

Abstract

In order to study the influence of topography coupled with wind speed and direction on pollution rates, a PEJA located in the South End of Albany, NY was chosen for study due to a moderately large hill and measuring station in the neighborhood. The data for the clustering study was collected from a year-long, hourly-averaged data set provided by the New York State Department of Environmental Conservation on a public access website, with the date interval ranging from 24 Sept. 2018 to 24 Sept 2019. The data comes from the measuring station located in the Ezra Prentice neighborhood as part of a network of statewide measuring stations. The main method of processing the directional correlation was K-Means Clustering. While the Nanoparticle distributions proved there is good mixing in the region, the K-Means clustering algorithm, using 15 clusters per chart, pointed towards a Northeastern correlation for origin of PM 2.5 and Black Carbon particles. However, further study is required in order to test for topography with a hillside with multiple measuring stations.

Set Up of the Data Set

- Year-long, hourly data from 24 Sept. 2018 to 24 Sept 2019.
- Observations between July 20th 2019 and August 18th 2018 were removed due to instrument issues.
- The final size of the data set was 5.2k observations

Motivation

This work is a continuation of a Summer Internship at the New York State Department of Environmental Conservation: Division of Air. Since signing up for the class would grant new tools to analyze air quality data in the Ezra Prentice neighborhood, it was believed that they would help the research. For comparison, a lot of the work presented will be compared to the DEC's report on the neighborhood published in October of 2019.

Initial Processing

- Shapiro test on the Normal and LogNormal Values of the pollutants (failed).
- Initial Plots of the data involves plotting of the cleaned data set, and looking for any patterns between amounts of pollutants and Wind Direction. (Figure 1)
- It should be noted that 0 degrees means the wind is coming from the North and going South

Functions were also established to create vectors of the Wind Speed and Wind Direction:

$$\begin{aligned} \text{X-Vector Adjustment} &= \text{WS} * \cos(\text{WD}/180*\pi) \\ \text{Y-Vector Adjustment} &= \text{WS} * \sin(\text{WD}/180*\pi) \end{aligned}$$

