

# **Urban aerosol: spatiotemporal variation & source characterization**

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Mechanical Engineering

Zhongju (Hugh) Li

B.S., Automotive Engineering, Wuhan University of Technology

Carnegie Mellon University  
Pittsburgh, PA

January 2018

Urban aerosol: spatiotemporal variation & source characterization

© Zhongju (Hugh) Li, 2018

All Rights Reserved

*~To my family and friends for the endless support~*

## Acknowledgements

I am forever thankful to my advisor Professor Albert Presto for everything during the last 5 years. I came with a naive mind for air pollution field, and the five-year training truly made me more inquisitive and insightful. The journey is truly pleasant, though ups and downs. I truly thank my advisor for the endless support and understanding all along.

I would like to acknowledge my committee members, Professors Allen Robinson, Peter Adams, and Jeremy Michalek. Thanks for taking time reading my proposal and thesis, and providing valuable feedback how to keep improving upon present.

I am grateful for previous and current research group members, the Center for Atmospheric Particle Studies. The weekly coffee hour is where many interesting conversations happened.

My deepest thank you goes to my wife and family. Thanks for the support and understanding for all these years.

The work was funded by Heinz Endowments grant number E0678, EPA STAR grant number RD83587301, and NSF grant number AGS1543786.

I wanna thank Yi Tan at the California Air Source Board for assistance with LUR source data. Sarah Rose Eilenberg at Carnegie Mellon University for biomass filter analysis, Katerina Liagou at University of Patras for cooking filter results. Rishabh Shah, Jiqiao Shi, Aliaksei Hauryliuk, and Dr. Aja Ellis for help with mobile sampling and distributed monitors.

## Abstract

Long and short-term exposure to particulate matter (PM) are linked to adverse health endpoints. Evidence indicates that PM composition such as metals and organic carbon (OC) might drive the health effects. As airborne pollutants show significant intracity spatiotemporal variation, mobile sampling and distributed monitors are utilized to capture the variation pattern. The measurements are then fed to develop models to better characterize the relationship between exposure and health outcomes.

Two sampling campaigns were conducted. One was sole mobile sampling in 2013 summer and winter in Pittsburgh, PA. Thirty-six sites were chosen based on three stratification variables: traffic density, proximity to point sources, and elevation. The other one was hybrid sampling network, incorporating a mobile sampling platform, 15 distributed monitors, and a supersite. We designed two case studies (transect and downtown), selected 14 neighborhoods ( $\sim 1 \text{ km}^2$ ), and conducted sampling in 2016 summer/fall and winter.

Spatial variation of  $\text{PM}_{2.5}$  mass and composition was studied in the 2013 campaign. X-ray fluorescence (XRF) was used to analyze concentrations of 26 elements: Na, Mg, Al, Si, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Br, Rb, Sr, Zr, Cd, Sb, and Pb. Trace elements had a broad range of concentrations from 0 to  $300 \text{ ng/m}^3$ . Comparison of data from mobile sampling with stationary monitors showed reasonable agreement. We developed Land use regression (LUR) models to describe spatial variation of  $\text{PM}_{2.5}$ , Si, S, Cl, K, Ca, Ti, Cr, Fe, Cu, and Zn. Independent variables included traffic influence, land-use type, and facility emissions. Models had an average  $R^2$  of 0.57 ( $\text{SD} = 0.16$ ). Traffic related variables explained the most variability with an average  $R^2$  contribution of 0.20 ( $\text{SD} = 0.20$ ). Overall, these results demonstrated significant intra-urban spatial variability of fine particle composition.

Spatial variation of OC was based on the 2013 campaign as well. We collected

organic carbon (OC) on quartz filters, quantified different OC components with thermal-optical analysis, and grouped them based on volatility in decreasing order (OC1, OC2, OC3, OC4, and pyrolyzed carbon (PC)). We compared our ambient OC concentrations (both gas and particle phase) to similar measurements from vehicle dynamometer tests, cooking emissions, biomass burning emissions, and a highway traffic tunnel. OC2 and OC3 loading on ambient filters showed a strong correlation with primary emissions while OC4 and PC were more spatially homogenous. While we tested our hypothesis of OC2 and OC3 as markers of fresh source exposure for Pittsburgh, the relationship seemed to hold at a national level. Land use regression (LUR) models were developed for the OC fractions, and models had an average  $R^2$  of 0.64 (SD = 0.09). We demonstrate that OC2 and OC3 can be useful markers for fresh emissions, OC4 is a secondary OC indicator, and PC represents both biomass burning and secondary aerosol. People with higher OC exposure are likely inhaling more fresh OC2 and OC3, since secondary OC4 and PC varies much less drastically in space or with local primary sources.

With the 2016 hybrid sampling campaign, we addressed the intracity exposure patterns, as they could be more complex than intercity ones because of local traffic, restaurants, land use, and point sources. This network studied a wide range of pollutants ( $\text{CO}_2$ , CO,  $\text{NO}_2$ ,  $\text{PM}_{10}$  mass and composition, and particle number PN). Mobile measurements and distributed monitors show good agreement. PN hotspots are strongly associated with restaurants and highway traffic. PN at sites with large local source impacts tends to have larger diurnal variation than daily variation, while CO in downtown center shows the opposite trend. PN exhibits the largest spatial and temporal variations. Spatial variation is generally larger than temporal variation among all five pollutants ( $\text{CO}_2$ ,  $\text{NO}_2$ , CO, PN, and  $\text{PM}_{10}$ ). These findings provide quantitative comparison between spatial and temporal variation

in different scales, and support the theoretical validity of developing long-term exposure models from short-term mobile measurement. A combined sampling network with mobile and distributed monitor could prove more valuable in studying intracity air pollution.

In the 2016 hybrid sampling campaign, we also studied spatial variability of air pollution in the vicinity of monitors. Monitoring network is essential for protecting public health, though evaluation is needed to assess spatial representativeness of monitors in different environments. Mobile sampling was conducted repeatedly around 15 distributed monitors. Substantial short-range spatial variability was observed. Spatial variation was consistently larger than temporal variation for NO<sub>2</sub> and CO at different sites. Ultrafine particles were highly dynamic both in space and time. PM<sub>1</sub> was less spatially and temporally variable. Urban locations had more frequent episodic source plume events compared with background sites. Using a single monitor measurement to represent surrounding ~1 km<sup>2</sup> areas could introduce an average daily exposure misclassification of 46 ppb (SD = 26) for CO (30% of regional background), 3 ppb (SD = 2) for NO<sub>2</sub> (43% of background), 4007 #/cm<sup>3</sup> (SD = 1909) for ultrafine particle number (64% of background), and 1.2 µg/m<sup>3</sup> (SD = 1.0) for PM<sub>1</sub> (13% of background). Exposure differences showed fair correlation with traditional land use covariates such as traffic and restaurant density, and the magnitude of misclassification could be even bigger for urban neighborhoods.

## Table of Contents

Chapter 1: Introduction	1
1.1 Atmospheric aerosols and air pollution	4
1.2 Intracity spatiotemporal variation	6
1.3 Land use regression model	7
1.4 Spatial representativeness of stationary monitors	8
1.5 Research objectives	9
1.6 Dissertation outline	12
1.7 References	14
Chapter 2: Application of mobile sampling to investigate spatial variation in fine particle composition	24
2.1 Introduction	26
2.2 Material and Methods	27
2.2.1 Mobile sampling lab setup	27
2.2.2 Mobile sampling overview	27
2.2.3 Data handling	28
2.2.4 LUR model construction and evaluation	30
2.3 Results and Discussion	33
2.3.1 Mobile sampling versus stationary sites	33
2.3.2 Abundance of selected particle components	36
2.3.3 Traffic indicators	38
2.3.4 LUR models	39
2.4 Limitations	52
2.5 Discussion and Implication	52
2.6 References	54
Chapter 3: Urban organic aerosol exposure: spatial variations in composition and source impacts	62
3.1 Introduction	64
3.2 Material and Methods	65
3.2.1 Mobile platform setup	66
3.2.2 Sampling overview	66
3.2.3 OC quantification	67
3.2.4 LUR development	69
3.3 Results and Discussion	72
3.3.1 OC spatial variation	73
3.3.2 Different fresh emission influence on OCX	73
3.3.3 From Pittsburgh to nationwide	75
3.3.4 LUR models	76



3.3.5 Exposure implications.....	79
3.4 Limitations.....	87
3.5 References.....	88
Chapter 4: Comparison of spatial and temporal variation of airborne pollutants using mobile and distributed sampling.....	101
4.1 Introduction.....	103
4.2 Methods.....	105
4.2.1 Sampling platform setup.....	105
4.2.2 Sampling overview.....	107
4.2.3 Data preparation and analysis.....	108
4.3 Results and Discussion.....	110
4.3.1 Sampling domain overview.....	110
4.3.2 Concurrent mobile vs. distributed monitors.....	111
4.3.3 Temporal variation in two scales – daily vs. diurnal.....	113
4.3.4 Spatial vs. temporal variation.....	118
4.3.5 COV comparison with other studies.....	120
4.4 Implication for land use regression modeling.....	128
4.5 Reference.....	130
Chapter 5: Spatial variability of air pollution near monitors and exposure misclassification in an eastern US city.....	145
5.1 Introduction.....	147
5.2 Material and Methods.....	148
5.2.1 Sampling setup.....	148
5.2.2 Sampling overview.....	150
5.2.3 Data treatment.....	151
5.3 Results and Discussion.....	152
5.3.1 Spatial variability at each site.....	152
5.3.2 Temporal variability during each trip.....	154
5.3.3 Spatial variability near urban monitors.....	155
5.3.4 Spatial variability near an urban background monitor.....	157
5.3.5 CDF of spatial vs. temporal variability.....	158
5.3.6 Exposure misclassification.....	159
5.4 Limitation.....	171
5.5 References.....	172
Chapter 6: Conclusions.....	183
6.1 Summary of major findings.....	185
6.2 Future studies.....	190

Appendices.....	195
Appendix A .....	197
Appendix B .....	201
Appendix C .....	206
Appendix D .....	210

## List of Figures

Fig. 1.1. Venn diagram for airborne pollutants included in the thesis .....	3
Fig. 2.1. Boxplot comparisons of PM composition, stationary vs. mobile .....	43
Fig. 2.2. Summer vs. winter PM composition .....	44
Fig. 2.3. Correlation of species in summer vs. winter .....	45
Fig. 2.4. Zn, EC and PAH based on traffic intensity and season .....	46
Fig. 2.5. LUR model relative $R^2$ explained by GIS groups .....	47
Fig. 2.6. Measured vs. LUR predicted concentrations .....	48
Fig. 2.7. LUR PM <sub>2.5</sub> prediction for mobile sites and 2 EPA monitors .....	49
Fig. 3.1. Average OC and EC at 36 sites .....	81
Fig. 3.2. Seasonal comparison of bare quartz OCX .....	82
Fig. 3.3. OCX on bare quartz filters from all CSN sites in 2013 .....	83
Fig. 3.4. LUR prediction of OCX .....	84
Fig. 3.5. CDF of population exposure to OCX .....	85
Fig. 4.1. 15 sampling sites overview with stratification information .....	121
Fig. 4.2. Concurrent NO <sub>2</sub> measurements from mobile and monitor .....	122
Fig. 4.3. CO measurements from monitors 6 and 14 .....	123
Fig. 4.4. 1 week PN measurements from sites 7 and 15 .....	124
Fig. 4.5. COV of temporal variations in two scales .....	125
Fig. 4.6. Coefficient of variation for spatial and temporal variations .....	126
Fig. 5.1. NO <sub>2</sub> spatial enhancement around 15 sites .....	163
Fig. 5.2. CDF of hourly concentration change from distributed monitors .....	164
Fig. 5.3. Spatial variation at downtown .....	165
Fig. 5.4. Spatial variation at urban background .....	166
Fig. 5.5. CDF of relative spatial differences at all sites .....	167
Fig. 5.6. Mean absolute relative pollutant concentration difference .....	168
Fig. 5.7. Exposure differences vs. land use covariates .....	169
Fig. A1. Mobile sampling domain overview .....	198
Fig. A2. Elevation map for Allegheny county, PA .....	199
Fig. B1. 36 mobile sampling locations .....	202
Fig. B2. LUR prediction versus measured values at 36 sampling sites .....	203
Fig. C1. Wind rose map .....	207
Fig. C2. High ICC value (left, 0.56) vs. low ICC (right, 0.02) .....	208

Fig. D1. Overview of sampling campaign in Allegheny County, PA	211
Fig. D2. Influence of grid box size (100 m vs. 50 m)	212
Fig. D3. Mobile (1 s) vs. distributed monitors (15 minutes)	213

List of Tables

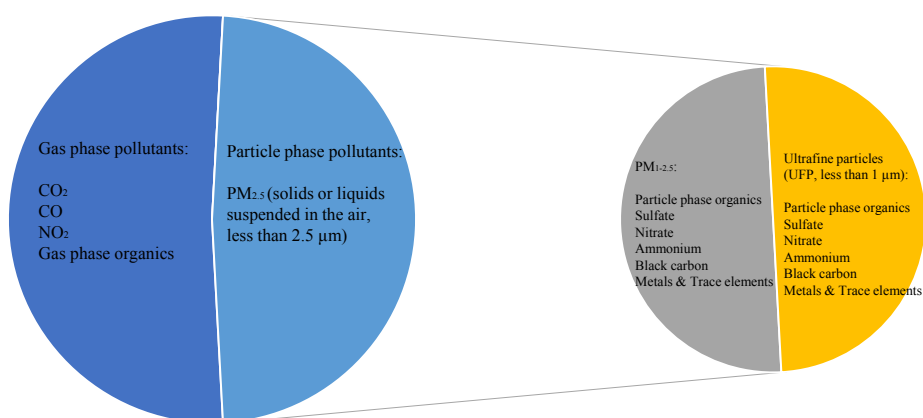
Table 2.1	50
Table 2.2	51
Table 3.1	86
Table 4.1	127
Table 5.1	170
Table A1	200
Table B1	204
Table B2	205
Table C1	209

## **Chapter 1: Introduction**

# **Chapter 1**

## **Introduction**

### Listed air pollutants in the thesis



\* Size of the pie does not necessarily mean the abundance of pollutants.

**Fig. 1.1.** Venn diagram for airborne pollutants included in the thesis. Airborne pollutants can be in the gas or particle phase.



## 1.1 Atmospheric aerosols and air pollution

Airborne fine particulate matter (PM<sub>2.5</sub>; particles smaller than 2.5  $\mu\text{m}$  diameter) continues to pose serious threats to human health, especially to vulnerable groups such as the elderly and children (Beelen et al., 2014; Brook et al., 2010; Di et al., 2017b; Dockery et al., 1993; Pope and Dockery, 2006). Worldwide, long-term exposure to PM results in 7 million deaths annually (Brauer et al., 2016). Recent findings suggest no short-term or long-term safe limit for breathing fine particulate matter and super-linear dose response functions when pollutant concentrations are below the U.S. National Ambient Air Quality Standards (Beelen et al., 2014; Di et al., 2017b, 2017a).

Figure 1.1 shows PM<sub>2.5</sub> is mainly composed of inorganic salts, organic carbon (OC), elemental carbon (EC), crustal components, and other trace elements (Rees et al., 2004). Sulfate ( $\text{SO}_4^{2-}$ ), nitrate ( $\text{NO}_3^-$ ), ammonium ( $\text{NH}_4^+$ ), and chloride ( $\text{Cl}^-$ ) are major inorganic ions. Crustal elements are relatively abundant species in the earth's crust such as silicon (Si), aluminum (Al), iron (Fe), potassium (K), calcium (Ca), titanium (Ti), sodium (Na), and magnesium (Mg).

Organic carbon (OC) is a major component of PM with a broad range of concentrations from 0.1 to 100  $\mu\text{g}/\text{m}^3$  worldwide (Jimenez et al., 2009). Ambient OC is a mixture of primary (directly emitted) and secondary (formed via chemical processing) components. Primary organic aerosol such as tailpipe exhaust is semi-volatile (Robinson et al., 2007). Fresh emissions are rapidly diluted in the ambient atmosphere. These emissions partially evaporate to form low-volatility vapors. The photo-oxidation of these newly created vapors, along with photo-oxidation of volatile organic gases, contributes to the burden of secondary ambient organic aerosol (Robinson et al., 2007). Secondary organic aerosol dominates the OC mass in most environments (Zhang et al., 2007), however in near source

regions (e.g., near roadways) emissions of primary OC can strongly influence population exposures (Donahue et al., 2016).

The mechanism how PM exposure reduces life expectancy remains unclear (Brook et al., 2010; Schwartz et al., 1996). While most air pollution epidemiology studies have focused on the association of health endpoints with total PM<sub>2.5</sub> mass (Brook et al., 2010; Pope et al., 2009), evidence suggests that PM<sub>2.5</sub> composition may drive some health effects.

Epidemiological studies by Krall et al. (2017, 2013) suggest that some PM components and PM from certain sources may be more toxic than others. Compositional differences may also explain the enhanced risks associated with PM exposure in urban versus rural areas (Bravo et al., 2017). Verma et al. (2015) observed that organic aerosol from biomass burning was more effective at generating reactive oxygen species (ROS) than organic aerosol from other sources, suggesting that these emissions may be more toxic. Strak et al. (2017) showed that the prevalence of diabetes was related to PM oxidative potential instead of PM<sub>2.5</sub> mass, suggesting a compositional dependence of human health response. Urch et al. (2008) showed strong association of blood pressure increase with particulate organic carbon concentration, but not with total PM<sub>2.5</sub> mass, as well as a significant negative association between particulate organic carbon and arterial diameter (Urch et al., 2005).

The US Environmental Protection Agency (EPA) regulates outdoor concentrations of CO, NO<sub>2</sub>, ozone (O<sub>3</sub>), and fine particulate matter (PM<sub>2.5</sub>) because of their known impacts on human health. Ultrafine particle number (UFP, particles with aerodynamic diameter less than 100 nm) are a pollutant of emerging concern (Kerckhoffs et al., 2016; Meier et al., 2015; Patton et al., 2015; Saraswat et al., 2013; Simon et al., 2017). UFP can penetrate deep into the respiratory system and have been linked to human health impacts (Health Effects Institute, 2013; Liu et al., 2015; Ostro et al., 2015; Stafoggia et al., 2017), though no current regulatory

standards exist. Particle number (PN) concentration is commonly used as a proxy for UFP (Eeftens et al., 2015; Meier et al., 2015).

Airborne pollutants can trigger both acute and chronic health outcomes (Di et al., 2017b, 2017a). PM<sub>2.5</sub> has been consistently found to link to both short-term (hours to days) and long-term (annual) health endpoints (Beelen et al., 2014; Brook et al., 2010; Di et al., 2017b, 2017a). All major air pollutants (carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, PM<sub>2.5</sub>, and PM<sub>10</sub>), with the exception of ozone, were significantly associated with a short-term increase in myocardial infarction and heart attack risk (Mustafić et al., 2012). Although no standards regulate particle concentration, particle number concentration, especial ultrafine particles, also showed statistically significant correlation with short-term and long-term cardiovascular morbidity (Ostro et al., 2015; Weichenthal, 2012).

## **1.2 Intracity spatiotemporal variation**

Numerous studies suggest substantial intra-urban variability of PM mass and composition (de Hoogh et al., 2013; Eeftens et al., 2012a; Tsai et al., 2015; Zhang et al., 2015). Many cities have only a few PM speciation monitors, and the spatial density of these sites may not be sufficient to characterize population exposures with high spatial resolution because of large spatial heterogeneity. Thus, understanding spatial variations in pollutant concentrations and subsequent human exposures often relies on targeted sampling campaigns. For example, the European Study of Cohorts for Air Pollution Effects (ESCAPE) sampled at 20 monitoring sites in each of 20 cities to resolve spatial patterns in PM<sub>2.5</sub> mass and composition (de Hoogh et al., 2013; Eeftens et al., 2012a). ESCAPE used Harvard impactors to collect PM samples in three 2-week periods over a year to analyze spatial patterns of pollutants. Similarly, Zhang et al. (2015) applied distributed sampling with filters to analyze metals associated with PM<sub>1.0</sub> in Calgary, Alberta, Canada. Recently, Zimmerman et al.

demonstrated the use of low-cost distributed monitors with 15-minute resolution (2017) to characterize spatial and temporal variations in pollutant concentrations.

Mobile sampling is an attractive alternative to distributed sampling for several reasons. High time resolution instruments can be deployed on mobile platforms (Larson et al., 2009; Tan et al., 2014a), enabling determination of specific source impacts, whereas distributed samplers often rely on filters or other time-integrated measurements. Mobile sampling may also allow investigation of a larger number of locations at lower monetary and logistic cost than distributed samplers. The primary downside of mobile sampling is that the sampling duration at a given measurement location is relatively short, ranging from minutes to several hours (Hankey and Marshall, 2015; Tan et al., 2014a). Thus, while mobile sampling can capture spatial variation in pollutant concentrations and particle composition, it is less adept at capturing long-term average concentrations (Tan et al., 2014b). Nonetheless, mobile measurement data have been used to build spatial maps of pollutant concentrations in several cities (Hankey and Marshall, 2015; Larson et al., 2009; Tan et al., 2016).

### **1.3 Land use regression model**

Data from distributed or mobile sampling campaigns are often used to build statistical models to represent concentrations across an entire city. Many models exist for modeling spatial variability such as kriging, inverse distance weighting, and land use regression (LUR). LUR is inexpensive to implement and can achieve reliable estimates when adequate land use, traffic information, and monitoring data are available (Health Effects Institute, 2010).

LUR links measured concentrations to land use type, traffic, point source, and other physical environmental variables. While not a formal source apportionment method, LURs can provide information on different source classes (e.g., traffic) and observed pollutant concentrations.

LUR can be used to resolve pollutant spatial patterns with high resolution (25 – 100 meters). However, while predictors in LUR models may link measured concentrations to specific sources, LUR models lack true physical insight. Thus LUR models are not easily transferrable outside of the region where they are built (Patton et al., 2015; Poplawski et al., 2009). This is in contrast to chemical transport models (CTM), which simulate all of the important chemical and physical processes in the atmosphere, but at much lower spatial resolution (e.g., 16x16 or 32x32 km grid cells) (Robinson et al., 2007). Successful applications of LUR include PM<sub>10</sub>, PM<sub>2.5</sub>, particle number, VOC, and PM composition (de Hoogh et al., 2013; Eeftens et al., 2012a; Jedynska et al., 2014; Zhang et al., 2015).

LUR models can be built with different data sources: distributed filters and passive samplers (Eeftens et al., 2012; Matte et al., 2013), mobile sampling (Hankey and Marshall, 2015; Li et al., 2017, 2016), and satellite data coupled to the EPA network (Bechle et al., 2015).

#### **1.4 Spatial representativeness of stationary monitors**

A regulatory monitoring network is an essential component for protecting public health from air pollution. Data from the network can be used to determine if an area is in compliance with current air pollution standards. The data are also important sources for studying the relationship between air pollution and health impacts with epidemiology. The ground-breaking Harvard Six City study used filter measurements from regulatory stationary monitors to represent residents' exposure to airborne pollutants in a county (Dockery et al., 1993). However, whether monitors could represent a county level air pollution pattern (~30 km) is subject to debate (Martin et al., 2014; Piersanti et al., 2015; Vitali et al., 2016).

Ground based monitor measurements are also important inputs to atmospheric models such as chemical transport models (Di et al., 2016; Ivey et al., 2017). The monitor data also

help to calibrate remote sensing models, e.g. deriving a reliable PM<sub>2.5</sub> – aerosol optical depth (AOD) relationship. However, for these modeling studies (chemistry transport model and remote sensing), mismatch in spatial resolution is an issue, as it can introduce significant uncertainty in final model output (Yin et al., 2016). Ground based monitors provide point measurements. Chemical transport model and remote sensing utilize a grid based approach, with grid size ranging from 3 km to 36 km (Donahue et al., 2016; Guo et al., 2017; Lamsal et al., 2017; Robinson et al., 2007). Using point measurements to represent surrounding areas as large as county scale casts doubts on uncertainty magnitude of final model output.

### **1.5 Research objectives and motivation**

The overarching goal of this thesis is to improve the characterization of spatial and temporal variation of multiple airborne pollutants with mobile sampling and distributed network. The thesis addressed the current knowledge gap in measuring and modeling particle composition within a city, shed new light in using OC fractions to characterize population exposure to organics, studied how pollutants varied both in space and time, proposed optimal sampling strategies to capture spatiotemporal pattern in different target pollutants, and quantified spatial representativeness of monitors in diverse neighborhoods (~1 km<sup>2</sup>). These findings provide guidance for future sampling designs in network optimization and avoiding redundant monitors, help to inform air quality planning, and advance understanding of pollutant variation and resulting population exposure pattern. Detailed research objectives and motivations are listed below.

- 1) How do PM compositions vary spatially? Limited studies described intracity spatial variation of PM compositions, partially due to heavy logistic requirements of setting up a large scale distributed monitor network (> 50 monitors) (de Hoogh et al., 2013; Zhang et al., 2015). In our study domain (Allegheny County, PA), only two regulatory

monitors reported PM speciation data every three or six days. PM health effects, however, showed compositional dependence (Krall et al., 2017, 2013). Studying the intracity spatial variation in PM composition could help answer if enough regulatory monitors existed in the domain for protecting public health.

- 2) What is the spatial variation in organics? Organics is a major component of fine PM, ranging from 20 to 90% in mass fraction (Zhang et al., 2007). Organics were also associated with adverse health impacts (Verma et al., 2015). Thus, it was important to quantify the spatial variation of organics, and assessed whether certain OC fractions were linked to specific sources. By identifying the source-receptor relationships, effective pollution control could be adopted to reduce population exposure to total organics, and finally PM as well. Two national speciation networks reported OC fractions data for decades (Solomon et al., 2014). Previous studies frequently incorporated total organics and simply disregarded OC fraction information. The potential answers to identify sources for OC fractions could contribute new information on using historical data to study organics exposure on a national scale.
- 3) How does multiple pollutant exposure vary spatially and temporally? Pollutants not only vary in space, but also in time. A comprehensive understanding of variation pattern for multiple pollutants help decide the optimal sampling strategy—mobile sampling only, distributed monitors only or both. In addition, studying temporal variation was especially important for acute health diseases. Numerous studies indicated statistically significant spatial variation of airborne pollutants in street level ( $\sim$  m) and neighborhood scales ( $\sim$ 1 km) (Apte et al., 2017; Donahue et al., 2016; Li et al., 2017, 2016). We suggested mobile sampling had the unique advantage of capturing spatial gradient with high resolution. With mobile measurements at different

locations as input, models could be built to predict concentrations at unmeasured spots. LUR was among the popular tools to predict long-term spatial pattern. Independent variables in LUR such as road length in surrounding areas were temporally invariant. Thus, it was critical to assess whether it was reliable to build long-term LUR models based on short-term mobile measurements. One way to address the knowledge gap is to compare relative magnitude of spatial and temporal variation in mobile sampling data. If temporal variation was orders of magnitude higher than spatial variation, even after temporal correction was applied to the raw mobile measurements, the corrected dataset would have substantial residual temporal variation. Little real spatial gradient could be inferred in this case. On the contrary, if spatial signal was significantly higher than temporal variation, spatial pattern would reveal itself even before we got to the point of having a stable annual average. This would justify the usage of LUR model based on short-term mobile measurements.

- 4) Do regulatory monitors represent air pollution in their neighborhood areas ( $\sim 1 \text{ km}^2$ ) well? Monitors can be used to determine if their surrounding regions meet regulatory standards of air pollution. Measurements at monitors describe surrounding residents' air pollution exposure, help derive exposure dose response functions in epidemiology studies, and are essential input for chemical transport model and remote sensing. The knowledge gap is to quantify the spatial representativeness of monitors at different environments for a variety of gas and particle phase pollutants. The answer can further help inform the spatial density of monitors needed to quantify residents' exposure within certain percentages of "true" population exposure based on residential address (home). This can be a valuable sampling planning tool, either to suggest frequency/density of future mobile sampling in a city, or the number of



distributed monitors in different areas.

## **1.6 Dissertation outline**

In Chapter 2, we conducted a mobile sampling campaign in 2013 summer and winter in Pittsburgh, PA to characterize spatial variation in PM<sub>2.5</sub> mass and composition. Thirty-six sites were chosen based on three stratification variables: traffic density, proximity to point sources, and elevation. We compared the mobile dataset to stationary monitors in the county to illustrate the difference between two sampling platforms. The effectiveness of using Zn as traffic marker in the ambient environment was investigated. We then developed LUR models to describe spatial variation of PM composition. The importance of different LUR predictors was discussed.

In Chapter 3, we specifically analyzed one major PM component collected in the same campaign -- OC. We compared our ambient OC concentrations (both gas and particle phase) to similar measurements from vehicle dynamometer tests, cooking emissions, biomass burning emissions, and a highway traffic tunnel. We then tested observed source relationship in national speciation network. LUR models were developed for these OC fractions. The models were then combined with census block population data to illustrate human exposure pattern to airborne organic carbon.

In Chapter 4, we use a hybrid sampling network to characterize spatiotemporal variations of multiple pollutants in Pittsburgh, PA. This network incorporates a mobile sampling platform, distributed monitors, and a supersite to investigate a wide range of pollutants (CO<sub>2</sub>, CO, NO<sub>2</sub>, PM<sub>1</sub> mass and composition, and particle number PN). We designed two case studies (transect and downtown), selected 14 neighborhoods (~1 km<sup>2</sup>), and conducted sampling in 2016 summer/fall and winter. We quantitatively compared spatial and temporal variation in different scales, and addressed the theoretical validity of developing

long term exposure models from short-term mobile measurement. We further suggested different sampling strategies to better capture intracity air pollution variation for different pollutant.

In Chapter 5, we used the same hybrid sampling network to study spatial variation of air pollution near stationary monitors. We conducted mobile sampling repeatedly around 15 distributed monitors. We compared observed concentration differences to hourly temporal variation during driving trips to answer whether differences were source driven or just temporal fluctuations. Short-term (daily) exposure misclassification was quantified, and the relationship between exposure difference and land use covariates was studied.

In Chapter 6, we summarized major research findings and discussed future research ideas.

## 1.7 References

- Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environ. Sci. Technol.* 51, 6999–7008.
- Bechle, M.J., Millet, D.B., Marshall, J.D., 2015. National Spatiotemporal Exposure Surface for NO<sub>2</sub>: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010. *Environ. Sci. Technol.* 49, 12297–12305. <https://doi.org/10.1021/acs.est.5b02882>
- Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A.S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., Hoek, G., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet* 383, 785–795. [https://doi.org/10.1016/S0140-6736\(13\)62158-3](https://doi.org/10.1016/S0140-6736(13)62158-3)

- Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L.T., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* 50, 79–88. <https://doi.org/10.1021/acs.est.5b03709>
- Bravo, M.A., Ebisu, K., Dominici, F., Wang, Y., Peng, R.D., Bell, M.L., 2017. Airborne Fine Particles and Risk of Hospital Admissions for Understudied Populations: Effects by Urbanicity and Short-Term Cumulative Exposures in 708 U.S. Counties. *Environ. Health Perspect.* 125, 594–601. <https://doi.org/10.1289/EHP257>
- Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. *Circulation* 121, 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbecel>
- de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., Dèdelè, A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gražulevičienė, R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsouyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Mosler, G., Nádor, G., Nieuwenhuijsen, M., Pershagen, G., Peters, A., Phuleria, H., Probst-Hensch, N., Raaschou-Nielsen, O., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Weinmayr, G.,

- Brunekreef, B., Hoek, G., 2013. Development of Land Use Regression Models for Particle Composition in Twenty Study Areas in Europe. *Environ. Sci. Technol.* 47, 5778–5786. <https://doi.org/10.1021/es400156t>
- Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017a. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. *JAMA* 318, 2446–2456. <https://doi.org/10.1001/jama.2017.17923>
- Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., Schwartz, J., 2016. Assessing PM<sub>2.5</sub> Exposures with High Spatiotemporal Resolution across the Continental United States. *Environ. Sci. Technol.* 50, 4712–4721. <https://doi.org/10.1021/acs.est.5b06121>
- Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.D., 2017b. Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* 376, 2513–2522. <https://doi.org/10.1056/NEJMoa1702747>
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G.J., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities. *N. Engl. J. Med.* 329, 1753–1759. <https://doi.org/10.1056/NEJM199312093292401>
- Donahue, N.M., Posner, L.N., Westervelt, D.M., Li, Z., Shrivastava, M., Presto, A.A., Sullivan, R.C., Adams, P.J., Pandis, S.N., Robinson, A.L., 2016. Where Did This Particle Come From? Sources of Particle Number and Mass for Human Exposure Estimates, in: *Airborne Particulate Matter*. Royal Society of Chemistry, Cambridge, pp. 35–71.
- Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dèdèlè, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M.,

- Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S. von, Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of Land Use Regression Models for PM<sub>2.5</sub>, PM<sub>2.5</sub> Absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European Study Areas; Results of the ESCAPE Project. *Environ. Sci. Technol.* 46, 11195–11205. <https://doi.org/10.1021/es301948k>
- Eeftens, M., Phuleria, H.C., Meier, R., Aguilera, I., Corradi, E., Davey, M., Ducret-Stich, R., Fierz, M., Gehrig, R., Ineichen, A., Keidel, D., Probst-Hensch, N., Ragettli, M.S., Schindler, C., Künzli, N., Tsai, M.-Y., 2015. Spatial and temporal variability of ultrafine particles, NO<sub>2</sub>, PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in Swiss study areas. *Atmos. Environ.* 111, 60–70. <https://doi.org/10.1016/j.atmosenv.2015.03.031>
- Guo, Y., Tang, Q., Gong, D.-Y., Zhang, Z., 2017. Estimating ground-level PM<sub>2.5</sub> concentrations in Beijing using a satellite-based geographically and temporally weighted regression model. *Remote Sens. Environ.* 198, 140–149. <https://doi.org/10.1016/j.rse.2017.06.001>
- Hankey, S., Marshall, J.D., 2015. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM<sub>2.5</sub>, Particle Size) Using Mobile Monitoring. *Environ. Sci. Technol.* 49, 9194–9202.
- HEI Review Panel on Ultrafine Particles, 2013. Understanding the Health Effects of Ambient Ultrafine Particles, HEI Perspectives 3. Health Effects Institute, Boston, MA.
- Ivey, C., Holmes, H., Shi, G.-L., Balachandran, S., Hu, Y., Russell, A.G., 2017. Development

- of PM<sub>2.5</sub> source profiles using a hybrid chemical transport-receptor modeling approach. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.7b03781>
- Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prevot, A.S.H., Zhang, Q., Kroll, J.H., DeCarlo, P.F., Allan, J.D., Coe, H., Ng, N.L., Aiken, A.C., Docherty, K.S., Ulbrich, I.M., Grieshop, A.P., Robinson, A.L., Duplissy, J., Smith, J.D., Wilson, K.R., Lanz, V.A., Hueglin, C., Sun, Y.L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J.M., Collins, D.R., Cubison, M.J., E, Dunlea, J., Huffman, J.A., Onasch, T.B., Alfarra, M.R., Williams, P.I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimojo, A., Sun, J.Y., Zhang, Y.M., Dzepina, K., Kimmel, J.R., Sueper, D., Jayne, J.T., Herndon, S.C., Trimborn, A.M., Williams, L.R., Wood, E.C., Middlebrook, A.M., Kolb, C.E., Baltensperger, U., Worsnop, D.R., 2009. Evolution of Organic Aerosols in the Atmosphere. *Science* 326, 1525–1529. <https://doi.org/10.1126/science.1180353>
- Kerckhoffs, J., Hoek, G., Messier, K.P., Brunekreef, B., Meliefste, K., Klompmaker, J.O., Vermeulen, R., 2016. Comparison of Ultrafine Particle and Black Carbon Concentration Predictions from a Mobile and Short-Term Stationary Land-Use Regression Model. *Environ. Sci. Technol.* 50, 12894–12902. <https://doi.org/10.1021/acs.est.6b03476>
- Krall, J.R., Anderson, G.B., Dominici, F., Bell, M.L., Peng, R.D., 2013. Short-term Exposure to Particulate Matter Constituents and Mortality in a National Study of U.S. Urban Communities. *Environ. Health Perspect.* <https://doi.org/10.1289/ehp.1206185>
- Krall, J.R., Mulholland, J.A., Russell, A.G., Balachandran, S., Winkquist, A., Tolbert, P.E., Waller, L.A., Sarnat, S.E., 2017. Associations between Source-Specific Fine

- Particulate Matter and Emergency Department Visits for Respiratory Disease in Four U.S. Cities. *Environ. Health Perspect.* 125, 97–103. <https://doi.org/10.1289/EHP271>
- Lamsal, L.N., Janz, S.J., Krotkov, N.A., Pickering, K.E., Spurr, R.J.D., Kowalewski, M.G., Loughner, C.P., Crawford, J.H., Swartz, W.H., Herman, J.R., 2017. High-resolution NO<sub>2</sub> observations from the Airborne Compact Atmospheric Mapper: Retrieval and validation. *J. Geophys. Res.-Atmospheres* 122, 1953–1970. <https://doi.org/10.1002/2016JD025483>
- Li, H.Z., Dallmann, T.R., Gu, P., Presto, A.A., 2016. Application of mobile sampling to investigate spatial variation in fine particle composition. *Atmos. Environ.* 142, 71–82. <https://doi.org/10.1016/j.atmosenv.2016.07.042>
- Li, H.Z., Dallmann, T.R., Li, X., Gu, P., Presto, A.A., 2017. Urban Organic Aerosol Exposure: Spatial Variations in Composition and Source Impacts. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.7b03674>
- Liu, L., Urch, B., Poon, R., Szyszkowicz, M., Speck, M., Gold, D.R., Wheeler, A.J., Scott, J.A., Brook, J.R., Thorne, P.S., Silverman, F.S., 2015. Effects of Ambient Coarse, Fine, and Ultrafine Particles and Their Biological Constituents on Systemic Biomarkers: A Controlled Human Exposure Study. *Environ. Health Perspect.* 123, 534–540. <https://doi.org/10.1289/ehp.1408387>
- Martin, F., Fileni, L., Palomino, I., Vivanco, M.G., Garrido, J.L., 2014. Analysis of the spatial representativeness of rural background monitoring stations in Spain. *Atmospheric Pollut. Res.* 5, 779–788. <https://doi.org/10.5094/APR.2014.087>
- Matte, T.D., Ross, Z., Kheirbek, I., Eisl, H., Johnson, S., Gorczynski, J.E., Kass, D., Markowitz, S., Pezeshki, G., Clougherty, J.E., 2013. Monitoring intraurban spatial patterns of multiple combustion air pollutants in New York City: Design and



- implementation. *J. Expo. Sci. Environ. Epidemiol.* 23, 223–231. <https://doi.org/10.1038/jes.2012.126>
- Meier, R., Eeftens, M., Aguilera, I., Phuleria, H.C., Ineichen, A., Davey, M., Ragettli, M.S., Fierz, M., Schindler, C., Probst-Hensch, N., Tsai, M.-Y., Kuenzli, N., 2015. Ambient Ultrafine Particle Levels at Residential and Reference Sites in Urban and Rural Switzerland. *Environ. Sci. Technol.* 49, 2709–2715. <https://doi.org/10.1021/es505246m>
- Mustafić, H., Jabre, P., Caussin, C., Murad, M.H., Escolano, S., Tafflet, M., Périer, M.-C., Marijon, E., Vernerey, D., Empana, J.-P., 2012. Main air pollutants and myocardial infarction: a systematic review and meta-analysis. *Jama* 307, 713–721.
- Ostro, B., Hu, J., Goldberg, D., Reynolds, P., Hertz, A., Bernstein, L., Kleeman, M.J., 2015. Associations of Mortality with Long-Term Exposures to Fine and Ultrafine Particles, Species and Sources: Results from the California Teachers Study Cohort. *Environ. Health Perspect.* 123, 549–556. <https://doi.org/10.1289/ehp.1408565>
- Patton, A.P., Zamore, W., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2015. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. *Environ. Sci. Technol.* 49, 6051–6060. <https://doi.org/10.1021/es5061676>
- Piersanti, A., Vitali, L., Righini, G., Cremona, G., Ciancarella, L., 2015. Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollut. Res.* 6, 953–960. <https://doi.org/10.1016/j.apr.2015.04.005>
- Pope, C.A.I., Dockery, D.W., 2006. Health Effects of Fine Particulate Air Pollution: Lines that Connect. *J. Air Waste Manag. Assoc.* 56, 709–742. <https://doi.org/10.1080/10473289.2006.10464485>

- Pope, C.A.I., Ezzati, M., Dockery, D.W., 2009. Fine-Particulate Air Pollution and Life Expectancy in the United States. *N. Engl. J. Med.* 360, 376–386.
- Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging. *Science* 315, 1259–1262. <https://doi.org/10.1126/science.1133061>
- Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon Particulate Matter in New Delhi, India. *Environ. Sci. Technol.* 47, 12903–12911. <https://doi.org/10.1021/es401489h>
- Schwartz, J., Dockery, D.W., Neas, L.M., 1996. Is daily mortality associated specifically with fine particles? *J. Air Waste Manag. Assoc.* 1995 46, 927–939.
- Simon, M.C., Hudda, N., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2017. Comparisons of traffic-related ultrafine particle number concentrations measured in two urban areas by central, residential, and mobile monitoring. *Atmos. Environ.* 169, 113–127. <https://doi.org/10.1016/j.atmosenv.2017.09.003>
- Solomon, P.A., Crumpler, D., Flanagan, J.B., Jayanty, R.K.M., Rickman, E.E., McDade, C.E., 2014. U.S. National PM<sub>2.5</sub> Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of networks. *J. Air Waste Manag. Assoc.* 64, 1410–1438. <https://doi.org/10.1080/10962247.2014.956904>
- Stafoggia, M., Schneider, A., Cyrus, J., Samoli, E., Andersen, Z.J., Bedada, G.B., Bellander, T., Cattani, G., Eleftheriadis, K., Faustini, A., Hoffmann, B., Jacquemin, B., Katsouyanni, K., Massling, A., Pekkanen, J., Perez, N., Peters, A., Quass, U., Yli-Tuomi, T., Forastiere, F., Group, on behalf of the U.S., 2017. Association Between

- Short-term Exposure to Ultrafine Particles and Mortality in Eight European Urban Areas. *Epidemiology* 28, 172–180. <https://doi.org/10.1097/EDE.0000000000000599>
- Strak, M., Janssen, N., Beelen, R., Schmitz, O., Vaartjes, I., Karssenberg, D., van den Brink, C., Bots, M.L., Dijst, M., Brunekreef, B., Hoek, G., 2017. Long-term exposure to particulate matter, NO<sub>2</sub> and the oxidative potential of particulates and diabetes prevalence in a large national health survey. *Environ. Int.* 108, 228–236. <https://doi.org/10.1016/j.envint.2017.08.017>
- Urch, B., Brook, J.R., Wasserstein, D., Wasserstein, D., Brook, R.D., Brook, R.D., Rajagopalan, S., Corey, P., Silverman, F., 2008. Relative Contributions of PM<sub>2.5</sub> Chemical Constituents to Acute Arterial Vasoconstriction in Humans. *Inhal. Toxicol.* 16, 345–352.
- Urch, B., Silverman, F., Corey, P., Brook, J.R., Lukic, K.Z., Rajagopalan, S., Brook, R.D., 2005. Acute Blood Pressure Responses in Healthy Adults During Controlled Air Pollution Exposures. *Environ. Health Perspect.* 113, 1052–1055.
- Verma, V., Fang, T., Xu, L., Peltier, R.E., Russell, A.G., Ng, N.L., Weber, R.J., 2015. Organic aerosols associated with the generation of reactive oxygen species (ROS) by water-soluble PM<sub>2.5</sub>. *Environ. Sci. Technol.* 49, 4646–4656. <https://doi.org/10.1021/es505577w>
- Vitali, L., Morabito, A., Adani, M., Assennato, G., Ciancarella, L., Cremona, G., Giua, R., Pastore, T., Piersanti, A., Righini, G., Russo, F., Spagnolo, S., Tanzarella, A., Tinarelli, G., Zanini, G., 2016. A Lagrangian modelling approach to assess the representativeness area of an industrial air quality monitoring station. *Atmospheric Pollut. Res.* 7, 990–1003. <https://doi.org/10.1016/j.apr.2016.06.002>
- Weichenthal, S., 2012. Selected physiological effects of ultrafine particles in acute

- cardiovascular morbidity. *Environ. Res.* 115, 26–36. <https://doi.org/10.1016/j.envres.2012.03.001>
- Yin, X., Dai, T., Schutgens, N.A.J., Goto, D., Nakajima, T., Shi, G., 2016. Effects of data assimilation on the global aerosol key optical properties simulations. *Atmospheric Res.* 178, 175–186. <https://doi.org/10.1016/j.atmosres.2016.03.016>
- Zhang, J.J.Y., Sun, L., Barrett, O., Bertazzon, S., Underwood, F.E., Johnson, M., 2015. Development of land-use regression models for metals associated with airborne particulate matter in a North American city. *Atmos. Environ.* 106, 165–177. <https://doi.org/10.1016/j.atmosenv.2015.01.008>
- Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Allan, J.D., Coe, H., Ulbrich, I., Alfarra, M.R., Takami, A., Middlebrook, A.M., Sun, Y.L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P.F., Salcedo, D., Onasch, T., Jayne, J.T., Miyoshi, T., Shimo, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R.J., Rautiainen, J., Sun, J.Y., Zhang, Y.M., Worsnop, D.R., 2007. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophys. Res. Lett.* 34, L13801. <https://doi.org/10.1029/2007GL029979>
- Zimmerman, N., Presto, A.A., Kumar, S.P.N., Gu, J., Hauryliuk, A., Robinson, E.S., Robinson, A.L., Subramanian, R., 2017. Closing the gap on lower cost air quality monitoring: machine learning calibration models to improve low-cost sensor performance. *Atmos Meas Tech Discuss* 2017, 1–36. <https://doi.org/10.5194/amt-2017-260>

## **Chapter 2: Application of mobile sampling to investigate spatial variation in fine particle composition**

## **Chapter 2**

# **Application of mobile sampling to investigate spatial variation in fine particle composition**

Published as Li, H.Z., Dallmann, T.R., Gu, P., Presto, A.A., 2016. Application of mobile sampling to investigate spatial variation in fine particle composition. *Atmospheric Environment* 142, 71–82. <https://doi.org/10.1016/j.atmosenv.2016.07.042>

## 2.1 Introduction

Airborne fine particulate matter (PM<sub>2.5</sub>; particles smaller than 2.5 µm diameter) continues to pose a serious threat to human health in the 21<sup>st</sup> century, especially to vulnerable groups such as the elderly and small children (Brunekreef and Holgate, 2002; Künzli et al., 2000; World Health Organization, 2006). Air pollution exposure caused deaths of around 7 million people worldwide in 2012, of which 95% were attributable to PM exposure (Brauer et al., 2012).

Metals make up a small but important part of PM<sub>2.5</sub>. Transition metals such as copper (Cu) and manganese (Mn) contribute to oxidative strengths of ambient PM<sub>2.5</sub> (Charrier and Anastasio, 2012), and may in turn drive health effects of PM<sub>2.5</sub> exposure. The US Environmental Protection Agency (EPA) identifies some PM metals as toxic species, including antimony (Sb), arsenic (As), beryllium (Be), cadmium (Cd), hexavalent chromium (Cr), cobalt (Co), lead (Pb), magnesium (Mg), nickel (Ni), and selenium (Se). PM metals identified as air toxics have known health effects associated with exposure. For example, Pb is a known neurotoxin.

This study presents measurements of PM<sub>2.5</sub> mass and composition using a mobile laboratory in Pittsburgh and surrounding Allegheny County, PA. The primary contributions of this work are: (1) to evaluate spatial and seasonal patterns of PM<sub>2.5</sub> mass and composition, (2) to compare mobile sampling data and stationary monitors, (3) to develop LUR models suitable for exposure estimation based on these data. Although LUR has been widely used to describe intra-city variations in airborne pollutants, fewer studies have applied it to PM<sub>2.5</sub> trace elements, partially because of the effort required to collect and analyze samples. This study suggests a cost-effective mobile sampling approach for collecting and mapping concentrations of PM<sub>2.5</sub> components.

## **2.2 Material and Methods**

Sampling was conducted at 36 sites in Pittsburgh and surrounding areas of Allegheny County, PA (Fig. A1, Supporting Information). The landscape is characterized by a plateau with three major river valleys (Allegheny, Monongahela, and Ohio Rivers). Many major roadways follow the river valleys. Industrial facilities are primarily located along rivers, as indicated in Fig. A1. Major industrial sources include the largest metallurgical coke oven in United States (which accounts for 31.9% emissions of all criteria pollutants in the county), a coal-fired power plant, and facilities related to the steel industry.

### **2.2.1 Mobile sampling lab setup**

The mobile laboratory was a gasoline-powered van; Tan et al. (Tan et al., 2014a) details instruments installed in the mobile laboratory. The mobile laboratory housed an on-board generator to power all instruments. Most sampling was conducted with the mobile laboratory parked at curbside. An 8 m tube was connected to the vehicle exhaust pipe to prevent self-pollution during sampling (Tan et al., 2014a).

