Improving the accuracy of bicycle AADT estimation: temporal patterns, weather and bicycle AADT estimation methods

Thomas Nosal

Thesis

Department of Civil Engineering and Applied Mechanics

McGill University, Montreal

February 2014

Thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

CONTRIBUTION OF AUTHORS

Several researchers have contributed to the work that comprises this thesis. The second chapter, as cited, is based heavily on a research paper that was written in collaboration with my supervisor, Dr. Luis Miranda-Moreno, Dr. Robert Schneider of the University of Wisconsin – Milwaukee, and Frank Proulx, a PhD student at the University of California – Berkeley. The improvements upon that work are explained in the text, and the version that appears in this document was written in collaboration with Dr. Luis Miranda-Moreno. The third chapter is based on a research paper that was again written in collaboration with Dr. Luis Miranda-Moreno, and the fourth chapter was written in collaboration with both Dr. Miranda-Moreno and Zlatko Krstulich, who works with the City of Ottawa.

ACKNOWLEDGEMENTS

I acknowledge and appreciate the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Foundation for Innovation (CFI) for providing funds as operating and equipment grants. I would also like to thank the corresponding transportation agencies of Montreal, Ottawa and Vancouver, as well as Vélo Québec for providing bicycle data. I owe tremendous thanks to David Stephens for his statistical assistance, to Jean-Francois Rheault from Eco-counter for his technical support, and to David Beitel, also of Eco-Counter, for his technical support and data-collection companionship. Finally, I would like to thank my supervisor, Luis Miranda-Moreno. There were many times throughout the two years I spent working on this thesis that I entered his office in despair, only to leave uplifted, motivated and with greater confidence in my capabilities.

ABSTRACT

Bicycle commuting, investments in bicycle infrastructure, and programs to promote cycling are becoming more common in North America. Thus, it is becoming increasingly necessary to understand how to properly integrate cycling into an urban transportation system. This requires studying how different infrastructure affects the safety and comfort of cyclists, how certain programs and policies affect cycling use, and so on. Cycling research often requires estimating average annual daily bicyclist traffic (bicycle AADT). Bicycle AADT at a given location is generally estimated either directly – by installing a permanent bicycle counter for one year and dividing the total number of cyclists by 365 – or indirectly, using data from a long-term data collection site to adjust short-term counts. The most common indirect AADT estimation method is the expansion factor method. It has long been applied to automobile traffic counts, but researchers have recently shown that the expansion factor method often produces inaccurate bicycle AADT estimates. There are two primary reasons for this inaccuracy. First, the temporal patterns at the long-term and short-term sites may not match each other. Second, the expansion factor method does not account for the weather conditions experienced during the short-term count. The objective of this thesis is to address these sources of bicycle AADT estimation error.

To address the first source of error, this work utilizes a data set comprised of data from multiple North American cities and bicycle facilities to develop four factor groups for bicycle traffic patterns. These groups range from primarily utilitarian to primarily recreational, with mixed groups in between. The relationship between these factor groups and land-use, the built environment and demographic information is investigated to help guide the matching of short-term sites to appropriate factor groups. To address the second source of error – weather conditions experienced during the short-term count – an analysis of the relationship between bicycle counts and weather is conducted, demonstrating how weather conditions affect bicycle counts across different North American cities and across factor groups. A model is developed for the purpose of adjusting short-term counts for weather conditions.

Two novel bicycle AADT estimation methods that account for the weather conditions experienced during the short-term count are proposed and evaluated. These methods produce bicycle AADT estimates with lower average errors than those obtained via traditional estimation methods. The effects of factors like the duration of the short-term count and the selection of different factor groups on bicycle AADT accuracy are investigated.

RESUME

Le nombre de voyages fait en vélos, ainsi que l'investissement dans les infrastructures et programmes visant à promouvoir le cyclisme, est en croissance en Amérique du Nord. Il est donc de plus en plus nécessaire d'étudier et de comprendre comment bien intégrer le cyclisme dans un système de transport urbain. Cela nécessite la compréhension des effets que les différentes infrastructures ont sur la sécurité et le confort des cyclistes, ainsi que l'effet que certains programmes et politiques ont sur la pratique du vélo, et ainsi de suite. La recherche concernant le cyclisme nécessite souvent des données de comptage, et l'un des indicateurs les plus importants et fréquemment utilisés lorsqu'il s'agit de données de comptage est la moyenne des cyclistes quotidiens annuels. À un endroit donné, cet indicateur est généralement estimée soit directement - par l'installation d'un compteur de vélo permanent pour une année et en divisant le nombre total de cyclistes par 365 - ou indirectement, en utilisant les données d'un site de collecte de données à long terme pour ajuster les comptages à court terme. La méthode d'estimation indirecte la plus courante est la méthode de facteur d'expansion. Il est utilisé depuis longtemps pour ajuster les comptages de la circulation automobile, mais les chercheurs ont récemment montré que les facteurs d'expansion ne parviennent souvent pas à produire des estimations précises quant aux comptages de cycliste. Il existe deux raisons principales pour cela. Tout d'abord, les tendances temporelles aux sites de collecte à long terme et à court terme peuvent ne pas correspondre l'un à l'autre. Deuxièmement, la méthode de facteur d'expansion ne tient pas compte des conditions météorologiques rencontrées au cours du comptage à court terme. L'objectif de cette thèse est de répondre à ces sources d'erreur dans l'estimation des comptages de cyclistes.

