AIR POLLUTION AND HEALTH:

TOWARD IMPROVING THE SPATIAL DEFINITION OF EXPOSURE, SUSCEPTIBILITY AND RISK

Marie-Pierre Parenteau

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the PhD degree in Geography

Department of Geography
Faculty of Arts
University of Ottawa

ABSTRACT

The role of the spatial representation in the relation between chronic exposure to NO_2 and respiratory health outcomes is studied through a spatial approach encompassing three conceptual components: the geography of susceptibility, the geography of exposure and the geography of risk. A spatially explicit methodology that defined natural neighbourhoods for the city of Ottawa is presented; it became the geography of analysis in this research. A LUR model for Ottawa is developed to study the geography of exposure. Model sensitivity to the spatial representation of population showed that dasymetric population mapping did not provide significant improvements to the LUR model over population at the dissemination block level. However, both the former were significantly better than population represented at the dissemination area. Spatial representation in the geography of exposure was also evaluated by comparing four kriging and cokriging interpolation models to the LUR. Geostatistically derived NO_2 concentration maps were weakly correlated with LUR model results. The relationship between mean NO_2 concentrations and respiratory health outcomes was assessed within the natural neighbourhoods. We find a statistically significant association between NO_2 concentrations and respiratory health outcomes as measured by global bivariate Moran's I. However, for regression model building, NO_2 had to be forced into the model, demonstrating that NO_2 is not one of the main contributing variables to respiratory health outcomes in Ottawa. The results point toward the importance of the socioeconomic status on the health condition of individuals. Finally, the role of spatial representation was assessed using three different spatial structures, which also permitted to better understand the role of the modifiable areal unit problem (MAUP) in the study of the relationship between exposure to NO_2 and health. The results confirm that NO_2 concentration is not a major contributing factor to the respiratory health in Ottawa but clearly demonstrate the implications that the use of opportunistic administrative boundaries can have on results of exposure studies. The effects of the MAUP, the scale effect and the zoning effect, were observed indicating that a spatial structure that embodies the scale of major social processes behind the health condition of individuals should be used when possible.

RÉSUMÉ

Le rôle de la représentation spatiale dans la relation entre l'exposition chronique au NO_2 et l'état de santé respiratoire est étudié à partir d'une approche spatiale comprenant trois composantes conceptuelles: la géographie de la vulnérabilité, la géographie de l'exposition et la géographie du risque. Une méthodologie spatialement explicite pour définir les quartiers naturels pour la ville d'Ottawa est présentée, ces derniers représentant la géographie d'analyse de cette étude. Le modèle de régression LUR a été développé pour Ottawa et portait sur la géographie de l'exposition. Un modèle de sensibilité de la représentation spatiale de la population a démontré que le dasymetric population mapping n'offrait aucune amélioration significative au modèle de régression LUR de la population au niveau de l'îlot de diffusion. Cependant ces deux derniers sont significativement meilleurs que la population représentée au niveau de l'aire de diffusion. Le rôle de la représentation spatiale dans la géographie de l'exposition a également été évalué en comparant les résultats pour quatre modèles d'interpolation de krigeage et de cokrigeage et le modèle LUR. La cartographie de la concentration de NO_2 obtenue d'une façon statistique est corrélée faiblement aux résultats du modèle LUR. La relation entre les concentrations moyennes de NO_2 et les conditions de santé respiratoire a été évaluée en utilisant la structure des quartiers naturels. Nous avons trouvé une association statistique significative entre la concentration de NO_2 et les conditions de santé respiratoire telle que mesurée par la *global bivariate Moran's I*. Cependant, pour l'élaboration du modèle de régression, la variable NO_2 a été imposée de force dans le modèle, indiquant que le NO_2 n'est pas un des

principaux facteurs contribuant à la santé respiratoire à Ottawa. Les résultats font également ressortir l'importance du statut socio-économique sur l'état de santé des individus. Finalement le rôle de la représentation spatiale a été évalué à partir de trois différentes structures spatiales, permettant ainsi de mieux comprendre le rôle de MAUP (Problèmes des unités spatiales modifiables) dans l'étude de la relation entre l'exposition au NO_2 et la santé. Les résultats obtenus corroborent que la concentration de NO_2 ne constitue pas un facteur qui contribue grandement à la santé respiratoire à Ottawa mais ils démontrent clairement les implications que l'utilisation de limites administratives opportunistes peuvent avoir sur les études d'exposition. Les effets du MAUP, l'effet d'échelle et l'effet de zone, ont été observés indiquant qu'une structure spatiale qui exprime les processus sociaux majeurs sous-tendant l'état de santé des individus devrait être utilisée lorsque cela est possible.

ACKNOWLEDGEMENTS

C'est avec bonheur que je tiens à remercier toutes les personnes qui m'ont permis de compléter cette dissertation.

This thesis would not have been possible without the support and dedication of Dr. Sawada, my Ph.D. supervisor. All along this journey, he provided encouragement, guidance and sound advices. Thank you for your support, both in my professional and my academic ventures.

J'aimerais souligner la participation et l'aide de Natividad Urquizo, de la ville d'Ottawa et de mes collègues du « Ottawa Neighbourhood Study ». Merci d'avoir collaboré à ce projet de recherche.

Je tiens aussi à remercier mes amis. Votre support moral a grandement contribué à la réalisation de cette thèse.

Un gros merci à mes parents qui ont été présents à chaque moment et qui m'ont encouragé à poursuivre la route.

… and finally my *rude bwoy, tanks man* for supporting me in all my endeavors. It's been a long road but the future looks *irie*.

"Fall seven times, stand up eight".

-*Japanese Proverb*

PREFACE

I am the first author on all the chapters/publications included in this thesis but I collaborated with other researchers from the University of Ottawa, in particular members of the Ottawa Neighbourhood Study. Chapters two and three have been published in peer-reviewed journals while chapters four to six are in preparation or have been submitted for publication.

CHAPTER 2

Parenteau, M.-P., Sawada, M., Krisjansson, E.A., Calhoun, M., Leclair, S., Labonte, R., Runnels, V., Musiol, A. & Herold S. (2008). Development of neighbourhoods to measure spatial indicators of health. *Journal of the Urban and Regional Information Systems Association*, *20*(2), 43-55.

In this chapter the methodology for the delineation of natural neighbourhood is explained. The geographic units delineated representing the spatial definition of susceptibility are used in this thesis but are also the geographical foundation of the Ottawa Neighbourhood Study. Key to this chapter is the work of Dr. Sawada and Dr. Krisjansson, co-main investigators of the Ottawa Neighbourhood Study, on the development of health indicators and most importantly on the creation of the Ottawa Neighbourhood Study itself.

CHAPTER 3

Parenteau, M.-P., & Sawada, M.C. (2010). The role of spatial representation in the development of a lur model for ottawa, canada. *Air quality, atmosphere and health*, *In press*, 1-13.

In this chapter are presented the results of a land-use regression (LUR) model developed for the city of Ottawa. The role of spatial representation in the geography of exposure is here assessed through different operationalizations of the population variable. A dasymetric mapping technique was used for the first time in a LUR model.

CHAPTER 4

Parenteau, M.-P., and M. Sawada. in preparation. Comparison of methodological approaches for the modeling of NO_2 in Ottawa, Canada. *In preparation*.

This chapter presents the results of different methodological approaches to the modeling of NO_2 in Ottawa and compares the NO_2 surfaces generated. Spatial representation is here studied in the context of exposure models.

CHAPTER 5

Parenteau, M.-P., and M. Sawada, in preparation. A Multi-scale analysis of the relationship between exposure to NO_2 and respiratory health. *In preparation*.

This chapter presents the results of the study of the relationship between exposure to NO_2 and health using different geographic structures. The emphasis is on the role of the MAUP in the assessment of this particular relationship. It is formulated by the use of three geographic structures: the census tract, the natural neighbourhood and the cluster structure.

CHAPTER 6

Parenteau, M.-P., and M. Sawada, in preparation. A spatial analysis of long-term exposure to NO_2 and respiratory morbidity in Ottawa, Canada. *In preparation*.

This chapter presents the results of the relationship between exposure to NO_2 and respiratory health using a spatial approach. The geography of risk is represented by the overlap between the geography of exposure, as defined in chapter three, and the geography of susceptibility as captured by the neighbourhoods delineated in chapter two.

List of Acronyms

AIC	Akaike information criteria
AR	Area ratio
CCHS	Canadian community health survey
CDA	Confirmatory data analysis
CI	Condition index
CMA	Census metropolitan area
CO	Carbon monoxide
COPD	Chronic obstructive pulmonary disease
CSDA	Confirmatory spatial data analysis
CT	Census tract
DA	Data augmentation
DA	Dissemination area
DAD	Discharge abstract database
DASYM	Dasymetric mapping
DISB	Dissemination block
EPOI	Enhanced points of interest
ESDA	Exploratory spatial data analysis
GEE	Generalized estimating equation
GIS	Geographic information system
GLM	Generalized linear model
GWR	Geographically weighted regression
ICD	International classification of diseases
LICO	Low income cutoff
LISA	Local indicator of spatial autocorrelation
LOOCV	Leave-one-out cross validation
LUR	Land-use regression
MAUP	Modifiable area unit problem

MLS	Multiple listing service
MOHLTC	Ministry of health and long-term care
NACRS	National ambulatory care reporting system
NAICS	North American industry classification system
NO_2	Nitrogen dioxide
NOx	Nitrogen oxide
O_3	Ozone
OLS	Ordinary least squares
OPH	Ottawa Public Health
PDF	Population density fraction
RMSE	Root-mean-square error
SES	Socioeconomic status
SIC	Standard industrial classification
SO_2	Sulfur dioxide
TF	Total fraction
TRAP	Traffic-related air pollution
VIF	Variation inflation factor

Table of Contents

List of Tables

List of Figures

Chapter 1 Exposure to air pollution and health impact: literature review, data and conceptual framework

1.1 Introduction

The literature on the relationship between health and exposure to air pollutants has grown considerably over the last decade (Brunekreef and Holgate 2002), with a large number of studies published in 2010 (Gurjar et al. 2010; Mölter et al. 2010; Pan et al. 2010; Patel et al. 2010; Guay et al. 2010; McConnell et al. 2010; Setton et al. 2010; Akinbami et al. 2010; Gehring et al. 2010; Miller et al. 2010; Grineski et al. 2010; Neupane et al. 2010; Szyszkowicz and Rowe 2010). However there is no consensus as to the degree to which long-term or short-term exposure to nitrogen dioxide ($NO_{2)}$ has a significant impact on human health within the context of an urban environment. One could hypothesize that the lack of agreement in the findings of health/NO_2 research is a consequence of the problem of spatial representation, which includes the MAUP but is not limited to it. Many have studied the relationship between exposure to NO_2 and respiratory health but the role of spatial representation on the analytical results has been neglected. The primary objective of this research is to bring attention to the different aspects of spatial representation and better understand its implications in health studies.

Spatial representation can be defined as how the spatial concepts of a research question are conceptualized and operationalized, which is closely linked to the MAUP (e.g. scale and zoning) but is not limited to it. For example, the use of different exposure models to map NO_2 concentrations can impact the spatial representation of NO_2 concentrations used in the assessment of the risk of exposure. These differences are the results of the unique characteristics that can be associated with each type of model. Considering the potential impact of spatial representation, including scale, on analytical results, too little research has explicitly accounted for this issue when studying the relationship between exposure to air pollutants and health. In a research preoccupied by the impact of exposure to air pollutants such as NO_2 on health, spatial representation is potentially a significant factor. The type of models used to generate maps of exposure from measurements as well as the conceptualization and operationalization of the variables used in the models are examples of variations that could impact the robustness of results and conclusions. In this research we refer to that as the role of spatial representation in the geography of exposure. The second level at which spatial representation can influence the results of health/NO_2 study pertains to the geography of susceptibility. Therein, the focus is on the impact of using different geographical delineations to capture social processes that could represent susceptibility. This specific question has not been addressed adequately in the literature, a complaint also voiced by the environmental justice community (Higgs and Langford 2009). This component of the research will make an important contribution to the discussions on the concept of neighbourhood. The last component of the research is the geography of risk, which is affected by spatial representation through the geography of exposure and the geography of susceptibility. The conceptual framework of this research is presented in more detail in section 1.5.

1.2 Research Objectives

The main objective of this research is to examine in detail the role of spatial representation in the study of the relationship between health and exposure to air pollutants. To meet the *raison d'être* of this research, the work performed is grouped into three domains (the geography of exposure, the geography of susceptibility and the geography of risk) which may be addressed by the following research questions:

- How can we operationalize pertinent study areas for research on the relationship between exposure to air pollution and health?
- How do different operationalizations of the variables modify the results of the modeling of exposure to NO_2?
- Do advanced interpolation methods such as the dasymetric mapping improve the results of the modeling of exposure to NO_2?
- What are the impacts of using different models on the results of exposure to NO_2?
- Does the use of statistical methods designed for the analysis of spatial data impact the strength of the relationship measured versus the use of traditional statistical methods?
- What are the repercussions of using varied spatial structures (the effect of the MAUP) on the analytical results?

1.3 Study Area

The city of Ottawa is located within the Quebec-Windsor corridor, a geographic area characterized by high NOx emissions (Figure 1.1). Transportation, the main contributor to air pollution in Ottawa, accounts for approximately 85% of NO_x

emissions (City of Ottawa, 2004). The air quality in Ottawa is considered relatively good with an annual mean NO_2 concentration well below the 24h 100 ppb standard set by the Province of Ontario's provincial ambient air quality criteria (Ontario Ministry of the Environment, 2008). Ottawa also has an air quality that is above the Canadian Standards (City of Ottawa, 2010a). The national ambient air quality objectives and guidelines for Canada are shown in Table 1.1.

Part of the Ottawa-Gatineau census metropolitan area (CMA; Ontario part), the city of Ottawa had a population of 812,129 in 2006 (Statistics Canada 2007a). Ottawa is characterized by a generally well-educated population with a higher median family income than the provincial median (Statistics Canada 2007a). Urban sprawl, known to have a negative effect on air quality by increasing traffic (Jerrett et al. 2009), has been recognized as a problem by the municipal government of Ottawa and for this particular reason city officials have adopted a new policy of not increasing the limits of the urban areas. Nevertheless, Ottawa is a relatively low density city - the population density has dropped from 57 people/ha in 1925 to approximately 28 people/ha in the 1980's and has remained constant (City of Ottawa 2010 b).

The City of Ottawa is an interesting area for research on the relationship between exposure to NO_2 and respiratory health. To our knowledge, this research is the first conducted in Ottawa for the study of this relationship at the intra-urban level. It is important to note that the land-use regression model developed is also the first one built for Ottawa. While other studies of this type have been conducted in a few other Canadian cities, Ottawa differs from the latter by its small industrial sector, its generally good air quality and its smaller population. Finally by conducting this research in Ottawa, it will be possible to share the findings with the Ottawa Neighbourhood Study (ONS). The ONS is a study group based at the Institute of Population Health at the University of Ottawa and aims at developing a better understanding of the physical and social pathways through which neighbourhoods affect health.

While the Ottawa-Gatineau census metropolitan area (CMA) crosses provincial boundaries and includes a few municipalities in the province of Quebec, the scope of this research is limited to the City of Ottawa. The Quebec section was excluded from this research because of issues related to collection of data on air pollutants within the different national, regional, and local administrative structures.

1.4 NO_2 and health

According to Environment Canada (Environment Canada 2009a), the natural characteristics of the atmosphere may be changed in reaction to any chemical, physical, or biological agent and is referred to as air pollution. NO_2, which is the pollutant of interest in this study, is a criteria contaminant, meaning that it is a pollutant for which acceptable thresholds of exposure are defined and so acts as a quality standard (California Environmental Protection Agency 2010). NO_2 is strongly associated with vehicle emissions (Figure 1.2) (Setton et al. 2010; Miller et al. 2010).

Significant effects of exposure to air pollutants on health have been demonstrated at levels that are below the acceptable thresholds (O'Connor et al. 2008), as well as when exposure levels exceed established thresholds (Environment Canada 2010; Wilson et al. 2004). Many authors have confirmed the effects of exposure to air pollution on health using measures of hospitalization and mortality resulting from diseases of the respiratory and circulatory systems (Brook et al. 2004; Jerrett et al. 2004; Gulliver and Briggs 2005; Holguin et al. 2007; Gauderman et al. 2007; Kim et al. 2007; Brook et al. 2007; Brook et al. 2008; Gehring et al. 2010). Clinical studies have shown that exposure to increased levels of NO_2 augments airway inflammation (Peden 1999; Villeneuve et al. 2006; Schikowski et al. 2007; Lefrance et al. 2009; Akinbami et al. 2010). In the case of short-term exposure to NO_2, this inflammation may precipitate asthma attacks and long-term exposure has been associated to chronic obstructive pulmonary disease (COPD), cardiopulmonary mortality/morbidity, and respiratory mortality/morbidity (Brauer et al. 2002; Luginaah et al. 2005; Arain et al. 2009; Burra et al. 2009; Gurjar et al. 2010; Grineski et al. 2010; Pan et al. 2010; Thach et al. 2010; Clark et al. 2010).

Many researchers have demonstrated the existence of a significant relationship between exposure to NO_2 and acute and chronic respiratory disease (Brook et al. 2004; Jerrett et al. 2004; Gulliver and Briggs 2005; Zanobetti and Schwartz 2006; Holguin et al. 2007; Gauderman et al. 2007; Kim et al. 2007; Brook et al. 2007; Brook et al. 2008; Jerrett et al. 2009; Lee and Ferguson 2009; Clougherty and Kubzansky 2009; Neupane et al. 2010). In Canada, following a time-series approach, Stieb et al. (2009) found a significant effect of air pollution on acute respiratory presentations through a large sample of emergency department visits and air quality data from multiple centers. The negative health effect following chronic exposure to high air pollutant concentrations has also been shown in European studies (Brauer et al. 2002; Sauerzapf et al. 2009; Lee and Ferguson 2009; Brunekreef 2007).

It has been demonstrated that subgroups of the population are more susceptible to increased NO_2 concentration exposure. Women, children, older adults and disadvantaged individuals have been identified as being more susceptible to the effect of NO_2 exposure (Villeneuve et al; 2007; Arbex et al. 2009; Sahsuvaroglu et al. 2009; Kan et al 2008; Neupane et al. 2010). For example, numerous studies found a statistically significant association between exposure to NO_2 and asthma in children (Brauer et al. 2002; McConnell et al. 2003; Clark et al. 2010; Grineski et al. 2010). The socioeconomic status (SES) of individuals has also been associated with a greater health effect of exposure to air pollution (Wheeler and Ben-Shlomo 2005; Health Canada 2010). This latter association could possibly be the result of an exposure to higher concentrations of air pollution by less affluent individuals but could also be the result of chronic health conditions that are linked to disadvantaged individuals (McLeod et al. 1998; Brainard et al. 2001; Forastiere et al 2007; Lee et al. 2006).

There is yet no positive and statistically significant association that has been unequivocally measured, although many studies have concluded that increased exposure to NO_2 is likely contributing to the negative health of individuals (Villeneuve et al. 2006; Linn and Gong 1999; Luginaah et al. 2005; Sahsuvaroglu et al. 2009; Clark et al. 2010; Barcelo et al. 2009; Akinbami et al. 2010). Thus, this research may contribute to the knowledge about health by determining if spatial representation could be a key factor in measuring effects. The exposure period (acute versus chronic), the variations in the study design (cross-sectional and time-series versus longitudinal) and the geography (local versus larger regions) are other potential factors responsible for divergent findings among researchers (Akinbami et al. 2010).

1.5 Conceptual Framework

This research is structured according to Jerrett and Finkelstein's (2005) proposed conceptual framework for the study of the relationship between human health and exposure to pollutants in the context of environmental justice (Figure 1.3). The framework defined by these authors is based on the work of Mayer (1983). In his article "The Role of Spatial Analysis and Geographic Data in the Detection of Disease Causation", Mayer described factors that can create different spatial patterns in the distribution of diseases: 1) physical-environmental factors; 2) social, economic and cultural factors; 3) genetic factors. Jerrett and Finkelstein (2005) translated the work of Mayer "into an operational framework that includes three underlying geographies: exposure, susceptibility, and adaptation. […] Areas where two or more of the circles in the Venn diagram overlap are termed geographies of risk" (Jerrett et al. 2010). In this dissertation, the geography of risk is defined as the overlap between exposure and susceptibility, excluding the geography of adaptation.

Jerrett and Finkelstein's framework (2005) is the base of this research as it is specifically designed to study the relationship between human health and exposure to pollutants. However, their proposed framework is an adaptation of the more common risk analysis framework where the risk is the product of the exposure and the susceptibility (Makri and Stilianakis 2008). In this research, the geography of exposure is represented by the exposure to NO_2 concentrations while the susceptibility is assessed through the spatial distribution of characteristics considered to be capable of modifying the level of risks to which humans are subjected. These susceptibilities can affect the level of risk through the level of exposure and the capacity to cope with the risk (Makri and Stilianakis 2008).

While this research applies an already existing conceptual framework, its focus remains on the question of spatial representation in the study of the relationship between exposure to air pollution and respiratory health. As such, the impact of alternative spatial representations in the geography of susceptibility and the geography of exposure are first assessed individually. In the following steps, the role of spatial representation in the study of the geography of the risk, which is the result of the overlap between the exposure and the susceptibility, is assessed.

The structure of the thesis follows the conceptual framework described in the next sub-sections. Different chapters of this research address specific elements of the conceptual framework, which is explained in more details in section 1.7.

1.5.1 The Geography of Exposure

Many elements can affect the geography of exposure when assessing the role of air pollution on health. A first element to consider is the scale at which the concentration of the pollutant of interest varies across space. In the case of NO_2, Hewitt (1991) reported that it could significantly vary within a distance as small as 50 metres. As a consequence, it is possible to conduct research on the relationship between exposure to NO_2 and respiratory health at a local scale. In the case of long-term exposure studies, the temporal component is very important. Typically data are collected by a combination of passive samplers and permanent stations over periods of two weeks (Miller et al. 2010). Data from multiple deployments are combined and an average concentration at each sampling location is calculated (Brauer et al. 2002; Henderson et al. 2007; Wheeler et al. 2008; Jerrett et al. 2009; Wilton et al. 2010). In geographic regions like Canada, where seasonality is an important factor, the spatial patterns have been shown to be relatively stable throughout the year, but concentrations vary from one season to the next (Wheeler et al. 2008).

The types of approaches for NO_2 modeling at the intraurban level can be classified into six categories (Jerrett et al. 2005a). In this research two of them, land-use regression model (LUR) and geostatistical interpolation, are reviewed and implemented. Although the LUR model has been widely used for the modeling of pollutants such as $PM_{2.5}$ and benzene, its principal application in health is for the modeling of NO_2 (Wheeler et al. 2008; Hoek et al. 2008; Atari and Luginaah 2009; Mukerjee et al. 2009; Gehring et al. 2010; Mölter et al. 2010). The LUR approach is based on two important assumptions: “1) environmental conditions for the variable of interest can be estimated from a small number of readily measurable predictor variables 2) the relation between the target variable and these predictors can be reliably assessed on the basis of a sample survey or training area” (Briggs et al. 2000). The LUR model allows modeling of local variations of a pollutant concentration more easily than with the statistical interpolation (Jerrett et al. 2007; Arain et al. 2009), using predictors such as land-use, the road network and population (among others) (Moore et al. 2007; Ryan and LeMasters 2007). In the case of NO_2 modeling, numerous variables describing land-use and other components of the environment located within buffer zones delineated around a sampling location correspond to the independent variables, while NO_2 is the dependent variable (Brauer et al. 2002). LUR is an effective and efficient approach to mapping air pollution (Su et al. 2009). LUR modeling has been used to create NO_2 surfaces with local area variability for large population concentration centers in Europe, North America and most recently Asia (Jerrett et al. 2005b; Jerrett et al. 2007; Marshall and al. 2008; Su et al. 2008; Wheeler et al. 2008; Atari and Luginaah 2009; Sahsuvaroglu et al. 2009; Su and al. 2009b; Crouse et al. 2009; Kashima et al. 2009; Clougherty et al. 2009; Larson et al. 2009).

The second modeling approach used in this research is the geostatistical interpolation. Geostatistical interpolation relies on air pollution data obtained throughout the study area in order to design concentration surfaces showing the continuous concentration of a pollutant in space. It has been used many times but it is now not as popular as the LUR (Mölter et al. 2010). The superiority of this approach resides in the fact that the data collected at the sampling sites are the only data required. This approach is very simple, however simplicity is its principal weakness: it makes unsophisticated assumptions about the spatial distribution of pollution (Briggs 2005). The geostatistical interpolation techniques used in this research are the kriging techniques and cokriging techniques, known as stochastic methods. These methods exploit spatial dependency, which is composed of two parts: first order effects (general trend) and second order effects (local trends) (Finkelstein et al. 2003; Jerrett et al. 2005a). Kriging methods are the most favorable because they provide the best linear unbiased estimate at any point in the study area (Jerrett et al. 2003). Because standard kriging methods do not use covariates or utilize a limited number of those in the modeling of air pollution they over-smooth concentration surfaces (Luginaah et al. 2006). Kriging can only account for variation in the pollutants as they are measured at the sampling locations (Beelen et al. 2009).

The cokriging technique presents a more advanced version of the algorithm than standard kriging for the modeling of NO_2. Cokriging uses covariates to direct the interpolation (Briggs 2005). This method can be utilized when auxiliary variables are associated to the pollutant. Auxiliary variables such as population density or traffic data are integrated in the modeling to allow the production of an improved version of the modeling of the pollutant resulting from the additional information offered by the auxiliary variable. Cokriging combines the spatial behavior of the auxiliary variables and the pollutant as a result of the cross-correlation between them. Although cokriging offers an acceptable approach to NO_2 modeling, there is little evidence that it has been used so far for this purpose with the exception of studies by Gallois et al. (2005), de Fouquet et al. (2007), Cardenas and Perdrix 2008 and Malherbe et al. (2008).

It is believed that unreliable or inaccurate modeling of NO_2 exposure could be a contributing factor to the lack of agreement in studies on the relationship between long-term exposure and health (Briggs et al. 2000). Another element related to the geography of exposure that could be behind the non-conformity to the general understanding are daily mobility patterns of individuals (Gehring et al. 2010; Setton et al. 2010; Mölter et al. 2010). Unfortunately this is not taken into account in this research due to a lack of data about commuting and because it is not one of the objectives of this research.

1.5.2 The Geography of Susceptibility

The conceptualization and operationalization of space is one element of public health studies in general, but more specifically of air pollution studies, that has been neglected (Stafford et al. 2008; Matthews 2008). Unfortunately, this weakness is crucial when it comes to assessing the role of susceptibility in the relationship between exposure to air pollutants and health (Jerrett and Finkelstein 2005). Numerous factors, such as socioeconomic status, are known to be potential confounders or modifying variables in the study of the relationship between exposure to NO_2 and health (Gouveia and Fletcher 2000; Finkelstein et al. 2005; Barcelo et al. 2009).

A spatial structure that does not capture underlying social processes may not depict the correct spatial scale of representation at which susceptibility is found in the relation between exposure to air pollutants and health (Stafford et al. 2008). These susceptibilities, that can confound or modify the relationship, are known as contextual and compositional effects in the field of population health. The compositional explanation implies that the characteristics of individuals living in the area are solely responsible for geographical differences in health outcomes. The spatial differences in health are also influenced by the environment where individuals live, also referred to as contextual effects. Contextual effects are the result of exposure to features and characteristics of the area where individuals live (Cummins et al. 2005; Jerrett and Finkelstein 2005; Chaix et al. 2008; Dragano et al. 2009; Chaix et al. 2010). Depending on the size of the geographical unit supporting the research, the effect of these variables on health or as modifiers of the air pollution effect may differ (Jerrett and Finkelstein 2005).

The problems associated with the geography of susceptibility are closely linked to the MAUP (O'Neill et al. 2003), a term coined by Openshaw (1984).

Numerous authors (Arbia 1989; Jelinski and Wu, 1996; Brindley et al. 2005) have studied the effects of the MAUP, which can be divided into the scale and the zoning effects. Scale effects express the fact that when studying a relationship, the scale at which it is conducted will impact the analytical results. The zoning effect occurs when given the same scale, different zoning systems (with the same number of units) do not provide the same results. More than 20 years ago Mayer (1983) stated that the scale effect was not considered in epidemiology and geographic pathology; today this is still a weakness of health studies. This research can make an important contribution to knowledge by looking at the effect of scale variability and the use of different zoning systems on analytical results.

Small-area health studies that use compositional and contextual data in the analysis have in most cases operationalized their geography at the census tract level. These statistical areas are used by Statistics Canada for the dissemination of census data on a large number of socioeconomic variables (Gauvin et al. 2007). Although the census tracts have been used as a proxy for the neighbourhood on numerous occasions, they were not specifically delineated with the objective of mimicking the subjective meaning of neighbourhoods held by many people (Gauvin et al. 2007) and they are not stable over time (Cummins et al. 2005; Dragano et al. 2009; Higgs and Langford 2009). Automated zone design techniques represent a solution to this problem. Using a set of constraints, automated delineation aggregates basic units together to form a custom geography (Cockings and Martin 2005). This approach can be used to delineate spatial structures that are optimal, for example, in terms of maximizing the homogeneity of certain socioeconomic variables (Gauvin et al. 2007). Since social stressors may be aggravating health effects caused by exposure to air pollution, the use of a geography that maximizes internal homogeneity could allow better identifying of communities more at risk (Clougherty and Kubzansky 2009).

According to some authors (Willis et al. 2003; Jerrett and Finkelstein 2005), new spatial units or zones that more correctly exemplify the underlying geographic susceptibility can emerge from air pollution studies. In order to evaluate how the MAUP can affect the results, researchers should at least run tests at multiple spatial scales. Among the few studies that have used a geographic structure other than statistical or administrative boundaries, some have reported no effect of using alternative spatial representations (Stafford et al. 2008; Ross et al. 2004) while others have reported some effects (Gauvin et al. 2007). Little work has been done in this field and as such this research may confirm that automated zone design techniques are a solution for small-area health studies.

1.5.2.1 The Neighbourhood

The neighbourhood is central to this research but the debate surrounding its conceptual definition, which could be the subject of a dissertation by itself, will not be addressed. The following presents a brief overview of approaches to the conceptualization of the neighbourhood.

The method used to define the neighbourhood is an outcome of the epistemological approach to which a researcher subscribes (Lebel et al. 2007). The neighbourhood may adhere to a definition based on the social character or on the functional character of the concept (Gagnon 1994; Galster 2001; Gauvin et al. 2007) According to the latter, it is often assumed that the neighbourhood corresponds to statistical units or administrative units (Ellen and Turner 1997), representing more or less those small local spatial regions to which individuals self-identify (Germain et Charbonneau 1994; Germain et Gagnon 1999; Clapp and Wang 2006; Benigeri, 2007; Gauvin et al. 2007). These spatial units are well defined and allow the integration of socioeconomic data in research (Coulton et al. 2001). Conversely, the social definition of neighbourhoood is focused on the relationships between residents of the neighbourhoods. For example, sociologists are interested in the social ties between individuals (frequency and strength of contacts) and questions of social isolation (McPherson et al. 2006; Ross et al. 2000; Ross and Jang 2000). The sociological approach is also interested in the perception residents have of their own neighbourhood. For the sociologist, the neighbourhood is perceived and studied as a community. For the geographer, under the functional construct, the neighbourhood represents a physical place; it is a geographic location where individuals have their residence (Germain Charbonneau, and Gagnon 1998).

In Canada and United States, the census tract is often referred to as a representation of the neighbourhood (Dubin, 1993; Statistics Canada, 2003; Lebel et al. 2005). On the other hand, it has been demonstrated that these units created by national statistical agencies for carrying out a census do not mirror the underlying social boundaries and as such may depict the artifacts of administrative rules or of a putative system (Martin, 2004). Therefore it is sometimes difficult to determine if the results of the analysis are representative of the reality or if they are the results of using a certain type of geographical unit (Mennis, 2003). To avoid this issue, a set of natural neighbourhoods are delineated in the context of this research. Borrowing from the functional approach, these units are delineated according to the spatial distribution of processes/characteristics of the environment and of the population that are able to influence the response variable, in this case the health of the population. These natural units are not intended to be aligned with the sociological definition of the concept. The natural neighbourhoods delineated in this research are units that maximize internal homogeneity while maximizing external heterogeneity for the variables used in the delineation process. As such, these neighbourhoods can be considered homogeneous in terms of these variables. Again this differs from the sociological approach where the focus is on the growing heterogeneity of the neighbourhood as they become more socially diverse (for example, in terms of multi-ethnic neighbourhoods) (Fong and Gulia 2000; Ray and Preston 2009).

1.5.3 The Geography of Risk

The geography of risk represents the overlap between the geography of exposure and the geography of susceptibility. Because spatial processes can occur at different scales and because data are at times compiled under different spatial structures, there is potential for a spatial mismatch (Jerrett and Finkelstein 2005). In studies preoccupied by the relationship between health and exposure to air pollutants, there is a possibility of encountering a spatial mismatch since the scale of variability of air pollutants such as NO_2 can vary greatly from the processes confounding the relationship under study. This issue may impede research based on direct observation (Jerrett and Finkelstein 2005).

Whenever there is potential for spatial mismatch, it is necessary to have a good comprehension of the spatial processes that beget geographic susceptibility and exposure, permitting the inclusion of the comprehension of the phenomena into the study design. Hence using a geographic structure that is optimal in terms of homogeneity of the contextual variables may reduce measurement error of exposure (Riva et al. 2009).

1.5.3.1 Population health

The factors explaining heterogeneities in health are many and not limited to the exposure to pollutants such as NO_2. Hertzman et al. (1994) have proposed a framework to explain differences in the health status of individuals. According to the latter, the population can be stratified based on its characteristics, with SES being one that has consistently indicated differences in the health status of individuals (Hertzman et al. 1994; Adler and Newman 2002; Chen et al. 2009). Socioeconomic status is a multi-faceted concept that can be measured through education, income and occupation (Braveman et al. 2005; Adler and Newman 2002).

The pathways behind the association between SES and health are still not well understood even though many studies have reported an association between the two (Subramanian et al. 2011; Eisner et al. 2011; Salm Ward et al. 2010; Minet Kinge and Morris 2010; Lorant et al. 2003; Lynch et al. 1997; Ross and Wu 1996; House et al. 1994; Winkleby et al. 1992). Adler and Newman (2002) suggest that the "pathways by which socioeconomic status influcences health should be those that affect health more generally". According to these authors, SES can generate heath problems through behaviour/lifestyle, environmental exposure and health care. This idea is expressed in Figure 1.4, which shows how the pathway through which SES influences health is by exposure to different environments and how individuals adapt to these environments. Even though the latter is complex and not well understood, it is important to consider it in the study of the relationship between exposure to NO_2 and respiratory health as it can be a confounding factor.

1.6 Study Design and Data

Data used in this research were gathered from numerous sources including municipal, provincial and federal governments were involved in the data acquisition process. The different datasets used to conduct this research will now be described and preprocessing, when applicable, will be explained.

1.6.1. Study Design

In this research, the relationship between chronic respiratory health problems and long-term exposure to NO_2, where exposure is equal to the location of the residence, is the fundamental element. The study design of this research guided the statistical methods used for the analysis (Dominici et al. 2003). In the case of studies on long-term exposure using the group level (ecological study) based on small-areas, “the annual numbers of health events in each area are being regressed against average pollution concentrations from preceding years” (Lee and Ferguson 2009). An ecological study has to be conducted since no data on personal exposure are available. As per previous research on the effect of NO_2 on health, compositional and contextual variables were taken into account in the analysis (Dominici et al. 2003; Willis et al. 2003). Ecological confounders included in this research were contextual risk factors for health related to the social environment (other contextual risk factors such as access to healthcare were not included).

1.6.2 Air Quality Data

In the fall of 2007, the National Capital Air Quality Mapping Project was launched by the City of Ottawa with the help of Environment Canada and Health Canada. A total of 30 Ogawa samplers were installed throughout the city of Ottawa. In order to have a better idea of the concentration of pollutants in the area at different periods, the samplers were placed at the same location on three occasions, each time for a period of two weeks. Data collected by 29 of 30 samplers over the period of 29 September – 13 October 2007 are used for this research. The monitor AQMP-09, located in a quiet residential area in Gatineau, Quebec, was not taken into consideration. There was no loss of key information as the samplers AQMP-17 and AQMP-26 were located in a very similar environment as AQMP-09. Data were also collected from two permanent stations.

The samplers and permanent stations measured the concentration in SO_2, O_3, and NO_2 at each location. This research focuses only on NO_2 modeling as it is an essential component of traffic-related air pollution. Through the literature it has been clearly demonstrated that NO_2 has a strong relation with common air quality indices (Piro et al. 2008; Crouse et al. 2009). The NO_2 has been used as a proxy to evaluate exposure to traffic emissions (Jerrett et al. 2007).

For the preprocessing of the data from the two permanent stations, the Data Augmentation (DA) algorithm (Tanner and Wong 1987) in the S-Plus Missing Data Library was used. This model-based approach is made up of two steps that are applied alternatively until the algorithm converges. First the imputation step simulates the missing data based on the observed data and the current parameter estimates. The results of the imputation are then fed into the posterior step where a new set of parameters are calculated (Lanza et al. 2005). The results of the DA algorithm were analyzed using the time-series plots and autocorrelation function plots (Schafer 1997; Lanza et al. 2005). Following the analysis of the plots the mean value at each permanent station was computed. These data were used as the mean NO_2 value for the two-week period.

1.6.3 Health Data

Because this research was using sensitive information collected by different organizations, the research proposal was reviewed by the Research Grants and Ethics Services of the University of Ottawa which granted a certificate of ethics approval (reference number 12-08-15). This certificate provided permission for secondary use of data. After the project was approved, an anonymous information agreement was signed on March 2nd 2010 between the Ministry of Health and Long-Term Care (MOHLTC) of Ontario, her Majesty the Queen in Right of Canada, and Dr. Michael Sawada from the University of Ottawa.

The main health measure used in this research is the respiratory morbidity rate for individuals 15 years and over within the Ottawa Public Health Unit. Morbidity was preferred as a measure over mortality due to the larger amount of data (set of information) that would be available, increasing the statistical power to characterize associations with air pollution (Villeneuve et al. 2006). Emergency room visits and hospital admissions are used to evaluate the effect of air pollution on the respiratory health (Lee and Ferguson 2009). Health conditions associated with exposure to NO_2 were identified from the 10th revision of the International Classification of Diseases (CIHI 2006) based on the literature of the health effects of air pollution (Kampa and Castanas 2008). The health outcomes considered in this research are from Chapter 9 (Diseases of the Respiratory System) and more specifically correspond to the block J40-J47 (Chronic Lower Respiratory Diseases).

Data on health conditions listed above were extracted from two different sources, the Discharge Abstract Database (DAD) and the National Ambulatory Care Reporting System (NACRS), maintained by the Ministry of Health and Long-Term Care (MOHLTC). The data extracted covered fiscal years 2005-2006, 2006-2007 and 2007-2008 and are compiled at the geographic level of natural neighbourhoods (chapter 2) and census tracts from the 2006 Census. The two datasets contain similar information, such as demographic and administrative data. The DAD dataset holds clinical data for hospital discharge and day procedures (CIHI 2010a) and the NACRS dataset is the source of information for clinical data for ambulatory care visits (CIHI 2010b). Data on hospitalization and ambulatory care were used along with some of the individual level data (sex, age, postal code of the residence) for each patient. The morbidity rates were directly sex and age standardized (Sridharan et al., 2007; Hu and Rao, 2009) for the age groups 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84 and 85 and over.

1.6.4 Census Data

Data from the 2001 and the 2006 Censuses of Population (Statistics Canada) were utilized in this research. Chapter 2, which is focused on the delineation of the natural neighbourhoods, is based on the 2001 Census data. Once the data from the 2006 Canadian Census of Population were disseminated, a request for custom tabulations to Statistics Canada was made by the Ottawa Neighbourhood Study in order to obtain updated information at the geographic level of the natural neighbourhood. All other analytical results are based on the use of the data from the 2006 Census of population.

1.6.5 Spatial Data

The two main sources of the spatial data used in this research are DMTI™ Spatial and Statistics Canada. The DMTI™ Spatial datasets were used for the delineation of the natural neighbourhoods. Following the delineation, more up-to-date spatial data were acquired from Statistics Canada through the Data Liberation Initiative. Through this program, data from Statistics Canada are made available to Canadian universities at no cost (Statistics Canada 2009). The City of Ottawa also collaborated in the data acquisition process by providing the zoning dataset. Finally, several other spatial layers, such as the greenspaces, were created by members of the Ottawa Neighbourhood Study.

1.7 Structure of the thesis

This thesis is composed of seven chapters. Chapter one is the foundation of the thesis; it sets the objectives of the research, describes and validates the study area and provides a literature review for the study of the relationship between exposure to NO_2 and respiratory health. It also defines the conceptual framework, the study design and the data.

The second chapter presents the methodological approach developed for the delineation of the natural neighbourhoods (section 2.4), the conceptual framework of the Ottawa Neighbourhood Study (ONS) (section 2.5) and presents indicators developed by the ONS (section 2.6). In this chapter GIS methods are applied in order to define a custom geography for the city of Ottawa. This component of the research is essential in the study of the relationship between exposure to NO_2 and respiratory health because the use of the census tracts (CTs) has been criticized by many authors for not representing social or other health-related processes that could have an effect on health outcomes (geography of risk). This custom geography delineated is important as it is used in the following chapters to summarize data associated to the vulnerability of individuals, summarize levels of NO_2 exposure and assess the risk to respiratory health associated with exposure to NO_2. As a consequence this chapter is crucial in the attainment of the main objective of this thesis.

In chapter three a LUR model representing the geography of exposure is developed for the city of Ottawa. This chapter provides a literature review of the LUR model as well as of the dasymetric mapping. The role of spatial representation in the development of a LUR model is the focus of this chapter. To our knowledge this is the first research to include dasymetric mapping into a LUR model and to assess the role of spatial representation in the model building exercise of the LUR. This model is important as it is used in chapter four to assess the importance of spatial representation in the geography of exposure. The NO_2 concentrations map (data) produced is also used in chapter five and six for the comparison of different exposure models and for the assessment of the risk associated with exposure to $NO_{2.}$.

A second aspect of the question of spatial representation in the geography of exposure is studied in chapter four. Spatial representation is examined in the context of exposure models; the results of different methodological approaches to the modeling of NO_2 are compared. The results, which include the identification of the best performing model are used as input into the last two analytical chapters of the thesis which address the relationship between exposure to NO_2 and respiratory health.

Chapter five focuses on the role of spatial representation on the geography of risk. The relationship between exposure to NO_2 and respiratory health is assessed under three different spatial structures: census tract, natural neighbourhood and a cluster structure. The latter is created through automated zone design. This chapter produces key findings in relation to the main objectives of this research. By using different spatial structures to summarize the information and assess the relationship it allows for a better understanding of the role of the MAUP in the study of the geography of risk.

The geography of risk is the main focus of the chapter six, the last analytical chapter of the thesis. Using the natural neighbourhoods delineated in chapter two and the NO_2 concentration map generated in chapter three, the effect of NO_2 exposure on respiratory health is assessed. In this chapter analytical methods designed specifically for the analysis of spatial data are applied. The use of multiple approaches to assess the relationship between exposure to NO_2 and respiratory health is considered a form of sensitivity analysis. The CCHS data are considered in this chapter, which is not the case for the other spatial structures discussed in chapter five.

Finally, chapter seven provides a summary of the findings along with suggestions for future research. Among these is the need to include daily mobility in future endeavors.

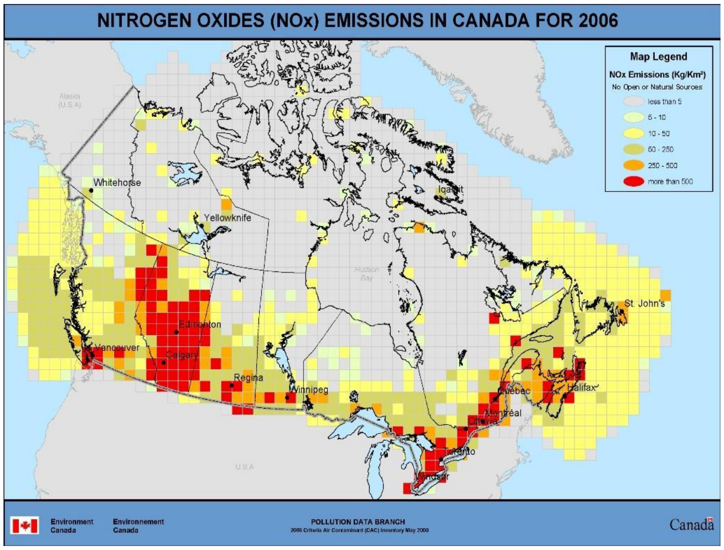


Figure 1.1 National map of NOx emissions for the 2006 reporting year distributed on a 100 km by 100 km grid. Emissions per square kilometer within the grid cell (Environment Canada 2010a).

Pollutant	Averaging Time	Maximum Desirable Level	Maximum Acceptable Level	Maximum Tolerable Level
Sulphur dioxide (SO_2)	annual	11 ppb	23 ppb	---
	24 hours	57 ppb	115 ppb	306 ppb
	1 hour	172 ppb	334 ppb	---
Total Suspended Particulate (TSP)	annual	60 µg/m3	70 µg/m3	---
	24 hours	---	120 µg/m3	400 µg/m3
Carbon Monoxide (CO)	8 hours	5 ppm	13 ppm	17 ppm
	1 hour	13 ppm	31 ppm	---
Nitrogen	annual	32 ppb	53 ppb	---
Dioxide (NO_2)	24 hours	---	106 ppb	160 ppb
	1 hour	---	213 ppb	532 ppb
Ozone (O_3)	annual	---	15 ppb	---
	24 hours	15 ppb	25 ppb	---
	1 hour	51 ppb	82 ppb	153 ppb

Table 1.1 National Ambient Air Quality objectives and guidelines in Canada (Health Canada 2006a).

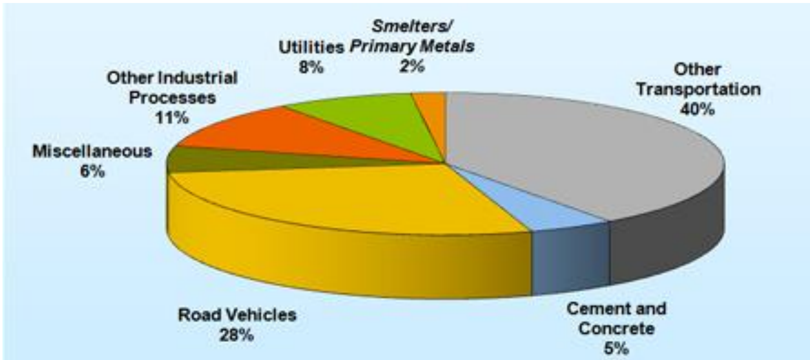


Figure 1.2 Ontario Nitrogen oxides emissions by sector (emissions from point/area/ transportation sources, 2006 estimates) (Ministry of the Environment Ontario (2010).

