Cross-calibration and visualization of data for a portable air quality monitor

Olivier Kepler Francois[1] (olivierkfrancois1@gmail.com), Ruifeng Song[1], Dr. Qi Wang[1], Dr. Amy Mueller[1,2]
[1]Civil and Environmental Engineering, Northeastern University, [2]Marine and Environmental Sciences, Northeastern University

Motivation

Urban air quality is indeed a critical concern that profoundly impacts the well-being of city dwellers worldwide. The challenges of monitoring air pollution in complex urban landscapes, including the high costs involved and the dynamic nature of pollution sources, have been significant.

Despite these formidable obstacles, this project has been initiated with the primary goal of addressing these issues and instigating positive change. The central objective of this project is to develop and produce a sensor that is not only cost-effective but also incorporates advanced technologies. This sensor aims to revolutionize the monitoring of urban air quality by making it more accessible and efficient.

The portability of the sensor is one of its key features. Being portable allows for greater flexibility in its deployment, enabling monitoring in various locations and at different times. This mobility empowers citizens and decision-makers with valuable data on the spatial and temporal heterogeneity of air pollution, which can serve as a foundation for making informed decisions and taking action toward creating a healthier and more sustainable urban environment.

Objectives

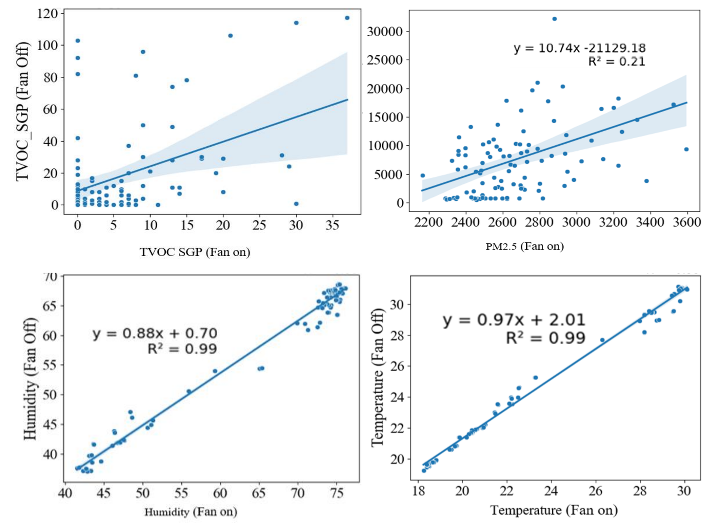
Conduct benchmarking of sensor data from both the sensor with a fan and the one without a fan by comparing it with data obtained from another reference sensor. This process will help evaluate their accuracy and identify any discrepancies between the readings.

Figure 1 Prototype sensor without a fan inside.

Figure 2 Prototype sensor without a fan inside.

Preliminary experiment

We conducted preliminary tests on the two sensors. These initial tests were performed before comparing them to another sensor, such as the TSI sensor, to evaluate their accuracy. During the comparison, we observed that both sensors exhibited greater regularity in the temperature and humidity data readings compared to the PM2.5 and TVOC readings.



Methods

For data collection, we traveled to various locations with different levels of concentration to gather the necessary data. The experiments were conducted at three main sites: the Northeastern courtyard, in front of Egan's lab, and at Northeastern and Ruggles stations. We ensured meticulous data analysis using Python codes and employed effective comparison methods to ensure the reliability of our results. Our sensors collected various data points, including Particulate Matter (PM) levels, Volatile Organic Compounds (VOCs) in the air, temperature, and humidity. Additionally, we integrated a GPS system to accurately track the exact locations during the experimental sessions.


Figure 7 In front of the Egan Lab

The collection of data for comparison

We collected data from different locations using two sensors and the TSI sensor. The purpose was to compare the data from the two sensors with that of the TSI sensor to assess their accuracy.

The TSI sensor is used as a reference for conducting comparison tests to evaluate the accuracy of the other two types of sensors.


Figure 3 TSI Sensor: Detecting Particulate Matter (PM)

The first sensor comparison experiment took place at Northeastern University, next to the Egan building. Data was collected for over an hour and subsequently analyzed in the lab.


Figure 4 Talking while collecting data at N.U.