Pour répondre à la première source d'erreur, ce travail utilise une base de données composé de données de plusieurs villes et pistes cyclables nord-américaines pour développer des groupes de facteurs pour les différents types de trafic de vélo. Ces groupes varient entre une classification principalement utilitaire et principalement récréative, avec des groupes mixtes entre les deux. Le rapport entre ces groupes de facteurs et l'utilisation des terres, l'environnement bâti et de l'information démographique est étudiée pour guider l'attribution des sites à court terme à des groupes de facteurs appropriés. Pour répondre à la deuxième source d'erreur (les conditions météorologiques pendant le comptage à court terme), une analyse approfondie est menée sur la

relation entre les comptages de cyclistes et la météo. Cela démontre comment les conditions météorologiques affectent le comptage de cyclistes dans les différentes villes d'Amérique du Nord et entre les différents groupes de facteurs. Un modèle a été développé dans le but d'ajuster les comptages à court terme pour les conditions météorologiques.

Deux nouvelles méthodes d'estimation de la moyenne des cyclistes quotidiens annuels qui tiennent compte des conditions météorologiques rencontrées pendant le comptage à court terme sont proposées et évaluées. Ces méthodes produisent des estimations avec des erreurs moyennes plus faibles. Les effets de facteurs tels que la durée du comptage à court terme et la sélection de différents groupes de facteurs sur la précision de l’estimation sont étudiées.

TABLE OF CONTENTS

CONTRIBUTION OF AUTHORS ii
ACKNOWLEDGEMENTS iii
ABSTRACT iv
RESUME v
LIST OF TABLES ix
LIST OF FIGURES x
CHAPTER 1: INTRODUCTION 1
1.1. CONTEXT AND MOTIVATION 2
1.2. OBJECTIVES AND THESIS STRUCTURE 4
1.3. BACKGROUND 5
CHAPTER 2: BICYCLE TRAFFIC PATTERNS: FACTOR GROUPS AND LAND-USE 8
2.1 INTRODUCTION 9
2.2 LITERATURE REVIEW 10
2.2.1 Temporal patterns and Factor Groups 10
2.2.2 Temporal Patterns and Land-Use 10
2.2.3 Measuring land-use, demographic and BE with GIS 11
2.3 METHODOLOGY 13
2.3.1. Study Locations 13
2.3.2. Data 13
2.3.3 Definition of standardized indices 14
2.3.4 Development of Factor Groups – Cluster Analysis 16
2.3.5 Measurement of Land-use, Built Environment, and Demographic Characteristics 16
2.4. RESULTS AND DISCUSSION 18
2.4.1. Cluster Analysis 18
2.4.2 Land-use, Built Environment, and Demographic Results 23
2.5 CONCLUSIONS 27
CHAPTER 3: THE RELATIONSHIP BETWEEN BICYCLE COUNTS AND WEATHER 28
3.1. INTRODUCTION 29
3.3 METHODOLOGY 32
3.3.1 Study Areas 32
3.3.2 Data 33

3.3.3 Counter classification according to temporal patterns ... 34
3.3.4 Absolute Bike Count Modeling ... 34
3.3.5 Relative Bicycle Count Modeling ... 36
3.4 RESULTS ... 38
3.4.1 Hourly Absolute Model Results ... 38
3.4.2 Daily Absolute Model Results ... 42
3.4.3. Daily Relative Model Results ... 45
3.5 DISCUSSION ... 47
3.6 CONCLUSION ... 49
CHAPTER 4: AADB ESTIMATION METHODS AND ACCURACY ... 50
4.1. INTRODUCTION ... 51
4.2. LITERATURE REVIEW ... 52
4.3 METHODOLOGY ... 53
4.3.1. Study Locations ... 54
4.3.2. Data ... 56
4.3.3. Weather Model Formulation ... 57
4.3.4. AADB Estimation Methods ... 58
4.3.5. Evaluation of Accuracy ... 62
4.4. RESULTS AND DISCUSSION ... 63
4.4.1. Weather Model ... 63
4.4.2. AADB Estimation ... 64
4.5. CONCLUSION ... 72
CHAPTER 5: CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH ... 73
5.1 FINAL COMMENTS ... 74
5.2 FUTURE RESEARCH ... 76
REFERENCES ... 77

LIST OF TABLES

Table 1. Summary of Bicycle Facility Factor Groups 12
Table 2. Bicycle Counter Locations 15
Table 3. Hourly and Daily Profiles of Final Factor Groups 19
Table 4. Summary of Final Factor Groups 22
Table 5. Summary of the urban and Green Route counting stations 32
Table 6. Hourly Modeling Results – Effects of Temperature, Humidity and Precipitation 41
Table 7. Weekday and Weekend Daily-Level Model Results – Effects of Temperature, Humidity and Precipitation 44
Table 8. Weekday and Weekend Relative Model Results for Montreal Counter Locations 46
Table 9. Short-Term and Long-Term Test Sites 57
Table 10. Weather Model Coefficients - Montreal 64