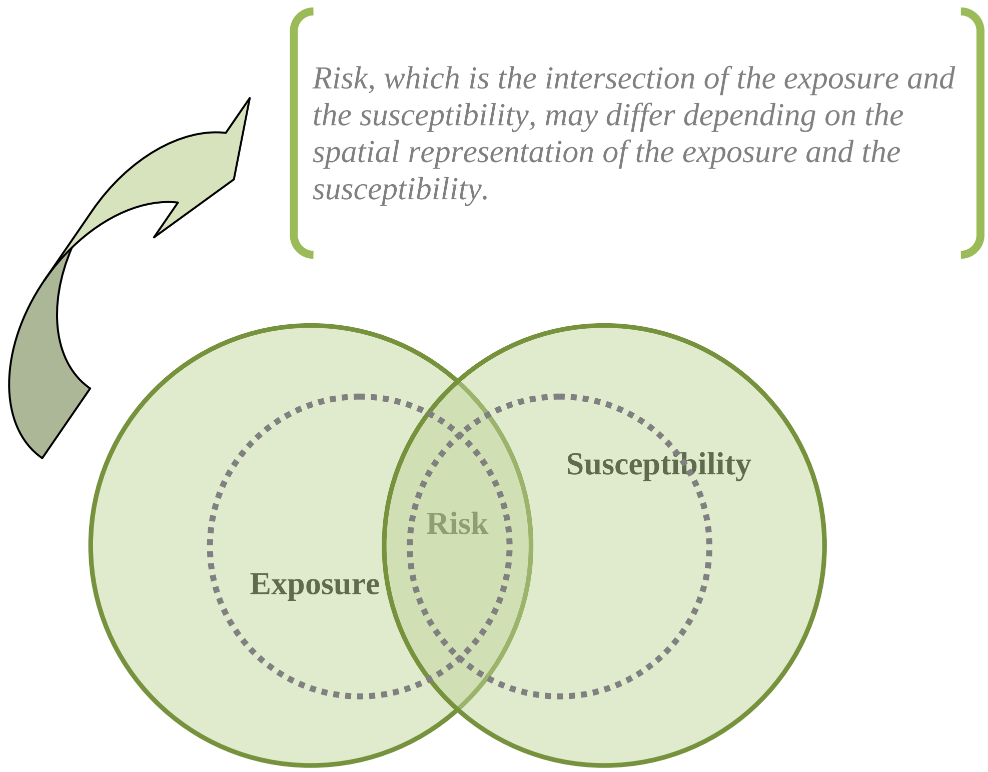


Figure 1.3 Conceptual Framework as defined by Jerrett and Finkelstein (2005). The overlap between the geography of exposure and the geography of susceptibility represents the geography of risk as defined for this research.

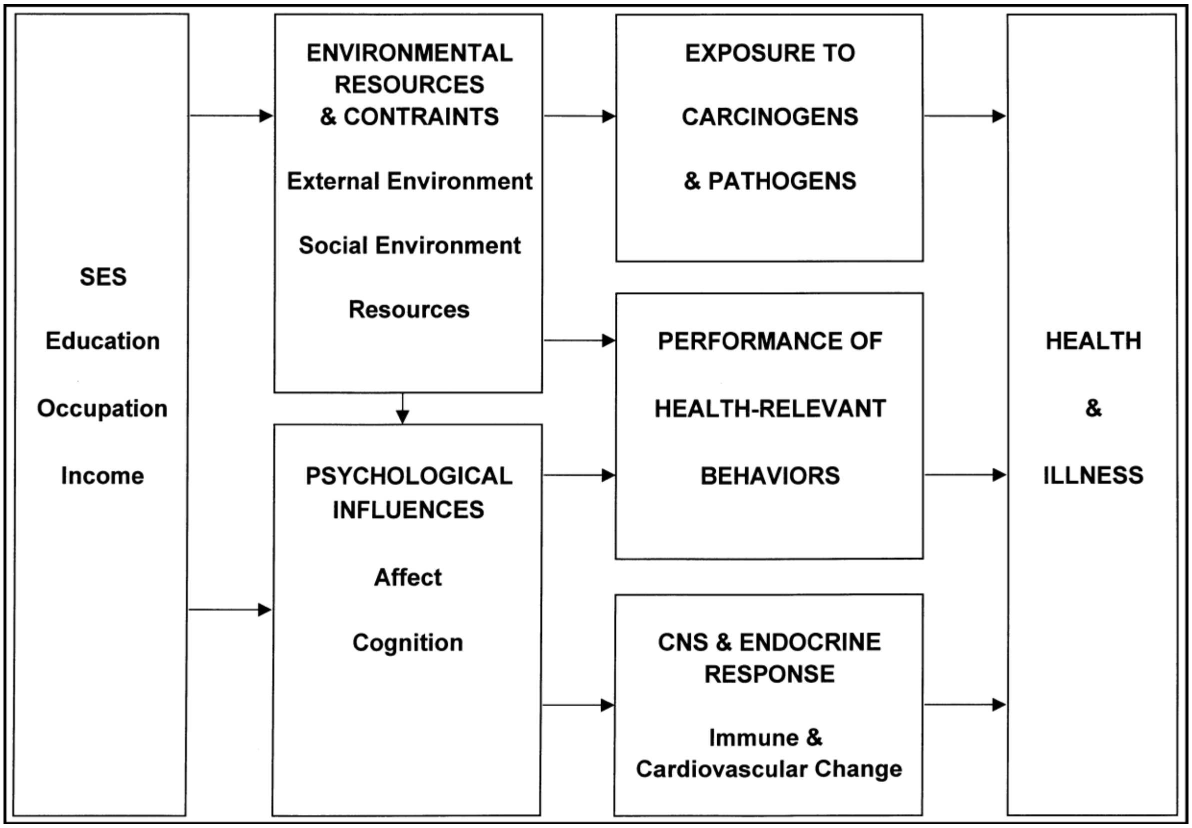


Figure 1.4 Possible pathways between SES and health (from Adler and Ostrove 2002).

Chapter 2 Development of Neighbourhoods to Measure Spatial Indicators of Health[1]

Marie-Pierre Parenteau[1]

Michael Sawada[1]

Elizabeth A. Krisjansson[2]

Melissa Calhoun[2]

Stephanie Leclair[2]

Ronald Labonte[3]

Vivien Runnels[4]

Anne Musiol[4]

Sam Herold[1]

(1) Department of Geography, University of Ottawa

(2) School of Psychology

(3) Canada Research Chair (Tier 1) in Globalization/Health Equity

(4) Institute of Population Health

Corresponding author: Marie-Pierre Parenteau

Laboratory for Applied Geomatics and GIS Science, Department of Geography, University of Ottawa, Ottawa, Ontario K1N 6N5

Tel:+1 613 562-5800 Ext. 1040; Fax: +1 613 562-5725

Published in: URISA Journal, Vol. 20, No.2, 43-55

[1] This chapter was written in collaboration with members of the Ottawa Neighbourhood Study. The sections of this article that pertain to the concept of neighbourhood and to the delineation of geographical units are the sections of interest for this dissertation.

2.1 Introduction

In place-based research, geographic information systems (GIS) can be used to derive the context of place and further our understanding of whether place influences health. The external context may include the quality of the physical environment, resources (material and social), and infrastructure that can affect individual health (Pearce et al. 2006). These contextual factors act directly in some instances and indirectly in others (Evans and Stoddart 1994). A strong relationship exists between individual socioeconomic status (SES) and the quality of the neighbourhood environment; this may amplify the disparities in health between the richer and the more deprived (Yen and Syme 1999, Fiscella and Williams 2004, Braveman 2006). Researchers only recently have begun to study the impact of various neighbourhood-level factors on individual health and health inequalities.

In place-based health research, GIS can be used to further an understanding of whether the characteristic context of a place influences health. In this research, natural neighbourhoods within Ottawa, Canada, were delineated, using data from DMTI™ Spatial, Statistics Canada, the City of Ottawa, the National Capital Commission (1998), the Ottawa Real Estate Board, DigitalGlobe satellite imagery, field-based observations, and expert and community knowledge. These neighbourhood units were used within a GIS to derive contextual health indicators in the natural environment, social environment, goods, services and amenities, and the built environment. These indicators were organized into a set of health-relevant domains inspired by Maslow's hierarchy of needs (Maslow 1968; Maslow 1970), which was the basis of the conceptual framework for this research[2]. The ultimate goal was to determine which, if any, contextual indicators act as predictors of health outcomes. In the subsequent sections, research goals are described and the methodology used to delineate the neighbourhoods and the conceptual framework and methods used to derive the indicators are provided. In conclusion, the initial results compare a measure of SES status within the neighbourhoods and neighbourhood health indicators.

2.2 Description of Research

This study was initiated by a multidisciplinary team from the University of Ottawa who engaged in collaborative community based research aimed at reducing regional health inequalities.

The practical objective was to work with city policy makers, planners, and program implementers to develop strategies and procedures to reduce health inequalities in Ottawa (Krisjansson et al. 2007). This project was focused on spatial inequalities in neighbourhood resources for health, which can lead to inequities from a social justice perspective. More specifically, this project had four objectives:

• To develop a methodology for defining “natural” neighbourhoods;

•To gather data on a number of neighbourhood social and physical resources/amenities; to essentially create a community inventory and subsequent measures of accessibility using GIS capabilities (Pearce et al. 2006);

[2] Ibid.

• To map the relationships between neighbourhood SES, the distribution of resources necessary for health, and health outcomes; and

• To share the evidence with decision-makers and relevant community organizations and to assess the usefulness of the GIS tools in a participatory process of neighbourhood delineation.

Because the project is still under way[3], an analysis of all community resource indicators with SES is not yet completed. However, the preliminary results suggest clear intra-urban variations in neighbourhood SES and relations with health indicators. As such, others should benefit from this experience and methods thus far. The health outcome indicator analysis is for a future publication[4].

2.3 Study Area

The City of Ottawa, Ontario, is the national capital of Canada (Figure 2.1), with a population of 812,129 and a population density per square kilometer of 292.3 in 2006 (Statistics Canada 2007a). While the Ottawa-Gatineau census metropolitan area (CMA) crosses a provincial boundary and includes the city of Gatineau in the province of Quebec to the north, it was beyond the scope of this current research. Gatineau was excluded because of issues of data collection within the different national, regional, and local administrative structures. However, work is ongoing to assimilate the Gatineau data and repeat the methods and study for the entire CMA.[5] Ottawa is characterized by a generally well-educated population (36.8% of the residents age 35 to 44 had a university certificate, diploma, or degree in 2001) with

[3] Ibid.
[4] Ibid.
[5] Ibid.

a higher median family income than the provincial median (with a median family income of $73,192 in 2000); Ottawa residents are $12,000 over the provincial median income (Statistics Canada 2002).

2.4 Defining Neighbourhoods

In neighbourhood studies, the areas studied often consist of political or statistical units (e.g., census tracts, wards, etc). In a study about social processes in neighbourhoods, Sampson et al. (2002) found that of 40 studies identified by a systematic search, less than five used a methodology that did not operationalize neighbourhoods according to political or statistical areas. Frequent use of the census tracts (CTs) has been criticized by many authors because they represent imposed, irregular boundaries (Bonham-Carter 1994) that have no effect on social or health-related processes that take place within and between them (Ellen and Turner 1997; Germain and Gagnon 1999; Kawachi and Berkman 2004; Martin 2004; Clapp and Wang 2006), but few studies have attempted to define units that represent residents' perceptions of their neighbourhoods (Dietz 2002; Diez-Roux 2001; Diez-Roux 2002). In this research, "natural" neighbourhood units were delineated, allowing confirmation that the obtained results were not artifacts of the boundaries that were utilized (Ross et al. 2004). The methodology developed to define the natural neighbourhoods was based on three considerations, specifically, the functional approach, the physical approach, and the Ottawa Multiple Listing Service (MLS) real estate board neighbourhood maps.

A functional approach, based on the Chicago School of Sociology (Park et al. 1967), was adopted in the first delineation of neighbourhoods. This approach considers the physical and demographic aspects of neighbourhoods (Martin 2003) and as such guided the selection of relevant socioeconomic and demographic variables. Data from the 2001 Canadian Census (see Table 2.1) at the geographic level of the dissemination area (DA) were used as input for spatially constrained clustering and wombling (Legendre and Fortin 1989; Fortin 1994; Fortin 1997; Fortin and Drapeau 1995; Lu and Carlin 2005). The DA is the smallest geographic unit at which 20% sample data from the Canadian Census are disseminated. Dissemination areas have a population count between 400 to 700 people and are used by Statistics Canada to generate census tracts (Statistics Canada 2001), which are used most often as a neighbourhood unit proxy (Pearce et al. 2006; Pearce et al. 2008; Ross et al. 2004).

Spatially constrained clustering identifies units that are similar and adjacent in space (shown in Figure 2.2). Clusters are computed using various clustering algorithms such as K-means, but the spatial constraint has to be respected. Only the sites or the group of sites that are contiguous according to a list of predetermined connections will be clustered (Fortin and Drapeau 1995). This technique results in areal boundaries that are closed and crisp (Jacquez et al. 2000). Wombling generates open boundaries (difference boundaries) by computing the slope (first partial derivative) of qualitative or quantitative data. Results from these two methods were integrated to operationalize the functional approach and provide the first quantitative approximation of Ottawa neighbourhoods.

Once the initial set of contiguous DA clusters was generated, a physical approach was used to refine the neighbourhood boundaries. Within the literature, physical features are considered important elements in the identification of neighbourhood boundaries. The underlying assumption is that these barriers mitigate the negative externalities for residents who prefer to not live near people who are different. Therefore, natural boundaries not only serve functional purposes such as transport or recreation, but they also have the capacity of working as a buffer zone between different groups (Aitken and Prosser 1990; Hoxby 2000; Noonan 2005).

In this work, elements of the environment that were considered to potentially act as physical barriers between neighbourhoods were overlaid on the results of the functional approach. The municipality identifies some boulevards, main streets, heritage conservation district streets, scenic parkways, transitways, railways, highways, bridge crossings, and waterways as barriers to movement and social and economic vitality (City of Ottawa 2006a). Road network data (Statistics Canada 2006; DMTI Spatial 2007d) were used; the juxtaposition of the functional boundaries with the results of the clustering and wombling allowed us to identify any anomalies, that is, areas where there was disagreement, and to rectify these accordingly. At this stage, two constraints were imposed to the neighbourhood physical approach:

(1) Boundaries must follow a dissemination area boundary and (2) boundaries must follow an ostensible feature, natural or imposed (Bonham-Carter 1994). From the combination of the functional and the physical approaches, the first preliminary set of natural boundaries was defined.

The preliminary set of natural boundaries then was compared to the Ottawa Multiple Listing Service (MLS) maps. These maps name and identify neighbourhood units that have been used by members of the real estate profession in Ottawa for more than a decade. While MLS units are somewhat ad hoc creations and unofficial, they are based on expert knowledge of local real estate interests and tend to be accepted by buyers, sellers, and residents and used by the provincial Municipal Property Assessment Corporation. The results of the combined functional and physical approaches agreed remarkably well with the MLS maps at a similar level of aggregation (Figure 2.3). As such, the results provide some evidence that time and familiarity with the units have made these MLS maps a standard from a socioeconomic and demographic perspective. The names of the neighbourhoods that were used in the MLS maps were retained for the current study where possible, to allow for better recognition of the neighbourhoods by citizens and city planners.

Rural DAs were assigned to their closest satellite village (e.g., Vars, Munster, etc.; shown in Figure 2.4). The assignment of rural areas to a given satellite village was based on nearest network travel time using the road network. The network was allowed to extend beyond the city boundary. The concept of nearest was based on the population-weighted centroid of each rural area and travel time to the satellite village coordinates provided by the city of Ottawa (2006). This method assumed that the majority of rural individuals will travel to the nearest population center for daily amenities such as food shopping and gasoline and thus more often than not tend to associate themselves with one of the rural centers.

Following the delineation of neighbourhood boundaries, consultations with team members representing city planning, public health, housing, community health centers, and grassroots organizations were undertaken and fieldwork was conducted. Finally, several neighbourhood units were aggregated to meet the minimum sampling requirements for health analysis of about 4,000 persons per neighbourhood (Ottawa Public Health, personal communication 2007). In the end, 89[6] neighbourhoods were delineated and approved by all of the investigators involved in this research.

2.5 Conceptual Framework

The organizational framework for the selection of health indicators was based on Maslow's hierarchy of needs (Maslow 1968; Maslow 1970; Figure 2.5). According to Maslow's work, an individual must satisfy his or her basic needs before he or she can focus on the higher needs (Shao et al. 2006). At the base of the hierarchy are four basic needs: physiological needs, safety needs, belongingness and love needs, and the need for self-esteem. In the middle of the hierarchy are two more advanced needs: the need for knowledge and understanding and the need for creativity and

[6] The final version of the natural neighbourhood boundary file, which was later completed, contains 95 units.

aesthetics (Hagerty 1999). Finally, the two most abstract needs, the need for self-actualization and the need for transcendence, are represented at the top of the needs hierarchy.

Maslow's theory was that people cannot grow and be physically, mentally, and spiritually healthy until basic needs are satisfied. For example, if one does not have adequate food for oneself and one's family, it is hard to think about higher-level things such as helping others, socializing, or cultural advancement. Looking at the hierarchy in terms of neighbourhoods, the question was asked: What should a neighbourhood have to make it a good place to live, to be healthy, to grow, and to raise a family?

2.6 The Indicators

In this research, a set of neighbourhood contextual indicators were developed to measure neighbourhood resources, to determine whether there were spatial inequalities in access to these resources, and to determine whether or not such inequalities correlated with variations in health. The indicators were conceptualized and operationalized based on the hierarchy of needs with most of the indicators representing basic levels with additional indicators inspired by the more advanced and abstract needs of the pyramid (Figure 2.5). As such, the contextual indicators focused on both physical and social resources affecting health and their linkage to Maslow's hierarchy are illustrated in Table 2.2. In the process of indicator selection, two criteria were applied: first, the indicator had to be established as important to health within the learned literature, and second, the indicator's context, composition, configuration, quality, or quantity had to be amenable to change or intervention, which would suggest that the city of Ottawa and other levels of government have the potential to modify them. The indicators were grouped into five domains representing potential for intervention (Table 2.2).

In total, 44 indicators were derived within the five domains. Table 2.3 defines the first four domains with one example of an indicator within each. The intent of this paper is to provide an example methodology for the selection of indicators and their operational implementation. Unfortunately, scope and space do not allow the provision of all 44 indicators. Collecting the spatial and attribute data to derive an indicator involved the integration of existing datasets, manual verification, and various research methods. For example, grocery stores or supermarkets (as shown in Table 2.3) are one indicator within the domain of Goods, Services, and Amenities. Different datasets and sources had different definitions of, for example, a grocery store, convenience store, or specialty food store. However, for purposes of examining food quality, the number of grocery stores per 1,000 households and the average distance to the four closest grocery stores in each neighbourhood was mapped. Referring to the North American Industry Classification System (NAICS) (U.S. Census Bureau 2007) to classify data that corresponded to an indicator of interest, a dataset containing only grocery stores was created. Definitions were refined to help distinguish these from convenience and specialty stores based on a widely accepted standard. The Standard Industrial Classification (SIC) (U.S. Dept of Labor) codes were used to create a list of food services from enhanced points of interest (DMTI Spatial 2007a). The EPOI dataset is a georeferenced Canada-wide inventory of industries that fall within the purview of the NAICS. The EPOI dataset from February 2007 was used to extract point locations for food services mapping. It quickly became apparent that there were inconsistencies, missing data, and misclassifications within the EPOI dataset. If one were interested in mapping the intensity of various SIC variables at the scale of a census metropolitan area (CMA) using the EPOI, for example, for a market-competition analysis, the patterns generally are well represented in terms of spatial intensity. However, at the neighbourhood level, the issues with the EPOI required a significant amount of research and fieldwork by the team to achieve a full enumeration of each indicator for the final dataset (conducted through fieldwork):

1. Classifications

•All classifications followed NAICS classification categories.
•Classification criteria were verified during telephone research.
•Grocery stores were required to provide a relatively full line of fruits and vegetables and fresh meats. If a store carried a limited line of fruits and vegetables or meat products, it was classified as a convenience store.
•Convenience store lacked the previous criteria.
•Specialty stores were classified as ethnic, meat, fish/seafood, fruit/vegetable, confectionary/nut, dairy, bakery, health food, other.
•Stores that could meet criteria for more than one category were classified according to their primary purpose.

2. Verification

• Checked Canada 411 Business (http://www.canada411.ca) for listing.

• Checked Canada 411 Person for listing as a residential line.

• Checked Ottawa Retail Survey (NAICS Codes 451 and 4452).

• Checked Web site search engines to verify name and address of store.

• Stores that were not on the DMTI Spatial dataset but were personally known to team members were added.

3. Top Grocers and Convenience Stores

• Top grocers and convenience stores were doubled-checked with store Web sites and Canada 411 and if not found were added; these included:

• Grocers:

• Loblaws

• Sobey's

• Loeb

• Food Basics

• Your Independent Grocer

• Real Canadian Superstore

• Convenience stores (also cross-checked with Ontario Convenience Stores Association):

• Mac's

• Quickie

• Quick Food Market

• 7-Eleven

• Hasty Market

• Pronto

• Ainee's

4. Transfers

• A number of stores that did not belong in this dataset were transferred to other datasets used in the study.

• These included fast food, restaurants, and pharmacies.

5. Additions

• Major department stores offering (limited) grocery counters were added to the list.

• These included Wal-Mart, Zellers, and Giant Tiger. Address verification was obtained from Canada 411 and retailer web sites.

6. Exclusions

• Listings from the dataset that were confirmed as residential numbers were removed from the list.

• In the case of duplicate entries (e.g., where the same store was listed twice at the same address but with separate phone numbers), only one entry was counted and the other was deleted.

• Food distributors were excluded.

• Any stores that were closed down/out of business at the time of calling were excluded. Store closure was determined via: (1) calling the number and being informed the store had closed, (2) working group knowledge of a store on the list that was known to have closed down, (3) a site visit.

• Phone numbers that were not in service were moved to another file for checking: if the existence of a store via Canada 411 or working group member knowledge could not be confirmed, the store was excluded.

• *Grocery store exclusions:* Online grocers and food delivery services were excluded because they are not neighbourhood-based resources.

• *Specialty stores:* Deli counters that were part of larger grocery stores and nutrition centers/superstores were excluded.

7. Final Verification

• Fieldwork was conducted (street searches for verification) where it was not possible to reach the business by telephone.

• With a complete set of entries for grocery stores, GIS concepts and capabilities were used to derive the neighbourhood indicator (Table 2.3). GIS functions allowed for the localization of facilities neighbourhoods (using point-in-polygon analysis), the calculation of summary statistics by neighbourhoods, and measurements of accessibility.

After address geocoding the list of grocery stores, the number in each neighbourhood was counted using point-in-polygon analysis, added to the attribute table of the neighbourhood layer, and standardized by population using the database functionality of the GIS. Subsequently, graduated symbols were used to represent the indicator as described in Table 2.3. A second food accessibility indicator required determining the average distance to the four closest grocery stores for each neighbourhood. Network analysis was used to find the closest four grocers either inside or outside each neighbourhood. Following Pearce et al. (2006), distances to grocery stores were measured from the population weighted centroid of each neighbourhood. The network used was based on the most recent road network files that contain the attributes of street name, type, direction, and address ranges. Road segment travel times were based on municipally mandated speed limits according to road type. These methods of data collection and subsequent GIS analysis were repeated to complete indicators for convenience stores, specialty stores, health services, community recreation, education, and financial services.

A composite socioeconomic indicator was developed and based on measures of SES, including educational attainment, occupational characteristics, income, living conditions, and immigration (Braveman et al. 2005; Braveman 2006; Gallo and Matthews 2003; Krieger et al. 1997; Krieger et al. 2003; Lynch and Kaplan 2000; Williams and Collins 1995). The index of neighbourhood advantage was based on neighbourhood census variables: % of residents with less than a high school education, % of lone-parent families, % of recent immigrants, % unemployed, % below the low income cutoff (LICO), and average income. Variables within the index were adjusted for age and sex distributions within the neighbourhood. The Principal Components Analysis (one strong component emerged) then was used to derive an overall index of socioeconomic advantage. The same methodology was used to derive other composite indicators such as healthy and unhealthy food indexes, healthy and unhealthy financial indexes, recreation indexes, etc.

The SES index was represented as a choropleth map by neighbourhood overlaid with different contextual indicators. For ease of comparison, different indexes were developed to represent the composite behavior of the contextual domains. For example, the quantitative data derived using network distances and counts within the GIS for indicators that represent accessibility to unhealthy food were created and shown as proportional symbols on top of the choropleth map of SES (Figure 2.6). The strength of GIS tools, therefore, is central to this project.

2.7 Health Outcomes

The health outcomes used to assess the relationship between inequalities in neighbourhood quality and health disparities were obtained from Ottawa Public Health (OPH) for the fiscal years 2004–2005 and 2005–2006. Lists of six-digit postal codes (DMTI Spatial 2007c) were provided to the Ottawa Public Health

(OPH)[7], which then tabulated the health outcome dataset for each unit. Four health outcomes were used to determine the health profiles of people in each neighbourhood.

Hospital Admission Rates for Ambulatory Care Sensitive Conditions

Ambulatory care sensitive conditions are those for which appropriate and timely outpatient care could reduce hospital admission by preventing the occurrence of the illness or condition, controlling the acute illness or condition, and/or managing the chronic disease or condition (Manitoba Centre for Health Policy 2007). Such ambulatory care sensitive conditions include, but are not limited to, asthma, diabetes, angina, pelvic inflammatory disease, gastroenteritis, congestive heart failure, severe ENT infections (ear-nose-throat), epilepsy, and cellulitis.

Rate of Emergency Room Use for Ambulatory Care Sensitive Conditions

Emergency departments have two core functions in an integrated primary care system: the provision of specialized clinical skills focused on the assessment and management of urgent or emergent medical needs, and the provision of continuous 24-hour access to primary care services. These are important primary care roles; recent Canadian estimates suggest that 15% to 25% of urban populations will use emergency department services at least once in a 12-month period. This health variable was measured using the number of emergency room visits by people in a given neighbourhood.

[7] The health outcome data was provided by OPH for the Ottawa Neighbourhood Study. For this dissertation, the health outcome data was obtained directly from the MOHLTC as described in section 1.6.3.

Rate of Smoking in Pregnancy

Smoking during pregnancy harms both mother and fetus. Aside from increased morbidity and mortality from cancer and cardiovascular and pulmonary disease in the mother, smoking has been implicated in the etiology of placenta abruption, placenta previa, spontaneous abortion, premature delivery, and stillbirth (Moner 1994). The health indicator used was maternal cigarette smoking during pregnancy.

Rate of Low Birth Weight

Low birth weight is an important population health outcome, for low birth weight babies are at a higher risk for physical and mental problems. A number of researchers have shown that neighbourhood socioeconomic status is related to birth weight, with poorer neighbourhoods having high rates of low birth weight (Joseph and Kramer 1997). Low birth weight was measured as the number of live newborn babies weighing between 500 and 2,499 grams at birth (Joseph and Kramer 1997).

2.8 Results and Discussion

The preliminary comparisons of food indicators indicate significantly more fast-food outlets in the lower two SES quintiles than in the highest quintile. Significantly more grocery and specialty stores were in the lowest and third SES quintiles when compared to the highest quintile. From a food quality and accessibility perspective, significantly more schools are closer to fast-food outlets in the lowest socioeconomic quintile (76%) than in the highest quintile (39%). While further analysis is forthcoming, it is clear that spatial variations appear in the indicators and that these variations may have socioeconomic implications and health impacts.

The delineation of the neighbourhoods was an important component of this project. The final units used in this study were defined and agreed on by the research team members. Eighty-nine neighbourhood units were delineated in total, all with a population of more than 4,000 individuals. The population count was fundamental to ensure a sample size sufficient for the next steps of statistical analysis of health outcomes. Nine neighbourhoods were classed as rural and were excluded from all analyses concerning amenities and socioeconomic status. (Note: rural neighbourhoods were included in the derivation of indicators and for neighbourhood profiling.) Three additional neighbourhoods were excluded because of insufficient population for statistical analysis. The neighbourhood approach in this research is novel and transferable to other studies that may have similar goals.

While food services such as grocery/supermarket, specialty, and convenience stores are well defined under the NAICS and SIC, frequent misclassifications appear in the datasets that were used. This observation held true for all of the indicators, making field observation as well as Internet and phone calls necessary to avoid errors of omission and commission as much as possible.

Alternate modes of travel to grocery stores or other amenities were not attempted. As such, it is uncertain that the optimal measure of accessibility was attained. A number of multimodal transportation scenarios could be tested using network analytical concepts and capabilities. Individuals in less affluent neighbourhoods may have limited access to automobiles, making cycling or walking a more feasible mode of transportation to the nearest amenity. These observations further suggest the possibility of standardization of average distances according to census-reported modes of transportation in the neighbourhoods. The extent to which the added complexity would provide additional explanatory power for health outcomes has yet to be explored.

2.9 Conclusion

Ultimately, the relationship between inequalities in neighbourhood quality and health disparities in Ottawa will be assessed[8]. This research project on the contextual influences of neighborhoods on health is one of the few projects of its type in Canada. "Natural" neighborhoods were defined and neighbourhood contextual indicators were conceptualized and operationalized, using the benefit of the experience of American and British counterparts. Another important aspect of this project is the high level at which community leaders and policy makers were involved. After two years into the research initiative, most of the core work has been completed. The assessment of the relationship between neighbourhood quality and health disparities is currently under way[9]. The results of these analyses then will be published and shared with decision makers and relevant community organizations to assess the usefulness of GIS tools as a means to understand the impact of

[8] In reference to the Ottawa Neighbourhood Study
[9] Ibid.

neighbourhoods on health. Readers can keep up-to-date by reference to the Ottawa Neighbourhood Project Web site, http:// neighbourhoodstudy.ca.

Acknowledgments

The authors thank all the members of the Ottawa Neighbourhood Study and appreciate the strong participation of Ottawa Public Health (OPH). Funding for this project was provided by CIHR in addition to the Canadian Foundation for Innovation and Ontario Innovation Trust. Stephen Perkins, Ian Cross and Russel Mawby at the City of Ottawa were instrumental in refining the datasets used in the definitions as was the participation of Amira Ali, Andrea Morissette, Caroline Andrew, Cliff Gazee, David Gibson, David Hole, David Salisbury, Holly Johnson, Simone Thibault, Moe Garahan, Nazeem Muhajarine, Peter Tugwell, Robin McAndrew, Ted Schrecker, James McKinnon and Katherine Haimes.

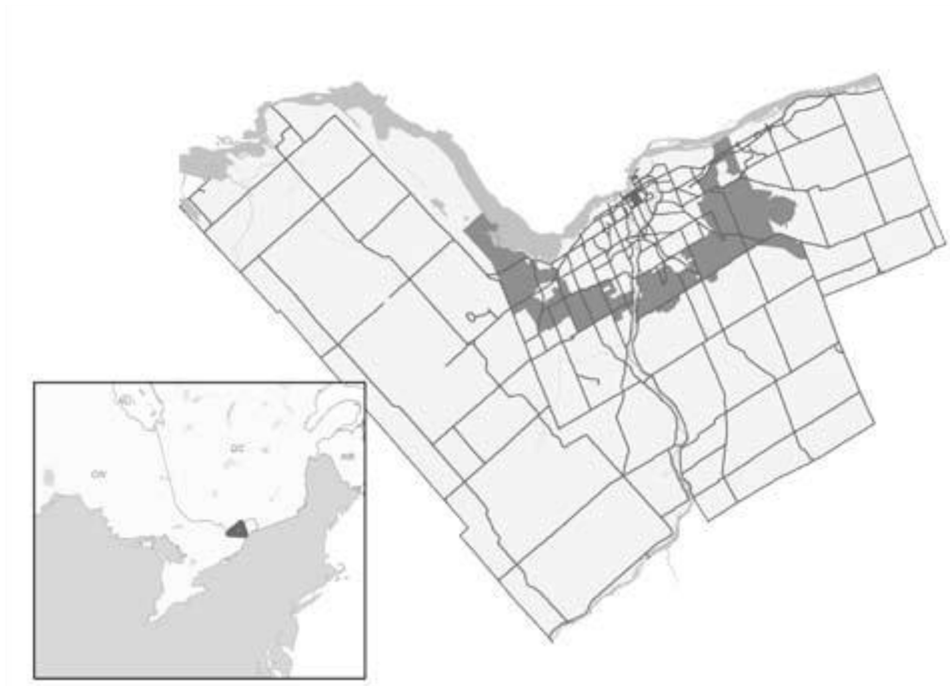


Figure 2.1 Study region in the context of North America.

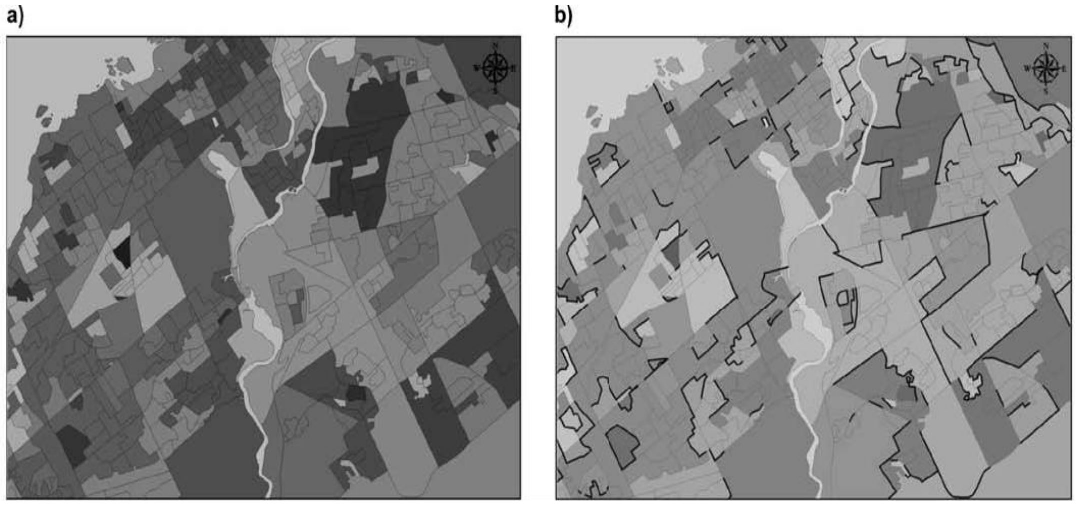


Figure 2.2 Section of Ottawa area: (A) Light gray line boundaries represent dissemination area polygons. Clusters of dissemination areas with the same shade of gray represent the results form spatially constrained clustering. (B) Illustration of the differences in boundaries from wombling that generally agree with socioeconomic changes at the cluster boundaries

.

a) b)

Figure 2.3 (A) Natural barriers – major roads – (heavy black lines) overlaid on neighbourhoods shown by regions of like grayscale; (B) illustration of MLS map for comparison of neighbourhood boundaries (MLS source http://orebweb1.oreb.ca/mlssearch/SearchMlsMap.aspx?x_map-53).

Factor	Variable
Economic	Median household income
	Unemployment rate
Housing	Affordability (more than 30% income spent on housing)
	% structures built before 1961
	% dwellings owned
	% single-detached
	Median value of dwelling
Social	% visible minority
	% pop. with a bachelor's degree

Table 2.1 Variables from the 2001 Canadian Census or population used as input for spatially constrained clustering and the wombling.

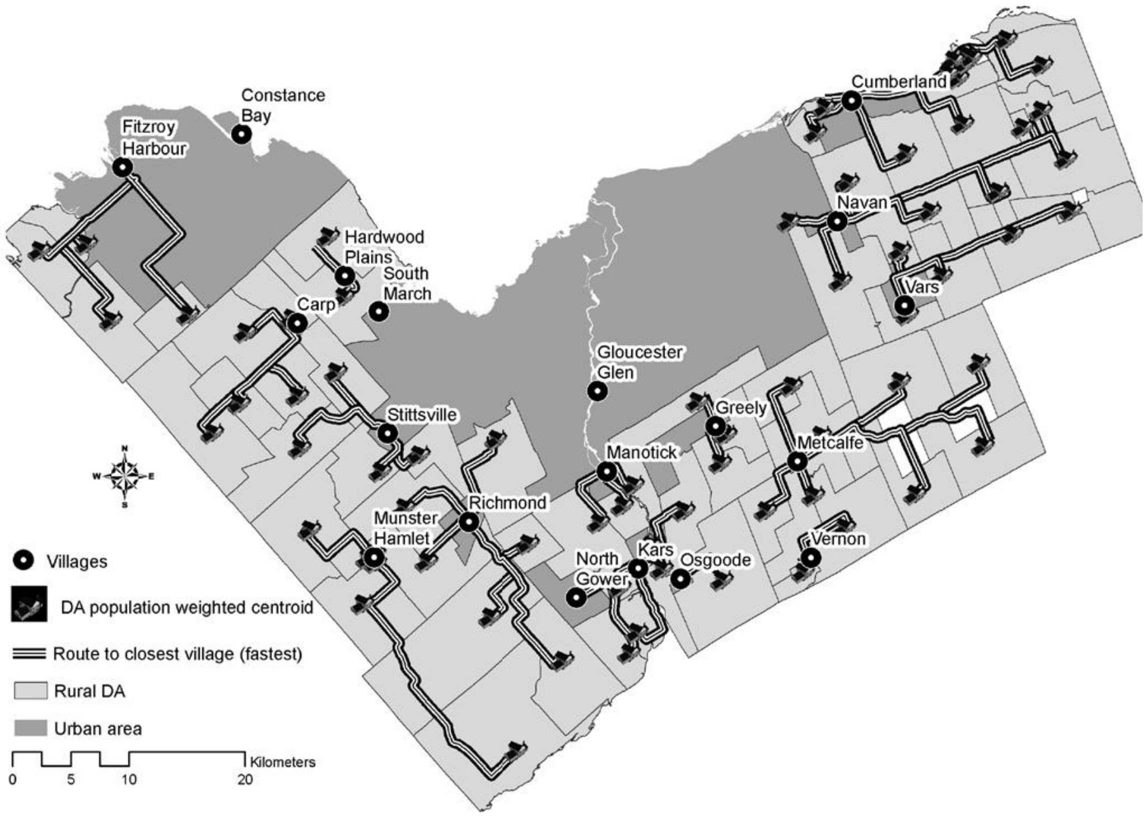


Figure 2.4 Assignment of rural DAs to the nearest satellite village to create rural neighbourhood affiliations.

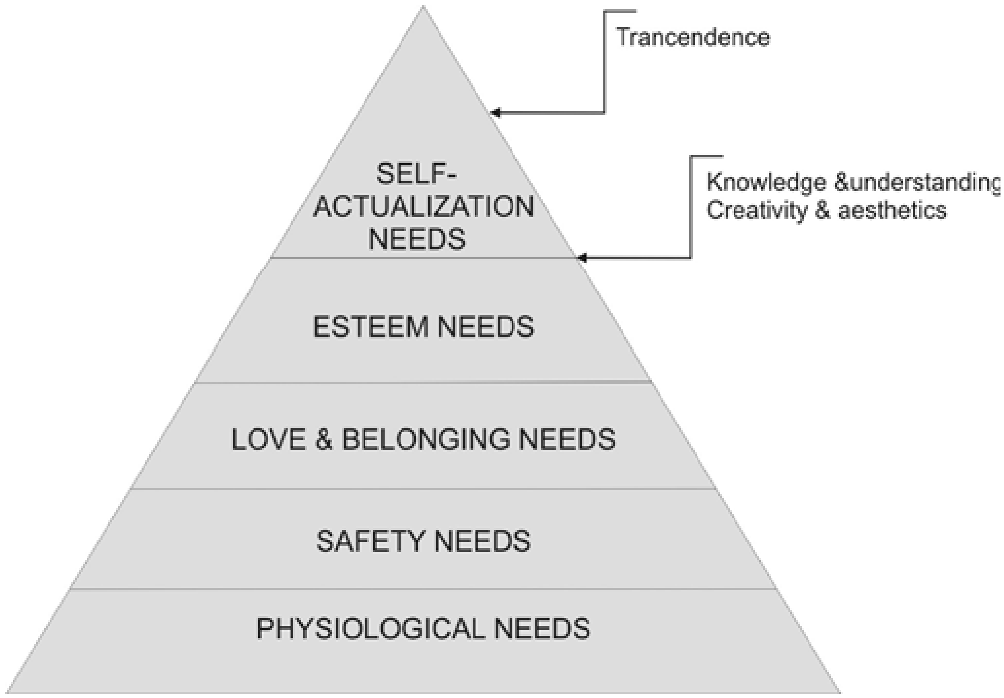


Figure 2.5 Maslow's (1968, 1970) hierarchy of needs (gray pyramid) with location of further needs as suggested by Hagerty (1999).

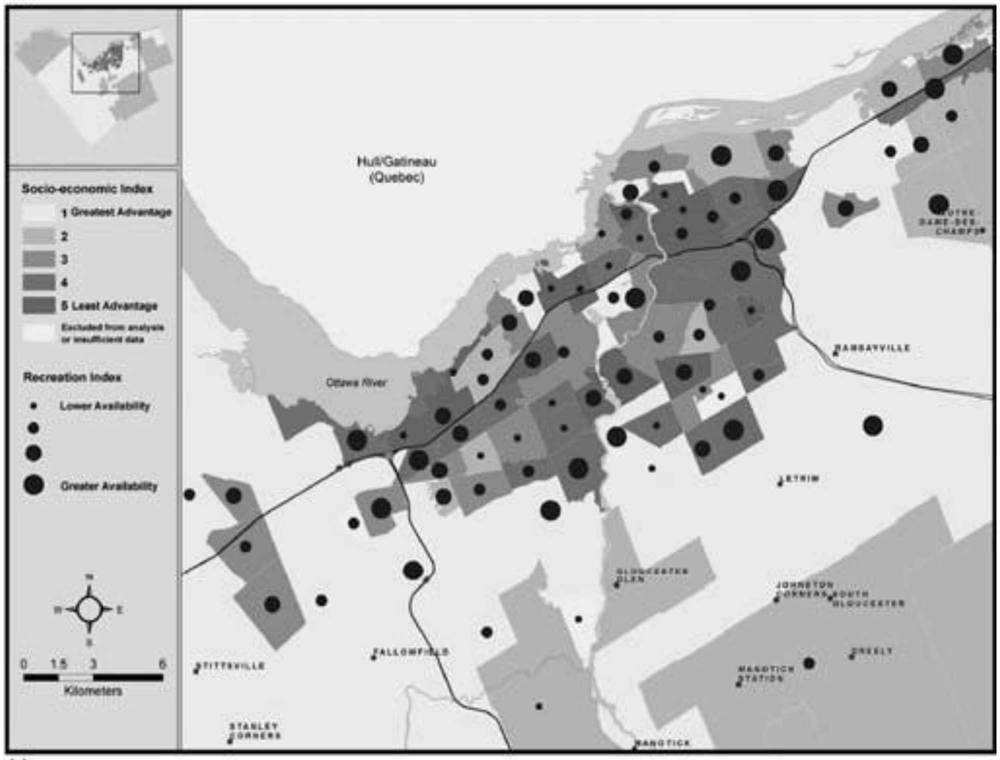
Domains and Indicator Category	Maslow's Hierarchy
Goods, Services and Amenities Food Recreation Education (schools, libraries) Health services Financial services	Physiological Needs
The Natural Environment Greenspace Parks	Aesthetic needs
The Social Environment Voting rates Crime (personal and property) Mobility Sense of belonging	Safety, belonging, need to know
The Built Environment Housing in need of repairs Crowding Affordability	Esteem needs, physiological needs (shelter)
Neighborhood Sociodemographics Families below low-income cutoff Education levels Lone-parent families Unemployment	Mix of needs

Table 2.2 Linkage of indicator domains organized around Maslow`s hierarchy of needs. Indented items represent categories within each domain for which specific indicators were derived.

Goods, Services, and Amenities	
Justification (domain)	Access to goods, services, and amenities promotes a healthy living by allowing residents to have good nutrition (Morland et al. 2002, Wrigley et al. 2002), be physically active, and obtain an education that will provide them with critical thinking skills that are useful in maintaining health (Grossman and Kaestner 1994).
Indicators (example)	Number of grocery stores per 1,000 households.
Definition (indicator)	Comprises establishments generally known as supermarkets and grocery stores (U.S. Census Bureau 2007).
Justification (indicator)	As the number of residents per available food store increases, the relative access to food decreases. A gradient in this ratio has been observed according to community SES (Bell and Burlin 1993).
Natural Environment	
Justification (domain)	A core amenity for the healthy functioning of a neighborhood. Greenspace provides beneficial environments for physical and mental health, on both the individual and community level (Diez-Roux 2001, Van Herzele and Wiedemann 2003).
Indicators (example)	Accessibility (average distance to all parks or greenspace, measured from weighted centroid of population).
Definition (indicator)	A city (or other political-level) designated area of open, publicly usable space provided for recreational use, usually infrastructure, measured by unit of population.
Justification (indicator)	Access to parks provides recreational opportunities (Douglas 2005) and promotes a healthy lifestyle by encouraging walking, cycling, and other leisure activities (City of Ottawa, 2006b).
Social Environment	
Justification (domain)	The social environment provides an outlet and location for community members to care for others, work collectively on social problems, participate in social policy debate, express their values and beliefs, enforce social control, and provide opportunities (Berkman and Kawachi 2000, McNeill et al. 2006).
Indicators (example)	Number of crimes against the person per 1,000 people.
Definition (indicator)	Crimes against the person are classified as abduction, assault, assault on child, breach of conditions, fraud, homicide, homicide attempt, robbery commercial, robbery other, senior abuse. Based on the Canadian Criminal Code Offender Category Definitions.
Justification (indicator)	Increasing evidence points to social cohesion as a vital ingredient for the maintenance of collective well-being, and crime is the mirror of the quality of social relationships among citizens (Kawachi et al. 1999).
Built Environment	
Justification (domain)	Housing is a significant engine of social inequality that has both material and psychosocial dimensions that may contribute to health differences. Housing factors may operate both directly and indirectly to modify the underlying factors shaping health status, such as social support and stress (Dunn 2002, Galea and Vlahov 2005).
Indicators (example)	Affordability: Percent renters/owners paying more than 30 percent of household income on shelter.
Definition (indicator)	Affordability was defined by Canada Mortgage and Housing Corporation as costing less than 30 percent of total cost before-tax household income. Depending on the source of the statistics, this also can include cost of utilities.
Justification (indicator)	Community affordability has been listed as a measure of quality of life (Seasons 2003). When families are forced not only to meet, but often to far exceed, standard spending on housing, other important needs suffer, such as food, health care, and insurance, as well as family activities that provide exercise and emotional stability (Bashir 2002).

