A walk in the park – identifying healthy greenspaces using scents

1. William T. Kay[a,*] https://orcid.org/0000-0002-2561-4856
2. Anya Lindström Battle[a] https://orcid.org/0000-0001-6356-4688
3. Mahal Humberstone[a]
4. Molly Tucker[a] https://orcid.org/0009-0002-5463-4239
5. Kieran E. Storer[a] https://orcid.org/0000-0001-9844-7174
6. Geoffrey Kite[c] https://orcid.org/0000-0003-4482-0409
7. Katherine Willis[a] https://orcid.org/0000-0002-6763-2489

[a] Department of Biology, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

[c] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK.

* Corresponding author: william.kay@biology.ox.ac.uk

1. Abstract

1. Biogenic volatile organic compounds (bVOCs) are carbon-based secondary metabolites emitted by all kingdoms of life, often playing roles in stress protection, growth modulation, and inter-organism communication. In plants, bVOCs such as terpenes (e.g., alpha-pinene) are highly prevalent and have been linked to diverse health benefits including stress reduction, enhanced cognitive function, and neuroprotection. In contrast, urban environments contain anthropogenic VOCs (aVOCs) produced from vehicles and industrial activity. Examples of aVOCs such as benzene, toluene, ethylbenzene and p-xylene (BTEX compounds) are all associated with serious health risks. Our study proposes a novel, health-oriented metric for site evaluation.

2. Primarily, we identified which volatiles appeared in the air across six urban greenspaces on a single date in Oxford Using Tenax™ filters and GCMS analysis. Data were processed using workflows optimised for large metabolomics-style datasets resulting in the identification of 245 unique compounds of biological interest – a significant increase from those identified in previous open-air studies. Using multivariate analysis, we show sites contain significantly different volatile profiles and further, by monitoring a single over 12 months, that temperature, humidity, wind-speed and rainfall all affect the presence of health-promoting compounds.

3. Target 12 of the Global Biodiversity Framework calls for expanding high-quality, health-promoting urban greenspaces. Our findings (i) support the integration of VOC profiling into urban planning and public health strategies to enhance the well-being of urban populations, and (ii) that environmental variables further enable us to prescribe the most beneficial periods for outdoor recreational activities.

2. Keywords

VOC, aVOC, bVOC, volatiles, health, terpenes, BTEX

3. Introduction

Biogenic volatile organic compounds (bVOCs) are a chemically diverse group of carbon-based secondary metabolites produced across all kingdoms of life (Wang et al. 2024). They function in stress protection, signalling, and communication within and between species (Weisskopf, Schulz, and Garbeva 2021) and many are released into the air via passive diffusion from different tissues. In plants, bVOCs such as isoprene, monoterpenes, and sesquiterpenes account for almost all biogenic emissions (Guenther et al. 2012). Among these, monoterpenes such as alpha-pinene, beta-pinene, and limonene are key contributors to both atmospheric chemistry and human well-being. Numerous studies link exposure to plant-emitted terpenes with lowered cortisol, blood pressure, and heart rate, as well as improved mood, cognition, and immune function (Zeng et al. 2022; Park et al. 2010; Li 2010; Bratman et al. 2015). By contrast, anthropogenic VOCs (aVOCs) derived from transport, plastics, and industrial sources - particularly benzene, toluene, ethylbenzene, and xylene (collectively known as the BTEX compounds) - are well-known urban pollutants associated with respiratory and cardiovascular diseases and cancer (Gao et al. 2023).

Because of their location the ‘scentscapes’ (the total volatile profile in a specific time and place) of urban greenspaces are likely to comprise both bVOCs and aVOCs. Understanding the balance between these “healthy” and “unhealthy” volatiles is therefore essential for assessing how urban greenspaces influence well-being, and for driving policy for positive changes in urban development.

Here, we present a new analytical pipeline for characterising ambient air VOCs *in situ*, integrating field sampling and thermal desorption gas chromatography mass spectrometry (TD-GC-MS) analysis. Although similar methods have been developed for headspace sampling (Yin et al. 2022; de Carvalho Couto et al. 2024; Reyrolle et al. 2024), none have yet been published for the analysis of ambient air. We apply this approach to six Oxford greenspaces to test whether site location determines the distribution of biogenic versus anthropogenic volatiles - and to identify those environments most likely to promote human health.

Additionally, it is well known that the release of bVOCs varies in time. For instance, plants coordinate their volatile production with the availability of substrates from photosynthesis, which varies considerably over the course of a day. Indeed, the production of most plant volatiles is under diurnal and/or circadian control; on a longer timescale, the effects of factors such as seasonality, phenology, and other environmental effects including temperature and humidity on the emission of volatiles have been well established (Joo et al. 2018; Yang et al. 2024; Mu et al. 2022; Li et al. 2023).

As well as looking at a volatile snapshot across these six sites we therefore ascertain a better understanding of the impacts of environmental effects and seasonality on a single greenspace scentscape.

Our work aimed to assess three main hypotheses: (i) that a sites location is a significant driver of volatile difference enabling the separation of greenspaces by their scentscape, (ii) that urban greenspace volatile profiles change over time dependent on environmental variables but remain characteristically distinctive, and iii) there is sufficient distinction between the presence and abundance of both bVOCs and aVOCs between urban green spaces that it is possible to determine which may be more healthy/unhealthy for human recreation.

4. Materials and Methods

4.1 Site locations and sampling strategy

Six urban greenspace sites around Oxford were chosen as sampling sites due to their (i) public accessibility, (ii) different ecologies, and (iii) contrasting distance from the city centre. We selected five outdoor greenspaces (Botanic Garden Outdoors, University Parks Dense, University Parks Open, Warneford Meadow, and Wytham Woods) and one indoor site (the Rainforest Glasshouse at the Oxford Botanical Gardens) (Figure 1A).

To investigate the characteristic VOC profiles of the selected sites, we employed a collection protocol adapted from Walker et al (2023) with the addition of housing the pump and apparatus within a bespoke metal box (VOKSBOX, B). Ambient air was captured using a GilAir® pump connected

simultaneously to 4 replicate Markes Tenax® TA tubes using a flow rate of 400ml/min for 2 hours (12 L of air per tube). Samples were collected at a height of 1.5m – based on the approximate level of the adult human nose while walking - between the hours of 11:00 - 13:00 (+/- 1 hour) GMT. Tubes were capped immediately after collection and stored at - 20 °C for up to 14 days.

A pilot dataset was collected on 27 November 2024 and used to assess the hypothesis that sampling sites have distinct volatile profiles (Figure 2). As well as looking at a snapshot of the scentscapes across the six urban greenspaces, we also wanted to explore how these scentscapes change across time and in the face of environmental variation (Figure 3). Because of the known impacts of seasonality on bVOC emissions we focused only on biogenic volatiles. Due to the higher numbers of biogenic volatiles present in the Botanical Outdoors location, and as this was likely to be the site with the highest footfall, we also sampled this site on 20 January 2025, 04 March 2025, 30 April 2025, 30 June 2025, 24 July 2025, and 04 September 2025. This allowed us to capture a range of temperature, humidity, precipitation, and wind conditions (Figure 3A). All Botanical Outdoors samples during the year were taken from 11:00 to 13:00 in exactly the same GPS location and height.

Site-specific temperature and relative humidity at the time of sampling were recorded using a Kestrel Drop D2 data logger. Historical meteorological data were obtained from multiple sources: air temperature data were retrieved from the Oxford Radcliffe Weather Station, and mean values recorded between 11:00 and 13:00 were used to coincide with the sampling window. Precipitation, humidity, wind speed, and wind direction data were obtained from the Kidlington weather station via Weather Underground (www.wunderground.com).

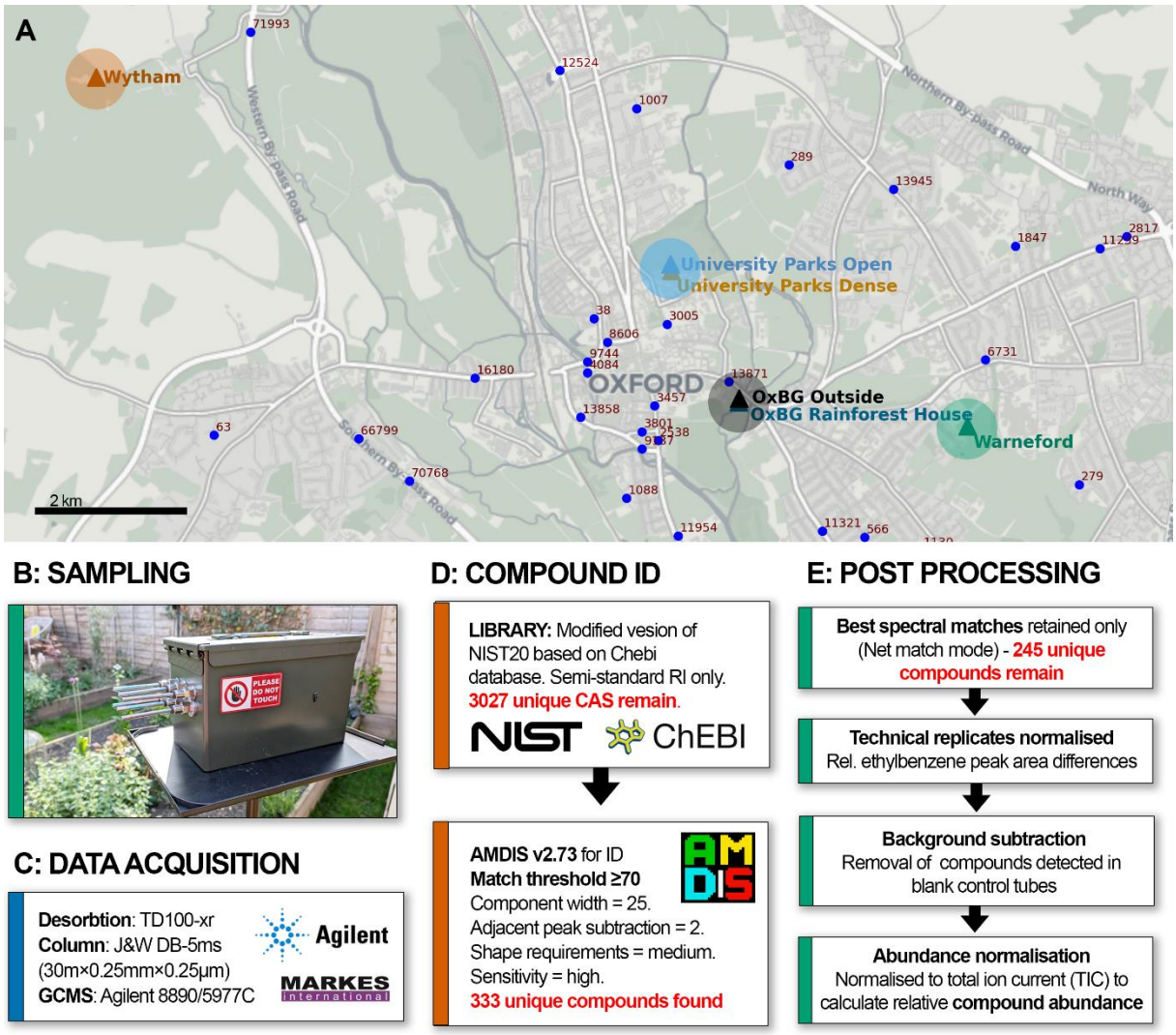

Figure 1: A: Map of the sampling locations used in this study. (Botanic Garden Outdoors; Botanic Garden Rainforest Glasshouse; University Parks Dense; University Parks Open; Warneford Meadow; Wytham Woods). Numbers on roads indicate Annual Average Daily Traffic (AADF) for 2023 as representative data for each site (GOV.UK 2023) B: Sampling box (VOKSBOX) developed for this work on a tripod at 1.5m height. Air from the GilAir® sample pump is split up to 6 ways using brass splitters with silicone tubing. C-E: A flow chart of the methods used in this paper for acquisition, identification, and processing of compounds. Red text shows number of compounds identified after each processing step.

