A Database of PurpleAir Particulate Matter Data for South Central Pennsylvania

Completed for the Clean Air Board of Central Pennsylvania

August 3, 2022

Timothy W. Hawkins and Kelsey Bortner
Professor and Chair
Geography-Earth Science Department
Shippensburg University
twhawk@ship.edu

CLUS graduate student fellow

Table of Contents

List of Figures and Tables

Funding Acknowledgement

This project was funded in part by a grant from the Clean Air Board of Central Pennsylvania.

Purpose

The purpose of this work is to create a readily accessible and easily updatable database of PurpleAir monitor data for Central Pennsylvania. Additionally, basic statistics are calculated for the data and a comparison is made with daily weather type data.

Rationale

Due to agriculture, industry, interstate highways, and surrounding mountains, the Harrisburg-Carlisle metropolitan area has the second worst air quality in the state of Pennsylvania with 114 days in 2018 when half or more of the monitoring locations reported elevated ozone and/or particulate matter pollution (Ridlington et al., 2020). Air pollution impacts are most pronounced in children, pregnant women, and the elderly and include respiratory illness, heart attack, stroke, cancer, dementia, mental health problems, and birth complications. Vehicle and power plant pollution were responsible for an estimated 4,800 premature deaths in Pennsylvania in 2018 (Goodkind et al., 2019).

To better understand air quality in south central Pennsylvania, individuals and organizations (including the Clean Air Board) have installed several low-cost PurpleAir particulate matter monitors. Data for these monitors are archived within the PurpleAir website. However, performing analyses with the data is cumbersome given the process to access the data. This project creates a spreadsheet of regional PurpleAir data that can be easily updated so that interested citizens and researchers can examine the data more readily.

Data Availability

PurpleAir data are available via the mapping tool on their website. To facilitate easy monthly updates of the database, at our request, PurpleAir created a custom map and data download pages for Central Pennsylvania. Figure 1 shows a screen shot of the map for June 21, 2022, at 12:36 PM and Table 1 shows basic location information. Locations range from Reading to Shippensburg in the east-west direction and Selinsgrove to Waynesboro in the north-south direction.

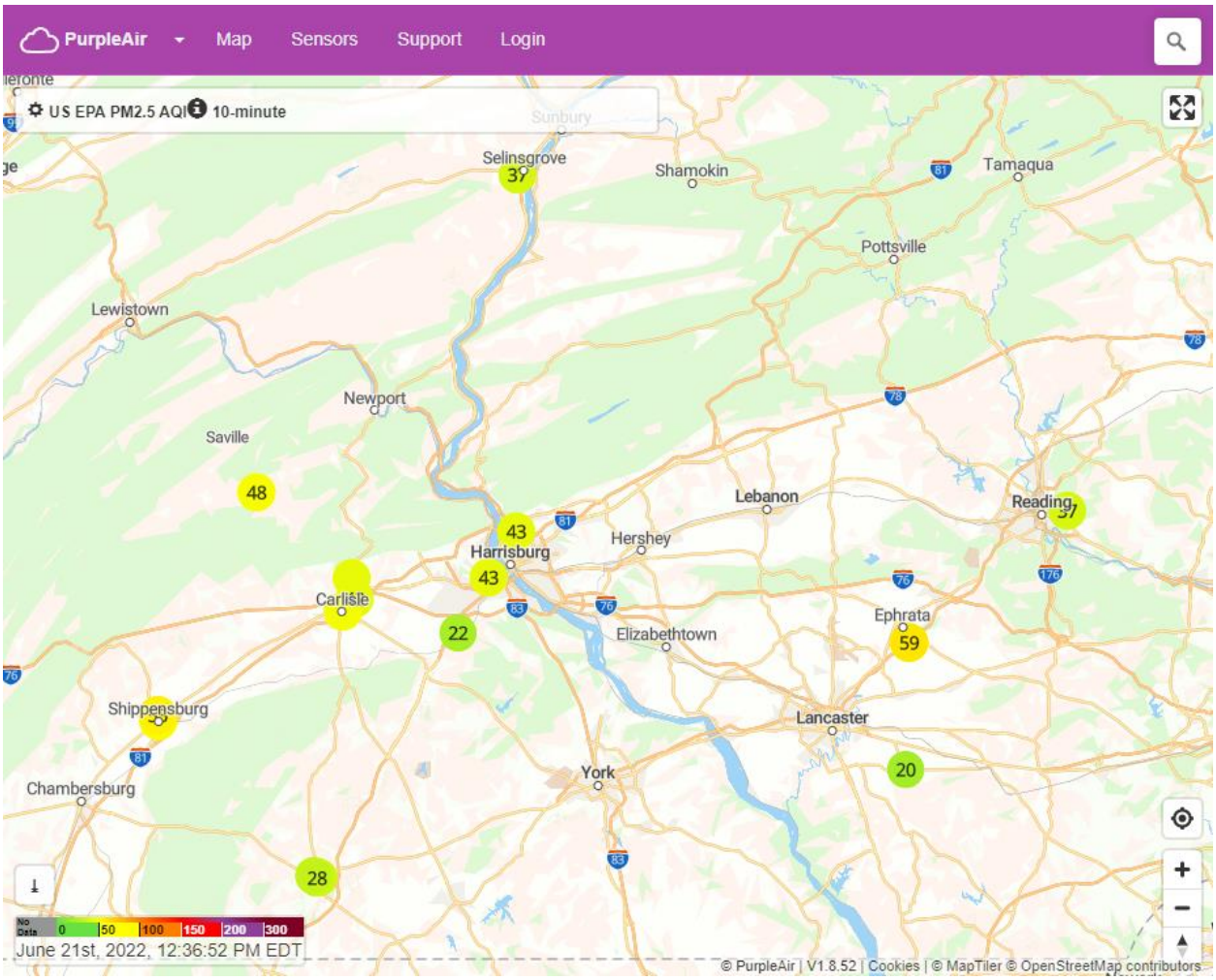


Figure 1. Screenshot of the custom-made Central Pennsylvania real-time PurpleAir data map.

Database Construction

Raw data from the PurpleAir download page consists of multiple files for each monitor and is inconsistent between the 18 monitors with respect to times of observation and observation frequency. Additionally, each monitor has a different start date and set of observations at the beginning of the records that were clearly taken before the monitor was installed in its intended location. Raw data are converted to hourly averages and maximums. The number of observations used to calculate the hourly average is recorded. Observations prior to installation are removed.

All PurpleAir monitors have an A and B sensor that each record particulate matter concentrations (µg/m^3) of particles ≤1, ≤2.5, and ≤10 µm (PM_1, $PM_{2.5}$, and PM_{10} respectively). Both hourly A and B values are recorded as well as the average of and difference between the A and B sensors.

Table 1. Location names and general location of monitors in Figure 1.

Monitor #	Name	General Location
1	115 Old Post Lane	SE of Lancaster
2	CAB Mechanicsburg	SE of Mechanicsburg
3	Clean Air Board	Carlisle
4	Clean Air Board - Susquehanna Twp	Susquehanna Township
5	Clean Air Board - CAMP HILL	Camp Hill
6	Gettysburg Carlisle St.	Gettysburg
7	Gettysburg Central Campus	Gettysburg
8	Gettysburg West Campus	Gettysburg
9	LeTort Spring Run Trail Head	Carlisle
10	Meglio - Loysville	Loysville
11	Middlesex Township	Middlesex Township
12	Middlesex Twp., Meadowbrook Farms	Middlesex Township
13	Mt Penn	E of Reading
14	Murrell	S of Ephrata
15	Shippensburg University - King Street	Shippensburg
16	Shippensburg University - Shearer Hall	Shippensburg
17	Susquehanna University	Selinsgrove
18	Wayne Heights - Old Forge Road	E of Waynesboro

No other quality controls are performed as it is deemed better for individuals to perform their own quality controls depending on their intent in accessing the data. As such, there are hours with no data and observation counts of zero. There are also hours with unreasonably high values or zero values for one or both sensors. Zero values may or may not be correct. Care should be taken when selecting individual or average sensor data for given hours. The analysis section below details a quality control procedure that was used for this report. Others may use these criteria as a guide.

Database Accessibility

The database is hosted on the Shippensburg University weather website (select Air Quality). The entire weather website is scheduled for a redesign so minimal effort was made on design aesthetics for the air quality page. The air quality page contains links to PurpleAir pages, including the custom maps and raw data download pages, along with links to *.csv

data files for each of the 18 monitors and a "Read Me" file to interpret the data files. Updating of the air quality data files will occur monthly as part of the workflow that currently exists for updating university weather data.

Basic Data Analysis

A basic analysis of hourly monitor data was completed to ensure that the data compilation process is accurate and the data are suitable for distribution. A quality control procedure was run on the hourly data that includes:

- Setting as missing data, all data for hours when either the A or B sensor have an average PM_{10} value > 1000 $\mu g/m^3$ or = 0 $\mu g/m^3$.
- Setting as missing data, all data for hours when either the A or B sensor have a maximum PM_{10} value > 2000 $\mu g/m^3$.
- Setting as missing data, all data for hours when either the A or B sensor have observation counts = 0.
- For all the above criteria, if only one sensor meets a specific criterion, the other sensor is not set to missing but the values for the average of and difference between both sensors are set to missing.

The resulting hourly time series of average and maximum values for each sensor and the sensor average are plotted for all 18 monitors (Appendix A). Overall average and average maximum PM concentrations are calculated for each particle size for each monitor (Table 2). Hourly average and maximum average PM_{10} concentrations are 16.34 and 22.46 $\mu g/m^3$ respectively. For perspective, the primary and secondary PM_{10} National Ambient Air Quality Standard is a 24-hour average concentration of 150 $\mu g/m^3$. Monitors 1 (SE of Lancaster), 10 (Loysville), and 14 (S of Ephrata) generally have the highest PM concentrations while monitors 2 (SE of Mechanicsburg) and 18 (E of Waynesboro) generally have the lowest. More detailed analyses are needed to determine the reasons for any geographic patterns in the data.

Seasonal average PM concentrations were calculated for each particle size and averaged across all 18 monitors (Figure 2). PM concentration is generally highest in winter (DJF) then summer (JAS) and lower in spring (MAMJ) and fall (ON). Winter and summer are the seasons where larger high pressure weather systems are likely to set up that promote the build-up of PM in the atmosphere over multiple days. Annual concentrations averaged across all 18 monitors are 9.79, 15.13 and 16.60 $\mu g/m^3$ for PM_1, $PM_{2.5}$, and PM_{10} respectively.

