The geography of being born too small and spatiotemporal relationships with the outdoor environment

by

Charlene Chris Nielsen

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Departments of Earth and Atmospheric Sciences and Medical Sciences – Paediatrics
University of Alberta

Abstract

Background: The geography of where pregnant mothers live is important in furthering our understanding of babies born too small. Many causes are suspected to restrict growth or incur premature delivery, and include prenatal exposures from natural, social, and built habitats. More research is needed on the shared hazards of the outdoor environment. Studies have identified industrial activities for exposure, but few on the actual chemicals that include developmental toxicants.

In Canada, conditions of short gestation and low birth weight are the second leading cause of infant mortality, linked to adult chronic disease, and increasing. Also, in Canada, industrial emissions are documented by the National Pollutant Release Inventory (NPRI) and readily available for space-time analyses to better understand the contribution of ambient health hazards to babies born too small.

Objectives: (1) Examine distributions and patterns across Canada to identify associations with outdoor environment and differences among provinces and territories; (2) Develop a multi-hazard index for Alberta to determine whether mothers living in environments with a higher accumulation of outdoor hazards had more small newborns; (3) Contrast hot spots of non-critical and critically ill small newborns; and (4) Compare hot spots of critically ill small newborns with hot spots of industrial pollutants for major Canadian cities.

Methods: For all objectives, I classified birth records with maternal residential postal codes as small for gestational age (SGA) and low birth weight at term (LBWT). I accessed three secondary databases for all of Canada (Statistics Canada's Vital Statistics–Births Database, 2006-2012), the province of Alberta (Alberta Perinatal Health Program, 2006-2012), and nineteen major metropolitan areas (Canadian Neonatal Network, 2006-2010). Births from the latter represented critically ill SGA (ciSGA) and critically ill LBWT (ciLBWT).

I used publicly available spatial data on outdoor hazards (including transportation and energy infrastructure, and the NPRI chemicals emitted to air), to calculate potential exposure estimates using easily accessible spatial tools in a Geographical Information System (GIS). Main methods that I applied were kernel density, focal statistics, overlay, and space-time pattern mining, which involved creating space-time cubes from locations and birthdates, then statistically categorizing emerging hot spot patterns. To model monthly chemical emissions from the NPRI, I developed a simple wind-directed dispersion mapping technique. For the statistical associations, I used correlation and logistic regression, with covariates on area-level socioeconomic status (SES), land use, nitrogen dioxide (from an existing land use regression model), traffic, maternal age, migration, sex, urban, total number of births, and season.

Results: Provinces and territories showed higher percentages of SGA and LBWT where there were more industrial emissions; however, there were more associations with land hazards (dumps/waste depots, gas stations, powerlines, and transformer stations). For the province of Alberta, the chemical emission-weighted index was associated with SGA and LBWT, but individual chemical emissions provided complementary information, especially sub-provincially. The emerging hot spots identified where there were more small newborns in space and time and revealed scattered areas throughout Calgary and Edmonton. The non-critical and critically ill hot spots were not located in the same areas of each city. Low SES was associated with SGA and LBWT, industrial land use was associated with ciLBWT, and no associations with ciSGA. Among all major cities, the larger ones had more and larger areas of hot spots of ciSGA. Seventy-eight wind-directed industrial chemical hot spots were associated with ciSGA hot spots.

The greatest number of positive associations were observed for 28 different pollutants, mostly in Edmonton, Halifax, Montréal, Toronto, Vancouver, and Winnipeg. Of the identified industrial chemicals, many are suspected or known developmental toxicants, including ammonia, benzene, carbon monoxide, methyl ethyl ketone, particulate matter, heavy metals, and VOCs.

Conclusions: More industrial chemicals were discovered to be related to small newborns when the study areas were smaller (i.e. sub-provincial administrative units or metropolitan areas) and each area had a unique chemical signature. My focus was on the outdoor environmental habitat because it is a public source of exposures susceptible to regulation. I hope that my research may assist everyone connected to health – from medical professionals to policy makers – to understand potential impacts the environment has on early life, learn what location-based variables may be associated, inform the public that where they live is important to their future family health, and implement preventive interventions.

Preface

This interdisciplinary dissertation is the original work of Charlene Nielsen, carried out under the co-supervision of Dr. Carl Amrhein (Earth and Atmospheric Sciences) and Dr. Alvaro Osornio-Vargas (Pediatrics).

Chapter 2 will be submitted for publication as: Nielsen CC, Amrhein CG, Osornio-Vargas AR, and the DoMiNO Team. "A geographic information assessment of maternal ambient health hazards and babies born too small." I was responsible for study design, acquiring, preparing, and analyzing data (note: the birth data required securing access to the University of Alberta's Research Data Centre [RDC]), and writing the original manuscript; Dr. Amrhein and Dr. Osornio-Vargas guided the research question and study design, provided theoretical advice, and edited the manuscript.

Chapter 3 has been published as: Nielsen CC, Amrhein CG, Osornio-Vargas AR. 2017. "Mapping outdoor habitat and abnormally small newborns to develop an ambient health hazard index." Int. J. Health Geogr. 16:1–21; doi:10.1186/s12942-017-0117-5. I was responsible for study design, acquiring, preparing and analyzing data, and writing the original manuscript; Dr. Amrhein and Dr. Osornio-Vargas provided theoretical advice and edited the manuscript.

Chapter 4 will be submitted for publication as: Nielsen CC, Amrhein CG, Shah PS, Aziz K, Osornio-Vargas AR, Canadian Neonatal Network, and the DoMiNO Team. "Spatiotemporal patterns of small newborns and associations with land use and socioeconomic status." I was responsible for study design, acquiring, preparing, and analyzing data (note: the birth data required securing access to the Canadian Neonatal Network [CNN] coordinating center in Toronto), and writing the original manuscript; Dr. Amrhein, Dr. Shah, Dr. Aziz, and Dr.

Osornio-Vargas provided theoretical advice and edited the manuscript. Dr. Shah also provided access to the CNN data.

Chapter 5 will be submitted for publication as: Nielsen CC, Amrhein CG, Shah PS, Stieb DM, Osornio-Vargas AR, Canadian Neonatal Network, and the DoMiNO Team. “Space-time hot spots of critically ill small for gestational age and industrial air pollutants.” I was responsible for study design, acquiring, preparing, and analyzing data (note: the birth data required securing access to the Canadian Neonatal Network [CNN] coordinating center in Toronto), and writing the original manuscript; Dr. Amrhein, Dr. Shah, and Dr. Osornio-Vargas provided theoretical advice and edited the manuscript. Dr. Shah also provided access to the CNN data. Dr. Stieb also provided the weather station data.

This study is part of the Data Mining and Neonatal Outcomes (DoMiNO) project, CIHR/NSERC Funding Reference Number (FRN) 127789 entitled “Spatial data mining exploring co-location of adverse birth outcomes and environmental variables.” Ethics approval was obtained from the Research Ethics Board at the University of Alberta, ID Pro00039545 (see Appendix III – last page for convenience) and approval from the Alberta Perinatal Health Program (APHP) and the Canadian Neonatal Network (CNN) coordinating center in Toronto.

Dedication

To my dad, Terry – you were an important part of my character and are dearly missed.

To my mom, Linda – you are my inspiration.

To my husband and best friend, Scott – you are my life.

To my children, Natalie and Jorgen – you are my purpose.

Acknowledgements

I am extremely privileged to have had the mentorship, generosity, and friendship of Dr. Alvaro Osornio-Vargas and Dr. Carl Amrhein. Because of them, I have greatly grown and truly enjoyed this health geography project. It fit so perfectly with my dedication to defend the "health of the land" from the most important aspect of environmental health – pediatrics.

I am indebted to the collaborations of all the DoMiNO Team members: Dr. Jesus Serrano-Lomelin, for sharing his intelligence in quantitative analyses and epidemiology; Osnat Wine, for her amazing strengths in coordinating us all and reminding us about what is important; Shazan Jabbar, for his persistence and being my field-trip partner in Toronto; Dr. Colin Bellinger, for his excellent syntheses; Dr. Osmar Zaiane, for his foresight to co-investigate the DoMiNO Project and drill the concepts of spatial data mining and lift in to our very beings; Dr. Prakesh Shah, for his breadth of experience and providing data; Dr. Khalid Aziz, for his knowledge; Dr. Sue Chandra, for her wisdom; Dr. Manoj Kumar, for his insight; Dr. Dave Stieb, for his remarkable contributions to air pollution and health and for providing data; Dr. Perry Hystad, for sharing his aptitude in spatial epidemiology and providing land-use regression modelling data; Dr. Anders Erickson, for his keen perspectives; Susan Crawford, Nancy Aelicks, and Kendra Malainey, for their helpful explanations and providing data; Drs. Paul Villeneuve, Paul Demers, and Yan Yuan for their expertise in epidemiology/biostatistics and scientific advice; Erica Phipps, for her ability to view the process and results from all angles; Dr. Emily Chan, for paving the way for the rest of us students; Dr. Irena Buka, for her practical guidance; and our unofficial but irreplaceable team member, Dr. Deliwe Ngwezi, for her interconnected contributions to children's environmental health.

Special thanks to my committee members: Dr. Theresa Garvin, for her expertise in health geography and invaluable perspectives on the PhD journey, which have enriched my project and my life; Dr. Po-Yin Cheung, for his exceptional contributions to neonatal research. Both have enhanced my dissertation with their encouragement and constructive feedback. Extra thanks to the external examiners Dr. Zulfiqar Bhutta and Dr. Manish Shirgaokar for their valuable suggestions.

Additional people I wish to acknowledge are Dr. Dan Griffith, Sonny Yeh, Irene Wong, Ralph Young, Saeed Hojjati, Dr. Lauren Scott-Griffin, Dr. Majid Nabipoor Sanjebad, Dr. Laura Rodriguez, Bridget Rusk, Elliott Silverman, Dr. Anita Kozyrskyj and lab, Dr. Maria Ospina, and Larry Laliberté.

I also wish to acknowledge the CIHR/NSERC funding that made this project possible.

Finalmente, muchas gracias a mi familia.

Table of Contents

Chapter 1 Introduction 1

Chapter 2 Geographic information assessment of maternal ambient health hazards and babies born too small 14

2.1 Abstract 14

2.2 Background 15

2.3 Methods 17

2.4 Results 25

2.5 Discussion 35

Chapter 3 Mapping outdoor habitat and abnormally small newborns to develop an ambient health hazard index 39

3.1 Abstract 39

3.2 Background 40

3.3 Methods 44

3.4 Results 54

3.5 Discussion 73

Chapter 4 Spatiotemporal patterns of small newborns and associations with land use and socioeconomic status 80

4.1 Abstract 80

4.2 Background 81

4.3 Methods 83

4.4 Results ... 95

4.5 Discussion ... 109

Chapter 5 Space-time hot spots of critically ill small for gestational age newborns and industrial air pollutants ... 114

5.1 Abstract ... 114

5.2 Background ... 115

5.3 Methods ... 117

5.4 Results ... 128

5.5 Discussion ... 144

Chapter 6 Conclusion ... 149

References ... 158

Appendix I ... 174

Appendix II ... 192

Appendix III ... 218

List of Tables

Table 1.1. Selected studies involved in spatial associations among environmental variables and adverse birth outcomes: SGA=small for gestational age; LBW=low birth weight; BW=birth weight; IUGR=intrauterine growth restriction. Table continues. 10

Table 2.1. Definitions of variables and concepts important to the study of maternal ambient hazards with babies born too small. 19

Table 2.2. Outdoor environmental factors mapped for association with adverse birth outcomes. The time (year), spatial method, units, and source are indicated for each. 21

Table 2.3. Descriptive statistics of the study population from Census Canada 2011 and the Vital Statistics–Birth Database. 26

Table 2.4. Descriptive statistics of the industrial air emissions from the National Pollutant Release Inventory. 29

Table 2.5. Descriptive statistics of the land hazards. Unless measurement unit is indicated in parentheses, values are counts. 30

Table 2.6. Logistic regression coefficients, adjusted by covariates, by each province for small for gestational age (SGA) or low birth weight at term (LBWT). p<0.05 in bold. See Appendix I: Table S1 for all results. Table continues. 32

Table 3.1. Alberta's sub-provincial units and descriptive statistics, in descending order of birth number. 46

Table 3.2. Outdoor environmental factors mapped for association with adverse birth outcomes. The time (year), distance threshold (radius in meters), units, and source are indicated for each. 50

Table 3.3. Spearman's rank correlations of small for gestational age (SGA) and low birth weight at term (LBWT) with air substances and land sources (*), in descending correlation rho values. In the right half of the table, the count of units exceeding rho>0.40 and the range are shown for the data aggregated by health regions and airshed zones. Variables having a rho>0.4 for the province or for 4 or more sub-provincial units are indicated by bold font. Table continues........ 60

Table 3.4. Spearman's rank correlations of small for gestational age (SGA) and low birth weight at term (LBWT) with air substances, land sources, and weighted sum overlay indices for the entire province of Alberta. 69

Table 4.1. Census Metropolitan Area (CMA) characteristics are from the 2011 Census for Canada; customized map projection parameters are based on the centroids of the CMA, designated in decimal degrees of longitude (central meridian) and latitude (origin): X,Y for the Azimuthal Equidistant projection with North American Datum (NAD) 1983. 83

Table 4.2. Statistically significant hot spot categories for the 5- and 3-year study periods are defined in terms of the total months aggregated by 3-month time steps. Table continues. 92

Table 4.3. Census Metropolitan Area (CMA) number of records are from the Alberta Perinatal Health Program (APHP) and Canadian Neonatal Network (CNN) databases for only the records having valid 6-character postal codes. Note: SGA=small for gestational age; LBWT= low birthweight at term. 95

Table 4.4. Space-time cubes and emerging hot spot analyses exhibit increasing trends across Alberta Perinatal Health Program (APHP) all births, small for gestational age (SGA), low birthweight at term (LBWT) and Canadian Neonatal Network (CNN) critically ill SGA and critically ill LBWT. Proportion of each hot/cold spot category shown; pattern categories defined in Table 4.2. 97

Table 4.5. Spearman's correlation (rho) statistics compare emerging hot spot patterns for all births, SGA/LBWT, and critically ill SGA/LBWT, by Census Metropolitan Area (CMA). Patterns are also correlated with proportions of each land use and SES category. Significant *rho* values ($p<0.05$) are marked with an asterisk (*). 106

Table 4.6. Spearman's correlation matrix of the regression model covariates: proportions of land use and socioeconomic status categories, by Census Metropolitan Area (CMA). Significant rho values ($p<0.05$) are marked with an asterisk (*). 107

Table 4.7. Logistic regression β coefficients for all births, SGA/LBWT, and critically ill SGA/LBWT, modelled with proportions of surrounding land use categories and level of socioeconomic status (SES). Residential and high SES are the reference categories; Likelihood Ratio (LR) chi^2 significance is $p<0.001$; significant coefficients ($p<0.05$) marked by an asterisk (*); number of locations are indicated in Table 3. 108

Table 5.1. Characteristics of the Census Metropolitan Area (CMA) are from the 2011 Census for Canada and the Canadian Neonatal Network (CNN) database for critically ill small for gestational age (ciSGA). 120

Table 5.2. Provincial-level characteristics of population and the National Pollutant Release Inventory (NPRI). 130

Table 5.3. Census Metropolitan Area (CMA) characteristics of the National Pollutant Release Inventory, mean wind speed values from Environment and Climate Change Canada weather stations (centrally located in each CMA), and the number of maps of monthly wind-dispersed emissions. Individual chemical amounts are shown in Appendix II: Figures S2.1-S2.4. Table continues. 130

Table 5.4. Space-time trends and emerging hot spots across Canadian Neonatal Network (CNN) critically ill small for gestational age (ciSGA) newborns, by Census Metropolitan Area (CMA). Note: ↑=increasing; ↓=decreasing; ns=not significant trend. 135

Table 5.5. List of 28 chemicals having hot spot associations with ciSGA in three or more Census Metropolitan Areas (CMAs); * identifies the 23 statistically significant chemicals (p<0.05). Of these, there are 21 known or suspected developmental toxicants according to the Canadian Environmental Protection Act (CEPA), California's Office of Environmental Health Hazard Assessment (OEHHA), or GoodGuide's Scorecard. Appendix II: Figure S2.6 graphically shows the beta coefficients for the list. Table continues. 142

List of Figures

Figure 1.1. The set of birth weight–for–gestational age standards with the 10th percentile birth weights describes small for gestational age (SGA) in the purple curve; low birth weight at term (LBWT) is a subset of SGA in the green shaded rectangle. 2

Figure 1.2. Three-year averages of small for gestational age (SGA) in Canada, 2005-2007. Alberta's provincial percentage is labeled in white. 4

Figure 1.3. Three-year average of low birth weight (LBW [all gestational ages]) in Canada, 2005-2007. Alberta's provincial percentage is labeled in white.. 4

Figure 1.4. In Canada and Alberta, small for gestational age (SGA) and low birth weight (LBW [all gestational ages]) have been increasing since the year 2000... 5

Figure 1.5. Meade's triangle of human ecology for maternal exposures and small for gestational age (SGA) and low birth weight (LBWT). 7

Figure 2.1. Study population of babies born too small in Canada, 2006-2012............................ 18

Figure 2.2. Methods flow chart for associating small newborns and the outdoor environment... 22

Figure 2.3. Distribution of small for gestational age (SGA) and low birth weight at term (LBWT) for all, male (m), and female (f) babies born in Canada, 2006-2012.. 27

Figure 2.4. Distribution of industrial air emissions by province shown as percentage of the national total for number of facilities (gold), number of chemicals (orange), and tonnes emitted (red)... 28

Figure 2.5. Distribution of chemical emissions (orange) and land hazards (blue) having relatively stronger and statistically significant associations (β> 0.2 and p<0.05) with small newborns...... 35

Figure 3.1. Meade's triangle of human ecology for maternal exposures and small for gestational age (SGA) and low birth weight at term (LBWT): dashed arrow indicates hypothesized mechanisms.. 41

Figure 3.2. Flow chart of standard GIS commands for constructing the indices (colored boxes).53

Figure 3.3. Percentages of births having small for gestational age (SGA) or low birth weight at term (LBWT) in sub-provincial units (* indicates value for the whole province). 55

Figure 3.4. Births per area ratios in sub-provincial units... 56

Figure 3.5. Small for gestational age (SGA) or low birth weight at term (LBWT) ratios in sub-provincial units.. 57

Figure 3.6. Double kernel density (DKD) distributions of small for gestational age (SGA) and low birth weight at term (LBWT) are ratios of the adverse birth outcome per area divided by total births per area, each within a 25 km radius; DKD is dimensionless. 58

Figure 3.7. Weighted sum overlays for air substances and land-based sources were combined as equal and 0.7/0.3 weighted indices to identify the most hazardous locations. 68

Figure 3.8. Spearman's rank correlations of each adverse birth outcome with indices – shown as bar charts in each health region or airshed. Background maps show ratios. 72

Figure 4.1. The study focused on the Calgary and Edmonton Census Metropolitan Areas (CMA: orange areas), in the province of Alberta, Canada, served by hospitals with neonatal intensive care units (NICUs: red crosses) participating in the Canadian Neonatal Network....................... 84

Figure 4.2. The Canadian Neonatal Network (CNN) and Alberta Perinatal Health Program (APHP) data were subset to valid postal codes within the extent of Census Metropolitan Areas (CMA): Calgary (2006-2010) and Edmonton (2008-2010)... 85

Figure 4.3. Land use and socioeconomic status (SES) maps of the Calgary CMA..................... 87

Figure 4.4. Land use and socioeconomic status (SES) maps of the Edmonton CMA................. 88

Figure 4.5. Flow chart of GIS commands for analyzing small newborns in space and time. 90

Figure 4.6. Emerging hot spots of all births in the Calgary CMA... 98

Figure 4.7. Emerging hot spots of SGA and critically ill SGA in the Calgary CMA.................. 99

Figure 4.8. Emerging hot spots LBWT and critically ill LBWT in the Calgary CMA. 100

Figure 4.9. Emerging hot spots of all births in the Edmonton CMA... 101

Figure 4.10. Emerging hot spots of SGA and critically ill SGA in the Edmonton CMA........... 102

Figure 4.11. Emerging hot spots of LBWT, and critically ill LBWT SGA in the Edmonton CMA. 103

Figure 5.1. The study areas across Canada in the Canadian Neonatal Network (CNN) were defined by the 19 Census Metropolitan Areas (CMA: orange areas) that are served by 27 participating hospitals (red crosses) with neonatal intensive care units. 118

Figure 5.2. The Canadian Neonatal Network (CNN) data were subset to valid postal codes within the extent of 19 Census Metropolitan Areas (CMA): Edmonton (2008-2010), Ottawa (2007-2010), and all others (2006-2010). Critically ill small for gestational age (ciSGA) was defined as a binary variable.. 119

Figure 5.3. Flow chart of GIS commands for analyzing critically ill small newborns in space and time, plus detail on wind-dispersed air pollutants using kernel density and focal sum wedge modelling. The analyses were replicated for 19 Census Metropolitan Areas (CMAs) using customized azimuthal equidistant map projections for each. 124

Figure 5.4. Percentages of critically ill small for gestational age (ciSGA): admissions to neonatal intensive care units (NICU) and percentages of space-time hot spot patterns in the 19 Canadian Neonatal Network (CNN) study areas, identified by Census Metropolitan Area (CMA). 129

Figure 5.5. The Edmonton Census Metropolitan Area (CMA) is shown as an example distribution of infants during the study period, location of the Environment Canada weather station that provided the wind measurements, and the National Pollutant Release Inventory (NPRI) point sources of the industrial facilities that emitted chemicals to the air. All CMAs are in Appendix II: Figure S2.5. 131

Figure 5.6. Example of monthly wind dispersed pollutant – methanol – January and July 2010, Edmonton CMA. 133

Figure 5.7. Emerging hot spots of critically ill small for gestational age (ciSGA), represented as binary values (red=hot spot; grey=not hot), in the Edmonton CMA. 136

Figure 5.8. Example of emerging hot spots of methanol, represented as binary, in the Edmonton CMA. 136

Figure 5.9. Graduated symbols represent the logistic regression coefficients for critically ill small for gestational age (ciSGA) binary hot spots and each industrial air pollutant, modelled with proportions of surrounding low socioeconomic status, total number of infants, and NO_2 pollution from land use regression. The size of the 'bubble' represents the strength of the coefficient and the color indicates the direction: blue is negative, and red is positive. Chemicals are alphabetized within groupings of: (A) large volume gases, particulate matter, and polycyclic aromatic hydrocarbons; (B) volatile organic compounds; and (C) heavy metals and other organics/inorganics. 140

Figure 6.1. Western Canada, 2006-2012, double kernel density map of small for gestational age (SGA). Darker shades of purple indicate areas having relatively more SGA. The overlay of National Pollutant Release Inventory (NPRI=grey) facilities and National Air Pollution Surveillance Program (NAPS=orange) stations helps to identify missing gaps for monitoring. 153

Figure 6.2. Eastern Canada, 2006-2012, double kernel density map of small for gestational age (SGA). Darker shades of purple indicate areas having relatively more SGA. The overlay of National Pollutant Release Inventory (NPRI=grey) facilities and National Air Pollution Surveillance Program (NAPS=orange) stations helps to identify missing gaps for monitoring. 154

List of Abbreviations

Acronym	Explanation
AB	Alberta
ABO	adverse birth outcomes
APHP	Alberta Perinatal Health Program
BC	British Columbia
BW	birth weight
CAC	criteria air contaminant
CEC	Commission for Environmental Cooperation
CEPA	Canadian Environmental Protection Act
ciLBWT	critically ill low birth weight at term
ciSGA	critically ill small for gestational age
CMA	census metropolitan area
CNN	Canadian Neonatal Network
CO	carbon monoxide
DA	dissemination area
DKD	double kernel density
DMTI	Digital Mapping Technology Inc.
DoMiNO	Data Mining and Neonatal Outcomes
ECCC	Environment and Climate Change Canada
EHSA	emerging hot spot analysis
ESRI	Environmental Systems Research Institute
GIS	Geographical Information System
IUGR	intrauterine growth restriction
LBW	low birth weight
LBWT	low birth weight at term
LUR	land use regression
MB	Manitoba
NAD83	North American Datum 1983
NB	New Brunswick

Acronym	**Explanation**
NDVI	Normalized Difference Vegetation Index
netCDF	network Common Data Form
NICU	neonatal intensive care units
NL	Newfoundland
NO	nitrogen oxide
NO_2	nitrogen dioxide
NPRI	National Pollutant Release Inventory
NS	Nova Scotia
NT	Northwest Territories
NU	Nunavut
O_3	ozone
OEHHA	Office of Environmental Health Hazard Assessment
ON	Ontario
PAHs	polycyclic aromatic hydrocarbons
PE	Prince Edward Island
PM	particulate matter
PM_{10}	particulate matter with aerodynamic diameter < 10 µm
$PM_{2.5}$	particulate matter with aerodynamic diameter < 2.5 µm
PRTR	Pollutant Release and Transfer Registers
QC	Quebec
RDC	Research Data Centre
SES	socioeconomic status
SGA	small for gestational age
SK	Saskatchewan
SO_2	sulfur dioxide
VOC	volatile organic compounds
VSBD	Vital Statistics–Birth Database
YT	Yukon

Chapter 1 Introduction

An underlying premise of environmental health involves place – where one lives and where one starts out in life, even during in utero development, ultimately determines lifelong health [1, 2].

The embryo and fetus are susceptible to toxicant exposure and other environmental influences on the mother during crucial stages of pregnancy [3–6], which may lead to babies being born too small, or too early. Because they are important markers of infant survival, development, and future health, newborns that are too small are a serious emotional and economic stress on society – hundreds of millions of dollars are spent on specialized equipment and treatments within the first several years of life [7, 8]. The Barker hypothesis [9] evolved from studies on low birth weight (as well as premature birth and intrauterine growth retardation) that found significant associations with adult hypertension, coronary heart disease, and non-insulin-dependent diabetes [10–12]. The suspected exposures associated with these birth outcomes are widespread, thus heightening the importance of early life health impacts.

As depicted in Figure 1.1, born too small is defined as:

- Small for gestational age (SGA), which are infants born with a birth weight $<10^{th}$ percentile of a Canadian reference population for sex-based gestational age (22 to 42 weeks gestation) [13], to enable comparisons with previous Canada-based studies;
- Low birth weight at term (LBWT), which are full-term infants (37-42 weeks gestation) with a birth weight <2,500 g [14]; and

- Critically ill SGA (ciSGA) or critically ill LBWT (ciLBWT) apply to infants described above and who are immediately admitted to neonatal intensive care units (NICU).

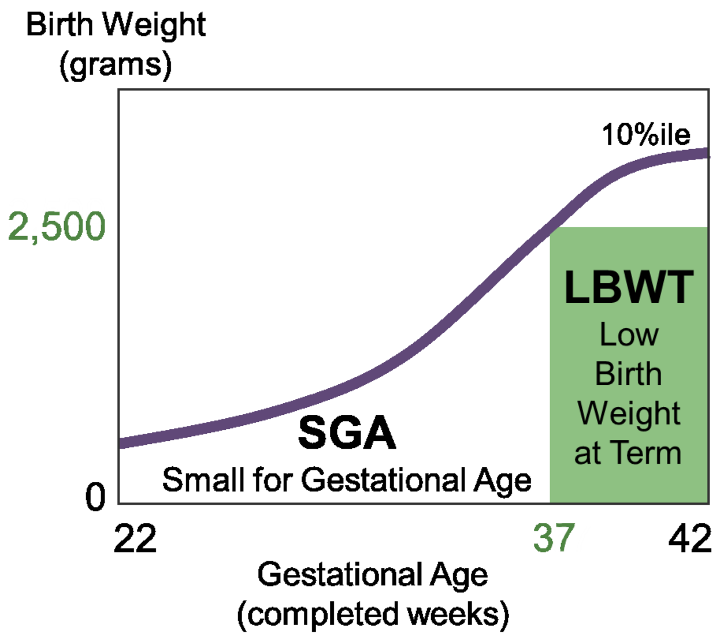


Figure 1.1. The set of birth weight–for–gestational age standards with the 10th percentile birth weights describes small for gestational age (SGA) in the purple curve; low birth weight at term (LBWT) is a subset of SGA in the green shaded rectangle.

SGA and LBWT are not homogeneous pregnancy outcomes, because they may consist of both infants born too early (known as preterm birth) or too small, (typically due to fetal growth restriction) [14, 15]. The etiologies are multifactorial, where the most important maternal risk factors are tobacco smoking, nutrition, pre-pregnancy weight, ethnic origin, short maternal stature, and pre-existing health conditions [15–18]. Other risks include genetic and constitutional, demographic and psychosocial (e.g. socioeconomic status and stress), obstetric, antenatal care, and toxic exposures. Critically ill SGA/LBWT likely result from the same risk factors but they have complications requiring higher costs of care.

Globally, the rate of SGA in low- and middle-income countries is around 27% of all live births. In 2010, 32.4 million babies were small for their gestational age [19]. LBW (all gestational ages) occurs in 15% of all births, mostly in low and middle-income countries [20]. Of 18 million low-birthweight babies, 10.6 million of them were born at term. In the United States SGA was 10% in 2005 [21] and LBW averaged 6.4% from 2006-2016 [22].

The Canadian average of SGA was 8.4% (Figure 1.2) and LBW was 6.0% (Figure 1.3) [23] during 2005-2007. Although Canada is lower than the world and U.S., these disorders related to short gestation and low birth weight are consistently ranked 2nd out of the 71 leading causes of infant death (congenital malformation is the leading cause [24]) – and has been increasing since the year 2000 (Figure 1.4) [23].

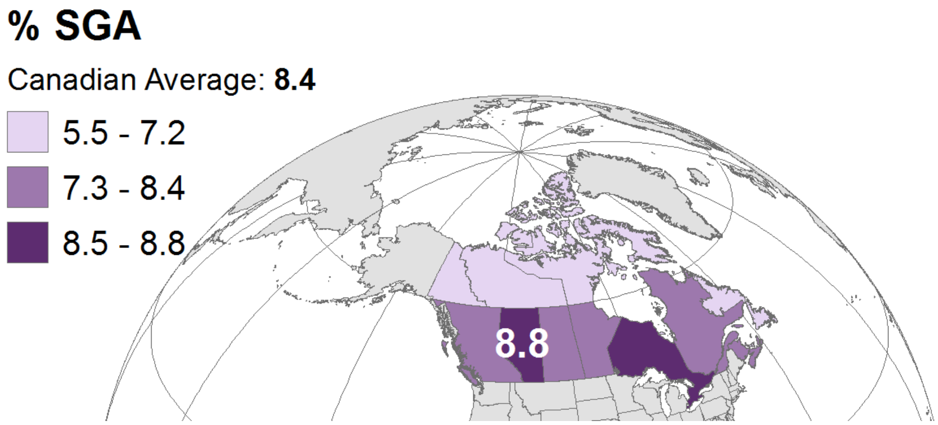


Figure 1.2. Three-year averages of small for gestational age (SGA) in Canada, 2005-2007. Alberta's provincial percentage is labeled in white.

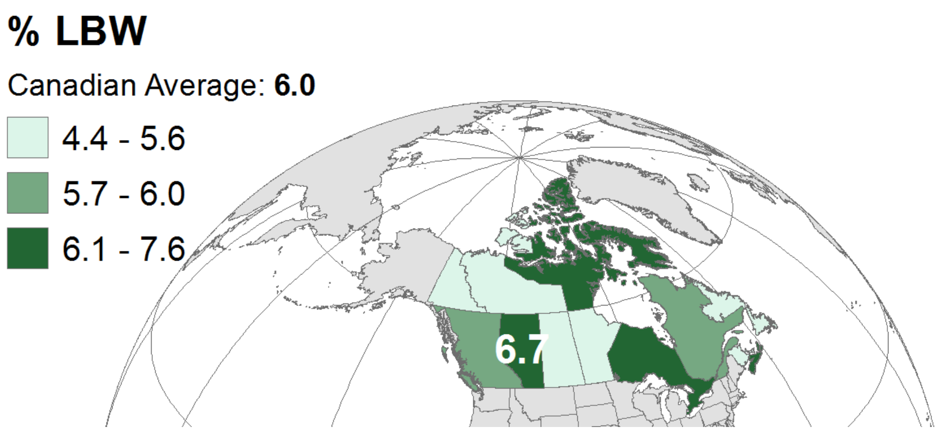


Figure 1.3. Three-year average of low birth weight (LBW [all gestational ages]) in Canada, 2005-2007. Alberta's provincial percentage is labeled in white.

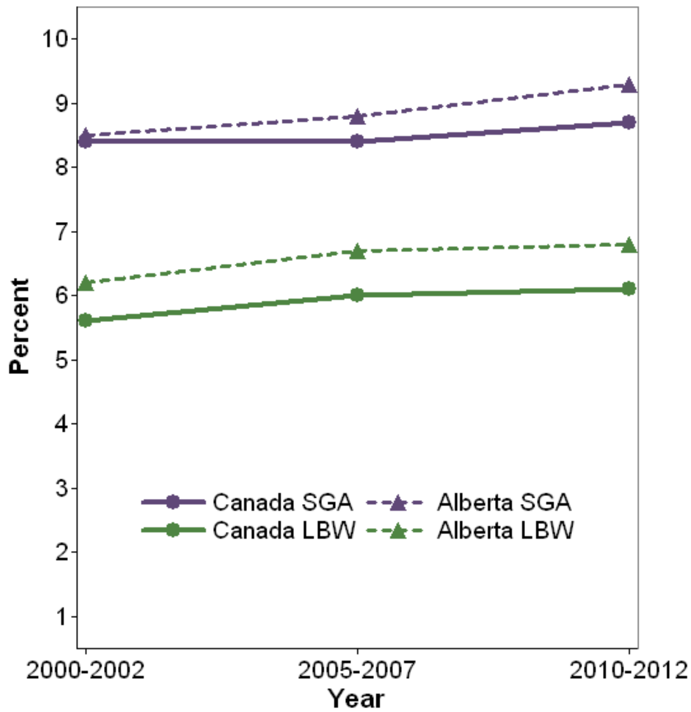


Figure 1.4. In Canada and Alberta, small for gestational age (SGA) and low birth weight (LBW [all gestational ages]) have been increasing since the year 2000.

The province of Alberta has been consistently higher than the national averages (Figure 1.4). Similar to the rest of Canada, 80% of the Alberta population lives in or near urban areas, but overall it has been rapidly increasing, in part due to an average of 50,000 births per year [25], from 2006 to 2012. In 2011, the Alberta population increased 10.8% from the 2006 census; the national increase was 5.9% [26]. Given that Alberta rates are close to or exceeding the U.S percentages, and Canada is increasing over time, it is valuable from a public health perspective to understand the patterns and processes involved in being born too small.

SGA/LBWT and their association with the environment necessitate an interdisciplinary approach with integration of knowledge from medicine and geography. Medical geography is a holistic investigation of health using concepts and methodologies from geography, which also encompasses the social, physical, and biological sciences [27]. Potential linkages between the maternal environment and SGA/LBWT is what I endeavor to understand as an important state of human health.

Informed by the earlier work of May – who stated that to understand disease as a biological expression of maladjustment, an ecological (i.e. ecosystem-based) study must involve the environment, the host, and the culture [28] – Meade proposed the **triangle of human ecology** as the framework for the state of human health [27, 29]. Meade's vertices are therefore anchored to:

- *Habitat* – the natural, social, and built *environments* where people live.
- *Population* – people (*hosts*) as biological organisms structured by age, gender, and genetics.
- *Behavior* – visible part of *culture* including beliefs, social organization, and technology.

These three points influence each other and the state of health, as can be seen when modelling and summarizing what is known about neonatal outcomes and maternal exposure to outdoor pollution (Figure 1.5).The primary population consists of pregnant mothers and their defining individual characteristics of varied ages, pre-existing health conditions and genetic makeup, with the location of where they live and work depending on their social and economic behaviors (i.e. nutritional status, access to quality health services). More research is needed that focuses on the lesser studied habitat vertex, more specifically the outdoor environment, since much less attention has been given to integrating ecological factors for understanding disease [27]. The location aspect of habitat (i.e. geography) – where mothers live, where industry and services are situated, where demographic groups congregate, and for many scales – is important to clinicians, specialists in environmental health and exposure assessment, epidemiologists, biostatisticians, and computing scientists.

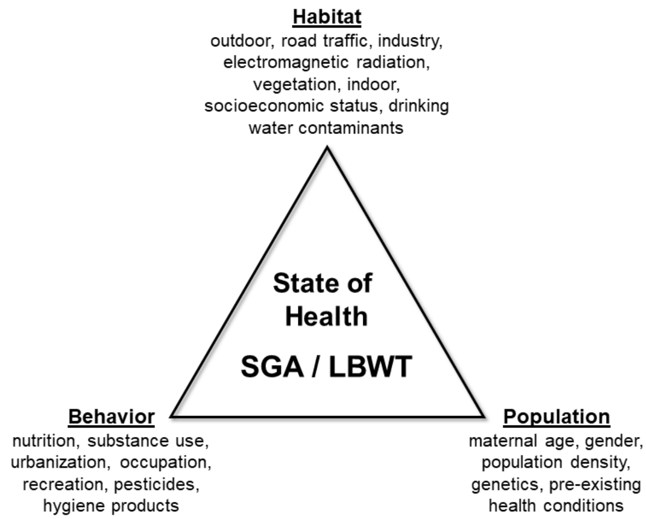


Figure 1.5. Meade's triangle of human ecology for maternal exposures and small for gestational age (SGA) and low birth weight (LBWT).

Geography and environmental health are inextricably linked. Environmental health, as defined by the World Health Organization, "comprises those aspects of human health and disease that are determined by factors in the environment, and includes both the direct pathological effects of chemicals, radiation and some biological agents, and the effects (often indirect) on health and wellbeing of the broad physical, psychological, social and aesthetic environment, which includes housing, urban development, land use and transport" [30]. These concepts are not new – Hippocrates, the father of medicine, c. 460 – c. 370 BC, understood the important role that the environment plays in health, in his "Airs, Waters, and Places" [31]:

Whoever wishes to investigate medicine properly, should proceed thus: in the first place to consider the seasons of the year, and what each of them produces (for they are not all alike, but differ much from themselves in regard to their changes). Then the winds, the hot and the cold, especially such as are common to all countries, and then such as are peculiar to each locality. We must also consider the qualities of the waters, for as they differ from one another in taste and weight, so also do they differ much in their qualities. In the same manner, when one comes into a city to which one is a stranger, one ought to consider its situation, how it lies as to the winds and the rising of the sun; for its influence is not the same whether it lies to the north or the south, to the rising or to the setting sun. These things one ought to consider most attentively, and concerning the waters, which the inhabitants use, whether they be marshy and soft, or hard, and running from elevated and rocky situations, and then if saltish and unfit for cooking; and the ground, whether it be naked and deficient in water, or wooded and well watered, and whether it lies in a hollow, confined situation, or is elevated and cold; and the mode in which inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labour, and not given to excess in eating and drinking.

From these things one must proceed to investigate everything else. For if one knows all these things well, or at least the greater part of them, he cannot miss knowing, when he comes into a strange city, either the diseases peculiar to the place, or the particular nature of common diseases, or commit mistakes, as is likely to be the case provided one had not previously considered these matters.

Hazards encompass those airs, waters, and places that comprise the chemical, physical, and biological aspects that insult human health [27]. Many hazards have been known for centuries (lead, radiation, microorganisms), but they are only effective in altering health if an individual is exposed to them.

Exposure is the occurrence a person comes into contact (via air, water, or skin) with a dose (requisite amount) of a hazard or toxicant (substance that produces a health effect) and may be one time, repeated, or continual [32]. The health outcome can only occur if a person is exposed to the integral dose of a hazard for the crucial amount of time. These ideas are directly applicable to being born too small, and the system can be simplified as follows:

$$Hazard_{environment} \rightarrow Exposure_{prenatal} \rightarrow Outcome_{SGA/LBWT}$$

An efficient way to associate the hazards and the health outcomes is in a geographical information system (GIS). A GIS is a system that automates, visualizes, analyzes, manages, and delivers information through geographic presentation (www.gis.com). A GIS integrates computer hardware, software, and digital data, in analytical methods performed by knowledgeable users to answer complex questions and support the modelling of complex hazard-exposure-dose-response processes in space and time as those included relevant for my research. Coupled with spatial statistics, GIS has been transforming how to answer complex questions across a myriad of disciplines since most human activity can be tied to a location or place, and more recently, time and space. Spatial science and technology inspire and enable a deeper comprehension of the world, while also facilitating spatial statistical analyses that quantify the distribution, pattern, process, and relationships of the environment and society.

GIS has been used in many environmental health studies to define epidemiologic study populations, identify source and potential routes of exposure, estimate environmental levels of target contaminants, and estimate personal exposure [33–40]. The integration of spatial databases, standardized from a variety of sources, using the appropriate spatial and temporal scales, and applying methods that incorporate coincidence, proximity, and surface analysis, are instrumental in mapping exposure and diseases, to reveal patterns that explore changes, associations, and risks [32].

Geographic inquiry has been applied specifically to adverse birth outcomes and associations with the environment. Table 1.1 highlights a list of environmental hazards studied using GIS or similar spatial analyses. The 39 selected studies identified the following variables: agricultural related chemicals, ambient air pollution, built environment, waste sites, transmission lines, natural gas activities, lead, mines, greenness, wildfires, socioeconomic status, roads and traffic-related pollution.

Table 1.1. Selected studies involved in spatial associations among environmental variables and adverse birth outcomes: SGA=small for gestational age; LBW=low birth weight; BW=birth weight; IUGR=intrauterine growth restriction. Table continues.

Study	Outcome studied	Spatial variables
Xiang et al. 2000 [41]	LBW	agriculture: crop production patterns
Ochoa-Acuna et al. 2009 [42]	SGA	agriculture: herbicides
Sathyanarayana et al. 2010 [43]	LBW	agriculture: pesticides
Weselak et al. 2007 [44]	IUGR	agriculture: pesticides
Coker et al. 2016 [45]	LBW	air pollution: NO, NO_2, $PM_{2.5}$
Malmquist et al. 2017 [46]	BW	air pollution: NO_2
Svechkina et al. 2018 [47]	LBW	air pollution: NO_2
Liu et al. 2007 [48]	IUGR	air pollution: NO_2, CO, SO_2, O_3, $PM_{2.5}$
Stieb et al. 2016 [49]	SGA, LBW, BW	air pollution: NO_2, $PM_{2.5}$
Choi et al. 2008 [50]	SGA	air pollution: PAHs
Huang et al. 2015 [51]	LBW	air pollution: PM_{10}, NO_2
Dugandzic et al. 2006 [52]	LBW	air pollution: PM_{10}, SO_2, O_3
Basu et al. 2014 [53]	LBW	air pollution: $PM_{2.5}$
Erickson et al. 2016 [54]	BW	air pollution: $PM_{2.5}$
Harris et al. 2014 [55]	LBW	air pollution: $PM_{2.5}$
Stieb et al. 2016 [56]	SGA, LBW, BW	air pollution: $PM_{2.5}$
Wilhelm et al. 2012 [57]	LBW	air pollution: $PM_{2.5}$, NO, NO_2, PAHs
Liu et al. 2003 [58]	LBW, IUGR	air pollution: SO_2, NO_2, CO, O_3, PM_{10}
Rich et al. 2015 [59]	BW	air pollution: SO_2, NO_2, CO, $PM_{2.5}$
Miranda et al. 2012 [60]	SGA, LBW	built environment
Woods et al. 2017 [61]	SGA, LBW	built environment
Zeka et al. 2008 [62]	BW, SGA	land use
Svechkina et al. 2018 [47]	LBW	petrochemical industry
Baibergenova et al. 2003 [63]	LBW	waste site
Goldberg et al. 1995 [64]	LBW, SGA	waste site
Auger et al. 2011 [65]	LBW, SGA	transmission lines
de Vocht et al. 2014 [66]	LBW, SGA	transmission lines
Casey et al. 2016 [67]	SGA, LBW	natural gas
McKenzie et al. 2014 [68]	LBW	natural gas
Stacy et al. 2015 [69]	SGA, BW	natural gas

Study	Outcome studied	Spatial variables
Philion et al. 1997 [70]	IUGR, SGA	lead
Rabito et al. 2014 [71]	LBW, SGA	lead
Ahern et al. 2011 [72]	LBW	mine site
Henn et al. 2016 [73]	BW	mine site
Hystad et al. 2014 [74]	SGA, BW	greenness
Holstius et al. 2012 [75]	BW	wildfires
Erickson et al. 2016 [54]	BW	socioeconomic status
Habermann and Gouveia 2014 [76]	LBW	socioeconomic status
Zeka et al. 2008 [62]	BW, SGA	socioeconomic status
Généreux et al. 2008 [77]	LBW, SGA	roads
Svechkina et al. 2018 [47]	LBW	roads
Zeka et al. 2008 [62]	BW, SGA,	roads
Brauer et al. 2008 [78]	SGA, LBW	traffic-related air pollution
Habermann and Gouveia 2014 [76]	LBW	traffic-related air pollution
Meng et al. 2013 [79]	LBW	traffic-related air pollution
Wilhelm et al. 2003 [80]	LBW	traffic-related air pollution

The strength of association in the studies varied greatly and had limitations due to sampling, spatial resolution, availability of confounding factors, adjusting for residential mobility, and inability to quantify duration and intensity of exposures. Overall, the studies contributed to the evolving evidence that maternal exposure during pregnancy to varying levels of outdoor environmental hazards are associated with adverse birth outcomes. Associations of air pollution with SGA/LBWT are accumulating the most in the published literature. Anthropogenic air pollution originates from industrial/traffic emissions and includes gaseous components – sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen oxide (NO), nitrogen dioxide (NO_2), ozone (O_3) – and particulate matter (PM) – $PM_{2.5}$ particles with aerodynamic diameter ≤ 2.5 μm and PM_{10} particles ≤ 10 μm.

Mechanisms that trigger adverse birth outcomes, such as born too small, among mothers exposed to pollutants are not well understood, but are suspected to include inflammation, direct toxic effects on the placenta and the fetus, interruption of oxygen-hemoglobin interaction, and damage to DNA [81–83]. Environmental associations differ among SGA and LBWT, enhanced by temporal variations in exposures, personal characteristics (mothers' health, nutrition, and demographics) and external factors such as region and socioeconomic status (SES), [3, 4, 84].

Many of the previous studies linked individual or small subsets of factors; however all factors can be modelled as vertices of the triangle of human ecology, synthesizing the complex disease ecology and advancing hypotheses [27]. As Table 1.1 exemplifies, the majority of air pollutants under investigation consisted of traffic-related criteria air contaminants. A handful of studies have targeted heavy metals and/or industrial activities. More research is needed on assessing the spatial relationships of the actual chemicals involved in those industrial activities, especially the known or suspected developmental toxicants. Similarly, the combined effect of multipollutant exposures are still relatively unknown.

I hypothesize that there were more small newborns from mothers living with more and larger amounts of ambient health hazards than newborns from mothers living in relatively healthier environments, where small newborns are defined by SGA and/or LBWT. Essentially, prenatally exposed newborns have lower birth weight:

$$H_0: birth\ weight_{exposed} = birth\ weight_{nonexposed}$$

$$H_1: birth\ weight_{exposed} < birth\ weight_{nonexposed}$$

My specific investigations involve the following interrelated questions:

i. What are the important outdoor environmental exposures on pregnant mothers associated with the occurrence of small newborns in Canada? Specifically, which provinces and territories have more outdoor environmental hazards spatially associated with SGA/LBWT? Do they differ among land sources and industrial chemicals?

ii. Can I apply methods to reduce/aggregate environmental variables into a hazard map/index? How do the separate and combined exposures to the outdoor environments of pregnant mothers coincide with patterns in SGA/LBWT? Does region matter?

iii. Where are the space and time "hot spots" of SGA/LBWT for two Alberta cities? Do the patterns compare with critically ill SGA/LBWT? Does area-level socioeconomic status and surrounding land use coincide with the hot spots?

iv. Do the hot spots of critically ill SGA collocate with hot spots of industrial air emission in space and time for nineteen Canadian cities?

To address the above and in an effort to better understand the contribution of ambient health hazards to babies born too small, I examine the distributions and patterns across Canada (Chapter 2), develop an index and test regional variations for Alberta (Chapter 3), compare non-critical and critically ill patterns (Chapter 4), and then focus on industrial emissions and critically ill small newborns in the major cities (Chapter 5). The geographic pathway from national to provincial to major cities facilitates regional understanding; the medical pathway from non-critical to critically ill small newborns helps identify patterns of the costlier events. Six percent of neonatal conditions may be attributed to modifiable environmental risks, i.e. maternal occupation, chemicals, air pollution, water, sanitation, and hygiene [85]. I use GIS and spatiotemporal analysis to recognize where and when environmental factors associate with SGA/LBWT in the expectation to promote and support awareness of this important children's environmental health issue, highlight research gaps, and contribute to improving quality of life through geography.

Chapter 2 Geographic information assessment of maternal ambient health hazards and babies born too small

2.1 Abstract

Background: Small newborns, defined as small for gestational age (SGA: birth weight below 10th centile) or low birth weight at term (LBWT: birth weight below 2,500 g at 37 or more weeks gestation), have been increasing in Canada since 2000 and are the second leading cause of infant mortality. Recently, associations of SGA and LBWT have been linked to maternal exposure to environmental hazards. My research assesses which provinces have more associations with industrial air pollutants or land activities.

Methods: I classified SGA and LBWT events from Statistics Canada's Vital Statistics–Birth Database (2006-2012). I calculated spatial proxies of exposures to 228 industrial chemicals released to air and seven land-based hazards and assigned the values to the 6-character postal codes of the maternal residence at birth. I used logistic regression, with covariates on area-level socioeconomic status, traffic, maternal age, migration, sex, urban, total number of births, and season.

Results: Of the 2,525,645 births meeting my criteria (single, live, between 22 to 42 weeks gestation, and having a valid 6-character postal code), 8.55% were SGA and 1.54% were LBWT. Maps of the provincial patterns showed higher adverse birth outcomes where there were more industrial emissions. More provinces had associations with land hazards, especially dumps/waste depots, gas stations, powerlines and transformer stations. Of the 12 identified industrial

chemicals, nine are suspected or known developmental toxicants, including ammonia, benzene, carbon monoxide, methyl ethyl ketone, and particulate matter.

Discussion: Maternal exposures to ambient health hazards and the identified associations with small newborns differed by province, reflecting the multifactorial nature of SGA and LBWT. The large geographical scale and population-level exposure assignment preclude causation, but the use of publicly available data and accessible tools identify associations and facilitate a more holistic environmental health approach in understanding SGA and LBWT risk factors.

2.2 Background

Where one lives and starts out in life, even during in utero development, contributes to lifelong health [1]. Numerous studies have found associations of toxicant exposures on mothers during vital phases of pregnancy and resulting adverse birth outcomes (ABOs), such as being born too small or too early [3, 5]. Since ABOs are important indicators of infant survival, development, and future health, they are a serious emotional and economic stress on society [7, 8].

Babies born too small are clinically described as small for gestational age (SGA), or the more objective subset of SGA, low birthweight at term (LBWT). In Canada, SGA defines infants with birth weights below the 10^{th} centile according to sex and gestational age, based on Kramer's compiled statistics on female and male birth weights at each week of pregnancy [13]. LBWT are full-term (more than 37 weeks gestation) infants with birth weights less than 2,500 grams [7, 14].

The percentage of small newborns in Canada has increased over time: the 2000-2002 average for SGA was 8.4% and the 2010-2012 average was 8.7%; LBW (infants with birth weights <2,500

g) was 5.6% in 2000-2002 and increased to 6.1% in 2010-2012 [23]. Disorders related to short gestation and low birth weight are the second leading cause of infant mortality in Canada [24].

SGA and LBWT are multifactorial health issues with risks including individual (maternal), socio-cultural, and environmental [18, 81, 86]. The individual factors do not completely explain all events [87, 88]. And because the associated environmental exposures are often widespread, they may be a substantial burden for SGA/LBWT. Environmental health research and SGA/LBWT have been growing and has primarily focused on traffic-related air-pollution exposures from estimates measured by air pollution monitors or land use regression models [78, 89, 90]. There are many challenges in modelling exposures, especially in the context of space and time [33, 91]. Seasonal effects (e.g. warmer months when residential windows might be open) are further complexities [91–96]. Additional environmental factors are increasingly being considered, such as greenness [74, 97], the built environment [45, 67, 68, 98], and agricultural/industrial chemicals [4]. Not every single chemical released from industrial facilities are monitored, but estimates are available from pollutant release and transfer registers that may be adapted for human health research [99].

I hypothesized that larger quantities of ambient health hazards – shared sources of maternal exposures – may be spatially related to more SGA and/or LBWT. I assessed which provinces and territories have more outdoor environmental hazards spatially associated with babies born too small, whether they differ among land hazards and industrial chemicals (analyzed individually), and determine how season may affect the associations, in an effort to better understand the contribution of ambient health hazards to babies born too small.

2.3 Methods

I conducted a retrospective cohort study using data from Statistics Canada's confidential Master Data File for the Vital Statistics–Birth Database (VSBD), record number 3231, between 2006 and 2012 [100]. For the spatial analyses, I used Esri's ArcGIS Desktop 10.6 [101], and for the statistical analyses, I used STATA 15 [102], at the individual level by 6-character postal code.

Because distance was important, I customized an Azimuthal Equidistant map projection to center on Canada at -95° W longitude and 55° N latitude. All analyses used a 1 km cell size, 3 km radius (except NPRI was 10 km), and geodesic parameter whenever available (to account for the earth's actual shape in the calculations, especially important for the large territories and provinces).

Dependent variables

The VSBD contained the necessary variables – birth weight, child's sex, and gestational age (i.e. pregnancy duration) – to classify birth events according to adverse birth outcome for all provinces and territories across Canada. Small for gestational age (SGA) defined newborns between 22-42 weeks having birthweights below the 10^{th} centile, based on Canadian normative data [13]. A subset of SGA is low birth weight at term (LBWT), which defined newborns having birthweights below 2,500 g at 37-42 weeks gestation (i.e. at term). The variables important for selecting the study population were gestational age, type of birth (single or multiple), stillborn, and maternal place of residence (i.e. 6-character postal code). As shown in Figure 2.1, I analyzed all records according to my criteria of single, live newborns between 22 and 42 weeks gestation, having valid postal codes. Variable definitions are summarized in Table 2.1.

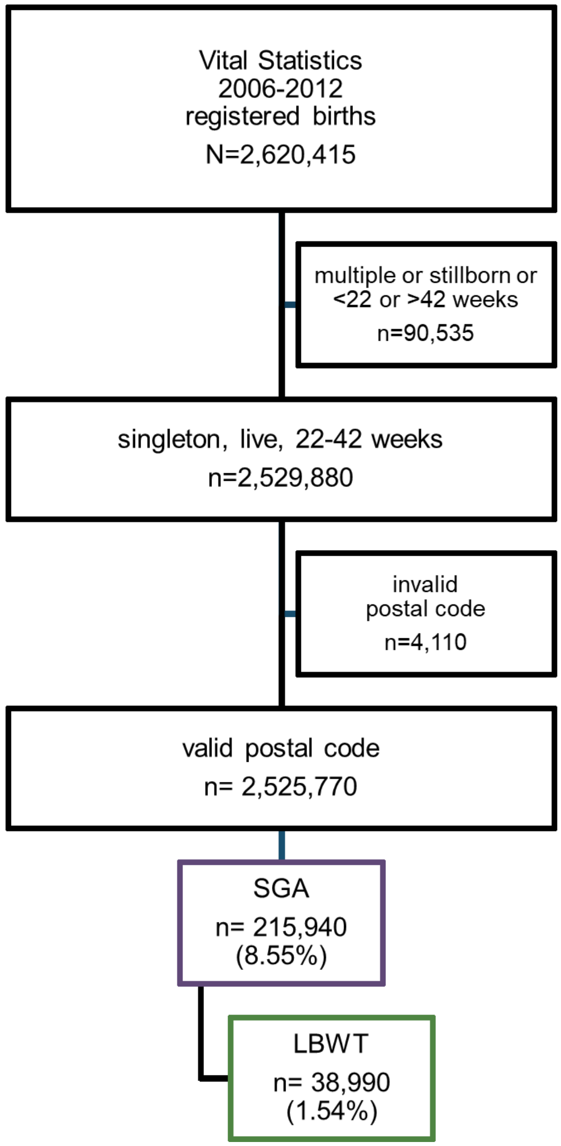


Figure 2.1. Study population of babies born too small in Canada, 2006-2012.

Table 2.1. Definitions of variables and concepts important to the study of maternal ambient hazards with babies born too small.

Variable	Definition	Value	Role
Small for gestational age (SGA)	birthweight ≤ Kramer's 10th centile (based on sex and pregnancy duration)	Binary	Dependent variable
Low birth weight at term (LBWT)	birthweight < 2500 g AND pregnancy duration ≥ 37 weeks		
Criteria for selecting the study population	residence postal code is not null AND born total this event = 1 AND stillborn total this event = 0 AND pregnancy duration ≥ 22 AND pregnancy duration ≤ 42 AND longitude is not null		Select data
Baby is female	sex = 2		Covariate
Mother age group 19 years and younger	maternal age ≤ 19		
Mother age group 40 years and older	maternal age ≥ 40		
Mother's birthplace different from baby's	maternal birthplace ≠ baby birthplace		
Urban residence	postal code's second character ≠ 0		
Warm season in first trimester	conception month ≥ 4 AND conception month ≤ 7; 2/3 or more of trimester was during May to August		
Warm season in second trimester	conception month ≥ 1 AND conception month ≤ 6; 2/3 or more of trimester was during May to August		
Warm season in 3 months prior to birth	birth month ≥ 4 AND birth month ≤ 7; 2/3 or more of trimester was during May to August (this equates to third trimester for term births)		
Number of births for the year	sum of births for postal code for same year as the birth	Integer	
Low socioeconomic status neighborhood	proportion of low SES within 10 km radius	Decimal	
Road density	meters of road per km^2 within 10 km radius; traffic pollution surrogate		
Birth year proportion	number weeks pregnancy during birth year / pregnancy duration		Calculate exposure
Conception year proportion	number weeks pregnancy during conception year / pregnancy duration		

Covariates

From the VSBD data on maternal age, I categorized binary records as mothers ≤ 19 years or mothers ≥ 40 years. I compared the maternal and baby's birthplace by province, and where they differed, I identified records as maternal migration. I also categorized binary records as female (sex) and urban birthplace (if the second character of the postal code ≠ 0). To account for higher numbers of births for some locations, I summed the total births meeting my criteria at each postal code, by year. To investigate the effect of season, I categorized binary variables that depended on whether the warm season (May through August) occurred during 2/3 or more of the first trimester, second trimester, or 3 months prior to birth (i.e. third trimester for term births). Variable definitions are summarized in Table 2.1. I did not have individual health or behavioral variables available for this study.

From Chan et al.'s [103] comprehensive index of Canadian socioeconomic status (SES), I categorized binary records by grouping the provided quintile values 0 (no data), 1, and 2 as low SES, and all else (3, 4, and 5) were grouped as medium-high SES. To allocate the area-level SES index, I calculated the proportion surrounding each postal code (see Focal statistics below).

From Statistic Canada's [104] road network files, I calculated a proxy of traffic pollution surrounding each postal code by calculating the density (see Kernel density below).

Independent variables

My variables of interest were outdoor-related environmental exposures that were from air or land hazards, summarized in Table 2.2. Figure 2.2 shows what standard GIS commands were used and where each dataset fits in to the overall analysis.

Canada's National Pollutant Release Inventory (NPRI) provided annual estimates of chemicals released by industrial facilities mandated to report under the 1999 Canadian Environmental Protection Act [105]. NPRI pollutants included core and alternate threshold substances, polycyclic aromatic hydrocarbons (PAH), dioxins, criteria air contaminants, and speciated volatile organic compounds (VOC). I selected all substances released to air for each of the years 2005 to 2012 and converted all units to tonnes.

For land hazards, I did not have access to standardized data for all study years. I chose the datasets that matched closest to the final year of the study (see Table 2.2). From the Commission for Environmental Cooperation (CEC)'s 2010 Land Cover of North America (30 meters), I reclassified the crop land [106]. For all other land hazards, I extracted the relevant datasets from Digital Mapping Technology Inc. (DMTI) Spatial's CanMap Content Suite [107]: dumps/solid waste depots (region centroids); gas stations; petroleum well pads, power lines; mines (merged points with region centroids); and transformer stations (merged points with region centroids).

Table 2.2. Outdoor environmental factors mapped for association with adverse birth outcomes. The time (year), spatial method, units, and source are indicated for each.

Variable	Year	Feature	Method	Units	Source
228 chemical emissions	2005-12	point	Kernel Density	tonnes / km^2	NPRI [105]
Roads	2005-12	line		km / km^2	StatsCan [104]
Electrical power lines	2015	line		km / km^2	DMTI Spatial [107]
Dumps/solid waste depots	2015	point		# / km^2	
Gas stations	2015	point		# / km^2	
Mine sites	2015	point		# / km^2	
Oil/Gas well pads	2015	point		# / km^2	
Transformer stations	2015	point		# / km^2	
Crop lands	2010	raster	Focal Statistics	km^2 / km^2	CEC [106]
Socioeconomic status – area-level index	2006	raster		index	Chan et al. [103]

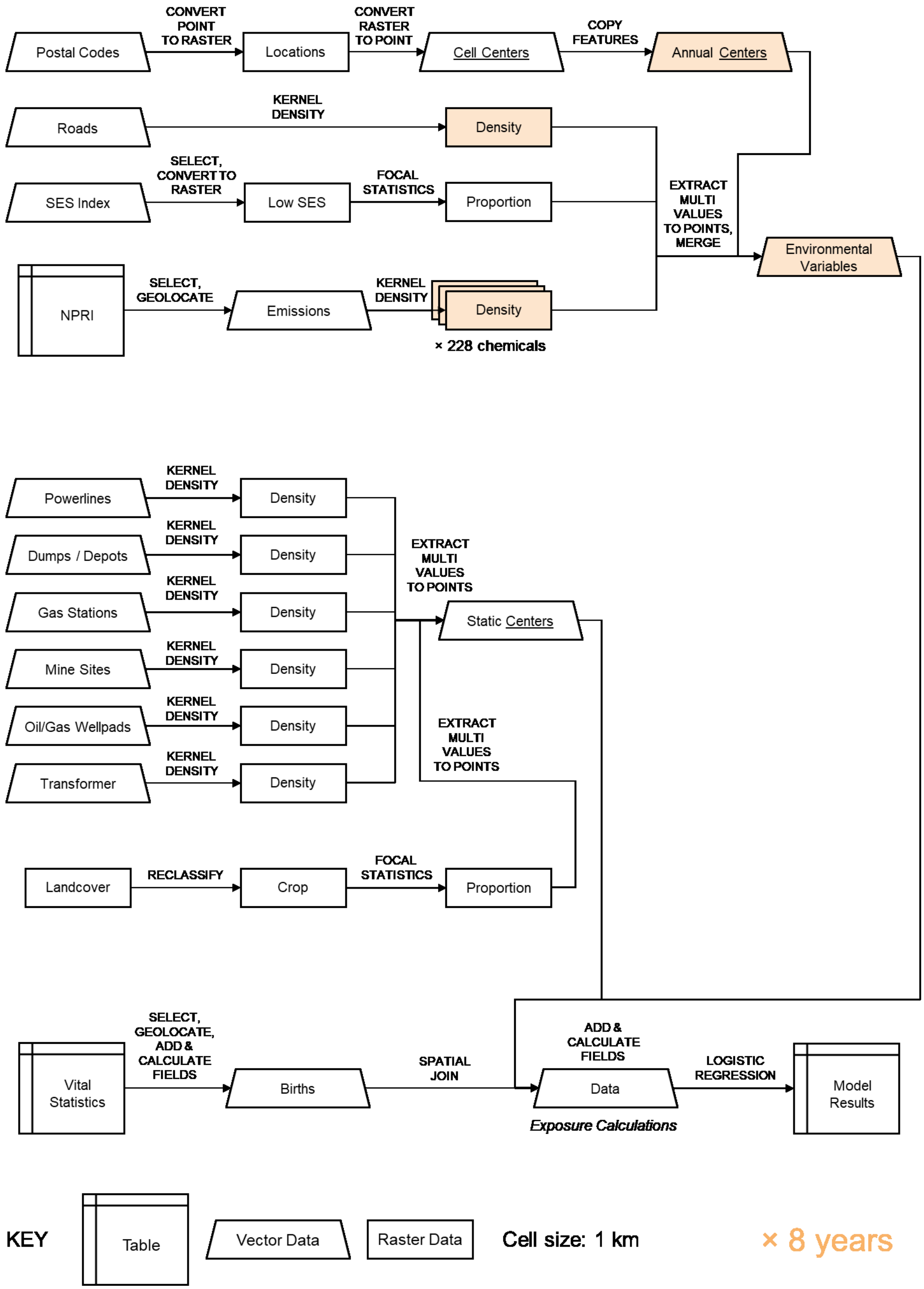


Figure 2.2. Methods flow chart for associating small newborns and the outdoor environment.

Kernel density

As depicted in Figure 2.2, I calculated the kernel densities for many of the input datasets. Kernel density is a non-parametric computation that spreads the values from points (or lines) across a surface by calculating the magnitude-per-unit area [108]. It fits to a smoothly tapered function within a specified distance around each point, which allows for distance decay, meaning features have less influence further away. For the industrial emissions, I specified the magnitude as the annual tonnes emitted. Roads and powerlines were specified using lengths. All points (see Table 2.2) were specified using counts. My independent variables were calculated as tonnes (10 km radius to correspond with my overall project research [109]), kilometers (3 km radius), or numbers (3 km radius) per square kilometer, using a 1 km cell size.

Focal statistics

For low SES and crop lands, I reclassified the categorical values in to binary surfaces where 1 indicated presence and 0 indicated absence. I then applied focal statistics, also known as moving-window or neighborhood analyses, with the mean statistic on binary values to give proportions. SES and crop calculations had a 3-km radius and 1 km cell size, resulting in the proportion of the category surrounding each cell.

Postal code assignment

Working with the VDSB required accessing the protected file from within the Statistics Canada's Research Data Centre (RDC) at the University of Alberta. The large size of the environmental variables required full processing outside and then importing to the RDC for spatially joining to the birth records. To achieve this as effectively as possible, I merged postal codes for the years

2001 to 2013 from DMTI Spatial's Postal Code Suite [110] and Statistics Canada's Postal Code Conversion File to ensure I had all possible locations (due to additions and retirements throughout the study period, this merging ensured I did not miss any). Because the assignment of the raster data to each location was restricted to within the 1 km cell size of the densities and proportions (described above), I converted the merged postal codes to 1-km raster and then converted the cell centers back to points (Figure 2.2). This reduced the set of coordinates that represented all possible 2 million+ postal codes to 239,711 records. To these 1-km spaced unique locations, I assigned the values from the densities and proportions: NPRI and roads were annual (i.e. I made 8 copies for 2005 to 2012); and all other environmental variables were static.

Exposure calculations

Exposure was assigned as the value of the environmental variable that occurred closest to the maternal residence. I spatially joined the 1-km environmental variables to every valid location in the VSBD. The values of the static variables (see the Year column in Table 2.2) were used directly. The values of the annual variables (NPRI emissions and roads) were calculated based on the proportion of the pregnancy that occurred in one or two years, according to the equation:

$$\mathrm{T} = \mathrm{bP} \times \mathrm{bEXP} + \mathrm{cP} \times \mathrm{cEXP}$$

where, T is the total exposure, bP is the proportion of gestation in the birth year (calculated as number weeks in birth year / total number weeks gestation), bEXP is the exposure in the birth year, cP is the proportion of gestation in the conception year (calculated as number weeks in conception year / total number weeks gestation), and cEXP is the exposure in conception year. Therefore, when the birth year = conception year, then bP = 1 and cP = 0, and when the actual birth date is during the 1st week of January, then bP = 0 and cP = 1. The calculation assumes homogeneous emissions throughout each year. Python programing language [111] automated the cumbersome calculations for the 2.5 million birth records.

Logistic regression

I performed logistic regression on the binary values of the health outcomes (SGA and LBWT) for each environmental variable: NPRI emission or land source. I calculated the beta coefficients for the crude models (SGA/LBWT ~ exposure) and controlled (SGA/LBWT ~ exposure + covariates) for total births for the year, female baby, mothers 19 years and younger, mothers 40 years and older, migration, urban, proportion of low SES neighborhood, and road density (i.e. traffic pollution surrogate). I also calculated the same models by including the three variables indicating whether the warm season occurred during the majority of the first trimester, second trimester, or the 3 months prior to birth. A Canada-wide analysis was not done because the country is an area deemed too large and variable for such an analysis to be meaningful, and an objective was to compare provinces and territories. Because I was interested only in the significance of the effect of one independent variable (X) on the response (Y), and there was no need of interpreting the coefficients, the coefficients were calculated (i.e. logarithm of the odds ratios).

2.4 Results

From a total of 2,620,415 births in Canada during 2006 to 2012, there were 2,525,645 births (96%) meeting my criteria: single, live, between 22 to 42 weeks gestation, and having a valid 6-character postal code (Figure 2.1). Of these, 215,940 (8.55%) were SGA and 38,990 (1.54%) were LBWT. Table 2.3 shows the provincial counts for births, SGA, and LBWT: Ontario had the most (941,085, 84,370, and 15,810) and Yukon had the least (2,615, 160, and 30). For context, Table 2.3 also provides the land areas and total population from the 2011 Canada Census [26].

Table 2.3. Descriptive statistics of the study population from Census Canada 2011 and the Vital Statistics–Birth Database.