4.2 Identification and analysis of VOC profiles

GC-MS Protocol: For identification of VOC profiles (Figure 1C-E), we employed gas chromatography mass spectrometry (GC-MS) analysis at the Jodrell laboratory, Royal Botanic Gardens, Kew. Samples were stored at room temperature for at least 2 hours before desorption for 10 mins at 230 °C using a TD100-xr automated thermal desorption (TD) unit from Markes. Samples were prefocused on a Tenax® TA packed cold trap at 4 °C for 0.5 mins followed by heating from 4 °C to 300 °C over 5 mins at a purge flow rate of 50ml/min onto a J&W DB-5ms GC column (30 m × 0.25 mm × 0.25 µm) within an Agilent 8890 GC System. Chromatography was achieved by heating the column at an initial

temperature of 40 °C at a rate of 3 °C/min to 220 °C at a flow rate of 1 ml/min Helium. Data were collected in electron ionization (EI) mode using an Agilent 5977C GC/ MSD over a mass range of 50–400 amu.

Library Development: The GC-MS library used for identification of compounds was a modified version of the NIST20 Wiley Registry 12th edition mass spectral library minimised to compounds existing in the Chemical Entities of Biological Interest (ChEBI) database (Degtyarenko et al. 2007). Only compounds with entries for retention indexes (RI) on semi-standard non-polar columns (to match the J&W DB-5ms column used) were selected. To avoid mismatches, isomers with the same RI were combined (e.g. (−)-alpha-pinene / (+)-alpha-pinene, became alpha-pinene). The final library used (W12N20MAIN_chebi_WK_edit.MSP) contained 3,027 unique compounds.

Pre-Analysis: To aid in the correct identification of compounds, a non-isothermal Kovats index (Kovats 1958) was built using retention times from C8 – C20 alkane standards (Merck, 04070-1ML). This index uses known compounds to produce accurate estimations of where all compounds will elute from the GC column (retention times) thus allowing precise identification. We validated this index using 60 pure terpene standards (Restek, 34142 and 34143).

Compound Identification: Sample analysis was performed using the Automated Mass Spectral Deconvolution and Identification System (AMDIS) version 2.73. Compound identification relied on spectral and RI matching, with a minimum match factor threshold of 70 (Net), a RI window of +/- 10, and a very strong match factor penalty. Deconvolution settings were component width: 25; adjacent peak subtraction: two; shape requirements: medium; sensitivity: high; and resolution: medium. The analysis was restricted to a single identification per library compound.

Post Processing: The relative amount (%) of each compound within a sample was used for all analysis in this paper. This was calculated as based on the Total Ion Count (TIC) of the sample / peak area of each compound. No absolute quantification was carried out in this work due to not having concentration curves for all compounds identified. Best-match compounds for each retention time and replicate tube were then chosen based on Net match score. Peak areas of any compounds identified in blank control tubes were removed from the peak areas of all samples within a GC-MS

run. Missing values for compounds (i.e. those compounds found only in some technical replicate tubes) were zero-filled.

4.3 Statistics and data visualisation

For the single November 2024 *snapshot* sampling event across six sites, multivariate analyses were used to explore patterns in volatile organic compound (VOC) composition across the sampling locations. Exploratory ordinations were performed using all 245 retained compounds following post-processing and filtering. Prior to multivariate analysis, compound abundance data were square-root transformed to reduce the influence of highly abundant compounds and to mitigate zero inflation, following established ecological recommendations (Legendre and Legendre 2012; 2013; Chang and Lin 2011).

For Principal Component Analysis (PCA, Figure 2A) transformed data were mean-centred and scaled to unit variance to ensure comparability among compounds with differing absolute abundances. PCA was conducted using the FactoMineR package with visualisation provided by factoextra in R (Kassambara and Mundt 2017; Lê, Josse, and Husson 2008). To assess compositional differences in VOC profiles among sampling locations, a permutational multivariate analysis of variance (PERMANOVA) was performed using the adonis2 function from the vegan R package, based on Euclidean distances and 9,999 permutations (Oksanen 2015).

Hierarchical Cluster Analysis (HCA, Figure 2B) was conducted using Ward's minimum variance linkage method with Euclidean distance, a combination that favours the formation of compact and internally coherent clusters. Cluster structure identified by HCA was subsequently used to define groups for downstream exploratory analyses. Partial Least Squares Discriminant Analysis (PLS-DA) was applied as a supervised exploratory method to characterise compounds contributing most strongly to differences among the three major clades identified by HCA. PLS-DA was implemented using the pls and mixOmics R packages (Rohart et al. 2017). To minimise over-interpretation, PLS-DA results were not used for hypothesis testing but to identify patterns in variable contributions. The top 10 compounds with the highest Variable Importance in Projection (VIP) scores for each clade are reported in Table 1.

To examine differences in VOC composition using ecologically interpretable groupings, compounds were classified into six functional classes: aliphatics, aromatics, monoterpenes, sesquiterpenes, terpenoids, and others (including esters and nitrogen- and sulphur-containing compounds) (Figure 2C and D). For each class, compound richness (number of compounds) and summed relative abundance were calculated per location. One-way ANOVAs were used to test for differences among locations, with post-hoc Tukey's HSD tests applied to identify pairwise differences in both compound richness (Table S 4 & Table S 5), or summed abundance (Table S 2 & Table S 3).

Spatial variation in terpene abundance was visualised using a hierarchical heatmap (Figure 2E) in which colour intensity represents compound abundance scaled within individual locations.

To contextualise VOC profiles in relation to anthropogenic influence, the relative abundance of all terpenes and terpenoids was expressed as a ratio to the summed relative abundance of anthropogenic BTEX compounds (benzene, toluene, ethylbenzene, and p-xylene; ∑terpenes/∑BTEX), following established approaches for assessing biogenic versus anthropogenic dominance (Malik et al. 2025).

All statistics were run using R version 4.4.1 (2024-06-14 ucrt) within RStudio 2024.09.0+375 "Cranberry Hibiscus" release for windows. All graphs were generated using the ggplot2 (Wickham 2011) package.

4.4 Statistical modelling of meteorological drivers of bVOC emissions

To investigate environmental drivers of bVOC emissions within a single greenspace (Botanical Gardens Outdoors) across the calendar year, linear mixed-effects models were fitted to compound class–level emission data from the Oxford Botanic Garden Outdoor site (BO) using R (v.4.x). Peak areas were averaged across technical replicates for each sampling date prior to analysis. Compounds were grouped into five functional classes: monoterpenes, sesquiterpenes, terpenoids, aromatics, aliphatics, and others (esters, nitrogen-, sulphur- and halogen-containing compounds), and summed abundances were calculated for each class per sampling date. Response variables were log-transformed using log(1 + x) to stabilise variance and reduce right skew. Predictor variables included mean daily temperature (°C), relative humidity (%), daily precipitation (mm), and wind speed (m s^{-1}).

All predictors were centred and scaled (z-score transformation) prior to modelling to facilitate comparison of effect sizes. For each compound class, a linear mixed-effects model was fitted using the lme4 package, with sampling date included as a random intercept to account for repeated measurements over time:

$$log(1 + yi) = \beta 0 + k\sum \beta kXk, i + bi + \varepsilon i$$

where $y_i$is the summed abundance of a given compound class on sampling date i, $X_k$are scaled meteorological predictors, $b_i$is the random effect of sampling date, and $\varepsilon_i$is the residual error.

Statistical significance of fixed effects was assessed using Wald tests implemented via the lmerTest package. Model coefficients and standard errors were extracted for visualisation as coefficient plots and heatmaps, with colour intensity representing the direction and magnitude of standardised effects. Model structure was defined a priori based on hypothesised meteorological drivers, and model fit was assessed to ensure consistency across compound classes.

5. Results

5.1 A snapshot in time: distinctive fingerprints of six urban greenspaces

Our analysis pipeline resulted in 245 unique compounds being identified within our 6 Oxford sampling sites (Table S 1). A PCA (Figure 2A) using all identified compounds as predictor variables revealed significant site-specific clustering (PERMANOVA, $F_{5,16} = 8.889$, $p = 0.001$) with pair-wise separation of all sites ($p < 0.05$) except Wytham Woods (WY) in the Warneford Meadow (WA, PERMANOVA, $p = 0.846$). Further HCA analysis further revealed site-level structuring of most VOC profiles (Figure 2B). Primarily, University Parks Open (UO) and Botanical Gardens Outdoors (BO) separate from all other sites (Figure 2B Clade i). PLS-DA VIP analysis (Table 1 and Table S7), due to higher amounts of many anthropogenic emission-related compounds such as 2,3,3-trimethylpentane, p-xylene, 2-ethylhexan-1-ol, indane, and ethylbenzene. Plant related compounds (beta-ocimene, dihydromyrcenol, sabinene, diacetone alcohol, beta-pinene, pinocarvone) were chiefly responsible for the secondary clustering (Figure 2B Clade ii) due to presence (or higher

proportions) in the Rainforest Glasshouse (RH) compared with University Parks Dense (UD), Wytham Woods (WY), and Warneford Meadow (WA). The third most significant clade (Figure 2B Clade iii) splits University Parks Open (UO) and Botanical Gardens Outdoors (BO), again due to the higher proportion/presence of terpenes/terpenoids such as 2-thujene, farnesane, and bornyl acetate found in Botanical Gardens Outdoors. Full results of PLS-DA analysis can be found in Table S7.

Table 1: Top 10 compounds from PLS-DA VIP analysis of the first 3 clades of the HCA analysis in Figure 2B: Bold compounds highlight terpenes/terpenoids.

Clade i
P-xylene (Aromatics), Ethylbenzene (Aromatics), Isooctane (Aliphatics), 2,3,4-trimethylpentane (Aliphatics), Toluene (Aromatics), 3-methylhexane (Aliphatics), Methylcyclohexane (Aliphatics), 2-dodecene (Aliphatics), 1-ethyl-2-methylbenzene (Aromatics), 3-methylheptane (Aliphatics)
Clade ii
Beta-ocimene (Monoterpenes), Dihydromyrcenol (Terpenoids), Sabinene (Monoterpenes), Beta-pinene (Monoterpenes), Pinocarvone (Terpenoids), Hydrindane (Aliphatics), **Alpha-pinene (Monoterpenes), Copaene (Sesquiterpenes), 4-terpineol (Terpenoids), Beta-phellandrene (Monoterpenes**)
Clade iii
2-thujene (Monoterpenes), Prehnitene (Aromatics), **Farnesane (Sesquiterpenes),** P-isopropenylmethylbenzene (Aromatics), Bromodichloromethane (Others (Esters, N, S, Halo)), 1,4-diethylbenzene (Aromatics), **Tricyclene (Monoterpenes),** 2,2,4,6,6-pentamethylheptane (Aliphatics), **Bornyl acetate (Terpenoids),** Propylbenzene (Aromatics)

5.1.1 Urban greenspaces show distinctive mix of compound classes

The number of identified terpenoids was found to be highest in the Rainforest Glasshouse (13), whereas monoterpenes and sesquiterpenes were found to be significantly higher in both the Rainforest Glasshouse (15, 10) and Botanical Gardens Outdoors (13, 2) compared to the other 4 sites (one-way ANOVAs followed by Tukey's HSD post-hoc test $p < 0.05$, Figure 2C - full statistics in Table S 4 & Table S 5). In terms of relative abundance (Figure 2B - full statistics in Table S 2 & Table S 3), Botanic Garden Outdoors contained a significantly higher proportion of aromatic compounds compared to other sites suggesting a strong influence from anthropogenic sources. The Rainforest Glasshouse showed significantly higher proportion of monoterpenes, whereas sesquiterpenes were shared by the Rainforest Glasshouse and Botanical Gardens Outdoors (One-way ANOVAs followed by Tukey's HSD post-hoc test $p < 0.05$).

