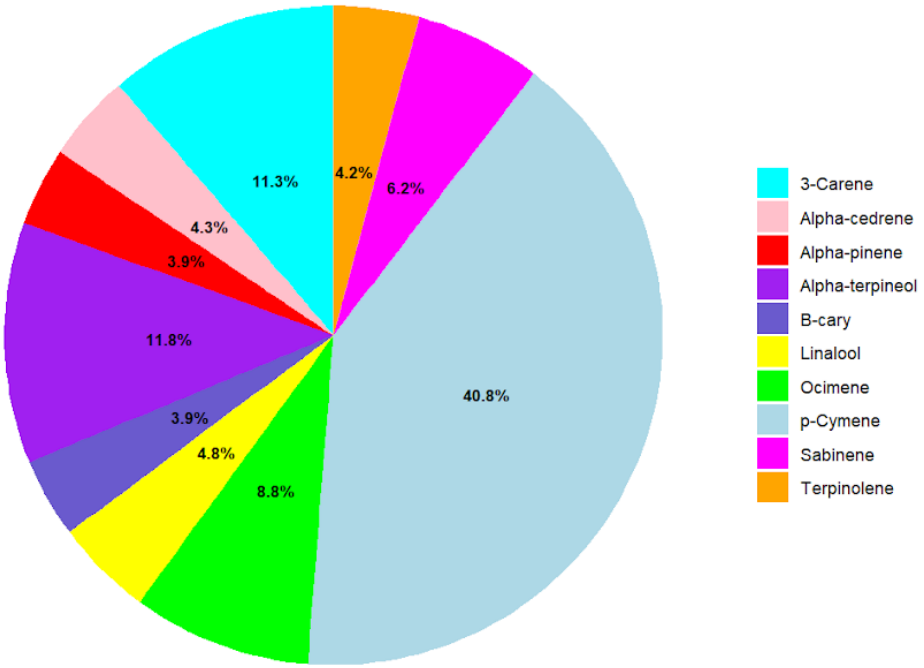


Figure 22: Average contributions of each precursor compound to SOA mass in Beijing. Percentages represent each compound's share of the total measured SOA mass.

Figure 22 indicates that the BSOA composition was dominated by *p*-cymene, representing, on average, 40.8 % of the total observed BSOA tracers. While only 31 *p*-cymene BSOA tracers were observed, they were present at high concentrations. The next highest contributions to the total BSOA peak area were α-terpineol (11.8 %) and 3-carene (11.3 %), which each contributed a much smaller share than *p*-cymene. Lowest contributions came from α-pinene (3.9 %), and more surprisingly, β-caryophyllene (3.9 %), confirming that while a large number of tracers were observed, they were present in low concentrations. This could potentially suggest the SOA produced from β-caryophyllene is extremely reactive and therefore short-lived, leading to low concentrations. Since these BVOCs are typically emitted by coniferous vegetation, it is likely that forested areas in or near Beijing are important sources.[47]

To examine how the relative contributions of each precursor changed over the study period, a stacked bar chart was generated for each filter sample, as shown in Figure 22. Each bar was normalised to 100 % to allow comparison of the proportional contribution of each BVOC precursor to the BSOA tracers observed across the 98 samples. It is important to note

that, due to the use of hourly sampling on the 20th and 21st of June, Figure 23 (plotted against Sample ID) is not perfectly aligned with Figure 20 (plotted against midpoint times).

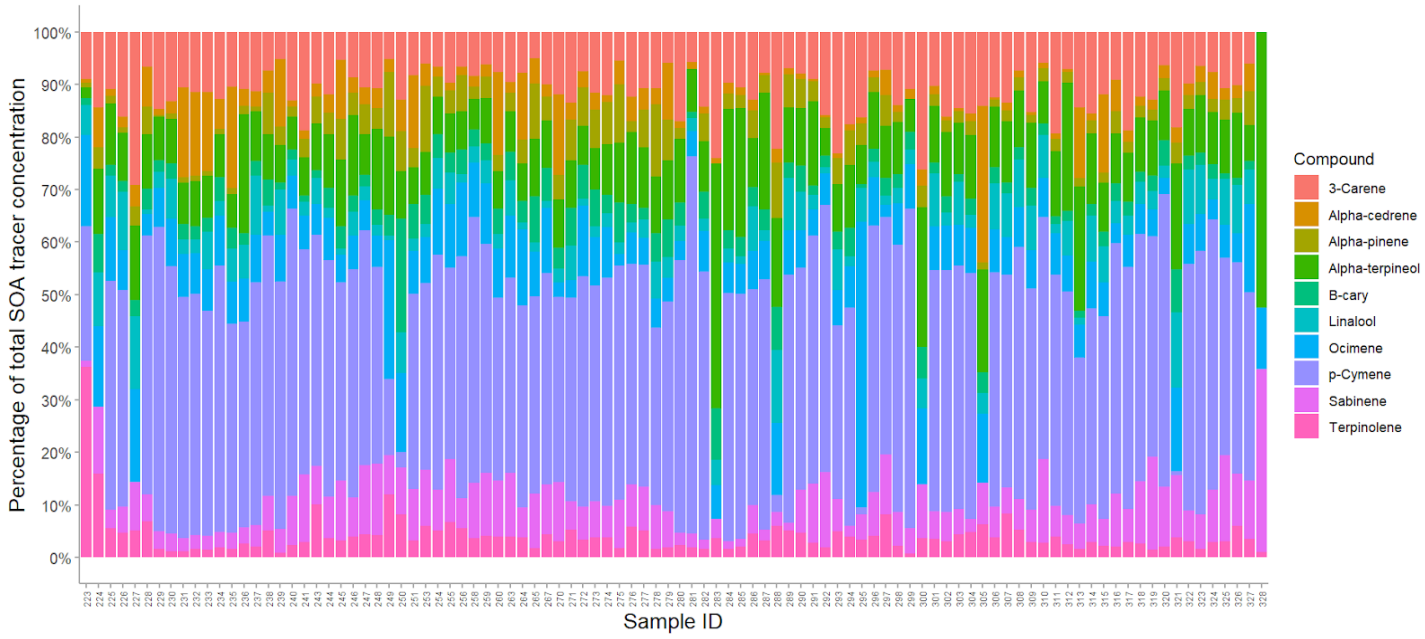


Figure 23: Normalised stacked bar chart showing changes in the contributions of each precursor to SOA mass across 98 samples in Beijing.

Figure 23 shows generally stable proportional contributions across the study period. Most samples exhibit a strong dominance of *p*-cymene, while the lower contributions of β-caryophyllene and α-pinene observed in the pie chart are consistently reflected across individual samples in the stacked bar chart.

Major differences in the BSOA composition are observed in Samples 227 (11/06/2017, 17:00), 250 (16/06/2017, 09:59), 283 (20/06/2017, 10:45), 288 (20/06/2017, 13:15), 313 (21/06/2017, 16:15), 321 (23/06/2017, 09:59) and 328 (24/06/2017, 23:59). These samples all show a distinct drop in the relative proportion of *p*-cymene BSOA tracers, in some cases down to zero. On these days, the most abundant BSOA precursor is α-terpineol. This trend is not a result of night-time emission changes, with an even split between daytime and night-time samples.

A different compositional change is observed in Sample 295, collected in the afternoon of the 20th of June, which exhibits an unusually high fraction of ocimene BSOA tracers. This is notable because other samples taken at similar times on different days do not show this pattern.

As demonstrated in Figure 20, there is a pronounced SOA concentration peak observed on the 20th of June, corresponding to Sample 279. Examination of Figure 23 indicates no substantial change in the relative contributions of precursors in this sample, suggesting that the event causing the concentration spike affected the precursors in a similar way.

3.4.2. p-Cymene BSOA in Beijing

One possible explanation for the high contribution of *p*-cymene tracers observed in Figure 22, alongside the comparatively low concentrations of other tracers, is the presence of anthropogenic sources. Unlike the other BVOCs studied, *p*-cymene is a substituted C_{10} mono-aromatic. It has previously been identified in cooking emissions, particularly during the frying of meat and herbs, and in vehicle emissions, having been detected in measurements of gases from traffic in tunnels. [68,69]

Examination of Figure 23 makes this explanation less likely, however, as samples with low *p*-cymene fractions occur irregularly with no clear association to specific times or days of the week and appear as isolated deviations rather than in consecutive groups. This suggests that if variations in anthropogenic sources were responsible, they would have to be extremely short-lived, allowing *p*-cymene tracer concentrations to return to typical levels by the time the next sample was collected. Another possible explanation is an analytical issue, such as the software failing to correctly identify the main *p*-cymene tracers. However, this seems unlikely, as the software consistently detects these tracers in other samples.

As well as this, lower concentrations of *p*-cymene tracers would be expected to coincide with lower concentrations of other anthropogenic markers, such as NO_x. However, comparison to observed mixing ratios of NO_x (shown in **Section 3.11. Comparison of BSOA trends to other pollutants**) shows no significant correlation between the drops in *p*-cymene tracers and NO_x concentrations.

3.5. Diurnal trends

3.5.1. Diurnal variability of total BSOA tracers

To better understand trends in BSOA tracer production, it is useful to examine diurnal cycles to determine whether precursor and, therefore, tracer concentrations are generally

higher at specific times of the day. For this analysis, samples were divided into four time intervals: 18:00-09:00 (representing night-time samples), 09:00-12:00, 12:00-15:00, and 15:00-18:00. Within each interval, the concentrations of all tracers in each sample were used to determine median, lower and upper quartiles, and maximum and minimum values in order to generate the box plot in Figure 24.

Sample 279 (collected overnight) was excluded from the dataset because it is an extreme outlier that affects the overall trend analysis. In addition, all zero values were removed from every box plot to prevent them from artificially lowering the minimum values. As shown in Figure 25, where the zero values and Sample 279 were not excluded, the appearance of the plot was significantly affected. Every time interval included at least one zero value, and Sample 279 stood out as an extreme outlier.

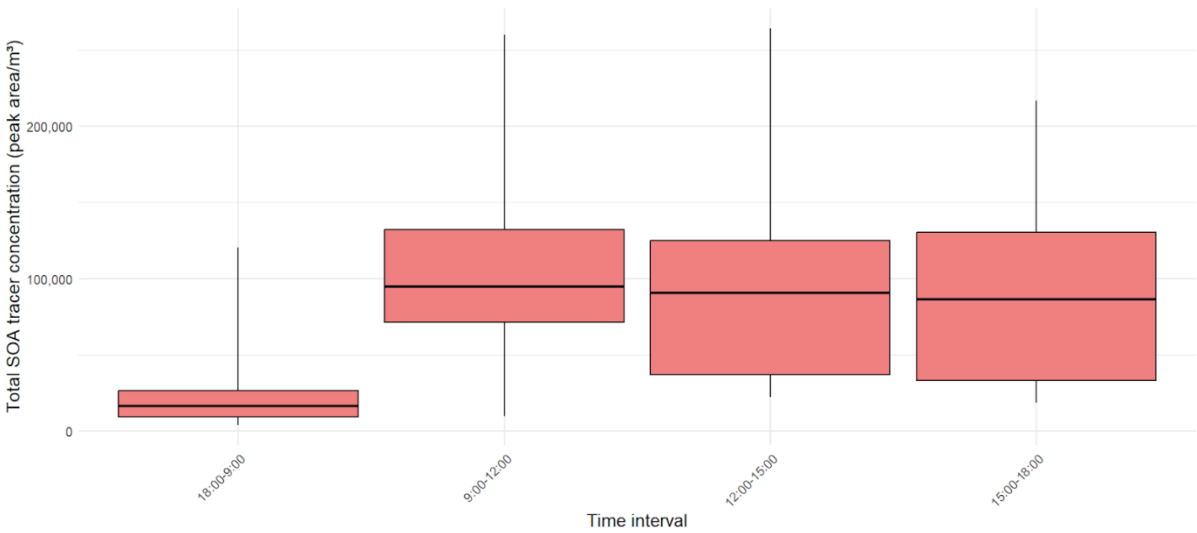


Figure 24: Diurnal variations in total SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. Sample 279 and all zero values were removed.

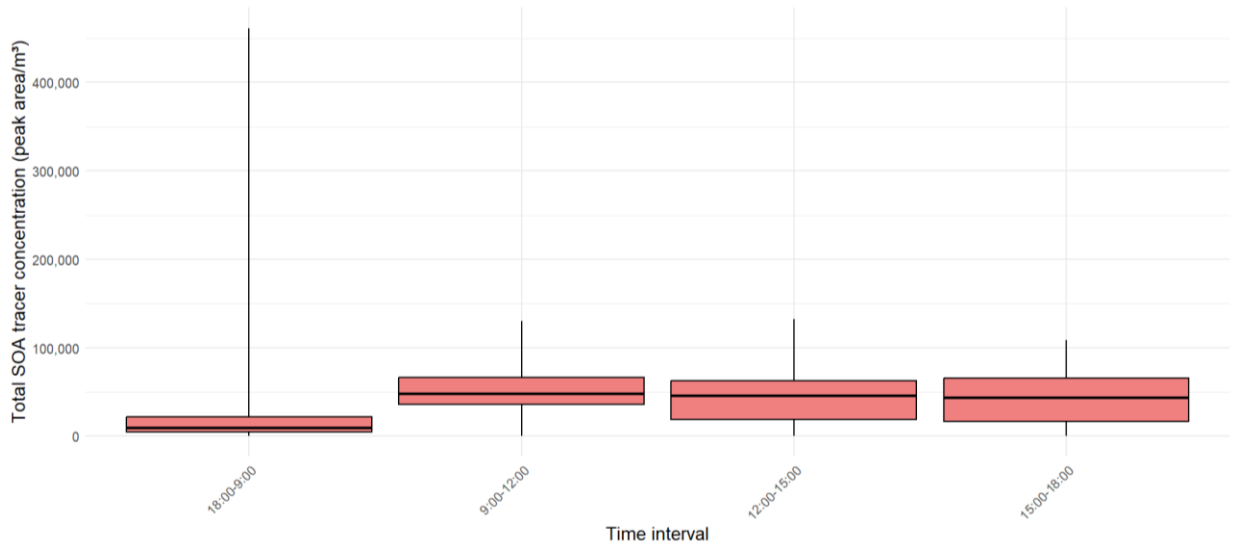


Figure 25: Diurnal variations in total SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. Sample 279 and all 0 values were retained.

Figure 24 shows that the total BSOA tracer peak area/m^3 for the night-time samples (18:00-09:00) is much lower than that of the daytime samples, with the median a factor of 10 lower than the daytime median. The narrow interquartile range in the night-time samples indicates stable, consistently low concentrations, likely as a result of both lower emissions, due to lower temperature, and reduced oxidation by OH radicals. It should be noted that NO_3 chemistry was not investigated in this study, and so the nocturnal BSOA is likely to be underestimated.

The morning samples (09:00-12:00 interval) show a sharp increase in concentrations, with the interquartile range broadening to reflect greater variability. The maximum value rises significantly, exceeding 250,000, while the minimum remains close to that of the night-time interval.

In the early afternoon samples (12:00-15:00 interval), the median, maximum, and minimum values are similar to those in the morning, but the interquartile range broadens further due to a substantial drop in the 25th percentile.

The late afternoon samples (15:00-18:00 interval) remain consistent with the previous two intervals in terms of median and minimum values, but the maximum decreases to just over 200,000. This interval shows the widest interquartile range of all, indicating the highest variability among the middle 50 % of samples. Overall, the observed median value is consistent across daytime samples.

3.5.2. Individual diurnal plots

Furthermore, individual tracer box plots were generated to identify which specific precursors contribute to the patterns observed in Figure 24. The same time intervals were used, and once again, all 0 values were excluded, as well as Sample 279 from eight of the ten datasets. The two datasets retaining this sample (ocimene and terpinolene) did not exhibit peaks in Sample 279, but ocimene exhibited a peak in Sample 295, and it was subsequently removed. The box plots below illustrate these results, starting with the five most abundant precursors, followed by the five less abundant precursors.

