ASSESSMENT OF PERSONAL EXPOSURE TO AIR POLLUTION BASED ON TRAJECTORY DATA

by

Guixing Wei, M.S., B.S.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Geographic Information Science
August 2018

Committee Members:

F. Benjamin Zhan, Chair

Yongmei Lu

Alexander Savelyev

Michael Kramer

ACKNOWLEDGEMENTS

First, I would specifically like to thank my advisor Dr. F. Benjamin Zhan for his infinite patience, parental encouragement, and expert guidance. I greatly appreciate the "freedom" Dr. Zhan has granted me for exploring my research interests. This work would not have been possible without his tremendous support. Dr. Zhan is also an ideal role model for me to look up to as I try to discover myself. All the things I have learned from him will be a great treasure for my future career.

I would also like to thank my committee members: Dr. Yongmei Lu, Dr. Alexander Savelyev, and Dr. Michael Kramer. I am grateful for their insightful comments and dedicated time and efforts. I have been so happy for having an opportunity to work with Dr. Lu. Especially, her previous research is largely relevant to this work. I sincerely thank Dr. Alexander Savelyev who has encouraged me to think critically in developing research questions. Special thanks go to Dr. Michael Kramer for his dedicated time and efforts of reviewing and commenting on my proposal and dissertation.

I would like to heartfully thank Dr. Yuan. For four years of working with her, her professionalism and diligence have motivated me to be a responsible scholar. I greatly appreciate all her sincere advice on my study and life.

I would like to thank my wife Xujiao (Joy) Wang for taking care of our family. Her love, understanding, and company have been important during my doctoral study. She is always my biggest motivation to move forward fearlessly.

Many thanks also go to the department of geography. Allison Glass-Smith, Pat Hell-Jones, Dan Hemenway, Stella LoPachin, Charles Robinson, Angelika Wahl, and Joyce Wilkerson are greatly helpful throughout my doctoral study.

I also thank all my dear friends herein: Md Niaz Morshed, David Mills, Shade Dixon, Han Lu, Wei Peng, Ugochukwu Francis Umeokafor, He Jin, Zhaochen Song, Xiaowen Cui, Wen Lu, Hao (Andy) Zhang and his parents, Boran Zhao, Xinyu Cai, Haizhong Qian, Xuelin Wu, Fu Ren, Mogahid Hussein, Christine Ames, Mark Deka, Monica Medel Valdivia, Ryan Schuermann, Nathaniel Dede-Bamfo, Yahan Teng, Francesco Zignol, and Erin Dorothea Dascher.

More importantly, to my family, Mom, Dad, my two lovely sisters, my two cute nephews, my upcoming daughter and those unnamed, thank you all so much!

Finally, I would like to thank Texas State University Graduate College, Texas State University Student Government, Sandra E. Estaville and Dr. Lawrence Estaville, Dr. F. Benjamin Zhan for the financial support during my doctoral study.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS vi

LIST OF TABLES viii

LIST OF FIGURES ix

ABSTRACT xi

1. INTRODUCTION 1
    1.1 Background 1
    1.2 Problem Statement 2
    1.3 Objective and Specific Aims 3
    1.4 Significance 5

2. LITERATURE REVIEW 7
    2.1 Literature Review of Air Pollution Exposure Assessment 7
        2.1.1 Bibliometrics of Air Pollution Exposure Assessment 7
        2.1.2 Categories of Air Pollution Exposure Assessment 10
    2.2 Literature Review of Time-activity Integrated Exposure Assessment Study 16
        2.2.1 The Spatiotemporal Essence of Exposure 16
        2.2.2 Diary-based Methods 18
        2.2.3 GPS-enabled Methods 20
        2.2.4 Limitations and Critical Needs 26

3. A MODEL FOR ASSESSING PERSONAL EXPOSURE TO AMBIENT AIR POLLUTION USING TRAJECTORY DATA 28
    3.1 Introduction 28
    3.2 Method 28
        3.2.1 A Conceptualization of Trajectories Documenting Human Activities in Geographic Space and Time 28
        3.2.2 Model Development 29
    3.3 Experiments based on Simulated Data 36
        3.3.1 Data 36
        3.3.2 Experimental Results 40
    3.4 Experiments based on Empirical Data 51

3.4.1 Pre-processing of Trajectory Data 51
3.4.2 Model Implementation Process.. 51
3.4.3 Results of Exposure Assessment on Empirical Data 53
3.4.4 Difference between Estimated Exposures based on Home Locations Only and Locations of All Activities on Empirical data 57
3.5 Conclusion and Discussions 58

4. QUANTIFYING PERSONAL EXPOSURE TO TRAFFIC EMISSION WHEN HUMAN TRAJECTORY AND HOURLY TRAFFIC VARIATION ARE CONSIDERED 61
4.1 Introduction.. 61
4.2 Study Area and Data 63
4.2.1 Study Area 63
4.2.2 Traffic Count Data 63
4.2.3 Simulated Individual Trajectory Data.. 65
4.3 Methods.. 66
4.3.1 Determination of Hourly Temporal Allocation Factors......................... 66
4.3.2 Construction of an Annual Hourly Traffic Emission Model 67
4.3.3 Linking of the Hourly Traffic Emission Model and Trajectories 69
4.3.4 Exposure Assessment.. 69
4.4 Results and Discussions 69
4.4.1 Results of TAFs 69
4.4.2 Results of the Annual Average Hourly Traffic Emission Model 71
4.4.3 Results of Personal Exposure to Traffic Emissions by Hours 74
4.5 Conclusions.. 78

5. CONCLUSIONS.. 80
5.1 Summary of Findings and Discussion 80
5.2 Limitations... 82
5.3 Future Research Direction 83

LITERATURE CITED 86

LIST OF TABLES

Table **Page**

2.1 Database search strategies (Web of Science, 1970 - 2017). 8

2.2 Comparison of exposure assessment methods.. 12

2.3 A review of GPS-enabled exposure assessment studies.. 22

3.1 Descriptive information of the speed profile 31

3.2 Descriptive information of time allocation in simulation.. 38

3.3 Time spent on visits of all MEs by four individuals.. 42

3.4 Average hourly exposure (AHE) associated with each VISIT 45

3.5 Average hourly exposure (AHE) and total exposure (TE) associated with each ME.. 46

3.6 Summary statistic of estimated exposures in different MEs associated with 100 simulated trajectories 47

3.7 Wilcoxon Signed-rank test of three trajectory datasets 50

3.8 Characteristics of 100 trajectories.. 51

3.9 Average hourly exposure (AHE) associated with all VISITs.................................... 54

3.10 Average hourly exposure (AHE) and total exposure (TE) associated with each ME 55

3.11 Summary statistic of estimated exposures in different MEs associated with 100 trajectories.. 56

3.12 Wilcoxon Signed-rank test of three trajectory datasets 58

4.1 The descriptive statistics of personal exposures of 100 trajectories........................... 74

LIST OF FIGURES

Figure **Page**

2.1 Publication records per year from Web of Science 9

2.2 Geographic distribution of retrieved study related to air pollution exposure assessment from Web of Science 10

3.1 Four different microenvironments (MEs) and different modes of travel associated with a person's activity patterns in geographic space and time 29

3.2 The workflow of the proposed trajectory-based exposure assessment model 30

3.3 Linking trajectory data with estimated air pollution data represented by a 3D space-time cube. 34

3.4 Geographical distribution of monitoring sites in the Beijing area 36

3.5 An example simulated trajectory documenting an individual's movement patterns in one day. 40

3.6 The results of trajectory data processing for the trajectory of Person 1. 41

3.7 Map layers showing the geographic distribution of the estimated intensity of $PM_{2.5}$ at each hour in Beijing over the 24-hour period. 44

3.8 Descriptive statistics of personal exposure for 4 hypothetical people in Beijing during a day... 47

3.9 Summary statistic of estimated exposures in different MEs associated with 100 simulated trajectories in Beijing over a 24-hour period... 48

3.10 The results of trajectory data processing of one empirical trajectory 53

3.11 Descriptive statistics of personal exposure for four people in Beijing during a day 55

3.12 Summary information of estimated exposures in different MEs based on an analysis of 100 trajectories in Beijing over a 24-hour period............... 57

4.1 Study area: Minneapolis-Saint Paul.. 63

4.2 Geographic distribution of traffic recorder stations and roadway system 64

4.3 Simulated trajectory of one person's activity patterns in the
Minneapolis-Saint Paul metropolitan area in one day 66

4.4 Hourly TAFs across thirty-five traffic recorder stations on workdays in 2016 70

4.5 Interpolated maps of TAFs in twenty-four hours in the study area 71

4.6 Hourly traffic density maps in twenty-four hours.. 72

4.7 A close-up view of the traffic density map in hour 8 along with
highlighted traffic density hotspots.. 73

4.8 Traffic density map in hour 17 along with highlighted traffic density hotspots......... 74

4.9 Trajectories 56, 49 and 45 along with traffic density hotspots 75

4.10 Exposure profiles of persons 56, 49 and 45 over the course of 24 hours 77

ABSTRACT

Air pollution has been among the biggest environmental risks to human health. Exposure assessment to air pollution is essentially a procedure to quantify the degree to which people get exposed to hazardous air pollution. Exposure assessment is also a critical step in health-related studies exploring the relationship between personal exposure to environmental stressors and adverse health outcomes. Given the critical role of exposure assessment, it is important to accurately quantify and characterize personal exposure in geographic space and time.

For years numerous exposure assessment methods have been developed with respect to a wide spectrum of air pollutants. Of all the methods, the most commonly used one is to use a representative geographic unit as the surrogate location to estimate the potential impact from hazardous air pollution from differing sources on that location. The representative unit is one person's home location in most cases. Such studies, however, have failed to recognize the significance of both the dynamics of human activities and the variation of air pollution in geographic space and time.

It is believed that personal exposure is essentially a function of space and time as an individual's time-activity patterns and intensities of air pollutant in question vary over space and time. It is therefore imperative to account for the spatiotemporal dynamics of both in exposure assessment. To this end, the goal of this study is to account for the spatiotemporal dynamics of both human time-activity patterns and air pollution for

assessing personal exposure. More specifically this dissertation aims to achieve three objectives as summarized below.

First, in light of the deficiency of existing home-based exposure assessment methods, this study proposes an innovative trajectory-based model for assessing personal exposure to ambient air pollution. This model provides a computational framework for assessing personal exposure when trajectories, documenting human spatiotemporal activities, are modeled into a series of tours, microenvironments (MEs), and visits. A set of individual-level trajectories was simulated to test the performance of the proposed model, in conjunction with one-day air pollution ($PM_{2.5}$) data in Beijing, China. The results from the test demonstrated that the trajectory-based model is capable of capturing the spatiotemporal variation of personal exposure, thus providing more accurate, detailed and enriched information to better understand personal exposure. The findings indicate that there is considerable variation in intra-microenvironment and inter-microenvironment exposure, which identified the importance of distinguishing between different MEs. Moreover, this study tested the proposed model using an empirical dataset.

Second, little is known about the difference between the estimated exposure based on home locations only and that considering the locations of all human activities. To fill this gap, this study aims to test whether the exposure calculated from the home-based method is statistically significantly different from the exposure estimated by the newly developed trajectory-based model. A Dataset containing 4,000 individual-level one-day trajectories (Dataset 1) was simulated to test the aforementioned hypothesis. The exposure estimates in comparison are the average hourly exposure over a 24-hour period from two exposure assessment methods. The 4,000 trajectories were split into another two subsets (Datasets 2, 3) according to the difference between home-based exposure estimates and trajectory-based exposure estimates. The Wilcoxon Signed-rank test was used to evaluate whether the difference between the two models is significant. The results show that the statistically significant difference was found only in Dataset 3. The same test was also applied to a set of empirical trajectories. The significant difference exists in the results from the empirical data. The mixed results suggest that additional research is needed to verify the difference between the two exposure assessment methods.

Third, little research has taken into consideration of hourly traffic variation and human activities simultaneously in a model for assessing personal exposure to traffic emissions. To fill this gap, this study develops a new trajectory-based model to quantify personal exposure to traffic emissions. The hourly share of daily traffic volume of each roadway in the study area was estimated by calculating the traffic allocation factors (TAFs) of each roadway. Next, the hourly traffic emission surfaces were built using the hourly shares and a kernel density algorithm. A 3-D cube representing the spatiotemporal distribution of traffic emission was constructed, which overlaid the simulated individual-level trajectory data for assessing personal exposure to traffic emissions. The results showed that people's time-activity patterns (e.g., where an individual lives/works, where an individual travels) were significant factors in exposure assessment. This study suggests that people's time activities and hourly variation of traffic emission should be simultaneously addressed when assessing personal exposure to traffic emissions.

To sum up, this study has devoted a large effort in quantifying and characterizing personal exposure in geographic space and time. A few of contributions to the knowledge of exposure science are listed as follows. First, this study contributes two exposure assessment models in characterizing personal spatiotemporal exposure using trajectory data. One is developed for assessing personal exposure to ambient air pollution, and the other one is for assessing personal exposure to traffic emissions. Second, this study demonstrates the intra- and inter-microenvironment variation of personal exposure and reveals the significance of people's time-activity patterns in exposure assessment. Third, this study investigates the difference in exposure estimates between conventional home-based methods considering home locations only and trajectory-based methods accounting for the locations of all activities. The mixed findings from Wilcoxon Signed-rank tests suggest more research is needed to explore how personal exposure varies with time-activity patterns. All these contributions will have important implications in exposure science, environment science, and epidemiology.

1. INTRODUCTION

1.1 Background

According to a World Health Organization (WHO)'s report, 92% of the world population lived and got exposed to unsafe polluted air in 2014 (World Health Organization 2017). Ambient and indoor air pollution were estimated to have caused 6.5 million deaths worldwide in 2012, which accounted for 11.6% of all deaths (World Health Organization 2017). In addition to mortality, numerous epidemiological studies have demonstrated that air pollution is associated with a series of adverse health effects including cardiovascular illness (Fiordelisi et al. 2017; Luben et al. 2017), respiratory symptoms (Zhang et al. 2015), and impaired neurological function (Heusinkveld et al. 2016; Lee et al. 2017).

One key aspect of studying the health effect of air pollution is human exposure assessment. There has been a myriad of methods developed for exposure assessment, either by measuring, known as direct exposure assessment methods, or modeling, known as indirect exposure assessment methods. Given the prohibitive cost and being hard to apply to a large population study of direct methods, most health-related studies go with modeling techniques for assessing exposure. Of all the indirect methods, a widely used method is to employ a person's home location to estimate the exposure to any air pollutant of interest. For instance, Leal and Chaix (2011) found that 90% of investigated papers solely used study participants' home addresses to evaluate exposure in cardiometabolic studies. Personal exposure, however, is essentially a phenomenon varying over space and time. In other words, in reality, people go to different places and come into contact with varying air pollutants when going about their daily activities. The mere consideration of the home location, therefore, appears not to reflect the true scenarios of personal exposure, thus introducing a considerable amount of error into the exposure estimates. The introduced error ultimately biases the reliability of the subsequent epidemiological findings.

Given the deficiency of home-based exposure assessment methods, a currently emerging need is to simultaneously consider personal time-activity patterns and the spatiotemporal variation of air pollution in the development of appropriate exposure assessment models. Moreover, it is imperative to characterize how personal exposure varies in geographic space and time for better understanding the health impact of the hazardous air pollutant in question. To this end, this study aims to develop models to integrate the spatiotemporal dynamics of both human activities and air pollution into the process of exposure assessment. In addition, this study seeks to investigate whether the consideration of these two dynamics would result in a statistically significant difference in exposure estimates.

1.2 Problem Statement

Exposure assessment modeling has been extensively studied. However, there are several limitations in the literature. First, there is a tendency among health practitioners to utilize home-based methods to quantify personal exposure. Second, most studies considering the spatiotemporal dynamics of both human activities and air pollution heavily rely on expensive equipment to measure personal exposure instead of through modeling techniques, which limits the application of exposure assessment in subsequent epidemiological studies (e.g., a large population cohort study). Third, little research concentrates on developing models to integrate trajectory data and air pollution data for exposure assessment. Fourth, few studies have investigated the difference between home-based exposure assessment methods and dynamics-resolved models regardings how estimated exposure varies in geographic space and time.

It is imperative to account for the human time-activity patterns and spatiotemporal dynamics of air pollutant in question in exposure assessment modeling, which can introduce many beneficial effects in understanding the associated health effect as well as the exposure itself. For instance, some studies reveal that the development of models that can capture personal time-activity patterns and spatiotemporally varying air pollution is an urgent and significant research challenge in exposure science, public health, and epidemiology field (Gerharz, Kruger, and Klemm 2009). Moreover, the spatiotemporal portrait of personal exposure can be beneficial for designing an effective intervention and control strategy (Adams, Riggs, and Volckens 2009). This study suggests that where, when, how long, and how much of the air pollutant appears are crucial factors in assessing personal exposure.

1.3 Objective and Specific Aims

The overarching objective of this study is twofold. The first is to develop trajectory-based models for assessing personal exposure by taking account of spatiotemporal dynamics of both human activities and air pollution. The second focus is to characterize and understand how personal exposure varies over space and time. This study is comprised of 3 core sections. The first section focuses on the development of a new exposure assessment model for ambient air pollution using individual-level trajectory data. The second section aims to develop a personal exposure assessment model for quantifying personal exposure to the traffic emission using individual-level trajectory data. Unlike the first two sections that concentrate on the modeling, the third section is to explore whether a trajectory-based exposure assessment approach would produce significantly different exposure results when compared to a home-based approach.

The specific aims are listed as follows:

Aim 1: To develop a trajectory-based model for assessing personal exposure to ambient air pollution. The air pollution data collected by a ground monitoring network is commonly used in environmental science to study the phenomena related to ambient air pollution. Given the widespread presence and use of this type of data, this study builds up an innovative exposure assessment model integrating individual-level human trajectory data and the ambient air pollution data. Instead of using home location alone, this newly proposed exposure assessment model analyzes the human time-activity patterns from trajectory data and characterizes the spatiotemporally varying air pollutant in question from the acquired ambient air pollution data. As a result, the proposed model is tested using the simulated trajectory data and obtained air pollution data.

Aim 2: To evaluate whether the consideration of spatiotemporal dynamics of human activities and air pollution creates significantly different exposure results when compared to home-based approaches. It is believed that models considering these two dynamics mentioned above outplay the home-based methods, but few studies investigated the difference between the two approaches. This study simulates a large number of individual-level trajectory data with differing time-activity patterns. For each trajectory, it is associated with two exposure estimates quantified from the aforementioned two methods. The difference between the two estimates of each trajectory is statistically examined. This examination is important to understand how human activities and the variation of air pollutant in question affect the assessment of personal exposure.

Aim 3: To develop a trajectory-based model for assessing personal exposure to traffic emissions. Traffic emissions are primary sources of air pollution in an urban setting. It is imperative to quantify how many traffic emissions people get exposed to when they go about their daily activities. This newly built model simultaneously considers both hourly variation of traffic emission and individual-level human trajectory in assessing personal traffic exposure to traffic. This model is helpful in investigating how personal traffic exposure varies in geographic space and time.

1.4 Significance

First, it is believed that our newly developed models can greatly improve exposure assessment as the models incorporate human time-activity patterns and spatiotemporal variation of air pollutant in question. For instance, several studies have revealed that the spatiotemporal time-activity pattern is a significant determinant of personal exposure (Dons, Int Panis, et al. 2011; Setton et al. 2011; Park and Kwan 2017).

Second, this study will help mitigate uncertain geographic context problem (UGCoP) and exposure misclassification to some extent. UGCoP is the problem in geography field that the spatial configuration of appropriate contextual influence is not fully understood and the timing and duration of personal exposure to these contextual stressors is uncertain (Kwan 2012). Exposure misclassification, in exposure science, refers to a situation in which the estimated exposure is not the correct one (Savitz 2003). Both problems introduce bias or uncertainty to the subsequent epidemiological studies since estimated exposure is not representative of the actual exposure. This study captures the full course of human activities and spatiotemporal variation of air pollution when assessing personal exposure, thus offering a more accurate, reliable estimate of exposure.

Finally, this study responds to the calling of National Research Council (NRC) about the development of mixed models to link Global Positioning System (GPS) data with environmental exposure (National Research Council 2012; Dias and Tchepel 2014). To this end, this study contributes two personal exposure assessment models linking different sources of air pollution data with individual-level trajectory data. The proposed models have important implications for many exposure- and health-related applications.

2. LITERATURE REVIEW

2.1 Literature Review of Air Pollution Exposure Assessment

This section presents the literature review of air pollution exposure assessment. This review is beneficial to understand the current status and trends of air pollution exposure assessment. The advantages and disadvantages of the different exposure assessment methods are also discussed in this section.

2.1.1 Bibliometrics of Air Pollution Exposure Assessment

This study conducted two bibliometric analysis to understand the trend of the exposure assessment research focused on air pollution. The literature search was conducted in the database of Web of Science in August 2017. The targeted research was articles written in the English language that were published from 1970 to 2016. The first search focused on the research of air pollution exposure assessment, whereas the second search concentrated on time-activity involved study related to air pollution exposure assessment. The search strategy and search keywords used in both queries are listed in Table 2.1.

Table 2.1 Database search strategies (Web of Science, 1970 - 2017).

No.	Search terms - 1	Search terms - 2
1	exposure	exposure
2	air*	air*
3	assess*	assess*
4	quantif*	quantif*
5	estimat*	estimat*
6		time-activity
7		time
8		activit*
9		mobility
10		microenvironment*
11		micro-environment*
12	#1 AND #2 AND (#3 OR #4 OR #5)	#1 AND #2 AND (#3 OR #4 OR #5) AND (#6 OR (#7 AND #8) OR #9 OR # 10 OR #11)

Figure 2.1 illustrates the number of papers per year for both searches. Both searches took place on August 10, 2017. A total of 19,980 records were returned for the first search, which covers all studies involving air pollution exposure assessment. However, only 1,872 records, were retrieved from the second search, which focuses on the time-activity related study of air pollution exposure assessment. As can be seen, both areas had a growing trend of literature regardless of the differences in the speed of growth. Before 1990, studies about air pollution exposure assessment were almost nonexistent. There were only 20 publications on air pollution exposure assessment studies. However, after 1991, the research of air pollution exposure assessment proliferates greatly. This rapid development coincided with the significant report, *Human Exposure Assessment for Airborne Pollutants* (National Research Council 1991a), published by the National Research Council in 1991. This report has laid the foundation for exposure assessment and motivated more research

to be devoted to this scientific discipline. One notable difference between the two trends is that, unlike the rapid development of air pollution exposure assessment discipline, the proportion of time-activity related study in exposure assessment field is steadily low, comprising only 10% of the total air pollution exposure assessment studies on average. This indicates that air pollution exposure assessment studies have paid insufficient attention to the significance of the time-activity information.