Province / Territory	**Abbreviation**	**Area (km^2)**	**Total Population**	**Births**	**SGA**	**LBWT**
Alberta	AB	640,082	3,645,257	338,535	30,600	5,515
British Columbia	BC	922,509	4,400,057	296,025	22,880	3,815
Manitoba	MB	552,330	1,208,268	105,660	8,545	1,530
New Brunswick	NB	71,377	751,171	48,975	3,795	675
Newfoundland	NL	370,511	514,536	31,525	2,215	415
Northwest Territories	NT	1,143,793	41,462	4,770	290	55
Nova Scotia	NS	52,939	921,727	59,925	5,130	965
Nunavut	NU	1,877,788	31,906	5,515	340	95
Ontario	ON	908,608	12,851,821	941,085	84,370	15,810
Prince Edward Island	PE	5,686	140,204	9,530	605	110
Quebec	QC	1,356,547	7,903,001	587,845	49,855	8,690
Saskatchewan	SK	588,239	1,033,381	93,645	7,155	1,275
Yukon	YT	474,713	33,897	2,615	160	30

Figure 2.3 maps out the provincial distributions of SGA and LBWT percentages. Male SGA was typically higher and female LBWT was consistently lower in all the provinces and territories. For SGA overall, Alberta had the highest (9.04%) and Northwest Territory had the lowest (6.08%). For male SGA, Ontario had the highest (9.10%) and Yukon had the lowest (5.84%). For female SGA, Alberta had the highest (9.01%) and Northwest Territory had the lowest (5.77%). For overall LBWT, Nunavut had the highest (1.68%) and Prince Edward Island had the lowest (1.15%). For male LBWT, Nunavut had the highest (1.40%) and Yukon had the lowest (0.73%). For female LBWT, Nunavut had the highest (2.06%) and Northwest Territory had the lowest (1.28%).

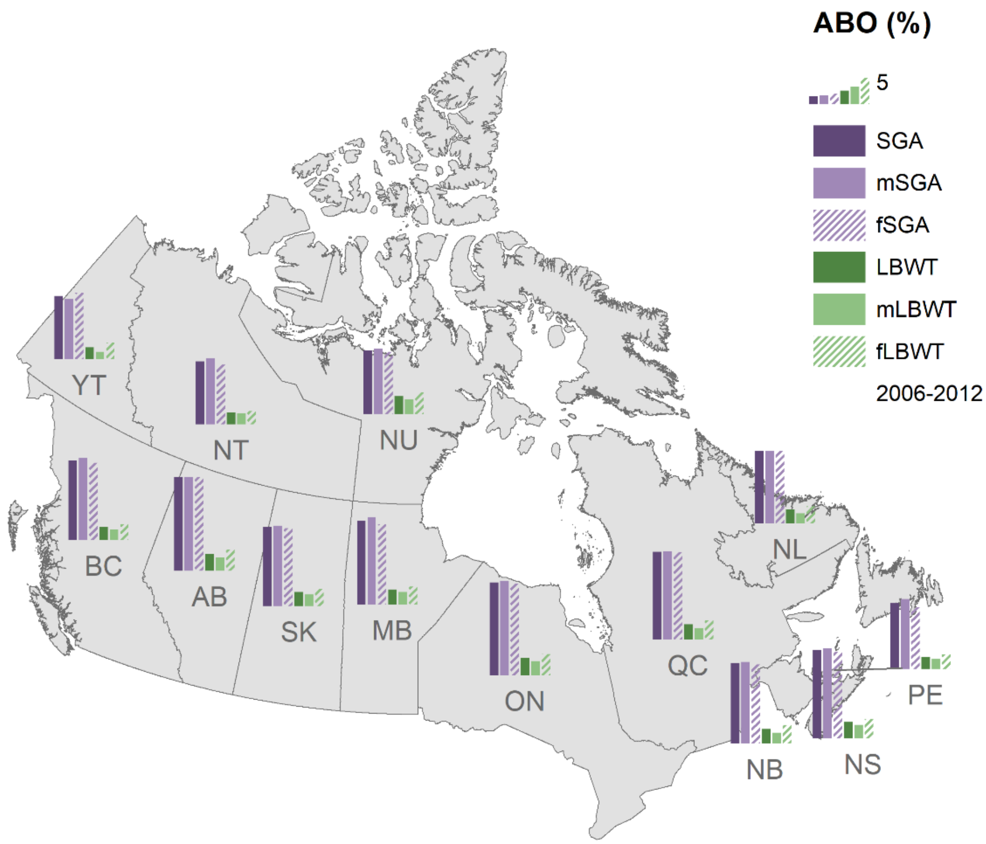


Figure 2.3. Distribution of small for gestational age (SGA) and low birth weight at term (LBWT) for all, male (m), and female (f) babies born in Canada, 2006-2012.

A total of 13,558 facilities released 30,855,608 tonnes of 228 unique chemicals to the air, primarily from the energy (electricity and oil/gas) and mining-related sectors (Appendix I: Figure S1.1). Table 2.4 shows that Alberta had the most facilities (n=6,643) and Yukon had the least (n=4). Ontario had the most chemicals (n=201) and Yukon had the least (n=5). Alberta emitted the most (9,004,138 tonnes) and Yukon emitted the least (5,147 tonnes). Figure 2.4 maps out the provincial distributions as percentages. Alberta had 49% of the facilities, followed by Ontario (20%) and Saskatchewan (10%). Alberta emitted 29% of the total national tonnes, followed by

Québec (18%) and Ontario (17%). Ontario had 88% of the total unique chemicals, followed by Québec (71%) and then Alberta (61%). Relative amounts of each chemical emitted for each province is available in Appendix I: Figures S1.2 through S1.5.

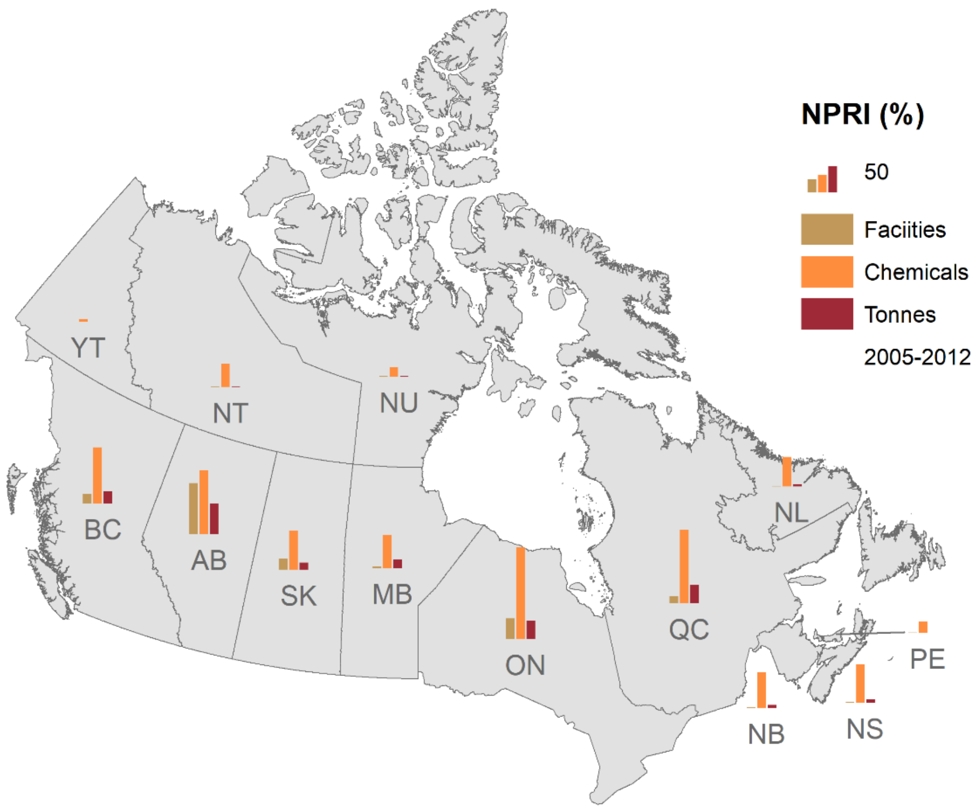


Figure 2.4. Distribution of industrial air emissions by province shown as percentage of the national total for number of facilities (gold), number of chemicals (orange), and tonnes emitted (red).

Table 2.4. Descriptive statistics of the industrial air emissions from the National Pollutant Release Inventory.

Province / Territory	Abbreviation	Number of Facilities	Number of Chemicals	Total Tonnes
Alberta	AB	6,643	140	9,044,138
British Columbia	BC	1,271	123	3,612,378
Manitoba	MB	198	72	2,531,479
New Brunswick	NB	107	78	795,615
Newfoundland	NL	91	65	664,952
Nova Scotia	NS	116	85	1,188,724
Northwest Territories	NT	42	51	96,493
Nunavut	NU	31	20	43,092
Ontario	ON	2,687	201	5,395,499
Prince Edward Island	PE	13	25	14,649
Quebec	QC	932	162	5,445,992
Saskatchewan	SK	1,423	85	2,017,450
Yukon	YT	4	5	5,147

Table 2.5 shows the provincial land-based hazards for powerlines, dumps, gas stations, mines, oil/gas well pads, transformer stations, and crop lands. The percentage of each is shown in Appendix I: Figure S1.6. Québec had the most powerlines (24%). Ontario had the most dumps (32%), gas stations (32%), mines (24%), and transformer stations (46%). Alberta had the most oil/gas well pads (72%). Saskatchewan had the most crop lands (42%).

Table 2.5. Descriptive statistics of the land hazards. Unless measurement unit is indicated in parentheses, values are counts.

Province / Territory	**Abbreviation**	**Powerlines (km)**	**Dumps**	**Gas Stations**	**Mines**	**Well Pads**	**Transformers**	**Crop land (km^2)**
Alberta	AB	9,366	486	1,453	1,956	59,785	435	143,226
British Columbia	BC	10,715	415	1,678	1,938	2,743	289	8,836
Manitoba	MB	8,773	396	517	1,559	851	332	52,757
New Brunswick	NB	4,298	281	231	1,616		165	3,275
Newfoundland	NL	5,242	284	188	1,028	3	139	
Nova Scotia	NS	3,814	171	336	963		166	2,135
Northwest Territories	NT	589	63	6	165	30	8	
Nunavut	NU	22	10	2	29	14		
Ontario	ON	20,428	1,887	3,374	5,442	15	2,028	52,232
Prince Edward Island	PE	241	28	57	137	1	14	2,480
Quebec	QC	22,744	1,207	2,165	5,382	2	611	27,636
Saskatchewan	SK	9,127	599	493	1,964	20,158	257	210,113
Yukon	YT	642	151	20	277	2	4	

The adjusted beta coefficients and p-values ($p<0.05$ are in bold) from the logistic regressions are shown in Table 2.6. Crude values were almost identical and therefore not included. For simplicity, SGA and LBWT having associations greater than 0.2 with the industrial emissions and land hazards are highlighted. Results for all exposures in each province is in Appendix I: Figures S1.2-S1.5. The highest statistically significant associations ($p<0.05$) for Alberta SGA was styrene ($\beta=0.721$, $p<0.009$) and LBWT was carbonyl sulphide ($\beta=0.959$, $p<0.026$). British Columbia SGA was most associated with dumps ($\beta=0.522$, $p<0.002$) and LBWT was well pads ($\beta=1.039$, $p<0.052$). Manitoba SGA was most associated with dumps ($\beta=0.639$, $p<0.009$) and LBWT was with gas stations ($\beta=0.328$, $p<0.001$). New Brunswick SGA was most associated with transformer stations ($\beta=0.585$, $p<0.044$) and LBWT was with gas stations ($\beta=0.831$, $p<0.000$). Newfoundland SGA was most associated with ammonia ($\beta=0.432$, $p<0.024$) and LBWT had no significant associations. Northwest Territory SGA and LBWT had no significant associations. Nova Scotia SGA was most associated with gas stations ($\beta=0.205$, 0.002) and LBWT was not significant. Nunavut SGA was most associated with carbon monoxide ($\beta=1.400$, $p<0.032$) and LBWT was with $PM_{2.5}$ ($\beta=8.804$, $p<0.038$). Ontario SGA was most associated with methyl ethyl ketone ($\beta=0.085$, $p<0.009$) and LBWT was with benzene ($\beta=1.093$, $p<0.032$). Prince Edward Island SGA was most associated with gas stations ($\beta=0.838$, $p<0.004$) and LBWT was not significant. Québec SGA was most associated with transformer stations ($\beta=0.442$, $p<0.000$) and LBWT was with i-Butyl alcohol ($\beta=0.934$, $p<0.024$). Saskatchewan SGA was most associated with transformer stations ($\beta=0.732$, $p<0.000$) and LBWT was with dumps ($\beta=1.247$, $p<0.020$). Yukon SGA had no significant associations and LBWT was most associated with PM_{10} ($\beta=3.785$, $p<0.053$).

Table 2.6. Logistic regression coefficients, adjusted by covariates, by each province for small for gestational age (SGA) or low birth weight at term (LBWT). p<0.05 in bold. See Appendix I: Table S1 for all results. Table continues.

Variables	**SGA**		**LBWT**	
	Beta	**p**	**Beta**	**p**
Alberta				
Carbonyl sulphide	0.445	0.074	**0.959**	**0.026**
HCFC-142b	-0.176	0.402	0.326	0.404
Hydrogen fluoride	0.239	0.107	0.411	0.115
Styrene	**0.721**	**0.009**	0.728	0.226
Total Reduced Sulphur (TRS) - [MOE]	0.180	0.200	0.377	0.178
Dump	**0.515**	**0.001**	0.317	0.352
British Columbia				
Carbonyl sulphide	0.417	0.589	1.626	0.122
Hydrogen fluoride	0.171	0.702	0.823	0.269
Nickel (and its compounds)	0.471	0.651		
Sulphuric acid	**0.362**	**0.001**	-0.061	0.871
Toluene	-0.437	0.888	1.893	0.526
Vanadium (fume or dust)	1.342	0.089		
Xylene (mixed isomers)	-0.053	0.966	1.702	0.224
Dump	**0.522**	**0.002**	**0.823**	**0.039**
Well pad	0.391	0.272	**1.039**	**0.052**
Powerline	**0.145**	**0.000**	**0.299**	**0.000**
Transformer	0.218	0.141	0.068	0.846
Manitoba				
Dump	**0.639**	**0.009**	0.965	0.076
Gas station	**0.234**	**0.000**	**0.328**	**0.001**
Transformer	**0.587**	**0.000**	0.543	0.071
New Brunswick				
Sulphuric acid	0.479	0.128	0.581	0.356
Crop	0.223	0.136	-0.065	0.854
Dump	**0.903**	**0.060**	1.559	0.145
Gas station	**0.509**	**0.000**	**0.831**	**0.000**
Transformer	**0.585**	**0.044**	0.540	0.415
Newfoundland				
Ammonia (total)	**0.432**	**0.024**	0.454	0.274
Methanol	0.228	0.678	**1.354**	**0.071**
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	0.127	0.256	0.265	0.262
Total Reduced Sulphur (TRS) - [MOE]	-0.141	0.849	0.906	0.381
Dump	0.563	0.383	0.682	0.636
Gas station	**0.290**	**0.039**	**0.216**	**0.487**
Well pad	0.929	0.876		
Transformer	-0.079	0.858	1.302	0.163

Variables	SGA		LBWT	
Northwest Territory				
Carbon monoxide	0.248	0.441	0.127	0.858
PM_{10} - Particulate Matter ≤ 10 Microns	1.458	0.411	-3.342	0.461
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	1.481	0.403	-3.381	0.456
Volatile Organic Compounds (VOCs)	0.317	0.108		
Dump	0.477	0.265	1.329	0.191
Gas station	1.941	0.115	2.607	0.415
Mine	-0.259	0.879	2.636	0.410
Nova Scotia				
Dump	0.334	0.445	0.662	0.487
Gas station	**0.205**	**0.002**	0.078	0.591
Transformer	0.267	0.109	0.109	0.768
Nunavut				
Carbon monoxide	**1.400**	**0.032**	**3.436**	**0.003**
Nitrogen oxides (expressed as NO_2)	0.118	0.472	0.396	0.195
PM_{10} - Particulate Matter ≤ 10 Microns	2.969	0.208	**8.778**	**0.039**
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	2.980	0.206	**8.804**	**0.038**
Sulphur dioxide	0.969	0.304	1.827	0.164
Well pad	1.324	0.304	-0.433	0.855
Ontario				
2-Butoxyethanol	-1.159	0.497	1.424	0.361
Benzene	0.487	0.128	**1.093**	**0.032**
Calcium oxide	0.304	0.160	-0.117	0.840
GE - Diethylene glycol butyl ether (DEGBE)	0.211	0.271	0.313	0.416
HCFC-22	1.418	0.220		
i-Butyl alcohol	-0.274	0.481	0.418	0.284
Manganese (and its compounds)	1.001	0.078		
Methyl ethyl ketone	**0.085**	**0.009**	**0.204**	**0.002**
MSG#1 - Hydrotreated heavy naphtha	0.195	0.343	0.523	0.061
Styrene	1.027	0.113	1.608	0.126
Tetrahydrofuran	0.487	0.659		
Trichloroethylene	-0.281	0.522	0.586	0.417
White mineral oil	2.073	0.237		
Prince Edward Island				
Dump	0.325	0.882	-6.480	0.273
Gas station	**0.838**	**0.004**	0.866	0.159
Transformer	1.107	0.517	-2.368	0.541
Québec				
Formaldehyde	0.886	0.197		
HCFC-142b	**0.320**	**0.000**	-0.177	0.488
i-Butyl alcohol	-0.120	0.788	**0.934**	**0.024**
Isopropyl alcohol	-0.120	0.788	**0.934**	**0.024**
Methyl ethyl ketone	0.062	0.657	0.316	0.207
n-Hexane	0.156	0.615	0.583	0.261
Toluene	0.091	0.566	0.429	0.159

Variables	SGA		LBWT	
Trichloroethylene	0.337	0.330	-0.720	0.554
Well pad	6.675	0.098	**17.603**	**0.001**
Powerline	**-0.056**	**0.000**	**-0.065**	**0.025**
Transformer	**0.442**	**0.000**	0.350	0.079
Saskatchewan				
Hydrogen fluoride	0.448	0.747		
Total Reduced Sulphur (TRS) - [MOE]	-0.181	0.776	0.540	0.615
Dump	**0.586**	**0.018**	**1.247**	**0.020**
Gas station	**0.203**	**0.000**	0.137	0.155
Transformer	**0.732**	**0.000**	0.233	0.599
Yukon				
Carbon monoxide	0.428	0.265	1.090	0.085
Nitrogen oxides (expressed as NO_2)	0.085	0.322	0.250	0.072
PM_{10} - Particulate Matter ≤ 10 Microns	1.527	0.211	**3.785**	**0.053**
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	1.546	0.204	**3.771**	**0.055**
Mine	-2.326	0.162	2.408	0.469
Transformer	5.213	0.489		

The exposures having the most associations with provincial SGA were gas stations (n=6), dumps (n=4), transformer stations (n=4), and powerlines (n=1). Ammonia, carbon monoxide, HCFC-142b, methyl ethyl ketone, styrene, and sulphuric acid were the industrial emissions associated with SGA. Provincial LBWT was associated with PM_{10} (n=2), dumps, (n=2), gas stations (n=2), and powerlines (n=1). Benzene, carbon monoxide, carbonyl sulphide, i-butyl alcohol, isopropyl alcohol, methyl ethyl ketone, $PM_{2.5}$ were also associated with LBWT.

The provinces having statistically significant associations with only chemical exposures were Ontario, Nunavut, and Yukon (Figure 2.5). Manitoba, New Brunswick, Nova Scotia, Prince Edward Island, and Saskatchewan had associations with only land hazards. Alberta, Newfoundland, and Québec were primarily associated with chemicals and British Columbia was mostly land hazards. Figure 2.5 only shows statistically significant patterns for $\beta>0.2$ ($p<0.05$).

The differences between beta coefficients for models that included the seasonal variables (i.e. if the warm season occurred during the first trimester, second trimester, or 3 months prior to birth) were negligible (±0.001; see Appendix I: Table S1.1).

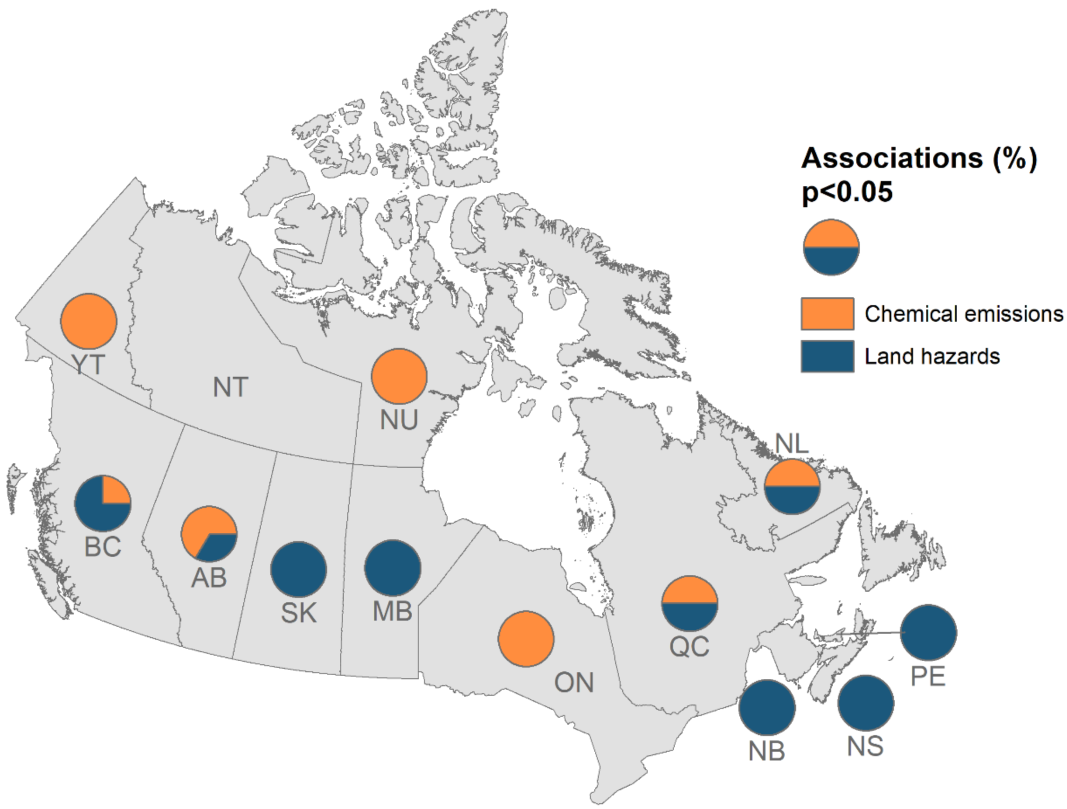


Figure 2.5. Distribution of chemical emissions (orange) and land hazards (blue) having relatively stronger and statistically significant associations (β> 0.2 and p<0.05) with small newborns.

2.5 Discussion

Maternal proximity to environmental variables potentially representing ambient health hazards and their associations with small newborns differed by province. Ontario, Québec, Alberta, and British Columbia had more associations, almost certainly due to the higher levels of industrial activity. There were no industrial chemicals having associations with more than one province, but dumps, gas stations, powerlines, transformer stations, and well pads had associations with two or more provinces. Overall, the land hazards had associations with more provinces, especially in the prairie and Atlantic provinces. Chemical emission associations were stronger in Ontario, Alberta, and the territories; the latter was surprising since the territories have fewer facilities, chemicals, and emissions compared to the rest of Canada. The lower land hazard associations in the territories were likely due to less infrastructure, such as dumps, gas stations, etc. The lack of land hazard associations in Ontario was surprising and may be because this province has the largest overall chemical emission pattern. I suggest that these differences in associations may reflect the multifactorial nature of the adverse birth outcomes. SGA includes preterm births making up the majority, and the potential exposures likely differ from LBWT, which are full term births.

The environmental variables were carefully chosen based on published studies [112] and accessibility of standard map data for the entire country. Previous research on dumps/waste sites [113, 114], gas stations [115], and power lines/electrical infrastructure [65, 66] have also indicated associations with a variety of birth outcomes. The majority of the industrial chemicals discovered – ammonia, benzene, carbon monoxide, isopropyl alcohol, methyl ethyl ketone, PM_{10} (particulate matter ≤ 10 microns), $PM_{2.5}$ (particulate matter ≤ 2.5 microns, styrene, and sulphuric acid – are suspected or known developmental toxicants [116, 117].

Season did not have an effect on either SGA or LBWT. Perhaps, the static warm season definition did not fit for all, mostly due to the vast latitudinal differences among and within the provinces/territories. Also, a finer temporal resolution conducive to seasonal estimations of the ambient health hazards is likely needed.

Limitations

The provincial scale is likely too large to identify all potential associations because the settlement and land use patterns differ within. Future research on smaller geographic areas is suggested, especially for assessing which industrial chemicals are important. Facilities likely cluster by industrial sector according to regional geographies, emitting different groups of chemicals in different parts of the provinces. Province-wide analysis using global regression methods cannot detect this. Thus, scale may also be the reason that industrial chemicals were not as strongly associated as would be expected for the more industrial areas.

Rural postal codes are too spatially inaccurate to give reliable results. This is particularly problematic when assigning density values of typically rural land hazards (e.g. crop land and well pads) where actual residences may be situated much closer to these land hazards than the postal code delivery locations.

The use of areal units of analysis underscores the risk of ecological fallacy (aggregation bias) [9]; i.e. it must be remembered that not all births in the postal code areas may have been SGA/LBWT.

Depending on alternative objectives (e.g. in epidemiology or planning policy), the reporting of coefficients (log of odds ratios) from the logistic regression model may not be suitable. Odds ratios are more easily interpretable as how much the levels of one variable (X = 1 = exposure) affects Y in relation to a reference for X (i.e. X = 0 = no exposure). The beta coefficients were useful for investigating whether any associations existed. More sophisticated statistical analyses to explore interactions of the environmental variables may be performed in the future.

My methods relied on proximity, but the greater amount or density of outdoor environmental hazards closer to populations did not equate to individual-level exposures. Future research is encouraged to include biomonitoring and lab experiments in non-human models to investigate potential causation of the factors identified here.

Strengths

The identified associations lead toward a more holistic environmental health approach to understanding SGA and LBWT factors. I have applied easily accessible GIS tools and developed methods to aid in population-level exposures of maternal ambient health hazards and really small newborns. My use of publicly available spatial databases means the research may also be applied to any health issue where exposure to the outdoor environment may be of concern.

Conclusion

This is the first study to demonstrate provincial differences in associations of multiple maternal ambient health hazards with SGA and LBWT in Canada. My calculation of the Canadian percentage of SGA, based on my delimited population, was still very similar to the government published Canadian three-year averages over the study period. Because I only included newborns at term, my calculation of Canadian percentage of LBWT was lower. Standard mapping of the hazards allowed for comparisons among different geographic areas, further recognizing that adverse birth outcomes are indeed multifactorial. Pollutants released to the air or land-based hazards may be more important depending on where one lives. Further research is needed to determine causation, but my results invoke the precautionary principle to do no harm. Therefore, increasing our understanding of these possible risk factors from the shared outdoor environment will aid planners and health professionals in identifying these potential issues for prevention and development of intervention solutions to reduce future occurrence of babies born too small.

Chapter 3 Mapping outdoor habitat and abnormally small newborns to develop an ambient health hazard index

3.1 Abstract

Background: The geography of where pregnant mothers live is important for understanding outdoor environmental habitat that may result in adverse birth outcomes. I investigated whether more babies were born small for gestational age or low birth weight at term to mothers living in environments with a higher accumulation of outdoor hazards.

Methods: Live singleton births from the Alberta Perinatal Health Program, 2006-2012, were classified according to birth outcome, and used in a double kernel density estimation to determine ratios of each outcome per total births. Individual and overlay indices of spatial models of 136 air emissions and 18 land variables were correlated with the small for gestational age and low birth weight at term, for the entire province and sub-provincially.

Results: There were 24 air substances and land sources correlated with both small for gestational age and low birth weight at term density ratios. On the provincial scale, there were 13 air substances and 2 land factors; sub-provincial analysis found 8 additional air substances and 1 land source. Air substances having rho>0.4 for both SGA and LBWT were i-Butyl alcohol, Asbestos (friable form), Toluenediisocyanate (mixed isomers), Toluene-2,4-diisocyanate, Toluene-2,6-diisocyanate, Chromium (and its compounds), Aluminum (fume or dust), Hydrogen sulphide, 2-Ethoxyethanol, Nickel (and its compounds), Quinoline (and its salts), Aniline (and its salts), Cyclohexane, Acetaldehyde; land sources having rho>0.4 were nighttime lights and roads.

Conclusion: This study used a combination of multiple outdoor variables over a large geographic area in an objective model, which may be repeated over time or in other study areas.

The air substance-weighted index best identified where mothers having abnormally small newborns lived within areas of potential outdoor hazards. However, individual air substances and the weighted index provide complementary information.

3.2 Background

A truly ecologically-based study of health integrates habitat, population, and behavior – encompassing a more complete geography as framed by Meade's triangle of human ecology [27, 29]. Three vertices conceptualize what is known about an important pediatric topic: maternal exposure to outdoor pollution and neonatal outcomes (Figure 3.1). Here I focus on the lesser studied habitat vertex, specifically the outdoor environment, since less attention is traditionally given to incorporating ecological factors in understanding disease [27]. The location aspect of habitat – where pregnant women live, where industry and services are situated, where demographic groups congregate – is important because where one lives and where one starts out in life, even during fetal development, ultimately influences lifelong health [1, 10–12].

Toxicant exposures and environmental influences on mothers during crucial stages of pregnancy may result in newborns that are too small or born too early. Adverse birth outcomes (ABO) are important markers of infant survival, development and future health. My research focuses on being born too small, clinically defined as Small for Gestational Age (SGA) when newborns are below the 10th percentile weight based on sex and weeks of pregnancy, or Low Birth Weight at Term (LBWT) when newborns are less than 2,500 grams weight at term, 37 or more weeks gestation [14].

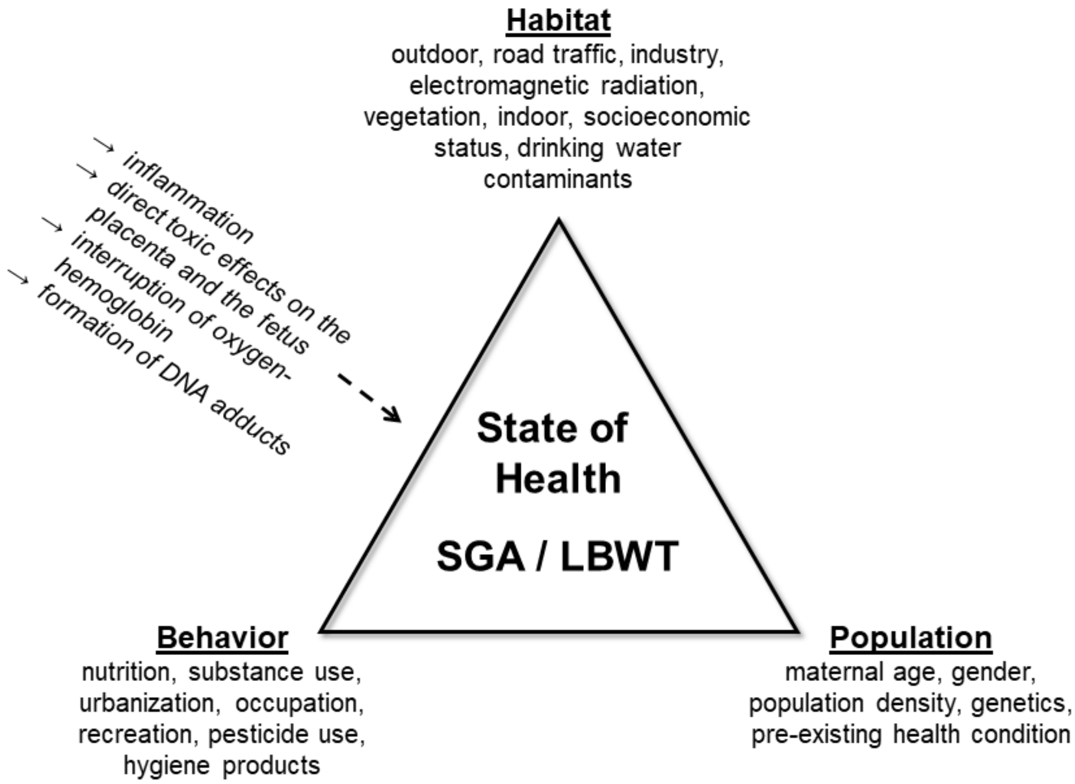


Figure 3.1. Meade's triangle of human ecology for maternal exposures and small for gestational age (SGA) and low birth weight at term (LBWT): dashed arrow indicates hypothesized mechanisms.

The province of Alberta, Canada, had a population of 3,645,257 at the 2011 Census [118]. That was a 10.8% increase from 2006 while the national average increase was 5.9%. For a land area of 640,082 km^2, the population density was 5.7 persons/km^2, where 83% of the population lived in or near urban centers. Alberta's economic activities were focused on agriculture, natural resources, and nonrenewable energy – having a higher number of industrial facilities reporting to the National Pollutant Release Inventory (NPRI) than any other province or territory [105]. The NPRI is a valuable data source on industrial-based pollutants [99]. Alberta also has higher

ABOs: SGA was 8.8% (Canada was 8.4%); and low birth weight for all gestational ages was 6.7% (Canada was 6.0%) [23]. Alberta rates also increased during 2000-2014: SGA from 10% to 11.5%; and low birth weight for all gestational ages from 6.1% to 6.7% [119].

SGA/LBWT complications include death, physical and cognitive disabilities, and chronic health problems later in life, costing emotional stress and the majority of the health care expenses among all newborns [7]. Disorders related to short gestation and low birth weight are consistently ranked as the 2nd leading cause of infant death (congenital deformation is the leading cause) [24] – and have increased in Canada since the year 2000 [23, 119].

Abnormally small newborns are the result of growth restriction, which may be due to environmental pollutants thought to cause inflammation in mothers, direct toxic effects on the placenta and the fetus, interruption of oxygen-hemoglobin interaction, or DNA damage represented by the formation of DNA adducts [81, 82].

The environment includes social, built, and natural features. Individual risks are also very important to ABOs but are often not available nor easily mapped. These include personal, behavioral, social, and indoor exposures, such as: adequate prenatal care; food type and contaminants; rest, stress, and pre-existing health conditions; occupation and socioeconomic status; smoking and other substance use; drinking water contaminants. My focus is on the outdoor environmental habitat because it is a common source of shared exposures susceptible to regulation (biology and behavior are not). These include air, water, human-constructed, and natural outdoor hazards, such as: industrial emissions; traffic pollution; agricultural chemical inputs of pesticides, herbicides, and fertilizers; electromagnetic radiation; proximity to oil and gas extraction activities; wildfire smoke.

Environmental health research has found many environmental factors to be associated with various health outcomes [44, 79, 120–126]. However these are typically explored singly: one exposure or category of exposure at a time. A unified environmental measure may be constructed across multiple variables to encompass the complex nature of the overall environment.

Environmental indices have history: Inhaber had proposed an integrated national index for Canada in the 1970s [127]. Rather than relying on individual pollutants to reflect the state-of-the-environment, Inhaber mathematically combined such indicators for the purpose of resource allocation, ranking of locations, enforcement of standards, trend analysis, public information, or scientific research [128]. Under that premise, Messer et al. [129] developed a California county-level environmental quality index using principal components analysis (PCA) to calculate five environmental domains (air, water, land, built, and sociodemographic), which were then combined into a single index using PCA on the first components, and stratified by rural-urban continuum codes. Similarly, Messer et al.'s CalEnviroScreen 2.0 [130, 131] superimposed 19 individual indicators that related to pollution exposures, environmental conditions, and population characteristics, weighted and summed each set of indicators, and then multiplied together pollution and population (i.e. Threat × Vulnerability = Risk). I have not found similar environmental health indices available for Canada, or the province of Alberta, and especially none focused primarily on maternal exposures associated with ABO.

Through the use of a Geographical Information System (GIS), I developed a simplified and reproducible index for Alberta by estimating and aggregating pollutants from communal outdoor factors. GIS supports the inclusion of diverse data and enables modelling of hazard-exposure-dose-response processes in space [32, 33]. To capture the relevant pollutant estimates, spatially and temporally appropriate GIS data files were overlaid to develop a vulnerability map of combined disparate (in theme and measurement units) environmental factors, similar to Messer et al. [129, 130]. The index will aid my examination of maternal ambient health hazards and abnormally small newborns by providing a relative ranking of locations across the province that are not limited by administrative boundaries.

My research is part of the Data Mining and Neonatal Outcomes (DoMiNO) project that is exploring the colocation of adverse birth outcomes and environmental variables in Canada [109]. For my geographical perspective on the project I hypothesized that SGA or LBWT babies are more likely to be born to mothers living in environments with a higher number of outdoor hazards (especially pollutants) than in relatively healthier habitats with fewer exposure hazards. My objective was to examine how the separate and combined exposures to the outdoor built-up, natural, and social environments of pregnant mothers coincided with patterns in SGA/LBWT. I also expected that the large Alberta province would have regional variations in the outdoor environment and investigated this effect on the associations.

3.3 Methods

GIS parameters

I used Esri's ArcGIS Desktop 10.5 software to perform all spatial database processing, management, distribution analyses, hazard estimations, and index calculations [101]. Proximity was extremely important in my spatial analysis; therefore, I customized an Alberta-focused map projection, based on the following parameters: name Azimuthal Equidistant; central meridian -113.5; latitude of origin 53.5; linear unit meter (1.0); and geographic coordinate system (GCS) with North American Datum 1983 (NAD 1983). I projected all GIS data to this distance-preserving spatial reference.

For raster files I used a 250 m by 250 m cell size to reasonably represent both urban and rural areas in the very large study area, and to match the coarsest dataset: MODIS Terra satellite [132].

Because Alberta is landlocked, I included data features within 50 km surrounding the provincial boundary where available: by doing so, any potential pollutant source close to the outer edge of the province was included.

Regional attribution

I produced sub-provincial maps of the percent ratios for each SGA/LBWT to facilitate comparisons that are more meaningful to health care and environmental management. I assigned administrative attributes to postal code locations. This allowed grouping by health region [133] or airshed zone [134] because both are health-related administrative boundaries that help identify where there may be different outdoor factors of importance.

Health regions are designated by the provincial Ministry of Health to identify geographic areas where hospital boards or regional health authorities administer and deliver public health care, and are subject to change [135]. At the start of my study period, there were nine health regions for Alberta (Table 3.1): Chinook Regional Health Authority (4821); Palliser Health Region (4822); Calgary Health Region (4823); David Thompson Regional Health Authority (4824); East Central Health (4825); Capital Health (4826); Aspen Regional Health Authority (4827); Peace Country Health (4828); and Northern Lights Health Region (4829).

Airshed zones are endorsed by the multi-stakeholder Clean Air Strategic Alliance to identify geographic areas where the air quality is similar in emission sources, volumes, impacts, dispersion and administrative characteristics [136]. Because Alberta has several unique topographical, meteorological, or ecological conditions for resolving air quality, there are nine airsheds currently recognized (Table 3.1): Alberta Capital Airshed Alliance (ACAA); Calgary Region Airshed Zone (CRAZ); Fort Air Partnership (FAP); Lakeland Industry and Community Association (LICA); Palliser Airshed Society (PAS); Parkland Airshed Management Zone (PAMZ); Peace Airshed Zone Association (PASZA); West Central Airshed Society (WCAS); and Wood Buffalo Environmental Association (WBEA). It is important to note that the entire province is not monitored by airshed zones, with the southwest corner, east-central, and majority of the north having no airshed (NA).

Table 3.1. Alberta's sub-provincial units and descriptive statistics, in descending order of birth number.

Unit	Map Code	Name	Area (km^2)	Postal Codes	Geolocated Births	SGA	LBWT
Province	none	Alberta	663,563	53,399	333,247	29,679	5,485
Health Region	4823	Calgary Health Region	39,350	20,537	121,965	12,543	2,339
	4826	Capital Health	11,883	20,004	99,691	8,596	1,566
	4824	David Thompson Regional Health Authority	61,578	3,325	29,766	2,394	476
	4827	Aspen Regional Health Authority	137,639	1,440	18,004	1,252	222
	4821	Chinook Regional Health Authority	26,062	2,406	16,639	1,342	233
	4828	Peace Country Health	123,870	1,580	16,428	1,188	215
	4829	Northern Lights Health Region	189,696	1,073	11,097	808	147
	4822	Palliser Health Region	39,772	1,723	9,920	858	147
	4825	East Central Health	33,812	1,311	9,737	698	140

Unit	Map Code	Name	Area (km^2)	Postal Codes	Geolocated Births	SGA	LBWT
Airshed Zone	CRAZ	Calgary Regional Airshed Zone	32,372	20,530	120,392	12,409	2,310
	ACAA	Alberta Capital Airshed Alliance	4,933	19,474	95,085	8,284	1,503
	NA	No Airshed zone	362,439	4,867	47,527	3,509	647
	PAMZ	Parkland Airshed Management Zone	40,936	2,774	24,896	1,978	387
	PASZA	Peace Airshed Zone Association	45,892	1,409	12,475	927	175
	PAS	Palliser Airshed Society	39,900	1,723	9,920	858	147
	WBEA	Wood Buffalo Environmental Association	69,214	1,061	7,540	627	115
	WCAS	West Central Airshed Society	47,142	612	7,386	559	107
	LICA	Lakeland Industrial Community Association	16,215	455	4,479	293	43
	FAP	Fort Air Partnership	4,519	494	3,547	235	51

Dependent variables

The Alberta Perinatal Health Program (APHP) provided anonymized data for the province of Alberta, from 2006 to 2012 [25].

I selected for live single births between 22 and 42 weeks gestation and geocoded them to the centroid of the 6-character postal code of the mothers' residences at the time of the birth registration. DMTI Spatial's Platinum Postal Code Suite [110] provided the longitude and latitude coordinates for the years 2001 to 2013, which I uniquely selected to guarantee static locations through the entire study period. 95% of the original data had valid coordinates for use in spatial analyses. Using the previous definitions, I classified the birth records as binary variables identifying SGA or LBWT. Details are available in Serrano-Lomelin [88].

To eliminate the confines of arbitrary administrative boundaries, I followed the double kernel density (DKD) method [137–142] to calculate distributions of SGA and LBWT, normalized by all births. DKD involves kernel density estimation – a non-parametric method that spreads point values across a surface by calculating the magnitude-per-unit area from points (representing the counts of birth events), fitted to a smoothly tapered function that spreads the values within a specified distance (25 km for this study) around each point [108]. Points within the radius that are further from the center are weighted lower than those closer and helps indicate "hot spots." Dividing each SGA/LBWT by the kernel density of total births yielded ratios of the birth outcome that also masked locations of the residences, helping protect privacy.

Independent variables

Personal maternal monitoring data were not available for this retrospective study. I used landscape features as spatial proxies of exposure hazards, as done in previously published research [33]. In total, I chose 18 outdoor sources, identified in published studies [44, 120–125] or added for novel exploration (10 built; five social; three natural) plus 136 industrial air substance emissions. Table 3.2 lists the environmental variables and indicates specific characteristics and processing details.

I applied kernel density to spread industrial emissions from the NPRI database as tonnes per area within a 10 km radius (based on distances determined from the project's data mining algorithm [109]). I used the count of other point features – industrial facilities, gas stations, waste/landfills, oil/gas well pads, food stores, and health care/hospitals – in kernel density to calculate the number per area within a 3 km radius. I also applied kernel density to roads and electrical power lines to calculate length per area within a 3 km radius. A main advantage of using kernel density is it accounts for distance decay (features have less influence further away). When linear features are the input it also helps to approximate the number of intersections – important when analyzing pollution sources from roads because vehicles idle at intersections.

For areal features, I used focal statistics, also known as moving-window or neighborhood analyses, on binary surfaces of feedlots, mine sites, cultivated lands, aboriginal lands, water/blue space, and wildfires. The mean statistic on binary values of 1, indicating presence of the feature, and 0, indicating absence, yielded proportions. For vegetation/naturalness, the mean statistic returned the mean Normalized Difference Vegetation Index (NDVI), where higher values identify more chlorophyll-producing healthy green vegetation captured by the satellite imagery pixels. Except for the 50 km wildfire radius, all others had a 3 km radius. I accepted the original values for the coarser resolution nighttime lights and area-based, neighborhood-level socioeconomic index.

Table 3.2. Outdoor environmental factors mapped for association with adverse birth outcomes. The time (year), distance threshold (radius in meters), units, and source are indicated for each.

Category	Variable	Year	Feature	Method	Radius (km)	Units	Source
Built	136 air substances	average 2006-12	point	Kernel Density	10	tonnes or kg/km^2	EnvCan [105]
	Industrial facilities	unique 2006-12	point		3	#/km^2	
	Roads	2012	line			km/km^2	StatsCan [104]
	Electrical power lines	2012	line			km/km^2	AltaLIS [143]
	Gas stations	2015	point			#/km^2	DMTI Spatial [110]
	Waste/landfills	2015	point			#/km^2	
	Oil/gas well pads	2012	point			#/km^2	ABMI [144]
	High density livestock operations	2012	area	Focal Statistics		#/km^2	
	Mine sites	2012	area			km^2/km^2	
	Cultivated lands	2012	area			km^2/km^2	
	Nighttime lights	average 2006-12	raster	None	0	index	NOAA [145]
Social	Food stores	2015	point	Kernel Density	3	#/km^2	DMTI Spatial [110]
	Health care	2015	point			#/km^2	
	Hospitals	2015	point			#/km^2	
	Aboriginal lands	2016	area	Focal Statistics		km^2/km^2	NRCan [146]
	Neighborhood socioeconomic	2006	raster	None	0	index	Chan [103]
Natural	Vegetation/naturalness	maximum 2006-12	raster	Focal Statistics	3	index/km^2	NASA [132]
	Water	2013	area			km^2/km^2	NRCan [146]
	Wildfires	average 2006-12	area		50	km^2/km^2	AgFor [147]

Spearman's rank correlation

I joined values from the DKD distributions and each independent variable surface extracted to unique postal codes where births occurred. My data were non-normally distributed due to many zero values in both the dependent and independent variables. I used Python 2.7 software [111] with the pandas 0.16 site package [148] to calculate Spearman's rank correlations among SGA/LBWT and each environmental variable. To test the association of the combined environmental factors, I calculated a second set of Spearman's rank correlations using DKD values to test the indices. Correlation was calculated for the entire province and aggregated by sub-provincial unit.

Overlay analysis

Overlay analysis is a simple and reproducible method to combine several inputs into a single output [149]. It is most common for optimal site selection and suitability modeling, especially for mapping habitat. The class values represent rankings from higher to lower suitability or risk. In my study, I applied it to map "reverse suitability" to identify maternal ambient health hazards.

Because the values of continuous surfaces varied in measurement units, I standardized them into a similar ratio scale by reclassifying the environmental variables into 5 standard classes using quintiles. The ordering of the reclassification corresponded with the direction of the correlation: most were straightforward but if the variable was negatively correlated then the reclassification was applied in a backwards fashion; e.g. vegetation, water and socioeconomic status classes were ranked 5 to 1 because lower original values were considered to be more hazardous. I calculated the sub-indices as weighted sum overlays with equal weightings on air substances and land-based sources separately, which were then overlaid together. I was interested in preserving the combined effects of the industrial air substances; therefore, in addition to an equal weighted sum of both, I also approximated a conservative two-thirds (0.7) weighting to the air substances summed with a one-third (0.3) weighting of the land-based sources. In the two different indices – Overlay Equal and Overlay 0.7/0.3 – the class rankings were accumulations that represent where the study area had more environmental hazards.

Overall, the reverse suitability indices were calculated by modeling each individual pollutant surface using distance-centered analyses (i.e. kernel densities and focal statistics), reclassified in to quintiles of class rankings, and overlaid as weighted sums. The detailed GIS methods for the map-based calculations of all the independent variables and subsequent indices are specified in Table 3.2 (i.e. features, methods, and radii) and shown graphically in Figure 3.2.

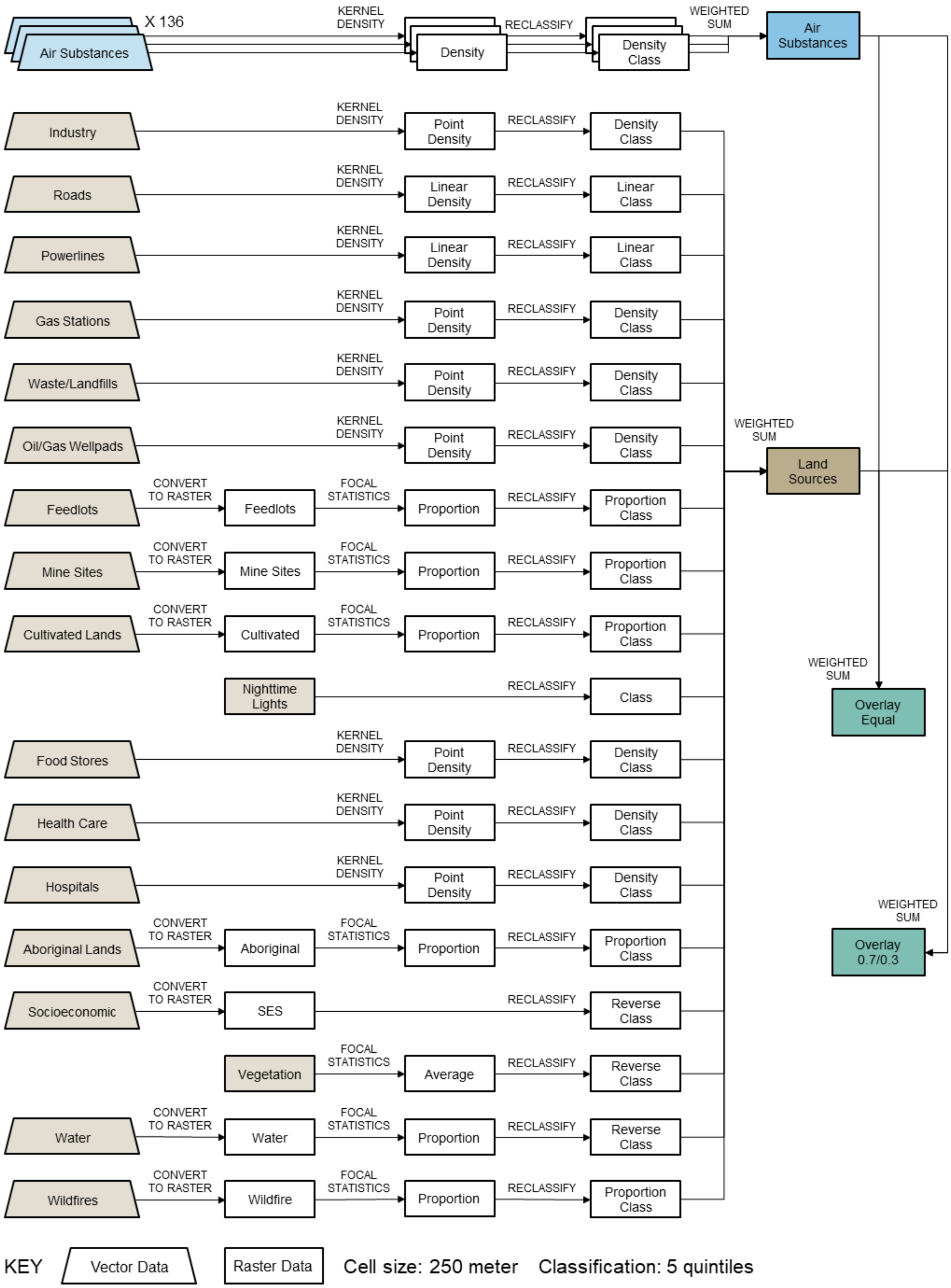


Figure 3.2. Flow chart of standard GIS commands for constructing the indices (colored boxes).

3.4 Results

Spatial distribution of adverse birth outcomes

Table 3.1 shows raw counts of births, SGA, and LBWT, based on valid postal codes. For 2006-2012, the entire province of Alberta had 333,247 births with a valid spatial location (95% of total registered), allocated to 53,399 postal codes. 29,679 geocoded births were classified as SGA (8.9%) and 5,485 were classified as LBWT at term (1.6%).

Figure 3.3 depicts the percentages of SGA/LBWT for each sub-provincial unit relative to Alberta (marked by *). For health regions, SGA ranged from 7.0-10.3% and LBWT ranged from 1.2-1.9%. Health region 4823 had the highest number of births (n=121,965), highest SGA (n=12,543, 10.3%), and highest LBWT (n=2,339, 1.9%); 4825 had the lowest number of births (n=9,737), SGA (n=698, 7.2%), and LBWT (n=140, 1.4%); but 4827 had the lowest SGA (n=1,252, 7.0%) and LBWT (n=222, 1.2%). For airshed zones, SGA ranged from 6.5-10.3% and LBWT ranged from1.0-1.9%. Airshed zone CRAZ had the highest number of births (n=120,392), SGA (n=12,409, 10.3%), and LBWT (n=2,310, 1.9%); FAP had the lowest number of births (n=3.547) and LICA had the lowest SGA (n=293, 6.5%) and lowest LBWT (n=43, 1.0%).

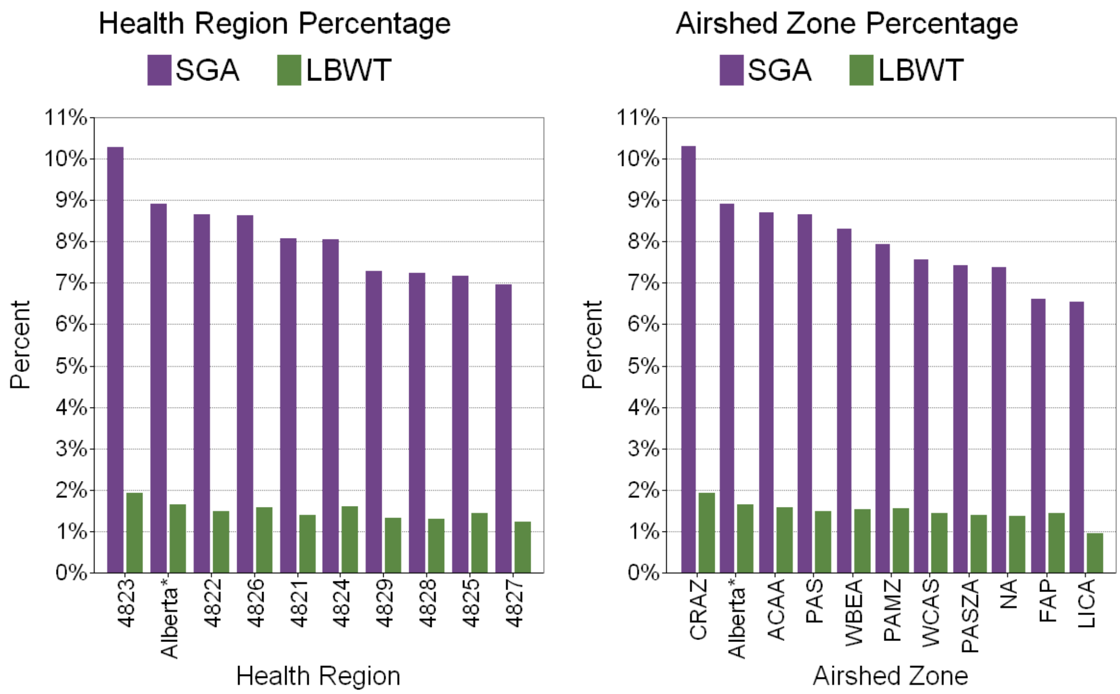


Figure 3.3. Percentages of births having small for gestational age (SGA) or low birth weight at term (LBWT) in sub-provincial units (* indicates value for the whole province).

The distributions of births per area in Figure 3.4 show higher concentrations of more than 3 births per km^2 in the sub-provincial units containing the major cities of Edmonton and Calgary, with medium densities in the adjacent units and in the airshed zones containing Grande Prairie (west-central) and Cold Lake (east-central).

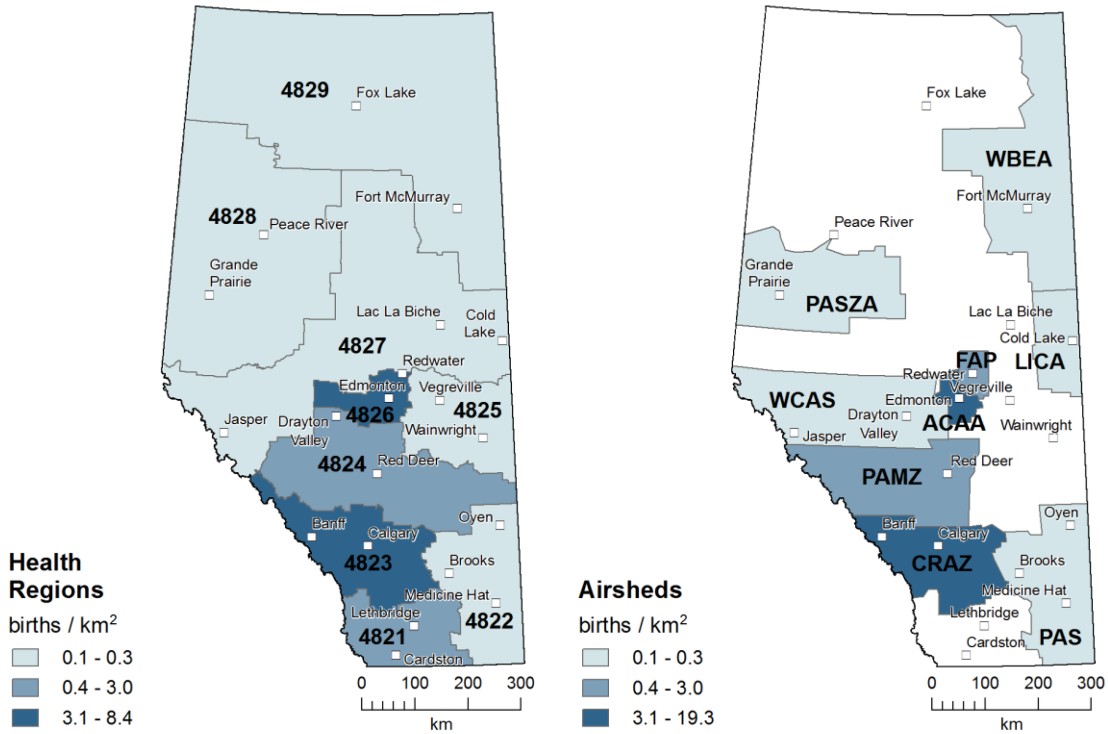


Figure 3.4. Births per area ratios in sub-provincial units.

The patterns differ by sub-provincial unit for SGA/LBWT mapped as numbers per births (Figure 3.5). SGA is highest in the units containing Edmonton and along the west-east Banff-Calgary-Brooks corridor. Health regions have medium SGA adjacent to the high SGA. Airsheds also show medium SGA in the west and north-east. LBWT is highest in the north-south Edmonton-Red Deer-Calgary corridor. Medium LBWT is adjacent to the higher units, with the exception of the northern health regions containing Grande Prairie-Peace River and Fort McMurray-Fox Lake. The lower LBWT in the central health region 4827 separates the province; LBWT in the airshed containing Cold Lake is the lowest in the province.

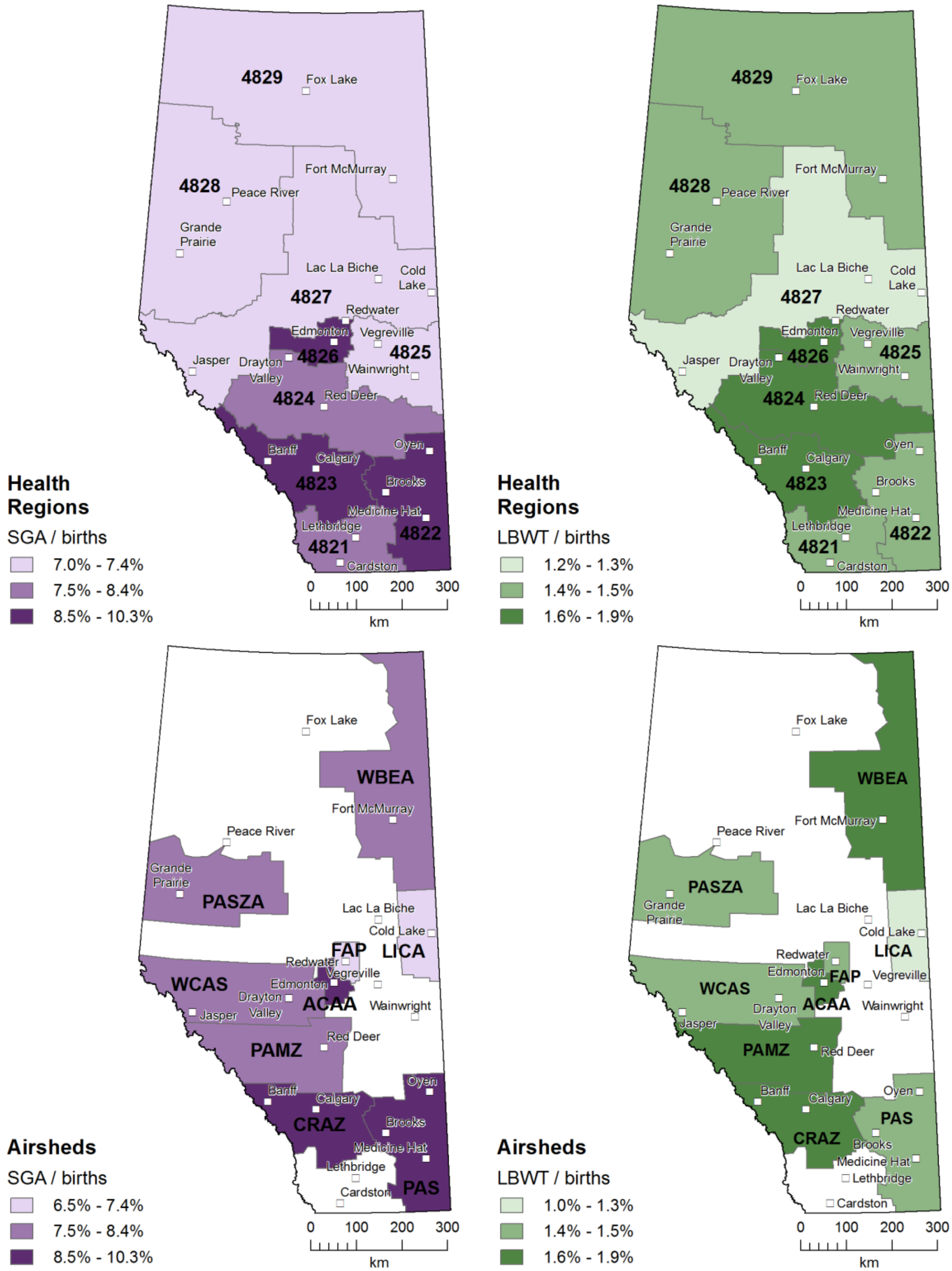


Figure 3.5. Small for gestational age (SGA) or low birth weight at term (LBWT) ratios in sub-provincial units.

Figure 3.6 maps the results of the DKD method for each SGA/LBWT. Both SGA/LBWT cover the same areas of the province and the darker colors indicate higher values for SGA (purple) and LBWT (green). The result of DKD is a continuous value, but the maps classified with tertiles visually enhance the slightly different distributions for SGA and LBWT: urban (Edmonton and Calgary) areas shared highest values for both SGA/LBWT; central areas had more LBWT; and southeast areas had more SGA.

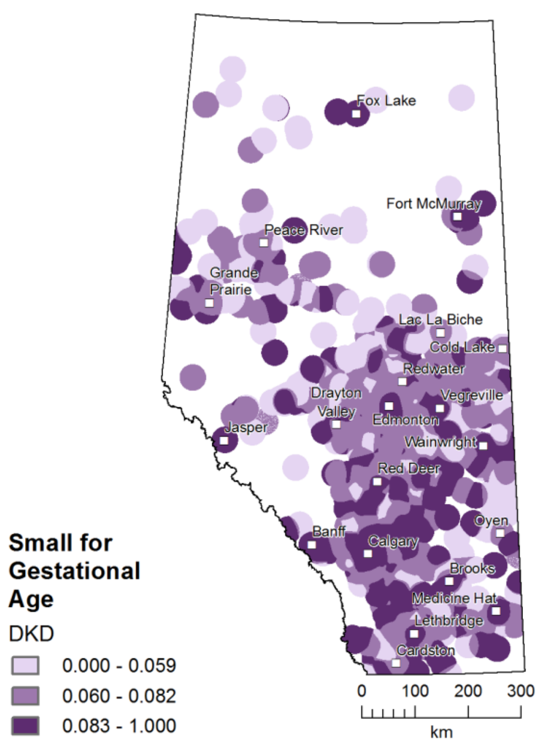


Figure 3.6. Double kernel density (DKD) distributions of small for gestational age (SGA) and low birth weight at term (LBWT) are ratios of the adverse birth outcome per area divided by total births per area, each within a 25 km radius; DKD is dimensionless.

Hazard mapping

The Spearman's rank correlation values were sorted in descending order for each of the independent variables (Table 3.3). Provincially, variables having correlations greater than 0.40 (low value accepted since data were not adjusted for epidemiological factors because they were not available for mapping) with SGA were: i-Butyl alcohol (*rho*=0.56); Asbestos; Nighttime Light; Toluenediisocyanate; Toluene-2,4-diisocyanate; Toluene-2,6-diisocyanate; Chromium Aluminum; Hydrogen sulphide; Road; 2-Ethoxyethanol; *Nickel; Quinoline; Aniline; Cyclohexane; Acetaldehyde; and *Phosphorus (*rho*=0.42). Variables with correlations greater than 0.40 with LBWT were: i-Butyl alcohol (*rho*=0.54); Asbestos; Toluenediisocyanate; Toluene-2,4-diisocyanate; Toluene-2,6-diisocyanate; Aluminum; Chromium; Nighttime Light; Hydrogen sulphide; 2-Ethoxyethanol; Quinoline; Aniline; Road; Cyclohexane; Acetaldehyde; *Isopropyl alcohol; and *Ethylene oxide (*rho*=0.41). Both SGA/LBWT were strongly associated with 15 air substances (the asterisk * marks those that differed: Nickel and Phosphorous for SGA; Ethylene oxide and Isopropyl alcohol for LBWT) and 2 land sources (both Nighttime Light and Road). Both SGA/LBWT had negative correlations (< -0.40) with Vegetation (SGA *rho*=-0.56; LBWT *rho*=-0.48), Oil/Gas Wellpad (SGA *rho*=-0.53; LBWT *rho*=-0.49), and Cultivated Land (SGA *rho*=-0.47; LBWT *rho*=-0.41).

The dilution effect of spreading the hazards across the large study area highlighted regional importance. Using the criteria of four or more health regions having a *rho* greater than 0.40 indicated the importance of Nitrogen oxides, Sulphur dioxide, Particulate Matter less than or equal to 2.5 microns ($PM_{2.5}$), and Acetaldehyde with SGA. The same criteria identified Xylene, Mine Site, Manganese, and Lead for LBWT. Four or more airshed zones having a *rho* greater than 0.40 highlighted Sulphur dioxide and Acetaldehyde with SGA, and Xylene, Particulate Matter less than or equal to 10 microns (PM_{10}), and $PM_{2.5}$ for LBWT.

The number of unique environmental variables having *rho* values greater than 0.40 province-wide or within four or more sub-provincial units totaled 30 (24 air substances and 6 land-based).

Table 3.3. Spearman's rank correlations of small for gestational age (SGA) and low birth weight at term (LBWT) with air substances and land sources (*), in descending correlation rho values. In the right half of the table, the count of units exceeding rho>0.40 and the range are shown for the data aggregated by health regions and airshed zones. Variables having a rho>0.4 for the province or for 4 or more sub-provincial units are indicated by bold font. Table continues.

