Spatial and Temporal Distribution of Airborne Ultrafine Particles, PM2.5, and PM10 in Three Communities in the Greater Boston Area

A thesis submitted by

Camille L. Gimilaro

in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil and Environmental Engineering

Tufts University

May 2025

Advisor: John Durant, Ph.D.

Reader: Shan Jiang, Ph.D.

Abstract

Exposure to pollution in the form of microscopic airborne particles contributes to cardiovascular and pulmonary health risks. In urban areas, residents are exposed to high concentrations of airborne particles in combustion emissions from industries, power plants, and motor vehicles. These sources tend to be intermittent and spatially heterogeneous across cities; thus, it is important to make measurements at multiple sites for long periods of time in urban areas to better understand source contributions and to more accurately quantify exposures. This thesis describes the results of a year-long campaign to measure airborne particles of different sizes in the three adjacent communities in the Boston (MA) area – Malden, Everett, and Charlestown. Except for Malden, these communities have received relatively little attention from the air quality monitoring community. In this study we measured the number concentration of particles ranging in size from 7 – 10,000 nm in diameter as well as the mass concentrations of particles smaller than 10,000 nm at stationary sites centrally located in the three communities. The results indicate that the concentrations of ultrafine particles (<100 nm in diameter) were generally lower in Malden compared to more urbanized Everett, and highest in Charlestown. Similar temporal patterns were observed in the three communities, but differences were observed in terms of source contributions by wind direction. The concentrations in Charlestown were unexpectedly elevated when winds were from the east, suggesting a strong source to the east. These results are discussed in the context of providing interpretable data products and information to stimulate community engagement.

Acknowledgements

I would like to acknowledge my advisor Dr. John Durant, and my reader Dr. Shan Jiang for taking time to edit my writing. I would also like to thank them for teaching me applicable skills necessary to complete this project.

I would also like to thank the CLEANAIR team including Patrick Herron, Greer Hamilton, and Jennifer Delgado for collaboration on this project. They went above and beyond to provide feedback and guidance during the data processing effort. Jennifer Delgado deserves a distinct acknowledgement for her fieldwork efforts as she collected data year-round during this air monitoring campaign. I would not have been able to complete this work without her persistence.

Lastly, I would like to acknowledge my friends and family for their unwavering support during the last two years. A special acknowledgement to Kaustuv Ray, a prior student, who helped me learn how to process condensation particle counter data and began the work on the Python codes that were used in the cleaning and visualization process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

A. General Overview of UFP Impact on Human Health

Ultrafine particles (UFPs) are particulate matter less than one hundred nanometers in diameter. Their small size along with a higher number concentration and a larger surface area to volume ratio than larger sized particles, enable them to penetrate deeper into the lungs and become more biologically reactive than larger-sized particles. First, the UFPs are activated in the airway, which results in the activation of endothelial and mononuclear cells (Utell et al., 2000). Endothelial cells line the blood vessels and regulate exchanges between the blood vessels and surrounding tissue (Alberts, 1970). Meanwhile, mononuclear cells are blood cells with a single round nucleus (LifeSci, n.d). Once endothelial cells are activated, more inducible nitric oxide synthase (iNOS) is produced, which continuously produces nitric oxide (NO) until the enzyme is degraded. This production of NO is involved in the physiological process of many human diseases including asthma (Science Direct, 2007). Additionally, activation of the endothelial and mononuclear cells can also result in a hypercoagulable state, as part of a delayed response, which increases the tendency for blood to clot outside of bleeding (Senst, 2023). In turn, this increases the likelihood of a cardiac event occurring. Additionally, mononuclear cell activation increases the production of serum amyloid A (SAA), as part of an acute phase response, which increases inflammation and can result in increased blood clotting and coronary events (Utell et al., 2000).

Past literature affirms the association between increases in UFPs and an increased presence of health issues. For instance, prenatal exposure to UFPs is associated with asthma development in children, independent of temperature and NO_2, through monitoring 376 mother-child dyads since pregnancy (Wright et al., 2021). This evidence presents an intergenerational

pulmonary health effect occurring from UFP exposure. Additionally, airport UFP exposure is associated with increased acute systemic inflammation for adults with asthma (Habre et al., 2018). This study monitored 22 adults near the Los Angeles Airport (LAX) for IL-6, a biomarker for increases in the acute phase response. The result of this study is evidence of short-term lung inflammation associated with UFP exposure. Lastly, a third study demonstrates the effects of the frequency and intensity of UFP peaks on biomarkers for inflammatory and blood lipids in the Greater Boston area (Lin et al., 2022). Findings show increases in peak exposures are correlated with increases in TNF-R11, which is an inflammatory biomarker, and decreases in high-density lipoprotein (HDL) and triglycerides (biomarkers for healthy blood lipids). Additionally, the increases in intensity in UFP peaks were correlated with increases in IL-6 and decreases in total cholesterol, which are biomarkers for inflammation and blood lipids respectively, (Lin et al., 2024). This literature further affirms the correlation between UFP exposure, higher short-term inflammation, and worsening blood-related health.

B. Spatial Variation of UFPs

A common source of UFPs is combustion emissions (Cooper and Ally, 2015), and several studies demonstrate an increase in UFPs with proximity to vehicular traffic and flight arrivals. A study in the Seattle-Tacoma metropolises region found that the highest concentration of UFPs was on a major roadway (Austin et al., 2021). This study is backed by two studies in the Greater Boston area. One study demonstrated a higher 1-minute mean particle number concentration (PNC), a proxy for UFP measurements, on roads than at central sites in Chelsea and Revere and at nearby residences (Simon et al., 2017). Another study found that although there is variance by neighborhood and time of day, pollutant levels increase with highway proximity (Patton et al., 2014). Furthermore, in addition to vehicular traffic-related UFPs,

another study found that aircraft arrivals per hour are associated with measured PNC (Chung et al., 2024). This group found that the arrival aircrafts contribute up to 50% of total UFPs, during hours with aircraft activity, at a site three kilometers from the airport. Previous literature recognizes the impact that both air and vehicular traffic have on the UFP concentrations in nearby communities.

C. Temporal Variation of UFPs

Previous research on UFP monitoring campaigns provide valuable insight into temporal pollution dispersion within urban areas. For instance, models have been employed to evaluate the spatial and temporal trends of particulate matter (Pakbin et al., 2010), highlighting periods and locations of elevated UFP concentrations, including daily, seasonal, and spatial trends. Karumanchi et al., studied the seasonal effects of UFPs in Montreal, Canada by monitoring 249 sites in 20-minute campaigns in the summer and winter, finding mean concentrations to be 16,593 #/cm^3 in the winter and 8,919 #/cm^3 in the summer, of 7-1000 nm particles. Although these findings indicate higher UFP concentrations in the winter, wind-direction also changed throughout their study; winds blew from the west, south-west, and northeast in the winter and south and southwest in summer. Findings by Karumanchi et al. also lack a control on the day of the week or time of day which could have influenced the results. Another study by Wang et al., also examined the spatial and temporal variation in UFPs through 12 monitoring sites in Rochester, NY and found that concentrations of UFPs are elevated in the fall, and winter near highways, off-road diesel engines and residential wood stoves. However, wind patterns are not controlled for and differ considerably for the duration of this study as each site was monitored for less than 20 minutes at each site. A fourth study in Copenhagen, Denmark, by Bergmann, et. al. developed a model to approximate the annual mean particle number concentrations in 37

residential short-term based on the temporal variation between (72-hour) monitoring campaigns and a year-long centrally located reference site. Temporal variation was determined through varying measurements in winter and summer, and Monday-Thursday vs. Thursday- Monday campaigns. However, although Bergmann, et. al expected the PNC at the residential sites to be temporally correlated with the PNC at the reference site, this was only sometimes the case, indicating a flaw in their model. Bergmann, et. al discovered the highest PNC in the spring, and did not find elevated PNC at sites less than 5 km from the airport. They also did not factor wind direction into their model which may be an oversight. The air monitoring campaigns aimed to understand temporal variation within a city have the potential to equip communities with critical information about when air pollution levels are expected to be the highest when all critical variables are considered.

D. Larger Particle Sizes

Another metric to measure air pollution is using particulate matter mass concentration in the form of PM2.5 and PM10. PM2.5 measures particulate matter mass concentration, in units of micrograms per cubic meter less than 2.5 micrometers in diameter, and PM10 is micrograms per cubic meter of particulate matter less than 10 micrometers in diameter. Generally, PM2.5 accounts for more anthropogenic pollution sources such as combustion emissions from transportation, while PM10 accounts for a larger proportion of natural particles such as pollen and dust. However, mass concentration is not the most important metric of potential health effects. Stronger associations have been observed between fine particles and health, making particle number concentration a better metric for potential health effects (Lighly et al., 2000). This is because particle number concentration puts a higher weight on ultrafine particles, which are high in number, while their individual mass is almost negligible. Low-Cost Particulate Matter

Sensors (LCPMS) of less than 1,000 dollars can provide mass concentration data to supplement the availability of fine particle data in budget restricted areas. In a laboratory setting, with controlled relative humidity and temperature, the Dylos DC1700 LCPMS was found to have a linear slope of 1.03 and a (R^2) of 0.94 when compared to the $PM_{2.5}$ mass concentrations for the research grade instrument: Grimm 11R, for particles greater than 2 microns in median aerodynamic diameter (Oluwadiro 2022). The close agreement of these two instruments suggests that the PM2.5 dataset from the Dylos DC1700 LCPMS may be accurate even though the device is a low-cost sensor.

E. UFP Clustering Techniques

Researchers have used modeling methods such as K-Means clustering to identify clusters representing different particle sizes. K-Means clustering is an Expectation Maximum (EM) algorithm that produces clusters with a minimum sum of the distance from each point to the cluster centroids and a maximum distance between cluster centroids, grouping data based on a number of variables. Clustering can reduce the dimensionality of the data set, which can allow further characterization of spatial and temporal trends through comparing the cluster means as opposed to each data point. Beddows et al. studied cluster patterns grouped by particle diameter and found that the patterns varied between sites and were different depending on whether it was an urban or a rural area (Beddows et al., 2009). In rural areas, the clusters exhibited distinctive diurnal behavior, wind speed, and pollutants, while in urban areas some clusters were unique to each sampling location. Another study in downtown Toronto grouped ultrafine particle data from a Scanning Mobility Particle Sizer (SMPS) into eight clusters (Sabaliauskas et al., 2013). Through examining the relationship between clusters, meteorological factors, seasonality, relationship to PM2.5, and diurnal patterns, Sabaliauskas et al.were able to identify spatial and

temporal patterns near a road. Findings indicate that clusters differed based on traffic patterns, long range transport, time of day, and seasonality. A third study used K-Means clustering to distinguish air pollution in a rural area 35-km downwind of Seoul, Korea. In this study 7 clusters were created to split condensation particle counters and SMPS data based on particle size. The following categories for pollutants were determined: background, local emissions, large particle traffic, small particle traffic, fresh nucleation (ultrafine particles less than 25 nm), aged nucleation and transported nucleation (Lee et al, 2021).

F. Environmental Justice

According to EJScreen, Charlestown, Everett and Malden face significant environmental challenges and include historically vulnerable and underserved populations. These localities rank in the 86th to 98th percentile in the state for Diesel Particulate Matter. Additionally, all localities are in the 99th percentile for Air Toxics Cancer Risk and in the 96th to 99th percentile for Air Toxics Respiratory Risk. These environmental burdens align with demographic indicators showing that these areas rank in the 57th to 90th percentile for populations of color and the 68th to 89th percentile of linguistically isolated individuals (Workplan: Lower Mystic Air Quality).

G. CLEANAIR Study

This project is a partnership with the Community Led Improvement of Air Quality and Health in the Lower Mystic (CLEANAIR) study. CLEANAIR is a funded study submitted by Patrick Herron, the Executive Director of the Mystic River Watershed Association (MyRWA). The goal of CLEANAIR is to improve the air quality and health of the communities in the Watershed most burdened by Traffic Related Air Pollution (TRAP). The CLEANAIR study aims to do this through deploying air quality monitors, informing infrastructure decisions, limiting exposure to pollutants and deploying a community advisory board to build capacity among

residents. The communities that were chosen for the CLEANAIR study were the cities of Everett, Malden, Charlestown, and East Boston in the city of Boston. CLEANAIR is a three-year project (2023-2026) centered on three long-term ambient air pollution monitoring campaigns in Malden, Everett, and Charlestown, and short-term studies focused on a specific area of concern in all four communities. The partnership with CLEANAIR involves CLEANAIR sharing their data and feedback with me for this project, and in turn, I provide them with a year's worth of analyzed data of their long-term monitoring sites, information on how to process the data, and summary graphics of their data.

H. Goals and Objectives

The goal of my research is to analyze the spatial and temporal trends of ambient airborne ultrafine particles, PM2.5 and PM10 in Malden, Everett and Charlestown. To do this, I will first visualize temporal and spatial trends in UFPs, small and large particle number concentration, PM2.5, and PM10 to find potential source attribution using boxplots and polar scatterplots. Then, I will explore which meteorological factors and time of day contribute to the highest 95^{th} percentile of UFP pollution. Lastly, I will use a clustering method to understand which meteorological conditions are associated with different particle sizes.

CHAPTER 2: DATA AND METHODOLOGY

A. Study Area

This study involves a year-long campaign to monitor ambient air pollution in distinct localities of Everett, Malden, and Charlestown. Monitors were placed close to the center of the community to get an estimate of the air pollution in said community. For Malden, the monitoring station is at 350 Main Street next to the community's growing gaming district, many food establishments, and the city's famous New England Coffee Roasters. The Everett monitoring station is placed at the Everett City Hall, nearby a main commercial strip of Broadway Avenue. Lastly, the Charlestown site is at the historic Bunker Hill Museum across the street from the Bunker Hill monument. All three sites are among city centers enclosed by primarily commercial and residential areas. The monitor in each city center can serve as an approximation for nearby residences because it is centrally located and away from major roads or highways. The monitors are also located on the roofs of the buildings to account for well-mixed air. To see the three sites in proximity to each other see Figure 1.

Each monitoring station contains a condensation particle counter (CPC) (TSI Model 3783; Shoreview, MN), a Dylos 1700-PM, a Dylos 1100, a pump, a heater, an air conditioning (AC), a CPC shelter, a Dylos shelter, and a pump shelter. The CPC, which is located within the largest shelter, measures the particle number concentrations which serves as a proxy for ultrafine particles. Similarly, the Dylos 1100 measures particle number concentration above 500 nm in diameter (small) and particle number concentration above 2,500 nm in diameter (large). Meanwhile the Dylos 1700-PM measures PM2.5 and PM10. Both Dylos are inexpensive devices that use a laser to optically count the number of particles that pass through the vent. Other than the air quality measuring devices the station has a pump which supplies fresh water through the

CPC, an AC and a heater (located inside the box, see Figure 4). The heater and AC are used with a temperature dependent power strip which is used to regulate the shelter's temperature between 50-100 °F so it does not freeze or overheat. Meanwhile, the Dylos are in a separate shelter on top of the CPC shelter. The Dylos shelter covers the Dylos on five sides, allowing air to flow in one open side. The Dylos shelter contains reflective tape which serves as temperature regulation. Bricks and cement blocks are used to keep the monitoring shelters in place. See Figure 3 for an example of the monitoring campaign at the Everett site.

