Spectral analysis of low-cost sensor network data

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

[1]Department of Mathematics, [2]Department of Computer Science, [3]Department of Mechanical and Aerospace Engineering, [4]Department of Biology, [5]Department of Civil and Environmental Engineering, Clarkson University and Department of Computer Science, SUNY Potsdam

Background

Motivations:

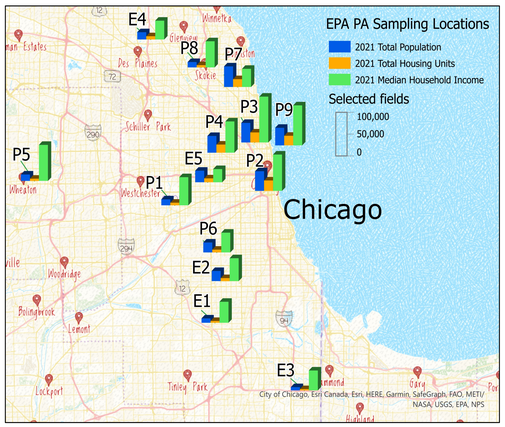
- Low-cost air sensor data increasingly used in health studies
- Sensitivity, noise, and accuracy of data of sensors remain a concern
- Differences exist between data from EPA and low-cost sensors
- What is driving these differences, and can that be ”calibrated” out?

Figure 1: Purple Air (PA) sensor

Research objectives:

- Analyze the temporal differences in $PM_{2.5}$ data from EPA and purple air (PA) low-cost sensor network
- Characterize the influence of different aerosol sources on the EPA and PA data

Methodology

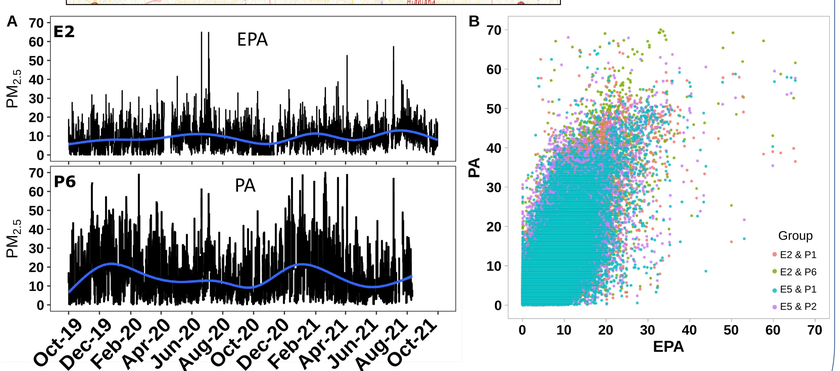


Data source:

EPA and Purple air (PA) data;

Cook county, IL;

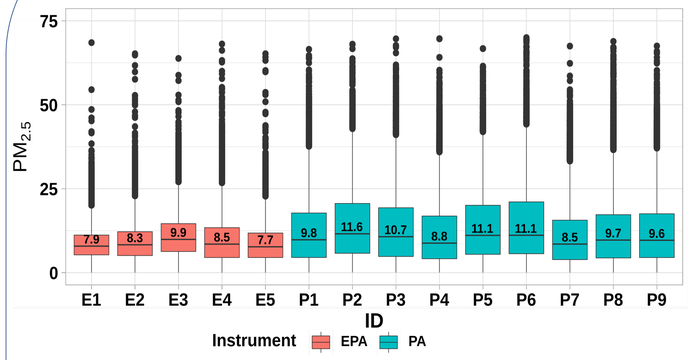
October 2019 to September 2021



(A) Time series data of selected EPA & PA sites for the study duration (B) Scatter plot showing relation between selected EPA & PA sites in close vicinity

Results

Monitoring data:

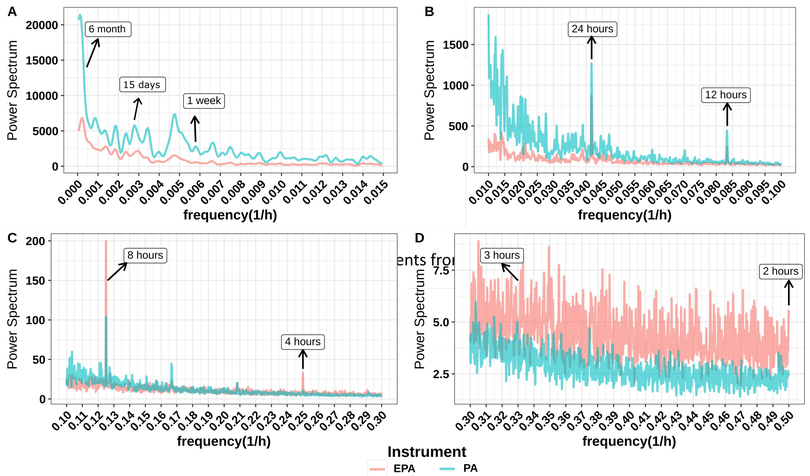


Distribution of EPA and PA $PM_{2.5}$ over the analysis time period

- PA data on average higher median values than EPA for all sites
- PA data has a broader distribution