LIST OF FIGURES

Figure 1. Daily Bicycle Totals and AADB 5
Figure 2. Daily expansion factors from typical utilitarian and recreational facilities 7
Figure 3. Average daily bicyclists in May by maximum daily temperature 7
Figure 4. Example of Bike Network Buffer 17
Figure 5. Hourly and Daily Profiles within Mixed Cluster 18
Figure 6. Monthly Expansion Factors 20
Figure 7. Average land use composition by factor group and for ORP counter 23
Figure 8. Typical Utilitarian Location 24
Figure 9. Typical recreational and mixed-recreational counter locations 25
Figure 10. Intersection, population and employee density by factor group 25
Figure 11. Ottawa Bicycle Counter Locations 26
Figure 12.Elasticity of temperature and humidity - hourly models 39
Figure 13. Autocorrelation plots for regression model with and without ARMA errors 40
Figure 14. Elasticity of temperature and humidity – daily models 43
Figure 15. Montreal and Ottawa bicycle counter locations 55
Figure 16. AAPE by Estimation method and Short-Term Count Location 65
Figure 17. Monthly Expansion Factors at Montreal Locations in 2008 66
Figure 18. Effect of Long-Term Site Location on AAPE of O_S1 67
Figure 19. Day of Week Expansion Factors for Select Ottawa Locations 67
Figure 20. Effect of Weather Conditions on AAPE by AADB Estimation Method 69
Figure 21. Effect of Time of Short-Term Count on AAPE by AADB Estimation Method 70
Figure 22. Effect of Duration of Short-Term Count on AAPE by AADB Estimation Method ... 71

CHAPTER 1:
INTRODUCTION

1.1. CONTEXT AND MOTIVATION

In order to both promote cycling and accommodate the growing number of cyclists, many North American cities are investing in bicycle infrastructure (e.g., cycle tracks, bicycle lanes, bicycle parking) and are implementing new policies and programs (e.g., bicycle sharing, complete streets, bicycle integration with transit) (Pucher et al. 2010; Pucher and Beuhler, 2011). As cities continue to invest and the number of cyclists grows further, it is becoming increasingly necessary to study the safety, operations and efficacy of cycling infrastructure and programs. Further study will help researchers and practitioners identify what factors increase or decrease cyclist injury risk, evaluate the impacts of programs designed to promote cycling, monitor the use of new and existing infrastructure, and so on. This will guide the future investments of planners, engineers and government agencies.

Research regarding cyclist infrastructure and programs often requires cyclist count data. For example, a city cannot determine whether the installation of a cycle track increased bicycle traffic on a given street without knowing how many cyclists rode there before and after it was installed. An important and often essential metric for analyzing bicycle count data is average annual daily bicycle traffic, or bicycle AADT. Simply put, the bicycle AADT at a given location is the total number of cyclists that have passed over the course of a year divided by 365 (FHWA 2001). It is used in a wide range of studies because it removes temporal biases inherent in traffic counts, such as seasonal or day-of-week biases.

Bicycle AADT can be estimated with either direct or indirect methods. The direct method is applicable at long-term data collection sites which have permanent, automatic bicycle counters, such as inductive loop counters. If the counter has been installed for at least one year, bicycle AADT can be computed directly from the total cyclist count. While straightforward and generally most accurate, this is both cost and time consuming; it requires that a counter be installed for at least one year in all locations for which an AADT estimate is desired. Therefore, indirect AADT estimation methods have been developed which use data from long-term counting sites to extrapolate short-term counts taken at other locations. Short-term counts can be obtained manually by an observer or with temporary counter installations such as pneumatic tube counters. The strategy that many cities, such as Portland, OR, then employ is to maintain a set of long-term bicycle counters and to supplement those sites with short-term counts at other locations (City of Portland 2012).

The most common indirect AADT estimation method is the expansion factor method. Though specific adaptations vary, in general, data from a long-term counter site is used to develop factors that adjust a short-term count for hourly, day-of-week, and monthly bias. Research and development regarding expansion factors has mainly taken place in the context of motor vehicle traffic. However, recent studies have demonstrated that when applied to bicycle counts, expansion factors often produce inaccurate bicycle AADT estimates (Nordback et al. 2013; Esawey et al. 2013; Figliozzi et al. 2014). There are two primary reasons for inaccurate bicycle AADT estimates. First, the temporal patterns of the long-term site, from which the expansion factors are derived, may not match the temporal patterns of the short-term site. This would be the case if expansion factors from a long-term site that is used mainly by commuters were used to adjust a short-term count taken at a site used by recreational riders. The second reason is that expansion factors generally do not account for the weather conditions experienced during the short-term count. Weather conditions that increase or decrease cycling during the short-term count period can result in high or low AADT estimates, respectively.

Developing reliable means of estimating bicycle AADT would benefit many cities in North America, like Portland, Oregon, which have counting programs that include short-term counts (City of Portland 2012). It would increase the utility of the National Bicycle and Pedestrian Documentation Project, which facilitates the collection of short-term counts across the US (Jones 2009). It would also help improve the quality of a wide range of bicycle studies and projects, such as level-of-service studies (Allen et al. 1998), injury studies (Strauss et al. 2013; Lusk et al. 2013), before-and-after cyclist count studies (Parker et al. 2011), and so on.

Note that throughout the rest of the document, the term "average annual daily bicyclists" (AADB) will be used in place of "bicycle AADT". While "bicycle AADT" may be a more appropriate term – considering that bicycles are a component of traffic, and that methods to estimate bicycle AADT have largely been developed for and are applied to other modes – "AADB" is used for simplicity.