Table 2. Overall average and average maximum PM concentration (µg/m^3). Orange and blue shading are column maximum and minimum values respectively.

	Average			Maximum		
	PM_1	$PM_{2.5}$	PM_{10}	PM_1	$PM_{2.5}$	PM_{10}
Average	9.51	14.81	16.34	12.77	20.06	22.46
Monitor 1	13.73	18.18	19.28	18.18	23.90	25.50
Monitor 2	5.36	10.03	11.35	7.69	13.64	15.46
Monitor 3	10.34	16.21	18.83	12.87	20.38	24.16
Monitor 4	10.27	15.88	16.87	12.19	18.96	20.28
Monitor 5	10.06	15.42	16.36	13.14	19.96	21.06
Monitor 6	9.59	14.34	15.80	12.35	18.95	21.69
Monitor 7	7.92	12.42	13.64	11.17	16.78	18.25
Monitor 8	8.51	14.01	15.42	10.27	17.08	18.85
Monitor 9	9.51	15.20	16.20	12.29	19.98	21.38
Monitor 10	11.58	19.14	21.93	21.12	37.05	44.19
Monitor 11	8.72	13.64	15.35	11.43	17.91	20.22
Monitor 12	10.27	15.56	16.80	13.62	21.08	23.49
Monitor 13	8.90	13.01	13.93	11.09	16.30	17.49
Monitor 14	13.23	21.23	24.23	17.60	28.87	33.69
Monitor 15	5.65	10.96	12.34	8.37	15.08	16.94
Monitor 16	11.67	16.92	18.48	14.28	20.69	22.78
Monitor 17	10.48	15.64	17.15	13.23	19.46	21.25
Monitor 18	5.35	8.86	10.18	8.93	15.00	17.67

To begin to assess the relationship between weather and PM concentration in south central Pennsylvania, daily weather type data were analyzed for Harrisburg International Airport (HIA). Weather types are based on meteorological data recorded at HIA and represent an integrative method for assigning each day a weather type based on temperature, humidity, pressure, and other weather data (Sheridan, 2002). While not a perfect analog, weather types can be thought of as synonymous with airmasses. The different weather types are Dry Moderate (DM), Dry Polar (DP), Dry Tropical (DT), Moist Moderate (MM), Moist Polar (MP), Moist Tropical (MT), and Transition (T). These weather types exist across all seasons, but the characteristics of each weather type changes based on the season.

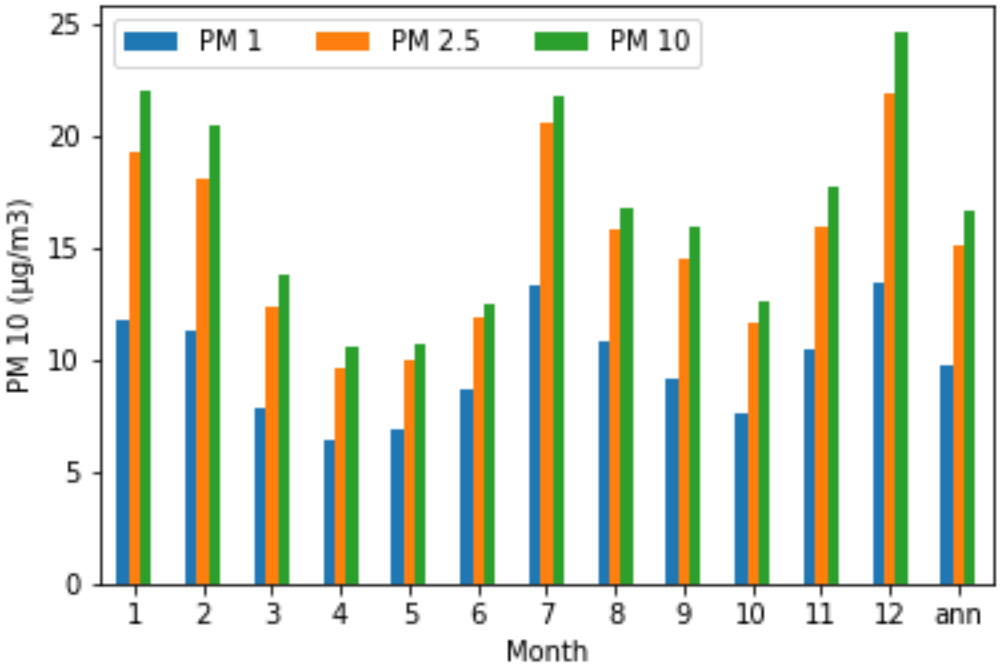


Figure 2. Seasonal average PM concentration (µg/m^3) averaged across all 18 monitors.

Daily weather type data were compared with daily average PM_{10} concentration averaged across all 18 monitors for each season (Table 3). For spring, summer, and fall, highest concentrations occur during Dry Tropical days. DT days during these seasons are characterized by large high-pressure systems and little rain. These air masses are likely from the central plains of the US. In winter, similar conditions are found with the Moist Polar weather type which results in the highest PM concentrations. Across all seasons, DT and MP weather types likely occur before storm events.

Lowest PM concentrations for all seasons are associated with Dry Polar weather types. DP days typically occur after a significant storm system and cold frontal passage and are characterized by precipitation in the days before and windy conditions.

Table 3. Seasonal PM_{10} concentration and number of days for each weather type. Orange and blue shading are seasonal maximum and minimum concentrations respectively.

Season	Weather Type	PM_{10} ($\mu g/m^3$)	Days
Winter (DJF)	Dry Moderate	18.46	112
	Dry Polar	14.25	47
	Dry Tropical	18.79	9
	Moist Moderate	19.46	36
	Moist Polar	28.67	31
	Moist Tropical	17.02	8
	Transition	14.43	28
Spring (MAM)	Dry Moderate	10.66	87
	Dry Polar	6.24	41
	Dry Tropical	14.32	25
	Moist Moderate	10.63	35
	Moist Polar	8.43	14
	Moist Tropical	12.69	43
	Transition	9.11	31
Summer (JJA)	Dry Moderate	10.68	40
	Dry Polar	3.61	3
	Dry Tropical	20.22	18
	Moist Moderate	16.91	20
	Moist Polar	8.02	2
	Moist Tropical	17.44	94
	Transition	14.36	7
Fall (SON)	Dry Moderate	11.60	112
	Dry Polar	8.37	25
	Dry Tropical	19.11	6
	Moist Moderate	11.46	38
	Moist Polar	11.11	10
	Moist Tropical	16.91	50
	Transition	12.33	27

References

Goodkind, AL, Tessum, CW, Coggins, JS, Hill, JD, Marshall, JD. (2019). Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. Proceedings of the National Academy of Sciences. 116(18): 8775-8780. doi: https://doi.org/10.1073/pnas.1816102116

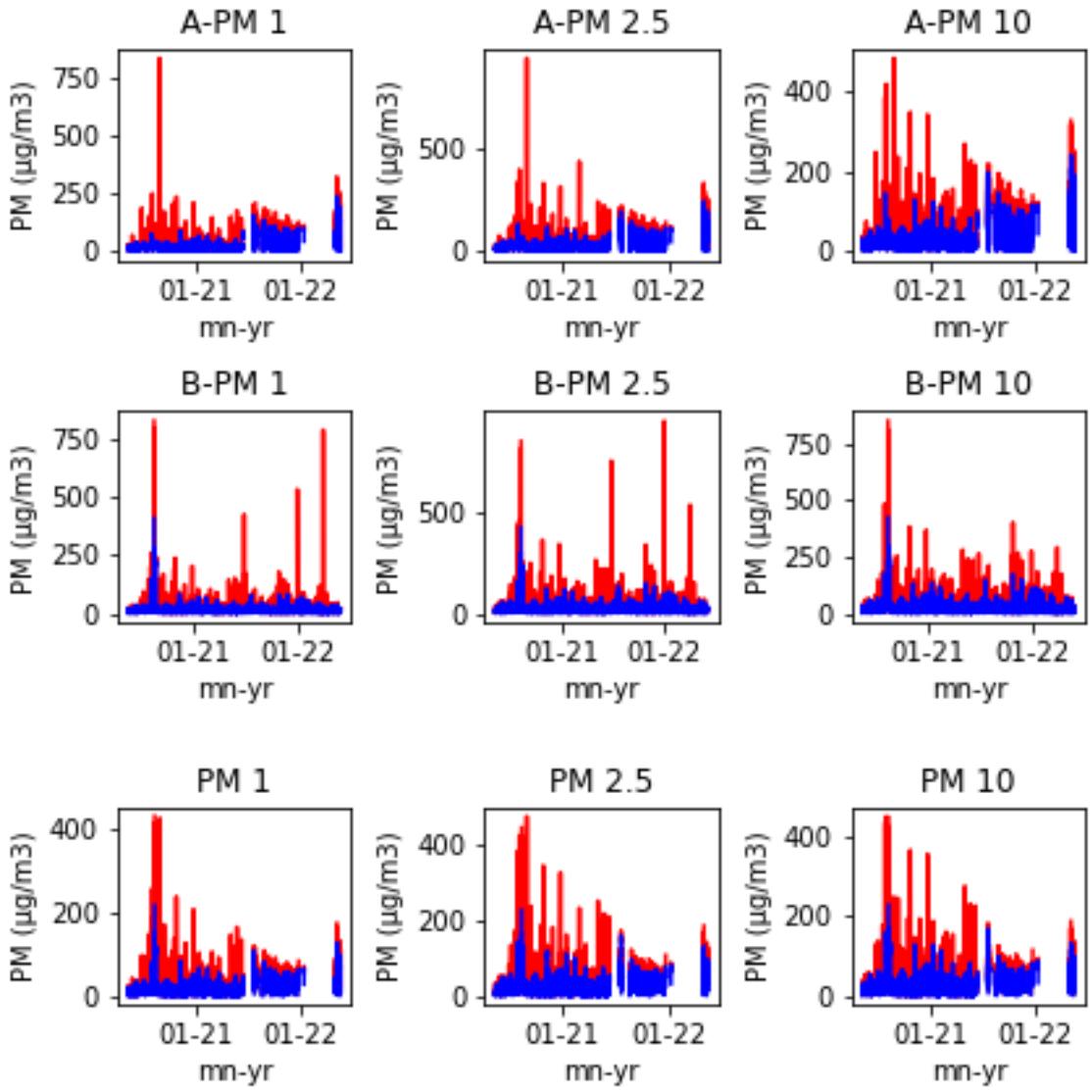
Ridlington, E, Weissman, G, Folger, M. (2020). Trouble in the Air: Millions of Americans Breathed Polluted Air in 2018. PennEnvironment Research & Policy Center and PennPIRG Education Fund.