This study focuses primarily on PM<sub>2.5</sub> mass and composition as determined from filter sampling. The sampling inlet was a 1/2" O.D. stainless steel tube installed on top of the van. Ambient air was sampled through a PM<sub>2.5</sub> cyclone. The sample stream was then divided into two lines with flow rate controlled by two identical pumps at 46 SLPM. One line had a Teflon filter (47 mm, Teflo R2PJ047, Pall-Gelman) followed by a quartz filter (47 mm, Tissuquartz 2500 QAOUP), and the other line had a bare quartz filter.

### **2.2.2 Mobile sampling overview**

We used stratification methods to select representative sites based on three factors: traffic density, proximity to industrial point sources, and elevation (Tan et al., 2014a). We identified sites as having high or low traffic based on annual average daily traffic volume



using traffic count data from the Pennsylvania Department of Transportation (PennDOT, 2012). Proximity to point sources was used to indicate the combined influence of distance to major surrounding facilities and the magnitude of their emissions. Facilities under consideration have annual PM<sub>2.5</sub> emissions larger than 50 tons (Fig. A1). Point source influenced sites were within an average distance of 1500 m to the nearest influential source. The elevation strata divided sites as either being in the river valley (elevation < 250 m) or on the plateau (upland) according to 2006 Allegheny County, PA contour data (PASDA, 2006; Fig. A2 the Supporting Information). Overall, we chose 36 sites to represent diverse urban features. Nineteen sites were labeled as high traffic, 11 heavily influenced by point source, and 12 in the valley.

We used the mobile laboratory to collect our dataset in two seasons, one in 2013 summer (August) and the other in 2013-2014 winter (December and January). We conducted sampling at each site three times in each of the two seasons. In each season, we visited every site once in separate morning (6 AM to 12 PM), evening (4 PM to 10 PM), and overnight (12 AM to 5 AM) periods (Tan et al., 2014a). Each time we would park the mobile laboratory at curbside and collect data for one hour. Ideally, we could conduct a total of 6 hours sampling at each site and collect 108 filter sets per season. For the summer campaign, we collected 103 out of a possible 108 filter sets. In the winter we collected all possible 108 filter sets. Missing filter sets were the result of operation failures (e.g., pump failures).

### **2.2.3 Data handling**

We quantified PM<sub>2.5</sub> mass and PM<sub>2.5</sub> inorganic compositions with Teflon filters, and quartz filters provided concentrations of OC and EC. The quartz filter behind Teflon filter (QBT) provided an estimate of positive artifacts on the bare quartz filter (Subramanian et al., 2004). We used gravimetric mass balance to measure PM<sub>2.5</sub> mass. Each Teflon filter was

weighed before and after sampling. Filters were allowed to equilibrate at  $34.3 \pm 1.6$  % RH and  $23.6 \pm 1.6$  °C for at least 12 hours prior to weighing.

OC and EC were analyzed via thermal-optical analysis (Sunset Laboratory Inc.). Quartz filters were baked at 550° C for at least six hours before sampling. In the OC-EC Aerosol Analyzer, samples are thermally desorbed from the filter medium under an inert helium atmosphere followed by an oxidizing atmosphere using carefully controlled heating ramps. The heating temperature profile followed the IMPROVE\_A protocol (Chow et al., 2007), and we used thermal-optical transmittance to account for pyrolysis.

We sent Teflon filters to the Crocker Nuclear Lab (University of California, Davis) for XRF analysis. XRF utilizes emission of characteristic secondary X-rays from materials, which are excited by incoming X-rays or gamma rays. XRF analysis reported the following species: Na, Mg, Al, Si, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Br, Rb, Sr, Zr, Cd, Sb, and Pb.

We collected two types of blank filters. One dynamic blank test was conducted each season. During the dynamic blank test, a HEPA filter was placed on the sample inlet of the mobile lab, and samples were collected for one hour. We also collected 12 handling blanks in summer and 18 in winter, roughly 10% of the total sample size. Handling blanks underwent all of the same handling (e.g., filter loading, unloading, and analysis) as the sampled filters, with the exception that no air sample was drawn through them. Handling and dynamic blanks had similar magnitudes. We corrected all sampling data using handling blanks.

We calculated the method detection limit (MDL) for PM components with unsampled filter blanks. MDL was defined as 3 times the standard deviation of reported filter blank concentrations. The field detection limit (FDL) was calculated as 3 times the standard deviation of handling blank concentrations. The species considered here did not differ much

between these two detection limits, and we chose FDL to be the more representative detection limit.

Concentrations below FDL were replaced with half of FDL. Extreme concentrations were defined as values less than the first quartile (25<sup>th</sup> percentile) minus four times the interquartile range or greater than the third quartile (75<sup>th</sup> percentile) plus four times the interquartile range (de Hoogh et al., 2013).

#### **2.2.4 LUR model construction and evaluation**

We developed land use regression models for PM<sub>2.5</sub>, Si, S, Cl, K, Ca, Ti, Cr, Fe, Cu, and Zn. These elements were chosen from the larger list of particle components identified with XRF based on their high abundance. Si, Ti, and Fe are major crustal elements. S is primarily present as secondary sulfate and conveys long-range transport information. K is a biomass-burning tracer. Most of ambient Cr sources in the urban atmosphere are industrial emissions (Seigneur and Constantinou, 1995). Fe, Cu, and Zn indicate traffic or industry effects (Pekney et al., 2006; Schauer et al., 2006; Sternbeck et al., 2002).

To better distinguish different kinds of predictors and help explain LUR model results, we regrouped variables extracted from ArcGIS into the following groups: traffic, industry, elevation, central reference site measurements (at corresponding sampling time), and other environmental factors (Table 2.1). Traffic variables include utility and transport land use zoning, road length in a certain buffer area, traffic density on the nearest roadway, and inverse distance to the nearest road. The industry group includes industrial land use zoning, annual pollutant emissions within a buffer, inverse distance weighted emissions, the nearest facility emission of specific pollutants, and inverse distance to the nearest facility. Elevation is the height above mean sea level of the sampling site. The central reference site group includes stationary site measurements at the corresponding mobile sampling time.

Meteorological variables such as wind speed and wind direction were used in some studies, while Abernethy et al. (Abernethy et al., 2013) suggested that adding these elements would not greatly increase model performance. We did not include these meteorological factors in our models.

We used ArcGIS-10.3 (ESRI, Redlands, CA) for geospatial analysis. Table 2.1 outlines land use variables used in LUR building. Traffic counts on roads were obtained from the Pennsylvania Department of Transportation. We classified major roads as having annual average daily traffic (AADT) greater than 5000 vehicles per day. For local roads without AADT recorded, we assigned them 100 vehicles per day (Tan et al., 2014a). Road length data came from Pennsylvania spatial data access (PASDA). Traffic type variables included diesel trucks, transit bus fuel consumption, and railways. We included the multiplication of AADT and inverse distance to the nearest road variable as a potential descriptor for near-road variations in PM component concentrations. Vehicle density was the product of AADT and road length in a certain buffer. It represented combined traffic influence similar to the product of AADT and inverse distance. We included both inverse distance and inverse square distance variables.

Elevation data came from USGS National Elevation Dataset. Land use type information came from the Allegheny County GIS group. Areas were classified as commercial, residential, industrial, agricultural, utility/transport, or vacant/forest. Minimum buffer size for these zoning variables was 100 m. For the point source group, we introduced point density variables with and without emission weighting.

We used circular buffers for all land use variables. Choices of buffer sizes were based on other LUR studies (Eeftens et al., 2012a; Zhang et al., 2015) and our understanding of pollutant distributions near sources. Large gradients can exist near roadways due to traffic

influences (Karner et al., 2010), thus buffers for traffic variables ranged from 25 m to 1,000 m. Land use zoning variables used buffers from 100 m to 5,000 m. The industry group had buffers ranging from 1,000 m to 30,000 m. The raster map for each predictor used a 5x5 m grid.

Central reference sites can be used to correct distributed or mobile samples for temporal variability via either addition or multiplication (Clougherty et al., 2013; Eeftens et al., 2012b). However, Jedynska et al. (Jedynska et al., 2014) suggested data before and after central site correction was highly correlated. Tan et al. (Tan et al., 2014b) used a Monte Carlo method to simulate the mobile sampling strategy described in this paper. In each attempt of 10,000 simulations, a random set of temporally distributed data points was sampled from each of the monitoring stations in the EPA air quality system (AQS) for a given pollutant (e.g., PM<sub>2.5</sub>) to estimate average concentration. The simulation was deemed accurate if the estimate was within 1/1.2 to 1.2 times the true mean concentration. Tan et al. (Tan et al., 2014b) found that the central site correction did not significantly decrease uncertainty in estimating long-term averages of pollutants.

An alternative approach for resolving temporal variability is to include central reference sites measurements as variables for use in LUR building (Saraswat et al., 2013). We used this approach. A reference site was operated at an urban background location on the Carnegie Mellon University (CMU) campus to measure PM<sub>2.5</sub> and black carbon (BC). PM<sub>2.5</sub> was measured with a TEOM (tapered element oscillating microbalance, Thermo Fisher Scientific 1405), and BC was monitored with an aethalometer (Magee Scientific AE31). The two reference monitors provide information on regional and local variations. PM<sub>2.5</sub> concentrations are dominated by secondary species and are often regional in nature (Subramanian et al., 2007), thus the PM<sub>2.5</sub> reference provides information on changes in

pollutant concentrations driven by large-scale factors (e.g., weather). BC concentrations are strongly influenced by local emissions (e.g., traffic) (Tan et al., 2014a). The BC reference measurements therefore provide information on more local variations and emissions.

The methodology for LUR development is the same as the ESCAPE project (Eeftens et al., 2012a). It is a supervised regression. Each variable is assumed a prior direction, either a positive or negative regression coefficient. For example, we expect road length to correlate positively with pollutant concentrations, and pollutants will generally be more abundant in the valley than in the upland. A new predictor is added to the model if it yields the highest adjusted- $R^2$  improvement and at the same time has the correct correlation direction with dependent variables. This adding process continues until no variable can increase adjusted- $R^2$  by more than 1%. Variables are removed if their p values are larger than 0.1 or variance inflation factor (VIF) is larger than 3. We used Cook's D to investigate potential outlier events. Samples with Cook's D larger than 1 were further examined.

We addressed LUR model robustness by the mean studentized prediction residual (MSPR), root mean square of studentized residuals (RMS) produced by leave-one-out cross validation (LOOCV), and LOOCV  $R^2$  (Mukerjee et al., 2009). The LOOCV approach tests model goodness at predicting test datasets. In LOOCV, one observation is deleted each time. New models are developed using remaining observations, and then used to predict the deleted one. Moran's I detects possible spatial autocorrelation of residuals. Statistical analysis was done using R (RStudio, Inc., Boston, Massachusetts, USA).

## **2.3 Results and Discussion**

### **2.3.1 Mobile sampling versus stationary sites**

One critical question we needed to address was how our mobile sampling strategy compared with stationary monitors both during the study period and in predicting long-term

average concentrations. Tan et al. (Tan et al., 2014b) simulated the sampling design used here and indicated that mobile sampling strategies can reproduce spatial patterns of pollutant concentrations but have an inherent deficiency to reproduce long-term mean concentrations due to a relatively small sampling time at each site. Additionally, there is concern that exclusively sampling on or adjacent to roadways may bias the dataset to higher concentrations, especially for traffic-related species.

Fig. 2.1 shows boxplot comparisons between mobile sampling data collected as part of this study and measurements from stationary monitors during concurrent sampling periods. There are eight regulatory PM<sub>2.5</sub> monitors in the study domain. Three (Lawrenceville, Harrison, and North Park) are urban or suburban population-based monitors. Four (Avalon, Clairton, North Braddock, and South Allegheny) are industrially dominated near source monitors. One (South Fayette) is the regional transport site. Mobile sampling was conducted at 36 sites as described above, and these samples spanned the spatial range covered by the stationary sites. Table A1 (Supplemental Information) lists statistical p values (Mann-Whitney U test) for comparisons of concurrent measurements between mobile and stationary measurements.

Overall, average PM<sub>2.5</sub> concentrations (mean = 14.6 µg/m<sup>3</sup>) measured with mobile sampling were higher than the annual mean concentrations (10.2 µg/m<sup>3</sup>) measured by the stationary monitoring sites. The data therefore suggest that the mobile measurements using the sampling strategy described here overestimate long-term average PM<sub>2.5</sub> mass concentrations for the study domain. As noted by Tan et al. (Tan et al., 2014b), this is in part due to the short sampling time at each site (6 hours) distributed across two seasons (summer and winter). PM<sub>2.5</sub> concentrations in Pittsburgh are highest during the summer months, driven largely by secondary sulfate (Wittig et al., 2004). Thus constructing an annual average

concentration from the mean of summer and winter mobile samples should be expected to lead to a slightly higher estimate.

The differences between mobile PM<sub>2.5</sub> samples and concurrent measurements at the stationary sites are not statistically significant ( $p>0.05$ ; Fig. 2.1 and Table A1). Thus, while the sampling strategy used here overestimates annual average PM<sub>2.5</sub> mass, there is not a systematic error between mobile and stationary measurements.

Fig. 2.1 also shows comparisons between mobile and stationary measurements for four particle components: sulfate, OC, EC, and Zn. Sulfate loading in the mobile samples is inferred from S measured by XRF and assuming that all S exists as sulfate. OC concentrations are determined from bare quartz minus QBT filters. There are only two PM speciation sites in the study domain: Lawrenceville (urban background) and South Allegheny (industrially dominated). South Allegheny exhibits consistently higher concentrations of sulfate, OC, EC, and Zn than Lawrenceville due to nearby emissions associated with steelmaking, metallurgical coke production, and diesel trucks.

Mobile measurements of sulfate, OC, and Zn are not statistically different than measurements of these particle components at either Lawrenceville or South Allegheny.

Agreement between mobile and stationary PM<sub>2.5</sub>, OC, sulfate, and Zn mass concentrations is not surprising. PM<sub>2.5</sub>, OC, and sulfate are dominated by secondary production, rather than primary emissions, in the ambient environment (Tang et al., 2004). Thus these pollutants are more regional. Zn is expected to be strongly influenced by local emissions, though some of the major sources, such as tire wear, are ubiquitous.

Mobile measurements of EC were not statistically different from EC measured at South Allegheny ( $p>0.05$ ), but were statistically higher than Lawrenceville ( $p=0.02$ ). EC concentrations at Lawrenceville are expected to be dominated by traffic emissions, whereas



the South Allegheny site is strongly influenced by local industrial emissions and associated diesel traffic (Cabada et al., 2002). Mobile sampling EC reflects contributions from both traffic and industrial emissions. Mobile measurements of EC may be larger relative to Lawrenceville because of higher traffic contributions on roads.

### **2.3.2 Abundance of selected particle components**

Metals and crustal species have low concentrations compared with OC or sulfate in PM<sub>2.5</sub> (Fig. 2.2). The most abundant trace element was Fe, with concentrations less than 0.1  $\mu\text{g}/\text{m}^3$  on average. Other crustal elements associated with suspended dust such as Si, K, and Ca were less abundant than Fe, but more abundant than non-crustal trace metals such as Cu, Cr, and Zn.

Fig. 2.2 compares concentrations of trace species in summer and winter. In terms of seasonal difference, Si, Ca, Ti, and Cr had statistically higher concentrations in summer using the Mann-Whitney U test, but Cl, K, Fe, Cu, and Zn were not statistically different between seasons. This suggests different source categories or source strengths for Si, Ca, Ti, and Cr in different seasons.

Fig. 2.3 shows correlations between particle components in summer and winter. The crustal elements Si, Ca, and Ti were correlated in both seasons. The highest Pearson's  $r$  was between Si and Ca in both seasons. This suggests that the higher concentrations of Si, Ca, and Ti observed in the summer are the result of a larger source strength (e.g., suspended dust) rather than a different set of sources in two seasons.

Cr concentrations were also significantly higher in summer than in winter. However, Cr is poorly correlated with all other species shown in Fig. 2.3 ( $r < 0.4$ ). The primary sources of Cr in the study domain are specialty steelmaking and metal working (Pekney et al., 2006), so higher Cr concentrations in the summer may indicate increased emissions from or activity

at these facilities in summer months.

K is another crustal element. It is also a tracer for biomass burning. K has a modest correlation with other crustal species in the summer ( $r \sim 0.5$  for Si and Ca), but this correlation is weakened in the winter ( $r < 0.3$ ). The weaker correlation between K and other crustal elements is likely due to biomass burning for heat or recreation during winter (Robinson et al., 2006). K concentrations were not higher in winter than summer (though the IQR is larger), suggesting that biomass burning is not a major contributor to PM<sub>2.5</sub> mass in Pittsburgh in the winter (Cabada et al., 2002). The apparent lack of seasonal variation in K concentrations could also indicate competing effects – less dust in winter (as shown by Si and Ca) but extra K from biomass burning.

Iron is a crustal element, and is therefore correlated with other crustal elements such as Si and Ca. Fe concentrations are higher in summer (though not statistically significant), and this may be due to higher dust concentrations as with Si and Ca. In Pittsburgh, there are also important industrial sources of Fe. Three steelmaking plants along the Monongahela River were identified as important iron sources (Pekney et al., 2006).

Non-tailpipe traffic emissions (e.g., brake wear) and steel production are the expected major sources for Zn. Fe had the highest Pearson's  $r$  with Zn in both seasons ( $\sim 0.5$ ), suggesting similar sources (e.g., steel production) and/or source regions (e.g., roadways) for these species.

Cl had relatively high Pearson's  $r$  with Si and Ca in both seasons, with higher correlations in winter. One possible explanation could be road salt, a source for Cl in winter. Road salt mixed with road dust would be airborne together in winter.

Brake wear is a major source for Cu (Sternbeck et al., 2002), and there are no major Cu point sources in the study area. Cu has a weak correlation with Zn, which is also emitted

from traffic, but is otherwise poorly correlated with other species.

Fig. 2.3 also includes S, which is assumed to exist exclusively as sulfate in the ambient atmosphere. Sulfate is dominated by secondary production, and strong correlation with other species listed in Fig. 2.3, which are dominated by primary emissions, was not observed.

### 2.3.3 Traffic indicators

Some metals are known to have traffic sources, such as Cu, Ba, Sb, and Zn. The first three are mainly derived from brake wear emissions instead of fuel combustion, and tire wear and tailpipe emissions are important sources for Zn. These metals have therefore been used as traffic indicator species in previous studies (Councell et al., 2004; Lough et al., 2005; Sternbeck et al., 2002). However, the data presented here suggest that metals such as these are weak traffic indicators in the ambient environment.

Fig. 2.4 shows a boxplot comparison of three traffic related PM components at high and low traffic sampling sites as a function of season: Zn, EC, and particle-bound polycyclic aromatic hydrocarbons (PAH). PAHs are a strong indicator for vehicle emissions (Tan et al., 2014a). The boxplot and Mann-Whitney U test indicate that PAH concentrations are statistically elevated at high traffic sites in both summer and winter.

Average EC concentrations were higher at high traffic sites, but the difference between high and low traffic sites was not statistically significant. Tan et al. (Tan et al., 2014a) showed that black carbon (BC) has contributions from both traffic and industrial sources in the study region. Thus there is an elevated BC/EC background and near-road gradients are weaker than for PAHs.

Zn concentrations at high and low traffic sites were nearly identical, and did not exhibit a statistically significant difference. Zn did not have relatively high concentrations at

high traffic sites. Thus, our data indicate that Zn is a poor indicator for traffic emissions. Zn is not purely an indicator for traffic emissions, as Zn is also emitted from steelmaking industries in the study domain.

Seasonal variations for the three traffic indicators investigated in Fig. 2.4 were also different. EC and PAH exhibited higher concentrations in summer than winter, suggesting higher traffic emissions of these species in the summer. Higher concentrations were observed at both high and low traffic sites in the summer. On the other hand, Zn showed consistent ambient concentrations across the whole year. Both high and low traffic sites had similar concentrations in summer and winter. Thus, Zn had both a different spatial and seasonal pattern than two known traffic markers (EC and PAH), and itself seems to be a poor traffic indicator.

### 2.3.4 LUR models

**Model results.** We developed LUR models for PM<sub>2.5</sub>, Si, S, Cl, K, Ca, Ti, Cr, Fe, Cu, and Zn. Fig. 2.5 and Table 2.2 summarize the performance of the LUR models presented here. Models with  $R^2$  larger than 0.5 are considered adequate, and models with  $R^2$  of 0.7 or above are considered good (de Hoogh et al., 2013).

With the exception of Cu, Ti, Cl, and PM<sub>2.5</sub>, all models had an  $R^2$  of at least 0.5. To better visualize relationship between measured and LUR predicted concentration, Figs. 2.6 and 2.7 show scatter plots of measured versus predicted concentrations for each modeled species. We had good  $R^2$  ( $> 0.7$ ) for S, K, Fe, and Zn models. Si, Ca, and Cr models had acceptable  $R^2$  ( $0.5 < R^2 < 0.7$ ) results. The Cu model performed the worst ( $R^2 = 0.22$ ). The PM<sub>2.5</sub> model has a fair  $R^2$  of 0.46.

We examined model performance using MSPR and RMS in Table 2.2. The ideal result is 0 for MSPR, and 1 for RMS. MSPR for all models was close to 0 (-0.01 – 0.04). The

highest RMS was in the Ca model, which was 1.08.

We also calculated the LOOCV  $R^2$  for each model. A  $R^2$  difference of 0.20 or smaller between the original and LOOCV model implies stable performance (Eeftens et al., 2012a). Every model was stable except Cr and Ti, which had 0.24 and 0.23  $R^2$  reductions, respectively. Based on Moran's I, we only saw spatial autocorrelation of residuals in S and Cr models ( $p < 0.05$ ).

Fig. 2.5 shows the model  $R^2$  and relative  $R^2$  contribution of different predictor groups. Overall, traffic related variables are the biggest contributor to model spatial variability with an average  $R^2$  of 0.20 (SD=0.20). The industry group comes second in importance ( $R^2 \sim 0.15$ ).

**LUR model discussion.** Overall, LUR generated good results for modeled species. Traffic related variables explained the biggest portion of spatial heterogeneity for most metals. This result was consistent with ESCAPE (de Hoogh et al., 2013), and conformed to expectations for traffic related pollutants. The industry group was the second largest contributor to overall  $R^2$ . ESCAPE studies did not have such high industry significance. This could be explained by the fact that ESCAPE studies did not incorporate detailed facility emissions in model development, or industrial emissions being more important in our study domain. A similar magnitude of industry influence was observed in LUR studies in Calgary, Alberta, Canada (Zhang et al., 2015), another area with significant industrial emissions.

PM<sub>2.5</sub> in eastern US cities is affected by long-range transport from the Ohio Valley plus urban excess (Tang et al., 2004). The PM model illustrated the combined effects of regional transport, traffic, and industry on Pittsburgh air quality. The central reference site had the largest  $R^2$  contribution to the PM<sub>2.5</sub> LUR. The model also included one industrial

variable and one traffic variable.

In addition to the model validation described above, we compared predictions from the PM<sub>2.5</sub> model to an independent data set. Fig. 2.7 shows the LUR predictions of PM<sub>2.5</sub> at the stationary EPA monitors in the model domain. The data in Fig. 2.7 show a consistent over prediction of the LUR model, reflecting the higher PM measured in mobile samples. In general, the over prediction is  $\sim 1.5 \mu\text{g m}^{-3}$  in suburban and urban background locations, and higher in areas near industrial sources. One exception is the industrially dominated Liberty site, where the LUR model under predicts by approximately  $1.5 \mu\text{g m}^{-3}$ .

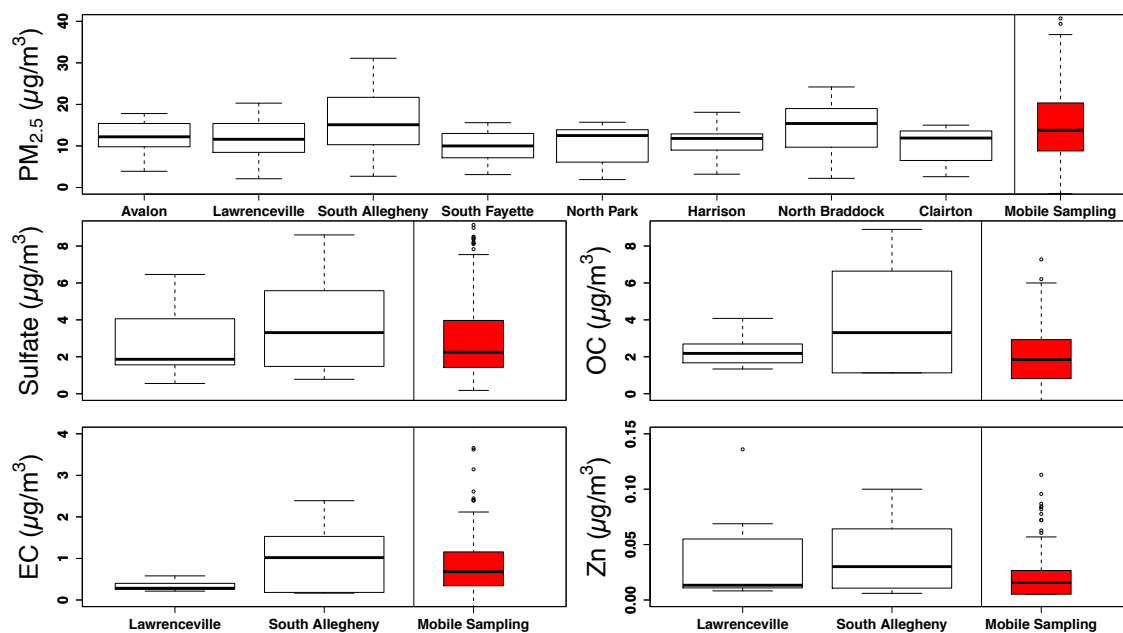
The LUR model preserves the spatial pattern observed in the EPA monitors. The Spearman rank correlation coefficient ( $r$ ) is 0.60 for predictions at the eight EPA sites, and increases to 0.93 if the Liberty site is excluded. Thus, the PM<sub>2.5</sub> LUR captures the pattern of spatial variations, but over predicts the absolute PM<sub>2.5</sub> concentrations and the magnitude of spatial gradients. This over prediction results in part from the short sample time at each site and the subsequent difficulty in determining an accurate annual average concentration from data collected in two seasons, as well as from the mobile PM<sub>2.5</sub> measurements used to build the LUR model being slightly higher than measurements at the stationary sites (Fig. 2.1).

The Zn model had an  $R^2$  of 0.7. Both traffic and industry group variables had a small  $R^2$  contribution (0.03). Most industrial facilities are located along rivers at low elevation, and 0.18  $R^2$  from the elevation factor may be a sign of industry influence. The other variables group was the largest contributor of  $R^2$ , which suggested complexity of Zn sources. There was 0.17  $R^2$  from central reference site BC. Steel mills emit Zn, and steel production is also a major source for BC in Pittsburgh (Cabada et al., 2002; Pekney et al., 2006). The link of Zn with reference site BC in the LUR model may indicate a common origin of steel mills.

K models performed worse than other species in ESCAPE studies (de Hoogh et al., 2013), and researchers explained that the addition of a biomass burning predictor would potentially improve model performance. In our work, K was closely correlated with central reference site BC, suggesting an influence from local sources. The central reference site group explained 0.43  $R^2$ . Source apportionment with positive matrix factorization (PMF) in Pittsburgh identified two important BC and K source factors as biomass burning and traffic (Pekney et al., 2006). The biomass burning PMF factor was prevalent in October and November while we collected summer filters in August and winter filters in December and January. The traffic factor was rich in K and BC, and did not show a seasonal trend (Pekney et al., 2006). As noted above, the Mann-Whitney U test did not find significant difference between summer and winter K in our data, suggesting a weak influence of biomass burning on K.

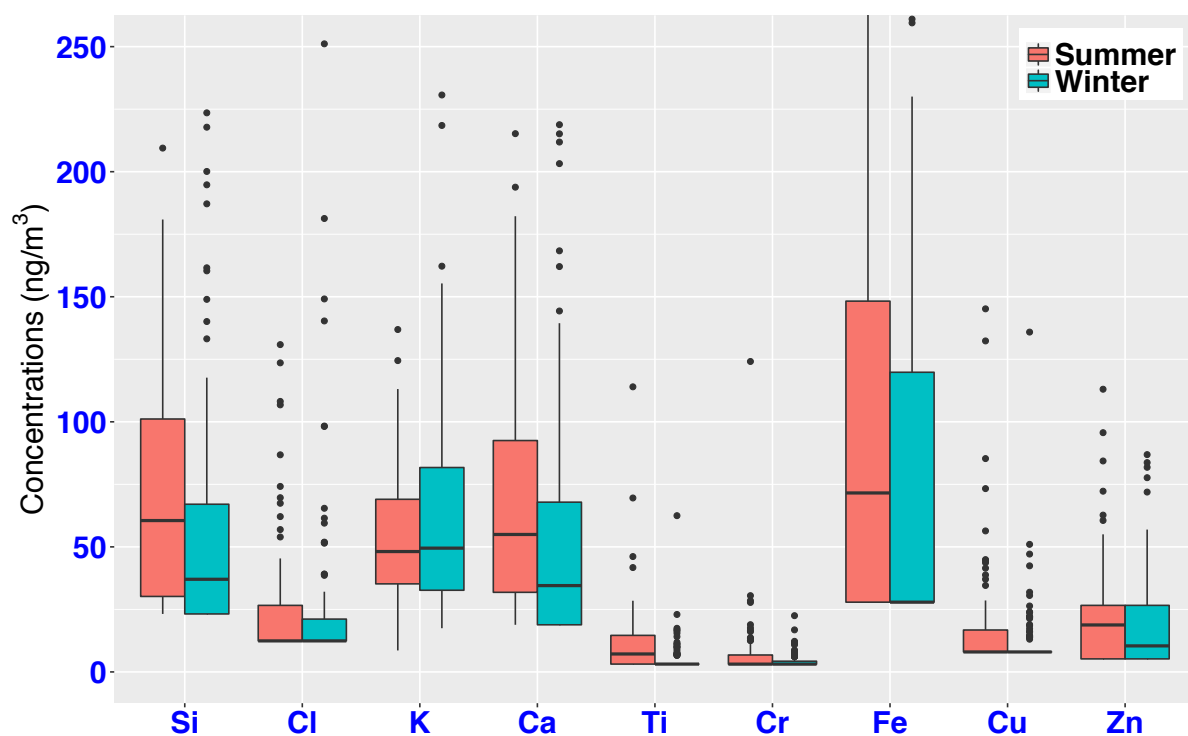
Considering the relatively small intra-urban variability of the secondary PM component sulfate, our S model performed well ( $R^2 = 0.7$ ). Central reference site PM was the most important predictor, reflecting the regional nature of secondary sulfate in Pittsburgh (Wittig et al., 2004).

Cr sources in the urban environment are mostly due to industrial emissions (Seigneur and Constantinou, 1995). In our Cr model, industry factors accounted for 0.15  $R^2$ , around one fourth of the total  $R^2$ . Traffic explained more spatial variability. However, the utility transport land use variable in the traffic group referred to railways, and industrial facilities in the valley are major users of the railroad network to transport raw materials or products. Thus, the traffic variables in the Cr model may also suggest facility emissions.

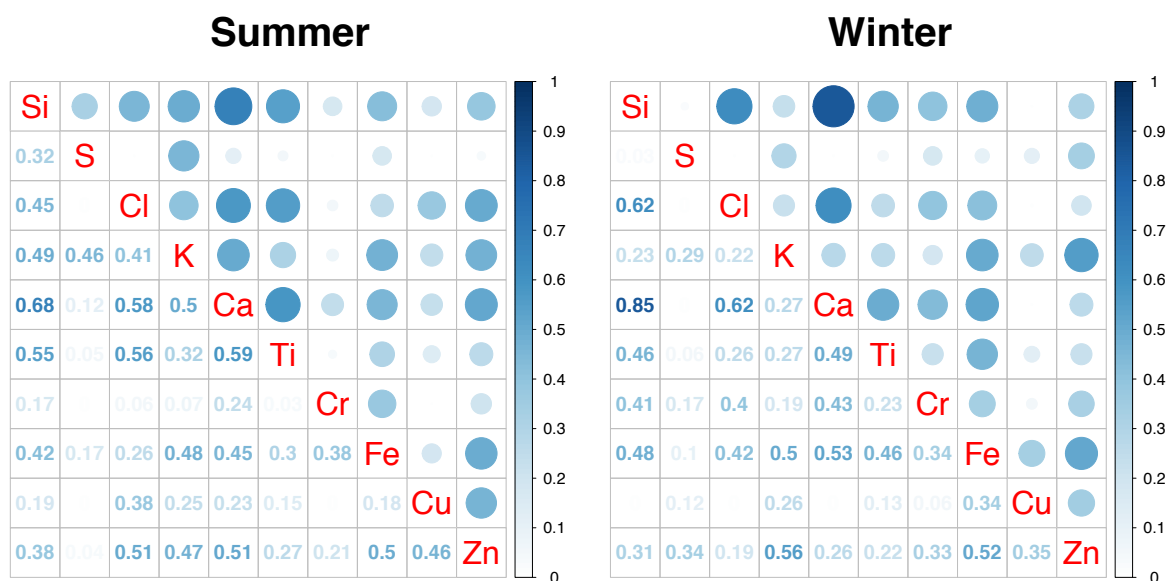


**Fig. 2.1.** Boxplot comparisons of  $PM_{2.5}$  mass and major fine particle components between mobile sampling and stationary sites during concurrent sampling periods. Mobile sampling data are in red. The top and the bottom of the box represent the 75th and 25th percentiles. The line inside the box is the median. The outer line extends to the most extreme concentrations not classified as outliers. Dots indicate outliers. There are 8 EPA monitors reporting  $PM_{2.5}$  mass while only 2 (Lawrenceville, South Allegheny) report PM speciation data.

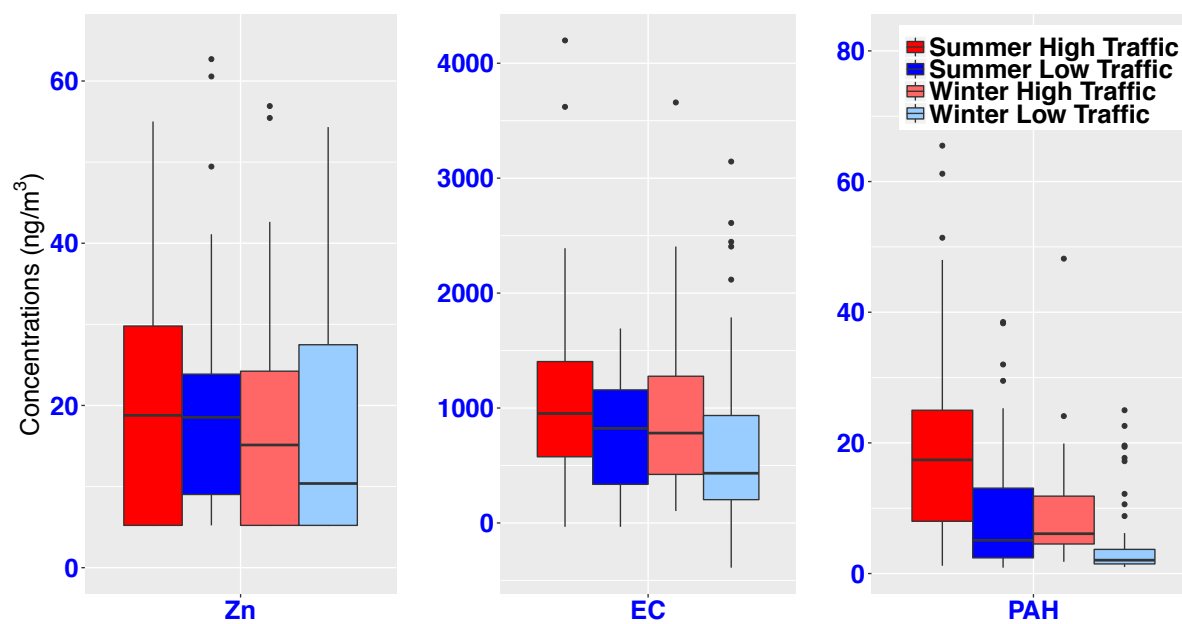




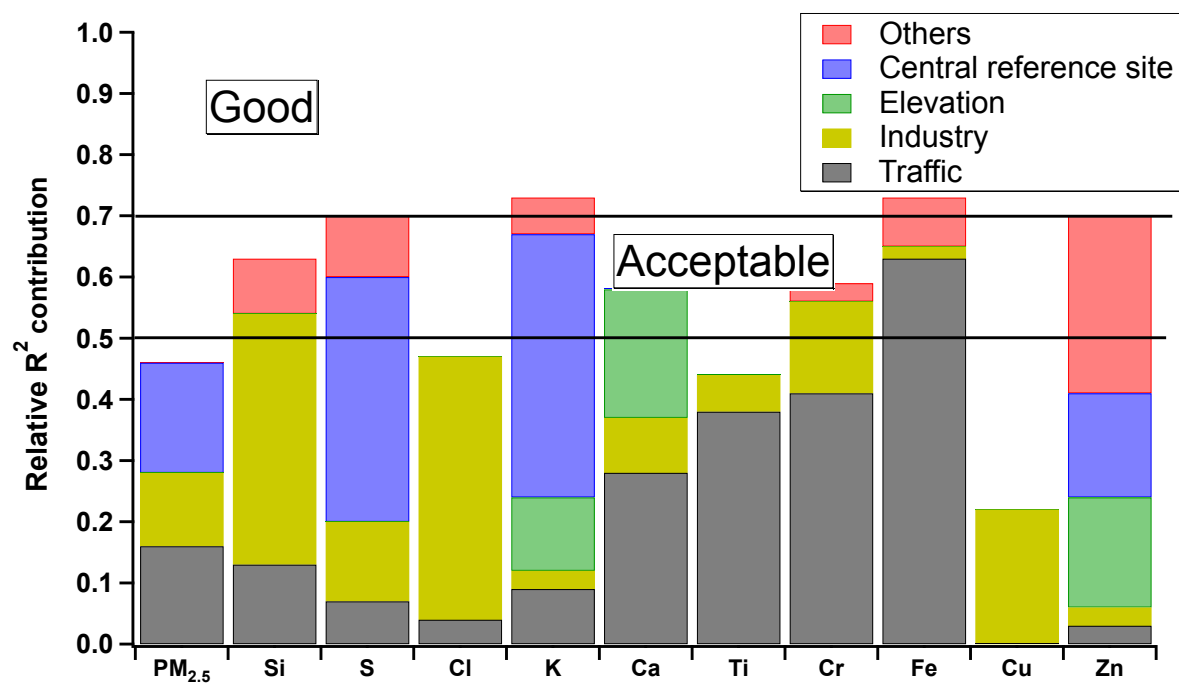
**Fig. 2.2.** Box-whisker plot of PM component concentrations measured in two seasons. Boxes, whiskers, and dots represent the same quantities as in Figure 2.1.



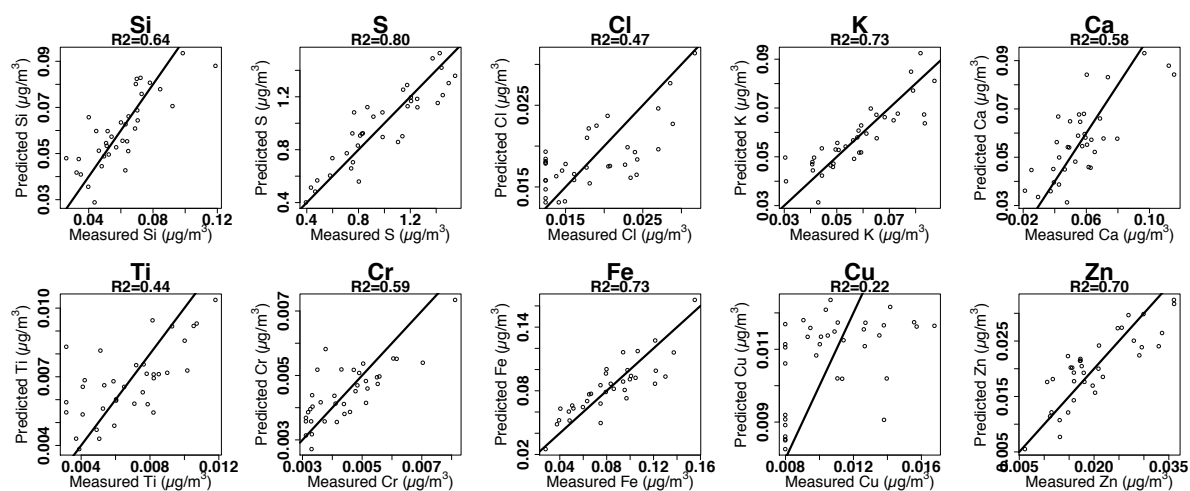
**Fig. 2.3.** Correlation of species in two mobile sampling sessions (summer and winter). Color and darkness of circles in the upper triangle represent correlation direction and magnitude. Numbers in the lower triangle are the corresponding Pearson's  $r$ . Both plots have the same legend.



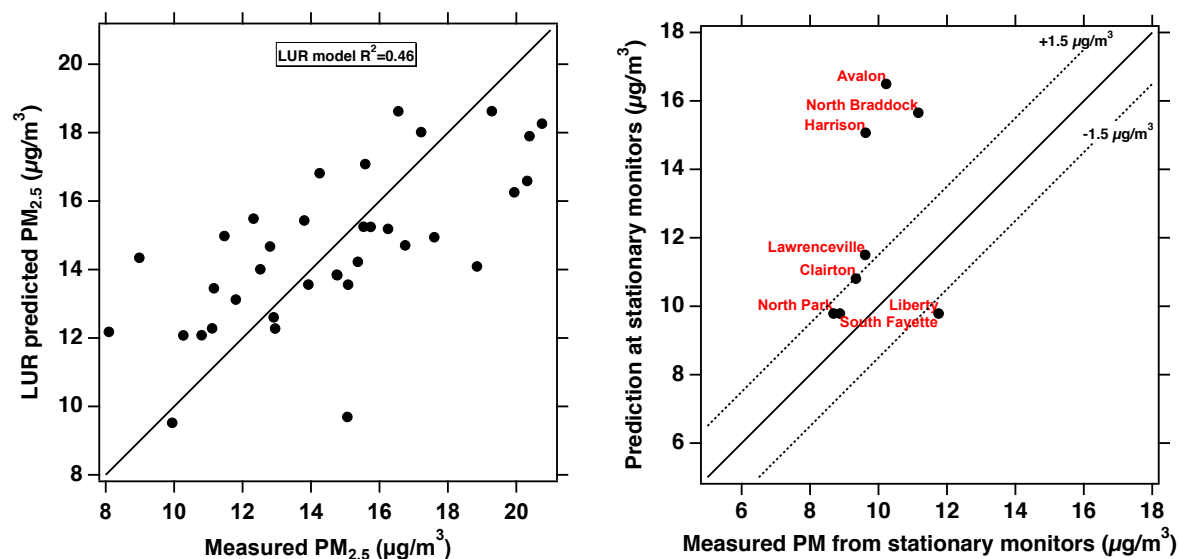
**Fig. 2.4.** Seasonal comparison of Zn, EC, and PAH based on traffic intensity.



**Fig. 2.5.** LUR Model  $R^2$  and relative portion of variability explained by different groups of sources.



**Fig. 2.6.** Measured and predicted concentrations for LUR modeled species. The bold line in each plot is the 1:1 line, and  $R^2$  values are listed under the element names.



**Fig. 2.7.** LUR prediction for mobile sampling  $PM_{2.5}$  (left) and 8 EPA monitors in the study area (right). Black solid lines in both plots are 1:1 line.

**Table 2.1.** General description of predictors used in LUR models.

Major Category	ArcGIS Category	Units	Circular Buffer Radius (m)	Description	Variable Abbreviations
Traffic	Road length	m	25, 50, 100, 300, 500, 1000	Length of all roads	RDALL
				Length of major roads	RDMAJ
	Inverse distance to the nearest road	m <sup>-1</sup>	NA	Inverse distance to nearest road	DISTINVALL
				Inverse square distance to nearest road	DISTINVALL2
				Inverse distance to nearest major road	DISTINVMAJ
				Inverse square distance to nearest major road	DISTINVMAJ2
				AADT on nearest road	ALLAADT
				AADT on nearest major road	MAJAAADT
	Annual average daily traffic (AADT) on nearest road	veh/day	NA	Diesel truck AADT on nearest road	ALLDIESAADT
				Diesel truck AADT on nearest major road	MAJDIESALLAADT
				Vehicle density on all roads	VEHDENSALL
				Vehicle density on major roads	VEHDENSMAJ
	Vehicle density	veh m/day	25, 50, 100, 300, 500, 1000	Diesel truck density on all roads	TRKDENSALL
				Diesel truck density on major roads	TRKDENSMAJ
				Bus fuel consumption	BUSFC
	Rail length	m	25, 50, 100, 300, 500, 1000	Rail length	RAIL
	Traffic land use zoning	m <sup>2</sup>	100, 300, 500, 1000, 5000	Utility/transport land use area	LUUfTr
Industry	Point density of industry sources (NEI)	km <sup>-2</sup>	1000, 1500, 3000, 5000, 7500, 10000, 15000, 20000, 30000	Number of facilities per unit area	PointDe_NEI
		lb km <sup>-2</sup>		Annual pollutant emissions per unit area	PointDe_NEI_Popu
	Inverse distance to nearest industrial source	m <sup>-1</sup>	NA	Euclidean inverse distance to nearest facility	EucDistinv
				Euclidean inverse square distance to nearest facility	EucDistinv2
	Pollution emission	lb	NA	Pollution emission at nearest corresponding facility	EucAllo
	Inverse distance weighted annual emissions	lb	NA	Inverse distance weighted annual emissions	IDW IDW2
	Industry land use zoning	m <sup>2</sup>	100, 300, 500, 1000, 5000	Industrial land use area	LUINDUS
Central reference site	Central reference site	μg/m <sup>3</sup>	NA	Corresponding PM measurements at the central site	CSMPM
				Corresponding BC measurements at the central site	CSMBC
Others	Land use zoning variables	m <sup>2</sup>	100, 300, 500, 1000, 5000	Elevation	Elevation
				Residential land use area	LURES
				Commercial land use area	LUCOMM
				Agricultural land use area	LUAGRI
	Population		100, 300, 500, 1000, 5000	Vacant/Forest land use area	LUVaFo
				Number of inhabitants	POP
				Number of households	HOUS

**Table 2.2.** Detailed LUR model validation results. In Moran's I column, values in parentheses are p values. For other columns, they indicate coefficients in LUR models.

Element	R <sup>2</sup>	adjusted R <sup>2</sup>	LOOCV R <sup>2</sup>	Traffic	Industry	Elevation	Central reference site	Others	Intercept	Moran's I	MSPR	RMS
PM <sub>2.5</sub>	0.46	0.4	0.34	RAIL500(2.07E-03)	LUINDUS1000(8.65E-06)		CSMPM(5.86E-01)		4.51	-0.11(0.41)	0.00352	1.02
PointDe_NEI_PM_Popu_5000(1.44E-06)												
Si	0.64	0.57	0.49	LUU0T1000(1.32E-07)	LUINDUS1000(5.17E-08)			LUVaFs5000(-2.63E-09)	8.22E-02	0.12(0.14)	0.00228	1.05
PointDe_NEI_PM_Popu_7500(5.16E-06)												
S	0.7	0.63	0.53	RAIL50(3.62E-03)	LUINDUS100(1.47E-05)		CSMPM(6.38E-02)	LURES5000(1.60E-08)	-3.62E-01	0.18(0.04)	-0.00122	1.04
Idw_PM_1(2.25E-06)												
PointDe_NEI_PM_Popu_3000(1.93E-07)												
Cl	0.47	0.39	0.28	RAIL300(5.62E-06)	LUINDUS300(6.95E-08)				0.013	-0.10(0.48)	0.00348	1.01
Euclideanv2_PM_2(2.88E-02)												
Euclideanv2_PM_2(2.88E-02)												
K	0.73	0.68	0.6	DISTINVAL1(6.85E-02)	LUINDUS500(1.33E-07)	Elevation(-9.84E-05)	CSMB(4.30E-02)	LUCOMM100(6.97E-07)	0.0374	-0.16(0.20)	0.01	1.05
RAIL50(1.50E-04)												
Ca	0.58	0.53	0.36	LUU0T100(3.63E-06)	LUINDUS1000(5.06E-08)	Elevation(-1.92E-04)			0.0983	0.09(0.22)	0.0095	1.08
MAIDIESAADT_DIS(1.19E-04)												
RAIL50(3.73E-05)												
RAIL300(2.37E-06)												
Ti	0.44	0.35	0.21	ALLDHESAADT(1.21E-06)	LUINDUS300(3.52E-08)				0.00381	-0.12(0.36)	-0.00417	1.05
RDMAL100(4.90E-06)												
MAIDIESAADT_DIS2(4.77E-05)												
Cr	0.59	0.52	0.35	LUU0T500(3.86E-08)	LUINDUS300(2.48E-08)			LURES300(6.66E-09)	2.26E-03	0.28(0.00)	-0.00722	1.08
MAIDIESAADT(5.78E-07)												
LUU0T1000(2.66E-07)												
Fe	0.73	0.69	0.62	VEHDENSMAL100(2.16E-04)	LUINDUS500(1.45E-07)			LUAGRI5000(-3.15E-09)	5.05E-02	0.17(0.05)	0.0358	1.03
LUU0T100(3.93E-06)												
Cu	0.22	0.19	0.14		PointDe_NEI_PM_Popu_20000(1.60E-06)				7.16E-03	0.15(0.07)	0.00995	1.02
PointDe_NEI_PM_Popu_20000(1.60E-06)												
Zn	0.7	0.63	0.56	ALLAADT(3.09E-07)	LUINDUS500(4.27E-08)	Elevation(-9.08E-05)	CSMB(1.37E-02)	LUVaFs300(-6.80E-08)	5.23E-02	-0.06(0.77)	0.0116	1.01
LUVaFs5000(-6.82E-10)												



## 2.4 Limitations

As noted above and by Tan et al. (Tan et al., 2014b), mobile sampling strategies that collect relatively small amounts of data at each sampling site, including the sampling strategy used here, may not represent long-term average concentrations of all PM species at each site. Our data reflect this. While Fig. 2.1 shows that mobile measurements of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components are not consistently over or underestimating concurrent stationary measurements, there may be inaccuracy when constructing an annual average concentration from mobile data collected in two seasons (Fig. 2.7). The error in calculated annual average concentrations is likely larger for species that have strong seasonal variations (e.g., PM<sub>2.5</sub>) than for species with relatively constant concentrations throughout the year (e.g., Zn). Nonetheless, the mobile measurements capture the spatial concentration pattern, and the mobile data are sufficient for constructing LUR models.