5.1.2 Presence/absence of plant-based terpenes and terpenoids

A total of 54 terpenes/terpenoids were identified in this work (Figure 2G). Ten compounds - 1-8 cineole (eucalyptol), 3-carene, alpha-calacorene, alpha-pinene, beta-pinene, camphene, camphor, limonene, menthol, and p-cymene - were detected at every site, with a further two – menthone and tricyclene – detected in five of the six sampling sites. Again, the Rainforest Glasshouse exhibited the largest number of identified terpenes/terpenoid compounds overall (38 compounds) of which 16 were not detected at other sites. Of the outdoor sites, Botanical Gardens Outdoors contained the highest number of identified terpenes/terpenoid (23), of which 6 were uniquely present at this site. University Parks Dense (17), University Parks Open (14), Warneford Meadow (14) and Wytham Forest (22) all showed an overlapping but distinctive terpene mix. A full breakdown of compounds unique to individual sites is shown in Table S 6.

5.1.3 BTEX compounds as predictors of pollution

The ratio of biogenic terpenes/terpenoids to anthropogenic BTEX compounds was significantly higher in the Rainforest Glasshouse compared to the other sites (Figure 2E, one-way ANOVA followed by Tukey's HSD post-hoc test $p < 0.001$). This is in part due to the increased prevalence of terpenes/terpenoids and in part to the lower proportion of BTEX compounds (Figure 2F) in this indoor space. The outdoor sites all had a significantly lower ratios but were not significantly different from each other.

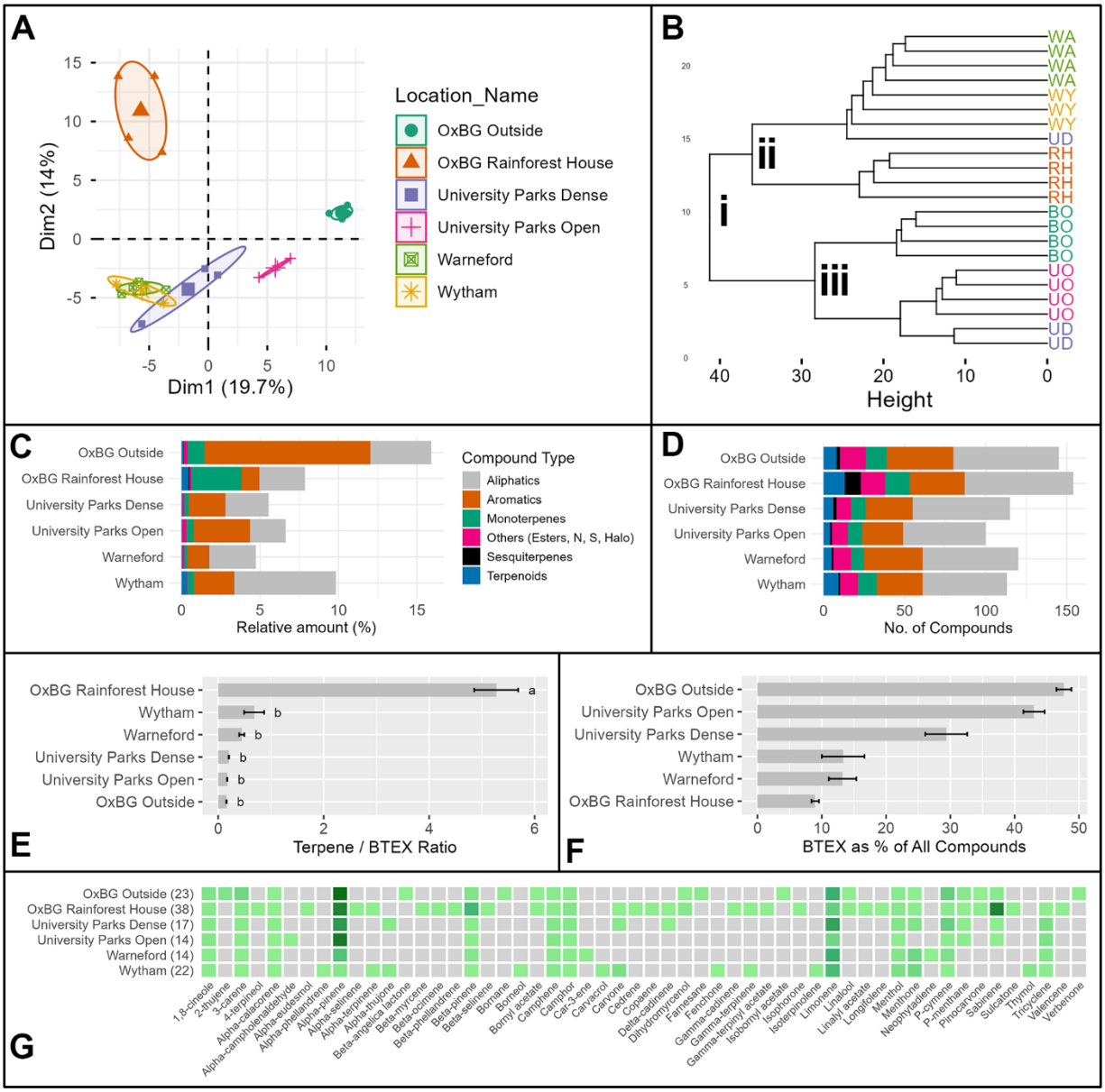

Figure 2: A: PCA which used all 245 identified compounds as predictor variables revealed significant variation between sites, each showing distinctive site-specific clustering (PERMANOVA, $F_{5,16} = 6.557$, $p = 0.001$). B: HCA conducted using Ward's linkage and Euclidean distances using all technical replicate samples over 6 urban greenspaces. C: Relative abundance (means) of detected volatile compounds split by classes across sites, expressed as a percentage of the total sample. D: Total number of detected volatile compounds per site, categorized by compound class. Results show that the number of all compounds groups differed between sites. E: Terpene: BTEX ratio of sampling sites (∑terpenes/∑BTEX). Relative amounts of all terpenes/terpenoids as a ratio to the anthropogenic BTEX compound. F: ∑BTEX as % of all identified compounds between sampling sites. G: Terpenes/terpenoid compounds identified across sampling locations. Heatmap showing the scaled relative amount (light green > dark green) or absence (grey) of any terpenes/terpenoid VOCs detected across six sampling locations. Scaling was carried out within single locations. C, D, F: Bars represent the mean (n = 4 technical replicate sampling tubes per site). Letters indicate statistically distinct sampling locations within chemical classes based on one-way ANOVA followed by Tukey's HSD post-hoc test ($p < 0.05$).

5.2 How do urban greenspace bVOC profiles vary through time under real-world conditions?

Using the same classification system for our identified compounds we counted both compound numbers and summed abundances to visualise each class over time. In terms of the relative abundances, all compound classes peaked during late spring to summer (May to August 2025) with two notable outliers - monoterpenes had an early peak on November 27th, 2024, and terpenoids had a late peak on October 29th, 2025. Interestingly, in terms of the absolute number of compounds identified, it was early spring (04 March 2025) which showed peaks for aliphatics (42), aromatics (14), terpenoids (10) and others (9), with monoterpene counts being more stable over the year.

Mixed-effects modelling revealed that different components of the biogenic scentscape responded in contrasting ways to short-term environmental variation (Figure 3C). Temperature emerged as the most consistent positive driver of terpene emissions, with higher temperatures associated with increased relative abundances of both monoterpenes and sesquiterpenes. Humidity also showed compound-class-specific effects with terpenes exhibiting a significant positive association, whereas aromatic and aliphatic compounds were weaker or neutral. Wind speed contributed to increases in both aromatic and aliphatic fractions.

