REVISITING THE 2013 SOUTHERN OXIDANT AEROSOL STUDY WITH EXPLORATORY DATA ANALYSIS AND MACHINE LEARNING

By

Ricardo Pfleider Tolentino

A capstone project submitted for Graduation with University Honors

May 05, 2021

University Honors
University of California, Riverside

APPROVED

Dr. Haofei Zhang
Department of Chemistry

Dr. Richard Cardullo, Howard H Hays Jr. Chair
University Honors

Abstract

Atmospheric organic aerosols are highly misunderstood and have broad implications on our climate, air quality, and the health of humanity. Understanding the origins of organic aerosols and the mechanisms by which they are transformed is difficult due to the complexity of the multiphase chemistry within the species. Achieving such understanding will greatly advance the scientific community in tackling our global climate crisis. Recent advances in technology have allowed researchers to be more precise in categorization and characterization of organic aerosols and their sources. This project will be revisiting the data set of Zhang et al., 2018 using exploratory data analysis to examine: (1) if there exist more categories than have already been determined under monoterpene SOA; (2) under the previously determined categories, whether sub-groups of species could be created which inform unique formation pathways rather than just the sources; (3) claims of anthropogenic influence on biogenic secondary organic aerosol formation; and (4) mechanistic understandings of absence of new particle formation (highly relevant to cloud formation).

Acknowledgements

Thank you, Professor Haofei Zhang, for allowing me the opportunity to learn about atmospheric chemistry under your guidance.

Thank you, Professor Richard Cardullo, for allowing me the opportunity to pursue a capstone project in UC Riverside's University Honors program.

Thank you, Gladis Hererra-Berkowitz, for taking me on as an Undergraduate Research Ambassador and allowing me the opportunity to work with amazing students.

Thank you PhD student Tiffany Zhao for the guidance, support, and general mentorship from the moment I joined the lab.

Thank you to my friends and family for the emotional support throughout the entire process.

Personal Capstone Reflection

While we were able to reveal novel information through EDA of our 830 x 254 dataset, the impetus of this project was to perform such analysis using machine learning. At the time of beginning the project, I had little to no coding experience, and even less with machine learning! I had simply been curious about the emerging field, and I felt that pursuing this topic for my Capstone project would allow me the opportunity to learn about these unfamiliar topics in the scope of my research interests.

My first attempts at machine learning wished to utilize the K-Means Clustering algorithm to hopefully create new groupings that could inform subgroups of OAs or different formation pathways. While I was able to implement this algorithm on a variety of practice datasets, I found that this algorithm was not appropriate for my dataset for a variety of reasons. This method relies first on dimensionality reduction of my dataset from being 830 samples x 254 time points to being 830 samples x 2 principal components. For this method to work properly, the data reduced principal components should be able to describe at least 85% of the variance in the dataset, whereas my analysis could only attain ~60%. As a result, the PCA plot generated had datapoints extremely close to each other, whereas the K-Means Algorithm was unable to create meaningful separations.

Aside from using machine learning, I also attempted to use neural networks on my dataset, without any cemented idea on what exactly I would be creating. I eventually was able to create a model that could semi-successfully predict which category an OA belonged to based on its timeseries, but such a model did not provide any insight into new subgroups nor formation pathways. This is commonly referred to as the black box problem in the field of data science.

While I could have optimized this model to give better predictions, I found this pursuit irrelevant to this project.

While I did in fact learn many Python methods in exploratory data analysis and machine learning, I feel I learned more about myself as an independent researcher. This was a humbling experience that came with many moments of self-doubt. While I have only dipped my toes into the field of exploratory data analysis, machine learning, and neural networks, I am very grateful to have had this experience in trying to solve problems that are relevant to my research group using these novel methods. I undoubtedly will continue this pursuit, and I am confident that one day I will develop methods to remedy the drawbacks of this study.

Table of Contents

Background

Aerosol relevance

Organic aerosols (OAs) and secondary organic aerosols (SOAs) contribute greatly to atmospheric fine particulate matter and have broad implications on our climate, air quality, and the health of humanity. Understanding the physical and chemical transformation of such aerosols is significant regarding research on cloud formation, the earth's energy balance, and a variety of aerosol related illnesses.

Cloud Formation

Cloud formation occurs when atmospheric water vapor condenses into suspensions of gaseous and particulate water, referred to as aerosols. This process is referred to as homogeneous water nucleation. While homogeneous water nucleation typically occurs under conditions of gaseous water supersaturation, OAs and SOAs can act as cloud condensation nuclei (CCN) and thus form clouds at supersaturations lower than those typically required for homogeneous water nucleation.[1] Size distributions of cloud droplets can be predicted from the chemical composition and size distribution of the pre-cloud aerosols. Understanding CCN activity of OAs and SOAs is important in the quantification of their effects on the formation, lifetime, and radiative properties of clouds. Therefore, aerosol composition and size distribution have significant implications on the lifetime and radiative effects of clouds. While the cloud forming potential of the inorganic constituents are generally well known, the same cannot be said for the organic portions that contribute to 20-50% of the mass of fine aerosol particles.[2] Additionally, SOA from photo oxidation of volatile organic compounds (VOCs) has been shown to contribute greatly to the organic portions.[3]

Earth's Energy Balance

The radiative properties of clouds formed from OAs and SOAs have grand implications on the energy balance of the earth. Such effects depend on the light absorption and scattering properties of certain aerosols.