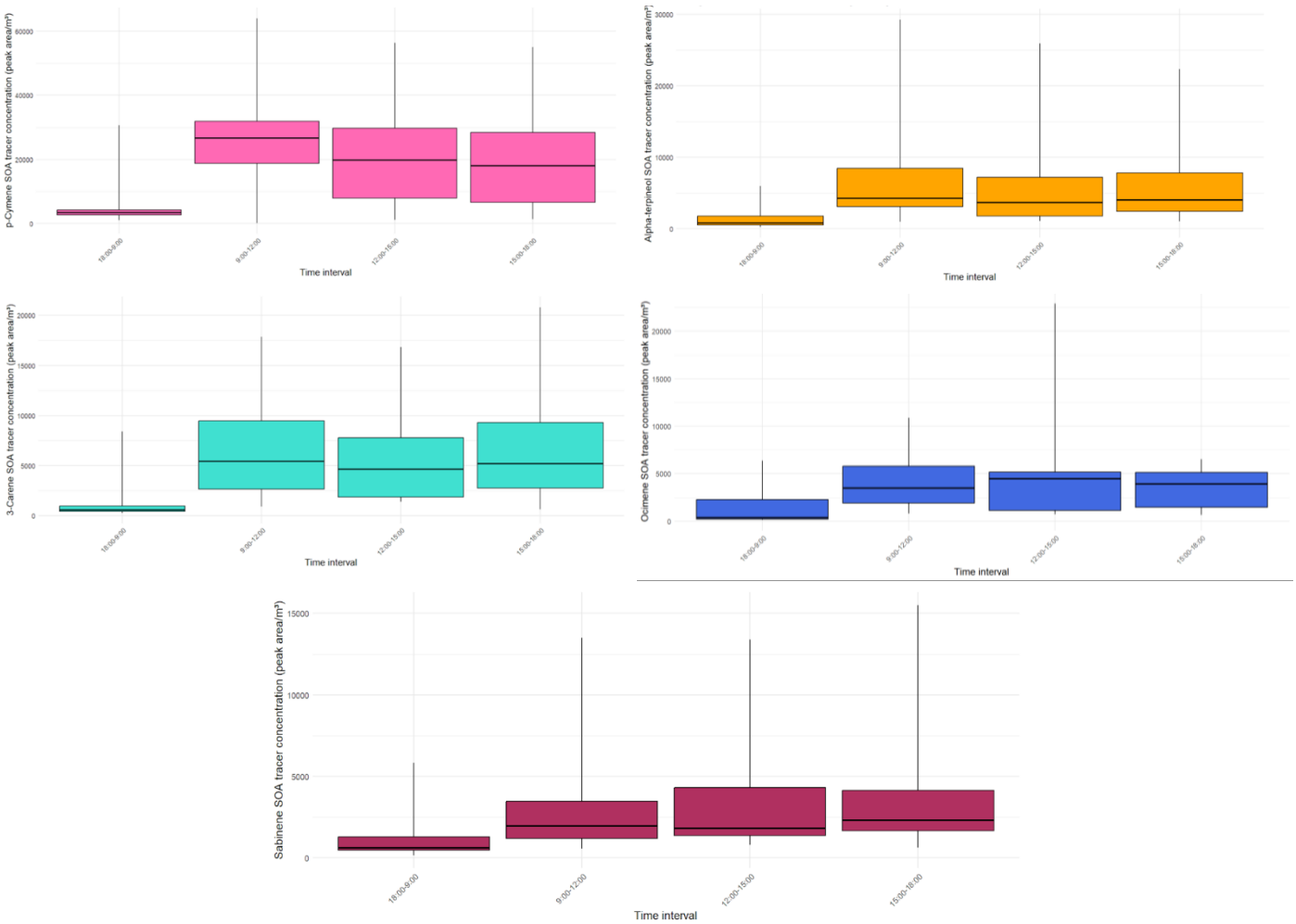


Figure 26: Diurnal variations in p-cymene (top left), α-terpineol (top right), 3-carene (middle left), ocimene (middle right) and sabinene (bottom). SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. All zero values were removed, and for p-cymene, α-terpineol, 3-carene and sabinene, Sample 279 was excluded. For ocimene, Sample 295 was removed.

Diurnal variations in α-terpineol, 3-carene and *p*-cymene tracers show trends similar to those observed in the total SOA tracer box plot (Figure 24). The three daytime samples exhibit comparable median, maximum, and minimum values as well as similar interquartile ranges. In general, the highest median values occur during the 09:00-12:00 interval.

p-Cymene and 3-carene display proportionally larger interquartile ranges than the other precursors, indicating greater variability within their concentrations compared to the other compounds. They also show higher day-to-night ratios than the other precursors.

Despite these differences, for all precursors, concentrations during the 18:00-09:00 interval remain extremely low, with narrow interquartile ranges reflecting consistently low night-time values.

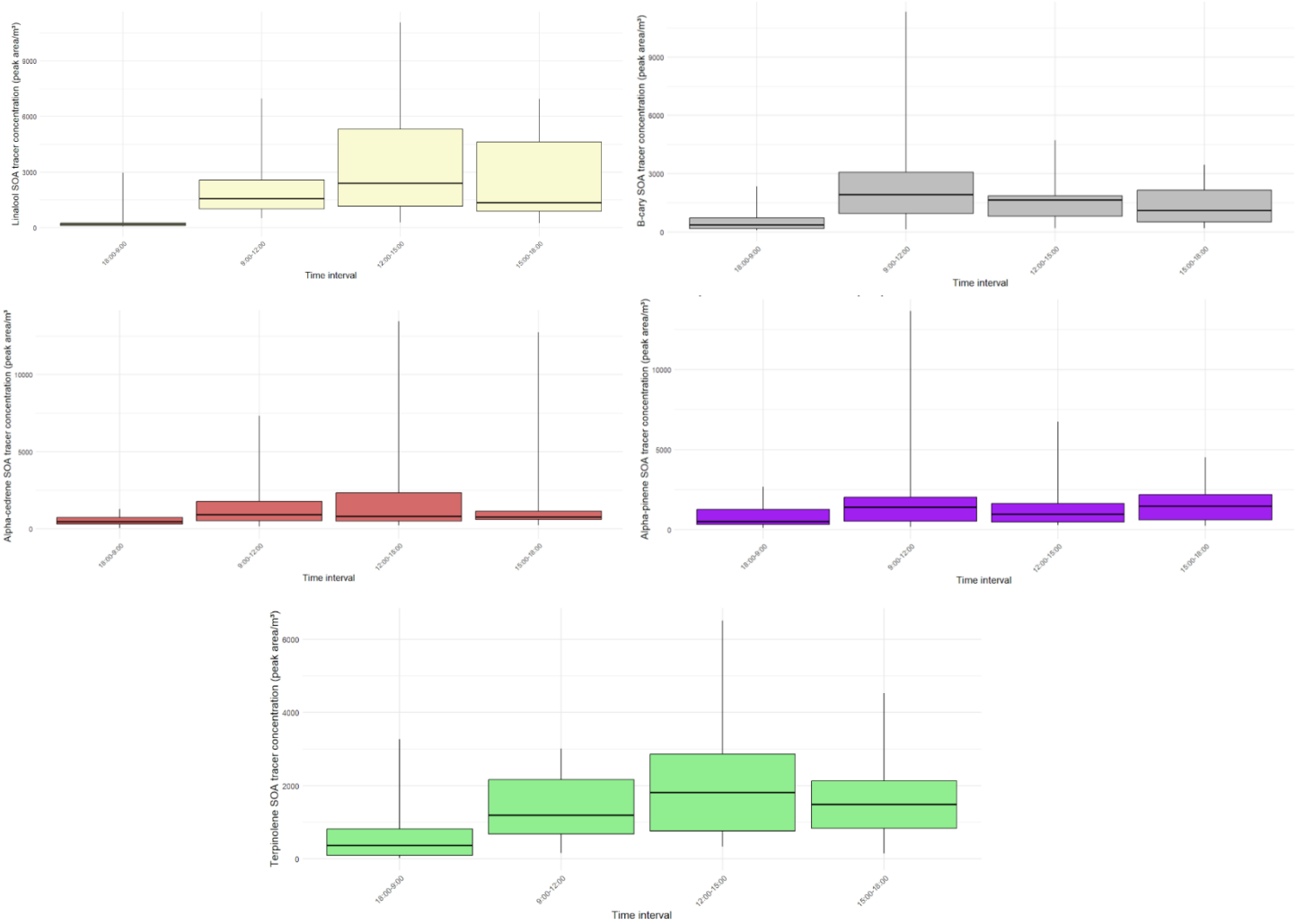


Figure 27: Diurnal variations in linalool (top left), β-caryophyllene (top right), α-cedrene (middle left), α-pinene (middle right) and terpinolene (bottom). SOA tracer concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00. All zero values were removed, and for linalool, β-caryophyllene, α-cedrene and α-pinene, Sample 279 was excluded.

Consistent with previous figures, the diurnal profiles for linalool, β-caryophyllene, and α-cedrene show consistently low concentrations during the night. In contrast, the terpinolene and α-pinene profiles show a greater contribution to SOA production during night-time periods. Additionally, the interquartile ranges are notably wide for terpinolene and α-pinene between 12:00-15:00, and for linalool between 15:00-18:00, significantly wider than any interquartile ranges observed in the other three compounds.

Another point of similarity is that three out of the five diurnal profiles display their highest median concentrations between 09:00-12:00, which is consistent with the trend seen in most of the earlier figures.

A significant trend throughout all 10 diurnal profiles is that most compounds (6 out of 10: *p*-cymene, α-terpineol, 3-carene, β-caryophyllene, α-cedrene, α-pinene) show the highest median SOA tracer levels between 9:00 and 12:00, likely because their structures include strained bicyclic rings or exocyclic alkenes, making their double bonds more reactive to OH radicals.[70] OH radicals are typically most abundant in the late morning, therefore, rapid formation of SOA from these precursors is seen during this time window.[71]

Contrastingly, linalool, terpinolene, and ocimene peak between 12:00 and 15:00 because their structures contain multiple double bonds, making them highly reactive to O_3 as well as OH radicals, which become abundant in the afternoon (see Figure 29 below) so these compounds are likely oxidised via a combination of ozonolysis and OH radicals, resulting in peak SOA formation later on.[72,73]

Sabinene, on the other hand, shows its highest median tracer concentrations in the 15:00-18:00 interval. This delayed peak may be attributed to its relatively slower reaction rates with both OH and O_3 compared to other terpene precursors. The presence of a strained bicyclic structure may influence its formation pathways, possibly leading to delayed SOA production.[74]

3.6. Causes of Diurnal Variations

To address the question of what drives the observed diurnal variations in BSOA tracer concentrations, the patterns shown in Figures 26-27 and the total tracer diurnal box plot (Figure 24) were examined. The most apparent trend observed is that SOA tracer concentrations are consistently higher during the daytime compared to the night-time. This pattern can be attributed to several factors. Firstly, BVOC emissions are strongly temperature

dependent, with previous studies demonstrating that emissions increase with rising temperature and with more sunlight.[75]

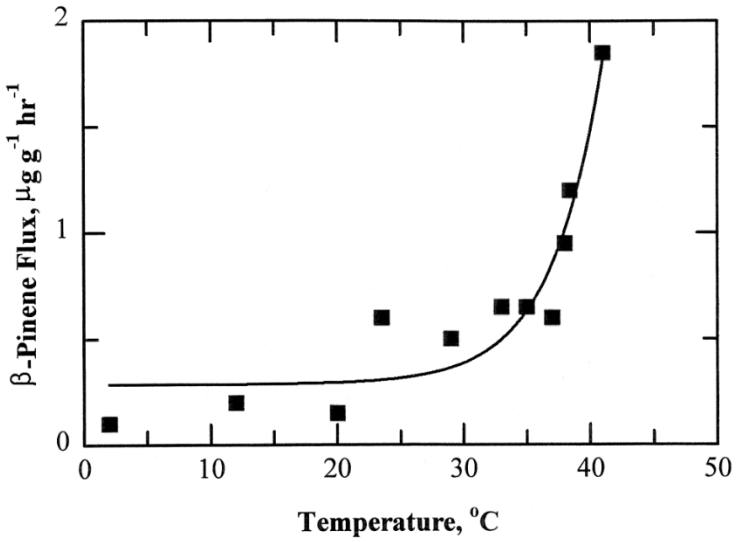


Figure 28: Variation of monoterpene (β-pinene) emissions as a function of temperature. The line represents the best fit to the data points and demonstrates an exponential increase in β-pinene emissions with respect to temperature.[76]

The warmer conditions during the day are expected to enhance precursor availability, leading to greater oxidation and SOA tracer production. In addition, higher daytime temperatures increase the kinetic energy of atmospheric molecules, therefore facilitating more efficient chemical reactions. Further temperature analysis is carried out in **Section 3.8. Temperature.**

Secondly, as outlined in Section **2.1. Chamber Experiments**, ozonolysis plays a key role, since O_3 is introduced as an oxidant in the chamber. Consequently, the photochemical reactions responsible for O_3 formation influence tracer concentrations. O_3 concentrations are significantly higher during daylight hours because it is generated through the photolysis of NO_2, as shown in R1 and R2 below.[77] Because this process requires UV radiation, it predominantly occurs during daytime, resulting in increased precursor oxidation and consequently, greater SOA tracer production.

$$NO_2 + h\nu\ (< 410\ \text{nm}) \rightarrow NO + O\ (^3P) \qquad \text{(R1)}$$

$$O_2 + O\ (^3P) + M \rightarrow O_3 + M \qquad \text{(R2)}$$

In addition to O_3, hydroxyl radicals (OH) are another key oxidant in the chamber. OH is primarily produced through the photolysis of O_3 followed by its reaction with water vapour, as shown in R3 and R4. This process also depends on solar radiation, and therefore, OH concentrations also peak during daylight hours.[78]

$$O_3 + hv\ (< 320\ \text{nm}) \rightarrow O\ (^3P) + O_2 \qquad \text{(R3)}$$

$$O\ (^3P) + H_2O \rightarrow 2OH \qquad \text{(R4)}$$

Low night-time SOA tracer concentrations could also be attributed to processes actively removing O_3 at night. As shown in Figure 29, O_3 concentrations increase throughout the day but decline overnight due to removal processes, resulting in lower levels by the following morning.

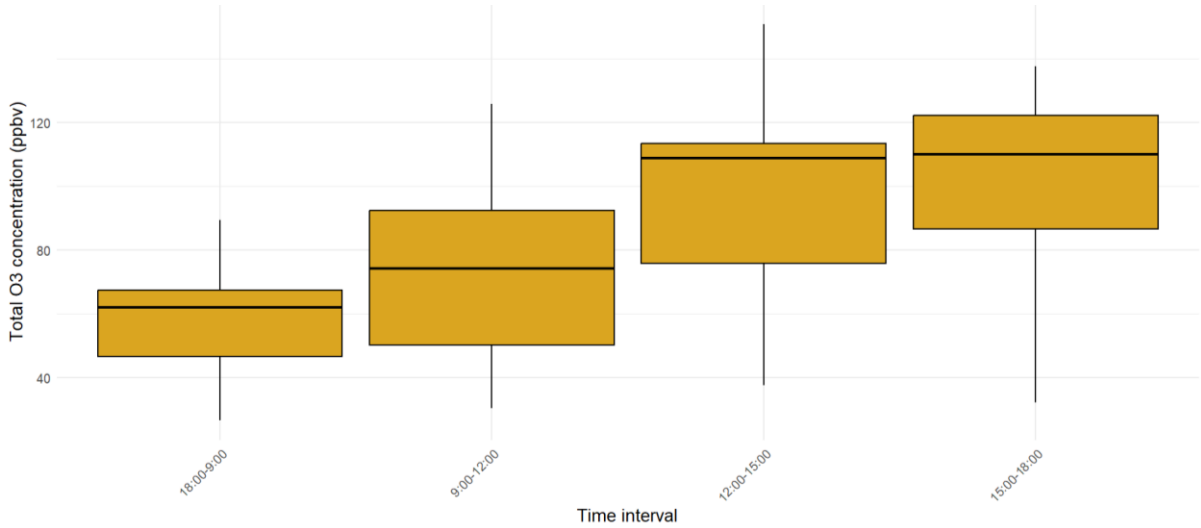


Figure 29: Diurnal variations in O_3 concentrations across four time intervals: 18:00-09:00, 09:00-12:00, 12:00-15:00, and 15:00-18:00.

An example of a possible O_3 removal process is the reaction with NO (a process known as ozone titration), where NO from vehicle emissions reacts with O_3, resulting in low O_3 concentrations near roads (see R5 below).[79]

$$NO + O_3 \rightarrow NO_2 + O_2 \qquad \text{(R5)}$$

Furthermore, O_3 can react with NO_2 radicals to form NO_3 (see R6 below) and can also be removed via dry deposition onto surfaces, resulting in very high loss rates of O_3 throughout the night.[80]

$$NO_2 + O_3 \rightarrow NO_3 + O_2 \qquad \text{(R6)}$$

As well as this, any residual OH at night is rapidly removed through reactions with nitrogen oxides such as NO and NO_2 (see R7 and R8 below).

$$OH + NO \rightarrow HONO \qquad \text{(R7)}$$

$$OH + NO_2 \rightarrow HNO_3 \qquad \text{(R8)}$$

The second trend observed is that the highest tracer concentrations occur during the 9:00-12:00 interval. Among the three daytime samples, 6 out of 10 precursors showed their highest median values within this timeframe, while 5 out of 10 precursors exhibited the highest upper quartile in the same interval.

This pattern is likely due to a combination of compound buildup overnight and fresh emissions in the morning. Although plants emit lower concentrations of precursors at night, some emissions still occur. However, because oxidant concentrations are lower during this time, with only some O_3 and NO_3 being present, the precursors are oxidised more slowly, therefore leading to their accumulation.

By the 9:00-12:00 interval, precursor emissions increase further as temperatures rise. Additionally, O_3 production resumes due to the resumed photolysis of NO_2, of which concentrations are typically elevated at this time, partly due to morning rush hour emissions.[81] As well as this, the photolysis of O_3 begins again, leading to the formation of OH radicals.

