Characterization of Nitrogen Dioxide and Sulphur Dioxide Ambient Concentrations over an Urban Area through Land use Regression Modeling: An Application to Hamilton, Ontario

Characterization of Nitrogen Dioxide and Sulphur Dioxide Ambient Concentrations over an Urban Area through Land use Regression Modeling: An Application to Hamilton, Ontari

By

Farhana Ferdous, (B. Urban and Regional Planning)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

For the Degree

Masters of Arts

McMaster University

School of Geography and Earth Sciences

MASTER OF ARTS (2008) MCMASTER UNIVERSITY

(School of Geography and Earth Sciences) Hamilton, Ontario

TITLE: Characterization of Nitrogen Dioxide and Sulphur Dioxide Ambient Concentrations over an Urban Area through Land use Regression Modeling: An application to Hamilton, Ontario

AUTHOR: Farhana Ferdous

B. Urban and Regional Planning

(Bangladesh University of Engineering and Technology)

SUPERVISOR: Dr. Pavlos S. Kanaroglou

NUMBER OF PAGES: 139

ABSTRACT

This study presents statistical land use regression models to characterize the spatial distribution of two air pollutants: nitrogen dioxide (NO_2) and sulphur dioxide (SO_2). While nitrogen dioxide (NO_2) is considered to be a marker of vehicular pollution, sulphur dioxide (SO_2) is a major industrial emission. The research uses mobile field-monitored pollution measurement data at over 2000 locations across the Hamilton Census Metropolitan area (CMA) during November 2005–March 2006. Various land use characteristics, transportation attributes and meteorological characteristics are used in the model to explain the spatial variations in the concentration of the pollutants. A key contribution of the paper is that it explicitly includes wind directions (i.e. north-east, north-west, south-west and south-east) into the models.

The results reveal that proximity to the major roads and highways, total length of roads nearby, traffic volume, nearby park and commercial areas, distance to the lakeshore industries and wind speed have major impact on NO_2 concentration. While proximity to the roads and highways and higher total length of roads nearby increases the concentration of NO_2, higher amount of nearby park areas and wind speed decreases the level of concentration of the pollutant. For north-east direction distance to major roads and highways, lakeshore industries, and wind speed has significant impact on the concentration level. While for the north-west direction, commercial area within 350 meter, distance to QEW and Centennial Parkway North intersection contribute significantly to the concentration of NO_2, total length of the roads within 50 meter from the observation points and distance to the railway are significantly associated with the concentration level in case of south-east wind direction model. Lastly, the significance of a dummy variable indicates that the upper city has lower concentration of NO_2 compared to the lower city according to the south-west wind direction model.

For the SO_2 models, the results suggest that proximity to downwind locations from the industries, proximity to commercial and park areas, and proximity to the Bay Front Area play a role to the concentration of the pollutant. While downwind location with

closeness to the industries increases the level of concentration of SO_2, higher wind speed and proximity to the Bay Front Area reduces the level of pollution in general.

All models for NO_2 and SO_2 provide reasonable model fit in terms of coefficient of determination (R^2) of the estimated land use regression models for different major wind directions. Almost all parameters are statistically significant at least at the 5% level of significance. The estimated models are used to generate pollution surfaces for the Hamilton CMA. The resulting spatial distribution of pollution concentration can be useful for informing policy.

ACKNOWLEDGEMENTS

At first all praises belong to the almighty, the most merciful, the most beneficial to man and His action.

I wish to express my sincere gratitude to my supervisor Dr. Pavlos S. Kanaroglou, Professor, School of Geography and Earth Sciences, McMaster University for his guidance, invaluable suggestions and strong support towards the successful completion of the study.

I am also grateful to Dr. Julie Wallace for her constant support, ideas and stimulating discussion throughout the whole study. I am thankful to Dr. Dennis Corr for data support. Special thanks to Dora , Rutba and Tarana for answering my questions time to time.

Thanks to my friends at CSpA: Kevin, Manolis, Hejun, Hanna, Justin, Kosta, Karen, Ceciley, Yefei, Patric Deluca and Laura for making my academic life enjoyable.

Finally, I am grateful to Muhammad Ahsanul Habib, without his support and encouragement I would have never been able to carry forward my works.

Thanks to my parents who always believe in me and taught me to believe in myself.

Table of Content

List of Tables

List of Figures

1. CHAPTER ONE: INTRODUCTION

1.1 Background

Since the 1950s air pollution has become an issue of major public concern in North America and has been linked to a wide variety of health effects (Fenger, 2004). As a matter of fact, air pollution creates problems in several dimensions, both short and long-term, and is continuously destabilizing environmental systems through complex chemical reactions that pose serious threat to both humans and wildlife. Given the significant impact of air quality on human health and environmental conditions, the importance of predicting concentration of different pollutants over space is well recognized (Ferdous and Kanaroglou, 2007). This study uses data collected from a mobile unit in a land use regression model to predict the spatial variability of pollutant concentrations in a particular urban area.

The World Health Organisation (WHO) reports that air pollution alone causes three million deaths annually worldwide (Solaiman, 2007). Recently, the Canadian Medical Association (CMA) released a summary report (National Illness Costs of Air Pollution) that estimates as many as 21,000 Canadians will die prematurely from the effects of air pollution in the year 2008. Although most of these deaths will occur due to chronic exposure to pollution, almost 3,000 will occur as a direct result of acute short-term exposure. The report predicts that air-pollution related deaths would rise over the next two decades, which will cost $1.3 billion annually in healthcare and loss in productivity. While the Provinces of Ontario and Quebec will bear the brunt, the number of premature deaths could double, from 1,200 a year to 2,200 by 2031 in these two provinces. In addition, the hospital admissions could jump by as much as 70%, and the annual healthcare and economic costs could rise by as much as 30%, from $570 million to $740 million (The National Medical Association, 2008). Therefore, it is important to study factors of pollution and level of pollution in relation to locations of human activities in order to reduce health risks.

Hamilton, a city in the Province of Ontario, has many medium and large steel industries, and is well known as the "steel city" of Canada. Two large North American steel industries are located in the area along the shore of lake Ontario. The city also has other heavy industries in its northeastern part, and is bounded by some of the busiest highways in Ontario. In addition, other geographical features of the city, such as the Niagara Escarpment that runs through the middle of the city play a role in the pollution problem. In summary, longstanding stationary sources, steady mobile sources, and its geographical setting make Hamilton one of the most polluted cities in Canada. The population of the city is exposed to detrimental air pollution and faces several health hazards. According to the Clean Air Hamilton (2007) nitrogen dioxide, sulphur dioxide, ground level ozone, fine particulate matter, and carbon monoxide are harmful pollutants that contribute to approximately 100 premature deaths and 620 hospital admissions annually in Hamilton. Therefore, analysis of concentration of air pollutants in the Hamilton Census Metropolitan Area (CMA) is important to help combat the pollution problem in the region.

Recently, a significant body of research has investigated the spatial context of air pollution, and applied spatial analytical techniques to understand impacts of land use, transportation infrastructure, location of industries and other key activity centers (Su et al., 2007; Sahsuvaroglu et al., 2006; Gilbert et al., 2005; Ross et al., 2006; Wheeler et al., 2008; Moore et al., 2007 among others). Since several studies have demonstrated that land use regression models can generate reasonably accurate small-area estimates of concentrations of air pollutants without using computationally intensive dispersion modeling (Brauer et al., 2003; Briggs et al. 2000), this study employs land use regression models to characterize the concentration of air pollutants over the Hamilton CMA. The study looks into the spatial relationship between the air pollutants and surrounding land uses, transportation factors and major activity centers. Since nitrogen dioxide (NO_2) is often called the marker of vehicular pollution and sulphur dioxide (SO_2) is a major industrial emission, this research focuses on these two pollutants, which are also major contributors to the health problems in Hamilton (Clean Air Hamilton, 2007).

1.2 Objectives

The key objective of the research is to explore underlying factors that contribute to the concentration of pollutants using the land use regression method. As level of pollution varies over space due to various land use characteristics, transportation attributes and meteorological conditions, this spatial analytic technique helps in understanding the influence of these spatial features to the ambient concentration of NO_2 and SO_2. This study uses mobile field-monitored pollution data to characterize the pollution concentration, and generates pollution surface that represents distribution of pollutants in the Hamilton CMA for four wind directions.

The specific objectives of this research can be summarized as follows:

1. To develop, estimate, and analyze relationships between the concentrations of NO_2 and SO_2, and various urban factors, including land use characteristics, transportation attributes and meteorological conditions using land use regression modeling techniques

2. To develop models for four major wind directions for each pollutant in the Hamilton CMA.

1.3 Contributions of the thesis

This thesis contributes to the existing body of evidence of spatial relationships of pollution concentration. It explains spatial variability of the pollutants (NO_2 and SO_2) in relation to land use, transportation and other key locational factors for the Hamilton Census Metropolitan Area. It also characterizes concentration of pollution level in terms of four major wind directions, which is often ignored in most of the existing literature. In addition, in earlier studies, land use regression models are applied to the daily/monthly/yearly averaged pollution data, whereas this study uses mobile field-observed data that are collected at a particular point in time. Finally, this study generates pollution surfaces for the Hamilton CMA, which might be useful for informed decision-making.

1.4 Organization of the Study

Chapter two reviews the literature on adverse effects of the pollutants (NO_2 and SO_2), and various approaches of land use regression models to explain the variability of air pollutants. Chapter three discusses the data sources such as pollutant concentrations, meteorology, land uses, transportation attributes, and demographic information. This chapter also describes the basic concepts of the land use regression models. Chapter 4 estimates the spatial relationship between NO_2 concentration and various land uses, transportation and locational attributes for four different wind directions. Chapter 5 analyzes the spatial relationship between SO_2 concentration and land uses, transportation and locational attributes for four different wind directions. Chapter 6 concludes the research and presents recommendations for future research directions.

2. CHAPTER TWO: LITERATURE REVIEW

Although technological advances and developments in different sectors have brought comfort and increased quality of life in modern societies, the progress has increased the emission footprint of each individual. These emissions cause a variety of health-hazards and threaten environmental sustainability. Air pollution is one of the major environmental concerns, which has received considerable attention in recent years. Rapid industrialization, continuous urbanization, population growth and automobile use play crucial roles in contributing to air pollution. In recent years, a substantial amount of research has investigated causes of air pollution, relationships between land use and emissions, environmental impacts and effects of pollution on human health. Consequently, spatial analysis of air pollution has become an important research stream that has helped in understanding the effects of air pollution in urban areas. Although there are many types of air pollutants that significantly affect human health, this research project focuses on two pollutants: (a) nitrogen dioxide (NO_2) and sulphur dioxide (SO_2). While nitrogen dioxide (NO_2) is one of the major vehicular emissions, sulphur dioxide (SO_2) is mainly an industrial emission. This chapter describes key characteristics of these two pollutants, their adverse health effects and spatial analysis techniques that are used to understand relationships between land use, transportation infrastructure and the urban environment. It also discusses the application of land use regression models in different parts of the world.

2.1 *Nitrogen dioxide (NO_2)*

NO_2 is an inorganic gas, which is reddish brown in colour with a pungent and irritating odour. It is formed by the oxidation of NO, which is emitted when fuel is burned at high temperature (Eq 2.1)

$$2NO + O_2 \rightarrow 2NO_2 \qquad \text{(Eq 2.1)}$$

Primary sources of outdoor NO are motor vehicle exhaust, and industrial sources such as electric utilities and industrial boilers. The pollutant NO_2 is very reactive and may combine with other atmospheric components to form nitric acid (HNO_3) and nitrates, which contribute to acid rain and particulate matter in the atmosphere. It is highly corrosive to metals, and frequent and long-term exposure to the NO_2 may cause to lower resistance to respiratory infection. NO_2 is an important pre-cursor emission for the formation of ground-level ozone and stands as one of the principal components of smog (Srivastava, 2004).

2.2 *Adverse Effects of NO_2*

With infrastructure development and advanced technology, cities are expanding daily with high capacity roads and highways. In addition, sub-urban living coupled with busy schedules and time constraints, people are becoming more and more auto-dependent. This ultimately results in more traffic emissions and further degradation of air quality in urban areas. Transportation contributes significantly to the emission of NO_2, which is often considered as a good indicator of traffic emissions (Rotko, 2002). It is one of the harmful pollutants that significantly deteriorate the urban air quality (Ross et al. 2006). Environment Canada reports that transportation sector alone emitted 1,263,145 tones NOx in 2005 annually.

Hoek et al. (2002) found that there is a high correlation between the mortality and the concentration of NO_2. In a recent study Shao, et al. (2002) found that children living within 200 meters of roads with heavy traffic density and heavy truck movements have higher risk of hospitalization due to asthma in New York. In addition, it is found that higher level of NO_2 causes to slower lung growth in Southern California. Children from highly polluted areas experienced slower development of lung compared to children who live in a less polluted environment.

In a recent study, specifically done for Hamilton by Finkelestein et al., (2005) found a strong association between the mortality and the exposure to traffic pollution. According to the research, people living closer to the roads, particularly residents living within 50 meters of the major roads and 100 meters of the highways, have higher risk of mortality. Another study for Hamilton–Burlington area shows that there is a positive relationship between the mortality rate and the level of the pollution (Finkelestein et al., 2003).

2.3 *Sulphur dioxide (SO_2)*

Sulphur dioxide is a colourless gaseous pollutant that smells like burnt matches. During the combustion of coal, sulphur burns in the presence of oxygen and forms sulphur dioxide.

$$S + O_2 = SO_2 \qquad \text{(Eq 2.2)}$$

The oxidation of SO_2 leads to the formation of sulphur trioxide (SO3) and sulphates particulate matter (SO_4^{2-}). SO_3 combines with water to create sulphuric acid (H2SO4), a major contributor to acid rain.

$$SO_3 + H_2O \rightarrow H_2SO_4 \qquad \text{(Eq 2.3)}$$

These sulphates and SO_2 create several health issues including breathing problems, respiratory illness and cardiovascular diseases. Long term exposure to the SO_2 may cause to serious resporitory diseases.

2.4 *Adverse Effects of SO_2*

Durham (1974) found that bronchitis, tonsillitis, colds and sore throats were associated with SO_2 and NO_2. The study showed that high concentrations of SO_2 were associated with reduced respiratory capacity which ultimately became conducive to more severe respiratory diseases. Shy et al. (1973) performed a study on school children aged beween five to ten years in New York City and concluded that exposure to high concentrations of SO_2 resulted in prolonged decrement in ventilatory function. According to the Clean Air Hamiton's (2007) report, 62 deaths out of 714 happened due to the exposure to SO_2 alone.

The pollutants SO_2 and NO_2 significantly contribute to form acid rain. Acid rain has harmful effects on lakes, rivers, forestry, soils and buildings. It also affects the fish and wild life. In addition, these pollutants also form particulate matters that cause severe health complications such as respiratory illness, visibility reductions. Environment Canada reports that this problem is more pronounced in central and eastern Canada.

2.5 *Sources of NO_2 and SO_2 in Hamilton*

Hamilton, which is known as a steel city in Canada, is a hub of two largest heavy metal-producing plants in Canada. These industries produce sixty percent of Canada's steel and stand as one of the top employers in Hamilton, contributing to the economy since 1910 (Schneider, 2006). In addition to these top steel industries, there are other large and medium industries in the region. These industries on one hand act as employment generators, and on the other hand stand as the major industrial pollution sources in the city. Sulphur dioxide is one of the main harmful pollutants for which the heavy metal industries are primarily responsible. Heavy truck traffic for loading and unloading of the materials on the industrial sites is another source of such pollution.

Additionally, several major roads and highways traverse the city, carrying enormous amount of vehicles daily. These vehicles emit significant quantities of pollution into the urban environment. Consequently, the city experiences higher air pollution near the highways and major roads (Wallace and Kanaroglou, 2008). Daily auto commute and frequent truck movements are the leading causes for increased NO_2. Hence, both the industrial and transportation sectors play a significant role in the degradation of the air quality in Hamilton.

It is found that elevated levels of SO_2 and NO_2 are a persistent problem for the city. A study done by Clean Air Hamilton (2005-2006) shows there has been little or no change in the trends of NO_2 along the downtown and the industrial core of the city over the past decade. A comparison of NO_2 trends in Hamilton and in the surrounding urban centers indicated that Hamilton had the least improvement in the reduction of NO_2 (Clean Air Hamilton Progress Report, 2005-2006). Figure 2.1 shows the trend of nitrogen dioxide over the last decade.

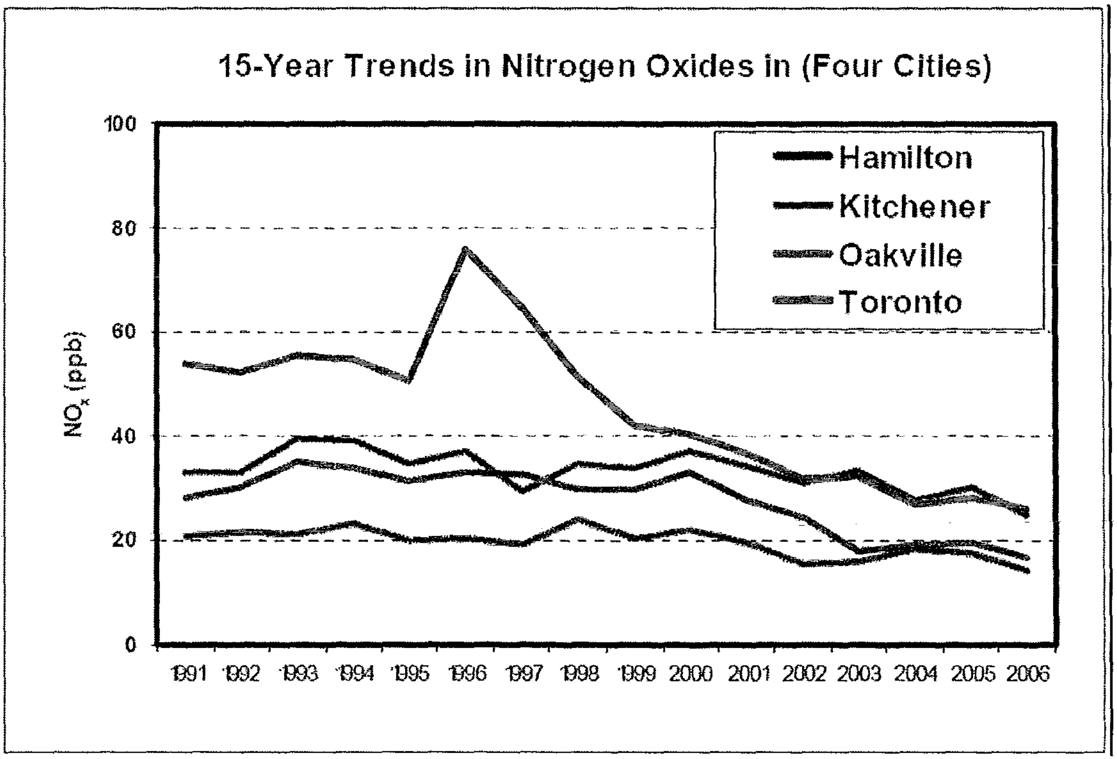

Figure 2.1: Trend of NO_2 over the last decade in the city of Hamilton, Kitchener, Oakville and Toronto

Source: Clean Air Hamilton Progress Report, 2005-2006

Similar results have been documented in the case of SO_2. The level of the pollution of SO_2 compared to other cities in Ontario is very high. Although the city experienced the highest level of pollution of SO_2 during early 90's, level of SO_2 has reduced during late 90's. However, the high concentration of the pollutant is still a major problem for the area. Figure 2.2 presents the distribution of annual average of SO_2 in Hamilton and the cities around Hamilton.

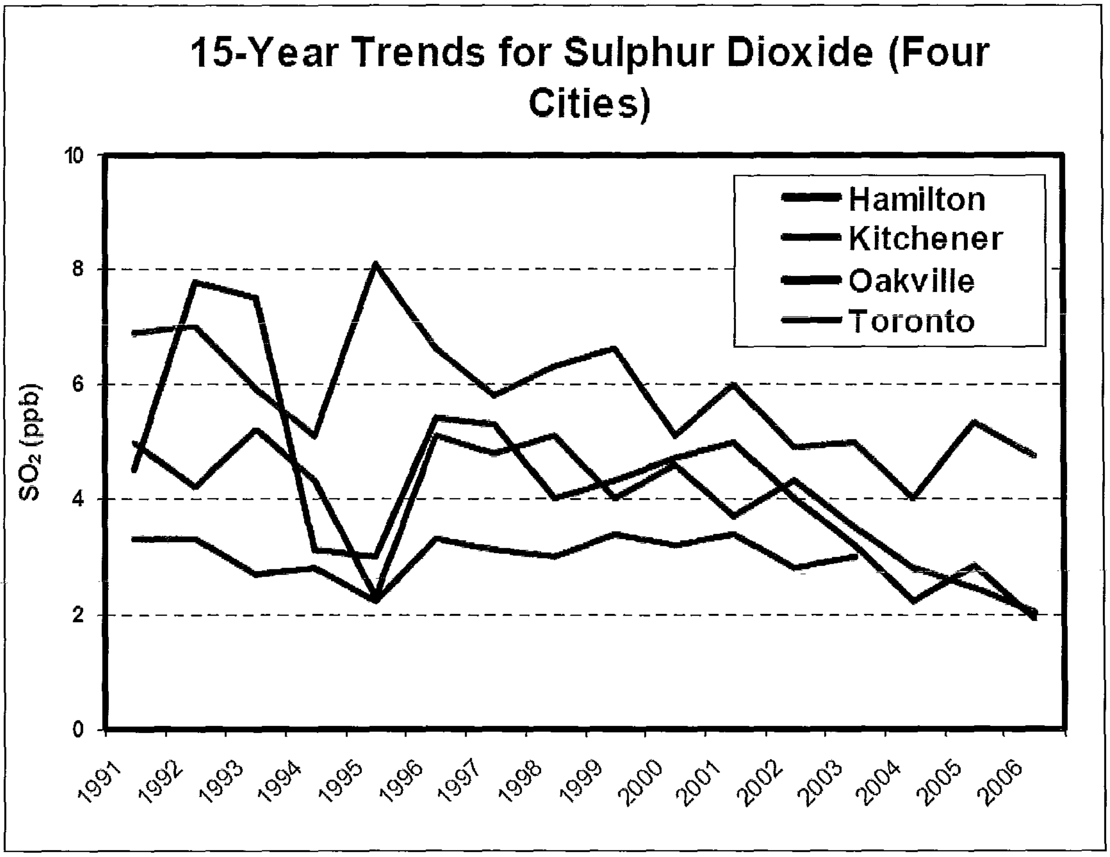

Figure 2.2 : Trend of SO_2 over the last decade in the city of Hamilton, Kitchener, Oakville and Toronto

Source: Clean Air Hamilton Progress Report, 2005-2006

SO_2 is usually found in higher concentrations near the source industries. The city's heavy industrial zone, which includes the two largest steel industries as well as other large and medium industries, is located in the north-east section of the city. This area is therefore exposed to higher industrial emissions. The spatial distribution of SO_2 reveals clusters of high concentrations in close proximity to and immediately downwind of this industrial sector. Concentrations dissipate with increasing distance away from and upwind of this zone. On the other hand, the spatial distribution of NO_2 is quite different as the source is primarily vehicular traffic with secondary contribution from industry. Therefore this study focuses on these two particular pollutants (NO_2 and SO_2), which might help in understanding impacts of both the industrial and transportation sector on urban air pollution, particularly in the Hamilton Census Metropolitan Area.

2.6 *Spatial analysis for air pollution*

Spatial analytical methods involve analyzing geographic data. These methods use statistical techniques to investigate the spatial or geographic component of a particular problem in question. Spatial models help to understand underlying relationships between a phenomenon, and the geographical features and/or their locations.

Air pollution is sometimes treated and analyzed as a temporal problem. Several extensive research projects and practical works have described air pollution from a temporal perspectives. Air pollution, however, has strong spatial ties. For instance, air pollution in built up areas is more acute compared to that in open spaces. Within the built up area, the level of pollution varies from pollutant to pollutant due to different land uses, transportation characteristics and human activities (McGregor, 1996). Therefore, spatial modeling is essential to characterize the problem from a spatial perspective.

2.7 Land use regression modeling

Several authors used land use regression to understand spatial relationships of different pollutants (for example, Briggs et al. 2000, Gilbert et al, 2005, Henderson et al. 2007, Sahsuvaroglu et al. 2006). Land use regression can be defined as "the stochastic, or regression modeling approach, which is based on statistics of land use, traffic counts, and other relevant input data to predict the spatial distribution of the pollutants in a specific region (Briggs et al., 2000; Brauer et al., 2002, 2003; Gilbert et al., 2005). This modeling approach has received a significant amount of attention in air pollution estimation research in the last decade, as they are flexible and produce relatively accurate results. Hence many researchers around the world utilize this modeling technique (Ross et al., 2006). Briggs et al. (1997) first employed this model at the intra-urban scale to model small-scale variation in air pollution, and to map the traffic-related pollutant NO_2 in Amsterdam, Netherlands, Huddersfield, UK and the Prague, Czech Republic (Ross et al. 2006). Later this modeling technique has been applied in numerous European countries, and has drawn significant interest in North America (Sahsuvaroglu et al, 2006; Gilbert et al., 2005; Jerrett et al., 2007).

Briggs et al. (1997) examined a number of factors including traffic volume, land cover (area of built up land) within 300 meters, elevation and the height, access roads of different lengths for different sizes of residential areas, and estimated the distribution of NO_2 in Huddersfield, U.K, Amsterdam, Netherlands and Prague, Czech Republic. The authors used the measurements of the pollutant (i.e. NO_2) from 80 different sites.

On the other hand, Brauer et al. (2003) estimated multivariate regression models to predict the concentration of particulate matter (PM) 2.5. Using traffic related variables the estimated regression models explained 73%, 56% and 50% of the variability in annual average fine particle concentrations for the Netherlands, Munich and Stockholm County, respectively. Hochadel (2006) applied a regression model to estimate the concentrations of NO_2 and PM 2.5 in the westerly end of the Ruhr-area in North-Rhine Westphalia, Germany, and found that NO_2 concentration and PM 2.5 absorbance were strongly correlated with the traffic-based variables (such as daily traffic flow and maximum traffic intensity of buffers with radii from 50 to 1000 meters, distances to main roads and highways).

Briggs et al (2000) tested different methodologies including dispersion modeling, spatial interpolation such as contouring, kriging and trend surface analysis and regression mapping to estimate the distribution of NO_2 in Hammersmith and Ealing, U.K., Northampton, U.K., and Sheffield, U.K. The authors compared the results and finally selected the regression mapping method for its reasonably accurate results (R-square values ranges from 0.8 to o.9).

2.8 LUR model in North America

Gilbert et al. (2005) first applied the land use regression modeling technique in North America to explain the spatial distribution of NO_2. The study conducted in Montreal revealed that the distance to highway, traffic counts on the nearest highway, length of highway, major roads within 100 meters, length of minor roads within 500 meters, open land use within 100 meters, and dwelling density within two kilometers significantly contribute to the concentration of NO_2. The model explained 54% variability in the sample data.

Sahsuvaroglu et al. (2006) utilized the land use regression technique to predict the concentration of NO_2 for the city of Hamilton. The study found that traffic density within 300 meters, distance highway within 50 meters, 1500 meters downwind to Highway 403, open land use within 500 meters, industrial land use within 200 meters, downtown industrial core within 1000 meters, and distance to lake have significant impact on the level of pollution. The study modeled the outcome of observed NO_2 concentrations from 102 monitoring points. They were monitored for a two-week period and average concentration for this period was used as the dependent variable in the land use regression model.

Again, Jerrett et al. (2006) found that industrial land use, highways within 200 meters, major roads within 50 meters, industry within 750 meters, dwelling density within 2000 meters, longitude, locations up to 1500 meters downwind of highways, and traffic density within 500 meters have a profound influence on the NO_2 levels in the city of Toronto.

Moore et al (2007) attempted to predict the concentration of ambient particulate matter for Los Angeles, California using the same approach. They found that traffic within a 300-meter buffer, and percentage of industrial and Government areas have a significant effect on NO_2.

Ross et al. (2006) estimated the similar type of model to analyze the concentration of NO_2 for Southern California. They found road length around 40 meters, traffic volumes and distance to the coast to have a noticeable effect on concentration of NO_2.

Ross et al (2007) also used land use regression analysis to predict the concentrations of PM2.5 for New York City Region. The predictors in the final model specification were traffic volume in the 500 meters buffer, population in the 1000 meters buffer and the industrial land use within a 300 meter buffer.

Rosenlund (2007) estimated a land use regression model to predict the concentration of NO_2 and compared the results with an emission model using information on the types of vehicles circulating, traffic counts and driving patterns. For the regression model the research utilize following characteristics of the area: geographic locations within the circular traffic zones (i.e. main ring road, green strip, inner ring road, and traffic-limited zone), distance from busy streets, distance from parks, distance from sea, size of the census block, population density and altitude. The study concludes that a land use regression model explains a substantial fraction of traffic related air pollution concentration and that inclusion of emission data do not significantly improve the model.

In a recent study, Arain et al. (2007) attempted to improve the predictive capabilities of land use regression models by adding the wind flow effects to display the temporal and spatial distribution of NO_2 along the Toronto-Hamilton urban air shed in Ontario, Canada. The authors used wind fields at different times of day in the models and demonstrated that NO_2 has diurnal and spatial variation in the study area where higher concentration in NO_2 is observed during the morning peak and evening peak.

Henderson et al. (2007) modeled NO, NO_2, and PM 2.5 for Vancouver, Canada using land use regression method. The variables used in the model are length of highways and major roads, vehicle density, commercial and residential land use within 750 meters buffers, industrial land use within 300 meters buffer, population density and elevation. The study found that the method is sensitive to differences in spatial distribution between the pollutants and can be applied in different context. Table 2.1 shows summary of various studies that use land use regression modeling techniques for different pollutants at different places.

Table 2.1 Summary of land use regression methods used in pollution concentration modeling

Description (Study)	Pollutant	Study area (Sample)	Variables
Briggs et al. (1997)	NO_2	Huddersfield, UK (80); Amsterdam, Netherlands (80), Prague, Czech Republic(80), Huddersfield(80)	<u>Amsterdam</u> All access roads within the study area, Classification of road on basis of population served, Distance to nearest road serving >25 000 people, Area of built up land <u>Hudderfields</u> All roads (stored as 10m grid), Mean 18 hour traffic flow (vehicles/hour) for each road segment, Land cover class (20 classes, stored as 10m grid), Altitude

			Prague All traffic routes in study area, Mean daytime traffic flow (vehicles/hour), Land cover class, Altitude
Gilbert et al (2005)	NO_2	Montreal, Canada (67)	Distance from nearest highway, traffic count on nearest highway, length of highways within 100 m, Length of major roads within 100 m, Length of minor highways within 500m, Area of open space within 100 m, population density within 2000m
Ross et al. (2006)	NO_2	San Diego County, Southern California (1300)	Road length-40m, traffic volume within 40-300m, traffic volume within 300-1000m, distance to coast
Jerret et al. (2006)	NO_2	Toronto, Canada (95)	Highways within 200m, major roads within 50m, industry within 750m, dwelling density within 2000m, longitude, downwind of all highways 1500m, traffic density within 500m
Sahsuvaroglu et al. (2006)	NO_2	Hamilton, Canada (102)	Traffic density at 300m, 50 m from highway, 1500m downwind of highway 403, open land use within 500m, industrial land use within 200m, downtown industrial core within 1000m, distance to lake
Ross et al. (2007)	PM 2.5	New York (62)	Traffic around 500m, population around 1000m and industrial land use 300m
Rosenlund et al. (2007)	NO_2	Rome, Italy (68)	Circular traffic zone, distance to busy road: 1) >500m, 150-500m or <150m, size of the census block, inverse population density, and the altitude
Moore et al. (2007)	PM2.5	Los Angelas, CA (23)	Traffic density at 300m, Industrial area within 5000m and government land area within 5000m

2.9 *Summary*

Since NO_2 and SO_2 have significant heath impacts it is important to study distribution of pollutants for urban areas. Most of the studies reviewed in this research shows that the spatial distribution of the pollutants can reasonably be explained using land use regression techniques. Although these studies were conducted for different areas, many of them examine similar types of variables to explain the spatial distribution of the pollution. Notable common types of variables are amount of different types of land uses; proximity to roads, industries and other activity centers; traffic flow, road length etc; and population/dwelling density, elevation etc. However, meteorological factors such as wind direction and wind speed are rarely considered for such kinds of studies. Since level of concentration may vary with respect to wind direction, wind speed etc., it is important to test hypothesis in relation to these meteorological elements. Since most existing literature ignores effects of wind direction and speed, it is possible that estimated values of pollution might have some biases. Therefore this research attempts to extend land use regression modeling approach by considering wind-related effects. It also tests hypotheses in relation to land use, transportation attributes and other locational factors in the context of Hamilton Census Metropolitan Area.

Again, although many of these studies use the observed concentration level of the pollutants for regression application, mostly they are averaged over an extended period of time (for example Henderson et al., 2007; Wheeler et al., 2008, Ross et al., 2007 among others), which basically obscure the temporal dynamics in the models. However, this study uses mobile field-monitored pollution measurement data that provides instantaneous measurement of pollution at a given point in time in a given location. It provides a unique dataset to test hypotheses with respect to temporarily varying variables such as wind speed, wind direction etc.

Finally, review of existing literature suggests that investigation of SO_2 in relation to spatial characteristics is rare. Since SO_2 pollution is a persistent problem in the Hamilton

Metropolitan Area, this research attempts to model concentration of SO_2 applying land use regression techniques. The next chapter discusses the research design and methodology used in this research.

3. CHAPTER THREE: RESEARCH DESIGN AND METHODS

As revealed by the literature review it was observed that the previous studies focus on the pollution data that shows the pollution concentration over a specific period of time whereas this research deals with the pollution concentration data that have been collected at a specific point in time, and estimates the spatial relationship between the pollution concentration and various urban characteristics. This chapter outlines the geographical settings of the study area, pollution data collection and preparation, basic concepts of the land use regression (LUR) modeling, LUR modeling procedure, and selection of various explanatory variables, including land uses, transportation and locational attributes etc. The research focuses on two pollutants: NO_2 and SO_2. One is mainly a traffic emission and the other is predominantly an industrial emission. The pollutants have different emission sources, possess different formation processes and show different spatial patterns in the study area. Hence, different model specifications were developed and tested to relate concentration of the pollutants with potential sources, land use characteristics, locational factors and meteorological conditions.

3.1 The study area

The study area of this research is Hamilton Census Metropolitan Area (CMA), which is situated in the southern central Ontario at the west-end of the Lake Ontario. Hamilton Harbor is the western tip of the lake that forms the northern boundary of the CMA. The Niagara Escarpment (approximately 100-120 meter high) divides the city into two parts. The segment which is located above the escarpment is known as the upper Hamilton or the Mountain area, and other portion that remains below the escarpment is generally termed as lower city that contains dowtown area.

The geography of the Hamilton itself plays a key role for the growth of the city and economic activities. The city is located at the focal point of the Greater Golden Horseshoe Area (GGHA) as well as within the Windsor-Quebec City Corridor, which is a major USA-Canada trade route. The city is bounded by two busy highways, Highway 403 and Queen Elizabeth Way (QEW). The northwest part of the city is linked with the Highway 403 that connects Toronto, and eastern part of the city is linked with QEW, one of the oldest highways in North America that carries over 200,000 average daily trips (Solaiman, 2007). Current transportation routes and economic activities as well as being a trade hub make Hamilton one of the fastest growing cities in Canada.

Figure 3.1: Geographic settings of the study area (Hamilton CMA)

Source: Google map, 2008

In addition, Hamilton is often known as a "steel city" of Canada, which has many medium and large steel and other industries, including two largest North American steel mills – Stelco and Dofasco (see Figure 3.2).