1.2. OBJECTIVES AND THESIS STRUCTURE

Failure to account for temporal and weather-related variation in bicycle counts is the primary source of error when estimating AADB from short-term counts. Most studies that have incorporated the temporal patterns of bicycle traffic have done so on a case-specific basis, or not with the goal of improving AADB estimates. Furthermore, no known research has been conducted that attempts to guide the linkage between short-term sites and long-term sites with similar temporal patterns. With regards to weather conditions and bicycle counts, researchers have demonstrated that bicycle counts are dependent upon weather, but very little work has been done to incorporate that dependence into AADB estimation methods. In light of these short-comings, this thesis has the following objectives:

i. Develop bicycle traffic factor groups to provide different sets of expansion factors for sites with different temporal patterns,
ii. Explore the connection between bicycle traffic patterns and land-use and the built-environment in order to guide the matching of short-term and long-term count sites,
iii. Explore the relationship between bicycle counts and weather conditions to understand how it differs across North American locations and across factor groups,
iv. Develop and evaluate AADB estimation methods that account for variation due to weather conditions and temporal patterns, and
v. Compare the accuracy of proposed AADB estimation methods to traditional methods.

The following subsection presents the background necessary for understanding the framework of this thesis. Each of the following three chapters includes a literature review of research pertaining to that chapter. The first of the three addresses the first and second objectives, presenting factor groups and exploring the relationship between bicycle traffic volume patterns and land-use and the built environment. The next chapter is about the relationship between cycling and weather and addresses the third and fourth objectives. The next chapter presents methods to estimate AADB and addresses the fourth objective as well, in addition to the fifth. The final chapter presents the conclusions, as well as some direction for future research.

1.3. BACKGROUND

As noted earlier, AADB at a given location is equivalent to the total number of cyclists who pass in a year, divided by 365. For example, in **Figure 1,** the grey line represents the daily cyclist totals at a site in Montreal, and the red line represents the average of the daily totals, the AADB. The clear variation throughout the year with respect to the AADB highlights the utility of AADB. If a researcher divided the number of injuries observed at a site by the number of cyclists observed in the spring, he or she would obtain a much higher injury risk estimate than if he or she had used counts observed in the summer. Likewise, if one wanted to determine whether the use of a given facility had changed from one year to the next and compared counts from the fall of the first year to counts from the summer of the second year, one would incorrectly conclude that there had been a large increase. Some counting programs attempt to account for this by counting year after year in the same month, but a cursory examination of **Figure 1** reveals that considerable variation can take place even within each month.

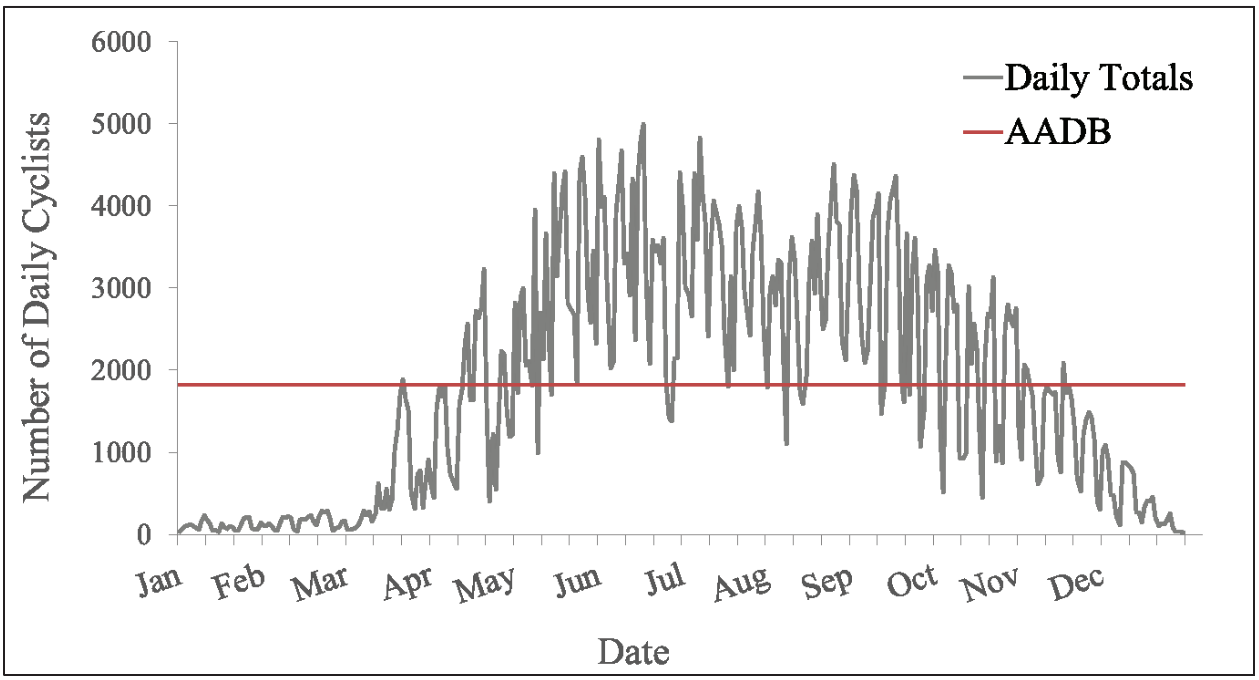


Figure 1. Daily Bicycle Totals and AADB

Accounting for this variation throughout the year is one of the primary motivations for estimating AADB, and as noted earlier, the most common indirect method for doing so is the expansion factor method. Expansion factors are generally calculated using data from a long-term site as follows (FHWA 2001; PIARC 2003):

$$I_h = (\bar{v}_h / AADB) \quad \textbf{(Equation 1)}$$

$$I_d = (\bar{v}_d / AADB) \quad \textbf{(Equation 2)}$$

$$I_m = (\bar{v}_m / AADB) \quad \textbf{(Equation 3)}$$

where:

I_h, I_d, I_m = hourly, daily and monthly expansion factors, respectfully

$\bar{v}_h, \bar{v}_d, \bar{v}_m$ = averages for a given hour *h*, day of the week *d*, or month *m*, respectively.

