Spatiotemporal Analysis of $PM_{2.5}$ in Chicago using Data from EPA and Low-Cost Sensor Network

Vijay Kumar

Department of Mathematics, Clarkson University, Potsdam

Date: October 20, 2021

Co-authors:

Prof. Suresh Dhaniyala, Clarkson University

Prof. Supraja Gurajala, SUNY Potsdam

Prof. Shantanu Sur Clarkson University

Prof. Sumona Mondal, Clarkson University

Dinushani Senarathna, Clarkson University

CLARKSON UNIVERSITY 18 · TECHNOLOGIA · 96 ™

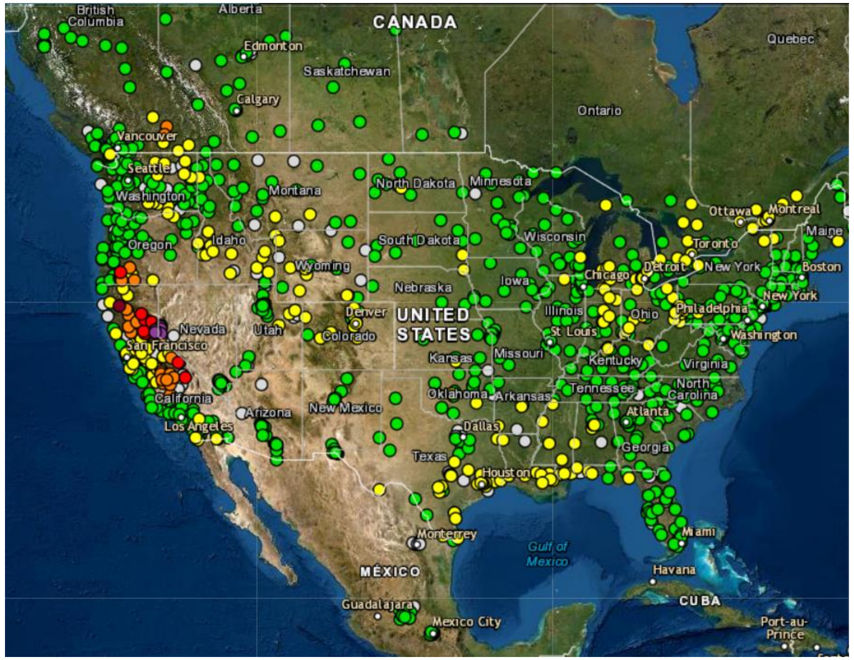
AIR LAB

AEROSOL INSTRUMENTAL RESEARCH LAB

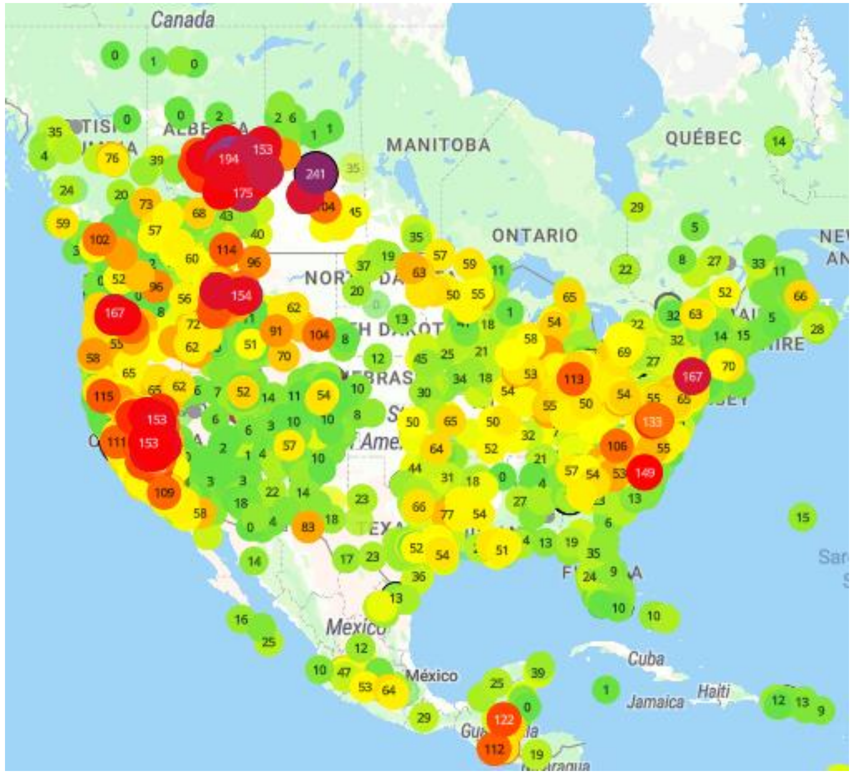
Introduction

- Air quality across the US is monitored through **EPA** sites distributed across the country
- Spatial resolution of air quality data is restricted by the limited number EPA monitoring sites
- Low-cost sensor networks (e.g., Purple air, **PA**) are considered an attractive alternative
- Sensitivity, noise and accuracy of data acquired by low-cost sensors remain to be a concern
- Ongoing effort to build models to correct low-cost sensor data using EPA data as reference
- Need for better understanding of the differences between air quality data acquired by EPA and low-cost sensors