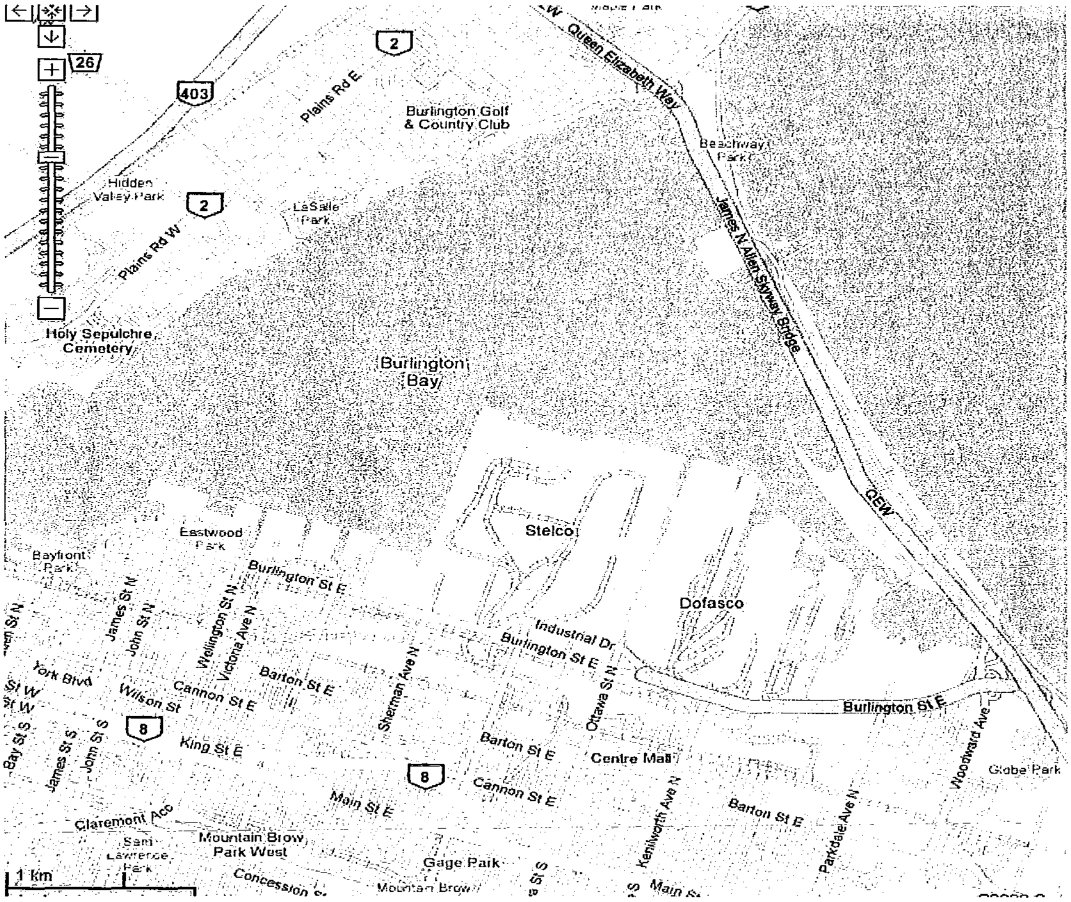

Figure 3.2 Stelco and Dofaso- the major steel industries in Hamilton

Source: Google map, 2008

3.2 *Pollution data collection*

The study employs a primary data set collected through a mobile field monitoring system operated by Rotek International. The survey used a truck (which was equipped with several pollution monitors) that traveled through different streets of the Hamilton

CMA to collect the pollution data. Measurements were taken at 2118 different locations across the city. The locations were mostly near the industrial areas, along the highways and at the intersections of major roads and highways. Measurements of the pollutant were also taken from the areas located near train lines and the train stations across the city. A GPS device was used to record the exact location of the monitored observation point. The data were mostly collected from 10.00 AM to 4.00 PM. They were collected in the months of November 2005, December 2005, January 2006 and March 2006.

3.3 *Pollution data preparation*

At the beginning of the study, the data were divided into four wind quadrants beginning from the north and moving clockwise. The quadrants are as follows: 0 to 89 degrees, 90 to 179 degrees, 180 to 269 degrees and 270 to 359 degrees. It should be noted that the quadrants represents the direction of the wind from which it is blowing. As such 0 to 89, 90 to 179, 180 to 269 and 270 to 359 degrees indicate the winds are blowing from north-east (N-E), south-east (S-E), south-west (S-W) and north-west (N-W), respectively. Table 3.1 presents summary statistics of the pollution data. In addition, while Figure 3.3, 3.4, 3.5 and 3.6 show locations of the pollution measurement points for NO_2, Figure 3.7, 3.8, 3.9 and 3.10 exhibit measurement locations for SO_2 for four wind directions.

Table 3.1 Summary statistics of the pollution measurements in the Hamilton CMA

Pollutant	Wind Direction	North-east	North-west	South-east	South-west
NO_2	Number of observations	384	454	230	861
	Min	2	2	3	2
	Mean	28.57	20.91	28.61	23.59
	Max	169	138	128	233
	Standard Deviation	20.7	17.1	18.21	19.21
SO_2	Number of observations	370	447	232	957
	Min	1.4	1	2	2
	Mean	16.27	11.33	17.74	16.27
	Max	123	62	40	109
	Standard Deviation	14.22	10.89	7.8	12.98

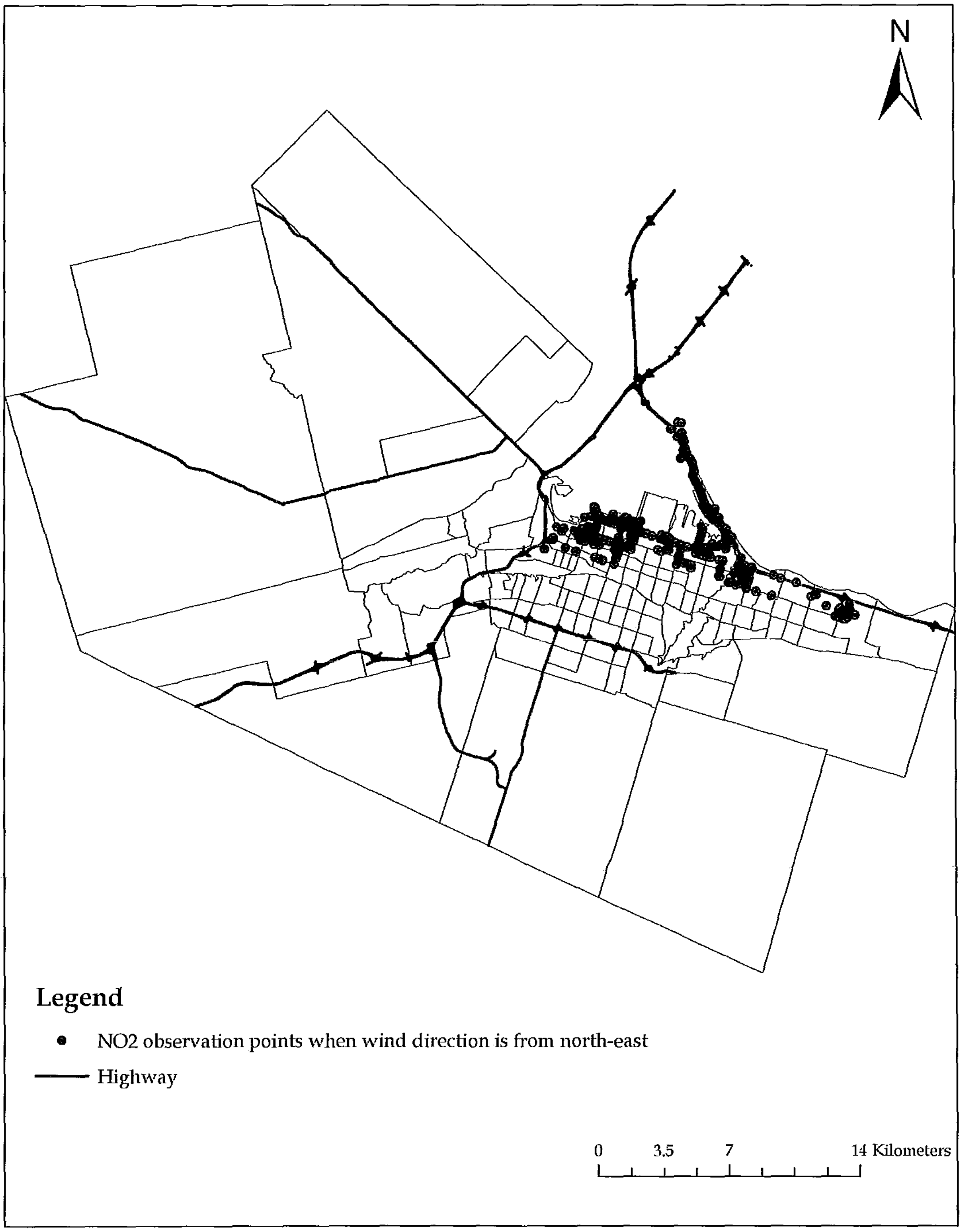

Figure 3.3 Locations of the NO_2 measurement points for north-east wind direction

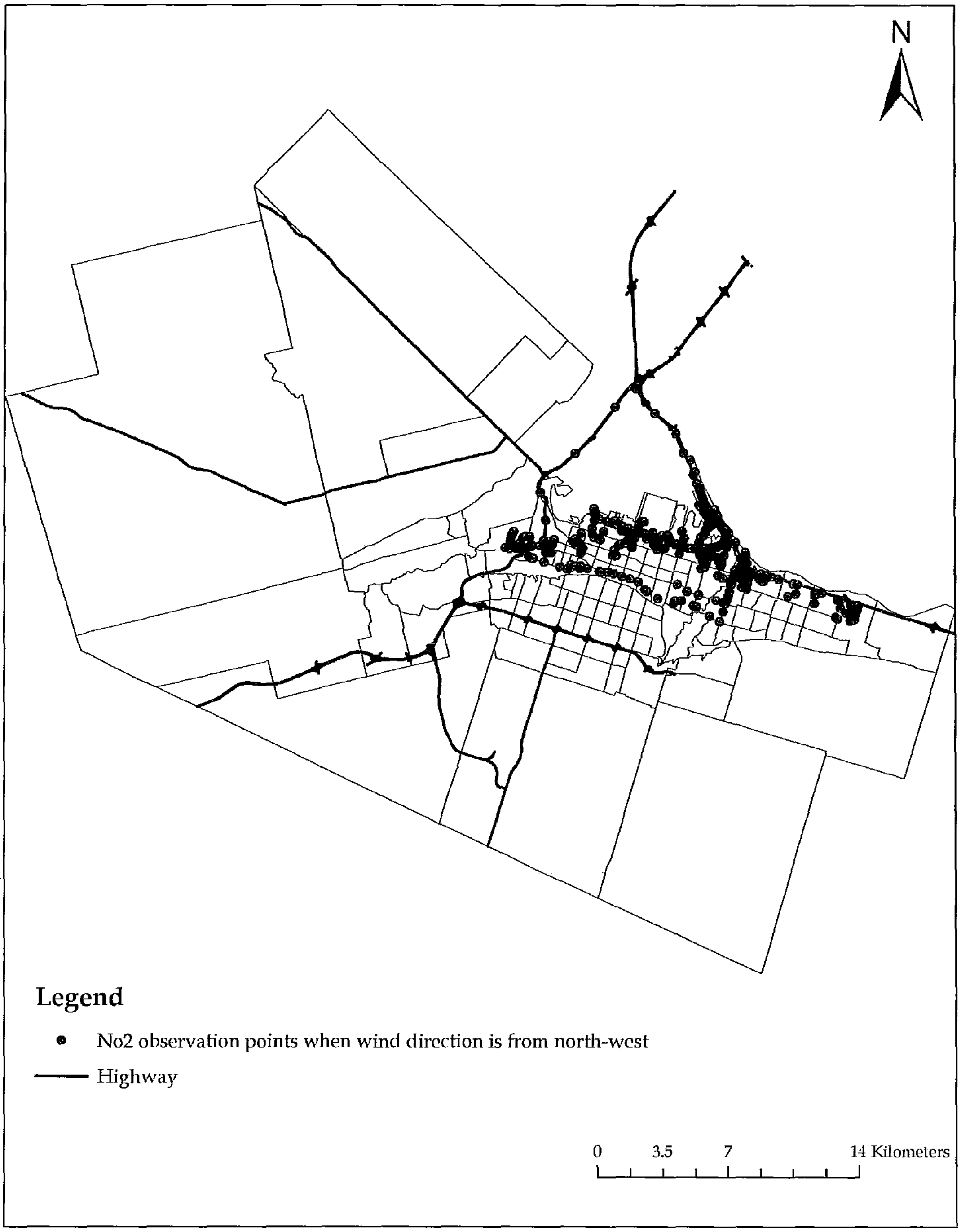

Figure 3.4 Locations of the NO_2 measurement points for north-west wind direction

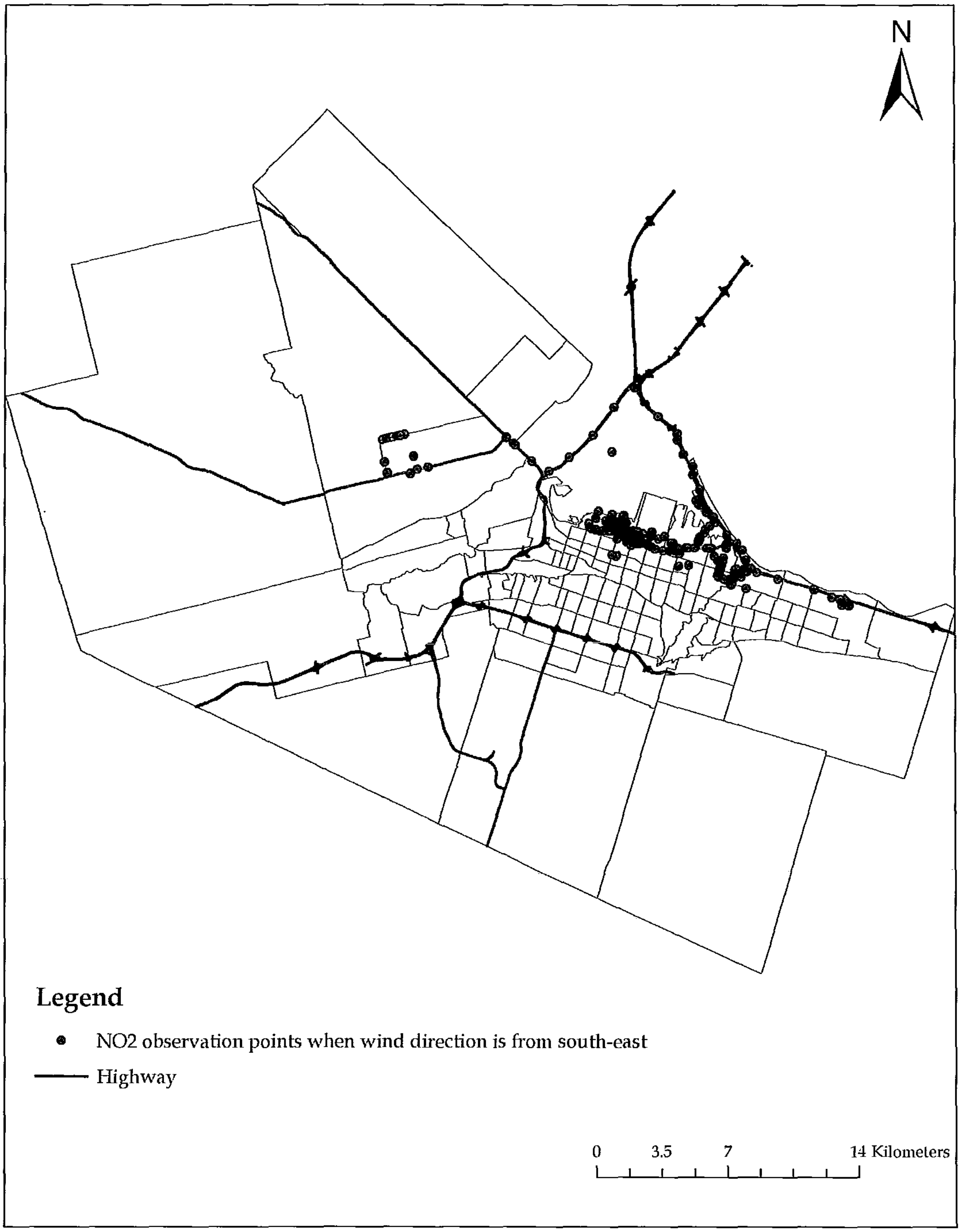

Figure 3.5 Locations of the NO_2 measurement points for south-east wind direction

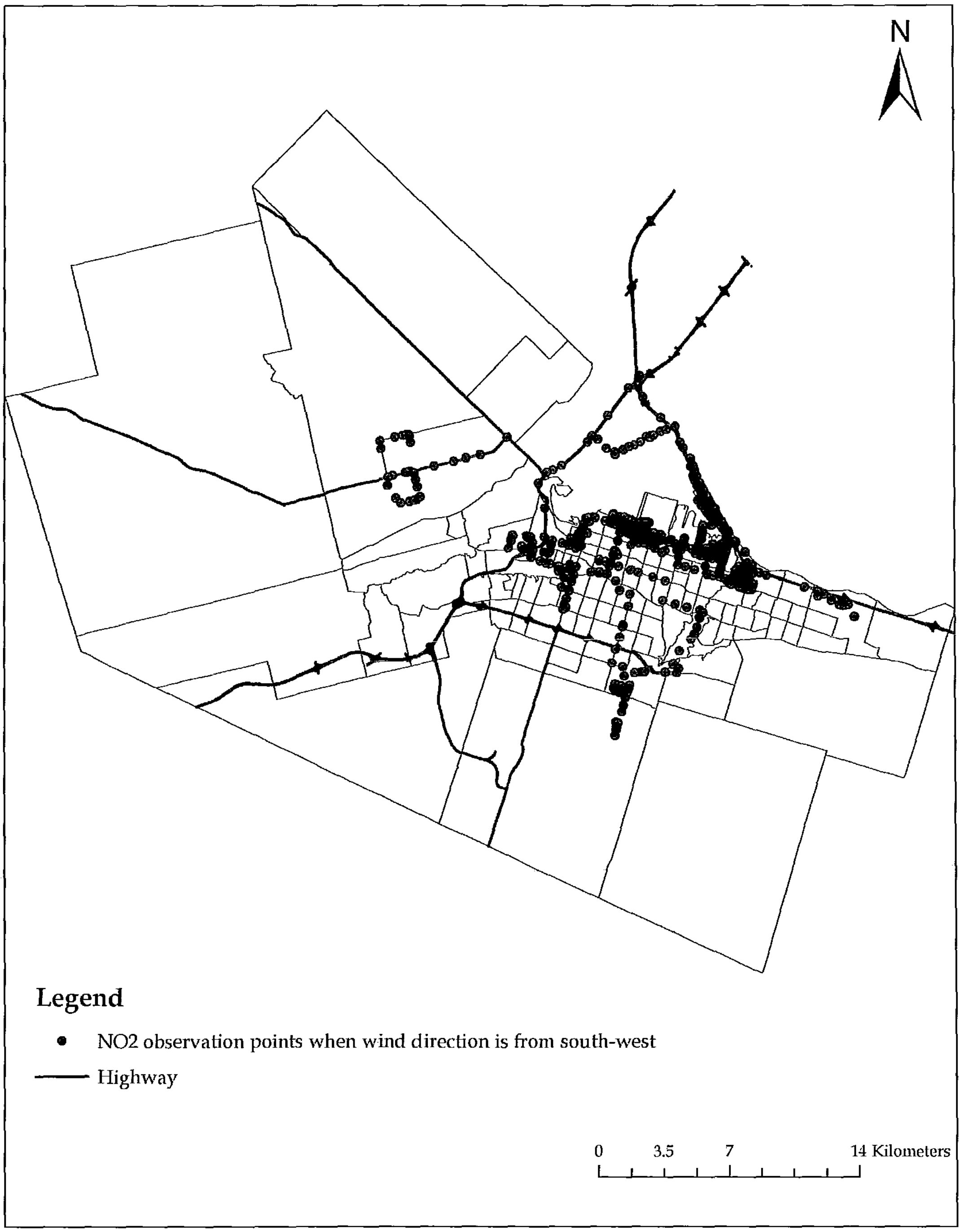

Figure 3.6 Locations of the NO_2 measurement points for south-west wind direction

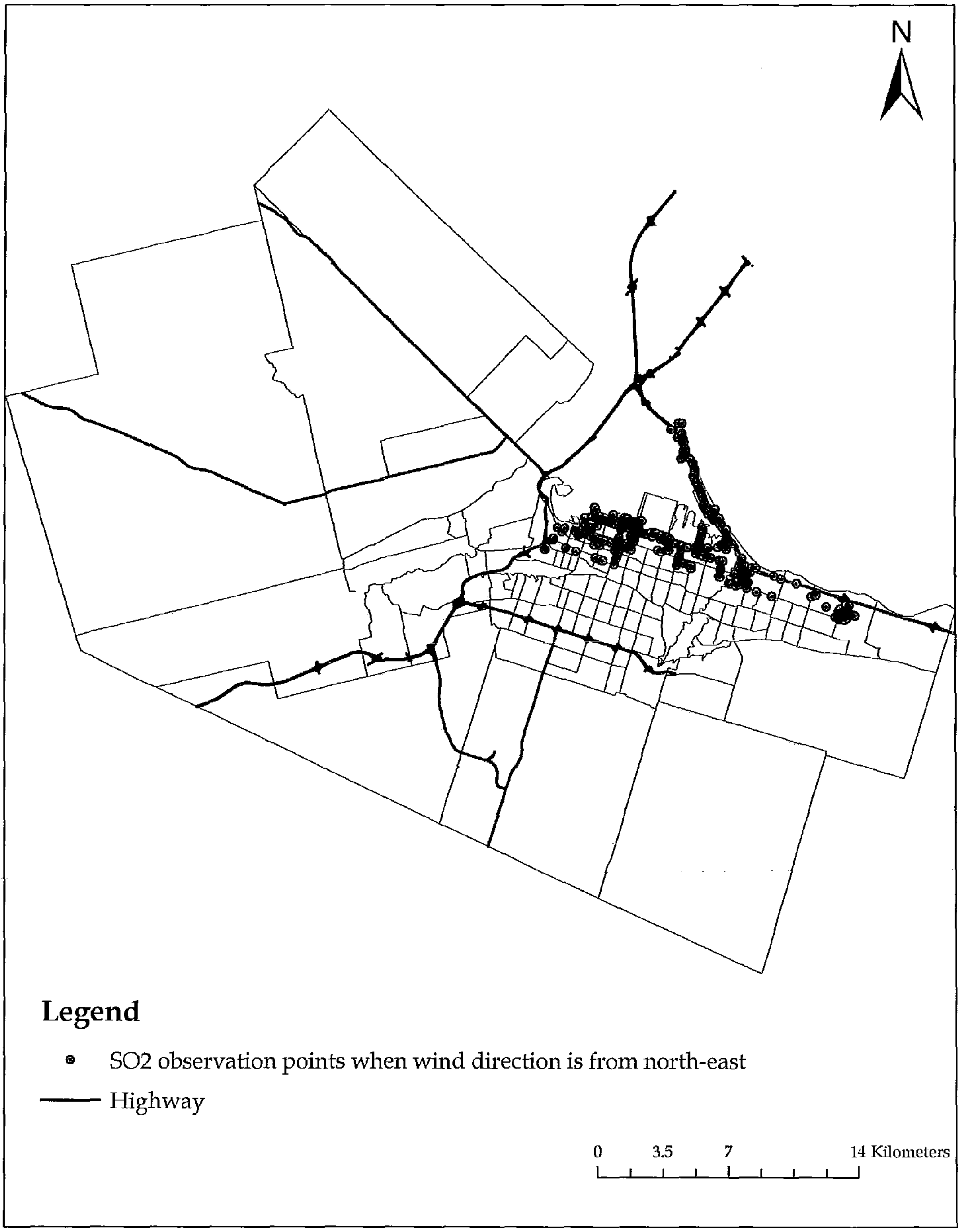

Figure 3.7 Locations of the SO_2 measurement points for north-east wind direction

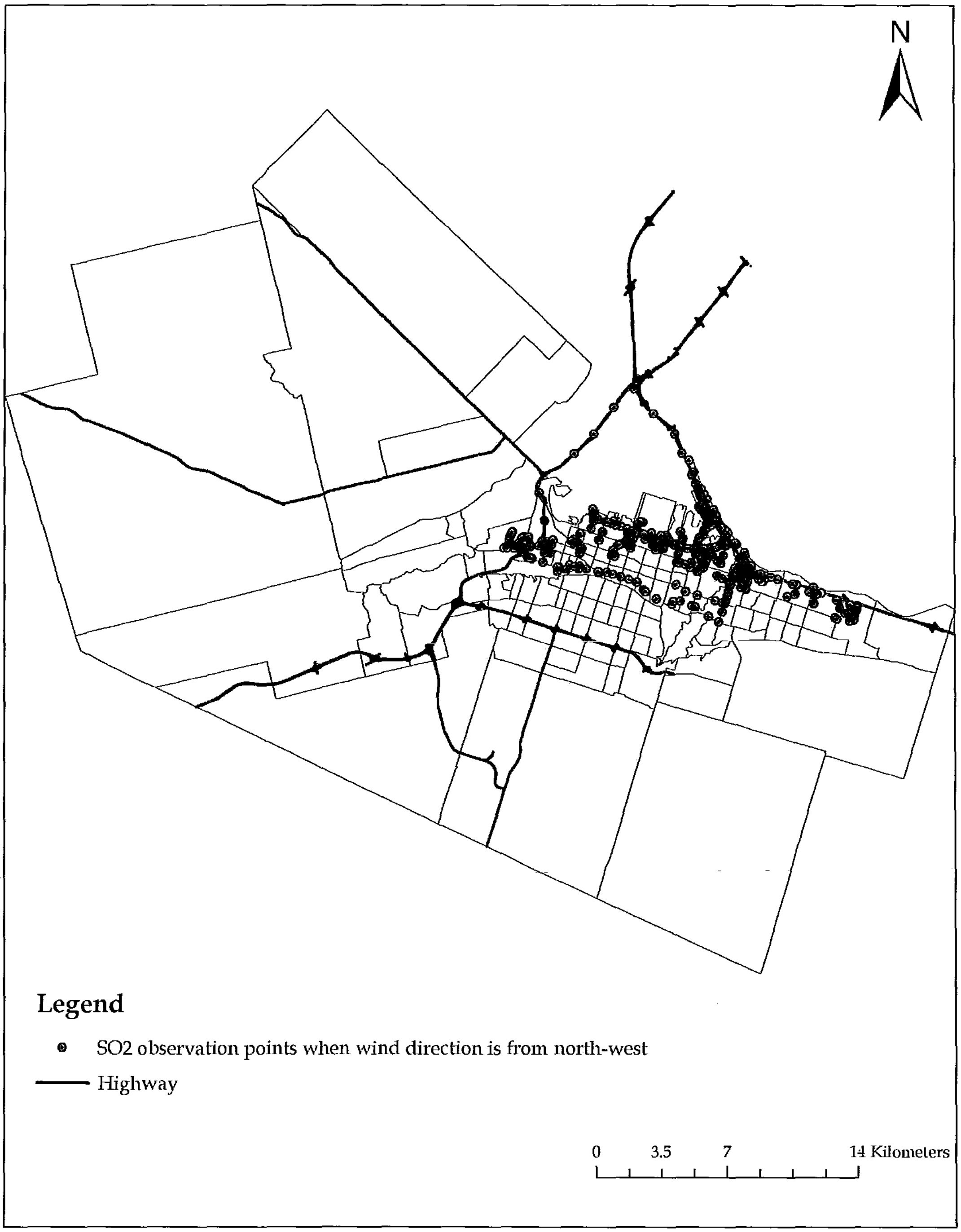

Figure 3.8 Locations of the SO_2 measurement points for north-west wind direction

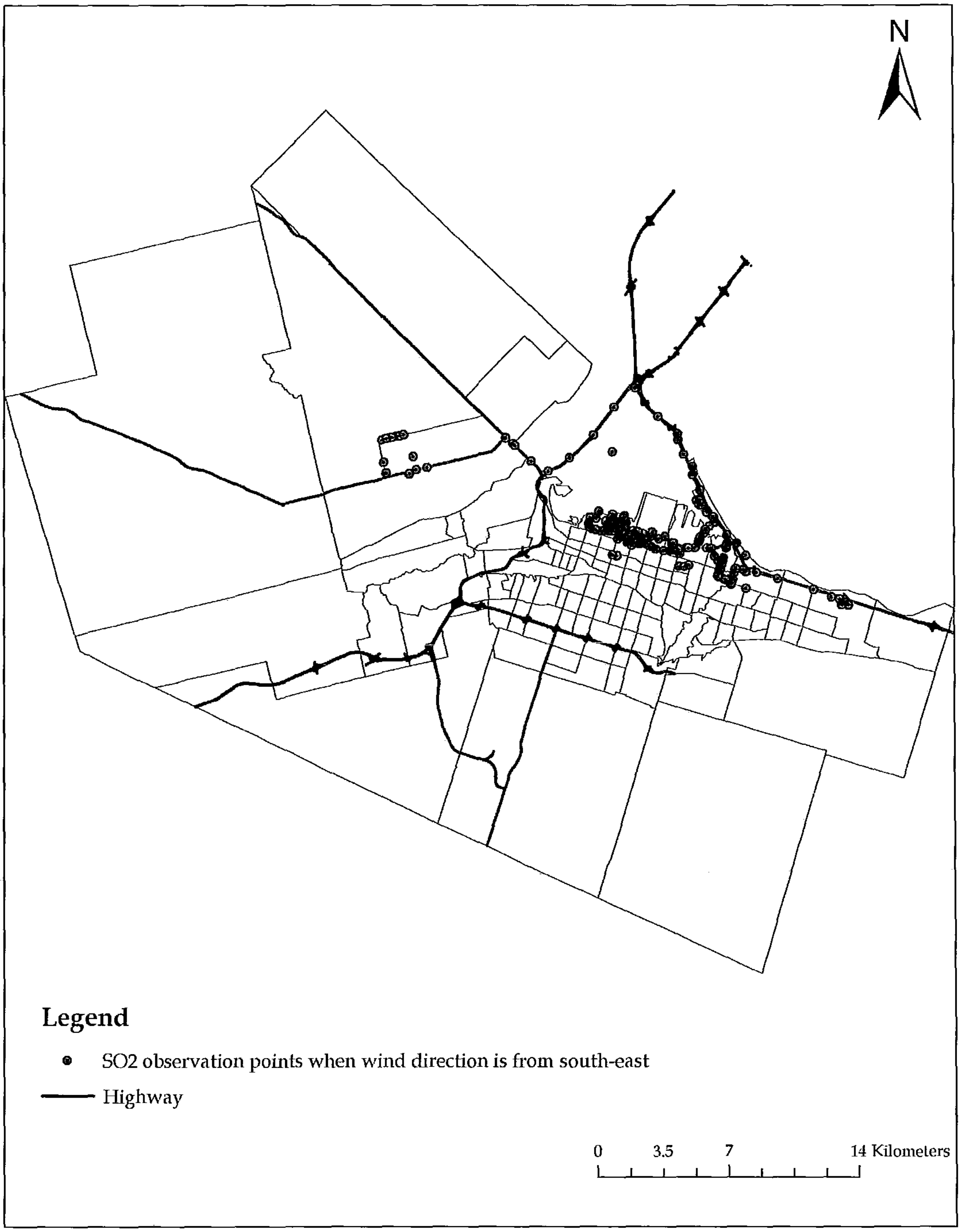

Figure 3.9 Locations of the SO_2 measurement points for south-east wind direction

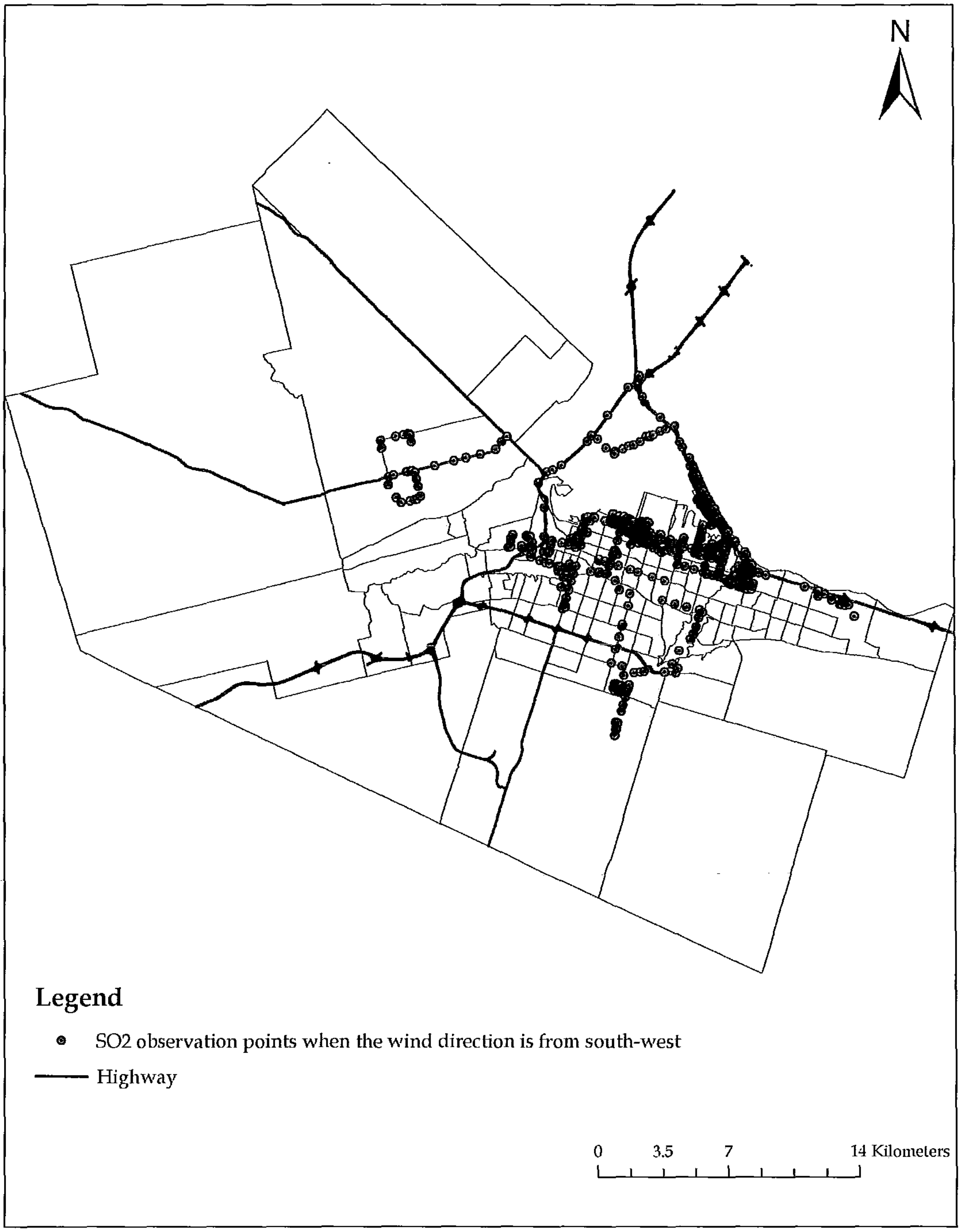

Figure 3.10 Locations of the SO_2 measurement points for south-west wind direction

3.4 Land use regression modeling method

According to Briggs et al., 1997, the land use regression model (LUR) is defined as "the stochastic, or regression modeling approach, which is based on statistics of land use, traffic counts, and other relevant input data to predict the spatial distribution in a specific region". It follows the same principles of regression modeling techniques.

The regression model is one of the most well-known and widely used statistical modeling techniques. The model is simple in its form, and explores the relationship between a dependent variable and independent variables. In a simple univariate linear regression model (i.e. the dependent variable is explained by one independent variable) takes the following form:

$$y = a + bx + \varepsilon \qquad \text{(Eq. 3.1)}$$

where y is the response variable, x is the explanatory variable and a is a constant (intercept). On the other hand, b is the slope of the regression line, which is called the regression coefficient or the parameter of the explanatory variable. It represents the average amount by which the dependent variable increases or decreases when the independent variable increases by one unit and other independents are held constant. ε is the random error term, which represents unexplained portion of the model.