Sheridan, SC. (2002). The redevelopment of a weather-type classification scheme for North America. International Journal of Climatology. 22: 51–68. DOI: 10.1002/joc.709

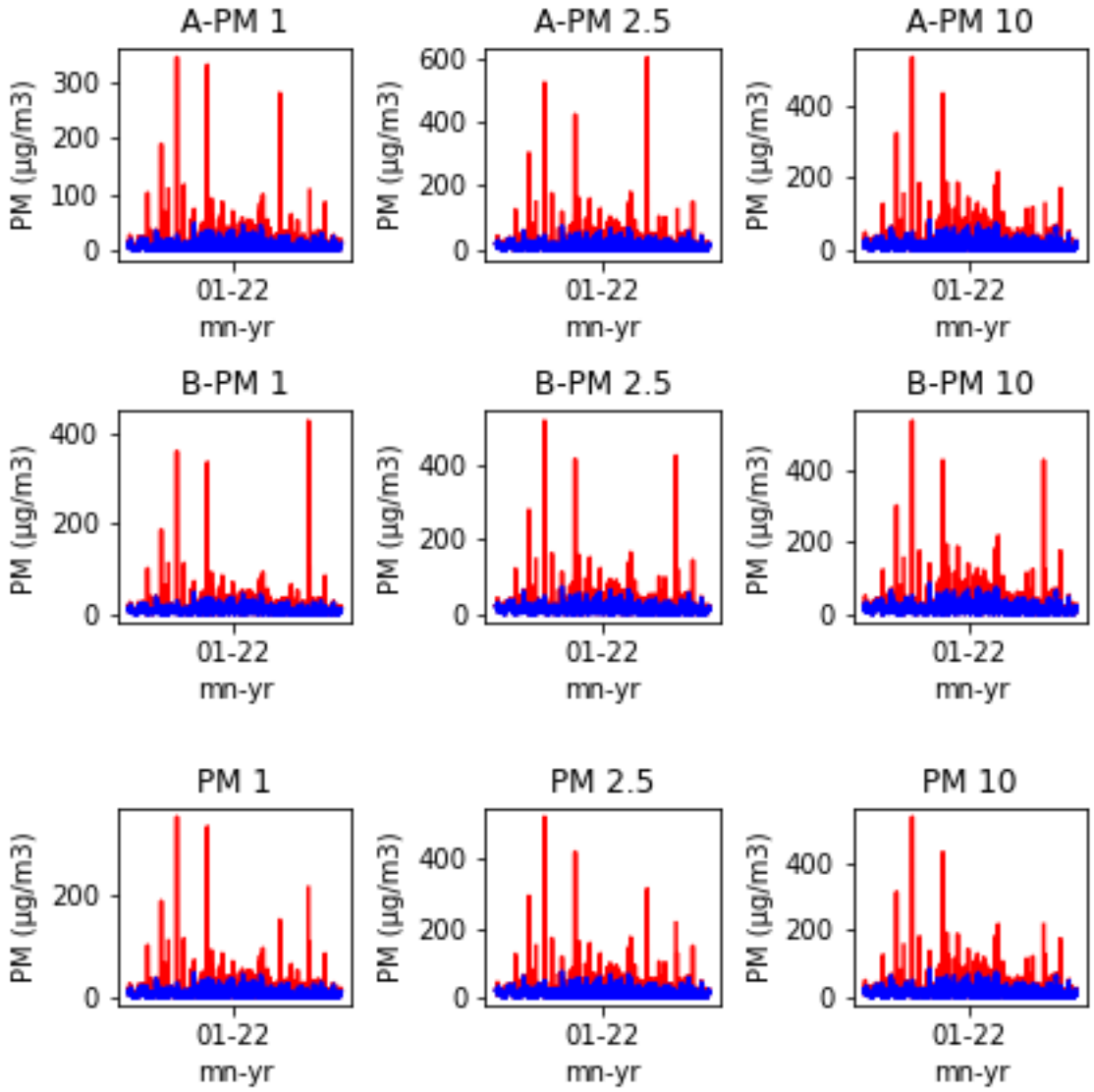
Appendix A

Time series of quality controlled hourly data for all 18 PurpleAir monitors. Columns are PM_1, $PM_{2.5}$, and PM_{10} respectively. Rows 1 and 2 are sensors A and B respectively. Row 3 is the average of sensors A and B. Red and blue are the hourly maximum and average respectively.