The changing boundary layer depth also plays a role, with it being relatively shallow in the early morning, helping to maintain higher concentrations of newly formed SOA tracers and increasing the likelihood of their oxidation.[82] Additionally, as the boundary layer collapses overnight, there is mixing of regionally produced BSOA into the local atmosphere. Together, these factors result in high concentrations of VOC precursors that are rapidly oxidised as O_3 production restarts.

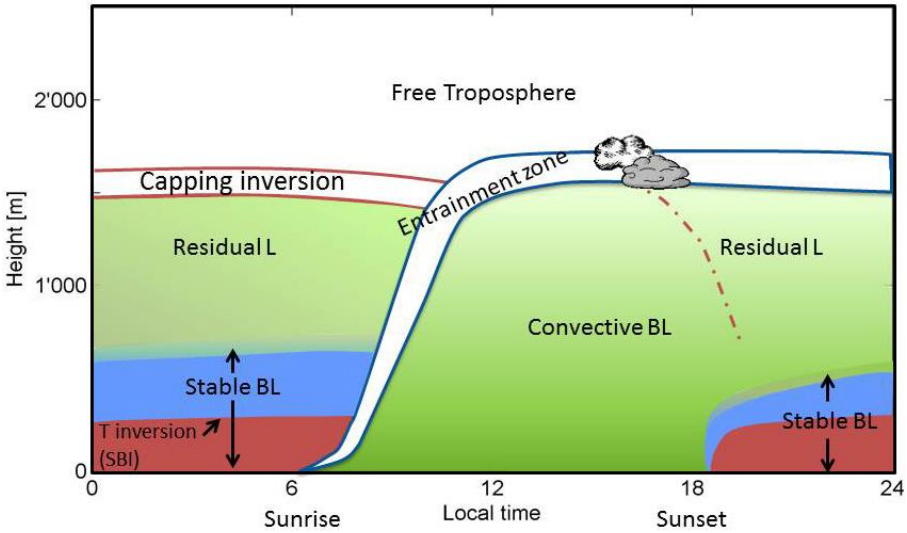


Figure 30: Diagram showing the typical diurnal changes in the boundary layer (BL) depth. At night, a shallow stable boundary layer (SBL) forms near the surface due to cooling, suppressing turbulence. After sunrise, surface heating creates turbulence that grows the convective boundary layer (CBL) upward throughout the day. The residual layer (RL) above the SBL consists of the previous day's CBL. After sunset, the CBL collapses and the SBL reforms.[83]

Due to the absence of photolysis processes at night, SOA formation occurs at much lower concentrations and via different pathways compared to daytime. While O_3 generation stops after sunset, residual O_3 produced during the day may last into the night due to its relatively long atmospheric lifetime.[84] In addition, although plant VOC emissions decrease during the night, they do not stop entirely, allowing for continued precursor availability.

As mentioned previously, the depth of the boundary layer varies throughout the day.[83] During the night, this layer becomes significantly shallower, which reduces the dilution of trace gases. Consequently, even low concentrations of VOCs and oxidants can lead to locally elevated SOA tracer levels. For example, studies in Beijing have shown that the boundary layer height drops to around 500 meters at approximately 6:00 a.m., then rises to about 1500 meters by 3:00 p.m.[85]

In addition, during night-time conditions, O_3 can react with NO_2 to form nitrate radicals, or with alkenes to generate hydroxyl radicals via the formation and decomposition of Criegee intermediates (see Reactions R9 and R10 below).[86]

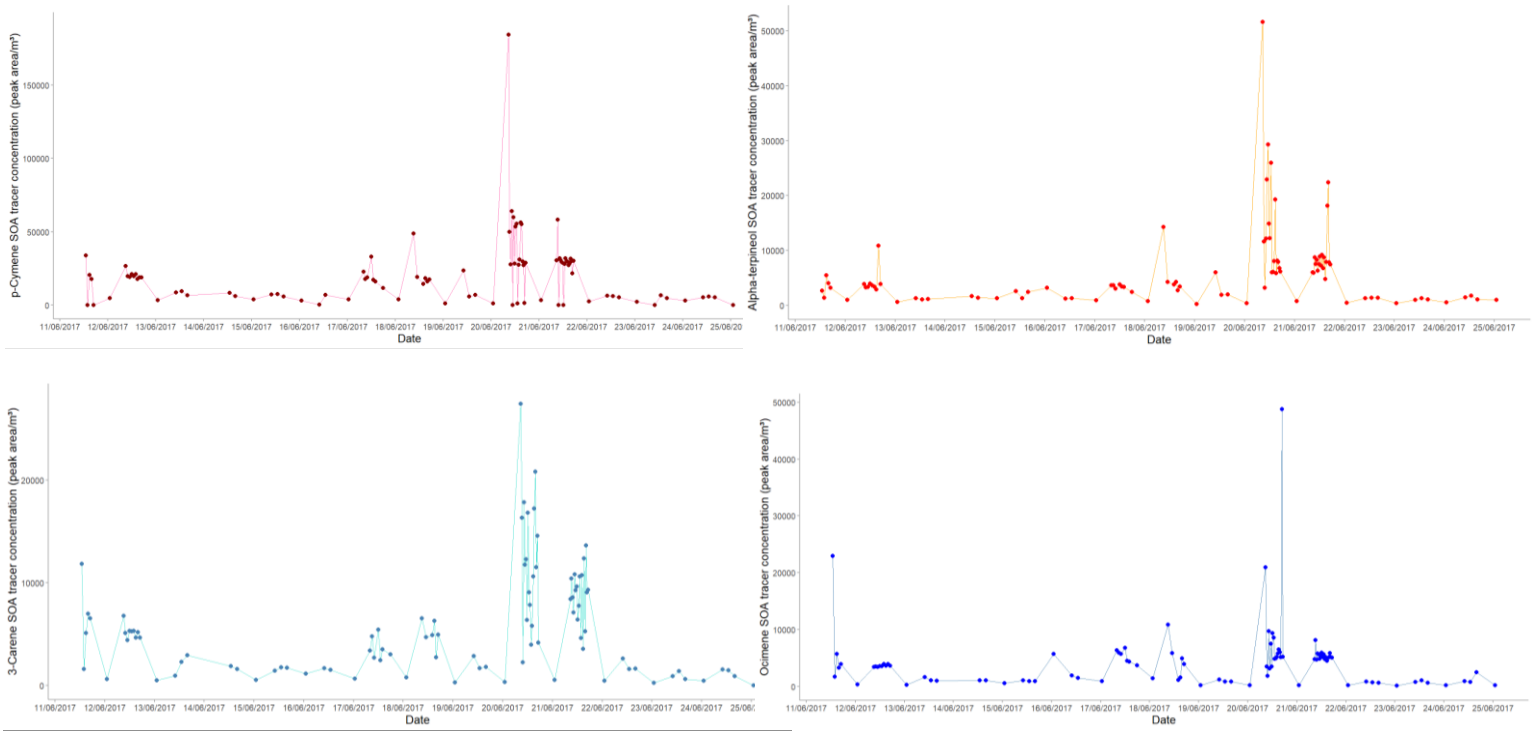
$$NO_2 + O_3 \rightarrow NO_3 + O_2 \qquad \text{(R9)}$$

O_3 + alkene → Criegee → OH + fragments (R10)

Both NO_3 and OH can serve as important night-time oxidants and contribute to the formation of SOA. NO_3, however, typically reacts via addition to double bonds, forming organic nitrates, which are unlikely to be present from the NO_x-free chamber experiments in this study. As a result, NO_3 chemistry may produce SOA that differs from the tracers identified here. In contrast, other oxidants, such as chlorine radicals, which can be present in polluted urban environments, particularly in the morning, may follow reaction pathways more similar to OH oxidation and therefore could contribute to the formation of products more like those observed in the chamber experiments.

3.7. Time series trends

To address the question of how individual BVOC precursors vary over time and contribute to temporal trends in SOA formation, individual time series graphs were generated for each precursor library by summing the concentrations of all tracers associated with that library for each sample. Figure 31 presents time series graphs for the five most abundant precursors, showing their concentrations plotted against the midpoint time of each sampling interval.



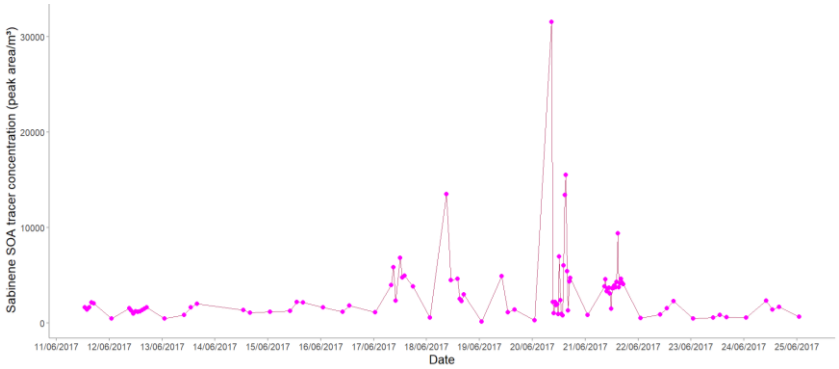
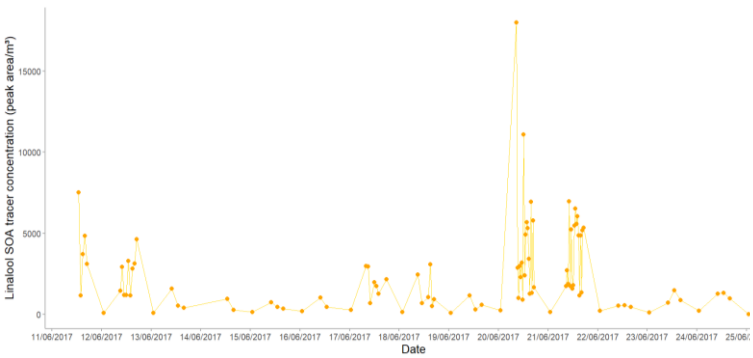


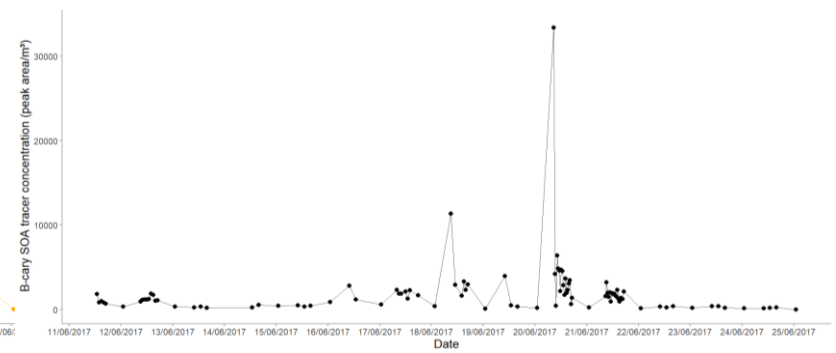
Figure 31: Time series graphs of p-cymene (top left), α-terpineol (top right), 3-carene (middle left), ocimene (middle right) and sabinene (bottom) SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017, demonstrating overall variations in tracer levels during the monitoring period.

The time series graphs in Figure 31 show considerable variation, largely driven by differences in tracer concentrations observed on the 20th and 21st. With the exception of 3-carene, all compounds exhibit lower tracer production on the 21st and significantly higher levels on the 20th. While 3-carene shows elevated concentrations on both days, its peak still occurs on the 20th.

p-Cymene, α-terpineol, and sabinene display trends that most closely resemble the overall SOA tracer profile shown in Figure 20. These compounds are characterised by consistently low and stable concentrations, followed by a sharp, distinct peak on the 20th, specifically in Sample 279. After this peak, concentrations drop significantly but remain elevated compared to levels before the 20th.

3-Carene displays continued activity later in the day, with a smaller secondary peak observed in Sample 295. In contrast, ocimene shows no peak in Sample 279 and only exhibits a single peak in Sample 295.





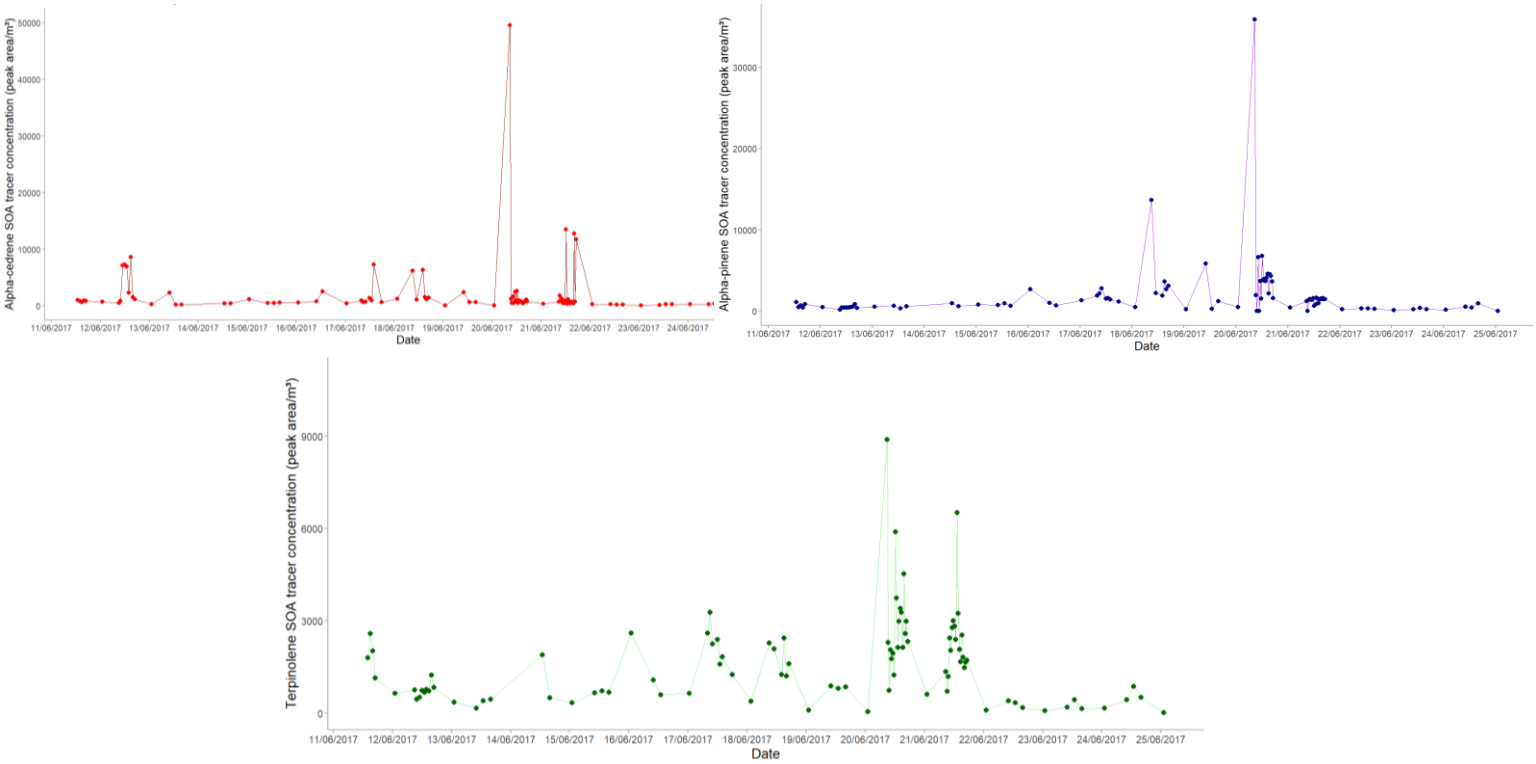


Figure 32: Time series graphs of linalool (top left), β-caryophyllene (top right), α-cedrene (middle left), α-pinene (middle right) and terpinolene (bottom) SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017, demonstrating overall variations in tracer levels during the monitoring period.

Linalool and terpinolene exhibit time series patterns closely resembling those of 3-carene, with their dominant peaks occurring on the 20th in Sample 279, along with lower-intensity fluctuations on the 21st. α-cedrene shows similar trends; however, in most cases, the data is close to the limit of detection, and therefore, a signal is only observed under certain conditions.

B-caryophyllene and α-pinene exhibit trends similar to each other and to those of *p*-cymene, α-terpineol, and sabinene, with a pronounced peak on the 20th in Sample 279 and very little activity on the 21st. In the case of β-caryophyllene and α-pinene, concentrations on the 21st are even lower than those of the previous three compounds.

Overall, all compounds exhibit notably elevated concentrations between the 18th-22nd compared to the 13th-18th. Further investigation is needed to investigate the chemical and atmospheric conditions to understand these trends.

3.8. Temperature

To investigate the factors influencing BSOA formation in Beijing, this section examines temperature, and assesses whether it could be a potential driver of observed tracer variability. A key observation in the BSOA tracer data is the clear shift in trends between the 13th-18th and the 18th-22nd. The earlier period shows relatively low and stable tracer concentrations, while the latter period exhibits a pronounced increase in activity, including a significant peak in Sample 279 on the 20th. This change suggests a shift in atmospheric conditions or emission sources, warranting further investigation into factors driving these differences.

