Toronto Residents’ Exposure to Ultrafine Particles

by

Kelly Maria Sabaliauskas

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Chemical Engineering and Applied Chemistry
University of Toronto

© Copyright by Kelly Sabaliauskas 2015
In urban areas, ultrafine particles (UFP: defined as particulate matter with diameters less than 100nm) are emitted in significant quantities from vehicles and form through a complex series of secondary reactions in the atmosphere. Large uncertainties surrounding the long-term behaviour and spatial distribution of UFP in urban areas have been a significant obstacle for exposure assessment. This research examined one of the longest existing urban UFP data sets, collected at a roadside location in downtown Toronto. Between 2006 and 2011, the concentration of particles with diameters <50nm and 50-100nm decreased by 21% and 17%, respectively. This reduction in concentration was attributed to changes in the vehicle fleet and reduced usage of coal-fired power plants for electricity generation. In addition, this research found that the shape of the particle size distribution exhibited distinct temporal and spatial behaviour suggesting that a single monitoring station does not provide sufficient information about UFP for an entire urban area. This investigation also produced a land-use regression model that was used to estimate the range of concentrations that exist across Toronto during the summer months. The highest concentrations were consistently observed near the downtown core and around highways and industrial areas. Finally, this work provides a foundation for future field studies in Toronto.
Acknowledgments

Funding for SOCAAR was provided by the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund. This work was supported by the Ontario Ministry of the Environment's Best in Science Research Program and the Ontario Graduate Scholarships in Science and Technology.

This work was made possible by:

The scientific guidance and financial support from Health Canada, Environment Canada and the Ontario Ministry of the Environment

The support from my colleagues in the Southern Ontario Centre for Atmospheric Aerosol Research

The continuous scientific expertise and mentorship from Professor Greg Evans

The enthusiasm and assistance from the summer research technicians: Christopher Reali, Tim Sun, Elki Tsang, Amy Senn, Tasko Olevski, Abby Eldib, Tanin Amerian and Scott Orr

The study participants that hosted the air quality monitoring equipment in their backyards

The encouragement from my family and friends
# Table of Contents

Acknowledgments ................................................................................................................................. iii

Table of Contents ..................................................................................................................................... iv

List of Tables ........................................................................................................................................... x

List of Figures ......................................................................................................................................... xii

List of Appendices ................................................................................................................................. xv

List of Abbreviations .............................................................................................................................. xvi

Chapter 1 Introduction ............................................................................................................................ 1

1.1 Background and motivation ............................................................................................................... 1

1.2 Particulate matter ............................................................................................................................. 2

1.3 Ultrafine Particles ............................................................................................................................ 3

1.4 Exposure Assessment and PM .......................................................................................................... 4

1.4.1 UFP Exposure Assessment ......................................................................................................... 6

1.5 Study Location: Toronto, Canada ...................................................................................................... 7

1.5.1 Air Quality in Toronto ............................................................................................................... 9

1.6 Scope of Work ................................................................................................................................ 11

1.7 Thesis structure ............................................................................................................................... 11

Chapter 2 Theoretical Principles: Ultrafine Particles ............................................................................... 12

2.1 Executive Summary ......................................................................................................................... 12

2.1.1 Author Contributions .............................................................................................................. 12

2.2 Introduction .................................................................................................................................. 12

2.2.1 Processes Governing Ultrafine Particle Characteristics ......................................................... 13

2.3 Sources of Ultrafine Particles in Urban Centres ............................................................................ 15

2.4 Ultrafine Particles Emissions from Vehicles .................................................................................. 15

2.4.1 Vehicle Fleet Characteristics .................................................................................................... 16
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.2 Fuel and Lubricating Oil Characteristics</td>
<td>16</td>
</tr>
<tr>
<td>2.4.3 Engine Operating Conditions</td>
<td>17</td>
</tr>
<tr>
<td>2.5 Urban Scale Variability in Ultrafine Particle Concentrations</td>
<td>17</td>
</tr>
<tr>
<td>2.6 Factors influencing UFP exposure near roadways</td>
<td>19</td>
</tr>
<tr>
<td>2.7 Summary</td>
<td>20</td>
</tr>
<tr>
<td>Chapter 3 Five-Year Roadside Measurements of Ultrafine Particles in a Major Canadian City</td>
<td>21</td>
</tr>
<tr>
<td>3.1 Executive Summary</td>
<td>21</td>
</tr>
<tr>
<td>3.1.1 Author Contributions</td>
<td>22</td>
</tr>
<tr>
<td>3.2 Abstract</td>
<td>22</td>
</tr>
<tr>
<td>3.3 Introduction</td>
<td>23</td>
</tr>
<tr>
<td>3.3.1 Urban UFP Ambient Measurements</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Methodology</td>
<td>25</td>
</tr>
<tr>
<td>3.4.1 Study Location Characteristics</td>
<td>25</td>
</tr>
<tr>
<td>3.4.2 Sampling Location</td>
<td>25</td>
</tr>
<tr>
<td>3.4.3 Instrumentation</td>
<td>26</td>
</tr>
<tr>
<td>3.5 Results and Discussion</td>
<td>28</td>
</tr>
<tr>
<td>3.5.1 Meteorology</td>
<td>28</td>
</tr>
<tr>
<td>3.5.2 Daily Traffic Variation</td>
<td>29</td>
</tr>
<tr>
<td>3.5.3 Long-Term and Seasonal Behaviour of Ultrafine Particles</td>
<td>30</td>
</tr>
<tr>
<td>3.5.4 Diurnal Behaviour</td>
<td>38</td>
</tr>
<tr>
<td>3.5.5 Particle Size Distribution</td>
<td>40</td>
</tr>
<tr>
<td>3.5.6 Effect of Changing Vehicle Fleet on UFP Emissions</td>
<td>41</td>
</tr>
<tr>
<td>3.5.7 Correlation with Criteria Pollutants and Meteorological Parameters</td>
<td>43</td>
</tr>
<tr>
<td>3.5.8 Predicting UFP Concentrations Using a Multiple Linear Regression Model</td>
<td>45</td>
</tr>
<tr>
<td>3.6 Conclusions</td>
<td>50</td>
</tr>
<tr>
<td>3.7 Acknowledgements</td>
<td>50</td>
</tr>
</tbody>
</table>
Chapter 5 The Application of Wavelet Decomposition to Quantify the Local and Regional Sources of Ultrafine Particles in Cities

5.1 Executive Summary

5.1.1 Author Contributions

5.2 Abstract

5.3 Introduction

5.4 Methodology

5.4.1 Study Location

5.4.2 Vehicle Fleet

5.4.3 Field Sites

5.4.4 Instrumentation

5.4.5 Other Data Sources:

5.4.6 Data Analysis: Wavelet Decomposition

5.4.7 Coefficient of Divergence

5.4.8 Traffic and Geographical Modelling

5.5 Results

5.5.1 Wavelet Decomposition: Background Subtraction

5.5.2 PN Concentration Behaviour at the Central Site

5.5.3 PN Concentration Behaviour at the Field Sites

5.5.4 Correlation and Coefficient of Divergence

5.5.5 UFP and Traffic Volume

5.6 Conclusions

5.7 Acknowledgements

Chapter 6 The Development of a Land-Use Regression Model for Ultrafine Particles in Toronto, Canada

6.1 Executive Summary

6.1.1 Author Contributions
7.1.3  Objective 3: Quantify the contributions of local scale and regional scale sources of UFP

7.1.4  Objective 4: Determine if land-use regression model can be applied to UFP data to quantify the range of possible concentrations that exist across Toronto

7.2  Insights from Data Mining

7.3  Other Perspectives

7.4  Future Research Directions

References
List of Tables

Table 2-1 Effect of each formation, transformation and removal process on the size and number of observed UFP ................................................................. 14

Table 3-1 Summarized data sources and availability ........................................... 27

Table 3-2 Average ± 95% confidence interval for PN_{50}, PN_{50-100} and PN_{100-300} on a seasonal basis ................................................................................................. 32

Table 3-3 Summary of recent UFP results from urban areas ..................................... 37

Table 3-4 Spearman Rank Order correlation coefficients of 24-hour averaged PN concentrations, criteria pollutant concentrations, and meteorological conditions. ................................................. 43

Table 3-5 Beta values ± standard error for the multiple linear regression model .............. 49

Table 4-1 Summary of the source of each PSD type, its frequency of occurrence and its percentage contribution to the total observed PN concentration. ................................................. 58

Table 4-2 The overall PSD type counts and their % contribution to the overall PSD mix ....... 61

Table 4-3 Seasonal Spearman’s Rank Order correlation coefficients for 24-hour PSD type counts and 24-hour PM_{2.5} filter data obtained from the NAPS network and meteorological data ............ 71

Table 4-4 Average cosine similarity between the centroid of each PSD type and the two residential field sites and the Central site ............................................................................. 72

Table 4-5 Characteristics of each PSD type throughout the measurement period. LRT refers to long-range transport. a: denotes the diameter of the maximum mode ........................................... 77

Table 5-1 The characteristics of the downtown and residential field sites ....................... 85

Table 5-2 Summary of measurements collected at the central site between June 13th and August 19th 2008 ............................................................................................................. 92
Table 5-3 The Spearman Rank Order Correlation coefficients between the 24-hour average PN concentration and pollutants collected at the Ontario Ministry of the Environment station located 1km from the central site.

Table 5-4 PN concentration ± 95% confidence interval for the original PN concentration data and the calculated contribution of the urban-regional (U-R) and local-neighbourhood (L-N) scale sources at each field site and the central site.

Table 5-5 Spearman’s Rank Order correlation coefficients (r_s) coefficients of divergence (COD) are reported for each field site at 1-minute time resolution (C2 and D2 = 2-minute time resolution).

Table 6-1 Examined variables for the PN concentration LUR for Toronto.

Table 6-2 Mean ± 95% confidence interval for major arterial (>20,000 vehicles per day) and local roads (<2,000 vehicles per day).

Table 6-3 Predictor variables, corresponding buffers, coefficients, standardized coefficients, p-values, R² values and root mean square error (RMSE) values for the Toronto UFP LUR model.
List of Figures

Figure 1-1 Theoretical mass distribution of particulate matter in the atmosphere ...................... 3
Figure 1-2 Theoretical number distribution of ultrafine particles in the atmosphere ................. 4
Figure 1-3 Urbanization in the Greater Toronto Area between 1984 and 2012 ....................... 8
Figure 1-4 Variation of the six criteria pollutants in downtown Toronto between 2006 and 2012 ............................................................................................................. 10

Figure 2-1 The dominant formation and transformation processes governing the growth of UFP .................................................................................................................. 14
Figure 2-2 The range of UFP concentrations reported in the literature for each spatial scale ..... 18
Figure 2-3 The variation of UFP number concentration while walking along high and low traffic roads in Toronto, Canada ................................................................................. 19

Figure 3-1: A map of the sampling locations in downtown Toronto ........................................ 26
Figure 3-2 Wind speed and wind direction recorded at Pearson International Airport between March 2006 and May 2011 ......................................................................................... 29
Figure 3-3 Traffic variation on College St. Data was collected by the SmartEye Traffic Sensor from the roof of the Gage Institute building between June and December 2010 .................. 30
Figure 3-4 Long-term behaviour of UFP in downtown Toronto .................................................. 31
Figure 3-5 Seasonal behaviour of UFP in downtown Toronto .................................................... 33
Figure 3-6 Wind directionality of UFP in downtown Toronto .................................................... 34
Figure 3-7 Diurnal behaviour of PN$_{50}$, PN$_{50-100}$, PN$_{100-300}$, and the GMD on weekdays and weekends ......................................................................................................................... 39
Figure 3-8 Ratio of winter PN concentration and Summer PN concentration ......................... 40
Figure 3-9: Evolution of the particle size distribution on weekdays during the summer and winter. ......................................................................................................................... 41

Figure 3-10 Average PN concentration between 7:00 and 9:00 on weekdays and weekends between 2006 and 2010. ................................................................................................................................. 42

Figure 3-11 Variation in the Spearman Rank Order Correlation Coefficient between the size resolved PN concentrations and NO₂, SO₂, and PM2.5 concentrations ................................................. 44

Figure 3-12 Comparison between the measured and predicted PN concentration ............... 46

Figure 3-13 Results from the multiple linear regression model .............................................. 47

Figure 3-14 Monthly average PN concentration values between 2003 and 2010 ................. 48

Figure 4-1 Particle size distribution for each PSD type................................................................. 60

Figure 4-2 Fourier analysis of the PSD types on a daily, weekly, monthly, semi-annual, and annual basis........................................................................................................................................ 63

Figure 4-3 Diurnal trends for each PSD type............................................................................. 65

Figure 4-4 Ratio of the contribution of each PSD during the summer to the contribution of each PSD during the winter..................................................................................................... 66

Figure 4-5 The contribution of each PSD type (%) as a function of wind directionality. ........ 68

Figure 4-6 The wintertime and summertime weekday diurnal trend......................................... 74

Figure 4-7 Example of long-range transport impacting Toronto between June 27th and July 1st 2006...................................................................................................................................................... 75

Figure 5-1 Map of Toronto and location of downtown sites (central, A, B) and residential sites (C, D, E, F) monitoring sites......................................................................................................................... 84

Figure 5-2 The sampling dates and instruments used at each site ........................................... 86
Figure 5-3 Variation of PN concentration (upper panel), urban-regional scale (middle panel) and local-neighbourhood scale (lower panel) at the central site between August 1st 2008 and August 19th 2008 ................................................................. 90

Figure 5-4 Wind direction dependence of the A) Original PN concentration data, B) urban-regional scale PN concentration, C) local-neighbourhood scale PN concentration, and D) the local-neighbourhood scale ratio (%) ........................................................................................................ 93

Figure 5-5 Upper panel: median local-neighbourhood scale PN concentration ratios regressed against the number of vehicles within a 2.5km buffer of the field site ............................................. 98

Figure 6-1 Map of the field sites and mobile monitoring routes ......................................................... 106

Figure 6-2 The distribution of particle counting instruments during the fixed and mobile phase of the study ................................................................................................................................. 107

Figure 6-3 A typical walking route in downtown Toronto. ................................................................. 113

Figure 6-4 The diurnal behaviour of UFP observed at the central site and fixed residential monitoring sites ........................................................................................................................ 115

Figure 6-5 The probability that the average of three short-duration PN concentrations are within x% of the average PN concentration collected between 12:00 and 15:00 on weekdays throughout the summer of 2008. ........................................................................................................ 116

Figure 6-6 The modelled spatial distribution of UFP in Toronto. .................................................... 119

Figure 6-7 The measured PN concentration at each fixed field site and the predicted PN concentration by UFP LUR model. ........................................................................................................ 120

Figure 7-1 Ratio of the PN concentration measured in GTA parks and downtown Toronto .... 132
List of Appendices

Appendix A .......................................................................................................................... 153
Appendix B .......................................................................................................................... 161
Appendix C .......................................................................................................................... 165
Appendix D .......................................................................................................................... 172
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CO</td>
<td>Carbon Monoxide</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon Dioxide</td>
</tr>
<tr>
<td>COD</td>
<td>Coefficient of Divergence</td>
</tr>
<tr>
<td>CPC</td>
<td>Condensation Particle Counter</td>
</tr>
<tr>
<td>CWS</td>
<td>Canada-Wide Standard</td>
</tr>
<tr>
<td>DMA</td>
<td>Differential Mobility Analyzer</td>
</tr>
<tr>
<td>dₚ or Dₚ</td>
<td>Diameter of a particle</td>
</tr>
<tr>
<td>EC</td>
<td>Environment Canada</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FMPS</td>
<td>Fast Mobility Particle Sizer</td>
</tr>
<tr>
<td>GMD</td>
<td>Geometric Mean Diameter</td>
</tr>
<tr>
<td>GTA</td>
<td>Greater Toronto Area</td>
</tr>
<tr>
<td>HEI</td>
<td>Health Effects Institute</td>
</tr>
<tr>
<td>HNO₃</td>
<td>Nitric Acid</td>
</tr>
<tr>
<td>H₂O</td>
<td>Water</td>
</tr>
<tr>
<td>H₂SO₄</td>
<td>Sulphuric Acid</td>
</tr>
<tr>
<td>Hₚ</td>
<td>Hopkins’ Statistic</td>
</tr>
<tr>
<td>LOO</td>
<td>Leave-one-out</td>
</tr>
<tr>
<td>LUR</td>
<td>Land-Use Regression</td>
</tr>
<tr>
<td>MLR</td>
<td>Multiple Linear Regression</td>
</tr>
<tr>
<td>MOE</td>
<td>Ontario Ministry of the Environment</td>
</tr>
<tr>
<td>NAPS</td>
<td>National Air Pollution Surveillance</td>
</tr>
<tr>
<td>NH₃</td>
<td>Ammonia</td>
</tr>
<tr>
<td>NH₄</td>
<td>Ammonium</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric Oxide</td>
</tr>
<tr>
<td>NO₂</td>
<td>Nitrogen Dioxide</td>
</tr>
<tr>
<td>NO₃⁻</td>
<td>Nitrate</td>
</tr>
<tr>
<td>NOₓ</td>
<td>Oxides of Nitrogen</td>
</tr>
<tr>
<td>OPG</td>
<td>Ontario Power Generation</td>
</tr>
</tbody>
</table>
pt = particle
PAHs = Polycyclic Aromatic Hydrocarbon
PM = Particulate Matter
PM$_{2.5}$ = Particulate Matter with diameter $\leq$ 2.5 microns
PM$_{10}$ = Particulate Matter with diameter $\leq$ 10 microns
PN = Particle Number
PN$_{50}$ = Particles with $d_p < 50$nm
PN$_{50-100}$ = Particles with $50 < d_p < 100$nm
PN$_{100-300}$ = Particles with diameters $100 < d_p < 300$nm
PSD = Particle Size Distribution
RH = Relative Humidity
RMSE = Root Mean Square Error
$r_s$ = Spearman’s Rank Order Correlation Coefficient
SE = Standard Error
SMPS = Scanning Mobility Particle Sizer
SO$_2$ = Sulphur Dioxide
SO$_4^{2-}$ = Sulphate
SOCAAR = Southern Ontario Centre for Atmospheric Aerosol Research
SR = Solar Radiation
T = Temperature
TDS = Traffic Data Sensor
UFP = Ultrafine Particles
UTMMS = University of Toronto Mississauga Meteorological Station
VOCs = Volatile Organic Compounds
WCPC = Water Condensation Particle Counter
WD = Weekday
WDir = Wind Direction
WHO = World Health Organization
WE = Weekend
WS = Wind Speed
YYZMS = Meteorological station located at Pearson International Airport
1.1 Background and motivation

Historically, poor urban air quality has been the end result of rapid economic growth coupled with inadequate and unenforced environmental regulation. For example, throughout North America and Europe during the 19\textsuperscript{th} century, a hazy cloud of pollutants enshrouded many urban areas due to the uncontrolled fossil fuel combustion that powered the Industrial Revolution. However, at that time, air pollution was synonymous with progress and its health consequences were unknown. During the first half of the 20\textsuperscript{th} century, two widely publicized extreme air pollution episodes occurred in Meuse Valley, Belgium (Nemery, et al., 2001) and in Donora, Pennsylvania (Schrenk et al., 1949) and were associated with 60 and 17 deaths, respectively. Evidence supporting the link between air pollution and human health was further strengthened as a result of an incident in December 1952. During a relatively cold period in London, England, a stagnant air mass hung over the city. Increased coal burning coupled with existing factory emissions created a toxic mixture of particulate matter (PM) and sulphur dioxide (SO\textsubscript{2}) that caused approximately 4,000 deaths over the course of a week and an additional 8,000 deaths during the months immediately following the episode (Bell et al., 2004).

These extreme air pollution episodes led governments in industrialized countries to enact legislation regulating the emission of pollutants such as PM and SO\textsubscript{2}. Further regulations and more stringent standards for industrial and vehicle emissions were implemented in subsequent decades, resulting in increased improvements in air quality. However, air quality continues to be a problem in developing countries. Recent extreme air pollution episodes in China and India have been subject to significant media scrutiny and public outrage. The World Health Organization (WHO) estimates air pollution exposure contributes to 7 million premature deaths worldwide (WHO, 2014). Thus, urban air pollution is a complex environmental issue that has existed for centuries and continues to impact urban residents today.
1.2 Particulate matter

Airborne PM is defined as a complex mixture of solid particles and liquid droplets that varies in size, shape and chemical composition (Seinfeld and Pandis 2006). PM can be emitted from both natural and manmade sources including forest fires, dust storms, volcanic eruptions, and fossil fuel combustion or form through a series of reactions between gas phase pollutants. PM is ubiquitous in the atmosphere and can be found in remote environments such as the arctic and highly polluted urban centers. Due to its multiple sources, the physical and chemical characteristics of PM can vary significantly between regions, seasons, and at different times of the day.

Particles can be quantified in terms of their number, surface area, volume, or mass. PM is regulated on a size basis described by the metric of mass of particles per unit volume (Figure 1-1). PM\textsubscript{10} describes PM with diameters (d\textsubscript{p}) <10 μm. PM\textsubscript{10} includes the coarse (2.5 <d\textsubscript{p} <10 μm), fine (0.1 < d\textsubscript{p} <2.5 μm) and ultrafine (d\textsubscript{p} < 0.1 μm) size fractions of PM. PM\textsubscript{2.5} describes PM with d\textsubscript{p} < 2.5 μm. PM\textsubscript{2.5} includes the fine and ultrafine fractions of PM. Particles in the coarse and fine size ranges make up the majority of the mass, but are few in number while particles in the ultrafine range contribute the majority of the number and only a small percentage to the overall mass. Ultrafine particles (UFP) are theoretically captured within the PM\textsubscript{10} and PM\textsubscript{2.5} ranges; however, UFP possess little mass and contribute little to the overall mass reported in PM\textsubscript{10} and PM\textsubscript{2.5} mass concentrations.
Historically, all particles with $d_p < 30$ $\mu$m, or Total Suspended Particulates (TSP), were regulated. In recent years, PM regulations have changed to better characterize the size fractions of PM and include particles capable of entering the human respiratory tract. The fine and ultrafine size ranges are considered respirable (i.e. can enter the respiratory tract) and many jurisdictions have ambient standards for PM$_{2.5}$ quantified by its particle mass concentration. Within the Canadian context, the Canada Wide Standard for PM$_{2.5}$ is 30$\mu$g/m$^3$ based on a 24-hour averaging time. A jurisdiction is within compliance of this standard if the annual 98$^{th}$ percentile ambient measurement, averaged over 3 years is less than 30$\mu$g/m$^3$ by 2010. Recently, an updated ambient standard has been proposed for PM$_{2.5}$ to limit the annual average to 8.0$\mu$g/m$^3$. No ambient standards exist for ultrafine fraction of PM in Canada.

### 1.3 Ultrafine particles

UFP describe the continuum of PM that has $d_p < 100$nm. The particle number (PN) concentration is a commonly used metric to describe the variation of UFP in the atmosphere. Unlike the
theoretical mass distribution of PM depicted in Figure 1-1, there exist no natural splits in the PN distribution (Figure 1-2). Several terms are used to describe particles within the ultrafine range including nuclei (1 < \(d_p\) ≤ 10nm), Aitken (10 < \(d_p\) ≤ 100nm), accumulation (100 < \(d_p\) ≤ 2,500nm).

![Figure 1-2 Theoretical number distribution of ultrafine particles in the atmosphere](image)

In urban areas, nuclei and Aiken mode particles dominate the particle size distribution (PSD). Nuclei mode particles tend to be spherical, composed of sulphuric acid and semi-volatile organic compounds and have short atmospheric lifetimes. In contrast, Aitken mode particles tend to be irregular in shape, composed of black carbon and have long atmospheric lifetimes. Thus, particles in the ultrafine range may differ significantly in shape, composition and geographical range. A more detailed description of the sources and processes impacting the observed concentration of UFP will be presented in Chapter 2.

1.4 Exposure Assessment and PM

To assess exposure to PM, high quality, long-term and spatially resolved data is required. For TSP, \(\text{PM}_{10}\) and \(\text{PM}_{2.5}\), air quality monitoring networks relying on gravimetric (filters) or optical sampling techniques have been established in several countries and have operated for decades. In
Canada, the National Air Pollution Surveillance (NAPS) Network, has measured TSP since 1970, and added PM$_{10}$ and PM$_{2.5}$ in 1984 (Environment Canada, 2013a). The NAPS network uses a consistent sampling protocol using filters and well defined analysis procedures to achieve its stated objective to “monitor and assess the quality of ambient air in urban regions of Canada continuously” (Environment Canada, 2013a). The network began with six sites and expanded to include 21 active sites across the country. In total, 37 locations have at least two years of data available (Environment Canada, 2013b). In most jurisdictions, air quality monitoring networks have been established in order to determine if a given region meets ambient air quality standards. Since these networks were designed for compliance purposes rather than health studies, the data that is provided is more likely to reflect an ambient background concentration to which the population is exposed while outdoors rather than the typical likely concentration that an individual may encounter in their daily lives.

Although not explicitly designed for health studies, these long-term data sets enable epidemiologists to examine the impacts of both chronic and acute exposure to PM$_{2.5}$. To date, numerous population-based studies have found increased PM concentrations in the atmosphere are associated with higher incidence of mortality from cardiovascular and respiratory disease causes (Dockery et al., 1993; Samet et al., 2000; Schwartz et al., 1996). In addition, long-term monitoring enabled a number of studies to observe reductions in mortality in locations with improved air quality due to policy interventions or other factors that result in the closing of industrial facilities. For example, a year after banning coal sales in Dublin, Ireland, non-trauma deaths decreased by 6%, respiratory deaths by 15% and cardiovascular deaths by 10% (Clancy et al., 2002). In Hong Kong, after the sulphur content of fuel oil was limited to 0.5% by weight for all power plants and vehicles, Hedley et al., (2002) observed a reduction in all-cause mortality by 2.1%, deaths from respiratory causes by 3.9% and deaths cardiovascular causes by 2.0%. After an eight month long national labour disruption at copper smelters in the United States, Pope et al., (2007) observed a reduction in nationwide deaths from cardiovascular and other respiratory diseases causes of 2.5%. Finally, Pope et al., (2009) examined changes in PM concentrations and mortality across 51 cities in the United States and found an increase in life expectancy of 7.2 months per 10 μg-m$^{-3}$ reduction in PM$_{2.5}$ concentration.
Only a small portion of the available literature on the health impacts of PM$_{2.5}$ are presented above; however, these findings have been remarkably consistent across multiple countries with different study populations and pollutant mixes. These findings are also paradoxical because most have relied on measurements collected at a centralized location to assess population exposure across a study area. PM$_{2.5}$ is an example of a regional pollutant because it can exist in the atmosphere for multiple days (on the order of 3-5 days) and be transported from one jurisdiction to another. Studies that have evaluated PM$_{2.5}$ mass concentration gradients across urban areas have found high temporal correlation between field sites (reviewed by Wilson et al., 2005). As a result, interpolating between field sites is a viable method to estimate the spatial distribution of PM$_{2.5}$ across a study area. In addition, once the spatial relationships are established, a single monitoring location may be sufficient to estimate the population’s exposure to PM$_{2.5}$.

1.4.1 UFP Exposure Assessment

UFP monitoring is a relatively new area of investigation. UFP measurements are inherently more difficult to collect than the larger fractions of PM because UFP cannot be reliably captured on a filter and are too small to scatter light. As a result, methods that detect UFP are both more complex and expensive than those that can reliably measure PM$_{2.5}$. UFP measurements are further complicated by the lack of consistent sampling protocols, calibration standards, and instrumentation specifications. There exist no ambient standards for UFP and no routine data collection as part of air quality monitoring networks. However, despite these challenges, short-term measurements on the order of weeks to months have been collected in many urban areas around the world (reviewed by Kumar et al., 2014). To date, there are only a handful of urban locations with long-term data (>5 years) (Mejía et al., 2007; Sabaliauskas et al., 2012, Wåhlin et al., 2009; Wang et al., 2011b). As a consequence, there exists significant uncertainty surrounding the long-term behaviour of UFP and the factors that may influence its variation.

Further, unlike PM$_{2.5}$, UFP are emitted in significant quantities from vehicles and have short atmospheric lifetimes (seconds to hours). The localized nature of UFP can result in large concentration gradients around roadways and across urban areas. A number of studies have examined UFP concentrations around highways and found elevated concentrations within the
first 300m from the roadway (Zhu et al., 2002a, 2002b). In addition, only moderate correlation has been observed between multiple field sites across an urban area (Buzorius et al., 1999; Cyrys et al., 2008). In addition, the few studies that examined UFP concentrations at multiple locations were equipped with a small number of instruments and could not sample at more than a few sites simultaneously. Therefore, establishing spatial relationships between a centralized site and multiple temporary sites remains challenging.

The lack of long-term and spatially resolved UFP measurements makes assessing UFP exposure difficult and limits the types of epidemiological studies that can be performed to assess potential health impacts (Young, 2005). For example, the long-term and spatially resolved data that is available for PM$_{2.5}$ enables population cohorts to be followed over extended periods of time (prospective study) or studied retrospectively with low risk of exposure misclassification. With existing UFP measurements, the abovementioned study designs are not possible given the large uncertainties surrounding long-term behaviour of UFP and its potential spatial variation. To date, no long-term epidemiological studies with comparable study design as those that reported the initial health impacts of PM$_{2.5}$ have been conducted.

Existing epidemiological studies examining the potential impacts of UFP exposure have relied on short-term UFP measurements collected at a central site. To date, some studies have found associations between UFP concentrations and increased mortality and morbidity from cardiovascular and respiratory causes (Belleudi et al., 2010; Breitner et al., 2011; Braniš et al., 2010) while others have not (Andersen et al., 207; Iskandar, et al., 2012). A recent assessment examining the potential health impacts of UFP concluded that despite significant research activity investigating the variation of UFP, the lack of spatially resolved UFP data is a significant source of uncertainty for exposure assessment (HEI, 2014).

1.5 Study Location: Toronto, Canada

The geographical focus of this dissertation is Canada’s largest and most populous city. Toronto, Ontario has a land area of 630km$^2$ and population of 2.6 million people (Statistics Canada, 2011). When combined with the municipalities that make up the Greater Toronto Area (GTA), the population grows to 5.5 million people covering a region of 7,124km$^2$ (Statistics Canada,
During the last thirty years, the GTA has undergone rapid urbanization to accommodate the growing housing needs of its population (Figure 1-3). However, investment in transportation infrastructure has not kept pace with the growing mobility demands of GTA residents. Consequently, a large majority of GTA residents use their personal vehicles to travel to and from their workplaces (Statistics Canada, 2011). The number of vehicles travelling into and out of Toronto has increased by 75% since 1987 (Toronto Public Health, 2007).

![Figure 1-3 Urbanization in the Greater Toronto Area between 1984 (upper) and 2012 (lower). Time lapsed satellite imagery courtesy of Google Earth.](image)

The Organization for Economic Cooperation and Development (OECD) undertook a study of Toronto and identified the lack of investment in transportation infrastructure and integration between transit networks across the GTA as significant issues that limits the region’s capacity to compete with other metropolitan areas. The OECD estimated that traffic congestion cost the Toronto economy $3.3 billion and the GTA economy an additional $2.7 billion in 2006 (OECD, 2009). Further, the demand on the region’s transportation networks is expected to increase over the next 20 years as the GTA’s population increases by another 2.5 million people (Ontario
This daily mass movement of people and heavy dependence on personal vehicles has consequences for Toronto residents. The impacts of traffic are not uniformly distributed across the city. For example, a large proportion of downtown Toronto residents walk, cycle or ride public transit to their workplaces (Statistics Canada, 2011) and yet the downtown residents are disproportionately affected by emissions created by those that reside outside of the city. In addition, over half of GTA residents live within 250m of a major roadway (Evans et al., 2011) and increasing traffic volumes across the region may impact a significant portion of the population.