In a multiple regression model, the response variable is explained by two or more variables. Similar to many applications of regression analysis that involve more than one explanatory variable, land use regression (LUR) model utilizes multiple regression method that takes the following form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots\dots\dots + \beta_k x_k + \varepsilon \qquad \text{(Eq. 3.2)}$$

Here y is the dependent variable, which represents the measurements of pollution concentration for this study. The variables, x_1, x_2 and x_k represent the land use, transportation, locational and other factors used in the study. Lastly, β_0 denotes the constant, and β_1, β_2 and β_k are the parameters to be estimated, which shows nature and magnitude of the relationship of the explanatory variables.

Commonly Ordinary Least Square (OLS) method is used for the parameter estimation. It can be shown that linear regression generates best possible estimate of linear model parameters. However, the model specification should make sense from a behavioral or theoretical point of view as well as from a practical perspective. That means the variables included should have a causal influence on the dependent variable. This study used LIMDEP 8.0 statistical package (Greene, 2002) to estimate the regression models.

Finally, the goodness of fit of the model is evaluated in terms of R^2 statistics, which is defined as:

R^2 = [Explained variance/Total variance]=[1 - (Unexplained variance/Total variance)]

$$=1-\frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \qquad \text{(Eq. 3.3)}$$

where, y_i is an actual observation of the dependent variable ($i = 1, 2, \ldots\ldots n,$), $\hat{y}_i$ is the corresponding fitted value from the regression model and $\bar{y}_i$ is the average value of observations. The term, $\sum_i (y_i - \hat{y}_i)^2$ is the error sum of squares and $\sum_i (y_i - \bar{y}_i)^2$ is the total variance to be explained. R^2 values lie between 0 and 1 in value, the closer to 1 in value the R^2 is, the better the model's fit.

3.5 LUR modeling procedure

3.5.1 Data pre-processing

- Four databases for each pollutant were created, one for each wind direction

- Land use, transportation, meteorology and other locational variables were derived using GIS techniques (details in section 3.6).

- Since frequency distribution plots of NO_2 concentration for four separate wind directions are right skewed (Figure 0.1, 0.2, 0.3 and 0.4 in Appendix A), the dependent variables (i.e. concentration level of NO_2) were transformed using a natural logarithm transformation. Figure 0.5, 0.6, 0.7 and 0.8 show the distribution of the data after natural log transformation (Appendix A).

- Similarly, since the frequency distributions of SO_2 concentration for four wind directions are right skewed (Figure 0.9, 0.10, 0.11 and 0.12 in Appendix A), the dependent variables were log-transformed for SO_2 model development (Figure 0.13, 0.14, 0.15 and 0.16 in Appendix A)

- Extensive exploratory analysis and investigation of different sources of pollution were performed to formulate hypotheses for the LUR modeling application.

- Potential variables were identified and created, including interaction variables such as downwind to the source of pollution (for example, industry location). Table 1 in Appendix A shows a comprehensive list of variables tested in this research.

3.5.2 Model development

- A manual forward selection method was employed to identify the significant independent variables that describe the spatial relationship with the concentration of NO_2 and SO_2.

- Each of the explanatory variables was first tested through bi-variate or simple univariate regression analysis.

- Variables were added stepwise to test different hypotheses in the model.

- Numerous model specifications were used to select a final model with better model fit. The best fit was determined by the R^2 value.

- The residuals from the final regression models were analyzed using a frequency histogram as an approximate check of normality

- Diagnosis tests for spatial autocorrelation were performed using GEODA 9.5i (Anselin et al., 2004) in terms of Moran's *I* statistics

3.6 *Selection of explanatory variables*

3.6.1 Land use variables

Since land use generates various activities that contribute to air pollution, the study takes into account land use variables explicitly to build the relationship between the pollution concentration and land uses.

Land use data were collected from Desktop Mapping Technologies Inc. (DMTI). The available land use data were categorized into seven broad categories: residential, commercial, government and institutional, resource and industrial, water body, open area and parks and recreational. Concentric circle buffers of different radii (e.g. 50, 100, 200, 300, and 350 meters) were created around the observation points and overlaid on the land use map to calculate the area of different land uses within the buffer. The study used GIS operations of ArcGIS® 9.2 to perform this process.

3.6.2 Transportation-related variables

Road network data were collected from DMTI containing information on highways, major roads and local roads, train line etc. All transportation related variables were created from this street network file by using ArcGIS® 9.2. Proximity to highways, major roads, major intersections, train lines, rail station etc. were calculated using GIS techniques. Several dummy variables were created to test different hypotheses during model development. Figure 3.11 shows major roads and highways in the Hamilton Metropolitan Area.

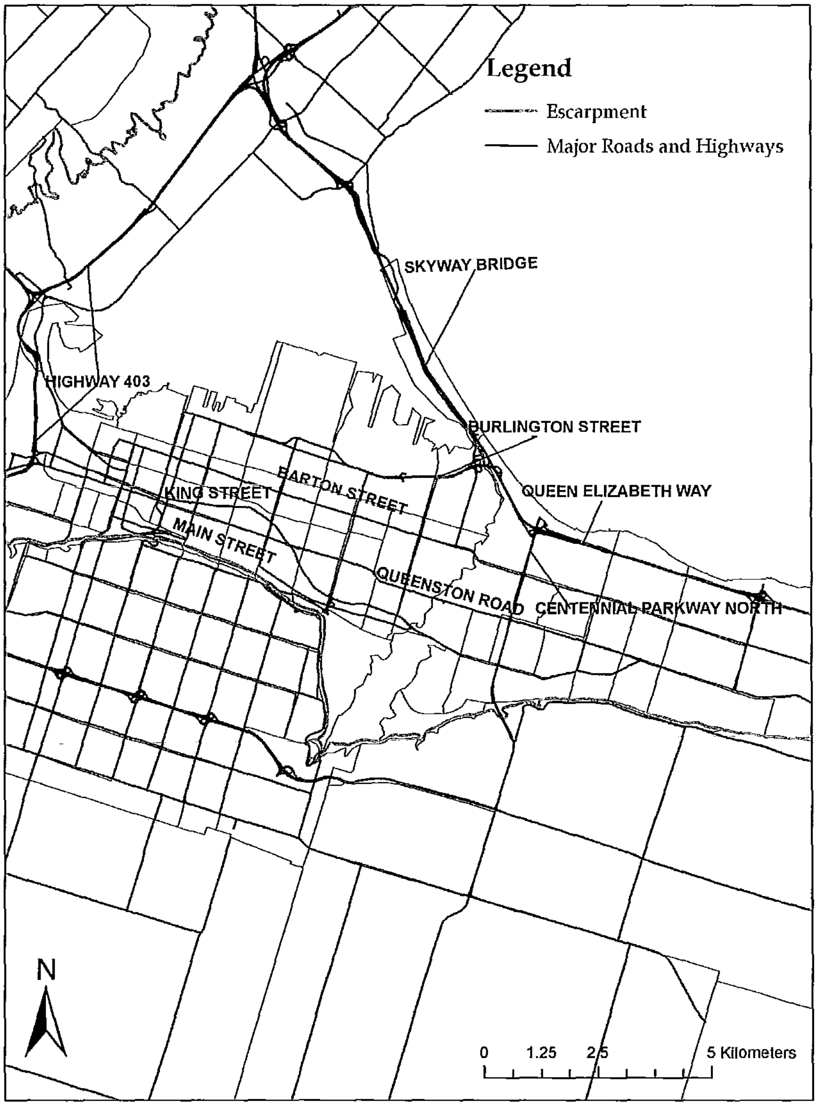

Figure 3.11 Major roads and highways in the Hamilton CMA

Additionally, buffers of radii 50, 100, 200, 300, 350, 500, 750 and 1000 meters were created around the observation points. These buffers were then overlaid on the road network to calculate total length of different types of roads within the specified areas.

In addition, the Integrated Model of Urban Land use and Transportation for Environmental Analysis (IMULATE) was used to simulate link-by-link traffic volume in the Hamilton CMA for the year 2006. Details of the model are documented elsewhere (Anderson et al., 1996; Maoh and Behan, 2006). Circular buffers of different radii (50m, 100m, 200m, 300m, 350m, 750m and 1000m) were overlaid on the road network to obtain the traffic volume around the observation points.

3.6.3 Other locational variables

Several other location-specific variables were created, including distances to the major parks (for example, the Bay Front Park) and different emission sources (such as Stelco and Dofasco). Figure 3.11 shows the major steel industries in the Hamilton CMA. Since industries are a major source of SO_2, several point source-specific variables were considered for the investigation (for example, proximity to lakeshore industries and other metal-producing industries).

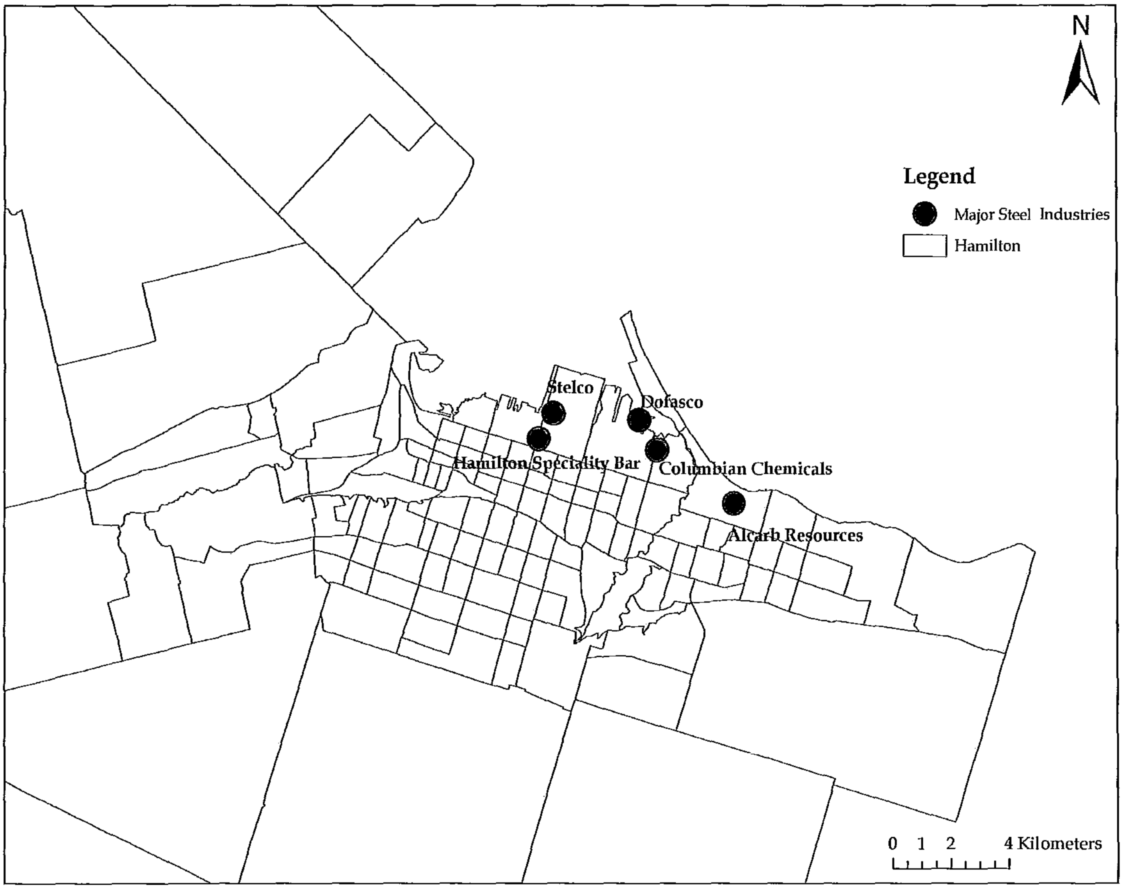

Figure 3.12 Major steel industries in the Hamilton CMA

In addition, since this study explicitly considered wind direction, several variables are created to identify downwind location from various industry and key activity locations. It also considered distance bands within different locations. Therefore several interaction variables are created to investigate during model development (for example, downwind from lakeshore industrial core and located within 1.5 kilometers; downwind and located within 2 km of the steel industry near city center etc.).

Some other variables were also constructed to represent different commercial activities. Such variables include proximity to Hamilton Central Business District (CBD), regional shopping centers, other local services, and parking lots.

3.6.4 Demographic variables

Demographic and other neighborhood characteristics were collected from the 2001 Census tabulations of Statistics Canada, which were the latest available data at the time when this study was conducted. The census attribute data were then joined to the census tract geographic boundary GIS layer. Pollution observation points were spatially joined to the census tracts, and variables such as population density, dwelling density etc. were tested during model development.

3.6.5 Meteorological variables

Meteorological data were collected from a weather station (located at the intersection of Niagara Street and Land Street) that includes wind direction measurements and wind speed for each observation point. As discussed earlier, wind directions were divided into four quadrants that begins from the north and moves clockwise. The quadrants are: 0 to 89 degrees (north-east), 90 to 179 degrees (south-east), 180 to 269 degrees (south-west) and 270 to 359 degrees (north-west). Other hourly meteorological data such as temperature, solar radiation and relative humidity data were taken from the Environment Canada to test different hypotheses during model development.

3.7 Diagnosis tests for spatial autocorrelation

While the land use regression model is useful in explaining pollution concentration with respect to land use and other characteristics, some portion of the concentration variability may remain unexplained, which might violate standard assumptions of the classical regression theory (Goodchild, 1993). For example, the unexplained portion of the regression model might be interdependent with a systematic pattern among observations in geographic space attributable to their relative location. Therefore, this study performed a spatial autocorrelation test for the residuals of the estimated model. It used GEODA 9.5i (Anselin et al., 2004) to estimate Moran's I statistics, which takes the following form:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{s^2 \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}} \qquad \text{(Eq. 3.4)}$$

where x_i and x_j are the value at point i and j respectively. n is the number of points and s^2 is the variance of x values with a mean of $\bar{x}$. w_{ij} is the weight for measuring spatial autocorrelation.

Moran's I has a value between -1 and 1. The expected value of Moran's I is $(-1)/(n-1)$, where n is the number of observations. If the Moran's I is equal to the expected value (which approximate to zero for large n), no spatial autocorrelation is present. On the other hand, if Moran's $I > (-1)/(n-1)$, positive spatial autocorrelation exist; and if Moran's $I < (-1)/(n-1)$, negative spatial autocorrelation is present. However, the magnitude of the Moran's I does not indicate significance, which is determined by a p value.

3.8 Generation of pollution surfaces

Finally, the estimated models for both pollutants were used to generate pollution surfaces for the Hamilton CMA. First, the study area was divided into 50m X 50m grids. Then, separate surfaces were created for each variable. Lastly, The surfaces were multiplied by the coefficient values obtained from the estimated LUR models using the raster calculator function in ArcGIS® 9.2 to get the pollution surfaces for NO_2 and SO_2.

4. CHAPTER FOUR: LAND USE REGRESSION MODELS OF NITROGEN DIOXIDE

The primary task of the land use regression model is to characterize a spatial relation between the concentration of NO_2 and the different urban factors - land use, transportation, major activity centers, and meteorological factors. As discussed in the previous section, four models were estimated to predict the concentration of nitrogen dioxide (NO_2) in the Hamilton Census Metropolitan Area (CMA). The major wind directions considered in the study were north-east, north-west, south-east and south-west.

4.1 *North-East model*

The north-east model was estimated using 384 observation points for which measurements of NO_2 were taken when the direction of wind was north-east. The final model specification has six explanatory variables that explain 23.5% variability in the estimated model. The variable definition is shown in the Table 4.1.

Table 4.1 Variable definition for NE model

Name of the variables	Description of the variable
WS567	Wind speed
DTIND	Distance from lakeshore industries
MJRD	Distance from major roads and highways
PARK100	Park area within 100 meters from the observation points
TVOL300	Traffic volume within 300 meters from the observation points
RAILDN	1 if the observation points are downwind and within a 500 meter buffer from the rail station, 0 otherwise

The model used explanatory variables that are not critically correlated with each other. The correlation matrix and the variables are shown below in Table 4.2.

Table 4.2 Correlation matrix NE model

	LNMNO2	WS567	DTIND	MJRD	RAILDN	PARK100
LNMNO2	1.0000	-0.2887	-0.3172	-0.1851	0.0813	-0.1124
WS567	-0.2887	1.0000	0.0470	0.1094	0.0561	0.1161
DTIND	-0.3172	0.0470	1.0000	0.0082	-0.0274	0.0034
MJRD	-0.1851	0.1094	0.0082	1.0000	0.2718	-0.1273
RAILDN	0.0813	0.0561	-0.0274	0.2718	1.0000	0.0088
PARK100	-0.1124	0.1161	0.0034	-0.1273	0.0088	1.0000

4.1.1 Variable analysis

4.1.1.1 Wind speed:

The first variable used in the model represents the meteorological attribute, wind speed. It was assumed that higher wind speed would reduce the pollution level. Since the high wind speed accelerates the mixing and dispersion, concentration of the pollutant reduces. The model presents a similar result, a negative association between the wind speed and the concentration of NO_2, which means as the wind speed increases the concentration decreases. The results comply with the assumption made for the variable.

4.1.1.2 Distance from Lakeshore industries:

The model introduced another variable that captures the relationship between the concentration of NO_2 and heavy metal producing industries. The distance between the observation points and the lakeshore industrial area was calculated. It was presumed that as the distance from the industries increases the concentration should decrease. This is because high- temperature fuel combustion in industrial and utility boiler is one of the major man-made sources of NO_2 emissions (Muzio and Quartucy, 1998). In the model results it has been found that higher NO_2 values are located near the lakeshore industrial core, which consists of heavy steel and metal industries. In other words, it can be said that locations near to the industrial areas experience a higher level of NO_2 concentration. Therefore, the variable that calculates the distance from the observation points to the industrial downtown core is negatively related with the concentration of NO_2. The coefficient value of the variable is -0.00009 and is significant at a 99% confidence level.

4.1.1.3 Distance from major roads and highways:

To shed some light on the transportation characteristics and their effect on the concentration of NO_2, a variable that measures and describes the distance from the observation points to the nearest major road was created. It was hypothesized that as the distance from the major road increases, the pollution concentration should decrease. This is because the major roads accommodate high volumes of vehicular traffic, which is the principal source for NO_2 (Ross et al., 2006). This transportation variable represents a negative relationship with the concentration of NO_2. It is found that as the distance from the major roads increases the concentration of NO_2 decreases. The result is confirmed at the 99% confidence level.

4.1.1.4 Traffic volume within a 300 meters buffer from the observation points:

As NO_2 is the marker of traffic emission, a variable that represents the traffic volume within 300 meters of the observation points was introduced. The *a priori* assumption for the variable is that locations, which are within 300 meters buffer of the observation points and experience high traffic volume are exposed to higher levels of NO_2. The model confirms the hypothesis showing the expected sign of a positive association between traffic volumes and NO_2 concentration. This means that increased traffic volume nearby contributes to the rise in the concentration level of NO_2.

4.1.1.5 Downwind to Hamilton rail station:

It was expected that locations which are within 500 meters of the Hamilton rail station experience higher concentration of NO_2. Trains in the Hamilton region are generally used for freight transportation. When they stop at the train station they usually do not turn off the engine and NO_2 emission continues at that location. Therefore a dummy variable representing observation points that are downwind and within 500 meters from the rail station is created and tested during model development. The model results show a positive relationship between the variable and the level of NO_2, which means that locations within 500 meters and downwind to the Hamilton rail station are exposed to higher level of pollution.

4.1.1.6 Park area within a 100 meter buffer from the observation points:

The variable, park100 in the model represents the amount of park area within a 100 meter buffer of the observation points. It was hypothesized that as the parks are not a source of NO_2 emissions, the values of NO_2 close to the park area will be lower. The model results comply with the expected assumption and exhibit that the variable park100 is negatively associated with the response variable. Hence, the higher the portion of the park area within a 100 meter buffer area of the monitoring points, the lower the NO_2 pollution concentration.

Table 4.3 NE model results

Variable	Coefficient	T stat	P Value	VIF
CONSTANT	3.7416900	36.46	0.0000	
WS567	-0.0559291	-4.50	0.0000	1.1
DTIND	-0.0000937	-6.50	0.0000	1.01
MJRD	-0.0006594	-3.63	0.0000	1.17
PARK100	-0.0000241	-2.96	0.0030	1.08
TVOL300	0.0000002	3.03	0.0030	1.24
RAILDN	0.3727846	3.55	0.0000	1.11
Adjusted R2	0.2356			

Finally, all the variables included in the model have the expected sign for the Beta coefficient with significant *t*-statistic values. The variables are significant at least at a 5% level of significance. Final model results are shown in Table 4.3.

4.1.2 Residual analysis

Once the model was finalized the residuals were tested to see if any significant spatial autocorrelation existed within the residuals. One of the major assumptions of OLS method is that the residuals should be randomly distributed over the space and the residuals should not take any specific pattern. The diagnosis test for the residuals was performed using GEODA 0.9.5i (Anselin et al., 2004). The Moran's I statistics is 0.0035 with a p value of 0.4. Since this statistic is not statistically significant it can be concluded that there is no spatial autocorrelation among the residuals (see Figure 0.21 for Moran's I scatter plot in Appendix B). Moreover, the distribution of the residuals shows that mean value of the residuals is close to zero (Figure 0.17 in Appendix B).

4.1.3 Surface generation

The final regression model was used to prepare a concentration of NO_2 surface for the Hamilton CMA. Figure 4.1 shows the distribution of NO_2 in the Hamilton CMA for north-east wind direction, and Figure 4.2 presents a detail map of some highly polluted areas according to the estimated model.

Since NO_2 is basically a traffic emission, higher concentrations are usually found near the major roads and highways. As such, highest concentrations (20-54 ppb) of NO_2 are observed around the Skyway Bridge and central Hamilton. High values of pollution (20-54 ppb) are seen at the intersection point of QEW and Centennial Parkway North as well. The Burlington Street and Industrial Drive area also show higher NO_2 levels. This is due to industrial emissions from the lakeshore industries. Due to a downwind location from lakeshore industries, the areas north of Barton Street (see Figure 3.11), specifically around downtown, also experience higher pollution levels.

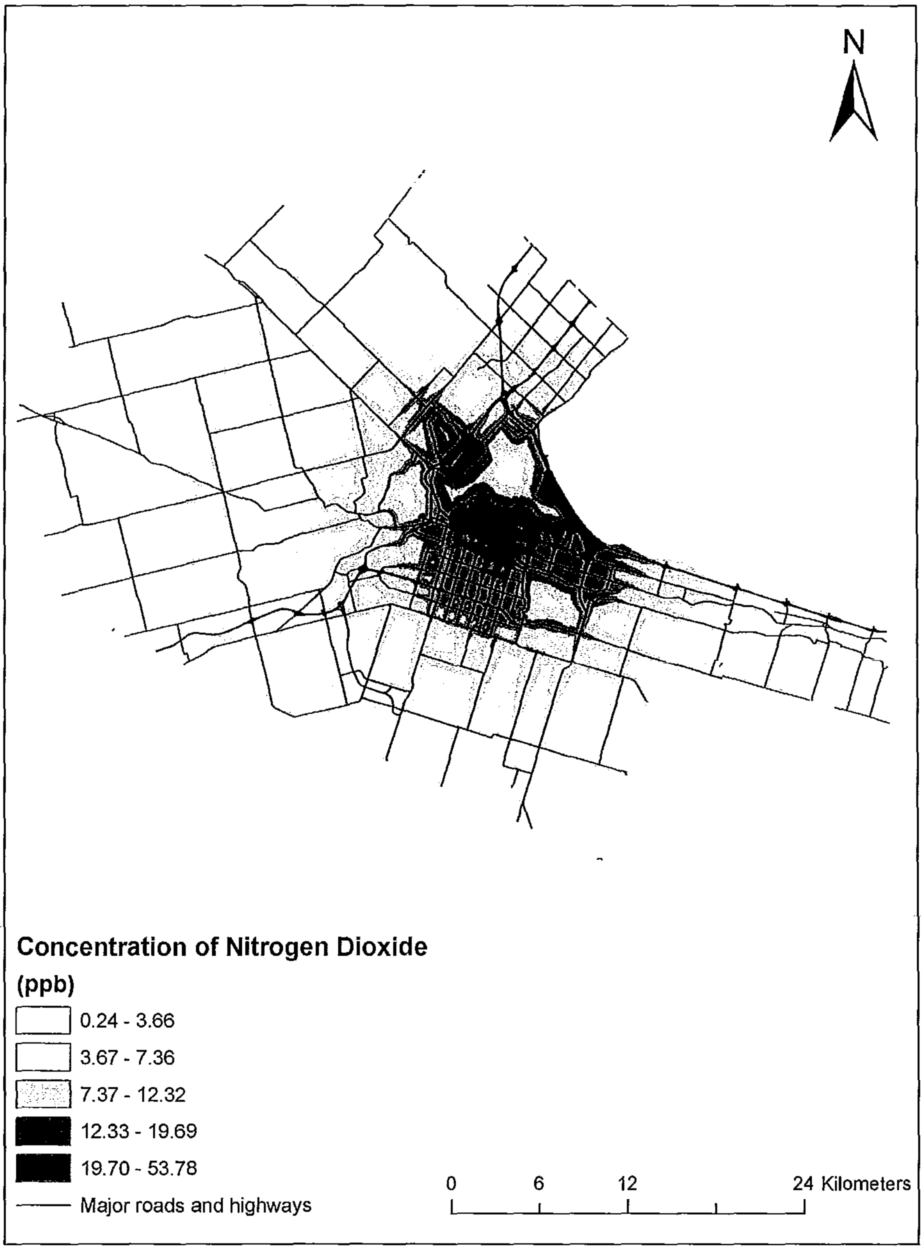

Figure 4.1 Nitrogen Dioxide distribution in the Hamilton CMA for NE wind direction

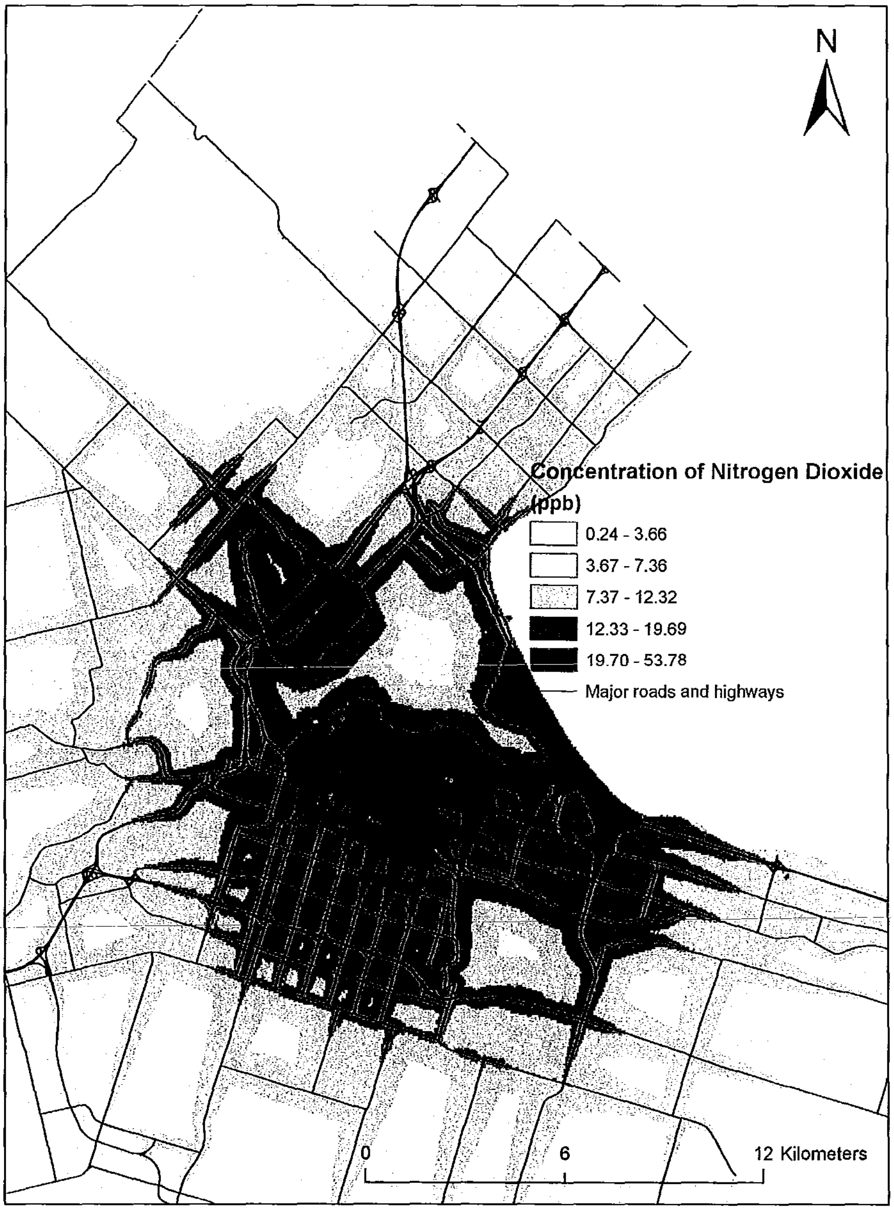

Figure 4.2 Nitrogen Dioxide distribution around downtown area (corresponds to observation points) for NE wind direction

4.2 North-West model

In total, 454 observations were used to estimate north-east wind direction model of NO_2. The observed level of pollution has a minimum value of 2 ppb and maximum value of 138 ppb with a mean 20.6 ppb. The significant variables used to estimate the model are described below in the Table 4.4 and their correlation matrix is shown in Table 4.5.

Table 4.4 Variable definition for NW model

Name of the variables	Description of the variable
WS567	Wind speed
COM350	Amount of commercial area within 350 meters of the observation points
HIWY	Distance from highways
DTIND	Distance from lakeshore industries
RDL50	Length of road within 50 meters of the observation points
QEWCEN	1 if the observation points are within 1 kilometer from the intersection of QEW and Centennial Parkway North, 0 otherwise

Table 4.5 Correlation matrix NW model

	LNMNO2	WS567	COM350	RDL50	HIWY	DTIND	QEWCEN
LNMNO2	1.0000	-0.4040	0.1559	0.1639	-0.0811	0.0021	0.1764
WS567	-0.4040	1.0000	-0.0922	-0.0262	-0.0101	-0.4120	-0.1746
COM350	0.1559	-0.0922	1.0000	0.0627	0.1799	0.0760	-0.0657
RDL50	0.1639	-0.0262	0.0627	1.0000	-0.2111	-0.1264	-0.1066
HIWY	-0.0811	-0.0101	0.1799	-0.2111	1.0000	-0.1333	-0.1552
DTIND	0.0021	-0.4120	0.0760	-0.1264	-0.1333	1.0000	0.1120
QEWCEN	0.1764	-0.1746	-0.0657	-0.1066	-0.1552	0.1120	1.0000

4.2.1 Variable analysis

4.2.1.1 Wind speed:

As high wind speed facilitates the dispersion, it was assumed that the higher the wind speed, the lower the concentration of NO_2. The model results suggest that there is a negative correlation between wind speed and the concentration of NO_2. It shows that high wind speed reduces the pollution level.

4.2.1.2 Commercial area within a 350 meter buffer from the observation points:

The variable defines the commercial area within a 350 meters buffer of the monitoring points. It is expected that larger parcels of commercial area incur higher probabilities of vehicle movement. Moreover, designated commercial areas are usually located near roads, which again is an NO_2 emission-contributing factor. Commercial buildings also represent the built up area and attracts numerous activities. The results suggest that the higher the commercial area the higher the concentration of the pollutant as expected.

4.2.1.3 Road length within a 50 meter buffer from the observation points:

The variable measures the length of major roads and highways within circular buffered areas of 50 meter radius. The larger the total lengths of the roads around an observation point, the higher the possibility of traffic movements in that area. Therefore, it was hypothesized that high concentrations of NO_2 exist if the length of road is larger around the observation points. It is found that when the total length of the road increases, the concentration of NO_2 rises.

4.2.1.4 Distance from highways:

The study used another transportation variable that calculates the distance from observation points to the nearest highway. Since highways experience higher frequency of motor vehicle movement, it is expected to get a higher concentration of NO_2 closer to highways. The estimated land use regression model indicates a negative association with the distance from highway and concentration of NO_2. The result is confirmed at a 95% confidence level.

4.2.1.5 Distance from lakeshore industries:

DTIND, the variable for calculating the distance between the observation points and lakeshore industries is another continuous variable, which proved highly significant for the model. As the steel industries are significant source of NO_2, the assumption for the variable was that observation points close to lakeshore industries possess higher concentrations of NO_2. The model results show a similar result, determining a negative relation between distance from the lakeshore and the level of NO_2. It can be said that as the distance increases from the lakeshore industries the pollution decreases. This variable is significant at a 99%confidence level.

4.2.1.6 Dummy indicating 1 kilometer from the intersection of QEW and Centennial Parkway North:

This variable takes account of the points, which are within 1 kilometer of the intersection of two major highways, QEW and Centennial Parkway North. This major intersection point provides a connection between the city and Stoney Creek, Grimsby, and the Niagara Region. Heavy traffic flow is the common characteristic of that area. As the vehicular traffic is a major source of NO_2, it was assumed that the concentration of NO_2 is high close to the intersection. Therefore, a categorical variable was created to consider the points within 1 kilometer of the intersection point. The variable assigns the value 1 if sample points are situated within 1 kilometer of the intersection and 0 otherwise. The model estimation shows that locations within a 1 kilometer buffer from the intersection of QEW and Centennial Parkway North experience higher concentration of NO_2. This variable is confirmed by a 95% confidence level.

The estimated model for the NW wind direction explained 24.37% variability. Table 4.6 shows detail model results.

Table 4.6 NW model results

Variable	Coefficient	t stat	P Value	VIF
Cons	3.43769	30.95	0.0000	
ws567	-0.03813	-9.85	0.0000	1.26
Com350	0.00001	3.54	0.0000	1.06
hiwy	-0.00007	-2.21	0.0280	1.17
dtind	-0.00006	-4.51	0.0000	1.29
rdl50	0.00081	2.52	0.0120	1.13
qewcen	0.41770	2.99	0.0030	1.09
Adjusted R2	0.2437			

4.2.2 Residual analysis

After estimating the model the residuals were tested to determine if they satisfy the assumptions of no significant spatial autocorrelation within the residuals. Moran's *I* statistics is found to be 0.0312 with a *p* value of 0.15. The statistics is not statistically significant with a reasonable level of significant (see Figure 0.22 for scatter plot in Appendix B). In addition, the distribution of the residuals shows that the mean value of the residuals is close to zero (Figure 0.18 in Appendix B).

4.2.3 Surface generation

As discussed and explained earlier, a NO_2 concentration pollution surface is generated on a 50m X 50m grid. The surface shows the distribution of NO_2 within the Hamilton CMA when the wind blows from the north-west (Figure 4.3, detail map in 4.4). Higher pollution is mostly found near the highways and down wind locations of lakeshore industries. It has been observed that 20-57 ppb NO_2 concentration prevails around the exit point of Highway 403 and King Street West. The downtown part as well as north of Barton Street is also exposed to a higher (11-19 ppb) concentration of NO_2. Specifically, Burlington Street East and Industrial Drive exhibit elevated values in level of NO_2. The intersection point of QEW and Centennial Parkway North and Queenston Road and Centennial Parkway North experience the highest (20-57 ppb) concentration of NO_2. This is because the areas are close to major highways, downwind to lakeshore industries and also very close a commercial area, including East Gate Shopping Mall. All these activities invite a large amount of traffic causing rise to NO_2 .