The number of sites for LUR development was 36. Cross validation might overestimate  $R^2$  for small training size; however, Zhang et al. (Zhang et al., 2015) concluded that a total of 25 sites could give acceptable cross validation performance for element models. Our various model validation processes also yielded reasonable results.

## 2.5 Discussion and Implication

We conducted a mobile sampling campaign in two seasons to characterize spatial variation of fine particle composition in Pittsburgh. We developed LUR models for PM<sub>2.5</sub> and ten trace species. Most elements displayed significant intra-urban variation with influence from traffic emissions and industrial production.

This manuscript presents the first successful LUR study for particle composition in North America using a cost effective mobile sampling approach. Mobile sampling is convenient to implement to study spatial variability with high resolution. Despite the inherent

deficiency of mobile sampling to estimate long-term average concentrations, we found that our dataset was not statistically significant compared with stationary reference monitors and we were able to derive acceptable LUR models. This study provides support for future LUR attempts with mobile sampling.

## 2.6 References

- Abernethy, R.C., Allen, R.W., McKendry, I.G., Brauer, M., 2013. A land use regression model for ultrafine particles in Vancouver, Canada. *Environ. Sci. Technol.* 47, 5217–5225. doi:10.1021/es304495s
- Brauer, M., Amann, M., Burnett, R.T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S.B., Krzyzanowski, M., Martin, R.V., Van Dingenen, R., van Donkelaar, A., Thurston, G.D., 2012. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. *Environ. Sci. Technol.* 46, 652–660. doi:10.1021/es2025752
- Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. *Lancet* 360, 1233–1242. doi:10.1016/S0140-6736(02)11274-8
- Cabada, J.C., Pandis, S.N., Robinson, A.L., 2002. Sources of atmospheric carbonaceous particulate matter in Pittsburgh, Pennsylvania. *J. Air Waste Manag. Assoc.* 52, 732–741. doi:10.1080/10473289.2002.10470811
- Charrier, J.G., Anastasio, C., 2012. On dithiothreitol (DTT) as a measure of oxidative potential for ambient particles: evidence for the importance of soluble transition metals. *Atmos. Chem. Phys.* 12, 9321–9333. doi:10.5194/acp-12-9321-2012
- Chow, J.C., Watson, J.G., Chen, L.W.A., Chang, M.C.O., Robinson, N.F., Trimble, D., Kohl, S., 2007. The IMPROVE\_A temperature protocol for thermal/optical carbon analysis: maintaining consistency with a long-term database. *J. Air Waste Manag. Assoc.* 57, 1014–1023. doi:10.3155/1047-3289.57.9.1014
- Clougherty, J.E., Kheirbek, I., Eisl, H.M., Ross, Z., Pezeshki, G., Gorczynski, J.E., Johnson, S., Markowitz, S., Kass, D., Matte, T., 2013. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants:

- The New York City Community Air Survey (NYCCAS). *J. Expo. Sci. Environ. Epidemiol.* 23, 232–240. doi:10.1038/jes.2012.125
- Council, T.B., Duckenfield, K.U., Landa, E.R., Callender, E., 2004. Tire-wear particles as a source of zinc to the environment. *Environ. Sci. Technol.* 38, 4206–4214. doi: 10.1021/es034631f
- de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gražulevičienė, R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsouyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Mosler, G., Nador, G., Nieuwenhuijsen, M., Pershagen, G., Peters, A., Phuleria, H., Probst-Hensch, N., Raaschou-Nielsen, O., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Weinmayr, G., Brunekreef, B., Hoek, G., 2013. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ. Sci. Technol.* 47, 5778–5786. doi:10.1021/es400156t
- Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nador, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S.V., Wolf, K., Brunekreef, B., Hoek, G., 2012a. Development of Land Use Regression models for PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European study areas; results

- of the ESCAPE project. *Environ. Sci. Technol.* 46, 11195–11205. doi:10.1021/es301948k
- Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., de Nazelle, A., de Vocht, F., Declercq, C., Dedele, A., Eriksen, K., Galassi, C., Gražulevičienė, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Kraemer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Moelter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskoevi, E., Weinmayr, G., Brunekreef, B., Hoek, G., 2012b. Spatial variation of PM<sub>2.5</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> absorbance and PM<sub>coarse</sub> concentrations between and within 20 European study areas and the relationship with NO<sub>2</sub> - Results of the ESCAPE project. *Atmos. Environ.* 62, 303–317. doi:10.1016/j.atmosenv.2012.08.038
- Hankey, S., Marshall, J.D., 2015. Land use regression models of on-road particulate air pollution (particle number, black carbon, PM<sub>2.5</sub>, particle size) using mobile monitoring. *Environ. Sci. Technol.* 49, 9194–9202. doi:10.1021/acs.est.5b01209
- Health Effects Institute, 2010. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. HEI Special Report 17, Health Effects Institute, Boston, MA.
- Henderson, S.B., Beckerman, B., Jerrett, M., Brauer, M., 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ. Sci. Technol.* 41, 2422–2428. doi:10.1021/es0606780
- Jedynska, A., Hoek, G., Wang, M., Eeftens, M., Cyrys, J., Keuken, M., Ampe, C., Beelen, R.,

- Cesaroni, G., Forastiere, F., Cirach, M., de Hoogh, K., de Nazelle, A., Nystad, W., Declercq, C., Eriksen, K.T., Dimakopoulou, K., Lanki, T., Meliefste, K., Nieuwenhuijsen, M.J., Yli-Tuomi, T., Raaschou-Nielsen, O., Brunekreef, B., Kooter, I.M., 2014. Development of land use regression models for elemental, organic carbon, PAH, and hopanes/steranes in 10 ESCAPE/TRANSPHORM European study areas. *Environ. Sci. Technol.* 48, 14435–14444. doi:10.1021/es502568z
- Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-roadway air quality: synthesizing the findings from real-world data. *Environ. Sci. Technol.* 44, 5334–5344. doi:10.1021/es100008x
- Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., Puybonnieux-Textier, V., Quénel, P., Schneider, J., Seethaler, R., Vergnaud, J.-C., Sommer, H., 2000. Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lancet* 356, 795–801. doi:10.1016/S0140-6736(00)02653-2
- Larson, T., Henderson, S.B., Brauer, M., 2009. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.* 43, 4672–4678. doi:10.1021/es803068e
- Lough, G.C., Schauer, J.J., Park, J.-S., Shafer, M.M., DeMinter, J.T., Weinstein, J.P., 2005. Emissions of metals associated with motor vehicle roadways. *Environ. Sci. Technol.* 39, 826–836. doi:10.1021/es048715f
- Mukerjee, S., Smith, L.A., Johnson, M.M., Neas, L.M., Stallings, C.A., 2009. Spatial analysis and land use regression of VOCs and NO<sub>2</sub> from school-based urban air monitoring in Detroit/Dearborn, USA. *Sci. Total Environ.* 407, 4642–4651. doi:10.1016/j.scitotenv.2009.04.030
- PASDA, 2006. Allegheny County - Contours.

[ftp://www.pasda.psu.edu/pub/pasda/allegHENYcounty/AlleghenyCounty\\_Contours2006.zip](ftp://www.pasda.psu.edu/pub/pasda/allegHENYcounty/AlleghenyCounty_Contours2006.zip)

Patton, A.P., Zamore, W., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2015.

Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area. *Environ. Sci. Technol.* 49, 6051–6060. doi: 10.1021/es5061676

Pekney, N., Davidson, C., Robinson, A., Zhou, L., Hopke, P., Eatough, D., Rogge, W., 2006.

Major source categories for PM<sub>2.5</sub> in Pittsburgh using PMF and UNMIX. *Aerosol Sci. Technol.* 40, 910–924. doi:10.1080/02786820500380271

PennDOT, 2012. Pennsylvania Traffic Counts. <http://www.pasda.psu.edu/data/padot/state/historic/PaTraffic>

Poplawski, K., Gould, T., Setton, E., Allen, R., Su, J., Larson, T., Henderson, S., Brauer, M.,

Hystad, P., Lightowlers, C., Keller, P., Cohen, M., Silva, C., Buzzelli, M., 2009.

Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. *J. Expo. Sci. Environ. Epidemiol.* 19, 107–117. doi:10.1038/jes.2008.15

Rees, S.L., Robinson, A.L., Khlystov, A., Stanier, C.O., Pandis, S.N., 2004. Mass balance closure and the Federal Reference Method for PM<sub>2.5</sub> in Pittsburgh, Pennsylvania.

*Atmos. Environ.* 38, 3305–3318. doi:10.1016/j.atmosenv.2004.03.016

Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop,

A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking organic aerosols: semivolatile emissions and photochemical aging. *Science* 315, 1259–1262. doi: 10.1126/science.1133061

Robinson, A.L., Subramanian, R., Donahue, N.M., Bernardo-Bricker, A., Rogge, W.F., 2006.

Source apportionment of molecular markers and organic aerosol. 2. Biomass smoke.

- Environ. Sci. Technol. 40, 7811–7819. doi:10.1021/es060782h
- Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. Environ. Sci. Technol. 47, 12903–12911. doi: 10.1021/es401489h
- Schauer, J.J., Lough, G.C., Shafer, M.M., Christensen, W.F., Arndt, M.F., DeMinter, J.T., Park, J.-S., 2006. Characterization of metals emitted from motor vehicles. Research Report 133, Health Effects Institute, Boston, MA.
- Seigneur, C., Constantinou, E., 1995. Chemical kinetic mechanism for atmospheric chromium. Environ. Sci. Technol. 29, 222–231. doi:10.1021/es00001a029
- Sternbeck, J., Sjödin, Å., Andréasson, K., 2002. Metal emissions from road traffic and the influence of resuspension - results from two tunnel studies. Atmos. Environ. 36, 4735–4744. doi:10.1016/S1352-2310(02)00561-7
- Subramanian, R., Donahue, N.M., Bernardo-Bricker, A., Rogge, W.F., Robinson, A.L., 2007. Insights into the primary–secondary and regional–local contributions to organic aerosol and PM<sub>2.5</sub> mass in Pittsburgh, Pennsylvania. Atmos. Environ. 41, 7414–7433. doi:10.1016/j.atmosenv.2007.05.058
- Subramanian, R., Khlystov, A.Y., Cabada, J.C., Robinson, A.L., 2004. Positive and negative artifacts in particulate organic carbon measurements with denuded and undenuded sampler configurations special issue of Aerosol Science and Technology on findings from the fine particulate matter supersites Program. Aerosol Sci. Technol. 38, 27–48. doi:10.1080/02786820390229354
- Tan, Y., Dallmann, T.R., Robinson, A.L., Presto, A.A., 2016. Application of plume analysis to build land use regression models from mobile sampling to improve model



- transferability. *Atmos. Environ.* 134, 51–60. doi:10.1016/j.atmosenv.2016.03.032
- Tan, Y., Lipsky, E.M., Saleh, R., Robinson, A.L., Presto, A.A., 2014a. Characterizing the spatial variation of air pollutants and the contributions of high emitting vehicles in Pittsburgh, PA. *Environ. Sci. Technol.* 48, 14186–14194. doi:10.1021/es5034074
- Tan, Y., Robinson, A.L., Presto, A.A., 2014b. Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. *Atmos. Environ.* 99, 333–340. doi:10.1016/j.atmosenv.2014.10.003
- Tang, W., Raymond, T., Wittig, B., Davidson, C., Pandis, S., Robinson, A., Crist, K., 2004. Spatial variations of PM<sub>2.5</sub> during the Pittsburgh air quality study. *Aerosol Sci. Technol.* 38, 80–90. doi:10.1080/02786820490442833
- Tsai, M.-Y., Hoek, G., Eeftens, M., de Hoogh, K., Beelen, R., Beregszászi, T., Cesaroni, G., Cirach, M., Cyrys, J., de Nazelle, A., de Vocht, F., Ducret-Stich, R., Eriksen, K., Galassi, C., Gražulevičienė, R., Gražulevicius, T., Grivas, G., Gryparis, A., Heinrich, J., Hoffmann, B., Iakovides, M., Keuken, M., Krämer, U., Künzli, N., Lanki, T., Madsen, C., Meliefste, K., Merritt, A.-S., Mölter, A., Mosler, G., Nieuwenhuijsen, M.J., Pershagen, G., Phuleria, H., Quass, U., Ranzi, A., Schaffner, E., Sokhi, R., Stempfelet, M., Stephanou, E., Sugiri, D., Taimisto, P., Tewis, M., Udvardy, O., Wang, M., Brunekreef, B., 2015. Spatial variation of PM elemental composition between and within 20 European study areas--Results of the ESCAPE project. *Environ. Int.* 84, 181–192. doi:10.1016/j.envint.2015.04.015
- Wittig, A.E., Anderson, N., Khlystov, A.Y., Pandis, S.N., Davidson, C., Robinson, A.L., 2004. Pittsburgh air quality study overview. *Atmos. Environ.* 38, 3107–3125. doi:10.1016/j.atmosenv.2004.03.003
- World Health Organization, 2006. WHO Air quality guidelines for particulate matter, ozone,

nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment.

Zhang, J.J.Y., Sun, L., Barrett, O., Bertazzon, S., Underwood, F.E., Johnson, M., 2015.

Development of land-use regression models for metals associated with airborne particulate matter in a North American city. *Atmos. Environ.* 122, 165–177. doi:

10.1016/j.atmosenv.2015.01.008

### **Chapter 3: Urban organic aerosol exposure: spatial variations in composition and source impacts**

# **Chapter 3**

## **Urban organic aerosol exposure: spatial variations in composition and source impacts**

Published as Li, H.Z., Dallmann, T.R., Li, X., Gu, P., Presto, A.A., 2017. Urban organic aerosol exposure: Spatial variations in composition and source impacts. Environ. Sci. Technol. <https://doi.org/10.1021/acs.est.7b03674>

### 3.1 Introduction

Organic carbon (OC) is a major component of PM with a broad range of concentrations from 0.1 to 100  $\mu\text{g}/\text{m}^3$  worldwide (Jimenez et al., 2009). National speciation networks, including the Chemical Speciation Network (CSN) and the Interagency Monitoring of Protected Visual Environments (IMPROVE), generally use thermal-optical OC/EC analysis for OC measurements. This method thermally desorbs OC captured by quartz filters at specified temperature ranges in a helium environment (Chow et al., 2007). While not chemically specific, the lumped OC fractions evolved at different temperatures – OC1, OC2, OC3, OC4, and pyrolyzed carbon (PC) – are associated with different volatility and may therefore indicate different composition, OC sources, or extents of photochemical processing. Ma et al. (2016) found that OC fractions (OCX) are related to volatility, with the most volatile material (OC1) existing almost exclusively in the vapor phase. PC has the lowest volatility and exists almost entirely in the particle phase (Ma et al., 2016; Zhu et al., 2014). OC2, OC3, and OC4 are more volatile than PC and partition between the particle and vapor phase.

Previous work performed source apportionment of OCX in order to link these OC fractions with specific sources (Kim and Hopke, 2004a, 2004b; Lee et al., 2003). Kim et al. (2004a) used positive matrix factorization (PMF) to resolve sources for data from a rural IMPROVE site. They found that PC was mostly from a secondary PM factor, and attributed the PC to water-soluble secondary organic carbon produced by oxidation of volatile precursors in the atmosphere. The gasoline factor and diesel factor identified by Kim et al. had high loadings of lower temperature carbon fractions OC2 and OC3. In a separate study, Kim et al. (2004b) performed PMF analysis for an urban IMPROVE site. The derived diesel factor was rich in OC2. Lee et al. (2003) attributed the high OC4 loading in their secondary

coal PMF factor to reflect a chemical aging process because the carbon associated with the secondary coal source would have travelled much farther than local traffic emitted OC.

Many LURs have been built for PM<sub>2.5</sub> mass concentrations (Chen et al., 2017; Eeftens et al., 2012; Hankey and Marshall, 2015; Li et al., 2016; Shi et al., 2016), and for PM components such as black carbon and specific PM metals (Brokamp et al., 2017; de Hoogh et al., 2013; Hankey and Marshall, 2015; Li et al., 2016). Few LURs have been built for organic PM constituents. Jedynska et al. (2014) built LURs for OC, polycyclic aromatic hydrocarbons (PAHs), and organic molecular markers such as steranes and hopanes in ten European cities. Hopanes, steranes, and PAHs can all be used as markers for motor vehicle emissions (Robinson et al., 2006; Tan et al., 2014a), and therefore these LURs could be used to estimate exposures to source-specific PM emissions. We are not aware of other LURs for OC, nor of any LURs built for lumped OC fractions from OC/EC analysis.

In this paper, we present measurements of OC and its thermally-resolved fractions using mobile sampling in Pittsburgh and surrounding Allegheny County, PA. The primary objectives of this work are: (1) to evaluate intracity spatial variation of OC and its composition, (2) to characterize potential sources of OC fractions, and (3) to derive LUR models of OC and OC fractions for high spatial resolution exposure estimation. We also identify certain OC fractions as indicators of exposure to fresh emissions versus more aged aerosol.

### **3.2 Material and Methods**

We chose 36 sites for mobile sampling, which were distributed in Pittsburgh and surrounding cities in Allegheny County, PA. The sampling campaign is described in detail in Li et al (2016), and the sampling locations are shown in Figure B1 of the Supporting Information. The landscape is characterized by a plateau with three major river valleys

(Allegheny, Monongahela, and Ohio Rivers). A variety of industrial facilities are located along the rivers, such as coke plants, steel manufacturing, and coal-fired electricity stations (Fig. B1, Supporting Information).

### **3.2.1 Mobile platform setup**

We used a gasoline powered van, equipped with diverse instruments (Li et al., 2016). Ambient air was pulled by mechanical pumps from a ½” O.D. stainless steel inlet on top of the van approximately 3 m above ground level.

Ambient air was sampled through a PM<sub>2.5</sub> cyclone for size selection. Then it was divided into two streams with flow rate controlled at 46 SLPM. We used Teflon filters (47 mm, Teflo R2PJ047, Pall-Gelman) and quartz filters (47 mm, Tissuquartz 2500 QAOUP). One sample line had a Teflon filter in front and a quartz filter behind (QBT – quartz behind Teflon), and the other line had one bare quartz filter (BQ).

Samples were collected with the mobile laboratory parked at curbside. An 8 m hose was connected to the vehicle tailpipe and placed in the downwind direction to avoid self-contamination. Filter sets were also collected from a highway tunnel in Pittsburgh, PA. The sampling configuration in the tunnel is the same as the setup in the mobile van (Li et al., 2016).

### **3.2.2 Sampling overview**

The sampling strategy is described in detail in Li et al (2016) and Tan et al (2014a). Briefly, we selected 36 sites based on three stratification variables – traffic intensity, proximity to major industrial sources, and elevation (Li et al., 2016). The traffic group criteria is annual average daily traffic (AADT) volume obtained from the Pennsylvania Department of Transportation (PASDA, 2017a). Locations with AADT > 2800 vehicles/day are classified as high traffic (Li et al., 2016). Proximity to industrial sources takes into account both the

distance to the nearest major industrial facility and their emissions. Facilities under consideration have annual PM<sub>2.5</sub> emissions larger than 50 tons (Fig. B1). The average distance to the nearest facility for source-impacted sites is about 1500 m. The elevation threshold of 250 m divides sites as either valley (< 250 m) or highland (> 250 m) according to 2006 Allegheny County, PA contour data (PASDA, 2017). After classification, 19 sites are high traffic density sites, 11 near point sources, and 12 in the valley. Table B2 in the Supporting Information lists the sampling sites and the strata for each site.

We visited sites in two seasons – 2013 summer and winter. For each season, we visited every site at three different time sessions on random days – morning (6 a.m. to 12 p.m.), evening (4 p.m. to 10 p.m.), and night (12 a.m. to 5 a.m.) (Li et al., 2016). For every sampling session, we parked the van at curbside and sampled for 1 hour. Ideally, we would collect 216 quartz filters for each season. We collected 206 out of 216 samples in summer because of pump failures or other errors and all 216 samples in the winter campaign.

We collected 18 tunnel filters in 2013 winter. Sampling is described in detail in Li et al (2016). The sampling time for 14 of the filter sets was 45 minutes and the remainder were sampled for 90 minutes. Filters were collected either during midday (12 p.m. – 2 p.m.) or in the afternoon rush hour (3 p.m. – 6 p.m.).

### **3.2.3 OC quantification**

We use the quartz behind Teflon (QBT) method to correct for sampling artifacts on bare quartz (BQ) filters and to determine particulate OC concentrations (Ma et al., 2016; Presto et al., 2011; Turpin et al., 2000, 1994). Two sampling artifacts are involved in this setup – positive and negative artifacts. Positive artifact is when vapors are captured by the bare quartz filter. Negative artifact is when particles are first captured by the filter but then re-evaporate. The BQ filter collects both particle and gas phase OC. The Teflon filter removes



the particle phase OC, and the quartz behind Teflon (QBT) filter collects the remaining gas phase OC. The difference between bare quartz OC loading and QBT (BQ-QBT) gives an estimate of the artifact-corrected particle phase OC. In the ambient, there is typically a larger positive artifact than negative artifact, and BQ-QBT approach provides a reasonable artifact-corrected particle phase OC concentration for our sampling setup (Subramanian et al., 2004). Before each sampling session, quartz filters are prebaked in the oven at 550<sup>0</sup> C for at least six hours to remove residual OC.

We used a Sunset OC/EC analyzer (Sunset Labs, Inc.) to measure OC concentrations using filter transmittance with the IMPROVE\_A protocol (Chow et al., 2007). OC desorbs from quartz filters at four different temperature stages in an inert helium environment and are classified as OC1 (< 140<sup>0</sup> C), OC2 (140<sup>0</sup> C – 280<sup>0</sup> C), OC3 (280<sup>0</sup> C – 480<sup>0</sup> C), and OC4 (480<sup>0</sup> C – 580<sup>0</sup>). Some OC chars and forms pyrolyzed carbon (PC). PC and elemental carbon (EC) are measured after the carrier gas changes to He/O<sub>2</sub> (O<sub>2</sub>: 10%). We used sucrose standards to calibrate the instrument prior to use every time.

We collected approximately 30 field blanks each season. Field blanks went through the same handling procedure as normal samples – filter loading, unloading, and analysis – but no sample was collected. Total carbon loading on all field blanks was lower than 1 µg/cm<sup>2</sup>, which is EPA's standard for clean blank quartz filters (Chow et al., 2010). OC on most blanks was lower than the instrument detection limit of 0.2 µg/cm<sup>2</sup>. No blank correction was made because of low signals on field blanks.

The collection area on the quartz filters was 11.34 cm<sup>2</sup>. We transformed instrument-reporting values in µg/cm<sup>2</sup> to ambient concentration in µg/m<sup>3</sup> by first multiplying by the collection area, and then dividing by the sampling volume. Under typical sampling conditions

(46 SLPM for one hour), the instrument detection limit of  $0.2 \mu\text{g}/\text{cm}^2$  would be converted to  $0.8 \mu\text{g}/\text{m}^3$ , and measurement resolution would be half the detection limit ( $0.4 \mu\text{g}/\text{m}^3$ ) (Lubin et al., 2004).

In addition to our dataset collected in Pittsburgh, we extracted OC data from the national CSN speciation network (Solomon et al., 2014). CSN data for OCX and total OC on bare quartz filters were retrieved for 2013.

### 3.2.4 LUR development

We developed land use regression (LUR) models for OC fractions on bare quartz filters (OC2, OC3, OC4, PC and total OC). These OC fractions include both particle and gas phase OC. However, quartz filters do not capture all OC vapors with unit efficiency; thus we exclude OC1, which exists almost entirely as vapors in the atmosphere and is poorly captured by quartz filters, from LUR model building (Subramanian et al., 2004; Zhao et al., 2016, 2015). OCX mentioned afterwards in this paper means OC loading on the bare quartz filter. We also derived LUR models for total particle phase OC (BQ-QBT).

LUR is an application of multilinear regression. It associates dependent variables (pollutant concentrations) with multiple independent ones (land use information) using a linear model. Formally, the model format, given  $p$  final predictors and  $n$  observations at different locations, is

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_p x_{pi} + \varepsilon_i \text{ for } i = 1, 2, \dots, n$$

Following our previous work, the LUR predictor categories are traffic, restaurants, industry, elevation, and other environmental variables (Li et al., 2016). Detailed predictor information is in Table B1 (Supporting Information). In the traffic group, we include variables such as total road length, inverse distance to the nearest road, annual average daily

traffic on the nearest road (AADT), bus fuel consumption, and traffic land use zoning area.

The restaurant and industry groups both include point density and inverse distance to the nearest source variables. The industry group also contains inverse distance weighted annual emissions and industrial land use zoning. Other land use variables include commercial and residential land use area, population density, and housing density. Elevation is the altitude of the sampling location above mean sea level.

Since we did not visit all 36 sites at same time on the same day, we need to account for temporal variations that are convolved with our spatially distributed measurements. One approach is to adjust the measured OC by either an additive or multiplicative factor based on measurements at a central reference site (Eeftens et al., 2012; Wang et al., 2013). An alternative approach is to include the central reference site measurements among the LUR predictor variables (Saraswat et al., 2013). We used the latter approach.

As in our previous LUR study (Li et al., 2016), we use hourly-resolved PM<sub>2.5</sub> measurements at an urban EPA monitoring site (Lawrenceville, Figure B1) during concurrent mobile sampling periods as an explanatory variable in our LUR models. Since PM<sub>2.5</sub> mass concentrations are dominated by secondary species and often regional in nature, variations in PM<sub>2.5</sub> captured at the central reference site reflect influences of large-scale factors (e.g., weather) that impact the entire sampling domain. The Lawrenceville site was selected as the reference because it represents typical urban background concentrations.

We obtained the road network shapefile from Pennsylvania spatial data access and traffic counts data from Pennsylvania Department of Transportation (PASDA, 2017b, 2017a). We classified roads with annual average daily traffic (AADT) greater than 5,000 vehicles per day as major roads. We also define vehicle density (Table B1) as the product of AADT and road length.

Locations of restaurants were retrieved from the Allegheny County Health Department. From an air pollution perspective, our primary interest was meat cooking (Schauer et al., 1999). We therefore excluded restaurants without an obvious cooking source, such as ice cream and coffee shops.

Industry emission and location information came from the National Emission Inventory (2011 NEI) (US EPA, 2015). Elevation data were from the USGS National Elevation Dataset (National Map, 2017). Land use type included utility/transport, industrial, commercial, residential, agricultural, or vacant/forest defined by Allegheny County GIS group (Tan et al., 2016).

Road and traffic intensity variables were extracted in circular buffers of 25, 50, 100, 300, 500, and 1,000 m around each sampling site. The traffic group had the smallest buffer size, starting from 25 m, as traffic related pollutants decrease to background concentrations within 100-500 meters of the road edge (Karner et al., 2010). The small buffer size indicated the importance of local source emissions. Buffers for land use zoning variables and the restaurant group began with 100 m. The industry group had buffers starting from 1,000 m due to transport of pollutants after emission from high stacks.

For LUR model building, we adopted the forward selection approach used by the European Study of Cohorts for Air Pollution Effects (ESCAPE) (Eeftens et al., 2012). Potential variables were assigned a prior regression coefficient sign as either positive or negative before they were examined in the variable selection process. The assignment was intuitive, e.g., traffic emissions tended to have a positive influence on pollutant concentrations. Thus, traffic related variables would have a positive regression coefficient ( $\beta_i$ ) in the model.

In the forward selection process, potential variables were added to the model

iteratively. First, univariate regression models were developed for the measured pollutants using individual predictor variables, and the predictor with the highest  $R^2$  was chosen for the initial model. The remaining variables were added to the model separately to create an intermediate model, and adjusted- $R^2$  was calculated. If the added predictor yielded an increase in adjusted- $R^2$  greater than 1% and had the same regression coefficient sign as prior assigned, the intermediate model was considered valid and was used as the base model for next round of iteration. This process continued until no more variables could meet the inclusion criteria.

We examined the final model for predictor significance, collinearity of predictors, and influential observations (potential outliers). Variables were removed if their p values based on an F-test were larger than 0.1 or had a variation inflation factor (VIF) larger than 3. We used Cook's D to signify influential observations. Observations with Cook's D bigger than 1 were further examined by developing and comparing models with this observation or without.

Other model diagnostics included leave-one-out cross validation (LOOCV)  $R^2$ , mean studentized prediction residual (MSPR), root mean square of studentized residuals (RMS) produced by LOOCV, and Moran's I (Mukerjee et al., 2009). LOOCV examined model goodness at predicting test datasets. Each observation was deleted iteratively, and new models were developed using the remaining data. The new model then predicted the deleted observation. Moran's I detected spatial autocorrelation of residuals. LUR, as a linear model, relied on the assumptions that observations were independent of each other. If spatial autocorrelation existed in the final model (p value less than 0.05), then this important assumption was violated.

### **3.3 Results and Discussion**

### 3.3.1 OC spatial variation

Figure 3.1 shows the spatial variation of OC and EC between our 36 sampling sites. Sampling locations are sorted by inverse distance from the city center, and sites near point sources are indicated with an asterisk. The detailed sampling site description is listed in Table B2 (Supporting Information).

Particle phase OC in Pittsburgh (Ma et al., 2016; Robinson et al., 2007; Subramanian et al., 2007) increases from  $0.8 \mu\text{g}/\text{m}^3$  at urban background and upwind locations to  $\sim 4 \mu\text{g}/\text{m}^3$  at downtown areas or industrial sites. While BQ-QBT OC (purple symbols in Fig. 3.1) generally increases from rural to urban locations, the increase is not monotonic. This suggests that the spatial variability of particulate OC is multi-factorial and does not simply increase along the rural-urban gradient. The highest OC concentrations are typically observed at sites near point sources (mean =  $2.25 \mu\text{g}/\text{m}^3$ ).

Overall, OC approximately triples from upwind to near source locations. This spatial variation has important human exposure implications, as OC occupies a large fraction of ambient fine PM. EC is primary in nature and its fraction spans from 4% to 18% of total carbon. The sites with the highest EC fraction are heavily influenced by local industrial emissions and heavy-duty vehicle emissions (Tan et al., 2014a).

### 3.3.2 Different fresh emission influence on OCX

Figure 3.2 shows that the abundance of OCX varies with total OC loading on the bare quartz filter for both ambient and tunnel samples. BQ OC2 has the largest range of concentrations from 0 to  $8 \mu\text{g}/\text{m}^3$ , OC3 from 0 to  $6 \mu\text{g}/\text{m}^3$ , and OC4 and PC tend to be  $< 2 \mu\text{g}/\text{m}^3$ . Data organize into rows because of the instrument resolution, especially for OC4 and PC. Ambient OC2 and OC3 seem to increase linearly with total OC loading on the bare quartz filter, and concentrations of both are clearly elevated in the tunnel. OC4 also increases

with BQ OC, but seems to level off in the ambient samples for BQ OC >  $\sim 5\text{-}10\ \mu\text{g}/\text{m}^3$  (though OC4 is slightly elevated in the tunnel relative to ambient). PC levels off in the ambient filters for BQ OC >  $\sim 5\ \mu\text{g}/\text{m}^3$  and is not elevated in the tunnel relative to ambient. The comparison of ambient to tunnel data in Figure 3.2 suggests that as volatility decreases from OC2 to PC, the contribution of fresh emissions also decreases. The ambient and tunnel data also suggest that primary traffic emissions are not a major source of PC in Pittsburgh.

Figure 3.2 also compares the ambient and tunnel measurements to published OCX/OC ratios derived from source emissions tests for gasoline and diesel vehicles, cooking, and biomass burning (May et al., 2013a, 2013b). For OC2, all four source relationships describe the ambient and tunnel observations with high  $R^2$  values. This suggests that OC2 is a marker for fresh emissions, but is not source specific.

The observed ambient and tunnel OC3/OC ratios are most similar to diesel emissions ( $R^2 = 0.7$ ), and gasoline emissions seem to set the lower bound of our Pittsburgh dataset. Cooking and biomass burning do not describe the observed OC3 in Pittsburgh. This suggests that OC3 may be a good traffic emissions marker, but it is unclear from our data whether we can definitively claim OC3 as a marker for gasoline or diesel emissions.

Observed ambient and tunnel OC4 are poorly correlated with source test data. Gasoline source tests are omitted from the OC4 panel of Figure 3.2 because of poor correlation between OC4 and BQ OC. OC4/OC ratios for diesel vehicles, cooking, and biomass burning do not explain the observed ambient and tunnel ratios. The poor correlation between ambient OC4 and the source tests, along with the minimal enhancement of OC4 in the traffic tunnel, suggests that OC4 may be an indicator for secondary organic aerosol.

PC also appears to be a marker of secondary organic aerosol for the Pittsburgh data. This is consistent with charring of ambient secondary OC contributing to measured PC. For

example, Yu et al. (2002) found that water-soluble OC accounted for a large fraction of charring on quartz filter samples, while hexane extractable organic compounds (non-polar) produced little charring. The ratio of PC to total water-soluble OC was found to increase with water-soluble OC loading; in other words, the abundance of PC was largely dependent on the fraction of water-soluble secondary OC.

PC was not observed in the gasoline (Chow et al., 2001; Schauer et al., 2003) and diesel (May et al., 2013a, 2013b) exhaust filters shown in Figure 3.2, but it is observed in wood smoke source tests (Chow et al., 2001; Li et al., 2016; Schauer et al., 2003). Figure 3.2 shows that the PC/OC trend from biomass burning source tests does not describe the Pittsburgh ambient samples, and this is likely because biomass burning is not a major source to PM<sub>2.5</sub> in Pittsburgh (Robinson et al., 2006).

Figure 3.2 suggests that OC2 and OC3 collected on bare quartz filters are good fresh emission markers, with OC3 specifically being a marker for traffic emissions. OC4 and PC are the secondary OC in the study region. The different OCX therefore suggest different source implications.

### **3.3.3 From Pittsburgh to nationwide**

To broaden study impacts and test our hypothesis of OCX's different fresh emission dependence nationwide instead of just inside the study region, we extracted OC reports from all CSN sites in 2013 and present their relationship with our Pittsburgh dataset and source test results in Figure 3.3.

In Pittsburgh, more volatile OCX (OC2 and OC3) are associated with fresh emissions, and this trend also applies to CSN sites nationwide. The national OC2 data in Figure 3.3 are linearly correlated with total BQ OC, and the trend can reasonably be described with OC2/OC ratios from source tests. The Pittsburgh OC2/OC relationship falls along the upper bound



of the nationwide OC2 data. The national OC2 data have the strongest agreement with diesel emissions ( $R^2 = 0.94$ ), whereas for our Pittsburgh dataset the strongest agreement was with gasoline vehicles. Nonetheless, the national CSN data suggest that BQ OC2 is an indicator of fresh emissions, though the specific identify of those emissions may be uncertain.

For OC3, the Pittsburgh and national CSN data largely overlap. Again, agreement is strongest with diesel emissions. Cooking and biomass burning do not describe the observed OC3/OC trend in either the Pittsburgh or national CSN data. Thus, the national CSN data seem to agree with our observation from Figure 3.2 that OC3 is a marker for traffic emissions.

For OC4 the national data show a similar trend (increasing at low BQ OC and plateauing at  $\sim 1 \mu\text{g}/\text{m}^3$  OC4 for BQ OC  $> \sim 5\text{-}10 \mu\text{g}/\text{m}^3$ ) as the Pittsburgh data and poor agreement with source test results. This again suggests that OC4 may be a marker for secondary organic aerosol (Leskinen et al., 2007a, 2007b).

PC is very different between Pittsburgh and the national CSN data. Whereas PC peaks at  $\sim 2 \mu\text{g}/\text{m}^3$  in Pittsburgh, some CSN sites have much higher PC. Biomass burning is a major source for global/nationwide ambient  $\text{PM}_{2.5}$  and PC, while it is only a minor source in Pittsburgh (Robinson et al., 2006; Zhou et al., 2017; Zhou et al., 2017). The implication is that PC in Pittsburgh seems to be an indicator of secondary PM, but in the nationwide dataset there seem to be some sites where PC is dominated by biomass burning. These sites fall near the biomass line in Figure 3.3. Thus, when we consider the Pittsburgh and CSN data together, PC is of mixed origin and interpreting it as either secondary OA or a biomass marker may depend on where the samples are collected, and knowing about sources near sampling sites.

### 3.3.4 LUR models

By using mobile sampling measurements (Figure 3.1) as input, we developed LUR

models for bare quartz OC, OC2, OC3, OC4, PC, and particle phase OC (BQ-QBT). The detailed model results are in Table 3.1. Variable names are briefly described in Table 3.1 and are described in detail in Table B1 of the Supporting Information.

Our LUR models achieve good performance as their  $R^2$  values are all above 0.5 (de Hoogh et al., 2013; Zhang et al., 2015). Comparison of filter measurements averaged for each of the 36 sampling sites and LUR model predictions at the sampling locations are shown in Figure B2 in the Supplemental Materials. Measurements and predictions generally agree with each other and are distributed along the 1:1 line.

The number of independent predictors in LUR models ranges from 3 to 6. The central reference site group appears in every model except BQ OC3, and this indicates strong association between OC and regional background  $PM_{2.5}$  concentrations, as expected (Robinson et al., 2007). Predictors from the elevation and restaurant groups are not selected in the final models.

As described above, we evaluated LUR models using LOOCV, MSPR, RMS, and Moran's I. A  $R^2$  shrinkage of 0.15 or smaller between original and LOOCV model indicates stable models (de Hoogh et al., 2013; Zhang et al., 2015). All derived models are stable. The ideal value for MSPR is 0, and that for RMS is 1. All models have MSPR value close to 0, and the largest RMS is 1.07. We do not detect any spatial autocorrelation of residuals based on Moran's I.

We used the LUR models to generate pollutant prediction maps in ArcGIS-10.3 (ESRI, Redlands, CA) with the spatial analyst tool. Since the LUR model purely extrapolates outside the sampling domain (de Hoogh et al., 2013; Zhang et al., 2015), we limit the minimum BQ-QBT OC predictions to be  $1.5 \mu\text{g}/\text{m}^3$  based on an upwind EPA monitor. Artifact-corrected OC concentrations are elevated in urban and near-source regions in Figure

3.4(a). Primary emission sources are more concentrated in the city center than in rural regions, and these emissions drive local enhancements in particulate (BQ-QBT) OC.

We retrieved annual average particulate (BQ-QBT) OC concentrations from two regulatory PM speciation monitors (Lawrenceville and Liberty; Figure B1). These two sites served as an independent test set to evaluate our LUR model, as they were not included as inputs during the LUR model building process. Our BQ-QBT OC prediction at Lawrenceville is  $0.04 \mu\text{g}/\text{m}^3$  larger than the monitor measurement. The LUR model under predicted OC by  $0.5 \mu\text{g}/\text{m}^3$  at Liberty. This under prediction at Liberty was also observed in our LUR  $\text{PM}_{2.5}$  model in the same study domain, suggesting LUR did not fully capture the spatial gradient induced by a nearby major industrial point source (Li et al., 2016).

BQ OCX exhibit different spatial patterns, as shown in Figure 3.4. Both OC2 and OC3 show significant spatial variability associated with fresh emissions, primarily from traffic sources. The road network is evident as yellow areas in both Figure 3.4(b) and (c). OC2 and OC3 are both elevated in the Pittsburgh central business district (center of the map), where there is both high traffic and large emissions from other human activities (e.g., restaurants). On the other hand, the concentration surface of OC4 is less variable in Figure 3.4(d).

The plot under each LUR map is projected from the black line in the map and quantifies the spatial variation of BQ-QBT OC, BQ OC2, OC3 or OC4 across a transect. BQ-QBT OC increases from  $1.5 \mu\text{g}/\text{m}^3$  to  $2\text{-}3 \mu\text{g}/\text{m}^3$  on a neighborhood scale (1-2 km) but otherwise is mostly at the background – most of the observed OC is regional. BQ OC2 and OC3 are enhanced near busy roads and the city center. People living near roadways have higher OC2 and OC3 exposures, though since OC2 and OC3 partition between phases, not all of this enhanced OC2 or OC3 shows up in the BQ-QBT map. OC4 is nearly invariant, and

oscillates around  $0.5 \mu\text{g}/\text{m}^3$  across the whole model prediction domain.

### 3.3.5 Exposure implications

Figure 3.2 shows that BQ OC2 has the largest concentration range of the OC fractions in Pittsburgh ( $0\text{--}6 \mu\text{g}/\text{m}^3$ ), whereas OC4 and PC are less variable ( $0\text{--}2 \mu\text{g}/\text{m}^3$ ). OC2 and OC3 seem to be associated with fresh emissions. Higher total BQ and BQ-QBT OC is therefore driven by changes in OC2 and OC3, whereas OC4 and PC are more spatially homogeneous. This means that variations in exposure to total OC mass are convolved with variations in the OC composition.

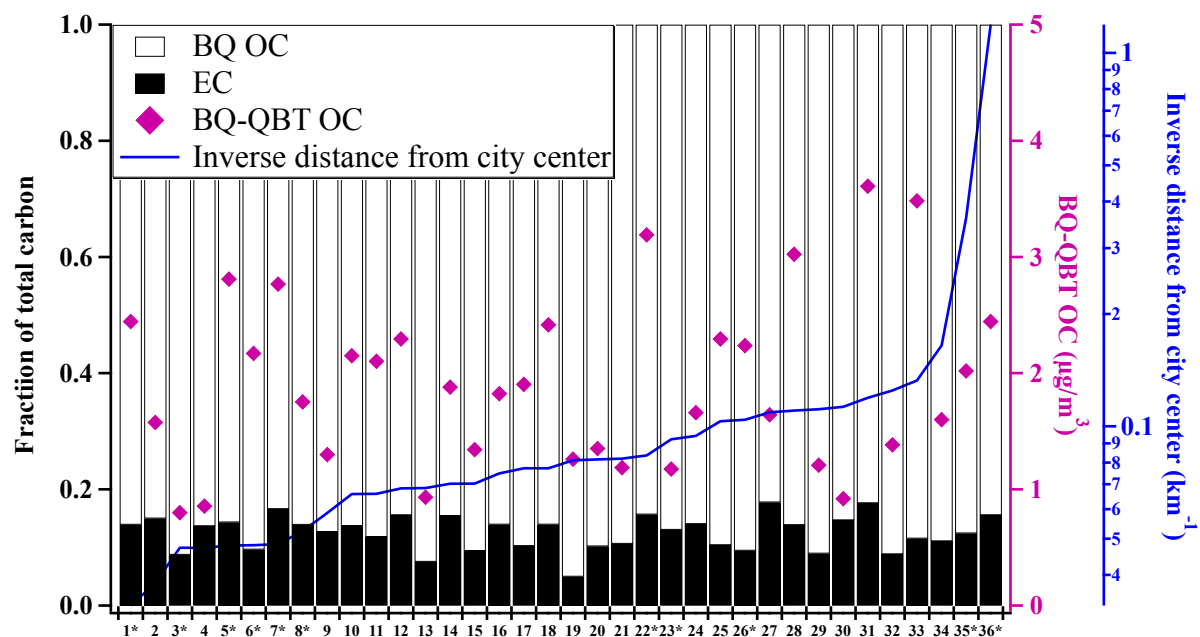
Figure 3.5 quantifies this variation and provides an estimate of human exposure considering the spatial heterogeneity of OCX and varying population density. We first obtained population of every census block in Allegheny county ( $\sim 30,000$  blocks) (US Census Bureau, 2010), and assumed the population was evenly distributed in the block. We then assigned a block-averaged LUR OCX prediction to each block.

The black line in Figure 3.5 shows the cumulative distribution function (CDF) of population-weighted BQ OC exposure. The stacked colors show the distribution of BQ OC among OC2, OC3, OC4, and PC. As noted above, we do not directly consider OC1, which exists almost exclusively as vapor in the atmosphere and is not captured efficiently by quartz filters. Therefore, the sum of OC2, OC3, OC4, and PC (right y-axis) is less than the total BQ OC (x-axis).

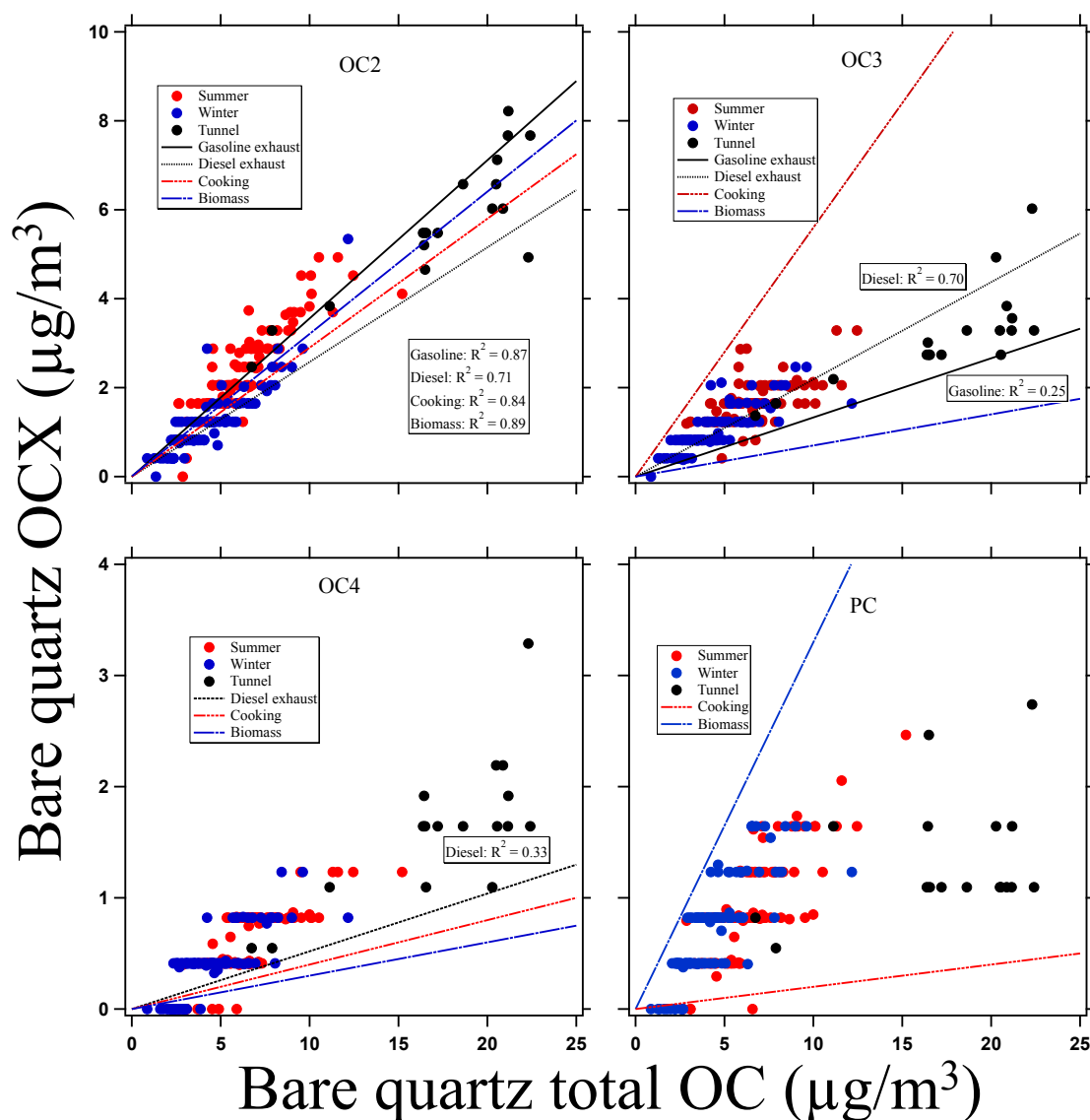
Figure 3.5 shows that variations in BQ OC exposure are driven almost entirely by the more volatile OC2 and OC3 fractions. For example, population-weighted BQ OC increases from  $\sim 3.5$  to  $\sim 4.5 \mu\text{g}/\text{m}^3$  from the 10<sup>th</sup> to the 90<sup>th</sup> percentile. OC4 and PC account for approximately 10% of the difference, increasing from  $\sim 1$  to  $\sim 1.1 \mu\text{g}/\text{m}^3$ , whereas OC2 and OC3 account for the remaining 90% of the additional exposure.

