University of Alberta

Determinants of Airborne Concentrations of Volatile Organic Compounds in Rural Areas of Western Canada

by

Xiaoqing You ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Medical Sciences – Public Health Sciences

Edmonton, Alberta
Spring 2006

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-81311-9
Our file Notre référence
ISBN: 978-0-494-81311-9

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

This study estimated the level and determinants of air concentrations of 26 selected volatile organic compounds (VOC) in rural Western Canada. A multi-site, multi-month unbalanced two-factorial design was used to collect air samples at fixed sites from April 2001 to December 2002.

Our results indicate that VOC levels were low. Three mixtures of VOC were identified in factor analysis. Factor I was a group of short-living VOC mainly emitted from oil and gas operation and combustion. Factor II was a group of vegetation-related VOC. Factor III was a group of chlorinated VOC. Proximity to batteries was most influential in determining concentrations of components of Factor I, followed by gas and oil wells.

This thesis presents the first comprehensive statistical models of VOC concentrations and their determinants in rural Western Canada. These models can help in development of air pollution control measures and design of studies investigating health effects of VOC.

ACKNOWLEGEMENTS

This thesis could not have been completed without the support and collaboration of many people and organizations. I would like to thank my husband and my son for their unconditional love and support. Without their understanding and encouragement, I could not come back to school.

Many thanks go to my supervisor, Dr. Igor Burstyn, for providing consistent guidance and support. All of his encouragement and understanding were much appreciated. Working with him was fun. I would especially like to thank my supervisor, Dr. Ambikaipakan Senthilselvan, who always provided me with guidance and assistance with statistical procedures. And I would like to thank Dr. Nicola Cherry for insights throughout the whole project.

I would also like to thank the Western Inter-Provincial Scientific Studies Association for providing funding for this project, Rowan Williams Davies & Irwin Inc. who collected air samples, AirZone Inc. who conducted laboratory analyses, Alberta Energy and Utilities Board, Oil and Gas Commission of British Columbia and Petroleum Development Branch of Government of Saskatchewan who provide regulatory data on location and type of oil and gas facilities. Researchers at the University of Saskatchewan (Elise Pietroniro and Cheryl Waldner) have my gratitude for computing distances between air monitoring stations and oil and gas facilities, and creating the map used in this thesis. Researcher assistansts at the University of Alberta (Jianduan Liang and Narmatha Thanigasalam) also have my gratitude for creating SAS program to calculate distance weights and for (Hyang-Mi Kim) helping implement them.

TABLE OF CONTENTS

Chapter 1: Introduction 1

Chapter 2: Background 5

2.1 Overview of health and environmental effects of VOC 5

2.2 Overview of Emission Sources of VOC 9

2.3 Factors influencing airborne VOC concentrations 11

2.4 Air sampling: methodological considerations 13

2.5 Research questions 15

Chapter 3: Methods 16

3.1 Air monitoring strategy 16

3.2 Sampling devices and chemical analyses 17

Descriptive analysis 17

3.3.1 Descriptive statistics 17

3.3.2 Factor analysis 19

3.3.3 Temporal trends 21

3.3.4 Semi-variogram 21

3.4 Potential determinants of VOC exposure 23

3.4.1 Potential determinants of VOC exposure from the site documentation data 23

3.4.2 Potential determinants of VOC exposure from the regulatory data 25

3.4.3 Selected potential determinants of VOC exposure for model building 27

3.5 Basic mixed effects models 27

3.6 Statistical models of determinants of environmental VOC concentrations 30

Chapter 4: Results 33

4.1 Sampling frequency 33

4.2 Observed concentrations and variability 36

4.3 Factor Analysis 40

4.4 Time Trends 43

4.5 Spatial correlation: semivariograms 47

4.6 Description of factor scores 51

4.7 Correlations among nearby potential industrial sources and their corresponding distances from site documentation data 53

Correlations between logarithms of distance weights from regulatory data 54

4.9 Exposure variability partitioned between location, months and repeats 54

4.10 Univariate analysis 57

4. 11 Final mixed effects models: Analysis of the effect of proximity to putative sources on observed concentrations of VOC in the air 63

4.11.1 The scores of Factor I 63

4.11.2 The scores of Factor II 68

4.11.3 The scores of Factor III 72

Chapter 5: Discussion and Conclusion 73

5.1 VOC concentrations *73*

5.2 Variability of VOC concentrations *78*

Determinants of concentrations of mixtures of VOC *78*

5.3.1 The scores of Factor I 78

5.3.2 The scores of Factor II 81

5.3.3 The scores of Factor III 82

5.4 Conclusions *83*

References 85

Appendix 1:Map of spatial distribution of VOC monitoring sites 93

Appendix 2: Correlation between distance weights* 94

Appendix 3: Correlations between potential industrial sources and their corresponding distances* 96

Appendix 4: The Phi correlation coefficients between potential industrial sources and vegetation 99

Appendix 5: Time-trends of VOC 100

Appendix 6: Output of SAS PROC SEMIVARIOGRAM for Factor I 112

LIST OF TABLES

Table 1 Details of classification of sources in site documentation 24

Table 2 Sampling frequency by year and month 34

Table 3 Frequencies of potential determinants of exposure and observations 35

Table 4 Observed concentrations of the 26 VOC in rural areas of Western Canada (N=11,399)* 37

Table 5 Observed concentrations of the 26 VOC from November to April in rural areas of Western Canada (N=3,169)* 38

Table 6 Observed concentrations of the 26 VOC from May to October in rural areas of Western Canada (N=8,230)* 39

Table 7 Results of factor analysis for 26 VOC measured at 1206 locations in rural Western Canada* 42

Table 8 Variance components from basic mixed models (N=11,375) 56

Table 9 The results of univariate analysis for the scores of Factor I (N=11,375) 58

Table 10 The results in univariate analysis for the scores of Factor II (N=11,375) 60

Table 11 The results of univariate analysis for the scores of Factor III (N=11,375) 62

Table 12 Variance components of Factor I from final mixed effects model (N=11,375) 64

Table 13 Mixed-effects multivariate model of the scores of Factor I (N=11,375) 65

Table 14 Variance components of Factor II from final mixed effects model (N=11,375) 69

Table 15 Mixed-effects multivariate model of the scores of Factor II (N=11,375) 70

Table 16 Concentrations of selected VOC in air at various urban and industrial sampling sites (in $\mu g/m^3$) * 74

Table 17 Concentrations of selected VOC in air at various rural and remote sampling sites (in µg/m3)* 75

LIST OF FIGURES

Figure 1 Observed time-trend of air concentrations of benzene with and without non-detectable (nd) values in rural areas of Western Canada 44

Figure 2 Observed time-trend of the scores of Factor I with non-detectable (nd) values in rural areas of Western Canada 44

Figure 3 Observed time-trend of air concentrations of n-decane with and without non-detectable values (nd) in rural areas of Western Canada 45

Figure 4 Observed time-trend of the scores of Factor II with non-detectable (nd) values in rural areas of Western Canada 45

Figure 5 Observed time-trend of air concentrations of chloroform with and without non-detectable (nd) values in rural areas of Western Canada 46

Figure 6 Observed time-trend of the scores of Factor III with non-detectable (nd) values in rural areas of Western Canada 46

Figure 7 Semivariograms for the score of Factor I using lag distance of 5 km, output distance of 2000 km 48

Figure 8 Semivariograms for the scores of Factor I using lag distance of 5 km, output distance of 600 km 49

Figure 9 Semivariograms for the scores of Factor II using lag distance of 5 km, output distance of 2000 km 50

Figure 10 Semivariograms for the scores of Factor III using lag distance of 5 km, output distance of 2000 km 51

Figure 11 Distribution of the scores of Factor I 52

Figure 12 Distribution of the scores of Factor II 52

Figure 13 Distribution of the scores of Factor III 53

Figure 14 Qqplot of standardized residuals vs. normal score residuals: Final mixed effects model of the scores of Factor I 66

Figure 15 Plot of residuals vs. individual predicted values: Final mixed effects model of the scores of Factor I 66

Figure 16 Plot of residuals vs. predicted means: Final mixed effects model of the scores of Factor I 67

Figure 17 Observed scores of Factor I vs. predicted score generated from the final mixed effects model 68

Figure 18 Qqplot of standardized residuals vs. normal score residuals: Final mixed effects model of the scores of Factor II 70

Figure 19 Plot of residuals vs. individual predicted values: Final mixed effects model of the scores of Factor II 71

Figure 20 Plot of residuals vs. predicted means: Final mixed effects model of the scores of Factor II 71

Figure 21 Measured the scores of Factor II vs. predicted score generated from the final mixed effects model 72

Chapter 1: Introduction

The upstream oil and gas industry plays a very important economic role in Western Canada, but the impact of emissions from upstream oil and gas industry on animal and human health raises substantial concerns. This is especially the case for beef cattle producers, since cattle pastures and primary oil and gas facilities are scattered across the rural areas in Western Canada and often overlap. In response, the Western Inter-Provincial Scientific Studies Association (WISSA) was formed in 2000. The WISSA has initiated a study to evaluate the impact of exposure to emissions from oil and natural gas facilities on animal health. As a part of the study, exposure measurements for volatile organic compounds (VOC) have been collected at fixed locations throughout Alberta, north-eastern British Columbia, and central and southern Saskatchewan over a period from April 2001 to December 2002. Rowan Williams Davies & Irwin Inc., a firm contracted by WISSA, collected air samples. AirZone Inc. (Mississauga, Ontario), a commercial laboratory contracted by WISSA, conducted laboratory analyses.

Volatile organic compounds are present in the atmosphere as gas, but under normal conditions of temperature and pressure would be liquids or solids (1). A volatile organic compound is, by definition, a compound containing at least one carbon atom (excluding carbon dioxide and carbon monoxide) and with a vapor pressure of 0.01 kPa or greater at 25 °C (2). In general, methane is excluded from this definition, due to its lack of photochemical reactivity in the atmosphere.

The purpose of the air quality monitoring campaign that is the basis for this thesis was to select VOC that were most representative of the products of combustion, and oil and gas operations. The BTEX (benzene, toluene, ethylbenzene, xylene) and hexane were of greatest interest to the WISSA study, while AirZone Inc. had a number of 'panels' that they offered their clients, and the BTEX and hexane were in a panel of 26 VOC. As the costs of completing the whole panel vs. just BTEX and hexane were almost the same, the total of 26 VOC in the same panel were analyzed.

VOC are an important class of air pollutants. Exposures to many of these compounds have been linked to carcinogenic, mutagenic, teratogenic, neurological and developmental health effects, irritation and skin reactions (3-9). Many VOC in the atmosphere are unstable. In the presence of sunlight, they react to form secondary organic compounds or degrade to provide the free radical species that participate in the formation of ground level ozone (10,11). Photochemical reaction products, such as ozone, peroxyacetyl nitrate and acid, can often be of more concern than the originally emitted VOC in affecting human, vegetation and wildlife health (12,13).

There is a wide range of human activities and natural biogenic sources that result in emissions of VOC. The major sources of VOC emissions are the upstream oil and gas industry, motor vehicle exhaust, combustion sources, chemical manufacturing, paints, solvents, and biogenic emissions (1,14). Meanwhile, the coexisting system of VOC is complex. A small change in environmental conditions may result in changes of VOC concentrations, and some VOC may be involved in chemical reaction in the atmosphere. Airborne concentrations of VOC are controlled by a balance between those factors that lead to accumulation and those that lead to pollutant dispersal and removal. The processes of emission, dispersion, and removal are subject to influences by a wide range of temporal and spatial factors (1). The identification of the significant and modifiable sources of VOC concentration in the air is crucial for controlling VOC emissions and protecting air quality and public health.

The National Pollution Release Inventory (NPRI) is an important source of VOC emission data for the provinces and Canada as a whole. Approximately 70 VOC were reportable to the NPRI in 2000. According to the inventory, the upstream oil and gas industry was the largest anthropogenic VOC contributor to VOC emissions in Alberta and in Canada in year 2000 (15). Flaring, venting, and fugitive leaks from production and storage facilities were suggested as the major sources of VOC in the oil and gas industry (16). Although the Canadian Association of Petroleum Producers research assessed VOC emissions in Alberta and Canada by industrial sector of the upstream oil and gas industry in 1995, there is no published data on VOC emissions for specific oil and gas facilities such as oil wells, gas wells, batteries, compressors, and gas plants.

Studies of concentrations and source identification of VOC in urban city and rural area have been conducted recently (17-23). No such studies have been reported for the rural areas of Western Canada. Limited knowledge of determinants of airborne VOC concentrations and association between major oil and gas facilities and airborne VOC concentrations was available prior to this study.

The following three datasets were explored in this research: 1. VOC monitoring data measured during the animal health study over the period April 2001 and December 2002; 2. Regulatory data on location and type of oil and gas facilities in 2001 and 2002 supplied by provincial regulatory agencies in British Columbia, Alberta and Saskatchewan; and 3. Site documentation, in which field technicians documented any potential sources of air pollution around monitoring stations that they were able to observe. The specific aims of the research were to describe the levels and variability of airborne concentrations for the 26 VOC measured during the animal health study, identify factors that determine concentrations of VOC in the ambient air, and develop exposure models that can be used to gain insight into the impact of major oil and gas facilities on the airborne VOC concentrations in the rural area of Western Canada after consideration of temporal and spatial variability, and that are also potentially useful for predicting VOC contamination in areas without monitoring.

Chapter 2: Background

2.1 Overview of health and environmental effects of VOC

VOC are an important class of air pollutants. They are complex and co-exist in the atmosphere as a mixture. The components of the mixture may vary with time and location. No data are available on how the whole mixture is likely to act in producing adverse effects on human and animal health. For individual VOC, their toxicities vary widely between compounds and exposure pathways. The effects of exposure are influenced by a number of factors: toxicity of the chemical, dose, duration, exposure pathway, personal traits and habits, and whether other chemicals are present to have additive or interaction effect with the chemical. (24)

Human and animal health can be negatively impacted through exposure to VOC via direct pathways, such as inhalation, ingestion, or dermal contacts. Several VOC have been shown, or are expected, to have carcinogenic effects (24). Epidemiological and case studies provide clear evidence of a causal association between exposure to benzene and leukemia (25). Styrene and naphthalene are possibly carcinogenic to humans (24). Several studies of workers exposed to styrene in the reinforced plastic industry, have found an excess of lung or respiratory cancer or pancreatic cancer (26-28). Abdo et al. reported that naphthalene was carcinogenic to male and female rats, causing increased incidences of respiratory epithelial adenoma (29). The US Department of Health and Human Services (DHHS) has determined that tetrachloroethylene may reasonably be anticipated to be a carcinogen (24). Tetrachloroethylene has been shown to cause liver tumors in mice and kidney tumors in male rats, and increase incidences of bladder cancer, leukemia, and other cancer in human (30-33).

The International Agency for Research on Cancer (IARC) classifications for these VOC are listed in Table 4. According to the United States Environmental Protection Agency (EPA), VOC account for 35 to 55% of outdoor air cancer risk in the United States, although this assessment excludes diesel exhaust particles (34).

The following summary of VOC toxicities is based on chemical toxicological profiles published by U.S. Agency for Toxic Substances and Disease Registry (24). Certain VOC have been shown to have non-carcinogenic effects, which include respiratory and central nervous system problems, exacerbate allergies, and have adverse neurological, reproductive, and developmental effects. Long-term exposure to hexane may be associated with peripheral nerve disorders and central nervous system effects. Exposure to high concentrations of xylene and toluene may increase the number of fetal death, delay fetal growth and development. Cumene has a potent central nervous system depression action. Repeated or prolonged exposure to 1,3,5-trimethybenzene or 1,2,4-trimethybenzene may affect lung functions and result in chronic bronchitis. Toluene has neurotoxic effects. Neurotoxic symptoms and reduced ability in tests of cognitive and neuromuscular function have been observed in humans occupationally exposed to toluene. Several VOC, such as 1,3-dichlorobenzene, p-cymene, and chloroform can cause liver and kidney damage. Other VOC, such as benzene, ethylbenzene, and 1,3,5-trimethybenzene can irritate eyes, nose, respiratory tract, or skin of animals and humans. In addition, the malodorous properties of certain VOC such as styrene, tetrachloroethylene and toluene may lead to irritation or annoyance in animals and humans.

Another important pathway that may ultimately lead to negative impact of VOC on human health and the environment is the formation of ground-level ozone (O_3) (1,14). It has long been established that non-methane hydrocarbons play an important role as precursors to ozone and other secondary photochemical pollutants (2,10,35). Ozone is formed in the complex reaction mechanism that involves some VOC and oxides of nitrogen (NO_x) in the presence of sunlight. In this situation, VOC react with hydroxyl radicals (produced photochemically) to produce peroxy-radicals; peroxy-radicals, in turn, react with NO to produce NO_2. The resulting NO_2 then reacts further to produce ozone (2,11). Ground-level ozone has been linked to negative impacts on human health, including lung damage, chest pains, coughing, nausea, throat irritation, and congestion; significant increases in symptoms have been observed following exposure to ozone as low as 60 ppb over a period of 16 to 28 minutes of heavy exercise (35). Ground-level ozone has been linked to negative vegetative effects, such as the reduced ability of plants to produce and store food, making them more susceptible to disease, insects, other pollutants, and harsh weather (36).

Moreover, volatile organic compounds in the ambient atmosphere can affect human health and the environment both directly and indirectly through the formation of particulate matter (1,14). Exposure to particulate pollution has been linked to premature death, difficulty in breathing, aggravated asthma, increased hospital admissions and emergency room visits, and increased respiratory symptoms in children (37). Environmental effects of increased particulate matter include reduced visibility and decreased growth and productivity of vegetation, due to interference with photosynthesis (37).

Reproductive and developmental effects are of special interest to the epidemiological study that the data is derived from. The reproductive and developmental toxicity of xylenes, toluene and styrene have been extensively studied (9,38-50). Decreases in fetal weight and delays in skeletal development have been reported in several studies of exposing pregnant animals to different concentrations of xylenes. Ungvary et al. observed weight retarded fetuses in rats at all air concentrations after exposure of rats by inhalation to air concentration of xylene (60 ppm, 440 ppm, 800 ppm) for 24h/day on days 7-15 of gestation (49). The same research group also observed increased incidences of weight-retarded fetuses and increased skeletal retarded fetuses at 230 ppm after exposure of mice by inhalation to air concentration of xylene (120 ppm, 230 ppm) for 24 h/day on days 7-15 of gestation (50). Marks et al. noted that 2060mg/kg/day of mixed xylene administered orally was associated with cleft palate and decreased fetal weight in mice (42). However, Rosen et al. reported that when they exposed pregnant Sprague-Dawley rats by inhalation to high concentration of p-xylene (800 ppm, 1600 ppm) on the days 7-16 of gestation, no effects were seen on litter size or weight at birth or on the subsequent growth rates of the pups (46). Hass et al. reported that xylene exposure did not show signs of maternal toxicity and no difference in the number of live or dead fetus; and the mean birth weight in exposed litters was about 5% lower compared to control litters, but the difference was not statistically significant (9). There is a large variation in the concentration of xylene producing developmental effects and of those producing no developmental effects. Reproductive effects were also documented by Taskinen et al. who found increased incidence of spontaneous abortions in 37 pathology and histology workers exposed to xylene and formaldehyde in the work place. However, the multiple chemical exposures and the small number of subjects in this study limit the conclusions that can be drawn as to reproductive effects of xylene in humans (51).

A number of inhalation studies have been performed in laboratory animals (rats, mice, rabbits) to address the developmental toxicity of toluene. In a fertility study, Ono et al. reported fetal mortality and the number of dams with dead fetuses were higher than controls after Sprague-Dawley rats were exposure to 2000 ppm of toluene; and among males, fertility was unaffected, but sperm counts decreased 20-25% at 2000 ppm and 10% at 600 ppm; absolute and relative epididymis weight were decreased at 2000 ppm (43,44). Dalgaard et al. reported that neither pre- and postnatal exposure of Wistar rats to 1200 ppm toluene, nor prenatal exposure to 1800 ppm toluene induced significant effects on the semen quality of exposed rats (40). Thiel et al. also reported that no differences in mating, fertility or pregnancy indices were found after *in utero* exposure of Wister rats to 1200 ppm toluene (52). Several epidemiological studies suggested toluene might have reproductive effects in humans. Arnold et al. and Pearson et al. reported that the intentional abuse of toluene (sniffing) during pregnancy can results in children with deformities similar to those described for fetal alcohol syndrome (53,54).

Developmental and reproductive toxicity of styrene has been studied in animals and humans. Brown et al. concluded in a recent review that styrene does not affect fertility or reproductive function in animals and humans, and that styrene is neither an endocrine-active substance, nor an endocrine disrupter, although some study authors have concluded that styrene is either a human, or an animal reproductive/developmental toxicant (39). Developmental and reproductive toxicity studies have been conducted in rats, mice, rabbits and hamsters. In most cases, high doses are required to induce health effect, and these effects are not unique to reproduction or development (39). Human studies often suffer from either inadequate exposure data or exposure to a wide variety of materials (39).

2.2 Overview of Emission Sources of VOC

Volatile organic compounds are emitted to ambient air from both anthropogenic and biogenic sources. A wide range of anthropogenic sources, primarily fuel production, distribution and combustion, as well as solvent usage, industrial process, land-filled waste, food manufacture, and agriculture, are related to VOC emissions (1,15). Natural biogenic emissions from plants, trees, wild animals, natural forest fires, and anaerobic processes in bogs and marches also give rise to substantial ambient concentration of VOC (14,55). Small amounts may also be released from ocean surfaces. Vegetation is the primary source of biogenic VOC emissions; emissions from land vegetation typically have the greatest significance in terms of total emitted VOC (14,55,56). For example, the total anthropogenic and biogenic emissions estimated by Environment Canada in 1995 in Canada were 3,575,202 and 12,769,510 tones, respectively; these figures are similar for Alberta: 762,732 and 1,235,410 tones, respectively (55,56). However, the nature of emitted VOC may be quite different between anthropogenic sources and biogenic sources.

Emissions estimates provided by Environment Canada indicated that the five largest contributors of anthropogenic VOC emissions in Canada in 1995 were the upstream oil and gas industry (739,760 tones), general solvent use (309,452 tones), off-road use of gasoline (251,274 tones), light-duty gasoline vehicles (219,152 tones), and light-duty gasoline trucks (148,494 tones) (56). In Alberta in 1995 the major sources of VOC emissions were report to be the upstream oil and gas industry (484,788 tones), forest fires (61,356 tones), light-duty gasoline vehicles (44,123 tones) and light-duty gasoline trucks (28,982 tones), residential fuel wood combustion (36,033 tones), and general solvent use (26,584 tones) (56). The upstream oil and gas industry is the largest anthropogenic contributor of VOC emissions both in Alberta and in Canada.

The upstream oil and gas industry includes exploration, production and basic processing of crude oil and natural gas. The industry in Canada has operations in seven provinces and two territories with the majority of activities taking place in Alberta, Saskatchewan and British Columbia. It comprises all infrastructure used to find, produce, gather, treat/process and transport natural gas, liquefied petroleum gas, condensate, crude oil, heavy oil and crude bitumen to market (16). The infrastructure varies in size, age, design and operating characteristics. The VOC emission sources range from small fugitive leaks on valves and fitting to large point sources such as flaring, process vents, and occasional well blowouts and pipeline release. Flaring, evaporation and fugitive leaks are the major sources (16).

Among the upstream oil and gas industry, conventional oil production sector is supposed to be the major source of VOC emissions. The Canadian Association of Petroleum Producers research found that, in 1995, the conventional oil production sector accounted for 66.4% of the total VOC emissions by the industry, while the gas production sector accounted for 16.1% (16). Other minor VOC emission sources include heavy oil production sector, gas processing sector, oil transportation sector, and accident and equipment failure (16). In the conventional oil production sector, the major facilities include oil well, oil battery (single-well battery, satellite battery, and central battery), and the associated flow line. The gas production sector is comprised of the following major facilities: gas well, gathering system, compressor station, metering station, and gas batteries (16). However, there is no data available for VOC emissions by specific oil and gas facilities.

2.3 Factors influencing airborne VOC concentrations

Volatile organic compounds are a large number of hydrocarbon compounds with variation of emission sources, physical and chemical properties, and reactivity in the atmosphere. Most studies were focused on only a small group of compounds. Only 50-60 VOC are well investigated. The factors affecting individual VOC vary widely.

Studies show there is considerable temporal and spatial variation of anthropogenic VOC emissions. Several studies already showed that in rural or remote environments worldwide, VOC exhibit pronounced seasonal cycles (20,22,57,58). The cycle of anthropogenic VOC concentrations with maxima in winter and minima in summer not only reflects enhanced photochemically driven processes in summer, but also the strengthening of emission rates in winter. For instance, wintertime heating and vehicle "cold start" effect may be responsible for higher emissions of combustion related products (59). The same seasonal pattern of VOC concentrations was also observed in urban areas. Cheng et al. reported that seasonal variations with maxima in winter and minima in summer were observed in a downtown site in Edmonton, Alberta, Canada and they attributed seasonal changes to variation in weather (17). Na et al. observed the same seasonal variation in Seoul, S. Korea and they attributed this to the variations in temperature, source strengths, and photochemical reactivity of reactive compounds (59).

VOC concentrations vary from one location to another. Derwent summarized concentrations measured at six representative sites along a pollution gradient across Europe (1). He found that concentrations of VOC from urban roadside site were the highest, followed by urban background site, rural site, and remote maritime background sites. A study of roadside microenvironments of metropolitan Hong Kong found that the highest VOC concentrations were found in the industrial district, which were followed by those in the commercial district, then central business district and finally the residential district (21). Site-to-site VOC concentrations differences were also observed in the Kanawha Valley of West Virginia (USA) and they were attributed the combination of topography, local meteorological conditions, and the presence of chemical industry (60). A number of studies showed the spatial variation of anthropogenic VOC concentration were influenced by traffic intensity, distance from emission sources, wind direction and speed, type of sources and strengths of sources, local climate, weather, and nearby land use (19,59-62).

Most hydrocarbons are removed from the atmosphere primarily by reaction with hydroxyl radicals (OH) (1,19). At northern middle latitudes, rates for removal vary markedly with season, while emissions are largely invariant. Rates of removal depend on the concentration of OH (19). In summer, the concentration of OH stays near its peak value for approximately 70 days, while OH stays near its minimum for approximately 105 days in winter (19).

Coniferous trees such as pines, cedars, redwoods and firs are predominant emitters of monoterpenes such as d-limonene, alpha- and beta-pinene, while many crop and grass species also emit small proportion of monoterpenes (63). Monoterpene emission rates are primarily affected by leaf temperature, and there have been contradictory reports of a light dependency of monoterpene (63). Alpha-pinene emission in *Quercus ilex* was found to increase three times when temperature raise from 20 to 30 °C (64). Tingey et al. reported no increase in monoterpene emissions from Slash pine under conditions of constant temperature (65). However, emissions from some plant species were found to increase with an increase in light intensity (66-68).

The air concentration of monoterpenes shows diurnal and seasonal variation (63). They mainly depend on emission rates from biogenic sources and removal rates by photochemical processes and dispersion. In a polluted area, monoterpenes can be destroyed by reactions with O_3 and OH radicals during the day, and during the night they can also react with NO_3 radicals, in addition to O_3. In 'clean air', dispersion rather than chemistry is the most important factor determining air concentration of the monoterpenes (63).

2.4 Air sampling: methodological considerations

Adsorbent type, sampling equipment and analytical technique are the key parameters that affect precision of VOC measurements. The choice of adsorbent defines which of the hundreds of the VOC in air can be quantitatively determined. The sampling technique can be either passive, based on diffusion of the molecules from an area of higher concentration in the environment to an area of lower or zero concentration in the collection device, or active, pumping a known air volume through the adsorbent.

Passive samplers such as charcoal-based organic vapor monitors (OVMs) have been used for many years for sampling VOC in occupational settings with relatively high concentrations. They have recently grown in popularity for indoor and personal sampling, because they are relatively inexpensive, non-intrusive, and more easily deployed than active samplers, and because analytical methods have been developed for determining low concentrations of VOC sampled with passive samplers (69). The term OVM is typically applied specifically to the 3M Corp. (Maplewood, MN) passive sampler.

Evaluations of the 3M OVM for sampling of VOC have been performed in several studies in indoor and outdoor settings (69-73). In terms of field comparisons, OVMs were found to be comparable to canister and continuous gas chromatography monitoring, with apparent over- and under-estimation of concentrations (71). A field comparison between 3M 3520 and automated-gas chromatograph (auto-GC) measurements found that the reproducibility was lower from duplicate OVMs for BTEX (benzene, toluene, ethylbenzene, and xylene isomers) with better precision for longer sampling periods; for BTEX species, generally good agreement was obtained between OVMs and auto-GC measures; except for toluene, OVM BTEX measurements generally exceed their continuous counterparts with a mean bias of 5-10% (72). Higher humidity has been found to result in an increasing negative bias for OVM measurements, because water adsorbed onto the surface of the activated carbon adsorbent tends to preclude adsorption of other compounds onto the charcoal active sides (73).

There are very diverse VOC air-sampling designs in terms of the duration of air sampling and height of sampling points. No data was found on effects of different durations of air sampling or effects of different heights of sampling points on VOC measurements.

2.5 Research questions

This thesis only focused on analysis of data already collected and that the sampling strategy was already implemented by study sponsors. Original research presented in this thesis will address three research questions.

1. What are the level and variability of environment concentrations of VOC in the rural area of Western Canada?
2. What are the determinants of environmental concentrations of VOC in the rural area of Western Canada?
3. Do oil and gas facilities contribute to environmental concentrations of VOC in the rural area of Western Canada?

Chapter 3: Methods

3.1 Air monitoring strategy

The monitoring sites were located in the vicinity of approximately 33,000 beef cattle in some 200 herds selected across a geographic area associated with primary oil and gas industry in the Canadian provinces of British Columbia, Alberta and Saskatchewan. The 3M 3500 Organic Vapor Monitors (OVMs), passive samplers, were deployed for one-month periods between April 2001 and December 2002. They were systematically relocated to follow the herds as they were transferred to new pastures. Each monitoring site included a shelter to protect the samplers from rain and snow. The samplers were installed upside down to allow undisturbed air movement across the reactive surface. For quality assurance and control purposes, 10% of samplers were used as blanks to determine potential contamination sources associated with field handling and transportation. An additional 10% of the sample sites contained replicate (2-3) month-long average samples.

The following criteria were used to select monitoring sites:

1. Adjacent to the pasture where the cattle were spending most of their time.
2. Away from minor sources, such as road, farm vehicle emission, fuel and farm equipment storage areas, and local oil field equipment.
3. More than 10 meters from roadways and other areas where vehicles were expected to be stopping or idling.
4. Away from fuel and farm equipment storage areas; a separation distance of 100 meters was recommended.
5. Not in the immediate vicinity of local oil and gas facilities.
6. Monitor heights were set from 1.5 to 1.8 meters above ground.