Spectral analysis:

- Obtained power spectrum of the two data sources EPA and PA; For simplicity just averaged the data from the different sites for each of the sources

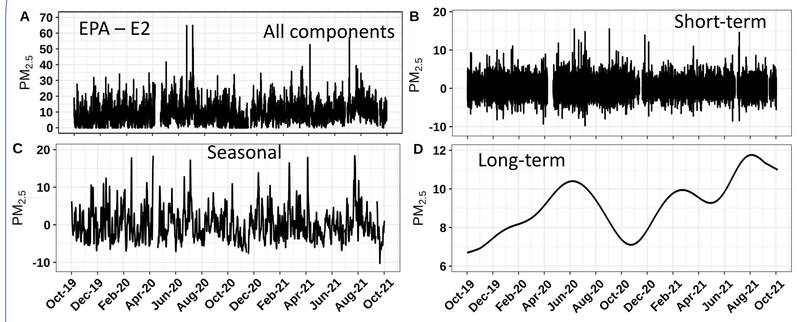


- Identified that there are selective frequencies (e.g., at 4,8, 12, 24 hours); that influence both data sets possibly corresponding to traffic patterns, local industrial emission, etc.
- Power spectrum of the PA and EPA different at two ends of spectrum
  - PA has a higher relative power spectrum at low frequency (i.e., for time periods >1 day). Possibly suggesting an over contribution of regional aerosol to the signal
  - EPA has a higher power spectrum at short time periods – probably indicating its relatively better sensitivity to local sources e.g., traffic

Results

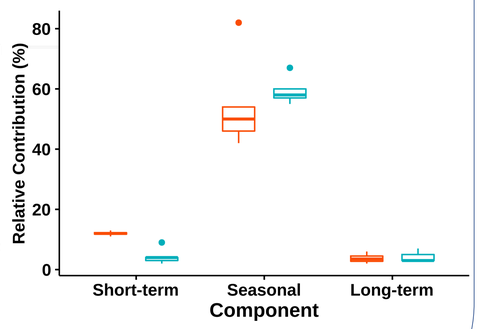
Relative contributions of temporal components of $PM_{2.5}$

Time series decomposition using Kolmogorov–Zurbenko (KZ) filter



EPA E2 site $PM_{2.5}$ time series (A), decomposition of EPA site E2 $PM_{2.5}$ time series into: (B) Short-term component (C) Seasonal component (D) Long-term component

- Decomposed signal from the data sets into short term (< 24 hrs), seasonal (< 60 days), and long-term.
- EPA Short-term component contribution to total signal is 3 times that of PA – suggesting significant loss of sensitivity to fast sources such as traffic
- Long-term components have similar contribution



Relative contribution of temporal components of $PM_{2.5}$

Conclusion

- Lower accuracy of PA data relative to EPA data has been previously identified, but reasons are uncertain
- Most correction models based on regression; Models will not be valid if the distribution of sources change
- Our study shows PA data comparable in efficiency to EPA for long-term variations in the PM2.5 data but less efficient in detecting short term variations (<12 hours) suggesting strong source dependence in sensitivity
- Models based on frequency-decomposition of data could enable corrections accounting for different sensitivity to different aerosol sources

References

[1] Kuula, J., Mäkelä, T., Aurela, M., Teinilä, K., Varjonen, S., González, Ó., and Timonen, H.: Laboratory evaluation of particle-size selectivity of optical low-cost particulate matter sensors, Atmospheric Measurement Techniques, 13, 2413–2423, 2020, https://doi.org/10.5194/amt-13-2413-2020

[2] Tryner, J., Mehaffy, J., Miller-Lionberg, D., and Volckens, J.: Effects of aerosol type and simulated aging on performance of low- cost PM sensors, Journal of Aerosol Science, 150, 105 654, 2020, https://doi.org/10.1016/j.jaerosci.2020.105654

[3] Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., and Bartonova, A.: Can commercial low-cost sensor, platforms contribute to air quality monitoring and exposure estimates?, Environment international, 99, 293–302, 2017, https://doi.org/10.1016/j.envint.2016.12.007

[4] Tchepel, O. and Borrego, C.: Frequency analysis of air quality time series for traffic related pollutants, Journal of Environmental Monitoring, 12, 544–550, 2010, https://doi.org/10.1039/B906717E

[5] Zhang, Z., Kim, S.-J., and Ma, Z.: Significant decrease of PM2. 5 in Beijing based on long-term records and Kolmogorov-Zurbenko filter approach, 2018, https://doi.org/10.4209/aaqr.2017.01.0011

[6] Bai, H., Gao, W., Zhang, Y., and Wang, L.: Assessment of health benefit of PM2. 5 reduction during COVID-19 lockdown in China and separating contributions from anthropogenic emissions and meteorology, Journal of Environmental Sciences, 115, 422–431, 2022, https://doi.org/10.1016/j.jes.2021.01.022