	Province *rho*		**Health Region count (*rho* range)**		**Airshed Zone count (*rho* range)**	
Variable	**SGA**	**LBWT**	**SGA**	**LBWT**	**SGA**	**LBWT**
i-Butyl alcohol	**0.56**	**0.54**	1 (0.02 to 0.81)	3 (0.42 to 0.80)	1 (0.32 to 0.81)	2 (0.45 to 0.80)
Asbestos (friable form)	**0.54**	**0.52**	1 (0.73 to 0.73)	1 (0.67 to 0.67)	1 (0.73 to 0.73)	1 (0.67 to 0.67)
***Nighttime Light**	**0.51**	**0.47**	2 (-0.17 to 0.48)	1 (-0.50 to 0.42)	2 (-0.40 to 0.51)	3 (-0.50 to 0.52)
Toluenediisocyanate (mixed isomers)	**0.49**	**0.51**	0 (0.26 to 0.26)	1 (0.42 to 0.42)	0 (0.26 to 0.26)	1 (0.42 to 0.42)
Toluene-2,4-diisocyanate	**0.49**	**0.50**	0 (0.26 to 0.26)	1 (0.41 to 0.41)	0 (0.26 to 0.26)	1 (0.41 to 0.41)
Toluene-2,6-diisocyanate	**0.49**	**0.50**	0 (0.26 to 0.26)	1 (0.41 to 0.41)	0 (0.26 to 0.26)	1 (0.41 to 0.41)
Chromium (and its compounds)	**0.48**	**0.47**	2 (-0.07 to 0.68)	2 (-0.07 to 0.77)	2 (-0.23 to 0.68)	2 (-0.26 to 0.77)
Aluminum (fume or dust)	**0.48**	**0.49**	1 (0.20 to 0.67)	1 (0.25 to 0.76)	1 (-0.18 to 0.67)	1 (-0.20 to 0.76)
Hydrogen sulphide	**0.47**	**0.47**	**4 (-0.09 to 0.59)**	3 (-0.06 to 0.67)	3 (-0.34 to 0.59)	3 (-0.35 to 0.67)
***Road**	**0.46**	**0.42**	2 (-0.11 to 0.47)	0 (-0.37 to 0.38)	3 (-0.01 to 0.60)	2 (-0.37 to 0.51)
2-Ethoxyethanol	**0.44**	**0.46**	0 (0.25 to 0.25)	1 (0.41 to 0.41)	0 (0.25 to 0.25)	1 (0.41 to 0.41)
Nickel (and its compounds)	**0.44**	**0.39**	2 (-0.20 to 0.66)	3 (-0.88 to 0.75)	2 (-0.24 to 0.66)	2 (-0.88 to 0.75)
Quinoline (and its salts)	**0.43**	**0.46**	0 (0.25	1 (0.40	0 (0.25	1 (0.40

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
			to 0.25)	to 0.40)	to 0.25)	to 0.40)
Aniline (and its salts)	**0.43**	**0.45**	0 (0.25 to 0.25)	1 (0.40 to 0.40)	0 (0.25 to 0.25)	1 (0.40 to 0.40)
Cyclohexane	**0.42**	**0.42**	3 (-0.07 to 0.81)	3 (-0.07 to 0.81)	3 (-0.27 to 0.81)	3 (-0.30 to 0.81)
Acetaldehyde	**0.42**	**0.42**	**4 (-0.41 to 0.60)**	3 (-0.07 to 0.59)	**4 (-0.34 to 0.60)**	3 (-0.34 to 0.59)
Phosphorus (total)	**0.42**	0.38	1 (-0.20 to 0.49)	1 (-0.88 to 0.44)	1 (-0.49 to 0.52)	1 (-0.88 to 0.46)
Isopropyl alcohol	0.40	**0.41**	2 (-0.40 to 0.52)	3 (-0.47 to 0.53)	2 (-0.60 to 0.77)	2 (-0.67 to 0.52)
PAHs, total unspeciated	0.40	0.40	0 (0.30 to 0.32)	1 (0.33 to 0.43)	0 (0.26 to 0.32)	1 (0.29 to 0.43)
Ethylene oxide	0.36	**0.41**	0 (-0.43 to 0.25)	1 (-0.20 to 0.41)	0 (-0.30 to 0.25)	1 (-0.32 to 0.41)
Ammonia (total)	0.36	0.34	1 (-0.41 to 0.63)	1 (-0.75 to 0.73)	1 (-0.50 to 0.53)	1 (-0.75 to 0.66)
Phosphorus (yellow or white)	0.35	0.37	0 (-0.14 to 0.05)	0 (-0.14 to 0.23)	0 (-0.15 to 0.05)	0 (-0.14 to 0.23)
Methylenebis(phenylisocyanate)	0.34	0.38	0 (-0.43 to 0.23)	0 (-0.49 to 0.38)	0 (0.13 to 0.23)	1 (0.38 to 0.72)
PM_{10} - Particulate Matter ≤ 10 Microns	0.33	0.30	3 (-0.33 to 0.89)	3 (-0.83 to 0.62)	3 (-0.75 to 0.93)	**4 (-0.83 to 0.68)**
n-Butyl alcohol	0.31	0.32	1 (-0.71 to 0.79)	1 (-0.75 to 0.81)	1 (0.19 to 0.79)	1 (0.36 to 0.81)
Dichloromethane	0.31	0.31	1 (0.27 to 0.72)	2 (0.42 to 0.74)	1 (-0.22 to 0.71)	2 (-0.25 to 0.73)
Ethylene	0.30	0.33	3 (0.02 to 0.80)	3 (0.02 to 0.79)	3 (-0.32 to 0.80)	3 (-0.33 to 0.79)
Styrene	0.30	0.31	1 (0.00 to 0.81)	1 (-0.01 to 0.82)	1 (-0.32 to 0.83)	1 (-0.32 to 0.85)
Lead (and its compounds)	0.30	0.30	3 (-0.07 to 0.68)	**4 (-0.26 to 0.76)**	3 (-0.23 to 0.87)	3 (-0.65 to 0.76)
*Food Store	0.28	0.28	1 (-0.18 to 0.58)	1 (-0.23 to 0.57)	2 (-0.18 to 0.58)	1 (-0.17 to 0.57)
Cumene	0.27	0.27	1 (0.29 to 0.58)	2 (0.44 to 0.61)	1 (-0.14 to 0.57)	2 (-0.09 to 0.60)
Methyl isobutyl ketone	0.25	0.26	1 (0.07 to 0.69)	1 (0.25 to 0.73)	1 (0.07 to 0.67)	1 (0.08 to 0.71)
Xylene (mixed isomers)	0.24	0.26	1 (-0.71 to 0.54)	**5 (-0.74 to 0.59)**	2 (-0.65 to 0.74)	**5 (-0.47 to 0.87)**
Sulphur dioxide	0.24	0.20	**5 (-0.27 to 0.88)**	3 (-0.87 to 0.74)	**4 (-0.35 to 0.91)**	2 (-0.87 to 0.68)
Manganese (and its	0.24	0.21	3 (-0.03	**4 (-0.34**	3 (-0.50	3 (-0.71

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
compounds)			to 0.68)	**to 0.72)**	to 0.65)	to 0.70)
Fluorene - PAH	0.23	0.13	2 (-0.32 to 0.50)	2 (-0.89 to 0.55)	2 (-0.49 to 0.52)	2 (-0.89 to 0.53)
2-Butoxyethanol	0.22	0.23	1 (-0.69 to 0.66)	1 (-0.70 to 0.66)	1 (0.11 to 0.64)	1 (0.11 to 0.64)
*Gas Station	0.20	0.20	1 (-0.23 to 0.58)	0 (-0.25 to 0.30)	2 (-0.41 to 0.56)	1 (-0.42 to 0.58)
Naphthalene	0.20	0.14	1 (-0.21 to 0.59)	2 (-0.89 to 0.63)	1 (-0.40 to 0.61)	2 (-0.89 to 0.65)
Propylene	0.18	0.19	1 (-0.11 to 0.68)	1 (-0.01 to 0.71)	1 (-0.30 to 0.70)	1 (-0.31 to 0.73)
*Health Care	0.17	0.17	2 (-0.21 to 0.51)	0 (-0.36 to 0.24)	2 (-0.22 to 0.53)	0 (-0.49 to 0.28)
Volatile Organic Compounds (VOCs)	0.17	0.15	3 (-0.45 to 0.88)	2 (-0.74 to 0.67)	2 (-0.45 to 0.92)	1 (-0.74 to 0.69)
Toluene	0.17	0.19	1 (-0.05 to 0.68)	2 (-0.35 to 0.72)	2 (-0.66 to 0.74)	3 (-0.46 to 0.89)
Formic acid	0.14	0.15	0 (-0.50 to 0.39)	1 (-0.50 to 0.52)	0 (-0.49 to 0.39)	1 (-0.50 to 0.52)
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	0.14	0.11	**4 (-0.57 to 0.88)**	3 (-0.73 to 0.59)	3 (-0.57 to 0.91)	**4 (-0.73 to 0.61)**
1,2,4-Trimethylbenzene	0.13	0.18	1 (-0.05 to 0.57)	2 (-0.07 to 0.60)	2 (-0.38 to 0.67)	1 (-0.47 to 0.60)
Formaldehyde	0.12	0.10	2 (-0.41 to 0.70)	3 (-0.81 to 0.71)	3 (-0.29 to 0.83)	3 (-0.81 to 0.70)
Vanadium (except when in an alloy) and its compounds	0.11	0.12	1 (-0.05 to 0.49)	1 (0.05 to 0.54)	1 (-0.21 to 0.47)	1 (-0.16 to 0.52)
Carbon disulphide	0.10	0.10	1 (-0.21 to 0.41)	1 (-0.28 to 0.52)	1 (-0.26 to 0.75)	0 (-0.32 to 0.30)
Benzo(g,h,i)perylene - PAH	0.10	0.05	2 (-0.27 to 0.60)	2 (-0.89 to 0.65)	2 (-0.49 to 0.59)	2 (-0.89 to 0.64)
Indeno(1,2,3-c,d)pyrene - PAH	0.09	0.05	2 (-0.27 to 0.61)	2 (-0.89 to 0.65)	2 (-0.49 to 0.59)	2 (-0.89 to 0.64)
Pyrene - PAH	0.09	0.03	2 (-0.27 to 0.62)	2 (-0.89 to 0.66)	2 (-0.49 to 0.61)	2 (-0.89 to 0.65)
Perylene - PAH	0.09	0.04	2 (-0.27 to 0.60)	2 (-0.89 to 0.64)	2 (-0.49 to 0.59)	2 (-0.89 to 0.64)
Benzo(a)phenanthrene - PAH	0.09	0.04	2 (-0.27 to 0.61)	2 (-0.89 to 0.64)	2 (-0.49 to 0.60)	2 (-0.89 to 0.64)
Benzo(e)pyrene - PAH	0.08	0.04	2 (-0.27 to 0.61)	2 (-0.89 to 0.65)	2 (-0.27 to 0.59)	2 (-0.89 to 0.64)
Benzo(a)anthracene - PAH	0.08	0.04	2 (-0.27 to 0.61)	2 (-0.89 to 0.65)	2 (-0.27 to 0.59)	2 (-0.89 to 0.64)
Fluoranthene - PAH	0.08	0.02	2 (-0.27	2 (-0.89	2 (-0.27	2 (-0.89

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
			to 0.61)	to 0.65)	to 0.60)	to 0.64)
Methyl ethyl ketone	0.08	0.08	1 (-0.21 to 0.80)	1 (-0.31 to 0.82)	1 (-0.32 to 0.79)	1 (-0.32 to 0.81)
Benzene	0.07	0.09	0 (-0.12 to 0.29)	1 (-0.08 to 0.54)	1 (-0.66 to 0.70)	2 (-0.31 to 0.88)
Benzo(k)fluoranthene - PAH	0.07	0.02	2 (-0.27 to 0.60)	2 (-0.89 to 0.65)	2 (-0.49 to 0.59)	2 (-0.89 to 0.64)
*Aboriginal Land	0.07	0.05	0 (-0.35 to 0.00)	0 (-0.29 to 0.03)	0 (-0.64 to -0.04)	0 (-0.29 to 0.23)
Diethanolamine (and its salts)	0.07	0.06	1 (0.00 to 0.41)	1 (-0.01 to 0.46)	1 (-0.25 to 0.44)	1 (-0.25 to 0.49)
Benzo(a)pyrene - PAH	0.06	0.02	2 (-0.27 to 0.57)	2 (-0.89 to 0.61)	2 (-0.49 to 0.56)	2 (-0.89 to 0.60)
Aluminum oxide (fibrous forms)	0.06	0.05	1 (0.81 to 0.81)	1 (0.80 to 0.80)	1 (0.81 to 0.81)	1 (0.81 to 0.81)
Benzo(j)fluoranthene - PAH	0.06	0.01	2 (-0.27 to 0.57)	2 (-0.89 to 0.61)	2 (-0.49 to 0.56)	2 (-0.89 to 0.60)
Benzo(b)fluoranthene - PAH	0.06	0.01	2 (-0.27 to 0.57)	2 (-0.89 to 0.61)	2 (-0.49 to 0.56)	2 (-0.89 to 0.60)
n-Hexane	0.04	0.05	2 (-0.09 to 0.63)	2 (-0.31 to 0.65)	2 (-0.36 to 0.53)	2 (-0.47 to 0.85)
Calcium fluoride	0.03	0.02	1 (0.67 to 0.67)	1 (0.71 to 0.71)	1 (0.08 to 0.67)	1 (0.08 to 0.70)
Carbonyl sulphide	0.03	0.03	1 (0.04 to 0.41)	2 (-0.28 to 0.52)	2 (-0.26 to 0.75)	1 (-0.31 to 0.48)
***Mine Site**	0.02	0.00	2 (-0.35 to 0.43)	**4 (-0.41 to 0.50)**	1 (-0.21 to 0.53)	2 (-0.60 to 0.57)
Biphenyl	0.01	0.00	1 (0.60 to 0.60)	1 (0.64 to 0.64)	1 (-0.16 to 0.59)	1 (-0.11 to 0.63)
*Waste / Landfill	0.01	0.05	0 (-0.29 to 0.30)	1 (-0.31 to 0.42)	1 (-0.29 to 0.80)	1 (-0.27 to 0.42)
Ethylene glycol	0.00	0.00	1 (-0.69 to 0.41)	2 (-0.85 to 0.53)	1 (-0.44 to 0.75)	1 (-0.85 to 0.49)
Hydrogen fluoride	0.00	0.00	1 (0.04 to 0.55)	1 (-0.05 to 0.59)	1 (-0.09 to 0.54)	1 (-0.01 to 0.58)
Methyl tert-butyl ether	0.00	-0.01	1 (0.77 to 0.77)	1 (0.78 to 0.78)	1 (0.15 to 0.76)	1 (0.15 to 0.77)
n,n-Dimethylformamide	0.00	-0.02	1 (0.72 to 0.72)	1 (0.74 to 0.74)	1 (0.15 to 0.71)	1 (0.15 to 0.73)
Vinyl acetate	0.00	-0.01	1 (0.56 to 0.56)	1 (0.58 to 0.58)	1 (0.54 to 0.54)	1 (0.57 to 0.57)
N-Methyl-2-pyrrolidone	0.00	-0.01	1 (0.53 to 0.53)	1 (0.57 to 0.57)	1 (0.15 to 0.52)	1 (0.15 to 0.56)
Isoprene	0.00	0.00	0 (0.00	0 (-0.01	0 (-0.01	0 (-0.01

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
			to 0.00)	to -0.01)	to -0.01)	to -0.01)
Titanium tetrachloride	0.00	0.00	0 (0.00 to 0.00)	0 (-0.01 to -0.01)	0 (-0.01 to -0.01)	0 (-0.01 to -0.01)
Methanol	-0.01	-0.01	1 (-0.39 to 0.52)	1 (-0.79 to 0.48)	0 (-0.39 to 0.30)	1 (-0.79 to 0.48)
Cresol (all isomers and their salts)	-0.01	-0.02	1 (0.00 to 0.55)	1 (-0.32 to 0.59)	1 (-0.49 to 0.54)	1 (-0.72 to 0.58)
Carbon monoxide	-0.01	-0.04	2 (-0.41 to 0.89)	1 (-0.82 to 0.60)	2 (-0.41 to 0.92)	2 (-0.82 to 0.48)
Trichloroethylene	-0.01	-0.01	1 (0.49 to 0.49)	1 (0.53 to 0.53)	1 (-0.24 to 0.51)	1 (-0.27 to 0.56)
p-Phenylenediamine (and its salts)	-0.02	-0.01	0 (-0.03 to -0.03)	0 (-0.03 to -0.03)	0 (-0.18 to -0.18)	0 (-0.13 to -0.13)
Acrolein	-0.02	-0.10	1 (0.06 to 0.70)	1 (0.00 to 0.67)	1 (0.05 to 0.75)	0 (0.05 to 0.20)
Hexavalent chromium (and its compounds)	-0.02	0.02	0 (-0.07 to 0.31)	0 (-0.32 to 0.17)	0 (-0.50 to 0.08)	0 (-0.71 to 0.25)
Dibenzo(a,i)pyrene - PAH	-0.02	-0.09	1 (-0.21 to 0.54)	1 (-0.89 to 0.59)	1 (-0.21 to 0.53)	1 (-0.89 to 0.58)
7H-Dibenzo(c,g)carbazole - PAH	-0.02	-0.11	1 (-0.21 to 0.55)	1 (-0.89 to 0.59)	1 (-0.21 to 0.54)	1 (-0.89 to 0.58)
tert-Butyl alcohol	-0.02	-0.03	0 (0.31 to 0.31)	0 (0.35 to 0.35)	0 (0.17 to 0.29)	0 (0.17 to 0.34)
Molybdenum trioxide	-0.03	-0.11	1 (-0.20 to 0.44)	1 (-0.88 to 0.49)	1 (-0.20 to 0.43)	1 (-0.88 to 0.48)
Acenaphthene - PAH	-0.03	-0.11	2 (-0.21 to 0.63)	2 (-0.89 to 0.66)	2 (-0.49 to 0.61)	2 (-0.89 to 0.66)
Chlorine	-0.04	-0.02	2 (-0.55 to 0.54)	3 (-0.32 to 0.57)	2 (-0.55 to 0.56)	3 (-0.72 to 0.60)
Phenanthrene - PAH	-0.05	-0.11	1 (-0.27 to 0.49)	2 (-0.89 to 0.44)	1 (-0.49 to 0.52)	1 (-0.89 to 0.46)
*Industrial Facility	-0.05	-0.02	1 (-0.18 to 0.43)	0 (-0.46 to 0.29)	1 (-0.62 to 0.47)	1 (-0.59 to 0.40)
*Hospital	-0.05	-0.03	1 (-0.19 to 0.44)	0 (-0.30 to 0.19)	2 (-0.33 to 0.70)	0 (-0.47 to 0.19)
5-Methylchrysene - PAH	-0.06	-0.22	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)
1-Nitropyrene - PAH	-0.06	-0.22	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)
Dibenzo(a,e)fluoranthene - PAH	-0.06	-0.22	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)
Dibenzo(a,h)pyrene - PAH	-0.06	-0.22	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)	0 (-0.21 to -0.21)	0 (-0.89 to -0.89)
Dibenzo(a,l)pyrene - PAH	-0.06	-0.22	0 (-0.21	0 (-0.89	0 (-0.21	0 (-0.89

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
			to -0.21)	to -0.89)	to -0.21)	to -0.89)
Dibenz(a,h)acridine - PAH	-0.06	-0.22	0 (-0.21 to -0.02)	0 (-0.89 to -0.01)	0 (-0.21 to -0.02)	0 (-0.89 to -0.01)
Dibenzo(a,e)pyrene - PAH	-0.06	-0.22	0 (-0.21 to -0.02)	0 (-0.89 to -0.01)	0 (-0.21 to -0.02)	0 (-0.89 to -0.01)
Anthracene	-0.06	-0.22	0 (-0.21 to -0.03)	0 (-0.89 to -0.03)	0 (-0.21 to -0.19)	0 (-0.89 to -0.14)
Dibenz(a,j)acridine - PAH	-0.06	-0.13	1 (-0.21 to 0.55)	1 (-0.89 to 0.59)	1 (-0.21 to 0.54)	1 (-0.89 to 0.58)
Sulphuric acid	-0.06	-0.10	2 (-0.21 to 0.52)	2 (-0.89 to 0.56)	2 (-0.50 to 0.54)	2 (-0.89 to 0.58)
Ethylbenzene	-0.07	-0.05	2 (-0.71 to 0.46)	3 (-0.74 to 0.55)	2 (-0.37 to 0.69)	1 (-0.47 to 0.49)
Dibenzo(a,h)anthracene - PAH	-0.07	-0.14	2 (-0.21 to 0.55)	2 (-0.89 to 0.59)	2 (-0.49 to 0.54)	2 (-0.89 to 0.58)
Hydrochloric acid	-0.08	-0.07	1 (-0.01 to 0.49)	2 (-0.32 to 0.61)	0 (-0.50 to 0.39)	2 (-0.71 to 0.49)
*High Density Livestock Operation	-0.08	-0.08	0 (-0.41 to 0.14)	0 (-0.39 to 0.13)	0 (-0.41 to 0.14)	0 (-0.18 to 0.14)
Polymeric diphenylmethane diisocyanate	-0.09	-0.10	1 (-0.15 to 0.41)	1 (-0.15 to 0.53)	1 (-0.15 to 0.75)	0 (-0.15 to 0.20)
7,12-Dimethylbenz(a)anthracene - PAH	-0.10	-0.25	1 (-0.21 to 0.48)	1 (-0.89 to 0.43)	1 (-0.49 to 0.51)	1 (-0.89 to 0.45)
3-Methylcholanthrene - PAH	-0.10	-0.25	1 (-0.21 to 0.49)	1 (-0.89 to 0.44)	1 (-0.49 to 0.52)	1 (-0.89 to 0.46)
*Power Line	-0.10	-0.04	1 (-0.84 to 0.41)	3 (-0.31 to 0.67)	0 (-0.86 to 0.26)	2 (-0.31 to 0.53)
1,1,2-Trichloroethane	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
HCFC-142b	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
1,1,2,2-Tetrachloroethane	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
Carbon tetrachloride	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
Pentachloroethane	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
Dicyclopentadiene	-0.10	-0.08	0 (-0.20 to 0.00)	0 (-0.20 to -0.01)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
1,3-Butadiene	-0.10	-0.08	0 (-0.20 to 0.00)	0 (-0.20 to -0.01)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
Chloroethane	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.28 to -0.01)	0 (-0.31 to -0.01)

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
Chloroform	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.31 to -0.01)	0 (-0.31 to -0.01)
Vinyl chloride	-0.10	-0.08	0 (-0.20 to -0.20)	0 (-0.20 to -0.20)	0 (-0.32 to -0.01)	0 (-0.34 to -0.01)
Zinc (and its compounds)	-0.11	-0.17	2 (-0.19 to 0.54)	3 (-0.86 to 0.58)	2 (-0.50 to 0.54)	2 (-0.86 to 0.59)
Arsenic (and its compounds)	-0.11	-0.13	1 (-0.18 to 0.55)	1 (-0.32 to 0.59)	1 (-0.50 to 0.52)	1 (-0.71 to 0.57)
Tetrachloroethylene	-0.12	-0.12	1 (0.42 to 0.42)	1 (0.47 to 0.47)	1 (-0.24 to 0.41)	1 (-0.27 to 0.46)
Dioxins and furans - total	-0.12	-0.12	2 (-0.10 to 0.66)	1 (-0.32 to 0.70)	2 (-0.50 to 0.67)	1 (-0.71 to 0.71)
Nitrogen oxides (expressed as NO_2)	-0.13	-0.19	**6 (-0.26 to 0.90)**	2 (-0.83 to 0.61)	3 (-0.27 to 0.91)	1 (-0.83 to 0.48)
Chlorine dioxide	-0.13	-0.15	1 (-0.15 to 0.49)	1 (-0.32 to 0.44)	1 (-0.49 to 0.52)	1 (-0.72 to 0.46)
Hexachlorobenzene	-0.13	-0.14	2 (-0.03 to 0.49)	2 (-0.32 to 0.48)	2 (-0.50 to 0.52)	2 (-0.71 to 0.52)
*Socioeconomic Index	-0.14	-0.14	0 (-0.59 to 0.21)	0 (-0.58 to 0.17)	0 (-0.59 to 0.26)	0 (-0.58 to 0.24)
1,2-Dichloroethane	-0.15	-0.08	0 (-0.42 to -0.20)	0 (-0.20 to 0.12)	0 (-0.26 to -0.01)	0 (-0.29 to -0.01)
Acetonitrile	-0.15	-0.15	0 (0.13 to 0.13)	0 (0.18 to 0.18)	0 (0.08 to 0.15)	0 (0.14 to 0.15)
1,4-Dioxane	-0.15	-0.08	0 (-0.42 to -0.20)	0 (-0.20 to 0.12)	0 (-0.30 to -0.30)	0 (-0.32 to -0.32)
HCFC-22	-0.15	-0.08	0 (-0.42 to -0.20)	0 (-0.20 to 0.12)	0 (-0.30 to -0.01)	0 (-0.32 to -0.01)
*Water Body	-0.17	-0.12	1 (-0.62 to 0.49)	0 (-0.55 to 0.14)	1 (-0.70 to 0.49)	0 (-0.63 to 0.22)
Phenol (and its salts)	-0.17	-0.18	1 (-0.13 to 0.41)	1 (-0.10 to 0.53)	1 (-0.18 to 0.75)	0 (-0.19 to 0.20)
Triethylamine	-0.17	-0.18	0 (0.06 to 0.06)	0 (0.11 to 0.11)	0 (0.02 to 0.15)	0 (0.07 to 0.15)
Acenaphthylene - PAH	-0.18	-0.26	1 (-0.21 to 0.49)	1 (-0.89 to 0.44)	1 (-0.49 to 0.52)	1 (-0.89 to 0.46)
Nitrilotriacetic acid (and its salts)	-0.19	-0.19	0 (-0.35 to -0.35)	0 (-0.35 to -0.35)	0 (-0.40 to -0.40)	0 (-0.40 to -0.40)
Mercury (and its compounds)	-0.19	-0.21	0 (-0.18 to 0.37)	1 (-0.33 to 0.42)	0 (-0.50 to 0.23)	0 (-0.71 to 0.28)
Nitrate ion in solution at pH >= 6.0	-0.19	-0.19	0 (-0.33 to -0.33)	0 (-0.31 to -0.31)	0 (-0.38 to -0.38)	0 (-0.37 to -0.37)
Nitric acid	-0.19	-0.19	0 (-0.34 to -0.34)	0 (-0.34 to -0.34)	0 (-0.40 to -0.40)	0 (-0.39 to -0.39)

Variable	Province *rho*		Health Region count (*rho* range)		Airshed Zone count (*rho* range)	
Selenium (and its compounds)	-0.19	-0.21	1 (-0.06 to 0.56)	1 (-0.43 to 0.61)	1 (-0.16 to 0.53)	1 (-0.67 to 0.58)
Silver (and its compounds)	-0.20	-0.20	0 (-0.36 to -0.36)	0 (-0.35 to -0.35)	0 (-0.42 to -0.42)	0 (-0.41 to -0.41)
Antimony (and its compounds)	-0.20	-0.20	0 (-0.37 to -0.37)	0 (-0.36 to -0.36)	0 (-0.42 to -0.18)	0 (-0.41 to -0.13)
Cadmium (and its compounds)	-0.20	-0.21	3 (-0.09 to 0.53)	2 (-0.03 to 0.60)	1 (-0.34 to 0.44)	2 (-0.66 to 0.49)
Copper (and its compounds)	-0.20	-0.23	1 (-0.18 to 0.62)	2 (-0.34 to 0.65)	1 (-0.21 to 0.61)	1 (-0.35 to 0.64)
*Wildfire	-0.24	-0.28	1 (-0.35 to 0.57)	0 (-0.64 to 0.39)	3 (-0.47 to 0.57)	1 (-0.71 to 0.75)
Cobalt (and its compounds)	-0.30	-0.39	0 (-0.43 to -0.03)	1 (-0.88 to 0.46)	0 (-0.42 to -0.11)	0 (-0.88 to 0.00)
***Cultivated Land**	**-0.47**	**-0.41**	0 (-0.33 to 0.17)	1 (-0.34 to 0.49)	1 (-0.61 to 0.81)	2 (-0.62 to 0.54)
***Oil/Gas Wellpad**	**-0.53**	**-0.49**	0 (-0.45 to 0.31)	0 (-0.34 to 0.33)	2 (-0.81 to 0.80)	2 (-0.74 to 0.79)
***Vegetation**	**-0.56**	**-0.48**	2 (-0.50 to 0.80)	3 (-0.52 to 0.48)	1 (-0.48 to 0.83)	3 (-0.52 to 0.58)

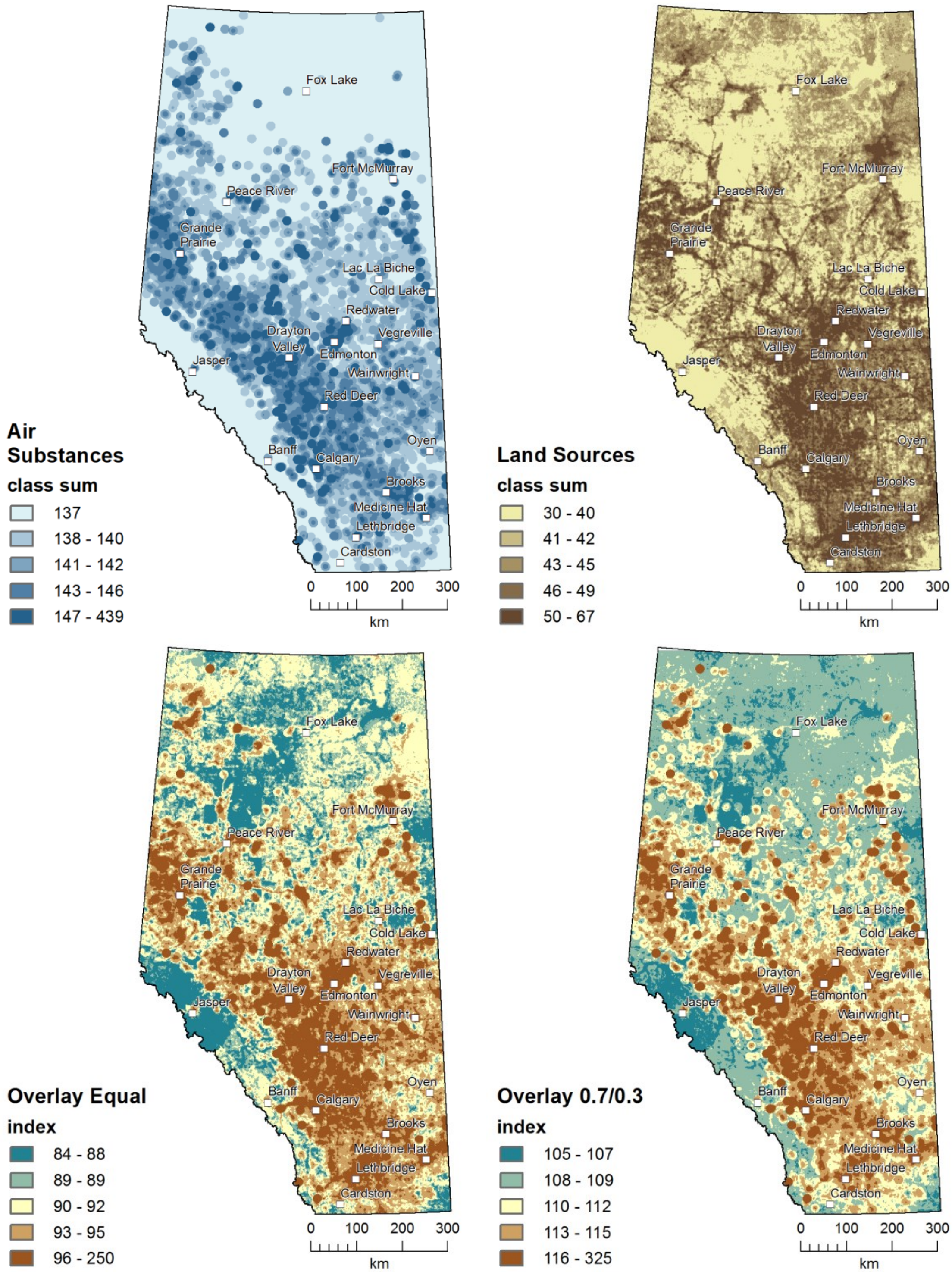


Figure 3.7. Weighted sum overlays for air substances and land-based sources were combined as equal and 0.7/0.3 weighted indices to identify the most hazardous locations.

Spatial distribution of the indices

Figure 3.7 maps the results from the weighted overlay sum of the five class rankings for 136 emitted Air Substances, 18 Land Sources, the Overlay Equal weighting of both, and the Overlay 0.7/0.3 weighting of air substances and land. The distribution of the higher rankings spatially coincides with Alberta's populated places, with the exception of higher values along the foothills, the Fort McMurray oil sands area in the north, and some scattered areas in the northeast. Quantile class breaks were used to visualize the contrast of higher to lower areas.

Associations with the hazards and indices

The actual index values were used for the correlations with SGA/LBWT DKD (Table 3.4). The correlations of the overlay indices with SGA/LBWT were very low for the entire province. The Air Substances were highest for both SGA (*rho*=0.21) and LBWT (*rho*=0.16). Land Factor correlations were slightly negative for SGA (*rho*=-0.26) and LBWT (*rho*=-0.23). Both overlay indices were lower than the Air Substances for SGA: Overlay Equal had a *rho*=0.18 and Overlay 0.7/0.3 had a *rho*=0.15. Overlay Equal was lower for LBWT (*rho*=0.13) but Overlay 0.7/0.3 was higher (*rho*=0.20).

Table 3.4. Spearman's rank correlations of small for gestational age (SGA) and low birth weight at term (LBWT) with air substances, land sources, and weighted sum overlay indices for the entire province of Alberta.

Index Name	**Inputs**	**SGA *rho***	**LBWT *rho***
Air Substances	sum of 136 variables classified to 5 quantiles	0.21	0.16
Land Sources	sum of 18 variables classified to 5 quantiles	-0.26	-0.23
Overlay Equal	air substances + land sources	0.18	0.13
Overlay 0.7/0.3	0.7 * air substances + 0.3 * land sources	0.15	0.20

Figure 3.8 displays index correlations with SGA/LBWT, by health region and airshed zone. In the graph symbols, longer bars mean greater association and bar direction designates positive (up) or negative (down). The air substances and land-based sources were included to demonstrate how much of an effect each had on the indices. The following indices had correlations greater than 0.40 with an SGA/LBWT:

- Air Substances with SGA in four health regions – 4829 (*rho*=0.85), 4828 (*rho*=0.67), 4826 (*rho* =0.55), 4823 (*rho*=0.42); and with LBWT in three health regions – 4828 (*rho*=0.73), 4826 (*rho*=0.59), and 4823 (*rho*=0.56).
- Air Substances with SGA in four airshed zones – WBEA (*rho*=0.89), PASZA (*rho*=0.57), ACAA (*rho*=0.55) and CRAZ (*rho*=0.42); and with LBWT in four airshed zones – LICA (*rho*=0.85), PASZA (*rho*=0.66), ACAA (*rho*=0.60), and CRAZ (*rho*=0.56).
- Land Sources were weakly associated with both SGA and LBWT in the majority of health regions and airshed zones.
- Overlay Equal index with SGA in four health regions – 4828 (*rho*=0.58), 4826 (*rho*=0.54), 4823 (*rho*=0.42), and 4827 (*rho*=0.42); and with LBWT in four health regions – 4828 (*rho*=0.63), 4826 (*rho*=0.59), 4821 (*rho*=0.57), and 4823 (*rho*=0.57).
- Overlay Equal index with SGA in three airshed zones – ACAA (*rho*=0.55), PASZA (*rho*=0.45), and CRAZ (*rho*=0.42); and with LBWT in three airshed zones – ACAA (*rho*=0.60), CRAZ (*rho*=0.57), and PASZA (*rho*=0.51).

- Overlay 0.7/0.3 index with SGA in four health regions – 4829 (*rho*=0.75), 4828 (*rho*=0.62), 4826 (*rho*=0.55), and 4823 (*rho*=0.42); and with LBWT in four health regions – 4828 (*rho*=0.68), 4826 (*rho*=0.59), 4823 (*rho*=0.57), and 4821 (*rho*=0.51).
- Overlay 0.7/0.3 index with SGA in four airshed zones – WBEA (*rho*=0.78), ACAA (*rho*=0.55), PASZA (*rho*=0.50), and CRAZ (*rho*=0.42); and with LBWT in four airshed zones – LICA (*rho*=0.60), ACAA (*rho*=0.60), PASZA (*rho*=0.59), and CRAZ (*rho*=0.57).

The health regions having the least association with SGA were 4821, 4822, 4824, and 4825; with LBWT these were 4822, 4824, 4827, and 4829. The airshed zones having the least association with SGA were FAP, PAMZ, PAS, and WCAS; with LBWT these were FAP, PAMZ, PAS, and WBEA.

SGA and LBWT were negatively correlated with *all* indices in health region 4822 and two airshed zones (PAS, FAP). The negative association also occurred in health region 4825 and airshed zone WCAS, but a higher positive correlation occurred with the Land Sources.

Using the criteria of correlations higher than 0.40, the Overlay 0.7/0.3 index had the highest overall count of sub-provincial units – both SGA/LBWT represented by at least 4 health regions and 4 airshed zones.

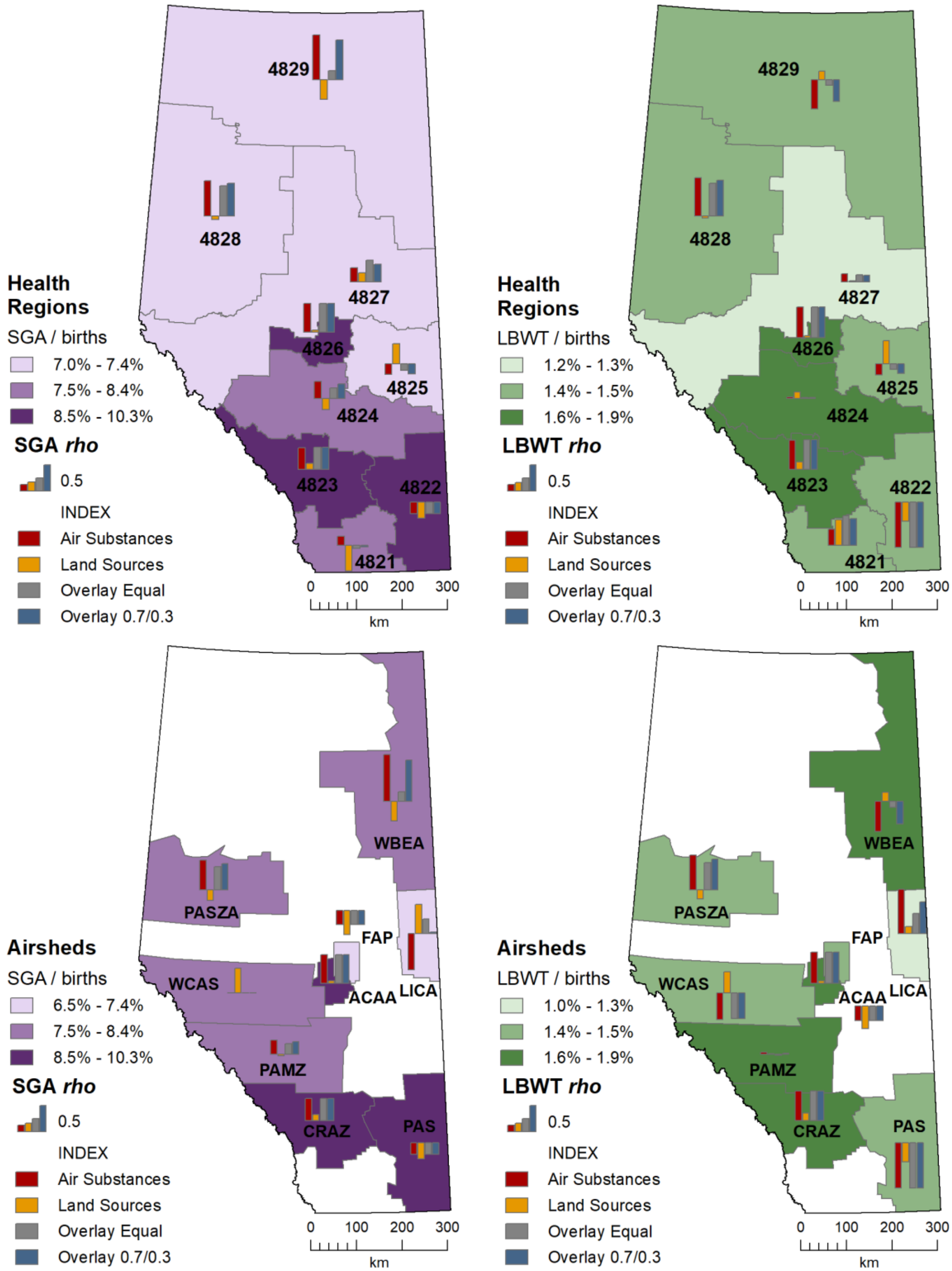


Figure 3.8. Spearman's rank correlations of each adverse birth outcome with indices – shown as bar charts in each health region or airshed. Background maps show ratios.

3.5 Discussion

Individual hazards

Of 136 NPRI substances reported in Alberta, 24 air-emitted substances had moderate correlations with one or both SGA/LBWT DKD ratios. Of these, 2-Ethoxyethanol and Lead are recognized developmental toxicants [150, 151]. Acetaldehyde, Aluminum, Ethylene oxide, Isopropyl alcohol, Nickel, Nitrogen oxides, PM_{10}, $PM_{2.5}$, Sulphur dioxide, Xylene, Chromium, Hydrogen sulphide, Manganese, Phosphorus, and Quinoline are suspected developmental toxicants, with more than half of the air substances associated with decreased fetal/offspring weight in animal studies [150, 151]. The following air substances are neither recognized or suspected as no studies were reported: Aniline, Asbestos, Cyclohexane, i-Butyl alcohol, Toluene-2,4-diisocyanate, Toluene-2,6-diisocyanate, and Toluenediisocyanate (note: the latter three have been combined in later versions of the NPRI database [105]).

Of the 18 land sources mapped, six had moderate correlations with one or both SGA/LBWT. Provincially, Cultivated Land was negatively associated with SGA and LBWT (likely because residences were not inside agricultural fields), but some regions were positive, similar to the Almberg et al. [152] study on proximity to pesticide-treated agricultural fields. Proximity to Mine Sites were associated for 2 to 3 health regions or airsheds; a related study found positive association for a single mine site indicating this is likely a more localized factor [72]. Nighttime Lights have not been explored with SGA/LBWT; however, breast cancer, which has other similar exposures, has a positive association [138, 153]. The smaller area airsheds showed high correlations of SGA/LBWT with Oil/Gas Wellpads, but was negative for the entire province and by health regions; mixed associations were also reported by Mckenzie et al. [68] and Casey et al. [67]. The moderate to higher correlations of Roads match much published research on the effect of maternal proximity to roads [80, 154]. Green or natural Vegetation was negatively correlated at the provincial level, but very mixed within health regions and airsheds; the sub-provincial dissimilarity with other studies [74, 155] was likely affected by the radii, resolution of the satellite sources, and the widely varying ecoregions in the province.

Ambient hazard indices

Both indices identified where there was an accumulation of hazards and therefore directly addressed the hypothesis that there were more small newborns where there were more outdoor hazards during the mothers' pregnancies. Since I was interested in preserving combined effects that the industrial air substances contributed to the outdoor environment, I weighted the sum of those more highly than the sum of all the land-based sources. Province-wide, the Overlay Equal index better identified SGA and the Overlay 0.7/0.3 better identified LBWT.

Differences in index associations were likely due to the spatial distributions (i.e. DKD) of SGA/LBWT. Both showed that hot spots did not occur strictly within the large urban centers. Calgary and Edmonton exhibited higher ratio classes, but not for their entire core. The peripheral edges of the Calgary-Red Deer corridor, the communities along the Banff-Calgary-Brooks corridor, the Fort McMurray surroundings, and the northern Fox Creek area were high for both SGA and LBWT. Jasper and south-east Alberta had higher SGA, while the communities west and east of Edmonton had higher LBWT. The distributions of the type of SGA/LBWT spatially varied across the province – differences that may have been due in part to population and behavior, but also visually collocated with the higher amounts of outdoor hazard mapping.

Separately, the air substances and land sources varied in association with the SGA/LBWT distributions. On the provincial scale, there were 13 hazards spatially related to both the SGA and LBWT ratios. Assessing the relationships sub-provincially found many more factors involved, including those already supported in the scientific literature, including: nitrogen oxides, particulate matter ($PM_{2.5}$ and PM_{10}), and sulphur dioxide.

Despite the disparate boundaries, spatially corresponding health regions (HR) and airshed zones (AZ) had comparable patterns in spatial relationships to the hazard indices. HR 4822 / AZ PAS had highly negative correlations with all indices, suggesting that factors other than the outdoor environment may be more important in these regions. HR 4829 / AZ WBEA exhibited opposite correlations with indices: SGA was positive and LBWT was negative. HR 4826 / AZ ACAA and FAP and HR 4823 / AS CRAZ for SGA and LBWT were positively correlated with the indices – these are the more populated regions. 4828 / PASZA also had positive index correlations with SGA and LBWT. HR 4824 / AZ PAMZ for SGA was positive with the indices, but for LBWT had no association. The reverse was found in HR 4821 (no corresponding AZ), where SGA was negative and LBWT was positive. HR 4825 had no relationship with SGA/LBWT, and AZ LICA had no association with SGA and a positive one with LBWT. HR 4827 and AZ WCAS are too large and diverse to compare. Inconsistent relationships for each SGA/LBWT with the indices may be due to: (i) the variable geography within the administrative boundaries; (ii) differences in etiology of SGA/LBWT; and/or (ii) the actual distribution of each SGA/LBWT exhibiting slightly different patterns: SGA and LBWT appear to be more of a heartland issue.

The combination of the outdoor hazards into a single index were very weakly associated with SGA and LBWT provincially. This was not surprising given Alberta includes forestry, agriculture, and energy extraction activities, thus yielding diverse “pockets” of different pollutants. Analyzing smaller geographic areas, based on health regions or airsheds, helped recognize possible differences in the outdoor environmental factors.

The large area of some units may capture populations that are more similar in size to the smaller units, but the environmental variability may have diluted the effects of hazards. The sub-provincial units that had negative correlations will need further analysis to determine the regionally important hazards. Relationships found here show that province-wide (i.e. large region) approaches to outdoor hazards may be inappropriate or inefficient. Where health regions and airshed zones are more similar, policy and monitoring may be more in sync.

Existing ambient hazard indices are not available for comparison. Environmental Quality Indices (EQI), such as those developed by Messer et al. [129] and Stieb et al. [156] depict the state-of-the-environment from actual measured conditions [128]. The Air Quality Health Index (AQHI) by Stieb et al. does a very good job at aggregating the monitored criteria air contaminants for risk communication. Messer's EQI was associated with pre-term birth [157], but still has the limitation of fixed administrative units. And because a main goal was a continuous index, I was unable to incorporate an effective rural classification without the introduction of administrative boundaries, as done by Messer et al. My more ecologically-encompassing index incorporated industrial air pollutants and land-based sources, similar to the holistic model developed for a single urban area by Tarocco et al. [158].

Limitations

I analyzed the entire registered birth population for the study period that had valid locations. The 6-character postal codes provided good accuracy for urban neighborhoods, especially within the context of the 250 m cell size, but the rural residences were not as exact. DMTI Spatial had applied algorithms to weight the postal code local delivery area centroid toward the more populated communities [110], but that did not guarantee an actual residence contained within the cell. The problem of rural resolution was exhibited by oil-gas well pads and agricultural land that may be closer to actual residences, but postal codes were not accurate enough due to too large of delivery areas for the centroids.

Although there is concern that the mother did not live at that postal code for the entire pregnancy, previous research determined low mobility during pregnancy and any relatively short distances moved did not substantially change the exposure assignment [159].

The spatial data for the independent variables were restricted to publicly available sources that may not have had the most temporally appropriate capture date of the mapped features. I also did not have access to reliable province-wide data for other possible environmental factors, such as water quality, noise, or non-industrial pollution sources. And as suitable as the NPRI data were, the values were annually reported estimates and not actual measurements [105]. Despite these shortcomings, the available data provided an as inclusive as possible foundation for the index.

Many of the GIS methods involved the selection of radius distances. The size of the radius used in calculating the DKD affected how "hot" an area appeared and may have exaggerated the extent for large distances; the 25 km radius may have been too large for rural communities with diverse topographies. When estimating air-emitted pollutants, wind would have varied by season and throughout the years; therefore the use of circular shapes in calculating the tonnes per area may not have accurately reflected wind-dispersion for some areas. The conservative 10 km radius for spreading the air substances may have remedied this for upwind locations, but potentially underrepresented it for downwind locations. For the index, not all variables may be equally important, but the use of expert judgment would have introduced subjectivity that was not reproducible. Therefore, the equal treatment of the air substances and land sources in the overlay analyses was used.

The correlation threshold value of 0.40 may have overrepresented the inclusion of some of the independent variables. The choice of this statistical threshold was based on inspection of the data to ensure that a wide variety of hazards would be represented and not erroneously overlooked due to the modifiable areal unit problem introduced by the boundaries of the sub-provincial units [160].

It is important to stress that my research was not able to find causal relationships but identified where outdoor environmental hazards collocate with residences of mothers who gave birth to abnormally small newborns.

Strengths

The calculations of the outdoor environmental variables were continuous and covered the entire study area. Therefore the DKD calculation of the SGA and LBWT ratios was appropriately consistent because it also was not confined to arbitrary geographical boundaries. Aggregation early in the analysis would have produced an inflexible distribution of SGA/LBWT. The introduction of health regions and airsheds afterward allowed for scenario investigations relevant to health care administration, policy implications, and airshed monitoring.

The primary outdoor pollutants associated with abnormally small newborns agreed with published research, but additional unstudied air substances were discovered. For many regions, the reduction of data into a single index was achievable.

The development and application of the ambient health hazard indices for any study area, any time period, and where relevant data are available is simplified by the reverse suitability approach in a standard GIS. The distance-centered methods and weighted sum overlay, commonly used in wildlife habitat studies, are also relevant to human habitat related to various environmental health outcomes.

Conclusion

This is to date the first study on abnormally small newborns that used a combination of multiple outdoor variables over a large geographic area. My results showed that SGA and LBWT varied sub-provincially with outdoor environmental factors, suggesting that provincial government should be aware of multiple sources of place-dependent exposures. Summing up class rankings of hazards provided a simple model for correlating with the sub-provincial distributions of SGA/LBWT. There were regions/airsheds that were higher than the national and provincial rates. The temporal nuances had been masked by combining all years: spatial patterns in the hazards and birth outcomes likely varied through time; therefore, future research should consider the timing of exposures. Research should also combine the vertices of habitat, population, and behavior to investigate the complex interactions of the outdoor hazards found here by including maternal characteristics revealed in traditional epidemiological studies. I found that the industrial air substances were important – and the Overlay 0.7/0.3 weighted index had the most associations in the sub-provincial units. Therefore, both the individual air substance associations and the convenient single-measure index provide complementary information to move us toward a better understanding of the links between the outdoor environment and birthweight. Mapping the outdoor environmental hazards for mothers giving birth to abnormally small newborns provides insight for preventative or remedial recommendations *where* they may be needed to help determine healthier futures.

Chapter 4 Spatiotemporal patterns of small newborns and associations with land use and socioeconomic status

4.1 Abstract

Background: In addition to small for gestational age (SGA) and low birth weight at term (LBWT), critically ill cases of SGA/LBWT are significant events from outcomes and economic perspectives that require further understanding of risk factors. My research aimed to assess the spatiotemporal distribution of the regions where there were consistently higher numbers of critically ill small newborns (hot spots) in comparison with all small newborns, and all births. I also assessed relationships with surrounding small-area-level land use and socioeconomic status (SES) in Calgary (2006-2010) and Edmonton (2008-2010), Alberta.

Methods: I created space-time cubes from residential locations of singleton newborns, small newborns, and those admitted to the neonatal intensive care unit, classified as SGA, LBWT, or the critically ill counterparts (ciSGA or ciLBWT) using geographical information systems (GIS). Then I applied emerging hot spot analyses to identify trends in the clustering of point densities. I compared the resulting aggregated categorical patterns with proportions of land use and socioeconomic status (SES) using Spearman's correlation and logistic regression.

Results: There was an overall increasing trend in all space-time clusters (Mann-Kendall statistics ranged from 1.86-6.72, $p<0.01-0.06$). Whole period emerging hot spot patterns among births and SGA generally coincided (Spearman's *rho* ranged from 0.47-0.48, $p<0.05$), but SGA and ciSGA did not (*rho*<0.20, $p<0.05$). Regression coefficients were highest for low SES with SGA (Calgary $\beta=4.9$; Edmonton $\beta=3.4$; $p<0.05$) and LBWT (Calgary $\beta=3.9$; Edmonton $\beta=4.5$;

p<0.05). Industry was most associated with ciLBWT (Calgary $\beta=3.4$; Edmonton $\beta=2.3$; $p<0.05$) but no statistically significant associations with ciSGA.

Discussion: The weak association among hot spot patterns for small newborns and critically ill small newborns was quantified by correlation and regression and supported by visualizations of hot spots occurring in different areas of the maps. The dominant area-level associations were low SES and industrial land use. The difference in space-time hot spot patterns and the associations with ciSGA and ciLBWT indicate further need to research the interplay of maternal and environmental influences. I demonstrated the novel application of mapping the space-time patterns of small newborns and spatially related them to the surrounding environment.

4.2 Background

Being born too small, such as low birth weight at term (LBWT) – defined as birth weight below 2,500 g for full term pregnancy – is considered an adverse birth outcome (ABO) because it is associated with infant mortality, physical and cognitive disabilities, and chronic health issues [1, 7, 153]. However, this absolute parameter does not take into consideration gestational age. To account for variability in birth weight at different gestation, another parameter called small for gestational age (SGA) is used. SGA is defined as birth weight below the 10^{th} centile weight, based on sex and weeks of gestation [13].

In Canada, the average rate of SGA was reported to be 8.4% and LBW (all gestational ages <2,500 g) was 6.0%, during 2005-2007 [23]; whereas in Alberta, the rate of SGA was 8.8% and LBW was 6.7%. Disorders related to short gestation and low birth weight are the 2^{nd} leading cause of infant death in Canada [24]. Both these outcomes are associated with adverse consequences with higher rates of admission to neonatal intensive care units (NICU), resulting in higher economic and social costs [7, 161].

Maternal conditions (e.g., pre-existing and pregnancy-related health conditions, behavior, nutrition) are important reasons for SGA/LBWT [18, 86–88], but they do not fully explain the occurrence of SGA/LBWT. The role of environmental factors in causation of SGA/LBWT has been suspected; however, no firm conclusion/attribution has been delineated in previous studies [61, 122, 125]. To reveal patterns and associations between SGA/LBWT and the environment that may not be evident in traditional spatial epidemiology, spatial statistics and geographic data mining in GIS allows for spatial-temporal variation because interactions of the environment are not constant [162]. Geographical Information Systems (GIS) are valuable for understanding patterns and the differences among births and SGA/LBWT because GIS provide various mapping techniques for public health data [32, 33, 37]. Using GIS to also analyze spatiotemporal patterns has the potential to identify priority areas for management and intervention, as has been established in other space-time pattern studies in health, crime, and conservation [163–166]. Birth events occur in space and time, so it is a natural application for emerging hot spot analysis.

Thus, my objective was to examine how hot spot patterns – in space and time – compare among pregnancies that resulted in SGA/LBWT and those that resulted in critically ill SGA/LBWT. In addition, I aimed to understand where the patterns coincide with surrounding land use and area-level socioeconomic status.

4.3 Methods

Study design and setting

I conducted a retrospective cohort study between the years 2006 and 2010 using Canadian Neonatal Network (CNN) and Alberta Perinatal Health Program (APHP) databases.

The Canadian Neonatal Network (CNN) maintains a standardized neonatal intensive care unit (NICU) database that includes all admissions to NICUs in Canada [167]. The database has shown a very high internal consistency and reliability [168]. I defined the primary areas served by the CNN NICUs as census metropolitan areas (CMA). A CMA is essentially urban core and surrounding municipalities integrated by commuting flows, and having a minimum total population of 100,000 [169]. According to census geography hierarchy, a CMA is composed of contiguous census subdivisions that may cross census division and provincial boundaries. My study area involved the Calgary and Edmonton CMAs, shown in Figure 4.1, and described in Table 4.1 in terms of size and population.

Table 4.1. Census Metropolitan Area (CMA) characteristics are from the 2011 Census for Canada; customized map projection parameters are based on the centroids of the CMA, designated in decimal degrees of longitude (central meridian) and latitude (origin): X,Y for the Azimuthal Equidistant projection with North American Datum (NAD) 1983.

Census Metropolitan Area (CMA)	Area (km^2)	Population			Central Meridian (X)	Latitude of Origin (Y)
		Total	Women 15 to 44 years	Infants 0 to 4 years		
Calgary	5,108	1,214,839	272,320	80,855	-114.078155	51.180782
Edmonton	9,427	1,159,869	252,085	73,645	-113.789137	53.512964

Figure 4.1. The study focused on the Calgary and Edmonton Census Metropolitan Areas (CMA: orange areas), in the province of Alberta, Canada, served by hospitals with neonatal intensive care units (NICUs: red crosses) participating in the Canadian Neonatal Network

The Alberta Perinatal Health Program (APHP) is an administrative clinical registry that collects and standardizes demographic information on all hospital births and out of hospital births (attended by registered midwives) for the province of Alberta [25]. The provincial data were subset to each CMA to compare with the CNN data. Calgary had five years (2006-2010) of data, but Edmonton had only three years because the participating hospital did not join the CNN until 2008.

Both CNN and APHP provided anonymized records of birthweight (grams), gestational age (completed weeks), sex, single/multiple, admission status (CNN only), pregnancy outcome (APHP only), and the residential postal code. As depicted in Figure 4.2, I selected singletons at first admission (CNN) and live births (APHP) with valid postal codes.

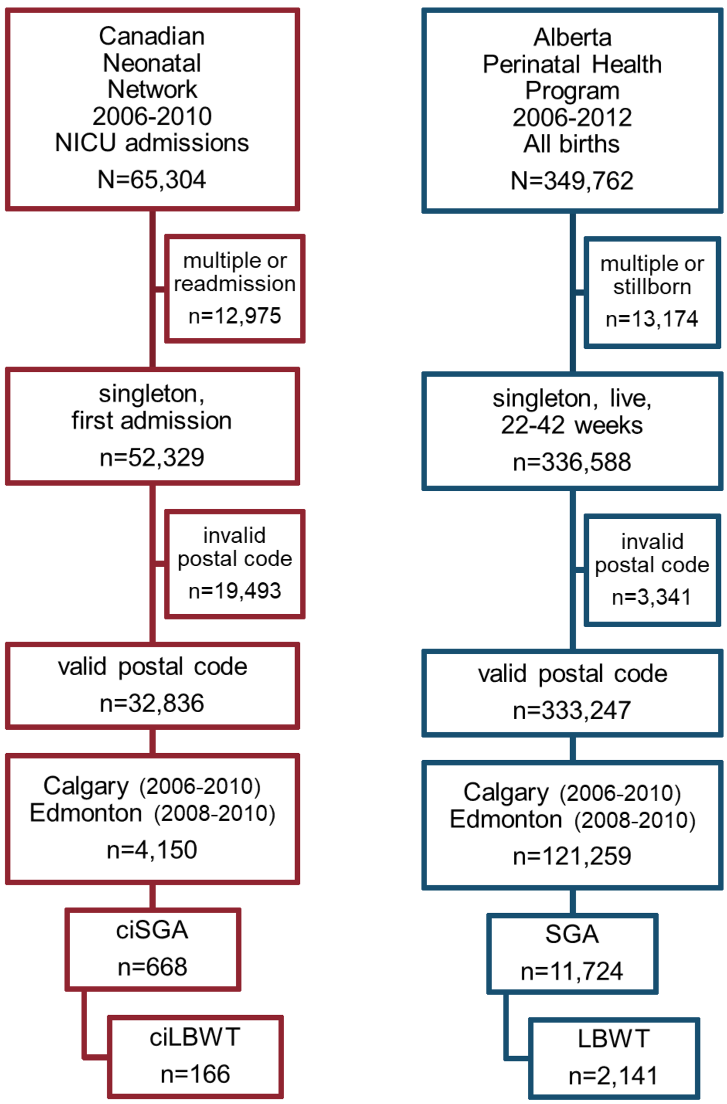


Figure 4.2. The Canadian Neonatal Network (CNN) and Alberta Perinatal Health Program (APHP) data were subset to valid postal codes within the extent of Census Metropolitan Areas (CMA): Calgary (2006-2010) and Edmonton (2008-2010).

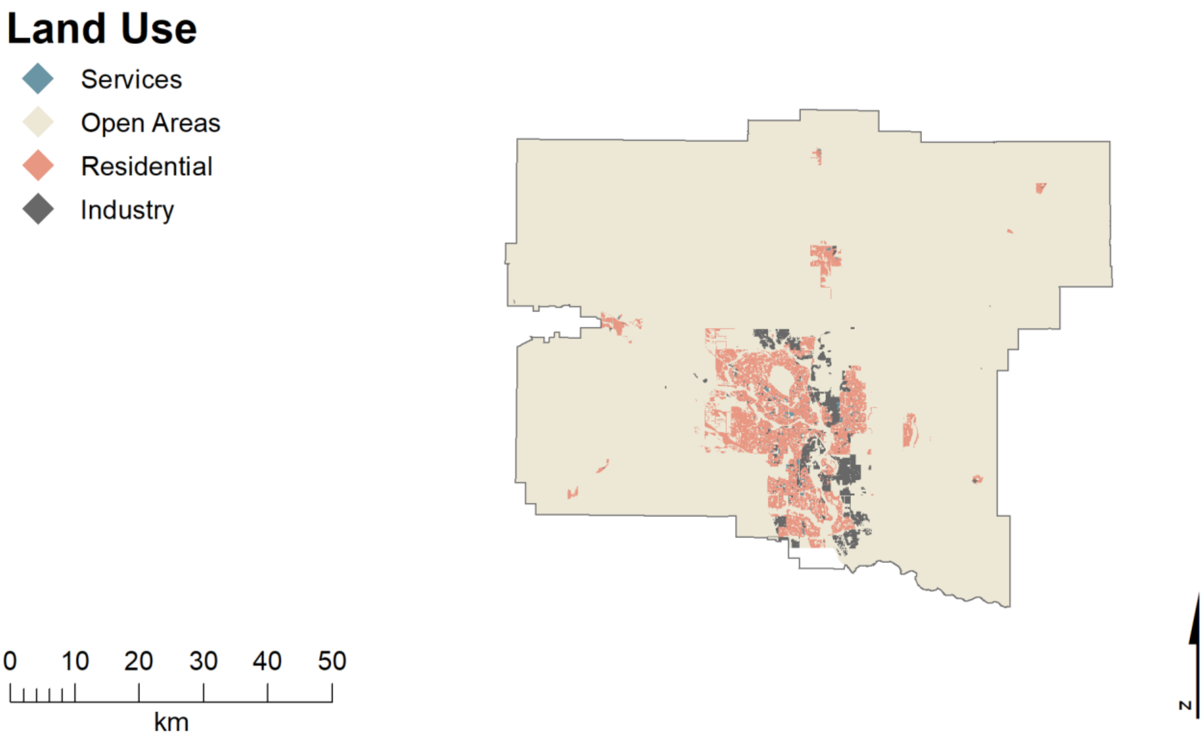
Dependent variables

Outcomes of interest were low birth weight at term (LBWT), defined as birth weight below 2,500 g at weeks 37 to 42, and small for gestational age (SGA), defined as birth weight below the 10^{th} centile for gestational age and sex according to Canadian normative data [13]. SGA and LBWT were from the APHP database. The critically ill SGA or LBWT were classified as those SGA and LBWT neonates who were also admitted to the NICU and were from the CNN database.

Independent variables

To help understand the SGA/LBWT patterns, I examined their relationships with landscape-level variables relevant to birth outcomes. These included the surrounding land use and the area-level socioeconomic status.

Digital Mapping Technologies Inc. (DMTI) Spatial provided a land use classification for the urban areas across Canada [170]. I grouped the seven standardized patterns of construction and activity that land was used for into four general categories: *services* (commercial, government/institution); *open areas* (open area, parks and recreation, waterbody); *residential*; and *industry* (resource and industry). Due to linkage with environmental pollutants, the primary category of interest was industry, defined as land occupied by establishments engaged in the mechanical or chemical transformation of materials or substances into new products, or land set aside for the extraction or production of renewable and non-renewable resources. The land use categories were mapped for Calgary in Figure 4.3 and Edmonton in Figure 4.4.



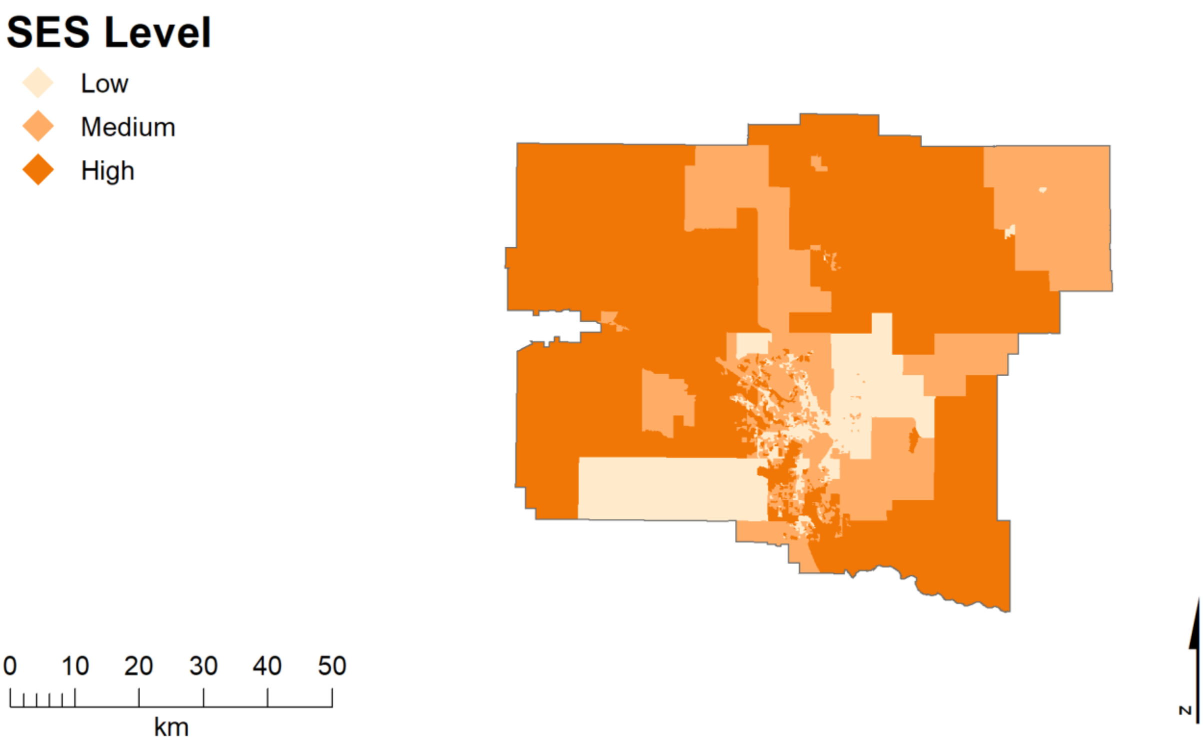


Figure 4.3. Land use and socioeconomic status (SES) maps of the Calgary CMA.

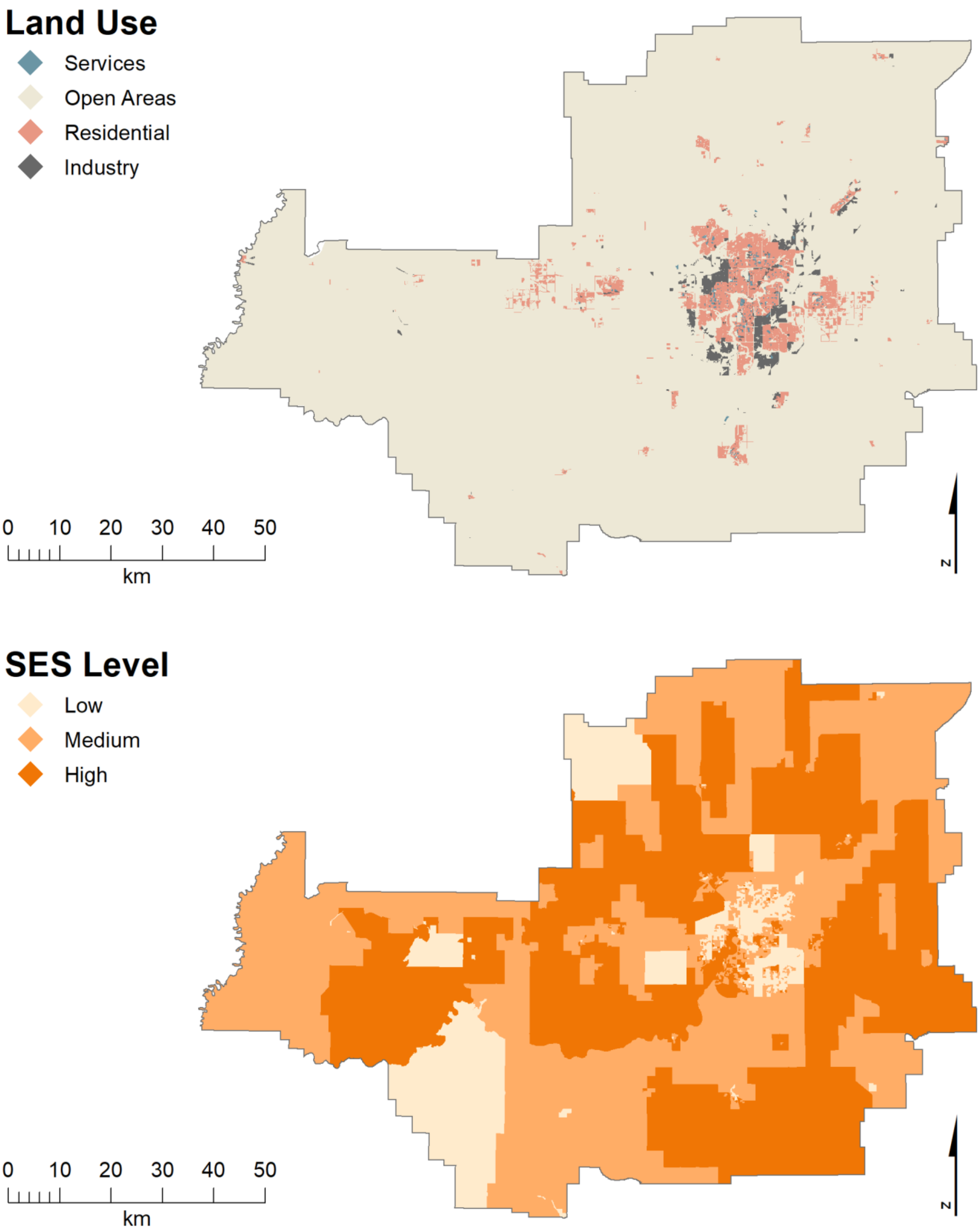


Figure 4.4. Land use and socioeconomic status (SES) maps of the Edmonton CMA.

Chan et al. provided a comprehensive index of Canadian socioeconomic status (SES) that is suitable for research in health and environmental pollutants [103]. The area-level SES index was developed from the 2006 Census Canada by incorporating 22 variables on culture, potential existence of indoor environmental pollutants, environmental injustice indicators, and deprivation variables in a principal components analysis for each dissemination area (DA). A DA was the smallest, relatively stable, geographic unit within which all census data were distributed, and was composed of contiguous dissemination blocks having a total population of 400 to 700 [169]. I grouped the SES reported as quintile values in to the following levels – *low* (0 [no data], 1, and 2), *medium* (3, 4), and *high* (5) – to indicate relative socioeconomic status for the DA. The SES levels were mapped for Calgary in Figure 4.3 and Edmonton in Figure 4.4.