7. Away from nearby obstructions such as buildings, hay storage, and trees that may obstruct airflow. Sites were located at least 20 meters away from the nearest tree canopy.
8. In a flat terrain (e.g. not at tops or bottoms of hills).

The number of sampling peaked in summer as herds were moved to pastures and as herds were split up. Over the period of July 2002 to December 2002, the number of total samples that were collected was reduced due to funding cutbacks.

Map in Appendix 1 illustrates the areas covered by 50 km radii around VOC monitoring sites.

3.2 Sampling devices and chemical analyses

The 3M 3500 OVMs, which were equipped with a Teflon membrane impregnated with charcoal, were used to collect air samples. VOC were extracted from the sampling media using a carbon disulphide solution (CS_2) spiked with deuterated compounds. Then the extracts were injected into a gas chromatograph equipped with a mass spectrometer operated in selected ion monitoring mode with three ions per compound monitored. The sampling rates, controlled by the diffusion barrier, were calculated for each of the 26 VOC based on extensive testing in chambers and in ambient air (further information on this matter is proprietary and was not released by AirZone Inc.). Monthly average concentrations for each of the 26 VOC were provided by AirZone Inc. and raw data were not used in this analysis. The VOC samplers were prepared and VOC samples were analyzed by AirZone Inc. (Mississauga, Ontario).

Descriptive analysis

3.3.1 Descriptive statistics

Histograms of monthly average concentrations of the 26 VOC were generated to examine the frequency distributions. Since histograms of VOC concentrations were quite right-skewed, natural logarithmic transformations were applied in order to obtain normally distributed variables for analyses. Non-detectable exposure measurements were replaced by one half of the lowest concentration reported for a given VOC. VOC measurements were described well by lognormal frequency distributions. Therefore, geometric means (GM) and standard deviations (GSD) were employed to describe the data. Geometric means and standard deviations were calculated by the following formulas: GM_x= exp(mean of log(X)); and GSDx=exp(standard deviation of log(X)).Other descriptive statistics, such as sample size, percentage of non-detectable measurements, arithmetic mean, and maximum values, were also generated.

The US Environmental Protection Agency has not developed standards for total or individual VOC compounds. Regulation of VOC and other atmospheric contaminants is considered the responsibility of individual states (14). In Canada, the Ontario Ministry of the Environment has developed ambient air quality criteria (AAQC) for more than 300 substances, including some 231 VOC species. However, AAQC were first developed in Ontario more than 20 years ago, and Ontario Ministry of the Environment is currently in the process of reviewing and updating existing air quality criteria to ensure that they are current and protective of human and ecosystem health (14). The World Health Organization (WHO) has also created guidelines for a number of atmospheric contaminants. Several VOC are included. However, the WHO guidelines are largely based on acceptable risk levels rather than absolute safety (zero risk), as insufficient data exists for the purpose of setting guidelines based on absolute safety (14). In Australia, the Environment Protection and Heritage Council has developed annual arithmetic mean standards for benzene 0.003 ppm (9,565 ng/m^3), toluene 0.1 ppm (376,000 ng/m^3), and total xylene 0.2 ppm (867,000 ng/m^3). These limits are based on the considerations of lifetime exposure that does not constitute a significant health risk, i.e. do not cause chronic diseases. (74). In this study, we adopted the standards developed by Australian Environment Protection and Heritage Council.

The arithmetic mean of benzene, toluene, and total xylenes (sum of m/p-xylene and o-xylene) were calculated for each monitoring site for each year, and then were compared to the standards developed by Australia Environment Protection and Heritage Council. After that, the percentages of measurements exceeding the limits were computed.

3.3.2 Factor analysis

Correlations among 26 VOC were examined (data not shown). Many of VOC were mutually correlated, especially the isomers. There is little advantage to construct statistical models for each chemical individually in such cases. Factor analysis can be employed to detect the underlying structure of mutually correlated variables and to reduce the number of dependent variables. Furthermore, factor analysis is widely used to categorize measured compounds into distinct source groups based on the covariance of their concentrations, creating an understanding of the variety of sources contributing to a broad range of measured species (75-80). Therefore, principle component factor analysis method was used in this study to reduce the number of dependent variables and to categorize the 26 VOC into 'source groups'.

Factor analysis is widely used to uncover the latent structure (dimensions) of a set of variables. There are several different types of factor analysis, with the most common being principal components analysis (PCA). PCA has been widely applied to investigate the complexities of workplace exposure to airborne contaminants (75-77,79), as well as environmental exposure to VOC (78,80). PCA allows the identification of groups of variables that are interrelated via phenomena that cannot be directly observed. The "factors" which are derived from PCA consist of patterns of variation or behaviour to which a significant number of input variables conform. Once a factor is determined, that pattern is removed from the data set and the residual data is re-examined to determine if any remaining covariance exists. PCA involves the successive abstraction of factors until all the covariance is assigned to the factors. The extent to which a particular factor explains multiple correlations is measured by its eigenvalue. The variability explained by a given factor is equal to its eigenvalue divided by the number of variables in the analysis (81). Eigenvalues and a screen plot (i.e. factor number versus eigenvalue) are the common ways to help determine how many of factors should be abstracted. Eigenvalues >1 and the slope of the "scree" begins to flatten out in the scree plot marks the location of the last factor that makes a significant contribution to explaining multiple correlation are the common criteria to determine the number of factors (81).

The VOC concentrations were obtained from repeated measurements with a multi-location, multi-month, two-level factorial design; the independent observations assumption of factor analysis was not met in this case. So, logarithmic means of VOC concentrations for each sampling site were estimated, and factor analyses were performed on them. VARIMAX rotation of the factors was used to obtain orthogonal factors that explain maximum amount of multiple correlations. Eigenvalue and scree plot were used to determine the number of factors to be extracted. A factor loading (i.e. correlation coefficient between factors and site-specific logarithmic means of specific VOC) of greater or equal to an absolute value of 0.5 was considered as a cut-off point for interpretation of the factors. These analyses were repeated for each month to observe whether the nature of factors was subject to temporal variability. In general, similar compositions of factors were observed for all months. In addition, factor analysis was also conducted on the actual logarithms of concentrations of 26 VOC. There was similar output between factor analysis based on the logarithms of concentrations and that based on logarithmic means of sites of the 26 VOC. However, the results of factor analysis based on logarithmic means of sites were easier to interpret. Therefore, we decided to adopt the results of factor analysis performed on logarithmic means of sites for presentation of results and calculation of factor scores. Standardized scoring coefficients generated from the factor analysis based on logarithmic means of sites were used to calculate factor scores. Factor scores of factors (new variables) were calculated using the following equation:

Factor score for factor F =Σ {(standardized scoring coefficient for chemical X and factor F) * (log (concentration of chemical X))} (1)

The new factors were used as dependent variables in spatial correlation analyses and mixed effects modeling. Factor analyses were conducted by using SAS **PROC FACTOR** procedure through *nfactor*, *rotation* and *scree* options (SAS version 9).

3.3.3 Temporal trends

Temporal trends in VOC concentrations were examined graphically. Two types of monthly geometric means were calculated for each VOC: 1. Non-detectable values were included in the calculation and replaced with half of the minimum detectable values. 2. Non-detectable values were excluded from the calculation. Time-trend graphs were generated separately by using two different monthly geometric means for each VOC.

3.3.4 Semi-variogram

The classical semi-variogram is an appropriate tool to measure spatial dependence for measurements sampled at irregularly spaced sites. Semi-variogram is a plot of semi-variance versus separation distance that allows one to detect and visualize spatial correlation in the data. It relies on the notion that measurements taken in close proximity to one another will be more similar if spatial correlation exists. The standard (regular) semi-variance is calculated using the following formula (2a):

$$2\gamma_z(h) = \frac{1}{|N(h)|}\sum_{N(h)}(Z(r_i) - Z(r_j))^2 \tag{2a}$$

where γ_Z (h) is semi-variance at separation distance h, z(r_i) and z(r_j) are observed values at distances r_i and r_j, h (lag) is the separation distance between two measurements, and N(h) is the number of z(r_i) and z(r_j) pairs within h and is given by

$$N(h, \delta h) = \{i, j :| r_i - r_j | \in [h - \delta h, h + \delta h)\} \tag{2b}$$

The robust semi-variance is computed using the following formula (2c):

$$2\,\gamma_Z(\mathrm{h}) = \psi^4(\mathrm{h}) \;/\; \{0.457 + 0.494 / \mathrm{N}(\mathrm{h})\} \tag{2c}$$

where

$$\psi(\mathrm{h}) = \{1/\,\mathrm{N(h)}\} * \Sigma_{\mathrm{all\ N(h)}} \{\mathrm{z}(\mathrm{r_i})-\mathrm{z}(\mathrm{r_j})\}^{0.5} \tag{2d}$$

Semi-variograms were constructed for three factor scores. Spatial correlation coefficients within lag distance of 5 kilometres were estimated. The average of spatial correlation coefficient within 5 km was computed by dividing the covariance for 5 km lag with the covariance for 0 km lag. The results of semi-variogram analyses were used to guide the selection and exploration of random components of mixed effects models.

The regular and robust semivariograms were graphed by using SAS **PROC VARIOGRAM** procedure through *compute* and *coordinates* statement and *lagdistance and maxlags* options (SAS version 9).

3.4 Potential determinants of VOC exposure

3.4.1 Potential determinants of VOC exposure from the site documentation data

Field technicians collected the site documentation data on nearby potential sources of VOC around the sampling sites and corresponding distances away from the sampling sites. We transferred this data into a relational database (Microsoft Access). Minor sources located within 2 km of the pastures and major sources located within 5 km of the pastures were required to be documented. Estimations were made to determine distances from nearby potential sources by visual observation. A total of 1241 sites were documented. Out of 1241 sites, 432 sites had no industrial source documented; 87 sites had no information about nearby potential sources, but other information such as site terrain, tree canopy were documented; 35 sites had no data; and 6 sites had nearby potential source information, but no VOC samples were collected at these sites. Among these 35 sites that had no data, only 6 sites had air samples (a total of 24 observations). These 35 sites and 24 observations were excluded from analyses of associations with potential sources of VOC. The 87 sites, which had missing information only about nearby potential sources, were coded as if there were no industrial sources of VOC. A total of 11,399 observations were used for descriptive analyses, factor analyses and semi-variograms, while a total of 1206 sites and 11,375 observations were used in statistical modelling of determinants of exposure.

We classified the nearby sources into 10 categories such as battery, compressor, flare, plant, well, other oil and gas facilities, other industries, highway, road, and no industrial source. The detailed classification scheme is showed in Table 1.

Table 1 Details of classification of sources in site documentation

Classification	Sources documented by field technicians (verbatim)
Well	Well, wells, gas well, oil well, wellhead, drilling rig, drilling well, pump, pump jack, pump tank
Plant	Gas plant, gas processing plant, plant
Battery	Battery, battery tank, batteries
Compressor	Compressor
Flare	Flaring, flare stack, flare
Other oil and gas facilities	Heater, pipeline riser, gas tank, diesel tank(s), diesel, purple clear gas, gas risers, gas valve, pipeline, tank(s), tank farm, X'mas tree, gas operations, many oil activity, oil field, pump station
Highway	Highway(s)
Road	Road, gravel road, railway (just used once or twice a week)
Other industries	Power sub-station, industrial park, town, city, energy corporation, power transformer, shop, welding shop
Non-industrial source	Acreage(s), forest reserve, yard, farmyard, power line, electric line, irrigation pump, lease site, pit, slough, water well, hay field, rittle range, river, salt lick, silage operation, water pipeline, farm(s), farmstead, homestead, corral(s), dugout, barn(s), farm erwip, feed lot, home, loading corrals, quarters, school, Sask Tel tower, bale stack, grain bins, grain terminal, gravel pile, gun club, hat stack, hay bales, microwave tower(s), shacks, Telus tower, tin shed

Each type of nearby potential industrial sources was dichotomized as "present" or "absent". The corresponding distances of nearby potential source were classified into "no distance", "close", "far", and "unqualified distance", since the distances were estimated by filed technicians using visual observation and the range of distances was limited. "No distance" was defined for sites where there was no corresponding source at given sites. "Close" distances were defined as distances that were less than or equal to 2 kilometres. "Far" distances were defined as distances that were further than 2 kilometres. When more than one similar potential source was present around the same sampling sites, the closest distance among these similar sources was used. "Unqualified distance" was defined for sites where field technicians just reported the nearby potential source(s), but they did not report the corresponding distances, or they wrote down numbers for distances without specifying units (e.g. meters, feet, km).

Considering biogenic sources of certain VOC in our study, we generated "vegetation" variable coded from the documentation of "tree canopy" from the site documentation data. The "vegetation" variable was dichotomized as "present" or "absent". "Present" was defined to mean that there were some tree(s) or bushes near the sampling site. "Absent" was defined to mean that there were neither trees, nor bushes near the sampling site.

There were some limitations to the site documentation data. Because of lack of clear definition of nearby potential sources, different technicians may have had different interpretation of this notion. Furthermore, some technicians did not write down exact numbers of nearby potential sources and exact distances of nearby sources. For example, they just wrote down "lots of wells around". Due to confidentiality issue, we could not contract the field technicians to verify the information in site documentation database.

3.4.2 Potential determinants of VOC exposure from the regulatory data

Data on location and type of oil and gas facilities in 2001 and 2002 were supplied by provincial regulatory agencies in British Columbia, Alberta and Saskatchewan. Information on coordinates of monitoring stations and oil and gas facilities was combined to estimate distances between each putative source and monitoring station. Effects of sources on measured concentrations were considered for two distance classes within monitoring stations: 0-2 km and 2-50 km. For each distance class around each monitoring station and year (2001 or 2002), we estimated the number of facilities and a weighted sum of their proximity to the monitoring station. The following types of sources were considered separately: oil wells, gas wells, bitumen wells, oil batteries, gas batteries, all "large" facilities, and gas plants. Only facilities classified as "active" by the regulatory agencies in a given year were considered in analysis.

The weighted sum of distances of similar sources (e.g. gas wells) in a given distance class (e.g. 2-50 km) were computed as sum of 1/ (distance) $^{2/3}$. This is based on dispersion model proposed by Strosher (82), which states that concentration of air pollutants in the air is directly proportional to emission rate and 1/ (distance) $^{2/3}$. Only those types of sources that can be identified in all provinces were used in subsequent analysis (except for bitumen wells, which were assumed to exist almost exclusively in Alberta, and were coded as non-existent elsewhere). Thus, proximity (P_{ijkl}) of kth monitoring station (1, ..., K) at lth time (1, ..., L) to s sources (1,...S_{ijkl}) of ith type (1, ..., I) within jth distance class (1, ..., J)

Can be quantified as the following equation (3):

$$P_{ijkl} = \Sigma_{\mathrm{all}\ s(ijkl)} [D_{s(ijkl)}{}^{-2/3}] \qquad (3)$$

Where D is distance in kilometres (km) from the source to monitoring station, and each sum is specific to fixed combination of source type, distance class, location and time (i.e. month) of air quality measurement. If value of proximity could not be determined, as when there were no bitumen wells within 2 km of a monitoring station, the value of P_{ijkl} was set to zero.

3.4.3 Selected potential determinants of VOC exposure for model building

The above two data sets both provided similar information about potential determinants of VOC exposure around sampling sites. However, the two data sets both have advantages and disadvantages. The site documentation data provided more nearby potential sources of VOC such as other industry, highway, road, and trees. On the other hand, the oil and gas facilities such as well, plant, and battery and their corresponding distance were estimated by field technicians through visual observation, we could not get accurate information about the number of facilities and their distances from sample sites. Furthermore, we could not classify wells into oil wells, gas wells and bitumen wells and batteries into oil batteries and gas batteries. In addition, we could not know whether these oil and gas facilities were active or not in a given year. Compared to site documentation data, the regulatory data provided more accurate information about the number of main oil and gas facilities, their distances from sample sites (whose location was determined by GPS) and their active statues in a given year. Nonetheless, regulatory data use Dominion Land Survey (DLS) system, exact location of oil and gas facilities was uncertain and was assumed to be the center of the smallest DLS units. However, the regulatory data only included the potential VOC sources from main oil and gas facilities; it did not include other important potential VOC sources such as highway and city or town. So, we combined the two data sets together to select potential predictors of VOC exposure. We chose weighted sum of distances of oil wells, gas wells, bitumen wells, batteries, all "large" facilities, and gas plants from the regulatory data set, and flare, compressor, other oil and gas facilities, other industry, highway, road, and vegetation from the site documentation data set as potential determinants of VOC exposure.

3.5 Basic mixed effects models

The following basic mixed effects model (4) was used to define seasonal effect, spatial correlation, and random effects of months, sampling sites, and duplicates:

$$Y_{hijm} = \mu_y + \delta^*\text{Season}_m + \tau_h + \lambda_i + \varepsilon_{j\,(hi)} \tag{4}$$

Y_{hijm} represents factor score at the ith sampling site and the jth replicate in the hth month and mth season; μ_y represents the true unknown mean of factor scores; δ represents the regression coefficient (fixed effect) for the dummy variable season (winter or summer); τ_h represents the random effects of the hth month; λ_i represents the random effects of ith sampling site; $\varepsilon_{j\,(hi)}$ represents the random effect of the jth replicate nested in the ith sampling site and the hth month. It was assumed that τ_h and $\varepsilon_{j\,(hi)}$ are each normally distributed, and mutually independent, with means of zero and variances of $\sigma_m{}^2$ and σr^2, respectively. Since the empirical semi-variograms of the scores of Factor II and Factor III indicated that there were negligible spatial correlations existed (Figure 9 and 10), we further assumed that τ_h, λ_i, $\varepsilon_{j\,(hi)}$ are each normally distributed, and mutually independent, with means of zero and variances of $\sigma_m{}^2$, $\sigma_s{}^2$ and $\sigma_r{}^2$, respectively. The covariance structure was specified as compound symmetry for each random effect for the scores of Factor II and Factor III.

Only for the scores of Factor I did semi-variogram showed that moderate spatial correlation existed (Figures 7 and 8). Spatial random effect λ_i has expectation zero and covariance $\gamma_0(d)$ (see below for details); it is also independent of τ_h and $\varepsilon_{j\,(hi)}$. Based on the empirical semi-variogram features of Factor I, the covariance structure of spatial correlated errors within month was specified as spherical covariance structure. The spherical covariance structure describes a relatively rapid increase in variance as distance between points increases followed by a plateau that represents the variance for the points far enough apart to be independent. The description of spherical covariance structure is as following equation (5):

$$\gamma_0(d) = \left\{ \begin{array}{cc} 0, & d = 0, \\ c_0 + c_1 \left\{ \frac{3}{2}\frac{d}{R} - \frac{1}{2}\left(\frac{d}{R}\right)^3 \right\}, & 0 < d \leq R, \\ c_0 + c_1 & d \geq R \end{array} \right\} \tag{5}$$

where, $\gamma_0(d)$ is the covariance between points with distance d. d is the absolute distance between points. $R>0$, a scale parameter, is the range. $C_0 \neq 0$, is the nugget effect, and C_1 is the sill. The terms "range", "sill", and "nugget effect" are from geostatistics. Range is the distance between points where the covariance is approximately zero. Sill is the covariance value where distances between points equal or exceed the range value. Nugget effect represents variations at a much smaller scale than any of measured pairwise distances. When the distance d between points reaches to the range R, the points will be far enough apart to be independent (SAS version 8 online help).

After comparing the model fit statistics of using different classification of seasons, such as four seasons, three seasons and different two seasons in the above basic model, we adopted two seasons that from May to October were classified as summer, from November to April were classified as winter. In this study, the model fit statistics used to choose better models are Akaike information criteria (AIC) and the Bayesian information criteria (BIC). Smaller AIC and BIC indicate better model fit.

The basic mixed effects model was performed using **PROC MIXED** procedure through the *random* and *repeated* statement (SAS version 9). It was fitted using restricted maximum likelihood (REML) method. Non-significant effects ($p>0.05$) were removed from final basic mixed effects models.

The ratios of the 97.5th percentile to 2.5th percentile of the corresponding scores of Factors ($R_{0.95}$) were computed to assess the variability of between months, sampling sites, and repeats. $R_{0.95}$ is also called fold-range of variation. Under the basic mixed effects models, $R_{0.95}$ of between repeats were estimated by $R_{0.95} = \exp(3.92^* \sigma_r)$, $R_{0.95}$ of across-sampling site were estimated by $R_{0.95} = \exp(3.92^* \sigma_s)$, and $R_{0.95}$ of between months were estimated by $R_{0.95} = \exp(3.92^* \sigma_m)$. For the scores of Factor I, the variance of across sampling sites vary with distance between sampling sites, $R_{0.95}$ was only computed for sampling sites which their distance reached the range where the scores of Factor I could be considered independent.

3.6 Statistical models of determinants of environmental VOC concentrations

Determinants of concentration models were constructed through univariate analyses and multivariate analyses. Univariate analyses were performed by adding each fixed effect of potential determinants of exposure selected from the site documentation data set and the regulatory data set to the basic mixed effects models once a time. We evaluated statistical significance ($p \leq 0.05$) and direction of effect estimates. Only those variables, which were significant during the step of univariate analyses and their estimations were plausible, were considered further in multivariate models.

In order to use distance weights as fixed effects in statistical models, we had to log-transform them. This was necessary because logarithmic transformation was applied to concentrations of all 26 VOC in order to normalize their frequency distributions, and therefore it is logarithmic distance weight that can be expect to be linearly related to measured concentrations. Furthermore, a constant value of 1 was added to all distance weights (P_{ijkl})

before log-transformation to avoid the problem with log(0) being undefined. For these continuous variables of logarithm of distance weights, scatter plots were graphed to examine univariate relations. Correlations between all logarithms of distance weights (i.e. $\log(P_{ijkl}+1)$) were examined and Pearson correlation coefficient were calculated (see Appendix 2). If two variables had correlation of at least 0.7, only one of them was considered further in multivariate models. The choice between the two was based primarily on plausibility and model fit. For example, the variable of logarithms of distance weights for all large facilities (incl. gas plants) within 2-50 km and the variable of gas plant within 2-50 km had a Pearson correlation coefficient of 0.93. We chose the variable of logarithm of distance weights for gas plant 2-50 km as a more specific variable into multivariate analyses. Variables of logarithms of distance weights for oil well within 2-50 km and battery within 2-50 km had a Pearson correlation coefficient of 0.82. We conducted two models which one used the variable of logarithms of distance weights for oil well within 2-50 km and another one used the variable of logarithms of distance weights for battery within 2-50 km. Then we chose a better-fit model judged by AIC and BIC.

Correlations between indicator variables of nearby potential sources of VOC from site documentation data set and their corresponding distances were examined and Spearman correlation coefficients were calculated (Appendix 3). Correlations between all indicator variables were examined using Phi correlation coefficient (ϕ) (Appendix 4). The Phi coefficient is a measure of association derived from the Pearson chi-square statistic. It has the range $-1 \leq \phi \leq 1$ for 2×2 tables. High correlations were found between indicator variables of presence of potential sources of VOC and their corresponding distances (Spearman correlation coefficients were all greater than 0.9). Since indicator variables of potential sources of VOC variables provided more accurate information than their corresponding distance (estimated by field technicians through visual observations), only indicator variables were considered further in multivariate models.

For multivariate analyses of determinants of exposure, we used mixed effects model of the following general form:

$$Y_{hijm} = \mu_y + \delta^{*}Season_m + \beta_1 X_{1hi} + \beta_2 X_{2hi} + \ldots \beta_k X_{khi} + \tau_h + \lambda_i + \varepsilon_{j\,(hi)} \qquad (6)$$

Here, $\beta_1, \beta_2, \ldots \beta_k$ represent regression coefficient of fixed effects of determinants of VOC exposure. X_{1hi}, X_{2hi}, $\ldots X_{khi}$ represent variables of fixed effects of determinants of VOC exposure at i^{th} sampling site and in h^{th} month. Y_{hijm}, μ_y, δ, τ_h, λ_I, $\varepsilon_{j\,(hi)}$ have the same meaning as in basic mixed model (Equation 4).

Stepwise variable selection approach was used to generate parsimonious models. The criteria of variable selection were p-value ≤0.05 and the direction of estimate that is plausible. For variables of logarithms of distance weights of oil and gas facilities, if variables of logarithms of distance weights within 2 km are not significant, even if variables of logarithms of distance weights within 2-50 km are significant, they are not considered in further modelling. Assumptions of mixed effects models were normality, linearity, and homoscedasticity. Whether those assumptions were satisfied was examined by standard residual plots (i.e. q-q plot of standardized residuals vs. normal score residuals, residuals vs. individual predicted values, residuals vs. predicted means).

Chapter 4: Results

4.1 Sampling frequency

For 26 individual VOC each, there were 11,399 measurements, and 60% of them (6846 measurements) were obtained in 2001 and 40% of them (4553 measurements) were obtained in 2002; out of the 11,399 measurements, there were 1,470 replicates (13%). The sampling frequencies by year and month are in Table 2.

Table 2 Sampling frequency by year and month

Year	Month	Air samples	Percent	Number of sites
2001	April	161	1.41	152
	May	637	5.59	599
	June	994	8.72	904
	July	1023	8.97	930
	August	1029	9.03	930
	September	1031	9.04	947
	October	898	7.88	791
	November	606	5.32	507
	December	467	4.10	371
2002	January	431	3.78	339
	February	409	3.59	319
	March	379	3.32	306
	April	391	3.43	317
	May	607	5.33	530
	June	888	7.79	801
	July	265	2.32	236
	August	292	2.56	260
	September	292	2.56	259
	October	274	2.40	238
	November	180	1.58	149
	December	145	1.27	117

Well(s) (including oil well, gas well) was the most common type of oil and gas facility in the vicinity of monitoring station, followed by road, plant, flare, and battery, etc. The frequencies of potential determinants of exposure and observations are shown in Table 3.

Table 3 Frequencies of potential determinants of exposure and observations

Source type	Source type	%	Count of samples	%
No industrial sources documented				
no industrial sources	432	34.81	4695**	41.19
source information missing	87	7.01		
No data	35(6)*	2.82	24	0.2
Battery	87	7.01	1079	9.47
Compressor	56	4.51	609	5.34
Flare	110	8.86	1350	11.84
Highway	74	5.96	831	7.29
Other industry	26	2.1	282	2.47
Other oil and gas facilities	57	4.59	667	5.85
Plant	124	9.99	1313	11.52
Road	218	17.57	1841	16.15
Well	311	25.06	3182	27.91
Vegetation	699	56.33	6793	59.59

Total documented sites=1241 Total observations=11399

*6 sites had a total of 24 measurements, but no data about sites, so these 24 observations were excluded from regression modeling.

** The number of samples measured at sites with no industrial sources and at sites with source information missing.

4.2 Observed concentrations and variability

26 VOC concentrations were in ng/m^3 quantities. Four VOC, pentachloroethane, 1,2,4-trichlorobenzene, styrene, and hexachloroethane, were mostly non-detectable. The percentage of non-detectable values of the above four VOC were 83.3%, 83.3%, 82.0%, and 66.9%, respectively. Table 4 lists the percentage of non-detectable values, arithmetic mean, geometric mean, geometric standard deviation, maximum concentrations, and IARC classifications of the 26 VOC. Among the 26 VOC, the GM of hexane concentration was the highest (520.1 ng/m^3), followed by toluene (206.9 ng/m^3), dichloromethane (167.3 ng/m^3), and benzene (158.0 ng/m^3). There was considerable variability exhibited in the air concentrations of VOC: GSD up to 15.7 for dichloromethane and on the order of 3 to 7 for the majority of VOC. Table 5 and Table 6 list the descriptive statistics by winter and summer separately. Geometric means of benzene, hexane, and toluene in winter were substantial higher than in summer, while geometric means of a-pinene, n-decane, and d-limonene in summer were substantial higher than in winter.

Table 4 Observed concentrations of the 26 VOC in rural areas of Western Canada (N=11,399)*

Volatile Organic Compound	Non-detectable (%)	Arithmetic Mean (ng/m^3)	Geometric Mean (ng/m^3)	Geometric Standard Deviation	Maximum Concentration (ng/m^3)	IARC Classification
Dichloromethane	12.72	1293.76	167.32	15.65	166049	2B
Hexane	2.98	1443.22	520.07	6.15	111827	3
Chloroform	2.1	79.1	62.67	2.42	8029	2B
Benzene	5.3	288.28	158.03	4.93	9036	1
1,2-Dichloroethane	10.4	25.2	14.49	3.99	298	2B
Trichloroethylene	25.1	19.49	5.87	5.54	4115	2A
Toluene	2.47	437.27	206.89	3.87	83509	2B
Tetrachloroethylene	7.86	37.99	22.89	3.62	6815	2A
Ethylbenzene	1.39	54.39	35.32	2.66	6217	2B
m/p-Xylene	0.6	97.93	57	2.82	18675	3
o-Xylene	3.78	58.24	32.45	3.46	3915	3
a-Pinene	17.95	76.56	16.89	7.95	7144	**
Styrene	81.96	2.58	0.79	2.92	808	2B
Cumene	30.04	8.11	3.16	4.26	397	**
n-Decane	25.94	64.99	13.63	9.05	6130	**
1,3,5-Trimethylbenzene	23.41	25.14	7.14	6.3	1577	3
1,1,2,2-Tetrachloroethane	48.21	9.67	2.49	5.59	1183	3
1,2,4-Trimethylbenzene	5.11	49.73	23.94	3.96	2129	3
d-Limonene	29.2	53.13	10.89	9.21	5264	3
Pentachloroethane	83.31	2.69	0.79	3.04	220	3
p-Cymene	7.63	18.57	11.47	3.27	1032	**
1,3-Dichlorobenzene	49.17	17.16	3.16	7.35	533	3
1,4-Dichlorobenzene	39.01	25.04	5.3	8.01	1134	3
Hexachloroethane	66.87	15.58	1.2	5.07	1480	2B
1,2,4-Trichlorobenzene	83.31	1.4	0.68	2.23	83	**
Naphthalene	29.59	8	3.15	4.23	1677	2B

* The minimum detectable concentration was 1 ng/m^3 for all compounds.