Table 2.3 Justification of domains and one example indicator for each domain is defined and justified.

a)



b)

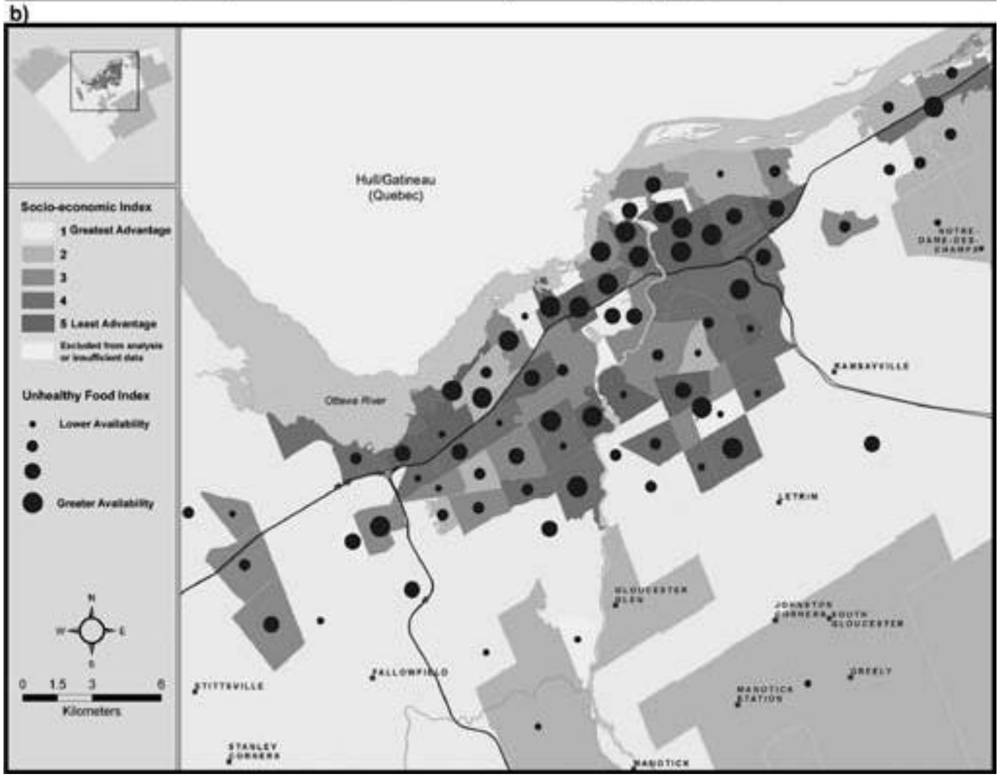


Figure 2.6 (a) City of Ottawa socioeconomic status (SES) and GIS derived recreation index; (b) same with unhealthy food index.

Chapter 3 The Role of Spatial Representation in the Development of a LUR Model for Ottawa, Canada

Marie-Pierre Parenteau[(1)]
Michael C. Sawada[(1)]

[(1)] Department of Geography, University of Ottawa

Corresponding author: Michael Sawada
Laboratory for Applied Geomatics and GIS Science, Department of Geography, University of Ottawa, Ottawa, Ontario K1N 6N5

Tel:+1 613 562-5800 Ext. 1040; Fax: +1 613 562-5725

Published in: Air Quality, Atmosphere and Health, Article in press, pp 1-13.
DOI: 10.1007/s11869-010-0094-3

3.1 Introduction

Modeling chronic air pollution exposure to constituents like nitrogen dioxide (NO_2) at an intra-urban scale is fundamental for health planning and intervention within cities. The land-use regression (LUR) model was first introduced in 1997 by a team of European researchers (Briggs et al. 1997) but it was not until 2005 that a first attempt at using this methodology in North America was tried (Gilbert et al. 2005). Since then, LUR models have been developed for only a limited number of large centers in Canada (Gilbert et al. 2005; Henderson et al. 2007; Jerrett et al. 2007; Marshall et al. 2008; Su et al. 2008; Wheeler et al. 2008; Sahsuvaroglu et al. 2009; Poplawski et al. 2009). The modeling of air quality based on LUR requires accurate data on a number of human and environmental factors such as land-use, street networks, location of greenspace and population distribution. Each of these variables can, and has been, integrated within LUR models using a wide variety of spatial representations and spatial scales. As the number of articles published that employ a land-use regression model has been increasing, a research agenda that focuses on the role of spatial representation and scale in LUR model performance is warranted.

In order to improve LUR model development, our principal objective is to examine the role of spatial representation of the LUR independent variables used to model atmospheric NO_2 concentrations. A second objective of this research aims at developing a reasonably accurate LUR model for Ottawa, Canada, which has not been attempted before. Ottawa is often considered a unique city because of its small manufacturing base and large government and technology sector activities and so developing a LUR for this city is challenging considering the size of Ottawa and the low industrial activity found within its boundaries (Wentz et al. 2002; Jerrett et al. 2007). The use of a population independent predictor is the key element that many published European and North American LUR models have in common. There are numerous ways in which population has been represented in LUR modeling efforts and these commonly include the number of dwellings per unit area, the population count per unit area and population density (Henderson et al. 2007; Ryan et al. 2008; Beelen et al. 2009). For the most part, operationalizing the population variable is achieved by using available census data at different geographic levels. The use of multiple population representations found in the learned literature begs the question: how robust are LUR results to the use of different population representations at different spatial scales? Since the question of spatial representation is a fundamental consideration, the modifiable areal unit problem (MAUP) (Openshaw 1984) in population representation is a concern (Andresen and Brantingham 2008). With regard to our present LUR undertaking, we have tried to limit our scope by looking more specifically at the spatial representation of the population variable and its ensuing effects on LUR model output performance. We thus address the question: how robust are LUR models to different population representations as independent variables?

To our knowledge, no research has yet studied the role of spatial representation in the development of a LUR model, more specifically for the population variable. With this research we propose to address this issue for the first time by developing regression models based on three different representations of population from the Canadian Census of Population: the population count at the dissemination area (DA) level, the population count at the dissemination block (DISB) level and the population count at a sub-dissemination block level using dasymetric mapping (DASYM).

The need for data integration arises when one wants to use data collected under a different spatial division (e.g., non census tract or non dissemination area level but a finer or custom geographic boundary set) than the one used by the census (Fisher and Langford 1996) or when wanting to understand or intervene at a scale that is finer than that collected by the census. This would be the case when the goal is to examine natural socioeconomic processes that are indifferent to the imposed non-physical boundaries. As such, spatial units are often incompatible with respect to the required or intended needs of the researcher and so areal interpolation techniques are required (Langford 2007). Solving this problem of incompatibility requires the assignment of one aggregated dataset to another incompatible dataset using various available spatial algorithms (Sadahiro 1999; Mennis 2003; Reibel and Bufalino 2005; Reibel and Agrawal 2007; Langford 2007). The approaches developed to solve the problem of incompatible spatial units have the capability of generating a more precise map of population distribution or many other census derived variables. Dasymetric mapping, which can also be pycnophylactic (Tobler 1979) in nature, is the spatial interpolation method used in this research. It is a method that is based on the integration of ancillary spatial data. Ancillary datasets like roads, greenspace, water, land-use, cadastral data can help to define both where people could live as well as where they cannot live within a predefined area. As such, a dasymetric approach provides a method by which the original dataset representing, for example, population counts in census tracts can be disaggregated and redistributed to a finer spatial scale. The use of this approach also corresponds to the last goal of this research, which is to work toward the development of a standardized methodology for dasymetric mapping (Langford and Higgs 2006).

We use dasymetric mapping in the context of LUR but it could also have numerous other applications where population data at fine spatial resolutions are required. For example, the availability of an accurate representation of the population distribution for governance (*the governance of oneself and of others*) is a very important task for administrating services (Crampton 2004). It can be argued that an accurate map of the human population is essential to municipal planning; even more so for public health planning and healthcare provisions (Hay et al. 2005). Global disaster management for the developing world has given rise to projects like LandScan (Bhaduri et al. 2007) and others that are producing dasymetric gridded global population estimates at fine spatial resolutions that compare well with known population distributions in the developed world (Sutton et al. 2003; Sabesan et al. 2007; Patterson et al. 2009). It is clear that accurate data on the spatial distribution of population is fundamental to a number of endeavors (Liu et al. 2008) but few studies have focused on the question of spatial representation in terms of population distribution and its impact on policies. One of the few examples of research on the subject is the work of Langford and Higgs (2006) who investigated the influence of alternative spatial population representations to the measure of potential access to primary healthcare services. The authors found that the modeling method for population impacted the results. The authors concluded that the use of dasymetric mapping consistently provide lower estimates of accessibility to healthcare, which in terms of policy and planning could have a significant impact. Research in the field of environmental justice has also started addressing the question of using alternative population representations (Most et al. 2004; Brindley et al. 2005; Mohai and Saha 2006; Mohai and Saha 2007). Hence, this research will also contribute to the advancement of these other fields of studies.

In Europe and North America, LUR modeling efforts to map exposure to NO_2 have performed generally well with R^2 values varying from approximately 0.5 to 0.9. In general Canadian research has yielded acceptable results, with R^2 values between 0.54 and 0.77. Even though the predictor variables have been generally the same for most studies, their specifications have been significantly different (Jerrett et al. 2007; Ryan and LeMasters 2007). Hence, not only it is very important to understand the sensitivity of the models to spatial representation in order to obtain consistency in the results but it is also very important as these exposure models are in most cases one of the first steps in the study of the relationship between exposure and health. Exposure models with an improved spatial resolution have been found to produce more robust associations (Sahsuvaroglu et al. 2009) when compared to health conditions. We expect dasymetric mapping will aid in obtaining better LUR models. Hence, better LUR models will contribute to the literature on the relationship between exposure and health. Still, the main objective of this research is to give a first approximation of the role of spatial representation in the development of LUR models and contribute to the research agenda on the role of spatial representation more generally.

3.2 Land-Use Regression Model

Land-use regression models were developed for application in European cities and research on the subject was first published by Briggs et al (1997). Those authors were interested in exposure models that would allow the study of the relationship between health and air pollution at a local scale. Work by these authors also corresponds to the period when academia started to be interested in the spatial distribution of pollutants, not only in the context of inter-urban studies but also for intra-urban studies (Jerrett et al. 2005b).

The general approach to modeling aims at predicting concentration (NO_2 or other pollutants) at an arbitrary location using observed concentration at selected sampling sites together with a number of spatial explanatory variables characterizing the environment in the proximity of those sampling sites (Jerrett et al. 2005a; Ryan and LeMasters 2007; Hoek et al. 2008). LUR is hence based on two important principals: "1) environmental conditions for the variable of interest can be estimated from a small number of readily measurable predictor variables 2) the relationship between the target variable and these predictors can be reliably assessed on the basis of a small sample survey or training area" (Briggs et al.

2000). Among the main advantages of this approach are the facts that it can easily be adapted to the environment in which the study is taking place and that the development of such a model is less costly than increasing the observed sampling network (Jerrett et al. 2005a).

Over the last few years a number of North American studies have been published and the results of these studies confirm the potential of LUR (Kanaroglou et al. 2005; Gilbert et al. 2005; Jerrett et al. 2007; Wheeler et al. 2008; Jerrett et al. 2007). An element that makes the development of a LUR model for Ottawa, Canada challenging is the reality that most published studies took place in large urban centres (Hoek et al. 2008). The City of Ottawa, part of Ottawa-Gatineau census metropolitan area (CMA) (Ontario part) is a large Canadian city with a population count of 812,129 in 2006 (Statistics Canada 2007a), but it is not comparable to large European and American cities in terms of population. Ottawa also lacks significant industrial activities, and hence air quality is relatively good.

In a literature review looking at the published LUR models, Hoek et al. (2008) found that a number of explanatory variables were more frequently used then others. Variables for population, traffic, land-use, altitude, meteorology and location tended to be included in the LUR models. In addition, Hoek et al. (2008) found that problems frequently associated with spatial data, such as completeness and precision, have not been addressed in an appropriate manner within LUR methodologies. In particular, the role of spatial representation of the explanatory spatial variables in LUR modeling requires more attention.

3.3 Areal Interpolation

When data are compiled according to different geographic boundaries, one is faced with a problem of geographic mismatch. Areal interpolation methods have been developed as a solution to this problem. Different methods of areal interpolation are available, the majority of which are based on a spatial overlay algorithm, and each method is based on a number of different assumptions. Areal interpolation methods are generally classified into two broad categories: 1) techniques not using ancillary data; 2) techniques using ancillary data. In this article dasymetric mapping, which falls in the category of techniques that use ancillary data, is considered. Dasymetric mapping can be implemented without a lot of additional data (Hay et al. 2005) and has the ability to provide results that are closer to the underlying spatial processes (Langford et al. 2008).

3.3.1 Dasymetric Mapping

Dasymetric mapping is an areal interpolation method that takes advantage of ancillary data and so considers the heterogeneous distribution of a phenomenon through space. Poulsen and Kennedy (2004) describe dasymetric mapping as “a technique that involves estimating the distribution of aggregated data within the units of analysis, by adding additional information that provides insights on how these data are potentially distributed”. Through the use of ancillary data, dasymetric mapping is capable of creating subzones with a higher level of homogeneity than the initial source zones, and so more accurately represents the underlying geographic patterns of population when compared to the finest available enumeration areas (Fisher and Langford 1996; Mennis 2003; Holt et al. 2004; Langford and Higgs 2006; Fiedler et al. 2006; Mennis and Hultgren 2006). To date, this approach has relied in most cases on classified multi-spectral satellite imagery as a source of ancillary data for the identification of non-residential discontinuities such as parks and cemeteries in the urban environment (Reibel and Bufalino 2005; Langford and Higgs 2006; Langford 2007; Langford et al. 2008). The general reliance on remotely sensed imagery may have been an obstacle to the implementation of this areal interpolation method as it requires expertise in the field of remote sensing (Langford 2007) and the cost of image acquisition. An alternative to the use of satellite imagery, and raster data more generally, has been proposed. A non-image based technique uses street network data (Reibel and Bufalino 2005) and has been successfully implemented in a few studies. In its simplest form, dasymetric mapping distinguishes between inhabited and uninhabited areas, which is called binary dasymetric mapping (Langford and Higgs 2006; Mennis and Hultgren 2006). It is the traditional approach to dasymetric mapping and has proven to perform better than area weighted interpolation or other methods of areal interpolation that do not exploit ancillary data. The second type of dasymetric mapping is a polycategorical approach where three or more population density (land-use) classes/categories are defined and is considered a more advanced and complex method to implement (Langford 2007; Eicher and Brewer 2001).

For the modeling of air quality, dasymetric mapping has been used only in research by Beelen et al. (2009). This study is quite different from the one presented here as it is interested in the possibility of modeling air quality over a large geographical area, namely the European Union using the LUR. Dasymetric mapping was used to refine population distribution from NUTS-5 level using CORINE land cover and light emissions data. Conceptually our work resembles the research by Beelen et al. (2009) but our goal is not to determine how well the LUR can perform at a large scale using data at a coarse spatial resolution but rather how using different spatial representations, all at a fairly high resolution, will have an impact on the performance of the LUR.

3.3.1.1 Sources of ancillary data

As mentioned, most implementation cases of dasymetric mapping have been based on the use of satellite imagery as a source of ancillary data. A study by Fisher and Langford (1996) showed that population estimates by dasymetric mapping based on Landsat imagery are robust. Given the wide availability of free Landsat imagery their approach is feasible if the analyst has the remote sensing expertise and the available imagery is sufficient. However, in their research the authors found that dasymetric mapping still outperforms other areal interpolation techniques in those instances where low image classification accuracy is involved. In fact, they found that it takes classification accuracy as low as 60% for other areal interpolation methods to perform better. Where there is a lack of ancillary data, a remote sensing approach is necessary and will provide a best estimate.

An alternative to the use of raster data for dasymetric mapping is the use of street network data (Xie 1995), making this approach more accessible to analysts who may not have the raster GIS skills (Reibel and Bufalino 2005). This use of the vector data is based on the assumption that population distribution is closely linked to the configuration of roads across the landscape (Hawley and Moellering 2005). Results of research show that the use of the street network data, which is a simple solution to the problem in comparison to the use of satellite imagery, provides better results than areal weighted interpolation (Xie 1995; Reibel and Bufalino 2005). Since areal weighting redistributes population based solely on the intersection of the target zone and the source zone, a homogeneous distribution of the population is assumed. The methodology implemented in this research takes into account both this aspect of the problem but adds another component that allows us to account for the differences in population density of the target zones.

3.4 Methods

3.4.1 Study Design

We developed a LUR model for the mapping of NO_2 concentrations in Ottawa, Canada. Our focus is on the role of different spatial representations on model performance. Air pollution data were obtained from the City of Ottawa through the National Capital Air Quality Mapping Project. In order to develop the LUR model, a spatial database that contained information on land-use, roads, population, zoning, greenspaces and elevation was created. This database was also used for the mapping of population using the dasymetric mapping approach and will be discussed in the following sections.

3.4.2 Data

The National Capital Air Quality Mapping Project was launched in the fall of 2007 by the city of Ottawa with the help of Environment Canada and Health Canada. Through this project 30 Ogawa passive samplers were installed throughout the city of Ottawa on three different occasions, each time for a period of two weeks (Figure 3.1). Unfortunately, due to logistical problems only the data collected from 29 samplers over the period of 29 September 2007 to 13 October 2007 could be used in the development of the LUR model. The monitor removed from the analysis (AQMP-09) was located in Gatineau, Quebec. The information collected at this location would have had a minor impact on the variation of the NO_2 concentrations (output of LUR model) in Ottawa. The monitor was located in a quiet residential area, very similar to the environmental setting of AQMP-17 and AQMP-26. As a consequence, removing AQMP-09 did not translate into the loss of key information.

The samplers measured the concentration in NO_2, O_3 and SO_2 at each location. Data collected at two permanent stations were also added to the dataset. Even though data on O_3 and SO_2 were available, only the NO_2 modeling is part of this ongoing research as it is an essential marker for traffic-related air pollution and strongly correlated with other common air quality indices. This pollutant has been used as a proxy to evaluate exposure to traffic emissions (Jerrett et al. 2007).

Spatial data from Statistics Canada are made available to Canadian universities through the Data Liberation Initiative (Statistics Canada 2009). For this particular reason, the 2008 street network file from Statistics Canada was used as it represents the 2007 streets in Ottawa. The 2006 dissemination block spatial boundary file from Statistics Canada was also used as it corresponds to the most recent Census of Population in Canada. The dissemination block is a spatial unit defined on all sides by roads and/or boundaries and is the smallest geographic area for which population and dwelling counts are publicly reported (Statistics Canada 2007d). Another important dataset used in the current research is the zoning data obtained from the city of Ottawa. The original dataset contained 39 zoning classes/categories which were reclassified into 23 classes according to the permitted uses for each of the main zoning types as described in the city of Ottawa Zoning By-law 2008-250 Consolidation (Table 3.1).

The zoning data eliminate the need to use remote sensing data, which in many cases is an obstacle to the implementation of dasymetric mapping (Langford 2007; Reibel and Agrawal 2007). Zoning data provide information that allows the identification of uninhabited areas as well as detailed information on the different levels of population density throughout the region. We attributed a population count value of zero to zoning types where residential habitation use is not allowed and the other zoning types were set aside for the evaluation of the population density fraction (PDF). The calculation of the PDF is one of the main steps of dasymetric mapping.

3.5 Dasymetric Mapping

Dasymetric mapping was undertaken at three scales, first at the scale of the dissemination block, second using the zoning data and third, the fusion of both datasets provided us with the necessary information to further refine the spatial resolution of the Census population data.

Because the zoning data were reclassified into 23 classes of land-use, the dasymetric approach used in this research is polycategorical. This approach should represent an improvement in comparison to the traditional binary approach where information on the different levels of population densities within a census reporting zone are not accounted for (Langford and Higgs 2006). Our methodology also facilitates the implementation of dasymetric mapping, requiring no data or knowledge of remote sensing. This approach also has the advantage of overcoming the misclassification problem that can occur between residential apartment buildings and commercial facilities (Langford 2007).

The methodology used in this research is based on the work of Mennis (2003). Mennis' approach has four main steps: 1) calculation of the PDF (population density fraction); 2) calculation of the area ratio (AR); 3) calculation of the total fraction (TF); and 4) the mapping of population (or other variable). The main difficulty with the implementation of the polycategorical approach is the identification/calculation of the different relative ratios of the PDF values that are assigned to each population density class/category (land-use class). Our approach to overcoming that issue is accomplished using selective sampling within the dataset (for more details see Mennis 2003). The Mennis' method was selected for its simplicity, as one of our goals is to define a methodology that can be implemented by local governments without the need for considerable expertise. The down side of this approach is that it requires that some source zones are totally nested into one single population density class in order to be able to calculate the relative ratios of population density values (Mennis 2003; Langford and Higgs 2006; Reibel and Agrawal 2007). For example, in this study the source zones were the dissemination blocks. Hence, it was necessary to have at least several blocks that were completely contained within one single land-use class for the calculation of the PDF. This approach was first implemented by Mennis (2003); alternative approaches are found in Eicher and Brewer (2001) and Yuan et al. (1997). These may not have the problem associated with selective sampling but present other issues such as vulnerability to outliers and incorrect sampling in the case of the centroid method (Mennis and Hultgren 2006).

The use of dasymetric mapping allowed us to refine the census geography, defining a smaller geographic scale than the dissemination block. Considering that population data are not disseminated at a lower geographic level than the dissemination block, our methodology represents a solution for the production of a population map at a sub-census unit level. For example, with the use of dasymetric mapping, we were able to split a dissemination block into a number of parts (Figure 3.2). Some of these parts, for example, do not allow for residential land-use and so they were attributed a population value of zero (see Table 3.1), and the total population count was reassigned within the larger census zone to the new sub-units according to the PDF and the AR.

3.6 LUR

The measured levels of NO_2 at the sampling locations (dependent variable) were used in a multiple least-squares regression model (Ryan and LeMasters 2007). Multiple explanatory variables were identified based on the literature, with an emphasis on the representation of the population variable. As such, population variables were created based on the population data at the dissemination area (DA), the dissemination block (DISB), and the refined dissemination block using dasymetric mapping (DASYM). Buffers ranging from 50 - 2000 m at an interval of 50 m were created for each of the sampling sites. At each interval, each explanatory variable was compared with NO_2 concentrations (40 comparisons per covariate). The use of the series of buffers provided a mean by which to determine an optimized distance for each spatial covariate (Su et al. 2008). As NO_2 concentrations are known to fluctuate over distances as small as 50 m, the use of multiple buffers allowed us to identify the distance at which the measured correlation between the pollutant (mean NO_2 concentrations for the two week period) and the different covariates was the strongest. This preliminary analysis used Pearson's r coefficient of correlation and bivariate regressions to look at the absolute strength of the variables and possible deviations from linear relationships with NO_2 by examination of residual plots (Poplawski et al. 2009). This analysis provided us with initial information on the relationship between the variables and measured NO_2 concentrations as well as a way to confirm the anticipated direction of the effect (positive or negative) (Henderson et al. 2007). For each variable, the distance of maximum correlation was used to define the explanatory variable used in the model building exercise. Because we were undertaking multiple tests (n = 40) for each variable we interpreted statistical significance at a Bonferroni corrected level of $p = 0.05/40$ (Cooper 1968).

All three representations of the population variable had the strongest correlation with NO_2 at a distance of 1750 meters, showing coherence. Pearson's r for the correlation between the measured NO_2 concentration and population at the DA level only has a value of 0.28719 and is not statistically significant. The population at the DISB level and DASYM has similar Pearson's r values for the same relationship, with the dasymetric mapping having a slightly higher level (0.78968 versus 0.78893). However, NO_2 and DISB and DASYM are statistically significant as indicated in Table 3.2 but not significantly different from each other.

Since this research is concerned with the role of spatial representation, three different LUR models were developed. The first model included a population variable derived from DASYM, the second model's population variable was derived by the use of the data at the geographic level of the DISB and the third model had population derived from the DA level. Stepwise selection was the method used to develop the different regression models (Jerrett et al. 2007; Su et al. 2008). Variables in the model had to be statistically significant at $p = 0.05$ and models were evaluated based on the R^2 values, RMSE, variation inflation factor (VIF) eigenvalues and the condition index (CI). The methodology used for the development of all the regression models was the same. For example, to develop the model using population at the DA level, DISB population and DASYM population representations were excluded from the list of covariates. All models were tested for spatial autocorrelation in the residuals using Moran's I statistic (Bailey and Gatrell 1995) and no significant spatial autocorrelation was present.

The model containing the DASYM population variable (model 1) provided an R^2 of 0.8055 (Table 3.3) while the model using the DISB variable provided a very similar value of 0.8038 (model 2). Using the population data at the geographic level of the DA (model 3) did perform poorly, yielding a result of R^2 of 0.6962, much lower than model 1 and model 2. Study of the correlation between the variables and NO_2 concentration as well as the development and evaluation of the regression models was achieved in the environment of SAS Enterprise (version 3.0).

The two models using the finest levels of aggregation as spatial representations of the population variable are very similar in terms of the variables included in the models, parameter estimates and statistics to evaluate the performance of the models. The model making use of spatial representation based on the data at the DA level had a different composition in terms of covariates with the Industrial_250m variable being excluded from the model. Most importantly, the population variable included in the DA model, PopDA_1750m, was not statistically significant at 95% level.

As mentioned earlier, all three model residuals were tested for spatial autocorrelation and none showed significant spatial dependency. As such, we are not concerned that spatial autocorrelation in the residuals will affect the size and significance of risk estimates from this analysis as others have found (Jerrett and Finkelstein 2005). Graphical representation of the residuals does not indicate presence of trends in any of the three models (Figure 3.3). Once again, plots of the residuals of model 1 and model 2 (Fig. 3.3a and 3.3b respectively) are very similar but the residual plot of model 3 (Fig. 3.3c) is clearly different.

3.7 Model Validation

Validation of the LUR model was achieved using the leave-one-out cross validation (LOOCV) approach (Hoek et al. 2008; Mukerjee et al. 2009). Using this methodology the regression model was re-estimated 31 times, each time leaving out one observation (n-1). The root-mean-square error and the mean absolute error were computed in order to assess the performance of the model. The results (RMSE = 1.05 ppb and MAE = 0.86 ppb) indicate that the NO_2 values predicted by the model are reliable.

3.8 Discussion

Our results indicate that the accuracy of the land-use regression model is affected by the use of different spatial representations of the population variable. The mapping of population at a sub-block geographic level using the dasymetric approach showed no significant differences over using population at the dissemination block. On the other hand both of these models outperformed the model based on dissemination area population data. As such, LUR model performance increases with mapping population distribution at a lower geographic level than the dissemination area. In addition, with the more detailed population estimates other socioeconomic and demographic covariates can potentially be better focused spatially to where the population resides.

With an R^2 of 0.8055, 0.8038 and 0.6962, all three LUR models are within the range found in the literature (Hoek et al. 2008).These values also demonstrate that it is possible to use the LUR model in a smaller city where the air quality is often considered generally good. This finding for Ottawa is important as it enlarges the applicability of the model to a number of additional centres, especially in Canada - where the number of large urban centres is limited. The NO_2 surface created through the implementation of the LUR model corroborates *a priori* knowledge of the air quality in Ottawa with pockets of higher NO_2 concentration located in areas known to have air quality problems. The variables included in the LUR model for Ottawa (length of major roads, population count, area of greenspace, and industrial land-use) are for the most part from the same categories as the ones included in the models developed for other Canadian cities. The size of the buffers is also within the same range as the ranges found in the literature on Canadian LUR models (Henderson et al. 2007; Jerrett et al. 2007; Marshall et al. 2008; Wheeler et al. 2008; Sahsuvaroglu et al. 2009). The main difference we found when building the LUR model for Ottawa in comparison to the other Canadian cities is the importance of the variable “Greenspace”. The only other model with similar variable was published by Sahsuvaroglu et al. (2009) where a LUR model for Hamilton, Ontario is developed using a variable “Open land-use”. The presence of greenspaces is an important feature of the landscape in Ottawa, explaining why it is not found in most other LUR models developed for Canadian cities.

The City of Ottawa has a unique landscape created by the existence of a large corridor of greenspace that represents the boundary between the urban and the suburban part of the city (Figure 3.4). Included in the planning of the city in 1950 by Jacques Greber, a French planner, the purpose of the 'Greenbelt' was to prevent urban sprawl and protect the rural land surrounding the city (National Capital Commission 2009). Due to population growth, residential and commercial developments were constructed beyond the greenbelt. The location of the Greenbelt can easily be identified in the LUR NO_2 concentration surface map; it corresponds to the large linear area where the levels of NO_2 are considerably lower than the surroundings. Hence, the existence of the Greenbelt potentially has a strong effect on residents' health.

We believe the combination of a number of factors is behind the fact that the use of dasymetric mapping did not significantly improve the results of NO_2 modeling. One first factor is potentially found in the level of heterogeneity of the DISBs within Ottawa. In general, population distribution within a particular dissemination block is highly homogeneous. As a consequence refinement of population distribution is not significant enough to considerably improve the results of the modeling.

Also, the scale at which NO_2 concentration varies is of importance. Refinement of the map of population distribution at a geographic level lower than the dissemination block may be much finer a scale than at which we can observe NO_2 variations. For this reason, spatial representation does not appear to be playing an important role when comparing population counts at the block level and dasymetric population. On the contrary the role of spatial representation is significant when comparing LUR with block or sub-block level data and DA level data.

Lastly, the poor contribution of dasymetric mapping to the LUR may be due to the fact that population count (all representations) is highly correlated to NO_2 concentration at a distance of 1,750 metres from the sampling sites. In a study on the evaluation of population at risk using dasymetric mapping, Higgs and Langford (2009) conclude that “as [the] buffer size increases with respect to census tracts, partial intersections proportionately decrease and correspondingly the results [of different population representations] tend to converge”. This suggests that if the strongest correlation of population with NO_2 would have been measured at a smaller distance, potentially the use of dasymetric mapping could have considerably improved the results. Such a hypothesis begs testing in this urban area by using data in different seasons for example.

Our results indicate that the use of dasymetric mapping to refine the population data at the sub-dissemination block level is possible. In terms of model improvement, the result is not significantly different from using block level population representations and we could not recommend the use of this approach when population data are available at a fine geographic level (as the dissemination block). On the other hand, we perceive this methodological approach to be of great value not only to redistribute population aggregated at a higher geographic level but also for a number of other applications. In addition, our approach would be beneficial for local level governments that have access to local zoning data but who may not have access or who do not want the added expense of utilizing block level population in Canada and elsewhere in urban areas where block level population may not be collected.

The main limitation of this study is the City of Ottawa's geography. Ottawa is located on the Ontario side of the Ottawa River, representing only one side of the Ottawa-Gatineau CMA. The other side of the Ottawa River is the City of Gatineau, located in the province of Quebec. The data used in the modeling of the NO_2 with the LUR model were acquired from the City of Ottawa and do not cover the City of Gatineau. An air shed does not follow political boundaries and it is realistic to assume that the integration of the Quebec portion would have resulted in a more accurate model.

Another limitation of this research is the fact that the NO_2 data used were collected through a limited air sampler deployment. Considering that Ottawa is characterized by large variability in temperatures measured at different times of the year, model improvements could certainly be made with NO_2 measurements taken during different seasons. As stated earlier, seasonal variations in the spatial correlation of NO_2 with covariates might improve model predictions. However Wheeler et al. (2008), in a study of the correlation between NO_2 measurements taken throughout different Canadian seasons with yearly averages, found that all seasons were highly correlated with the annual average. The problem would then be one of overestimation or underestimation, depending on the season at which measurements were taken, and would not affect the general spatial distribution of NO_2 concentrations.

Also, the location of the samplers did not include rural sites. As a large proportion of the territory of the city of Ottawa is rural, having sampling sites in these areas would have potentially improved the models beyond the urban core and in the suburbs. The location of the samplers may be more important than the number of samplers deployed (Ryan and LeMasters 2007). However, such was out of our control but should be considered in the future deployments by the city. Lastly, the importance of the temporal period of sampler deployment and resultant effects on NO_2 modeling in regard to the selection of covariates need attention and is beyond the scope of this study.

3.9 Conclusion

As the land-use regression model grows in popularity as a simple modeling approach, we believe that it is time that the research community focuses on the question of spatial representation and the impact on the mapping of pollutants. The secondary goal of this research was to develop a LUR model for Ottawa. Our results indicate that it is possible to develop such a model for a smaller urban centre where air quality is considered to be good.

Our results point to the fact that the output of LUR modeling for NO_2 mapping is affected by the use of different spatial representations of population. Where we had expected the LUR model using results of dasymetric mapping to outperform the one using block level population counts, we did find essentially identical results.

This confirms the work of Higgs and Langford (2009) who found that the best results with dasymetric mapping arise at small buffer sizes. Consequently we recommend the use of dasymetric mapping for LUR models if the strongest correlation between NO_2 and population is found in close proximity of the air quality sampling sites. Moreover, the availability of zoning datasets by municipalities can easily replace the need for block level population in cases where such data are not available or too costly to acquire. In general the use of zoning data obtained in a vector format did allow the refinement of population distribution with the dasymetric mapping methodology. This also has the advantage of easy adoption by local governments for the creation of population maps at a sub-block level without the need to use remote sensing data.

Acknowledgements

The authors would like to thank Natividad Urquizo from the Community Sustainability Department, City of Ottawa, for providing the air quality data.

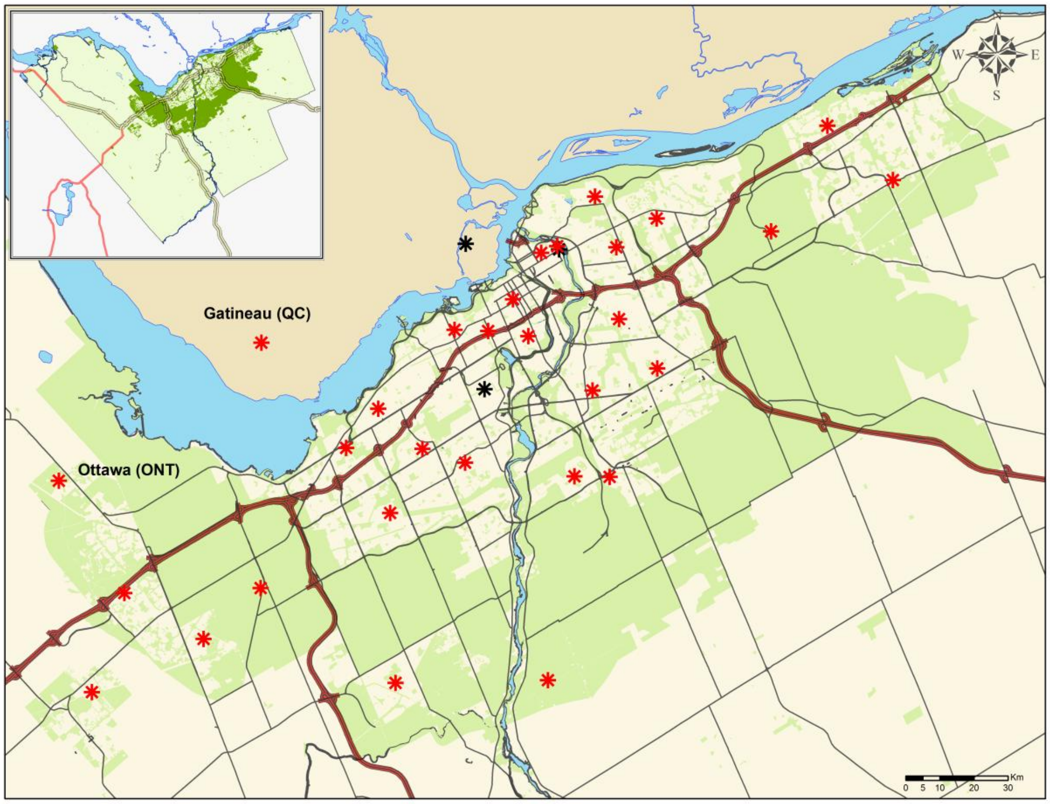


Figure 3.1 Distribution of the passive samplers (red) and permanent stations (black) in Ottawa.

Original Zoning Type	New Zoning Type	New Definition
Agricultural Zone	AG	Agricultural Zone
Arterial Mainstreet Zone	AM	Arterial Mainstreet Zone
Development Reserve Zone	DR	Development Reserve Zone
Environment Protection Zone	EP	Environment Protection Zone
General Mixed Use Zone	GM	General Mixed Use Zone
Minor Institutional Zone	NoPop	No population
Major Institutional Zone	NoPop	No population
General Industrial Zone	IG	General Industrial Zone
Heavy Industrial Zone	NoPop	No population
Light Industrial Zone	IL	Light Industrial Zone
Business Park Industrial Zone	IP	Business Park Industrial Zone
Community Leisure Facility Zone	NoPop	No population
Major Leisure Facility Zone	NoPop	No population
Central Experimental Farm Zone	NoPop	No population
Local Commercial Zone	R3	Local Commercial Zone
Mixed Use Centre Zone	MC	Mixed Use Centre Zone
Mixed Use Downtown Zone	MD	Mixed Use Downtown Zone
Mineral Extraction Zone	ME	Mineral Extraction Zone/ Mineral Aggregate Reserve Zone
Mineral Aggregate Reserve Zone	ME	Mineral Extraction Zone/ Mineral Aggregate Reserve Zone
Parks and Open Space Zone	O1	Parks and Open Space Zone
Residential First Density Zone	R1	Residential First Density Zone
Residential Second Density Zone	R2	Residential Second Density Zone
Residential Third Density Zone	R3	Residential Third Density Zone
Residential Fourth Density Zone	R4	Residential Fourth Density Zone
Residential Fifth Density Zone	R5	Residential Fifth Density Zone
Rural Commercial Zone	NoPop	No population
Rural General Industrial Zone	NoPop	No population
Rural Heavy Industrial Zone	NoPop	No population
Rural Institutional Zone	NoPop	No population
Mobile Home Park Zone	RM	Mobile Home Park Zone
Rural Residential Zone	RR	Rural Residential Zone
Rural Countryside Zone	RU	Rural Countryside Zone
Air Transportation Facility Zone	NoPop	No population
Ground Transportation Facility Zone	NoPop	No population
Traditional Main Street Zone	TM	Traditional Main Street Zone
Village Mixed Use Zone	VillRes	Village Residential
Village Residential First Density Zone	VillRes	Village Residential
Village Residential Second Density Zone	VillRes	Village Residential
Village Residential Third Density Zone	VillRes	Village Residential

Table 3.1 Aggregation of the zoning classes according to the city of Ottawa Zoning By-law 2008-250 Consolidation.

Figure 3.2 Example of dasymetric mapping using zoning data. The population was redistributed (b) within census zone (DISB) shown in (a) as some of the zoning types within the selected DISB do not allow residential use.

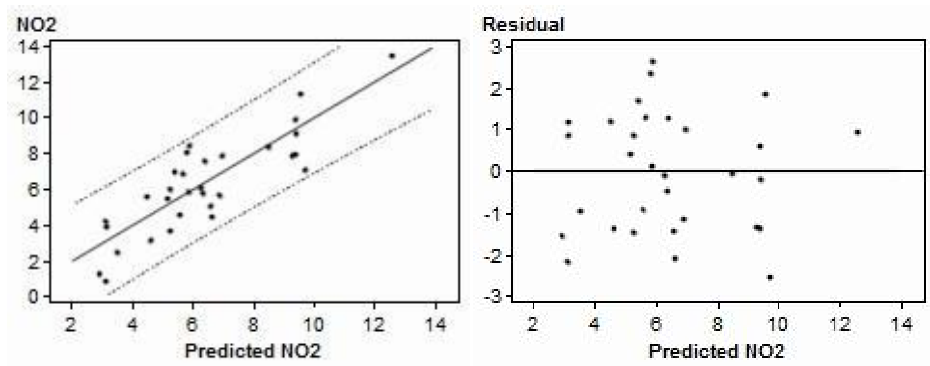
Spatial Representation	Pearson's r	Prob>|r|
DA	0.28719	0.1141
DISB	0.78893	<0.0001
DASYM	0.78968	<0.0001

Table 3.2 Correlation between NO_2 concentration and the different population representations.

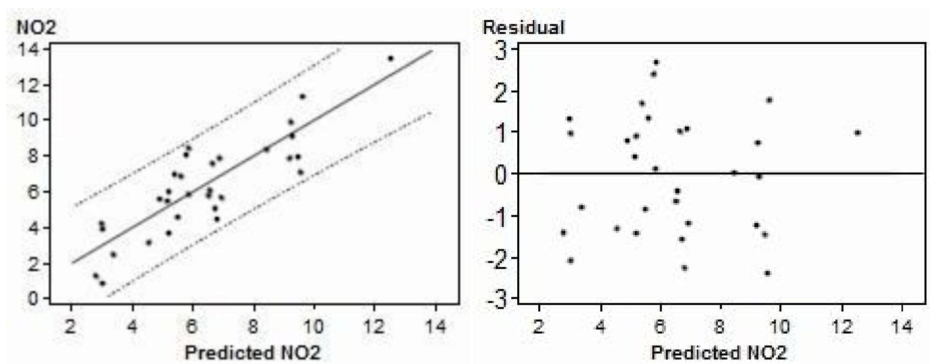
Variable	estimate	Pr > t	VIF
Model 1 (R square=0.8055)			
Intercept	3.66507	0.0027	0
Major Roads 250m	0.00041824	0.0116	1.47569
Popupation Dasymetric 1750m	0.00010586	0.001	2.24357
Greenspace (land use) 700m	-0.00000216	0.0229	2.39553
Industrial (land use) 250m	0.00005783	0.0398	1.28638
Model 2 (R square=0.8038)			
Intercept	3.49419	0.0058	0
Major Roads 250m	0.00041876	0.0119	1.47878
Populatio Dissemination Block 1750m	0.00010492	0.0011	2.35903
Greenspace (land use) 700m	-0.00000206	0.03046	2.53099
Industrial (land use) 250m	0.00002917	0.037	1.29209
Model 3 (R square=0.6962)			
Intercept	7.18141	<0.0001	0
Major Roads 250m	0.00065638	0.0009	1.2866
Populatoin Dissemination Area 1750m	0.00001609	0.2665	1.26329
Greenspace (land use) 700m	-0.00000488	<0.0001	1.02323

Table 3.3 Definition of the LUR models using different population representations. Probability (Pr) and variation inflation factor (VIF) values associated with each of the variables.

a)



b)



c)

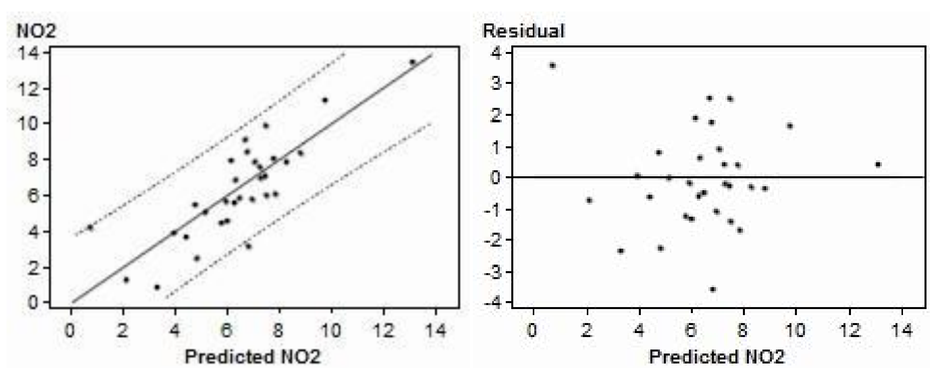


Figure 3.3 Regression (left) and residuals (right) of the three different models: (a) PopDASYM (b) PopDISB (c) PopDA. Predicted values of NO_2 level in ppb vs. measured NO_2 level in ppb (left). Predicted NO_2 level in ppb vs the residual in ppb (right).

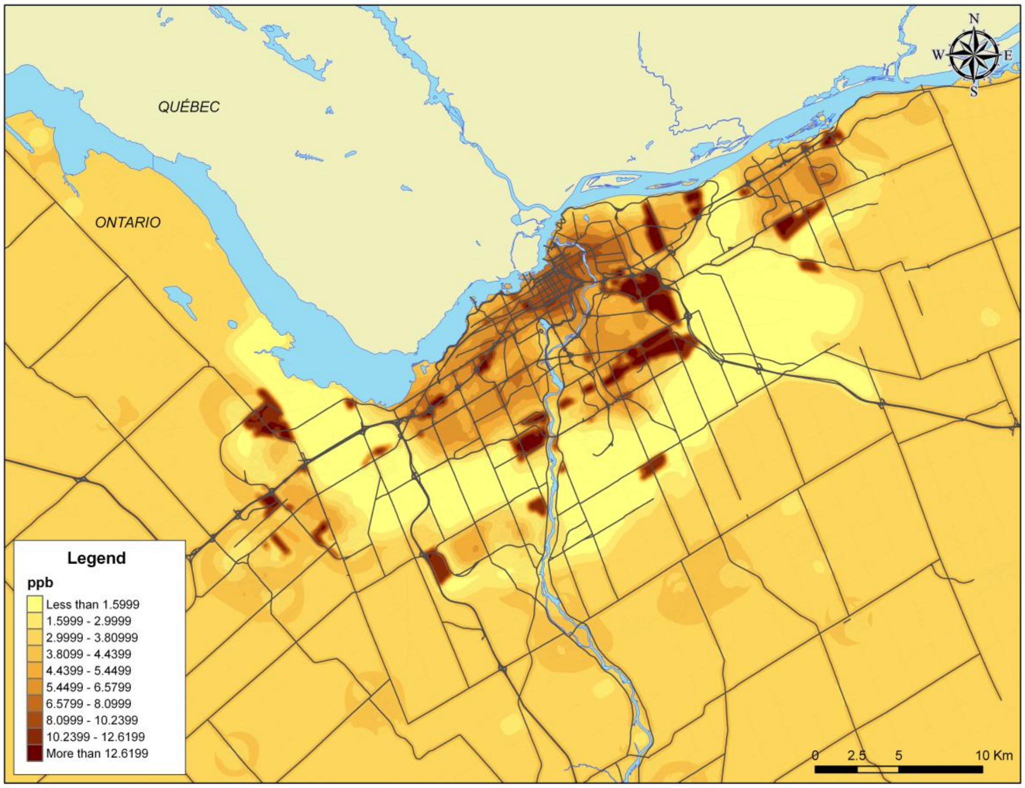


Figure 3.4 LUR for NO_2 concentrations (ppb) in Ottawa.

Chapter 4 Comparison of Methodological Approaches for the Modeling of NO_2 in Ottawa, Canada

Marie-Pierre Parenteau[(1)]
Michael Sawada[(1)]

[(1)] Department of Geography, University of Ottawa

Laboratory for Applied Geomatics and GIS Science, Department of Geography, University of Ottawa, Ottawa, Ontario K1N 6N5

Tel:+1 613 562-5800 Ext. 1040; Fax: +1 613 562-5725

In preparation.

4.1 Introduction

The number of publications on GIS-based approaches to the modeling of air pollutants at the intra-urban level has increased considerably over the last few years with land-use regression (LUR) modeling as the focus of many of these publications (Wheeler et al. 2008; Atari and Luginaah 2009; Mukerjee et al. 2009; Gehring et al. 2010; Mölter et al. 2010; Jerrett et al. 2005b; Moore et al. 2007; Ryan et al. 2008). In many cases such research is a core component in the study of the relationship between health outcomes and exposure to air pollution (Brook et al. 2007; Ryan and LeMasters, 2007; Ryan et al. 2008; Jerrett et al. 2009). Considering the lack of agreement in the results coming out of the studies on the relationship between health and exposure to NO_2 (Villeneuve et al. 2006; Sahsuvaroglu et al. 2009; Clark et al. 2010; Barcelo et al. 2009; Akinbami et al. 2010; Lee and Ferguson 2009) it is plausible that the use of different methodological approaches to the modeling of NO_2 concentrations is a contributing factor in the observed discrepancies.