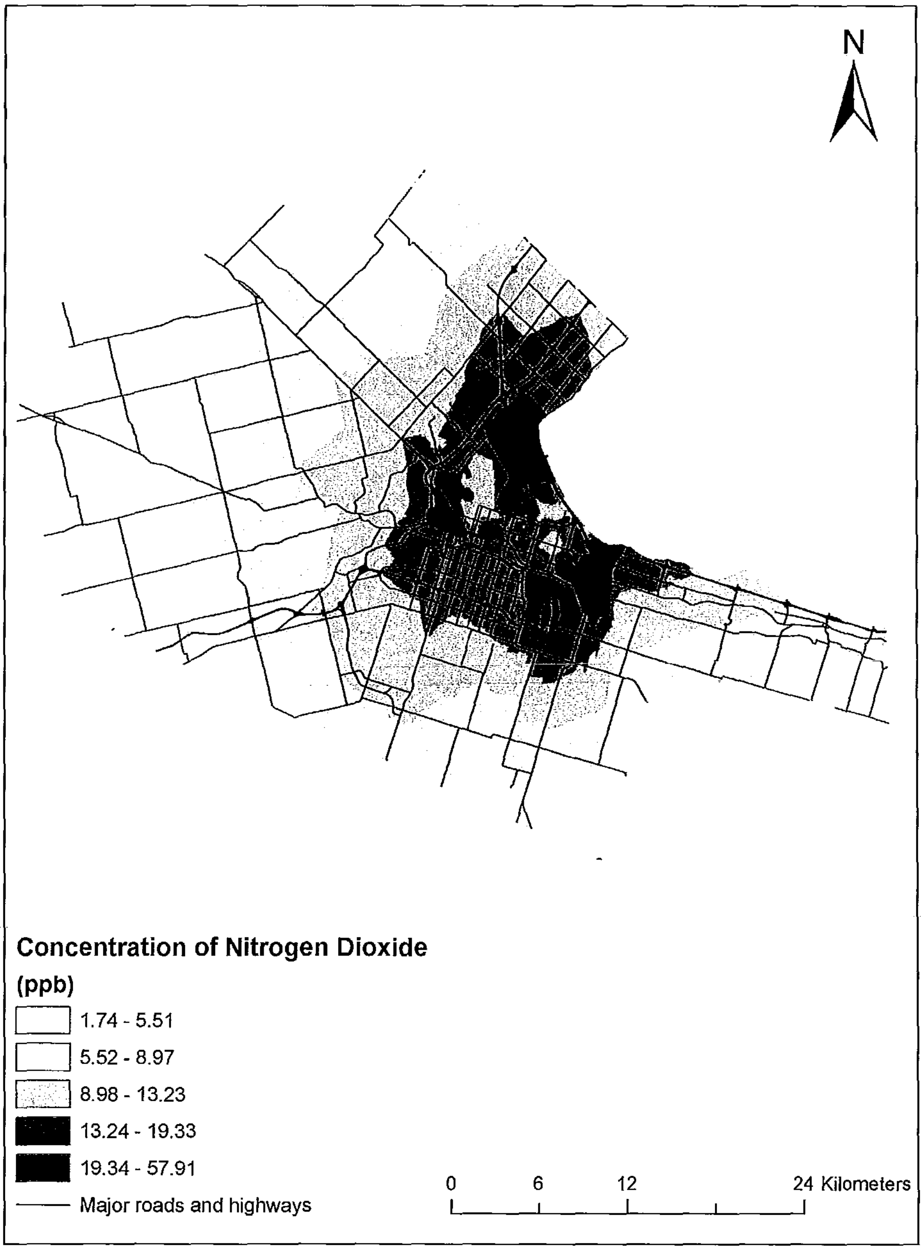

Figure 4.3 Nitrogen Dioxide distribution in the Hamilton CMA for NW wind direction

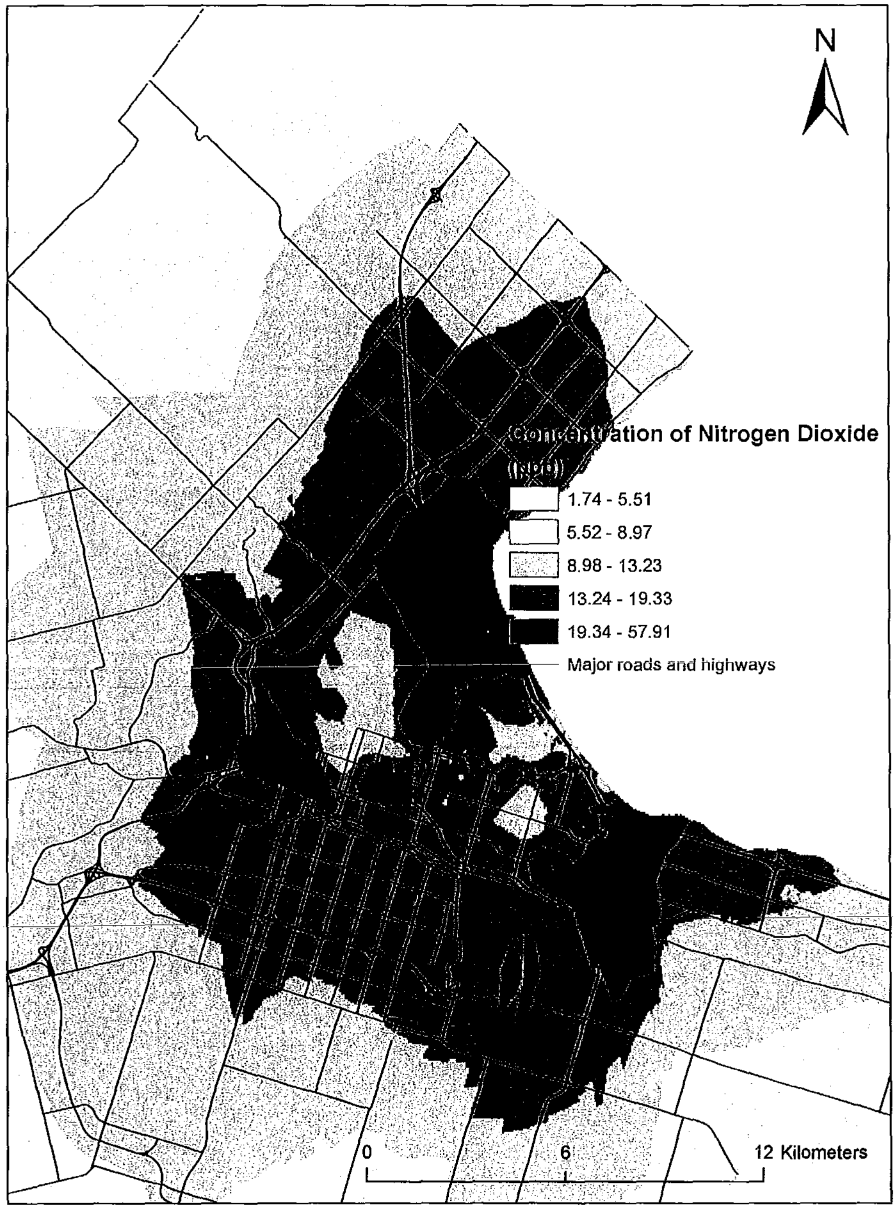

Figure 4.4 Nitrogen Dioxide distribution around downtown area (corresponds to observation points) for NW wind direction

4.3 *South-East model*

This portion of the research estimates the concentration of NO_2 when the wind direction is from south-east. The model was estimated with 230 observation points. The land use regression model explained 26.96% variability in the concentration of NO_2. Table 4.7 and 4.8 present the description of the variables used for modeling and the correlation matrix respectively.

Table 4.7 Variable definition for SE model

Name of the variables	Description of the variable
GOV100	Area of governmental area within 100 meter from the observation points
H403	Distance from highway 403
TRNLN	Distance from train line
POPDENSI	Population Density of the census tract
RDL50	Length of road within 50 meter from the observation point

Correlation matrix of the explanatory variables was also observed to identify if there is any critical correlation among the predictors or not. Table 4.8 shows that there is no significant correlation among the explanatory variables.

Table 4.8 Correlation matrix SE model

	LNMNO2	GOV100	H403	POPDENSI	TRNLN	RDL50
LNMNO2	1.0000	-0.1208	-0.1009	0.1952	-0.4104	0.2636
GOV100	-0.1208	1.0000	-0.2462	0.0623	-0.1051	-0.0936
H403	-0.1009	-0.2462	1.0000	-0.1558	0.1181	0.1318
POPDENSI	0.1952	0.0623	-0.1558	1.0000	-0.0702	0.0410
TRNLN	-0.4104	-0.1051	0.1181	-0.0702	1.0000	-0.0482
RDL50	0.2636	0.0936	0.1318	0.0410	-0.0482	1.0000

4.3.1 Variable analysis:

4.3.1.1 Governmental and institutional land use within a 100 meter buffer of the observation points:

A land use variable was introduced into the model to figure out the relation between the governmental land use and concentration of the pollutant NO_2. The variable computes the total area of governmental land use within 100 meter radius circular buffered area of the observation point. Governmental and institutional land use itself is not a source of NO_2. But this type of land use is considered as an activity center hence frequent vehicle movement and traffic circulation during peak hours are well expected from the mentioned land use. As the data shows non rush hour condition it was hypothesized that as the portion of governmental area gets larger, concentration of NO_2 reduces. Moreover the variable represents the built up environment of the urban area, which again helps to accelerate the concentration of the pollutant. The built up areas interrupt free flow of wind hence wind speed reduces and increases the concentration of the pollutant. The analysis and results confirm the previously made assumption and shows a negative association between the concentration of NO_2 and governmental land use within 100 buffers of the observation point. The variable is well significant above a 95% confidence level.

4.3.1.2 Population density:

Population density is another variable that was tested to see if any relation existed between the population density and the concentration of NO_2. The variable measures the population density of the census tracts dividing the total population by the total area. Population density indicates the amount of residential area. Higher population density means larger residential area and higher portion of built environment. As discussed earlier the built environment again helps to accelerate the concentration of pollutant. Hence, it was expected that the higher the population density the heavier the concentration of NO_2. The model results provide a similar positive relationship between the population density and concentration of NO_2.

4.3.1.3 Proximity to highway 403:

The study used a transportation variable that calculates the distance from observation points to highway 403. It is a major highway that runs close to the city and connects Hamilton with other cities. On the other hand it is one of the busiest highways in Ontario and heavy motor vehicle traffic, especially during peak hours. Hence it is a prominent source of NO_2. As such, it was hypothesized that points close to highways will experience higher concentration of NO_2 than the distant ones. The analysis ends up with similar result showing a negative relationship between the concentration of NO_2 and distance from highway 403.

4.3.1.4 Proximity to Rail line:

Railways in Hamilton are usually used for freight transport. The inter city trains are significant source of NO_2. Furthermore, when they stop at the rail stations they generally do not turn off their engine and the emission of different pollutants, including NO_2, continues. Hence, it was assumed that the monitoring points close to train lines should be exposed to higher pollution. The variable also works as proxy for proximity to major roads and highways as well as for proximity to lakeshore industries. The regression analysis conforms to previously made hypothesis, by presenting a negative relation between the concentration of NO_2 and distance from train lines. The variable is highly significant at a 100% confidence level.

4.3.1.5 Road Length within a 50 meter buffer of sample points:

The variable again calculates the length of major roads and highways within a 50 meter buffer of the monitoring points. As discussed earlier, the major roads and highways are one of the significant sources of NO_2. Hence, it was thought that the larger the length of the major roads and highways the heavier the concentration of NO_2. The estimated model suggests similar result. Model results are presented in Table 4.9.

Table 4.9 SE model results

Variable	Coefficient	t stat	P Value	VIF
CONSTANT	3.303298	33.94	0.0000	
GOV100	-0.0000325	-2.99	0.0030	1.08
H403	-0.0000218	-1.76	0.0800	1.11
TRNLN	-0.0002514	-6.89	0.0000	1.03
POPDENSI	0.0000699	2.66	0.0080	1.03
RDL50	0.0015487	4.11	0.0000	1.03
Adjusted R2	0.2696			

4.3.2 Residual analysis

After estimating the regression model, the residuals were tested to see if any specific pattern prevailed in the distribution or not. The global spatial autocorrelation test shows that Moran's I value is 0.0336 with a p value of 0.18, which suggest a statistically insignificant spatial autocorrelation. The scatter plot of Moran I is presented in Figure 0.23 in Appendix B. In addition, The distribution of the residuals is presented in Figure 0.19 in Appendix B. It shows that mean value of the residual is very close to zero.

4.3.3 Surface generation

The generated surfaces in Figure 4.5 and 4.6 show that south-east wind pushes the pollution to the north-west part of the city. Areas around Highway 403, west part of Hamilton city, exit point to highway 403 from Dundurn Street, exit point from highway 403 to Main street are highly (28-33 ppb) polluted. This is due to closeness of major roads as well as frequent movement of motor vehicles in that area. Lower north-east part of the city such as Sherman Avenue, Gage Avenue and Kenilworth Avenues and the Burlington Bay is exposed to moderate pollution level (14-22 ppb).

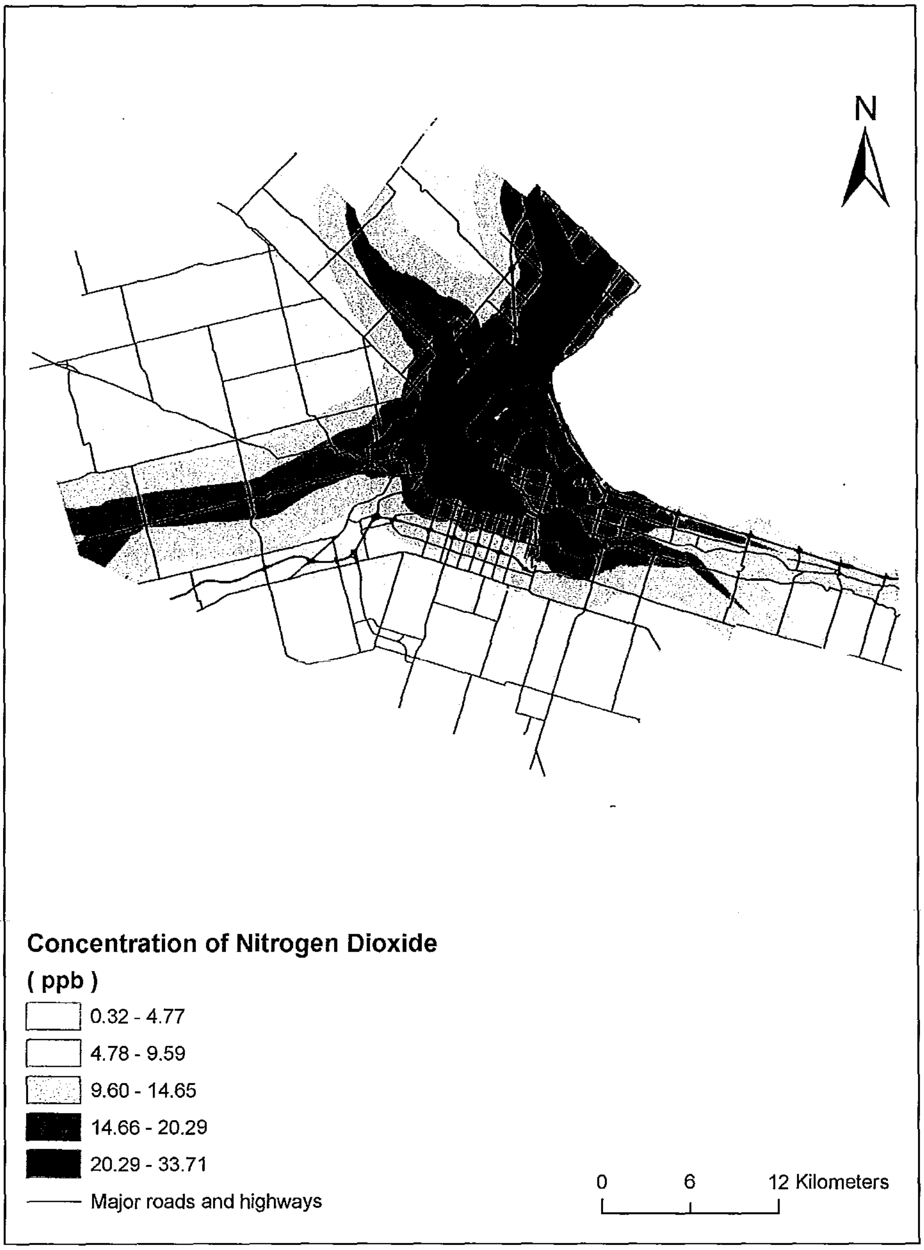

Figure 4.5 Nitrogen Dioxide distribution in the Hamilton CMA for SE wind direction

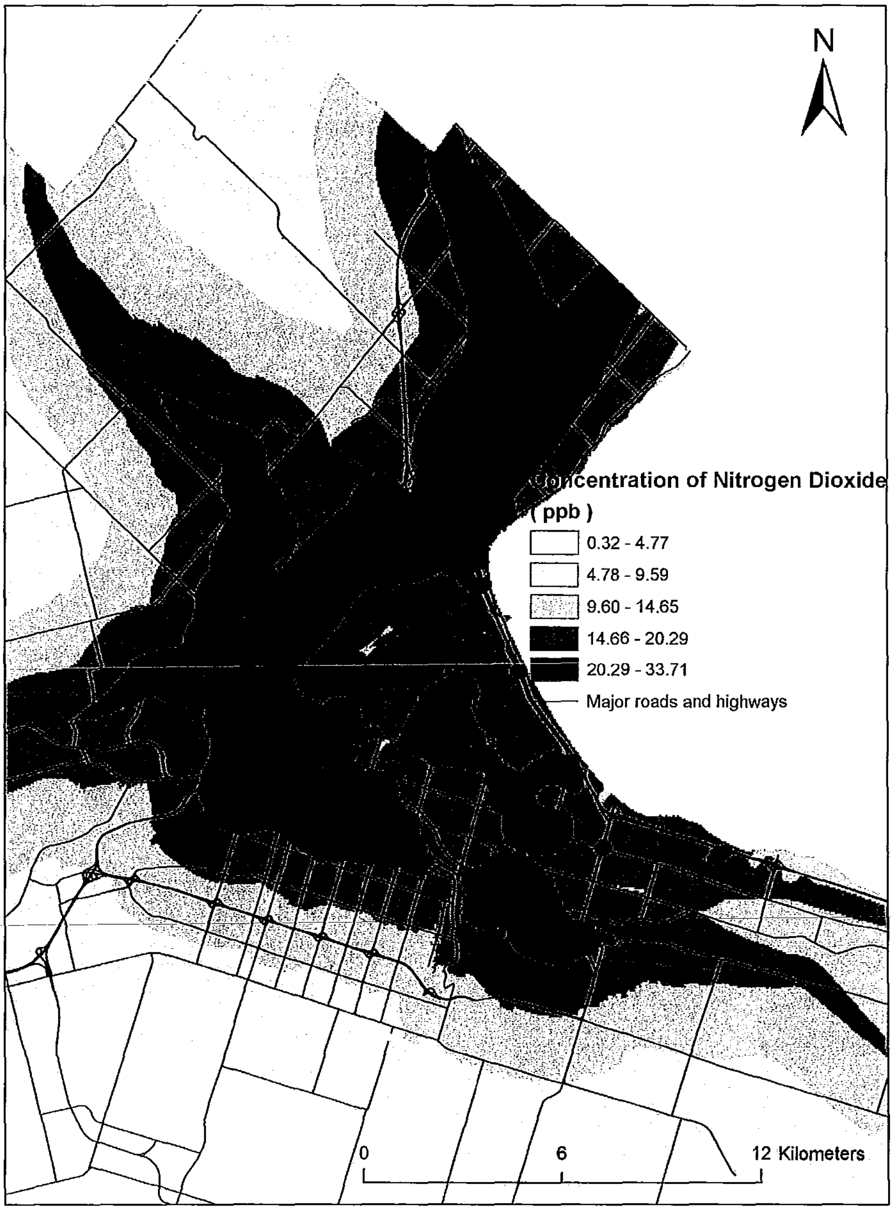

Figure 4.6 Nitrogen Dioxide distribution around downtown area (corresponds to observation points) for SE wind direction

4.4 *South-West model*

South-west is the prevailing wind direction for the city of Hamilton. Most of the data in this study belongs to this wind direction. The current model was estimated with 861 observation points. It has few distinguishable findings, which define the spatial relationship between the concentration of NO_2 and the different land use and transportation factors related to urban settings. The predictors used for the model estimation are shown in Table 4.10 with the correlation matrix in Table 4.11.

Table 4.10 Variable definition for SW model

Name of the variables	Description of the variable
WS567	Wind Speed
RDL50	Length of road within 50 meter of the observation point
TRNLN	Distance between train line and the observation points
COM350	Area of commercial area within 350 meter of the observation points
LOC_ESCT	1 if the observation points are above the Niagara escarpment, 0 otherwise
HIWY	Distance between highway and observation points
PARK50	Area of park area within 50 meter of the observation points

The correlation matrix indicated that no discernible collinearity exists among the explanatory variables.

Table 4.11 Correlation matrix SW model

	LNMNO2	WS567	COM350	LOC_ESCT	RDL50	HIWY	TRNLN	PARK50
LNMNO2	1.0000	-0.1827	0.0499	-0.2636	0.1139	-0.1289	-0.2123	-0.0662
WS567	-0.1827	1.0000	0.0882	-0.0352	0.1164	-0.0090	0.0701	0.0257
COM350	0.0499	0.0882	1.0000	0.0670	0.1319	0.0599	0.0863	-0.0327
LOC_ESCT	-0.2636	-0.0352	0.0670	1.0000	-0.0712	0.1776	0.4511	-0.1002
RDL50	0.1139	0.1164	0.1319	-0.0712	1.0000	-0.2480	0.0572	0.0126
HIWY	-0.1289	-0.0090	0.0599	0.1776	-0.2480	1.0000	-0.1863	-0.1189
TRNLN	-0.2123	0.0701	0.0863	0.4511	0.0572	-0.1863	1.0000	0.2090
PARK50	-0.0662	0.0257	-0.0327	-0.1002	0.0126	-0.1189	0.2090	1.0000

4.4.1 Variable analysis

4.4.1.1 Wind speed:

Wind speed is one of the major meteorological attributes that has significant control over the pollution concentration. It has been found from studies that as the wind speed increases, it facilitates the dispersion and ultimately reduces the concentration of pollution (similar to many previous studies for other pollutants, such as Degaetano and Doherty, 2004; Grivas et al., 2004; Martuzevicius et al., 2004; Mestayer et al., 2003; Roosli et al, 2001; Vecchi et al., 2004). The hypothesis for the variable is that higher wind speed would lower the concentration of the NO_2. The findings from the model result suggest a similar negative relation between the wind speed and concentration of pollutant, NO_2. It can be said that as the wind speed increases, the level of pollution decreases.

4.4.1.2 Commercial land use within a 350 meters buffer from the observation points:

Commercial land use is a center of human activity where through traffic and heavy vehicle movement takes place more often. Again, frequent motor vehicle movement means more emissions and higher concentration of NO_2. As such, it was assumed that if the portion of the commercial area gets larger, the level of NO_2 would get higher as well. The regression analysis confirms the previously made assumption and stands significant at a 95%confidence level. The larger portions of commercial areas again represent the built up areas in urban settings. The built up areas interrupt the regular wind flow hence the speed reduces, local wind field develops and chances of occurring high concentration of NO_2 rise.

4.4.1.3 Observation points above the Niagara escarpment:

As Hamilton has variation in its topography and the city is almost divided into two portions as lower and upper Hamilton, a variable was tested to check if there is any differences associated for the concentration of NO_2 in terms of its location or not.

Therefore, if the observation point is located above the mountain then it obtains the value of 1, or 0 otherwise. It has been found that higher concentration of NO_2 is found among the observation points that are located at the lower city. This is because major sources of NO_2 for the city are highway 403, lakeshore industries and QEW for the city. It was expected that monitoring points close to lakeshore industries and the highways will be exposed to higher pollution than the points which are situated away from the metal industries. Model result confirms the hypothesis with a 99% confidence level.

4.4.1.4 Road Length within a 50 meter buffer from the observation points:

The variable here calculates the length of the major roads and highways situated within a 50 meter buffer of the observation points. It is a continuous variable that determines how the lengths of the road influence on the concentration of NO_2. It was thought that longer roads provide higher possibility for traffic and motor vehicle movement. The higher amount of traffic will again facilitate the concentration of NO_2. The analysis from the regression model suggests the same, showing a positive relation with the concentration of NO_2. It reveals that the level of NO_2 gets higher as the length of the roads increase. The coefficient of the variable is 0.00061 and proves as highly significant.

4.4.1.5 Proximity to highways:

The study employed another transportation variable which determines the distance from the observation points to the nearest highway. This is a continuous variable and shows the relationship between the concentration of NO_2 and distance from the highway. The hypothesis for this variable was that points locating close to the highway have higher concentrations of NO_2. As NO_2 is the primary symbol of traffic emission and highway is one of the leading sources of NO_2 it is reasonable to show higher concentration of NO_2 near the highways. The regression analysis arrives at the same conclusion showing a negative association between the level of NO_2 and the proximity to highway. More generally it can be said that as the distance from the highway increases, the level of pollution decreases.

4.4.1.6 Proximity to train lines:

The research tested another continuous variable that calculates the distance from the observation points to the train lines. The train lines are mostly used for goods movement is a significant source of NO_2. So it has been tested whether the concentration of NO_2 and distance from train line has any significant relationship or not. The estimated model results show a negative relation between the concentration of NO_2 and proximity to train lines. The result has been confirmed by a 99% confidence level.

4.4.1.7 Amount of park area within a 50 meter buffer of the observation points:

This variable shows the total park area within a circular buffered area of 50 meter radius from the observation points. This variable has been introduced to capture the relationship between the park area and the NO_2 concentration. Since parks are not source of pollutants a negative association is hypothesized for this variable. The estimated model satisfies the previously made assumption and significantly shows that as the portion of park area increases, the level of NO_2 decreases.

Table 4.12 SW model results

Variable	Coefficient	T Stat	P value	VIF
Constant	3.34252	61.34	0.0000	
WS567	-0.02167	-6.19	0.0000	1.03
RDL50	0.00062	2.80	0.0052	1.10
TRNLN	-0.00011	-3.28	0.0010	1.52
COM350	0.000004	2.60	0.0093	1.04
LOC_ESCT	-0.35046	5.24	0.0000	1.45
HIWY	-0.00007	-3.13	0.0017	1.21
PARK50	-0.00004	-1.98	0.0473	1.10
Adjusted R2	0.14199			

4.4.2 Residual analysis

Spatial autocorrelation diagnosis test was performed to see whether there is any significant spatial autocorrelation among the residuals or not. The study generated Moran's *I* statistic using GEODA 9.5i. The Moran's *I* statistic appeared to be 0.025 (*p* value 0.12). Moran's *I* statistics is marginally significant. However, the value of the statistics 0.025 is very close to the expected value of Moran's *I*.

4.4.3 Surface generation

Figure 4.7 and 4.8 presents the distribution of generated surface created using the estimated model. Generally higher concentration of NO_2 is found near the highway and the heavy metal producing industries, which are the primary source of NO_2. Similar spatial characterization has been found in the generated surface as well. The estimated surface shows higher values (21-32 ppb) of NO_2 near highway 403, close to lakeshore industrial core, around Burlington Street. Higher level (21-32 ppb) of NO_2 is also observed near QEW and the major intersection of QEW and Centennial Parkway North. The other highways also hold higher level of NO_2, notably at the exit point of highway 403 from King Street West and the exit point of QEW from Burlington Street. Higher concentration (16-20 ppb) of NO_2 also prevails close the major roads. Higher level of NO_2 is also found at north of Barton Street close to the lakeshore. The downtown is not exceptional in this regard. It also experience higher concentration of NO_2, approximately 16-20 ppb.

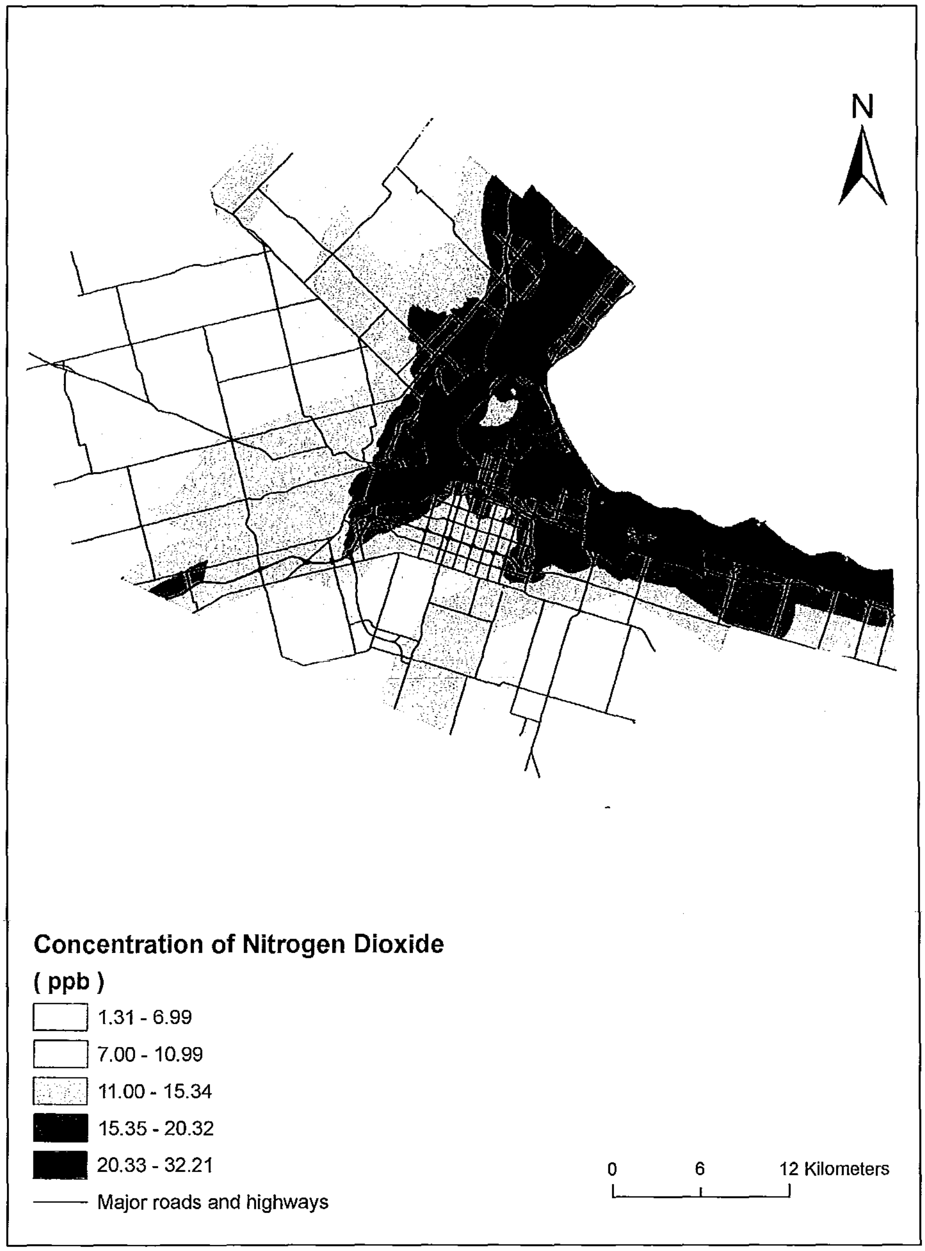

Figure 4.7 Nitrogen Dioxide distribution in the Hamilton CMA for SW wind direction

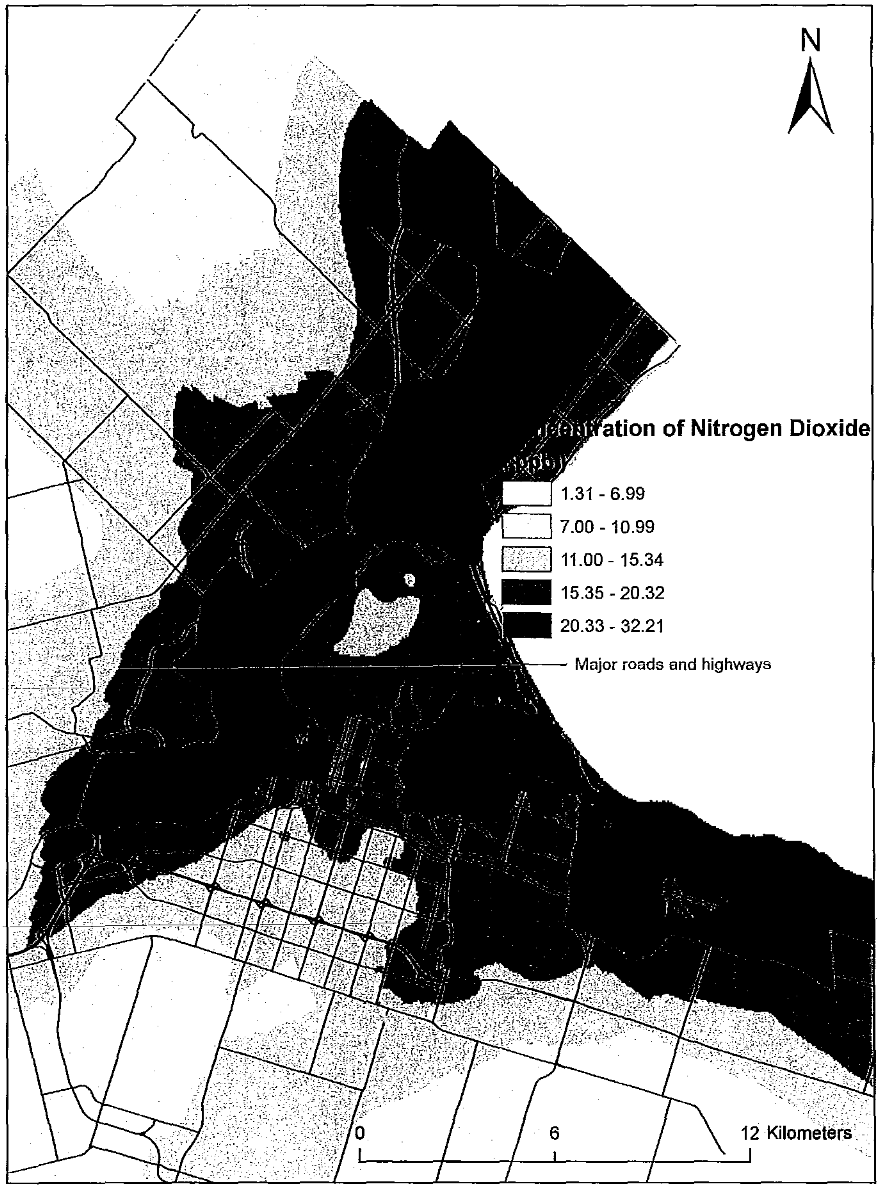

Figure 4.8 Nitrogen Dioxide distribution around downtown area (corresponds to observation points) for SW wind direction

4.5 Summary and validation

NO is a traffic emission that comes from motor vehicles and then oxidized to form NO_2. It is a highly reactive pollutant and causes several health problems for humans. In this study an effort has been made to observe the pollution condition of NO_2 in different wind directions. Mostly the areas close to highways and major roads are exposed to higher NO_2 level. When the wind blows from north-east the affected areas are mainly around QEW, Burlington Street, near the lakeshore industries and the south of Burlington Bay. Again when the wind comes from north-west higher concentration is observed around Centennial Parkway North, QEW and around Burlington Street. The southerly winds cause the higher pollution near highway 403, south part of Burlington Bay and along the QEW.

For comparing these wind direction-specific models with a single model including all direction parameters, an estimated model is presented in Table 2, Appendix B. The results suggest that models with specific wind direction provide better explanatory power in terms of model fit and parameters estimates.

Finally, the predicted values generated by the estimated land use regression models for four wind directions were assessed by evaluating observed level of concentration recorded by Ministry of Environment (MOE), Ontario at the downtown weather station (located at 43°15'30" N, 79°51'41"W). The historical hourly pollution data for November 2005 to March 2006 were obtained from archived databases. In addition, Environment Canada weather station (located at the Hamilton Airport) records were used to match corresponding hourly wind directions. Table 4.13 shows observed and predicted average concentration of NO_2 at the downtown location.

Table 4.13 Comparison of the predicted and observed pollution concentration for NO2

NO_2 Concentration			Average	Min.	Max.
North-east	Observed	Downtown location	24.95	4	61
	Predicted	Downtown location	23.22		
	Predicted	Hamilton CMA (whole)	5.92	1	54
North-west	Observed	Downtown location	13.8	3	54
	Predicted	Downtown location	12.89		
	Predicted	Hamilton CMA (whole)	8.61	1.74	57.91
South-east	Observed	Downtown location	26.17	4	63
	Predicted	Downtown location	25.27		
	Predicted	Hamilton CMA (whole)	9.86	1	33.71
South-west	Observed	Downtown location	17.79	4	53
	Predicted	Downtown location	21.22		
	Predicted	Hamilton CMA (whole)	12.32	1.31	32.21

The results reveal that the land use regression model predicts pollution concentration well since the observed and predicted values are very close. For all wind directions, except SW wind direction the model slightly underestimate pollution concentration, which could be an effect of using non-rush hour data in this study.