** Not evaluated by IARC Monograph program.

Table 5 Observed concentrations of the 26 VOC from November to April in rural areas of Western Canada (N=3,169)*

Volatile Organic Compound	Non-detectable (%)	Arithmetic Mean (ng/m^3)	Geometric Mean (ng/m^3)	Geometric Standard Deviation	Maximum Concentration (ng/m^3)
Dichloromethane	16.2	1700.18	133.71	21.79	150970
Hexane	2.7	2374.61	901.05	5.93	111827
Chloroform	1.4	84.96	72.55	2.08	641
Benzene	0.3	510.94	420.93	1.96	9036
1,2-Dichloroethane	11.6	32.08	18.13	4.43	236
Trichloroethylene	14.2	17.32	9.15	4.13	386
Toluene	2.4	720.9	278.25	4.24	83509
Tetrachloroethylene	7.4	48.4	32.41	3.52	773
Ethylbenzene	2.3	61.72	45.72	2.57	869
M/p-Xylene	0.5	90.58	59.00	2.66	1592
o-Xylene	4.5	57.24	35.12	3.44	877
a-Pinene	36.8	37.24	6.21	8.77	2048
Styrene	83.7	3.5	0.79	3.10	808
Cumene	35.2	8.24	3.10	4.54	193
n-Decane	56.8	28.47	3.01	9.07	2616
1,3,5-Trimethylbenzene	33.4	24.17	5.72	7.36	432
1,1,2,2-Tetrachloroethane	59.7	6.02	1.56	4.71	235
1,2,4-Trimethylbenzene	9.1	38.55	18.7	4.5	1378
d-Limonene	39.4	33.24	6.12	9.05	1341
Pentachloroethane	82.3	4.18	0.89	3.75	198
p-Cymene	12.3	15.92	9.02	3.74	245
1,3-Dichlorobenzene	64.2	14.59	1.72	6.56	415
1,4-Dichlorobenzene	61.5	18.35	1.98	7.23	503
Hexachloroethane	70.7	8.08	0.96	3.72	612
1,2,4-Trichlorobenzene	85.9	0.98	0.64	1.95	60
Naphthalene	32.4	7.98	3.31	4.49	180

* The minimum detectable concentration was 1 ng/m^3 for all compounds.

Table 6 Observed concentrations of the 26 VOC from May to October in rural areas of Western Canada (N=8,230)*

Volatile Organic Compound	Non-detectable (%)	Arithmetic Mean (ng/m^3)	Geometric Mean (ng/m^3)	Geometric Standard Deviation	Maximum Concentration (ng/m^3)
Dichloromethane	11.4	1137.27	182.4	13.55	166049
Hexane	3.1	1084.58	420.88	5.97	32041
Chloroform	2.4	76.84	59.24	2.53	8029
Benzene	7.2	202.54	108.37	5.39	5884
1,2-Dichloroethane	9.9	22.55	13.29	3.78	298
Trichloroethylene	29.3	20.33	4.95	5.95	4115
Toluene	2.5	328.06	184.58	3.66	76423
Tetrachloroethylene	8.0	33.98	20.03	3.57	6815
Ethylbenzene	1.0	51.56	31.98	2.65	6217
M/p-Xylene	0.6	100.75	56.25	2.88	18675
o-Xylene	3.5	58.62	31.47	3.46	3915
a-Pinene	10.7	91.70	24.81	6.68	7144
Styrene	81.3	2.22	0.79	2.85	424
Cumene	28.0	8.06	3.18	4.15	397
n-Decane	14.0	79.05	24.39	6.73	6130
1,3,5-Trimethylbenzene	19.6	25.51	7.77	5.86	1577
1,1,2,2-Tetrachloroethane	43.8	11.07	2.98	5.75	1183
1,2,4-Trimethylbenzene	3.6	54.03	26.33	3.72	2129
d-Limonene	25.2	60.78	13.59	8.91	5264
Pentachloroethane	83.7	2.12	0.75	2.76	220
p-Cymene	5.8	19.59	12.58	3.06	1032
1,3-Dichlorobenzene	43.4	18.16	3.99	7.30	533
1,4-Dichlorobenzene	30.4	27.62	7.75	7.33	1134
Hexachloroethane	65.4	18.47	1.31	5.59	1480
1,2,4-Trichlorobenzene	82.3	1.55	0.70	2.33	83
Naphthalene	28.5	8.00	3.09	4.13	1677

* The minimum detectable concentration was 1 ng/m^3 for all compounds.

4.3 Factor Analysis

Factor analysis of the 26 VOC concentrations of 1206 site-specific logarithmic means revealed that three factors accounted for a total of 47% of the multiple correlations among the 26 VOC. These factors (i.e. grouping of chemicals) appeared to be stable from month-to-month (data not shown). Factor I accounted for 29% of the total variance in the data set and appears to represent a mixture of hydrocarbons that are short-living pollutants from oil and gas operation and combustion. Factor I was associated most strongly with 8 aromatic species and hexane. The 8 aromatic VOC include BTEX, cumene, and trimethylbenzene isomers. Factor II accounted for 11% of the total variance and appears to represent a mixture of hydrocarbons that are associated with vegetation. Factor II was associated most strongly with 6 VOC: a-pinene, n-decane, d-limonene, p-cymene, 1,3-dichlorobenzene, and 1,4-dichlorobenzene. a-pinene, n-decane, d-limonene, and p-cymene are mainly emitted by vegetation. Dichlorobenzenes do not occur naturally. 1,3-dichlorobenzene and 1,4-dichlorobenzene are mainly used as fumigant and insecticide for moth control (24). Factor III accounted for 7% of the total variance in the data set and appears to represent a group of chlorinated VOC. Factor III was associated most strongly with 4 VOC: trichloroethylene, dichloromethane, tetrachloroethylene, and chloroform. They are used in a variety of business and industries. Trichloroethylene, dichloromethane, and tetrachloroethylene are mainly emitted into the air from solvent usage (14, 24) Chloroform is also used as solvent, but it is mainly emitted into the air from the chlorination of water in Northern America (14).

Seven VOC did not resolve into any factors (1,1,2,2-tetrachloroethane, naphthalene, 1,2-dichloroethane, styrene, pentachloroethane, hexachloroethane, and 1,2,4-trichlorobenzene). Out of these, five had a significant number of measurements below the limit of detection.

Among VOC that contributed to the three factors, only 1,3-dichlorobenzene (Factor II) had a significant proportion of non-detectable values (49.2%). Table 7 shows the factors and factor loadings (factor loadings ≥ | 0.5 | are highlighted).

Table 7 Results of factor analysis for 26 VOC measured at 1206 locations in rural Western Canada*

Volatile Organic Compound	Chemical formula	Non-detectable	loading		
		(%)	Factor I	Factor II	Factor III
Ethylbenzene	C8H10	1.39	**0.92**	-0.01	0.18
M/p-Xylene	C8H10	0.6	**0.94**	0.09	-0.08
o-Xylene	C8H10	3.78	**0.85**	0.24	0.05
Toluene	C7H8	2.47	**0.83**	0.04	0.26
1,2,4-Trimethylbenzene	C9H12	5.11	**0.80**	0.33	0.06
1,3,5-Trimethylbenzene	C9H12	23.41	**0.61**	0.35	0.10
Hexane	C6H14	2.98	**0.65**	-0.07	0.22
Cumene	C9H12	30.04	**0.59**	0.32	0.19
Benzene	C6H6	5.3	**0.52**	-0.42	0.24
1,3-Dichlorobenzene	C6H4CL2	49.17	-0.03	**0.68**	0.47
1,4-Dichlorobenzene	C6H4CL2	39.01	-0.01	**0.66**	0.34
a-Pinene	C10H16	17.95	0.20	**0.69**	0.17
n-Decane	C10H22	25.94	0.32	**0.66**	-0.10
d-Limonene	C10H16	29.2	0.05	**0.60**	0.30
p-Cymene	C10H14	7.63	0.47	**0.50**	0.17
Trichloroethylene	C2HCL3	25.1	0.02	0.09	**0.72**
Dichloromethane	CHCL2	12.72	0.08	0.13	**0.64**
Tetrachloroethylene	C2CL4	7.86	0.25	-0.10	**0.58**
Chloroform	CHCL3	2.1	0.33	0.04	**0.48**
1,1,2,2-Tetrachloroethane	C2H4CL4	48.21	0.26	0.39	0.01
Naphthalene	C10H8	29.59	0.36	0.07	0.11
1,2-Dichloroethane	C2H2CL2	10.4	0.20	-0.22	0.25
Styrene	C8H8	81.96	0.38	0.23	0.13
Pentachloroethane	C2HCL5	83.31	-0.13	-0.36	0.01
Hexachloroethane	C2CL6	66.87	-0.02	0.35	-0.14
1,2,4-Trichlorobenzene	C6H3CL3	83.31	-0.04	-0.06	-0.15

* Analyses were performed on location-specific logarithmic means; VARIMAX rotation employed. Proportion of variance explained by each factors: Factor I: 29%; Factor II: 11%; Factor III: 7%

4.4 Time Trends

In general, the time trends were not affected by exclusion of non-detectable values, except for the VOC with a significant proportion of non-detectable values.

There were no obvious patterns of seasonal time trends for VOC associated with different factors. Concentrations of toluene and benzene, which are strongly associated with Factor I, appeared to show clear seasonal effects with higher concentrations in winter months and lower concentrations in summer months. However, the other seven components of Factor I concentrations did not show clear patterns of seasonal effects. Concentration of n-decane, a-pinene, and 1,4-dichlorobenzene, which were strongly associated with Factor II, appeared to show clear seasonal effects with higher concentrations in summer months and lower concentrations in winter months, while the other three components of Factor II concentrations did not show clear patterns of seasonal effects. All the four chlorinated VOC, which were strongly associated with Factor III, did not show clear patters of seasonal effects. Figure 1, Figure 3, and Figure 5 illustrate benzene, n-decane, and chloroform (from the three factors) time trends. The other 23 VOC time trends graphs are shown in Appendix 5.

Monthly mean scores for each factor were calculated to generate time-trend for the three factors (Figure 2, 4, 6). The Factor I scores were higher in winter months from December 2001 to March 2002 and lower in summer months between April 2002 and August 2002, and higher again in October, November, and December, 2002. The scores of Factor II followed the similar time trends as n-decane, which the scores were much higher in summer months of 2001. The scores of Factor III did not show a clear pattern of seasonal variation, although time-trend was evident.

Figure 1 Observed time-trend of air concentrations of benzene with and without non-detectable (nd) values in rural areas of Western Canada

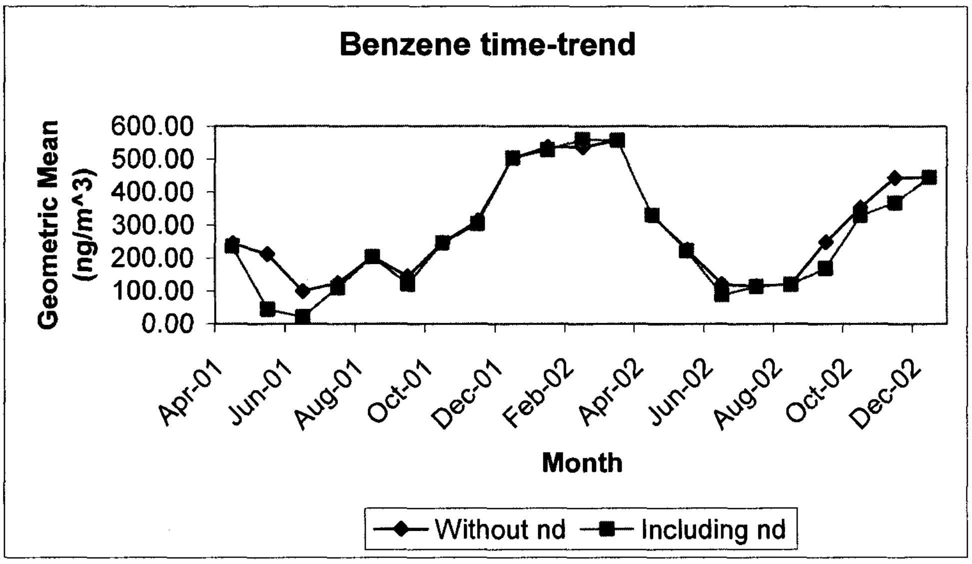


Figure 2 Observed time-trend of the scores of Factor I with non-detectable (nd) values in rural areas of Western Canada

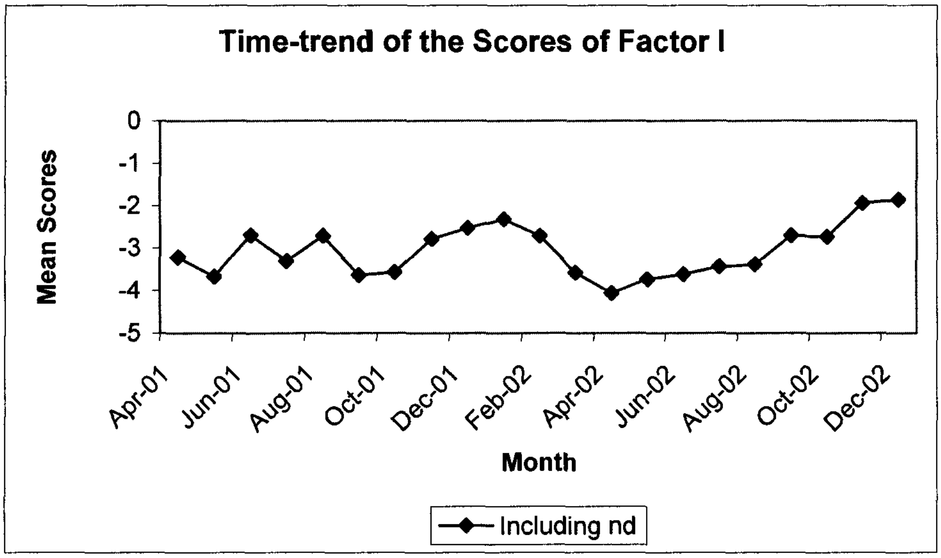


Figure 3 Observed time-trend of air concentrations of n-decane with and without non-detectable values (nd) in rural areas of Western Canada

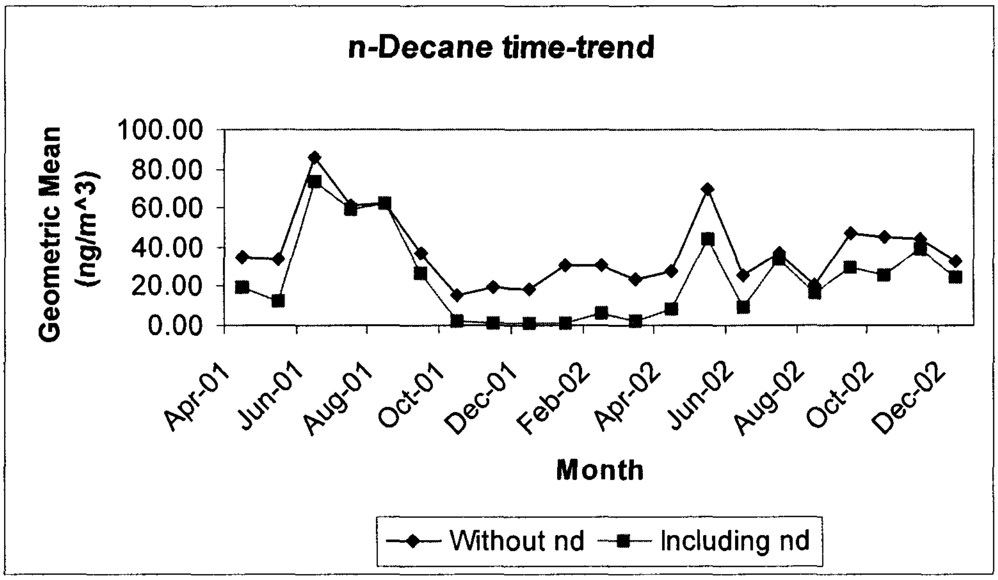


Figure 4 Observed time-trend of the scores of Factor II with non-detectable (nd) values in rural areas of Western Canada

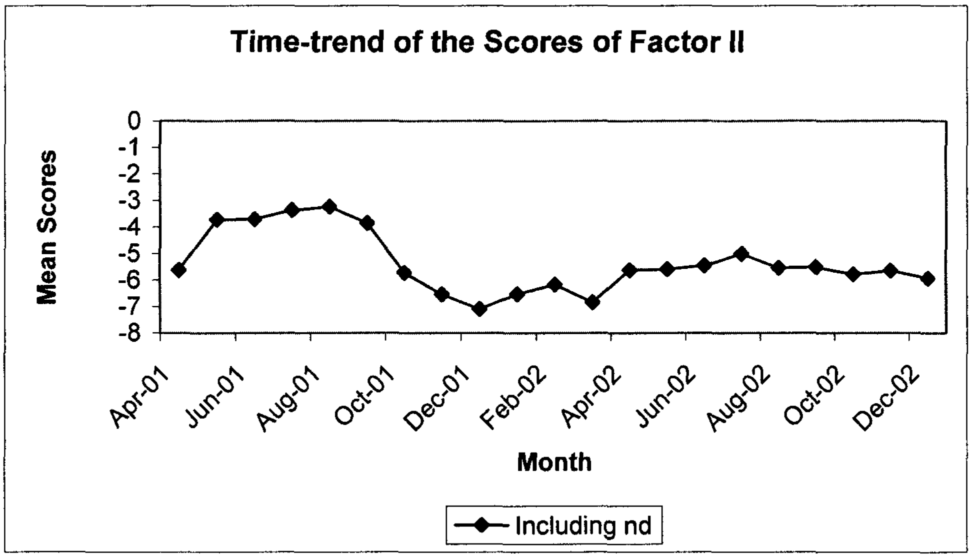


Figure 5 Observed time-trend of air concentrations of chloroform with and without non- detectable (nd) values in rural areas of Western Canada

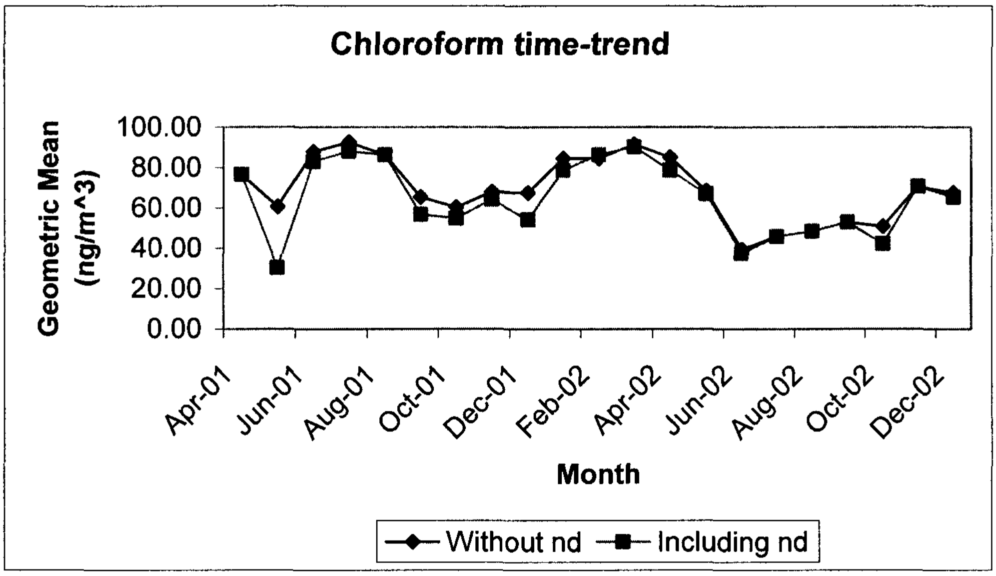
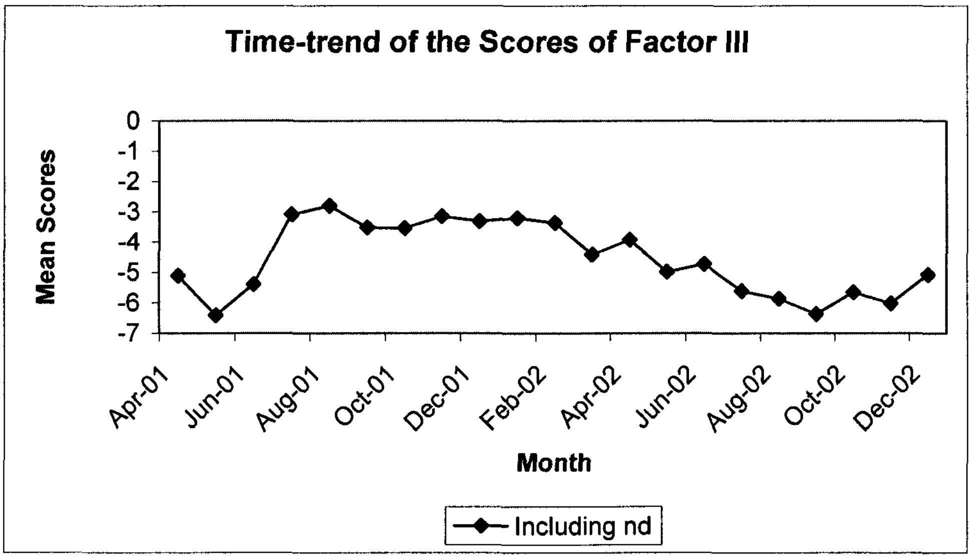


Figure 6 Observed time-trend of the scores of Factor III with non-detectable (nd) values in rural areas of Western Canada



4.5 Spatial correlation: semivariograms

Empirical semivariograms of the scores of Factor I revealed that spatial correlations existed within about 600 km range (Figures 7 and 8). The average spatial correlation coefficient within 5 km was 0.37. The example of detailed output of semi-variogram analysis for Factor I, such as lag, the number of measurements, average distance between measurements, regular semi-variance, covariance, and robust semi-variance within a given lag, is in Appendix 6 (only partial data (0-35 lag) are given). In Figure 7, we can see that the variance between measurements increase when the distances between measurements increase from 0 to around 600 km (range). After the distance reaches the range, the measurements become independent and the covariance between measurements approaches zero. However, there was another peak at around 1000 km range. What caused the peak needs further study. In general, it is clear that there is more variance between measurements of components of Factor I taken at locations separated by greater distance.

Empirical semivariograms (Figures 9 and 10) did not suggest significant spatial correlations in the scores of Factor II and Factor III (The average spatial correlation within 5 km for Factor II and Factor III was 0.06 and 0.07). Including spatial covariance matrix in basic mixed effects models and multivariate mixed effects models for Factor II and Factor III did not improve model fit (results not shown). Therefore, the spatial correlations were small and considered unimportant for Factor II and Factor III.

There was agreement among trends identified in regular and robust semivariograms.

Figure 7 Semivariograms for the score of Factor I using lag distance of 5 km, output distance of 2000 km

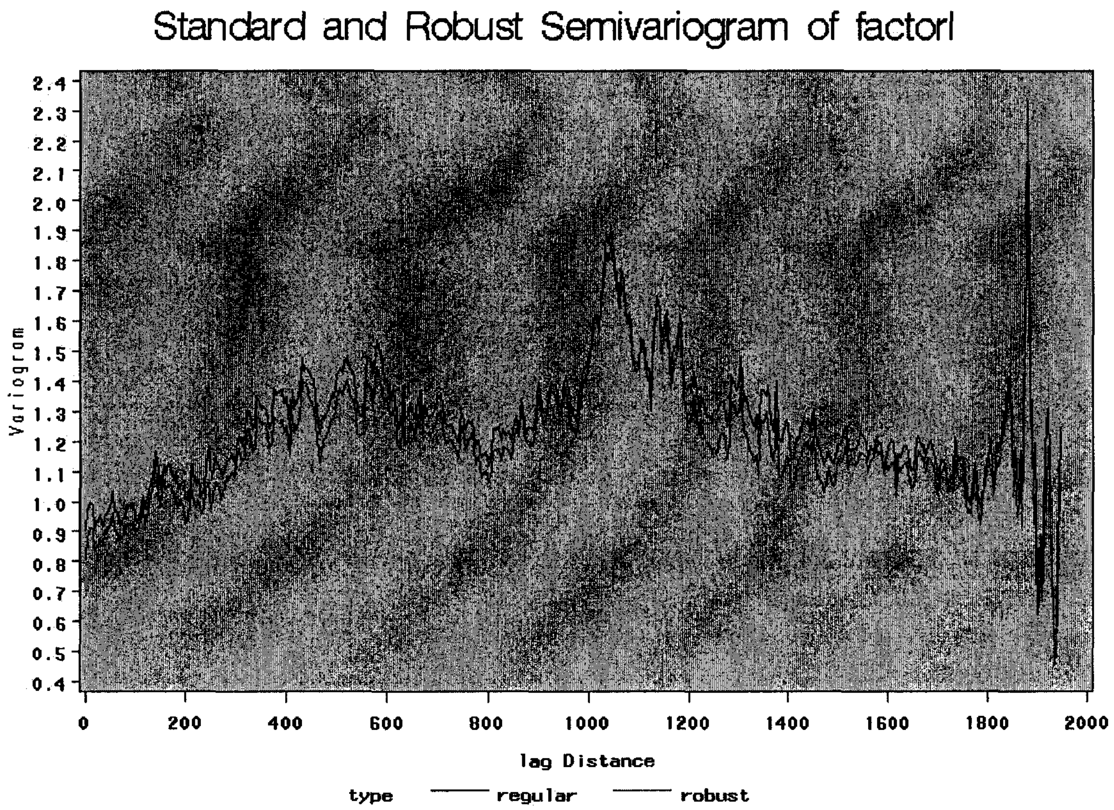


Figure 8 Semivariograms for the scores of Factor I using lag distance of 5 km, output distance of 600 km

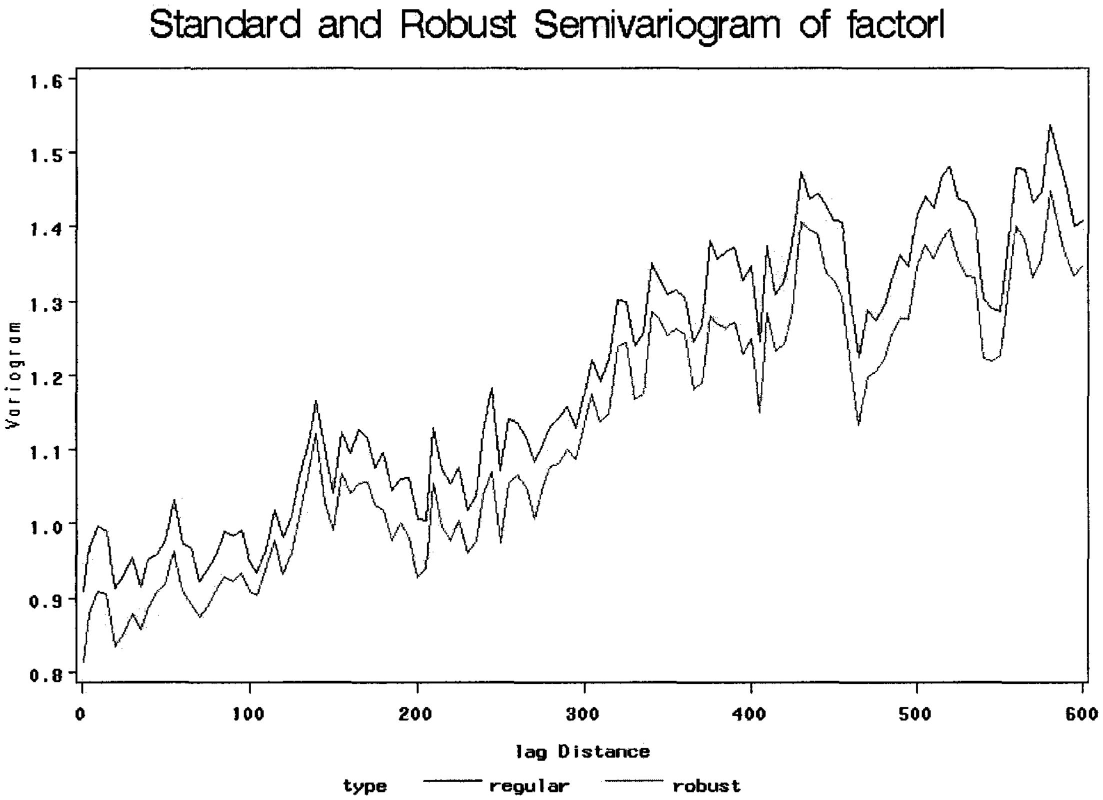


Figure 9 Semivariograms for the scores of Factor II using lag distance of 5 km, output distance of 2000 km

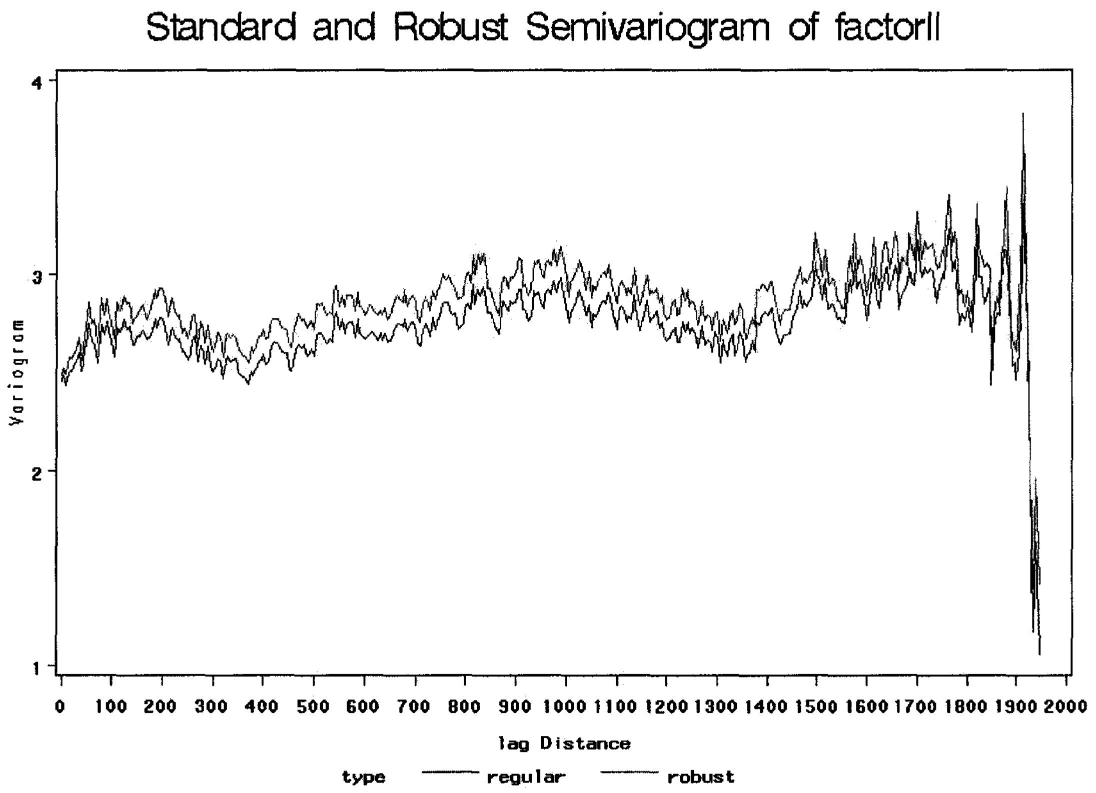
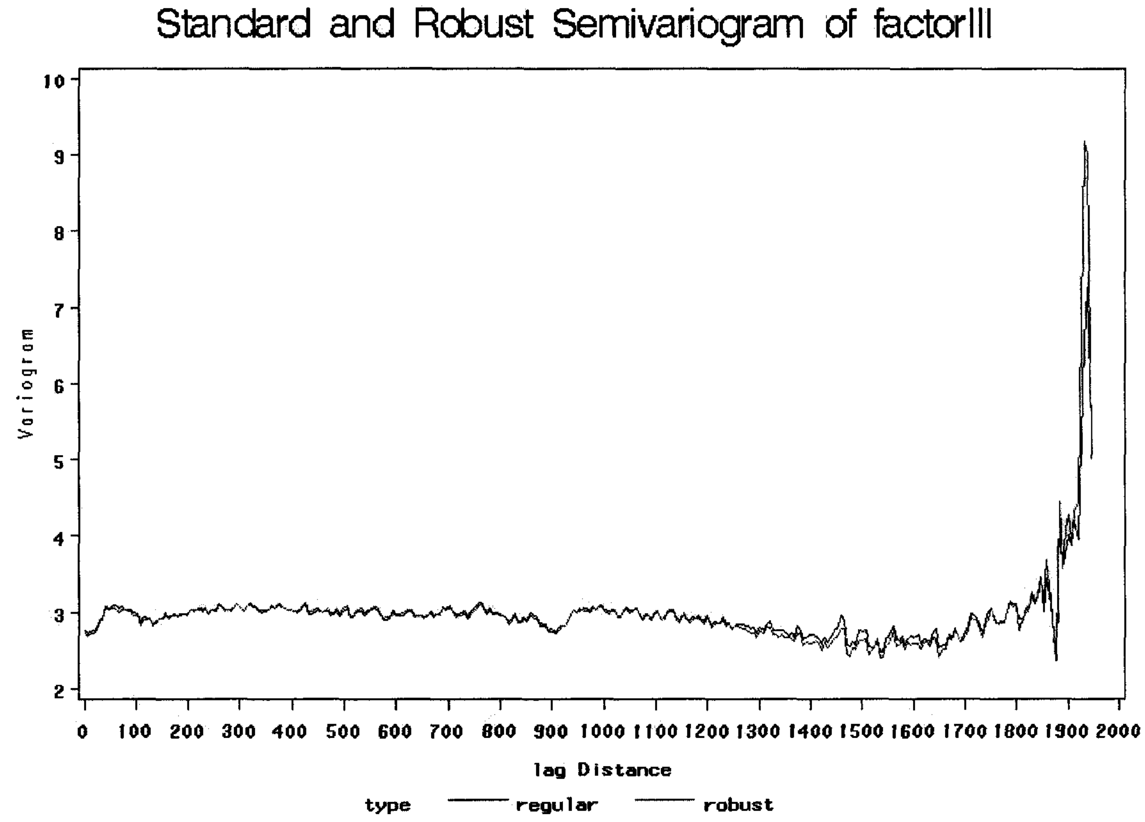


Figure 10 Semivariograms for the scores of Factor III using lag distance of 5 km, output distance of 2000 km



4.6 Description of factor scores

Each factor score is a linear combination of the observed 26 VOC and, in theory, represents a measure of an underlying process that is causing certain chemical species to behave similarly. The histograms (Figure 11, Figure 12, and Figure 13) revealed that the scores of Factor I, Factor II, and Factor III were approximately normally distributed.