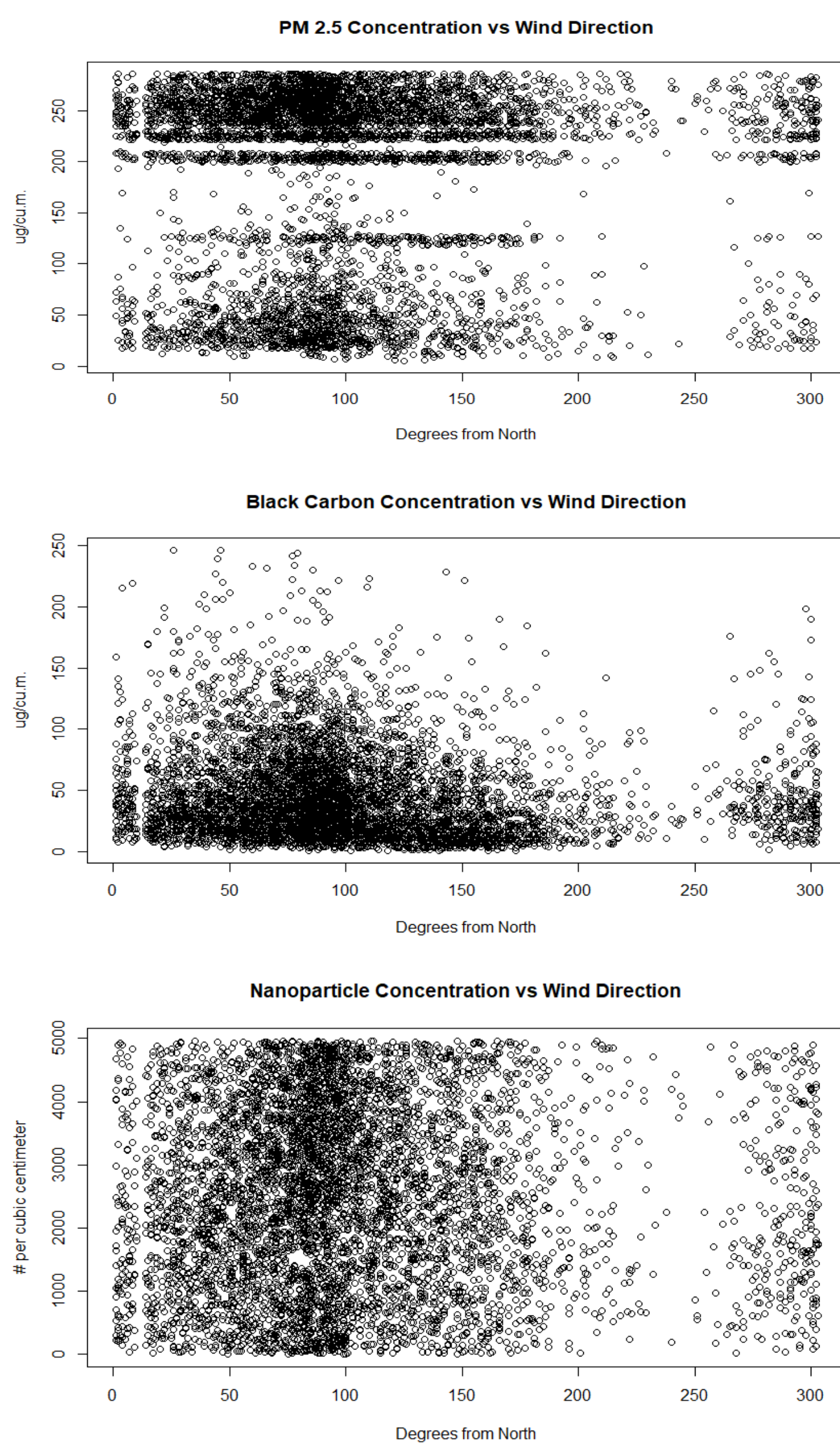


Figure 1: Initial plotting charts

Glossary:

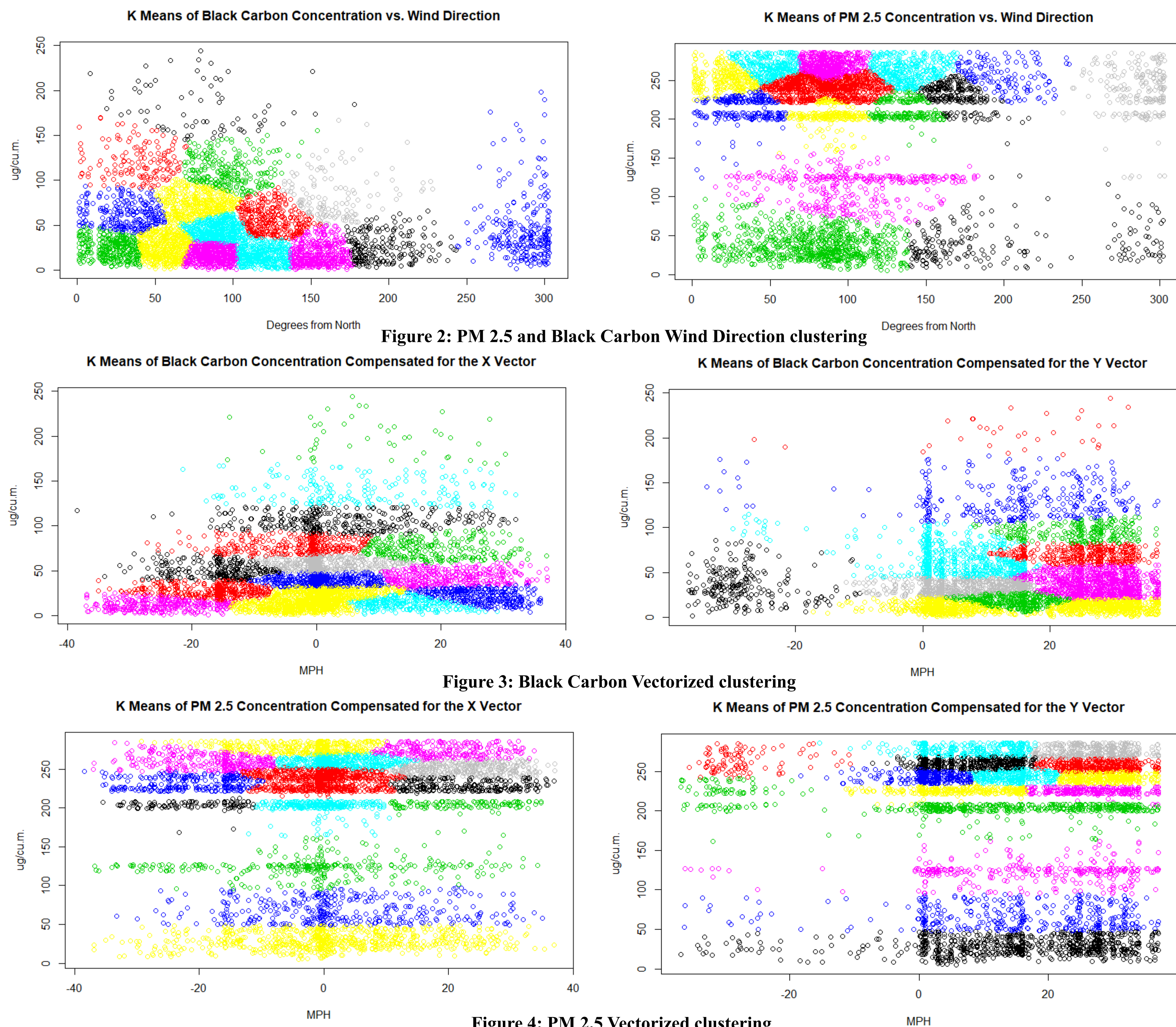
WD – Wind Direction
WS – Wind Speed
PC – Nanoparticle Counts
PM 2.5 – 2.5-micrometer Particle Concentrations
BC – Black Carbon Concentration
PEJA: Potential Environmental Justice Area