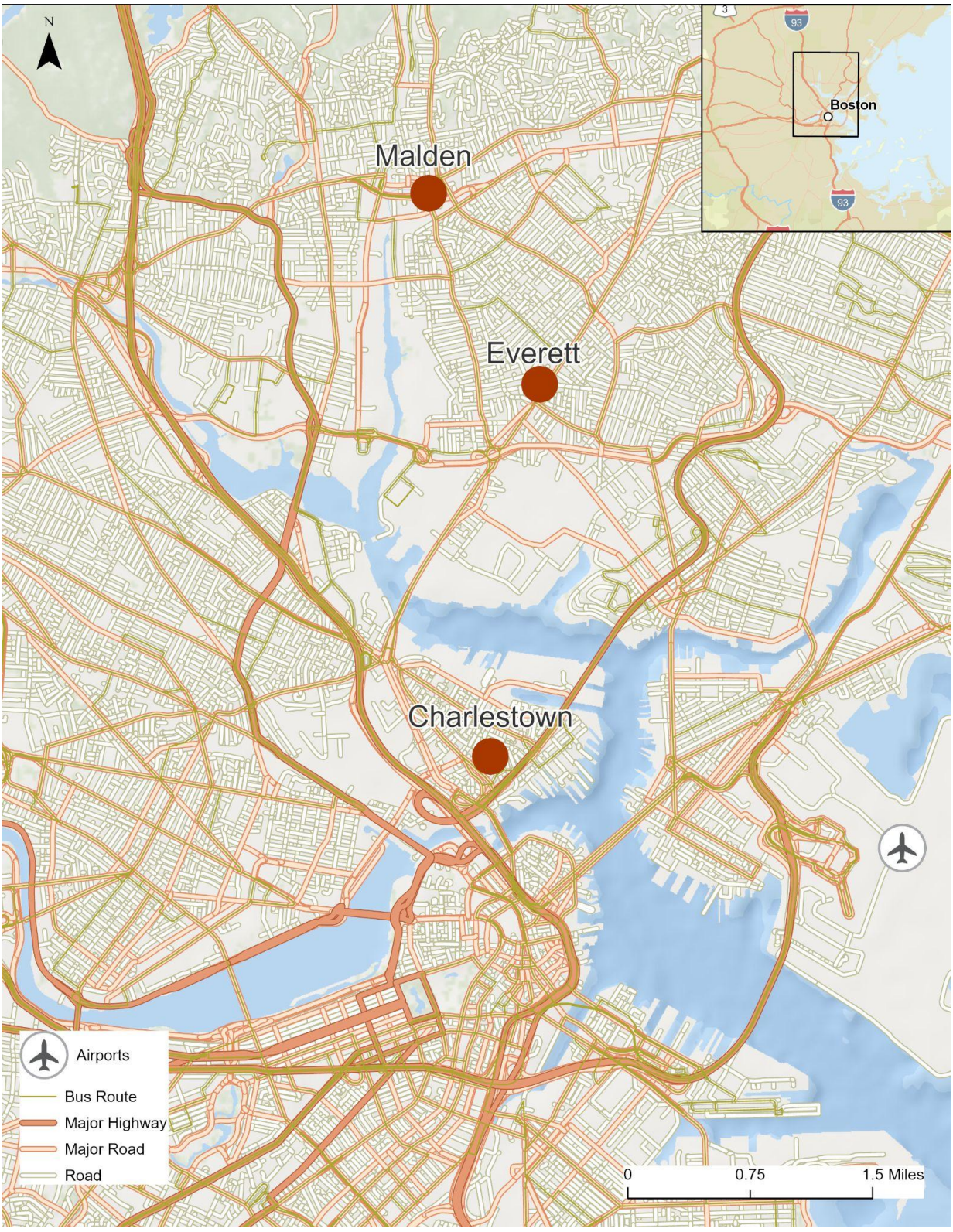

Figure 1 Study area. Each circle represents the community's monitoring site. From North to South, sites are for the communities of Malden, Everett, and Charlestown. Sites are located on roofs in the middle of each community. The sites are located at 350 Main Street, Malden, 484 Broadway, Everett, and 43 Monument Square, Charlestown.

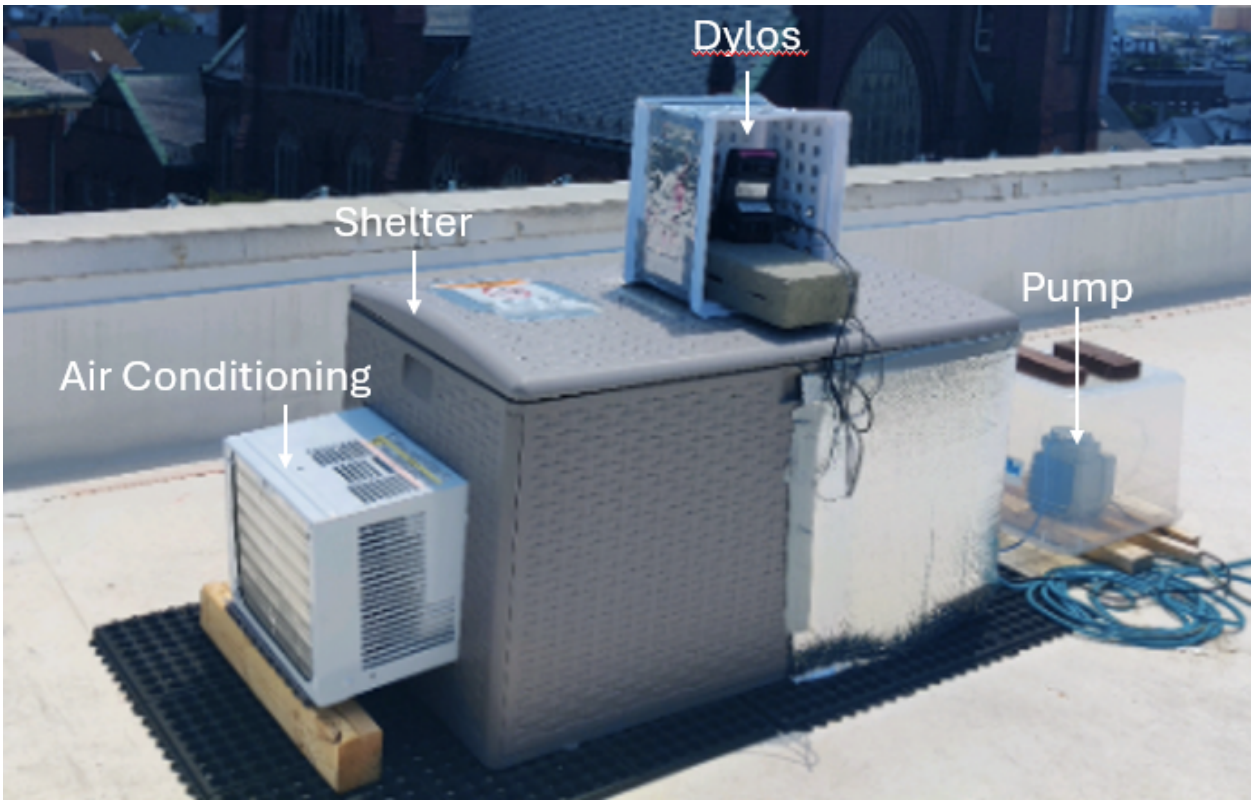

Figure 2. Site set-up in Everett.

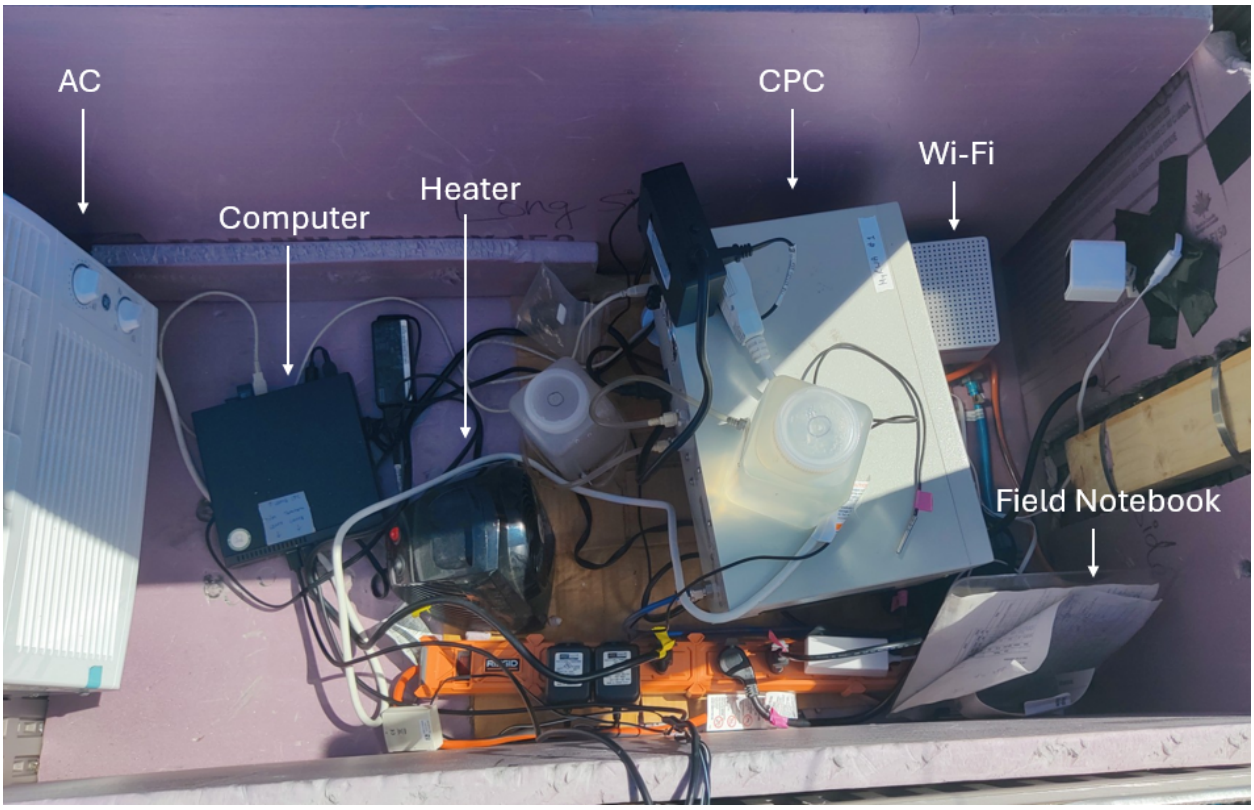

Figure 3. Inside view of CPC shelter in Everett.

B. Measurements

Ultrafine particle (UFP) measurements were made using condensation particle counters (CPC) (TSI Model 3783; Shoreview, MN). The CPCs function by condensing water vapor on particles so they grow larger and can be optically detected. The CPC can detect particles down to 7 nanometers, capturing a large majority of UFPs. This model of CPC will also measure particles larger than UFPs (up to ~1,000 nm), but compared to the number of UFPs, particles larger than 100 nm typically only represent a small percentage of total particle number concentration, making it an adequate proxy of UFP count. UFP data from the CPC model 3783 is collected at a frequency of one measurement per second. Each site was visited once a week for instrument maintenance and data collection. Data collected from the CPC was uploaded in a .DAT format

and subsequently cleaned using Python to remove errors in the dataset. The cleaning process is further explained in the data processing section.

In addition to a CPC, each site also had a Dylos DC 1100-PRO and a Dylos 1700-PM. The Dylos DC 1700-PM measures the mass concentrations of particles 2.5 micrometers and smaller (PM2.5) and particles 10 micrometers and smaller (PM10). Both particle sizes are regulated under the Clean Air Act. The Dylos DC 1100 PRO measures the number concentration of particles greater than 0.5 microns and greater than 2.5 microns in diameter. Both Dylos are inexpensive devices that use a laser to optically count the number of particles that pass through the vent.

C. PM Data Cleaning

Dylos and CPC data were cleaned with separate Python scripts (Appendix C). Data from the Dylos 1100s was recorded on the Dylos Logger software and exported each week onto a shared Box drive in a text file format. The minutely files were then aggregated on MATLAB and plotted to check for accuracy. Once the data was properly aggregated, duplicated data rows (if any) were removed, and the data was sorted based on ascending timestamp. Zero-values were also removed. At the end of this cleaning process the Dylos 1100 data was converted to hourly averages, and the concentration units were converted from #/0.01 ft^3 (reported by the instrument) to #/cm^3. Similarly, the data from the Dylos 1700-PM is collected on Tera Term and saved in a text format. The same steps were followed with the Dylos 1700-PM with a couple of additional steps. The Dylos 1700-PM underwent additional cleaning from a Python script to adjust for relative humidity and oversaturation. Relative humidity over 60% can lead to overestimation of particle mass concentration on the Dylos 1700-PM. Dylos recommended the following formula to correct for relative humidity:

$$Adjusted\ Mass\ Concentration = Initial\ Mass\ Concentration * \frac{60}{\%RH}^{6}$$

Another adjustment was recommended for oversaturation, which causes coincidence loss on the Dylos 1700-PM devices. Coincidence loss occurs when more than one particle is detected in the same counting zone at the same time, which would cause two particles to be counted as one. Coincidence loss occurs on the Dylos 1700-PM when there is over 300 ug/m^3 PM2.5 or PM10. Using a table provided by Dylos for DC 1700-PM and True PM (Table 5), I plotted the values in Excel and derived an adjustment formula for Dylos 1700-PM measurements using an exponential best fit line.

$$True\ PM = 56.5 * e^{0.0057*PMreading}$$

This over saturation formula was applied to PM readings between 300 µg/m^3 and 550 µg/m^3. Since Dylos disclosed that readings over 550 µg/m^3 are unreliable, and this is about 1,300 µg/m^3 after correction, all concentrations greater than 550 µg/m^3 were converted to 1,300 µg/m^3 representing the maximum reliable value for PM.

D. CPC Data Cleaning

The CPC data was cleaned using Python scripts. Most of the CPC data from the three sites was downloaded each week and saved in a .DAT file format to a secure drive at Tufts. All the .DAT files were aggregated at a secondly resolution. The date and time column were aggregated into a single column and the particle number concentration column and error code column were extracted into a smaller data frame. Because some of the CPC data was recorded directly from the CPC onto the AIM software, an additional Python script was used to convert the exported AIM files into a user-friendly format with three columns: date/time, concentration, and error message. All the resulting comma separated value files were concatenated to the

aggregated .DAT files to create one large secondly data file and duplicate data was removed. Because the error messages were different depending on whether the data was collected on AIM or on the USB directly connected to the CPC, another Python script was used to read all the error messages and write them to an Excel sheet. This Excel sheet was used to manually sort out fatal errors including “Optics Temperature”, “Laser Status”, “Pulse Height Fault”, and “Nozzle Pressure”. A new Excel sheet was manually created from all the printed errors that had fatal error messages. The error message column in the concatenated data was changed to a string type and Python code removed all rows of data containing a fatal error. After fatal errors were removed particle number concentrations below 300/cm^3 were removed. During the cleaning process 9% of secondly CPC data was removed from the Malden CPC out of the total seconds the CPC was in the field with power, 5% from the Charlestown CPC, and 4% from the Everett CPC. Lastly, the data was sorted by date ascending, plotted on MATLAB to ensure that the errors were removed, and converted to hourly averages.

E. Weather Data

Weather data was downloaded for 2023, 2024, and 2025 from the National Oceanic and Atmospheric Administration’s National Centers for Environmental Information website, Local Climatology Data. This data contains information on precipitation, pressure, and wind conditions from the Boston Logan International Airport. To use this data, three data files, one for each year, were merged. Then the Pandas Backfill (Pandas, 2024) function was used to fill data at a minutely resolution so empty data cells were assumed to be the next reading. This may add some error in the dataset but allows for every hour, on the hour to have data approximations for pressure, wind, relative humidity and temperature.