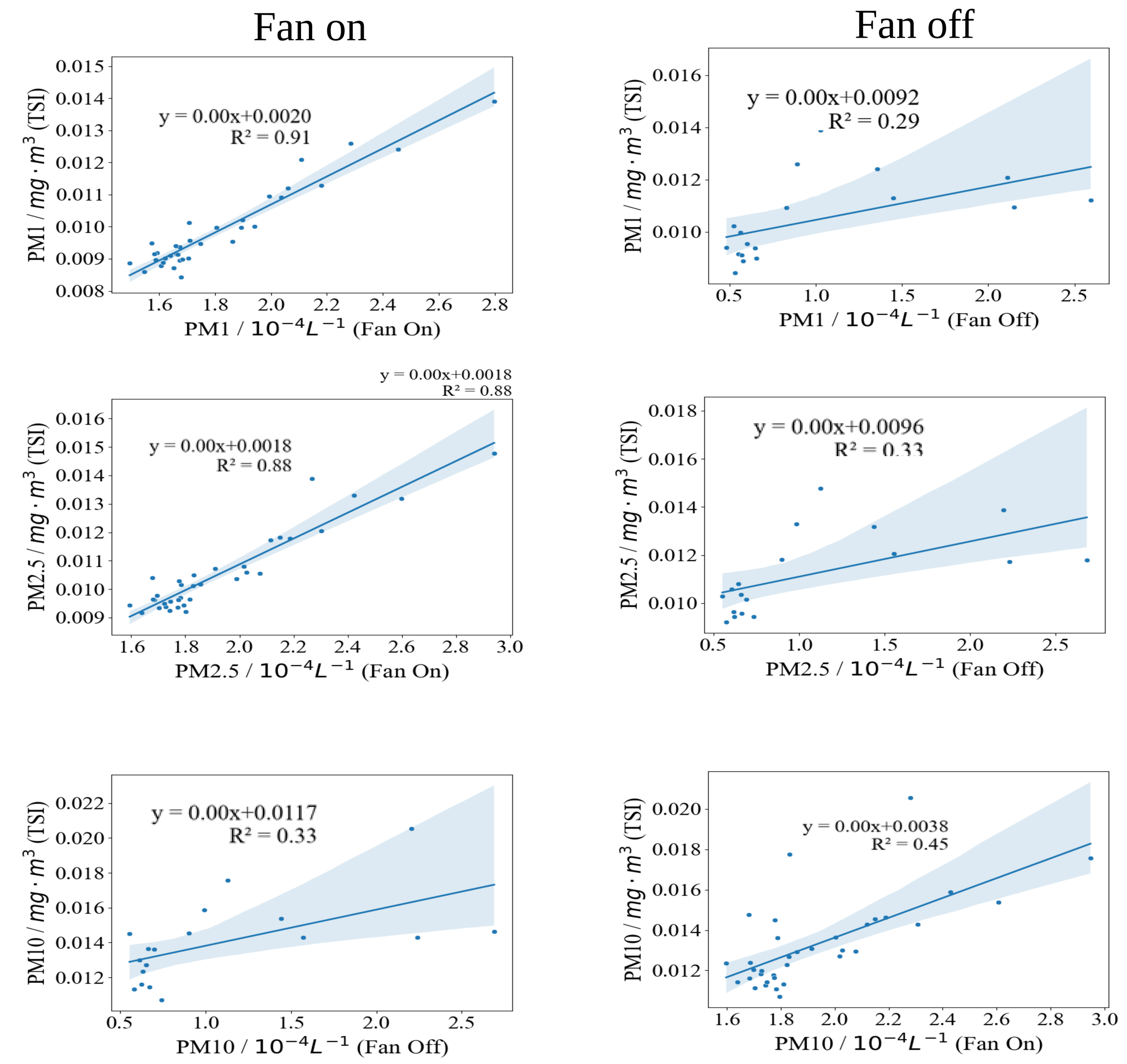
The second sensor comparison experiment was conducted at Northeastern Station, where data was also collected for more than an hour and later analyzed in the laboratory.


Figure 5 Talking while collecting data at N.U. station

The third sensor comparison experiment was at Ruggles Station. We collected data for over an hour, just like in the first two previous experiments, and then analyzed it in the lab.


Figure 6 Talking while collecting data at Ruggles station

Once compiled, these data were subjected to comparison tests using efficient comparison and visualization methods. Graphical representations of the test results are provided bellow.



Results

- The preliminary experiment aimed at evaluating two sensors: the one containing a fan and the other without a fan. The results demonstrated that the sensor with the fan is more accurate, particularly in the collection of Particulate Matter (PM) and volatile organic compounds (VOC).
- When comparing both sensors with the TSI sensor, the results show that the R^2 value is greater when the fan is on. The coefficient of determination, or R^2, measures how well the data fit the fitted regression line. This suggests that the lunch box with the fan is more accurate than the one without the fan when comparing their performance to the TSI sensor.

Challenges

- Finding suitable locations with varying concentrations to conduct experiments proved to be a challenging task. We needed to locate areas with different concentration levels, including those with low, medium, and high concentrations.
- Data collection was challenging due to irregularities and missing data. The data collected by the two sensors previously exhibited small irregularities, which prompted us to recalibrate them.
- Analyzing the data for comparison using Python also posed challenges that demanded careful thought and problem-solving.

Next steps

- As we conduct experiments in locations with low and medium pollutant concentrations, we must also do the same in locations with high concentrations of pollutants using different sensors for comparison purposes.
- Explore methods to improve accuracy by changing the placement of sensor modules and adjusting the location of the fan.
- Conduct similar comparison experiments for Volatile Organic Compounds (VOCs) as we previously conducted for Particulate Matter (PM). We have initiated some preliminary experiments with the VOCS sensor that we will utilize for comparison purposes later.

Acknowledgment

Supported by the Research Experiences for Undergraduates (REU)

National Science Foundation (NSF) Smart and Connected Communities grant #2125326.

And the support from to Dr. Amy Mueller and Ruifeng Song