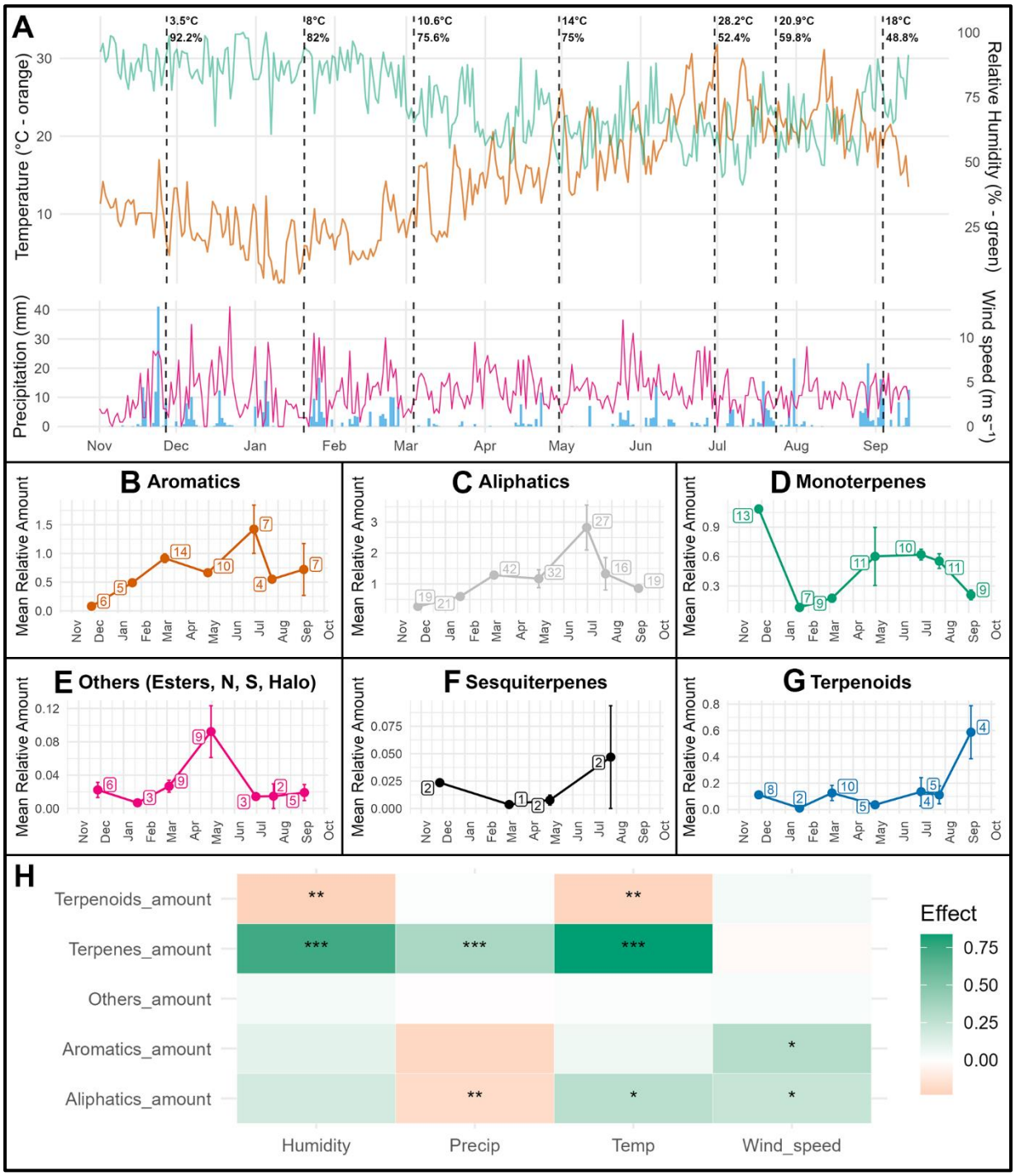

Figure 3: Seasonal dynamics and environmental drivers of biogenic volatile organic compounds (bVOCs) at the Botanical Garden outdoor site. (A) Daily temperature (orange) and relative humidity (green), and daily precipitation (bars) with wind speed (line), derived from local and regional meteorological datasets. Vertical dashed lines indicate sampling dates; values above denote site-specific temperature and relative humidity measured during sampling. (B-G) shows monthly mean relative abundances (± SE) of compound classes (aliphatics, aromatics, monoterpenes, sesquiterpenes, terpenoids, and "others" including esters and N-, S- and halogen-containing compounds) from November 2024 to September 2025. Numbers beside points indicate the number of detected compounds contributing to each class per sampling event. (H) Mixed-effects model coefficients linking standardised environmental predictors (temperature, precipitation, humidity and wind speed) to log-transformed relative abundances of each compound class. Tile colour represents effect size (positive = green, negative = brown), and symbols denote statistical significance (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Models were fitted with sampling date as a random effect to account for temporal structure.

6. Discussion

6.1 bVOCs and aVOCs shape urban greenspace scentscapes

An initial PCA of our six sampling sites on a single date (Nov 27th, 2024) revealed significant site-specific clustering. Further HCA revealed site structuring was due to both bVOCs and aVOCs (Table 1, Figure 2B). We chose to interrogate the first 3 clades which appeared. The first clade separated two outdoor sites - Botanical Gardens Outdoors and University Parks Open – from all other sampling sites. PLS-DA VIP analysis showed that this was due to many emission-related aVOCs (Table 1) including the traffic-related BTEX compounds (Figure 2F). These two sites are very close to a busy city centre roads (Figure 1A) suggesting that traffic is the likely main contributor. This is supported by BTEX presence measurements where over 40% of identified volatiles were BTEX compounds in the Botanical Gardens Outdoors and University Parks Open, but just 13% in Wytham Woods and Warneford Meadow which are more isolated from traffic. The relationship between proximity to traffic and BTEX prevalence has not been assessed in this paper but has been previously reported (Hoque et al. 2008).

Interestingly, despite its proximity to University Parks Open (~80 m), University Parks Dense showed markedly lower BTEX levels (43% to 29%). Localised differences in BTEX presence within urban greenspaces has been demonstrated before; for example, Upmanis et al. (2001) reported that concentrations of benzene and toluene were up to one-third lower just 40m into urban parkland within two large Scandinavian cities. A second study also showed decreases in particulate matter (PM_{10} and $PM_{2.5}$) of up to 50% at 200m (Gómez-Moreno et al. 2019). Additionally, the presence of exposed soil and leaf litter layer around trees in University Parks Dense compared to the grassy University Parks Open site may contribute to increased degradation of BTEX compounds through microbial metabolism. A 2012 study, for example, showed that by adding a mixed bacterial consortium to soil, BTEX compounds were more rapidly degraded (Mukherjee and Bordoloi 2012).

Biogenic volatiles, particularly terpenes (Table 1, Figure 2G), were chiefly responsible for the secondary clustering of samples from Rainforest House, and those of Wytham Woods and Warneford Meadow. In total across all sampling sites, we identified 54 unique terpenes/terpenoids. The

Rainforest Glasshouse contained the highest number of terpenes (38) with considerable spread across the outdoor sampling sites (14 to 23). The presence of these compounds is highly suggestive of multiple and specific health benefits within each location. For example, 10 terpenes/terpenoids were ubiquitous across all sampling sites - compounds which have also been found in air samples elsewhere using similar methods (Gu et al. 2024; Bryant et al. 2022; Walker et al. 2023; Hellén et al. 2024; Jones, Hopkins, and Lewis 2011; Hakola et al. 2003; Detournay et al. 2011). These compounds alone are cited to have anti-inflammatory, anxiolytic, antimicrobial, immune-boosting, antidepressant, anaesthetic, gastroprotective, antioxidant, anti-cancer (*in vitro*), immune-modulatory, sedative, hypolipidemic, analgesic, decongestant, bronchodilatory, mucolytic, and cognitive enhancing effects (Lee et al. 2022; Vallianou and Hadzopoulou-Cladaras 2016; Salehi et al. 2019; Oliveira et al. 2024; Sun 2007; Woo et al. 2019; Jo et al. 2021). It could be argued, therefore, that spending time in any of these sites could impart significant health gains. However, the relationships between these compounds and their benefits are likely concentration dependent, therefore further methodological work is needed to quantify and make precise recommendations such as when, and for how long, to visit specific greenspaces. Previously it was found that just 20 minutes is adequate to derive physiological and psychological benefits from time in woodland (Haluza et al. 2025; Park et al. 2010).

Another useful metric to separate our sites is a ratio of biogenic to anthropogenic compounds. An example of this can be found in study which assessed air over Covid lockdown in 2020 at an urban site in India (Malik et al. 2025) in which a sharp increase in the ∑α+beta--pinene/∑BTEX ratios during the strict lockdown phase was hypothesised to be due to a reduction in traffic related activity. We mirrored this study and found that, on a particular day, the Rainforest House had significantly higher terpene/BTEX ratio than all the other sites (Figure 2F). This result is likely due to the interplay of two different factors known to affect the volatile makeup of indoor spaces. On the one hand, this glasshouse is plant rich and enclosed and as such likely acts as an enclosure for those bVOCs produced within, and conversely, aVOC ingression may be limited. These results suggest that plant rich greenhouses may be the most beneficial place to spend time, however, in a 2006-2007 study, aVOCs were found to be significantly higher in indoor sites than outdoor sites in the vicinity of an

industrial park (Chang et al. 2019), showing that simply being inside is little protection from external sources of pollution.

6.2 The impacts of seasonality on bVOC emissions in the Botanical Gardens

As well as looking at a snapshot of the scentscapes across the six urban greenspaces, we also explored how bVOC change across time and in the face of environmental variation in the Botanical Outdoors location. Using the same classification system for our identified compounds we counted both compound numbers and summed abundances to visualise each class over time. In terms of the relative abundances, all compound classes peaked during late spring to summer (May to August 2025) with two notable outliers - monoterpenes had an early peak on November 27th, 2024, and terpenoids had a late peak on October 29th, 2025. Interestingly, in terms of the absolute number of compounds identified, it was early spring (March XX) which showed peaks for aliphatics (42), aromatics (14), terpenoids (10) and others (9), with monoterpene counts being more stable over the year.

Temperature and humidity were the most consistent positive predictors of biogenic emissions, particularly for terpenes. For temperature this is likely linked to the well-established positive relationship between emission rates and leaf temperatures via both biosynthesis through increased enzymatic activity and volatilisation vapour pressure (Guenther et al. 1993; Luo et al. 2025). Aliphatics also showed a positive temperature response, likely driven by enhanced volatilisation at higher temperatures but also resistance to biodegradation (Abbasian et al. 2015). In contrast, terpenoids, which are oxygenated terpenes and much more reactive, declined with increasing temperature. This is consistent with their greater susceptibility to photochemical degradation at higher temperatures (Atkinson and Arey 2003; Hallquist et al. 2009). The role of humidity in terpene/terpenoid presence may be linked to an increase in stomatal opening and subsequent environmental release (Arve and Torre 2015).

Precipitation effects also differed among VOC classes. Terpene abundances showed a positive association with precipitation. Wetting is known to give a short-term pulse release through mechanical displacement from soil air spaces after rain (Pugliese et al. 2023) and this was likely the

390 reason for the high amount recorded on November 27th, 2024, where 8.10 mm of rain was recorded.
391 In contrast, aromatics and aliphatics showed a negative association with precipitation on the day of
392 sampling, consistent with atmospheric dilution, wet deposition, or reduced resuspension of
393 background hydrocarbons rather than changes in emission strength (Casas et al. 2021). There was a
394 second rainfall event on September 4th, 2025 (9.4 mm), however, this followed 16.3 mm of rain on
395 the previous day which is likely to have negatively affected soil-based terpene release on the day of
396 sampling. Notably, only two sampling events coincided with measurable rainfall so these patterns
397 need further exploration in urban greenspaces.

398 Wind speed effects showed significant positive associations for aromatic and aliphatic compounds.
399 Given their relatively long atmospheric lifetimes compared to terpenes and terpenoids (Luo et al.
400 2025; Hellén et al. 2018), these increases are more likely to reflect regional background
401 concentrations than increases in local emissions. In contrast, terpene emissions showed little
402 sensitivity to wind speed, consistent with their predominantly local, vegetation-derived sources and
403 rapid atmospheric reactivity leading to degradation if produced distantly.

404 Overall, these results indicate that temperature, precipitation, wind speed and humidity all impact on
405 the detection of different bVOC chemical classes. The absence of a single dominant predictor across
406 all VOC classes highlights the importance of considering scentscapes as multidimensional chemical
407 systems, shaped by interacting biological and meteorological processes. However, this class-specific
408 sensitivity to environmental drivers adds understanding of when and where urban greenspaces may
409 be most chemically beneficial for human exposure. These results are suggestive that warm, humid,
410 and (unfortunately) rainy conditions are the most beneficial time to spend time in urban
411 greenspaces.

412 6.3 Methodological advancements for the analysis of greenspace 413 scentscapes

414 As well as providing an understanding of urban greenspace scentscapes over time, this study aimed
415 to develop a robust and repeatable protocol for measuring the presence and relative amounts of a
416 wide range of volatiles of interest. Although similar methods have been developed for headspace

sampling (Yin et al. 2022; de Carvalho Couto et al. 2024; Reyrolle et al. 2024), none have yet been published for the analysis of ambient air.

VOC diversity in ambient air is known to be vast and highly location-dependent, often comprising thousands of compounds. Many air sampling studies have historically employed targeted approaches that significantly limit the number of compounds identified. For instance, previous studies have used between 15 and 66 VOCs across various forest and urban environments (Walker et al. 2023; Pripdeevech et al. 2025; Mula et al. 2024). In contrast, our study adopted an untargeted metabolomics-style approach commonly used in biological matrices to enhance the depth and resolution of VOC profiling (Fernie et al. 2011; Van den Berg et al. 2006; Rosenthal et al. 2024) to confidently identify 245 unique compounds, of which 54 were unique terpenes/terpenoids. This is significantly higher than any studies on ambient air so far published (Walker et al. 2023; Sanaei et al. 2023; Mula et al. 2024). This improved identification capacity is likely attributable to several key enhancements in our analytical pipeline: greater technical replication, the use of a Kováts retention index, a more focused compound library, a slow GC temperature ramp, and the use of deconvolution for resolving co-eluting peaks. However, even with these methods we only identify around 10% of the compounds present in our samples. Longer sampling times, a further development of our GCMS library and the running of standards to increase available retention index values are all areas for future improvement.