5. CHAPTER FIVE: LAND USE REGRESSION MODELS OF SULPHUR DIOXIDE

Hamilton, which is often called a hub of steel industry in Canada, suffers from air pollution problems where the trend of SO_2 emission is always high compared to most Canadian cities. As discussed earlier, this study uses the land use regression method to explore the relationships between the concentration of SO_2 and land use factors, transportation characteristics and meteorological and other related attributes. This chapter discusses four land use regression models of SO_2 for four wind directions (north-east, north-west, south-east, south-west), in Hamilton CMA.

5.1 North-East model

The Hamilton CMA, which is located on the south shore of Burlington Bay and holds Canada's two largest steel mills, suffers most when the wind blows from the north-east. The central business district (CBD) and downtown part of the city experience higher pollution levels than other areas. It is found that wind speed, the amount of park area within 100 meters and locations related to industries and highways are major determinants for the concentration of SO_2 when the wind blows from the north-east (NE) direction. Table 5.1 shows the definition of the variables used in the final model specification. The correlation matrix of the variable is presented in Table 5.2.

Table 5.1 Variable definition for NE model

Variables	Description of the variable
WS567	Wind Speed
PARK100	Park area within 100m buffer from the observation points
DNSTDF	1 if downwind from lake shore industrial core and located within 1.5km, 0 otherwise
DNSPBR	1 if downwind and located within 2 km of the steel industry near city center, 0 otherwise
BURG	Proximity to the intersection of QEW and Burlington Street
HIWY	Proximity to highway

Table 5.2 Correlation matrix NE model

	$LNSO_2$	WS567	PARK10	HIWY	BURG	DNSTDF	DNLKBG
LN SO_2	1.000	-0.264	-0.261	0.366	-0.106	0.301	0.182
WS567	-0.264	1.000	0.087	-0.054	0.116	-0.007	-0.138
PARK10	-0.261	0.087	1.000	-0.182	0.031	-0.178	-0.021
HIWY	0.366	-0.054	-0.182	1.000	0.565	0.402	0.347
BURG	-0.106	0.116	0.031	0.565	1.000	0.133	0.256
DNSTDF	0.301	-0.007	-0.178	0.402	0.133	1.000	-0.206
DNSPBR	0.182	-0.138	-0.021	0.347	0.256	-0.206	1.000

The final model explains 34.4% variability of the concentration of the SO_2. Table 5.3 shows the results of the model. All the variables are statistically significant at the 99% confidence level. In the following subsection, each of the variables of the final model is discussed in detail.

5.1.1 Variable Analysis

5.1.1.1 Wind Speed:

Wind speed is an important meteorological variable that has a significant effect on the concentration of the pollutants. The variable presented here is a continuous variable and is measured in kilometers per hour. The wind speed ranges from 0 to 15 km/hr for the sample data. The hypothesis for this variable is that the higher the wind speed the lower the pollution concentration. Since higher wind speed facilitates dispersion, pollutants are generally transported from the source to other wider areas mostly towards the downwind locations. The model estimation results comply with the hypothesis and reveals that wind speed is negatively associated with the concentration of SO_2. Therefore, it can be said that higher wind speeds reduce the pollution concentration.

5.1.1.2 Amount of park area within a 100m buffer:

The land use variable, which defines the amount of park area within a 100 meter buffer of the observation points, appeared significant for our models. This is again a continuous variable and ranges between 0 and 30615m^2 for the sample. The hypothesis for this variable is that the higher the amount of the park area near the observation points, the lower the concentration of SO_2. The results satisfy the hypothesis and show a negative association with the concentration of SO_2 meaning that larger parcels of park area around the observation points result in reduced pollution levels.

5.1.1.3 Downwind and within 1.5 km from the lakeshore industries:

Stelco and Dofasco Inc., the two largest steel producing industries are the major sources of air pollution in Hamilton. These industries emitted 9572.167 tonnes of SO_2 in 2006 and pollute the urban environment regularly (NPRI, 2007). In addition, other metal industries have developed in close proximity to these heavy industries over time, and also emit enormous amounts of SO_2. This cluster of heavy industries is termed as the lakeshore industrial core. To test the significance of the industrial core on the concentration of SO_2, a categorical variable was introduced into the model that defines the observation points, which are at downwind locations from lakeshore industries and situated within a radius of 1.5 kilometers from the industries. That means, the dummy variable reflects the observation points (that assume a value 1) in the data set which are directly downwind from Stelco and Dofasco on Burlington Street. The hypothesis is that the concentration levels of the pollutant at these points (which are downwind from Stelco and Dofasco within a distance of 1.5 kilometers) would be high compared to other locations in the region. The model confirms the result with more than 99% confidence and a high coefficient value (0.383). Hence it suggests that when the wind comes from north-east direction, it collects the pollution emitted from the industries and pushes them towards the lower city (that lies in the downwind direction).

5.1.1.4 Downwind and within 2 km from the industry near city center:

Another categorical variable was created to capture the effect of industrial emissions near the city center. This variable defines the points, which are situated further south of Burlington Street, downwind from the steel producing industry, Hamilton Speciality Bar Inc., which is located near the CBD at Sherman Avenue and within a distance of 2 kilometers of the industry. The points are generally on and around Barton Street and near the downtown area. The *a priori* expectation for the variable was that near to the steel industry (Hamilton Specialty Bar) the pollution would be higher. The steel industry emitted 18 tonnes of SO2 in 2006 (NPRI, 2007). These points reflect the areas around the downtown part of the city. The variable obtained the value as 1 when it is down wind from the metal industry and within 2 kilometers from it, and 0 otherwise. The model result complies with the result with 99% confidence and a high coefficient value (0.457) demonstrating a positive relation with the concentration of SO_2.

5.1.1.5 Distance from the intersection of Burlington Street and QEW:

The continuous variable represents the distance from the intersection of Burlington Street and the QEW. This variable stands for the proximity to a major intersection (a major exit point for the city) and proximity to a major industry, Columbian Chemicals that generates huge amounts of SO_2 (according to NPRI report, 2007 the industry emitted 1318 tonnes SO_2 in 2006) and is located close to the intersection. It is assumed that the observation points near to this area would have higher concentrations of SO_2 because of the presence of this industry. The final model results also confirm the fact showing a negative relation with the distance from Burlington Street and the concentration of SO_2. Therefore it is found that the concentration of SO_2 increases as the distance from the intersection of Burlington Street and QEW decreases.

5.1.1.6 Distance from highways:

This is another continuous variable that represents the distance from the observation points to the nearest highway. Since SO_2 is mainly an industrial emission, it is expected that there is no discernible negative relationship exists between the proximity to the highways and the concentration of SO_2. Rather, the level of pollution is dependent on the location of industries. Since generally industries are located relatively distant from the highways in the Hamilton area (with the exception of the QEW) and most highways are at the upwind location when wind blows from the north-east direction, the *a priori* hypothesis is that there is a positive relationship between the variable and the concentration level. The model results show the expected positive B coefficient and reveals that the higher the distance from the highways (hence closer and downwind to the industries) the higher the concentration of SO_2.

Table 5.3 NE model results

Variables	Coefficient	T Stat	P Value	VIF
Constant	2.83191	24.438	0.0000	
WS567	-0.05478	-3.77	0.0002	1.06
PARK100	-0.00003	-2.755	0.0062	1.08
DNSTDF	0.38301	3.707	0.0002	1.45
DNSPBR	0.45733	2.946	0.0034	1.39
BURG	-0.00017	-7.308	0.0000	1.58
HIWY	0.00036	6.908	0.0000	2.15
Adjusted R2	0.3443			

5.1.2 Residual analysis

The residuals of the models were tested to confirm that they are not spatially correlated and are distributed randomly. The study used GEODA 9.5i to carry out the diagnostic test and found the Moran's *I* statistic as 0.06 with a *p* value 0.12 which states that there is no statistically significant spatial correlation among the residuals. The Moran's *I* scatter plot is presented in Figure 0.29 in Appendix B. Also, the distribution of the residuals is shown in Figure 0.25 (Appendix B). It is evident from the figure that mean values of the residuals are zero or very close to zero.

5.1.3 Surface generation

The surface of SO_2 distribution in Figure 5.1 clearly shows that the most affected area for SO_2 pollution is near the Burlington Street, Industrial Drive and north of Barton Street and near the downtown area. When the wind blows from north-east it collects the pollution from the lakeshore industries and pushes it towards the city center. These lakeshore industries which are mostly steel producing industries emit the bulk amount of SO_2 in the city. Close to the downtown area along Sherman Street North, the steel factory also contributes to SO_2 pollution. Therefore, higher pollution (approximately, 15 ppb) takes place at the northern part of downtown between Wellington Street and Gage Avenue. In general, the north-east wind affects mostly the residential area around the downtown area. In some cases strong wind driving the pollution towards the city center, is deflected by the Niagara Escarpment and circulates into the downtown residential areas. Mostly the north-east and the east part of the city are exposed to severe pollution levels due to this north-east wind. Figure 5.2 shows the affected areas of downtown Hamilton.

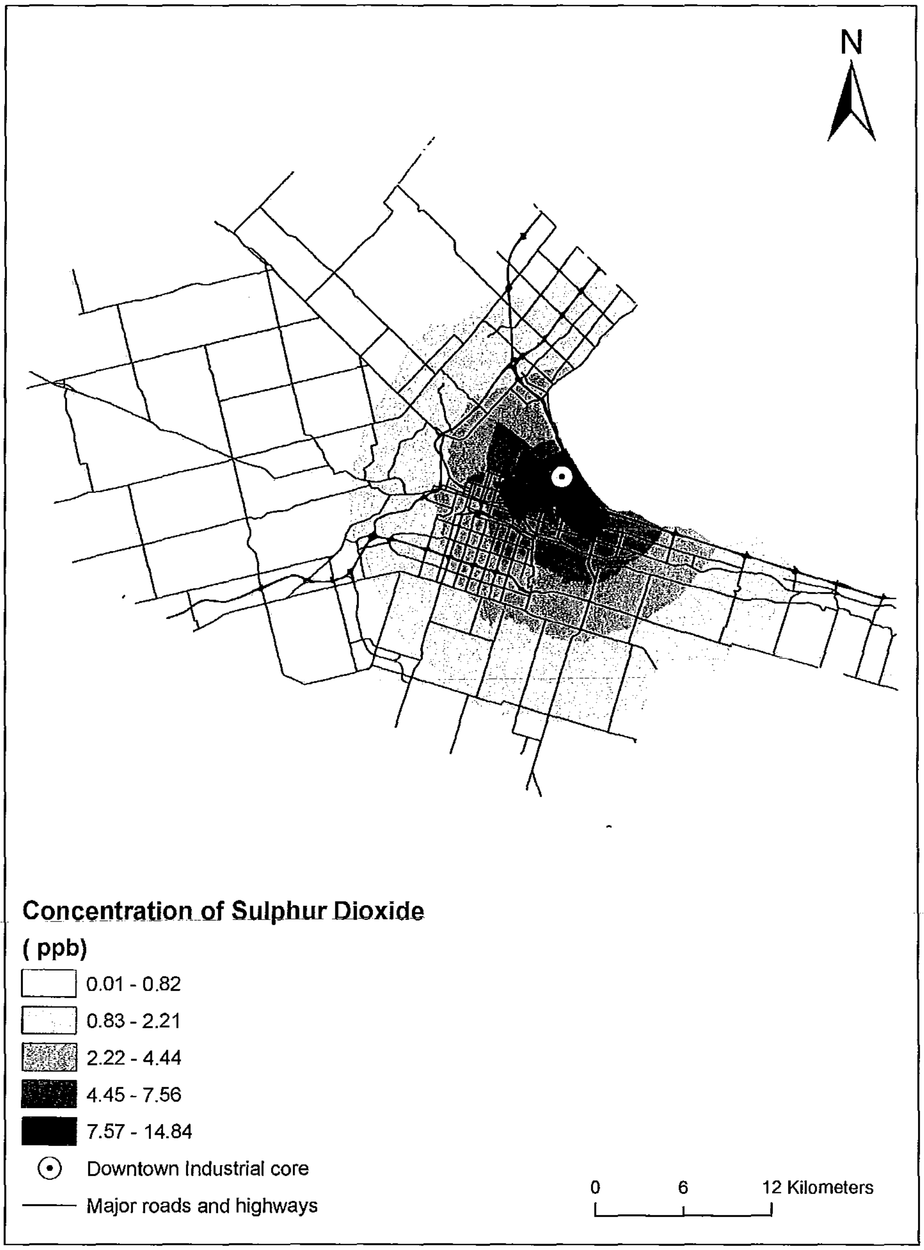

Figure 5.1 Sulphur Dioxide distribution in the Hamilton CMA for NE wind direction

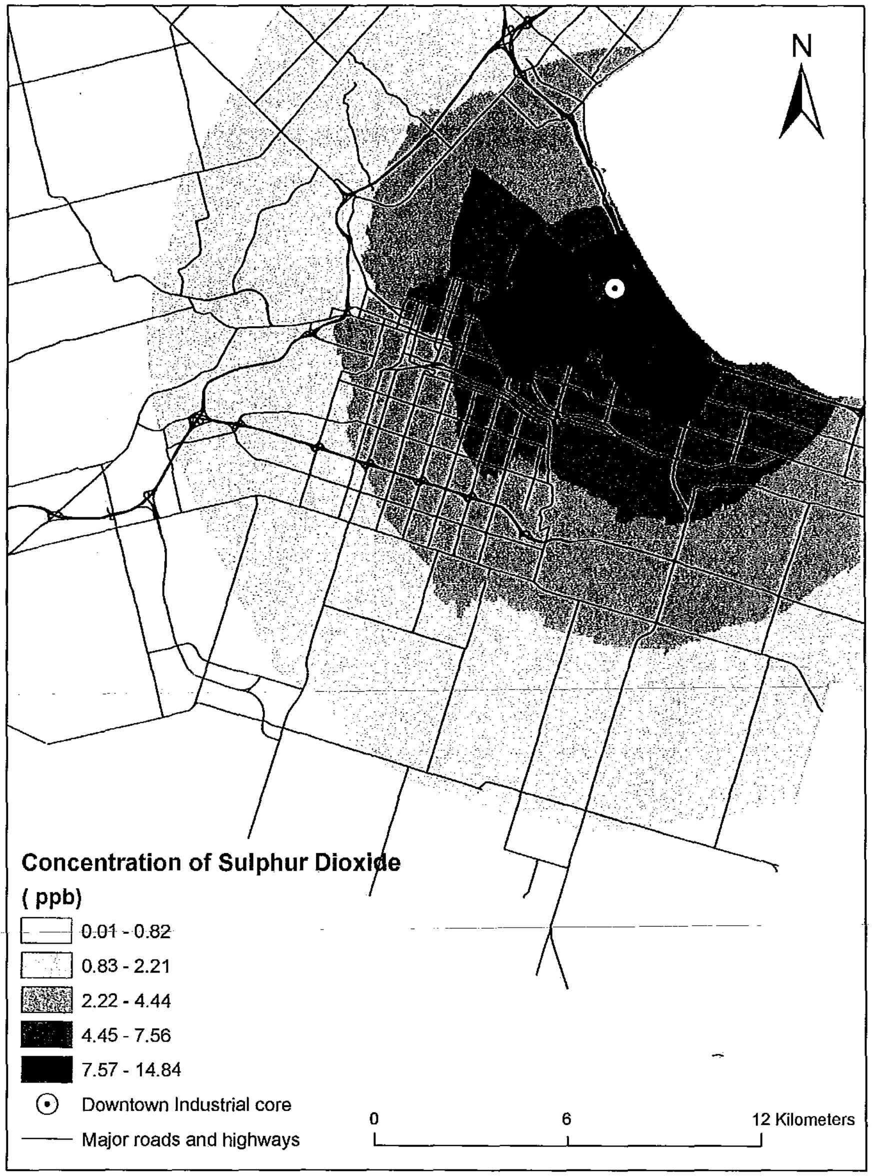

Figure 5.2 Sulphur Dioxide distribution around downtown area (corresponds to observation points) for NE wind direction

5.2 North-West model

When the wind flows from the north-west it pushes the pollution towards the south-east of the city. Major polluters are here again, the lake shore industrial core and a steel industry named Alcarb Resources which is close to the major intersection of Centennial Parkway North and QEW. Below is the analysis of the variables used for the model calibration. The coefficient of determination of the model is 35.5%. The variable definitions are shown in Table 5.4. Table 5.6 shows the correlation matrix.

Table 5.4 Variable definition for NW model

Variables	Definition
WS567	Wind Speed
COM350	Commercial area within 350m buffer from the observation points
DNSKYIND	1 if downwind from lakeshore industrial core and located within 2km, 0 otherwise
DNCENT	1 if located within 1 km from the steel industry (which is close to the intersection of QEW and Centennial Parkway North), 0 otherwise

Table 5.5 Correlation matrix NW model

	LN SO_2	WS567	COM350	DNSKYIND	DNCENT
LN SO_2	1.0000	0.0180	0.0214	0.5500	0.0805
WS567	0.0180	1.0000	-0.0552	0.3323	-0.1999
COM350	0.0214	-0.0552	1.0000	-0.1515	0.0775
DNSKYIND	0.5500	0.3323	-0.1515	1.0000	-0.1471
DNCENT	0.0805	-0.1999	0.0775	-0.1471	1.0000

5.2.1 Variable analysis

5.2.1.1 Wind speed:

The meteorological variable, wind speed, seems to have a strong effect on the concentration of SO_2. It was hypothesized that as the wind speed increases the concentration decreases. This is because the higher wind speed facilitates the dispersion which in turn reduces the concentration of SO_2. We have found the same result in our model showing that concentrations of the pollutant decrease with higher wind speeds. The model confirms this result with more than 99% confidence (Table 5.5).

5.2.1.2 Downwind from the lakeshore industries:

This is a categorical variable. The variable defines the situation if the monitoring points are located down-wind from the lake shore industries and remains within 2 kilometers of the steel industries, thus obtain the value as 1 or 0 otherwise. This variable has been created to explain the high SO_2 values, specifically those that lie on or around the Skyway Bridge and Burlington Street. When the wind blows from the north-west, the lake shore industries (specifically the industries of Stelco and Dofasco) play a major role in SO_2 pollution. With moderate or the higher wind speeds the pollution emitted from the industries are moved toward the intersection of Burlington Street and the QEW as well as on the Skyway Bridge. It is hypothesized that when wind comes from the north-west it carries the pollution from the industries and pushes the pollution towards the southern side. The model results correspond to the assumption showing a positive relation between the pollution concentration and downwind location from lake shore industries.

5.2.1.3 Downwind to the metal industry

The model includes another categorical variable that describes the state of the observation points when located within a 1 kilometer buffer of the metal industry near

QEW and Centennial Parkway North. The points within the selected region are assigned a value of 1 or 0 otherwise. According to National Pollutant Release Inventory 2007 report (NPRI) a large metal industry (Alcarb Resources) is located at Lanark Street, near the highway intersection of QEW and centennial Parkway North. This industry emits significant amounts of SO_2. It was expected that near the sources there will be a plume of high SO_2, so when the wind blows from north-west, the pollution towards the downwind location will also be high. Analysis of the variable from the regression model confirms this previously made assumption and places a positive relation with the concentration of SO_2. Thereby, it is evident that points within a 1kilometer buffer and downwind of the metal industry near the QEW and Centennial Parkway North experience elevated concentrations of SO_2 when the wind is blowing from the north-west.

5.2.1.4 Commercial area within 350meters

The variable defines the area of influence for the observation points. To create the variable, circular buffers with 350 meter radii were created around the monitoring points and intersected with the land use map to calculate the different areas of around the observation points. Then the commercial areas were separated to be used as a variable to see whether the commercial area has any profound effect on the SO_2 or not. It was expected that as the commercial area gets larger the pollution of SO_2 should increase because the commercial area around the observation points are mostly located near the lake shore industries and commercial heating is also a source of SO_2 (Schmandt, 1988). The results complied with the hypothesis. It shows that higher portion of commercial area around the monitoring points results in higher SO2 pollution. This variable is also a representation of built up area in urban environment. More commercial area represents more buildings. The built up area generally interrupts with the free flow of the wind. Obstruction of free wind flow works against the dispersion of pollutants which ultimately means high concentration of pollution. So the assumption stands as the more the commercial area, the higher the concentration of SO_2. The result satisfies the assumption with 99% confidence and a high t statistics of 4.124 (see Table 5.6).

Table 5.6 NW model results

Variable	Coefficient	T Stat	p value	VIF
Constant	1.866113	29.885	0.0000	
WS567	-0.016659	-3.965	0.0001	1.15
DNSKYIND	1.380305	15.637	0.0000	1.15
DNCENT	0.466482	3.448	0.0006	1.05
COM350	0.000005	2.566	0.0106	1.03
Adjusted R2	0.3555			

5.2.2 Residual Analysis

Finally, the residuals were tested to determine the level of spatial autocorrelation. The study used Geoda 9.5i and found Moran's *I* statistic as .003 with *p* value of 0.42, which is not statistically significant at a reasonable level of significance. Scatter plot is shown in Figure 0.30 in Appendix B). Also, the distribution of the residuals shows that residual mean value is close to zero (Figure 0.26 in Appendix B).

5.2.3 Surface generation

Using the final model specification a surface was created to present the distribution of SO_2 in Hamilton CMA. The surface for SO_2 distribution (Figure 5.3) exhibits that when the wind blows from the north-west, the east part of the city is generally affected. The intersection of Burlington Street and QEW experiences highest pollution (20-24 ppb) concentration. The wind from the north-west accumulates the pollution from lakeshore industries (particularly from Dofasco Inc. VFT Canada Inc. and Columbian Chemicals) and pushes it towards the south-east part of the city. Areas close to Parkdale Avenue North, Woodward Avenue North and the intersection of Burlington Street and QEW experience high pollution. The situation also gets worse near Warrington Street due to Alcarb Resources, a metal service center and producer of steel and heavy metals. Hence the downwind locations, which are mostly east of Centennial Parkway North, are more polluted (16-19 ppb). A detailed map of Sulphur Dioxide distribution for NW wind directions is presented in Figure 5.4.

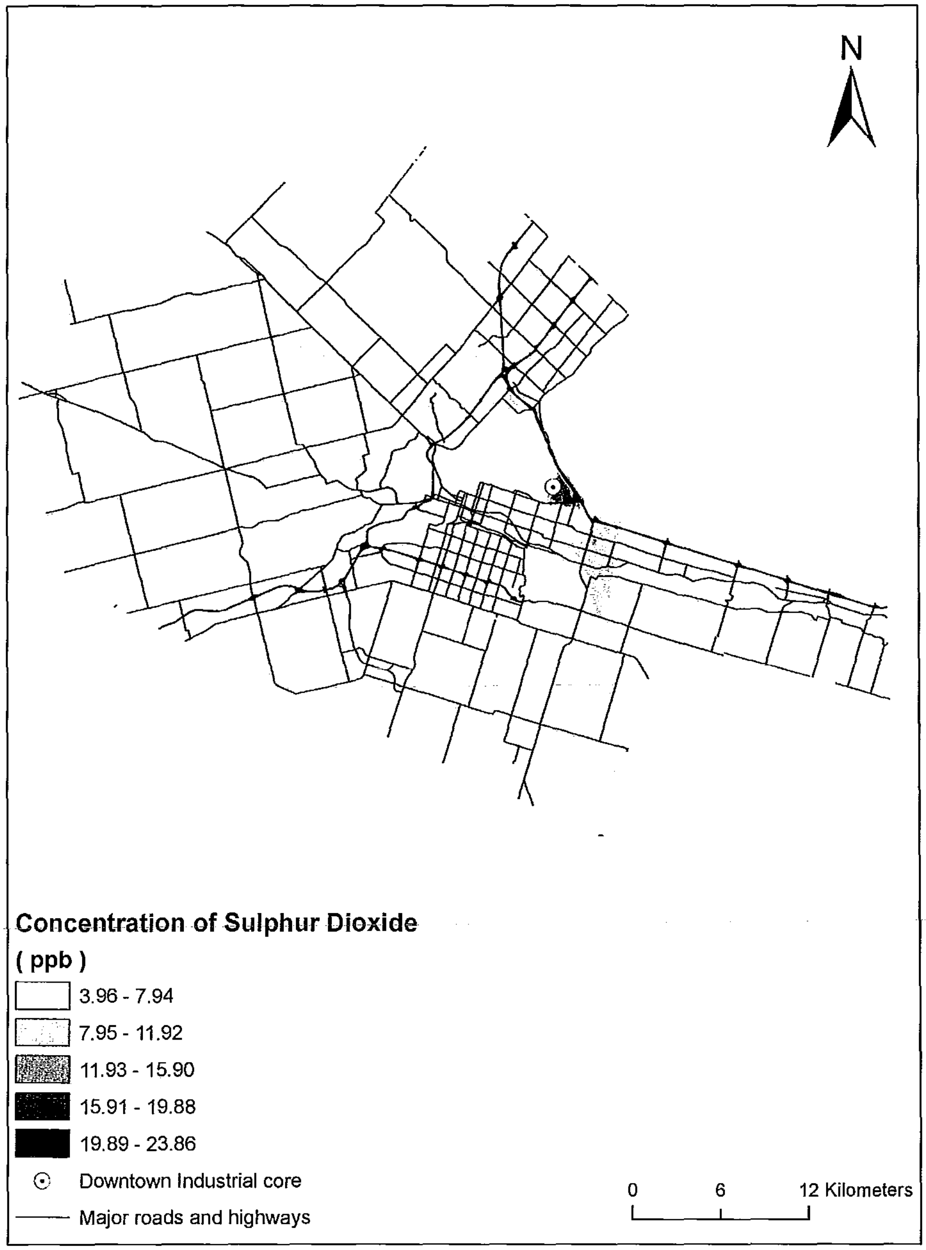

Figure 5.3 Sulphur Dioxide distribution in the Hamilton CMA for NW wind direction

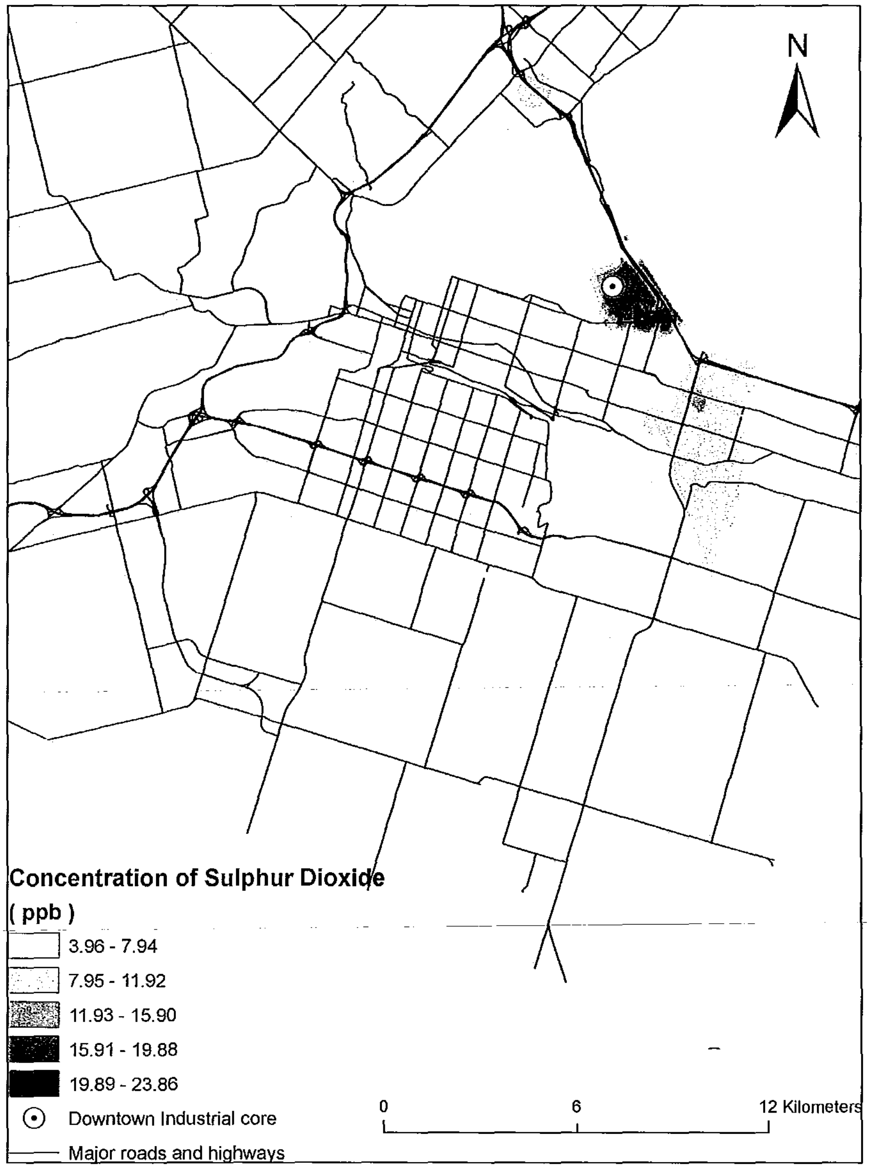

Figure 5.4 Sulphur Dioxide distribution around downtown area (corresponds to observation points) for NW wind direction

5.3 *South-East model*

Winds form the south-east mainly pollute the areas close to lake shore and Burlington Bay. The main sources of SO2 are in the industrial zone. This research estimated a land use regression model, which explains 33.4% of the variability in the concentration of SO_2. Table 5.7 presents the variable definitions.

Table 5.7 Variable definition for SE model

Variables	Definition
RES350	Residential areas within 350m buffer from the observation points
DNDTIND	1 if downwind from lakeshore industrial core, 0 otherwise
STLIND	1 if downwind from other major steel industries and locates within 1km of the industries, 0 otherwise
ESTPRT	1 if located within 750 meter from the East Port Industries, 0 otherwise
BAYFRONT	1 if located within 500 m from the Bay Front Park, 0 otherwise

The Pearson's Correlation matrix in Table 5.8 shows that there is no significant correlation between the explanatory variables.

Table 5.8 Correlation matrix SE model

	$LNSO_2$	RES350	DNDTIND	STLIND	ESTPRT	BAYFRONT
LN SO_2	1.0000	-0.2556	0.1799	0.2954	0.2254	-0.4044
RES350	-0.2556	1.0000	0.1252	-0.2730	0.1536	0.3016
DNDTIND	0.1799	0.1252	1.0000	-0.1467	-0.0474	-0.0521
STLIND	0.2954	-0.2730	-0.1467	1.0000	-0.1395	-0.1535
ESTPRT	0.2254	0.1536	-0.0474	-0.1395	1.0000	-0.0496
BAYFRONT	-0.4044	0.3016	-0.0521	-0.1535	-0.0496	1.0000

5.3.1 Variable Analysis

5.3.1.1 Residential areas within 350m buffer:

This variable defines the total residential area within the 350 meter buffer of the observation points. The hypothesis for this variable is that the pollution of SO_2 near the residential area will be lower in concentration. It has been found from the exploratory analysis that generally, there is a lower density of residential areas near to the heavy metal industries, particularly the steel mills which emit SO_2. As residential areas are not a significant source of SO_2, it is very likely that SO_2 pollution near that area will be lower. Similarly the estimated model produces a negative relation with the concentration of SO_2 confirming that the higher the portions of residential area within 350 meter buffer, the lower the pollution level.

5.3.1.2 Downwind from Lakeshore Industries:

The observation points, which are downwind from the lakeshore industries, Stelco and Dofasco, are specified by this variable. When the wind blows from south-east, the most effected areas are the south shore of Burlington Bay and some areas which are located at the southern part of Waterdown. This is because when the wind blows from south-east it pushes the pollution towards the north-west, hence Burlington Bay becomes mostly polluted. The model presents the variable as a categorical one. The study put the value as 1 when the points are downwind location from Stelco and Dofasco, 0 otherwise. The variable is significant at a 99% confidence level and presents the positive association with the concentration of SO_2. It is found that points downwind from the lakeshore industrial core experience more pollution concentration than other areas.

5.3.1.3 Downwind and within 1 km from all other steel-producing industries:

The points, which are downwind from steel industries other than Stelco and Dofasco and within 1 kilometer, were defined by this variable. This variable only considered the metal industries that emit substantial amounts of SO_2 according to the National Pollution Release Inventory (NPRI) 2007 data. The selected points are mainly on the Pier 24 Access-Burlington Street and on Industrial Drive. The industries responsible for the pollution considered here are the Columbian Chemicals, VFT Canada Inc and Hamilton Specialty Bar. The hypothesis behind this variable is that when the wind blows from the south-east, generally points which are downwind to these industries and located within a distance of 1 kilometer are subject to more pollution of SO_2 than the others. The variable appears very significant at 99% and describes a positive relation with the concentration of SO_2. The model confirms this result at approximately 99% confidence level.

5.3.1.4 Points within 750m from East Port industries:

The next variable for the model defines the points which are at distance of 750 meters from the industry at East Port Drive. The industry at East Port Drive (Bitumar Hamilton) produces bitumen and emits of SO_2 (according to NPRI report, 2007 the industry emitted 35 tonnes of SO_2 in 2006). On the other hand the area around East Port Drive experience higher pollution almost all the time as two big metal industrial facilities (Dofasco, Colombian Chemicals) are located close to that area. A dummy variable was placed here assigning the value at 1 when it is situated within 750m of the industries, 0 otherwise. The model shows a positive relation with the dependent variable and satisfying the assumption for the variable with a high coefficient value.

5.3.1.5 locations within 500m from the Bay Front Park

The model presents another categorical variable that captures the observation points, which are within 500m of the Bay Front Park. As there are no significant sources of SO_2 near the Bay front it was expected that the areas near Bay Front Park will have less pollution exposure. To account for this fact a variable was introduced as the observation points located within 500m from the Bay Front Park will receive the value of 1, or 0 otherwise. The model ended up with conclusion that there is a negative relation exists between the proximity of Bay Front Park and pollution concentration. It can be said from the analysis that points within 500m of Bay Front Park will have lower concentration of SO_2 than the points which are located further from the park and closer to the industrial core, when the wind flow is from south-east.

Table 5.9 SE model results

Variable	Coefficient	T Stat	P value	VIF
Constant	2.732388	48.426	0.0000	
DNDTIND	0.629998	4.381	0.0000	1.05
STLIND	0.340831	4.955	0.0000	1.12
ESTPRT	0.786884	5.214	0.0000	1.06
BAYFRONT	-0.71196	-4.964	0.0000	1.14
RES350	-0.000001	-2.832	0.0050	1.21
Adjusted R2	0.3346			

5.3.2 Residual analysis

Finally the distribution of residuals was observed to determine the difference between the observed values and the predicted values. The distribution shows that most of the residual values are close to zero (Figure 0.27 in Appendix B). The Moran's *I* statistic is 0.0313 with a *p* value of 0.2. The Moran's *I* scatter plot is shown in Figure 0.31 in Appendix B.

5.3.3 Surface generation

The final surface is presented in Figure 5.5. As the lakeshore heavy metal producing industries are the major sources of SO_2 emission, the north-west side of city experience higher pollution when the wind flows from south-east. At this time of wind direction Burlington Bay and north of Plains Road and Highway 403 are exposed to a significant amount of pollution (22-31 ppb) particularly. Some high values of SO_2 are also observed along the skyway bridge. The core downtown portion and the upper city have less pollution (1-6 ppb). Figure 5.6 shows a detail map of the surface of SO_2 pollution in the Hamilton CMA.