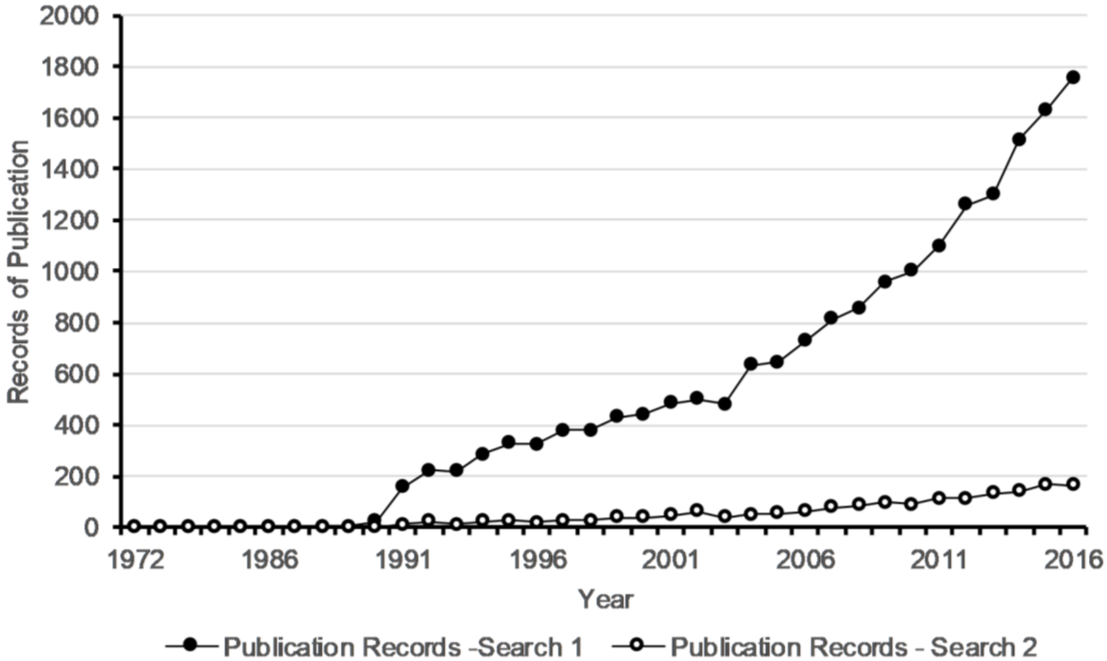

Figure 2.1 Publication records per year from Web of Science. (Search 1 covers all air pollution exposure assessment studies, whereas search 2 focuses on time-activity related air pollution exposure assessment studies)

Figure 2.2 depicts the global geographical distribution of air pollution exposure assessment research along with the time-activity related studies. In total, 126 countries have contributed to the air pollution exposure assessment research. It is notable that publications in this field were not evenly spatially distributed across the world. The top twenty countries accounted for up to 85% of the whole literature for both searches in terms of publication records. Most of the studies are located in North America (e.g., U.S.A, Canada) and European countries (e.g., Germany, Italy). Very few studies come from South Asia whereby air pollution has been a significant issue for a long time.

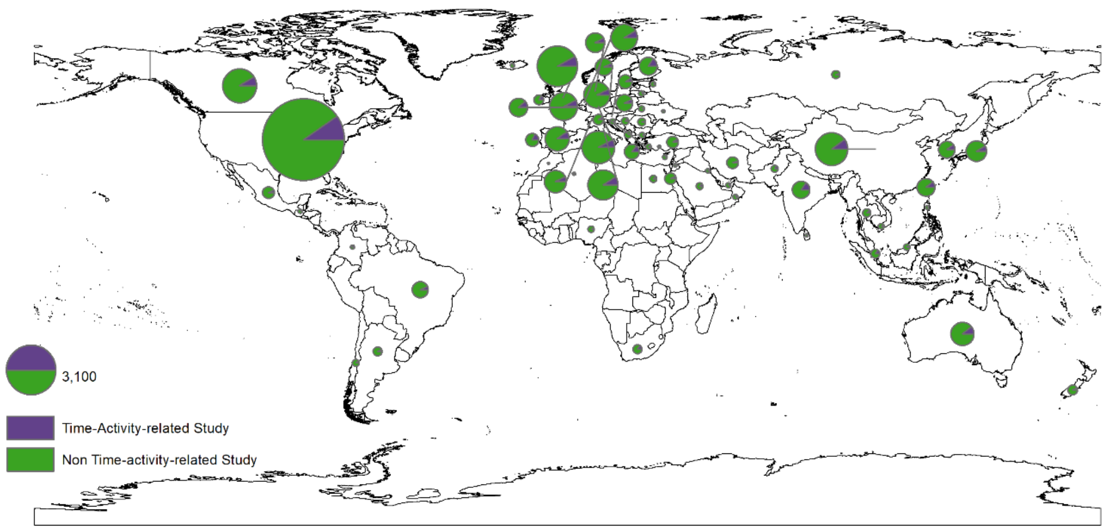

Figure 2.2 Geographic distribution of retrieved study related to air pollution exposure assessment from Web of Science. (Each chart is comprised of time-activity related and non-time-activity research in the air pollution exposure assessment field. The size of a graph is proportional to the publication records of air pollution exposure assessment studies.)

2.1.2 Categories of Air Pollution Exposure Assessment

The definition of human exposure was first presented in the 1980s as "the event that a person comes in contact with the pollutant" (Ott 1982). It can be simply thought of as a study of stressors, receptors, and their contacts. In the recent years, there has been an increasing focus on the field of exposure assessment, as illustrated in Figure 2.1. The principal goal of an exposure assessment is to obtain an accurate, precise estimate of the actual exposure that a person experienced.

Exposure assessment methods can be classified into different tracks. The most straightforward and cost-effective way to quantify personal exposure is to sort exposure estimates into a set of nominal groups. For example, the groups can be yes and no, or high, medium and low based on experts' suggestion. In contrast, the costliest way to quantify personal exposure is to directly measure it using various sets of equipment, known as direct methods. For instance, a personal biological monitoring device can be utilized to measure the response of an organism to infer how much of a specific toxic substance enters it. Alternatively, when measurement methods are not available or cost-prohibitive, personal exposure can be indirectly estimated through exposure modeling, also known as direct methods.

A detailed comparison is illustrated in Table 2.2. Each method has its advantages and disadvantages. It is believed that accuracy increases along with the cost of exposure assessment. Nevertheless, in many cases, the choice of exposure assessment method is a balancing act that considers feasibility (Armstrong 1996).

Table 2.2 Comparison of exposure assessment methods

Categories	Methods	Examples	Advantages	Disadvantages
Direct methods	Personal monitoring	• Mobile/portable monitoring devices	• Most accurate estimation • used as a benchmark to validate other models	• Labor intensive • Costly • hard to carry out • small sampling size
	Biomonitoring	• Biomarkers	• Considers all routes of exposure • Can measure internal dose and effective dose • It can reflect a long-time exposure in a retrospective way	• Data collection is expensive • Analytical methods are elaborate and difficult to reproduce • Not all biomarkers have reference values.
Indirect methods	Interpolation models	• Deterministic: IDW • Stochastic: Kriging, Spline	• Allow the use of real pollution measurements instead of surrogates • Create a distance-weighting continuous surface	• The interpolation method is mechanistic • Algorithms fall apart at the edge due to the lack of monitoring data • Variability might be exaggerated.
	Dispersion models	• AERMOD modeling system • CALPUFF modeling system	• Can integrate various atmospheric conditions and the atmospheric motion • Dynamic modeling capabilities • Represent complex pollutant pathways that lead to secondary pollutants	• Intensive data requirement • High need for the knowledge of meteorology and climatology • Extensive computational resources • Spatial resolution is low.
	Hybrid models	• Combination of personal monitoring and regional monitoring • Combination of models at regional, urban and local scales	• Combine the strengths of different models/data	• Personal data is not always accessible. • The integration of multiple models involving different scales is hard to surmount.

2.1.2.1 Direct Methods

Direct personal exposure assessment methods are comprised of personal monitoring and biomonitoring. Both methods can produce personal exposure measurements or estimates best without any prior assumptions. The exposure assessed by these methods is referred to as the "gold standard."

Personal monitoring methods usually ask participants to wear or carry mobile monitoring devices to record the "actual exposure" rather than "potential exposure" in a direct, explicit and continuous way. Usually personal monitoring methods also consider obtaining time-activity information from either diaries/surveys (Gauvin et al. 2001; Dennekamp et al. 2002; Lai et al. 2004; Piechocki-Minguy et al. 2006) or GPS-enabled devices (Elgethun et al. 2002; Greaves, Issarayangyun, and Liu 2008; Nethery et al. 2014; Fang 2012; Fang and Lu 2012) . Diary-based personal monitoring studies not only require participants to wear electronic monitoring equipment but also request them to record their 24-h time-activities during the study period. For instance, participants need to record their location when in different microenvironments (home, work, in-transit, other, etc.) and also the time spent in each microenvironment (Nethery, Teschke, and Brauer 2008). GPS-enabled studies record participants' movements continuously over a short time interval, usually in units of a few seconds, at a high accuracy.

Biomonitoring method is the use of various biomarkers to assess the actual air pollution concentration inhaled by an individual instead of the concentration to which an individual is exposed. Besides measuring exposure events and doses, biomonitoring methods can also accurately quantify individual metabolite(s) in biological media (e.g., urine, breath, blood, heart, lung), which better clarify the linkage in exposure-to-effect and reduce the associated uncertainties (Hubal et al. 2000; Sobus et al. 2011). For instance, a positive association was found between daily exposure to PM_{10} (atmospheric particulate matter having less than 10 micrometers in diameter) and oxidative stress and impaired cardiovascular (Liu et al. 2007).

2.1.2.2 Indirect Methods

Indirect personal exposure methods utilize various exposure models to quantify exposure. The data commonly used for modeling is the ambient air pollution data, which is collected by an air pollution monitoring network. Some other auxiliary data (e.g., land use data, temperature data, etc.) can also be utilized in the modeling. Some researchers have reviewed and classified the existing models into different categories (Klepeis et al. 2001; Jerrett et al. 2005; Zou et al. 2009).

When compared to direct methods, indirect methods offer many tangible benefits. First, it allows researchers to model the dynamics of human activities and air pollution separately (Fang and Lu 2012). Second, it saves experimental cost to a large extent, especially when dealing with a large population. It's not always possible to request all participants to carry a monitoring device. Third, with the aid of GIS, the indirect methods can easily incorporate other types of environmental data in modeling.

Of all the indirect methods, the microenvironment model is the one that is commonly used for modeling human time activities in exposure assessment. This model treats personal total exposure as a list of distinct exposure scenarios (Duan 1982; Klepeis et al. 2001). The typical microenvironments include home, school, workplace, and in-transit (Adgate et al. 2004; Lee et al. 2012; Wu et al. 2005; Zhang and Batterman 2009; Lazenby et al. 2012). Some studies have used finer microenvironment types such as kitchen room, living room, shopping, restaurant, or various transportation modes (Harrison et al. 2002; Rojas-Bracho et al. 2002; Briggs et al. 2003; Saksena et al. 2003; Shimada and Matsuoka 2011; Wheeler et al. 2011). Instead of solely considering the home location, microenvironment models can incorporate a number of significant places in modeling personal exposure.

2.1.2.3 Limitations and Critical Needs

Direct exposure assessment is believed to be the most accurate estimate of actual exposure. Nevertheless, it cannot be applied to any large cohort study or longitudinal studies due to the following limitations: (a) the cost of a large-population or longterm study is prohibitively high; (b) monitoring devices or biomarkers are highly intrusive and may bring in substantial burdens to participants' daily activities, especially for the elderly and children (Bricka et al. 2012); (c) The diary-based direct exposure methods heavily rely on the participants' subjective inputs.

Another challenge is that many health-related studies that took place before resulted in a large variety of data, some of which may be the diary data containing massive time-activity information or the air pollution data collected by a field monitoring campaign. Therefore, how to combine different sources of data is an emerging problem that needs to be solved in exposure science disciplines. Meanwhile, the National Research Council (NRC) has called for the development of models to link human GPS data with environmental data to seek a better way to characterize personal exposure (National Research Council 2012).

The limitations of direct methods, the challenge of data fusion, and the realization of imperativeness of spatiotemporal dynamics of personal exposure have encouraged this study to develop models to integrate human trajectory data with spatiotemporally varying environmental data in a unified platform. GIS provides a powerful means of joining diverse data in one georeferenced coordinate system. This study makes the full use of GIS techniques in developing integrative models, as well as characterizing personal exposure in a spatiotemporal manner.

2.2 Literature Review of Time-activity Integrated Exposure Assessment Study

This section provides a review of air exposure assessment research that incorporates the human time-activity information in modeling personal exposure. The importance of spatiotemporal dynamics of personal exposure is discussed. The two methods of collecting human time-activity information, diary-based methods, GPS-enabled methods are reviewed and evaluated.

2.2.1 The Spatiotemporal Essence of Exposure

Exposure was defined as "*an event that occurs when there is contact at a boundary between a human and the environment with a contaminant of a specific concentration for an interval of time*" (National Research Council 1991b). Air pollution is an ever-present phenomenon in which different pollutants interact and create heterogeneous concentration at different microenvironments; human activities also change over the course of a 24-hour period (Steinle, Reis, and Sabel 2013). The spatiotemporal dynamics of both air pollution and human activities lead the exposure to be a continuously changing process.

Numerous epidemiologic and toxicological studies have linked various air pollutants to adverse health outcomes to investigate the association. A conventional means of estimating personal exposure is to use participants' residential locations as spatial surrogates. This is problematic since it ignores the impact of human activities. The exposure estimates result in a significant difference in actual exposure. This difference increases when a person spends a greater proportion of time away from home.

Some studies have found that apart from the home places, other places also have a significant impact on personal exposure. In a meta-analysis, Avery et al. (2010) found that the median within-person residential outdoor–personal $PM_{2.5}$ correlation coefficient is 0.53 with a range of 0.25-0.79. In addition, the standard derivation of the correlation coefficient varies widely. Setton et al. (2008) conducted a simulation study to investigate census tract-specific exposure distributions for people in Metro Vancouver. This study found the time spent at workplaces contributes most to the within-census tract variability. Setton et al. (2011) suggested that ignoring daily mobility patterns could contribute to the null hypothesis in epidemiological studies by 1%, 16%, 30%, and 34% for different groups of people. de Nazelle et al. (2013) found that compared to home-based exposure, accounting for the spatiotemporal activity patterns increased personal exposure to NO_2 by 24%. These studies have provided evidence that that the lack of human time-activity patterns in exposure assessment would eventually bring in a bias to epidemiology research.

This study found that there were two principle methods to collect human time-activity information for exposure assessment in the literature. The first method oftentimes asks participants to carry over a diary to write down the detailed information about human activities. The second method relies on various GPS-enabled devices to automatically collect a string of trajectory points, which can be used to derive time-activity patterns. These two methods are briefly discussed as follows.

2.2.2 Diary-based Methods

Studies on human time-activities were first extensively investigated in the field of sociology whereby the term “time budget” was frequently used to represent the amount of time an individual spends in different activities/places (Klepeis et al. 2001). One way of obtaining time budget information from the participants is to ask the respondents to maintain a diary over a required period. The tool to record the time activities and the complementary information is commonly known as a time-activity diary (TAD) or time-location diary (TLD) (Hazlehurst et al. 2017). Participants are usually instructed to fill in the information on where and how they spent time. Some studies also require participants to detail transportation modes (e.g., on foot, train, bus, car driver, car passenger, etc.) whenever they go on a trip (Dons et al. 2013).

Starting from the 1980s, researchers in environmental health fields started collecting human time-activity information to support exposure assessment studies. Johnson (1984) and Akland et al. (1985) conducted a personal exposure project including 454 participants in Denver and Washington, DC. Over the data collection, each participant was required to carry a personal exposure monitor (PEM) to measure carbon monoxide (CO) and to fill out an activity diary on a 24-h basis. The participants had to fill in all detailed time-activity information on a certain page whenever they changed places or activities. Quackenboss et al. (1986) conducted an exposure study to measure nitrogen dioxides (NO_2) for approximately 350 individuals in Portage, Wisconsin. Participants were asked to record the time they spent in each of the five predefined activity categories. A more thorough review can be found in Ott (1989).

Since 1990, a few large-scale population studies, designed to collect exposure-relevant information on human activity patterns have emerged (Klepeis et al. 2001). Wiley et al. (1991) conducted a diary-based activity study for residents, including both adults and children, in California from 1987 to 1990. Following this, the national human activity pattern survey (NHAPS) took place and was patterned after Wiley's study. NHAPS is the first national-scope exposure-relevant study on human time-activity patterns, which began in September 1992 and ended on October 1994 (Klepeis et al. 2001). In addition to the national coverage, NHAPS study has a refined spectrum of microenvironments. The time spent at microenvironments (e.g., kitchens, restaurants, etc.) with an elevated air pollution level was required to be logged in the diary. Another national survey study of human activity focused on human exposure to soil, took place from 1994 to 1995 (Robinson and Silvers 2000). Unlike the NHAPS study that emphasizes the phenomenon of air quality, this study asked participants to answer some questions about the extent of contact with soil they had experienced. The Canadian Human Activity Pattern Survey (CHAPS) is another large-scale population survey study to collect Canadian activity pattern information for exposure modeling (Leech et al. 1996). Currently, a large cohort study, the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA AIR), is proceeding to evaluate the association between long-term air pollution exposure and the progression of subclinical atherosclerosis and the incidence of cardiovascular disease (Kaufman et al. 2012). This study recruited more than 7,000 participants in six US cities and obtained the typical time-location pattern of each participant via questionnaire. The MESA AIR study requires participants to record the time spent at home, work, in-transit and other locations when they visit.

It is noted that all the survey studies primarily used a diary or questionnaire to acquire the time activity information. A diary can take place in real time. That is, the participants detail all relevant information whenever an activity starts or ends. Alternatively, the diary can be filled out through a subsequent interview in which respondents have to recall all the activities in the past 24 hours (Leech et al. 2002).

The real-time data collection can be intrusive to participants' daily life. A diary can detail more entries of activities to some extent. In contrast, the retrospective methods result in a lower burden to participants but suffer from a recalling issue. Meanwhile, the introspective methods introduce uncertainties to further epidemiological studies.

2.2.3 GPS-enabled Methods

With the advent of GPS and GIS, the tools to collect time-activity patterns have expanded from TAD/TLD to the use of GPS-enabled devices (e.g., smartphones). The time, locations, and movements of whoever carries a GPS are automatically logged. In some cases, with the assistance of the accelerometer sensor, the related equipment can also record the energy expenditure data associated with physical activity (de Nazelle et al. 2013). The spatiotemporally logged points, over a certain period, constitute a string of trajectory. Each point along a trajectory contains information about location, time, and speed or acceleration attributes.

The first study that integrated GPS devices in the environmental exposure assessment field is the Oklahoma Urban Air Toxics Study (OUATS) (Phillips et al. 2001). This study asked each participant to wear a GPS recorder along with a personal monitoring device. The monitoring device collected volatile organic compounds (VOCs) near the breathing zone. Additionally, this study asked each participant to maintain a diary to detail his or her locations and activities. This study found that GPS data could confirm all travel events that were reported in diaries. Moreover, GPS data also detected additional activities that were not recorded in the diaries. As a result, Phillips et al. (2001), suggested that GPS techniques would be a promising tool to track people's time-activity information in exposure assessment related studies. All the select individual-level GPS-enabled exposure assessment studies are summarized and illustrated in Table 2.3.

Table 2.3 A review of GPS-enabled exposure assessment studies

Author & Year	Category	Pollutant	Sample Size	Data Collection
Phillips et al. (2001)	Direct	VOC	NA	GPS & Diary
Lee et al. (2005)	Direct	Noise	1	GPS
Greaves, Issarayangyun, and Liu (2008)	Direct	$PM_{2.5}$	1	GPS & Voice Recorder
(Nethery et al. 2008)	Direct & Indirect	NO & NO_2 & $PM_{2.5}$	62	GPS & Diary
Adams, Riggs, and Volckens (2009)	Direct	Fine PM	1	GPS
Gerharz, Kruger, and Klemm (2009)	Indirect	$PM_{2.5}$	6	GPS & Diary
Morabia et al. (2009)	Direct	$PM_{2.5}$	20	GPS & Diary
Morabia et al. (2010)	Direct	$PM_{2.5}$	21	GPS & Diary
Broich, Gerharz, and Klemm (2011)	Direct	PM_1 & $PM_{2.5}$ & PM_{10}	16	GPS & Video Recorder & Diary
Dons, Panis, et al. (2011)	Direct	BC	16	GPS & Diary
Houston et al. (2011)	Indirect	AADT	47	GPS & Diary
Lonati et al. (2011)	Direct	PM	1	GPS
Buonanno et al. (2012)	Direct	PM	103	GPS & Diary
Wu et al. (2012)	Direct	PB-PAH	28	GPS & Questionnaire
Buonanno et al. (2013)	Direct	UFP & BC	103	GPS & Diary
de Nazelle et al. (2013)	Indirect	NO_2	36	Smartphone & Diary
Dons et al. (2013)	Direct	BC	62	GPS & Diary
Gerharz et al. (2013)	Indirect	PM_{10} & $PM_{2.5}$	10	GPS & Diary & Video Recorder
Houston et al. (2013)	Direct	PB-PAH	24	GPS
Buonanno, Stabile, and Morawska (2014)	Direct	UFP	48	GPS & Diary
Dias and Tchepel (2014)	Indirect	$PM_{2.5}$	5	GPS
Moller et al. (2014)	Direct	UFP	30	GPS
Nethery et al. (2014)	Direct	$PM_{2.5}$	54	GPS
Nyhan, McNabola, and Misstear (2014)	Direct	PM_{10}	60	GPS

Author & Year	Category	Pollutants	Sample Size	Location-wise Acquisition
Tchepel et al. (2014)	Direct	Benzene	10	GPS
Arku et al. (2015)	Direct	$PM_{2.5}$	56	GPS
Beko et al. (2015)	Direct	UFP	59	GPS
Lu and Fang (2015)	Indirect	O_3	1	GPS
Nieuwenhuijsen et al. (2015)	Direct & Indirect	BC	54	Smartphone
Ouidir et al. (2015)	Direct & Indirect	NO_2 & $PM_{2.5}$	40	GPS & Diary
Pilla and Broderick (2015)	Direct & Indirect	PM_{10}	NA	GPS & Diary
Ryan et al. (2015)	Direct	UFP	20	GPS & Diary
Steinle et al. (2015)	Direct	$PM_{2.5}$	17	GPS & Diary & Interview
Su et al. (2015)	Indirect	NOx	1	Smartphone & WIFI
Yoo et al. (2015)	Indirect	$PM_{2.5}$	43	Smartphone & Diary
Adams, Yiannakoulias, and Kanaroglou (2016)	Indirect	$PM_{2.5}$	NA	Simulated Trajectory
Dewulf, Neutens, Lefebvre, et al. (2016)	Indirect	NO_2	180	GPS
Lei et al. (2016)	Direct	$PM_{2.5}$ & BC	51	GPS & Diary
Rabinovitch et al. (2016)	Direct	PM	30	GPS
Sloan, Philipp, et al. (2016)	Direct	$PM_{2.5}$	10	GPS
Sloan, Weber, et al. (2016)	Direct	Elements of PM	21	GPS
Williams and Knibbs (2016)	Direct	BC	1	GPS & Diary
Zhu, Marshall, and Levinson (2016)	Indirect	UFP	144	GPS
Huck et al. (2017)	Direct	NO_2	NA	GPS
Moller et al. (2017)	Direct	UFP	69,175	GPS
Panella et al. (2017)	Direct	UFP & BC	100	Smartphone

(Note: NA: not applicable; VOC: volatile organic compound; PM: particle matter; UFP: ultrafine particle; BC: black carbon; O3: ozone; NO: nitrogen oxides; PB–PAH: particle-bound polycyclic aromatic hydrocarbon; AADT: annual average daily traffic)

Most of the aforementioned studies are grouped within the track of direct methods. Only few studies attempted to develop approaches for estimating personal exposure using GPS data (Gerharz, Kruger, and Klemm 2009; Lu and Fang 2015; Gerharz et al. 2013). For instance, Gerharz, Kruger, and Klemm (2009) applied a dispersion model for outdoor $PM_{2.5}$ and a mass balance model for indoor concentration to model personal exposure with the assistance of GPS tracks and diary data. In addition, considering that the application of GPS is still at an early stage in environmental exposure assessment field, most of the reviewed studies merely use GPS generated data as a secondary data to complement diary data.