F. Data Recovery

The backfilled NOAA NCEI data was merged with hourly resolution air pollution data for PM2.5, PM10, small particles (<0.5 μm), large particles (< 2.5 μm), and ultrafine particles. Data recovery was calculated based on the number of days each device was in the field with power going to the instruments (Table 1).

Table 1. Data recovery based on days the instrument was in the field and electricity was on at the shelter.

	Start Date	End Date	Days instrument not in field or power outage	Notes	Days in the field	% Recovered from Fieldwork	% Recovered after removing Errored Data (CPC)	% Recovered after Removing Low Concentration Data (CPC)
Malden CPC	11/14/23	1/30/25	41	Power Outage and Instrument Repair	402	80%	72%	71%
Charlestown CPC	2/9/24	2/5/25	17	Brought back for Repair and Heatwave	345	85%	80%	80%
Everett CPC	10/26/23	1/30/25	12	Brought back for Repair and Heatwave	450	65%	61%	61%
Charlestown Dylos 1700	6/26/24	2/5/25	0		224	88%		
Everett Dylos 1700	6/26/24	2/4/25	0		223	95%		
Malden Dylos 1700	5/15/24	2/5/25	54	Not in field from 5/18 to 7/3; power outages	212	97%		
Malden Dylos 1100	3/26/24	1/31/25	87	Not in field from 4/12 to 6/26; power outages	224	80%		
Charlestown Dylos 1100	7/19/24	2/5/25	0		201	73%		
Everett Dylos 1100	6/26/24	2/4/25	25	Taken out of field due to a shortage of field computers	198	85%		

G. Analysis of Spatial and Temporal Trends

Spatial and temporal trends were analyzed using Seaborn (Watson, 2024), a Python data visualization library, to create boxplots and polar scatter plots. Boxplots were used to show the annual trends in hourly particulate matter at each site for PM2.5, PM10, and the number

concentration of ultrafine particles, particles > 0.5 microns, and particles > 2.5 microns. The boxplots include whiskers that represent 1.5 times the interquartile range. Additional boxplots were created representing each hour of data at each site for each pollutant size. These boxplots show the diurnal trends of air pollution. Polar scatter plots at each site represent the concentrations of particulate matter based on wind speed and direction. Points are plotted in the direction the wind is blowing from, and the radius represents the wind speed in m/s. These graphs give a spatial representation of where the highest concentrations of air pollution are coming from. Each point on the scatter plot represents an hourly average concentration.

H. 95th Percentile of UFPs

To understand the conditions contributing to the highest 5% of hourly average UFPs, I first merged the data from all three sites and found the 95th percentile UFP concentration. This benchmark was used to filter the data frame and make plots. Python Seaborn's (Watson, 2024) package was used to create histograms to highlight the UFP concentration at the three sites, along with the month of the year, day of the week and hour of the day that they occurred. The wind direction associated with these high hourly averages was also analyzed as was the hourly atmospheric pressure.

I. Particle Size Comparison

A Gaussian Mixing Model was used to explore the relationship between mass concentration and particle size at the three sites. A Gaussian Mixing Model (GMM), an unsupervised machine learning model, was applied to the merged dataset to cluster the data into four clusters. These clusters are labeled based on each cluster's mean concentration of ultrafine particles, PM2.5 and PM10 concentrations (Table 4). The dominant particle sizes of each cluster will be compared to the meteorological conditions at the time of the data to understand which weather conditions are associated with low pollution, medium pollution, high particulate matter and high ultrafine particles at these sites. The GMM is a probabilistic Expectation-Maximum (EM) model. Like K-Means it uses a two-step convergence algorithm that calculates the probability of each data point belonging to each cluster based on parameter estimates. The probability of each point belonging to a cluster is:

$$P(z_n = k \mid x_n) = \frac{\pi_k * N(x_n|\mu_k \sum k)}{\sum_{k-1}^{K} \pi_k * N(x_n|\mu_k \sum k)},$$

where $z_n = k$ indicates which Gaussian the point belongs to, $\pi_k$is the mixing probability of the k[th] Gaussian, and $N(x_n|\mu_k \sum k)$is the Gaussian distribution with mean $x_n|\mu_k$ and covariance $\sum k$.

The second step of the GMM calculates the overall likelihood of observing each data point in each assigned cluster:

$$P(x_n) = \sum_{k-1}^{K} \pi_k * N(x_n|\mu_k \sum k)$$

Where $P(x_n)$ is the likelihood of observing data point x_n

Next, the algorithm updates the meaning, covariance and mixing coefficient. This algorithm is also an expectation maximum algorithm similar to the K-Means clustering algorithm but instead of linking data points based on the nearest cluster centroid, GMM uses a density based Gaussian probability to create the ideal clusters. Both techniques were tested, but the GMM performed better at assessing the differences in particle sizes. K-means divided the air pollution data primarily based on ultrafine particle concentration. Meanwhile GMM divided the dataset into more meaningful categories (Low Pollution, High PM, Medium Pollution, and High UFPs). This is because the simple formula for K-means aims to minimize the spread within a cluster, and ultrafine particles exhibited the greatest spread within the dataset, therefore data was grouped primarily based on ultrafine particle concentrations. However, since GMM is a probabilistic normal distribution model, which is better suited for ellipsoid shapes with varying spread, the clusters were not limited to ultrafine particle concentration differences alone. The number of clusters was determined based on the Davies-Bouldin Index and the Silhouette Scores as well as a limit of 2-6 clusters that would be visually appealing for comparison in this project. Lastly, violin plots are created with each of the four categories to show the differences in wind speed, wind direction, precipitation, relative humidity, and station pressure between the four groups to further analyze what conditions contribute to various air pollution.

CHAPTER 3: RESULTS

Overall, there were clear differences in the ultrafine particle concentrations at the three sites, but no significant differences in the number concentrations of larger particles or in the mass concentrations of PM2.5 and PM10. The following sections describe spatial and temporal trends in the data.

A. Spatial Trends

The Charlestown site has an annual mean ultrafine particle concentration of 16,400/cm^3, which is higher than Everett: (11,800/cm^3) and Malden: (10,700/cm^3). The Malden site has 5.99 small particles (>500 nm) /cm^3 and 0.368 large particles (>2,500 nm) /cm^3. In contrast, Everett and Charlestown have lower large particle counts than Malden: 4.89-5.16 particles >500 nm/cm^3 and 0.299- 0.302 particles >2,500 nm. Additionally, the mean PM2.5 and PM10 mass concentrations are higher at the Malden site, which has 8.97 ug/m^3 PM2.5 and 30.7 ug/m^3 PM10, than at the Everett and Charlestown sites, which have 5.44-5.91 PM2.5 and 6.78-10.2 PM10 (Table 2). The median annual PM2.5 and PM10 concentrations are roughly the same at all three sites. The median PM2.5 at Malden is 3.64 ug/m^3, 3.60 ug/m^3 at Everett, and 3.23 at Charlestown while median PM10 is 6.96 ug/m^3 at Malden, 6.78 ug/m^3 at Everett and 6.44 ug/m^3 at Charlestown (Table 3). The highest 95th percentile hourly average ultrafine particle concentrations occurred more often at the Charlestown site than at the other two sites as depicted in Figure 16. While occasionally the highest hourly average ultrafine particle concentrations occur at the Malden or Everett sites, the highest hourly average seen at these sites is 80,000 /cm^3 as opposed to concentrations up to 140,000 /cm^3 at Charlestown.

As shown in Figure 15, at Charlestown the highest ultrafine particle concentration occurred during easterly winds. At the Everett site, the highest concentrations of ultrafine particles occurred during southeasterly winds (Figure 14). In contrast, at the Malden site the highest ultrafine particles occur during times of southwesterly winds, and there is evidence to suggest a local source near to the monitoring site (Figure 13). The polar scatter plots in Figures 13 through 15 show that at all three sites, the highest PM2.5 and PM10 occur when the wind is blowing from the northwest. The number concentrations of small and large particle counts do not show a clear association with wind direction; however, the polar scatter plots show an association between low wind speeds and higher small and large particle number concentrations.

B. Temporal Trends

As shown in Figures 9 through 13, at each site and for each particle size, there is consistently a bimodal peak in daily particulate matter. The peaks occur at around 7 am and 7 pm but vary by site and by particle size. For example, the evening peak ultrafine particle concentration in Charlestown begins earlier than the peak at the other two sites and lasts longer. Charlestown also has a higher small particle concentration on weekend mornings compared to weekdays as evidenced by the 75th percentile mark on the boxplots (Figure 10) being up to 12/cm^3 while only up to 6.5/cm^3 on weekday mornings. Typically, concentrations on weekdays are generally higher than on weekends. Additionally, ultrafine particle concentrations exhibit the most pronounced diurnal peaks compared to the larger particle sizes except for Malden PM2.5 and PM10. Furthermore, Malden PM2.5 and PM10 exhibit higher concentrations during the day than at night compared to the other two sites on both weekends and weekdays.

Among the top 5% UFP concentrations, there did not seem to be any significant day of week differences at any of the sites, though the Malden site rarely has a top 5% hourly UFP

concentration on the weekend (Figure 16). Additionally, while the Everett and Malden sites have hourly UFP concentrations in the top 5% during colder months (October through March), the Charlestown site has the highest number of hourly UFP concentrations in the top 5% in the summer months (June through August) (Figure 16). Another difference in the highest 5% hourly UFP concentrations is the time of day these occur. At both the Malden and Everett sites, the histogram showing the number of hours with the highest 5% of UFPs was bimodal (Figure 16) and had the same shape as Figure 9 showing the diurnal trends with peaks in the morning around 7 am and in the evening around 7 pm. However, at the Charlestown site, the highest 5% of UFPs shows one peak in the midafternoon around 5pm (Figure 16). In summary, while the higher concentrations of UFPs at Everett and Malden occur in colder months in the morning and at night, the highest concentrations of UFPs at Charlestown are occurring in midafternoon in the summer.

To further investigate the high ultrafine particle measurements at Charlestown the polar scatter plots for Charlestown were plotted based on season as well as the wind rose (Figure 18). Most of the easterly winds, which bring about higher ultrafine particle concentrations, occurred in the spring and the summer months.

C. Trends from GMM

The Gaussian Mixing Model was used to cluster the merged air pollution dataset of ultrafine particles, PM2.5 and PM10 into four groups based on ellipsoidal clusters. This clustering algorithm clusters the data into easily recognizable groups. One cluster has low overall air pollution that has a mean UFP concentration of 7,900 /cm^3, 2.8 ug/m^3 PM2.5 and 5.3 ug/m^3 PM10. The second cluster has a very high average UFP concentration of 30,853 /cm^3 but relatively low mean PM2.5 (4.3 ug/m^3) and low mean PM10 (7.7 ug/m^3). The third cluster has

medium number and mass concentration with an average UFP concentration of 13,566/cm^3, 12.5 ug/m^3 of PM2.5 and 21.4 ug/m^3 for PM10. Meanwhile the fourth cluster has medium to low average UFP concentrations of 9,281/cm^3, and very high PM2.5 and PM10 which are 53 and 290 ug/m^3, respectively. These four clusters of hourly averages are referred to herein as low pollution, high UFP, medium pollution, and high PM (Table 4).

Plotting each cluster against different meteorological variables in Figure 20 shows weather conditions that may trigger higher particle mass or particle number concentration. For example, relatively lower relative humidity and atmospheric pressure are associated with the high PM cluster which has very high PM2.5 and PM10 and low particle number concentration. The high PM cluster is also associated with a wind direction of 330 (North-Northwest) which is consistent with the previous findings. Meanwhile, the high UFP group has the highest station pressure and relative humidity. It also has more precipitation than any other cluster. When observing the violin plots the starkest difference between the high UFP group and the low pollution group is the wind direction. The highest UFP occurs when the wind is blowing from the East or the Southwest which is consistent among the previous results. The high UFP category also occurs among lower wind speeds than the other 3 categories. Lastly, the medium pollution cluster which has relatively medium PM2.5, PM10, and ultrafine particles has a similar wind direction to the no pollution group. The largest difference between these two groups is that the medium pollution group has a lower wind speed and a lower relative humidity than the no pollution group. In summary, the Gaussian Mixing model revealed that high amounts of ultrafine particles are associated with high pressure, relative humidity and precipitation while higher mass concentrations are associated with lower relative humidity and pressure.

CHAPTER 4: DISCUSSION

A. Dispersion and Transport

The results of the field campaign suggest that ultrafine particulate matter sources are more localized while larger particulate matter mass concentrations may derive from regional sources (or source areas) as the PM2.5 and PM10 concentrations at all three sites were associated with northwesterly winds. This explanation is supported by the literature (e.g., Seinfeld and Pandis, 2016; Hankley and Marshall, 2015) which has demonstrated that ultrafine particles may disperse more rapidly than larger particles as they do not spread as far since trends differ across the region. Seinfeld and Pandis indicate that in an urban area particle number concentration can increase by ten times within one hundred meters of a major road compared to an average urban concentration. In contrast, mass concentration next to a major road is only 10-20% of the urban background. This is because gas to particle conversion occurs preferably on smaller particles in the aitken and nucleation modes (10-100 nm), due to the availability of space on the particles for gas to particle conversions to occur. Mass distributions are dominated by coarse mode particles while particle number concentration is dominated by ultrafine particles. Similarly, Hankley and Marshall estimated that the share of on-bicycle exposure attributable to near traffic, as opposed to regional pollution, is 50% for particle number concentration while only 25% for PM2.5, in Minneapolis, MN. Given that ultrafine particles disperse quicker it makes sense that there are regional trends in PM2.5 and PM10 and local ultrafine particle trends from this study.

Jesus et al. (2019) examined ultrafine particles and PM2.5 around the world and found little correlation between PM2.5 and ultrafine particles. This is consistent with the results reported herein through spatial analysis pointing to different sources of PM2.5 and ultrafine particles. Jesus et al. (2019) also report a more distinct diurnal trend in particle number

concentration than PM2.5, which is consistent with my findings in Figures 9 through 13. Seeing less distinct diurnal trends in PM2.5 compared to ultrafine particles further supports the idea that ultrafine particles disperse faster and are more variable throughout the day. It also supports the idea that ultrafine particles primarily come from combustion emissions including traffic as the diurnal trends are consistent with rush hour traffic patterns.

B. Hypothesized Source Attribution

Higher concentrations of particulate air pollution occurred bimodally throughout the average day with the highest concentrations around the peak rush hour traffic hours. The only variance from this trend was at the Charlestown site, which experienced a higher particle number concentration than the other two sites and peak times in mid-afternoon in the summer months as opposed to cooler months. This may disclose a possible local source of ultrafine particles east of the Charlestown site. The list of possible sources includes Boston Logan International Airport. Figure 21 shows the Charlestown polar scatter plot overlaid on a map showing the major facilities near Boston that require a permit to discharge emissions into ambient air. Wind coming from the direction of Logan Airport has the highest concentrations of ultrafine particles. Air traffic contributing heavily to the ultrafine particle concentrations at Charlestown could explain why this site exhibits different temporal trends, and a significantly higher annual particle number concentrations than the other two sites, which likely have a large portion of ultrafine particles sourced from traffic. While it can be assumed that rush-hour traffic patterns continue consistently throughout the year, flight patterns do not. July and August of 2024 saw roughly 38,000 total flights each month. Meanwhile January and February of 2024 saw only 28,000 flights, a 36% difference (source: Massport.com). In addition, the easterly wind patterns occur more often in spring and summer (Figure 18), which explain why the Charlestown UFP concentrations are

higher in the summer while the other two sites and previous studies see higher UFPs throughout the winter months. Additionally, the higher particle number concentrations during the day as opposed to rush hour contribute to this theory that a major source of the ultrafine particles is the Boston Logan International Airport. While rush-hour is consistently between 4:00 PM and 7:00 PM on weekdays, airports are consistently busy between 5:30 AM and 10:00 PM (Knabe and Schultz, 2016). While the ultrafine particles are likely primarily coming from traffic at the Malden and Everett sites, the Charlestown site has a significant portion of the ultrafine particles from the nearby airport. This source may be hyperlocal as the Malden site and the Everett site location 8 km 6 km from the airport respectively, do not see pollution from this source to the same extent. Rather than the Logan Airport, the local emitter at the Malden site is likely southwest of the site, and the New England Coffee Roaster is a possible contributor to the high ultrafine particle concentrations in that area as it is directly southwest of the site and has been observed to generate visible plumes of emissions.