Using our methods we conclude that (i) at a snapshot in time, site location is a significant driver of volatile difference, and (ii) that environmental conditions affect volatile presence in a strongly compound-class-specific manner, allowing us to make recommendations to maximise human recreational benefit.

7. Author Contributions

William Kay was responsible for methods development, data collection and all data analyses. William Kay and Anya Lindstrom Battle were both involved in the writing and editing of the manuscript. Kathy Willis conceived of the initial ideas and methodologies and was also involved in the writing and

editing of the manuscript. Mahal Humberstone, Molly Tucker and Kieran Storer all contributed to in-field sample collection. Geoffrey Kite was involved in the GC-MS data collection, as well as providing useful guidance. All authors contributed critically to the drafts and gave final approval for publication.

8. Funding

This work was funded by the Leverhulme Centre for Nature Recovery, made possible thanks to the generous support of the Leverhulme Trust - Grant number RC-2021-76.

9. Data Availability Statement

Raw AGILENT files, and GCMS data files used to analyse data are available here: https://unioxfordnexus-my.sharepoint.com/:f:/r/personal/dops0827_ox_ac_uk/Documents/KW_GCMS_Oxford/_Papers/Paper%201%20-%20Kay%20et%20al.%20A%20Walk%20in%20the%20Park?csf=1&web=1&e=yjWJF2

10. Acknowledgments

We would firstly like to thank the University of Oxford Botanic Garden and Wytham Woods for permission to sample on their sites. We would also like to thank Dr Ben Martin who provided helpful advice for sampling and analysis as well as the development of early sampling methodologies, and Dr Luciana Carvalho who always had time to share thoughts and ideas relating to the project. Finally, we thank Dr Pedro Bota for his generosity of time, availability and equipment.

11. Conflicts of Interest

The authors declare no conflicts of interest.

12. References

Abbasian, F., Lockington, R., Mallavarapu, M., and Naidu, R. (2015). A Comprehensive Review of Aliphatic Hydrocarbon Biodegradation by Bacteria, *Applied Biochemistry and Biotechnology*, 176: 670–99. https://doi.org/10.1007/s12010-015-1603-5

Arve, L. E., and Torre, S. (2015). Ethylene is involved in high air humidity promoted stomatal opening of tomato (Lycopersicon esculentum) leaves, *Functional Plant Biology*, 42: 376–86. https://doi.org/10.1071/fp14247

Atkinson, R., and Arey, J. (2003). Atmospheric Degradation of Volatile Organic Compounds, *Chemical Reviews*, 103: 4605–38. https://doi.org/10.1021/cr0206420

Bratman, G. N., Daily, G. C., Levy, B. J., and Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition, *Landscape and urban planning*, 138: 41–50. https://doi.org/10.1016/j.landurbplan.2015.02.005

Bryant, D. J., Nelson, B. S., Swift, S. J., Budisulistiorini, S. H., Drysdale, W. S., Vaughan, A. R., Newland, M. J., Hopkins, J. R., Cash, J. M., and Langford, B. (2022). Biogenic and anthropogenic sources of isoprene and monoterpenes and their secondary organic aerosol in Delhi, India, *Atmospheric Chemistry and Physics Discussions*, 2022: 1–35. https://doi.org/10.1016/j.apr.2019.07.004

Casas, G., Martinez-Varela, A., Vila-Costa, M., Jiménez, B., and Dachs, J. (2021). Rain Amplification of Persistent Organic Pollutants, *Environmental Science & Technology*, 55: 12961–72. https://doi.org/10.1021/acs.est.1c03295

Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines, *ACM transactions on intelligent systems and technology (TIST)*, 2: 1–27. https://doi.org/10.1145/1961189.1961199

Chang, T.-Y., Liu, C.-L., Huang, K.-H., and Kuo, H.-W. (2019). Indoor and outdoor exposure to volatile organic compounds and health risk assessment in residents living near an optoelectronics industrial park, *Atmosphere*, 10: 380. https://doi.org/10.3390/atmos10070380

de Carvalho Couto, C., Chávez, D. W. H., Oliveira, E. M. M., Freitas-Silva, O., and Casal, S. (2024). SPME-GC-MS untargeted metabolomics approach to identify potential volatile compounds as markers for fraud detection in roasted and ground coffee, *Food Chemistry*, 446: 138862. https://doi.org/10.1016/j.foodchem.2024.138862

Degtyarenko, K., De Matos, P., Ennis, M., Hastings, J., Zbinden, M., McNaught, A., Alcántara, R., Darsow, M., Guedj, M., and Ashburner, M. (2007). ChEBI: a database and ontology for chemical entities of biological interest, *Nucleic acids research*, 36: D344–D50. https://doi.org/10.1093/nar/gkm791

Detournay, A., Sauvage, S., Locoge, N., Gaudion, V., Leonardis, T., Fronval, I., Kaluzny, P., and Galloo, J.-C. (2011). Development of a sampling method for the simultaneous monitoring of straight-chain alkanes, straight-chain saturated carbonyl compounds and monoterpenes in remote areas, *Journal of Environmental Monitoring*, 13: 983–90. https://doi.org/10.1039/c0em00354a

Fernie, A. R., Aharoni, A., Willmitzer, L., Stitt, M., Tohge, T., Kopka, J., Carroll, A. J., Saito, K., Fraser, P. D., and DeLuca, V. (2011). Recommendations for reporting metabolite data, *The Plant Cell*, 23: 2477–82. https://doi.org/10.1105/tpc.111.086272

Gao, X., Wang, Y., Wu, L., Zheng, F., Sun, N., Liu, G., Liu, Y., Meng, P., Sun, L., and Jing, B. (2023). The Impact of Anthropogenic VOC Emissions on Atmospheric Pollution: A Case Study of a Typical Industrialized Area in China, *Atmosphere*, 14: 1586. https://doi.org/10.3390/atmos14101586

Gómez-Moreno, F. J., Artíñano, B., Ramiro, E. D., Barreiro, M., Núñez, L., Coz, E., Dimitroulopoulou, C., Vardoulakis, S., Yagüe, C., Maqueda, G., Sastre, M., Román-Cascón, C., Santamaría, J. M., and Borge, R. (2019). Urban vegetation and particle air pollution: Experimental campaigns in a traffic hotspot, *Environmental Pollution*, 247: 195–205. https://doi.org/https://doi.org/10.1016/j.envpol.2019.01.016

GOV.UK. 2023. "Road traffic statistics - Local authority Oxfordshire." In, edited by Department for Transport. roadtraffic.dft.gov.uk: GOV.UK

Gu, S., Luo, W., Charmchi, A., McWhirter, K. J., Rosenstiel, T., Pankow, J., and Faiola, C. L. (2024). Limonene Enantiomeric Ratios from Anthropogenic and Biogenic Emission Sources, *Environ Sci Technol Lett*, 11: 130–35. https://doi.org/10.1021/acs.estlett.3c00794

Guenther, A., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T. a., Emmons, L., and Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions, *Geoscientific Model Development*, 5: 1471–92. https://doi.org/10.5194/gmd-5-1471-2012

Guenther, A. B., Zimmerman, P. R., Harley, P. C., Monson, R. K., and Fall, R. (1993). Isoprene and monoterpene emission rate variability: model evaluations and sensitivity analyses, *Journal of Geophysical Research: Atmospheres*, 98: 12609–17. https://doi.org/10.1029/93JD00527

Hakola, H., Tarvainen, V., Laurila, T., Hiltunen, V., Hellén, H., and Keronen, P. (2003). Seasonal variation of VOC concentrations above a boreal coniferous forest, *Atmospheric Environment*, 37: 1623–34. https://doi.org/10.1016/S1352-2310(03)00014-1

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: current and emerging issues, *Atmos. Chem. Phys.*, 9: 5155–236. https://doi.org/10.5194/acp-9-5155-2009

Haluza, D., Kersten, P., Lazic, T., Steinparzer, M., and Godbold, D. (2025). Unlocking the Power of Nature: Insights from a 20-Minute Forest Visit on Well-Being, *Forests*, 16: 792

Hellén, H., Praplan, A. P., Tykkä, T., Ylivinkka, I., Vakkari, V., Bäck, J., Petäjä, T., Kulmala, M., and Hakola, H. (2018). Long-term measurements of volatile organic compounds highlight the importance of sesquiterpenes for the atmospheric chemistry of a boreal forest, *Atmospheric Chemistry and Physics*, 18: 13839–63

Hellén, H., Tykkä, T., Schallhart, S., Stratigou, E., Salameh, T., and Iturrate-Garcia, M. (2024). Measurements of atmospheric C 10–C 15 biogenic volatile organic compounds (BVOCs) with sorbent tubes, *Atmospheric Measurement Techniques*, 17: 315–33. https://doi.org/10.5194/amt-17-315-2024

Hoque, R. R., Khillare, P., Agarwal, T., Shridhar, V., and Balachandran, S. (2008). Spatial and temporal variation of BTEX in the urban atmosphere of Delhi, India, *Science of the total environment*, 392: 30–40. https://doi.org/10.1016/j.scitotenv.2007.08.036

Jo, H., Cha, B., Kim, H., Brito, S., Kwak, B. M., Kim, S. T., Bin, B.-H., and Lee, M.-G. (2021). α-Pinene enhances the anticancer activity of natural killer cells via ERK/AKT pathway, *International Journal of Molecular Sciences*, 22: 656. https://doi.org/10.3390/ijms22020656

Jones, C. E., Hopkins, J., and Lewis, A. (2011). In situ measurements of isoprene and monoterpenes within a south-east Asian tropical rainforest, *Atmospheric Chemistry and Physics*, 11: 6971–84. https://doi.org/10.5194/acp-11-6971-2011

Joo, Y., Schuman, M. C., Goldberg, J. K., Kim, S. G., Yon, F., Brütting, C., and Baldwin, I. T. (2018). Herbivore-induced volatile blends with both “fast” and “slow” components provide robust indirect defence in nature, *Functional Ecology*, 32: 136–49. https://doi.org/10.1111/1365-2435.12947

Kassambara, A., and Mundt, F. 2017. "Package ‘factoextra’." In *Extract and visualize the results of multivariate data analyses*, 10.18637.

Kovats, v. E. (1958). Gas-chromatographische charakterisierung organischer verbindungen. Teil 1: retentionsindices aliphatischer halogenide, alkohole, aldehyde und ketone, *Helvetica Chimica Acta*, 41: 1915–32. https://doi.org/10.1002/hlca.19580410703

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis, *Journal of statistical software*, 25: 1–18. https://doi.org/10.18637/jss.v025.i01

Lee, S. H., Kim, D. S., Park, S. H., and Park, H. (2022). Phytochemistry and Applications of Cinnamomum camphora Essential Oils, *Molecules*, 27. https://doi.org/10.3390/molecules27092695

Legendre, P., and Legendre, L. 2012. *Numerical ecology* (Elsevier).