Geolocation

I geolocated the CNN and APHP records by joining the postal codes to DMTI Spatial's Platinum Postal Code Suite to assign longitude and latitude [110]. To ensure static locations throughout the study period, I uniquely selected postal codes from 2001 through 2013 (the time span was necessary because new postal codes were added, and old ones were retired during the study).

My space-time analyses required a distance-preserving spatial reference; therefore, I customized Azimuthal Equidistant map projections for each CMA (Table 4.1), and implemented in Esri's ArcGIS software [101]. Figure 4.5 shows the analytical steps that are described below.

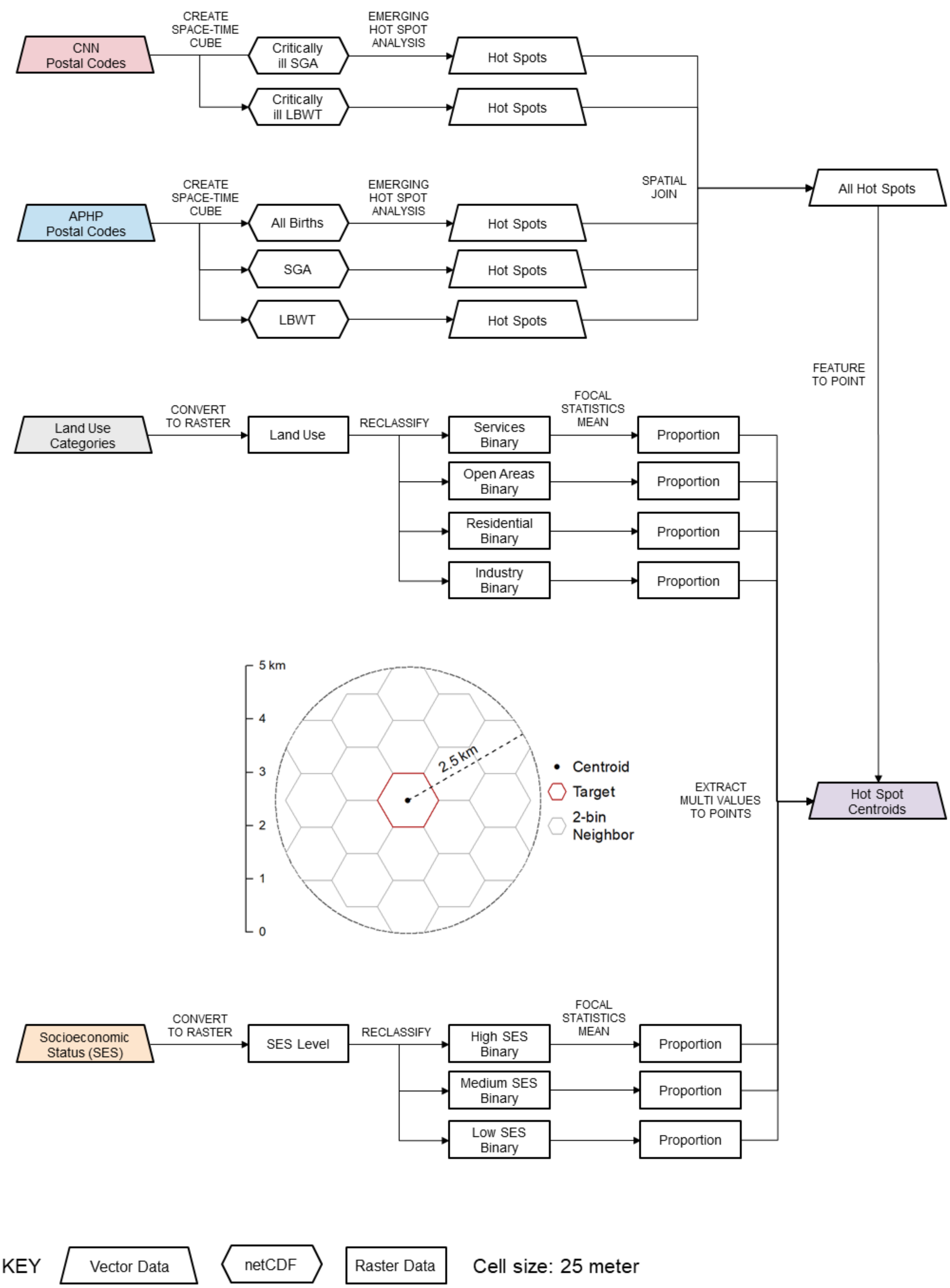


Figure 4.5. Flow chart of GIS commands for analyzing small newborns in space and time.

Spatial-temporal patterns

I analyzed the distributions and patterns of each SGA/LBWT and all births – for both the CNN and APHP data – in the context of both space and time using the ArcGIS space-time pattern mining tools [171]. For each CMA, I transformed the postal codes time-stamped by birthdate into multidimensional data cubes, stored as network Common Data Form (netCDF) files, by: (i) aggregating the points – spatially in 1-km high hexagon bins and temporally in 1-month time slices; (ii) summing the binary values of SGA or LBWT; (iii) filling empty bins with zeros; and (iv) aligning to a reference time equal to the beginning of the study (2006/01/01 for Calgary, and 2008/01/01 for the Edmonton CMA). The Mann-Kendall statistic evaluated the trend in SGA/LBWT point counts for each data cube.

The hexagon was chosen because it is more natural in shape, better represents connectivity, and minimizes edge effects [172]; the 1-km size fit within typical city neighborhoods and helped protect individual privacy. The 1-month time-step interval fit within a trimester. Bins were filled with zeros because SGA/LBWT are considered rare events, counted in whole numbers, and therefore interpolation would not be appropriate. The reference time ensured all SGA/LBWT would have the same start date for comparison purposes.

Emerging hot spot analysis (EHSA) analyzed each data cube by calculating statistically significant hot and cold spot trends in SGA and LBWT using two statistics. The Getis-Ord Gi* statistic assessed the location and degree of spatial clustering by calculating the z-score, p-value, and hot spot bin classification. The Mann-Kendall statistic evaluated these measures to assess temporal trends and then categorized locations according to Table 4.2. To simulate city neighborhood sizes, I used a fixed distance of 2001 m (note: the additional 1 m ensured that complete hexagons were included), which encompassed the current hexagon and two adjacent hexagons (2.5-3 km). To simulate a trimester, I used 2 time steps, which included the current month and previous two months (3 months). Hot spot maps were output to visualize the spatial-temporal significance of SGA, LBWT, and all births (from APHP only) in each CMA for the study period.

Table 4.2. Statistically significant hot spot categories for the 5- and 3-year study periods are defined in terms of the total months aggregated by 3-month time steps. Table continues.

Pattern Category	Emerging Hot Spot Definition
New Hot Spot	A hot spot location for the last 3 months of the time series (the final time-step interval) and has never been a hot spot before.
Consecutive Hot Spot	A *never-been-hot-before* location with a single uninterrupted run of hot spot bins in the final time-step intervals, and for <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months).
Intensifying Hot Spot	A hot spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months), including the last 3 months (final time step), and there is an *increase* in the intensity of clustering of *high* counts in each 3-month time step.
Persistent Hot Spot	A hot spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months) and has no increasing/decreasing trend in the intensity of clustering over time.
Diminishing Hot Spot	A hot spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months), including the last 3 months (final time step), and there is a *decrease* in the intensity of clustering of *high* counts in each 3-month time step.
Sporadic Hot Spot	A hot spot location that is *on-again then off-again* for <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months), and none of the time-step intervals have been cold spots.
Oscillating Hot Spot	A hot spot location for the last 3 months (the final time-step interval) that has previously been a cold spot, and <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months) have been hot spots.
Historical Hot Spot	A location that is *not* a hot spot for the last 3 months (the final time-step interval), but ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months) have been hot spots.
No Pattern Detected	Does not fall into any of the hot or cold spot patterns defined above or below.

Pattern Category	Emerging Hot Spot Definition
New Cold Spot	A cold spot location for the last 3 months of the time series (the final time-step interval) and has never been a cold spot before.
Consecutive Cold Spot	A *never-been-cold-before* location with a single uninterrupted run of cold spot bins in the final time-step intervals, and <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months).
Intensifying Cold Spot	A cold spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months), including the last 3 months (final time step), and there is an *increase* in the intensity of clustering of *low* counts in each 3-month time step.
Persistent Cold Spot	A cold spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months) and has no increasing/decreasing trend in the intensity of clustering over time.
Diminishing Cold Spot	A cold spot location for ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months), including the last 3 months (final time step), and there is a *decrease* in the intensity of clustering of *low* counts in each 3-month time step.
Sporadic Cold Spot	A cold spot location that is *on-again then off-again* for <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months), and none of the time-step intervals have been hot spots.
Oscillating Cold Spot	A cold spot location for the last 3 months (the final time-step interval) that has previously been a hot spot, and <90% of time-step intervals (Calgary: <54 months; Edmonton <32.4 months) have been cold spots.
Historical Cold Spot	A location that is *not* a cold spot for the last 3 months (the final time-step interval), but ≥90% of the time-step intervals (Calgary: 54 of the 60 months; Edmonton: 32.4 of the 36 months) have been cold spots.

Focal statistics

For both the independent variables, I reclassified the categorical values (land use, n=4; SES, n=3) in to separate binary surfaces, where "one" indicated presence and "zero" indicated absence. Then I applied a neighborhood moving-window analysis, called focal statistics. Calculating the mean statistic within a 2,500 m radius on the binary surfaces resulted in proportions. I assigned the proportions of land use and SES to the centroids of the hexagons that resulted from the emerging hot spot analysis for each SGA/LBWT. The 2,500 m neighborhood estimated the proportions of each land use or SES class within the distance defined for the emerging hot spot analysis described above.

Statistical analyses

For each CMA, I spatially joined all hot/cold spots maps, calculated Spearman's correlation on the pattern categories ranked from coldest to hottest and used the resulting statistics to determine the association of (i) SGA/LBWT with all births, or (ii) critically ill cases with all SGA/LBWT of the same type. The categories were also correlated with the land use and SES proportions to help determine any relationships with SGA/LBWT.

To better understand the relationship strength of each SGA/LBWT hot spot category and surrounding proportions of land use and SES, I used logistic regression. Binary variables were coded as "one" for all hot spot categories and as "zero" for cold spot categories and no pattern. Because the land use and SES categories were each mutually exclusive proportions, I specified residential and high SES as the references categories (i.e., left out of the model) to test my hypothesis that the target categories of industry and low SES have the highest associations with SGA/LBWT hot spot patterns. To account for areas having more births, I included the covariate sum of births (from APHP data) in each hexagon bin over the entire study period. I used STATA 12 statistical software [173]. Because I was interested only in the significance of the effect of one independent variable (X) on the response (Y), and there was no need of interpreting the coefficients, the coefficients were calculated (i.e. logarithm of the odds ratios).

4.4 Results

Characteristics of the study population

The two CMAs varied in the raw counts of all births, all small newborns (SGA or LBWT), and critically ill small newborns. As shown in Table 4.3, Calgary had 77,711 total births over five years, there were 7,907 (10.2%) SGA, 505 (0.7%) critically ill SGA, 1,462 (1.9%) LBWT, and 126 (0.2%) critically ill LBWT. For Edmonton's 43,548 births over three years, there were 3,817 (8.8%) SGA, 163 (0.4%) critically ill SGA, 679 (1.6%) LBWT, and 40 (0.1%) critically ill LBWT.

Table 4.3. Census Metropolitan Area (CMA) number of records are from the Alberta Perinatal Health Program (APHP) and Canadian Neonatal Network (CNN) databases for only the records having valid 6-character postal codes. Note: SGA=small for gestational age; LBWT= low birthweight at term.

Census Metropolitan Area (CMA)	Years	APHP			CNN		
		Births	**SGA**	**LBWT**	**NICU Admissions**	**Critically ill SGA**	**Critically ill LBWT**
Calgary	2006-2010	77,711	7,907	1,462	2,908	505	126
Edmonton	2008-2010	43,548	3,817	679	1,242	163	40
Both CMAs		**121,259**	**11,724**	**2,141**	**4,150**	**668**	**166**

Space-time cube trends

When the space-time cubes were created, information on the overall data trend was reported. The nonparametric Mann-Kendall statistic, an aspatial time-series analysis, indicated whether the events increased or decreased over time by evaluating count values for the locations in each three-month time-step interval for my study. Table 4.4 contains the trend statistics, which showed increasing trends for every SGA/LBWT and births, in both CMAs. The Mann-Kendall statistics ranged from 1.86 to 4.89 (p-values: <0.01 to 0.06) in Calgary, and 2.56 to 6.72 (p-values: <0.01 to 0.01) in Edmonton; both were positive and much higher than the expected zero value of no trend.

Emerging hot spot patterns

Table 4.4 identifies the patterns that resulted from the emerging hot spot analyses (EHSA) for each SGA/LBWT in the CMAs. Because the areal and temporal extents differed in each study area, the proportions of each category are shown. The EHSA pattern categories are defined in Table 4.2 within the context of Calgary's 60-month and Edmonton's 36-month time series. Calgary had more variability in hot/cold spots with two to 12 categories; Edmonton had two to five categories. The largest proportions of both CMAs had no patterns. Small amounts of new hot spots were present in SGA/LBWT, except Edmonton's ciLBWT. Consecutive hot spots occurred in all SGA/LBWT for Edmonton, but only for ciSGA/ciLBWT and all births in Calgary. Intensifying, persistent, and diminishing hot spots occurred in Calgary for all births and SGA. Sporadic hot spots were present in all births and every SGA/LBWT, with the highest proportion in Edmonton's SGA. Oscillating hot spots had the highest proportion in Edmonton but occurred in both CMAs for all births. Cold spots occurred in both CMAs (Calgary had six cold categories; Edmonton had two), but only for all births. Overall, the proportions of each pattern indicated that sporadic and consecutive hots spots dominated the trends; and births in both CMAs also exhibited cold spots.

Table 4.4. Space-time cubes and emerging hot spot analyses exhibit increasing trends across Alberta Perinatal Health Program (APHP) all births, small for gestational age (SGA), low birthweight at term (LBWT) and Canadian Neonatal Network (CNN) critically ill SGA and critically ill LBWT. Proportion of each hot/cold spot category shown; pattern categories defined in Table 4.2.

		Calgary					**Edmonton**				
		APHP=865 locations			CNN=568 locations		APHP=1032 locations			CNN=locations	
		Births	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**	**Births**	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**
Trend		↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Mann-Kendall Statistic		4.89	3.07	1.86	3.65	2.22	6.72	6.66	5.72	3.71	2.56
p-value		< 0.01	< 0.01	0.06	< 0.01	0.03	< 0.01	< 0.01	< 0.01	< 0.01	0.01
Sparseness (% non-zero)		52.75	12.8	2.70	1.46	0.36	27.57	5.38	1.07	0.56	0.14
No Pattern		0.508	0.874	0.939	0.979	0.944	0.421	0.684	0.898	0.937	0.939
Hot Spots	New	-	0.001	0.010	0.002	0.018	-	0.008	0.004	0.014	-
	Consecutive	0.003	-	-	0.004	0.018	0.002	0.045	0.002	0.011	0.009
	Intensifying	0.112	0.015	-	-	-	-	-	-	-	-
	Persistent	0.045	0.020	-	-	-	-	-	-	-	-
	Diminishing	0.013	0.003	-	-	-	-	-	-	-	-
	Sporadic	0.082	0.084	0.051	0.016	0.021	0.009	0.264	0.096	0.038	0.052
	Oscillating	0.006	-	-	-	-	0.513	-	-	-	-
	Historical	0.001	0.001	-	-	-	-	-	-	-	-
Cold Spots	New	0.001	-	-	-	-	-	-	-	-	-
	Consecutive	-	-	-	-	-	-	-	-	-	-
	Intensifying	0.043	-	-	-	-	-	-	-	-	-
	Persistent	0.090	-	-	-	-	-	-	-	-	-
	Diminishing	0.014	-	-	-	-	0.016	-	-	-	-
	Sporadic	0.082	-	-	-	-	0.040	-	-	-	-
	Oscillating	-	-	-	-	-	-	-	-	-	-
	Historical	-	-	-	-	-	-	-	-	-	-
Hot/Cold Trends		0.492	0.126	0.061	0.021	0.056	0.579	0.316	0.102	0.063	0.061
Category Count		12	6	2	3	3	5	3	3	3	2

Pattern comparisons among SGA/LBWT

In Calgary, there were six distinct areas of hot spot patterns for all births (indicated by red toned symbols in Figure 4.6). The largest patch was in the northeast, and smaller ones in the northwest, northcentral, central, southcentral, and southeast. The five distinct areas of SGA occurred in the northeastern (largest), northcentral, central, southcentral, and southeast (Figure 4.7). Much smaller areas were observed for critically ill SGA: central and scattered in the northwest (Figure 4.7). Figure 4.8 shows two separate hot spot patterns for LBWT in the northeast, one in the east, one central, and an outlying community. The distinct areas for critically ill LBWT were northeast, central (but expanded beyond LBWT), and in the southeast (Figure 4.8).

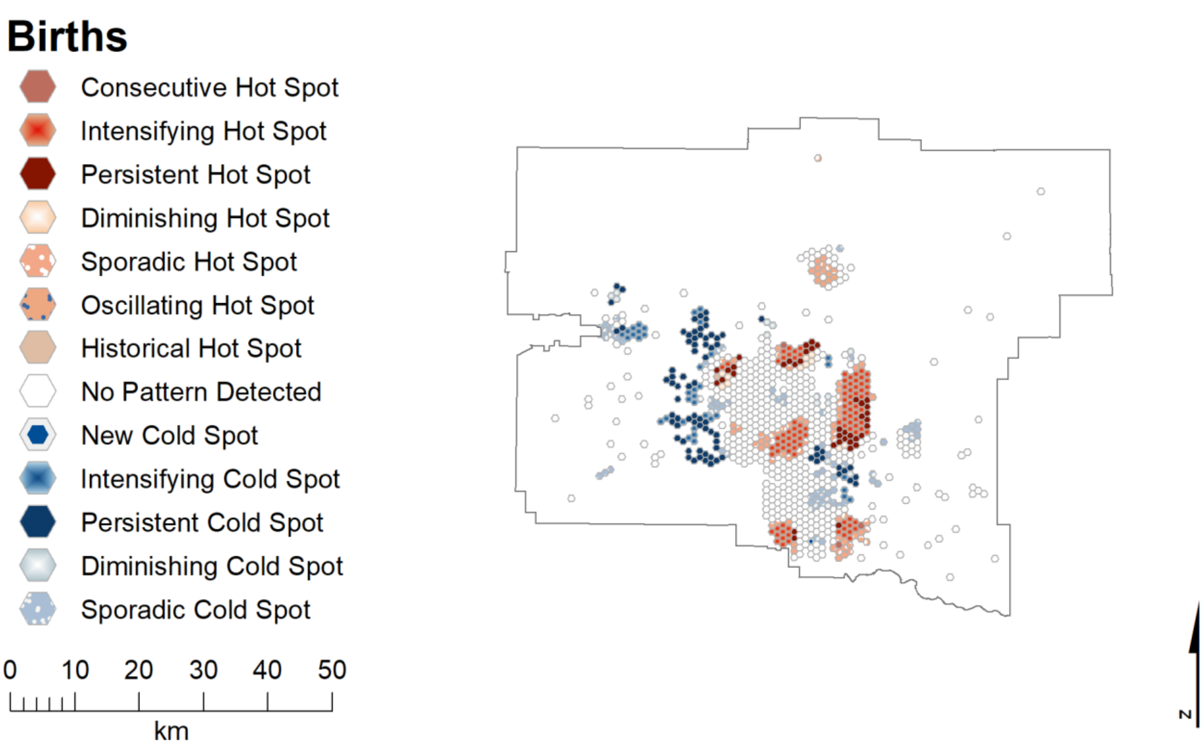


Figure 4.6. Emerging hot spots of all births in the Calgary CMA.

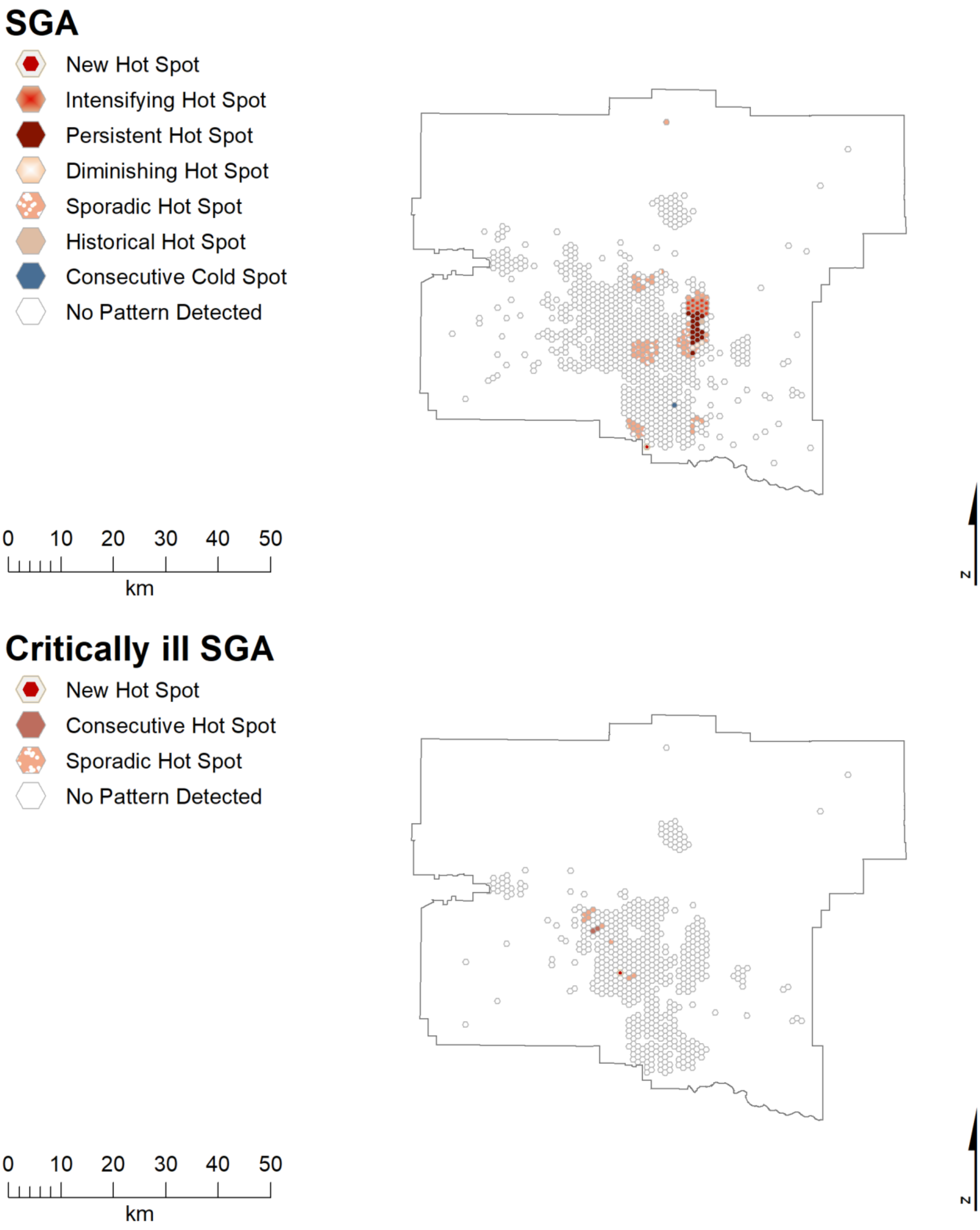


Figure 4.7. Emerging hot spots of SGA and critically ill SGA in the Calgary CMA.

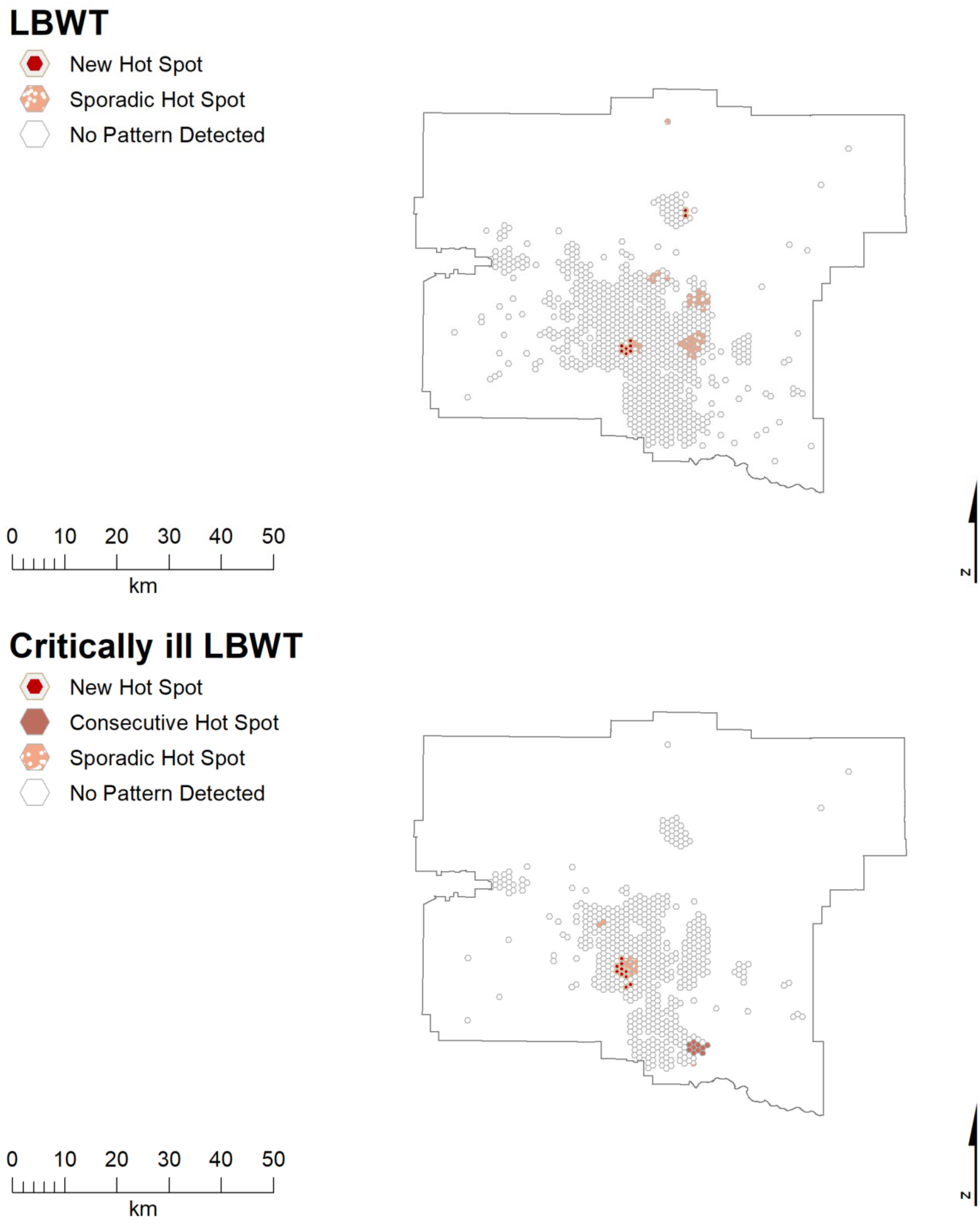


Figure 4.8. Emerging hot spots LBWT and critically ill LBWT in the Calgary CMA.

In Edmonton, there were oscillating hot spots for all births covering most of the core CMA (Figure 4.9). Figure 4.10 shows distinct areas of SGA occurred in a large band from the northeast through central to west, across the south, and in outlying communities. Much smaller areas were seen for critically ill SGA: northcentral, west, and southeast (Figure 4.10). Figure 4.11 shows hot spots for LBWT in the north-northwest, north-central, southeast, west of central, west, and south. Three distinct areas were seen for critically ill LBWT: northwest, south-southeast, and an outlying community (Figure 4.11).

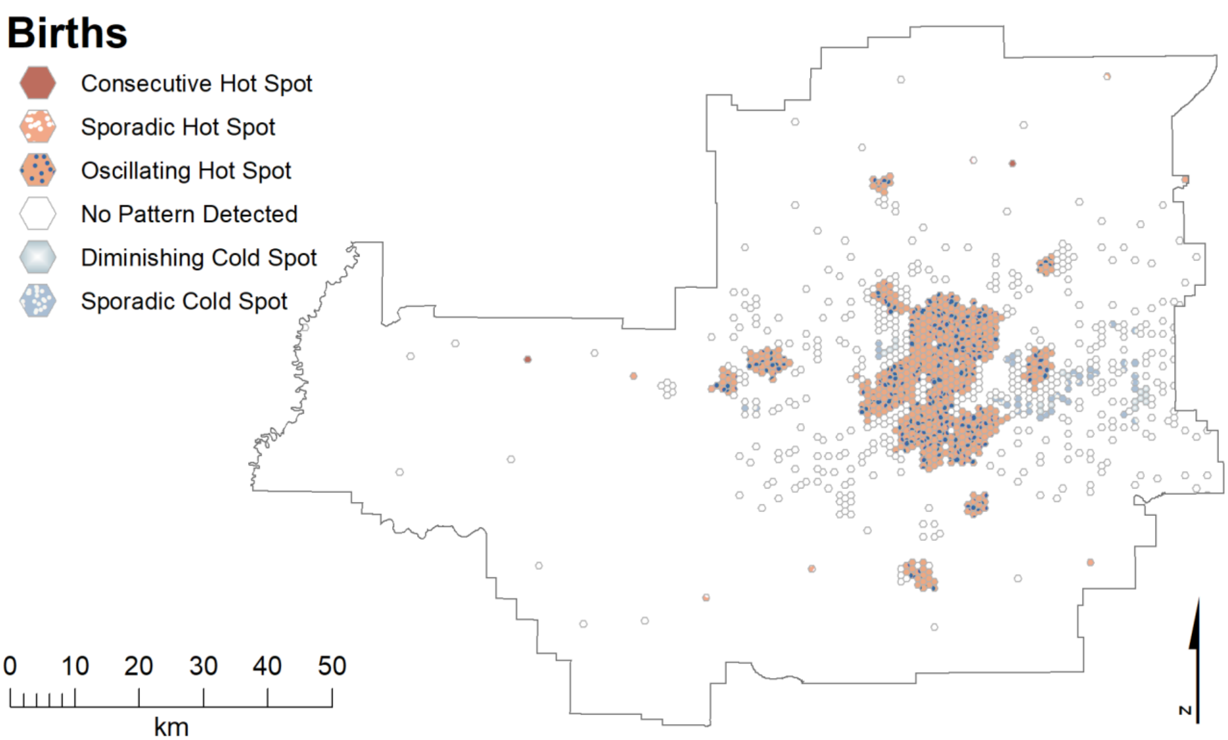


Figure 4.9. Emerging hot spots of all births in the Edmonton CMA.

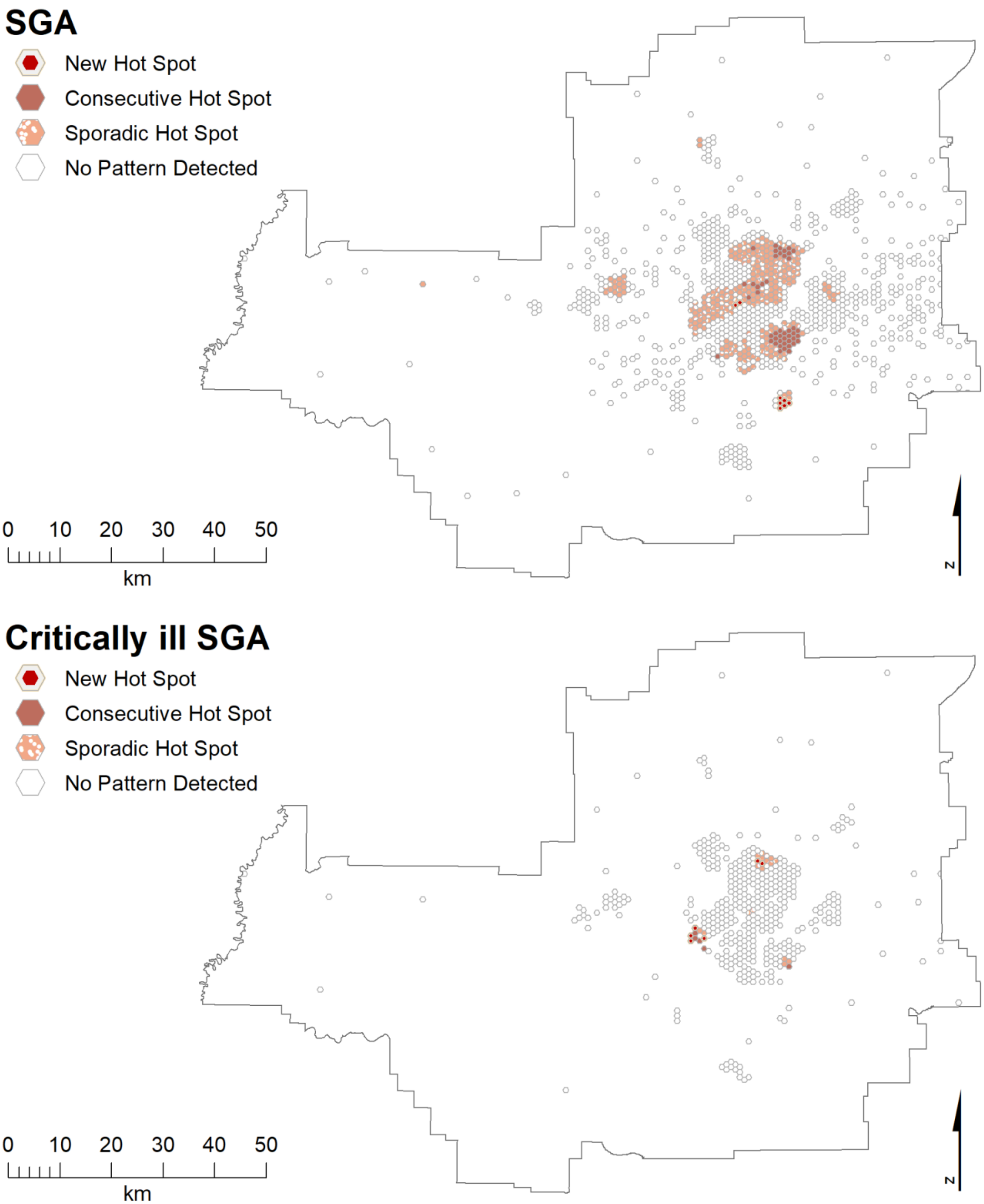


Figure 4.10. Emerging hot spots of SGA and critically ill SGA in the Edmonton CMA.

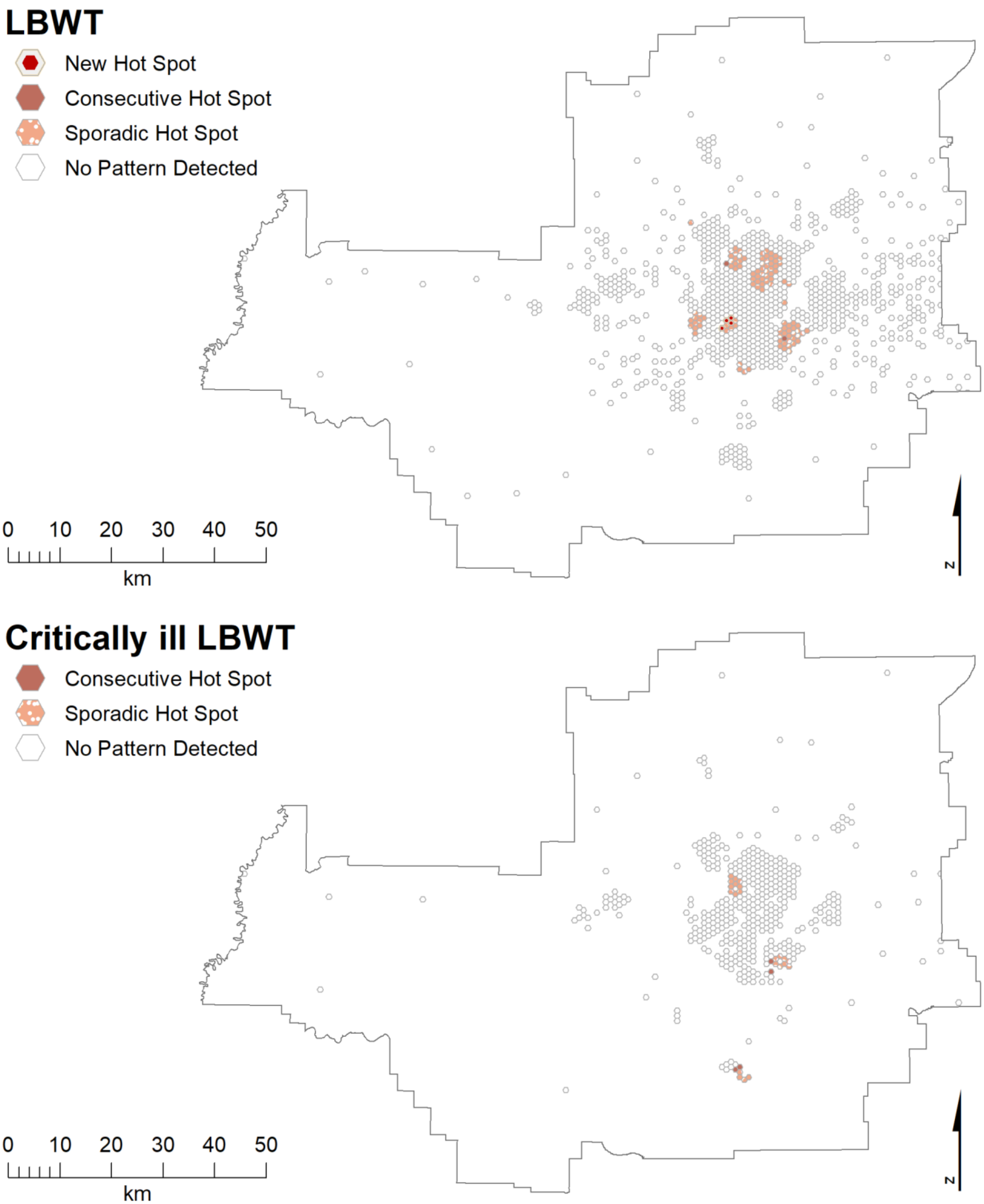


Figure 4.11. Emerging hot spots of LBWT, and critically ill LBWT SGA in the Edmonton CMA.

Table 4.5 reports the Spearman's correlations among all births, SGA/LBWT, and critically ill SGA/LBWT. A *rho* value closer to positive one indicates higher association among the hot spot category patterns; a minus sign (-) indicates a negative association. For both CMAs, there was very poor association overall (*rho* 0.09 to 0.48, $p<0.05$), with the highest between all births-SGA. The correlations decreased from SGA/LBWT to critically ill SGA/LBWT ($p<0.05$): in Calgary, all births-SGA was *rho*=0.47, SGA-critically ill SGA was *rho*=-0.03, all births-LBWT was *rho*=0.31, and LBWT-critically ill LBWT was *rho*=0.15; in Edmonton, all births-SGA was *rho*=0.48, SGA-critically ill SGA was *rho*=0.18, all births-LBWT was *rho*=0.18, and LBWT-critically ill LBWT was *rho*=0.13.

Associations of space-time patterns with land use and SES

The direction and relative *rho* values of Spearman's correlations ($p<0.05$) gave insight to which land use and SES categories had any relationships with the SGA/LBWT space-time hot spot patterns. As shown in Table 4.5, all births and SGA were associated the most with land use and SES categories for *rho*>|0.4|.

In Calgary, SGA hot spots were negatively associated with high SES (*rho*=-0.42); no strong associations were seen for all births, LBWT, or either critically ill SGA/LBWT.

In Edmonton, birth hot spots were positively associated with low SES (*rho*=0.41) and residential (*rho*=0.48), and negatively with open areas (*rho*=-0.52); SGA hot spots had similar associations (low SES: *rho*=0.43; residential: *rho*=0.44; open areas: *rho*=-0.40) but were also negatively associated with high SES (*rho*=-0.41); no strong associations were seen for LBWT, or either critically ill SGA/LBWT.

Table 4.6 indicates the correlation among land use and area-level SES, suggesting the variables of interest were independent in the Calgary CMA, but relatively less independent in the Edmonton CMA. Open areas and services were noticeably negatively correlated (Calgary *rho*=-0.66; Edmonton *rho*=-0.73), and the same negative relationship was seen for open areas and residential (Calgary *rho*=-0.85; Edmonton *rho*=-0.84).

The logistic regression model coefficients are displayed in Table 4.7, where residential land use and high SES were the reference variables. According to the Pseudo R^2 values, the model fit ranged from 0.30 (critically ill SGA, Edmonton) to 0.45 (SGA, Calgary and Edmonton), meaning 30-45% of the SGA/LBWT hot spot variations were explained by area-level land use and SES.

In Calgary ($p<0.05$), birth hot spots were surrounded by more industry (β=3.2 [95%CI: 1.5, 5.0]) and low SES (β=2.1 [95%CI: 1.4, 2.9]); SGA hot spots were surrounded by more area of low SES (β=4.9 [95%CI: 3.7, 6.2]); LBWT hot spots were surrounded by more area of low SES (β=3.9 [95%CI: 2.5, 5.4]); critically ill SGA hot spots were not significantly different from the reference; and critically ill LBWT hot spots had higher open areas (β=1.7 [95%CI: 0.6, 2.8) and industry (β=3.4 [95%CI: 1.6, 5.2]).

In Edmonton ($p<0.05$), birth hot spots were surrounded by more medium SES (β=3.4 [95%CI: 2.6, 4.3]); SGA hot spots were surrounded by low SES (β=3.4 [95%CI: 2.4, 4.4]) and medium SES (β=3.3 [95%CI: 2.4, 4.3]); LBWT hot spots were surrounded by low SES (β=4.5 [95%CI: 3.2, 5.7]); critically ill SGA hot spots had slightly more open areas (β=1.6 [95%CI: 0.5, 2.7]); and critically ill LBWT hot spots had more open areas (β=1.6 [95%CI: 0.5, 2.8]) and industry (β=2.3 [95%CI: 0.4, 4.2]).

Table 4.5. Spearman's correlation (rho) statistics compare emerging hot spot patterns for all births, SGA/LBWT, and critically ill SGA/LBWT, by Census Metropolitan Area (CMA). Patterns are also correlated with proportions of each land use and SES category. Significant *rho* values (p<0.05) are marked with an asterisk (*).

	Calgary					Edmonton				
Spearman's *rho*	**Births**	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**	**Births**	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**
Births	1					1				
Small for Gestational Age	0.47*	1				0.48*	1			
Low Birthweight at Term	0.31*	0.47*	1			0.18*	0.23*	1		
Critically ill SGA	0.09*	-0.03	0.08	1		0.10*	0.20*	0.19*	1	
Critically ill LBWT	0.17*	-0.01	0.15*	0.23*	1	0.12*	-0.13*	0.13*	0.09	1
Land Use										
Services	0.04	0.14*	0.11*	-0.06	0.02	0.33*	0.22*	0.00	-0.10*	0.00
Open Areas	-0.29*	-0.23*	-0.08	-0.02	-0.17*	-0.52*	-0.40*	-0.24*	-0.03	-0.09
Residential	0.20*	0.11*	0.07	0.08*	0.11*	0.48*	0.44*	0.26*	0.06	-0.03
Industry	0.28*	0.20*	0.05	-0.07	0.12*	0.35*	0.23*	0.21*	0.07	0.17*
Socioeconomic Status										
SES Low	0.16*	0.38*	0.30*	-0.07	-0.07	0.41*	0.43*	0.38*	0.07	0.13*
SES Medium	-0.07	-0.13*	-0.18*	0.01	0.00	0.13*	0.06	-0.14*	0.18*	0.05
SES High	-0.23*	-0.42*	-0.30*	0.07	0.11*	-0.38*	-0.41*	-0.27*	-0.15*	-0.12*

Table 4.6. Spearman's correlation matrix of the regression model covariates: proportions of land use and socioeconomic status categories, by Census Metropolitan Area (CMA). Significant rho values (p<0.05) are marked with an asterisk (*).

Spearman's *rho*	**Calgary**						
	SES Low	**SES Medium**	**SES High**	**Services**	**Open Areas**	**Residential**	**Industry**
SES Low	1						
SES Medium	0.00	1					
SES High	-0.78*	-0.40*	1				
Services	0.53*	0.04	-0.30*	1			
Open Areas	-0.44*	-0.24*	0.33*	-0.66*	1		
Residential	0.36*	0.13*	-0.17*	0.60*	-0.85*	1	
Industry	0.20*	0.33*	-0.35*	0.13*	-0.34*	-0.07	1

Spearman's *rho*	**Edmonton**						
	SES Low	**SES Medium**	**SES High**	**Services**	**Open Areas**	**Residential**	**Industry**
SES Low	1						
SES Medium	-0.24*	1					
SES High	-0.70*	-0.34*	1				
Services	0.57*	-0.20*	-0.24*	1			
Open Areas	-0.77*	0.22*	0.50*	-0.73*	1		
Residential	0.61*	-0.2*	-0.37*	0.67*	-0.84*	1	
Industry	0.48*	0.04	-0.45*	0.14*	-0.46*	0.05	1

Table 4.7. Logistic regression β coefficients for all births, SGA/LBWT, and critically ill SGA/LBWT, modelled with proportions of surrounding land use categories and level of socioeconomic status (SES). Residential and high SES are the reference categories; Likelihood Ratio (LR) chi^2 significance is $p<0.001$; significant coefficients ($p<0.05$) marked by an asterisk (*); number of locations are indicated in Table 3.

Model *β* coefficient (95% CI)	**Calgary**				
	Births	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**
Services	-17.3 (-32.9, -1.7)*	5.8 (-17.0, 28.6)	4.5 (-21.5, 30.6)	-18.6 (-37.8, 0.7)	-7.5 (-23.4, 8.5)
Open Areas	-2.2 (-3.6, -0.8)*	-1.4 (-3.8, 1.0)	-0.4 (-3.0, 2.1)	0.9 (-0.3, 2.2)	1.7 (0.6, 2.8)*
Industry	3.2 (1.5, 5.0)*	2.3 (-0.2, 4.7)	-3.5 (-7.5, 0.6)	0.8 (-1.3, 2.9)	3.4 (1.6, 5.2)*
SES Low	2.1 (1.4, 2.9)*	4.9 (3.7, 6.2)*	3.9 (2.5, 5.4)*	0.8 (-0.1, 1.8)	0.1 (-0.7, 0.9)
SES Medium	-1.6 (-2.6, -0.6)*	1.4 (-0.2, 3.0)	1.1 (-1.0, 3.2)	0.2 (-0.6, 1.0)	-0.4 (-1.1, 0.3)
Sum Births	0.01 (0.01, 0.02)*	0.01 (0.01, 0.01)*	-0.04 (-0.04, -0.03)*	0.01 (0.00, 0.01)*	-0.02 (-0.02, -0.02)*
Intercept	-1.4 (-2.5, -0.3)*	-5.4 (-7.4, -3.4)*	0.5 (-0.7, 1.7)	-5.1 (-7.3, -2.8)*	-0.3 (-1.3, 0.7)
LR chi^2	347.1	294.5	503.1	129.3	368.2
Pseudo R^2	0.35	0.45	0.45	0.32	0.32
Model *β* coefficient (95% CI)	**Edmonton**				
	Births	**SGA**	**LBWT**	**Critically ill SGA**	**Critically ill LBWT**
Services	-34.9 (-48.6, -21.3)*	-30.1 (-40.2, -20.1)*	-34.9 (-47.3, -22.4)*	-15.2 (-25.3, -5.1)*	-13.5 (-23.8, -3.1)*
Open Areas	-9.7 (-11.4, -8.0)*	-7.0 (-8.4, -5.5)*	-4.2 (-5.9, -2.6)*	1.6 (0.5, 2.7)*	1.6 (0.5, 2.8)*
Industry	-5.5 (-7.6, -3.3)*	-5.7 (-7.5, -3.9)*	-6.1 (-8.7, -3.6)*	1.1 (-0.7, 2.9)	2.3 (0.4, 4.2)*
SES Low	1.0 (-0.2, 2.3)	3.4 (2.4, 4.4)*	4.5 (3.2, 5.7)*	0.6 (-0.3, 1.6)	0.5 (-0.4, 1.5)
SES Medium	3.4 (2.6, 4.3)*	3.3 (2.4, 4.3)*	0.9 (-0.4, 2.2)	-0.3 (-0.9, 0.4)	-0.6 (-1.3, 0.1)
Sum Births	0.03 (0.02, 0.04)*	0.01 (0.01, 0.01)*	-0.03 (-0.03, -0.02)*	0.00 (0.00, 0.01)	-0.03 (-0.04, -0.03)*
Intercept	4.4 (3.1, 5.7)*	1.2 (0.1, 2.2)*	0.88 (-0.13, 1.89)	-0.7 (-2.3, 0.8)	1.0 (0.0, 2.0)
LR chi^2	857.5	579.5	494.2	203.6	537.2
Pseudo R^2	0.60	0.45	0.36	0.30	0.39

4.5 Discussion

Hot spots for critically ill SGA and critically ill LBWT occurred in different locations than the non-critically ill, but hot spots of both SGA and LBWT logically occurred in the same locations as hot spots for all births. The differing locations were counterintuitive for the critically ill hot spots, suggesting there may be neighborhood-level environmental influences unevenly distributed across the cities.

The increasing trends of SGA/LBWT in each CMA were supported by increasing trends of all births: SGA/LBWT hot spot space-time clusters were increasing because birth hot spots were increasing. However, the locations did not coincide across the study areas, and the low correlation values quantified this difference in hot spot patterns.

The regression coefficients supported that low SES and industrial land use had the highest associations, depending on the birth outcome. However, the low regression coefficients for the critically ill SGA/LBWT suggested maternal factors and/or other environmental exposures, such as urban air pollutants may be additionally important. Higher amounts of surrounding open spaces were associated with ciSGA and ciLBWT hot spots. As the negative correlations of open spaces with services suggest, less access to health services may potentially be implied.

In Canada, there is a paucity of published studies on the spatial and temporal trends of SGA/LBWT, especially for those critically ill. Statistics Canada has reported that small newborns are increasing over time for my geographical areas of interest. Nielsen et al. [174] published on only the spatial distribution of SGA and LBWT for the entire province. As for critically ill SGA/LBWT, there are no published temporal trends for each city participating in the

CNN to compare to. The space-time patterns demonstrated here agree with the increasing national trend, but additionally pinpoints the locations of where there are hot spots.

Limitations

Although I had access to all records from the APHP and CNN databases, the postal code locations may not have been as accurate for the non-urban areas in each CMA. Also, the critically ill SGA/LBWT were reduced by half due to the loss of CNN locations without valid postal codes. This subset provided an indication of hot spot patterns for NICU-admitted critically ill small babies that can be limited.

Similar to the postal code centroids, the SES index outside of urban areas does not have as accurate spatial resolution because the dissemination areas (DA) may be vast. Larger areas are encompassed by the postal delivery units and DA in rural areas.

The use of areal units of analysis underscores the risk for ecological fallacy (aggregation bias) [9]; i.e. it must be remembered that not all births in the postal code areas may have been SGA/LBWT.

More dynamic land use and SES maps, matching the temporal resolution of the health data, may possibly strengthen relationships. The SES data was based on the 2006 Canada Census, the starting year of the study only. The land use data was published in 2016, beyond the study period, but no metadata described the exact year of the mapping, which was likely earlier. There is a small possibility that the residential areas may be overrepresented where the suburban footprint of the Alberta cities had rapidly grown. I acknowledge that it was somewhat arbitrary to consider residential as the proxy for standard land use and high SES as the optimal socioeconomic level for where pregnancies should occur; however, the collinearity with the remaining land use and SES classes required me to exclude one of each category in the logistic regression.

The CNN data collection methods differed between Calgary and Edmonton, where the latter only reported critically-ill newborns having gestational ages <33 weeks. Therefore, the ciLBWT results are not entirely reliable for the Edmonton CMA, although they are similar to the Calgary CMA. Also, because of the data difference direct comparisons cannot be made. This study was not hospital-specific, meaning that the analysis was based on the maternal residential postal code and may include NICU admissions to hospitals not in the same CMA as the residences.

Depending on alternative objectives (e.g. in epidemiology or planning policy), the reporting of coefficients (log of odds ratios) from the logistic regression model may not be suitable. Odds ratios are more easily interpretable as how much the levels of one variable (X = 1 = exposure) affects Y in relation to a reference for X (i.e. X = 0 = no exposure). The beta coefficients were useful for investigating whether any associations existed. More sophisticated statistical analyses to explore interactions of the environmental variables may be performed in the future.

The observational study design precluded any casual relationships, but instead identified differences on where hot spot patterns corresponded in space and time for birth outcomes in the two main cities of Alberta.

Strengths

For this analysis, I prepared a static postal code file spanning beyond the minimum and maximum years of the study. This was necessary because growing communities received more postal delivery routes over time, so that later births were counted in the same spatial location as earlier births.

Instead of blindly assigning land use and SES values at the centroid, spatial inaccuracy is minimized by measuring the proportions of land use and SES categories surrounding the focal hot spot hexagons. And as mentioned above, hexagons have less edge effects than squares, and more closely match the circular neighborhood used in focal statistics [172].

The user-friendly space-time cube tools allow for rapid visualization and quantification of areas with statistically significant increasing or decreasing trends of SGA/LBWT. The choice of spatial and temporal aggregation can be changed to address different research questions that may inform policy decisions on where to focus on monitoring or mitigating potential risk factors at the identified hot spots.

I was able to map the spatiotemporal trends of babies born too small, which had the end result of 2-dimensional maps for the entire time period. Then I took the analysis to the next level by associating those patterns with the surrounding environment to discover potential processes.

Conclusion

The mapping of spatial-temporal hot spots indicated that critically ill small newborns admitted to NICUs occurred in different areas than all small newborns – not what would be expected, but it was clearly demonstrated by the low correlation. The dominant area-level associations with non-critically ill SGA and LBWT hot spot patterns were primarily higher proportions of surrounding industrial land use and low SES, directly answering my research objective to help understand why the patterns were different. Only surrounding land use was associated with critically ill LBWT. However, land use or SES were not related to the critically ill SGA hot spots indicating that further research is warranted on including environmental exposures (such as air pollution from traffic and industrial sources) and maternal factors in the hot spot analyses.

Space time cubes and emerging hot spot analyses promise to be useful for any public health investigation in space and time. This is the first known study examining spatial-temporal hot spots of adverse birth outcomes.

Chapter 5 Space-time hot spots of critically ill small for gestational age newborns and industrial air pollutants

5.1 Abstract

Background: Critically ill small for gestational age (ciSGA) newborns are those who are admitted to neonatal intensive care units (NICU) and have a birthweight below the 10th percentile for gestational age and sex according to Canadian normative data. These are life-threatening and costly events requiring further understanding of risk factors. I assessed spatiotemporal hot spots of ciSGA and industrial air emissions, an infrequently studied source of shared exposures.

Methods: Using neonatal admission data from participating NICUs in the Canadian Neonatal Network between 2006 and 2010, I aggregated the mother's residential postal codes from nineteen census metropolitan areas (CMA) into space-time cubes and applied emerging hot spot analyses. Using National Pollutant Release Inventory data and Environment Canada weather station data, I estimated monthly dispersion of air emissions in these areas. I compared the resulting patterns using logistic regression, with covariates for low socioeconomic status, traffic pollution, and the total number of infants during the study period.

Results: The larger CMAs had more and larger hot spots of ciSGA in space and time. Seventy-eight industrial chemical hot spots were associated with ciSGA hot spots. The greatest number of positive associations were observed for 28 different pollutants, mostly in Edmonton, Halifax, Montréal, Toronto, Vancouver, and Winnipeg. Twenty-one of those chemicals were known or

suspected developmental toxicants, such as particulate matter, carbon monoxide, heavy metals, and VOCs.

Discussion: Hot spot patterns of ciSGA differed among CMAs. Associations with hot spots of industrial chemical emissions were geographically specific and may help explain the space-time trends of ciSGA.

5.2 Background

Environmental influences and toxicant exposures of a pregnant mother may result in a neonate born “too small” or “too soon”. This is a significant health problem associated with infant mortality, physical and cognitive disabilities, and chronic diseases later in adulthood [1, 7, 153]. Preterm neonates and neonates who are small are associated with high resource utilization including admission to neonatal intensive care units (NICU) [7, 161]. One other group of neonates characterized as small for gestational age (SGA) is also at high risk of the above-mentioned complications. SGA applies to newborns whose birthweights are below the 10^{th} percentile, based on sex and gestational age at birth [13]. Many of these neonates require admission to NICUs. Such critically ill SGA (ciSGA) infants are a high priority for research because they incur higher economic and social costs.

Maternal risk factors (e.g. pre-existing and pregnancy-related health conditions, behavior, nutrition) are important determinants of perinatal outcomes [18, 86–88]; however, adverse outcomes are not solely attributable to these factors . Incorporating other explanatory variables, such as the environment, may improve our understanding of why SGA occurs and *where* it occurs [61, 122]. SGA generally reflects growth restriction, possibly from inflammation, direct toxic effects on the placenta and fetus, interruption of oxygen-hemoglobin interaction, or DNA damage [81, 82]. Urban pollution has been associated with adverse birth outcomes [49, 52, 56, 78, 125, 175–178] but the role of specific industrial air emissions is not well understood.

It is challenging to estimate exposures for association studies between SGA and unmonitored pollutants. Pollutant Release and Transfer Registers (PRTR) are publicly available reporting systems that provide much needed data for health-related studies, despite some limitations [99]. Building on previous research utilizing PRTR data [174, 179], I developed a strategy for estimating dispersion of pollutants in space and time and mapping these emissions to where pregnant mothers lived. Because the interactions of the environment are not constant in space and time, I take advantage of spatial statistics and geographic data mining in readily available Geographical Information System (GIS). These tools allow for spatiotemporal variation to reveal patterns and associations between SGA and the environment that may not be evident in traditional spatial epidemiology [162]. Spatiotemporal GIS has the potential to identify priority areas for management and intervention in crime, conservation, and more recently, health [163–166].

In Chapter 4, I found that hot spot patterns of ciSGA did not coincide with hot spot patterns of non-critical SGA and area-level socioeconomic status (SES). The patterns were unique to the area of interest and in some instances associated with surrounding land use. Given that there is much support of air pollution in the scientific literature, I decided to use this as more specific information than land use. This study aimed to: (i) examine spatiotemporal patterns of ciSGA across 19 Canadian metropolitan areas that were served by tertiary-level (i.e. large referral hospitals that provided specialized health care) NICUs; (ii) estimate monthly emissions of industrial air pollutants, dispersed by wind, and calculate their spatiotemporal patterns; and (iii) discover associations for space-time hot spots of critically ill small newborns with those of the industrial air pollutants.

5.3 Methods

Study design and setting

I conducted a retrospective cohort study using data from the Canadian Neonatal Network (CNN), which maintained a standardized neonatal intensive care unit (NICU) database, in collaboration with 27 hospitals across Canada [167]. The database has shown a very high internal consistency and reliability [168].

I delimited the primary areas served by the CNN NICUs as 19 census metropolitan areas (CMA) containing the participating hospitals from 2006-2010, shown in Figure 5.1, and described in Table 5.1 in terms of size and population [26]. A CMA has a minimum total population of 100,000 and is defined as urban core with its surrounding municipalities connected via commuting flows [169]. I used Esri's ArcGIS software to manage, process, analyze, and map the spatial data [101, 180].

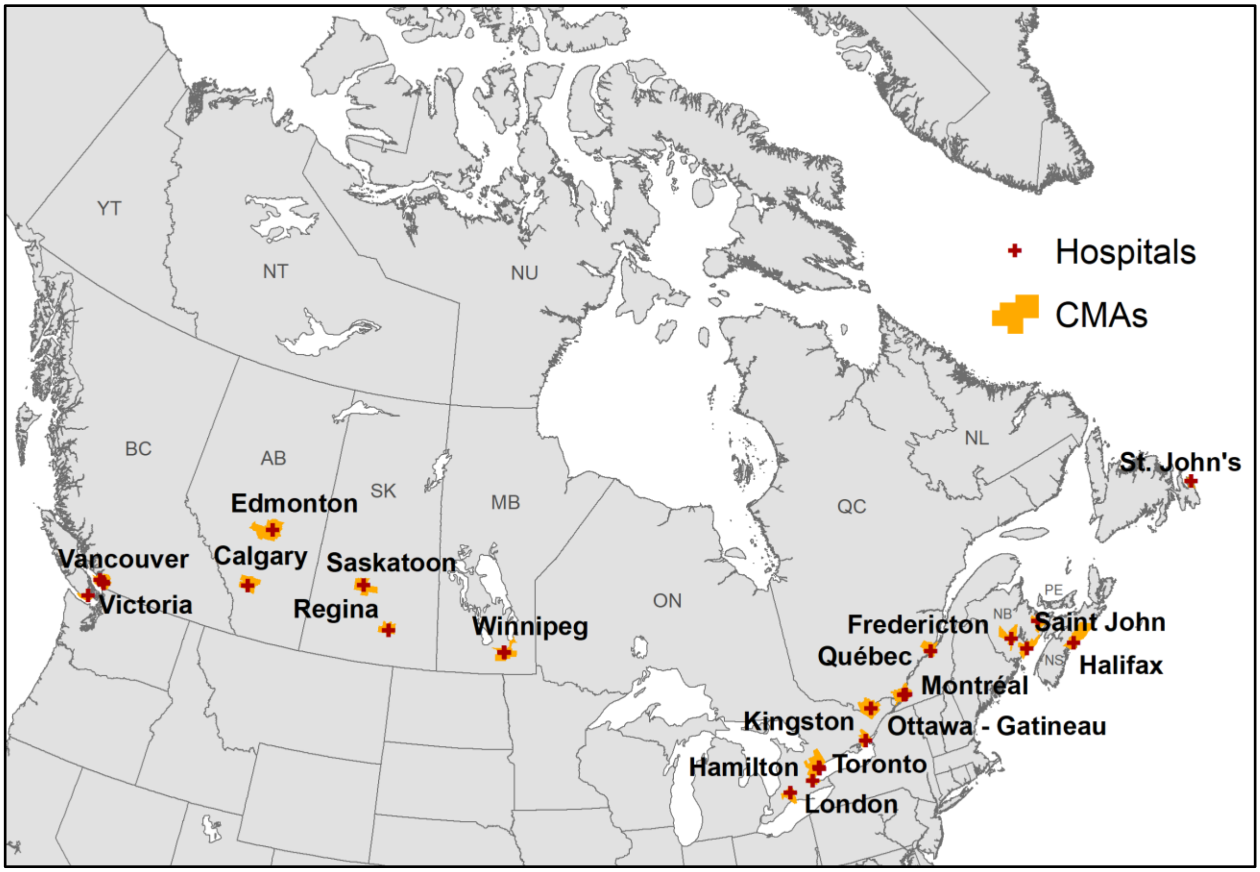


Figure 5.1. The study areas across Canada in the Canadian Neonatal Network (CNN) were defined by the 19 Census Metropolitan Areas (CMA: orange areas) that are served by 27 participating hospitals (red crosses) with neonatal intensive care units.

Dependent variables

My outcome was critically ill small for gestational age (ciSGA), based on Canadian normative data, and defined as birth weight below the 10th percentile for gestational age and sex [13] for newborns admitted to NICUs. The CNN provided anonymized records on infant birthdate, birthweight (grams), gestational age (completed weeks), sex, single/multiple, admission status, and the residential postal code. As depicted in Figure 5.2, I selected singletons at first admission to the NICU, having valid postal codes and classified all records as binary variables according to the SGA definition.

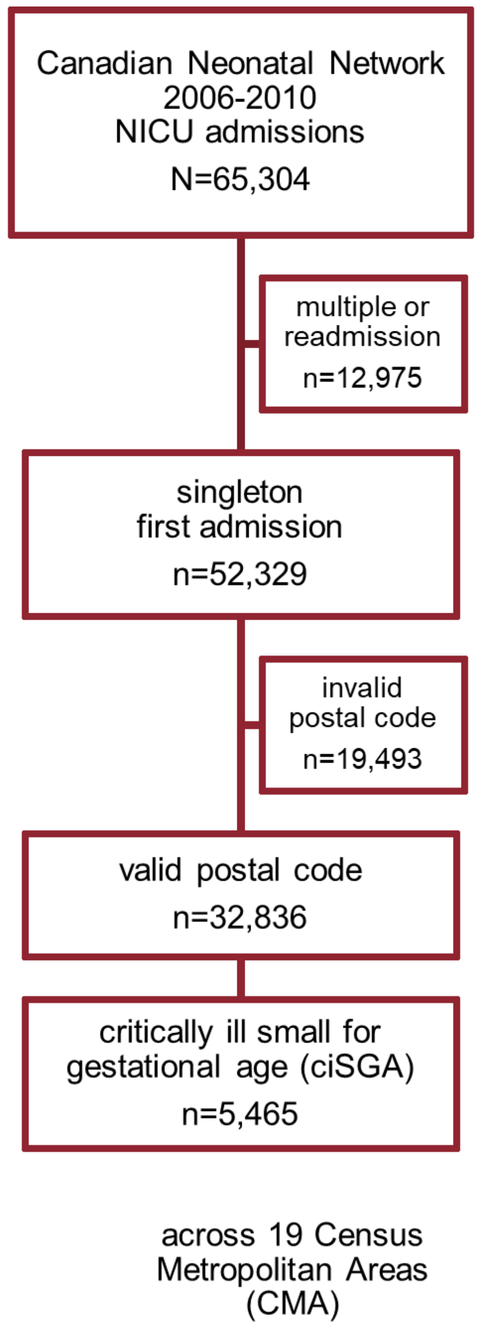


Figure 5.2. The Canadian Neonatal Network (CNN) data were subset to valid postal codes within the extent of 19 Census Metropolitan Areas (CMA): Edmonton (2008-2010), Ottawa (2007-2010), and all others (2006-2010). Critically ill small for gestational age (ciSGA) was defined as a binary variable.

Table 5.1. Characteristics of the Census Metropolitan Area (CMA) are from the 2011 Census for Canada and the Canadian Neonatal Network (CNN) database for critically ill small for gestational age (ciSGA).

CMA	Province	Area (km^2)	Total Population	Infants 0 to 4 Years	Study Start Year	CNN Hospitals	NICU Admissions*	ciSGA*
Calgary	Alberta	5,108	1,214,839	80,855	2006	1	2,908	505
Edmonton	Alberta	9,427	1,159,869	73,645	2008	1	1,242	163
Fredericton	New Brunswick	4,886	94,268	5,130	2006	1	141	24
Halifax	Nova Scotia	5,496	390,328	19,965	2006	1	1,816	291
Hamilton	Ontario	1,372	721,053	38,350	2006	1	1,486	195
Kingston	Ontario	1,939	159,561	7,865	2006	1	546	87
London	Ontario	2,666	474,786	26,150	2006	1	1,063	170
Moncton	New Brunswick	2,406	138,644	7,410	2006	1	589	92
Montréal	Québec	4,258	3,824,221	222,225	2006	4	4,645	838
Ottawa	Ontario	6,287	1,236,324	71,245	2007	1	82	19
Québec	Québec	3,349	765,706	40,775	2006	1	154	19
Regina	Saskatchewan	3,408	210,556	13,225	2006	1	1,297	188
Saint John	New Brunswick	3,363	127,761	6,740	2006	1	587	139
Saskatoon	Saskatchewan	5,215	260,600	16,625	2006	1	1,498	252
St. John's	Newfoundland	805	196,966	10,725	2006	1	735	102
Toronto	Ontario	5,906	5,583,064	318,900	2006	3	7,064	1,257
Vancouver	British Columbia	2,883	2,313,328	115,185	2006	3	3,259	496
Victoria	British Columbia	696	344,615	14,775	2006	1	1,011	141
Winnipeg	Manitoba	5,303	730,018	40,550	2006	2	2,697	487

* indicates records having valid postal codes.

Independent variables

Air pollution released by industry was my primary variable of interest. Canada's PRTR is the National Pollutant Release Inventory (NPRI), which provided the annual estimates reported by facilities mandated to do so under the Canadian Environmental Protection Act, 1999 [105]. NPRI pollutants included core and alternate substances (based on criteria of specific concentrations, quantities, number of employees), polycyclic aromatic hydrocarbons (PAH), dioxins, criteria air contaminants (CAC), and speciated volatile organic compounds (VOC). I divided the tonnes released per year by 12 to estimate the monthly average at each facility, and selected facilities within 100 km of each CMA boundary (note: the large distance guaranteed the inclusion of all potential pollutants emitted inside and outside the census boundary).

Wind direction and speed were important for calculating simple dispersion of the NPRI substances on a monthly basis. Environment and Climate Change Canada (ECCC) provided the wind measurements for centrally-located weather stations, having more than 75% observations, in each CMA [181]. For each CMA station, I aggregated hourly values to calculate: (i) overall and monthly averages of wind speed to use as an estimate of dispersal distance; and (ii) the monthly mean and standard deviation of the wind angular direction, to parameterize the extent of a wedge-shaped filter.

Because wind direction is an angular measurement, I applied trigonometry to calculate circular descriptive statistics, as described by Fisher, Rogerson, and Mardia and Jupp [182–184]. The calculations were made using Python 2.7 [111] and involved the following steps, repeated for each of the 19 CMA study areas:

1. Calculate overall mean wind speed, rounded to 1000s meters, for all years: D
2. Calculate monthly mean wind speeds, rounded to 1000s meters, for all months: D_m
3. Transform the wind direction from degrees to radians: θ
4. Calculate the sine of all angles: $S = \sin\theta$
5. Calculate the cosine of all angles: $C = \cos\theta$
6. Summarize by CMA, year, and month to calculate mean of sin and cos: $\overline{S}$ and $\overline{C}$
7. Calculate the resultant: $R = \tan^{-1}(\overline{S} / \overline{C})$
8. Apply the appropriate calculation based on the mathematical quadrant for $\overline{S}$ and $\overline{C}$, for the angular mean, A_m:
    a. Convert R to degrees, if $\overline{S} > 0$ and $\overline{C} > 0$
    b. Convert $R + \pi$ to degrees, if $\overline{C} < 0$
    c. Convert $R + 2\pi$ to degrees, if $\overline{S} < 0$ and $\overline{C} > 0$
9. Transform to mathematical angle using the modulo operator: $mA_m = (450 - A_m) \% 360$
10. Calculate the standard deviation: $A_{sd} = \sqrt{-2 \times \log(1 - (1 - R))}$
11. Calculate the start angle by subtracting a standard deviation: $A_{start} = mA_m - A_{sd}$
12. Calculate the end angle by adding a standard deviation: $A_{end} = mA_m + A_{sd}$

The values needed for computing the wind-dispersed pollutants described below, were overall average speed (for the radius distance), monthly average speed (i.e. distance), mathematical angular mean, start angle, and end angle (D, D_m, As_{tart}, and A_{end}).