Figure 11 Distribution of the scores of Factor I

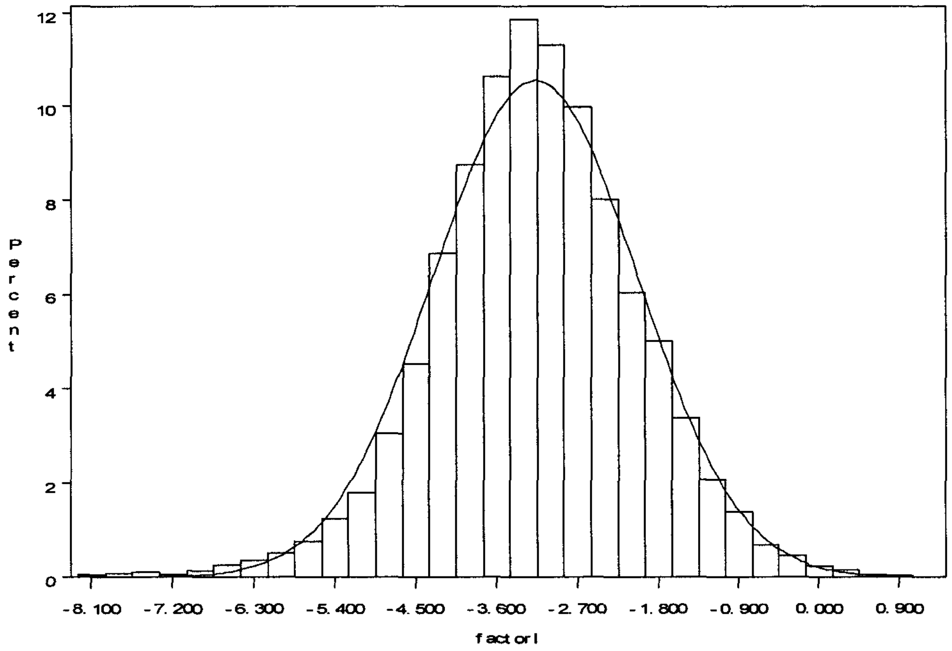


Figure 12 Distribution of the scores of Factor II

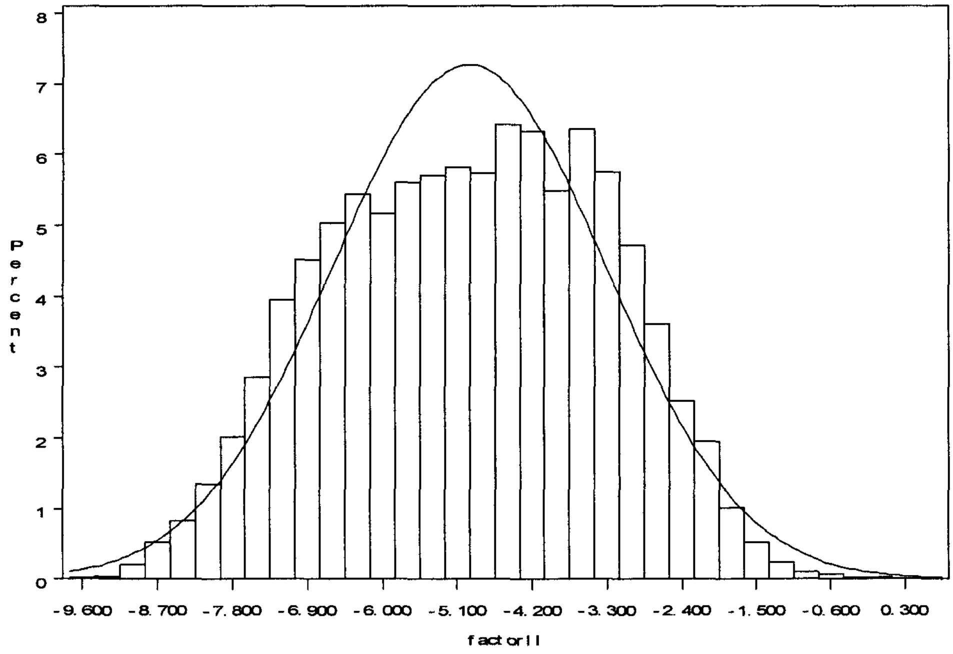
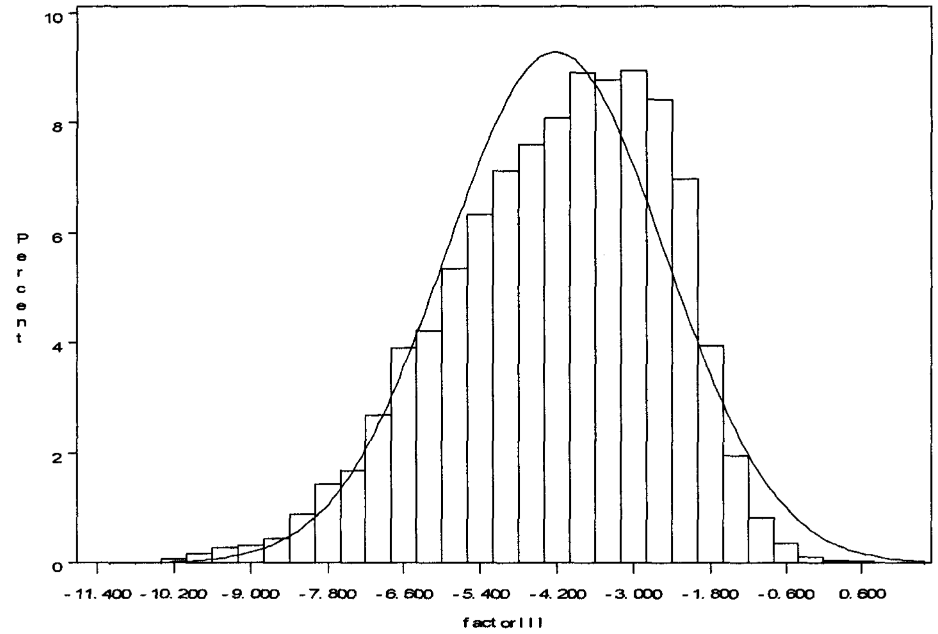


Figure 13 Distribution of the scores of Factor III



4.7 Correlations among nearby potential industrial sources and their corresponding distances from site documentation data

Spearman's correlation coefficients between presence of potential industrial sources and their corresponding distances and between presence of potential industrial sources and other sources' distances were computed. Spearman's correlation coefficients (r_s) are given in Appendix 3. Correlations between presence of potential industrial sources and their corresponding distances were large (r_s was greater than 0.9 for all cases). Correlations between presence of potential industrial sources and other sources' distances were small. Phi coefficients between presence of potential industrial sources and 'vegetation' were calculated. Phi coefficients (ϕ) are given in Appendix 4. Correlations between categorical variables representing potential industrial sources and 'vegetation' tended to be small, with the range of $|\phi|$ from 0.002 to 0.33.

Correlations between logarithms of distance weights from regulatory data

Correlations were computed for logarithms of distance weights for oil wells, gas wells, bitumen wells, batteries, all "large" facilities, and gas plants within 2 km and 2-50 km range. Pearson correlation coefficients (r) are given in Appendix 2. Some correlations were large. For example, the correlations between logarithms of distance weights of all large facilities 2-50 km and gas plant 2-50 km and all large facilities 2-50 km and batteries 2-50 km were r=0.92 and r=0.89 respectively.

4.9 Exposure variability partitioned between location, months and repeats

Spatial and temporal variability were quantified by fitting basic mixed effects models (described above). Table 8 summarizes variance component estimates obtained from basic mixed models of best fit. For the scores of Factor II and Factor III, substantial portion of total variability was due to errors in the measurement technique (as assessed by between repeated variance): 53% for Factor II, 52% for Factor III, respectively. For the scores of Factor I, it is complicated to assess what portion of total variability that was due to repeats. There was spatial correlation in the scores of Factor I, the spatial variability varied with distances between measurements. Temporal variability was assessed by seasonal effect and month-to-month variance. For the scores of Factor I, seasonal effect was significant ($p<0.04$), while month-to-month variance (within a season) accounted for a small portion of the total variability ($R_{0.95}=4$) and was not statistically significant. For the scores of Factor II, seasonal effect was significant ($p<0.0001$) and month-to-month variance (within a season) accounted for larger proportion (39%) of the total variability ($R_{0.95}=29$). For the scores of Factor III, seasonal effect was not significant ($p=0.20$), while month-to-month variance accounted for 46% of the total variability ($R_{0.95}=106$). Spatial variability was dominant for the scores of Factor I. However, spatial variability only accounted for a small portion of total variability for the scores of factor II and Factor III: 8% and 2%, respectively.

Table 8 Variance components from basic mixed models (N=11,375)

Dependent Variable	Variance Component (s^2)	Estimate	Standard Error	p- value	$R_{0.95}$ [a]	AIC [f]	BIC [f]
Scores of Factor I [d]	Between locations [b] (>=615 km)	1.429	0.086	<. 0001	108	25365.7	25369.9
	Month-to-month (within a season)	0.135	0.097	0.0815	4		
	Between repeats [c]	0.381	0.006	<. 0001	11		
Scores of Factor II [e]	Between locations	0.149	0.011	<. 0001	5	33464.5	33467.7
	Month-to-month (within a season)	0.730	0.238	0.0011	29		
	Between repeats [c]	1.001	0.014	<. 0001	51		
Scores of Factor III	Between locations	0.060	0.009	<. 0001	3	37889.4	37904.7
	Month-to-month	1.417	0.449	0.0008	106		
	Between repeats [c]	1.572	0.022	<. 0001	136		

[a] $R_{0.95}$ represents the fold range which includes 95% of the values.

[b] Because of existing spatial correlations in factor I, when the distance between measurements reaches 615 km range or more than 615 km, the variance equals to 1.429. If the distance between measurements is less than 615 km, the variance is smaller than 1.429.

[c] Repeats that were collected at the same location in the same month.

[d] The basic mixed effects model for the scores of Factor I included the fixed effect of season. The estimate of season (winter)=0.500, standard error=0.240, and P=0.037.

[e] The basic mixed effects model for the scores of Factor II included the fixed effect of season. The estimate of season (winter)=-1.513, standard error=0.378, and P<0.0001

[f] Fit statistics (Akaike and the Bayesian information criteria, respectively): smaller values indicate better model fit.

4.10 Univariate analysis

Univariate analyses were conducted by adding one predictor variable (as fixed effect) at a time to the basic mixed models once. Table 9 summarizes results of univariate analysis for the scores of Factor I. Variables of logarithms of distance weight of bitumen well 2-50 km, other oil and gas facilities, compressor, road and 'vegetation' were not statistically significant in univariate analysis. However, although the presence of road variable was not significant, the variable of distance of road was significant (p=0.016), therefore, variable of presence of road was considered for multivariate analyses. Among variables of logarithms of distance weights, high correlation was found between all large facilities (including gas plant) < 2 km and gas plant < 2 km (r=0.78); and high correlations were found between all large facilities (including gas plant) 2-50 km and gas plant 2-50 km (r=0.92) and between all large facilities 2-50 km and battery 2-50 km (r=0.89). All large facilities < 2 km and 2-50 km were not specific variables, so these two variables were not considered for multivariate analyses. Logarithms of distance weights of oil well 2-50 km and battery 2-50 km were highly correlated (r=0.82). It was hard to choose one from these two variables. Therefore, we decided to conduct separately models by using these two variables separately. Then we chose the best-fit overall model by examining fit statistics. However, neither the variable of logarithms of distance weights of oil well 2-50 km nor the variable of logarithms of distance weights of battery 2-50 km was statistically significant during multivariate analysis.

Table 9 The results of univariate analysis for the scores of Factor I (N=11,375)

Potential predictor	Estimate	Standard Error	p-value
Oil well <2 km	0.9148	0.07577	<0.0001
Oil well 2-50 km	0.2103	0.03793	<0.0001
Gas well < 2 km	2.0421	0.2168	<0.0001
Gas well 2-50 km	0.2598	0.05288	<0.0001
Gas plant <2 km	6.6238	3.3658	0.0491
Gas plant 2-50 km	4.7453	2.2028	0.0312
Battery <2 km	5.1540	0.5277	<0.0001
Battery 2-50 km	0.4748	0.1186	<0.0001
All large facilities (incl. Gas plants) <2 km	5.0485	1.3879	0.0003
All large facilities (incl. Gas plants) 2-50 km	4.9620	0.6726	<0.0001
Flare	0.08274	0.02238	0.0002
Highway	0.1098	0.02720	<0.0001
Other industry	0.2458	0.04544	<0.0001
Road*	0.01959	0.01887	0.2993
Bitumen well 2-50 km	-0.07271	0.1902	0.7788
Other oil and gas facilities	-0.01645	0.02929	0.5814
Compressor	0.05496	0.03180	0.0839
Vegetation	0.01087	0.01596	0.4958

*Variable of presence of road was not significant, but variable of distance of road was significant (p=0.0156), so variable of presence of road was put in next step of modeling.

Table 10 summarizes the results of the univariate analyses for the scores of Factor II. Even through variables of logarithms of distance weights of oil well, gas well, battery and all large facilities were statistical significant, their estimates were in the wrong direction. The presence of highway(s) was also negatively related to the scores of Factor II. Since Factor II VOC were vegetation-related, it is possible that the presence of highway(s) reduced the coverage of vegetation around the sample site area. Therefore, the variable of presence of highway was considered for multivariate analysis. Very small vegetation areas were cleared to install oil/gas facilities compared to highway, therefore the estimates for variables of logarithms of distance weights of oil well, gas well, battery and all large facilities were not plausible and we did not consider such variables in subsequent steps of model-building. An independent 'large' scale vegetation index would help interpret the models for the scores of Factor II.

Table 10 The results in univariate analysis for the scores of Factor II (N=11,375)

Potential predictor	Estimate	Standard Error	p-value
Oil well <2 km	-0.2716	0.1312	<0.0001
Oil well 2-50 km	-0.1928	0.0222	<0.0001
Gas well < 2 km	-0.7699	0.3256	0.0181
Gas well 2-50 km	-0.1783	0.0269	<0.0001
Gas plant <2 km	14.5011	7.6027	0.0565
Gas plant 2-50 km	-0.1298	1.0968	0.9058
Battery <2 km	-0.1868	0.9280	0.8405
Battery 2-50 km	-0.2545	0.0629	<0.0001
All large facilities (incl. Gas plants) <2 km	1.9564	3.0441	0.5204
All large facilities (incl. Gas plants) 2-50 km	-0.6513	0.2961	0.0279
Flare	0.0542	0.0510	0.2876
Highway	-0.1553	0.0622	0.0125
Other industry	0.0233	0.1057	0.8258
Road	0.0399	0.0412	0.3330
Bitumen well 2-50 km	-0.4220	0.2640	0.1100
Other oil and gas facilities	-0.0512	0.0703	0.4663
Compressor	0.05496	0.0318	0.0839
Vegetation	0.1377	0.0311	<0.0001

Table 11 summarizes the results of univariate analysis for the scores of Factor III. Even through those variables of logarithms of distance weights for oil well within 2-50 km, gas well within 2-50 km, battery within 2-50 km, gas plant within 2-50 km, all large facilities within 2 km and within 2-50 km were statistical significant, their estimates were in the wrong direction. This is not plausible, and such variables were not considered in subsequent model-building.

Table 11 The results of univariate analysis for the scores of Factor III (N=11,375)

Potential predictor	Estimate	Standard Error	p-value
Oil well <2 km	-0.2018	0.1167	0.0839
Oil well 2-50 km	-0.1265	0.02087	<0.0001
Gas well < 2 km	-0.3287	0.3033	0.2785
Gas well 2-50 km	-0.1122	0.0257	<0.0001
Gas plant <2 km	-5.5837	7.1020	0.4318
Gas plant 2-50 km	-5.7973	1.0057	<0.0001
Battery <2 km	-0.6801	0.8684	0.4336
Battery 2-50 km	-0.3228	0.0577	<0.0001
All large facilities (incl. Gas plants) <2 km	-7.1036	2.8879	0.0139
All large facilities (incl. Gas plants) 2-50 km	-1.5408	0.2701	<0.0001
Flare	-0.0029	0.0461	0.9504
Highway	-0.0255	0.0568	0.6537
Other industry	0.0276	0.0962	0.7739
Road	0.0617	0.0387	0.1109
Bitumen well 2-50 km	-0.2823	0.2436	0.2467
Other oil and gas facilities	-0.1150	0.0636	0.6537
Compressor	-0.0362	0.0658	0.5826
Vegetation	-0.0478	0.0293	0.1027

4. 11 Final mixed effects models: Analysis of the effect of proximity to putative sources on observed concentrations of VOC in the air

4.11.1 The scores of Factor I

The variance components from the final and parsimonious model for the scores of Factor I agree with the empirical semi-variogram very well (Table 12). Compared to the basic mixed model, addition of fixed effects reduced variance between locations from 1.429 to 1.325, where the distances between measurements reach about 600 km; the change in the estimate of random effect between repeats was negligible (reduced by 1%). Since the random effect of month was not significant in the basic mixed effect model, it was not included in the final mixed effects model.

The scores of Factor I appear to be positively related to proximity to oil well within 2 km, gas well within 2 km and 2-50 km, battery within 2 km, other industry, highways, and roads (Table 13). Seasonal variation was observed, with higher scores in winter and lower scores in summer. Standard residual plots, i.e. standardized residuals against their normal scores plot, residuals against the predicted values plot, and residuals against the predicted means, were generated for the final mixed effects model. Those residual plots did not indicate violations of model's normality, linearity, and homoscedasticity assumptions (Figures 14, 15 and 16). There was poor agreement between the observed scores of Factor I and predicted values (Figure 17): correlation coefficient of 0.29.

Table 12 Variance components of Factor I from final mixed effects model (N=11,375)

Dependent Variable	**Variance Component** (s^2)	**Estimate**	**Standard Error**	**p-value**	**$R_{0.95}$** [c]	**AIC** [d]	**BIC** [d]
Scores of Factor I	Between locations [a] (≥608 km)	1.325	0.086	<0.0001	91	25064.8	25067.9
	Between repeats [b]	0.377	0.006	<0.0001	11		

[a] Because of existing spatial correlation in factor I, when the distance between measurements reaches 608 km range or more than 608 km, the variance equals to 1.325. If the distance between measurements is less than 608 km, the variance is smaller than 1.325.

[b] Repeats that were collected at the same location in the same month.

[c] $R_{0.95}$ represents the fold range which includes 95% of the values.

[d] Fit statistics: smaller values indicate better fit (Akaike and the Bayesian information criteria, respectively)

Table 13 Mixed-effects multivariate model of the scores of Factor I (N=11,375)

Effect	Estimate	Standard Error	p Value
Intercept	-3.1959	0.1335	<0.0001
Winter	0.4780	0.1692	0.0108
Oil well <2 km	0.6872	0.07944	<0.0001
Gas well <2 km	1.1414	0.2524	<0.0001
Gas well 2-50 km	0.1262	0.05902	0.0324
Battery < 2 km	3.4012	0.5449	<0.0001
Highway	0.1051	0.02741	0.0001
Other industries	0.2047	0.04538	<0.0001
Road	0.03752	0.01891	0.0472

Figure 14 Qqplot of standardized residuals vs. normal score residuals: Final mixed effects model of the scores of Factor I

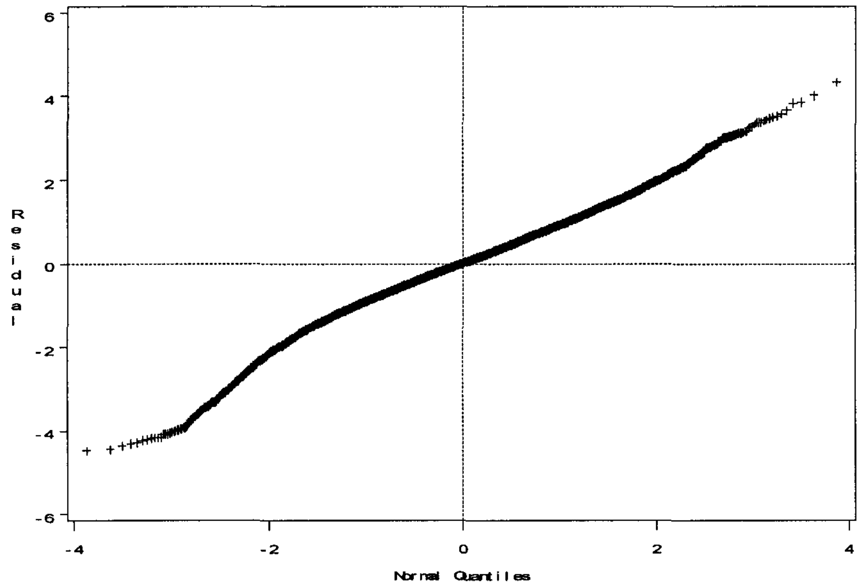


Figure 15 Plot of residuals vs. individual predicted values: Final mixed effects model of the scores of Factor I

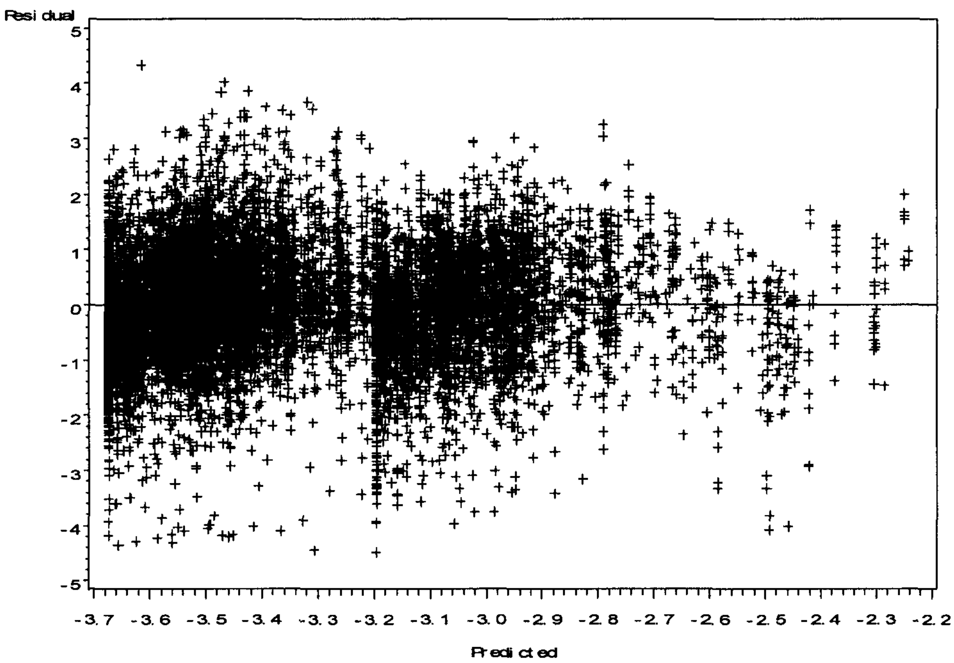


Figure 16 Plot of residuals vs. predicted means: Final mixed effects model of the scores of Factor I

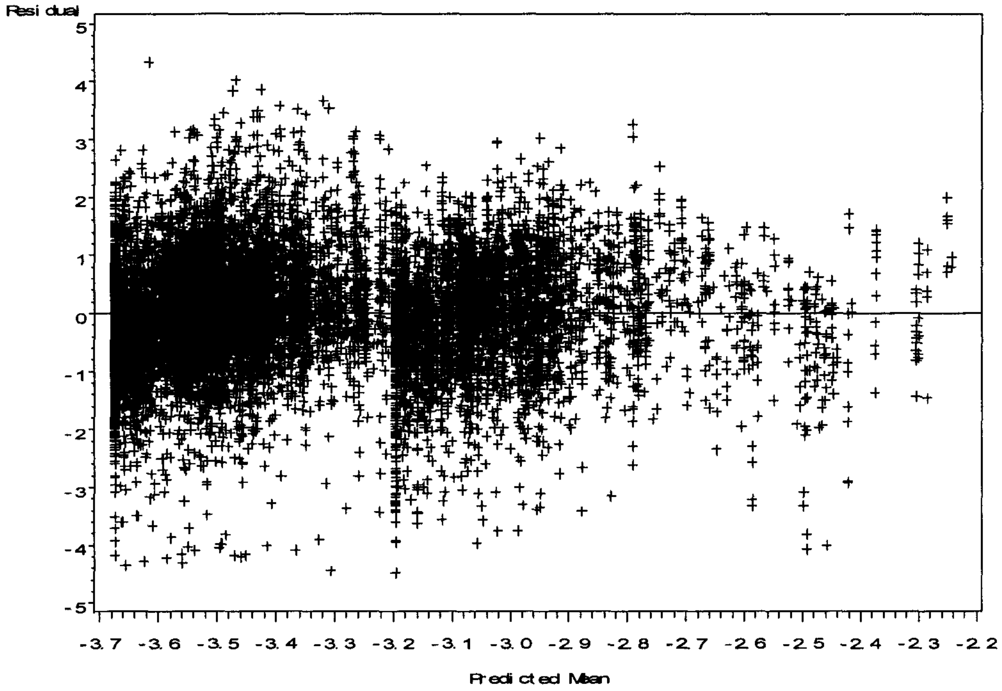
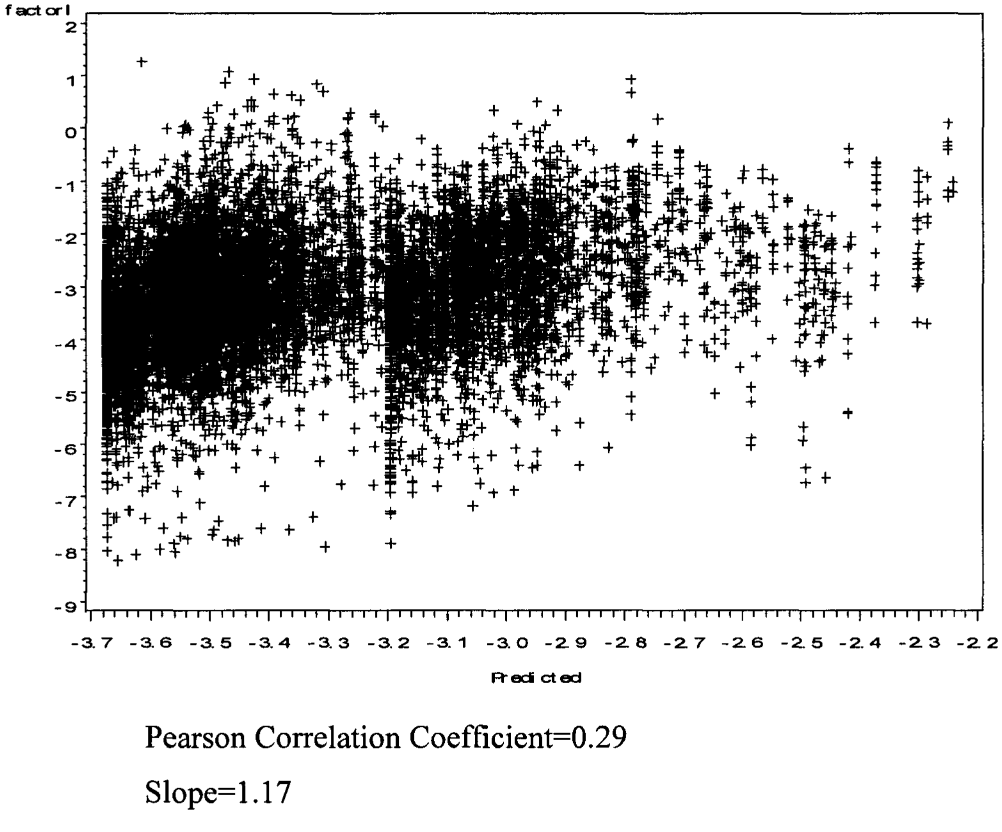


Figure 17 Observed scores of Factor I vs. predicted score generated from the final mixed effects model



Pearson Correlation Coefficient=0.29
Slope=1.17

4.11.2 The scores of Factor II

The main predictor of the scores of Factor II was the 'season'. Other minor predictors were the presence of vegetation ('trees or bushes') and highway(s). The factor scores were much higher in summer months than in winter months. The factor scores appear to be positively related to proximity to vegetation and negatively related to proximity to highway (Table 15). Addition of these fixed effects explains a very small portion of between-location variance (from 0.149 to 0.144). The random effect estimate of between months and repeats remained unchanged (Table 14). Standardized residuals against their normal scores plot, residuals against the predicted values plot, and residuals against the predicted means plot were generated for the final mixed effects model. Those residual plots did not indicate violations of model's assumption (Figures 18, 119, and 20). There was a good agreement between measured and predicted values of the scores of Factor II (Figure 21): the correlation coefficient of 0.81.

