1-1-2018

# HUMAN EXPOSURE TO AIR TOXICS IN URBAN ENVIRONMENTS: HEALTH RISKS, SOCIODEMOGRAPHIC DISPARITIES, AND MIXTURE PROFILES

Zhuqing Xue

# HUMAN EXPOSURE TO AIR TOXICS IN URBAN ENVIRONMENTS: HEALTH RISKS, SOCIODEMOGRAPHIC DISPARITIES, AND MIXTURE PROFILES

by

Zhuqing Xue

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of Doctor of Philosophy

Major: Epidemiology

The University of Memphis

May 2018

## Dedication

This is dedicated to my grandmother Guihua Sun, my grandfather Yushen Rao, my parents Qiaobo Xue and Yuanqing Rao, my aunt Yuanqiao Rao, my uncle Xin Jin and all the people who supported me and believed in me. Thank you for seeing my potential and raising me up to more than I can be. Thank you all for your love to me. Without you I could never go beyond my limits and be who I am today.

"Believe in your infinite potential. Your only limitations are those you set upon yourself."

― Roy T. Bennett, The Light in the Heart

## Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Chunrong Jia, for his guidance and inspiration to me ever since I entered the PhD program in Epidemiology. Dr. Jia is a fount of knowledge, wisdom, and advice. His expertise in air toxics and innovative ideas on researches always bring me motivation to achieve more. Dr. Jia's support to me encouraged me to overcome so many obstacles to achieve what I have today. I feel truly fortunate to have the opportunity to learn from and work with him.

I would also like to thank my committee members, Drs. Xinhua Yu, Fawaz Mzayek, Meredith Ray and Anzhelika Antipova for their guidance and inspiring advices to me. In the similar vein, I am thankful to Drs. Hongmei Zhang, Yu Jiang, George Relyea, Kenneth Ward, Marian Levy, and Wilfried Karmaus for their contribution to my intellectual development at school of public health in University of Memphis. I highly appreciate the support from Mr. Jim Holt and Larry Smith to the REACT monitoring program and studies.

Finally, I would love to thank all friends I made during graduate school. We shared some much in common and you know what I mean! Cheers to our friendship and what we have achieved along the journey!

## Abstract

Zhuqing Xue. PhD. The University of Memphis. May 2018. Human Exposure to Air Toxics in Urban Environments: Health Risks, Sociodemographic Disparities, and Mixture Profiles. Major Professor: Chunrong Jia, PhD.

Exposure to air toxics in urban environments may be of significant health concern because populations and emission sources are concentrated in the same geographic area. The overall objective of this study is to characterize the sources, variations, and mixture profiles of ambient air toxics in urban environments, and examine the sociodemographic disparities in exposures to air toxics in a typical U.S. metropolitan area.

A model-to-monitor comparison was performed to evaluate the validity of modeling air toxics data using national datasets. Modeled concentrations in the 2011 National-scale Air Toxics Assessment (NATA) moderately agreed with monitoring measurements, and a sizable portion showed underestimation. Results warranted the need for actual monitoring data to conduct air toxics exposure assessment.

Air toxics samples were collected in 106 census tracts in the Memphis area in 2014, and samples were analyzed for 71 volatile organic compounds (VOCs). Ambient VOC levels in Memphis were generally higher than the national averages in urban settings, but were mostly below the reference concentrations (RfCs). Factor analysis identified 5 major sources: manufacturing processes, vehicle exhaust, industrial solvents, refrigerants, and gasoline additives. The major non-cancer risks were from neurological, respiratory, and reproductive/developmental effects. The cumulative cancer risk was $5.9\pm 3.3 \times 10^{-4}$, with naphthalene and benzyl chloride as risk drivers.

Sociodemographic disparities in cancer risks were examined by regressing cancer risks against socioeconomic, racial, and spatial parameters at the census tract level. We conducted separate disparity analyses using modeling data from 2011 NATA and our air toxics monitoring data. Analysis using modeling data showed strong sociodemographic disparities but that using monitoring data did not show. The discrepancy brought cautions for use of modeling air pollution data in environmental disparity research.

We further assessed exposure to VOCs mixtures in five typical microenvironments (MEs): home, office, vehicle cabin, gas station, and community outdoors. The multivariate analysis of variance and pairwise analysis showed VOC profiles were distinguishingly different among MEs. The classification of profiles was achieved using the random forest. We anticipate wide applications of exposure profiles in epidemiologic research of exposure to air toxic mixtures.

## Table of Contents

## List of Tables

## List of Figures

## List of Abbreviations and Acronyms

HAPs - Hazardous Air Pollutants

LOQ - Limit of Quantitation

REACT - Reducing Exposure to Airborne Chemical Toxics

GC/MS - Gas Chromatography/ Mass Spectrometry

VOCs - Volatile Organic Compounds

TOSHI - Target Organ Specific Hazard Index

CAA - Clean Air Act

USEPA - U.S. Environmental Protection Agency

PM 2.5 - Particulate Matter 2.5 microns

PM 10 - Particulate Matter 10 microns

PAHs - Polycyclic Aromatic Hydrocarbons

NEI - National Emission Inventory

TRI - Toxic Release Inventory

NATA - National-scale Air Toxics Assessments

WHO - World Health Organization

HEI - Health Effects Institute

NIOSH - National Institute for Occupational Safety and Health

ATSDR - Agency for Toxic Substances & Disease Registry

EJSCREEN - Environmental Justice Screening and Mapping Tool

EJ - Environmental Justice

AQS - Air Quality System

ΔM - difference between the modeled and observed concentrations

DF - Sample Detection Frequency

UR – Unit Risk

RfC – Reference Concentration

BTEX - Benzene, Toluene, Ethyl benzene, and Xylenes

LASSO - Least Absolute Selection Operator

RF – Random Forest

PCA - Principle Component Analysis

PMF - Positive Matrix Factorization

CART - Classification and Regression Tree

MDL - Minimum Detection Limits

NPL - National Priorities List

RMP - Risk Management Plan

TSDF - Treatment Storage and Disposal

# Chapter 1

# Introduction

## 1.1 Air toxics

Air pollutants are various forms of agents in the air which contaminate the air and lead to adverse health effect when their level elevates above threshold (Godish, 2014; WHO, 2017a). Ambient air pollutions can be caused by natural processes such as volcano explosions, forest fires etc. (Kampa & Castanas, 2008). However, anthropogenic activities play the major role in the formation of air pollution. Emissions from combustion of fossil-fuel, industrial manufacturing, transportation, and households increase concentrations of particulate matter, persistent organic pollutants, heavy metals and gaseous-phase pollutants in the ambient air (Godish, 2014; Kampa & Castanas, 2008). These primary pollutants form secondary pollutants such as haze and tropospheric ozone through chemical and photochemical reactions among primary pollutants (Godish, 2014). The Clean Air Act (CAA) amended in 1990 requires regulation of 6 criteria air pollutants and 187 air toxics. Air toxics include volatile organic compounds (VOCs), semi-VOCs (SVOCs), and heavy metals in particular matter (PM) (McCarthy, O'Brien, Charrier, & Hather, 2009; U.S.EPA, 2009). Lead is regulated as both a criteria pollutant and an air toxic (Godish, 2014; Suh, Bahadori, Vallarino, & Spengler, 2000).

Air toxics are released to the ambient air from numerous sources. U.S. EPA summarized eight air toxics emission source sectors (U.S.EPA, 2017b): (1) agriculture: e.g., fertilizer application; (2) dust: e.g., pave road dust; (3) fires: e.g., wildfires; (4) fuel combustion: e.g., biomass, coal, natural gas combustion for electric generation ; (5) industrial processes: e.g.,

chemical manufacturing; (6) miscellaneous: e.g., commercial cooking; (7) mobile: e.g.; aircraft, commercial marine vessels, non-road equipment and on-road vehicles etc.; and (8) solvents: e.g. dry cleaning, degreasing. Currently, the U.S. EPA estimates air toxics emissions of air pollutions majorly on the basis of the National Emission Inventory (NEI), which is updated every three years and based on data collected from state, local, and tribal air agencies. The NEI covers five major emission sources (U.S.EPA, 2017f): (1) point sources: large groups of stationary facilities that are regulated for emission level, e.g., power plants and airports; (2) nonpoint sources: small magnitude individual stationary source, e.g., gas and service stations, dry cleaners, residential cooking and heating; (3) on-road sources: fuels combustions from automobiles during transportation or road-work on highways, streets and roads; (4) non-road sources: aircrafts, commercial marine vessels, rail transport vehicles, construction equipment, lawn and garden equipment; (5) "event" sources: wildfires and prescribed burnings. Additionally, another important emission inventory related to air emission and NEI is Toxic Release Inventory (TRI). TRI tracks pollutants emitted to air, water or in land disposal (U.S.EPA, 2017i). Therefore, TRI had overlap with NEI in air emissions. For emissions from some stationary facilities (about 10%), TRI is the only data source for NEI (U.S.EPA, 2002b).

Valid and representative data of hazardous air pollutants (HAPs) are required to evaluate emission compliance, air quality attainment, and population health risks. Chronic and acute exposure to HAPs may cause damages to multiple human organs (Kampa & Castanas, 2008), including respiratory (Chen, Salam, Eckel, Breton, & Gilliland, 2015), nervous (Novaes et al., 2010; Sunyer et al., 2015), circulatory (Brook et al., 2004), reproductive (Lewtas, 2007), immune (Nadeau et al., 2010), digestive (Kelishadi & Poursafa, 2011), and urinary systems (Jarup, 2003). The U.S. Environmental Protection Agency (EPA) aimed to reduce HAP emissions by 75% of

the 1993 level to meet the requirement by the Government Performance and Results Act. EPA has been working with state, local, and tribal air pollution control agencies to measure ambient HAP concentrations. However, the current monitoring efforts are inadequate for increasingly refined health and climate studies, and modeling programs are then developed to estimate exposures at high temporal and spatial resolutions (Jerrett et al., 2005).

EPA initiated the National-scale Air Toxics Assessments (NATA) in 1996 to serve as a geographical extension of the existing air monitoring network. NATA is designed to inform decision-making, e.g., prioritize pollutants and sources, identify locations for investigation, and design monitoring programs (USEPA, 2012). NATA models HAP concentrations at geographic resolutions down to the census tract level. The high spatial resolution data have a number of environmental applications. Environmental epidemiology studies have used NATA data to explore association between HAP exposure and health endpoints such as respiratory disease (deCastro, 2014; Stoner, Anderson, & Buckley, 2013), autism spectrum disorder in children (A. L. Roberts et al., 2013), and school performance (Clark-Reyna, Grineski, & Collins, 2016; Grineski, Clark-Reyna, & Collins, 2016). The cancer risk estimates in NATA often serve as bases for addressing environmental justice issues (Apelberg, Buckley, & White, 2005; Chakraborty, 2012; Grineski, Collins, & Chakraborty, 2013; James, Jia, & Kedia, 2012; Linder, Marko, & Sexton, 2008; Pastor, Morello-Frosch, & Sadd, 2005; Rice et al., 2014; Wilson et al., 2015). The methodology and data are also used to model population exposure (Ozkaynak, Palma, Touma, & Thurman, 2008), predict future exposures (Cook et al., 2007), estimate excess risks (Woodruff, Wells, Holt, Burgin, & Axelrad, 2007), and establish an emission-to-intake relationship (Marshall, Teoh, & Nazaroff, 2005).

Evaluating NATA model performance is imperative for its numerous applications. NATA modeling uses conservative assumptions that potentially lead to overestimation (U.S.EPA, 2016g); however, some comparison studies gave the opposite results (Garcia et al., 2014; Logue, Small, & Robinson, 2011; Lupo & Symanski, 2009). A few independent evaluation studies used local-scale monitoring in California (Garcia et al., 2014), Pittsburgh, Pennsylvania (Logue et al., 2011), Detroit, Michigan (George et al., 2011), Texas (Lupo & Symanski, 2009), and South Baltimore, Maryland (D. C. Payne-Sturges, Burke, Breysse, Diener-West, & Buckley, 2004). These model-to-monitor comparisons are often limited in terms of number of chemicals and geographic areas. EPA has conducted limited evaluations and encourages more studies (George et al., 2011; Rosenbaum et al., 1999).

The 2011 NATA yielded the latest available database that contains concentrations, exposures, and cancer and non-cancer risks for 180 HAPs, as well as their contributing sources. There has been no independent evaluation of 2011 NATA, although EPA has made limited model-to-monitor comparisons for selected compounds (U.S.EPA, 2016a). Methodologically, EPA used multiple comparison measures, e.g., linear regressions, factor of 2, and absolute biases; however, they often give inconsistent results. Measurement uncertainty is not considered in previous comparisons, which may lead to huge bias as many modeled concentrations are far below the detection limits. These limitations call for a systematic approach for model-to-monitor comparisons.

### 1.2 Health risks from exposure to air toxics

According to the most recent estimates from World Health Organization (WHO) in 2012, three million deaths, 5.4% of total deaths in the world, were due to ambient air pollution (WHO, 2017b). Chronic and acute exposure to air toxics may cause damages to multiple human organs

(Kampa & Castanas, 2008), including respiratory (Chen et al., 2015), nervous (Novaes et al., 2010; Sunyer et al., 2015), circulatory (Brook et al., 2004), reproductive (Lewtas, 2007), immune (Nadeau et al., 2010), digestive (Kelishadi & Poursafa, 2011), and urinary systems (Jarup, 2003). Environmental epidemiology studies have extensively explored the association between criteria air pollutants and various adverse health outcomes. More and more studies have been focusing on air toxics. Both modeled and monitored concentrations of air toxics have been utilized to: assess the association between adverse health effect and exposure to air toxics; and characterize the health risks.

Cancer risk is a great concern when addressing adverse effects from exposure to air toxics. Among the 187 air toxics, formaldehyde, benzene, acetaldehyde, 1, 3-butadiene, chloroform, naphthalene and most of PAHs are major contributor to cancer risks (Bostrom et al., 2002; M. M. Loh, J. I. Levy, J. D. Spengler, E. A. Houseman, & D. H. Bennett, 2007). Particularly, formaldehyde, benzene, acetaldehyde are the three leading cancer risk drivers based on 2011 NATA modeling results (U.S.EPA, 2016a). Formaldehyde was reported to have effects on modification of miRNA expression which may lead to lung cancer, breast cancer, and leukemia (Lu et al., 2005; Rager, Smeester, Jaspers, Sexton, & Fry, 2011). Additionally, exposure to formaldehyde was linked with nose and throat cancer. Long-term high level exposure to benzene was linked to acute myelogenous leukemia (AML) (Galbraith, Gross, & Paustenbach, 2010). Acetaldehyde from ambient air is associated with nose cancer (Y. Zhou, Li, Huijbregts, & Mumtaz, 2015). Early childhood exposure high levels of 1,3-butadine was linked to acute lymphocytic leukemia (Symanski et al., 2016). Exposure to mixture of PAHs is associated with elevated cancer risks targeting skin, lung, bladder, and gastrointestinal tracts (Kim, Jahan, Kabir, & Brown, 2013).

Non-cancer risks including respiratory disease, neurotoxicity, negative birth outcomes, renal toxicity, and cardiovascular disease have been linked with exposure to air toxics. Air toxics, such as benzene, 1,3 –butadiene and PAHs, have short term respiratory adverse effect including mucosal irritation, cough, wheeze, and shortness of breath (Goldizen, Sly, & Knibbs, 2016). VOCs, especially formaldehyde, and PAHs from traffic emissions are relevant to asthma exacerbation and development. However, epidemiologic evidence for causal relationship between air toxics and asthma is not sufficient (Delfino, 2002). Emerging evidence has been found targeting the association between air pollution and neurotoxicity (Block et al., 2012). Air toxics (e.g. VOCs, PAHs, and heavy metals) are associated with adverse effects on central nervous system (CNS), which may lead to autism spectrum disorder (ASD) (Volk, Hertz-Picciotto, Delwiche, Lurmann, & McConnell, 2011), delayed intelligence and cognitive development in children (Calderon-Garciduenas et al., 2008; S. Q. Wang et al., 2009), Parkinson's disease (Calderon-Garciduenas et al., 2002), and Alzheimer's disease (Calderon-Garciduenas et al., 2004). Exposure to air pollution is also associated with increased risk of negative birth outcomes including childhood mortality, intrauterine growth restriction (IUGR), premature birth, low birth weight (LBW), and birth defects (Lewtas, 2007; Sram, Binkova, Dejmek, & Bobak, 2005; Stieb, Chen, Eshoul, & Judek, 2012). PAH has been linked to IUGR (Dejmek, Solansky, Benes, Lenicek, & Sram, 2000; Sram et al., 2005). Exposure to heavy metals, such as cadmium and lead, is associated with renal dysfunction and failures (Damek-Poprawa & Sawicka-Kapusta, 2003; Jarup, 2003; Kampa & Castanas, 2008). Additionally, lead is also associated with hypertension, coronary heart disease (CHD), stroke, and peripheral arterial disease (Uzoigwe, Prum, Bresnahan, & Garelnabi, 2013).

Air toxics adversely affect human health in various ways not only at the individual level as it targets different human organs, but also at the community level due to the spatial and temporal variations of its concentrations. Only a few previous studies were focused on the spatial and temporal patterns of air toxic concentrations in areas clustering various emission sources based on ambient monitoring program. A study conducted in Camden, NJ pointed out the concentrations of ambient air toxic, e.g., VOCs, vary differently around different stationary emission sources in different neighborhoods (Zhu et al., 2008). Another study in Oakland, CA suggested that spatial variation of air toxics in urban areas was due to different traffic density (Fujita, Campbell, Arnott, Lau, & Martien, 2013). Concentrations of air toxics are expected to be higher in cold seasons than warm seasons due to more emissions from heating and lower air exchange rate (Rehwagen, Schlink, & Herbarth, 2003). However, although previous studies found concentrations of air toxics vary differently between cold and warm seasons, no consistent patterns were found (Fujita et al., 2013; C. Jia, Batterman, & Godwin, 2008b).

More studies are in need to characterize the health risks from air toxics. Spatial and temporal patterns of air toxic concentrations in areas where there are multiple emission sources need to be further studied to have a better understanding so that more efficient regulations could be applied addressing health risks from exposure to air toxic.

### 1.3 Exposure to mixture of air toxics

Exposure to air pollutants in real life is rarely to single pollutants but rather to mixtures. Mixtures of air pollutants reflect the integration of many sources, emission constituents, or ongoing photochemical processes in the atmosphere. Air pollutants exist in form of complex mixtures in indoor and ambient environments. Symptoms and diseases may arise from multipollutant exposure; however, the interactions between pollutants are still unclear. Both the

scientific community and regulatory agencies have been shifting from the traditional single-pollutant approach toward a multipollutant approach to quantify the health consequences of air pollution mixtures (Dominici, Peng, Barr, & Bell, 2010). The Health Effects Institute (HEI) has made research of health effects of the air pollution mixture a top priority in its 2010-2015 strategic plan (HEI, 2010). The U.S. Environmental Protection Agency (EPA) has initiated the Clean Air Program to address multipollutant issues(U.S. EPA, 2009a). The U.S. National Institute for Occupational Safety and Health (NIOSH) published a "Mixed Exposure Research Agenda" that suggested surveillance, evaluation and research, and controls and interventions for air pollutant mixtures(NIOSH, 2005). The U.S. Agency for Toxic Substances & Disease Registry (ATSDR) was engaged in a "Chemical Mixtures Program" to identify mixtures, estimate the joint toxic effects and develop new methodologies for evaluating the health effects of mixtures (De Rosa, El-Masri, Pohl, Cibulas, & Mumtaz, 2004).

Several terms are used for different aspects of mixture exposure. Regarding the exposure agents, "mixture" probably is the earliest and most widely used term, e.g., chemical mixture and complex mixture were used in early literature (Samet & Speizer, 1993). The ATSDR manual further categorized mixtures as (ATSDR, 2004): (1) intentional mixture, which are manufactured products, such as pesticide formulations, gasoline, or laundry detergent; (2) generated mixture, which are byproducts of such processes as smelting, drinking water disinfection, fuel combustion, and cigarette smoking; and (3) coincidental mixture which consist of unrelated chemicals from different sources, deposited separately at the site, but having the potential to reach the same "receptor population" by their presence in or migration into the same medium (commonly groundwater), or through a combination of media and pathways. To describe exposure pathway and route, mixed exposure (NIOSH, 2005), co-exposure and multipollutant

exposure are interchangeably used. Recently, “multipollutant” was used for describing methodology and science of chemical mixture exposure (Greenbaum & Shaikh, 2010).

Human health risks are affected by multiple pollutants, and may be enhanced by pollutant interactions. Although incompletely understood, exposures to mixtures can cause additive and synergistic effects (De Rosa et al., 2004). For example, VOC mixtures at sufficiently high concentrations and over a range of exposure periods have been demonstrated to induce adverse health effects in both indoor (Mendell, 2007; Rumchev K, Brown H, & Spickett J, 2007) and occupational (Hansen, De Rosa, Pohl, Fay, & Mumtaz, 1998; Lippy & Turner, 1991) settings.

Therefore, identifying common air pollutant mixtures, determining the mixture composition and categorizing the exposure to these mixtures into different exposure profiles based on exposure locations and emission sources of mixtures can help us gain better understanding and estimation of health risks from exposure to air toxic mixtures.

### 1.4 Health disparities from exposure to air toxics

Health disparity from exposure to air pollution has been reflected on individual-level vulnerability and community-level vulnerability (D. Payne-Sturges & Gee, 2006). At the individual level, empirical evidence on health disparities demonstrated that racial minorities and those with lower socioeconomic status tend to experience disproportionate adverse health effect from exposure to air pollution (Chakraborty, 2012; Chakraborty, Collins, & Grineski, 2017; Chakraborty, Collins, Grineski, & Maldonado, 2017; C. R. Jia, James, & Kedia, 2014; Linder et al., 2008; Morello-Frosch & Jesdale, 2006; Tyrrell, Melzer, Henley, Galloway, & Osborne, 2013). For instance, minority residents, compared to non-Hispanic whites in the United States, have higher risk perception of environmental risks and, specifically, more concerns on air

pollutions (Macias, 2016). Individuals with low income are more likely to move into more polluted areas with worse housing condition due to lower living and housing cost. These individuals cannot easily relocate even when they realize more emission facilities are moved to where they live (Wilkinson & Pickett, 2006). At the community level, communities with social disadvantages are more likely to be exposed to air toxics. For example, emission facilities tend to locate in disadvantaged area because the communities there lack of political power (Perlin, Sexton, & Wong, 1999). In general, residents living in minority's communities are more likely to 1) lack knowledge of health, 2) experience more social stress like crime, drugs and poverty exist, 3) have less insurance coverage, and 4) have less access to appropriate healthcare (Mayberry, Mili, & Ofili, 2000).

The two critical elements to address the health disparity from exposure to air pollution are: (1) Geographical extent and spatial resolution of measurement of both exposure to air pollution and sociodemographic indicators (e.g., race, ethnicity, and income); and (2) Analytical methodology of assessing the association between exposure and sociodemographic indicators (Chakraborty, Collins, & Grineski, 2017; Chakraborty, Maantay, & Brender, 2011; Maantay, 2007).

The measurement of exposure to air pollution from multiple emission sources in health disparities studies falls into two categories including: (1) Proximity or distance to emission sources, e.g., Utilizing circular buffer rings , constructed via geographic information system (GIS), around emission sources as a geographic boundary for population who are considered to be exposed to air pollution (Perlin, Wong, & Sexton, 2001); (2) modeled air pollution level and risk based on emission data of air pollutions and meteorological conditions, e.g., NATA modeled air toxic concentrations and cancer risks etc. (Chakraborty, Collins, & Grineski, 2017). The

primary data source of exposure to air toxics is from parameters based on ambient monitoring concentrations of air toxics. However, due to geographical limitation, secondary data on emissions of air pollution and related risk from NATA and TRI become major data sources of exposure to air toxics for health disparities studies. Particularly, U. S. EPA has developed the Environmental Justice Screening and Mapping Tool (EJSCREEN) providing high-resolution environmental and demographic data around United States for health disparities studies and environmental justice (EJ) studies(U.S.EPA, 2017h). Eleven prioritized environmental indexes (U.S.EPA, 2017c) for EJ studies were also listed as following:

- National Scale Air Toxics Assessment Air Toxics Cancer Risk
- National Scale Air Toxics Assessment Respiratory Hazard Index
- National Scale Air Toxics Assessment Diesel PM (DPM)
- Particulate Matter (PM2.5)
- Ozone
- Lead Paint Indicator
- Traffic Proximity and Volume
- Proximity to Risk Management Plan Sites
- Proximity to Treatment Storage and Disposal Facilities
- Proximity to National Priorities List Sites
- Proximity to Major Direct Water Dischargers

The sociodemographic measurement can be categorized into individual level and composite level of socioeconomic attributes (SEA) and racial/ethnic attributes. Individual level measurement include: (1) census; (2) sociodemographic indicators designed for EJSCREEN (U.S.EPA, 2017g): percent Low-income, percent minority, less than high school education, linguistic isolation, individuals under age 5, and individuals over age 64; (3) delayed rent or mortgage in  past year and food stamp program enrollment (Shmool et al., 2014); (4) social processes (D. Payne-Sturges & Gee, 2006). Composite level measurement include: (1) segregation indices: dissimilarity index, isolation index, delta, relative cluster, and Townsend

index (Rice et al., 2014); (2) HOUSE index (Bang et al., 2014). (3) Principle component of multiple sociodemographic indicators (C. R. Jia et al., 2014).

Three major analytical methods have been applied into environmental justice studies addressing health disparities from exposure to air pollutions. These methods include: (1) proximity analysis using pair-wised Kolmogorov-Smirnov test (Abel, 2008; Perlin et al., 2001) (2) chi-square tests (Linder et al., 2008); (3) regression analysis: linear regression (OLS), quantile regression (James et al., 2012), spatial regression (Simultaneous autoregressive models) accounting for spatial auto correlation (Anselin, 2014; Chakraborty, Collins, & Grineski, 2017; Raddatz & Mennis, 2013). Particularly, regression analysis is the most common approach for addressing health disparities (Gilbert & Chakraborty, 2011). These analytical approaches have advantages and disadvantages. Proximity analysis assumes that proximity to emission sources has same negative linear relationship to level of exposure when subjects were exposed to the emission sources from different direction but this relationship varies due to different environmental conditions, e.g., wind direction and speed, meteorological conditions (Chakraborty et al., 2011). Traditional linear regression, another commonly used approach, has the assumption of independence of errors. However, in most cases, spatial autocorrelation, reflecting neighboring effect geographically, violates this assumption. In contrast, spatial regression, e.g., geographically weighted regression is more appropriate in assessing sociodemographic disparities in exposure to air toxics if spatial autocorrelation was detected.

Previous studies have reported that minorities and communities of lower socioeconomic status were more likely to live close to emission sources (Stuart, Mudhasakul, & Sriwatanapongse, 2009) exposing to elevated level of air toxics and thus having increased health risks (Abel, 2008; Chakraborty, Collins, & Grineski, 2017; Grineski, Collins, Chakraborty, &

McDonald, 2013; C. R. Jia et al., 2014). For example, a study in Texas reported that census tracts with higher proportion of Hispanics and socioeconomic disadvantage in Houston area showed higher cancer risks burden from exposure to air toxics (Linder et al., 2008). Another study in South Carolina strengthened the evidence that non-white population and communities with low-income had higher cancer risks from exposure to air toxics (Wilson et al., 2015). However, these studies all utilized modeled air toxic concentrations to investigate the health disparities in health risks from exposure to air toxics. Their modeled air toxics concentrations and related health risks were from National Air Toxic Assessment (NATA) which was conducted by United States Environmental Protection Agency (U.S.EPA) to have an overview of national level of air toxics and to prioritize certain air toxics for regulation and reduction of air toxics emissions (U.S.EPA, 2016g). Particularly, NATA was announced as not appropriate for characterizing and comparing risk at local level (U.S.EPA, 2015d). However, due to high expenses and tremendous efforts acquired for monitoring at census tract level, census tract level monitoring measurement were usually surrogated with modeled estimates in disparity studies at local level. Nevertheless, several model assessment studies have reported that NATA estimated air toxics concentrations generally underestimate the air toxic concentrations (Garcia et al., 2014; Logue, Small, Stern, Maranche, & Robinson, 2010; Lupo & Symanski, 2009). Thus the uncertainty of modeled estimates might affect the true association between sociodemographic factors and exposure to air toxics or related health risks. So far, no previous studies have utilized monitored air toxic measurement to address the sociodemographic disparities in exposure to air toxics or related health risks. Therefore, evaluating the sociodemographic disparities with the monitor measurement might give us an alternative prospect on this issue.

### 1.5 Emission Sources of Air Toxics in the Memphis area

Shelby County, TN have heavy traffic burden including automobile traffic, railway transportation, barge traffic on Mississippi river and the busiest airport in the U.S.A. The county seat Memphis is the largest city in TN with 64% African-American population. Memphis clusters various industries, including transportation carriers, a petroleum refinery, petrochemical storage and transfer facilities, waste disposal facilities, a power plant, and etc. EPA's Toxic Registry Inventory (TRI) is on watch for 311 emission facilities in Memphis. Most of these emission facilities are located in low-income African American concentrated areas. Clustering and uneven distribution of emission sources raises great concerns of health risk and health disparities related to air pollution in Memphis Area.

### 1.6 Study Aims

The overall objective of this study is to characterize the sources, variations, and mixture profiles of ambient air toxics in urban environments, and examine the sociodemographic disparities in exposures to air toxics in a typical U.S. metropolitan area. There are four specific aims:

Specific aim 1: Evaluate the extent to which concentrations predicted by the 2011 National-scale Air Toxics Assessment (NATA) program represent the actual monitored concentrations. The hypothesis is that the modeled and monitored annual average concentrations of ambient air toxics agree at the census tract level.

Specific aim 2: Characterize the distribution and concentrations of ambient air toxics in Memphis, identify major sources, and estimate non-carcinogenic and carcinogenic risks.

Specific aim 3: Identify the social, economic, demographic, and special factors that determine ambient air toxic concentrations. We will explore these determinants using 2011 NATA data in Memphis and the actual monitoring data collected for Specific aim 2. There are two hypotheses: a) Air toxic concentrations display differential distributions by sociodemographic status, and that, under most conditions, people in lower SES and of minority are at greater risk. b) both modeling and monitoring data give similar results regarding sociodemographic disparities in air toxics concentrations.

Specific aim 4: Identify common air toxic mixtures and establish the exposure profiles in common microenvironments (MEs). We hypothesize that exposure profiles are distinguishable in various MEs.

To increase the representativeness of results, the samplers will be deployed in a variety of neighborhoods and MEs. State-of-the-art methods will be used to collect time-integrated air toxics samples, time-location information, and other data that represent exposure levels. Urban/industrial/ suburban areas of metropolitan Memphis are selected as field study sites given the significance of existing exposures, the diversity of source conditions, the gaps of air toxics information available, the strong partnerships with local community and governmental organizations, and the proximity to the investigators.

# Chapter 2

# A Model-to-Monitor Comparison of 2011 National-scale Air Toxics Assessment (NATA)

## 2.1 Introduction

Measuring air toxic levels are fundamental step precedes any other procedures in studying health risks from exposure to ambient air toxics. Reliable measurement of air toxics levels can assure the quality of study findings. Ambient monitoring and statistic modeling are two major approaches of measuring air toxic levels (McCarthy et al., 2009). Although ambient monitoring following lab analysis usually assures the quality of the measurement, tremendous labor efforts and financial costs are barricade for expanding monitoring area to extended geographic scales. In contrast, statistical modeling exceeds monitoring by using existing emission data to predict ambient air toxic levels saving both time and efforts. However, uncertainty in existing emission data and limitation of statistical model may affect the prediction generated from model approach (U.S.EPA, 2016e, 2017a).