Biogenic VOC emissions are known to be strongly temperature dependent, with previous studies demonstrating that emissions increase with rising temperatures.[75] In subarctic ecosystems, even a warming of 1-2 °C led to five to sixfold increases in isoprene and three to fourfold increases in monoterpene emissions over just a few years.[87]

It is therefore plausible that the increased BSOA tracer concentrations were caused by a sudden increase in temperature, given the well-documented variability of Beijing's summer climate. Situated on the northern edge of the East Asian Summer Monsoon (EASM) region, Beijing is subject to a complex mixture of climatic influences. These include not only tropical and subtropical monsoon systems but also mid to high latitude atmospheric circulations and interactions with the polar climate system.[88] These influences contribute to significant variability in local climate conditions, particularly during the summer months.

To enable assessment of this, a time series graph was generated to examine fluctuations in temperature over the study period, with a particular focus on detecting increases in temperature from the 18th-2nd . The temperature was recorded every 15 minutes, and the resulting time series is shown below, as well as the time series plotted alongside the α-pinene time series for comparison.

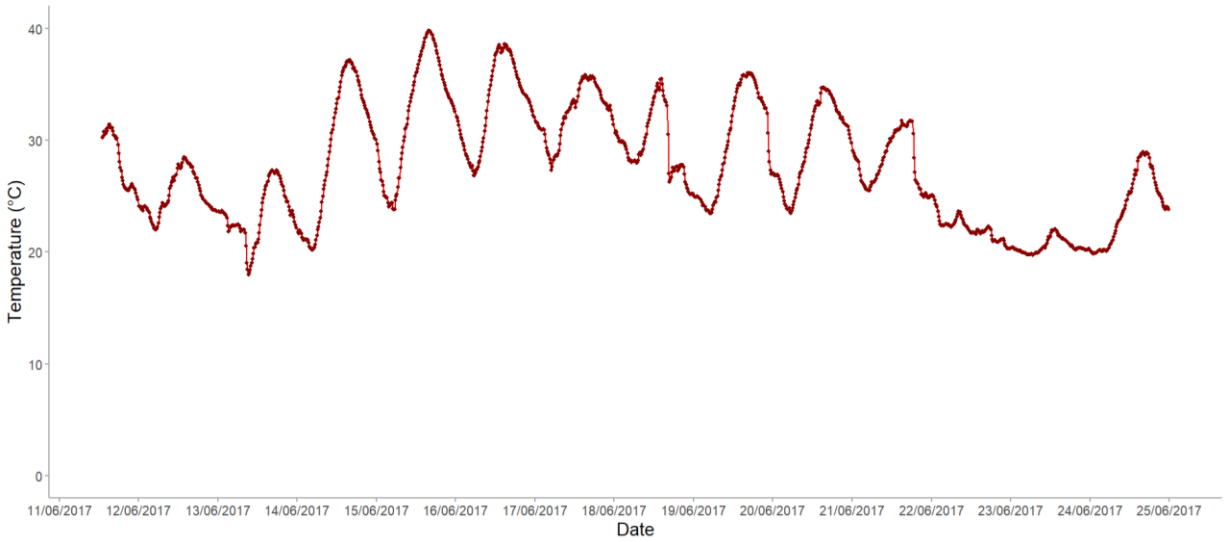


Figure 33: Time series graph of temperature (°C) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period.

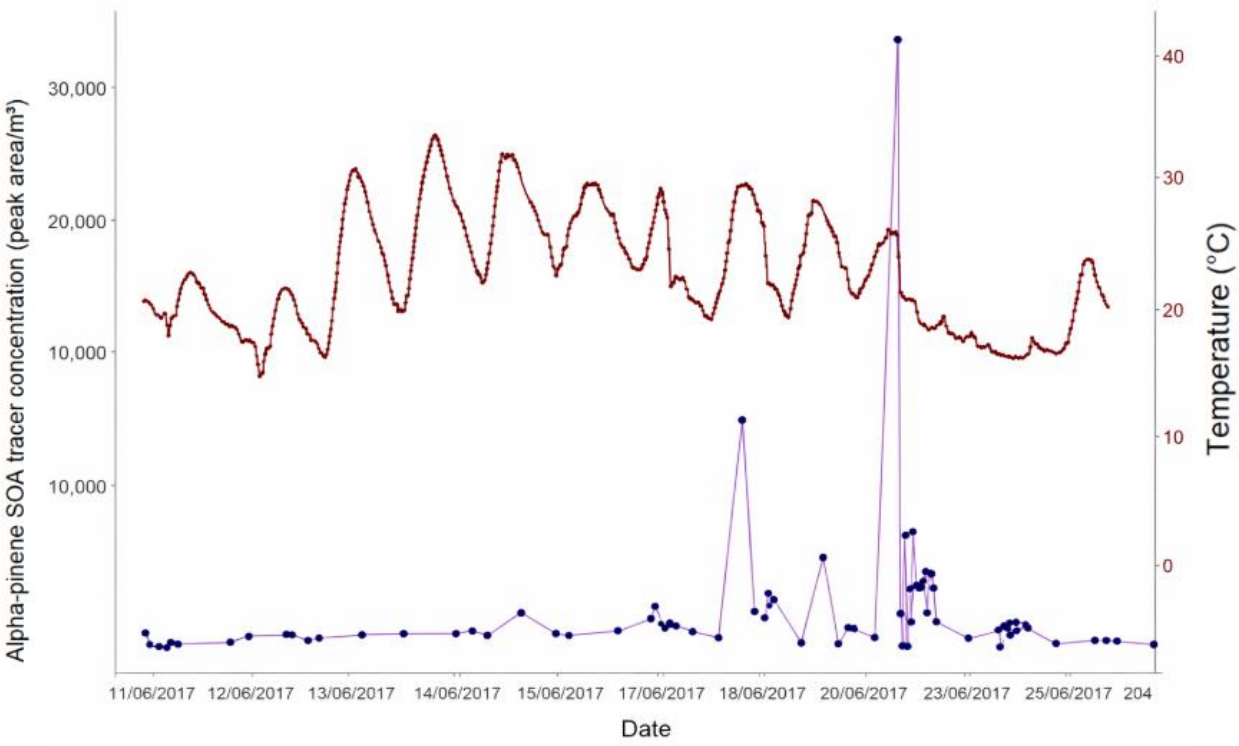


Figure 34: Time series graph of temperature (red, °C), and α-pinene concentrations (purple, peak area/m³) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period.

The temperature remains relatively stable throughout the sampling period, with the only notable variations corresponding to expected diurnal cycles and minor decreases observed on the 13th and between the 22nd-24th. Importantly, there is no indication of increases in temperature from the 18th-22nd, contradicting the initial hypothesis that the observed SOA tracer increases on those days could be caused by a sudden rise in temperature.

This raises the question of whether SOA tracer concentrations in the dataset exhibit any temperature dependence, which would suggest a temperature dependence in the emissions of their precursors. To investigate this, a box plot was created to examine SOA tracer concentrations across different temperature ranges, as shown in Figure 35.

For this analysis, samples were divided into four temperature intervals: 15-25 °C, 25-30 °C, 30-35 °C and 35-40 °C. Within each interval, the concentrations of all tracers in each sample were used to determine median, lower and upper quartiles, and maximum and minimum values.

Sample 279 was excluded from the dataset because it was an extreme outlier that affected the legibility. This sample fell within the 25-30 °C interval. In addition, all values below the detection limit were removed to prevent them from artificially lowering the minimum values, and the temperature measurements were averaged to the BSOA sample times.

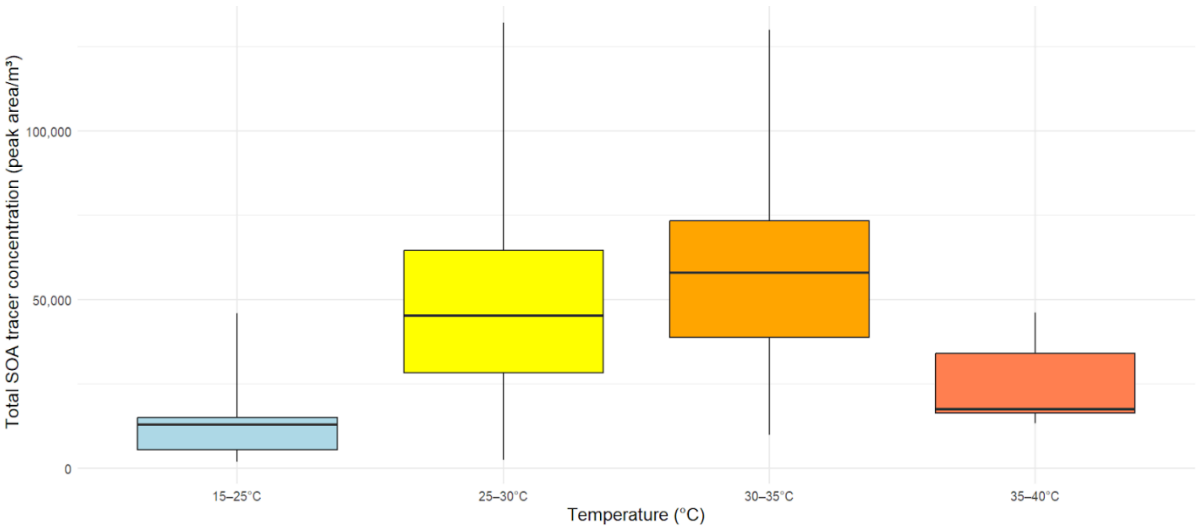


Figure 35: Boxplot showing total SOA tracer concentrations across four temperature intervals: 15-25°C, 25-30°C, 30-35°C and 35-40°C. Temperature measurements were averaged to filter times, and Sample 279 and all zero values were removed.

The observed increase in SOA tracer concentrations from 15-35 °C can be attributed to several temperature-dependent mechanisms. As temperatures rise, both enzymatic activity and vapour pressure in vegetation increase, leading to enhanced emissions of BVOCs.[89] Additionally, elevated temperatures accelerate reactions between BVOCs and atmospheric oxidants, resulting in more rapid formation of oxidation products that subsequently partition into the particle phase to form SOA.

As well as this, during the summer months, higher temperatures are generally accompanied by increased solar radiation. This intensifies photolysis-driven production of atmospheric oxidants, further promoting the oxidation of BVOCs and therefore contributing to enhanced SOA formation.

At temperatures exceeding 35 °C, semi-volatile and low-volatility oxidation products are less likely to partition into the particle phase due to increased saturation vapour pressures.[90] As a result, a greater fraction of these oxidation products remains in the gas phase, leading to reduced SOA tracer concentrations. Additionally, elevated temperatures can cause evaporation of already formed aerosols, further decreasing SOA levels.[91]

Extreme heat can also induce physiological stress in plants, particularly under prolonged drought or low soil moisture conditions.[92] This stress may decrease general plant performance, suppressing overall BVOC emissions, therefore contributing to the observed decline in SOA tracer concentrations at high temperatures.

3.9. Humidity

Humidity can play a significant role in influencing SOA tracer concentrations through both chemical and physical pathways. Increased humidity levels can enhance the formation of OH radicals via reactions involving water vapour, therefore accelerating the oxidation of VOCs.[78] Additionally, higher humidity promotes a more aqueous environment that can alter particle phase chemistry and influence tracer yields.[93] When humidity is high, existing particles absorb water and become more liquid-like, which can make them more reactive and better at taking up gases from the air, leading to higher SOA concentrations. They also have a higher surface area which increases the likelihood of gas-particle interactions. To explore this further, Figure 36 shows the relative humidity time series over the study period.

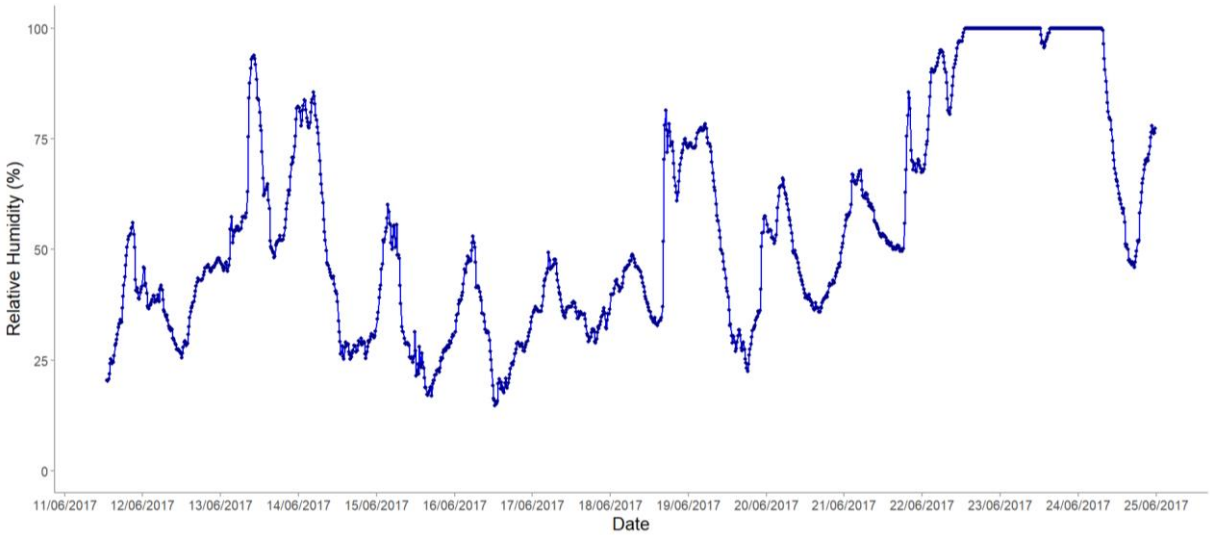


Figure 36: Time series graph of relative humidity (%) from the 11th-25th of June 2017, demonstrating overall variations during the monitoring period.

Figure 36 shows a peak in relative humidity around the 19th, reaching approximately 85 %. While this increase may have partially contributed to the SOA tracer peak observed on the morning of the 20th, it is unlikely to be the primary driver, as similar or higher humidity levels are observed at other points in the monitoring period without corresponding increases in tracer concentrations. Another notable feature is the period of 100 % humidity from the 22nd-24th, however, this occurs after the elevated tracer concentrations on the 18th-22nd and therefore could not have contributed.

3.10. Spatial variability of BSOA tracers

To investigate the potential causes of the pronounced BSOA tracer peak on 20th June, atmospheric transport patterns were considered alongside local meteorological factors. Previous research has shown that in Beijing, air pollution events are often influenced by polluted air masses transported from various directions. Regional transport has been identified as the dominant contributor to $PM_{2.5}$ levels in the city, accounting for up to 72 % of the total $PM_{2.5}$ burden.[94] The highest contribution from short-range transport occurs in summer (25 %), while long-range transport peaks in winter. Notably, the Hebei province has been shown to play a major role, contributing between 37-42 % of $PM_{2.5}$ pollution in Beijing across different seasons. It is therefore highly plausible that the high tracer concentrations between 17th-21st, and the peak on the 20th, were caused by a shift in wind direction or wind speed, allowing heavily polluted air to be carried from a particular region.

The polar plot in Figure 37 illustrates the distribution of the total measured BSOA tracer concentrations as a function of wind direction and wind speed, offering insight into the potential transport pathways responsible for the observed peak. To create the plot, the data was binned by wind direction (angular sectors) and wind speed (radial distance), and the mean tracer concentration within each bin was calculated and represented using a colour gradient. Warmer colours indicate higher tracer concentrations, and cooler colours indicate lower levels.

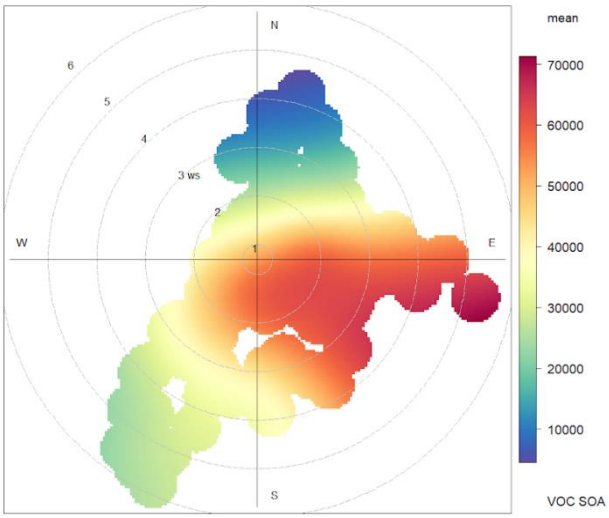


Figure 37: Polar plot showing the mean total SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed.

The distribution of SOA tracers is notably asymmetric, indicating that emission sources are likely directionally dependent. The highest concentrations are associated with winds originating from the east to southeast, suggesting significant sources from that sector. Additionally, tracer concentrations are not impacted by wind speed in this direction, which may imply the influence of both more distant sources and more local ones.

Moderate concentrations are observed in the southern and northeastern sectors, likely reflecting smaller or more dispersed emission sources in those areas.