For example, light absorption from dust and black carbon results in a warming effect on climate, while light scattering from reflective species like nitrate, sulfate, and sea salt aerosols result in a cooling effect.[4] A study in 2013 found that SOAs from a-pinene formed under typical atmospheric conditions generally do not absorb light, but also reports significant light absorption at 355 and 405 nm when the SOAs are formed from a-pinene + O3 + NO3 in the presence of highly acidic sulfate seed aerosols under dry conditions.[4] Further, they state that no absorption was found when the relative humidity was higher than 27%. Multiphase reactions of SOAs can also affect their refractive index (RI), affecting their light scattering and absorbing properties. Such reactions in the presence of liquid water result in oligomer formation from glyoxal and methylglyoxal intermediates which further enhances the RI and thus the light-scattering properties of the SOA.

Aerosol Illness

SOAs can potentially be more dangerous than their OA precursors and can lead to a variety of illnesses including cardiovascular disease, respiratory system damage, lung cancer, and preterm birth. A 2006 study showed that a-pinene SOA increased the release of the proinflammatory mediator interleukin-8 (IL8) by respiratory epithelial cells.[5] A 2019 study continued upon this and found that SOA derived from a-pinene, m-xylene, and trimethylbenzene increased the expression of both the heme oxygenase I (HMOXI) and IL8 genes, typical markers for oxidative stress and inflammatory responses, respectively.[6] Another recent 2019 study

examined isoprene derived SOA and found that such SOA also induced the expression of oxidative stress and inflammation genes in human lung cells.[7] These studies suggest that SOAs affect lung epithelial cells and macrophages through oxidative stress and inflammation and can contribute to a variety of illnesses.

SOAS 2013 Background

A field campaign in 2013, dubbed the Southern Oxidant and Aerosol Study (SOAS 2013), planned to quantify VOC, ozone, and NOx surface fluxes and reconcile differences with "blank-down" emission estimates to achieve a better understanding of Hox/NOx/ozone/organics/aerosol distributions, sources, and sinks.[8]

One paper by Professor Haofei Zhang from UC Riverside comprehensively characterized OA composition with molecular-level details by doing complementary analyses of OAs collected in Centreville, Alabama during SOAS 2013.[9] This study reports that SOA from monoterpene oxidation accounts for approximately half of summertime fine OA in Centreville, Al, a forested area in the southeastern United States which is influenced by anthropogenic pollution, and further, that different chemical processes occurring during days and nights, determine the mass of monoterpene SOA produced. The findings of this study shed light on the strong anthropogenic-biogenic interaction affecting ambient aerosol in the southeastern United States and highlight the importance in reducing anthropogenic emissions.

Recent analyses of the Southern Oxidant and Aerosol Study (SOAS) have demonstrated that anthropogenic (man-made) SO2 and NOx are responsible for 43-70% of total organic aerosol in the southeastern US during the summer[10], and that monoterpene secondary organic aerosol accounts for about half of total fine organic aerosol[9]. Additionally, the studies indicate

that reducing anthropogenic SO2 and NOx emissions will reduce secondary organic aerosol in the southeastern United States.

This goal of this research is to revisit the SOAS 2013 dataset to reveal (1) if there exist more categories than have already been determined under monoterpene SOA; (2) under the previously determined categories, whether sub-groups of species could be created which inform unique formation pathways rather than just the sources; (3) claims of anthropogenic influence on biogenic secondary organic aerosol formation; and (4) mechanistic understandings of absence of new particle formation (highly relevant to cloud formation).

Methods, Results and Discussion

Original Sample Collection and Analysis

During SOAS 2013, 254 submicron aerosol samples were collected on quartz filters (prebaked at 600 C for 12 hrs to remove organics) using a sequential sampler at ~120 L/min every 4 hours from June 3 to July 15.[9] Filters were stored in baked foil at -20 C before performing in-lab experiments.

Categorization of aerosol samples was done using linear regression analysis; field samples from the SOAS campaign were correlated to known lab samples of monoterpene secondary organic aerosol. Under this method, 89 OA samples were identified as originating from monoterpene oxidation. Linear regression analysis of the remaining aerosols with those of the 89 samples was performed and any aerosols with a correlation coefficient (r^2) greater than .60 were assigned to be monoterpene SOA (MTSOA).

Theory Behind New Groupings

Despite the utilization of state-of-the-art analytical instrumentation during SOAS 2013, the study's analytical method of linear regression, while standard in the scientific community, is

considered quite simple relative to emerging analytical methods. I hypothesize that there exists more information within the SOAS 2013 dataset that could be revealed through exploratory data analysis (EDA) and machine learning (ML) methods.

Python Data Analysis

Analysis of the SOAS 2013 dataset was powered by Python in Anaconda's Jupyter Notebook utilizing SciPy which includes the Pandas, NumPy, and matplotlib libraries.[11] While the original goal of this project was to perform machine learning analysis, such analysis was found to be inviable for a variety of reasons. This will be expanded upon in a later segment. This study instead focused on a simpler exploratory data analysis approach that still reveals novel information from the SOAS 2013 Dataset.

SOAS 2013 Dataset

The dataset I was provided with for this analysis is an 830 x 254 matrix with the 830 rows representing individually identified samples and the 254 columns representing each 4-hour time step from June 3 to July 15, 2013. Additionally, Professor Haofei Zhang provided me with a previously unused tracer dataset from SOAS 2013 which is a 52 x 254 matrix with the 52 rows representing various tracer species such as isoprene, a-pinene, nitrates, carbon monoxide, and ozone, and the 254 columns representing the time steps. From this supplemental dataset, I will be utilizing the sample time (the time the sample was taken) to perform a diurnal dataset split to examine differences in aerosol correlation during daytime and nighttime hours. Additionally, I will be utilizing the wind direction (direction of wind during time of sample collection) to perform a wind-direction dataset split to examine differences in aerosol correlation during varying wind directions.