The only notable exceptions are the few studies that utilized GPS data as the main source for deriving human time-activity patterns (Adams, Riggs, and Volckens 2009; Wu et al. 2011; Wu et al. 2012; de Nazelle et al. 2013; Dias and Tchepel 2014; Breen et al. 2014).

Adams, Riggs, and Volckens (2009) developed a classification algorithm to apportion GPS trajectory data into a set of microenvironments. This algorithm considered the circular distances to the physical footprints of residential structures. Additionally, this study took advantage of the temperature sensor to distinguish between indoor and outdoor activities. The accuracy of activity classification was reported to be up to 97%. Nevertheless, this study had only one single participant under examination, and that subject's home and workplace location were known beforehand. Therefore, technically the apportionment algorithm was not driven by trajectory data.

Wu et al. (2011) developed two automated approaches to classify GPS trajectory data into a set of predetermined activities. Unlike Adams, Riggs, and Volckens (2009), Wu et al. (2011) did not know participants' home and workplace locations ahead of time. This study systematically evaluated the two proposed approaches, that is, a rule-based classification method, and a random forest classification method. The study suggested that automated classification models would be a promising means to identify human time-activity information in air pollution health studies. In another study, Wu et al. (2012) deployed the proposed automated classification models to process GPS trajectory data for investigating personal exposure to particle-bound polycyclic aromatic hydrocarbon (PB-PAH).

Rather than using GPS instruments, de Nazelle et al. (2013) innovatively used smartphones with a designated application to collect people's trajectories and energy intensities. A map of NO_2 concentration was overlaid with participants' trajectories to evaluate personal exposure. One of the biggest benefits of using smartphones is that, when a GPS signal is absent, assisted GPS technique and WI-FI positioning system (e.g., Skyhook[1]) will be operable to improve the accuracy of location (Glasgow et al. 2016).

These trajectory-based methods provide tremendous opportunities for investigating human time-activity patterns. The trajectory data generated from the use of the GPS-enabled devices will ultimately enhance the understanding of personal exposure to environmental hazardous air pollution and the association between health outcomes and exposure.

[1] https://www.skyhookwireless.com/

2.2.4 Limitations and Critical Needs

It is apparent that the principal means of collecting human time-activity information is reliant on diaries or interviews. For a retrospective interview, one of the biggest problems is the issue of recalling in which participants cannot remember what they did at a specific moment. Similarly, for a diary-based study, participants may just forget to write down the ongoing activities at some point.

Many types of studies have made efforts to pinpoint the deficiency of diaries or questionnaires in collecting human time-activity information. Houston et al. (2011) did a comparison study regarding the performances of diaries and GPS logging. This study found that, overall, nearly 49% of the locations and trips identified by GPS trajectories were ignored by diary logs. Bricka and Bhat (2006) investigated the GPS trajectory dataset of 377 drivers In Kansas City and found that 29% had at least one instance of a trip that was not reported in the self-reported diaries. Elgethun et al. (2007) compared the GPS trajectory data and diary data in Seattle, Washington. Consistent with other studies, this study found about 48% time on diary was misreported. Furthermore, the missing diary logs problem became severe for Spanish-speaking groups.

Besides the recalling issue, a diary or interview is burdensome and intrusive to study participants. de Nazelle et al. (2013) reported that most participants complained about the nuisance of maintaining a travel log. Another problem is that a diary may not reliably reflect a person's short-term activity (Hazlehurst et al. 2017).

To sum up, the diary-based study is limited regarding the accuracy of recall, reliability, and compliance (Wu et al. 2011).

The benefits of using GPS-based techniques are primarily attributed to high and reliable spatiotemporal resolution and minimal burden to participants (Rainham et al. 2010). Compared to a diary-based study, a GPS-enabled study can provide objective logging data regarding a person's time-activity information with fewer burdens. Another promising feature is the pervasiveness of GPS-enabled devices, that is, smartphones, smartwatches, etc. All such devices offer an extraordinary opportunity to collect spatiotemporal data for deriving human time-activity information for exposure assessment purposes. Moreover, smartphones can offer improved locating solutions when not enough satellites are available. These enhanced locating solutions make a longitudinal time-activity-integrated exposure study possible. All the opportunities necessitate a greater focus on GPS-enabled exposure assessment studies. However, as can be seen in Table 2.3, this field is surprisingly under studied. Very few studies systematically discussed how to integrate and analyze trajectory data with various environmental models in the track of direct exposure assessment methods.

To sum up, personal exposure to air pollution is essentially a function of space and time as humans, and air pollution vary over space and time. Recent literature has realized the significance of the variation of the air pollution in estimating personal exposure (Setton et al. 2011). However, very few studies have paid adequate attention to the human time-activity patterns in the environmental exposure assessment discipline. To fix this gap, this study aims to construct trajectory-based approaches for assessing personal exposure to air pollution, as well as to explore how personal exposure varies over space and time.

3. A MODEL FOR ASSESSING PERSONAL EXPOSURE TO AMBIENT AIR POLLUTION USING TRAJECTORY DATA

3.1 Introduction

This chapter presents a new model for assessing personal exposure to ambient air pollution using trajectory data recording human time-activity patterns in geographic space and time.

3.2 Method

3.2.1 A Conceptualization of Trajectories Documenting Human Activities in Geographic Space and Time

The key of trajectory-based exposure assessment is to determine the specific location of a person at any given time, the duration the person spends at that location and the estimated intensity of the air pollutant in question at that location for the specific duration of time associated with the person in question. We follow the established terminology in network analysis in this discussion. A tour refers to a trip starting at location (e.g., a person's home), traveling along paths connecting different locations, stopping at several other locations, and ending at the starting location. Human activities in geographic space and time can be documented by GPS-recorded trajectories consisting of a sequence of points associated with one or more tours on a daily basis.

A tour can be decomposed into different microenvironments (MEs) representing different locations where activities of distinctively different nature are carried out (Duan, 1981, 1982). A microenvironment is either one location or a set of locations in close proximity where activities of the same nature are carried out by an individual. Figure 3.1 illustrates four distinctively different types of MEs associated with a tour. We distinguish the home ME, workplace ME, travel ME, and other MEs in this discussion. MEs associated with home and workplace locations are easy to understand. In some environmental analyses, it is important to distinguish between the home ME and workplace ME. This is why the home and workplace MEs are used in the model. Any path collecting two neighboring MEs is called the "travel" ME. All activities that are covered by these three MEs are collectively called the "other MEs." Without losing generality of the model, travel mode may be either walking, cycling, driving, or taking transit. As will be discussed in the next section, a key step in the development of the model is to classify points on a trajectory associated with a person's activities into these four MEs.

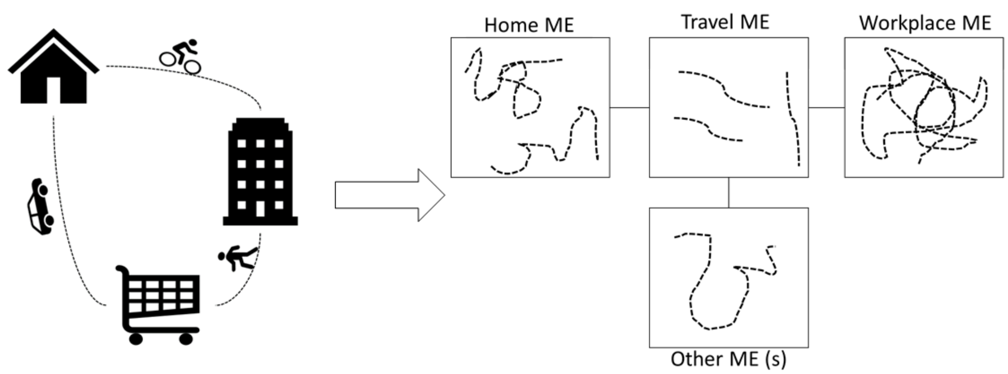

Figure 3.1 Four different microenvironments (MEs) and different modes of travel associated with a person's activity patterns in geographic space and time

3.2.2 Model Development

The model consists of three components: (1) trajectory data processing, (2) representation of air pollution in geographic space and time, (3) exposure assessment. First, the model uses trajectory data processing techniques to classify trajectory data into a set of MEs as defined above and determine the time that person spends at the ME in question. Second, the model represents the geographic distribution of estimated intensities of an air pollutant in question over time using a specific model. The model then links the processed trajectory data and the adequately represented air pollution data in geographic space and time. Third, the model calculates a person's exposure to the pollutant in question using a set of newly developed procedures. The workflow of this innovative model is illustrated in Figure 3.2. The three components are described in detail in the next section.

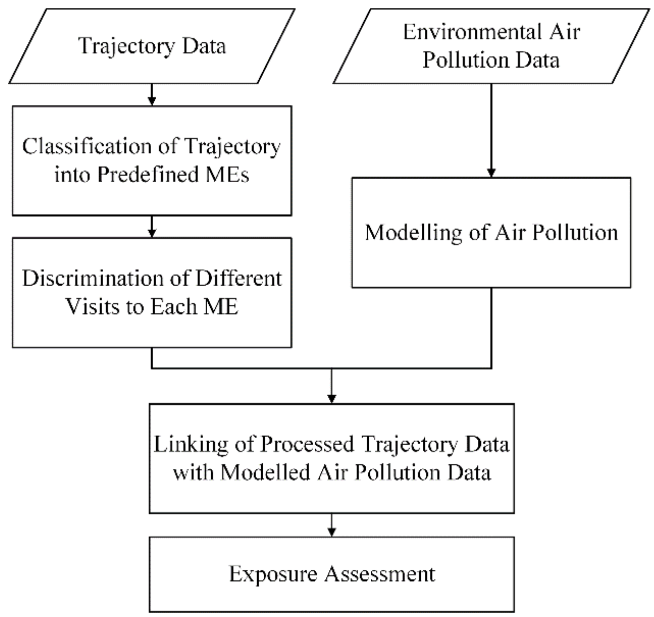

Figure 3.2 The workflow of the proposed trajectory-based exposure assessment model (Note: ME-Microenvironment)

3.2.2.1 Trajectory Data Processing

To classify points of a given trajectory into the four MEs, a series of data processing algorithms were developed and implemented in combination with some predefined rules determining the time windows of typical activities during a working day. For instance, for an average person, the time window in a workplace is typically from 9:00 am to 5:00 pm. A trajectory dataset typically consists of a series of data points. Each data point may contain at least three variables: a pair of geographic coordinates, time stamp, and movement speed. We developed a 3-step procedure to classify the data points.

First, the procedure classifies the points into either stationary or moving points based on a speed profile suggested in the literature (Table 3.1) (Reddy et al. 2010; Dewulf, Neutens, Van Dyck, et al. 2016). All points with speed slower than 1.5 km/h were classified as stationary points. These stationary points are potential candidate points belonging to either the home ME, workplace ME, or other MEs. The rest of the points with speed equal to or faster than 1.5 km/h were considered to be in the travel ME.

Table 3.1 Descriptive information of the speed profile

#	Travel mode	Speed (km/h)
1	Sedentary state	0 ~ 1.5
2	Walking	1.5 ~ 8
3	Cycling	8 ~ 25
4	Driving	25 ~ maximum allowable speed

(Note: In the case study discussed in this chapter, the maximum driving speed is limited to 80 km/h based on the typical speed limit of National Highways in China.)

Second, the classified points in close proximity are grouped into the same ME based on a given distance threshold. For example, a person may move a lot in his or her home ME. These movements result in a lot of stationary data points. All these data points belong to a single home ME. This grouping is achieved through a density-based spatial clustering algorithm DBSCAN (Ester et al. 1996). DBSCAN is a highly efficient spatial clustering algorithm, which can discover clusters of any arbitrary shapes and effectively handle noise points. Another advantage of DBSCAN over other spatial clustering techniques is that it does not require a prior number of clusters. Two important parameters of DBCSAN are the search distance, *MinDist*, and the minimum number of points, *MinPts*, for forming a cluster.

The DBSCAN algorithm consists of four steps. First, DBSCAN checks the neighborhood (defined by the radius of *MinDist*) of each point. If the neighborhood of center point *X* contains more than *MinPts* points, a new cluster is created. The center point *X* is called a core point. Second, DBSCAN iteratively checks whether some of the clusters are density-connected or density-reachable to ensure that all the points that should be in the same cluster are included in that cluster. Any two adjacent clusters sharing one or more points are considered to be density-connected, and any cluster in a sequence of density-connected clusters is defined as density-reachable with respect to any other cluster in this sequence of clusters. Third, the algorithm merges clusters that are density-connected or density-reachable, creating unified clusters with arbitrary shapes. Fourth, the algorithm terminates when no points can be included into a cluster, and there are no connected or reachable clusters.

Whether a stationary data point belongs to the home ME or workplace ME can be determined based on a set of predefined temporal filtering rules (Yuan, Raubal, and Liu 2012). The workplace ME is typically defined as the largest cluster of stationary points recorded during 8:00 am – 5:00 pm on weekdays, and the home ME is typically considered to include the largest cluster of stationary points during other hours on weekdays. All remaining stationary data points that are not categorized into either the home ME or workplace ME are placed into the other MEs. As a result, for each trajectory corresponding to the daily activities of an individual, the two steps described above discover one home ME, one workplace ME, one travel ME, and multiple other MEs when the trajectory is complete in a 24-hour period.

The third step of the procedure is to derive time-related information associated with all points in a single ME in any time window during which the individual in question remains in the same ME. The time-related information is important because we need it to link location data in each ME to the air pollution data in the subsequent exposure assessment procedure. For the convenience of discussion, each stay in an ME without a time lapse is called a visit. An individual may visit a single ME during a day several times. A single visit is defined as a chain of line segments connecting points with sequential time stamps in the same ME. This study implemented a line detection algorithm to identify the chain of line segments associated with each visit. For each ME, the line detection algorithm starts from the first point in the point cluster and keeps adding the next point to the existing chain if the time difference between two consecutive points is less than a predefined time interval. The line detection algorithm continues expanding the chain of line segments by adding a line segment to the chain until no points can be found within the predefined time difference. Next, a new chain is created and initialized by the first point in the remaining point sets. The algorithm terminates when it exhausts all points in a cluster.

3.2.2.2 Representation of Air Pollution in Geographic Space and Time

The intensity of a specific air pollutant in a given location at a particular time can be determined through air pollution modeling. Many air pollution models have been developed in the last few decades. Because the development of air pollution models is not the focus of this study, this topic is not discussed further in this discussion. Once the distribution of the estimated intensity of an air pollutant in geographic space and time is determined, the spatiotemporal variations of the intensities can be represented by a 3D space-time cube model (Fang and Lu 2011) as shown in Figure 3.3. In this 3D cube, the geographic distribution of the estimated intensities of the air pollutant in question at each time step is represented as a single map layer as shown in Figure 3.3. The representation of the intensities in geographic space and time is achieved by a stack of map layers in the cube. An example trajectory documenting the movements of an individual's activity patterns in geographic space and time relative to the 3D cube is shown in Figure 3.3. The linkage of the 3D cube and the trajectory can be achieved using the location and time stamp information obtained during trajectory data processing.

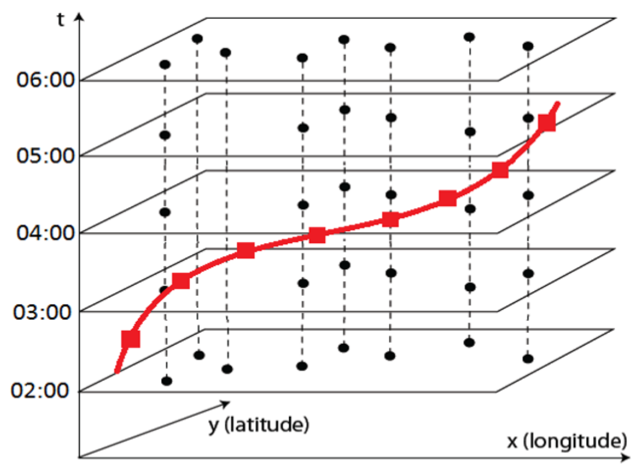

Figure 3.3 Linking trajectory data with estimated air pollution data represented by a 3D space-time cube. The trajectory in red contains a sequence of red square points representing logged points with both location and time stamp information. The black dots mark the locations of monitoring sites where hourly air pollution data are recorded.

3.2.2.3 Exposure Assessment

We are ready to calculate an individual's exposure to a pollutant based on the activity patterns documented by a trajectory. Recall that a person's daily activities may include one or more tours. The travel patterns of a tour or tours recorded by trajectories can be decomposed into a set of MEs. Each ME, in turn, may consist of a number of visits.

This hierarchy of tour-ME-visit related to an individual's spatiotemporal activity patterns can be summarized by four vectors shown in Equation 3.1 below.

$$Personal\ exposure = \{Tour_1, Tour_2, Tour_3, \dots \dots, Tour_q\}$$
$$Tour\ exposure = \{ME_1, ME_2, ME_3 \dots \dots, ME_n\}$$
$$ME\ exposure = \{Visit_1, Visit_2, Visit_3, \dots \dots, Visit_m\}$$
$$Visit\ exposure = \{Point_1, Point_2, Point_3, \dots \dots, Point_w\} \quad (3.1)$$

Based on discussions in the last subsection, each visit consists of a chain of consecutive line segments connecting points recorded sequentially in time. The duration associated with each visit is the time difference between the time associated with the last and first data points of the visit in question. An individual's exposure to the pollutant in question during a visit then can be easily determined by multiplying the average estimated intensity of the pollutant at the locations associated with the visit and the duration of the visit. For each visit, the total exposure (TE) and average hourly exposure (AHE) are distinguished. Formulas for calculating TE and AHE are given in Equation 3.2.

$$Visit\ TE = \sum_{i=1}^{m}(c_{ij} \times \Delta t_{ij}),\ Visit\ AHE = \frac{\sum_{i=1}^{m}(c_{ij} \times \Delta t_{ij})}{\sum_{i=1}^{m}\Delta t_{ij}} \quad (3.2)$$

where c_{ij} is the average intensity of the air pollutant in question at the locations of data points i and j, and Δt_{ij} is the time difference (duration) associated with trajectory points i and j. Once the exposure associated with a visit is determined, it is straightforward to calculate the exposure associated with an ME, a tour, and an individual over a given period of several days through simple summations of the relevant estimated exposure at different levels of the hierarchy.

3.3 Experiments based on Simulated Data

3.3.1 Data

3.3.1.1 Air Pollution Data

This study collected the hourly readings of $PM_{2.5}$ from thirty monitoring sites over a 24-hour period on February 28, 2016, in Beijing, China. There were a total of 750 readings. The geographic distribution of the monitoring sites is illustrated in Figure 3.4. The data were obtained from AQICN.org which is a third-party air quality data source that is published by the World Air Quality Index project. All air quality data available on AQICN.org are compiled from ground air quality monitoring sites operated by different levels of government in the respective countries.

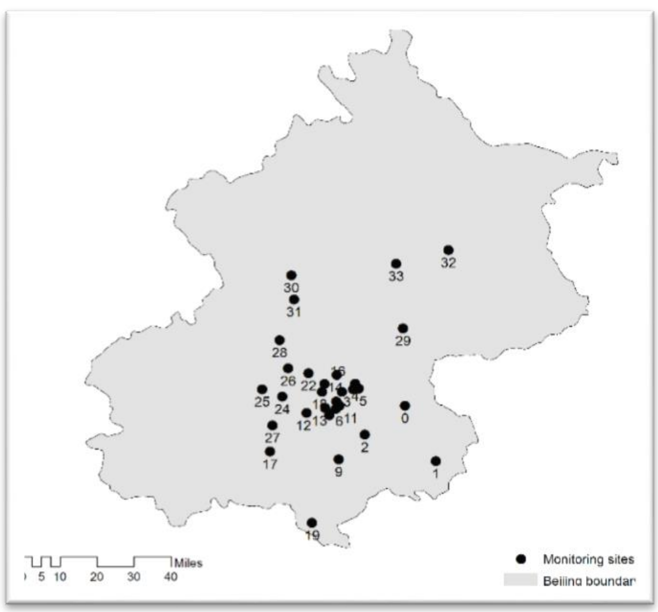

Figure 3.4 Geographical distribution of monitoring sites in the Beijing area

3.3.1.2 Simulated Trajectory Data

Because of the lack of public domain individual-level trajectory data in the study area - Beijing - where large-scale air quality monitoring data were available, this study developed a trajectory simulation platform to generate a set of trajectories mimicking the movement patterns of an adult in geographic space and time on weekdays. The simulated trajectory data have the three attributes mentioned above that a GPS device commonly recorded, that is, location, timestamp, and speed. Therefore, the simulated trajectory data have all the necessary characteristics of real data, and the simulated trajectories were used to test the performance of the model.

It is necessary to make a few assumptions when generating the simulated trajectory data. First, an individual's daily activities are assumed to include only the home ME, workplace ME, travel ME, and other MEs as discussed in Section 2 of this chapter. Second, people are assumed to use the fastest route when traveling between two locations using a randomly assigned travel mode from the options shown in Table 3.1. Third, it is assumed that people can move freely within a confined 50-m circular area centered at a predefined location in a single sedentary ME (i.e., home, workplace and other MEs) (Beko et al. 2015). Fourth, people's activities are constrained by a series of time allocation rules as explained below.

The time allocation to different MEs is assumed to resemble a typical adult's spatiotemporal activity patterns on a weekday: (1) a day has 24 hours, (2) a person stays at a workplace from 9:00 a.m. to 5:00 p.m., (3) the time spent in the travel ME connecting the home ME and the workplace ME is assumed to be less than 4 hours, (4) the time consumed in the travel ME from the workplace ME to other MEs is also less than 4 hours,

(5) the time spent in other MEs ranges from 0.5 to 2 hours. Information about the time allocation is provided in Table 3.2.

Table 3.2 Descriptive information of time allocation in simulation

#	Criteria	Value (hours)
1	Total time	24
2	Time spent at workplace ME	8 (i.e. 9:00 a.m.-5:00 p.m.)
3	Travel time from home to workplace ME	$\{x \mid 0 < x < 4\}$
4	Travel time from workplace to other MEs	$\{x \mid 0 < x < 4\}$
5	Time spent at other MEs	$\{x \mid 0.5 < x < 2\}$

The trajectory dataset was generated based on the road network in Beijing, China, which was obtained from OpenStreetMap (OSM). Given that the network dataset did not have speed limit information, we randomly assigned a speed limit to each road segment. The speed attribute of a roadway was used in the subsequent analysis to distinguish points of travel ME from stationary MEs (i.e., home ME, workplace ME, and other MEs). The data points in the simulated trajectory data were assigned a logged location and time stamp every 5 seconds.

This study followed two procedures described below when generating the simulated trajectory data. The first procedure was to obtain 100 sets of three nodes that met all the predefined time allocation rules. The three nodes were used for the center locations of the home ME, workplace ME, and one other ME when simulating trajectories. The second procedure was to generate a series of trajectory points based on the assumptions, the time allocation rules, and adequate algorithms mentioned below.

In selecting 100 sets of three qualified nodes, first, we randomly selected three nodes from the network dataset as candidate locations for the home ME, workplace ME, and the other ME. Second, we computed the fastest path between any pairs of locations among the three notes using the Dijkstra's algorithm (Dijkstra 1959) on the Beijing network. Third, during each selection, we made sure that the two nodes of the home ME and the workplace ME can be reached within four hours as defined by the rules shown in Table 3.2. Also, the two nodes of the workplace ME and other ME can be reached within four hours. Last, the steps described above were repeated until we obtained 100 sets, each containing three nodes representing different locations.

Once the 100 sets of nodes were determined, we started to generate the point locations on a trajectory. For points in the travel ME, a linear interpolation algorithm was applied to each link on the determined routes. The sedentary points that were within the home ME, workplace ME and other MEs were generated following a random walk algorithm (Marsh and Jones 1988) within the specified 50-m radius of the three selected nodes. The time stamp associated with each point on a trajectory was assigned based on a data logging frequency of one point every 5 seconds, the time allocation rules, and the travel speeds.