It is interesting that the spatial polar scatter plots of PM2.5 show that the highest concentrations are occurring primarily when winds are from the north-northwest direction. Considering PM2.5 comes primarily from urban sources, originally PM2.5 was predicted to be associated with southern winds. However, it is possible that PM2.5 is higher further from the coast where PM2.5 is not able to mix and disperse over the sea. A study in Gdynia, a coastal city in Poland, reveals that while fuel burning contributes 83.7% of PM2.5, agriculture contributes 9.1% of PM2.5 and seawater only contributes 4.6% of PM2.5 (Gorka et al., 2025). Horizontal air mixing with cleaner air above the Atlantic Ocean may contribute to lower PM2.5 concentrations coming from Boston compared to northwest of the city. In addition to traffic, some particulate matter in the form of PM2.5 and PM10 can originate from coal-fired power plants. Although

Massachusetts does not have any coal power plants in the state, there are two power plants in New Hampshire and many more in Pennsylvania and West Virginia polluting the air (Out of Control: The Deadly Impact of Coal Pollution). It is possible that the PM2.5 experienced in the Boston area primarily comes from fuel burning activities such as traffic emissions and coal fired power plants in other states.

C. Meteorology: Rain, Pressure, Humidity

The Gaussian Mixing Model revealed that ultrafine particles are more likely to occur in high pressure and high humidity and can occur in precipitation events while high PM2.5 and PM10 occur when there is lower humidity, pressure, and very rarely in a precipitation event. A study in Bachok has similar findings contributing to high particle number concentrations which found that relative humidity, pressure, and wind direction are associated with an increase in particles. This study notes that meteorology is a driving factor to the concentration of air particles (Rahim et al., 2021). It is possible that the higher-pressure system is associated with a greater concentration of ultrafine particles because the air flow is traveling down from higher elevations so ultrafine particles from smokestacks and airplanes have greater impacts. Likewise, it is possible that a lower-pressure system is associated with higher PM2.5 and PM10 because air coming from the higher-pressure regions may blow PM2.5 and PM10 particles more regionally. These larger particles are less likely to disperse due to Brownian physics than smaller particles and some ultrafine particles may coagulate locally to grow larger into fine particles in a low-pressure system.

A study in northern Thailand agrees with these results on the inverse relationship between relative humidity and PM2.5 and PM10. In this study relative humidity is significantly related to a decrease of PM2.5 and PM10 ($r = -0.722$) (Sirithian and Thanatrakolski, 2022).

Seinfeld and Pandis (2016) indicate that as relative humidity increases, particles remain solid until the relative humidity reaches a threshold value called deliquescence relative humidity where the particle becomes a saturated aqueous solution. Meanwhile nucleation scavenging of aerosols in clouds refers to growth of aerosols within clouds. It is possible that ultrafine particles are more susceptible to nucleating in clouds (Figure 8.10) while larger particles (PM2.5 and PM10) have the tendency to become a saturated aqueous solution at the deliquescence relative humidity. This would result in larger particles such as PM2.5 and PM10 being associated with lower relative humidity while high concentrations of ultrafine particles are associated with high relative humidity.

As for the precipitation phenomenon, according to Seinfeld and Pandis (2016) the scavenging coefficient, describing the rate of removal of particles by rain, decreases for particles between 0.01 micrometers in diameter to 10 micrometers in a U shape. Particles between 0.1 and 1 micron have the lowest removal efficiency rate. This characteristic minimum is when particles are too large for an appreciable Brownian diffusivity and too small to be collected by impaction or interception. While PM2.5 and PM10 in this study are large enough to be impacted, particle number concentration collected in this study on the CPC 3783 accounts for primarily diameters 0.007 to 0.11 micrometers (Figure 8.10 Atmospheric Chemistry and Physics). These ultrafine particles may be too small to be impacted and removed by rainfall while high PM2.5 and PM10 will be impacted and removed by rain, leading to lower amounts of fine particles during a rainstorm, but a possibility of higher amounts of ultrafine particles. The phenomenon of ultrafine particle concentrations increasing as rain hits the ground and disturbs previously settled ultrafine particles is called petrichor and is associated with the smell of rainfall.

D. Comparing Results

The world cities study also found that the average annual concentrations ranged from 8,000 to 19,500 /cm^3 for ultrafine particles and 7.0 to 65.8 ug/m^3 for PM2.5. The average annual particle number concentrations were in this range at all three of the sites. Meanwhile, the average PM2.5 concentrations were lower than this range, except for at the Malden site.

As World Health Organization guidelines recommend keeping the hourly ultrafine particle concentrations below 20,000 /cm^3, this was exceeded during the study duration. Roughly 25% of the Charlestown site data exceeded this parameter as did some of the hourly data at Everett and Malden. The World Health Organization also recommended limiting the annual mean PM2.5 exposure to 5 ug/m^3 which is exceeded at all three sites. Lastly the same organization recommends the annual average PM10 exposure to be limited to 15 ug/m^3 which has only been exceeded at the Malden site. Overall, the communities at Everett, Malden, and Charlestown should put in place additional measures to reduce air pollution to keep residents safe.

E. Limitations

A limitation of this study is in the low budget air quality sensors. Unfortunately, the Dylos 1100, which read counts of small and large particles, did not produce any interesting results. There are limits to instrument quality as the Dylos 1100 was made with inexpensive technology. This instrument did not come with detailed data post-processing instructions further limiting its accuracy. Also, compared with the CPC 3783, which counted to 140,000 /cm^3 at Charlestown, the Dylos 1100 only counts to 20 /cm^3 representing only 0.01% of the particle number concentration. This may miss a large portion of the story which is detected using the CPC 3783. In turn, the Dylos 1100 devices may not be too helpful for communitywide air pollution monitoring efforts.

F. Future Work

If allotted more time on this project, it would be interesting to research and test more expectation-maximum methods using Machine Learning to further investigate meteorological trends in air pollution. Using the GMM method would be interesting with a fast mobility particle sizer (FMPS) to further explore the relationship between meteorology and particle size. Furthermore, in terms of understanding source attribution it would be beneficial to study the chemical composition of the particles collected at Malden, Everett and Charlestown to further test the hypothesis that ultrafine air pollution comes from the traffic and from the airport at Charlestown. It would also be interesting to work on this project longer to work with the community advisory board to figure out the best ways to communicate the study results with the public and spread this information effectively to communities. This project could be expanded to set up a framework to help communities set up and execute their own air pollution studies as much of the air pollution concerns are hyper localized.

CHAPTER 5: CONCLUSION

Understanding the sources of particulate matter air pollution as well as which meteorological conditions are associated with higher concentrations of pollution is key to protecting community health. This study is novel because it is focused on ultrafine particles in communities that have never been studied and uses novel approaches such as the Gaussian Mixture Model. Because air pollution is localized it is important to conduct studies at local levels to empower residents with data they can use to advocate for change. Graphics that were made for this report are also shared with a local community advisory board in the form of a presentation. Later, they will be presented in a report card and report out session. Additionally, testing a new approach to understanding air pollution: the Gaussian Mixture Model, helps the broader scientific community with a new tool to understand the relationship and weather conditions relating to higher mass concentration and ultrafine particles. Lastly, testing the low-cost Dylos devices, advances the goal of understanding which low-cost air pollution monitoring equipment is effective for communities to launch their own grass-roots campaigns to collect air pollution data for advocacy and community awareness.

Appendix

A. Tables and Figures

Table 2. Overall mean concentrations over the study duration.

Mean Concentrations

	Malden	Everett	Charlestown
UFP	10,700	11,800	16,400
PM2.5	8.97	5.91	5.44
PM10	30.7	6.78	10.2
Small	5.99	5.16	4.89
Large	0.368	0.299	0.302

Table 3. Overall median concentrations over the study duration.

Median Concentrations

	Malden	Everett	Charlestown
UFP	7,760	9,360	10,800
PM2.5	3.64	3.60	3.23
PM10	6.96	6.78	6.44
Small	3.53	3.24	3.23
Large	0.182	0.191	0.199

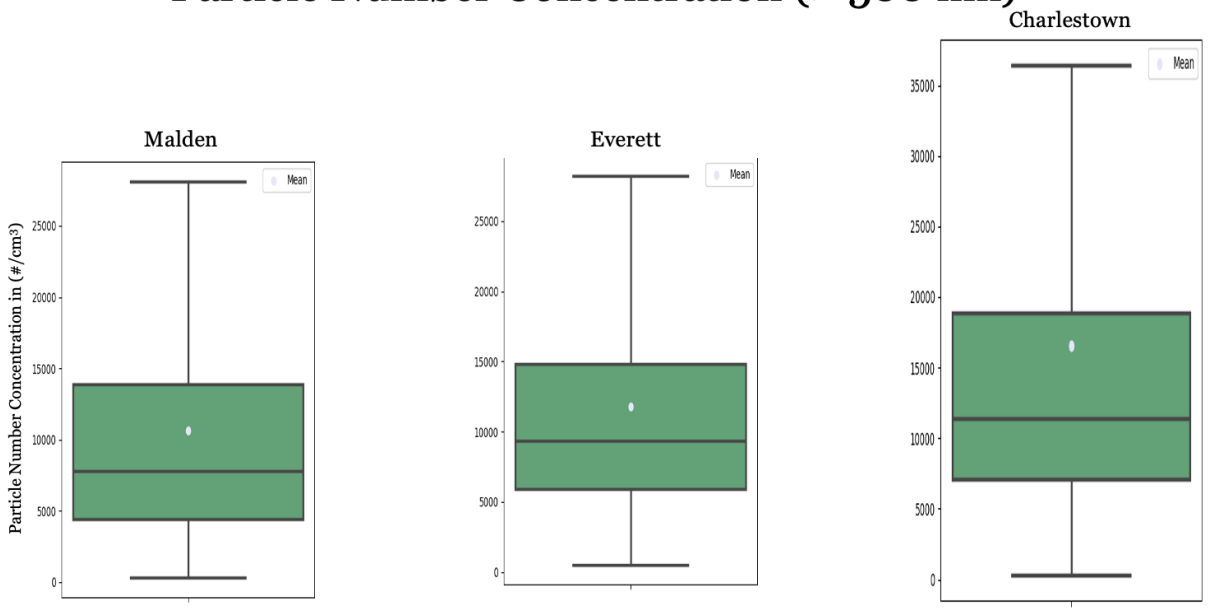

Figure 4. Particle number concentration boxplots of hourly average ultrafine particle concentration (#/cm^3). Measurements were made between October 26, 2024, and January 31, 2025. Each plot shows the 25th, 50th, and 75th quartiles of the data; the whiskers represent 1.5 times the interquartile range. The average is shown with a circular symbol.

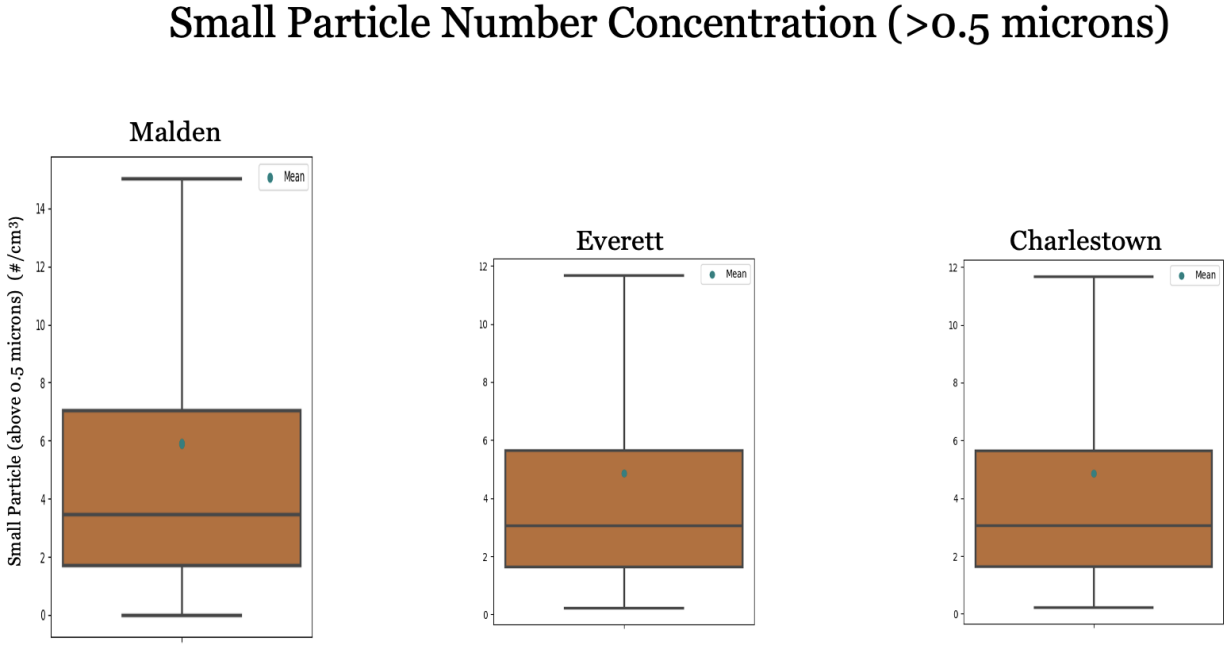

Figure 5. Small particles number boxplots of hourly average concentration of particle above 0.5 microns (#/cm^3), Measurements were made between October 26, 2024, and January 31, 2025. Each plot shows the 25th, 50th, and 75th quartiles of the data; the whiskers represent 1.5 times the interquartile range. The average is shown with a circular symbol.

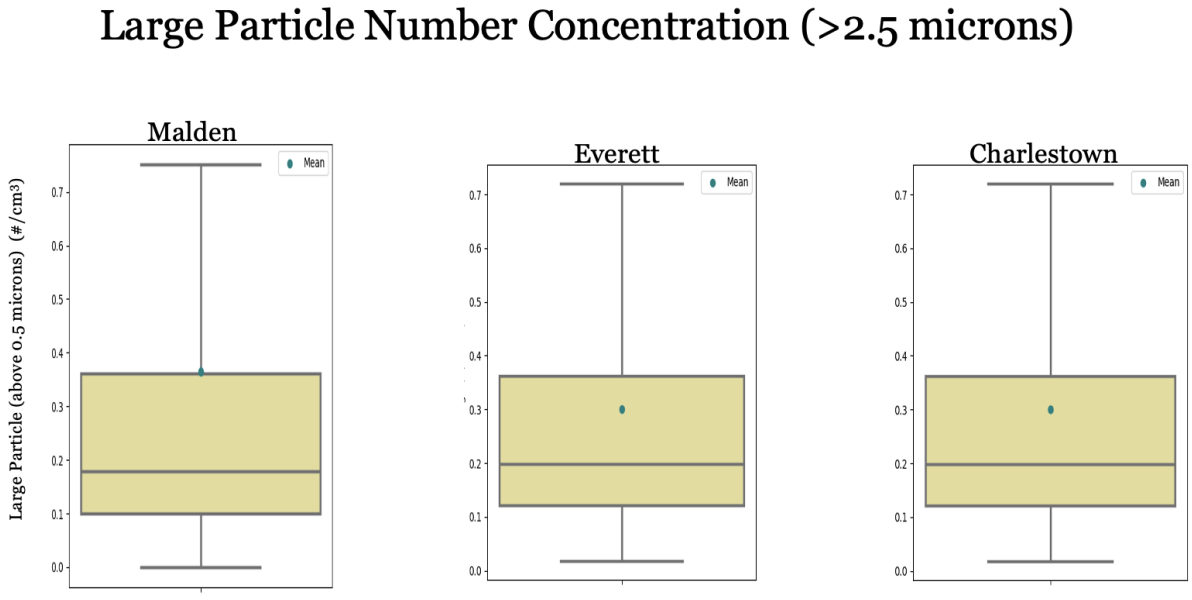

Figure 6. Boxplots of hourly average concentration of particle above 2.5 microns (#/cm^3), Measurements were made between October 26, 2024, and January 31, 2025. Each plot shows the 25th, 50th, and 75th quartiles of the data; the whiskers represent 1.5 times the interquartile range. The average is shown with a circular symbol.