In contrast, the lowest SOA tracer concentrations occur under winds from the northwest to north, particularly at lower wind speeds. This suggests that air masses from these directions are relatively clean or originate from regions with minimal emissions. This aligns with previous findings, saying when the wind originated from the north, northeast, west, and northwest of Beijing, the annual hourly increases in $PM_{2.5}$ were negative, signifying that these winds were sufficiently clean to facilitate $PM_{2.5}$ removal from Beijing.[95]

To better understand the individual precursors contributing to these trends, separate polar plots were generated for the SOA tracers corresponding to each precursor. Half of these plots (5 out of 10: *p*-cymene, α-terpineol, 3-carene, linalool, β-caryophyllene) exhibited patterns consistent with those observed in Figure 37, with the highest concentrations associated with winds originating from the east to southeast, and these plots are shown below.

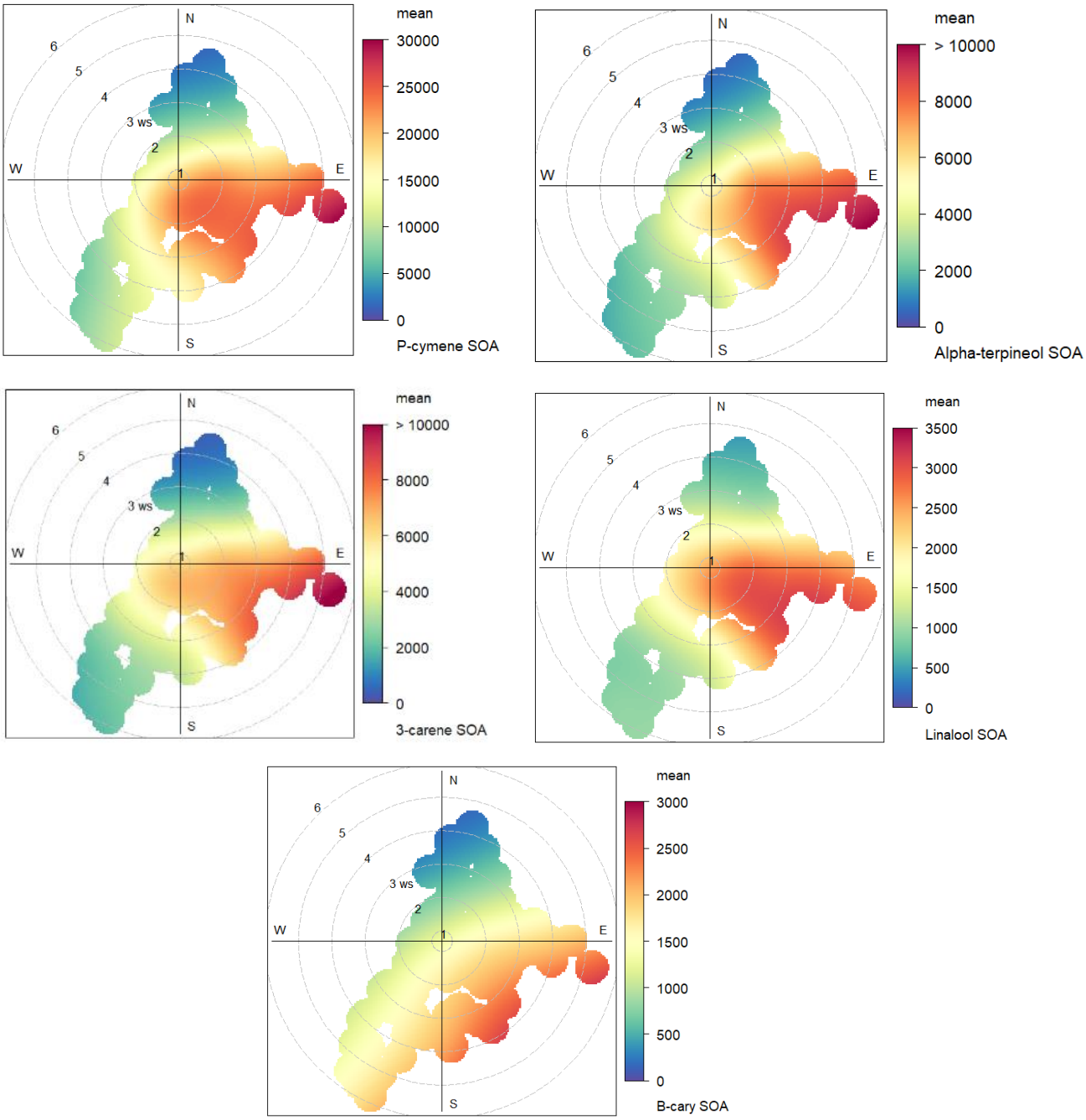


Figure 38: Polar plots showing the mean p-cymene (top left), α-terpineol (top right), 3-carene (middle left), linalool (middle right) and β-caryophyllene (bottom) SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed.

These individual polar plots all display the same directional trend, differing only in the magnitude of concentrations. The three precursors that contributed the largest fraction of total SOA formation are included here (*p*-cymene, α-terpineol and 3-carene), suggesting that the spatial pattern observed in the total tracer polar plot is primarily driven by a small number of dominant precursors.

Additionally, all five of these precursors exhibited distinct peaks at Sample 279 and high concentrations between the 18th-22nd in their respective time series graphs, further supporting the indication that a gust of wind from the east or southeast containing high BVOC concentrations was responsible for the elevated concentrations observed across these compounds during this time.

3.11. Comparison of BSOA trends to other pollutants

In addition to the SOA tracers, three additional factors were examined: ozone (O_3), nitrogen oxides (NO_x), and oxidised organic aerosol (OOA). If an air mass containing both anthropogenic and biogenic components was responsible for the observed increase in SOA tracer concentrations over the period of 18th-22nd, corresponding trends in these co-pollutants would be expected. Alternatively, biogenic precursors may have been transported into the region and subsequently reacted with locally present oxidants, which will likely mean trends won't be similar between biogenic and anthropogenic sources.

To explore these possibilities, time series plots and polar plots were generated to assess the hypothesis that the increased SOA tracer concentrations were associated with polluted easterly or southeasterly winds, and assess whether O_3, NO_x, and OOA were similarly transported by air masses from the east. Polar plots were used as they incorporate wind speed and direction, allowing the influence of atmospheric transport to be assessed alongside chemical relationships. In addition, using the same approach as for the SOA tracers enables direct comparison.

Each pollutant time series was plotted using 15-minute averaged data, and also plotted using mixing ratios or concentrations averaged to BSOA filter sample times. Polar plots were generated using 15-minute averaged data. The plots of O_3 are shown below.

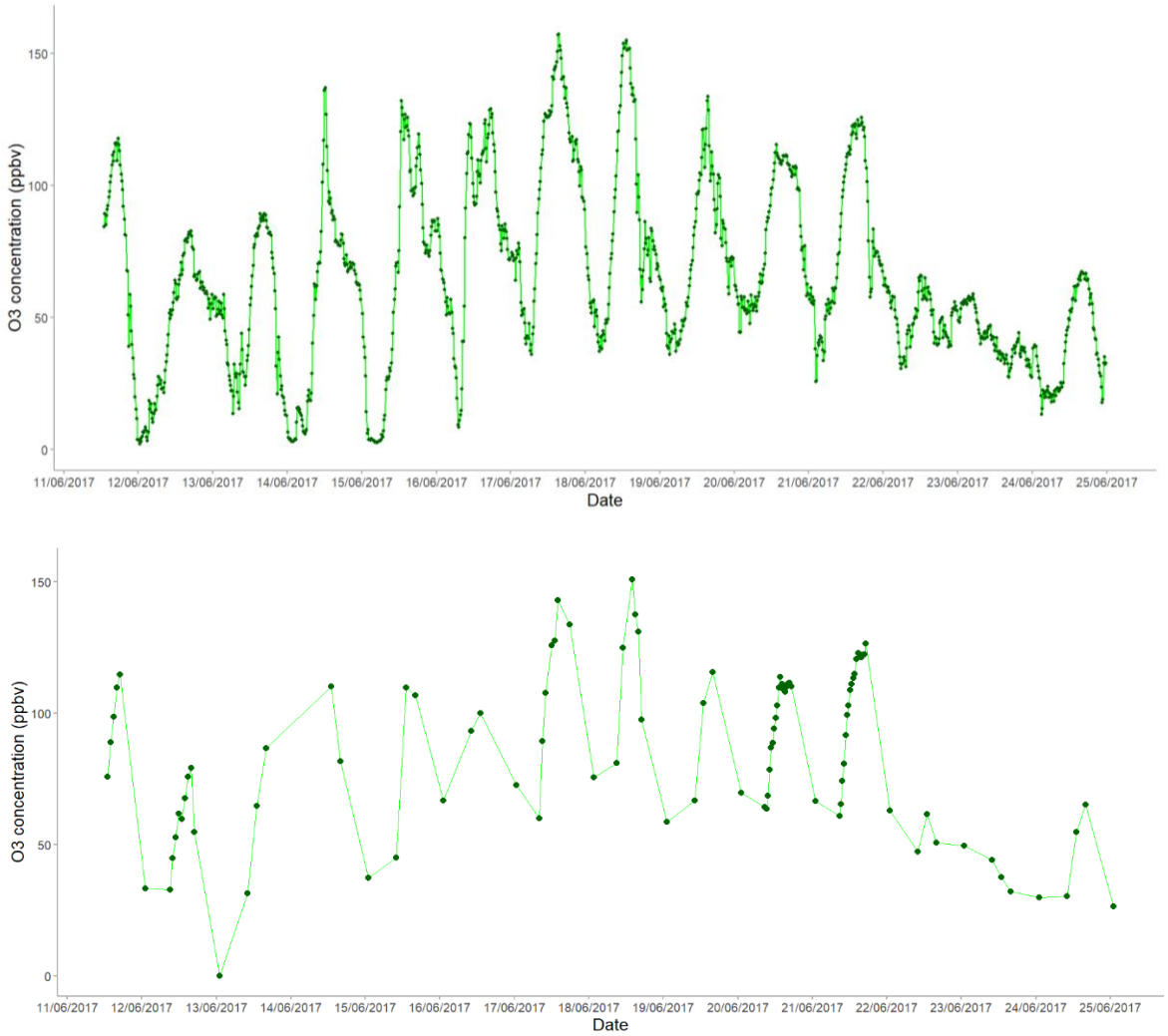


Figure 39: Time series graph of O_3 concentrations (ppbv) from the 11th-25th of June 2017, plotted using 15-minute data (top) and averaged to BSOA filter times (bottom).

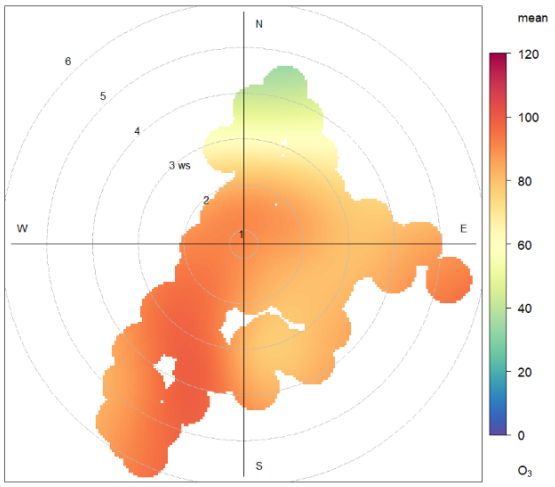


Figure 40: Polar plot showing the mean O_3 concentrations (ppbv) as a function of wind direction and wind speed using 15-minute O_3 data.

Figure 39, the time series plots of O_3 concentrations, reflect reasonably consistent daytime and night-time concentrations, with clear diurnal trends with high O_3 concentrations during the day and low concentrations at night. Notably, there is nothing unusual about the data collected during Sample 279. Significantly, as shown in Figure 29, O_3 levels would be expected to be very low at night, partially due to removal by NO. This is observed until the 17th, after which nocturnal O_3 levels remain higher than expected, lasting until the 22nd. This shows there is a definite shift in concentrations from the 17th onwards.

Additionally, the O_3 polar plot (Figure 40) shows a contrasting pattern to the tracer polar plots (Figures 37-38), with no elevated O_3 concentrations associated with southeasterly winds. The polar plot shows little to no directional dependence, apart from lower concentrations seen coming from the north.

Following on from this, time series graphs of NO_x were also plotted, once again using 15-minute averaged data and then using mixing ratios or concentrations averaged to BSOA filter sample times. A polar plot was also generated using 15-minute averaged data.

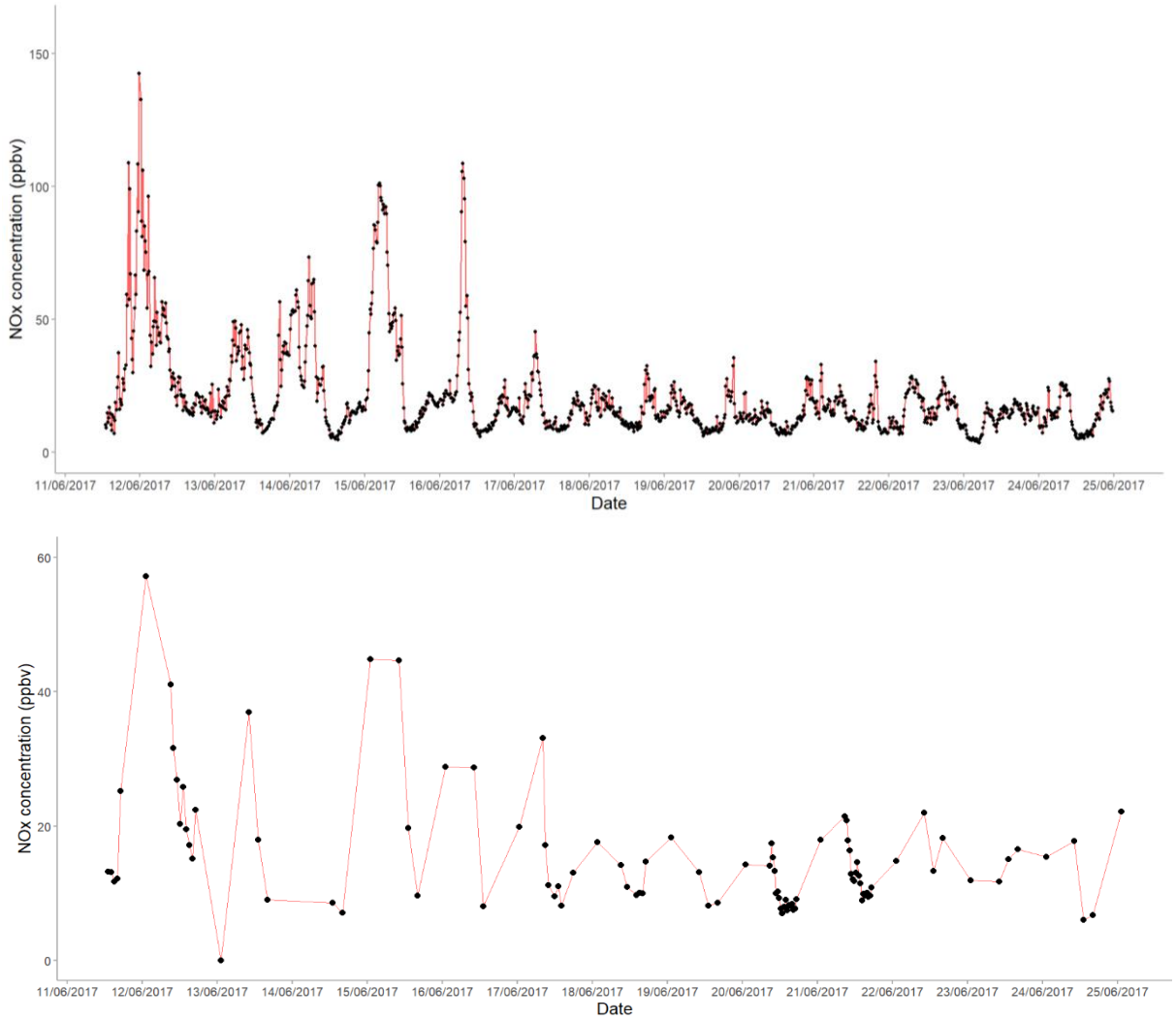


Figure 41: Time series graph of NO_x ($NO + NO_2$) concentrations (ppbv) from the 11th-25th of June 2017, plotted using 15 minute data (top) and averaged to SOA filter times (bottom).

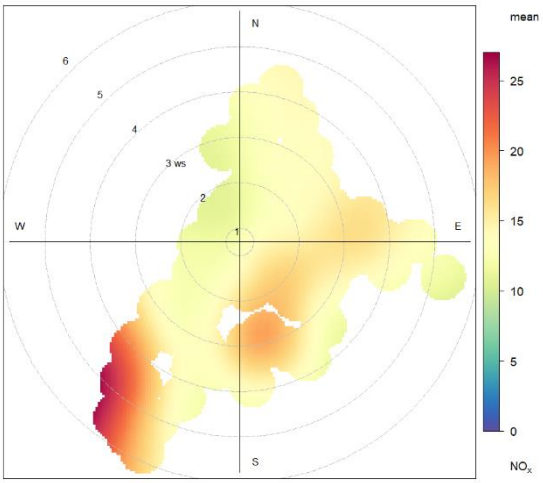


Figure 42: Polar plot showing the mean NO_x ($NO + NO_2$) concentrations (ppbv) as a function of wind direction and wind speed using 15-minute NO_x data.