This study is a component of a broader research project focused on the role of spatial representation in the study of the relationship between exposure to NO_2 and respiratory health. In chapter three, it was found that the way in which the population variable class is represented in a land-use regression (LUR) model has an impact on the results of NO_2 modeling. In this chapter, the three same spatial representations as in chapter three are used as input into NO_2 models: the population distribution based on a dasymetric approach (DASYM), the population at the dissemination block level (DISB), and the population at the geographic level of the dissemination area (DA).

In chapter three, the use of the population representation at the DISB level in the LUR model provided results similar to the ones with the population representation based on DASYM mapping. The LUR model using the population representation at the DA level did not perform as well as when using the population at the DISB or the DASYM levels. If the simple use of different spatial representations of the population variable has an impact on the results of exposure, we hypothesize that the use of a different methodological approach to modeling NO_2 could potentially provide variable NO_2 concentration patterns and as a consequence impact the assessment of the relationship between exposure to NO_2 and health. These differences could perhaps be one of the reasons behind the equivocal relationship between health and exposure to air pollution that has been reported over the years (Jerrett et al. 2005b; Jerrett et al. 2007; Beelen et al. 2008; Jerrett et al. 2008; Sahsuvaroglu et al. 2009). This paper aims to gain a better understanding of the impact of using different air pollution models on the geography of exposure. A secondary goal of this research is to determine if the use of different spatial representations of a variable can, as in the case of the LUR, affect the results of geostatistical interpolation when modeling NO_2 concentrations.

4.2 NO_2 modeling

Two types of GIS-based approaches to the modeling of NO_2 are discussed in the current research: the LUR approach and the geostatistical interpolation approach. Both types of models are considerably inexpensive to build and have been used for NO_2 modeling at the intra-urban level (Jerrett et al. 2005a; Jerrett et al. 2005b; Leem et al. 2006; Hoek et al. 2008). The air pollution data requirements are the same for the two approaches and require direct measurements of the pollutant being modeled.

The LUR model has been used mostly for the modeling of NO_2 but more recently for the modeling of other pollutants such as $PM_{2.5}$, $PM_{10,}$ benzene, toluene, black carbon and woodsmoke (Wheeler et al. 2008; Hoek et al. 2008; Atari and Luginaah 2009; Mukerjee et al. 2009; Mölter et al. 2010; Allen et al. 2010; Crouse et al. 2010; MacIntyre et al. 2010). The LUR approach is based on two important principals: "1) environmental conditions for the variable of interest can be estimated from a small number of readily measurable predictor variables 2) the relationship between the target variable and these predictors can be reliably assessed on the basis of a small sample survey or training area" (Briggs et al. 2000). By using data on land-use, the road network and population (among others) (Moore et al. 2007; Ryan et al. 2008), it is possible to predict local variations of a pollutant concentration more easily than with statistical interpolation of only the pollutant of interest (Jerrett et al. 2007; Arain et al. 2009). In the case of NO_2 modeling, the pollutant is the dependent variable, whereas the independent variables are derived from characterizing land-use and other elements of the environment within buffer zones calculated from the sampling location (Brauer et al. 2002). Recent research suggests that LUR is a valid approach to the mapping of air pollutants and provides good results at a low cost (Su et al. 2009a; Su et al. 2009b). This methodology has been shown to create NO_2 surfaces with local variability for urban centres in Europe, North America (United States and Canada) and more recently in Asia

(Jerrett et al. 2005b; Jerrett et al. 2007; Marshall et al. 2008; Su et al. 2008; Wheeler et al. 2008; Atari and Luginaah, 2009; Sahsuvaroglu et al. 2009; Su et al. 2009b; Crouse et al. 2009; Kashima et al. 2009; Clougherty et al. 2009; Larson et al. 2009).

The second methodological approach used in the research is geostatistical interpolation. Geostatistical interpolation techniques, such as kriging and cokriging, make use of air pollution data acquired throughout a study area in order to create concentration surfaces representing the continuous distribution of an air pollutant in space. The kriging and cokriging techniques used are known as stochastic methods. These methods exploit spatial dependency, which have two components: first order effects (general trend) and second order effects (local trends) (Finkelstein et al. 2003; Jerrett et al. 2005a). The ordinary kriging applied in this research is one of the most common forms of kriging used (O'Sullivan and Unwin 2010) and accounts for local variability of the variable of interest (Beelen et al. 2009). Ordinary kriging methods are optimal because they provide the best linear unbiased estimate at any point in the study area (Jerrett et al. 2003; Waller and Gotway 2004). The selection of the search neighbourhood for the prediction, which defines how the data points for each target nodes are selected, is an important step in the creation of a kriged surface. Depending on the spatial distribution of the data points, a different strategy may be required. Waller and Gotway (2004) recommend testing different scenarios before finalizing the search neighbourhood. The main advantage of kriging is that it requires no data aside from the data collected at the sampling sites. This simplicity is also its main weakness; it supposes an unsophisticated distribution of pollutants in space (Briggs, 2005; Luginaah et al. 2006). This can be explained by the fact that ordinary kriging does not use auxiliary data. For example, in the case of air pollution modeling, ordinary kriging has difficulties representing local variations and consequently tends to over-smooth the output surface (Luginaah et al. 2006).

We implement the cokriging technique, a more advanced version of the kriging algorithm, in order to compare its performance to the LUR. The cokriging technique is also a stochastic method but it is considered more “intelligent” than kriging as it uses covariates to guide the interpolation (Briggs 2005). These variables are integrated into the modeling in order to provide additional information that could improve the modeling of the air pollutant throughout the study area but more specifically in areas that were not part of the sampling (Briggs 2005). This method can be implemented when auxiliary variables such as population density or traffic data are associated to the pollutant being modeled since cokriging combines the spatial behavior of the pollutant and of the auxiliary variable based on the cross-correlation between the two variables. To our knowledge, this GIS-based approach to the modeling of air pollutants has yet to be used in the context of a North American study although it has been used by a limited numbers of researchers in Europe (Gallois et al. 2005; de Fouquet et al. 2007; Cardenas and Perdrix 2008; Malherbe et al. 2008).

4.3 Methods

4.3.1 Study Area

Transportation is the main contributor to air pollution in Ottawa: > 50% of O_3 at ground level, 90% of CO, > 85% of NO_x and approximately 60% of SO_2 measured are caused by transportation activities (City of Ottawa, 2004). SO_2 concentrations are comparable to the average rate for Ontario, while NO_x levels are less than the provincial average (City of Ottawa, 2004). Despite a relatively good air quality level, the Ontario Medical Association estimates to $65 million a year the cost of health services dispensed because of air pollution in Ottawa (City of Ottawa, 2004).

4.3.2 Data Acquisition

The National Capital Air Quality Mapping project is the result of cooperative work between the City of Ottawa, Environment Canada and Health Canada. Launched in the fall of 2007, this project permitted the deployment of 30 passive samplers (from which 29 were used for this research) over the period of 29 September 2007 to 13 October 2007 (Figure 3.1). Data collected at two permanent stations were also added to the dataset. The passive samplers and the permanent stations measured the concentration in NO_2 at each location. For this research NO_2 concentration was used as it is an essential marker for traffic-related air pollution and strongly correlated with other common air quality indices (Piro et al. 2008; Crouse et al. 2009). This pollutant has been used as a proxy to evaluate exposure to traffic emissions (Jerrett et al. 2007).

The 2008 street network file from Statistics Canada was used as it corresponds to the state of the infrastructure in Ottawa in 2007. The 2006 spatial boundary files at the geographic level of the dissemination block (DISB) and the dissemination area (DA) were also used (as they are from the vintage associated to the most recent Canadian Census of population). Through the Data Liberation Initiative the data are made available to Canadian universities at no cost (Statistics Canada 2009). The dissemination block "is an area bounded on all sides by roads and/or boundaries of standard geographic areas. [It is] the smallest geographic area for which population and dwelling counts are disseminated" (Statistics Canada, 2007d). The geographic level of the dissemination area is the next level in the hierarchy of the standard geographic area. It is defined as "a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all census data are disseminated" (Statistics Canada 2007c). Zoning data provided by the city of Ottawa were used in the development of the LUR model and for the dasymetric mapping of population. Finally, the greenspace layer used in the development of the LUR model originated by the digitizing of aerial photography

4.3.3 NO_2 modeling

In Chapter Three, LUR models were developed based on three different spatial representations of the population variable class. The LUR model used for the comparison work in this chapter is the best performing model presented in Chapter Three, which is the model developed using the population representation based on dasymetric mapping (DASYM) at a 20 meters resolution.

The kriging model and all of the cokriging models adopted a spherical semivariogram with a nugget effect (Table 4.1). Due to the distribution of the samplers and permanent stations, which are irregularly distributed through space, all four models used four quadrants to divide up the search neighbourhood (Waller and Gotway 2004). Using this approach not only the nearest neighbours are included in the search neighbourhood but the nearest neighbours in all directions. Each of the quadrants was set to include two to three neighbours.

4.4 Results

4.4.1 LUR Model

As explained in detail in chapter three, the measured levels of NO_2 at the sampling locations (dependent variable) were used in a multiple least-squares regression model (Ryan and LeMasters 2007). Buffers ranging from 50 - 2000 meters at an interval of 50 meters were created for each of the sampling sites and used to create over 300 variables (10 variables using 40 different buffer sizes in most cases). Stepwise selection was the method used to develop the different regression models (Jerrett et al. 2007; Su et al. 2008). No statistically significant spatial autocorrelation was found in the residuals when tested using the modified Moran's I test specifically for OLS regression residuals (Cliff and Ord 1981) in the 'spdep' package of the R language (v 2.10.0). The OLS residuals were tested using both distance and k-nearest neighbour (coerced to a symmetric matrix) derived spatial weights matrices for reasons of robustness. The model containing the dasymetrically derived population variable provided an R^2 of 0.8055, which is within the range of the LUR performances in Canada (Table 4.2). The model includes the population count within 1,750 metres using the dasymetric representation, the area of greenspace within 700 m, the area of industrial land-use within 250 m and the total length of major roads within 250 m which serves as a proxy for traffic data. The length of major road was used as a proxy for traffic as the latter data were not available (Hoek et al. 2008). The model was validated using leave-one-out cross validation (LOOCV) (Hoek et al. 2008; Mukerjee et al. 2009). The root-mean-square error (RMSE) and the mean absolute error (MAE) were computed in order to assess the performance of the model (Cassell 2007). The results (RMSE = 1.05 ppb and MAE = 0.86 ppb) indicate that the NO_2 values predicted by the model are reasonable.

4.4.2 Geostatistical Interpolation

The first step to geostatistical interpolation was to conduct exploratory spatial data analysis in order to develop a better understanding of the data. The NO_2 data used for this research are characterized by a unimodal distribution with a mean of 6.43 ppb. The air pollution data exhibit two spatial trends: a West-East trend (low to high concentrations) and a South-North trend (low to high concentrations). Global Moran's I, a measure of global spatial autocorrelation, produced as a distance based correlogram, indicates that the NO_2 data are significantly autocorrelated to a certain distance, after which the spatial autocorrelation becomes non-significant. This distance, which corresponds to the range, is found for each of the interpolation method in Table 4.1.

In order to determine if representations of population at different spatial scales, as a covariate with NO_2, has an impact on the results of geostatistical interpolation, three cokriging models were developed using observed NO_2 and the following representations of the population variable: dasymetrically derived population (DASYM), block level population (DISB), and dissemination area population (DA). The population density was calculated for all three spatial representations of population. Density was used instead of the raw population count for the development of the models because the varying size of the geographical units among the different representations of population would have influenced the results of the cokiriging. A kriging surface, which does not make use of covariates, was also generated using ordinary kriging for comparison. All geostatistical interpolations were conducted in ArcGIS 9.2 and created NO_2 surfaces with 20 metres resolution - the same spatial resolution as the surface obtained through the implementation of the LUR model.

Metrics on the quality of the four geostatistical models developed for this research are found in Table 4.3 and Figure 4.1. All four models provided comparable results, indicating that the kriging model and the three cokriging models are similar. The main difference is found in the mean prediction error; the results of the kriging approach are characterized by a slightly higher mean error than the three cokriging methods. The results of the cross-validation confirmed that the kriging and the cokriging models developed are good representations of the study area.

4.5 Discussion

The NO_2 concentration surfaces created using both kriging and the LUR model were compared through visual and statistical analysis. The goal of each comparison technique presented hereafter is different and complementary. The first comparison is focused on the spatial patterns and the second is focused on the distribution of the NO_2 values without any consideration to their spatial distribution within the study area.

4.5.1 Visual Comparison

In order to conduct proper comparison between the five NO_2 surfaces, the results of all five models were reclassified into 20 classes using the Jenks' natural classification algorithm in ArcGIS 9.2 (Figure 4.2).

The first comparison encompasses the three different cokriging models in order to assess the effect of using different representations of the population variable class on the results of geostatistical interpolation. The results exhibit the same general patterns: a small downtown area with high NO_2 concentrations surrounded by rings of decreasing levels of NO_2 concentrations. Another key trend is the higher NO_2 concentrations in the East of the city in comparison to the West of the city, a trend also identified using exploratory spatial data analysis (ESDA) on the air pollution data. Dasymetric mapping, which is the method underlying the cokriging DASYM model, is a spatial interpolation method with potential to improve the precision of many GIS applications but the results obtained indicate that the use of the population data at the DA and DISB geographic levels provides very similar results. As a consequence, it appears that the spatial scale of population representation does not have a large impact on the general patterns of the NO_2 surfaces created using geostatistical interpolation. In fact, the results of the kriging approach display similar general trends as the three cokriging models but by using covariates, the NO_2 surfaces generated with the cokriging models are less generalized and thus provide more spatial detail when compared to ordinary kriging.

The key difference between the results of the LUR and the geostatistical interpolation surfaces is found in the level of detail that can be obtained through the LUR approach (Ryan and LeMasters 2007). The LUR is capable of representing variations associated with linear sources of emissions and these can be the main source of emissions in urban area (Briggs et al. 1997). NO_2 is known to vary significantly at a scale as small as 50 metres (Hewitt 1991) and in Ottawa the concentration of the pollutant is strongly associated with traffic. Capturing these variations is key to improved modeling of the geography of exposure. As with geostatistical interpolation, the LUR identifies an area of higher NO_2 concentrations in the downtown area of the city, that is however, strongly spatially delineated compared to the geostatistical interpolation outputs. The areas with the highest NO_2 concentrations on the LUR surface are for the most part associated to the location of small industrial parks, which is one of the covariates included in the model. Another important element observed on the LUR surface is the large linear area characterized by lower NO_2 concentrations that surrounds the urban core of the city. This area is known as the “greenbelt”. The greenbelt is a conservation area made up of different types of greenspaces and characterized by the presence of scant infrastructure and a very low population density. Consequently the NO_2 concentrations in the greenbelt and in its proximity are lower than in the rest of the urban regions of the study area. The last difference between the LUR surface and the geostatistical interpolation surfaces is the variation of NO_2 concentrations in the suburban and the rural areas (areas outside of the greenbelt). With the LUR surface NO_2 concentrations follow the road network and the location of settlements while the use of geostatistical interpolation did not provide that level of detail. These observed differences between the geostatistical interpolation models and the LUR model exemplify the benefit of using spatially detailed covariates: the LUR approach can infer pollutant concentrations in areas that lack air quality sampling and as such provide a far more detailed map than geostatistical interpolation approaches. Finally, the LUR does not confirm the generally higher NO_2 concentrations in the East of the city as it was the case with the geostatistical interpolation surfaces. The results of the LUR and the geostatistical interpolation exhibit numerous differences in terms of the geography of exposure. These differences could be a sign of exposure misclassification and hence have an impact on the strength of the relationship between exposure to NO_2 and health.

4.5.2 Statistical Comparison

The NO_2 surfaces obtained with the different modeling approaches were statistically compared. A regular point layer at a resolution of 20 metres was generated from the raster data and each of the points was attributed the NO_2 concentration values at that specific location through spatial overlay. This overlay approach was applied for each of the five NO_2 surfaces in order to build the database for the statistical comparison. From the regular point layer, 400 points were randomly selected for our statistical analysis (Figure 4.3).

Table 4.4 provides the statistical measures for comparison of the NO_2 concentration surfaces extracted at the location of the 400 random points. All measures obtained for the geostatistical interpolation models are similar indicating that the use of different representations of the population variable does not have a strong effect on the modeled NO_2 values. In terms of geostatistical interpolation, the main differences are found between the kriging and the three cokriging models. With the LUR model, the central half of the data is found within a 0.1327 ppb range, which corresponds to the interquartile range but the other half is found within a wider range then with geostatistical interpolation models. The median value of the LUR model and the geostatistical interpolation models are similar. On the other hand, the value for the third quartile is considerably lower for the LUR model than for the other models while the first quartile of the geostatistical interpolation models is considerably lower than for the LUR model, making the interquartile range value also smaller.

The distribution of the NO_2 values for the LUR, cokriging and kriging models are compared in Figures 4.4 and 4.5. The majority of the values for the LUR model are found around the mean, which is not the case with the other models. A first comparison can be made between the results of the LUR model and the cokriging DASYM using comparative histograms (Figure 4.5). The histogram of the cokriging values is skewed to the left in comparison to the LUR model. Figure 4.4 shows a box plot confirming a different distribution of NO_2 values in the LUR model. A second comparison between the kriging and the cokriging DA emphasizes the similarities between the different geostatistical interpolation models. Both models have a distribution that is skewed to the left. As in Table 4.4 the histograms show that the minimum value of kriging is larger than cokriging DA and as a consequence the frequency of other NO_2 concentrations within the distribution is higher for kriging. The distribution of the data for all geostatistical approaches share similar characteristics.

The resulting database obtained through spatial overlays was analyzed using a T-Test for Pearson correlation coefficient corrected for spatial autocorrelation (Dutilleul 1993) as implemented by Legendre (2000) to compare the outcomes of the different NO_2 surfaces. A series of modified T-Tests was performed to determine if the different models provided NO_2 surfaces that are statistically different. The results of each NO_2 model were compared to the results of the four other models (Table 4.5). The modified T-Test indicates that the result of the NO_2 modeling using the LUR is not statistically different ($p < 0.0001$) than all the other models tested but has a weak correlation with the other models. This result is possibly a consequence of the fact that the LUR is capable of capturing small scale variations of NO_2 concentrations using auxiliary variables. The implementation of the modified T-Test procedure for the comparison of the different cokriging models confirms the observations made through visual analysis. The use of different spatial representations of the population variable class does not significantly affect the results of cokriging. All the comparisons between the cokriging models are statistically significant and display very strong Pearson's r values. In chapter three it was found that the use of different representations of the population variable class had a modifying affect on the results of LUR modeling. It appears that the spatial representation can be an element affecting the results in the study of the relationship between exposure to NO_2 and health depending on the type of model being used. Finally, the results of the modified T-Tests between kriging and all other geostatistical models point to the fact that this simple approach to modeling of NO_2 concentrations does provide results that are very similar to the cokriging using any of the population representations. On the other hand, the result of the modified T-Test indicates a weak, but statistically significant, correlation between the LUR and the kriging models.

The visual interpretation of the NO_2 surfaces has shown that the results of the LUR modeling provide more precise results, being able to represent variations at the local scale both in rural and urban environments. Completion of a series of modified T-Tests corrected for spatial autocorrelation confirmed that the LUR surface is not statistically different than all other NO_2 surfaces created through geostatistical interpolation methods but the correlation is weak. The findings of our analysis point to the fact that the LUR performs best for modeling of NO_2 which confirms findings reported in the literature (Jerrett et al. 2005b; Hoek et al. 2008; Sahsuvaroglu et al. 2009). Finally our results also demonstrate that the use of different representations of the population variable class does not significantly affect the results of cokriging due to the fact that the algorithm is not sensitive to information stored at a small geographic level. It is possible, however, that the use of a variable other than population would have provided different results. For consistency and to allow for the comparison of the results obtained through the LUR model with the results of the geostatistical interpolation models the population variable was selected.

The main weakness associated with the methodological approach developed in this research is the limited number of data points (sampling locations) available for the modeling of NO_2. As it is the case with the vast majority of statistical analytical methods, a larger number of data will allow for the development of a better model (O'Sullivan and Unwin 2010).

4.6 Conclusion

Two elements were tested in this research, both centered on the role of spatial representation in the modeling of NO_2 concentrations. The first issue addressed was the importance of using different spatial representations (scales) of the population variable on the modeling of NO_2 using cokriging. Contrary to the results obtained in the study of spatial representation in the development of a LUR model, the use of different spatial representations of the population variable class does not significantly affect the result of cokriging. As such, the use of alternative scale representations may have an impact on the modeling of air pollution exposure depending on the approach used (Brindley et al. 2005).

The second issue investigated was to evaluate the results of different approaches to NO_2 modeling through the comparison of LUR and geostatistical interpolation methods. The results obtained indicate that the use of different methodological approaches could potentially be one reason behind the equivocal relationship between health and exposure to air pollutants. A series of T-tests did not confirm statistically significant differences between the results of the different approaches but the correlation between the LUR and the geostatistical interpolation methods was weak. Our next step will be to assess the relationship between exposure to NO_2 and health using natural neighbourhoods. A similar analysis will measure the same relationship based on different operationalizations of the units of analysis. This will represent the last component of the three-tier conceptual framework; it will aim at investigating the impact of spatial representation on the geography of risk.

Parameter	Kriging	Cokriging DA	Cokriging DISB	Cokriging DASYM
Semivariogram type	spherical	spherical	spherical	spherical
Neighbourhood	4 sectors (2-3 neighbours)	4 sectors (2-3 neighbours)	4 sectors (2-3 neighbours)	4 sectors (2-3 neighbours)
Range	30797	30147.1	29927.7	30543.5
Nugget	4.3283	4.3148	4.3099	4.3233
Sill	5.9939	5.9254	5.9018	5.9664

Table 4.1 Parameters in the kriging and the cokriging models.

Variable	Parameter estimate	PR> t	VIF
Intercept	3.66507	0.0027	0
Major Roads 250m	0.00041824	0.0116	1.47569
Population Dasymetric 1750m	0.00010586	0.001	2.24357
Greenspace (land use) 700m	-0.00000216	0.0229	2.39553
Industrial (land use) 250m	0.00005783	0.0398	1.28638

Table 4.2 Variables included in the LUR model.

Prediction Errors	Cokriging DA	Cokriging DISB	Cokriging DASYM	Kriging
Mean	0.05137	0.04994	0.03918	0.08863
Root-Mean-Square	2.297	2.265	2.268	2.297
Average Standard Error	2.472	2.482	2.483	2.472
Mean Standardized	0.01645	0.00122	-0.003009	0.01645
Root-Mean-Square Standardized	0.9338	0.9187	0.9197	0.9338

Table 4.3 Statistics on the geospatial models developed for Ottawa.

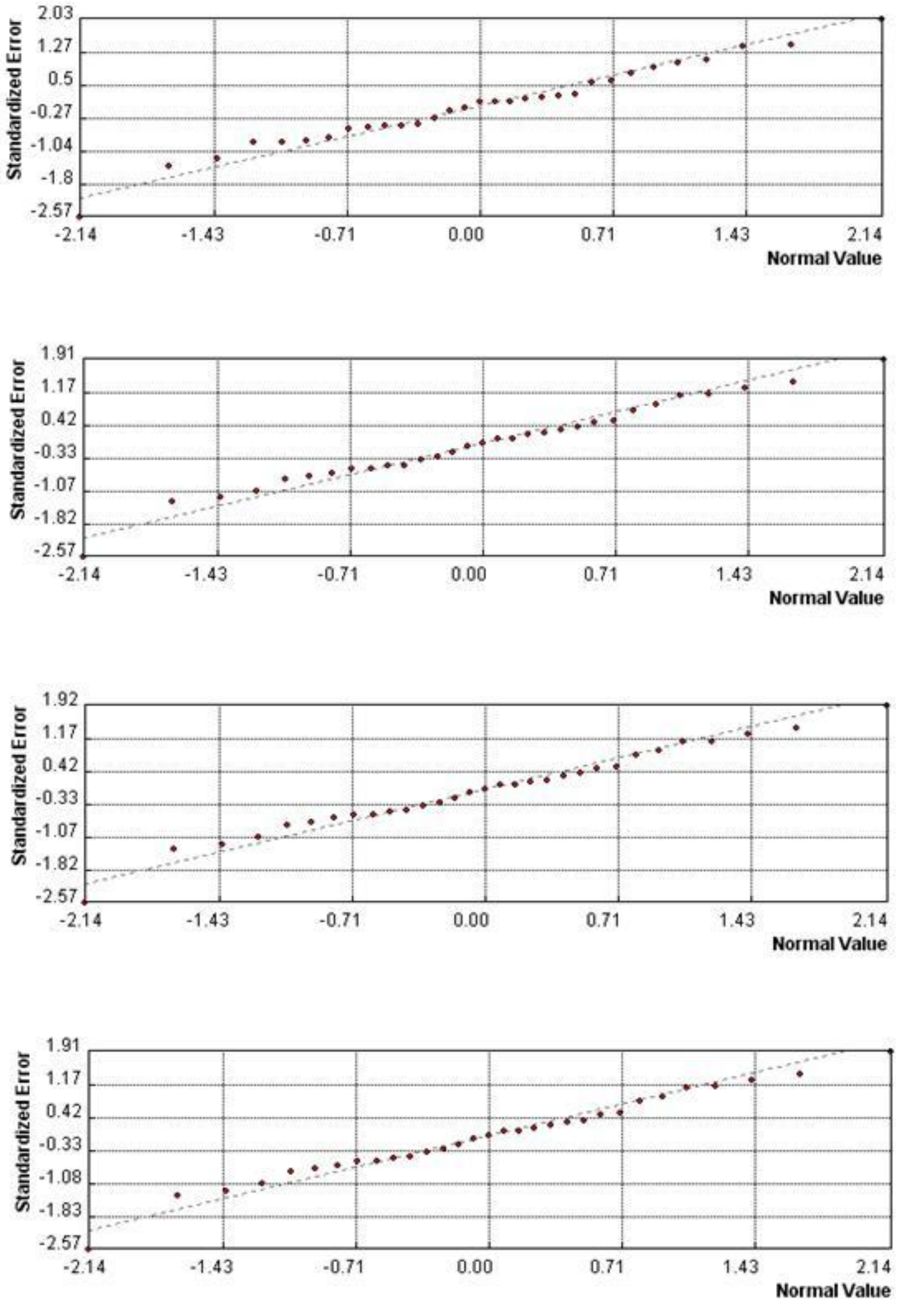


Figure 4.1 QQ plots from the cross-validation step. From the top to the bottom: kriging, cokriging DA, cokriging DISB, and cokriging DASYM.

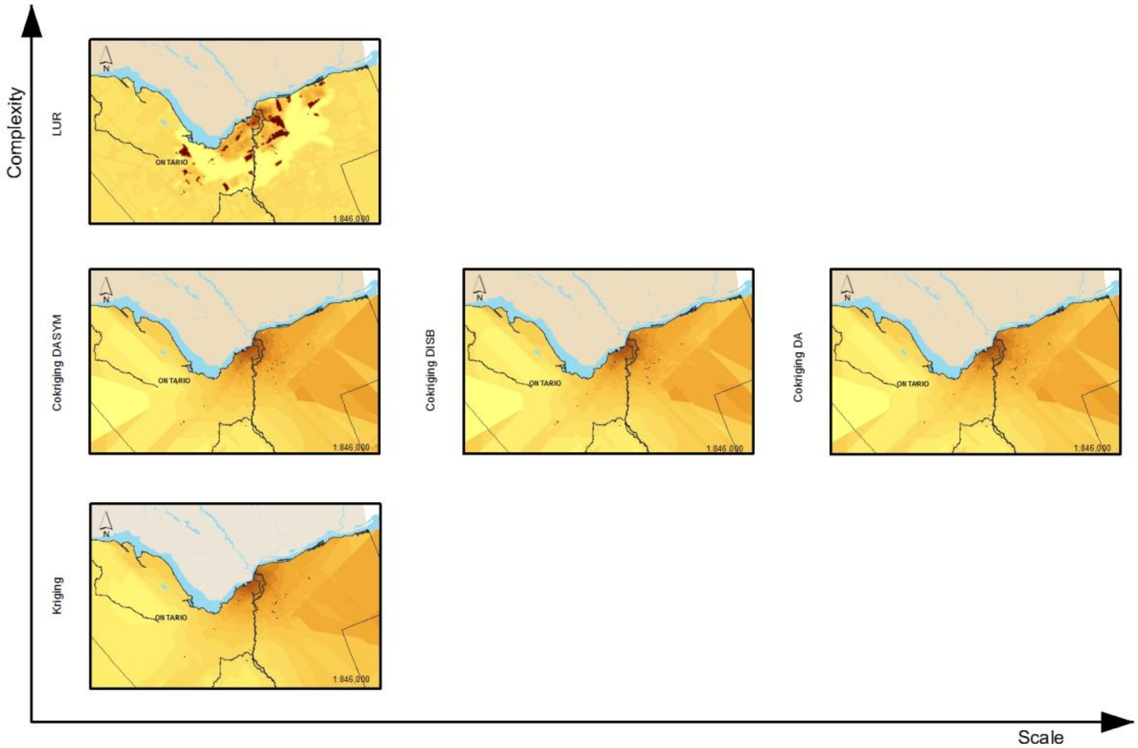


Figure 4.2 Spatial distributions of the NO_2 concentrations using geostatistical models and the LUR (higher concentrations of the pollutant are represented by darker shades).

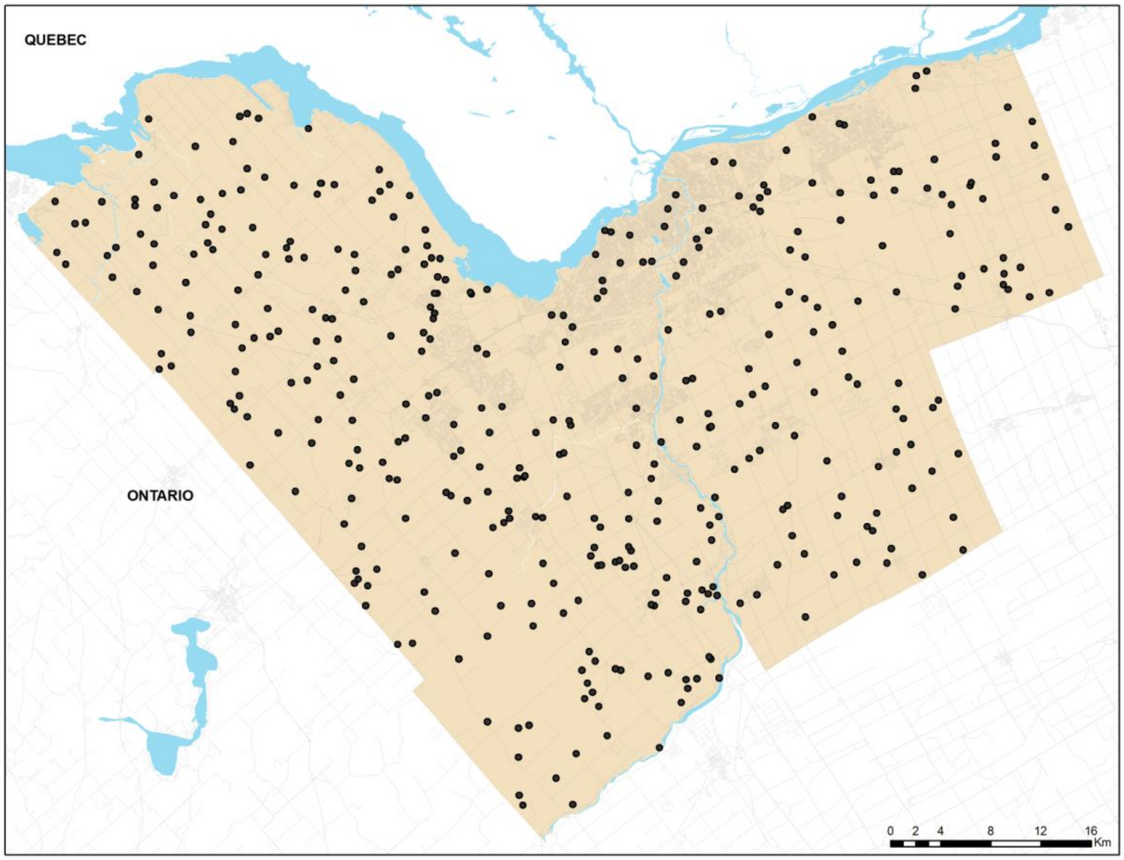


Figure 4.3 Location of the 400 random points utilized for the spatial comparison of the NO_2 surfaces.

	LUR	Cokriging DASYM	Cokriging DISB	Cokriging DA	Kriging
Max	13.1135	8.8101	8.8294	8.8188	8.6537
3rd quartile	3.8055	4.6353	4.6397	4.6391	4.9064
Median	3.6856	3.4827	3.4845	3.4795	3.6690
1st quartile	3.6728	2.994	2.9926	2.9756	2.959
Min	13.1135	1.6654	1.6655	1.6628	1.7334
Interquartile Range	0.1327	1.6413	1.6471	1.6635	1.9474

Table 4.4 Comparison of the distribution of the NO_2 concentrations (ppb) using the four geostatistical models and the LUR model.

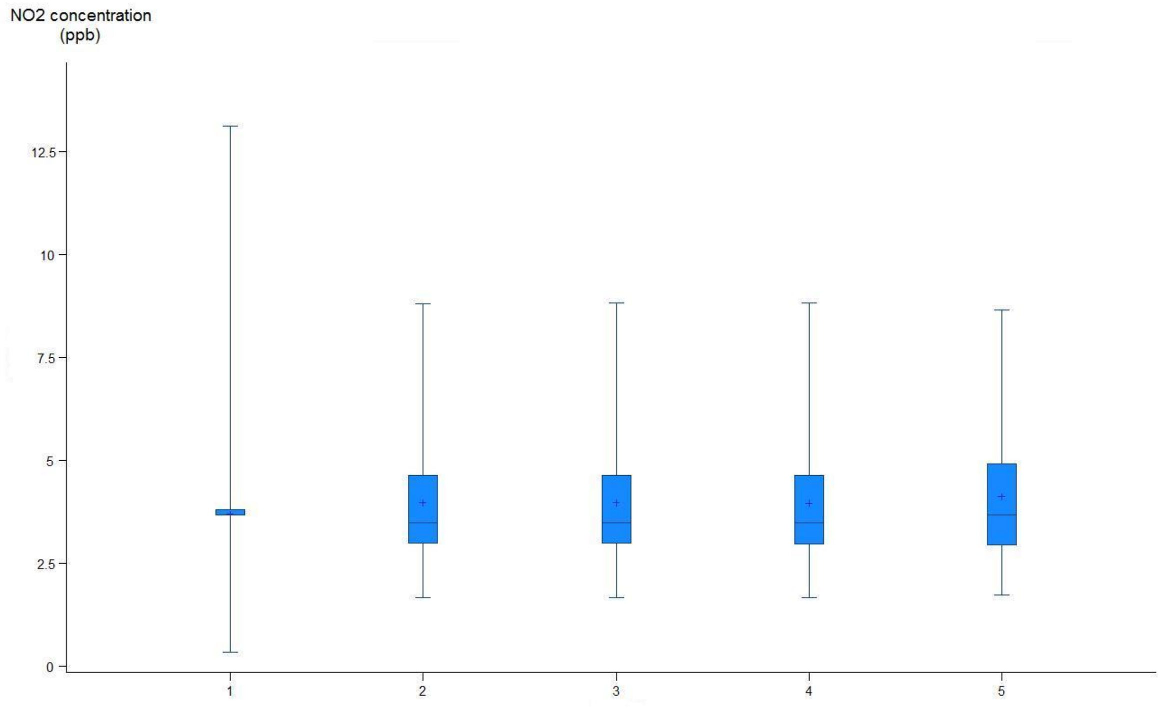


Figure 4.4 Box plots showing the distribution of the NO_2 values (ppb) for each modeling approach: 1) LUR, 2) cokriging DASYM, 3) cokriging DISB, 4) cokriging DA, and 5) kriging.

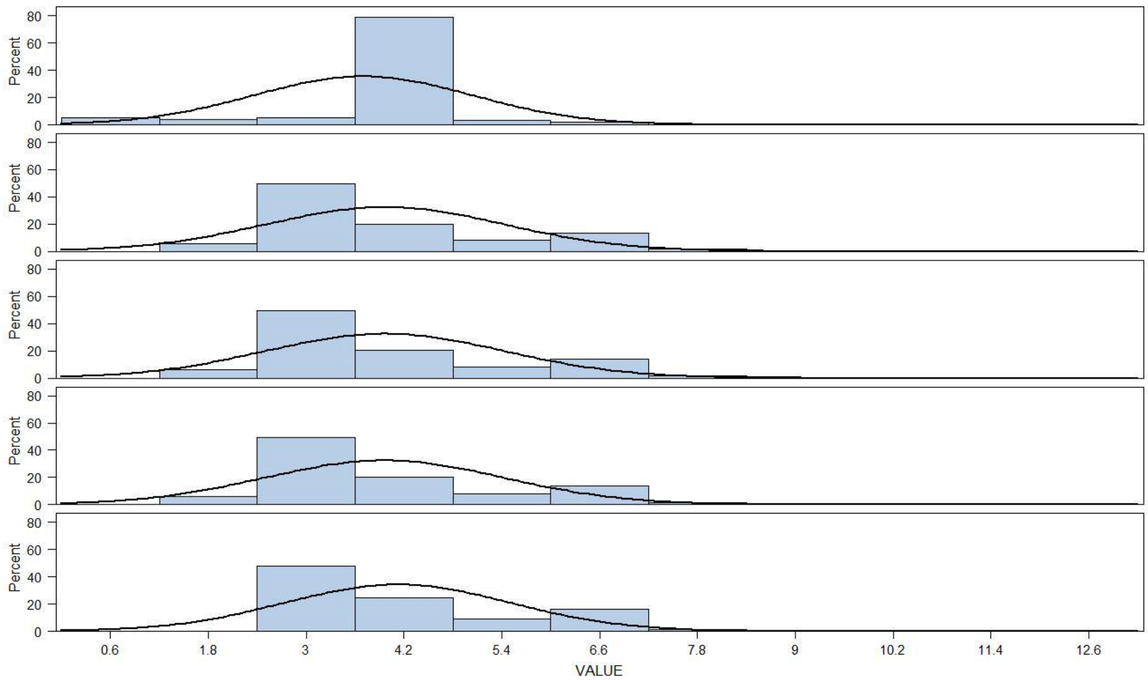


Figure 4.5 Comparison histograms of the distribution of the NO_2 values. From top to bottom: 1) LUR, 2) cokriging DASYM, 3) cokriging DISB, 4) cokriging DA, and 5) kriging.

Model 1	Model 2	Pearson's r	Pr > |t|
LUR	Kriging	0.15199	0.00889
LUR	Cokriging DA	0.14107	0.01038
LUR	Cokriging DISB	0.13927	0.01131
LUR	Cokriging DASYM	0.13883	0.01144
Cokriging DASYM	Cokriging DISB	0.99994	<0.0001
Cokriging DASYM	Cokriging DA	0.99979	<0.0001
Cokriging DASYM	Kriging	0.94415	0.0006
Cokriging DISB	Cokriging DA	0.99977	<0.0001
Cokriging DISB	Kriging	0.94383	0.0006
Cokriging DA	Kriging	0.9437	0.0006

Table 4.5 Results of the modified T-Tests for Pearson correlation coefficient corrected for spatial autocorrelation.

Chapter 5 A Multi-Scale Analysis of the Relationship Between Exposure to NO_2 and Respiratory Health

Marie-Pierre Parenteau[(1)]
Michael Sawada[(1)]

[(1)] Department of Geography, University of Ottawa

Laboratory for Applied Geomatics and GIS Science, Department of Geography, University of Ottawa, Ottawa, Ontario K1N 6N5

Tel:+1 613 562-5800 Ext. 1040; Fax: +1 613 562-5725

In preparation.

5.1 Introduction

The concept of neighbourhood is ill defined and there is no consensus on its definition. In most cases these units are defined as small-areas that share some predefined set of characteristics (Galster 2001; Jones et al. 2010). Neighbourhood definition is a problem because many health studies at the intra-urban level rely on this geographical concept. In Canada, the neighbourhood has been operationalized as the census tract in many studies (Buzzelli et al. 2003; Glazier et al. 2000; James et al. 2007; Gauvin et al. 2008; Walks and Maaranen 2008; Généreux et al. 2008; Urquia et al. 2009), even if the use of this geographic unit has been criticized. As a consequence, it can be said that the way place is conceptualized at the local scale is one of the weakest theoretical aspects in health studies, (Matthews 2008). The study of the relationship between exposure to criteria pollutants, such as nitrogen dioxide (NO_2), and health outcomes is no exception to this problem. Unfortunately few health studies have focused on this issue (Stafford et al. 2008) even though the literature clearly identifies the possibility of measuring different levels of association under different zoning systems (Openshaw 1977; Opensaw 1984; Holt et al. 1996; Reynolds and Amrhein 1997). Considering the fact that the studies on the relationship between NO_2 and health outcomes are divided on the role of exposure to NO_2 on health, it is surprising that the study of the role of spatial representation in the analysis of this relationship has not received more attention.

Standard geographical units from the Canadian Census, most particularly the census tract, are commonly used to operationalize the concept of neighbourhood.

The main advantage of this approach is the fact that census data are readily available (Gauvin et al. 2007). The census tract is delineated based on optimal population counts, the compactness of the shape, visible boundaries and input from local experts (Statistic Canada 2007b). The census tract boundaries must follow visible features when it is possible but may in some cases be bounded by administrative boundaries, making these ambiguous and hard to define (Lee et al. 2008; Schaefer-McDaniel et al. 2010). Homogeneity in terms of socioeconomic status is not part of the delineation criteria for the census tract or any other standard geographical classification levels of the 2006 Canadian Census geography (Schuurman et al. 2007; Lee et al. 2008). Considering the fact that neighbourhoods are expected to be homogeneous units (Galster 2001), the use of the census tract as a proxy for the neighbourhood is problematic from a conceptual point of view. Its use is even more troublesome when no other alternative spatial structure is used for sensitivity analysis and thereby not allowing for the assessment of MAUP effects on the results (Openshaw 1984; Lee et al. 2008).

The Modifiable Area Unit Problem (MAUP) can bring differences in the analytical results of the same input data compiled under different zoning systems (Openshaw 1984; Riva et al. 2009; Briant et al. 2010). Openshaw (1984) explains that the MAUP is made up of two components: the scale effect and the zoning effect. The first one, the scale effect, causes spatial aggregation to affect analytical results. On the other hand, the zoning effect is one where the number of units is stable, but the use of a different spatial structure provides different results (Jelinski and Wu 1996; Stafford et al. 2008). The importance of the MAUP is such that according to Schuurman et al. (2007), any study about the association between health and place will be influenced by the scale and zoning design used to conduct the study. In general, the scale effect is recognized as the most troublesome component of the MAUP; the shape (zoning effect) also matters but to a lesser extent (Schuurman et al. 2007; Briant et al. 2010).

The MAUP is also known to impact the results of univariate and multivariate regressions (Briant et al. 2010). Calibration of the model can be affected to such a level that according to Jelinksi and Wu (1996) it can lead to unreliable results. Fotheringham et al. (2002) provide insights into the cause of MAUP in regression analysis. According to the authors, it may be caused by the spatial non-stationarity of multiple predictors that together may be factors for a response variable. Because the health status of an individual is generally the result of multiple factors, health studies are at risk of being affected by the MAUP (Swift et al. 2008; Diniz-Filho et al. 2009). One approach to mitigate the effect of the MAUP on analytical results is to create a geographical structure where the units have high levels of internal homogeneity and are externally heterogeneous (Haynes et al. 2007). According to Mu and Wang (2008) this approach could represent a solution to the errors in the model building process that are induced by spatial autocorrelation. Another solution that has been proposed is the analysis of the relationship under different spatial structures as a way to conduct sensitivity analysis (Jelinski and Wu 1996).

An automated zone design methodology was first developed by Openshaw (1977) for the study of the zoning and scale effects on analytical results. These automated approaches, based on computer algorithms, regroup basic spatial units (BSU) into N zones so that each BSU is linked to one zone only (Openshaw 1977; Daras and Alvanides 2006). One delineation constraint generally used in these models is a contiguity constraint (Openshaw 1977). Examples of other constraints that can be used for the delineation are the number of units created, the compactness (shape), population count, area size and internal homogeneity. Automated zone design algorithms have been used by Cockings and Martin (2005) for the study of the relationship between morbidity and deprivation. The authors found that zoning systems automatically delineated produced stronger relationships than when using census tract. They also used this approach to study the scale effect; their results confirmed the work of Openshaw where stronger relationships were measured using more aggregated units.

Since the main objective of this research is to assess the effect of the MAUP on the study of the relationship between exposure to NO_2 and respiratory health, three different zoning systems (geographic structures) were incorporated into our framework. The census tracts from the 2006 Canadian Census of population are used as statistical units, the natural neighbourhoods delineated based on a homogeneity criteria represent an optimal zoning design and a cluster structure created through automated delineation represents a zoning design structure delineated using a simple contiguity constraint. The use of these three structures will allow furthering the understanding of the effect of the MAUP on analytical results.

5.2 Methods

5.2.1 Health Data

The effect of air pollution on respiratory health can be measured through emergency room visits and hospital admissions (Lee and Ferguson 2009). The respiratory morbidity rate for individuals of age 15 and over within the Ottawa Public Health Unit is the main health outcome measure in this research. For this study, morbidity was preferred over mortality due to the larger amount of data available, thus increasing the statistical power to characterize associations with air pollution (Villeneuve et al. 2006).

Health conditions associated to exposure to NO_2 were identified from the 10^{th} revision of the ICD (CIHI 2006) based on literature regarding the health effects of air pollution (Kampa and Castanas 2008). The health outcomes considered are from Chapter 9 (Diseases of the Respiratory System) and more specifically correspond to the block J40-J47 (Chronic Lower Respiratory Diseases). The *Discharge Abstract Database* (DAD) and *National Ambulatory Care Reporting System* (NACRS) data, for the health conditions listed above, for fiscal years 2005-2006, 2006-2007 and 2007-2008 were extracted by the Ministry of Health and Long-Term Care (MOHLTC) and compiled at the geographic level of the 95 Ottawa neighbourhoods and census tracts. The DAD dataset contains demographic, administrative, and clinical data for hospital discharges and day procedures (CIHI 2010a). The second source of information for this research is the NACRS, which contains demographic, administrative and clinical data for ambulatory care visits (CIHI 2010b).

The morbidity rates were directly sex and age standardized (Sridharan et al. 2007; Hu and Rao 2009) for the age groups 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84 and 85 and over. Data on hospitalization and ambulatory care were used along with some of the individual level data (sex, age, and postal code of the residence) for each patient. The spatial distribution of the morbidity rates for each of the spatial structures is shown in Figure 5.1.