Table 14 Variance components of Factor II from final mixed effects model (N=11,375)

Dependent Variable	Variance component (s^2)	Estimate	Standard Error	p-value	$R_{0.95}$ [b]	AIC [c]	BIC [c]
Scores of Factor II	Between locations	0.144	0.011	<0.0001	2	33372.9	33388.2
	Month-to-month (within season)	0.730	0.237	0.0011	29		
	Between repeats [a]	1.001	0.014	<0.0001	51		

[a] Repeats that were collected at the same location in the same month.

[b] $R_{0.95}$ represents the fold range which includes 95% of the values.

[c] Fit statistics: smaller values indicate better fit (Akaike and the Bayesian information criteria, respectively)

Table 15 Mixed-effects multivariate model of the scores of Factor II (N=11,375)

Effect	Estimate	Standard Error	P Value
intercept	-6.2892	0.286	<0.0001
Winter	-1.5161	0.377	<0.0001
Vegetation	0.1333	0.031	<0.0001
Highway	-0.1384	0.062	0.0252

Figure 18 Qqplot of standardized residuals vs. normal score residuals: Final mixed effects model of the scores of Factor II

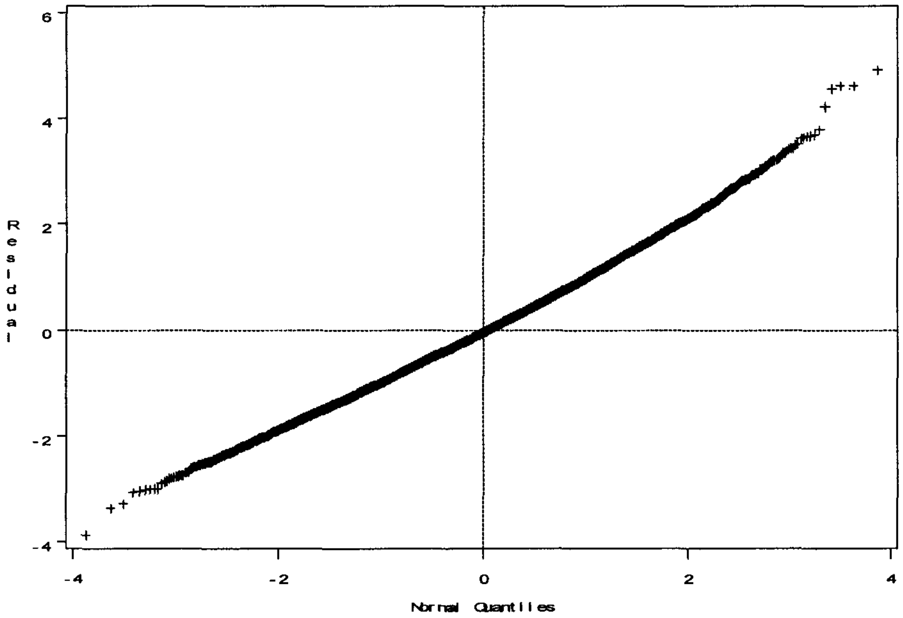


Figure 19 Plot of residuals vs. individual predicted values: Final mixed effects model of the scores of Factor II

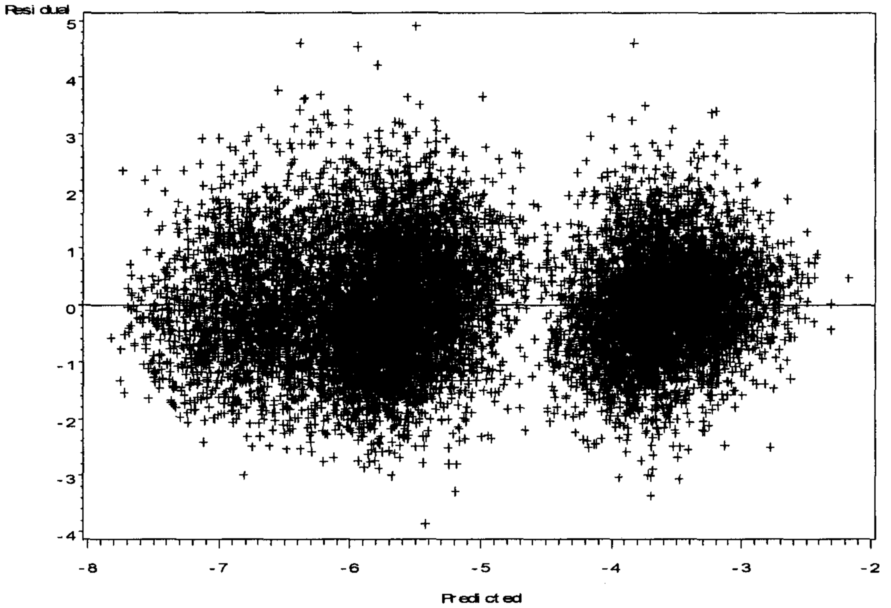


Figure 20 Plot of residuals vs. predicted means: Final mixed effects model of the scores of Factor II

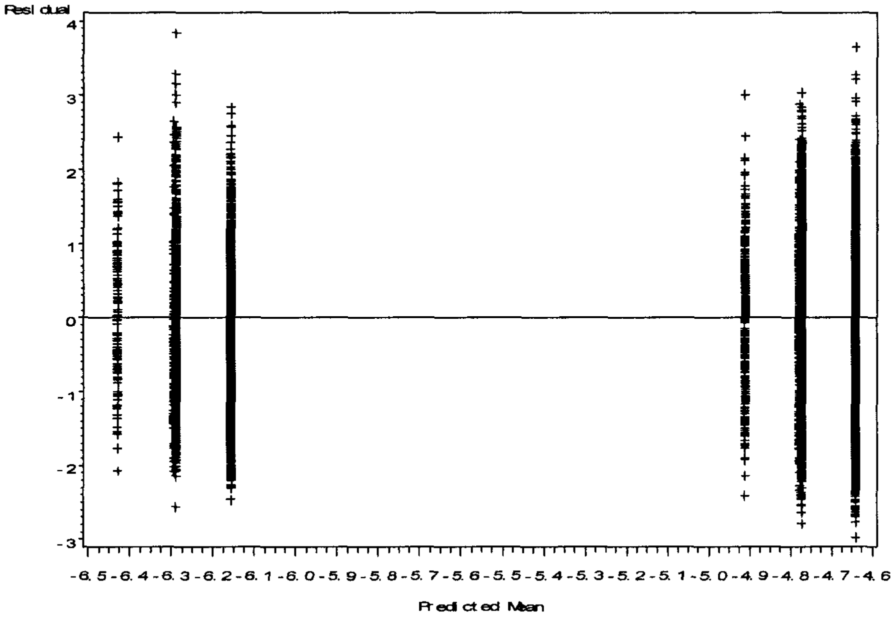
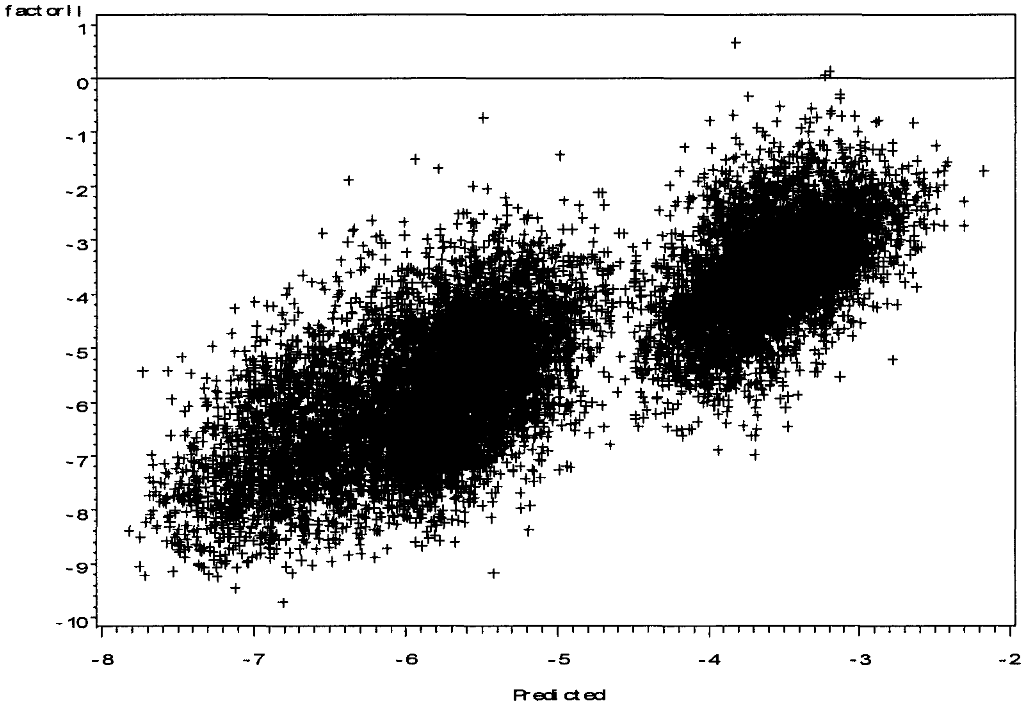


Figure 21 Measured the scores of Factor II vs. predicted score generated from the final mixed effects model



Correlation Coefficient=0.81

Slope=1.03

4.11.3 The scores of Factor III

No meaningful predictors were found for the scores of Factor III. Seasonal variation was not observed. The mixed effects model was not better than the random model that included random effects of months and random effects of locations. Potential predictors of the scores of Factor III maybe not have been captured in our data.

Chapter 5: Discussion and Conclusion

5.1 VOC concentrations

Hexane, benzene, toluene, ethylbenzene, m/p-xylene, o-xylene are common VOC measured in rural areas of Western Canada. The non-detectable percentages for those compounds were low, from 0.60% to 5.30%. Their geometric means of monthly concentrations were in the order of 30 to 500 ng/m^3. These VOC concentrations are much lower than those measured in urban and industrial sites. Cheng et al. reported VOC median of 24-hour average concentrations measured at downtown and industrial area monitoring sites in Edmonton, Alberta, over the period from December 1991 to November 1993. The median concentration of benzene, toluene, m/p-xylene, o-xylene, and hexane: in downtown of Edmonton (Alberta) were 3.38 μg/m^3, 7.09 μg/m^3, 4.56 μg/m^3, 1.94 μg/m^3 and 1.97 μg/m^3, respectively; in the industrial area: 2.60 μg/m^3, 4.56 μg/m^3, 2.84 μg/m^3, and 5.14 μg/m^3, respectively (17). Table 16, adopted from a recent review by Kindzierski, summarized median level or mean level of BTEX, styrene and naphthalene from urban and industrial monitoring sites in Canada, US and Italy (83). In our study, the geometric means of styrene and naphthalene were 0.79 ng/m^3 and 3.15 ng/m^3 respectively. The concentration of styrene and naphthalene measured in rural areas of Western Canada were about 10-20 times lower than median or mean concentrations typically measured in urban areas.

Table 16 Concentrations of selected VOC in air at various urban and industrial sampling sites (in μg/m^3) *

Site description	Ben	Tol	Eth	Xyls	Styr	Naph
Results reported as median concentrations						
Edmonton, Alta. 24-h samples every 6th day, 1991 to 1993 (n=106) [reference: (17)]						
Downtown business section	3.4	7.1	nr	4.6**	nr	nm
Industrial corridor	2.6	4.6	nr	2.8**	nr	nm
Various US monitoring sites (n>8000) [reference: (84,85)]	5.1	8.6	1.1	1.4**	0.6	1.2
18 Canadian sites, samples every 6th day, 1989 [reference: (18)]	3.2	8.4	nr	5.2	0.2	nr
Results reported as mean concentrations						
Columbus, Ohio, 3-h samples at 6 sites, summer 1989 (n=288) [reference: (86)]	1.5	5.1	1.1	4.9	nr	nm
Kanawha Valley, W.V., 12-h samples at 3 sites, 1987/88 [reference: (60)]	4.7	7	1.5	6.1	2.1	nm
Turin, Italy, 1991, urban (n=116) [reference: (61)]	44	65	nm	65	nm	nm

Note: Benzene (Ben), toluene (Tol), ehtylbenzene (Eth), total o/m/p-xylenes (Xyls), styrene (Styr), naphthalene (Naph), not reported (nr), not measured (nm).
*Adapted from reference (83)
**m/p-xylenes only

In our study, BTEX and hexane mean concentrations were in the order of 50 to 1500 ng/m^3. Table 17 summarized annual mean concentrations of BTEX and hexane measured in rural areas of Canada, Germany, and United Kingdom. The concentrations of BTEX measured in rural areas of Western Canada were much lower than those measured in the above rural areas. However, the mean concentration of hexane in our study was 1443 ng/m^3. It was 2-3 times higher than annual mean concentrations measured in Langenbrugge, Germany (57) and

West Beckham, UK (10). Hagerman et al. also reported seasonal mean and median concentrations of hydrocarbon compounds measured at the rural southeast United States from September 1992 to October 1993. They reported that the seasonal average concentrations of benzene were in the range of 0.54 to 12.86 ppbC. (22).

Table 17 Concentrations of selected VOC in air at various rural and remote sampling sites (in μg/m3)*

Site description	Ben	Tol	Eth	Xyls	Hex
Results reported as annual mean concentrations					
Langenbrugge, Germany [reference: (57)]	1.1	1.7	0.2	0.9	0.4
West Beckham, UK [reference: (10)]	2.3	3.8	1.1	3.4	0.5
Violet Grove, Alberta, Canada [reference: (14)]	1.7	13.9	1.0	3.0	nr

Note: Benzene (Ben), toluene (Tol), ehtylbenzene (Eth), total o/m/p-xylenes (Xyls), hexane (Hex), not reported (nr), not measured (nm).
*Violet Grove, Alberta, Canada is a rural area with a lot of oil and gas industrial activity.

It is worthwhile to point out that direct comparison of VOC levels between different sites should be treated with caution, because of differences in weather, topography and other conditions during sampling (14,17,60-61,84-85). Also, the sampling time and durations can vary between studies, as do the sampling and laboratory analysis methods used in different investigations.

In our study, among chlorinated VOC, the dichloromethane concentration was the highest, with arithmetic mean of 1293.8 ng/m^3 and geometric mean of 167.3 ng/m^3, and was followed by chloroform and tetrachloroethylene. It should be noted that the variability of dichloromethane concentration was very large (GSD=16) with a maximum concentration of 166,049 ng/m^3. In general, chlorinated VOC concentrations measured in rural area of Western Canada were slightly lower than those measured in Calgary-Central, Edmonton-Central and Edmonton-East monitoring sites from 1990 to 2000. For example, the mean concentration of chloroform in our study was 79.1 ng/m^3. The mean concentrations of chloroform measured in Calgary-Central, Edmonton-Central and Edmonton-East monitoring stations in 2000 were 130 ng/m^3, 140 ng/m^3, and 110 ng/m^3, respectively (14). The level of chloroform concentration measured in our study was comparable to the chloroform concentration measured in Pittsburgh area (Pennsylvania, USA) in 2002. The geometric means of chloroform measured in our study in winter and in summer were 72.55 ng/m^3 and 59.24 ng/m^3. The median concentrations in winter and in summer in Pittsburgh area were 53.8 ng/m^3 and 83.1 ng/m^3, respectively (80).

The geometric means of a-pinene, d-limonene and n-decane and p-cymene in our study were in the order of 10 to 17 ng/m^3, and the arithmetic means of the above monoterpenes were in the order of 10 to 80 ng/m^3. Millet et al. reported that a-pinene was below detection limit in winter and the median concentration in summer was 89.1 ng/m^3 in Pittsburgh area in 2002 (80). Hewitt et al. reported a-pinene and d-limonene day-time and night-time concentrations measured at Southwest Scotland Sitka Spruce Forest in 1994 autumn: a-pinene day-time and night-time concentrations were 150 ng/m^3 and 139 ng/m^3, respectively; d-limonene day-time and night-time concentrations were 89 ng/m^3 and 84 ng/m^3, respectively (63). Alberta Environment also reported average ambient a-pinene concentrations measured at Athabasca valley (Alberta, Canada) and Barge Landing (Alberta, Canada). The average concentrations of a-pinene measured at Athabasca valley in 1999 was 3342 ng/m^3, and measured at Barge

Landing in 2000 was 4234 ng/m^3 (14). In general, monoterpenes' concentrations measured at rural area of Western Canada were comparable to concentrations measured at Pittsburgh area (80) and Southwest Scotland Sitka Spruce Forest (63). However, they were lower than concentrations measured in other rural areas of Alberta (14).

The level of 1,4-dichlorobenzene in our study was lower than those measured at Calgary-Central, Edmonton-Central and Edmonton-East monitoring stations. The mean concentrations of 1,3-dichlorobenzene and 1,4-dichlorobenzene in our study were 17.16 ng/m^3 and 25.04 ng/m^3. 1,3-dichlorobenzene was not detectable at Calgary-Central, Edmonton-Central and Edmonton-East monitoring stations in 2000, and 1,4-dichlorobenzene mean concentrations at the above stations in 2000 were 180 ng/m^3, 130 ng/m^3, and 70 ng/m^3 respectively (14).

For the study period of year 2001 and 2002, the observed annual arithmetic means for benzene were 240 ng/m^3 and 361 ng/m^3, for toluene were 398 ng/m^3 and 496 ng/m^3, and for total xylene were 178 ng/m^3 and 123 ng/m^3, respectively. In comparison to Australian Environment Protection and Heritage Council limits (74) (i.e. annual arithmetic mean standards for benzene 0.003 ppm (9565 ng/m^3), toluene 0.1 ppm (376,000 ng/m^3), and total xylene 0.2 ppm (867,000 ng/m^3), the concentration of benzene, toluene and total xylene measured in the rural areas of Western Canada were far below these limits. Furthermore, no single observation exceeded the limits. There were 12 sites (1%) in 2001 and 13 sites (1%) in 2002 that the annual arithmetic mean of benzene exceeded 1/10 of benzene limit (956.5 ng/m^3). There was 1 site in 2001 and no sites in 2002 that the annual arithmetic means of toluene exceeded 1/10 of toluene limit (37,600 ng/m^3). There were no sites both in 2001 and 2002 that the annual arithmetic mean of total xylene exceeded 1/10 of total xylene limit (86,700 ng/m^3).

5.2 Variability of VOC concentrations

For the scores of Factor I, the variability of across locations was dominant. In contrast, for scores of Factor II and Factor III, variability across locations was much smaller than temporal variability; the variability of between repeats was the largest variance component. There was limited knowledge about relative importance of spatial and temporal variability in the airborne concentrations of VOC prior to this study. Our results suggest that the dominance of one variance component over another depends on VOC species. Importance of obtaining repeated measurements is highlighted by the fact that for some VOC between-repeat variance was substantial and any inference about determinants and levels of VOC has to take this into account.

Determinants of concentrations of mixtures of VOC

5.3.1 The scores of Factor I

The central research question of this study was whether emissions from oil and gas facilities contribute to environmental concentrations of VOC. Aromatic hydrocarbons such as BTEX, and hexane, previously known to be mainly emitted by upstream oil and gas industry operation and oil and gas combustion, clustered to form Factor I. Mixed-effects multivariate models were developed to investigate the relationship between proximity to oil and gas facilities and the scores of Factor I. The model suggests that oil wells (within 2 km), gas wells (within 2 km and 2-50 km), and battery (within 2 km) contribute to environmental concentrations of VOC that are strongly associated with Factor I, even after accounting for temporal and spatial variation and effects of contribution from motor vehicles (highway and road) and proximity to a city or town. The model indicates that as the number of oil wells, gas wells and batteries within a given radius increase, the expected score of Factor I and the concentrations of components of Factor I also increase. Emissions from batteries (within 2 km) were most influential in determining monthly airborne concentrations of components of Factor I, followed by gas wells within 2 km. The model also suggests that oil and gas facilities located at distances within 2 km were more influential in determining airborne concentrations of components of Factor I, compared to those located further away.

Limited knowledge of effects of emissions from *specific* oil and gas facilities on environmental air concentrations of VOC was available prior to this study. The available reports only provided the amount of total hydrocarbons emissions from oil and industry sectors and the percentages of total hydrocarbons emitted by different infrastructures within industry sectors. Canadian Association of Petroleum Producers observed that battery had much greater hydrocarbon emission strengths than other oil and gas facilities such as oil well, gas well and gas compressor (16). This is consistent with our findings. It should be pointed out that our findings should not be taken to mean that batteries, gas wells and oil wells are the most important sources for explaining the variation in the components of Factor I. They indicate that at the position of monitoring stations, the effects of emissions from batteries, gas wells and oil wells were more influential in determining monthly airborne concentrations of components in Factor I than other studied sources.

The models also indicate that emissions from cities and/or towns, highways and roads contribute to environment concentrations of components of Factor I. Within this group of determinants, proximity to a city or town was more influential in determining airborne concentrations of Factor I VOC, followed by highway. This reflects source emission strengths and is in accordance with the literature. Elevated airborne VOC concentrations in a city or town by numerous emission sources such as manufacturing factories, shops, motor vehicles were observed worldwide (1,21,59,60,87). VOC can dilute from higher concentration in urban to surrounding rural areas. Road vehicle emissions contribute significantly to air pollution, because vehicle exhaust is a major source of BTEX and styrene on roadways (83). Studies report that concentrations of VOC related to vehicle combustion are influenced by vehicle volume, composition, and speed (62,78,88). This is consistent with our findings that proximity to a highway was more influential than proximity to a road in determining airborne concentrations of components in Factor I

The empirical semi-variograms indicated that there existed spatial correlation for the observed scores of Factor I within approximately 600 km in this rural study area. Variance component derived from the final mixed effects model of the scores of Factor I agreed with the semi-variogram. Little was known about spatial correlation among environmental air concentrations of VOC prior to this study. Seasonal variation was observed with higher scores in winter and lower scores in summer. For individual VOC, BTEX and hexane concentrations displayed higher concentrations in winter months and lower concentration in summer. These observed seasonal variations of BTEX and hexane not only reflect enhanced photo-chemically driven processes in summer, but also greater emission rates in winter. For instance, winter-time heating and vehicle 'cold start' effect may be responsible for higher emissions of combustion related products. However, cumene, 1,2,4-trimethylbenzene, and 1,3,5-trimethylbenzene did not show the same seasonal pattern. They displayed slightly higher concentrations in summer months. Potential higher evaporation rate and leakage rate form oil and gas facilities in summer may be responsible for this, but we lack the data to investigate these matters further.

The seasonal variation of anthropogenic non-methane hydrocarbons was previously observed in rural and urban environments worldwide. Jobson et al. reported that alkane compounds and acetylene concentrations displayed a winter maxima and summer minima at a remote boreal site in Canada (58). Penkett et al. observed similar seasonal trends in hydrocarbon concentrations in ambient air over the North Atlantic Ocean (89). Seasonal variations of ambient air hydrocarbons were also observed in rural eastern France (20) and rural New England (19), as well as in city of Edmonton, Alberta, Canada (17) and in city of Seoul, S. Korea (59). The winter maxima and summer minima of hydrocarbon concentration has been attributed to hydroxyl chemistry and the seasonal abundance of the OH radical (19,20,58,59,89). Seasonal variation in hydrocarbon source strengths, and differences in atmospheric behaviour such as increased convection and vertical mixing in the summer, and differences in air mass climatology, also play a role in the hydrocarbon seasonal variation (20,58,59).

5.3.2 The scores of Factor II

Factor II is most strongly associated with monoterpenes and dichlorobenzene, which are related to vegetation. Our results suggest that oil and gas facilities do not contribute to environmental concentrations of these vegetation-related VOC. As expected, the models indicate that proximity to trees and/or bushes is influential in determining environmental concentrations of components in Factor II. Seasonal variation with minima in winter and maxima in summer was observed for the scores of Factor II and concentrations of all components in Factor II. The seasonal pattern of a-pinene, n-decane, d-limonene and p-cymene may be attributed to higher temperature and sunlight intensity in summer than in winter in Western Canada; 1,4-dichlorobenzene and 1,3-dichlorobenzene may be mainly applied to vegetation as pesticides in summer, and this may be responsible for their seasonal pattern. However, we do not have the information on types of pesticides used on the vegetation around the sampling sites. The pesticide application information will be very useful for future studies that investigate seasonal variation of dichlorobenzenes. The seasonal variation of monoterpenes with highest concentration in summer and lowest concentration in winter is in accordance with the literature. Hagerman et al. reported that naturally emitted monoterpenes such as a-pinene, b-pinene and d-limonene showed a seasonal distribution with lowest concentrations in the winter and highest in summer; and, unlike isoprene, the variation of monoterpenes concentration levels were less dramatic than isoprene in the rural southeast United States (22). Emissions of monoterpenes have been studied extensively because of their role in atmospheric photochemistry and the formation of tropospheric ozone. The monoterpene emission rates from different plants depend on temperature and light intensity (64,67,68).

5.3.3 The scores of Factor III

Factor III is most strongly associated with chlorinated VOC. Our results suggest that oil and gas facilities do not contribute to environmental concentrations of chlorinated VOC. No clear seasonal pattern was observed for chlorinated VOC in this study. In the atmosphere, chlorinated VOC have long lifetime and may be transported over great distances before ultimately being degraded. For example, the dominant degradation process of chloroform in the atmosphere is reaction with hydroxyl radical, and the half-life of chloroform is about 80 days based upon a 12-hour sunlit day in a typical atmosphere containing $1\text{x } 10^6$ hydroxyl (90). The relative non-reactivity and long lifetime may be contributed to small spatial variation and no obvious seasonal variation for chlorinated VOC. The concentrations of chlorinated VOC reflect not only emissions from local sources, also emissions that have taken place elsewhere and have become integrated into the regional background.

5.4 Conclusions

The analysis of the multi-location and multi-month VOC concentration data set has provided estimates of the 26 VOC levels and their variability in rural Western Canada. In general, the concentration levels of 26 VOC were low (in ng/m^3 quantities) and, for selected VOC (benzene, toluene, and total xylene), with exposure levels below some regulatory guidelines. However, the variability of VOC concentrations was substantial. Three factors were extracted from factor analysis. Factor I is most strongly associated with compounds previously suspected to be emitted from oil and gas operation and oil and gas combustion. Factor II is most strongly associated with vegetation-related VOC. Factor III is most strongly associated with chlorinated VOC. Different factors represent distinct emission sources of VOC groups. Linear mixed-effects multivariate models were developed to investigate the relationship between proximity to oil and gas facilities and the scores of factors. They have provided important insights on factors that influence temporal and spatial distributions of VOC in rural Western Canada. However, most of the determinants available to us did not vary in time and therefore we were better equipped to identify determinants of spatial, rather than temporal, variability. BTEX and hexane showed a seasonal variation with maxima in winter and minima in summer, while monoterpenes and dichlorobenzene displayed that opposite seasonal trend. Chlorinated VOC did not show a clear seasonal pattern. We observed that oil and gas facilities contribute to airborne concentrations of BTEX and hexane and associated VOC, but they do not contribute to airborne concentrations of monoterpenes, dichlorobenzene and chlorinated VOC. Emissions form batteries are most influential in determining monthly airborne concentrations of BTEX, hexane, and associated VOC, followed by gas and oil wells. Modification of batteries to reduce evaporation and leakage from batteries may be considered as an effective and efficient measures of control airborne concentrations of the most influential components of Factor I: BTEX and hexane. Building batteries in a different location (>2 km from houses or pastures) may also be considered to help reduce exposure to VOC emissions from batteries.

This thesis presents the first comprehensive statistical models of VOC air concentrations and their determinants in rural Western Canada. These models can help in the development of air pollution control measures and design of studies investigating the health effects of VOC.

References

(1) Derwent RG. Sources, Distributions, and Fates of VOCs in the Atmosphere. Issues in Environmental Science and Technology (4). Volatile Organic Compounds in the Atmosphere. 1st ed. Thomas Graham House, Science Park, Milton Road, Cambridge, UK: The Royal Society of Chemistry; 1995. p. 1.

(2) Canadian Council of Ministers of the Environment (CCME). Canadian 1996 NOx/VOC Science Assessment. Ground-Level Ozone and Its Precursors, 1980-1993---Report of the Data Analysis Working Group. 1997.

(3) Buffler PA, Crane M, Key MM. Possibilities of detecting health effects by studies of populations exposed to chemicals from waste disposal sites. Environmental health perspectives 1985 Oct; 62:423-456.

(4) Enterline PE. Review of new evidence regarding the relationship of gasoline exposure to kidney cancer and leukemia. Environmental health perspectives 1993 Dec;101 Suppl 6:101-103.

(5) Gamble JF, Pearlman ED, Nicolich MJ. A nested case-control study of kidney cancer among refinery/petrochemical workers. Environmental health perspectives 1996 Jun; 104(6):642-650.

(6) Geschwind SA, Stolwijk JA, Bracken M, Fitzgerald E, Stark A, Olsen C, et al. Risk of congenital malformations associated with proximity to hazardous waste sites. American Journal of Epidemiology 1992 Jun 1; 135(11):1197-1207.

(7) Mehlman MA. Dangerous and cancer-causing properties of products and chemicals in the oil refining and petrochemical industry. VIII. Health effects of motor fuels: carcinogenicity of gasoline--scientific update. Environmental research 1992 Oct;59(1):238-249.