The simulation was implemented in PostgreSQL database bundled with PostGIS (version 2.3.0) and pgRouting (version 2.2.0). We simulated the trajectories of activity patterns of 100 hypothetical individuals over a 24-hour period based on the road network data in Beijing. Figure 3.5 shows one of the simulated daily trajectories.

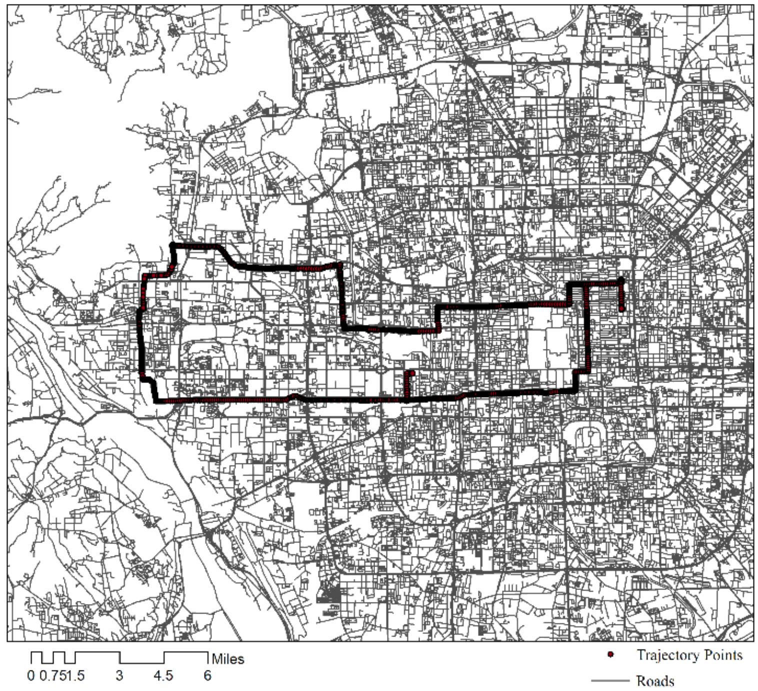

Figure 3.5 An example simulated trajectory documenting an individual's movement patterns in one day.

3.3.2 Experimental Results

3.3.2.1 Results of Trajectory Data Processing

Recall that this part of the analysis of the model is to detect the different MEs and the visits within each ME. Figure 3.6 presents the result of trajectory data processing for the simulated trajectory of Person 1. In processing the data, all data points with speed faster than 1.5 km/h were classified to belong to the travel ME first. Next, all remaining data points were treated as stationary points and became the input points of the clustering analysis algorithm. After different clusters were detected, the home, workplace, and other MEs were determined. The line segment detection algorithm was applied to detect different visits within each ME.

In Figure 3.6, the home ME (Figure 3.6(a)), workplace ME (Figure 3.6(b)), and the other ME (Figure 3.6(c)) are displayed to illustrate the results of clustering analysis. As is shown in the figure, the points in the home ME are rendered in two different colors, indicating that two distinctively different visits were detected in the home ME from the simulated trajectory of Person 1. The travel ME contained three visits and the workplace and other MEs each contained only one visit. For demonstration purposes, we present the results associated with four out of the 100 simulated trajectories below. Information about the MEs, visits in each ME, and the duration associated with each visit in the respective MEs corresponding to the simulated trajectory of each of the four individuals are provided in Table 3.3.

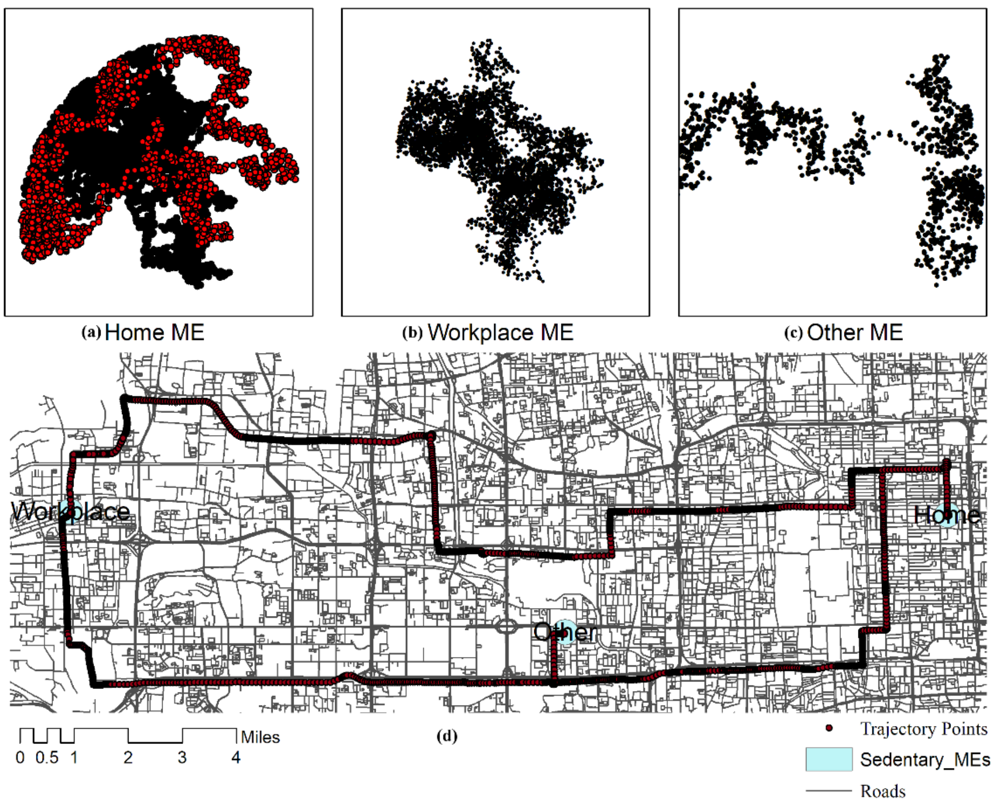

Figure 3.6 The results of trajectory data processing for the trajectory of Person 1.

Table 3.3 Time spent on visits of all MEs by four individuals

	Home ME		Workplace ME	Travel ME			Other ME
Person	H-1	H-2	W-1	T-1	T-2	T-3	O-1
1	7.83	3.78	8.00	1.17	0.72	0.58	1.92
2	7.02	3.43	8.00	1.98	0.95	1.97	0.65
3	6.93	2.42	8.00	2.07	0.53	2.11	1.94
4	7.47	2.11	8.00	1.53	0.79	2.32	1.78

(Unit: hours)

3.3.2.2 Representation of Air Pollution in Geographic Space and Time in Beijing

The key in this part of the analysis is to build a 3D space-time cube to represent the geographic distribution of estimated intensities of $PM_{2.5}$ and link the location points in each trajectory to their corresponding points in the 3D cube. The construction of the 3D cube contains two steps. The first step is to model the spatial variation of $PM_{2.5}$ at any given time based on ground monitoring data. This study used three spatial interpolation methods, the inverse distance weighting (IDW), the radial basis functions (RBF), and the ordinary kriging method, to model the spatial variations of $PM_{2.5}$ at each hour of the day based on the hourly ground monitoring data in Beijing. We used a 10-fold cross-validation method (Stone 1974) to evaluate the accuracy of the estimated intensities at each hour from these three methods and selected the results of the method with the smallest error to represent the spatial variation of $PM_{2.5}$ at that specific hour. The results from this modeling and cross-validation process is a set of twenty-five map layers representing the geographic distribution of $PM_{2.5}$ at each hour during the 24-hour period (Figure 3.7).

Based on the hourly map layers, we then used a simple linear interpolation approach to obtain the estimated intensity for any given time between the two adjacent hours with estimated intensity. Equation 3.3 illustrates the linear interpolation approach.

$$C_{h\sim m\sim s} = C_{h\sim m} = C_h + \frac{m}{60} \times (C_{h+1} - C_h) \qquad (3.3)$$

Where $C_{h\sim m\sim s}$ is the estimated intensity of the air pollutant in question at hour h, minute m, and second s. For ease of computation, this study simply assumes there is no transitional change of intensity within one minute. Theoretically, the overall approach described above generates a seamless 3D space-time cube representing the estimated intensity of an air pollutant in question in every location at any given time.

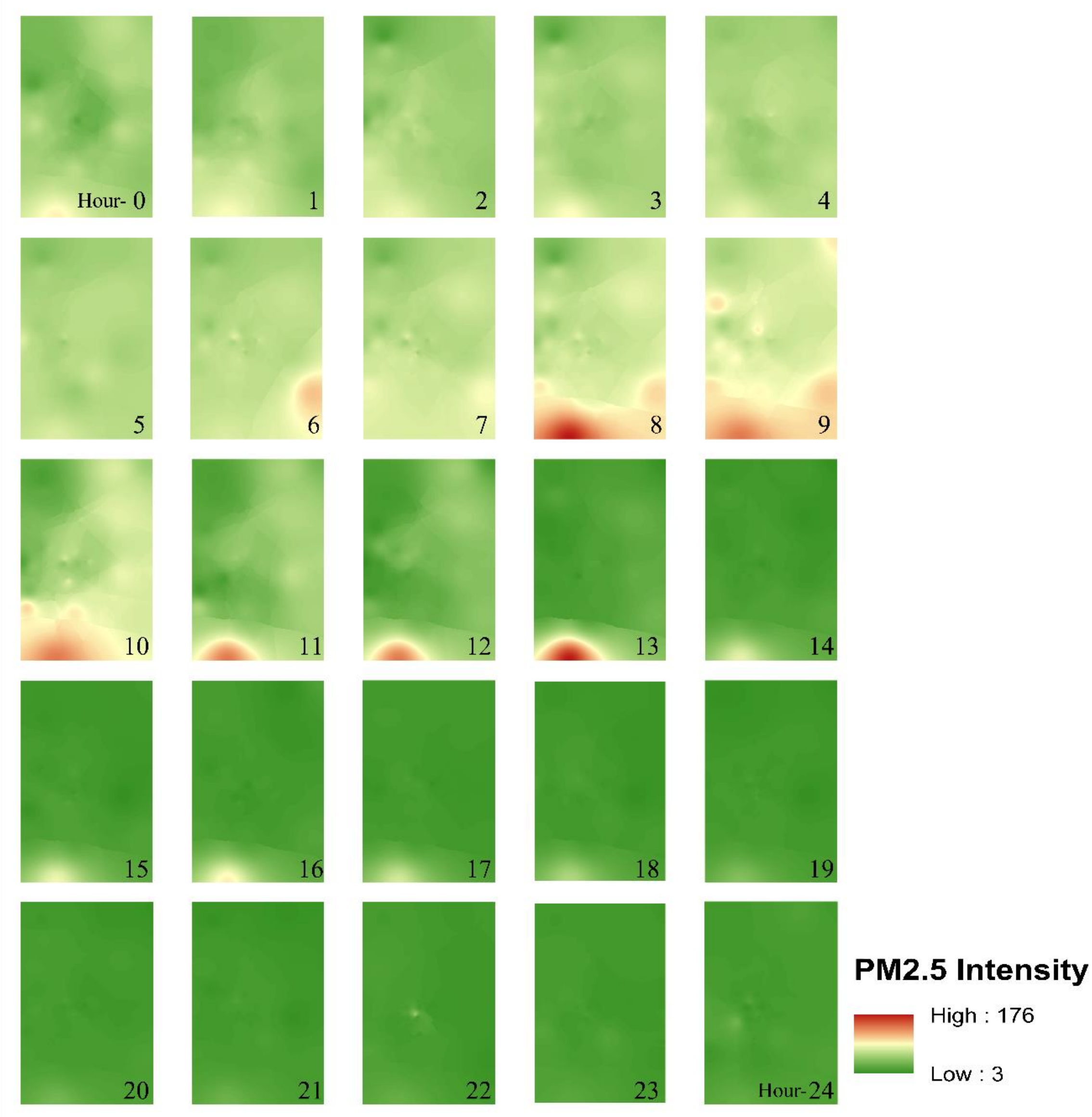

Figure 3.7 Map layers showing the geographic distribution of the estimated intensity of $PM_{2.5}$ at each hour in Beijing over the 24-hour period.

As is shown in Figure 3.7, higher intensities were observed before 9:00 am with a peak value of 82.62 $\mu g/m^3$. From 9:00 to 15:00, the intensity decreased sharply. After 15:00, the intensity stayed at a low level with a mean value of 11.65 $\mu g/m^3$. The density of $PM_{2.5}$ in areas surrounding monitoring site 19 (see Figure 3.5) had much higher intensities than those at other monitoring sites. For instance, at 13:00, the difference of intensities between site 19 and other sites was as high as 162.07 μg/m^3. This significant variation of intensity between areas surrounding site 19 and areas surrounding other sites was evident for most of the day before 19:00. It was expected that a person who traveled through areas around site 19 would have received a much higher exposure.

3.3.2.3 Results of Exposure Assessment

Table 3.4 is a summary of the AHE for the four people. As shown in the table, significant variation exists among the level of exposure associated with different visits within one ME. For all four people, the first visit to the home ME happened in the morning before 9:00 am when the areal $PM_{2.5}$ intensity was higher than 50 μg/m^3 on average. The second visit to the home ME happened after work while the average areal $PM_{2.5}$ intensity dropped to around 10 μg/m^3. The travel ME consisted of three distinct visits. In the experiment, the first travel visit occurred when the four people moved from home MEs to workplace MEs; the second visit took place when people moved from workplace MEs to other MEs; the third visit happened on the way back home MEs. It was evident that even for the same ME, people experienced differing levels of exposure while visiting the ME at a different time. This result confirms the observation that the time of visiting a specific ME is indeed a key factor in affecting exposure due to the variations of the intensity of an air pollutant in geographic space and time.

Table 3.4 Average hourly exposure (AHE) associated with each VISIT

	Home		Workplace	Travel			Other
Person	H-1	H-2	W-1	T-1	T-2	T-3	O-1
1	52.90	9.14	22.43	68.20	8.13	10.53	10.47
2	56.38	11.66	25.59	75.12	11.73	10.94	13.49
3	68.10	13.84	25.43	75.01	12.51	10.68	11.78
4	63.28	12.02	46.52	93.18	20.15	13.80	17.65

(Unit: μg/m^3)

Table 3.5 shows the personal total exposure and the AHE and TE of each ME for each person in a 24-hour period. A few observations can be made from the results shown in Table 3.5. First, it is clear from the results that there are considerable variations in the exposures associated with different MEs. For the simulated trajectories of the four people, the personal total exposures are 740.04, 830.65, 916.12, and 1092.09 μg*h/m^3, respectively, giving a personal average hourly exposure of 30.84 μg/m^3 for Person 1, 34.61 for Person 2, 38.17 for Person 3, and 45.50 μg/m^3 for Person 4 over the 24-hour period.

Table 3.5 Average hourly exposure (AHE) and total exposure (TE) associated with each ME

Person (total)	Home		Workplace		Travel		Other	
	TE	AHE	TE	AHE	TE	AHE	TE	AHE
1 (899.42)	448.9	38.70	179.44	22.43	91.64	37.11	179.44	10.47
2 (830.65)	435.82	41.72	204.73	25.59	181.22	37.04	8.88	13.49
3 (916.13)	505.60	54.04	203.46	25.43	184.22	39.16	22.85	11.78
4 (1092.09)	498.20	51.99	372.21	46.52	190.17	41.07	31.51	17.65

(Unit: TE - μg*h/m^3 ; AHE –μg/m^3)

To gain some insights about the variation associated with each ME and their contributions to the overall exposure of each individual, the mean, median, first quartile, third quartile, minimum and maximum values of exposure associate with each ME for the four individuals are illustrated in Figure 3.8. For persons 1, 2 and 3, the exposure associated with the home ME had the highest mean and median, followed by traffic ME, workplace ME, and other MEs. The exception is Person 4 whose workplace ME is in close proximity to the site 19 which had much higher intensities than other sites did.

These results are echoed by the results of 100 simulated trajectories with randomly selected home, workplace and other locations. Table 3.6 and Figure 3.9 provide some summary statistics of the estimated exposures associated with these 100 trajectories. The exposure associated with the Home ME in the 100 trajectories had the highest mean and median values, followed by the Workplace ME, and then the Travel ME.

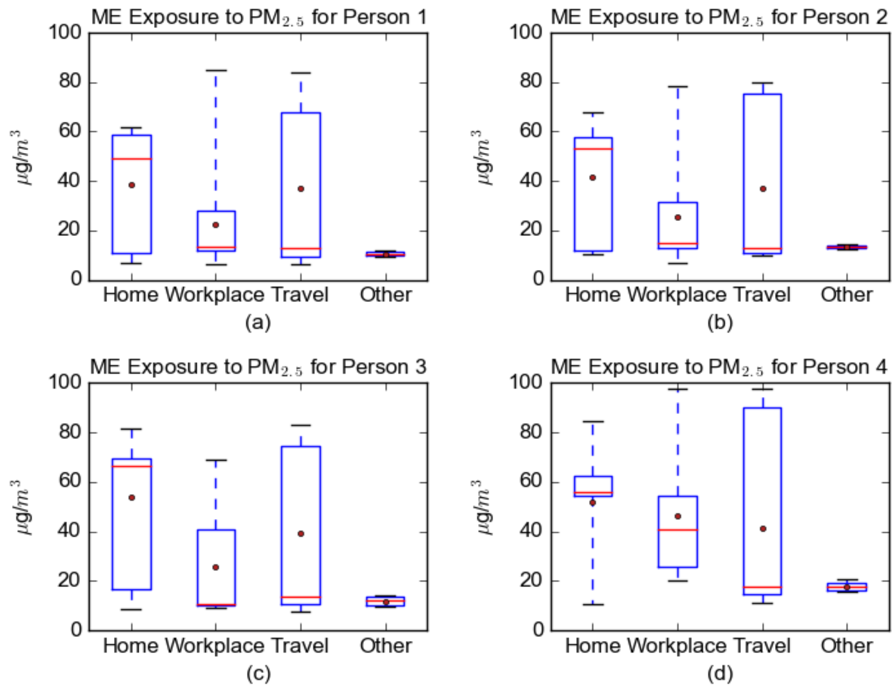

Figure 3.8 Descriptive statistics of personal exposure for 4 hypothetical people in Beijing during a day (Notes: the central rectangle spans the first quartile and the third quartile; the red line shows the median; the red dot is the average for that exposure; the top and bottom bar show the maximum and minimum exposures.)

Table 3.6 Summary statistic of estimated exposures in different MEs associated with 100 simulated trajectories

	TE over 24 hours				**AHE over 24 hours**			
	Home ME	**Workplace ME**	**Travel ME**	**Other ME**	**Home ME**	**Workplace ME**	**Travel ME**	**Other ME**
Max	534.60	372.21	347.11	31.51	57.52	46.52	44.67	17.65
Min	205.92	137.54	86.56	4.32	30.99	17.19	19.98	6.40
Median	423.78	210.36	162.64	12.85	42.66	26.29	35.67	10.30
Mean	411.78	215.75	178.94	12.66	43.01	26.96	35.27	10.28
SD	75.07	32.19	70.28	5.12	4.37	4.02	3.84	1.74

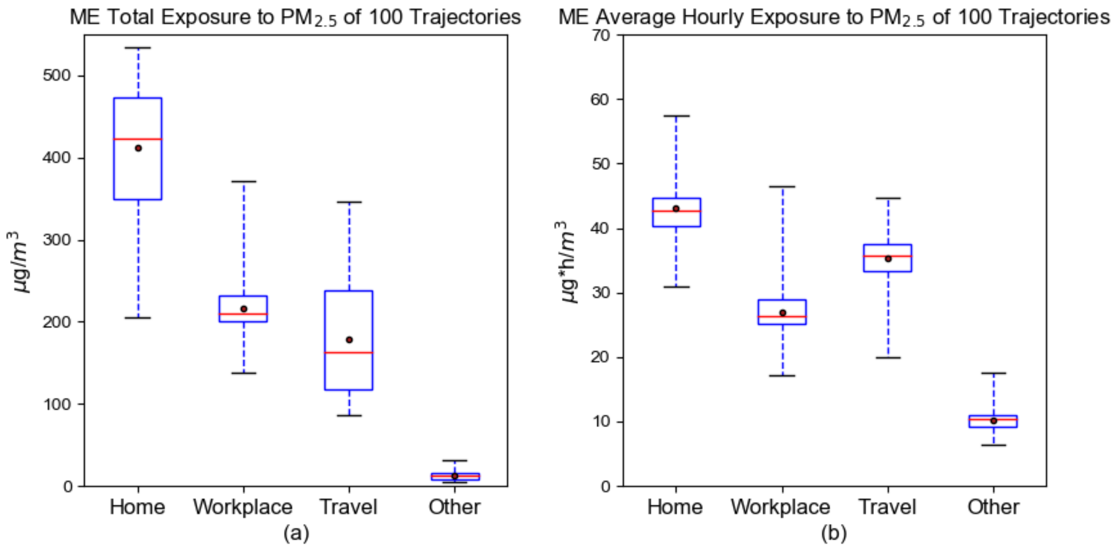

Figure 3.9 Summary statistic of estimated exposures in different MEs associated with 100 simulated trajectories in Beijing over a 24-hour period. (Notes: the central rectangle spans the first quartile and the third quartile; the red line shows the median; the red dot is the average for that exposure; the top and bottom bar show the maximum and minimum exposures.)

The home ME was found to have the highest TE and HAE out of these 100 simulated trajectories for all descriptive statistics. This peak can be explained by the high intensity in the morning hours, before 9 a.m., when hypothetical people were programmed to stay at home ME. Another factor controlling the total exposure is the duration. In our experiment, it was assumed that hypothetical people spend a majority of time staying at home. Therefore, the total exposure of the home ME appears higher than the workplace, travel, and other MEs.

The workplace ME is the second highest boxplot in Figure 3.9 (a), while it becomes slightly lower than Travel ME in Figure 3.9 (b). This discrepancy is caused by the differing duration among different MEs. We can tell that, in this experiment, people have a tendency of staying longer in the workplace ME than in travel ME, thus leading to a lower average hourly exposure in workplace ME. The other ME is the opposite of the home ME in terms of intensity of air pollutant in question and the duration. The home ME has a high level of exposure and long duration, whereas the other ME has a situation whereby a low level of exposure and a short duration appear, which makes the other ME have the lowest total exposure and average hourly exposure.

3.3.2.4 Difference between Estimated Exposures based on Home Locations Only and Locations of All Activities

As discussed in the literature review chapter, one long-standing problem in environmental exposure assessment is that most studies reported in the literature used people's residence locations as the sole place for exposure assessment. The limitation of that approach is obvious because exposure in other locations related to a person's activities is not accounted for. The trade-off, of course, is that home-based exposure assessment is much simpler and less costly to implement. The trajectory-based exposure assessment method presented in this chapter is far more complicated and costlier to implement. Therefore, it is important to understand weather the differences between home-based exposures and trajectory-based exposures are indeed significantly different.

To address this concern, we simulated 4,000 trajectories using the simulation platform described in this chapter first. We estimated the trajectory-based exposure associated with each of these 4,000 trajectories and then calculated the home-based exposures associated with the Home MEs of the 4,000 trajectories. The average hourly exposure over a 24-hour period was calculated in both cases when calculating the home-based exposures; we used the average estimated hourly intensity of $PM_{2.5}$ in the center of each Home ME associated with each of the 4,000 trajectories.

We compiled Dataset 1 consisting of the 4,000 trajectory-based exposures and the 4,000 home-based exposures. We also generated two additional datasets, Dataset 2 and 3. Dataset 2 contained 100 trajectories selected from Dataset 1 such that the difference between the home-based exposures and trajectory-based exposures are the smallest. Dataset 3 is composed of the 100 trajectories with the largest differences.