5.2.2 Natural neighbourhood, cluster structure and census tract

The natural neighbourhoods were delineated through a semi-automatic approach with the intention of being used as the geography of reference for this research as well as for the Ottawa Neighbourhood Study (ONS) project. The objective was to delineate homogeneous units in terms of socioeconomic status (SES) that would also maximize external heterogeneity (the list of variables used is found in Table 5.1). The first step was to automatically aggregate dissemination areas (DAs) through spatially constrained clustering and wombling using the software BoundarySeer. The clustering and wombling algorithms were applied to SES and housing data from the 2001 Canadian Census at the geographic level of the DA. Once the automated delineation completed the natural neighbourhoods were manually refined. Some boundaries were later manually updated following the dissemination of the 2006 Census data to better represent Ottawa's neighbourhoods. A total of 95 neighbourhood units make up this boundary file. This work was achieved through consultations with the City of Ottawa and local grassroots organizations. Details on the methodological approach can be found in Chapter 2.

The cluster structure was created through the grouping of census tracts using ArcGIS 9.2 and a simple contiguity constraint. The development of this aggregated set was based on the selection of a random census tract as an initial seed, followed by unioning this census tract with a neighbor(s) and then moving to the next non-aggregated census tract and repeating the process until all census tracts were visited. Using this approach we automatically delineated 95 units to compare with the 95 natural neighbourhoods, thus allowing for the testing of the zoning effect. The last spatial structure used in this research is from the 2006 Canadian Census of population. From the latter, 184 census tracts were extracted to cover the study area. The use of these three spatial structures allows for the assessment of the scale effect and the zoning effect in the study of the relationship between exposure to NO_2 and health.

Contextual variables are characteristics of the local environment (e.g. natural, built, social) that can influence the health of individuals in ecologic studies such as this one (Sahsuvaroglu et al. 2009; Jerrett et al. 2005b; Diez-Roux 2002; Piro et al. 2008). Compositional variables are aggregates of resident characteristics (e.g. income, age, education) that could have an effect on the measured association between air pollution and health (Diez-Roux 2002; Macintyre et al. 2002). Both contextual and compositional data were extracted from the 2006 Canadian Census of population. Data at the geographic level of the natural neighbourhood were obtained through custom tabulations from Statistics Canada through the Ottawa Neighbourhood Study.

5.2.3 Exposure Measures

A land-use regression (LUR) model was developed for the mapping of NO_2 concentrations in Ottawa, Canada (Figure 3.4). Details about the methodological approach for the development of the LUR model for Ottawa can be found in chapter three. The model, which included data on the road network, population, green spaces and industrial land-use, yielded an R^2 of 0.8055. Zonal statistics were applied to the NO_2 raster obtained with the implementation of the LUR model to derive mean NO_2 concentrations by census tract, neighbourhood and aggregated cluster (Hu and Rao 2009).

5.3 Results

Preliminary data analysis was first conducted to assess the role of spatial representation on summary statistics, global spatial autocorrelation and bivariate Moran's I. Bivariate Ordinary Least Square (OLS) and multivariate OLS regressions where then applied to the three zoning systems to assess the effect of the MAUP in the study of the relationship between respiratory health outcome and exposure to NO_2. Spatial regressions were not used for this component of the research as a comparison between spatial regression and OLS will be explored in chapter six. SAS (version 9.2) and GEODA were used for the statistical and spatial analysis of the data.

5.3.1 Summary statistics

The mean value of the explanatory variables for the three spatial structures is similar (Table 5.2). The variables of educational attainment and % low income have the highest levels of variability. The variable Mean NO_2 concentration displays a much smaller level of variability under the different structures, with mean values ranging from 5.29 to 5.39 ppb. On the other hand, the mean value for the respiratory health outcome rate is significantly affected by the use of different spatial structures. The minimum average rate is 1,275.35 per 100,000 at the natural neighbourhood level in comparison to 2,346.89 per 100,000 for the census tract structure. Standard deviations are also similar for most variables under the three different zoning systems but as with the mean values, there are exceptions. Finally in terms of variance, the census tract structure generally has higher variance values than the two other representations.

5.3.2 Global Spatial Autocorrelation

Moran's I is one of the most widespread measure of global spatial autocorrelation. Its implementation is the result of combining indicators of similarity: one for the value and the other for the spatial relationship (Anselin 2005b). Using a first-order Queen's case spatial weight matrix and 999 permutations, we found statistically significant spatial autocorrelation at a pseudo-significance level of $p \leq 0.05$ for all the explanatory variables for each of the spatial structures with the exception of the average income under the natural neighbourhood structure (Table 5.3). Exposure data (Mean NO_2) is characterized by strong statistically significant spatial autocorrelation under all spatial structures. The respiratory health outcome data are characterized by statistically significant global spatial autocorrelation under the census tract structure but using the natural neighbourhood boundaries, the respiratory data no longer exhibit statistically significant spatial autocorrelation.

The comparison of the Moran's I statistic under the natural neighbourhood and the cluster structure suggests a zoning effect. In general the clusters tend to be characterized by higher levels of spatial autocorrelation. For both data structures, the spatial autocorrelation is statistically significant for all the variables with the exception of the respiratory health outcome and the average income (under the neighbourhood structure). The variable respiratory health outcome is characterized by statistically significant positive spatial autocorrelation under the cluster structure while, as mentioned, the level of spatial autocorrelation using the neighbourhood is close to 0 and is not statistically significant.

A scale effect is also observed when comparing results using the clusters and the census tracts. The clusters, which are based on an aggregation of census tracts, display stronger spatial autocorrelation than the census tract. On the other hand, a comparison between the census tracts and the natural neighbourhoods reveals that in approximately half the cases, the global spatial autocorrelation is stronger within the census tracts units. The use of homogeneous areas is here compensating for larger aggregation zones.

5.3.3 Bivariate Moran's I

Bivariate Moran's I was used as an exploratory spatial data analysis (ESDA), providing information on the strength of the associations between NO_2

concentrations and the other covariates as well as in between the health outcomes and all explanatory variables (Sridharan et al. 2007). This approach also provided initial information on the anticipated direction of the effect (Henderson et al. 2007).

The correlation between the health variable and the explanatory variables appears to be affected by both the scaling and the zoning effects as the strength of the associations measured vary from one structure to the other (Table 5.4). For most relationships, the direction of the correlation is the same for all three structures. The variables average value of owned dwelling and % retail trade industry are the two exceptions. The variable average value of owned dwelling has a positive and not statistically significant spatial correlation with the respiratory outcomes under the neighbourhood structure but is negative and not statistically significant under the cluster and the census tract structures. The variable % retail trade industry is negatively correlated to the respiratory outcomes under the natural neighbourhood structure (statistically significant) and census tract structure (not statistically significant) but is positive under the cluster structure (statistically significant).

The strength of the correlation between NO_2 concentration and respiratory health outcome varies according to the spatial structure. For the natural neighbourhoods, bivariate Moran's I has a value of 0.1905 and is statistically significant. The bivariate Morans' I value for the same relationship according to the cluster structure is 0.0656 and is statistically significant. Finally using the census tract, the bivariate Moran's I value is low (I= 0.0365) and not statistically significant.

5.3.4 Bivariate regression

An ordinary least square (OLS) bivariate regression model was developed in GeoDA to measure the relationship between the variable Mean NO_2 concentration and the respiratory health outcomes for each of the three spatial structures (Table 5.5), thereby allowing the assessment of the scale and the zoning effects. The R^2 measure is low for each of the spatial structures: 0.0492 for the census tract structure, 0.0494 for the cluster structure and 0.0307 with the neighbourhood structure. The value of the mean NO_2 coefficient is positive for the three spatial structures; it is statistically significant in the cluster and the census tract models but not in the natural neighbourhood model.

The OLS regression model of the natural neighbourhood structure is characterized by a non-normal distribution of the error term (Jarque-Bera) and non-stationarity between the explanatory variables and the health outcomes (Breusch-Pagan and Koenker-Bassett tests). The OLS models for the cluster and the census tract structures are also characterized by a non-normal distribution of the error term (Jarque-Bera) but passes the tests for stationarity (Breusch-Pagan and Koenker-Bassett). Finally, the census tract and the cluster structures are characterized by a statistically significant spatial autocorrelation in the residuals using Moran's I test for regression residuals. Worth mentioning is the fact that R^2 and the adjusted R^2 values are smaller for the neighbourhood model than the census tract but according to the log likelihood, Akaike criterion and the Schwarz criterion the neighbourhood model is a better fitted OLS model.

5.3.5 Multivariate regression

Explanatory variables were introduced into a stepwise regression for each of the three spatial structures using SAS 9.2 to define the best fitting model for each of the zoning systems (Table 5.6). Since the objective of this research was to study the effect of the MAUP in the relationship between exposure to NO_2 and health, the variable mean NO_2 was included in each of the models even if it is not statistically significant. Models were to include variables statistically significant at 95%.

The model developed for the census tract structure contains six explanatory variables: % occupied private dwellings in need of major repairs, % occupied private dwellings built before 1946, educational attainment, average income, % retail trade industry and mean NO_2 concentration and yielded an R^2 value of 0.43 and the adjusted R^2 of 0.41. The model for the cluster structure is a subset of the census tract model; it contains four of the variables found in the census tract model and explains less of the variability in the health outcome data (R^2= 0.39 and Adj R^2= 0.37). The model building exercise provided a different set of explanatory variables using the neighbourhood structure. The model is made up of seven variables (educational attainment, participation rate of the total population 15 years and over, % management occupations, % walk - mode of transportation to work, % occupied private dwellings of type apartment, % lone-parent families, and mean NO_2 concentration). The only common variable to the three models developed, with the exception of Mean NO_2, is educational attainment. The R^2 (0.43) and the adjusted R^2 (0.38) for the neighbourhood structure are between the values generated by the OLS models for the census tract and the cluster structure. The log likelihood is larger for the natural neighbourhood than for the census tract and cluster structures while the AIC and Schwarz criterion have smaller values for the natural neighbourhood structure than the two other ones, all characterizing the neighbourhood model as a better fitted model. The model developed under the natural neighbourhood structure is also the only one of the three models not characterized by a non-normal distribution of the error term (Jarque-Bera) and non-stationarity between the explanatory variables and the dependent variables (Breusch-Pagan and Koenker-Bassett tests). In terms of global spatial autocorrelation, the only model that is not characterized by statistically significant spatial autocorrelation is based on the use of the natural neighbourhood structure.

5.4 Discussion

Numerous studies have concluded that an increased exposure to NO_2 is likely contributing to the negative health of individuals but there is yet no positive and statistically significant association that has been unequivocally measured (Villeneuve et al. 2006; Sahsuvaroglu et al. 2009; Clark et al. 2010). The main objective of this research was to study the role of spatial representation in the study of the relationship between exposure to NO_2 and health, which is closely tied to the MAUP, as a way to determine how scale and zoning are factors in non agreement on the role of NO_2 on health.

5.4.1 Understanding the effects of the MAUP

Some authors have recommended approaches to gain a better insight of the role of the MAUP on analytical results (Openshaw and Taylor 1981; Openshaw 1984; Fotheringham 1989) but few published studies that incorporated any of these in their analysis. In this research we implemented two of the recommended approaches (Jelinski and Wu 1996). The first approach is based on the use of an optimal zoning approach while the second suggests conducting sensitivity analysis. In this study the natural neighbourhoods delineated through a semi-automated approach were used as the “optimal” zoning system. As mentioned in section 2.4, the main component of the delineation process was based on the concept of internal homogeneity. Neighbourhoods were delineated in order to create homogeneous units in terms of SES. These units may not be appropriate for all studies conducted in Ottawa but because they capture many social processes associated to health, we believe they could be used for research looking into other health outcomes aside from respiratory morbidity or even other social processes related to SES.

This research also used sensitivity analysis to mitigate the effect of the MAUP (see Openshaw and Taylor 1981; Openshaw 1984; Fotheringham 1989). This study was conducted using three spatial representations: the natural neighbourhood, the cluster structure and the census tract. The reason for using this sensitivity analysis approach was two-fold. First, it allowed addressing a number of questions such as the effect of the MAUP on spatial autocorrelation (univariate and bivariate and regressions). Second, by studying this relationship under three different zoning systems it would permit corroboration of the results obtained in chapter five where NO_2 concentration was not found to be a main contributing factor to the respiratory health outcomes in Ottawa. That research was based only on the neighbourhood structure.

5.4.2 Lessons learned

Exploratory and confirmatory data analysis methods were used to assess the role of spatial representation in the study of the relationship between exposure to NO_2 and the respiratory health outcomes. The results obtained from the different analytical approaches converge towards three main conclusions, which will be discussed in more detail:

1. Exploratory data analytical methods, such as univariate Moran's I and bivariate Moran's I, can serve as an indication of the potential effect of the MAUP in the study of the relationship between exposure to NO_2 and health;

2. Bivariate and multivariate regressions show that the use of different spatial representations in the study of the relationship between exposure to NO_2 and health could be a contributing factor to the lack of agreement found in the literature about the specific relationship;

3. The use of three different spatial representations confirms that the lack of a measurable effect of NO_2 exposure on respiratory health in Ottawa is not due to a lack of attention to the spatial unit of measurement.

5.4.2.1 Explanatory data analysis

The results obtained confirm the documented effect of the MAUP (Arbia 1989; Amrhein and Reynolds 1997) on summary statistics. The analysis of the summary statistics shows that the MAUP does not have a strong effect on the mean of the explanatory variables, confirming the results of Jelinski and Wu (1996). Our results also corroborate the work of the same authors in which they found that the variance increases with the scale but can also vary when the same number of zones is used along with a different partitioning of space.

Global Moran's I was calculated for all the explanatory variables as well as for the dependent variable under the three spatial structures. All the explanatory variables are characterized by positive spatial autocorrelation using the three zoning systems. We also observe that Moran's I tends to be lower for the census tract structure than with the two other spatial structures. As explained by Jelinski and Wu (1996), as the level of aggregation increases Moran's I is also expected to increase due to the increased homogeneity in the "landscape structure". While not definitive, our results concur.

The variable respiratory health outcome rate, which is the dependent variable, displays low levels of statistically significant spatial autocorrelation using the census tract and the cluster structure. Global Moran's I, calculated for the dependent variable based on the natural neighbourhood spatial structure, shows a value close to zero and is not statistically significant. Instead of seeing the level of spatial autocorrelation increase with aggregation, it is here lowered and is no longer statistically significant. We believe this can be explained by the fact that the variability of the heath outcome is better captured through the variability of the census variables used in the delineation process. Hence, the natural neighbourhoods are not only internally homogeneous in terms of SES but also for the respiratory health outcomes, reducing spatial dependence between neighbourhoods. A zoning effect can also be observed; levels of spatial autocorrelation using the neighbourhood structure are generally lower than when using the cluster structure. These observations on the effect of Moran's I can serve as one more justification for the use of a custom geography that is delineated based on the processes associated with the dependent variable.

Under the MAUP, we expected both the cluster and the neighbourhood structures (n=95) to be characterized by generally stronger bivariate Moran's I values between the explanatory variables and the NO_2 concentrations than under the census tract structure (n=184) because aggregation has been showed to increase correlations (Openshaw 1984). The use of an optimal zone design based on the concept of homogeneity appears to be reducing the scale effect of the MAUP. For half of the explanatory variables the correlation with NO_2 concentrations is stronger at the census tract level than at the neighbourhood level. On the other hand the correlations measured using the cluster structure are generally stronger than under the census tract structure, which is the expected result (Openshaw 1984). Consequently, a relatively strong zoning effect is also observed when comparing the bivariate Moran's I values for the correlation between the NO_2 and health for the neighbourhood and the cluster structures. Similar results were obtained for the correlations between the explanatory variables and the health outcomes. This confirms the fact that homogeneous units capture social processes linked to respiratory health outcomes. We also found that the variable % retail trade industry has different directions of association depending on the spatial structure, confirming the work of Fotheringham and Wong (1991) who suggested that "correlation inference is not robust to the aggregation process".

The use of the exploratory analytical approach in the context of sensitivity analysis can be seen as a tool to assess the role of the MAUP prior to conducting confirmatory data analysis. It allows us to gain a more in-depth understanding of the potential effect of the MAUP on analytical results as the latter can have an effect on possibly most statistical modeling used in the study of the relationships between health and a variety of factors (Waller and Gotway 2004).

5.4.2.2 Bivariate and multivariate regressions

Bivariate and multivariate regressions confirm that the use of different spatial representations in the study of the relationship between exposure to NO_2 and health could be a contributing factor to the lack of consensus found in the literature regarding this relationship. Bivariate regressions between the NO_2 concentrations and the dependent variable once again confirmed both the zoning effect and the scale effect of the MAUP. According to the census tract and the cluster structures, there is a low but statistically significant relationship between NO_2 and respiratory health. As expected from earlier work on the MAUP, the R^2 value is higher under the cluster structure (higher level of aggregation) than with the census tract (Swift et al. 2008). The relationship measured between NO_2 and health cannot be confirmed using the neighbourhood structure (not statistically significant). The OLS model for the neighbourhood structure is also the only model that is not characterized by statistically significant positive spatial autocorrelation in the residuals and passes all tests for non-normality of the error term and heteroskedasticity.

The effect of the MAUP on multivariate analyses has been described as “complex and unpredictable” (Jelinski and Wu 1996). Considering the fact that the health of an individual is the result of many factors and not only exposure to NO_2, (Goovaerts 2010), it becomes a relationship that is complicated to measure. The question of the variables included in the multivariate regression models developed herein is an important element to be addressed since only one variable is common to all three models with the exception of the mean NO_2 which was forced into the models. The variable educational attainment is an explanatory variable for the variations of the morbidity rate for all three structures, meaning that changes in the value due to the modification of the scale and/or the zoning design are being filtered out. The geographic scale of the variable educational attainment is hence larger than the scale of the census tract. Another comparison to be made is between the variables used as input into the delineation of the optimal zoning design and the variables included in the different models. We observe that the census tract model includes both a housing period variable and an income variable, which were used in the delineation of the natural neighbourhoods. By creating an optimal zoning system based on these variables and other SES variables, housing and income covariates are no longer confounding variables in the relationship between exposure to NO_2 and health and hence are not found in the neighbourhood model.

The variable mean NO_2 is still a not statistically significant factor in the model but it gave room for other variables to show their association with respiratory health. For example, the variable % of lone-parents, a measure of the economic structure, is part of the optimal zoning design model but not the census tract or the cluster model. Briant et al. (2010) found that if the process of aggregating has the same impact on the independent variables and the dependent variable then the effect of the MAUP is lessened. These results are demonstrated when the independent and dependent variables are spatially autocorrelated and averaged. Since the level of global spatial autocorrelation was found to be different for the dependent variable depending on the structure, the differences observed between the models are justified. It is important to note that among the variables included in the neighbourhood model, two variables have a coefficient with a direction opposite to what was suggested by bivariate Moran's I (mean NO_2 and % lone-parent families). One last element that needs to be given attention is the resemblance between the census tract and the cluster models; the cluster model is a subset of the census tract model. According to Swift et al. (2008), aggregation process can introduce collinearities between independent variables, which would explain the composition of the cluster model as a subset of the census tract model.

Because of the effect of the MAUP, an OLS model developed for a specific spatial structure cannot be transferred to any other structure. Among others, this finding has implications for research that may use index mapping techniques to assess vulnerability to air pollutants. The sensitivity analysis conducted in this research clearly demonstrates that the explanatory factors of respiratory health will vary according to the geographical structure used to conduct the study. Since the research on the relationship between exposure to NO_2 and health has used a variety of geographical units to conduct the analysis, it is possible that some of the differences in the results reported are caused by the MAUP. We believe the delineation of a custom geography that is optimal in regards of the phenomenon studied can be justified since administrative and statistical boundaries do not grasp the fundamental nature of economic and social phenomena that often cross boundaries (Briant et al. 2010). The use of an optimal zoning design then becomes a way to reconcile the consequences of both the geographic and methodological scale, the former describing the geography at which social processes are identifiable and the latter the scale at which the data were collected and tabulated (Reardon et al. 2008).

5.4.2.3 The effect of NO_2 on respiratory health

The use of three different spatial representations confirms there is no measurable effect of NO_2 exposure in Ottawa. Bivariate OLS regression between the NO_2 concentrations and the health outcomes indicates a very low correlation between the independent and the dependent variable under all three spatial structures. In chapter five, no statistically significant association between NO_2 and health was found, but this finding does not necessarily mean that exposure to NO_2 in Ottawa is not affecting the health of residents. We had hypothesized that the lack of a relationship between exposure to NO_2 and respiratory morbidity may have been the result of inexact correspondence between the neighbourhood boundaries and the ecological properties that shape health processes (Sampson and Morenoff 2004; Sridharan et al. 2007). Based on the results obtained through sensitivity analysis, we can confirm that for Ottawa the lack of a significant statistical association is probably not induced by the use of a particular geography.

The methodological approach used also demonstrates that many factors are explaining the differences observed in the respiratory morbidity rates in the study area. Considering that respiratory health has already been associated to with SES in many studies (McElmurry et al. 1999), which was discussed in chapter five, our research serves as another example of the importance of the social and the environmental context on health.

There have been few studies on the role of spatial representation and to our knowledge this is the first one specifically interested by its effect on relationship between exposure to NO_2 and health. Among the modest research on the question of spatial representation, Briant et al. (2010) found similar results to this present study. The authors found that “the use of different specifications to assess spatial concentration, agglomerations economies, and trade determinants produces substantial variation in the estimated coefficients”. Cockings and Martin (2005), in a study on the relation between morbidity and deprivation also found that the use of spatial representations other than the census tracts provided different analytical results. On the other hand few studies on the role of spatial representation concluded that the way in which the neighbourhood was defined did not modify significantly the relationship and its strength. One of these studies was by Jones et al. (2010) who studied the role of area definition on neighbourhood variations in child physical activity. Ross et al. (2004) came to a similar conclusion when studying neighbourhood effects on health. Additional studies on the impact of using different operationalizations of the neighbourhood on analytical results are required to better understand the role of spatial representation. In return, a thorough understanding of the role of MAUP on the study of the relationship between NO_2 concentration and health will allow decision-makers to make interventions where it is the most needed. Policy makers' decision about how to improve the health of communities should be strongly influenced by the conclusions that neighbourhood quality affects health. This is particularly of interest for Canada where the population at large believes that the government has a responsibility for the health of citizens (Heyman and Fisher 2003).

5.4.3 Strengths and weaknesses

One of the strengths of this current study is the use of an optimal zone design that has been reviewed and approved by city planners, public health agencies, community health centers and grassroots organizations. These natural neighbourhood units have also been used for the Ottawa Neighbourhood Study and have been updated since their creation to reflect changes in Ottawa's communities.

Another strong point of this research is the use of automated zone design to create a cluster structure that we could use to compare to the census tract structure and our natural neighbourhood structure. One issue with the cluster structure is the fact that it had to be created by grouping census tracts due to the availability of the health data at that level. As a consequence, we did not have as much flexibility when creating the cluster structure as we could have if a smaller geographic level was used as the basic spatial unit for grouping. Another element that makes this study more complicated to conduct is the fact the NO_2 concentration is not a statistically significant explanatory variable in the multivariate regressions. If the situation would have been different, we could have got a better insight of how NO_2's association with respiratory health varies at different scales.

5.5 Conclusion

The natural neighbourhoods used in this research can be viewed as "exposure areas" as they were delineated with the objective of being homogeneous units from an SES perspective (Chaix 2009). The use of area level data such as income, education and housing variables from the Canadian Census, created units where environmental and SES conditions are similar. Area units should be delineated with a purpose to represent the expected relationship between neighbourhood and health. If this relationship is well defined, the modifiable area unit issue is can be lessened as problem (Haynes et al. 2007).

Our results indicate that both components of the MAUP, the scale effect and the zoning effect, can impact the analytical results when studying the relationship between exposure to NO_2 and respiratory health. Our results should be considered as a call for sensitivity analysis when using index mapping techniques to assess vulnerability to air pollutants as they demonstrate that explanatory factors of respiratory health will vary according to the scale and zoning of the structure used to conduct the study. Finally, this research confirms the work of other authors on the role of the MAUP on multivariate analyses. As described by Jelinski and Wu (1996) the effect of the MAUP is complex and more research on its effect in health studies is necessary.

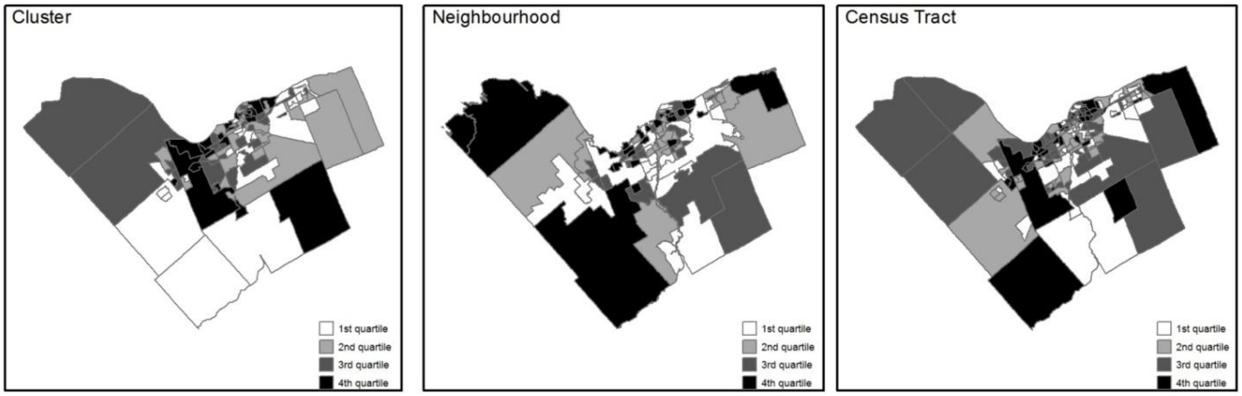


Figure 5.1 Morbidity rates for each spatial structure (using the quartile for classification).

Variables of the 2001 Canadian Census of Population
Variable
Median household income
Unemployment rate
Housing affordability
% Structures built before 1961
% Dwelling owned
Median value of dwelling
% Visible minority
% Population with a bachelor's degree

Table 5.1 Variables used for the automated component of the delineation process of the natural neighbourhoods.

	Natural Neighbourhoods (n=95)					Cluster (n=95)					Census tract (n=184)				
Variable	Mean	Std Dev	Minimum	Maximum	Variance	Mean	Std Dev	Minimum	Maximum	Variance	Mean	Std Dev	Minimum	Maximum	Variance
% Occupied private dwellings in need of major	5.8601	3.2902	0.3712	13.4766	10.8255	6.0687	3.4103	0.9600	13.5500	11.6298	6.0613	3.8870	0.0000	15.7300	15.1084
% Occupied private dwellings built before 1946	7.6504	13.0441	0.0000	72.8673	170.1490	9.0836	14.3778	0.0000	68.4300	206.7214	9.3770	16.5159	0.0000	77.9800	272.7753
Unemployment rate of the total population 15 years and over	6.1158	1.6008	2.9000	9.9000	2.5626	6.0383	1.9898	1.9500	11.6600	3.9595	5.9647	2.3903	0.0000	15.5000	5.7134
Educational attainment - no certificate, diploma or degree	15.6249	4.7142	6.4791	31.5308	22.2234	8.0593	4.5666	1.0500	22.1900	20.8540	7.9707	5.3707	0.0000	27.6600	28.8446
Median income (2005)	33062.3800	5677.3100	20229.0000	45670.0000	32231814.7500	34320.1300	7219.8600	14364.0000	48480.9100	52126423.9000	34169.7700	9242.0900	0.0000	59831.0000	85416209.5400
Average income (2005)	43376.4600	9764.1000	22736.8200	81390.4400	95337573.3800	44067.3600	9636.2200	19994.0000	76383.9100	92856748.0900	43967.1000	14575.1600	0.0000	134403.0000	212435368.0000
% low income	14.9365	10.7941	1.5479	49.9107	116.5124	9.2259	7.5727	0.0000	40.2000	57.3453	8.8060	8.7050	0.0000	51.6000	75.7776
% Lone-parent families	17.1233	7.0923	6.3417	38.9327	50.3010	16.5398	6.7321	5.7377	37.4468	45.3213	16.1917	7.7104	0.0000	48.3516	59.4505
% Owned dwellings	67.1669	24.6118	2.4416	98.3227	605.7429	68.7802	24.9516	5.2198	98.0200	622.5806	68.0649	27.3035	0.0000	99.2337	745.4809
% Rented dwellings	32.6802	24.6534	1.4513	97.6043	607.7895	31.1976	24.9167	2.1200	94.5055	620.8401	30.8565	26.5568	0.0000	95.3125	705.2636
% Occupied private dwellings of type single	44.4073	26.4335	0.1853	96.1708	698.7319	45.2950	26.1407	0.2747	97.3510	683.3360	45.1351	30.1448	0.0000	100.0000	908.7076
% Occupied private dwellings of type row house	20.7212	16.6610	0.0000	63.4444	277.5880	19.7553	15.8864	0.0000	83.0827	252.3767	18.8983	17.7605	0.0000	83.0827	315.4342
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	18.2560	21.1654	0.0000	84.9109	447.9725	16.1207	18.4369	0.0000	73.9300	339.9181	15.7290	21.3870	0.0000	96.7320	457.4020
% Did not lived at the same address one year ago	14.4924	6.1582	4.4657	36.4394	37.9238	14.6832	6.4719	5.4900	35.0300	41.8859	14.5430	7.8960	0.0000	55.6522	62.3461
% Did not lived at the same address five years ago	42.8642	11.2726	19.7927	78.4816	127.0707	43.3513	12.4817	22.3100	82.9200	155.7938	42.6931	15.5471	0.0000	90.5437	241.7131
% Walk - mode of transportation to work	7.2959	8.8507	0.0040	52.0332	78.3347	8.3850	10.5996	1.0100	52.5800	112.3524	8.2017	10.8372	0.0000	53.3613	117.4446
% Bicycle - mode of transportation to work	2.1788	2.0129	0.0000	9.7765	4.0516	2.3685	2.1737	0.0000	12.0200	4.7251	2.4026	2.3467	0.0000	12.8954	5.5072
Average value of owned dwelling	299264.5500	68185.8800	197876.3600	590363.4500	4649314326.0000	298357.7200	64832.1300	168309.0000	552521.7700	4203205353.0000	298744.3400	92328.6600	0.0000	935866.0000	8524582093.0000
Participation rate of the total population 15 years and over	68.3834	6.4545	47.7365	80.8734	41.6599	69.2452	6.3040	53.5200	81.1200	39.7401	68.1728	10.2091	0.0000	82.7000	104.2254
% Management occupations	11.3692	3.2586	4.9933	19.9972	10.6187	11.7220	3.2298	5.3400	21.6867	10.4317	11.5749	4.0322	0.0000	21.6867	16.2587
% Retail trade industry	10.2252	2.8070	2.7252	17.3572	7.8794	10.2328	2.6193	5.2800	19.0700	6.8605	10.0886	3.3894	0.0000	20.8791	11.4883
Mean nitrogen dioxide concentration	5.3196	1.7992	1.0530	10.5570	3.2370	5.2955	1.9103	1.0328	9.5151	3.6494	5.3989	1.9333	0.9563	10.1816	3.7378
Respiratory health outcome rate	1275.3500	782.6471	0.0000	3515.1000	612536.5500	2345.2600	1153.9900	550.3206	6207.1800	1331685.3700	2346.8900	1561.5600	0.0000	12023.9400	2438471.7500

Table 5.2 Summary statistics for the explanatory variables and the dependent variable.

Variable	Natural Neighbourhood (n=95)	Cluster (n = 95)	Census Tract (n=184)
% Occupied private dwellings in need of major repair	0.4337 (0.001)	0.5604 (0.001)	0.4800 (0.001)
% Occupied private dwellings built before 1946	0.5196 (0.001)	0.6182 (0.001)	0.5948 (0.001)
Unemployment rate of the total population 15 years and over	0.4695 (0.001)	0.3192 (0.001)	0.2367 (0.001)
Educational attainment - no certificate, diploma or degree	0.2518 (0.002)	0.3468 (0.001)	0.3397 (0.001)
Median income (2005)	0.3932 (0.001)	0.2489 (0.001)	0.2397 (0.001)
Average income (2005)	0.0561 (0.145)	0.1022 (0.045)	0.1139 (0.005)
% low income	0.5134 (0.001)	0.4518 (0.001)	0.2859 (0.001)
Mean nitrogen dioxide concentration	0.5215 (0.001)	0.5463 (0.001)	0.6656 (0.001)
Respiratory health outcome rate	0.0261 (0.288)	0.1775 (0.009)	0.1315 (0.002)
% Lone-parent families	0.3329 (0.001)	0.4183 (0.001)	0.3206 (0.001)
% Owned dwellings	0.5931 (0.001)	0.6714 (0.001)	0.5796 (0.001)
% Rented dwellings	0.5943 (0.001)	0.6715 (0.001)	0.6029 (0.001)
% Occupied private dwellings of type single house	0.4824 (0.001)	0.5444 (0.001)	0.4164 (0.001)
% Occupied private dwellings of type row house	0.1829 (0.007)	0.2871 (0.001)	0.2814 (0.001)
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	0.4607 (0.001)	0.5457 (0.001)	0.4177 (0.001)
% Did not lived at the same address one year ago	0.4538 (0.001)	0.4535 (0.001)	0.3808 (0.001)
% Did not lived at the same address five years ago	0.2851 (0.001)	0.3049 (0.001)	0.3151 (0.001)
% Walk - mode of transportation to work	0.8129 (0.001)	0.7489 (0.001)	0.7801 (0.001)
% Bicycle - mode of transportation to work	0.5865 (0.001)	0.6383 (0.001)	0.6757 (0.001)
Average value of owned dwelling	0.1418 (0.002)	0.1440 (0.013)	0.1547 (0.001)
Participation rate of the total population 15 years and over	0.5157 (0.001)	0.5141 (0.001)	0.2063 (0.001)
% Management occupations	0.1138 (0.045)	0.1763 (0.005)	0.2000 (0.001)
% Retail trade industry	0.1989 (0.002)	0.3199 (0.001)	0.2652 (0.001)

Table 5.3 Global Moran's I for the explanatory variables.

	Natural Neighbourhood (n=95)		Cluster (n = 95)		Census Tract (n=184)	
Variable	Mean NO2	Resp health	Mean NO2	Resp health	Mean NO2	Resp health
% Occupied private dwellings in need of major repair	0.3971 (0.001)	0.1329 (0.001)	0.4776 (0.001)	0.1798 (0.001)	0.4515 (0.001)	0.0899 (0.001)
% Occupied private dwellings built before 1946	0.3235 (0.001)	0.1045 (0.001)	0.4180 (0.001)	0.0636 (0.001)	0.3939 (0.001)	0.0001 (0.999)
Unemployment rate of the total population 15 years and over	0.3741 (0.001)	0.1799 (0.001)	0.3982 (0.001)	0.1446 (0.001)	0.3266 (0.001)	0.0363 (0.087)
Educational attainment - no certificate, diploma or degree	0.1518 (0.001)	0.0589(0.001)	0.2843 (0.001)	0.1609 (0.001)	0.2420 (0.001)	0.0663 (0.001)
Median income (2005)	-0.4019 (0.001)	-0.1598 (0.001)	-0.3765 (0.001)	-0.1111 (0.001)	-0.3031 (0.001)	-0.0583 (0.001)
Average income (2005)	-0.1739 (0.001)	-0.0598 (0.001)	-0.2272 (0.001)	-0.0435 (0.001)	-0.1613 (0.001)	-0.2200 (0.509)
% low income	0.5027 (0.001)	0.1685 (0.001)	0.4933 (0.001)	0.1215 (0.001)	0.4158 (0.001)	0.0302 (0.512)
Mean nitrogen dioxide concentration	1 ()	0.1905 (0.001)	1 ()	0.0656 (0.001)	1 ()	0.0365 (0.085)
Respiratory health outcome rate	0.0794 (0.001)	1 ()	0.1056 (0.001)	1 ()	0.1517 (0.001)	1 ()
% Lone-parent families	0.3353 (0.001)	0.1486 (0.001)	0.4181 (0.001)	0.1716 (0.001)	0.3481 (0.001)	0.0454 (0.700)
% Owned dwellings	-0.5492 (0.001)	-0.2271 (0.001)	-0.6191 (0.001)	-0.1823 (0.001)	-0.5872 (0.001)	-0.0946 (0.001)
% Rented dwellings	0.5507 (0.001)	0.2278 (0.001)	0.6188 (0.001)	0.1824 (0.001)	0.6008 (0.001)	0.0877 (0.001)
% Occupied private dwellings of type single house	-0.4173 (0.001)	-0.2118 (0.001)	-0.4982 (0.001)	-0.2071 (0.001)	-0.4592 (0.001)	-0.0779 (0.001)
% Occupied private dwellings of type row house	-0.2593 (0.001)	-0.0785 (0.001)	-0.2433 (0.001)	-0.0119 (0.999)	-0.2354 (0.001)	-0.0244 (0.087)
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	0.4756 (0.001)	0.2116 (0.001)	0.5433 (0.001)	0.1441 (0.001)	0.4728 (0.001)	0.0469 (0.110)
% Did not lived at the same address one year ago	0.4212 (0.001)	0.1737 (0.001)	0.4830 (0.001)	0.1223 (0.001)	0.4114 (0.001)	0.0599 (0.001)
% Did not lived at the same address five years ago	0.2344 (0.001)	0.1044 (0.001)	0.2876 (0.001)	0.0889 (0.001)	0.2155 (0.001)	0.0257 (0.484)
% Walk - mode of transportation to work	0.4795 (0.001)	0.1590 (0.001)	0.5175 (0.001)	0.0858 (0.001)	0.5531 (0.001)	0.0390 (0.117)
% Bicycle - mode of transportation to work	0.4458 (0.001)	0.1576 (0.001)	0.4949 (0.001)	0.1683 (0.001)	0.4553 (0.001)	0.0263 (0.001)
Average value of owned dwelling	0.1167 (0.001)	0.0292 (0.606)	0.0641 (0.001)	-0.0287 (0.0980)	0.0509 (0.001)	-0.0360 (0.096)
Participation rate of the total population 15 years and over	-0.3280 (0.001)	-0.1754 (0.001)	-0.3140 (0.001)	-0.1791 (0.001)	-0.1788 (0.001)	-0.1788 (0.001)
% Management occupations	-0.1845 (0.001)	-0.0967 (0.001)	-0.2724 (0.001)	-0.1577 (0.001)	-0.2309 (0.001)	-0.0597 (0.001)
% Retail trade industry	-0.2428 (0.001)	-0.0993 (0.001)	-0.1929 (0.001)	0.0377 (0.0440)	-0.1629 (0.001)	-0.0084 (0.999)

Table 5.4 Bivariate Moran's I under the three spatial structures.

	Natural Neighbourhood (n=95)	Cluster (n=95)	Census Tract (n=184)
	Bivariate RESP	Bivariate RESP	Bivariate RESP
R-squared	0.0307	0.0494	0.0492
Adjusted R-squared	0.0202	0.0392	0.0440
Sum squared residual	55810400.0000	118990000.0000	424291000.0000
Sigma-square	600112.0000	1279470.0000	2330000.0000
S.E. of regression	774.6690	1131.1400	1526.8500
Sigma-square ML	587478.0000	1252530.0000	2305000.0000
S.E of regression ML	766.4710	1119.1700	1518.5300
F-statistic	2.9462	4.8363	9.4154
Prob(F-statistic)	0.0894	0.0303	0.0025
Log likelihood	-765.7700	-801.7310	-1608.9800
Akaike	1535.5400	1607.4600	3221.9500
Schwarz criterion	1540.6500	1612.5700	3228.3800
Intercept coefficient	869.8488 (0.0007)	1634.0250 (<0.0001)	1379.7439 (<0.0001)
Mean NO2 coefficient	76.2276 (0.0894)	134.3078 (0.0303)	179.136 (0.0024)
Jarque-Bera	6.3023 (0.0428)	26.1702 (<0.0001)	923.152 (<0.0001)
Breusch-Pagan Test	5.1764 (0.0228)	1.5039 (0.2200)	0.1623 (0.6870)
Koenker-Bassett test	4.1263 (0.0422)	1.0065 (0.3157)	0.0270 (0.8693)
Moran's I (error)	-0.0039 (0.8723)	0.2454 (<0.0001)	0.1080 (0.0096)
Lagrange Multiplier (lag)	0.0003 (0.9842)	12.2297 (0.0004)	5.4663 (0.0193)
Robust LM (lag)	0.0112 (0.9155)	0.1264 (0.7221)	0.0668 (0.7959)
Lagrange Multiplier (error)	0.0025 (0.9599)	1.0239 (0.3115)	5.4253 (0.0198)
Robust LM (error)	0.0133 (0.9079)	13.2537 (0.0013)	0.0257 (0.8724)

Table 5.5 Results of the bivariate regressions (OLS) between the NO_2 concentrations (mean NO_2) and the respiratory health outcomes (RESP).

	Natural Neighbourhood (n=95)	Cluster (n=95)	Census Tract (n=184)
	Multivariate OLS	Multivariate OLS	Multivariate OLS
R-squared	0.4306	0.3947	0.4341
Adjusted R-squared	0.3848	0.3678	0.4150
Sum squared residual	32783700.0000	576190000000.0000	252486000.0000
Sigma-square	376824.0000	841798.0000	1426470.0000
S.E. of regression	613.8600	917.4960	1194.3500
Sigma-square ML	345091.0000	797493.0000	1372210.0000
S.E of regression ML	587.4450	893.0250	1171.4100
F-statistic	9.3999	14.6759	22.6379
Prob(F-statistic)	0.0000	0.0000	0.0000
Log likelihood	-740.4980	-780.2880	-1561.2200
Akaike	1497.0000	1570.5800	3136.4400
Schwarz criterion	1517.4300	1583.3400	3158.9500
Variables	n.a.	n.a.	n.a.
	Intercept 5327.4610 (0.0001)	Intercept 738.3682 (0.0192)	Intercept -1013.495 (0.0905)
	Mean NO2 -18.0183 (0.6855)	Mean NO2 36.4373 (0.5349)	Mean NO2 30.9425 (0.5706)
	Educational Attainment 78.4534 (0.0002)	% Occupied private dwellings in need of major repairs 179.7687 (0.0007)	% Occupied private dwellings in need of major repairs 177.3213 (<0.0001)
	Participation Rate -53.0824 (0.0003)	% Occupied private dwellings built before 1946 -29.6505 (0.0118)	% Occupied private dwellings built before 1946 -20.91902 (0.0163)
	% Management occupations -84.1282 (0.004)	Educational Attainment 73.4928 (0.0111)	Educational Attainment 113.1868 (<0.0001)
	% Walk - mode of transportation to work 40.1355 (<0.0001)		Average income 0.0151 (0.0442)
	% Occupied private dwellings of type apartment -16.7916 (0.0015)		% Retail trade industry 73.8654 (0.0161)
	% Lone-parent families -34.7820 (0.0157)		
Jarque-Bera	0.9239 (0.6300)	44.1697 (<0.0001)	1173.241 (<0.0001)
Breusch-Pagan Test	3.0898 (0.8765)	8.5867 (0.0723)	149.1566 (<0.0001)
Koenker-Bassett test	3.1269 (0.8730)	3.8557 (0.4258)	21.8956 (0.0012)
Moran's I (error)	0.0498 (0.2527)	0.1892 (0.0006)	0.1091 (0.0051)
Lagrange Multiplier (lag)	0.1953 (0.6584)	7.8384 (0.0051)	2.7432 (0.0976)
Robust LM (lag)	3.4593 (0.0628)	0.6030 (0.4374)	0.2288 (0.6324)
Lagrange Multiplier (error)	0.4047 (0.0554)	7.8074 (0.0052)	5.5355 (0.0186)
Robust LM (error)	3.6687 (0.0628)	0.5720 (0.4494)	3.0211 (0.0821)

Table 5.6 Results of the multivariate regressions (OLS).

Chapter 6 A Spatial Analysis of Long-Term Exposure to NO_2 and Respiratory Morbidity in Ottawa, Canada

Marie-Pierre Parenteau[(1)]

Michael Sawada[(1)]

[(1)] Department of Geography, University of Ottawa

Laboratory for Applied Geomatics and GIS Science, Department of Geography, University of Ottawa, Ottawa, Ontario K1N 6N5

Tel:+1 613 562-5800 Ext. 1040; Fax: +1 613 562-5725

In preparation.

6.1 Introduction

An increase in exposure to criteria air pollutant such as NO_2 can have a negative effect on the health of the lungs and other organs of the human body (Health Canada 2006). Both short-term and long-term exposure to NO_2 has been studied in a number of different locations around the world, mostly in North America and Europe. While short-term exposure has been associated with asthma attacks through the inflammation of the airways which confirms effects measured in clinical studies, long-term exposure has been linked to an increase in chronic obstructive pulmonary disease (COPD), cardiopulmonary mortality, asthma and more generally respiratory morbidity and mortality (Peden 1999; Thurston et al. 2009; Arain et al. 2009; Clark et al. 2010). Most importantly, the significance of exposure to NO_2 on the health status of Canadians remains unclear (Health Canada 2006).

In many studies, NO_2 is a proxy measure for traffic-related air pollution (Brook et al. 2007). Despite variations in methodological design, many studies have measured a positive and statistically significant association between NO_2 and acute and chronic respiratory disease (Brook et al. 2004; Jerrett et al. 2004; Gulliver and Briggs 2005; Zanobetti and Schwartz 2006; Holguin et al. 2007; Gauderman et al. 2007; Kim et al. 2007; Brook et al. 2007; Brook et al. 2008; Jerrett et al. 2009; Neupane et al. 2010). In Canada, using a time-series approach and a large sample of air pollution and emergency department (ED) visits from multiple centers in Canada, Stieb et al. (2009) found significant effects of air pollution on acute cardiac and respiratory presentations. In Europe, researchers have additionally found negative health effects during long exposure to air pollution (Sauerzapf et al. 2009).

Despite these findings, there is no general agreement on the role of NO_2 exposure on health status. Numerous studies have concluded that although increased exposure is likely contributing to the negative health of individuals, there is yet no unequivocally positive and statistically significant association (Villeneuve et al. 2006; Sahsuvaroglu et al. 2009; Clark et al. 2010). In some cases, a statistically significant positive effect was measured but only for a specific segment of the population (Luginaah et al. 2005; Sahsuvaroglu et al. 2009) while others have found no association between NO_2 concentrations and respiratory health outcomes (Hinwood et al. 2006; Lee and Ferguson 2009).

Much of the published work on the relationship between NO_2 exposure and health has focused on the differences between metropolitan centers, herein referred to as inter-urban differences. As such, at the opposite end, there is a poor understanding of the role that variations in NO_2 have on health across a given metropolitan area, herein referred to as intra-urban scale (Beelen 2008; Neupane et al. 2010). Likewise, in Canada there is little research focusing on the intra-urban association between NO_2 and health. Moreover, where such research exists, it is conducted at the geographic level of the census tract, which is generally used as a proxy for the neighbourhood. Another often neglected aspect of the association between NO_2 and health involves the type of environment in which the study is conducted. Most studies have taken place in large urban centers while few have focused on smaller centers for the obvious reason - smaller urban environments have less overall pollution production in most cases.