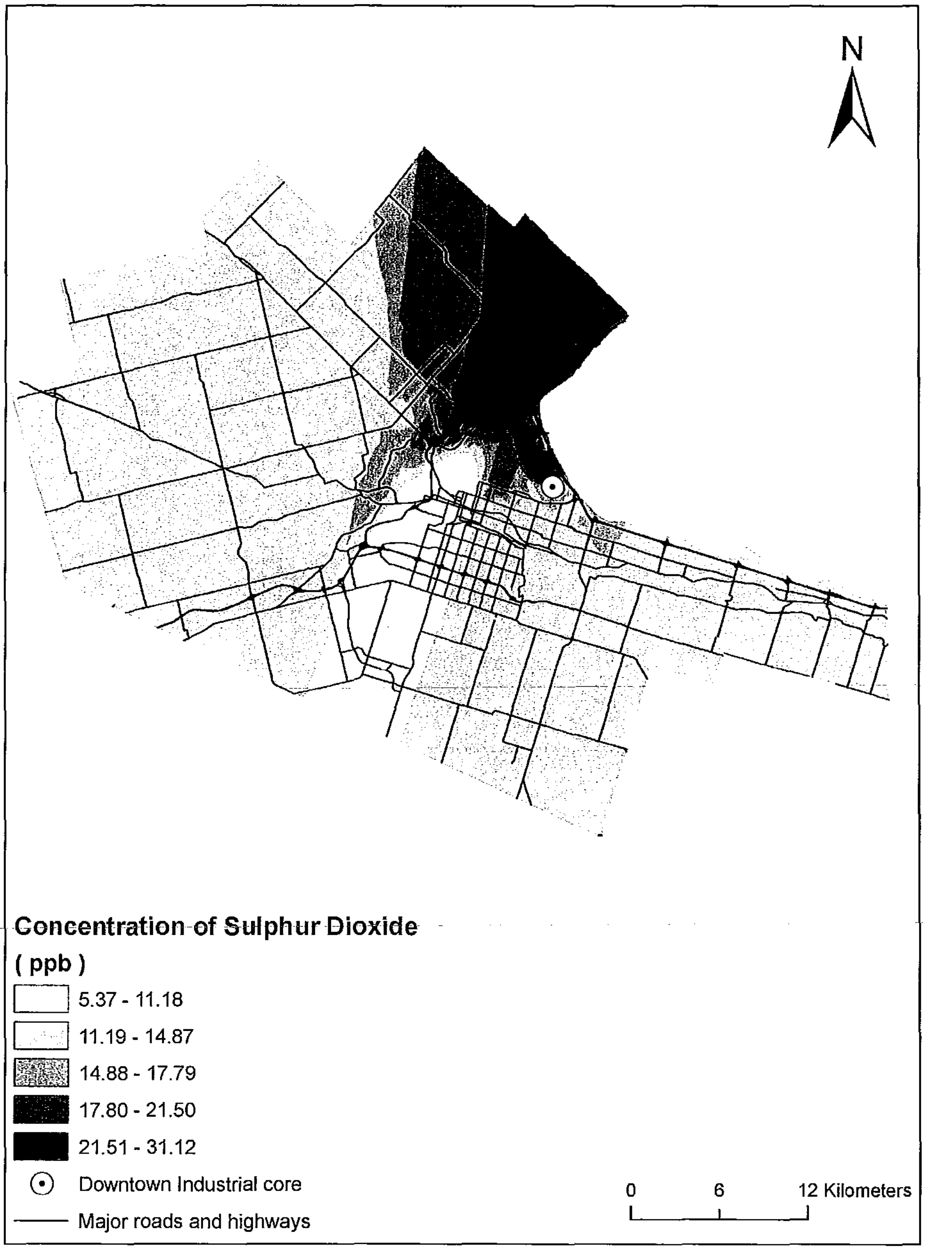

Figure 5.5 Sulphur Dioxide distribution in the Hamilton CMA for SE wind direction

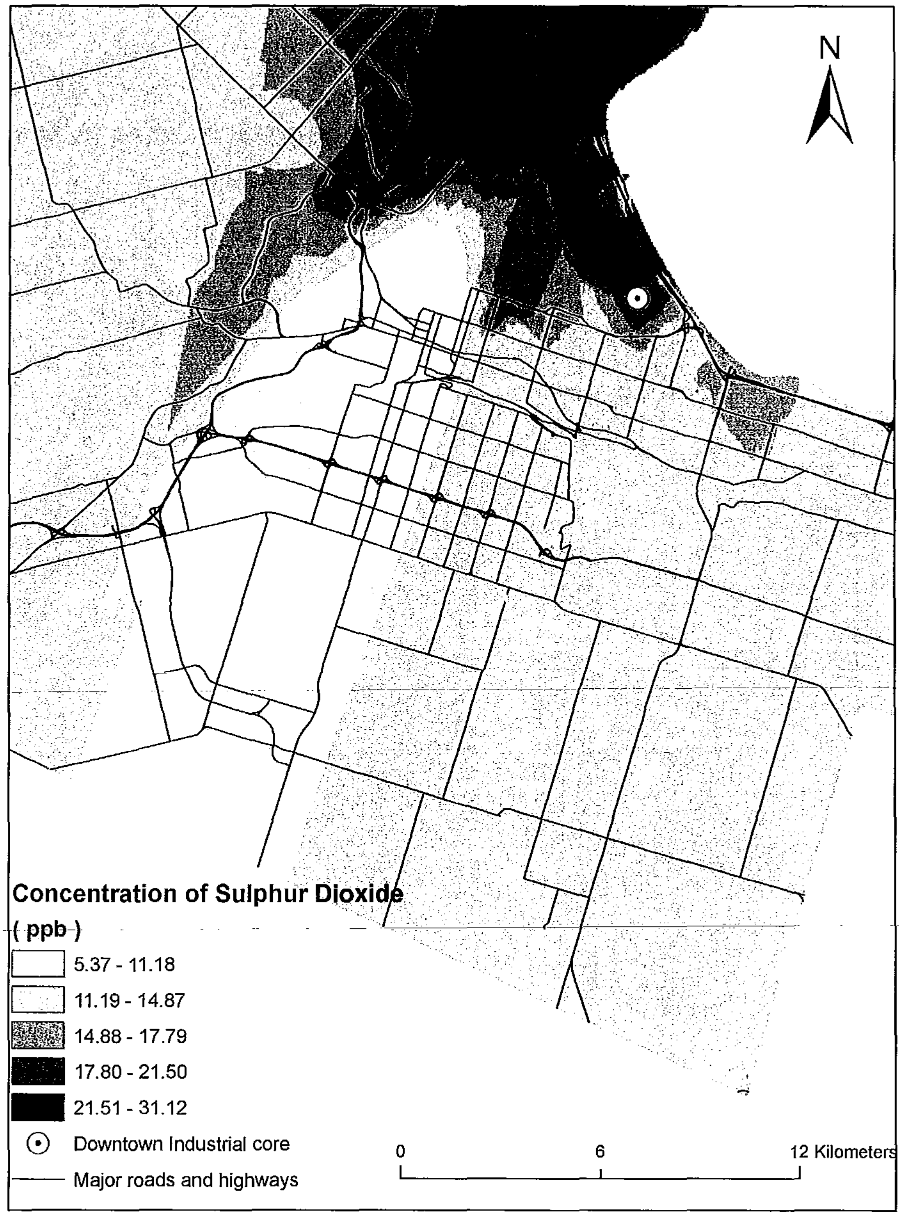

Figure 5.6 Sulphur Dioxide distribution around downtown area (corresponds to observation points) for SE wind direction

5.4 *South-West model*

South west is the prevailing wind direction for Hamilton. When the wind blows from the South-West, the pollutants are usually moved towards the outer city around the Skyway Bridge. Here again, the wind speed and other land use factors play a crucial role. Our model explains the 33.3% variability in the concentration of SO_2. The next section of this research explains the variables used in the model, their validity and finally the model results.

Table 5.10 Variable definition for SW model

Variables	Definition
WS567	Wind Speed
ESTPRT	1 if located within 750 meter from the East Port Industries
DNBAY	1 if located within 1500 m from the Bay Front Park, 0 otherwise
QEWCEN	1 if upwind from steel industries around QEW and Centennial Parkway North and locates within 1.5 km, 0 otherwise
DNDTIND	1 if downwind from lake shore industries, 0 otherwise

The correlation matrix is presented in the Table 5.11

Table 5.11 Correlation matrix SW model

	$LNSO_2$	WS567	ESTPRT	DNBAY	QEWCEN	DNDTIND
$LNSO_2$	1.0000	-0.0625	0.4164	-0.1908	-0.2138	0.4813
WS567	-0.0625	1.0000	0.1147	-0.0684	0.1008	0.1151
ESTPRT	0.4164	0.1147	1.0000	-0.0572	-0.1012	0.4116
DNBAY	-0.1908	-0.0684	-0.0572	1.0000	-0.0652	-0.1389
QEWCEN	-0.2138	0.1008	-0.1012	-0.0652	1.0000	-0.2459
DNDTIND	0.4813	0.1151	0.4116	-0.1389	-0.2459	1.0000

5.4.1 Variable analysis

5.4.1.1 Wind speed:

Wind speed is one of the important factors that describe the variability in the concentration of SO_2. As higher wind speed facilitates the dispersion a negative relationship between the dependent variable, level of SO_2 and the explanatory variable, wind speed was expected. The result found from the model estimation confirmed the *a prioi* hypothesis, that the strong wind speed facilitates the dispersion and lowers the concentration of the pollutant.

5.4.1.2 Downwind from Lake shore Industries:

The variable down wind from steel industries is the second and one of the important variables for this model estimation when the wind comes from south-west. It is a categorical variable and has two values, either 1 or 0. The value 1 denotes the points which are downwind from the lakeshore industrial core, Stelco and Dofasco. The points are mostly located around the intersection of Burlington Street and QEW, on the Skyway Bridge and on North Shore Boulevard East. These are the locations which are primarily affected due to industrial pollution from Stelco and Dofasco when the wind blows from South-West. The *a priori* expectation for this variable was that when the wind blows from south-west, the most affected area for SO_2 will be the downwind location from Stelco and Dofasco, mainly Burlington Bay and areas around it. The model result confirms the hypothesis with 99.9% confidence with a 't' statistics of 11.22.

5.4.1.3 Locations within 750m from the East Port Drive Industries:

This is another dummy variable that has been introduced to point out the hot spots of the air pollution of SO_2. The hypothesis behind this variable is to identify and specify the location points which are around the skyway bridge, on East Port Drive and which lie within 750m of the industry (Bitumar Hamilton) on the East Port Drive. The industry produces liquid bitumen and asphalt and emits SO_2. Also, this area is approximately within 1.5 kilometers of Dofasco and Columbian Chemicals, hence in general, is exposed to higher pollution levels. Thus, the analysis reveals that the points, which are located on the East Port Drive and within 750m of the industries, have higher concentrations of SO_2 confirming the previously made hypothesis.

5.4.1.4 Upwind from the metal producing industry (near QEW):

The study introduced this variable by specifying the points which are at upwind wind locations and within 1500m of a steel producing industry. The selected points are assigned a value 1, the rest of the data points get the value 0 otherwise. The hypothesis behind this variable is that the upwind location points from the industries will experience lower pollution compared to the downwind locations. The results suggest the same, identifying a negative relation with the concentration of SO_2. Hence, it can be said that the points which are at upwind locations from Alcarb Resources and remain within the 1500m buffer are exposed to lower pollution than the other downwind location points.

5.4.1.5 Locations within 1500m of Bay Front Park

This variable specifies the location points which are within 1500m of the Harbor Front Park. Accordingly, points within that defined area were given the value 1, or 0 otherwise. The assumption for this variable is that as there is no major metal industry located at the south side of Bay Front Park, thus the possibility of SO_2 pollution near this area is very low. That is why the points near the Bay Front park show very low values. Model results conclude that areas close to Bay Front Park have low levels of SO_2 when the wind comes from the south-west. The variable is negatively related with a coefficient value of -.5449 and a 't' statistics of -5.334.

All parameters are statistically significant at least at the 5% level of significance. Table 5.12 shows the final model results.

Table 5.12 SW model results

Variables	Coefficient	T Stat	P Value	VIF
Constant	2.5574	54.87	0.0000	
WS567	-0.0172	-4.926	0.0000	1.04
ESTPRT	0.7033	9.397	0.0000	1.21
DNDTIND	0.5033	11.228	0.0000	1.31
QEWCEN	-0.2275	-3.559	0.0004	1.09
DNBAY	-0.5449	-5.334	0.0000	1.03
Adjusted R2	0.3307			

5.4.2 Residual analysis

The residuals of the model were tested to confirm the random distribution of the residuals. The study performed a spatial autocorrelation test using Geoda 9.5i. The *Moran's I* statistics is found to be 0.05 with a p value of 0.4. The scatter plot of the test is shown in Figure 0.32 in Appendix B. In addition, the distributions of the residuals are shown in Figure 0.28 in Appendix B.

5.4.3 Surface generation

Figure 5.7 shows the land use regressed surface of SO_2 pollution in the Hamilton CMA when the wind direction is south-west, which is a the dominant direction of wind in the area. When the wind blows from south-west, it normally transports the pollution from the lakeshore industries to the outer part of region. Although the residential area in the lower city is not affected when wind blows south-west, the Burlington Bay and areas around the Skyway Bridge are greatly exposed to the high SO_2 pollution (24-37 ppb). On the other hand, the upper city which is often called the Mountain area experience fairly lower SO_2 pollution levels. Figure 5.8 shows a detail map of Sulphur Dioxide distribution for SW wind direction.

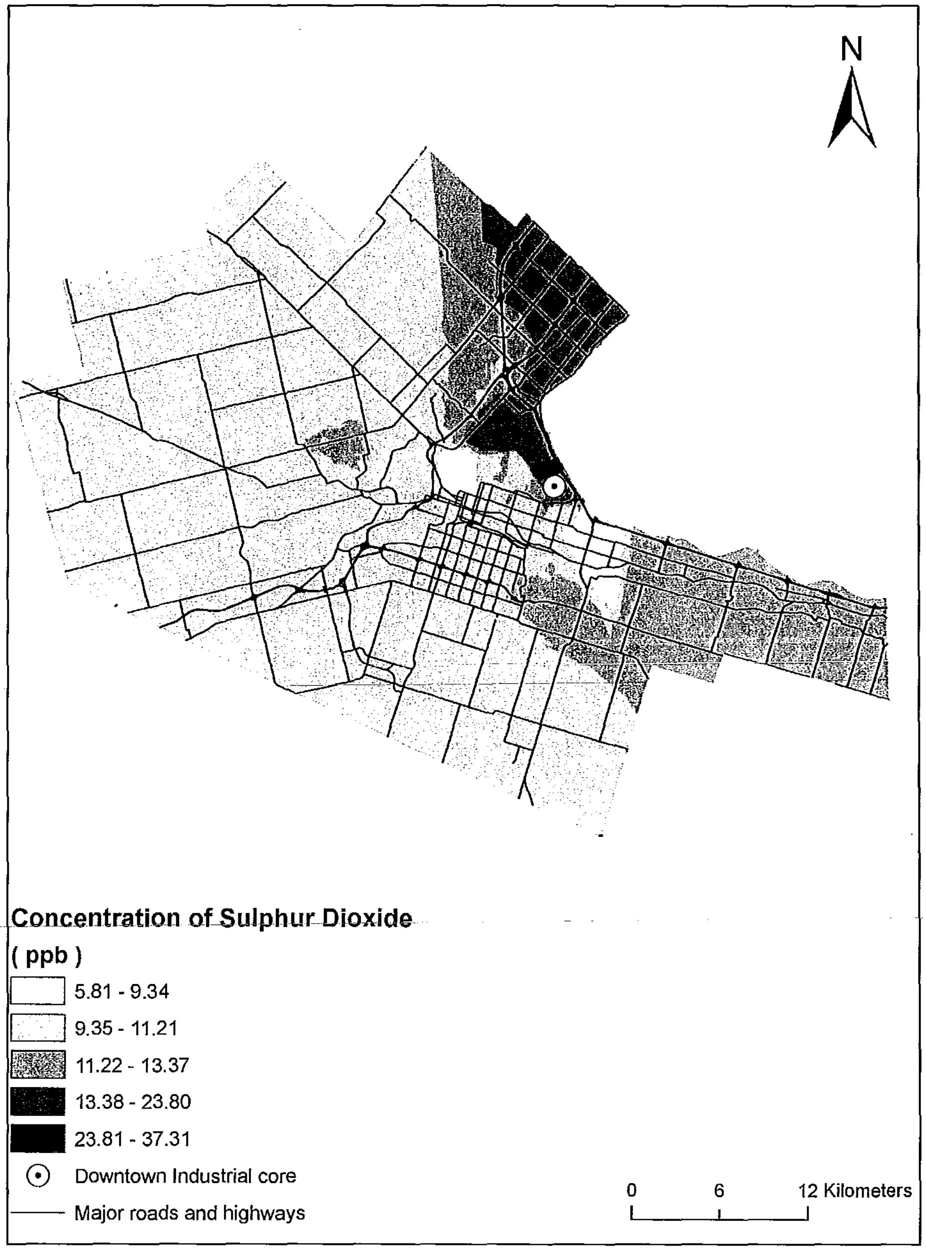

Figure 5.7 Sulphur Dioxide distribution in the Hamilton CMA for SW wind direction

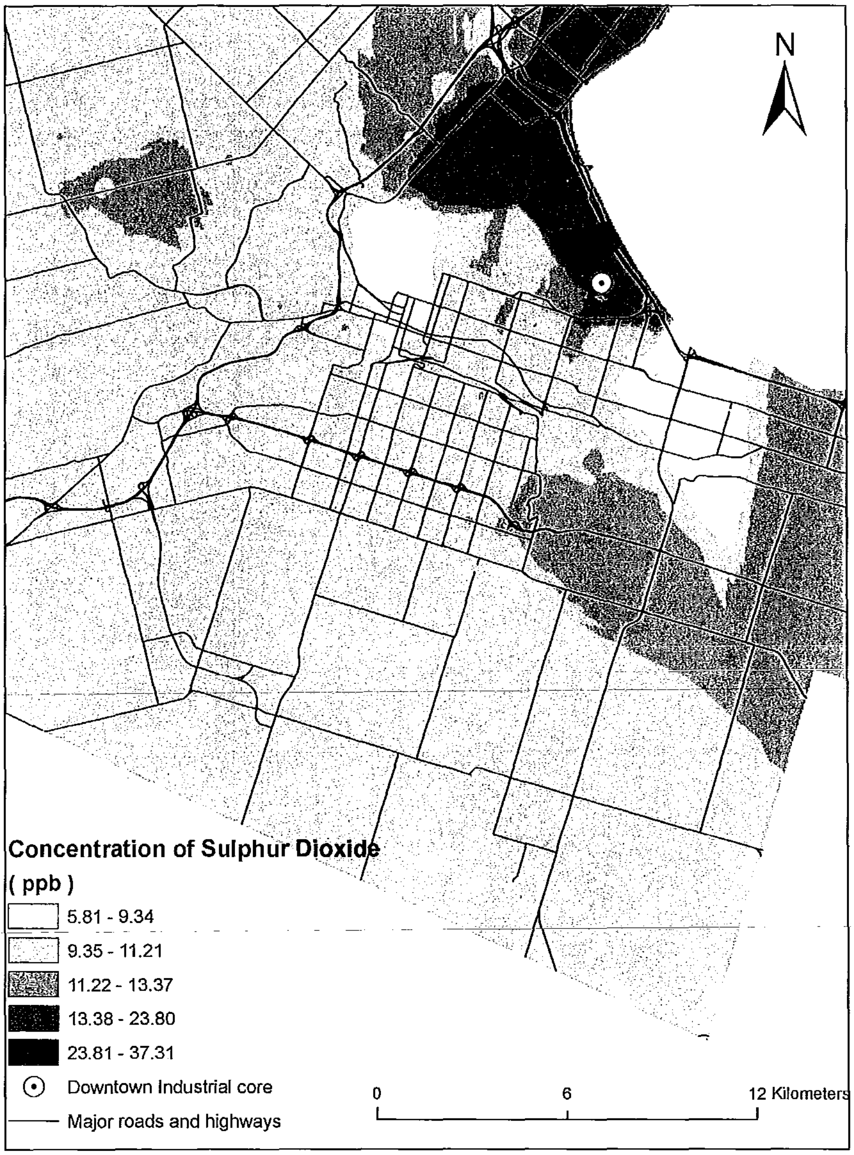

Figure 5.8 Sulphur Dioxide distribution around downtown area (corresponds to observation points) for SW wind direction

5.5 *Summary and validation*

The SO_2 pollution intensity and the distribution change with the wind direction. When the wind direction is from north-east it has been found that the lower city and central business district suffer most in comparison to the upper city. For north-west wind direction, the south-east part of the city, particularly areas around Burlington Street and QEW, and Centennial Parkway North and QEW are highly affected. However, the pollution concentration at Hamilton downtown is fairly low. With south-west prevailing winds the pollution is moved out from the city center. Burlington Bay, Skyway Bridge and southern part of the Lake Ontario are mostly affected for this wind direction. Lastly, for the south-east wind direction, the Burlington Bay and areas near Waterdown are the most affected regions. In general, the SO_2 pollution is profound near the east and north-east part of the lower city compared to west part and the upper city.

In addition to four wind direction-specific models a single model with all direction inclusive were estimated and presented in Table 3, Appendix B. It shows that the four wind direction-based model perform better than the aggregated model in terms of goodness-of-fit and variability to be explained.

The research also validated the results with the observed pollution concentration obtained from the Ministry of Environment (MOE), Ontario. Table 5.13 shows observed and predicted concentration of SO_2 in the downtown area.

Table 5.13 Comparison of the predicted and observed pollution concentration for SO_2

SO_2 Concentration			Average	Min.	Max.
North-east	Observed	Downtown location	3.90	0	76
	Predicted	Downtown location	4.97	----	----
	Predicted	Hamilton CMA (whole)	1.14	1	14.84
North-west	Observed	Downtown location	3.17	0	71
	Predicted	Downtown location	6.31	----	----
	Predicted	Hamilton CMA (whole)	5.94	4	24
South-east	Observed	Downtown location	4.92	0	49
	Predicted	Downtown location	7.17	----	----
	Predicted	Hamilton CMA (whole)	14.53	5.37	31.14
South-west	Observed	Downtown location	4.23	0	73
	Predicted	Downtown location	6.47	----	----
	Predicted	Hamilton CMA (whole)	11.17	5.81	37.31

Although the predicted values are close to the observed pollution concentration for the MOE downtown weather station for four major wind directions, it slightly overestimates level of concentration for all wind directions. Since this study mostly used observation data that are close to the sources (for example, steel industries), the estimated values are higher than the observed values. But the models provides a general sense of accuracy for all four wind directions.

6. CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS

6.1 *Conclusion*

The research employs land use regression models to characterize the concentration of two air pollutants, NO_2 and SO_2, in the Hamilton CMA. It looks into the spatial relationship between urban air pollutants (NO_2 and SO_2) and surrounding land uses, transportation factors and major activities, with a special consideration of wind parameters, wind direction and wind speed. The research uses a land use regression approach and generates pollution surfaces that show the spatial patterns of NO_2 (which is marker of traffic related pollution) and SO_2 (which is a major industrial pollutants) for different wind directions.

Since these air pollutants have significant impact on human health, it is very important to understand distribution of pollution over space. Although many studies used land use regression model to derive spatial distribution of pollutants, most of them ignored the effect of wind direction that might generate biases in the estimated pollution concentration. A key contribution of this research is that it explicitly incorporates wind directions into the models in addition to common attributes used in previous studies. Four wind directions considered in this research are north-east, north-west, south-east and south-west for both pollutants. Separate models were run for each wind direction. The models were significant and the parameters of the variables used for model estimation are statistically significant at least at the 5% level of significance.

The research also uses a unique dataset collected though a mobile field monitoring system, which represents pollution measurements at the one point in time at a given location as opposed averaged pollution concentrations used in many earlier studies. It provides a unique opportunity to include temporally varying explanatory variables (such as wind speed and direction) into the models. In addition, spatial investigations of SO_2 are rare in literature. This study attempts to characterize spatial distribution of this pollutant, which is a major problem in the Hamilton CMA.

The results for NO_2 models suggest that proximity to the major roads and highways, total length of roads nearby, traffic volume, nearby park and commercial areas, distance to the lakeshore industries and wind speed have significant contribution to the concentration of the pollutants. The four NO_2 models for north-east, north-west, south-east and south-west wind directions explain 23.5%, 24.35%, 26.96% and 14.19% of the variability in the concentration level respectively, with statistically significant regression coefficients. The models reveal that there are variations in the concentration of the pollutant, depending on the location. In general, while proximity to the roads and highways and higher total length of roads nearby increases the concentration of NO_2, higher amount of nearby park areas and wind speed decreases the level of concentration of the pollutant. In terms of wind directions, for north-east wind, distance to major roads and highways, lakeshore industries, and wind speed has significant impact on the concentration level. While for the north-west direction, commercial area within 350 meter, distance to QEW and Centennial Parkway North intersection contribute significantly to the concentration of NO_2, total length of the roads within 50 meter from the observation points and distance to the railway are significantly associated with the concentration level in case of south-east wind direction model. In south-west wind direction model, the significance of a dummy variable indicates that the upper city has lower concentration of NO_2 compared to the lower city.

In addition to proximity to the traffic-related attributes and other locational factors, the level of concentration is also varied by the prevailing direction of the wind. It is found that locations that are downwind to the potential sources are exposed to higher pollution compared to the upwind location points. As such, it is found that when the wind blows from the north-east, the lower city or areas around the CBD is mostly affected, whereas when the wind flows from south-west, the outer city is mainly affected. While south-east wind carries the pollution to the Lake Ontario, the north-west wind pushes the pollution to the south-east part of the city. While the predicted average concentration of NO_2 for the Hamilton CMA is 5.92 ppb with a minimum 1ppb and maximum 54 ppb for the north-east wind direction, average concentration is 8.61 ppb for the north-west wind direction (min, 1.74 ppb and max, 57.91 ppb). On the other hand, average concentration of the pollutant is 9.86 ppb (min, 1 ppb and max, 33.71 ppb) for the south-east wind, and 12.32 ppb for the south-west wind direction with minimum of 1.31 ppb and maximum of 32.21 (see Table 4.13).

The results are validated against observed pollution data at the downtown area. Since most of the observation points that are used in this study are near downtown and/or lakeshore industries, it is expected that a downtown weather station monitoring data will provide a valid test for prediction accuracy. It is found that the models generate very close estimate of the pollution concentration at the downtown location in comparison to observed concentration.

Since it is observed that distance from highway, distance from lakeshore industries, traffic volume, park area around the observation points and wind speed play an important role in increasing the concentration levels of NO_2, introducing the carpool policy, improving transit system and by encouraging the use of transit system as well as preserving the green areas might be useful tools to reduce the NO_2 pollution level.

This research also investigates concentration level of SO_2, which is primarily an industrial emission and causes severe problems for the Hamilton CMA. Similar to NO_2, four regression models were estimated to observe the spatial pattern of the pollutant over Hamilton. The estimated models explains 34.43%, 35.55%, 33.46% and 33.07% variability for north-east, north-west, south-east and south-west wind directions respectively.

Overall the results suggest that proximity to downwind locations from the industries, proximity to commercial and park areas, and proximity to the Bay Front Area play a role to the concentration of the pollutant. While downwind location with closeness to the industries increases the level of concentration of SO_2, higher wind speed and proximity to the Bay Front Area reduces the level of pollution in general. All models provide a good model fit in terms of coefficient of determination (R^2). Almost all parameters included in the models are statistically significant at least at the 5% level of significance. The estimated models are used to generate pollution surfaces for the Hamilton CMA.

The spatial distribution of SO_2 is mostly influenced by the location of steel and other industries. Particularly downwind locations from the lakeshore industrial core are exposed to high pollution levels. The results also reveal that the distribution of concentration is very much source-dependent for the pollutant SO_2 . Surfaces suggest that for south-west winds, the higher pollution occurs at the outer city compared to the inner city. On the other hand, when the wind direction is from south-east, the higher pollution is found near Lake Ontario and lakeshore areas. In general, northern wind carries the pollution towards the city. If the wind blows from north-east, the lakeshore industries play crucial roles in this regard. Under such conditions mostly the downtown part of the city experiences heavier SO_2 pollution. When the wind flows from north-west, the south-east part, specifically the areas close to lakeshore industries, face intense pollution of SO_2. Overall, the models predicts that average concentrations of SO_2 are 1.17 ppb, 5.94 ppb, 14.53 and 11.17 ppb for north-east, north-west, south-east and south-west wind directions respectively for the Hamilton CMA.

The predicted concentrations are validated using Ministry of Environment, Ontario data for a downtown weather station. The results show a reasonable accuracy for all four wind directions. Since distance from the steel industries, lakeshore industrial core, amount of park areas mostly influence the concentration of SO_2, it might be useful to take mitigation measures, including reducing the use of coal-fired plants to reduce SO_2 pollution problem.

6.2 Recommendations for future work:

This research investigates concentration of two major pollutants NO_2 and SO_2. Particularly, it examines relationships between these pollutants with the land uses, transportation and other locational factors and meteorological conditions using a land use regression method. Following are the recommendations, which might be worth considering for future research:

1. In addition to testing other explanatory variables (such as observed traffic count), this research can also be extended using rush-hour data and compare with these non-rush hour based estimated models.

2. This research mostly used observations points near or around downtown areas and lakeshore industries, it will be better to estimate models using data that covers entire Hamilton CMA.

3. Other methods of investigating spatial and temporal distribution of pollution concentration should be investigated, including dispersion modeling

4. Finally, the results obtained from this study could be used for exposure studies and health impact assessments.

References

Anderson, W.P., Kanaroglou, P.S., Miller E.J., Buliung, R.N. (1996), "Simulating automobile emissions in an integrated urban model", *Transportation Research Record,* **1520,** pp. 71-80

Anselin, L., Syabri, I., and Kho., Y. (2004), "Geoda: An Introduction to Spatial Data Analysis", Working paper, Department of Agricultural and Consumer Economics, University of Illinois, Urbana Champaign.

Arain, M.A., Blair, R., Finkelestein, N., Brook, J.R., Sahsuvaroglu, T., Beckerman, B., Zhang, L., Jerret, M. (2007), "The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies", *Atmospheric Environment,* **41,** pp. 3453-3464.

Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., Brunekreef, B., (2003), "Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems", *Epidemiology,* **14(2),** pp. 228-39.

Brauer, M. , Hoek, G., van Vliet P, Meliefste K, Fischer P, Wijga A, Koopman, L.P., Neijens, H.J., Gerritsen, J., Kerkhof, M., Heinrich, J., Bellander, T. and Brunekreef, B. (2002) "Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children", *American Journal of Respiratory Critical Care Medecine,* **166(8),** pp.1092-1098.

Briggs, D. Collins, S. Elliot, P. P. F., Kingham, S., Lebert, E., Pryl, K., Van Reeuwijk, H. Smallbone, K. and Van der veen, A. (1997) "Mapping urban air pollution using GIS: a regression based approach", *International Journal of Geographical Information Sciences,* **11,** pp. 699-718.

Briggs, D., de Hoogh, C., Gulliver, J., Wills, 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", *Sciences of Total Environment*, **253**, 151-167

Clean Air Hamilton (2007), "2005-2006 Progress Report", Online resource, URL: http://www.cleanair.hamilton.ca/downloads/CAH-Report-2005-2006-FINAL.pdf

Degaetano, A.T and Doherty, O. M. (2004). "Temporal, spatial and meteorological variations in hourly PM 2.5 concentration extremes in New York City", *Atmospheric Environment*, **38 (11)**, pp. 1547-1558.

Durham, W.H., (1974), "Air pollution and student health", *Archives of Environmental Health*, **28**, pp. 241-254

Fenger, J. (2004) *Air pollution Science for the 21st Century*, Edited by Austin, J., Brimblecombe,P., Sturges, W., Elsevier Ltd.

Ferdous, F. and Kanaroglou, P. (2007), "Land Use Regression Modeling for Predicting Ambient Concentration of Sulphur Dioxide: An Application to Hamilton, Ontario, Canada, Presented in the 54th Annual North American Meetings of the Regional Science Association International, Savannah, Georgia, Nov 07-10.

Finkelstein, M., M., Jerrett, M., and Sears, M.R. (2005). "Environmental inequality and circulatory disease mortality gradient", *Journal of epidemiology and Community Health*, **169**, pp. 397-402

Finkelstein, M.,M., Jerrett,M., Deluca, P., Finkelstein, N., verma, D.K., Chapman, K. and Sears, M.R. (2003). "A cohorot study of income, air pollution and mortality", *Canadian Medical Association Journal*, **169**, pp. 397-402

Gilbert, N.L.; Goldberg, M.S.; Beckerman, B.; Brook, J.; Jerrett, M. (2005), "Assessing Spatial Variability of Ambient Nitrogen Dioxide in Montreal, Canada, with a Land Use Regression Model", *Journal of Air & Waste Management. Association*, **55**, pp. 1059-1063.

Goodchild, M.F., Parks, O.B., Steyaert, L.T. (1993), Environmental Modelling with GIS, New York, Oxford University Press

Greene, W.H. (2002) *LIMDEP 8.0 Econometric Modeling Guide*, Plainview, New York: Econometric Software, Inc.

Grivas, G., Chaloulakao, A., Samara, C. & Spyrellis, N. (2004). "Spatial and temporal variation of PM10 Mass concentration within the greater area of Athense, Greece", *Water*, **158 (1)**, pp. 357-371.

Henderson, S.B., Beckerman, B., Jerrett, M., Brauer, M., (2007) "Application of Land use regression to estimate lomg-term concentrations of traffic related nitrogen oxides and fine particulate matters", *Environment, Science Technology*, **41**, pp. 2422-2428

Hoek, G., Brunrkreef,B., Goldbohm.S., Fischer,P., & Van den Brandt, P.A. (2002) "Association between mortality and indicators of traffic –related air pollution in the Netherlands: a cohort study", *The Lancet*, **360 (9341)**:pp 1203-1209

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N. & Finkelstein, M.M. (2007), "Modelling the intra-urban variability of ambient traffic pollution in Toronto, Canada", *Journal of Toxicology and Environmental Health*, **70**, pp. 200-212

Lin. S., Munsie, J.P., Hwang, S.A., Fitzgerald, E. and Cayo, M.R., (2002) "Childhood Asthma Hospitalization and Residential Exposure to State Route Traffic" Environmental Research, **88(2)**, pp. 73-81

Maoh, H and Behan, K.(2006), IMULATE: Integrated Model for urban land use , transportation and environmental analysis, McMaster University.

Martuzevicius, D., Grinshpun, S.A., Reponen, T., Gorny, R.l., Shukla, R., Lockey , J., Hu,S., McDonald,R., Biswas, P., Kliucininkas, L. & LeMasters, G. (2004). "Spatial and temporal variations of PM2.5 concentration and composition throughout an urban area with high freeway density - the Greater Cincinnati study" *Atmospheric Environrment*, **38** (8): 1091-1105.

Matthias Hochadel,M, Heinrich,J.,Gehring,U., Verena Morgenstern, V., Kuhlbusch,T., Link,E., Wichmann,H.E., Kra¨mer, U, (2005), "Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information", *Atmospheric Environment*, **40**, pp. 542-553

McGregor, G.R. (1996)," Identification of air quality affinity areas in Birmingham, UK", *Applied Geography*, **16(2)**, pp. 109-122

Mestayer. P. Alambuer, R. & Tchepel, O. (2003). Urban Field champaigns. In: N. Moussiopoulos (ed.), Air Quality in Cities. Berlin: Springer -Verlag, pp. 51-89.

Moore, D.K., Jerret, M. & Mack W.J. (2007), "A land use regression model for predicting ambient fine particulate matter across Loas Angeles, CA", *Journal of Environmental Monitoring*, **9**, pp. 246-252.

Muzio, L. J. and Quartucy, G.C.(1997), "Implementing NO_x control: Research to application, Progress in Energy and Combustion Science", **23 (13)**, pp. 233-266

Roosli, M., Theis, G., Kunzli, N. Staehelin,J., Mathys,P.Oglesby,L., Camenzind,M. & Braun-Fahrlander, Ch. (2001), "Temporal and Spatial variation of the chemocal composition of PM10 at urban and rural sites in the Basel area, Switzerland", *Atmospheric Environment*, **35 (21)**, pp. 3701-3713.