Exploratory Data Analysis

To explore different conditions within the full SOAS 2013 dataset, various subsets representing different environmental conditions were created. Such datasets and their respective dimensions are as follows: full dataset (830 x 254), daytime (830 x 127), night time (830 x 127), northeast winds (830 x 36), southeast winds (830 x 88), southwest winds (830 x 99), northwest winds (830 x 28), east winds (830 x 127), west winds (830 x 127), north winds (830 x 67), and south winds (830 x 187).

An important thing to notice is that the number of aerosols analyzed does not change, but rather the number of time steps. **Essentially, we are selecting which quartz filters to perform analysis on depending on environmental conditions of interest.** Additionally, individual OA samples were correlated with each tracer within each dataset to examine how such environmental conditions affect OA correlation individually.

Results

To examine correlations between OAs and tracer species, this study used Pearson correlation (r^2) from the SciPy package as the primary metric. As in Zhang et. al., 2018, An r^2 of +.6 or greater indicated correlation between the two species. The r^2 values are calculated between two time series: a (1 x 254) sample series, and a (1 x 254) tracer series from the (830 x 254) and (52 x 254) datasets respectively. If at least 50 OAs reach this .6 correlation threshold within a given tracer group, such OAs are considered to be a part of such tracer group. Further, such group is considered to be significant in regard to elucidating new OA groupings. In the following distributions, the x-axis is populated with each tracer species, while the y-axis represents frequency. Such distributions do not show individual OA activity, but rather OA to tracer correlations on a macro scale. Individual OA tracer correlations within each dataset for the first

110 species were also examined. The changes in OA tracer correlation are quite chaotic on a micro scale and could be the topic of future studies. The table following the distributions will demonstrate this novel activity.

Tracer Name (unit)	Description
AMS-BBOA factor (ug/m3)	AMS PMF factor of biomass burning OA (BBOA)
AMS-IEPOX factor (ug/m3)	AMS PMF factor of IEPOX-derived OA (a major pathway in isoprene oxidation)
AMS-LOOOA factor (ug/m3)	AMS PMF factor of less oxidized (LO) oxygenated OA (OOA)
AMS-MOOOA factor (ug/m3)	AMS PMF factor of more-oxidized (MO) oxygenated OA (OOA)
AMS-NH4 (ug/m3)	total aerosol ammonium (NH4) mass measured by AMS
AMS-SO4 (ug/m3)	total aerosol sulfate (SO4) mass measured by AMS
BrC (ug/m3)	brown carbon
HNO3 (ppb)	nitric acid in the gas phase
AMS-IEPOX factor (ug/m3)	AMS PMF factor of IEPOX-derived OA (a major pathway in isoprene oxidation)
AMS-ISOPOOH factor (ug/m3)	AMS PMF factor of ISOPOOH-derived OA (a minor pathway in isoprene oxidation)
LWC (ug/m3)	liquid water content in the particle phase
MTR	monoterpene reactivity
NO (ppb)	nitric oxide
NO2 (ppb)	nitrogen dioxide
NOx (ppb)	generic nitrogen oxides
NOy (ppb)	NOy = NO + NO2 + NOz
NOz (ppb)	NOz = HNO3 + HONO + 2N2O5 + HO2NO2 + PAN + NO3 + Organic Nitrates
O3 (ppb)	ozone
SO2 (ppb)	sulfur dioxide
WSOC (ug/m3)	total water-soluble organic compounds in the particle phase
a-pinene (ppb)	alpha-pinene
b-pinene (ppb)	beta-pinene
benzene (ppb)	benzene
black carbon (ug/m3)	black carbon
H_2SO_4 (ppb)	Sulfuric acid
orgNO3 (ppb)	total organic nitrate in the gas phase
total monoterpene (ppb)	total monoterpene, including alpha-pinene, beta-pinene, limonene, and several other minor monoterpenes

Table 1. Tracer names and descriptions. Only the most relevant tracers are shown.

Dataset split distributions:

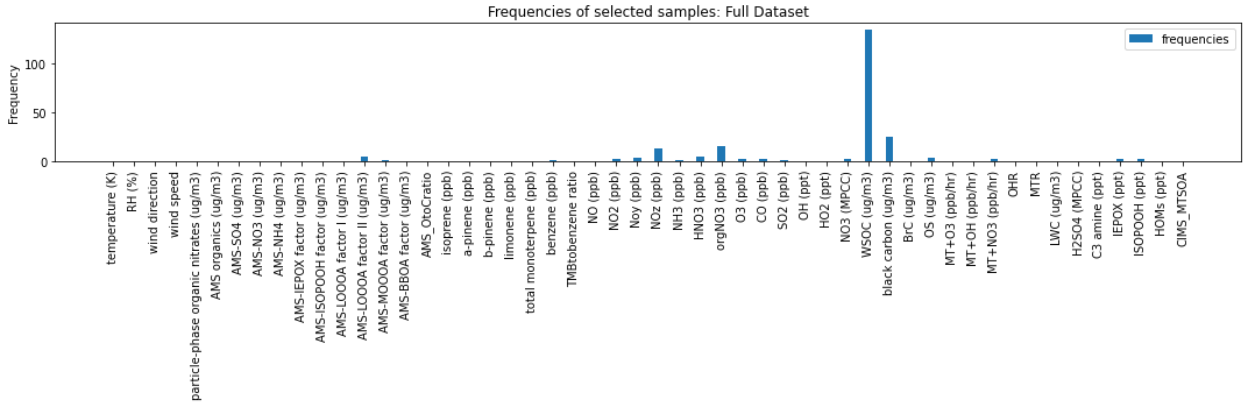

Fig 1. Distribution of selected aerosols within the entire dataset (830 x 254).