The time series plots of NO_x (Figure 41) exhibit clear diurnal trends. However, in contrast to O_3, the peaks in NO_x concentrations occur at night, with troughs during the day. Additionally, elevated NO_x levels are observed between the 11th and 17th, which contrasts with the SOA tracer time series data (Figures 31 and 32), where higher concentrations are recorded during the 18th to 22nd.

Lower concentrations during the 18th to 22nd also explain trends observed in Figure 39, the O_3 time series data, as low levels of NO_x observed in Figure 41 led to higher nocturnal O_3 concentrations due to reduced O_3 removal at night.

Additionally, the polar plot (Figure 42) shows very little directional influence apart from higher concentrations coming from the southwest, further confirming that whatever caused the high BSOA tracer concentrations on the 18th-22nd was not associated with a large increase in anthropogenic pollutants.

Following on from this, time series graphs of OOA were also plotted, once again using 15-minute averaged data and then using mixing ratios or concentrations averaged to BSOA filter sample times. A polar plot was also generated using 15-minute averaged data.

Oxidised organic aerosol (OOA) is expected to show the most similar trends to the BSOA, as it is typically dominated by SOA from the oxidation of volatile organic compounds (VOCs) in the atmosphere. However, this will be dependent on the relative role of anthropogenic versus biogenic emissions and chemical oxidation pathways on the OOA signal. While the BSOA tracers measured here are specific chemical markers from known biogenic precursors, OOA represents a broader, less chemically specific measure of oxidised organic aerosol mass.

In this case, the total OOA is made up of three AMS factors determined by positive matrix factorisation, separated based on their degree of oxidation: OOA1 (O:C = 0.53), OOA2 (O:C = 0.74), and OOA3 (O:C = 1.18), with OOA1 being the least oxidised and OOA3 the most.[96] These fractions were summed to produce the total OOA concentrations used for comparison.

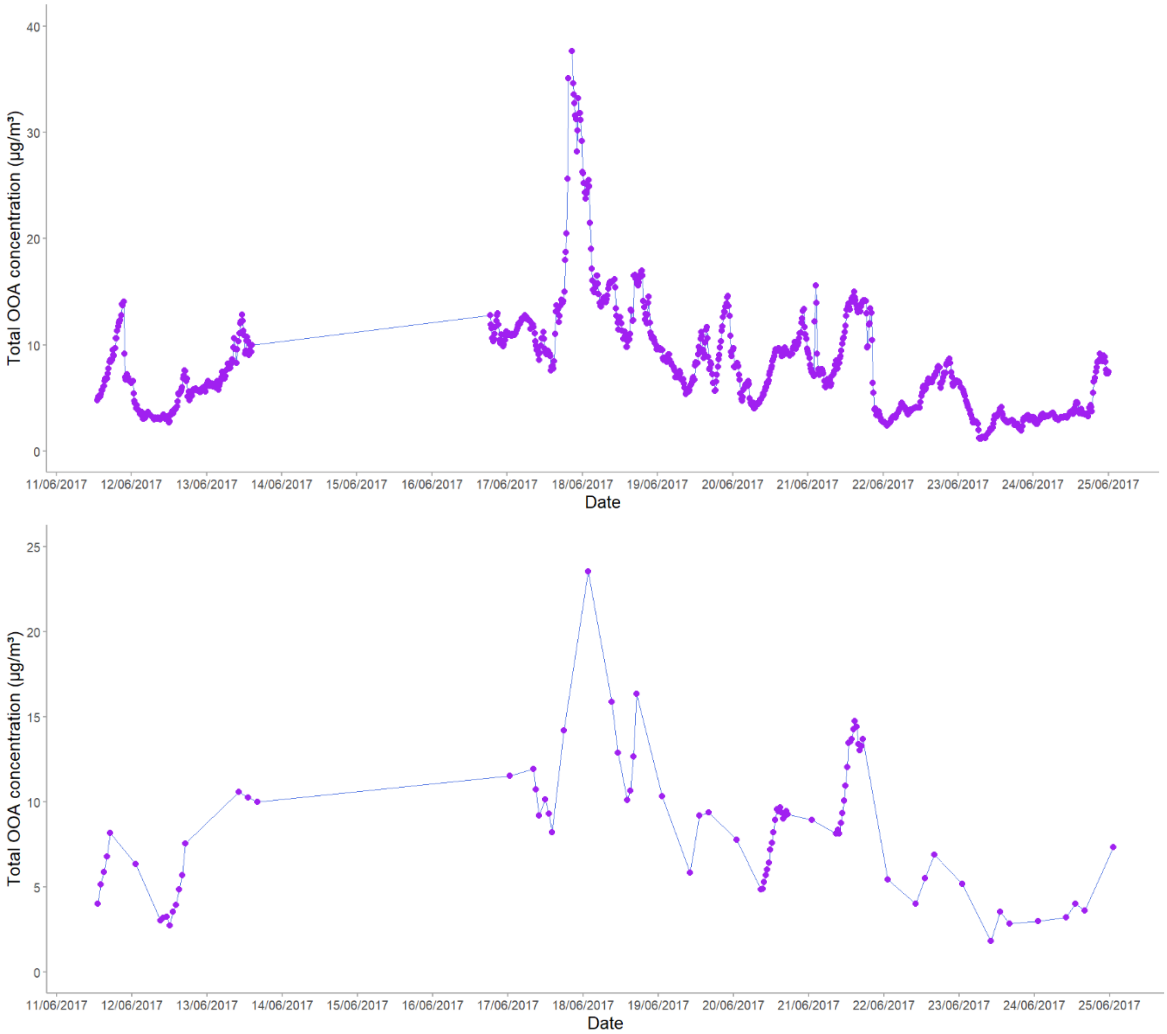


Figure 43: Time series graph of OOA (OOA1 + OOA2 + OOA3) concentrations (µg/m^3) from the 11th-25th of June 2017, plotted using 15-minute data (top) and averaged to SOA filter times (bottom). Data between the 14th-17th is unavailable due to instrument issues.

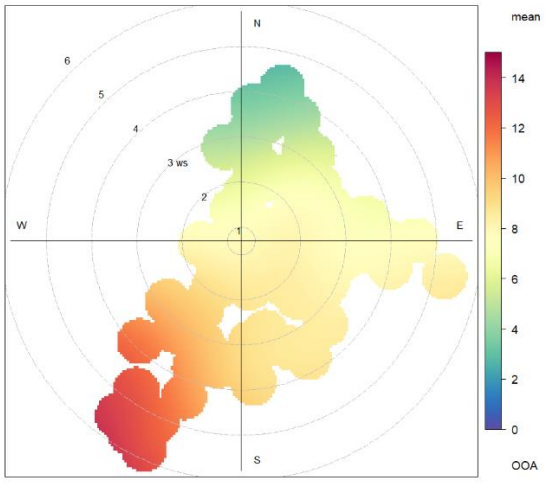


Figure 44: Polar plot showing the mean total OOA (OOA1 + OOA2 + OOA3) concentrations (µg/m^3) as a function of wind direction and wind speed using 15-minute OOA data.

The OOA time series shown in Figure 43 shows a prominent peak on the 18th of June, however, this does not align with the BSOA tracer peak observed on the 20th. Following this peak, concentrations return to levels similar to those seen prior, therefore, they do not correspond with the increase in SOA tracer concentrations observed between the 18th and 22nd. There is, however, a period of missing data from the AMS due to instrument issues, and so it is hard to draw many conclusions based on this comparison.

The polar plot for OOA shown in Figure 44 has elevated concentrations under high wind speeds from the southwest. This differing pattern in comparison to that of BSOA polar plots may be a result of BSOA constituting only a small fraction, and that other SOA precursors may have a greater influence.

The analysis of O_3, NO_x, and OOA data provides strong evidence against the hypothesis that a polluted air mass containing both biogenic and anthropogenic pollutants was responsible for the elevated concentrations observed between the 18th-22nd. None of the three pollutants exhibited peaks in their time series during this period, with O_3 and NO_x only displaying diurnal trends.

Additionally, none of the polar plots showed elevated pollutant concentrations associated with easterly winds. While the OOA polar plot indicated some influence from the southwest, the other two pollutants demonstrated minimal directional dependence. These findings collectively suggest that the high BSOA tracer concentrations observed between the 18th and 22nd were isolated events, likely limited to BSOA tracers only, rather than reflecting broader pollution events involving anthropogenic factors and other oxidised aerosol species.

The elevated concentrations of SOA tracers observed under southeasterly wind conditions suggest that the polluted air mass originated from vegetated areas east and southeast of Beijing, which likely provided the precursors necessary for the formation of BSOA tracers. These BVOCs may originate from forested regions in Hebei Province, as mentioned previously.[97] Located just southeast of Beijing, this area contains extensive rural forest landscapes that are recognised as BVOC emission hotspots, particularly during the summer months when photochemical activity peaks, making them significant sources of isoprene and monoterpenes.

3.12. Contrasting individual polar plots

As previously noted, five out of the ten individual precursor polar plots showed elevated concentrations associated with easterly or southeasterly wind directions. However, this also means that the remaining five did not.

The absence of high concentrations coming from the east or southeast in the polar plots for ocimene and terpinolene is aligned with expectations, as these compounds showed different time series trends and did not exhibit a pronounced peak in Sample 279. More unexpectedly, however, the polar plots for α-pinene, α-cedrene, and sabinene also lack high concentrations from easterly or southeasterly directions, despite all three showing distinct peaks in Sample 279. This discrepancy raises questions about possible differences in reactivity or formation processes influencing these compounds. The corresponding polar plots are shown below:

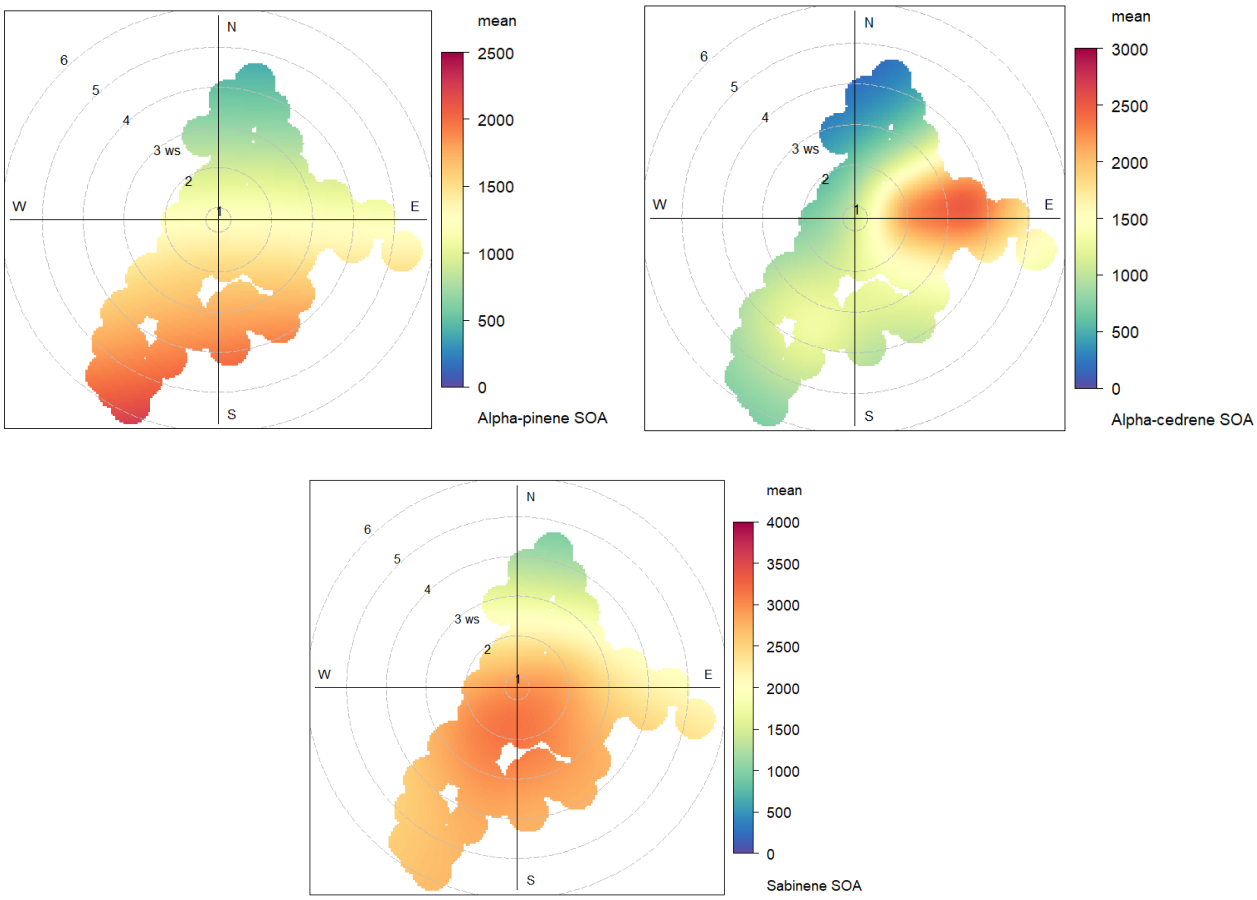


Figure 45: Polar plots showing the mean α-pinene (top left), α-cedrene (top right) and sabinene (bottom) SOA tracer concentrations (peak area/m^3) as a function of wind direction and wind speed.

Of the three polar plots, α-pinene and α-cedrene exhibit the strongest correlations between tracer concentrations and both wind speed and direction, with elevated concentrations observed under northeasterly winds for α-cedrene, and southwesterly winds for α-pinene. This suggests that these SOA tracers likely were derived from a relatively consistent, directional source, possibly a forested area northeast and southwest of the sampling site.

In contrast, the polar plot for sabinene displays less directional influence with only a slight decrease in concentrations from the northeast. This pattern resembles that of α-pinene, however, less pronounced, suggesting a more complex mixture of sources. The polar plots of sabinene and α-pinene are also comparable to the OOA polar plot in Figure 44, with only a slightly weaker directional dependence. The α-cedrene polar plot, however, shows the greatest similarity to the other BSOA polar plots.

These tracer compounds may have been emitted locally from non-easterly sources and only coincidentally peaked during Sample 279 due to local atmospheric mixing. It is also possible that emissions originated from multiple directions and sources, resulting in a more spatially diffuse signal.

3.13. Trends between tracers

3.13.1. Filtering process

Having examined the relationships between SOA tracers and key atmospheric pollutants such as O_3, NO_x, and OOA, further analysis now needs to be carried out on the relationships among the SOA tracers themselves to better understand the potential common sources or formation pathways of SOA. To explore which tracers behave similarly and identify positive and negative relationships, a Corr plot was constructed.

A Corr plot is a visual tool used to show the strength of relationships between pairs of variables. It is in the form of a matrix, where each row and column represent a variable, and the cells display the correlation coefficient between the pair. These coefficients range from -1 to +1: a value of -1 indicates a perfect negative correlation, 0 indicates no correlation, and +1 indicates a perfect positive correlation.

To ensure reliable correlation coefficients are calculated, each factor needs to have a sufficient amount of data. Some tracers appeared in only a few samples, making their correlations unreliable. To address this, compounds detected in fewer than 30 % of samples

were removed, reducing the dataset from 311 to 44 compounds. However, because night-time samples were collected less frequently (once every four samples), this step excluded all compounds produced only at night.

To correct for this, the removed compounds were re-examined against the night-time subset, and any compound present in more than 50 % of those samples was reintroduced. This added 15 compounds back into the dataset, resulting in a final set of 59 compounds. The figure below shows the full time series of SOA tracers before and after filtering.

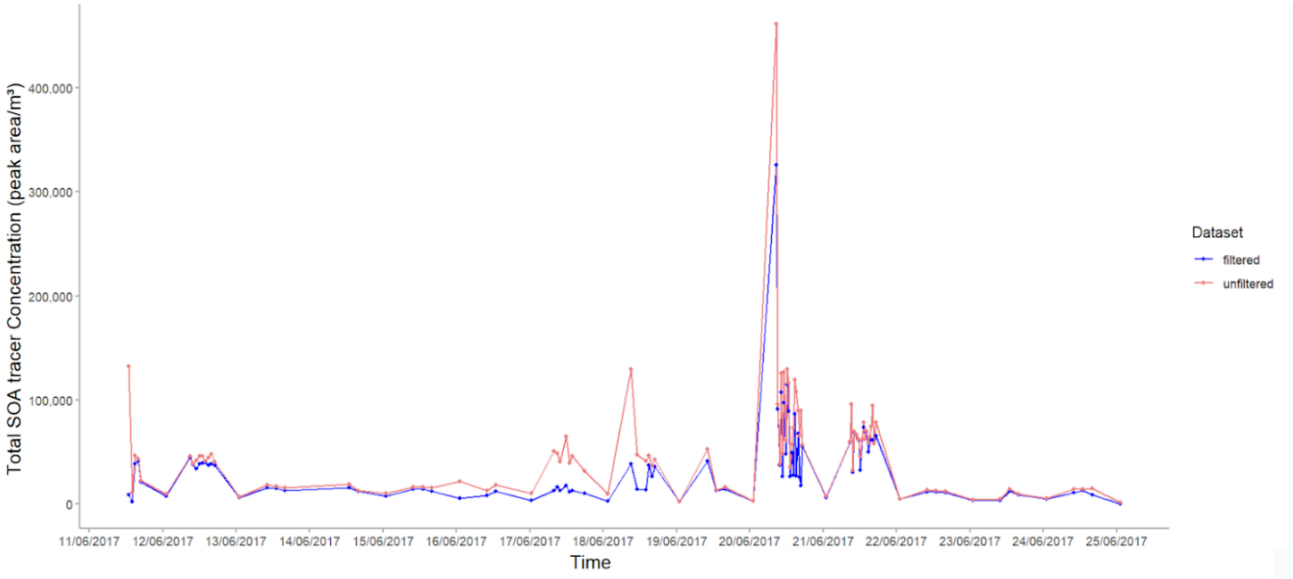


Figure 46: Total SOA tracer concentrations (peak area/m³) from the 11th-25th of June 2017 before (peach) and after (blue) filtering.