Li, Q. (2010). Effect of forest bathing trips on human immune function, *Environmental health and preventive medicine*, 15: 9–17. https://doi.org/10.1007/s12199-008-0068-3

Li, T., Baggesen, N., Seco, R., and Rinnan, R. (2023). Seasonal and diel patterns of biogenic volatile organic compound fluxes in a subarctic tundra, *Atmospheric Environment*, 292: 119430. https://doi.org/10.1016/j.atmosenv.2022.119430

Luo, R., Lun, X., Gao, R., Wang, L., Yang, Y., Su, X., Habibullah-Al-Mamun, M., Xu, X., Li, H., and Li, J. (2025). A Review of Biogenic Volatile Organic Compounds from Plants: Research Progress and Future Prospects, *Toxics*, 13: 364

Malik, T. G., Gupta, M., Tripathi, N., and Sahu, L. K. (2025). Change in monoterpene concentrations during winter-to-summer transition period and impact of COVID-19 lockdown at an urban site in India, *Atmospheric Environment*, 350: 121141. https://doi.org/10.1016/j.atmosenv.2025.121141

Mu, Z., Asensio, D., Llusia, J., Filella, I., Ogaya, R., Yi, Z., and Penuelas, J. (2022). Annual and seasonal variations in soil volatile organic compound concentrations in a Mediterranean shrubland and holm oak forest, *Geoderma*, 405: 115401. https://doi.org/10.1016/j.geoderma.2021.115401

Mukherjee, A. K., and Bordoloi, N. K. (2012). Biodegradation of benzene, toluene, and xylene (BTX) in liquid culture and in soil by Bacillus subtilis and Pseudomonas aeruginosa strains and a formulated bacterial consortium, *Environmental Science and Pollution Research*, 19: 3380–88. https://doi.org/10.1007/s00449-015-1364-0

Mula, V., Bogdanov, J., Stanoeva, J. P., Zeneli, L., and Zdravkovski, Z. (2024). Monitoring Volatile Organic Compounds in Air Using Passive Sampling: Regional Cross-Border Study between N. Macedonia and Kosovo, *Aerosol and Air Quality Research*, 24: 230170. https://doi.org/10.4209/aaqr.230170

Müller, A., Faubert, P., Hagen, M., Zu Castell, W., Polle, A., Schnitzler, J.-P., and Rosenkranz, M. (2013). Volatile profiles of fungi–chemotyping of species and ecological functions, *Fungal genetics and biology*, 54: 25–33. https://doi.org/10.1016/j.fgb.2013.02.005

Oksanen, J. 2015. "Vegan: community ecology package." In *R package version*, 3.

Oliveira, C. D., Cardoso, M. d. G., Batista, L. R., Alves, E., Rosa, M. B. P., Ferreira, V. R. F., de Souza, L., Pineda, M., Fernandes, A. I., and Nelson, D. L. (2024). The antibacterial, antioxidant, and insecticidal activities of essential oils from Thymus vulgaris L., Salvia officinalis L., and Ocimum basilicum L, *Journal of Food Safety*, 44: e13145. https://doi.org/10.1111/jfs.13145

Park, B. J., Tsunetsugu, Y., Kasetani, T., Kagawa, T., and Miyazaki, Y. (2010). The physiological effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan, *Environmental health and preventive medicine*, 15: 18–26. https://doi.org/10.1007/s12199-009-0086-9

Pripdeevech, P., Janta, R., Sripahco, T., Meesang, W., Aiyathiti, C., Prabamroong, T., Mahatheeranont, S., Poshyachinda, S., Pongpiachan, S., and Khruengsai, S. (2025). Seasonal volatile organic compound dynamics in urban and forest environments in Thailand: Implications for air quality and secondary pollutants, *Environmental Pollution*, 367: 125565. https://doi.org/10.1016/j.envpol.2024.125565

Pugliese, G., Ingrisch, J., Meredith, L. K., Pfannerstill, E. Y., Klüpfel, T., Meeran, K., Byron, J., Purser, G., Gil-Loaiza, J., and Van Haren, J. (2023). Effects of drought and recovery on soil volatile organic compound fluxes in an experimental rainforest, *nature communications*, 14: 5064. https://doi.org/10.1038/s41467-023-40661-8

Reyrolle, M., Desauziers, V., Pigot, T., Gautier, L., and Le Bechec, M. (2024). Comparison of Untargeted and Markers Analysis of Volatile Organic Compounds with SIFT-MS and SPME-GC-MS to Assess Tea Traceability, *Foods*, 13: 3996. https://doi.org/10.3390/foods13243996

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration, *PLoS computational biology*, 13: e1005752. https://doi.org/10.1371/journal.pcbi.1005752

Rosenthal, K., Lindley, M., Turner, M., Ratcliffe, E., and Hunsicker, E. (2024). Current data processing methods and reporting standards for untargeted analysis of volatile organic compounds using direct mass spectrometry: a systematic review, *Metabolomics*, 20: 42. https://doi.org/10.1007/s11306-024-02104-3

Salehi, B., Upadhyay, S., Erdogan Orhan, I., Kumar Jugran, A., LD Jayaweera, S., A. Dias, D., Sharopov, F., Taheri, Y., Martins, N., and Baghalpour, N. (2019). Therapeutic potential of α- and β-pinene: A miracle gift of nature, *Biomolecules*, 9: 738. https://doi.org/10.3390/biom9110738

Sanaei, A., Herrmann, H., Alshaabi, L., Beck, J., Ferlian, O., Fomba, K. W., Haferkorn, S., van Pinxteren, M., Quaas, J., and Quosh, J. (2023). Changes in biodiversity impact atmospheric chemistry and climate through plant volatiles and particles, *Communications Earth & Environment*, 4: 445. https://doi.org/10.1038/s43247-023-01113-9

Sun, J. 2007. "D-Limonene: safety and clinical applications." In *Alternative medicine review*, 259.

Upmanis, H., Eliasson, I., and Andersson-Sköld, Y. (2001). Case studies of the spatial variation of benzene and toluene concentrations in parks and adjacent built-up areas, *Water, air, and soil pollution*, 129: 61–81. https://doi.org/10.1023/A:1010357914047

Vallianou, I., and Hadzopoulou-Cladaras, M. (2016). Camphene, a Plant Derived Monoterpene, Exerts Its Hypolipidemic Action by Affecting SREBP-1 and MTP Expression, *PloS one*, 11: e0147117. https://doi.org/10.1371/journal.pone.0147117

Van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and Van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data, *BMC genomics*, 7: 1–15. https://doi.org/10.1186/1471-2164-7-142

Walker, H., Jena, A., McEwan, K., Evans, G., and Campbell, S. (2023). Natural volatile organic compounds (NVOCs) are greater and more diverse in UK forests compared with a public garden, *Forests*, 14: 92. https://doi.org/10.3390/f14010092

Wang, L., Lun, X., Wang, Q., and Wu, J. (2024). Biogenic volatile organic compounds emissions, atmospheric chemistry, and environmental implications: a review, *Environmental Chemistry Letters*, 22: 3033–58. https://doi.org/10.1007/s10311-024-01785-5

Weisskopf, L., Schulz, S., and Garbeva, P. (2021). Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions, *Nature Reviews Microbiology*, 19: 391–404. https://doi.org/10.1038/s41579-020-00508-1

Wickham, H. (2011). ggplot2, *Wiley interdisciplinary reviews: computational statistics*, 3: 180–85. https://doi.org/10.1002/wics.147

Woo, J., Yang, H., Yoon, M., Gadhe, C. G., Pae, A. N., Cho, S., and Lee, C. J. (2019). 3-Carene, a Phytoncide from Pine Tree Has a Sleep-enhancing Effect by Targeting the GABA_A-benzodiazepine Receptors, *Experimental neurobiology*, 28: 593–601. https://doi.org/10.5607/en.2019.28.5.593

Yang, K., Llusià, J., Preece, C., Ogaya, R., Tur, L. M., Mu, Z., You, C., Xu, Z., Tan, Y., and Peñuelas, J. (2024). Impacts of seasonality, drought, nitrogen fertilization, and litter on soil fluxes of biogenic volatile organic compounds in a Mediterranean forest, *Science of the total environment*, 906: 167354. https://doi.org/10.1016/j.scitotenv.2023.167354

Yin, X.-L., Fu, W.-J., Chen, Y., Zhou, R.-F., Sun, W., Ding, B., Peng, X.-T., and Gu, H.-W. (2022). GC-MS-based untargeted metabolomics reveals the key volatile organic compounds for discriminating grades of Yichang big-leaf green tea, *LWT*, 171: 114148. https://doi.org/10.1016/j.lwt.2022.114148

Zeng, W., Liu, H., Hou, S., Qiu, X., Chen, X., Liu, M., Wu, D., and Liu, L. (2022). Explore the Benefits of Natural Air: New Insights from Field and Chamber Tests on Cognitive Performance, *Atmosphere*, 13: 1006. https://doi.org/10.3390/atmos13071006

13. Supplementary Information

Table S 1: A full list of the compounds detected on 27th November, 2024. List is sorted by retention time (RT).