Relevant tracer: WSOC

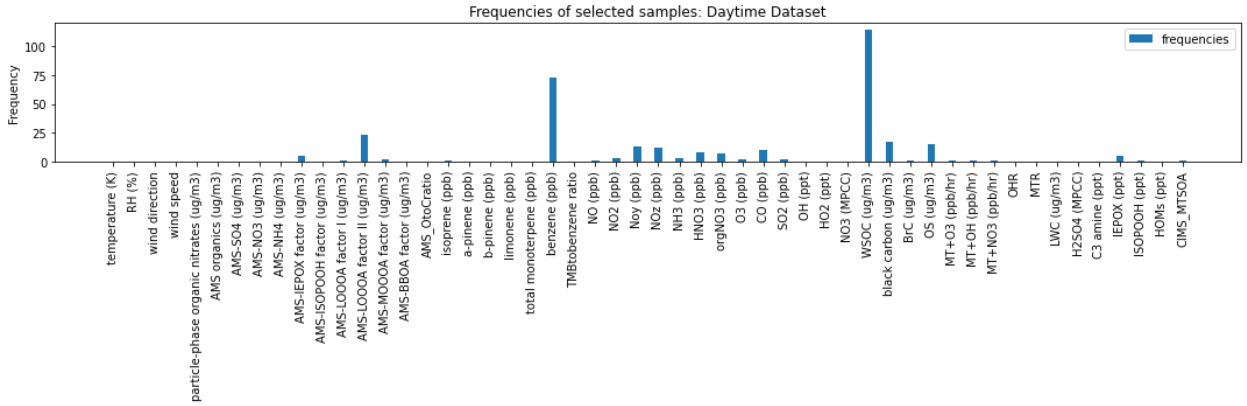

Fig 2. Distribution of selected aerosols within the daytime dataset (830 x 127).

Relevant tracers: benzene, WSOC

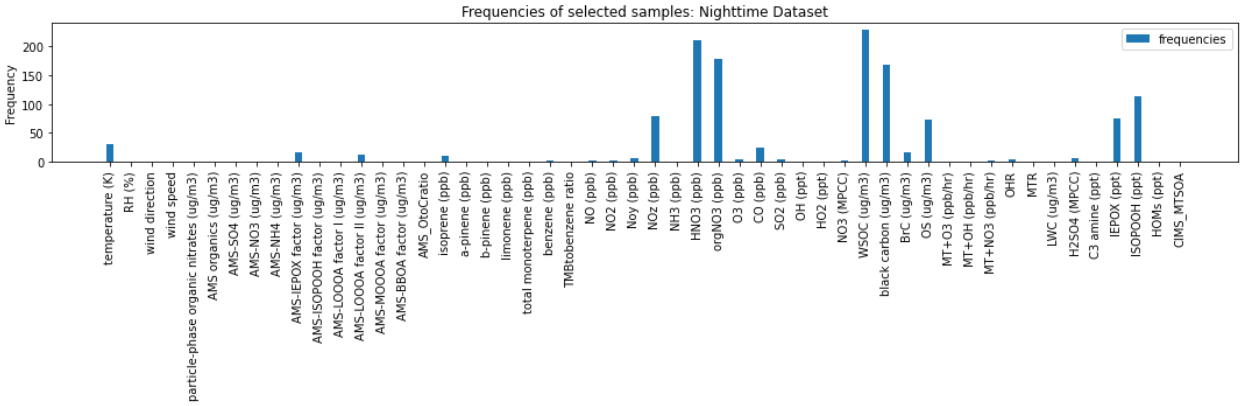

Fig 3. Distribution of selected aerosols within the nighttime dataset (830 x 127).

Relevant tracers: NO_z, HNO_3, $orgNO_3$, WSOC, black carbon, organosulfates (OS), IEPOX, ISOPOOH

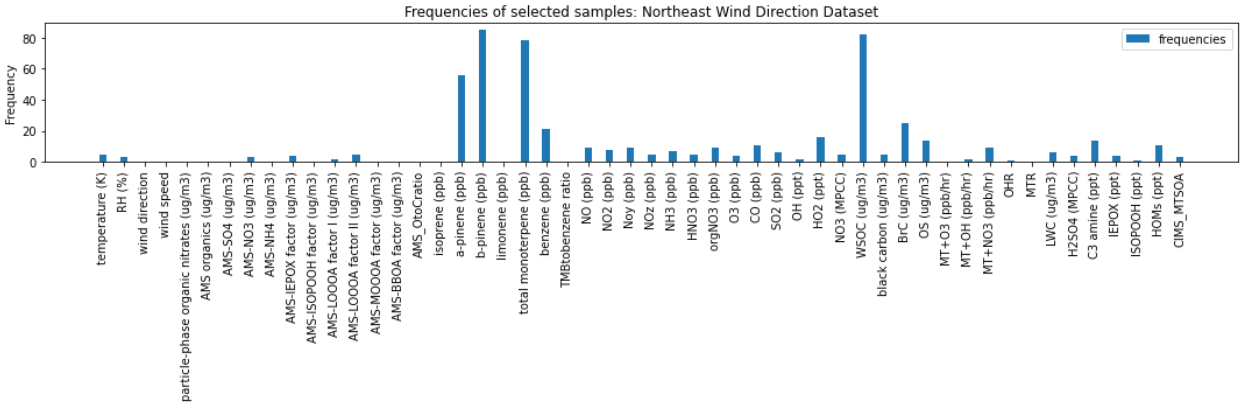

Fig 4. Distribution of selected aerosols within the northeast wind direction dataset (830 x 36).