Additional covariates were low socioeconomic status (SES) from Chan et al. [103], the total number of infants during the study from Census Canada 2011 (detailed below), and NO_2 emissions from Hystad et al.'s land use regression (LUR) modelling [185].

Geolocation

I assigned longitude and latitude to the CNN records by joining the postal codes to the DMTI Spatial's Platinum Postal Code Suite [110]. To ensure static locations throughout the study period, I uniquely selected postal codes from 2001 through 2013 (the time span was necessary due to the addition and retirement of postal codes during the study).

I also geolocated the NPRI facilities by joining the unique facility identifiers to the provided table of longitude and latitudes.

My space-time analyses required distance-preserving spatial references; therefore, I customized Azimuthal Equidistant map projections centered on each CMA (detailed in Appendix II: Table S2.1) and implemented in Esri's ArcGIS software for all my spatial analyses.

Wind-dispersed pollutants

For each CMA, I used the mean wind speed and direction to specify kernel density and focal wedge-shaped parameters (described above in the section on independent variables and shown in Figure 5.3).

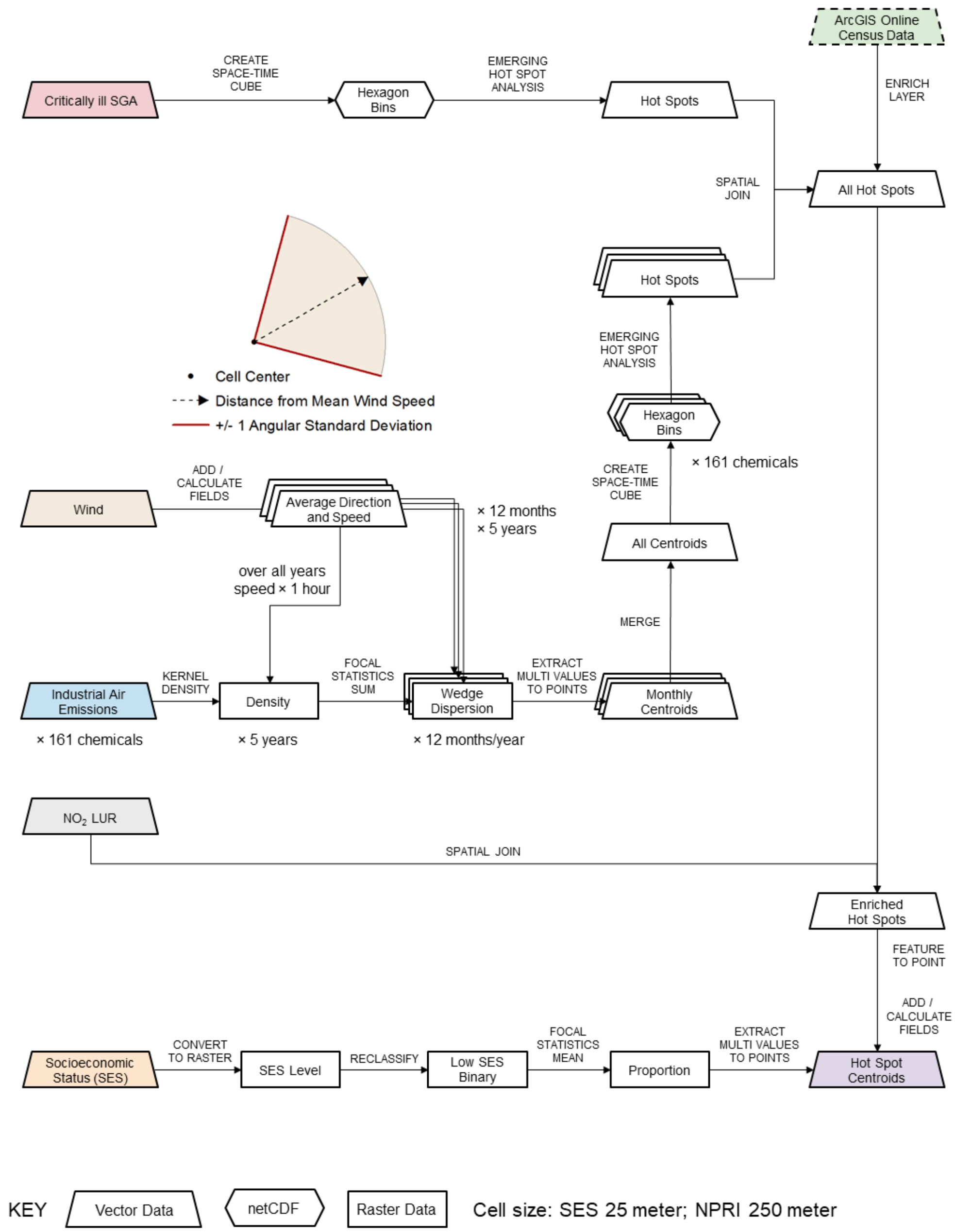


Figure 5.3. Flow chart of GIS commands for analyzing critically ill small newborns in space and time, plus detail on wind-dispersed air pollutants using kernel density and focal sum wedge modelling. The analyses were replicated for 19 Census Metropolitan Areas (CMAs) using customized azimuthal equidistant map projections for each.

I established the base emissions as tonnes per square kilometer, by applying kernel density. Kernel density is a non-parametric method that extends point source values across a surface by calculating the magnitude-per-unit area, fitted to a smoothly tapered function that spreads the values within a specified distance around each point [108]. I set the magnitude as the mean monthly NPRI emissions, in tonnes (i.e. the annual total divided by 12 months), and the radius distances were the overall mean wind speed (described above), with a 250 m raster cell size. These surfaces became the input to the corresponding monthly calculations using focal statistics, which is a moving-window or "filtering" operation that computes statistics (e.g., sum) on values encountered in the neighborhood.

I applied focal statistics using a wedge, which is a pie-shaped neighborhood that I defined as ±1 standard deviation on either side of the mean monthly wind direction, extending the radius distance by the mean monthly wind speed. The wedge defined by 1 standard deviation followed dispersion modelling work by Qiu et al. [186]. To estimate the relative cumulative exposure, I calculated the focal sum as the statistic [187]. Python 2.7 in ArcGIS automated the iterative process for each CMA, year, month, and NPRI chemical combination.

Hexagon centroids from the space-time pattern mining (described next) were used to extract the modelled pollution values for each month. I merged the monthly data for each chemical into complete temporal files matching the CNN study period and CMA to then be input to their own space-time pattern mining.

Spatiotemporal patterns

I analyzed the spatial and temporal clustering of ciSGA distributions using the ArcGIS space-time pattern mining tools [171]. I transformed the postal codes of each CMA into multidimensional data cubes, known as network Common Data Form (netCDF) files, through aggregation spatially by 1-km high hexagon bins and temporally by 1-month birthdate time slices; summing the binary values of ciSGA; filling empty bins with zeros; and aligning to a reference time equal to the beginning of the study (2007/01/01 for Ottawa, 2008/01/01 for Edmonton, and 2006/01/01 for all other CMA). The Mann-Kendall statistic evaluated the trend in ciSGA point counts for each data cube.

The hexagon was chosen because it is more natural in shape, better represents connectivity, and minimizes edge effects [172]; the 1-km size fit within typical city neighborhoods and helped protect individual privacy. The 1-month time-step interval fit within a trimester. Bins were filled with zeros because ciSGA are considered rare events, counted in whole numbers, and therefore interpolation would not be appropriate.

Emerging hot spot analysis (EHSA) analyzed each data cube by calculating statistically significant hot and cold spot trends of ciSGA locations over the entire time period using two statistics. The Getis-Ord Gi* statistic assessed the location and degree of spatial clustering by calculating the z-score, p-value, and hot spot bin classification. The Mann-Kendall statistic evaluated these measures to assess temporal trends and then categorized locations as hot or cold spots. To simulate city neighborhood sizes, I used a fixed distance of 2001 m (note: the additional 1 m ensured that complete hexagons were included), which encompassed the current hexagon and 2 adjacent hexagons (i.e. 2.5 km). To simulate a trimester, I used 2-time steps, which included the current month and previous two months (i.e. 3 months). Hot spot maps representing the entire time period were produced to visualize the spatiotemporal significance of ciSGA in each CMA.

I applied the same space-time parameters to the merged files of the NPRI chemicals. For each CMA, I first created a base hexagon grid that matched the ciSGA hot spot maps, and then used the centroids to extract the dispersed chemicals, by year and month. I transformed the centroids into 1-km hexagons by 1-month space-time cubes and applied the neighbor parameters for EHSA relevant to each CMA. This was automated in the GIS by Python [111]. The resulting hot spot maps represented the entire time period for each chemical.

All EHSA maps were reclassified to binary values, where 1=hot spot and 0=not.

Demographic enrichment

To include a proxy of the total number of births, I assigned to the hexagons the total number of infants for the entire 2006-2010 study period. The ArcGIS Online services of the Enrich Layer tool provided the males and females aged 0 to 4 years old from the 2011 Canadian Census. [188]. I summed these to obtain total infants for each hexagon.

Statistical analyses

For each CMA, I spatially joined all resulting hot spot maps. The original hot spot categories were reclassified as binary values; 1=hot; 0=not. Logistic regression was applied using the ciSGA binary hot spots as the dependent variable and each NPRI chemical binary hot spot as the independent variable. Because I was interested only in the significance of the effect of one independent variable (X) on the response (Y), and there was no need of interpreting the coefficients, the coefficients were calculated (i.e. logarithm of the odds ratios).

I included the proportion of low SES because it is associated with adverse birth outcomes. To account for major traffic-related pollution, I incorporated average values from the LUR model. Since I did not have total number of births at the postal code level, I included total births based on the enriched hexagons as a covariate; there may have not been hot spots for some of the hexagons but there may still have been births potentially exposed that I needed to control for. It is important to note that it was the hot spots of ciSGA and hot spots of chemical emissions that were associated – there was no assignment of the potential exposure to individual locations because EHSA aggregated the data in space and time. I used STATA 15 statistical software [102].

5.4 Results

Characteristics

Critically ill SGA – The nineteen CMAs varied in the raw counts of ciSGA. As shown in Table 5.1, Toronto had the most (n=1,257) and Ottawa/Québec were tied for the least (n=19). The next three highest CMAs were Montréal, Calgary, and Vancouver (n=838, 505, and 496 respectively). The next three lowest were Fredericton, Kingston, and Moncton (n=24, 87, and 92). Percentages of critically ill SGA per total number of admissions in each CMA are shown as green bars in Figure 5.4. Saint John had the most (2.1%) and Ottawa had the least (0.03%). Saskatoon, Halifax, and Regina were among the highest (1.52%, 1.46%, and 1.42%). Québec, Edmonton, and Montréal were among the lowest (0.05%, 0.22%, and 0.38%).

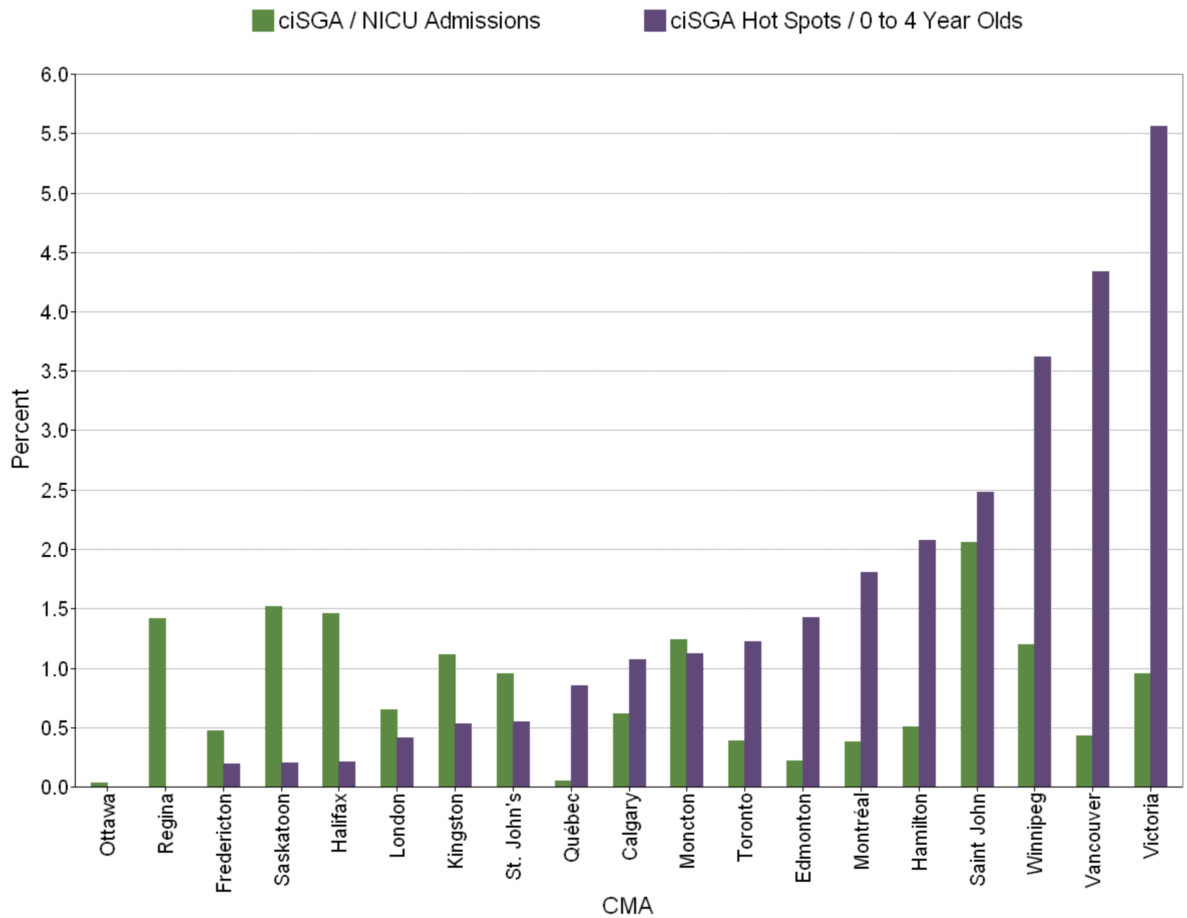


Figure 5.4. Percentages of critically ill small for gestational age (ciSGA): admissions to neonatal intensive care units (NICU) and percentages of space-time hot spot patterns in the 19 Canadian Neonatal Network (CNN) study areas, identified by Census Metropolitan Area (CMA).

NPRI – For comparison with the CMAs, Table 5.2 shows the industrial emissions for each of the nine provinces: Alberta had the greatest number of facilities (n=6,331) and emissions (7,832,008 tonnes) and Ontario had the largest number of chemicals (n=207). Table 5.3 shows that the number of substances released to air within 100 km differed among the 19 CMAs. The number of chemicals ranged from 27 (Saskatoon) to 132 (Toronto), the number of industrial facilities ranged from 16 (St. John’s) to 1,310 (Toronto), and the tonnes emitted ranged from 64,014

(Saskatoon) to 1,472,193 (London). Appendix II: Figures S2.1-S2.4 shows the log of total tonnes for all 161 chemicals emitted by facilities within each CMA. Figure 5.5 shows the distribution of industrial facilities as blue dots for an example CMA: Edmonton (all other maps are in Appendix II: Figure S2.5).

Table 5.2. Provincial-level characteristics of population and the National Pollutant Release Inventory (NPRI).

Province	Area (km^2)	Total Population	Infants 0 to 4 Years	Number of Facilities	Number of Chemicals	Emissions (tonnes)
Alberta	661,848	3,790,191	250,316	6,331	143	7,832,008
British Columbia	944,735	4,499,139	222,822	1,217	131	3,134,962
Manitoba	647,797	1,233,728	78,182	193	76	2,103,010
New Brunswick	72,908	755,530	36,484	105	80	645,999
Newfoundland	405,212	525,037	24,802	85	63	567,807
Nova Scotia	55,284	944,469	44,296	115	86	1,012,630
Ontario	1,076,395	13,263,544	718,240	2,603	207	4,434,770
Québec	1,542,056	8,007,656	436,844	928	163	4,792,390
Saskatchewan	651,036	1,066,349	70,119	1,350	88	1,750,251

Table 5.3. Census Metropolitan Area (CMA) characteristics of the National Pollutant Release Inventory, mean wind speed values from Environment and Climate Change Canada weather stations (centrally located in each CMA), and the number of maps of monthly wind-dispersed emissions. Individual chemical amounts are shown in Appendix II: Figures S2.1-S2.4. Table continues.

CMA	Number of Facilities	Number of Chemicals	Emissions (tonnes)	Mean Wind Speed (m/h)	Number of Monthly Maps
Calgary	996	51	961,989	15,000	2,556
Edmonton	1,063	73	1,121,450	13,000	3,948
Fredericton	67	60	248,705	12,000	2,976
Halifax	87	64	455,116	12,000	3,108
Hamilton	1,274	130	1,262,835	16,000	5,952
Kingston	154	63	153,592	16,000	2,568
London	557	120	1,472,193	14,000	5,964
Moncton	78	64	213,172	18,000	3,048

CMA	Number of Facilities	Number of Chemicals	Emissions (tonnes)	Mean Wind Speed (m/h)	Number of Monthly Maps
Montréal	495	112	1,044,608	7,000	5,580
Ottawa	215	79	175,350	11,000	3,996
Québec	164	73	766,218	10,000	3,948
Regina	77	47	134,173	19,000	2,184
Saint John	77	61	241,930	18,000	3,084
Saskatoon	70	27	64,014	16,000	1,392
St. John's	16	34	152,558	16,000	1,776
Toronto	1,310	132	1,278,689	17,000	6,000
Vancouver	224	64	203,251	6,000	3,024
Victoria	203	61	184,631	11,000	2,964
Winnipeg	101	49	90,314	8,000	2,208

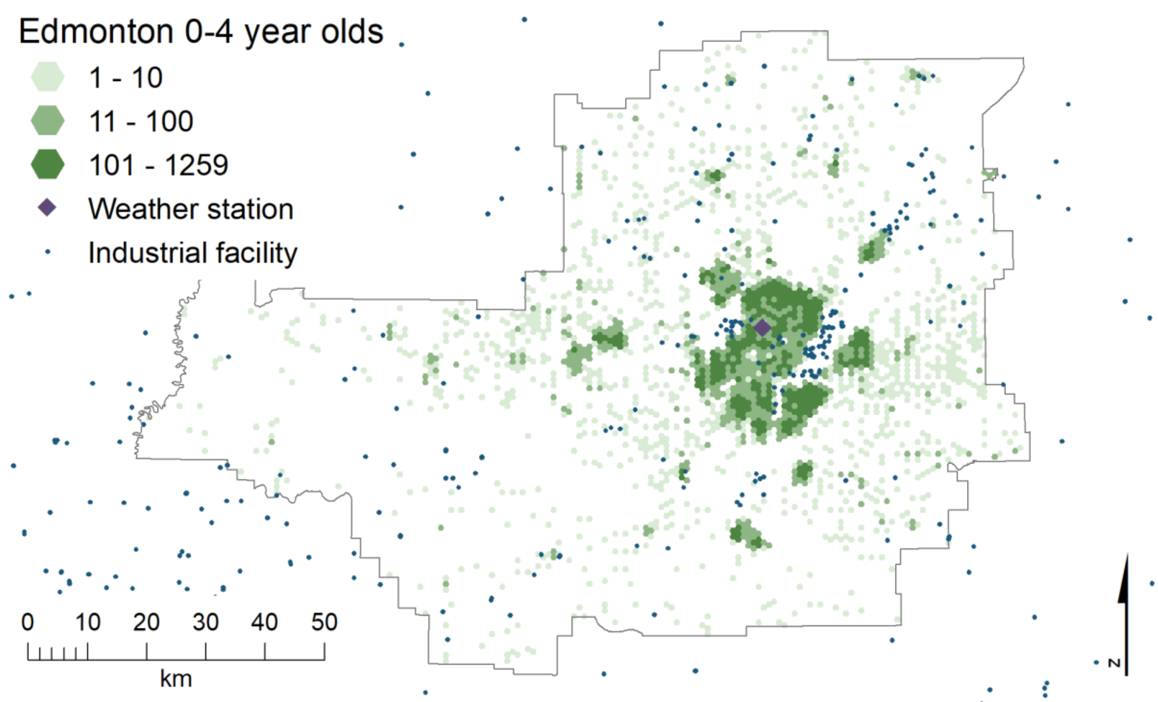


Figure 5.5. The Edmonton Census Metropolitan Area (CMA) is shown as an example distribution of infants during the study period, location of the Environment Canada weather station that provided the wind measurements, and the National Pollutant Release Inventory (NPRI) point sources of the industrial facilities that emitted chemicals to the air. All CMAs are in Appendix II: Figure S2.5.

Wind – Table 5.3 shows that the mean wind speed ranged from 6,000 m/h (Vancouver) to 19,000 m/h (Regina). The speed values were multiplied by one hour to convert to distances for the dispersion mapping.

Dispersion mapping

A total of 66,276 maps were generated for the 19 CMAs, by 5 years × 12 months × number of chemicals (map counts are shown in Table 5.3). The number of chemicals varied slightly from year to year within the same CMAs and the mean monthly wind direction dispersed emissions to accumulate in different areas. Figure 5.6 shows an example of a chemical common to all CMAs, for two time periods: methanol dispersion in Edmonton for January and July 2010.

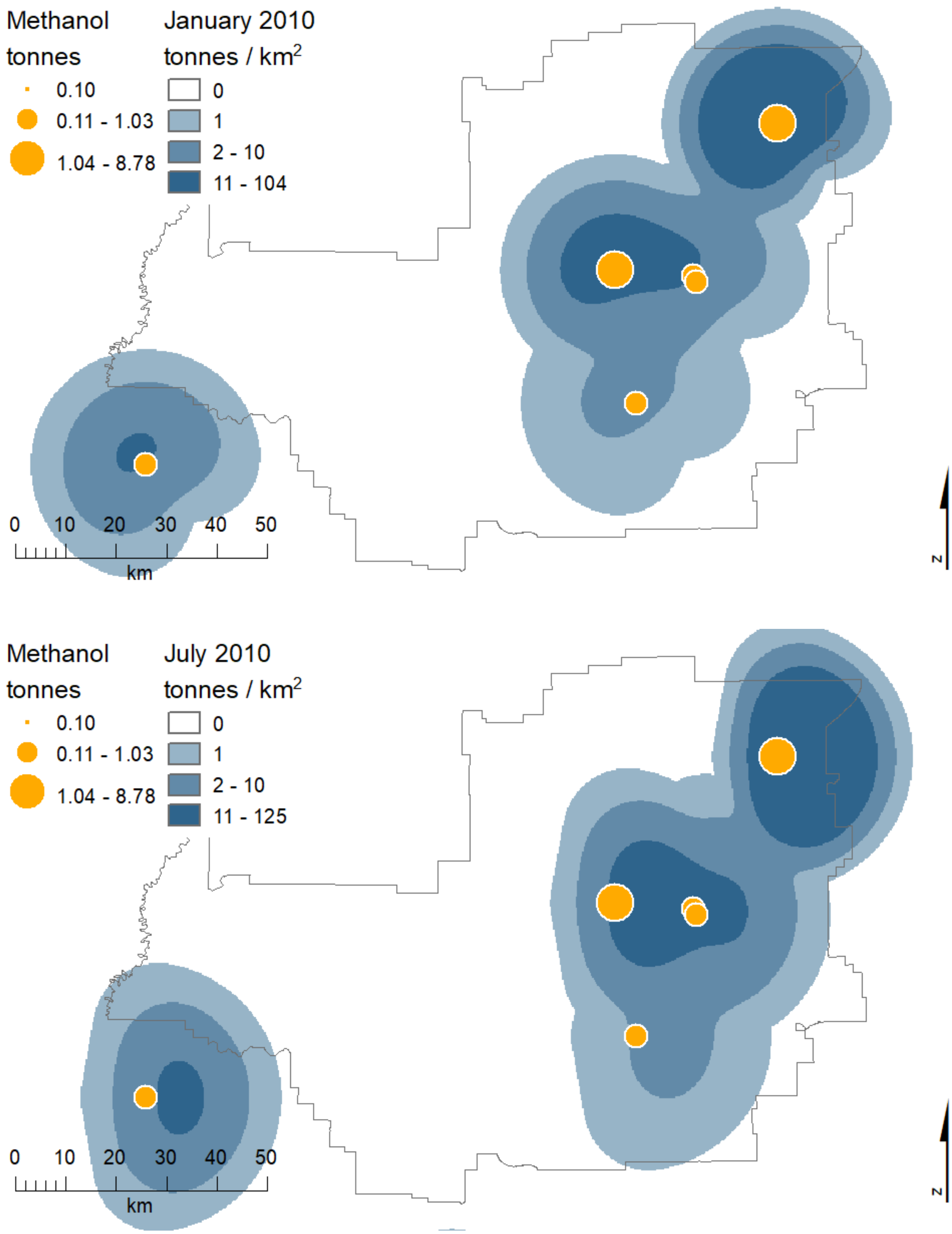


Figure 5.6. Example of monthly wind dispersed pollutant – methanol – January and July 2010, Edmonton CMA.

Space-time cube trends and emerging hot spot patterns

Spatiotemporal trends for ciSGA were increasing for seven CMAs, decreasing for one CMA (Regina), and no statistically significant change for 11 CMAs, as shown in Table 5.4. The raw counts of hot spot locations ranged from the lowest in Ottawa and Regina (n=0) to the most in Vancouver (n=66). Standardized by population, the percentage of ciSGA hot spots per total number of infants is shown in Figure 5.4 as purple bars (Percent Hot Spots = Number of Hot Spots ÷ Number of Hexagons with 0-4 Year Olds × 100%). For those CMAs having hot spots, the lowest percentages were Fredericton, Saskatoon, and Halifax (0.19%, 0.20%, and 0.21%) and the highest were Victoria, Vancouver, and Winnipeg (5.56%, 4.34%, and 3.62%).

Figure 5.7 shows an example hot spot map for the Edmonton CMA. Appendix II: Figure S2.5 has all CMAs. Hamilton, Moncton, Québec, Victoria, and Winnipeg were the only CMAs to have one large central core. The others had smaller, scattered space-time patterns.

The variations in spatiotemporal trends and hot spots of the chemical emissions are too numerous to describe for all CMA and chemical combinations. Figure 5.8 shows an example of methanol emission hot spots in Edmonton. The distributions were less patchy than the ciSGA because the emission point sources were fewer and more localized than residential postal codes.

Table 5.4. Space-time trends and emerging hot spots across Canadian Neonatal Network (CNN) critically ill small for gestational age (ciSGA) newborns, by Census Metropolitan Area (CMA). Note: ↑=increasing; ↓=decreasing; ns=not significant trend.

CMA	Trend	Mann-Kendall	p	Hexagons with Hot Spots	Hexagons with NICU Admissions	Hexagons with 0-4 Year Olds
Calgary	↑	3.65	0.00	12	568	1,117
Edmonton	↑	3.71	0.00	28	443	1,960
Fredericton	↑	2.67	0.01	1	91	533
Halifax	ns	0.12	0.91	2	434	936
Hamilton	ns	1.25	0.21	17	317	816
Kingston	ns	-0.09	0.93	2	93	378
London	ns	1.61	0.11	3	241	729
Moncton	ns	0.39	0.70	6	206	538
Montréal	↑	5.53	0.00	47	1,149	2,603
Ottawa	ns	1.59	0.11	0	61	1,842
Québec	↑	1.68	0.09	8	123	944
Regina	↓	-3.00	0.00	0	131	265
Saint John	ns	0.30	0.77	16	220	646
Saskatoon	ns	0.62	0.54	1	144	500
St. John's	↑	1.83	0.07	2	184	362
Toronto	ns	0.84	0.40	42	1,549	3,442
Vancouver	ns	1.20	0.23	66	790	1,519
Victoria	↑	1.77	0.08	28	240	504
Winnipeg	ns	1.35	0.18	41	369	1,133

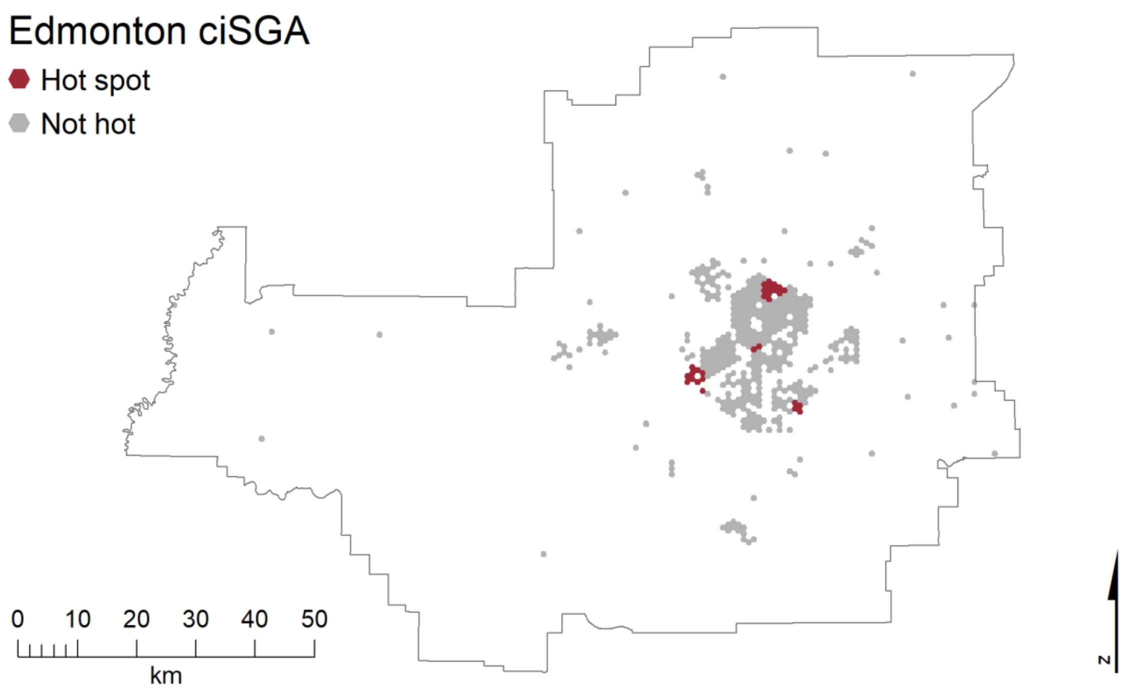


Figure 5.7. Emerging hot spots of critically ill small for gestational age (ciSGA), represented as binary values (red=hot spot; grey=not hot), in the Edmonton CMA.

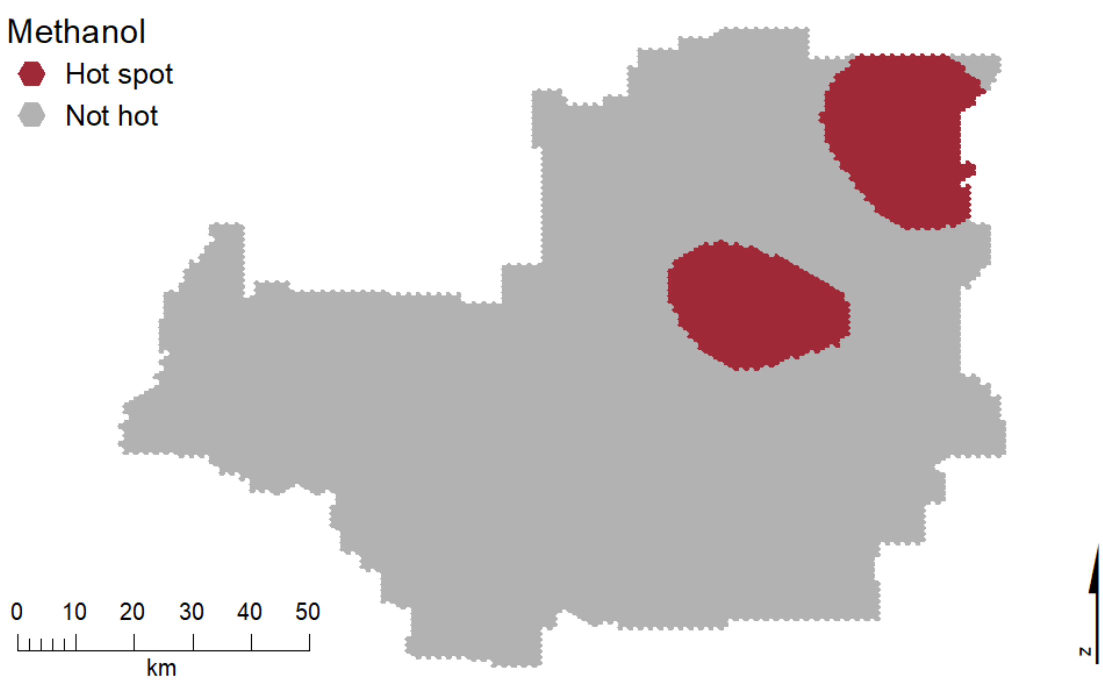


Figure 5.8. Example of emerging hot spots of methanol, represented as binary, in the Edmonton CMA.

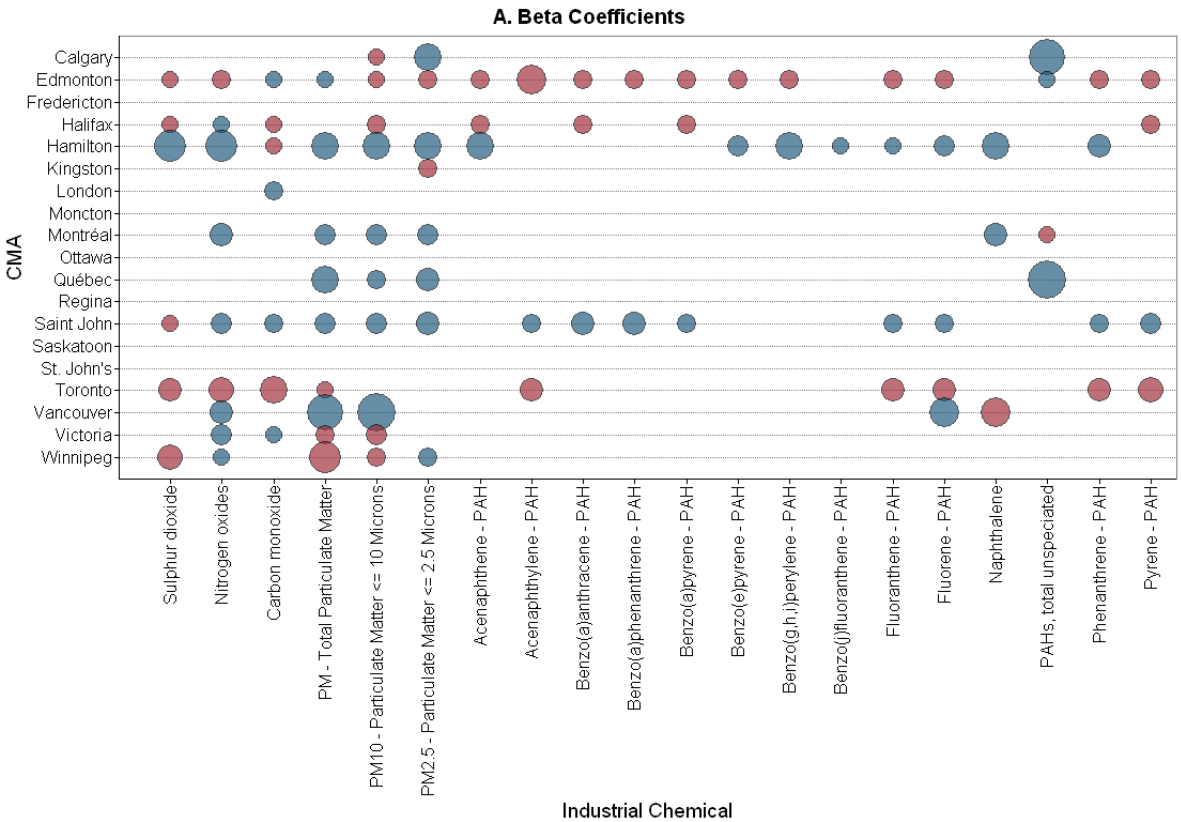
Pattern associations between ciSGA and NPRI

Out of the total 161 relevant NPRI chemical emissions impacting all CMAs (see Appendix II: Figures S2.1-S2.4 for the chemical names), 78 exhibited associations with ciSGA and the covariates low SES, total number of infants, and NO_2 LUR. The logistic regression coefficients for each chemical, by CMA, are shown by relative sizes in Figure 5.9. Note that the blue symbols represent negative coefficients and the red symbols represent positive.

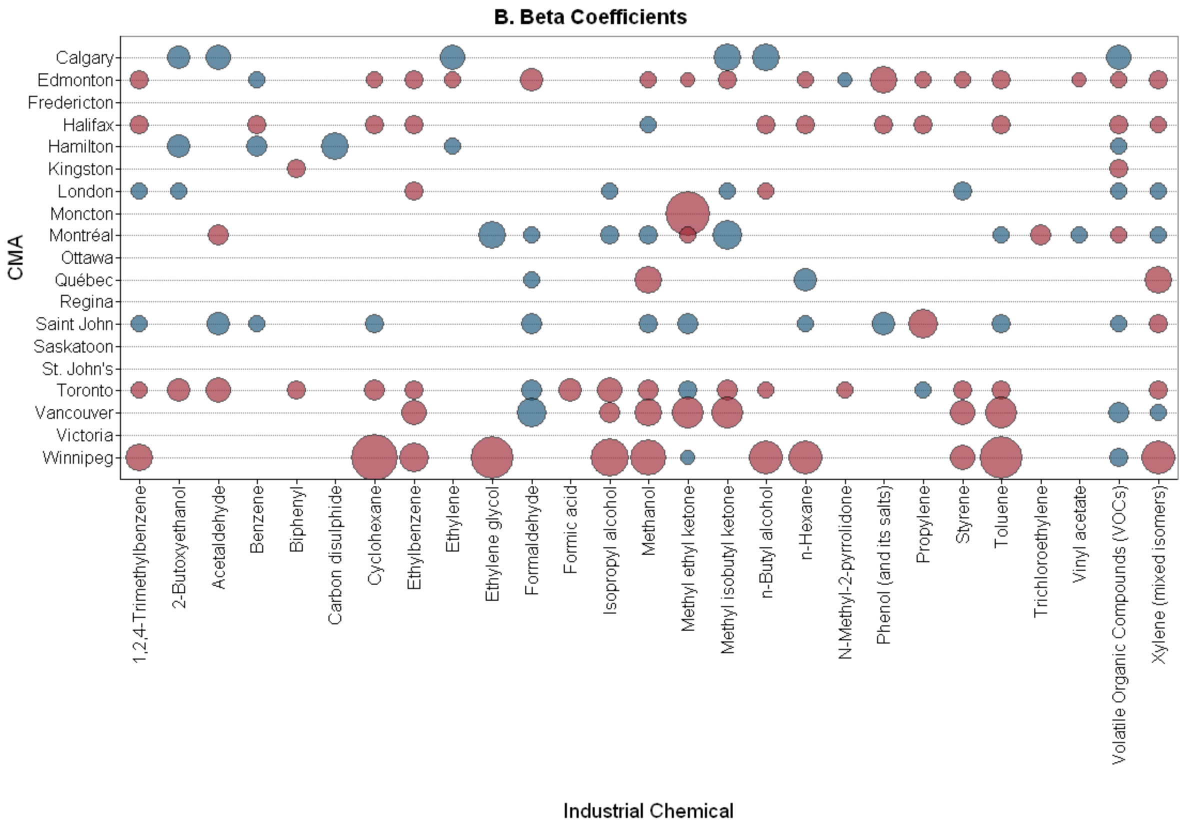
No associations between ciSGA and emissions were observed in Fredericton, Ottawa, Regina, Saskatoon, or St. John's. Either because there were no hot spots for ciSGA or for the chemicals. Logistic regression coefficients for all chemical associations with ciSGA in Calgary, Edmonton, Halifax, Hamilton, Kingston, London, Moncton, Montréal, Québec, Saint John, Toronto, Vancouver, Victoria, and Winnipeg ranged from β=-4.14 to β=6.133.

By *number of chemicals*, Moncton (n=1) exhibited the fewest positive associations with any chemical hot spots and Edmonton (n=46) exhibited the most. In descending order, the other CMAs exhibiting positive hot spot associations were Toronto (n=40), Halifax (n=26), Winnipeg (n=20), Montréal (n=8), and Vancouver (n=8). The other CMAs ranged from one to five positive chemical associations. Where there were negative associations (β<0 indicated by blue symbols in Figure 5.9 A-C), it did not mean protective, but rather the chemical hot spots did not coincide with ciSGA hot spots.

A. Large volume gases, particulate matter, and polycyclic aromatic hydrocarbons (PAHs)



B. Volatile Organic Compounds (VOCs)



C. Heavy metals and other organics/inorganics

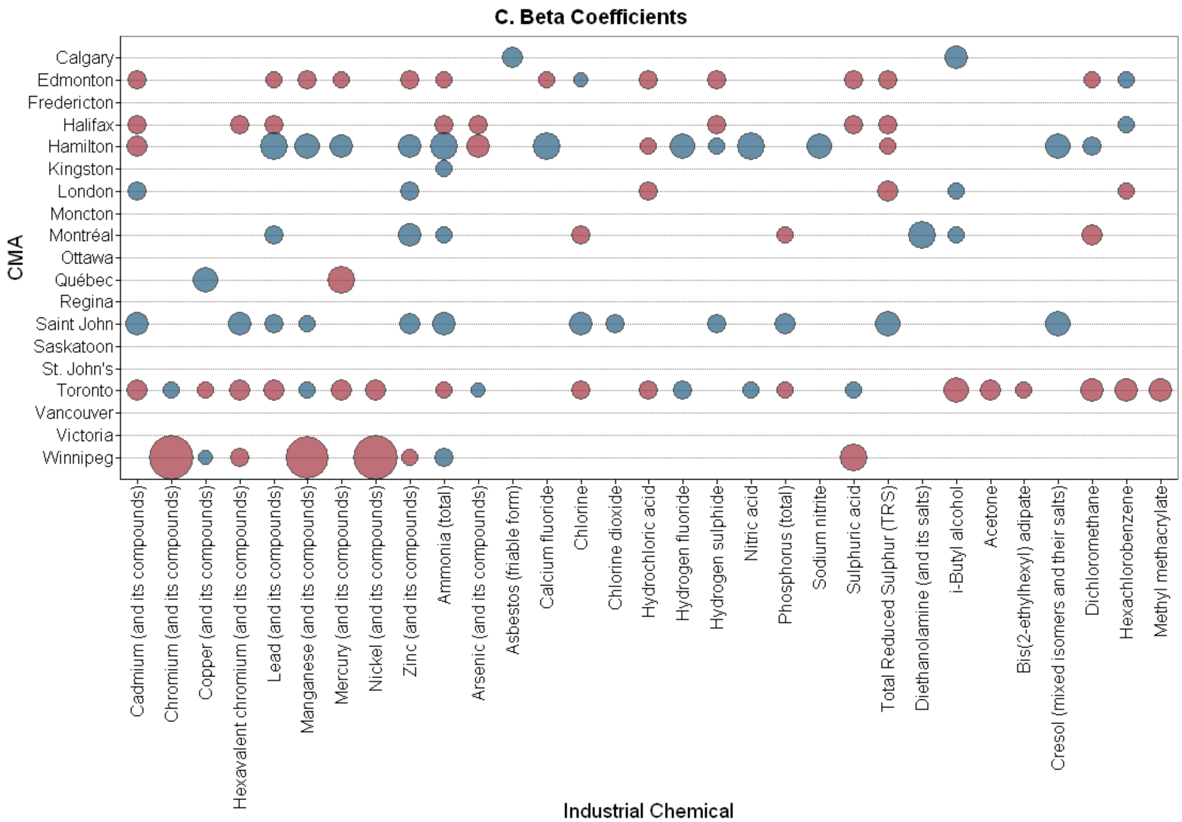


Figure 5.9. Graduated symbols represent the logistic regression coefficients for critically ill small for gestational age (ciSGA) binary hot spots and each industrial air pollutant, modelled with proportions of surrounding low socioeconomic status, total number of infants, and NO_2 pollution from land use regression. The size of the 'bubble' represents the strength of the coefficient and the color indicates the direction: blue is negative, and red is positive. Chemicals are alphabetized within groupings of: (A) large volume gases, particulate matter, and polycyclic aromatic hydrocarbons; (B) volatile organic compounds; and (C) heavy metals and other organics/inorganics.

By *number of CMAs*, there were 28 chemicals having positive associations with ciSGA in three or more CMAs: 1,2,4-Trimethylbenzene, Ammonia (total), Cadmium (and its compounds), Carbon monoxide, Cyclohexane, Dichloromethane, Ethylbenzene, Hexavalent chromium (and its compounds), Hydrochloric acid, Isopropyl alcohol, Lead (and its compounds), Mercury (and its compounds), Methanol, Methyl ethyl ketone, Methyl isobutyl ketone, n-Butyl alcohol, n-Hexane, PM - Total Particulate Matter, PM_{10} - Particulate Matter ≤ 10 Microns, Propylene, Pyrene, Styrene, Sulphur dioxide, Sulphuric acid, Toluene, Total Reduced Sulphur (TRS), Volatile Organic Compounds (VOCs), and Xylene (mixed isomers). Table 5.5 shows that 23 associations were statistically significant, and highlights 21 chemicals that are known or suspected developmental toxicants according to the Canadian Environmental Protection Act [189], California's Office of Environmental Health Hazard Assessment [116], and/or GoodGuide's Scorecard [117].

Table 5.5. List of 28 chemicals having hot spot associations with ciSGA in three or more Census Metropolitan Areas (CMAs); * identifies the 23 statistically significant chemicals (p<0.05). Of these, there are 21 known or suspected developmental toxicants according to the Canadian Environmental Protection Act (CEPA), California's Office of Environmental Health Hazard Assessment (OEHHA), or GoodGuide's Scorecard. Appendix II: Figure S2.6 graphically shows the beta coefficients for the list. Table continues.

Chemical	**Number of CMAs**	**Class**	**CEPA** [189]	**OEHHA** [116]	**Scorecard** [117]
1,2,4-Trimethylbenzene*	4	Organics			
Ammonia (total)	3	Inorganics	listed		
Cadmium (and its compounds)*	4	Heavy Metals	listed	developmental	recognized
Carbon monoxide*	3	Gases	listed	developmental	recognized
Cyclohexane*	4	Organics			
Dichloromethane*	3	Organics	listed		
Ethylbenzene*	6	Organics			suspected
Hexavalent chromium (and its compounds)*	3	Heavy Metals	listed		
Hydrochloric acid	4	Inorganics			
Isopropyl alcohol*	3	Organics			suspected
Lead (and its compounds)*	3	Heavy Metals	listed	developmental	recognized
Mercury (and its compounds)*	3	Heavy Metals	listed	developmental	recognized
Methanol*	5	Organics		developmental	suspected
Methyl ethyl ketone*	4	Organics			suspected
Methyl isobutyl ketone*	3	Organics		developmental	suspected
n-Butyl alcohol*	4	Organics			
n-Hexane*	3	Organics		developmental	suspected
PM - Total Particulate Matter*	3	PMs	listed		suspected
PM_{10} - Particulate Matter ≤ 10 Microns	5	PMs	listed		suspected

Chemical	Number of CMAs	Class	CEPA [189]	OEHHA [116]	Scorecard [117]
Propylene*	3	Organics			
Pyrene*	3	PAHs			
Styrene*	4	Organics			suspected
Sulphur dioxide*	5	Gases	listed	developmental	suspected
Sulphuric acid*	3	Inorganics	listed		
Toluene*	5	Organics		developmental	recognized
Total Reduced Sulphur (TRS)	4	Inorganics			
Volatile Organic Compounds (VOCs)	4	Organics	listed		suspected
Xylene (mixed isomers)*	6	Organics			suspected

5.5 Discussion

I discovered hot spot patterns of ciSGA spatially related to those of the industrial air pollutants. Each of the CMAs had different chemicals that were statistically associated. My findings on space-time relationships for 28 chemicals with ciSGA may be good candidates for future toxicity research to investigate causal relationships. I found that 21 of those are known or suspected developmental toxicants and I encourage health and environmental policy makers to continue to prioritize them for emission reductions.

The space-time hot spots of ciSGA for the nineteen major metropolitan areas across Canada showed distinct patterns. Most CMAs with higher populations had larger hot spots, but the majority had several smaller, scattered hot spots. All patterns may be useful for knowing where to target public health interventions.

My industrial emission estimations resulted in many sporadic hot spots, which is logical considering that the average wind speeds and directions varied widely throughout the seasons. The space-time hot spots for industrial chemicals became unique signatures for each CMA and the numbers and types of chemicals were not the same. Because the industrial facilities were localized according to land use, using the wind parameters to estimate dispersion in the spatial modeling helped assign cumulative emissions to the maternal residences within more probable distances and directions from sources. For other study areas that experience large seasonal variability, simple circular neighborhoods would be sufficient for modelling because the standard deviations of the wind direction angles would be very large. The edges of the dispersion patterns would be more constant, resulting in less sporadic patterns.

Hot spots not only show where health events exist in space and time but provide an opportunity to examine their determinants. My focus was on industrial emissions, but it is important to note that not all hot spots were likely related to these. In Chapter 4, I did not find associations between ciSGA hot spots and industrial land use for both Calgary and Edmonton. Here, after incorporating wind parameters, I found that Edmonton ciSGA had associations with 35 chemicals, but Calgary ciSGA only had one. Edmonton, Halifax, Montréal, Toronto, Vancouver, and Winnipeg exhibited many more associations with industrial chemicals than Calgary, Hamilton, Kingston, Moncton, Québec, Saint John, and Victoria. However, since SGA/LBWT are multifactorial, these cities may also have additional environmental variables that should be accounted for. Critically ill SGA in Fredericton, Ottawa, Regina, Saskatoon, and St. John's could not be linked to the air pollution patterns. For those CMAs that actually had space-time patterns but no association with emission hot spots, reflects the multifactorial nature of ciSGA and the inclusion of additional factors may be needed. For those not having any hot spot patterns alternative methods [47, 174] may be needed to identify any potential associations with industrial emissions because there may not be sufficient data for space-time analysis.

Limitations

First, I limited the analysis to inside the CMA boundaries. However, this resulted in the loss of half the CNN records, some due to data entry errors, incomplete or unknown postal codes, but mostly because they were outside of the CMA boundaries. Also, the postal code locations may not be as spatially accurate for the less-urban areas. Second, the NPRI data were reported as annual estimates and not actual measurements, so more rigorous modelling, such as interpolation, could not be applied. The focal summation within the "wind" wedge resulted in relative, and not true estimates of dispersion. Third, the simplicity of using a single weather station for each CMA did not allow for localized spatial variation of the wind data, but most of the CMAs were small enough for this to be acceptable. Also, the monthly averaging was confined to calendar groupings and not necessarily by true seasonal variation for each of the CMAs, which spanned 42.5°N to 54°N latitude. Fourth, the hexagon size is subject to the modifiable areal unit problem [160]. Although the positioning of the hexagon grid may not be optimal for all areas of each CMA, the 1-km dimension was found by experimentation to be appropriate for urban neighborhood analysis.

The CNN data collection years and methods differed among CMAs (e.g. Edmonton began participating in 2008 and only reported critically-ill newborns having gestational ages <33 weeks. Therefore, direct comparisons cannot be made.

The use of areal units of analysis underscores the risk for ecological fallacy inference (aggregation bias) [9]; i.e. it must be remembered that not all births in the postal code areas may have been SGA/LBWT.

My access to the SES and traffic-related NO_2 covariates was limited to one static time period, 2006, which matched conditions at the beginning of the study. More temporally matched data might strengthen the associations.

Depending on alternative objectives (e.g. in epidemiology or planning policy), the reporting of coefficients (log of odds ratios) from the logistic regression model may not be suitable. Odds ratios are more easily interpretable as how much the levels of one variable (X = 1 = exposure) affects Y in relation to a reference for X (i.e. X = 0 = no exposure). The beta coefficients were useful for investigating whether any associations existed. More sophisticated statistical analyses to explore interactions of the environmental variables may be performed in the future.

I did not have individual maternal variables that are important risk factors for adverse birth outcomes. Due to aggregation in space and time, it may be meaningless to incorporate the maternal factors. However, future research should consider them in a more complete assessment of the identified hot spots. Finally, causal relationships could not be determined from the observational study design. Instead I identified where hot spot patterns corresponded in space and time for birth outcomes and industrial pollutants.

Strengths

The spatial and temporal aspects of the analysis accounted for privacy and mobility. I was able to protect individual privacy by binning the locations in both space and time, where the results masked the true locations. The EHSA tool uses the spatial neighbors' parameter, automatically accounting for a mobility area, which is important for investigations of the outdoor environment.

It should be noted that hexagons have less edge effects than squares, and more closely matched the circular neighborhood used for assigning the covariates via focal statistics [172]. By applying focal statistics, I calculated the proportions of the SES categorical covariates surrounding the hot spot locations, which was more plausible than accepting the exact values that coincided with the hexagon centroids.

The user-friendly space-time cube tools allowed for rapid visualization and quantification of areas with statistically significant increasing or decreasing trends of ciSGA. For future work that may inform policy decisions on where to focus on monitoring or mitigating potential risk factors, the spatial and temporal dimensions may be adjusted to address different research questions in other study areas.

Access to existing data filled the void due to unmonitored industrial air pollutants. I applied readily-available GIS tools to estimate wind-dispersed air pollutant models at a temporal resolution matching that of the health outcome data. The NPRI is available for other time periods and all across Canada, and similar registries exist in other countries. Therefore, all spatial analyses are accessible and reproducible, meaning the methods developed here are easily applicable to other studies.

My space-time hot spots exhibited associations with 21 known or suspected developmental toxicants that included already described chemicals, such as particulate matter, carbon monoxide, heavy metals, and VOCs [49, 56, 190–196]. These results continue to support the link between air pollution and adverse birth outcomes.

Conclusion

The mapping of spatiotemporal hot spots for critically ill small for gestational age showed that the patterns differed among major Canadian metropolitan areas. The incorporation of publicly available databases on industrial emissions allowed me to discover space-time pattern associations with critically ill small newborns. The application of space-time cubes and emerging hot spot analysis promises to be useful for assessing patterns of health outcomes and exposures; they were integral to handling multiple study areas over multiple historical time periods in space and time. I anticipate that the results will inform health professionals and policy makers in the study areas for identifying areas and emissions of interest. To my knowledge, this is the first application for investigating space-time patterns of adverse birth outcomes and industrial air pollutants.

Chapter 6 Conclusion

The goal of my research was to better understand the relationships of the outdoor environment and babies born small for gestational age (SGA) and low birth weight at term (LBWT). Because my research was part of the larger DoMiNO project, I focused on babies born too small. Another researcher investigated preterm birth. Stillbirth and mortality numbers were too low to analyze [88].

In an interdisciplinary approach, I determined that potential prenatal exposures to the outdoor built-up, social, and natural environments were associated with patterns in the birth outcomes, singly and combined. I used secondary health databases and publicly-available sources of spatial data on air pollution, human-built environment, and natural factors, and estimated these as maternal exposures that are hypothesized to affect birth weight. I mapped and quantified where and what environmental emissions and other outdoor habitat coincided with adverse birth outcomes using Geographical Information Systems (GIS) and spatial statistics.

As mentioned in the introduction, the importance of understating the relationships between the outdoor environment and birth outcomes was first documented by Hippocrates when he wrote detailed accounts on how "airs, waters, places" mattered in prenatal health [31]. He observed that women became barren where waters were hard, indigestible, and cold. Babies had low birth weight where waters were warm and stagnant. Hippocrates alluded that the nomadic lifestyles in hot winds resulted in women having difficulty conceiving. And he explained that cities exposed to good quality winds and waters had inhabitants who were of "good complexion and blooming," where women were very prolific and had easy deliveries. One could argue that Hippocrates was indeed the originator of the idea of medical geography, but later Greek scholars (especially, Eratosthenes) brought geography and the system for mapping locations to the world.

Hippocrates also informed us that not only must physicians know matters of “airs, waters, places” to not make any mistakes in treating disease, but he also gave us the Hippocratic Oath. The oath includes the promise to either help or do not harm to the patient (“First do no harm” [Latin: *Primum non nocere*]) and protect the privacy of the patient.

These are wise words in the modern age of medical geography, where maps can reveal extremely helpful information, as long as they do not betray individual locations or cause injury. In my observational study, I worked with de-identified records that were stored in a fire-wall protected server. However, the ethics of sharing my dissertation results with the populations and communities at risk must be carefully balanced. It seems that the broader public are unaware that prenatal exposure to ubiquitous environmental pollutants may cause intrauterine growth restriction and developmental delays in their offspring [197]. And they may be further unaware that preventive interventions are possible. Because pregnant women have physiological differences and a greater consumption of air and food, most of which cross the placenta, Knudsen et al. [197] recommend regulation on occupational and other exposures during pregnancy and before reproductive age, as well as providing health information on preventive measures to couples.

Information is power. Maps are information. And maps that indicate where environmental hazards exist open up a whole host of potential solutions, from future research to air pollution monitoring to regulation and policy making. It is my hope that my maps may be useful for paving the way in this regard. The maps in Chapters 2 through 5 highlight areas and patterns of concern. Health professionals may consult them to help determine which of their patients may be at risk. Clinicians and epidemiologists may be guided by them to select optimal participants for a longitudinal study with biomonitoring to get closer to causation. Urban planners may incorporate them in land use zoning to design cities that minimize exposures by ensuring industrial facilities are situated and kept farther away from areas where people live and work.

The reader is reminded that my research does not imply causation. The Bradford Hill's criteria are used to distinguish statistically significant associations from causation [9]; therefore my research only adds evidence supporting that outdoor exposures are more than just spatiotemporally associated with SGA/LBWT. Some strong associations were discovered, especially with the dominance of chemical emissions in chapters 2 and 3, and there is coherence with known traffic pollutants. The variability in the study areas, the reporting of coefficients instead of odds ratios, and the lack of true chemical measurements mean that additional research is needed to support true causal inferences.

The environmental variables representing ambient health hazards were tested for associations, which according to the principle of "First do no harm" may assist further research and potential risk mitigation. A large number of variables can be problematic in terms of multiple comparisons. Methods to adjust for multiple testing were not needed in my research according to Bender et al. [198] because one final confirmatory conclusion was not sought in my research. Individual environmental variables were examined using multiple categorical variables (e.g. area-level SES, land use, and maternal age) as covariates in logistic regression. Comparisons of these controlling factors, which would have required multiple test adjustments, were not evaluated. The environment was the focus.

A noteworthy and prospective contribution is demonstrated in Figures 6.1 and 6.2, which identify gaps in the current National Air Pollutant System (NAPS) that monitors criteria air pollutants [199]. These maps highlight areas where there are more small newborns. I used the double kernel density method, outlined in Chapter 3, to show spatial “hot spots” of where we have greater percentages of SGA/LBWT (calculated as density of SGA/LBWT divided by density of all births, where density represents count per square kilometer). Many areas are surrounded by industrial facilities, and they emit more chemicals than what NAPS monitors. These maps are meant to educate and empower, and future work is needed to disseminate the knowledge.

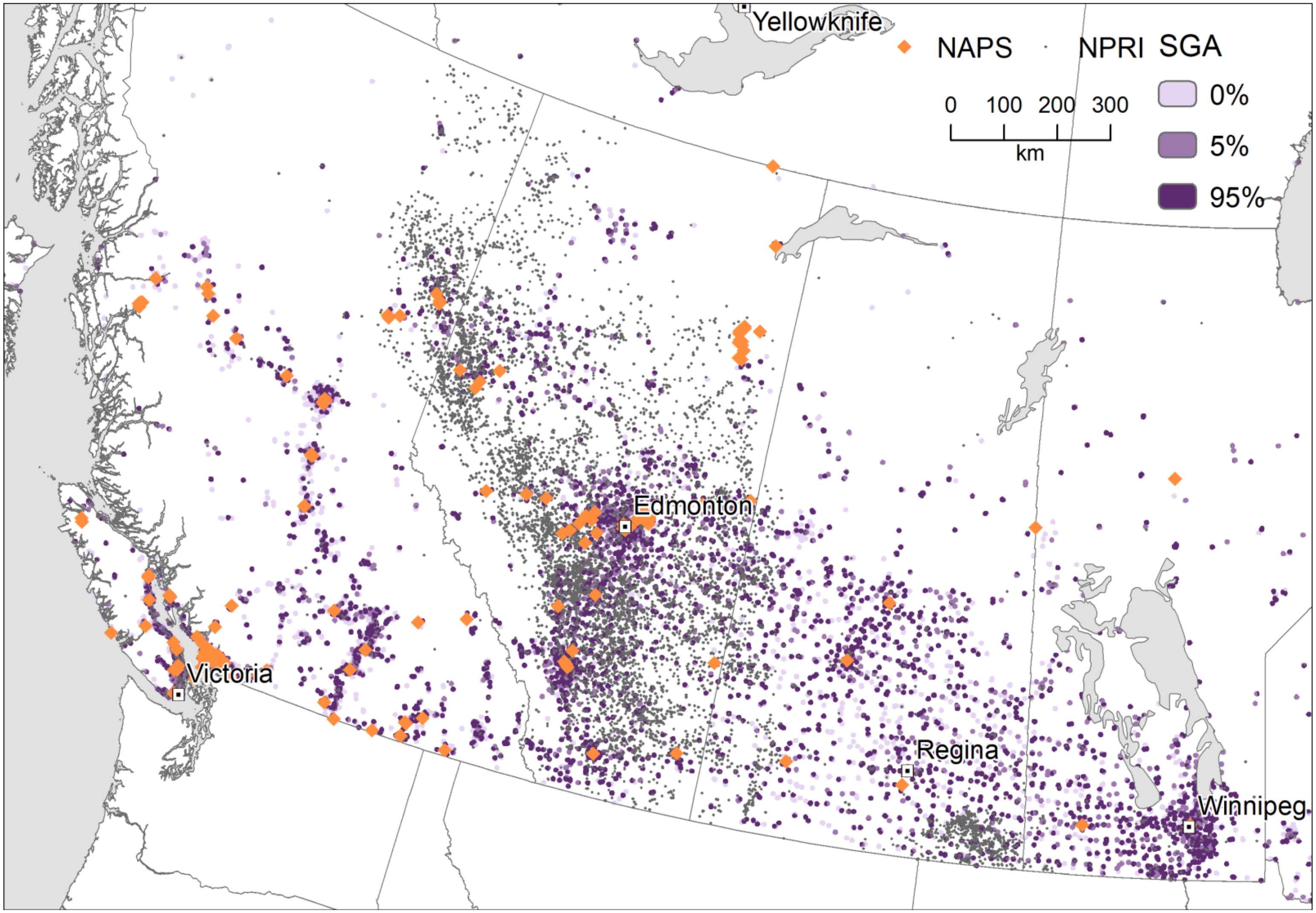


Figure 6.1. Western Canada, 2006-2012, double kernel density map of small for gestational age (SGA). Darker shades of purple indicate areas having relatively more SGA. The overlay of National Pollutant Release Inventory (NPRI=grey) facilities and National Air Pollution Surveillance Program (NAPS=orange) stations helps to identify missing gaps for monitoring.

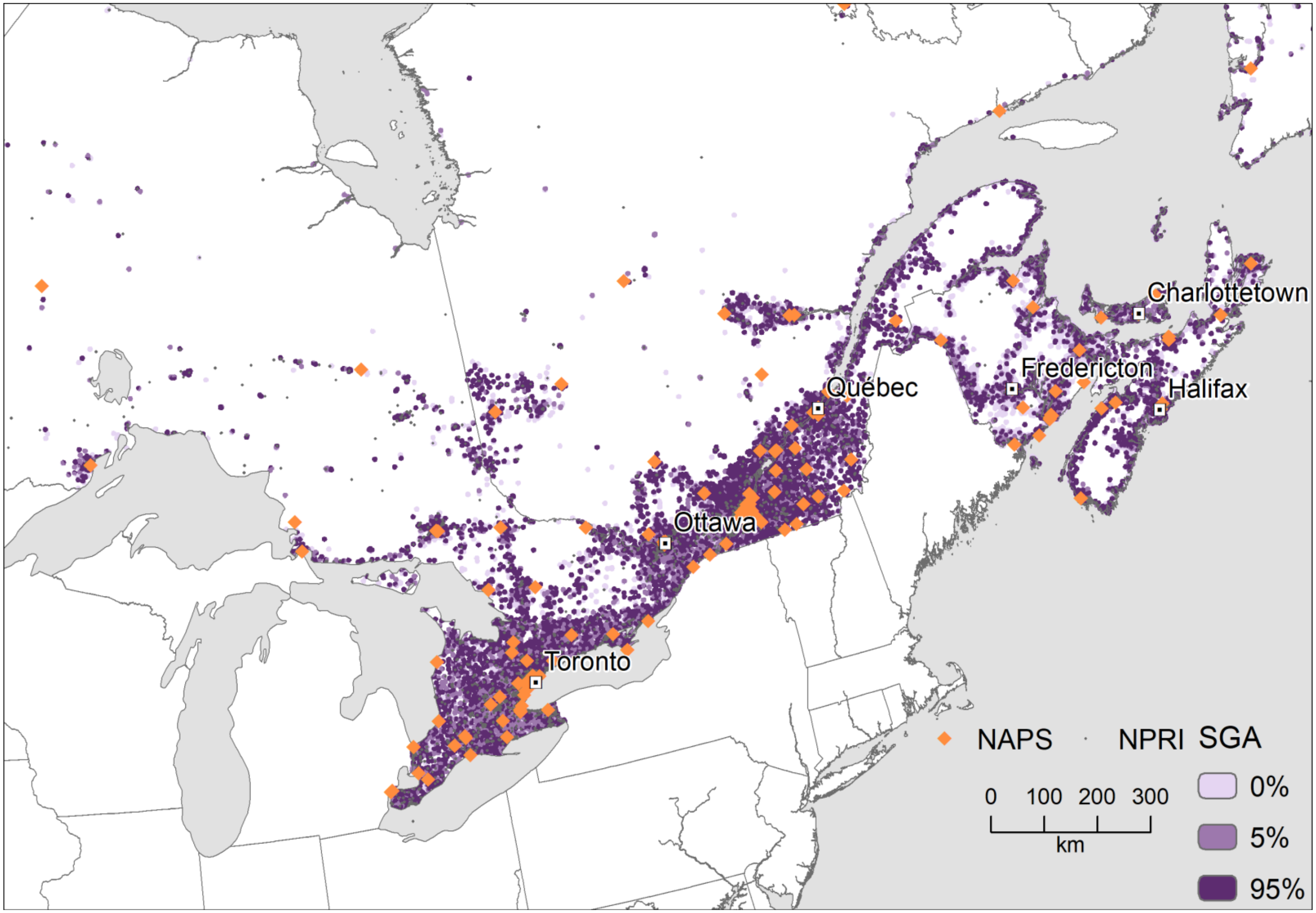


Figure 6.2. Eastern Canada, 2006-2012, double kernel density map of small for gestational age (SGA). Darker shades of purple indicate areas having relatively more SGA. The overlay of National Pollutant Release Inventory (NPRI=grey) facilities and National Air Pollution Surveillance Program (NAPS=orange) stations helps to identify missing gaps for monitoring.

Additional contributions of my dissertation are prevalent in all chapters – I used existing public databases (i.e. the National Pollutant Release Inventory [NPRI]), land-based features, and weather station data were available via public web pages) and easily accessible GIS tools. These may be adapted for future research in understanding neonatal environmental health, as well as addressing any research question involving exposures to outdoor hazards.

Several firsts have evolved from this work as I conducted research from the national to regional scales. In Chapter 2, I offer Canada-wide mapping of SGA, LBWT, and a variety of environmental hazards for 2006-2012. These years align with the national census, deeming it an ideal baseline for future surveillance. It was clear that the province of Alberta had percentages of SGA and LBWT higher than the national average, prompting further investigation. Also, the smaller geographical area and more consistent spatial data of a single province was more amenable for use as a pilot for developing a multi-hazard index, which shows promise for smaller regional analyses. Noting that regional analyses yielded stronger associations and given the availability of the neonatal intensive care database from CNN, in Chapter 4, I applied space-time pattern mining to non-critical and critically ill SGA and LBWT. I discovered that the hot spots do not match and surrounding low SES and industry land use were more strongly associated with SGA and LBWT. Critically ill SGA (ciSGA) and critically ill LBWT (ciLBWT) hot spots were associated with open spaces, but only ciLBWT was associated with industrial land use. To investigate non-static environmental hazard associations with ciSGA, in Chapter 5, I developed a simple method to disperse industrial pollutants from their point sources. This was instrumental in estimating finer temporal resolution from the annual reports – very helpful for space-time analyses. To have stronger findings, ideally there would have been actual measurements from air pollution monitors, as previously discussed, but these maps will still be advantageous for a variety of environmental health research.

From all the chapters, it is evident that the different study areas have their own unique signature of ambient health hazards and pollutants. Not all provinces, territories, metropolitan or administrative units had identical lists of pollutants and hazards. For many study areas there were more associations with land hazards (dumps/waste depots, gas stations, powerlines, and transformer stations). Of the identified industrial chemicals, many are suspected or known developmental toxicants, including ammonia, benzene, carbon monoxide, methyl ethyl ketone, particulate matter, heavy metals, and VOCs. There were some commonalities, but not consistently, supporting the multifactorial nature of SGA and LBWT.