Figure 46 shows that although fewer than 20 % of the compounds were retained, they accounted for such a large proportion of the total tracer mass that in many parts of the time series, the filtered and unfiltered lines overlap almost completely. The main differences occur on the 17th, the 18th, and at the peak on the 20th. With the exception of the peaks on the 17th and 18th, all other peaks are preserved in the post-filtering dataset.

3.13.2. Corr plot

The Corr plot was then generated with the dataset of the remaining 59 compounds.

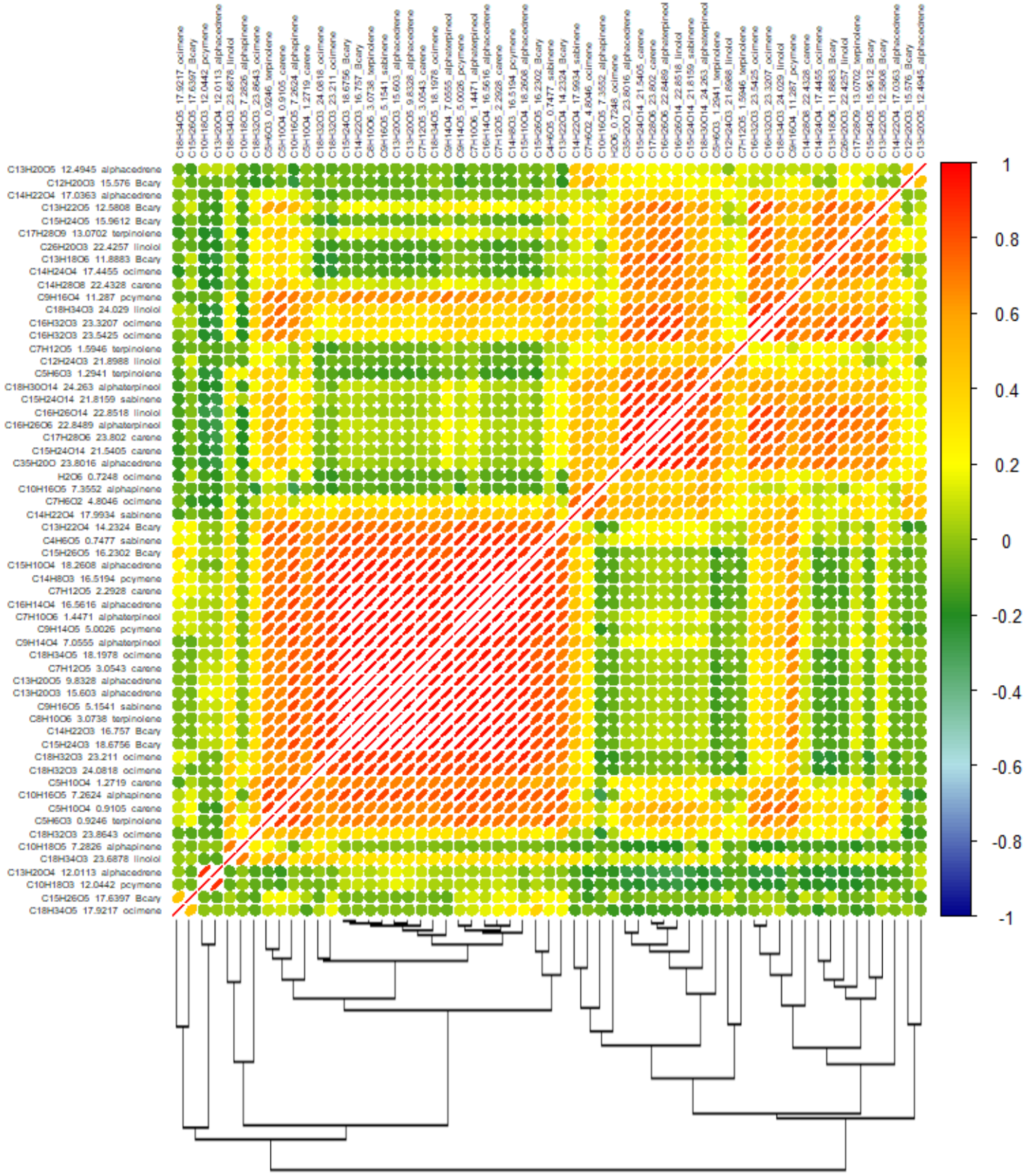

C18H34O5 17.9217 ocimene
C15H26O5 17.6397 Bcary
C10H18O3 12.0442 pcymene
C13H20O4 12.0113 alphacedrene
C18H34O3 23.6878 linolol
C10H18O5 7.2826 alphapinene
C18H32O3 23.8643 ocimene
C5H6O3 0.9246 terpinolene
C5H10O4 0.9105 carene
C10H16O5 7.2624 alphapinene
C5H10O4 1.2719 carene
C18H32O3 24.0818 ocimene
C18H32O3 23.211 ocimene
C15H24O3 18.6756 Bcary
C14H22O3 16.757 Bcary
C8H10O6 3.0738 terpinolene
C9H16O5 5.1541 sabinene
C13H20O3 15.603 alphacedrene
C13H20O5 9.8328 alphacedrene
C7H12O5 3.0543 carene
C18H34O5 18.1978 ocimene
C9H14O4 7.0555 alphaterpineol
C7H10O6 1.4471 alphaterpineol
C16H14O4 16.5616 alphacedrene
C7H12O5 2.2928 carene
C14H8O3 16.5194 pcymene
C10H18O4 18.2608 alphacedrene
C15H26O5 16.2302 Bcary
C4H6O5 0.7477 sabinene
C13H22O4 14.2324 Bcary
C14H22O4 17.9934 sabinene
C7H6O2 4.8046 ocimene
C10H16O5 7.3552 alphapinene
H2O6 0.7248 ocimene
C35H20O 23.8016 alphacedrene
C15H24O14 21.5405 carene
C17H28O6 23.802 carene
C16H26O6 22.8489 alphaterpineol
C16H26O14 22.8518 linolol
C15H24O14 21.8159 sabinene
C18H30O14 24.263 alphaterpineol
C5H6O3 1.2941 terpinolene
C12H24O3 21.8988 linolol
C7H12O5 1.5946 terpinolene
C16H32O3 23.5425 ocimene
C16H32O3 23.3207 ocimene
C18H34O3 24.029 linolol
C9H16O4 11.287 pcymene
C14H28O8 22.4328 carene
C14H24O4 17.4455 ocimene
C13H18O6 11.8883 Bcary
C26H20O3 22.4257 linolol
C17H28O9 13.0702 terpinolene
C15H24O5 15.9612 Bcary
C13H22O5 12.5808 Bcary
C14H22O4 17.0363 alphacedrene
C12H20O3 15.576 Bcary
C13H20O5 12.4945 alphacedrene
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Figure 47: Correlation heatmap with the 59 SOA tracers on each axis. Red shows a +1 correlation coefficient, yellow indicates a 0 correlation coefficient, and blue indicates a -1 correlation coefficient. The dark line across the middle represents +1 correlations of each compound's relationship with itself. Dendrons show hierarchical clustering of SOA tracers, with branches being short within clusters to show tight relationships and long between clusters of more distinct groups.

The most striking feature in Figure 47 is a large red cluster of around 25 compounds that are strongly correlated with each other. Tracers from nearly all precursors (9 out of 10, excluding linalool) are represented, and there is a distinct lack of a consistent trend in O:C ratios, retention times, or carbon number. A smaller but similarly strong red cluster of seven compounds is also present, also without obvious links in precursor type, carbon number, or O:C ratio. This may represent another subset of tracers that respond to a specific external control.

The dendrogram highlights a major split between two broad groups of tracers, once again, these groups do not show obvious structural similarities. There are mid-strength correlations between the groups.

Finally, the absence of strong negative correlations below correlation coefficients of -0.4 suggests that no set of tracers is exclusively formed at the expense of another, and the formation pathways do not generate mutually exclusive tracer sets. Instead, most tracers appear to arise under overlapping or complementary conditions.

3.13.3. Causes of Corr plot trends

Upon first inspection, the large cluster of around 25 compounds could suggest that these compounds belong to the same precursor class. However, since tracers from nearly all precursors are represented, the correlation is more likely linked to shared atmospheric processes. One possibility is that these compounds share common oxidation pathways. The lack of a consistent trend in O:C ratios, retention times, or carbon number, however, points to external controls such as oxidant availability or photochemistry. The smaller cluster of around 5 compounds may represent another subset of tracers that respond to a specific external control.

The major split between two broad groups of tracers is likely once again caused by environmental conditions rather than properties of the compounds driving the separation, due to the lack of obvious structural similarities. The mid-strength correlations between the

groups could point to overlapping formation conditions, such as atmospheric mixing or oxidation that produces intermediates linking the two processes. The lack of strong negative correlations suggests the formation pathways do not generate mutually exclusive tracer sets. Instead, most tracers appear to arise under overlapping or complementary conditions. These trends point to atmospheric processes rather than precursor identity as the dominant driver of clustering.

A possibility is also that the NO_x regime plays a role. Whether it is a high or a low NO_x regime strongly affects yields of SOA. In High NO_x regimes, peroxy radicals (RO_2) usually react with NO in Reaction R11 below.[98]

$$RO_2 + NO \rightarrow RO + NO_2 \qquad \text{(R11)}$$

This then leads to the early termination of oxidation, limiting the formation of highly oxygenated, low volatility products which partition effectively into the particle phase.[99] It also promotes fragmentation, which produces smaller, more volatile molecules that don't partition efficiently into the particle phase. As a result, overall SOA yields are often lower in high NO_x regimes.

In low NO_x air, the concentration of NO is too low for Reaction R11 to dominate. RO_2 mainly reacts with HO_2 or with other RO_2 molecules in reactions R12 and R13.[100]

$$RO_2 + HO_2 \rightarrow ROOH + O_2 \qquad \text{(R12)}$$

$$RO_2 + RO_2 \rightarrow ROOR + O_2 \qquad \text{(R13)}$$

These reactions can be relatively slow, so oxidation has time to proceed over several steps, producing oxidised products that efficiently partition into the particle phase.[101]

The large red cluster may therefore correspond to tracers formed predominantly under low or high NO_x conditions, while the dendrogram split could reflect the contrast between the two regimes. The moderate correlations bridging the clusters are consistent with fluctuating NO_x levels, under which products of both regimes can co-exist and partially overlap.

While trends observed in Figure 47 offer insight into the relationships between SOA tracers and their potential dependence on processes such as NO_x regime, additional work is needed to confirm these interpretations. In particular, focused chamber experiments that directly link SOA formation to atmospheric conditions could help clarify the cause of the

clustering observed in Figure 47. However, using this approach to develop predictive models is outside the scope of this thesis.

3.14. Conclusion

This study investigates SOA formation pathways and composition from biogenic precursors via photo-oxidation, and their contribution to PM pollution in Beijing. Spectral libraries were developed for eight monoterpenoids and two sesquiterpenes through chamber experiments and UHPLC-MS2. These revealed both shared and precursor-specific products, including many common tracers. The study saw production of fragmentation-driven C_7-C_{10} compounds from monoterpenoids, production of higher-carbon sesquiterpene fragmentation products, and signs of functionalisation and oligomerisation.

Applying the libraries to the Beijing filter samples identified 311 SOA tracers. All monoterpenoids, with the exception of linalool, mainly produced low carbon number fragmentation products, while linalool was functionalised to produce medium-sized C_{11}-C_{20} tracers. Sesquiterpenes yielded mainly C_1-C_{15} products with little functionalisation and oligomerisation. Overlap between chamber experiments and ambient data confirmed the relevance of the libraries and the detection of lower abundance compounds, though atmospheric SOA proved to be more complex.

SOA levels remained mostly stable, with high concentrations from the 18th-22nd, and a sharp peak on the 20th of June dominated by *p*-cymene. Diurnal analysis showed daytime concentrations that were significantly higher than night-time, peaking at mid-morning due to precursor buildup, high oxidant concentrations, temperature and shallow boundary layers. α-pinene and terpinolene, however, showed additional night-time production from residual oxidants, leading to night-time oxidation.

The high concentrations from the 18th-22nd and the peak on the 20th were shown not to be temperature-driven but likely linked to southeasterly winds transporting air with high BVOC concentrations from forested regions in the Hebei province. O_3, NO_X and OOA analysis showed contrasting trends, with mostly just diurnal variations, and very little directional influence, suggesting the high tracer concentrations were caused by a factor that didn't influence anthropogenic sources.

This study demonstrates the value of combining controlled chamber experiments with ambient filter analysis to examine the complexity of SOA formation in an urban environment. By building and applying tracer libraries, it was possible to distinguish between precursor-specific oxidation products, identify the dominant role of *p*-cymene in driving SOA pollution in Beijing, and discover both diurnal and month-long patterns in precursor behaviour.

The results show that SOA in Beijing is not shaped solely by local emissions or individual biogenic precursors, but by the interaction of sources under varying atmospheric conditions, often introduced through regional transport. In particular, the high tracer concentrations from the 18th-22nd and the peak on the 20th demonstrate how significant the regional transport of polluted air masses can be, with trends likely being due to southeasterly winds carrying high concentrations of BVOC emissions into the city.

The clustering of tracers across different precursors highlights that environmental factors such as NO_x regimes often control SOA formation pathways, rather than the precursor. These findings provide new insight into the causes of particulate matter pollution in Beijing, putting emphasis on the fact that effective air quality management must account for regional transport and environmental factors.

Future work should aim to improve the atmospheric relevance of chamber experiments by incorporating more realistic conditions, including varying sunlight intensities and interactions with anthropogenic emissions. The use of mixed VOC systems rather than single precursors would allow cross-reactions and competing pathways to be captured, and expanding the range of oxidants to include species such as NO_3 would also help account for combined radical oxidation processes observed in the atmosphere.

In addition, using VOC concentrations closer to ambient levels would ensure more representative reaction pathways. These experiments would have to be longer and at lower, more atmospherically relevant concentrations, and use seed particles to better simulate atmospheric partitioning processes.

Future work should also consider how a changing climate may alter BVOC emissions and subsequent SOA formation in urban environments such as Beijing. Rising temperatures are expected to enhance BVOC emission rates, particularly for temperature-sensitive compounds such as isoprene and monoterpenes. Incorporating these factors into chamber experiments will be essential for predicting how BSOA formation impacts air quality under future climate scenarios. These improvements would help to understand the sources and chemistry of BSOA in urban environments.

References:

1. C. Tomasi, S. Fuzzi, A. A. Kokhanovsky, *Atmospheric Aerosols: Life Cycles and Effects on Air Quality and Climate*, Wiley-VCH, Weinheim, Germany, 2017.

2. J. G. Watson, J. C. Chow, L-W. A. Chen, *Measurement System Evaluation for Fugitive Dust Emissions Detection and Quantification*, Desert Research Institute, Reno, NV, 2010.

3. M. R. Heal, *The formation and characterisation of secondary organic aerosol and its contribution to PM_{10} concentrations in the UK atmosphere*, University of Edinburgh, 2007-2009, https://www.research.ed.ac.uk/en/projects/the-formation-and-characterisation-of-secondary-organic-aerosol-a/, (accessed 10 Oct 2025).

4. Emissions from human activity modify biogenic secondary organic aerosol formation, https://phys.org/news/2021-05-emissions-human-biogenic-secondary-aerosol.html, (accessed 5 Oct 2025).

5. Literature review on emissions of semi- and intermediate volatile organic compounds and formation of organic aerosols with focus on the refinery sector, https://www.concawe.eu/publication/literature-review-on-emissions-of-semi-and-intermediate-volatile-organic-compounds-and-formation-of-organic-aerosols-with-focus-on-the-refinery-sector/, (accessed 7 Oct 2025).