1.5.1 Air Quality in Toronto
Since 2006, a number of regulations aimed at reducing emissions from vehicles and electricity generation were enacted by the federal, provincial, and municipal governments. Regulations focused on reducing the sulphur content in gasoline and diesel fuels, increasing the ethanol content in gasoline, adopting more stringent emission controls for NO\textsubscript{x} on diesel engines, phasing out coal-fired power plants in Ontario and limiting vehicle idling. The results of this collective effort are encouraging. Long-term measurements collected at the downtown Toronto site are presented in Figure 1-4. Between 2006 and 2012, clear reductions in SO\textsubscript{2}, NO and NO\textsubscript{2} concentrations can be observed. CO concentrations decreased between 2006 and 2009 and subsequently increased between 2010 and 2011 and O\textsubscript{3} concentrations increased throughout the measurement period. While no policies were aimed directly at CO, the timing of the minimum concentration corresponds to the economic crisis that affected the United States in 2008 and 2009. The increase in O\textsubscript{3} concentrations was likely driven by the more favourable O\textsubscript{3} formation conditions that resulted as the concentration of NO and NO\textsubscript{2} decreased.
Figure 1-4 Variation of the six criteria pollutants in downtown Toronto between 2006 and 2012. All measurements were collected at the Ontario Ministry of the Environment downtown Toronto site. The MOE discontinued CO and SO$_2$ monitoring in downtown Toronto in 2012.

In 2004, Toronto Public Health reported that air pollution contributed 1,700 premature deaths and 6,000 hospitalizations in Toronto. Since then, air quality in Toronto has improved and revised estimates suggest 1,300 premature deaths and 3,550 hospitalizations are attributable to air pollution. Toronto Public Health estimates that half of the pollution in the city is generated locally from the vehicles and the rest is generated outside the boundaries of the city (Toronto Public Health, 2014).
1.6 Scope of Work
This research is motivated by the large uncertainty surrounding the long-term behaviour and spatial distribution of UFP in urban areas. Vehicles are a dominant source of UFP and over half of the GTA’s population lives within 250m of a major roadway. This research seeks to characterize the long-term and spatial variation of UFP in Toronto and develop effective data analysis techniques to further our understanding of this pollutant. This dissertation has four main objectives:

- **Objective 1**: Identify the factors that have influenced the temporal behaviour of UFP.
- **Objective 2**: Explore the use of alternative metrics to describe the temporal and spatial variation of UFP.
- **Objective 3**: Quantify the contributions of local scale and regional scale sources of UFP
- **Objective 4**: Determine if land-use regression modeling can be applied to UFP data to quantify the range of possible concentrations that exist across Toronto

1.7 Thesis structure
This dissertation is a compilation of published or to be submitted papers. At the beginning of each chapter, an executive summary highlights the research significance of the paper. This executive summary also outlines the contributions of my co-authors and includes an abstract of the findings. For each chapter, I have done the majority of the conceptualization, identification and design of the research strategy, data analysis and interpretation, creation of the figures and tables, and preparation of the manuscript.
Chapter 2
Theoretical Principles: Ultrafine Particles

2.1 Executive Summary
This chapter contains an excerpt from a chapter published in the book: “Exposure to Ultrafine Particles in Urban Centres” in Urban Airborne Particulate Matter: Origins, Chemistry, Fate and Health Impacts (reprinted with permission from Springer, license number: 3333171368774):


The original intent of this publication was to provide an overview of the state of exposure assessment to UFP. The excerpt presented in this chapter provides the technical background information required to understand the complex behaviour of UFP in the urban atmosphere that will be presented through the results in this thesis. Concepts relating to the formation, transformation and removal pathways of UFP are described along with the factors that may influence exposure to ambient levels of UFP.

2.1.1 Author Contributions
This review paper was compiled and written by me with input and guidance from Professor Greg Evans

2.2 Introduction
Ambient particulate matter (PM) is often classified based on particle size due in part to the size dependence of particle deposition within the respiratory system. Ultrafine particles (UFP) are the smallest particles, with diameters less than 100 nm. They can originate from anthropogenic sources such as fossil fuel combustion, or naturally through nucleation of biogenic sulphuric acid, water, and ammonia molecules. The mass of PM per cubic meter is the most commonly used metric to describe airborne particles. Despite being few in number, the largest particles contribute the majority of the PM mass. In contrast, the majority of all atmospheric particles are
found in the ultrafine size range. Although numerous, UFP contribute little to the overall mass of PM in the atmosphere.

From a health and policy perspective, a mass metric does not adequately describe exposure to UFP. The number or surface area of particles per unit volume may be a better metric. However, at this time, only limited exposure data based on number concentration exists and no number concentration based ambient standards have been established. Further, there is no consensus yet as to whether size, chemical composition, or some combination thereof provides the best measure of UFP toxicity.

UFP has been shown to vary spatially over small distances in relation to traffic patterns on nearby roadways. Significant sources of UFP also exist within homes. Consequently, there may be substantial exposure variability across an urban population. This paper will examine sources of outdoor exposure to UFP in cities. Sources and processes influencing UFP concentrations are also discussed along with their spatiotemporal variability.

### 2.2.1 Processes Governing Ultrafine Particle Characteristics

UFP can be emitted directly from anthropogenic sources such as vehicles, furnaces or other technologies involving the combustion of fossil fuels. They can also form through a series of secondary chemical reactions in the atmosphere. There are five main processes that govern the formation, transformation and removal of UFP: nucleation, condensation or evaporation of semi-volatile species, coagulation, deposition, and dilution. The effect of each of these processes on the total number of particles in the atmosphere and their size is summarized in Table 2-1. Each process affects the number and size of UFP and operates over vastly different characteristic times. Further, these processes have varying importance for different sized particles.
Table 2-1 Effect of each formation, transformation and removal process on the size and number of observed UFP

<table>
<thead>
<tr>
<th>Process</th>
<th>Effect on the number of particles</th>
<th>Effect on the average particle size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleation</td>
<td>Increase</td>
<td>Decrease</td>
</tr>
<tr>
<td>Dilution</td>
<td>Decrease</td>
<td>Preserved</td>
</tr>
<tr>
<td>Condensation</td>
<td>Conserved</td>
<td>Increase</td>
</tr>
<tr>
<td>Evaporation</td>
<td>Conserved</td>
<td>Decrease</td>
</tr>
<tr>
<td>Coagulation</td>
<td>Decreases number of small particles</td>
<td>Increase</td>
</tr>
<tr>
<td>Deposition</td>
<td>Decreases number of large particles</td>
<td>Decrease</td>
</tr>
</tbody>
</table>

The dominant formation processes of UFP are binary nucleation from interactions between sulphuric acid and water as well as ternary nucleation between sulphuric acid, ammonia, and water. Once formed, the particles ‘age’ by accumulating layers of organic and other semi-volatile compounds (see Figure 2-1). Particles are further transformed through coagulation whereby growth occurs through collisions between particles of the same or different sizes.

Figure 2-1 The dominant formation and transformation processes governing the growth of UFP

The relative importance of each transformation and removal processes varies by size. For the smallest particles, diffusion to surfaces, evaporation, and coagulation are favoured as
transformation and removal processes. As the particle gets smaller (<10 nm) the Kelvin effect (i.e. vapour pressure of a curved surface is greater than a non-curved surface) becomes more important thus making it easier for molecules to leave the particle’s surface by evaporation. It is suggested by Hinds’ (1999) research, Brownian motion for a 10 nm particle is 80 times faster than for particles greater than 100 nm. As a result, the smallest particles are more likely to collide with each other and diffuse to surfaces than their larger counterparts. Conversely, larger particles can be removed through impaction, or wet or dry deposition. Thus loss through deposition tends to be lowest for the intermediate accumulation mode (greater than 90 nm in diameter) particles.

Dilution and evaporation are posited by Zhang and Wexler (2002, 2004a) as being the most dominant processes affecting the size distribution of UFP. Evaporation plus coagulation appeared to explain a large portion of the observed evolution in nuclei mode (less than 50 nm in diameter) size distributions as shown by Jacobson et al. (2005). Zhu et al. (2002a) argue that dilution for the largest particles and coagulation for the smallest particles are the most important mechanisms for the transformation of the size distribution. It should be noted that the relative importance of coagulation as a removal process for the smallest particles remains a subject of debate in the literature.

### 2.3 Sources of Ultrafine Particles in Urban Centres

In urban settings, vehicles are the dominant source of UFP observed outdoors. Although regional scale nucleation and growth events can contribute to the particle number concentration, these events are believed to only contribute a small portion of the UFP in urban settings.

### 2.4 Ultrafine Particles Emissions from Vehicles

The size, chemistry, and emission rate of UFP from vehicles depends on the vehicle fleet’s characteristics, their fuel types, and the fleet operating conditions.
2.4.1 Vehicle Fleet Characteristics

Fleet characteristics refer to the different classes, ages, control technologies and fuel requirements of vehicles on a given roadway. The vehicles used in the majority of cities are powered using gasoline or diesel fuels. Each of these fuel types requires specific engine technologies and emission control strategies. Gasoline powered vehicles rely on a volatilized fuel and air mixture ignited by a spark plug. Diesel powered engines rely on the heat of compression to ignite and burn the fuel, which is injected into the combustion chamber during the final stage of compression. Since these engines each rely on different processes to ignite the fuel, the resulting emissions are very different.

The differences between these two methods of combustion lead to differing emissions depending on engine load and driving conditions. Zielinska et al. (2004) conclude that diesel vehicles typically emit more UFP mass than spark ignition vehicles. Particles derived from diesel-powered vehicles contain non-volatile and semi-volatile polycyclic aromatic hydrocarbons, elemental carbon or “soot”, and sulphate. In contrast, the UFP emitted from spark ignition vehicles have been shown to contain little sulphate but significantly more ammonium and calcium than diesel-powered vehicles.

The control technology that is applied to the vehicle itself also has consequences to the ultimate emission and formation of UFP. Diesel particulate traps are now commonly used to control the mass emissions of PM. There are new classes of continuously regenerating diesel particulate filters that allow for PM emissions to be controlled. One of the consequences of this technology is an increased formation of small particles (10 nm) due to the absence of elemental carbon downstream of the trap.

2.4.2 Fuel and Lubricating Oil Characteristics

Now that many jurisdictions regulate the sulphur and biodiesel content of diesel fuel, significant variability exists in diesel fuel composition and, consequently, UFP emissions. One investigation has demonstrated a 30% decrease in PM emissions with use of 100% biodiesel fuel but the soluble organic fraction increased by 40% (Sharp 1998). The sulphur content in diesel fuel also
greatly influences the observed UFP emissions. In jurisdictions that have implemented ultra-low sulphur fuel, significant reductions in ambient UFP have been observed (Wåhlin 2009).

Kittelton et al. (2006) demonstrated that for diesel emissions on roadways, most of the mass (97%) in the 12–30 nm range was volatile with properties consistent with C24–C32 n-alkane compounds typical of lubricating oil. The high molecular weight molecules used in lubricating oil typically have lower volatility than unburned hydrocarbons. As a result, compounds from lubricating oil are more likely to be found in the smallest particles because they are more likely to condense first during cooling.

2.4.3 Engine Operating Conditions

Traffic density and congestion in cities depends on a balance between the number of vehicles, roadway design, and traffic management. The resulting operating conditions of individual vehicles can be described in terms of idling, cruising, deceleration, and acceleration. An increase in particle number concentration and decrease in particle size with increasing vehicle speed was reported by Kittelson et al. (2004). The composition of UFP also varies due to engine operation. Kubo et al. (2006) concluded that hydrocarbons from oxygenated fuel dominated the UFP composition during idling conditions (>C19) and lubricating oil dominated during deceleration (>C35) while sulphate contributed most to UFP under high engine loads for diesel vehicles.

2.5 Urban Scale Variability in Ultrafine Particle Concentrations

When studying the factors that influence exposure to vehicle generated UFP in cities three spatial scales need to be considered. The first scale is on the order of 1–2 m where UFP is formed near the tailpipe of a vehicle. Scale two considers the dilution and transformation processes that occur within 100 m from the roadway. Third, the urban scale is on the order of kilometres. The time scales related to the abovementioned spatial scales are on the order of seconds to several hours. Meteorology is also an influence on the extent of mixing and dilution on a regional scale causing day-to-day variability in UFP concentrations. As a result, each component of this transformation cannot be studied independently. Figure 2-2 summarizes UFP concentration ranges that are found at each of the abovementioned spatial scales based on studies conducted in urban and rural...
settings throughout the world (Buzorius et al. 1999; Cyrys et al. 2008; Kaur et al. 2006; Ketzel et al. 2004; Mejía et al. 2008; Puustinen et al. 2007; Kim et al. 2002; Westerdahl et al. 2005; Zhu et al. 2002a, b)

Figure 2-2 The range of UFP concentrations reported in the literature for each spatial scale.

From an exposure assessment perspective, Figure 2–2 also demonstrates that, depending on an individual’s location within a city, the outdoor UFP number concentration can vary by as much as three orders of magnitude. Furthermore, it should be noted that the size distribution, and presumably the chemistry of UFP, at each spatial scale also differs. Near roadways, there are many small particles (<30 nm) that are derived from traffic sources. As distance from the roadway increases (100–300 m from the road), the number of small particles decreases due to evaporation and growth processes and the UFP size distribution is dominated by larger particles (>60 nm). One of the implications for these observations is that UFP not only exhibits spatial heterogeneity in terms of the total number of particles but also exhibits significant variability in size in relation to proximity to traffic.
2.6 Factors influencing UFP exposure near roadways

While individuals may only spend 7% of their time outdoors (Kleipis et al. 2001), they often spend that time outdoors near traffic sources such as when they walk along high traffic roads, wait at intersections, sit in vehicles with the windows open, or when working in buildings near traffic. This time spent near roadways results in exposure to elevated total UFP number concentrations and to smaller particles than is typically observed further away from roadways. There are, therefore, profound implications for individuals living near major arterial roadways because vehicle-generated UFP will also increase their indoor exposure.

Figure 2-3 The variation of UFP number concentration while walking along high and low traffic roads in Toronto, Canada.

Several studies have reported large spatial gradients in UFP concentrations beside major arterial roads and highways. However, the mechanisms underlying the associated evolution in particle size distribution are still under investigation. On major arterial roads, particles with diameters less than 50 nm (nuclei particles) are dominant. Both Zhu et al. (2002b) and Kittelson et al. (2004) observed no nuclei-sized particles at distances greater than 90 m away from the roadway.
Fushimi et al. (2008), in turn, observed a particle size distribution with a modal diameter of 20 nm that was present at a roadside site but absent at a background site.

UFP exposure near roadways in cities also varies temporally over very short time scales. When an individual walks along a major arterial roadway, he or she is exposed to a combination of an urban background concentration as well as a vehicle signal that consists of many spikes caused by passing diesel vehicles. The lifetime of these UFP concentration spikes is on the order of 1–2 min. Therefore, these UFP concentration spikes are most consequential for individuals sitting in vehicles with the windows open, and when walking or cycling along major arterial roadways. The intensity and frequency of these spikes relates directly to the nearby vehicle characteristics and operating conditions.

Illustrated in Figure 2-3 is a transition from a roadway with 25,000 vehicles per day to one with less than 2000 vehicles per day. On the high traffic road there are frequent spikes that correspond to passing diesel vehicles. On the low traffic road the frequency of these UFP concentration spikes decreases due to fewer diesel-powered vehicles. Despite moving away from a significant traffic source, a constant UFP baseline or urban background concentration is observed in both locations. Another interesting characteristic of note is that urban background concentration varies both temporally and spatially but over much larger and longer time scales than the spikes. This background is also strongly influenced by the traffic density within a few kilometres as well as by day-to-day changes in meteorology.

2.7 Summary

It has been established that numerous sources and processes govern the characteristics of UFP. As a result, UFP concentrations, size distributions, and presumably chemistry, exhibit large spatiotemporal variability in cities. Vehicles are the dominant source for UFP outside. Roadside concentrations vary over a range of characteristic times and distances. Nearby vehicles can produce short-lived concentration spikes while the concentration away from busy roads is related to neighbourhood traffic density and meteorology.
Chapter 3  
Five-Year Roadside Measurements of Ultrafine Particles in a Major Canadian City

3.1 Executive Summary

This chapter contains a paper published in the peer-reviewed journal Atmospheric Environment (reprinted with permission).


This chapter seeks to identify the factors that influenced the temporal behaviour of UFP (Objective 1) by examining the long-term variation of UFP in downtown Toronto. This five year size-resolved UFP dataset is one of the longest of its kind in the world. This work was motivated by the lack of long-term monitoring of UFP in urban areas and sought to answer the following questions:

1. In a major urban center, how do ultrafine particles vary on an annual, seasonal and diurnal basis?
2. Did the concentration of UFP change between 2006 and 2011?
3. Can UFP variation be predicted by the behaviour of other pollutants and meteorological patterns?

Although all questions are important, the most relevant for other urban areas that lack UFP data is question 3. Many jurisdictions monitor other pollutants such as NO, NO$_2$, O$_3$, CO, SO$_2$ and PM$_{2.5}$. Therefore, establishing relationships between these pollutants, meteorological patterns and UFP may provide an opportunity to extrapolate a short-term UFP data set to longer time horizons.
3.1.1 Author Contributions
This study benefited from the contributions from the following co-authors: Cheol-Heon Jeong, Xiaohong Yao, Jun Yun-Seok, Parnian Jadidian, and Greg J. Evans. Cheol-Heon Jeong maintained, operated and provided quality assurance for the Fast Mobility Particle Sizer throughout the five year measurement period and provided advice on the data interpretation. Xiaohong Yao provided advice on data interpretation and feedback on the initial drafts of the manuscript. Jun Yun-Seok provided input and access to a complimentary ultrafine particle dataset for hypothesis testing during the manuscript development stages. Parnian Jadidian validated and operated the Traffic Sensor and compiled these data. The study conceptualization, design, data compilation, validation, analysis and interpretation were performed by me. The article was entirely written by me with input, guidance and editing from Professor Evans.

3.2 Abstract
This study reports the diurnal, seasonal and annual variation of ultrafine particles (UFP) in a large city. Particle number (PN) concentrations were measured in Toronto, Canada on a major arterial roadway between March 2006 and May 2011 using a Fast Mobility Particle Sizer. The PN in the size range of 8 to 300nm decreased during the measurement period primarily due to changes in the vehicle fleet. PN_{50} (particles with diameters less than 50nm) decreased by 21% between 2006 and 2010. Notably, the majority of the reduction occurred during the winter months. PN_{50} exhibited the strongest seasonality and diurnal trend. PN_{50-100} (particles between 50 and 100nm) and PN_{100-300} (particles between 100 and 300nm) decreased by 17% and 24%, respectively. Correlation analysis between gas phase criteria pollutants showed good correlation between PN_{50-100} and NO_2, SO_2, and PM_{2.5}. In contrast, PN_{50} exhibited the highest correlation with temperature, NO and NO_2. A multiple linear regression model was developed for each size fraction. The model adequately explained the annual, seasonal and day-to-day variability of PN_{50-100} (R^2 = 0.64) and PN_{100-300} (R^2 = 0.83). The model captured the annual and seasonal variability of PN_{50} but only partially explained the day-to-day variability (R^2 = 0.52). The long-term reductions in PN_{50} indicate that policy interventions are having some success in slowly decreasing UFP concentrations in Toronto.
3.3 Introduction

Emerging epidemiological and toxicological evidence suggests ultrafine particles (UFP: particles with diameters less than 100nm) may be more toxic per unit mass than the larger fractions of particulate matter (PM) (Sioutas et al., 2005). UFP has been shown to preferentially deposit in the deepest portions of the respiratory tract as a function of size (Daigle et al. 2003), induce oxidative stress (Araujo et al., 2008), and translocate to secondary organs in the body (Oberdörster et al. 2004). Despite the attention that UFP has received, limited long-term monitoring has occurred in urban areas worldwide.

In addition to primary emissions, UFP can form in the atmosphere. A number of nucleation mechanisms have been proposed including interactions between ions (ion-induced nucleation) (Yu and Turco, 2000), sulphuric acid and water (binary nucleation) or sulphuric acid, water, and ammonia (ternary nucleation) (Korhonen et al., 1999). Once in the atmosphere, particles can grow or shrink as semi-volatile species condense onto or evaporate from their surface (Yao et al., 2010). The majority of all atmospheric particles are in the ultrafine range, yet they contribute little to the overall aerosol mass. Other investigations in urban areas have observed poor correlation between the particle number (PN) concentration and particle mass concentration (Jeong et al., 2006; Ketzel et al., 2004a). Therefore, the long-term behaviour of UFP cannot be inferred from existing PM mass measurements. To date, most PM emission and ambient standards are based on the mass of particles per unit volume rather than the number of particles per unit volume; although, particle number-based emission standards are being introduced for diesel-powered passenger cars and light-duty commercial vehicles in the European Union (European Union, 2008). Thus, existing ambient air quality monitoring networks and the current regulatory framework do not adequately address the potential risk that UFP exposure may pose for susceptible populations.

3.3.1 Urban UFP Ambient Measurements

Elevated PN concentrations have been reported in urban areas around the world (Aalto et al., 2005; Buzorius et al., 1999; Jeong et al., 2004; 2006; Mejía et al., 2007). The PN concentration exhibits strong seasonal and diurnal behaviour due to vehicle emissions, photochemical nucleation, meteorology, and building heating requirements during cooler periods (Aalto et al.,
The highest UFP concentrations have been reported to occur during the winter months due to reduced ambient temperatures, limited mixing, and increased combustion for residential heating (Aalto et al., 2005; Cyrys et al., 2008; Jeong et al., 2004). Finally, seasonal parameters such as temperature, humidity, solar radiation, and mixing can enhance or inhibit new particle formation (Jeong et al., 2004; 2006; Shi et al., 1999, Yao et al., 2007; Yao and Zhang, 2011).

UFP exhibits a characteristic diurnal behaviour where the concentration peaks during the morning and afternoon periods. The increase in UFP from 6:00 and 9:00 in the morning corresponds to the morning rush hour period and has been well documented in several urban centers (Buzorius et al., 1999; Jeong et al., 2004; 2006). The second peak that occurs during the afternoon period is associated with secondary particle formation (Jeong, et al., 2004; 2006; 2010; Kulmala et al., 2004) and may be correlated with the intensity of solar radiation (Shi et al., 1999). Finally, the diurnal behaviour of UFP has been reported to differ on working days and non-working days due to differences in driving patterns in the areas surrounding the monitoring sites (Hussein et al., 2004; Jeong et al., 2006).

Near roadways, the PN concentration and particle size distribution varies as a function of distance, traffic volume (particularly trucks), wind speed and wind direction (Zhu et al., 2002a; 2002b; Hu et al., 2009; Fushimi et al., 2008). This paper explores the annual, seasonal, and diurnal variation of UFP in Toronto, Canada, and explores the role of vehicle fleet changes, and reductions in coal-fired power plant usage. Vehicular emissions account for the vast majority of NO and NO\textsubscript{2} emissions in Toronto (MOE, 2009). In recent years, reductions in ambient NO\textsubscript{x} concentrations have occurred (Geddes et al., 2009) due to the introduction of more stringent vehicle emission standards (Tier II limits on NO\textsubscript{x} emissions: 0.07g/mile) and mandatory vehicle emission testing. Further, several coal-fired power plants have been closed, operated at reduced capacity, or retrofitted with scrubbers thereby reducing ambient SO\textsubscript{2} emissions. Correlation between combustion-related pollutants and UFP has been previously reported (Cyrys et al., 2008). Therefore, changes in vehicle emissions or emissions from electricity generation may be indicative of changes in the ambient PN concentration and will be examined in this study.
3.4 Methodology

3.4.1 Study Location Characteristics

Toronto, Ontario is the largest and most populous city in Canada with a population of 2.5 million people (Statistics Canada, 2007a). The city is situated on Lake Ontario and surrounded by four regional municipalities. Combined with the municipalities that make up the Greater Toronto Area, the population of the region rises to 5.5 million people (Statistics Canada, 2007a). Downtown Toronto has a well-developed public transportation system comprised of subways, streetcars (trams), and buses. In the densely populated core, 65 to 80% of all residents walk, cycle, or use public transportation to commute to their workplaces on a regular basis (Statistics Canada, 2007a). In the less densely populated areas of Toronto and within the regional municipalities, public transportation is less accessible and more than 80% of all residents use their personal vehicles to commute to their workplaces (Statistics Canada, 2007a). Recently, the Greater Toronto Area was identified as having the longest average commute time in North America (Toronto Board of Trade, 2011b).

3.4.2 Sampling Location

Continuous aerosol measurements were collected at the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR) located in downtown Toronto (Figure 3-1). SOCAAR is located on the north side of a four-lane major arterial roadway with traffic volumes ranging from 15,000 to 20,000 vehicles per day and vehicle speeds up to 50km/h. SOCAAR’s location can be described as a quasi-street canyon due to the presence of several two to four storey buildings that are taller than the width of the road. Under normal meteorological conditions, the potential for accumulation of vehicle emissions is limited by the site’s proximity to an intersection controlled with a traffic light. The SOCAAR sampling inlet is located 20m from College St. and 6m aboveground. A second monitoring site at the Gage Building located on the south side of College St. 150m to the west of SOCAAR provided traffic volume data.
Figure 3-1: The upper panel (A) shows a map of the sampling locations in Downtown Toronto. The distance between the SOCAAR Site and the Gage Site is 150m. The distance between the SOCAAR Site and the Ontario Ministry of the Environment (MOE) Site is 750m. The lower panel (B) shows the location of the SOCAAR lab within the quasi-street canyon. Satellite and aerial imagery is courtesy of Bing Maps.

3.4.3 Instrumentation

A summary of the available data sources is presented in Table 3-1. The PN concentration and particle size measurements were collected using a Fast Mobility Particle Sizer (FMPS, Model 3091, TSI Incorporated, Shoreview, MN, USA) between March 2006 and May 2011. The FMPS provided the number-based size distribution over the range of 5.6 to 560nm with 1-second time resolution. One minute averages were extracted using the FMPS Software Version 3.1.0.0. Hourly averages were calculated for each hour with a minimum of 70% of data available (42 minutes). The FMPS had low counting efficiency and frequent zero values for particles less than 8nm and greater than 300nm. As a result, only particles with diameters between 8 and 300nm were included in this study. The particle size distribution was corrected using the procedure described by Jeong and Evans (2009). The resulting size distribution was similar to those
obtained using a Scanning Mobility Particle Sizer (SMPS). The zero-reading on the FMPS was checked on a weekly basis and the electrometer column was cleaned on a monthly basis. Comparisons between the FMPS and SMPS were performed at times when the two instruments were collocated.

Table 3-1 Summarized data sources and availability. All data were sampled over one-hour time intervals and valid N refers to the number of hourly observations available between the start and end dates. The MOE site refers to the Ontario Ministry of the Environment’s Downtown Toronto monitoring site. The YYZMS refers to the meteorological station operated by Environment Canada located at Pearson International Airport. The UTMMS refers to the University of Toronto Mississauga Campus’ Meteorological Station.

<table>
<thead>
<tr>
<th>Data Source/Location</th>
<th>Measured Data</th>
<th>Start Date</th>
<th>End Date</th>
<th>Valid N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOCAAR Site</td>
<td>PN</td>
<td>18/03/2006</td>
<td>31/05/2011</td>
<td>36252 (79)</td>
</tr>
<tr>
<td>Gage Site</td>
<td>Traffic volume</td>
<td>19/06/2010</td>
<td>31/12/2010</td>
<td>4530 (97)</td>
</tr>
<tr>
<td>MOE Site</td>
<td>PM$_{2.5}$ mass</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>69610 (99)</td>
</tr>
<tr>
<td></td>
<td>CO</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>65200 (93)</td>
</tr>
<tr>
<td></td>
<td>NO and NO$_2$,</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>69589 (99)</td>
</tr>
<tr>
<td></td>
<td>O$_3$</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>69540 (99)</td>
</tr>
<tr>
<td></td>
<td>SO$_2$</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>68954 (98)</td>
</tr>
<tr>
<td>YYZMS</td>
<td>T</td>
<td>01/01/2003</td>
<td>31/05/2011</td>
<td>73806 (100)</td>
</tr>
<tr>
<td></td>
<td>RH</td>
<td>01/01/2003</td>
<td>31/05/2011</td>
<td>73806 (100)</td>
</tr>
<tr>
<td></td>
<td>WS</td>
<td>01/01/2003</td>
<td>31/05/2011</td>
<td>73806 (100)</td>
</tr>
<tr>
<td></td>
<td>WDir</td>
<td>01/01/2003</td>
<td>31/05/2011</td>
<td>72224 (98)</td>
</tr>
<tr>
<td>UTMMS</td>
<td>SR</td>
<td>01/01/2003</td>
<td>31/12/2010</td>
<td>67802 (97)</td>
</tr>
</tbody>
</table>

The SmartEye Traffic Data Sensor (TDS) was used to collect traffic counts along College St. between June and December 2010. The TDS uses edge-based scene information to count passenger cars and heavy-duty diesel-powered trucks across four lanes. Although the TDS was not available during the entire measurement period, the data collected is believed to be representative of the daily and seasonal traffic patterns on College St. The TDS is mounted on the roof of the four-storey building (Gage Site) 150m to the west of the SOCAAR sampling site (Figure 3-1).
Additional long-term air quality data was obtained from the Ontario Ministry of the Environment (MOE) Air Quality Network (MOE, 2010). The downtown Toronto MOE site is located 750m to the northeast of the SOCAAR site and provided hourly average concentrations for nitric oxide (NO), nitrogen dioxide (NO₂), sulphur dioxide (SO₂), carbon monoxide (CO), ozone (O₃) and fine particulate matter (PM₂.₅).

The meteorological station located at Pearson International Airport (YYZMS) provided hourly temperature (T), relative humidity (RH), wind speed (WS), wind direction (WDir), visibility and cloud cover data (Environment Canada, 2011). The WS and WD data obtained from Pearson International Airport are believed to reflect the macro-scale movement of air through Toronto. Solar radiation (SR) data was obtained from the University of Toronto Mississauga campus’ meteorological station (UTMMS) located 25km west of the SOCAAR site (UTMMS, 2010). The SR measurements from the UTMMS were the most continuous data available in the area and exhibited good correlation with a second SR measurement site operated on an intermittent basis at the Gage site in downtown Toronto.

### 3.5 Results and Discussion

#### 3.5.1 Meteorology

Toronto experienced monthly average ambient temperatures ranging from -4.6°C in February to 21.5°C in July. The relative humidity was between 59% in April and 77% in November. Higher wind speeds occurred in January-March (5.1m/s) compared to July-September (3.7m/s) of each year. Winds from north-easterly directions were relatively rare. Large year-to-year changes in each meteorological parameter were not observed. The dominant wind directions during the five-year measurement period were from the north-west, south-west and south-east directions (Figure 3-2). The highest wind speeds were observed from westerly directions. The north regions of Toronto are sparsely populated with few large industrial facilities; therefore, air masses from northerly directions bring clean air into the city. Air masses from the south traveled over the United States and often contained higher concentrations of PM₂.₅, and SO₂ due to the presence of large coal-powered electricity generation facilities and other industrial facilities.
Figure 3-2 Wind speed and wind direction recorded at Pearson International Airport between March 2006 and May 2011.

3.5.2 Daily Traffic Variation

Figure 3-3 shows the average daily traffic patterns for cars and trucks on weekdays and weekends. Downtown Toronto experienced an increase in vehicular traffic between 7:00 and 9:00 in the morning on weekdays. On weekends, the traffic volume is significantly lower, and the daily maximum occurs later in the day. Higher weekend traffic counts between midnight and 4:00 were observed due to the presence of several restaurants, bars, and clubs south of the sampling location. Maximum truck traffic occurred overnight and decreased significantly during periods with high car traffic. On weekdays, 19,100 cars and 1200 trucks pass the sampling location each day. On weekends, 15,300 cars and 1200 trucks pass the sampling location.
Figure 3-3 Traffic variation on College St. Data was collected by the SmartEye Traffic Sensor from the roof of the Gage Institute building between June and December 2010.