There are more research questions to be addressed, and I recommend a space-time comparison across Canada, especially using more ecological regions. Because the locations of industrial facilities are typically determined by the available natural resources, the use of ecoregions as the analytical unit may help strengthen the statistical modeling.

The importance of the lesser studied habitat vertex (Meade's triangle of human ecology) using spatial analysis has helped reveal outdoor environmental factors associated with the occurrence of underweight neonates in Canada, the province of Alberta, and major cities across the country. It will be challenging, but well worth it, to incorporate the other vertices in the space-time associations, for a more holistic discovery of what outdoor environmental factors are associated with the geography of being born too small. As mentioned earlier, other studies on nutrition, behavior, and pre-existing health conditions are all very important, and gene-environment interactions have been gaining more attention.

It is hoped that this research may assist healthcare givers, in Hippocrates-style, by providing them with what location-based variables may be associated with their patients' health issues, as well as informing the public that where they live is as important to their current and future family health as what they eat and do. My environmental associations were not able to account for nutrition or occupation, but neither have those studies accounted for outdoor exposures. Each contributes pieces of the puzzle. Medical researchers will be provided with more motivation for studying what components of outdoor environmental exposures may cause reduction in neonatal weight, a condition that if prevented will diminish future adverse health, such as adult cardiac disease, diabetes, and other non-communicable diseases that require a strong healthy start in life. Policy makers and planners (health, urban, transportation, industrial) may use this information for mitigating developments to reduce environmental effects on places where pregnant mothers (and everyone else) live. For example, existing land use may need to be altered over time depending on proximity of industrial activities and residential areas.

May this research add to the many needed arguments for reducing the most widespread source of hazardous exposures – outdoor environmental pollution – in the places where one lives and where one starts out in life, to promote a more positive state of human health for all.

"There can be no keener revelation of a society's soul than the way in which it treats its children."

"I dream of our vast deserts, of our forests, of all our great wildernesses. We must never forget that it is our duty to protect this environment."

"Each of us as citizens, has a role to play in creating a better world for our children."

-Nelson Mandela

References

[1] Barker DJ. The fetal and infant origins of adult disease. *BMJ Br Med J* 1990; 301: 1111.

[2] Aschengrau A, Seage III GR. *Essentials of Epidemiology in Public Health*. 3rd ed. Burlington, MA: Jones & Bartlett Learning, 2014.

[3] Backes CH, Nelin T, Gorr MW, et al. Early life exposure to air pollution: How bad is it? *Toxicol Lett* 2013; 216: 47–53.

[4] Stillerman KP, Mattison DR, Giudice LC, et al. Environmental exposures and adverse pregnancy outcomes: a review of the science. *Reprod Sci* 2008; 15: 631–50.

[5] Wigle DT, Arbuckle TE, Turner MC, et al. Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants. *J Toxicol Environ Heal part B, Crit Rev* 2008; 11: 373–517.

[6] Selevan SG, Kimmel CA, Mendola P. Identifying critical windows of exposure for children's health. *Environ Health Perspect* 2000; 108: 451–455.

[7] Canadian Institute for Health Information (CIHI). *Too early, too small: a profile of small babies across Canada*https://secure.cihi.ca/free_products/too_early_too_small_en.pdf (2009).

[8] Lim G, Tracey J, Boom N, et al. CIHI survey: Hospital costs for preterm and small-for-gestational age babies in Canada. *Healthc Q* 2009; 12: 20–24.

[9] Last JM. A Dictionary of Public Health. *Oxford University Press* 2007; 407.

[10] Barker DJP. Maternal nutrition, fetal nutrition, and disease in later life. *Nutrition* 1997; 13: 807–813.

[11] Barker DJP. The developmental origins of chronic adult disease. *Acta Paediatr Suppl* 2004; 93: 26–33.

[12] Barker DJP, Osmond C. Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales. *Lancet* 1986; 327: 1077–1081.

[13] Kramer MS, Platt RW, Wen SW, et al. A New and Improved Population-Based Canadian Reference for Birth Weight for Gestational Age. *Pediatrics* 2001; 108: e1–e7.

[14] Kramer MS. The epidemiology of adverse pregnancy outcomes: an overview. *J Nutr*

2003; 133: 1592S–1596S.

[15] Goldenberg RL, Culhane JF. Low birth weight in the United States. *Am J Clin Nutr* 2007; 85: 584S–90S.

[16] Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull World Health Organ* 1987; 65: 663–737.

[17] Lasker JN, Coyle B, Li K, et al. Assessment of risk factors for low birth weight deliveries. *Health Care Women Int* 2005; 26: 262–80.

[18] Shah PS, Shah V, on behalf of Knowledge Synthesis Group on Determinants of Preterm/LBW Births. Influence of the maternal birth status on offspring: A systematic review and meta-analysis. *Acta Obstet Gynecol Scand* 2009; 88: 1307–1318.

[19] Black RE. *Global Prevalence of Small for Gestational Age Births*. Epub ahead of print 2015. DOI: 10.1159/000365790.

[20] Lee ACC, Katz J, Blencowe H, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Glob Heal* 2013; 1: e26–e36.

[21] Center for Disease Control and Prevention (CDC). QuickStats: Percentage of Small-for-Gestational-Age Births, by Race and Hispanic Ethnicity---United States, 2005. *Morbidity and Mortality Weekly Report 57(50)*https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5750a5.htm (2008).

[22] Center for Disease Control and Prevention (CDC). Singleton Low Birthweight Rates, by Race and Hispanic Origin: United States, 2006–2016. *National Center for Health Statistics (NCHS) - National Vital Statistics System - NCHS Data Brief No. 306*https://www.cdc.gov/nchs/data/databriefs/db306.pdf (2018).

[23] Statistics Canada. Table 102-4318 - Birth-related indicators (low and high birth weight, small and large for gestational age, pre-term births), by sex, three-year average, Canada, provinces, territories, census metropolitan areas and metropolitan influence zones, occasional. *Canadian Socio-Economic Information Management System (CANSIM)*http://www5.statcan.gc.ca/cansim/a05?lang=eng&id=01024318 (2014).

[24] Statistics Canada. Table 102-0562 - Leading causes of death, infants, by sex, Canada, annual, 2006-2012 [digital data]. *Canadian Socio-Economic Information Management System (CANSIM)*http://www5.statcan.gc.ca/cansim/pick-choisir?lang=eng&searchTypeByValue=1&id=1020562 (2012).

[25] APHP. Alberta Perinatal Health Program, 2006-2012. *Information Management and*

*Research - Data Management*http://aphp.dapasoft.com (2014).

[26] Statistics Canada. Census Profile – Comprehensive download files for a selected geographic level: 98-316-XWE2011001-101 [digital data]. *2011 Census Profile*http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/download-telecharger/comprehensive/comp-csv-tab-dwnld-tlchrgr.cfm?Lang=E#tabs2011 (2011).

[27] Meade MS, Emch M. *Medical Geography*. 3rd ed. New York: Guilford Press, 2010.

[28] May JM. The ecology of human disease. *Ann N Y Acad Sci* 1958; 84: 789–794.

[29] Meade MS. Medical geography as human ecology: the dimension of population movement. *Geogr Rev* 1977; 64: 379–393.

[30] World Health Organization (WHO). *Environment and health: the European charter and commentary, WHO Regional Publications, European Series No. 35*. Copenhagenhttp://www.euro.who.int/data/assets/pdf_file/0011/116012/WA3095.pdf (1989).

[31] Hippocrates, Translated by Francis Adams. *Airs, Waters, Places.* London: Printed for the Sydenham societyhttps://archive.org/details/genuineworksofhi01hippuoft (1849).

[32] Cromley EK, McLafferty SL. *GIS and Public Health*. 2nd ed. New York: Guilford Press, 2012.

[33] Nuckols JR, Ward MH, Jarup L. Using Geographic Information Systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect* 2004; 112: 1007–15.

[34] Vine MF, Degnan D, Hanchette C. Geographic Information Systems: their use in environmental epidemiologic research. *Environ Health Perspect* 1997; 105: 598–605.

[35] Kistemann T, Dangendorf F, Schweikart J. New perspectives on the use of Geographical Information Systems (GIS) in environmental health sciences. *Int J Hyg Environ Health* 2002; 205: 169–81.

[36] Jarup L. Health and environment information systems for exposure and disease mapping, and risk assessment. *Environ Health Perspect* 2004; 112: 995–997.

[37] Jerrett M, Gale S, Kontgis C. Spatial modeling in environmental and public health research. *Int J Environ Res Public Health* 2010; 7: 1302–29.

[38] Dent AL, Fowler DA, Kaplan BM, et al. Using GIS to study the health impact of air emissions. *Drug Chem Toxicol* 2000; 23: 161–78.

[39] Meliker JR, Sloan CD. Spatio-temporal epidemiology: principles and opportunities. *Spat Spatiotemporal Epidemiol* 2011; 2: 1–9.

[40] Beale L, Abellan JJ, Hodgson S, et al. Methodologic issues and approaches to spatial epidemiology. *Environ Heal* 2008; 116: 1105–10.

[41] Xiang H, Nuckols JR, Stallones L. A geographic information assessment of birth weight and crop production patterns around mother's residence. *Environ Res* 2000; 82: 160–7.

[42] Ochoa-Acuña H, Frankenberger J, Hahn L, et al. Drinking-water herbicide exposure in Indiana and prevalence of small-for-gestational-age and preterm delivery. *Environ Health Perspect* 2009; 117: 1619–24.

[43] Sathyanarayana S, Basso O, Karr CJ, et al. Maternal pesticide use and birth weight in the agricultural health study. *J Agromedicine* 2010; 15: 127–36.

[44] Weselak M, Arbuckle TE, Foster W. Pesticide exposures and developmental outcomes: the epidemiological evidence. *J Toxicol Environ Heal part B, Crit Rev* 2007; 10: 41–80.

[45] Coker E, Liverani S, Ghosh JK, et al. Multi-pollutant exposure profiles associated with term low birth weight in Los Angeles County. *Environ Int* 2016; 91: 1–13.

[46] Malmqvist E, Liew Z, Källén K, et al. Fetal growth and air pollution - A study on ultrasound and birth measures. *Environ Res* 2017; 152: 73–80.

[47] Svechkina A, Dubnov J, Portnov BA. Environmental risk factors associated with low birth weight: The case study of the Haifa Bay Area in Israel. *Environ Res* 2018; 165: 337–348.

[48] Liu S, Krewski D, Shi Y, et al. Association between maternal exposure to ambient air pollutants during pregnancy and fetal growth restriction. *J Expo Sci Environ Epidemiol* 2007; 17: 426–32.

[49] Stieb DM, Chen L, Hystad P, et al. A national study of the association between traffic-related air pollution and adverse pregnancy outcomes in Canada, 1999-2008. *Environ Res* 2016; 148: 513–526.

[50] Choi H, Rauh V, Garfinkel R, et al. Prenatal exposure to airborne polycyclic aromatic hydrocarbons and risk of intrauterine growth restriction. *Environ Health Perspect* 2008; 116: 658–65.

[51] Huang C, Nichols C, Liu Y, et al. Ambient air pollution and adverse birth outcomes: A natural experiment study. *Popul Health Metr* 2015; 13: 1–7.

[52] Dugandzic R, Dodds L, Stieb D, et al. The association between low level exposures to

ambient air pollution and term low birth weight: a retrospective cohort study. *Environ Heal* 2006; 5: 1–8.

[53] Basu R, Harris M, Sie L, et al. Effects of fine particulate matter and its constituents on low birth weight among full-term infants in California. *Environ Res* 2014; 128: 42–51.

[54] Erickson AC, Ostry A, Chan LHM, et al. The reduction of birth weight by fine particulate matter and its modification by maternal and neighbourhood-level factors: A multilevel analysis in British Columbia, Canada. *Environ Heal* 2016; 15: 1–12.

[55] Harris G, Thompson WD, Fitzgerald E, et al. The association of PM2.5 with full term low birth weight at different spatial scales. *Environ Res* 2014; 134: 427–434.

[56] Stieb DM, Chen L, Beckerman BS, et al. Associations of pregnancy outcomes and PM in a national Canadian study. *Env Heal Perspect* 2015; 124: 243–249.

[57] Wilhelm M, Ghosh JK, Su J, et al. Traffic-related air toxics and term low birth weight in Los Angeles County, California. *Environ Health Perspect* 2012; 120: 132–138.

[58] Liu S, Krewski D, Shi Y, et al. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. *Environ Health Perspect* 2003; 111: 1773–8.

[59] Rich DQ, Liu K, Zhang J, et al. Differences in Birth Weight Associated with the 2008 Beijing Olympic Air Pollution Reduction: Results from a Natural Experiment. *Environ Health Perspect* 2015; 117: 1713–1717.

[60] Miranda ML, Messer LC, Kroeger GL. Associations between the quality of the residential built environment and pregnancy outcomes among women in North Carolina. *Environ Health Perspect* 2012; 120: 471–7.

[61] Woods N, Gilliland J, Seabrook JA. The influence of the built environment on adverse birth outcomes. *J Neonatal Perinatal Med* 2017; 10: 233–248.

[62] Zeka A, Melly SJ, Schwartz J. The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts. *Environ Heal* 2008; 7: 1–12.

[63] Baibergenova A, Kudyakov R, Zdeb M, et al. Low birth weight and residential proximity to PCB-contaminated waste sites. *Environ Heal* 2003; 1352–7.

[64] Goldberg MS, Goulet L, Riberdy H, et al. Low birth weight and preterm births among infants born to women living near a municipal solid waste landfill site in Montreal, Quebec. *Environ Res* 1995; 69: 37–50.

[65] Auger N, Joseph D, Goneau M, et al. The relationship between residential proximity to extremely low frequency power transmission lines and adverse birth outcomes. *J Epidemiol Community Health* 2011; 65: 83–5.

[66] de Vocht F, Hannam K, Baker P, et al. Maternal residential proximity to sources of extremely low frequency electromagnetic fields and adverse birth outcomes in a UK cohort. *Bioelectromagnetics* 2014; 35: 201–209.

[67] Casey JA, Savitz DA, Rasmussen SG, et al. Unconventional natural gas development and birth outcomes in Pennsylvania, USA. *Epidemiology* 2016; 27: 163–172.

[68] Mckenzie LM, Guo R, Witter RZ, et al. Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. *Environ Health Perspect* 2014; Advance publication.

[69] Stacy SL, Brink LAL, Larkin JC, et al. Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania. *PLoS One* 2015; 10: 1–15.

[70] Philion JJ, Schmitt N, Rowe J, et al. Effect of lead on fetal growth in a Canadian smelter city, 1961-1990. *Arch Environ Health* 1990; 52: 472–5.

[71] Rabito FA, Kocak M, Werthmann DW, et al. Changes in low levels of lead over the course of pregnancy and the association with birth outcomes. *Reprod Toxicol* 2014; 50: 138–144.

[72] Ahern M, Mullett M, MacKay K, et al. Residence in coal-mining areas and low-birth-weight outcomes. *Matern Child Health J* 2011; 15: 974–979.

[73] Claus Henn B, Ettinger AS, Hopkins MR, et al. Prenatal arsenic exposure and birth outcomes among a population residing near a mining-related superfund site. *Environ Health Perspect* 2016; 124: 1308–1315.

[74] Hystad P, Davies HW, Frank L, et al. Residential greenness and birth outcomes: evaluating the influence of spatially correlated built-environment factors. *Environ Health Perspect* 2014; 122: 1095–1102.

[75] Holstius DM, Reid CE, Jesdale BM, et al. Birth weight following pregnancy during the 2003 Southern California wildfires. *Environ Health Perspect* 2012; 120: 1340–1345.

[76] Habermann M, Gouveia N. Socioeconomic position and low birth weight among mothers exposed to traffic-related air pollution. *PLoS One* 2014; 9: 1–16.

[77] Généreux M, Auger N, Goneau M, et al. Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *J Epidemiol*

Community Health 2008; 62: 695–700.

[78] Brauer M, Lencar C, Tamburic L, et al. A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ Health Perspect* 2008; 116: 680–6.

[79] Meng G, Thompson ME, Hall GB. Pathways of neighbourhood-level socio-economic determinants of adverse birth outcomes. *Int J Health Geogr* 2013; 12: 1–16.

[80] Wilhelm M, Ritz B. Residential proximity to traffic and adverse birth outcomes in Los Angeles County, California, 1994-1996. *Environ Health Perspect* 2003; 111: 207–16.

[81] Shah PS, Balkhair T. Air pollution and birth outcomes: a systematic review. *Environ Int* 2011; 37: 498–516.

[82] Vadillo-Ortega F, Osornio-Vargas A, Buxton MA, et al. Air pollution, inflammation and preterm birth: a potential mechanistic link. *Med Hypotheses* 2014; 82: 219–24.

[83] Kannan S, Misra DP, Dvonch JT, et al. Exposures to airborne particulate matter and adverse perinatal outcomes: a biologically plausible mechanistic framework for exploring potential effect modification by nutrition. *Environ Health Perspect* 2006; 114: 1636–42.

[84] Bosetti C, Nieuwenhuijsen MJ, Gallus S, et al. Ambient particulate matter and preterm birth or birth weight: a review of the literature. *Arch Toxicol* 2010; 84: 447–60.

[85] Prüss-Üstün A, Wolf J, Corvalán C, et al. *Preventing disease through healthy environments: A global assessment of the burden of disease from environmental risks*. Genevahttp://apps.who.int/iris/bitstream/10665/204585/1/9789241565196_eng.pdf (2016).

[86] Shah PS. Parity and low birth weight and preterm birth: A systematic review and meta-analyses. *Acta Obstet Gynecol Scand* 2010; 89: 862–875.

[87] Tough SC, Svenson LW, Johnston DW, et al. Characteristics of preterm delivery and low birthweight among 113,994 infants in Alberta: 1994-1996. *Can J Public Heal* 2001; 92: 276–280.

[88] Serrano-Lomelin J. *Profiling industrial air-pollutant mixtures and their associations with preterm birth and small for gestational age in Alberta, Canada*. University of Albertahttps://era.library.ualberta.ca/items/3fdd6a2e-454c-406f-a531-cecf487d3700/view/f565dcfb-d046-4f6a-8a9d-af114c660496/Serrano_Jesus_A_201712_PhD.pdf (2017).

[89] Dadvand P, Ostro B, Figueras F, et al. Residential proximity to major roads and term low birth weight: the roles of air pollution, heat, noise, and road-adjacent trees. *Epidemiology*

2014; 25: 518–525.

[90] Fleischer NL, Merialdi M, Donkelaar A Van, et al. Outdoor air pollution, preterm birth, and low birth weight: analysis of the World Health Organization Global Survey on Maternal and Perinatal Health. *Environ Health Perspect* 2014; 122: 425–430.

[91] Jerrett M, Arain A, Kanaroglou P, et al. A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol* 2005; 15: 185–204.

[92] Ha S, Hu H, Roussos-Ross D, et al. The effects of air pollution on adverse birth outcomes. *Environ Res* 2014; 134: 198–204.

[93] Hartig T, Catalano R. Cold summer weather, constrained restoration, and very low birth weight in Sweden. *Health Place* 2013; 22: 68–74.

[94] Ritz B, Yu F. The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993. *Environ Health Perspect* 1999; 107: 17–25.

[95] Murray LJ, O'Reilly DPJ, Betts N, et al. Season and outdoor ambient temperature: Effects on birth weight. *Obstet Gynecol* 2000; 96: 689–695.

[96] Poeran J, Birnie E, Steegers EAP. The Impact of Extremes in Outdoor Temperature and Sunshine Exposure on Birth Weight. *J Environ Health* 2015; 78: 92–100.

[97] Dadvand P, de Nazelle A, Triguero-Mas M, et al. Surrounding greenness and exposure to air pollution during pregnancy: an analysis of personal monitoring data. *Environ Health Perspect* 2012; 120: 1286–90.

[98] Ferguson KK, O'Neill MS, Meeker JD. Environmental contaminant exposures and preterm birth: a comprehensive review. *J Toxicol Environ Heal part B, Crit Rev* 2013; 16: 69–113.

[99] Wine O, Hackett C, Campbell S, et al. Using pollutant release and transfer register data in human health research: a scoping review. *Environ Rev* 2014; 22: 51–65.

[100] Statistics Canada. 2012 Vital Statistics - Birth Database. *Surveys and statistical programs, data released February 10, 2016*http://www.statcan.gc.ca/imdb-bmdi/3231-eng.htm (2016).

[101] Esri. ArcGIS Desktop, Release 10.6 [software]www.esri.com (2017).

[102] StataCorp. Stata Statistical Software: Release 15 [software]www.stata.com (2017).

[103] Chan E, Serrano J, Chen L, et al. Development of a Canadian socioeconomic status index for the study of health outcomes related to environmental pollution. *BMC Public Health* 2015; 15: 714.

[104] Statistics Canada. Census and Intercensal Road Network. Statistics Canada Catalogue no. 92-500-X [digital data]. *92-500-X*http://www12.statcan.gc.ca/census-recensement/2011/geo/RNF-FRR/index-eng.cfm (2012).

[105] Environment Canada. National Pollutant Release Inventory, 2005-2012 [digital data]. *National Pollutant Release Inventory (NPRI)*https://www.ec.gc.ca/inrp-npri/ (2014).

[106] Commission for Environmental Cooperation (CEC). 2010 Land Cover of North America at 30 meters [digital data]. *North American Environmental Atlas and North American Land Change Monitoring System (NALCMS)*http://www.cec.org/tools-and-resources/map-files/land-cover-2010-landsat-30m (2017).

[107] DMTI Spatial. CanMap Content Suite [digital data]. *CanMap Content Suite*http://www.dmtispatial.com/canmap/ (2016).

[108] Silverman BW. *Density estimation for statistics and data analysis*. Chapman and Hall, 1997.

[109] Osornio-Vargas AR, Zaiaine O, Wine O. Domino Project: Data Mining and Newborn Outcomes Exploring Environmental Variables, Abstract Number 2187. In: *International Society for Environmental Epidemiology (ISEE)*. Seattle, WA: National Institute of Environmental Health Sciences, pp. p3-607.

[110] DMTI Spatial. Platinum Postal Code Suite 2001-2013 [digital data]. *CanMap Content Suite*http://www.dmtispatial.com/canmap/ (2014).

[111] Python Software Foundation. Python Language Reference, Version 2.7www.python.org (2016).

[112] Brender JD, Maantay JA, Chakraborty J. Residential proximity to environmental hazards and adverse health outcomes. *Am J Public Health* 2011; 101: 37–52.

[113] Forastiere F, Badaloni C, de Hoogh K, et al. Health impact assessment of waste management facilities in three European countries. *Environ Heal* 2011; 10: 1–13.

[114] Kihal-Talantikite W, Zmirou-Navier D, Padilla C, et al. *Systematic literature review of reproductive outcome associated with residential proximity to polluted sites*. BioMed Central, 2017. Epub ahead of print 2017. DOI: 10.1186/s12942-017-0091-y.

[115] Lin YY, Leon Guo YL, Chen PC, et al. Associations between petrol-station density and

manganese and lead in the cord blood of newborns living in Taiwan. *Environ Res* 2011; 111: 260–265.

[116] Office of Environmental Health Hazard Assessment (OEHHA). The Proposition 65 List. *May 25, 2018*https://oehha.ca.gov/proposition-65/proposition-65-list (2018).

[117] GoodGuide. Health Effects: Developmental Toxicants. *Scorecard*http://scorecard.goodguide.com/health-effects/explanation.tcl?short_hazard_name=devel (2011).

[118] Statistics Canada. Population, urban and rural, by province and territory, Alberta [digital data]http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo62j-eng.htm (2011).

[119] Alberta Health. Alberta Interactive Health Data Application - Reproductive Health - Singleton Small for Gestational Age Percent by Geography and Low Birth Weight Percent by Geography, 2000-2014 [digital data]. *Interactive Health Data Application (IHDA)*http://www.ahw.gov.ab.ca/IHDA_Retrieval/ihdaData.do (2016).

[120] Maisonet M, Correa A, Misra D, et al. A review of the literature on the effects of ambient air pollution on fetal growth. *Environ Res* 2004; 95: 106–15.

[121] Triche EW, Hossain N. Environmental factors implicated in the causation of adverse pregnancy outcome. *Semin Perinatol* 2007; 31: 240–2.

[122] Koranteng S, Osornio-Vargas AR, Buka I. Ambient air pollution and children’s health: A systematic review of Canadian epidemiological studies. *Paediatr Child Health* 2007; 12: 225–233.

[123] Windham G, Fenster L. Environmental contaminants and pregnancy outcomes. *Fertil Steril* 2008; 89: e111-6.

[124] Dadvand P, Parker J, Bell ML, et al. Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity. *Environ Health Perspect* 2013; 121: 367–373.

[125] Stieb DM, Chen L, Eshoul M, et al. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ Res* 2012; 117: 100–11.

[126] Meng G, Hall GB, Thompson ME, et al. Spatial and environmental impacts on adverse birth outcomes in Ontario. *Can Geogr* 2013; 57: 154–72.

[127] Inhaber H. Environmental quality: outline for a national index for Canada. *Science (80-)* 1974; 186: 798–805.

[128] Ott WR. *Environmental Indices: Theory and Practice*. Ann Arbor, MI: Ann Arbor Science Publishers, Inc., 1978.

[129] Messer LC, Jagai JS, Rappazzo KM, et al. Construction of an environmental quality index for public health research. *Environ Heal* 2014; 13: 1–22.

[130] Office of Environmental Health Hazard Assessment. *Draft California communities environmental health screening tool (CalEnviroScreen)*http://oehha.ca.gov/ej/pdf/CES20PublicReview04212014.pdf (2014).

[131] Messer LC, Vinikoor LC, Laraia B a, et al. Socioeconomic domains and associations with preterm birth. *Soc Sci Med* 2008; 67: 1247–57.

[132] NASA EOSDIS Land Processes DAAC, Didan K. MOD13Q1: MODIS/Terra Vegetation Indices 16-Day L3 Global 250m Grid SIN V006 [digital data]. *MOD13Q1 Version 6*. Epub ahead of print 2014. DOI: 10.5067/MODIS/MOD13Q1.006.

[133] Statistics Canada. Health region boundary files, ArcInfo, Alberta, 2007, Catalog no. 82-402-X [digital data]. *82-402-X*http://www.statcan.gc.ca/pub/82-402-x/82-402-x2009001-eng.htm (2008).

[134] Alberta Environment and Sustainable Resource Devlopment. Alberta Airsheds [digital data]. *GeoDiscover Alberta*https://genesis.srd.alberta.ca/genesis_tokenauth/rest/services/Air-Layers/Latest/MapServer/generatekml (2010).

[135] Statistics Canada. *Health regions: boundaries and correspondence with census geography 2007 (updates)*http://www.statcan.gc.ca/pub/82-402-x/82-402-x2009001-eng.pdf (2009).

[136] Clean Air Strategic Alliance. *Airshed Zones Guidelines*. Edmonton, Albertahttp://www.casahome.org/ (2004).

[137] Davarashvili S, Zusman M, Keinan-Boker L, et al. Application of the double kernel density approach to the analysis of cancer incidence in a major metropolitan area. *Environ Res* 2016; 150: 269–281.

[138] Kloog I, Haim A, Portnov BA. Using kernel density function as an urban analysis tool: Investigating the association between nightlight exposure and the incidence of breast cancer in Haifa, Israel. *Comput Environ Urban Syst* 2009; 33: 55–63.

[139] Müller AH, Stadtmüller U, Tabnak F, et al. Spatial Smoothing of Geographically Aggregated Data , With Application to the Construction of Incidence Maps. *J Am Stat Assoc* 2016; 92: 61–71.

[140] Portnov BA, Dubnov J, Barchana M. Studying the association between air pollution and lung cancer incidence in a large metropolitan area using a kernel density function. *Socioecon Plann Sci* 2009; 43: 141–150.

[141] Zusman M, Broitman D, Portnov BA. Application of the double kernel density approach to the multivariate analysis of attributeless event point datasets. *Lett Spat Resour Sci* 2015; 1–20.

[142] Zusman M, Dubnov J, Barchana M, et al. Residential proximity to petroleum storage tanks and associated cancer risks: Double Kernel Density approach vs. zonal estimates. *Sci Total Environ* 2012; 441: 265–276.

[143] AltaLIS. 20K Alberta Base Features [digital data]. *AltaLIS, Agent for Alberta Data Partnerships Ltd. (ADP)*http://www.altalis.com/products/base/20k_base_features.html (2012).

[144] Alberta Biodiversity Monitoring Institute. ABMI Human Footprint Inventory for 2012 conditions Version 3 [digital data]. *Alberta Biodiversity Monitoring Institute (ABMI)*http://www.abmi.ca/home/data/gis-data/human-footprint-download.html?scroll=true (2012).

[145] National Oceanic and Atmospheric Administration. DMSP-OLS Nighttime Lights Time Series, 2005-2012, Version 4 [digital data]. *Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS), National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI)*https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (2012).

[146] Natural Resources Canada. Atlas of Canada National Frameworks Data, Aboriginal Lands, and CanVec [digital data]. *Natural Resources Canada (NRCan), Government of Canada*https://www.nrcan.gc.ca/earth-sciences/geography/topographic-information/free-data-geogratis/download-directory-documentation/17215 (2013).

[147] Alberta Agriculture and Forestry. Historical Wildfire Information - Spatial Wildfire Data, 2005-2012 [digital data]. *Alberta Agriculture and Forestry (AgFor), Government of Alberta, Agriculture and Forestry, Provincial Forest Fire Centre*http://wildfire.alberta.ca/wildfire-maps/historical-wildfire-information/default.aspx (2016).

[148] PyData Development Team. pandas 0.16 [software]http://pandas.pydata.org/pandas-docs/version/0.16.0/ (2015).

[149] Mitchell A. *The Esri Guide to GIS Analysis, Volume 3: Modeling Suitability, Movement, and Interaction*. Redlands, CA: Esri Press, 2012.

[150] GoodGuide. Chemical Profiles. *Scorecard: The Pollution Information Site*http://scorecard.goodguide.com/chemical-profiles/ (2011).

[151] Air Toxics Assessment Group. Health Effects Notebook for Hazardous Air Pollutants. *Environmental Protection Agency*https://www.epa.gov/haps/health-effects-notebook-hazardous-air-pollutants (2017).

[152] Almberg KS, Turyk M, Jones RM, et al. A study of adverse birth outcomes and agricultural land use practices in Missouri. *Environ Res* 2014; 134: 420–426.

[153] Curtis A, Leitner M. *Geographic information systems and public health: eliminating perinatal disparity*. Hershey: IRM Press, 2006.

[154] Miranda ML, Edwards SE, Chang HH, et al. Proximity to roadways and pregnancy outcomes. *J Expo Sci Environ Epidemiol* 2012; 23: 32–8.

[155] Donovan GH, Michael YL, Butry DT, et al. Urban trees and the risk of poor birth outcomes. *Health Place* 2011; 17: 390–3.

[156] Stieb DM, Burnett RT, Smith-Doiron MH, et al. A New Multipollutant, No-Threshold Air Quality Health Index Based on Short-Term Associations Observed in Daily Time-Series Analyses. *J Air Waste Manage Assoc* 2008; 58: 435–450.

[157] Rappazzo KM, Messer LC, Jagai JS, et al. The associations between environmental quality and preterm birth in the United States, 2000-2005: a cross-sectional analysis. *Environ Health* 2015; 14: 50.

[158] Tarocco S, Amoruso I, Caravello G. Holistic model-based monitoring of the human health status in an urban environment system: Pilot study in Verona city, Italy. *J Prev Med Hyg* 2011; 52: 73–82.

[159] Chen L, Bell EM, Caton AR, et al. Residential mobility during pregnancy and the potential for ambient air pollution exposure misclassification. *Environ Res* 2010; 110: 162–8.

[160] Amrhein CG. Searching for the elusive aggregation effect: evidence from statistical simulations. *Environ Plan A* 1995; 27: 105–119.

[161] Qiu X1, Lodha A, Shah PS, Sankaran K, Seshia MM, Yee W, Jefferies A LSCNN. Neonatal outcomes of small for gestational age preterm infants in Canada. *Am J Perinatol* 2012; 29: 87–94.

[162] Mitchell A. *The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics*. Redlands, CA: Esri Press, 2005.

[163] Abdrakhmanov SK, Mukhanbetkaliyev YY, Korennoy FI, et al. Spatio-temporal analysis and visualisation of the anthrax epidemic situation in livestock in Kazakhstan over the period 1933-2016. *Geospat Health* 2017; 12: 316–329.

[164] Hosseini SM, Parvin M, Bahrami M, et al. Pattern mining analysis of pulmonary TB cases in Hamadan province : Using space-time cube. *Int J Epidemiol Res* 2017; 4: 111–117.

[165] Bunting RJ, Chang OY, Cowen C, et al. Spatial patterns of larceny and aggravated assault in Miami–Dade County, 2007–2015. *Prof Geogr* 2017; 69: 1–13.

[166] Harris NL, Goldman E, Gabris C, et al. Using spatial statistics to identify emerging hot spots of forest loss. *Environ Res Lett* 2017; 12: 1–13.

[167] CNN. *Canadian Neonatal Network - Annual Report 2010*http://www.canadianneonatalnetwork.org/Portal/LinkClick.aspx?fileticket=vis_K7gRBsc%3D&tabid=39 (2010).

[168] Shah PS, Seidlitz W, Chan P, et al. Internal Audit of the Canadian Neonatal Network Data Collection System. *Am J Perinatol* 2017; 34: 1241–1249.

[169] Statistics Canada. *Census Dictionary 2011*http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/98-301-X2011001-eng.pdf (2012).

[170] DMTI Spatial. Landcover Region [digital data]. *CanMap Content Suite*http://www.dmtispatial.com/canmap/ (2016).

[171] Esri. An overview of the Space Time Pattern Mining toolbox. *Documentation for ArcGIS*http://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining/an-overview-of-the-space-time-pattern-mining-toolbox.htm (2017).

[172] Birch CPD, Oom SP, Beecham JA. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecol Modell* 2007; 6: 347–359.

[173] StataCorp. Stata Statistical Software: Release 12 [software]www.stata.com (2011).

[174] Nielsen CC, Amrhein CG, Osornio-Vargas AR. Mapping outdoor habitat and abnormally small newborns to develop an ambient health hazard index. *Int J Health Geogr* 2017; 16: 1–21.

[175] Lavigne E, Yasseen AS, Stieb DM, et al. Ambient air pollution and adverse birth outcomes: differences by maternal comorbidities. *Environ Res* 2016; 148: 457–466.

[176] Olsson D, Mogren I, Eneroth K, et al. Traffic pollution at the home address and pregnancy outcomes in Stockholm, Sweden. *BMJ Open* 2015; 5: 1–9.

[177] Mannes T, Jalaludin B, Morgan G, et al. Impact of ambient air pollution on birth weight in Sydney, Australia. *Occup Environ Med* 2005; 62: 524–30.

[178] Le HQ, Batterman SA, Wirth JJ, et al. Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: long-term trends and associations. *Environ Int* 2012; 44: 7–17.

[179] Li J, Adilmagambetov A, Mohomed Jabbar MS, et al. *On discovering co-location patterns in datasets: a case study of pollutants and child cancers*. 2016. Epub ahead of print 2016. DOI: 10.1007/s10707-016-0254-1.

[180] Esri. ArcGIS Pro, Release 2.0 [software]www.esri.com (2017).

[181] Environment Canada. Historical Weather Data. *Historical Climate Data*http://climate.weather.gc.ca/ (2017).

[182] Fisher NI. *Statistical Analysis of Circular Data*. New York: Cambridge University Press, 1993.

[183] Rogerson PA. *Statistical Methods for Geography*. London: SAGE, 2015.

[184] Mardia K V., Jupp PE. *Directional Statistics*. New York: John Wiley and Sons, 2000.

[185] Hystad P, Setton E, Cervantes A, et al. Creating national air pollution models for population exposure assessment in Canada. *Environ Health Perspect* 2011; 119: 1123–1129.

[186] Qiu F, Li B, Chastain B, et al. A GIS based spatially explicit model of dispersal agent behavior. *For Ecol Manage* 2008; 254: 524–537.

[187] Vienneau D, de Hoogh K, Briggs D. A GIS-based method for modelling air pollution exposures across Europe. *Sci Total Environ* 2009; 408: 255–66.

[188] Esri. Enrich Layer. *Documentation for ArcGIS*https://pro.arcgis.com/en/pro-app/tool-reference/analysis/enrich-layer.htm (2018).

[189] Environment and Climate Change Canada. Canadian Environmental Protection Act, 1999, List of toxic substances, Schedule 1. *April 18, 2018*https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/toxic/schedule-1.html (2018).

[190] Currie J, Schmieder JF. Fetal exposures to toxic releases and infant health. *Am Econ Rev* 2009; 99: 177–183.

[191] Hudák A, Ungváry G. Embryotoxic effects of benzene and its methyl derivatives: toluene, xylene. *Toxicology* 1978; 11: 55–63.

[192] Gilliland F, Avol E, Kinney P, et al. Air pollution exposure assessment for epidemiologic studies of pregnant women and children: lessons learned from the Centers for Children's Environmental Health and Disease Prevention Research. *Environ Health Perspect* 2005; 113: 1447–1454.

[193] Salam MT, Millstein J, Li Y-F, et al. Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the Children's Health Study. *Environ Health Perspect* 2005; 113: 1638–1644.

[194] McDermott S, Bao W, Marjorie Aelion C, et al. When are fetuses and young children most susceptible to soil metal concentrations of arsenic, lead and mercury? *Spat Spatiotemporal Epidemiol* 2012; 3: 265–72.

[195] Shirai S, Suzuki Y, Yoshinaga J, et al. Maternal exposure to low-level heavy metals during pregnancy and birth size. *J Environ Sci Heal part A* 2010; 45: 1468–74.

[196] Cho S, Lee C-K, Kim B. The impacts of air pollution on low birth weight. *Appl Econ Lett* 2013; 20: 208–212.

[197] Knudsen LE, Hansen PW, Pedersen M, et al. *Environmental Health Ethics in Study of Children*. Elsevier Inc. Epub ahead of print 2017. DOI: 10.1016/B978-0-12-409548-9.10513-5.

[198] Bender R, Lange S. Adjusting for multiple testing - When and how? *J Clin Epidemiol* 2001; 54: 343–349.

[199] Environment and Climate Change Canada. National Air Pollution Surveillance Program (NAPS) Active Stations 2014 v4. *NAPS Data Products - Network Information*http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx?lang=en (2016).

Appendix I

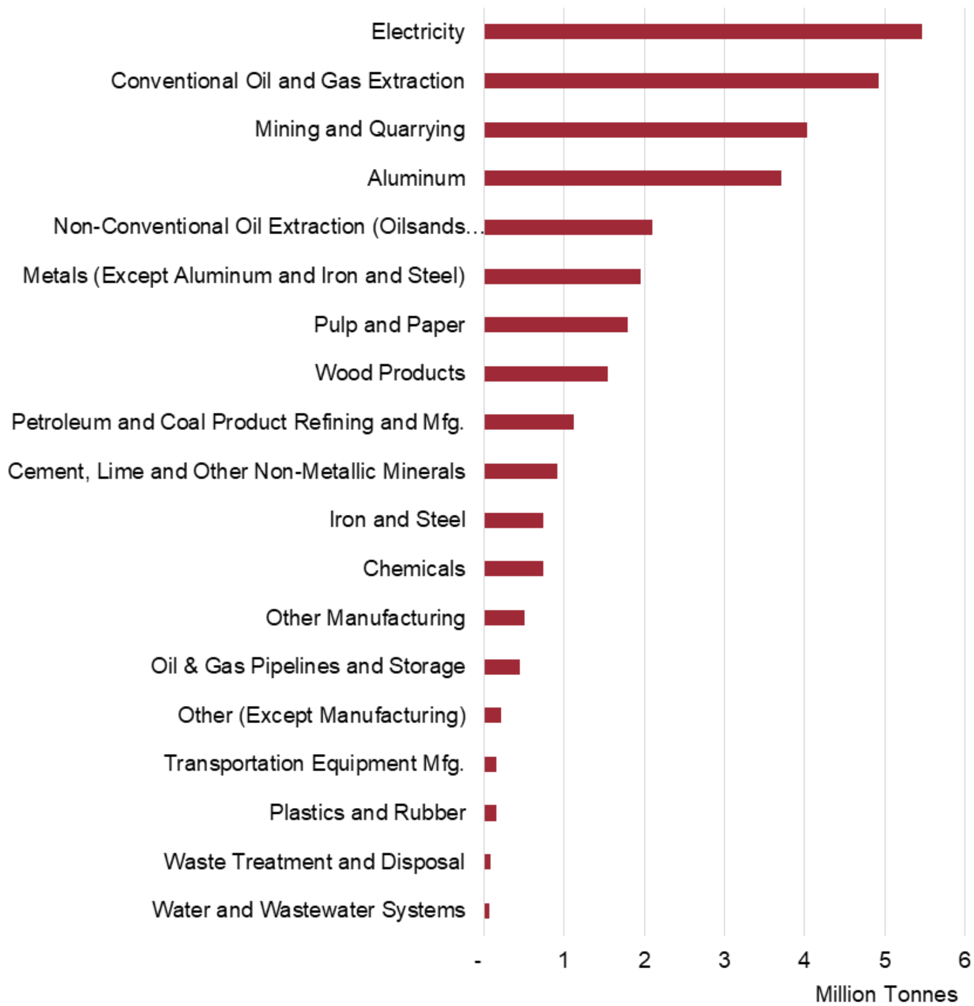


Figure S1.1. Sector distribution of industrial emissions from the National Pollutant Release Inventory (NPRI), 2005-2012.

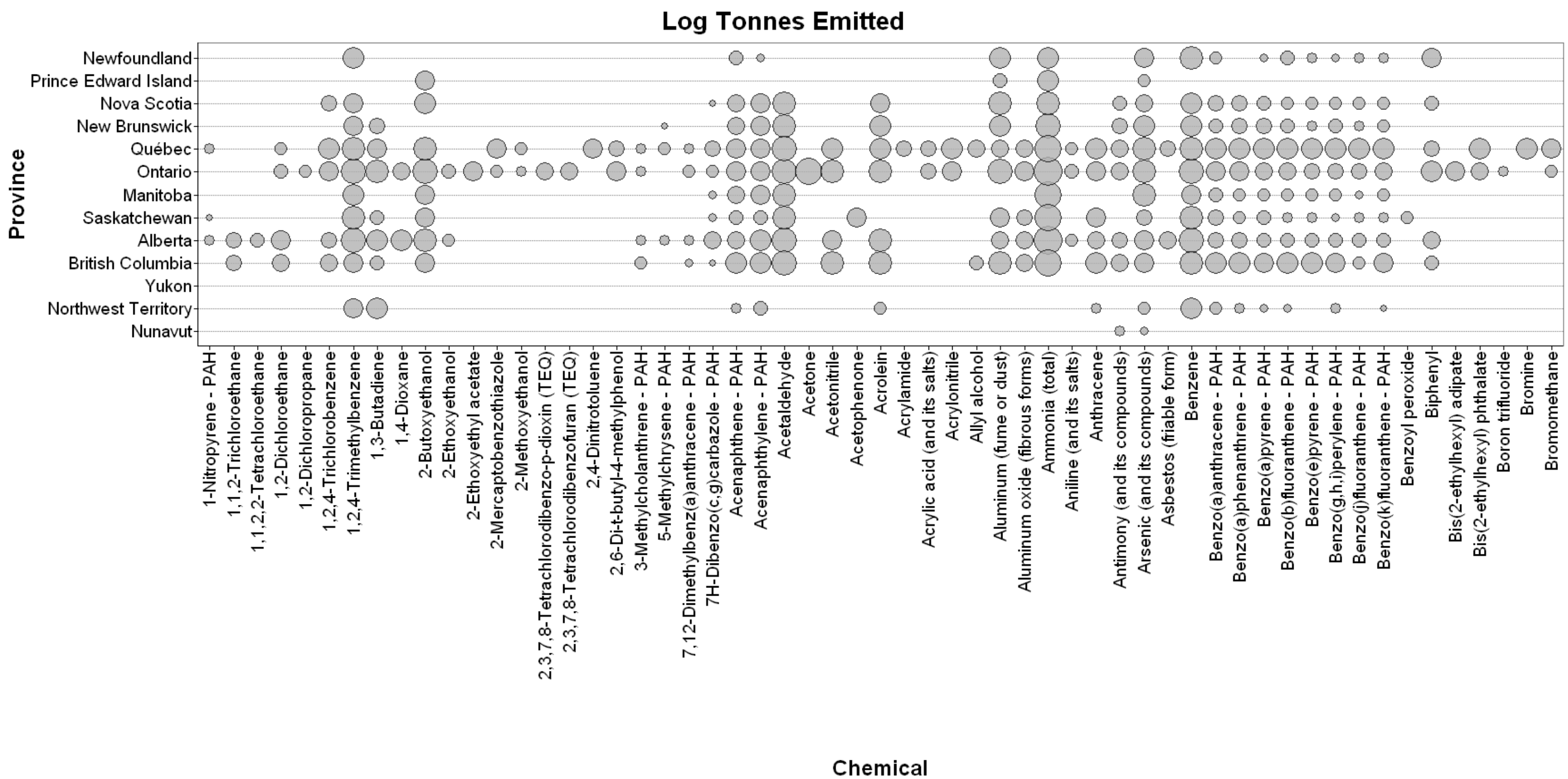


Figure S1.2. Log of the tonnes emitted by province

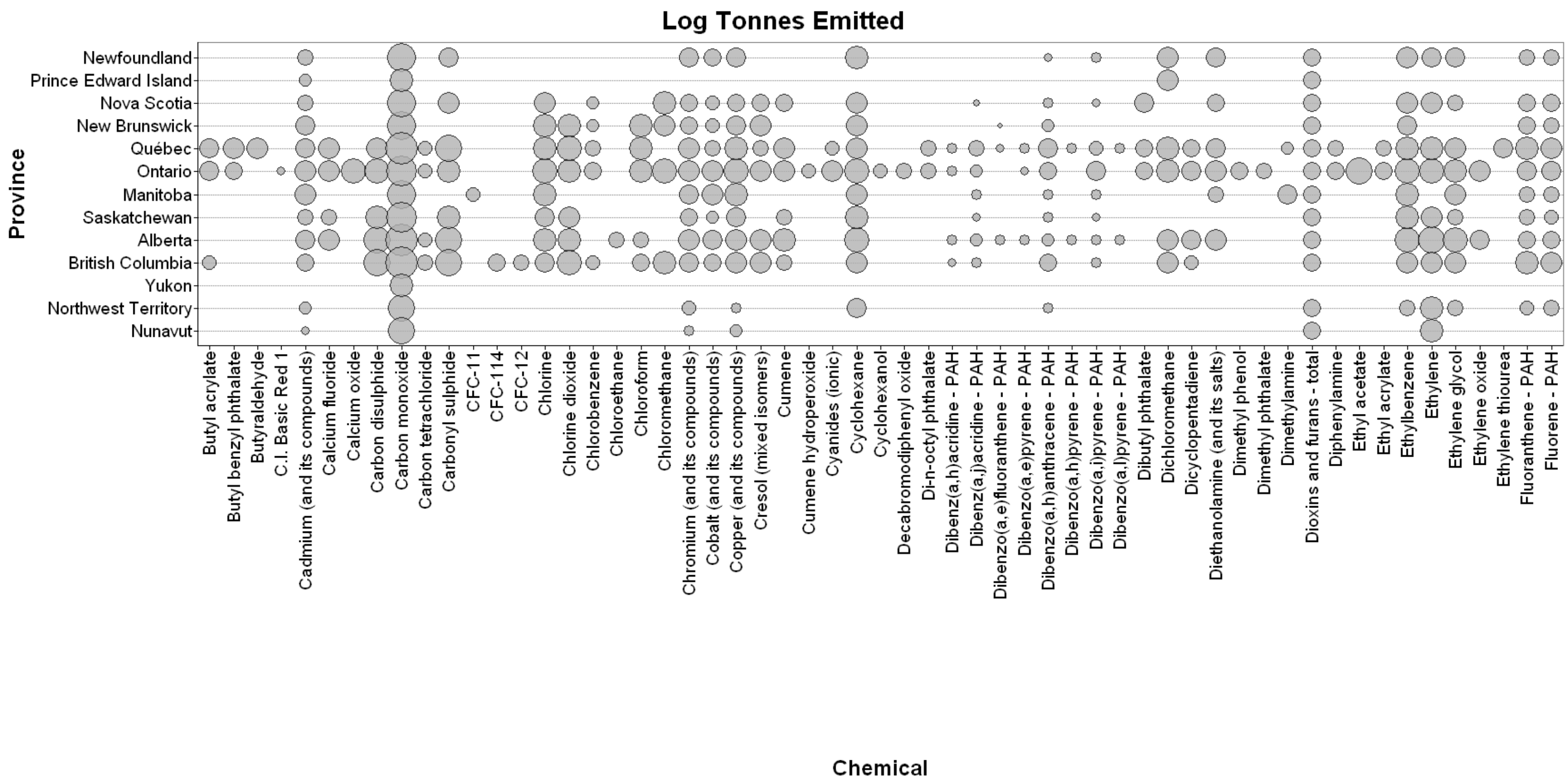


Figure S1.3. Log of the tonnes emitted by province

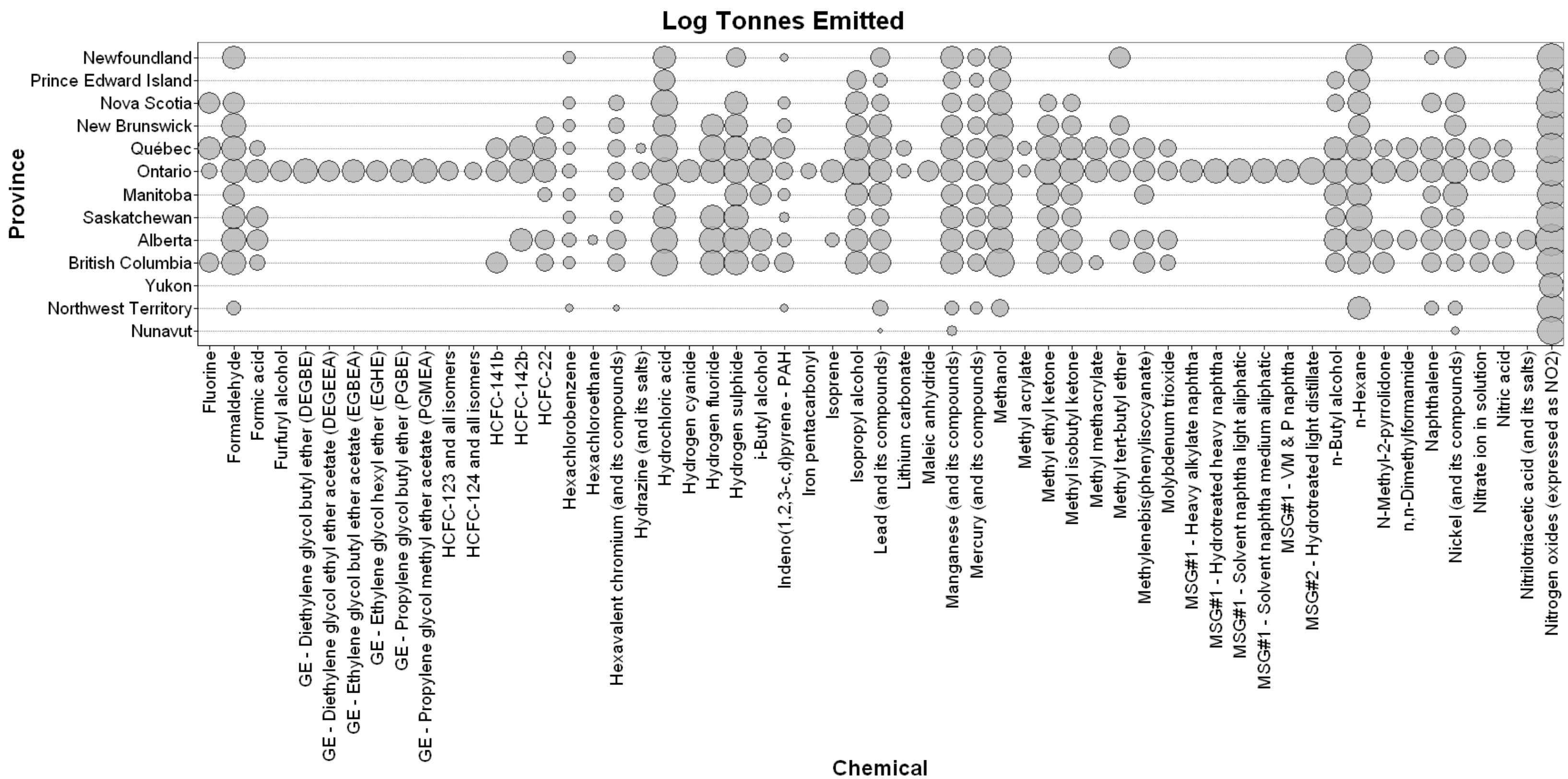


Figure S1.4. Log of the tonnes emitted by province

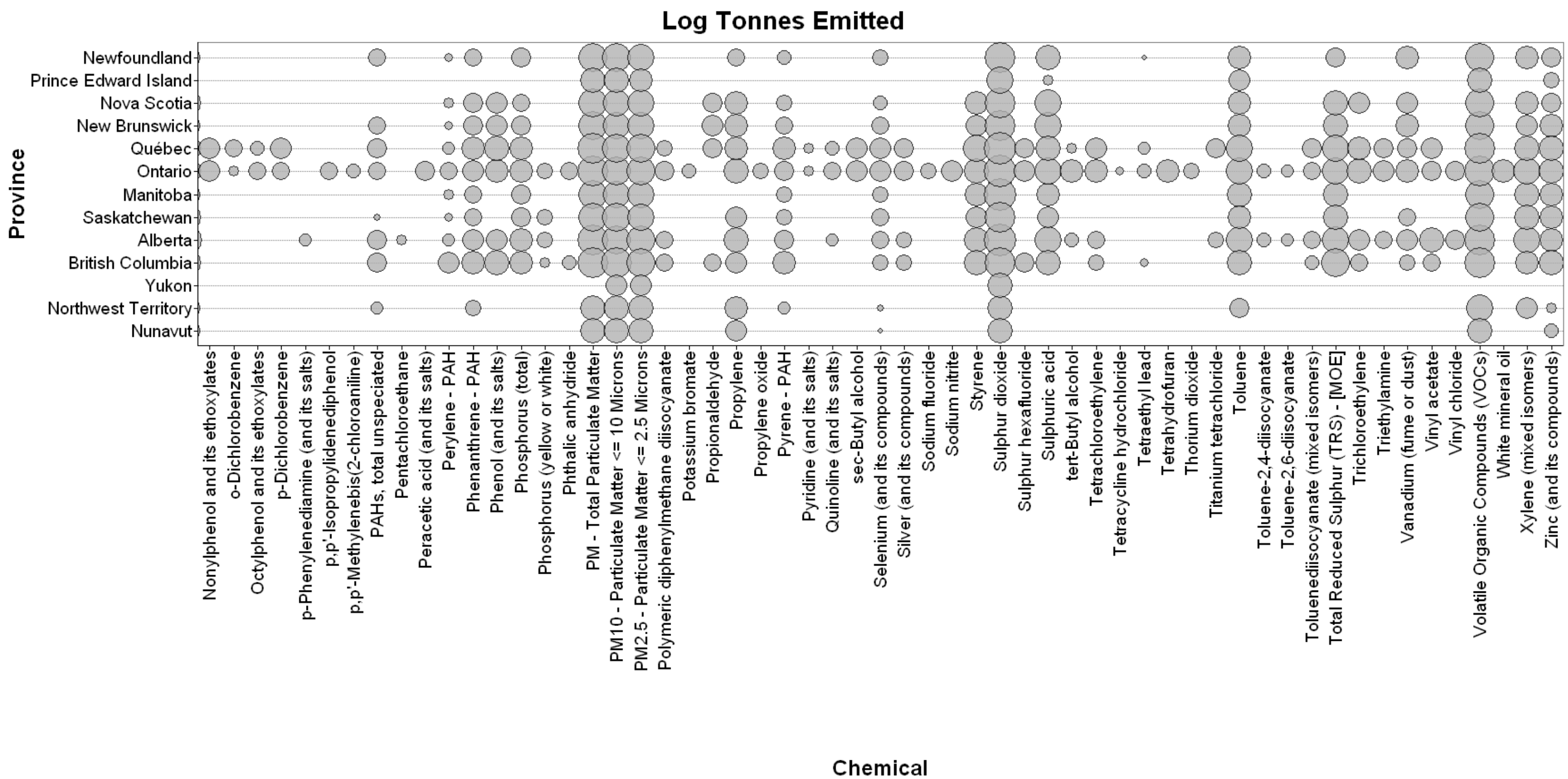


Figure S1.5. Log of the tonnes emitted by province

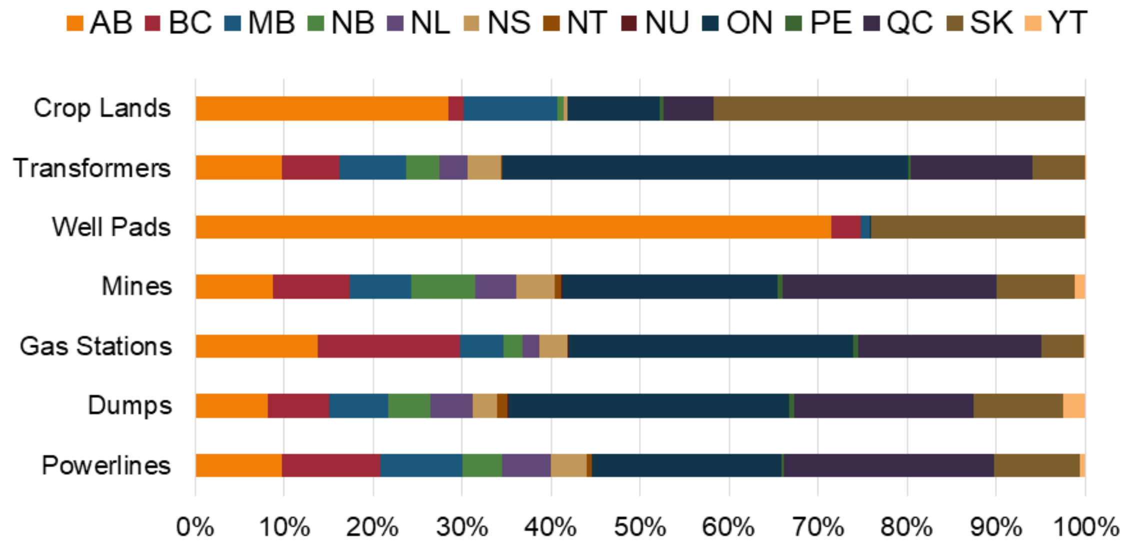


Figure S1.6. Percent of land hazards by province.

Table S1.1. All beta coefficients and p-values, with seasonal covariates: Tri1=warm season in first trimester; Tr2=warm season in second semester; Prior3=warm season in three months preceding birth; NONE=no seasonal variable was used in the model. Variables highlighted in Table 2.6 are in bold. Table continues.

Alberta Season	**SGA Tri1**		**Tri2**		**Prior3**		**NONE**		**LBWT Tri1**		**Tri2**		**Prior3**		**NONE**	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Ammonia (total)	-0.019	0.091	-0.019	0.091	-0.019	0.091	-0.019	0.091	-0.069	0.017	-0.069	0.017	-0.069	0.017	-0.069	0.017
Carbon disulphide	0.176	0.195	0.176	0.193	0.175	0.197	0.175	0.196	0.192	0.509	0.195	0.504	0.190	0.514	0.194	0.505
Carbon monoxide	-0.003	0.007	-0.003	0.007	-0.003	0.007	-0.003	0.007	-0.001	0.769	-0.001	0.771	-0.001	0.767	-0.001	0.771
Carbonyl sulphide	**0.446**	0.074	**0.444**	0.074	**0.445**	0.074	**0.445**	0.074	**0.958**	0.026	**0.958**	0.026	**0.959**	0.026	**0.959**	0.026
Ethylene	-0.010	0.676	-0.010	0.672	-0.010	0.677	-0.010	0.677	-0.084	0.169	-0.084	0.169	-0.084	0.170	-0.084	0.169
HCFC-142b	-0.176	0.403	-0.176	0.401	-0.176	0.401	-0.176	0.402	**0.326**	0.405	**0.326**	0.405	**0.324**	0.407	**0.326**	0.404
Hydrochloric acid	-0.027	0.397	-0.027	0.398	-0.027	0.397	-0.027	0.397	0.001	0.986	0.001	0.986	0.001	0.986	0.001	0.986
Hydrogen fluoride	**0.239**	0.107	**0.240**	0.106	**0.239**	0.107	**0.239**	0.107	**0.411**	0.115	**0.411**	0.114	**0.412**	0.114	**0.411**	0.115
Methanol	-0.034	0.317	-0.034	0.319	-0.034	0.316	-0.034	0.316	-0.025	0.747	-0.025	0.749	-0.025	0.747	-0.025	0.748
n-Hexane	-0.678	0.189	-0.677	0.189	-0.678	0.189	-0.678	0.189	-1.709	0.336	-1.705	0.337	-1.709	0.336	-1.708	0.337
Nitrogen oxides (expressed as NO_2)	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.003	0.002	-0.001	0.543	-0.001	0.545	-0.001	0.541	-0.001	0.544
PM - Total Particulate Matter	-0.005	0.036	-0.005	0.036	-0.005	0.036	-0.005	0.036	0.001	0.776	0.001	0.774	0.001	0.777	0.001	0.774
PM_{10} - Particulate Matter ≤ 10 Microns	-0.014	0.006	-0.014	0.006	-0.014	0.006	-0.014	0.006	-0.003	0.771	-0.003	0.773	-0.003	0.770	-0.003	0.773
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.029	0.001	-0.029	0.001	-0.029	0.001	-0.029	0.001	-0.023	0.238	-0.023	0.239	-0.023	0.238	-0.023	0.239
Styrene	**0.723**	0.009	**0.723**	0.009	**0.720**	0.010	**0.721**	0.009	**0.724**	0.229	**0.729**	0.226	**0.719**	0.232	**0.728**	0.226
Sulphur dioxide	0.000	0.047	0.000	0.048	0.000	0.047	0.000	0.047	0.000	0.881	0.000	0.882	0.000	0.879	0.000	0.882
Sulphuric acid	-0.037	0.263	-0.038	0.262	-0.038	0.261	-0.038	0.262	0.001	0.993	0.001	0.992	0.000	0.996	0.001	0.992
Total Reduced Sulphur (TRS) - [MOE]	0.180	0.200	0.180	0.201	0.181	0.200	0.180	0.200	**0.377**	0.178	**0.377**	0.178	**0.378**	0.177	**0.377**	0.178
Vinyl acetate	0.052	0.175	0.051	0.180	0.052	0.173	0.052	0.174	-0.026	0.781	-0.027	0.776	-0.026	0.784	-0.027	0.779
Volatile Organic Compounds (VOCs)	-0.004	0.005	-0.004	0.005	-0.004	0.005	-0.004	0.005	-0.002	0.415	-0.002	0.416	-0.002	0.415	-0.002	0.416
Zinc (and its compounds)	-0.359	0.772	-0.357	0.773	-0.361	0.770	-0.360	0.771								
Crop	-0.058	0.039	-0.058	0.038	-0.058	0.039	-0.058	0.039	0.015	0.818	0.014	0.820	0.015	0.817	0.014	0.819
Dump	**0.515**	0.001	**0.516**	0.001	**0.515**	0.001	**0.515**	0.001	**0.317**	0.352	**0.317**	0.351	**0.317**	0.352	**0.317**	0.352
Gas station	0.134	0.000	0.134	0.000	0.134	0.000	0.134	0.000	0.132	0.000	0.132	0.000	0.132	0.000	0.132	0.000
Mine	0.042	0.651	0.042	0.648	0.041	0.652	0.041	0.652	0.098	0.643	0.098	0.641	0.097	0.644	0.098	0.642

Alberta	SGA								LBWT							
Well pad	-0.030	0.033	-0.030	0.033	-0.030	0.033	-0.030	0.033	0.000	0.990	0.000	0.989	0.000	0.989	0.000	0.990
Powerline	0.107	0.007	0.107	0.007	0.107	0.007	0.107	0.007	0.051	0.569	0.051	0.571	0.051	0.567	0.051	0.571
Transformer	-0.072	0.663	-0.070	0.673	-0.072	0.662	-0.072	0.662	-0.044	0.906	-0.042	0.909	-0.044	0.905	-0.044	0.907

British Columbia	SGA								LBWT							
Season	Tri1		Tri2		Prior3		NONE		Tri1		Tri2		Prior3		NONE	
Variables	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p
Ammonia (total)	0.042	0.050	0.042	0.050	0.042	0.050	0.042	0.050	0.103	0.026	0.104	0.026	0.103	0.026	0.103	0.026
Carbon monoxide	0.000	0.841	0.000	0.841	0.000	0.841	0.000	0.840	0.001	0.669	0.001	0.670	0.001	0.667	0.001	0.670
Carbonyl sulphide	**0.417**	0.589	**0.417**	0.589	**0.417**	0.589	**0.417**	0.589	**1.626**	0.122	**1.623**	0.122	**1.626**	0.122	**1.626**	0.122
Hydrochloric acid	-0.019	0.670	-0.019	0.670	-0.019	0.670	-0.019	0.670	-0.009	0.932	-0.009	0.932	-0.009	0.929	-0.009	0.932
Hydrogen fluoride	0.171	0.702	0.171	0.702	0.171	0.702	0.171	0.702	**0.823**	0.269	**0.824**	0.269	**0.820**	0.270	**0.823**	0.269
Methanol	-0.005	0.796	-0.005	0.795	-0.005	0.796	-0.005	0.796	0.007	0.861	0.007	0.862	0.007	0.861	0.007	0.861
Nickel (and its compounds)	**0.471**	0.651	**0.471**	0.651	**0.471**	0.651	**0.471**	0.651								
Nitrogen oxides (expressed as NO_2)	-0.001	0.536	-0.001	0.536	-0.001	0.536	-0.001	0.536	0.000	0.940	0.000	0.941	0.000	0.940	0.000	0.940
Phenol (and its salts)	-0.099	0.608	-0.099	0.607	-0.099	0.607	-0.099	0.607	-0.062	0.890	-0.062	0.890	-0.061	0.891	-0.062	0.890
PM - Total Particulate Matter	-0.002	0.738	-0.002	0.738	-0.002	0.738	-0.002	0.738	0.000	0.982	0.000	0.983	0.000	0.980	0.000	0.982
PM_{10} - Particulate Matter ≤ 10 Microns	-0.002	0.822	-0.003	0.821	-0.003	0.822	-0.002	0.822	0.006	0.816	0.006	0.816	0.006	0.814	0.006	0.816
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.003	0.871	-0.003	0.871	-0.003	0.871	-0.003	0.871	0.012	0.760	0.012	0.760	0.012	0.758	0.012	0.760
Styrene	0.061	0.319	0.061	0.318	0.061	0.319	0.061	0.318	0.171	0.153	0.171	0.153	0.172	0.152	0.171	0.153
Sulphur dioxide	0.000	0.745	0.000	0.746	0.000	0.746	0.000	0.745	0.001	0.247	0.001	0.247	0.001	0.245	0.001	0.247
Sulphuric acid	**0.362**	0.001	**0.362**	0.001	**0.362**	0.001	**0.362**	0.001	-0.061	0.871	-0.061	0.872	-0.060	0.873	-0.061	0.871
Toluene	-0.438	0.888	-0.437	0.889	-0.438	0.888	-0.437	0.888	**1.893**	0.526	**1.895**	0.525	**1.901**	0.524	**1.893**	0.526
Total Reduced Sulphur (TRS) - [MOE]	-0.069	0.259	-0.069	0.259	-0.069	0.259	-0.069	0.259	-0.328	0.131	-0.328	0.131	-0.328	0.131	-0.328	0.131
Vanadium (fume or dust)	**1.341**	0.089	**1.342**	0.089	**1.341**	0.089	**1.342**	0.089								
Volatile Organic Compounds (VOCs)	-0.005	0.499	-0.005	0.499	-0.005	0.499	-0.005	0.499	0.006	0.716	0.006	0.717	0.006	0.714	0.006	0.716
Xylene (mixed isomers)	-0.053	0.966	-0.052	0.966	-0.053	0.966	-0.053	0.966	**1.702**	0.224	**1.707**	0.223	**1.701**	0.225	**1.702**	0.224
Crop	-0.020	0.775	-0.020	0.775	-0.020	0.775	-0.020	0.775	0.137	0.395	0.137	0.395	0.137	0.396	0.137	0.395
Dump	**0.522**	0.002	**0.522**	0.002	**0.522**	0.002	**0.522**	0.002	**0.823**	0.039	**0.823**	0.039	**0.824**	0.039	**0.823**	0.039

British Columbia	SGA								LBWT							
Gas station	0.048	0.018	0.048	0.018	0.048	0.018	0.048	0.018	0.037	0.440	0.037	0.440	0.037	0.440	0.037	0.440
Mine	-0.080	0.489	-0.080	0.489	-0.080	0.489	-0.080	0.489	-0.033	0.904	-0.033	0.904	-0.033	0.903	-0.033	0.904
Well pad	**0.391**	0.273	**0.391**	0.272	**0.391**	0.273	**0.391**	0.272	**1.040**	0.052	**1.040**	0.052	**1.042**	0.051	**1.039**	0.052
Powerline	0.145	0.000	0.145	0.000	0.145	0.000	0.145	0.000	**0.299**	0.000	**0.299**	0.000	**0.299**	0.000	**0.299**	0.000
Transformer	**0.218**	0.141	**0.218**	0.141	**0.218**	0.141	**0.218**	0.141	0.068	0.846	0.068	0.846	0.069	0.845	0.068	0.846