In urban areas with high population and source density, people are exposed to higher BQ OC because of the local enhancement of OC2 and OC3 from primary sources such as traffic and restaurants. On the other hand, the relatively constant concentrations of OC4 and PC indicate that people living in both urban and rural areas are inhaling similar concentrations of secondary OC4 and PC. When people move from rural areas to downtown, exposures to primary OC2 and OC3 increase while those of secondary OC4 and PC are roughly the same. Since OC4 and PC are effectively always in the particle phase (Ma et al., 2016), everyone is breathing the spatially homogeneous portion of the (mostly secondary) OC.

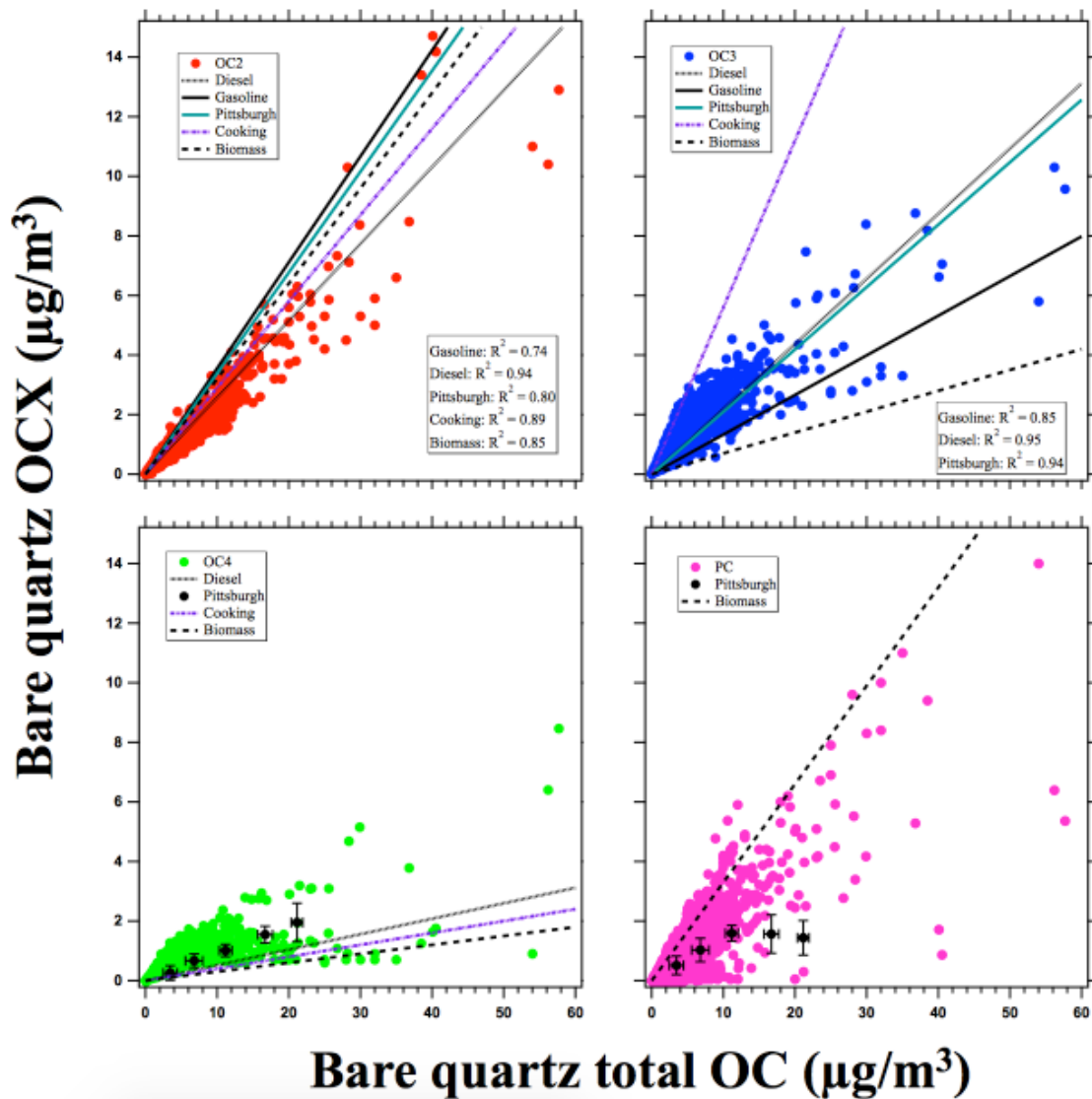
Figure 3.5 shows exposures for BQ OC, and bare quartz filters capture both particulate and vapor-phase organics. We would expect the exposure contrast for particulate (BQ-QBT) OC to be somewhat muted, because OC2 and OC3 are semi-volatile, and exist partially as vapors in the atmosphere. Nonetheless, a similar exposure pattern should hold for particulate OC: spatially homogeneous background aerosol (secondary OC4 and PC) punctuated by zones of higher exposure due to local emissions sources.



**Fig. 3.1.** Comparison of quartz filter measurements of average OC and EC at each sampling site. Solid and blank bars indicate relative fractions of total carbon on the left y axis. We order the thirty-six sites based on inverse distance from city center, namely from rural to downtown (blue line, log scale on right y axis). Sites near point sources are indicated with an asterisk.



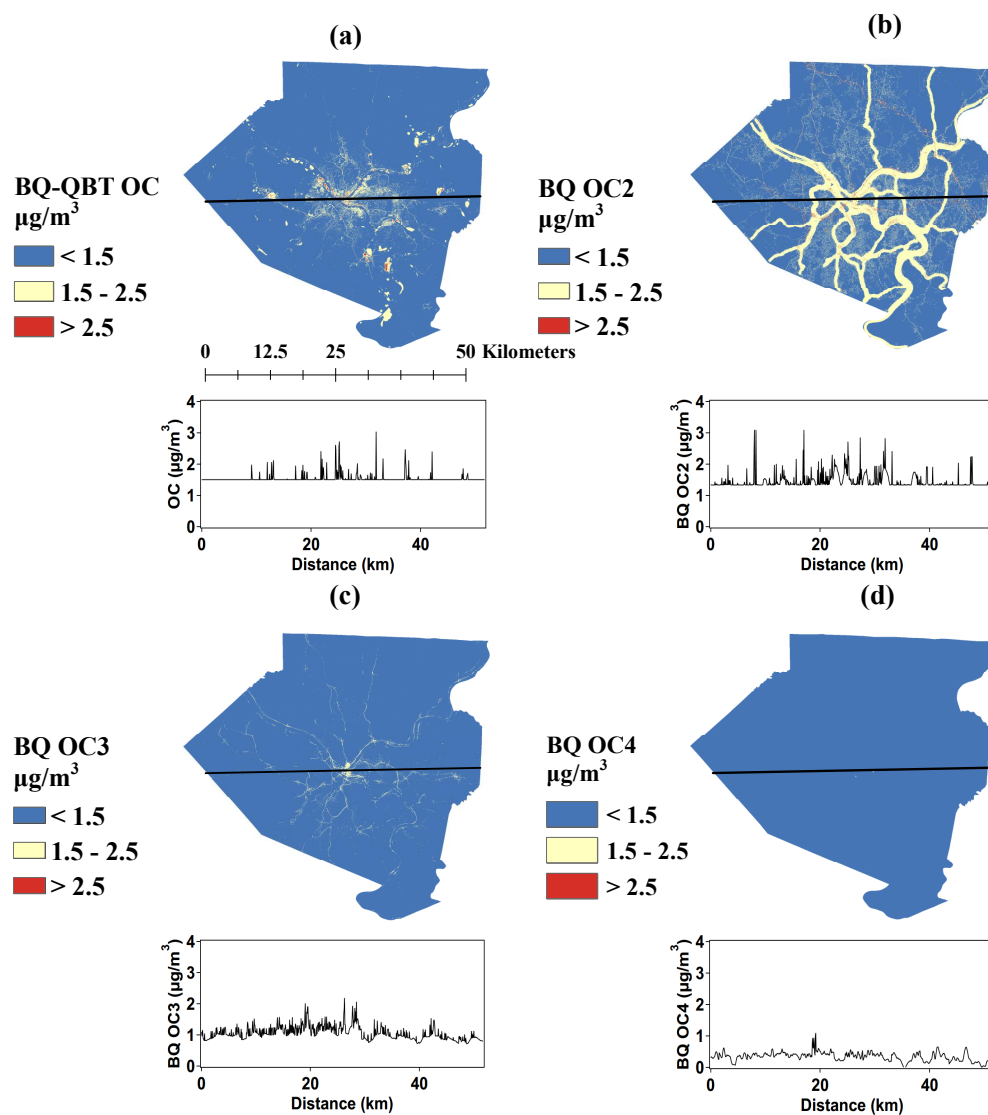
**Fig. 3.2.** Seasonal comparison of bare quartz OCX on ambient and tunnel filters. Regression lines are from dynamometer studies, cooking emission, and biomass burning filters. Diesel, gasoline, cooking and biomass burning lines are all present in OC2 and OC3. OC4 has the diesel, cooking, and biomass lines, and PC has cooking and biomass burning ones. The missing gasoline line in OC4 (and both vehicle lines in PC) results from poor correlation of OC4 (PC) with total OC on source filters. The top and bottom panels have different y scales.



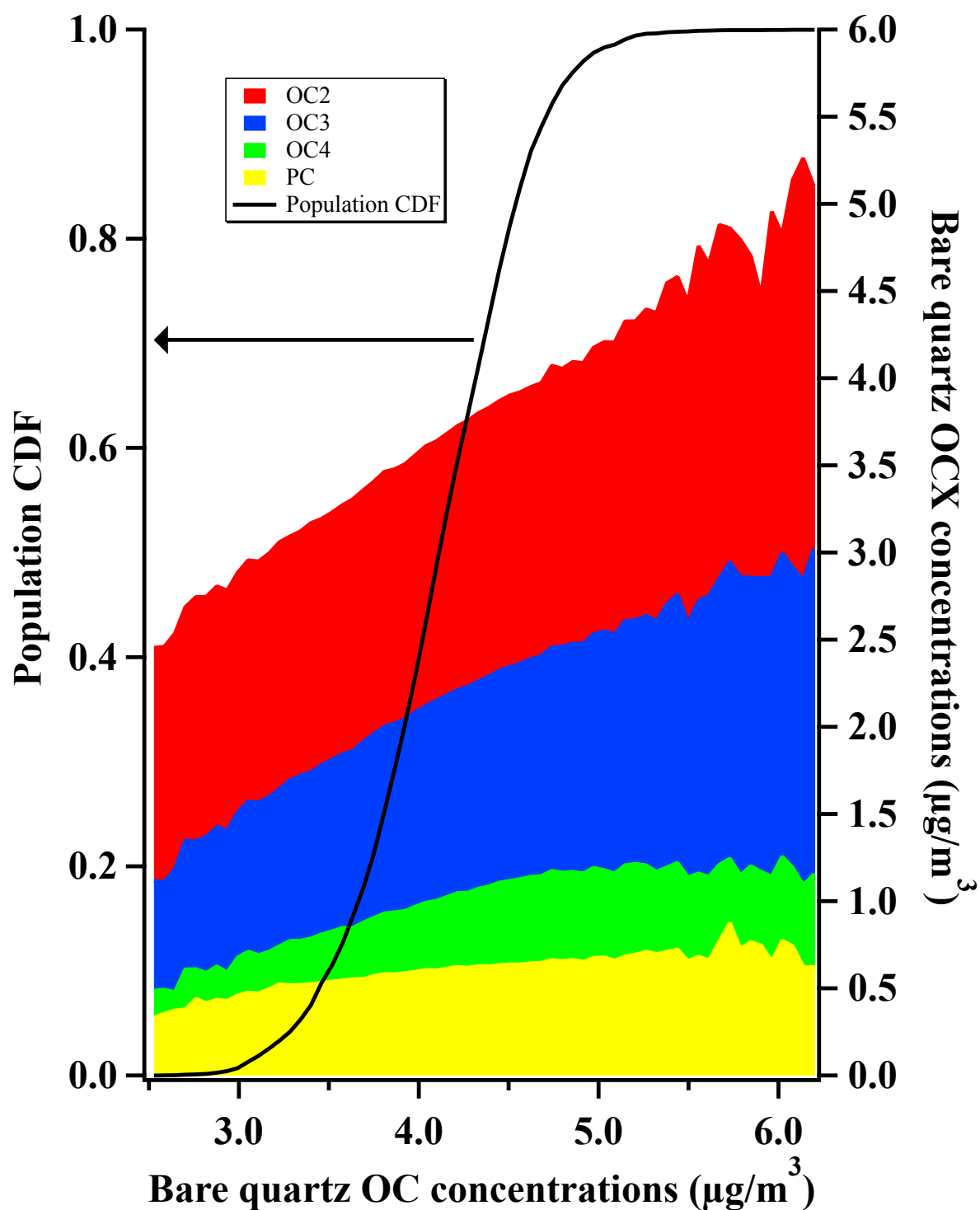
**Fig. 3.3.** OCX on bare quartz filters from all CSN sites in 2013. Regression lines for Pittsburgh OC2 and OC3 are based on the combined dataset of summer, winter, and tunnel.

For OC4 and PC, Pittsburgh data are grouped into bins of  $2 \mu\text{g}/\text{m}^3$ , and the error bars show one standard deviation.





**Fig. 3.4.** LUR predicted BQ-QBT OC concentration (a), BQ OC2 (b), OC3 (c), and OC4 (d) for the whole study region of Allegheny County, PA. Plots under the maps are the projected pollutant predictions along the black transect lines. The shapefile is downloaded from Pennsylvania Spatial Data Access.



**Fig. 3.5.** CDF of Population exposure to bare quartz OC (left y axis) in Allegheny County, PA. Stacked lines of different color show LUR predicted BQ OCX concentrations (right y axis) as a function of total BQ OC concentration.

**Table 3.1.** LUR model parameters and diagnostic test results. Values in parenthesis are coefficients ( $\beta$ ) for final predictors except for those in Moran's I column, which are p values. Elevation and restaurant groups are not included in the table as no final models include these variables. Detailed variable meanings are in Table S1 of the Supporting Information. Selected variables (\*) are explained in the footnote.

Model Performance				Model Variables						Model Evaluation		
Species	Quartz source	R <sup>2</sup>	Adjusted R <sup>2</sup>	LOOCV R <sup>2</sup>	Traffic	Industry	Central reference site	Others	Intercept	Moran's I	MSPR	RMS
OC_BQ	BQ	0.76	0.72	0.68	*RDALL300(1.69×10 <sup>-4</sup> ) *DISTINVALL(3.05) *MAJAAADT_DIS(2.56×10 <sup>-4</sup> )		*CSMPM(2.37×10 <sup>-1</sup> )	*LUVaFo500(-1.66×10 <sup>-6</sup> )	1.39	0.06(0.36)	-5.72×10 <sup>-4</sup>	1.03
OC2	BQ	0.50	0.45	0.38	RAIL500(1.94×10 <sup>-4</sup> ) ALLDIESAADT_DIS2(1.58×10 <sup>-2</sup> )		CSMPM(8.97×10 <sup>-2</sup> )		5.22×10 <sup>-1</sup>	-0.19(0.11)	-5.31×10 <sup>-4</sup>	1.03
OC3	BQ	0.61	0.56	0.40	RDALL25(4.79×10 <sup>-3</sup> ) RAIL50(3.21×10 <sup>-3</sup> ) VEHDENSALL500(1.56×10 <sup>-3</sup> )			LUVaFo1000(-2.01×10 <sup>-7</sup> )	1.11	0.04(0.48)	9.29×10 <sup>-3</sup>	1.07
OC4	BQ	0.70	0.64	0.57		*PointDe_NEI_PM_Popu_5000 (1.12×10 <sup>-5</sup> )	CSMPM(2.16×10 <sup>-2</sup> )	POP100(1.22×10 <sup>-3</sup> ) LUCOMM300(1.84×10 <sup>-6</sup> ) LUAGRI1000(-6.29×10 <sup>-7</sup> ) LURES500(3.25×10 <sup>-7</sup> )	-3.77×10 <sup>-2</sup>	0.02(0.62)	1.47×10 <sup>-3</sup>	1.01
PC	BQ	0.60	0.55	0.49	RAIL500(1.65×10 <sup>-4</sup> )	PointDe_NEI_PM_1500 (1.61×10 <sup>-1</sup> )	CSMPM(5.31×10 <sup>-2</sup> )	LURES500(5.97×10 <sup>-7</sup> )	-1.63×10 <sup>-1</sup>	-0.01(0.82)	4.80×10 <sup>-3</sup>	1.02
OC	BQ-QBT	0.71	0.66	0.59	RDALL25(1.06×10 <sup>-2</sup> )	PointDe_NEI_All_10000 (4.32)	CSMPM(2.06×10 <sup>-1</sup> )	LUINDUS100(5.02×10 <sup>-5</sup> ) LUCOMM1000(7.07×10 <sup>-7</sup> )	-2.00	0.02(0.61)	7.32×10 <sup>-3</sup>	1.05

\*RDALL300: total road length within surrounding 300 m circular buffers. DISTINVALL: inverse distance to nearest roads. MAJAADT\_DIS: nearest major road AADT (annual average daily traffic) times the inverse distance to it. CSMPM: PM measurement at reference site during mobile sampling periods. LUVaFo500: Vacant/Forest land-use areas within surrounding 500 m circular buffers. PointDe\_NEI\_PM\_Popu\_5000: PM point sources density in NEI.

### 3.4 Limitations

The major limitation of this paper is the short-duration quartz filter sampling. Tan et al. (2014b) showed that short term mobile sampling was less adept at estimating long term average pollutant concentrations compared with stationary monitors. Nonetheless, our recent study demonstrated that mobile measurements could capture spatial variation of PM<sub>2.5</sub> and its constituents and were adequate for developing LUR models (Li et al., 2016).

The BQ-QBT method can overcorrect OC artifacts and lead to smaller concentrations (Subramanian et al., 2004). We assume that BQ-QBT works well for artifact correction, and that filters do a good job of capturing particle plus vapors for OC2, OC3, OC4, and PC. This may not be a perfect assumption. However, it is worth noting that our sampling setup is a common one, which makes it easy to compare to data from nationwide networks such as CSN and IMPROVE.

To avoid unrealistic LUR prediction due to model extrapolation, we set a reference standard for our lowest BQ-QBT OC concentration using the upwind EPA Florence site. The Florence site reports bare quartz carbon speciation. We used the volatility basis set framework (Donahue et al., 2006) to calculate particle OC and treated it as the lowest particle OC prediction in our LUR maps.

### 3.5 References

- Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A.S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., Hoek, G., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet* 383, 785–795. [https://doi.org/10.1016/S0140-6736\(13\)62158-3](https://doi.org/10.1016/S0140-6736(13)62158-3)
- Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L.T., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* 50, 79–88. <https://doi.org/10.1021/acs.est.5b03709>

- Bravo, M.A., Ebisu, K., Dominici, F., Wang, Y., Peng, R.D., Bell, M.L., 2017. Airborne Fine Particles and Risk of Hospital Admissions for Understudied Populations: Effects by Urbanicity and Short-Term Cumulative Exposures in 708 U.S. Counties. *Environ. Health Perspect.* 125, 594–601. <https://doi.org/10.1289/EHP257>
- Brokamp, C., Jandarov, R., Rao, M.B., LeMasters, G., Ryan, P., 2017. Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches. *Atmos. Environ.* 151, 1–11.
- Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. *Circulation* 121, 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbecel>
- Chen, L., Shi, M., Li, S., Bai, Z., Wang, Z., 2017. Combined use of land use regression and BenMAP for estimating public health benefits of reducing PM<sub>2.5</sub> in Tianjin, China. *Atmos. Environ.* 152, 16–23.
- Chow, J.C., Watson, J.G., Chen, L.-W.A., Chang, M.C.O., Robinson, N.F., Trimble, D., Kohl, S., 2007. The IMPROVE\_A Temperature Protocol for Thermal/Optical Carbon Analysis: Maintaining Consistency with a Long-Term Database. *J. Air Waste Manag. Assoc.* 57, 1014–1023. <https://doi.org/10.3155/1047-3289.57.9.1014>
- Chow, J.C., Watson, J.G., Chen, L.-W.A., Rice, J., Frank, N.H., 2010. Quantification of PM<sub>2.5</sub> organic carbon sampling artifacts in US networks. *Atmos Chem Phys* 10, 5223–5239. <https://doi.org/10.5194/acp-10-5223-2010>
- Chow, J.C., Watson, J.G., Crow, D., Lowenthal, D.H., Merrifield, T., 2001. Comparison of

IMPROVE and NIOSH Carbon Measurements. *Aerosol Sci. Technol.* 34, 23–34.

<https://doi.org/10.1080/02786820119073>

- de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., D  del  , A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gra  ulevi  ien  , R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsouyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., M  lter, A., Mosler, G., N  dor, G., Nieuwenhuijsen, M., Pershagen, G., Peters, A., Phuleria, H., Probst-Hensch, N., Raaschou-Nielsen, O., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varr  , M.J., Vienneau, D., Weinmayr, G., Brunekreef, B., Hoek, G., 2013. Development of Land Use Regression Models for Particle Composition in Twenty Study Areas in Europe. *Environ. Sci. Technol.* 47, 5778–5786. <https://doi.org/10.1021/es400156t>
- Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.D., 2017. Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* 376, 2513–2522. <https://doi.org/10.1056/NEJMoal702747>
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G.J., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities. *N. Engl. J. Med.* 329, 1753–1759. <https://doi.org/10.1056/NEJM199312093292401>
- Donahue, N.M., Posner, L.N., Westervelt, D.M., Li, Z., Shrivastava, M., Presto, A.A., Sullivan, R.C., Adams, P.J., Pandis, S.N., Robinson, A.L., 2016. Where Did This Particle Come From? Sources of Particle Number and Mass for Human Exposure Estimates, in: *Airborne Particulate Matter*. Royal Society of Chemistry, Cambridge, pp. 35–71.

- Donahue, N.M., Robinson, A.L., Stanier, C.O., Pandis, S.N., 2006. Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics. *Environ. Sci. Technol.* 40, 2635–2643. <https://doi.org/10.1021/es052297c>
- Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dèdelè, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S. von, Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of Land Use Regression Models for PM<sub>2.5</sub>, PM<sub>2.5</sub> Absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European Study Areas; Results of the ESCAPE Project. *Environ. Sci. Technol.* 46, 11195–11205. <https://doi.org/10.1021/es301948k>
- Geography, U.C.B., 2010. 2010 Census - Census Tract Reference Maps [WWW Document]. URL <https://www.census.gov/geo/maps-data/maps/2010tract.html> (accessed 10.28.17).
- Hankey, S., Marshall, J.D., 2015. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM<sub>2.5</sub>, Particle Size) Using Mobile Monitoring. *Environ. Sci. Technol.* 49, 9194–9202.
- Jedynska, A., Hoek, G., Wang, M., Eeftens, M., Cyrys, J., Keuken, M., Ampe, C., Beelen, R., Cesaroni, G., Forastiere, F., Cirach, M., de Hoogh, K., De Nazelle, A., Nystad, W., Declercq, C., Eriksen, K.T., Dimakopoulou, K., Lanki, T., Meliefste, K., Nieuwenhuijsen, M.J., Yli-Tuomi, T., Raaschou-Nielsen, O., Brunekreef, B., Kooter,



- I.M., 2014. Development of Land Use Regression Models for Elemental, Organic Carbon, PAH, and Hopanes/Steranes in 10 ESCAPE/TRANSPHORM European Study Areas. *Environ. Sci. Technol.* 48, 14435–14444. <https://doi.org/10.1021/es502568z>
- Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prevot, A.S.H., Zhang, Q., Kroll, J.H., DeCarlo, P.F., Allan, J.D., Coe, H., Ng, N.L., Aiken, A.C., Docherty, K.S., Ulbrich, I.M., Grieshop, A.P., Robinson, A.L., Duplissy, J., Smith, J.D., Wilson, K.R., Lanz, V.A., Hueglin, C., Sun, Y.L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J.M., Collins, D.R., Cubison, M.J., E, Dunlea, J., Huffman, J.A., Onasch, T.B., Alfarra, M.R., Williams, P.I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J.Y., Zhang, Y.M., Dzepina, K., Kimmel, J.R., Sueper, D., Jayne, J.T., Herndon, S.C., Trimborn, A.M., Williams, L.R., Wood, E.C., Middlebrook, A.M., Kolb, C.E., Baltensperger, U., Worsnop, D.R., 2009. Evolution of Organic Aerosols in the Atmosphere. *Science* 326, 1525–1529. <https://doi.org/10.1126/science.1180353>
- Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. *Environ. Sci. Technol.* 44, 5334–5344. <https://doi.org/10.1021/es100008x>
- Kim, E., Hopke, P.K., 2004a. Improving source identification of fine particles in a rural northeastern U.S. area utilizing temperature-resolved carbon fractions. *J. Geophys. Res. Atmospheres* 109, D09204. <https://doi.org/10.1029/2003JD004199>
- Kim, E., Hopke, P.K., 2004b. Source Apportionment of Fine Particles in Washington, DC, Utilizing Temperature-Resolved Carbon Fractions. *J. Air Waste Manag. Assoc.* 54,

- 773–785. <https://doi.org/10.1080/10473289.2004.10470948>
- Krall, J.R., Anderson, G.B., Dominici, F., Bell, M.L., Peng, R.D., 2013. Short-term Exposure to Particulate Matter Constituents and Mortality in a National Study of U.S. Urban Communities. *Environ. Health Perspect.* <https://doi.org/10.1289/ehp.1206185>
- Krall, J.R., Mulholland, J.A., Russell, A.G., Balachandran, S., Winkquist, A., Tolbert, P.E., Waller, L.A., Sarnat, S.E., 2017. Associations between Source-Specific Fine Particulate Matter and Emergency Department Visits for Respiratory Disease in Four U.S. Cities. *Environ. Health Perspect.* 125, 97–103. <https://doi.org/10.1289/EHP271>
- Lee, P.K.H., Brook, J.R., Dabek-Zlotorzynska, E., Mabury, S.A., 2003. Identification of the Major Sources Contributing to PM<sub>2.5</sub> Observed in Toronto. *Environ. Sci. Technol.* 37, 4831–4840.
- Leskinen, A.P., Jokiniemi, J.K., Lehtinen, K.E.J., 2007a. Characterization of aging wood chip combustion aerosol in an environmental chamber. *Atmos. Environ.* 41, 3713–3721. <https://doi.org/10.1016/j.atmosenv.2006.12.016>
- Leskinen, A.P., Jokiniemi, J.K., Lehtinen, K.E.J., 2007b. Transformation of diesel engine exhaust in an environmental chamber. *Atmos. Environ.* 41, 8865–8873. <https://doi.org/10.1016/j.atmosenv.2007.08.021>
- Li, H.Z., Dallmann, T.R., Gu, P., Presto, A.A., 2016. Application of mobile sampling to investigate spatial variation in fine particle composition. *Atmos. Environ.* 142, 71–82. <https://doi.org/10.1016/j.atmosenv.2016.07.042>
- Li, X., Chen, Y., Bond, T.C., 2016. Light absorption of organic aerosol from pyrolysis of corn stalk. *Atmos. Environ.* 144, 249–256. <https://doi.org/10.1016/j.atmosenv.2016.09.006>
- Li, X., Dallmann, T.R., May, A.A., Tkacik, D.S., Lambe, A.T., Jayne, J.T., Croteau, P.L., Presto, A.A., 2016. Gas-Particle Partitioning of Vehicle Emitted Primary Organic

- Aerosol Measured in a Traffic Tunnel. *Environ. Sci. Technol.* 50, 12146–12155.  
<https://doi.org/10.1021/acs.est.6b01666>
- Lubin, J.H., Colt, J.S., Camann, D., Davis, S., Cerhan, J.R., Severson, R.K., Bernstein, L., Hartge, P., 2004. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ. Health Perspect.* 112, 1691–1696.
- Ma, J., Li, X., Gu, P., Dallmann, T.R., Presto, A.A., Donahue, N.M., 2016. Estimating ambient particulate organic carbon concentrations and partitioning using thermal optical measurements and the volatility basis set. *Aerosol Sci. Technol.* 50, 638–651.  
<https://doi.org/10.1080/02786826.2016.1158778>
- May, A.A., Presto, A.A., Hennigan, C.J., Nguyen, N.T., Gordon, T.D., Robinson, A.L., 2013a. Gas-particle partitioning of primary organic aerosol emissions: (1) Gasoline vehicle exhaust. *Atmos. Environ.* 77, 128–139. <https://doi.org/10.1016/j.atmosenv.2013.04.060>
- May, A.A., Presto, A.A., Hennigan, C.J., Nguyen, N.T., Gordon, T.D., Robinson, A.L., 2013b. Gas-Particle Partitioning of Primary Organic Aerosol Emissions: (2) Diesel Vehicles. *Environ. Sci. Technol.* 47, 8288–8296. <https://doi.org/10.1021/es400782j>
- Mukerjee, S., Smith, L.A., Johnson, M.M., Neas, L.M., Stallings, C.A., 2009. Spatial analysis and land use regression of VOCs and NO<sub>2</sub> from school-based urban air monitoring in Detroit/Dearborn, USA. *Sci. Total Environ.* 407, 4642–4651. <https://doi.org/10.1016/j.scitotenv.2009.04.030>
- Pennsylvania Spatial Data Access | Data Summary [WWW Document], 2017a. URL <http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=56> (accessed 10.28.17).
- Pennsylvania Spatial Data Access | Data Summary [WWW Document], 2017b. URL <http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=1200> (accessed 10.28.17).

- Pope, C. A.; Dockery, D. W. Health Effects of Fine Particulate Air Pollution: Lines that Connect. *J. Air Waste Manag. Assoc.* **2006**, *56* (6), 709–742.
- Pope, C. A.; Ezzati, M.; Dockery, D. W. Fine-Particulate Air Pollution and Life Expectancy in the United States. *N. Engl. J. Med.* **2009**, *360* (4), 376–386.
- Presto, A.A., Nguyen, N.T., Ranjan, M., Reeder, A.J., Lipsky, E.M., Hennigan, C.J., Miracolo, M.A., Riemer, D.D., Robinson, A.L., 2011. Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine. *Atmos. Environ.* *45*, 3603–3612. <https://doi.org/10.1016/j.atmosenv.2011.03.061>
- Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging. *Science* *315*, 1259–1262. <https://doi.org/10.1126/science.1133061>
- Robinson, A.L., Subramanian, R., Donahue, N.M., Bernardo-Bricker, A., Rogge, W.F., 2006. Source Apportionment of Molecular Markers and Organic Aerosol. 2. Biomass Smoke. *Environ. Sci. Technol.* *40*, 7811–7819. <https://doi.org/10.1021/es060782h>
- Robinson, A.L., Subramanian, R., Donahue, N.M., Rogge, W.F., 2006. Source Apportionment of Molecular Markers and Organic Aerosol1. Polycyclic Aromatic Hydrocarbons and Methodology for Data Visualization. *Environ. Sci. Technol.* *40*, 7803–7810. <https://doi.org/10.1021/es0510414>
- Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon Particulate Matter in New Delhi, India. *Environ. Sci. Technol.* *47*, 12903–12911. <https://doi.org/10.1021/es401489h>
- Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 1999. Measurement of Emissions

- from Air Pollution Sources. 1. C1 through C29 Organic Compounds from Meat Charbroiling. *Environ. Sci. Technol.* 33, 1566–1577. <https://doi.org/10.1021/es980076j>
- Schauer, J.J., Mader, B.T., DeMinter, J.T., Heidemann, G., Bae, M.S., Seinfeld, J.H., Flagan, R.C., Cary, R.A., Smith, D., Huebert, B.J., Bertram, T., Howell, S., Kline, J.T., Quinn, P., Bates, T., Turpin, B., Lim, H.J., Yu, J.Z., Yang, H., Keywood, M.D., 2003. ACE-Asia Intercomparison of a Thermal-Optical Method for the Determination of Particle-Phase Organic and Elemental Carbon. *Environ. Sci. Technol.* 37, 993–1001. <https://doi.org/10.1021/es020622f>
- Schwartz, J., Dockery, D.W., Neas, L.M., 1996. Is daily mortality associated specifically with fine particles? *J. Air Waste Manag. Assoc.* 1995 46, 927–939.
- Shi, Y., Lau, K.K.-L., Ng, E., 2016. Developing Street-Level PM<sub>2.5</sub> and PM<sub>10</sub> Land Use Regression Models in High-Density Hong Kong with Urban Morphological Factors. *Environ. Sci. Technol.* 50, 8178–8187. <https://doi.org/10.1021/acs.est.6b01807>
- Solomon, P.A., Crumpler, D., Flanagan, J.B., Jayanty, R.K.M., Rickman, E.E., McDade, C.E., 2014. U.S. National PM<sub>2.5</sub> Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of networks. *J. Air Waste Manag. Assoc.* 64, 1410–1438. <https://doi.org/10.1080/10962247.2014.956904>
- Strak, M., Janssen, N., Beelen, R., Schmitz, O., Vaartjes, I., Karssenbergh, D., van den Brink, C., Bots, M.L., Dijst, M., Brunekreef, B., Hoek, G., 2017. Long-term exposure to particulate matter, NO<sub>2</sub> and the oxidative potential of particulates and diabetes prevalence in a large national health survey. *Environ. Int.* 108, 228–236. <https://doi.org/10.1016/j.envint.2017.08.017>
- Subramanian, R., Donahue, N.M., Bernardo-Bricker, A., Rogge, W.F., Robinson, A.L., 2007.

- Insights into the primary–secondary and regional–local contributions to organic aerosol and PM<sub>2.5</sub> mass in Pittsburgh, Pennsylvania. *Atmos. Environ.* 41, 7414–7433. <https://doi.org/10.1016/j.atmosenv.2007.05.058>
- Subramanian, R., Khlystov, A.Y., Cabada, J.C., Robinson, A.L., 2004. Positive and Negative Artifacts in Particulate Organic Carbon Measurements with Denuded and Undenuded Sampler Configurations Special Issue of Aerosol Science and Technology on Findings from the Fine Particulate Matter Supersites Program. *Aerosol Sci. Technol.* 38, 27–48. <https://doi.org/10.1080/02786820390229354>
- Tan, Y., Dallmann, T.R., Robinson, A.L., Presto, A.A., 2016. Application of plume analysis to build land use regression models from mobile sampling to improve model transferability. *Atmos. Environ.* 134, 51–60. <https://doi.org/10.1016/j.atmosenv.2016.03.032>
- Tan, Y., Lipsky, E.M., Saleh, R., Robinson, A.L., Presto, A.A., 2014a. Characterizing the Spatial Variation of Air Pollutants and the Contributions of High Emitting Vehicles in Pittsburgh, PA. *Environ. Sci. Technol.* 48, 14186–14194. <https://doi.org/10.1021/es5034074>
- Tan, Y., Robinson, A.L., Presto, A.A., 2014b. Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. *Atmos. Environ.* 99, 333–340. <https://doi.org/10.1016/j.atmosenv.2014.10.003>
- The National Map: Elevation [WWW Document], 2017. URL <https://nationalmap.gov/elevation.html> (accessed 10.28.17).
- Turpin, B.J., Huntzicker, J.J., Hering, S.V., 1994. Investigation of organic aerosol sampling artifacts in the los angeles basin. *Atmos. Environ.* 28, 3061–3071. [https://doi.org/10.1016/1352-2310\(94\)00133-6](https://doi.org/10.1016/1352-2310(94)00133-6)

- Turpin, B.J., Saxena, P., Andrews, E., 2000. Measuring and simulating particulate organics in the atmosphere: problems and prospects. *Atmos. Environ.* 34, 2983–3013. [https://doi.org/10.1016/S1352-2310\(99\)00501-4](https://doi.org/10.1016/S1352-2310(99)00501-4)
- Urch, B., Brook, J.R., Wasserstein, D., Wasserstein, D., Brook, R.D., Brook, R.D., Rajagopalan, S., Corey, P., Silverman, F., 2008. Relative Contributions of PM<sub>2.5</sub> Chemical Constituents to Acute Arterial Vasoconstriction in Humans. *Inhal. Toxicol.* 16, 345–352.
- Urch, B., Silverman, F., Corey, P., Brook, J.R., Lukic, K.Z., Rajagopalan, S., Brook, R.D., 2005. Acute Blood Pressure Responses in Healthy Adults During Controlled Air Pollution Exposures. *Environ. Health Perspect.* 113, 1052–1055.
- US EPA, O., 2015. 2011 National Emissions Inventory (NEI) Data [WWW Document]. US EPA. URL <https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data> (accessed 10.28.17).
- Verma, V.; Fang, T.; Xu, L.; Peltier, R. E.; Russell, A. G.; Ng, N. L.; Weber, R. J. Organic aerosols associated with the generation of reactive oxygen species (ROS) by water-soluble PM<sub>2.5</sub>. *Environ. Sci. Technol.* **2015**, 49 (7), 4646–4656.
- Wang, M., Beelen, R., Basagana, X., Becker, T., Cesaroni, G., de Hoogh, K., Dedele, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H., Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M.J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., Brunekreef, B., 2013. Evaluation of land use regression models for NO<sub>2</sub> and particulate matter in 20 European study areas: the ESCAPE project.

- Environ. Sci. Technol. 47, 4357–4364. <https://doi.org/10.1021/es305129t>
- Yu, J.Z., Xu, J., Yang, H., 2002. Charring Characteristics of Atmospheric Organic Particulate Matter in Thermal Analysis. Environ. Sci. Technol. 36, 754–761. <https://doi.org/10.1021/es015540q>
- Zhang, J.J.Y., Sun, L., Barrett, O., Bertazzon, S., Underwood, F.E., Johnson, M., 2015. Development of land-use regression models for metals associated with airborne particulate matter in a North American city. Atmos. Environ. 106, 165–177. <https://doi.org/10.1016/j.atmosenv.2015.01.008>
- Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Allan, J.D., Coe, H., Ulbrich, I., Alfarra, M.R., Takami, A., Middlebrook, A.M., Sun, Y.L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P.F., Salcedo, D., Onasch, T., Jayne, J.T., Miyoshi, T., Shimojo, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R.J., Rautiainen, J., Sun, J.Y., Zhang, Y.M., Worsnop, D.R., 2007. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. Geophys. Res. Lett. 34, L13801. <https://doi.org/10.1029/2007GL029979>
- Zhao, Y., Nguyen, N.T., Presto, A.A., Hennigan, C.J., May, A.A., Robinson, A.L., 2016. Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines. Environ. Sci. Technol. 50, 4554–4563. <https://doi.org/10.1021/acs.est.5b06247>
- Zhao, Y., Nguyen, N.T., Presto, A.A., Hennigan, C.J., May, A.A., Robinson, A.L., 2015. Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol



Production. Environ. Sci. Technol. 49, 11516–11526. <https://doi.org/10.1021/acs.est.5b02841>

- Zhou, S., Collier, S., Jaffe, D.A., Briggs, N.L., Hee, J., Sedlacek III, A.J., Kleinman, L., Onasch, T.B., Zhang, Q., 2017. Regional influence of wildfires on aerosol chemistry in the western US and insights into atmospheric aging of biomass burning organic aerosol. *Atmos Chem Phys* 17, 2477–2493. <https://doi.org/10.5194/acp-17-2477-2017>
- Zhou, Y., Xing, X., Lang, J., Chen, D., Cheng, S., Wei, L., Wei, X., Liu, C., 2017. A comprehensive biomass burning emission inventory with high spatial and temporal resolution in China. *Atmospheric Chem. Phys.* 17, 2839–2864. <https://doi.org/10.5194/acp-17-2839-2017>
- Zhu, C.-S., Cao, J.-J., Tsai, C.-J., Shen, Z.-X., Han, Y.-M., Liu, S.-X., Zhao, Z.-Z., 2014. Comparison and implications of PM<sub>2.5</sub> carbon fractions in different environments. *Sci. Total Environ.* 466–467, 203–209. <https://doi.org/10.1016/j.scitotenv.2013.07.029>

**Chapter 4: Comparison of spatial and temporal variation of airborne pollutants using mobile and distributed sampling**

## **Chapter 4**

# **Comparison of spatial and temporal variation of airborne pollutants using mobile and distributed sampling**

## 4.1 Introduction

One major challenge with mobile sampling is that it is difficult to collect sufficient data to resolve long-term average (e.g., seasonal or annual) concentrations, because mobile sampling always convolves spatial and temporal variations in pollutant concentrations. Apte et al. (2017) showed that multiple, repeated drives over a small area can recreate long-term average concentrations if individual areas are sampled on >20 unique days; however, such an intense sampling strategy may not be feasible in all areas.

Since mobile sampling is conducted at different sites and different times of day, another major challenge is to accurately separate spatial variation (between site) from temporal variation (within site). This separation is often achieved by temporal data correction based on a continuous central reference site, however, few studies address the robustness of this methodology (Brantley et al., 2014), or quantify the relative magnitude of spatial and temporal variations (Sullivan and Pryor, 2014; Yu et al., 2016). Yu et al. used concurrent parallel mobile sampling to characterize spatial and temporal variation of traffic-related air pollutants in an urban community, and suggested separation of temporal variation from spatial variation was important for more accurately estimating human exposure (2016). Sullivan et al. (2014) used combined fixed and mobile measurements, and indicated PM<sub>2.5</sub> spatial variability was greater than temporal variability on short time scales (hourly). Apte et al. (2017) conducted a yearlong extensive mobile sampling with Google Street View cars, and concluded spatial variability generally dominated over temporal variability for NO, NO<sub>2</sub>, and BC. Additional studies are still needed to quantify relative magnitude of spatial and temporal variation, and help explain precision and robustness of long-term exposure models such as LUR based on short-term measurements (Allen et al., 2011; Vienneau et al., 2010).

LUR models quantify spatial variation in pollutant concentrations and are mainly used

for estimating chronic exposures. The use of LURs for exposure estimation assumes that spatial variations exist in addition to temporal variations. Exposure at a given location is the combination of background and local source impacts. The background can change day-to-day or within a day because of weather and regional transport. The local signal is more dependent on nearby sources (Tan et al., 2014a, 2016). Long-term epidemiology studies focus on spatial variation, and LUR proves an adequate tool to capture pollutant spatial gradients (Gerard et al., 2008). But temporal variation also exists in pollutant exposure. Research is needed to compare spatial and temporal trends for multiple pollutants, and answer questions whether spatial and temporal variations in human exposure are the same for all pollutants (e.g., does CO exposure behaves the same as exposure to NO<sub>2</sub> or PN).

People have been making LURs for ~20 years, but no optimal sampling strategy exists (Hatzopoulou et al., 2017). Most studies use one sampling strategy or another (e.g., distributed filters or mobile sampling) to collect the data needed to build LUR models. With a combined sampling network including mobile and distributed samplers with high time resolution, we can start probing these methods and think about refined sampling strategies. For example, different sampling strategies may be needed in urban versus rural areas.

In this study, we designed an extensive sampling network including both distributed monitors and mobile sampling in Pittsburgh, PA. The main objectives are to quantify: 1) spatial and temporal variation of gaseous and particle pollutants (CO<sub>2</sub>, CO, NO<sub>2</sub>, PN, and PM<sub>2.5</sub>); 2) temporal variations in different scales, e.g., diurnal vs. daily; 3) spatial vs. temporal variations. We use both mobile and stationary monitor data as a basis to compare spatial and temporal variations in different micro-environments, and make some inferences about future sampling strategies for building LURs.

## 4.2 Methods

### 4.2.1 Sampling platform setup

Our sampling network includes three components, a mobile sampling van, 14 distributed monitors, and a supersite. We synchronized clocks on all instruments, and data were all reported in eastern standard time.

The mobile laboratory was a gasoline powered van, equipped with gas and particle instruments. The mobile lab has been described in detail by Li et al (2016) and Tan et al (2014a). Ambient air is pulled through an 1/2" O.D. stainless tube installed on top of the van. The flow rate is controlled by mechanical pumps at 16.7 SLPM, and a PM<sub>2.5</sub> cyclone removes coarse particles in air flows before they reach instruments. A customized onboard generator dispatches power to all instruments. GPS information is logged by Bad Elf GPS Pro with 1 s resolution (Bad Elf, CT, USA). A High-Resolution Aerodyne Aerosol Mass Spectrometer (AMS, Aerodyne Research Inc., MA, USA) operating in fast-MS mode (Kimmel et al., 2011) measures submicron (less than 1  $\mu\text{m}$ ) particle composition with 20 s resolution (DeCarlo et al., 2006; Jayne et al., 2000). An Aethalometer AE-33 (Magee Scientific, CA, USA) measured black carbon (BC) with 1 min resolution. AMS output (organic and inorganic concentrations) plus BC yields PM<sub>1</sub> concentration. A fast mobility particle sizer spectrometer (FMPS, TSI Incorporated, MN, USA) measured PN concentration (5.6 to 560 nm) with 1 s resolution. Gas analyzers include T200 NO<sub>x</sub>, T300 CO (Teledyne Technology, CA, USA), and LI-820 CO<sub>2</sub> analyzer (LI-COR Biosciences, NE, USA). The gas analyzers all report data every 1 s.

Distributed monitors measured gaseous pollutants with Real-time Affordable Multi-Pollutant (RAMP) sensors (SenSevere LLC, PA, USA) (Zimmerman et al., 2017). The RAMP uses Alphasense electrochemical sensors (Alphasense Ltd., UK) to measure NO<sub>2</sub>, CO,

O<sub>3</sub>, and SO<sub>2</sub>. An additional non-dispersive infrared (NDIR) sensor measures CO<sub>2</sub>, temperature (method: bandgap) and RH (method: capacitive). Voltage signals from individual sensors installed on the RAMPs were converted to concentration using a machine learning calibration detailed by Zimmerman et al. All RAMP data are reported at 15-minute resolution. PM<sub>2.5</sub> at each distributed site was measured via nephelometry with a MetOne Neighborhood PM Monitor (MetOne Instruments, Inc., OR, USA). Eight sites had PN measurements from water-based condensation particle counters (CPC, Aerosol Dynamics Inc., CA, USA) (Hering et al., 2017). To ensure consistency, cleaned final data output from distributed monitors are reported at 15-minute resolution, including gaseous pollutants, PM<sub>2.5</sub>, and PN.

The supersite was on the Carnegie Mellon University Campus, in Pittsburgh, PA. It is ~500 m away from major roads, and at least 40 km away from any point sources with more than 50 tons of annual PM<sub>2.5</sub> emission in the predominant wind direction. Gaseous pollutants measured at the supersite are reported in 1 s resolution, including CO<sub>2</sub> (LICOR 820), NO<sub>2</sub> (2B Technologies Model 405 nm), CO (Teledyne T300U), O<sub>3</sub> (Teledyne T400), and SO<sub>2</sub> (Teledyne T100A). Particle number (7.5 nm to 316.2 nm) is measured with scanning mobility particle sizer (TSI Incorporated, MN, USA) every 3 minutes. The Supersite also had a RAMP with a MetOne PM<sub>2.5</sub> nephelometer. The supersite operated continuously for the entire sampling period, and therefore provides a good reference of urban background concentrations for multiple pollutants.

We performed calibration for instruments in all three sampling platforms. Gas monitors in the mobile van and supersite were calibrated weekly. The AMS ionization efficiency was calibrated monthly. The CPCs at distributed monitors undercount relative to the TSI SMPS, which we treat as a reference. Thus, raw CPC output was scaled based on

linear regressions from collocations conducted in our laboratory. Calibration of RAMP was based on machine learning technique random forest (Zimmerman et al., 2017). PM<sub>2.5</sub> measured by Met One are corrected for different RH environments based on hygroscopic growth factor and reference PM<sub>2.5</sub> monitor (Petters and Kreidenweis, 2007).

The CPCs at distributed monitors had weekly maintenance, mainly refilling water reservoirs. FMPS was zero checked prior to each mobile sampling trip with HEPA filtered air. AMS was zero checked after each trip with HEPA filtered air.