3.5.3 Long-Term and Seasonal Behaviour of Ultrafine Particles

The long-term PN concentration behaviour in downtown Toronto is presented in Figure 3-4. Each observation is a monthly average concentration ± 95% confidence interval. The average concentration for all size fractions decreased between March 2006 and May 2011. Particles less than 50nm (PN$_{50}$) decreased by 21% between 2006 and 2010. Notably, the majority of the reduction occurred during the winter months with no statistically significant trend observed during the summer months (Table 3-2). PN$_{50-100}$ (number of particles between 50 and 100nm) and PN$_{100-300}$ (number of particles between 100 and 300nm) decreased by 17% and 24%, respectively. No year-to-year trend was observed for the geometric mean diameter (GMD). PN$_{50}$ exhibited strong seasonality with maximums occurring between January and March of each year (Figure 3-5). PN$_{50-100}$ did not exhibit significant seasonal variation. Higher concentrations of PN$_{100-300}$ were observed during July and August due to increased ambient temperatures and
photochemical production of particle-phase sulphate. The GMD was smallest during the winter months due to increased PN$_{50}$ concentrations.

Figure 3-4 Long-term behaviour of UFP. Each value is an average ± 95% confidence interval. All size fractions of UFP decreased during the measurement period.
Table 3-2 Average ± 95% confidence interval for PN$_{50}$, PN$_{50-100}$ and PN$_{100-300}$ on a seasonal basis. Winter = January to March, Spring = April to June, Summer = July to September, Fall = October to December. Unless otherwise stated, all reported averages have N>1500 hours of observation. * denotes averages with 1000<N<1500 hours of observation and † denotes seasons with N<1000 hours of observation.

<table>
<thead>
<tr>
<th>Size Fraction</th>
<th>Season</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(cm$^{-3}$)</td>
<td>(cm$^{-3}$)</td>
<td>(cm$^{-3}$)</td>
<td>(cm$^{-3}$)</td>
<td>(cm$^{-3}$)</td>
<td>(cm$^{-3}$)</td>
</tr>
<tr>
<td>PN$_{50}$</td>
<td>Winter</td>
<td>27150 ± 1510†</td>
<td>23230 ± 700</td>
<td>18870 ± 560</td>
<td>18270 ± 500</td>
<td>14960 ± 580*</td>
<td>18160 ± 490</td>
</tr>
<tr>
<td></td>
<td>Spring</td>
<td>19140 ± 510</td>
<td>15900 ± 520*</td>
<td>14000 ± 430*</td>
<td>12160 ± 340</td>
<td>12300 ± 440*</td>
<td>13110 ± 520</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>12480 ± 330</td>
<td>10900 ± 500*</td>
<td>10630 ± 350*</td>
<td>10570 ± 290</td>
<td>9990 ± 420†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td>15290 ± 450</td>
<td>15940 ± 460</td>
<td>15140 ± 410</td>
<td>13140 ± 390</td>
<td>14680 ± 430</td>
<td></td>
</tr>
<tr>
<td>PN$_{50-100}$</td>
<td>Winter</td>
<td>4040 ± 290†</td>
<td>4040 ± 140</td>
<td>3970 ± 120</td>
<td>3360 ± 90</td>
<td>2270 ± 80*</td>
<td>3190 ± 100</td>
</tr>
<tr>
<td></td>
<td>Spring</td>
<td>3620 ± 120</td>
<td>3170 ± 120*</td>
<td>3210 ± 110*</td>
<td>2650 ± 80</td>
<td>3100 ± 100*</td>
<td>2430 ± 100</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>3610 ± 90</td>
<td>3620 ± 110*</td>
<td>3200 ± 90*</td>
<td>3020 ± 70</td>
<td>3170 ± 120†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td>3370 ± 100</td>
<td>3520 ± 110</td>
<td>2980 ± 80</td>
<td>2540 ± 80</td>
<td>2740 ± 80</td>
<td></td>
</tr>
<tr>
<td>PN$_{100-300}$</td>
<td>Winter</td>
<td>1420 ± 140†</td>
<td>1060 ± 40</td>
<td>1150 ± 40</td>
<td>1060 ± 30</td>
<td>800 ± 30*</td>
<td>860 ± 30</td>
</tr>
<tr>
<td></td>
<td>Spring</td>
<td>1290 ± 60</td>
<td>840 ± 40*</td>
<td>980 ± 40*</td>
<td>750 ± 20</td>
<td>930 ± 30*</td>
<td>700 ± 30</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>1540 ± 60</td>
<td>1530 ± 60*</td>
<td>1210 ± 50*</td>
<td>920 ± 30</td>
<td>1240 ± 50†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td>1130 ± 40</td>
<td>1090 ± 40</td>
<td>940 ± 30</td>
<td>850 ± 30</td>
<td>780 ± 20</td>
<td></td>
</tr>
</tbody>
</table>
The PN$_{50}$ size fraction exhibited the strongest wind directionality (Figure 3-6). The highest PN$_{50}$ concentrations were observed for northerly wind directions during the winter (January to March). Northerly winds were associated with lower ambient temperatures. Previous studies in Toronto have observed enhanced traffic-related particle formation as cooler air parcels enter the city (Jeong et al., 2006). In addition, the SOCAAR site is located on the north side of a quasi-street canyon. When northerly winds hit the taller building immediately opposite the SOCAAR site, it
can create a circulation air flow in the opposite direction at street level, from south to north. Thus air from the roadway is drawn into the sampling inlet. Lower PN$_{50}$ concentrations are observed from southerly wind directions. Higher PN$_{50-100}$ and PN$_{100-300}$ concentrations were observed from southerly wind directions suggesting that particles in these size fractions originated from different sources than PN$_{50}$. Further, these PN$_{50-100}$ and PN$_{100-300}$ may have scavenged PN$_{50}$ particles, or their precursor gases, contributing to the lower PN$_{50}$ concentrations when winds were from the south.

![Figure 3-6 Wind directionality of UFP. PN$_{50}$ was highest from northerly directions. PN$_{100-300}$ was highest from southerly wind directions.](image)

Few studies have reported continuous long-term UFP measurements in urban areas. Although several studies have examined the changes in PN concentrations over time horizons of less than two years, the effect of public policy initiatives such as changing transportation fuel composition, sources of electricity, and vehicle fleet characteristics on the PN concentration
cannot be identified during short monitoring periods. Recently, Wang et al. (2011) observed a reduction in the PN concentrations over an eight-year period at a site in Rochester N.Y. (located approximately 150km to the south-east of Toronto). Factors influencing the observed ~50% reduction of nuclei-mode particles included the closure of a large coal-fired power plant, reduction of emissions from an industrial facility and the transition from low-sulphur to ultra-low-sulphur diesel. In another investigation in Copenhagen, Denmark, a significant reduction of 27% in nuclei-mode particles was observed over a period of six years. The most pronounced reduction in PN concentration occurred after the transition to ultra-low-sulphur diesel and the introduction of modern diesel cars with oxidation catalysts (Wåhlin, 2009). Mejía et al. (2007) reported an increase in PN concentration in Brisbane, Australia, during a five-year measurement period due to increased population and vehicle ownership during the period. Finally, Hussein et al., (2004) reported a decrease of 17% during a six year measurement period in Helsinki, Finland, due to the changes in traffic density and the adoption of newer vehicles. The observed reduction in PN concentration in Toronto is attributable to improved emission control technology on gasoline-powered vehicles as the new NOx standard was implemented across Canada’s vehicle fleet, and reduced power output at the Nanticoke Electricity Generating Facility located ~100 km southwest of the monitoring site and several other coal-fired power plants located in the United States. In addition, between 2006 and 2009, the number of vehicles driven in Ontario increased by 5% (Statistics Canada, 2010). Therefore, the observed reductions in UFP occurred despite more vehicles being on the road. The transition to ultra-low sulphur diesel (maximum sulphur concentration = 15 mg/kg) occurred in the middle of 2006 and PN concentration data is unavailable for the years prior to the fuel change (Environment Canada, 2010).

UFP studies in other urban areas have reported a wide range of PN concentrations due to differences in meteorology, emission profiles (i.e. fleets dominated by diesel or gasoline powered vehicles), distance from the roadway, traffic density, population, land-use and instrumentation. Table 3-3 summarizes the range of concentrations observed in several cities across North America and Europe. The included studies best match Toronto’s climate, the urban sampling site and the size fractions of UFP examined in this investigation. In several European countries, a significant fraction of the passenger vehicle fleet is diesel-powered. Therefore cities with relatively small populations have PN concentrations that are similar in magnitude to
Toronto. Higher PN concentrations were consistently observed during the coldest periods of the year (Puustinen et al., 2007; Wang et al., 2011b) and regardless of location, the PN concentration increased between 7:00 and 10:00 on working days (Puustinen et al., 2007; Wang et al., 2011b).

Other investigations have identified several ambient temperature-related factors that may influence the seasonality of UFP (Aalto et al., 2005; Cyrys et al., 2008; Holmes et al., 2005; Ketzel et al., 2004a; Jeong et al., 2004; 2006): atmospheric mixing, driving patterns, residential heating patterns, the efficiency of vehicle emission control technologies, evaporation and condensation of volatile organic species, photochemistry, and ambient PM$_{2.5}$ concentrations. During the winter months, cooler ambient temperatures favour the formation of nuclei-mode particles in vehicle exhaust. Further, volatile organic compounds that would have remained in the gas phase during periods with warmer ambient temperatures are now available to condense onto the surface of the particles. Lower ambient temperatures can result less evaporation and photochemistry, thereby, increasing the atmospheric lifetime of the particles. Thus, the winter period is more conducive for nuclei-mode particle formation. All of the abovementioned factors are believed to influence the formation and fate of PN$_{50}$ in this study.

The distance between the instrumentation and traffic source has remained constant throughout the measurement period. The rate at which UFP decreases from the roadway has exhibited seasonal differences in previous studies (Zhu et al., 2002a; 2002b; 2004). Zhu et al., (2002a, 2002b) studied the transformation of UFP near the I-405 and I-710 highways during the summer. In both studies, particles in smallest size fractions decreased with increasing distance from the road at a much faster rate than the largest particles. Zhu et al., (2004) repeated the summertime measurements during the cooler winter months at the same sampling locations near the I-405 and I-710 highways. Although the observed decay pattern of black carbon and CO were similar between the two seasons, the particle size distribution behaved differently. During the winter study, more particles with diameters less than 50nm were observed at distances further away from the highway than during the summer suggesting that the atmospheric lifetime of UFP may be different between seasons.
Table 3-3 Summary of recent UFP results from urban areas. Included studies had a minimum of one month of continuous data, instrumentation with a diameter size cut-off of 7-10nm, and a sampling location with a climate similar to Toronto, Canada. * Passenger vehicle dieselization is expressed as a percentage of diesel vehicles in the total passenger vehicle fleet as of 2009 (EEA, 2009). † The average PN concentration was calculated from multiple annual values reported in the original study. ‡ The average PN concentration was estimated from a figure.

<table>
<thead>
<tr>
<th>Location</th>
<th>Population</th>
<th>Passenger Fleet %</th>
<th>Period</th>
<th>Size Range (nm)</th>
<th>PN (cm(^{-3}))</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helsinki, Finland</td>
<td>0.6 million</td>
<td>18</td>
<td>1998-2002 Oct 2002-Mar</td>
<td>8-400</td>
<td>16,310†</td>
<td>Hussein et al., 2004</td>
</tr>
<tr>
<td>Helsinki, Finland</td>
<td>0.6 million</td>
<td>18</td>
<td>2004 Oct 2002-Mar</td>
<td>7-3000</td>
<td>12,490</td>
<td>Puustinen et al., 2007</td>
</tr>
<tr>
<td>Birmingham, UK</td>
<td>1.0 million</td>
<td>24</td>
<td>2004 Oct 2002-Mar</td>
<td>7-3000</td>
<td>18,790</td>
<td>Puustinen et al., 2007</td>
</tr>
<tr>
<td>Amsterdam, Netherlands</td>
<td>1.2 million</td>
<td>24</td>
<td>2004 Oct 2002-Mar</td>
<td>7-3000</td>
<td>18,090</td>
<td>Puustinen et al., 2007</td>
</tr>
<tr>
<td>Copenhagen, Denmark</td>
<td>1.2 million</td>
<td>19</td>
<td>2002-2007</td>
<td>7-3000</td>
<td>26,490†</td>
<td>Wählin, 2009</td>
</tr>
<tr>
<td>Rochester, NY, US</td>
<td>0.2 million</td>
<td>&lt; 5</td>
<td>2002-2009</td>
<td>10-500</td>
<td>7,640††</td>
<td>Wang et al., 2011b</td>
</tr>
<tr>
<td>New York, NY, US</td>
<td>8.2 million</td>
<td>&lt; 5</td>
<td>Jan-Feb 2004</td>
<td>8-284</td>
<td>66,760</td>
<td>Bae et al., 2010</td>
</tr>
<tr>
<td>Toronto, Canada</td>
<td>2.5 million</td>
<td>&lt; 5</td>
<td>2006-2011</td>
<td>8-300</td>
<td>19,440†</td>
<td>This Study</td>
</tr>
</tbody>
</table>
3.5.4 Diurnal Behaviour

Each size fraction of UFP exhibited strong diurnal behaviour. In Figure 3-7, the diurnal behaviour on weekdays and weekends are compared. On each weekday, the initial increase between 6:00 and 9:00 in the morning corresponded with the morning rush hour period. On weekends, the PN concentration increased later in the morning and followed a similar pattern to the traffic profile on weekends. The change in geometric diameter from larger to smaller particles between 6:00 and 9:00 in the morning on weekdays was not observed at the same time on weekends suggesting that increased vehicular traffic is the source of UFP rather than some other photochemical process that occurs as solar intensity increases. Higher PN$_{50}$ and PN$_{50-100}$ concentrations were observed during the overnight period on weekends due to the presence of numerous restaurants, bars, and nightclubs located south of the sampling site.

Although a significant drop in traffic is known to occur between midnight and 6:00, and likewise, an increase in traffic during the morning rush hour period on weekdays, the observed PN concentration does not drop and increase by the same amount. During the overnight period, the PN$_{50}$ concentration decreased by 30% while traffic decreased by 73%. Between 6:00 and 9:00 on weekdays, the PN$_{50}$ concentration increased 2.5 times while the traffic volume increased by a factor of 4.6. During follow-up measurements conducted by Zhu et al. (2006) near the I-405 highway at night, the PN concentration decreased exponentially from the highway but at a much slower rate than during the day. Further, the night-time PN concentration on the road was 80% of the daytime value, which suggests that the colder ambient temperatures may increase the formation of nuclei-mode particles in vehicle exhaust (Zhu et al., 2006). Hu et al. (2009) measured the PN concentration near the I-10 highway in Santa Monica, California at night during the winter. Significant differences between the daytime and night-time decay of UFP were observed. In this study, the distance between the roadway and instrumentation remained constant and only the contribution of individual vehicles to the overall PN concentration differed between day and night. Lower ambient temperatures, increased condensation of semi-volatile species, and increased atmospheric stability at night may have increased the atmospheric lifetime of UFP. The higher number of trucks at night would also have played a role.
The diurnal trend from the winter months was compared to the diurnal trend from the summer months for all size fractions (Figure 3-8). Explanations relating to the observed seasonal differences in UFP concentrations often focus on the effect of mixing and photochemistry. In this study, a ratio of 1.02 ± 0.03 was observed for PN$_{50-100}$ between the summer and winter months. Particles in this fraction mostly consist of carbonaceous soot formed inside of diesel engines (Ban-Weiss et al., 2010). Therefore, ambient temperature has limited impact on their formation. Further, any seasonal differences observed between the roadside and the SOCAAR laboratory would be impacted by atmospheric mixing and vehicle-induced turbulence. Since the winter to summer ratio was close to 1.0 for PN$_{50-100}$, the air dynamics within the quasi-street...
canyon was similar in both seasons. In contrast, the observed average winter to summer ratio for PN$_{50}$ was $1.75 \pm 0.05$. Thus it appears that the formation of PN$_{50}$ was favoured, or the atmospheric lifetime of PN$_{50}$ increased during the colder period, even though the emission rate from vehicles, and mixing was apparently similar between the seasons. The average temperature and daytime-solar-radiation differentials between the summer and winter months were 22°C (winter = -2°C, summer = 20°C) and 110 W/m$^2$ (winter = 198 W/m$^2$, summer = 308 W/m$^2$). The relative impacts of colder temperatures on PN$_{50}$ formation and lifetime remains unknown but both are related to faster condensation and or slower evaporation during the winter. Finally, the winter to summer ratio for PN$_{100-300}$ was below 1.0. This difference was attributed to reduced photochemistry during the winter months resulting in lower concentrations of this pollutant.

Figure 3-8 Ratio of winter PN concentration and Summer PN concentration. The ratio of PN$_{50}$ was significantly higher than 1.0. The error bars are 95% confidence intervals.

3.5.5 Particle Size Distribution
The evolution of the particle size distribution is shown in Figure 3-9. The size distribution on weekdays and weekends were similar in shape but lower in overall number (not shown). The
modal diameter observed during the winter months was similar at all times of the day (19.1nm). During the summer months, the modal diameter decreased from 39.2nm during the overnight period to 16.5nm during the morning, and increased throughout the afternoon (22.1nm) and evening periods (29.4nm).

Figure 3-9: Evolution of the particle size distribution on weekdays during the summer and winter. Overnight: 0:00-5:00, Morning: 6:00-11:00, Afternoon: 12:00-17:00, and Evening: 18:00-23:00.

3.5.6 Effect of Changing Vehicle Fleet on UFP Emissions
In 2007, the Tier II NO\textsubscript{x} standard of 0.07g/mile was implemented in Canada’s vehicle fleet (Environment Canada, 2009). Vehicles that meet this standard increased from 20% of the fleet in 2006 to nearly 40% in 2009 (Statistics Canada, 2007b; 2010). Figure 3-10 compares the hours of 7:00 and 9:00 on weekdays and weekends; traffic volume is ~50% lower on weekends. Figure 3-10 shows a substantial decrease in PN\textsubscript{50} concentrations occurred during the morning rush hour period with little change on weekends. This suggests that the traffic emission or formation processes for PN\textsubscript{50} was affected by a vehicle fleet-related change. Thus the reduction in PN\textsubscript{50} was likely an indirect benefit of Canadian policies introduced to reduce vehicle
emissions. If so, similar reductions are occurring in other Canadian cities. Regrettably, long-term near-road UFP data does not yet exist to explore this hypothesis. Smaller differences between weekdays and weekends were observed for PN_{50-100} which suggests the presence of a regional source of particles in this size fraction in addition to a local traffic source. Finally, the year-on-year changes for PN_{100-300} were similar on weekdays and weekends suggesting that the source of this pollutant was predominantly regional.

Figure 3-10 Average PN concentration between 7:00 and 9:00 on weekdays and weekends between 2006 and 2010. Error bars are 95% confidence intervals.
3.5.7 Correlation with Criteria Pollutants and Meteorological Parameters

Table 3-4 shows the Spearman’s rank correlation coefficients between the 24-hour averaged PN concentration, criteria pollutants concentrations, and meteorological parameters. The strongest relationships for size fractions less than 50nm were between T, RH, NO, and NO₂. SR and SO₂ were not correlated with PN₁₀ and PN₁₀₋₂₀ suggesting that the temperature and NO₃ may be better predictors of these size fractions of UFP in cities. Size fractions between 50 and 100nm exhibited stronger correlation with combustion-related pollutants, PM₂.₅, and WS, and limited correlation with T and RH. The positive correlation between particles size fractions between 50 and 100nm with SO₂ suggests that particles within this size fraction originate from long-range transport in addition to local combustion. As discussed above, the dominant source of NO₂ in Toronto is vehicular traffic with no significant source of SO₂ in Toronto.

Table 3-4 Spearman Rank Order correlation coefficients of 24-hour averaged PN concentrations, criteria pollutant concentrations, and meteorological conditions. * Daytime solar radiation (SR) values were used in the correlation matrix. All values are significant to the 95% level of confidence. Case-wise deletion was used to handle blank or erroneous data.  T = temperature, RH = relative humidity, WS = wind speed, N.S. = not statistically significant.

<table>
<thead>
<tr>
<th>PN</th>
<th>PM₂.₅</th>
<th>CO</th>
<th>NO</th>
<th>NO₂</th>
<th>O₃</th>
<th>SO₂</th>
<th>T</th>
<th>RH</th>
<th>WS</th>
<th>SR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PN₁₀</td>
<td>-0.29</td>
<td>-0.07</td>
<td>0.34</td>
<td>0.35</td>
<td>0.22</td>
<td>N.S.</td>
<td>-0.53</td>
<td>-0.28</td>
<td>0.18</td>
<td>N.S.</td>
</tr>
<tr>
<td>PN₁₀₋₂₀</td>
<td>-0.22</td>
<td>N.S.</td>
<td>0.43</td>
<td>0.45</td>
<td>-0.26</td>
<td>N.S.</td>
<td>-0.54</td>
<td>-0.23</td>
<td>0.09</td>
<td>N.S.</td>
</tr>
<tr>
<td>PN₂₀₋₃₀</td>
<td>-0.10</td>
<td>0.06</td>
<td>0.52</td>
<td>0.55</td>
<td>-0.26</td>
<td>0.16</td>
<td>-0.46</td>
<td>-0.23</td>
<td>0.05</td>
<td>N.S.</td>
</tr>
<tr>
<td>PN₃₀₋₄₀</td>
<td>0.18</td>
<td>0.14</td>
<td>0.58</td>
<td>0.65</td>
<td>-0.21</td>
<td>0.35</td>
<td>-0.24</td>
<td>-0.2</td>
<td>0.29</td>
<td>0.06</td>
</tr>
<tr>
<td>PN₄₀₋₅₀</td>
<td>0.39</td>
<td>0.20</td>
<td>0.58</td>
<td>0.68</td>
<td>-0.16</td>
<td>0.46</td>
<td>-0.06</td>
<td>-0.14</td>
<td>0.43</td>
<td>0.09</td>
</tr>
<tr>
<td>PN₅₀₋₆₀</td>
<td>0.52</td>
<td>0.23</td>
<td>0.56</td>
<td>0.68</td>
<td>-0.13</td>
<td>0.52</td>
<td>0.06</td>
<td>-0.09</td>
<td>0.50</td>
<td>0.09</td>
</tr>
<tr>
<td>PN₆₀₋₇₀</td>
<td>0.67</td>
<td>0.26</td>
<td>0.49</td>
<td>0.64</td>
<td>-0.07</td>
<td>0.57</td>
<td>0.19</td>
<td>N.S.</td>
<td>-0.54</td>
<td>0.08</td>
</tr>
<tr>
<td>PN₇₀₋₁₀₀</td>
<td>0.79</td>
<td>0.29</td>
<td>0.39</td>
<td>0.57</td>
<td>N.S.</td>
<td>0.59</td>
<td>0.31</td>
<td>0.07</td>
<td>-0.52</td>
<td>0.06</td>
</tr>
<tr>
<td>PN₁₀₀₋₂₀₀</td>
<td>0.88</td>
<td>0.32</td>
<td>0.28</td>
<td>0.48</td>
<td>N.S.</td>
<td>0.58</td>
<td>0.35</td>
<td>0.18</td>
<td>-0.44</td>
<td>N.S.</td>
</tr>
<tr>
<td>PN₂₀₀₋₃₀₀</td>
<td>0.78</td>
<td>0.33</td>
<td>0.17</td>
<td>0.36</td>
<td>N.S.</td>
<td>0.44</td>
<td>0.25</td>
<td>0.25</td>
<td>-0.33</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

The variation of the Spearman rank order correlation coefficient was explored on a seasonal basis (Figure 3-11). For PN₅₀₋₁₀₀ and PN₁₀₀₋₃₀₀, the relationship between NO₂, SO₂ and PM₂.₅ was more consistent throughout the year than PN₅₀; however, a drop in overall correlation was observed during the summer months for NO₂ and SO₂. During the summer months, increased
photochemistry enhances the production of particle phase sulphate from SO$_2$ and the oxidation of NO$_2$. Although increased evaporation of volatile species affects the concentration of PN$_{50}$ during the summer months, the effect of evaporation on the particle size distribution is independent of the photochemical reactions impacting the NO$_2$ and SO$_2$ concentrations. A small drop in the correlation between all PN concentration size fractions and NO$_2$ was observed during the month of December. This may be attributable to changing driving patterns around the holiday season.

Figure 3-11 Variation in the Spearman Rank Order Correlation Coefficient between the size resolved PN concentrations and NO$_2$, SO$_2$, and PM$_{2.5}$ concentrations. The Spearman Rank Correlation Coefficient was calculated by grouping the 24-hour average values for each pollutant by month
3.5.8 Predicting UFP Concentrations Using a Multiple Linear Regression Model

The criteria pollutants and meteorological parameters have been continuously monitored since January 2003 and the PN concentration was measured continuously since March 2006. To capitalize on the existing long-term monitoring of the criteria pollutants and meteorological parameters, a multiple linear regression (MLR) model was developed to extrapolate the PN concentration back to 2003. Daily-averaged values of PN concentration, criteria pollutant concentrations (except CO), and meteorological parameters were log-normally distributed and log-transformed prior to performing the regression. CO was normally distributed and was not transformed. Data from alternating days between 2006 and 2010 were used in the regression model and the excluded data were used for validation purposes. Table 3-5 shows the statistically significant beta ± standard error of each regression coefficient. For the largest size fractions, the PM$_{2.5}$ concentration was the strongest predictor. In contrast, the strongest predictors for the smallest size fraction of UFP were meteorological parameters and NO$_2$.

The MLR model was validated by comparing the measured excluded PN concentration measurements with the corresponding predicted PN concentrations (Figure 3-12). The MLR model was able to explain 52% of the day-to-day variability observed in the PN$_{50}$ concentration. This suggests that the PN$_{50}$ concentration can be extrapolated using the MLR; however, the results must be interpreted with caution. The MLR model better predicted PN$_{50-100}$ ($R^2 = 0.63$) and PN$_{100-300}$ ($R^2 = 0.82$), allowing greater confidence in these estimates.

No continuous PN concentration monitoring occurred in downtown Toronto between 2003 and 2005. The results presented in Figure 3-13 compare the predicted and measured annual average PN concentration between 2006 and 2010 (for the days excluded when creating the MLR) and show the extrapolated annual average PN concentration back to 2003. The predicted annual averages were within the 95% confidence intervals of the measured PN concentrations for all size fractions between 2006 and 2010. The modelled PN concentrations suggest that a continuous decrease in the PN concentration has occurred from at least 2003 in Toronto with PN$_{50-100}$ and PN$_{100-300}$ showing a greater percentage decrease than PN$_{50}$. The highest PN concentration typically occurs between January and March. Since monitoring for PN...
concentration began in March 2006, the annual average PN$_{50}$ concentration for 2006 was likely slightly higher than reported in Figure 3-13.

Figure 3-12 Comparison between the measured and predicted PN concentration. Each observation is a 24-hour average value.
Figure 3-13 Results from the multiple linear regression model. Each observation is an annual average ± 95% confidence interval. The measured PN concentrations used in this figure were not included in the generation of the MLR model. The predictive power of the model was highest for PN\textsubscript{50-100} and PN\textsubscript{100-300}.

To validate the modelled annual averages, the 2003 modelled values were compared to the total number concentration obtained from a CPC (model 3020, TSI Inc.) during a year-long study of UFP behaviour at this site (Jeong et al., 2006). The 2003 measured annual PN concentration was 28,010 cm\textsuperscript{-3}. The CPC had a larger size range (7 to 1000nm) and different efficiency than the FMPS used in this study (8 to 300nm), and these differences were corrected for. Jeong and Evans (2009) compared the FMPS to a UWCPC (model 3786, TSI Inc.) with a size range between 2.5 and 1000nm. The resulting ratio between the two instruments was FMPS/UWCPC...
Previous measurements with a Scanning Mobility Particle Sizer with a nano-differential mobility analyzer suggest that at this site, particles between 3 and 7nm contribute ~4% to the total number concentration. Thus, the estimated FMPS to CPC (Model 3020) ratio was 0.83. The adjusted 2003 measured PN concentration is 23,250 pt/cm$^3$ and is within 3% of the modelled total number concentration of 23,950 cm$^{-3}$. Figure 3-14 shows the seasonal variation of the measured and modelled PN concentration. The modelled PN concentration was within the 95% confidence intervals for all measured monthly averages.

Figure 3-14 Monthly average PN concentration values between 2003 and 2010. Between 2006 and 2010, the measured averages were calculated from data not included in the generation of the MLR model. All predicted values were within the 95% confidence interval of the measured data (error bars not shown for clarity)
Table 3-5 Beta values ± standard error for the multiple linear regression model. All variables except CO were log-normally distributed and were log-transformed prior to performing the regression model. CO was normally distributed and was not transformed. * Daytime values were used in the daily average calculation of solar radiation (SR). For the smallest particles (diameters less than 50nm), the largest beta values were for meteorological conditions and NO\textsubscript{2}. For the larger particles, the largest beta values corresponded with PM\textsubscript{2.5}. T= temperature, RH = relative humidity, WS = wind speed.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>PN\textsubscript{10}</th>
<th>PN\textsubscript{10-20}</th>
<th>PN\textsubscript{20-30}</th>
<th>PN\textsubscript{30-40}</th>
<th>PN\textsubscript{40-50}</th>
<th>PN\textsubscript{50-60}</th>
<th>PN\textsubscript{60-70}</th>
<th>PN\textsubscript{70-100}</th>
<th>PN\textsubscript{100-200}</th>
<th>PN\textsubscript{200-300}</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM\textsubscript{2.5}</td>
<td>-0.22 ± 0.04</td>
<td>-0.21 ± 0.04</td>
<td>-0.16 ± 0.04</td>
<td>-</td>
<td>-</td>
<td>0.20 ± 0.03</td>
<td>0.26 ± 0.04</td>
<td>0.42 ± 0.03</td>
<td>0.59 ± 0.03</td>
<td>0.67 ± 0.03</td>
</tr>
<tr>
<td>CO</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NO</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.14 ± 0.04</td>
</tr>
<tr>
<td>NO\textsubscript{2}</td>
<td>0.55 ± 0.05</td>
<td>0.61 ± 0.04</td>
<td>0.66 ± 0.04</td>
<td>0.56 ± 0.04</td>
<td>0.59 ± 0.03</td>
<td>0.50 ± 0.03</td>
<td>0.49 ± 0.04</td>
<td>0.30 ± 0.04</td>
<td>0.33 ± 0.03</td>
<td>-0.14 ± 0.04</td>
</tr>
<tr>
<td>O\textsubscript{3}</td>
<td>-0.19 ± 0.04</td>
<td>-0.19 ± 0.04</td>
<td>-0.19 ± 0.04</td>
<td>-0.18 ± 0.04</td>
<td>-0.16 ± 0.03</td>
<td>-0.17 ± 0.03</td>
<td>-0.15 ± 0.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SO\textsubscript{2}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.11 ± 0.03</td>
<td>0.20 ± 0.03</td>
<td>0.19 ± 0.03</td>
<td>0.21 ± 0.03</td>
<td>0.20 ± 0.03</td>
<td>0.16 ± 0.02</td>
<td>-</td>
</tr>
<tr>
<td>T</td>
<td>-0.29 ± 0.04</td>
<td>-0.28 ± 0.04</td>
<td>-0.24 ± 0.04</td>
<td>-0.14 ± 0.03</td>
<td>-</td>
<td>-</td>
<td>0.19 ± 0.04</td>
<td>0.23 ± 0.03</td>
<td>0.20 ± 0.03</td>
<td>-</td>
</tr>
<tr>
<td>RH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.17 ± 0.02</td>
<td>0.20 ± 0.03</td>
</tr>
<tr>
<td>WS</td>
<td>0.35 ± 0.04</td>
<td>0.29 ± 0.04</td>
<td>0.19 ± 0.04</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.08 ± 0.02</td>
<td>-</td>
</tr>
<tr>
<td>SR*</td>
<td>0.42 ± 0.04</td>
<td>0.38 ± 0.04</td>
<td>0.37 ± 0.04</td>
<td>0.35 ± 0.04</td>
<td>0.30 ± 0.03</td>
<td>0.26 ± 0.03</td>
<td>0.14 ± 0.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R\textsuperscript{2}</td>
<td>0.53</td>
<td>0.55</td>
<td>0.54</td>
<td>0.52</td>
<td>0.56</td>
<td>0.59</td>
<td>0.65</td>
<td>0.72</td>
<td>0.81</td>
<td>0.65</td>
</tr>
</tbody>
</table>
3.6 Conclusions
To summarize, the concentration of UFP decreased between 2006 and 2011. PN$_{50}$ decreased by 21% during the measurement period, PN$_{50-100}$ decreased by 17% and PN$_{100-300}$ decreased by 24%. PN$_{50}$ exhibited significant negative correlation with temperature and relative humidity, and positive correlation with NO and NO$_2$. The other fractions of UFP exhibited stronger positive correlation with PM$_{2.5}$, NO$_2$, and SO$_2$. A MLR model captured the annual and seasonal variability of PN$_{50}$ but only partially explained the day-to-day variability ($R^2 = 0.52$). The MLR model more fully explained the day-to-day variability for PN$_{50-100}$ and PN$_{100-300}$ suggesting that long-term measurements for criteria pollutants can be used to estimate PN concentration for particles greater than 50nm. The long-term trends in PN$_{50}$ indicate that policy interventions are having some success in slowly reducing UFP concentrations in Toronto. Thus, parallel reductions may also be occurring in other Canadian cities.