Manitoba Season	**SGA Tri1**		**Tri2**		**Prior3**		**NONE**		**LBWT Tri1**		**Tri2**		**Prior3**		**NONE**	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Carbon monoxide	-0.001	0.034	-0.001	0.034	-0.001	0.034	-0.001	0.034	-0.002	0.095	-0.002	0.093	-0.002	0.094	-0.002	0.093
Carbonyl sulphide	-0.147	0.084	-0.148	0.083	-0.147	0.084	-0.147	0.083	-0.380	0.124	-0.383	0.122	-0.381	0.123	-0.383	0.122
Hydrogen fluoride	-0.208	0.392	-0.209	0.390	-0.208	0.392	-0.209	0.391	-0.400	0.506	-0.406	0.499	-0.403	0.502	-0.406	0.499
Hydrogen sulphide	-0.700	0.462	-0.710	0.456	-0.698	0.464	-0.703	0.460								
Methanol	-0.044	0.897	-0.047	0.888	-0.043	0.899	-0.045	0.893	-0.469	0.627	-0.480	0.617	-0.464	0.631	-0.480	0.617
Nitrogen oxides (expressed as NO_2)	-0.010	0.092	-0.010	0.093	-0.010	0.092	-0.010	0.092	-0.025	0.168	-0.025	0.166	-0.025	0.167	-0.025	0.166
PM - Total Particulate Matter	-0.093	0.045	-0.093	0.045	-0.093	0.045	-0.093	0.045	-0.245	0.059	-0.245	0.059	-0.246	0.059	-0.245	0.059
PM_{10} - Particulate Matter $\leq$ 10 Microns	-0.175	0.005	-0.175	0.005	-0.175	0.004	-0.175	0.004	-0.396	0.021	-0.397	0.020	-0.397	0.020	-0.397	0.020
$PM_{2.5}$ - Particulate Matter $\leq$ 2.5 Microns	-0.239	0.003	-0.238	0.003	-0.239	0.003	-0.239	0.003	-0.551	0.019	-0.552	0.018	-0.552	0.018	-0.552	0.018
Sulphur dioxide	-0.004	0.028	-0.004	0.028	-0.004	0.028	-0.004	0.028	-0.009	0.080	-0.009	0.078	-0.009	0.079	-0.009	0.078
Sulphuric acid	-0.132	0.180	-0.132	0.181	-0.132	0.180	-0.133	0.180	-0.309	0.268	-0.310	0.265	-0.309	0.266	-0.310	0.265
Total Reduced Sulphur (TRS) - [MOE]	-0.182	0.375	-0.182	0.374	-0.182	0.374	-0.182	0.374	-0.695	0.247	-0.699	0.244	-0.698	0.245	-0.699	0.244
Volatile Organic Compounds (VOCs)	-0.051	0.594	-0.052	0.587	-0.051	0.594	-0.051	0.590	-0.217	0.380	-0.223	0.368	-0.218	0.379	-0.223	0.368
Crop	-0.123	0.011	-0.123	0.011	-0.123	0.011	-0.123	0.011	-0.370	0.001	-0.369	0.001	-0.369	0.001	-0.369	0.001
Dump	**0.639**	0.009	**0.639**	0.009	**0.638**	0.009	**0.639**	0.009	**0.969**	0.074	**0.965**	0.076	**0.966**	0.075	**0.965**	0.076
Gas station	**0.235**	0.000	**0.235**	0.000	**0.234**	0.000	**0.234**	0.000	**0.329**	0.000	**0.328**	0.001	**0.328**	0.001	**0.328**	0.001
Mine	-0.334	0.309	-0.334	0.309	-0.335	0.308	-0.334	0.309	-0.748	0.354	-0.751	0.353	-0.755	0.350	-0.751	0.353
Well pad	-0.229	0.138	-0.230	0.135	-0.229	0.137	-0.229	0.138	0.006	0.984	0.004	0.989	0.003	0.993	0.004	0.989
Powerline	0.181	0.000	0.181	0.000	0.181	0.000	0.181	0.000	0.196	0.016	0.196	0.016	0.195	0.017	0.196	0.016
Transformer	**0.587**	0.000	**0.587**	0.000	**0.587**	0.000	**0.587**	0.000	**0.546**	0.069	**0.543**	0.071	**0.543**	0.071	**0.543**	0.071

New Brunswick	SGA								LBWT							
Season	Tri1		Tri2		Prior3		NONE		Tri1		Tri2		Prior3		NONE	
Variables	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p
Carbon monoxide	-0.006	0.122	-0.006	0.122	-0.006	0.122	-0.006	0.122	-0.011	0.362	-0.011	0.363	-0.011	0.362	-0.011	0.362
Carbonyl sulphide	-1.305	0.217	-1.307	0.217	-1.306	0.217	-1.307	0.217								
Methanol	-0.125	0.757	-0.124	0.758	-0.126	0.755	-0.124	0.757								
Nitrogen oxides (expressed as NO_2)	0.001	0.911	0.001	0.910	0.001	0.910	0.001	0.910	-0.050	0.156	-0.050	0.157	-0.050	0.156	-0.050	0.156
PM - Total Particulate Matter	-0.026	0.619	-0.026	0.617	-0.026	0.618	-0.026	0.617	-0.564	0.023	-0.564	0.023	-0.564	0.023	-0.564	0.023
PM_{10} - Particulate Matter ≤ 10 Microns	-0.015	0.820	-0.015	0.815	-0.015	0.816	-0.015	0.815	-0.401	0.050	-0.401	0.050	-0.401	0.050	-0.401	0.050
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.111	0.317	-0.111	0.317	-0.111	0.317	-0.111	0.317	-0.383	0.204	-0.382	0.204	-0.382	0.204	-0.382	0.204
Sulphur dioxide	0.000	0.943	0.000	0.940	0.000	0.937	0.000	0.940	-0.007	0.632	-0.007	0.634	-0.007	0.633	-0.007	0.633
Sulphuric acid	**0.480**	0.127	**0.479**	0.128	**0.480**	0.128	**0.479**	0.128	**0.583**	0.355	**0.581**	0.357	**0.581**	0.357	**0.581**	0.356
Volatile Organic Compounds (VOCs)	0.015	0.198	0.015	0.198	0.015	0.199	0.015	0.198	-0.015	0.679	-0.014	0.681	-0.015	0.679	-0.015	0.679
Crop	**0.224**	0.134	**0.223**	0.136	**0.222**	0.136	**0.223**	0.136	-0.064	0.857	-0.065	0.854	-0.065	0.855	-0.065	0.854
Dump	**0.903**	0.060	**0.903**	0.060	**0.905**	0.060	**0.903**	0.060	**1.559**	0.145	**1.554**	0.146	**1.558**	0.145	**1.559**	0.145
Gas station	**0.509**	0.000	**0.509**	0.000	**0.510**	0.000	**0.509**	0.000	**0.831**	0.000	**0.831**	0.000	**0.831**	0.000	**0.831**	0.000
Mine	-0.196	0.188	-0.196	0.190	-0.195	0.191	-0.196	0.190	-0.161	0.635	-0.160	0.637	-0.161	0.635	-0.161	0.636
Powerline	0.086	0.189	0.086	0.187	0.086	0.187	0.086	0.187	0.045	0.765	0.045	0.766	0.045	0.765	0.045	0.765
Transformer	**0.584**	0.044	**0.585**	0.044	**0.585**	0.044	**0.585**	0.044	**0.540**	0.416	**0.542**	0.413	**0.540**	0.415	**0.540**	0.415

Newfoundland Season	**SGA Tri1**		**Tri2**		**Prior3**		**NONE**		**LBWT Tri1**		**Tri2**		**Prior3**		**NONE**	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Ammonia (total)	**0.431**	0.024	**0.432**	0.024	**0.432**	0.023	**0.432**	0.024	**0.454**	0.274	**0.458**	0.270	**0.449**	0.279	**0.454**	0.274
Carbon monoxide	-0.006	0.320	-0.006	0.320	-0.006	0.320	-0.006	0.320	-0.026	0.175	-0.025	0.174	-0.025	0.175	-0.025	0.175
Methanol	**0.227**	0.679	**0.227**	0.679	**0.228**	0.678	**0.228**	0.678	**1.345**	0.072	**1.360**	0.069	**1.349**	0.072	**1.354**	0.071
n-Hexane	0.118	0.483	0.118	0.482	0.118	0.482	0.118	0.482	0.189	0.563	0.189	0.562	0.192	0.559	0.193	0.557
Nitrogen oxides (expressed as NO_2)	-0.005	0.541	-0.005	0.541	-0.005	0.542	-0.005	0.541	-0.024	0.341	-0.024	0.346	-0.024	0.341	-0.024	0.342
PM - Total Particulate Matter	-0.014	0.827	-0.014	0.828	-0.014	0.828	-0.014	0.828	0.009	0.950	0.010	0.944	0.009	0.947	0.010	0.945
PM_{10} - Particulate Matter ≤ 10 Microns	0.044	0.590	0.044	0.590	0.044	0.590	0.044	0.590	0.133	0.434	0.136	0.427	0.133	0.434	0.134	0.432
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	0.127	0.256	0.127	0.256	0.127	0.255	0.127	0.256	**0.265**	0.263	**0.266**	0.260	**0.264**	0.264	**0.265**	0.262
Sulphur dioxide	-0.003	0.741	-0.003	0.742	-0.003	0.743	-0.003	0.742	-0.012	0.556	-0.012	0.565	-0.012	0.557	-0.012	0.560
Total Reduced Sulphur (TRS) - [MOE]	-0.141	0.848	-0.141	0.849	-0.141	0.849	-0.141	0.849	**0.904**	0.382	**0.906**	0.381	**0.907**	0.380	**0.906**	0.381
Volatile Organic Compounds (VOCs)	-0.006	0.851	-0.006	0.852	-0.006	0.851	-0.006	0.851	0.008	0.895	0.008	0.899	0.008	0.895	0.008	0.894
Dump	**0.563**	0.383	**0.562**	0.383	**0.562**	0.383	**0.563**	0.383	**0.680**	0.637	**0.706**	0.624	**0.685**	0.634	**0.682**	0.636
Gas station	**0.290**	0.039	**0.290**	0.039	**0.291**	0.039	**0.290**	0.039	**0.217**	0.485	**0.219**	0.480	**0.213**	0.492	**0.216**	0.487
Mine	-0.217	0.444	-0.216	0.446	-0.216	0.446	-0.216	0.445	-0.666	0.305	-0.663	0.307	-0.666	0.305	-0.660	0.309
Well pad	**0.926**	0.877	**0.923**	0.877	**0.925**	0.877	**0.929**	0.876								
Powerline	-0.085	0.365	-0.085	0.364	-0.085	0.364	-0.085	0.364	-0.248	0.263	-0.248	0.263	-0.249	0.261	-0.250	0.260
Transformer	-0.078	0.859	-0.078	0.859	-0.079	0.856	-0.079	0.858	**1.311**	0.160	**1.295**	0.166	**1.313**	0.160	**1.302**	0.163

Nova Scotia Season	**SGA Tri1**		**Tri2**		**Prior3**		**NONE**		**LBWT Tri1**		**Tri2**		**Prior3**		**NONE**	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Ammonia (total)	-0.030	0.421	-0.030	0.420	-0.030	0.420	-0.030	0.421								
Carbon monoxide	-0.007	0.339	-0.007	0.340	-0.007	0.341	-0.007	0.340	-0.004	0.791	-0.004	0.786	-0.004	0.790	-0.004	0.789
Hydrochloric acid	-0.224	0.200	-0.224	0.200	-0.225	0.198	-0.224	0.200	-1.108	0.212	-1.105	0.213	-1.108	0.212	-1.105	0.213
Methanol	-0.234	0.209	-0.234	0.209	-0.234	0.210	-0.234	0.209	0.196	0.573	0.196	0.573	0.197	0.572	0.196	0.573

Nova Scotia	**SGA**								**LBWT**							
Nitrogen oxides (expressed as NO_2)	-0.003	0.440	-0.003	0.440	-0.003	0.442	-0.003	0.440	-0.001	0.955	-0.001	0.953	-0.001	0.956	-0.001	0.954
PM - Total Particulate Matter	-0.040	0.286	-0.040	0.287	-0.040	0.289	-0.040	0.287	0.037	0.627	0.036	0.632	0.037	0.627	0.037	0.629
PM_{10} - Particulate Matter ≤ 10 Microns	-0.039	0.359	-0.039	0.360	-0.038	0.364	-0.039	0.359	0.049	0.571	0.048	0.577	0.049	0.570	0.049	0.574
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.042	0.463	-0.042	0.464	-0.041	0.469	-0.042	0.464	0.089	0.443	0.088	0.449	0.089	0.443	0.089	0.446
Sulphur dioxide	-0.001	0.347	-0.001	0.347	-0.001	0.349	-0.001	0.347	-0.002	0.561	-0.002	0.559	-0.002	0.562	-0.002	0.561
Sulphuric acid	0.045	0.849	0.045	0.849	0.048	0.840	0.045	0.849	0.129	0.801	0.128	0.803	0.130	0.799	0.128	0.803
Total Reduced Sulphur (TRS) - [MOE]	-0.220	0.857	-0.220	0.857	-0.223	0.855	-0.220	0.857								
Volatile Organic Compounds (VOCs)	-0.007	0.512	-0.007	0.513	-0.007	0.518	-0.007	0.513	0.002	0.929	0.002	0.932	0.002	0.928	0.002	0.931
Crop	0.182	0.124	0.182	0.124	0.182	0.125	0.182	0.124	0.153	0.567	0.153	0.566	0.153	0.567	0.153	0.566
Dump	**0.334**	0.445	**0.334**	0.445	**0.334**	0.445	**0.334**	0.445	**0.662**	0.487	**0.664**	0.486	**0.662**	0.487	**0.662**	0.487
Gas station	**0.205**	0.002	**0.205**	0.002	**0.205**	0.002	**0.205**	0.002	0.078	0.590	0.078	0.591	0.078	0.590	0.078	0.591
Mine	-0.073	0.711	-0.073	0.711	-0.073	0.713	-0.073	0.711	0.029	0.948	0.028	0.948	0.029	0.947	0.029	0.948
Powerline	0.060	0.328	0.060	0.328	0.059	0.329	0.060	0.328	-0.010	0.939	-0.010	0.939	-0.010	0.940	-0.010	0.941
Transformer	**0.267**	0.109	**0.267**	0.109	**0.267**	0.108	**0.267**	0.109	0.109	0.767	0.108	0.769	0.109	0.767	0.109	0.768

Northwest Territory	**SGA**								**LBWT**							
Season	**Tri1**		**Tri2**		**Prior3**		**NONE**		**Tri1**		**Tri2**		**Prior3**		**NONE**	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Carbon monoxide	**0.249**	0.440	**0.247**	0.443	**0.248**	0.441	**0.248**	0.441	0.122	0.864	0.138	0.845	0.133	0.852	0.127	0.858
Nitrogen oxides (expressed as NO_2)	0.009	0.873	0.009	0.882	0.009	0.877	0.009	0.880	0.030	0.806	0.032	0.792	0.033	0.790	0.031	0.799
PM_{10} - Particulate Matter ≤ 10 Microns	**1.447**	0.414	**1.455**	0.412	**1.453**	0.412	**1.458**	0.411	-3.342	0.461	-3.251	0.471	-3.348	0.461	-3.342	0.461
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	**1.470**	0.407	**1.478**	0.404	**1.476**	0.405	**1.481**	0.403	-3.381	0.456	-3.290	0.466	-3.388	0.456	-3.381	0.456
Sulphur dioxide	-0.351	0.228	-0.351	0.229	-0.351	0.228	-0.350	0.229								
Volatile Organic Compounds (VOCs)	**0.316**	0.110	**0.317**	0.109	**0.317**	0.108	**0.317**	0.108								
Dump	**0.480**	0.263	**0.477**	0.265	**0.478**	0.264	**0.477**	0.265	**1.327**	0.192	**1.339**	0.188	**1.337**	0.189	**1.329**	0.191
Gas station	**1.921**	0.119	**1.940**	0.115	**1.937**	0.115	**1.941**	0.115	**2.625**	0.412	**2.614**	0.413	**2.592**	0.417	**2.607**	0.415
Mine	-0.269	0.874	-0.267	0.875	-0.259	0.879	-0.259	0.879	**2.640**	0.409	**2.800**	0.384	**2.634**	0.410	**2.636**	0.410
Powerline	-1.050	0.356	-1.050	0.357	-1.058	0.353	-1.051	0.356	-2.069	0.451	-2.102	0.444	-2.099	0.443	-2.068	0.451
Transformer	-1.010	0.518	-1.001	0.521	-1.002	0.521	-1.002	0.521	-2.409	0.513	-2.396	0.515	-2.369	0.520	-2.401	0.514

Nunavut Season	SGA Tri1		Tri2		Prior3		NONE		LBWT Tri1		Tri2		Prior3		NONE	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Carbon monoxide	**1.382**	0.034	**1.480**	0.023	**1.385**	0.034	**1.400**	0.032	**3.423**	0.003	**3.539**	0.002	**3.385**	0.003	**3.436**	0.003
Nitrogen oxides (expressed as NO_2)	0.117	0.474	0.126	0.445	0.118	0.474	0.118	0.472	**0.395**	0.195	**0.406**	0.184	**0.393**	0.197	**0.396**	0.195
PM_{10} - Particulate Matter ≤ 10 Microns	**2.907**	0.218	**3.186**	0.177	**2.922**	0.216	**2.969**	0.208	**8.724**	0.040	**9.048**	0.033	**8.621**	0.042	**8.778**	0.039
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	**2.918**	0.215	**3.196**	0.175	**2.932**	0.213	**2.980**	0.206	**8.750**	0.039	**9.074**	0.033	**8.647**	0.042	**8.804**	0.038
Sulphur dioxide	**0.968**	0.304	**0.889**	0.342	**0.951**	0.313	**0.969**	0.304	**1.818**	0.166	**1.674**	0.197	**1.743**	0.183	**1.827**	0.164
Dump	-8.547	0.001	-8.372	0.001	-8.498	0.001	-8.503	0.001	-6.768	0.159	-6.475	0.178	-6.705	0.163	-6.716	0.162
Gas station	-4.181	0.040	-4.105	0.044	-4.142	0.042	-4.146	0.042	-2.770	0.421	-2.614	0.447	-2.709	0.431	-2.741	0.426
Mine	-0.094	0.858	-0.126	0.812	-0.103	0.845	-0.101	0.849	-0.606	0.517	-0.641	0.493	-0.626	0.503	-0.611	0.513
Well pad	**1.317**	0.307	**1.258**	0.329	**1.309**	0.310	**1.324**	0.304	-0.433	0.856	-0.486	0.838	-0.487	0.838	-0.433	0.855
Powerline	-4.519	0.042	-4.466	0.044	-4.476	0.044	-4.480	0.044	-4.142	0.296	-4.010	0.311	-4.076	0.304	-4.107	0.300

Ontario Season	SGA Tri1		Tri2		Prior3		NONE		LBWT Tri1		Tri2		Prior3		NONE	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
1,2,4-Trimethylbenzene	-0.019	0.873	-0.018	0.873	-0.019	0.872	-0.019	0.872	-0.267	0.361	-0.266	0.362	-0.267	0.360	-0.266	0.362
2-Butoxyethanol	-1.159	0.497	-1.152	0.500	-1.159	0.497	-1.159	0.497	**1.421**	0.362	**1.426**	0.360	**1.416**	0.363	**1.424**	0.361
Aluminum (fume or dust)	0.017	0.858	0.017	0.859	0.017	0.859	0.017	0.859	0.089	0.662	0.089	0.661	0.088	0.667	0.089	0.661
Ammonia (total)	-0.006	0.713	-0.006	0.706	-0.006	0.718	-0.006	0.716	-0.032	0.381	-0.033	0.372	-0.032	0.383	-0.033	0.375
Benzene	**0.488**	0.128	**0.489**	0.127	**0.487**	0.128	**0.487**	0.128	**1.088**	0.033	**1.095**	0.032	**1.087**	0.033	**1.093**	0.032
Calcium oxide	**0.304**	0.159	**0.305**	0.158	**0.303**	0.160	**0.304**	0.160	-0.119	0.837	-0.115	0.843	-0.122	0.833	-0.117	0.840
Carbon monoxide	0.001	0.147	0.001	0.146	0.001	0.149	0.001	0.148	0.001	0.627	0.001	0.617	0.001	0.634	0.001	0.621
Chloromethane	-0.095	0.287	-0.095	0.286	-0.095	0.287	-0.095	0.287	-0.244	0.348	-0.244	0.349	-0.244	0.348	-0.244	0.349
Copper (and its compounds)	0.095	0.350	0.095	0.352	0.095	0.349	0.095	0.349	-1.845	0.339	-1.837	0.340	-1.851	0.339	-1.839	0.339
Cyclohexane	-0.212	0.315	-0.212	0.316	-0.212	0.315	-0.212	0.315	-0.220	0.657	-0.221	0.656	-0.220	0.658	-0.222	0.655
Ethyl acetate	0.022	0.034	0.022	0.033	0.022	0.035	0.022	0.035	0.037	0.084	0.038	0.076	0.037	0.089	0.038	0.078
Ethylene	-0.256	0.271	-0.255	0.273	-0.256	0.271	-0.256	0.271	-0.514	0.421	-0.512	0.423	-0.513	0.422	-0.513	0.422

Ontario	SGA								LBWT							
GE - Diethylene glycol butyl ether (DEGBE)	**0.211**	0.271	**0.212**	0.269	**0.211**	0.271	**0.211**	0.271	**0.314**	0.414	**0.313**	0.415	**0.313**	0.415	**0.313**	0.416
GE - Propylene glycol methyl ether acetate (PGMEA)	0.095	0.873	0.095	0.874	0.095	0.874	0.095	0.874								
HCFC-142b	0.031	0.563	0.031	0.558	0.031	0.566	0.031	0.565	0.001	0.996	0.002	0.986	-0.001	0.997	0.001	0.990
HCFC-22	**1.416**	0.220	**1.419**	0.219	**1.418**	0.220	**1.418**	0.220								
Hydrochloric acid	-0.010	0.556	-0.010	0.556	-0.010	0.552	-0.010	0.553	-0.014	0.742	-0.013	0.751	-0.014	0.735	-0.013	0.749
Hydrogen fluoride	-0.062	0.676	-0.062	0.676	-0.062	0.675	-0.062	0.675	0.130	0.657	0.131	0.654	0.128	0.661	0.131	0.655
Hydrogen sulphide	-0.740	0.222	-0.738	0.222	-0.740	0.221	-0.740	0.221	-1.081	0.513	-1.076	0.514	-1.079	0.513	-1.080	0.513
i-Butyl alcohol	-0.274	0.481	-0.275	0.480	-0.274	0.481	-0.274	0.481	**0.418**	0.284	**0.416**	0.287	**0.416**	0.286	**0.418**	0.284
Isopropyl alcohol	0.032	0.647	0.032	0.645	0.032	0.649	0.032	0.649	0.111	0.449	0.112	0.445	0.111	0.450	0.112	0.446
Manganese (and its compounds)	**1.000**	0.078	**0.996**	0.079	**1.002**	0.077	**1.001**	0.078								
Methanol	0.063	0.527	0.063	0.528	0.063	0.528	0.063	0.527	-0.025	0.918	-0.025	0.919	-0.026	0.916	-0.025	0.919
Methyl ethyl ketone	0.085	0.009	0.085	0.009	0.085	0.009	0.085	0.009	**0.203**	0.002	**0.204**	0.002	**0.202**	0.002	**0.204**	0.002
MSG#1 - Hydrotreated heavy naphtha	0.196	0.341	0.196	0.342	0.195	0.344	0.195	0.343	**0.517**	0.064	**0.522**	0.061	**0.515**	0.065	**0.523**	0.061
MSG#1 - Solvent naphtha light aliphatic	0.067	0.355	0.068	0.349	0.067	0.361	0.067	0.359	0.102	0.513	0.106	0.493	0.099	0.525	0.105	0.501
MSG#2 - Hydrotreated light distillate	0.027	0.261	0.027	0.262	0.027	0.263	0.027	0.262	0.069	0.041	0.069	0.041	0.068	0.042	0.069	0.041
n-Butyl alcohol	0.102	0.236	0.102	0.237	0.102	0.236	0.102	0.236	0.016	0.936	0.015	0.940	0.016	0.938	0.015	0.939
n-Hexane	0.008	0.790	0.008	0.792	0.008	0.792	0.008	0.791	0.104	0.059	0.105	0.058	0.104	0.059	0.105	0.058
Nitrogen oxides (expressed as NO_2)	0.002	0.054	0.002	0.054	0.002	0.055	0.002	0.055	0.003	0.143	0.003	0.139	0.002	0.147	0.003	0.140
PM - Total Particulate Matter	0.006	0.091	0.006	0.090	0.006	0.092	0.006	0.092	0.009	0.225	0.010	0.218	0.009	0.231	0.010	0.220
PM_{10} - Particulate Matter ≤ 10 Microns	0.018	0.002	0.018	0.002	0.018	0.003	0.018	0.003	0.023	0.077	0.023	0.072	0.023	0.081	0.023	0.073
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	0.036	0.000	0.036	0.000	0.036	0.000	0.036	0.000	0.046	0.014	0.047	0.013	0.046	0.015	0.046	0.013
Propylene	-0.026	0.877	-0.026	0.877	-0.026	0.878	-0.026	0.877	-0.002	0.997	-0.001	0.997	0.000	0.999	-0.002	0.997
Styrene	**1.027**	0.113	**1.031**	0.112	**1.026**	0.113	**1.027**	0.113	**1.605**	0.126	**1.612**	0.125	**1.603**	0.127	**1.608**	0.126
Sulphur dioxide	0.000	0.735	0.000	0.734	0.000	0.738	0.000	0.737	0.000	0.979	0.000	0.987	0.000	0.973	0.000	0.985
Sulphuric acid	-0.052	0.438	-0.052	0.440	-0.052	0.438	-0.052	0.438	-0.072	0.655	-0.071	0.658	-0.072	0.656	-0.072	0.656
Tetrahydrofuran	**0.486**	0.659	**0.488**	0.658	**0.487**	0.658	**0.487**	0.659								
Toluene	0.016	0.401	0.016	0.395	0.016	0.409	0.016	0.406	-0.015	0.726	-0.013	0.753	-0.016	0.709	-0.014	0.745
Total Reduced Sulphur (TRS) - [MOE]	-0.809	0.053	-0.810	0.053	-0.809	0.053	-0.809	0.053	-1.066	0.331	-1.065	0.331	-1.067	0.330	-1.065	0.331
Trichloroethylene	-0.278	0.525	-0.276	0.528	-0.282	0.520	-0.281	0.522	**0.570**	0.430	**0.590**	0.413	**0.559**	0.438	**0.586**	0.417
Volatile Organic Compounds (VOCs)	0.005	0.000	0.005	0.000	0.005	0.000	0.005	0.000	0.006	0.000	0.006	0.000	0.006	0.000	0.006	0.000
White mineral oil	**2.076**	0.237	**2.084**	0.235	**2.072**	0.238	**2.073**	0.237								

Ontario	SGA								LBWT							
Xylene (mixed isomers)	-0.021	0.586	-0.020	0.595	-0.021	0.575	-0.021	0.579	-0.060	0.478	-0.056	0.507	-0.062	0.465	-0.058	0.498
Zinc (and its compounds)	0.154	0.432	0.153	0.434	0.154	0.431	0.154	0.432	-0.984	0.340	-0.980	0.341	-0.986	0.340	-0.983	0.340
Crop	-0.035	0.115	-0.035	0.114	-0.035	0.115	-0.035	0.115	0.017	0.736	0.016	0.743	0.017	0.732	0.017	0.740
Dump	-0.007	0.917	-0.007	0.920	-0.007	0.918	-0.007	0.918	0.019	0.894	0.019	0.895	0.020	0.890	0.019	0.896
Gas station	-0.017	0.111	-0.017	0.112	-0.017	0.110	-0.017	0.110	-0.006	0.789	-0.006	0.797	-0.006	0.785	-0.006	0.793
Mine	0.074	0.378	0.074	0.379	0.074	0.378	0.074	0.378	-0.140	0.468	-0.141	0.468	-0.141	0.467	-0.141	0.468
Well pad	-0.007	0.986	-0.010	0.982	-0.007	0.987	-0.007	0.986	-0.079	0.936	-0.082	0.933	-0.076	0.938	-0.080	0.935
Powerline	0.041	0.000	0.041	0.000	0.041	0.000	0.041	0.000	0.046	0.024	0.046	0.024	0.046	0.023	0.046	0.023
Transformer	0.030	0.125	0.030	0.124	0.030	0.126	0.030	0.126	0.006	0.885	0.007	0.875	0.006	0.891	0.007	0.878

Prince Edward Island	SGA								LBWT							
Season	Tri1		Tri2		Prior3		NONE		Tri1		Tri2		Prior3		NONE	
Variables	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p
Carbon monoxide	-0.116	0.007	-0.116	0.007	-0.116	0.007	-0.116	0.007	-0.137	0.142	-0.136	0.142	-0.137	0.142	-0.136	0.142
Methanol	-0.450	0.037	-0.449	0.037	-0.450	0.037	-0.450	0.037								
Nitrogen oxides (expressed as NO_2)	-0.119	0.006	-0.119	0.006	-0.120	0.006	-0.120	0.006	-0.113	0.200	-0.113	0.201	-0.114	0.200	-0.113	0.201
PM - Total Particulate Matter	-0.292	0.278	-0.294	0.274	-0.296	0.272	-0.294	0.275	0.005	0.991	0.008	0.986	0.002	0.996	0.008	0.986
PM_{10} - Particulate Matter ≤ 10 Microns	-0.340	0.246	-0.342	0.244	-0.343	0.242	-0.342	0.244	0.132	0.803	0.134	0.800	0.131	0.805	0.134	0.800
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-1.494	0.010	-1.493	0.010	-1.495	0.010	-1.494	0.010	-0.717	0.491	-0.715	0.492	-0.718	0.491	-0.715	0.492
Sulphur dioxide	-0.104	0.008	-0.104	0.008	-0.104	0.008	-0.104	0.008	-0.069	0.306	-0.068	0.307	-0.069	0.305	-0.068	0.307
Volatile Organic Compounds (VOCs)	-0.098	0.034	-0.098	0.034	-0.099	0.034	-0.098	0.034	-0.023	0.673	-0.022	0.677	-0.023	0.670	-0.022	0.676
Crop	-0.126	0.621	-0.126	0.621	-0.124	0.625	-0.126	0.622	-0.383	0.533	-0.383	0.533	-0.380	0.536	-0.383	0.533
Dump	**0.322**	0.883	**0.345**	0.875	**0.325**	0.882	**0.325**	0.882	-6.478	0.274	-6.477	0.273	-6.482	0.273	-6.480	0.273
Gas station	**0.836**	0.004	**0.837**	0.004	**0.838**	0.004	**0.838**	0.004	**0.866**	0.159	**0.865**	0.159	**0.865**	0.160	**0.866**	0.159
Mine	-1.870	0.032	-1.867	0.032	-1.869	0.032	-1.869	0.032	-0.452	0.810	-0.451	0.811	-0.458	0.808	-0.451	0.811
Powerline	0.114	0.671	0.113	0.674	0.113	0.675	0.113	0.675	-0.971	0.168	-0.969	0.169	-0.969	0.169	-0.969	0.169
Transformer	**1.102**	0.519	**1.094**	0.522	**1.106**	0.517	**1.107**	0.517	-2.364	0.542	-2.373	0.540	-2.367	0.542	-2.368	0.541

Québec Season	SGA Tri1		Tri2		Prior3		NONE		LBWT Tri1		Tri2		Prior3		NONE	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Ammonia (total)	0.039	0.003	0.039	0.003	0.039	0.003	0.039	0.003	0.070	0.008	0.070	0.008	0.070	0.008	0.070	0.008
Carbon monoxide	-0.001	0.114	-0.001	0.115	-0.001	0.115	-0.001	0.115	-0.001	0.294	-0.001	0.293	-0.001	0.293	-0.001	0.293
Carbonyl sulphide	-0.142	0.020	-0.142	0.021	-0.142	0.020	-0.142	0.020	-0.173	0.235	-0.174	0.233	-0.173	0.235	-0.174	0.233
Copper (and its compounds)	-1.086	0.271	-1.086	0.271	-1.087	0.270	-1.086	0.271								
Ethylene	0.075	0.815	0.065	0.839	0.073	0.822	0.064	0.842	-0.803	0.430	-0.786	0.440	-0.803	0.429	-0.791	0.437
Formaldehyde	**0.886**	0.197	**0.887**	0.197	**0.886**	0.197	**0.886**	0.197								
HCFC-142b	**0.321**	0.000	**0.320**	0.000	**0.321**	0.000	**0.320**	0.000	-0.179	0.485	-0.177	0.489	-0.178	0.486	-0.177	0.488
Hydrochloric acid	0.016	0.725	0.015	0.727	0.015	0.728	0.015	0.727	0.119	0.135	0.119	0.135	0.120	0.134	0.119	0.135
Hydrogen fluoride	-0.352	0.087	-0.351	0.088	-0.351	0.087	-0.350	0.088	-0.284	0.535	-0.287	0.531	-0.285	0.534	-0.286	0.532
i-Butyl alcohol	-0.119	0.791	-0.121	0.788	-0.120	0.789	-0.120	0.788	**0.932**	0.024	**0.932**	0.024	**0.934**	0.024	**0.934**	0.024
Isopropyl alcohol	-0.119	0.791	-0.121	0.788	-0.120	0.789	-0.120	0.788	**0.932**	0.024	**0.932**	0.024	**0.934**	0.024	**0.934**	0.024
Methanol	-0.078	0.114	-0.077	0.116	-0.078	0.115	-0.077	0.116	-0.029	0.787	-0.030	0.783	-0.029	0.785	-0.030	0.783
Methyl ethyl ketone	0.062	0.655	0.062	0.656	0.062	0.657	0.062	0.657	**0.316**	0.208	**0.317**	0.206	**0.317**	0.206	**0.316**	0.207
n-Hexane	0.156	0.615	0.155	0.616	0.157	0.611	0.156	0.615	**0.582**	0.261	**0.582**	0.261	**0.579**	0.264	**0.583**	0.261
Nitrogen oxides (expressed as NO_2)	0.004	0.008	0.004	0.009	0.004	0.008	0.004	0.009	0.006	0.062	0.006	0.061	0.006	0.062	0.006	0.062
Phenol (and its salts)	0.060	0.912	0.058	0.915	0.056	0.917	0.057	0.916	0.080	0.949	0.085	0.945	0.085	0.946	0.084	0.946
PM - Total Particulate Matter	-0.004	0.473	-0.004	0.472	-0.004	0.471	-0.004	0.473	0.000	1.000	0.000	0.999	0.000	0.999	0.000	1.000
PM_{10} - Particulate Matter ≤ 10 Microns	-0.009	0.230	-0.009	0.228	-0.009	0.229	-0.009	0.228	-0.005	0.764	-0.005	0.767	-0.005	0.765	-0.005	0.767
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.018	0.106	-0.018	0.105	-0.018	0.105	-0.018	0.105	-0.009	0.719	-0.009	0.723	-0.009	0.720	-0.009	0.722
Propylene	0.134	0.189	0.134	0.190	0.134	0.188	0.134	0.191	-0.326	0.263	-0.324	0.265	-0.326	0.262	-0.325	0.264
Styrene	-0.421	0.154	-0.420	0.154	-0.420	0.154	-0.420	0.154	-0.482	0.492	-0.483	0.491	-0.483	0.491	-0.482	0.491
Sulphur dioxide	0.000	0.178	0.000	0.177	0.000	0.176	0.000	0.177	0.000	0.203	0.000	0.203	0.000	0.202	0.000	0.203
Sulphuric acid	-0.018	0.036	-0.018	0.036	-0.018	0.035	-0.018	0.036	0.014	0.393	0.014	0.392	0.014	0.391	0.014	0.392
Tetrahydrofuran	-0.801	0.458	-0.803	0.457	-0.802	0.458	-0.803	0.457	-2.740	0.640	-2.735	0.641	-2.741	0.640	-2.738	0.641
Toluene	0.092	0.563	0.091	0.566	0.091	0.566	0.091	0.566	**0.429**	0.160	**0.429**	0.159	**0.429**	0.159	**0.429**	0.159
Total Reduced Sulphur (TRS) - [MOE]	-0.426	0.001	-0.425	0.001	-0.426	0.001	-0.425	0.001	-0.404	0.157	-0.406	0.156	-0.404	0.158	-0.406	0.156
Trichloroethylene	**0.338**	0.329	**0.337**	0.330	**0.338**	0.330	**0.337**	0.330	-0.721	0.553	-0.719	0.554	-0.720	0.554	-0.720	0.554
Volatile Organic Compounds (VOCs)	0.001	0.703	0.001	0.705	0.001	0.702	0.001	0.706	-0.011	0.224	-0.011	0.225	-0.011	0.223	-0.011	0.225

Québec	SGA								LBWT							
Xylene (mixed isomers)	-0.006	0.962	-0.007	0.960	-0.006	0.962	-0.007	0.961	-0.551	0.198	-0.551	0.198	-0.551	0.198	-0.551	0.198
Crop	-0.033	0.218	-0.033	0.222	-0.033	0.219	-0.033	0.222	-0.176	0.004	-0.177	0.004	-0.176	0.004	-0.177	0.004
Dump	-0.177	0.031	-0.178	0.031	-0.178	0.031	-0.178	0.031	0.098	0.596	0.099	0.593	0.099	0.593	0.099	0.593
Gas station	0.084	0.000	0.084	0.000	0.084	0.000	0.084	0.000	0.105	0.001	0.106	0.001	0.105	0.001	0.105	0.001
Mine	0.014	0.836	0.014	0.839	0.014	0.839	0.014	0.838	0.043	0.782	0.043	0.783	0.044	0.779	0.044	0.781
Well pad	**6.674**	0.098	**6.670**	0.099	**6.698**	0.097	**6.675**	0.098	**17.602**	0.001	**17.590**	0.001	**17.553**	0.001	**17.603**	0.001
Powerline	-0.056	0.000	-0.056	0.000	-0.056	0.000	-0.056	0.000	-0.065	0.025	-0.065	0.025	-0.065	0.025	-0.065	0.025
Transformer	**0.442**	0.000	**0.442**	0.000	**0.442**	0.000	**0.442**	0.000	**0.349**	0.079	**0.349**	0.079	**0.349**	0.080	**0.350**	0.079

Saskatchewan Season	SGA Tri1		Tri2		Prior3		NONE		LBWT Tri1		Tri2		Prior3		NONE	
Variables	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Ammonia (total)	-0.068	0.097	-0.068	0.097	-0.068	0.097	-0.068	0.096	-0.082	0.403	-0.082	0.400	-0.082	0.402	-0.082	0.399
Carbon monoxide	-0.002	0.274	-0.002	0.274	-0.002	0.274	-0.002	0.273	0.000	0.990	0.000	0.989	0.000	0.991	0.000	0.988
Carbonyl sulphide	-0.345	0.258	-0.346	0.257	-0.346	0.257	-0.346	0.257	0.162	0.755	0.161	0.756	0.162	0.755	0.161	0.756
Hydrochloric acid	-0.944	0.109	-0.941	0.110	-0.944	0.109	-0.943	0.109	-0.367	0.653	-0.364	0.655	-0.367	0.653	-0.364	0.654
Hydrogen fluoride	**0.445**	0.749	**0.456**	0.742	**0.445**	0.749	**0.448**	0.747								
Nitrogen oxides (expressed as NO_2)	0.013	0.156	0.013	0.156	0.013	0.155	0.013	0.155	0.012	0.557	0.012	0.555	0.012	0.552	0.012	0.555
PM - Total Particulate Matter	0.007	0.844	0.007	0.840	0.007	0.843	0.007	0.845	-0.021	0.815	-0.021	0.815	-0.020	0.819	-0.021	0.814
PM_{10} - Particulate Matter ≤ 10 Microns	-0.037	0.318	-0.037	0.320	-0.037	0.319	-0.037	0.317	-0.098	0.284	-0.098	0.284	-0.098	0.287	-0.098	0.284
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	-0.031	0.705	-0.031	0.709	-0.031	0.703	-0.032	0.701	-0.112	0.579	-0.112	0.578	-0.112	0.581	-0.113	0.576
Sulphur dioxide	0.000	0.888	0.000	0.889	0.000	0.891	0.000	0.890	0.002	0.739	0.002	0.737	0.002	0.735	0.002	0.737
Total Reduced Sulphur (TRS) - [MOE]	-0.180	0.778	-0.180	0.778	-0.180	0.777	-0.181	0.776	**0.542**	0.614	**0.541**	0.615	**0.544**	0.613	**0.540**	0.615
Volatile Organic Compounds (VOCs)	-0.010	0.057	-0.010	0.057	-0.010	0.056	-0.010	0.056	-0.019	0.130	-0.019	0.129	-0.019	0.130	-0.019	0.129
Crop	0.073	0.264	0.073	0.264	0.073	0.263	0.073	0.262	0.072	0.626	0.072	0.624	0.072	0.625	0.072	0.623
Dump	**0.587**	0.018	**0.586**	0.018	**0.587**	0.018	**0.586**	0.018	**1.249**	0.019	**1.247**	0.020	**1.250**	0.019	**1.247**	0.020
Gas station	**0.203**	0.000	**0.203**	0.000	**0.203**	0.000	**0.203**	0.000	0.137	0.154	0.137	0.155	0.138	0.153	0.137	0.155
Mine	-0.217	0.336	-0.216	0.339	-0.218	0.335	-0.218	0.335	-0.338	0.531	-0.337	0.531	-0.338	0.530	-0.338	0.530
Well pad	0.098	0.044	0.098	0.043	0.098	0.044	0.098	0.044	0.084	0.442	0.084	0.440	0.084	0.440	0.084	0.441

Saskatchewan	SGA								LBWT							
Powerline	-0.228	0.018	-0.227	0.018	-0.228	0.018	-0.228	0.018	-0.636	0.010	-0.636	0.010	-0.637	0.010	-0.637	0.010
Transformer	**0.733**	0.000	**0.732**	0.000	**0.732**	0.000	**0.732**	0.000	**0.235**	0.596	**0.234**	0.598	**0.235**	0.596	**0.233**	0.599

Yukon	SGA								LBWT							
Season	Tri1		Tri2		Prior3		NONE		Tri1		Tri2		Prior3		NONE	
Variables	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p	Beta	p
Carbon monoxide	**0.433**	0.260	**0.439**	0.254	**0.428**	0.265	**0.428**	0.265	**1.142**	0.073	**1.082**	0.087	**1.088**	0.085	**1.090**	0.085
Nitrogen oxides (expressed as NO_2)	0.085	0.323	0.087	0.312	0.086	0.318	0.085	0.322	**0.256**	0.067	**0.249**	0.074	**0.250**	0.072	**0.250**	0.072
PM_{10} - Particulate Matter ≤ 10 Microns	**1.530**	0.210	**1.548**	0.205	**1.534**	0.208	**1.527**	0.211	**3.877**	0.050	**3.773**	0.054	**3.782**	0.053	**3.785**	0.053
$PM_{2.5}$ - Particulate Matter ≤ 2.5 Microns	**1.548**	0.204	**1.566**	0.199	**1.554**	0.201	**1.546**	0.204	**3.861**	0.051	**3.760**	0.055	**3.769**	0.055	**3.771**	0.055
Dump	0.033	0.990	-0.039	0.988	0.074	0.977	0.032	0.990	-6.761	0.350	-6.740	0.353	-6.794	0.348	-6.813	0.347
Gas station	-1.265	0.104	-1.282	0.099	-1.289	0.098	-1.272	0.102	-4.905	0.009	-4.889	0.009	-4.922	0.008	-4.904	0.009
Mine	-2.293	0.167	-2.319	0.163	-2.373	0.154	-2.326	0.162	**2.436**	0.458	**2.367**	0.476	**2.385**	0.474	**2.408**	0.469
Powerline	-0.199	0.795	-0.192	0.801	-0.178	0.816	-0.179	0.815	-1.712	0.309	-1.618	0.336	-1.619	0.335	-1.621	0.334
Transformer	**5.207**	0.490	**5.200**	0.491	**5.211**	0.490	**5.213**	0.489								

Appendix II

Table S2.1. Customized azimuthal equidistant map projection parameters were based on the centroids of each Census Metropolitan Area (CMA), designated in decimal degrees of longitude (central meridian) and latitude (origin).

Census Metropolitan Area (CMA)	Area (km^2)	Custom Map Projection Parameters	Central Meridian (X)	Latitude of Origin (Y)
Calgary	5,108		-114.078155	51.180782
Edmonton	9,427	Projection: Azimuthal Equidistant	-113.789137	53.512964
Fredericton	4,886	False Easting: 0.0	-66.706042	46.111538
Halifax	5,496	False Northing: 0.0	-63.125663	44.839496
Hamilton	1,372	Central Meridian: **(X)**	-79.922040	43.266381
Kingston	1,939	Latitude Of Origin: **(Y)**	-76.542139	44.340704
London	2,666	Linear Unit: Meter (1.0)	-81.312858	42.901574
Moncton	2,406		-64.864306	46.025620
Montréal	4,258	Geographic Coordinate System:	-73.734869	45.567259
Ottawa	6,287	GCS North American 1983	-75.765999	45.452371
Québec	3,349	Angular Unit: Degree	-71.305923	46.927416
Regina	3,408	(0.0174532925199433)	-104.718907	50.519386
Saint John	3,363	Prime Meridian: Greenwich (0.0)	-66.010434	45.375455
Saskatoon	5,215	Datum: D North American 1983	-106.603066	52.069258
St. John's	805	Spheroid: GRS 1980	-52.810751	47.517317
Toronto	5,906	Semimajor Axis: 6378137.0	-79.579271	43.900699
Vancouver	2,883	Semiminor Axis: 6356752.3141403	-122.950102	49.258531
Victoria	696	Inverse Flattening: 298.257222101	-123.552401	48.397793
Winnipeg	5,303		-97.047831	49.888087

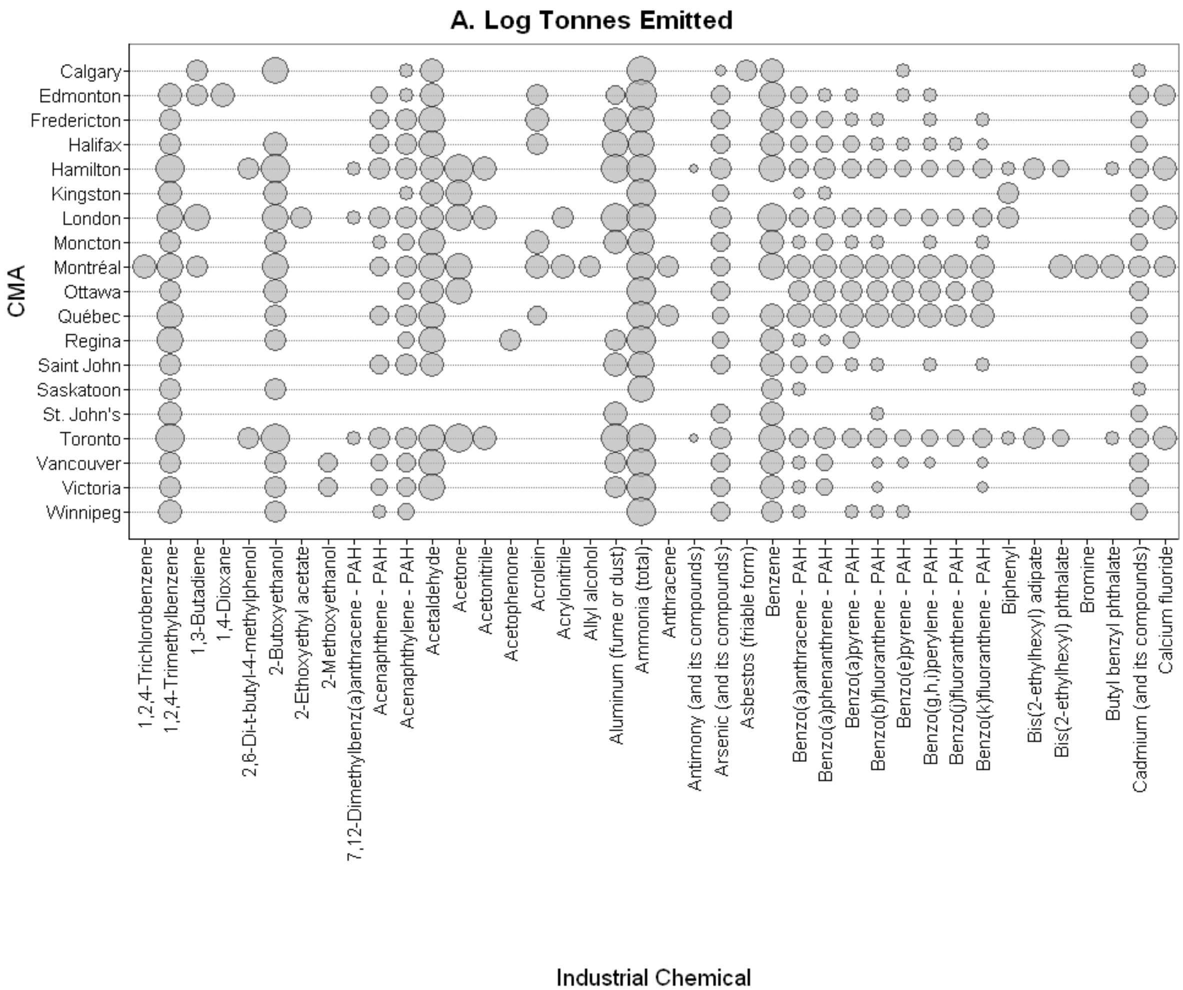


Figure S2.1. Total emissions of industrial chemical reported by the National Pollutant Release Inventory (NPRI) for each Census Metropolitan Area (CMA). Figures A through D indicate the relative tonnes emitted – in log scaled proportional symbols for ease of visualization; absence of a symbol indicates the CMA did not emit that particular chemical. Chemicals are listed in alphabetical order.

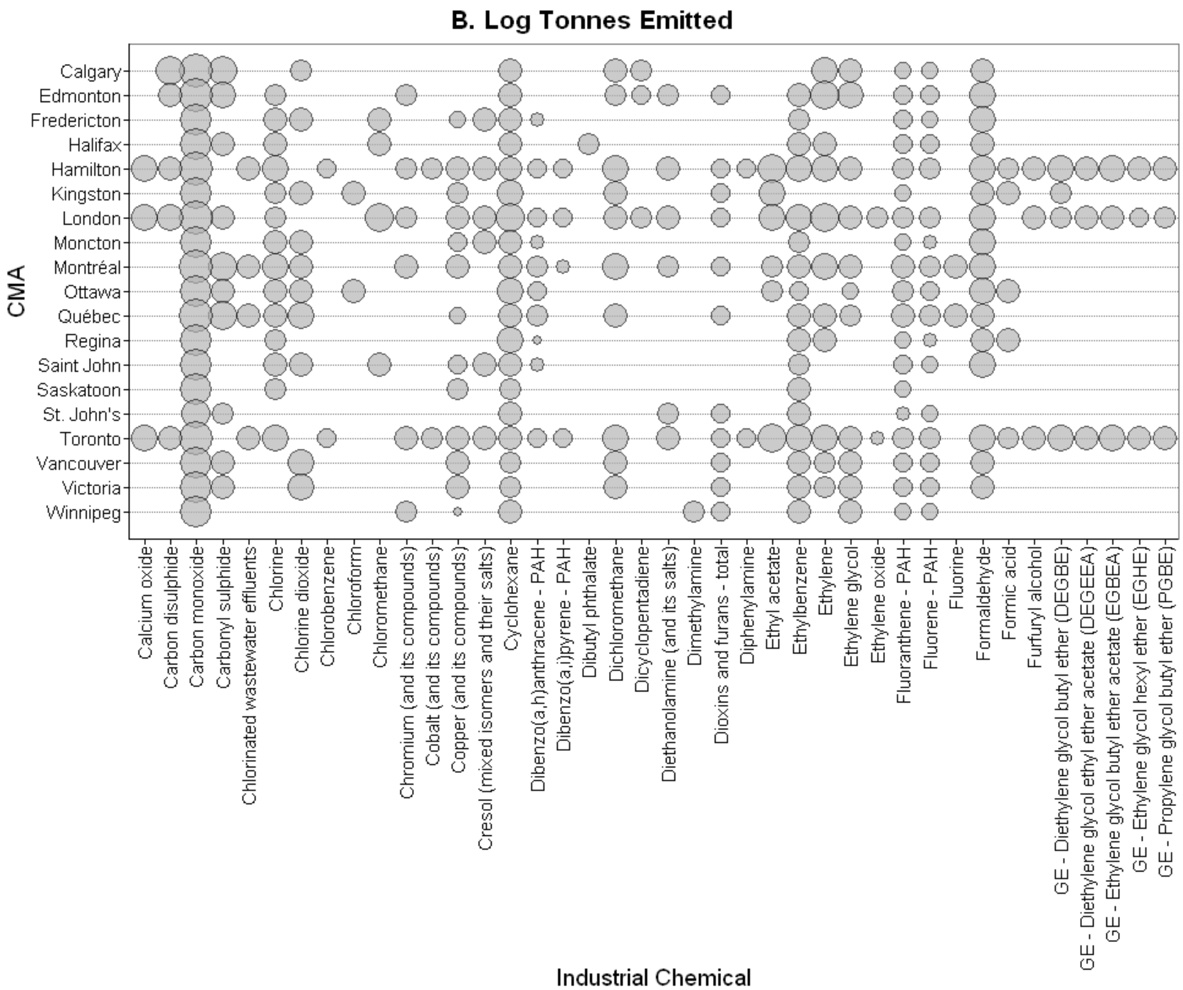


Figure S2.2. Total emissions of industrial chemical reported by the National Pollutant Release Inventory (NPRI) for each Census Metropolitan Area (CMA). Figures A through D indicate the relative tonnes emitted – in log scaled proportional symbols for ease of visualization; absence of a symbol indicates the CMA did not emit that particular chemical. Chemicals are listed in alphabetical order.

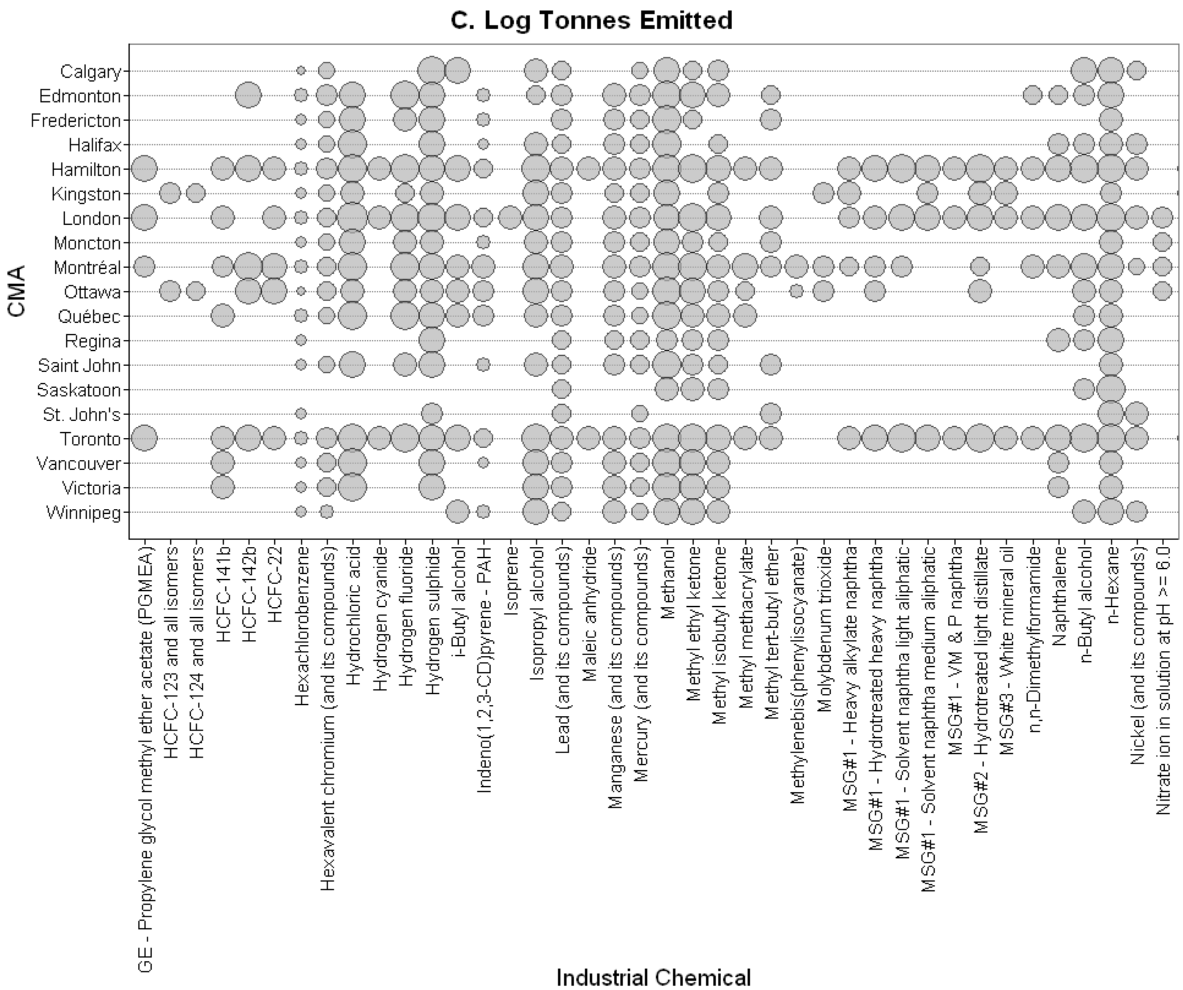


Figure S2.3. Total emissions of industrial chemical reported by the National Pollutant Release Inventory (NPRI) for each Census Metropolitan Area (CMA). Figures A through D indicate the relative tonnes emitted – in log scaled proportional symbols for ease of visualization; absence of a symbol indicates the CMA did not emit that particular chemical. Chemicals are listed in alphabetical order.

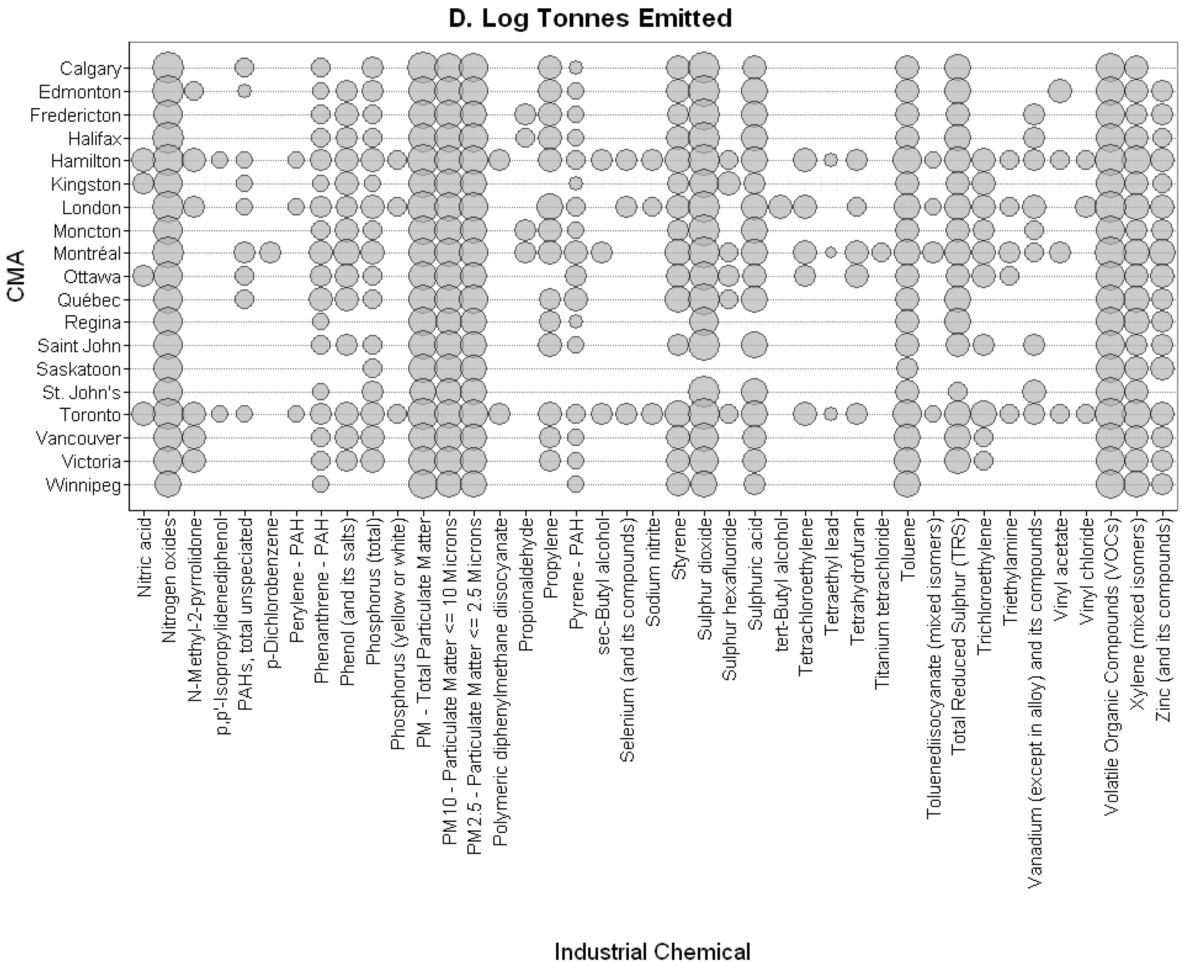
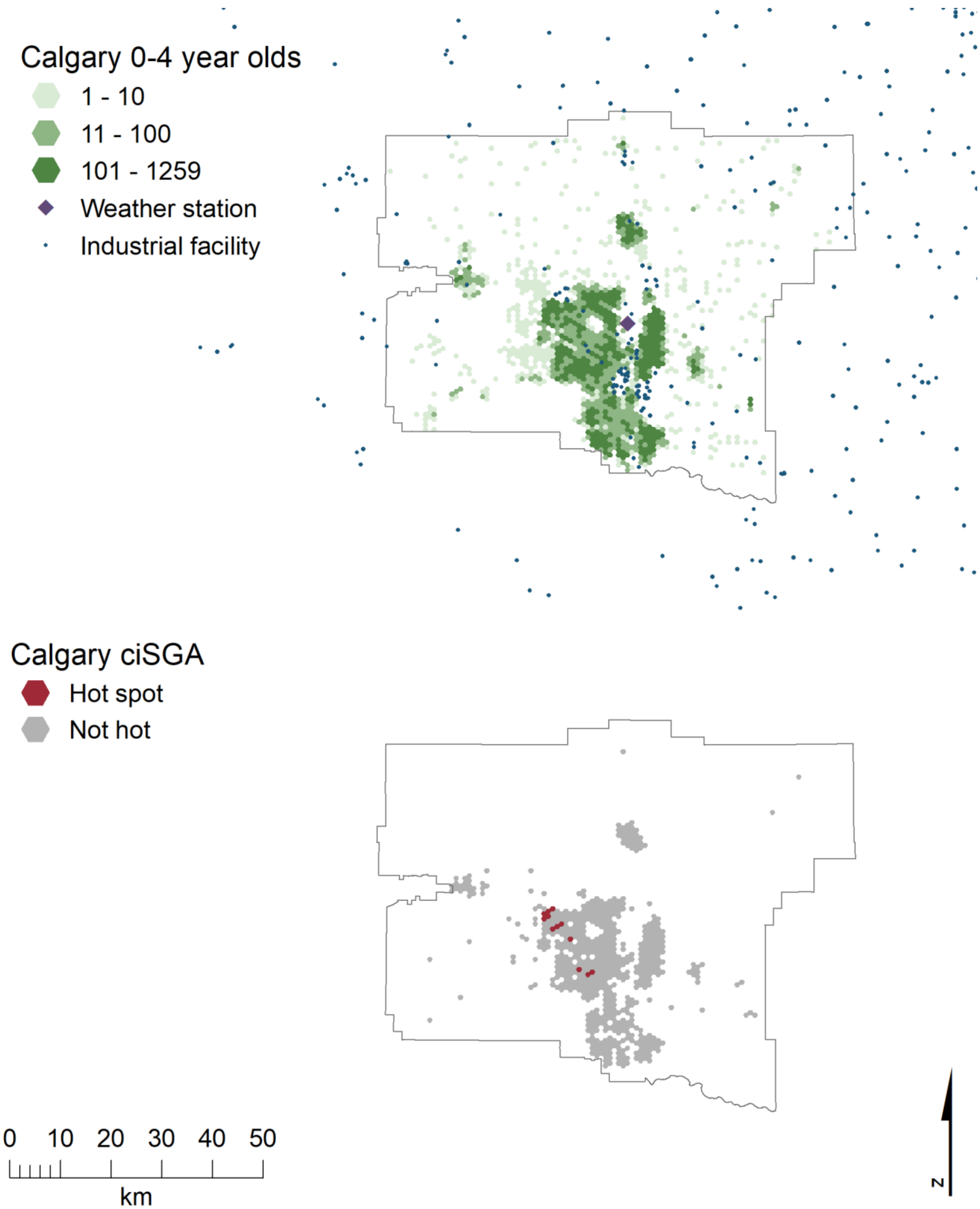


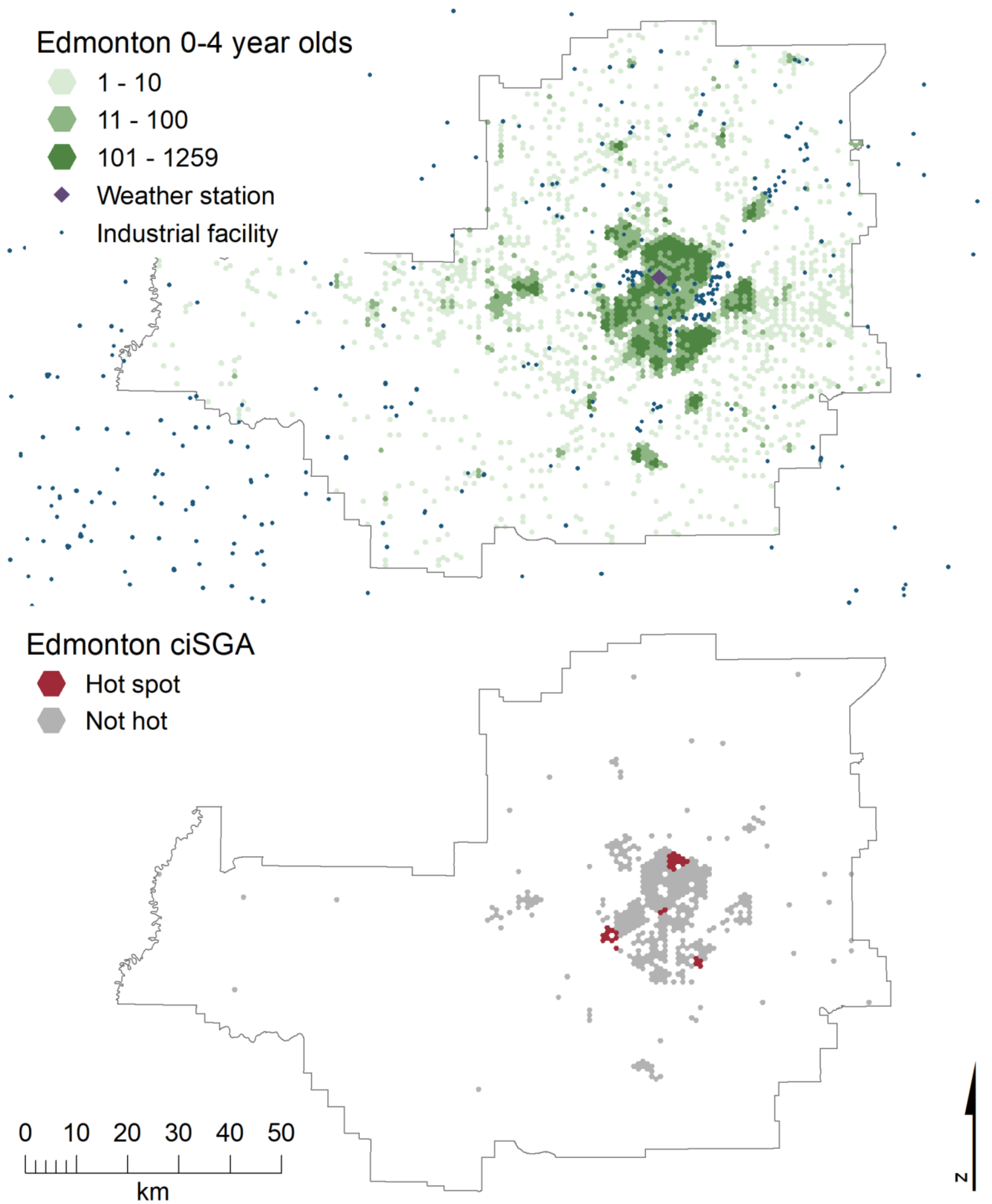
Figure S2.4. Total emissions of industrial chemical reported by the National Pollutant Release Inventory (NPRI) for each Census Metropolitan Area (CMA). Figures A through D indicate the relative tonnes emitted – in log scaled proportional symbols for ease of visualization; absence of a symbol indicates the CMA did not emit that particular chemical. Chemicals are listed in alphabetical order.