These expansion factors would be applied to a short-term count in order to estimate AADB in the following manner:

$$\widehat{AADB}_i = Y_{d,m} * \frac{1}{I_d} * \frac{1}{I_m} \quad \textbf{(Equation 4)}$$

where

$\widehat{AADB}_i$ is the estimated average annual daily bicyclists for site *i,*

$Y_{d,m}$ is an observed short-term (in this case 24-hour) cyclist count on day of the week *d* in month *m*,

Note that in this case, the short-term count would be 24 hours long; if an observed count is less than a full day, then an additional expansion factor for the hour of the day is introduced.

There are two primary reasons for inaccurate AADB estimates. First, the patterns at the long-term site from which the expansion factors are derived may not match the underlying behavior of the short-term site. For instance, day-of-week (I_d) expansion factors from typical utilitarian and recreational bicycle facilities are plotted in **Figure 2**. Utilitarian locations are used more heavily for commutes to work and school or for errands, and therefore experience higher use during the week than on weekends; recreational locations are used for leisure and exercise, and therefore experience greater use on the weekend. Say a short-term count was taken at a recreational location. It is clear that if the expansion factors from a utilitarian site were used to adjust that count, the resulting AADB estimate would be too low. Researchers have developed so-called factor groups, which provide separate sets of expansion factors for different types of locations. Miranda-Moreno (2013) identified four factor groups – *primarily utilitarian, mixed-utilitarian, mixed-recreational,* and *primarily recreational*. Using the set of factors that best corresponds to the temporal patterns of the short-term site is expected to improve the accuracy of AADB estimates. However, further research is necessary to validate and refine these factor groups, and help guide the process of matching short-term sites to appropriate long-term sites.

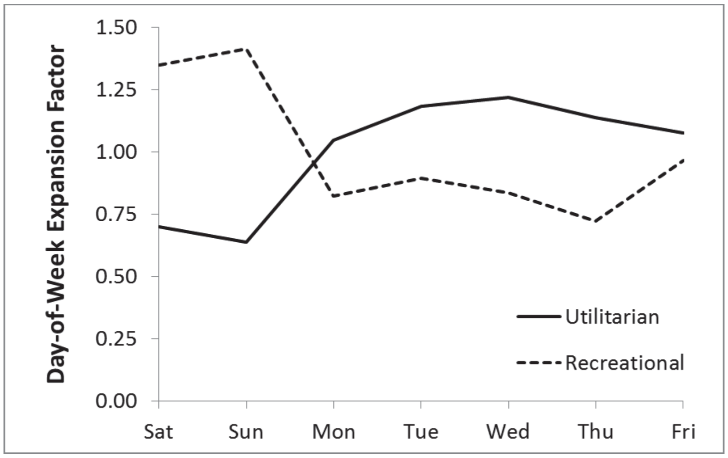


Figure 2. Daily expansion factors from typical utilitarian and recreational facilities

The second primary reason for inaccurate AADB estimates is that the expansion factor method does not adequately account for variation due to weather conditions. **Figure 3**, which utilizes data from a long-term count station in Montreal, shows the average daily bicyclists in May for days within two temperature ranges. It is clear that if a short-term count was collected on either a relatively cold or warm day in May, it would produce an AADB estimate that was lower or higher, respectively, than the true value. Further research is necessary to develop AADB estimation methods that account for weather-related variation.

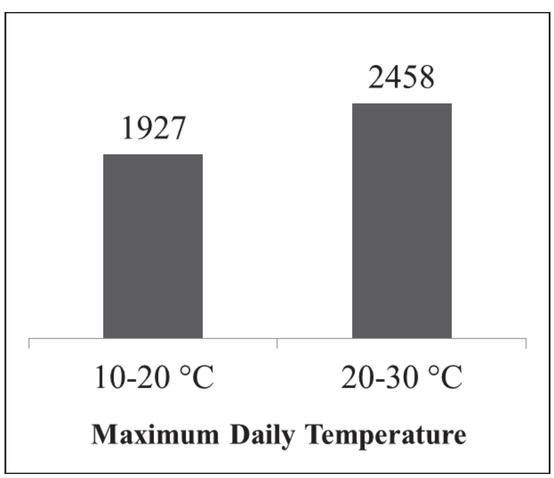


Figure 3. Average daily bicyclists in May by maximum daily temperature

CHAPTER 2:
BICYCLE TRAFFIC PATTERNS: FACTOR GROUPS AND LAND-USE

2.1 INTRODUCTION

As noted in the **Chapter 1**, the expansion factor method is complicated by the fact that different locations exhibit different temporal traffic patterns, necessitating multiple sets of hourly, daily, and monthly expansion factors. To address this issue, researchers have developed factor groups – groups of count locations that exhibit similar temporal traffic patterns. Though most research has been devoted to motorized traffic, Miranda-Moreno et al. (2013) defined four factor groups for bicycle traffic, which range from primarily recreational to primarily utilitarian, with mixed categories in between. If a short-term site fits into one of these groups, the corresponding factors can be used to estimate AADB for that site. However, their groups were defined using a relatively primitive procedure that involved visually observing the patterns of a large set of bicycle data. This chapter will use an iterative clustering technique, k-means clustering, to validate the factor groups determined by Miranda-Moreno et al. (2013).