3.7 Acknowledgements
Funding for SOCAAR was provided by the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund. This work was supported by the Ontario Ministry of the Environment's Best in Science Research Program. Student stipends were provided by the Ontario Graduate Scholarships in Science and Technology.
Chapter 4
Cluster Analysis of Roadside Ultrafine Particle Size Distributions

4.1 Executive Summary

This chapter contains a paper published in the peer-reviewed journal Atmospheric Environment (reprinted with permission). The supporting information from this paper is presented in Appendix C.


This chapter seeks to explore the use of alternative metrics to describe the temporal and spatial variation of UFP (Objective 2). In recent years, high-time resolution instruments that are capable of measuring both the number and size of ultrafine particles have become more prevalent. However, to capitalize on the wealth of information that is generated with these instruments, alternative data analysis techniques need to be developed. A cluster can be described as a group of objects that are similar in some way. There exist hundreds of clustering algorithms that capitalize on different data set properties to organize the objects into groups. In this chapter, the shape of the particle size distribution is examined and classified using the k-means clustering algorithm. A detailed explanation of the algorithm and cluster diagnostic tests is contained in Appendix B.

This work was undertaken to answer the following questions about the shape of the particle size distribution and its variation over the five year measurement period in downtown Toronto:

1. Mathematically, does the particle size distribution data contain natural clusters and can clustering analysis be applied to these data?
2. Are these clusters physically meaningful and what can they tell us about the variation of UFP in downtown Toronto?
3. What is the geographical range of these clusters?
4.1.1 Author Contributions

This work benefited from collaboration with the following co-authors: Cheol-Heon Jeong, Xiaohong Yao, Jun Yun-Seok and Greg J. Evans. Cheol-Heon Jeong calibrated, maintained and operated the Fast Mobility Particle Sizer and provided advice on the data interpretation aspects of this chapter. Xiaohong Yao provided data interpretation advice and initial feedback on the manuscript. Jun Yun-Seok provided access to a complimentary ultrafine particle data set for hypothesis testing. All field measurements, including instrument calibration and quality assurance, data compilation, validation, analysis, algorithm development, and interpretation were performed by me. The manuscript conceptualization and writing was performed by me with input, guidance, and editing from Professor Greg Evans.

4.2 Abstract

This study reports the diurnal, seasonal, and annual variation of ultrafine particle size distributions in downtown Toronto. The k-means clustering algorithm was applied to five years of size-resolved data for particles with diameters less than 100 nm. Continuous particle number concentrations were measured 16 m from a major arterial roadway between March 2006 and May 2011 using a Fast Mobility Particle Sizer. Eight particle size distribution (PSD) types were identified. The PSD types exhibited distinct weekday-weekend and diurnal patterns. The relative frequency that each PSD occurred varied with season and wind direction and was correlated with other pollutants. These temporal patterns and correlation helped in elucidating the sources and processes that each of the eight PSD represent. Finally, similar PSD types were observed in residential areas located 6 and 15 km away from the central monitoring site suggesting that these PSD types may be generalizable to other sites. Identification of PSD types was found to be a valuable tool to support the interpretation of PSD data so as to elucidate the sources and processes contributing to ultrafine particle concentrations.

4.3 Introduction

The proliferation of high-time resolution particle sizing instruments has greatly increased the rate at which long-term data on ultrafine particles (UFP: particles with diameters less than 100 nm) is being generated around the world (Mejía et al., 2007; Sabaliauskas et al., 2012; Wåhlin, 2009; Wang et al., 2011b) . Studies examining the behaviour of UFP have focused on reporting the
particle number (PN) concentration as a function of particle size, correlation of PN with gas and particle phase species, and back trajectory analysis to identify possible origins (Ketzel et al., 2004a; Mejía et al., 2008; Wang et al., 2011b). Although useful, the abovementioned analysis techniques do not fully capitalize on the wealth of information contained within the data. A complimentary representation is worth exploring as the health effects associated with exposure to particulate matter (PM) are widely believed to be dependent on particle size (Sioutas et al., 2005). The particle size distribution (PSD) of UFP present at any point in time and space arises from a combination of the contributing sources and the subsequent transformation processes. UFP are emitted in significant quantities from vehicles (Kittelson, 1998; Seigneur, 2009) and form as a result of interactions between sulphuric acid, water and ammonia (Korhonen et al., 1999). UFP in vehicular emissions differ in size and chemical composition by engine type (Kittelson, 1998), driving conditions (Ketzel et al., 2003), sulphur content of the fuel (Wåhlin, 2009), fuel composition (Turrio-Baldassarri et al., 2004) and emission control technology (Seigneur, 2009). Finally, UFP can shrink or grow rapidly in the atmosphere through evaporation and condensation processes (Yao et al., 2010).

The presence of certain PSD types on a given day may be indicative of the presence of combustion sources, occurrence of nucleation and growth events, or influence of long-range transport, photochemistry and meteorology. Factor analysis, Positive Matrix Factorization (PMF) and cluster analysis have previously been applied to PSD data (Zhou et al., 2004; Kim et al., 2004; Ogulei et al., 2007; Beddows et al., 2009; Kasumba et al., 2009). For factor analysis and PMF, the size bins within the PSD are treated as variables and are modelled as the linear combination of the potential factors plus a residual error term. Although a PSD may consist of a linear combination of multiple sources, interdependencies between the sources and subsequent transformation through condensation, evaporation, and coagulation processes can alter the shape of the PSD in a non-linear fashion. For example, coagulation could be a source of larger particles and a sink for smaller particles. However, particle sinks are not accounted for in the PMF model. Studies that have applied PMF have used monitoring sites located away from major traffic sources (Ketzel et al., 2004b). Zhou et al., (2004) applied PMF to two months of data from Pittsburgh, USA. However, the occurrence of numerous particle nucleation and growth events that interfered with the PMF analysis led to the removal of all but 17 days from the analysis. Kim et al., (2004) applied The Environmental Protection Agency’s Unmix and PMF to two months of
PSD data from Seattle, USA. The monitoring site was located far from traffic where no nucleation and growth events occurred during the measurement period that allowed for the majority of the data to be retained in the study. In both studies, factors associated with vehicular emissions, long-range transport and photochemical processes were identified.

In areas near traffic, or during nucleation and growth events, the PSD can be altered through coagulation, condensation, and evaporation (Ketzel et al., 2004b). In these cases, cluster analysis is appropriate because there are no assumptions that the PSD will behave as a linear combination of multiple factors without being impacted by non-linear transformative processes. Beddows et al., (2009) applied the k-means clustering algorithm to PSD data collected at multiple sites in London and at a rural site in the United Kingdom, so as to identify clusters representing the different PSD types at these sites. At the rural site, the clusters had distinctive diurnal behaviour and exhibited correlation with wind speed and combustion-related gas pollutants. At the urban sites in London, some clusters differed from those observed at the rural site and in some cases were unique to each sampling location. In addition to being location specific, some clusters appeared at certain times of the day (i.e., at night or midday). Therefore, cluster-based descriptions are likely to be site specific to an extent and only some PSD types may be common across large or distinctive geographical areas.

Clustering may offer a simple and efficient way to separate PSD shapes into common PSD types or clusters. The variability of these PSD types over time may provide insight into the temporal variability in underlying sources and processes contributing to the formation, growth and removal of UFP. Specifically, the type of PSD present at a site should change over time due to changing sources (gasoline vehicles, diesel vehicles, nucleation events, other regional combustion sources) and processes (condensational growth, evaporation, coagulation, removal). Furthermore, different PSD types could potentially arise from the same source, due to differing amounts of processing, while other PSD types could be from different sources after undergoing the same processing. This study applies cluster analysis to five years of UFP data collected at a roadside site in downtown Toronto, Canada. Previous studies investigating PSD factors or clusters have typically relied on UFP data collected over relatively short-time horizons (1 year or less). This paper examines seasonal and diurnal trends, the relationship between the PSD types and more traditional UFP metrics such as the PN concentration or GMD, the geographic range of
the clustering solution, and two case studies exploring long-range transport and seasonal diurnal trends.

## 4.4 Methodology

### 4.4.1 Study Location

The Greater Toronto Area is the most populous region in Canada with a population of 5.5 million people (Statistics Canada, 2012). Downtown Toronto experiences an increase in vehicular traffic between 7:00 and 9:00 in the morning on weekdays that remains constant throughout the day and gradually decreases between 19:00 and 21:00. On weekends, the traffic volume is lower, and the daily maximum occurs later in the day (11:00-13:00). Due to the presence of numerous restaurants, bars, and clubs in the downtown core, overnight traffic volumes are significantly higher on weekends than on weekdays. The traffic fleet is dominated by gasoline-powered spark ignition passenger vehicles. The sulphur content in gasoline and diesel is less than 30 ppm and 15 ppm, respectively.

### 4.4.2 Data Sources

Ultrafine particle measurements were collected at the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR) located on a four-lane major arterial roadway in downtown Toronto (Figure C-1, upper panel). The SOCAAR sampling inlet is located 20 m from the roadway and 6 m aboveground. Traffic volumes range from 15,000 to 20,000 vehicles per day with speeds up to 50 km/h. The UFP data consisted of 36,252 hourly-averaged UFP size distributions measured by a Fast Mobility Particle Sizer (FMPS: Model 3091, TSI Incorporated, Shoreview, MN, USA). Hourly averages were calculated from 1-minute averages for hours that had a minimum of 70% of data reporting (42 minutes). The resulting hourly-averaged size distributions used as the input data encompassed 83% of the five year period and the missing hours were uniformly distributed across years and seasons. The FMPS reported a number-based size distribution of $dN/d\log D_p$ across 32 size bins between 5.6 and 560 nm. The FMPS had variable counting efficiency for particles less than 8 nm and greater than 300 nm throughout the measurement period. To ensure the reported number concentration was consistent, only size bins between 8 and 300 nm (26 size bins) were included in this study. The PSD was corrected using the procedure described by Jeong and Evans (2009). The resulting size distributions were similar to those obtained using a Scanning Mobility Particle Sizer.
Hourly average concentrations for nitric oxide (NO), nitrogen dioxide (NO$_2$), ozone (O$_3$), carbon monoxide (CO), sulphur dioxide (SO$_2$) and fine particulate matter (PM$_{2.5}$) were available from the Ontario Ministry of the Environment’s Downtown Toronto monitoring station located 1 km from the SOCAAR monitoring site (MOE, 2011). Hourly temperature, relative humidity, wind speed and wind direction data was obtained from Pearson International Airport’s meteorological station operated by Environment Canada (Environment Canada, 2011). The National Air Pollution Surveillance Network (NAPS) provided speciation data for PM$_{10}$, PM$_{2.5}$, gas and particle phase polycyclic aromatic hydrocarbons (PAHs), and volatile organic compounds (VOCs) on a 24 hour sampling interval (Environment Canada, 2012). The detailed chemical sampling methodology for each species is reported by (Dabek-Zlotorzynska et al., 2011).

### 4.4.3 Data Treatment

Particle size distribution data has two components: the total number concentration and its shape. To normalize for this day-to-day variability in total concentration, the size distributions were normalized by treating each observation as a vector and dividing the number concentration in each size bin by the length of the vector as described by Beddows et al., (2009) (Equation 4-1):

\[
PSD_n = \frac{PN_i}{\sqrt{\sum_{i=1}^{n} (PN_i)^2}}
\]

where PSD$_n$ is the normalized particle size distribution, PN is the particle number concentration, i is the size bin, and n is the number of size bins. These normalized particle size distributions were used for the clustering so as to identify the clusters (PSD type). The un-normalized size distributions for all the hours when a given cluster was identified were then averaged to yield average PSD shapes to describe each PSD type (e.g., Figure 4-1).

### 4.4.4 Measuring Clustering Tendency

The validity of applying cluster analysis as a means of reducing the dimensionality of the dataset was assessed using a clustering tendency test. The Hopkins’ statistic ($H_p$) was previously applied to PSD data suspected to contain natural clusters (Beddows et al., 2009). Five subsets of 2000
size distributions were randomly selected from the dataset. The Euclidean distance was calculated between each normalized observation within each subset. The nearest neighbour distance (or minimum distance between a single size distribution and any of the other 1999 size distributions) was assigned to each observation \((W)\). Next, five sets of 2000 random size distributions were generated with a random number generator that possessed both the same range and log-normal distribution as the original PSD data. The Euclidean distance was calculated between each random point and its nearest neighbour \((U)\). Finally, the minimum distances assigned to the PSD data \((W)\) and random data \((U)\) were summed. The Hopkins’ Statistic was calculated according to Equation 4-2:

\[
H_p = \frac{\sum W}{\sum U + \sum W}
\]

The \(H_p\) value for the PSD data was 0.089 ± 0.001 (mean ± standard deviation). Values of \(H_p\) close to zero suggest that the measured data forms natural clusters. Values close to 0.5 suggest the experimental data is similarly distributed as the randomly generated data.

### 4.4.5 Data Analysis

All data analysis was performed using the R software environment Version 2.12 (R Development Core Team, 2008). The k-means clustering algorithm from the Flexible Procedures for Clustering (fpc) package (Henning, 2010) was applied to the PSD data. To select the optimal number of clusters, cluster diagnostic metrics from the ClValid package (Brock et al., 2011) were applied to the PSD data.

### 4.4.6 Solution Validation

To determine the optimal number of clusters, multiple cluster diagnostic metrics were calculated including the average diameter of each cluster, average distance between clusters, within-between ratio (average within distance/average distance between clusters), average silhouette width, Huber Gamma Index, and the Dunn Index for solutions with 2 to 20 clusters. As for Beddows et al., (2009), only the Dunn Index varied significantly between different solutions.
The Dunn index is the ratio of the minimum inter-cluster distance and the maximum intra-cluster distance. The Dunn index is maximized when a clustering solution contains the most compact and separated clusters. Due to the size of the dataset, the Dunn index was calculated using three subsets of the 10,000 PSD to arrive at the optimum number of clusters. When applied to the PSD data, the Dunn index was maximized at eight clusters implying that the clusters in this solution were more compact and separated than a solution consisting of seven or nine clusters (Figure C-2).

4.5 Results and Discussion
An eight cluster solution was selected for the PSD data. Table 4-1 lists the likely sources of each PSD type, its frequency of occurrence, and its % contribution to the observed PN concentration. The main sources are traffic emissions, regional nucleation, regional background, and long-range transport. Although traffic was the dominant source (C1-C6), only two PSD types could be attributed to traffic alone (C2 and C4). The presence of C1, C3, C5 and C6 indicated traffic emission influences in conjunction with other processes such as regional nucleation or long-range transport. Finally, the PSD types have distinct shapes (Figure 4-1) and exhibit unique diurnal (Figure 4-3), weekday-weekend (Table 4-2), and seasonal behaviour (Figure 4-4) that will be further discussed in this the paper.

Table 4-1 Summary of the source of each PSD type, its frequency of occurrence and its percentage contribution to the total observed PN concentration.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Sources</th>
<th>Frequency of occurrence (%)</th>
<th>% Contribution to PN Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Traffic, regional nucleation</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>C2</td>
<td>Winter traffic</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>C3</td>
<td>Traffic with low regional background, long-range</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>Night traffic</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>C5</td>
<td>Night traffic, long-range transport</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>C6</td>
<td>Summer regional with high traffic</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>C7</td>
<td>Summer regional with low traffic</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>C8</td>
<td>Summer long-range transport with low traffic</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>
4.5.1 Particle Size Distribution Clustering Solution

The average PSD in dN/dlogDp for each cluster is presented in Figure 4-1. The clusters were dominated by either nuclei-mode (Dp<50nm) or Aitken mode (Dp>50nm) particles. The nuclei-mode clusters tended to be unimodal (C1-C4) while the Aitken mode clusters (C5, C7, C8) were bimodal or unimodal with a shoulder. C6 was a mixture of both nuclei-mode and Aitken mode particles. The PN concentrations on weekdays (working days) were higher than on weekends (non-working days) for all clusters. However, smaller weekday-weekend differences were observed for clusters dominated by Aitken mode particles (C5-C8). The average PSD observed in Toronto during the five year measurement period is shown in Figure C-4.
Figure 4-1 Particle size distribution for each PSD type. C1-C4 were dominated by nuclei-mode particles, C5 had both nuclei-mode and Aitken mode particles, and C6-C8 were dominated by Aitken mode particles.
Table 4-2 The overall PSD type counts and their % contribution to the overall PSD mix. a: The weekday and weekend n refer to the number of instances that a PSD type was observed on a weekday or weekend. b: The theoretical split between weekdays and weekends is 71.4% and 28.6% respectively. Any weekday % contribution greater than 71.4% implies the PSD type is more likely to appear on weekdays. Any weekend % contribution greater than 28.6% implies the PSD type is more likely to appear on weekends. c: Average particle number (PN) concentration during hours when the given particle size distribution (C1-C8) was present.

<table>
<thead>
<tr>
<th>PSD Type</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall n</td>
<td>4,639</td>
<td>7,025</td>
<td>6,368</td>
<td>4,874</td>
<td>4,482</td>
<td>4,539</td>
<td>2,688</td>
<td>1,637</td>
</tr>
<tr>
<td>(% contribution to PSD mix)</td>
<td>(13)</td>
<td>(19)</td>
<td>(18)</td>
<td>(13)</td>
<td>(12)</td>
<td>(13)</td>
<td>(7)</td>
<td>(5)</td>
</tr>
<tr>
<td>Weekday n^a</td>
<td>3,145</td>
<td>5,553</td>
<td>4,648</td>
<td>3,734</td>
<td>31,24</td>
<td>3,114</td>
<td>1,778</td>
<td>1,133</td>
</tr>
<tr>
<td>(weekday split %)^b</td>
<td>(68)</td>
<td>(79)</td>
<td>(73)</td>
<td>(77)</td>
<td>(70)</td>
<td>(69)</td>
<td>(66)</td>
<td>(69)</td>
</tr>
<tr>
<td>Weekend n^a</td>
<td>1,494</td>
<td>1,472</td>
<td>1,720</td>
<td>1,140</td>
<td>1,358</td>
<td>1,425</td>
<td>910</td>
<td>504</td>
</tr>
<tr>
<td>(weekend split %)^b</td>
<td>(32)</td>
<td>(21)</td>
<td>(27)</td>
<td>(23)</td>
<td>(30)</td>
<td>(31)</td>
<td>(34)</td>
<td>(31)</td>
</tr>
<tr>
<td>PN concentration weekdays^c</td>
<td>21,610</td>
<td>27,080</td>
<td>18,520</td>
<td>26,530</td>
<td>22,550</td>
<td>16,650</td>
<td>16,940</td>
<td>13,290</td>
</tr>
<tr>
<td>PN concentration on weekends^c</td>
<td>15,290</td>
<td>19,060</td>
<td>13,850</td>
<td>19,250</td>
<td>18,770</td>
<td>13,890</td>
<td>15,710</td>
<td>11,720</td>
</tr>
<tr>
<td>PN concentration Ratio Weekday/Weekend</td>
<td>1.41</td>
<td>1.42</td>
<td>1.33</td>
<td>1.37</td>
<td>1.20</td>
<td>1.20</td>
<td>1.08</td>
<td>1.13</td>
</tr>
<tr>
<td>GMD (nm)</td>
<td>19.9</td>
<td>22.4</td>
<td>26.4</td>
<td>26.6</td>
<td>31.3</td>
<td>33.2</td>
<td>39.2</td>
<td>46.4</td>
</tr>
</tbody>
</table>
4.5.2 Cluster Contribution

The frequency of occurrence of each PSD type was determined in terms of the contribution of each PSD type to the overall PSD mix (Table 4-2). The PSD types dominated by nuclei-mode particles (C1-C4) contributed 63% of all observed PSD types. PSD types dominated by Aitken mode particles (C7 and C8) contributed the least to the overall PSD mix (total of 12%). In addition, the PSD types occurred with different frequencies on weekdays and weekends (Table 4-2). Notably, C2 and C4 were more likely to be observed on weekdays while C7 was more likely to be observed on weekends. The ratio of PN concentrations on weekdays and weekends was consistently above 1.0 for all PSD types. However, PSD types dominated by Aitken mode particles (C7 and C8) had ratios closer to 1.0.

4.5.3 Time Series Analysis

Power spectral density is a commonly used signal processing technique that describes how the energy of a time series is distributed in the frequency domain. This method allows the relevant time scales governing the temporal variation of each PSD type to be explored and for the identification of daily, weekly, semi-annual and annual patterns (Figure 4-2). At period = 24 hours, each PSD type possessed a strong spectral density signal indicating that the relative frequency with which all the PSD types occurred exhibited diurnal behaviour. C2 had the strongest spectral density signals (y-axis scale from 0 to 100 versus 0 to 20 for other PSD types). C2 also had a relatively strong spectral density signal at period = 168 indicating a possible weekly pattern. Consistent with this, C2 exhibited the strongest weekday-weekend distribution split (weekday = 79% and weekends = 21%) and highest weekday-weekend PN concentration ratio of 1.42 (Table 4-2). The temporal patterns of the PSD types are based on the frequency of occurrence. The presence of one PSD type means that other PSD types are absent. Therefore, a strong temporal trend created by one PSD type has the potential to create an opposing or mirroring temporal trend that corresponds to the behaviour of the absent PSD types. For example, the strong diurnal pattern of C2 may have produced weak patterns in other PSD types such as C1 and C6.
Figure 4-2 Fourier analysis of the PSD types on a daily, weekly, monthly, semi-annual, and annual basis. Notably, the spectral density at period = 24 hours contained a strong signal suggesting that the PSD all possess strong diurnal behaviour. All PSD types except C3 and C5 had strong spectral density signals at period = 1 year. All peaks possessing statistical significance at the 95% level of confidence are denoted by *. 

Further, wind patterns in Toronto are impacted by both mesoscale and synoptic scale processes (Figure C-3). Lake breeze influences occur daily while other larger scale shifts in wind direction occur over time horizons of 3-5 days. Weak and distributed spectral density signals were present between 72 and 120 hours (3 to 5 days) for C3, C4, C5, C7 and C6 possibly due to these synoptic scale changes in wind direction. In addition, C1-C4 had weak spectral density signals at period = 4369 (~ 6 months) suggesting that seasonal factors influenced the occurrence of these PSD types, such as similar meteorological conditions observed during some spring and autumn months (i.e.,
April and October). Finally, all PSD types except C3 and C5 had spectral density signals at period = 9362 hours (~1 year) suggesting that each PSD type exhibit annual patterns, presumably due to differences between summer and winter.

4.5.4 Diurnal Behaviour
Each PSD type had distinct diurnal patterns (Figure 4-3) and in some cases possessed strong weekday-weekend variation. For example, the diurnal patterns of the C3, C4, C5 and C6 PSD types were similar on weekdays and weekends, while C1, C2, C7 and C8 exhibited different weekday-weekend patterns. The relative frequency, with which C2 occurred, increased during the morning and afternoon periods and decreased overnight on both weekdays and weekends; however, the frequency of C2 was higher during the afternoon on weekdays. C1 had a relatively flat weekday diurnal pattern and a large peak during the afternoon periods on weekends. C7 and C8 peaked between 4:00 and 6:00 and dropped back to background levels by 8:00 on weekdays. On weekends, C7 peaked later in the morning (between 6:00 and 8:00) and returned to background levels by 12:00. Further, C2 and C3 increased starting at 6:00 on weekdays, corresponding to the morning rush-hour period, and remained elevated throughout the afternoon.
Figure 4-3 Diurnal trends for each PSD type. The diurnal trend was calculated by summing the total number of times a given PSD type appeared at given hour on a weekday divided by the total number of times all types of PSD were observed in that given hour on a weekday. A similar calculation method was used for weekends.
4.5.5 Seasonal Variation

Figure 4-4 shows the ratio of each PSD types’ summertime contribution (July – September) and wintertime contribution (January – March). PSD types dominated by larger particles (C7-C8) were more commonly observed during the summer months (ratios = 1.8 - 6.0), perhaps due to increased production of particle phase sulphate from several coal-fired power plants located to the south and southwest of Toronto (Figure C-1, lower panel). During the winter months, cooler temperatures and reduced photochemistry favour the formation and increase the atmospheric lifetime of nuclei-mode particles. For example, although C2 and C3 exhibited similar weekday diurnal behaviour, C2 was more likely to be present during the winter months (ratio < 1.0). Similarly, C1 had a strong weekend diurnal pattern and occurred more often in winter. This suggests that the formation or removal processes for particles within the C1 and C2 PSD types may be temperature dependent. In addition, C3 and C5 were shown to exhibit limited seasonal variation in the Fourier analysis and had summertime to wintertime ratios of approximately 1.0.

Figure 4-4 Ratio of the contribution of each PSD during the summer to the contribution of each PSD during the winter. PSD types associated with nuclei-mode particles (C1, C2, C4) contributed more to the overall PSD type mix during the winter. PSD types associated with Aitken mode particles (C6-C8) were more likely to be present during the summer months. The contribution of C3 and C5 was relatively constant between seasons.
4.5.6 Wind Directionality

The predominant wind directions in Toronto are north-westerly and south-westerly (Figure C-3). Multiple coal-fired electricity generation facilities and industrial sources are located to the south and south-west of Toronto (Figure C-1, lower panel). In addition to higher ambient temperatures when Toronto experiences southerly winds, there is an increased likelihood of observing elevated concentrations of PM$_{2.5}$ and other combustion-related pollutants (Godri et al., 2009). The areas to the north of Toronto have few large industrial emitters and northerly winds typically transport cooler and cleaner air into the city. Figure 4-5 shows each PSD types’ behaviour as a function of wind direction and season (winter and summer). All PSD types showed some form of wind directionality dependence. Generally, PSD types dominated by nuclei-mode particles exhibited north-westerly dependence while PSD types dominated by Aitken mode particles were more likely to be observed from southerly wind directions. This difference may have reflected the greater number of distant sources to the south-west than the north-west, resulting in older and hence larger particles arriving with winds from the south-west direction. Not all PSD types exhibited the same behaviour between seasons. For example, C5, C6, and C7 showed much stronger south-westerly wind dependence during the winter than during the summer.
Figure 4-5 The contribution of each PSD type (%) as a function of wind directionality. Regardless of season, PSD types dominated by nuclei-mode particles (C1, C2, C4) were more likely to originate from north-westerly directions. PSD types dominated by Aitken mode particles (C5-C7) were more likely to originate from southerly directions during the winter.
4.5.7 Correlation with PM$_{2.5}$, Gas Phase Pollutants and Meteorological Data

The strength of the relationship between the occurrence of each PSD type and the concentration of gas and particle phase species may be indicative of the sources or processes that are associated with their appearance. Although not reflective of the chemical composition of UFP, relationships between PM$_{2.5}$ and the occurrence of each PSD type can provide insight into the factors that lead to the presence or absence of certain PSD types. For example, PM$_{2.5}$ provides an indication of the availability of secondary compounds, often produced through photochemical processes that may contribute larger particles from regional or distant sources, or to local particle growth. The dominant PM$_{2.5}$ species observed in Toronto are sulphate, nitrate, ammonium, organic carbon and elemental carbon. However, the composition of PM$_{2.5}$ varies by season due to differences in photochemistry, local emissions, and long-range transport (Lee et al., 2003; Buset et al., 2006; Jeong et al., 2011). During the summer months, higher peak electricity demand results in increased electricity generation from coal-fired power plants located to the south and south-west of Toronto. Higher SO$_2$ emissions coupled with photochemistry lead to higher ambient PM$_{2.5}$ concentrations dominated by secondary sulphate. During the winter months, lower ambient temperatures allow for more nitrate species to remain in the particle phase. Previous studies applying receptor analysis (PMF) in Toronto have observed four dominant factors contributing ~85% of the PM$_{2.5}$ mass: secondary sulphate, secondary nitrate, traffic, and elemental carbon rich (Jeong et al., 2011). Minor factors contributing ~15% of the PM$_{2.5}$ mass include biomass burning, salt, road dust, soil dust, and metallurgy.

The correlation matrix presented in Table 4-3 shows the Spearman Rank Order correlation coefficients for each PSD type and PM$_{2.5}$, indicator gas phase pollutants, and meteorological parameters on a seasonal basis. Additional correlation analysis with the main constituents of PM$_{2.5}$ and specific PAH compounds can be found in Table C-1. C1 and C2 exhibited negative correlation with PM$_{2.5}$, NO, and NO$_2$ and positive correlation with wind speed. Thus, these two PSD types were characteristic of “new” particles, with very limited growth or contribution from “older” particles of more regional origins. Based primarily on their temporal patterns, C2 was attributed to vehicle emissions and C1 to nucleation events adding to particles from vehicles. In contrast, although C3 was dominated by nuclei-mode particles, a slight shoulder at 60 nm was observed. C3 exhibited positive correlation with PM$_{2.5}$ during the winter and negative correlation
during the summer. This finding is further supported by the wind directionality analysis. C3 did not show strong wind direction dependence during the summer, but exhibited southerly dependence during the winter. Thus, C3 was attributed to new particles from traffic emissions adding to aged particles from local to regional sources during the summer or long-range transport during the winter. Effectively, the slower oxidation to form secondary pollutants within polluted air masses from the south-west in the winter would cause limited growth, yielding C3. In the summer, similar growth could occur due to more local emissions. Further, the rapid oxidation within air masses from the south-west in the summer would cause more growth, presumably yielding PSD types with larger modes such as C6 and C7, or in extreme cases C8. Although the nuclei-mode PSD types (C1-C4) are believed to be traffic induced, only C4 and C5 exhibited positive correlation with NO and NO₂. This correlation was governed by the day-to-day variability in NOₓ, rather than diurnal patterns as 24-hour averaged data was used to create Table 4-3. The weekday diurnal pattern for the contribution of C2 (Figure 4-3) did in fact have similarities to that of NOₓ, yet C2 was negatively correlated with NOₓ in terms of daily-averaged values. NOₓ emissions from vehicles vary little between weekdays and thus daily-averaged NOₓ concentrations are inversely correlated with wind speed. C1 and C2 were positively correlated with wind speed; presumably less time is available for growth on windy days, which also have lower NOₓ. Conversely, C4 and C5 (winter) were negatively correlated with wind speed suggesting that stagnant conditions promote the formation of these PSD types. The occurrence of C7 and C8, which was much more frequent in the summer, was correlated with PM₉.₅. These PSD types, which occurred most frequently during the early morning in summer, were believed to be characteristic of strong regional or long range transport contributions with a small contribution from local vehicle emissions. In contrast, PSD type C6, which occurred more evenly throughout the day and was also correlated with PM₂.₅, was believed to be characteristic of periods when vehicles emissions were present but overwhelmed by particles from long range transport and regional sources.
Table 4-3 Seasonal Spearman’s Rank Order correlation coefficients for 24-hour PSD type counts and 24-hour PM$_{2.5}$ filter data obtained from the NAPS network and meteorological data. All presented correlation coefficients are significant to 95% confidence. Correlation coefficients are bolded to emphasize the changes in the relationships between the PSD types and the PM$_{2.5}$ between seasons.