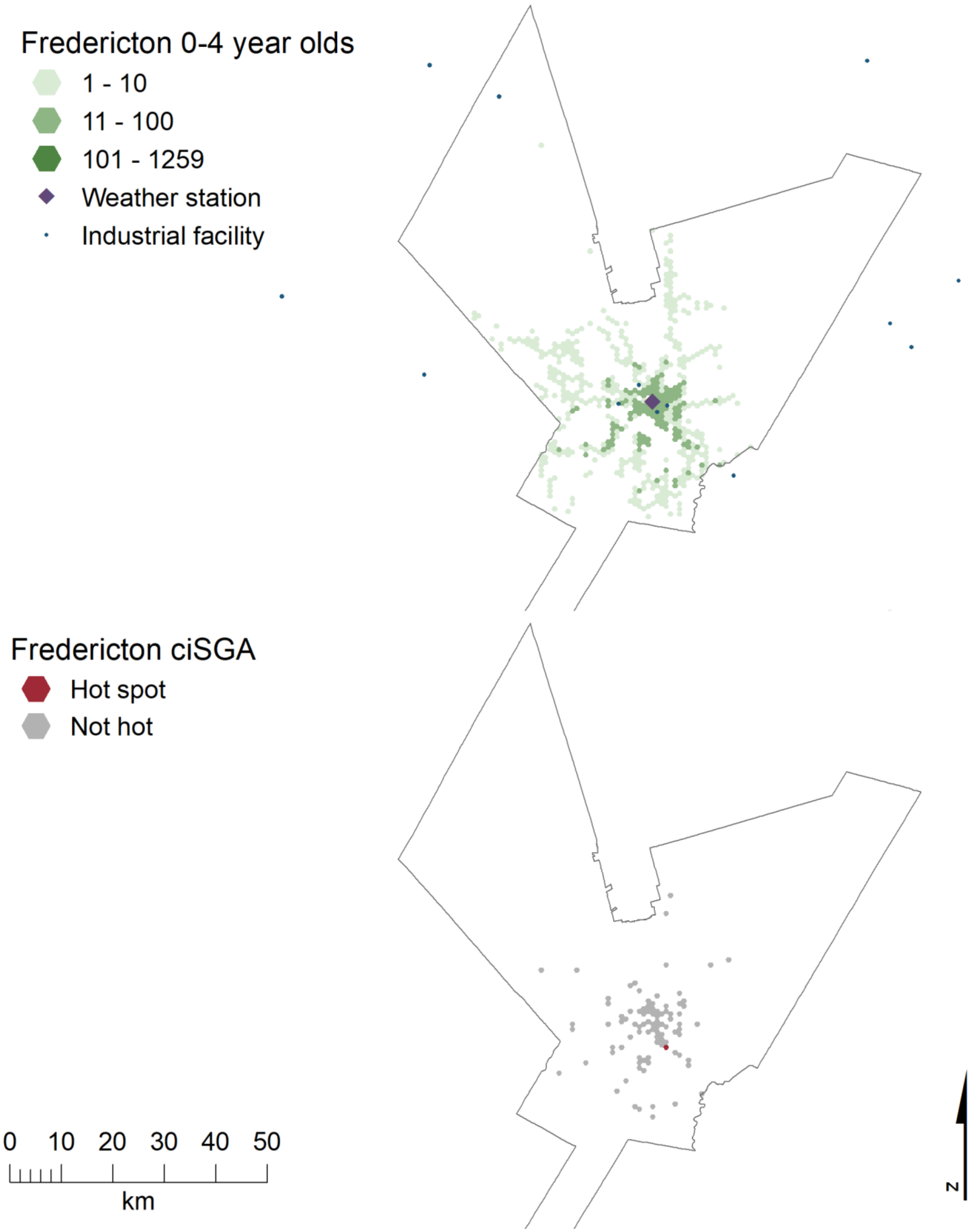
[MULTIPLE FIGURES FOLLOW BELOW AND ARE IDENTIFIED BY CMA NAME]

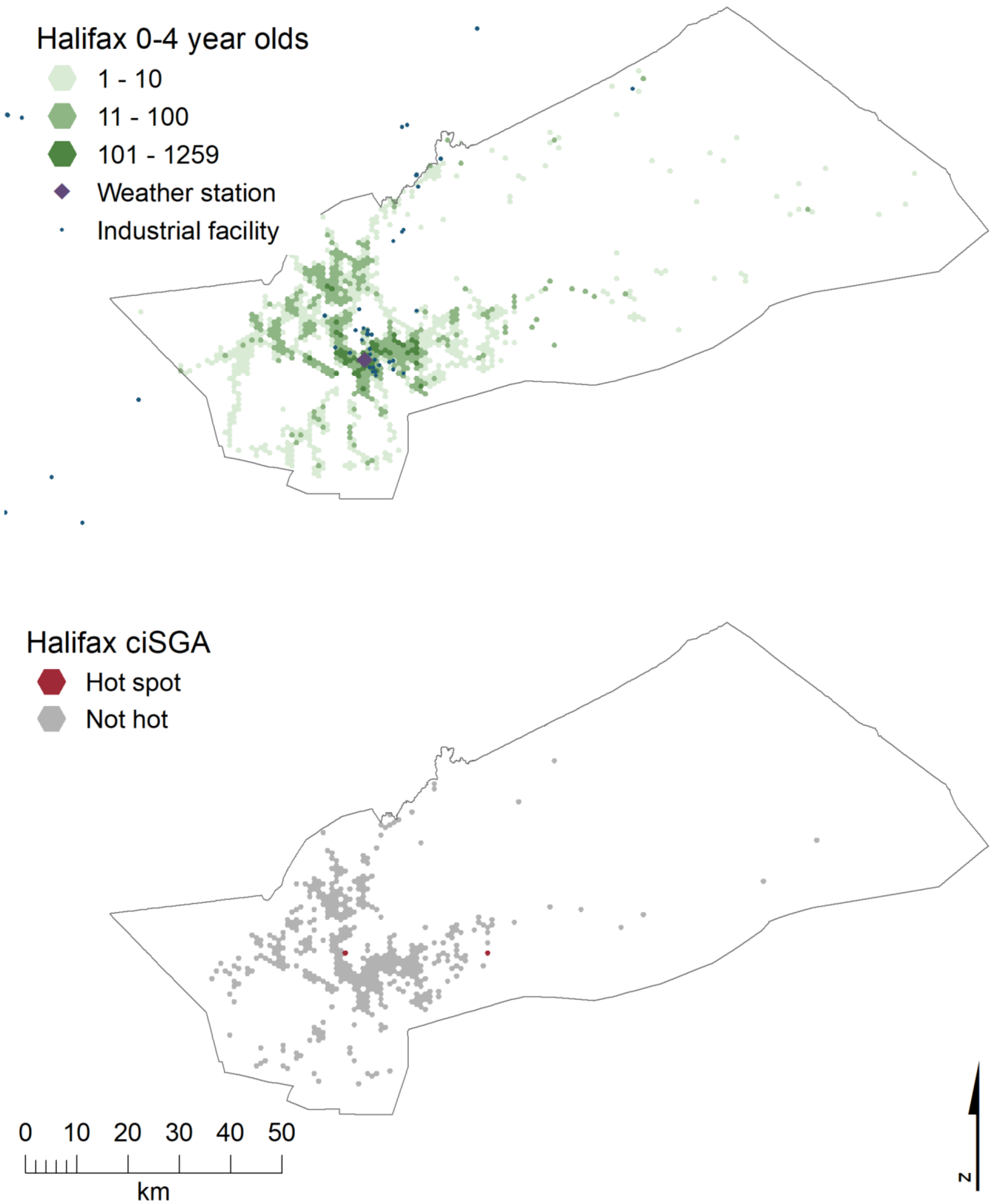
Figure S2.5. Distribution of infant counts (green hexagons), weather stations (purple diamonds), industrial facilities (blue dots), and space-time hot spot maps of critically ill SGA (ciSGA; red=hot, grey=not). Refer to Figure 5.1 for the location in Canada of each Census Metropolitan Area (CMA).

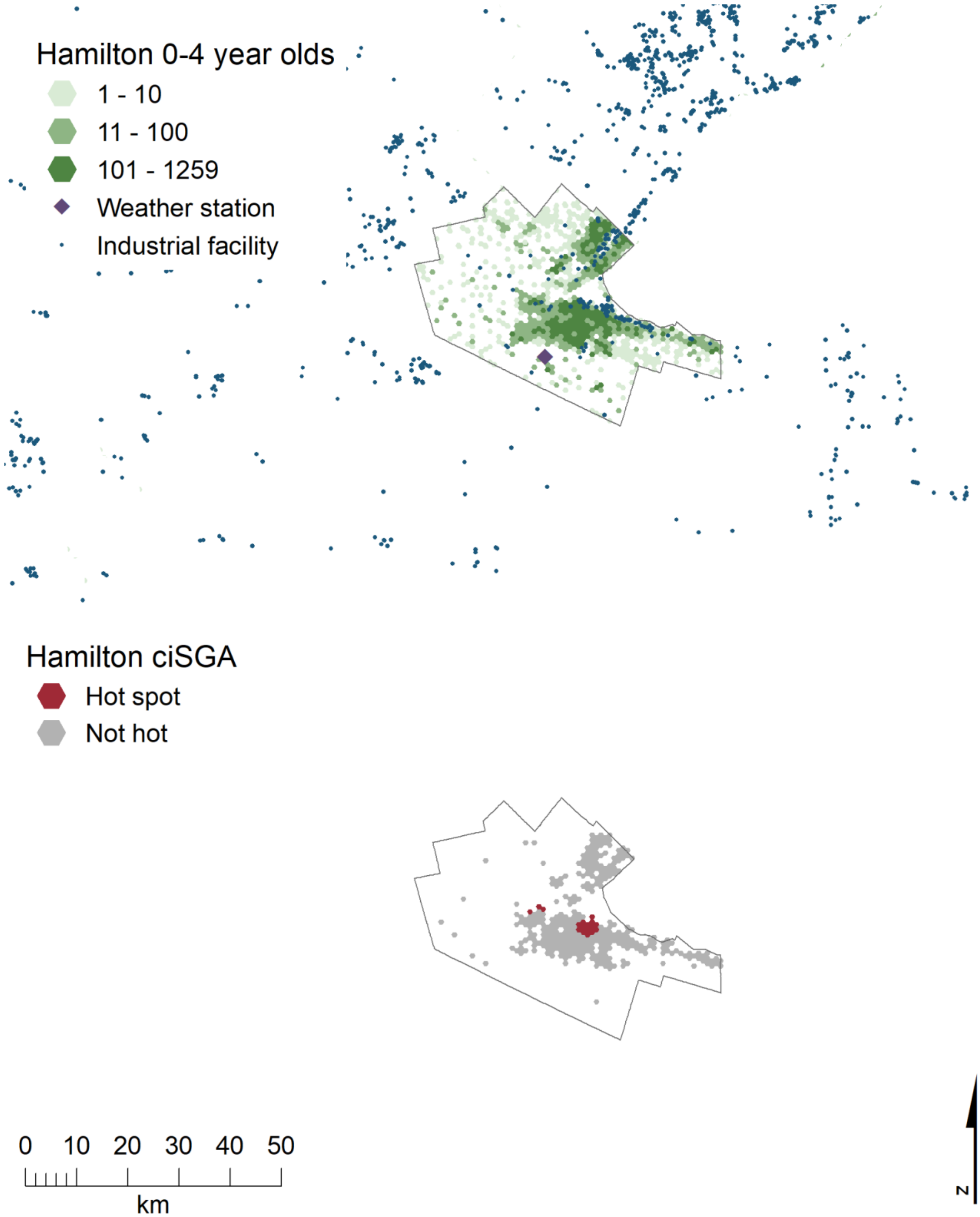


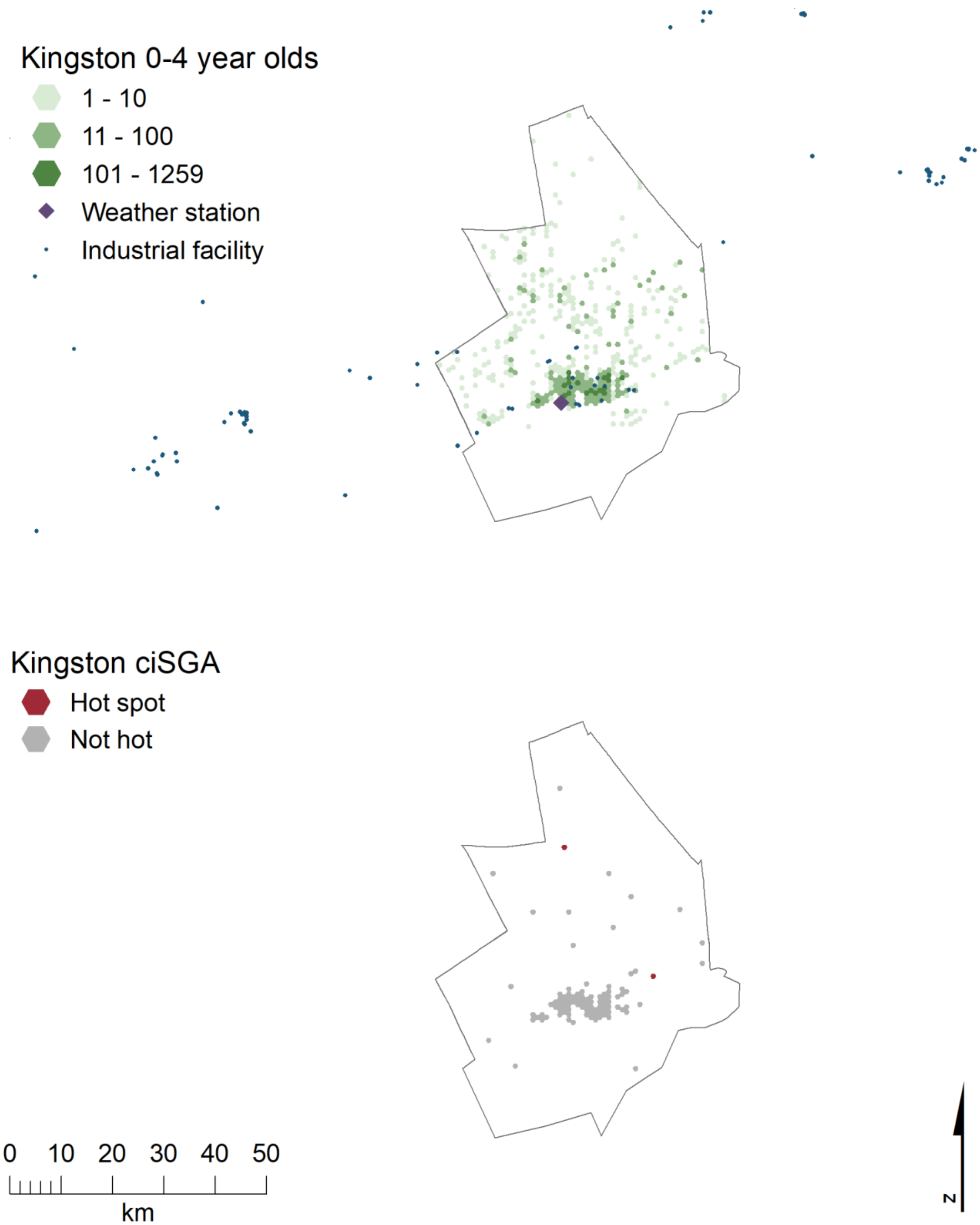
Calgary 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Calgary ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

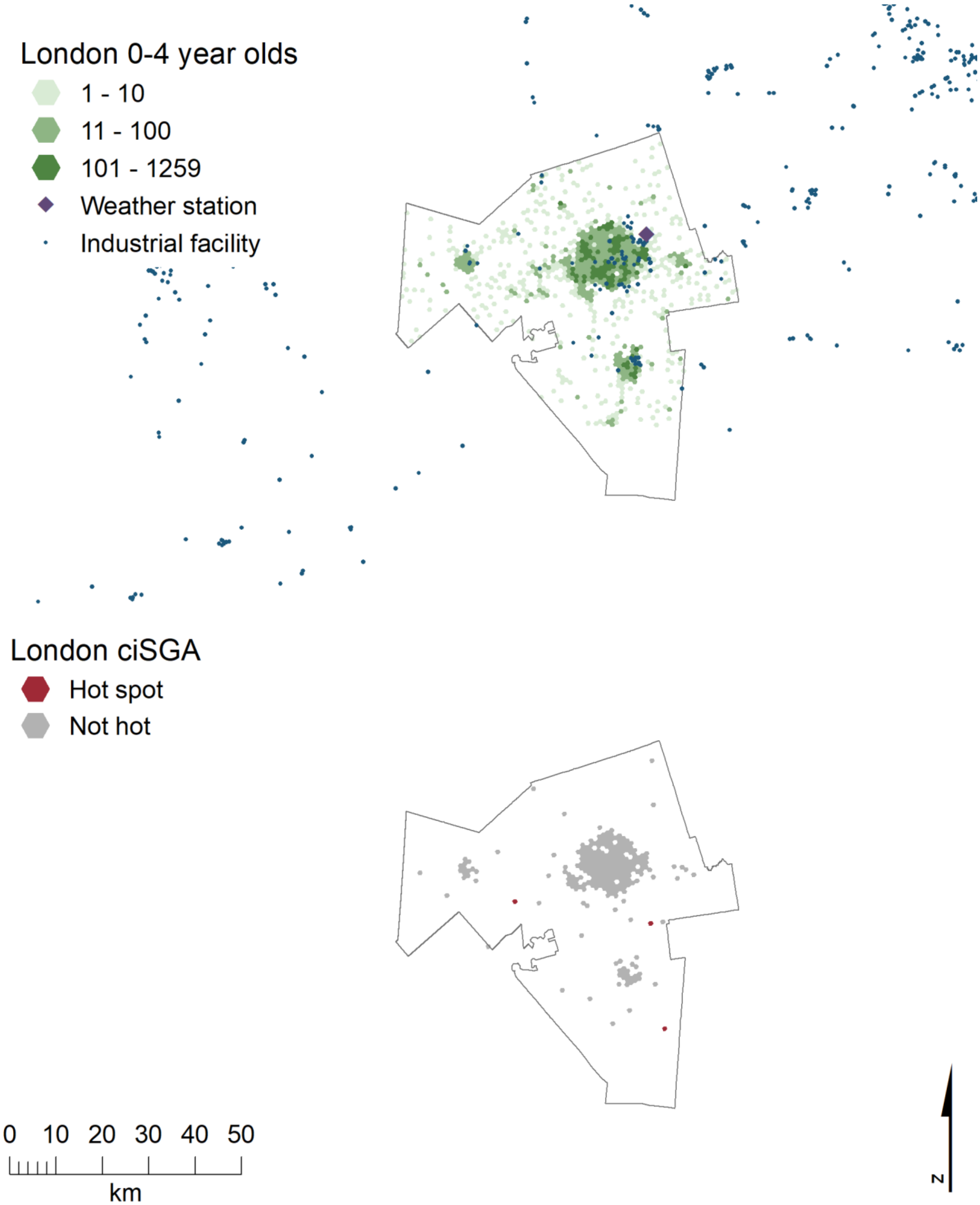

Edmonton 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Edmonton ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

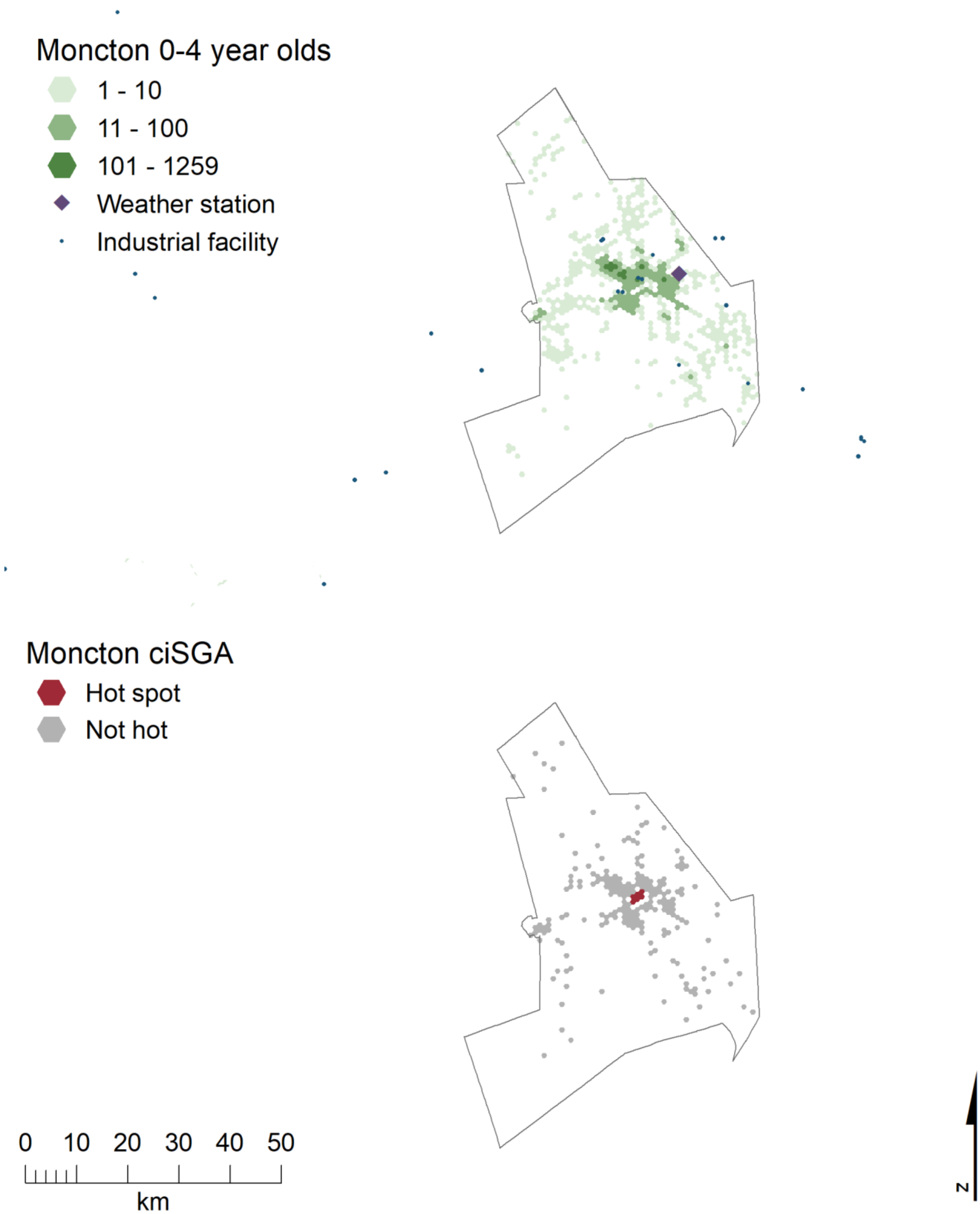

Fredericton 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Fredericton ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


Halifax 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Halifax ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

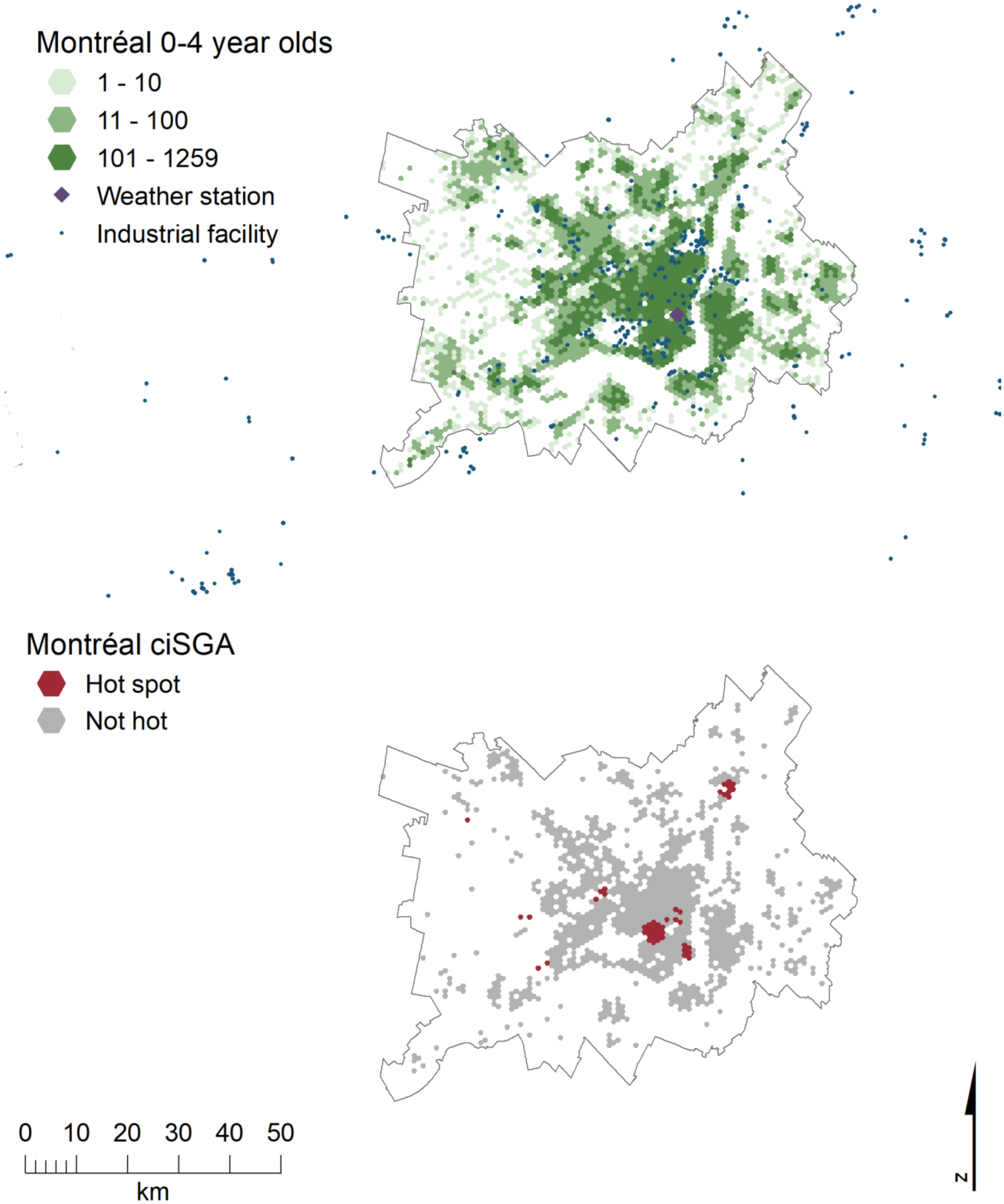

Hamilton 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Hamilton ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


Kingston 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Kingston ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

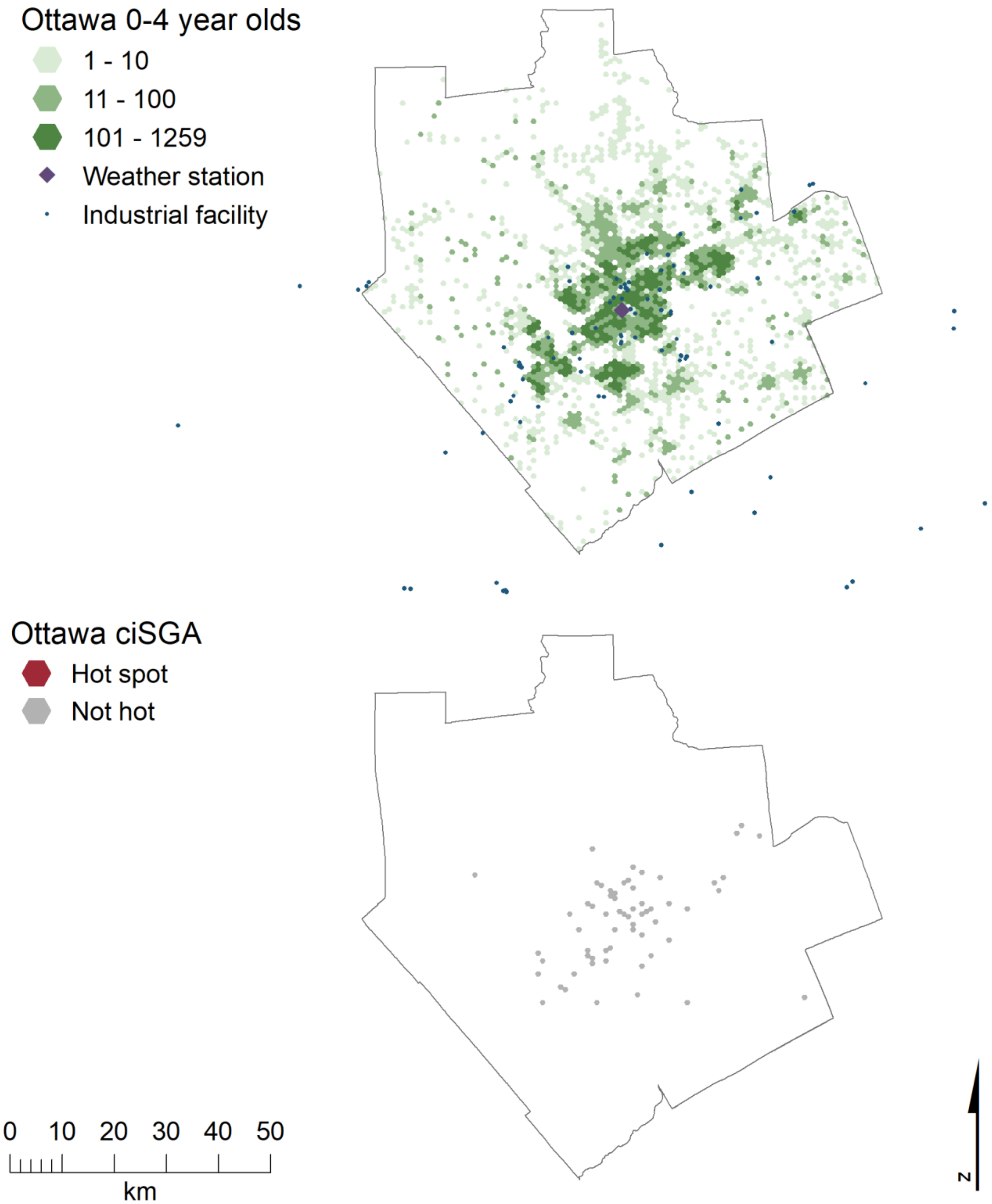

London 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
London ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

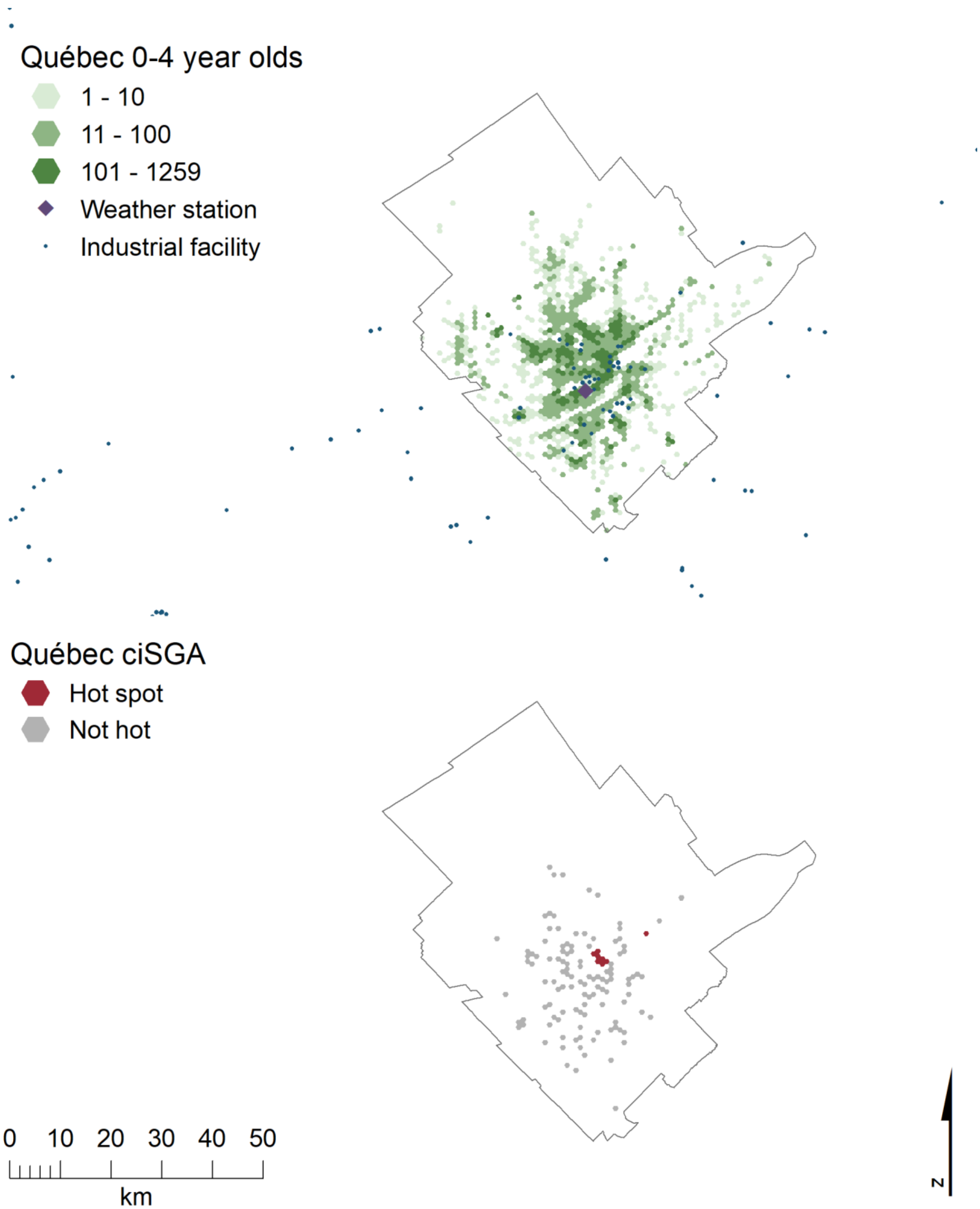


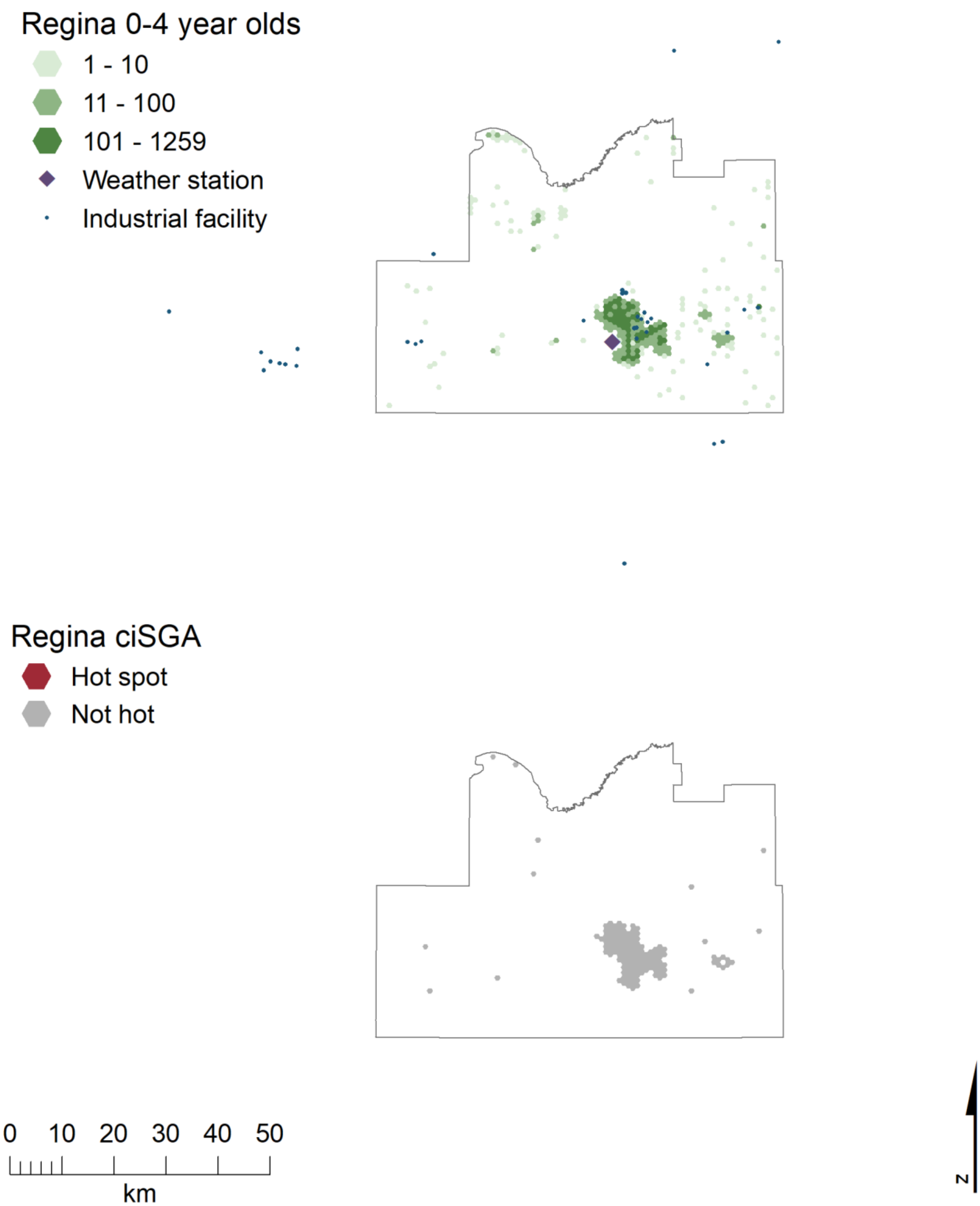
Moncton 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Moncton ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

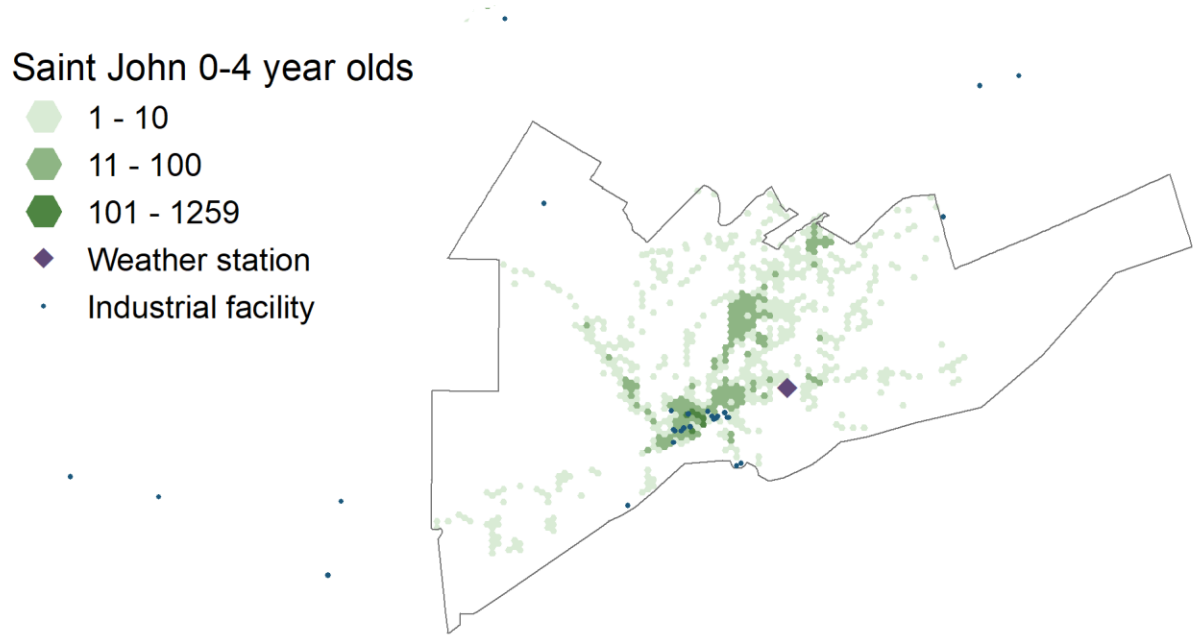


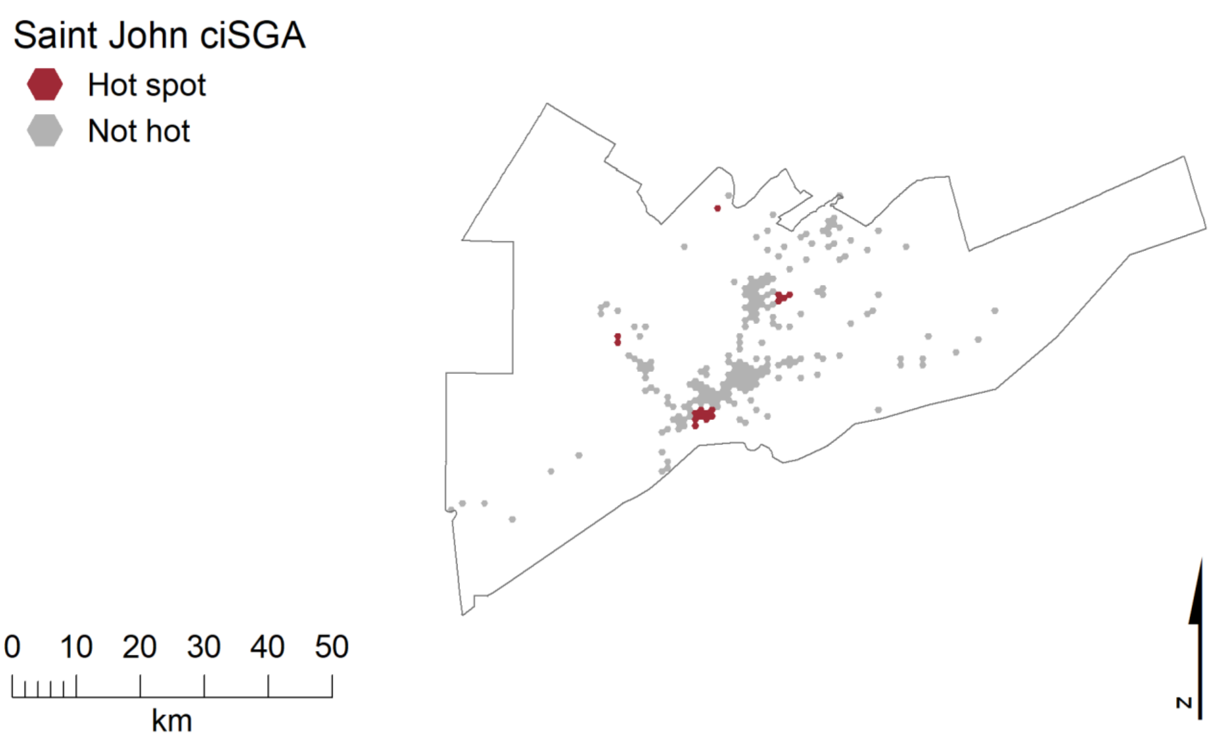
Montréal 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Montréal ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

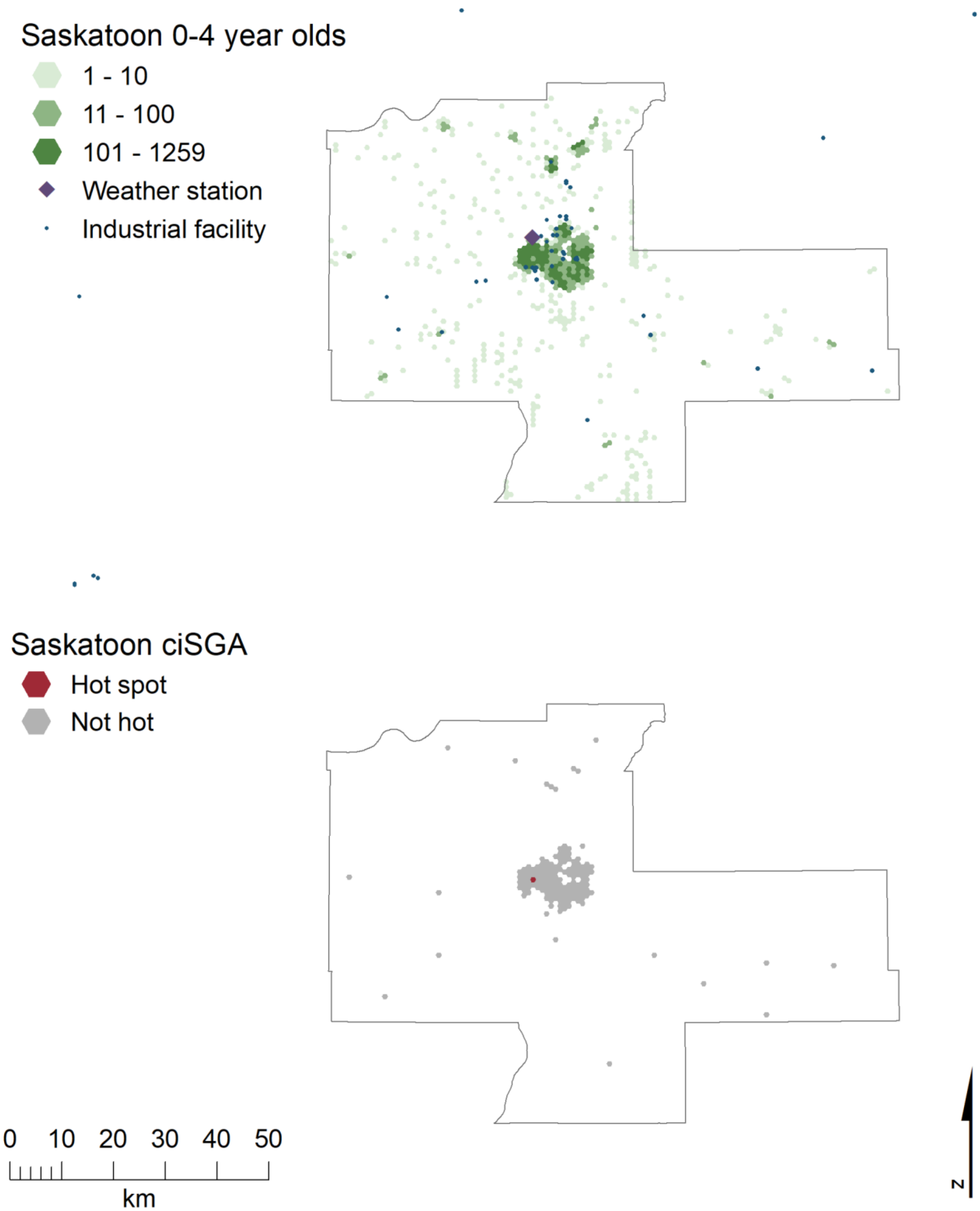

Ottawa 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Ottawa ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

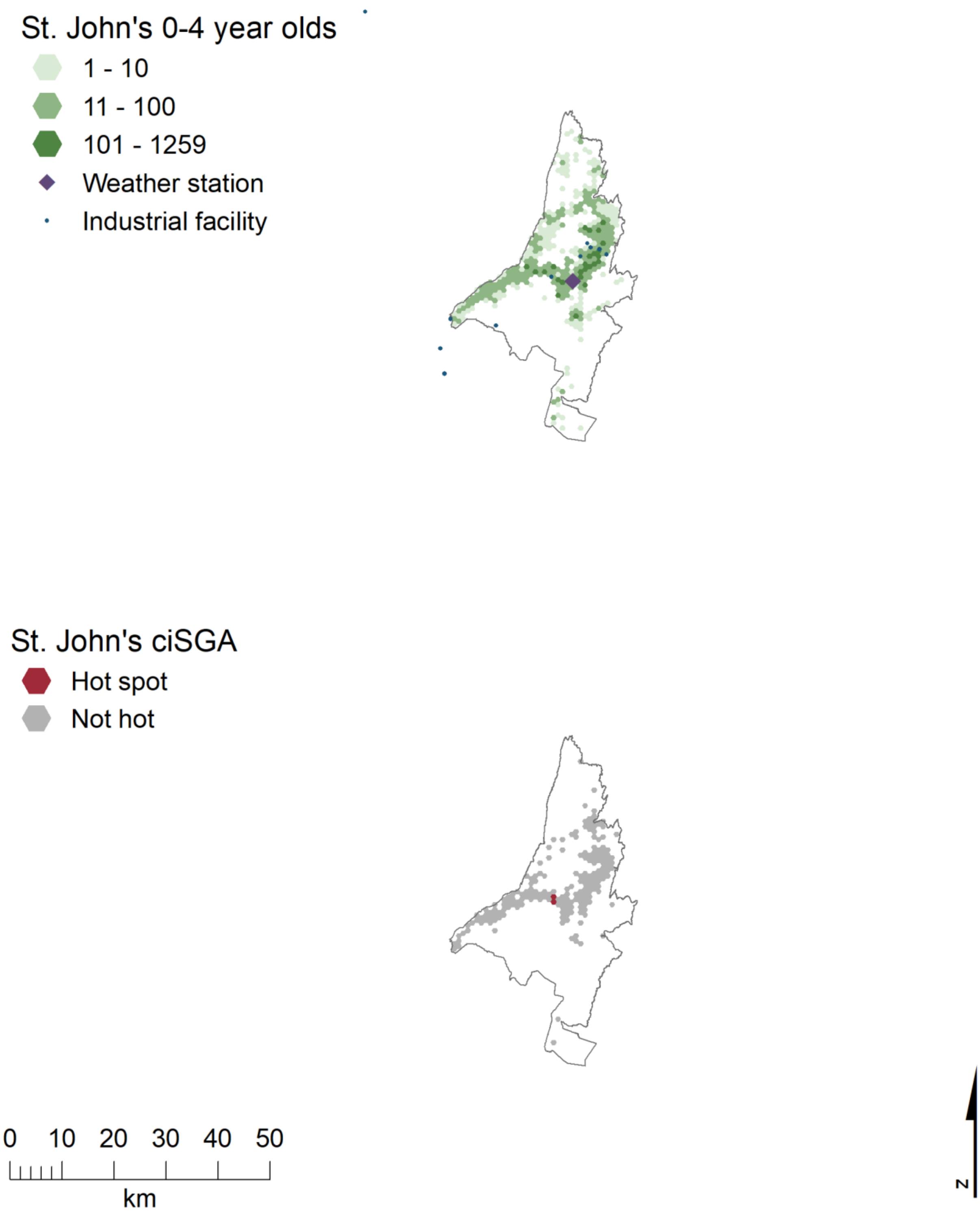

Québec 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Québec ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

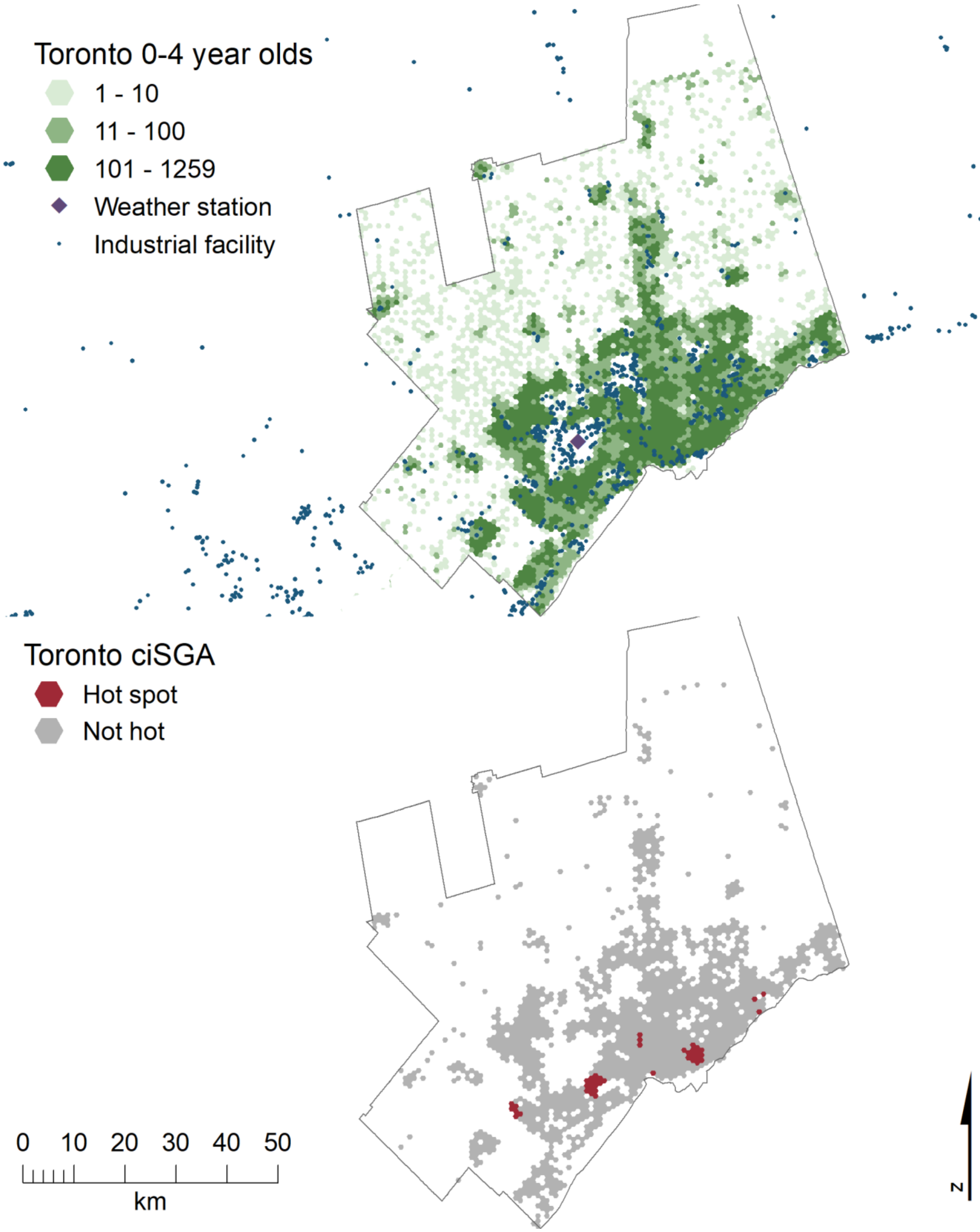

Regina 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Regina ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

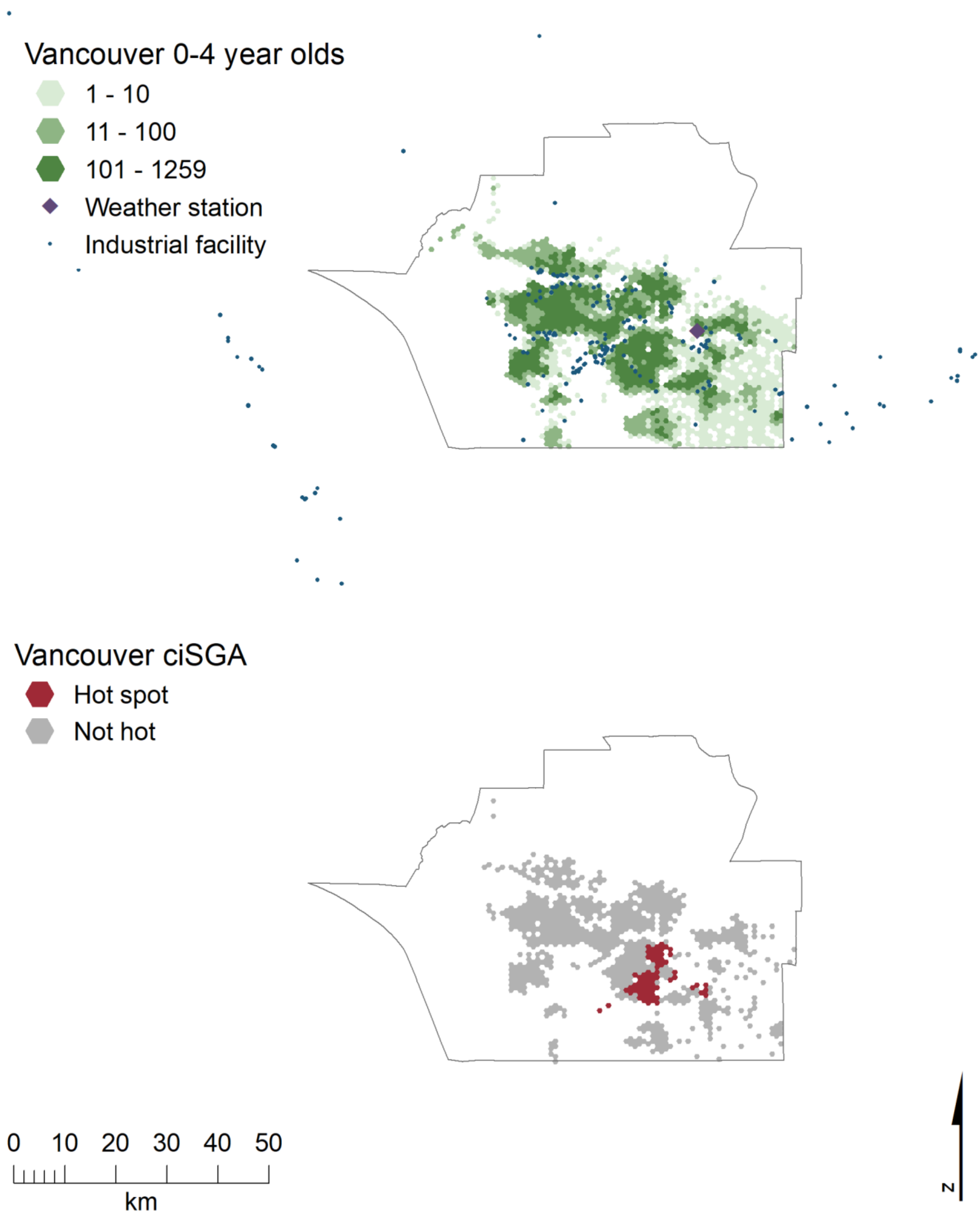

Saint John 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility

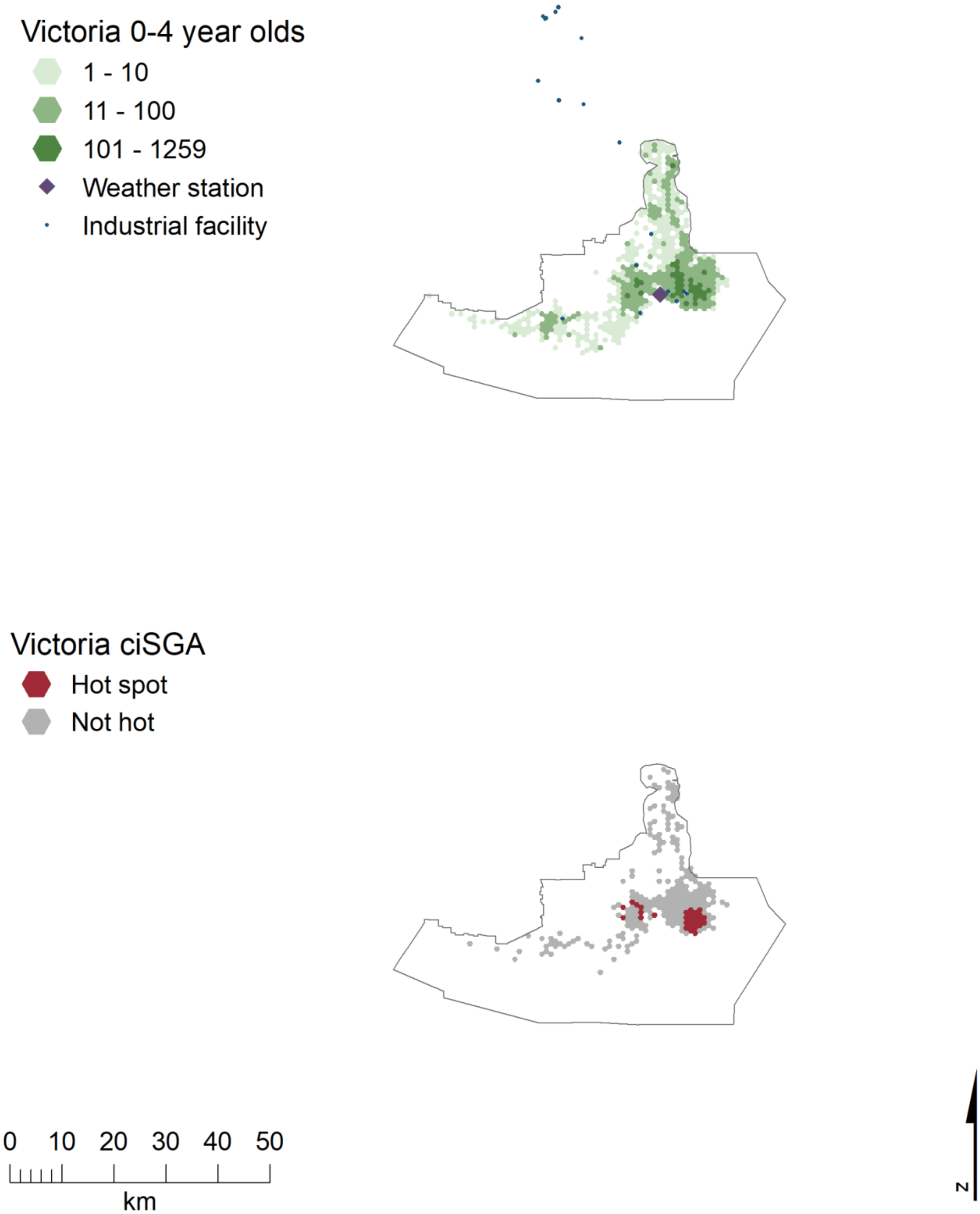

Saint John ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

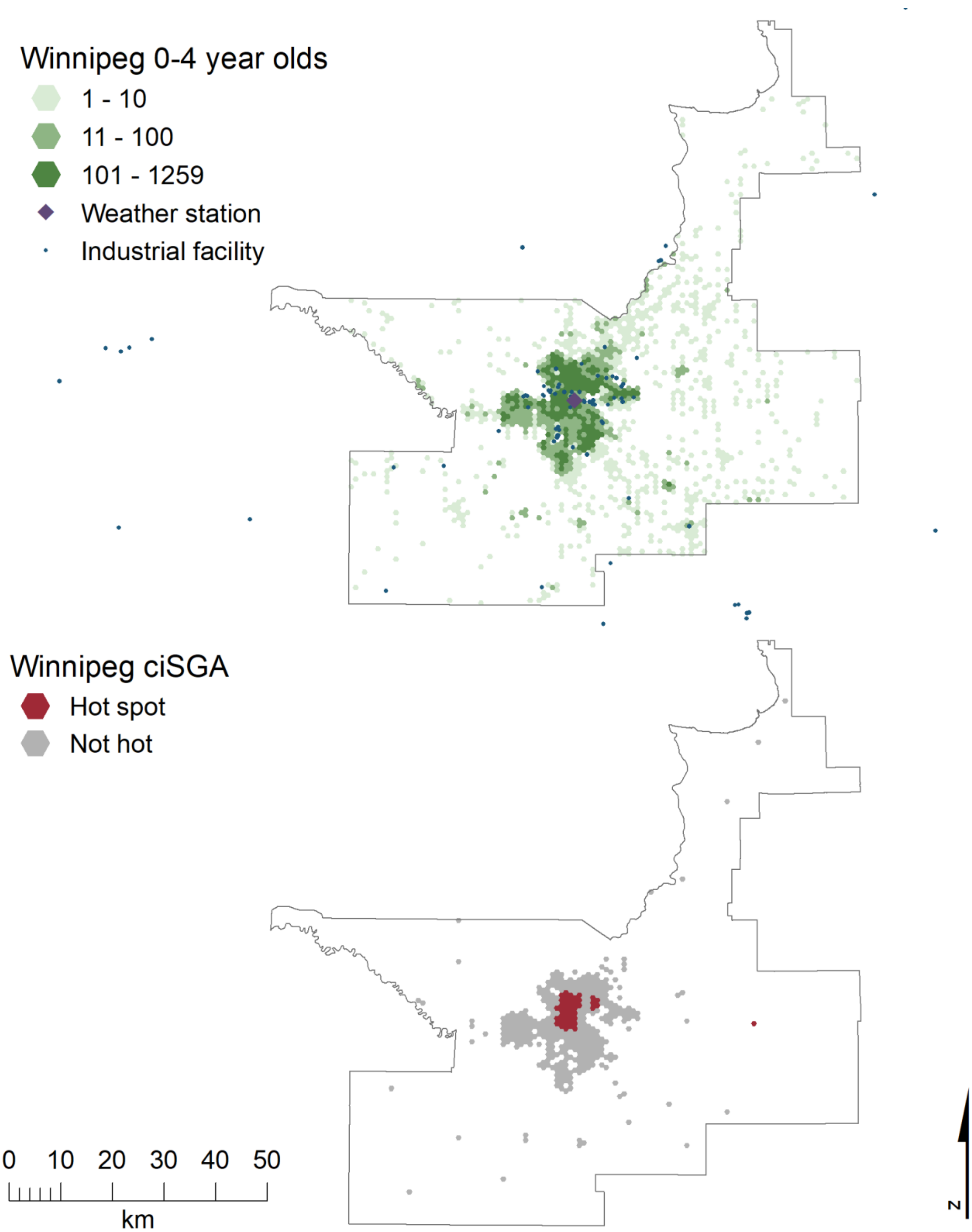

Saskatoon 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Saskatoon ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


St. John's 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
St. John's ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


Toronto 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Toronto ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


Vancouver 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Vancouver ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N


Victoria 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Victoria ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N



Winnipeg 0-4 year olds
1 - 10
11 - 100
101 - 1259
Weather station
Industrial facility
Winnipeg ciSGA
Hot spot
Not hot
0 10 20 30 40 50
km
N

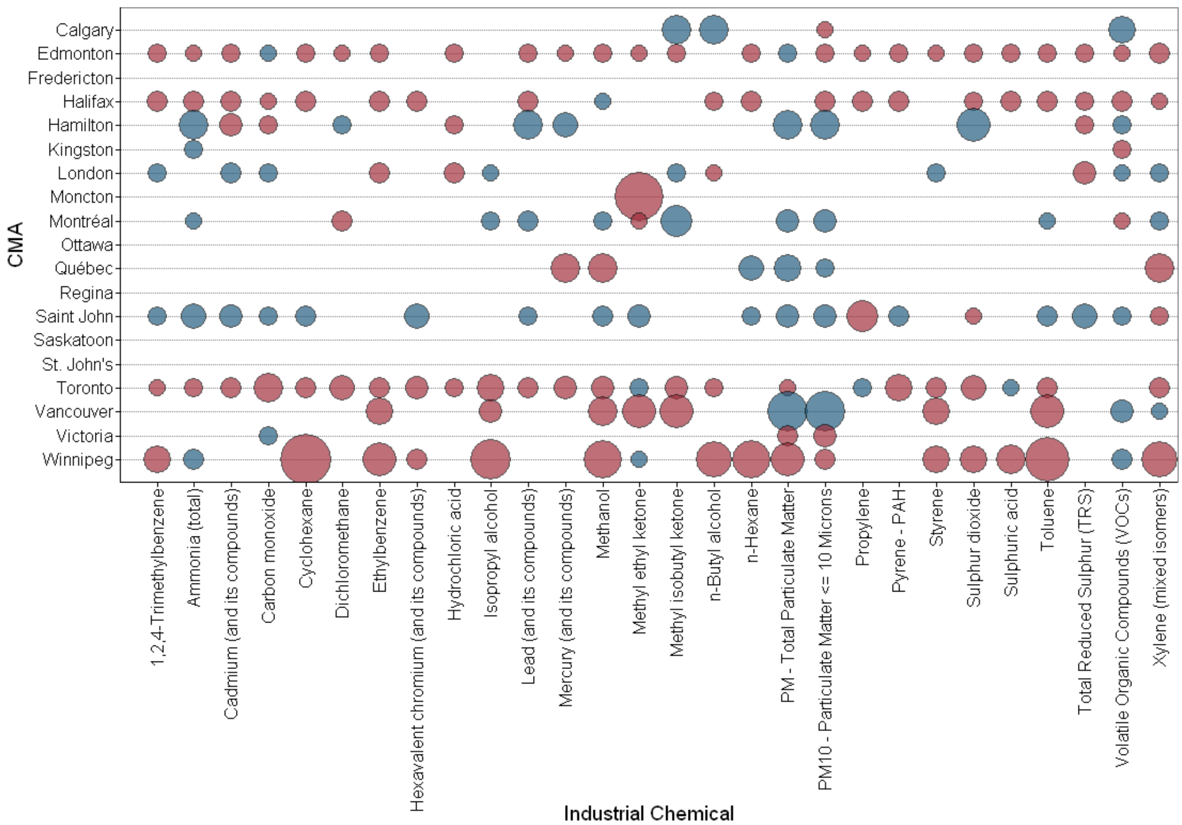


Figure S2.3. Subset of beta coefficients from Figure 5.9 (symbol key: red=positive; blue=negative; size=strength or relative magnitude of coefficient) showing the 28 industrial chemicals listed in Table 5.5 that had positive associations with critically ill small for gestational age (ciSGA) in 3 or more Census Metropolitan Areas (CMAs).

Appendix III

ETHICS LETTER

hero@ualberta.ca 25 July, 2013 3:02 PM
To: osornio@ualberta.ca
Reply-To: DoNotReply@ais.ualberta.ca
HERO: Your Ethics Application is Approved Pro00039545

ORIGINAL

Ethics Application has been Approved

ID:	Pro00039545
Title:	Spatial data mining exploring co-location of adverse birth outcomes and environmental variables
Study Investigator:	Alvaro Osornio Vargas
Description:	This is to inform you that the above study has been approved.
	Click on the link(s) above to navigate to the HERO workspace.
	Note: Please be reminded that the REMO system works best with Internet Explorer or Firefox.
	Please do not reply to this message. This is a system-generated email that cannot receive replies.

University of Alberta
Edmonton Alberta
Canada T6G 2E1

Contact Us | Privacy Policy | City of Edmonton

From: hero@ualberta.ca
Subject: HERO: An Amendment or Renewal has been Approved Pro00039545_REN6
Date: May 31, 2018 at 15:15
To: osornio@ualberta.ca

RENEWAL

Amendment/Renewal to Study has been Approved

Amendment/Renewal ID:	Pro00039545_REN6
Study ID:	MS11_Pro00039545
Study Title:	Spatial data mining exploring co-location of adverse birth outcomes and environmental variables
Study Investigator:	Alvaro Osornio Vargas
Description:	The amendment/renewal to the above study has been approved.
	Click on the link(s) above to navigate to the HERO workspace.
	Please do not reply to this message. This is a system-generated email that cannot receive replies.

University of Alberta
Edmonton Alberta
Canada T6G 2E1

Contact Us | Privacy Policy | City of Edmonton