EPA network



PA network



https://map.purpleair.com/1/mAQI/a10/p604800/cC0#2.33/42.94/-98.5

https://gispub.epa.gov/airnow/

Motivation and Objectives

- The different detection and sampling approaches of EPA and PA instruments are likely to impact quality of collecting the $PM_{2.5}$ data
- $PM_{2.5}$ in the air is influenced by multiple factors that vary widely in their time-course
- Literature studies have shown that the two measurements differ, but the reasons for differences are not fully established

Objectives:

- Investigate the differences in temporal course of $PM_{2.5}$ data from EPA and PA using spectral analysis
- Characterize the reason for differences between EPA and PA measurements
- Study the influence of meteorological parameters on the EPA and PA data

PA sensor

EPA instrument

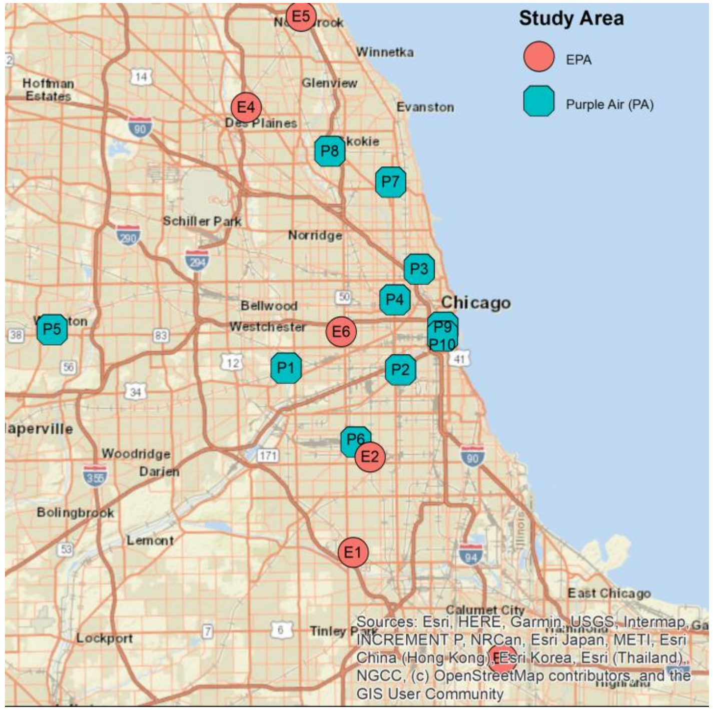
http://www.bishoptribeemo.com/ambient.htm

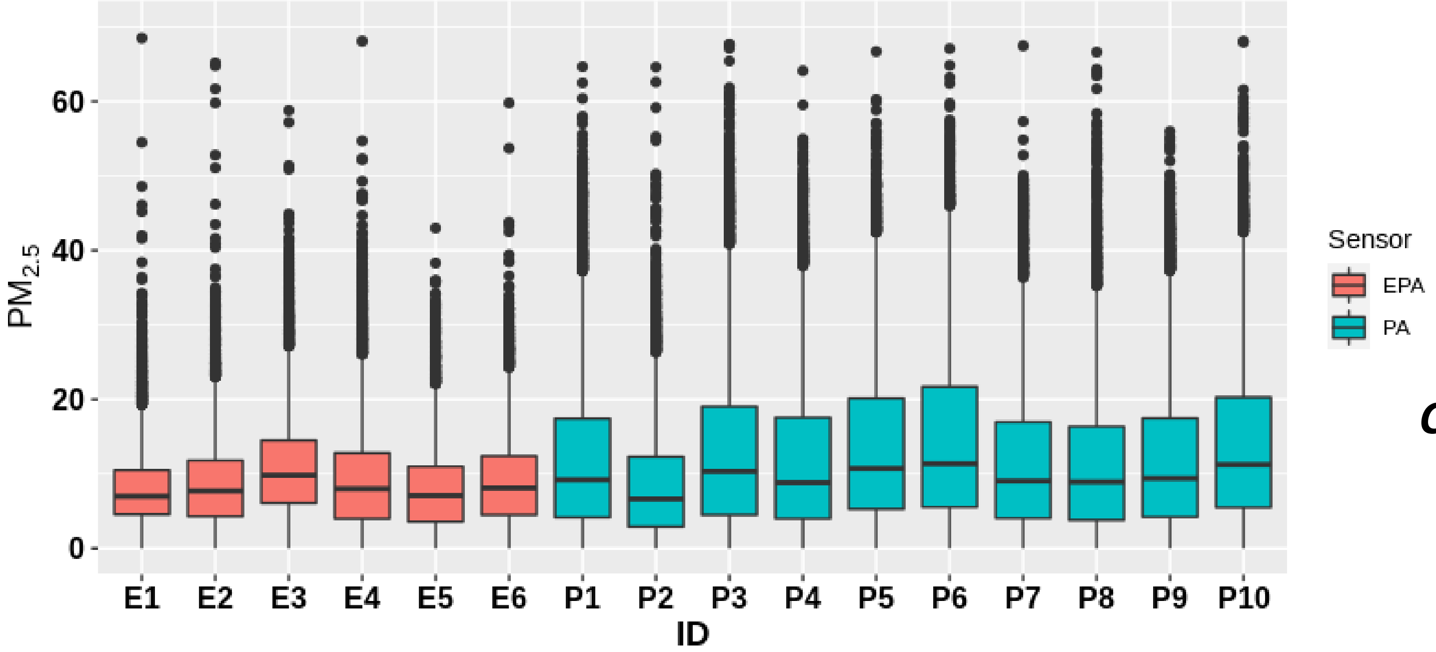
https://www2.purpleair.com/collections/air-quality-sensors/products/purpleair-pa-ii-sd