Relevant tracers: a-pinene, b-pinene, total monoterpene, WSOC

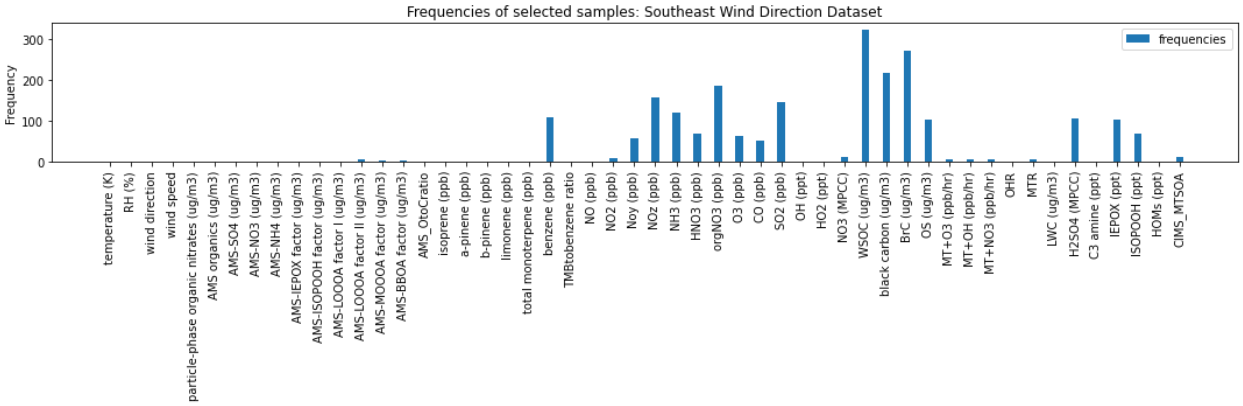

Fig 5. Distribution of selected aerosols within the southeast wind direction dataset (830 x 88).

Relevant tracers: benzene, NO_y, NO_z, NH_3, HNO_3, $orgNO_3$, O_3, CO, SO_2, WSOC, black carbon, BrC, OS, H_2SO_4, IEPOX, ISOPOOH

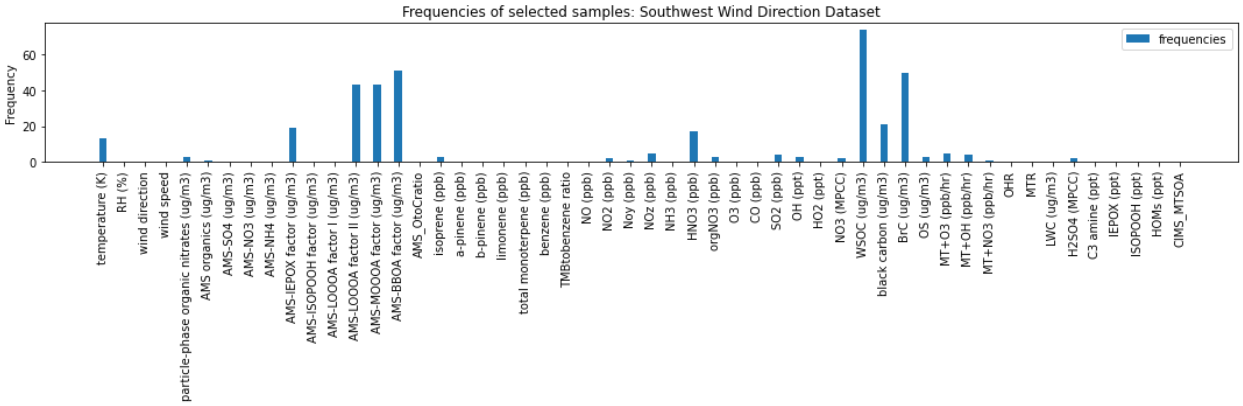

Fig 6. Distribution of selected aerosols within the southwest wind direction dataset (830 x 99).

Relevant tracers: AMS-LOOOA, AMS-MOOOA, AMS-BBOA, WSOC, BrC

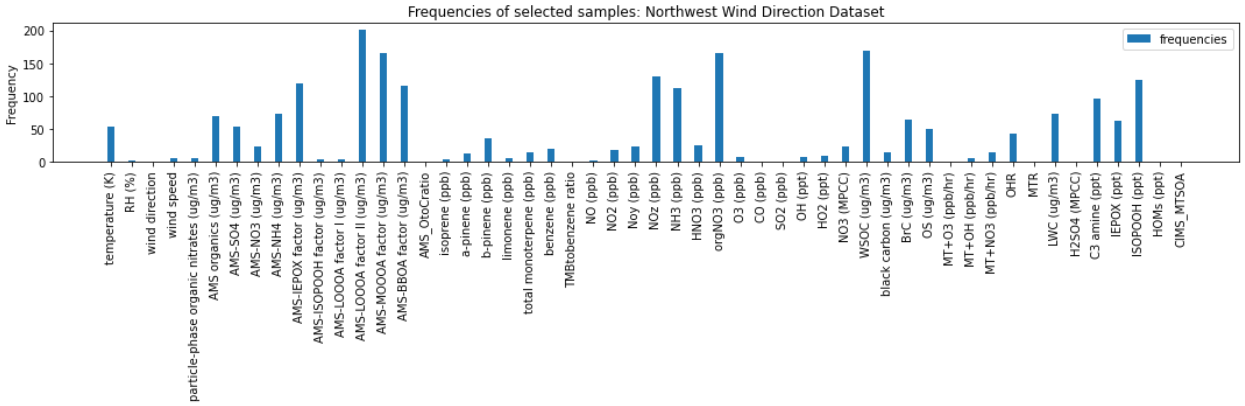

Fig 7. Distribution of selected aerosols within the northwest wind direction dataset (830 x 28). Relevant tracers: temperature, AMS organics, AMS-SO_4, AMS-NH_4, AMS-IEPOX factor, AMS-LOOOA factor II, AMS-MOOOA factor, AMS-BBOA factor, NO_z, NH_3, orgNO_3, WSOC, BrC, OS, OHR, LWC, C3 amine, IEPOX, ISOPOOH