<table>
<thead>
<tr>
<th>Season</th>
<th>Compound</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>PM$_{2.5}$</td>
<td>-0.71</td>
<td>-0.68</td>
<td>0.34</td>
<td>-</td>
<td>0.45</td>
<td>0.78</td>
<td>0.58</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>-0.42</td>
<td>-</td>
<td>-</td>
<td>0.31</td>
<td>0.38</td>
<td>-</td>
<td>0.12</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NO$_2$</td>
<td>-0.55</td>
<td>-0.21</td>
<td>0.14</td>
<td>0.31</td>
<td>0.46</td>
<td>0.19</td>
<td>0.20</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>O$_3$</td>
<td>0.41</td>
<td>0.12</td>
<td>-0.14</td>
<td>-0.17</td>
<td>-0.27</td>
<td>-0.13</td>
<td>-0.24</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Temperature Relative</td>
<td>-0.26</td>
<td>-0.34</td>
<td>0.37</td>
<td>-</td>
<td>-</td>
<td>0.32</td>
<td>0.20</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Humidity</td>
<td>-0.41</td>
<td>-0.18</td>
<td>0.27</td>
<td>-</td>
<td>-</td>
<td>0.31</td>
<td>0.30</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Wind Speed</td>
<td>0.57</td>
<td>0.32</td>
<td>0.16</td>
<td>-0.32</td>
<td>-0.44</td>
<td>-0.17</td>
<td>-0.27</td>
<td>-</td>
</tr>
<tr>
<td>Summer</td>
<td>PM$_{2.5}$</td>
<td>-0.55</td>
<td>-0.63</td>
<td>-0.46</td>
<td>-0.47</td>
<td>-0.29</td>
<td>0.59</td>
<td>0.32</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>-</td>
<td>-</td>
<td>0.13</td>
<td>0.26</td>
<td>0.31</td>
<td>-0.19</td>
<td>-</td>
<td>-0.28</td>
</tr>
<tr>
<td></td>
<td>NO$_2$</td>
<td>-0.40</td>
<td>-0.24</td>
<td>-0.15</td>
<td>-</td>
<td>0.12</td>
<td>0.16</td>
<td>0.18</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>O$_3$</td>
<td>-0.34</td>
<td>-0.47</td>
<td>-0.38</td>
<td>-0.51</td>
<td>-0.29</td>
<td>0.48</td>
<td>0.34</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>Temperature Relative</td>
<td>-0.28</td>
<td>-0.41</td>
<td>-0.19</td>
<td>-0.44</td>
<td>-0.40</td>
<td>0.53</td>
<td>0.21</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Humidity</td>
<td>-0.17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Wind Speed</td>
<td>0.38</td>
<td>0.31</td>
<td>0.17</td>
<td>-0.24</td>
<td>-</td>
<td>-0.21</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4.5.8 Geographical Range

The validity of extrapolating the results from this clustering solution to areas outside of Toronto’s downtown core was explored by comparing the shape of the PSD measured at two residential field sites located 6 and 15 km west of the SOCAAR monitoring site. Measurements were collected with a Scanning Mobility Particle Sizer with a nano-differential mobility analyzer over a period of two weeks at Site A and one week at Site B. Using the average centroid of each PSD type, cosine similarity was calculated for each hourly normalized PSD observed at the residential site and central site. Essentially, this calculation evaluated how similar the PSD at a residential site was to the known PSD type observed simultaneously at the central site. Each hourly similarity value was then stratified by PSD type and averaged, with a high average value indicating that the PSD types at the two sites were generally the same. Average values of the
similarity between the hourly PSD at the central site and the average centroid for the PSD type were calculated to give an indication of the expected variability within any PSD type.

Table 4-4 Average cosine similarity between the centroid of each PSD type and the two residential field sites and the Central site. a: Site A is located 6km west of the SOCAAR site and 280m from a major arterial roadway (n=226). b: Site B is located 16km west of the SOCAAR site, 800m from a major arterial roadway and near a large park and ravine area (n = 156). c: PSD type was not observed during the measurement period. d: n=2.

<table>
<thead>
<tr>
<th>PSD Type</th>
<th>Site A^a</th>
<th>Site B^b</th>
<th>Central Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>0.76</td>
<td>N.A. c</td>
<td>0.91</td>
</tr>
<tr>
<td>C2</td>
<td>0.74</td>
<td>0.77</td>
<td>0.94</td>
</tr>
<tr>
<td>C3</td>
<td>0.74</td>
<td>0.76</td>
<td>0.96</td>
</tr>
<tr>
<td>C4</td>
<td>0.80</td>
<td>0.80</td>
<td>0.94</td>
</tr>
<tr>
<td>C5</td>
<td>0.79</td>
<td>0.82</td>
<td>0.92</td>
</tr>
<tr>
<td>C6</td>
<td>0.76</td>
<td>0.78</td>
<td>0.93</td>
</tr>
<tr>
<td>C7</td>
<td>0.85</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td>C8</td>
<td>N.A. c</td>
<td>0.85 d</td>
<td>0.82 d</td>
</tr>
</tbody>
</table>

The shape of the PSD types observed at the residential sites showed a relatively high degree of similarity (min = 0.74) with the average centroid from the clustering solution (Table 4-4). The values at the residential sites were lower than those for the central site, indicating that the same PSD types did not always occur simultaneously at the two sites. In addition, cosine similarity values were slightly lower at Site A than at Site B. Similarity was highest for C4, C5 and C7 implying that the presence of these PSD types in the downtown core may be indicative of more frequent or similar PSD shapes in these residential areas. However, the PN concentrations corresponding to C4 and C7 at the residential sites were lower than at the downtown SOCAAR site. The baseline PN concentration observed in the downtown core is typically higher than at the residential sites due to larger traffic influences. This similarity suggests that the sources and processes yielding these PSD types were similar at both location but less intense at the residential site. The relatively high degree of similarity for all PSD types implies that similar PSD shapes may exist in areas away from the downtown core and in areas away from significant traffic sources across the city even though the PN concentration may differ. During this measurement period, the shapes of the observed size distributions at the SOCAAR site were similar to the centroid values obtained from the clustering solution. For C1-C6, high
cosine similarity was observed (>0.91). PSD types C7 (n=25), and C8 (n=2) had fewer observations that may have impacted the averaging yielding greater variability. Notably, the cosine similarity values obtained for C7 and C8 at the residential sites were close to those obtained at the SOCAAR site.

4.5.9 Case Studies
Two cases were developed to illustrate how the PSD type approach could be used to help understand and describe different sources and processes contributing to UFP.

4.5.9.1 Case Study 1: Diurnal Trends
At this near-road site in a major city, vehicle emissions were believed to always contribute UFP to an extent that varied during the day. The presence of nuclei-mode particles from vehicles was strongly evident in some PSD (e.g., C2) and essentially absent in others (e.g., C8). Figure 4-6 examines the evolution of the weekday wintertime and summertime diurnal trends. The normalized contribution of each PSD type was calculated by summing the total number of particles observed while each PSD type was present, and dividing this total by total number of particles observed during the season. In fact, the shape of the diurnal trend was the same as that for the average weekday particle number concentration and was similar between seasons (although PN is higher in winter). However, each PSD type was present in different proportions throughout the day and between seasons. During the winter, C2 contributed the most to the diurnal trend while PSD types dominated by Aitken mode particles (C5-C8) produced a “background” that was relatively constant throughout the day. In contrast, the summertime diurnal trend had more C3 than during the winter and had higher contributions of C6, C7, and C8. The higher contribution of C3 during the summer further supports the hypothesis that C3 represents traffic particles superimposed on a regional background concentration. A version of Figure 4-6 plotted in terms of the percentage contribution of each PSD type can be found in Figure C-4.
Figure 4-6 The wintertime (upper panel) and summertime (lower panel) weekday diurnal trend. The normalized contribution is the total number of particles observed while each PSD was present divided by the total number of particles observed during the season. The time-of-day percentage contribution of each PSD type can be found in Supplementary Figure 3.

4.5.9.2 Case Study 2: Long-Range Transport
Toronto frequently experiences summertime PM$_{2.5}$ episodes due to production of particle phase sulphate from SO$_2$ emitted by coal-fired power plants located to the south and south-west of the city. The impact of long-range transport on the PSD type mix was explored using data collected
between June 27th and July 1st 2006 (Figure 4-7). During this four day period, the concentration of PM$_{2.5}$ increased after the wind direction shifted from an easterly to southerly direction. In addition, the PN concentration of particles between 100 and 300nm (PN$_{100-300}$) increased and the PSD types transitioned from types dominated by nuclei-mode particles to PSD types dominated by Aitken mode particles (C6 and C8).

![Graph showing PM$_{2.5}$ and wind direction](image)

Figure 4-7 Example of long-range transport impacting Toronto between June 27th and July 1st 2006. The upper panel shows the variation of PM$_{2.5}$ (left axis) and the wind direction (right axis). The lower panel shows the variation of PN$_{100-300}$. As the wind direction shifted from easterly to southerly directions, the concentration of PM$_{2.5}$ and PN$_{100-300}$ increased and the PSD types transitioned from PSD types dominated by nuclei-mode particles (C3 and C4) to PSD types dominated by Aitken mode particles (C6 and C8).
4.5.10 Summary

The characteristics and possible sources associated with each PSD type are summarized in Table 4-5. C1 was more likely to be observed on weekends and exhibited a weekend diurnal trend that was consistent with regional nucleation and growth events. Thus, C1 was characteristic of periods when regional nucleation added to particles routinely present due to local vehicle emissions. C2 likely originated from “fresh” traffic emissions due to its increased contribution throughout the daytime period on weekdays. Thus, C2 was characteristic of periods with higher traffic when little photochemical growth, or evaporation, occurred and regional aerosol was absent, conditions that most often occur during the winter. C3 had a shoulder at 60 nm and likely represented traffic emissions added to a low background of aged aerosol. This aged aerosol originated from local-to-regional emissions from the northwest in the summer or long range transport from the southwest in the winter. C4 was more likely to be observed on weekdays during the evening suggesting an association with night-time traffic and was present in similar proportions during the winter and summer. C5 was present in higher quantities overnight and was positively correlated with NOx also suggesting a possible traffic origin. C5 differed from C4 in terms of its association with PM_{2.5} and southwest winds, suggesting that it was characteristic of night-time traffic emissions on top of older particles from regional sources or long range transport. C6, C7, and C8 exhibited positive correlation with PM_{2.5} mass and occurred when higher numbers of older/larger particles were present as a result of long-range transport. C6 was characteristic of small particles from vehicle emissions in combination with a high concentration of these older particles (cf. C3). In contrast, C7 and C8 occurred most often in the early morning when traffic is low, with C8 having larger particles indicative of periods with greater growth, presumably due to increased photochemical production of secondary pollutants in the summer.
Table 4-5 Characteristics of each PSD type throughout the measurement period. LRT refers to long-range transport. a: denotes the diameter of the maximum mode.

<table>
<thead>
<tr>
<th>PSD Type</th>
<th>Modes Dp (nm)</th>
<th>Highest Contribution</th>
<th>Defining Characteristics</th>
<th>Possible Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>14.3</td>
<td>Winter, Weekends</td>
<td>Negative correlation with PM$_{2.5}$ mass species and positive correlation with wind speed</td>
<td>Traffic, regional nucleation</td>
</tr>
<tr>
<td>C2</td>
<td>19.1</td>
<td>Winter, Weekdays</td>
<td>Strong weekday diurnal trend Negative correlation with PM$_{2.5}$ mass and positive correlation with wind speed</td>
<td>Winter traffic</td>
</tr>
<tr>
<td>C3</td>
<td>16.5</td>
<td>Afternoon</td>
<td>Similar weekday-weekend diurnal trends, Limited seasonal variation, Seasonal wind direction dependence, Positive correlation with PM$<em>{2.5}$ during the winter and negative correlation with PM$</em>{2.5}$ during the summer</td>
<td>Traffic with low regional background, LRT</td>
</tr>
<tr>
<td>C4</td>
<td>25.5</td>
<td>Evening, Weekdays</td>
<td>Similar weekday-weekend diurnal trends, Limited seasonal variation, Positive correlation with NO and NO$<em>2$ in winter, Negative correlation with wind speed in winter; PM$</em>{2.5}$ mass and temperature during the summer</td>
<td>Night traffic</td>
</tr>
<tr>
<td>C5</td>
<td>39.2</td>
<td>Overnight</td>
<td>Similar weekday-weekend diurnal trends, Positive correlation with PM$_{2.5}$ mass during the winter</td>
<td>Night traffic, with LRT</td>
</tr>
<tr>
<td>C6</td>
<td>19.1a, 45.3</td>
<td>Summer</td>
<td>Positive correlation with PM$_{2.5}$ mass, Relatively flat diurnal trend</td>
<td>Summer regional with high traffic</td>
</tr>
<tr>
<td>C7</td>
<td>19.1, 52.3a</td>
<td>Summer, Early morning</td>
<td>Positive correlation with PM$_{2.5}$ mass, Large peaks in early morning (4:00-6:00)</td>
<td>Summer regional with low traffic</td>
</tr>
<tr>
<td>C8</td>
<td>16.1, 80.6a</td>
<td>Summer, Early morning</td>
<td>Strong positive correlation with PM$_{2.5}$ mass, Large peaks in early morning (4:00-6:00)</td>
<td>Summer LRT with low traffic</td>
</tr>
</tbody>
</table>

4.6 Conclusions

This paper applied cluster analysis to five years of size resolved UFP data collected at a roadside site in downtown Toronto. Eight particle size distribution (PSD) types were identified with distinct shapes, arising from different combination of vehicular emissions, meteorology, long-range transport, and photochemistry. PSD types dominated by nuclei-mode particles tended to be unimodal and PSD types dominated by Aitken mode particles were bimodal. In addition, some
PSD types exhibited strong weekday-weekend differences along with distinct diurnal trends. Further, the frequency of occurrence of some PSD types varied by season as did their correlation with other pollutants and meteorological parameters. Finally, similar PSD types were observed at other nearby sites suggesting that PSD types identified through clustering may be generalizable over larger geographical areas. Identification of PSD types was found to be a valuable tool to support the interpretation of PSD data so as to elucidate the sources and processes contributing to ultrafine particle concentrations.

4.7 Acknowledgements

Funding for SOCAAR was provided by the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund. This work was supported by the Ontario Ministry of the Environment's Best in Science Research Program. Student stipends were provided by the Ontario Graduate Scholarships in Science and Technology.
Chapter 5
The Application of Wavelet Decomposition to Quantify the Local and Regional Sources of Ultrafine Particles in Cities

5.1 Executive Summary

The paper presented in this chapter is accepted for publication in Atmospheric Environment (reprinted with permission). The supporting information for this paper is presented in Appendix D.


This chapter seeks to explore the use of alternative metrics to describe the temporal and spatial variation of UFP (Objective 2) and to quantify the contributions of the local scale and regional scale sources of UFP (Objective 3).

This chapter continues along the theme of developing alternative analysis methods for assessing exposure to UFP. One of the main features of a roadside UFP time series are high frequency spikes in PN concentration that occur vehicles pass the monitoring site. These PN concentration spikes typically persist for seconds to a few minutes and appear to be superimposed on a background PN concentration that changes over the course of several minutes to hours. The contribution of these individual vehicle-related spikes to the total PN concentration potentially varies significantly across an urban area as a function of local traffic. In addition, drivers and pedestrians travelling along major roadways may also be exposed to significantly more of these “spikes” than those that select less trafficked routes.

Wavelet decomposition is a commonly used signal processing technique that separates a time series by frequency using a series of high and low pass filters. In the case of the PN concentration time series, the high frequency portion relates to the passing vehicles while the low frequency portion relates to the slowly changing background concentration. This chapter
examines the PN concentration at a central downtown site located at the Southern Ontario Centre for Atmospheric Aerosol Research, two nearby downtown sites and four residential sites located across Toronto. To separate the PN concentration spikes from the background concentration, a wavelet based background subtraction algorithm was adapted from Klems et al., (2010) and applied to the time series data collected at each field site.

This study seeks to answer the following questions:

1. Can the UFP time series be separated into its high and low frequency components to quantify the impacts of sources and processes that are local vs. regional in nature
2. At the residential field sites, is the high frequency portion of the time series related to nearby vehicles?
3. What are the factors influencing the urban to regional portion of the time series?

5.1.1 Author Contributions
This study benefited from the contributions from the following co-authors: Cheol-Heon Jeong, Xiaohong Yao and Greg J. Evans. Cheol-Heon Jeong maintained and operated the Fast Mobility Particle Sizer throughout the study period and assisted with data interpretation. Xiaohong Yao assisted with data interpretation and provided feedback during manuscript development. The study design, field site selection, operation and calibration of field instruments, data analysis, and algorithm development were performed by me. All manuscript writing was performed by me with input, guidance and editing from Professor Greg Evans.

5.2 Abstract
This study explores the application of wavelet decomposition as a means to distinguish between local and regional sources of ultrafine particles (UFP). Particle number concentrations were measured at a central site, two downtown sites, and four residential sites located across Toronto, Canada. Using a wavelet decomposition algorithm, particle concentration time series were separated into two signals: high frequency local-to-neighbourhood scale sources and low frequency urban-to-regional scale sources and processes. At the field sites, local-neighbourhood sources contributed between 13 and 32% of the total particle concentration. The urban-regional signal at each field site exhibited stronger correlation and greater homogeneity with respect to the central site than the original concentration time series. In contrast, the high frequency local-
neighbourhood source signals exhibited limited correlation and high heterogeneity with respect to the central site. Traffic volume within a 2.5km buffer explained 87% of the variability in the local-neighbourhood level signal observed between field sites while no significant association with traffic was found for the original particle number concentration data. This study has demonstrated that wavelet decomposition can be a useful tool for estimating exposure to UFP from local-neighbourhood and urban-regional scale sources and processes.

5.3 Introduction

Ultrafine particles (UFP) are defined as particulate matter (PM) with diameters ($d_p$) less than 100nm. In urban areas, UFP are emitted from combustion sources such as traffic, and form through secondary reactions between sulphuric acid, ammonia, and water (Korhonen et al., 1999; Kulmala et al., 1998; Sihto et al., 2006). The relative contribution of each of these sources varies due to human behaviour, meteorological conditions, and the presence or absence of large particles and precursor gases. As a result, the UFP concentration can vary significantly with time of year, time of day, and between working and non-working days (Aalto et al., 2005; Cyrys et al. 2008; Johansson et al. 2007; Ketzel et al. 2004a; Mejía et al. 2008).

Unlike the larger fractions of PM, UFP are not yet routinely monitored for regulatory purposes nor included in existing air quality monitoring networks in most countries. However, a growing number of urban areas have long-term (>5 years) UFP datasets (Cyrys et al., 2008; Mejía et al., 2007; Sabaliauskas et al., 2012; Wählin et al., 2009; Wang et al., 2011b). Long-term UFP trends suggest a gradual reduction in particle number (PN) concentrations in North American and European cities largely attributable to changes in both the vehicle fleet and fuel composition (Sabaliauskas et al., 2012; Wählin et al., 2009; Wang et al., 2011b). A recent assessment examining the impacts of UFP on health concluded that UFP exposure estimates have greater uncertainty than the larger fractions of PM due to a lack of spatially resolved data (HEI, 2013). Although the PN concentration has been shown to be highly variable between cities (Aalto et al., 2005; Kim et al., 2002; Puustinen, et al., 2007), moderate temporal correlation has been observed between field sites separated by up to 30 km (Buzorius et al., 1999; Ketzel et al., 2004a; Cyrys et al., 2008).
The observed UFP concentration at any given urban site can be described by sources and processes that operate on local, neighbourhood, urban and regional spatial scales. At the urban and regional spatial scales, meteorological conditions can dramatically impact the observed concentration of UFP. Temperature, humidity, solar radiation, atmospheric mixing, and pre-existing aerosol may enhance or inhibit new particle formation (Jeong et al., 2010; Johansson et al., 2007; Kim et al. 2002; Shi et al. 2001; Yao et al., 2007). For example, regional scale nucleation events can impact the observed UFP concentration across entire urban areas, last several hours, and have spatial scales on the order of hundreds of kilometers (Jeong et al., 2010). From an exposure perspective, regional and urban scale processes can influence temporal patterns by increasing or decreasing background UFP concentrations in a spatially homogenous way across the entire urban area.

Local and neighbourhood scale sources are believed to have the largest influence on intra-urban temporal and spatial patterns in UFP exposure. For example, UFP studies near roadways have consistently demonstrated that nuclei mode particles ($d_p<50$nm) are emitted in significant quantities from vehicles (Pirjola et al., 2006; Zhang and Wexler 2004a; Zhu et al. 2002a; 2002b). In addition passing vehicles can produce large spikes in PN concentration that can persist for seconds and are not typically observed within tens of meters from major roadways (Fushimi et al., 2008; Kaur et al., 2006; Klems et al., 2010). The magnitude of these PN concentration spikes differs by vehicle age and type (i.e. diesel or gasoline powered), and traffic conditions. These local UFP emissions can be unique to individual roadways within an urban area.

One third of all Canadians, about 10 million people, live within 250m of a major roadway. In the more densely populated Greater Toronto Area, 56% of the population live within 250m from a major roadway and 24% of the population live within a distance of 100m (Evans et al., 2011;). This paper explores spatiotemporal variation of UFP in five neighbourhoods with different traffic influences in Toronto, Canada. Wavelet decomposition is applied to express observed PN concentrations as two separate signals, reflecting local-to-neighbourhood and the urban-to-regional scales contributions. Klems et al., (2010) applied wavelet decomposition as a means to separate the PN concentration “spikes” from the urban-regional signal at a single roadside monitoring site. Here, this approach is applied to longer term monitoring data conducted at
multiple sampling sites enabling the spatial dependence of the urban to regional signal and PN concentration “spikes” to be explored.

5.4 Methodology

5.4.1 Study Location
The Greater Toronto area (GTA) is the largest urbanized area in Canada and is home to 5.5 million people or about one sixth of Canada’s population. Due to its significant population and economic growth, rapid urbanization has occurred throughout the GTA. However, investment in public transportation networks and other transportation infrastructure has not kept pace with the mobility demands of the GTA’s population. In recent years, traffic congestion on regional highways and arterial roadways has become a major issue. GTA residents have the longest average commute times in North America (Toronto Board of Trade, 2011b). While the majority of Toronto residents that live near the downtown core regularly walk, cycle or use public transportation to commute to their workplaces, those living outside the downtown core have few alternatives to driving (Statistics Canada, 2007a).

5.4.2 Vehicle Fleet
Passenger vehicles in Ontario are gasoline spark-ignition vehicles (>97%), have a median age of 7 years, and are equipped with three-way catalytic converters (Statistics Canada, 2009). Larger vehicles such as trucks and buses are diesel-powered and can be equipped with a wide range of emission control technologies depending on the age of the vehicle.

5.4.3 Field Sites
UFP concentrations were measured at a central site, two downtown sites (A, B) and four residential sites (C, D, E, F) in Toronto for 1-3 week periods (Figure 5-1). Each neighbourhood was selected based on its traffic volume, distance from the central site and distance from major roadways (Table 5-1).
Figure 5-1 Map of Toronto and location of downtown sites (central, A, B) and residential sites (C, D, E, F) monitoring sites. All highways and major arterial roadways are shown.

The central site was located at the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR) in downtown Toronto. The air sampling inlet was 6m above the ground and 20m from a major arterial roadway with a weekday traffic volume of 20,000 vehicles per day. UFP has been continuously measured at the central site since March 2006 (Sabaliauskas et al., 2012). Measurements at the two downtown (A, B) and four residential (C, D, E, F) sites were collected at the rear of the participants’ homes. Participants were instructed to abstain from activities that can produce smoke such as cooking with a barbecue or operating gasoline powered lawn maintenance equipment near the sampling inlet.
Table 5-1 The characteristics of the downtown and residential field sites. Traffic volumes: major arterial roadways > 20,000, collector roadways = 2,500 – 8,000, and local roadways < 2,500 vehicles per day (City of Toronto Transportation Services Division, 2010). 24-hour traffic volumes for major arterial and collector roadways were obtained from traffic counts conducted by City of Toronto. a: measurements were collected from a 2nd floor apartment. All measurements were collected at the rear of the property.

<table>
<thead>
<tr>
<th>Field Site</th>
<th>Neighbourhood Type</th>
<th>Roadway Type</th>
<th>Traffic Volume (vehicles-day-1)</th>
<th>Nearest Major Arterial Road (m)</th>
<th>Distance to Central Site (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Downtown</td>
<td>Major Arterial</td>
<td>22,500</td>
<td>25</td>
<td>0.4</td>
</tr>
<tr>
<td>B</td>
<td>Downtown</td>
<td>Local</td>
<td>&lt;2,500</td>
<td>90</td>
<td>0.5</td>
</tr>
<tr>
<td>C</td>
<td>Residential</td>
<td>Local</td>
<td>&lt;2,500</td>
<td>235</td>
<td>6.2</td>
</tr>
<tr>
<td>D</td>
<td>Residential</td>
<td>Collector</td>
<td>6,500</td>
<td>800</td>
<td>15.7</td>
</tr>
<tr>
<td>E</td>
<td>Residential</td>
<td>Local</td>
<td>&lt;2,500</td>
<td>765</td>
<td>15.2</td>
</tr>
<tr>
<td>F</td>
<td>Residential</td>
<td>Local</td>
<td>&lt;2,500</td>
<td>175</td>
<td>9.6</td>
</tr>
<tr>
<td>Central Site</td>
<td>Downtown</td>
<td>Major Arterial</td>
<td>20,000</td>
<td>20</td>
<td>-</td>
</tr>
</tbody>
</table>

5.4.4 Instrumentation

Multiple particle sizing and counting instruments were deployed at the sites for the study (Figure 5-2). UFP was monitored continuously at the central site using a Fast Mobility Particle Sizer (FMPS, Model 3091, TSI Incorporated, Shoreview, MN, USA). The FMPS provided particle number based size distributions over the range of 5.6 to 560nm with 1-minute time resolution. Due to low counting efficiencies in the upper and lower size bins of the FMPS, only size bins between 8 and 300nm were included in this study. The size distribution was corrected by applying the procedure described by Jeong and Evans (2009). The resulting corrected size distributions were similar to those obtained using a Scanning Mobility Particle Sizer (SMPS).
The residential field sites were equipped with one of three Water Condensation Particle Counters (WCPC, Model 3781, TSI, Shoreview, MN, USA). The WCPC reported the total number of particles per cubic centimeter of air for particles between 6 and 1000 nm. The instrument was operated at 1-minute time resolution. The FMPS and all three WCPCs showed excellent correlation ($R^2=0.92-0.95$) for particle size bins between 8 and 300 nm; however, the FMPS consistently measured 1.4 times higher than the WCPCs, despite the WCPCs higher size cut-off. Thus the WCPC data were scaled so as to match that from the FMPS. The critical orifice was cleaned after every seven days of field operation. Studies that have previously used this WCPC near traffic sources have not reported significant issues with its operation (Barone and Zhu, 2008; Buonocore et al., 2009).

Measurements at residential sites C and D were repeated in May and June 2009 using a Scanning Mobility Particle Sizer (SMPS, Model 3936, TSI Incorporated, Shoreview, MN, USA) with a nano-Differential Mobility Analyzer. The SMPS measured the total number of particles between 2.9 and 107 nm at 2-minute time resolution. To match the size range reported by the FMPS, only size bins between 8 and 107 nm were included in this study. All measurements were corrected for diffusion and multiple charging losses. The corrected FMPS and SMPS were compared before
and after field deployment and showed excellent agreement in both total number concentration (slope = 1.05, \(R^2 = 0.96\)) and geometric mean diameter values (slope = 0.98, \(R^2 = 0.84\)).

5.4.5 Other Data Sources:
Hourly temperature (T), relative humidity (RH), wind speed (WS), wind direction (WDir), and cloud cover measurements were obtained from the meteorological station located at Pearson International Airport (Environment Canada, 2011). Additional air pollutant data was obtained from the Ontario Ministry of the Environment Air Quality Monitoring Network (MOE). The downtown Toronto site located 1km to the northeast of the central site provided hourly PM\(_{2.5}\), CO, NO, NO\(_2\), NO\(_x\), O\(_3\) and SO\(_2\) concentration data (MOE 2010).

5.4.6 Data Analysis: Wavelet Decomposition
The local-neighbourhood and urban-regional contributions to the observed PN concentration were calculated using the wavelet decomposition background subtraction algorithm developed by Klems et al., (2010). Wavelet decomposition is a widely used signal processing technique that separates a time series (or signal) into its high frequency (level of detail) and low frequency (level of approximation) components. This method is advantageous because it provides the ability to partition and analyze the signal based on frequency bands, and reconstruct the portions of the signal that are of interest. In the case of PN concentration data, the high frequency portion represents short lived PN concentration spikes that are believed to originate from local sources, such as nearby high emitting vehicles, while the low frequency portion represents contributions that vary over longer time scales, such as by time of day and between days. This longer timescale signal is believed to correspond to sources with larger urban –regional spatial scales, such as city wide traffic patterns or regional nucleation events.

Wavelet decomposition in its simplest form involves passing a time series of data through a single high frequency and a single low frequency filter. This first level of decomposition creates the first level of detail (d\(_1\)) and approximation (a\(_1\)) representations of the time series data. To further decompose the data, a\(_1\) can be passed through a second pair of high and low pass filters to yield a second level of detail (d\(_2\)) and approximation (a\(_2\)). The process can be repeated multiple times (n) until the specified level of signal decomposition is reached or the signal cannot be further decomposed. Each n-level of decomposition corresponds to events that have different
time scales that increase by 2n. The original data can be reconstructed by summing each level of detail (d₁-dₙ) and the final level of approximation (aₙ).

The background subtraction algorithm developed by Klems et al., (2010) decomposed the PN concentration to the d₈ level. The corresponding level approximation, a₈, was subtracted from the original signal resulting in the first approximation of the PN concentration spikes S₁. In cases where the value of S₁ was less than zero, the corresponding S₁ values were replaced by the PN concentration values from the original signal. This corrected S₁, redefined as baseline estimation b₁, was decomposed to d₈ and the above correction procedure was repeated. The process was repeated until the baseline estimation bₙ and bₙ₊₁ differed by less than 1%. The PN concentration spikes contribution was calculated by subtracting bₙ₊₁ from the original signal.

This study applied the same algorithm as Klems et al., (2010) with slight differences in the input parameters. First, Klems et al., (2010) sought to reduce the instrument noise associated with the FMPS operating at 1-second time-resolution by calculating an 11-point moving average. In this study, the time resolution was significantly lower and the data was simply input into the algorithm. Second, the time scales that yielded the best baseline estimation in the Klems et al., (2010) study were between 256 and 512 seconds (level 8 decomposition). In this study, multiple levels of decomposition were explored (level 2 to 10) and the best baseline fit occurred using a level 5 decomposition for the 2-minute time resolution data (SMPS and matching FMPS) and a level 6 decomposition for the 1-minute time resolution data (WCPC and matching FMPS). Using these two different decomposition levels (5 vs. 6) accounted for the different time resolutions of the measurements (2 vs. 1 min) and thereby yielded a baseline corresponding to events with time constants longer than 1 to 2 hours. All wavelet decomposition calculations were performed using the Wavelets Toolbox in the MATLAB (R2011b) programming environment.

5.4.7 Coefficient of Divergence

The degree of spatial heterogeneity between different field sites was represented using the coefficient of divergence (COD) (Equation 5-1). Mathematically, COD is defined as
Equation 5-1

\[ COD_{jk} = \sqrt{\frac{1}{p} \sum_{i=1}^{p} \left( \frac{x_{ij} - x_{ik}}{x_{ij} + x_{ik}} \right)^2} \]

where \( x_{ij} \) and \( x_{ik} \) represent the concentration of the pollutant of interest for a sampling interval \( i \) at locations \( j \) and \( k \), and where \( p \) is the number of observations (Wongphatarakul et al., 1998). A COD value of 0 means that there was no difference between the two sampling locations and a value near 1 would mean that there was large heterogeneity between the two locations. If the COD value is less than 0.2, then the two field sites are considered to be homogeneous. If the COD value is greater than 0.4, then the two field sites exhibit heterogeneity.

5.4.8 Traffic and Geographical Modelling

All traffic counts were extracted from a traffic micro-trip simulation model that encompassed the entire GTA during the morning and afternoon rush hour periods (Roorda et al., 2010). The model was developed using a combination of time activity patterns of GTA residents and logs from commercial distribution centers. Vehicle counts were available for all roadways classified as expressways, major arterial and minor arterial. Available vehicle categories included passenger cars, light duty diesel, medium duty diesel, and heavy duty diesel. Traffic counts were calibrated against existing cordon counting stations located across the GTA. All data extraction and geographical modelling was performed in ArcGIS Version 10.1 (ESRI, Redlands, CA).