The majority of studies about social processes, environmental justice, and health operationalize neighbourhoods that closely follow existing political and statistical geographies like the census tract (Sampson et al. 2002; Stafford et al. 2008). Many authors have criticized this conceptual operationalization of the neighbourhood because census tracts represent imposed, irregular boundaries (Bonham-Carter 1994), within and beyond which social or health-related processes take place independently of those arbitrary boundaries (Ellen and Turner 1997; Germain and Gagnon 1999; Kawachi and Berkman 2004; Martin 2004; Clapp and Wang 2006; Willis et al. 2003). By definition, such arbitrary boundaries may have no *a priori* relationship to the spatial scale of variation for socioeconomic or health conditions within an urban area. At the same time however, some area unit or neighbourhood must be defined in order to measure and compare health and environmental variables. Recent research suggests that such static representations of a neighbourhood and its milieu, while convenient for analysis, do not capture the extent and variation of spatial health determinants (Rainham et al. 2010). These representation issues with environmental exposures have been known for some time and recent research on ‘healthscapes’ has quantitatively shown how representations like the census tract are most often not representative of individual level exposures in urban areas (Rainham et al. 2010). However, it is worth nothing here that there have been very few health studies using alternative spatial representations to census tracts or other arbitrary static spatial units. While concepts like healthscapes may capture exposure more accurately, their delineation requires considerable data on individual and group mobility that are beyond the scope of most modest population health research endeavors and so their application may not be feasible but remains a promising area of research. While the census tract is popular in Canada for comparative health research, the census tract is not ideal and at the same time the application and development of healthscapes is not equivocal. As such, a middle ground can be found that involves the creation of so-called natural neighbourhoods whose boundaries presumably coincide better with the spatial scale of socioeconomic processes within a city.

In this research we delineated natural neighbourhoods in Ottawa, Canada using a combination of functional and physical approaches with the addition of local real estate knowledge (e.g., Ottawa Multiple Listing Service (MLS) real estate board neighbourhood maps). Following numerous reviews of the boundaries by local community leaders, groups and municipal authorities, 95 neighbourhoods were delineated[10]. These neighbourhoods have been delineated in the context of a broader research project on the inequalities in health in Ottawa (www.neighbourhoodstudy.ca). Chapter two provides the methodological approach for the delineation of the natural neighbourhood that we use. As socioeconomic (SES) factors have been linked to a number of health conditions (Crouse et al. 2009), we hypothesize that the use of natural neighbourhoods that are more homogeneous in terms of SES will allow improved control for confounding effects. Also, in accord with Briant et al. (2010), we believe that zoning systems delineated

[10] A preliminary version of the neighbourhood boundary file was delineated in chapter two. This file was later updated and a new boundary file containing 95 units was created.

specifically to answer questions at a local scale should be favored over administrative or statistical zoning systems.

This research uses exploratory spatial data analysis (ESDA) and confirmatory spatial data analysis (CSDA) to measure the relationship between exposure to NO_2 and health. These methods have been developed specifically for the analysis of spatial data, which can be characterized by spatial autocorrelation (dependence) and non-stationarity (Anselin and Getis 1992; Unwin and Unwin 1998; Haining and al. 2010) of the processes over space. Poisson regression, a statistical approach commonly used in the study of the relationship between exposure to air pollution and health (Linn et al. 2000) was also used as a way to conduct sensitivity analysis. The ultimate goal of this research is to elucidate the relationship between exposure to NO_2 and respiratory health outcomes while also advancing one of the weakest theoretical areas of health and environment research, namely the conceptualization of place (Matthews 2008).

6.2 Methods

6.2.1 Health Data

The effect of air pollution on respiratory health can be measured through emergency room visits and hospital admissions (Lee and Ferguson 2009). The respiratory morbidity rate for individuals of age 15 and over within the Ottawa Public Health Unit is the main health outcome measure in this research. For this study, morbidity was preferred over mortality due to the larger amount of data available, thus increasing the statistical power to characterize associations with air pollution (Villeneuve et al. 2006).

Health conditions associated to exposure to NO_2 were identified from the 10th revision of the ICD (CIHI 2006) based on literature regarding the health effects of air pollution (Kampa and Castanas 2008). The health outcomes considered are from Chapter 9 (Diseases of the Respiratory System) and more specifically correspond to the block J40-J47 (Chronic Lower Respiratory Diseases). The *Discharge Abstract Database* (DAD) and *National Ambulatory Care Reporting System* (NACRS) data, for the health conditions listed above, for fiscal years 2005-2006, 2006-2007 and 2007-2008 were extracted by the Ministry of Health and Long-Term Care (MOHLTC) and compiled at the geographic level of the 95 Ottawa neighbourhoods. The DAD dataset contains demographic, administrative, and clinical data for hospital discharges and day procedures (CIHI 2010a). The second source of information for this research is the NACRS, which contains demographic, administrative and clinical data for ambulatory care visits (CIHI 2010b). Data on hospitalization and ambulatory care were used along with some of the individual level data (sex, age, postal code of the residence) for each patient.

The morbidity rates were directly sex and age standardized (Sridharan et al. 2007; Hu and Rao 2009) for the age groups 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84 and 85 and over.

6.2.2 Exposure Measures

In Ottawa, transportation accounts for approximately 85% of NO_x emissions (City of Ottawa, 2004). The air quality in Ottawa is considered relatively good with an annual mean NO_2 concentration well below the 24h 100 ppb standard set by the Province of Ontario's provincial ambient air quality criteria (Ontario Ministry of the Environment, 2008). Nevertheless, the Ontario Medical Association estimates a cost of $65 million a year for health services dispensed because of air pollution in Ottawa, Canada (City of Ottawa 2004).

In chapter three a land-use regression (LUR) model was developed for the mapping of NO_2 concentrations in Ottawa based on air pollution data obtained through the National Capital Air Quality Mapping Project. The LUR model used 29 Ogawa passive samplers that were deployed throughout the study area for a period of two weeks in the fall 2007 as well as the data collected at two permanent stations administered by the National Air Pollution Surveillance program of Environment Canada. A model-based approach was used for the data collected at the permanent stations prior to their integration into the LUR model because of missing data (from 0% to 10% depending on the station). The Data Augmentation (DA) algorithm (Tanner and Wong 1987) in the S+ MissingData library (Schimert et al. 2001) was used for the modeling. This Monte-Carlo model-based approach is made up of two steps that are applied iteratively until the algorithm converges. In a first part, the imputation step simulates the missing data based on the observed data and the current parameter estimates. The results of the imputation are then fed into the posterior step where a new set of parameters are calculated (Lanza et al. 2005).

The results of the DA algorithm were analyzed using the time-series plots and the autocorrelation function plots (Schafer 1997; Lanza et al. 2005). Following the analysis, the mean value at each permanent station was computed. These data were used as the mean NO_2 value for the two week period. From the measurements obtained with the passive samplers and an additional two permanent monitoring stations in Ottawa, along with auxiliary data, the LUR model for NO_2 concentrations was developed. While data on other pollutants were available, NO_2 was selected as the pollutant of interest because it is an essential marker for traffic-related air pollution that is strongly correlated with other common air quality indices (Jerrett et al. 2007; Piro et al. 2008; Crouse et al. 2009).

The measured concentrations of NO_2 at the sampling locations were used in a multiple least-squares regression model (Ryan and LeMasters 2007). A preliminary analysis of the data was conducted using global Moran's I and bivariate regressions to look at the absolute strength of the variables and possible deviations from linear relationships with NO_2 by examination of residual plots (Poplawski et al. 2009). This analysis provided us with initial information on the relationship between the variables and measured NO_2 concentrations as well as a way to confirm the anticipated direction of effect (positive or negative) (Henderson et al. 2007). Stepwise selection was used to develop the different regression models (Jerrett et al. 2007; Su et al. 2008). Variables in the model had to be statistically significant and models were evaluated based on the R^2 values, RMSE, variation inflation factor, eigenvalues and the condition index. All models were tested for spatial autocorrelation in the residuals using Moran's I (Bailey and Gatrell 1995) and results indicated that no significant spatial autocorrelation was present. Details on the methodological approach for the development of the LUR model for Ottawa can be found in chapter three. The model, which included data on the road network, population, greenspaces and industrial land-use, yielded an R^2 of 0.8055. Figure 3.4 shows the output of the regression model at a spatial resolution of 20 meters. Average NO_2 concentrations by neighbourhood were obtained using zonal statistics within a GIS (Hu and Rao 2009).

6.3 Results

When there is spatial dependence or spatial autocorrelation in two spatially coincident datasets, the correlation between observations can often lead to a statistically significant relationship being uncovered when in fact no such relationship exists (a Type-I error). Correlated observations are non-independent and so contain less statistical information than an equal number of independent observations would imply and, as such, leads to the null hypothesis being rejected when it is actually true (Waller and Gotway 2004). Because of the particular characteristics of the data used in this research, which are spatial in nature, the association between exposure to NO_2 and health was assessed using a set of methods specifically developed for spatial data that take into account spatial autocorrelation. These methods can be divided into two groups: exploratory spatial data analysis (ESDA) and confirmatory spatial data analysis (CSDA). Following CSDA, a Poisson regression was used to assess the relationship between long-term exposure and respiratory health as a way to conduct sensitivity analysis.

6.3.1 Factors affecting the susceptibility of the population

The World Health Organization (2004) identifies population susceptible to air pollution based on innate factors, acquired environmental, social or behavioural factors, and levels of exposure. Of importance is the fact that both innate and acquired factors can have an effect on the level of exposure and the physiological response to the exposure (Makri and Stilianakis 2008). This implies that susceptibility is a function of a set of indefinite environmental and individual variables, not all of which are easily identified outside clinical studies (*c.f.* Hasterberg et al. 2009). In accord with such reasoning and based on review of previous research on respiratory health in general and environmental interactions, 26 neighbourhood level variables associated with susceptibility were considered for this research (Lwebuga-Mukasa et al. 2004; Jerrett et al. 2008; Jerrett et al. 2005b; Neupane et al. 2010; Beelen et al. 2008; Godley et al. 2010; Benzeval 1998; Rahkonen et al. 2005; Oyana and Rivers 2005; Buzzelli and Jerrett 2007; Makri and Stilianakis). These 26 variables are listed in Table 6.1 along with a justification for their inclusion.

Contextual variables at the geographic level of the neighbourhood, which are characteristics of the local environment (e.g. natural, built, social) that can influence the health of individuals in ecologic studies such as this one (Sahsuvaroglu et al. 2009; Jerrett et al. 2005b; Diez-Roux 2002; Piro et al. 2008) were obtained through custom tabulations of data from the 2006 Canadian Census of population. Compositional variables, which are aggregates of resident characteristics (e.g. income, age, education) that could have an effect on the measured association between air pollution and health (Diez-Roux 2002; Macintyre et al. 2002), were obtained from the 2006 Canadian Census of population and the Canadian Community Health Survey (CCHS). With a sample size of 65,000 respondents across Canada, the CCHS is a cross-sectional survey designed to “provide reliable estimates at the geographic level of the health region”, an area much larger than the neighbourhood (Statistics Canada 2010). Since the sample of each cycle of CCHS is small, a database was created for the City of Ottawa Health Unit using data from cycle 1.1 (2001), cycle 2.1 (2003), cycle 3.1 (2005) and cycle 4.1 (2007-2008) at the geographic level of individual Ottawa neighbourhoods. From the CCHS, the compositional variables considered were the % of smoker, the % of overweight individuals, % of obese individuals and % of obese or overweight individuals (Richmond and Subramanian 2009). After removing neighbourhoods with a sample size smaller than 10 individuals, 88 neighbourhoods were retained for analysis.

Listed in Table 6.2 are the summary statistics for the contextual and the compositional variables based on the neighbourhood level data. The range of the mean NO_2 concentrations is small, especially relative to other studies, with a minimum value of 1.1 ppb and a maximum value of 10.6 ppb (Jerrett et al. 2008; Neupane et al. 2010). Standardized respiratory rates for the neighbourhoods range between 101.42 – 3,515 cases per 100,000 of respiratory diseases with an average of 1,369 cases.

6.3.2 ESDA

ESDA methods were used to identify spatial outliers, clusters and hot spots (Anselin 1998) in the health outcome data, exposure data and covariates. Using GeoDa (Anselin et al. 2006), global and local measures of spatial autocorrelation were applied to all the covariates under study as well as to the health outcome rates and the NO_2 concentrations. ArcGIS 9.2 was used to perform Geographically Weighted Regression (GWR) (Fotheringham et al. 1998; Anselin 1995).

6.3.2.1 Univariate Global Spatial Autocorrelation

A test for global spatial autocorrelation was first conducted on each of the 26 covariates to determine if values separated by an arbitrary definition of adjacency (e.g., distance) are more similar to each other than we would expect by chance, that is to say under a random distribution by calculating a pseudo p-value (Anselin et al. 2007; Loughnan et al. 2008; Jerrett et al. 2010). This approach is based on random permutations of the data for which Moran's I is calculated for each permutation. The result is a set of Moran's I values that forms a reference distribution against which the pseudo-significance level is calculated (Sridharan et al. 2007). Assessing the presence and strength of spatial autocorrelation is crucial because traditional statistical tests should not be applied sensu stricto in the presence of spatial dependency. Using a first-order Queen's case row-standardized spatial weight matrix for the neighbourhoods and 999 permutations, we found statistically significant spatial autocorrelation at a pseudo-significance level of $p \leq 0.05$ for all the contextual variables, with the exception of % occupied private dwellings in need or major repair and % occupied private dwellings built before 1946 (Table 6.3). However, Moran's I for the respiratory outcome was not statistically significant.

Compositional data from the CCHS (% smoker, % overweight individuals, % obese individuals, and % overweight or obese individuals) exhibited negative spatial autocorrelation in the case of the % of overweight individuals and was not statistically significant (Table 6.3). The % of smoker variable was characterized by positive and statistically significant spatial autocorrelation. Finally the variable average income, a compositional factor from the Census, was not statistically significant.

6.3.2.2 Univariate Local Spatial Autocorrelation

A measure of local spatial autocorrelation was applied to the respiratory health outcome data and the NO_2 concentration data in order to identify local clusters and local outliers. Local Moran's I is considered a LISA (local indicator of spatial association) and allows for the identification of the degree of influence that individual neighbourhoods have on the magnitude of the global Moran's I statistic as well as the identification of statistically significant outliers or 'hot/cold spots' when assessed using a conditional permutation approach (Anselin et al. 2007; Anselin 1995). The hot/cold spots are represented spatially by a core (Anselin 2005a) neighbourhood around which the values of a variable (including the neighbourhood's value) tend to be higher (hot) or lower (cold) than average - indicating positive spatial autocorrelation - or mixed (high-low or low-high) - indicating negative spatial autocorrelation.

As with the global measure of spatial autocorrelation, local Moran's I used 999 permutations at a pseudo-significance level of $p \leq 0.05$ to assess the presence and strength of local spatial autocorrelation. In the case of the exposure data, the significant high value clusters (high-high) units are located in the downtown area and the low value clusters (low-low) units are found outside the urban core in the protected greenbelt region and beyond (Figure 6.1). As such, the significant positive spatial autocorrelation in NO_2 is dominated by the similarity in values downtown and outside of the Ottawa greenbelt. Three outliers are found within the territory for the exposure data; two low-high values (low surrounded by high values) and one high-low (high value surrounded by low values). In the health outcome data, which did not exhibit significant first-order spatial autocorrelation, four high-high clusters or hot spots and three low-low clusters or cool spots are found. In terms of outliers, there are three low-high units and five high-low outliers.

6.3.2.3 Bivariate Global and Local Moran's I

Global bivariate Moran's I was used to look at the absolute strength of the relations between the confounding variables, health outcomes, and NO_2 concentrations (Hu and Rao 2009; Uthman 2008). Bivariate Moran's I was preferred over Pearson's r product-moment correlation coefficient to assess the strength of the relationship between variables because the data exhibit spatial autocorrelation, meaning that there is violation of the assumption of independence. Using global Moran's I, a positive and significant association (I = 0.2458, pseudo-$p < 0.001$) is found between the NO_2 concentrations and the age-sex standardized rates of the respiratory outcomes (Table 6.4). Results suggest that the mean NO_2 concentrations at the neighbourhood level have the strongest association with variables that measure the socioeconomic status (SES) of the household (% low income, unemployment rate, median income, % lone parent families), housing (% owned dwellings, % rented dwellings, % occupied private dwellings of type apartment and % occupied private dwellings of type single house, % did not live at the same address one year ago) and transportation to work (% Walk – mode of transportation to work, % bicycle – mode of transportation to work). The confounding variables showing the strongest correlation with the respiratory morbidity rates are measures of the SES (unemployment rate, % low income) and housing (% rented dwellings, % owned dwellings, % occupied private dwellings of type single house, % occupied private dwellings of type apartment). Again the variables % bicycle – mode of transportation to work and % did not live at the same address one year ago display a strong association. While all the aforementioned variable relations are statistically significant, the majority shows strong positive or strong negative relations with NO_2 but weak positive or weak negative associations with respiratory health. Many of the contextual and compositional variables have an association both with the NO_2 concentrations as well as with the respiratory health outcomes, indicating that they could be confounding as they exert an effect both on the dependent variable and the independent variable mean NO_2 (Jerrett et al. 2008).

The bivariate LISA measures using local Moran's I - with 999 permutations for each neighbourhood and a pseudo-$p < 0.05$ - indicate that the linear association between the NO_2 concentrations and the respiratory health outcome rates at neighboring locations is high-high in the east part of the downtown area as well as in three other locations west of the downtown area (Figure 6.2). The association is low-high in two locations and high-low in three locations east of the downtown area. The association is low-low in some of the neighbourhoods contiguous to the greenbelt and high-high in a few other locations within the greenbelt area.

6.3.2.4 Geographically Weighted Regression

Geographically Weighted Regression (GWR) can be considered another exploratory approach for assessing non-stationary spatial relations (Fotheringham et al. 1998; Brunsdon et al. 1996). Using this methodological approach, the regression parameters are not fixed through space, they are allowed to very spatially and as a consequence R^2 is fitted locally. We used this approach to explore whether the significant positive bivariate spatial relationship between NO_2 and respiratory health outcomes is spatially stable or whether it fluctuates across the study area. Figure 6.3 shows a clear West-East trend in the variation of the local R^2 where the relationship is stronger in the West. The coefficient values of the (NO_2) concentrations are largest east of the city but the range of variation is small, between 76.226 and 76.263.

6.3.3 CSDA & CDA

Confirmatory spatial data analysis (CSDA) and Confirmatory Data Analysis (CDA) were the next steps in the assessment of the relationship between exposure to NO_2 and respiratory health. The contextual and compositional covariates from the Canadian Census of population and the CCHS were considered as explanatory variables in the relationship under study. CSDA methods were used as they allow the statistical analysis of data with spatial dependencies. A Poisson regression, a

CDA method, was also included in this assessment for the relationship between NO_2 and health.

6.3.3.1 Bivariate Regression

Ordinary Least Squares (OLS) was used to measure the bivariate relationship between the mean NO_2 concentrations and the respiratory health outcomes (Table 6.5). The measured R^2 is small (0.0652) but statistically significant ($p < 0.016$) and the coefficient of the mean NO_2 concentration is positive (104.3807), as expected from global bivariate Moran's I. The Jarque-Bera test indicates a non-normal distribution of the error term (statistically significant) while the Breusch-Pagan and the White tests (both statistically significant) indicate non-stationarity between the explanatory variable and the respiratory health outcomes (dependent variable). Global Moran's I shows no statistically significant spatial autocorrelation in the model residuals. As the analysis of the OLS points toward the existence of heteroskedasticity, spatial regression was performed despite the lack of statistically significant spatial autocorrelation in the residuals. A spatial regression with spatial lag was undertaken because the analysis of the OLS suggests that a spatial lag is more pronounced than the spatial error. The spatial lag used was a first-order Queen's case contiguity for the neighbourhoods. This type of regression is used when the dependent variable in one location influences the value of the variable in surrounding locations (Anselin 2002). An analysis of the results indicates a better fit of the model using spatial regression in the case of the log likelihood as it increased. The Akaike Information criterion (AIC) also increased along with the Schwarz criterion while the coefficient of NO_2 decreased and remained statistically significant (p < 0.001). Finally, Rho is low and positive but not statistically significant.

6.3.3.2 Multivariate regression

Because SES can be associated with multiple health conditions (Adler et al. 1994; Adler and Newman 2002; Evans and Kantrowitz 2002; Sridharan et al. 2007) we evaluated the contribution of SES and other covariates through multivariate regression for the relationships uncovered. SAS (9.2) was used to perform backward, forward and stepwise selections for the OLS model building exercise (Table 6.6). The mean NO_2 concentration variable had to be forced into these models since it was not statistically significant. Forcing in the NO_2 variable had the effect of making the variable % Rented not statistically significant at the 95% confidence level. The final OLS model, which included the mean NO_2 variable, % Rented dwellings, educational attainment, and % smoker, yielded an R^2 of 0.38 and an adjusted R^2 of 0.35. The multivariate spatial lag model was then executed in GeoDa (Anselin et al. 2006). Rho, the spatial lag, has a negative and statistically significant value. The NO_2 variable was still not statistically significant and was forced into this model. Other characteristics of the spatial regression model are very similar to the OLS model.

A second statistical approach was used to assess the relationship between exposure to NO_2 and respiratory health as a way to conduct sensitivity analysis. In the literature on long-term exposure to NO_2, Generalized Linear Models (GLMs) are a class of model that is common and are defined as “an extension of traditional linear models that allows the mean of a population to depend on a linear predictor through a nonlinear link function" (SAS 2008). In order to confirm that our results were not caused by the analytical approach we used, which is not as common as GLMs, a Poisson regression was run in SAS 9.2 using PROC GENMOD and a log link function (Table 6.7). To take into account that multiple observations for the same person may be correlated, we used a Generalized Estimating Equation (GEE). The Poisson regression provided the same results as those obtained with the spatial regression approach; when included in the model the mean NO_2 variable was not statistically significant and thus did not contribute significantly to the model.

6.4 Discussion

ESDA methods indicated that there is both global and local spatial autocorrelation for the mean NO_2 concentrations within Ottawa; local Moran's I identified significant hot spots in the downtown area (high-high neighbourhood units). In terms of the respiratory health outcome data global Moran's I indicated no statistically significant spatial autocorrelation but local Moran's I did identify hot spots of higher morbidity rates, but they were fewer and all located within a small geographic area. We consider the lack of global spatial autocorrelation in the dependent variable as an indication that our natural neighbourhood boundaries adequately captured the within-unit homogeneity while maximizing the heterogeneity between-units in terms of the health outcome data. Furthermore, the spatial autocorrelation present in the NO_2 concentrations suggests that this variable has a larger spatial range that naturally spans multiple neighbourhoods. Using bivariate Moran's I (global) a statistically significant association between the mean NO_2 concentrations and the respiratory outcomes was measured. As explained by Sridharan et al. (2007), it is important to note that this result does not imply an association between NO_2 and health but indicates that the patterns of NO_2 concentrations in Ottawa may be involved in the respiratory outcomes for the period under study. GWR, our last method of ESDA, was used in a bivariate regression (NO_2 and respiratory health). It indicated a West-East trend in local R^2 but the range of the values was small. The values of the local R^2 were low, indicating again an overall weak association between exposure to NO_2 and health.

Global bivariate Moran's I between the mean NO_2 concentrations and the respiratory health outcome is weakly positive (I = 0.2458) and statistically significant (pseudo-p < 0.001). On the other hand, we observe strong associations between the NO_2 and the confounding variables, especially the ones representing SES measures (economic activity and housing). These findings are repeated with the respiratory outcomes but tend to be weaker relations. The association between NO_2 and the respiratory health outcomes may in fact be the result of the SES variables being strongly correlated with NO_2, causing a false positive relation (Lindgren et al. 2009). Bivariate OLS suggests the same results; a weak statistically significant association is measured between NO_2 and respiratory health outcomes (R^2 = 0.0652). As spatial data often require the use of statistical approaches that take into account adjacency relations, two spatial regression models were tested even though the results of the OLS did not indicate the presence of spatial autocorrelation in the residuals. Finally, multivariate Poisson regression was also performed for completeness. In both these cases, the mean NO_2 variable had to be forced into the model and was not statistically significant.

6.4.1 The role of SES

Our present results appear to be identifying SES as an important confounding factor in the differences in the respiratory health outcomes in the neighbourhoods of Ottawa. Again, numerous studies have reported an association between SES and health (Shankardass et al. 2007; Chaix et al. 2010; Subramanian et al. 2011; Eisner et al. 2011; Salm Ward et al. 2010; Minet Kinge and Morris 2010; Lorant et al. 2003; Lynch et al. 1997; Ross and Wu 1996; House et al. 1994; Winkleby et al. 1992). In a study by CIHI (2008) using age standardized hospitalization data from 2003-2004 and 2005-2006, "hospitalization rates for CODP in the low-SES group were found to be 2.7 times of those of the high-SES group". Jerrett et al. (2004), Martin (2004) and Crouse et al. (2009) studied intra-urban variation in air pollution and mortality. In all cases SES was found to have a modifying effect on the results. Modifiable risk factors such as smoking also tend to be correlated with the socioeconomic status of individuals (Bertazzon et al. 2010).

The mechanisms explaining the larger effect of NO_2 for individuals with lower SES have yet to be confirmed (Adler and Newman 2002) but O'Neill et al. (2003) have formulated three possible pathways:

1) individuals in neighbourhoods of a lower SES may be more exposed to air pollution.

2) individuals in neighbourhoods of a lower SES may be more susceptible to air pollution because of already compromised health.

3) individuals in neighbourhoods of a lower SES are more likely to have conditions and traits that increase their vulnerability to air pollution.

This study cannot confirm the pathways described above but the results of this research indicate that the role of SES is more important on health status than the level of exposure to NO_2, which would confirm the results of Ramos et al. (2006). In terms of the biological pathways between SES and health, the mechanisms explaining this relationship are still not fully understood but research on the subject is ongoing. One area of explanation that is growing in popularity is the role of psychosocial factors (Chen et al. 2006; Clougherty et al. 2007; Chen et al. 2009. This explanation hypothesizes an indirect pathway between SES and health where: 1) SES is associated with stress and; 2) stress is associated with the immune dysregulation. For example, according to this explanation, stress will amplify the development of asthma through “activation of autonomic, hypothalamic-pituitary-adrenal, and immune pathways” (Chen et al 2006).

The multivariate regression model, which is made of a combination of contextual variables, compositional variables and NO_2 concentrations, explains 38% of the variation in the respiratory morbidity rates. The variables included in the model are the % Rented dwellings, % Smoker, educational attainment and Mean NO_2. Variables on the educational attainment and the housing situation can be considered measures of the area SES. Housing characteristics such as the average value of the owned dwelling and the tenure type are measures of the economic situation of the households in the different neighbourhoods (Adler and Newman 2002; Lwebuga-Mukasa et al. 2004). On the other hand, education is considered one of the main indicators of SES, potentially affecting future outcomes such as occupation and income (Adler and Newman 2002; Willis et al. 2003). Finally, smoking has been associated both with health status (Mokdad et al. 2004; Rabe et al. 2007; Peat et al. 1987; Mannino 2002; Mannino and Buist 2007) and SES (Winkleby et al. 1992; Sundquist et al. 1999; Reijneveld et al. 1998; Wardle et al. 2003; Gilman et al. 2003).

Our results does not necessarily mean that exposure to NO_2 in Ottawa is not affecting the health of its residents. Many confounding factors are at play in the study of the relationship between NO_2 and health, which makes it very complex to measure (Goovaerts 2010). For example, neighbourhoods of lower SES are more likely to be exposed to higher NO_2 concentrations (Buzzelli and Jerrett 2007) but there is also a higher likelihood of engaging behaviours with a detrimental effect on health. Hence, the pathway between exposure to NO_2 and respiratory health outcomes is intricate. Furthermore, not everybody whose health is affected by air quality will be visiting the hospital. Consequently not only is the statistical power decreased by the smaller volume of data but some neighbourhoods may be underrepresented if the constituent population is less mobile.

6.4.2 NO_2 concentrations

Another important element that could potentially explain the low level of association that we have measured between NO_2 and respiratory morbidity is the level of variation in the NO_2 concentrations (Dominici et al. 2003). The mean NO_2 concentrations in Ottawa neighbourhoods range from a minimum value of 1.1 ppb to a maximum value of 10.6 ppb. In a study on the association between long-term exposure to NO_2 and respiratory hospital admissions in Scotland, Lee and Ferguson (2009) found evidence of an association in Edinburgh and Glasgow but not in two other smaller urban centres. They believe that the low level of variation of the NO_2 concentrations in the two smaller centres could be one of the reasons explaining their results. There is also a lack of understanding on the nature of the relation between NO_2 and health, for example, is the health response linear with increasing NO_2 or is the relationship quantized as NO_2 thresholds are surpassed as a function of health outcome. Such stimulus-response research needs to be addressed and will become possible through meta-analyses as the body of research grows on NO_2 exposure and population health across different urban areas with variable air pollution concentrations.

6.4.3 Study Design

This research was designed as an ecologic study and as a consequence the level of interest was the area (group). Since variations in the health outcome have been found to be less at the group level than at the individual level (Oakes 2004; Villeneuve et al. 2006), it is possible that our approach repeated at the individual level would produce different results. Our study can be considered opportunistic (Dominici et al. 2003); it had to be designed according to the data that were available. Also, we did not take into account individual exposures which could be an important factor (Kwan 2009). The lack of individual level data is one of the main limitations of this study. Another potential issue with this research is the use of the CCHS as the source of information for compositional data. CCHS was the only source of data available but it is possible that because of its sampling design, even when using all available cycles, the sample size in Ottawa is not large enough for study at the scale of the neighbourhood.

One element that was not explicitly considered in this research is the role of the social environment on respiratory health. Previous studies have found an association between greater social capital and social cohesion and lower mortality/morbidity rates (Kawachi et al. 1997; Kawachi et al. 1990; McNeill et al. 2006; Marmot 2005). The development of ecometric measures, allowing the incorporation information on the social environment at the neighbourhood level, appears to be a logical step for further research.

Finally, it is important to note that our results are not sensitive to the statistical approach used to assess the relationship between exposure to NO_2 and health. Both the spatial regression and the Poisson regression identify NO_2 as a small contributing factor that is not statistically significant in explaining the different rates of respiratory health outcomes.

6.4.4 Temporal Discontinuity and Misrepresentation

Temporal discontinuity and misrepresentation could be one of the reasons behind the lack of a statistically significant association between NO_2 and respiratory health outcomes. Air quality data were acquired in the fall 2007 while the health data were collected between 2005 and 2008. The problem is even more important as the Census data were collected in 2006 and the CCHS data between 2001 and 2008. The different periods of reference could be affecting the results of this study. All variables are temporally aggregated and therefore suffer from relative issues of under or over-representation. Moreover, the NO_2 data may not be representative of the longer term spatially averaged situation in Ottawa. We do however believe that the spatial pattern is representative of the NO_2 concentrations over time but the absolute magnitudes, minimum and maximum differences in space and extreme events could be misrepresented. Further monitoring over longer time periods is necessary within the city to better characterize NO_2 long-term exposure along with more spatially detailed health outcome data collection.

6.4.5 Policy implications

As mentioned earlier, our results do not mean that we can rule out any association between NO_2 and respiratory health outcomes in Ottawa. It means that with the data available for this research we could not confirm that NO_2 is a significant factor affecting respiratory health in Ottawa (Wilson et al. 2004). Still, our findings are of interest in terms of policy. Ottawa is a city that has been witnessing a lot of changes and like other cities across Canada, the level of inequalities in health is a continuing process that is likely increasing in range. Our results indicate that there are differences in the health as a function of socioeconomic levels. Perhaps access to health care and other services for individuals within lower SES neighbourhoods could reduce differences observed in this study. However, given the lack of any causal mechanisms uncovered, it is equally likely that targeted health living campaigns, interventions and community support could equally or in combination aid in reducing the inequalities observed. City planning and regulations could further aid in that goal if further research into why exposure tends to be higher in lower SES neighbourhoods is undertaken.

6.5 Conclusion

As research attempts to uncover the effect of exposure to NO_2 on health, few have investigated the relationship at the intra-urban level. A spatial approach was used to measure the association between NO_2 and respiratory health outcomes at the neighbourhood level in Ottawa for individuals 15 years old and over. A regression model was developed with the objective of identifying the factors explaining the variability of the health outcomes. This model used data from the Canadian Census of population, the Canadian Community Health Survey and NO_2 monitoring data obtained from the City of Ottawa. The final model (R^2 = 0.38) is made up of four variables: % Rented dwellings, % Smoker, educational attainment and Mean NO_2. The variable NO_2 concentration was forced into the model as it is not statistically significant. It is important to note that our results are not sensitive to the statistical approach used to assess the relationship between exposure to NO_2 and health as evidenced by similar results across models. Our results suggest that, at this time, area SES has a more important role to play in the health status in Ottawa then exposure to NO_2.

Variable Category	Justification	Variables
Housing	The quality of the housing stock can directly affect respiratory health. For example, the presence of mold has been associated with an increase in respiratory health outcomes (Strachan 1988; Jaakkola et al 1993; Howden-Chapman 2004). Housing tenure is also associated with differences in health status of individuals (Sundquiest and Johansson 1997; Macintyre et al 1998). This association is explained by the fact that tenure is a maker for income but it may also have health promoting/damaging effects (Macintyre et al 1998; Ellaway and Macintyre 1998; Howden-Chapman 2004). Certain housing characteristics of the housing stock, such as its value, may limit the ability to minimize exposure (Spengler et al 2004). Location of the residence may have a direct impact on respiratory health (Brunekreef et al 1997; Oyana et al 2004; Lwebuga-Mukasa et al 2004).	% Owned dwellings % Rented dwellings % Occupied private dwellings of type single house % Occupied private dwellings of type row house % Occupied private dwellings of type apartment (building that has 5 or more storeys) % Occupied private dwellings in need of major repair % Occupied private dwellings built before 1946 Average value of owned dwelling
Socioeconomic status	More health problems, including lung functions, have been reported by individuals of lower social classes (Lang and Polansky 1994; Gottlieb et al. 1995; Marcon et al 2009). The same patterns have been found for education status and socioeconomic variables such as income and employment status (Carr et al. 1992; Kind et al 1998; Sundquist and Johansson 1997). Manual occupational social class and educational qualifications have also been found to be predictors of lower lung functions (Shohaimi et al 2004; Bohadama et al 2011; Kogenivas et al. 1999; Blan and Toren 1999; Trupin et al 2003; Bang et al. 2009). Marital status (widowed, separated, divorced) has also been associated with the health status of population; it may indirectly affect the respiratory health of individuals (Kind et al. 1998).	Unemployment rate of the total population 15 years and over Educational attainment - no certificate, diploma or degree Median income Average income % low income Participation rate of the total population 15 years and over % Management occupations % Retail trade industry % Lone-parent families
Mobility	Mobility can negatively affect the general health status of individuals (Abada et al. 2007), making them more susceptible of developing respiratory health issues when exposed to air pollution. The lack of social cohesion associated with high mobility could be affecting the health status of individuals (Elliott 2000), making them more vulnerable to exposure.	% Did not live at the same address one year ago % Did not live at the same address five years ago
Transportation	Depending on the duration and the intensity of the activity, walking and cycling to work can increase exposure in comparison to other modes of transportation by increasing the amount of pollutant absorbed by the body (Chertok et al 2004; Dirk et al 2009; Marshall et al 2009).	% Walk - mode of transportation to work % Bicycle - mode of transportation to work
Behavioural	Smoking can have a direct effect on the respiratory health of individuals (Kind et al 1998; Mullahy and Portney 1990; Jaakkola et al 1991; Cazzoletti et al 2009; Lwebuga-Mukasa et al 2002). Overweight and obese status have been significantly associated with asthma and poor health status (Mokdad et al 2003; Shaheen et al 1999; Stenius-Aarniala et al 2000; Chen et al 2002; Guerra et al 2002).	% Smoker % Overweight individuals % Obese individuals % Overweight or obese individuals
Exposure	Many studies have measured a positive and statistically significant association between NO_2 and respiratory health (Brook et al. 2004; Jerrett et al. 2004; Gulliver and Briggs 2005; Zanobetti and Schwartz 2006; Holguin et al. 2007; Gauderman et al. 2007; Brook et al. 2007; Brook et al 2008; Jerrett et al. 2009; Neupane et al. 2010).	Mean nitrogen dioxide concentration

Table 6.1 Justifications for the variables included in the model building exercise.

Variable Name	Abbreviation	Mean	Std Dev	Minimum	Maximum	Variance
% Occupied private dwellings in need of major repair	PerMajRep	6.0	3.2	0.8	13.5	10.2
% Occupied private dwellings built before 1946	PerBef1946	7.9	13.5	0.0	72.9	182.2
Unemployment rate of the total population 15 years and over	UnempRate	6.0	1.6	2.9	9.9	2.5
Educational attainment - no certificate, diploma or degree	PerNoDeg	15.7	4.8	6.5	31.5	22.6
Median income (2005)	MedInc	33,279.9	5,719.7	20,229.0	45,670.0	32,714,503.1
Average income (2005)	AvgInc	43,622.0	9,587.6	22,736.8	81,390.4	91,921,690.9
% low income	PrevLowInc	14.1	9.8	1.5	39.9	96.0
% Lone-parent families	PerLonePar	16.5	6.5	6.3	33.8	41.9
% Owned dwellings	PerOwned	69.1	23.1	20.4	98.3	532.6
% Rented dwellings	PerRent	30.8	23.1	1.5	79.8	534.6
% Occupied private dwellings of type single house	PerSingHou	46.1	25.9	2.0	96.2	670.4
% Occupied private dwellings of type row house	PerRowHous	20.2	16.1	0.0	63.4	259.4
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	PerApart5m	16.8	19.0	0.0	81.3	362.3
% Did not lived at the same address one year ago	PerMove1ye	14.0	5.6	4.5	30.5	30.9
% Did not lived at the same address five years ago	PerMove5ye	42.0	10.5	19.8	69.3	111.2
% Walk - mode of transportation to work	PerWalk	7.1	8.9	0.8	52.0	78.4
% Bicycle - mode of transportation to work	PerBike	2.2	2.0	0.0	9.8	4.2
Average value of owned dwelling	AvgValDwel	296,288.2	66,634.2	197,876.4	590,363.5	4,440,112,133.0
Participation rate of the total population 15 years and over	ParticipRa	68.3	6.6	47.7	80.9	43.0
% Management occupations	PerManagem	11.5	3.1	5.0	18.5	9.8
% Retail trade industry	PerRetail	10.2	2.6	3.8	17.4	6.9
Mean nitrogen dioxide concentration	MEAN NO2	5.3	1.8	1.1	10.6	3.2
% Smoker	PercSmoker	18.7	7.9	0.0	42.9	62.4
% Overweight individuals	PercOverwe	27.8	6.0	16.2	47.4	36.2
% Obese individuals	PercObese	11.7	5.5	2.2	31.3	30.2
% Overweight or obese individuals	PercOverWE	39.4	7.8	23.8	60.5	60.3

Table 6.2 Summary statistics of the explanatory variables. The unit of the variable mean nitrogen dioxide concentration is ppb while the variables median income, average income and average value of owned dwelling are in Canadian dollars.

Variable Definition	Global Moran'S I
% Occupied private dwellings in need of major repair	-0.0037 (0.930)
% Occupied private dwellings built before 1946	-0.0019 (0.980)
Unemployment rate of the total population 15 years and over	0.4668 (0.001)
Educational attainment - no certificate, diploma or degree	0.2386 (0.002)
Median income (2005)	0.4089 (0.001)
Average income (2005)	0.0645 (0.139)
% low income	0.5221 (0.001)
% Lone-parent families	0.3447 (0.001)
% Owned dwellings	0.5882 (0.001)
% Rented dwellings	0.5889 (0.001)
% Occupied private dwellings of type single house	0.5052 (0.001)
% Occupied private dwellings of type row house	0.2106 (0.002)
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	0.4478 (0.001)
% Did not live at the same address one year ago	0.4260 (0.001)
% Did not live at the same address five years ago	0.2573 (0.002)
% Walk - mode of transportation to work	0.8048 (0.001)
% Bicycle - mode of transportation to work	0.6494 (0.001)
Average value of owned dwelling	0.1395 (0.022)
Participation rate of the total population 15 years and over	0.5201 (0.001)
% Management occupations	0.1100 (0.043)
% Retail trade industry	0.2694 (0.001)
% Smoker	0.1730 (0.005)
% Overweight individuals	-0.0129 (0.516)
% Obese individuals	0.0199 (0.319)
% Overweight or obese individuals	0.0533 (0.180)
Mean nitrogen dioxide concentration	0.6104 (0.001)
Respiratory health outcome rate	0.0957 (0.071)

Table 6.3 Global Moran's I test for global spatial autocorrelation with pseudo-*p* values in parentheses.

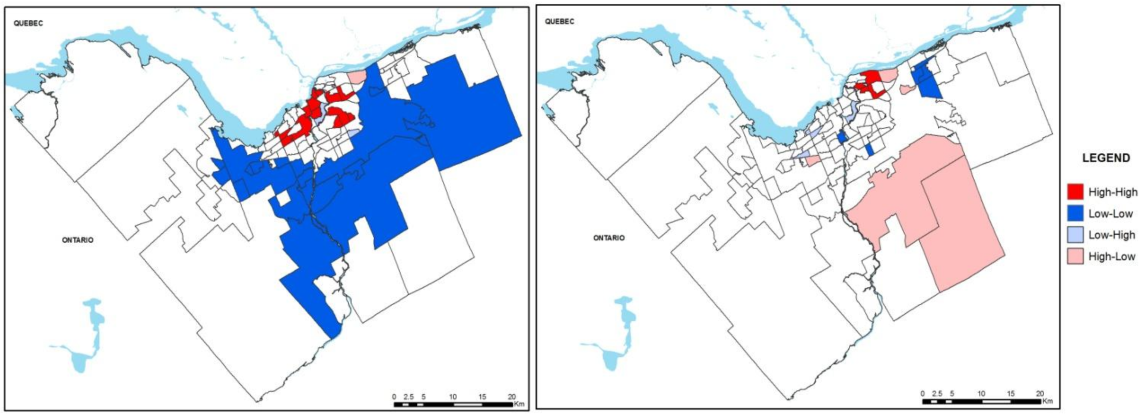


Figure 6.1 Local Moran's I core clusters. Left: NO_2 concentrations; Right: Respiratory outcomes.

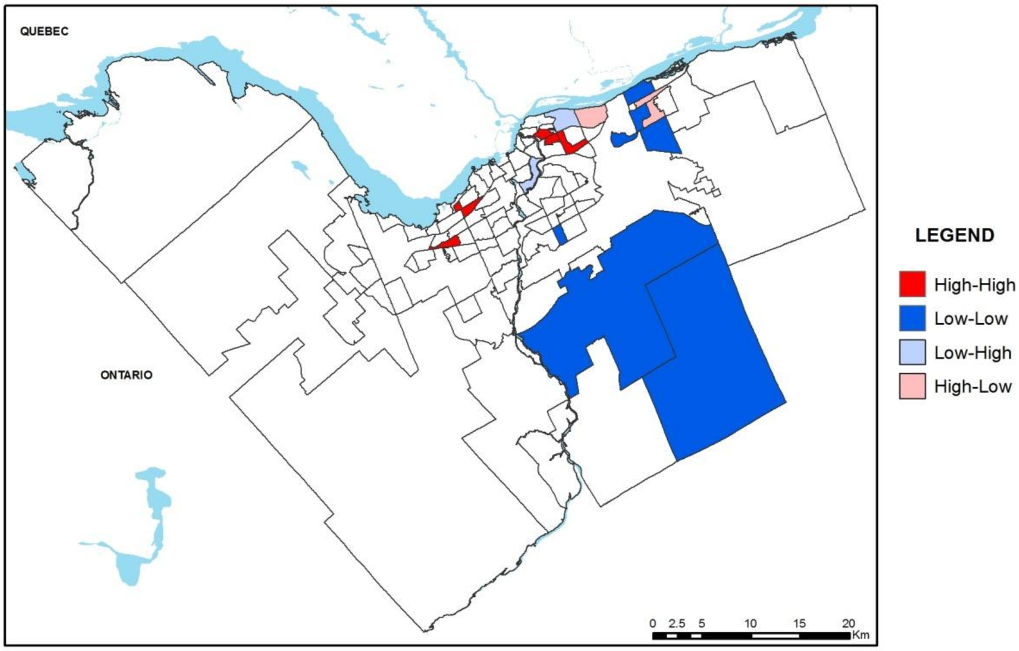


Figure 6.2 Bivariate LISA between the NO_2 concentrations and the respiratory health outcomes

	Natural Neighbourhood (n=88)	
Variable Definition	Mean NO2	Resp
% Occupied private dwellings in need of major repair	0.1912 (0.592)	-0.0108 (0.009)
% Occupied private dwellings built before 1946	0.1931 (0.599)	-0.0099 (0.999)
Unemployment rate of the total population 15 years and over	0.4331 (0.001)	0.2211(0.001)
Educational attainment - no certificate, diploma or degree	0.1655 (0.001)	0.0715 (0.001)
Median income (2005)	-0.4104 (0.001)	0.1060 (0.001)
Average income (2005)	-0.1476 (0.001)	-0.0215 (0.524)
% low income	0.5828 (0.001)	0.1982 (0.001)
% Lone-parent families	0.4014 (0.001)	0.1370 (0.001)
% Owned dwellings	-0.5950 (0.001)	-0.2423 (0.001)
% Rented dwellings	0.5862 (0.001)	0.2424 (0.001)
% Occupied private dwellings of type single house	-0.4539 (0.001)	-0.2180 (0.001)
% Occupied private dwellings of type row house	-0.2412 (0.001)	-0.1122 (0.001)
% Occupied private dwellings of type apartment (building that has 5 or more storeys)	0.5069 (0.001)	0.2118 (0.001)
% Did not live at the same address one year ago	0.4709 (0.001)	0.2100 (0.001)
% Did not live at the same address five years ago	0.2681 (0.001)	0.1157 (0.001)
% Walk - mode of transportation to work	0.5157 (0.001)	0.1846 (0.001)
% Bicycle - mode of transportation to work	0.5508 (0.001)	0.2254 (0.001)
Average value of owned dwelling	0.1165 (0.001)	0.0533 (0.001)
Participation rate of the total population 15 years and over	-0.3308 (0.001)	-0.2047 (0.001)
% Management occupations	-0.1575 (0.001)	-0.0644 (0.001)
% Retail trade industry	-0.2467 (0.001)	-0.0899 (0.001)
% Smoker	0.2136 (0.001)	0.1532 (0.001)
% Overweight individuals	-0.0672 (0.001)	-0.0025 (0.999)
% Obese individuals	-0.0927 (0.001)	0.0088 (0.990)
% Overweight or obese individuals	-0.1176 (0.001)	0.0042 (0.990)
Mean nitrogen dioxide concentration	1 ()	0.2458 (0.001)

Table 6.4 Global bivariate Moran's I between the explanatory variables and both the mean nitrogen dioxide concentration (Mean NO_2) and the respiratory health outcome rate (Resp).