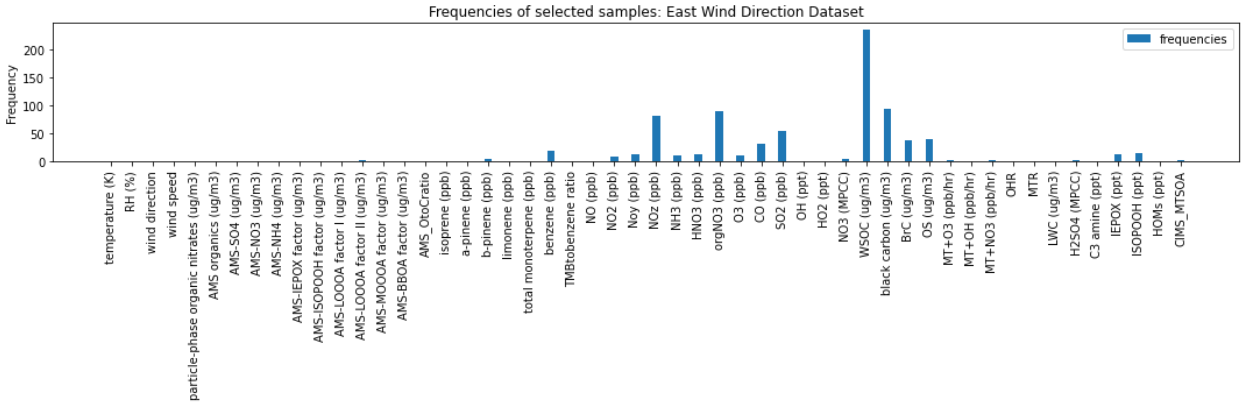

Fig 8. Distribution of selected aerosols within the east wind direction dataset (830 x 127). Relevant tracers: NO_z, orgNO_3, SO_2, WSOC, black carbon

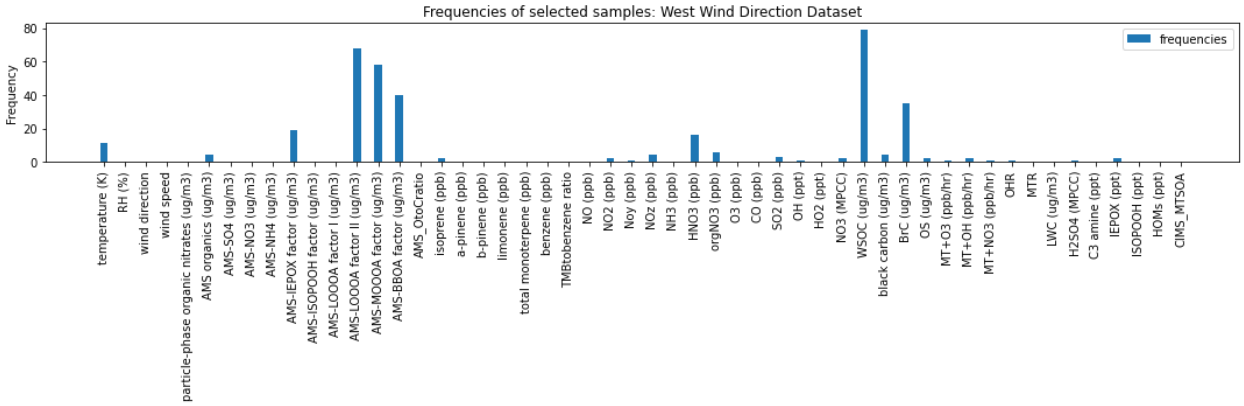

Fig 9. Distribution of selected aerosols within the west wind direction dataset (830 x 127).

Relevant tracers: AMS-LOOOA factor II, AMS-MOOOA factor, WSOC

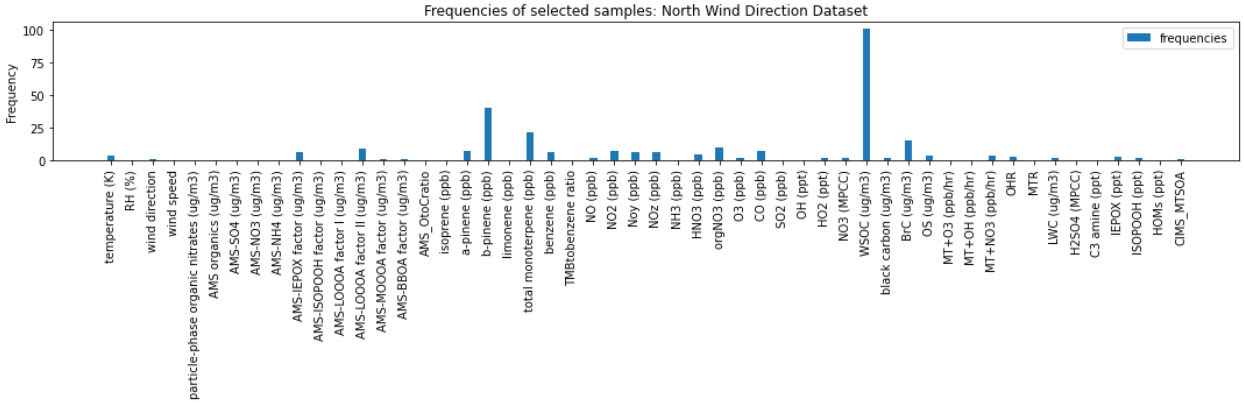

Fig 10. Distribution of selected aerosols within the north wind direction dataset (830 x 67).