Study Location: Cook County, Chicago

- $PM_{2.5}$ data collected from EPA and PA network from Sep 2019 to Dec 2020
- Sites with >30% missing observations were excluded from the study
- Data from 6 EPA and 10 PA sensors were considered in the final analysis
- $PM_{2.5}$ data range was set as [0,70] μg/m³ to exclude outliers

EPA & PA sites in Cook County

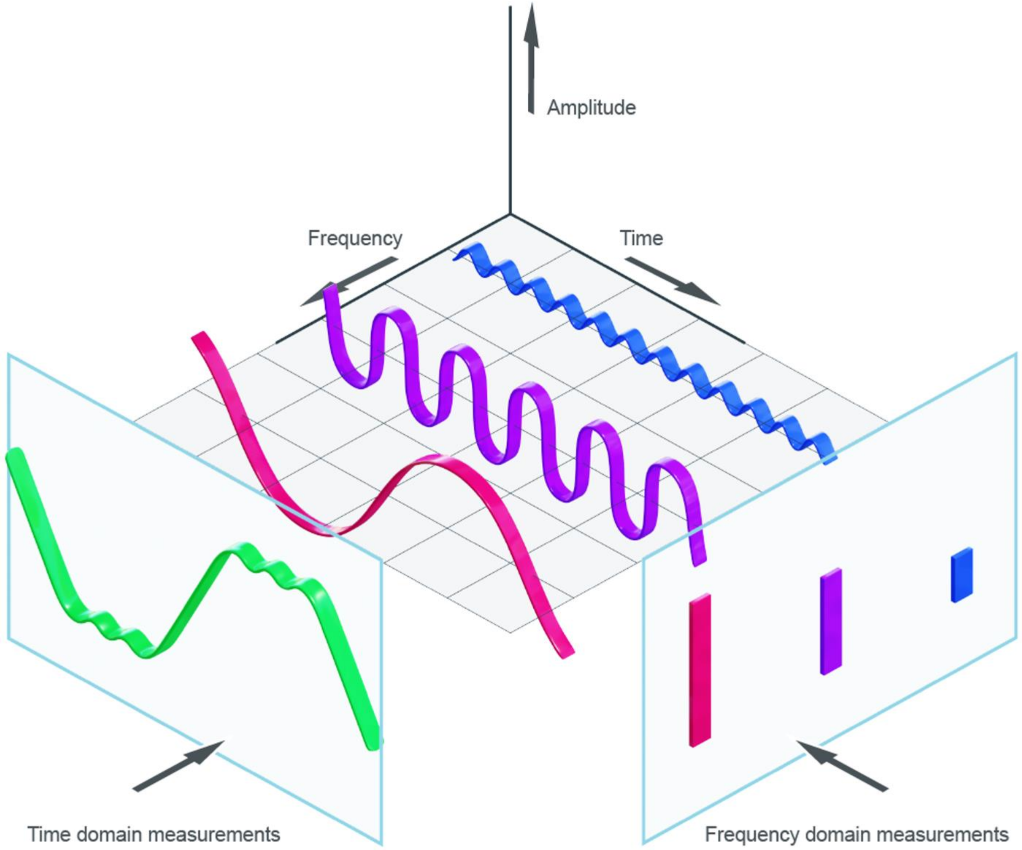




Comparison of $PM_{2.5}$ data from EPA & PA

Frequency-based Approach to Analyze $PM_{2.5}$ Data

- Frequency domain analysis provides a better understanding of periodic behaviors
- A broad spectrum of frequencies are found in the time series data of air pollutants
- Short-term fluctuations of $PM_{2.5}$ are related to local scale phenomena
- Seasonal changes and long-range transport of the particulate matter will contribute to long-term fluctuations
- Decomposition data into long-term, seasonal, and short-term fluctuations would allow the contributions from different sources[4]