Wilcoxon Signed-rank Test was used on the three datasets. The results of the tests are summarized in Table 3.7. The results in Table 3.7 show that only the differences in the exposures from the two types of the method in Dataset 3 are statistically significant with a Z score of -2.060 and a *p*-value of 0.04. The differences in Datasets 1 and 2 are not statistically significant. These mixed results suggest that additional research is needed to verify the results.

Table 3.7 Wilcoxon Signed-rank test of three trajectory datasets

Dataset	Trajectory-based AHE[a] Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)	Home-based AHE[a], Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)	Wilcoxon Signed-rank test z score (*p*-value)
Trajectories (N = 4,000)	33.92/2.62 (22.90-72.94) 33.81 (32.72/35.03)	33.99/4.98 (18.37-85.20) 34.17 (32.40/35.64)	-0.36 (0.74)
Trajectories with smallest differences (N = 100)	33.83/1.20 (20.15-38.40) 33.61 (33.20/34.24)	33.83/1.19 (30.20-38.37) 33.67 (33.19/34.24)	-1.15 (0.25)
Trajectories with largest differences (N = 100)	36.52/10.07 (25.74-72.94) 36.15 (28.87/40.53)	41.41/21.48 (18.37-85.20) 42.85 (20.52/49.82)	-2.06 (0.04)

[a]AHE: the average hourly exposure over the period of 24 hours, $\mu g/m^3$

3.4 Experiments based on Empirical Data

3.4.1 Pre-processing of Trajectory Data

This study used the Geolife GPS trajectory dataset (Zheng et al. 2011) to test the exposure assessment model. The Geolife GPS trajectory dataset was collected by the Microsoft Research Asia Geolife project, which covers 178 users over a period of four years from April 2007 to October 2011. Most (91%) of the trajectories were logged at either a time interval of 1~5 seconds or a distance interval of every 5~10 meters.

After a preliminary examination of the Geolife dataset, we selected a sample of 100 trajectories in the Beijing area that contain sufficient data points. Each selected trajectory includes more than 10,000 trajectory points. It is noted some trajectories span more than one day. To ensure all the trajectories used in the analysis are within a 24-hour period (hours 1-24), this study first determined the day in each trajectory that contained the most data points and used those data points to represent that trajectory within the 24-hour time window. Some characteristics of these 100 trajectories are summarized in Table 3.8.

Table 3.8 Characteristics of 100 trajectories

Geolife Dataset	Number of Data Points Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)	Cumulative Time (hours) Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)
Trajectories (N = 100)	9192/3739 (2222-19739) 9456 (6706/11104)	5.88/3.05 (0.81-15.44) 5.66 (3.95/7.98)

3.4.2 Model Implementation Process

Remember that the model consists of three core components, that is, trajectory data processing, representation of air pollution in geographic space and time, and exposure assessment (Figure 3.2). Some challenging issues of working with empirical data are that the logged trajectories often have missing data points in some time intervals and a person may visit an ME multiple times at any time windows during a day.

We added two additional steps in the workflow of the model to enhance the model to process empirical data. In the first added step, all trajectory data points that are logged out of the predefined time windows of home ME but fall within in the enclosed convex boundary of home ME points are also classified as data points in the Home ME. This step helps find the home visits that take place out of the typical time windows for a person to be home. The second added step detects whether the convex boundary of data points in the home ME overlaps with the boundary of data points in the workplace ME. If the workplace ME overlaps the home ME, then the workplace ME labels are removed from the trajectory. The second step is added to ease the activity misclassification issues for home dwellers who mainly stay at home during typical work hours.

It is important to have a procedure in the model to differentiate between different visits within an ME when processing the trajectory data. Recall that a segment is defined as a stay in an ME without any time lapse in Section 3.2.2.1. This study follows this definition when working with empirical data. The time lapse, however, was extended to a longer period to overcome the problem of missing data for a short time interval. A segment is redefined as a series of consecutive data points in an ME without a time lapse of one minute or more in that ME.

Figure 3.10 shows the trajectory data processing results for one of the 100 trajectories. One biggest difference, when compared to the results of simulated data (Figure 3.6), is the number of segments. In the simulated dataset, the home and workplace MEs have only two and one segments, respectively. However, due to the data missing issues and the complexity of human time-activity patterns, an individual may visit an ME multiple times during a day, which results in multiple segments in an ME. For instance, this trajectory shown in Figure 3.10 contains thirty-nine segments in the home ME and seven segments in the workplace ME.

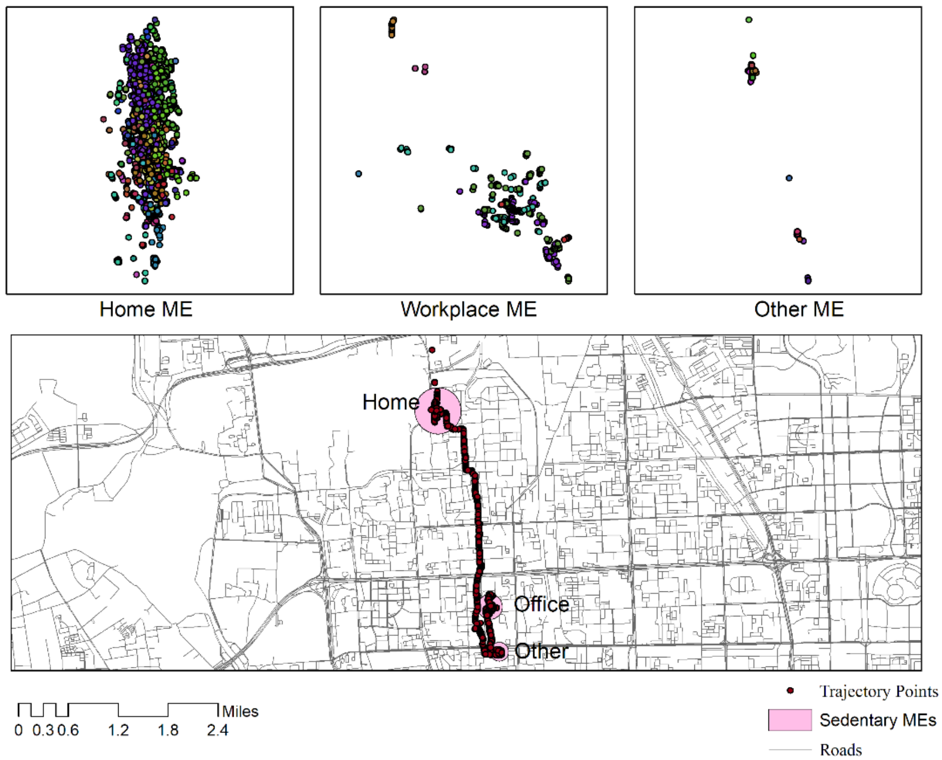

Figure 3.10 The results of trajectory data processing of one empirical trajectory

3.4.3 Results of Exposure Assessment on Empirical Data

Table 3.9 presents the summary of the AHE of all segments of each ME for four randomly selected trajectories. This study reports the descriptive statistics of segments in an ME if the ME has more than three segments. As shown in Table 3.9, the segments in different MEs feature differing characteristics regarding AHE. For instance, the standard deviation of segment AHEs within an ME is as small as 0.1 μg/m^3 (for Person 3) whereas it is up to 23.85 μg/m^3 for Person 2. Regarding the extent of personal exposure, the segment AHEs range from 63.41 to 66.40 μg/m^3 for Person 4 in workplace ME while Person 1 has a wide range of 58.93 μg/m^3 (from 9.22 to 68.15 μg/m^3). The heterogeneity of segment AHE in an ME suggests that people experience differing levels of exposure at different visits, even within the same ME, which is consistent with the findings of Table 3.4.

Table 3.9 Average hourly exposure (AHE) associated with all VISITs

	Home Mean/SD (minimum-maximum) Median (25th/75th percentiles)	Workplace Mean/SD (minimum-maximum) Median (25th/75th percentiles)	Travel Mean/SD (minimum-maximum) Median (25th/75th percentiles)	Other Mean/SD (minimum-maximum) Median (25th/75th percentiles)
1	29.81/2.84 (10.00-55.94) 12.05 (11.19/53.93)	40.32/3.03 (9.22-68.15) 51.21 (10.81/62.44)	31.82/2.24 (8.95-68.15) 12.76 (10.68/54.67)	32.70/3.88 (8.92-65.79) 32.82 (10.50/52.86)
2	57.02	51.83/7.75 (11.14-66.66) 64.65 (53.36/64.99)	44.68/2.81 (9.58-71.68) 50.17 (38.96/55.63)	40.95/23.85 (9.37-65.70) 49.60 (11.26/64.19)
3	10.57/0.10 (9.10-12.84) 10.52 (9.87/11.02)	62.81	23.61/2.18 (9.18-72.04) 10.97 (10.06/36.95)	36.79/2.87 (9.23-65.51) 39.59 (11.90/62.22)
4	67.79, 68.41	65.30/0.95 (63.41-66.40) 66.09 (64.75/66.24)	59.91/1.15 (48.06-68.12) 58.88 (56.66/65.32)	59.15/1.31 (48.02-68.07) 58.75 (53.70/65.06)

(Unit: μg/m^3) (Notes: If there are less than three segments in an ME, we choose to report the AHEs of all segments)

Table 3.10 shows personal total exposure and the AHE and TE of each ME for these four persons. It is noted that TE may not reflect the true scenarios of personal cumulative exposure due to the incompleteness issue. It is, therefore, more reasonable to analyze exposure profiles via AHE. A few observations can be made based on the results shown in Table 3.10. First, the AHEs are not uniform throughout all MEs, meaning that exposure differences exist between MEs. Second, there is not a consistent pattern suggesting that the

AHE estimates associated with one particular ME are higher/lower than those associated with other MEs.

Table 3.10 Average hourly exposure (AHE) and total exposure (TE) associated with each ME

	Home		Workplace		Travel		Other	
Person (total)	TE	AHE	TE	AHE	TE	AHE	TE	AHE
1 (634.91)	90.10	35.82	160.74	44.02	371.69	35.62	12.38	18.14
2 (488.75)	105.81	57.02	131.32	36.87	67.80	36.53	183.82	47.79
3 (395.62)	42.54	10.37	10.90	62.81	280.81	27.75	61.37	36.88
4 (287.63)	13.58	67.80	13.23	65.69	139.22	59.84	121.60	59.40

(Unit: TE - $\mu g*h/m^3$; AHE –$\mu g/m^3$)

Figure 3.11 reports the mean, median, first quartile, third quartile, minimum and maximum values of the estimated exposures associated with each ME for the four trajectories.

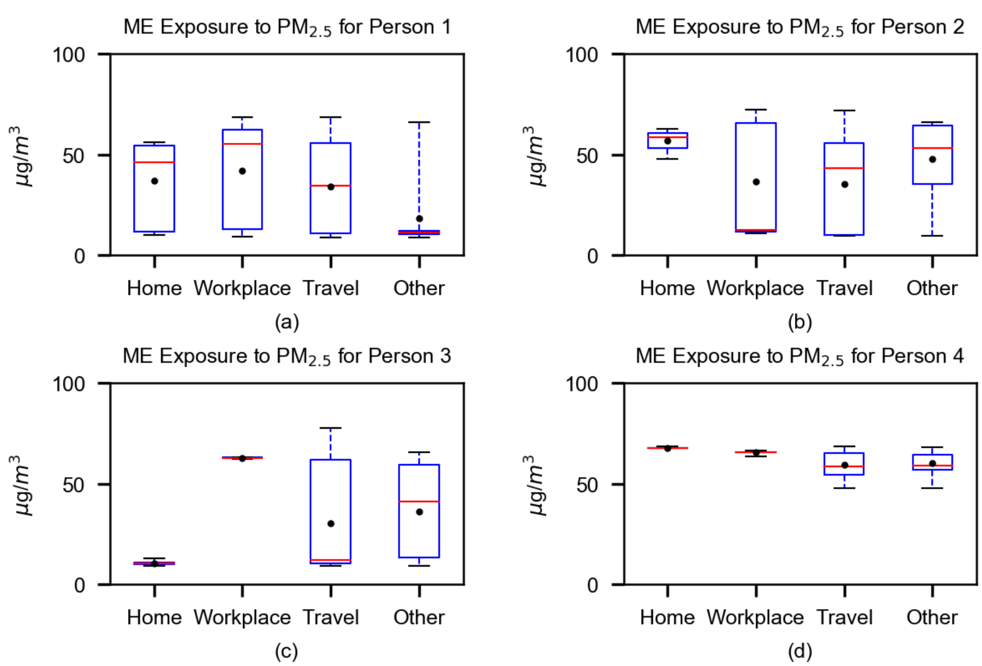

Figure 3.11 Descriptive statistics of personal exposure for four people in Beijing during a day (Notes: the central rectangle spans the first quartile and the third quartile; the red line shows the median; the dot is the average for that exposure; the top and bottom bar show the maximum and minimum exposures.)

To gain additional insights about the variations of the estimated exposures associated with different MEs among these 100 trajectories, Table 3.11 and Figure 3.12 show the descriptive statistics of the exposures. Speaking of the distribution of values of TE in different MEs, there is no consistent pattern observed. The travel ME is found to have the highest TE for all statistics. TE is the product of $PM_{2.5}$ intensity and duration. As mentioned earlier, the empirical dataset has missing data points that happened mainly in sedentary MEs (i.e., home, workplace and other MEs). As a result, most trajectories have more logged points in travel ME than in the home, workplace and other MEs. In terms of the distribution of AHE, different MEs tend to have largely similar patterns.

One of the findings from simulated data suggested that home ME have the highest TE and HAE. However, this finding does not hold in the empirical data.

Table 3.11 Summary statistic of estimated exposures in different MEs associated with 100 trajectories

	TE over 24 hours				**AHE over 24 hours**			
	Home ME	**Workplace ME**	**Travel ME**	**Other ME**	**Home ME**	**Workplace ME**	**Travel ME**	**Other ME**
Max	341.07	185.71	521.64	255.06	69.35	72.18	71.42	71.57
Min	2.62	2.68	22.92	0.34	7.41	10.14	9.66	8.75
Median	48.92	16.29	158.80	17.62	16.79	37.67	37.31	46.76
Mean	68.74	39.96	169.83	38.52	29.00	40.31	37.33	41.18
SD	62.25	53.21	106.50	51.13	21.10	20.44	17.56	18.08

(Unit: TE - $\mu g*h/m^3$; AHE –$\mu g/m^3$)

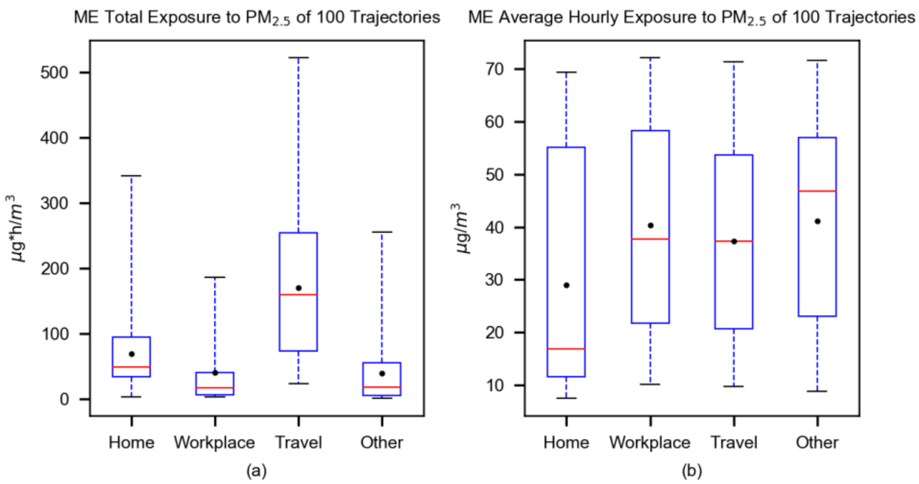

Figure 3.12 Summary information of estimated exposures in different MEs based on an analysis of 100 trajectories in Beijing over a 24-hour period. (Notes: the central rectangle spans the first quartile and the third quartile; the red line shows the median; the dot is the average for that exposure; the top and bottom bar show the maximum and minimum exposures.)

3.4.4 Difference between Estimated Exposures based on Home Locations Only and Locations of All Activities on Empirical data

This study also replicated the experiment we did in Section 3.5 which explored the statistical difference between estimated exposures based on home locations only and locations of all activities. The same statistical test, Wilcoxon Signed-rank Test, was utilized for seeking the difference, but the data were changed to the empirical dataset. This study selected thirty-one trajectories that have classified home ME points, and have at least ten hours recorded on trajectories. The selection is to ensure that the home location can be inferred from the trajectory and have sufficient points.

When calculating the trajectory-based average hourly exposure for a person, this study used the TWAE of $PM_{2.5}$ of the trajectory at a person level. The home-based exposure of a person is calculated by using the average of hourly intensities of $PM_{2.5}$ at the center of home ME. It is noted that due to the incompleteness of trajectory, this study only counts the hours recorded in the trajectory instead of 24 hours when calculating a person's home-based exposure.

The results of the Wilcoxon Signed-rank Test are summarized in Table 3.12. The findings show that a statistically significant difference, with a Z score of -4.00 and a *p*-value of 0.000, was found between two exposure measures.

Table 3.12 Wilcoxon Signed-rank test of three trajectory datasets

Dataset	Trajectory-based AHE[a] Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)	Home-based AHE[a], Mean/SD (minimum-maximum) Median ($25^{th}/75^{th}$ percentiles)	Wilcoxon Signed-rank test z score (*p*-value)
Trajectories (N = 31)	26.86/10.45 (10.40-51.57) 24.82 (17.79/34.02)	29.67/9.00 (11.86-51.69) 30.04(22.86/35.69) of	-4.00 (0.000)

[a]AHE: the average hourly exposure over the period of 24 hours, μg/m^3

3.5 Conclusion and Discussions

This study developed a new model for assessing personal exposure to air pollution based on trajectory data. When the geographic distribution of an air pollutant over a given period of time is known, and when the trajectories of an individual's movement patterns are recorded with both the location and time stamp attributes, this new model can be used to fully assess this individual's exposure to a certain air pollutant. This model, therefore, contributes to the literature of environmental exposure assessment.

The spatiotemporal patterns of human activities in geographic space and time can be conceptualized as a hierarchy that consists of a series of tours starting from a person's home, traveling to other locations such as the workplace and places for other purposes such as recreation, staying at these locations for some time, and coming back to this person's home. A tour can be decomposed into a number of microenvironments (MEs). Typical MEs include the home ME, the workplace ME, the travel ME, and other MEs. The travel

ME refers to travel from one location to another location. The home ME, workplace ME, and other MEs are usually places or locations where a person stays for a few hours or less and moves slowly within an ME. A person may stay at an ME for some time, leave it for a while, and come back and stay in that ME for another while. Each stay of this nature is called a visit to the ME. Therefore, a person may visit an ME a number of times during a day of 24 hours. In other words, an ME may contain one or more than one visit. It is worth noting that it is important and necessary to distinguish the home ME, workplace ME, and travel ME in environmental exposure assessment because many research projects have been and will be designed to test hypotheses aiming to verify how environmental conditions at home, in workplaces, and exposure to traffic emissions would affect human health.

The contribution of this study was the development of a computational procedure that determines the MEs and visits in a tour documented by a trajectory consisting of a sequence of points with locations and time stamps. After each visit in an ME is determined, the exposure to the air pollutant in question associated with that visit then can be calculated. In turn, the exposure for each ME and tour then can be easily computed. For any given period of time, an individual's exposure to the air pollutant for any given period of time can be assessed as long as we have access to the trajectory data documenting the movement and activity patterns in geographic space and time for that period of time.

An experiment was carried out to demonstrate the computational power of the model. The experiment used ground monitoring data of $PM_{2.5}$ in Beijing, China and simulated trajectories of the activity patterns of 100 people during a 24-hour period in Beijing. In simulating the trajectories, this study used randomly assigned home and workplace locations on the Beijing road network with randomly assigned speed limit to each link. The results from the experiment demonstrate that the model works well. Another experiment took place using a real GPS trajectory dataset data to test the model and refine the algorithms used to process the trajectory data.

An additional experiment was conducted to test whether the estimated exposure based on home locations only is statistically significantly different from the exposure considering the locations of all activities. We simulated 4,000 trajectories and created three datasets from the 4,000 trajectories. A Wilcoxon Signed-rank test was used to analyze each of the three datasets. The results of the analysis suggest that only the difference is statistically significant when we selected the 100 trajectories where the differences between home-based exposure and trajectory-based exposure are the largest among the 4,000 trajectories. Moreover, the same test was applied to an empirical dataset containing thirty-one distinctive trajectories. The results from the experiment on the empirical dataset showed a significant difference between the exposure estimates from two exposure assessment methods.

4. QUANTIFYING PERSONAL EXPOSURE TO TRAFFIC EMISSION WHEN HUMAN TRAJECTORY AND HOURLY TRAFFIC VARIATION ARE CONSIDERED

4.1 Introduction

Air pollution has been among the biggest environmental risks to human health. In 2012, 11.11% of deaths worldwide were attributable to air pollution-related conditions (World Health Organization 2016). Numerous epidemiological studies have suggested that air pollution is a contributing factor in causing morbidity, mortality and various health problems such as respiratory symptoms (Zhang et al. 2015), neurological disorders (Heusinkveld et al. 2016; Lee et al. 2017), and cardiovascular diseases (Luben et al. 2017; Fiordelisi et al. 2017).

People come in contact with different air pollutants in a large variety of places (e.g., home, workplace, roadways, among other sources) during the course of their daily activities. Most people in developed countries live near roadways and spend a significant amount of time driving on roadways. Moreover, the concentration of traffic-related air pollutants (e.g., nitrogen dioxide (NO_2), particle matter (PM_*)) on roadways is disproportionately high compared to that of other places. Traffic emissions from vehicles on highways accounted for 47%, 33%, and 20% of total CO, NO_x, and VOC emissions (Health Effects Institute 2010). The percentage of contribution is even greater in metropolitan areas (Health Effects Institute 2010). Therefore, exposure to traffic emissions constitutes a significant portion of personal total exposure to air pollutants. To more completely assess personal exposure to air pollutants, it is important to quantify personal exposure to traffic emissions.

The most commonly used approach to assess an individual's exposure to air pollutants from traffic emissions is to employ a stationary geographic location (e.g., a person's residence location) and use the estimated exposure to traffic emissions at that location as a proxy to represent the exposure of the individual in question. Other approaches used census tracts and zipcode polygons as the area units as the locations. These approaches have obvious limitations because the dynamics of both human activities and the variations of traffic emissions in geographic space and time are not fully accounted for in these exposure assessment methods (Kwan 2004; Elgethun et al. 2007)

To overcome these limitations, this study develops an innovative approach to quantifying personal exposure to traffic emissions based on trajectory data documenting the movement patterns of individuals in geographic space and time and estimated hourly traffic emissions based on a newly developed model.

The remainder of this paper is structured as follows. Section 2 presents the study area, traffic count data, and simulated trajectory data. Section 3 describes the development of a model that can be used to estimate hourly traffic emissions based on traffic count data and procedures that can be used to assess personal exposure based on the estimated hourly traffic emissions and trajectory data. Section 4 analyzes the results of hourly traffic emission model, personal exposure estimates, and the disagreement between two traffic emission models. The final section concludes the chapter and discusses topics for future studies.

4.2 Study Area and Data

4.2.1 Study Area

Minneapolis-Saint Paul was chosen for this study because the Minnesota Department of Transportation (MnDOT) had good traffic count data in the city. The twin-city is a metropolitan area in east-central Minnesota. It is comprised of seven counties. It is the 16th largest urban area in the United States according to the 2010 census, with a population of 2,650,890. Figure 4.1 shows the geographic location of the city in the United States and road networks covering the area.