(8) Mehlman MA. Dangerous and cancer-causing properties of products and chemicals in the oil-refining and petrochemical industry--Part XXII: Health hazards from exposure to gasoline containing methyl tertiary butyl ether: study of New Jersey residents. Toxicology and industrial health 1996 Sep-Oct;12(5):613-627.

(9) Hass U, Lund SP, Simonsen L, Fries S. Effects of Prenatal Exposure to Xylene on Postnatal Development and Behavior in Rats. Neurotoxicology and Teratology 1995;17:341-349.

(10) United Kingdom Photochemical Oxidants Review Group. Ozone in the United Kingdom. 1993.

(11) Uno I, Wakamatsu S, Wadden RA, Konno S, Koshio H. Evaluation of hydrocarbon reactivity in urban air. Atmospheric Environment 1985; 19:1283-1293.

(12) Burnett RT, Dales RE, Raizenne ME, Krewski D, Summers PW, Roberts GR, et al. Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. Environmental research 1994 May; 65(2):172-194.

(13) Thurston GD, Ito K, Kinney PL, Lippmann M. A multi-year study of air pollution and respiratory hospital admissions in three New York State metropolitan areas: results for 1988 and 1989 summers. Journal of exposure analysis and environmental epidemiology 1992 Oct-Dec; 2(4):429-450.

(14) Alberta Environment. Approaches to a Total (or Grouped) VOC Guideline
Final report. 2002; Available at: http://www3.gov.ab.ca/env/protenf/publications/approaches_VOC_guideline.pdf. Accessed June 7, 2005.

(15) Environment Canada. 2000 Criteria Air Contaminant Emissions for Canada (tonnes). 2004; Available at: www.ec.gc.ca/pdb/ape/ape_tables/canada2000_e.cfm. Accessed October 12, 2005, 2005.

(16) Canadian Association of Petroleum Producers. CH4 and VOC Emissions from the Canadian Upstream Oil and Gas Industry-Volume 1. 1999; Available at: http://www.capp.ca/raw.asp?x=1&dt=ntv&e=pdf&dn=84180. Accessed August 8, 2005, 2005.

(17) Cheng L, Fu L, Angle RP, Sandhu HS. Seasonal Variations of Volatile Organic Compounds in Edmonton, Alberta. Atmospheric Environment 1996; 31:239-246.

(18) Volatile organic compound measurements in Canadian urban and rural areas. In proceeding 85th Air Waste Management Association Annual Meet: Air &Waste Management Association; 1992.

(19) Goldstein AH, Wofsy SC, Spivakovsky CM. Seasonal variations of nonmethane hydrocarbons in rural New England: Constraints on OH concentrations in northern midlatitudes. Journal of Geophysical Research 1995; 100:21023-21033.

(20) Borbon A, Coddeville P, Locoge N, Galloo JC. Characterising sources and sinks of rural VOC in eastern France. Chemosphere 2004; 57:931-942.

(21) Chan CY, Chan LY, Wang XM, Liu YM, Lee SC, Zou SC, et al. Volatile Organic Compounds in roadside microenvironments of metropolitan Hong Kong. Atmospheric Environment 2002; 36(12): 2039-2047.

(22) Hagerman LM, Aneja VP, Lonneman WA. Characterization of Non-Methane Hydrocarbons in the Rural Southeast United States. Atmosphere Environment 1997; 31:4017-4038.

(23) Zhao W, Hopke PK, Karl T. Source Identification of Volatile Organic Compounds in Houston, Texas. Environmental Science & Technology 2004; 38:1338-1347.

(24) ATSDR. Finalized Toxicological Profiles. 2005; Available at: http://www.atsdr.cdc.gov/toxprofiles. Accessed July 2005, 2005.

(25) EPA. Carcinogenic Effects of Benzene: An Update. 1998; Available at: http://www.epa.gov/ncea/pdfs/benzenef.pdf. Accessed November 7, 2005, 2005.

(26) Kogevinas M, Ferro G, Andersen A, Bellander T, Biocca M, Coggon D, et al. Cancer mortality in a historical cohort study of workers exposed to styrene. Scandinavian journal of work, environment & health 1994 Aug; 20(4): 251-261.

(27) Wong O, Trent LS, Whorton MD. An updated cohort mortality study of workers exposed to styrene in the reinforced plastics and composites industry. Occupational and environmental medicine 1994 Jun; 51(6): 386-396.

(28) Kolstad HA, Juel K, Olsen J, Lynge E. Exposure to styrene and chronic health effects: mortality and incidence of solid cancers in the Danish reinforced plastics industry. Occupational and environmental medicine 1995 May; 52(5): 320-327.

(29) Abdo KM, Grumbein S, Chou BJ, Herbert R. Toxicity and carcinogenicity study in F344 rats following 2 years of whole-body exposure to naththalene vapors. Inhalation Toxicology 2001; 13:931-950.

(30) Mennear J, Maronpot R, Boorman G, Eustis S, Huff J, Haseman J, et al. Toxicologic and carcinogenic effects of inhaled tetrachloroethylene in rats and mice. Developments in toxicology and environmental science 1986; 12:201-210.

(31) Aschengrau A, Ozonoff D, Paulu C, Coogan P, Vezina R, Heeren T, et al. Cancer risk and tetrachloroethylene-contaminated drinking water in Massachusetts. Archives of Environmental Health 1993 Sep-Oct;48(5):284-292.

(32) McGeehin MA, Reif JS, Becher JC, Mangione EJ. Case-control study of bladder cancer and water disinfection methods in Colorado. American Journal of Epidemiology 1993 Oct 1; 138(7): 492-501.

(33) Paulu C, Aschengrau A, Ozonoff D. Tetrachloroethylene-contaminated drinking water in Massachusetts and the risk of colon-rectum, lung, and other cancers. Environmental health perspectives 1999 Apr; 107(4): 265-271.

(34) United Staties Environment Protection Agency (EPA). Cancer risk from outdoor exposure to air toxics. 1990; EPA-450/1-90-004a.

(35) Health Canada. National Ambient Air Quality Objectives for Ground-Level Ozone Summary Science Assessment Document. A report by the Federal-Provincial Working Group on Air Quality Objectives and Guidelines. 1999 July 1999:1-64.

(36) U.S. Environmental Protection Agency (EPA). Good up high bad nearby. 2003; Available at: www.epa.gov/oar/oaqps/gooduphigh/ozone.html#4. Accessed October 12, 2005, 2005.

(37) U.S. Environmental Protection Agency (EPA). Regional Approach to improving Air Quality---Particulate Matter. 2001; Available at: www.epa.gov/oar/oapqs/airtrans/pm.html. Accessed October 12, 2005, 2005.

(38) Brown-Woodman PD, Webster WS, Picker K, Ritchie HE. Embryotoxicity of xylene and toluene: an in vitro study. Industrial health 1991;29(4):139-152.

(39) Brown NA, Lamb JC, Brown SM, Neal BH. A Review of the Developmental and Reproductive Toxicity of Styrene. Regulatory Toxicology and Pharmacology 2000; 32(3): 228-247.

(40) Dalgaard M, Hossaini A, Hougaard KS, Hass U, Ladefoged O. Developmental toxicity of toluene in male rats: effects on semen quality, testis morphology, and apoptotic neurodegeneration. Archives of Toxicology 2001 Apr; 75(2): 103-109.

(41) Hudak A, Ungvary G. Embryotoxic effects of benzene and its methyl derivatives: toluene, xylene. Toxicology 1978 Sep; 11(1): 55-63.

(42) Marks TA, Ledoux TA, Moore JA. Teratogenicity of a commercial xylene mixture in the mouse. Journal of toxicology and environmental health 1982 Jan; 9(1): 97-105.

(43) Ono A, Sekita K, Ogawa Y, Hirose A, Suzuki S, Saito M, et al. Reproductive and developmental toxicity studies of toluene. II. Effects of inhalation exposure on fertility in rats. Journal of environmental pathology, toxicology and oncology: official organ of the International Society for Environmental Toxicology and Cancer 1996; 15(1): 9-20.

(44) Ono A, Sekita K, Ohno K, Hirose A, Ogawa Y, Saito M, et al. Reproductive and developmental toxicity studies of toluene. I. Teratogenicity study of inhalation exposure in pregnant rats. The Journal of toxicological sciences 1995 May; 20(2): 109-134.

(45) Roberts LG, Bevans AC, Schreiner CA. Developmental and reproductive toxicity evaluation of toluene vapor in the rat. Reproductive Toxicology 2003; 17(6): 649-658.

(46) Rosen MB, Crofton KM, Chernoff N. Postnatal evaluation of prenatal exposure to p-xylene in the rat. Toxicology letters 1986 Dec; 34(2-3): 223-229.

(47) Saillenfait AM, Gallissot F, Morel G, Bonnet P. Developmental toxicities of ethylbenzene, ortho-, meta-, para-xylene and technical xylene in rats following inhalation exposure. Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association 2003 Mar; 41(3): 415-429.

(48) Shigeta S, Aikawa H, Misawa T. Effects of maternal exposure to toluene during pregnancy on mouse embryos and fetuses. The Tokai journal of experimental and clinical medicine 1982 Mar; 7(2): 265-270.

(49) Ungvary G, Tatrai E, Hudak A, Barcza G, Lorincz M. Studies on the embryotoxic effects of ortho-, meta- and para-xylene. Toxicology 1980; 18:61-74.

(50) Ungvary G, Tatrai E. On the embryotoxic effects of benzene and its alkyl derivatives in mice, rats and rabbits. Archives of Toxicology. Supplement. 1985; 8:425-430.

(51) Taskinen H, Kyyronen P, Hemminki K, Hoikkala M, Lajunen K, Lindbohm ML. Laboratory work and pregnancy outcome. Journal of occupational medicine. : official publication of the Industrial Medical Association 1994 Mar;36(3):311-319.

(52) Thiel R, Chahoud I. Postnatal development and behaviour of Wistar rats after prenatal toluene exposure. Archives of Toxicology 1997; 71(4): 258-265.

(53) Arnold GL, Kirby RS, Langendoerfer S, Wilkins-Haug L. Toluene embryopathy: clinical delineation and developmental follow-up. Pediatrics 1994 Feb; 93(2): 216-220.

(54) Pearson MA, Hoyme HE, Seaver LH, Rimsza ME. Toluene embryopathy: delineation of the phenotype and comparison with fetal alcohol syndrome. Pediatrics 1994 Feb; 93(2): 211-215.

(55) Environment Canada. Total Biogenic VOC Emissions for 1995. Available at: www.ec.gc.ca/pdb/ape/ape_tables/bioVOC95_e.cfm. Accessed October 12,2005, 2005.

(56) Environment Canada. 1995 Criteria Air Contaminant Emissions for Canada. Available at: www.ec.gc.ca/pdb/ape/ape_tables/canada95_e.cfm. Accessed October 12, 2005, 2005.

(57) Solberg S, Schmidbauer N, Pedersen U, Schaug J. VOC measurements August 1992-June 1993. 1993;EMEP/CCC.

(58) Jobson BT, Wu Z, Niki H. Seasonal trends of isoprene, c_2-c_5 alkanes, and acetylene at a remote boreal site in Canada. Journal of Geophysical Research 1994; 99:1589-1599.

(59) Na K, Kim Y. Seasonal Characteristics of Ambient Volatile Organic Compounds in Seoul, Korea. Atmospheric Environment 2001; 35(15): 2603-2614.

(60) Cohen MA, Ryan PB, Spengler JD, Ozkaynak H. Source-Receptor Study of Volatile Organic Compounds and Particulate Matter in the Kanawha Valley, WV-I. Methods and Descriptive Statistics. Atmospheric Environment 1991; 25B: 79-93.

(61) Gilli G, Scursation E, Bono R. Benzene, toluene and xylenes in air, geographical distribution in the Piedmont region (Italy) and personal exposure. The Science of the Total Environment 1994; 148:49-56.

(62) Upmanis H, Eliasson I, Andersson-skold Y. Case Studies of the Spatial Variation of Benzene and and Toluene Concentrations in Parks and Adjacent Built-up Areas. Water, Air, and Soil Pollution 2001; 129:61-81.

(63) Hewitt CN, Cao X, Boissard C, Duckhum SC. Atmospheric VOCs from Natural Sources. In: Hester RE, Harrison RM, editors. Issues in Environmental Science and Technology (4).Volatile Organic Compounds in the Atmosphere. 1st ed. Thomas Graham House, Science Park, Milton Road, Cambridge CB4 4WF, UK: The Royal Society of Chemistry; 1995. p. 17-36.

(64) Bertin N, Staudt M, Hansen U, Seufert G, Ciccioli P, Foster P, et al. Diurnal and seasonal course of monoterpene emissions from *Quercus ilex* (L.) under natural conditions---applications of light and temperature algorithms. Atmospheric Environment 1997; 31:135-144.

(65) Tingey DT, Manning M, Grothaus LC, Burns WF. Influence of light and temperature on mototerpene emission rate from slash pine. Plant Physiology 1980; 65:797-801.

(66) Owen SM, Harley P, Guenther A, Hewitt CN. Light dependency of VOC emissions from selected Mediterranean plant species. Atmospheric Environment 2002; 36(19): 3147-3159.

(67) Schuh G, Heiden AC, Hoffmann T, Kahl J, Rockel P, Rudolph J, et al. Emissions of volatile organic compounds from sunflower and beech: dependence on temperature and light intensity. Journal of Atmospheric Chemistry 1997; 27:291-318.

(68) Guenther AB, Hewitt CN, Erickson D, mFall R, Geron C, Graedel T, et al. A global model of natural volatile organic compound emissions. Journal of Geophysical Research 1995; 100:8873-8892.

(69) Pratt GC, Bock D, Stock TH, Morandi M, Adgate JL, Ramachandran G, et al. A Field Comparison of Volatile Organic Compound Measurements Using Passive Organic Vapor Monitors and Stainless Steel Canisters. Environmental Science & Technology 2005; 39:3261-3268.

(70) Cohen MA, Ryan PB, Yanagisawa Y, Spengler JD, Ozkaynak H, Epstein PS. The Validation of Passive Sampler for Indoor and Outdoor Concentration of Volatile Organic Compounds. Journal of Air & Waste Management Association 1990; 40:993-997.

(71) Morandi MT, Stock TH. Personal Exposure to Toxic Air Pollutants
Feasibility Study: Results of the Laboratory Evaluation, Field Pilot Exposure Study, Modification of the 3500 OVM Passive Dosimeter. 1997; 3-43.

(72) Mukerjee S, Smith LA, Norris G.A., Morandi MT, Gonzales M, et al. Fiels Method Comparison between Passive Air Samplers and Continuous Monitors for VOCs and NO2 in EI Paso, Texas. Journal of the Air & Waste Management Association 2004; 54:307-319.

(73) Chung CW, Morandi MT, Stock TH, Afshar M. Evaluation of a Passive Sampler for Volatile Organic Compounds at ppb Concentrations, Varying Temperatures, and Humidities with 24-hr exposure. Enviroment Science & Technology 1999; 33:3666-3671.

(74) The Australia National Environment Protection Council. National Environment Protection (Air Toxics) Measure Explanatory document. 2004; Available at: http://www.ephc.gov.au/nepms/air/air_toxics.html. Accessed October 19, 2005, 2005.

(75) Burstyn I, Boffetta P, Burr GA, Cenni A, Knecht U, Sciarra G, et al. Validity of empirical models of exposure in asphalt paving. Occupational and Environmental Medicine 2002; 59:620-624.

(76) Burstyn I, Kromhout H, Kauppinen T, et al. Statistical modeling of the determinants of historical exposure to bitumen and polycyclic aromatic hydrocarbons among paving workers. The Annals of Occupational Hygiene 2000; 44:43-56.

(77) Burstyn I, Ferrari P, Wegh H, Heederik D, Kromhout H. Characterizing worker exposure to bitumen during hot mix paving and asphalt mixing operations. AIHA J.(Fairfax, Va) 2002 May-Jun;63(3):293-299.

(78) Derwent RG, Middleton DR, Field RA, Goldstone ME, Lester JN, Perry R. Analysis and Interpretation of Air Quality Data from An Urban Roadside Location in Central London over the Period from July 1991 to July 1992. Atmospheric Enviroment 1995; 29:923-946.

(79) Meijster T, Burstyn I, Van Wendel de Joode, B., Posthumus M, Kromhout H. Evaluating exposures to complex mixtures of chemicals during a new production process in the plastics industry. The Annals of Occupational Hygiene 2004; 48:499-507.

(80) Millet DB, Donahue NM, Pandiis SN, Polidori A, Stanier CO, Turpin BJ, et al. Atmospheric Volatile Organic Compound Measurements during the Pittsburgh Air Quality Study: Results, interpretation, and quantification of primary and secondary contributions. Journal of Geophysical Research 2005; 110: D07S07.

(81) Burstyn I. Principal Component Analysis is a Powerful Instrument in Occupational Hygiene Inquiries. The Annals of Occupational Hygiene 2004; 48:655-661.

(82) Strosher M. Investigations of flare gas emissions in Alberta
Final report to Environment Canada, Conservation and Protection, Alberta Energy and Utilities Board, and Canadian Association of Petroleum Producers. 1996.

(83) Kindzierski WB. Importance of human environmental exposure to hazardous air pollutants from gas flares. Env. Rev. 2000; 8:41-62.

(84) Kelly TJ, Mukund R, Spicer CW, Pollack AJ. Concentrations and transformations of hazardous air pollutants: What we know and don't know about the CAAA's 189 hazardous air pollutants. Environmental Science & Technology 1994; 28:378A-387A.

(85) Shaw JJ, Joseph DW. National ambient VOC data base update
Prepared for US Environmental Protection Agency by G2 Environmental, Inc., Washington, D.C. 1993; 3.0. EPA-600/R-94-089.

(86) Spicer CW, Buxton BE, Holdren MW, Smith DL, Kelly TJ, Rust SW, et al. Variability of hazardous air pollutants in an urban area. Atmospheric Environment 1996; 30:3443-3456.

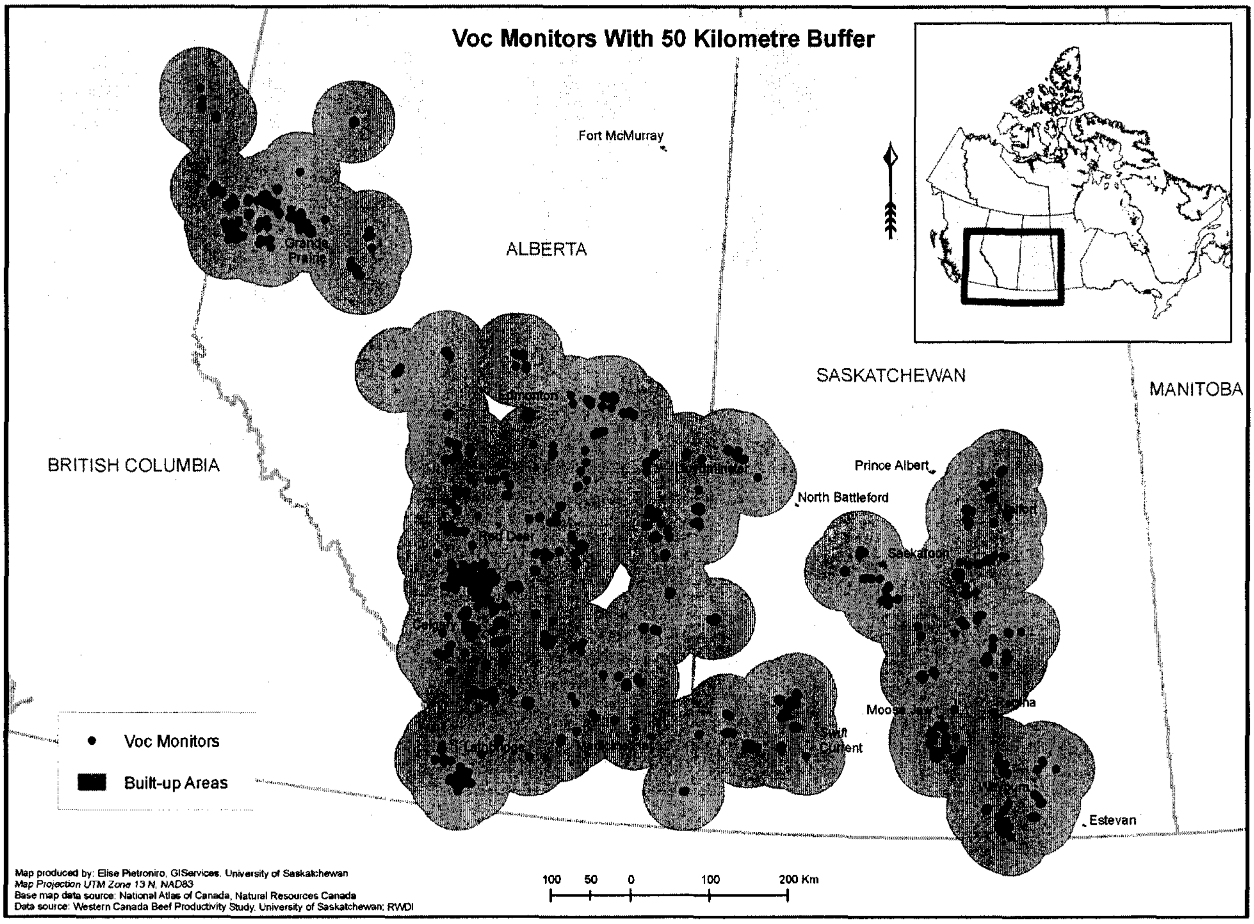
(87) Papadopoulos A, Vassilakos C, Hatzianestis J, Maggos T, Bartzis JG. Volatile organic compounds concentration levels on a day without traffic in the centre of Athens. Fresenius environmental Bulletin 2005; 14(6): 498-502.

(88) Esteves-Booth A, Muneer T, Kubie J, Kirby H. A review of vehicular emission models and driving cycles. The Journal of Mechanical Engineering Science 2002(216): 777-797.

(89) Penkett SA, Blake NJ, Lightman P, Marsh ARW, Anwyl P, Butcher G. The seasonal variation of nonmethane hydrocarbons in the free troposphere over the North Atlantic Ocean: possible evidence for extensive reactions of hydrocarbons with the nitrate radical. Journal of Geophysical Research 1993; 48:2865-2885.

(90) ATSDR. Chloroform: Potential for Human Exposure. 1998; Available at: http://www.atsdr.cdc.gov/toxprofiles/tp6-c5.pdf. Accessed October 11, 2005, 2005.

Appendix 1:Map of spatial distribution of VOC monitoring sites



Voc Monitors With 50 Kilometre Buffer

Map produced by: Elise Pietroniro, GIServices, University of Saskatchewan
Map Projection UTM Zone 13 N, NAD83
Base map data source: National Atlas of Canada, Natural Resources Canada
Data source: Western Canada Beef Productivity Study, University of Saskatchewan; RWDI

Appendix 2: Correlation between distance weights*

	lw4wloil	lw7wloil	lw4wlgas	lw7wlgas	lw7wlbit	lw4tpgpl
lw4wloil	1.00000	0.48849	0.14238	0.08340	0.01176	0.04866
lw7wloil	0.48849	1.00000	0.22849	0.49792	0.04081	0.04553
lw4wlgas	0.14238	0.22849	1.00000	0.66367	0.04295	0.04070
lw7wlgas	0.08340	0.49792	0.66367	1.00000	0.04219	0.04392
lw7wlbit	0.01176	0.04081	0.04295	0.04219	1.00000	-0.02112
lw4tpgpl	0.04866	0.04553	0.04070	0.04392	-0.02112	1.00000
lw7tpgpl	0.07570	0.56732	0.16455	0.47398	-0.12466	0.05802
lw4tpall	0.12047	0.16311	0.16797	0.14101	-0.03062	0.77878
lw7tpall	0.15248	0.68460	0.25527	0.60145	-0.00759	0.04280
lw4btall	0.45541	0.42797	0.18737	0.15833	0.01071	0.08900
lw7btall	0.31573	0.82245	0.26685	0.55173	0.03541	0.05325

	lw7tpgpl	lw4tpall	lw7tpall	lw4btall	lw7btall
lw4wloil	0.07570	0.12047	0.15248	0.45541	0.31573
lw7wloil	0.56732	0.16311	0.68460	0.42797	0.82245
lw4wlgas	0.16455	0.16797	0.25527	0.18737	0.26685
lw7wlgas	0.47398	0.14101	0.60145	0.15833	0.55173
lw7wlbit	-0.12466	-0.03062	-0.00759	0.01071	0.03541
lw4tpgpl	0.05802	0.77878	0.04280	0.08900	0.05325
lw7tpgpl	1.00000	0.13877	0.92397	0.28190	0.78837
lw4tpall	0.13877	1.00000	0.16101	0.24274	0.18723
lw7tpall	0.92397	0.16101	1.00000	0.34008	0.88720
lw4btall	0.28190	0.24274	0.34008	1.00000	0.46599
lw7btall	0.78837	0.18723	0.88720	0.46599	1.00000

* Pearson Correlation Coefficient, n=11,381

** lw4wloil represents log(distance weights) of oil well <2km; lw7wloil represents log(distance weights) of oil well 2-50 km; lw4wlgas represents log(distance weights) of gas well <2 km; lw7wlgas represents log(distance weights) of gas well 2-50 km; lw7wlbit represents log(distance weights) of bitumen well 2-50 km; lw4tpgpl represents log(distance weights) of gas plant <2km; lw7tpgpl represents log(distance weights) of 2-50 km; lw4tpall represents log(distance weights) of all large facilities (incl. Gas plants) < 2km; lw7tpall represents log(distance weights) of all large facilities (incl. Gas plants) 2-50km; lw4btall represents log(distance weights) of batteries <2 km; lw7btall represents log(distance weights) of batteries 2-50 km.

Appendix 3: Correlations between potential industrial sources and their corresponding distances*

	well_i	well_d	plant_i	plant_d	flare_i	flare_d	battery_i
well_i	1.00000	0.98907	0.01141	0.00462	0.26675	0.26520	0.29628
well_d	0.98907	1.00000	0.01417	0.00781	0.27788	0.27780	0.29529
plant_i	0.01141	0.01417	1.00000	0.99831	0.12001	0.11881	0.05671
plant_d	0.00462	0.00781	0.99831	1.00000	0.12029	0.11938	0.05145
flare_i	0.26675	0.27788	0.12001	0.12029	1.00000	0.99888	0.13359
flare_d	0.26520	0.27780	0.11881	0.11938	0.99888	1.00000	0.13136
battery_i	0.29628	0.29529	0.05671	0.05145	0.13359	0.13136	1.00000
battery_d	0.29669	0.29717	0.05427	0.04909	0.13235	0.13037	0.99938
highway_i	0.08979	0.08316	0.00205	0.00054	-0.02578	-0.02729	-0.05863
highway_d	0.08836	0.08213	0.00102	-0.00047	-0.02576	-0.02727	-0.05807
comp_i	0.16332	0.16684	0.09980	0.09587	0.33177	0.32772	0.00968
comp_d	0.16212	0.16564	0.09934	0.09544	0.33109	0.32714	0.00990
others_i	0.14963	0.14157	0.02340	0.02320	0.11671	0.11671	-0.06798
others_d	0.14716	0.13936	0.02371	0.02362	0.11690	0.11711	-0.06727
otherin_i	-0.00518	-0.00635	0.09769	0.09147	-0.03939	-0.03972	-0.05168
otherin_d	-0.00590	-0.00700	0.09619	0.09004	-0.03957	-0.03990	-0.05168
road_i	-0.09262	-0.08752	-0.04014	-0.04334	-0.10446	-0.10311	-0.11051
road_d	-0.09956	-0.09304	-0.03852	-0.04174	-0.10485	-0.10324	-0.11247

	battery_d	highway_i	highway_d	comp_i	comp_d	others_i
well_i	0.29669	0.08979	0.08836	0.16332	0.16212	0.14963
well_d	0.29717	0.08316	0.08213	0.16684	0.16564	0.14157

plant_i	0.05427	0.00205	0.00102	0.09980	0.09934	0.02340
plant_d	0.04909	0.00054	-0.00047	0.09587	0.09544	0.02320
flare_i	0.13235	-0.02578	-0.02576	0.33177	0.33109	0.11671
flare_d	0.13037	-0.02729	-0.02727	0.32772	0.32714	0.11671
battery_i	0.99938	-0.05863	-0.05807	0.00968	0.00990	-0.06798
battery_d	1.00000	-0.05784	-0.05721	0.00943	0.00970	-0.06679
highway_i	-0.05784	1.00000	0.99966	0.05172	0.05126	0.12824
highway_d	-0.05721	0.99966	1.00000	0.05052	0.05007	0.12730
comp_i	0.00943	0.05172	0.05052	1.00000	0.99990	0.15011
comp_d	0.00970	0.05126	0.05007	0.99990	1.00000	0.14931
others_i	-0.06679	0.12824	0.12730	0.15011	0.14931	1.00000
others_d	-0.06601	0.12975	0.12885	0.14783	0.14703	0.99972
otherin_i	-0.05165	0.12640	0.12466	-0.03797	-0.03797	0.06346
otherin_d	-0.05164	0.12475	0.12303	-0.03797	-0.03796	0.06246
road_i	-0.11054	-0.10879	-0.10890	-0.06319	-0.06334	-0.00204
road_d	-0.11245	-0.11067	-0.11077	-0.06437	-0.06452	-0.00092

	others_d	otherin_i	otherin_d	road_i	road_d
well_i	0.14716	-0.00518	-0.00590	-0.09262	-0.09956
well_d	0.13936	-0.00635	-0.00700	-0.08752	-0.09304
plant_i	0.02371	0.09769	0.09619	-0.04014	-0.03852
plant_d	0.02362	0.09147	0.09004	-0.04334	-0.04174
flare_i	0.11690	-0.03939	-0.03957	-0.10446	-0.10485
flare_d	0.11711	-0.03972	-0.03990	-0.10311	-0.10324
battery_i	-0.06727	-0.05168	-0.05168	-0.11051	-0.11247
battery_d	-0.06601	-0.05165	-0.05164	-0.11054	-0.11245
highway_i	0.12975	0.12640	0.12475	-0.10879	-0.11067
highway_d	0.12885	0.12466	0.12303	-0.10890	-0.11077

comp_i	0.14783	-0.03797	-0.03797	-0.06319	-0.06437
comp_d	0.14703	-0.03797	-0.03796	-0.06334	-0.06452
others_i	0.99972	0.06346	0.06246	-0.00204	-0.00092
others_d	1.00000	0.06233	0.06135	-0.00253	-0.00137
otherin_i	0.06233	1.00000	0.99992	-0.03037	-0.03100
otherin_d	0.06135	0.99992	1.00000	-0.03075	-0.03138
road_i	-0.00253	-0.03037	-0.03075	1.00000	0.98618
road_d	-0.00137	-0.03100	-0.03138	0.98618	1.00000

*Spearman Correlation Coefficients, N=11381

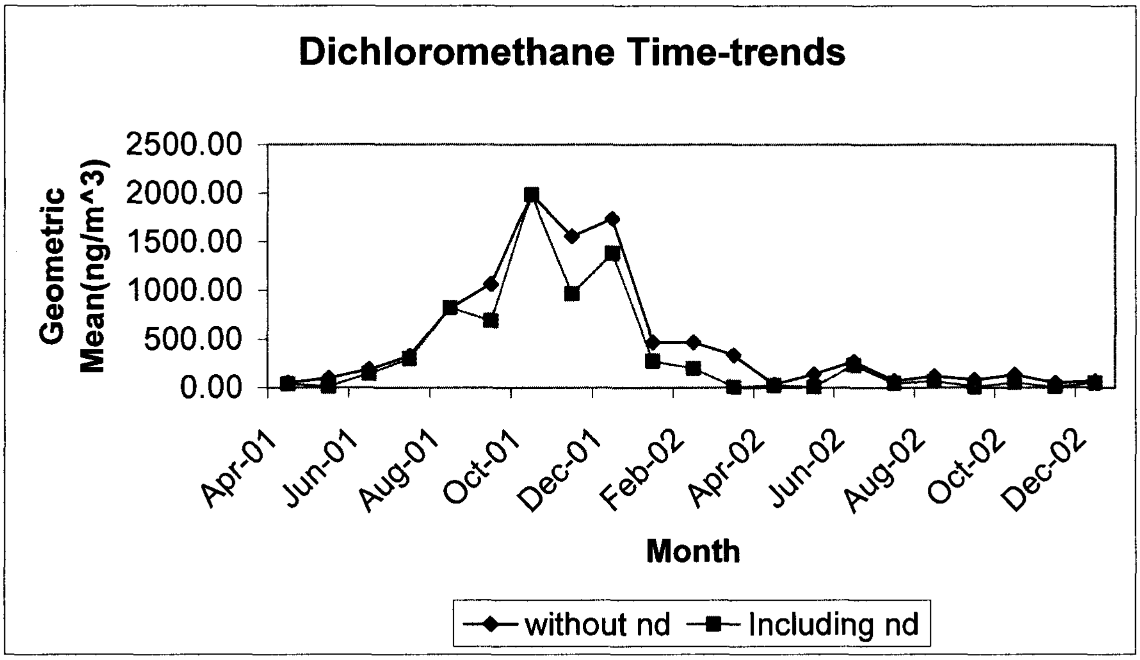
** well_i represents well identification, well_d represents well distance. The same symbols were applied for other nearby sources. Comp, others, and otherin represent compressor, other oil and gas facilities, other industry, respectively.