https://www.ifm.com/us/en/us/real-time-maintenance/vibration/technology/frequency-domain

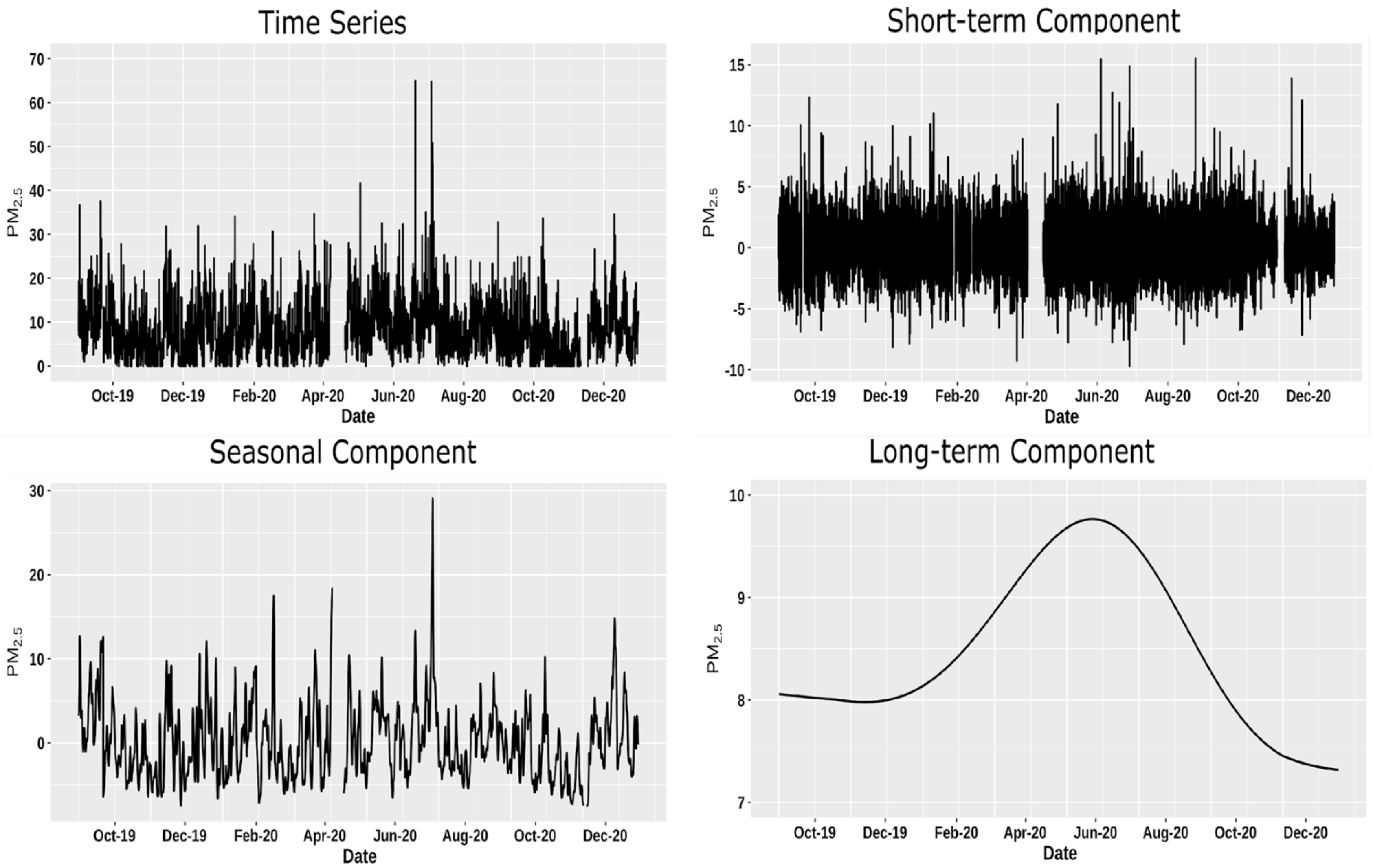
Decomposition of Data: Kolmogorov–Zurbenko (KZ) Filter