The most commonly used model predictions in previous researches were estimates from NATA which estimates the air toxic level by modeling emission data from national emissions inventory (NEI) which was built upon the toxics release inventory (TRI)(U.S.EPA, 2016f). However, TRI mainly focus on large stationary emission sources, some local small scale emission sources such as dry-cleaning store, auto body and paint shop, and local restaurants are usually not included (U.S.EPA, 2016f). Additionally, for urban area, emissions from mobile sources are major contributor to elevated air toxic levels (Jakubiak-Lasocka, Lasocki, Siekmeier, & Chlopek, 2015). Traffic related emissions vary temporally and spatially which NATA has limitation in capturing(U.S.EPA, 2017a). Furthermore, local meteorological condition is hard to

model and its uncertainty is part of the limitation of model approach. Air toxic levels varies due to different wind direction and speed, temperatures are common difficulties for model approach (U.S.EPA, 2017a).

Despite NATA model assessments completed on behalf of U.S.EPA, limited number of evaluations has been conducted to assess the model approach using monitored measurement as criteria for comparison. A common finding from previous assessment was that NATA usually underestimate the monitoring measurement (Garcia et al., 2014; Pratt et al., 2004; U.S.EPA, 2016h). However, the model approach is being improved along time.

This chapter aims to provide an independent model performance evaluation for 2011 NATA. We propose a model-to-monitor assessment framework by incorporating measurement uncertainty. We compile the real measurements collected at 274 sites throughout the U.S. in 2011, and merge modeling and monitoring datasets. We then assess the agreement using the proposed framework.

## 2.2 Method

### 2.2.1 Data sources and compilation

The monitoring HAP data were extracted from U.S. EPA's Air Quality System (AQS). AQS is a web-based air pollution database accessible to the public. It contains ambient air pollution data and sampling condition information collected from tribal, local and state agencies through consistent and strict quality assurance (QA) processes. The HAPs were monitored following EPA's Air Toxics Monitoring Methods (U.S.EPA, 2016d). In brief, VOCs were measured by TO-15 method, aldehydes by TO-11A, polycyclic aromatic hydrocarbons (PAHs) by TO-13A, and heavy metals by IO 3.5 method. Most HAP samples were analyzed at central

laboratories and their typical limits of quantitation (LOQs) are available (Little, 2015; U.S.EPA, 2015c). We downloaded daily (24-hour) HAP concentrations measured in 2011 (U.S.EPA, 2015a). Conventional units, such as part per billion (ppb), part per million (ppm), ppb carbon (ppbC), or ppm carbon (ppmC), were converted to the standard unit $\mu g/m^3$ to match that used in NATA. Locations of the monitoring sites in AQS were geocoded and assigned the census tract number in ArcGIS 10.3.1 (ESRI, Inc.).

The modeling HAP concentrations at the census tract level were downloaded from 2011 NATA database. The 2011 NATA contained 78,000 census tracts in the continental U.S. AQS and NATA data were then merged by census tract. The merged dataset contained up to 274 monitoring stations from AQS but only 274 matched census tracts from NATA. This subset of NATA data was representative of the entire 2011 NATA dataset, as their key descriptive statistics were very similar (Table 2.1). Thus, NATA in the following text means the matched sub-dataset. In any census tract, NATA gives a single annual average concentration of a compound, and AQS gives 5-162 measurements of the same compound taken around Year 2011.

**Table 2.1 NATAQS and NATAall comparison**

| HAPs | NATAQS | | | | | | NATAall | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | M | SD | Min | Mdn | Max | N | M | SD | Min | Mdn | Max |
| | | | | μg/m$^3$ | | | | | | μg/m$^3$ | | |
| Benzene | 274 | 0.79 | 0.47 | 0.09 | 0.69 | 3.02 | 74034 | 0.70 | 0.43 | 0 | 0.63 | 7.50 |
| Toluene | 257 | 2.62 | 3.01 | 0.09 | 2.08 | 30.37 | 74034 | 2.61 | 3.46 | 0 | 1.84 | 39.67 |
| Styrene | 236 | 0.05 | 0.16 | 0.00 | 0.02 | 1.57 | 75023 | 0.02 | 0.06 | 0 | 0.02 | 6.72 |
| Ethylbenzene | 252 | 0.28 | 0.21 | 0.02 | 0.25 | 1.66 | 75025 | 0.23 | 0.19 | 0 | 0.20 | 3.09 |
| Cumene | 113 | 0.01 | 0.03 | 0.00 | 0.00 | 0.20 | 75013 | 0.00 | 0.01 | 0 | 0.00 | 1.98 |
| Naphthalene | 52 | 0.05 | 0.04 | 0.00 | 0.04 | 0.22 | 74034 | 0.04 | 0.03 | 0 | 0.03 | 0.94 |
| 1,4-Dichlorobenzene | 153 | 0.01 | 0.04 | 0.00 | 0.00 | 0.27 | 74034 | 0.02 | 0.05 | 0 | 0.00 | 1.13 |
| Methyl chloride | 203 | 1.10 | 0.07 | 1.09 | 1.09 | 2.06 | 74591 | 1.07 | 0.13 | 0 | 1.09 | 2.06 |
| Vinyl chloride | 240 | 0.00 | 0.02 | 0.00 | 0.00 | 0.15 | 74034 | 0.00 | 0.00 | 0 | 0.00 | 0.73 |
| Bromomethane | 200 | 0.04 | 0.09 | 0.03 | 0.03 | 1.34 | 74438 | 0.04 | 0.05 | 0 | 0.03 | 2.00 |
| Ethylene Dichloride | 230 | 0.00 | 0.01 | 0.00 | 0.00 | 0.08 | 74034 | 0.00 | 0.00 | 0 | 0.00 | 0.31 |
| Chloroform | 252 | 0.01 | 0.03 | 0.00 | 0.00 | 0.23 | 74034 | 0.00 | 0.01 | 0 | 0.00 | 1.57 |
| Trichloroethylene | 247 | 0.02 | 0.04 | 0.00 | 0.01 | 0.44 | 74034 | 0.02 | 0.05 | 0 | 0.01 | 5.49 |
| 1,1,2-Trichloroethane | 185 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 74871 | 0.00 | 0.00 | 0 | 0.00 | 0.01 |
| Carbon tetrachloride | 247 | 0.55 | 0.00 | 0.55 | 0.55 | 0.55 | 74917 | 0.54 | 0.08 | 0 | 0.55 | 0.64 |
| Tetrachloroethylene | 252 | 0.09 | 0.17 | 0.00 | 0.03 | 1.12 | 74034 | 0.11 | 0.24 | 0 | 0.03 | 5.07 |
| Formaldehyde | 128 | 1.66 | 0.52 | 0.55 | 1.61 | 2.94 | 74034 | 1.59 | 0.55 | 0 | 1.57 | 5.56 |
| Acetaldehyde | 133 | 1.95 | 0.55 | 0.91 | 1.85 | 3.28 | 74034 | 1.94 | 0.66 | 0 | 1.88 | 4.15 |
| Methyl Isobutyl ketone | 93 | 0.10 | 0.15 | 0.00 | 0.07 | 1.27 | 74968 | 0.07 | 0.08 | 0 | 0.05 | 2.12 |
| Chromium VI (TSP) | 27 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 74034 | 0.00 | 0.00 | 0 | 0.00 | 0.00 |
| Acrylonitrile | 81 | 0.00 | 0.01 | 0.00 | 0.00 | 0.09 | 74034 | 0.00 | 0.01 | 0 | 0.00 | 1.24 |
| 1,3-butadiene | 258 | 0.08 | 0.06 | 0.00 | 0.06 | 0.44 | 74034 | 0.06 | 0.05 | 0 | 0.05 | 0.79 |
| Carbon disulfide | 93 | 0.05 | 0.34 | 0.01 | 0.01 | 3.27 | 74906 | 0.01 | 0.05 | 0 | 0.01 | 4.91 |
| n-Hexane | 159 | 0.91 | 0.60 | 0.12 | 0.83 | 2.87 | 75020 | 0.76 | 0.66 | 0 | 0.60 | 16.05 |
| Methyl tert-butyl ether | 127 | 0.00 | 0.01 | 0.00 | 0.00 | 0.12 | 73815 | 0.00 | 0.01 | 0 | 0.00 | 0.31 |
| 2,2,4-trimethylpentane | 105 | 0.36 | 0.19 | 0.10 | 0.34 | 0.92 | 75009 | 0.37 | 0.23 | 0 | 0.33 | 3.59 |
| Lead (TSP) | 51 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 74034 | 0.00 | 0.00 | 0 | 0.00 | 0.10 |

### 2.2.2 HAPs of interest

We selected 27 HAPs to evaluate the model-to-monitor agreement. The selection was based on four criteria: (1) They had high ranking in both cancer and respiratory risks in NATA; (2) They were measured at ≥25 monitoring sites in AQS; (3) They were prioritized in previous NATA reports; and (4) They represented different chemical groups. We then divided the 27 HAPs into four groups mainly based on their chemical structures: (1) Aromatic compounds: benzene, isopropyl benzene, ethyl benzene, styrene, toluene, 1,4-dichlorobenzene, and naphthalene; (2) Halogenated compounds: 1,1,2-trichloroethane, bromoethane, carbon tetrachloride, chloroform, methyl chloride, ethylene dichloride, trichloroethylene, tetrachloroethylene, and vinyl chloride; (3) Carbonyl compounds: methyl isobutyl ketone,

acetaldehyde, and formaldehyde; and (4) Other compounds: chromium, lead, carbon disulfide, 2,2,4-trimethylpentane, *n*-hexane, 1,3-butadiene, methyl *tert*-butyl ether, and acrylonitrile. Not all the sites measured all the compounds, and thus the site numbers varied from 27 to 274 depending on compound.

### 2.2.3 Model-to-monitor comparison methods

The 2011 NATA modeling results contained extremely low concentration for certain compounds, e.g., the annual average concentrations of 1,1,2-trichloroethane and chromium were 0.00041 and 0.00003 μg/m$^3$, respectively. In practice, the measurement method has a LOQ for a specific chemical, defined as the lowest concentration that can be accurately measured during regular laboratory analyzing conditions (Little, 2015). Following the concept of LOQ, the difference between the modeled and observed concentrations, i.e., $\Delta M = |M - O|$, was unquantifiable if ΔM was less than LOQ. In this case, M and O were considered to be in agreement (U.S.EPA, 2017d). When ΔM was ≥LOQ, we conducted the following analyses.

To determine the national-level agreement, we compared national medians considering the large spatial heterogeneity among monitoring sites. For each target HAP, we first determined whether the difference of modeled and observed medians was quantifiable or not, and then compared two medians using Wilcoxon signed-rank test if quantifiable. A p-value of ≥0.05 was considered agreement.

At individual sites, we compared annual modeled and observed averages using statistical methods if ΔM was quantifiable. We calculated the 95% confidence interval (CI) of multiple AQS measurements of a chemical, and determine if the single NATA annual average value fell within the 95% CI. We log-transformed AQS data as they followed a skewed lognormal

distribution, and then calculated the 95% CI using Cox's method (X. H. Zhou & Gao, 1997). This is a strict statistical comparison method and applies the widely accepted criterion of 95% CI or p-value of 0.05.

At each site, if NATA agreed with AQS, the site was defined as an agreement site for that chemical; otherwise, it was defined as underestimation or overestimation site. These steps were repeated for the 27 HAPs and all the available sites. The percentages of underestimation, agreement, and overestimation sites were calculated for all the sites in the U.S. Percentages of agreement sites were further calculated by EPA region, based on just difference and screening analyses.

EPA has long been using a factor of 2 as the criterion for model-to-monitor comparisons (Eastern Research Group, 2010; Garcia et al., 2014; U.S.EPA, 2002a). That is, an M/O ratio of 0.5-2 could be considered agreement. To verify if this empirical method gave the same results as statistical methods, we then repeated the comparisons using M/O ratios. These were side analyses, and the results were presented as supplemental information.

All the analyses were performed SAS (v9.4, SAS Institute Inc., Cary, NC), Microsoft Excel (2010) and Arc GIS 10.3.1 (ESRI, Inc.).

## 2.3 Results

### 2.3.1 Comparison of national statistics

Ambient HAP concentrations were low in the U.S. in 2011 (Table 1). Eighteen compounds had median concentrations below 0.1 $\mu g/m^3$ in both two datasets, and these low concentrations were slightly above or far below their LOQs. According to AQS, the median concentrations ranged from near 0 $\mu g/m^3$ (1,1,2-trichloroethane) to 1.44 $\mu g/m^3$ (acetaldehyde).

Ethylbenzene (12.6 μg/m$^3$), trichloroethylene (16.9 μg/m$^3$), and carbon sulfide (29.4 μg/m$^3$) were the only three compounds with maxima that exceeded 10 μg/m$^3$.

NATA can moderately predict national median HAP levels. In total, sixteen HAPs had their ΔM less than LOQs indicating agreement between NATA and AQS. One out of the rest 11 compounds did not show significant difference of medians between NATA and AQS in Wilcoxon Signed rank test (Table 1). In general, seventeen HAPs reached modeled-to-monitored agreement, while ten compounds showed poor agreement. Nine of these ten compounds were underestimated by modeling predictions and one was overestimated. Carbon disulfide was consistently and the most underestimated while acetaldehyde was the most overestimated. In combination we can consider that the medians of all compounds moderately agreed between two datasets.

NATA is unable to capture extreme concentrations. The maximum concentrations in AQS were much higher than those in NATA (Table 2.2). Each air monitoring station is a point monitor in nature. Although representing a surrounding area, the monitoring station may be impacted by the local sources such as highways, industrial facilities, non-road vehicles, and episodic emissions, all of which could cause high concentrations. In contrast, NATA estimated the annual average concentration using the Gaussian dispersion models. The dispersion models often are unable to simulate extrema (Rosenbaum et al., 1999; Scheffe et al., 2016).

## Table 2.2 National statistics of HAP concentrations

| HAPs | | | Sites | AQS | | | | | | NATAQS | | | | | | Mdn | Wilcoxon |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | LOQ | ΔM<LOQ? | ΔM<LOQ | M | SD | Min | Mdn | Max | <LOQ | M | SD | Min | Mdn | Max | <LOQ | M/O | Signed rank test |
| | | µg/m$^3$ | Yes/No | % | µg/m$^3$ | | | | | % | µg/m$^3$ | | | | | % | Ratio | *p*-value |
| **Benzene** | 274 | 0.06 | Yes | 12.04 | 0.95 | 0.78 | 0.00 | 0.73 | 5.57 | 0.7 | 0.79 | 0.47 | 0.09 | 0.69 | 3.02 | 0.0 | 0.9 | 0.01 |
| **Toluene** | 257 | 0.10 | No | 8.17 | 1.75 | 1.27 | 0.06 | 1.41 | 6.98 | 0.4 | 2.62 | 3.01 | 0.09 | 2.08 | 30.37 | 0.4 | 0.7 | <.001 |
| **Styrene** | 236 | 0.04 | No | 45.34 | 0.14 | 0.19 | 0.00 | 0.07 | 1.45 | 35.6 | 0.05 | 0.16 | 0.00 | 0.02 | 1.57 | 81.4 | ***0.3*** | <.001 |
| **Ethylbenzene** | 252 | 0.05 | Yes | 18.25 | 0.32 | 0.80 | 0.00 | 0.21 | 12.57 | 9.1 | 0.28 | 0.21 | 0.02 | 0.25 | 1.66 | 10.7 | 0.9 | 0.60 |
| **Cumene** | 113 | 0.50 | Yes | 94.69 | 0.11 | 0.40 | 0.00 | 0.03 | 3.79 | 93.8 | 0.01 | 0.03 | 0.00 | 0.00 | 0.20 | 100 | 0.1 | <.001 |
| **Naphthalene** | 52 | 0.00 | No | 3.85 | 0.07 | 0.11 | 0.01 | 0.04 | 0.56 | 0.0 | 0.05 | 0.04 | 0.00 | 0.04 | 0.22 | 0.0 | 0.6 | 0.89 |
| **1,4-Dichlorobenzene** | 153 | 0.07 | Yes | 62.75 | 0.18 | 0.68 | 0.00 | 0.05 | 7.40 | 63.4 | 0.01 | 0.04 | 0.00 | 0.00 | 0.27 | 94.1 | 0.0 | <.001 |
| **Methyl chloride** | 203 | 0.04 | No | 12.32 | 1.23 | 0.25 | 0.00 | 1.23 | 2.5 | 0.5 | 1.10 | 0.07 | 1.09 | 1.09 | 2.06 | 0.0 | 0.9 | <.001 |
| **Vinyl chloride** | 240 | 0.01 | Yes | 91.67 | 0.01 | 0.04 | 0.00 | 0.00 | 0.51 | 89.6 | 0.00 | 0.02 | 0.00 | 0.00 | 0.15 | 95.0 | 0.1 | 0.01 |
| **Bromomethane** | 200 | 0.02 | Yes | 59.50 | 0.02 | 0.04 | 0.00 | 0.01 | 0.46 | 69.5 | 0.04 | 0.09 | 0.03 | 0.03 | 1.34 | 0.0 | 1.7 | <.001 |
| **Ethylene Dichloride** | 230 | 0.02 | Yes | 54.78 | 0.07 | 0.34 | 0.00 | 0.02 | 4.73 | 53.0 | 0.00 | 0.01 | 0.00 | 0.00 | 0.08 | 95.7 | 0.0 | <.001 |
| **Chloroform** | 252 | 0.04 | No | 26.19 | 0.13 | 0.39 | 0.00 | 0.09 | 6.05 | 25.0 | 0.01 | 0.03 | 0.00 | 0.00 | 0.23 | 93.7 | ***0.0*** | <.001 |
| **Trichloroethylene** | 247 | 0.03 | Yes | 72.06 | 0.11 | 1.07 | 0.00 | 0.01 | 16.85 | 62.8 | 0.02 | 0.04 | 0.00 | 0.01 | 0.44 | 76.9 | 0.6 | <.001 |
| **1,1,2-Trichloroethane** | 185 | 0.05 | Yes | 93.51 | 0.01 | 0.02 | 0.00 | 0.00 | 0.15 | 93.5 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.1 | 0.03 |
| **Carbon tetrachloride** | 247 | 0.04 | No | 25.10 | 0.54 | 0.19 | 0.00 | 0.59 | 1.68 | 3.6 | 0.55 | 0.00 | 0.55 | 0.55 | 0.55 | 0.0 | 0.9 | <.001 |
| **Tetrachloroethylene** | 252 | 0.06 | No | 39.68 | 0.13 | 0.14 | 0.00 | 0.09 | 1.10 | 29.0 | 0.09 | 0.17 | 0.00 | 0.03 | 1.12 | 67.1 | ***0.3*** | <.001 |
| **Formaldehyde** | 128 | 0.03 | No | 4.69 | 1.29 | 0.71 | 0.20 | 1.16 | 6.77 | 0.0 | 1.66 | 0.52 | 0.55 | 1.61 | 2.94 | 0.0 | 0.7 | <.001 |
| **Acetaldehyde** | 133 | 0.03 | No | 5.26 | 1.56 | 0.63 | 0.43 | 1.44 | 3.73 | 0.0 | 1.95 | 0.55 | 0.91 | 1.85 | 3.28 | 0.0 | 1.3 | <.001 |
| **Methyl Isobutyl ketone** | 93 | 0.08 | Yes | 65.59 | 0.12 | 0.10 | 0.00 | 0.10 | 0.60 | 38.7 | 0.10 | 0.15 | 0.00 | 0.07 | 1.27 | 53.8 | 0.7 | 0.00 |
| **Chromium VI (TSP)** | 27 | 0.00 | Yes | 100.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100.0 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 1.6 | 0.00 |
| **Acrylonitrile** | 81 | 0.10 | Yes | 79.01 | 0.11 | 0.24 | 0.00 | 0.01 | 1.06 | 79.0 | 0.00 | 0.01 | 0.00 | 0.00 | 0.09 | 100 | 0.0 | <.001 |
| **1,3-butadiene** | 258 | 0.02 | Yes | 29.46 | 0.08 | 0.14 | 0.00 | 0.05 | 1.74 | 34.1 | 0.08 | 0.06 | 0.00 | 0.06 | 0.44 | 12.0 | ***1.1*** | <.001 |
| **Carbon disulfide** | 93 | 0.02 | No | 20.43 | 0.92 | 3.44 | 0.00 | 0.07 | 29.44 | 19.4 | 0.05 | 0.34 | 0.01 | 0.01 | 3.27 | 95.7 | ***0.1*** | <.001 |
| **n-Hexane** | 159 | 0.31 | Yes | 43.40 | 0.81 | 0.75 | 0.06 | 0.57 | 4.39 | 19.5 | 0.91 | 0.60 | 0.12 | 0.83 | 2.87 | 16.4 | 0.8 | <.001 |
| **Methyl tert-butyl ether** | 127 | 0.16 | Yes | 98.43 | 0.01 | 0.07 | 0.00 | 0.00 | 0.72 | 98.4 | 0.00 | 0.01 | 0.00 | 0.00 | 0.12 | 100 | 0.0 | 0.08 |
| **2,2,4-trimethylpentane** | 105 | 0.28 | Yes | 79.05 | 0.50 | 0.53 | 0.01 | 0.34 | 3.1 | 39.0 | 0.36 | 0.19 | 0.10 | 0.34 | 0.92 | 38.1 | 1.0 | 0.16 |
| **Lead (TSP)** | 51 | 0.00 | No | 7.84 | 0.01 | 0.01 | 0.00 | 0.00 | 0.07 | 7.8 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 5.9 | ***0.2*** | <.001 |

Note: HAPs –hazardous air pollutants; AQS – AQS data; NATAQS – NATA data in which sites were matched with AQS data; LOQ – limit of quantitation; M – mean; SD – standard deviation; Mdn – median; Min – minimum; Max – maximum

### 2.3.2 Comparisons at individual sites

At individual monitoring sites, the ΔM between NATA and AQS annual average was first examined for each compound (Table 2.3). It was noticeable that ΔM (chromium) was below the LOQ at all the sites, indicating that NATA chromium estimates agreed with AQS measurements. Similarly, ΔM was below the LOQ at over 90% of sites for cumene, vinyl chloride, 1,1,2-Trichloroethane, and methyl tert-butyl ether, and at 50-90% of sites for another seven compounds. A total of 12 compounds showed agreement at ≥50% of sites by comparing ΔM to LOQ.

**Table 2.3 Agreement of NATA sites with AQS sites in the U.S.**

| HAPs | N | ΔM<LOQ | <95% LCL | Within 95% CL | >95% UCL | Total |
|---|---|---|---|---|---|---|
| | | | | % | | |
| Benzene | 274 | 12.0 | 43.8 | 9.9 | 34.3 | 21.9 |
| Toluene | 257 | 8.2 | 18.7 | 20.6 | 52.5 | 28.8 |
| Styrene | 236 | 45.3 | 45.3 | 5.9 | 3.4 | 51.3 |
| Ethylbenzene | 252 | 18.3 | 36.5 | 9.5 | 35.7 | 27.8 |
| Cumene | 113 | 94.7 | 5.3 | 0.0 | 0.0 | 94.7 |
| Naphthalene | 52 | 3.8 | 30.8 | 26.9 | 38.5 | 30.8 |
| 1,4-Dichlorobenzene | 153 | 62.7 | 30.7 | 3.3 | 3.3 | 66.0 |
| Methyl chloride | 203 | 12.3 | 63.1 | 10.3 | 14.3 | 22.7 |
| Vinyl chloride | 240 | 91.7 | 5.0 | 2.9 | 0.4 | 94.6 |
| Bromomethane | 200 | 59.5 | 2.5 | 3.5 | 34.5 | 63.0 |
| Ethylene Dichloride | 230 | 54.8 | 41.3 | 3.5 | 0.4 | 58.3 |
| Chloroform | 252 | 26.2 | 68.7 | 5.2 | 0.0 | 31.3 |
| Trichloroethylene | 247 | 72.1 | 15.0 | 4.9 | 8.1 | 76.9 |
| 1,1,2-Trichloroethane | 185 | 93.5 | 4.9 | 1.6 | 0.0 | 95.1 |
| Carbon tetrachloride | 247 | 25.1 | 49.8 | 2.8 | 22.3 | 27.9 |
| Tetrachloroethylene | 252 | 39.7 | 29.8 | 13.9 | 16.7 | 53.6 |
| Formaldehyde | 128 | 4.7 | 7.0 | 21.1 | 67.2 | 25.8 |
| Acetaldehyde | 133 | 5.3 | 10.5 | 21.8 | 62.4 | 27.1 |
| Methyl Isobutyl ketone | 93 | 65.6 | 23.7 | 3.2 | 7.5 | 68.8 |
| Chromium VI (TSP) | 27 | 100.0 | 0.0 | 0.0 | 0.0 | 100.0 |
| Acrylonitrile | 81 | 79.0 | 19.8 | 1.2 | 0.0 | 80.2 |
| 1,3-butadiene | 258 | 29.5 | 18.2 | 10.5 | 41.9 | 39.9 |
| Carbon disulfide | 93 | 20.4 | 60.2 | 17.2 | 2.2 | 37.6 |
| n-Hexane | 159 | 43.4 | 13.2 | 5.0 | 38.4 | 48.4 |
| Methyl tert-butyl ether | 127 | 98.4 | 1.6 | 0.0 | 0.0 | 98.4 |
| 2,2,4-trimethylpentane | 105 | 79.0 | 17.1 | 0.0 | 3.8 | 79.0 |
| Lead (TSP) | 51 | 7.8 | 74.5 | 5.9 | 11.8 | 13.7 |

Note: LOQ – limit of quantitation; CI– confidence interval; LCL – lower confidence limit; UCL – upper confidence limit

When ΔM was quantifiable, toluene, formaldehyde, acetaldehyde, naphthalene showed agreement at 20-27% of sites, methyl chloride, 1,3-butadiene, tetrachloroethylene, and carbon disulfide showed agreement at 10-20% of sites, and the remaining 20 chemicals all showed

agreement at <10% of sites. Therefore, NATA agreed with AQS at a small portion (<30%) of sites nationally at the quantifiable concentration ranges (Table 2.3).

Taken together, 14 compounds had NATA-AQS agreement at >50% of sites. Methyl chloride, chloroform, carbon disulfide, and lead were nationally underestimated, and toluene, formaldehyde, and acetaldehyde were nationally overestimated. Benzene, ethylbenzene, naphthalene, carbon tetrachloride, 1,3-butadiene, and n-hexane did not show strong patterns, meaning agreement, underestimation and overestimation occurred at roughly similar numbers of sites.

EPA's factor of 2 criterion gave better agreement results (Table 2.4). A total of 21 compounds showed agreement at ≥50% of sites. Significant increase of agreement sites occurred for benzene, methyl chloride, carbon tetrachloride, formaldehyde, and acetaldehyde. Styrene, tetrachloroethylene and 1,3-butadiene showed agreement at 44-48% of sites. Chloroform, carbon disulfide, and lead were underestimated at 72-75% of sites. NATA overestimated concentrations of all compounds at a small portion of sites. These indicated that the "factor of 2" criterion was looser than statistical analyses, and the resulting disagreement inclined to underestimation.

**Table 2.4 Agreement of NATA sites with AQS sites in the U.S. (Fractional Bias)**

| HAPs | N | | FB | | | Total |
|---|---|---|---|---|---|---|
| | | ΔM<LOQ | (-2,-0.67) | (-0.67,0.67) | (0.67,2) | |
| | | | % | | | |
| Benzene | 274 | 12.0 | 20.1 | 55.8 | 12.0 | 67.9 |
| Toluene | 257 | 8.2 | 11.7 | 48.2 | 31.9 | 56.4 |
| Styrene | 236 | 45.3 | 49.2 | 2.5 | 3.0 | 47.9 |
| Ethylbenzene | 252 | 18.3 | 19.4 | 40.1 | 22.2 | 58.3 |
| Cumene | 113 | 94.7 | 5.3 | 0.0 | 0.0 | 94.7 |
| Naphthalene | 52 | 3.8 | 21.2 | 55.8 | 19.2 | 59.6 |
| 1,4-Dichlorobenzene | 153 | 62.7 | 34.0 | 1.3 | 2.0 | 64.1 |
| Methyl chloride | 203 | 12.3 | 0.5 | 85.7 | 1.5 | 98.0 |
| Vinyl chloride | 240 | 91.7 | 6.3 | 2.1 | 0.0 | 93.8 |
| Bromomethane | 200 | 59.5 | 4.0 | 2.0 | 34.5 | 61.5 |
| Ethylene Dichloride | 230 | 54.8 | 44.8 | 0.4 | 0.0 | 55.2 |
| Chloroform | 252 | 26.2 | 71.8 | 2.0 | 0.0 | 28.2 |
| Trichloroethylene | 247 | 72.1 | 18.6 | 2.0 | 7.3 | 74.1 |
| 1,1,2-Trichloroethane | 185 | 93.5 | 6.5 | 0.0 | 0.0 | 93.5 |
| Carbon tetrachloride | 247 | 25.1 | 0.4 | 63.6 | 10.9 | 88.7 |
| Tetrachloroethylene | 252 | 39.7 | 42.5 | 4.8 | 13.1 | 44.4 |
| Formaldehyde | 128 | 4.7 | 1.6 | 78.1 | 15.6 | 82.8 |
| Acetaldehyde | 133 | 5.3 | 0.0 | 81.2 | 13.5 | 86.5 |
| Methyl Isobutyl ketone | 93 | 65.6 | 23.7 | 4.3 | 6.5 | 69.9 |
| Chromium VI (TSP) | 27 | 100.0 | 0.0 | 0.0 | 0.0 | 100.0 |
| Acrylonitrile | 81 | 79.0 | 21.0 | 0.0 | 0.0 | 79.0 |
| 1,3-butadiene | 258 | 29.5 | 18.6 | 16.7 | 35.3 | 46.1 |
| Carbon disulfide | 93 | 20.4 | 75.3 | 2.2 | 2.2 | 22.6 |
| n-Hexane | 159 | 43.4 | 8.8 | 20.8 | 27.0 | 64.2 |
| Methyl tert-butyl ether | 127 | 98.4 | 1.6 | 0.0 | 0.0 | 98.4 |
| 2,2,4-trimethylpentane | 105 | 79.0 | 13.3 | 5.7 | 1.9 | 84.8 |
| Lead (TSP) | 51 | 7.8 | 68.6 | 13.7 | 9.8 | 21.6 |

Note: LOQ – limit of quantitation; FB– fractional bias

### 2.3.3 Agreement assessment by EPA region

The agreement between NATA estimates and AQS measurements could be further examined by EPA regions, as shown in Figure 2.1. Checking by row, chloroform, tetrachloroethylene, 1,3-butadiene, carbon disulfide, n-hexane, and lead notably had poor agreement in most or all regions. In contrast, seven compounds showed good agreement across regions, including methyl chloride, vinyl chloride, trichloroethylene, 1,1,2-trichloroethane, acetaldehyde, chromium VI (TSP), and methyl tert-butyl ether. Checking by column in Figure 2.1, the majority of regions showed good agreement. In addition to poor agreement compounds mentioned above, Regions 3, 8 and 10 showed poor agreement for aromatic compounds. For example, styrene and ethylbenzene did not show model-to-monitor agreement at any sites in Region 10. Combining all the regions, these agreement and disagreement reflect the results in Table 2.3.


**Figure 2.1 EPA regional modeled-to-monitored comparison**

Note: R – EPA region; Under – underestimation; Agree – agreement; Over – overestimation

## 2.4 Discussion

### 2.4.1 Similar findings from national and local studies

Our results confirmed previous national and local comparisons. Previous NATA evaluations found good agreement for only a few compounds and underestimation for most compounds (Garcia et al., 2014; U.S.EPA, 2016g). The 2005 NATA model assessment reported only 8 out of 68 compounds showed agreement at the national level and other compounds were all underestimated (Eastern Research Group, 2010). At state and local levels, Lupo and Symanski (Lupo & Symanski, 2009) found 1996 NATA underestimated 8 out 15 HAPs and 1999 NATA underestimated 18 out of 27 HAPs in Texas. Wang et al. found general agreement for benzene and toluene concentrations modeled by 1999 NATA in Camden, New Jersey (S. W. Wang et al., 2009). The 2002 NATA underestimated 32 out of 49 HAPs measured at 7 sites in and around Pittsburgh, Pennsylvania (Logue et al., 2011). The Detroit Exposure and Aerosol Research Study (DEARS) reported that benzene concentrations in 2002 NATA generally agreed with field measurements during 2004 to 2007 (George et al., 2011). Garcia et al. found that all 12 HAPs were underestimated by 1996 NATA, 8 out of 9 were underestimated by 1999 NATA, 10 out of 12 were underestimated by 2002 NATA, and 6 out of 10 were underestimated by 2005 NATA (Garcia et al., 2014). These findings indicate that model-to-monitor agreement was inconsistent by region and chemical, and underestimation was more frequent (Garcia et al., 2014; George et al., 2011; S. W. Wang et al., 2009).