There is no clear way to match short-term count sites to factor groups. Sharma et al. (1996) noted that even more important to the accuracy of an AADT estimate than the duration of the short-term count is the correct assignment of expansion factors. Currently, the most common method for matching short-term count locations to factor groups is educated guessing based on knowledge of the study area (Sprinkle 2011). Research efforts are needed to help guide factor group assignment. Therefore, this chapter also investigates whether a link can be made between the obtained bicycle factor groups and land-use, demographics, and the built environment. In particular, this work seeks to investigate what factors may be associated with bicycle traffic patterns and how they can be measured using GIS. This is intended to serve as a springboard for further investigation, which may ultimately result in a formal assignment method for short-term counts to factor groups.

Note that this research only considers the relationship between land-use and factor groups as defined by hourly and daily patterns. Though monthly and seasonal factors differ across factor groups, this variation is related more to climate and is not considered at present.

2.2 LITERATURE REVIEW

2.2.1 Temporal patterns and Factor Groups

Based on temporal traffic patterns and expansion factors presented in **Section 1.3**, Miranda-Moreno et al. (2013) classified 37 counter locations into four groups – *recreational, mixed-recreational, mixed-utilitarian,* and *utilitarian* - which are described in **Table 1**. Their procedure involved plotting the hourly and daily expansion factors of the 37 locations to identify how the facilities are used temporally and to group similar counting locations together. In general, *primarily utilitarian* locations exhibit AM and PM traffic peaks on weekdays, and have higher ridership on weekdays than over the weekend; conversely, primarily recreational locations exhibit only one peak around noon and have higher ridership on the weekend than on weekdays. Mixed-utilitarian facilities exhibit AM and PM commuting peaks, but they have greater use on the weekends. Mixed-recreational facilities do not exhibit distinctive AM and PM commuting peaks, but they also exhibit similar levels of use on weekdays and on the weekend.

2.2.2 Temporal Patterns and Land-Use

While other studies have classified bicycle counter locations as either utilitarian or recreational based on temporal patterns or functional classifications, the only known research that has associated temporal traffic patterns with land-use, demographics and socioeconomics has concerned motorized traffic. Using a sample of motorized traffic counter sites in Florida, Yang et al (2009) used regression analysis to directly relate monthly expansion factors to variables including the percentage of seasonal households, the percentage of museum, art gallery or garden employees, and specific land-uses like universities or amusement parks. They demonstrated the power of associating traffic patterns to these variables by matching test sites to control sites, based on the similarity of variables they found to be significant, and estimating AADT with less than 10% average error.

Using a similar regression analysis, Li et al. (2004) showed significant relationships between seasonal expansion factors on Florida roads and hotel and motel density, seasonal households, retired population density and retail employment. They also noted that roadway functional classification was not significantly linked to the seasonal factors, which conflicts with advice provided in several traffic monitoring guides (Robichau and Gordon 2003).

With a sample of over 1000 short-term motorized traffic count sites, Hernandez (2012) used cluster analysis to identify 7 factor groups within the City of Winnipeg. He determined the composition of the land use surrounding each site using simple buffers. Based on the average land-use composition, he demonstrated apparent associations between certain groups and commercial, residential, industrial, and scholastic parcels. However, his research was examining variation strictly within an urban area, with sites that behave primarily like the utilitarian bicycle category. With a limited number of sites, such a nuanced analysis is not feasible here, but Hernandez's results do suggest the ability to show links between patterns and land use.