- The time series $A(t)$ can be decomposed into three components

$$A(t) = W(t) + S(t) + e(t)$$

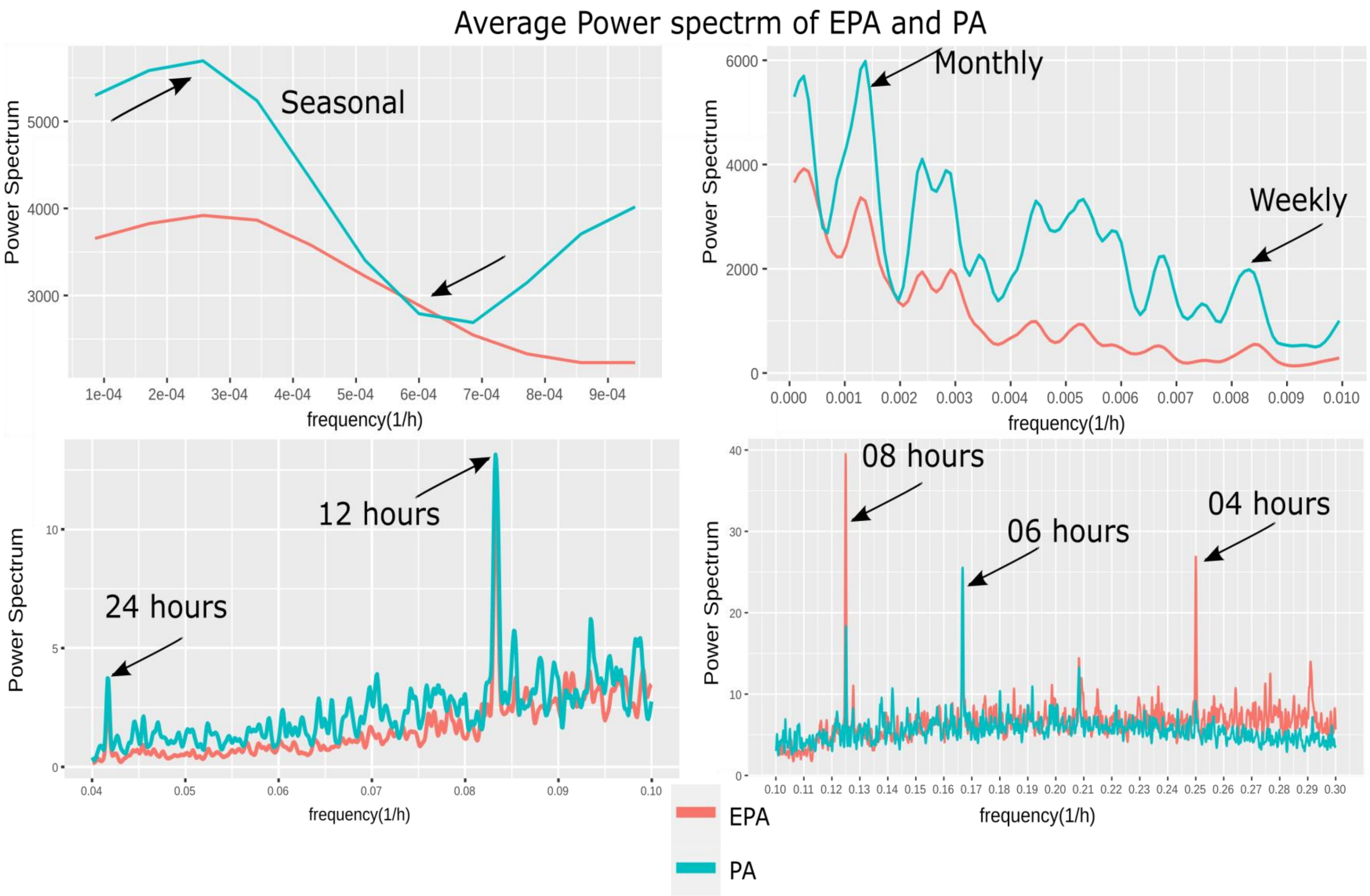
- The short time component is $W(t)$, seasonal $S(t)$, and long term $e(t)$