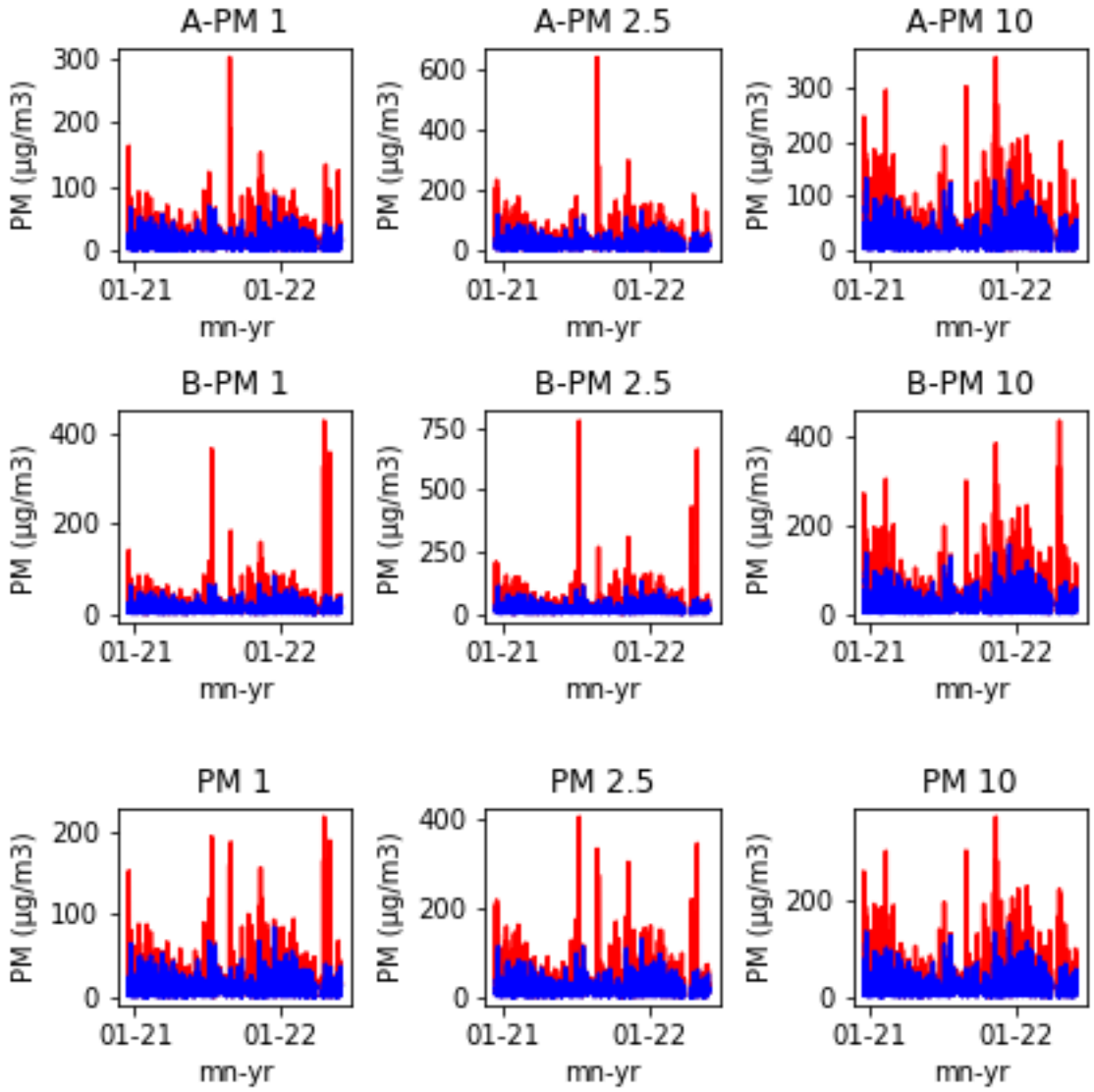
1. 115 Old Post Lane



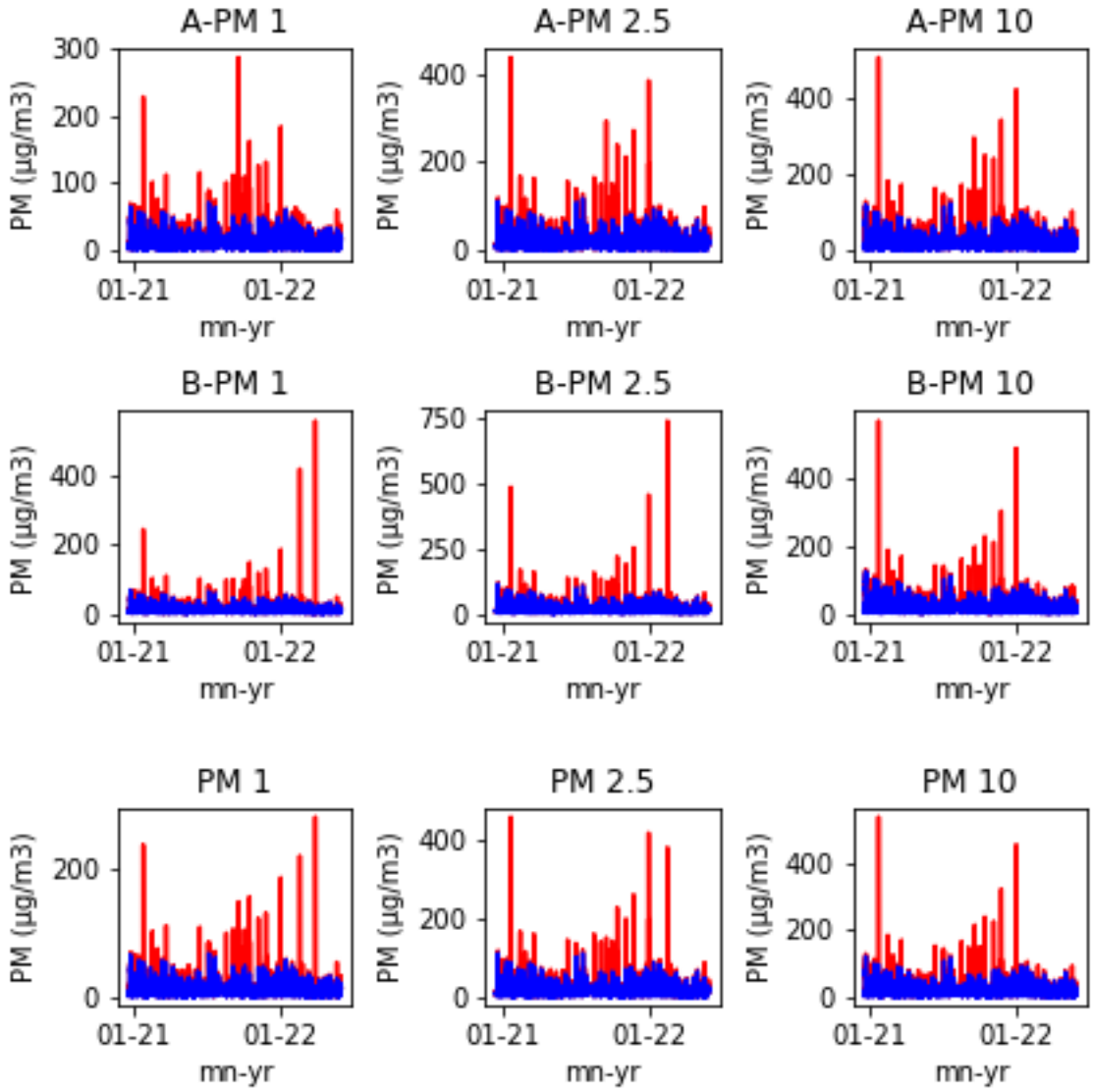
2. CAB Mechanicsburg



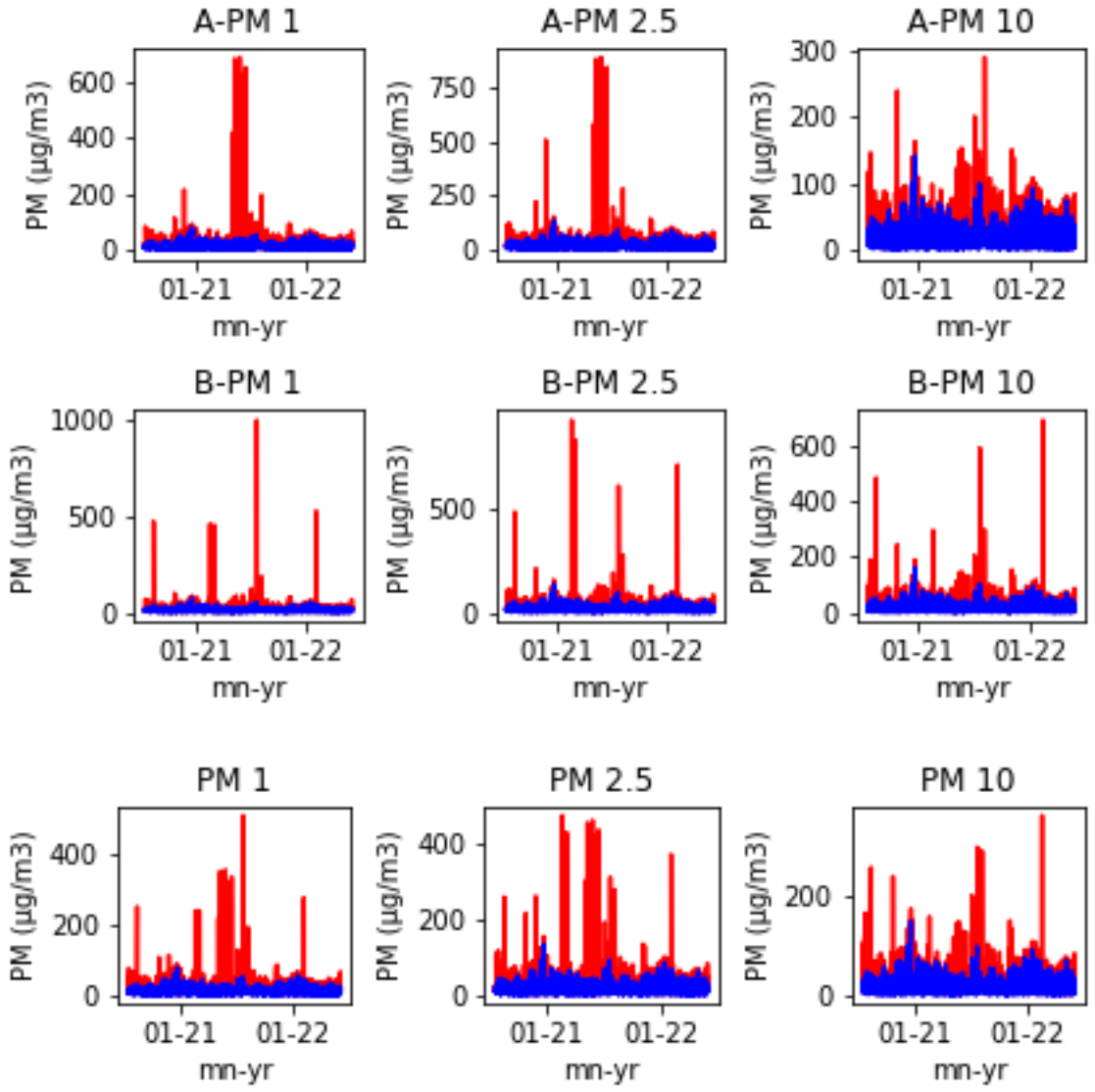
3. Clean Air Board - Susquehanna Twp



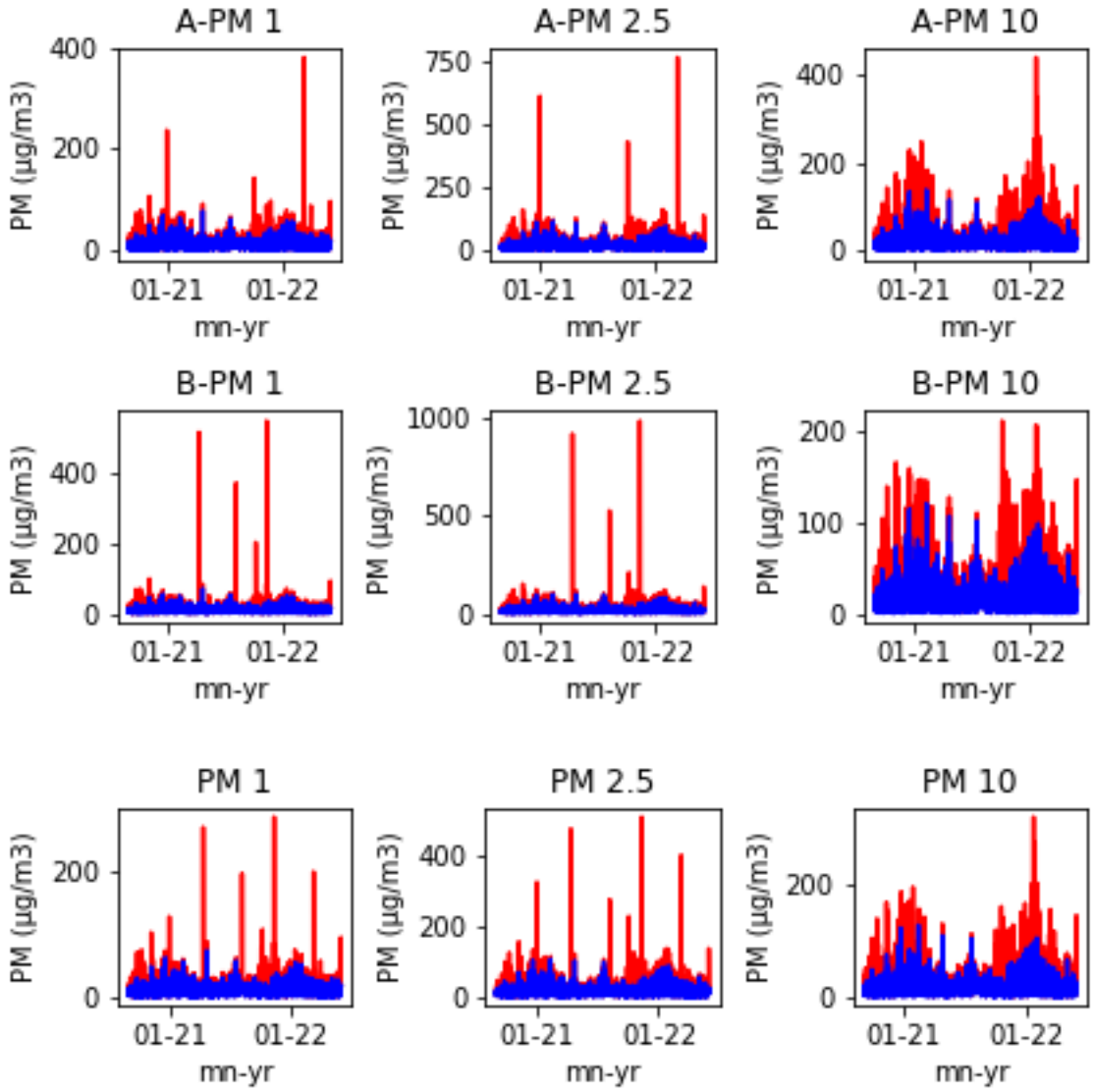
4. Clean Air Board - CAMP HILL