5.5 Results

5.5.1 Wavelet Decomposition: Background Subtraction

Figure 5-3 demonstrates the efficacy of applying the wavelet decomposition algorithm to separate the original PN concentration time series into its low (urban-regional scale) and high (local-neighbourhood scale) frequency components. The upper panel shows the original PN concentration data collected at the central site. The middle panel presents the variation of the urban-regional scale PN concentration as calculated through the wavelet decomposition algorithm. The lower panel shows the variation of the local-neighbourhood scale PN concentration calculated by subtracting the background PN concentration (middle panel) from
the original data (upper panel). The performance of the wavelet decomposition algorithm can be further examined in Figure D-1.

![Graph showing variation of PN concentration](image)

**Figure 5-3** Variation of PN concentration (upper panel), urban-regional scale (middle panel) and local-neighbourhood scale (lower panel) at the central site between August 1st 2008 and August 19th 2008. The middle panel shows the output from the wavelet decomposition algorithm. The lower panel is the urban-regional scale PN concentration subtracted from the original data.

### 5.5.2 PN Concentration Behaviour at the Central Site

Table 5-2 shows the average, standard deviation, minimum, 25th percentile, median, 75th percentile and maximum values for the original data, the urban-regional scale and the local-neighbourhood scale on both working days (weekdays) and non-working days (weekends) at the
central site. Notably, the PN concentration on weekends was only 10% lower than those observed on weekdays. Although traffic volume is typically 20-25% lower in the downtown core of Toronto on weekends vs. weekdays PN concentrations did not always follow the same pattern. Numerous photochemical nucleation events occurred during the summer of 2008 and these events occurred more often in Toronto on weekends. Consequently, weekends often had at least one day with elevated PN concentrations that were not directly related to the traffic and as a result, weekends and weekdays had similar average UFP concentrations for both the original and the urban-regional scale PN concentration data. In contrast, the local-neighbourhood scale PN concentration was 28% lower on weekends than on weekdays (p<0.001) at the central site. On average, the local-neighbourhood scale PN concentration contributed 17% to the total observed PN concentration at the central site. However, this contribution varied throughout the measurement period likely due to variation in local traffic.

Throughout the summer of 2008, the predominant wind directions were northwesterly, westerly and southwesterly. Toronto experienced only a few brief periods with winds originating from the east or northeast (~10% of the time). Figure 5-4 shows the wind directionality dependence of the original PN concentration data, the urban-regional scale, and the local-neighbourhood scale. Notably, the original data (panel A) and the urban-regional scale (panel B) PN concentration exhibited a stronger southerly bias while the local-neighbourhood scale PN (panel C) concentration exhibited a northerly wind bias. The local-neighbourhood scale PN concentration ratio (panel D) ranged in value between 9% and 25% with the largest contribution occurring when Toronto experienced northerly winds.
Table 5-2 Summary of measurements collected at the central site between June 13th and August 19th 2008. U-R: urban-regional scale. L-N: local-neighbourhood scale.

<table>
<thead>
<tr>
<th>Day Type</th>
<th>Parameter</th>
<th>Average</th>
<th>Stdev</th>
<th>Min</th>
<th>25th</th>
<th>Median</th>
<th>75th</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (N=78,476 minutes)</td>
<td>Original (cm$^3$)</td>
<td>17,230</td>
<td>10,090</td>
<td>990</td>
<td>10,870</td>
<td>15,690</td>
<td>21,670</td>
<td>537,260</td>
</tr>
<tr>
<td></td>
<td>U-R (cm$^3$)</td>
<td>13,800</td>
<td>6,670</td>
<td>840</td>
<td>8,740</td>
<td>13,250</td>
<td>17,720</td>
<td>45,830</td>
</tr>
<tr>
<td></td>
<td>L-N (cm$^3$)</td>
<td>3,430</td>
<td>6,900</td>
<td>0</td>
<td>460</td>
<td>1,670</td>
<td>4,030</td>
<td>521,040</td>
</tr>
<tr>
<td></td>
<td>U-R Contribution (%)</td>
<td>83</td>
<td>16</td>
<td>3</td>
<td>76</td>
<td>88</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>L-N Contribution (%)</td>
<td>17</td>
<td>16</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>24</td>
<td>97</td>
</tr>
<tr>
<td>Weekdays (N=52,555 minutes)</td>
<td>Original (cm$^3$)</td>
<td>17,780</td>
<td>10,820</td>
<td>990</td>
<td>11,290</td>
<td>16,010</td>
<td>22,220</td>
<td>537,260</td>
</tr>
<tr>
<td></td>
<td>U-R (cm$^3$)</td>
<td>13,990</td>
<td>6,590</td>
<td>840</td>
<td>9,270</td>
<td>13,360</td>
<td>17,800</td>
<td>43,270</td>
</tr>
<tr>
<td></td>
<td>L-N (cm$^3$)</td>
<td>3,790</td>
<td>7,980</td>
<td>0</td>
<td>460</td>
<td>1,750</td>
<td>4,300</td>
<td>521,040</td>
</tr>
<tr>
<td></td>
<td>U-R Contribution (%)</td>
<td>83</td>
<td>17</td>
<td>3</td>
<td>75</td>
<td>88</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>L-N Contribution (%)</td>
<td>17</td>
<td>17</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>25</td>
<td>74</td>
</tr>
<tr>
<td>Weekends (N=25,921 minutes)</td>
<td>Original (cm$^3$)</td>
<td>16,120</td>
<td>8,320</td>
<td>2,100</td>
<td>9,870</td>
<td>15,090</td>
<td>20,690</td>
<td>111,100</td>
</tr>
<tr>
<td></td>
<td>U-R (cm$^3$)</td>
<td>13,410</td>
<td>6,830</td>
<td>2,100</td>
<td>7,740</td>
<td>13,010</td>
<td>17,590</td>
<td>45,830</td>
</tr>
<tr>
<td></td>
<td>L-N (cm$^3$)</td>
<td>2,710</td>
<td>3,810</td>
<td>0</td>
<td>450</td>
<td>1,560</td>
<td>3,550</td>
<td>99,590</td>
</tr>
<tr>
<td></td>
<td>U-R Contribution (%)</td>
<td>85</td>
<td>15</td>
<td>8</td>
<td>77</td>
<td>88</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>L-N Contribution (%)</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>23</td>
<td>92</td>
</tr>
</tbody>
</table>
Table 5-3 shows the Spearman Rank Order correlation coefficients between the 24-hour PN concentrations from the original data, urban-regional scale, and local-neighbourhood scale and other monitored air pollutants at the nearby MOE monitoring station. During the summer months, a dominant source of SO$_2$ in Toronto was coal combustion from power plants located to the south and southwest of the city. In conjunction with elevated concentrations of SO$_2$, southerly winds are typically associated with elevated concentrations of secondary pollutants such as PM$_{2.5}$ and O$_3$. Therefore, elevated concentrations of SO$_2$, PM$_{2.5}$ and O$_3$ imply Toronto is experiencing regional transport and/or enhanced photochemistry. The local-neighbourhood scale...
PN concentration exhibited positive correlation with NO ($r_s = 0.53$) and negative correlation with O$_3$ ($r_s = -0.69$), SO$_2$ ($r_s = -0.49$), CO ($r_s = -0.48$) and PM$_{2.5}$ ($r_s = -0.40$).

Table 5-3 The Spearman Rank Order Correlation coefficients between the 24-hour average PN concentration and pollutants collected at the Ontario Ministry of the Environment station located 1km from the central site. All reported coefficients are statistically significant to a confidence level of 95%. U-R: PN concentration attributable to the urban-regional scale. L-N: PN concentration attributable to the local-neighbourhood scale.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Original Data</th>
<th>U-R</th>
<th>L-N</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM$_{2.5}$</td>
<td>-</td>
<td>-</td>
<td>-0.40</td>
</tr>
<tr>
<td>CO</td>
<td>0.54</td>
<td>0.63</td>
<td>-0.48</td>
</tr>
<tr>
<td>NO</td>
<td>0.29</td>
<td>-</td>
<td>0.53</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>0.49</td>
<td>0.47</td>
<td>-</td>
</tr>
<tr>
<td>NO$_x$</td>
<td>0.47</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>O$_3$</td>
<td>-</td>
<td>0.37</td>
<td>-0.69</td>
</tr>
<tr>
<td>SO$_2$</td>
<td>0.38</td>
<td>0.48</td>
<td>-0.49</td>
</tr>
<tr>
<td>T</td>
<td>-</td>
<td>-</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

5.5.3 PN Concentration Behaviour at the Field Sites

Table 5-4 shows the paired mean PN concentration ± 95% confidence interval for the original PN concentration data and the percent contribution of the urban-regional and local-neighbourhood scale sources to the total PN concentration. The average PN concentrations for the original, urban-regional scale and local-neighbourhood scale are presented in Supporting Information Table D-2. The PN concentration was consistently higher at the central site than the field sites on both weekdays and weekends. No significant differences were observed between the observed local-neighbourhood scale contribution on weekday and weekends.
Table 5-4 PN concentration ± 95% confidence interval for the original PN concentration data and the calculated contribution of the urban-regional (U-R) and local-neighbourhood (L-N) scale sources at each field site and the central site. WD: weekdays, WE: weekends. C1 and D1 = Sites C and D during the 2008 measurement period. C2 and D2 = Sites C and D during the 2009 measurement period with a SMPS. *Not statistically different from weekend % contribution ratio. Note 1: Only periods with valid data at both the field site and the central site were included. Note 2: The U-R and L-N ratios at all the field sites were statistically different from those observed at the Central Site at the 95% confidence level.

<table>
<thead>
<tr>
<th>Site</th>
<th>Day Type</th>
<th>N (hours)</th>
<th>Field Site Original Data (cm$^{-3}$)</th>
<th>U-R (%)</th>
<th>L-N (%)</th>
<th>Central Site Original Data (cm$^{-3}$)</th>
<th>U-R (%)</th>
<th>L-N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WD</td>
<td>418</td>
<td>15390 ± 120</td>
<td>81*</td>
<td>19*</td>
<td>20800 ± 120</td>
<td>89</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>191</td>
<td>15500 ± 170</td>
<td>81</td>
<td>19</td>
<td>20350 ± 120</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>WD</td>
<td>118</td>
<td>15610 ± 220</td>
<td>83</td>
<td>17</td>
<td>20060 ± 220</td>
<td>86</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>43</td>
<td>21040 ± 680</td>
<td>81</td>
<td>19</td>
<td>24000 ± 200</td>
<td>89</td>
<td>11</td>
</tr>
<tr>
<td>C1</td>
<td>WD</td>
<td>184</td>
<td>8820 ± 140</td>
<td>85</td>
<td>15</td>
<td>13150 ± 250</td>
<td>79*</td>
<td>21*</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>48</td>
<td>6450 ± 150</td>
<td>83</td>
<td>17</td>
<td>8940 ± 170</td>
<td>79</td>
<td>21</td>
</tr>
<tr>
<td>C2</td>
<td>WD</td>
<td>132</td>
<td>8410 ± 320</td>
<td>86</td>
<td>14</td>
<td>14320 ± 340</td>
<td>79*</td>
<td>21*</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>94</td>
<td>8890 ± 310</td>
<td>84</td>
<td>16</td>
<td>14030 ± 380</td>
<td>77</td>
<td>23</td>
</tr>
<tr>
<td>D1</td>
<td>WD</td>
<td>175</td>
<td>9450 ± 190</td>
<td>77</td>
<td>23</td>
<td>14310 ± 220</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>65</td>
<td>5750 ± 190</td>
<td>68</td>
<td>32</td>
<td>9380 ± 130</td>
<td>81</td>
<td>19</td>
</tr>
<tr>
<td>D2</td>
<td>WD</td>
<td>120</td>
<td>11460 ± 260</td>
<td>85</td>
<td>15</td>
<td>16060 ± 400</td>
<td>81*</td>
<td>19*</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>38</td>
<td>10550 ± 400</td>
<td>83</td>
<td>17</td>
<td>13470 ± 450</td>
<td>81</td>
<td>19</td>
</tr>
<tr>
<td>E</td>
<td>WD</td>
<td>125</td>
<td>3950 ± 50</td>
<td>86</td>
<td>14</td>
<td>15170 ± 300</td>
<td>79*</td>
<td>21*</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>68</td>
<td>4180 ± 90</td>
<td>84</td>
<td>16</td>
<td>12820 ± 290</td>
<td>79</td>
<td>21</td>
</tr>
<tr>
<td>F</td>
<td>WD</td>
<td>169</td>
<td>13790 ± 230</td>
<td>85</td>
<td>15</td>
<td>16810 ± 220</td>
<td>78</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>96</td>
<td>14560 ± 260</td>
<td>87</td>
<td>13</td>
<td>16610 ± 220</td>
<td>81</td>
<td>19</td>
</tr>
</tbody>
</table>

On weekdays, the percent contribution of the local–neighbourhood signal at the central site was generally higher than that at the residential sites (C, D, E, F), consistent with the higher traffic at the central site. However, the percent contribution of the local-neighbourhood scale sources was not always lower on weekends than on weekdays at the field sites. Household activities such as lawn maintenance with gasoline-powered equipment, cooking with a barbeque, burning candles, and smoking are common in residential areas. Although the study participants abstained from these activities while the equipment was in their backyard, the activities of their neighbours and the surrounding neighbourhood could not be controlled. The presence of these additional UFP sources in the residential areas and lack of similar sources near the central site may account for
the observed differences in the contribution of the local-neighbourhood scale sources. In addition to local sources in residential areas, photochemical nucleation events occurred on several weekend days throughout the summer of 2008. This effect was most pronounced at the sites with the shortest sampling times (Site B and F).

5.5.4 Correlation and Coefficient of Divergence
Table 5-5 presents the Spearman’s Rank Order correlation coefficients (r_s) and coefficients of divergence (COD) between the central site and residential sites at 1-minute time resolution (at 2 minute time resolution at sites C2 and D2). In all cases (except Site D), higher correlation was observed between the urban-regional PN concentration data than the original time series data. The local-neighbourhood scale sources observed at the field sites exhibited little correlation with the central site. The majority of the correlation observed between sites in the original data was driven primarily by the urban-regional signal. The local-neighbourhood scale sources exhibited little correlation with the PN concentration spikes observed in downtown Toronto. In most cases, the COD value for the urban-regional PN concentration data was similar to or lower than the corresponding value for the original PN concentration time series. In most cases, the COD value ranged from 0.2 and 0.4 for the background data suggesting each field site exhibited both homogeneous and heterogeneous tendencies. The local-neighbourhood scale portion exhibited the most heterogeneity when compared against the central site and ranged from 0.64 to 0.78.
Table 5-5 Spearman’s Rank Order correlation coefficients ($r_s$) coefficients of divergence (COD) are reported for each field site at 1-minute time resolution (C2 and D2 = 2-minute time resolution). All correlation coefficients were statistically significant at the 95% confidence level. COD<0.2: PN concentrations at the field site and central site were homogeneous, 0.2<COD<0.4: PN concentration at both sites exhibited homogeneous and heterogeneous tendencies, COD>0.4: PN concentration at both sites were heterogeneous. U-R: urban-regional scale, L-N: local-neighbourhood scale.

<table>
<thead>
<tr>
<th>Site</th>
<th>Original Data $r_s$</th>
<th>U-R Data $r_s$</th>
<th>L-N Data $r_s$</th>
<th>Original Data COD</th>
<th>U-R Data COD</th>
<th>L-N Data COD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.71</td>
<td>0.84</td>
<td>0.28</td>
<td>0.27</td>
<td>0.28</td>
<td>0.64</td>
</tr>
<tr>
<td>B</td>
<td>0.70</td>
<td>0.85</td>
<td>0.26</td>
<td>0.27</td>
<td>0.25</td>
<td>0.63</td>
</tr>
<tr>
<td>C1</td>
<td>0.61</td>
<td>0.74</td>
<td>0.12</td>
<td>0.29</td>
<td>0.23</td>
<td>0.69</td>
</tr>
<tr>
<td>C2</td>
<td>0.65</td>
<td>0.79</td>
<td>0.14</td>
<td>0.34</td>
<td>0.26</td>
<td>0.81</td>
</tr>
<tr>
<td>D1</td>
<td>0.22</td>
<td>0.13</td>
<td>0.10</td>
<td>0.43</td>
<td>0.46</td>
<td>0.67</td>
</tr>
<tr>
<td>D2</td>
<td>0.35</td>
<td>0.44</td>
<td>0.05</td>
<td>0.31</td>
<td>0.25</td>
<td>0.70</td>
</tr>
<tr>
<td>E</td>
<td>0.37</td>
<td>0.40</td>
<td>0.12</td>
<td>0.55</td>
<td>0.52</td>
<td>0.73</td>
</tr>
<tr>
<td>F</td>
<td>0.41</td>
<td>0.57</td>
<td>0.08</td>
<td>0.27</td>
<td>0.21</td>
<td>0.67</td>
</tr>
</tbody>
</table>

5.5.5 UFP and Traffic Volume

For each field site, traffic volumes for passenger vehicles, light duty diesel, medium duty diesel and heavy duty diesel were extracted using multiple buffer sizes (50, 100, 200, 300, 500, 750, 1000, 1500, 2000, 2500, 3000 and 4000m). With the exception of heavy-duty diesel vehicles, all traffic class counts were highly correlated ($R^2>0.93$); therefore, all vehicle classes were summed (including heavy-duty diesel trucks) and a single traffic parameter was assigned to each field site. PN concentration ratios were calculated between paired measurements at each field site and the central site for the original data, urban-regional scale and local-neighbourhood scale to normalize for the temporal variability. From these a single median field-to-central-site ratio was calculated for each site for all the weekdays.

The strongest relationship was observed between the median PN concentration ratios associated with the local-neighbourhood signals and the total number of vehicles (including heavy-duty diesel) within a 2.5km buffer of the field site (Figure 5-5, upper panel) and between the local-neighbourhood scale PN concentration percent contribution and traffic counts within the same 2.5km buffer (Figure 5-5, lower panel). The median urban-regional scale ratios and the original
PN concentration ratios exhibited no statistically significant correlation with total traffic counts at 2500m or any other buffer size. The absence of correlation between traffic and the original PN concentration data speaks to the value of isolating the local-neighbourhood signal when exploring the spatial patterns of traffic related pollutants, such as ultrafine particles, across urban areas.

Figure 5-5 Upper panel: median local-neighbourhood scale PN concentration ratios regressed against the number of vehicles within a 2.5km buffer of the field site. Relationship between the local-neighbourhood scale and total traffic counts: \( y = 1.35 \times 10^{-6} x + 0.10 \). Lower panel: the local-neighbourhood scale percent contribution regressed against the number of vehicles within a 2.5km buffer of the field site. The relationship between the local-neighbourhood scale percent contribution and total traffic counts \( y = 5.45 \times 10^{-6} x + 12.99 \). * Site D1 was not included in the regression due to suspected influences of non-traffic related sources.
UFP has been monitored continuously at the central site (located 16m from a major roadway) since March 2006 with intermittent video recording of the traffic. From this visual record at the central site, the majority of the PN concentration spikes occur when trucks and buses, presumably diesel-powered, pass the monitoring site. However, not all large vehicles produce these intense spikes and occasional passenger vehicles also produce spikes. Therefore, the local-neighbourhood scale signal is believed to represent only particles from “heavy-emitter” vehicles that pass within the vicinity of the monitoring sites and thus only a portion of the overall vehicle related emissions. The majority of vehicles do not produce resolvable spikes and thus their contribution is likely mixed with the urban-regional scale signal. Further, at the residential sites, other combustion sources may result in location specific PN concentration spikes that can vary through time (i.e. site D). Additional data mining approaches are being explored to isolate and quantify the contributions of low-emitting vehicles and other combustion related sources to the overall ultrafine PN concentrations in cities.

Finally, the coefficient of divergence was calculated to quantify the spatial heterogeneity of the local-neighbourhood and urban to regional scale PN concentrations between the field sites and the central site. Due to the intermittent nature of the PN concentration spikes at the roadside site, it may prove to be beneficial to pair the residential sites with a less traffic influenced background site.

5.6 Conclusions

This study explored the application of wavelet decomposition as a means to quantify the impacts of local sources on communities across Toronto. The PN concentration time series was separated into two signals that operate over different temporal and spatial scales: the low frequency urban-regional scale processes and the high frequency local-neighbourhood scale sources signals. At the field sites, local-neighbourhood scale sources and signals contribute between 13 and 32% to the total PN concentration. This high frequency portion of the time series exhibited low correlation and high heterogeneity (COD>0.4) with the central site. In contrast, the urban-regional scale portion of the time series exhibited high correlation with the central site. Meteorological parameters such as temperature, relative humidity and wind speed were not good predictors of the relative contributions of each of the abovementioned sources. However, traffic volume within a 2.5km buffer of each field site explained 87% of the variability observed.
between field sites. This study has demonstrated that wavelet decomposition may be a useful tool to estimate the contribution of local-neighbourhood scale and urban-regional scale sources and processes to urban residents’ UFP exposure. By separating the PN concentration time series into its high and low frequency components, a more defined population exposure assessment can be performed.

5.7 Acknowledgements
Funding for SOCAAR was provided by the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund. This work was supported by the Ontario Ministry of the Environment's Best in Science Research Program. Student stipends were provided by the Ontario Graduate Scholarships in Science and Technology. Special thanks to Matt Roorda for providing access to the micro-trip simulation traffic model.
Chapter 6
The Development of a Land-Use Regression Model for Ultrafine Particles in Toronto, Canada

6.1 Executive Summary
The paper presented in this chapter was submitted to Atmospheric Environment in July 2014 and will differ from the final published version:


Land-use regression is an empirical modeling technique that relates geographical characteristics to outcome of interest at a study location. Land-use regression models apply the fundamental assumption that the air quality within a given study area is affected by the how the surrounding land is used. For example, if the land is zoned as residential, the emissions profile will differ from that of an industrial area or that of a heavily trafficked highway. This chapter seeks to determine if land-use regression modeling can be applied to UFP data to quantify the range of possible concentrations that exist across Toronto (Objective 4).

In this chapter, the first UFP land-use regression model for Toronto is described. This work aims to provide a first approximation of the potential spatial variation of UFP across the city. This study is unique because it capitalizes on a hybrid fixed-mobile monitoring strategy to quantify the UFP concentration gradients. This work attempts to answer the following questions:

1. Can land-use regression be applied to a mobile UFP data set to provide information about the spatial gradients of UFP across Toronto?
2. What is the range of possible UFP concentrations across the city?
3. How representative are these concentrations of longer term measurements?
6.1.1 Author Contributions

This paper benefited from the contributions of the following co-authors Cheol-Heon Jeong, Xiaohong Yao, Christopher Reali, Tim Sun and Greg Evans. Cheol-Heon Jeong provided instrumentation support by operating the Fast Mobility Particle Sizer and assisted with data interpretation during the development phases of the manuscript. Xiaohong Yao provided sampling advice and feedback on the initial manuscript. Christopher Reali and Tim Sun provided field support for the mobile phase of the study. My contributions included operating the equipment at the fixed monitoring sites, collecting mobile field measurements, analyzing the data set and performing the geographical modeling. All manuscript writing was performed by me with input from Professor Evans.

6.2 Abstract

This study applies land-use regression (LUR) to characterize the spatial distribution of ultrafine particles (UFP) in a large city. Particle number (PN) concentrations were measured in residential areas around Toronto, Canada, between June – August 2008. A combination of fixed and mobile monitoring was used to assess spatial gradients between and within communities. The fixed monitoring locations included a central site, two downtown sites, and four residential sites located 6 to 15km from the downtown core. The mobile data included average PN concentrations collected on 112 roadways from 10 study routes that were repeated on three separate days. The mobile data was used to create the land-use regression model while the fixed sites were used for validation purposes. The predictor variables that best described the spatial variation of UFP ($R^2 = 0.72$, validated $R^2 = 0.68$) included population density within 300m, total resource and industrial area within 1000m, total residential area within 3000m, and major roadway and highway length within 3000m. The LUR model successfully predicted the afternoon PN concentration (slope = 0.96, $R^2 = 0.86$) and 24-hour average PN concentration (slope = 1.28, $R^2 = 0.72$) measured at seven fixed monitoring sites.

6.3 Introduction

Particulate matter (PM) exposure has been identified as a significant risk factor for the development of lung cancer and adverse health outcomes from cardiovascular and respiratory
causes (Pope et al., 2002). These associations have been observed worldwide despite different sources of PM and different pollutant mixtures. More recently, ultrafine particles (UFP: particles with diameters less than 100nm) have been scrutinized due to their preferential deposition in the deepest regions of the human respiratory tract (Daigle et al., 2003), ability to promote the production of inflammatory biomarkers (Araujo et al., 2008; Bräuner et al., 2007; Delfino et al., 2010), and possible translocation to secondary organs within the body (Elder et al., 2006; Oberdörster, 2004). The health consequences resulting from UFP exposure remain uncertain due to the lack of long-term and spatially resolved data (HEI, 2014).

UFP concentrations are influenced by numerous combustion related sources, secondary formation pathways and transformation processes that alter their number, shape, size and chemical composition. In urban areas, vehicles, residential heating, cooking and industrial activities are dominant sources of UFP (Aalto et al., 2005; Buzorius, et al., 1999; Wang et al., 2011a; Sabaliauskas et al., 2012). Previous investigations have found both the particle number (PN) concentration and size of UFP exhibit significant seasonal, diurnal and weekday-weekend variation (Cyrys et al., 2008; Puustinen et al., 2007; Wang et al., 2011b). Although the number of urban areas with long-term UFP datasets is growing, the majority of these data are from single monitoring sites. Assessing UFP concentration gradients across an urban area remains challenging. Efforts to characterize UFP concentration gradients in urban areas have measured UFP simultaneously at multiple sites (Buzorius et al., 1999; Cyrys et al., 2008; Johansson et al., 2007; Kim et al., 2002; Krudysz et al., 2009; Moore et al., 2009; Puustinen et al., 2007; Sabaliauskas et al., 2014), while walking along roadways with different traffic intensities (Kaur et al., 2006), and while driving (Kittelson et al., 2006; Pijola et al., 2006).

Land-use regression (LUR) is a commonly used modelling technique that establishes empirical relationships between air pollutant concentrations and geographical predictor variables. This technique has been applied to NO/NO\(_2\), PM\(_{2.5}\) and VOCs (reviewed by Hoek et al., 2008) and recently applied to UFP (Abernethy et al., 2013; Hoek et al., 2011; Patton et al., 2014; Rivera et al., 2012). Although the application of LUR methods presents an opportunity to characterize the range of UFP concentrations that may exist within a city, numerous challenges relating to instrumentation, logistics and statistics remain. To develop a reliable LUR model for any
pollutant, the number of sites is an extremely important parameter. Unlike the passive NO/NO₂ badges that enable reliable simultaneous sample collection over long time horizons (weeks) and at many sites (at least 20, but ~ 100 is common), existing UFP instrumentation is expensive and often requires substantial user intervention to collect accurate results. Consequently, existing UFP LUR models developed for large urban areas rely on measurements collected at many sites (>20) over short time horizons (hours) (Abernethy et al., 2013; Rivera et al., 2012) or at many sites (50) over longer time scales (weeks) but across multiple seasons (Hoek et al., 2011).

The abovementioned model development approaches result in a few unavoidable analysis challenges. Firstly, UFP exhibits a strong diurnal pattern from a combination of vehicle emissions and secondary particle formation (Jeong et al., 2010; Kulmala et al., 2004). Studies that rely on a short-duration measurement to characterize the UFP concentration at each site need to either collect measurements at all sites at roughly the same time of day or apply a correction factor to make the mornings and afternoons comparable. However, the shape of the diurnal trend can vary substantially between days and even between sites due to localized differences in meteorology and emissions (Jeong et al., 2010; Ketzel et al., 2004a; Sabaliauskas et al., 2012). Therefore, simply applying the same diurnal correction factor across all sites may not be appropriate. Secondly, UFP exhibits an inverse relationship with temperature with the highest PN concentrations typically observed during the cooler winter months (Sabaliauskas et al., 2012; Wang et al., 2011b). As a consequence, measurements need to be either collected during periods with similar meteorological conditions or corrected. Thirdly, the PN concentration can be impacted by the presence of larger particles (Kulmala et al., 2001). The particle mass concentration can vary on a seasonal basis with higher temperatures favouring the formation of secondary aerosol. As a result, despite measuring the UFP concentration under similar meteorological conditions, the PN concentration may be suppressed due to the presence of large particles on one day and enhanced by their absence on another. Finally, nucleation and growth events can dramatically impact the observed PN concentration during the afternoon periods and may dwarf any impacts from localized emissions.

This study describes the development of an LUR model for UFP in Toronto, Canada. UFP has been measured continuously in downtown Toronto since 2006. Long-term trends suggest a gradual reduction in the total number concentration between 2006 and 2011, likely due to
changes in the vehicle fleet, reductions in coal-fired power plant usage and the economic downturn that impacted the United States and Southern Ontario (Sabaliauskas et al., 2012). A recent analysis of UFP measurements collected at six fixed residential field sites in Toronto suggests that vehicle counts may be an important predictor of only a portion of the UFP concentration (Sabaliauskas et al., 2014). This study applies a hybrid methodology and relies on the measurements collected at the same six fixed field sites supplemented with mobile monitoring to better characterize and establish spatial relationships across the city. Specifically, an LUR model was created based upon mobile measurements and tested against measurements form the fixed sites.

6.4 Methodology

6.4.1 Study Design

UFP measurements were collected using a combination of fixed site and mobile monitoring (Figure 6-1) in Toronto, Canada. The fixed monitoring portion of the study was executed at a central site, two downtown sites (A, B) and four residential sites (C, D, E, F) for 1-3 weeks. The residential sites were located between 6 and 15km away from the downtown core (Sabaliauskas et al., 2014). The mobile phase of the study assessed spatial gradients within neighbourhoods by collecting UFP measurements while walking along roadways with different traffic densities near and between the fixed field sites. Two routes were designed within each neighbourhood to ensure that locations both upwind and downwind from major sources such as expressways, major arterial roads and railroads were visited. The maximum distance between the paired measurements varied by route but ranged between 0 and 6 km. Each route was followed on three separate days.
Figure 6-1 Map of the field sites and mobile monitoring routes

6.4.2 Instrumentation

Multiple particle sizing and counting instruments were deployed at the sites for the study (Figure 6-2). The Central Site was located at the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR) near downtown Toronto. The air sampling inlet was 6m above the ground and 20m from a major arterial roadway with a weekday traffic volume of 20,000 vehicles per day. UFP was monitored continuously at the Central site using a Fast Mobility Particle Sizer (FMPS, Model 3091, TSI Incorporated, Shoreview, MN, USA). The FMPS provided particle number based size distributions over the range of 5.6 to 560nm with 1-second time resolution. One-minute averaged data were extracted and hourly averages were calculated for hours with more than 70% of the data availability. The size distribution was corrected to match those obtained by a Scanning Mobility Particle Sizer by applying the procedure described by Jeong and Evans (2009). The counting efficiency of the FMPS for size bins less than 8nm and greater than 300nm was low. To ensure consistent results, only particles between 8 and 300nm were included in this study.
The residential field sites (A-F) were equipped with one of three Water Condensation Particle Counter (WCPC, Model 3781, TSI, Shoreview, MN, USA). This portable WCPC measured particles between 6 and 1000nm. The instrument was operated at 1-minute time resolution. Any observation flagged as erroneous due to out of range temperature or flow rate values were removed and not included in the analysis. Hourly average concentrations were calculated for hours with more than 70% data availability. Side-by-side comparisons between the three WCPCs and FMPS (size bins 8-300nm) showed excellent agreement ($R^2 = 0.92-0.95$). All WCPC measurements were scaled by a factor of 1.4 to ensure equivalency with the FMPS. The critical orifice was cleaned after every seven days of field operation.