6. M. Hallquist, J. Wenger, U. Baltensperger, et al., *Atmos. Chem. Phys.*, 2009, 9, 5155-5236.

7. H-J. Lim, B. J. Turpin, *Environ. Sci. Technol.*, 2002, 36, 4489-4496.

8. M. Shrivastava *et al.*, *Rev. Geophys.*, 2017, 55, 509-559

9. M. Thoma, F. Bachmeier, F. L. Gottwald, M. Simon, A. L. Vogel, *Atmos. Meas. Tech.*, 2022, 15, 7137-7154.

10. J. Zhu, J. E. Penner, G. Lin, C. Zhou, L. Xu, B. Zhuang, Proc. *Natl. Acad. Sci. U.S.A.*, 2017, 114, 12685-12690.

11. Y. Zheng, L. W. Horowitz, R. Menzel, D. J. Paynter, V. Naik, J. Li, J. Mao, *Atmos. Chem. Phys*., 2023, 23(15), 8993-9007.

12. J. L. Jimenez *et al.*, *Science*, 2009, 326 (5959), 1525-1529.

13. G. Chen et al., *Environ. Int*., 2022, 166, 107325.

14. Y. Mancilla, P. Herckes, M. P. Fraser, A. Mendoza, *Atmos. Res*., 2014, 153, 348-359.

15. S. Zhang, Y. Li, W. Wang, J. Chen, X. Liu, Y. Zhou, H. Wang, *Atmosphere*, 2021, 12(7), 854, DOI: 10.3390/atmos12070854.

16. W. Guo, Z. Li, R. Zhu, Z. Zhou, H. Xiao, H. Xiao, EGUsphere, 2025, 25, 13221-13243

17. C. E. Scott, A. Rap, D. V. Spracklen, P. M. Forster, K. S. Carslaw, G. W. Mann, K. J. Pringle, N. Kivekäs, M. Kulmala, H. Lihavainen, P. Tunved, *Atmos. Chem. Phys.*, 2014, 14(1), 447-470.

18. S. A. Barnhart, MSc. Thesis, University of Dayton, 2020.

19. X. Ma, J. Huang, T. Zhao, C. Liu, K. Zhao, J. Xing, W. Xiao, *Atmos. Chem. Phys.*, 2021, 21, 1-16,

20. C.-S. Zhu, Y. Qu, Y. Zhou, H. Huang, H.-K. Liu, L. Yang, Q.-Y. Wang, A. D. A. Hansen and J.-J. Cao, *Gondwana Res.*, 2021, 90, 159-164

21. Intergovernmental Panel on Climate Change (IPCC), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the IPCC*, Cambridge, UK and New York, NY, USA,
2013, https://www.ipcc.ch/site/assets/uploads/2017/09/WG1AR5_Frontmatter_FINAL.pdf (accessed 10 Oct 2025).

22. T. Subba, M. M. Gogoi, B. Pathak, P. K. Bhuyan, S. S. Babu, *Atmos. Sci. Lett.*, 2020, 21, e975, DOI: 10.1002/asl.975.

23. C. R. Hoyle, G. Myhre, T. K. Berntsen. I. S. A. Isaksen, *Atmos. Chem. Phys.*, 2009, 9(8), 2715-2728.

24. M. Paramonov, PhD Thesis, University of Helsinki, 2015

25. T. Déméautis, M. Delles, S. Tomaz, G. Monneret, O. Glehen, G. Devouassoux, C. George, A. Bentaher, *Chem. Res. Toxicol*., 2022, 35(7), 1146-1161.

26. T. Déméautis, A. Bouyssi, A. Chapalain, J. Guillemot, P. Doublet, A. Geloen, C. George, J. Menotti, O. Glehen, G. Devouassoux, A. Bentaher, *Environ. Sci. Technol*., 2023, 57(15), 6085-6094.

27. X. Wang, B. J. Dewancker, D. Tian, S. Zhuang, Toxics, 2024, 12(6), 377.

28. H. O. T. Pye, C. K. Ward-Caviness, B. N. Murphy, K. W. Appel, K. M. Seltzer, Nat. Commun., 2021, 12, 7215, DOI: 10.1038/s41467-021-27484-1.

29. B. Zhang, L. Qiao, H. Han, W. Xie, L. Li, *Toxics*, 2023, 11(8), 645, DOI: 10.3390/toxics11080645.

30. K. Sindelarova, C. Granier, I. Bouarar, A. Guenther, S. Tilmes, T. Stavrakou, J.-F. Müller, U. Kuhn, P. Stefani, W. Knorr, *Atmos. Chem. Phys.*, 2014, 14(17), 9317-9341.

31. National Research Council, *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, The National Academies Press, Washington, DC, 1991.

32. L. Gao, J. Song, C. Mohr, W. Huang, M. Vallon, F. Jiang, T. Leisner, H. Saathoff, *Atmos. Chem. Phys.*, 2022, 22(9), 6001-6020.

33. J. L. Jimenez *et al.*, *Science*, 2009, 326(5959), 1525-1529.

34. Y. Ma, G. Marston, *Phys. Chem. Chem. Phys.*, 2009, 11(21), 4198-4209.

35. M. J. Newland, A. R. Rickard, T. Sherwen, M. J. Evans, L. Vereecken, A. Muñoz, M. Ródenas, W. J. Bloss, *Atmos. Chem. Phys.*, 2018, 18, 6095-6120.

36. A. Kahnt, R. Vermeylen, Y. Iinuma, M. Safi Shalamzari, W. Maenhaut, M. Claeys, *Atmos. Chem. Phys.*, 2018, 18, 8453-8467.

37. H. Oppata, D. Shimokuri, A. Miyoshi, *Int. J. Chem. Kinet.*, 2024, 57(3), 164-184

38. A. H. Hallward-Driemeier, J. R. Hall, K. A. Spence, B. P. Telicki, A. F. Schaeffer, J. C. Avila, I. S. Albores, A. J. Carrasquillo, *ACS Earth Space Chem.*, 2024, 8(5), 907-919.

39. P. J. Ziemann, R. Atkinson, *Chem. Soc. Rev.*, 2012, **41**, 6582-6605.

40. B. H. Lee, S. Iyer, T. Kurtén, J. G. Varelas, J. Luo, R. J. Thomson, J. A. Thornton, *Environ. Sci.: Atmos.*, 2023, 3, 399-407.

41. D. J. Bryant et al., *Atmos. Chem. Phys.*, 2020, 20, 7531-7552.

42. M. Huo, K. Yamashita, F. Chen, K. Sato, *Atmosphere*, 2022, 13(11), 1813

43. M. Paglione, S. Gilardoni, M. Rinaldi, S. Decesari, N. Zanca, S. Sandrini, L. Giulianelli, D. Bacco, S. Ferrari, V. Poluzzi, F. Scotto, A. Trentini, L. Poulain, H. Herrmann, A. Wiedensohler, F. Canonaco, A. S. H. Prévôt, P. Massoli, C. Carbone, M. C. Facchini, S. Fuzzi, *Atmos. Chem. Phys.*, 2020, 20, 1233-1254.

44. 14-C Analysis | Laboratory of Atmospheric Chemistry, https://www.psi.ch/en/lac/14-c-analysis, (accessed 14 Oct 2025).

45. J. Wu, Q. Zhang, L. Wang, L. Li, X. Lun, W. Chen, Y. Gao, L. Huang, Q. Wang, B. Liu, *Environ. Pollut*., 2024, 361, 124895, doi: 10.1016/j.envpol.2024.124895.

46. H. Wang, Q. Wu, H. Liu, Y. Wang, H. Cheng, R. Wang, L. Wang, H. Xiao, X. Yang, Sensitivity of Biogenic Volatile Organic Compound Emissions to Leaf Area Index and Land Cover in Beijing, Atmos. Chem. Phys., 2018, 18, 9583-9596.

47. X. Jing, X. Lun, C. Fan, W. Ma, *China, J. Environ. Sci.*, 2020, 95, 73-81.

48. J. Wu, Q. Zhang, L. Wang, L. Li, X. Lun, W. Chen, Y. Gao, L. Huang, Q. Wang, B. Liu, Environ. Pollut., 2024, 361, 124895, DOI: 10.1016/j.envpol.2024.124895

49. X. Yang, D. Xiao, H. Bai, J. Tang, W. Wang, *Front. Environ. Sci*., 2022, 10, 842237, DOI: 10.3389/fenvs.2022.842237.

50. H. Ni, R. J. Huang, P. Yao, M. M. Cosijn, N. Kairys, H. Zhong, U. Dusek, *Environ. Int.*, 2022, 158, 106890, DOI: 10.1016/j.envint.2021.106890.

51. S. Hou, D. Liu, J. Xu, T. V. Vu, X. Wu, D. Srivastava, P. Fu, L. Li, Y. Sun, A. Vlachou, V. Moschos, G. Salazar, S. Szidat, A. S. H. Prévôt, R. M. Harrison, Z. Shi, *Atmos. Chem. Phys.*, 2021, 21(10), 8273-8292.

52. J. Cao, S. Situ, Y. Hao, S. Xie, L. Li, *Atmos. Chem. Phys.*, 2022, 22(4), 2351-2364.

53. Z. Shi *et al.*, *Atmos. Chem. Phys.*, 2019, 19(11), 7519-7546.

54. K. L. Pereira, G. Rovelli, Y. C. Song, A. W. Mayhew, J. P. Reid, J. F. Hamilton, *Atmos. Meas. Tech.*, 2019, 12(8), 4519-4541.

55. D. Bryant, PhD thesis, University of York, 2022.

56. J. Xu, W. Chen, Y. Gao, L. Huang, L. Li, B. Liu, J. Wu, Q. Wang, Q. Zhang, S. Tao, *Atmos. Chem. Phys.*, 2021, 21(9), 7321-7341.

57. D. Bryant, Unpublished work.

58. A. Mehra *et al.*, *Faraday Discuss.*, 2021, 226, 382-408.

59. M. J. Newland *et al.*, *Atmos. Chem. Phys.*, 2021, 21(3), 1613-1625

60. N. Sugarman, P. M. Daugherty, *Industrial & Engineering Chemistry*, 1956, *48*(10), 1831-1835.

61. A. Ylisirniö, I. Kurzydym, A. Błaziak, K. Podgórniak, K. Kułacz, K. Błaziak, *Atmos. Chem. Phys.*, 2024, 24, 9309-9331.

62. Z. Fu, S. Guo, Y. Yu, H.-B. Xie, S. Li, D. Lv, P. Zhou, K. Song, Z. Chen, R. Tan, M. Yao, M. Hu, *Environ. Health*, 2024, 2, 486-498.

63. M. Sköld, A.-T. Karlberg, M. Matura, A. Börje, *Food Chem. Toxicol.*, 2006, 44, 538-545.

64. M. R. Alfarra, J. F. Hamilton, K. P. Wyche, N. Good, M. W. Ward, T. Carr, M. H. Barley, P. S. Monks, M. E. Jenkin, A. C. Lewis & G. B. McFiggans, *Atmos. Chem. Phys.*, 2012, 12(14), 6417-6436.

65. OpenAI, *ChatGPT* (image generated via AI), https://chat.openai.com/, (accessed 7 September 2025).

66. Y. Ma, R. A. Porter, D. Chappell, A. T. Russell, G. Marston, *Phys. Chem. Chem. Phys.*, 2009, 11, 4184-4197.

67. N. L. Ng, J. H. Kroll, A. W. H. Chan, P. S. Chhabra, R. C. Flagan, J. H. Seinfeld, *Atmos. Chem. Phys*., 2007, 7(14), 3909-3922.

68. F. Klein, N. J. Farren, C. Bozzetti, K. R. Daellenbach, D. Kilic, N. K. Kumar, S. M. Pieber, J. G. Slowik, R. N. Tuthill, J. F. Hamilton, U. Baltensperger, A. S. H. Prévôt, I. El Haddad, *Sci. Rep*., 2016, 6, 36623, DOI: 10.1038/srep36623.

69. S. N. Rogak, U. Pott, T. Dann, D. Wang, J. *Air Waste Manage. Assoc*., 1997, 48(7), 604-615.

70. I. R. Piletic, T. E. Kleindienst, *J. Phys. Chem. A*, 2022, 126, 88-100.

71. H. Satsumabayashi, H. Kurita, Y.-S. Chang, G. R. Carmichael, H. Ueda, *Atmos. Environ. Part A*, 1992, 26(15), 2835-2844.

72. F. Bernard, V. Daële, A. Mellouki, H. Sidebottom, *J. Phys. Chem. A*, 2012, 116(24), 6113-6126.

73. J. E. Ham, S. R. Jackson, J. C. Harrison, J. R. Wells, *Atmos. Environ.*, 2015, 122, 520

74. A. Calogirou, B. R. Larsen, D. Kotzias, *Atmos. Environ.*, 1999, **33**, 1423-1439.

75. Y. Yang, F. Sun, Y. Chen, S. Yang, J. Environ. Sci., 2025, 159, DOI: 10.1016/j.jes.2025.03.054.

76. J. D. Fuentes, M. Lerdau, R. Atkinson, D. Baldocchi, J. W. Bottenheim, P. Ciccioli, B. Lamb, C. Geron, L. Gu, A. Guenther, T. D. Sharkey, W. Stockwell, *J. Atmos. Sci.*, 2000, 57, 123-150.

77. C. Reed, M. J. Evans, P. Di Carlo, J. D. Lee, L. J. Carpenter, *Atmos. Chem. Phys.*, 2016, *16*, 4707-4724.

78. B. Alicke, U. Platt, J. Stutz, *J. Geophys. Res.*, 2002, 107, 8196.

79. G. A. Novak, M. P. Vermeuel, T. H. Bertram, *Atmos. Meas. Tech.*, 2020, 13(4), 1887-1907.

80. S. Albertin, J. Savarino, S. Bekki, A. Barbero, R. Grilli, Q. Fournier, I. Ventrillard, N. Caillon, K. Law, *Atmos. Chem. Phys.*, 2024, 24(2), 1361-1388.

81. E. Kim, G. H. Moore, *Young Sci. J.*, 2024, 19, available at https://wp0.vanderbilt.edu/youngscientistjournal/article/investigating-the-effect-of-the-ulez-on-nox-concentrations-in-central-london.

82. S. K. Mehta, M. V. Ratnam, S. V. Sunilkumar, D. N. Rao, B. V. Krishna Murthy, *Atmos. Chem. Phys.*, 2017, 17(1), 531-549.

83. R. B. Stull, *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, 1988.

84. M. J. Prather, X. Zhu, *Elementa: Sci. Anthropocene*, 2024, 12, 00112. DOI: 10.1525/elementa.00112.

85. T. Yang, Z. Wang, W. Zhang, A. Gbaguidi, N. Sugimoto, X. Wang, I. Matsui, Y. Sun, *Atmos. Chem. Phys.*, 2017, 17, 6215-6225.

86. W. R. Stockwell, C. V. Lawson, E. Saunders, W. S. Goliff, *Atmosphere*, 2012, 3, 1-32.

87. J. Tang, H. Valolahti, M. Kivimäenpää, A. Michelsen, R. Rinnan, J. *Geophys. Res.: Biogeosci*., 2018, 123, DOI: 10.1002/2017jg004139.

88. L. Yang, J. Zhao, P. Hu, S. Qu and G. Feng, *Atmos.-Ocean*, **2019**, *57*(4), 295-306.

89. A. Grira, PhD thesis, Université de Rennes, 2021.

90. N. M. Donahue, A. L. Robinson, S. N. Pandis, J. H. Seinfeld, *Environ. Sci. Technol*., 2006, 40, 2635-2643.

91. K. R. Kolesar, Z. Li, K. R. Wilson, C. D. Cappa, *Environ. Sci. Technol.*, 2015, 49(20), 12242-12252.

92. J. Feldner, M. O. P. Ramacher, M. Karl, M. Quante, M. L. Luttkus, *Environ. Sci.: Atmos.*, 2022, 2, 1132-1151.

93. Z. Li, K. A. Smith, C. D. Cappa, *Atmos. Chem. Phys.*, 2018, 18, 14585-14608.

94. Z. Zhang, S. Wang, P. Wang, S. Zhu, H. Zhang, S. Wang, *Atmos. Environ.*, 2024, *323*, 120395. DOI: 10.1016/j.atmosenv.2024.120395.

95. L. Wang, X. An, Y. Li, K. Jing, Q. Wang, F. Sun, B. Liu, *Atmos. Environ.*, 2024, *327*, 120493. DOI: 10.1016/j.atmosenv.2024.120493.

96. D. S. Gross, M. E. Gälli, M. Kalberer, A. S. H. Prévôt, J. Dommen, M. R. Alfarra, J. Duplissy, K. Gaeggeler, A. Gascho, A. Metzger, U. Baltensperger, Anal. Chem., 2006, 78, 2130-2137.

97. Y. Lin, X. Lun, W. Tang, Z. Zhang, X. Jing, C. Fan, Q. Wang, For. Ecosyst., 2021, 8, 52, DOI: 10.1186/s40663-021-00322-y.

98. C. Fittschen, *Chem. Phys. Lett.*, submitted 21 March 2019.

99. M. Sarrafzadeh, J. Wildt, I. Pullinen, M. Springer, E. Kleist, R. Tillmann, S. H. Schmitt, C. Wu, T. F. Mentel, D. Zhao, D. R. Hastie, A. Kiendler-Scharr, *Atmos. Chem. Phys.*, 2016, 16, 11237-11248.

100. N. Illmann, *Environ. Sci.: Atmos*., 2025, 5, 737-755.

101. M. Schervish, N. M. Donahue, *Environ. Sci.: Atmos*., 2021, 1, 79-92.