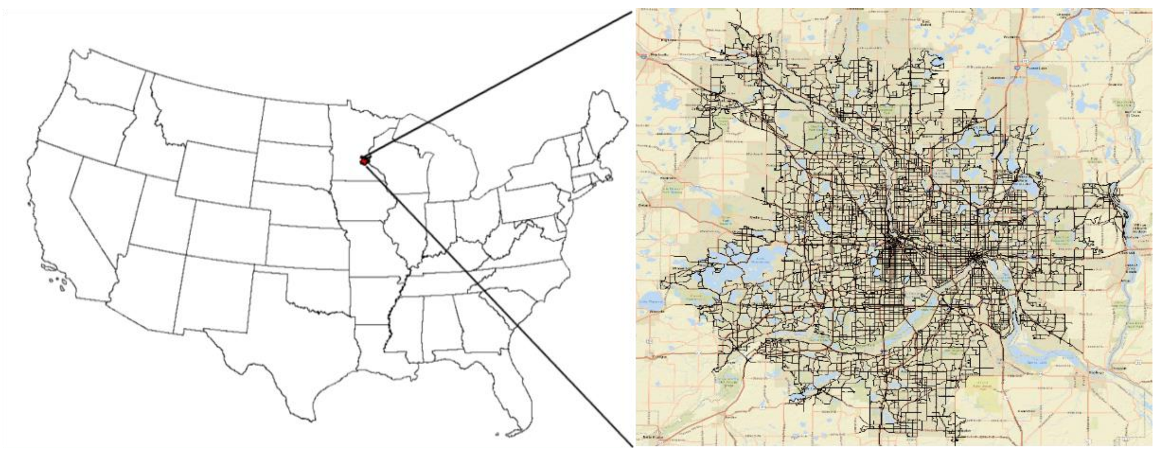

Figure 4.1 Study area: Minneapolis-Saint Paul

4.2.2 Traffic Count Data

This study used three types of traffic count datasets to estimate the annual hourly traffic emissions. These three datasets are annual average daily traffic (AADT) data, annual average daily traffic of heavy commercial cars data (HCAADT), and hourly continuous traffic count data in 2016. The AADT is a measure of average daily traffic volume of a roadway calculated by dividing the total annual volume of that roadway by 365 days. The HCAADT is the average daily count of heavy commercial vehicles, which is calculated similarly. The continuous traffic count data were collected from a campaign of Weigh-in-

Motion (WIM), and Automatic Traffic Recorder (ATR) sites located on or next to roadways. There were a total of 35 recorder stations located in the study area. Figure 4.2 illustrates the geographic distribution of traffic recorder stations, as well as the roadways.

All three traffic count datasets were obtained from MnDOT. The AADT and HCAADT datasets were in a GIS shapefile format, and hourly continuous traffic count data were organized in a set of text files. When preprocessing the AADT and HCAAT datasets, if there was a record missing for any roadway segment, the records in the preceding year was used to replace the missing value. For continuous traffic data, this study developed a script to extract the traffic counts by hours across all workdays in 2016 for each traffic record station.

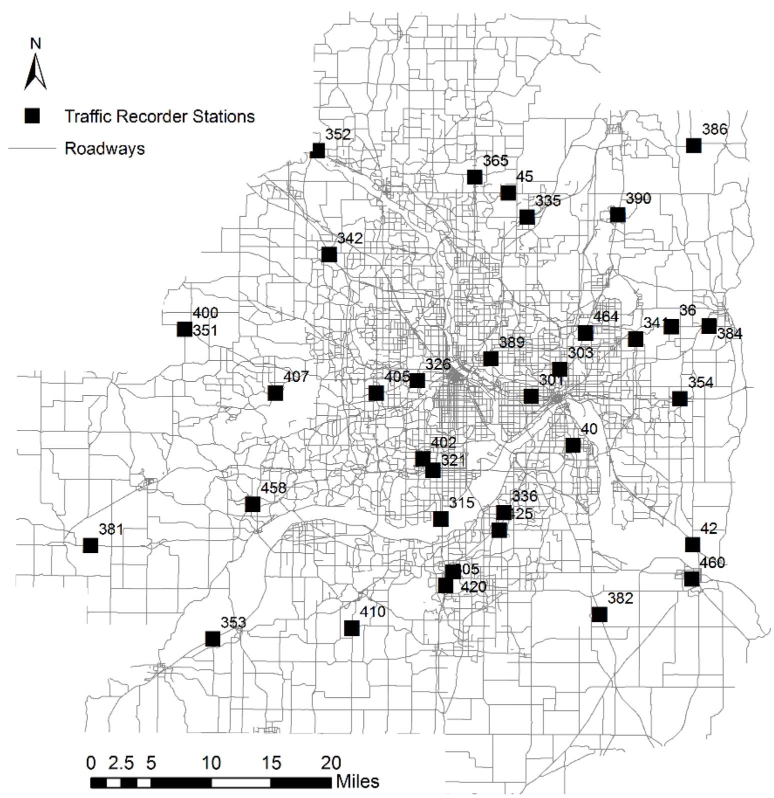

Figure 4.2 Geographic distribution of traffic recorder stations and roadway system

4.2.3 Simulated Individual Trajectory Data

This study used a set of simulated individual-level one-day trajectories. The trajectories were used to overlay traffic emission fields for assessing personal exposure to traffic emissions. Given the importance of on-road travel behaviors, two assumptions were made when generating the simulated trajectory data. First, a person must travel along a series of roadways to get to a destination when he or she travels. Second, the amount of time spent in travel microenvironments and the corresponding transportation modes are significant determinants of personal traffic exposure (Dons, Panis, et al. 2011). In this study, simulated trajectories of 100 persons were constrained to travel along roadways determined by the road network data (source: MnDOT). Four transportation modes, stationary state, walking, cycling, and driving, were modeled in the simulation (Dewulf, Neutens, Van Dyck, et al. 2016), which allows the simulated trajectories of 100 persons to mimic different time-activity patterns. Additional implementation specifications related to the simulation can be found in Chapter 3. Figure 4.3 illustrates one person's one-day trajectory, which is comprised of a string of points.

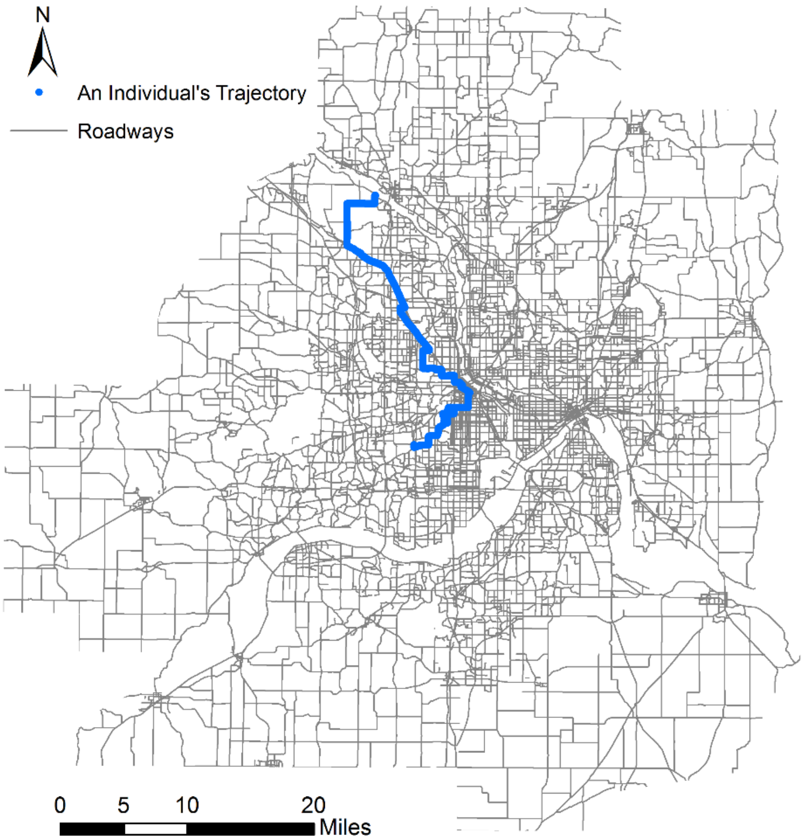

Figure 4.3 Simulated trajectory of one person's activity patterns in the Minneapolis-Saint Paul metropolitan area in one day

4.3 Methods

4.3.1 Determination of Hourly Temporal Allocation Factors

Temporal allocation factors (TAFs) provide a simple means to apportion annual average traffic volume to hourly estimates. Multiplying the annual average traffic count (e.g., AADT, HCAADT) by hourly TAFs results in hourly estimates. Unlike studies that utilize a single national or local TAFs profile across the whole region (Dons et al. 2013; Batterman, Cook, and Justin 2015), this study presents a new method to calculate hourly TAFs. This method considers the spatiotemporal characteristics of traffic activities based on the continuous measurements of traffic counts from the campaign of traffic monitors in the study area.

Traffic activity patterns differ on workdays, Saturdays, Sundays and holidays (Batterman, Ganguly, and Harbin 2015). This study focuses only on traffic patterns on workdays. Traffic patterns during other days can be processed in similar ways. For this purpose, only hourly records from the recorder stations on workdays were used for the estimation of hourly TAFs. Another reason for focusing on workday traffic patterns is that the simulated trajectories mimic the behaviors of "9-to-5" office workers on a workday as described in Chapter 3.

For each traffic recorder station, the total annual daily traffic volume in all workdays throughout the year of 2016 was aggregated first regardless of traffic directions. Then, for each hour, the total annual hourly traffic volume recorded at each recorder site was aggregated throughout the year of 2016. Finally, the ratios of the traffic volume of each recorder station by hours were determined using the following equation:

$$TAF_{i,h} = \frac{V_{i,h}}{V_i} \qquad (4.1)$$

Where $TAF_{i,h}$ = estimated ratio of traffic volume in a particular hour h at site i; $V_{i,h}$ = total annual traffic volume in a particular hour h at site i; V_i = total annual daily traffic volume at site i.

4.3.2 Construction of an Annual Hourly Traffic Emission Model

The construction of the annual hourly traffic emission model was based on a study by Pratt et al. (2014) that developed an innovative traffic emission model using traffic count data. The modeled traffic density was validated in correspondence to the intensity of the traffic-related pollutant NO_2. This study, therefore, uses the terms traffic emission and traffic density interchangeably for ease of discussion in the rest of this chapter.

First, the TAFs of each roadway was determined. It is assumed that the traffic activities of a roadway are spatiotemporally affected by nearby roadways. Hence, this study utilized a spatial interpolation algorithm, inverse distance weighting (IDW), to capture the spatial variation of traffic activities in a particular hour. This study considered three nearest stations when running the IDW algorithm for estimating the traffic TAFs at unknown locations. This procedure resulted in twenty-four traffic TAFs raster surfaces. The value in a cell of a raster layer represents the ratio of hourly traffic share to daily total traffic volume during a specific hour.

Second, a toxicity-weighted traffic count layer was generated by combining the AADT and HCAADT data (Pratt et al. 2014). The toxicity-weighted traffic count layer was then converted into a point layer that is at a resolution of 100 meters. This point layer overlaid each traffic TAFs surface for calculating the hourly share of traffic volume. Within the overlay analysis, the hourly traffic count of a point of a particular roadway was determined by multiplying the weighted traffic count by the corresponding hour-specific ratio. This procedure resulted in twenty-four layers that represent the respective hourly traffic volume of all roadways.

Third, for each layer of hourly traffic volume, a kernel density estimation algorithm was applied to all roadway points. The parameters to run the algorithm were set up with a scaling factor of 1,000,000, the kernel set to bivariate normal, and a smoothing factor of 300 meters (Pratt et al. 2014). This step resulted in twenty-four raster surfaces at a resolution of fifty meters. Each cell on a raster layer represents the value of the estimated traffic density in a particular hour. The unit of the traffic density was traffic counts/m^2.

4.3.3 Linking of the Hourly Traffic Emission Model and Trajectories

The simulated individual-level trajectories were superimposed to the twenty-four traffic density surfaces. The 24-hour surfaces constituted a 3-D spatiotemporal cube carrying the assumption that there is no transitional change regarding traffic counts within one hour. The linking procedure was implemented in the PostgreSQL (version 9.5) with the PostGIS (version 2.3.5) extension. Each point of a trajectory was assigned a traffic density estimate after the linking procedure.

4.3.4 Exposure Assessment

This study used the average exposure over 24 hours to represent personal exposure to traffic emission. Specifically, the summation of the traffic density estimates of all points of a particular trajectory divided by the total number of points equals to the average exposure.

4.4 Results and Discussions

4.4.1 Results of TAFs

Figure 4.4 shows the traffic allocation factors of all recorder stations across all workdays in 2016 by hours. The overall distribution appears to be a bimodal shape with one peek at 7-8 am and another one at 4-5 pm. It is noted that there is a considerable amount of the variation across all stations by hours. Specifically, on an hourly basis, the range of traffic allocation factors was up to 6.14%. Also, Site 464 and site 402 does not follow the common bimodal trend. These findings suggested that the utilization of one single TAFs profile across the whole area may not reflect the spatiotemporal variation of traffic activities well.

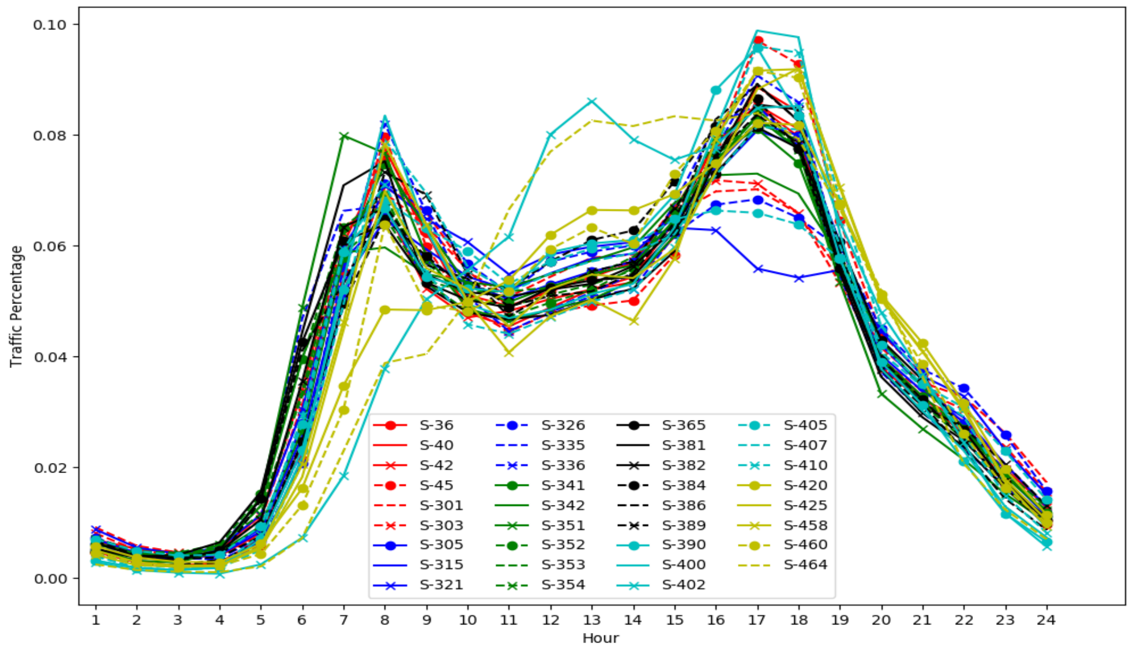

Figure 4.4 Hourly TAFs across thirty-five traffic recorder stations on workdays in 2016

Within each hour there were thirty-five station-specific values of traffic allocation factors across the study area. The thirty-five factors served as model inputs for the spatial interpolation algorithm. This interpolation procedure resulted in thirty-five raster layers at a resolution of fifty meters. Each raster layer represented the geographic distribution of traffic allocation factors in a particular hour. A cell on a particular raster surface stored the value of allocation factor at its location for a certain hour. Figure 4.5 shows the interpolated TAFs maps in 24 hours. As can be seen clearly, most layers do not have a uniform distribution except for hours 2, 3 and 4. The uniform distribution from these hours was due to a lack of traffic across the majority of the study area. The non-uniform distribution suggested that a single TAFs profile across the whole area may not reflect the spatial and temporal dynamics of traffic activity patterns well.

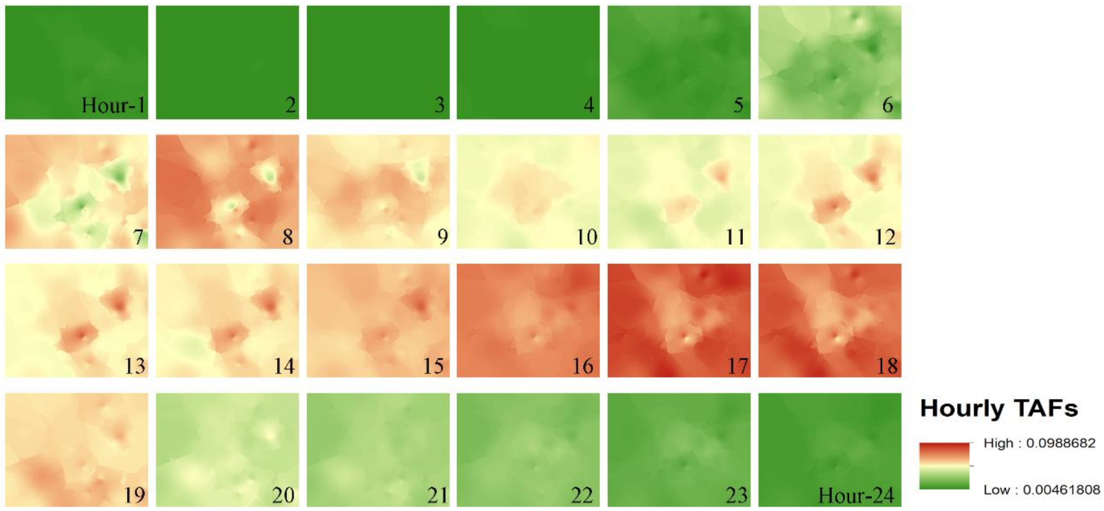

Figure 4.5 Interpolated maps of TAFs in twenty-four hours in the study area

4.4.2 Results of the Annual Average Hourly Traffic Emission Model

Figure 4.6 shows the annual hourly traffic density layers for 24 hours. The traffic density layers indicated that people within the study area largely began commuting around 5 a.m. Light traffic activity was observed between 1 a.m. to 5 a.m., and traffic activities occurred only on some highways (e.g., I-94, I-494, I-694, and I-35). After 5 a.m., traffic activities took place on a majority of the non-highway roads within the metropolitan area.

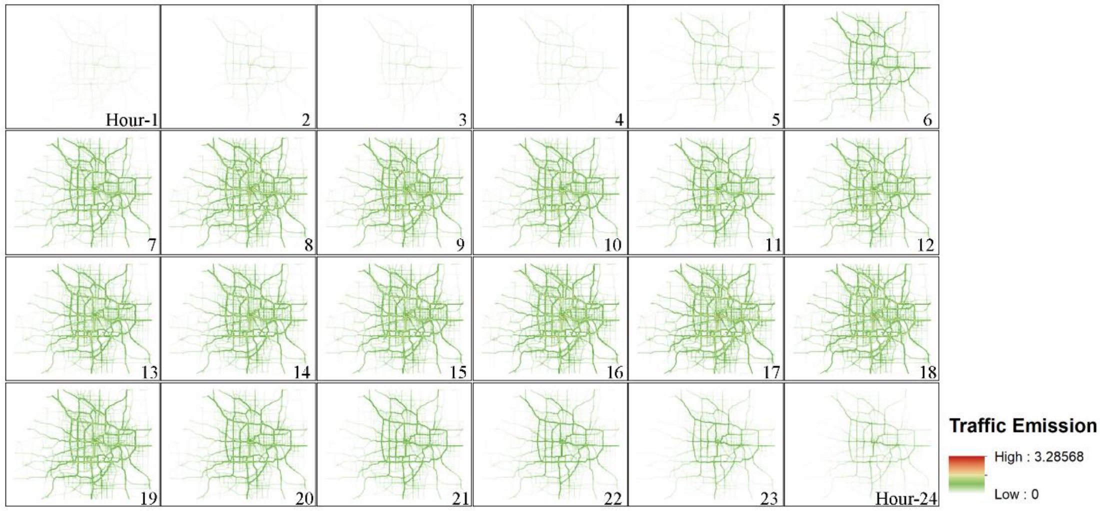

Figure 4.6 Hourly traffic density maps in twenty-four hours

A close-up view of the traffic density surface for hour 8 is illustrated in Figure 4.7. The zones with heavy traffic loads are highlighted by rectangles. Areas 1, 5, and 6 appear to be the most crowded roadways in hour 8. Area 1 is located at the intersection of I-35 and I-494. Areas 5 and 6 were the regions close to the downtown areas of Minneapolis and St Paul, respectively.

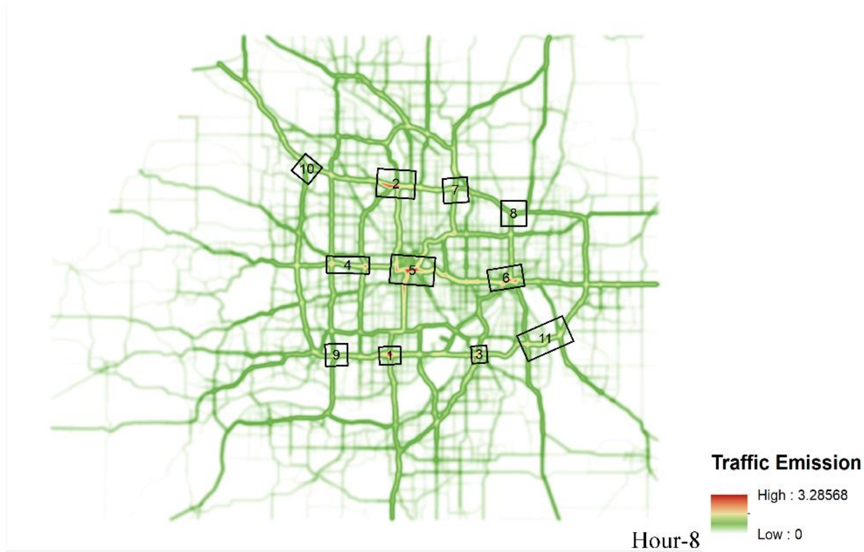

Figure 4.7 A close-up view of the traffic density map in hour 8 along with highlighted traffic density hotspots

A close-up view of traffic density in hour 17 is illustrated in Figure 4.8, which shows the second peek of daily traffic activities. When compared to hour 8, one noticeable difference is that hour 17 has slightly higher traffic densities, thus leading to a more severe traffic emission scenario.