Decomposition of $PM_{2.5}$ data from EPA site E2

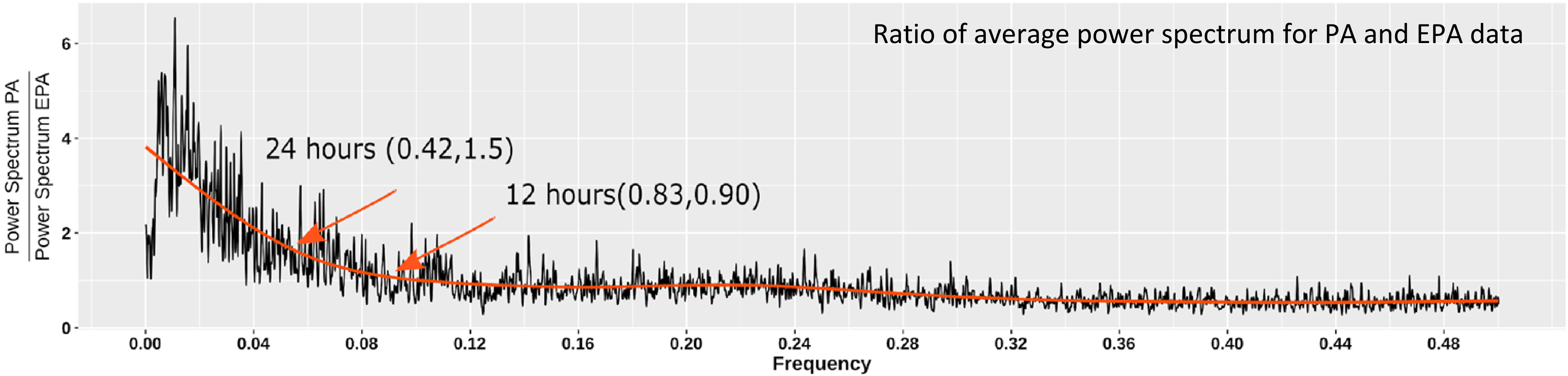


Comparison of EPA and PA Power Spectra

- Power spectrum of EPA and PA data from each site were averaged and plotted
- The spectra were divided into different frequency range for easier comparison
- Distinct peaks for variations occurring in monthly, weekly, daily, diurnal, and shorter time-scale could be identified
- EPA sites show larger power spectral peaks for variations in the time-scale of **8 hours** and less



Frequency Dependence of PA-EPA Power Spectral Ratio



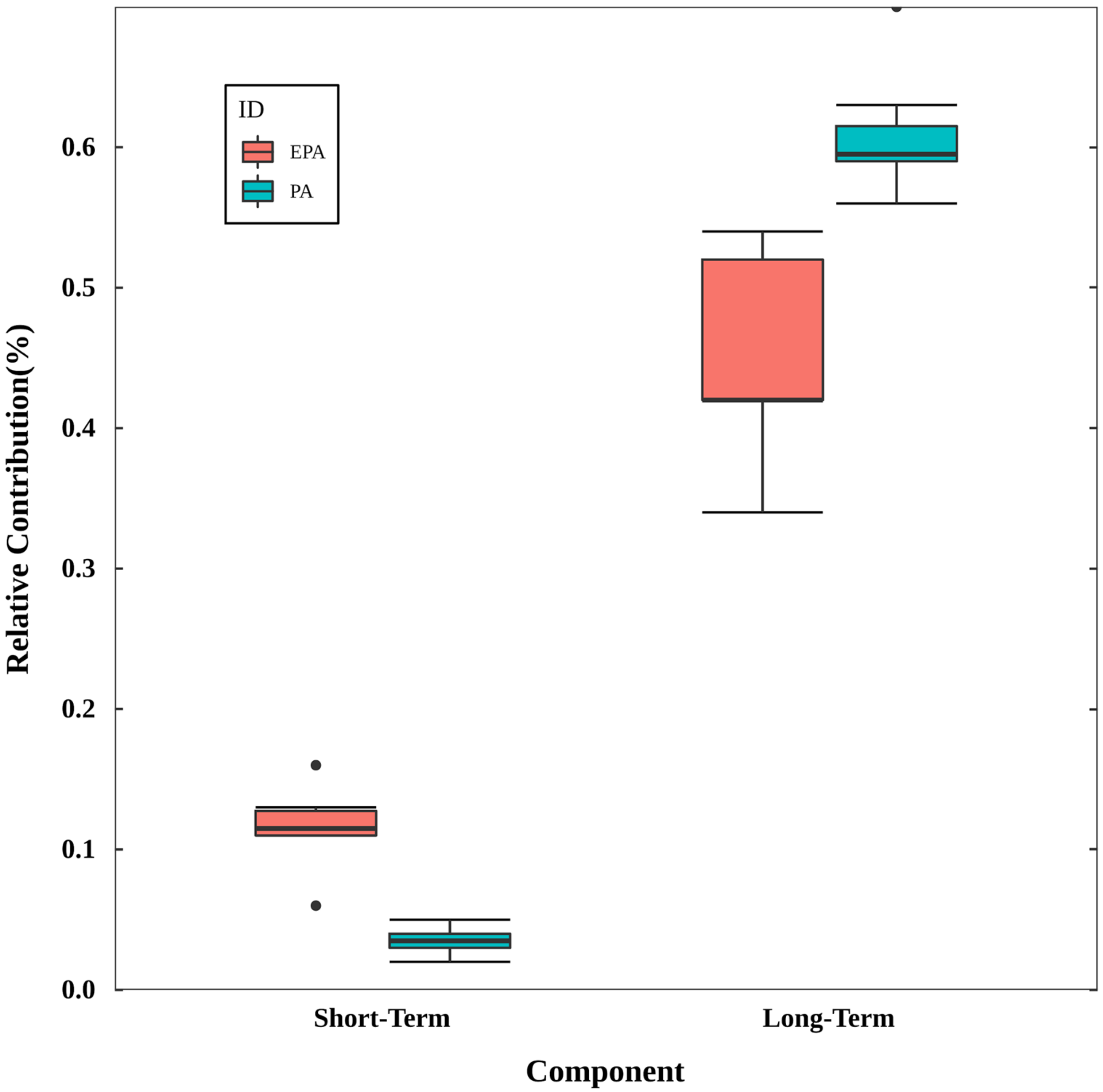
- PA/EPA power spectrum ratio drops significantly at higher frequencies corresponding to diurnal and faster variations
- The finding suggests PA sensors seem to be less sensitive to the local temporal sources such as traffic