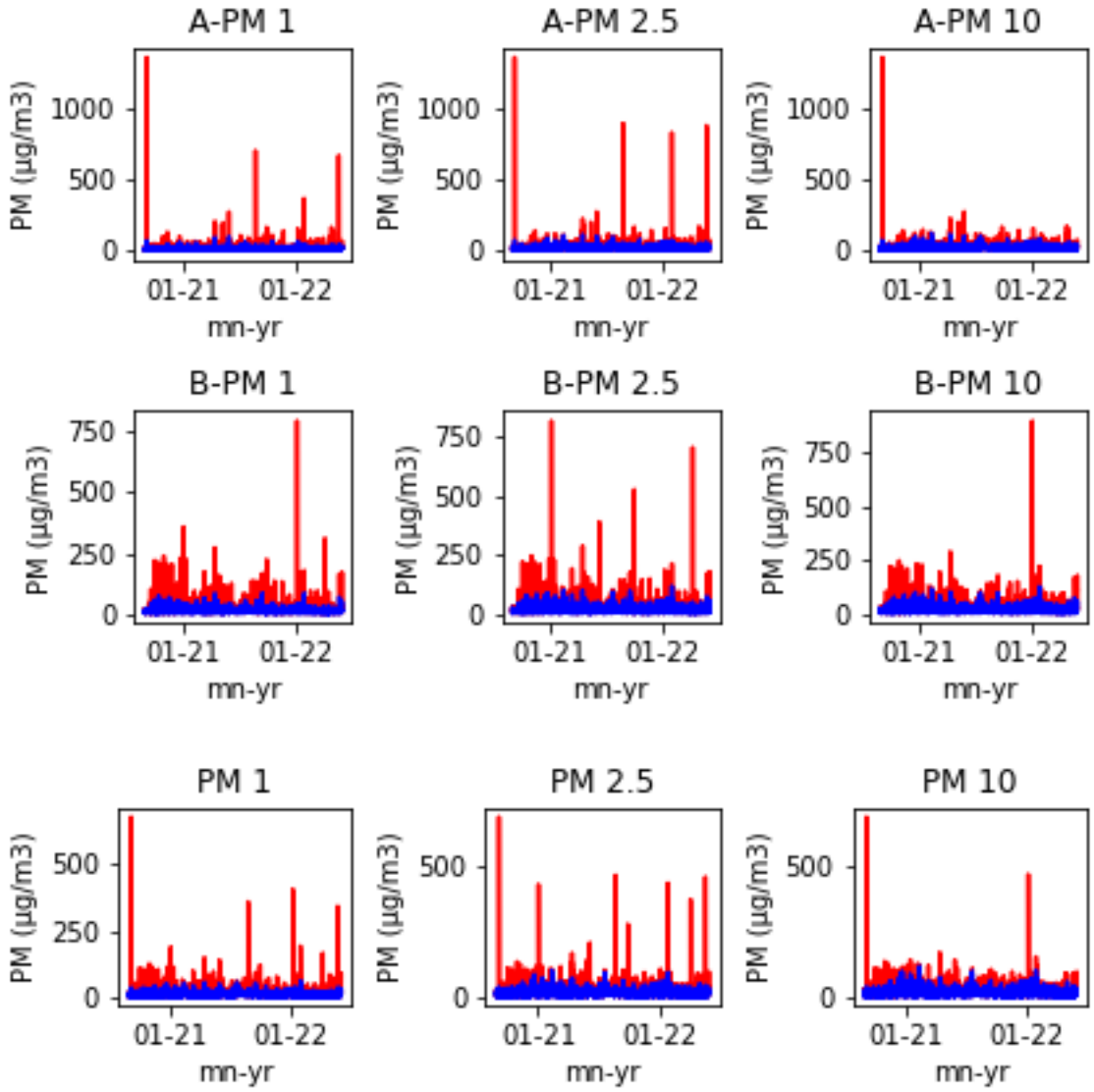
5. Clean Air Board



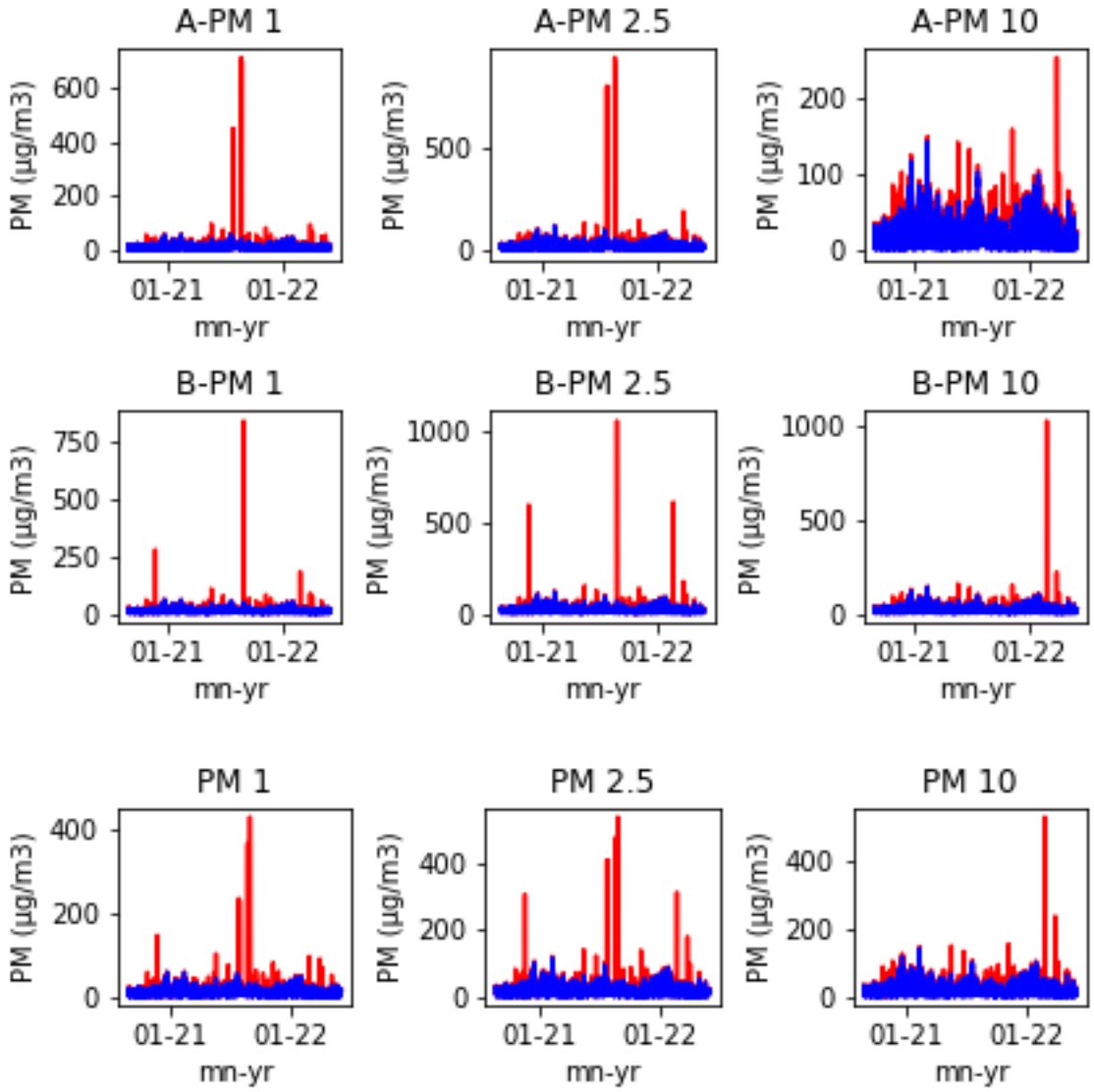
6. Gettysburg Carlisle St



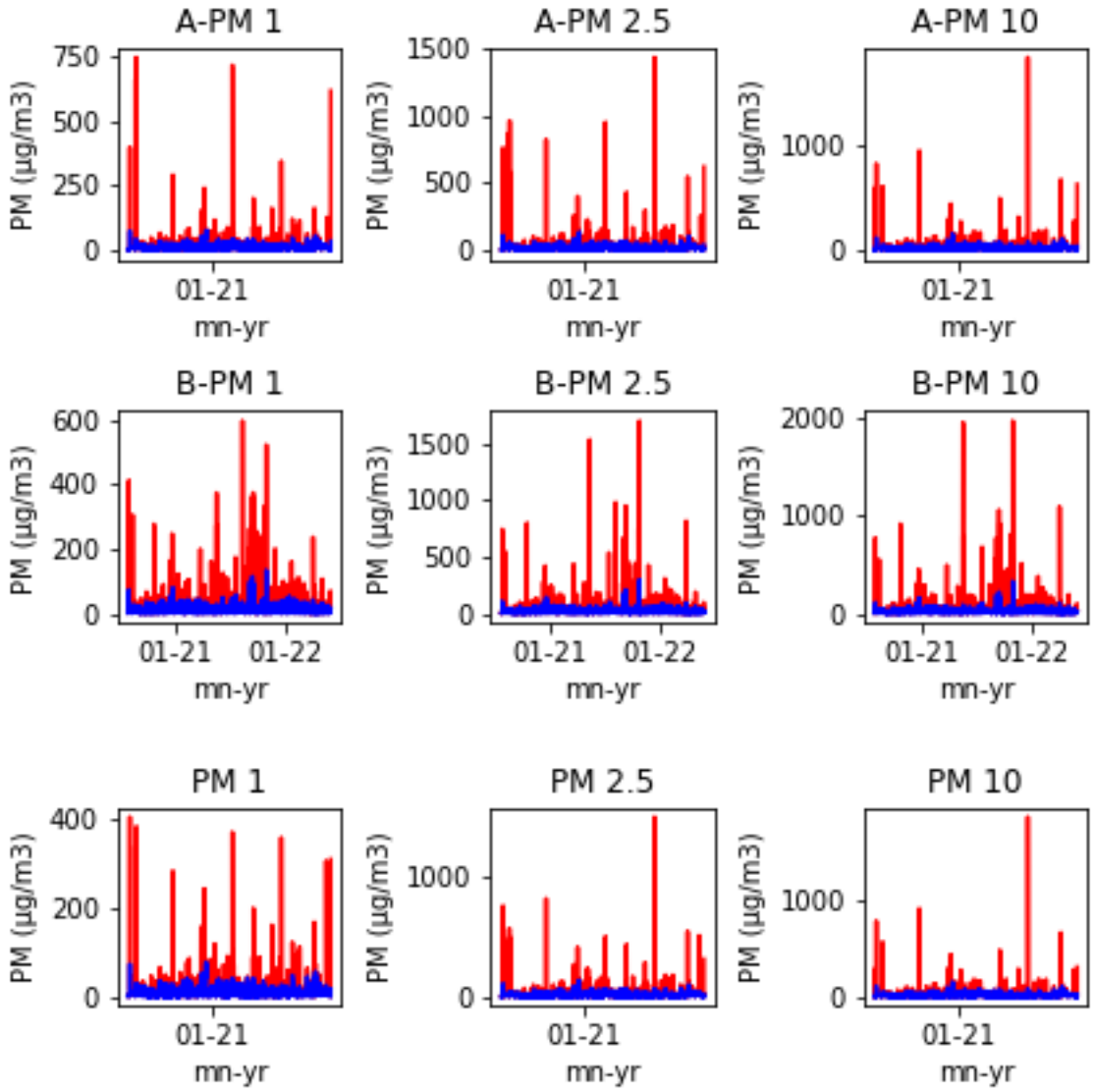
7. Gettysburg Central Campus



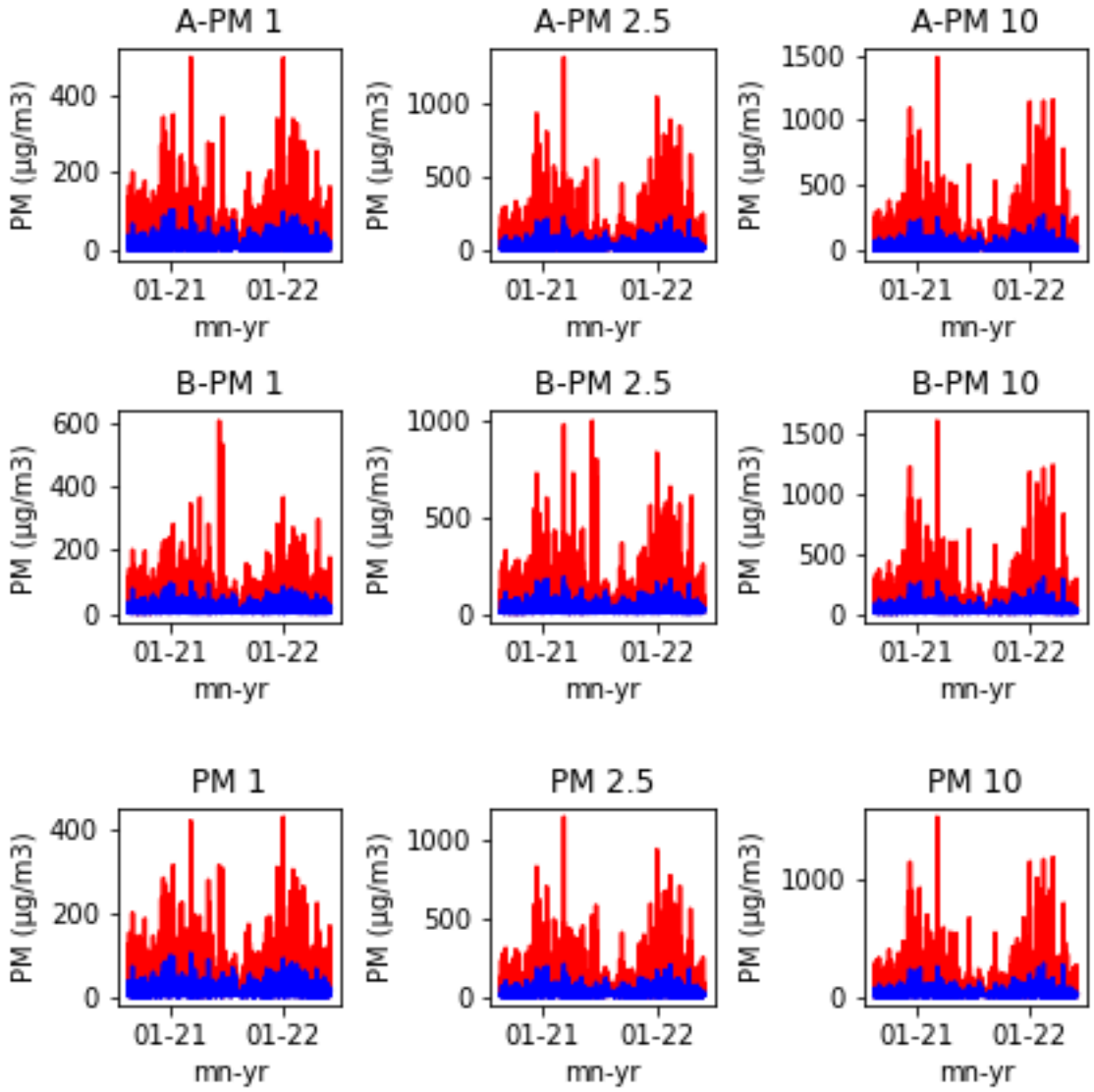