UFP spatial gradients were measured using a handheld P-Trak Condensation Particle Counter (Model 8285, TSI Incorporated, Shoreview, Minnesota, USA). The P-Trak measured the total particle number concentrations (pt/cm$^3$) for particles between 20 and 1000nm. The P-Trak had an upper limit of 500,000 pt/cm$^3$ and logged data at 10-second time resolution. The P-Trak and FMPS exhibited good correlation for particles greater than 20nm when sampling 4m from a
A slope of 1.0 between the FMPS and P-Trak was observed for particles with diameters greater than 24nm suggesting that the size-cut off may be slightly higher than the factory reported value of 20nm. Given this slope of 1.0, the P-Trak data were used without any scaling. Matson et al. (2004) compared a P-Trak to a Condensation Particle Counter with a size cut-off of 10nm and observed good agreement to within 20% near traffic sources. Zhu et al. (2006) demonstrated that the P-Trak underestimates the total particle number concentration when sampling freshly emitted particles from traffic sources. The P-Trak was operated for no more than 3 hours on a single charge of alcohol under high relative humidity conditions (greater than 70%) to minimize any interference.

6.4.3 Data Analysis
6.4.3.1 Mobile Monitoring

A total of 112 roadways located along ten sampling routes were visited on three separate days between June 13th and August 20th 2008. To ensure sampling occurred under similar meteorological conditions, all measurements were collected on precipitation free days and on days with forecasted afternoon temperatures between 25 and 32°C. To minimize diurnal influences, measurements were collected on 15 weekday afternoons between 12:00 and 15:00. Two study routes were followed by simultaneously. Each sampler carried a GPS unit and log sheets to record their position relative to street addresses and other local features at 1-minute intervals. This manual position logging ensured that the sample locations were well defined in areas with tall buildings.

The midpoint on each visited road segment was identified and used as an average position on the roadway. Routes that followed the same road past multiple major arterial roads were divided to multiple segments at the major arterial road. The average UFP concentration was calculated for each road segment and matched to the average position on the road. The resulting UFP samples were between 5 and 10 minutes long per road segment. The UFP concentration that was used to produce the LUR model was a three visit average PN concentration for each road segment.
6.4.3.2 Central Site and Fixed Sites

A detailed description of the data treatment at the central site and fixed sites is available in Sabaliauskas et al., (2014). Briefly, at the Central Site and each fixed monitoring site, all valid paired data were included in the analysis. In each neighbourhood, intermittent sources of UFP were likely present but could not be removed from the data set and may have included emissions from garbage trucks, street cleaning, landscaping services, cooking, individual vehicles and construction. Data were removed from the residential field sites for times during which it was known that the site was influenced by barbequing or lawn maintenance activities as reported by the study volunteers. However, these activities tended to occur during the evenings and on weekends and were for the most part outside of the mobile sampling period of 12:00 to 15:00.

6.4.4 Study Method Validation

A Monte Carlo algorithm was developed to determine the probability that the short duration mobile data is representative of data collected at the fixed sites over longer time horizons. The algorithm was implemented as follows:

1) An overall mean PN concentration was calculated for each of the fixed sites based on all the five minute averaged data collected on weekdays between 12:00 and 15:00.

2) Three five minute averages were randomly selected for a given site from three separate days.

3) The mean value of the three random samples was calculated.

4) This process was repeated 100,000 times to ensure that a representative distribution of all possible combinations of the data was generated.

5) A probability distribution of simulated averages was obtained from the resulting from step 4.

6) Based on the probability distribution from step 5, the likelihood that a simulated average would be within ± 5%, 10%, … 95%, 100% of the overall average at each fixed site was calculated.
6.4.5 Geographical Variables

The UFP LUR was developed from the mobile measurement data using geographical and socioeconomic data from multiple sources (Table 6-1) including variables for roadway length, roadway distance, traffic density, land use classification, population characteristics, restaurant density and address density. Circular buffer sizes of 50, 100, 200, 300, 500, 750, 1000, 1500, 2000, 2500 and 3000 m were examined. All variables were extracted using ArcGIS software (Version 10.1, ESRI, Redlands, CA, 2011). Most predictor variables were derived for different buffer sizes by summing the length of lines (e.g. roads and highways), area (e.g. park area) or the number of points (e.g. number of restaurants) within the circle.

6.4.6 Model Development and Validation

Variable inclusion and exclusion parameters were determined using the model-building algorithm described in Henderson et al., (2007). 1) all variables were ranked by absolute strength of their correlation with the measured pollutant; 2) the highest ranking were identified in each subcategory; 3) variables exhibiting correlation with the highest ranked variable in each subcategory (Pearson’s r>0.6) were removed; 4) all remaining variables were entered into a forward stepwise linear regression; 5) variables with insignificant t-statistics (p>0.05) were removed; 6) Steps 4 and 5 were repeated until convergence and variables that contributed less than 1% to the R² were removed. The outcome was a multiple linear regression model of the form (Equation 6-1):

Equation 6-1

\[
\text{UFP Concentration} = \text{intercept} + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n + \varepsilon
\]

where \( \beta \) is the coefficient for each predictor variable X and \( \varepsilon \) is the random error term.

The LUR model was evaluated using leave-one-out (LOO) cross validation recreating the LUR model with n-1 sites. The predictive power of the LUR model was tested by predicting the UFP concentration measured at the central site and the six residential field sites.
Table 6-1 Examined variables for the PN concentration LUR for Toronto. Explored buffers included 50, 100, 200, 300, 500, 750, 1000, 1500, 2000, 2500 and 3000 m.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variable Type</th>
<th>Variable Categories</th>
<th>Data Source</th>
<th>Data Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic</td>
<td>Road Length</td>
<td>Expressway, major arterial, minor arterial, collector, local</td>
<td>CanMap Route Logistics (Ontario, 2008)</td>
<td>DMTI Spatial</td>
</tr>
<tr>
<td></td>
<td>Road Area</td>
<td>Expressway, major arterial, minor arterial, collector, local</td>
<td>CanMap Route Logistics (Ontario, 2008)</td>
<td>DMTI Spatial</td>
</tr>
<tr>
<td></td>
<td>Distance to nearest road</td>
<td>Expressways, major arterial</td>
<td>CanMap Route Logistics (Ontario, 2008)</td>
<td>DMTI Spatial</td>
</tr>
<tr>
<td></td>
<td>Traffic Density</td>
<td>Counts: passenger, light duty diesel, medium duty diesel, heavy duty diesel vehicles</td>
<td>Regional micro-trip simulation model</td>
<td>Roorda et al., 2010</td>
</tr>
<tr>
<td>Land Use</td>
<td>Land Use Area</td>
<td>Commercial, government &amp; institutional, open area, parks &amp; recreational, residential, resource &amp; industrial, waterbody</td>
<td>CanMap Route Logistics (Ontario, 2008)</td>
<td>DMTI Spatial</td>
</tr>
<tr>
<td>Restaurants</td>
<td>Restaurant Density</td>
<td>Counts: all restaurants, fast food restaurants, hotdog stands</td>
<td>DineSafe</td>
<td>City of Toronto, Open Data Project</td>
</tr>
<tr>
<td>Population</td>
<td>Population Density</td>
<td>people per km$^2$, dwellings per km$^2$</td>
<td>Canadian Census 2006</td>
<td>Statistics Canada (2007a)</td>
</tr>
</tbody>
</table>
6.4.7 UFP Surface Generation

To generate the UFP surface, universal kriging was performed using a sample of 4000 random points generated within the city limits of Toronto. Random points were separated by minimum distances of 200m. Maps were truncated to the maximum measured PN concentration during the mobile phase of the study.

6.5 Results

6.5.1 Mobile Measurements

An example of two simultaneous sampling routes in the downtown neighbourhood is shown in Figure 6-3. The starting location on this route was the downtown central site. When the distance between the two samplers was less than 500m, the PN concentration exhibited high correlation. As the distance between the two samplers increased, lower temporal correlation was observed. The average ± 95% confidence interval for major arterial (>20,000 vehicles per day) and local roads (<2000 vehicles per day) for each neighbourhood are reported in Table 6-2. In all neighbourhoods, the mean PN concentration on major arterial roadways was higher than on local roads (except Midtown 1, Route 1). The UFP concentration measured on downtown major arterial roadways were often 3 to 5 times higher than those in residential areas. Further, major arterial and local roadways located in downtown and mid-town neighbourhoods were significantly higher than those located in areas away from the downtown core.
Figure 6-3 A typical walking route in downtown Toronto. Route 1 visited a combination of major arterial (time = 0-40 minutes) and local roadways (time = 40-60 minutes) while Route 2 remained on major arterial roadways. The maximum distance between samplers was 2,800m.
Table 6-2 Mean ± 95% confidence interval for major arterial (>20,000 vehicles per day) and local roads (<2,000 vehicles per day). Routes 1 and 2 in each neighbourhood were followed on the same day and each reported mean includes three separate sampling days.

<table>
<thead>
<tr>
<th>Walking Route</th>
<th>Sample ID</th>
<th>Local Roads (pt/cm$^3$)</th>
<th>Major Arterial (pt/cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Toronto</td>
<td>1</td>
<td>9710 ± 440</td>
<td>15960 ± 1960</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9200 ± 530</td>
<td>12670 ± 1150</td>
</tr>
<tr>
<td>East Toronto</td>
<td>1</td>
<td>4970 ± 150</td>
<td>6620 ± 590</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7610 ± 430</td>
<td>13010 ± 3300</td>
</tr>
<tr>
<td>Mid-Town 1</td>
<td>1</td>
<td>17530 ± 500</td>
<td>17400 ± 610</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>20140 ± 570</td>
<td>23460 ± 1050</td>
</tr>
<tr>
<td>Midtown 2</td>
<td>1</td>
<td>10790 ± 330</td>
<td>17900 ± 1090</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11070 ± 670</td>
<td>20790 ± 1550</td>
</tr>
<tr>
<td>Downtown</td>
<td>1</td>
<td>15880 ± 1630</td>
<td>20070 ± 1040</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>N.A.</td>
<td>21910 ± 1190</td>
</tr>
</tbody>
</table>

6.5.2 Method Validation: Selected Measurement Period

The average diurnal trend observed at the central site and six residential sites is presented for weekdays in Figure 6-4. The diurnal trend presented for the central site encompasses the entire measurement period; however, the shape and magnitude were consistent even when truncated to match the sampling periods of the individual residential sites. The downtown sites (A, B) had similar diurnal behaviour as the central site, while the residential sites (C, D, E, F) exhibited distinctly different behaviour. Across the city, the minimum PN concentration occurred between 4:00 and 6:00 in the morning. The initial increase in PN concentration between 6:00 and 9:00 corresponded to the morning rush hour period. At most sites, the daily maximum PN concentration was observed between 12:00 and 15:00. Although afternoon PN concentration peaks are often associated with particle formation and growth events, simultaneous events of the same magnitude were not observed at the residential sites. On some days, slight increases in PN concentration were observed during the afternoon periods at the residential sites; however, the contribution of these formation events to the overall PN concentration appears to be small.
Figure 6-4 The diurnal behaviour of UFP observed at the central site and fixed residential monitoring sites

The mobile measurements were collected between 12:00 and 15:00 (Figure 6-4). At the residential sites (C, D, E, and F), the PN concentrations were significantly lower than those observed at the central site and sites A and B. These concentration differences were likely due to different emissions profiles surrounding the monitoring sites. Further, the PN concentration exhibited relatively stable behaviour between 12:00 and 15:00 (within 20%) suggesting that the daily maximums have been reached and this period offers an opportunity to study the spatial variation of UFP without the need to correct for diurnal variation. However, the resulting predictive model likely overestimates the 24-hour average.

6.5.3 Method Validation: Sample Time

Each mobile route was followed on three separate days to capture a range of possible conditions in each neighbourhood. The combination of mobile and fixed site monitoring enabled better quantification of the uncertainties surrounding the mobile measurements. Using the Monte Carlo simulation method described in the methods section, a probability distribution of averages was
obtained based random subsamples from the fixed monitoring sites. The percent of simulated averages that were within ±5%, 10% ... 100% of the measured mean are presented in Figure 6-5 for each field site. Notably, at the downtown sites (A, B), there exists high probability (>90%) that the simulated means are within 30% of the measured PN concentration average. However, at the residential sites, the probability that the simulated means were within 30% of the measured PN concentration average was between 54 and 74%.

![Graph](image)

**Figure 6-5** The probability that the average of three short-duration PN concentrations are within x% of the average PN concentration collected between 12:00 and 15:00 on weekdays throughout the summer of 2008.

Although the uncertainty at the residential sites (C, D, E, F) was considerably higher than observed at the downtown sites (A, B), this simulation may represent a worst case scenario for mobile study. At the downtown sites, the traffic intensity is fairly constant. Consequently, when measuring the PN concentration near the downtown core, the likelihood that an individual measurement is skewed due to a local source is low because the source is continuous. In contrast, at the residential sites, significantly fewer vehicles pass the monitoring sites. Since the flow of
vehicles is not continuous, these few vehicles have the ability to skew the PN concentration upwards over short time horizons. When walking in the residential areas, times when trucks passed the mobile samplers or times when lawn maintenance was occurring were recorded and subsequently removed from the data set. Since an equivalent record is unavailable for the fixed sites, intermittent sources may broaden the simulated probability distribution. As a result, the likelihood that the measured PN concentration data within the 30% of the longer term average is possibly higher than the simulated means suggest. In the interest of future studies, the probability short-term measurements collected in residential areas would be within 30% of the measured PN concentration increases to between 71 and 88% with the addition of two extra sampling days (five days instead of three).

6.5.4 The UFP LUR Model
Table 6-3 summarizes the LUR model obtained from the mobile PN concentration data. The model has an $R^2 = 0.72$ and a leave one out (LOO) cross validation $R^2 = 0.68$. The parameters included in the model were related to population density, dwelling density, roadway length and resource and industrial activities. Two traffic related parameters were included in the model: length of major arterial roadways and highways within 300m and within 300-3000m. In the initial development of stages of the LUR model, a single road length parameter was included (3000m buffer); however, the PN concentrations on major arterial roadways were consistently underestimated. Upon separating the two traffic-related terms, the total model $R^2$ increased by 2% and the predictive power of the model for PN concentration data collected on major roadways improved. Although resource and industrial area within 1000m was a good predictor variable, Toronto does not have many large emitting facilities. The roadways in and near areas with resource and industrial land use are known to have more heavy duty diesel truck traffic than other areas in the city. Thus the resource and industrial land area parameter may be a partial proxy variable for freight traffic.
Table 6-3 Predictor variables, corresponding buffers, coefficients, standardized coefficients, p-values, \( R^2 \) values and root mean square error (RMSE) values for the Toronto UFP LUR model. SE: Standard Error, LOO: leave-one-out.

<table>
<thead>
<tr>
<th>Variables (unit)</th>
<th>Buffer (m)</th>
<th>Coefficients (SE)</th>
<th>Beta (SE)</th>
<th>p-value</th>
<th>Partial ( R^2 )</th>
<th>Model ( R^2 )</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>In population density (ln (people-km(^{-2})))</td>
<td>300</td>
<td>2084 (927)</td>
<td>0.23 (0.05)</td>
<td>0.027</td>
<td>0.56</td>
<td>4380</td>
<td></td>
</tr>
<tr>
<td>Resource &amp; industrial area (m(^2))</td>
<td>1000</td>
<td>8.40 x 10(^{-3}) (1.60 x 10(^{-3}))</td>
<td>0.31 (0.06)</td>
<td>&lt;0.001</td>
<td>0.09</td>
<td>3896</td>
<td></td>
</tr>
<tr>
<td>Length of major roads and expressways (m)</td>
<td>300-3000</td>
<td>0.121 (0.03)</td>
<td>0.40 (0.09)</td>
<td>&lt;0.001</td>
<td>0.04</td>
<td>3662</td>
<td></td>
</tr>
<tr>
<td>Residential area (m(^2))</td>
<td>3000</td>
<td>4.70 x 10(^{-4}) (1.50 x 10(^{-4}))</td>
<td>0.24 (0.08)</td>
<td>0.003</td>
<td>0.01</td>
<td>3573</td>
<td></td>
</tr>
<tr>
<td>Length of major roads and expressways (m)</td>
<td>300</td>
<td>1.51 (0.63)</td>
<td>0.15 (0.06)</td>
<td>0.017</td>
<td>0.01</td>
<td>3495</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-21288 (5366)</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|  | Total Model | 0.72 | 3495 |
|--------------------------|-------------|------|
| LOO Cross Validation | 0.68 | |

6.5.5 Mapping UFP

Figure 6-6 shows PN concentration distribution across Toronto. The map suggests that the PN concentration in Toronto may vary by as much as a factor of 10 during the afternoons. The highest PN concentrations were observed in the downtown core (near the central site, sites A and B), near the highway junction points and around industrialized areas in the city. The PN concentrations were slightly higher on the west side of the city. The lowest PN concentrations were found along the shore of Lake Ontario to the east of the city.
Figure 6-6  The modelled spatial distribution of UFP in Toronto.
6.5.6 Model Evaluation

The LUR model developed for Toronto had a total model $R^2 = 0.72$. The individual models developed for the LOO cross-validation had $R^2$ values ranging from 0.72 to 0.76 and RMSE values ranging from 3,150 to 3,380 cm$^{-3}$ or between 23 and 25% of the average measured PN concentration. This study benefited by collecting measurements at approximately the same time every day to avoid the need for applying temporal correction factors and by visiting each road segment on multiple days to avoid the need to correct for day-to-day variability.

![Graph](image.png)

Figure 6-7 The measured PN concentration at each fixed field site and the predicted PN concentration by UFP LUR model. The upper panel shows the PN concentration collected during the mobile measurement period between 12:00 and 15:00 while the lower panel shows the 24-hour average PN concentration. All values were extracted from the surface generated by kriging.
Figure 6-7 presents the predictive power of the UFP LUR model (upper: measurement period 12:00 – 15:00, lower: 24-hour average). Overall excellent agreement was observed between the measured and predicted PN concentrations. The LUR model had stronger predictive power for measurements collected between 12:00 – 15:00 ($R^2 = 0.86$) than the 24-hour average regression ($R^2 = 0.72$). The LUR model overestimated the 24-hour average PN concentration at the fixed site by a factor of 1.28. This is unsurprising since the mobile measurements were collected during the afternoon when the PN concentration is likely to be highest.

6.5.7 Discussion
Initial efforts to develop a UFP LUR model focused on normalizing the mobile PN concentration data with paired data collected at the central site. This effort was ultimately abandoned because the central site was located 20m from a major arterial roadway and the impacts of individual vehicles disproportionally skewed the results. In future studies, an urban background monitoring site might prove to be appropriate to minimize the impacts of day-to-day variability in PN concentration. This study demonstrates that LUR modelling can be successfully applied to raw mobile UFP concentration data ($R^2 = 0.72$, validated $R^2 = 0.68$) provided that the measurements are collected at the same time of day and under similar meteorological conditions. The Toronto LUR model has a similar $R^2$ value as the Hoek et al., (2011) for Amsterdam ($R^2 = 0.65$) and considerably higher $R^2$ values than the LUR models developed by Abernethy et al., (2012) for Vancouver, Canada ($R^2 = 0.34 – 0.53$) and by Rivera et al., (2012) for multiple urban areas in Spain ($R^2 = 0.13 - 0.63$).

Abernethy et al., (2011) observed correlation between PN concentration and the density of fast food restaurants. This study also observed strong correlation between the number of restaurants and PN concentration; however, restaurant density was highly correlated with the length of major roadways and highways within 300-3000m predictor variable. Consequently, restaurant density was not included in this LUR model due to the likelihood that vehicles were the more dominant source of UFP and the current study likely cannot separate the impacts of restaurants and traffic. However, with greater spatial coverage in future studies, the relationship between UFP and restaurants may be further elucidated.
This study relied on the P-Trak condensation counter with a lower size cut-off of 20nm. Previous investigation of UFP in Toronto and other urban areas suggests that the particle size distribution in areas away from traffic is dominated by larger particles (Ketzel et al., 2004a; Sabaliauskas et al., 2012). When walking along roadways far from major traffic sources, a large majority of all the particles are likely captured by the P-Trak. In areas close to major traffic sources, the UFP size distribution shifts towards smaller particles and the P-Trak likely underestimates the total PN concentration. Even though the size cut-off of the P-Trak (>20nm) was higher than the CPC 3786 (>6nm), only a small fraction of the UFP size distribution was likely less than 20nm at the residential sites (C, D, E, F) making the measurements from the two instruments directly comparable. However, in locations closer to major traffic sources, such as the central site and the downtown sites (A and B), particles <20nm were likely present. At the central site (20m from a major arterial road), the average geometric mean diameter was smallest between the hours of 12:00 and 15:00 (mean = 28nm) and 40% of all particles were <20nm. Sites A and B were within 30 and 100m of a major arterial roadway, respectively. Given that the developed LUR model only captures particles >20nm, the predicted PN concentration at the central site, site A and possibly site B are likely underestimated.

This study presents a first approximation of the range of PN concentrations that are likely present in Toronto. The ability of the UFP LUR model to accurately predict the PN concentration in other seasons remains an open question. At the central site, the PN concentration is between 2 and 3 times higher during the winter than the summer due to changes in vehicle emissions, increased residential heating and the increased atmospheric lifetime of UFP (Sabaliauskas et al., 2012). Consequently, the spatial relationships established between PN concentrations and traffic-related and residential area variables will likely differ in during the winter. Further, meteorological patterns between summers can vary substantially in Toronto. Consequently, the ability to predict summer PN concentrations at different locations in different years may not be possible.

### 6.6 Conclusions

This study successfully developed an LUR model for UFP in Toronto, Canada using a combination of fixed site and mobile monitoring (R² = 0.72, validated R² = 0.68). The predictor variables included population density within 300mn, resource and industrial area within 1000m,
residential area within 3000m and major road and highway length within 300m and within 300 to 3000m. In addition, this study benefited from a hybrid sampling protocol that balanced the competing needs of maximizing sampling time and spatial coverage to develop a robust LUR model. The need to apply temporal correction factors to the mobile PN concentration data were minimized by collecting measurements at the same time and by selecting days with similar meteorological conditions. Assessing the ability of the LUR model to predict PN concentrations in other seasons and in other years remains an open question.

6.7 Acknowledgements
Funding for SOCAAR was provided by the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund. This work was supported by the Ontario Ministry of the Environment's Best in Science Research Program. Student stipends were provided by the Ontario Graduate Scholarships in Science and Technology. Special thanks to Matt Roorda for providing access to the micro-trip simulation traffic model and the study participants that volunteered to host the particle counting equipment in their backyards.
Chapter 7
Concluding Remarks and Future Directions

7.1 Summary
This research was motivated by the large gaps in our scientific knowledge surrounding the long-term behaviour and spatial distribution of UFP in urban areas. This work examined one of the longest existing urban UFP data sets collected over a five year period in downtown Toronto. Prior to this analysis, knowledge of UFP behaviour in cities was dependent on intermittent UFP measurements. This short-term data collection strategy allowed for weekday-weekend and diurnal behaviour to be studied but lacked the long-term perspective required to make conclusions about urban residents’ exposure to UFP. The long-term, continuous and size-resolved data collected in this study enabled for annual, seasonal, weekday-weekend and diurnal behaviour of UFP to be studied. These data are the first reported for a major Canadian city and among the first in North America.

Next, this work also examined a spatially resolved UFP dataset collected at six field sites and while walking along roadways with different traffic volumes. In major North American cities, only limited spatially resolved UFP data exists. This work combined both fixed site and mobile monitoring strategies to balance the competing need for temporal and spatial data. While this work relied on intermittent UFP data collection strategies, the resulting data allowed for the first LUR model to be developed for Toronto and one of the first LUR models developed for UFP worldwide. Although a few other UFP LUR models have been developed, their predictive power has been low. The LUR model developed as part of this work is the only the second model that explains a majority spatial variability of UFP.

This chapter will summarize the key findings of this research in terms of the objectives listed in Chapter 1 followed by a discussion focused on the insights obtained from applying data mining to these UFP data and possible future research directions.
7.1.1 Objective 1: Identify the factors that have influenced the temporal behaviour of UFP

In Chapter 3, the factors that influenced the temporal behaviour of UFP over a five year measurement period in downtown Toronto were explored. The nuclei (PN$_{50}$) and Aiken (PN$_{50-100}$) size fractions were examined separately and their variation over four time scales was investigated: annual, seasonal, diurnal, and weekday-weekend. The core observations from this study are summarized below:

- Between 2006 and 2011, a reduction of 21% for PN$_{50}$ and 17% for PN$_{50-100}$ was observed.
- The concentration of PN$_{50}$ was a factor of 2.2 times higher during the winter than the summer while PN$_{50-100}$ was only marginally higher during the winter months.
- Both size fractions of UFP exhibited distinct diurnal behaviour that differed on weekdays and weekends.
- Both PN$_{50}$ and PN$_{50-100}$ had lower concentrations on weekends than on weekdays.

Between 2006 and 2011, no government policies actively sought to reduce the ambient concentration of UFP. However, several policies were implemented aimed at reducing NO, NO$_2$, SO$_2$ and PM$_{2.5}$ emissions including changes to the composition of diesel and gasoline fuels, the adoption of new engine and emission control technologies, and reduced usage of coal-fired power plants. Each of the abovementioned policies was implemented gradually and occurred at roughly the same time as the economic downturn that affected the United States and Canada. As a result, the impacts of individual policies on UFP concentrations could not be assessed. However, these types of policies were not unique to Toronto and other jurisdictions may have had similar reductions in UFP concentrations.

Meteorological conditions were shown to dramatically impact the concentration of UFP on a seasonal and day-to-day basis. PN$_{50}$ was inversely correlated with temperature while PN$_{50-100}$ showed no dependence. Further, wind direction was also shown to influence both size fractions differently. PN$_{50}$ exhibited a strong wind directionality bias towards the northwest during the winter months and no wind direction dependence during the summer. PN$_{50-100}$ consistently showed a wind direction bias towards the southwest suggesting that long-range transport may be an important source of particles with diameters between 50 and 100nm.
Macro-scale human behaviour (e.g. driving patterns) dramatically impacted the diurnal and weekday-weekend variation of UFP. Each weekday morning, a large portion of the GTA’s working population drives their personal vehicles to downtown Toronto. This mass movement of people corresponded with a consistent increase in PN$_{50}$ and PN$_{50-100}$ concentrations between 6:00 and 10:00am on weekdays. In contrast, on weekends, fewer people commute into the downtown core to begin work at 9:00am and lower overall traffic volumes are typically observed. These differences in driving patterns matched the temporal behaviour of both size fractions of UFP. The concentration of both PN$_{50}$ and PN$_{50-100}$ increased later in the morning and had lower concentrations on weekends than on weekdays.

In summary, this work identified three main factors that influenced temporal patterns in the concentrations of PN$_{50}$ and PN$_{50-100}$: government policy changes surrounding fuel composition, vehicle technologies and electricity generation, meteorology and macro-scale human behaviour (e.g. driving patterns). Further, this study demonstrated the need to report UFP on a size-resolved basis. While a large portion of PN$_{50}$ and PN$_{50-100}$ are likely from vehicle emissions, both size fractions of UFP exhibited distinct seasonal behaviour, varying correlation with other pollutants and different wind direction dependence. This divergent behaviour suggested that a portion of PN$_{50}$ and PN$_{50-100}$ likely originates from different sources. However, from this analysis alone, quantifying the impacts of different UFP sources was not possible.

7.1.2 Objective 2: Explore the use of alternative metrics to describe the temporal and spatial variation of UFP

In Chapter 3, the PN concentration and GMD were used to describe the temporal variation of UFP in downtown Toronto. The above mentioned traditional air quality metrics do not fully capitalize on the wealth of size-resolved information that the FMPS instrument can provide. Firstly, by reporting only the concentration of PN$_{50}$ or PN$_{50-100}$, the shape of the PSD was unknown (e.g. unimodal or bimodal). Secondly, while the factors that influenced the PN concentration throughout the measurement period were identified, the contribution of the known sources of UFP (i.e. traffic emissions, long-range transport or photochemistry) remained
unquantifiable. For example, it would be incorrect to identify all periods with elevated PN$_{50}$ concentrations as traffic influenced or all periods with elevated PN$_{50-100}$ concentration as long-range transport related. In Chapter 4, cluster analysis was applied to the same five year UFP data set to determine if the shape of the PSD could be grouped into physically meaningful clusters and if greater insight into the behaviour of UFP could be elucidated. The core observations from this study are summarized below:

- The PSD data could be grouped into eight physically meaningful clusters that possessed distinct seasonal, diurnal and weekday-weekend behaviour.
- The eight PSD types exhibited variable correlation with other pollutants and wind direction dependence

This work demonstrated the value of applying cluster analysis to PSD data. Cluster analysis did provide greater insight into the behaviour of UFP especially when the PSD types’ temporal behaviour was explored on a seasonal, diurnal and weekday-weekend basis, relationships with other pollutants were established and dependencies on meteorological parameters were defined. However, matching PSD types to individual sources remained challenging. For example, C8 was identified as being related to long-range transport due to its correlation with PM$_{2.5}$ and strong southerly wind direction dependence. However, this PSD type was relatively rare (5% occurrence) and a small contributor to the total PN concentration (3%). The PSD types were more likely to be a combination of the regional background with a superimposed traffic signal of varying intensity (C3, C6, C7) or related to vehicle emissions and photochemical nucleation (C1, C2, and C4). Although separating these PSD types into distinct sources was not possible, cluster analysis allowed for behaviour of UFP to be better described.

7.1.3 Objective 3: Quantify the contributions of local scale and regional scale sources of UFP

Consistently throughout this work, frequent PN concentration spikes were observed as vehicles passed roadside measurement locations. Further away from major traffic sources, the PN concentration spikes although present, were less frequent and intense. In Chapter 5, a wavelet based background subtraction algorithm was applied to the total PN concentration time series
collected at the SOCAAR facility and six field sites located across Toronto. This method successfully separated the PN concentration spikes from the background concentration thereby enabling the impacts of local scale and regional scale sources of UFP to be quantified. The key observations from this study are summarized below:

- The local to neighbourhood scale sources of UFP operate over spatial scales of 2.5km
- The PN concentration attributable to the local to neighbourhood scale sources contributed between 13 and 32% to the total observed PN concentration at the six residential sites during the summer.

This study suggested that at least two spatial scales need to be considered when assessing exposure across a population. The portion of the PN concentration attributable to sources and processes that operated at urban to regional spatial scales exhibited moderate temporal correlation with the central site and was not related to the total number of vehicles surrounding the monitoring site. The PN concentration attributable to the local to neighbourhood scale sources exhibited no correlation with the central site but was correlated with the number of vehicles within a 2.5km buffer of the monitoring site. Further, the contribution of local scale sources could be higher than 32% at high traffic locations and higher still on-road when in traffic. These results suggested that Toronto residents are exposed to a large range of PN concentrations by virtue of their proximity to major traffic sources.

7.1.4 Objective 4: Determine if land-use regression model can be applied to UFP data to quantify the range of possible concentrations that exist across Toronto

In Chapter 6, land-use regression modeling was successfully applied UFP data collected while walking in neighbourhoods with different geographical characteristics. This LUR model was intended to serve as a first approximation of the potential range of UFP concentrations that exist across Toronto. The key outcomes from this study are summarized below:
The predictor variables that best described the spatial variation of UFP included population density within 300m, total resource and industrial area within 1000m, total residential area within 3000m, and major roadway and highway length within 3000m.

The LUR model successfully predicted the afternoon PN concentration (slope = 0.96, $R^2 = 0.86$) and 24-hour average PN concentration (slope = 1.28, $R^2 = 0.72$) measured at seven fixed monitoring sites.

The LUR model developed in this study suggested that the range of possible PN concentrations that Toronto residents encounter may vary by a factor of 10 during the summer. The highest PN concentrations in Toronto arise around highway junction points, industrial areas and the downtown core. While the LUR model was able to predict the summertime PN concentration at seven fixed monitoring sites, the predictive power of the LUR in other seasons remains uncertain. The seasonal trends from Chapter 3 suggested that cooler temperatures impact both the formation and atmospheric lifetime of UFP. As a consequence, the relationships established between traffic-related variables may not remain constant between seasons.

7.2 Insights from Data Mining

Long-term and spatially resolved UFP data is needed for reliable population exposure assessment. The data mining techniques that were further developed and validated as part of this thesis represent a new and complimentary way of thinking about UFP exposure. This research demonstrated that reporting traditional metrics would limit the information that can be uncovered about the factors influencing the long and short term behaviour of UFP and shape of the PSD.