Relative Contribution of Temporal Components on $PM_{2.5}$ Data

The relative contribution of temporal components

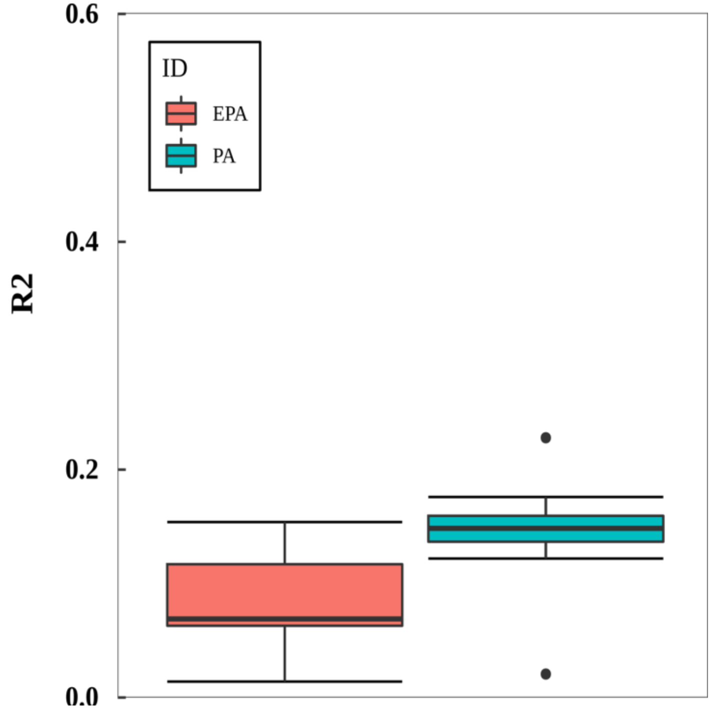
$$\boldsymbol{i}(\boldsymbol{t})) = \frac{\boldsymbol{Var}(\boldsymbol{i}(\boldsymbol{t}))}{\boldsymbol{Var}(\boldsymbol{A}(\boldsymbol{t}))}$$

- Short-term and long-term temporal components are defined as variations in the time-scale of 24 hours and less and seasonal respectively
- Short term EPA data has a higher relative contribution for short term fluctuations (~15%) compared with PA data (<1%)
- PA data is comparable to EPA data in capturing the long-term temporal components