#### **4.2.2 Sampling overview**

We selected 14 sites to deploy our distributed monitors. Sites were selected using stratified sampling with traffic, restaurant density, and building height as the selection variables (Allegheny County Information Portal, 2017; Department of City Planning, 2017; City of Pittsburgh GIS Data, 2015; US Census Tiger, 2010). High traffic density sites were classified based on vehicle miles travelled, number of traffic lights, and road length within a 200 m by 200 m grid cell. High restaurant density locations were designated as having 3 or more restaurants within a 200 m by 200 m grid cell. We excluded restaurants with no obvious emissions source, such as ice cream and coffee shops. Building height suggests street canyon effect. The criteria to select tall buildings (those that likely influence surrounding air patterns) was when the average height of buildings in a 200 m by 200 m grid cell was over 40 ft.

The distributed monitors were deployed on two separate case studies—urban-rural transect and downtown (Table 4.1, Figure 4.1). The nine sites in the transect campaign were distributed from upwind, through the urban center, to downwind locations. The predominant wind direction is from the southwest (Figure D1, Supplemental Materials). The six sites in the downtown campaign are much closer in space compared with ones along the transect (Figure 4.1). The transect campaign spanned summer/fall seasons from August 2016 to



December 2016. The downtown campaign started from January 2017 to December 2017. Site 9 and the Supersite (Site 15) were operated continuously during both campaigns.

Mobile sampling was conducted in a  $\sim 1 \text{ km}^2$  surrounding box near each distributed monitor (Figure 4.1). Mobile sampling at each site was spread across the day. Total visits during morning rush hour (5 am to 9 am) or afternoon/evening (11 am to 9 pm) were at least 3. Downtown sites have more visits as sites were closer to each other, and less labor required to cover all sites. For each mobile sampling trip near a monitor, we tended to cover all main roads in the corresponding box at least once. The average required driving time around each distributed monitor was one hour.

#### **4.2.3 Data preparation and analysis**

Various data wrangling processes and QA procedures were performed. Extreme measurements are defined as values either larger than 75<sup>th</sup> percentile plus 4 times the interquartile range or less than 25<sup>th</sup> percentile minus 4 times the interquartile range. Data from the stationary monitors were filtered for extreme and null values. Mobile measurements are expected to show larger variability, as it is more frequently impacted by local sources. No extreme value cleaning procedure is applied on it. We removed any CPC data with error codes, and counts less than  $100 \text{ \#/cm}^3$  or bigger than  $1,000,000 \text{ \#/cm}^3$  (random spikes sometimes due to power cycling). Any  $\text{PM}_{2.5}$  data with error flags or  $\text{PM}_{2.5}$  concentrations larger than  $600 \text{ \mu g/m}^3$  were filtered. GPS data were aligned with mobile pollutant measurements after accounting for residence time in the sampling lines. The FMPS tended to undersize accumulation mode particles ( $>80 \text{ nm}$ ) and we applied a calibration relationship based on Zimmerman et al (2015).

We applied temporal correction to our mobile sampling dataset as background concentrations could change diurnally or from day to day (Klompniaker et al., 2015; Masiol

et al., 2017). Previous research commonly used a single background supersite for temporal correction via addition or multiplication method (de Hoogh et al., 2013; Tan et al., 2014b; Wang et al., 2013). However, this method relied on the assumption that temporal variation was not spatially variable (Apte et al., 2017; Van den Bossche et al., 2015). Here we first applied wavelet decomposition to 15-minute average measurements at each distributed monitor located across the whole county (Klems et al., 2010). We defined signals with frequency larger than 8 hours to be the regional background represented by the monitor. We then compared the 8-hour frequency signals from all distributed monitors in each 15-minute window and selected the smallest one as the regional background for that specific 15 minutes period. Raw mobile measurements were compared with the corresponding regional background, and we defined local enhancement as raw mobile measurements minus the background. The enhancement represented the portion of spatial signal in the raw measurement. As only supersite has consistent PN measurement throughout campaigns, we instead used the rolling hourly fifth percentile of mobile measurements as the PN regional background.

Coefficient of variation (COV) is used to compare the relative magnitude of spatial and temporal variation (R. C. Sullivan and Pryor, 2014). It is calculated as the standard deviation divided by the mean.

Intraclass correlation (ICC) compares the variability between groups versus within each group. It is based on analysis of variance (ANOVA) framework, and is calculated as the ratio of variation between groups (mean square between groups, MSP) to the sum of MSP and variation within groups (mean square errors, MSE). By definition, ICC ranges from 0 to 1. An upper limit 1 means that within-group variation is much smaller than between-group variations ( $MSP \gg MSE$ ). In contrast, 0 means significant variability within groups relative

to differences between groups ( $MSE \gg MSP$ ). Figure D2 in Supplemental Materials provides an example showing the difference between high and low ICC values. The interpretation for ICC is as follows: less than 0.4 – small or little between group difference, between 0.40 and 0.59 – adequate between group difference, between 0.60 to 1 – substantial between group difference (Cicchetti, 1994).

$$ICC = \frac{MSP}{MSP + MSE}$$

Mobile measurements are grouped into 50 m fishnet boxes following Apte et al (2017). After we corrected mobile measurements' temporal variation by wavelet decomposition, the derived enhancement values were then averaged to represent each grid box.

### 4.3 Results and Discussion

#### 4.3.1 Sampling domain overview

Figure 4.1 shows the sampling sites' stratification overview. Sites 1, 2, 3, 4 are essentially upwind background sites in Table 4.1. Sites 5, 6, 7, 10 are under combined influences of traffic, street canyon, and restaurant emissions. Site 8 is an urban background location in a residential neighborhood. Sites 9 and 11 have high restaurant density. Site 12 has restaurants and green space mixed inside. Site 13 has high traffic volume, and site 14 is a downwind background. Site 15 is the supersite in university campus.

Figure 4.1 also shows UFP spatial variation in the upwind background site 2, urban core sites 5, 6, 7, and a residential area with high density of restaurants site 11. Background UFP concentration is around  $6257 \text{ \#}/\text{cm}^3$  based on supersite measurements. The mean temporally corrected mobile measurements (local enhancement) are calculated for 50 m square grid boxes. The legend on the middle right indicates magnitude of the averages.

Upwind sites such as 1, 2, 3 and 4 are essentially background sites (blue) as no obvious UFP sources are near these sites. Sites in downtown (5, 6 and 7) have large UFP enhancement, up to a factor of 7 above background due to combined influence of vehicle emissions, restaurants, and street canyon effect. Not only is there enhancement, the enhancement is spatially variable indicated by mix of red and blue grid boxes. Site 11 has the highest average UFP enhancement. Hotspots (red grid boxes) indicate a dense restaurant district with high traffic. Air pollution from restaurants has an influencing spatial extent around neighborhood scales (~1 km). Site 11 is both a residential and recreational area. This could exacerbate air pollution health impacts. People usually go to site 11 for dining at restaurants or bars and spend a bit portion of their daily time there, but they are still exposed to a large fraction of their daily accumulated UFP exposure during this time window.

#### **4.3.2 Concurrent mobile vs. distributed monitors**

Concurrent NO<sub>2</sub> from mobile sampling (1 s measurement) are generally comparable to that from distributed monitors (15-min resolution) as shown in Figure 4.2. Site 15 is the urban background supersite, and there was no mobile sampling around this location. Site 14 is omitted as measurements fail QAQC check. Site 13 has the largest discrepancy of NO<sub>2</sub> measurement between mobile and distributed sampling. Site 13 is surrounded by a highway (North) and one major road (South). The distributed monitor (blue boxplot) was placed on the edge of the highway, and measures higher concentrations of NO<sub>2</sub> than mobile on-road measurements nearby. Even though mobile sampling was conducted quite near the highway, it did not capture high emissions from vehicle sources. This illustrates the quick decay of traffic related pollutants within hundreds of meters away from roads (Karner et al., 2010). The complex river valley topography and predominant southwestern wind could also help explain the lower NO<sub>2</sub> concentrations in the driving domain, as the mobile sampling route

was often upwind of the highway.

As sites 5, 6, 7 are within a single 1 km<sup>2</sup> box, they were assigned the same mobile sampling boxplot stats in Figure 4.2. Site 7 measures the highest NO<sub>2</sub> concentrations among these 3 clustered downtown sites. It is located ~10 m away from a pizza restaurant exhaust fan in a narrow street canyon. Under this local source impact, it measures high NO<sub>2</sub> with a similar magnitude as the monitor installed next to a highway (site 13). Site 6 is near a bus stop, and the RAMP monitor is placed at about the same height as bus exhaust tailpipe relative to the ground. It is under consistent diesel exhaust pollution during daytime hours, and this partially explains why site 6 measures higher NO<sub>2</sub> compared with site 5. Site 5 is ~2 m above the ground, placed next to a 2-lane major road in downtown.

Site 8 is ~2 km downwind from the urban center sites (5, 6, 7). It has spatial nonuniformity in terms of traffic volume on roads inside the 1 km<sup>2</sup> driving box. The distributed monitor at site 8 measures similar NO<sub>2</sub> as the urban center sites (5, 6, and 7). However, other driving roads inside the 1-km<sup>2</sup> box used for mobile sampling have significantly less traffic volume. NO<sub>2</sub> from the mobile measurement represents the average traffic intensity in the whole box instead of a specific major road, and is thus slightly lower than values from the distributed monitor.

Sites 1, 2 and 3 are upwind background, and have lower NO<sub>2</sub> compared with downtown sites. Site 1 has lower NO<sub>2</sub> mobile measurements compared to monitor values. A portable storage company is 200 m upwind of site 1. Thus, the NO<sub>2</sub> discrepancy might be due to the distributed monitor capturing off road or on road vehicle emissions from the point source while mobile van does not.

Spatial variation exists between sites and within sites (~1 km). Within site variability is higher in downtown locations indicated by the interquartile range of the boxplots in Figure

4.2, especially sites 5, 6, 7, and 11, which are all impacted by large numbers of restaurants and high traffic volume. Figure 4.2 shows significant spatial variability down to the street level, and this suggests spatial representativeness of stationary urban monitors could be quite limited (Piersanti et al., 2015; Righini et al., 2014; Vitali et al., 2016).

#### **4.3.3 Temporal variation in two scales – daily vs. diurnal**

Pollutant concentrations vary over certain relevant timescales – hourly, daily, seasonal, and annual. This section focuses on variations over hourly and daily timescales using 15-minute resolution measurements from the RAMPs.

Figure 4.3 shows temporal variation of CO at downtown site 6 and background site 14. For downtown site 6, the day-to-day variation of median CO concentration is larger than within-day variation. Interquartile range (IQR: 75th percentile minus 25th percentile) of mean daily CO is 231 ppb. Among all 45 sampling days at site 6, only 22 % of days have a larger IQR value of within-day CO. On the contrary, for the background site 14, within-day variation is bigger than daily variation. The IQR of daily CO at site 14 is 45 ppb. Among a total of 50 sampling day, 74% of days have larger IQR of within-day CO.

Figure 4.3(a) indicates Tuesday has the highest CO concentrations at downtown site 6. CO during weekdays are generally higher than CO during weekends as expected by large decrease in traffic volume during weekends (Alghamdi et al., 2014). Boundary layer shifting and traffic flow change are the main causes for the diurnal temporal variation in Figure 4.3(b). One CO peak period is the morning rush hour ~8 a.m., and the other is around 5 p.m. Downtown site 6 is consistently impacted by traffic and other primary sources throughout the whole day, while traffic activity changes significantly between weekdays and weekends. We computed ICC of CO<sub>2</sub>, CO, NO<sub>2</sub>, PN, and PM<sub>2.5</sub> measurements grouped by each day. ICC for temporal CO difference at site 6 is 0.56 in Table C1 (Supplemental Materials). ICC values

between 0.4 and 1.0 indicate more intergroup differences and more similarity within groups. Daily CO variation has higher ICC values, indicating more difference between days than variation within a single day (diurnal variation). Overall, day-to-day CO variation is higher than diurnal CO variation in urban center.

Figure 4.3(c) shows throughout the whole week, no specific days have higher CO concentrations than others at background site 14. Figure 4.3(d) shows that peaks around 7 a.m., likely reflecting a combination of boundary layer effects and regional traffic contributions.

Figure 4.4 compares PN at two sites – site 7 (downtown) is heavily influenced by local traffic and cooking sources, and site 15 (supersite) represents the urban background.

Site 7 is located in a narrow pedestrian alley near a restaurant exhaust fan; emissions from the restaurant significantly impact measurements at this location. Site 7 shows large within-day changes in PN due to cooking activities. There are periodic spikes due to the non-continuous nature of cooking at the restaurant, though there is a consistent peak starting at midday and running through evening. At this site, daily variation of PN was smaller than its diurnal variation. When PN was compared across days, PN did not change much between weekday vs. weekend, and the business working hours for restaurants were consistent from day to day.

On the contrary, PN at the supersite was around  $8000 \text{ \#}/\text{cm}^3$ , on average nearly an order of magnitude smaller than PN measured near the restaurants. PN concentrations at the supersite was only close to PN at site 7 during overnight periods (~midnight to 1 a.m.), when cooking and traffic activity near site 7 fell and the site 7 looked more like background. Day-to-day variation of PN at the supersite was larger than diurnal variations, in contrast to PN temporal pattern near restaurants at downtown.

Overall, the relative magnitude of the difference between two temporal variations for a specific pollutant depends on sources and land use near the monitor locations. We used coefficient of variation (COV) to quantify temporal variation in daily or diurnal scale. Daily COV values compares daily mean concentrations across multiple days. Diurnal COV value compares hourly averages within a day.

Figure 4.5 shows temporal variation in two scales for five pollutants. CO<sub>2</sub> daily and diurnal variations are both small and nearly identical, since CO<sub>2</sub> is dominated by the global background.

CO is a primary pollutant. Incomplete fuel combustion is major source of ambient CO (US EPA, 2016). As shown in Figure 4.3 for site 6, CO tended to have higher daily variation than diurnal variation near the urban center. This pattern was observed in sites 4, 5, 6, 7, and 10. Significant traffic pattern and volume changes between weekday and weekend might partially explain the difference. However, as shown in Figure 4.3(a), there are also significant day-to-day differences in CO at urban sites between weekdays, suggesting that weekday/weekend differences in traffic patterns are not the only driver for the observed daily variations. Figure 4.3(a) suggests that traffic differences between weekdays can be as significant as weekday/weekend differences.

Background and urban background sites 1, 2, 12, 14, and 15 all have larger within-day (diurnal) variations than day-to-day variations of CO. These sites follow the same general pattern as shown in Figure 4.3(c) and 4.3(d), though these sites are a mixture of suburban/rural (sites 1, 2, and 14) and urban background (sites 12 and 15).

Sites 8 and 11 are urban sites where the daily and diurnal CO variations are similar in magnitude. Site 8 is in a residential neighborhood adjacent to downtown, and site 11 is in a neighborhood with high restaurant and traffic density.



Overall, the general trend for CO is that urban, source-impacted sites have slightly higher daily variations than diurnal variations, and background sites have higher diurnal variations. Two sites do not fit this narrative. Site 9 is in an urban neighborhood with tall buildings and high restaurant density. Based on nearby sites 5, 6, and 7, we might expect larger daily variations, but the opposite is observed. Similarly, site 13, which is adjacent to a highway, might be expected to have larger daily variations, but daily and diurnal variations are nearly identical.

NO<sub>2</sub> has a similar pattern as CO, and overall the magnitude of COVs is similar for NO<sub>2</sub> and CO. Traffic impacted urban sites have larger NO<sub>2</sub> daily variability than diurnal variability, whereas background sites have larger diurnal variability. Site 9, which has lower than expected daily variation in CO, also has low daily variation for NO<sub>2</sub>. One difference is for the near highway site 13, which has larger daily variation than diurnal variation in NO<sub>2</sub>, similar to traffic impacted urban sites.

PN has the largest temporal variation of the measured pollutants. PN, especially UFP, is highly dynamic (Cattani et al., 2017; Donahue et al., 2016; Kerckhoffs et al., 2017, 2016). For sites 3, 14 and 15, suburban and urban background sites, daily PN variation is larger than diurnal variation. These sites have minimal impact from nearby sources. Nucleation might be an important contributor to day-to-day variations at these sites. During our campaign from September 2016 to February 2017, supersite PN measurement indicates 18% nucleation frequency, and these nucleation events contribute around 21% of PN concentration increase at the supersite. For these background sites without nearby primary sources such as traffic or restaurants as places in downtown, PN will be higher mostly due to nucleation events occurring. For a day without nucleation events, PN concentration would fluctuate much less and change smoothly as a regional background. The source impacted sites are influenced by

nucleation too, but have additional primary source influences.

Sites 4, 5, 6, 8 and 11 have similar daily and diurnal variations. These sites all have nearby source impacts from traffic and/or cooking. The only site where diurnal variation is larger than daily variation is site 7, which as described above is heavily impacted by a cooking source. These data suggest that PN at source-impacted sites is very dynamic, with concentrations changing significantly on both daily and sub-daily scales.

PM<sub>2.5</sub> is mostly secondary (Robinson et al., 2007; Seinfeld and Pandis, 2016), and shows the least intersite difference. Daily COV is nearly identical for all the sites, confirming the largely regional nature of PM<sub>2.5</sub>. It also suggests that most of the variation can be described as temporal variation with a single monitor, and may be why many PM<sub>2.5</sub> LUR models have a central site as an important predictor (Eeftens et al., 2012; Li et al., 2017, 2016). Within-day PM variations tend to be small except in cases with a large nearby source like site 7. Even some of the other source impacted sites in downtown that are heavily impacted by bus traffic (e.g., site 6) are dominated by the day-to-day variations.

ICC magnitude also supports the results presented in Figure 4.5 (Table C1, Supplemental materials). ICC of PM<sub>2.5</sub> at most sites indicates either adequate (0.40 to 0.59) or substantial (0.60 to 1) between group differences. These higher ICC values mean less variability within the same day (diurnal variation), but more variability between days (daily variation). Thus, in Figure 4.5, high ICC specifically means daily COV (black dot) will be above median or upper hinge of the boxplot (diurnal COV). On the other hand, low ICC values (less than 0.15) mean daily COV (black dot) will be well below the median or lower hinge of the boxplot, as CO and NO<sub>2</sub> at background sites 1 and 2.

Overall, there is not a single temporal pattern across pollutants at different sites. Looking just at the median values across all sites, CO and NO<sub>2</sub> have slightly larger hourly

variation than day-to-day variation. PM<sub>2.5</sub> and PN have larger day-to-day variations. But patterns are not consistent across all sites. Background sites (1, 2, 12, 14) tend to have larger diurnal than daily variations for CO and NO<sub>2</sub>. For downtown sites, the reverse is true. But lots of sites have similar hourly and daily variation. PN is consistently the most variable across all sites. In general, daily and hourly PN variations are similar – exceptions are background sites (daily > hourly) and the site with the restaurant (source dominates, hourly >> daily). For PM<sub>2.5</sub>, daily variations are consistent across all sites, and daily variation is bigger than hourly except for the restaurant site.

One implication of the results shown in Figure 4.5 is that pollutant spatial patterns change with time over the day (or between days), and that the changes are pollutant specific. A simple background correction that lifts or drops all sites may be too simple to fully capture between-day or within-day variations, especially in cases where time- or location-weighted exposures are a desired endpoint.

#### **4.3.4 Spatial vs. temporal variation**

Figure 4.6 shows the comparison between spatial and temporal variation. Spatial variation was calculated based on temporally corrected mobile measurements inside each ~1 km<sup>2</sup> driving box. The resolution of temporal variation was set to daily. Average days for temporal variation calculations are around 60 days, or 2 months. Average sampling days for spatial calculation are around 10 days. Note here, mobile measurements are generally in 1 s resolution, while distributed monitors report data in 15 minutes resolution.

For almost every pollutant at every sampling location, spatial variations as measured by COV were larger than temporal variations. CO<sub>2</sub> spatial variations are nearly an order of magnitude larger than temporal variations. This is likely due to the mobile laboratory being sensitive to CO<sub>2</sub> emissions from nearby vehicles.

CO, NO<sub>2</sub> and PN are more spatially variable than temporally variable at every single site. This reflects variations in source impacts. As shown in Figures 4.1 and 4.4, spatial differences of PN are on the order of a factor of 2-3 or more, but the temporal variations are smaller. Similarly, Figure 4.3 shows large spatial differences in CO (~100 ppb) between sites 6 (downtown) and 14 (suburban). Mobile sampling routinely identifies variations of similar magnitude within the individual 1 km<sup>2</sup> driving domains.

Several sites (2, 3, 4, 9, 10, and 13) have PM temporal variations that are larger than spatial variations. We cannot determine an obvious reason for this observation, as these sites span a range from background to near-highway. For example, site 4, while located near downtown, is largely classified as “low” for all of the land use variables. Thus, it might be reasonable to expect a lack of PM sources and this spatial uniformity at this site. However, we also observe smaller spatial than temporal variations at sites 9 and 10, which have much more variation in the land use in the 1 km<sup>2</sup> mobile sampling domain.

One possible explanation for the PM results is that the mobile laboratory measures PM<sub>1</sub> whereas the stationary sites measure PM<sub>2.5</sub>. However, since most PM<sub>1</sub> constitutes a large fraction of PM<sub>2.5</sub> mass (Buczyńska et al., 2014), this seems unlikely.

As mobile sampling is conducted in different locations and times of day, this sampling platform collects spatial signals as well as temporal variations. If the temporal variations dominate over spatial variations, then it will be very hard to separate these two variations, and get the target spatial signals in this sampling approach. On the other hand, if spatial signals are much higher than temporal ones, the implications are, with an appropriate temporal correction method, the dataset can be distilled to provide us with the true spatial variations.

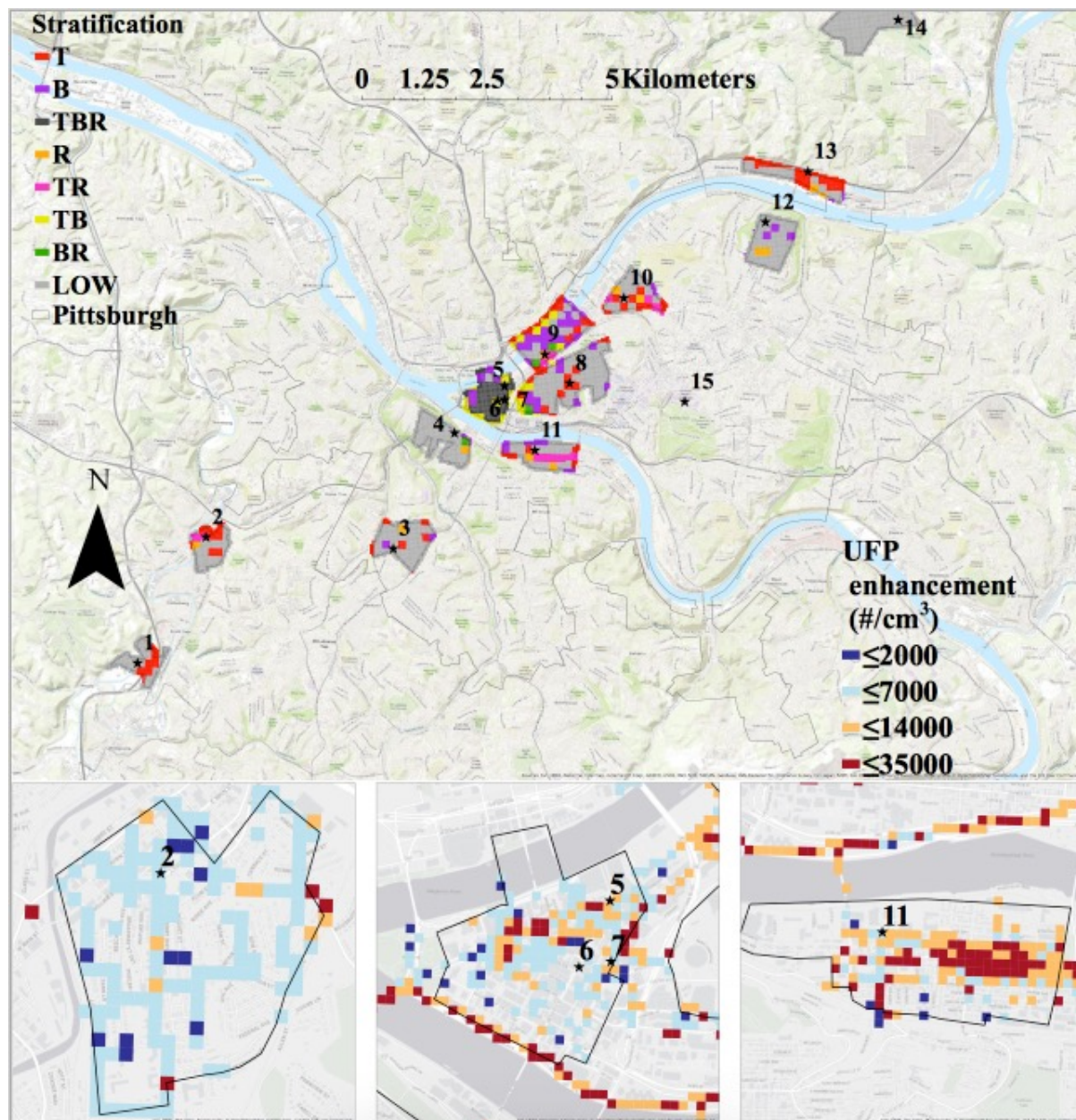
Overall, all five pollutants show higher spatial variability than temporal variability.

This suggests that when we build LUR based on short term mobile measurements, the potential harmful effect of temporal fluctuations in short term mobile sampling is limited. With an appropriate background correction method, mobile sampling is reliable to infer long term spatial contrasts.

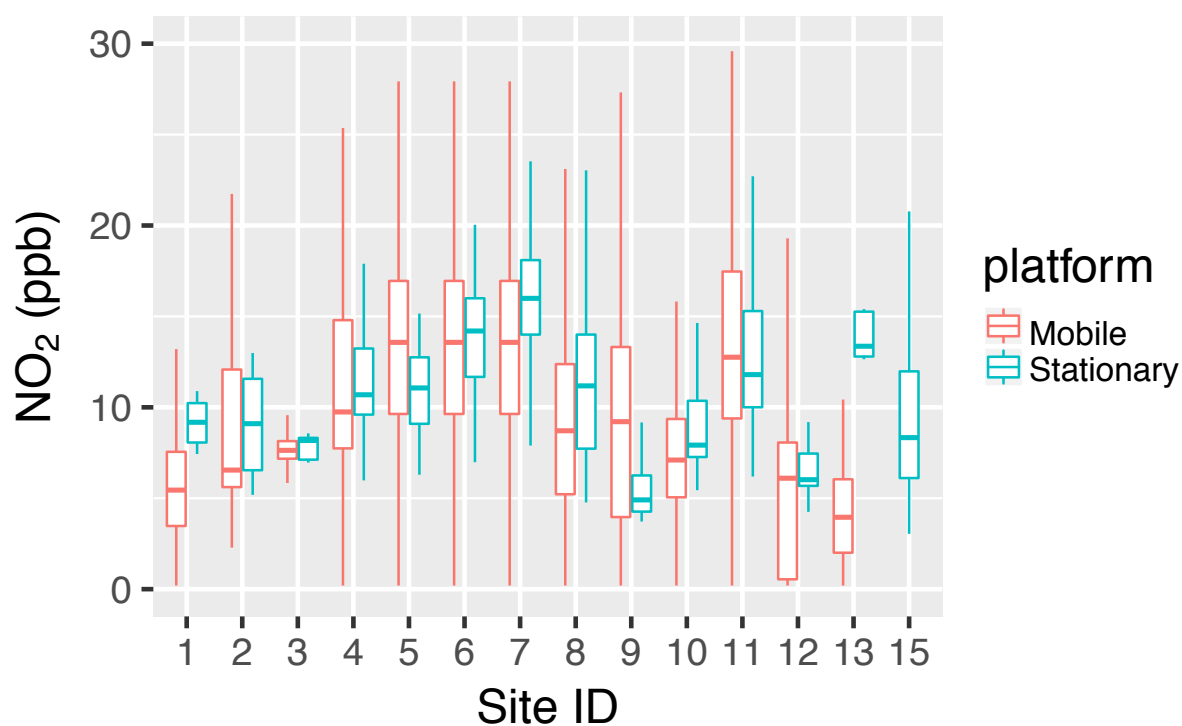
#### **4.3.5 COV comparison with other studies**

Figure 4.5 and 4.6 used COV metric to compare relative magnitude of spatial and temporal variation. Other studies such as Sullivan et al. (2014) showed diurnal COV of PM<sub>2.5</sub> temporal variation was 0.20 (SD = 0.05). Our diurnal PM<sub>2.5</sub> COV is around 0.22. Their studies also showed spatial variability of PM<sub>2.5</sub> was a factor of 2 of the temporal variability.

Yu et al. (2016) showed COVs in concurrent measurements from two parallel routes were around 0.21 for all traffic related air pollutants (BC, PN, PM<sub>2.5</sub>, and CO). Their COV was calculated by the standard deviation of two paired same time measurements divided by the mean of the two. The relative small sample size (2) might not give a good estimate of the general spatial or temporal variation.

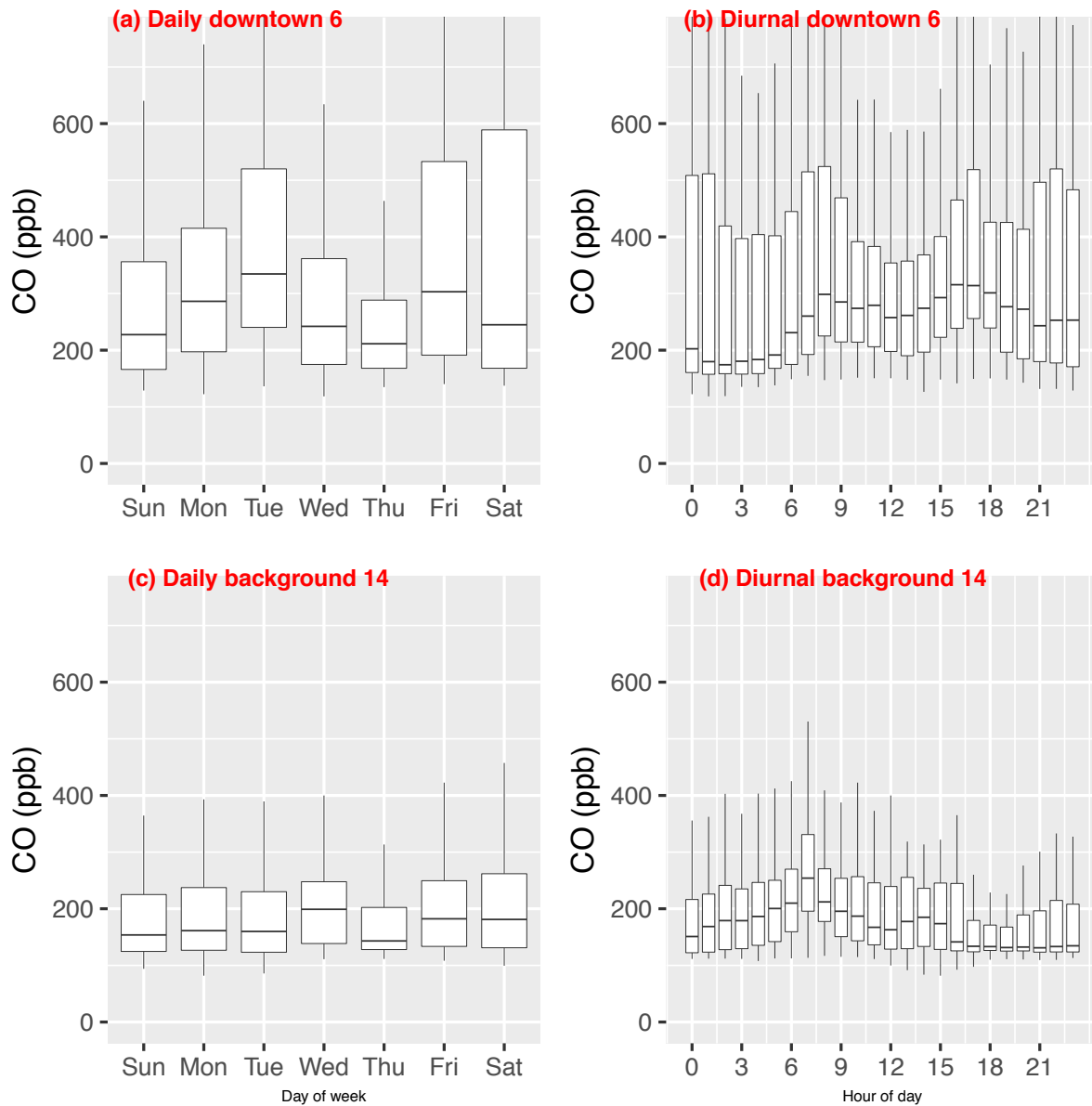


**Fig. 4.1.** 15 sampling sites overview with stratification information labelled based on 50 m grid boxes. T: high traffic density, B: high building height, R: high density of restaurants, LOW: low source density. Stars are the locations of distributed monitors. Polygon areas are around  $\sim 1 \text{ km}^2$ , and mobile sampling is conducted repeatedly inside the polygons to study spatial variability around each monitor. UFP spatial variation in 3 polygons with 50 m spatial resolution are shown in the bottom panels.



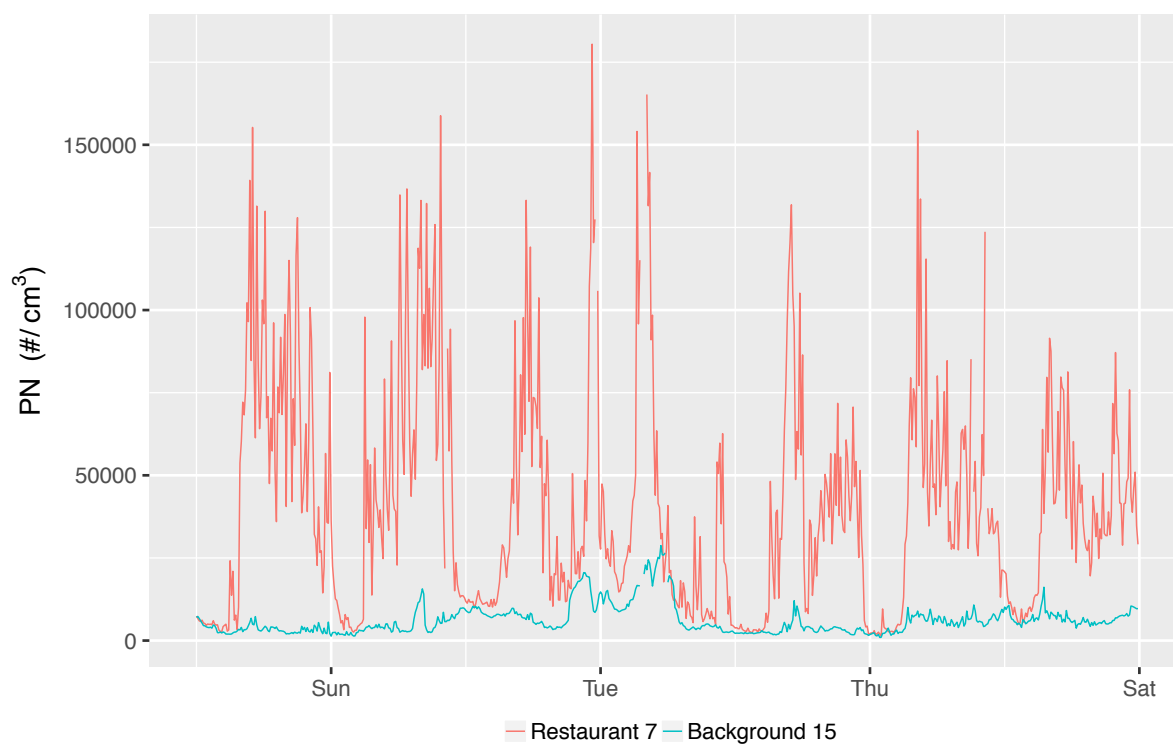
**Fig. 4.2.** Concurrent NO<sub>2</sub> measurement comparison between distributed monitors and mobile

measurements inside corresponding 1 km<sup>2</sup> box. The top and the bottom of the box represent the 75th and 25th percentiles. The line inside the box is the median. The outer line extends to the most extreme concentrations not classified as outliers.

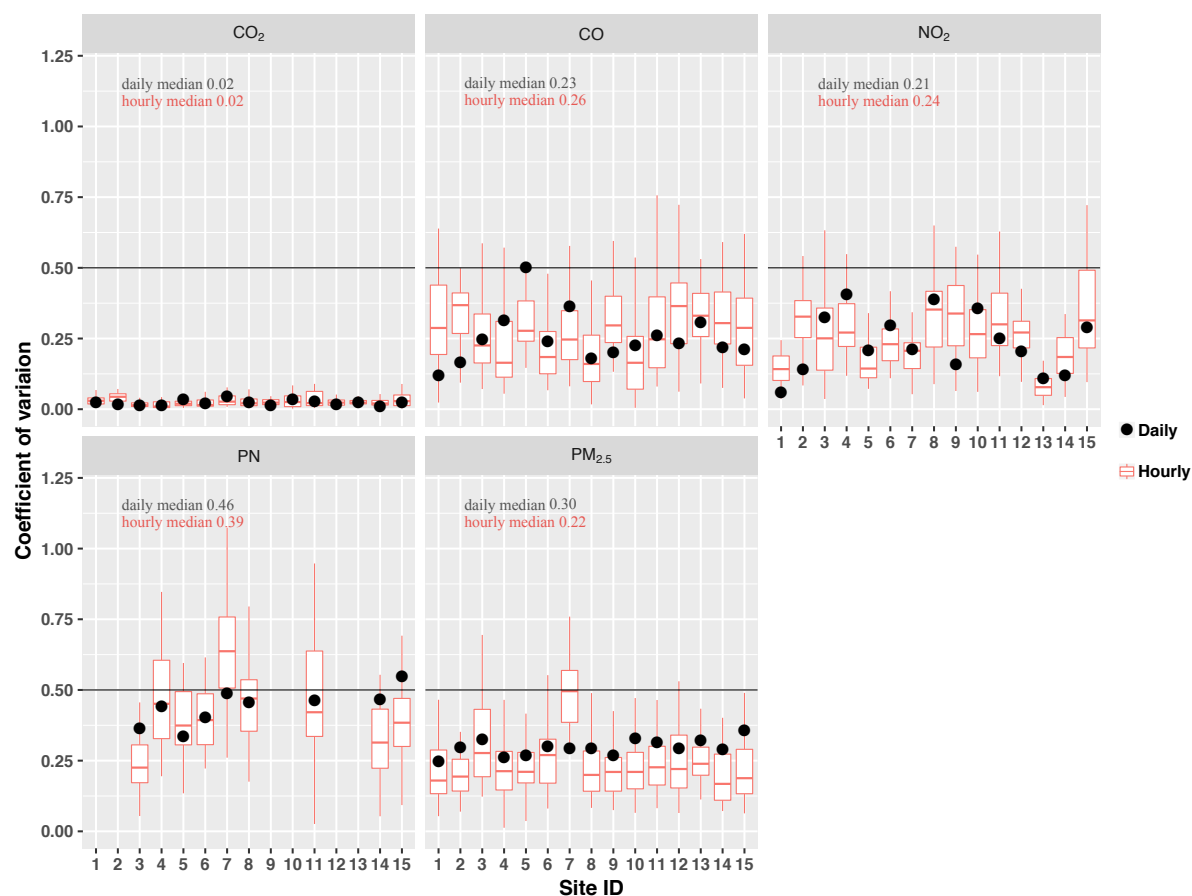


**Fig. 4.3.** CO measurements from monitor 6 in downtown and a downwind background monitor 14. Measurements are plotted with respect to day of the week or hour of the day.

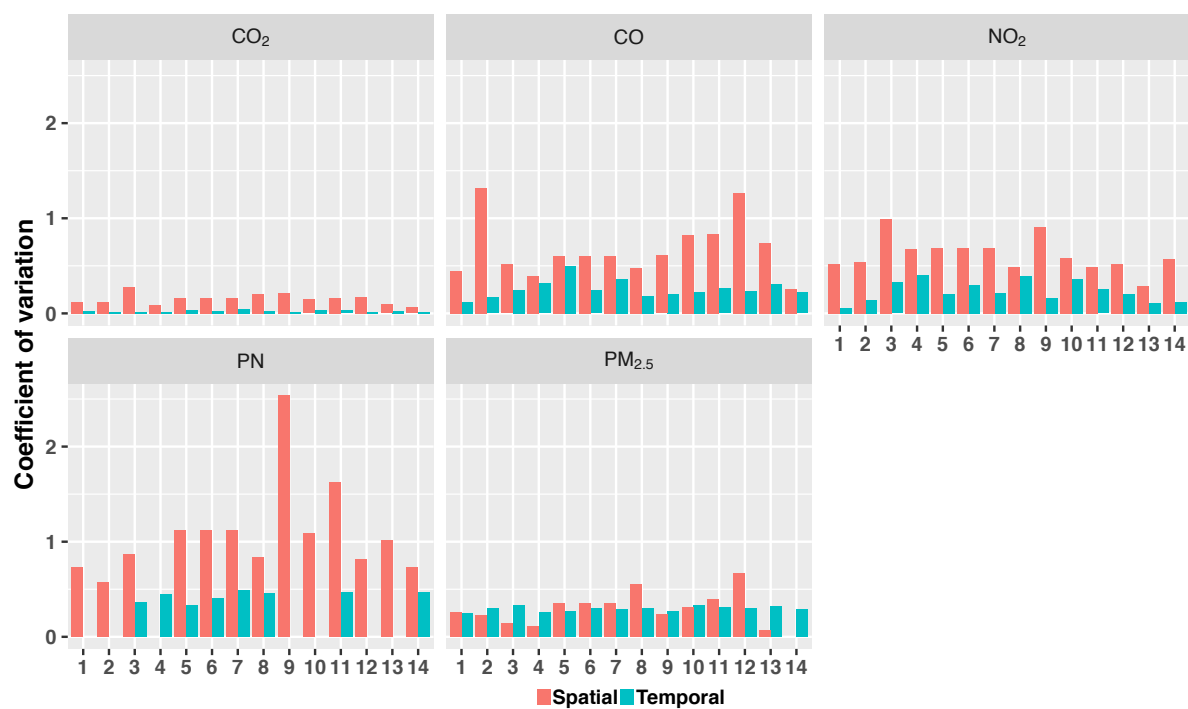




**Fig. 4.4.** 1 week (02/25/17 to 03/03/17) PN measurements from a restaurant emission influenced site 7 and background supersite 15.



**Fig. 4.5.** Temporal variations indicated by coefficient of variation (COV) in two scales (daily or hourly) for CO<sub>2</sub>, CO, NO<sub>2</sub>, PN, and PM<sub>2.5</sub>. A COV value of 0.5 is drawn to show difference between pollutants. Median COV values are labelled on top left.



**Fig. 4.6.** Coefficient of variation for spatial and temporal variations.

**Table 4.1.** Overview of sampling network: site strata, operating month period. Inside each of three strata columns--traffic, restaurant and building height, 'H' means high density with threshold defined in Methods section. Site ID is defined based on geographic locations from southwest to northeast, consistent with predominant wind direction.

Site ID	Campaign	Site name	Traffic	Restaurant	Building height	Site Type	Distributed monitor operating dates	Mobile driving dates
1	Transect	Chartiers				Upwind Background		
2	Transect	Carnegie				Upwind Background	08/2016 - 09/2016	08/2016 - 12/2016
3	Transect	Beechview				Upwind Background		
4	Downtown	Mt Washington				Upwind Background		
5	Downtown	Penn	H	H	H	Downtown		
6	Downtown	Mellon	H	H	H	Downtown	01/2017 - 03/2017	01/2017 - 02/2017
7	Downtown	Church	H	H	H	Downtown		
11	Downtown	Southside	H	H	H	Restaurants and traffic		
8	Transect/Downtown	Hill				Urban Background	08/2016 - 03/2017	08/2016 - 12/2016 01/2017 - 02/2017
9	Transect	Strip		H	H	Restaurants		
10	Transect	ACHD	H	H	H	City center		
12	Transect	Zoo		H		Restaurant and green space mixed	08/2016 - 09/2016	08/2016 - 12/2016
13	Transect	Aspinwall	H			Highway surrounded		
14	Transect	Fox Chapel				Downwind Background		
15	Transect/Downtown	Supersite				Urban Background	08/2016 - 12/2017	

#### 4.4 Implication for land use regression modeling

Land use regression is a multilinear regression. The dependent variable is pollutant concentration, usually averages of all measurements. The independent variables are GIS covariates, such as traffic, land use, and population density. Some of these predictors are temporally invariant. They only inform differences between sites at multiple locations. Insufficient temporal correction to mobile sampling dataset (both spatial and temporal variation) can result in confusing temporal variability as spatial gradient. This mis-assignment is accounted as measurement error. For epidemiological studies of environmental exposure, this random error would not result in biased regression coefficient, but indeed derive models with less precision and power (Armstrong, 1998). We deployed multiple stationary sites across the county, and did signal processing to raw measurements at all available monitors to calculate the regional background. Our results showed spatial variation was generally higher than temporal variation. This supported the applicability of long term exposure modeling based on short term mobile sampling. LUR proves the right way to go. Capturing spatial variations that are imposed on top of the regional background is important. However, day to day temporal variations are bigger than within day temporal variations for pollutants such as CO and NO<sub>2</sub>. This means that, in order to get a robust annual average concentration at a specific location, we need to sample enough days or hours to characterize the day-to-day variability to get a robust estimate of the annual average. Since the spatial variation is larger than the temporal variation and that spatial variation can often be associated with sources, for example, CO spatial variations are expected to be strongly tied to vehicle density, this means that the spatial pattern should reveal itself with mobile sampling well before we get to a point of having a stable annual average. So even if we don't have enough mobile sampling to get a stable annual average, we should be able to pretty quickly

resolve high versus low concentration sites or ranks sites. This idea is shown in our group's previous paper (Tan et al., 2014b).

As for building LUR models for specific pollutants, different sampling strategy is suggested based on findings in this paper. For  $PM_{2.5}$ , Figure 4.5 shows temporal pattern of  $PM_{2.5}$  is quite consistent across different sites unless a massive local source is in the vicinity like site 7. Figure 4.6 shows spatial variation of  $PM_1$  is only slightly higher than temporal variability. Thus, distributed monitors, with one or two near major sources, probably does a good job of estimating  $PM_{2.5}$  exposures within a given city.

For the other pollutants, spatial variation is much more important to the overall picture, so a network of monitors or trying to mix stationary sites with mobile sampling add more value. UFP (or PN) is the toughest – it has the largest temporal and spatial variations (therefore presumably the largest data needs to get representative averages). This matches with LUR results that traditionally do badly for UFP (Cattani et al., 2017; Farrell et al., 2016; Kerckhoffs et al., 2017; Saraswat et al., 2013). However, since UFP seems to be location-specific, the conventional UFP sampling strategy of going to a location and measuring for a short period of time (~10 minutes) could be sufficient with careful selection of sites.

## 4.5 Reference

ACHD - Air Quality Reports, 2016.