### 2.4.2 Impacts of comparison methods and metrics

Model-to-monitor comparison results were significantly impacted by the comparison methods. Previous studies have applied a number of model-to-monitor comparison metrics and methods, including biases and root mean square error (Scheffe et al., 2016; Stroud et al., 2016;

Vennam, Vizuete, & Arunachalam, 2015; S. W. Wang et al., 2009), Kendall rank correlation (Rosenbaum et al., 1999), ratios (Garcia et al., 2014; Logue et al., 2011; Lupo & Symanski, 2009), regressions (S. W. Wang et al., 2009), and even complex metrics (S. C. Yu, Eder, Dennis, Chu, & Schwartz, 2006). One difficulty was in selecting a commonly accepted criterion for the metric, for example, EPA uses a relative bias of within ±30% and median ratio of 0.5-2 for agreement. The goodness-of-fit of a regression line, indicated by $R^2$, is often arbitrary. Median ratio of modeled-to-monitored concentrations is the most commonly used metric; however, it may become extremely small or large when concentrations are too small to be practically quantifiable.

We introduced LOQ to conquer the measurement uncertainty issue ignored in previous studies. It turned out model vs. monitor differences were unquantifiable at a large portion of sites for many compounds. An unquantifiable difference should mean an agreement; however, statistical analyses of these uncertain small numbers often lead to significant differences. For example, we found 100% agreement for chromium due to its extremely low concentrations (median=0.00001 $\mu g/m^3$) estimated by NATA, while the 2005 NATA evaluation reported a 0% agreement when just using ratios (U.S.EPA, 2010). This and other examples suggest ignorance of measurement uncertainty, in particular the LOQs, would lead to distinctly different results.

### 2.4.3 Causes of disagreement

HAP modeling has been affected by a number of factors and uncertainties, including monitoring station siting, sampling frequency, emission inventory, measurement uncertainty, model uncertainty, and comparison method. The 1996 NATA evaluation attributed the general underestimation to four reasons (EPA, 2016): (1) missing emissions sources; (2) underestimated emission rates; (3) sites intended to find peak concentrations; and (4) measurement accuracy. As

seen in Table 2.2, NATA model was unable to capture extreme concentrations. Average concentrations measured from monitors within a census tract might be affected by extrema due to nearby short-term strong emissions, which could not be captured by the census-tract averages in NATA. Similarly, the NEI, which NATA estimates were based on, might miss local emission sources (Scheffe et al., 2016). Lack of stable estimates on meteorological conditions and photochemical reactions is another factor leading to disagreement. For example, unstable estimates on wind conditions and secondary formation of chemicals were major weakness of the NATA model (Rosenbaum et al., 1999). The uncertainty in monitored measurement due to insufficient and unbalanced geographic coverage of monitoring sites also contributed to the discrepancies between monitored and modeled estimates (Garcia et al., 2014; Scheffe et al., 2016).

### 2.4.4 Study limitations

The strengths of this study were utilization of national data from EPA's monitoring sites, consideration of measurement uncertainty, and statistical comparisons with commonly used p-value of 0.05 criterion. We have also acknowledged limitations in data sources and the comparison methodology.

We restricted our assessment to 27 HAPs rather than the 180 HAPs EPA promoted. We targeted on these 27 compounds because they were measured at more than 50 sites and were prioritized for their risk on cancer and respiratory disease. Chromium VI (TSP) was measured at only 27 sites, but it was assessed because it has high carcinogenicity with inhalation unit risk of $1.2 \times 10^{-2}$ per $mg/m^3$ (U.S.EPA, 2017e). , Those compounds not assessed in our study still need to be evaluated carefully before utilized in any further studies. The annual average monitored measurement extracted from AQS was considered more accurate when it is calculated based on

complete data which has at least two seasons' data in a year (EPA, 2016). Due to limited data for certain target compounds in our study, we extracted annual average reported in AQS without restriction to datasets which have at least two seasons' data.

In our comparison methodology, LOQs were adopted from ERG (U.S.EPA, 2015c) where the LOQ for each measurement is not specified. However, the introduction of LOQ into comparison methods is innovative. Given limited data source, the LOQ from ERG served well for the goal of identifying the quantifiable difference of modeled and observed medians. We conducted point-to-range comparison at a very small geographic unit, i.e. census tract. 1996 NATA evaluation concluded that NATA estimates was unreliable at census tract but at county level or above (EPA, 2016). A more reasonable comparison should use larger geographic scale and the technique MAXTOMON which compares the average monitored measurement with the maximum modeled estimate within certain distance of the monitor (EPA, 2016).

### 2.4.5 Implications for environmental health disparity research

Environment plays a critical role in determining people's health (ODPHP, 2018). Environmental health disparity is the difference of health risks that people have when they experience both uneven exposure to various environmental risk factors and social inequality (Gee & Payne-Sturges, 2005).Therefore, environmental justice requires that people with various sociodemographic characteristics have equally distributed health risks from the environment given appropriate environmental policy making and law enforcement (Schlosberg, 2004).

Environmental health disparity is usually examined at census tract level because census tract is considered as a geographic area roughly representative for a neighborhood where the sociodemographic characteristics are homogeneous among a stable size (1500 housing units and

4000 people in average) of population within that census tract (Farley, 2004). A few local studies assessed model performance with census tract level NATA data in the past decade (Garcia et al., 2014; George et al., 2011; Logue et al., 2011; Lupo & Symanski, 2009). Lupo and Symanski (Lupo & Symanski, 2009) found 1996 NATA underestimated 8 out 15 HAPs and 1999 NATA underestimated 18 out of 27 HAPs. Wang et al. compared benzene and toluene concentrations modeled by 1999 NATA with monitoring data collected during 2004 to 2006 in Camden, New Jersey, and found general agreement and underestimation for high-end percentiles (S. W. Wang et al., 2009). Logue et al. found 2002 NATA underestimated 32 out of 49 HAPs measured at 7 sites in and around Pittsburgh, Pennsylvania (Logue et al., 2011). The Detroit Exposure and Aerosol Research Study (DEARS) reported that 2002 NATA obtained ambient benzene concentrations generally agreed with that measured during 2004 to 2007 (George et al., 2011). Garcia et al. assessed all previous NATA (1996, 1999, 2002, and 2005) for 12 HAPs in California. They found that in 1996, all 12 HAPs were underestimated in California; 8 out of 9 HAPs were underestimated in 1999; 10 out 12 HAPs were underestimated in 2002 ; and 6 out of 10 HAPs were underestimated in 2005 (Garcia et al., 2014). All these studies concluded that modeling data tended to underestimate ambient HAP concentrations and certain compounds (e.g., benzene) always had good agreement (Garcia et al., 2014; George et al., 2011; S. W. Wang et al., 2009).

NATA data have values for prioritizing high risk areas; however, it may not be useful for EJ research because NATA data is only at census tract level and EJ analysis is often at the neighborhood level which has wider range of geographical level, e.g., census block group level, zip code level. Additionally, census tract sometimes exceed city boundaries and have rural area

included (Schlosberg, 2004).Therefore, it is recommended being cautious about using NATA for environmental health disparity analysis.

## 2.5 Conclusion

This study provides an independent model-to-monitor assessment for the latest NATA. Given the findings from our assessment, we concluded that modeling estimates of evaluated air toxics from NATA 2011 could moderately estimate the ambient measurements from monitors. Our results strengthened the idea that NATA modeling estimates tend to underestimate the monitored measurement and cautious need to be taken when using NATA to assess environmental health disparity. Overall, modeled-to-monitored assessment studies would be useful for future environmental epidemiology and justice studies given that model estimates serves as practical alternatives for monitoring measurement which might not be available due to resources and time it costs. Future assessment with expanded number of pollutants and various environments (e.g., urban vs. rural) utilizing robust evaluation method would provide researchers with more options and information when conducting environmental epidemiology and justice studies.

# Chapter 3

## Characterization of health risks from air toxics in Memphis Area

### 3.1 Introduction

Memphis has many health issues that may be related to air toxics exposure. Memphis was among the nation's top three "Asthma Capitals" from 2010-2015 (AAFA 2016). Shelby County has infant mortality twice of the national level (Community Commons 2014). Shelby County has many health indicators ranked top in TN, such as infant mortality (#1), hypertension (#1), obesity (#2), and stroke mortality (#3) (Tennessee Department of Health 2011). Cardiovascular disease and cancer are the top two leading causes of death in Shelby (Tennessee Department of Health 2011). Many schools are located near freeways, which may cause childhood asthma (Gale et al. 2012). Diseases prevalence also displays strong spatial patterns: mortalities of cardiovascular disease, cancer, and chronic lower respiratory disease are all elevated in the western part of Memphis, an area consisting predominantly of low-income African American residents. As air pollution is linked to these diseases (Suh et al. 2000), communities have high concerns about air pollution and the health impacts.

Volatile organic compounds (VOCs) represent the majority of air toxics and may pose serious health risks on human populations. Of the 187 air toxics listed under the 1990 Clean Air Act Amendments (CAAA 1990), 88% are organic chemicals or mixture of organic chemicals. VOCs are defined as organic compounds having a vapor pressure greater than 0.1 mm Hg at 25 °C and 760 mm Hg (US EPA 1999). Human exposure to VOCs is linked to many adverse health effects, ranging from respiratory diseases, immune and neurological damage, reproductive and endocrine disorders, cardiovascular diseases to cancers (Kampa and Castanas 2008; Shin et al. 2015; Suh et al. 2000). For example, benzene, toluene, ethylbenzene and xylenes (BTEX) are

common constituents of ambient air. A recent review show that even at very low concentrations, BTEX exposure is associated with effects on immune, metabolic, respiratory, and reproductive functioning, as well as development (Bolden et al. 2015).

No previous monitoring program has measured VOCs at county level in Memphis area. REACT is the first census tract level community scale monitoring program targeting on VOCs. The work in this chapter aims to characterize the health risks from exposure to VOCs in Memphis area.

### 3.2 Method

The REACT air toxics monitoring program provided ambient monitoring concentrations of 71 VOCs at 112 sites in 106 census tracts throughout Shelby County, TN (Figure 3.1). The sites were selected to have the largest coverage of urbanization, emission sources, land-use type, and socioeconomic status (SES) for exploring spatial patterns for air toxic levels. From January 2014 through December 2014, air samples were collected at individual once every season to examine possible seasonal effects. Each measurement session was a continuous 24-hour period, and a 24-hour integrated air sample was obtained at the end of the session. All the VOCs were originally reported in part per billion (ppb), and this unit was converted to $\mu g/m^3$ following the U.S. EPA's "Air Toxics Data Analysis Workbook" (U.S. EPA, 2009b). Concentrations below the minimum detection limits (MDL) were replaced with half MDL. Duplicate samples were averaged if the percent difference was acceptable.

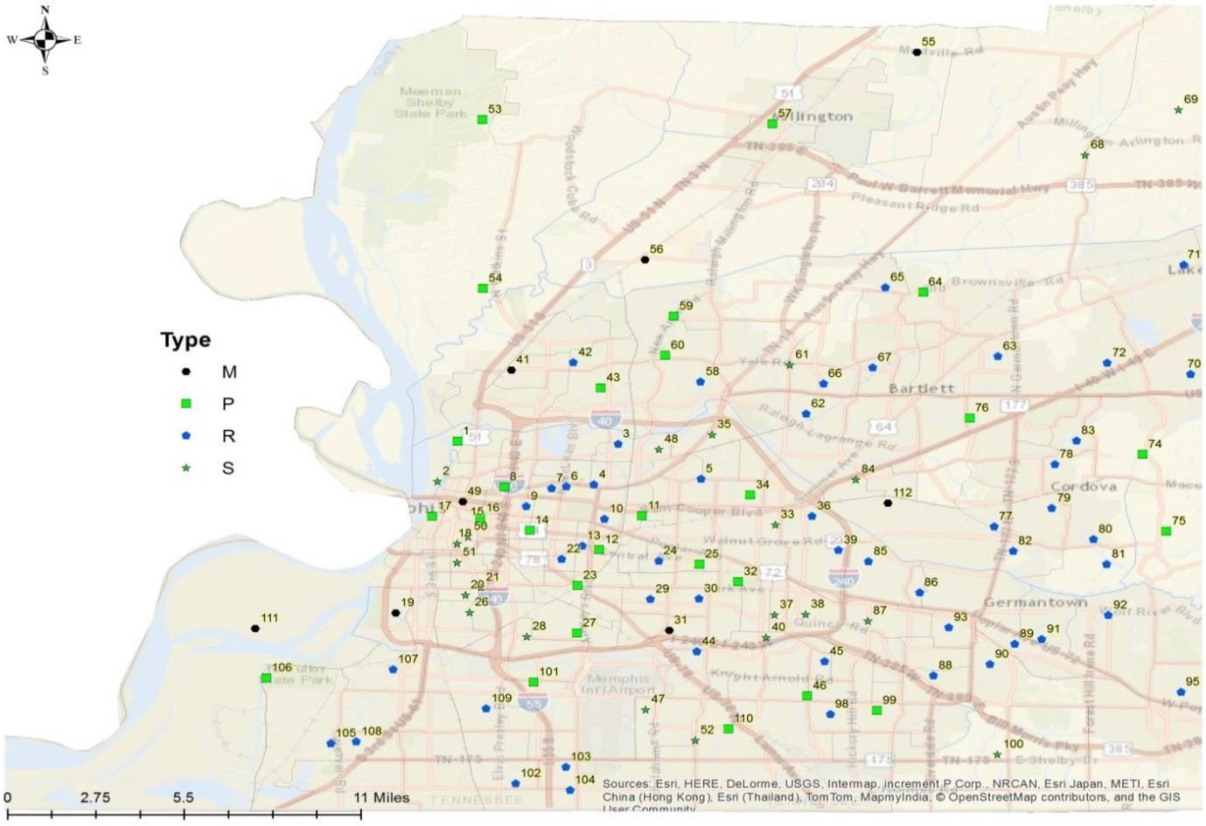


**Figure 3.1 REACT study air toxics monitoring sites.**

### 3.2.1 Descriptive statistics

Sample detection frequency (DF) for a certain VOC was defined as the percent of measurements above its MDL out of the total sample size. Site DF for a certain VOC was defined as the percent of sites that had detectable level of this compound out of 112 sites. Sample DFs were calculated for seasonal and annual data, and site DFs were calculated for the annual data. Other descriptive statistics included mean, standard deviation and percentiles were calculated for seasonal and annual data. Annual concentrations of all the detected compounds in this study were then compared with the national data. The U.S. EPA's Urban Air Toxics Monitoring Program (UATMP) measures VOC concentrations at 59 monitoring stations nationally. We obtained the annual data for Year 2014 from EPA's Air Quality System (AQS)

(U.S.EPA, 2014). We made two comparisons: the county average vs. the national average; and the county median vs. the national median. The former comparison used two-sample t-tests, and the latter used Mann-Whitney tests.

### 3.2.2 Spatial and temporal variation

In Shelby County, the city of Memphis is largely an urban setting, and the rest areas of the county are suburban. Thus, we classified a census tract as "urban" if it is within Memphis, and "suburban" if it is outside of Memphis. The mean and median of each compound were then compared using t-test and Wilcoxon rank sum test, respectively, to examine the difference between two settings. Differences were considered significant at $\alpha<0.05$. Seasonal means and medians of air toxic concentrations were tabulated and ranked to obtain an overall picture of the seasonal variation.

### 3.2.3 Source identification

The concentrations of many air toxics were highly correlated, implying that they were possibly emitted from the common sources. We used principle component and factor analysis to identify the common sources. Specifically, we applied varimax rotation to optimize factor analysis, and extracted factors for eigenvalues >1 with at least one variable with a loading >0.5. The emission sources, the factors, grouped in the analysis were then identified based on the grouping types of consisting air toxics (C. Jia, Batterman, & Godwin, 2008a).

### 3.2.4 Health risks estimation

Cancer and non-cancer risk of air toxics were estimated using inhalation risk assessment methods recommended in Air Toxics Risk Assessment Reference Library (Chapter 3) (U.S.EPA, 2015b), For inhalation exposures, chronic cancer risks for individual air toxics are typically

estimated by multiplying the estimate of long-term (70 years) exposure concentration (EC) by the corresponding Inhalation Unit Risk (IUR) for each pollutant to estimate the potential incremental cancer risk for an individual:

$$\text{Cancer Risk} = EC \times IUR \quad (1)$$

Where:

Cancer risk = the probability of getting cancer given a level and duration of exposure; unitless with value between 0 and 1;

Exposure Concentration = Personal exposure to a specific air toxic (in $\mu g/m^3$);

IUR = the upper-bound of the excess cancer risk estimated to result from continuous exposure to a concentration of 1 $\mu g/m^3$ of a compound over a 70-year lifetime.

Cancer risk is a probability, e.g., a risk level of 1 in a million ($10^{-6}$) implies a likelihood that up to one person, out of one million equally exposed people would contract cancer if exposed continuously (24 hours per day) to the specific concentration over 70 years (an assumed lifetime). This risk would be an excess cancer risk that is in addition to any cancer risk borne by a person not exposed to these air toxics. The use of upper-bound estimates for IURs overestimates the actual risks. Thus, this analysis is intended as a screening tool for risk managers and cannot make realistic predictions of biological effects.

Risks from simultaneous exposure to more than one carcinogenic substance are typically estimated by assuming that the individual risks are additive. The additive approach also treats all carcinogens as equal, despite potential differences in the underlying database (e.g., animal versus human data) or the weight of evidence for human carcinogenicity, i.e.,

$$Risk_T = Risk_1 + Risk_2 + ... + Risk_i \quad (2)$$

Where:

$Risk_T$ = total risk from exposure to all the chemicals of concern;

$Risk_i$ = individual risk estimate for the *i*th substance in the inhalation pathway.

Estimates of non-cancer risk are based on the assumption that there is a level of exposure below which it is unlikely to experience adverse health effects. A common method of evaluating non-cancer risks is to generate a "hazard quotient" (HQ), which represents the ratio of the exposure to the toxicity.

$$HQ = EC / RfC \quad (3)$$

Where:

HQ = non-cancer hazard quotient;

EC = estimate of chronic inhalation exposure to that air toxic (in $\mu g/m^3$);

RfC = the corresponding reference concentration for that air toxic (in $\mu g/m^3$).

Based on the definition of the RfC, an HQ less than or equal to one indicates that adverse non-cancer effects are not likely to occur, and thus can be considered to have negligible hazard. Unlike cancer risks, however, HQs should not be interpreted as a statistical probability of harm occurring. Instead, they are a simple statement of whether (and by how much) an exposure concentration exceeds the RfC. Moreover, the level of concern does not increase linearly or to the same extent as HQs increase above one for different chemicals because RfCs do not generally have equal accuracy or precision and are generally not based on the same severity of

effect. Thus, we can only say that with exposures increasingly greater than the RfC, (i.e., HQs increasingly greater than 1), the potential for adverse effects increases, but we do not know by how much. An HQ of 100 does not mean that the hazard is 10 times greater than an HQ of 10. Also an HQ of 10 for one substance may not have the same meaning (in terms of hazard) as another substance resulting in the same HQ.

Similar to carcinogens, risks from simultaneous exposure to more than one non-carcinogenic substance or from multiple exposure pathways are generally assumed to be additive by regulatory agencies. Specifically, these effects can be evaluated by summing the individual estimated HQs. The assumption of dose additivity is most appropriate to compounds that induce the same effect by similar modes of action. Thus, EPA guidance for chemical mixtures suggests subgrouping pollutant-specific HQs by toxicological similarity of the pollutants for subsequent calculations; that is, calculating a target-organ-specific-hazard index (TOSHI) for each subgrouping of pollutants. This calculation allows for a more appropriate estimate of overall hazard (USEPA, 2006).

$$TOSHI = HQ_1 + HQ_2 + \ldots + HQ_i \quad (4)$$

Where: $HQ_i$ = hazard quotient for the *i*th air toxic.

## 3.3 Results and discussion

### 3.3.1 Detection of air toxics in Memphis

All of the 71 target VOCs were detected in Shelby County during the 1-year monitoring period (Table 3.1). The most frequently detected VOCs in Memphis were acetone, ethanol, Freon 112, and propene, with DFs of 93%, 78%, 68%, and 60%, respectively. Other than these four compounds, 30 compounds were detected in 10-60% of the samples. Most of these

compounds were aromatic compounds, ketones, and alkanes, e.g., benzene, methyl butyl ketone, and 2,2,4-trimethylpentane. Another 37 compounds were found in less than 10% of samples in the entire study (Figure 3.2). Seasonally, DFs were the highest in summer, followed by spring, fall, and winter.

**Table 3.1 Sample and site detection frequencies (DFs, %) of target VOCs**

| VOCs | Sample DFs | | | | | Site DFs |
|---|---|---|---|---|---|---|
| | Winter | Spring | Summer | Fall | Annual | |
| 1,1,1-Trichloroethane | 0 | 0 | 0 | 0 | 0 | 2 |
| 1,1,2,2-Tetrachloroethane | 0 | 0 | 4.62 | 2.17 | 2.46 | 9 |
| 1,1,2-Trichloroethane | 0 | 0 | 0 | 0 | 0 | 2 |
| 1,1-Dibromoethane | 0 | 0 | 2.31 | 0 | 1.12 | 4 |
| 1,1-Dichloroethane | 0 | 0 | 0 | 0 | 0 | 2 |
| 1,1-Dichloroethene | 0 | 0 | 0.77 | 0 | 0.67 | 3 |
| 1,2,4-Trichlorobenzene | 0 | 100 | 15.4 | 0.72 | 36.2 | 97 |
| 1,2,4-Trimethylbenzene | 0 | 18.7 | 86.2 | 33.3 | 35.3 | 95 |
| 1,2-Dichlorobenzene | 0.75 | 1.49 | 8.46 | 2.90 | 4.24 | 16 |
| 1,2-Dichloroethane | 0.75 | 0 | 0 | 0 | 0.67 | 3 |
| 1,2-Dichloropropane | 0 | 0 | 0 | 0 | 0 | 2 |
| 1,3,5-Trimethylbenzene | 0 | 0.75 | 15.4 | 0 | 4.46 | 18 |
| 1,3-Butadiene | 0 | 0 | 0 | 35.5 | 10.5 | 42 |
| 1,3-Dichlorobenzene | 0 | 2.99 | 40.0 | 1.45 | 11.6 | 44 |
| 1,4-Dichlorobenzene | 0 | 0.75 | 7.69 | 2.90 | 3.57 | 13 |
| 2,2,4-Trimethylpentane | 0 | 1.49 | 90.8 | 10.1 | 26.3 | 92 |
| 2-Chlorotoulene | 0 | 0 | 2.31 | 0 | 1.12 | 4 |
| 4-Ethyltoulene | 0 | 6.72 | 76.9 | 4.35 | 22.5 | 78 |
| Acetone | 88.8 | 88.8 | 100 | 83.3 | 93.5 | 100 |
| Allyl chloride | 13.4 | 36.6 | 97.7 | 0 | 37.3 | 98 |
| Benzene | 0.75 | 1.49 | 97.7 | 24.6 | 31.3 | 95 |
| Benzyl chloride | 1.49 | 100 | 3.85 | 2.90 | 26.1 | 95 |
| Bromodichloromethane | 0 | 0 | 0 | 0 | 0 | 2 |
| Bromoethene | 0 | 0 | 0 | 0 | 0 | 2 |
| Bromoform | 0 | 0 | 3.08 | 0 | 1.34 | 2 |
| Bromomethane | 0 | 0 | 0 | 0 | 0 | 4 |
| Carbon disulfide | 1.49 | 3.73 | 2.31 | 7.25 | 4.46 | 18 |
| Carbon tetrachloride | 0 | 1.49 | 0.77 | 36.2 | 11.8 | 46 |
| Chlorobenzene | 0 | 0 | 6.15 | 0 | 2.01 | 3 |
| Chlorodibromomethane | 0 | 0 | 0.77 | 0 | 0.67 | 2 |

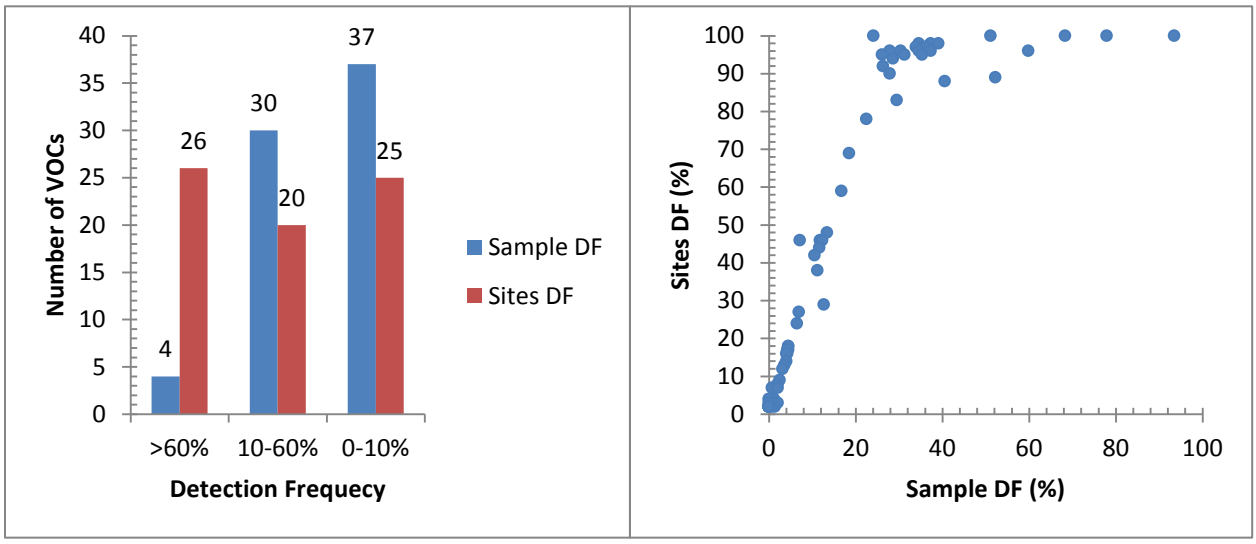


**Figure 3.2 Numbers of VOCs detected at the sampling sites in Shelby County**

### 3.3.2 Concentrations of air toxics in Memphis

The average total VOC (TVOC) concentration (± standard deviation) was 132 (±46.5) $\mu g/m^3$, ranging from 86.2 $\mu g/m^3$ to 482 $\mu g/m^3$. Specifically, the top five compounds with the highest county-average concentrations were ethanol, acetone, isopropyl alcohol, allyl chloride, and naphthalene. Ethanol had the highest mean concentration of 14.4±16.8 $\mu g/m^3$ and the second highest annul detection frequency (78%). Ethanol is commonly mixed with gasoline for car engine fuel. Acetone had the second highest mean concentration of 10.7± 5.45 $\mu g/m^3$ with the highest annual detection frequency (94%). Acetone exists naturally in plants, trees, but it is also from anthropogenic sources such as manufacture releases and vehicle exhaust. Isopropyl alcohol had a mean concentration of 6.8 ± 5.89 $\mu g/m^3$ annually and had highest average concentration in spring among all target compounds. Isopropyl alcohol is a major ingredient in fuel additives to reduce moisture in engine and is also used as intermediate and solvent in industry. Allyl chloride had a mean concentration of 4.42 ± 22.39 $\mu g/m^3$.It is usually used in synthesis as an alkylating agent or catalyst and modifier. It is also useful in the manufacture of pharmaceuticals and

pesticides. Naphthalene had a mean concentration of 4.2 ± 1.99 μg/$m^3$. Naphthalene in ambient air is majorly due to combustion from production and process facilities, open burning, tailpipe emissions, and cigarettes. Naphthalene's use as a deodorizer, repellent and fumigant is another source. Ten compounds had average concentrations between 2 μg/$m^3$ and 4 μg/$m^3$, including 1,2,4-trichlorobenzene (3.91± 1.84 μg/$m^3$), Freon 113 (3.05± 0.29μg/$m^3$), 1,1,1-trichloroethane (2.94± 0.29μg/$m^3$), 4-ethyltoulene (2.76 ± 0.42 μg/$m^3$), Freon 112 (2.74 ± 0.42 μg/$m^3$), hexachloro-1,3-butadiene (2.46 ± 1.93 μg/$m^3$), benzyl chloride (2.27± 2.85 μg/$m^3$), methyl butyl ketone (2.24 ± 2.48 μg/$m^3$), chloroform (2.03 ± 2.68 μg/$m^3$), methylene chloride (2.01 ± 4.69 μg/$m^3$). Out of the rest 56 compounds, 35 had concentrations of 1-2 μg/$m^3$, and 21 below 1 μg/$m^3$.

Memphis has higher air toxics levels than the national urban levels. We obtained VOC concentrations measured in 2014 from EPA's Air Quality System (AQS). There are 123 monitoring stations in the U.S. that measured VOCs in 2014. We determined whether each station was located in urban or rural setting based on 1999 National-scale Air Toxics Assessment (NATA) table (NATA, 1999), and identified 83 urban stations. The mean and median concentrations of 15 key VOCs were computed based on data from these 83 stations. Figure 3.3 displays the comparisons of key VOC concentrations between AQS and REACT.

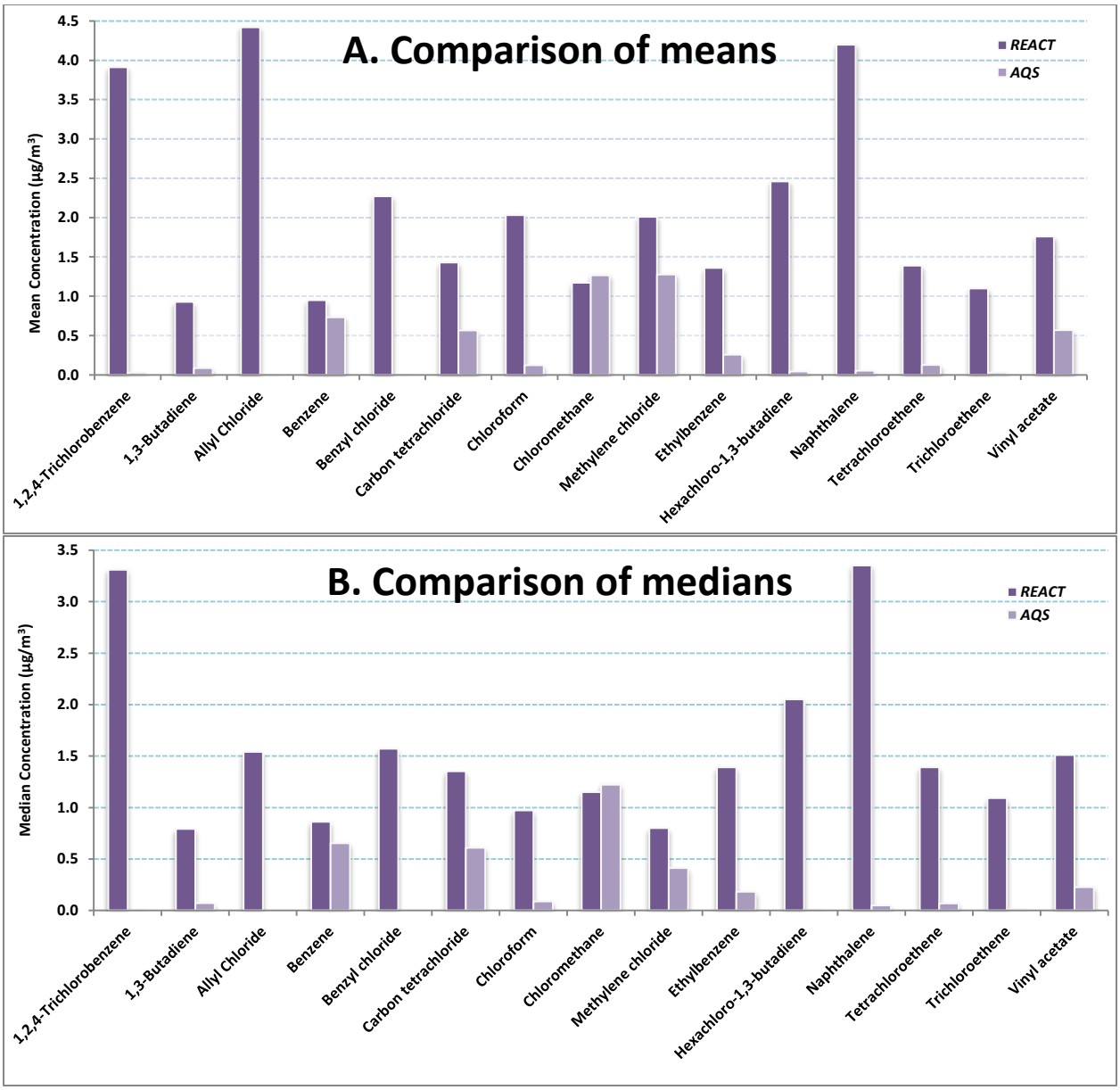


**Figure 3.3 Comparison of air toxics levels between the REACT study and AQS.**

The first comparison was based on mean concentrations. The mean concentrations in REACT were all higher than the national levels, except for benzene and chloromethane (Figure 3.3 A). REACT levels were 2 times (methylene chloride) to 2,035 times (allyl chloride) than AQS levels. In particular, two groups of compounds showed much higher levels in the REACT study. The first group is chlorinated compounds, including allyl chloride (2,035 times), benzyl chloride

(707 times), 1,2,4-trichlorobenzene (147 times), trichloroethene (39 times), chloroform (17 times) and methylene chloride (2 times).The second group is mainly associated with vehicle combustion, including naphthalene (76 times) and 1,3-butadiene (11 times).