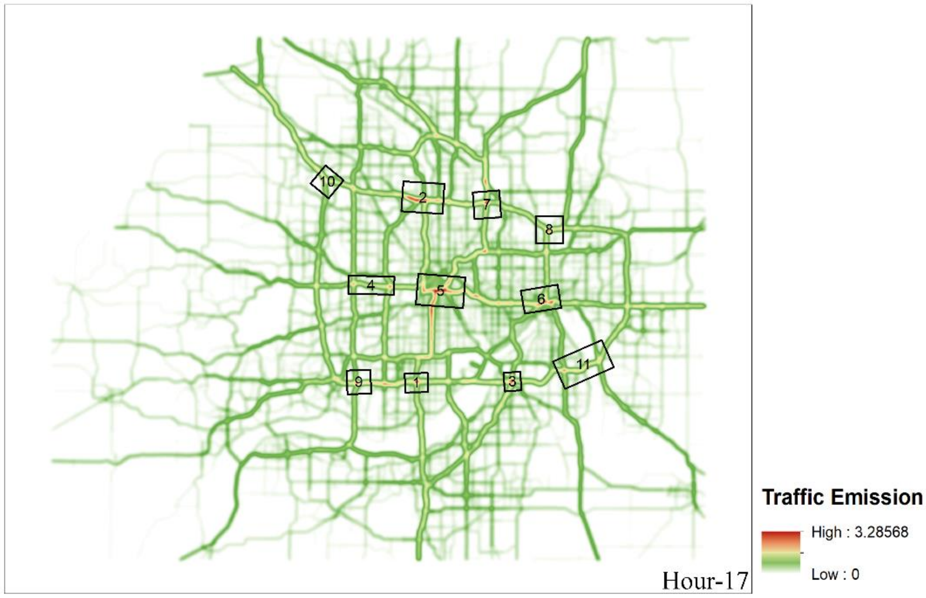

Figure 4.8 Traffic density map in hour 17 along with highlighted traffic density hotspots

Between hour 18 and hour 23, a few roadways are found to be highly congested. That is, the roads between areas 5 and 6 (i.e., the downtown areas), the roads between areas 5 and 1, the roads between areas 9 and 1, and the roads between areas 10 and 2. After hour 23, only the two downtown areas have relatively more vehicles passing, and there is no obvious difference between major highways concerning traffic densities.

4.4.3 Results of Personal Exposure to Traffic Emissions by Hours

Table 4.1 shows the maximum, minimum, median, mean and standard deviation of the personal exposure associated with 100 simulated trajectories.

Table 4.1 The descriptive statistics of personal exposures of 100 trajectories

	Maximum	Minimum	Median	Mean	Standard deviation
Trajectories (N=100)	1.088	0.044	0.188	0.238	0.159

(Unit: traffic counts/m^2)

Trajectories 56, 49 and 45 were found to have the maximum, median and minimum exposure respectively among the 100 trajectories. Figure 4.9 shows the three trajectories over the course of 24 hours. The home locations of these three trajectories are highlighted using exaggerated dotted circles.

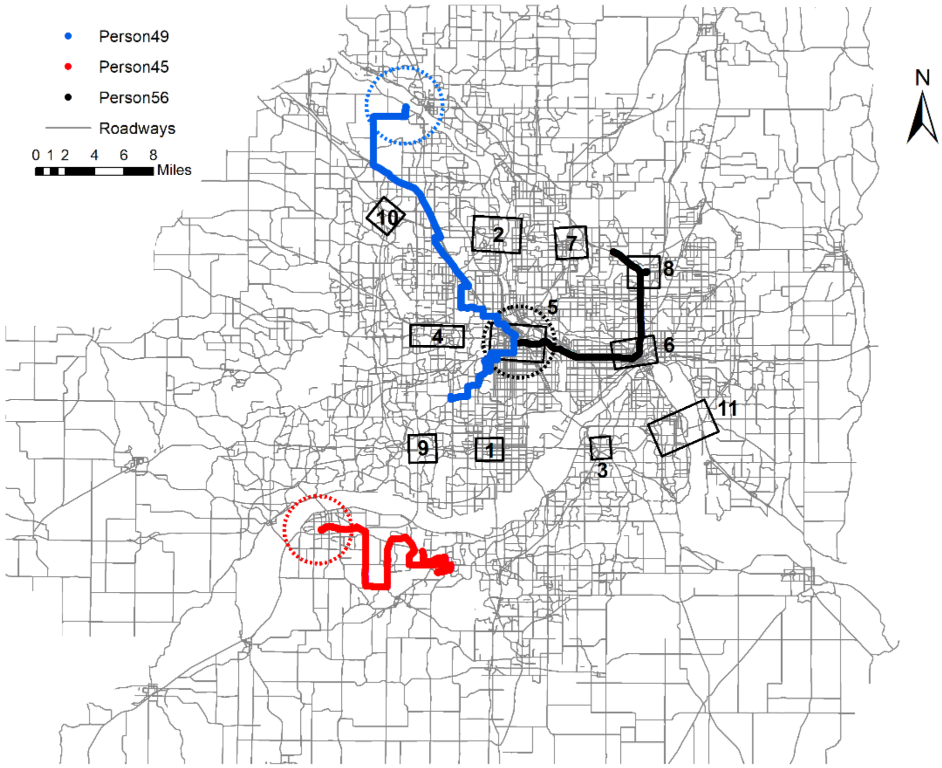

Figure 4.9 Trajectories 56, 49 and 45 along with traffic density hotspots (Note: the home locations of these three persons are marked by exaggerated dotted circles)

Trajectory 56, which was observed to have the highest exposure, began the commute at 6 a.m. The home was located within the traffic hotspot 5, which was in the downtown Minneapolis area. The person of trajectory 56 passed through hotspot 6 in the morning rush-hour while on route to the workplace that was in hotspot 8. After 5 p.m., this person went northwest and stayed at a location between hotspots 7 and 8. Then, this person returned to the home location. All routes of person 56 were along congested roadways and near traffic hotspots.

Trajectory 49, which had the median average exposure, began the commute from the home location that was far away from the downtown areas. This person did not commute on a heavily congested roadway for most of the time. However, this person experienced relatively high exposure while traveling toward and staying in the workplace.

Trajectory 45, which had the least exposure, did not travel on major highways or stay in close proximity to the two downtown areas. Person 45's time-activity patterns featured a suburban residential place, a workplace outside of downtown, and "local" commute routes. This resulted in person 45 being exposed to the least traffic emissions.

Figure 4.10 shows the detailed exposure profiles of these three trajectories over the course of 24 hours. Before 6 a.m., all three persons were exposed to low traffic emissions because there was very little traffic during these time periods. Person 56 had relatively higher exposure than persons 49 and 45 though. This was in part due to person 56's home location being close to the Minneapolis downtown area (see Figure 4.9). The first exposure peaks for these three persons were observed between hours 6 and 9. These exposure peaks coincided with a peak in traffic activities. From hours 9 to 17, the exposure magnitudes of all persons decreased significantly because all persons were programmed to stay at their workplace instead of traveling on roadways. The exception was that person 56 still got exposed to an elevated level of traffic emission. The reason is that its workplace is located in a traffic hotspot (i.e., area 8). After hour 17, all simulated trajectories of 100 persons began their commute home, which caused the second exposure peak. After hour 21, all persons' profiles decreased to a low level.

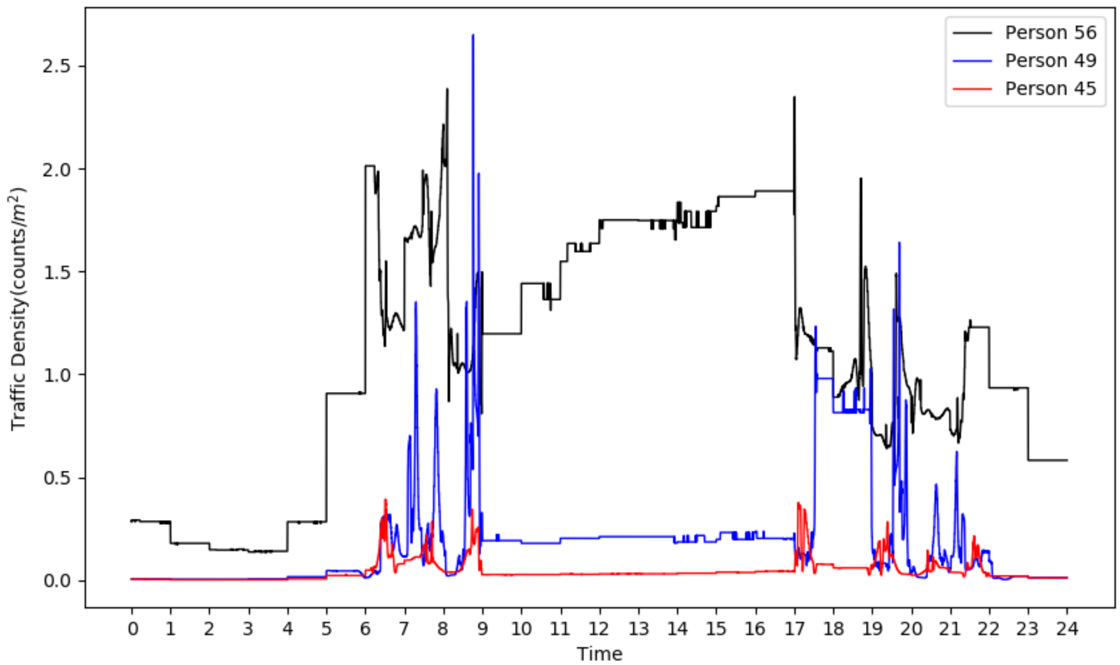

Figure 4.10 Exposure profiles of persons 56, 49 and 45 over the course of 24 hours

The highest "instantaneous" exposure out of the three select persons was observed on person 49. Person 49's home and workplace are not close to highways or traffic hotspots. But, this person commuted through traffic hotspot 5 during the morning rush hours. A similar pattern occurred during the afternoon rush hours. In conclusion, person 49 always travel through areas having the most elevated emissions at rush hours. This finding suggests that personal exposure to traffic emissions is indeed a complex spatiotemporal function. The finding also evidences that using single location (e.g., residential place) to assess personal exposure to traffic emissions cannot represent the actual exposure that a person is confronted to.

4.5 Conclusions

The objective of this part of the research was to develop a new approach to quantifying personal exposure to traffic emissions by accounting for a person's individual-level trajectory and spatiotemporal variations of traffic activities. Moreover, this study also investigated how personal traffic emissions varied across geographic space and time. This study used continuous traffic count data from traffic recorder stations to derive spatiotemporally resolved TAFs. Then, this study built up an hourly traffic emission model to assign hourly traffic emission estimates to a set of individual-level human trajectories. For demonstration purposes, this study selected three trajectories to explore how personal exposure was affected by personal time-activity patterns and spatiotemporally varying traffic emissions. The findings of this study are summarized as follows.

First, this study derived station-specific TAFs by hours using the continuous traffic count data from a campaign of traffic recorder stations. It was found that the hourly TAFs did not show a spatially homogeneous pattern. This finding supports our initiative to apply the spatial interpolation analysis to capture the spatial variations of traffic activities. A set of hourly TAFs surfaces were then generated.

Second, this study produced 24-hour traffic emission surfaces in the study area. A total of 100 simulated one-day trajectories were superimposed onto each hourly emission surface for calculating exposure estimates of all trajectory points. The three trajectories were selected to examine the exposure profiles. This study found that a person's time-activity pattern (e.g., the location of a person's home and workplace, the commute routes) were important factors in determining personal exposure to traffic emissions. Another finding was that personal exposure to traffic emissions varied greatly in geographic space and time. Using single location to estimate personal exposure to traffic emissions may lead to exposure misclassification due to spatiotemporal variation.

This model developed in this chapter has a few implications in exposure science, public health, and epidemiology. First, it is believed that it will improve the accuracy of exposure estimates by capturing more “activity space” instead of using only one location. Second, this model can provide enriched information about how personal exposure to traffic emissions varies across space and time. This information may ultimately promote the understanding of the association between traffic emissions and health effects in epidemiological studies, as well as help reduce the exposure risk.

Several limitations can be observed. First, traffic activity variations between daily, weekly and seasonal patterns were not accounted for. This could be overcome by using hierarchical temporal TAFs accurately (Batterman, Cook, and Justin 2015). Second, when producing the spatial distribution maps of hourly TAFs, this study assumed nearby roadways were more related regarding traffic activities. However, accidents, congestions, meteorological influences, natural hazards (e.g., flooding or storms), and other events that affect traffic activity were not addressed in this study.

5. CONCLUSIONS

This chapter provides a summary of major findings and their corresponding discussion for this study. It also discusses some limitations of this study, and some topics for future research.

5.1 Summary of Findings and Discussion

The overarching objective of this study is twofold. The first is to model both human activities and spatiotemporally varying air pollution in geographic space and time for assessing personal exposure. The second objective of this study is to characterize how personal exposure varies over space and time. To achieve these objectives, this study attempted to achieve three specific research aims listed below.

- To develop a trajectory-based model for assessing personal exposure to ambient air pollution.
- To evaluate if the consideration of the spatiotemporal dynamics of both human activities and air pollution will bring in a significant difference to exposure assessment.
- To develop a trajectory-based model for assessing personal exposure to traffic emission.

The first specific research aim is addressed in Chapter 3. This chapter proposes an innovative model to assess personal exposure to ambient air pollution. This model couples two types of data, individual-level human trajectory data and hourly air pollution data collected by a ground monitoring network. Without foreknowledge of the locations of significant places (e.g., home, workplace), this model takes advantage of spatiotemporal trajectory mining techniques to apportion trajectory points into a set of significant places (called microenvironments (MEs)). Next, we construct a 3D spatiotemporal cube to represent the spatiotemporal distribution of $PM_{2.5}$ in geographic space and time. The constructed 3D cube is then overlaid with processed trajectory data, using the matched geographic coordinates and timestamp attributes, in a GIS environment. This newly proposed model is evaluated using both simulated and empirical trajectories to demonstrate the computational power of the model.

The findings from Chapter 3 show that the developed model works well. The model that considers human activities and spatiotemporal variation of air pollution can offer more accurate, detailed, and enriched information to exposure assessment. This chapter also confirms the inter- and intra-microenvironmental variation with respect to personal exposure to ambient air pollution. Another finding is that time-activity patterns of a person are significant factors for personal exposure. Therefore, a model that considers the home location of a person only may not fully reflect the true exposure of an individual.

The second specific research aim is achieved by the work described in the last part of Chapter 3. At the end of Chapter 3, an exploratory endeavor was conducted to answer the question of whether the estimated exposure accounting for home location only is statistically significantly different from that considering the locations of all activities. We used both simulated (organized in three datasets) and empirical trajectories with differing time-activity patterns to test the methods. For both methods, the average hourly average exposure was calculated for comparison purposes.

The findings regarding the second aim are that there is no consistent evidence that a trajectory-based method can produce statistically significantly different results than a home-based method with respect to personal exposure. The inconsistence is attributed to many factors, such as the differing characteristics of human activities, the spatiotemporal variation of the air pollutant in question, and the low statistical power. The mixed results, therefore, suggest more studies are needed to verify how the two exposure results differ when confronted with different groups of people and differing spatiotemporal distribution patterns of air pollutants.

Chapter 4 illustrates a new model to quantify traffic emission when human activities are accounted for. The key question in Chapter 4 is how to model the spatiotemporal variation of traffic emissions on an hourly basis prior to exposure assessment. First, this study derived station-specific traffic allocation factors using the traffic count data recorded by a number of traffic recorder stations in the study area. This study utilized a spatial interpolation algorithm to capture the spatial variation of traffic activities. The average hourly count of each roadway was determined by multiplying the average daily count by its corresponding TAF at a specific hour. A traffic emission model is then applied to each roadway for representing the traffic emission at each hour of the day. By assuming there are no transitional changes of traffic emissions in consecutive hours, a 3D spatiotemporal cube is built using the twenty-four hourly traffic emission layers. Finally, the traffic emission cube was overlaid with simulated individual-level trajectory data for assessing personal exposure to traffic emissions.

5.2 Limitations

The first limitation is that this study assumes that simulated individual-level trajectory data follows certain spatiotemporal constraints from standard groups of "9-to-5" people. This study utilizes the simulated trajectory data complying with these assumed spatiotemporal constraints. In reality, many groups of people may not conform to these standard constraints. However, this study does not seek to develop an advanced trajectory classification algorithm to apportion trajectory points into the appropriate microenvironments. Nor does this study aim to explain how specific characteristics of human activities influence the assessment of personal exposure.

The second limitation is that, in the process of modeling ambient air pollution, due to data availability, this study does not lean on advanced algorithms (e.g., dispersion models, land-use regression models) that usually have demanding requirements on data inputs. Alternatively, this study compares a row of spatial interpolation algorithms and accepts the one with the least error in different temporal instances. In addition, the Wilcoxon Signed-rank test in Chapter 3 is impacted by the performance of the air pollution modeling also to some extent.

Third, in deriving the TAFs of traffic activities, this study assumes that, at a particular hour, the traffic activity of any location is affected by only its three nearest traffic recorder stations. This assumption seems to have an impact on the quantification of hourly traffic emissions, thus affecting personal exposure assessment.

Finally, the quality of the road network data may have some influences on the quantification of traffic emissions, as well as the simulated trajectory data. For example, the topology of the road network data might affect the route choices in the trajectory simulation.

5.3 Future Research Direction

This study has certain implications for exposure- and health-related studies. As noted throughout the study, one of the main motivations is to mitigate exposure misclassification by offering better and more accurate exposure estimates. Compared to the commonly used methods that solely use home locations for exposure assessment, this study considers the dynamics of both human activities and air pollution in geographic space and time for assessing personal exposure to ambient air pollution and traffic emission. The models developed in this study and the findings would greatly benefit epidemiological studies. Some important topics for future research are highlighted below.

First, the trajectory simulation procedures this study developed can turn the existing time-activity data (e.g., diary data) into geographic trajectories for exposure assessment purposes. In most cases, the endeavor of collecting time-activity information was not designated for exposure assessment. Time-activity focused study, and exposure-centered study is usually carried out in separate experiments and for differing research purposes. The pre-existing data (e.g., California Activity Pattern Survey (CAPS), Consolidated Human Activity Database (CHAD), Multinational Time Use Study (MTUS), etc.) will create tremendous opportunities for the application of this research. Future research can use the developed trajectory simulation platform to generate human's trajectory following the realistic patterns for some region. Then, the simulated trajectory data can be coupled with different sources of environmental air pollution data to assess individual- or population-level exposure. This research will improve the understanding of human exposure to various air pollutants in a region when human activities are well represented in the modeling.

Another potential research direction is to investigate how human time-activity patterns affect the assessment of personal exposure through sensitivity analysis. Human time-activity patterns feature the visited places, the duration of staying at some places, the travel distances, and some other relevant characteristics. All of these characteristics, treated as parameters of some defined distribution, can be modeled in Monte Carlo analysis. Then, the permutations of different settings of parameters will generate as diverse trajectories as possible. This research can help better understand the relationship between human activities and personal exposure assessment, as well as how personal exposure varies over geographic space and time for different groups.

In addition, this study can be used to evaluate the evidence of the association between air pollutant in question and its related health outcomes in epidemiology. As noted in the literature review section, many epidemiology studies only use home-based exposure to understand the association between the air pollutant in question and its related human health outcomes. Future research will replicate some epidemiological studies in certain areas using the method presented in this study. The exposure estimates for a certain area or a group of people will be quantified by trajectory-based models rather than the home-based methods. The improved exposure estimates will be used to seek the association between human exposure to the air pollutant in question and adverse health outcomes.

Given that more and more people intend to choose an environmentally-friendly mode of transportation (e.g., cycling, running), this future research expanding on Chapter 4 will develop a web-based route planning tool for the public in Minneapolis–Saint Paul area to plan routes regarding personal exposure to traffic emissions. This tool expected to allow users to specify the origins, destinations, preferred transportation mode, and the travel start time using an interactive web map. Based on the users' inputs, that tool will return users the optimal routes having the lowest exposure to traffic emissions.

LITERATURE CITED

Adams, C., P. Riggs, and J. Volckens. 2009. Development of a method for personal, spatiotemporal exposure assessment. *J Environ Monit* 11 (7):1331-9.

Adams, M. D., N. Yiannakoulias, and P. S. Kanaroglou. 2016. Air pollution exposure: An activity pattern approach for active transportation. *Atmospheric Environment* 140:52-59.

Adgate, J. L., T. R. Church, A. D. Ryan, G. Ramachandran, A. L. Fredrickson, T. H. Stock, M. T. Morandi, and K. Sexton. 2004. Outdoor, indoor, and personal exposure to VOCs in children. *Environ Health Perspect* 112 (14):1386-92.

Akland, G. G., T. D. Hartwell, T. R. Johnson, and R. W. Whitmore. 1985. Measuring human exposure to carbon monoxide in Washington, DC, and Denver, Colorado, during the winter of 1982-1983. *Environmental Science & Technology* 19 (10):911-918.

Arku, R. E., K. L. Dionisio, A. F. Hughes, J. Vallarino, J. D. Spengler, M. C. Castro, S. Agyei-Mensah, and M. Ezzati. 2015. Personal particulate matter exposures and locations of students in four neighborhoods in Accra, Ghana. *J Expo Sci Environ Epidemiol* 25 (6):557-66.

Armstrong, B. G. 1996. Optimizing power in allocating resources to exposure assessment in an epidemiologic study. *American Journal of Epidemiology* 144 (2):192-197.

Avery, C. L., K. T. Mills, R. Williams, K. A. McGraw, C. Poole, R. L. Smith, and E. A. Whitsel. 2010. Estimating error in using residential outdoor PM2.5 concentrations as proxies for personal exposures: a meta-analysis. *Environ Health Perspect* 118 (5):673-8.

Batterman, S., R. Cook, and T. Justin. 2015. Temporal variation of traffic on highways and the development of accurate temporal allocation factors for air pollution analyses. *Atmospheric Environment* 107:351-363.

Batterman, S., R. Ganguly, and P. Harbin. 2015. High Resolution Spatial and Temporal Mapping of Traffic-Related Air Pollutants. *International Journal of Environmental Research and Public Health* 12 (4):3646-3666.

Beko, G., B. U. Kjeldsen, Y. Olsen, J. Schipperijn, A. Wierzbicka, D. G. Karottki, J. Toftum, S. Loft, and G. Clausen. 2015. Contribution of various microenvironments to the daily personal exposure to ultrafine particles: Personal monitoring coupled with GPS tracking. *Atmospheric Environment* 110:122-129.

Breen, M. S., T. C. Long, B. D. Schultz, J. Crooks, M. Breen, J. E. Langstaff, K. K. Isaacs, Y. M. Tan, R. W. Williams, Y. Cao, A. M. Geller, R. B. Devlin, S. A. Batterman, and T. J. Buckley. 2014. GPS-based microenvironment tracker (MicroTrac) model to estimate time-location of individuals for air pollution exposure assessments: model evaluation in central North Carolina. *J Expo Sci Environ Epidemiol* 24 (4):412-20.

Bricka, S., and C. Bhat. 2006. Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data. *Transportation Research Record: Journal of the Transportation Research Board* 1972:9-20.

Bricka, S. G., S. Sen, R. Paleti, and C. R. Bhat. 2012. An analysis of the factors influencing differences in survey-reported and GPS-recorded trips. *Transportation Research Part C: Emerging Technologies* 21 (1):67-88.

Briggs, D. J., A. R. Denman, J. Gulliver, R. F. Marley, C. A. Kennedy, P. S. Philips, K. Field, and R. M. Crockett. 2003. Time activity modelling of domestic exposures to radon. *J Environ Manage* 67 (2):107-20.

Broich, A. V., L. E. Gerharz, and O. Klemm. 2011. Personal monitoring of exposure to particulate matter with a high temporal resolution. *Environ Sci Pollut Res Int* 19 (7):2959-72.

Buonanno, G., S. Marini, L. Morawska, and F. C. Fuoco. 2012. Individual dose and exposure of Italian children to ultrafine particles. *Sci Total Environ* 438:271-7.

Buonanno, G., L. Stabile, and L. Morawska. 2014. Personal exposure to ultrafine particles: the influence of time-activity patterns. *Sci Total Environ* 468-469:903-7.

Buonanno, G., L. Stabile, L. Morawska, and A. Russi. 2013. Children exposure assessment to ultrafine particles and black carbon: The role of transport and cooking activities. *Atmospheric Environment* 79:53-58.

de Nazelle, A., E. Seto, D. Donaire-Gonzalez, M. Mendez, J. Matamala, M. J. Nieuwenhuijsen, and M. Jerrett. 2013. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environ Pollut* 176:92-9.