CAS	Name	RT
107-06-2	1,2-dichloroethane	4.5965
590-86-3	3-methylbutanal	4.6278
563-80-4	3-methylbutan-2-one	4.6306
71-36-3	Butan-1-ol	4.6986
71-43-2	Benzene	4.7208
56-23-5	Tetrachloromethane	4.7596
110-82-7	Cyclohexane	4.7915
565-59-3	2,3-dimethylpentane	4.8001
589-34-4	3-methylhexane	4.8718
107-87-9	Pentan-2-one	4.9875
110-83-8	Cyclohexene	5.0306
617-78-7	3-ethylpentane	5.0625
123-75-1	Pyrrolidine	5.0664
540-84-1	Isooctane	5.0978
96-22-0	Pentan-3-one	5.2349
142-82-5	Heptane	5.2872
10152-76-8	Sulfide, allyl methyl	5.3094
3208-16-0	2-ethylfuran	5.3103
79-01-6	Trichloroethene	5.3625
110-62-3	Pentanal	5.4846
80-62-6	Methyl methacrylate	5.501
109-60-4	Propyl acetate	5.5457
625-86-5	2,5-dimethylfuran	5.588
75-27-4	Bromodichloromethane	5.6272
18936-17-9	2-methylbutanenitrile	5.7888
108-87-2	Methylcyclohexane	5.9164
589-43-5	2,4-dimethylhexane	5.9782
137-32-6	2-methylbutan-1-ol	6.0631
123-51-3	Isoamylol	6.0637
108-10-1	Methyl isobutyl ketone	6.152
563-16-6	3,3-dimethylhexane	6.2117
624-92-0	Dimethyl disulfide	6.3419
565-69-5	3-pentanone, 2-methyl -	6.4284
565-75-3	2,3,4-trimethylpentane	6.4356
556-52-5	Glycidol	6.5762
560-21-4	2,3,3-trimethylpentane	6.6197
123-15-9	2-methyl valeraldehyde	6.7101
589-53-7	4-methylheptane	6.8579
108-88-3	Toluene	6.9301
544-25-2	Cyclohepta-1,3,5-triene	6.9428
589-81-1	3-methylheptane	7.044
71-41-0	Pentan-1-ol	7.2035
123-54-6	Acetylacetone	7.2443
1632-16-2	3-methyleneheptane	7.3184
591-78-6	2-oxohexane	7.5002
111-66-0	1-octene	7.5843
3848-24-6	Hexane-2,3-dione	7.63
565-80-0	2,4-dimethyl-3-pentanone	7.6375
120-92-3	Cyclopentanone	7.6439
141-79-7	Mesityl oxide	7.8172
111-65-9	Octane	7.8792
66-25-1	Hexanal	7.9063
589-90-2	1,4-dimethylcyclohexane	7.9708
124-48-1	Chlorodibromomethane	8.0102
127-18-4	Tetrachloroethene	8.1167
106-36-5	Propyl propionate	8.1745
1069-53-0	2,3,5-trimethylhexane	8.3623
123-86-4	Butyl acetate	8.3834
2213-23-2	2,4-dimethylheptane	8.6027
123-42-2	Diacetone alcohol	9.281
19549-87-2	2,4-dimethyl-1-heptene	9.3129
108-90-7	Chlorobenzene	9.581
7452-79-1	Ethyl 2-methylbutyrate	9.6666
100-41-4	Ethylbenzene	10.15
106-42-3	P-xylene	10.5036
123-92-2	Isoamyl acetate	10.7472
2918-13-0	1-hepten-3-one	10.7485
106-35-4	Heptan-3-one	11.1228
124-11-8	1-nonene	11.1885
110-43-0	Heptan-2-one	11.2844
100-42-5	Styrene	11.405
108-94-1	Cyclohexanone	11.543
4466-24-4	2-butylfuran	11.5933
111-84-2	Nonane	11.7717
111-71-7	Heptanal	11.8431
142-83-6	2,4-hexadienal	11.8594
111-76-2	2-butoxyethanol	11.9543
590-01-2	Butyl propionate	12.0473
815-17-8	3,3-dimethyl-2-oxobutanoi…	12.1797
67-71-0	Sulfonyldimethane	12.3492
15869-93-9	3,5-dimethyloctane	12.7351
542-18-7	Monochlorocyclohexane	12.7752
508-32-7	Tricyclene	12.8184
98-82-8	Cumene	12.8383
930-68-7	Cyclohex-2-enone	13.1709
4110-44-5	3,3-dimethyloctane	13.2882
80-56-8	Alpha-pinene	13.329
591-11-7	Beta-angelica lactone	13.7166
79-92-5	Camphene	14.0329
103-65-1	Propylbenzene	14.2312
123-05-7	2-ethylhexanal	14.242
620-14-4	3-ethyltoluene	14.5803
100-52-7	Benzaldehyde	14.6746
108-67-8	1,3,5-trimethylbenzene	14.9639
3387-41-5	Sabinene	15.1728
28634-89-1	2-thujene	15.1858
99-82-1	P-menthane	15.2135
464-15-3	Bornane	15.2882
611-14-3	1-ethyl-2-methylbenzene	15.3914
18172-67-3	Beta-pinene	15.4356

108-95-2	Phenol	15.5177
110-93-0	Sulcatone	15.6842
872-05-9	1-decene	15.7281
100-47-0	Benzonitrile	15.7838
123-35-3	Beta-myrcene	15.9591
111-13-7	2-octanone	15.9665
13475-82-6	2,2,4,6,6-pentamethylhep…	15.98
3777-69-3	2-pentylfuran	16.0288
4551-51-3	Hydrindane	16.0893
95-63-6	1,2,4-trimethylbenzene	16.1626
288-88-0	1h-1,2,4-triazole	16.3004
124-18-5	Decane	16.4837
124-13-0	Octanal	16.6386
538-93-2	Isobutylbenzene	16.8503
13466-78-9	3-carene	16.9046
99-83-2	Alpha-phellandrene	16.9243
13466-78-9	Car-3-ene	16.9305
111-90-0	Diethylene glycol monoeth…	17.0222
99-86-5	Alpha-terpinene	17.333
527-84-4	O-cymene	17.4599
535-77-3	M-cymene	17.4614
526-73-8	1,2,3-trimethylbenzene	17.494
99-87-6	P-cymene	17.7111
104-76-7	2-ethylhexan-1-ol	17.8409
5989-27-5	Limonene	17.929
555-10-2	Beta-phellandrene	18.0159
470-82-6	1,8-cineole	18.1048
496-11-7	Indane	18.1741
13877-91-3	Beta-ocimene	18.2286
872-50-4	N-methylpyrrolidin-2-one	18.2303
95-13-6	1h-indene	18.6427
122-78-1	Phenylacetaldehyde	18.6691
90-02-8	Salicylaldehyde	18.6901
105-05-5	1,4-diethylbenzene	18.8056
99-85-4	Gamma-terpinene	19.3498
493-02-7	Decahydronaphthalene	19.4416
98-86-2	Acetophenone	19.7478
111-87-5	Octan-1-ol	20.0322
18479-58-8	Dihydromyrcenol	20.043
1195-32-0	P-isopropenylmethylbenzene	20.4872
617-94-7	2-phenylpropan-2-ol	20.7469
586-63-0	Isoterpinolene	20.7523
1195-79-5	Fenchone	20.9363
821-55-6	Nonan-2-one	20.9501
78-69-3	Linalool	21.4118
1120-21-4	Undecane	21.4394
62108-26-3	2,6,8-trimethyldecane	21.487
493-01-6	Decalin	21.6431
124-19-6	Nonanal	21.6711
546-80-5	Alpha-thujone	21.798
3878-55-5	Monomethyl succinate	22.0457
95-93-2	Durene	22.3823
78-59-1	Isophorone	22.4917
4501-58-0	Alpha-campholenaldehyde	22.7966

76-22-2	Camphor	23.815
488-23-3	Prehnitene	23.9092
1196-31-2	Menthone	24.229
598-42-5	2-hydroxyacetamide	24.3384
119-64-2	Tetralin	24.4536
140-39-6	Tolylacetate	24.5413
30460-92-5	Pinocarvone	24.5466
65-85-0	Benzoic acid	24.7442
143-08-8	Nonan-1-ol	24.9554
464-45-9	Borneol	25.0638
2216-51-5	Menthol	25.2807
150-78-7	1,4-dimethoxybenzene	25.3022
562-74-3	4-terpineol	25.4396
91-20-3	Naphthalene	25.6484
693-54-9	Decan-2-one	25.8984
119-36-8	Methyl salicylate	25.9934
112-40-3	Dodecane	26.3335
112-31-2	Decanal	26.612
1196-01-6	Verbenone	26.7204
7206-26-0	2-dodecene	26.8779
17301-22-3	2,5-dimethylundecane	26.9308
95-16-9	Benzothiazole	27.6234
99-49-0	Carvone	28.4312
115-95-7	Linalyl acetate	28.6212
105-60-2	Epsilon-caprolactam	28.8731
112-05-0	Nonanoic acid	29.3346
31295-56-4	2,6,11-trimethyldodecane	29.7224
112-30-1	Decan-1-ol	29.7332
5655-61-8	Bornyl acetate	30.3449
125-12-2	Isobornyl acetate	30.4542
98-54-4	4-tert-butylphenol	30.5982
112-12-9	Undecan-2-one	30.6369
90-12-0	1-methylnaphthalene	30.8758
91-57-6	2-methylnaphthalene	30.888
89-83-8	Thymol	30.93
499-75-2	Carvacrol	30.97
629-50-5	Tridecane	30.998
112-44-7	Undecanal	31.3361
10235-63-9	Gamma-terpinyl acetate	33.131
334-48-5	Decanoic acid	33.7862
26730-12-1	4-methyltridecane	33.8909
3891-98-3	Farnesane	34.3715
3856-25-5	Copaene	34.4868
92-52-4	Biphenyl	34.6362
1120-36-1	1-tetradecene	35.1113
80-46-6	P-t-amylphenol	35.1525
629-59-4	Tetradecane	35.4711
581-42-0	2,6-dimethylnaphthalene	35.7427
112-54-9	Dodecanal	35.8519
475-20-7	Longifolene	36.0339
11028-42-5	Cedrene	36.3204
575-43-9	1,6-dimethylnaphthalene	36.3352
575-41-7	1,3-dimethylnaphthalene	36.3538
575-37-1	1,7-dimethylnaphthalene	36.537

101-81-5	Diphenylmethane	36.8255
131-11-3	Dimethyl phthalate	37.446
112-53-8	Dodecan-1-ol	38.6094
473-13-2	Alpha-selinene	39.334
17066-67-0	Beta-selinene	39.3537
13360-61-7	1-pentadecene	39.361
4630-07-3	Valencene	39.6733
629-62-9	Pentadecane	39.6802
96-76-4	2,4-di-tert-butylphenol	39.8689
5875-45-6	2,5-di-tert-butylphenol	39.9219
10486-19-8	Tridecanal	40.1288
39029-41-9	Gamma-cadinene	40.3191
132-64-9	Dibenzofuran	40.4459
483-76-1	Delta-cadinene	40.46
21391-99-1	Alpha-calacorene	41.4053
4621-36-7	3-phenyldecane	42.1512
134-62-3	N,n-diethyl-m-toluamide	42.592
2050-08-0	Amyl salicylate	42.6522
84-66-2	Diethyl phthalate	43.1152

544-76-3	Hexadecane	43.6921
629-73-2	1-hexadecene	43.7255
10233-13-3	Isopropyl laurate	44.6351
4537-14-8	6-phenylundecane	44.718
4537-15-9	5-phenylundecane	44.8585
473-16-5	Alpha-eudesmol	45.4636
6765-39-5	1-heptadecene	47.181
2156-97-0	Lauryl acrylate	47.2019
629-78-7	Heptadecane	47.4797
24157-81-1	2,6-diisopropylnaphthalene	48.1924
2719-62-2	6-phenyldodecane	48.405
593-45-3	Octadecane	51.1069
118-60-5	Octyl salicylate	51.1733
638-36-8	Phytane	51.2331
110-27-0	Isopropyl myristate	51.8993
504-96-1	Neophytadiene	52.3804
1222-05-5	Musk gx	52.5787
629-92-5	Nonadecane	54.6432

Table S 2: ANOVA results for differences in the summed relative amount of each compound class identified between sampling sites on 27th November, 2024.

term	df	sumsq	meansq	statistic	p.value	Type_7
Location Name	**5**	**0.243**	**0.049**	**1.367**	**0.288**	**Aliphatics**
Location Name	**5**	**0.290**	**0.058**	**10.486**	**<0.001**	**Aromatics**
Location Name	**5**	**0.159**	**0.032**	**3.343**	**0.030**	**Monoterpenes**
Location Name	**5**	**0.003**	**0.001**	**23.940**	**<0.001**	**Others (Esters, N, S, Halo)**
Location Name	**4**	**0.000**	**0.000**	**3.860**	**0.043**	**Sesquiterpenes**
Location Name	**5**	**0.029**	**0.006**	**1.594**	**0.218**	**Terpenoids**

Table S 3: Summed relative amount of compounds identified within each compound class between sampling sites on 27th November, 2024. Letters indicate groups differences based on Tukey post-hoc tests (p<0.05).