Rosenlund, M., Forastiere, F., Stafoggia, M., Porta, D., Perucci, M., Ranzi, A., Nussio, F., Perucci, C. (2007), "Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome", *Journal of Exposure Science and Environmental Epidemiology,* 1-8

Ross, Z. English P. B., Scale, R., Gunier, R. Smorodinsky, S.W. & Jerret M. (2006) "Nitrogen dioxide prediction in Southern California using land use regression modeling : potential for environmental health analyses", *Journal of Exposure Science and environmental Epidemiology,* **16**, pp. 106-114

Ross, Z., Jerrett, M., Ito, K.,Tempalski, B., and Thurston, G.D. (2007) "A land use regression for predicting fine particulate matter concentrations in the New York City region", *Atmospheric Environment,* **41(11)**, pp. 2255-2269

Rotko T; Oglesby L; Kunzli N; Carrer P; Nieuwenhuijsen MJ; Jantunen M. (2002). "Determinants of perceived air pollution annoyance and association between annoyance scores and air pollution (PM2.5, NO2) concentrations in the European EXPOLIS study", *Atmospheric Environment,* **36**, pp 4593-4602.

Sahsuvaroglu T. Arain, A. Kanaoglou P. Finkelstein, N. Newbold, B. Jerret, M. Beckerman, B. Brook, J. Finkelstein, M. and Gilbert, N.L. (2006), "A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada", *Journal of Air & Waste Management. Association,* **56**, pp. 1056-1069.

Schneider, J. (2006), Canadian steel town hopes for turnaround, too, Online resource, URL: http://www.iht.com/articles/2006/01/23/business/bxsteel.php

Shy, C.M., Hasselbad, V., Burton, R.M., Nelson, C.J. and Coles, A.A.,(1973) "Air pollution effects on ventilatoru function of US school- children", *Archives of Environmental Health,* **27**, pp. 124-128

Solaiman, T. (2007), Air pollution modelling and forecasting in Hamilton using data - driven methods, M. A. Sc. Thesis, Department of Civil Engineering, McMaster University, Hamilton, Canada

Srivastava, A (2004), "Source apportionment of ambient VOCS in Mumbai city", *Atmospheric Environment*, **38**, pp 6829-6843

Su, J.G, Brauer, M., Ainslie, B., Steyn, D., Larson, T., Buzzelli, M., "An innovation land use regression model incorporating meteorology for exposure analysis", *Science of Total Environment*, **390**, pp 520-529

The National Medical Association (2008), "The National Illness Costs of Air Pollution Summary Report ", The National Illness Costs of Air Pollution Study.

Vecchi, R., Marcazzan, G. VAlli, G., Ceriani, M. & Antoniazzi,C. (2004). "The role of atmospheric dispersion in the seasonal variation of PM1 and PM2.5 concentration and composition in the yrban area of Milan (Italy)", *Atmospheric Environment*, **38** **(27)**, pp. 4437-4446.

Wallace, J and Kanaroglu, P. (2008), "Modeling NOx and NO2 emissions from mobile sources: A case study for Hamilton, Ontario, Canada", *Transportation Research* part D, **(13)**, pp. 323-333

Wheeler, A.J., Doiron, M.S., 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**, pp 7-16

Appendix A

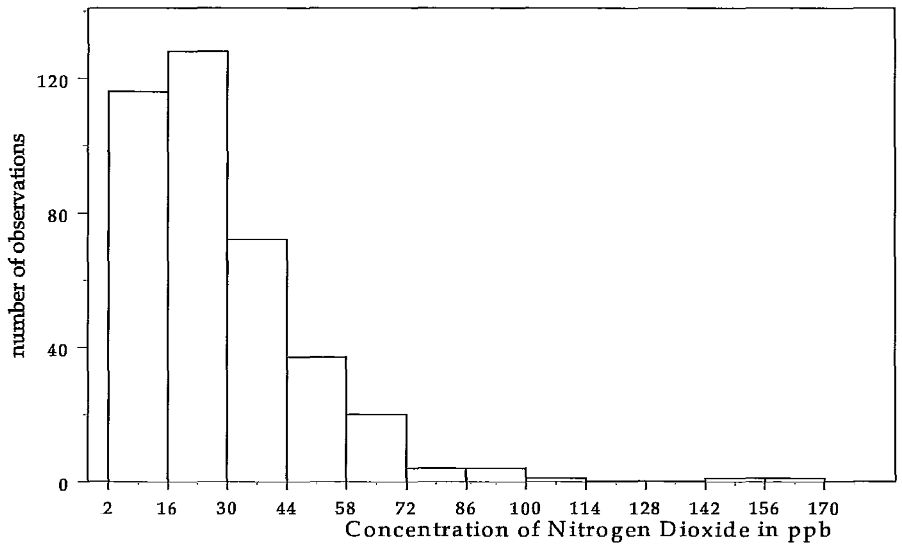

Figure 0.1 Distribution of NO2 level for the north-east wind direction at the Hamilton CMA

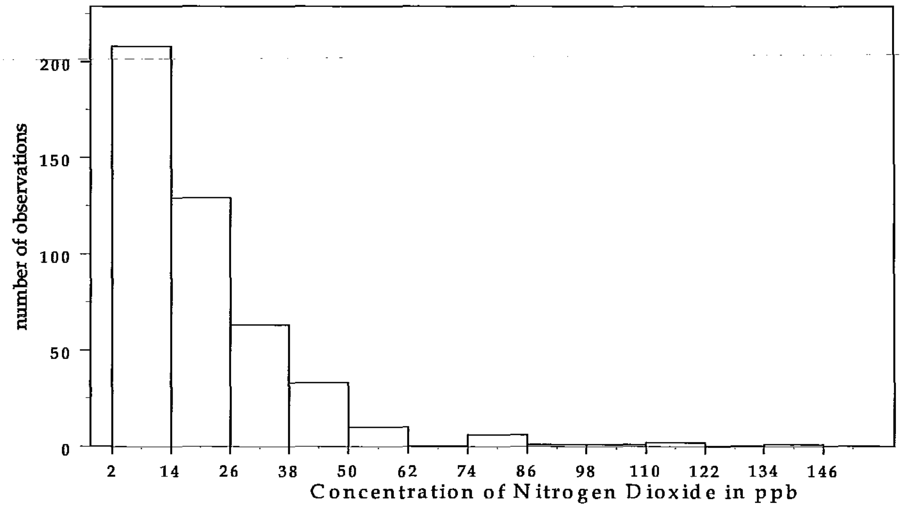

Figure 0.2 Distribution of NO2 level for the north-west wind direction at the Hamilton CMA

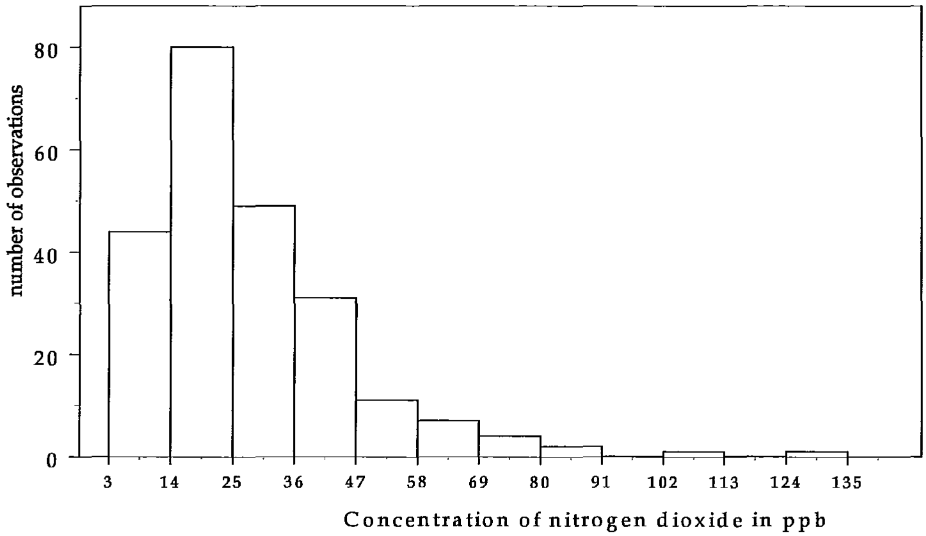

Figure 0.3 Distribution of NO2 level for the south-east wind direction at the Hamilton CMA

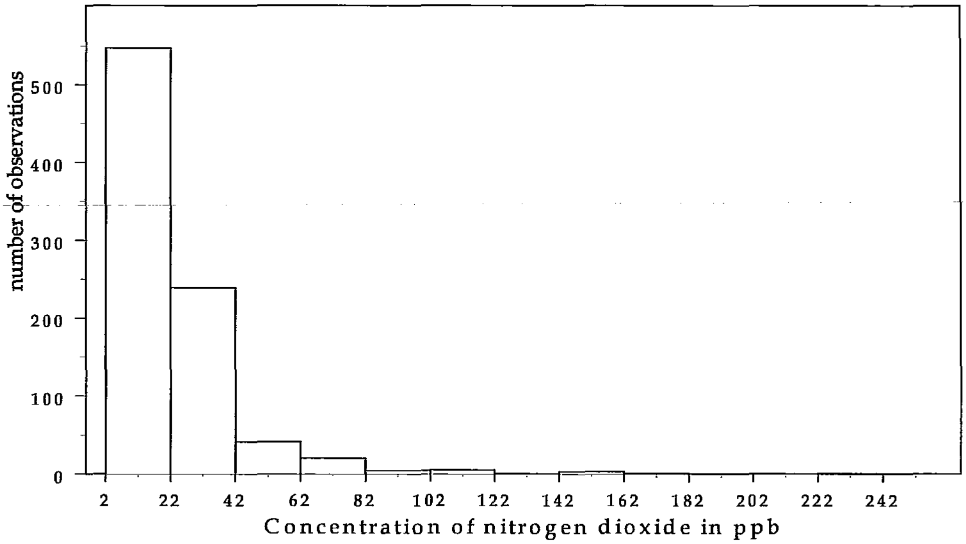

Figure 0.4 Distribution of NO2 level for the south-west wind direction at the Hamilton CMA

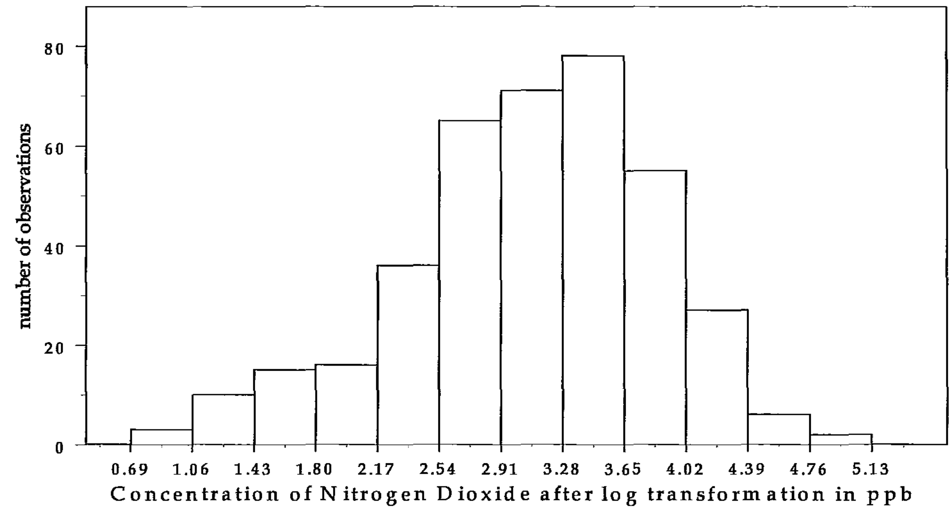

Figure 0.5 Distribution of log-transformed NO2 level for the north-east wind direction at the Hamilton CMA

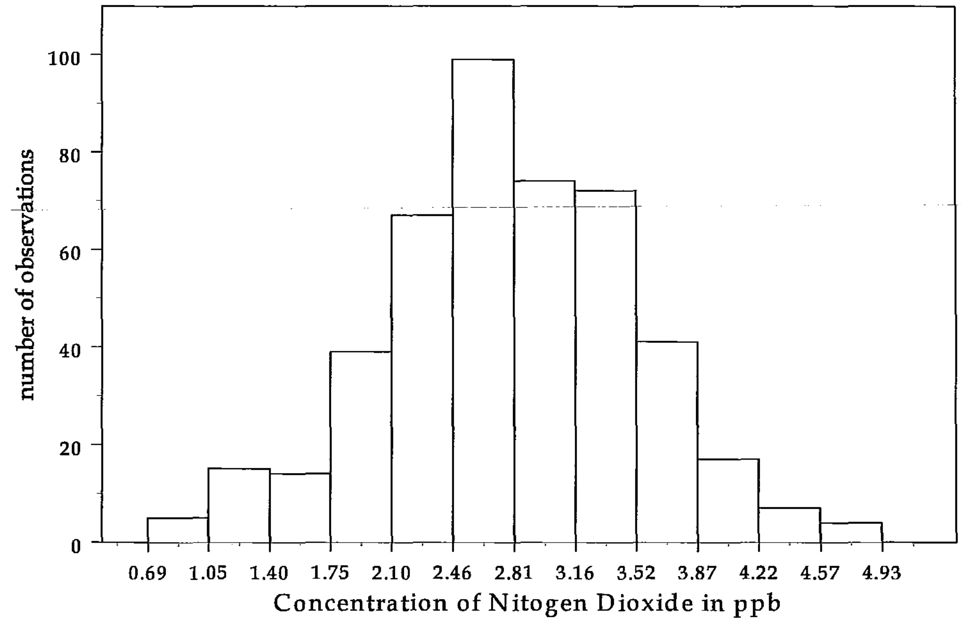

Figure 0.6 Distribution of log-transformed NO2 level for the north-west wind direction at the Hamilton CMA

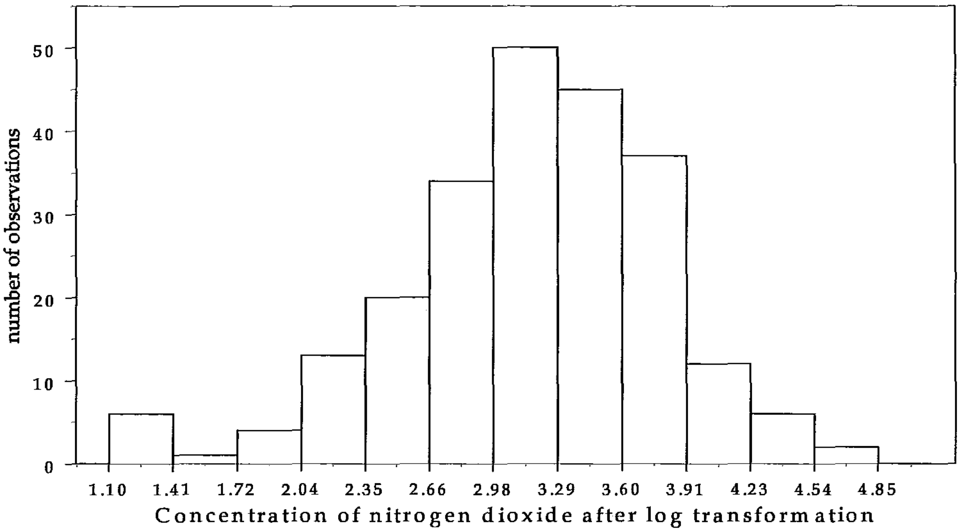

Figure 0.7 Distribution of log-transformed NO2 level for the south-east wind direction at the Hamilton CMA

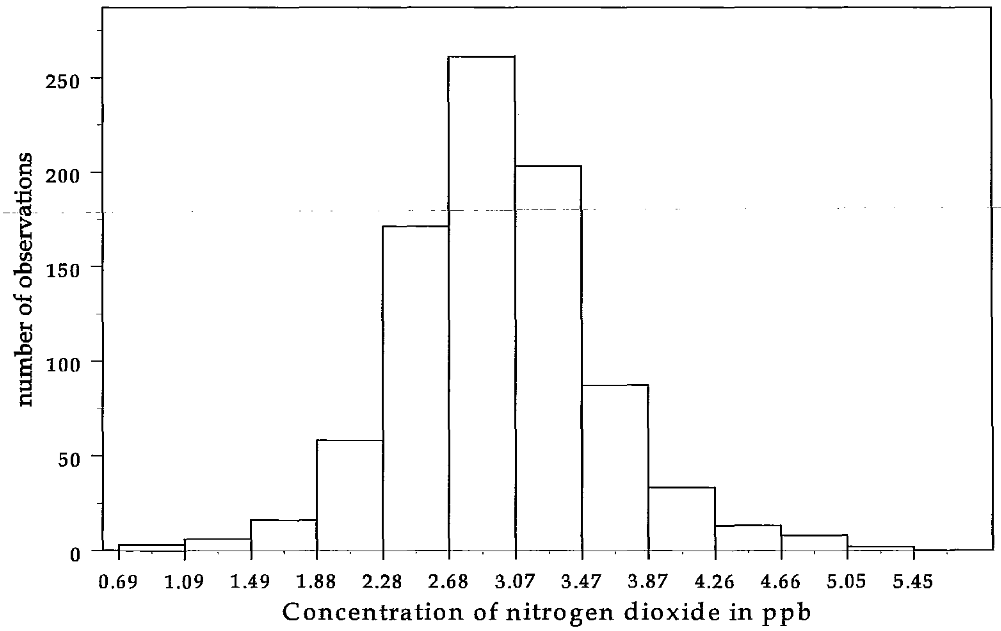

Figure 0.8 Distribution of log-transformed NO2 level for the south-west wind direction at the Hamilton CMA

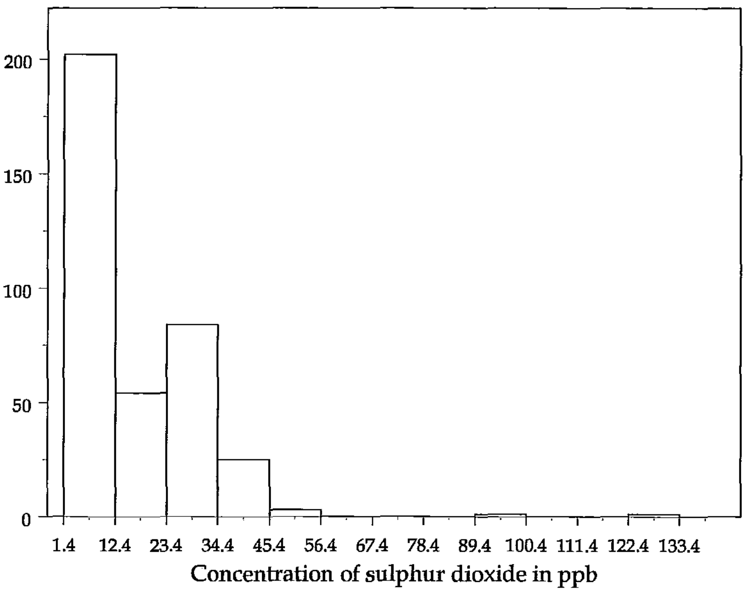

Figure 0.9 Distribution of SO2 level for the north-east wind direction at the Hamilton CMA

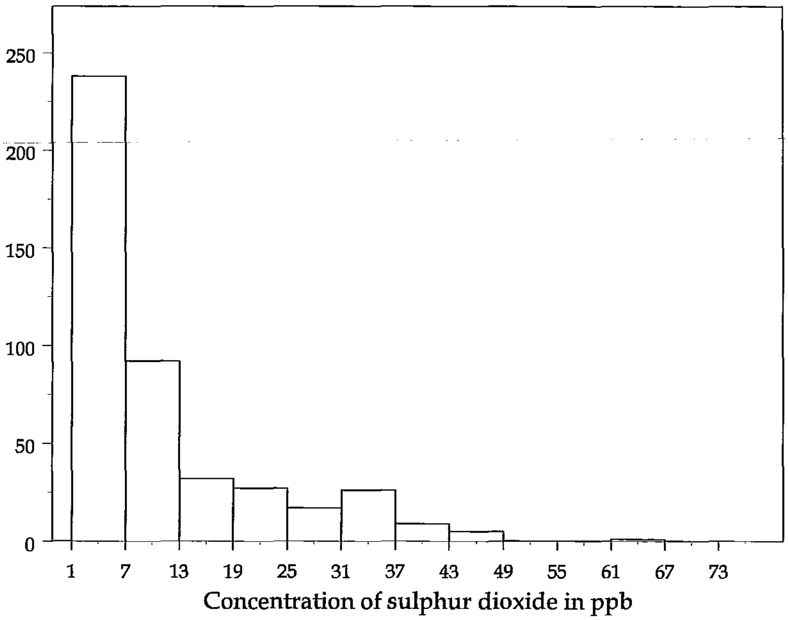

Figure 0.10 Distribution of SO2 level for the north-west wind direction at the Hamilton CMA

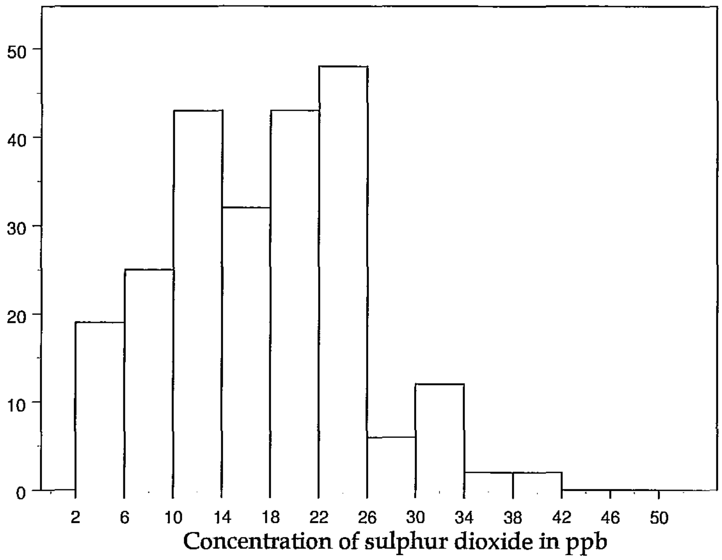

Figure 0.11 Distribution of SO2 level for the south-east wind direction at the Hamilton CMA

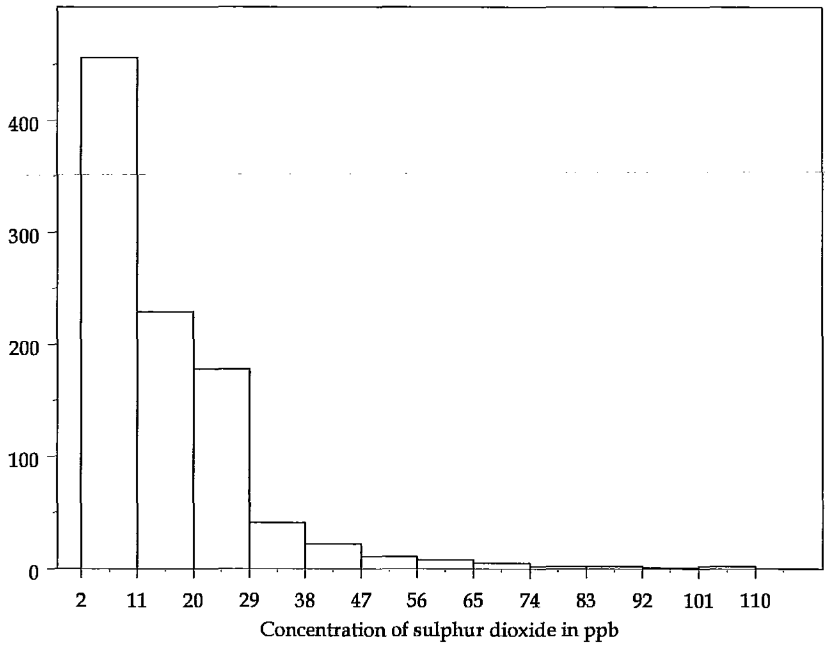

Figure 0.12 Distribution of SO2 level for the south-west wind direction at the Hamilton CMA

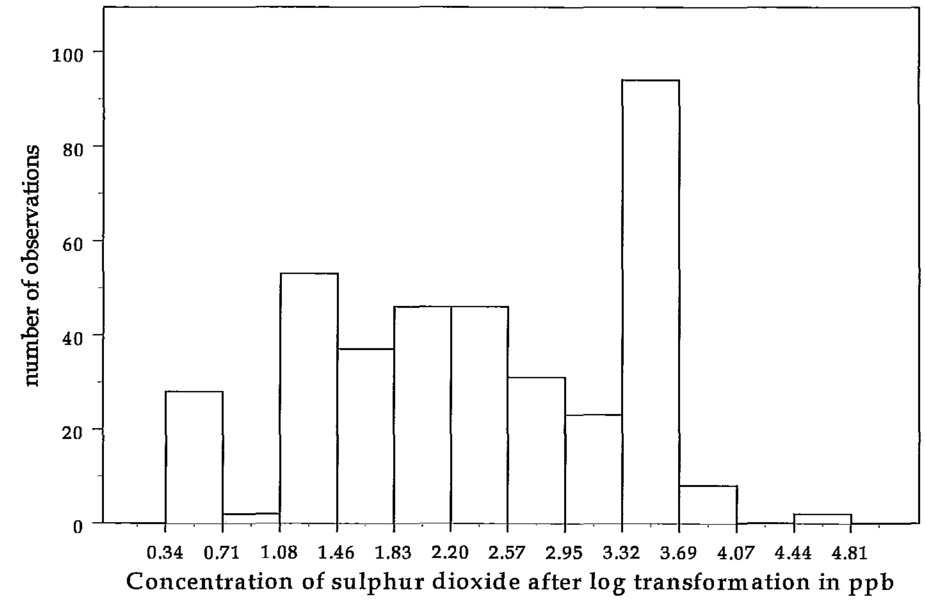

Figure 0.13 Distribution of log-transformed SO2 level for the north-east wind direction at the Hamilton CMA

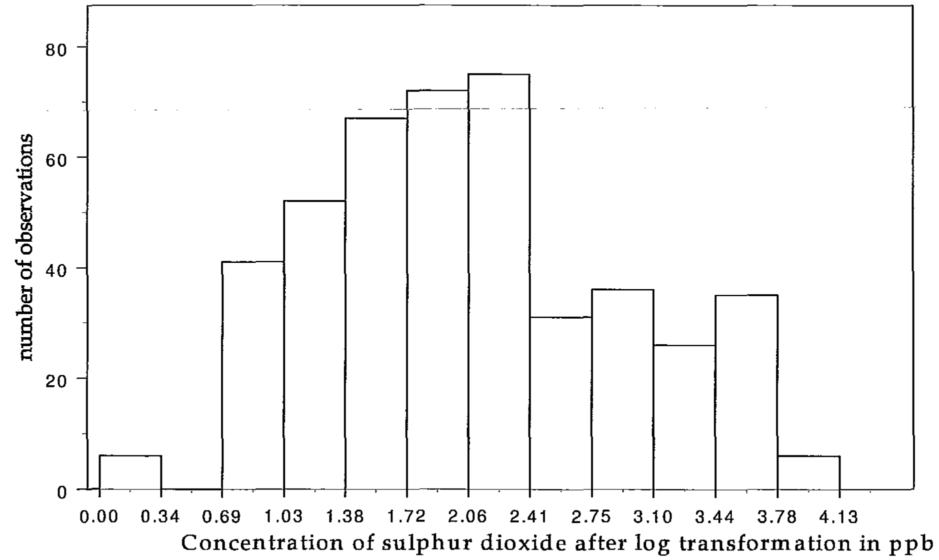

Figure 0.14 Distribution of log-transformed SO2 level for the north-west wind direction at the Hamilton CMA

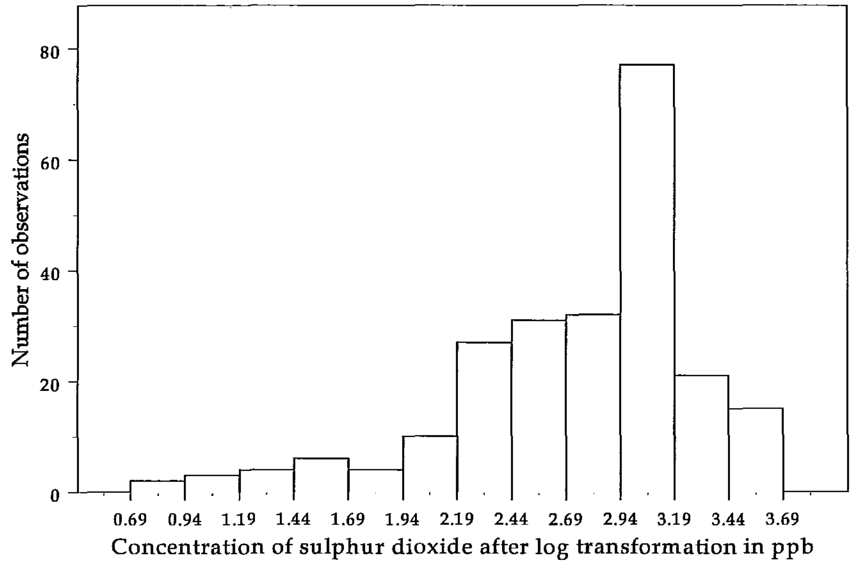

Figure 0.15 Distribution of log-transformed SO2 level for the south-east wind direction at the Hamilton CMA

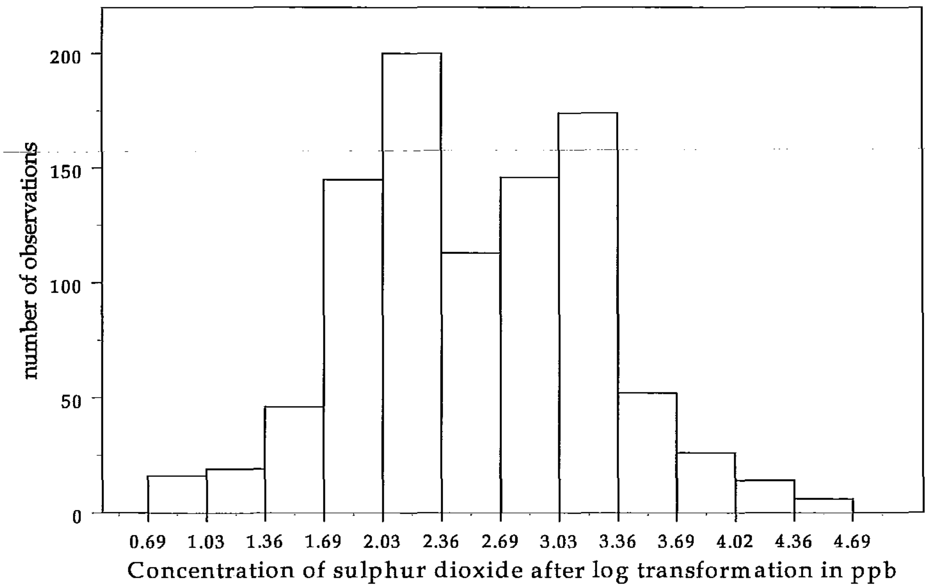

Figure 0.16 Distribution of log-transformed SO2 level for the south-west wind direction at the Hamilton CMA

Table 1: A comprehensive list of variables tested in the regression analysis

VARIABLE	TYPE	DEFINITION
WS567	Num	Wind Speed km/hr
STN29567WD1	Num	Wind Direction
COM50	Num	Commercial land use area within 50 m buffer from the monitoring station (m^2)
GOV50	Num	Governmental land use area within 50 m buffer from the monitoring station (m^2)
OPEN50	Num	Open land use area within 50 m buffer from the monitoring station (m^2)
PARK50	Num	Parks land use area within 50 m buffer from the monitoring station (m^2)
RES50	Num	Residential land use area within 50 m buffer from the monitoring station. (m^2)
IND50	Num	Industrial land use area within 50 m buffer from the monitoring station. (m^2)
WATER50	Num	Waterbody area within 50 m buffer from the monitoring station. (m^2)
COM100	Num	Commercial land use area within 100 m buffer from the monitoring station. (m^2)
GOV100	Num	Governmental land use area within 100 m buffer from the monitoring station. (m^2)
OPEN100	Num	Open land use area within 100 m buffer from the monitoring station. (m^2)
PARK10	Num	Parks land use area within 100 m buffer from the monitoring station (m^2)
RES100	Num	Residential land use area within 100 m buffer from the monitoring station (m^2)
IND100	Num	Industrial land use area within 100 m buffer from the monitoring station (m^2)
WATER100	Num	Waterbody area within 100 m buffer from the monitoring station
COM200	Num	Commercial land use area within 200 m buffer from the monitoring station (m^2)
GOV200	Num	Governmental land use area within 200 m buffer from the monitoring station (m^2)
OPEN200	Num	Open land use area within 200 m buffer from the monitoring station (m^2)
PARK200	Num	Parks land use area within 200 m buffer from the monitoring station (m^2)
RES200	Num	Residential land use area within 200 m buffer from the monitoring station. (m^2)
IND200	Num	Industrial land use area within 200 m buffer from the monitoring station (m^2)
WATER200	Num	Waterbody area within 200 m buffer from the monitoring station (m^2)
COM300	Num	Commercial land use area within 300 m buffer from the monitoring station (m^2)
GOV300	Num	Governmental land use area within 300 m buffer from the monitoring station (m^2)

OPEN300	Num	Open land use area within 300 m buffer from the monitoring station (m^2)
PARK300	Num	Parks land use area within 300 m buffer from the monitoring station (m^2)
RES300	Num	Residential land use area within 300 m buffer from the monitoring station. (m^2)
IND300	Num	Industrial land use area within 300 m buffer from the monitoring station (m^2)
WATER300	Num	Waterbody area within 300 m buffer from the monitoring station (m^2)
COM350	Num	Commercial land use area within 350 m buffer from the monitoring station (m^2)
GOV350	Num	Governmental land use area within 350 m buffer from the monitoring station (m^2)
OPEN350	Num	Open land use area within 350 m buffer from the monitoring station (m^2)
PARK350	Num	Parks land use area within 350 m buffer from the monitoring station (m^2)
RES350	Num	Residential land use area within 350 m buffer from the monitoring station (m^2)
IND350	Num	Industrial land use area within 350 m buffer from the monitoring station (m^2)
WATER350	Num	Waterbody area within 350 m buffer from the monitoring station (m^2)
HIWY	Num	Proximity to highway (m)
MJRD	Num	Proximity to major roads (m)
TRNLN	Num	Proximity to train line (m)
ABDN	Num	Proximity to the Aberdeen rail station (m)
PARK	Num	Proximity to park (m)
IND	Num	Proximity to industries (m)
CBDHAM	Num	Proximity to CBD (m)
GASSTN	Num	Proximity to gas station (m)
REGSHOP	Num	Proximity to Regional Shop (m)
OLOCSHOP	Num	Proximity to other local shop (m)
OSERV	Num	Proximity to other services (m)
PARKING	Num	Proximity to parking lot (m)
POPDENSITY	Num	Population density
HQEW	Num	Proximity to QEW (m)
H403	Num	Proximity to Highway 403 (m)
HLAP	Num	Proximity to Lincoln Alexander highway (m)
DTIND	Num	Proximity to Downtown industrial core (m)
RDL50	Num	Total lengths of the roads within 50 m buffer from the observation points (m)
RDL100	Num	Total lengths of the roads within 100 m buffer from the observation points (m)
RDL200	Num	Total lengths of the roads within 200 m buffer from the observation points (m)