- Alghamdi, M.A., Khoder, M., Harrison, R.M., Hyvärinen, A.-P., Hussein, T., Al-Jeelani, H., Abdelmaksoud, A.S., Goknil, M.H., Shabbaj, I.I., Almehmadi, F.M., Lihavainen, H., Kulmala, M., Hämeri, K., 2014. Temporal variations of O<sub>3</sub> and NO<sub>x</sub> in the urban background atmosphere of the coastal city Jeddah, Saudi Arabia. *Atmos. Environ.* 94, 205–214. <https://doi.org/10.1016/j.atmosenv.2014.03.029>
- Allen, R.W., Amram, O., Wheeler, A.J., Brauer, M., 2011. The transferability of NO and NO<sub>2</sub> land use regression models between cities and pollutants. *Atmos. Environ.* 45, 369–378. <https://doi.org/10.1016/j.atmosenv.2010.10.002>
- Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environ. Sci. Technol.* 51, 6999–7008.
- Armstrong, B.G., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup. Environ. Med.* 55, 651–656. <https://doi.org/10.1136/oem.55.10.651>
- Bechle, M.J., Millet, D.B., Marshall, J.D., 2015. National Spatiotemporal Exposure Surface for NO<sub>2</sub>: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010. *Environ. Sci. Technol.* 49, 12297–12305. <https://doi.org/10.1021/acs.est.5b02882>
- Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A.S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U.,

- Pedersen, N.L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., Hoek, G., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet* 383, 785–795. [https://doi.org/10.1016/S0140-6736\(13\)62158-3](https://doi.org/10.1016/S0140-6736(13)62158-3)
- Brantley, H.L., Hagler, G.S.W., Kimbrough, E.S., Williams, R.W., Mukerjee, S., Neas, L.M., 2014. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. *Atmospheric Meas. Tech.* 7, 2169–2183. <https://doi.org/10.5194/amt-7-2169-2014>
- Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L.T., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* 50, 79–88. <https://doi.org/10.1021/acs.est.5b03709>
- Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V.,



- Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. *Circulation* 121, 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbece1>
- Buczyńska, A.J., Krata, A., Van Grieken, R., Brown, A., Polezer, G., De Wael, K., Potgieter-Vermaak, S., 2014. Composition of PM<sub>2.5</sub> and PM<sub>1</sub> on high and low pollution event days and its relation to indoor air quality in a home for the elderly. *Sci. Total Environ.* 490, 134–143. <https://doi.org/10.1016/j.scitotenv.2014.04.102>
- Canagaratna, M.R., Onasch, T.B., Wood, E.C., Herndon, S.C., Jayne, J.T., Cross, E.S., Miake-Lye, R.C., Kolb, C.E., Worsnop, D.R., 2010. Evolution of Vehicle Exhaust Particles in the Atmosphere. *J. Air Waste Manag. Assoc.* 60, 1192–1203. <https://doi.org/10.3155/1047-3289.60.10.1192>
- Cattani, G., Gaeta, A., Di Menno di Bucchianico, A., De Santis, A., Gaddi, R., Cusano, M., Ancona, C., Badaloni, C., Forastiere, F., Gariazzo, C., Sozzi, R., Inglessis, M., Silibello, C., Salvatori, E., Manes, F., Cesaroni, G., 2017. Development of land-use regression models for exposure assessment to ultrafine particles in Rome, Italy. *Atmos. Environ.* 156, 52–60. <https://doi.org/10.1016/j.atmosenv.2017.02.028>
- Cicchetti, D.V., 1994. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychol. Assess.* 6, 284–90.
- Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos,

T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 389, 1907–1918. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6)

Data | Information Portal | Allegheny County [WWW Document], n.d. URL <http://infoportal.alleghenycounty.us/data.html> (accessed 11.13.17).

de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., Dèdelè, A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gražulevičienė, R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsouyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Mosler, G., Nádor, G., Nieuwenhuijsen, M., Pershagen, G., Peters, A., Phuleria, H., Probst-Hensch, N., Raaschou-Nielsen, O., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Weinmayr, G., Brunekreef, B., Hoek, G., 2013. Development of Land Use Regression Models for Particle Composition in Twenty Study Areas in Europe. *Environ. Sci. Technol.* 47, 5778–5786. <https://doi.org/10.1021/es400156t>

DeCarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-Deployable, High-Resolution, Time-of-Flight Aerosol Mass Spectrometer. *Anal. Chem.* 78, 8281–8289. <https://doi.org/10.1021/ac061249n>

Department of City Planning [WWW Document], 2017. URL <http://pittsburghpa.gov/dcp> (accessed 11.13.17).

Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017a. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults.

- JAMA 318, 2446–2456. <https://doi.org/10.1001/jama.2017.17923>
- Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.D., 2017b. Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* 376, 2513–2522. <https://doi.org/10.1056/NEJMoal702747>
- Donahue, N.M., Posner, L.N., Westervelt, D.M., Li, Z., Shrivastava, M., Presto, A.A., Sullivan, R.C., Adams, P.J., Pandis, S.N., Robinson, A.L., 2016. Where Did This Particle Come From? Sources of Particle Number and Mass for Human Exposure Estimates, in: *Airborne Particulate Matter*. Royal Society of Chemistry, Cambridge, pp. 35–71.
- Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dèdelè, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičienė, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S. von, Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of Land Use Regression Models for PM<sub>2.5</sub>, PM<sub>2.5</sub> Absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European Study Areas; Results of the ESCAPE Project. *Environ. Sci. Technol.* 46, 11195–11205. <https://doi.org/10.1021/es301948k>
- Eeftens, M., Phuleria, H.C., Meier, R., Aguilera, I., Corradi, E., Davey, M., Ducret-Stich, R., Fierz, M., Gehrig, R., Ineichen, A., Keidel, D., Probst-Hensch, N., Ragettli, M.S., Schindler, C., Künzli, N., Tsai, M.-Y., 2015. Spatial and temporal variability of ultrafine particles, NO<sub>2</sub>, PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in Swiss

- study areas. *Atmos. Environ.* 111, 60–70. <https://doi.org/10.1016/j.atmosenv.2015.03.031>
- Elser, M., Bozzetti, C., El-Haddad, I., Maasikmets, M., Teinemaa, E., Richter, R., Wolf, R., Slowik, J.G., Baltensperger, U., Prévôt, A.S.H., 2016. Urban increments of gaseous and aerosol pollutants and their sources using mobile aerosol mass spectrometry measurements. *Atmos Chem Phys* 16, 7117–7134. <https://doi.org/10.5194/acp-16-7117-2016>
- Farrell, W., Weichenthal, S., Goldberg, M., Valois, M.-F., Shekarizfard, M., Hatzopoulou, M., 2016. Near roadway air pollution across a spatially extensive road and cycling network. *Environ. Pollut.* 212, 498–507. <https://doi.org/10.1016/j.envpol.2016.02.041>
- Geography, U.C.B., n.d. TIGER Products [WWW Document]. URL <https://www.census.gov/geo/maps-data/data/tiger.html> (accessed 11.13.17).
- Gerard, H., Rob, B., Kees, de H., Danielle, V., John, G., Paul, F., David, B., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* 42, 7561–7578.
- Hankey, S., Marshall, J.D., 2015. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM<sub>2.5</sub>, Particle Size) Using Mobile Monitoring. *Environ. Sci. Technol.* 49, 9194–9202.
- Hatzopoulou, M., Valois, M.F., Levy, I., Mihele, C., Lu, G., Bagg, S., Minet, L., Brook, J., 2017. Robustness of Land-Use Regression Models Developed from Mobile Air Pollutant Measurements. *Environ. Sci. Technol.* 51, 3938–3947. <https://doi.org/10.1021/acs.est.7b00366>
- HEI Review Panel on Ultrafine Particles, 2013. Understanding the Health Effects of Ambient Ultrafine Particles, HEI Perspectives 3. Health Effects Institute, Boston, MA.

- Hering, S.V., Lewis, G.S., Spielman, S.R., Eiguren-Fernandez, A., Kreisberg, N.M., Kuang, C., Attoui, M., 2017. Detection near 1-nm with a laminar-flow, water-based condensation particle counter. *Aerosol Sci. Technol.* 51, 354–362. <https://doi.org/10.1080/02786826.2016.1262531>
- Home | City of Pittsburgh GIS Data [WWW Document], 2015. URL <http://pghgis-pittsburghpa.opendata.arcgis.com/> (accessed 11.13.17).
- Jayne, J.T., Leard, D.C., Zhang, X., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an Aerosol Mass Spectrometer for Size and Composition Analysis of Submicron Particles. *Aerosol Sci. Technol.* 33, 49–70. <https://doi.org/10.1080/027868200410840>
- Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. *Environ. Sci. Technol.* 44, 5334–5344. <https://doi.org/10.1021/es100008x>
- Keller, J.P., Olives, C., Kim, S.-Y., Sheppard, L., Sampson, P.D., Szpiro, A.A., Oron, A.P., Lindstrom, J., Vedal, S., Kaufman, J.D., 2015. A Unified Spatiotemporal Modeling Approach for Predicting Concentrations of Multiple Air Pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. *Environ. Health Perspect.* 123, 301–309. <https://doi.org/10.1289/ehp.1408145>
- Kerckhoffs, J., Hoek, G., Messier, K.P., Brunekreef, B., Meliefste, K., Klompmaker, J.O., Vermeulen, R., 2016. Comparison of Ultrafine Particle and Black Carbon Concentration Predictions from a Mobile and Short-Term Stationary Land-Use Regression Model. *Environ. Sci. Technol.* 50, 12894–12902. <https://doi.org/10.1021/acs.est.6b03476>
- Kerckhoffs, J., Hoek, G., Vlaanderen, J., van Nunen, E., Messier, K., Brunekreef, B.,

- Gulliver, J., Vermeulen, R., 2017. Robustness of intra urban land-use regression models for ultrafine particles and black carbon based on mobile monitoring. *Environ. Res.* 159, 500–508. <https://doi.org/10.1016/j.envres.2017.08.040>
- Kimmel, J.R., Farmer, D.K., Cubison, M.J., Sueper, D., Tanner, C., Nemitz, E., Worsnop, D.R., Gonin, M., Jimenez, J.L., 2011. Real-time aerosol mass spectrometry with millisecond resolution. *Int. J. Mass Spectrom.* 303, 15–26. <https://doi.org/10.1016/j.ijms.2010.12.004>
- Klems, J.P., Pennington, M.R., Zordan, C.A., Johnston, M.V., 2010. Ultrafine Particles Near a Roadway Intersection: Origin and Apportionment of Fast Changes in Concentration. *Environ. Sci. Technol.* 44, 7903–7907. <https://doi.org/10.1021/es102009e>
- Klompniaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. *Sci. Total Environ.* 508, 266–275. <https://doi.org/10.1016/j.scitotenv.2014.11.088>
- Li, H.Z., Dallmann, T.R., Gu, P., Presto, A.A., 2016. Application of mobile sampling to investigate spatial variation in fine particle composition. *Atmos. Environ.* 142, 71–82. <https://doi.org/10.1016/j.atmosenv.2016.07.042>
- Li, H.Z., Dallmann, T.R., Li, X., Gu, P., Presto, A.A., 2017. Urban Organic Aerosol Exposure: Spatial Variations in Composition and Source Impacts. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.7b03674>
- Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M.A., Amann, M., Anderson, H.R., Andrews, K.G., Aryee, M., Atkinson, C., Bacchus, L.J., Bahalim, A.N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M.L., Blore, J.D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M.,

Brauer, M., Brooks, P., Bruce, N.G., Brunekreef, B., Bryan-Hancock, C., Bucello, C.,  
 Buchbinder, R., Bull, F., Burnett, R.T., Byers, T.E., Calabria, B., Carapetis, J.,  
 Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J.S., Cheng, A.T.-A., Child,  
 J.C., Cohen, A., Colson, K.E., Cowie, B.C., Darby, S., Darling, S., Davis, A.,  
 Degenhardt, L., Dentener, F., Jarlais, D.C.D., Devries, K., Dherani, M., Ding, E.L.,  
 Dorsey, E.R., Driscoll, T., Edmond, K., Ali, S.E., Engell, R.E., Erwin, P.J., Fahimi, S.,  
 Falder, G., Farzadfar, F., Ferrari, A., Finucane, M.M., Flaxman, S., Fowkes, F.G.R.,  
 Freedman, G., Freeman, M.K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G.,  
 Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H.R., Hall, W., Hoek,  
 H.W., Hogan, A., Hosgood, H.D., Hoy, D., Hu, H., Hubbell, B.J., Hutchings, S.J.,  
 Ibeanusi, S.E., Jacklyn, G.L., Jasrasaria, R., Jonas, J.B., Kan, H., Kanis, J.A.,  
 Kassebaum, N., Kawakami, N., Khang, Y.-H., Khatibzadeh, S., Khoo, J.-P., Kok, C.,  
 Laden, F., Lalloo, R., Lan, Q., Lathlean, T., Leasher, J.L., Leigh, J., Li, Y., Lin, J.K.,  
 Lipshultz, S.E., London, S., Lozano, R., Lu, Y., Mak, J., Malekzadeh, R., Mallinger,  
 L., Marcenes, W., March, L., Marks, R., Martin, R., McGale, P., McGrath, J., Mehta,  
 S., Memish, Z.A., Mensah, G.A., Merriman, T.R., Micha, R., Michaud, C., Mishra, V.,  
 Hanafiah, K.M., Mokdad, A.A., Morawska, L., Mozaffarian, D., Murphy, T., Naghavi,  
 M., Neal, B., Nelson, P.K., Nolla, J.M., Norman, R., Olives, C., Omer, S.B., Orchard,  
 J., Osborne, R., Ostro, B., Page, A., Pandey, K.D., Parry, C.D., Passmore, E., Patra, J.,  
 Pearce, N., Pelizzari, P.M., Petzold, M., Phillips, M.R., Pope, D., Pope, C.A., Powles,  
 J., Rao, M., Razavi, H., Rehfuess, E.A., Rehm, J.T., Ritz, B., Rivara, F.P., Roberts, T.,  
 Robinson, C., Rodriguez-Portales, J.A., Romieu, I., Room, R., Rosenfeld, L.C., Roy,  
 A., Rushton, L., Salomon, J.A., Sampson, U., Sanchez-Riera, L., Sanman, E.,  
 Sapkota, A., Seedat, S., Shi, P., Shield, K., Shivakoti, R., Singh, G.M., Sleet, D.A.,

- Smith, E., Smith, K.R., Stapelberg, N.J., Steenland, K., Stöckl, H., Stovner, L.J., Straif, K., Straney, L., Thurston, G.D., Tran, J.H., Dingenen, R.V., Donkelaar, A. van, Veerman, J.L., Vijayakumar, L., Weintraub, R., Weissman, M.M., White, R.A., Whiteford, H., Wiersma, S.T., Wilkinson, J.D., Williams, H.C., Williams, W., Wilson, N., Woolf, A.D., Yip, P., Zielinski, J.M., Lopez, A.D., Murray, C.J., Ezzati, M., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 380, 2224–2260. [https://doi.org/10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)
- Liu, L., Urch, B., Poon, R., Szyszkowicz, M., Speck, M., Gold, D.R., Wheeler, A.J., Scott, J.A., Brook, J.R., Thorne, P.S., Silverman, F.S., 2015. Effects of Ambient Coarse, Fine, and Ultrafine Particles and Their Biological Constituents on Systemic Biomarkers: A Controlled Human Exposure Study. *Environ. Health Perspect.* 123, 534–540. <https://doi.org/10.1289/ehp.1408387>
- Masey, N., Gillespie, J., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Influence of wind-speed on short-duration NO<sub>2</sub> measurements using Palmes and Ogawa passive diffusion samplers. *Atmos. Environ.* 160, 70–76. <https://doi.org/10.1016/j.atmosenv.2017.04.008>
- Masiol, M., Squizzato, S., Formenton, G., Harrison, R.M., Agostinelli, C., 2017. Air quality across a European hotspot: Spatial gradients, seasonality, diurnal cycles and trends in the Veneto region, NE Italy. *Sci. Total Environ.* 576, 210–224. <https://doi.org/10.1016/j.scitotenv.2016.10.042>
- Matte, T.D., Ross, Z., Kheirbek, I., Eisl, H., Johnson, S., Gorczynski, J.E., Kass, D., Markowitz, S., Pezeshki, G., Clougherty, J.E., 2013. Monitoring intraurban spatial



- patterns of multiple combustion air pollutants in New York City: Design and implementation. *J. Expo. Sci. Environ. Epidemiol.* 23, 223–231. <https://doi.org/10.1038/jes.2012.126>
- Meier, R., Eeftens, M., Aguilera, I., Phuleria, H.C., Ineichen, A., Davey, M., Ragettli, M.S., Fierz, M., Schindler, C., Probst-Hensch, N., Tsai, M.-Y., Kuenzli, N., 2015. Ambient Ultrafine Particle Levels at Residential and Reference Sites in Urban and Rural Switzerland. *Environ. Sci. Technol.* 49, 2709–2715. <https://doi.org/10.1021/es505246m>
- Ostro, B., Hu, J., Goldberg, D., Reynolds, P., Hertz, A., Bernstein, L., Kleeman, M.J., 2015. Associations of Mortality with Long-Term Exposures to Fine and Ultrafine Particles, Species and Sources: Results from the California Teachers Study Cohort. *Environ. Health Perspect.* 123, 549–556. <https://doi.org/10.1289/ehp.1408565>
- Patton, A.P., Zamore, W., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2015. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. *Environ. Sci. Technol.* 49, 6051–6060. <https://doi.org/10.1021/es5061676>
- Petters, M.D., Kreidenweis, S.M., 2007. A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. *Atmos Chem Phys* 7, 1961–1971. <https://doi.org/10.5194/acp-7-1961-2007>
- Piersanti, A., Vitali, L., Righini, G., Cremona, G., Ciancarella, L., 2015. Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollut. Res.* 6, 953–960. <https://doi.org/10.1016/j.apr.2015.04.005>
- Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris Jr., B.G., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities 329, 1753–

1759.

Pope, C.A.I., Ezzati, M., Dockery, D.W., 2009. Fine-Particulate Air Pollution and Life

Expectancy in the United States. *N. Engl. J. Med.* 360, 376–386.

Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G.,

Hoffmann, B., Fischer, P., Nieuwenhuijsen, M.J., Brunekreef, B., Xun, W.W.,

Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A.,

Oftedal, B., Schwarze, P.E., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.-

G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Sørensen, M.,

Tjønneland, A., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M.,

Bueno-de-Mesquita, B., Key, T.J., de Hoogh, K., Concini, H., Nagel, G., Vilier, A.,

Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E.,

Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P.,

Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., Hoek, G., 2013. Air

pollution and lung cancer incidence in 17 European cohorts: prospective analyses

from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet*

*Oncol.* 14, 813–822. [https://doi.org/10.1016/S1470-2045\(13\)70279-1](https://doi.org/10.1016/S1470-2045(13)70279-1)

Righini, G., Cappelletti, A., Ciucci, A., Cremona, G., Piersanti, A., Vitali, L., Ciancarella, L.,

2014. GIS based assessment of the spatial representativeness of air quality monitoring

stations using pollutant emissions data. *Atmos. Environ.* 97, 121–129. [https://doi.org/](https://doi.org/10.1016/j.atmosenv.2014.08.015)

[10.1016/j.atmosenv.2014.08.015](https://doi.org/10.1016/j.atmosenv.2014.08.015)

Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop,

A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking Organic Aerosols:

Semivolatile Emissions and Photochemical Aging. *Science* 315, 1259–1262. [https://](https://doi.org/10.1126/science.1133061)

[doi.org/10.1126/science.1133061](https://doi.org/10.1126/science.1133061)

Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013.

Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon

Particulate Matter in New Delhi, India. *Environ. Sci. Technol.* 47, 12903–12911.

<https://doi.org/10.1021/es401489h>

Seinfeld, J.H., Pandis, S.N., 2016. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons.

Simon, M.C., Hudda, N., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2017.

Comparisons of traffic-related ultrafine particle number concentrations measured in

two urban areas by central, residential, and mobile monitoring. *Atmos. Environ.* 169,

113–127. <https://doi.org/10.1016/j.atmosenv.2017.09.003>

Stafoggia, M., Schneider, A., Cyrys, J., Samoli, E., Andersen, Z.J., Bedada, G.B., Bellander,

T., Cattani, G., Eleftheriadis, K., Faustini, A., Hoffmann, B., Jacquemin, B.,

Katsouyanni, K., Massling, A., Pekkanen, J., Perez, N., Peters, A., Quass, U., Yli-

Tuomi, T., Forastiere, F., Group, on behalf of the U.S., 2017. Association Between

Short-term Exposure to Ultrafine Particles and Mortality in Eight European Urban

Areas. *Epidemiology* 28, 172–180. <https://doi.org/10.1097/EDE.0000000000000599>

Sullivan, R.C., Pryor, S.C., 2014. Quantifying spatiotemporal variability of fine particles in

an urban environment using combined fixed and mobile measurements. *Atmos.*

*Environ.* 89, 664–671.

Sullivan, R.C., Pryor, S.C., 2014. Quantifying spatiotemporal variability of fine particles in

an urban environment using combined fixed and mobile measurements. *Atmos.*

*Environ.* 89, 664–671. <https://doi.org/10.1016/j.atmosenv.2014.03.007>

Tan, Y., Dallmann, T.R., Robinson, A.L., Presto, A.A., 2016. Application of plume analysis to

build land use regression models from mobile sampling to improve model

- transferability. *Atmos. Environ.* 134, 51–60. <https://doi.org/10.1016/j.atmosenv.2016.03.032>
- Tan, Y., Lipsky, E.M., Saleh, R., Robinson, A.L., Presto, A.A., 2014a. Characterizing the Spatial Variation of Air Pollutants and the Contributions of High Emitting Vehicles in Pittsburgh, PA. *Environ. Sci. Technol.* 48, 14186–14194. <https://doi.org/10.1021/es5034074>
- Tan, Y., Robinson, A.L., Presto, A.A., 2014b. Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. *Atmos. Environ.* 99, 333–340. <https://doi.org/10.1016/j.atmosenv.2014.10.003>
- US EPA, O., 2016. 2014 National Emissions Inventory (NEI) Data [WWW Document]. US EPA. URL <https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data> (accessed 11.13.17).
- Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., 2015. Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmos. Environ.* 105, 148–161. <https://doi.org/10.1016/j.atmosenv.2015.01.017>
- Vienneau, D., de Hoogh, K., Beelen, R., Fischer, P., Hoek, G., Briggs, D., 2010. Comparison of land-use regression models between Great Britain and the Netherlands. *Atmos. Environ.* 44, 688–696. <https://doi.org/10.1016/j.atmosenv.2009.11.016>
- Vitali, L., Morabito, A., Adani, M., Assennato, G., Ciancarella, L., Cremona, G., Giua, R., Pastore, T., Piersanti, A., Righini, G., Russo, F., Spagnolo, S., Tanzarella, A., Tinarelli, G., Zanini, G., 2016. A Lagrangian modelling approach to assess the representativeness area of an industrial air quality monitoring station. *Atmospheric Pollut. Res.* 7, 990–1003. <https://doi.org/10.1016/j.apr.2016.06.002>

- Wang, M., Beelen, R., Basagana, X., Becker, T., Cesaroni, G., de Hoogh, K., Dedele, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H., Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M.J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., Brunekreef, B., 2013. Evaluation of land use regression models for NO<sub>2</sub> and particulate matter in 20 European study areas: the ESCAPE project. *Environ. Sci. Technol.* 47, 4357–4364. <https://doi.org/10.1021/es305129t>
- Yu, C.H., Fan, Z., Li, P.J., Baptista, A., Greenberg, M., Laumbach, R.J., 2016. A novel mobile monitoring approach to characterize spatial and temporal variation in traffic-related air pollutants in an urban community. *Atmos. Environ.* 141, 161–173. <https://doi.org/10.1016/j.atmosenv.2016.06.044>
- Zimmerman, N., Jeong, C.-H., Wang, J.M., Ramos, M., Wallace, J.S., Evans, G.J., 2015. A source-independent empirical correction procedure for the fast mobility and engine exhaust particle sizers. *Atmos. Environ.* 100, 178–184. <https://doi.org/10.1016/j.atmosenv.2014.10.054>
- Zimmerman, N., Presto, A.A., Kumar, S.P.N., Gu, J., Hauryliuk, A., Robinson, E.S., Robinson, A.L., Subramanian, R., 2017. Closing the gap on lower cost air quality monitoring: machine learning calibration models to improve low-cost sensor performance. *Atmos Meas Tech Discuss* 2017, 1–36. <https://doi.org/10.5194/amt-2017-260>

**Chapter 5: Spatial variability of air pollution near monitors and exposure misclassification in an eastern US city**

## **Chapter 5**

### **Spatial variability of air pollution near monitors and exposure misclassification in an eastern US city**

## 5.1 Introduction

Some studies pointed out the limited spatial representativeness of stationary monitors (Apte et al., 2017; Piersanti et al., 2015; Shi et al., 2016; Vardoulakis et al., 2005; Vitali et al., 2016). Vitali et al. (2016) used dispersion modeling and concluded an industrial air quality monitoring station was representative of a surrounding area of 0.07 km<sup>2</sup> (roughly 250 m by 250 m) based on relative difference between monitor measurements and predicted concentrations in locations nearby. Shi et al. (2016) used geostatistical methods to find an ideal pixel size for spatially aggregating mobile PM<sub>2.5</sub> measurements in Hong Kong. They indicated 300 m resolution-based aggregation preserved the short-range features of intracity street-level air pollution variation. Monitors tended to represent a larger spatial extent when compared in the context of long term exposure than short term exposure, and monitors in rural or suburbs could be more spatially representative than urban monitors (Piersanti et al., 2015). Vardoulakis et al. (2005) conducted mobile sampling in the vicinity of a monitoring station located near busy intersections in central Paris, and observed the monitoring station was consistently recording higher air pollutant concentrations compared to mobile measurements one block away. Given the siting of the monitor and the pronounced spatial variability near the monitor, they concluded the monitor was inadequate to indicate the overall air pollution picture in the study area, and thus not a good reference for population exposure estimate. Recently, Apte et al. (2017) deployed instruments on a Google StreetView car and conducted a yearlong extensive sampling in Oakland, CA. They found substantial spatial variation in street level and neighborhood scale (~ km) for nitrogen oxides and black carbon.

Studying the spatial representativeness of monitors can help design improved monitoring networks, maximizing spatial coverage and avoiding redundant sites (Martin et



al., 2014). However, the aforementioned studies either studied only one or two pollutants (Lightowlers et al., 2008; Piersanti et al., 2015; Vitali et al., 2016), or they used model prediction at one location instead of real instrument measurements to compare with the reference monitor measurement (Martin et al., 2014).

In this study, we conducted a campaign with both mobile sampling and distributed monitors. The goals are to study: 1) spatial variability of multiple pollutants including CO, NO<sub>2</sub>, UFP, and PM<sub>1</sub> near stationary monitors, 2) short-term exposure misclassification of using point measurement to represent the whole surrounding neighborhoods, 3) spatial representativeness of monitors in different microenvironments.

## **5.2 Material and Methods**

The sampling domain is Allegheny County, PA. The city of Pittsburgh is in the center of the county. The county's landscape is characterized by three river valleys and a vast plateau. The air quality is influenced by the interaction of regional pollutants transported from power plants and other industrial emissions in the Ohio River valley and local industrial and traffic sources (ACHD - Air Quality Reports, 2016). Diverse point sources are distributed in the whole county, including the largest eastern U.S. coke plant, steel manufacturing, power plants, and specialty steel production.

### **5.2.1 Sampling setup**

The sampling network is composed of a mobile sampling van and fifteen distributed monitors. Mobile sampling is conducted around each monitor repeatedly across different times of day in 2016 summer/fall and winter.

The mobile van is a gasoline powered high roof van. It is described in detail in Li et al and Tan et al (Li et al., 2016; Tan et al., 2014a). Power is drawn from the van engine via a converter and then dispatched to different gas or particle instruments. Two half-inch O.D.

stainless sampling lines are installed on top of the van. The sampling inlet is about 4 m above the ground, with an extended stainless bowl cover to keep out rain or snow. A Bad Elf GPS logger (Bad Elf, CT, USA) records the mobile sampling location every 1 s.

One sampling line is for particle measurements including  $PM_{10}$  composition measured by a Time of Flight Aerosol Mass Spectrometer (AMS) (Aerodyne Research Inc., Boston, MA) (DeCarlo et al., 2006; Jayne et al., 2000), black carbon (BC) by Aethalometer (Magee Scientific, CA, USA), and particle number (PN) from Fast Mobility Particle Sizer (FMPS, size range 5.6 to 560 nm) (TSI Incorporated, MN, USA). Before the air flow reaches these particle instruments, it is size selected by a cyclone with flow rate controlled at 16.7 SLPM. The cutoff size is 2.5  $\mu m$ . The sum of particle composition from the AMS (organic and inorganic) plus the BC from Aethalometer yields the  $PM_{10}$ . AMS is operated at fast mass spectra mode and outputs data in 20 s resolution (Kimmel et al., 2011). BC data are in 1-min resolution. PN is reported every 1 s.

The other sampling line is for gas instruments, including a  $NO_x$  analyzer T200 and CO analyzer T300 (Teledyne Technology, CA, USA). A HEPA filter is placed behind the top sampling inlet to remove particles inside air flow before the air reaches the gas instruments. All gas instruments on the mobile van report data every 1 s.

15 distributed monitors are all equipped with a low-cost sensor package—the real-time affordable multipollutant sensor (RAMP) (SenSevere LLC, PA, USA). The RAMP is described in detail in Zimmermann et al (2017). Briefly, the core component of the RAMP is Alphasense electrochemical sensors (Alphasense Ltd., UK) for measuring CO and  $NO_2$ . Voltage difference during the reaction of incoming gas pollutants with reagents is converted to ambient concentration using a previously developed machine learning calibration (Zimmerman et al, 2017). The RAMPs report  $NO_2$  and CO every 15 minutes.

The distributed monitors are also equipped with MetOne Neighborhood PM monitors (MetOne Instruments, Inc., OR, USA) for PM<sub>2.5</sub> measurement. Some of the distributed monitors had a water based condensation particle counter (Aerosol Dynamics Inc., Berkeley, CA) for PN concentration. To ensure consistency, PM<sub>2.5</sub> and PN measurements are reported along with gas pollutants NO<sub>2</sub> and CO at 15-min resolution.

Routine calibration and maintenance are performed for both sampling platforms. On the mobile platform, the AMS ionization efficiency calibration was performed every time the AMS was unloaded and reloaded from the van. The AMS collects one HEPA filtered blank sample after each sampling trip. The FMPS is zero cleaned via HEPA filter air before each sampling trip. Calibration of NO<sub>x</sub> and CO analyzers includes zero and span check with reference gases every week. On the distributed monitors, the RAMP output is calibrated with machine learning random forest model based approach (Zimmerman et al., 2017). MetOne PM<sub>2.5</sub> output is first corrected based on hygroscopic growth factor (Petters and Kreidenweis, 2007), and then calibrated against reference PM monitors. The raw output from the distributed CPCs was adjusted based on co-location tests with a butanol CPC (TSI Incorporated, MN, USA) in the lab. All instruments were synchronized, and reported data in eastern standard time.

### **5.2.2 Sampling overview**

The whole campaign is comprised of two case studies, and 15 distributed sites are selected based on different land use and source characteristics. Sites in the transect case study span much larger spatial extent (~ 50 km) compared to downtown sites (within ~5 km) (Table 5.1; Figure D1, Supplemental Materials). The transect case study was conducted in 2016 summer/fall, and the downtown study was in 2016 winter. Sites are stratified based on traffic density, restaurants, and building height (Allegheny County Information Portal, 2017;

Department of City Planning, 2017; City of Pittsburgh GIS Data, 2015; US Census Tiger, 2010). Traffic density criteria includes road length, number of traffic lights, and vehicles miles travelled (VMT) in defined buffer areas (200 m by 200 m grid boxes). Building height to road width ratio is a common metric to describe pollution trapping by street canyon effects in downtown (Kwak et al., 2016). Here we simplified the methodology and used building height to indicate whether air flow could be influenced by tall buildings. A 200 m grid box is labelled as high building if the average building height is above 40 feet. Cooking emissions from restaurants are a major source of urban organic aerosol (Liu et al., 2017). Dining at restaurants is usually an important part of people daily life. The resulting pollutant exposure problems raise concern. A 200 m grid box is regarded as high restaurant density if the box has 3 or more restaurants inside.

Mobile sampling is conducted inside a predetermined 1 km<sup>2</sup> box around each monitor. For each sampling trip, we tried to cover all major roads inside the box at least once. Sampling time spent per site was around 1 hour for downtown sites, and 40 mins for suburban sites. Repeated mobile sampling at each site covered different times of day to capture both spatial and temporal variation. Each transect site was visited on at least 3 different days, and downtown sites each had more than 10 unique visits. The data difference was due to the different spatial range in two case studies.

### **5.2.3 Data treatment**

Various data wrangling processes were performed including data import and export, data spatial joining, outlier filtering, null value replacement, etc. 1 Hz mobile measurements were spatially allocated based on 50 or 100 m grid boxes following Apte et al (2017). GPS information was joined to concurrent pollutant measurements. For PM<sub>1</sub> reported at 20 s resolution, the mean of GPS latitude and longitude measurements was joined to the PM<sub>1</sub>

measurement in that 20-s time window. FMPS undersizes accumulation mode particles and is corrected based on Zimmerman et al (2015). FMPS occasionally reported unrealistic size distributions (no particles in middle size bins) when the van was driving on bumpy roads. Such data were filtered. Extreme concentrations were defined as values either larger than the 75th percentile plus 4 times the interquartile range or less than 25th percentile minus 4 times interquartile range. This outlier filtering only applied to distributed monitor data. Mobile sampling data are expected to show more episodic events and broader concentration range compared with monitor measurements (Tan et al., 2014a).

We used mobile measurements to study spatial patterns within the predetermined driving domains. For each 1 km<sup>2</sup> driving box, we assigned the 50 or 100 m grid with the most total measurements from all sampling trips as the reference grid. The reference grid would be fixed and not change from day to day. Then the raw mobile data were grouped by each day and site. For one mobile sampling trip on a specific day at one site, the median of all measurements inside the reference grid was calculated and treated as the reference value for that trip at the site. Measurements in other grids at that site were then subtracted from the reference value. This process was repeated for every sampling date, and all the subtracted concentrations were then averaged to represent each grid's relative difference to the reference grid.

## **5.3 Results and Discussion**

### **5.3.1 Spatial variability at each site**

Pollutants are known to exhibit significant intracity variation (Apte et al., 2017; Donahue et al., 2016; Li et al., 2016, 2017). Figure 5.1 shows NO<sub>2</sub> spatial variability mapped with 50 m spatial resolution inside each 1 km<sup>2</sup> mobile sampling box at different sites in our sampling domain. Enhancement is calculated by raw mobile measurement minus the

corresponding background derived from distributed monitors using wavelet decomposition (Klems et al., 2010). We first decomposed 15-minute resolution time series of distributed monitor measurements into different frequencies, selected components with longer than 8-hour frequency at each site, and then compare and select the smallest one as the true regional background for that 15-minute window. Urban background NO<sub>2</sub> in the study region is 7 ppb. Grey boxes are within 50% of the regional background, and yellow or red means a factor of 2-3 difference with respect to background. Each 50 m grid box in Figure 5.1 has at least 20 mobile measurements before averaging, so the spatial pattern tends to inform a long-term trend (Apte et al., 2017; Van den Bossche et al., 2015, 2016).

Upwind sites such as 1, 2, 3, 4 (Figure D1. Supplemental Materials) are essentially background with little or no NO<sub>2</sub> enhancement. Urban sites 5-12 show different degrees of spatial enhancement. Downtown sites 6, 7, 8 have the largest NO<sub>2</sub> enhancement due to combined influence of traffic, building height induced street canyon effects, and point sources. Site 13 has mixed residential area (south) and green space (north) inside, and NO<sub>2</sub> enhancement increases from south to north. Site 14 has one highway in the north, only 50 meters away from the sampling box, though NO<sub>2</sub> enhancement is barely caught. One reason might be due to the quick decay of traffic related pollutants to background within hundreds of meters from roadside (Karner et al., 2010). Another reason might be due to the southwestern predominant wind direction. Downwind background site 15 has little NO<sub>2</sub> enhancement as there are no obvious primary sources inside the domain. Red or purple grid boxes frequently are located near major road intersections such as at sites 6, 11 and 14. Traffic density has a significant impact on driving NO<sub>2</sub> spatial pattern in different sites.

Overall, Figure 5.1 indicates different spatial variability in different micro-environments. Urban sites tend to have large spatial heterogeneity, which means that a

reference monitor located in urban center might be less representative of residents living nearby. As less air pollutant variability is observed around suburban or background monitors, these monitors could represent a larger spatial extent of surrounding population exposure.

### 5.3.2 Temporal variability during each trip

To identify whether concentration difference in space is due to temporal variation or source impacts, Figure 5.2 shows cumulative distribution function (CDF) of sequential hourly pollutant concentration differences based on data from all continuously operating distributed monitors. Downtown has more episodic events, and histogram of site 6 is shown as an inset in top left of each panel as an example.

Figure 5.2(A) shows that 50% of the population of CO sequential differences lie between -13 and 13 ppb. 75% lie between -36 and 36 ppb. Since each mobile sampling visit to a 1 km<sup>2</sup> box takes about one hour, we interpret absolute CO differences relative to a chosen reference grid (like a stationary monitor) larger than 13 ppb to indicate real spatial variation instead of mere temporal fluctuation during the sampling drive. The histogram of CO differences at downtown shows less than 50% or 75% values lying between thresholds based on the whole population. This means, temporal variation at downtown seems to be bigger than background sites.

Urban background CO is around 154 ppb. 36 ppb (75% threshold) is about 23% of the background value. Unlike the log-normal distribution commonly seen for pollutants from mobile sampling (Tan et al., 2014b), the histogram of CO at downtown shows a Gaussian like distribution. This indicates different pattern of temporal and spatial variations.

For NO<sub>2</sub> in Figure 5.2(B), similarity, 50% of the population differences lie between -0.6 to 0.6 ppb. 75% values are between -1.5 and 1.5 ppb. Urban background NO<sub>2</sub> is around 7 ppb. 1.5 ppb is around 20% of this background value.

Urban background PN is around  $6257 \text{ \#/cm}^3$ . The 75% threshold of 4300 is around 70% of the background value. This indicates PN is highly temporally dynamic.

Urban background  $\text{PM}_{2.5}$  is  $9 \text{ \mu g/m}^3$ . The 75% threshold -- 1.1 is only 12% of the value.  $\text{PM}_{2.5}$  shows the least temporal variability among all four pollutants listed.

### 5.3.3 Spatial variability near urban monitors

In comparison to the magnitude of temporal variation during each sampling trip shown in Figure 5.2, Figure 5.3 shows the spatial variability inside the downtown driving box. Each 50 m grid box is colored by the concentration difference relative to the reference grid (black dot). Grey legend values correspond to differences smaller than the 50% temporal threshold for each pollutant determined in Figure 5.2. Red or blue grids therefore indicate spatial differences larger than typical hourly temporal variations. Hot colors show that a particular grid box has larger concentrations than the reference grid, and cool colors show boxes with concentrations smaller than the reference grid. Similar to Figure 5.2, the histogram of the spatial difference is shown as an insert on top right.

The logic is to treat the reference grid as an imaginary monitor. The exact location of the monitor does not influence the spatial variation pattern, as we are using relative concentration difference in this study. We do not use the distributed monitor data as the reference for several reasons. Comparing mobile data to mobile data, rather than mobile to stationary, avoids potential issues associated with temporal mismatch (15 min resolution for RAMPs versus 1 Hz for mobile) and potential confounding factors due to different detection methods for mobile sampling versus RAMPs. Additionally, not every distributed site was equipped with a CPC, so obtaining PN spatial differences at every site requires comparisons of the mobile data to itself.

CO at downtown shows significant spatial variation as indicated by mixture of red or



yellow grids. The maximum CO difference was 731 ppb, and the regional CO background was around 154 ppb, around a factor of 5 difference. For the whole 1 km<sup>2</sup> box, average absolute CO difference was 32 ppb (SD = 67 ppb). Assuming population were evenly distributed in downtown, CO exposure was underestimated by > 40% for 15% of the whole population (the portions at least one SD larger than mean). Incomplete fuel combustion from vehicles is the major source of CO in urban environments (US EPA, 2016). High traffic density and frequent traffic congestion led to higher CO emission. The street canyon effect further trapped emitted CO in downtown. These combined to contribute to the significant CO spatial variability shown in Figure 5.3(A).

Another traffic related pollutant NO<sub>2</sub> also shows street-level spatial differences. Few grey grids mean more difference relative to reference grid. The maximum NO<sub>2</sub> difference was 28 ppb, and the regional NO<sub>2</sub> background was around 7 ppb, around a factor of 4 difference. For the whole 1 km<sup>2</sup> box, average absolute NO<sub>2</sub> difference was 2 ppb (SD = 3 ppb).

UFP has relatively more gray grids compared with CO and NO<sub>2</sub>. The magnitude of the grey grid was 1800 #/cm<sup>3</sup>, 1/4 of the regional background UFP 6257 #/cm<sup>3</sup>. In comparison, Grey grid of CO and NO<sub>2</sub> were around 1/10 the background. This contrast means that UFP has both significant spatial and temporal variation. Thus, the potential number of monitors studying UFP spatiotemporal pattern in downtown locations may be more than the amount needed for CO and NO<sub>2</sub>. The maximum UFP difference was 27455 #/cm<sup>3</sup>, 3 times more than the regional UFP background. For the whole 1 km<sup>2</sup> box, average absolute UFP difference was 3257 #/cm<sup>3</sup> (SD = 4098 #/cm<sup>3</sup>).

PM<sub>1</sub> mass is mostly secondary (Robinson et al., 2007), therefore less spatial variability is expected. As PM<sub>1</sub> data were reported at 20 s resolution, grid box size was

increased from 50 m to 100 m in Figure 5.3(D). The grid filtering threshold changed from at least 20 mobile measurements to 5. The maximum  $\text{PM}_{10}$  difference was  $5 \mu\text{g}/\text{m}^3$ , around half the regional  $\text{PM}_{2.5}$  background. For the whole  $1 \text{ km}^2$  box, average absolute  $\text{PM}_{10}$  difference was  $1.4 \mu\text{g}/\text{m}^3$  ( $\text{SD} = 1.3 \mu\text{g}/\text{m}^3$ ).

Overall, CO,  $\text{NO}_2$ , UFP and  $\text{PM}_{10}$  all showed different degrees of spatial variability in downtown locations. High traffic density, points sources, and street canyon effect are attributable to this variation.

### 5.3.4 Spatial variability near an urban background monitor

Figure 5.4 shows the same four pollutants mapped in an urban background location (site 9). The site 9 was located  $\sim 500 \text{ m}$  downwind of downtown center (sites 6-8). This site had less traffic or other primary source activity, and slightly decreasing spatial variability of pollutants was observed.

More grids are in grey for CO in Figure 5.4(A) than for the downtown sites. The maximum CO difference was 541 ppb, while the maximum one for downtown was 731 ppb. Mean absolute CO difference was 40 ppb ( $\text{SD} = 60 \text{ ppb}$ ), roughly the same magnitude as downtown.

For  $\text{NO}_2$ , the reference grid tended to have higher  $\text{NO}_2$  concentrations compared with most of the areas. This grid was located near major road intersections, and traffic sources were related to this  $\text{NO}_2$  hotspot. The maximum  $\text{NO}_2$  difference was 11 ppb, close to the regional background 7 ppb, and half the maximum  $\text{NO}_2$  difference at downtown. Mean absolute  $\text{NO}_2$  difference was 2 ppb ( $\text{SD} = 2 \text{ ppb}$ ), slightly less variable compared to the downtown driving box.

For UFP, the vast majority of grids were grey, meaning less difference compared with temporal variation during driving. As temporal variation of UFP was large, these gray areas

in the maps also had large concentration changes during the driving. The maximum UFP difference was  $14315 \text{ \#/cm}^3$ , 2 times the regional background. The mean absolute UFP difference was  $2440 \text{ \#/cm}^3$  ( $\text{SD} = 2422 \text{ \#/cm}^3$ ), and both metrics were lower than downtown ones.

For  $\text{PM}_{10}$ , less spatial variability was observed compared to downtown  $\text{PM}_{10}$ . The maximum  $\text{PM}_{10}$  difference was  $1.6 \text{ \mu g/m}^3$ . The mean absolute  $\text{PM}_{10}$  difference was  $0.5 \text{ \mu g/m}^3$  ( $\text{SD} = 0.4 \text{ \mu g/m}^3$ ). Both mean and SD were around  $1 \text{ \mu g/m}^3$  less than the corresponding downtown metrics.

Overall, urban background site showed less spatial variability compared to downtown site. A monitor near the background neighborhoods could possibly represent a larger area population exposure compared to downtown monitors, though the spatial extent might still be within  $\sim 1 \text{ km}$  when addressing short-term, such as daily, exposures.

### 5.3.5 CDF of spatial vs. temporal variability

Figures 5.3 and 5.4 showed comparison between downtown and urban background sites, and their relative spatial and temporal variation. Figure 5.5 combines all relative spatial differences and presents them in a CDF plot. Red lines correspond to 50% of the relative difference for mobile sampling, and black lines mean 75% of the population of sampled grid boxes. Dashed lines show the corresponding 50% and 75% of temporal differences from Figure 5.2.

Relative CO spatial differences are generally larger than temporal fluctuations during sampling trips. The spatial differences were skewed, while temporal differences were more Gaussian. 50% of spatial differences almost encompass 75% of temporal differences. As shown in Figures 5.3 and 5.4, urban sites tended to have larger CO spatial heterogeneity compared to background sites. The magnitude of 75% spatial difference was 88 ppb, around

½ the regional background (154 ppb).

A similar trend was observed for NO<sub>2</sub>. 75% of temporal differences nearly fall between 50% of spatial differences. Traffic and other primary source drives spatial pattern of NO<sub>2</sub> on top of the boundary layer induced temporal differences. The magnitude of 75% spatial difference was 5 ppb, about the same as regional NO<sub>2</sub> (7 ppb).

For UFP, temporal and spatial difference seemed to overlap more compared to CO and NO<sub>2</sub>. The magnitude of 75% spatial difference was around 5912 #/cm<sup>3</sup>, similar to regional background (6257 #/cm<sup>3</sup>). The same 75% metric for temporal difference was 4300 #/cm<sup>3</sup>. UFP has substantial spatial and temporal variations. For urban sites, UFP was influenced by both sources and boundary layer shifting. For suburban or background sites, regional nucleation, sources and boundary layer shifting could also drive significant variation, and lead to more background baseline change in mobile sampling time series signal.

Spatial and temporal differences of PM<sub>1</sub> also overlapped with each other. The spatial difference was based on PM<sub>1</sub> from AMS measurement, while the temporal difference was based on PM<sub>2.5</sub> from MetOne monitors. PM<sub>1</sub> occupies a large portion, around 51%, of PM<sub>2.5</sub> (Buczyńska et al., 2014). The 75% magnitude of spatial difference was 1.4 µg/m<sup>3</sup>, and the regional PM<sub>2.5</sub> concentration was 9 µg/m<sup>3</sup>. PM<sub>1</sub> or PM<sub>2.5</sub> are both less variable in space and time compared with the other three pollutants.

### **5.3.6 Exposure misclassification**

Plots 5.1, 5.3, and 5.4, all illustrate spatial variations within the 1 km<sup>2</sup> driving domains. These spatial variations can be translated to short-term and/or long-term exposure misclassification. Figure 5.6 shows the mean absolute difference for each pollutant, relative

to the reference grid, for each sampling location. The right y-axis transforms left y-axis to show concentration difference as the fraction of corresponding regional background. The numbers below the bars are the number of 50 or 100 m grids inside the  $\sim 1 \text{ km}^2$  driving boxes.

The mean absolute relative CO difference shows intersite difference in Figure 5.6(A). Downtown sites 6, 7, 8 (which are grouped together) had an average difference from the reference grid of 30 ppb, 1/5 of regional background of 154 ppb. The other urban sites 10, 11, 12 showed larger average differences, 60-75 ppb. The sampling box in site 10 had two distinct sampling regions, separated by a river. The sampling area in the north of the box was river valley with a bike trail as a relatively clean environment. The other half was a commercial area with large amounts of restaurants and major traffic roads. The larger exposure difference was attributable to the significant difference in land use. Site 14 was considered to be an urban site as it was surrounding by highways. Unfavorable wind direction, low elevation topography and rapid decay of traffic emission away from road edge could explain the small spatial variation and less exposure difference.

Background sites 1, 4, 15 showed less spatial variability, and therefore smaller differences from the reference grid. This means that mobile measurement inside these backgrounds showed little variation. Putting a single monitor inside these areas introduced little exposure misclassification and could represent the surrounding air pollution exposure. However, the other background sites 2, 3, 5, 13 showed similar exposure differences as urban sites.

Site 5 was an upwind background with high altitude, separated from the downtown by a river. Although no obvious primary CO sources were located inside site 5, the site was just above a major traffic tunnel. Site 13 had green space, residential areas, and restaurants mixed inside. Similar to site 10, the significant difference in land use could lead to high exposure

differences.

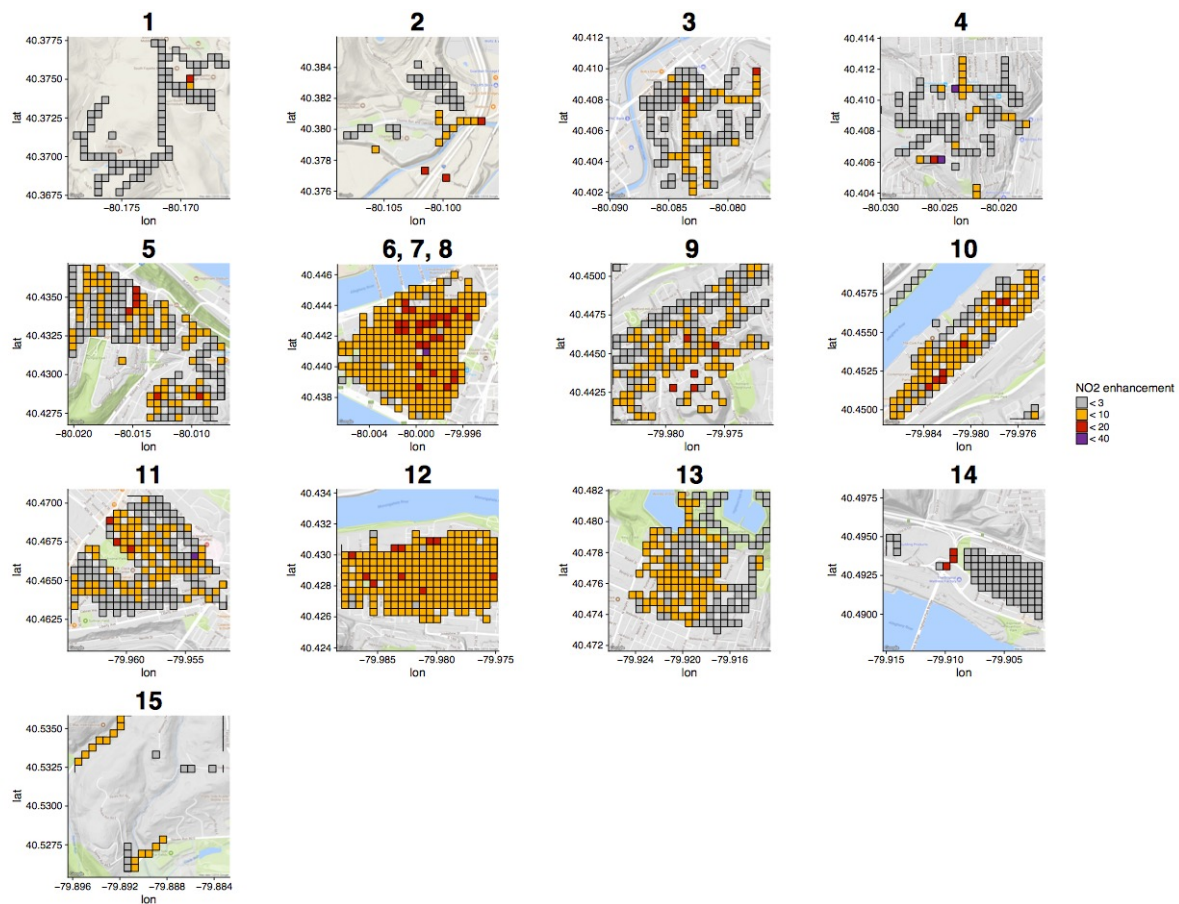
Some of the sites with larger exposure contrasts (e.g., sites 2 and 3) had fewer 50 m grids for use in the calculation of mean exposure difference than the urban sites. Additionally, fewer mobile sampling trips were made to these sites than the urban sites. Both of these factors may contribute to the higher than expected exposure differences at these sites.