	Aliphatics		Aromatics		Monoterpenes		Others		Sesquiterpenes		Terpenoids	
Location	**Sum**	**T**	**Sum**	**T**	**Sum**	**T**	**Sum**	**T**	**Sum**	**T**	**Sum**	**T**
OxBG Outside	2.41	a	10.67	a	1.08	ab	0.27	a	0.03	a	0.14	a
OxBG Rainforest House	1.79	a	1.21	b	3.26	a	0.17	a	0.03	ab	0.52	a
University Parks Dense	1.13	a	2.39	bc	0.31	ab	0.13	a	0.01	ab	0.09	a
University Parks Open	1.28	a	3.73	c	0.49	ab	0.26	b	0.00	b	0.09	a
Warneford	1.99	a	1.41	b	0.25	b	0.12	a	0.01	b	0.14	a
Wytham	6.60	a	3.15	bc	0.51	ab	0.03	a			0.53	a

Table S 4: ANOVA results for differences in the number of compounds identified within each compound class between sampling sites on 27th November, 2024.

term	df	sumsq	meansq	statistic	p.value	Type_7
Location Name	**5**	**320.364**	**64.073**	**2.044**	**0.127**	**Aliphatics**
Location Name	**5**	**331.833**	**66.367**	**4.496**	**0.009**	**Aromatics**

Location Name	**5**	**105.538**	**21.108**	**7.436**	**0.001**	**Monoterpenes**
Location Name	**5**	**69.167**	**13.833**	**3.125**	**0.037**	**Others (Esters, N, S, Halo)**
Location Name	**4**	**24.357**	**6.089**	**4.567**	**0.027**	**Sesquiterpenes**
Location Name	**5**	**177.333**	**35.467**	**21.280**	**<0.001**	**Terpenoids**

Table S 5: Summed number of compounds identified within each compound class between sampling sites on 27th November, 2024. Letters indicate groups differences based on Tukey post-hoc tests (p<0.05).

	Aliphatics		Aromatics		Monoterpenes		Others		Sesquiterpenes		Terpenoids	
Location	**n**	**Tuk**	**n**	**Tuk**	**n**	**Tuk**	**n**	**Tuk**	**n**	**Tuk**	**n**	**Tuk**
OxBG Outside	36	a	31	a	10	ab	12	a	2	a	7	a
OxBG Rainforest House	42	a	25	ab	13	a	11	ab	3	a	13	b
University Parks Dense	24	a	16	b	6	b	6	ab	1	a	5	ac
University Parks Open	27	a	18	b	7	b	5	ab	1	a	3	c
Warneford	35	a	21	ab	7	b	7	ab	1	a	5	ac
Wytham	26	a	20	b	4	b	2	b	0		5	ac

Table S 6: All compounds identified only at a single site during the 27 November 2024 sampling.

Name	Loc	Type_7
2,4-hexadienal	**BO**	**Aliphatics**
2,5-dimethylundecane	**BO**	**Aliphatics**
2-methyl valeraldehyde	**BO**	**Aliphatics**
3,5-dimethyloctane	**BO**	**Aliphatics**
4-methyltridecane	**BO**	**Aliphatics**
Acetylacetone	**BO**	**Aliphatics**
Decalin	**BO**	**Aliphatics**
Glycidol	**BO**	**Aliphatics**
Mesityl oxide	**BO**	**Aliphatics**
Pyrrolidine	**BO**	**Aliphatics**
1,4-diethylbenzene	**BO**	**Aromatics**
1h-indene	**BO**	**Aromatics**
2-butylfuran	**BO**	**Aromatics**
Cumene	**BO**	**Aromatics**
Isobutylbenzene	**BO**	**Aromatics**
P-isopropenylmethylbenzene	**BO**	**Aromatics**
Prehnitene	**BO**	**Aromatics**
Styrene	**BO**	**Aromatics**
Tetralin	**BO**	**Aromatics**
2-thujene	**BO**	**Monoterpenes**
Beta-angelica lactone	**BO**	**Monoterpenes**
Bornane	**BO**	**Monoterpenes**
Isobornyl acetate	**BO**	**Monoterpenes**
Verbenone	**BO**	**Monoterpenes**
Ethyl 2-methylbutyrate	**BO**	**Others**
Tolylacetate	**BO**	**Others**
Farnesane	**BO**	**Sesquiterpenes**
1-hepten-3-one	**RH**	**Aliphatics**
2,4-dimethyl-3-pentanone	**RH**	**Aliphatics**
2-hydroxyacetamide	**RH**	**Aliphatics**
3-methylbutanal	**RH**	**Aliphatics**
3-pentanone, 2-methyl -	**RH**	**Aliphatics**
Epsilon-caprolactam	**RH**	**Aliphatics**
6-phenyldodecane	**RH**	**Aromatics**
Musk gx	**RH**	**Aromatics**
Beta-myrcene	**RH**	**Monoterpenes**
Beta-ocimene	**RH**	**Monoterpenes**
Beta-phellandrene	**RH**	**Monoterpenes**
Gamma-terpinyl acetate	**RH**	**Monoterpenes**
Isoamyl acetate	**RH**	**Others**
N,n-diethyl-m-toluamide	**RH**	**Others**
Sulfide, allyl methyl	**RH**	**Others**
Alpha-eudesmol	**RH**	**Sesquiterpenes**
Alpha-selinene	**RH**	**Sesquiterpenes**
Beta-selinene	**RH**	**Sesquiterpenes**
Cedrene	**RH**	**Sesquiterpenes**
Copaene	**RH**	**Sesquiterpenes**
Gamma-cadinene	**RH**	**Sesquiterpenes**
Longifolene	**RH**	**Sesquiterpenes**
Valencene	**RH**	**Sesquiterpenes**
4-terpineol	**RH**	**Terpenoids**
Isophorone	**RH**	**Terpenoids**
Linalyl acetate	**RH**	**Terpenoids**
Sulcatone	**RH**	**Terpenoids**

2-octanone	**UD**	**Aliphatics**
Decan-2-one	**UD**	**Aliphatics**
Nonadecane	**UD**	**Aliphatics**
Nonan-2-one	**UD**	**Aliphatics**
2,6-dimethylnaphthalene	**UD**	**Aromatics**
Diphenylmethane	**UD**	**Aromatics**
Cyclohex-2-enone	**UD**	**Others**
Pentan-1-ol	**UO**	**Aliphatics**
1,4-dimethoxybenzene	**UO**	**Aromatics**
3,3-dimethyl-2-oxobutanoic..	**UO**	**Others**
Alpha-campholenaldehyde	**UO**	**Terpenoids**
1-hexadecene	**WA**	**Aliphatics**
Butan-1-ol	**WA**	**Aliphatics**
Cyclohexene	**WA**	**Aliphatics**
Nonan-1-ol	**WA**	**Aliphatics**
2,5-di-tert-butylphenol	**WA**	**Aromatics**
2-pentylfuran	**WA**	**Aromatics**
2-phenylpropan-2-ol	**WA**	**Aromatics**
Benzothiazole	**WA**	**Aromatics**
Phenol	**WA**	**Aromatics**
Car-3-ene	**WA**	**Monoterpenes**
Lauryl acrylate	**WA**	**Others**
Neophytadiene	**WA**	**Terpenoids**
1-pentadecene	**WY**	**Aliphatics**
2,3-dimethylpentane	**WY**	**Aliphatics**
2,6,8-trimethyldecane	**WY**	**Aliphatics**
Cyclopentanone	**WY**	**Aliphatics**
Pentanal	**WY**	**Aliphatics**
Alpha-phellandrene	**WY**	**Monoterpenes**
Isoterpinolene	**WY**	**Monoterpenes**
Thymol	**WY**	**Monoterpenes**
Butyl acetate	**WY**	**Others**
Butyl propionate	**WY**	**Others**
Monomethyl succinate	**WY**	**Others**
Borneol	**WY**	**Terpenoids**
Carvacrol	**WY**	**Terpenoids**
Fenchone	**WY**	**Terpenoids**

692

Table S 7: Top-ranking compounds from PLS-DA analyses with explicit directionality. VIP scores (> ranked top 10) identify compounds contributing most strongly to group separation. Group-wise mean abundances are provided to support directional interpretation.

Comparison	Direction	Compound	VIP	Loading	Mean_Group1	Mean_Group2
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	P-xylene (Aromatics)	2.296	-0.148	0.186	0.367
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	Ethylbenzene (Aromatics)	2.268	-0.146	0.108	0.209
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	Isooctane (Aliphatics)	2.206	-0.142	0.035	0.133
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	2,3,4-trimethylpentane (Aliphatics)	2.150	-0.138	0.032	0.087
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	Toluene (Aromatics)	2.138	-0.137	0.239	0.424
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	3-methylhexane (Aliphatics)	2.121	-0.136	0.064	0.136
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	Methylcyclohexane (Aliphatics)	2.065	-0.133	0.065	0.131
WAH/WYH/UDH/RHH v BOH/UOH	> in WAH/WYH/UDH/RHH	2-dodecene (Aliphatics)	2.060	0.132	0.198	0.000
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	1-ethyl-2-methylbenzene (Aromatics)	2.014	-0.129	0.027	0.110
WAH/WYH/UDH/RHH v BOH/UOH	> in BOH/UOH	3-methylheptane (Aliphatics)	2.005	-0.129	0.041	0.074
WAH/WYH/UDH v RHH	> in RHH	Beta-ocimene (Monoterpenes)	2.285	-0.147	0.000	0.072
WAH/WYH/UDH v RHH	> in RHH	Dihydromyrcenol (Terpenoids)	2.281	-0.147	0.000	0.105
WAH/WYH/UDH v RHH	> in RHH	Sabinene (Monoterpenes)	2.273	-0.146	0.005	0.374
WAH/WYH/UDH v RHH	> in RHH	Beta-pinene (Monoterpenes)	2.244	-0.144	0.033	0.287
WAH/WYH/UDH v RHH	> in RHH	Pinocarvone (Terpenoids)	2.232	-0.143	0.000	0.043
WAH/WYH/UDH v RHH	> in RHH	Hydrindane (Aliphatics)	2.216	-0.142	0.000	0.028
WAH/WYH/UDH v RHH	> in RHH	Alpha-pinene (Monoterpenes)	2.196	-0.141	0.104	0.376
WAH/WYH/UDH v RHH	> in RHH	Copaene (Sesquiterpenes)	2.195	-0.141	0.000	0.048
WAH/WYH/UDH v RHH	> in RHH	4-terpineol (Terpenoids)	2.195	-0.141	0.000	0.034
WAH/WYH/UDH v RHH	> in RHH	Beta-phellandrene (Monoterpenes)	2.169	-0.139	0.000	0.045
BOH v UOH/UDH	> in BOH	2-thujene (Monoterpenes)	2.300	-0.148	0.055	0.000
BOH v UOH/UDH	> in BOH	Prehnitene (Aromatics)	2.288	-0.147	0.035	0.000
BOH v UOH/UDH	> in BOH	Farnesane (Sesquiterpenes)	2.284	-0.147	0.035	0.000
BOH v UOH/UDH	> in BOH	P-isopropenylmethylbenzene (Aromatics)	2.274	-0.146	0.020	0.000
BOH v UOH/UDH	> in BOH	Bromodichloromethane (Others (Esters, N…)	2.268	-0.146	0.013	0.000
BOH v UOH/UDH	> in BOH	1,4-diethylbenzene (Aromatics)	2.259	-0.145	0.045	0.000
BOH v UOH/UDH	> in UOH/UDH	Tricyclene (Monoterpenes)	2.251	0.145	0.000	0.076
BOH v UOH/UDH	> in BOH	2,2,4,6,6-pentamethylheptane (Aliphatics)	2.201	-0.142	0.099	0.062
BOH v UOH/UDH	> in BOH	Bornyl acetate (Terpenoids)	2.183	-0.140	0.017	0.000
BOH v UOH/UDH	> in BOH	Propylbenzene (Aromatics)	2.146	-0.138	0.093	0.009