8. Gettysburg West Campus



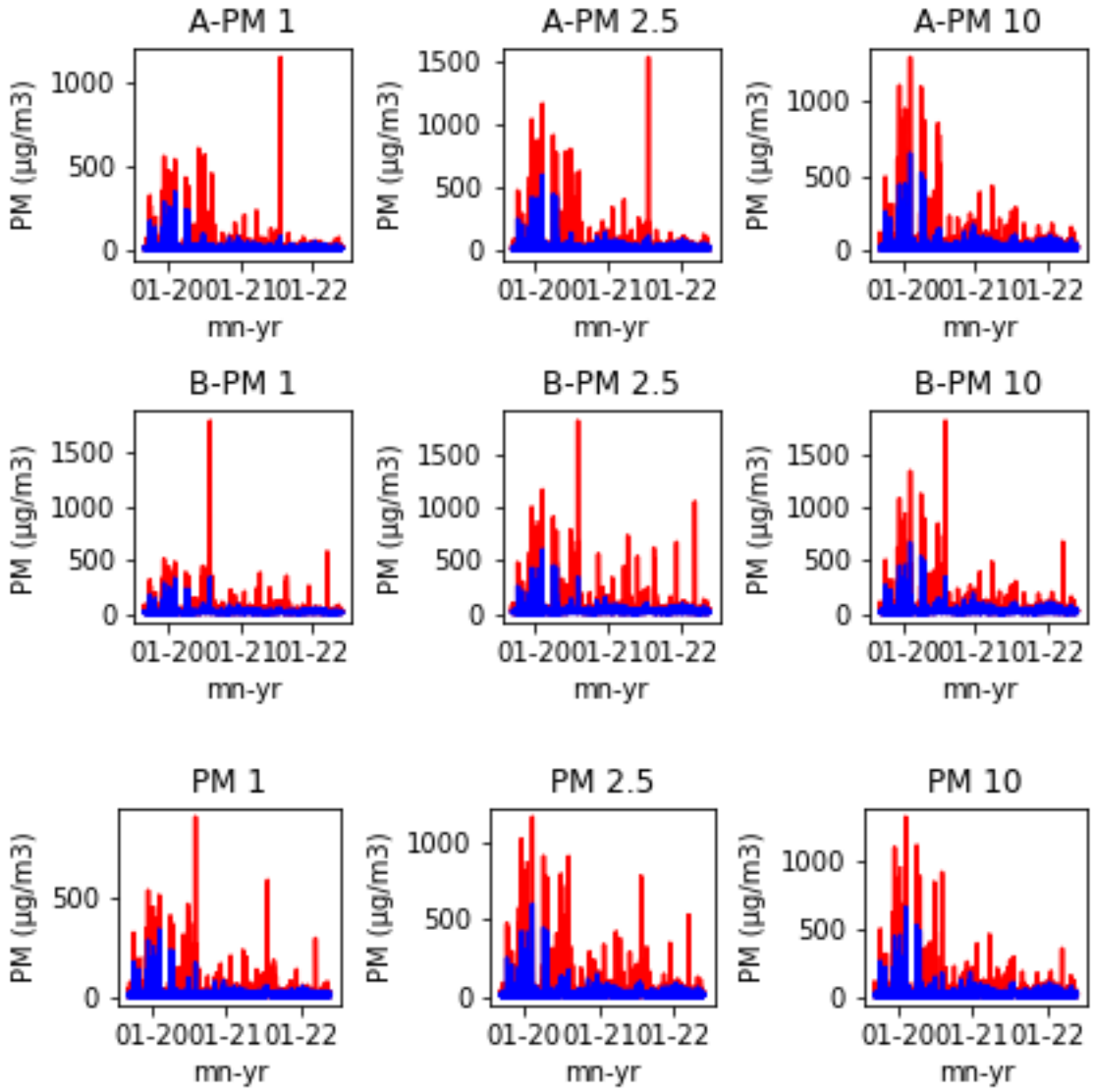
9. LeTort Spring Run Trail Head



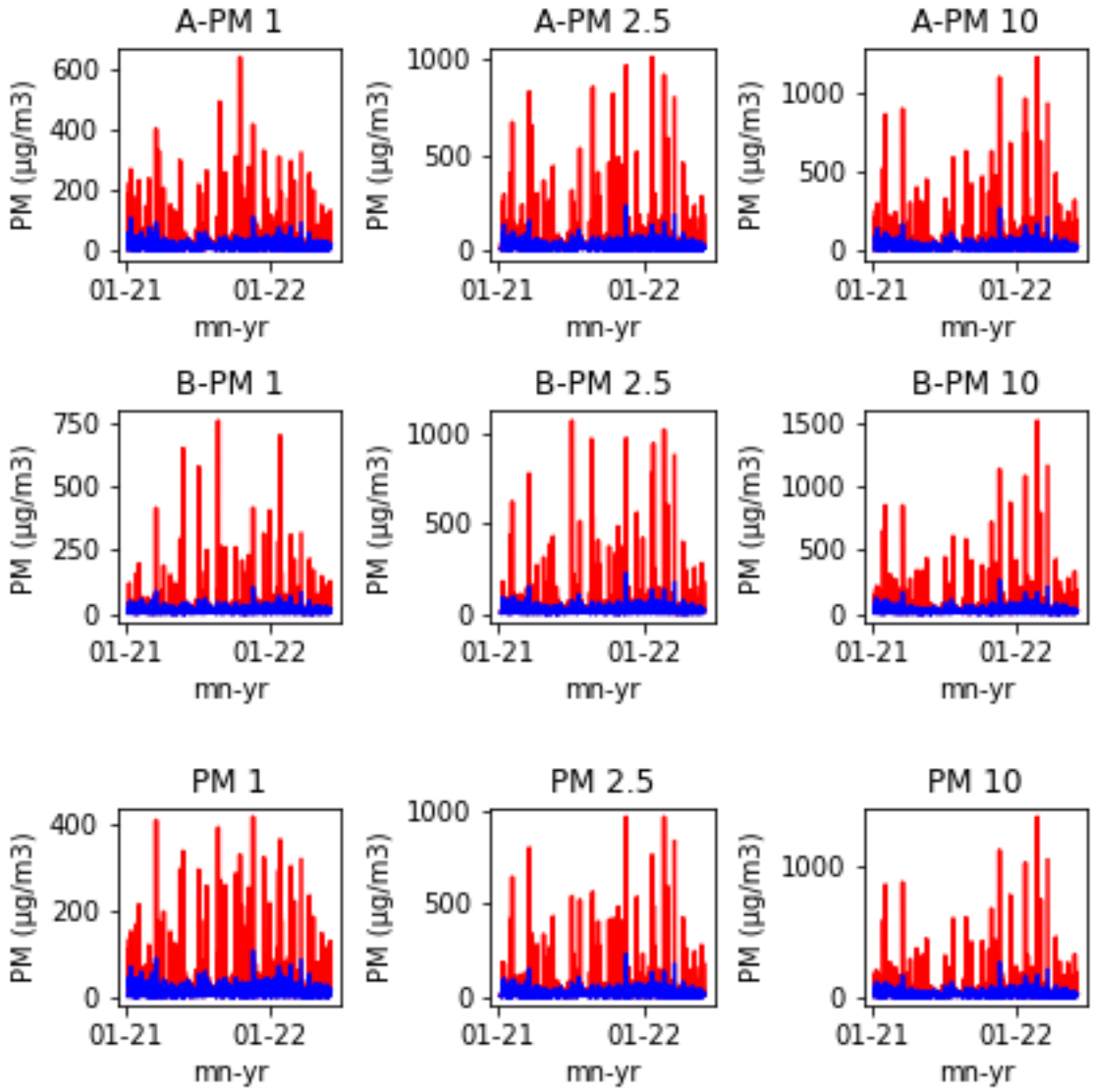
10. Meglio - Loysville



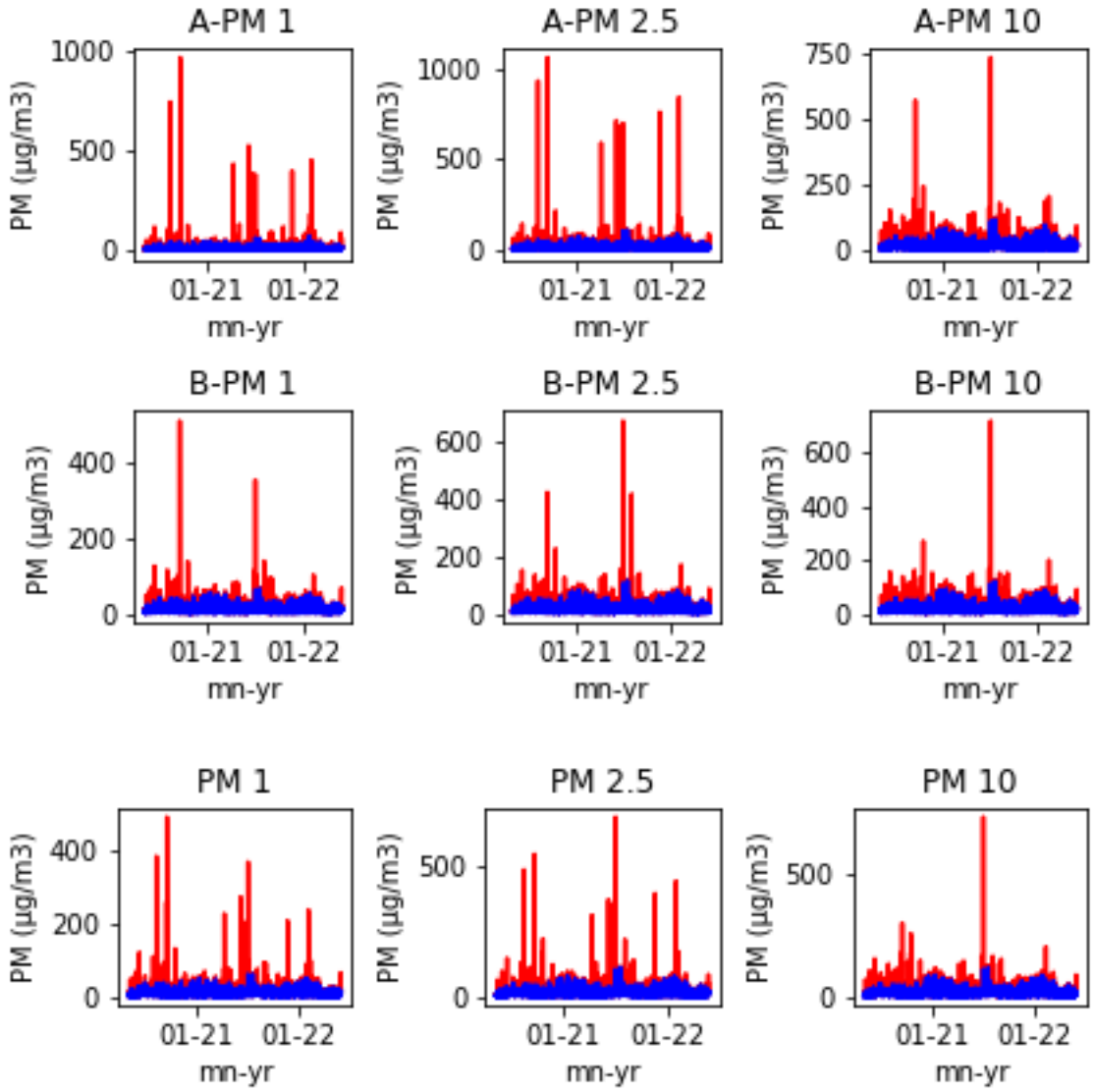
11. Middlesex Township



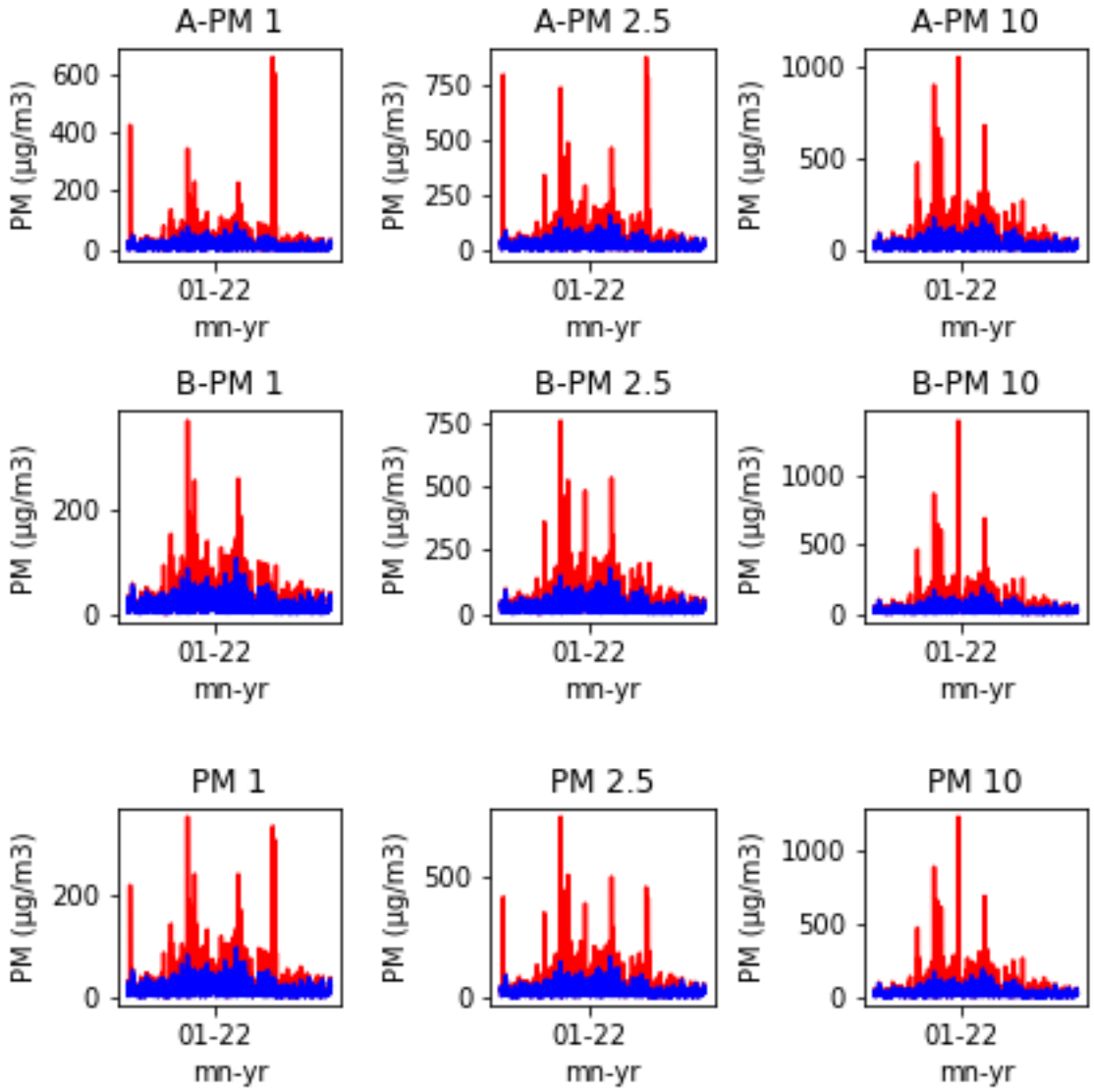