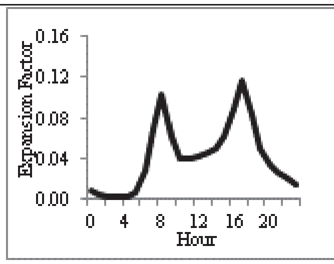
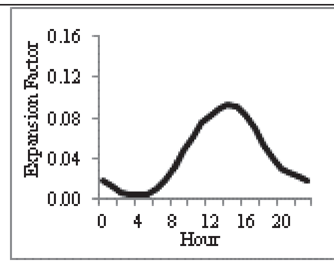
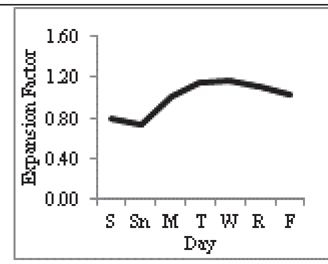
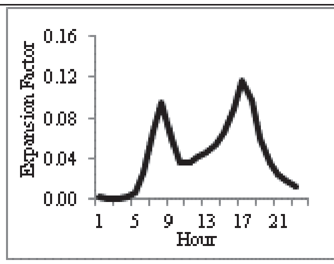
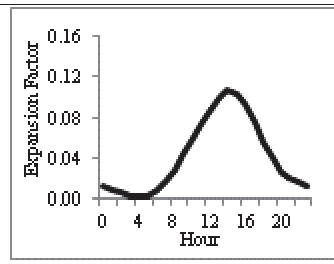
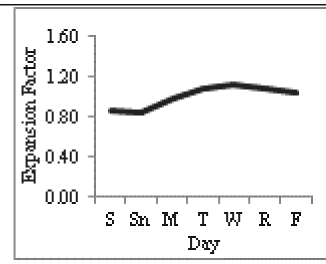
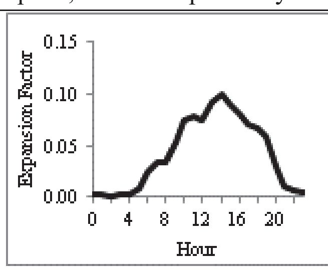
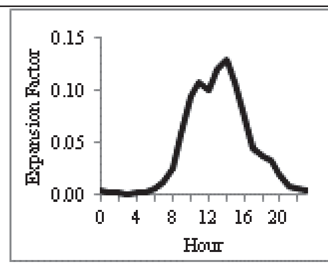
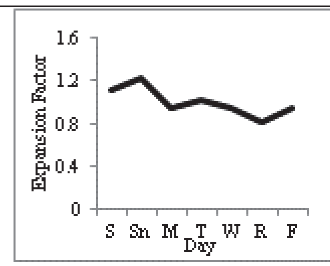
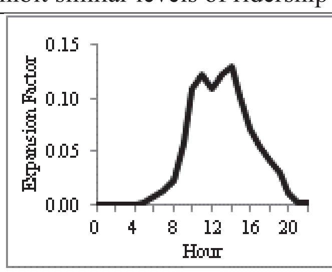
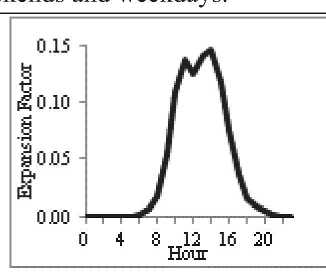
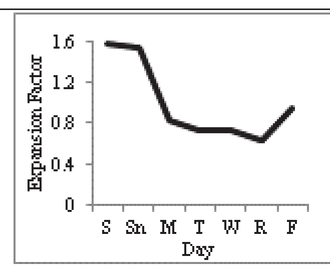
2.2.3 Measuring land-use, demographic and BE with GIS

A literature review was conducted to identify measures of land-use, demographics and the built environment that are suitable for the available data. Entropy indexes were considered for measuring the composition or diversity of the built environment. However, several researchers have noted that such indexes can mask relevant information, that land use mix is less important than land use complementarity, and that the implications of particular land uses are more clear when they are kept separate (Hess et al. 2001; Brown 2009). Therefore, composite indexes in general were not utilized for this analysis. Measures that capture demographics, or can serve as proxies for land-use, are population, household, and employee density (Robitaille 2009; Song and Rodriguez 2004). Indications of built environment and design elements include intersection and block density (Robitaille 2009; Song and Rodriguez 2004).

With regards to unit of analysis, simple circular buffers centered on the counter location were considered, but Nicholls (2001) found that network-based buffers produce more reliable results, as they are a more realistic representation of what people can actually access from a given location. In their study described above, Li et al (2004) measured land-use and demographics using both circular and linear buffers (along the roadway) that emanated from the counter locations, finding that R^2 and the significance and impact of variables did not vary much between models using either type.

Finally, it has been shown that accessibility to destinations is strongly related to travel demand (Ewing and Cervero 2010). It is not a stretch to assume that accessibility can also affect traffic patterns. Because buffers provide only information regarding the area they contain, basic gravity-based accessibility measures may provide a better measure of the land-uses surrounding and influencing a counter site (Zhang et al 2011).

Table 1. Summary of Bicycle Facility Factor Groups

<table>
<tr><th></th><th colspan="2">Hourly Profiles[1]</th><th rowspan="2">Daily Profile[1]</th></tr>
<tr><th>Type</th><th>Weekday</th><th>Weekend</th></tr>
<tr><td rowspan="2">Primarily Utilitarian</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Utilitarian locations exhibit distinct AM and PM peaks on weekdays, one PM peak on the weekend. They have much higher ridership during the week than on the weekend.</td></tr>
<tr><td rowspan="2">Mixed – Utilitarian</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Mixed-utilitarian locations still exhibit AM and PM peaks on weekdays, though the midday ridership may be slightly higher than at primarily utilitarian locations. The difference between weekday and weekend ridership is much less pronounced, and may even be negligible. Weekend ridership builds gradually to a PM peak, similar to primarily utilitarian locations.</td></tr>
<tr><td rowspan="2">Mixed-Recreational</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Mixed-recreational locations exhibit one PM peak, as opposed to both AM and PM peaks. They may experience slightly higher AM ridership on weekdays than primiarily recreational locations. They also exhibit similar levels of ridership on weekends and weekdays.</td></tr>
<tr><td rowspan="2">Primarily Recreational</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Primarily recreational locations exhibit one PM peak on weekdays and the weekend. They exhibit considerably higher ridership on the weekend than during the week. The workweek hourly profile closely resembles the weekend profile, which exhibits higher AM ridership than either utilitarian group and slightly higher AM ridership than the mixed recreational group, producing more of a midday plateau.</td></tr>
</table>

2.3 METHODOLOGY

2.3.1. Study Locations

This chapter utilizes count data from counting stations in five North American cities, as well as along the *Route Verte* in Quebec. The *Route Verte* is primarily operated by Velo Quebec (VQ), a non-profit cycling advocacy and research organization. This is the same data set used by Miranda-Moreno et al. (2013) with some additional data. There are six locations in Montreal (referred to as Mon1-Mon6), four in Ottawa (Ott1-Ott4), one in Portland (Port), eight in San Francisco (SF1-SF8), six in Vancouver (Van1-Van6), and sixteen along the *Route Verte*, in cities and towns from Montreal to Quebec City (VQ1-VQ16). The locations chosen for this study have some of the most extensive sets of automated bicycle count data in North America. A brief description of each counter location involved in the analysis is presented in **Table 2**.