The second comparison was based on median concentrations. The median concentrations in REACT were all higher than the national levels, except that benzene and chloromethane were similar (Figure 3.3 B). The REACT study had allyl chloride (1.54 $\mu g/m^3$) and benzyl chloride (1.57 $\mu g/m^3$) with median concentrations at the detectable levels, while these compounds had median concentrations below the detection limits in AQS. Among the rest 11 compounds, REACT levels were 2 times (methylene chloride) to 4,156 times (hexachloro-1,3-butadiene) than AQS levels.

Both comparisons indicate Memphis has higher air toxics levels than the national urban levels. The elevated concentrations of these compounds likely reflect the increasing intensity of industrial and mobile sources in the urban setting in Memphis.

### 3.3.3 Spatial and temporal variation

The comparison between urban (Memphis) and suburban (Shelby excluding Memphis) showed that 11 compounds had significantly higher concentrations in urban areas. (Table 3.2). Other target compounds did not display urban/suburban differences. The 11 compounds with significant differences could be classified into two groups: those from gasoline vapors, vehicle exhausts, and gasoline additive (e.g., n-heptane, 1,3-butadiene, and MTBE), and those from industrial solvents (e.g., 1,1-dichloroethane and methyl methacrylate).This difference might be the result of more intensive traffic and industrial sources in the inner city of Memphis.

**Table 3.2 Comparison of air toxic concentrations between urban and suburban areas**

| VOCs | Urban | | Suburban | | t-test |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | p-value |
| Propene | 1.47 | 0.75 | 1.17 | 0.32 | <0.01 |
| Freon 112 | 2.70 | 0.66 | 2.89 | 0.70 | 0.20 |
| Chloromethane | 1.07 | 0.42 | 1.54 | 1.79 | 0.15 |
| Freon 114 | 1.77 | 0.05 | 1.72 | 0.17 | 0.10 |
| Chloroethene | 1.11 | 0.11 | 1.06 | 0.15 | 0.09 |
| 1,3-Butadiene | 0.98 | 0.23 | 0.83 | 0.17 | <0.01 |
| Bromomethane | 0.89 | 0.02 | 0.86 | 0.08 | 0.11 |
| Chloroethane | 0.60 | 0.01 | 0.58 | 0.05 | 0.10 |
| Ethanol | 16.6 | 20.80 | 10.4 | 7.11 | 0.02 |
| Bromoethene | 0.98 | 0.02 | 0.95 | 0.09 | 0.10 |
| Acetone | 11.0 | 5.81 | 10.7 | 5.77 | 0.84 |
| Freon 11 | 1.60 | 0.20 | 1.63 | 0.29 | 0.58 |
| Isopropyl alcohol | 7.01 | 5.26 | 6.71 | 9.52 | 0.87 |
| 1,1-Dichloroethene | 0.86 | 0.01 | 0.84 | 0.08 | 0.18 |
| Methylene chloride | 1.41 | 1.74 | 3.51 | 8.23 | 0.16 |
| Allyl chloride | 1.86 | 1.45 | 13.3 | 54.90 | 0.25 |
| Freon 113 | 3.04 | 0.30 | 2.90 | 0.62 | 0.22 |
| Carbon disulfide | 0.69 | 0.19 | 2.35 | 8.14 | 0.26 |
| cis-1,2-Dichloroethene | 0.99 | 0.03 | 0.96 | 0.10 | 0.13 |
| 1,1-Dichloroethane | 0.86 | 0.02 | 0.83 | 0.08 | 0.10 |
| Methyl tert-butyl ether | 0.83 | 0.25 | 0.67 | 0.06 | <0.01 |
| Vinyl acetate | 1.81 | 1.12 | 1.70 | 0.83 | 0.62 |
| Ethyl Methyl Ketone | 2.09 | 1.22 | 1.82 | 0.59 | 0.13 |
| trans-1,2-Dichloroethene | 0.93 | 0.04 | 0.90 | 0.11 | 0.13 |
| n-Hexane | 1.10 | 0.48 | 0.99 | 0.48 | 0.29 |
| Ethyl acetate | 1.48 | 2.07 | 1.06 | 0.64 | 0.11 |
| Chloroform | 2.16 | 2.94 | 1.74 | 2.02 | 0.39 |
| Tetrahydrofuran | 0.71 | 0.29 | 0.64 | 0.15 | 0.10 |
| Naphthalene | 4.08 | 2.12 | 3.78 | 1.80 | 0.48 |
| Hexachloro-1,3-butadiene | 2.29 | 2.06 | 2.12 | 1.92 | 0.69 |
| TVOCs | 132 | 35.8 | 135 | 91.2 | 0.88 |

Notes: p-values <0.05 were highlighted.

The effects of urbanization/industrialization have been observed as an influential factor of ambient VOC levels (C. Jia et al., 2008a, 2008b). Traffic density is a major factor influencing concentrations of especially aromatic VOCs (Ilgen et al., 2001; Mohamed, Kang, & Aneja, 2002). The RIOPA study found that concentrations of BTEX, MTBE and PERC were inversely associated with distances to major roadways, gas stations and dry cleaning facilities, respectively, and that levels were also inversely associated with atmospheric stability, wind speed, temperature and humidity (Kwon et al., 2006). Dramatically elevated VOC concentrations have been found near large industrial facilities such as oil refineries (Cetin, Odabasi, & Seyfioglu, 2003) and industrial complexes (Park & Jo, 2004). We inferred that proximity to general traffic and industrial emissions in Memphis elevated ambient VOC concentrations.

Concentrations of ambient air toxics displayed significant seasonal variations (Tables 3.3 & 3.4). In Table 3.3, 56 and 52 compounds had the highest or 2$^{nd}$ highest mean concentrations in winter and spring, respectively. In contrast, 48 and 60 compounds had the lowest or 2$^{nd}$ lowest mean concentrations in summer and fall, respectively. We can group winter and spring into the cold season, and summer and fall into the warm season. Thus, most compounds showed higher average concentrations in the cold season than in the warm season. If we examine the median concentrations, very similar results were obtained (Table 3.4). The bar charts in Figure 3.4 clearly indicate higher concentrations of key air toxics in winter and spring.

**Table 3.3 Mean and rank of air toxic concentrations by season**

| VOCs | Mean (µg/m3) | | | | Rank of Means | | | |
|---|---|---|---|---|---|---|---|---|
| | Win | Spr | Sum | Fall | Win | Spr | Sum | Fall |
| Propene | 1.8 | 0.6 | 0.8 | 2.3 | 2 | 4 | 3 | 1 |
| Freon 112 | 3.2 | 1.4 | 1.5 | 4.8 | 2 | 4 | 3 | 1 |
| Chloromethane | 1.4 | 0.8 | 0.8 | 1.8 | 2 | 3 | 4 | 1 |
| Freon 114 | 2.8 | 2.6 | 0.8 | 0.9 | 1 | 2 | 4 | 3 |
| Chloroethene | 1.5 | 2.2 | 0.4 | 0.3 | 2 | 1 | 3 | 4 |
| 1,3-Butadiene | 1.2 | 1.2 | 0.4 | 0.9 | 2 | 1 | 4 | 3 |
| Bromomethane | 1.3 | 1.3 | 0.6 | 0.4 | 2 | 1 | 3 | 4 |
| Chloroethane | 0.8 | 0.8 | 0.4 | 0.4 | 1 | 2 | 3 | 4 |
| Ethanol | 9.8 | 3.2 | 31.1 | 13.6 | 3 | 4 | 1 | 2 |
| Bromoethene | 1.4 | 1.4 | 0.8 | 0.2 | 2 | 1 | 3 | 4 |
| Acetone | 7.7 | 6.7 | 18.6 | 9.9 | 3 | 4 | 1 | 2 |
| Freon 11 | 1.9 | 2.4 | 0.5 | 1.6 | 2 | 1 | 4 | 3 |
| Isopropyl alcohol | 3.1 | 20.2 | 3.3 | 0.6 | 3 | 1 | 2 | 4 |
| 1,1-Dichloroethene | 1.3 | 1.1 | 0.5 | 0.4 | 1 | 2 | 3 | 4 |
| Methylene chloride | 5.3 | 1.6 | 0.5 | 0.6 | 1 | 2 | 4 | 3 |
| Allyl chloride | 1.7 | 10.1 | 5.8 | 0.2 | 3 | 1 | 2 | 4 |
| Freon 113 | 2.6 | 8.6 | 0.3 | 0.8 | 2 | 1 | 4 | 3 |
| Carbon disulfide | 1.0 | 2.2 | 0.7 | 0.3 | 2 | 1 | 3 | 4 |
| cis-1,2-Dichloroethene | 1.2 | 1.6 | 0.7 | 0.4 | 2 | 1 | 3 | 4 |
| 1,1-Dichloroethane | 1.3 | 1.3 | 0.4 | 0.4 | 1 | 2 | 4 | 3 |
| Methyl tert-butyl ether | 1.2 | 0.7 | 0.9 | 0.3 | 1 | 3 | 2 | 4 |
| Vinyl acetate | 1.1 | 1.0 | 2.1 | 2.8 | 3 | 4 | 2 | 1 |
| Ethyl Methyl Ketone | 0.8 | 1.0 | 4.1 | 2.0 | 4 | 3 | 1 | 2 |
| trans-1,2-Dichloroethene | 1.3 | 1.6 | 0.4 | 0.4 | 2 | 1 | 4 | 3 |
| n-Hexane | 1.5 | 1.2 | 1.3 | 0.3 | 1 | 3 | 2 | 4 |
| Ethyl acetate | 2.1 | 1.7 | 1.2 | 0.3 | 1 | 2 | 3 | 4 |
| Chloroform | 4.5 | 2.3 | 0.5 | 0.8 | 1 | 2 | 4 | 3 |
| Tetrahydrofuran | 0.9 | 0.8 | 0.8 | 0.2 | 1 | 2 | 3 | 4 |
| Hexachloro-1,3-butadiene | 3.1 | 4.2 | 1.1 | 1.4 | 2 | 1 | 4 | 3 |
| TVOCs | 141.4 | 171.7 | 136.4 | 79.7 | 2 | 1 | 3 | 4 |
| Sum of Rank #1 | | | | | 33 | 20 | 14 | 5 |
| Sum of Rank #2 | | | | | 23 | 32 | 10 | 7 |
| Sum of Rank #3 | | | | | 13 | 11 | 29 | 19 |
| Sum of Rank #4 | | | | | 3 | 9 | 19 | 41 |

**Table 3.4 Median and rank of air toxic concentrations by season**

| VOCs | Median (µg/m3) | | | | Median Rank | | | |
|---|---|---|---|---|---|---|---|---|
| | Win | Spr | Sum | Fall | Win | Spr | Sum | Fall |
| Propene | 1.6 | 0.5 | 0.7 | 2.3 | 2 | 4 | 3 | 1 |
| Freon 112 | 3.3 | 1.4 | 1.6 | 4.7 | 2 | 4 | 3 | 1 |
| Chloromethane | 1.0 | 0.5 | 0.7 | 2.1 | 2 | 4 | 3 | 1 |
| Freon 114 | 2.8 | 2.7 | 0.8 | 0.9 | 1 | 2 | 4 | 3 |
| Chloroethene | 1.5 | 2.3 | 0.4 | 0.3 | 2 | 1 | 3 | 4 |
| 1,3-Butadiene | 1.2 | 1.3 | 0.4 | 0.3 | 2 | 1 | 3 | 4 |
| Bromomethane | 1.3 | 1.3 | 0.6 | 0.4 | 2 | 1 | 3 | 4 |
| Chloroethane | 0.8 | 0.8 | 0.4 | 0.4 | 1 | 2 | 3 | 4 |
| Ethanol | 1.5 | 0.7 | 16.7 | 12.8 | 3 | 4 | 1 | 2 |
| Bromoethene | 1.4 | 1.5 | 0.8 | 0.2 | 2 | 1 | 3 | 4 |
| Acetone | 5.2 | 2.2 | 13.8 | 8.6 | 3 | 4 | 1 | 2 |
| Freon 11 | 1.9 | 2.4 | 0.5 | 1.9 | 2 | 1 | 4 | 3 |
| Isopropyl alcohol | 0.4 | 12.0 | 2.2 | 0.3 | 3 | 1 | 2 | 4 |
| 1,1-Dichloroethene | 1.3 | 1.1 | 0.5 | 0.4 | 1 | 2 | 3 | 4 |
| Methylene chloride | 1.1 | 1.2 | 0.5 | 0.4 | 2 | 1 | 3 | 4 |
| Allyl chloride | 1.0 | 1.0 | 3.4 | 0.2 | 2 | 3 | 1 | 4 |
| Freon 113 | 2.6 | 8.8 | 0.3 | 0.8 | 2 | 1 | 4 | 3 |
| Carbon disulfide | 1.0 | 1.0 | 0.4 | 0.3 | 1 | 2 | 3 | 4 |
| cis-1,2-Dichloroethene | 1.2 | 1.6 | 0.7 | 0.4 | 2 | 1 | 3 | 4 |
| 1,1-Dichloroethane | 1.3 | 1.3 | 0.4 | 0.4 | 1 | 2 | 4 | 3 |
| Methyl tert-butyl ether | 1.2 | 0.7 | 0.5 | 0.3 | 1 | 2 | 3 | 4 |
| Vinyl acetate | 1.1 | 1.1 | 0.6 | 2.9 | 2 | 3 | 4 | 1 |
| Ethyl Methyl Ketone | 0.8 | 0.8 | 3.5 | 2.1 | 3 | 4 | 1 | 2 |
| Naphthalene | 6.8 | 2.4 | 2.5 | 1.5 | 1 | 3 | 2 | 4 |
| Hexachloro-1,3-butadiene | 3.1 | 2.7 | 0.9 | 1.4 | 1 | 2 | 4 | 3 |
| TVOCs | 121.3 | 137.6 | 110.3 | 75.3 | 2 | 1 | 3 | 4 |
| Sum of Rank #1 | | | | | 34 | 19 | 15 | 5 |
| Sum of Rank #2 | | | | | 26 | 35 | 6 | 5 |
| Sum of Rank #3 | | | | | 11 | 11 | 31 | 19 |
| Sum of Rank #4 | | | | | 1 | 7 | 20 | 43 |

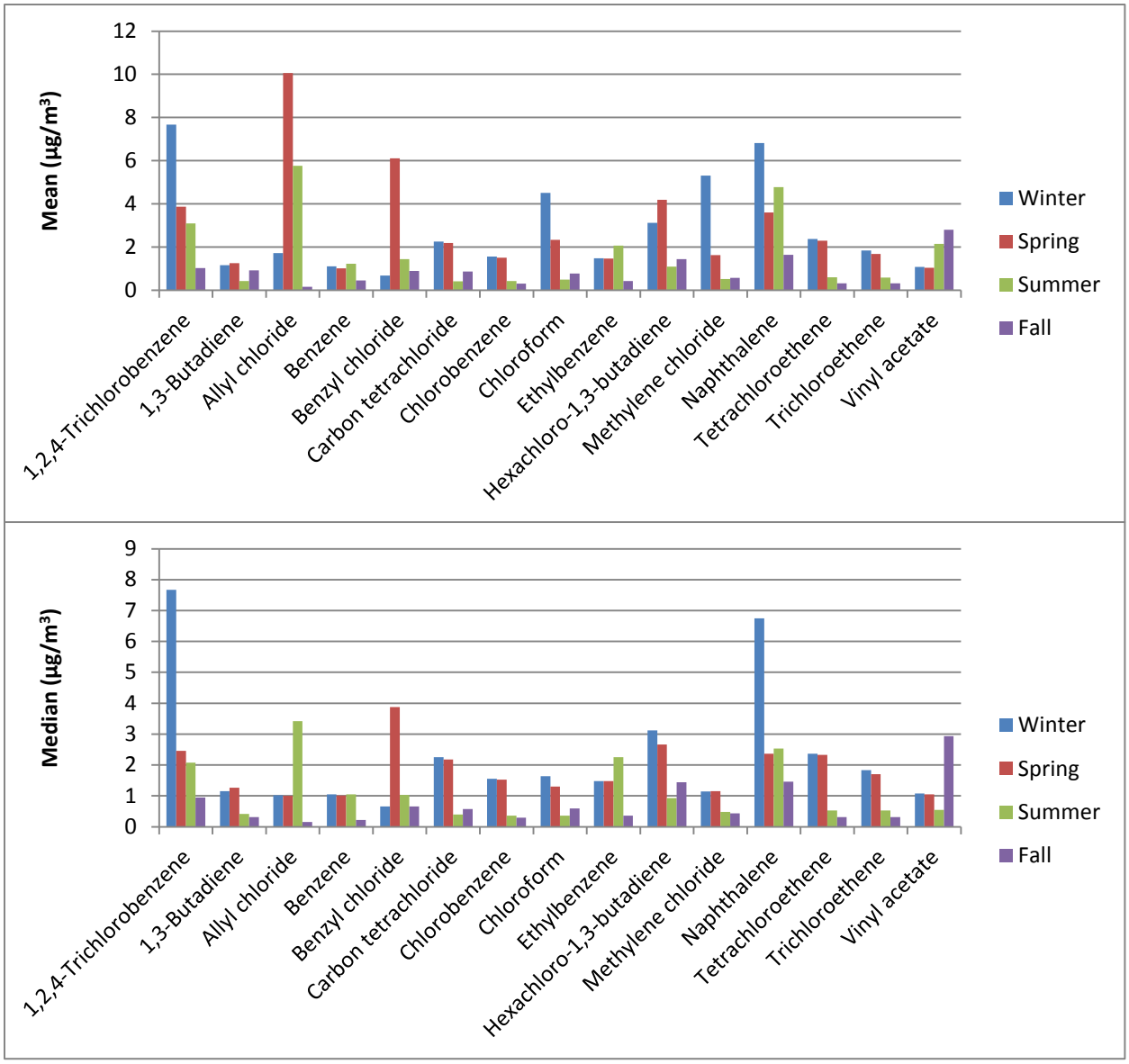


**Figure 3.4 Comparison of seasonal concentrations of key air toxics**

Elevated VOC concentrations in cold weather have been reported in many studies (Cheng, Fu, Angle, & Sandhu, 1997; Das & Aneja, 2003; Ho, Lee, Guo, & Tsai, 2004; Mohamed et al., 2002). Many factors affect seasonal patterns, but generally cooler temperatures slow rates of photochemical reactions (Mohamed et al., 2002), increase emissions from heating sources and vehicles (Mohamed et al., 2002), and lower mixing heights (Cheng et al., 1997). Seasonal changes may be negligible for VOCs that result primarily from nearby industrial

sources (Cheng et al., 1997), especially for slowly reacting compounds, e.g., many halogens (Mohamed et al., 2002). Our results matched the seasonal patterns observed in other U.S. studies.

### 3.3.4 Source identification

The factor analysis of annual concentrations of 24 frequently detected (DF > 80%) air toxics revealed 5 factors (Table 3.5). These five factors explained 72% of the total variance of the data. The first factor included hexachloro-1,3-butadiene, 1,2,4-trimethylbenzene, n-butylbenzene, benzyl chloride, methyl butyl ketone, naphthalene, and methyl isobutyl ketone. This factor was likely to reflect emissions from organic synthesis processes. Factor 2 contained aromatic compounds including 2,2,4-trimethylpentane, 1,3,5-trimethylbenzene, toluene, benzene, o-xylene, m,p-xylene, ethylbenzene, and may reflect vehicle exhaust. The third factor included propene, acetone, ethyl methyl ketone, vinyl acetate and could be industrial solvent and precursors. Factor 4 contained chloromethane and Freon 112 which are refrigerants. Factor 5 included isopropyl alcohol, ethanol, and n-hexane, which are gasoline additives.

**Table 3.5 Factor patterns of major air toxics in Memphis**

| VOCs/Factors | F1 | F2 | F3 | F4 | F5 |
|---|---|---|---|---|---|
| Hexachloro-1,3-butadiene | **0.96** | -0.03 | -0.04 | 0.11 | -0.01 |
| 1,2,4-Trimethylbenzene | **0.95** | -0.02 | 0.05 | 0.09 | -0.02 |
| n- Butylbenzene | **0.94** | 0.04 | 0.06 | -0.03 | 0.07 |
| Benzyl chloride | **0.94** | -0.05 | 0.01 | 0.06 | -0.06 |
| Methyl butyl ketone | **0.92** | -0.02 | 0.07 | 0.08 | 0.00 |
| Naphthalene | **0.84** | 0.02 | 0.15 | 0.06 | 0.07 |
| Methyl isobutyl ketone | **0.72** | 0.08 | 0.06 | 0.11 | 0.39 |
| 2,2,4-Trimethylpentane | 0.07 | **0.80** | 0.12 | -0.15 | 0.01 |
| 1,3,5-Trimethylbenzene | 0.06 | **0.67** | 0.15 | -0.19 | 0.07 |
| Toulene | 0.05 | **0.77** | 0.20 | -0.05 | 0.34 |
| Benzene | 0.05 | **0.52** | 0.23 | 0.05 | 0.31 |
| o-Xylene | 0.00 | **0.90** | -0.05 | 0.03 | -0.04 |
| m,p-Xylene | -0.04 | **0.92** | -0.02 | 0.00 | -0.01 |
| Ethylbenzene | -0.16 | **0.86** | -0.11 | 0.11 | -0.09 |
| Propene | 0.10 | 0.32 | **0.55** | 0.41 | 0.04 |
| Acetone | 0.09 | -0.03 | **0.80** | 0.21 | 0.20 |
| Ethyl Methyl Ketone | 0.07 | 0.02 | **0.83** | -0.01 | 0.23 |
| Vinyl acetate | 0.02 | 0.16 | **0.77** | -0.17 | -0.06 |
| Chloromethane | 0.23 | -0.14 | 0.08 | **0.87** | -0.03 |
| Freon 112 | 0.00 | -0.13 | 0.07 | **0.89** | -0.04 |
| Isopropyl alcohol | 0.21 | -0.02 | -0.01 | -0.06 | **0.74** |
| Ethanol | 0.07 | 0.03 | 0.50 | 0.01 | **0.61** |
| n-Hexane | -0.15 | 0.18 | 0.24 | 0.06 | **0.58** |
| Allyl chloride | 0.37 | 0.13 | -0.16 | 0.44 | 0.27 |
| Variance explained (%) | 26.5 | 20.6 | 11.8 | 7.9 | 5.1 |
| Cumulative (%) | 26.5 | 47.1 | 58.9 | 66.8 | 71.9 |

### 3.3.5 Health risks from exposure to air toxics

According to EPA's guidelines, an excess lifetime cancer risk below the range of $10^{-6}$ to $10^{-4}$ can be considered acceptable. Twenty-two carcinogenic VOCs were detected during 2014 (Table 3.6). At the county level, the cumulative cancer risk based on mean ambient concentration at each site was 5.9± 3.3 ×$10^{-4}$, meaning that 590 persons out of 1 million are expected to contract cancers due to the exposure to all carcinogenic air toxics over their lifetime. Given the

population size of around 1 million in Memphis, this represents approximately the additional number of cancer cases resulting from air toxics pollution. The average ambient levels of eight compounds exceeded the $10^{-6}$ criterion, nine compounds exceeded the $10^{-5}$ criterion, and two compounds (naphthalene and benzyl chloride) exceeded the $10^{-4}$ criterion. The rest three carcinogenic compounds (tetrachloroethene, methylene chloride, and vinyl acetate) had both average and maximum concentrations that were associated with $<10^{-6}$ risk levels.

**Table 3.6 Cancer risks from exposure to carcinogenic air toxics in Memphis**

| Carcinogenic VOCs | UR | Mean | SD | Min | P50 | P95 | Max | Contrib. |
|---|---|---|---|---|---|---|---|---|
| | $(1/(\mu g/m^3)*10^6)$ | $(1/10^6)$ | | | | | | (%) |
| Naphthalene | 34 | 143 | 68 | 103 | 114 | 310 | 440 | 24.35 |
| Benzyl chloride | 49 | 111 | 139 | 59.3 | 77.1 | 215 | 974 | 18.99 |
| 1,1,2,2-Tetrachloroethane | 58 | 55.7 | 8.75 | 36.5 | 53.5 | 77.3 | 97.7 | 9.495 |
| Chloroform | 23 | 46.6 | 61.6 | 15.2 | 22.4 | 171 | 389 | 7.935 |
| Hexachloro-1,3-butadiene | 22 | 54.1 | 42.5 | 39.5 | 45.0 | 68.9 | 321 | 9.217 |
| Bromoethene | 32 | 31.2 | 1.50 | 20.2 | 31.4 | 31.4 | 31.4 | 5.322 |
| Allyl chloride | 6 | 26.5 | 134 | 4.04 | 9.22 | 28.5 | 1392 | 4.522 |
| 1,3-Butadiene | 30 | 28.0 | 6.75 | 14.3 | 23.6 | 41.9 | 50.9 | 4.778 |
| 1,2-Dichloroethane | 26 | 19.1 | 1.47 | 14.7 | 19.0 | 19.0 | 33.2 | 3.253 |
| 1,1,2-Trichloroethane | 16 | 17.6 | 0.92 | 10.8 | 17.7 | 17.7 | 17.7 | 2.998 |
| 1,4-Dichlorobenzene | 11 | 12.3 | 3.39 | 8.2 | 11.1 | 22.1 | 24.8 | 2.093 |
| Chloroethene | 8.8 | 9.74 | 0.93 | 4.82 | 9.77 | 9.77 | 16.7 | 1.660 |
| Carbon tetrachloride | 6 | 8.56 | 1.13 | 4.94 | 8.12 | 9.44 | 17.6 | 1.458 |
| Benzene | 7.8 | 7.39 | 2.00 | 5.15 | 6.74 | 11.3 | 20.0 | 1.258 |
| Trichloroethene | 4.1 | 4.50 | 0.27 | 2.79 | 4.48 | 4.94 | 5.13 | 0.766 |
| trans-1,3-Dichloropropene | 4 | 4.04 | 0.19 | 2.61 | 4.07 | 4.07 | 4.07 | 0.689 |
| Ethylbenzene | 2.5 | 3.39 | 0.44 | 2.34 | 3.48 | 3.88 | 4.31 | 0.578 |
| Bromoform | 1.1 | 2.03 | 0.18 | 1.22 | 2.01 | 2.01 | 3.18 | 0.346 |
| 1,1-Dichloroethane | 1.6 | 1.36 | 0.07 | 0.88 | 1.37 | 1.37 | 1.37 | 0.232 |
| Tetrachloroethene | 0.26 | 0.36 | 0.02 | 0.21 | 0.36 | 0.40 | 0.41 | 0.062 |
| Methylene chloride | 0.01 | 0.02 | 0.05 | 0.01 | 0.01 | 0.08 | 0.43 | 0.003 |
| Vinyl acetate | 0.01 | 0.02 | 0.01 | 0.01 | 0.02 | 0.03 | 0.10 | 0.003 |
| **Cumulative** | | 587 | 327 | 420 | 494 | 871 | 2888 | |

Notes: The unit for cancer risk is $1/10^6$; UR – Unit risk; SD– Standard deviation; P– percentage; Contrib. – contribution.

Thirty-one out of the 71 target compounds has available reference concentrations for estimating target organ specific hazard index (TOSHI). The top TOSHIs were 6.71, 1.84, and 1.03 for neurological, respiratory, and reproductive/developmental effects, respectively. These effects are the major non-cancer health concerns from exposure to mixtures of air toxics in Memphis. It should be noted that the neurological TOSHI was driven by allyl chloride (76%) and naphthalene (20%), respiratory TOSHI by naphthalene (72%), and reproductive/developmental TOSHI by trichloroethene (53%) and 1,3-butadiene (45%). Allyl chloride and trichoroethene are both widely used industrial solvents; the former is used in the manufacture of pharmaceuticals and pesticides, and the latter is used as a degreaser. However, the 2014 TRI data did not report any release of these two chemicals in Shelby County. Vehicle exhaust is a constant source of naphthalene and 1,3-butadiene. Thus, unreported industrial emissions and vehicle exhaust are the two major sources contributing to non-cancer effects in Memphis. On average, TOSHIs for effects in ocular system, immune system, cardiovascular system, hematologic system, liver and kidney were below 1. It was also encouraging that the maxima of these TOSHIs were all below 1. These facts indicate that long exposure is unlikely to cause these adverse health effects among all the populations in Shelby County (Table 3.7).

**Table 3.7 Target-organ-specific-hazard indices (TOSHIs)**

| Target organs | N | Mean | SD | Min | P5 | P25 | P50 | P75 | P95 | P99 | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Neurological | 112 | 6.71 | 29.7 | 1.85 | 2.25 | 2.66 | 2.95 | 3.60 | 6.57 | 60.4 | 313 |
| Respiratory | 112 | 1.84 | 0.68 | 1.41 | 1.43 | 1.48 | 1.61 | 1.93 | 3.44 | 4.54 | 4.77 |
| Reproductive/Developmental | 112 | 1.03 | 0.13 | 0.60 | 0.96 | 0.96 | 0.97 | 1.11 | 1.26 | 1.41 | 1.41 |
| Ocular | 112 | 0.61 | 0.04 | 0.38 | 0.51 | 0.61 | 0.61 | 0.61 | 0.67 | 0.69 | 0.69 |
| Immune | 112 | 0.58 | 0.04 | 0.36 | 0.55 | 0.57 | 0.58 | 0.58 | 0.63 | 0.66 | 0.66 |
| Cardiovascular | 112 | 0.55 | 0.04 | 0.34 | 0.45 | 0.55 | 0.55 | 0.55 | 0.61 | 0.62 | 0.63 |
| Liver and Kidney | 112 | 0.40 | 0.04 | 0.26 | 0.39 | 0.39 | 0.40 | 0.40 | 0.46 | 0.55 | 0.57 |
| Hematologic | 112 | 0.03 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.05 | 0.06 | 0.10 |
| **Total** | 112 | 11.8 | 29.9 | 6.44 | 6.97 | 7.44 | 7.74 | 8.84 | 14.5 | 65.6 | 320 |

**Notes:** ***Respiratory system*** – toluene, 1,2,4-trichlorobenzenem,p-xylene, o-xylene, trans-1,3-dichloropropene, bromomethane, 1,2-dichloropropane, naphthalene; ***Hematologic system*** – benzene; ***Cardiovascular system*** – methylene chloride, trichloroethene; ***Reproductive/Developmental system*** –chloroethane, methyl isobutyl ketone, ethylbenzene, isopropylbenzene, 1,2,4-trichlorobenzene, trichloroethene, 1,3-butadiene***; Liver and kidney*** – 1,1,1-trichloroethane, methyl tert-butyl ether, 1,2-dichloroethane, chlorobenzene, 1,4-dichlorobenzene, methyl methacrylate, methylene chloride, 1,1,2-trichloroethane, 1,2,4-trichlorobenzene, 1,1-dichloroethene, carbon tetrachloride, chloroethene, chloroform, bromoethene; ***Ocular*** – methyl tert-butyl ether, m,p-xylene, o-xylene, tetrachloroethene, trichloroethene; ***Neurological system*** – 1,1,1-trichloroethane, toluene, styrene, n-hexane, methylene chloride, 1,1,2-trichloroethane, m,p-xylene, o-xylene, chloromethane, tetrachloroethene, bromomethane, naphthalene, allyl chloride; ***Immune system*** – benzene, trichloroethene

### 3.3.6 Study limitations

The risk assessment is a process that integrates analysis and information of toxicity and exposure into a risk characterization that provides risk estimates. Risk assessments are often not as definitive in all aspects as would be desirable. Uncertainty exists in all aspects of a risk assessment process:

#### *3.3.6.1 Uncertainty in sampling*

The sampling locations were still limited considering the large area of Shelby County (784 square miles). Samples were not collected at several large census tracts. Within each tract, we could not always use the centroid due to accessibility issue. The sampling period was only one day which might not represent long-term exposure, although we repeated sampling seasonally. The field samples underwent inclement weather in certain days, e.g., storm, fog, rain,

and extremely cold and hot days. All these might influence the canister samples. We also did not know the unexpected sources during the sampling period, such as mowing, car idling, spray of pesticides, or other unknown human interference.