12. Middlesex Twp., Meadowbrook Farms



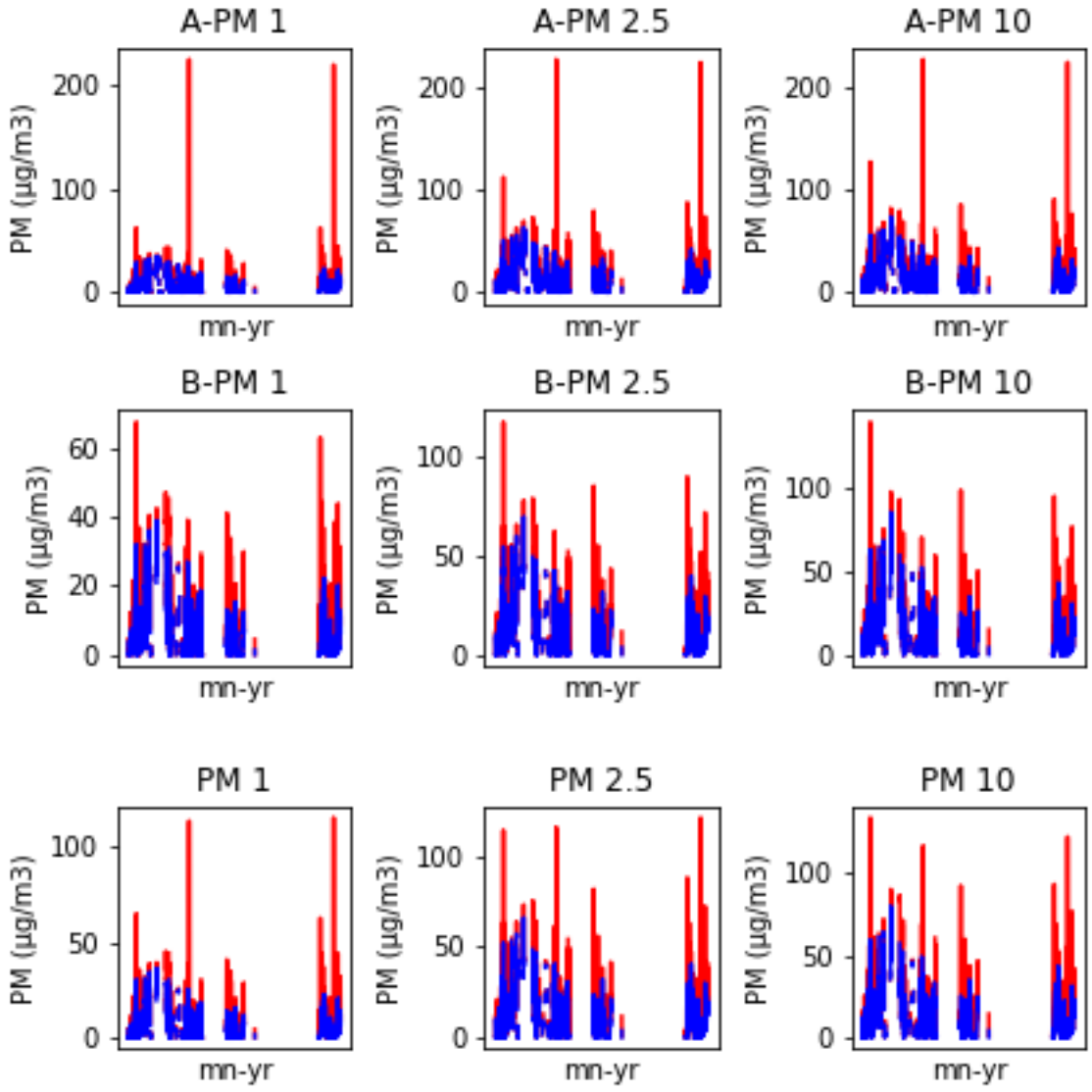
13. Mt Penn


A-PM 1
A-PM 2.5
A-PM 10
B-PM 1
B-PM 2.5
B-PM 10
PM 1
PM 2.5
PM 10
PM (µg/m3)
mn-yr
01-21
01-22

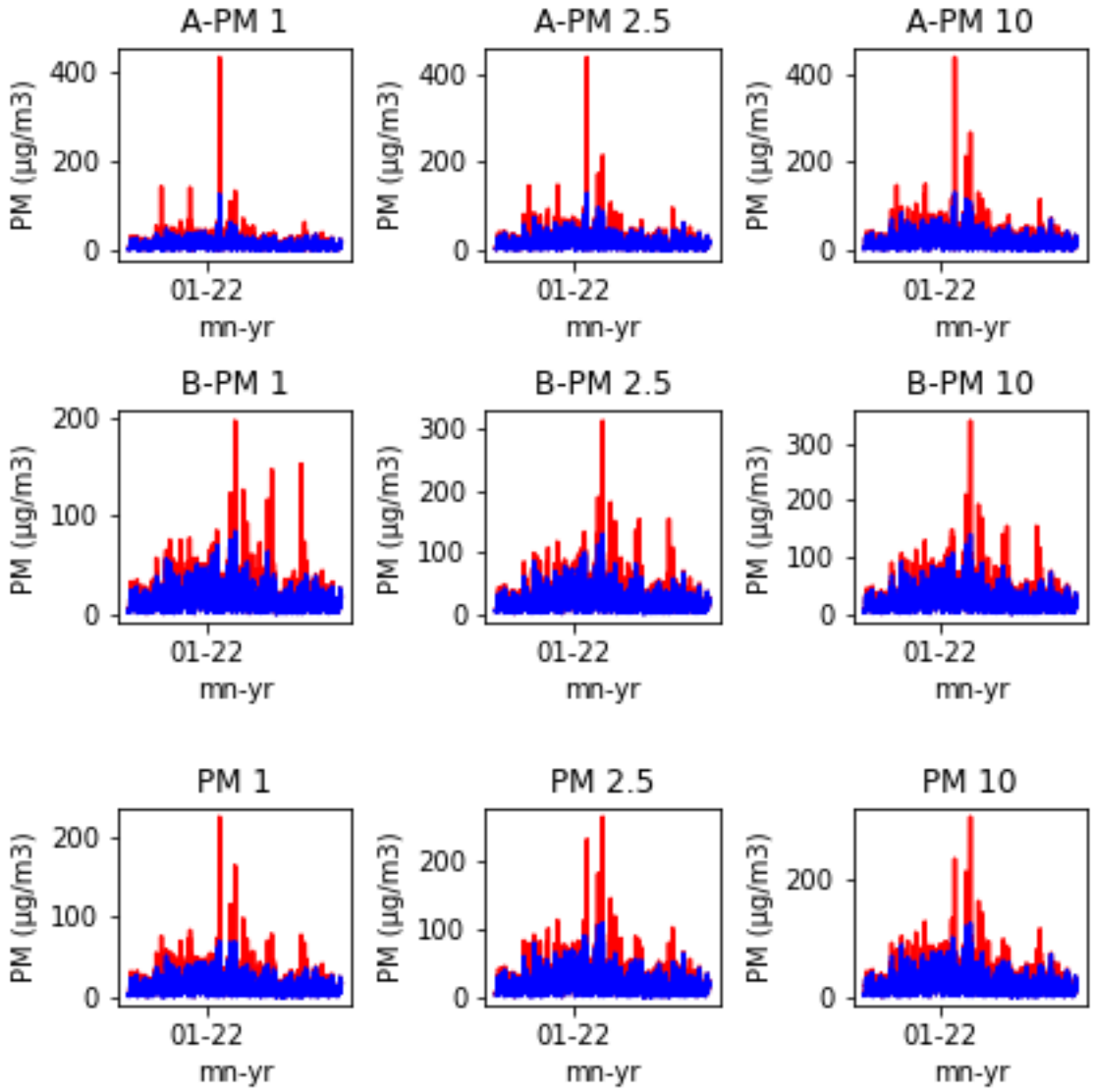
14. Murrell


A-PM 1
A-PM 2.5