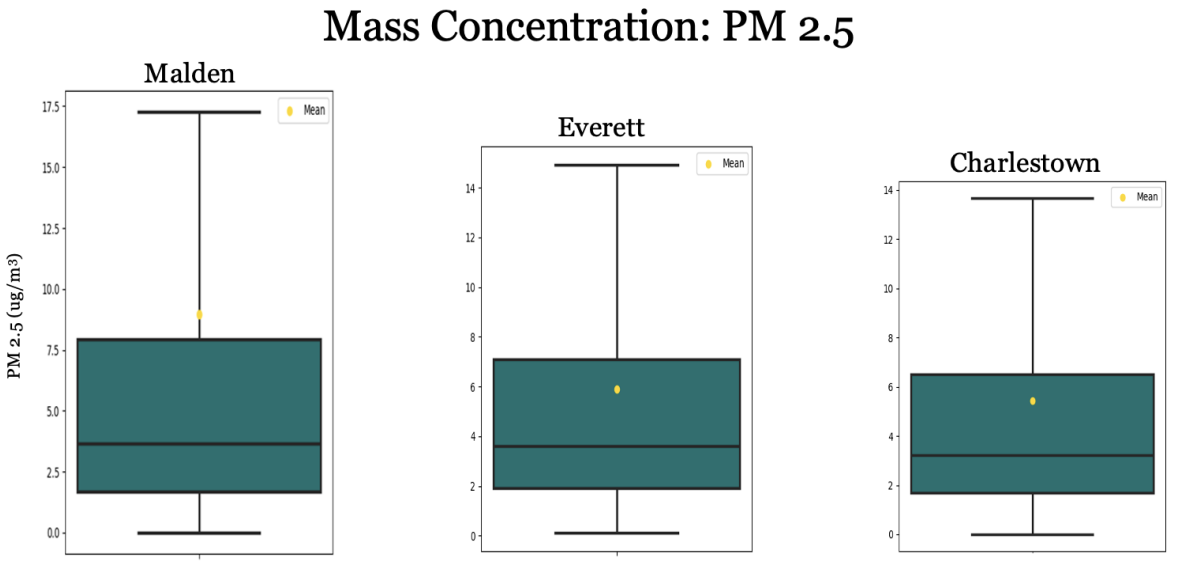

Figure 7. Boxplots of hourly average PM2.5 (ug/m^3). Measurements were made between October 26, 2024, and January 31, 2025. Each plot shows the 25th, 50th, and 75th quartiles of the data; the whiskers represent 1.5 times the interquartile range. The average is shown with a circular symbol.

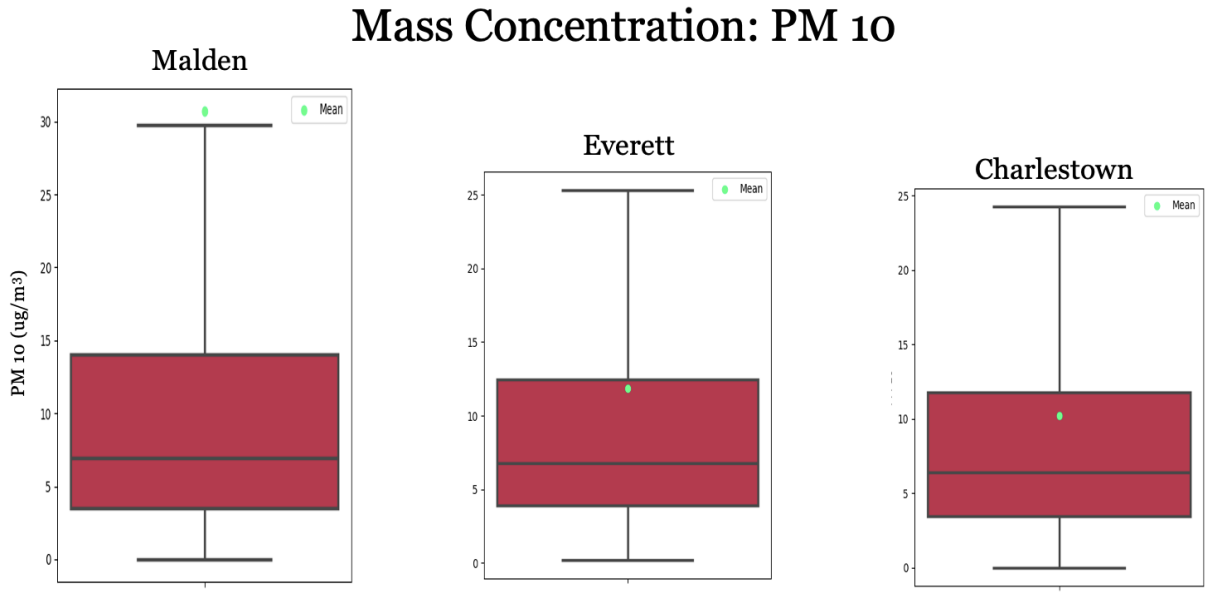

Figure 8. Boxplots of hourly average PM10 (ug/m^3). Measurements were made between October 26, 2024, and January 31, 2025. Each plot shows the 25th, 50th, and 75th quartiles of the data; the whiskers represent 1.5 times the interquartile range. The average is shown with a circular symbol.

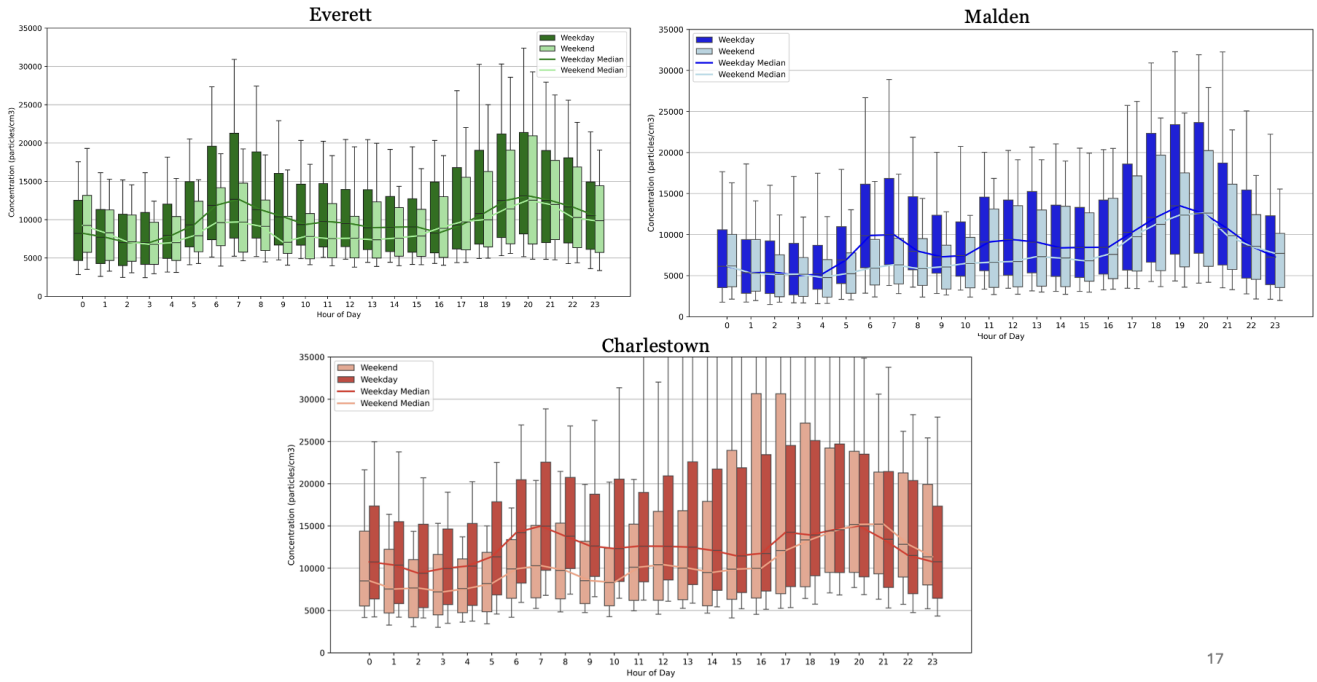

Figure 9. Diurnal Ultrafine Particle Trends at Everett, Malden and Charlestown.

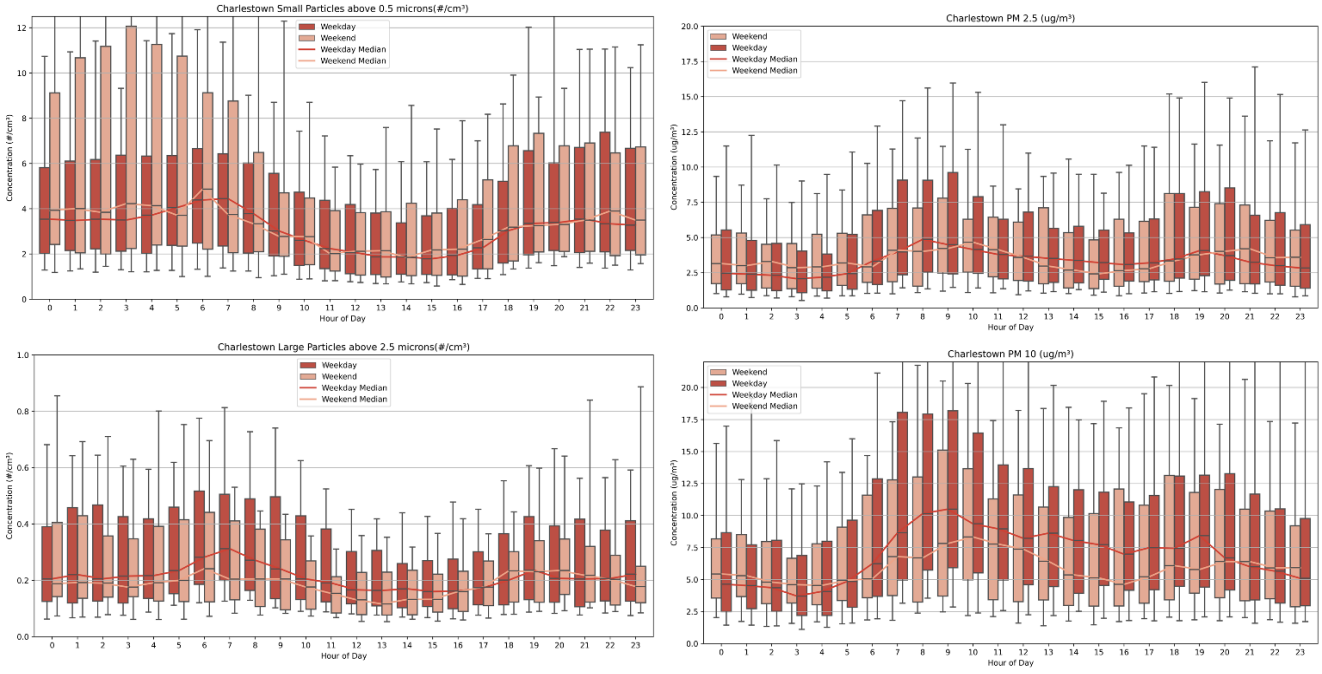

Figure 10. Diurnal trends Small Particles, Large Particles, PM2.5 and PM10 at Charlestown.

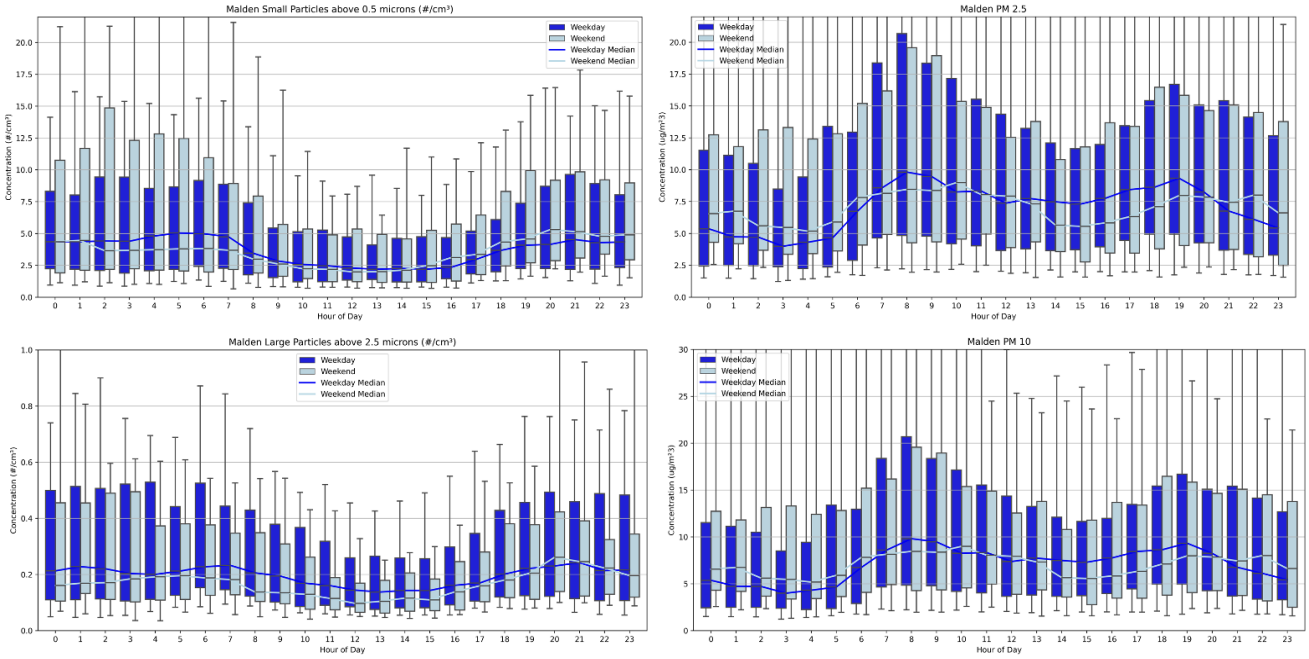

Figure 11. Diurnal trends in Small Particles, Large Particles, PM2.5 and PM10 at Malden.