#### *3.3.6.2 Uncertainty in laboratory analysis*

There are many factors that could influence the analytical performance in laboratory. These could be the concentration of the analyte and proximity to MDL, the performance of the instrumentation on a given day, the consistency and skill of the analyst, the most recent instrument maintenance, influence of moisture in samples, and chromatographic data analysis skills. The uncertainty of each measurement varied by day, sample and analyst, precluding definitive estimates of uncertainty for each parameter in each sample for each day. Perhaps the greatest uncertainty associated with laboratory results occurred when the concentration of the analyte was close to the detection limit. As the concentration of the analyte decreases, the level of uncertainty will increase as the MDL is approached (C. Jia, Batterman, & Chernyak, 2006). A simple replacement of non-detects with half MDL might bias the concentration estimation.

#### *3.3.6.3 Uncertainty in exposure assessment*

The chronic exposure parameter was the central tendency estimate, obtained as the arithmetic average of the seasonal monitoring data. The intention was to base the exposure and risk estimates on the estimated long-term average exposure levels. The exposure assumptions were extremely simplified, and did not consider other exposure scenarios, e.g., indoor exposure. There was no accounting for the variability of human activity, or for the reasonableness of actual human presence at the monitoring sites. This approach is characteristic of a “screening level” type of approach.

#### *3.3.6.4 Uncertainty in toxicity assessment*

The health protective benchmarks used for chronic noncancer and cancer risk assessment were taken from the best available information from U.S.EPA (RfCs, RfDs, cancer UREs or slope factors). When those preferred U.S.EPA benchmark sources were not available, we used California EPA's toxicological data, which is a common practice in the U.S. (Miranda M. Loh, Jonathan I. Levy, John D. Spengler, E. Andres Houseman, & Deborah H. Bennett, 2007). U.S.EPA acknowledges the conservatism of these methodologies and the assumption of no threshold, by noting that the risk is unlikely to be higher than estimated, but is likely to be lower, and may be as low as zero at the extrapolated low doses for the target risk. Not all toxic VOCs were considered. Not all toxic VOCs had the health benchmarks available, and thus only a fraction of air toxics were considered in the risk assessment.

#### *3.3.6.5 Uncertainty in risk characterization*

Comparisons to absolute risk levels of $10^{-4}$, $10^{-5}$ and $10^{-6}$, or to HI values of 1, should be viewed qualitatively since these values function as protective guidelines, not absolute standards. The analysis only assumed additivity when evaluating risk from mixtures. The nature of the ambient air monitoring approach, rather than emissions modeling, does not enable the characterization of the population size or populated area that are represented by the risk estimates. Monitoring results are valid for the point of measurement and some surrounding area, but the size and shape of areas represented by the data are not known and cannot be reliably estimated**.**

### 3.4 Conclusion

This comprehensive field study expands the ambient VOC data routinely collected by existing fixed monitoring sites, and updates the old VOC databases. Few community-scale air

toxics monitoring studies have characterized the distribution and variability of VOC levels in a comprehensive and validated manner. None have explored ambient air toxic concentrations across suburban, urban and industrial communities. The uniqueness of this study is to deploy samplers at the census tract level, which allows linking exposure to socioeconomic status to address environmental justice. Thus, this work fills the data gap by delineating the spatial patterns and temporal trends and by exploring factors that affect VOC concentrations.

This study represents an evolutionary advance from previous studies designed to assess human exposures to ambient air toxics. It is one of the few comprehensive studies examining VOC exposures in the US since 2010. The representativeness of the study is enhanced by having the target communities located along an urban/industrial gradient, using multiple sites (112) at the census tract level, and employing a multi-factor (indoor/outdoor, seasonal, etc) sampling plan. The results of this study will help fill gaps in our understanding of VOC exposures, improve exposure assessments for urban air toxics, more realistically evaluate public health risks, potentially improve the management of hazard air pollutants and other pollutants, and help provide the basis for setting standards.

# Chapter 4

# Sociodemographic disparities in cancer risks from exposure to air toxics in Memphis Area

## 4.1 Introduction

Ambient air quality has been prioritized for environment and human health. Healthy People 2020 set outdoor air quality as the first priority for creating health-promoting environment (ODPHP, 2017). Elevated level of air pollution from various emission sources has been affecting both environment and human health. Short- and long- term exposures to elevated level of ambient hazardous air pollutants (HAPs), or air toxics, can trigger various, acute or chronic, adverse health effects including cancer (Bostrom et al., 2002; M. M. Loh et al., 2007), respiratory disease (Goldizen et al., 2016), neurological toxicity (Block et al., 2012; Calderon-Garciduenas et al., 2004; Volk et al., 2011), cardiovascular disease (Uzoigwe et al., 2013), negative reproductive effects or birth defects (Lewtas, 2007; Sram et al., 2005; Stieb et al., 2012) and renal toxicity (Damek-Poprawa & Sawicka-Kapusta, 2003; Jarup, 2003; Kampa & Castanas, 2008).

Previous studies have raised the concerns that minorities and communities of lower socioeconomic status were more likely to live close to emission sources (Stuart et al., 2009) exposing to elevated level of air toxics and thus having increased health risks (Abel, 2008; Chakraborty, Collins, & Grineski, 2017; Grineski, Collins, Chakraborty, et al., 2013; C. R. Jia et al., 2014). For example, a study in Texas reported that census tracts with higher proportion of Hispanics and socioeconomic disadvantage in Houston area showed higher cancer risks burden from exposure to air toxics (Linder et al., 2008). Another study in South Carolina strengthened the evidence that non-white population and communities with low-income had higher cancer

risks from exposure to air toxics (Wilson et al., 2015). However, these studies all utilized modeled air toxic concentrations to investigate the health disparities in health risks from exposure to air toxics. Their modeled air toxics concentrations and related health risks were from National Air Toxic Assessment (NATA) which was conducted by United States Environmental Protection Agency (U.S.EPA) to have an overview of national level of air toxics and to prioritize certain air toxics for regulation and reduction of air toxics emissions (U.S.EPA, 2016g). Particularly, NATA was announced as not appropriate for characterizing and comparing risk at local level (U.S.EPA, 2015d). However, due to high expenses and tremendous efforts acquired for monitoring at census tract level, census tract level monitoring measurement were usually surrogated with modeled estimates in disparity studies at local level. Nevertheless, several model assessment studies have reported that NATA estimated air toxics concentrations generally underestimate the air toxic concentrations (Garcia et al., 2014; Logue et al., 2010; Lupo & Symanski, 2009). Thus the uncertainty of modeled estimates might affect the true association between sociodemographic factors and exposure to air toxics or related health risks. So far, no previous studies have utilized monitored air toxic measurement to address the sociodemographic disparities in exposure to air toxics or related health risks. Therefore, evaluating the sociodemographic disparities with the monitor measurement might give us an alternative prospect on this issue.

Our study geographically focused on Memphis area at Shelby County in the State of Tennessee (TN). Memphis, the county seat, is the largest city in TN and 64% of the residents are African-Americans. Memphis area manifested urban and sub-urban pattern and clusters various types of air toxic emission sources. Automobile traffic, railway transportation, barge traffic on Mississippi river and the busiest airport in the U.S.A. are major mobile sources of air toxics in

Memphis area. Different types of industries, including petroleum refinery, power plant, transportation carriers, petrochemical storage and transfer facilities, and waste disposal facilities are major stationary sources of air toxics in Memphis area. EPA's Toxic Registry Inventory (TRI) is on watch for 311 emission facilities in Memphis. Furthermore, most of these emission facilities are located in low-income African American concentrated areas. Clustering and uneven distribution of emission sources in Memphis area has raised great concerns of health risk and health disparities related to air toxics. In responding to these concerns, the Reducing Exposure to Airborne Chemical Toxics (REACT) air toxics monitoring program was developed to measure ambient air toxic concentrations at census tract level in Memphis area through 2014. This monitoring program collected air samples of 71 target compounds for four different seasons throughout the year at 112 sites in 106 census tracts in Memphis area. With these monitored measurement from REACT and modeled estimates from EPA's 2011 NATA, we investigated the sociodemographic disparities in cancer risks from exposure to air toxics in Memphis area. In our assessment, we expected to provide an alternative and comprehensive view in assessing sociodemographic disparities in risks from exposure to air toxics by introducing monitored measurement into analysis.

## 4.2 Method

### 4.2.1 Cancer risks from exposure to air toxics

Cancer risks from exposure to carcinogenic air toxics at census tract level on the basis of modeled and monitored estimates were both involved in our assessments. Model predicted cumulative cancer risks from exposure to each carcinogenic air toxics at census tract level were extracted from 2011 National-Scale Air Toxics Assessment (NATA). Monitored concentration of 22 carcinogenic Volatile Organic Compounds (VOCs) from the 71 targeted VOCs was

extracted from REACT database. At census tract level, cancer risks from these 22 VOCs were then calculated on the basis of the inhalation risk assessment methods: Risk = Exposure (monitored concentration) × IUR (inhalation unit risk) (Cook et al., 2007). Cumulative cancer risks at each census tract was estimated by summing individual cancer risks from each VOC because cancer risks from air toxics were assumed to be additive (Morello-Frosch & Jesdale, 2006). Additionally, based on cancer risk contribution rank from NATA 2011 and REACT, we also extracted the census tract level concentrations of 4 prioritized individual compounds which were both assessed in NATA and REACT. These 4 compounds were 1,3-butadiene, benzene, benzyl chloride and naphthalene. Specifically, benzene were reported as one of the top 3 cancer risks contributor in both national and regional scale by NATA 2011 (U.S.EPA, 2016b, 2016c) 1,3-butadiene and naphthalene were reported as two of the 8 major national cancer risks contributor (U.S.EPA, 2016c). Naphthalene and benzyl chloride were also identified as the first and second cancer risk contributor in Memphis area by REACT monitoring program. The sociodemographic disparities in exposure to these 4 individual compounds were assessed in this study.

### 4.2.2 Sociodemographic predictors

Sociodemographic variables at census tract level were extracted and derived from census 2010 and EPA EJSCREEN database. To be noticed, the newly developed EJSCREEN is designed for addressing health disparities in environmental justice studies. The sociodemographic variables considered in our study were: population density (person per square mile), median inflation adjusted household income, percent of minority, percent less than high school education, traffic proximity and volume, proximity to National Priorities List (NPL) sites, proximity to Risk Management Plan (RMP) facilities, and proximity to Treatment Storage and

Disposal (TSDF) facilities. Specifically, population density was calculated via dividing total population by land area in each individual census tract. Both total population and land area at census tract level was extracted from census 2010. All the other sociodemographic variables were extracted from EJSCREEN. Median inflation adjusted household income was the level of household income at census tract level where half the households have income above this level and the other half of households have income less than this level. Percent of minority was percent of individuals in a census tract who are not non-Hispanic white. Percent less than high school education was the percent of people age 25 or older without high school diploma. Traffic proximity and volume indicated average annual daily count of vehicles at the nearest (within 500 meters) major roads divided by distance between the major roads to the location (community or residents) of interest. National Priorities List (NPL) sites are a key subset of superfund sites. Risk Management Plan (RMP) facilities are facilities which have potential chemical accident risk according to Clear Air Act (CAA) and thus was managed as RMP facilities. Treatment Storage and Disposal (TSDF) facilities are hazardous waste management facilities that treat, store and dispose hazardous waste. The proximity to NPL, RMP and TSDF facilities are the counts of those facilities (within 5km) divided by the distance from location of interest to those facilities.

### 4.2.3 Data compilation and predictor selection

Cumulative cancer risks and sociodemographic variables were matched by census tract number in Memphis area using ArcGIS (ESRI Inc.). Pearson correlation test were conducted to identify any strong correlation between sociodemographic predictors using correlation coefficient of 0.65 as a rule of thumb. Percent of population who have less than high school education, percent of minorities and median household income showed strong correlation.

Therefore, we selected median household income as a representative predictor for all these three predictors because it gave us the best regression model fit.

Multicollinearity was then evaluated for all sociodemographic variables that were included in the statistical model before any further analysis. The variation inflation (VIF) values of all tested variables were less than 5 indicating no multicollinearity among those variables.

### 4.2.4 Statistical analysis

Moran's I statistic were tested separately on both modeled and monitored census tract level cancer risks to see if there is spatial autocorrelation of cancer risks among neighboring census tract. Ordinary least square (OLS) regression was applied to address the sociodemographic disparity in cancer risks from air toxics when cancer risks have no spatial auto correlation. Otherwise, spatial regression analysis was utilized to assess this association by implementing geographic weight into the statistical model.

In our preliminary analysis, Moran's I test p-value was 0.24 for modeled cancer risks indicating no spatial auto-regression. For modeled cancer risks, the Moran's I test p-value was 0.001indicating existence of spatial autocorrelation. Therefore, we assessed the monitored data using OLS regression and we utilized spatial regression for modeled data. Particularly, spatial regression has two major types including spatial lag model and spatial error model. Spatial regression model introduces a "spatial lag" term into the regression to combine the effect of certain variable at nearby location into the effect of its current assessing location so that spatial interaction can be simultaneously assessed. When spatial lag is included in the dependent variable, the model is spatial lag model. In contrast, a spatial lag in error term is the spatial error model (Raddatz & Mennis, 2013). We assessed the disparity via spatial error model for our

analysis because it had greater R-squared and Log likelihood values and a smaller AIC than the spatial lag model. Our spatial error model is as followings:

$$\text{Cancer Risk} = \beta_0 + \beta_1 \text{ population density} + \beta_2 \text{ income} + \beta_3 \text{ traffic density} + \beta_4 \text{ proximity to NPL} + \beta_5 \text{ proximity to RMP} + \beta_6 \text{ proximity to TSDF} + \lambda Wu + e$$

In this model, $\lambda$ represents the coefficient for spatially auto-correlated errors (spatial autoregressive coefficient), W is the spatial weights matrix based on tract centroids, e represents the random error term in the OLS model, and u is the spatially independent error term. The distance-based spatial weights were constructed using a threshold distance of 2 km. The critical distance of 2 km was selected because air pollution and its associated health risks are negligible beyond this distance. The OLS regression analysis was completed using SAS (SAS 9.4, SAS Institute Inc.) The spatial regression analyses were completed in OpenGeoDa (Version 1.4.6, GeoDa Center, Tempe, AZ, USA).

## 4.3 Results

### 4.3.1 Cancer risk and sociodemographic distribution in Memphis area

Population in Memphis generally had high level of cancer risk from air toxics (Table 4.1). The average census tract level lifetime cancer risk due to exposure to air toxics in Shelby County was 50 people in 10 million which was 50 times the benchmark of 1 in one million. Furthermore, REACT monitored cancer risks in Memphis area were 587 times the benchmark. The manifested discrepancy of sociodemographic status among the population in Memphis area was also presented in Table 4.1. On average, more than half (64.3%) of the population were minorities. Almost half (45.2%) of the population were categorized to low-income group. One in third of the people who age 25 or older did not complete high school with diploma. The sampling

sites selected in REACT monitoring program is representative of Shelby County. As showed in Table 4.1, the average level of sociodemographic variables summarized from 112 REACT monitoring sites in 106 census tracts was similar to the statistics summarized from 220 census tract of entire Shelby County.

**Table 4.1 Summary statistic of variables**

| | | N | M | SD | Min | Mdn | Max |
|---|---|---|---|---|---|---|---|
| **Modeled** | | | | | | | |
| | Cancer risk (person/million) | 220 | 50 | 4 | 41 | 49 | 62 |
| | Population density (person/mile$^2$) | 220 | 3077 | 1794 | 0 | 2934 | 11561 |
| | Less than high school education (%) | 220 | 28.9 | 10.8 | 0.0 | 30.5 | 59.2 |
| | Minority (%) | 220 | 64.3 | 33.1 | 0.0 | 76.3 | 100.0 |
| | Median household income (10K) | 220 | 4.5 | 2.7 | 0.0 | 3.8 | 14.9 |
| | Traffic proximity and volume | 220 | 185 | 282 | 1 | 95 | 2165 |
| | Proximity to NPL | 220 | 0.1 | 0.1 | 0.0 | 0.1 | 1.3 |
| | Proximity to RMP | 220 | 0.5 | 0.7 | 0.0 | 0.2 | 4.0 |
| | Proximity to TSDF | 220 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **Monitored** | | | | | | | |
| | Cancer risk (person/million) | 112 | 587 | 327 | 420 | 494 | 2888 |
| | Population density (person/mile$^2$) | 112 | 2770 | 1744 | 0 | 2749 | 7028 |
| | Less than high school education (%) | 112 | 27.2 | 11.5 | 0.0 | 28.5 | 51.8 |
| | Minority (%) | 112 | 58.0 | 34.0 | 0.0 | 62.7 | 100.0 |
| | Median household income (10K) | 112 | 5.0 | 3.1 | 0.0 | 4.3 | 14.9 |
| | Traffic proximity and volume | 112 | 154 | 194 | 1 | 94 | 1005 |
| | Proximity to NPL | 112 | 0.1 | 0.2 | 0.0 | 0.1 | 1.3 |
| | Proximity to RMP | 112 | 0.5 | 0.6 | 0.0 | 0.2 | 3.7 |
| | Proximity to TSDF | 112 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

NPL- National Priorities List sites; RMP-Risk Management Plan facilities; TSDF-Treatment Storage and Disposal facilities

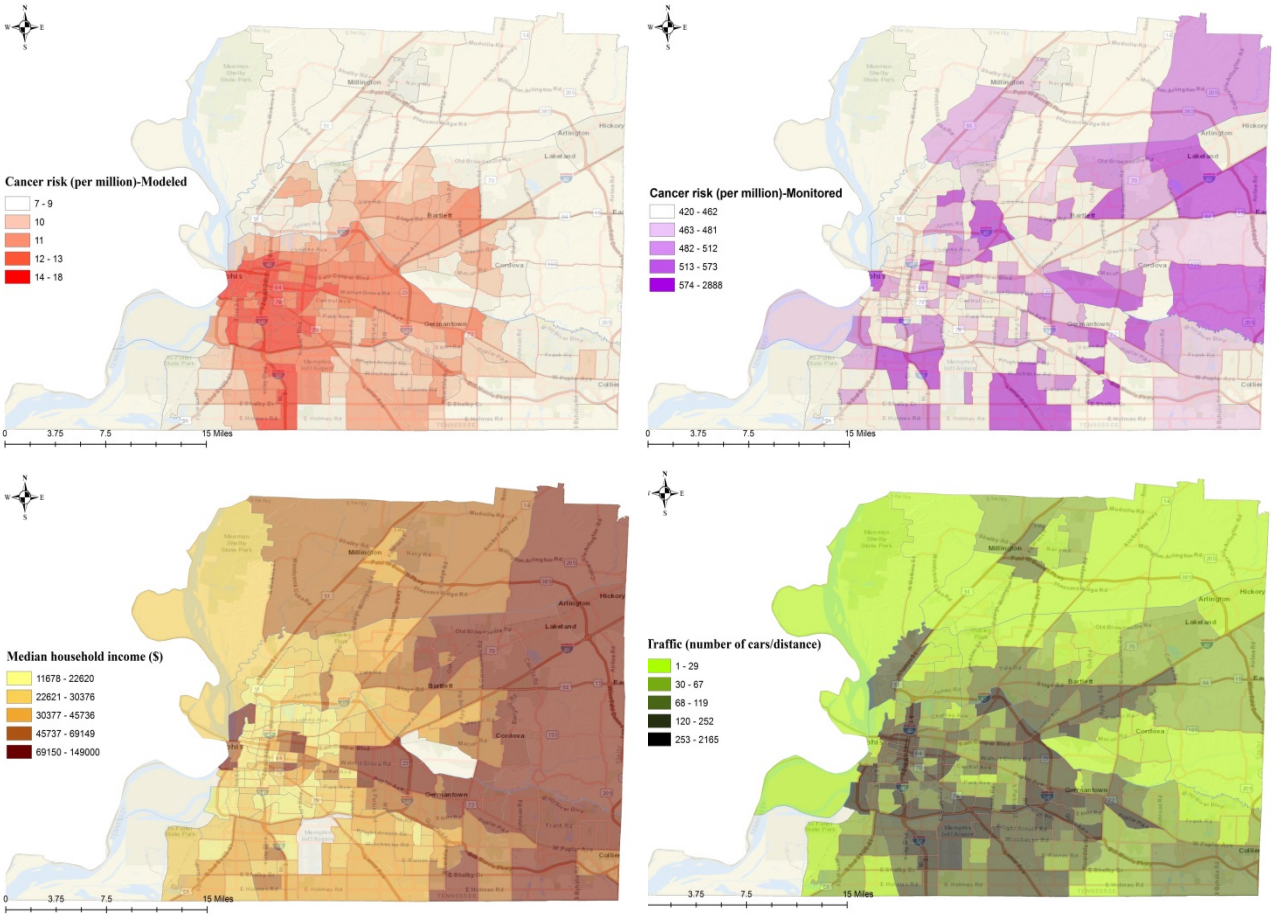


**Figure 4.1 cancer risks and risk factors distributions**

### 4.3.2 Sociodemographic disparities in cancer risks

For modeled data, significant association between sociodemographic factors and cancer risks were found in our analysis (Table 4.2). The spatial regression model was statistically significant with $R^2$ of 0.83 indicating that sociodemographic predictors can explain the distribution of cancer risks from exposure to air toxics. Specifically, census tract with higher population density were more likely to have higher cancer risks from exposure to air toxics. The coefficient 0.31 indicated that every increase of 10K people per square miles in a census tract was positively and significantly associated 31% increasing in cumulative cancer risks ($p$<0.001). People with higher income tended to have less cancer risks from exposure to air toxics than those

with lower income. Every 10K increase of household income will lower than cancer risks from exposure to air toxics by 1% ($p<0.01$). Residents who live close to heavy traffic were also exposed to higher cancer risks. Within a year, every increase of 1000 cars on the near road within 500 meters of the residents' house will elevate the cumulative cancer risks from exposure to air toxics by 26% ($p<0.001$). Residents who live closer to the NPL sites had higher cancer risks from exposure to air toxics than those who live far away from NPL sites Specifically, given an resident who live at the center of a circle area with a 5kilometers radius, one more NPL site set within this area is significantly associated with 15 percent increase in cancer risks to this resident ($p=0.03$). Alternatively, if the number of NPL sites within the circle area is set, 1km closer the resident's house is to the NPL site is significantly associated with 15% increase in cancer risks. Proximity to RMP and TSDF positively associated with cancer risks but the associations were not statistically significant.

In contrast, for monitored data, we did not found significant association between sociodemographic factors and cancer risks. Furthermore, the direction of associations was not as expected for some of the sociodemographic factors and cancer risks. Median household income was positively associated with cancer risks indicating people with lower income would have lower cancer risks. Traffic density, proximity to NPL sites, and proximity to RMP facilities were negatively associated with cancer risks indicating protective effect.

**Table 4.2 Sociodemographic disparities in cancer risks**

| | Modeled | | Monitored | |
|---|---|---|---|---|
| | **Coefficient** | **p-Value** | **Coefficient** | **p-value** |
| Population density (10K person/ mile$^2$) | 0.31 | <.001 | 0.21 | 0.32 |
| Median household income (10K) | -0.01 | <.01 | 0.01 | 0.36 |
| Traffic density (1000 cars/distance) | 0.26 | <.001 | -0.03 | 0.87 |
| Proximity to NPL | 0.15 | 0.03 | -0.20 | 0.36 |
| Proximity to RMP | 0.01 | 0.42 | -0.001 | 0.98 |
| Proximity to TSDF | 184.86 | 0.10 | 466.98 | 0.20 |

### 4.3.3 Sociodemographic disparities in exposure to major carcinogenic VOCs

For the 4 VOCs targeted as prioritized cancer risks contributors, the strength of the association between their ambient concentrations and sociodemographic factors varies largely between modeled and monitored data. The results from analyzing modeled data showed generally statistically significant associations while the results generated based on monitored data gave us non-significant associations (Table 4.3). Additionally, the direction of associations was consistent in results from modeled data while was not consistent in results from monitored data. For example, traffic density was positively statistically significantly associated with the ambient concentration of naphthalene when analyzing modeled data. However, when analyzing monitored data, traffic density was negatively associated with the ambient concentration of naphthalene.

**Table 4.3 Sociodemographic disparities in VOCs' level**

| VOCs | V1 | V2 | V3 | V4 | V5 | V6 |
|---|---|---|---|---|---|---|
| **Modeled** | | | | | | |
| 1,3-Butadiene | 0.03 ** | -2.19E-03 ** | 0.03 ** | 0.03 * | 1.65E-03 | 56.64 ** |
| Benzene | 0.25 ** | -3.62E-03 | 0.15 ** | 0.09 * | 3.48E-03 | 2.33 |
| Benzyl chloride | 0.00 | -1.81E-07 ** | -5.04E-07 | -1.23E-06 | -1.74E-07 | -1.69E-04 |
| Naphthalene | 0.02 ** | -1.32E-03 ** | 0.02 ** | 0.01 * | 1.06E-03 | 19.82 ** |
| **Monitored** | | | | | | |
| 1,3-Butadiene | 0.18 | -0.01 | 0.19 | 0.28 * | 0.02 | -261.12 |
| Benzene | -0.10 | -0.02 * | 0.04 | 0.09 | 0.07 * | -344.04 |
| Benzyl chloride | 0.46 | 0.02 | -0.12 | -0.25 | -0.03 | 1123.57 * |
| Naphthalene | 0.13 | -3.15E-03 | -0.05 | -0.14 | 0.02 | 194.02 |

** $p<0.01$; *$p<0.05$; V1- Population density (1000 people/mile$^2$); V2- Median household income (10K); V3-Traffic density (1000 cars/distance); V4-Proximity to NPL; V5-Proximity to RMP; V6-Proximity to TSDF

## 4.4 Discussion

### 4.4.1 Previous sociodemographic and environmental factors involved in disparity studies

To address the socioeconomic and demographic disparities in risks from exposure to air toxics, various sociodemographic factors have been assessed in previous studies. Race and ethnicity was considered one of the most important sociodemographic factors that were associated with people's exposure of air toxics (Gilbert & Chakraborty, 2011; Morello-Frosch & Jesdale, 2006). In a local study conducted in Maryland, census tract where residents had high fraction of African-Americans yielded three times higher cancer risks than those census tract with less faction of African-Americans (Apelberg et al., 2005). In Florida, race and ethnicity were reported to be significantly associated with cancer risks. Specifically, proportions of black, Hispanic and Asian were all significantly associated with cancer risks (Gilbert & Chakraborty, 2011). In summary, percentage of minorities is a common factor that has been involved in many disparities studies. Another important sociodemographic factor that has been commonly assessed in previous studies was socioeconomic status (James et al., 2012; D. Payne-Sturges & Gee,

2006). Several different variables were used previously to represent the socioeconomic status. The most common variables for socioeconomic status were median household income and percentage of poverty. A study in Louisiana reported that residents live in low-income census tracts had 12% more of cancer risks than those live in high-income census tracts (James et al., 2012). In Houston, Texas, a study reported that residents live in census tracts where the proportion of poverty reach 25% experience 4 to 10 times cancer risks from exposure to air pollution than those who live in census tracts with lower proportion of poverty (Linder et al., 2008). In addition to these sociodemographic factors, environmental factors were also assessed in previous disparity studies. Traffic impact such as proximity to traffic or traffic density was reported to be an important factor which was associated with variation of the exposure to air toxics in different communities (Fujita et al., 2013; Wu & Batterman, 2006). A study in California reported that air toxics concentrations showed consistent pattern with proximity to traffic. Closer to heavy traffic indicated higher exposure to air toxics (Fujita et al., 2013). There were also multiple studies in which the proximity to stationary emission sources was utilized as one of the factors to address disparity issues (Abel, 2008; Perlin et al., 2001). In the study conducted in St. Louis, the proximity to nearest TRI facilities was assessed, the results indicated that minorities who lives closer to TRI facilities tend to expose to higher level of air toxics (Abel, 2008). In another disparity study conducted in West Virginia, Louisiana and Maryland, minorities were reported to be more likely to live in poor area which was close to multiple industrial emission sources of air pollution (Perlin et al., 2001). In general, three major factors of sociodemographic disparity in exposure to air toxics can be summarized from previous studies. The first one is socioeconomic status, the second one is traffic and the third one is the proximity to stationary emission source. In our study, we utilized the U.S.EPA's EJSCREEN database and

census data as data source for these three major factors of sociodemographic disparities in exposure to air toxics. EJSCREEN had advantage on providing consistent geographically matched environmental and demographic indicators at census tract level. Particularly, these environmental and demographic indicators are prioritized for addressing health disparity issues. Previously, EJSCREEN database was rarely used in air toxic related health disparity issues. Our study is the first study to implement this new tool in exploring the air toxic related health disparity issues so that all the sociodemographic and environmental factors utilized in previous studies could be examined together in this study.

### 4.4.2 Consistent findings from modeled data

In previous studies, only model estimated air toxic levels from NATA were used to assess the sociodemographic and environmental disparities in risks from exposure to air toxics. Therefore, we firstly assessed the health disparity using NATA modeled estimates. Our results indicate that there was statistically significant association among sociodemographic factors, environmental factors and the cancer risks from exposure to air toxics. The 4 significant factors included population density, median household income, traffic density and proximity to NPL sites. Compared to previous studies, some of these factors indeed had been reported as significant factors that affect health disparities (Fujita et al., 2013; James et al., 2012). Proximity to NPL sites was significantly associated with cancer risks from exposure to air toxic while proximity to RMP and TSDF were not. This finding indicated that regulation on emissions from NPL sites has larger impact on reducing cancer risks from exposure to air toxics comparing with regulations to RMP and TSDF facilities. Alternatively, regulation to RMP and TSDF might be better implemented than regulation to NPL as the proximity to RMP and TSDF was not

significantly associated with cancer risks from exposure to air toxics. Therefore, regulation to NPL sites could be furtherly improved.

### 4.4.3 Contradictory findings from monitored data

Our results from analyzing monitored data indicated that the association between sociodemographic and environmental factors and cancer risks from exposure to air toxics was not statistically significant. Additionally, the direction of association was not consistent with that when analyzing modeled data. Although no previous studies have assessed the sociodemographic and environmental disparity on cancer risk from exposure to air toxics on the basis of monitored measurement, a few previous studies addressing disparity in exposure air toxics using sociodemographic segregation index pointed out that the association between air pollution levels and segregation index might be not be strong and could be uncertain (Downey, 2007; Downey, Dubois, Hawkins, & Walker, 2008; Rice et al., 2014) . The association might vary by air pollutants and the direction of association was inconsistent (Morello-Frosch & Lopez, 2006).