Acknowledgments:

Dr. Brian Frank, Marilyn Wurth, Gil LaDuke, Thilanka Munasinghe

Further Processing

- K-Means Clustering was used on data frames of pairs coupled with pollutants and wind directions. (Figure 2)
- The Second Stage of K-Means was conducted on wind directions vectorized to account for wind speed. (Figures 3 and 4)
- 15 clusters were used for each K-Means as that is when groupings appeared to stop changing.



Overview of Ezra Prentice Neighborhood

- Located in South End of Albany.
- Lots of nearby industry. (Figure 5)
- Path denotes a 105 ft rise in hillside in just 0.2 miles
- Black denotes north while the light pink denotes the direction of pollutant origin direction. (Figure 6)

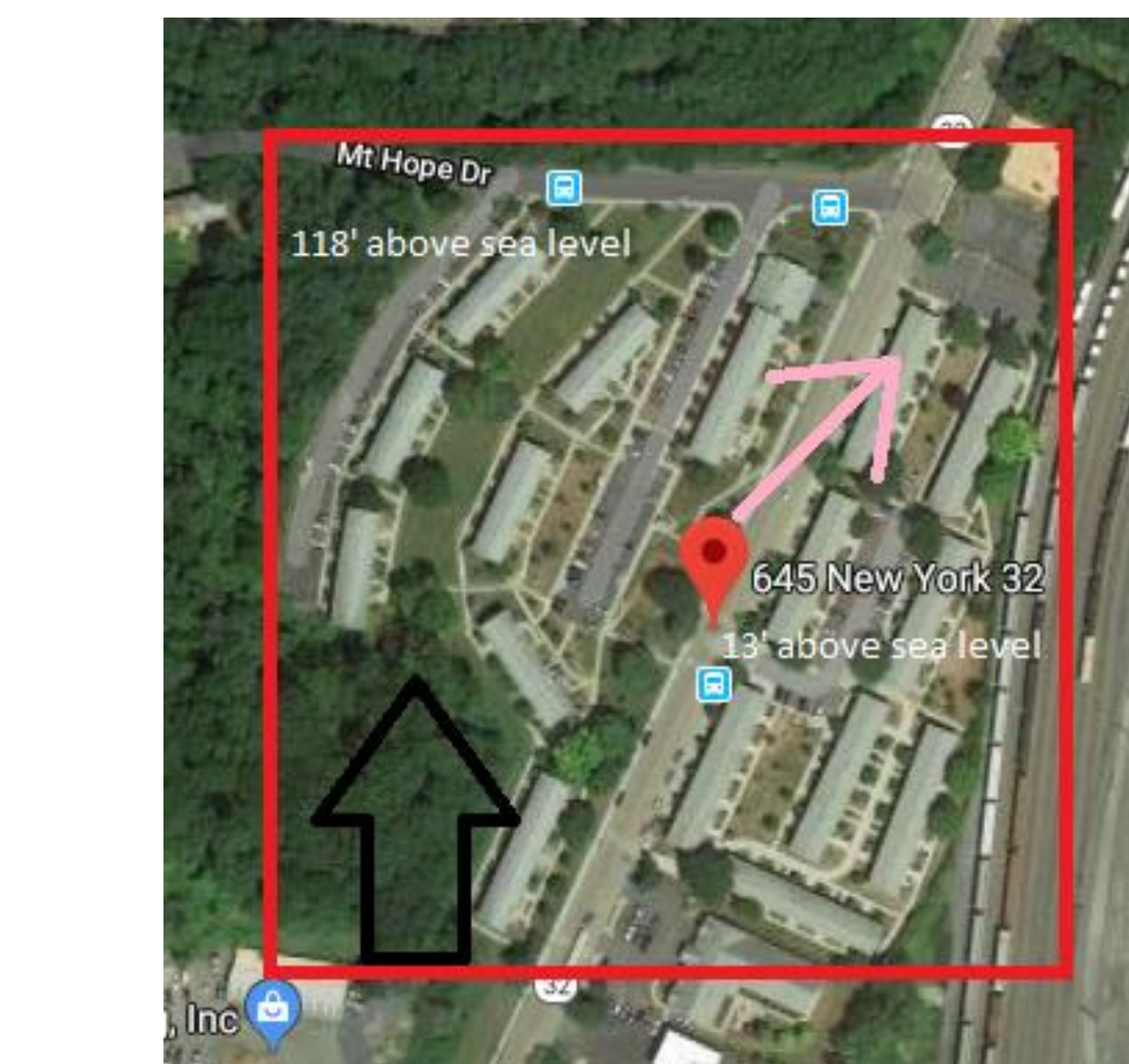


Figure 6: Localized Map

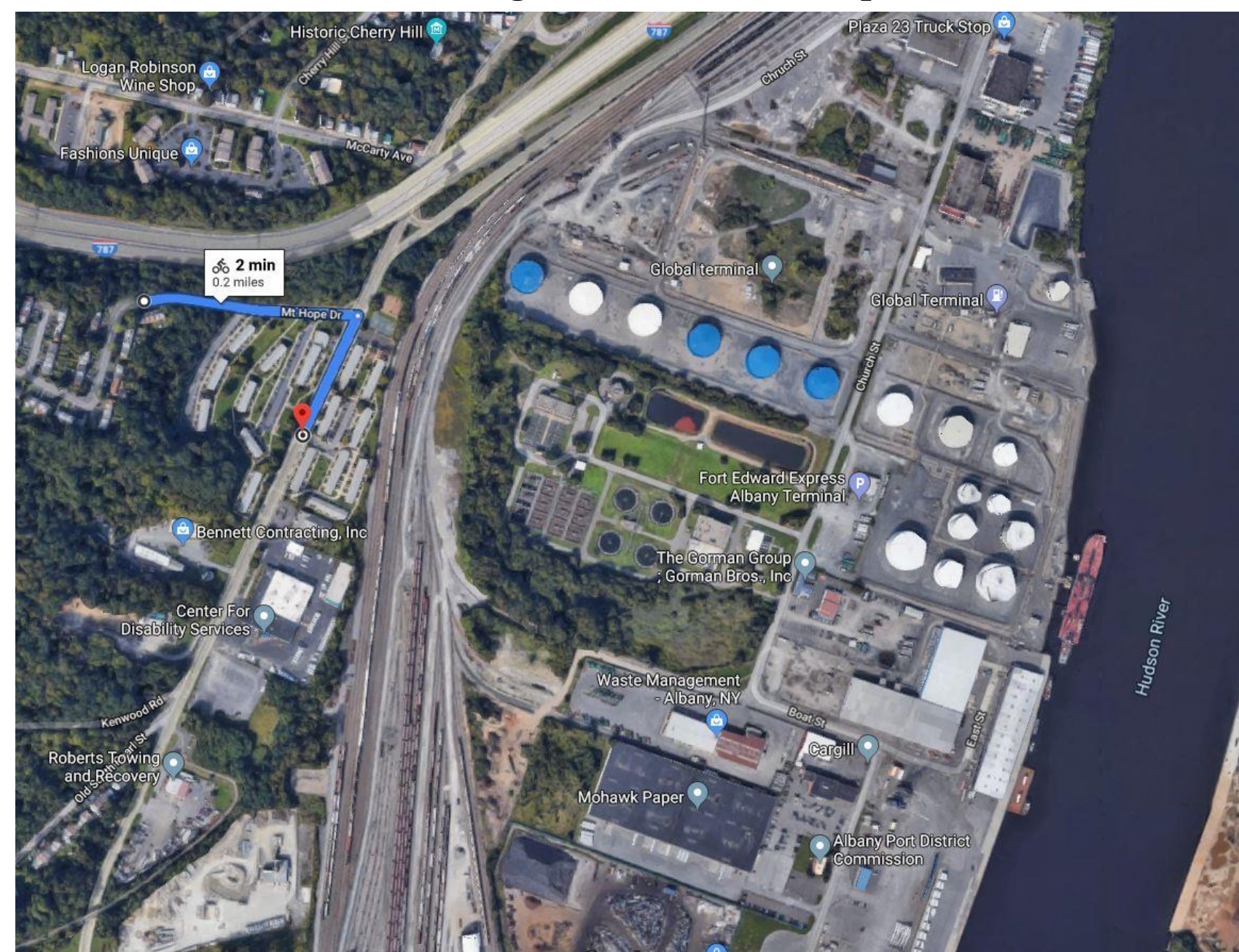


Figure 5: Area Map with Elevation rise

Observations

- There is a less pollution at around 250 degrees from North.
- A peaking that appears to occur around 90 degrees in each of the pollutant's wind direction distributions.
- Clustering of the data set revealed that there is a Northeastern tendency of pollutant origins.
- The clustering data also revealed that the Nanoparticle distribution in the area is well-mixed.
- Wind speed averages recorded were on the scale of 26MPH.

Conclusions, Further Study, Difficulties

- Initial plotting points to an Eastern correlation with regards to origin of pollutants.
- Depression in the pollution rates coming from the hillside of the residential area
- No way to control the averaging system used and accepting the hourly-averaged data set as provided by NYS DEC.
 - Very high readings of wind speeds of 25+MPH and gaps in the PM 2.5 initial plots.