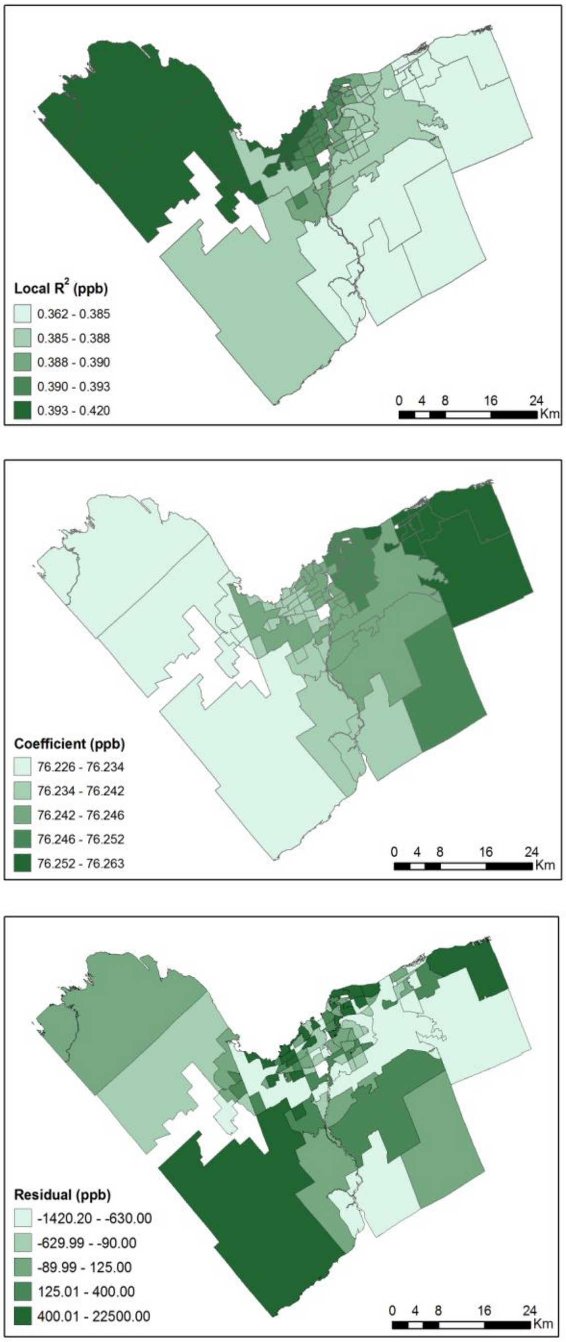


Figure 6.3 Geographically Weighted Regression; Top – Spatial distribution of the local R^2; Middle – Spatial distribution of the coefficient for the variable NO_2 mean; Bottom – Spatial distribution of the residuals.

	Natural Neighbourhood (n=88*)	
	Bivariate RESP (OLS)	Bivariate RESP (Spatial Lag)
R-squared	0.0600	0.0600
Adjusted R-squared	0.0500	n.a.
RMSE	713.4200	703.9300
Rho coefficient	n.a.	0.07 (0.61)
Intercept coefficient	819.20(<0.01)	768.96 (<0.01)
Mean nitrogen dioxide concentration	104.38 (0.02)	96.97 (0.02)
Jarque-Bera	12.23 (<0.01)	n.a.
Breusch-Pagan Test	4.01 (0.04)	3.66 (0.06)
Koenker-Bassett test	2.80 (0.09)	n.a.
Moran's I (error)	-0.04 (0.42)	n.a.
* Only neighbourhoods with CCHS sample > 9 included.		

Table 6.5 Bivariate ordinary least squares (OLS) and spatial regressions between NO_2 and the respiratory health outcomes.

	Natural Neighbourhood (n=88*)	
	Multivariate RES OLS	Multivariate RES Spatial Lag
R-squared	0.38	n.a.
Adjusted R-squared	0.35	n.a.
RMSE	591.22	728.8
Rho	n.a.	-1.21 (<0.01)
Intercept	-54.37 (0.84)	836.82 (0.02)
Educational attainment	58.82 (<0.01)	52.14 (<0.01)
% Rented dwellings	5.95 (0.11)	13.26 (<0.01)
% Smoker	18.91 (0.05)	28.74 (0.02)
concentration	-6.51 (0.889)	40.06 (0.47)
Jarque-Bera	10.36 (<0.01)	n.a.
Breusch-Pagan Test	11.13 (0.02)	22.86 (0.02)
Koenker-Bassett test	7.33 (0.12)	n.a.
Moran's I (error)	-0.02 (0.95)	n.a.

* Only neighbourhoods with CCHS sample > 9 included.

Table 6.6 Multivariate ordinary least squares (OLS) and spatial regressions between NO_2 and the respiratory health outcomes.

Parameter	Estimate	Standard Error	Pr > |Z|
Intercept	6.1551	0.2405	< 0.0001
Educational Attainment	0.0376	0.0118	0.0014
% Rented dwellings	0.0011	0.0035	0.7479
% Smoker	0.0223	0.0077	0.0037
Mean nitrogen dioxide concentration	-0.0155	0.0467	0.7403

Table 6.7 Analysis of GEE parameter estimates in the assessment of the relationship between exposure to NO_2 and respiratory health.

Chapter 7 Summary of Conclusions

7.1 Introduction

The role of spatial representation was studied in the context of the assessment of the spatial relationship between exposure to NO_2 and respiratory health. The research used a three-tier conceptual framework, based on the work of Jerrett and Finkelstein (2005), to gain a better understanding of how spatial representation can impact analytical results. This is of importance as results of research on the subject are not in agreement when it comes to the role of NO_2 exposure on health. Using innovative spatial analytical techniques, it was concluded that elements such as the geographic structure of analysis and the conceptualization and operationalization of variables used in modeling could be important factors explaining these inconsistencies in the literature. As such, this dissertation has shed light on some methodological elements that act as modifying factors and have not explicitly been addressed before when studying the relationship between exposure to air pollutants and health.

7.2 Summary of findings

In most cases, neighbourhood studies or small-area studies have used political or statistical units (often referred to as neighbourhoods) as the base geography to conduct their work. This research shows that GIS and spatial analytical techniques can be used to develop a better understanding of the importance of the context of place in place-based health research. Using a novel approach, natural neighbourhoods were delineated for the City of Ottawa. Wombling and spatially constrained clustering techniques were applied to socioeconomic data from the 2001 Canadian Census of population to delineate homogeneous units in terms of the geography of susceptibility. The original process resulted in the delineation of 89 units but with the availability of the 2006 Census data, additional consultations with partners were held and modifications were made. The final version of the boundary file contains 95 natural neighbourhoods. The methodology developed is easily reproducible and transferable to other population centers of different sizes in Canada but also to other countries where the required input data are available. Following the delineation of the natural neighbourhoods several local organizations have adopted these units and are collecting and compiling information based on this geography. For the purpose of this study, the natural neighbourhoods maximized the internal homogeneity and hence permitted a better identification of communities at risk.

The geography of exposure to NO_2 concentrations was modeled using geostatistical interpolation techniques and a land-use regression (LUR) model. Through the use of these approaches, we assessed model sensitivity against three spatial representations of the population. One of the population representations was based on the use of a polycategorical dasymetric mapping approach, a novel method that allows the redistribution of the population at a sub-block level. In the case of the LUR, we found no significant improvement using dasymetric mapping in comparison to the population representation at the dissemination block level but these models performed better than when using the population representation at the dissemination area level in the LUR model. The results obtained indicate that the way variables are operationalized in terms of the scale of representation in the LUR will affect the results. We used population representations as an example but other variables could in the same manner affect modeled results. The best performing model, based on dasymetric mapping, yielded an R^2 of 0.8055, and showed that LUR can be used in population centers the size of Ottawa. The LUR model developed for Ottawa was the first model of this type elaborated for this city.

As one of the most common spatial approaches to the modeling of NO_2, geostatistical interpolation techniques were used to generate NO_2 surfaces. Four geostatistical interpolation models were compared with the NO_2 surface generated through the LUR model. The surface generated through the LUR is not statistically different than the four other surfaces but has a low spatial correlation to all of them. We also observed that the LUR model is the only surface that provides detail at the street level. Hence, we were able to confirm that the spatial approach used to model exposure could affect the measured relationship between exposure to NO_2 and health, as errors could be introduced through a less precise mapping of the concentration of the pollutant through space.

The importance of the geography of susceptibility on the geography of risk was then studied. The relationship was examined using three different types of spatial structures. The NO_2 exposure data and the health outcome data were compiled according to census tract structure boundaries (which in the literature is commonly used as a way to operationalize the neighbourhood), the natural neighbourhoods delineated using our novel approach, and a cluster structure obtained through automated zone design. This component of the research highlighted three key findings, two of which are related to the effect of the MAUP on the study of the relationship between exposure to NO_2 and respiratory health and the third one related to the role of NO_2 on health. Our results indicate that exploratory analytical methods, both spatial and traditional, can serve as an indication of the potential effect of the MAUP in the study of this relationship. The MAUP has not been directly addressed in most studies on the relationship between exposure to NO_2 and health, and according to our results its effects should no longer be neglected. Mayer (1983) stated more than 20 years ago that scale was not taken into account in health studies. Having integrated this idea in our research we can confirm that it is a crucial element that needs to become an integral consideration in spatially-centric health research. The observed effect of the MAUP on Moran's I can serve as a rationale for the use of a geography that is delineated based on the processes associated with the dependent variable. Our results also confirm that the use of different spatial structures is a factor contributing to the lack of agreement found in the literature as the effect of the MAUP on multivariate analyses is complex and unpredictable. Lastly, this component of the research confirmed that NO_2 is not a key factor in the health of the residents of Ottawa, but depending on the operationalization of the spatial units, the processes explaining the differences in the morbidity rates will not be the same. Instead the processes will share a link to local socioeconomic status.

Having assessed the impact of using different spatial structures to study the relationship between exposure to NO_2 and respiratory health, we now focus on the natural neighbourhood structure. A spatial approach needed to be implemented considering that the use of traditional statistical methods can generate erroneous results when the data are characterized by spatial autocorrelation and non-stationarity. The results indicate that there is a weak association between exposure to NO_2 concentrations and respiratory health. According to our results socioeconomic status is the dominant variable contributing to the respiratory health of the population. Our results were not sensitive to the methodological approach implemented; the use of a Poisson regression, which is commonly applied for this type of study, concurred with the results obtained with the spatial approach. We conclude that the lack of agreement in the studies published on the relationship between exposure to an air pollutant and health are less the result of the use of different statistical methods to assess the relationship than the effect of spatial representation.

The results obtained in this research are contradictory to some obtained in similar studies where a statistically significant relationship between exposure to NO_2 and health outcomes have been measured. On the other hand, other studies conducted in the U.S. and in Europe have published similar findings. The fact that we did not find NO_2 concentrations to be a main factor explaining differences in the respiratory health of residents could potentially be the result of the generally good air quality found in Ottawa in comparison to the other Canadian cities where similar studies took place. The fact that our results do not identify air quality as a main contributing factor in the health status of the Ottawa population do not in any way lessen the importance of the results obtained. This research demonstrates the necessity of addressing scale more strategically in health research as it can have an effect on the representation of exposure, susceptibility and consequently on the assessment of the relationship between exposure and health. In this sense, this research meets its objectives and makes an important contribution to knowledge.

7.3 Future research

The geography of exposure is a domain of research where the availability of additional data would most likely strengthen the analytical results obtained. If the National Capital Air Quality Mapping project was to be extended, which may be the case, it would be important to collect data through multiple field deployments. Even if the patterns of distribution of NO_2 concentration are the same from one season to the other (Wheeler et al. 2008), and as a consequence it is possible to use data from a unique two week deployment, it is recommended to compile data from multiple field deployments since the effect of the period of deployment could not be tested in this study due to limited data. In a future study the spatial distribution of the passive samplers should be done through a location-allocation model (Kanaroglou et al. 2005) and cover both the rural and the urban environment of the city of Ottawa.

This study focused on the relationship between exposure to NO_2 and respiratory morbidity rate for the population of age 15 and over without distinction for gender and age groups. Since the effect of NO_2 exposure on health has been reported to vary as a function of gender and age (Lin et al. 2004; Annesi-Maesano et al. 2003; Clougherty et al. 2010), a study of this relationship by subgroup of the population may provide different results. Such an approach may reveal that NO_2 exposure is an important factor contributing to the health of individuals for some segments of the population. For example, studies focused on children have in many cases reported statistically significant relationships (Nicolari et al. 2003; Gauderman et al. 2005; McConnell et al. 1999; Brauer et al. 2007; Mortimer et al. 2002; McConnell et al. 2003).

The association between respiratory health and air quality has been studied for pollutants other than NO_2. Since SO_2 and O_3 data were collected using the Ogawa samplers and the permanent stations, it could be of value to reproduce some of the methods developed in this dissertation for these other pollutants (Strickland et al. 2010; Clark et al. 2010; Stieb et al. 2009; Peel et al. 2005; Brunekreef and Holgate 2002; Fusco et al. 2001). Measures of association between NO_2, SO_2 and O_3 could also be conducted for a better understanding of the interaction between different air pollutants.

Finally an effort should be made to include daily mobility into the estimate of exposure to NO_2. One approach would be the use of personal GPS systems

(Rainham et al. 2010) to develop a more accurate measure of exposure. As costly as it can be to conduct such studies, this information could correct errors induced by the fact that individuals may actually spend more time away from home. The latter information would allow the implementation of a more advanced type of exposure measurement based on the synergy between air pollution modeling using a spatial approach and mobility data. Another approach, less costly then the aforementioned approach, could rely on data tabulated according to the usual place of work instead of the location of the residence. A study conducted in Sweden developed a similar methodology (Lindgren et al. 2010), which allowed for the assessment of the relationship between exposure to NOx and the prevalence of asthma and asthma symptoms. The authors found no statistically significant relationship between traffic exposure at workplace address and health or between total exposure (sum of the exposure at home, work, and while commuting) and health. A similar methodology could be developed using Census data in order to confirm the findings of Lindgren et al (2010). Such a research should be preceded by a thorough review of data quality issues associated with the collection of the data on the usual place of work.

Bibliography

Abada, T., Hou, F., & Ram, B. (2007). Racially mixed neighborhoods, perceived neighborhood social cohesion, and adolescent health in canada. *Social Science and Medicine, 65*(10), 2004-2017.

Adler, N.E., Boyce, T., Chesney, M.A., Cohen, S., Folkman, S., Kahn, R.L., & Syme, S.L. (1994). Socioeconomic status and health; the challenge of the gradient. *American Psychological Association*, *49*(1), 15-24.

Adler, N.E., & Newman, K. (2002). Socioeconomic disparities in health: pathways and policies. *Health Affairs*, *21*(2), 60-76.

Aitken, S.C., & Prosser, R. (1990). Residents' spatial knowledge of neighborhood continuity and form. *Geographical Analysis*, *22*(4), 301-326.

Akinbami, L.J., Lynch, C.D., Parker, J.D., & Woodruff, T.J. (2010). The association between childhood asthma prevalence and monitored air pollutants in metropolitan areas, united states, 2001-2004. *Environmental Research*, *110*(3), 294-301.

Allen, R. W., Amram, O., Wheeler, A. J., & Brauer, M. (2010). The transferability of NO and NO_2 land use regression models between cities and pollutants. *Atmospheric Environment. In press.*

Amrhein, C.G., & Reynolds, H. (1997). Using the getis statistic to explore aggregation effects in metropolitan toronto census data. *The Canadian Geographer*, *41*(2), 137-149.

Andresen, M.A., & Brantingham, P.L. (2008). Visualizing ambient population data within census boundaries: a dasymetric mapping procedure. *Cartographica*, *43*, 267-275.

Annesi-Maesano, I., Agabiti, N., Pistelli, R., Couilliot, M. -., & Forastiere, F. (2003). Subpopulations at increased risk of adverse health outcomes from air pollution. *European Respiratory Journal, Supplement, 21*(40), 57S-63S.

Anselin, L. (1995). Local indicators of spatial association-lisa. *Geographical Analysis*, *27*, 93-115.

Anselin, L. (1998). Exploratory spatial data analysis in a geocomputational. In P. l Longley, S. Brooks, B. Macmillan, & R. McDonnell (Eds.), *GeoComputation, a primer* (pp. 77-94). New York, NY: Wiley.

Anselin, L. (2002). Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, *17*, 247-267.

Anselin, L. (2005a). *Exploring spatial data with geoda tm: a workbook*. Retrieved from http://www.csiss.org/

Anselin, L. (2005b). Spatial statistical modeling in a gis environment. In D. Maguire, Batty, M., & M.F. Goodchild (Eds.) *Gis, spatial analysis, and modeling* (pp. 93-111). Redlands, California: ESRI Press.

Anselin, L., & Getis, A. (1992). Spatial statistical and geographic information systems. *Ann Reg Sci*, *26*, 19-33.

Anselin, L., Sridharan, S., & Gholston, A. (2007). Using exploratory spatial data analysis to leverage social indicator databases: the discovery of interesting patterns. *Social Indicators Research*, *82*, 287-309.

Anselin, L., Syabri, I., & Kho, Y. (2006). Geoda: an introduction to spatial data analysis. *Geographical Analysis*, *38*(1), 5-22.

Arain, M.A., Blair, R., Finkelstein, N., Brook, J., & Jerrett, M. (2009). Meteorological influences on the spatial and temporal variability of no_2 in toronto and hamilton. *The Canadian Geographer*, *53*(2), 165-190.

Arbex, M.A., De Souza Conceição, G.M., Cendon, S.P., Arbex, F.F., Lopes, A.C., Moysés, E.P., & et al. (2009). Urban air pollution and chronic obstructive pulmonary disease-related emergency department visits. *Journal of Epidemiology and Community Health*, *63*(10), 777-783.

Arbia, G. (1989). *Spatial data configuration in statistical analysis of regional economic and related problems*. Netherlands: Kluwer Academic Publishers.

ArcGIS (version 9.2) [Computer software]. Redlands, CA: ESRI

Atari, D.O., & Luginaah, I.N. (2009). Assessing the distribution of volatile organic compounds using land use regression in sarnia, "chemical valley", ontario, canada. *Environmental Health*, *8*(1).

Bailey, T.C., & Gatrell, A.C. (1995). *Interactive spatial analysis*. New York, NY: John Wiley and Sons.

Bang, K. M., Syamlal, G., & Mazurek, J. M. (2009). Prevalence of chronic obstructive pulmonary disease in the U.S. working population: An analysis of data from the 19972004 national health interview survey. *COPD: Journal of Chronic Obstructive Pulmonary Disease, 6*(5), 380-387.

Barcelo, M.A., Sacz, M., & Saurina, C. (2009). Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the barcelona metropolitan region, spain. *Science of the Total Environment*, *407*, 5501-5523.

Bashir, S.A. (2002). Home is where the harm is: inadequate housing as a public health crisis. *American Journal of Public Health*, *92*(5), 733-738.

Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., & Briggs, D.J. (2009). Mapping of background air pollution at a fine spatial scale across the european union. *Science of the Total Environment*, *407*, 1852-1867.

Beelen, R., Hoek, G., van den Brandt, P.A., Goldbohm, R.A., Fisher, P., Schouten, L.J., Jerrett, M., Hughes, E., Armstrong, B., & Brunekreef, B. (2008). Long-term effects of traffic-related air pollution on mortality in a dutch cohort (nlcs-air study). *Environmental Health Perspectives*, *116*(2), 196-202.

Bell, J., & Burlin, B.M. (1993). In urban areas: many of the poor still pay more for food. *Journal of Public Policy and Marketing*, *12*(2), 268-270.

Benzeval, M. (1998). The self-reported health status of lone parents. *Social Science and Medicine, 46*(10), 1337-1353.

Berkman, L.F., & Kawachi, I. (2000). A historical framework for social epidemiology. In L.F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 3-12). New York, NY: Oxford University Press.

Bertazzon, S., Olson, S., & Knudtson, M. (2010). A spatial analysis of the demographic and socio-economic variables associated with cardiovascular disease in calgary (canada). *Appl. Spatial Analysis*, *3*, 1-23.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M.L. (2007). Landscan u.s.a.: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal*, *69*(1-2), 103-117.

Blanc, P. D., & Toren, K. (1999). How much adult asthma can be attributed to occupational factors? *American Journal of Medicine, 107*(6), 580-587.

Bohadana, A. B., Hannhart, B., Ghezzo, H., Teculescu, D., & Zmirou-Navier, D. (2011). Exhaled nitric oxide and spirometry in respiratory health surveillance. *Occupational Medicine, 61*(2), 108-114.

Bonham-Carter, G.F. (1994). *Geographic information systems for geoscientists: modelling with gis*. Oxford; New York, NY: Pergamon.

BoundarySeer [Computer Software]. Ann Harbor, MI: Terraseer.

Brainard, J., Jones, A.P., Bateman, I.J., Lovett, A.A., & Fallon, P.J. (2001). Modelling the relationship between socio-economic indicators and air pollution.

Brauer, M., Hoek, G., Smit, H. A., de Jongste, J. C., Gerritsen, J., Postma, D. S., et al. (2007). Air pollution and development of asthma, allergy and infections in a birth cohort. *European Respiratory Journal, 29*(5), 879-888.

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fisher, P.H., Wijga, A., Koopman, L.P., Neijens, H.J., Gerritsen, J., Kerkhof, M., Heinrich, J., Bellander, T., & Brunekreef, B. (2002). Air pollution from traffic and the development of respiratory infections and asthmatic and allergies symptoms in children. *Am J Respir Cirt Care Med*, *166*, 1092-1098.

Braveman, P. (2006). Health disparities and health equity: concepts and measurement. *Annual Review of Public Health*, *27*, 167-194.

Braveman, P.A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K.S., Metzler, M., & Posner, S. (2005). Socioeconomic status in health research; one size does not fit all. *Journal of the American Medical Association*, *294*(22), 2879-2888.

Briant, A., Combes, P.-P., & Lafourcade, M. (2010). Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*, 287-302.

Briggs, D. (2005). The role of gis: coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health, Part A*, *68*, 1243-1261.

Briggs, D.J., Collins, S., Elliott, P., Fisher, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., & Van Der Veen, A. (1997). Mapping urban air pollution using gis: a regression-based approach. *International Journal of Geographical Information Science*, *11*(7), 699-718.

Briggs, D.J., de Hoogh, C., Gulliver, J., Willis, J., Elliott, P., Kingham, S., & Smallbone, K. (2000). A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, *253*, 151-167.

Brindley, P., Wise, S.M., Maheswaran, R., & Haining, R.P. (2005). The effect of alternative representations of population location on the areal interpolation of air pollution exposure. *Computers, Environment and Urban Systems*, *29*, 455-469.

Brook, J.R., Burnett, R.T., Dann, T.F., Cakmak, S., Goldberg, M.S., Fan, X., & Wheeler, A.J. (2007). Further interpretation of the acute effect of nitrogen dioxide observed in canadian time-series studies. *Journal of Exposure Science and Environmental Epidemiology*, *17*, S36-S44.

Brook, R.D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., Luepker, R., Mittleman, M., Samet, J., Smith, S.C., & Tager, I. (2004). Air pollution and cardiovascular disease; a statement for healthcare professionals from the expert panel on population and prevention science of the american heart association. *Circulation*, *109*(21), 2655-2671.

Brook, R.D., Jerrett, M., Brook, J.R., Bard, R.L., & Finkelstein, M.M. (2008). The relationship between diabeters mellitus and traffic-related air pollution. *Journal of Occupational & Environmental Medicine*, *50*(1), 32-38.

Brunekreef, B. (2007). Health effects of air pollution observed in cohort studies in europe. *Journal of Exposure Science and Environmental Epidemiology*, *17*, S61-S65.

Brunekreef, B., & Holgate, S.T. (2002). Air pollution and health. *Lancet*, *360*(9341), 1233-1242.

Brunekreef, B., Janssen, N. A. H., De Hartog, J., Harssema, H., Knape, M., & Van Vliet, P. (1997). Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology, 8*(3), 298-303.

Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281-298.

Burra, T.A., Moineddin, R., Agha, M.A., & Glazier, R.H. (2009). Social disadvantage, air pollution, and asthma physician visits in toronto, canada. *Environmental Research*, *109*, 567-574.

Buzzelli, M., Jerrett, M., Burnett, R., & Finklestein, N. (2003). Spatiotemporal perspectives on air pollution and environmental justice in hamilton, canada, 1985-1996. *Annals of the Association of American Geographers, 93*(3), 557-573.

Buzzelli, M., & Jerrett, M. (2007). Geographies of susceptibility and exposure in the city: environmental inequity of traffic-related air pollution in toronto. *Canadian Journal of Regional Science*, *2*, 195-210.

California Environmental Protection Agency. (2010). *Glossary of air pollution terms*. Retrieved from http://www.arb.ca.gov/html/gloss.htm#C

Cardenas, G., & Perdrix, E. (2008). Automatic mapping algorithm of nitrogen dioxide levels from monitoring air pollution data using classical geostatistical approach: application to the french lille city. *Quantitative Geology and Geostatistics*, *15*(1), 95-1010.

Carr, W., Zeitel, L., & Weiss, K. (1992). Variations in asthma hospitalizations and deaths in new york city. *American Journal of Public Health, 82*(1), 59-65.

Cassell, D.L. (2007). Don't be loopy: re-sampling and simulation the sas way. *Proceedings of the SAS Global Forum 2007*, http://www2.sas.com/proceedings/forum2007/183-2007.pdf

Cazzoletti, L., Marcon, A., Corsico, A., Janson, C., Jarvis, D., Pin, I., et al. (2009). Asthma severity according to global initiative for asthma and its determinants: An international study. *International Archives of Allergy and Immunology, 151*(1), 70-79.

Chaix, B. (2009). Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. *Annu. Rev. Public Health*, *30*, 81-105.

Chaix, B., Leal, C., & Evans, D. (2010). Neighborhood-level confounding in epidemiologic studies; unavoidable challenges, uncertain solutions. *Epidemiology*, *21*(1), 124-127.

Chaix, B., Lindstrom, M., Rosvall, M., & Merlo, J. (2008). Neighbourhood social interactions and risk of acute myocardial infarction. *J Epidemiol Community Health*, *62*, 62-68.

Chaix, B., Merlo, J., Evans, D., Leal, C., & Havard, S. (2009). Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. a response to riva, gauvin, apparicio and brodeur. *Social Science & Medicine*, *69*, 1306-1310.

Chang, H. (2008). Spatial analysis of water quality trends in the han river basin, south korea. *Water Research*, *42*, 3285-3304.

Chen, Y., Dales, R., Tang, M., & Krewski, D. (2002). Obesity may increase the incidence of asthma in women but not in men: Longitudinal observations from the canadian national population health surveys. *American Journal of Epidemiology, 155*(3), 191-197.

Chen, E., Hanson, M.D., Paterson, L.Q., Griffin, M.J., Walker, H.A., & Miller, G.E. (2006). Socioeconomic status and inflammatory processes in childhood asthma: the role of psychological stress. *J Allergy Clin Immunol*, *117*(5), 1014-1020.

Chen, E., Miller, G.E., Walker, H.A., Arevalo, J.M., Sung, C.Y., Cole, S.W. (2009). Genome-wide transcriptional profiling linked to social class in asthma. *Thorax*, *64*, 38-43.

Chertok, M., Voukelatos, A., Sheppeard, V, & Rissel, C. (2004). Comparison of air pollution exposure for five commuting. *Health Promotion Journal of Australia* , *15*(1), 63-67.

CIHI - Canadian Institute for Health Information. (2006). International statistical classification of diseases and related health problems, tenth revision, canada. *CIHI (2006) ICD-10-CA/CCI*.

CIHI - Canadian Institute for Health Information. *(2008) Summary report, reducing gaps in health; a focus on socio-economic status in urban canada*. (2008). Retrieved from http://www.cihi.ca/cihiweb/en/downloads/Reducing_Gaps_in_Health_Summary_Report_EN_080816.pdf

CIHI - Canadian Institute for Health Information. *(2010a) Discharge Abstract Database.* Retrieved from http://www.cihi.ca/CIHI-ext-portal/internet/en/document/types+of+care/hospital+care/acute+care/services_dad

CIHI - Canadian Institute for Health Information. *(2010b) National Ambulatory Care Reporting System (NACRS).* Retrieved from http://www.cihi.ca/CIHI-ext-portal/internet/en/document/types+of+care/hospital+care/emergency+care/services_nacrs

City of Ottawa. (2004). Air quality and climate change management plan. *Air and Energy Initiatives, Environmental Management Division, Planning and Growth Management Department*, 107 p.

City of Ottawa. (2006). City of ottawa vector gis data. [*Computer file*]. *Available from City of Ottawa GIS services*.

City of Ottawa. (2006a). *Building blocks for a multi-layered city: neighborhoods, precincts, corridors, waterways*. Retrieved from http://www.ottawa.ca/city_services/planning/resources/publications/neighborhood_Design_Options_Study.pdf

City of Ottawa. (2006b). *City of ottawa parks and recreation*. Retrieved from http://www.ottawa.ca/residents/parks_recreation/parks/index_en.html

City of Ottawa. (2010a). *National capital air quality mapping pilot project; frequently asked questions*. Retrieved from http://www.ottawa.ca/residents/environment/community/airquality/faq_en.html

City of Ottawa. (2010b). *Ottawa 2020 environmental strategy-annex 2*. Retrieved from http://www.city.ottawa.on.ca/city_services/planningzoning/2020/enviro/annex2_en.shtml

Clapp, J.M., & Wang, Y. (2006). Defining neighborhood boundaries: are census tracts obsolete? *Journal of Urban Economics*, *59*, 259-284.

Clark, N.A., Demers, P.A., Karr, C.J., Koehoorn, M., Lencar, C., Tamburic, L., & Brauer, M. (2010). Effect of early life exposure to air pollution on development of childhood asthma. *Environmental Health Perspectives*, *118*(2), 284-290.

Cliff, A.D., & Ord, J.K. (1981). *Spatial processes : models & applications* . London: Pion.

Clougherty, J. E. (2010). A growing role for gender analysis in air pollution epidemiology. *Environmental Health Perspectives, 118*(2), 167-176.

Clougherty, J.E., Houseman, E.A., & Levy, J.I. (2009). Examining intra-urban variation in fine particle mass constituents using gis and constrained factor analysis. *Atmospheric Environment*, *43*, 5545-5555.

Clougherty, J.E., & Kubzansky, L.D. (2009). A framework for examining social stress and susceptibility to air pollution in respiratory health. *Environmental Health Perspectives*, *117*(9), 1351-1358.

Clougherty, J.E., Levy, J.I., Kubzansky, L.D., Ryan, P.B., Suglia, S.F. , Jacobson Canner, M., & Wright, R.J. (2007). Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. *Environmental health perspectives*, *115*(8), 1140-1146.

Cockings, S., & Martin, D. (2005). Zone design for environment and health studies using pre-aggregated data. *Social Science & Medicine*, *60*, 2729-2742.

Cooper, D.W. (1968). The significance level in multiple tests made simultaneously. *Heredity*, *23*, 614-617.

Crampton, J.W. (2004). Gis and geographic governance: reconstructing the choropleth map. *Cartographica*, *39*(1), 42-53.

Crouse, D.L., Goldberg, M.S., & Ross, N.A. (2009). A prediction-based approach to modeling temporal and spatial variability of traffic-related air pollution in montreal, canada. *Atmospheric Environment*, *43*, 5075-5084.

Crouse, D. L., Goldberg, M. S., Ross, N. A., Chen, H., & Labrèche, F. (2010). Postmenopausal breast cancer is associated with exposure to traffic-related air pollution in montreal, canada: A case-control study. *Environmental Health Perspectives, 118*(11), 1578-1583.

Crouse, D.L., Ross, N.A., & Goldberg, M.S. (2009). Double burden of deprivation and high concentrations of ambient air pollution at the neighbourhood scale in montreal, canada. *Social Science & Medicine*, *69*, 971-981.

Cummins, S., Macintyre, S., Davidson, S., & Ellaway, A. (2005). Measuring neighbourhood social and material context: generation and interpretation of ecological data from routine and non-routine sources. *Health & Place*, *11*, 249-260.

Daras, K., & Alvanides, S. (2006). Zone design in public health policy. In M. Campagna (Ed.), *GIS for sustainable development* (pp. 247-265). United Kingdom: Taylor & Francis.

de Fouquet, C., Gallois, D., & Perron, G. (2007). Geostatistical characterization of the nitrogen dioxide concentration in an urban area: part 1: spatial variability and cartography of the annual concentration. *Atmospheric Environment*, *41*(32), 6701-6714.

Dietz, R.D. (2002). The estimation of neighborhood effects in the social sciences: an interdisciplinary approach. *Social Science Research*, *31*, 539-575.

Diez-Roux, A. (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health*, *91*(11), 1783-1789.

Diez-Roux, A. (2002). Invited: commentary: places, people and health. *American Journal of Epidemiology*, *155*(6), 516-519.

Diez-Roux, A.V. (2002). A glossary for multilevel analysis. *Journal of Epidemiology and Community Health 2002*, *56*, 588-594.

Diniz-Filho, J.A.F., Nabout, J.C., Telles, M.P., Soares, T.N., & Rangel, T.F.L.V.B. (2009). A review of techniques for spatial modeling in geographical, conservation and landscape genetics. *Genetics and molecular Biology*, *32*(2), 203-211.

Dirks, K.N., Longley, I., Salmond, J., Kingham, S., Fernandes, K. 2009. Blood Carboxyhemoglobin levels as a biomarker for urban air pollution exposure. Paper presented at the International Conference on Southern Hemisphere Meteorology and Oceanography, Melbourne, Melbourne, 9-13 February.

DMTI Spatial. (2007a). Enhanced points of interest (epoi) V2007 3. [Data File]. Available from DMTI™ Spatial.

DMTI Spatial. (2007b). Canmap parks & recreation V2007 3. [Data File]. Available from DMTI™ Spatial.

DMTI Spatial. (2007c). Platinum postal suite V2007 3. [Data File]. Available from DMTI™ Spatial.

DMTI Spatial. (2007d) Canmap route logistics V2007 3. [Data File]. Available from DMTI™ Spatial.

Dominici, F., Sheppard, L., & Clyde, M. (2003). Health effects of air pollution: a statistical review. *International Statistical Review*, *71*(2), 243-276.

Douglas, I. (2005). Urban greenspace and mental health. *Proceedings of the UK MAB Urban Forum*, http://www.ukmaburbanforum.co.uk/docunents/papers/ukmabgrnspcepap2.pdf

Dragano, N., Hoffmann, B., Stang, A., Moebus, S., Verde, P.E., Weyers, S., Mohlenkamp, S., Schmermund, A., Mann, K., Jockel, K.H., Erbel, R., Siegrist, J., & Heinz Nixdorf Recall Study Investigative Group. (2009). Sublinical coronary atherosclerosis and neighbourhood deprivation in an urban design. *Eur J Epidemiol*, *24*, 25-35.

Dunn, J.R. (2002). Housing and inequalities in health: a study of socioeconomic dimensions of housing and self reported health from a survey of vancouver residents. *Journal of Epidemiology and Community Health*, *56*, 671-681.

Dutilleul, P. 1993. Modifying the t test for assessing the correlation between two spatial processes. Biometrics 49: 305-314.

Eicher, C.L., & Brewer, C.A. (2001). Dasymetric mapping and areal interpolation: implementation and evaluation. *Cartography and Geographic Information Systems*, *28*(2), 125-138.

Eisner, M. D., Blanc, P. D., Omachi, T. A., Yelin, E. H., Sidney, S., Katz, P. P., et al. (2011). Socioeconomic status, race and COPD health outcomes. *Journal of Epidemiology and Community Health, 65*(1), 26-34.

Ellaway, A., & Macintyre, S. (1998). Does housing tenure predict health in the UK because it exposes people to different levels of housing related hazards in the home or its surroundings? *Health and Place, 4*(2), 141-150.

Ellen, I.G., & Turner, M.A. (1997). Does neighborhood matter? assessing recent evidence. *Housing Policy Debate*, *8*(4), 833-866.

Elliott, M. (2000). The stress process in neighborhood context. *Health and Place, 6*(4), 287-299.

Environment Canada. (2009a). *Air pollution*. Retrieved from http://www.ec.gc.ca/default.asp?lang=En&n=499D6B13-1

Environment Canada. (2010). *Air quality*. Retrieved from http://www.ec.gc.ca/Air/default.asp?lang=En&n=04104DB7-1

Environment Canada. (2010a). *Criteria air contaminant (cac) emissions density maps*. Retrieved from http://www.ec.gc.ca/inrp-npri/default.asp?lang=En&n=1567DA21-1

Evans, G.W., & Kantrowitz, E. (2002). Socioeconomic status and health: the potential role of environmental risk exposure. *Annu. Rev. Public Health*, *23*, 303-331.

Evans, R.G., & Stoddart, G.L. (1994). Producing health, consuming health care. In R.G. Evans, M.L. Barer, & T.R. Marmor (Eds.), *Why are some people healthy and others not? The determinants of health of populations* (pp. 27-64). New York, NY: Aldine De Gruyter.

Fiedler, R., Schuurman, N., & Hyndman, J. (2006). Improving census-based socioeconomic gis for public policy: recent immigrants, spatially concentrated poverty and housing need in vancouver. *ACME*, *4*, 145-171.

Finkelstein, M.M., Jerrett, M., DeLuca, P., Finkelstein, N., Verma, D.K., Chapman, K., & Sears, M.R. (2003). Relation between income, air pollution and mortality: a cohort study. *CMAJ*, *169*(5), 397-402.

Finkelstein, M.M., Jerrett, M., & Sears, M.R. (2005). Environmental inequality and circulatory disease mortality gradients. *J. Epidemiol Community Health*, *59*(6), 481-487.

Fiscella, K., & Williams, D.R. (2004). Health disparities based on socioeconomic inequities: implications for urban health care. *Academic Medicine*, *79*(12), 1139-1147.

Fisher, P.F., & Langford, M. (1996). Modeling sensitivity to accuracy in classified imagery: a study of areal interpolation by dasymetric mapping. *The Professional Geographer*, *48*(3), 299-309.

Fong, E., & Gulia, M. (2000). Neighborhood change within the canadian ethnic mosaic, 1986-1991. *Population Research and Policy Review, 19*(2), 155-177.

Forastiere, F., Stafoggia, M., Tasco, C., Picciotto, S., Agabiti, N., Cesaroni, G., & et al. (2007). Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or differential susceptibility. *American Journal of Industrial Medicine*, *50*(3), 208-216.

Fortin, M.-J. (1994). Edge detection algorithms for two dimensional ecological data. *Ecology*, *75*, 956-965.

Fortin, M.-J. (1997). Effects of data types on vegetation boundary delineation. *Canadian Journal of Forest Research*, *27*, 1851-1858.

Fortin, M.-J., & Drapeau, P. (1995). Delineation of ecological boundaries: comparison of approaches and significance tests. *OIKOS*, *72*, 323-332.

Fotheringham, A.S. (1989). Scale-independent spatial analysis. In M. Goodchild & S. Gopal (Eds.), *Accuracy of spatial databases* (pp. 221-228). United Kingdom: Taylor & Francis.

Fotheringham, A.S., Brunsdon, C., & Charlton, M.E. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. United Kingdom: Wiley.

Fotheringham, A.S., Charlton, M.E., & Brunsdon, C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, *30*(11), 1905-1927.

Fotheringham, A.S., & Wong, D.W.S. (1991). The modifiable area unit problem in multivariate analysis. *Environment and Planning A*, *23*, 1025-1044.

Fusco, D., Forastiere, F., Michelozzi, P., Spadea, T., Ostro, B., Arcà, M., et al. (2001). Air pollution and hospital admissions for respiratory conditions in rome, italy. *European Respiratory Journal, 17*(6), 1143-1150.

Galea, S., & Vlahov, D. (2005). Urban health: evidence, challenges, and directions. *Annual Review of Public Health*, *26*, 341-365.

Gallo, L.C., & Matthews, K.A. (2003). Understanding the association between socioeconomic status and physical health: do negative emotions play a role? *Psychological Bulletin*, *129*(1), 10-51.

Gallois, D., de Fouquet, C., Le Loc'h, G., Malherbe, L., & Cardenas, G. (2005). Mapping annual nitrogen dioxide concentrations in urban areas. *Quantitative Geology and Geostatistics*, *14*(5), 1087-1096.

Galster, G.C. (2001). On the nature of neighbourhood. *Urban Studies*, *38*(12), 2111-2124.

Gauderman, W. J., Avol, E., Lurmann, F., Kuenzli, N., Gilliland, F., Peters, J., et al. (2005). Childhood asthma and exposure to traffic and nitrogen dioxide. *Epidemiology, 16*(6), 737-743.

Gauderman, J., Vara, H., McConnell, R., Berhane, K., Gilliland, F., Thomas, D., Lurmann, F., Avol, E., Kunzli, N., Jerrett, M., & Peters, J. (2007). Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study. *Lancet*, *369*, 571-577.

Gauvin, L., Riva, M., Barnett, T., Richard, L., Craig, C. L., Spivock, M., et al. (2008). Association between neighborhood active living potential and walking. *American Journal of Epidemiology, 167*(8), 944-953.

Gauvin, L., Robitaille, E., Riva, M., McLaren, L., Dassa, C., & Potvin, L. (2007). Conceptualizing and operationalizing neighbourhoods. *Canadian Journal of Public Health*, *98*(1), S18-S26.

Gehring, U., Wijga, A.H., Brauer, M., Fisher, P., De Jongste, J.C., Kerkhof, M., & et al. (2010). Traffic-related air pollution and the development of asthma and allergies during the first 8 years of life. *American Journal of Respiratory and Critical Care Medicine*, *181*(6), 596-603.

Généreux, M., Auger, N., Goneau, M., & Daniel, M. (2008). Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *Journal of Epidemiology and Community Health, 62*(8), 695-700.

Germain, A., & Gagnon, J.E. (1999). Is neighborhood a black box? a reply to galster, metzger, and waite. *Canadian Journal of Urban Research*, *8*(2), 172.

Gilbert, N.L., Goldberg, M.S., Beckerman, B., Brook, J.R., & Jerrett, M. (2005). Assesssing spatial variability of ambient nitrogen dioxide in montreal, canada, with a land-use regression model. *Journal of Air & Waste Management Association*, *55*, 1059-1063.

Gilman, S. E., Abrams, D. B., & Buka, S. L. (2003). Socioeconomic status over the life course and stages of cigarette use: Initiation, regular use, and cessation. *Journal of Epidemiology and Community Health, 57*(10), 802-808.

Glazier, R. H., Badley, E. M., Gilbert, J. E., & Rothman, L. (2000). The nature of increased hospital use in poor neighbourhoods: Findings from a canadian inner city. *Canadian Journal of Public Health, 91*(4), 268-273.

Godley, J., Haines, V.A., Hawe, P., & Shiell, A. (2010). Small area contextual effects on self-reported health: evidence from riverside, calgary. BMC Public Health, 10, 264.

Goovaerts, P. (2010). Geostatistical analysis of county-level lung cancer mortality rates in the southeastern united states. Geographical Analysis, 42, 32-52.

Gottlieb, D. J., Beiser, A. S., & O'Connor, G. T. (1995). Poverty, race, and medication use are correlates of asthma hospitalization rates: A small area analysis in boston. *Chest, 108*(1), 28-35.

Gouveia, N., & Fletcher, T. (2000). Respiratory diseases in children and outdoor air pollution in sao paulo, brazil: a time series analysis. Occup Environ Med, 57, 477-483.

Grineski, S.E., Staniswalis, J.G., Peng, Y., & Atkinson-Palombo, C. (2010). Children's asthma hospitalizations and relative risk due to nitrogen dioxide (no_2): effect modification by race, ethnicity, and insurance status. Environmental Research, 110(2), 178-188.

Grossman, M., & Kaestner, R. (1994). The effects of education on health. In J. Behrman & N. Stacey (Eds.), *Social benefits of education* (pp. 69-124). Ann Arbor, MI: University of Michigan.

Guay, M., Stieb, D.M., & Smith-Doiron, M. (2010). Assessment of long-term exposure to air pollution in a longitudinal national health survey. *Journal of Exposure Science and Environmental Epidemiology*.

Guerra, S., Sherrill, D. L., Bobadilla, A., Martinez, F. D., & Barbee, R. A. (2002). The relation of body mass index to asthma, chronic bronchitis, and emphysema. *Chest, 122*(4), 1256-1263.

Gulliver, J., & Briggs, D.J. (2005). Time-space modeling of journey-time exposure to traffic-related air pollution using gis. *Environmental Research*, *97*, 10-25.

Gurjar, B.R., Jain, A., Sharma, A., Agarwal, A., Gupta, P., Nagpure, A.S., & et al. (2010). Human health risks in megacities due to air pollution. *Atmospheric Environment*, *44*(36), 4606-4613.

Hagerty, M.R. (1999). Testing maslow's hierarchy of needs: national quality-of-life across time. *Social Indicators Research*, *46*, 249-271.

Haining, R.P., Kerry, R., & Oliver, M.A. (2010). Geography, spatial data analysis and geostatistics: an overview. *Geographical Analysis*, *42*, 7-31.

Hawley, K., & Moellering, H. (2005). A comparative analysis of areal interpolation method. *Cartography and Geographic Information Science*, *32*(4), 411-423.

Hay, S.I., Noor, A.M., Nelson, A., & Tatem, A.J. (2005). The accuracy of human population maps for public health application. *Tropical Medicine and International Health*, *10*(1), 1073-1086.

Haynes, R., Daras, K., Reading, R., & Jones, A. (2007). Modifiable neighbourhood units, zone design and residents' perceptions. *Health & Place*, *13*, 812-825.

Health Canada. (2006). *Health effects of air pollution*. Retrieved from http://www.hc-sc.gc.ca/ewh-semt/air/out-ext/effe/health_effects-effets_sante-eng.php

Health Canada. (2006a). *National ambient air quality and objectives (naaqs)*. Retrieved from http://www.hc-sc.gc.ca/ewh-semt/air/out-ext/reg-eng.php#a3

Health Canada. (2010). *People, place and health*. Retrieved from http://www.hc-sc.gc.ca/sr-sr/pubs/hpr-rpms/bull/2007-people-place-gens-lieux/pollution-eng.php

Henderson, S.B., Beckerman, B., Jerrett, M., & Brauer, M. (2007). Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology*, *41*(7), 2422-2428.

Hertzman, C., Frank, J., & Evans, R.G. (1994). Heterogeneities in health status and the determinants of population health. In R.G. Evans, M.L. Barer, & T.R. Marmor (Eds.), *Why are some people healthy and others not?* (pp. 67-92). New York, NY: Aldine de Gruyter.

Hesterberg, T. W., Bunn, W. B., McClellan, R. O., Hamade, A. K., Long, C. M., & Valberg, P. A. (2009). Critical review of the human data on short-term nitrogen dioxide (NO2) exposures: Evidence for NO2 no-effect levels. *Critical Reviews in Toxicology, 39*(9), 743-781.

Hewitt, C.N. (1991). Spatial variations in nitrogen dioxide concentrations in an urban area. *Atmos Environ*, *25*, 429-434.

Heyman, J., & Fisher, A. (2003). Neighbourhoods, health research, and its relevance to public policy. In I. Kawachi & L.F. Berkman (Eds.), *Neighbourhoods and Health* New York, NY: Oxford University Press.

Higgs, G., & Langford, M. (2009). Giscience, environmental justice, & estimating populations at risk: the case of landfills in wales. *Applied Geography*, *29*, 63-76.

Hinwood, A.L., De Klerk, N., Rodriguez, C., Jacoby, P., Runnion, T., Rye, P., Landau, L., Murray, F., Feldwick, M., & Spickett, J. (2006). The relationship between changes in daily air pollution and hospitalizations in perth, australia 1992-1998: a case cross-over study. *International Journal of Environmental Health Research*, *16*(1), 27-46.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., et al. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment, 42*(33), 7561-7578.