Relevant tracer: WSOC

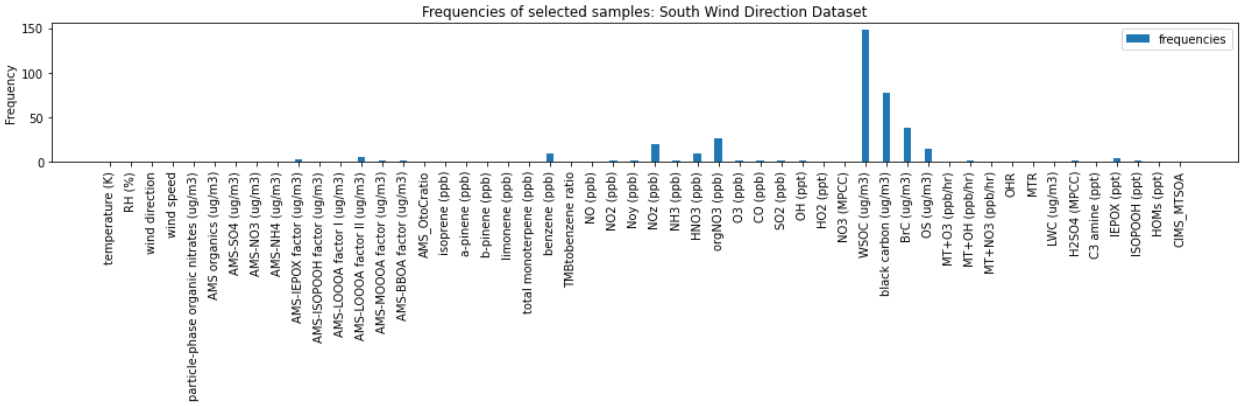

Fig 11. Distribution of selected aerosols within the south wind direction dataset (830 x 187).

Relevant tracers: WSOC, black carbon

Wind Direction	Relevant Tracers
Northeast	a-pinene, b-pinene, total monoterpene, WSOC
Southeast	benzene, NOy, NOz, NH3, HNO3, orgNO3, O3, CO, SO2, WSOC, black carbon, BrC, OS, H2SO4, IEPOX, ISOPOOH
Southwest	AMS-LOOOA, AMS-MOOOA, AMS-BBOA, WSOC, BrC
Northwest	temperature, AMS organics, AMS-SO4, AMS-NH4, AMS-IEPOX factor, AMS-LOOOA factor II, AMS-MOOOA factor, AMS-BBOA factor, NOz, NH3, orgNO3, WSOC, BrC, OS, OHR, LWC, C3 amine, IEPOX, ISOPOOH
East	NOz, orgNO3, SO2, WSOC, black carbon
West	MS-LOOOA factor II, AMS-MOOOA factor, WSOC
North	WSOC
South	WSOC, black carbon

Table 2. Relevant tracers for each wind direction condition.

Sample Index	NE	SE	SW	NW	E	W	N	S
sample index: 7	[]	[]	[]	[]	[]	[]	[]	[]
sample index: 8	['WSOC (ug/m3)']	['NOz (ppb)', 'orgNO3 (ppb)', 'WSOC (ug/m3)', 'black carbon (ug/m3)', 'BrC (ug/m3)', 'OS (ug/m3)', 'IEPOX (ppt)', 'ISOPOOH (ppt)']	['AMS-MOOC factor (ug/m3)']	['AMS-MOOC factor (ug/m3)', 'AMS-BBOA factor (ug/m3)', 'BrC (ug/m3)']	['orgNO3 (ppb)', 'WSOC (ug/m3)', 'black carbon (ug/m3)', 'BrC (ug/m3)', 'OS (ug/m3)']	['AMS-MOOC factor (ug/m3)']	['WSOC (ug/m3)']	['WSOC (ug/m3)', 'black carbon (ug/m3)']
sample index: 9	[]	['BrC (ug/m3)']	[]	[]	[]	[]	[]	[]
sample index: 10	['WSOC (ug/m3)']	['WSOC (ug/m3)', 'BrC (ug/m3)']	[]	['LWC (ug/m3)']	['WSOC (ug/m3)', 'BrC (ug/m3)']	[]	['WSOC (ug/m3)']	['WSOC (ug/m3)']

Table 3. Table showing index 7-10 OAs and their tracer correlations in different wind direction datasets. Empty brackets indicate no tracer correlations. Notice how samples 8 and 10 both correlate with the WSOC tracer in NE, SE, E, N, and S winds, and have different tracer correlations in the SW, NW, and W winds.

Discussion

In examining the distributions of OA to tracer correlations within our different datasets, we find that the distribution of aerosols that significantly correlate with certain tracers varies under different environmental conditions.

Diurnal Differences

Our exploratory data analysis into daytime and nighttime hours verifies diurnal activity reported by Zhang et al., 2018. Relevant day time tracers include benzene and WSOC (**Fig 2**), while relevant nighttime tracers include NO_z, HNO_3, $orgNO_3$, WSOC, black carbon, OS, IEPOX,

and ISOPOOH (**Fig 3)**. The differences in these distributions demonstrate the enhanced nighttime activity of nitrates, which has been previously observed and described in the field of atmospheric chemistry. Aside from increased activity of nitrate correlations, the nighttime dataset shows that OS, IEPOX, and ISOPOOH are also relevant. Therefore, further investigation into nighttime aerosol correlations for such tracers could be the topic of future research.

Wind Direction Differences

Investigation into our aerosol to tracer correlations as a function of wind direction reveals that tracer correlations vary under different wind direction conditions. **Table 2** summarizes relevant tracers under each wind condition, and **Figures 4 – 11** illustrate the unique tracer distributions under each differing wind condition. The differences between distributions indicate that wind direction affects formation of OAs and therefore plays a role in the reaction dynamics of such OAs. Therefore, future research should account for wind direction during sampling and analysis of OAs to confirm this novel activity.

