Understanding the source components captured by the Purple Air Network

Vijay Kumar[1], *S. Dinushani Senarathna[1], Supraja Gurajala[4], William Olsen[5], Shantanu Sur[3], Sumona Mondal[1], Suresh Dhaniyala[2]*

[1]Department of Mathematics, Clarkson University
[2]Department of Mechanical and Aeronautical Engineering, Clarkson University
[3]Department of Biology, Clarkson University
[4]Department of Computer Science, State University of New York, Potsdam
[5]Department of Civil and Environmental Engineering, Clarkson University

$PM_{2.5}$ has been linked to numerous pollution-mediated adverse health effects and their monitoring is key for taking preventative and mitigative measures. Accurate measurements of $PM_{2.5}$ concentrations are available at EPA sites, but such data lacks spatial resolution due to a limited number of monitoring locations. In recent years the deployment of low-cost sensor networks has opened up the possibility of acquiring air quality data at a high spatiotemporal resolution. However, the sensitivity, noise, and accuracy of data acquired by low-cost sensors remain a concern. Here, we studied $PM_{2.5}$ measurements made from EPA and Purple Air (PA) sensor networks in the Chicago area to understand the parameters influencing the performance characteristics of the low-cost sensor network. We decomposed the $PM_{2.5}$ time series data into short-term, seasonal, and long-term components using the Kolmogorov–Zurbenko (KZ) filter. We then extracted different frequency signals of $PM_{2.5}$ data for each of the filtered components. The PA sensor data accurately capture long-term variations in the $PM_{2.5}$ data but not short-term variations (<24 hours). We hypothesize that the sensor networks may have different sensitivity to aerosol from different sources and hence care must be taken in their use for evaluating the impact of air quality mitigation policies. Therefore, calibrations should be performed in short-term, seasonal, and long-term components of $PM_{2.5}$ measurements captured by from low-cost sensors using reference monitors before using them into epidemiological studies.