In Chapter 3, MLR models were developed between UFP, commonly monitored air pollutants and meteorological parameters. These models are the first and only known attempt to estimate the historical UFP concentrations in an urban area. The models suggest that between 2003 and 2006, the concentration of $PN_{50}$ and $PN_{50-100}$ decreased by 10% and 14%, respectively. Many urban areas have short-term UFP data sets and collect routine measurements for NO, NO$_2$, SO$_2$, CO, O$_3$ and PM$_{2.5}$. The relationships that were established between the size-resolved UFP concentration and the commonly monitored pollutants and meteorological patterns may allow other jurisdictions to model the long-term variation of UFP.
In Chapter 4, the temporal variation of the PSD shape was explored using the k-means clustering algorithm. One of the most remarkable findings from this analysis was that eight clusters can reliably describe the temporal variation of the PSD shape over a five year time horizon. In addition, despite the observed reduction in PN_{50} and PN_{50-100} concentrations between 2006 and 2011, the frequency of occurrence of each cluster did not change between years. Finally, while only examined over a single season, the application of the clustering and wavelet decomposition algorithms (Chapter 5) yielded complimentary results. At the central site during the summer of 2008, PSDs that were identified as “traffic-related” by the clustering algorithm appeared at times with higher percent contribution from the local-neighbourhood signal (22-26%) than when “long-range transport” related particle size distributions were present (13-15%).

From an exposure perspective, the MLR models, cluster analysis and wavelet based background subtraction algorithm provide complimentary information at the central site. The MLR models provided evidence suggesting that in addition to the observed reduction in UFP concentrations between 2006 and 2011, the downward trend likely existed since at least 2003. The cluster analysis demonstrated that despite the gradual reduction in PN concentrations, the shapes of the PSD remained constant. The wavelet based background subtraction algorithm showed that periods with more PN concentration spikes were more likely to be impacted by clusters related to traffic.

In addition to enhancing the temporal analysis of UFP, the data mining techniques presented in this thesis also further characterize the spatial variation of UFP. The wavelet decomposition algorithm was developed for use at a roadside monitoring station with a FMPS operating at 1-second time resolution by Klems et al., (2010). In Chapter 5, this research modified and implemented the algorithm for use on lower time resolution data and at urban background locations. To date, this is the first attempt to decompose the UFP time series at multiple sites and describe the spatial variability of UFP in terms of two spatial scales: the local-to-neighbourhood and urban-to-regional. This analysis provided two important insights: 1) the urban-to-regional scale PN concentration is correlated across Toronto and 2) the local-to-neighbourhood scale PN concentration is likely related to traffic.
7.3 Other Perspectives

One of the core hypotheses many studies is that UFP exhibit large concentration gradients around roadways. Despite walking on major roadways and local roadways around Toronto throughout the summer of 2005, 2006 and 2008, large UFP concentration gradients were not observed while walking from an “upwind” to a “downwind” location. Within a typical city block, local weather conditions are nearly impossible to characterize. Vehicle induced turbulence, local topography, buildings and changes in wind direction and speed can affect the dispersion patterns of pollutants around roadways. Further, the limited availability of reliable traffic data for Toronto is an ongoing challenge. In the event that the dispersion patterns around the roadway could be successfully modelled, traffic volumes remain uncertain.

Further, distance to the nearest major roadway or highway is a frequently used metric to estimate traffic exposure in epidemiological studies. While developing the UFP LUR model, distance to the nearest major roadway and highway were poor predictors of PN concentrations. Roadways are classified based on a number of parameters including the number of lanes, traffic volume and legal speed limit. In Toronto, major arterial roadways have at least 2 lanes, traffic volumes > 20,000 vehicles per day, and legal speed limits of 50-80km/h. Therefore, across the city, roadways categorized as major arterial roadways can have a wide range of characteristics and may not be directly comparable to each other without incorporating other descriptors such as traffic or lane counts.

In Chapter 5, two broadly defined spatial scales were used to describe the variation of UFP: local-to-neighbourhood and urban-to-regional. From years of observation at the SOCAAR facility, PN concentration spikes occur as vehicles pass the roadside monitoring site (~20m from the road). From the geographical modeling in Chapter 5, the PN concentrations spikes had the strongest relationship with traffic counts within a 2.5km buffer area. This relationship between local and neighbourhood scale traffic is not necessarily contradictory. Existing traffic data for individual roadways consists of either a manual count conducted for a single day every few years or as a series of model simulations. Both approaches result in a single value for vehicles along each road segment. While these data are useful for identifying areas with high and low vehicle counts, a single value is unlikely to reflect the full range of traffic patterns that can exist on each roadway. Local traffic patterns are often related to social and geographical parameters such as
population density, restaurant density and employment opportunity. Therefore, by increasing the buffer area to include more roadways, the odds of describing traffic behaviour at a location of interest is likely improved.

Further, the urban to regional scale PN concentration exhibited a correlation with the central site. However, the geographical range of the urban to regional scale PN concentration remained unknown. The relevant geographical range may be elucidated by examining the measurements collected during a 2006 field study. UFP measurements were collected at 23 parks located within 7 and 53km from the central site. Within a city, large parks represent the best opportunity to maximize distance from traffic sources; therefore, a large majority of the observed PN concentration likely originates from sources and processes operating at the urban to regional spatial scale.

![Figure 7-1 Ratio of the PN concentration measured in GTA parks and downtown Toronto](image)

Figure 7-1 shows the ratio between the PN concentration measured in the parks and those measured at the central site. The parks located the furthest from the downtown core (>30km) were more likely to be surrounded by agricultural land while those within the GTA (<30km) tended to be surrounded by roadways and residential areas. A weak downward trend can be observed as a function of distance from the central site. This suggests that the urban scale PN concentration may exhibit variability over distances spanning 30-50 km and the regional scale operates at spatial scales >50km from Toronto.
7.4 Future Research Directions

1. This work demonstrated that measurements collected at a centralized location are unlikely to reflect the range of concentrations that may exist across an urbanized area. Although short duration monitoring was conducted in five neighbourhoods, more long-term, spatially resolved and size resolved UFP data is needed to solidify our understanding of the potential range of concentrations and particle sizes that exist in Toronto and between seasons.

2. This work focused on residential areas and did not examine the range of possible UFP concentrations that may exist in and around industrial areas. In Toronto, areas zoned for industrial and resource activities consist of a unique combination of facilities that may have variable UFP emissions. The LUR model developed in this work treated all industrialized areas equally and therefore may have over or underestimated their impacts on UFP concentrations across Toronto. For future LUR model development, identifying large emitters within individual industrial areas may reduce the existing model uncertainty.

3. In 2013, the Spatial Characterization of Ultrafine Particles in Toronto study measured UFP, black carbon and NOx at 22 locations for 2 weeks in the summer and 2 weeks in the winter. During this study, measurements were repeated at sites C, D and F from Chapters 5 and 6. Although the data analysis is ongoing, the addition of other traffic-related pollutants (NOx and black carbon) may enhance our understanding of the longer term trends of UFP at these three monitoring sites.

4. Since its publication in 2013, the clustering methods presented in Chapter 5 have been applied by other research groups in different locations around the world. Recently, Bismarck-Osten and Weber (2014) applied cluster to PSD data collected at nine monitoring sites in three European cities and Beddows et al., (2014) applied cluster analysis to PSD data collected in 24 European cities. The application of other data mining strategies to these large data sets will continue to be an area of active research as more locations begin to collect size resolved UFP data sets. So far, each study has relied on
different UFP instruments, time resolutions and data treatment protocols. This rapidly
developing area of research would benefit by adopting a series of standardized analysis
protocols to enable inter-study comparisons.

5. Finally, the vehicle fleet continues to evolve to improve fuel economy and reduce
emissions of regulated pollutants. The reductions in UFP concentrations that occurred in
recent years could be undone by the adoption of new engine technologies or emission
control technologies. For example, gasoline direct injection engines are becoming more
common on new cars to meet tightening fuel economy standards. A few studies suggest
that these vehicles emit substantial quantities of UFP (Bonatesta et al., 2014; Myung et
al., 2014). UFP remains unregulated and is not currently included in existing air quality
monitoring networks. Continuous reductions in ambient UFP concentrations are not
guaranteed. To examine the impacts of new technologies and policies on the vehicle fleet,
surveillance of UFP is required.
References


Brock, G., Pihur, V., Datta, S., Datta, S., 2011. clValid: Validation of Clustering Results. R package version 0.6-4.


City of Toronto Open Data Project. Dine Safe Data available at: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=b54a5f9cd70bb210VgnVCM10000071d60f89RCRD&vgnextchannel=1a66e03bb8d1e310VgnVCM10000071d60f89RCRD. Last Accessed: April 2014.


Perspectives, 114, 1172-1178.


Jeong, C.-H., Evans, G.J. 2009. Inter-Comparison of a Fast Mobility Particle Sizer and a Scanning Mobility Particle Sizer Incorporating an Ultrafine Water-Based Condensation Particle Counter. Aerosol Science and Technology, 43, 364-373.


Ketzel, M., Berkowicz, R., 2004b. Modelling the fate of ultrafine particles from exhaust pipe to rural background: an analysis of time scales for dilution, coagulation and deposition. Atmospheric Environment, 38, 2639-2652.


Roorda, M., M. Hain, G. Amirjamshidi, R. Cavalcante, B. Abdulhai, C. Woudsma (2010). Exclusive Truck Facilities in the Toronto Area: Analysis of Truck and Automobile Demand. Transportation Research Record: Journal of the Transportation Research Board 2168, pp. 114-128


Appendix A

Ultrafine Particle Instruments

A.1 Measurement Principles

PM monitoring instruments reliably measure the particle mass concentration by capitalizing on the tendency of airborne particles to scatter light in a predictable manner or by directly weighing the particles over a certain time interval. Due to instrument advances, many monitors can now operate at high time resolution. The particles in the ultrafine range (≤100nm) possess very little mass and are small in size. This presents several challenges for instrument designers. Firstly, UFP is smaller than the wavelength of ultraviolet light used in most low-powered lasers. As a result, particles pass through the laser beam without scattering light. Secondly, if UFP were successfully captured onto a filter or microbalance, the time interval required for enough UFP to accumulate would be too long and a significant portion of the particles may evaporate. As a result, other methods have been developed to capitalize on the properties of the smallest particles to allow for reliable detection of both the total number of particles and particle size distribution.

Particles in the ultrafine range can be grown through condensation by passing the air sample through a chamber saturated with water or alcohol vapour. Once the particles are greater than 800nm in size, optical methods can be used to count the number of particles within a given volume of air. An alternative method of detection involves applying an electrical charge to the particles and separating them based on mobility diameter. This allows for the total number of particles and the size distribution of particles within a given volume of air to be simultaneously detected at high time resolution. One final detection method can be used to resolve the size distribution of UFP. By producing a mono-dispersed aerosol sample, the total number of particles within a given size range can be counted.
A.2 UFP Instrumentation

A.2.1 Fast Mobility Particle Sizer

The Fast Mobility Particle Sizer (FMPS) Spectrometer (Model 3091, TSI Incorporated, Shoreview, MN, USA) measures particles ranging from 5.6 to 560nm across 32 size bins. This instrument is capable of resolving the size distribution and total number of particles at 1-second time resolution. Air is drawn into the instrument continuously through the sample inlet. Particles in the air sample are charged using a corona charger. The charged particles are then separated according to mobility diameter by passing through a column of electrometers (Figure A-1).

![Figure A-7-2 Operation of the Fast Mobility Particle Sizer (TSI Incorporated)](image)

A.2.2 Scanning Mobility Particle Sizer

The Scanning Mobility Particle Sizer (SMPS) Spectrometer (Model 3936, TSI Incorporated, Shoreview, MN, USA) measures particles ranging from 2.9 to 1000 nm depending on the instrument configuration at 1 to 5 minute time resolution. The SMPS relies on producing a
mono-dispersed aerosol sample using a Differential Mobility Analyzer (DMA). Each aerosol sample is divided into “packets” and is sent to a condensation particle counter that counts particles within a given size range individually (Figure A-2). Over the sampling period, the size spectrum is scanned from the smallest size bin and is verified by scanning from the largest size bin to the smallest. The resolved size distribution range depends on the instrument configuration. When the Nano-DMA is used, particles between 2.9 and 109nm can be resolved. When the long-DMA is used, particles between 10 and 1000nm can be reliably measured.

![Diagram of a scanning mobility particle sizer spectrometer](image)

**Figure A-7-3** Operation principles of the scanning mobility particle sizer spectrometer (TSI incorporated)

### A.2.3 Water Condensation Particle Counter

The Water Condensation Particle Counter (WCPC) is a portable particle condensation counter that measures particles between 6 and 1000nm (Model 3781, TSI, Shoreview, MN, USA). The instrument detects the total number of particles per cubic centimeter of air at up to 1-second time resolution to a maximum concentration of 500,000pt/cm³. Particles are grown to 1000nm in a chamber saturated with water vapour. This instrument was developed and tested between 2004
and 2006 (Hering, 2006) and became commercially available in 2007. To date, this instrument has not been extensively used in field deployments. Studies that have used this CPC near traffic sources have not reported significant issues with its operation (Barone and Zhu, 2008; Buonocore et al., 2009).

**A.2.4 P-Trak Condensation Particle Counter**

The P-Trak Condensation Particle Counter (Model 8525, TSI Incorporated, Shoreview, MN, USA) is a handheld monitor that measures the total number of particles between 20nm and 1000nm per cubic centimeter of air. The P-Trak has a concentration range between 0 and 500,000 particles per cubic centimeter (pt/cm³) and can log data at 1-second time resolution. The P-Trak can operate for up to 8 hours in the field. This instrument condenses isopropyl alcohol onto the particles to allow them to be detected optically. Air enters the inlet and is passed through a wick containing a mixture of liquid and gaseous isopropyl alcohol. The air is then passed through a photodetector and individual particles are counted (Figure A-3).
A.1.5.1 Validation of P-Trak

The portable nature of the P-Trak allows for microenvironments to be studied relatively easily; however, the instrument both underestimates and overestimates the total number of particles under certain meteorological conditions and UFP emission situations. At high relative humidity (greater than 70%), the isopropyl alcohol used to grow particles can become saturated with water. This causes the efficiency at which the instrument can grow particles to decrease the total

Figure A-7-4 The operating principles of the P-Trak. Adapted from material provided by TSI Incorporated.
number of particles to be underestimated and reduces the total number of hours that the instrument can be operated in the field reliably Hammond et al. (2007). Secondly, at concentrations in excess of 100,000pt/cm$^3$, particle coincidence errors will cause the actual number of particles to be underestimated (Kaur et al. 2006).

There is no laboratory-based standard reference material available to calibrate the P-Traks. The manufacturer (TSI Incorporated) claims that the P-Trak measures the total number of particles between 20 and 1000nm up to a maximum concentration of 500,000 particles per cubic centimeter. The claimed size range and accuracy of the instrument over the 0-500,000 pt/cm$^3$ concentration range were examined using four approaches as listed in Table A-1
**Table A-1 Validation of the P-Trak condensation particle counters**

<table>
<thead>
<tr>
<th>Test</th>
<th>Approach</th>
<th>Concentration Range</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement of instruments</td>
<td>Two P-Traks sampled ambient air from the sampling duct located in the SOCAAR lab (20m from traffic)</td>
<td>8000-70,000 pt/cm³</td>
<td>Instruments exhibited good agreement and were well correlated (slope=1.02-1.07, R²&gt;0.95)</td>
</tr>
<tr>
<td>Size cut-off of instruments</td>
<td>Two P-Traks sampled ambient air from the sampling duct and compared to the Fast Mobility particle Sizer</td>
<td>8000-70,000 pt/cm³</td>
<td>The FMPS and P-Traks showed good agreement. The best correlation was observed for the total number of particles greater than 25nm</td>
</tr>
<tr>
<td>Agreement of instruments over larger concentration range</td>
<td>Side-by-side field measurements near traffic</td>
<td>30,000 – 300,000 pt/cm³</td>
<td>Good agreement was observed at low concentrations. At high concentrations, the two instruments disagreed by a factor of 2.0</td>
</tr>
<tr>
<td>Agreement of instruments when sampling different sources</td>
<td>Two P-Traks compared to the Fast Mobility Particle Sizer using particles derived from cooking sources</td>
<td>3000-500,000 pt/cm³</td>
<td>Good agreement between the two P-Traks and FMPS at low concentrations Poor agreement above 300,000 pt/cm³</td>
</tr>
</tbody>
</table>

**A.3 References**


Appendix B
The Application of Clustering Analysis and Solution Validation Procedures

B.1 Clustering Analysis
A cluster can be described as a group of objects that are similar in some way. There exist hundreds of clustering algorithms that capitalize on different data set properties to organize the objects into groups. Clustering algorithms assess “similarity” or “dissimilarity” using different metrics that can include distance, cosine similarity or relative entropy. For example some algorithms rely identifying the centroids of the clusters while others build sequential trees to merge similar data into groups. In this thesis, the k-means clustering algorithm is applied to particle size distribution data. In the following sections, the theoretical basis for assessing clustering tendency, determining the number of clusters in a data set, and measuring clustering quality will be discussed.

B.1.1 Assessing Clustering Tendency
Cluster analysis can only be meaningful when non-random structure exists in the data. Therefore, prior to performing cluster analysis, the tendency of the data to form natural clusters must be assessed. Figure B-1 illustrates three possible ways data can be structured in two dimensional space: randomly, uniformly, or clustered. In two dimensional space, clustering tendency can be assessed visually; however, in more complex data sets, clustering tendency is not always readily apparent.
The Hopkins’ Statistic is an easily calculated metric that can determine how the data hypothesized to contain clusters compares to randomly generated data. The Hopkins’ statistic seeks to accept or reject the null hypothesis: the homogeneous hypothesis – that the data is uniformly distributed and contains no meaningful clusters. To calculate the Hopkins’ Statistic, random data (grey circles, Figure B-2) is generated within the range of the measured data (black circles, Figure B-2). Next, the distance Euclidean distance is calculated between each observation and between each randomly generated datum.

The nearest neighbour distance was assigned to each observation (W). The Euclidean distance was calculated between each random point and its nearest neighbour (U). Finally, the minimum
distances assigned to the PSD data (W) and random data (U) were summed. The Hopkins’ Statistic was calculated according to Equation B-1:

Equation B-1

\[ H_p = \frac{\sum W}{\sum U + \sum W} \]

Values of \( H_p \) close to zero suggest that the measured data forms natural clusters. Values close to 0.5 suggest the experimental data is similarly distributed as the randomly generated data.

B.1.2 K-Means Clustering Algorithm

The k-means clustering algorithm is a vector quantization method that separates \( n \) observations into \( k \) clusters. To apply the algorithm, the number of clusters must be defined a priori. Figure B-3 (left) shows the initialization of the k-means clustering algorithm for \( k = 3 \) clusters. First, \( k = 3 \) centroids are randomly generated within the range of the data. Next, the Euclidean distance is calculated between each centroid and the data. Membership of each cluster is determined based on the distance to the nearest centroid. The centroids are then recalculated based on the data contained within each cluster. This process is repeated until the position of the centroids no longer changes (Figure B-3, right) and the observations are no longer assigned to new groups.

Figure B-3 Clustered data in two dimensional space. On the left, the initial guess for a three cluster solution. On the right, the final iteration of the k-means clustering algorithm in relation to the centroids.
B.1.3 Cluster Diagnostics: The Dunn Index

One of the major disadvantages of the k-means clustering algorithm is that the number of clusters needs to be defined a priori. By selecting an arbitrary number of clusters, the optimal solution is unlikely to be found. Although several solution validation metrics have been developed that describe the clusters in terms of their shape, similarity of members and compactness, this thesis applied the Dunn Index.

The Dunn index is the ratio of the minimum inter-cluster distance and the maximum intra-cluster distance. The Dunn index is maximized when a clustering solution contains the most compact and separated clusters. To evaluate the Dunn Index, multiple iterations of the k-means clustering algorithm need to be performed with different numbers of clusters (e.g. $k = 2 – 20$ clusters). The solution that results in the highest Dunn Index value likely contains the most separated and most compact clusters.
Appendix C
Supplementary Material for Cluster Analysis of Roadside Ultrafine Particle Size Distributions

Table C-1: Seasonal Spearman’s Rank Order correlation coefficients for 24-hour PSD type counts and 24-hour PM$_{2.5}$ filter data obtained from the NAPS network and meteorological data. All presented correlation coefficients are significant to 95% confidence. Correlation coefficients are bolded for C3, C6 and C8 to emphasize the changes in the relationships between the PSD types and the PM$_{2.5}$ species between seasons.

<table>
<thead>
<tr>
<th>Season</th>
<th>Compound</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>SO$_4^{2-}$</td>
<td>-0.67</td>
<td>-0.73</td>
<td><strong>0.32</strong></td>
<td>-</td>
<td>0.41</td>
<td><strong>0.77</strong></td>
<td>0.59</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>NO$_3^-$</td>
<td>-0.77</td>
<td>-0.66</td>
<td><strong>0.37</strong></td>
<td>-0.23</td>
<td>0.47</td>
<td><strong>0.74</strong></td>
<td>0.61</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>NH$_4^+$</td>
<td>-0.74</td>
<td>-0.74</td>
<td><strong>0.36</strong></td>
<td>-</td>
<td>0.48</td>
<td><strong>0.79</strong></td>
<td>0.64</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>OC1</td>
<td>-0.30</td>
<td>-0.33</td>
<td><strong>0.23</strong></td>
<td>-</td>
<td>-</td>
<td><strong>0.43</strong></td>
<td>-</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Total Carbon</td>
<td>-0.71</td>
<td>-0.65</td>
<td><strong>0.35</strong></td>
<td>-</td>
<td>0.47</td>
<td><strong>0.78</strong></td>
<td>0.55</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>OC3</td>
<td>-0.61</td>
<td>-0.59</td>
<td><strong>0.27</strong></td>
<td>-</td>
<td>0.44</td>
<td><strong>0.71</strong></td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>OC2</td>
<td>-0.56</td>
<td>-0.48</td>
<td><strong>0.34</strong></td>
<td>-</td>
<td>0.29</td>
<td><strong>0.59</strong></td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>OC4</td>
<td>-0.71</td>
<td>-0.60</td>
<td><strong>0.28</strong></td>
<td>-</td>
<td>0.49</td>
<td><strong>0.73</strong></td>
<td>0.52</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>EC1</td>
<td>-0.77</td>
<td>-0.67</td>
<td><strong>0.32</strong></td>
<td>-</td>
<td>0.50</td>
<td><strong>0.78</strong></td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>-0.67</td>
<td>-0.60</td>
<td><strong>0.37</strong></td>
<td>-</td>
<td>0.45</td>
<td><strong>0.73</strong></td>
<td>0.51</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Total PAH</td>
<td>-0.76</td>
<td>-0.64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.63</strong></td>
<td>0.51</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Fluorene</td>
<td>-0.75</td>
<td>-0.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Phenathrene</td>
<td>-0.69</td>
<td>-0.68</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.82</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Fluoranthene</td>
<td>-0.76</td>
<td>-0.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>SO$_2$</td>
<td>-0.52</td>
<td>-0.56</td>
<td>-0.23</td>
<td>-</td>
<td>0.46</td>
<td><strong>0.47</strong></td>
<td>0.51</td>
<td>0.25</td>
</tr>
<tr>
<td>Summer</td>
<td>SO$_4^{2-}$</td>
<td>-0.54</td>
<td>-0.63</td>
<td>-<strong>0.44</strong></td>
<td>-0.46</td>
<td>-0.29</td>
<td>0.51</td>
<td>0.32</td>
<td><strong>0.76</strong></td>
</tr>
<tr>
<td></td>
<td>NO$_3^-$</td>
<td>-0.33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NH$_4^+$</td>
<td>-0.55</td>
<td>-0.64</td>
<td>-<strong>0.45</strong></td>
<td>-0.49</td>
<td>-0.28</td>
<td>0.51</td>
<td>0.35</td>
<td><strong>0.75</strong></td>
</tr>
<tr>
<td></td>
<td>Total Carbon</td>
<td>-0.48</td>
<td>-0.65</td>
<td>-<strong>0.46</strong></td>
<td>-0.48</td>
<td>-</td>
<td>0.61</td>
<td>0.27</td>
<td><strong>0.70</strong></td>
</tr>
<tr>
<td></td>
<td>OC1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.22</strong></td>
</tr>
<tr>
<td></td>
<td>OC3</td>
<td>-0.54</td>
<td>-0.68</td>
<td>-<strong>0.44</strong></td>
<td>-0.51</td>
<td>-</td>
<td>0.56</td>
<td>0.28</td>
<td><strong>0.65</strong></td>
</tr>
<tr>
<td></td>
<td>OC2</td>
<td>-</td>
<td>-0.29</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>OC4</td>
<td>-0.39</td>
<td>-0.59</td>
<td>-<strong>0.40</strong></td>
<td>-0.42</td>
<td>-</td>
<td>0.58</td>
<td>0.29</td>
<td><strong>0.59</strong></td>
</tr>
<tr>
<td></td>
<td>EC1</td>
<td>-0.51</td>
<td>-0.67</td>
<td>-<strong>0.54</strong></td>
<td>-0.51</td>
<td>-</td>
<td>0.57</td>
<td>0.32</td>
<td><strong>0.76</strong></td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>-0.49</td>
<td>-0.58</td>
<td>-<strong>0.38</strong></td>
<td>-0.44</td>
<td>-</td>
<td>0.51</td>
<td>0.28</td>
<td><strong>0.58</strong></td>
</tr>
<tr>
<td></td>
<td>Total PAH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.48</td>
<td>-</td>
<td><strong>0.70</strong></td>
</tr>
<tr>
<td></td>
<td>Fluorene</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Phenathrene</td>
<td>-</td>
<td>-0.51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Fluoranthene</td>
<td>-</td>
<td>-0.52</td>
<td>-</td>
<td>-0.54</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>SO$_2$</td>
<td>-0.34</td>
<td>-0.38</td>
<td>-0.36</td>
<td>-0.28</td>
<td>-</td>
<td>0.23</td>
<td>0.36</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Figure C-1: A) Map of the sampling area including the National Air Pollution Surveillance Network (NAPS) Site located at the Gage Institute, the SOCAAR UFP monitoring site and the Ontario Ministry of the Environment (MOE) Site. B) A map of Toronto relative to the major coal-powered electricity generating sites located to the south and southwest of Toronto.
Figure C-2 Variation of the Dunn Index across the n=2-20 cluster space. The maximum value of the Dunn index occurs at Number of Clusters = 8.
Figure C-3 Wind direction and wind speed in Toronto. Notably, winds from the east and north east are relatively rare. The dominant directions are north westerly and southerly.
Figure C-4: The 5-year average, the 5-year wintertime average and 5-year summertime average PSD observed in Toronto. Lower PN concentrations were observed during the summer months.
Figure C-5: Fourier analysis of wind direction on a daily, weekly, and monthly basis. Notably, the spectral density at period = 24 hours for northerly and southerly directions contained a strong signal suggesting the lake breeze influences from Lake Ontario. Peaks at period = 80 to 100 hours for all directions suggests impacts from synoptic scale meteorological patterns.
Figure C-6 The wintertime (upper panel) and summertime (lower panel) contribution of each cluster by the time of the day.
Appendix D

Supplementary Material for the Application of Wavelet Decomposition to Quantify the Local and Regional Sources of Ultrafine Particles in Cities

Supporting Information: Wavelets

Table D1: Summary of all 39 wavelets used in the baseline separation algorithm.

<table>
<thead>
<tr>
<th>Family</th>
<th>Wavelets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daubechies (db)</td>
<td>db1, db2, db3, db4, db5, db6, db7, db8, db9, db10</td>
</tr>
<tr>
<td>Symlet (sym)</td>
<td>sym1, sym2, sym3, sym4, sym5, sym6, sym7, sym8, sym9, sym10</td>
</tr>
<tr>
<td>Coiflet (coif)</td>
<td>coif1, coif2, coif3, coif4, coif5</td>
</tr>
<tr>
<td>Biorthogonal (bior)</td>
<td>bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8</td>
</tr>
</tbody>
</table>
Table D2: The original PN concentration (mean ± 95% confidence) and the calculated PN concentrations attributable to urban-regional (U-R) scale and local-neighbourhood (L-N) scale sources. WD: weekdays, WE: weekends. C1 and D1 = Sites C and D during the 2008 measurement period. C2 and D2 = Sites C and D during the 2009 measurement period with a SMPS.

<table>
<thead>
<tr>
<th>Site</th>
<th>Field Site</th>
<th>Original Data</th>
<th>U-R (cm$^3$)</th>
<th>L-N (cm$^3$)</th>
<th>Central Site</th>
<th>Original Data</th>
<th>U-R (cm$^3$)</th>
<th>L-N (cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WD</td>
<td>15390 ± 120</td>
<td>11810 ± 80</td>
<td>3580 ± 70</td>
<td>20800 ± 120</td>
<td>17740 ± 80</td>
<td>3060 ± 70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>15500 ± 170</td>
<td>11800 ± 110</td>
<td>3700 ± 90</td>
<td>20350 ± 120</td>
<td>18070 ± 100</td>
<td>2280 ± 50</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>WD</td>
<td>15610 ± 220</td>
<td>12550 ± 160</td>
<td>3050 ± 100</td>
<td>20060 ± 220</td>
<td>16440 ± 140</td>
<td>3620 ± 160</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>21040 ± 680</td>
<td>15070 ± 290</td>
<td>5970 ± 510</td>
<td>24000 ± 200</td>
<td>21030 ± 140</td>
<td>2960 ± 120</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>WD</td>
<td>8820 ± 140</td>
<td>7060 ± 80</td>
<td>1770 ± 100</td>
<td>13150 ± 250</td>
<td>9630 ± 150</td>
<td>3520 ± 160</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>6450 ± 150</td>
<td>5030 ± 70</td>
<td>1420 ± 110</td>
<td>8940 ± 170</td>
<td>6640 ± 100</td>
<td>2300 ± 100</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>WD</td>
<td>8410 ± 320</td>
<td>6550 ± 100</td>
<td>1870 ± 290</td>
<td>14320 ± 340</td>
<td>10080 ± 180</td>
<td>4230 ± 240</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>8890 ± 310</td>
<td>7210 ± 190</td>
<td>1680 ± 230</td>
<td>14030 ± 380</td>
<td>10560 ± 300</td>
<td>3470 ± 170</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>WD</td>
<td>9450 ± 190</td>
<td>6600 ± 110</td>
<td>2850 ± 130</td>
<td>14310 ± 220</td>
<td>10740 ± 150</td>
<td>3570 ± 140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>5750 ± 190</td>
<td>3180 ± 90</td>
<td>2570 ± 150</td>
<td>9380 ± 130</td>
<td>7260 ± 90</td>
<td>2130 ± 90</td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>WD</td>
<td>11460 ± 260</td>
<td>9280 ± 150</td>
<td>2190 ± 190</td>
<td>16060 ± 400</td>
<td>11890 ± 170</td>
<td>4340 ± 350</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>10550 ± 400</td>
<td>8040 ± 220</td>
<td>2500 ± 270</td>
<td>13470 ± 450</td>
<td>10250 ± 280</td>
<td>3320 ± 300</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>WD</td>
<td>3950 ± 50</td>
<td>3270 ± 40</td>
<td>680 ± 30</td>
<td>15170 ± 300</td>
<td>10780 ± 100</td>
<td>4390 ± 270</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>4180 ± 90</td>
<td>3320 ± 60</td>
<td>860 ± 40</td>
<td>12820 ± 290</td>
<td>9470 ± 200</td>
<td>3340 ± 160</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>WD</td>
<td>13790 ± 230</td>
<td>10980 ± 100</td>
<td>2810 ± 200</td>
<td>16810 ± 220</td>
<td>11770 ± 90</td>
<td>5040 ± 180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>14560 ± 260</td>
<td>12150 ± 160</td>
<td>2410 ± 180</td>
<td>16610 ± 220</td>
<td>12850 ± 150</td>
<td>3760 ± 130</td>
<td></td>
</tr>
</tbody>
</table>
Figure D-1: Variation of PN concentration (black), urban-regional scale (red) during the morning rush hour on June 19\textsuperscript{th} 2008.