RDL250	Num	Total lengths of the roads within 250 m buffer from the observation points (m)
RDL500	Num	Total lengths of the roads within 500 m buffer from the observation points (m)
RDL750	Num	Total lengths of the roads within 750 m buffer from the observation points (m)
RDL1000	Num	Total lengths of the roads within 1000 m buffer from the observation points (m)
TVOL50	Num	Traffic Volume within 50 m buffer from the observation point
TVOL100	Num	Traffic Volume within 100 m buffer from the observation point
TVOL200	Num	Traffic Volume within 200 m buffer from the observation point
TVOL250	Num	Traffic Volume within 250 m buffer from the observation point
TVOL300	Num	Traffic Volume within 300 m buffer from the observation point
TVOL350	Num	Traffic Volume within 350 m buffer from the observation point
TVOL500	Num	Traffic Volume within 500 m buffer from the observation point
TVOL750	Num	Traffic Volume within 750 m buffer from the observation point
DNEQEW	Binary	Downwind locations from QEW when the wind comes from north-east =1, 0 otherwise
DNW403	Binary	Downwind locations from highway 403 when the wind comes from north-west =1, 0 otherwise
DSW403	Binary	Downwind locations from highway 403 when the wind comes from south-west =1, 0 otherwise
TEMP	Num	Temperature during pollution measurement (celcius)
PRHUM	Num	Percentage of Relative Humidity during pollution measurement
SRADIATION	Num	Solar Radiation during pollution measurement
DNW	Binary	1 if wind direction is from N-W, 0 otherwise.
DSW	Binary	1 if wind direction is from S-W, 0 otherwise.
DNE	Binary	1 if wind direction is from N-E, 0 otherwise
INDDNE	Num	Distance from industries when the wind direction is from north -east. (m)
INDDSW	Num	Distance from industries when the wind direction is from south -west. (m)
INDDNW	Num	Distance from industries when the wind direction is from north -west. (m)
CBDDNE	Num	Distance from central business district when the wind direction is from North -East. (m)
HIWYDSW	Num	Distance from highway when the wind direction is from south -west. (m)
HIWYDNW	Num	Distance from highway when the wind direction is from north -west (m)

Appendix B

Table 2: Results of all wind-direction inclusive model for NO2

Variable	Coefficient	*t* statistics	*P* value
Constant	3.60286	64.19	0.00
WS567	-0.03439	-13.93	0.00
DTIND	-0.00005	-9.26	0.00
RDL50	0.00091	5.82	0.00
HIWY	-0.00003	-2.36	0.02
TRNLN	-0.00008	-3.13	0.00
DNE	-0.13782	-2.65	0.01
DNW	-0.19661	-3.65	0.00
DSW	0.00070	0.01	0.99
Adjusted R^2	0.187		

Table 2: Results of all wind-direction inclusive model for SO2

Variable	Coefficient	*t*-statistic	*P* value
Constant	2.79885	50.118	0
WS567	-0.01119	-3.714	0.0002
DNW	-0.85348	-9.748	0
DSW	-0.25009	-3.341	0.0008
DNE	0.33616	3.538	0.0004
INDDNE	-0.00035	-3.078	0.0021
INDDSW	0.00037	8.736	0
INDDNW	0.00049	5.064	0
CBDDNE	-0.0001	-8.838	0
HIWYDSW	-0.00014	-4.97	0
HIWYDNW	-0.00017	-4.142	0
PARK50	-0.00006	-3.271	0.0011
Adjusted R^2	0.2159		

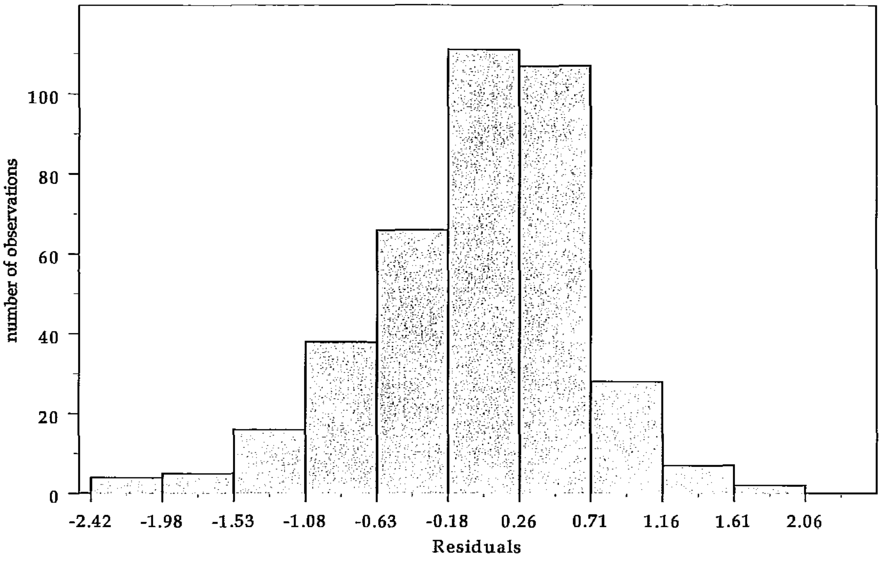

Figure 0.17: Distribution of the Residuals NO2 (NE)

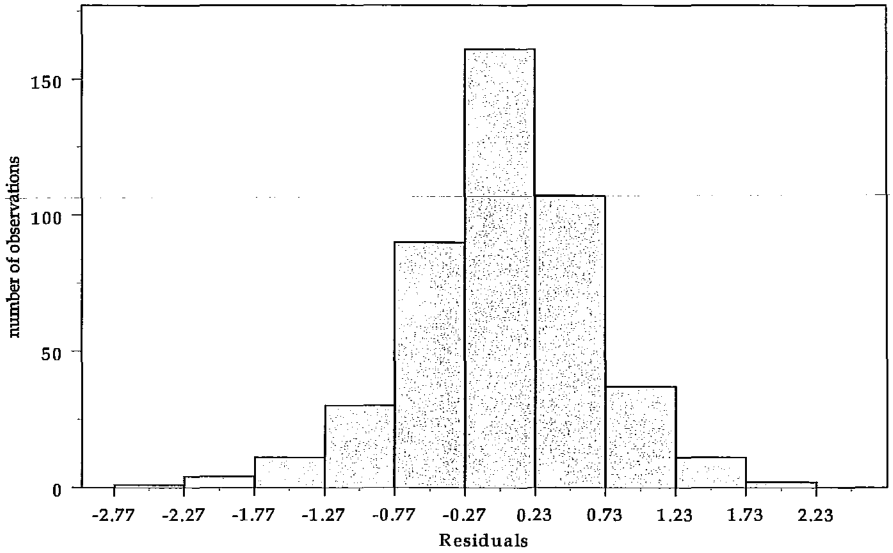

Figure 0.18: Distribution of the residuals of NO2 (NW)

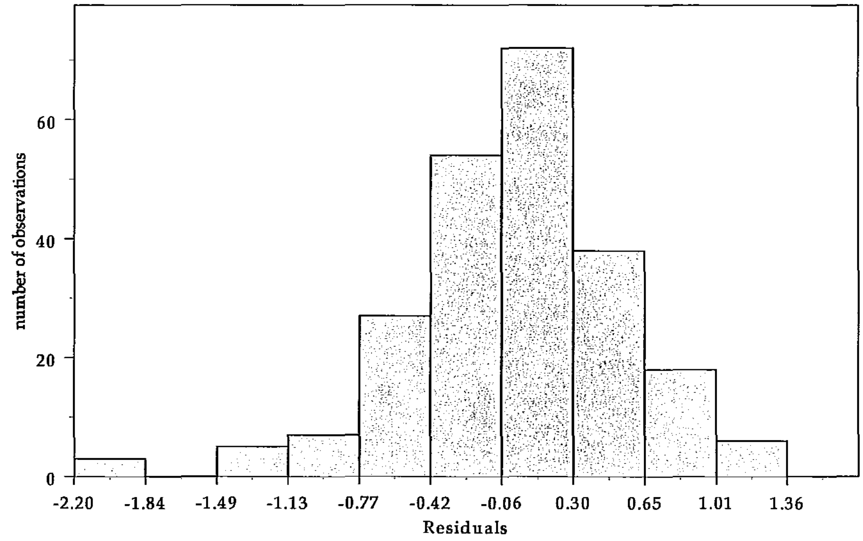

Figure: 0.19: Distribution of the Residuals of NO2 (SE)

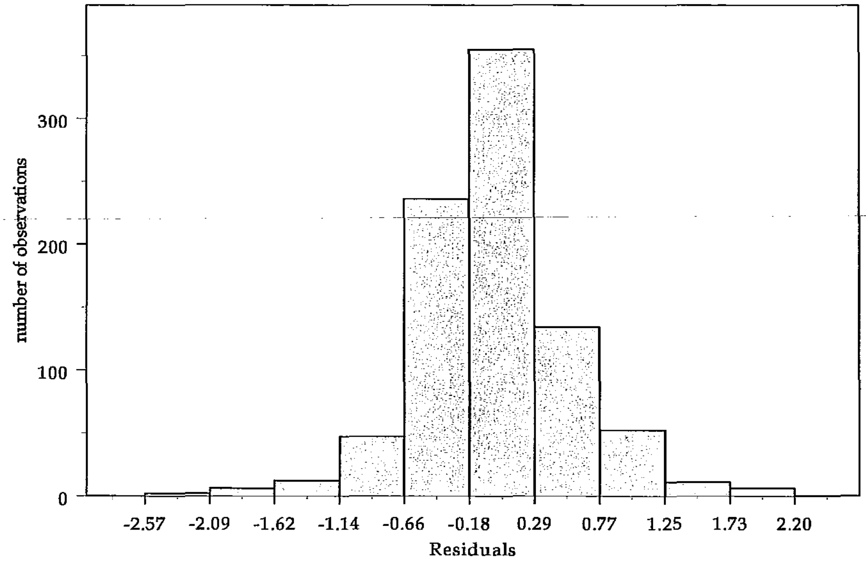

Figure 0.20: Distribution of the residuals NO2 (SW)

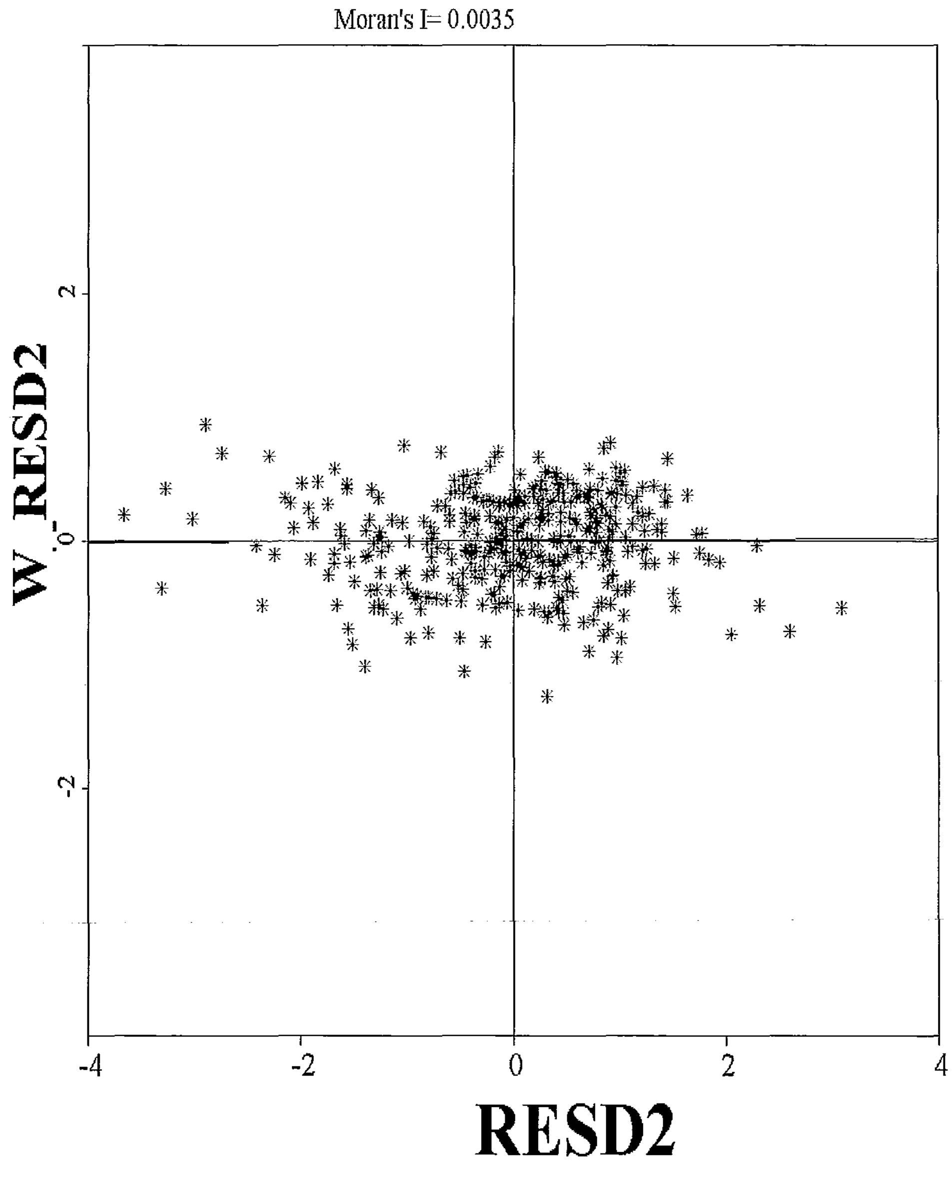

Figure 0.21: Moran's I Scatter plot, residual vs. spatial lag (NO2, NE model)

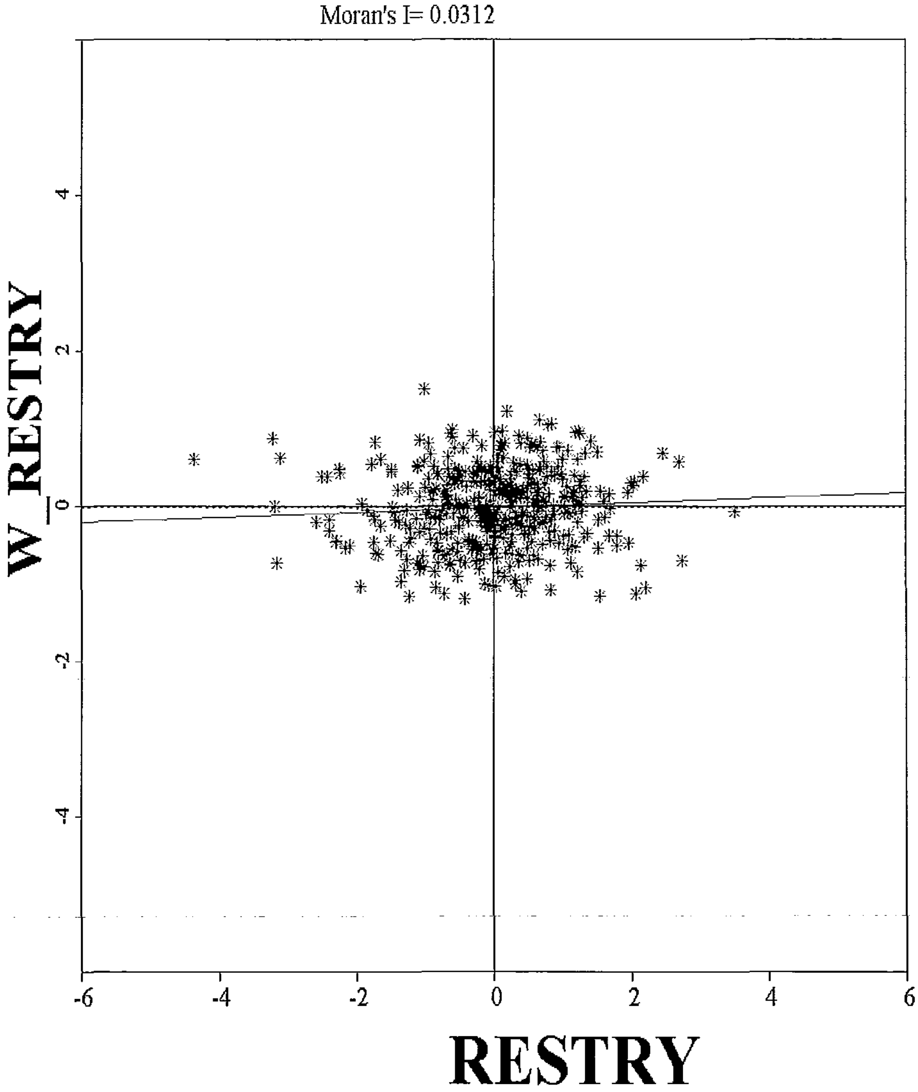

Figure 0.22: Moran's I Scatter plot, residual vs. spatial lag (NO2, NW model)

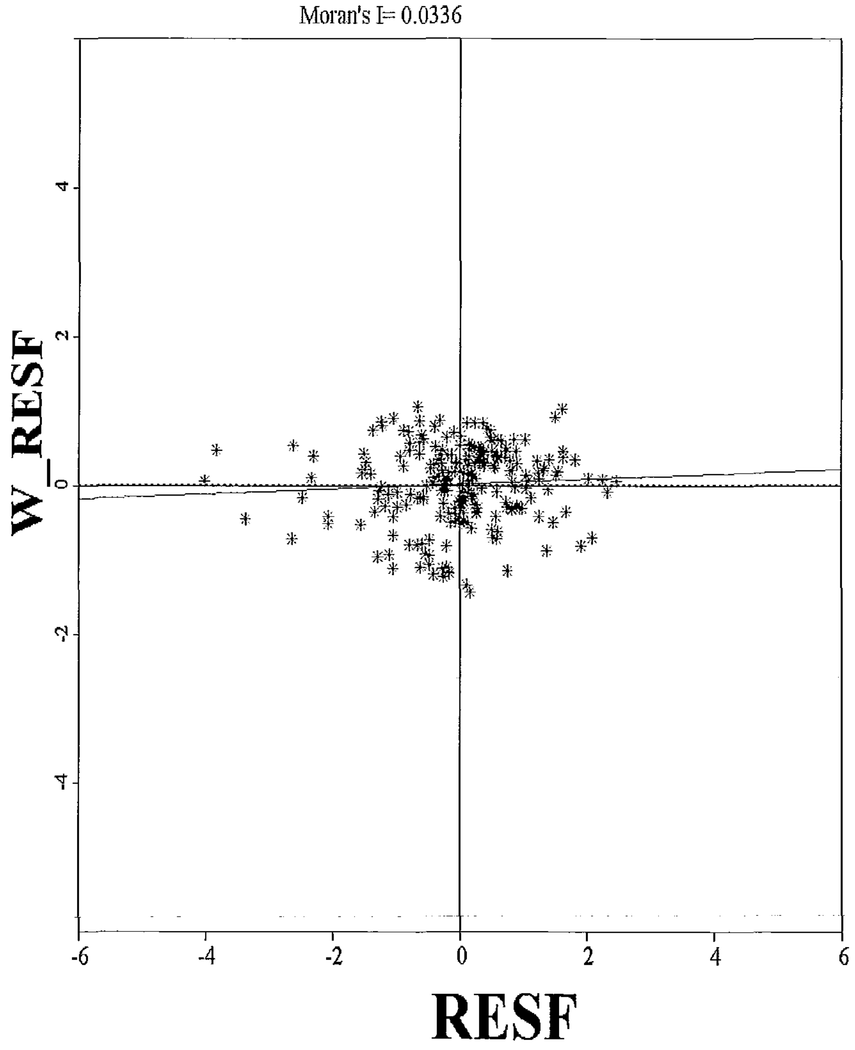

Figure 0.23: Moran's I Scatter plot, residual vs. spatial lag (NO2, SE model)

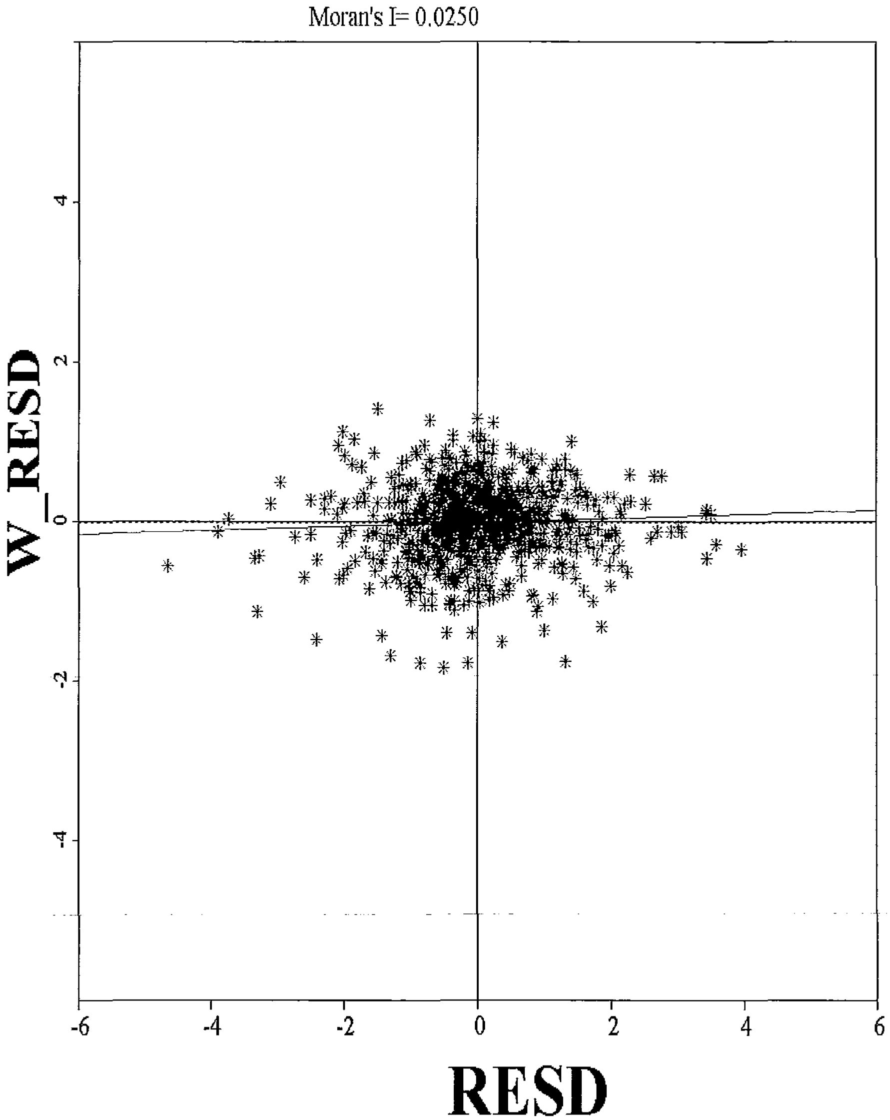

Figure 0.24: Moran's I Scatter plot, residual vs. spatial lag (NO2, SW model)

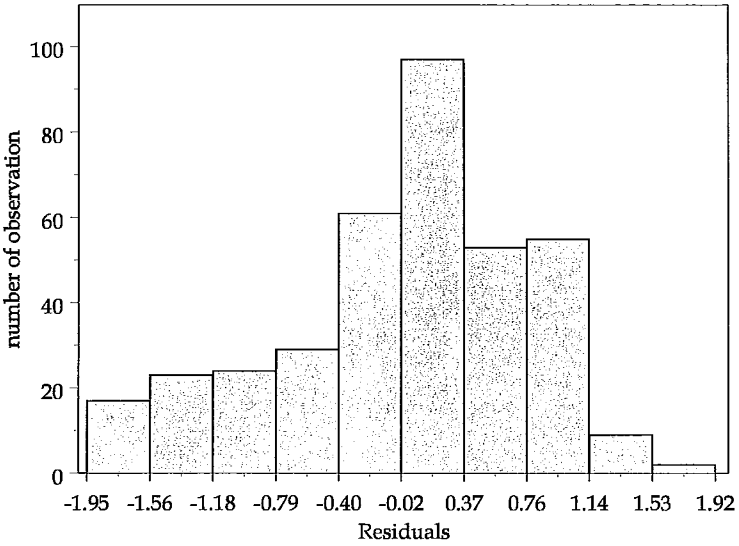

Figure: 0.25 Distribution of residuals of SO2 (NE)

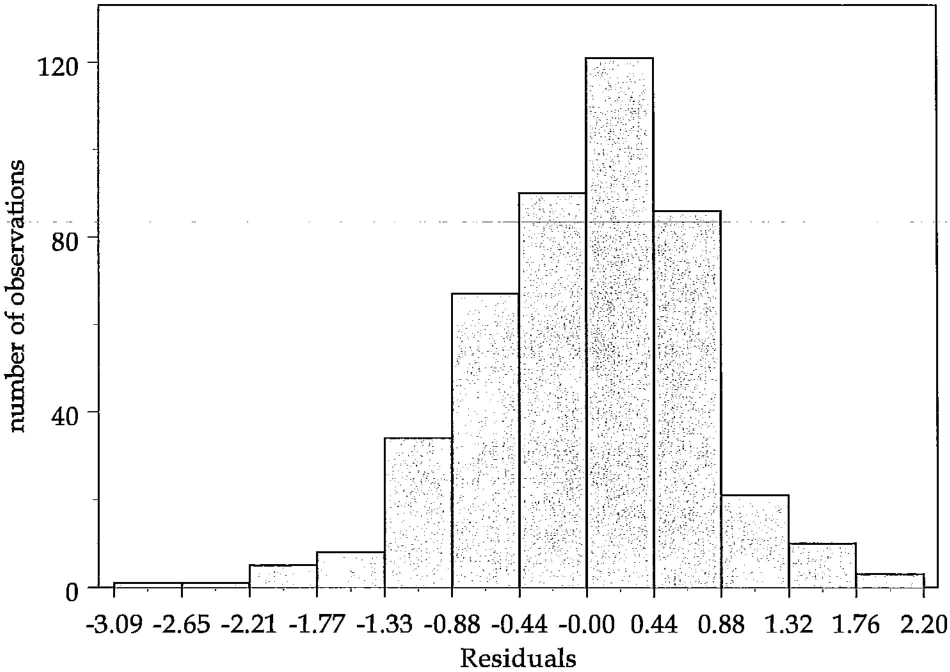

Figure 0.26: Distribution of residuals of SO2 (NW)

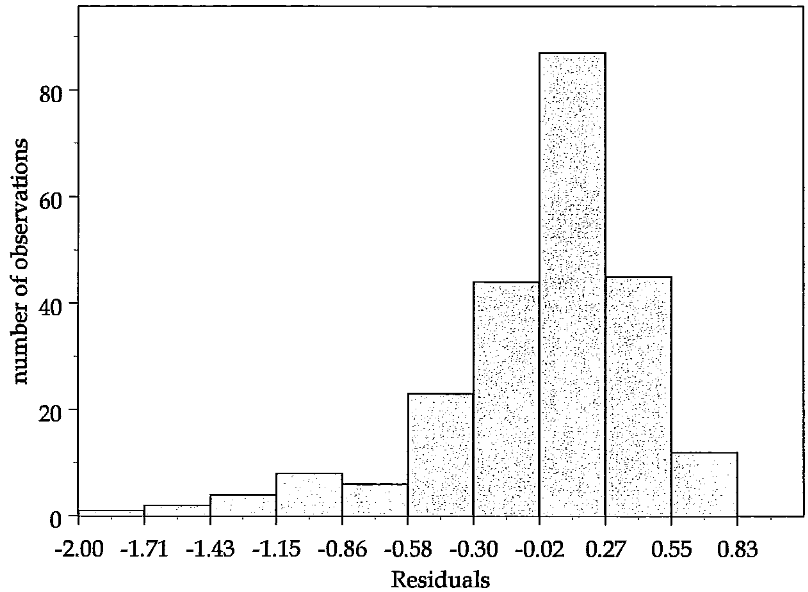

Figure 0.27 : Distribution of residuals of SO2 (SE)

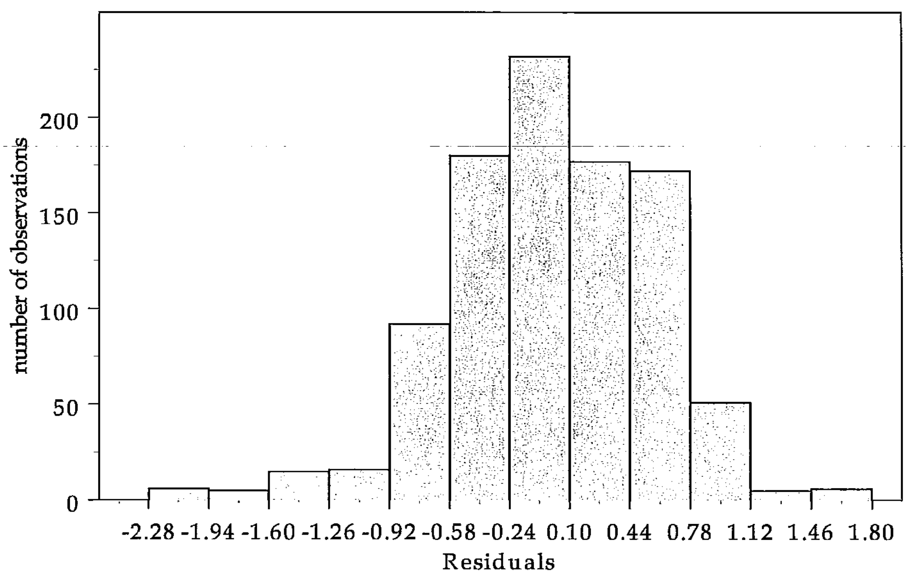

Figure 0.28 : Distribution of residuals of SO2 (SW)

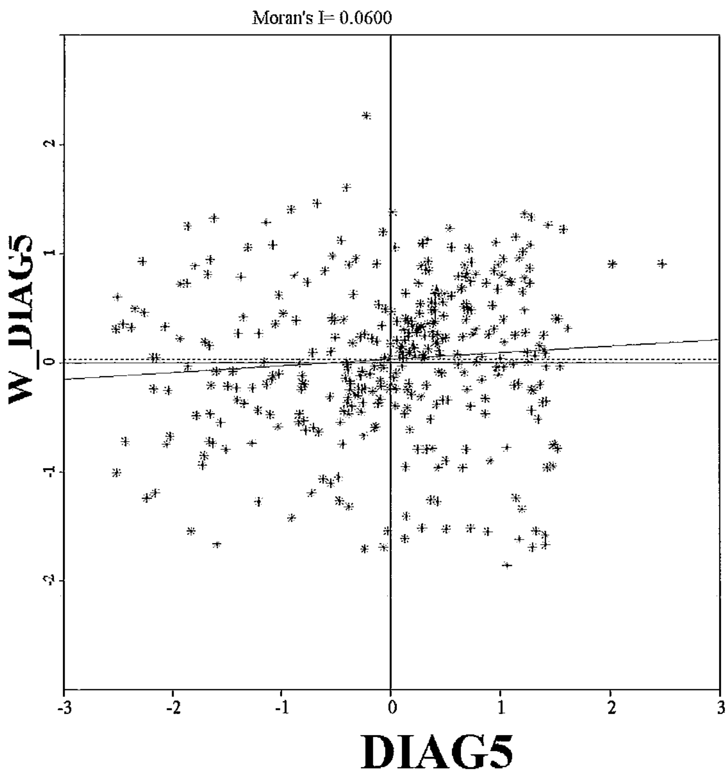

Figure 0.29 Moran's I Scatter plot, residual vs. spatial lag (SO2, NE model)

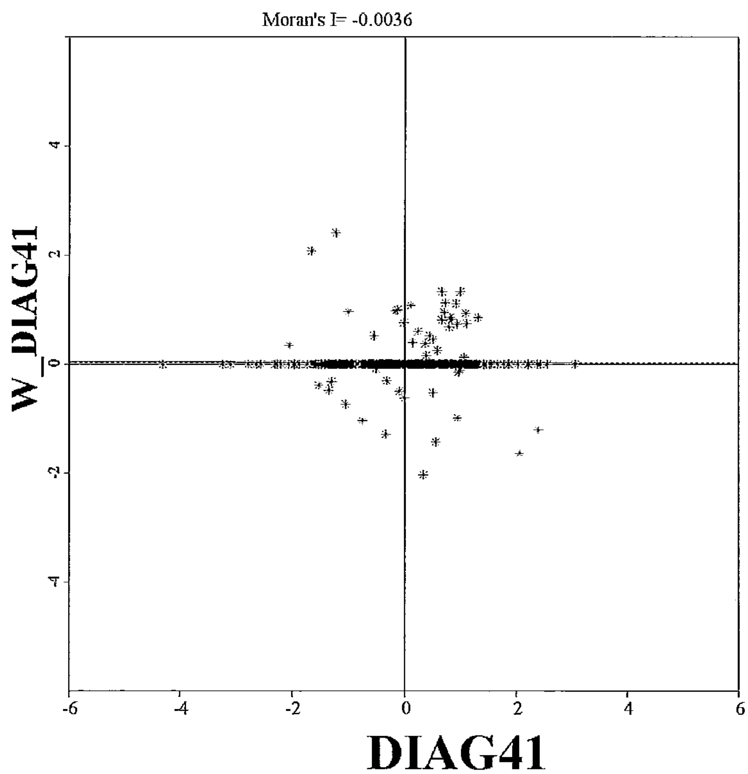

Figure 0.30: Moran's I Scatter plot, residual vs. spatial lag (SO2, NW model)

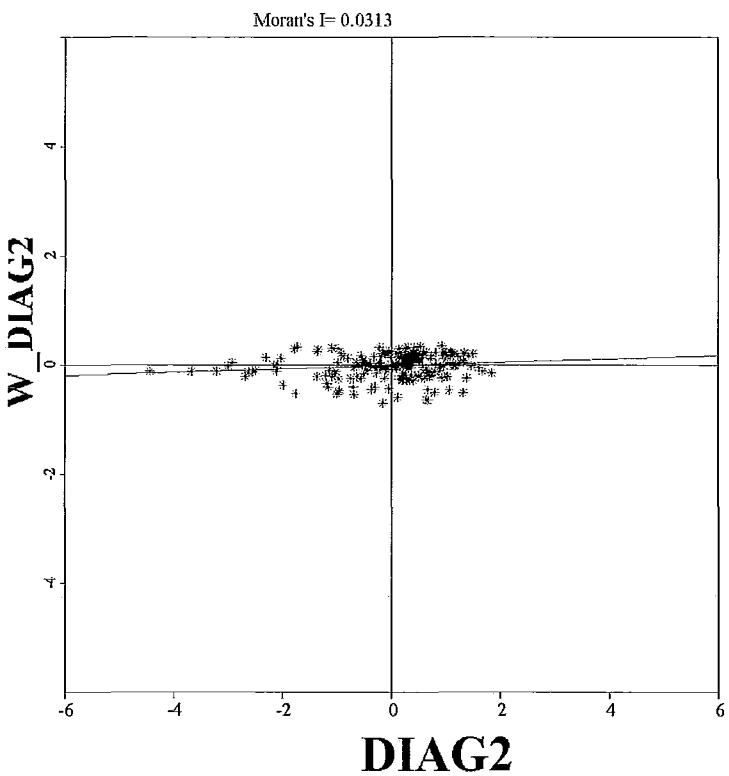

Figure 0.31: Moran's I Scatter plot, residual vs. spatial lag (SO2, SE model)

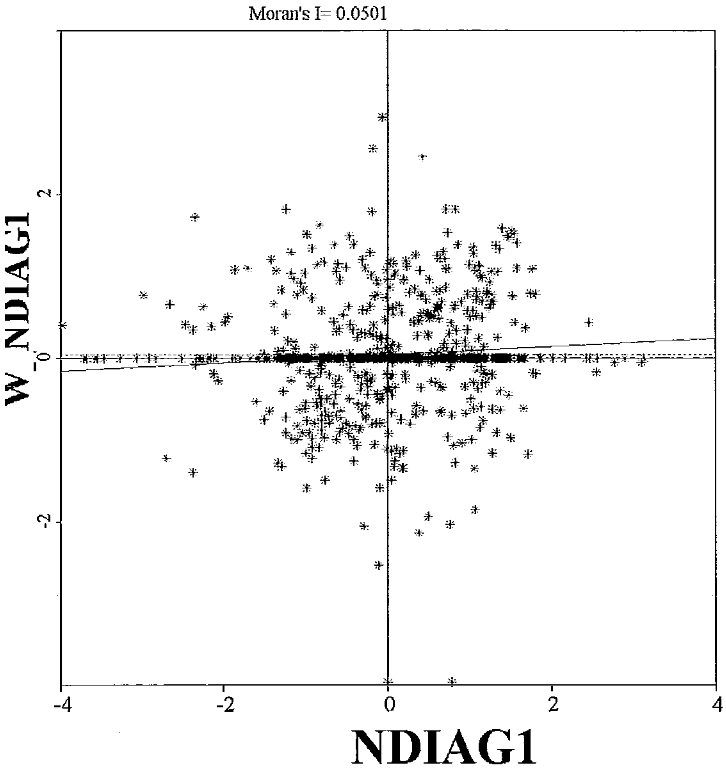

Figure 0.32: Moran's I Scatter plot, residual vs. spatial lag (SO2, SW model)