NO<sub>2</sub> in Figure 5.6(B) shows a similar trend as CO. Urban sites 10 and 11 had NO<sub>2</sub> exposure differences around 3 ppb, half the urban background 7 ppb. Background sites 1, 4 and 15 had little exposure misclassification. The biggest differences were observed in sites 3 and 5. The implication is, an ideal background might not seem quite easy to identify, and a monitor located at a seemingly background neighborhood could also lead to high exposure misclassification.

UFP sites showed consistent differences of 2000 #/cm<sup>3</sup> across urban or background site except for sites 10, 12, and 13, partially due to significant temporal variation during driving at different sites. These three sites all had high density of restaurants compared with the other urban sites. Site 10 and 13 also had significant spatial heterogeneity in land use characteristics. As dining at restaurants was an important part of people daily life, the actual exposure differences could be even bigger than presented here.

The largest PM<sub>1</sub> exposure difference was at site 10, though only four 100 m grids might seem too small a number. Site 13 had an average exposure difference of 3 µg/m<sup>3</sup>. Di et al. (2017a) indicated each short-term increase of 10 µg/m<sup>3</sup> in fine PM was statistically associated with a relative increase of 1.05% (95% CI, 0.95 % -- 1.15%) in daily mortality rate. Thus, the magnitude of exposure misclassification at site 13 could lead to large uncertainty for estimating short-term exposure. Downtown center sites 6-8 had an average difference of 1.5 µg/m<sup>3</sup>, and background sites frequently had differences below 0.5 µg/m<sup>3</sup>.

Figure 5.7 shows the relationship between mean absolute difference in Figure 5.6 and representative land use covariates, (A) CO, (B) NO<sub>2</sub>, (C) UFP, and (D) PM<sub>1</sub>. Equations are labelled in bottom right. CO, NO<sub>2</sub> and PM<sub>1</sub> exposure difference showed a good correlation with vehicle density. Downtown sites (6, 7, 8) has the largest traffic density, though not the highest exposure misclassification. Figure 5.1 showed the downtown NO<sub>2</sub> spatial enhancement pattern was mostly yellow, with a few deep red or purple grids. The whole downtown areas have elevated pollutant concentrations compared with background. Site 14 was near highway, though unlikely capturing emissions from traffic on the highway, as described above. The regression slope overestimated the exposure difference in site 14. Sites 13 and 10 tended to have high exposure difference considering their moderate average traffic density. As discussed above, substantial spatial heterogeneity in land use type inside these two sites could contribute to high exposure misclassification. UFP (C) showed more exponential decay like relationship instead of linear trend with restaurant density. Many sites were clustered in the bottom left area in UFP plot (C). These sites had a nearly constant exposure misclassification regardless of land use characteristics. As land use based exposure models such as land use regression (LUR) tried to link spatial variation with changes in land use, Figure 5.7 partially explained why UFP LUR model had low  $R^2$  as changes in dependent variable UFP variation sometimes were not related to traditional LUR predictors (Kerckhoffs et al., 2017; Patton et al., 2015; Saraswat et al., 2013).

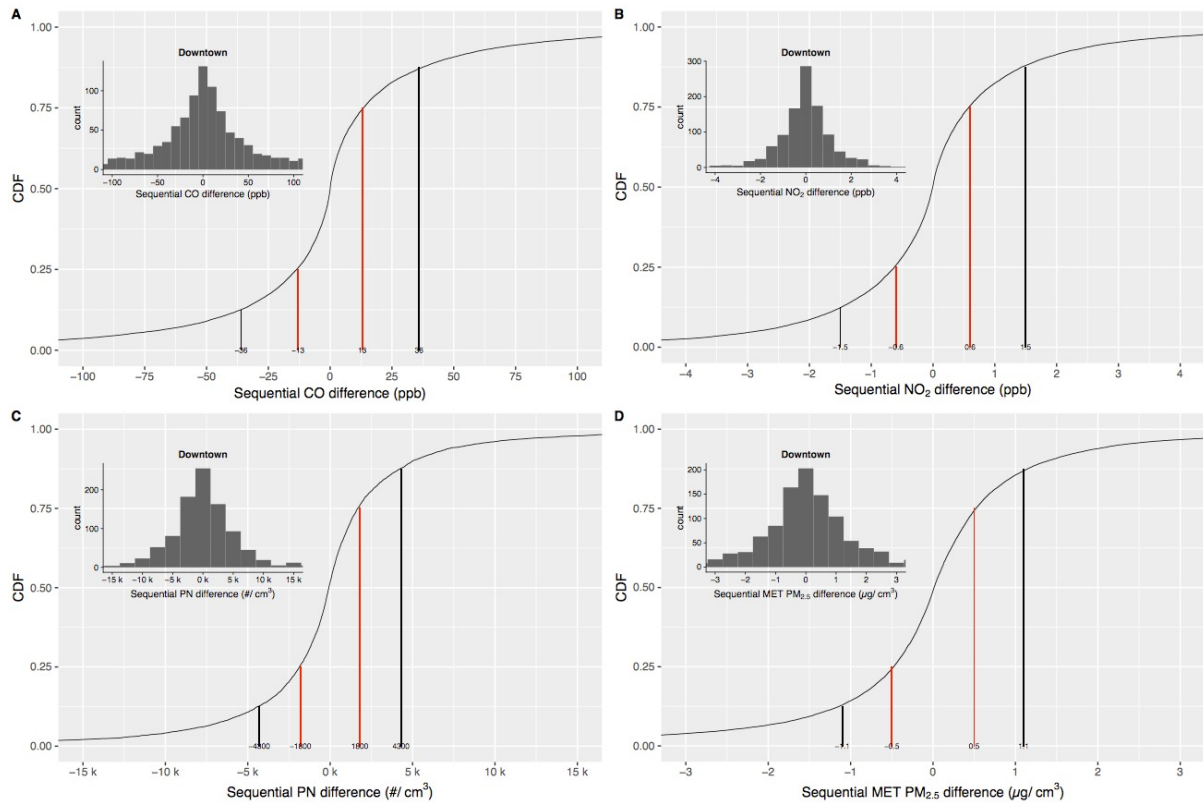


**Fig. 5.1.** Temporally corrected NO<sub>2</sub> spatial variation averaged across all sampling days

mapped with 50 m resolution inside fifteen  $\sim 1 \text{ km}^2$  mobile driving boxes. Sites 6, 7, 8 are clustered inside one driving box. Site number is labeled on top with details listed in Table 5.1.

Each 50 m box has more than twenty 1 s NO<sub>2</sub> measurements.

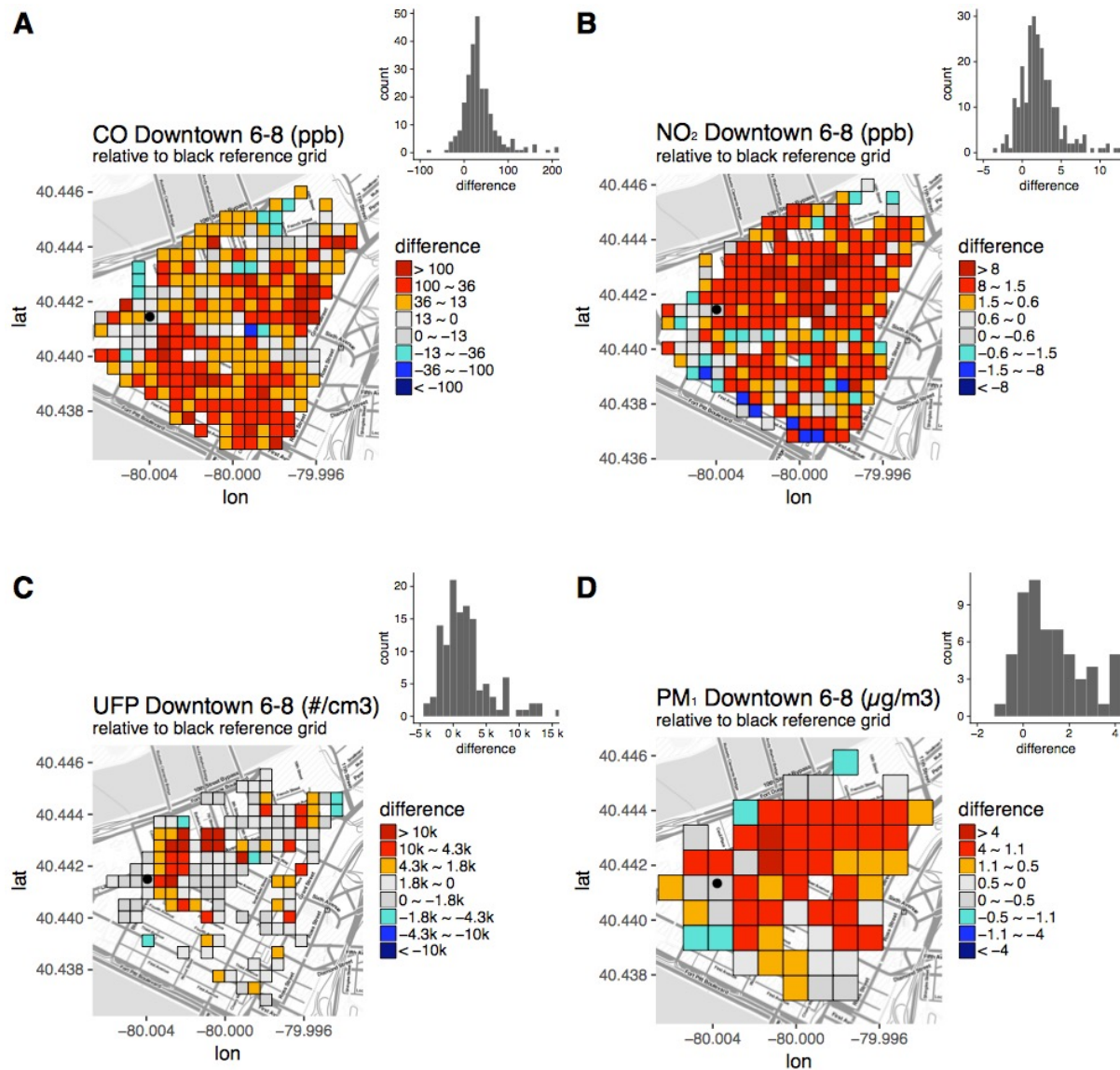




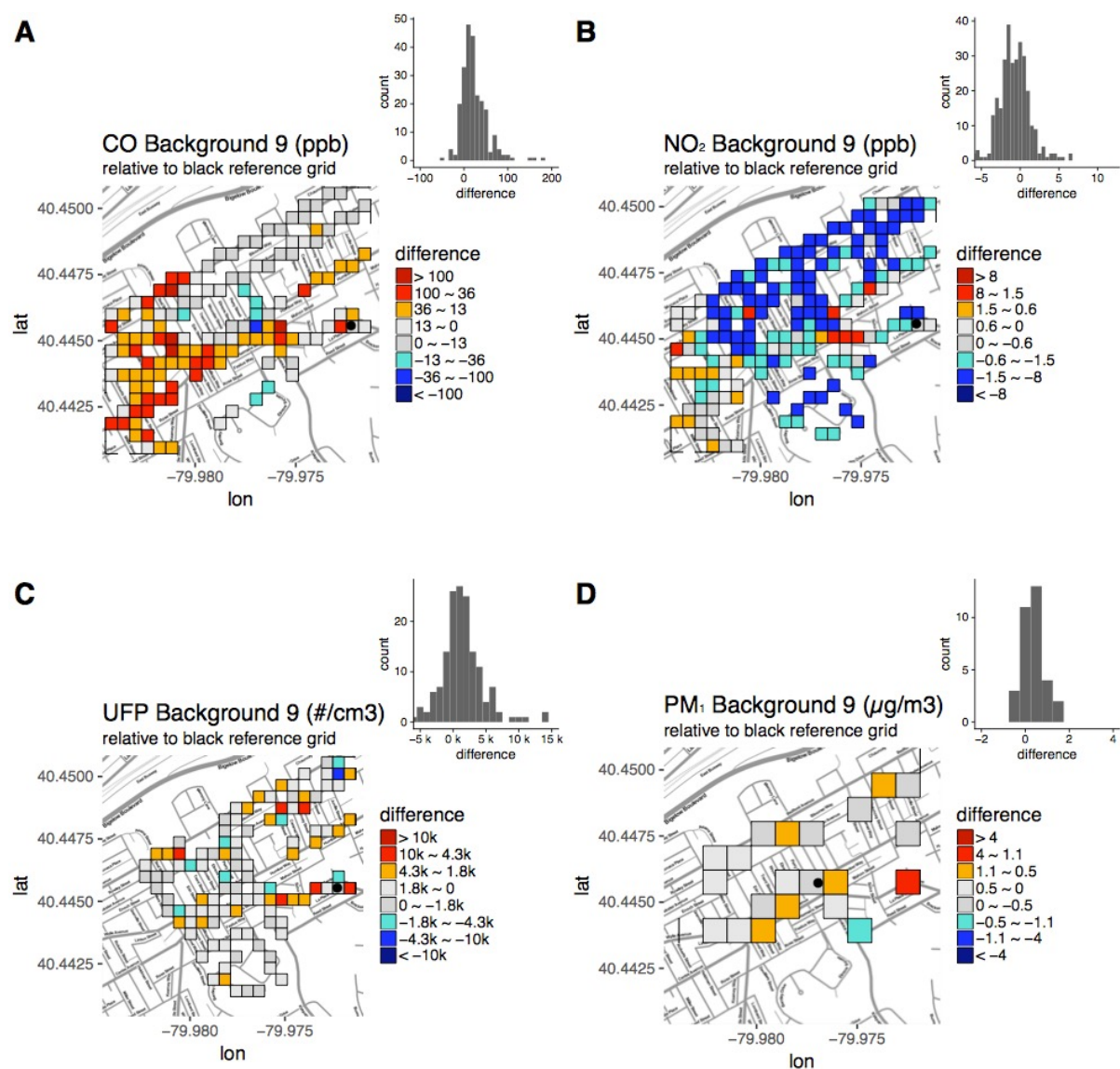
**Fig. 5.2.** Cumulative distribution function of sequential hourly pollutant concentration

difference from distributed monitors—(A) CO, (B) NO<sub>2</sub>, (C) PN, (D) PM<sub>2.5</sub>. Red vertical

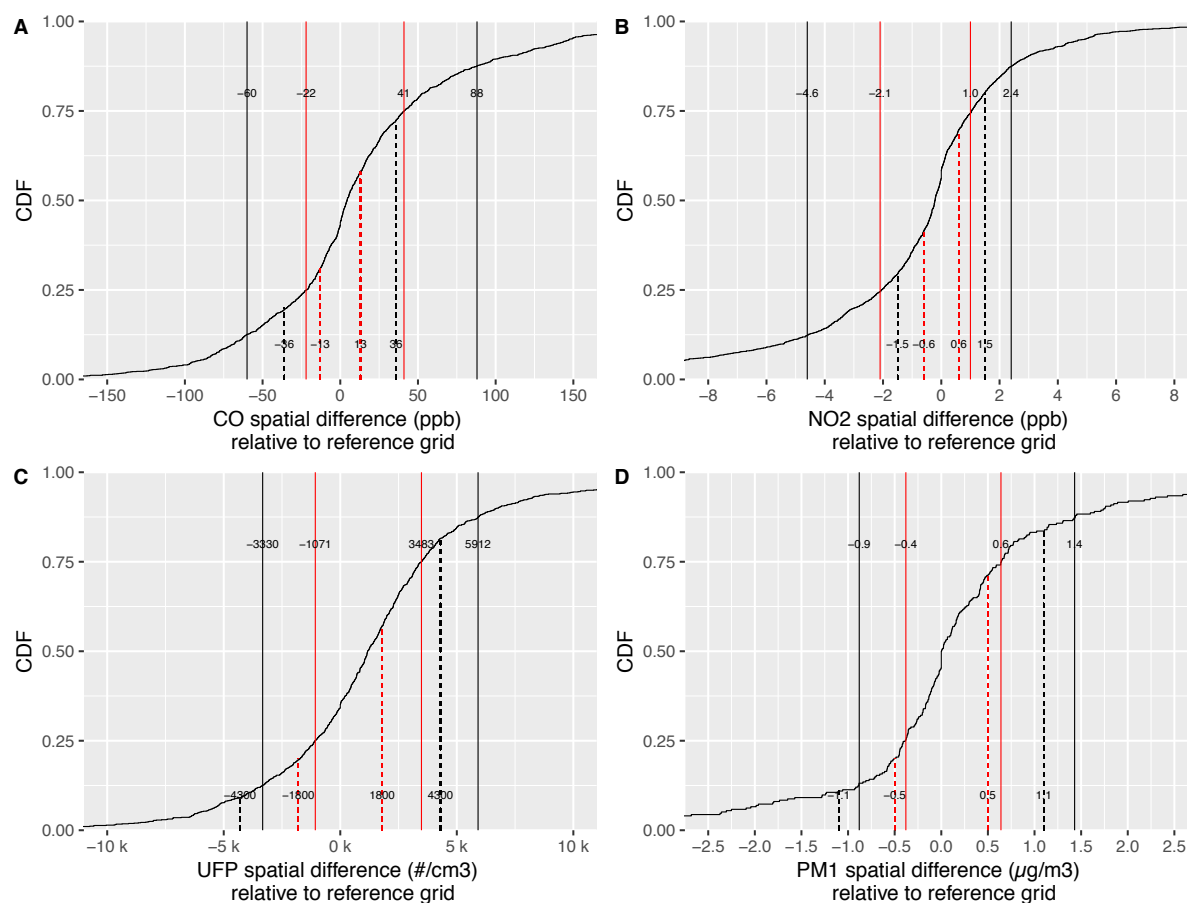
lines indicate more than 50% of values are within this range. Black means more than 75% of population. Histogram examples of pollutants at downtown are shown as an inset in top left.



**Fig. 5.3.** Spatial variation of (A) CO, (B) NO<sub>2</sub>, (C) UFP, and (D) PM<sub>1</sub> in the 1 km<sup>2</sup> mobile driving box at downtown. Grid boxes are filled with concentration difference relative to same time reference grid values (black dot). Choice of grey legend is based on cumulative distribution function of distributed monitors in Figure 5.2 (50% cutoff of temporal variation). Thus, red or blue grids represent truly different pollutant concentrations due to emission source instead of temporal fluctuations during the sampling trip.

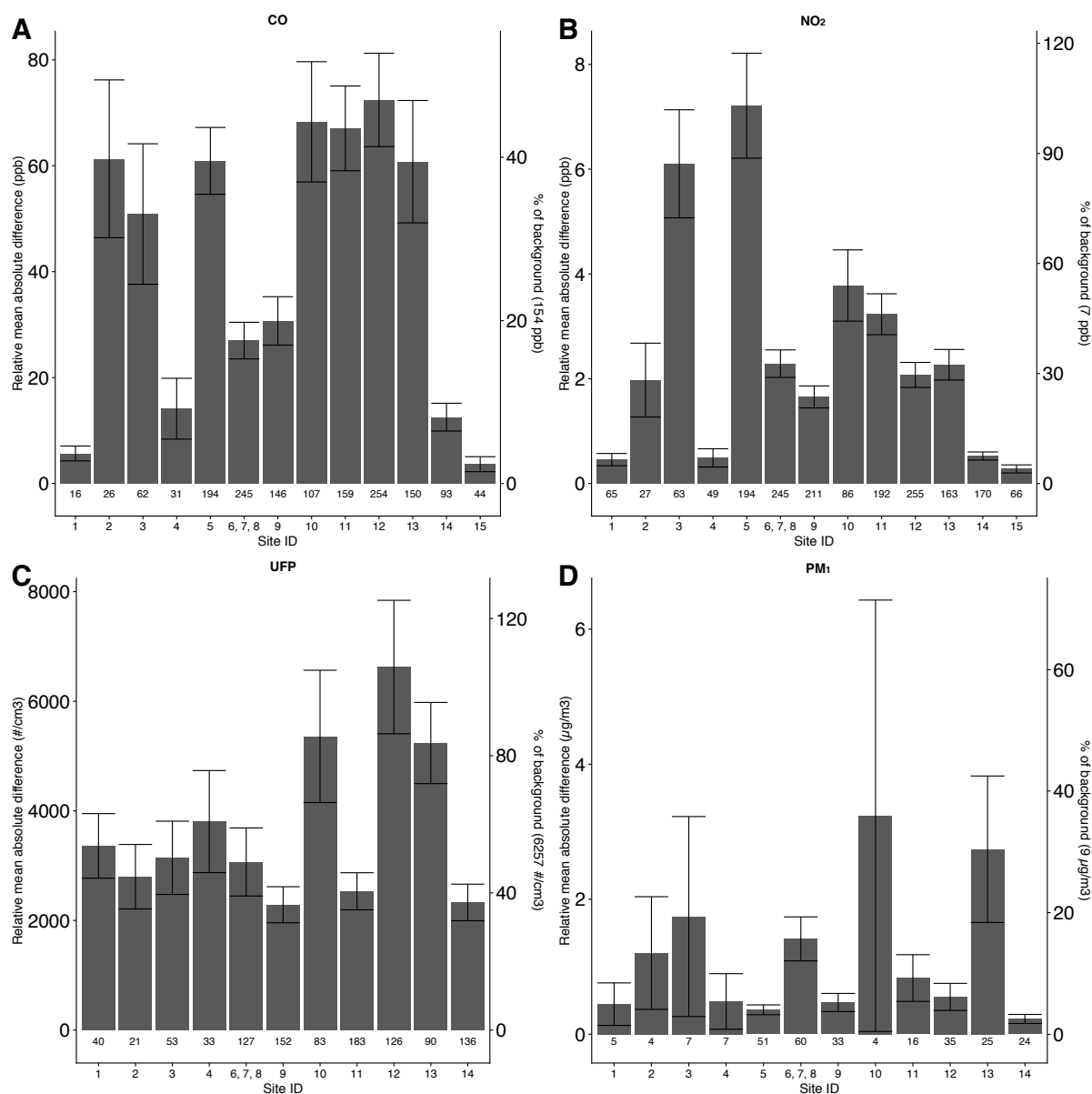


**Fig. 5.4.** Replicate of Fig. 3. based on an urban background site 9.

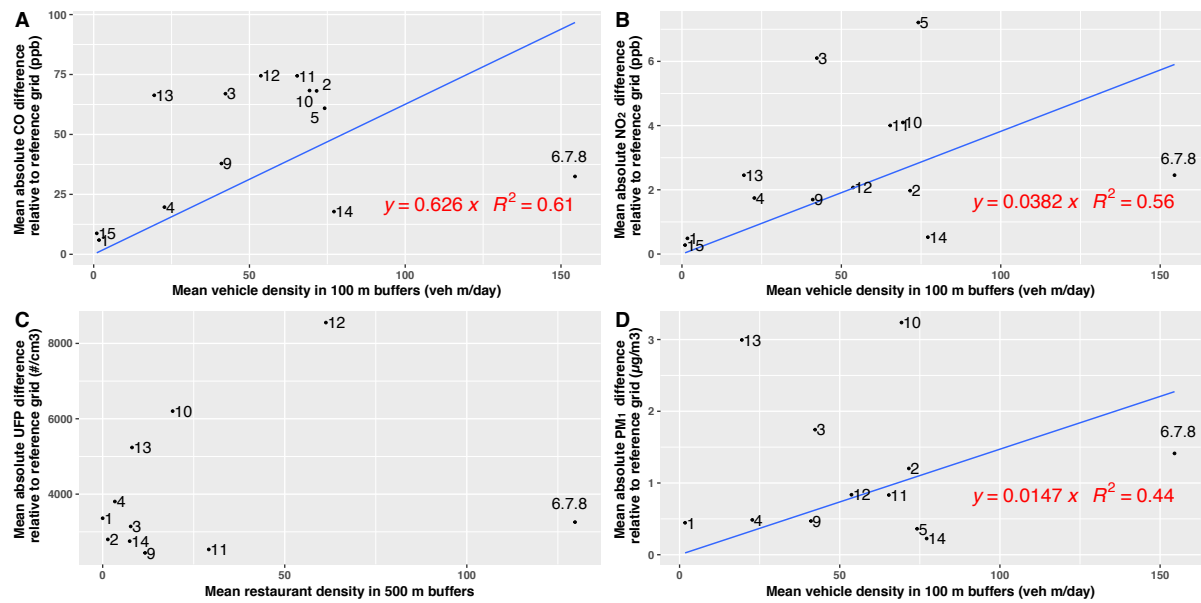


**Fig. 5.5.** CDF of relative spatial difference at all sites—(A) CO, (B) NO<sub>2</sub>, (C) PN, (D) PM<sub>1</sub>.

Solid lines, black and red, correspond to 75% and 50% of population. Dashed lines are based on temporal variability in Fig. 5.2.



**Fig. 5.6.** Mean absolute relative pollutant concentration difference (exposure misclassification) across different sites, A: CO, B: NO<sub>2</sub>, C: UFP, D: PM<sub>1</sub>. Error bars are two times the standard error. Site 5 and 15 missing from UFP or PM<sub>1</sub> panel are due to data quality failing QA/QC check.



**Fig. 5.7.** Scatterplot of mean absolute relative differences in each 1 km<sup>2</sup> and mean traffic or restaurant density. Each dot is labeled with their site ID. The regression equation is shown in bottom right.

**Table 5.1.** Overview of sampling network: site strata, operating month period. Inside each of three strata columns--traffic, restaurant and building height, 'H' means high density with threshold defined in Methods section. Site ID is defined based on geographic locations from southwest to northeast, consistent with predominant wind direction.

Site ID	Campaign	Site name	Traffic	Restaurant	Building height	Distributed monitor operating dates	Mobile driving dates
1	Transect	South Fayette				Malfunction	08/2016 - 12/2016
2	Transect	Chartiers					
3	Transect	Carnegie					
4	Transect	Beechview					
5	Downtown	Mt Washington				01/2017 - 03/2017	01/2017 - 02/2017
6	Downtown	Penn	H	H	H		
7	Downtown	Mellon	H	H	H		
8	Downtown	Church	H	H	H		
12	Downtown	Southside	H	H	H		
9	Transect/Downtown	Hill				08/2016 - 03/2017	08/2016 - 12/2016 01/2017 - 02/2017
10	Transect	Strip		H	H	08/2016 - 09/2016	08/2016 - 12/2016
11	Transect	ACHD	H	H	H		
13	Transect	Zoo		H			
14	Transect	Aspinwall	H				
15	Transect	Fox Chapel					

#### **5.4 Limitation**

One limitation of this study is the difficulty to infer exposure error for long-term health impacts analysis. The large-scale sampling network limits total visits at each site due to labor requirement. Our exposure misclassification was based on short-term spatial pattern.

Another limitation is the way we define the reference grid. We pick the grid with the most data points and implicitly assume that grid to be most representative of the long-term average. From an analytical standpoint this is defensible, but there could be other ways to pick the reference grid.



## 5.5 References

ACHD - Air Quality Reports, 2016.

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environ. Sci. Technol.* 51, 6999–7008.

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L.T., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ. Sci. Technol.* 50, 79–88. <https://doi.org/10.1021/acs.est.5b03709>

Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsett, L., Kaufman, J.D., 2010. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. *Circulation* 121, 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbee1>

Buczyńska, A.J., Krata, A., Van Grieken, R., Brown, A., Polezer, G., De Wael, K., Potgieter-Vermaak, S., 2014. Composition of PM<sub>2.5</sub> and PM<sub>1</sub> on high and low pollution event days and its relation to indoor air quality in a home for the elderly. *Sci. Total Environ.* 490, 134–143. <https://doi.org/10.1016/j.scitotenv.2014.04.102>

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K.,

- Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 389, 1907–1918. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6)
- Data | Information Portal | Allegheny County [WWW Document], n.d. URL <http://infoportal.alleghenycounty.us/data.html> (accessed 11.13.17).
- DeCarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-Deployable, High-Resolution, Time-of-Flight Aerosol Mass Spectrometer. *Anal. Chem.* 78, 8281–8289. <https://doi.org/10.1021/ac061249n>
- Department of City Planning [WWW Document], 2017. URL <http://pittsburghpa.gov/dcp> (accessed 11.13.17).
- Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017a. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. *JAMA* 318, 2446–2456. <https://doi.org/10.1001/jama.2017.17923>
- Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., Schwartz, J., 2016. Assessing PM<sub>2.5</sub> Exposures with High Spatiotemporal Resolution across the Continental United States. *Environ. Sci. Technol.* 50, 4712–4721. <https://doi.org/10.1021/acs.est.5b06121>
- Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.D., 2017b. Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* 376, 2513–2522. <https://doi.org/10.1056/NEJMoa1702747>

- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G.J., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities. *N. Engl. J. Med.* 329, 1753–1759. <https://doi.org/10.1056/NEJM199312093292401>
- Donahue, N.M., Posner, L.N., Westervelt, D.M., Li, Z., Shrivastava, M., Presto, A.A., Sullivan, R.C., Adams, P.J., Pandis, S.N., Robinson, A.L., 2016. Where Did This Particle Come From? Sources of Particle Number and Mass for Human Exposure Estimates, in: *Airborne Particulate Matter*. Royal Society of Chemistry, Cambridge, pp. 35–71.
- Geography, U.C.B., n.d. TIGER Products [WWW Document]. URL <https://www.census.gov/geo/maps-data/data/tiger.html> (accessed 11.13.17).
- Guo, Y., Tang, Q., Gong, D.-Y., Zhang, Z., 2017. Estimating ground-level PM<sub>2.5</sub> concentrations in Beijing using a satellite-based geographically and temporally weighted regression model. *Remote Sens. Environ.* 198, 140–149. <https://doi.org/10.1016/j.rse.2017.06.001>
- Hama, S.M.L., Cordell, R.L., Kos, G.P.A., Weijers, E.P., Monks, P.S., 2017. Sub-micron particle number size distribution characteristics at two urban locations in Leicester. *Atmospheric Res.* 194, 1–16. <https://doi.org/10.1016/j.atmosres.2017.04.021>
- HEI Review Panel on Ultrafine Particles, 2013. *Understanding the Health Effects of Ambient Ultrafine Particles*, HEI Perspectives 3. Health Effects Institute, Boston, MA.
- Home | City of Pittsburgh GIS Data [WWW Document], 2015. URL <http://pghgis-pittsburghpa.opendata.arcgis.com/> (accessed 11.13.17).
- Ivey, C., Holmes, H., Shi, G.-L., Balachandran, S., Hu, Y., Russell, A.G., 2017. Development of PM<sub>2.5</sub> source profiles using a hybrid chemical transport-receptor modeling

- approach. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.7b03781>
- Jayne, J.T., Leard, D.C., Zhang, X., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an Aerosol Mass Spectrometer for Size and Composition Analysis of Submicron Particles. *Aerosol Sci. Technol.* 33, 49–70. <https://doi.org/10.1080/027868200410840>
- Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. *Environ. Sci. Technol.* 44, 5334–5344. <https://doi.org/10.1021/es100008x>
- Kerckhoffs, J., Hoek, G., Messier, K.P., Brunekreef, B., Meliefste, K., Klompmaker, J.O., Vermeulen, R., 2016. Comparison of Ultrafine Particle and Black Carbon Concentration Predictions from a Mobile and Short-Term Stationary Land-Use Regression Model. *Environ. Sci. Technol.* 50, 12894–12902. <https://doi.org/10.1021/acs.est.6b03476>
- Kerckhoffs, J., Hoek, G., Vlaanderen, J., van Nunen, E., Messier, K., Brunekreef, B., Gulliver, J., Vermeulen, R., 2017. Robustness of intra urban land-use regression models for ultrafine particles and black carbon based on mobile monitoring. *Environ. Res.* 159, 500–508. <https://doi.org/10.1016/j.envres.2017.08.040>
- Kimmel, J.R., Farmer, D.K., Cubison, M.J., Sueper, D., Tanner, C., Nemitz, E., Worsnop, D.R., Gonin, M., Jimenez, J.L., 2011. Real-time aerosol mass spectrometry with millisecond resolution. *Int. J. Mass Spectrom.* 303, 15–26. <https://doi.org/10.1016/j.ijms.2010.12.004>
- Klems, J.P., Pennington, M.R., Zordan, C.A., Johnston, M.V., 2010. Ultrafine Particles Near a Roadway Intersection: Origin and Apportionment of Fast Changes in Concentration. *Environ. Sci. Technol.* 44, 7903–7907. <https://doi.org/10.1021/es102009e>

- Klompniaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. *Sci. Total Environ.* 508, 266–275. <https://doi.org/10.1016/j.scitotenv.2014.11.088>
- Kwak, K.-H., Lee, S.-H., Seo, J.M., Park, S.-B., Baik, J.-J., 2016. Relationship between rooftop and on-road concentrations of traffic-related pollutants in a busy street canyon: Ambient wind effects. *Environ. Pollut.* 208, 185–197. <https://doi.org/10.1016/j.envpol.2015.07.030>
- Lamsal, L.N., Janz, S.J., Krotkov, N.A., Pickering, K.E., Spurr, R.J.D., Kowalewski, M.G., Loughner, C.P., Crawford, J.H., Swartz, W.H., Herman, J.R., 2017. High-resolution NO<sub>2</sub> observations from the Airborne Compact Atmospheric Mapper: Retrieval and validation. *J. Geophys. Res.-Atmospheres* 122, 1953–1970. <https://doi.org/10.1002/2016JD025483>
- Leoni, C., Hovorka, J., Dočekalová, V., Cajthaml, T., Marvanová, S., 2016. Source Impact Determination using Airborne and Ground Measurements of Industrial Plumes. *Environ. Sci. Technol.* 50, 9881–9888. <https://doi.org/10.1021/acs.est.6b02304>
- Li, H.Z., Dallmann, T.R., Gu, P., Presto, A.A., 2016. Application of mobile sampling to investigate spatial variation in fine particle composition. *Atmos. Environ.* 142, 71–82. <https://doi.org/10.1016/j.atmosenv.2016.07.042>
- Li, H.Z., Dallmann, T.R., Li, X., Gu, P., Presto, A.A., 2017. Urban Organic Aerosol Exposure: Spatial Variations in Composition and Source Impacts. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.7b03674>
- Lightowlers, C., Nelson, T., Setton, E., Peter Keller, C., 2008. Determining the spatial scale for analysing mobile measurements of air pollution. *Atmos. Environ., Selected Papers*

from the First International Conference on Atmospheric Chemical Mechanisms 42,  
5933–5937. <https://doi.org/10.1016/j.atmosenv.2008.03.033>

Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M.A., Amann, M., Anderson, H.R., Andrews, K.G., Aryee, M., Atkinson, C., Bacchus, L.J., Bahalim, A.N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M.L., Blore, J.D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N.G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R.T., Byers, T.E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J.S., Cheng, A.T.-A., Child, J.C., Cohen, A., Colson, K.E., Cowie, B.C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Jarlais, D.C.D., Devries, K., Dherani, M., Ding, E.L., Dorsey, E.R., Driscoll, T., Edmond, K., Ali, S.E., Engell, R.E., Erwin, P.J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M.M., Flaxman, S., Fowkes, F.G.R., Freedman, G., Freeman, M.K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H.R., Hall, W., Hoek, H.W., Hogan, A., Hosgood, H.D., Hoy, D., Hu, H., Hubbell, B.J., Hutchings, S.J., Ibeanusi, S.E., Jacklyn, G.L., Jasrasaria, R., Jonas, J.B., Kan, H., Kanis, J.A., Kassebaum, N., Kawakami, N., Khang, Y.-H., Khatibzadeh, S., Khoo, J.-P., Kok, C., Laden, F., Lalloo, R., Lan, Q., Lathlean, T., Leasher, J.L., Leigh, J., Li, Y., Lin, J.K., Lipshultz, S.E., London, S., Lozano, R., Lu, Y., Mak, J., Malekzadeh, R., Mallinger, L., Marcenes, W., March, L., Marks, R., Martin, R., McGale, P., McGrath, J., Mehta, S., Memish, Z.A., Mensah, G.A., Merriman, T.R., Micha, R., Michaud, C., Mishra, V., Hanafiah, K.M., Mokdad, A.A., Morawska, L., Mozaffarian, D., Murphy, T., Naghavi, M., Neal, B., Nelson, P.K., Nolla, J.M., Norman, R., Olives, C., Omer, S.B., Orchard,

- J., Osborne, R., Ostro, B., Page, A., Pandey, K.D., Parry, C.D., Passmore, E., Patra, J., Pearce, N., Pelizzari, P.M., Petzold, M., Phillips, M.R., Pope, D., Pope, C.A., Powles, J., Rao, M., Razavi, H., Rehfuess, E.A., Rehm, J.T., Ritz, B., Rivara, F.P., Roberts, T., Robinson, C., Rodriguez-Portales, J.A., Romieu, I., Room, R., Rosenfeld, L.C., Roy, A., Rushton, L., Salomon, J.A., Sampson, U., Sanchez-Riera, L., Sanman, E., Sapkota, A., Seedat, S., Shi, P., Shield, K., Shivakoti, R., Singh, G.M., Sleet, D.A., Smith, E., Smith, K.R., Stapelberg, N.J., Steenland, K., Stöckl, H., Stovner, L.J., Straif, K., Straney, L., Thurston, G.D., Tran, J.H., Dingenen, R.V., Donkelaar, A. van, Veerman, J.L., Vijayakumar, L., Weintraub, R., Weissman, M.M., White, R.A., Whiteford, H., Wiersma, S.T., Wilkinson, J.D., Williams, H.C., Williams, W., Wilson, N., Woolf, A.D., Yip, P., Zielinski, J.M., Lopez, A.D., Murray, C.J., Ezzati, M., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 380, 2224–2260. [https://doi.org/10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)
- Liu, T., Wang, Z., Huang, D.D., Wang, X., Chan, C.K., 2017. Significant Production of Secondary Organic Aerosol from Emissions of Heated Cooking Oils. *Environ. Sci. Technol. Lett.* <https://doi.org/10.1021/acs.estlett.7b00530>
- Martin, F., Fileni, L., Palomino, I., Vivanco, M.G., Garrido, J.L., 2014. Analysis of the spatial representativeness of rural background monitoring stations in Spain. *Atmospheric Pollut. Res.* 5, 779–788. <https://doi.org/10.5094/APR.2014.087>
- Patton, A.P., Zamore, W., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2015. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. *Environ. Sci. Technol.* 49, 6051–6060.

<https://doi.org/10.1021/es5061676>

Petters, M.D., Kreidenweis, S.M., 2007. A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. *Atmos Chem Phys* 7, 1961–1971.

<https://doi.org/10.5194/acp-7-1961-2007>

Piersanti, A., Vitali, L., Righini, G., Cremona, G., Ciancarella, L., 2015. Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollut. Res.* 6, 953–960. <https://doi.org/10.1016/j.apr.2015.04.005>

Pope, C.A.I., Dockery, D.W., 2006. Health Effects of Fine Particulate Air Pollution: Lines that Connect. *J. Air Waste Manag. Assoc.* 56, 709–742. <https://doi.org/10.1080/10473289.2006.10464485>

Pope, C.A.I., Ezzati, M., Dockery, D.W., 2009. Fine-Particulate Air Pollution and Life Expectancy in the United States. *N. Engl. J. Med.* 360, 376–386.

Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M.J., Brunekreef, B., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P.E., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Sørensen, M., Tjønneland, A., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Key, T.J., de Hoogh, K., Concini, H., Nagel, G., Vilier, A., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., Hoek, G., 2013. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet*



- Oncol. 14, 813–822. [https://doi.org/10.1016/S1470-2045\(13\)70279-1](https://doi.org/10.1016/S1470-2045(13)70279-1)
- Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging. *Science* 315, 1259–1262. <https://doi.org/10.1126/science.1133061>
- Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon Particulate Matter in New Delhi, India. *Environ. Sci. Technol.* 47, 12903–12911. <https://doi.org/10.1021/es401489h>
- Seinfeld, J.H., Pandis, S.N., 2016. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons.
- Shi, Y., Lau, K.K.-L., Ng, E., 2016. Developing Street-Level PM<sub>2.5</sub> and PM<sub>10</sub> Land Use Regression Models in High-Density Hong Kong with Urban Morphological Factors. *Environ. Sci. Technol.* 50, 8178–8187. <https://doi.org/10.1021/acs.est.6b01807>
- Stafoggia, M., Schneider, A., Cyrys, J., Samoli, E., Andersen, Z.J., Bedada, G.B., Bellander, T., Cattani, G., Eleftheriadis, K., Faustini, A., Hoffmann, B., Jacquemin, B., Katsouyanni, K., Massling, A., Pekkanen, J., Perez, N., Peters, A., Quass, U., Yli-Tuomi, T., Forastiere, F., Group, on behalf of the U.S., 2017. Association Between Short-term Exposure to Ultrafine Particles and Mortality in Eight European Urban Areas. *Epidemiology* 28, 172–180. <https://doi.org/10.1097/EDE.0000000000000599>
- Strak, M., Janssen, N., Beelen, R., Schmitz, O., Vaartjes, I., Karssenbergh, D., van den Brink, C., Bots, M.L., Dijst, M., Brunekreef, B., Hoek, G., 2017. Long-term exposure to particulate matter, NO<sub>2</sub> and the oxidative potential of particulates and diabetes prevalence in a large national health survey. *Environ. Int.* 108, 228–236. <https://doi.org/10.1016/j.envint.2017.07.011>

[doi.org/10.1016/j.envint.2017.08.017](https://doi.org/10.1016/j.envint.2017.08.017)

- Tan, Y., Lipsky, E.M., Saleh, R., Robinson, A.L., Presto, A.A., 2014a. Characterizing the Spatial Variation of Air Pollutants and the Contributions of High Emitting Vehicles in Pittsburgh, PA. *Environ. Sci. Technol.* 48, 14186–14194. <https://doi.org/10.1021/es5034074>
- Tan, Y., Robinson, A.L., Presto, A.A., 2014b. Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. *Atmos. Environ.* 99, 333–340. <https://doi.org/10.1016/j.atmosenv.2014.10.003>
- US EPA, O., 2016. 2014 National Emissions Inventory (NEI) Data [WWW Document]. US EPA. URL <https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data> (accessed 11.13.17).
- Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., 2015. Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmos. Environ.* 105, 148–161. <https://doi.org/10.1016/j.atmosenv.2015.01.017>
- Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., De Baets, B., 2016. Opportunistic mobile air pollution monitoring: A case study with city wardens in Antwerp. *Atmos. Environ.* 141, 408–421.
- Vardoulakis, S., Gonzalez-Flesca, N., Fisher, B.E.A., Pericleous, K., 2005. Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmos. Environ.* 39, 2725–2736. <https://doi.org/10.1016/j.atmosenv.2004.05.067>
- Vitali, L., Morabito, A., Adani, M., Assennato, G., Ciancarella, L., Cremona, G., Giua, R., Pastore, T., Piersanti, A., Righini, G., Russo, F., Spagnolo, S., Tanzarella, A., Tinarelli, G., Zanini, G., 2016. A Lagrangian modelling approach to assess the

- representativeness area of an industrial air quality monitoring station. *Atmospheric Pollut. Res.* 7, 990–1003. <https://doi.org/10.1016/j.apr.2016.06.002>
- Yin, X., Dai, T., Schutgens, N.A.J., Goto, D., Nakajima, T., Shi, G., 2016. Effects of data assimilation on the global aerosol key optical properties simulations. *Atmospheric Res.* 178, 175–186. <https://doi.org/10.1016/j.atmosres.2016.03.016>
- Zimmerman, N., Jeong, C.-H., Wang, J.M., Ramos, M., Wallace, J.S., Evans, G.J., 2015. A source-independent empirical correction procedure for the fast mobility and engine exhaust particle sizers. *Atmos. Environ.* 100, 178–184. <https://doi.org/10.1016/j.atmosenv.2014.10.054>
- Zimmerman, N., Presto, A.A., Kumar, S.P.N., Gu, J., Hauryliuk, A., Robinson, E.S., Robinson, A.L., Subramanian, R., 2017. Closing the gap on lower cost air quality monitoring: machine learning calibration models to improve low-cost sensor performance. *Atmos Meas Tech Discuss* 2017, 1–36. <https://doi.org/10.5194/amt-2017-260>

**Chapter 6: Conclusions**

## **Chapter 6**

### **Conclusions**

## 6.1 Summary of major findings

Long and short-term exposure to particulate matter are linked to adverse health impacts. Single or few stationary monitors are inadequate to capture substantial intracity spatial variation of airborne pollutants, thus leading to high uncertainty in describing population exposure. This dissertation addresses urban aerosol exposure focusing on spatiotemporal variation and source characterization. Main achievements of the thesis were on improving the characterization of spatial and temporal variation of multiple airborne pollutants. Parts of the improvements were about sampling design, choosing the appropriate sampling network for different target pollutants. The thesis also dealt with current knowledge gap regarding exposure model building, specifically the robustness of feeding LUR models with mobile measurements. We investigated and compared all available sampling approaches – mobile sampling (moving or short-term stationary, 1 hr), distributed monitors, and supersite. Novel parts of the thesis included deploying low-cost sensors on distributed monitors to provide high temporal resolution (15 min) data, using aerosol mass spectrometry to get spatially resolved PM<sub>1</sub> with the mobile sampling van, providing a cost-effective framework to measure spatial variation in particle composition, identifying certain OC fractions to be a marker for fresh vs. aged organics exposure, and quantifying spatial variability near stationary monitors with special sampling design (mobile sampling around monitors in diverse areas).

As reported in Chapter 2, application of mobile sampling was conducted to investigate spatial variation in particle composition. This research highlighted the growing importance of measuring and modeling spatial variation in PM trace elements, as they exhibited intracity variation and could drive health effects of particulate matter. However, limited studies existed partially due to the logistic requirement of establishing distributed

network and monetary cost for filter analysis in the lab. We suggested a practical mobile sampling approach to alleviate the drawbacks and captured substantial intracity spatial variation in particle composition. 36 sites were selected in diverse microenvironments based on traffic density, proximity to industrial sources, and elevation. Trace elements showed a broad range of concentrations from 0 to 300 ng/m<sup>3</sup>. We showed our mobile sampling data were not statistically significant different from same time regulatory monitor measurements. This finding indicated we could use mobile sampling to capture spatial variation in particle composition while in the same time yielding a representative dataset. Correlation strength between PM composition in summer and winter indicated different source intensity or source categories. Crustal elements such as Si, Ca, and Ti showed good correlation in both seasons. Biomass burning tracer K showed stronger correlation with other elements in summer than winter, likely because of the influence of biomass burning during winter months. Sulfate is dominated by secondary production, and little correlation was observed with other primary trace metals. The correlation results could help classify source profiles in source apportionment studies such as positive matrix factorization. We examined the effectiveness of using Zn (from brake and tire wear) as traffic intensity marker in the ambient environment, and Zn proved less useful compared with elemental carbon or polycyclic aromatic hydrocarbons. Source apportionment studies relied on good tracers (chemical inert and source specific) to accurately quantify source contributions at receptor sites. Our finding indicated using Zn as the sole tracer for traffic emissions in ambient sampling environment was not robust. With mobile measurements as model input, we compared and selected land use regression to predict long-term spatial gradient in the whole study domain (Allegheny County, PA). We demonstrated traffic related predictors accounted for major variability observed. This manuscript presents the first successful LUR study for particle composition in

North America using a cost-effective mobile sampling approach.