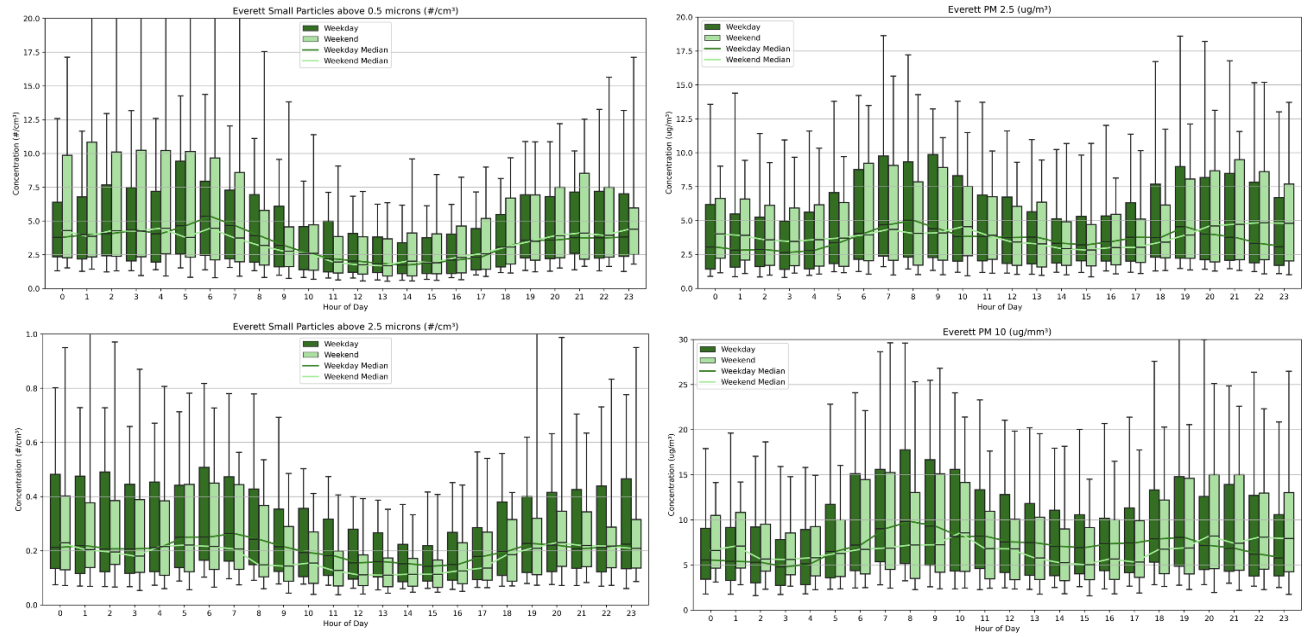

Figure 12. Diurnal trends in Small Particles, Large Particles, PM2.5 and PM10 at Everett.

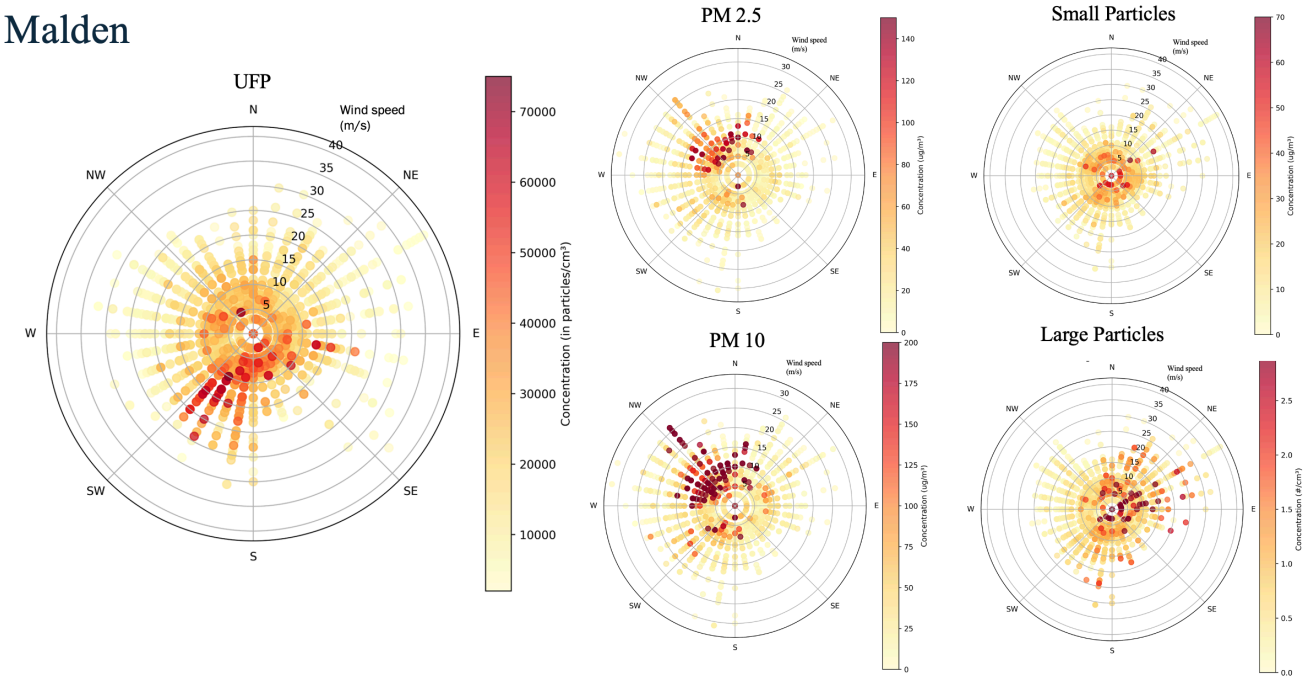

Figure 13. Hourly average polar plots from the Malden site (350 Main Street) showing 5 pollutant sizes: PM10, PM2.5, small particle concentration (>0.5 microns diameter), large particle concentration (>2.5 microns diameter), and ultrafine particles. Each point represents an hourly average of pollutant concentration plotted based on the direction the wind was blowing from and the speed of the wind. Higher concentrations of the pollutants overlay lower concentrations at the same wind speed and wind direction. The dark points represent higher concentrations, and lighter points represent lower concentrations.

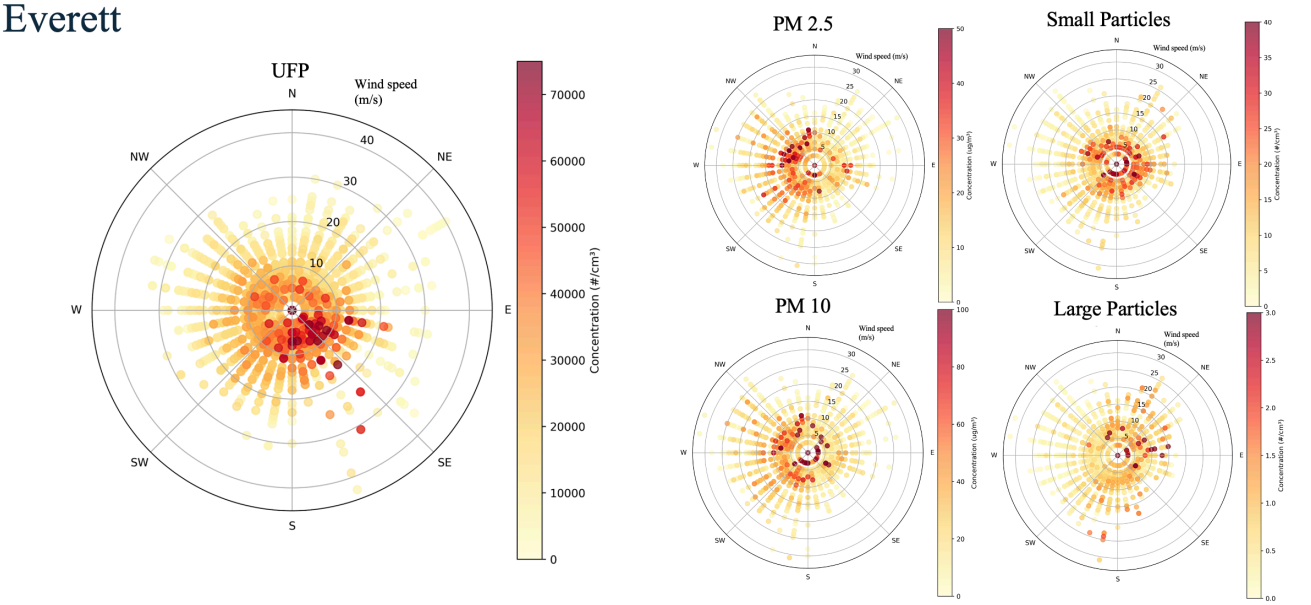

Figure 14. Hourly average polar plots from the Everett site (484 Broadway) showing 5 pollutant sizes: PM10, PM2.5, Small Particle Concentration (>0.5 microns diameter), Large Particle Concentration (>2.5 microns diameter), and Ultrafine particles. Each point represents an hourly average of pollutant concentration plotted based on the direction the wind was blowing from

and the speed of the wind. Higher concentrations of the pollutants overlay lower concentrations at the same wind speed and wind direction. The dark points represent higher concentrations, and lighter points represent lower concentrations.

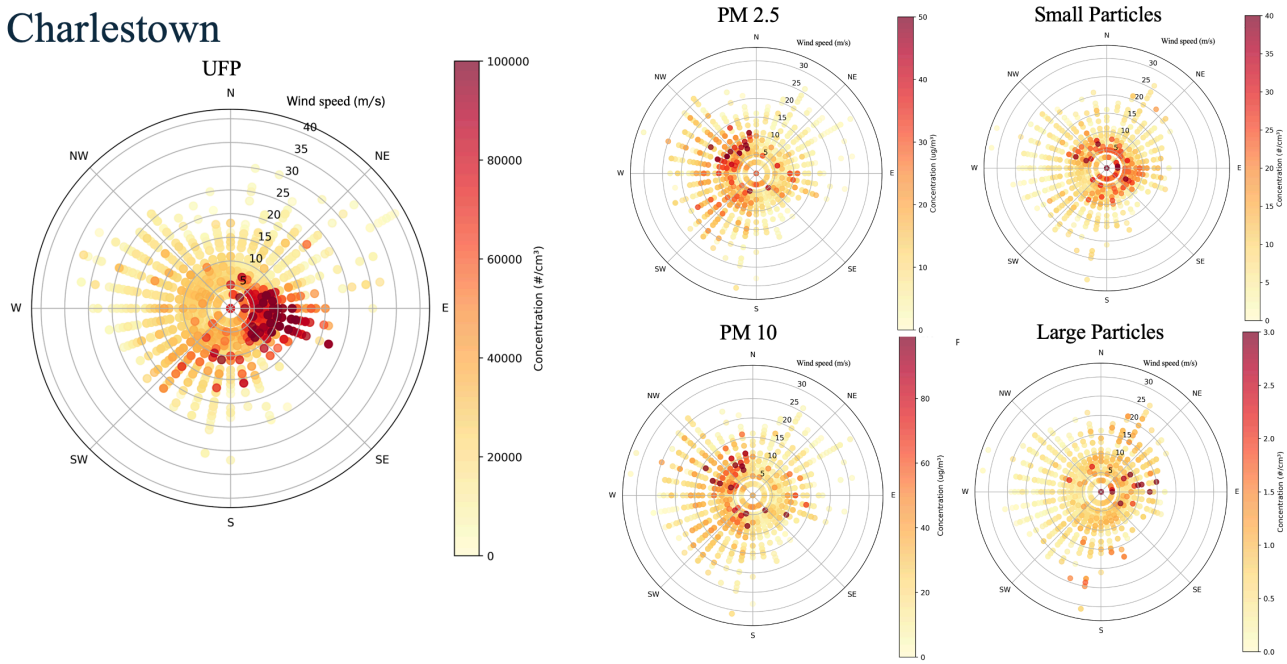

Figure 15. Hourly average polar plots from the Charlestown site (43 Monument Sq.) showing 5 pollutant sizes: PM10, PM2.5, Small Particle Concentration (>0.5 microns diameter), Large Particle Concentration (>2.5 microns diameter), and Ultrafine particles. Each point represents an hourly average of pollutant concentration plotted based on the direction the wind was blowing from and the speed of the wind. Higher concentrations of the pollutants overlay lower concentrations at the same wind speed and wind direction. The dark points represent higher concentrations, and lighter points represent lower concentrations.

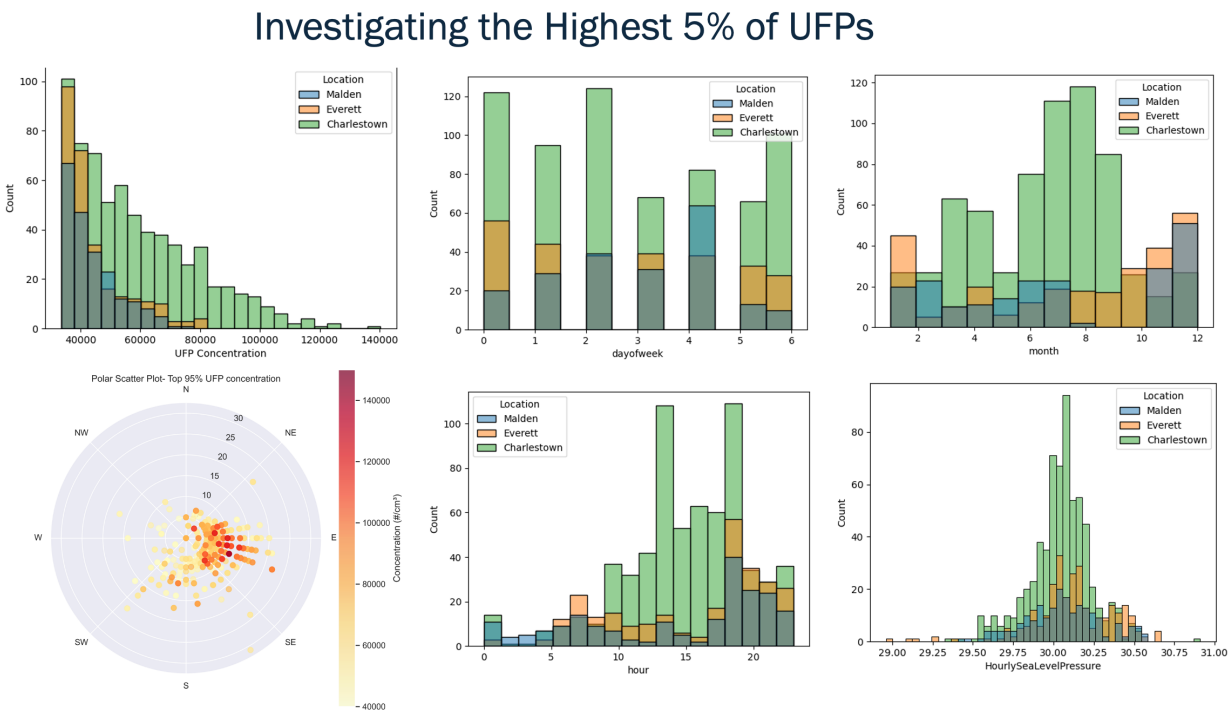

Figure 16. Histograms and polar scatter plots of the highest 95th percentile of hourly UFP concentration. From left to right the plots show hourly UFP concentration, day of the week where 0 is Monday and 6 is Sunday, month of the year, wind direction, hour of the day, and sea level pressure.