Holguin, F., Flores, S., Ross, Z., Cortez, M., Molina, M., Molina, L., Rincon, L., Jerrett, M., Berhane, K., Granados, A., & Romieu, I. (2007). Traffic-related exposures, airway function, inflammation, and respiratory symptoms in children. *Am J Respir Cirt Care Med*, *176*, 1236-1242.

Holt, D., Steel, D. G., & Tranmer, M. (1996). Area homogeneity and the modifiable areal unit problem. *Geographical Systems, 3*(2-3), 181-200.

Holt, J.B., Lo, C.P., & Hodler, T.W. (2004). Dasymetric estimation of population density and areal interpolation of census data. *Cartography and Geographic Information Centre*, *31*(2), 103-121.

House, J. S., Lepkowski, J. M., Kinney, A. M., Mero, R. P., Kessler, R. C., & Herzog, A. R. (1994). The social stratification of aging and health. *Journal of Health and Social Behavior, 35*(3), 213-234.

Howden-Chapmam, P. (2004). Housing standards: A glossary of housing and health. *Journal of Epidemiology and Community Health, 58*(3), 162-168.

Hoxby, C.M. (2000). Does competition among public schools benefit students and taxpayers? *The American Economic Review*, *90*(5), 1209-1238.

Hu, Z., & Rao, K.R. (2009). Particulate air pollution and chronic ischemic heart disease in the eastern united states: a county level ecological study using satellite aerosol data. *Environmental Health*, *8*, 26.

Jaakkola, J. J., Jaakkola, N., & Ruotsalainen, R. (1993). Home dampness and molds as determinants of respiratory symptoms and asthma in pre-school children. *Journal of Exposure Analysis and Environmental Epidemiology, 3 Suppl 1*, 129-142.

Jaakkola, M. S., Ernst, P., Jaakkola, J. J. K., N'Gan'ga, L. W., & Becklake, M. R. (1991). Effect of cigarette smoking on evolution of ventilatory lung function in young adults: An eight year longitudinal study. *Thorax, 46*(12), 907-913.

Jaakkola, M. S., Jaakkola, J. J. K., Ernst, P., & Becklake, M. R. (1993). Respiratory symptoms in young adults should not be overlooked. *American Review of Respiratory Disease, 147*(2), 359-366.

Jacquez, G.M., Maruca, S., & Fortin, M.-J. (2000). From fields to objects: a review of geographic boundary analysis. *Journal of Geographic Systems*, *2*, 221-241.

James, P. D., Wilkins, R., Detsky, A. S., Tugwell, P., & Manuel, D. G. (2007). Avoidable mortality by neighbourhood income in canada: 25 years after the establishment of universal health insurance. *Journal of Epidemiology and Community Health, 61*(4), 287-296.

Jelinski, D.E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, *11*(3), 129-140.

Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N., & Finkelstein, M.M. (2007). Modeling in intraurban variability of ambient traffic pollution in toronto, canada. *Journal of Toxicology and Environmental Health, Part A*, *70*, 200-212.

Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., & Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, *15*, 185-204.

Jerrett, M., Burnett, R.T., Brook, J., Kanaroglou, P., Giovis, C., Finkelstein, N., & Hutchinson, B. (2004). Do socioeconomic characteristics modify the short term association between air pollution and mortality? evidence from a zonal time series in hamilton, canada. *Journal of Epidemiology & Community Health*, *58*(1), 31-41.

Jerrett, M., Burnett, R.T., Goldberg, M.S., Sears, M., Krewski, D., Catalan, R., Kanaroglou, P., Giovis, C., & Finkelstein, N. (2003). Spatial analysis for environmental health research: concepts, methods and examples. *Journal of Toxicology and Environmental Health, Part A*, *66*, 1783-1810.

Jerrett, M., Burnett, R.T., Ma, R., Pope III, C.A., Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E., & Thun, M.J. (2005). Spatial analysis of air pollution and mortality in los angeles. *Epidemiology*, *16*(6), 727-736.

Jerrett, M., & Finkelstein, M. (2005). Geographies of risk in studies linking chronic air pollution exposure to health outcomes. *Journal of Toxicology and Environmental Health Part A*, *68*, 1207-1242.

Jerrett, M., Finkelstein, M.M., Brook, J.R., Arain, M.A., Kanaroglou, P., Stieb, D.M., Gilbert, N.L., Verma, D., Finkelstein, N., Chapman, K.R., & Sears, M.R. (2009). A cohort study of traffic-related air pollution and mortality in toronto, ontario, canada. *Environmental Health Perspectives*, *117*(5), 772-777.

Jerrett, M., Gale, S., & Kontgis, C. (2010). Spatial modeling in environmental and public health research. *Int. J. Environ. Res. Public Health*, *7*, 1302-1329.

Jerrett, M., Shankardass, K., Berhane, K., Gauderman, W.J., Kunzli, N., Avol, E., Gilliland, F., Lurmann, F., Molitor, J.N., Molitor, J.T., Thomas, D.C., Peters, J., & McConnell, R. (2008). Traffic-related air pollution and asthma onset in children: a prospective cohort study with individual exposure measurement. *Environmental Health Perspectives*, *116*(10), 1433-1438.

Jones, A.P., van Sluijs, E.M.F., Ness, A.R., Haynes, R., & Riddoch, C.J. (2010). Physical activity in children: does how we define neighbourhood matter? *Health & Place*, *16*, 236-241.

Joseph, K.S., & Kramer, M.S. (1997). Recent trends in infant mortality rates and proportions of low birth weight live births in canada. *Canadian Medical Association Journal, 157*, 535-541.

Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, *151*(2), 362-367.

Kan, H., London, S.J., Chen, G., Zhang, Y., Song, G., Zhao, N., & et al. (2008). Season, sex, age, and education as modifiers of the effects of outdoor air pollution on daily mortality in shanghai, china: the public health and air pollution in asia (papa) study. *Environmental Health Perspectives*, *116*(9), 1183-1188.

Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., & Brook, J.R. (2005). Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. *Atmospheric Environment*, *39*, 2399-2409.

Kashima, S., Yorifuji, T., Tsuda, T., & Doi, H. (2009). Application of land use regression to regulatory air quality data in japan. *Science of the Total Environment*, *407*, 3055-3062.

Kawachi, I., & Berkman, L.F. (2004). Introduction in neighborhoods and health. In I. Kawachi & L.F. Berkman (Eds.), *Neighborhoods and health* (pp. 1-19). New York, NY: Oxford University.

Kawachi, I., & Kennedy, B. P. (1999). Income inequality and health: Pathways and mechanisms. *Health Services Research, 34*(1 II), 215-227.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health, 87*(9), 1491-1498.

Kim, J.Y., Burnett, R.T., Neas, L., Thurston, G.D., Schwartz, J., Tolbert, P.E., Brunekreef, B., Goldberg, M.S., & Romieu, I. (2007). Panel discussion review: session two - interpretation of observed associations between multiple ambient air pollutants and health effects in epidemiologic analyses. *Journal of Exposure Science and Environmental Epidemiology*, *17*, S83-S89.

Kind, P., Dolan, P., Gudex, C., & Williams, A. (1998). Variations in population health status: Results from a united kingdom national questionnaire survey. *British Medical Journal, 316*(7133), 736-741.

Kogevinas, M., Antó, J. M., Sunyer, J., Tobias, A., Kromhout, H., & Burney, P. (1999). Occupational asthma in europe and other industrialised areas: A population-based study. *Lancet, 353*(9166), 1750-1754.

Krieger, N., Williams, D.R., & Moss, N.E. (1997). Measuring social class in u.s. public health research: concepts, methodologies, and guidelines. *Annual Review of Public Health*, *18*, 341-378.

Krieger, N., Zierler, S., Hogan, J.W., Waterman, P., Chen, J., Lemieux, K., & Gjelsvik, A.(2003). Geocoding and measurement of neighborhood socioeconomic position: a U.S. perspective. In I. Kawachi & L.F. Berkman (Eds.), *Neighborhoods and health* (pp. 147-178). New York, NY: Oxford University Press.

Kristjansson, E.A., Sawada, M., Labonte, R., Parenteau, M.-P., Calhoun, M., Leclair, S., Runnels, V., & Musiol, A. (2007). Neighborhood resources, socioeconomic status and health outcomes in ottawa. *Proceedings of the 6th International Conference on Urban Health*, http://www.be-md.ncbi.nlm.nih.gov/pmc/articles/PMC2232040/

Kwan, M.P. (2009). From place-based to people-based exposure measures. *Social Science & Medicine*, *69*, 1311-1313.

Lang, D. M., & Polansky, M. (1994). Patterns of asthma mortality in philadelphia from 1969 to 1991. *New England Journal of Medicine, 331*(23), 1542-1546.

Langford, M. (2007). Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems*, *31*, 19-32.

Langford, M., & Higgs, G. (2006). Measuring potential access to primary healthcare services: the influence of alternative spatial representations of population. *The Professional Geographer*, *58*(3), 294-306.

Langford, M., Higgs, G., Radcliff, J., & White, S. (2008). Urban population distribution models and service accessibility estimation. *Computers, Environment and Urban Systems*, *32*, 66-80.

Lanza, S.T., Collins, L.M., Schafer, J.L., & Flaherty, B.P. (2005). Using data augmentation to obtain standard errors and conduct hypothesis tests in latent class and latent transition analysis. *Psychological Methods*, *10*(1), 84-100.

Larson, T., Henderson, S.B., & Brauer, M. (2009). Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.*, *43*, 4672-4678.

Lebel, A., Pampalon, R., & Villeneuve, P. Y. (2007). A multi-perspective approach for defining neighbourhood units in the context of a study on health inequalities in the quebec city region. *International Journal of Health Geographics, 6.*

Lee, B.A., Reardon, S.F., Firebaugh, G., Farrell, C.R., Matthews, S.A., & O'Sullivan, D. (2008). Beyond the census tract: patterns of determinants of racial segregation at multiple geographic scales. *American Sociological Review: Oct. 2008*, *73*(5), 766-791.

Lee, D., & Ferguson, C. (2009). Air pollution and health in scotland: a multicity study. *Biostatistics*, *10*(3), 409-423.

Lee, J. -., Son, J. -., Kim, H., & Kim, S. -. (2006). Effect of air pollution on asthma-related hospital admissions for children by socioeconomic status associated with area of residence. Archives of Environmental and Occupational Health, 61(3), 123-130.

Leem, J. -., Kaplan, B. M., Shim, Y. K., Pohl, H. R., Gotway, C. A., Bullard, S. M., et al. (2006). Exposures to air pollutants during pregnancy and preterm delivery. *Environmental Health Perspectives, 114*(6), 905-910.

Lefrance, A., Pascal, L., Larrieu, S., Blanchard, M., Wagner, V., Declercq, C. (2009). Pollution atmosphérique et maladies cardiovasculaires: éléments apportés par le programme de surveillance air et santé. *Archives des Maladies Professionnelles et de l'Environnement*, *70*, 339-345.

Legendre, P. 2000. Program Mod_t_test. Departement de sciences biologiques, Universite de Montreal. Available on http://www.fas.umontreal.ca/BIOL/legendre/

Legendre, P., & Fortin, M.J. (1989). Spatial pattern and ecological analysis. *Vegetation*, *80*, 107-138.

Lin, M., Chen, Y., Villeneuve, P. J., Burnett, R. T., Lemyre, L., Hertzman, C., et al. (2004). Gaseous air pollutants and asthma hospitalization of children with low household income in vancouver, british columbia, canada. *American Journal of Epidemiology, 159*(3), 294-303.

Lindgren, A., Stroh, E., Montnémery, P., Nihlén, U., Jakobsson, K., & Axmon, A. (2009). Traffic-related air pollution associated with prevalence of asthma and COPD/chronic bronchitis. A cross-sectional study in southern sweden. *International Journal of Health Geographics, 8*(1).

Linn, W. S., & Gong Jr., H. (1999). The 21st century environment and air quality influences on asthma. *Current Opinion in Pulmonary Medicine, 5*(1), 21-26.

Linn, W.S., Szlachcic, Y., Gong Jr., H., Kinney, P.L., & Berhane, K.T. (2000). Air pollution and daily hospital admissions in metropolitan los angeles. *Environmental Health Perspectives*, *105*(5), 427-434.

Liu, X.H., Kyriadis, P.C., & Goodchild, M.F. (2008). Population-density estimation using regression and area-to-point residual kriging. *International Journal of Geographical Information Science*, *22*(4), 431-447.

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic inequalities in depression: A meta-analysis. *American Journal of Epidemiology, 157*(2), 98-112.

Loughnan, M.E., Nicholls, N., & Tapper, N.J. (2008). Demographic, seasonal, and spatial differences in acute myocardial infarction admissions to hospital in melbourne, australia. *International Journal of Health Geographics*, *7*, 42.

Lu, H., & Carlin, B.P. (2005). Bayesian areal wombling for geographical boundary analysis. *Geographical Analysis*, *37*(3), 265-285.

Luginaah, I., Xu, X., Fung, K., Grgicak-Mannions, A., Wintermute, J., Wheeler, A., & Brook, J. (2006). Establishing the spatial variability of ambient nitrogen dioxide in windsor, ontario. *International Journal of Environment Studies*, *63*(4), 487-500.

Luginaah, I.N., Fung, K.Y., Gorey, K.M., Webster, G., & Wills, C. (2005). Association of ambient air pollution with respiratory hospitalization in a government-designated "area of concern": the case of windsor, ontario. *Environmental Health Perspectives*, *113*(3), 290-296.

Lwebuga-Mukasa, J. S., Oyana, T., Thenappan, A., & Ayirookuzhi, S. J. (2004). Association between traffic volume and health care use for asthma among residents at a U.S.-canadian border crossing point. *Journal of Asthma, 41*(3), 289-304.

Lwebuga-Mukasa, J.S., Oyana, T.S., & Wydro, P. (2004). Risk factors for asthma prevalence and chronic respiratory illnesses among residents of different neighbourhoods in buffalo, new york. *J. Epidemiol Community Health*, *58*, 951-957.

Lwebuga-Mukasa, J. S., Wojcik, R., Dunn-Georgiou, E., & Johnson, C. (2002). Home environmental factors associated with asthma prevalence in two buffalo inner-city neighborhoods. *Journal of Health Care for the Poor and Underserved, 13*(2), 214-228.

Lynch, J. W., Kaplan, G. A., & Salonen, J. T. (1997). Why do poor people behave poorly? variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic lifecourse. *Social Science and Medicine, 44*(6), 809-819.

Lynch, J., & Kaplan, G. (2000). Socioeconomic position. In L.F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 13-35). New York, NY: Oxford University Press.

Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: How can we conceptualise, operationalise and measure them? *Social Science and Medicine, 55*(1), 125-139.

Macintyre, S., Ellaway, A., Der, G., Ford, G., & Hunt, K. (1998). Do housing tenure and car access predict health because they are simply markers of income or self esteem? A scottish study. *Journal of Epidemiology and Community Health, 52*(10), 657-664.

MacIntyre, E. A., Karr, C. J., Koehoorn, M., Demers, P. A., Tamburic, L., Lencar, C., et al. (2010). Residential air pollution and otitis media during the first two years of life. *Epidemiology. In press.*

Makri, A., & Stilianakis, N.I. (2008). Vulnerability to air pollution health effects. *International Journal of Hyginene and Environmental Health*, *211*, 326-336

Malherbe, L., Cardenas, G., Colin, P., & Yahyaoui, A. (2008). Using different spatial scale measurements in a geostatistically based approach for mapping atmospheric nitrogen dioxide concentrations; applications to the french centre region. *Environmetrics*, *19*(7), 751-764.

Manitoba Centre for Health Policy, University of Manitoba. (2007). *Term: ambulatory care sensitive (acs) conditions* Winnipeg: University of Manitoba . Retrieved from http://mchp-appserv.cpe.umanitoba.ca/viewDefinition.php?definitionID=102237

Mannino, D. M. (2002). COPD: Epidemiology, prevalence, morbidity and mortality, and disease heterogeneity. *Chest, 121*(5 SUPPL.), 121S-126S.

Mannino, D. M., & Buist, A. S. (2007). Global burden of COPD: Risk factors, prevalence, and future trends. *Lancet, 370*(9589), 765-773.

Marcon, A., Corsico, A., Cazzoletti, L., Bugiani, M., Accordini, S., Almar, E., et al. (2009). Body mass index, weight gain, and other determinants of lung function decline in adult asthma. *Journal of Allergy and Clinical Immunology, 123*(5), 1069-1074.e4.

Marmot, M. (2005). The social environment and health. *Clinical Medicine, Journal of the Royal College of Physicians of London, 5*(3), 244-248.

Marshall, J. D., Brauer, M., & Frank, L. D. (2009). Healthy neighborhoods: Walkability and air pollution. *Environmental Health Perspectives, 117*(11), 1752-1759.

Marshall, J.D., Nethery, E., & Brauer, M. (2008). Within-urban variability in ambient air pollution: comparison of estimation methods. *Atmospheric Environment*, *42*, 1359-1369.

Martin, D. (2004). Neighborhoods and area statistics in the post 2001 census era. *Area*, *36*(2), 136-145.

Martin, D.G. (2003). Enacting neighborhood. *Urban Geography*, *24*(5), 361-385.

Maslow, A.H. (1968). *Toward a psychology of being, 2nd ed*. New York, NY: Van Nostrand.

Maslow, A.H. (1970). *Motivation and personality*. New York, NY: Harper & Row.

Matthews, S.A. (2008). The salience of neighbourhood; some lessons from sociology. *Am J Prev Med*, *34*(3), 257-259.

Mayer, J.D. (1983). The role of spatial analysis and geographic data in the detection of disease causation. *Soc. Sci. Med*, *17*(16), 1213-1221.

McConnell, R., Berhane, K., Gilliland, F., London, S. J., Vora, H., Avol, E., et al. (1999). Air pollution and bronchitic symptoms in southern california children with asthma. *Environmental Health Perspectives, 107*(9), 757-760.

McConnell, R., Berhane, K., Gilliland, F., Molitor, J., Thomas, D., Lurmann, F., Avol, E., Gauderman, W.J., & Peters, J.M. (2003). Prospective study of air pollution and bronchitic symptoms in children with asthma. *Am J Respir Crit Care Med*, *168(7)*, 790-797.
McConnell, R., Islam, T., Shankardass, K., Jerrett, M., Lurmann, F., Gilliland, F., & et al. (2010). Childhood incident asthma and traffic-related air pollution at home and school. *Environmental Health Perspectives*, *118*(7), 1021-1026.

McElmurry, B., Buseh, A., & Dublin, M. (1999). Health education program to control asthma in multiethnic, low-income urban communities: the chicago health corps asthma program. *Chest*, *116*, 198S-199S.

McLeod, H., Langford, I. H., Jones, A. P., Stedman, J. R., Day, R. J., Lorenzoni, I., et al. (1998). *Regional variations in the socio-economic distribution of air pollutants in england and wales: Implications for environmental justice.*

McNeill, L. H., Kreuter, M. W., & Subramanian, S. V. (2006). Social environment and physical activity: A review of concepts and evidence. *Social Science and Medicine, 63*(4), 1011-1022.

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in america: Changes in core discussion networks over two decades. *American Sociological Review, 71*(3), 353-375.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer*, *55*(1), 31-42.

Mennis, J., & Hultgren, T. (2006). Intelligent dasymetric mapping and its application to areal interpolation. *Cartography and Geographic Information Science*, *33*(3), 179-194.

Miller, L., Lemke, L.D., Xu, X., Molaroni, S.M., You, H., Wheeler, A.J., & et al (2010). Intra-urban correlation and spatial variability of air toxics across an international airshed in detroit, michigan (u.s.a.) and windsor, ontario (canada). *Atmospheric Environment*, *44*(9), 1162-1174.

Minet Kinge, J., & Morris, S. (2010). Socioeconomic variation in the impact of obesity on health-related quality of life. *Social Science and Medicine, 71*(10), 1864-1871.

Ministry of the Environment. (2010). *Nitrogen dioxide*. Retrieved from http://www.airqualityontario.com/science/pollutants/nitrogen.cfm

Mohai, P., & Saha, R. (2006). Reassessing racial and socioeconomic disparities in environmental justice research. *Demography*, *43*(2), 383-399.

Mohai, P., & Saha, R. (2007). Racial inequality in the distribution of hazardous water: a national-level reassessment. *Social Problems*, *54*(3), 343-370.

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Bales, V. S., et al. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Journal of the American Medical Association, 289*(1), 76-79.

Mokdad, A.H., Marks, J.S., Stroup, D.F., & Gerberding, J.L. (2004). Actual causes of death in the united states, 2000. *JAMA*, *291*(10), 1238-1245.

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., & Agius, R. (2010). Modelling air pollution for epidemiologic research - part I: A novel approach combining land use regression and air dispersion. *Science of the Total Environment, 408*(23), 5862-5869.

Moner, S.E. Public Health Agency of Canada, (1994). *Smoking and pregnancy* Ottawa: Public Health Agency of Canada. Retrieved from http://www.phac-aspc.gc.ca/publicat/clinic-clinique/pdf/s1c3e.pdf

Moore, D.K., Jerrett, M., Mack, W.J., & Kunzli, N. (2007). A land use regression model for predicting ambient fine particulate matter across los angeles, ca. *J. Environ. Monit.*, *9*, 246-252.

Morland, K., Wing, S., Diez-Roux, A., & Poole, C. (2002). Neighborhood characteristics associated with the location of food stores and food service places. *American Journal of Preventive Medicine*, *22*(1), 23-28.

Mortimer, K. M., Neas, L. M., Dockery, D. W., Redline, S., & Tager, I. B. (2002). The effect of air pollution on inner-city children with asthma. *European Respiratory Journal, 19*(4), 699-705.

Most, M.T., Sengupta, R., & Burgener, M.A. (2004). Spatial scale and population assignment choices in environmental justice analyses. *The Professional Geographer*, *56*(4), 574-586.

Mu, L., & Wang, F. (2008). A scale-space clustering method: mitigating the effect of scale in the analysis zone-based data. *Annals of the Association of American Geographers*, *98*(1), 85-101.

Mukerjee, S., Smith, L.A., Johnson, M.M., Neas, L.M., & Stallings, C.A. (2009). Spatial analysis and land use regression of vocs and no2 from school-based urban air monitoring in detroit/dearborn, u.s.a.. *Science of the Total Environment*, *407*, 4642-4651.

Mullahy, J., & Portney, P. R. (1990). Air pollution, cigarette smoking, and the production of respiratory health. *Journal of Health Economics, 9*(2), 193-205.

National Capital Commission, Ottawa. (1998). National capital region gis vector data. [Data File].Available from the National Capital Commission.

National Capital Commission. (2009, January 12). *The national capital greenbelt; history and culture*. Retrieved from http://www.canadascapital.gc.ca/bins/ncc_web_content_page.asp?cid=16297-16299-9735-113846-9742&lang&bhcp=1&lang=1

Neupane, B., Jerrett, M., Burnett, R.T., Marrie, T., Arain, A., & Loeb, M. (2010). Long-term exposure to ambient air pollution and risk of hospitalization with community-acquired pneumonia in older adults. *American Journal of Respiratory and Critical Care Medicine*, *181*(1), 47-53.

Nicolai, T., Carr, D., Weiland, S. K., Duhme, H., von Ehrenstein, O., Wagner, C., et al. (2003). Urban traffic and pollutant exposure related to respiratory outcomes and atopy in a large sample of children. *European Respiratory Journal, 21*(6), 956-963.

Noonan, D.S. (2005). Neighbors, barriers and urban environments: are things "different on the other side of the tracks"? *Urban Studies*, *42*(10), 1817-1835.

Oakes, J.M. (2004). The (mis)estimation of neighbourhood effects: causal inference for a practicable social epidemiology. *Social Science and Medicine*, *58*(10), 1929-1952.

O'Connor, G.T., Neas, L., Vaughn, B., Kattan, M., Mitchell, H., Crain, E.F., Evans III, R., Gruchalla, R., Morgan, W., Stout, J., Adams, G.K., & Lippmahnn, M. (2008). Acute respiratory health effects of air pollution on children with asthma in us inner cities. *J. Allergy Clin. Immunol*, *121*, 1133-1139.

O'Neill, M.S., Jerrett, M., Kawachi, I., Levy, J.I., Cohen, A.J., Gouveia, N., Wilkinson, P., Fletcher, T., Cifuentes, L., & Schwartz, J. (2003). Health, wealth, and air pollution: advancing theory and methods. *Environmental Health Perspectives*, *111*(16), 1861-1870.

Ontario Ministry of the Environment. (2006). *Air quality in ontario; 2005 report*. Retrieved from http://www.ene.gov.on.ca/en/publications/air/6041e.pdf

Ontario Ministry of the Environment, Standards Development Branch. (2008). *Ontario's ambient air quality criteria*. Retrieved from http://www.ene.gov.on.ca/publications/6570e-chem.pdf

Ontario Ministry of the Environment. (2010). *Nitrogen dioxide (no$_2$)*. Retrieved from http://www.airqualityontario.com/science/pollutants/nitrogen.cfm

Openshaw, S. (1977). A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling. *Transactions of the Institute of British Geographers, 2*(4), 459-472.

Openshaw, S. (1984). *The modifiable areal unit problem* [CATMOG 38]. (Adobe Digital Editions version), Retrieved from http://qmrg.org.uk/files/2008/11/38-maup-openshaw.pdf

Openshaw, S., & Taylor, P.J. (1981). The modifiable areal unit problem. In N. Wrigley & R. Bennett (Eds.), *Quantitative geography: a british view* (pp. 60-69). United Kingdom: Routledge & Kegan Paul.

O'Sullivan, D., & Unwin, D.J. (2010). *Geographic information analysis (second edition)*. U.S.A.: John Wiley & Sons, Inc.

Oyana, T.J., & Rivers, P.A. (2005). Geographic variations of childhood asthma hospitalization and outpatient visits and proximity to ambient pollution sources at a u.s.-canada border crossing. *International Journal of Health Geographics*, *4*, 14.

Oyana, T. J., Rogerson, P., & Lwebuga-Mukasa, J. S. (2004). Geographic clustering of adult asthma hospitalization and residential exposure to pollution at a united states-canada border crossing. *American Journal of Public Health, 94*(7), 1250-1257.

Pan, G., Zhang, S., Feng, Y., Takahashi, K., Kagawa, J., Yu, L., Wang, P., Liu, M., Liu, Q., Hou, S., Pan, B., & Li, J. (2010). Air pollution and children's respiratory symptoms in six cities of northen china. *Respiratory Medicine, In Press*.

Park, R.E., Burgess, E.W., & McKenzie, R.D. (1967). *The city*. Chicago, IL: University of Chicago press.

Patel, M.M., Chillrud, S.N., Correa, J.C., Hazi, Y., Feinberg, M., Deepti, K., & et al. (2010). Traffic-related particulate matter and acute respiratory symptoms among new york city area adolescents. *Environmental Health Perspectives*, *118*(9), 1338-1343.

Patterson, L., Urban, M., Myers, A., Bhaduri, B., Bright, E., & Coleman, P. (2009). The effects of quality control on decreasing error propagation in the landscan u.s.a. population distribution model: a case study of philadelphia county. *Transactions in GIS*, *13*(2), 215-228.

Pearce, J., Hiscock, R., Blakely, T., & Witten, K. (2008). The contextual effects of neighborhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. *Journal of Epidemiology and Community Health*, *62*(3), 198-201.

Pearce, J., Witten, K., & Bartie, P. (2006). Neighborhoods and health: a gis approach to measuring community resource accessibility. *Journal of Epidemiology and Community Health*, *60*(5), 389-395.

Peat, J. K., Woolcock, A. J., & Cullen, K. (1987). Rate of decline of lung function in subjects with asthma. *European Journal of Respiratory Diseases, 70*(3), 171-179.

Peden, D. B. (1999). Controlled exposures of asthmatics to air pollutants. *Air Pollution and Health*, 865-880.

Peel, J. L., Tolbert, P. E., Klein, M., Metzger, K. B., Flanders, W. D., Todd, K., et al. (2005). Ambient air pollution and respiratory emergency department visits. *Epidemiology, 16*(2), 164-174.

Piro, F.N., Madsen, C., Naess, O., Nafstad, P., & Claussen, B. (2008). A comparison of self reported air pollution in people with and without chronic diseases. *Environmental Health*, *7*, 9.

Poplawski, K., Gold, T., Setton, E., Allen, R., Su, J., Larson, T., Henderson, S., Brauer, M., Hystad, P., Lightowlers, C., Keller, P., Cohen, M., Silva, C., & Buzzelli, M. (2009). Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. *Journal of Exposure and Environmental Epidemiology*, *19*, 107-117.

Poulsen, E., & Kennedy, L.W. (2004). Using dasymetric mapping for spatially aggregated crime data. *Journal of Quantitative Criminology*, *20*(3), 243-262.

Rabe, K.F., Hurd, S., Anzueto, A., Barnes, P.J., Buist, S.A., Calverley, P., Fukuchi, Y., Jenkins, C., Rodriguez-Roisin, R., van Weel, C., & Zielinski, J. (2007). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease gold executive summary. *Am. J. Respir Crit Care Med*, *176*, 532-555.

Rahkonen, O., Laaksonen, M., & Karvonen, S. (2005). The contribution of lone parenthood and economic difficulties to smoking. *Social Science and Medicine*, *61*, 211-216.

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

Ramos, R.G., Talbot, E.O., Youk, A., & Karol, M.H. (2006). Community urbanization and hospitalization of adults for asthma. *J. Environ Health*, *68*, 26-32.

Raudenbush, S. W., & Sampson, R. J. (1999). Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology, 29*(1), 1-41.

Ray, B., & Preston, V. (2009). Geographies of discrimination: Variations in perceived discomfort and discrimination in canada's gateway cities. *Journal of Immigrant and Refugee Studies, 7*(3), 228-249.

Reardon, S.F., Matthews, S.A., O'Sullivan, D., Lee, B.A., Firebaugh, G., Farrell, C.R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45*(3), 489-514.

Reibel, M., & Agrawal, A. (2007). Areal interpolation of population counts using pre-classified land cover data. *Population Research and Policy Review*, *26*, 619-633.

Reibel, M., & Bufalino, M.E. (2005). Street-weighted interpolation techniques for demographic count estimation in incompatible zone systems. *Environment and Planning A*, *37*, 127-139.

Reijneveld, S. A. (1998). The impact of individual and area characteristics on urban socioeconomic differences in health and smoking. *International Journal of Epidemiology, 27*(1), 33-40.

Reynolds, H., & Amrhein, C. (1997). Using a spatial data set generator in an empirical analysis of aggregation effects on univariate statistics. *Geographical and Environmental Modelling, 1*(2), 199-219.

Richmond, T.K., & Subramanian, S.V. (2009). *School variation in asthma: compositional or contextual? plos one, 4(2)*. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0008512

Riva, M., Gauvin, L., Apparicio, P., & Brodeur, J.-M. (2009). Disentangling the relative influence of built and socioeconomic environments on walking: the contribution of areas homogeneous along exposures of interest. *Social Science & Medicine*, *69*, 1296-1305.

Ross, C. E., & Jang, S. J. (2000). Neighborhood disorder, fear, and mistrust: The buffering role of social ties with neighbors. *American Journal of Community Psychology, 28*(4), 401-420.

Ross, C. E., Reynolds, J. R., & Geis, K. J. (2000). The contingent meaning of neighborhood stability for residents' psychological well-being. *American Sociological Review, 65*(4), 581-597.

Ross, C. E., & Wu, C. -. (1996). Education, age, and the cumulative advantage in health. *Journal of Health and Social Behavior, 37*(1), 104-120.

Ross, N.A., Tremblay, S., & Graham, K. (2004). Neighborhood influences on health in montreal, canada. *Social Science and Medicine*, *59*, 1485-1494.

Ryan, P.H., & LeMasters, G.K. (2007). A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicology*, *19*((Suppl 1)), 127-133.

Ryan, P.H., LeMasters, G.K., Levin, L., Burkle, J., Biswas, P., Hu, S., Grishpun, S., & Reponen, T. (2008). A land-use regression model for estimating microenvironmental diesel exposure given multiple addresses from birth through childhood. *Science of the Total Environment*, *404*, 139-147.

Sabesan, A., Abercrombie, K., Ganguly, A.R., Bhaduri, B., Bright, E.A., & Coleman, P.R. (2007). Metrics for the comparative analysis of geospatial datasets with applications to high-resolution grid-based population data. *GeoJournal*, *69*(1-2), 81-91.

Sadahiro, Y. (1999). Accuracy of areal interpolation: a comparison of alternative methods. *Journal of Geographical Systems*, *1*, 323-346.

Sahsuvaroglu, T., Jerrett, M., Sears, M.R., McConnell, R., Finkelstein, N., Arain, A., Newbold, B., & Burnett, R. (2009). Spatial analysis of air pollution and childhood asthma in hamilton, canada: comparing exposure methods in sensitive subgroups. *Environmental Health*, *8*(14).

Salm Ward, T. C., Mori, N., Patrick, T. B., Madsen, M. K., & Cisler, R. A. (2010). Influence of socioeconomic factors and race on birth outcomes in urban milwaukee. *Wisconsin Medical Journal, 109*(5), 254-260.

Sampson, R.J., & Morenoff, J.D. (2004). Spatial (dis) advantage and homicide in chicago neighborhoods. In M. Goodchild & D. Janelle (Eds.), *Spatially integrated social science* (pp.145-170). New York, NY: Oxford University Press.

Sampson, R.J., Morenoff, J.D., & Gannon-Rowley, T. (2002). Assessing "neighborhood effects": social processes and new directions in research. *Annual Review of Sociology*, *28*, 443-478.

SAS Institute Inc. (2008). Sas/stat 9.2 user's guide, cary. *NC:SAS Institute INC.*

SAS Institute Inc. (version 9,2) [Computer software]. Cary. NC : SAS Institute INC.

Sauerzapf, V., Jones, A.P., & Cross, J. (2009). Environmental factors and hospitalizations for chronic obstructive pulmonary disease in a rural county of england. *J. Epidemiol Community Health*, *63*, 324-328.

Schaefer-McDaniel,N., O'Brien Caughy, M., O'Campo, P., & Geary, W. (2010). Examining methodological details of neighbourhood observations and the relationship to health: a literature review. *Social Science & Medicine*, *70*, 277-292

Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. United Kingdom: Chapman & Hall.

Schikowski, T., Sugiri, D., Ranft, U., Gehring, U., Heinrich, J., Wichmann, H. E., et al. (2007). Does respiratory health contribute to the effects of long-term air pollution exposure on cardiovascular mortality? *Respiratory Research, 8*, 20.

Schimert, J., Schafer, J. L., Hesterberg, T., Fraley, C., & Clarkson, D. B. (2001). *Analyzing Missing Values in S-PLUS.*

Schuurman, N., Bell, N., Dunn, J.R., & Oliver, L. (2007). Deprivation indices, population health and geography: an evaluation of the spatial effectiveness of indices at multiple scales. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *84*(4), 591-603.

Seasons, M. (2003). Indicators and core area planning: application in canada's mid-sized cities. *Planning Practice and Research*, *18*(1), 63-80.

Setton, E., Marshall, J. D., Brauer, M., Lundquist, K. R., Hystad, P., Keller, P., et al. (2010). The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *Journal of Exposure Science and Environmental Epidemiology, Article in Press.*

Shaheen, S. O., Sterne, J. A. C., Montgomery, S. M., & Azima, H. (1999). Birth weight, body mass index and asthma in young adults. *Thorax, 54*(5), 396-402.

Shankardass, K., McConnell, R.S., Milam, J., Berhane, K., Tatalovich, Z., Wilson, J.P., & Jerrett, M. (2007). The association between contextual socioeconomic factors and prevalent asthma in a cohort of southern california school children. *Social Science & Medicine*, *65*, 1792-1806.

Shao, S.J., Su, S.C., & Kanarek, N. (2006). Country-level basics of living affects overall mortality rate in the united states - 2000. *Applied Research in Quality of Life*, *1*(1), 97-106.

Shohaimi, S., Welch, A., Bingham, S., Luben, R., Day, N., Wareham, N., et al. (2004). Area deprivation predicts lung function independently of education and social class. *European Respiratory Journal, 24*(1), 157-161.

Spengler, J. D., Jaakkola, J. J. K., Parise, H., Katsnelson, B. A., Privalova, L. I., & Kosheleva, A. A. (2004). Housing characteristics and children's respiratory health in the russian federation. *American Journal of Public Health, 94*(4), 657-662.

Sridharan, S., Tunstall, H., Lawder, R., & Mitchell, R. (2007). An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in scotland. *Social Science & Medicine*, *65*, 1942-1952.

Stafford, M., Duke-Williams, O., & Shelton, N. (2008). Small are inequalities in health: are we underestimating them? *Social Science & Medicine*, *67*, 891-899.

Statistics Canada. (2001). *Census dictionary*. Retrieved from http://www12.statcan.ca/english/census01/Products/Reference/dict/appendices/92-378-XIE02002.pdf

Statistics Canada. (2002). *2001 community profiles: ottawa-hull (ontario part)*. Retrieved fromhttp://www12.statcan.ca/english/profi101/CP01/Details/Page.cfm?Lang=E&Geo1=CMA&Code1=50535&Geo2=PR&Code2=35&Data=Count&SearchText=Ottawa&SearchType=Begins&SearchPR=01&B1=All&Custom=

Statistics Canada. (2006). *Road network file (rnf) 92-500-gwe, [Data File].*Available from Statistics Canada.

Statistics Canada. (2007a). *2006 community profiles: ottawa-hull (ontario part)*. Retrieved from http://www12.statcan.ca/english/census06/data/profiles/community/Details/Page.cfm?Lang=E&Geo1=CMA&Code1=50535&Geo2=PR&Code2=35&Data=Count&SearchText=Ottawa&SearchType=Begins&SearchPR=01&B1=All&Custom=

Statistics Canada (2007b) Census tract: Detailed Definition. http://geodepot.statcan.gc.ca/2006/180506051805140305/03150707/1215140704051903 18091620091514_05-eng.jsp?REFCODE=10&LANG=E&GEO_LEVEL=12&TYPE=L

Statistics Canada. (2007c). *Dissemination area: detailed definition*. Retrieved from http://geodepot.statcan.gc.ca/2006/180506051805140305/03150707/121514070405190318091620091514_05-eng.jsp?REFCODE=10&LANG=E&GEO_LEVEL=35&TYPE=L.com

Statistics Canada. (2007d). *Dissemination block: detailed definition*. Retrieved from http://geodepot.statcan.gc.ca/2006/180506051805140305/03150707/121514070405190318091620091514_05-eng.jsp?REFCODE=10&LANG=E&GEO_LEVEL=102&TYPE=L

Statistics Canada. (2009). *Data liberation initiative (dli)*. Retrieved from http://www.statcan.gc.ca/dli-ild/dli-idd-eng.htm

Statistics Canada. (2010). Canada Community Health Survey (CCHS). Retrieved from http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3226&lang=en&db=imdb&adm=8&dis=2

Stenius-Aarniala, B., Poussa, T., Kvarnström, J., Grönlund, E. -., Ylikahri, M., & Mustajoki, P. (2000). Immediate and long term effects of weight reduction in obese people with asthma: Randomised controlled study. *British Medical Journal, 320*(7238), 827-832.

Stieb, D. M., Szyszkowicz, M., Rowe, B. H., & Leech, J. A. (2009). Air pollution and emergency department visits for cardiac and respiratory conditions: A multi-city time-series analysis. *Environmental Health: A Global Access Science Source, 8*(1).

Strachan, D. P. (1988). Damp housing and childhood asthma: Validation of reporting of symptoms. *British Medical Journal, 297*(6658), 1223-1226.

Strickland, M. J., Darrow, L. A., Klein, M., Flanders, W. D., Sarnat, J. A., Waller, L. A., et al. (2010). Short-term associations between ambient air pollutants and pediatric asthma emergency department visits. *American Journal of Respiratory and Critical Care Medicine, 182*(3), 307-316.

Su, J.G., Brauer, M., & Buzzelli, M. (2008). Estimating urban morphometry at the neighborhood scale for improvement in modeling long-term average air pollution concentrations. *Atmospheric Environment*, *42*, 7884-7893.

Su, J.G., Jerrett, M., & Beckerman, B. (2009). A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Science of the Total Environment*, *407*, 3890-3898.

Su, J.G., Jerrett, M., Beckerman, B., Wilhelm, M., Ghosh, J.K., & Ritz, B. (2009b). Predicting traffic-related air pollution in los angeles using a distance decay regression selection strategy. *Environmental Research*, *109*, 657-670.

Subramanian, S. V., Perkins, J. M., Özaltin, E., & Smith, G. D. (2011). Weight of nations: A socioeconomic analysis of women in low- to middle-income countries. *American Journal of Clinical Nutrition, 93*(2), 413-421.

Sundquist, J., & Johansson, S. (1997). Self reported poor health and low educational level predictors for mortality: A population based follow up study of 39,156 people in sweden. *Journal of Epidemiology and Community Health, 51*(1), 35-40.

Sundquist, J., Malmström, M., & Johansson, S. (1999). Cardiovascular risk factors and the neighbourhood environment: A multilevel analysis. *International Journal of Epidemiology, 28*(5), 841-845.

Sutton, P.C., Elvidge, C., & Obremski, T. (2003). Building and evaluating models to estimate ambient population density. *Photogrammetric Engineering and Remote Sensing*, *69*(5), 545-553.

Swift, A., Liu, L., & Uber, J. (2008). Reducing maup bias of correlation statistics between water quality and gi illness. *Computers, Environment and Urban Systems*, *32*, 134-148.

Szyszkowicz, M., & Rowe, B. (2010). Air pollution and emergency department visits for chest pain and weakness in edmonton, canada. *International Journal of Occupational Medicine and Environmental Health*, *23*(1), 15-19.

Tanner, M.A., & Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc*, *82*, 528-540.

Thurston, G.D., Bekkedal, M.Y.V., Roberts, E.M., Ito, K., Pope III, A., Glenn, B.S., Ozkaynak, H., & Utell, M. (2009). Use of health information in air pollution health research: past successes and emerging trends. *Journal of Exposure Science and Environmental Epidemiology*, *19*, 45-58.

Tobler, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, *74*, 519-530.

Trupin, L., Earnest, G., San Pedro, M., Balmes, J. R., Eisner, M. D., Yelin, E., et al. (2003). The occupational burden of chronic obstructive pulmonary disease. *European Respiratory Journal, 22*(3), 462-469.

Unwin, A., & Unwin, D. (1998). Exploratory spatial data analysis with local statistics. *The Statistician*, *47*(3), 415-421.

Urquia, M. L., Frank, J. W., Glazier, R. H., Moineddin, R., Matheson, F. I., & Gagnon, A. J. (2009). Neighborhood context and infant birthweight among recent immigrant mothers: A multilevel analysis. *American Journal of Public Health, 99*(2), 285-293.

U.S. Census Bureau. (2007). *Naics - north american industry classification system*. Retrieved from http://www.census.gov/epcd/www/naics.html

Uthman, O.A. (2008). Geographical variations and contextual effects on age of initiation of sexual intercourse among women in nigeria: a multilevel and spatial analysis. *International Journal of Health Geographics*, *7*, 27.

VanHerzele, A., & Wiedemann, T. (2003). A monitoring tool for the provision of accessible and attractive urban green spaces. *Landscape and Urban Planning*, *62*, 109-126.

Villeneuve, P.J., Chen, L., Rowe, B.H., & Coates, F. (2007). Outdoor air pollution and emergency department visits for asthma among children and adults: a case-crossover study in northern alberta, canada. *Environmental Health: A Global Access Science Source*, *6*.

Villeneuve, P.J., Chen, L., Stieb, D., & Rowe, B.H. (2006). Associations between outdoor air pollution and emergency department visits for stroke in edmonton, canada. *European Journal of Epidemiology*, *21*, 689-700.

Walks, R. A., & Maaranen, R. (2008). Gentrification, social mix, and social polarization: Testing the linkages in large canadian cities. *Urban Geography, 29*(4), 293-326.

Waller, L.A., & Gotway, C.A. (2004). *Applied spatial statistics for public health data*. U.S.A.: Wiley.

Walters, S., Phupinyokul, M., & Ayres, J. (1995). Hospital admission rates for asthma and respiratory disease in the west midlands: Their relationship to air pollution levels. *Thorax, 50*(9), 948-954.

Wardle, J., & Steptoe, A. (2003). Socioeconomic differences in attitudes and beliefs about healthy lifestyles. *Journal of Epidemiology and Community Health, 57*(6), 440-443.

Wentz, E., Gober, P., Balling, R.C., & Day, T.A. (2002). Spatial patterns and determinants of winter atmospheric carbon dioxide concentrations in an urban environment. *Annals of the Association of American Geographers*, *92*, 15-28.

Wheeler, A.J., Smith-Doiron, M., Xu, X., Gilbert, N.L., & Brook, J.R. (2008). Intra-urban variability of air pollution in windsor, ontario-measurement and modeling for human exposure assessment. *Environmental Research*, *106*, 7-16.

Wheeler, B.W., & Ben-Shlomo, Y. (2005). Environmental equity, air quality, socioeconomic status, and respiratory health: a linkage analysis of routine data from the health survey for england. *Journal of Epidemiology and Community Health*, *59*(11), 948-954.

Williams, D.R., & Collins, C. (1995). U.S. socioeconomic and racial differences in health: patterns and explanations. *Annual Review of Sociology*, *21*, 349-386.

Willis, A., Krewski, D., Jerrett, M., Goldberg, M.S., & Burnett, R.T. (2003). Selection of ecologic covariates in the american cancer society study. *Journal of Toxicology and Environmental Health, Part A*, *66*, 1563-1589.

Wilson, A., Salloway, J.C., Wake, C.P., & Kelly, T. (2004). Air pollution and the demand for hospital services: a review. *Environment International*, *30*, 1109-1118.

Wilton, D., Szpiro, A., Gould, T., & Larson, T. (2010). Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in los angeles, ca and seattle, wa. *Science of the Total Environment*, *408*, 1120-1130.

Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health, 82*(6), 816-820.

Wrigley, N., Warm, D., Margetts, B., & Whelan, A. (2002). Assessing the impact of improved retail access on diet in a "FOOD DESERT": a preliminary report. *Urban Studies*, *39*, 2061-2082.

WHO. (2004). Health aspects of air pollution. *Health Aspects of Air Pollution,*Retrieved from www.scopus.com

Xie, Y. (1995). The overlaid network algorithms for areal interpolation problem. *Computers, Environment and Urban Systems*, *19*(4), 287-306.

Yen, I.H., & Syme, S.L. (1999). The social environment and health: a discussion of the epidemiologic literature. *Annual Review of the Public Health*, *20*, 287-308.

Yuan, Y., Smith, R. M., & Limp, W. F. (1997). Remodeling census population with spatial information from landsat TM imagery. *Computers, Environment and Urban Systems, 21*(3-4), 245-258.

Zanobetti, A., & Schwartz, J. (2006). Air pollution and emergency admissions in boston, ma. *J Epidemiol Community Health*, *60*, 890-895.

Zemp, E., Elsasser, S., Schindler, C., Künzli, N., Perruchoud, A. P., Domenighetti, G., et al. (1999). Long-term ambient air pollution and respiratory symptoms in adults (SAPALDIA study). *American Journal of Respiratory and Critical Care Medicine, 159*(4 I), 1257-1266.