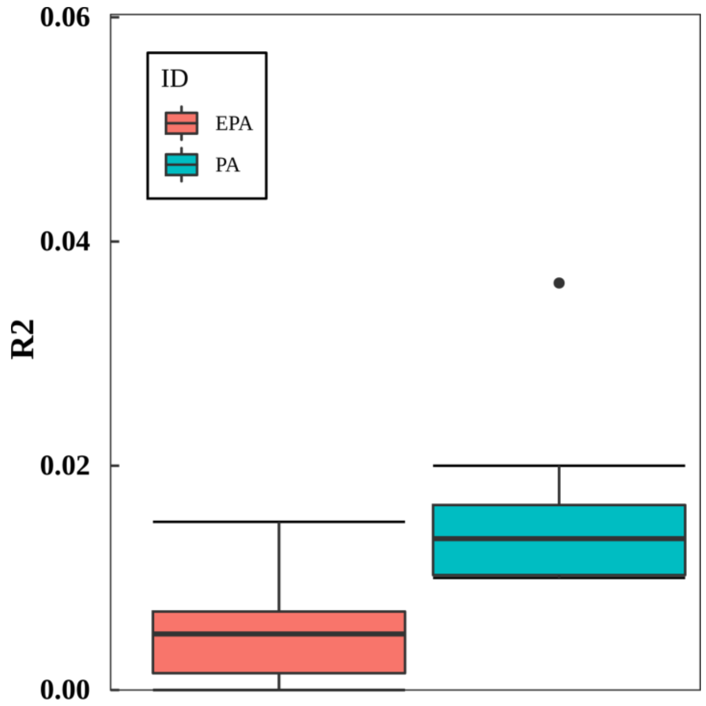


Impact of Meteorological Parameters on $PM_{2.5}$ Data

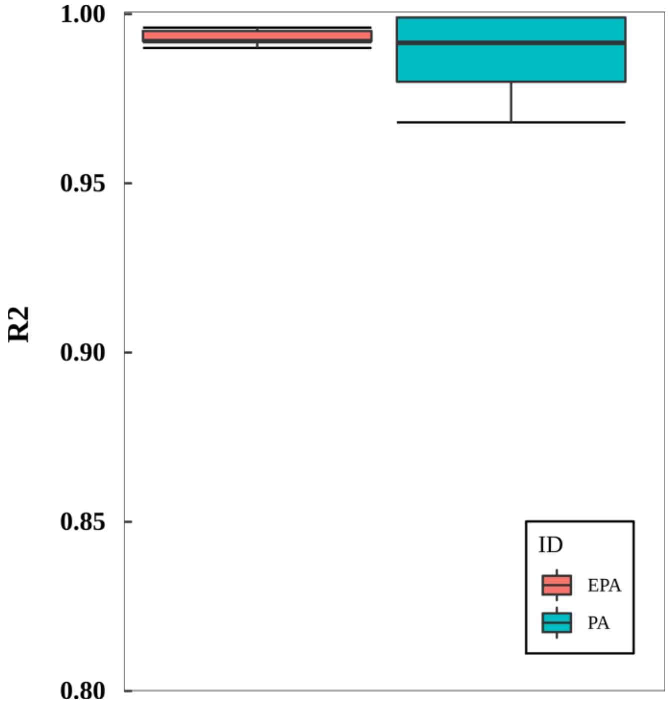
- Linear regression model was built using temperature, relative humidity, wind speed, and wind direction as meteorological parameters
- Temporal components of $PM_{2.5}$ data were used as independent variables
- Relative contribution of short-term and long-term components were compared against total data
- Model R^2 values were used to compare the goodness of fit



All data



Short-term component



Long-term component

Meteorological parameters better explains the variability in the long-term component of $PM_{2.5}$ data

Conclusions, and Future Work

- Analysis of $PM_{2.5}$ time-series data after decomposition into various frequency components could be useful to understand the contribution from various sources
- PA sensor data are comparable in efficiency to EPA in capturing the long-term variations in the $PM_{2.5}$ data but less efficient in detecting short term variations (<12 hours)
- Long term variations in the $PM_{2.5}$ data could be explained by the meteorological parameters (temperature, relative humidity, wind speed, and wind direction)
- Future studies will focus on modeling the short-term variations with local sources of $PM_{2.5}$ such as traffic information

References

[1] K. Vohra, et al. Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: results from GEOS-Chem Environ. Res., 195 (2021), p. 110754

[2] Rao, S. T., and I. Zurbenko, 1994: Detecting and tracking changes in ozone air quality. J. Air Waste Manag. Assoc., 44, 1089–1095.

[3] Yong-Sang Choi, Chang-Hoi Ho, Deliang Chen, Yeon-Hee Noh, Chang-Keun Song,Spectral analysis of weekly variation in PM10 mass concentration and meteorological conditions over China, Atmospheric Environment, Volume 42, Issue 4, 2008, Pages 655-666 https://doi.org/10.1016/j.atmosenv.2007.09.075.

[4] O. Tchepel, Frequency analysis of air quality time series for traffic related pollutants, journal of environmental monitoring, 2010

[5] Poojan Upadhaya, Meteorological Detrending of Ozone at Three Sites in the Dallas-Fort Worth Area: Application of KZ Filter Method, atmosphere, 2020

[6] Edward Valachovic, Separation of Spatio-Temporal Frequencies, Quest Journals Journal of Research in Applied Mathematics Volume 3 ~ Issue 7 (2017) pp: 29-37

[7] Erika K. Wise & Andrew C. Comrie (2005) Extending the Kolmogorov–Zurbenko Filter: Application to Ozone, Particulate Matter, and Meteorological Trends, Journal of the Air & Waste Management Association, 55:8, 1208-1216, DOI: 10.1080/10473289.2005.10464718

Thank You!