Individual Aerosol Activity

In addition to performing macroanalysis on aerosol to tracer correlations within each dataset, microanalysis on individual aerosols and their tracer correlations was also performed. In this method, we are examining a specific OA's tracers within each wind direction. This analysis was performed for the first 110 OAs in the provided dataset, as these OAs are more atmospherically relevant. **Table 3** provides an example of varying individual aerosol to tracer correlations for aerosols index 7-10. It is important to notice that while some aerosols may correlate together under one condition, they do not necessarily correlate in others. This demonstrates the complex nature of the transformation of organic aerosols. With that said, there are certain aerosols that partially or completely share relevant tracers in certain wind directions,

but gaining insight into this activity requires further advanced analysis and is outside of the scope of this project.

Answering original goals of SOAS 2013 / Conclusion

This research contributes to answering a fundamental question posed by SOAS 2013: What are the chemical and physical processes that control the oxidation of biogenic volatile organic carbons? This research confirmed diurnal differences found between our distributions which results from the presence of different oxidants during daytime and nighttime. This research additionally sheds light on the relevance of atmospheric dynamics toward our understanding of multiphase OA chemistry and hopefully aids in building more robust models of our atmosphere.

Drawbacks of this Study

In generating datasets to represent different environmental conditions, certain conditions resulted in datasets with a severely reduced number of columns. For example, since the condition of northwest winds was very sparse, the number of time points used for this analysis was only 28 (**Fig 7)**. Since our primary metric of correlation was the Pearson coefficient, it's important to note that as the number of time points decreases, the Pearson coefficient is artificially inflated toward 1. To combat this issue, future research should perform sampling over a longer timescale, or more frequently, to ensure a satisfactory number of datapoints for analysis.

It was previously stated that our investigation revealed that wind direction affects formation of OAs and therefore plays a role in the reaction dynamics of such OAs. It is possible that wind direction does not affect such, but rather affects our OA sampling process. For example, if a filter were in one static position and the wind direction was directly perpendicular to the surface of the filter, the described filter would have a greater amount of analyte deposited

onto in comparison to a similar filter in a static position with the wind direction parallel to the surface of the filter. This could explain the correlation activity observed, rather than wind direction having direct implications on OA multiphase chemistry. Future research should utilize filters that are capable of rotation in such a way that they always face perpendicular to the direction of the wind.

References

1. Cruz, C. N.; Pandis, S. N. A Study of the Ability of Pure Secondary Organic Aerosol to Act as Cloud Condensation Nuclei. Atmospheric Environment 1997, 31 (15), 2205–2214.
2. Kanakidou, M.; Seinfeld, J. H.; Pandis, S. N.; Barnes, I.; Dentener, F. J.; Facchini, M. C.; Dingenen, R. V.; Ervens, B.; Nenes, A.; Nielsen, C. J. Organic Aerosol and Global Climate Modelling: A Review. Atmospheric Chemistry and Physics 2005, 5 (4), 1053–1123.
3. Robinson, A. L.; Donahue, N. M.; Shrivastava, M. K.; Weitkamp, E. A.; Sage, A. M.; Grieshop, A. P.; Lane, T. E.; Pierce, J. R.; Pandis, S. N. Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging. Science 2007, 315 (5816), 1259–1262.
4. Song, C.; Gyawali, M.; Zaveri, R. A.; Shilling, J. E.; Arnott, W. P. Light Absorption by Secondary Organic Aerosol from α-Pinene: Effects of Oxidants, Seed Aerosol Acidity, and Relative Humidity. Journal of Geophysical Research: Atmospheres 2013, 118 (20), 11–741.
5. Jang, M.; Ghio, A. J.; Cao, G. Exposure of BEAS-2B Cells to Secondary Organic Aerosol Coated on Magnetic Nanoparticles. Chemical research in toxicology 2006, 19 (8), 1044–1050.
6. Ito, T.; Bekki, K.; Fujitani, Y.; Hirano, S. The Toxicological Analysis of Secondary Organic Aerosol in Human Lung Epithelial Cells and Macrophages. Environmental Science and Pollution Research 2019, 26 (22), 22747–22755.
7. Eaves, L. A.; Smeester, L.; Hartwell, H. J.; Lin, Y.-H.; Arashiro, M.; Zhang, Z.; Gold, A.; Surratt, J. D.; Fry, R. C. Isoprene-Derived Secondary Organic Aerosol Induces the

Expression of MicroRNAs Associated with Inflammatory/Oxidative Stress Response in Lung Cells. Chemical research in toxicology 2019, 33 (2), 381–387.

8. Eol.ucar.edu. 2021. SOAS | Earth Observing Laboratory. [online] Available at: <https://www.eol.ucar.edu/field_projects/soas> [Accessed 3 May 2021].
9. Zhang, H.; Yee, L. D.; Lee, B. H.; Curtis, M. P.; Worton, D. R.; Isaacman-VanWertz, G.; Offenberg, J. H.; Lewandowski, M.; Kleindienst, T. E.; Beaver, M. R. Monoterpenes Are the Largest Source of Summertime Organic Aerosol in the Southeastern United States. Proceedings of the National Academy of Sciences 2018, 115 (9), 2038–2043.
10. Xu, L.; Guo, H.; Boyd, C. M.; Klein, M.; Bougiatioti, A.; Cerully, K. M.; Hite, J. R.; Isaacman-VanWertz, G.; Kreisberg, N. M.; Knote, C. Effects of Anthropogenic Emissions on Aerosol Formation from Isoprene and Monoterpenes in the Southeastern United States. Proceedings of the National Academy of Sciences 2015, 112 (1), 37–42.
11. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B. E.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J. B.; Grout, J.; Corlay, S. Jupyter Notebooks-a Publishing Format for Reproducible Computational Workflows.; 2016; Vol. 2016.