A-PM 10
B-PM 1
B-PM 2.5
B-PM 10
PM 1
PM 2.5
PM 10
PM (µg/m3)
01-22
mn-yr

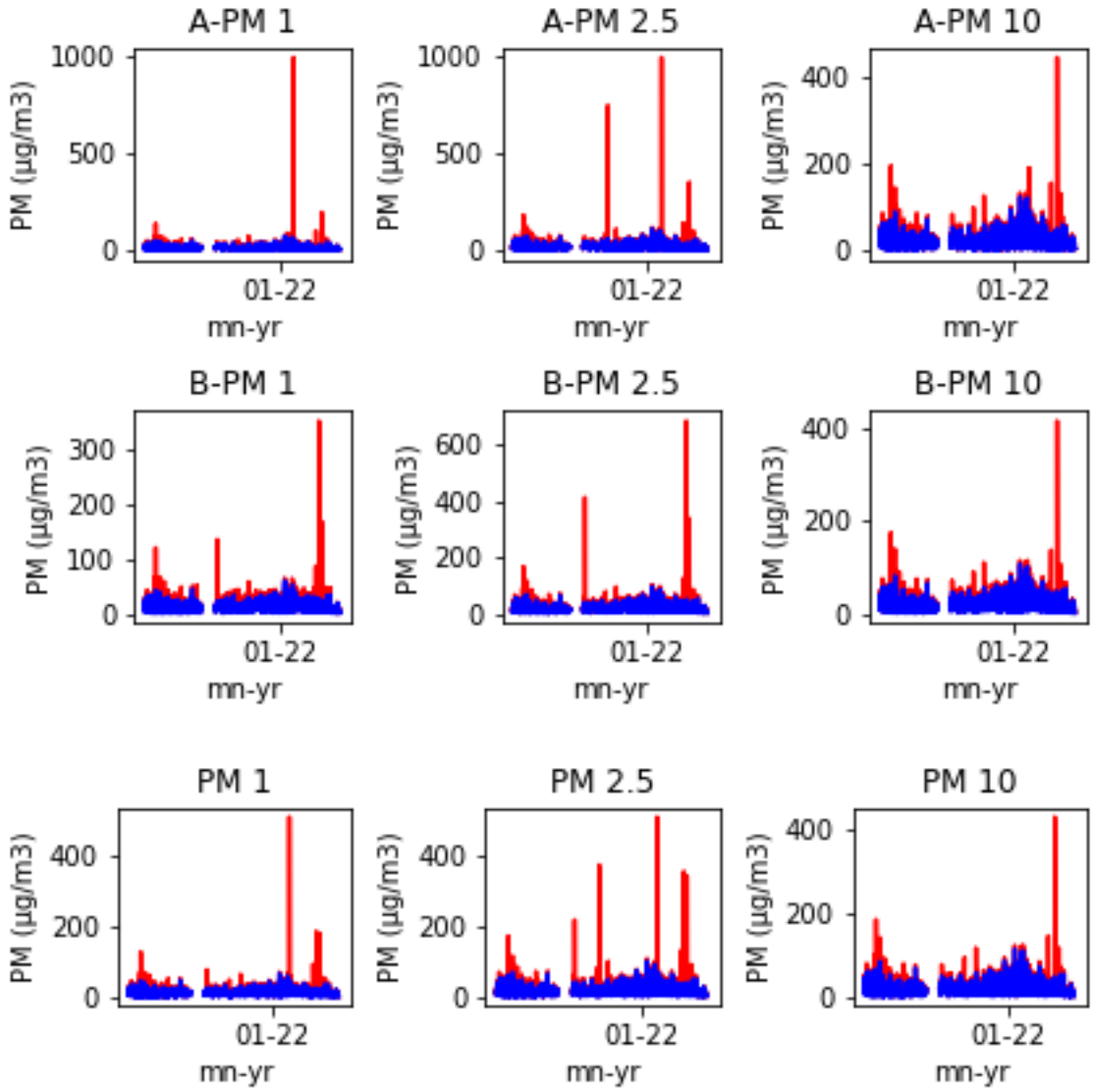
15. Shippensburg University - King Street



16. Shippensburg University - Shearer Hall



17. Susquehanna University



18. Wayne Heights - Old Forge Road