Chapter 3 continued to study spatial variation of PM compositions, and specifically analyzed one major PM component (20 to 90% in mass fraction) (Zhang et al., 2007), organic carbon (OC), based on the same mobile sampling campaign. OC is also linked to adverse health outcomes (Urch et al., 2008, 2005). Though secondary organic aerosol dominates the OC mass in most environments (Zhang et al., 2007), in near source regions (e.g., near roadways) emissions of primary OC can strongly influence population exposures (Donahue et al., 2016). We collected OC on quartz filters and quantified different OC fractions OCX using thermo-optical methods. Particle phase OC tripled from upwind to near source locations, and this indicated substantial intracity OC spatial variation and the need to accurately quantify human exposure to OC. We illustrated the different fresh emission dependence of OCX collected on the bare quartz. OC2 and OC3 are good fresh emission markers, with OC3 specifically being a marker for traffic emissions. OC4 and PC are the secondary OC in the study region. The different OCXs therefore suggest different source implications. These implications are also applicable to national speciation network. This finding shed new light on using existing and historical monitor measurements to investigate population exposure status on a national scale. Ridley et al. (2017) indicated OC declined by 25-50% between 1990 and 2012 based on national speciation network data. They showed 2/3 OC reduction could be explained by vehicle emissions and residential fuel burning. Such a large OC decrease was not anticipated in the 2011 EPA report to congress. This implied the Clean Air Act might be more beneficial than what we think. Our finding could further support this study by analyzing how primary source marker OC2 and OC3 varied in the last few decades. Similar to Chapter 2, we derived long-term LUR models for OCX and total OC with an average  $R^2$  0.64 (SD=0.09). We further combined LUR predictions with census block



population data to quantify long-term exposure spatial pattern. We showed primary source emitted OC2 and OC3 were the main drivers for enriched urban OC exposure. This finding implied that controlling OC2 and OC3 might be more effective in reducing total organics exposure and thus gaining more public health benefits.

Chapters 2 and 3 investigated different PM components' spatial variation with mobile sampling, and used land use based exposure models to describe the long-term spatial variation. Population exposure in real world included temporal variation besides spatial variation. Especially for acute health impacts, we needed to understand temporal variability.

Chapter 4 studied temporal variation of both gas and particle pollutants, addressed current knowledge gap regarding the theoretical validity of developing long-term exposure models based on short-term mobile measurement, and provided suggestions on better capturing intracity spatiotemporal variation of multiple pollutants. We used a hybrid sampling network including a mobile sampling platform, 14 distributed monitors, and a supersite to investigate a wide range of pollutants ( $\text{CO}_2$ , CO,  $\text{NO}_2$ ,  $\text{PM}_{10}$  mass and composition, and particle number PN). Mobile sampling was conducted repeatedly near distributed monitors ( $\sim 1 \text{ km}^2$ ), and results from both platforms showed good agreement. PN was highly dynamic, and hotspots were strongly associated with restaurants and highway traffic. Spatial variation existed both between and within driving boxes ( $\sim 1 \text{ km}^2$ ). Within box spatial variability was higher in downtown areas due to combined influence of traffic, street canyon, and point sources. Distributed monitors recorded different temporal pattern across pollutants. A single background correction that lifts or drops all sites together may be too simple to fully capture between-day or within-day variations, especially in cases where time- or location-weighted exposures are a desired endpoint. Spatial variation was generally larger than temporal variation among all five pollutants ( $\text{CO}_2$ ,  $\text{NO}_2$ , CO, PN, and  $\text{PM}_{10}$ ). This suggested that when

we built long-term LUR based on short-term mobile measurements, the potential harmful effect of temporal fluctuations in short-term mobile sampling data was limited. As for measuring spatiotemporal variation of specific pollutants, different sampling strategy was suggested as followed. Distributed monitors, with one or two near major sources, probably did a good job of estimating PM<sub>2.5</sub> exposures within a given city. For the other pollutants such as CO, NO<sub>2</sub>, and PN, spatial variation tended to be more important to the overall picture, so a network of monitors or trying to mix stationary sites with mobile sampling added more value.

Chapter 5 used the same hybrid sampling network in Chapter 4 to study spatial representativeness of stationary monitors in different environments. Monitoring network is essential for protecting public health, though studies pointed out the limited spatial representativeness of stationary monitors (Apte et al., 2017; Piersanti et al., 2015; Shi et al., 2016; Vardoulakis et al., 2005; Vitali et al., 2016). This limitation could create additional uncertainty in epidemiology results, chemical transport model output, and remote sensing studies. CO and NO<sub>2</sub> exhibited within-neighborhood ( $\sim 1 \text{ km}^2$ ) spatial variation, with hotspots elevated by up to a factor of 5 above the regional background. UFP was the most variable, with spatial variations up to an order of magnitude higher than background. PM<sub>1</sub> showed the least spatial variability. Urban neighborhoods showed larger spatial variability compared to suburban and background ones. The magnitude of relative concentration differences was commonly larger than the interquartile range of hourly temporal differences. This suggested concentration variations captured by mobile sampling were more indicative of source impacts. We quantified exposure misclassification within each  $1 \text{ km}^2$  neighborhood. Using a single monitor measurement to represent surrounding  $\sim 1 \text{ km}^2$  areas could introduce an average daily exposure misclassification of 46 ppb (SD = 26) for CO (30% of regional

background), 3 ppb (SD = 2) for NO<sub>2</sub> (43% of background), 4000 #/cm<sup>3</sup> (SD = 1900) for ultrafine particle number (64% of background), and 1.2 µg/m<sup>3</sup> (SD = 1) for PM<sub>1</sub> (13% of background). We identified sites with significant land use spatial heterogeneity seemed to have larger exposure misclassification, and background sites tended to have smaller exposure differences. We regressed exposure misclassification values against representative land use predictors such as traffic or restaurant density, and found fair correlation between them. These findings addressed limited representative spatial extent of urban monitors measuring multiple pollutants, and provided quantitative short-term exposure misclassification.

## 6.2 Future studies

This dissertation studied spatiotemporal variation of airborne pollutants in the urban environment. The findings can be used as direct input for epidemiology study. We built local LUR models for PM composition and organic aerosol, and we can try using our LUR model output coupled with other statistical tools such as Environmental Benefits Mapping and Analysis Program (BenMAP) to analyze the relationship between hospital admittance and air pollution events in the region. The findings can be an interesting case study in the context of living in an old industrialized U.S. city.

We showed different fresh emission dependence of OC fractions (OCX), and the relationships held applicable in national specification network. This finding shed new light on utilizing existing thermo-optical OC fractions data. We can try building a nationwide OCX LUR to start testing if OCX exposures are associated with certain health effects.

We mainly relied on LUR to model spatial variation of pollutant. The major drawback of land use regression is transferability. One model built in one place commonly cannot be transferred to another location. So solely relying on land use regression and trying to analyze large population means a multi-city campaign like MESA-air (Kim et al., 2016), which is

hard. Recently, people combined LUR with machine learning random forest approach to improve prediction accuracy and transferability (Brokamp et al., 2017). Traditional LUR assumes a linear relationship between GIS predictors and concentrations, and this assumption can be too simple. Random forest, on the other hand, accounted for the non-linear relationships and interaction between predictors (Brokamp et al., 2017). We can apply machine learning to our existing mobile sampling dataset, and address whether random forest can improve our LUR model transferability.

Emission control proves successful for pollutants such as NO<sub>x</sub> in urban area. However, Zhao et al. (2017) indicates though NO<sub>x</sub> emission is consistently decreasing with new control technology and more strict law enforcement adopted, the interaction between organics, NO<sub>x</sub>, and ozone could lead to a possible higher ambient organic aerosol in future simulating scenario. This finding is backed up by Praske et al. (2017), which indicates the autoxidation of peroxy radicals is increasing important in urban and suburban North America. Studying spatiotemporal variation of pollutants and source characterization can get us a good first step identifying which might be the optimal air pollution control strategy. But the evolving nature of primary emissions in the atmosphere make it especially hard for us to be sure we are doing definitely the right thing. Ambient measurement and lab chemistry experiments need to be combined together to illustrate the complex air pollution nature and contribute to human health.

### 6.3 References

- Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environ. Sci. Technol.* 51, 6999–7008.
- Brokamp, C., Jandarov, R., Rao, M.B., LeMasters, G., Ryan, P., 2017. Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches. *Atmos. Environ.* 151, 1–11.
- Kim, S.-Y., Sheppard, L., Bergen, S., Szpiro, A.A., Sampson, P.D., Kaufman, J.D., Vedal, S., 2016. Prediction of fine particulate matter chemical components with a spatio-temporal model for the Multi-Ethnic Study of Atherosclerosis cohort. *J. Expo. Sci. Environ. Epidemiol.* 26, 520–528. <https://doi.org/10.1038/jes.2016.29>
- Piersanti, A., Vitali, L., Righini, G., Cremona, G., Ciancarella, L., 2015. Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollut. Res.* 6, 953–960. <https://doi.org/10.1016/j.apr.2015.04.005>
- Praske, E., Otkjær, R.V., Crounse, J.D., Hethcox, J.C., Stoltz, B.M., Kjaergaard, H.G., Wennberg, P.O., 2017. Atmospheric autoxidation is increasingly important in urban and suburban North America. *Proc. Natl. Acad. Sci.* 201715540. <https://doi.org/10.1073/pnas.1715540115>
- Ridley, D.A., Heald, C.L., Ridley, K.J., Kroll, J.H., 2017. Causes and consequences of decreasing atmospheric organic aerosol in the United States. *Proc. Natl. Acad. Sci.* 201700387. <https://doi.org/10.1073/pnas.1700387115>
- Shi, Y., Lau, K.K.-L., Ng, E., 2016. Developing Street-Level PM<sub>2.5</sub> and PM<sub>10</sub> Land Use Regression Models in High-Density Hong Kong with Urban Morphological Factors.

- Environ. Sci. Technol. 50, 8178–8187. <https://doi.org/10.1021/acs.est.6b01807>
- Urch, B., Brook, J.R., Wasserstein, D., Wasserstein, D., Brook, R.D., Brook, R.D., Rajagopalan, S., Corey, P., Silverman, F., 2008. Relative Contributions of PM<sub>2.5</sub> Chemical Constituents to Acute Arterial Vasoconstriction in Humans. *Inhal. Toxicol.* 16, 345–352.
- Urch, B., Silverman, F., Corey, P., Brook, J.R., Lukic, K.Z., Rajagopalan, S., Brook, R.D., 2005. Acute Blood Pressure Responses in Healthy Adults During Controlled Air Pollution Exposures. *Environ. Health Perspect.* 113, 1052–1055.
- Vardoulakis, S., Gonzalez-Flesca, N., Fisher, B.E.A., Pericleous, K., 2005. Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmos. Environ.* 39, 2725–2736. <https://doi.org/10.1016/j.atmosenv.2004.05.067>
- Vitali, L., Morabito, A., Adani, M., Assennato, G., Ciancarella, L., Cremona, G., Giua, R., Pastore, T., Piersanti, A., Righini, G., Russo, F., Spagnolo, S., Tanzarella, A., Tinarelli, G., Zanini, G., 2016. A Lagrangian modelling approach to assess the representativeness area of an industrial air quality monitoring station. *Atmospheric Pollut. Res.* 7, 990–1003. <https://doi.org/10.1016/j.apr.2016.06.002>
- Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Allan, J.D., Coe, H., Ulbrich, I., Alfarra, M.R., Takami, A., Middlebrook, A.M., Sun, Y.L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P.F., Salcedo, D., Onasch, T., Jayne, J.T., Miyoshi, T., Shimojo, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R.J., Rautiainen, J., Sun, J.Y., Zhang, Y.M., Worsnop, D.R., 2007. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophys. Res. Lett.* 34, L13801. <https://doi.org/>

10.1029/2007GL029979

Zhao, Y., Saleh, R., Saliba, G., Presto, A.A., Gordon, T.D., Drozd, G.T., Goldstein, A.H.,  
Donahue, N.M., Robinson, A.L., 2017. Reducing secondary organic aerosol formation  
from gasoline vehicle exhaust. *Proc. Natl. Acad. Sci.* 114, 6984–6989. [https://doi.org/](https://doi.org/10.1073/pnas.1620911114)  
10.1073/pnas.1620911114

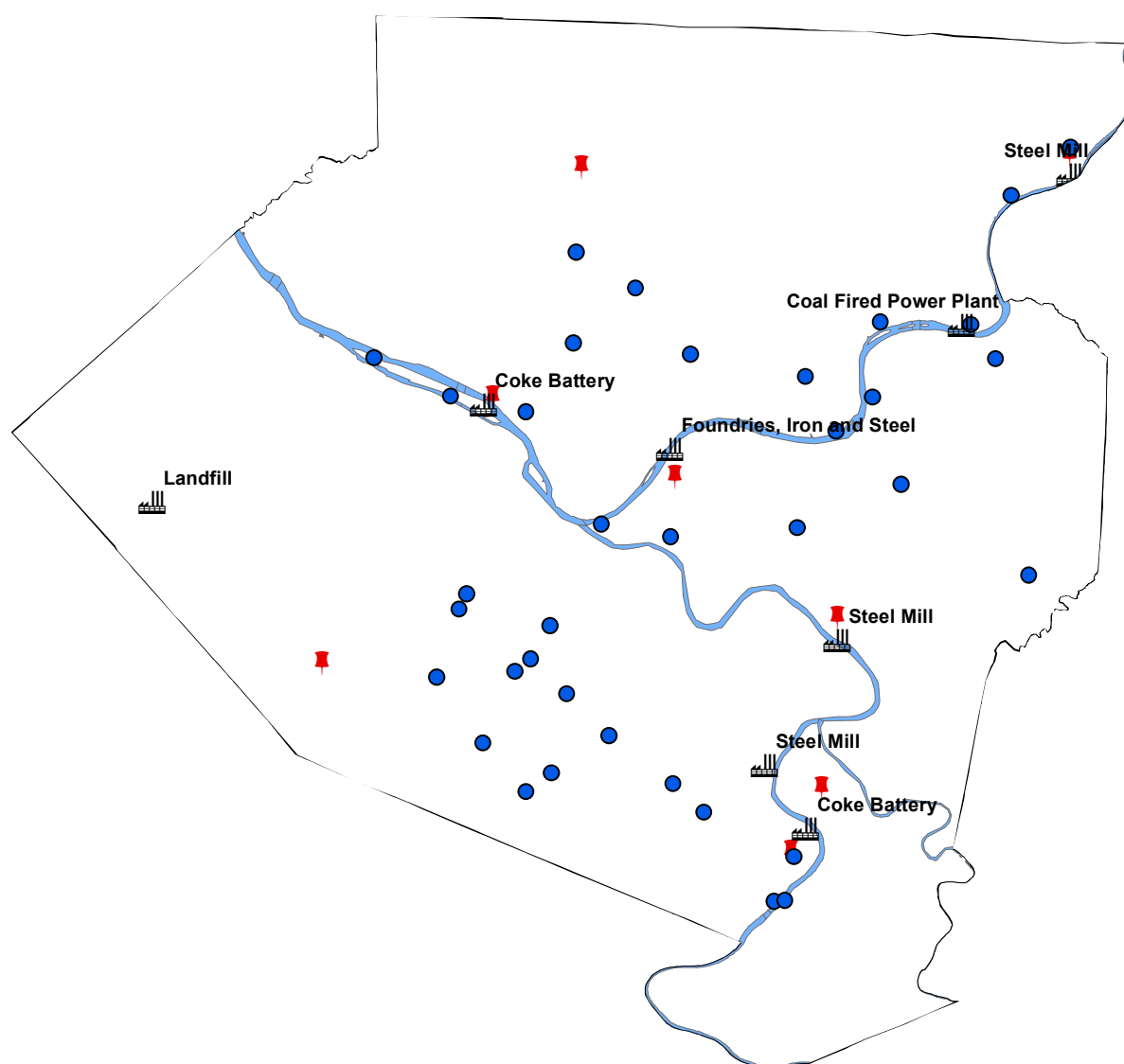
## Appendices



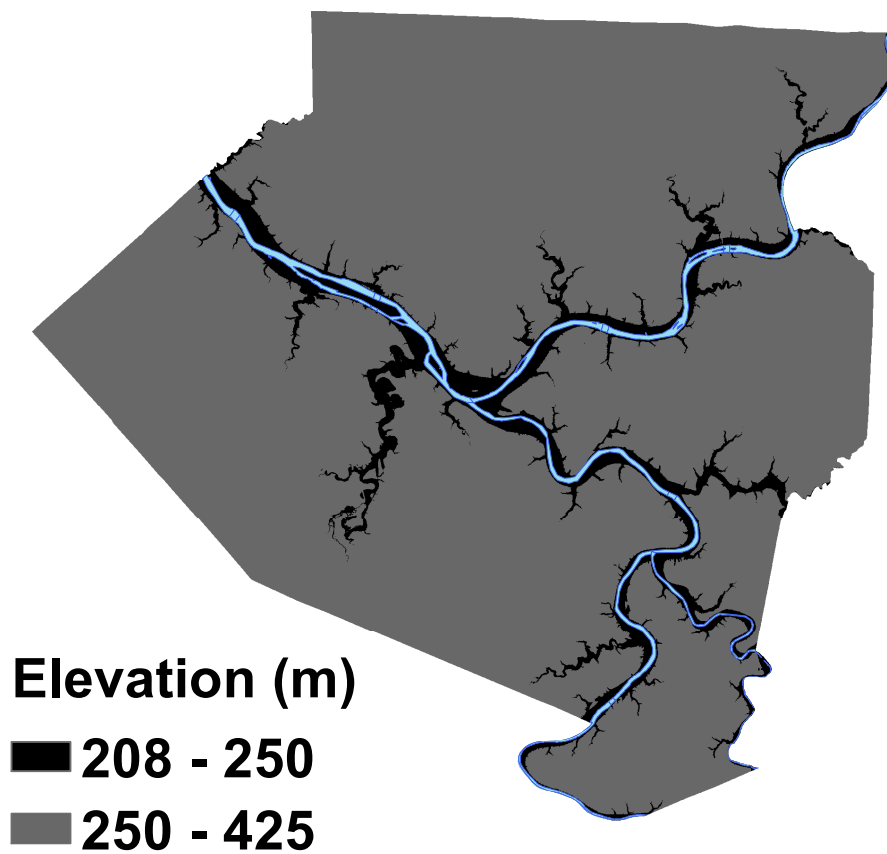
## Appendices

## **Appendix A**

### **Application of mobile sampling to investigate spatial variation in fine particle composition**



**Fig. A1.** Blue dots indicate mobile sampling locations. Red pins are stationary regulatory monitors. Industrial facilities with annual  $\text{PM}_{2.5}$  emissions greater than 50 tons are also shown.



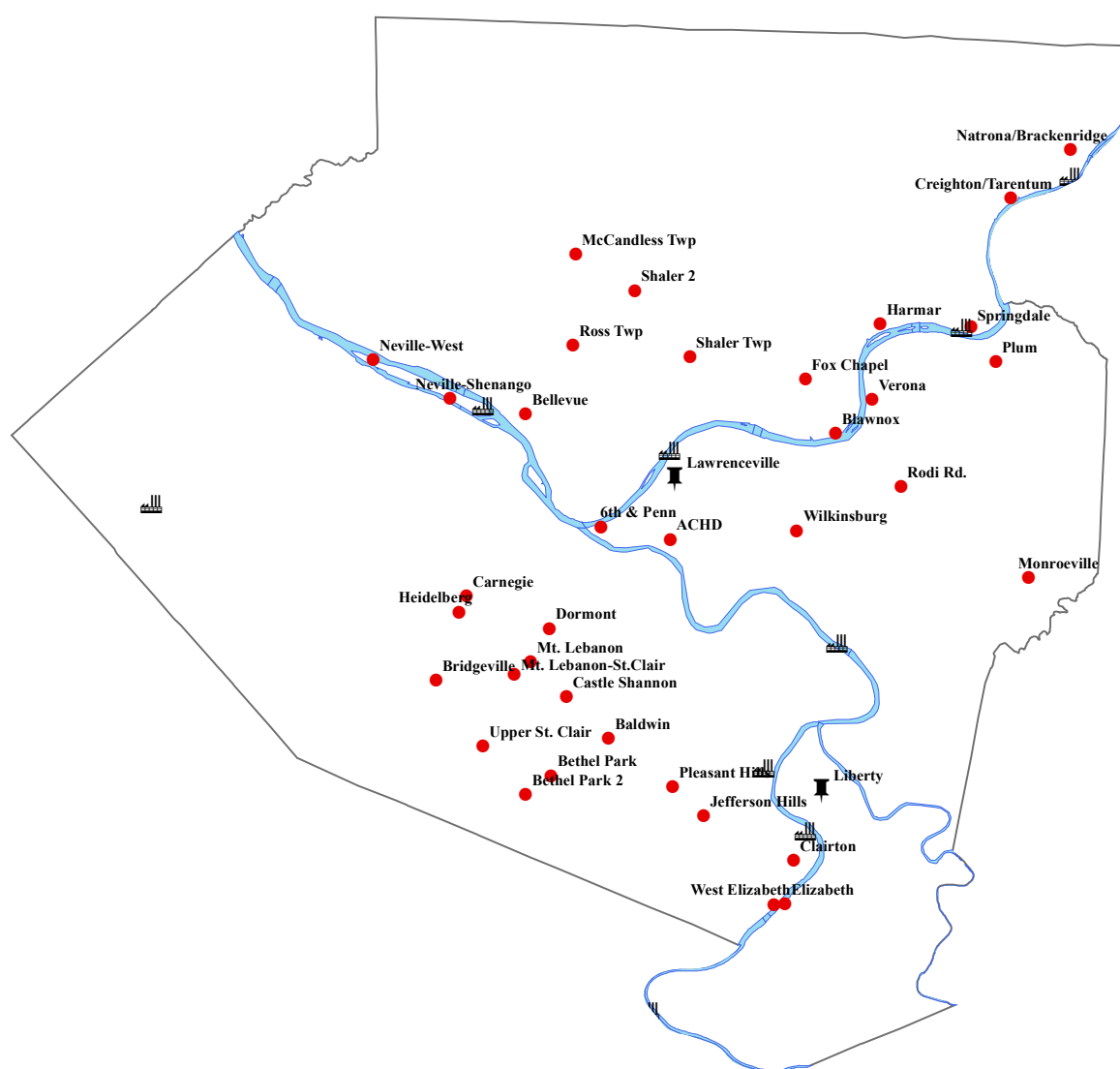
**Fig. A2.** Elevation map (meters above sea level) for Allegheny county, PA. Blue lines are the rivers. 250 m is the distinction value between river valleys and upland locations.

**Table A1.** Group comparison versus mobile sampling data at concurrent periods using the Mann-Whitney U test.

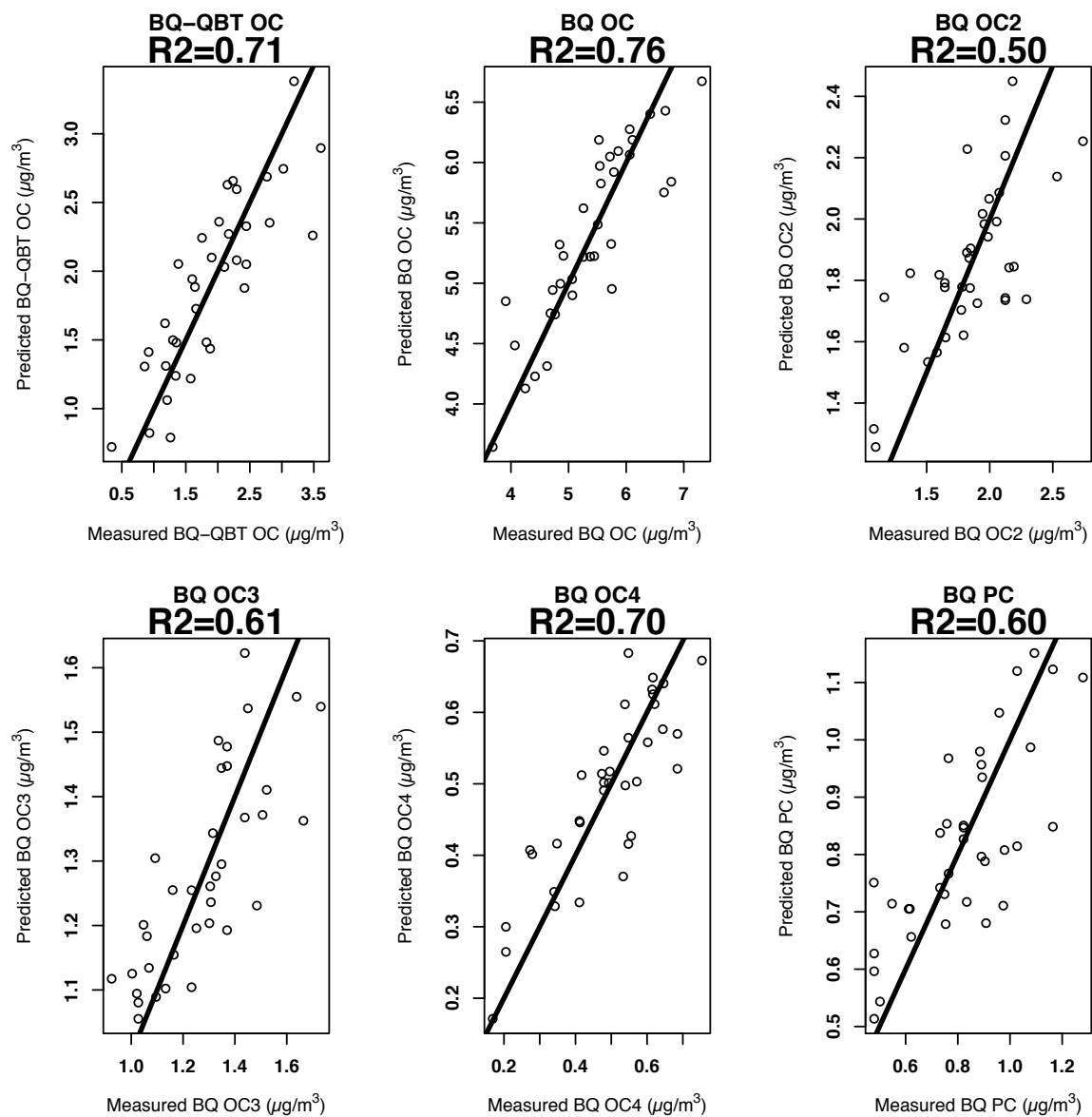
p value	Lawrenceville	South Allegheny	Clairton	Harrison	North Braddock	South Fayette	Avalon	North Park
PM <sub>2.5</sub>	0.07	0.4	0.26	0.27	0.92	0.11	0.41	0.28
Sulfate	0.78	0.62						
OC	0.39	0.18						
EC	0.02	0.06						
OC	0.39	0.18						
Zn	0.27	0.12						

## **Appendix B**

### **Urban organic aerosol exposure: spatial variations in composition and source impacts**



**Fig. B1.** Red dots indicate 36 mobile sampling locations. Industrial facilities with annual PM<sub>2.5</sub> emissions greater than 50 tons are shown as factory labels. Two regulatory monitors (Lawrenceville and Liberty) are displayed as black pins. The shapefile is downloaded from Pennsylvania Spatial Data Access (PASDA, 2017). Factory label and pins are from the software ArcGIS-10.3 (ESRI, Redlands, CA).



**Fig. B2.** LUR prediction versus measured values at 36 sampling sites. The black line is each

plot is the 1:1 line. LUR model  $R^2$  value is under the title.



**Table B1.** General description of predictors used in LUR models.

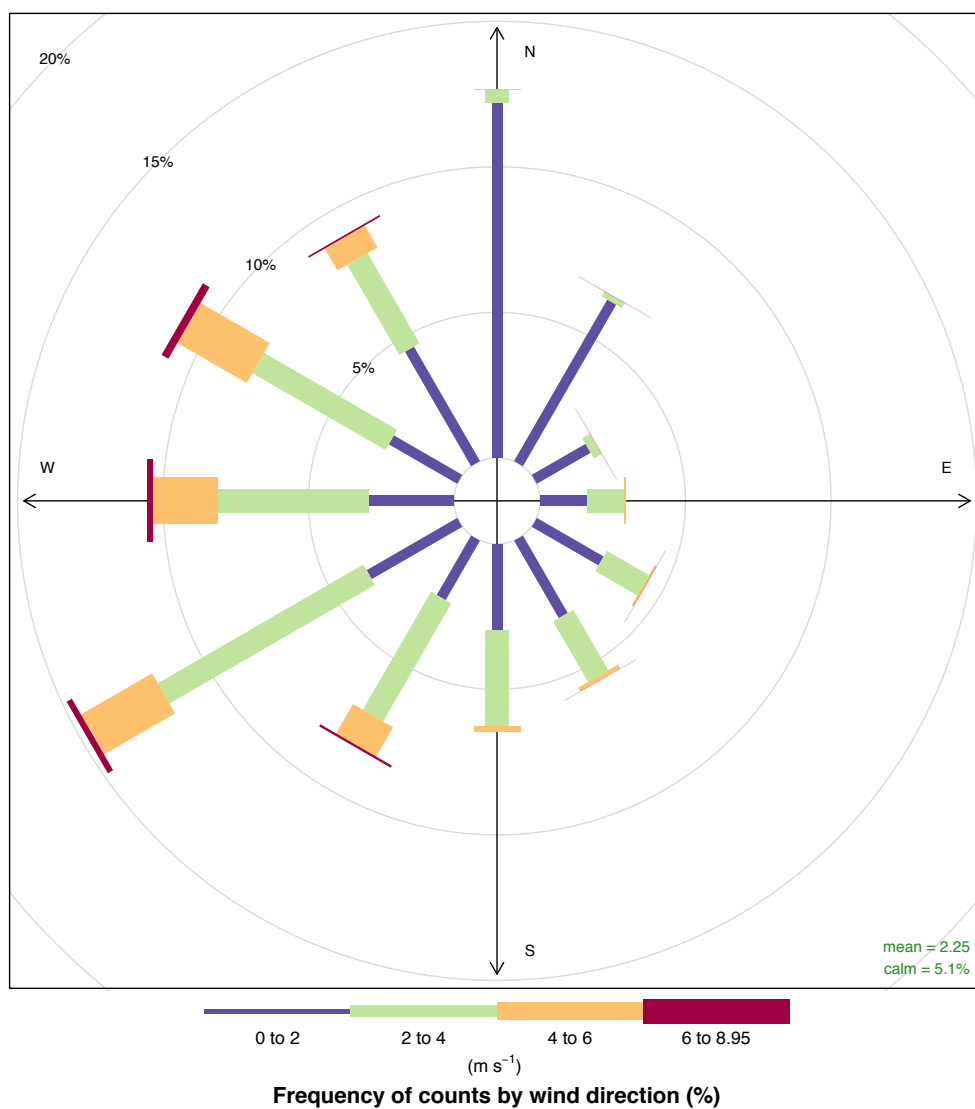
Major Category	ArcGIS Category	Units	Circular Buffer Radius (m)	Description	Variable Abbreviations
Traffic	Road length	m	25, 50, 100, 300, 500, 1000	Length of all roads	RDALL
				Length of major roads	RDMAJ
				Inverse distance to nearest road	DISTINVALL
	Inverse distance to the nearest road	m <sup>-1</sup>	NA	Inverse square distance to nearest road	DISTINVALL2
				Inverse distance to nearest major road	DISTINVMAJ
				Inverse square distance to nearest major road	DISTINVMAJ2
	Annual average daily traffic (AADT) on nearest road	veh/day	NA	AADT on nearest road	ALLAADT
				AADT on nearest major road	MAJAAADT
				Diesel truck AADT on nearest road	ALLDIESAADT
				Diesel truck AADT on nearest major road	MAJDIESALLAADT
	Vehicle density	veh m/day	25, 50, 100, 300, 500, 1000	Vehicle density on all roads	VEHDENSALL
				Vehicle density on major roads	VEHDENSMAJ
				Diesel truck density on all roads	TRKDENSALL
				Diesel truck density on major roads	TRKDENSMAJ
	Bus fuel consumption	kg fuel/day	25, 50, 100, 300, 500, 1000	Bus fuel consumption	BUSFC
	Rail length	m	25, 50, 100, 300, 500, 1000	Rail length	RAIL
	Traffic land use zoning	m <sup>2</sup>	100, 300, 500, 1000, 5000	Utility/transport land use area	LUUTr
Restaurant	Point density of restaurants	km <sup>-2</sup>	100, 300, 500, 1000, 5000	Number of restaurants per unit area	PointDe_Rest
	Inverse distance to the nearest restaurants	m <sup>-1</sup>	NA	Euclidean inverse distance to nearest restaurant	EucDistinv
			NA	Euclidean inverse square distance to nearest restaurant	EucDistinv2
Industry	Point density of industry sources (NEI)	km <sup>-2</sup>	1000, 1500, 3000, 5000, 7500, 10000, 15000, 20000, 30000	Number of facilities per unit area	PointDe_NEI
		lb km <sup>-2</sup>		Annual pollutant emissions per unit area	PointDe_NEI_Popu
	Inverse distance to nearest industrial source	m <sup>-1</sup>	NA	Euclidean inverse distance to nearest facility	EucDistinv
				Euclidean inverse square distance to nearest facility	EucDistinv2
	Pollution emission	lb	NA	Pollution emission at nearest corresponding facility	EucAllo
	Inverse distance weighted annual emissions	lb	NA	Inverse distance weighted annual emissions	IDW IDW2
	Industry land use zoning	m <sup>2</sup>	100, 300, 500, 1000, 5000	Industrial land use area	LUINDUS
Central reference site	Central reference site	µg/m <sup>3</sup>	NA	Corresponding PM measurements at the central	CSMPM
Elevation	Elevation	m	NA	Elevation	Elevation
Others	Land use zoning variables	m <sup>2</sup>	100, 300, 500, 1000, 5000	Residential land use area	LURES
				Commercial land use area	LUCOMM
				Agricultural land use area	LUAGRI
				Vacant/Forest land use area	LUVaFo
	Population		100, 300, 500, 1000, 5000	Number of inhabitants	POP
	Housing		100, 300, 500, 1000, 5000	Number of households	HOUS

**Table B2.** Classification of 36 sampling site based on three criteria (T: high traffic density; P: point source influenced; V (U): valley or upland).

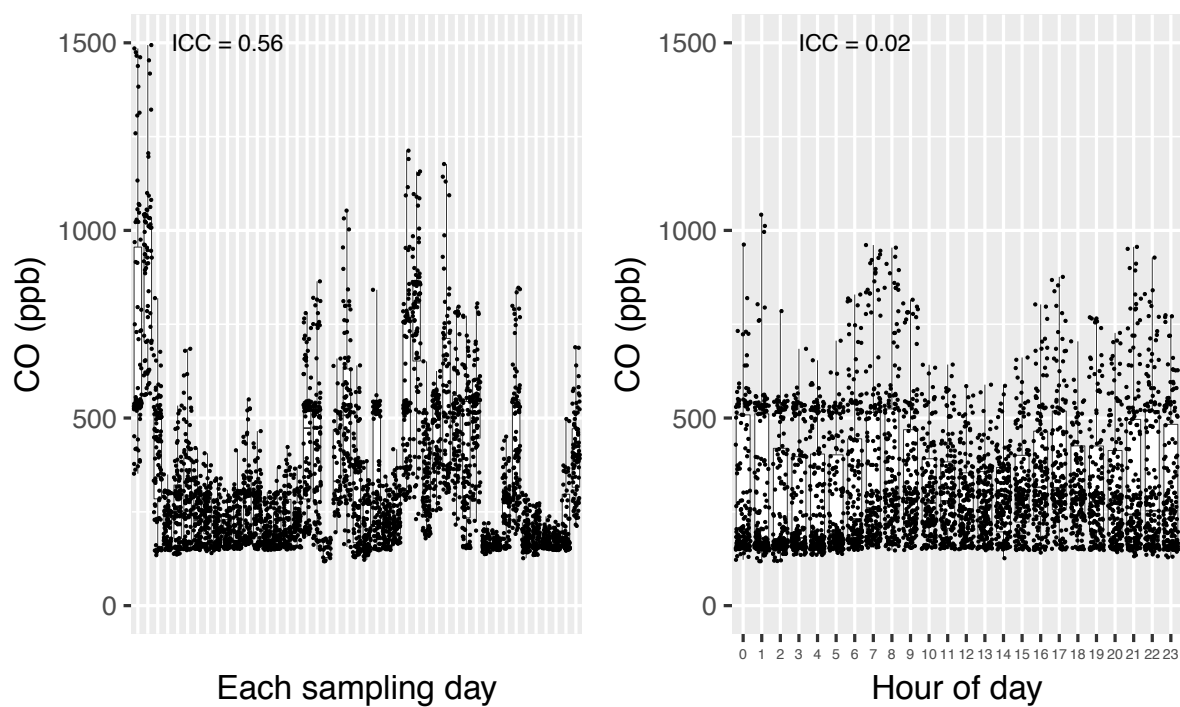
Sites	Strata	Corresponding number in Fig. 3.1
Natrona/Brackenridge	U+P+T	1
Creighton/Tarentum	V	2
Plum	U+P	3
Monroeville	U	4
Elizabeth	V+P	5
Springdale	V+P	6
West Elizabeth	V+P	7
Clairton	U+P	8
Harmar	V+T	9
Jefferson Hills	U+T	10
Neville-West	V	11
Rodi Rd.	U+T	12
Verona	V+T	13
Bethel Park 2	U	14
McCandless Twp	U+T	15
Pleasant Hills	U+T	16
Upper St. Clair	U	17
Bethel Park	U+T	18
Fox Chapel	U	19
Shaler 2	U	20
Blawnox	V	21
Bridgeville	U+P+T	22
Neville-Shenango	V+P	23
Baldwin	U	24
Ross Twp	U	25
Shaler Twp	U+P	26
Wilkinsburg	U+T	27
Heidelberg	V	28
Mt. Lebanon-St.Clair	U+T	29
Castle Shannon	U+T	30
Carnegie	V+T	31
Mt. Lebanon	U+T	32
Bellevue	U+T	33
Dormont	U+T	34
ACHD	U+P+T	35
6th & Penn	V+P+T	36

## **Appendix C**

### **Comparison of spatial and temporal variation of airborne pollutants using mobile and distributed sampling**



**Fig. C1.** Wind rose map based on yearlong meteorology data from an urban regulatory monitor at Lawrenceville, PA. The predominant wind direction is southwest.



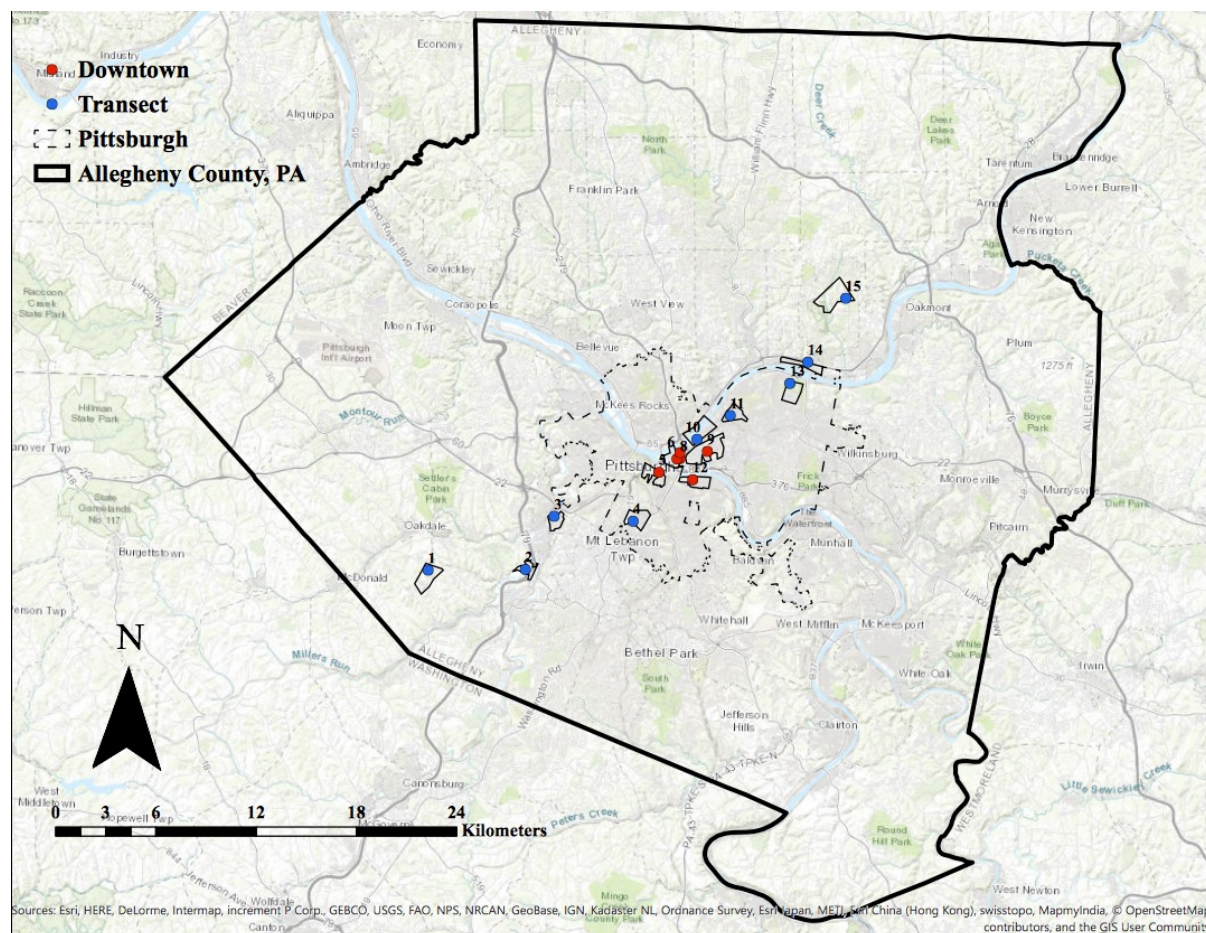
**Fig. C2.** High ICC value (left, 0.56) vs. low ICC (right, 0.02). High ICC means larger between group difference compared to within group variability. Low ICC means the opposite, i.e., smaller between group variability.

**Table C1.** ICC value for measured pollutants across all sites based on distributed monitor data. Data are grouped by each day. The interpretation for ICC is: less than 0.4 -- small or little between group difference, between 0.40 and 0.59 -- adequate between group difference, between 0.60 to 1 -- substantial between group difference (Cicchetti et al., 1994). Substantial ICC values are in orange. Adequate ICC ones are in blue.

Site ID	Site name	CO <sub>2</sub>	NO <sub>2</sub>	CO	PN	PM <sub>2.5</sub>
1	Chartiers	0.32	0.11	0.09	NA	0.44
2	Carnegie	0.10	0.13	0.15	NA	0.58
3	Beechview	0.28	0.50	0.35	0.54	0.35
4	Mt Washington	0.34	0.46	0.46	0.37	0.47
5	Penn	0.60	0.56	0.56	NA	0.52
6	Mellon	0.39	0.56	0.41	0.35	0.44
7	Church	0.54	0.42	0.46	0.21	0.16
8	Hill	0.38	0.49	0.37	0.33	0.55
9	Strip	0.18	0.15	0.20	NA	0.53
10	ACHD	0.44	0.53	0.46	NA	0.63
11	Southside	0.31	0.31	0.32	0.31	0.52
12	Zoo	0.27	0.29	0.24	NA	0.44
13	Aspinwall	0.38	0.46	0.34	NA	0.53
14	Shadyside	0.17	0.21	0.27	0.55	0.59
15	Supersite	0.26	0.30	0.24	0.54	0.65

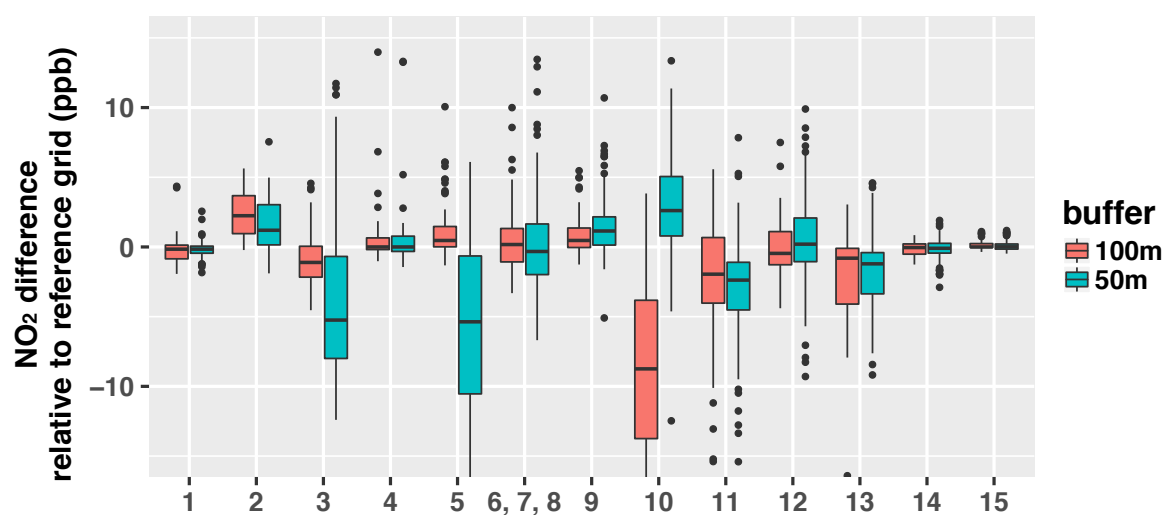
## **Appendix D**

### **Spatial variability of air pollution near monitors and exposure misclassification in an eastern US city**

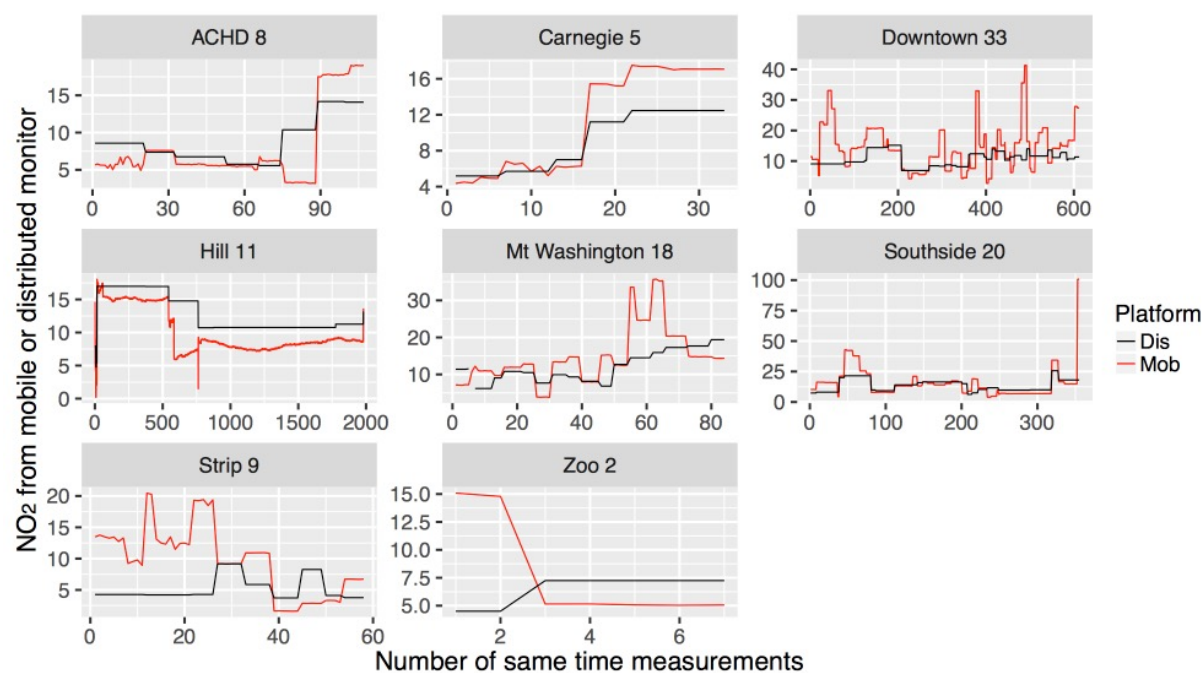


**Fig. D1.** Overview of the whole sampling campaign in Allegheny County, PA. Transect case study (blue) covers larger areas compared with Downtown one (red). Sites are arranged and labeled according to predominant southwestern wind direction. Mobile sampling is conducted in a  $\sim 1 \text{ km}^2$  surrounding area near the monitor at each site.





**Fig. D2.** Influence of grid box size (100 m vs. 50 m) on NO<sub>2</sub> relative difference pattern. The top and the bottom of the box represent the 75th and 25th percentile. The line inside the box is the median. The outer line extends to the most extreme concentrations not classified as outliers.



**Fig. D3.** Comparison of mobile (1 s) and distributed monitors (15 minutes) NO<sub>2</sub> measurements at same time and location. Numbers in the title mean the number of unique 15-minute window. X-axis is the number of occurrences when van is inside the 50-m grid where the monitor is located.