The differences between modeled and monitored estimations were the major contributor to the discrepancies in findings from modeled and monitored data. As aforementioned, although the conservative assumptions used in NATA modeling were assumed to lead to overestimation of ambient air toxic level (U.S.EPA, 2016g), underestimation and uncertainty in modeled data was generally concluded in the previous third-party studies which compared EPA's NATA estimates with monitored measurement of air toxics (Garcia et al., 2014; Logue et al., 2011; Lupo & Symanski, 2009). Insufficient information from local small scale emissions sources is one of the factors that lead to underestimations of certain air toxics in NATA estimation (Scheffe et al., 2016). NATA has been collecting information of air toxic emissions mainly from National Emissions Inventory (NEI) and Toxics Release Inventory (TRI); however, various local sources

of air toxic emissions may not be regulated or reported to NEI or TRI. For example, personal smoking, household Barbecue cooking (Memphis area), dry-cleaning stores and some auto body shops were not listed in NEI. As showed in figure 5.1 and 5.2, cancer risks estimated from modeled data were higher majorly in southwest Memphis area where more large point emission sources clustered. In contrast, cancer risks level predicted from monitored measurement were similar around Memphis area, which reflected that air toxic level was almost similar around Memphis area. Another difference between modeled estimates and monitored measurement was that modeled data was estimated in a way that extreme concentrations could not be captured. In comparison, monitored measurement can provide average concentrations after dealing with extrema. Additionally, modeled estimates cannot accurately reflect the meteorological conditions and photochemical reactions in real environment (Rosenbaum et al., 1999). To be noticed, monitored measurement might also be different from modeled estimation due to insufficient number of monitors in study area and variation in the distance from monitors to emission sources in different census tracts (Garcia et al., 2014; Scheffe et al., 2016).

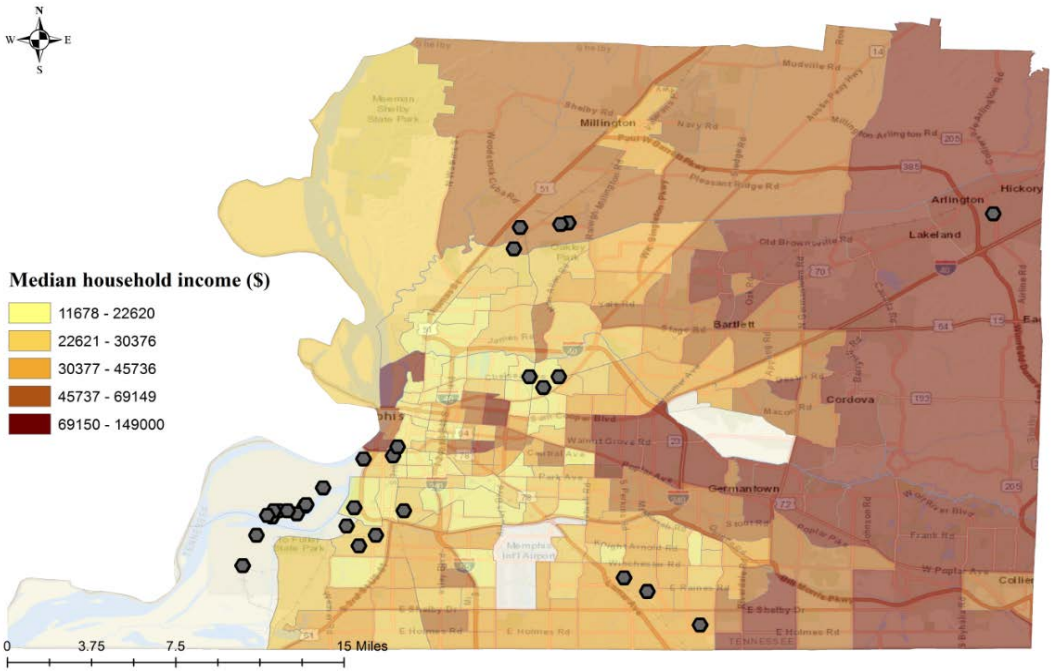


**Figure 4.2 Point emission sources distribution in Memphis area**

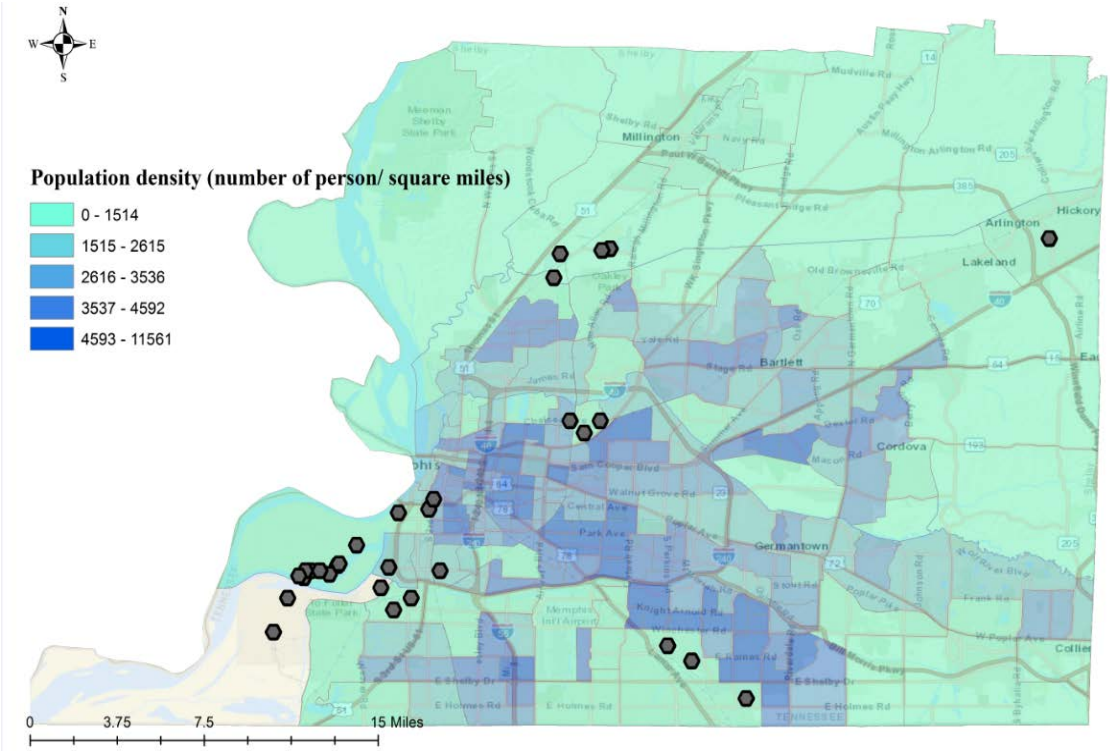


**Figure 4.3 Population density distributions in Memphis area**

### 4.4.4 Limitations

There are several limitations in our study. For our monitored measurement, we had air sample collected one time in each season in 2014 at each of the 112 sampling site within 106 census tract. Although we collected duplicate sample to assure the quality of our sample collection, but if samples were more frequently collected at each site, we may capture the ambient air toxics level better. However, due to limited resources, we did our best to collect as much samples in as many sites as possible. Another limitation is that NATA modeled air toxic level and the monitored air toxic level are not in the same year. NATA modeled estimates was extracted in the year of 2011 while the monitored air toxics were in 2014. The gaps between years might lead to difference in distribution of air toxics in Memphis area. The health disparities in cancer risks from exposure to air toxics might be different when addressing different mixture of air toxics. (H. F. Yu & Stuart, 2016). In our analysis on NATA modeled estimates, all carcinogenic air toxics were included. While when we analyzed the monitored measurement, only 15 targeted carcinogenic VOCs were assessed.

We assessed the modeled data using both ordinary least square linear regression (OLSR) and spatial regression. The adjusted R square in OLSR was 0.4814 and the residual sum of square was 4.5. In the spatial regression, the adjusted R square was 0.58 and the sigma square was 0.02. Although the model was improved by using spatial regression, spatial dependency still showed in the model as the Breusch-Pagan test and likelihood ratio test showed significance. We also analyzed monitored data using spatial regression and OLSR. As monitored data did not show spatial autocorrelation, the spatial regression and OLSR showed same results indicating non-significant association between sociodemographic factors and cancer risks from exposure to air toxics. The results from analyzing monitored data might indicate that the entire Memphis area

shared similar exposure pattern to air toxics. Therefore, our findings might not be generalized to other area other than similar metropolitan area due to geographical limitation.

### 4.5 Conclusion

Our study is the first one in the U.S.A. utilizing local monitored measurement to explore the sociodemographic disparity on cancer risk from exposure to air toxics. The discrepancy between results from modeled and monitored data indicated that conclusions made in previous studies purely based on modeled data might be misleading when addressing sociodemographic disparities issues. The uncertainty in modeled data was critical in this discrepancy. Analyzing monitored data is an alternative option to previous approach and is more appropriate for revealing the true association between sociodemographic factors and risks from exposure to air toxics. The air toxics might be clustered as a whole in the air of entire Memphis area so that there are not clear sociodemographic disparities in exposure to air toxics. This study provided alternative aspect in addressing disparity issues and provided evidence for future air pollution regulation urban planning and public administration. More analysis on monitored measurement from local monitoring programs is needed to provide more evidence regarding health disparity issues.

# Chapter 5

## Exposure profiles of air toxics mixtures in microenvironments

### 5.1 Introduction

Air pollution has been one of the most important environmental health concerns because air pollutants are ubiquitous and exposure to air pollutants is associated with various health risks (Kampa & Castanas, 2008; ODPHP, 2017) . It is well known that air pollutants exist as mixtures in the air and people are exposed to multipollutant mixtures in daily life (Dominici et al., 2010; Oakes, Baxter, & Long, 2014). National Academies of Science (NAS), National Institute of Environmental Science (NIEHS) and United States Environment Protection Agency (U.S.EPA) have prioritized researches in human's adverse health effect from exposure to multipollutant mixtures (NRC, 2004; Rider, Carlin, DeVito, Thompson, & Walker, 2013; U.S.EPA, 2008). However, most of previous studies assessed the association between air pollutants and various health outcomes via a single pollutant approach omitting the fact that people are exposed to air pollutants mixtures. More studies via multipollutant approach are in need to further examine the exposure to multipollutant mixtures (Dominici et al., 2010; Hidy & Pennell, 2010; M. Oakes et al., 2014).

The highly correlated and interactive individual air toxics within the mixtures can create joint effect which negatively affects human health in a different way from purely individual effect of single pollutants (Billionnet, Sherrill, Annesi-Maesano, & Study, 2012; Coker et al., 2016). In a few recent studies, exposure to air toxics mixtures has been associated with adverse health effects such as low birth weight (Coker et al., 2016) and cognitive development of children (Stingone, Pandey, Claudio, & Pandey, 2017). However, current scientific evidence is still not adequate to explicitly demonstrate the joint effect from mixture of air toxics on human

health. Mixture of air toxics varies distinctly in different environment due to variation in terms of types of air toxics and the formation of the mixture. The uncertainties in toxicity of air toxic mixtures have been major concerns when addressing health risks from exposure to air toxic mixtures (Dominici et al., 2010). Additionally, multicollinearity among multiple pollutants and difficulties in interpretation of results are critical challenges to researchers in terms of analytical methodology (Davalos, Luben, Herring, & Sacks, 2017; Dominici et al., 2010).

Traditional stepwise linear regression model with interaction terms has limitations in estimating the health risks from exposure to multipollutant mixtures when there are three or more pollutants that are highly correlated (Cantuaria, Brandt, Lofstrom, & Blanes-Vidal, 2017; Dominici et al., 2010). Therefore, to estimate the join effect from the interaction among the highly correlated multipollutant mixtures, various statistical and analytical approaches have been adopted in previous studies. For example, Bayesian approach (Bobb et al., 2015), recursive partitioning (Stingone et al., 2017), dimension reduction and shrinkage methods (Qian, Zhang, Korn, Wei, & Chapman, 2004), classification and clustering metrics, source apportionment such as positive matrix factorization (PMF)(Sarnat et al., 2008) and chemical mass balance (CMB) techniques have all been applied to explore the joint effect of multipollutant mixtures (Billionnet et al., 2012; Davalos et al., 2017).

To study the exposure to multipollutant mixtures in different area, establishing an exposure profile is valuable. The spatial specific exposure profile of mixtures enable researchers identify the sources of the mixture, target the communities exposed to mixtures and further assess adverse health effect from exposure to mixtures (Stingone et al., 2017). Furthermore, exposure profile can be established through multipollutant approach which is promising for assessing exposure from multipollutant mixtures (Cantuaria et al., 2017; Dominici et al., 2010)

Random forest, one of the recursive partitioning methods, has advantages in analyzing high dimension data with interaction among multiple variables (Billionnet et al., 2012; Strobl, Malley, & Tutz, 2009) and can classify the exposure profiles of different mixtures (Stingone et al., 2017). Previously, most studies using multipollutant approach to assess the exposure to criteria air pollutants or indoor volatile organic compounds (VOCs). Exposure to outdoor VOCs mixtures was rarely assessed (Billionnet et al., 2012; Coker et al., 2016; Davalos et al., 2017). Our study is the first study that adopted random forest approach to establish spatial specific exposure profiles of VOC mixtures based on both ambient and indoor monitoring measurement in Memphis area where clusters various emission sources. The goal of our study is to classify and eastablish the spatial specific exposure profiles for VOC mixtures and prioritize VOCs that distinguish the exposure profiles in Memphis area.

## 5.2 Method

### 5.2.1 Sampling procedure

From July to November 2016, a total of 130 air samples were collected at Memphis area in five different microenvironments. Air samples were collected at 51 different gas stations, 20 different locations within community, 14 offices at different locations, 30 different resident homes and 15 different vehicles. Both active and passive sampling techniques were utilized to collect those samples. Active sampling were conducted to collect air samples at 200 ml/min at gas stations and inside vehicles for short period time  by paring SKC air sampling pump with stainless-steel Tenax TA tube given the average short exposure time in these two micro-environment in daily life. Air samples collected in communities, resident's home and offices were achieved using passive sampling techniques by installing stainless-steel Tenax TA tube for 24 hours sampling time.

### 5.2.2 Laboratory analysis

Collected air samples were analyzed on an automated thermal desorption-gas chromatography/mass spectrometry (TD-GC/MS) system for 70 target compounds using the scan mode. Among these 70 compounds, 16 VOCs were detected in all the 5 micro-environments and had concentrations that were significant different from non-detection level. These 16 VOCs, with known health risks, were benzene, carbontetrachloride, toluene, ethylbenzene, m,p-xylene, styrene, o-xylene, isopropylbenzene, propylbenzene, 1,3,5-trimethylbenzene, 1,2,4-trimethylbenzene, 1,3-dichlorobenzene, 1,2,3-trimethylbenzene, p-isopropyltoluene, d-limonene, naphthalene.

### 5.2.3 Statistical analysis

Descriptive statistics were generated for all 16 VOCs for each micro-environment and the entire Memphis area. Multivariate analysis of variance (MANOVA) was conducted to detect whether the VOCs profile is different for all the micro-environments. Furthermore, pairwise analysis was utilized to compare each pair of micro-environments and determine if the 16 VOCs' profile in these micro-environments were significantly different. Sample similarity within each micro-environment was tested using Euclidean distance via R package "ade4" (Dray & Dufour, 2007; Thioulouse & Dray, 2007). Random forest (RF) algorithm, one of the recursive partition methods (Strobl et al., 2009), was applied in the supervised classification machining learning mechanism to classify the sampling sites in order to establish spatial-specific exposure profiles of VOCs for different microenvironment. At first, we utilized K-fold (K=5) cross-validation to examine the reliability of RF methods in our VOCs monitoring data. Specifically, among all 130 samples, 100 samples were randomly selected as training data. A total of 2000 trees were generated in each supervised RF process. The rest of 30 samples were used as validation data to

verify the classification. The cross validation was repeated for 100 times. The average error rate was calculated to be 0.2. Secondly, we applied random forest method to the entire 130 samples to identify VOCs which played leading roles in distinguishing exposure profiles of VOC mixtures in different micro-Environments including community, gas station, home, office, and vehicle. As we considered that the monitoring recorded classification of the sampling sites might be misclassified due to the variance of surrounding environment, RF classification results were used to establish spatial-specific exposure profiles. The range and average of concentrations of the 16 detected VOCs were calculated to demonstrate the spatial-specific exposure profiles of different micro-environments. All analysis was conducted using SAS (v9.4, SAS Institute Inc., Cary, NC), R (3.4.1& 3.2.2, The R Foundation), Microsoft Excel (2010) and Arc GIS 10.3.1 (ESRI, Inc.).

**5.3 Results**

The average concentrations of the 16 VOCs in different micro-environments and Memphis area were showed in Table 5.1. Toluene and d-Limonene had higher average and concentrations than other VOCs. Toluene is one of the compounds in a typical air toxic mixture related to gasoline. This mixture consists of benzene, toluene, ethyl benzene, and xylenes (BTEX). The concentration of BTEX sampled from gas station ranged from 2.05 to 15 $\mu g/m^3$, which was 20 times higher than those in communities. However, the level of BTEX measured at gas station was still far below the corresponding minimum risk levels (MRLs) for acute adverse health effects.

**Table 5.1 Descriptive statistics of target VOCs**

| VOCs | Mean (μg/m$^3$) | | | | | |
|---|---|---|---|---|---|---|
| | **Memphis (N=130)** | **Community (N=20)** | **Gas Station (N=51)** | **Home (N=30)** | **Office (N=14)** | **Vehicle (N=15)** |
| 1,2,3-Trimethylbenzene | 0.63 | 0.09 | 0.82 | 0.70 | 0.38 | 0.80 |
| 1,2,4-Trimethylbenzene | 1.89 | 0.23 | 2.72 | 1.78 | 1.05 | 2.30 |
| 1,3,5-Trimethylbenzene | 0.45 | 0.07 | 0.61 | 0.46 | 0.23 | 0.60 |
| 1,3-Dichlorobenzene | 0.72 | 0.03 | 0.04 | 2.85 | 0.26 | 0.17 |
| Benzene | 2.09 | 0.69 | 3.25 | 1.74 | 0.81 | 1.86 |
| Carbon tetrachloride | 0.84 | 0.66 | 0.72 | 0.89 | 1.19 | 1.09 |
| d-Limonene | 10.67 | 0.89 | 0.87 | 33.95 | 20.57 | 1.28 |
| Ethylbenzene | 1.50 | 0.18 | 2.05 | 1.89 | 0.75 | 1.29 |
| Isopropyl benzene | 0.10 | 0.02 | 0.13 | 0.11 | 0.05 | 0.09 |
| m, p-Xylene | 3.99 | 0.31 | 4.10 | 7.47 | 2.81 | 2.63 |
| Naphthalene | 0.93 | 0.09 | 0.33 | 3.00 | 0.35 | 0.45 |
| o-Xylene | 2.21 | 0.26 | 3.02 | 2.65 | 1.08 | 2.21 |
| p-Isopropyl toluene | 0.37 | 0.05 | 0.04 | 1.15 | 0.65 | 0.13 |
| Propyl benzene | 0.29 | 0.04 | 0.44 | 0.29 | 0.12 | 0.28 |
| Styrene | 0.31 | 0.10 | 0.13 | 0.79 | 0.19 | 0.38 |
| Toluene | 10.50 | 1.20 | 15.00 | 11.37 | 6.67 | 9.41 |

### 5.3.1 Variance of exposure in different sampled microenvironments

There was significant difference in VOCs levels among different sampled microenvironment as the global null hypothesis test of MANOVA was significant ($p$<.001). Furthermore, significant difference in paired comparison indicated that concentrations of different VOCs in each microenvironment are significantly different from that in any other microenvironments (Table 5.2).

**Table 5.2 Paired comparison of VOCs' level in different Microenvironment**

| Pair of Micro-Environment | DF | Pillai | Approx F | Num DF | Den DF | Pr(>F) |
|---|---|---|---|---|---|---|
| Gas Station vs Community | 1 | 0.57 | 4.50 | 16 | 54 | <.001 |
| Gas Station vs Vehicle | 1 | 0.48 | 2.79 | 16 | 49 | 0.001 |
| Gas Station vs Office | 1 | 0.98 | 157.91 | 16 | 48 | <.001 |
| Gas Station vs Home | 1 | 0.97 | 119.83 | 16 | 64 | <.001 |
| Community vs Vehicle | 1 | 0.77 | 3.75 | 16 | 18 | 0.001 |
| Community vs Office | 1 | 0.99 | 121.80 | 16 | 17 | <.001 |
| Community vs Home | 1 | 0.97 | 57.29 | 16 | 33 | <.001 |
| Vehicle vs Office | 1 | 0.99 | 91.60 | 16 | 12 | <.001 |
| Vehicle vs Home | 1 | 0.97 | 60.24 | 16 | 28 | <.001 |
| Office vs Home | 1 | 0.82 | 7.66 | 16 | 27 | <.001 |

The level of VOCs also varied within same sampled micro-environment indicating the potential interaction among multiple micro-environment and emission sources. Specifically, air samples collected in vehicles had the highest similarity (90%) indicating that there was a major exposure pattern to VOCs in vehicle due to the separated space inside a vehicle. Gas stations had the second highest similarities (89%) showing that the exposure pattern to VOCs were quite similar due to common emission sources such as gasoline. Air sample collected at office and residents' homes were 77% and 64% similar indicating that more personal exposure pattern existed at homes than offices. Air sample collected in communities had the least similarities (59%) indicating that multiple emission sources within communities were contributors to the existing air toxics and surrounding environment played an important role affecting the air toxic level in different communities.

### 5.3.2 Supervised random forest classification

We conducted K-fold (K=5) cross validation to verify the supervised random forest approach for our monitoring data in this study. Among one hundred RF, 80 RF have error rate less than 0.3. The average and median error rate is 0.2. Therefore, the RF was an appropriate classification method for our study.

Exposure profiles in different micro-environments were partially reflected on the variance in concentrations of same VOCs in different micro-environments. Certain VOCs were prioritized as indicators for categorizing exposure patterns based on the supervised Random forest. The top 5 VOCs that determined the categories of exposure profiles are p-isopropyl toluene, carbon tetrachloride, benzene, d-limonene, and styrene.

We also estimated the microenvironment misclassification error rate (Table 5.3). Gas station, Office, Home are less likely to be misclassified. Community has a misclassification error rate of 0.45 and was likely to be misclassified as Gas Station. Vehicle has a misclassification error rate of 0.47 and was likely to be misclassified as Gas Station, home and community. An explanation for this is that vehicle's emissions are very similar from identified emissions from gas station. In the community, residents usually have several vehicles at their house and some gas stations were close to communities. Therefore, misclassification is more likely to happen.

**Table 5.3 Classification of different microenvironment**

| | Community | Gas Station | Home | Office | Vehicle | Class Error |
|---|---|---|---|---|---|---|
| **Community** | 11 | 9 | 0 | 0 | 0 | 0.45 |
| **Gas Station** | 7 | 44 | 0 | 0 | 0 | 0.14 |
| **Home** | 2 | 0 | 27 | 1 | 0 | 0.10 |
| **Office** | 0 | 0 | 2 | 12 | 0 | 0.14 |
| **Vehicle** | 1 | 5 | 1 | 0 | 8 | 0.47 |

The spatial specific exposure profile for community, office, home, gas station, and vehicle were demonstrated via the concentration range of those 16 VOCs in each micro-environments classified by random forest . In general, the concentrations of 16 detected VOCs were low in Memphis area. Home had more VOCs with high concentrations compared with other microenvironment, while community had the least number of VOCs with high concentrations. However, the toxicity of VOCs varies. Although some VOCs had low concentration, exposure to these VOCs might lead to more health risks than exposure to VOCs with higher concentration but lower toxicity. For example, D-limonene had highest maximum concentration (246.32 μg/$m^3$) and highest median concentration (19.87 μg/$m^3$) in home environment where toluene had second highest maximum concentration (76.25 μg/$m^3$) and median concertation (5.23 μg/$m^3$). Exposure to toluene might lead to more health risks than exposure to d-limonene at home because d-limonene has been studied and considered as with low toxicity based on animal experiments (J. Sun, 2007) while toluene is considered more toxic due to its target organ such as central nervous system (Greenberg, 1997; Sarigiannis, Karakitsios, Gotti, Liakos, & Katsoyiannis, 2011). Therefore, it is important to know the range of concentrations of VOCs to estimate possible health risks of exposure.

Some VOCs stood out in particular microenvironment as they had wider range and higher maximum concertation in the mixtures featuring certain microenvironment. The exposure profiles were presented in figure 4.1. The VOC 1,3-Dichlorobenzene (0.01~32.6 μg/$m^3$, mean=1.54μg/$m^3$), naphthalene (0.01~15.78 μg/$m^3$, mean=0.84μg/$m^3$) , toluene (0.18~2.69 μg/$m^3$, mean=0.91μg/$m^3$) and d-limonene (0~2.55 μg/$m^3$, mean=0.72μg/$m^3$) were the four VOCs that showed higher concentrations in community environment. Toluene (0.57~147.29 μg/$m^3$, mean=12.23μg/$m^3$) showed highest average concentrations and extrema compared with other

VOCs in office. Three VOCs including *m, p*-xylene (0.21~102.43 μg/m$^3$, mean=8.28 μg/m$^3$), d-Limonene (0.13~49.71 μg/m$^3$, mean=5.77μg/m$^3$) and *o*-xylene (0.13~35.05 μg/m$^3$, mean=5.77μg/m$^3$) had higher level of concentrations than the rest VOCs. D-limonene (3.44~246.32 μg/m$^3$, mean=41.27μg/m$^3$), toluene (0.63~76.25 μg/m$^3$, mean=9.12 μg/m$^3$), *m, p*-xylene (0.38~33.04 μg/m$^3$, mean=5.27 μg/m$^3$), and naphthalene (0.04~24.53 μg/m$^3$, mean=2.13μg/m$^3$) were representative VOCs at home. Benzene (0.5~35.41 μg/m$^3$, mean=3.28μg/m$^3$), toluene (0.32~177.19 μg/m$^3$, mean=14.81μg/m$^3$), ethylbenzene (0.05~40.23 μg/m$^3$, mean=1.93 μg/m$^3$), m,p-xylene (0.08~77.74 μg/m$^3$, mean=3.84μg/m$^3$), and o-xlyene (0.06~58.68 μg/m$^3$, mean=2.82μg/m$^3$) are typical mixture known as BTEX. BTEX presented higher concentration in the microenvironment of gas station. To be noticed, 1,2,4-trimethylbenzene (0.07~66.87μg/m$^3$, mean=2.53 μg/m$^3$) also showed high concentrations at gas station. Toluene (0.69~37.49 μg/m$^3$, mean=8.51 μg/m$^3$), *m, p*-xylene (0.24~15.03 μg/m$^3$, mean=3.47μg/m$^3$), 1,2,4-trimethylbenzene (0.17~14.85 μg/m$^3$, mean=3.47μg/m$^3$) , and o-xlyene (0.17~13.58 μg/m$^3$, mean=3.05μg/m$^3$) were the four VOCs showed higher concentrations among the VOCs in vehicles.

Although p-isopropyl toluene had relative low concentrations (mean=0.4 μg/m$^3$, median=0.04 μg/m$^3$) in these five microenvironments exposure profiles, its concentration and the correlated concentration of other VOCs varied distinctly. Thus p-isopropyl toluene was ranked as the top VOC to distinguish these five exposure profiles. The VOC p-isopropyl toluene or p-cymene was known to have neurotoxicity for animals (Lam, Ladefoged, Ostergaard, Lund, & Simonsen, 1996). Carbon tetrachloride and benzene were the No.2 and 3 for distinguish the spatial-specific exposure profile of VOCs mixtures. Carbon tetrachloride was known to be toxic to liver, kidney, lung, testis, blood and central nervous system (Huang, Lei, Wei, & Zeng, 2014).

Benzene was known to be carcinogenic and exposure to benzene can cause leukemia (Bollati et al., 2007; Jex & Wyman, 1996). In comparison, toluene ranked as 7$^{th}$ VOC for distinguishing exposure profiles, but it had consistent relatively high median concentration in all microenvironments. Toluene was known to negatively affect the central nervous system of human beings (Greenberg, 1997; Sarigiannis et al., 2011).

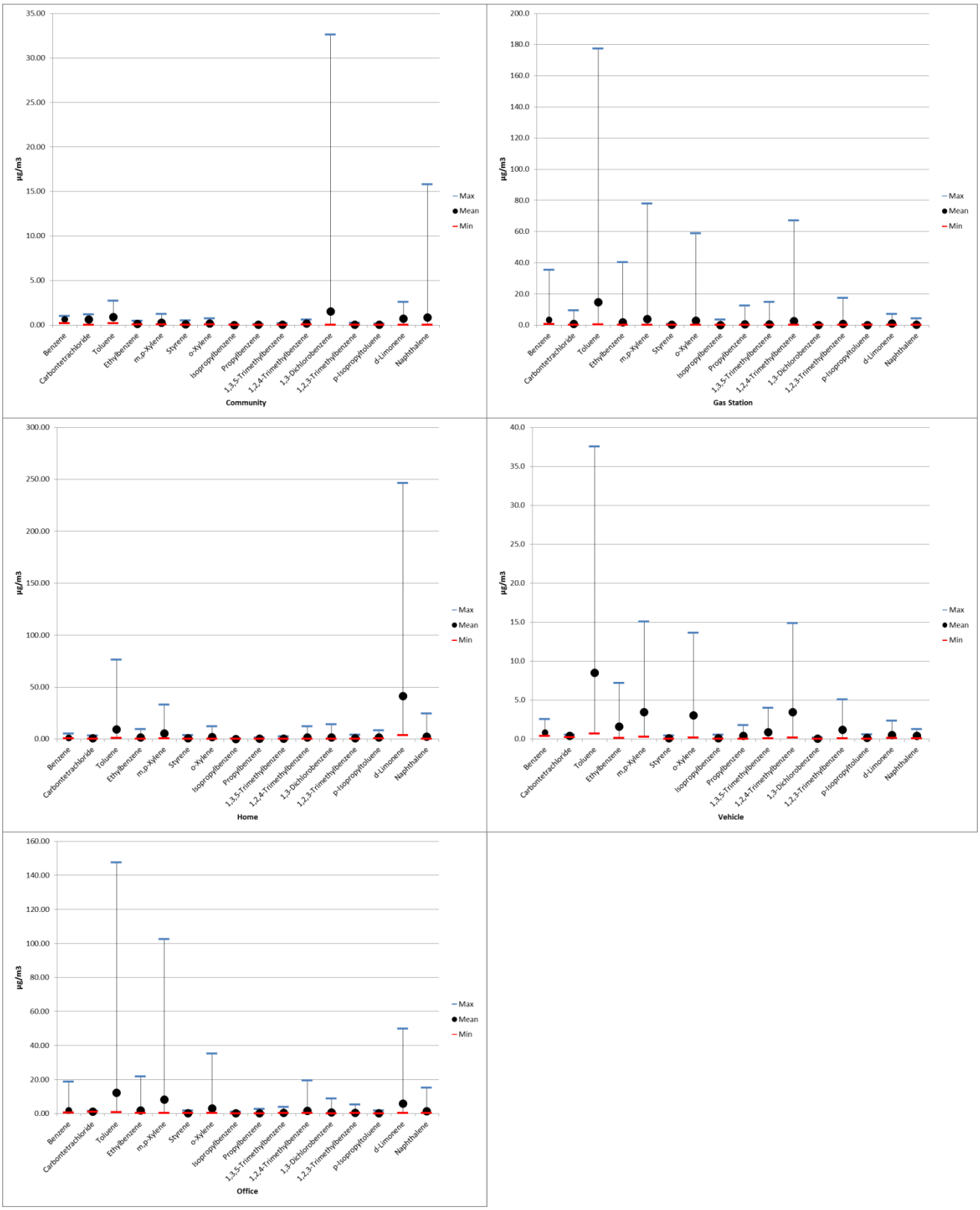


**Figure 5.1 Exposure profiles of different micro-environments in Memphis area**

## 5.4 Discussion

### 5.4.1 Exposure profile is beneficial for assessing exposure to multipollutant mixtures

It is important to quantify the exposure from multipollutant mixtures through multipollutant approach because single or co-pollutant approach has weakness in estimating high dimensional interactions among highly correlated pollutants in the mixtures (Dominici et al., 2010). Studying exposure profile for geographical area such as city or County is beneficial for assess exposure to multipollutant mixtures in two major aspects. First, well-established spatial-specified exposure profiles can help local administration to target major air pollution emission sources for affected communities. Second, exposure profile generated through multipollutant approach can serve well in a two-stage multipollutant approach involving exposure profile identification as the first stage and the second stage as assessing association between exposure profiles and health risks (Davalos et al., 2017; Dominici et al., 2010; Stingone et al., 2017).