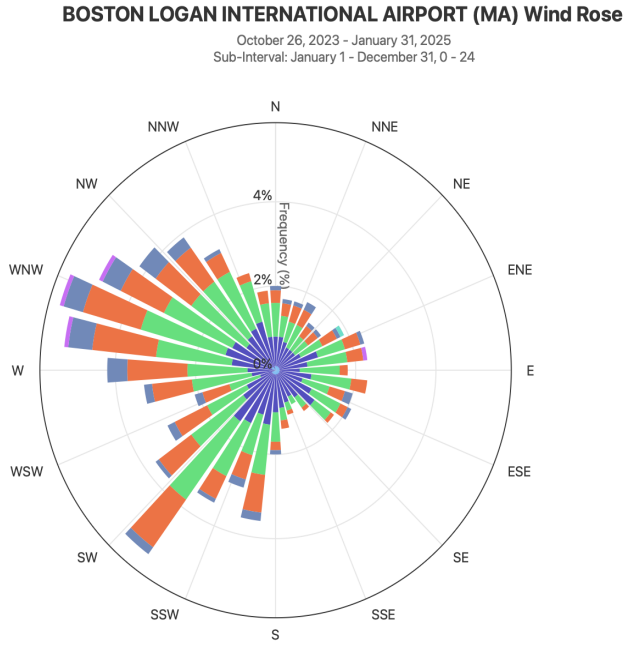

Figure 17. Windrose depicts wind speed and where the wind is blowing from at Boston Logan International Airport during the study duration (October 26, 2023, to January 31, 2025).

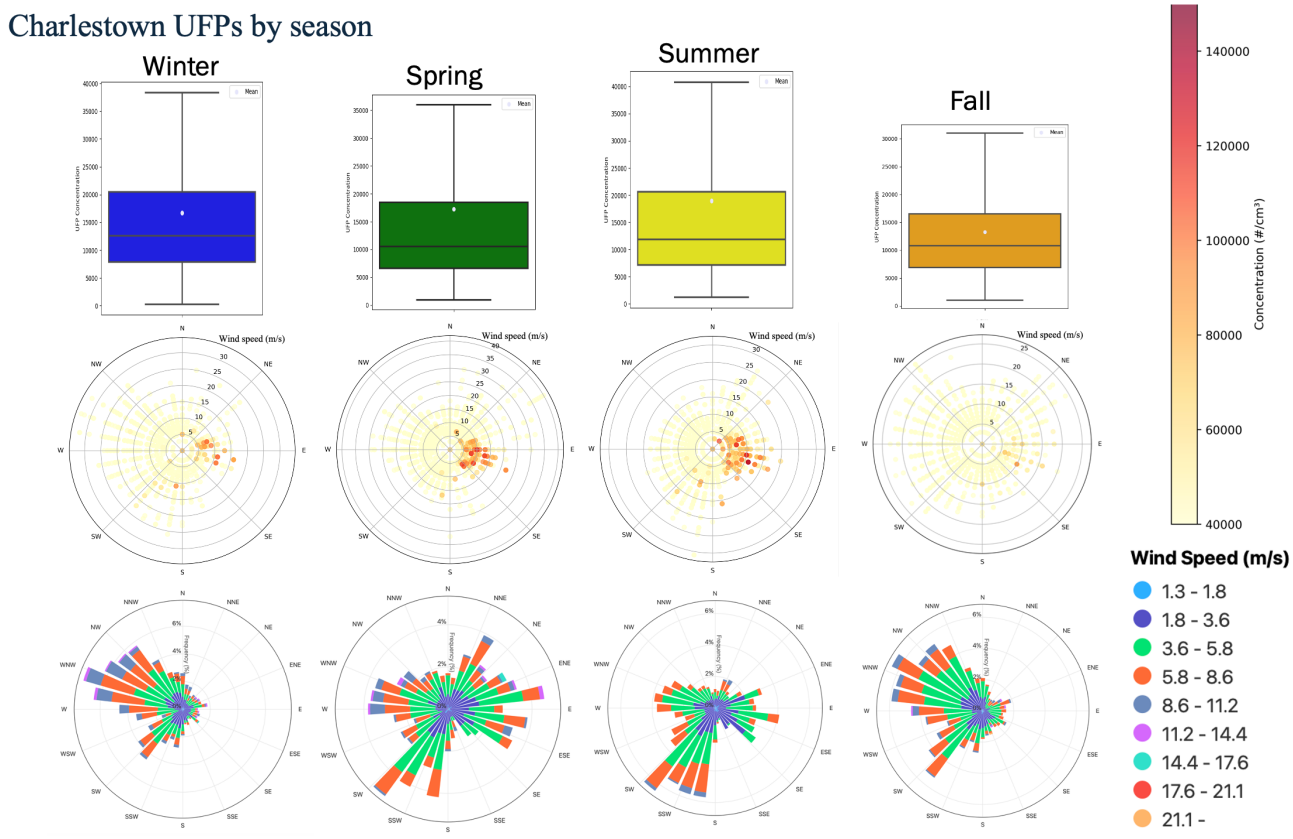

Figure 18. Charlestown ultrafine particle counts by season. The middle row shows the polar scatter plots split by each season, and the bottom row shows the wind rose at Boston Logan Airport based on the season.

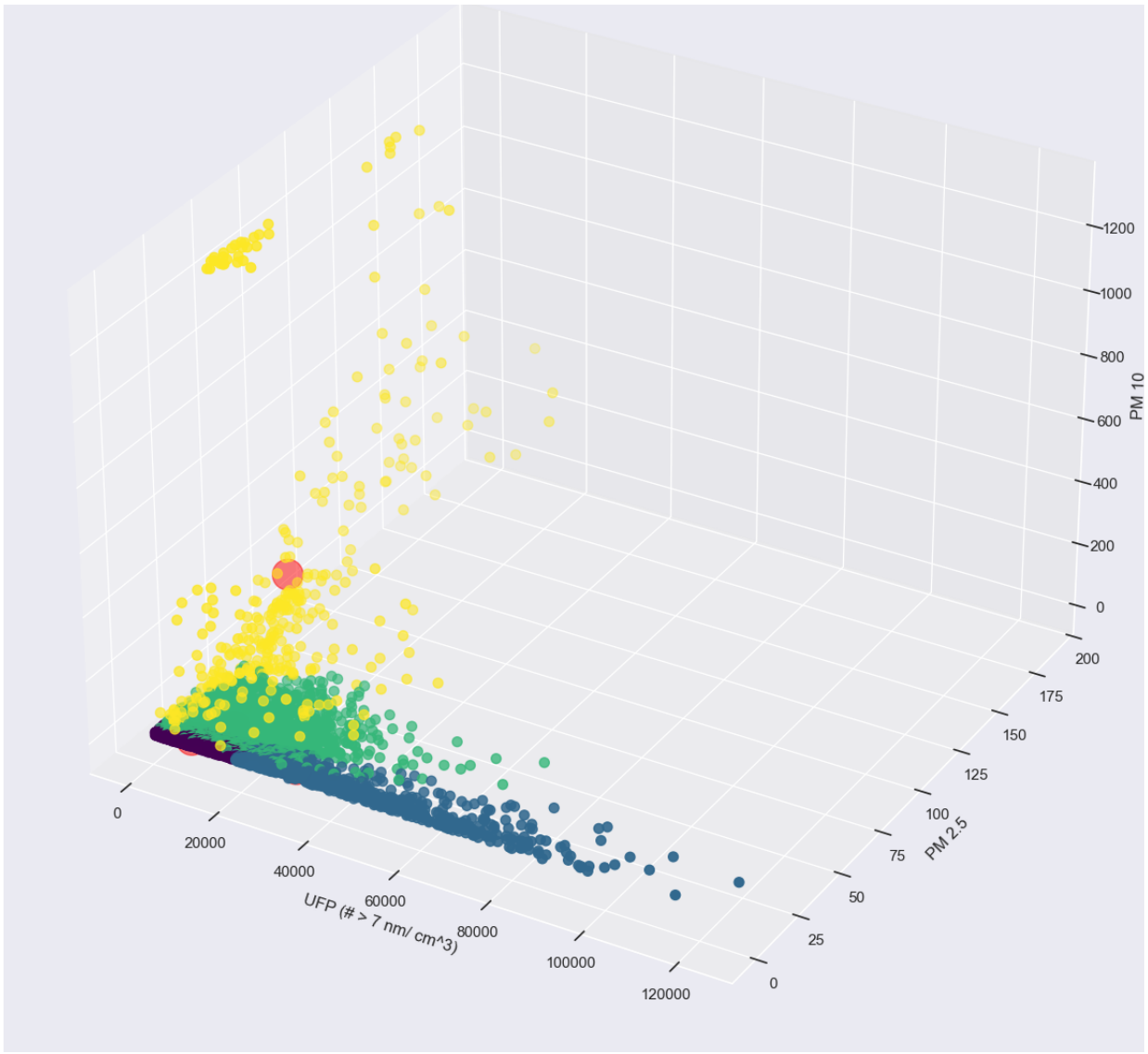

Figure 19. Gaussian Mixing Model results with hourly data from all three sites.

Table 4. Gaussian Mixing Model results. Average concentrations between of each cluster.

Nickname	*Color*	*Cluster*	*UFP (#/cm³)*	*PM2.5 (ug/m³)*	*PM10 (ug/m³)*
Low Pollution	Purple	0	7,863	2.8	5.3
High UFP	Blue	1	30,853	4.3	7.7
Medium Pollution	Green	2	13,566	12.5	21.4
High PM	Yellow	3	9,281	53	290

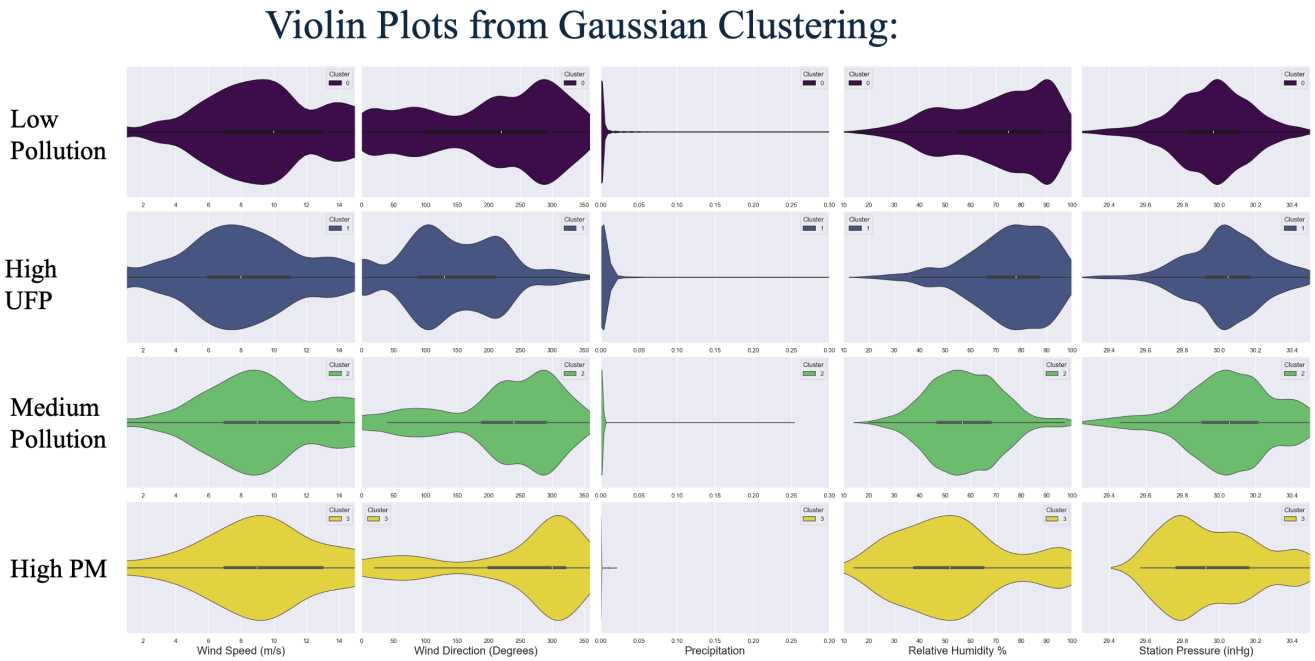

Figure 20. Violin plot representing wind speed, wind direction, precipitation, relative humidity, and station pressure by cluster.

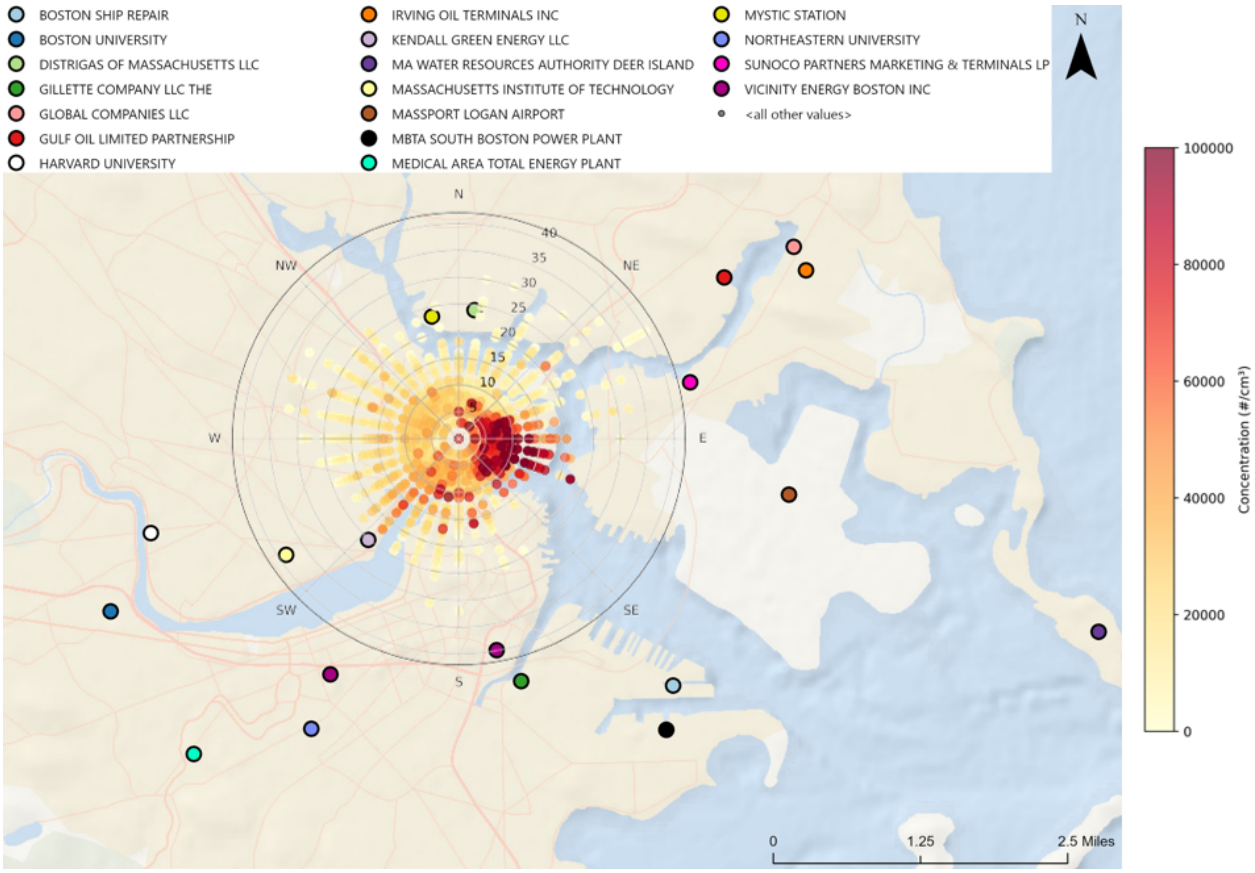

Figure 21. Major facilities that have an air permit and the Charlestown ultrafine particle concentration wind rose is overlaid on top.

Table 5. Typical over saturation adjustment table from Dylos. This was used to derive the equation to adjust for particle mass above 300 ug/m3.