Dennekamp, M., O. Mehenni, J. Cherrie, and A. Seaton. 2002. Exposure to ultrafine particles and PM2. 5 in different micro-environments. *Annals of Occupational Hygiene* 46 (suppl_1):412-414.

Dewulf, B., T. Neutens, W. Lefebvre, G. Seynaeve, C. Vanpoucke, C. Beckx, and N. Van de Weghe. 2016. Dynamic assessment of exposure to air pollution using mobile phone data. *Int J Health Geogr* 15:14.

Dewulf, B., T. Neutens, D. Van Dyck, I. de Bourdeaudhuij, L. I. Panis, C. Beckx, and N. Van de Weghe. 2016. Dynamic assessment of inhaled air pollution using GPS and accelerometer data. *JOURNAL OF TRANSPORT & HEALTH* 3 (1):114-123.

Dias, D., and O. Tchepel. 2014. Modelling of human exposure to air pollution in the urban environment: a GPS-based approach. *Environ Sci Pollut Res Int* 21 (5):3558-71.

Dijkstra, E. W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1:269-271.

Dons, E., L. Int Panis, M. Van Poppel, J. Theunis, H. Willems, R. Torfs, and G. Wets. 2011. Impact of time-activity patterns on personal exposure to black carbon. *Atmospheric Environment* 45 (21):3594-3602.

Dons, E., L. I. Panis, M. Van Poppel, J. Theunis, H. Willems, R. Torfs, and G. Wets. 2011. Impact of time-activity patterns on personal exposure to black carbon. *Atmospheric Environment* 45 (21):3594-3602.

Dons, E., P. Temmerman, M. Van Poppel, T. Bellemans, G. Wets, and L. Int Panis. 2013. Street characteristics and traffic factors determining road users' exposure to black carbon. *Sci Total Environ* 447:72-9.

Duan, N. 1982. Models for human exposure to air pollution. *Environment International* 8 (1-6):305-309.

Elgethun, K., R. A. Fenske, M. G. Yost, and G. J. Palcisko. 2002. Time-Location Analysis for Exposure Assessment Studies of Children Using a Novel Global Positioning System Instrument. *Environmental Health Perspectives* 111 (1):115-122.

Elgethun, K., M. G. Yost, C. T. Fitzpatrick, T. L. Nyerges, and R. A. Fenske. 2007. Comparison of global positioning system (GPS) tracking and parent-report diaries to characterize children's time-location patterns. *J Expo Sci Environ Epidemiol* 17 (2):196-206.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Paper read at Kdd.

Fang, T. B. 2012. *A pseudo individual near real-time measurement for assessing air pollution exposure in selected Texas cities*: Texas State University-San Marcos.

Fang, T. B., and Y. Lu. 2012. Personal real-time air pollution exposure assessment methods promoted by information technological advances. *Annals of GIS* 18 (4):279-288.

Fang, T. F. B., and Y. M. Lu. 2011. Constructing a Near Real-time Space-time Cube to Depict Urban Ambient Air Pollution Scenario. *Transactions in GIS* 15 (5):635-649.

Fiordelisi, A., P. Piscitelli, B. Trimarco, E. Coscioni, G. Iaccarino, and D. Sorriento. 2017. The mechanisms of air pollution and particulate matter in cardiovascular diseases. *Heart Failure Reviews* 22 (3):337-347.

Gauvin, S., Y. Le Moullec, F. Bremont, I. Momas, F. Balducci, F. Ciognard, M. P. Poilve, and D. Zmirou. 2001. Relationships between nitrogen dioxide personal exposure and ambient air monitoring measurements among children in three French metropolitan areas: VESTA study. *Arch Environ Health* 56 (4):336-41.

Gerharz, L. E., O. Klemm, A. V. Broich, and E. Pebesma. 2013. Spatio-temporal modelling of individual exposure to air pollution and its uncertainty. *Atmospheric Environment* 64:56-65.

Gerharz, L. E., A. Kruger, and O. Klemm. 2009. Applying indoor and outdoor modeling techniques to estimate individual exposure to PM2.5 from personal GPS profiles and diaries: a pilot study. *Sci Total Environ* 407 (18):5184-93.

Glasgow, M. L., C. B. Rudra, E. H. Yoo, M. Demirbas, J. Merriman, P. Nayak, C. Crabtree-Ide, A. A. Szpiro, A. Rudra, J. Wactawski-Wende, and L. Mu. 2016. Using smartphones to collect time-activity data for long-term personal-level air pollution exposure assessment. *J Expo Sci Environ Epidemiol* 26 (4):356-64.

Greaves, S., T. Issarayangyun, and Q. Liu. 2008. Exploring variability in pedestrian exposure to fine particulates (PM2.5) along a busy road. *Atmospheric Environment* 42 (8):1665-1676.

Harrison, R. M., C. A. Thornton, R. G. Lawrence, D. Mark, R. P. Kinnersley, and J. G. Ayres. 2002. Personal exposure monitoring of particulate matter, nitrogen dioxide, and carbon monoxide, including susceptible groups. *Occup Environ Med* 59 (10):671-9.

Hazlehurst, M. F., E. W. Spalt, C. L. Curl, M. E. Davey, S. Vedal, G. L. Burke, and J. D. Kaufman. 2017. Integrating data from multiple time-location measurement methods for use in exposure assessment: the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *J Expo Sci Environ Epidemiol*.

Health Effects Institute. 2010. *Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects*: Health Effects Institute,.

Heusinkveld, H. J., T. Wahle, A. Campbell, R. H. S. Westerink, L. Tran, H. Johnston, V. Stone, F. R. Cassee, and R. P. F. Schins. 2016. Neurodegenerative and neurological disorders by small inhaled particles. *Neurotoxicology* 56:94-106.

Houston, D., P. Ong, G. Jaimes, and A. Winer. 2011. Traffic exposure near the Los Angeles-Long Beach port complex: using GPS-enhanced tracking to assess the implications of unreported travel and locations. *Journal of Transport Geography* 19 (6):1399-1409.

Houston, D., J. Wu, D. Yang, and G. Jaimes. 2013. Particle-bound polycyclic aromatic hydrocarbon concentrations in transportation microenvironments. *Atmospheric Environment* 71:148-157.

Hubal, E. A. C., L. S. Sheldon, J. M. Burke, T. R. McCurdy, M. R. Berry, M. L. Rigas, V. G. Zartarian, and N. C. G. Freeman. 2000. Children's Exposure Assessment: A Review of Factors Influencing Children's Exposure, and the Data Available to Characterize and Assess That Exposure. *Environmental Health Perspectives* 108 (6):475-486.

Huck, J. J., J. D. Whyatt, P. Coulton, B. Davison, and A. Gradinar. 2017. Combining physiological, environmental and locational sensors for citizen-oriented health applications. *Environ Monit Assess* 189 (3):114.

Jerrett, M., A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison, and C. Giovis. 2005. A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol* 15 (2):185-204.

Johnson, T. 1984. Study of Personal Exposure to Carbon Monoxide in Denver Colorado. Washington, D.C.: U.S. Environmental Protection Agency.

Kaufman, J. D., S. D. Adar, R. W. Allen, R. G. Barr, M. J. Budoff, G. L. Burke, A. M. Casillas, M. A. Cohen, C. L. Curl, M. L. Daviglus, A. V. Diez Roux, D. R. Jacobs, Jr., R. A. Kronmal, T. V. Larson, S. L. Liu, T. Lumley, A. Navas-Acien, D. H. O'Leary, J. I. Rotter, P. D. Sampson, L. Sheppard, D. S. Siscovick, J. H. Stein, A. A. Szpiro, and R. P. Tracy. 2012. Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Am J Epidemiol* 176 (9):825-37.

Klepeis, N. E., W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, and W. H. Engelmann. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of exposure analysis and environmental epidemiology* 11 (3):231-252.

Kwan, M.-P. 2012. The Uncertain Geographic Context Problem. *Annals of the Association of American Geographers* 102 (5):958-968.

Kwan, M. P. 2004. Gis methods in time-geographic research: geocomputation and geovisualization of human activity patterns. *Geografiska Annaler: Series B, Human Geography* 86 (4):267-280.

Lai, H. K., M. Kendall, H. Ferrier, I. Lindup, S. Alm, O. Hänninen, M. Jantunen, P. Mathys, R. Colvile, M. R. Ashmore, P. Cullinan, and M. J. Nieuwenhuijsen. 2004. Personal exposures and microenvironment concentrations of PM2.5, VOC, NO2 and CO in Oxford, UK. *Atmospheric Environment* 38 (37):6399-6410.

Lazenby, V., A. Hinwood, A. Callan, and P. Franklin. 2012. Formaldehyde personal exposure measurements and time weighted exposure estimates in children. *Chemosphere* 88 (8):966-73.

Leal, C., and B. Chaix. 2011. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. *Obes Rev* 12 (3):217-30.

Lee, H., W. Myung, D. K. Kim, S. E. Kim, C. T. Kim, and H. Kim. 2017. Short-term air pollution exposure aggravates Parkinson's disease in a population-based cohort. *Scientific Reports* 7:10.

Lee, K., W. J. Parkhurst, J. Xue, H. Özkaynak, D. Neuberg, and J. D. Spengler. 2012. Outdoor/Indoor/Personal Ozone Exposures of Children in Nashville, Tennessee. *Journal of the Air & Waste Management Association* 54 (3):352-359.

Lee, L. A., S. C. Soderholm, M. M. Flemmer, and J. L. Hornsby-Myers. 2005. Field test results of an automated exposure assessment tool, the local positioning system (LPS). *J Environ Monit* 7 (7):736-42.

Leech, J. A., W. C. Nelson, R. T. Burnett, S. Aaron, and M. E. Raizenne. 2002. It's about time: A comparison of Canadian and American time-activity patterns. *Journal of exposure analysis and environmental epidemiology* 12 (6):427-432.

Leech, J. A., K. Wilby, E. McMullen, and K. Laporte. 1996. The Canadian Human Activity Pattern Survey: report of methods and population surveyed. *Chronic Dis Can* 17 (3-4):118-23.

Lei, X., G. Xiu, B. Li, K. Zhang, and M. Zhao. 2016. Individual exposure of graduate students to PM2.5 and black carbon in Shanghai, China. *Environ Sci Pollut Res Int* 23 (12):12120-7.

Liu, L., T. D. Ruddy, M. Dalipaj, M. Szyszkowicz, H. You, R. Poon, A. Wheeler, and R. Dales. 2007. Influence of personal exposure to particulate air pollution on cardiovascular physiology and biomarkers of inflammation and oxidative stress in subjects with diabetes. *J Occup Environ Med* 49 (3):258-65.

Lonati, G., S. Ozgen, G. Ripamonti, S. Cernuschi, and M. Giugliano. 2011. Pedestrian exposure to size-resolved particles in Milan. *J Air Waste Manag Assoc* 61 (11):1273-80.

Lu, Y. M., and T. B. Fang. 2015. Examining Personal Air Pollution Exposure, Intake, and Health Danger Zone Using Time Geography and 3D Geovisualization. *ISPRS International Journal of Geo-Information* 4 (1):32-46.

Luben, T. J., J. L. Nichols, S. J. Dutton, E. Kirrane, E. O. Owens, L. Datko-Williams, M. Madden, and J. D. Sacks. 2017. A systematic review of cardiovascular emergency department visits, hospital admissions and mortality associated with ambient black carbon. *Environment International* 107:154-162.

Marsh, L. M., and R. E. Jones. 1988. The form and consequences of random walk movement models. *Journal of Theoretical Biology* 133 (1):113-131.

Moller, K. L., C. Brauer, S. Mikkelsen, S. Loft, E. B. Simonsen, H. Koblauch, S. H. Bern, T. Alkjaer, O. Hertel, T. Becker, K. H. Larsen, J. P. Bonde, and L. C. Thygesen. 2017. Copenhagen Airport Cohort: air pollution, manual baggage handling and health. *BMJ Open* 7 (5):e012651.

Moller, K. L., L. C. Thygesen, J. Schipperijn, S. Loft, J. P. Bonde, S. Mikkelsen, and C. Brauer. 2014. Occupational exposure to ultrafine particles among airport employees--combining personal monitoring and global positioning system. *PLoS ONE* 9 (9):e106671.

Morabia, A., P. N. Amstislavski, F. E. Mirer, T. M. Amstislavski, H. Eisl, M. S. Wolff, and S. B. Markowitz. 2009. Air pollution and activity during transportation by car, subway, and walking. *Am J Prev Med* 37 (1):72-7.

Morabia, A., F. E. Mirer, T. M. Amstislavski, H. M. Eisl, J. Werbe-Fuentes, J. Gorczynski, C. Goranson, M. S. Wolff, and S. B. Markowitz. 2010. Potential health impact of switching from car to public transportation when commuting to work. *Am J Public Health* 100 (12):2388-91.

National Research Council. 1991a. *Human exposure assessment for airborne pollutants: advances and opportunities*. Washington, DC: National Academies Press.

———. 1991b. *Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities*. Washington, DC: The National Academies Press.

———. 2012. *Exposure Science in the 21st Century: a Vision and a Strategy*: National Academies Press.

Nethery, E., S. E. Leckie, K. Teschke, and M. Brauer. 2008. From measures to models: an evaluation of air pollution exposure assessment for epidemiological studies of pregnant women. *Occupational and Environmental Medicine* 65 (9):579-86.

Nethery, E., G. Mallach, D. Rainham, M. S. Goldberg, and A. J. Wheeler. 2014. Using Global Positioning Systems (GPS) and temperature data to generate time-activity classifications for estimating personal exposure in air monitoring studies: an automated method. *Environ Health* 13 (1):33.

Nethery, E., K. Teschke, and M. Brauer. 2008. Predicting personal exposure of pregnant women to traffic-related air pollutants. *Sci Total Environ* 395 (1):11-22.

Nieuwenhuijsen, M. J., D. Donaire-Gonzalez, I. Rivas, M. de Castro, M. Cirach, G. Hoek, E. Seto, M. Jerrett, and J. Sunyer. 2015. Variability in and agreement between modeled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. *Environ Sci Technol* 49 (5):2977-82.

Nyhan, M., A. McNabola, and B. Misstear. 2014. Evaluating artificial neural networks for predicting minute ventilation and lung deposited dose in commuting cyclists. *JOURNAL OF TRANSPORT & HEALTH* 1 (4):305-315.

Ott, W. 1989. Human activity patterns: a review of the literature for estimating time spent indoors, outdoors, and in transit. In *Proceedins of the Research Planning Conference on Human Activity Patterns*, 3-1 to 3-38. Las Vegas, NV: EPA National Exposure Research Laboratory.

Ott, W. R. 1982. Concepts of human exposure to air pollution. *Environment International* 7 (3):179-196.

Ouidir, M., L. Giorgis-Allemand, S. Lyon-Caen, X. Morelli, C. Cracowski, S. Pontet, I. Pin, J. Lepeule, V. Siroux, and R. Slama. 2015. Estimation of exposure to atmospheric pollutants during pregnancy integrating space-time activity and indoor air levels: Does it make a difference? *Environment International* 84:161-73.

Panella, P., M. Casas, D. Donaire-Gonzalez, R. Garcia-Esteban, O. Robinson, A. Valentin, J. Gulliver, I. Momas, M. Nieuwenhuijsen, M. Vrijheid, and J. Sunyer. 2017. Ultrafine particles and black carbon personal exposures in asthmatic and non-asthmatic children at school age. *Indoor Air*.

Park, Y. M., and M. P. Kwan. 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* 43:85-94.

Phillips, M. L., T. A. Hall, N. A. Esmen, R. Lynch, and D. L. Johnson. 2001. Use of global positioning system technology to track subject's location during environmental exposure sampling. *J Expo Anal Environ Epidemiol* 11 (3):207-15.

Piechocki-Minguy, A., H. Plaisance, C. Schadkowski, I. Sagnier, J. Y. Saison, J. C. Galloo, and R. Guillermo. 2006. A case study of personal exposure to nitrogen dioxide using a new high sensitive diffusive sampler. *Sci Total Environ* 366 (1):55-64.

Pilla, F., and B. Broderick. 2015. A GIS model for personal exposure to PM10 for Dublin commuters. *Sustainable Cities and Society* 15:1-10.

Pratt, G. C., K. Parson, N. Shinoda, P. Lindgren, S. Dunlap, B. Yawn, P. Wollan, and J. Johnson. 2014. Quantifying traffic exposure. *J Expo Sci Environ Epidemiol* 24 (3):290-6.

Quackenboss, J. J., J. D. Spengler, M. S. Kanarek, R. Letz, and C. P. Duffy. 1986. Personal exposure to nitrogen dioxide: relationship to indoor/outdoor air quality and activity patterns. *Environ Sci Technol* 20 (8):775-83.

Rabinovitch, N., C. D. Adams, M. Strand, K. Koehler, and J. Volckens. 2016. Within-microenvironment exposure to particulate matter and health effects in children with asthma: a pilot study utilizing real-time personal monitoring with GPS interface. *Environ Health* 15 (1):96.

Rainham, D., I. McDowell, D. Krewski, and M. Sawada. 2010. Conceptualizing the healthscape: Contributions of time geography, location technologies and spatial ecology to place and health research. *Social Science & Medicine* 70 (5):668-676.

Reddy, S., M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. 2010. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks* 6 (2):1-27.

Robinson, J. P., and A. Silvers. 2000. Measuring potential exposure to environmental pollutants: time spent with soil and time spent outdoors. *J Expo Anal Environ Epidemiol* 10 (4):341-54.

Rojas-Bracho, L., H. H. Suh, P. Oyola, and P. Koutrakis. 2002. Measurements of children's exposures to particles and nitrogen dioxide in Santiago, Chile. *Sci Total Environ* 287 (3):249-64.

Ryan, P. H., S. Y. Son, C. Wolfe, J. Lockey, C. Brokamp, and G. LeMasters. 2015. A field application of a personal sensor for ultrafine particle exposure in children. *Sci Total Environ* 508:366-73.

Saksena, S., P. B. Singh, R. K. Prasad, R. Prasad, P. Malhotra, V. Joshi, and R. S. Patil. 2003. Exposure of infants to outdoor and indoor air pollution in low-income urban areas - a case study of Delhi. *J Expo Anal Environ Epidemiol* 13 (3):219-30.

Savitz, D. A. 2003. *Interpreting Epidemiologic Evidence: Strategy for Study Design and Analysis*: Oxford, UK: Oxford University Press.

Setton, E., J. D. Marshall, M. Brauer, K. R. Lundquist, P. Hystad, P. Keller, and D. Cloutier-Fisher. 2011. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *J Expo Sci Environ Epidemiol* 21 (1):42-8.

Setton, E. M., C. P. Keller, D. Cloutier-Fisher, and P. W. Hystad. 2008. Spatial variations in estimated chronic exposure to traffic-related air pollution in working populations: a simulation. *Int J Health Geogr* 7:39.

Shimada, Y., and Y. Matsuoka. 2011. Analysis of indoor PM2.5 exposure in Asian countries using time use survey. *Sci Total Environ* 409 (24):5243-52.

Sloan, C. D., T. J. Philipp, R. K. Bradshaw, S. Chronister, W. B. Barber, and J. D. Johnston. 2016. Applications of GPS-tracked personal and fixed-location PM(2.5) continuous exposure monitoring. *J Air Waste Manag Assoc* 66 (1):53-65.

Sloan, C. D., F. X. Weber, R. K. Bradshaw, T. J. Philipp, W. B. Barber, V. L. Palmer, R. J. Graul, S. C. Tuttle, R. T. Chartier, and J. D. Johnston. 2016. Elemental analysis of infant airborne particulate exposures. *J Expo Sci Environ Epidemiol*.

Sobus, J. R., Y. M. Tan, J. D. Pleil, and L. S. Sheldon. 2011. A biomonitoring framework to support exposure and risk assessments. *Sci Total Environ* 409 (22):4875-84.

Steinle, S., S. Reis, and C. E. Sabel. 2013. Quantifying human exposure to air pollution--moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Sci Total Environ* 443:184-93.

Steinle, S., S. Reis, C. E. Sabel, S. Semple, M. M. Twigg, C. F. Braban, S. R. Leeson, M. R. Heal, D. Harrison, C. Lin, and H. Wu. 2015. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. *Sci Total Environ* 508:383-94.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological)*:111-147.

Su, J. G., M. Jerrett, Y. Y. Meng, M. Pickett, and B. Ritz. 2015. Integrating smart-phone based momentary location tracking with fixed site air quality monitoring for personal exposure assessment. *Sci Total Environ* 506-507:518-26.

Tchepel, O., D. Dias, C. Costa, B. F. Santos, and J. P. Teixeira. 2014. Modeling of human exposure to benzene in urban environments. *J Toxicol Environ Health A* 77 (14-16):777-95.

Wheeler, A. J., L. A. Wallace, J. Kearney, K. Van Ryswyk, H. You, R. Kulka, J. R. Brook, and X. Xu. 2011. Personal, Indoor, and Outdoor Concentrations of Fine and Ultrafine Particles Using Continuous Monitors in Multiple Residences. *Aerosol Science and Technology* 45 (9):1078-1089.

Wiley, J. A., J. P. Robinson, Y. T. Cheng, T. Piazza, L. Stork, and K. Pladsen. 1991. Study of children's activity patterns: Final report. Sacramento, CA: California Air Resources Board.

Williams, R. D., and L. D. Knibbs. 2016. Daily personal exposure to black carbon: A pilot study. *Atmospheric Environment* 132:296-299.

World Health Organization. 2016. *Ambient air pollution: a global assessment of exposure and burden of disease*. Geneva.

———. 2017. *World Health Statistics 2017: Monitoring Health for the SDGs Sustainable Development Goals*.

Wu, J., C. Jiang, D. Houston, D. Baker, and R. Delfino. 2011. Automated time activity classification based on global positioning system (GPS) tracking data. *Environ Health* 10:101.

Wu, J., F. Lurmann, A. Winer, R. Lu, R. Turco, and T. Funk. 2005. Development of an individual exposure model for application to the Southern California children's health study. *Atmospheric Environment* 39 (2):259-273.

Wu, J., T. Tjoa, L. Li, G. Jaimes, and R. J. Delfino. 2012. Modeling personal particle-bound polycyclic aromatic hydrocarbon (pb-pah) exposure in human subjects in Southern California. *Environ Health* 11:47.

Yoo, E., C. Rudra, M. Glasgow, and L. Mu. 2015. Geospatial Estimation of Individual Exposure to Air Pollutants: Moving from Static Monitoring to Activity-Based Dynamic Exposure Assessment. *Annals of the Association of American Geographers* 105 (5):915-926.

Yuan, Y., M. Raubal, and Y. Liu. 2012. Correlating mobile phone usage and travel behavior – A case study of Harbin, China. *Computers, Environment and Urban Systems* 36 (2):118-130.

Zhang, K., and S. A. Batterman. 2009. Time allocation shifts and pollutant exposure due to traffic congestion: an analysis using the national human activity pattern survey. *Sci Total Environ* 407 (21):5493-500.

Zhang, Q. L., Z. M. Qiu, K. F. Chung, and S. K. Huang. 2015. Link between environmental air pollution and allergic asthma: East meets West. *Journal of Thoracic Disease* 7 (1):14-22.

Zheng, Y., H. Fu, X. Xie, W.-Y. Ma, and Q. Li. 2011. Geolife GPS trajectory dataset - User Guide.

Zhu, S. J., J. D. Marshall, and D. Levinson. 2016. Population exposure to ultrafine particles: Size-resolved and real-time models for highways. *Transportation Research Part D-Transport and Environment* 49:323-336.

Zou, B., J. G. Wilson, F. B. Zhan, and Y. Zeng. 2009. Air pollution exposure assessment methods utilized in epidemiological studies. *J Environ Monit* 11 (3):475-90.