### 5.4.2 Previous multipollutant approaches for assessing exposure to mixtures

Uncertainty in components of various multipollutant mixtures, multicollinearity, sophisticated interaction among pollutants and difficulties in interpreting results are major concerns and drawbacks of studies on exposure to multipollutant mixtures (Billionnet et al., 2012; Davalos et al., 2017; Dominici et al., 2010; M. Oakes et al., 2014). Beyond traditional stepwise algorithm in model selection, multipollutant approach via different statistical approaches has been applied to study exposure to multipollutant mixtures. In general, these statistical approaches can be categorized into four categories as deletion/substitution/addition

(DSA) algorithm, linear regression awith shrinkage feature, dimension reduction approach and nonparametric approach (Billionnet et al., 2012; Davalos et al., 2017).

Deletion/substitution/addition (DSA) algorithm was built upon the traditional stepwise model selection procedure and was applied as one of the multipollutant approaches previously (Beckerman et al., 2013; Dominici, Wang, Crainiceanu, & Parmigiani, 2008; Mortimer et al., 2008). For example, DSA was applied in a study assessing the association between exposure to criteria air pollutants and lung function among children who have asthma (Mortimer et al., 2008). DSA is a more optimal model selection method than the suboptimal stepwise model selection method (Fernandes, Geeven, Soetens, & Klontza-Jaklova, 2011; Sinisi & van der Laan, 2004) and has advantage in dealing with outlier via cross validation during model selection but is only appropriate for non-nested model search (Dominici et al., 2008).

Linear regression with shrinkage feature can be categorized into two sub categories as non-effect modification and effect modification approach (Davalos et al., 2017). Given linear shrinkage approach with no effect modification, two typical methods used are multi-level model (Suh, Zanobetti, Schwartz, & Coull, 2011) and penalized regression methods such as least absolute selection operator (LASSO) and ridge regression (S. Roberts & Martin, 2005). Alternatively, linear shrinkage approach with effect modification can be implemented via Bayesian methods or LASSO etc. (Z. C. Sun et al., 2013). The linear shrinkage approach without interaction term in the model may miss the multipollutant interaction while the linear shrinkage approach with interaction term may be inappropriate due to model uncertainty and collinearity. Alternatively, linear shrinkage approach has advantage on ease of interpreting results (Davalos et al., 2017).

Dimension reduction approach also has two sub categories as non-supervised and supervised dimension reduction approach (Davalos et al., 2017). Principle component analysis (PCA)(Sacks, Ito, Wilson, & Neas, 2012), factor analysis, Positive Matrix Factorization (PMF)(M. M. Oakes et al., 2014), cluster analysis (Qian, Chapman, et al., 2004; Qian, Zhang, et al., 2004) and partial least square (PLS) (Seagrave et al., 2006) are major methods applied in both unsupervised and supervised dimension reduction approach when assessing exposure to multipollutant mixtures. Both unsupervised and supervised dimension reduction approach have advantages on reducing overfitting of models when number of predictors are more than sample size. However, the unsupervised dimension reduction may have difficulties in determining number of factors or clusters and interpreting the meaning of each factor or cluster when lack of scientific justification(M. Oakes et al., 2014). Supervised dimension reduction approach may lose data features due to assumption made for pollutant inclusion criteria (Billionnet et al., 2012). For example, source apportionment approach such as PMF requires prior knowledge of grouping chemicals. In our study, the VOCs' grouping is not previously known. Furthermore, the VOCs mixture varies in different micro-environments.

Nonparametric approach includes various statistical methods (Billionnet et al., 2012; Davalos et al., 2017). Classification and regression tree (CART)(Gass, Klein, Chang, Flanders, & Strickland, 2014; Stingone et al., 2017), random forest (Billionnet et al., 2012; Breiman, 2001), Bayesian Kernel machine regression (KMR) (Bobb et al., 2015), and Bayesian profile regression (Coker et al., 2016) are major methods applied in as nonparametric approach when assessing exposure to multipollutant mixtures. Nonparametric approach has advantages over previous approaches in assessing sophisticated non-linear interaction and is also adaptive to high dimensional data analysis (Billionnet et al., 2012; Davalos et al., 2017).

Among these multipollutant approaches, K-means and hierarchical clustering (Austin, Coull, Thomas, & Koutrakis, 2012), Bayesian profile regression (Coker et al., 2016; Molitor et al., 2011) and CART (Gass et al., 2014; Stingone et al., 2017) were utilized to generate exposure profiles of criteria air pollutants mixtures. Although these methods demonstrated their strength in establishing exposure profiles, they have their limitations. Bayesian method has limitation in dealing with high dimensional data (Coker et al., 2016; M. Oakes et al., 2014). Selection of number of clusters and interpretabilities of clusters are major concerns when using clustering methods (Austin et al., 2012; Billionnet et al., 2012). The results from decision trees algorithm (e.g., CART) may vary quite differently if the training data (different subset of data as training data) changes (Prasad, Iverson, & Liaw, 2006). Random forest randomly select predictors at each node to reduce the correlation of subtrees and reduce the overfitting of model when generate large number of trees (Berk, 2006; Breiman, 2001).

### 5.4.3 Strength and limitations of current study

Most of previous studies using multipollutant approaches were focusing on criteria air pollutants (Zanobetti, Austin, Coull, Schwartz, & Koutrakis, 2014) and rarely utilized multipollutant approach to study ambient air toxics such as outdoor VOCs (Davalos et al., 2017). No previous studies established exposure profiles of VOCs mixtures using random forest (Billionnet et al., 2012; Davalos et al., 2017; M. Oakes et al., 2014). Furthermore, most previous studies via multipollutant approaches utilized modeled data such as that from National Air Toxic Assessment (NATA) which has limitation to capture extrema (Rosenbaum et al., 1999; Scheffe et al., 2016). The strength of our current study can be generalized into three major aspects. First, our study is the first study targeting on exposure profiles of both indoor and outdoor VOC mixtures via multipollutant approach. Second, our study is the first study adopted random forest,

one of the very best classifier (Berk, 2006; Breiman, 2001), to classify and establish exposure profiles of VOCs mixtures. Thirdly, our analysis was on the basis of monitoring measurement. Particularly, supervised random forest utilized on monitoring measurement in our study was cross validated and was innovative in establishing exposure profiles. Additionally, compared to CART which create one decision tree with many subtrees that are potentially correlated, random forest can create less correlated subtrees and reduce bias via the random selection of predictors. Furthermore, RF generated large number of trees (2000 trees in our case), it also reduced the risks of overfitting data (Berk, 2006; Prasad et al., 2006). Rather than directly calculated concentration range of VOCs of each microenvironment based on documented microenvironment types, our study utilized cross-validated RF to reduce the random error in order to reflect the true exposure profile as close as possible. In terms of interpretability, as our monitoring measurement was from sample we collected in documented micro-environment, the general categories of micro-environment were known. Thus, after we utilized RF to classify the samples, we could easily interpret the categories with the five type of microenvironment we documented.

Our study also had limitations. Uncertainty in sampling was one of the limitations of the current study. The number of sampling sites was still limited considering the entire Memphis area. However, with limited funding and accessibility to certain locations, we chose the most representative locations for each microenvironment. We were not able to collect repeated samples for same locations due to limited funding and restricted time schedule. Our measurement might not reflect the long term exposure to VOCs mixtures. However, we were trying to estimate the range of concentrations of VOCs mixtures in different types of microenvironment and we collected multiple samples for same microenvironment from July to

November which captured the characteristics of warm and cold seasons. Significant weather changes during outdoor sampling and acute emission of VOCs from unexpected sources might also affect the measurement. Uncertainty in lab analysis was another limitation and was contributed by multiple factors. The major factor is the difference between the concertation of the analyte and the minimum detection limits (MDL) of the VOCs. Uncertainty increases when the concertation of the analyte approaches the MDL (C. Jia et al., 2006). Despite the uncertainties of samplings and lab analysis, our results might not be generalized to estimate exposure profiles of other counties because the exposure profile established based on our monitoring measurement was spatially specific. However, the cross-validated RF algorithm we applied can be used to classify VOCs samples with unknown category of microenvironment.

### 5.4.4 Further applications of current study

Assessing the exposure to multipollutant mixtures via multipollutant approach is the current and future trend in environmental health studies. Multipollutant approach shall be adopted to assess the exposure to air toxics mixtures as well. Exposure profiles is useful for studies which further explore the association between health risks and exposure (Austin et al., 2012; Molitor et al., 2016; M. M. Oakes et al., 2014; Pearce et al., 2016; Stingone et al., 2017). One of the goals of our study was to serve as an inspiration for future studies to further explore the exposure profile of air toxic mixtures with different statistical methods. The cross validated supervised RF can be further tested with data collected by future studies. Another goal of our study is to establish the exposure profile of VOCs mixtures for Memphis area. Our exposure profile can be utilized to prioritize VOCs subgroups for ad hoc regulation. Furthermore, the exposure profile established by current study can be further applied into studies examining the association between VOCs mixtures and various health risks in Memphis area.

## 5.5 Conclusion

Our study classified VOCs mixtures by 5 microenvironments and established the exposure profile of VOCs mixtures for each microenvironment via cross-validated supervised random forest machine learning. The VOC p-Isopropyltoluene, carbon tetrachloride and benzene were the top 3 VOCs to distinguish the spatial-specific exposure profile of VOCs mixtures. Toluene had consistent relative high average concentrations and wide range of concentrations in all profiled microenvironments. Given the level of toxicity, carbon tetrachloride, benzene and toluene need to be prioritized in Memphis area considering potential associated health risks. In conclusion, exposure profile is important in exposure assessment as the exposure to air toxics is not to individual air toxics but to mixture of air toxics. An exposure profile can highlight the exposure patterns in different microenvironment which is useful in determine the level and characteristics of the exposure. Additionally, exposure profiles served well for future studies that focus on the exposure assessment and investigating health risks from exposure to air toxics mixtures. A well-established exposure profile is useful to prioritize specific mixtures for strengthening air pollution regulations, facilitating decision making and improvement of health policies.

# Chapter 6

## Conclusion

Our living environment is closely associated with our health in many aspects. Air quality is one of the most important indexes that reflect the quality of our living environment. Air pollution affects our health every time we breathe. Both acute and chronic exposure to air pollution are associated with various adverse health outcome (Kampa & Castanas, 2008). To assess the exposure to air pollutants and examine the health disparity from exposure to air pollutants, the essential work is to have reliable measurement or appropriate estimation of air pollutants level in our living environment. Monitoring measurement and model estimation are two major source of air pollutants level. Model estimation from EPA NATA was widely used to examine health disparities from exposure of air pollutants due to the extended geographic coverage and convenient access of the data. NATA model have conservative assumption which tend to lead to overestimation of actual air pollutant level (U.S.EPA, 2016g). However, previous studies compared the model estimation to monitoring measurement and concluded that model estimation generally underestimate the monitoring measurement. Given that the model estimation is improved along time, assessing the latest model estimation from NATA 2011 is imperative for its numerous applications. On the basis of accurate measurement of air pollutant level, exposure to air toxics mixtures can be assessed through multipollutant approach rather than single pollutant approach because people expose to air toxic mixtures in reality. Exploring the interaction among multipollutant and assess the exposure to mixture of air pollutants have been prioritized (Dominici et al., 2010). People in different microenvironment usually expose to various mixtures of air pollutants. The variance of air pollutants mixture can be reflected in exposure profiles of air pollutants (Stingone et al., 2017). No previous studies established

exposure profiles via multipollutant approach for air toxics such as VOCs. Therefore, study exploring the exposure profiles of VOCs via multipollutant approach is in need. Variance in exposure to air pollutants could also be reflected on health disparities among people. As most previous studies on health disparities have limitations on only utilizing NATA modeled estimations for assessment, conducting a health disparity study adopting both NATA modeled estimations and monitored measurement can bring in alternative thoughts on assessing health disparity. Given that Memphis area clusters various air pollution emission sources, assessing the exposure to air pollutants and examine the health disparities from exposure to air pollutants are important for improving environmental health.

This thesis firstly evaluated the model estimation of air pollutants level from the latest NATA (NATA 2011) through comparison with monitored measurement from AQS at both national and EPA regional geographical scale. Secondly, the risk assessment was conducted based on the measurement from our local monitoring program REACT study in Memphis area. Thirdly, the health disparity from exposure to air toxics was examined using both NATA 2011 modeled measurements and monitoring measurements from REACT study. Finally, exposure profiles of VOCs mixtures in 5 typical microenvironments of Memphis area were established via supervised random forest machine learning.

The results from this thesis provided several implications for public health policies. In Memphis area, cancer risk from VOCs needs to be noticed for regulation because it is much higher than that at national level. Naphthalene need to be prioritized for regulation as it was the major cancer risk contributor. Toxicity on neurological system need to be flagged as it ranked as top non- cancer risks from exposure to VOCs for residents in Memphis area. p-Isopropyltoluene, carbon tetrachloride and benzene should also be prioritized for regulation as they were the top

three VOCs that distinguish the exposure profiles of different micro-environment in Memphis area.

This thesis provided strong evidence that comparison between modeled estimation and monitored measurement via the new framework with LOQ provided a method introducing laboratory analysis into statistical analysis for future assessment. Cross validated supervised random forest machining learning can be applied in future studies to establish exposure profiles of multipollutant mixtures. NATA might be utilized cautiously to serve as a practical alternative of monitoring measurement in assessment of health disparity from exposure to air pollutants but underestimation of certain air pollutants is the major drawback of NATA modeled estimation. More analysis on monitored measurement via local monitoring program is in need to provide more evidence on health disparity issues. In conclusion, conducting monitoring program for measurement of local VOCs level is important for assessing health risks from exposure to VOCs and health disparities from exposure to VOCs.

## References


Abel, T. D. (2008). Skewed riskscapes and environmental injustice: A case study of metropolitan St. Louis. *Environmental Management, 42*(2), 232-248. doi: 10.1007/s00267-008-9126-2

Anselin, L., Rey, S.J. (2014). *Modern spatial econometrics in practice: : A Guide to GeoDa, GeoDaSpace and PySAL* Chicago, IL: Geoda Press LLC.

Apelberg, B. J., Buckley, T. J., & White, R. H. (2005). Socioeconomic and racial disparities in cancer risk from air toxics in Maryland. *Environmental Health Perspectives, 113*(6), 693-699. doi: Doi 10.1289/Ehp.7609

ATSDR. (2004). Guidance Manual for the Assessment of Joint Toxic Action of Chemical Mixtures. Atlanta, GA: Agency for Toxic Substances and Disease Registry.

Austin, E., Coull, B., Thomas, D., & Koutrakis, P. (2012). A framework for identifying distinct multipollutant profiles in air pollution data. *Environment International, 45*, 112-121. doi: 10.1016/j.envint.2012.04.003

Bang, D. W., Manemann, S. M., Gerber, Y., Roger, V. L., Lohse, C. M., Rand-Weaver, J., . . . Juhn, Y. J. (2014). A Novel Socioeconomic Measure Using Individual Housing Data in Cardiovascular Outcome Research. *International Journal of Environmental Research and Public Health, 11*(11), 11597-11615. doi: 10.3390/ijerph111111597

Beckerman, B. S., Jerrett, M., Martin, R. V., van Donkelaar, A., Ross, Z., & Burnett, R. T. (2013). Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. *Atmospheric Environment, 77*, 172-177. doi: 10.1016/j.atmosenv.2013.04.024

Berk, R. A. (2006). An introduction to ensemble methods for data analysis. *Sociological Methods & Research, 34*(3), 263-295. doi: 10.1177/0049124105283119

Billionnet, C., Sherrill, D., Annesi-Maesano, I., & Study, G. (2012). Estimating the Health Effects of Exposure to Multi-Pollutant Mixture. *Annals of Epidemiology, 22*(2), 126-141. doi: 10.1016/j.annepidem.2011.11.004

Block, M. L., Elder, A., Auten, R. L., Bilbo, S. D., Chen, H. L., Chen, J. C., . . . Wright, R. J. (2012). The outdoor air pollution and brain health workshop. *Neurotoxicology, 33*(5), 972-984. doi: 10.1016/j.neuro.2012.08.014

Bobb, J. F., Valeri, L., Claus Henn, B., Christiani, D. C., Wright, R. O., Mazumdar, M., . . . Coull, B. A. (2015). Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics, 16*(3), 493-508. doi: 10.1093/biostatistics/kxu058

Bollati, V., Baccarelli, A., Hou, L. F., Bonzini, M., Fustinoni, S., Cavallo, D., . . . Yang, A. S. (2007). Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Research, 67*(3), 876-880. doi: 10.1158/0008-5472.CAN-06-2995

Bostrom, C. E., Gerde, P., Hanberg, A., Jernstrom, B., Johansson, C., Kyrklund, T., . . . Westerholm, R. (2002). Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environmental Health Perspectives, 110*, 451-488.

Breiman, L. (2001). Random Forest. Retrieved 3/4, 2018, from https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

Brook, R. D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., . . . Prevention Science of the American Heart, A. (2004). Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and

Prevention Science of the American Heart Association. *Circulation, 109*(21), 2655-2671. doi: 10.1161/01.CIR.0000128587.30041.C8

Calderon-Garciduenas, L., Azzarelli, B., Acuna, H., Garcia, R., Gambling, T. M., Osnaya, N., . . . Rewcastle, B. (2002). Air pollution and brain damage. *Toxicologic Pathology, 30*(3), 373-389. doi: Doi 10.1080/01926230252929954

Calderon-Garciduenas, L., Mora-Tiscareno, A., Ontiveros, E., Gomez-Garza, G., Barragan-Mejia, G., Broadway, J., . . . Engle, R. W. (2008). Air pollution, cognitive deficits and brain abnormalities: A pilot study with children and dogs. *Brain and Cognition, 68*(2), 117-127. doi: 10.1016/j.bandc.2008.04.008

Calderon-Garciduenas, L., Reed, W., Maronpot, R. R., Henriquez-Roldan, C., Delgado-Chavez, R., Calderon-Garciduenas, A., . . . Swenberg, J. A. (2004). Brain inflammation and Alzheimer's-like pathology in individuals exposed to severe air pollution. *Toxicologic Pathology, 32*(6), 650-658. doi: 10.1080/01926230490520232

Cantuaria, M. L., Brandt, J., Lofstrom, P., & Blanes-Vidal, V. (2017). Public perception of rural environmental quality: Moving towards a multi-pollutant approach. *Atmospheric Environment, 170*, 234-244. doi: 10.1016/j.atmosenv.2017.09.051

Cetin, E., Odabasi, M., & Seyfioglu, R. (2003). Ambient volatile organic compound (VOC) concentrations around a petrochemical complex and a petroleum refinery. *Science of the Total Environment, 312*(1-3), 103-112. doi: 10.1016/s0048-9697(03)00197-9

Chakraborty, J. (2012). Cancer risk from exposure to hazardous air pollutants: spatial and social inequities in Tampa Bay, Florida. *International Journal of Environmental Health Research, 22*(2), 165-183. doi: 10.1080/09603123.2011.628643

Chakraborty, J., Collins, T. W., & Grineski, S. E. (2017). Cancer risks from exposure to vehicular air pollution: a household level analysis of intra-ethnic heterogeneity in Miami, Florida. *Urban Geography, 38*(1), 112-136. doi: 10.1080/02723638.2016.1150112

Chakraborty, J., Collins, T. W., Grineski, S. E., & Maldonado, A. (2017). Racial Differences in Perceptions of Air Pollution Health Risk: Does Environmental Exposure Matter? *International Journal of Environmental Research and Public Health, 14*(2). doi: Artn 11610.3390/Ijerph14020116

Chakraborty, J., Maantay, J. A., & Brender, J. D. (2011). Disproportionate Proximity to Environmental Health Hazards: Methods, Models, and Measurement. *American Journal of Public Health, 101*, S27-S36. doi: 10.2105/Ajph.2010.300109

Chen, Z. H., Salam, M. T., Eckel, S. P., Breton, C. V., & Gilliland, F. D. (2015). Chronic effects of air pollution on respiratory health in Southern California children: findings from the Southern California Children's Health Study. *Journal of Thoracic Disease, 7*(1), 46-+. doi: 10.3978/j.issn.2072-1439.2014.12.20

Cheng, L., Fu, L., Angle, R. P., & Sandhu, H. S. (1997). Seasonal variations of volatile organic compounds in Edmonton, Alberta. *Atmospheric Environment, 31*(2), 239-246. doi: 10.1016/1352-2310(96)00170-7

Clark-Reyna, S. E., Grineski, S. E., & Collins, T. W. (2016). Residential exposure to air toxics is linked to lower grade point averages among school children in El Paso, Texas, USA. *Population and Environment, 37*(3), 319-340. doi: 10.1007/s11111-015-0241-8

Coker, E., Liverani, S., Ghosh, J. K., Jerrett, M., Beckerman, B., Li, A., . . . Molitor, J. (2016). Multi-pollutant exposure profiles associated with term low birth weight in Los Angeles County. *Environment International, 91*, 1-13. doi: 10.1016/j.envint.2016.02.011

Cook, R., Strum, M., Touma, J. S., Palma, T., Thurman, J., Ensley, D., & Smith, R. (2007). Inhalation exposure and risk from mobile source air toxics in future years. *Journal of Exposure Science and Environmental Epidemiology, 17*(1), 95-105. doi: 10.1038/sj.jes.7500529

Damek-Poprawa, M., & Sawicka-Kapusta, K. (2003). Damage to the liver, kidney, and testis with reference to burden of heavy metals in yellow-necked mice from areas around steelworks and zinc smelters in Poland. *Toxicology, 186*(1-2), 1-10. doi: Pii S0300-483x(02)00595-4Doi 10.1016/S0300-483x(02)00595-4

Das, M., & Aneja, V. P. (2003). Regional analysis of nonmethane volatile organic compounds in the lower troposphere of the Southeast United States. *Journal of Environmental Engineering-Asce, 129*(12), 1085-1103. doi: 10.1061/(asce)0733-9372(2003)129:12(1085)

Davalos, A. D., Luben, T. J., Herring, A. H., & Sacks, J. D. (2017). Current approaches used in epidemiologic studies to examine short-term multipollutant air pollution exposures. *Annals of Epidemiology, 27*(2), 145-153. doi: 10.1016/j.annepidem.2016.11.016

De Rosa, C. T., El-Masri, H. A., Pohl, H., Cibulas, W., & Mumtaz, M. M. (2004). Implications of chemical mixtures in public health practice. *Journal of Toxicology and Environmental Health-Part B-Critical Reviews, 7*(5), 339-350. doi: 10.1080/10937400490498075

deCastro, B. R. (2014). Acrolein and Asthma Attack Prevalence in a Representative Sample of the United States Adult Population 2000-2009. *Plos One, 9*(5). doi: ARTN e9692610.1371/journal.pone.0096926

Dejmek, J., Solansky, I., Benes, I., Lenicek, J., & Sram, R. J. (2000). The impact of polycyclic aromatic hydrocarbons and fine particles on pregnancy outcome. *Environmental Health Perspectives, 108*(12), 1159-1164. doi: Doi 10.2307/3434828

Delfino, R. J. (2002). Epidemiologic evidence for asthma and exposure to air toxics: Linkages between occupational, indoor, and community air pollution research. *Environmental Health Perspectives, 110*, 573-589.

Dominici, F., Peng, R. D., Barr, C. D., & Bell, M. L. (2010). Protecting Human Health From Air Pollution Shifting From a Single-pollutant to a Multipollutant Approach. *Epidemiology, 21*(2), 187-194. doi: 10.1097/EDE.0b013e3181cc86e8

Dominici, F., Wang, C., Crainiceanu, C., & Parmigiani, G. (2008). Model selection and health effect estimation in environmental epidemiology. *Epidemiology, 19*(4), 558-560. doi: 10.1097/EDE.0b013e31817307dc

Downey, L. (2007). US metropolitan-area variation in environmental inequality outcomes. *Urban Studies, 44*(5-6), 953-977. doi: 10.1080/00420980701256013

Downey, L., Dubois, S., Hawkins, B., & Walker, M. (2008). Environmental inequality in metropolitan America. *Organization & Environment, 21*(3), 270-294. doi: 10.1177/1086026608321327

Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software, 22*(4), 1-20.

Eastern Research Group, I. (2010, 12/17/2010). Results of the 2005 NATA Model-to-Monitor Comparison, Final Report. Retrieved 9/17, 2016, from http://www.epa.gov/ttn/atw/nata2005/05pdf/nata2005_model2monitor.pdf

EPA, U. S. (2016, 2/21/2016). Comparison of ASPEN Modeling System Results to Monitored Concentrations. Retrieved 12/19, 2017, from https://archive.epa.gov/airtoxics/nata/web/html/draft6.html#other_gases

Farley, R. (2004). Racial and ethnic residential segregation in the United States: 1980 to 2000. *Contemporary Sociology-a Journal of Reviews, 33*(1), 21-23.

Fernandes, R., Geeven, G., Soetens, S., & Klontza-Jaklova, V. (2011). Deletion/Substitution/Addition (DSA) model selection algorithm applied to the study of archaeological settlement patterning. *Journal of Archaeological Science, 38*(9), 2293-2300. doi: 10.1016/j.jas.2011.03.035

Fujita, E. M., Campbell, D. E., Arnott, W. P., Lau, V., & Martien, P. T. (2013). Spatial variations of particulate matter and air toxics in communities adjacent to the Port of Oakland. *Journal of the Air & Waste Management Association, 63*(12), 1399-1411. doi: 10.1080/10962247.2013.824393

Galbraith, D., Gross, S. A., & Paustenbach, D. (2010). Benzene and human health: A historical review and appraisal of associations with various diseases. *Critical Reviews in Toxicology, 40*, 1-46. doi: 10.3109/10408444.2010.508162

Garcia, E., Hurley, S., Nelson, D. O., Gunier, R. B., Hertz, A., & Reynolds, P. (2014). Evaluation of the agreement between modeled and monitored ambient hazardous air pollutants in California. *International Journal of Environmental Health Research, 24*(4), 363-377. doi: 10.1080/09603123.2013.835031

Gass, K., Klein, M., Chang, H. H., Flanders, W. D., & Strickland, M. J. (2014). Classification and regression trees for epidemiologic research: an air pollution example. *Environmental Health, 13*. doi: Artn 1710.1186/1476-069x-13-17

Gee, G. C., & Payne-Sturges, D. C. (2005). Environmental health disparities: A framework integrating psychosocial and environmental concepts (vol 112, pg 1645, 2004). *Environmental Health Perspectives, 113*(1), A18-A18.

George, B. J., Schultz, B. D., Palma, T., Vette, A. F., Whitaker, D. A., & Williams, R. W. (2011). An evaluation of EPA's National-Scale Air Toxics Assessment (NATA): Comparison with benzene measurements in Detroit, Michigan. *Atmospheric Environment, 45*(19), 3301-3308. doi: 10.1016/j.atmosenv.2011.03.031

Gilbert, A., & Chakraborty, J. (2011). Using geographically weighted regression for environmental justice analysis: Cumulative cancer risks from air toxics in Florida. *Social Science Research, 40*(1), 273-286. doi: 10.1016/j.ssresearch.2010.08.006

Godish, T., Davis, W. T., Fu, J. S. (2014). *Air Quality* (5th ed.). Boca Raton, FL: CRC press.

Goldizen, F. C., Sly, P. D., & Knibbs, L. D. (2016). Respiratory effects of air pollution on children. *Pediatric Pulmonology, 51*(1), 94-108. doi: 10.1002/ppul.23262

Greenbaum, D., & Shaikh, R. (2010). First Steps Toward Multipollutant Science for Air Quality Decisions. *Epidemiology, 21*(2), 195-197. doi: 10.1097/EDE.0b013e3181ccc52a

Greenberg, M. M. (1997). The central nervous system and exposure to toluene: A risk characterization. *Environmental Research, 72*(1), 1-7. doi: DOI 10.1006/enrs.1996.3686

Grineski, S. E., Clark-Reyna, S. E., & Collins, T. W. (2016). School-based exposure to hazardous air pollutants and grade point average: A multi-level study. *Environmental Research, 147*, 164-171. doi: 10.1016/j.envres.2016.02.004

Grineski, S. E., Collins, T. W., & Chakraborty, J. (2013). Hispanic heterogeneity and environmental injustice: intra-ethnic patterns of exposure to cancer risks from traffic-

related air pollution in Miami. *Population and Environment, 35*(1), 26-44. doi: 10.1007/s11111-012-0184-2

Grineski, S. E., Collins, T. W., Chakraborty, J., & McDonald, Y. J. (2013). Environmental Health Injustice: Exposure to Air Toxics and Children's Respiratory Hospital Admissions in El Paso, Texas. *Professional Geographer, 65*(1), 31-46. doi: 10.1080/00330124.2011.639625

Hansen, H., De Rosa, C. T., Pohl, H., Fay, M., & Mumtaz, M. M. (1998). Public health challenges posed by chemical mixtures. *Environmental Health Perspectives, 106*, 1271-1280.

HEI. (2010). HEI Strategic Plan for Understanding the Health Effects of Air Pollution. Boston, MA: Health Effects Institute.

Hidy, G. M., & Pennell, W. T. (2010). Multipollutant Air Quality Management. *Journal of the Air & Waste Management Association, 60*(6), 645-674. doi: 10.3155/1047.3289.60.6.645

Ho, K. E., Lee, S. C., Guo, H., & Tsai, W. (2004). Seasonal and diurnal variations of volatile organic compounds (VOCs) in the atmosphere of Hong Kong. *Science of the Total Environment, 322*(1-3), 155-166. doi: 10.1016/j.scitotenv.2003.10.004

Huang, B. B., Lei, C., Wei, C. H., & Zeng, G. M. (2014). Chlorinated volatile organic compounds (Cl-VOCs) in environment - sources, potential human health impacts, and current remediation technologies. *Environment International, 71*, 118-138. doi: 10.1016/j.envint.2014.06.013

Ilgen, E., Karfich, N., Levsen, K., Angerer, J., Schneider, P., Heinrich, J., . . . Begerow, J. (2001). Aromatic hydrocarbons in the atmospheric environment: Part I. Indoor versus outdoor sources, the influence of traffic. *Atmospheric Environment, 35*(7), 1235-1252. doi: 10.1016/s1352-2310(00)00388-5

Jakubiak-Lasocka, J., Lasocki, J., Siekmeier, R., & Chlopek, Z. (2015). Impact of Traffic-Related Air Pollution on Health. *Environment Exposure to Pollutants, 834*, 21-29. doi: 10.1007/5584_2014_14

James, W., Jia, C. R., & Kedia, S. (2012). Uneven Magnitude of Disparities in Cancer Risks from Air Toxics. *International Journal of Environmental Research and Public Health, 9*(12), 4365-4385. doi: 10.3390/ijerph9124365

Jarup, L. (2003). Hazards of heavy metal contamination. *British Medical Bulletin, 68*, 167-182. doi: 10.1093/bmb/ldg032

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., . . . Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology, 15*(2), 185-204. doi: 10.1038/sj.jea.7500388

Jex, T. T., & Wyman, D. O. (1996). A minireview of benzene. *Toxic Substance Mechanisms, 15*(2), 135-143.

Jia, C., Batterman, S., & Chernyak, S. (2006). Development and comparison of methods using MS scan and selective ion monitoring modes for a wide range of airborne VOCs. *Journal of Environmental Monitoring, 8*(10), 1029-1042. doi: 10.1039/b607042f

Jia, C., Batterman, S., & Godwin, C. (2008a). VOCs in industrial, urban and suburban neighborhoods - Part 2: Factors affecting indoor and outdoor concentrations. *Atmospheric Environment, 42*(9), 2101-2116. doi: 10.1016/j.atmosenv.2007.11.047

Jia, C., Batterman, S., & Godwin, C. (2008b). VOCs in industrial, urban and suburban neighborhoods, Part 1: Indoor and outdoor concentrations, variation, and risk drivers. *Atmospheric Environment, 42*(9), 2083-2100.