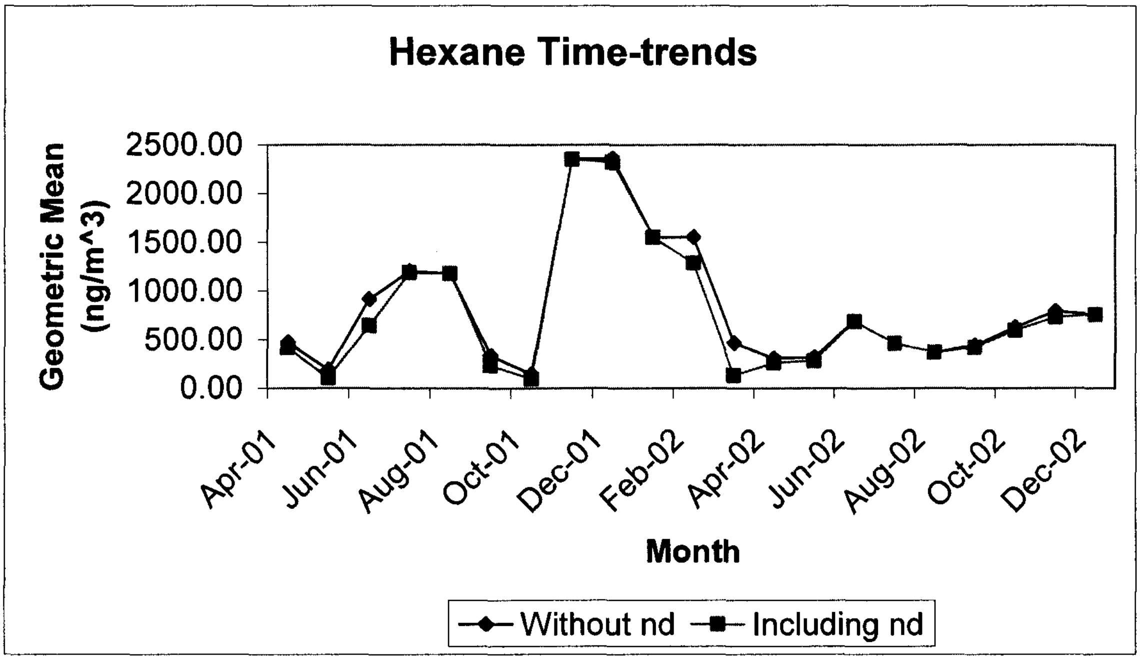
Appendix 4: The Phi correlation coefficients between potential industrial sources and vegetation

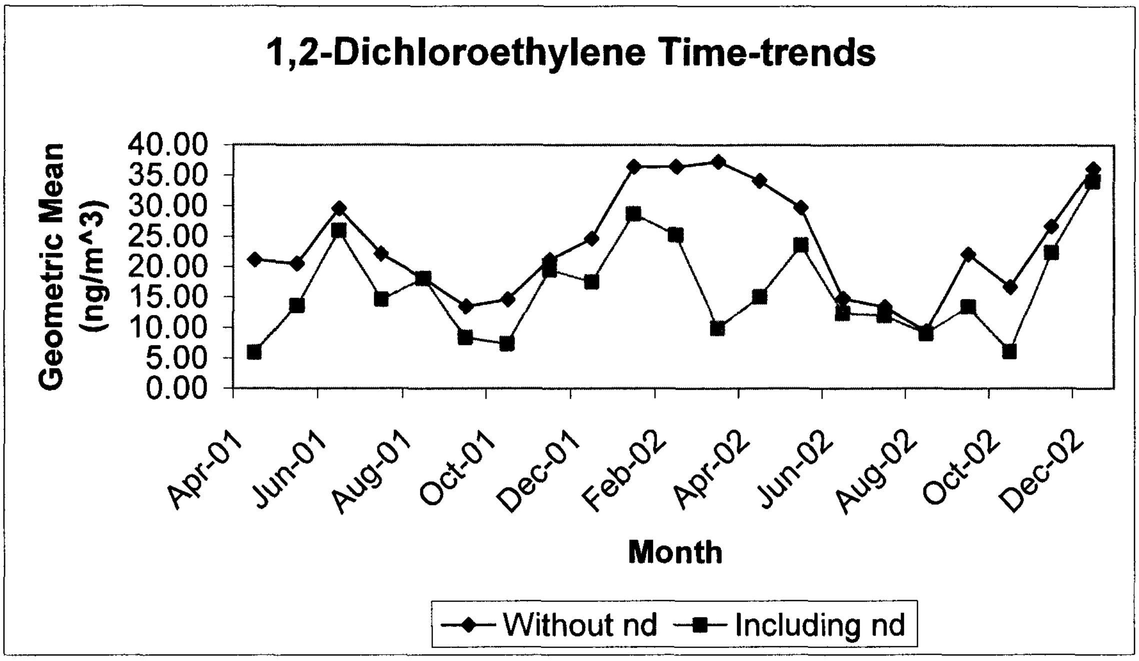
	Well	Plant	Flare	Battery	Highway
Well	—	0.01	0.27	0.30	0.09
Plant	0.01	—	0.12	0.06	0.002
Flare	0.27	0.12	—	0.13	-0.03
Battery	0.30	0.06	0.13	—	-0.06
Highway	0.09	0.002	-0.03	-0.06	—
Others*	0.15	0.02	0.12	-0.07	0.13
Road	-0.09	-0.04	-0.11	-0.11	-0.002
Otherin*	-0.01	0.10	-0.04	-0.05	0.13
Compressor	0.16	0.10	0.33	0.01	0.05
Vegetation	0.002	0.04	0.09	0.05	-0.07
	Others*	Road	Ohterin*	Compressor	Vegetation
Well	0.15	-0.09	-0.01	0.16	0.002
Plant	0.02	-0.04	0.10	0.10	0.04
Flare	0.12	-0.10	-0.04	0.33	0.09
Battery	-0.07	-0.11	-0.05	0.01	0.05
Highway	0.13	-0.11	0.13	0.05	-0.07
Others*	—	-0.002	0.06	0.15	0.01
Road	-0.002	—	-0.03	-0.06	-0.07
Otherin*	0.06	-0.03	—	-0.04	-0.05
Compressor	0.15	-0.06	-0.04	—	0.01
Vegetation	0.01	-0.07	-0.05	0.01	—

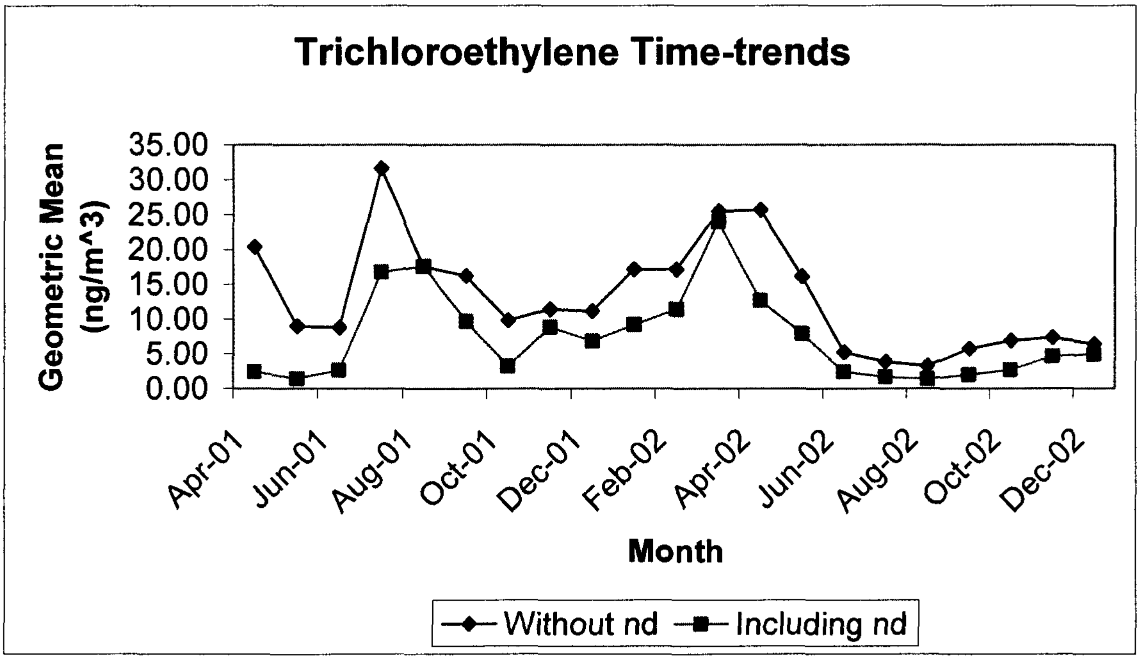
* others and otherin represent other oil and gas facilities and other industry respectively.

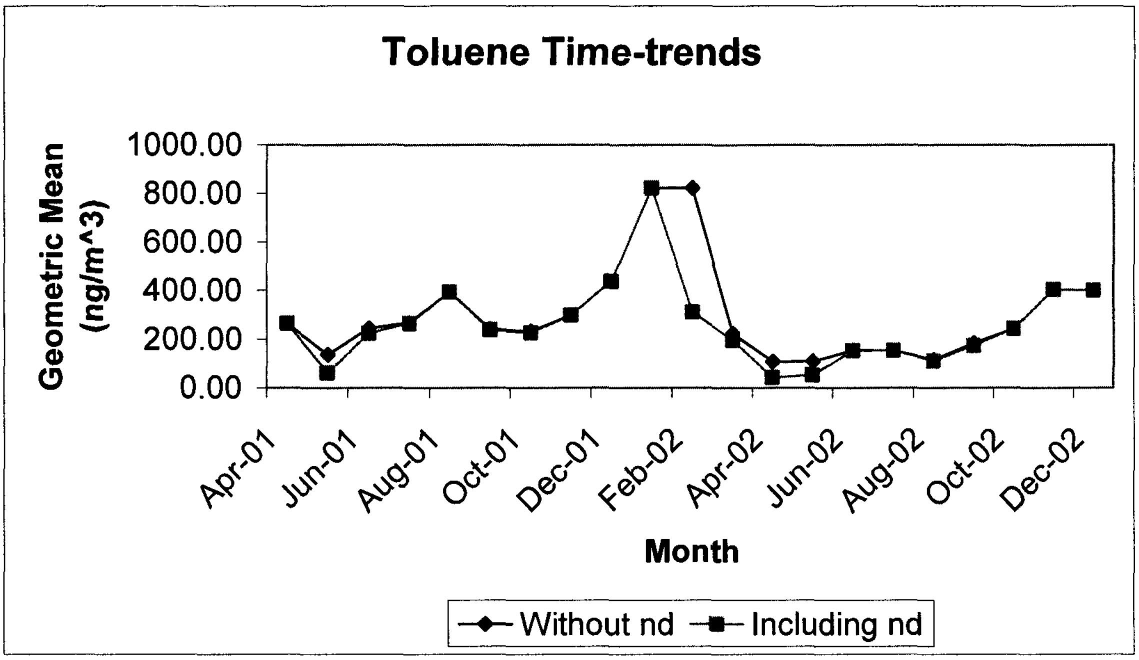
Appendix 5: Time-trends of VOC

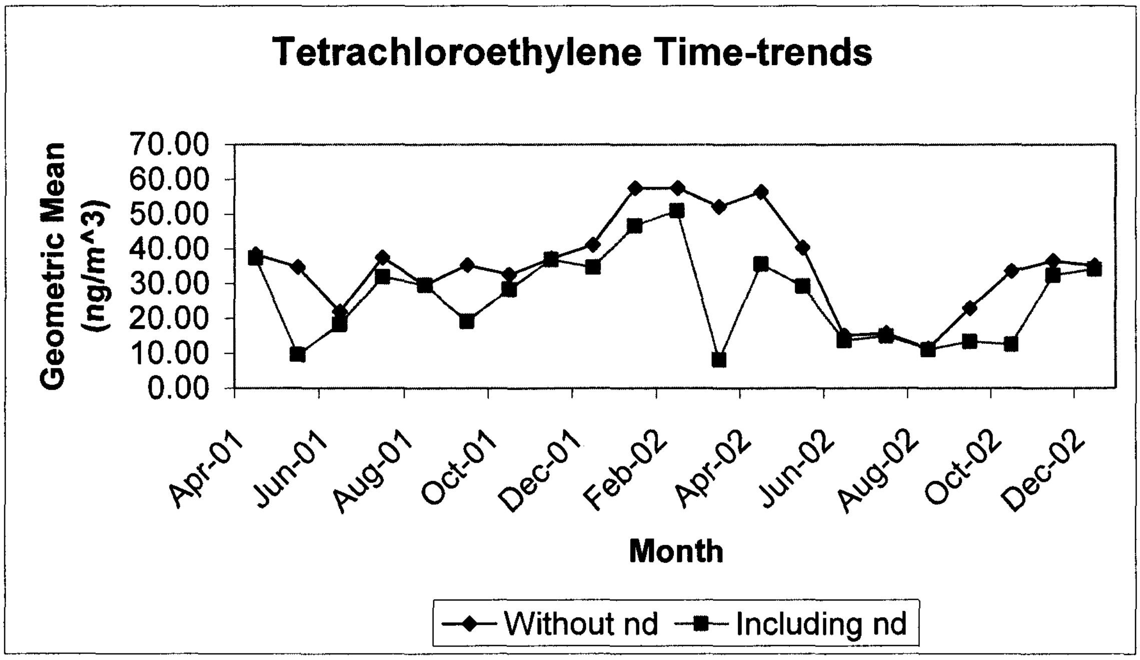


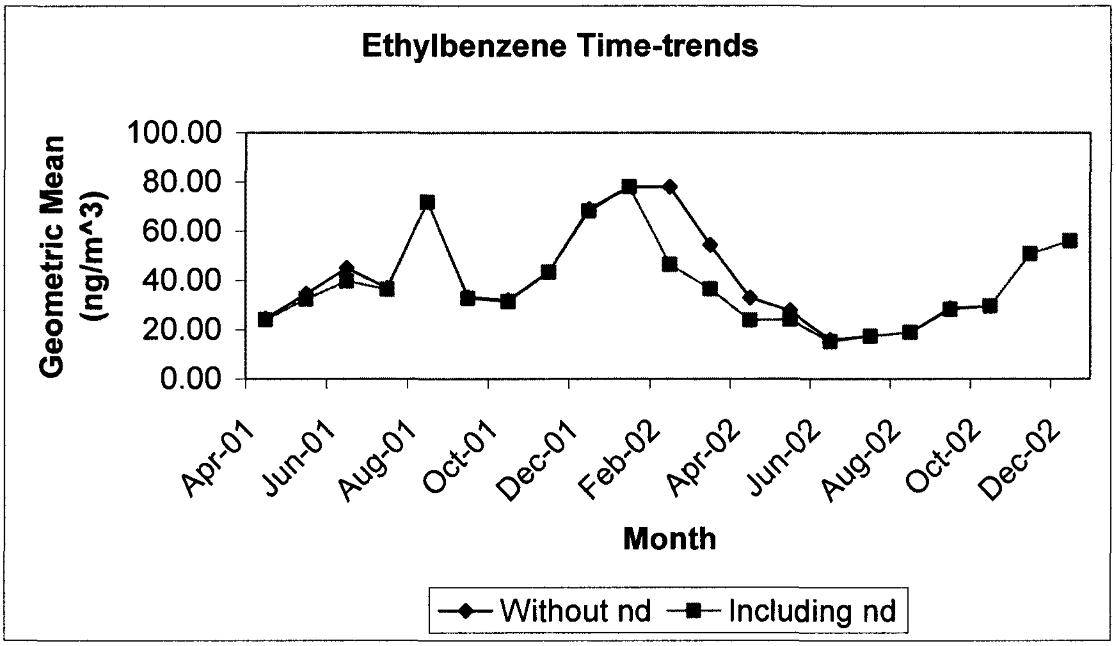


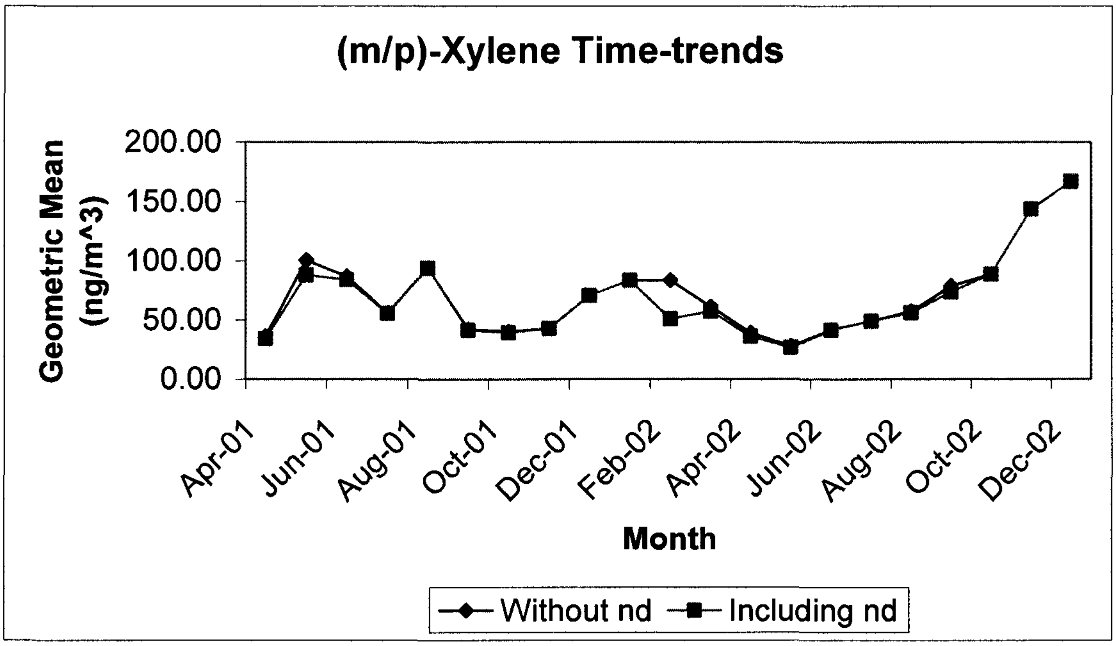

1,2-Dichloroethylene Time-trends
Geometric Mean (ng/m^3)
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

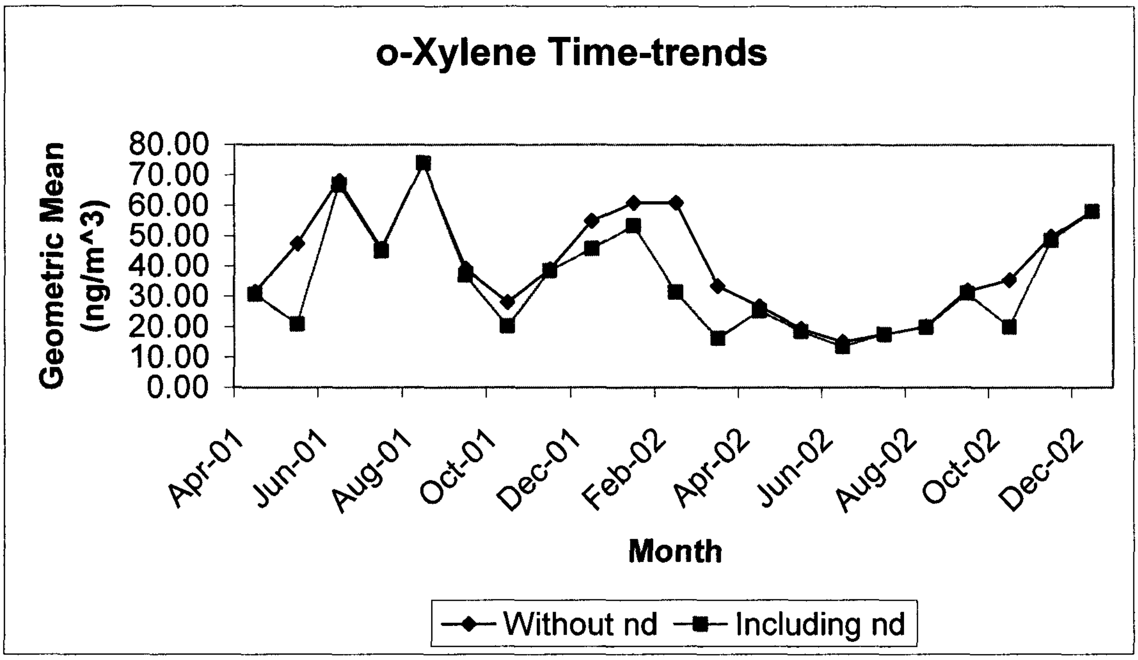

Trichloroethylene Time-trends
Geometric Mean (ng/m^3)
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

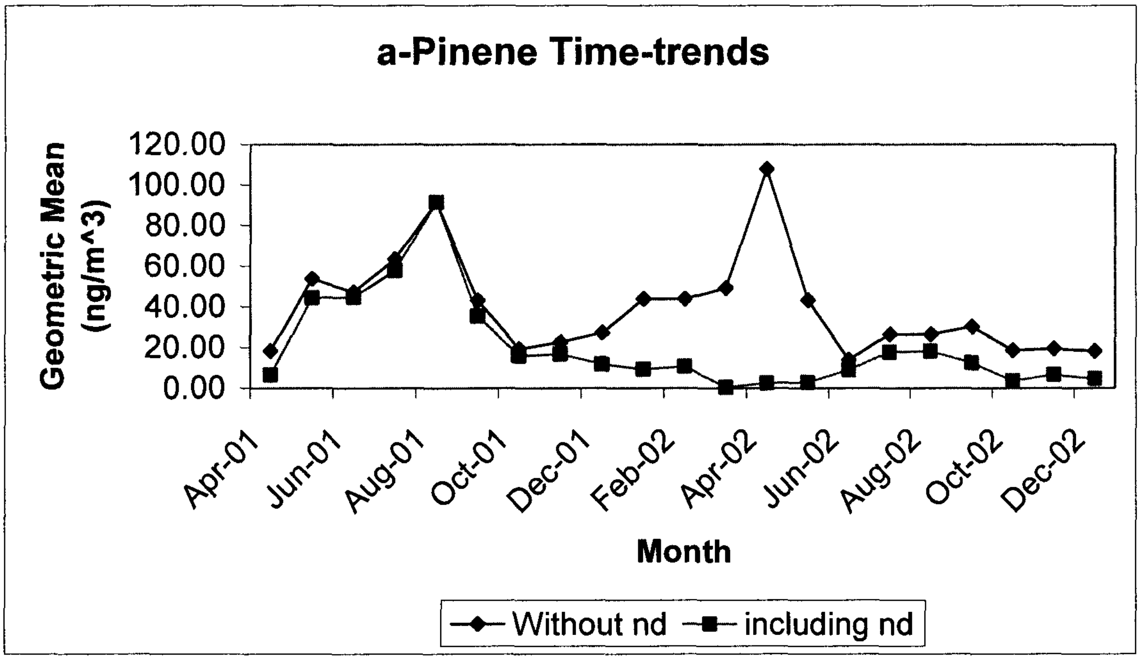

Toluene Time-trends
Geometric Mean (ng/m^3)
1000.00
800.00
600.00
400.00
200.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

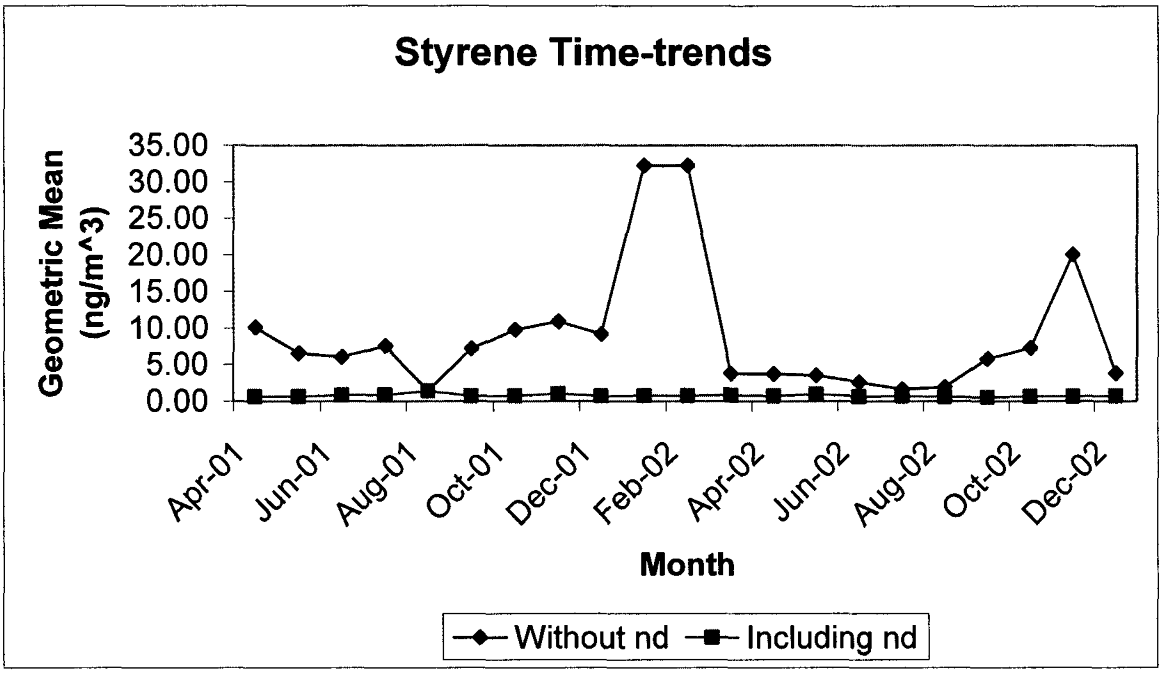

Tetrachloroethylene Time-trends
Geometric Mean (ng/m^3)
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

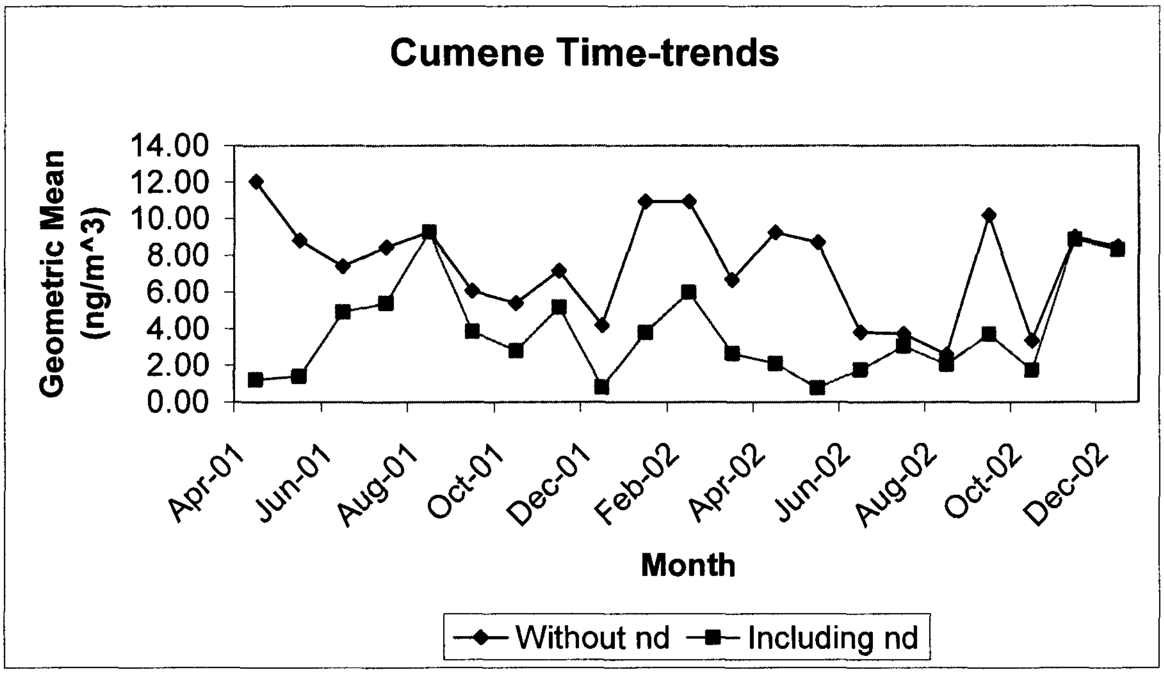

Ethylbenzene Time-trends
Geometric Mean (ng/m^3)
100.00
80.00
60.00
40.00
20.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