DC 1700- PM2.5 ug/m3	True PM2.5
300	300
350	420
400	560
450	760
500	930
550	1280

B. References

"AIRNOW FIRE AND SMOKE MAP - QUESTIONS AND ANSWERS." AIR NOW, EPA, DOCUMENT.AIRNOW.GOV/AIRNOW-FIRE-AND-SMOKE-MAP-QUESTIONS-AND-ANSWERS.PDF. ACCESSED 27 AUG. 2024.

"AIRPORT STATISTICS." MASSPORT, MASSPORT, www.massport.com/logan-airport/about LOGAN/AIRPORT-STATISTICS. ACCESSED 17 APR. 2025.

ALBERTS, B. (1970, JANUARY 1). BLOOD VESSELS AND ENDOTHELIAL CELLS. MOLECULAR BIOLOGY OF THE CELL. 4TH EDITION. HTTPS://WWW.NCBI.NLM.NIH.GOV/BOOKS/NBK26848/#:~:TEXT=ENDOTHELIAL%20CELLS%20FORM%20A%20SINGLE,OF%20THE%20BLOOD%2DVESSEL%20WALL.

ARTER, C. A., BUONOCORE, J. J., MONIRUZZAMAN, C., YANG, D., HUANG, J., & ARUNACHALAM, S. (2022). AIR QUALITY AND HEALTH-RELATED IMPACTS OF TRADITIONAL AND ALTERNATE JET FUELS FROM AIRPORT AIRCRAFT OPERATIONS IN THE U.S. ENVIRONMENT INTERNATIONAL, 158, 106958. HTTPS://DOI.ORG/10.1016/J.ENVINT.2021.106958

AUSTIN, E., XIANG, J., GOULD, T. R., SHIRAI, J. H., YUN, S., YOST, M. G., LARSON, T. V., & SETO, E. (2021). DISTINCT ULTRAFINE PARTICLE PROFILES ASSOCIATED WITH AIRCRAFT AND ROADWAY TRAFFIC. ENVIRONMENTAL SCIENCE AND TECHNOLOGY, 55(5), 2847–2858. https://doi.org/10.1021/acs.est.0c05933

BEDDOWS, D.C.S., DALL'OSTO, M., HARRISON, R.M., 2009. CLUSTER ANALYSIS OF RURAL, URBAN, AND CURBSIDE ATMOSPHERIC PARTICLE SIZE DATA. ENVIRON. SCI. TECHNOL. 43 (13), 4694–4700. HTTPS://DOI.ORG/10.1021/ES803121T.

BERGMANN, MARIE L., ET AL. "SPATIAL AND TEMPORAL VARIATION OF FAÇADE-LEVEL PARTICLE NUMBER CONCENTRATIONS USING PORTABLE MONITORS IN COPENHAGEN, DENMARK." *ENVIRONMENTAL POLLUTION (1987)*, VOL. 365, 2025, PP. 125398-, HTTPS://DOI.ORG/10.1016/J.ENVPOL.2024.125398.

BLANCO, M. N., GASSETT, A., GOULD, T., DOUBLEDAY, A., SLAGER, D. L., AUSTIN, E., SETO, E., LARSON, T. V., MARSHALL, J. D., & SHEPPARD, L. (2022). CHARACTERIZATION OF ANNUAL AVERAGE TRAFFIC-RELATED AIR POLLUTION CONCENTRATIONS IN THE GREATER SEATTLE AREA FROM A YEAR-LONG MOBILE MONITORING CAMPAIGN. ENVIRONMENTAL SCIENCE AND TECHNOLOGY. HTTPS://DOI.ORG/10.1021/ACS.EST.2C01077

CHUNG, C. S., LANE, K. J., BLACK-INGERSOLL, F., KOLACZYK, E., SCHOLLAERT, C., LI, S., SIMON, M. C., & LEVY, J. I. (2023). ASSESSING THE IMPACT OF AIRCRAFT ARRIVAL ON AMBIENT ULTRAFINE PARTICLE NUMBER CONCENTRATIONS IN NEAR-AIRPORT COMMUNITIES IN BOSTON, MASSACHUSETTS. ENVIRONMENTAL RESEARCH, 225. HTTPS://DOI.ORG/10.1016/J.ENVRES.2023.115584

COOPER, C. D., & ALLEY, F. C. (2015). AIR POLLUTION CONTROL: A DESIGN APPROACH. SCIENTIFIC INTERNATIONAL.

De Jesus, Alma Lorelei, et al. "Ultrafine particles and PM2.5 in the air of cities around the world: Are they representative of each other?" Environment International, vol. 129, Aug. 2019, pp. 118–135, https://doi.org/10.1016/j.envint.2019.05.021.

EPA, Environmental Protection Agency, www.epa.gov/smartgrowth/smart-location-mapping. Accessed 13 Dec. 2024.

Fuller, C. H., Brugge, D., Williams, P. L., Mittleman, M. A., Durant, J. L., & Spengler, J. D. (2012). Estimation of ultrafine particle concentrations at near-highway residences using data from local and central monitors. Atmospheric Environment, 57, 257–265. https://doi.org/10.1016/j.atmosenv.2012.04.004

GÓRKA, Maciej, et al. "Chemical and isotopic genesis of fine and ultra-fine aerosols in the coastal urban atmosphere." *Science of The Total Environment*, vol. 967, Mar. 2025, p. 178823, https://doi.org/10.1016/j.scitotenv.2025.178823.

Habre, R., Zhou, H., Eckel, S. P., Enebish, T., Fruin, S., Bastain, T., Rappaport, E., & Gilliland, F. (2018). Short-term effects of airport-associated ultrafine particle exposure on lung function and inflammation in adults with asthma. Environment International, 118, 48–59. https://doi.org/10.1016/J.ENVINT.2018.05.031

Hankey, Steve, and Julian D. Marshall. "On-Bicycle Exposure to Particulate Air Pollution: Particle Number, Black Carbon, PM2.5, and Particle Size." *Atmospheric Environment (1994)*, vol. 122, 2015, pp. 65–73, https://doi.org/10.1016/j.atmosenv.2015.09.025.

Herron, Patrick. "Community-Led Improvement of Air Quality and Health in the Lower Mystic (CLEANAIR)." 2022.

Jung, C. R., Chen, W. T., Young, L. H., & Hsiao, T. C. (2023). A hybrid model for estimating the number concentration of ultrafine particles based on machine learning algorithms in central Taiwan. Environment International, 175. https://doi.org/10.1016/j.envint.2023.107937

Karumanchi, Shilpa, et al. "Spatial and Temporal Variability of Airborne Ultrafine Particles in the Greater Montreal Area: Results of Monitoring Campaigns in Two Seasons." *The Science of the Total Environment*, vol. 771, no. 4, 2021, pp. 144652–144652, https://doi.org/10.1016/j.scitotenv.2020.144652.

Kerckhoffs, J., Hoek, G., Gehring, U., & Vermeulen, R. (2021). Modeling nationwide spatial variation of ultrafine particles based on mobile monitoring. Environment International, 154, 106569. https://doi.org/10.1016/J.ENVINT.2021.106569

Kerckhoffs, J., Hoek, G., Portengen, L., Brunekreef, B., & Vermeulen, R. C. H. (2019). Performance of Prediction Algorithms for Modeling Outdoor Air Pollution Spatial Surfaces. Environmental Science and Technology, 53(3), 1413–1421. https://doi.org/10.1021/acs.est.8b06038

KNABE, FRANZ, AND MICHAEL SCHULTZ. "A NEW WAY TO INDICATE AIRPORT AIRSIDE PERFORMANCE FROM AN ECONOMIC PERSPECTIVE." TRANSPORTATION RESEARCH PROCEDIA, VOL. 14, 2016, PP. 3771–3780, HTTPS://DOI.ORG/10.1016/J.TRPRO.2016.05.462.

LEE, YONGHWAN, ET AL. "CLUSTER ANALYSIS OF ATMOSPHERIC PARTICLE NUMBER SIZE DISTRIBUTIONS AT A RURAL SITE DOWNWIND OF SEOUL, KOREA." ATMOSPHERIC POLLUTION RESEARCH, VOL. 12, NO. 6, JUNE 2021, P. 101086, HTTPS://DOI.ORG/10.1016/J.APR.2021.101086.

LIFESCI. (N.D.). MONONUCLEAR CELLS. BIOCOMPARE. HTTPS://WWW.BIOCOMPARE.COM/PFU/111694/SOIDS/2263418/CELLS_AND_MICROORGANISMS/MONONUCLEAR_CELLS

LIN, C., LANE, K. J., GRIFFITHS, J. K., & BRUGGE, D. (2022). A NEW EXPOSURE METRIC FOR THE CUMULATIVE EFFECT OF SHORT-TERM EXPOSURE PEAKS OF TRAFFIC-RELATED ULTRAFINE PARTICLES. JOURNAL OF EXPOSURE SCIENCE AND ENVIRONMENTAL EPIDEMIOLOGY, 32(4). HTTPS://DOI.ORG/10.1038/S41370-021-00397-3

"U.S. LOCAL CLIMATOLOGICAL DATA (LCD)." *NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION (NCEI)*, NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA), WWW.NCEI.NOAA.GOV/ACCESS/SEARCH/DATA-SEARCH/LOCAL-CLIMATOLOGICAL-DATA?STATIONS=72509014739&PAGENUM=1&STARTDATE=2023-01-01T00%3A00%3A00&ENDDATE=2025-02-01T23%3A59%3A59. ACCESSED 23 APR. 2025.

MOORE, K., KRUDYSZ, M., PAKBIN, P., HUDDA, N., & SIOUTAS, C. (2009). INTRA-COMMUNITY VARIABILITY IN TOTAL PARTICLE NUMBER CONCENTRATIONS IN THE SAN PEDRO HARBOR AREA (LOS ANGELES, CALIFORNIA). AEROSOL SCIENCE AND TECHNOLOGY, 43(6), 587–603. HTTPS://DOI.ORG/10.1080/02786820902800900

"MASSGIS (BUREAU OF GEOGRAPHIC INFORMATION)." MASS.GOV, WWW.MASS.GOV/ORGS/MASSGIS-BUREAU-OF-GEOGRAPHIC-INFORMATION. ACCESSED 13 DEC. 2024.

WASKOM, MICHAEL. "SEABORN: STATISTICAL DATA VISUALIZATION." SEABORN, 2024, SEABORN.PYDATA.ORG/.

OLUWADAIRO, T., WHITEHEAD, L., SYMANSKI, E. ET AL. EFFECTS OF AEROSOL PARTICLE SIZE ON THE MEASUREMENT OF AIRBORNE PM2.5 WITH A LOW-COST PARTICULATE MATTER SENSOR (LCPMS) IN A LABORATORY CHAMBER. ENVIRON MONIT ASSESS 194, 56 (2022). HTTPS://DOI.ORG/10.1007/S10661-021-09715-6

PAKBIN, P., HUDDA, N., CHEUNG, K. L., MOORE, K. F., & SIOUTAS, C. (2010). SPATIAL AND TEMPORAL VARIABILITY OF COARSE (PM10-2.5) PARTICULATE MATTER CONCENTRATIONS IN THE LOS ANGELES AREA. AEROSOL SCIENCE AND TECHNOLOGY, 44(7), 514–525. HTTPS://DOI.ORG/10.1080/02786821003749509

"Pandas.Dataframe.Backfill." Pandas.DataFrame.Backfill - Pandas 2.2.3 Documentation, NumFocus, Inc., 2024, pandas.pydata.org/docs/reference/api/pandas.DataFrame.backfill.html.

Patton, A. P., Perkins, J., Zamore, W., Levy, J. I., Brugge, D., & Durant, J. L. (2014). Spatial and temporal differences in traffic-related air pollution in three urban neighborhoods near an interstate highway. Atmospheric Environment, 99, 309–321. https://doi.org/10.1016/j.atmosenv.2014.09.072

Rahim, Haasyimah Ab, et al. "Coastal Meteorology on the Dispersion of Air Particles at the Bachok GAW Station." *The Science of the Total Environment*, vol. 782, 2021, pp. 146783–146783, https://doi.org/10.1016/j.scitotenv.2021.146783.

Sabaliauskas, Kelly, et al. "Cluster analysis of roadside ultrafine particle size distributions." Atmospheric Environment, vol. 70, May 2013, pp. 64–74, https://doi.org/10.1016/j.atmosenv.2012.12.025.

Science Direct. (2007). Inducible nitric oxide synthase. Inducible Nitric Oxide Synthase - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/neuroscience/inducible-nitric-oxide-synthase

Seinfeld, John H. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. John Wiley & Sons, Incorporated, 2016.

Senst, B. (2023, August 22). Hypercoagulability. StatPearls [Internet]. https://www.ncbi.nlm.nih.gov/books/NBK538251/

Simon, M. C., Hudda, N., Naumova, E. N., Levy, J. I., Brugge, D., & Durant, J. L. (2017). Comparisons of traffic-related ultrafine particle number concentrations measured in two urban areas by central, residential, and mobile monitoring. Atmospheric Environment, 169, 113–127. https://doi.org/10.1016/J.ATMOSENV.2017.09.003

Sirithian, Duanpen, and Pantitcha Thanatrakolsri. "Relationships between Meteorological and Particulate Matter Concentrations (PM2.5 and PM10) during the Haze Period in Urban and Rural Areas, Northern Thailand." *Air, Soil and Water Research*, vol. 15, 2022, https://doi.org/10.1177/11786221221117264.

"The Deadly Impact of Coal Pollution - beyond Coal - Sierra Club." Sierra Club, coal.sierraclub.org/sites/nat-coal/files/Out%20of%20Control%20coal%20mortality%20report%20FINAL.pdf. Accessed 17 Apr. 2025.

"US-EPA New England Region 1: Quality Assurance Staff." Office of Environmental Measurement and Evaluation, Oct. 1999.

Utell, M. J., & Frampton, M. W. (2000). Acute Health Effects of Ambient Air Pollution: The Ultrafine Particle Hypothesis. In JOURNAL OF AEROSOL MEDICINE (Vol. 13, Issue 4). Mary Ann Liebert, Inc. Pp. www.liebertpub.com

WANG, YUNGANG, ET AL. "URBAN-SCALE SEASONAL AND SPATIAL VARIABILITY OF ULTRAFINE PARTICLE NUMBER CONCENTRATIONS." *WATER, AIR, AND SOIL POLLUTION*, VOL. 223, NO. 5, 2012, PP. 2223–35, HTTPS://DOI.ORG/10.1007/S11270-011-1018-Z.

WHO GLOBAL AIR QUALITY GUIDELINES. PARTICULATE MATTER (PM2.5 AND PM10), OZONE, NITROGEN DIOXIDE, SULFUR DIOXIDE AND CARBON MONOXIDE. GENEVA: WORLD HEALTH ORGANIZATION; 2021. LICENCE: CC BY□NC□SA 3.0 IGO.

WRIGHT, R. J., HSU, H. H. L., CHIU, Y. H. M., COULL, B. A., SIMON, M. C., HUDDA, N., SCHWARTZ, J., KLOOG, I., & DURANT, J. L. (2021). PRENATAL AMBIENT ULTRAFINE PARTICLE EXPOSURE AND CHILDHOOD ASTHMA IN THE NORTHEASTERN UNITED STATES. AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE, 204(7), 788–796. HTTPS://DOI.ORG/10.1164/RCCM.202010-3743OC

YUAN, Y., ZHU, Y., & WU, J. (2011). MODELING TRAFFIC-EMITTED ULTRAFINE PARTICLE CONCENTRATION AND INTAKE FRACTION IN CORPUS CHRISTI, TEXAS. CHEMICAL PRODUCT AND PROCESS MODELING, 6(1). HTTPS://DOI.ORG/10.2202/1934-2659.1510

C. GitHub Repository

https://github.com/cgimilaro/Air-Pollution