 - When compared to a DEC report published in October of 2019, wind speeds from the monitoring station were recorded at and average 8.8 MPH from June to July of 2018, while the publicly-exported hourly data had an average of 4.4 MPH.
- Pollution roses of Nanoparticles in the DEC report agree with the Nanoparticle Clustering that they are well-mixed in the area.
- There is a directional correlation on the pollution rates of PM 2.5 particles and Black Carbon.
- A hillside with a few traffic pathways and some source of industrial pollution would be ideal in further testing this theory.
- CFD with this specific geography in mind would also help the study.

References:

Neeldip Barman and Sharad Gokhale, "Urban black carbon - source apportionment, emissions and long-range transport over the Brahmaputra River Valley", *Science of The Total Environment*, Volume 693, 2019

Lu, X., Zhu, T., Chen, C., & Liu, Y. (2014). Right or left: the role of nanoparticles in pulmonary diseases. *International journal of molecular sciences*, 15(10), 17577–17600. doi:10.3390/ijms151017577

Ana Fernández Tena, Pere Casan Clara, "Deposition of Inhaled Particles in the Lungs", *Archivos de Bronconeumologia (English Edition)*, Volume 48, Issue 7, 2012.

"What are the Air Quality Standards for PM?" Retrieved November 27, 2019, from <https://www3.epa.gov/region1/airquality/pm-aq-standards.html>.

Junyan Yang, Beixiang Shi, Yi Shi, Simon Marvin, Yi Zheng, Geyang Xia, "Air pollution dispersal in high density urban areas: Research on the triadic relation of wind, air pollution, and urban form", *Sustainable Cities and Society*, 2019.

New York State Department of Environmental Conservation (NYS DEC), Division of Air Resources. *Albany South End Community Air Quality Study*. October 2019.