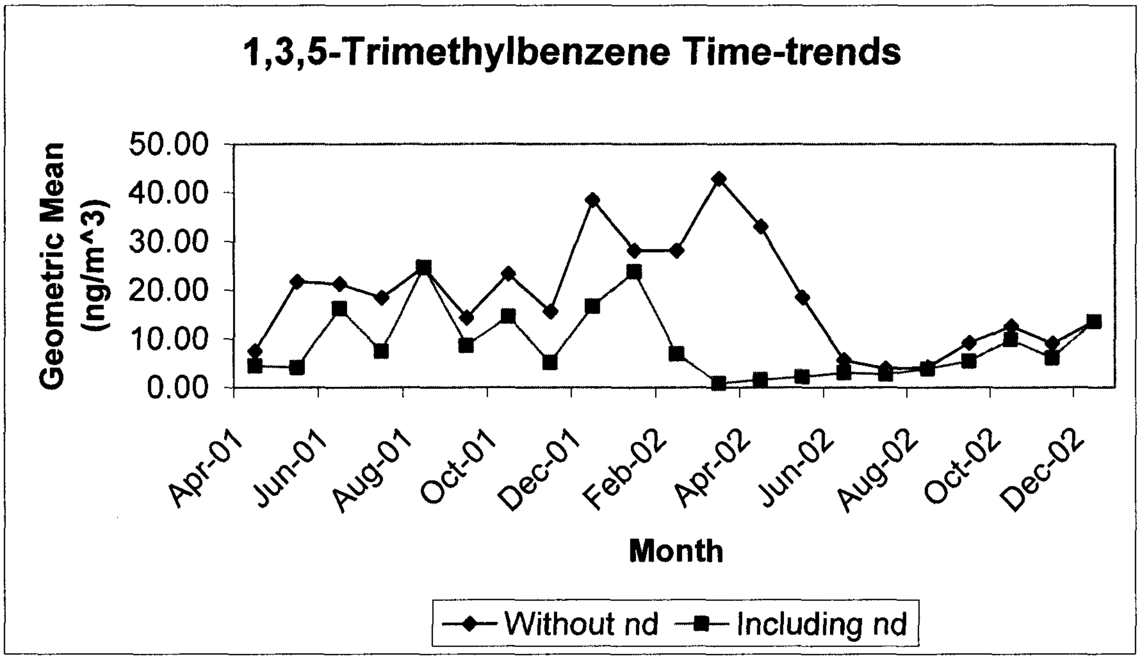

(m/p)-Xylene Time-trends
Geometric Mean (ng/m^3)
200.00
150.00
100.00
50.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

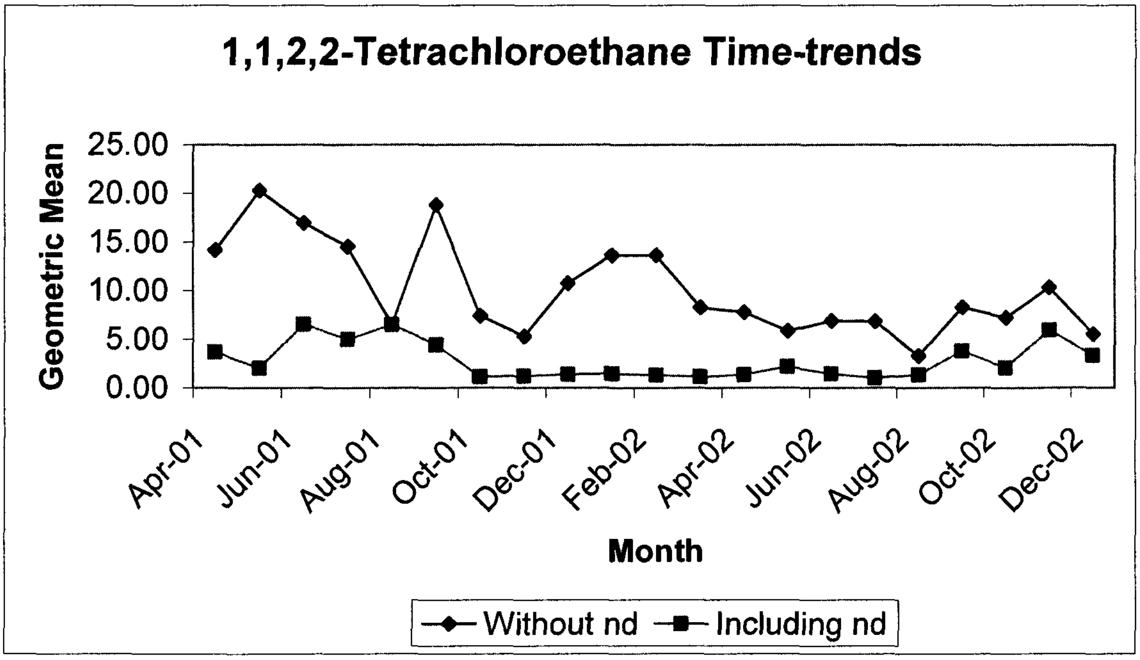

o-Xylene Time-trends
Geometric Mean (ng/m^3)
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

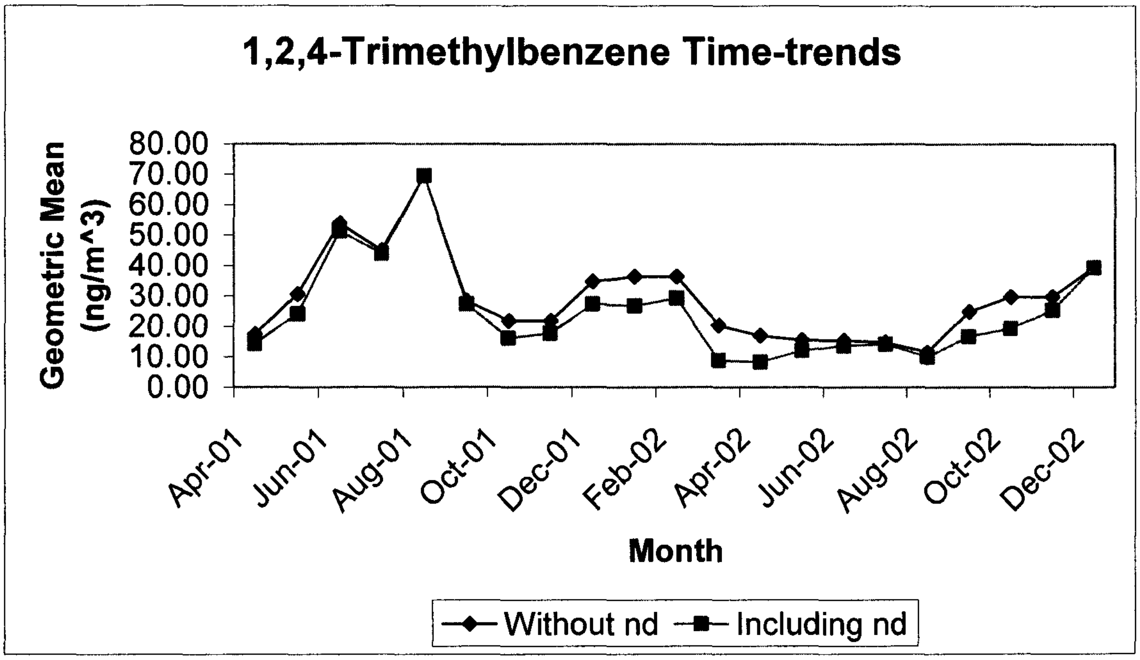

a-Pinene Time-trends
Geometric Mean (ng/m^3)
120.00
100.00
80.00
60.00
40.00
20.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
including nd

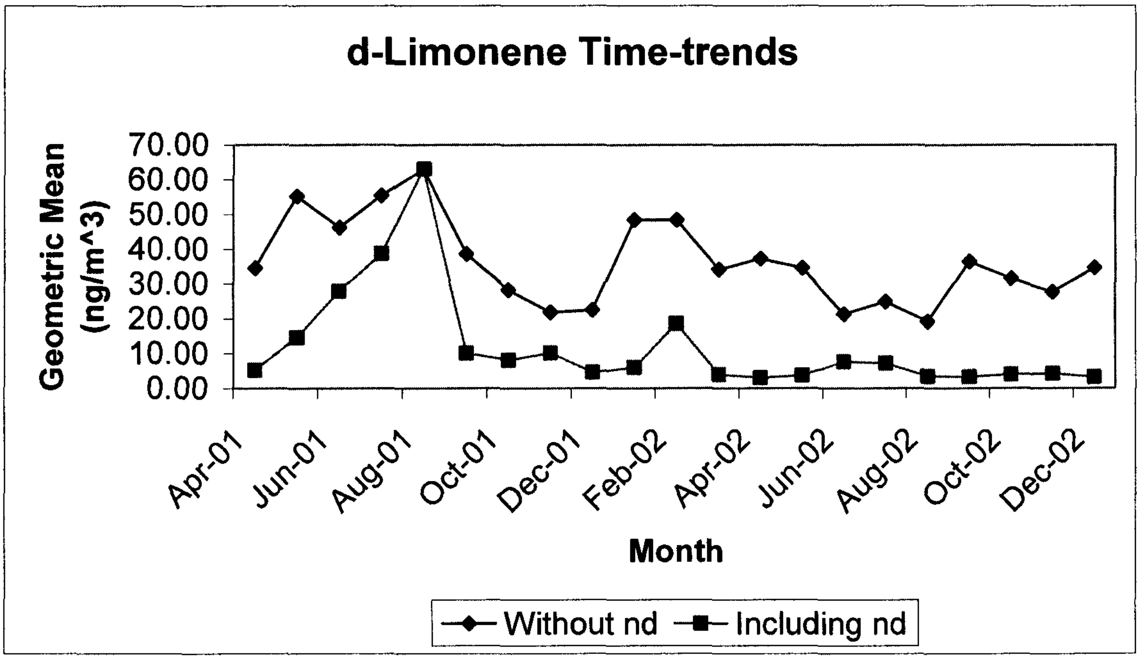

Styrene Time-trends
Geometric Mean (ng/m^3)
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

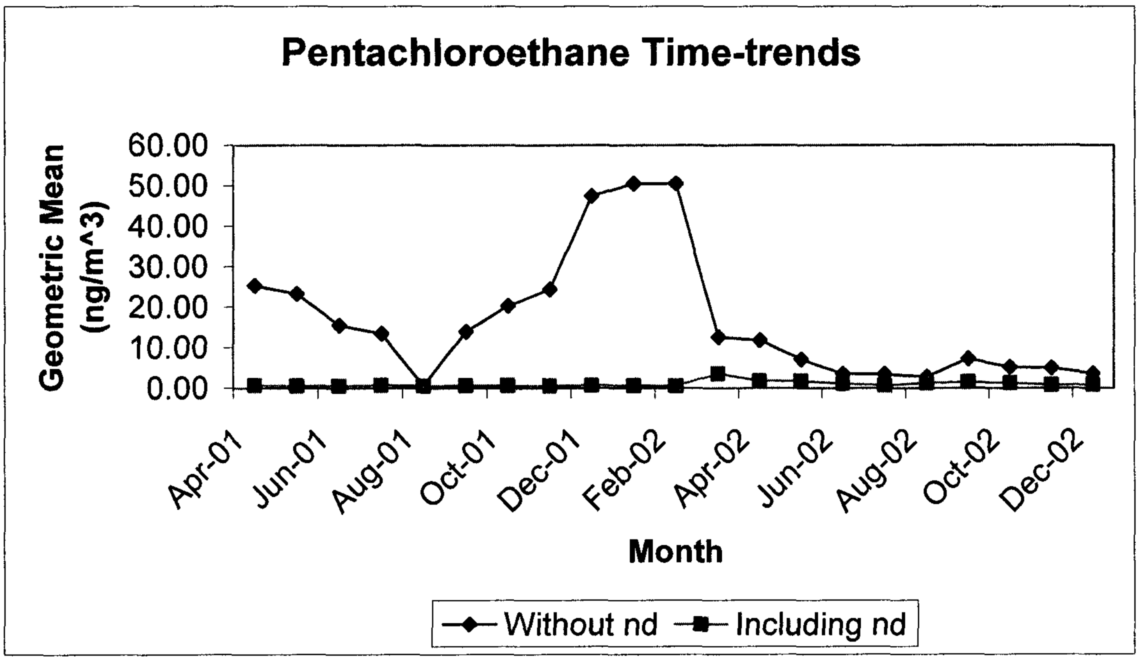

Cumene Time-trends
Geometric Mean (ng/m^3)
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

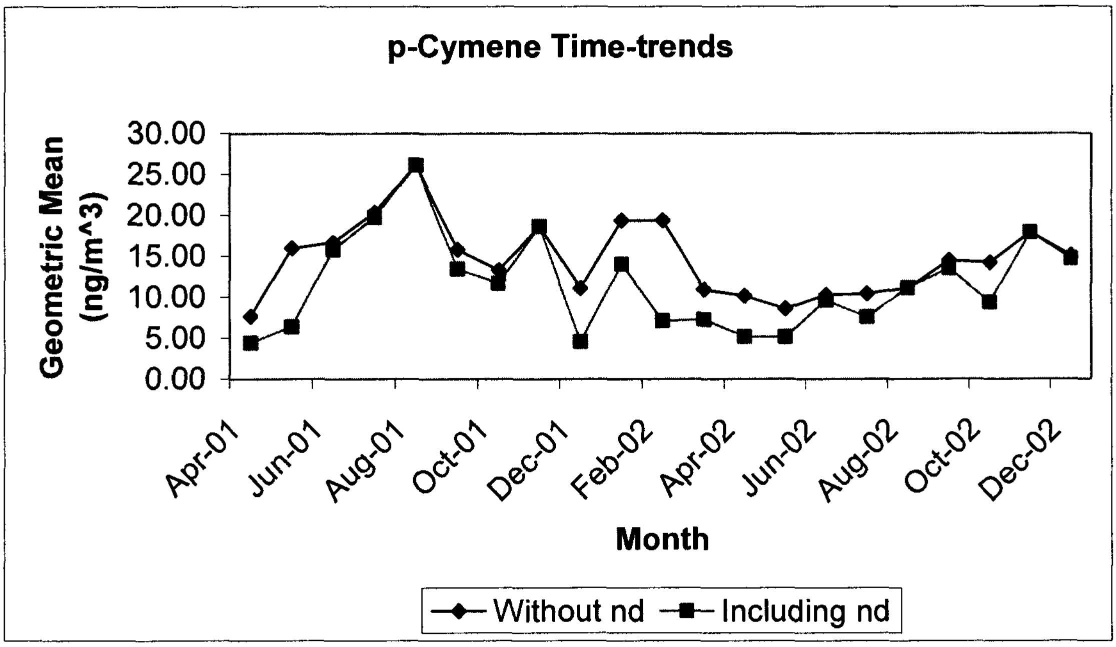

1,3,5-Trimethylbenzene Time-trends
Geometric Mean (ng/m^3)
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

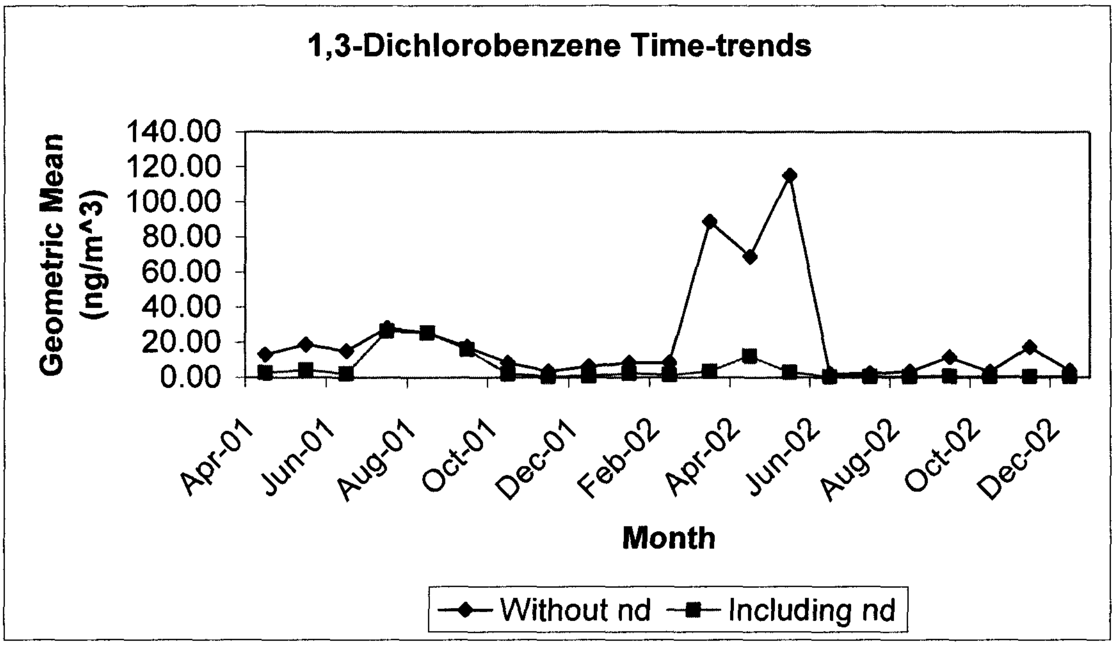

1,1,2,2-Tetrachloroethane Time-trends
Geometric Mean
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

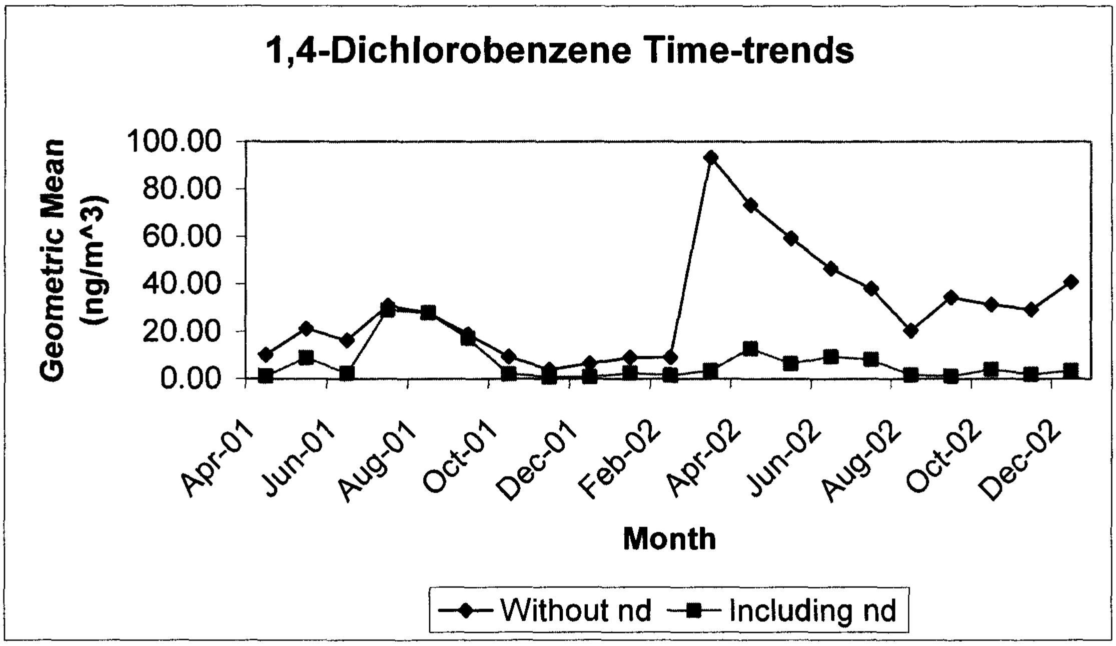

1,2,4-Trimethylbenzene Time-trends
Geometric Mean (ng/m^3)
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

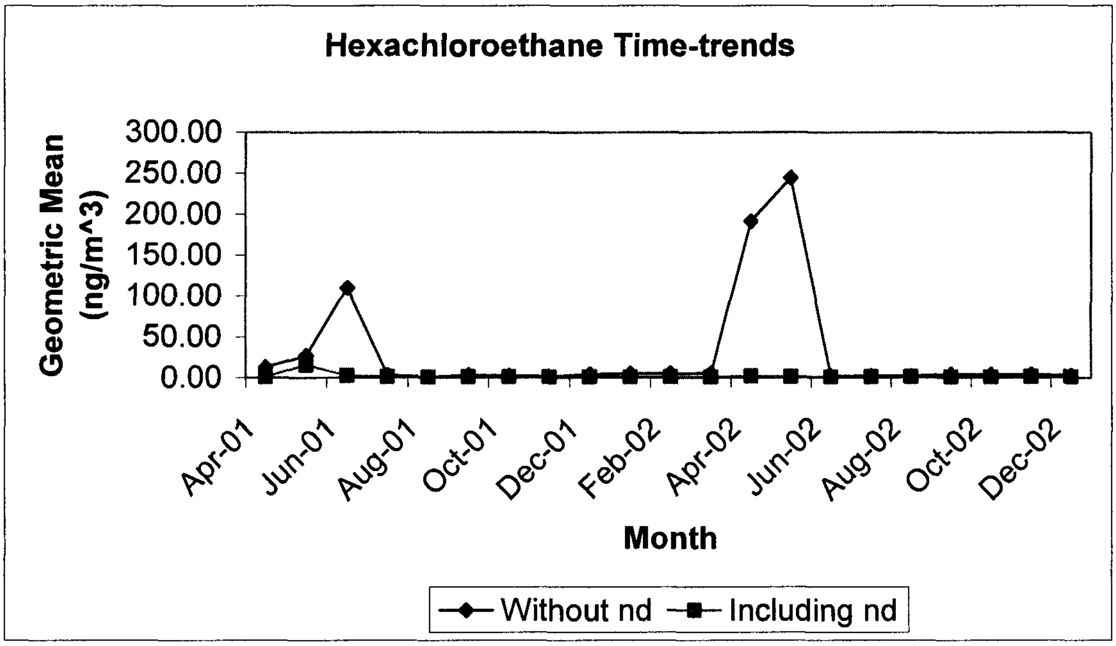

d-Limonene Time-trends
Geometric Mean (ng/m^3)
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

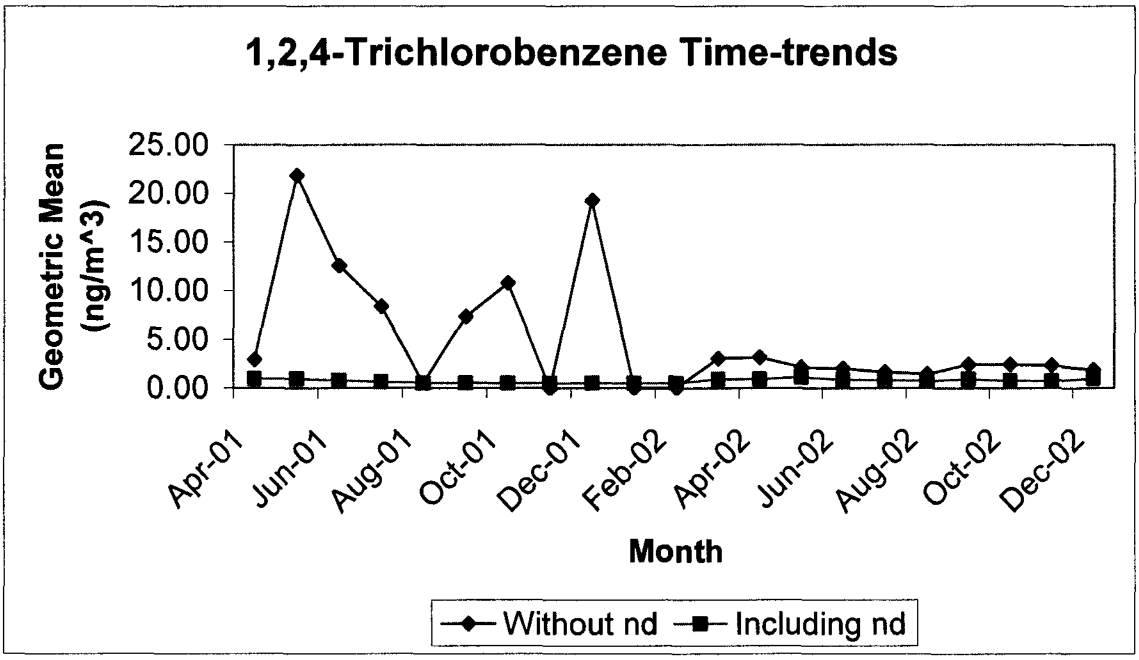

Pentachloroethane Time-trends
Geometric Mean (ng/m^3)
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

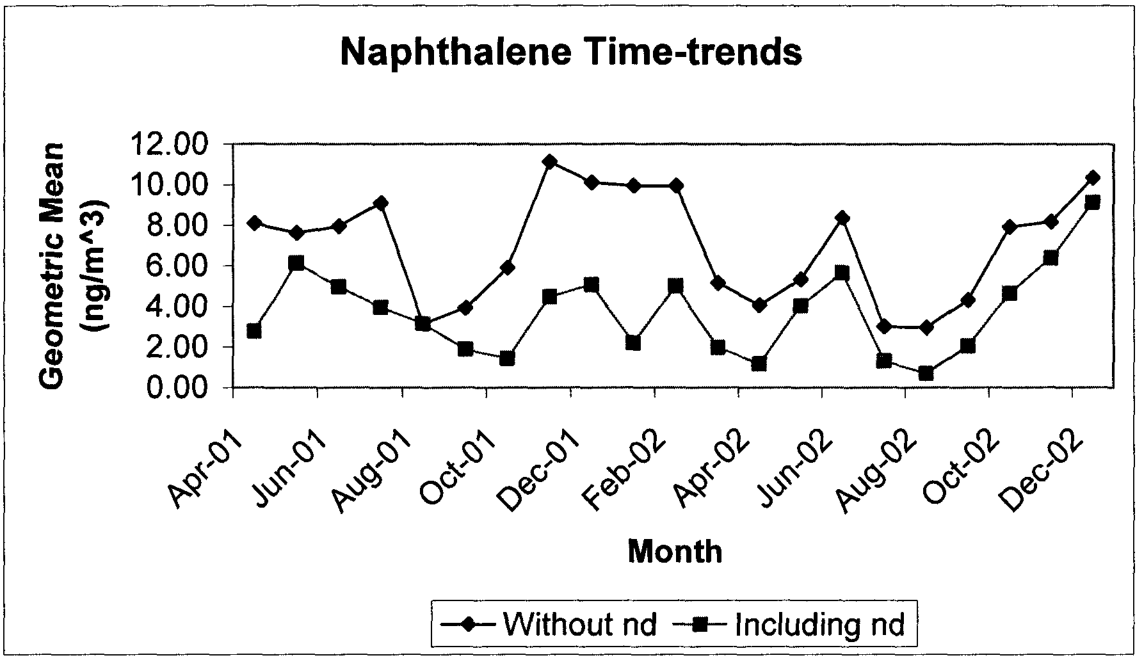

p-Cymene Time-trends
Geometric Mean (ng/m^3)
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd


1,3-Dichlorobenzene Time-trends
Geometric Mean (ng/m^3)
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd


1,4-Dichlorobenzene Time-trends
Geometric Mean (ng/m^3)
100.00
80.00
60.00
40.00
20.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd


Hexachloroethane Time-trends
Geometric Mean (ng/m^3)
300.00
250.00
200.00
150.00
100.00
50.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd


1,2,4-Trichlorobenzene Time-trends
Geometric Mean (ng/m^3)
25.00
20.00
15.00
10.00
5.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd


Naphthalene Time-trends
Geometric Mean (ng/m^3)
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Apr-01
Jun-01
Aug-01
Oct-01
Dec-01
Feb-02
Apr-02
Jun-02
Aug-02
Oct-02
Dec-02
Month
Without nd
Including nd

Appendix 6: Output of SAS PROC SEMIVARIOGRAM for Factor I

```
VARNAME         LAG          COUNT       DISTANCE        VARIOG      COVAR          RVARIO
factorI                    -1        11434                                          1.286
factorI                     0        73525           1.636       0.906          0.481            0.810
factorI                     1       209265           4.967       0.969          0.408            0.880
factorI                     2       168252           9.985       0.998          0.327            0.912
factorI                     3       176149          14.960       0.989          0.356            0.907
factorI                     4       182757          19.855       0.910          0.315            0.835
factorI                     5       174340          25.018       0.930          0.281            0.852
factorI                     6       170050          30.169       0.956          0.310            0.880
factorI                     7       213861          35.045       0.915          0.317            0.858
factorI                     8       169113          39.927       0.951          0.251            0.887
factorI                     9       180651          45.067       0.958          0.259            0.905
factorI                    10       203009          49.883       0.979          0.337            0.920
factorI                    11       177262          54.947       1.030          0.281            0.961
factorI                    12       135889          59.999       0.971          0.230            0.910
factorI                    13       162494          65.167       0.967          0.248            0.894
factorI                    14       187063          70.039       0.922          0.315            0.875
factorI                    15       215649          74.997       0.939          0.370            0.889
factorI                    16       204114          79.993       0.957          0.310            0.908
factorI                    17       215193          84.981       0.991          0.308            0.929
factorI                    18       190006          89.892       0.983          0.255            0.923
factorI                    19       173320          94.913       0.992          0.221            0.933
factorI                    20       156595          99.987       0.950          0.244            0.909
factorI                    21       170787         105.015       0.933          0.353            0.903
factorI                    22       185137         109.976       0.965          0.305            0.937
factorI                    23       193719         114.947       1.019          0.296            0.979
factorI                    24       186669         120.066       0.982          0.266            0.932
factorI                    25       169608         125.034       1.011          0.213            0.961
factorI                    26       167202         130.043       1.067          0.313            1.017
factorI                    27       179559         135.026       1.110          0.281            1.071
factorI                    28       223170         140.025       1.165          0.296            1.122
factorI                    29       213806         144.879       1.102          0.298            1.027
factorI                    30       213175         150.083       1.040          0.226            0.991
factorI                    31       213885         154.917       1.123          0.228            1.067
factorI                    32       236611         159.994       1.093          0.227            1.041
```

* Count = the number of observations within a given lag; Distance = average distance between observations within a given lag; Variog =regular semi-variance within a given lag; Covar=covariance within a given lag; Rvario= robust semi-variance within a given lag.