Jia, C. R., James, W., & Kedia, S. (2014). Relationship of Racial Composition and Cancer Risks from Air Toxics Exposure in Memphis, Tennessee, USA. *International Journal of Environmental Research and Public Health, 11*(8), 7713-7724. doi: 10.3390/ijerph110807713

Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution, 151*(2), 362-367. doi: 10.1016/j.envpol.2007.06.012

Kelishadi, R., & Poursafa, P. (2011). Obesity and Air Pollution: Global Risk Factors for Pediatric Non-alcoholic Fatty Liver Disease. *Hepatitis Monthly, 11*(10), 794-802. doi: 10.5812/kowsar.1735143X.746

Kim, K. H., Jahan, S. A., Kabir, E., & Brown, R. J. C. (2013). A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. *Environment International, 60*, 71-80. doi: 10.1016/j.envint.2013.07.019

Kwon, J., Weisel, C. P., Turpin, B. J., Zhang, J. F., Korn, L. R., Morandi, M. T., . . . Colome, S. (2006). Source proximity and outdoor-residential VOC concentrations: Results from the RIOPA study. *Environmental Science & Technology, 40*(13), 4074-4082. doi: 10.1021/es051828u

Lam, H. R., Ladefoged, O., Ostergaard, G., Lund, S. P., & Simonsen, L. (1996). Four weeks' inhalation exposure of rats to p-cymene affects regional and synaptosomal neurochemistry. *Pharmacology & Toxicology, 79*(5), 225-230. doi: DOI 10.1111/j.1600-0773.1996.tb00264.x

Lewtas, J. (2007). Air pollution combustion emissions: Characterization of causative agents and mechanisms associated with cancer, reproductive, and cardiovascular effects. *Mutation Research-Reviews in Mutation Research, 636*(1-3), 95-133. doi: 10.1016/j.mrrev.2007.08.003

Linder, S. H., Marko, D., & Sexton, K. (2008). Cumulative cancer risk from air pollution in Houston: Disparities in risk burden and social disadvantage. *Environmental Science & Technology, 42*(12), 4312-4322. doi: 10.1021/es072042u

Lippy, B. E., & Turner, R. W. (1991). Complex-Mixtures in Industrial Workspaces - Lessons for Indoor Air-Quality Evaluations. *Environmental Health Perspectives, 95*, 81-83.

Little, T. A. (2015). Method Validation Essentials, Limit of Blank, Limit of Detection, and Limit of Quantitation. *Biopharm International, 28*(4), 48-51.

Logue, J. M., Small, M. J., & Robinson, A. L. (2011). Evaluating the national air toxics assessment (NATA): Comparison of predicted and measured air toxics concentrations, risks, and sources in Pittsburgh, Pennsylvania. *Atmospheric Environment, 45*(2), 476-484. doi: 10.1016/j.atmosenv.2010.09.053

Logue, J. M., Small, M. J., Stern, D., Maranche, J., & Robinson, A. L. (2010). Spatial Variation in Ambient Air Toxics Concentrations and Health Risks between Industrial-Influenced, Urban, and Rural Sites. *Journal of the Air & Waste Management Association, 60*(3), 271-286. doi: 10.3155/1047-3289.60.3.271

Loh, M. M., Levy, J. I., Spengler, J. D., Houseman, E. A., & Bennett, D. H. (2007). Ranking cancer risks of organic hazardous air pollutants in the United States. *Environmental Health Perspectives, 115*(8), 1160-1168. doi: 10.1289/ehp.9884

Loh, M. M., Levy, J. I., Spengler, J. D., Houseman, E. A., & Bennett, D. H. (2007). Ranking cancer risks of organic hazardous air pollutants in the United States. *Environ Health Perspect, 115*(8), 1160-1168. doi: 10.1289/ehp.9884

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., . . . Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature, 435*(7043), 834-838. doi: 10.1038/nature03702

Lupo, P. J., & Symanski, E. (2009). A Comparative Analysis of Modeled and Monitored Ambient Hazardous Air Pollutants in Texas: A Novel Approach Using Concordance Correlation. *Journal of the Air & Waste Management Association, 59*(11), 1278-1286. doi: 10.3155/1047-3289.59.11.1278

Maantay, J. (2007). Asthma and air pollution in the Bronx: Methodological and data considerations in using GIS for environmental justice and health research. *Health & Place, 13*(1), 32-56. doi: 10.1016/j.healthplace.2005.09.009

Macias, T. (2016). Environmental risk perception among race and ethnic groups in the United States. *Ethnicities, 16*(1), 111-129. doi: 10.1177/1468796815575382

Marshall, J. D., Teoh, S. K., & Nazaroff, W. W. (2005). Intake fraction of nonreactive vehicle emissions in US urban areas. *Atmospheric Environment, 39*(7), 1363-1371. doi: 10.1016/j.atmosenv.2004.11.006

Mayberry, R. M., Mili, F., & Ofili, E. (2000). Racial and ethnic differences in access to medical care. *Medical Care Research and Review, 57*, 108-145. doi: Doi 10.1177/107755800773743628

McCarthy, M. C., O'Brien, T. E., Charrier, J. G., & Hather, H. R. (2009). Characterization of the Chronic Risk and Hazard of Hazardous Air Pollutants in the United States Using Ambient Monitoring Data. *Environmental Health Perspectives, 117*(5), 790-796. doi: 10.1289/ehp.11861

Mendell, M. J. (2007). Indoor residential chemical emissions as risk factors for-respiratory and allergic effects in children: a review. *Indoor Air, 17*(4), 259-277.

Mohamed, M. F., Kang, D. W., & Aneja, V. P. (2002). Volatile organic compounds in some urban locations in United States. *Chemosphere, 47*(8), 863-882. doi: Pii S0045-6535(02)00107-8Doi 10.1016/S0045-6535(02)00107-8

Molitor, J., Coker, E., Jerrett, M., Ritz, B., Li, A., & Health Review, C. (2016). Part 3. Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes. *Res Rep Health Eff Inst*(183 Pt 3), 3-47.

Molitor, J., Su, J. G., Molitor, N. T., Rubio, V. G., Richardson, S., Hastie, D., . . . Jerrett, M. (2011). Identifying Vulnerable Populations through an Examination of the Association Between Multipollutant Profiles and Poverty. *Environmental Science & Technology, 45*(18), 7754-7760. doi: 10.1021/es104017x

Morello-Frosch, R., & Jesdale, B. M. (2006). Separate and unequal: Residential segregation and estimated cancer risks associated with ambient air toxics in US metropolitan areas. *Environmental Health Perspectives, 114*(3), 386-393. doi: 10.1289/ehp.8500

Morello-Frosch, R., & Lopez, R. (2006). The riskscape and the color line: Examining the role of segregation in environmental health disparities. *Environmental Research, 102*(2), 181-196. doi: 10.1016/j.envres.2006.05.007

Mortimer, K., Neugebauer, R., Lurmann, F., Alcorn, S., Balmes, J., & Tager, I. (2008). Air pollution and pulmonary function in asthmatic children - Effects of prenatal and lifetime exposures. *Epidemiology, 19*(4), 550-557. doi: 10.1097/EDE.0b013e31816a9dcb

Nadeau, K., McDonald-Hyman, C., Noth, E. M., Pratt, B., Hammond, S. K., Balmes, J., & Tager, I. (2010). Ambient air pollution impairs regulatory T-cell function in asthma. *Journal of Allergy and Clinical Immunology, 126*(4), 845-U280. doi: 10.1016/j.jaci.2010.08.008

NATA. (1999). 1999 National-Scale Air Toxics Assessment. Retrieved 3.15, 2016, from https://archive.epa.gov/airtoxics/nata1999/web/html/index.html

NIOSH. (2005). Mixed Exposures Research Agenda. Cincinnati, OH: National Institute for Occupational Safety and Health.

Novaes, P., Saldiva, P. H., Matsuda, M., Macchione, M., Rangel, M. P., Kara-Jose, N., & Berra, A. (2010). The effects of chronic exposure to traffic derived air pollution on the ocular surface. *Environ Res, 110*(4), 372-374. doi: 10.1016/j.envres.2010.03.003

Air quality management in the United States (2004).

Oakes, M., Baxter, L., & Long, T. C. (2014). Evaluating the application of multipollutant exposure metrics in air pollution health studies. *Environment International, 69*, 90-99. doi: 10.1016/j.envint.2014.03.030

Oakes, M. M., Baxter, L. K., Duvall, R. M., Madden, M., Xie, M. J., Hannigan, M. P., . . . Long, T. C. (2014). Comparing Multipollutant Emissions-Based Mobile Source Indicators to Other Single Pollutant and Multipollutant Indicators in Different Urban Areas. *International Journal of Environmental Research and Public Health, 11*(11), 11727-11752. doi: 10.3390/ijerph111111727

ODPHP. (2017, 05/25/17). Envrionmental Health. Retrieved 5/26, 2017, from https://www.healthypeople.gov/2020/topics-objectives/topic/environmental-health

ODPHP. (2018, 1/3/2018). Determinants of Health. Retrieved 1/3, 2018, from https://www.healthypeople.gov/2020/about/foundation-health-measures/Determinants-of-Health

Ozkaynak, H., Palma, T., Touma, J. S., & Thurman, J. (2008). Modeling population exposures to outdoor sources of hazardous air pollutants. *Journal of Exposure Science and Environmental Epidemiology, 18*(1), 45-58. doi: 10.1038/sj.jes.7500612

Park, K. H., & Jo, W. K. (2004). Personal volatile organic compound (VOC) exposure of children attending elementary schools adjacent to industrial complex. *Atmospheric Environment, 38*(9), 1303-1312. doi: 10.1016/j.atmosenv.2003.11.032

Pastor, M., Morello-Frosch, R., & Sadd, J. L. (2005). The air is always cleaner on the other side: Race, space, and ambient air toxics exposures in California. *Journal of Urban Affairs, 27*(2), 127-148. doi: 10.1111/j.0735-2166.2005.00228.x

Payne-Sturges, D., & Gee, G. C. (2006). National environmental health measures for minority and low-income populations: Tracking social disparities in environmental health. *Environmental Research, 102*(2), 154-171. doi: DOI 10.1016/j.envres.2006.05.014

Payne-Sturges, D. C., Burke, T. A., Breysse, P., Diener-West, M., & Buckley, T. J. (2004). Personal exposure meets risk assessment: A comparison of measured and modeled exposures and risks in an urban community. *Environ Health Perspect, 112*(5), 589-598.

Pearce, J. L., Waller, L. A., Sarnat, S. E., Chang, H. H., Klein, M., Mulholland, J. A., & Tolbert, P. E. (2016). Characterizing the spatial distribution of multiple pollutants and populations at risk in Atlanta, Georgia. *Spatial and Spatio-Temporal Epidemiology, 18*, 13-23. doi: 10.1016/j.sste.2016.02.002

Perlin, S. A., Sexton, K., & Wong, D. W. S. (1999). An examination of race and poverty for populations living near industrial sources of air pollution. *J Expo Anal Environ Epidemiol, 9*(1), 29-48. doi: DOI 10.1038/sj.jea.7500024

Perlin, S. A., Wong, D., & Sexton, K. (2001). Residential proximity to industrial sources of air pollution: Interrelationships among race, poverty, and age. *Journal of the Air & Waste Management Association, 51*(3), 406-421.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems, 9*(2), 181-199. doi: 10.1007/s10021-005-0054-1

Pratt, G. C., Wu, C. Y., Bock, D., Adgate, J. L., Ramachandran, G., Stock, T. H., . . . Sexton, K. (2004). Comparing air dispersion model predictions with measured concentrations of VOCs in urban communities. *Environmental Science & Technology, 38*(7), 1949-1959. doi: 10.1021/es0306381

Qian, Z. M., Chapman, R. S., Hu, W., Wei, F. S., Korn, L. R., & Zhang, J. F. J. (2004). Using air pollution based community clusters to explore air pollution health effects in children. *Environment International, 30*(5), 611-620. doi: 10.1016/j.envint.2003.11.003

Qian, Z. M., Zhang, J. F., Korn, L. R., Wei, F. S., & Chapman, R. S. (2004). Factor analysis of household factors: are they associated with respiratory conditions in Chinese children? *International Journal of Epidemiology, 33*(3), 582-588. doi: 10.1093/ije/dyg278

Raddatz, L., & Mennis, J. (2013). Environmental Justice in Hamburg, Germany. *Professional Geographer, 65*(3), 495-511. doi: 10.1080/00330124.2012.700500

Rager, J. E., Smeester, L., Jaspers, I., Sexton, K. G., & Fry, R. C. (2011). Epigenetic Changes Induced by Air Toxics: Formaldehyde Exposure Alters miRNA Expression Profiles in Human Lung Cells. *Environmental Health Perspectives, 119*(4), 494-500. doi: 10.1289/ehp.1002614

Rehwagen, M., Schlink, U., & Herbarth, O. (2003). Seasonal cycle of VOCs in apartments. *Indoor Air, 13*(3), 283-291. doi: DOI 10.1034/j.1600-0668.2003.00206.x

Rice, L. J., Jiang, C. S., Wilson, S. M., Burwell-Naney, K., Samantapudi, A., & Zhang, H. M. (2014). Use of Segregation Indices, Townsend Index, and Air Toxics Data to Assess Lifetime Cancer Risk Disparities in Metropolitan Charleston, South Carolina, USA. *International Journal of Environmental Research and Public Health, 11*(5), 5510-5526. doi: 10.3390/ijerph110505510

Rider, C. V., Carlin, D. J., DeVito, M. J., Thompson, C. L., & Walker, N. J. (2013). Mixtures research at NIEHS: An evolving program. *Toxicology, 313*(2-3), 94-102. doi: 10.1016/j.tox.2012.10.017

Roberts, A. L., Lyall, K., Hart, J. E., Laden, F., Just, A. C., Bobb, J. F., . . . Weisskopf, M. G. (2013). Perinatal air pollutant exposures and autism spectrum disorder in the children of Nurses' Health Study II participants. *Environ Health Perspect, 121*(8), 978-984. doi: 10.1289/ehp.1206187

Roberts, S., & Martin, M. (2005). A critical assessment of shrinkage-based regression approaches for estimating the adverse health effects of multiple air pollutants. *Atmospheric Environment, 39*(33), 6223-6230. doi: 10.1016/j.atmosenv.2005.07.004

Rosenbaum, A. S., Axelrad, D. A., Woodruff, T. J., Wei, Y. H., Ligocki, M. P., & Cohen, J. P. (1999). National estimates of outdoor air toxics concentrations. *Journal of the Air & Waste Management Association, 49*(10), 1138-1152.

Rumchev K, Brown H, & Spickett J. (2007). Volatile organic compounds: Do they present a risk to our health? *Reviews on Environmental Health, 22*, 39-55.

Sacks, J. D., Ito, K., Wilson, W. E., & Neas, L. M. (2012). Impact of Covariate Models on the Assessment of the Air Pollution-Mortality Association in a Single- and Multipollutant Context. *American Journal of Epidemiology, 176*(7), 622-634. doi: 10.1093/aje/kws135

Samet, J. M., & Speizer, F. E. (1993). Introduction and recommendations - working group on indoor air and other complex-mixtures. *Environmental Health Perspectives, 101*, 143-147.

Sarigiannis, D. A., Karakitsios, S. P., Gotti, A., Liakos, I. L., & Katsoyiannis, A. (2011). Exposure to major volatile organic compounds and carbonyls in European indoor environments and associated health risk. *Environment International, 37*(4), 743-765. doi: 10.1016/j.envint.2011.01.005

Sarnat, J. A., Marmur, A., Klein, M., Kim, E., Russell, A. G., Sarnat, S. E., . . . Tolbert, P. E. (2008). Fine particle sources and cardiorespiratory morbidity: An application of chemical mass balance and factor analytical source-apportionment methods. *Environmental Health Perspectives, 116*(4), 459-466. doi: 10.1289/ehp.10873

Scheffe, R. D., Strum, M., Phillips, S. B., Thurman, J., Eyth, A., Fudge, S., . . . Cook, R. (2016). Hybrid Modeling Approach to Estimate Exposures of Hazardous Air Pollutants (HAPs) for the National Air Toxics Assessment (NATA). *Environmental Science & Technology, 50*(22), 12356-12364. doi: 10.1021/acs.est.6b04752

Schlosberg, D. (2004). Reconceiving environmental justice: Global movements and political theories. *Environmental Politics, 13*(3), 517-540. doi: 10.1080/0964401042000229025

Seagrave, J., McDonald, J. D., Bedrick, E., Edgerton, E. S., Gigliotti, A. P., Jansen, J. J., . . . Mauderly, J. L. (2006). Lung toxicity of ambient particulate matter from southeastern US sites with different contributing sources: Relationships between composition and effects. *Environmental Health Perspectives, 114*(9), 1387-1393. doi: 10.1289/ehp.9234

Shmool, J. L. C., Kubzansky, L. D., Newman, O. D., Spengler, J., Shepard, P., & Clougherty, J. E. (2014). Social stressors and air pollution across New York City communities: a spatial approach for assessing correlations among multiple exposures. *Environmental Health, 13*. doi: Artn 9110.1186/1476-069x-13-91

Sinisi, S. E., & van der Laan, M. J. (2004). Deletion/substitution/addition algorithm in learning with applications in genomics. *Stat Appl Genet Mol Biol, 3*, Article18. doi: 10.2202/1544-6115.1069

Sram, R. J., Binkova, B. B., Dejmek, J., & Bobak, M. (2005). Ambient air pollution and pregnancy outcomes: A review of the literature. *Environmental Health Perspectives, 113*(4), 375-382. doi: Doi 10.1289/Ehp.6362

Stieb, D. M., Chen, L., Eshoul, M., & Judek, S. (2012). Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. *Environmental Research, 117*, 100-111. doi: 10.1016/j.envres.2012.05.007

Stingone, J. A., Pandey, O. P., Claudio, L., & Pandey, G. (2017). Using machine learning to identify air pollution exposure profiles associated with early cognitive skills among US children. *Environmental Pollution, 230*, 730-740. doi: 10.1016/j.envpol.2017.07.023

Stoner, A. M., Anderson, S. E., & Buckley, T. J. (2013). Ambient Air Toxics and Asthma Prevalence among a Representative Sample of US Kindergarten-Age Children. *Plos One, 8*(9). doi: UNSP e7517610.1371/journal.pone.0075176

Strobl, C., Malley, J., & Tutz, G. (2009). An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. *Psychological Methods, 14*(4), 323-348. doi: 10.1037/a0016973

Stroud, C. A., Zaganescu, C., Chen, J., McLinden, C. A., Zhang, J. H., & Wang, D. (2016). Toxic volatile organic air pollutants across Canada: multi-year concentration trends, regional air quality modelling and source apportionment. *Journal of Atmospheric Chemistry, 73*(2), 137-164. doi: 10.1007/s10874-015-9319-z

Stuart, A. L., Mudhasakul, S., & Sriwatanapongse, W. (2009). The Social Distribution of Neighborhood-Scale Air Pollution and Monitoring Protection. *Journal of the Air & Waste Management Association, 59*(5), 591-602. doi: 10.3155/1047-3289.59.5.591

Suh, H. H., Bahadori, T., Vallarino, J., & Spengler, J. D. (2000). Criteria air pollutants and toxic air pollutants. *Environmental Health Perspectives, 108*, 625-633. doi: Doi 10.2307/3454398

Suh, H. H., Zanobetti, A., Schwartz, J., & Coull, B. A. (2011). Chemical Properties of Air Pollutants and Cause-Specific Hospital Admissions among the Elderly in Atlanta, Georgia. *Environmental Health Perspectives, 119*(10), 1421-1428. doi: 10.1289/ehp.1002646

Sun, J. (2007). D-limonene: Safety and clinical applications. *Alternative Medicine Review, 12*(3), 259-264.

Sun, Z. C., Tao, Y. B., Li, S., Ferguson, K. K., Meeker, J. D., Park, S. K., . . . Mukherjee, B. (2013). Statistical strategies for constructing health risk models with multiple pollutants and their interactions: possible choices and comparisons. *Environmental Health, 12*. doi: Artn 8510.1186/1476-069x-12-85

Sunyer, J., Esnaola, M., Alvarez-Pedrerol, M., Forns, J., Rivas, I., Lopez-Vicente, M., . . . Querol, X. (2015). Association between Traffic-Related Air Pollution in Schools and Cognitive Development in Primary School Children: A Prospective Cohort Study. *Plos Medicine, 12*(3), 1792-1792.

Symanski, E., Tee Lewis, P. G., Chen, T. Y., Chan, W., Lai, D., & Ma, X. (2016). Air toxics and early childhood acute lymphocytic leukemia in Texas, a population based case control study. *Environ Health, 15*(1), 70. doi: 10.1186/s12940-016-0154-8

Thioulouse, J., & Dray, S. (2007). Interactive multivariate data analysis in R with the ade4 and ade4TkGUI packages. *Journal of Statistical Software, 22*(5), 1-14.

Tyrrell, J., Melzer, D., Henley, W., Galloway, T. S., & Osborne, N. J. (2013). Associations between socioeconomic status and environmental toxicant concentrations in adults in the USA: NHANES 2001-2010. *Environment International, 59*, 328-335. doi: 10.1016/j.envint.2013.06.017

U.S. EPA. (2009a). Air Pollution Research Improves Understanding of Multipollutant Impacts on Human Health. Washington D.C.: U.S. Environmental Protect Agency

U.S. EPA. (2009b). Air Toxics Data Analysis Workbook. Retrieved 1/28, 2016, from http://www3.epa.gov/ttnamti1/files/ambient/airtox/workbook/AirToxicsWorkbook6-09.pdf

U.S.EPA. (2002a, 2/21/2016). Comparison of 2002 Model-Predicted Concentrations to Monitored Data. Retrieved 4/13, 2016, from https://archive.epa.gov/nata2002/web/pdf/2002compare.pdf

U.S.EPA. (2002b). Compiling the Point Source 2002 NEI: Whose Data Get Priority? Retrieved 4/14, 2017, from https://www3.epa.gov/ttnchie1/conference/ei14/session9/pope_pres.pdf

U.S.EPA. (2008). The multi-pollutant report: technical concepts and examples. Retrieved 1/12, 2018, from https://archive.epa.gov/airquality/aqmp/web/pdf/20080702_multipoll.pdf
U.S.EPA. (2009). Air Toxics Data Analysis Workbook. Retrieved 4/12, 2017, from https://www3.epa.gov/ttnamti1/files/ambient/airtox/workbook/AirToxicsWorkbook6-09.pdf
U.S.EPA. (2010). Results of the 2005 NATA Model-to-Monitor Comparison, Final Report. Retrieved 6/6, 2016, from http://www.epa.gov/ttn/atw/nata2005/05pdf/nata2005_model2monitor.pdf
U.S.EPA. (2014). Air Quality System 2016, from http://www.epa.gov/aqs
U.S.EPA. (2015a, 5/18/2016). Annual Summary Data. Retrieved 6/14, 2016, from http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html
U.S.EPA. (2015b, 10/29/2015). Risk Assessment and Modeling - Air Toxics Risk Assessment Reference Library. from https://www.epa.gov/fera/risk-assessment-and-modeling-air-toxics-risk-assessment-reference-library
U.S.EPA. (2015c). SUPPORT FOR THE EPA NATIONAL CONTRACT FOR LEAD ANALYSIS Retrieved 6/18, 2017, from https://www3.epa.gov/ttnamti1/files/ambient/pb/nationalcontractforleadanalysis.pdf
U.S.EPA. (2015d, 12). Technical Support Document EPA's 2011 National-scale Air Toxics Assessment Retrieved 6/7, 2016, from https://www.epa.gov/sites/production/files/2015-12/documents/2011-nata-tsd.pdf
U.S.EPA. (2016a). 2011 NATA: Assessment Results. Retrieved 6/14, 2016, from https://www.epa.gov/national-air-toxics-assessment/2011-nata-assessment-results
U.S.EPA. (2016b, 12/26/2016). 2011 NATA: Fact Sheet. Retrieved 10/30/2017, 2017, from https://www.epa.gov/national-air-toxics-assessment/2011-nata-fact-sheet
U.S.EPA. (2016c, 12/26/2017). 2011 NATA: Summary of Results. Retrieved 10/30, 2017, from https://www.epa.gov/national-air-toxics-assessment/2011-nata-summary-results
U.S.EPA. (2016d, 11/1/2016). Air Toxics - Monitoring Methods. Retrieved 6/18, 2017, from https://www3.epa.gov/ttnamti1/airtox.html
U.S.EPA. (2016e). Ambient Monitoring Technology Information Center (AMTIC). Retrieved 3/11, 2018, from https://www3.epa.gov/ttnamti1/airtoxpg.html
U.S.EPA. (2016f). EPA's National Air Toxics Assessment (NATA) and the role of TRI data. Retrieved 3/11, 2018, from https://www.epa.gov/sites/production/files/2016-11/documents/strum_tri_training_conference_nata.pdf
U.S.EPA. (2016g, 3/30/2016). NATA overview. Retrieved 6/10, 2016, from https://www.epa.gov/national-air-toxics-assessment/nata-overview
U.S.EPA. (2016h, 3/30/2016). Previous NATA Assessments. Retrieved 6/6, 2016, from https://www.epa.gov/national-air-toxics-assessment/previous-nata-assessments
U.S.EPA. (2017a, 6/16/2016). 2011 NATA: Assessment Methods. Retrieved 4/12, 2017, from https://www.epa.gov/national-air-toxics-assessment/2011-nata-assessment-methods
U.S.EPA. (2017b, 11/2/2016). Air Emissions Sources. Retrieved 4/13, 2017, from https://www.epa.gov/air-emissions-inventories/air-emissions-sources
U.S.EPA. (2017c, 6/27/2016). Environmental Justice Indexes in EJSCREEN. Retrieved 4/22, 2017, from https://www.epa.gov/ejscreen/environmental-justice-indexes-ejscreen
U.S.EPA. (2017d, 9/23/2016). EPA Region 3 Guidance on Handling Chemical Concentration Data Near the Detection Limit in Risk Assessments. Retrieved 4/26, 2017, from

https://www.epa.gov/risk/epa-region-3-guidance-handling-chemical-concentration-data-near-detection-limit-risk
U.S.EPA. (2017e, 11/9/2016). IRIS chromium (VI). Retrieved 9/25, 2017, from https://cfpub.epa.gov/ncea/iris2/chemicallanding.cfm?substance_nmbr=144
U.S.EPA. (2017f, 10/26/2016). National Emissions Inventory (NEI). Retrieved 4/13, 2017, from https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei
U.S.EPA. (2017g, 12/15/2016). Overview of Demographic Indicators in EJSCREEN. Retrieved 4/22, 2017, from https://www.epa.gov/ejscreen/overview-demographic-indicators-ejscreen
U.S.EPA. (2017h, 6/3/2016). Purposes and Uses of EJSCREEN. Retrieved 4/22, 2017, from https://www.epa.gov/ejscreen/purposes-and-uses-ejscreen
U.S.EPA. (2017i, 3/27/2017). Toxics Release Inventory (TRI) Program Retrieved 4/14, 2017, from https://www.epa.gov/toxics-release-inventory-tri-program/tri-data-quality

USEPA. (2006). Total Risk Integrated Methodology (TRIM) Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4). *Office Of Air Quality Planning And Standards., U.S. Environmental Protection Agency*.
USEPA. (2012). National Air Toxics Assessments. Retrieved 08/12, 2012, from http://www.epa.gov/nata/
Uzoigwe, J. C., Prum, T., Bresnahan, E., & Garelnabi, M. (2013). The emerging role of outdoor and indoor air pollution in cardiovascular disease. *N Am J Med Sci, 5*(8), 445-453. doi: 10.4103/1947-2714.117290
Vennam, L. P., Vizuete, W., & Arunachalam, S. (2015). Evaluation of model-predicted hazardous air pollutants (HAPs) near a mid-sized US airport. *Atmospheric Environment, 119*, 107-117. doi: 10.1016/j.atmosenv.2015.08.015
Volk, H. E., Hertz-Picciotto, I., Delwiche, L., Lurmann, F., & McConnell, R. (2011). Residential Proximity to Freeways and Autism in the CHARGE Study. *Environmental Health Perspectives, 119*(6), 873-877. doi: 10.1289/ehp.1002835
Wang, S. Q., Zhang, J. L., Zeng, X. D., Zeng, Y. M., Wang, S. C., & Chen, S. Y. (2009). Association of Traffic-Related Air Pollution with Children's Neurobehavioral Functions in Quanzhou, China. *Environmental Health Perspectives, 117*(10), 1612-1618. doi: 10.1289/ehp.0800023
Wang, S. W., Tang, X. G., Fan, Z. H., Wu, X. M., Lioy, P. J., & Georgopoulos, P. G. (2009). Modeling of Personal Exposures to Ambient Air Toxics in Camden, New Jersey: An Evaluation Study. *Journal of the Air & Waste Management Association, 59*(6), 733-746. doi: 10.3155/1047-3289.59.6.733
WHO. (2017a). Air Pollution. Retrieved 4/11, 2017, from http://www.who.int/topics/air_pollution/en/
WHO. (2017b). Mortality and burden of disease from ambient air pollution. Retrieved 4/18/2017, 2017, from http://www.who.int/gho/phe/outdoor_air_pollution/burden_text/en/
Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine, 62*(7), 1768-1784. doi: 10.1016/j.socscimed.2005.08.036
Wilson, S., Burwell-Naney, K., Jiang, C. S., Zhang, H. M., Samantapudi, A., Murray, R., . . . Williams, E. (2015). Assessment of sociodemographic and geographic disparities in

cancer risk from air toxics in South Carolina. *Environ Res, 140*, 562-568. doi: 10.1016/j.envres.2015.05.016

Woodruff, T. J., Wells, E. M., Holt, E. W., Burgin, D. E., & Axelrad, D. A. (2007). Estimating risk from ambient concentrations of acrolein across the United States. *Environmental Health Perspectives, 115*(3), 410-415. doi: 10.1289/ehp.9467

Wu, Y. C., & Batterman, S. A. (2006). Proximity of schools in Detroit, Michigan to automobile and truck traffic. *Journal of Exposure Science and Environmental Epidemiology, 16*(5), 457-470. doi: 10.1038/sj.jes.7500484

Yu, H. F., & Stuart, A. L. (2016). Exposure and inequality for select urban air pollutants in the Tampa Bay area. *Science of the Total Environment, 551*, 474-483. doi: 10.1016/j.scitotenv.2016.01.157

Yu, S. C., Eder, B., Dennis, R., Chu, S. H., & Schwartz, S. E. (2006). New unbiased symmetric metrics for evaluation of air quality models. *Atmospheric Science Letters, 7*(1), 26-34. doi: 10.1002/asl.125

Zanobetti, A., Austin, E., Coull, B. A., Schwartz, J., & Koutrakis, P. (2014). Health effects of multi-pollutant profiles. *Environment International, 71*, 13-19. doi: 10.1016/j.envint.2014.05.023

Zhou, X. H., & Gao, S. (1997). Confidence intervals for the log-normal mean. *Statistics in Medicine, 16*(7), 783-790. doi: Doi 10.1002/(Sici)1097-0258(19970415)16:7<783::Aid-Sim488>3.0.Co;2-2

Zhou, Y., Li, C. Y., Huijbregts, M. A. J., & Mumtaz, M. M. (2015). Carcinogenic Air Toxics Exposure and Their Cancer-Related Health Impacts in the United States. *Plos One, 10*(10). doi: ARTN e014001310.1371/journal.pone.0140013

Zhu, X. L., Fan, Z. H., Wu, X. M., Meng, Q. Y., Wang, S. W., Tang, X. G., . . . Lioy, P. (2008). Spatial variation of volatile organic compounds in a "Hot Spot" for air pollution. *Atmospheric Environment, 42*(32), 7329-7338. doi: 10.1016/j.atmosenv.2008.07.039