2.3.2. Data

Bicycle Count Data

Count data for the 37 long-term count locations was collected using inductive loop bicycle counters, which count bicycles by detecting changes in the electric current in sub-pavement loops of cable. They can distinguish cyclists from automobiles in mixed traffic, making them suitable for both segregated and non-segregated facilities. The data are continuously logged in 15-minute intervals and for the purpose of this research were aggregated into hourly and daily totals. The performance of this technology has been documented by Nordback et al (2011), who reported that, on separated facilities, the absolute error between the true number of cyclists and the counted number is generally less than 3% when the counters are operating properly. Some bicycle facilities in Montreal, Ottawa, and along the *Route Verte* are not maintained during winter and therefore do not have count data. In order to be consistent, this analysis only incorporated data from the cycling season, April through November (inclusive). All data were collected between 2008 and 2012, and each facility had at least one season of data.

Before analysis, each dataset was reviewed thoroughly to identify missing values, which can be caused by routine maintenance, counter malfunction, construction, or other factors. In addition, an effort was made to identify and remove extremely high values, which, for instance, can be caused by large bicycle races or group rides on the *Route Verte*.

Geospatial Data

Geospatial data was obtained from the Transportation Research at McGill (TRAM) GIS data archive. Base maps were made for each geographic location that consisted of shapefiles containing political outlines, street centerlines, and bicycle infrastructure centerlines. Demographic data (population, employment, etc.) are derived from the 2006 Canadian census and are aggregated at the census tract level. The land use data set, from 2007, identifies parcels as commercial, governmental and institutional, open space, parks and recreational, resource and industrial, or waterbody. Geospatial data was not available for all location, so this portion of the analysis was restricted to a subset of the sites, which is denoted in **Table 2**. All of the sites for which geospatial data were available are in Montreal, Ottawa, and along the *Route Verte* in Quebec. No land-use data was available for Portland, San Francisco, or Vancouver.

2.3.3 Definition of standardized indices

Two traffic distribution indices, WWI and AMI, are used later to quickly summarize the distribution of bicycle traffic throughout the day, week or year (**Equations 5 – 6**).

$$I_{we/wd} = (\bar{v}_{we}/\bar{v}_{wd}) \qquad \textbf{(Equation 5)}$$

where:

$I_{we/wd}$ = relative index of weekend vs. weekday cycling traffic (WWI).

$\bar{v}_{we}, \bar{v}_{wd}$ = seasonal average daily weekend and weekday traffic, respectively.

$$I_{AM/Mid} = \frac{\delta_i^{AM}}{\delta_i^{Mid}} \qquad \textbf{(Equation 6)}$$

where:

$I_{AM/Mid}$ = relative index of morning (hours 7:00 to 9:00) to midday (hours 11:00 to 13:00) cycling traffic (AMI).

$\delta_i^{AM} = \sum_{h=7}^{9} \bar{V}_h$

$\delta_i^{Mid} = \sum_{h=11}^{13} \bar{V}_h$

Table 2. Bicycle Counter Locations

Region	Facility	Location	Facility Type	AADB[1]
Montreal, QC, Canada	**Mon1**	Maisonneuve at Peel[2]	Cycle Track	2200
	Mon2	Maisonneuve at Berri[2]	Cycle Track	4324
	Mon3	Brebeuf at Rachel[2]	Cycle Track	3736
	Mon4	Berri at Maisonneuve[2]	Cycle Track	3735
	Mon5	Cote Ste Catherine at Stuart	Cycle Track	1489
	Mon6	Jacques Cartier Bridge	Multiuse Path on Bridge	1724
Ottawa, ON, Canada	**Ott1**	Ottawa River Pathway[2]	Multiuse Path	1637
	Ott2	Colonel By Pathway[2]	Multiuse Path	832
	Ott3	Laurier[2]	Segregated Bike Lanes	1390
	Ott4	Alexandra Bridge	Separated Bikeway	1250
Portland,	**Port**	Hawthorne Bridge	Separated Bikeway	4869
San Francisco, CA, USA	**SF1**	Northpoint at Polk	Paired Bicycle Lanes	421
	SF2	Polk at Grove	Unidirectional Bicycle Lane	404
	SF3	Potrero at 23rd St.	Paired Bicycle Lanes	259
	SF4	Valencia at 14th St.	Paired Bicycle Lanes	2475
	SF5	Seventh Ave. at Kirkham	Paired Bicycle Lanes	156
	SF6	Panhandle at Masonic	Multiuse Path in Park	3452
	SF7	Lake at Arguello	Paired Bicycle Lanes	188
	SF8	Arguello at Lake	Paired Bicycle Lanes	511
Vancouver, BC, Canada	**Van1**	Burrard St. Bridge	Bicycle Lane Separated Path	3004
	Van2	Canada Line Bridge	Bike/Ped Bridge	336
	Van3	Cambie St. Bridge	Separated Bikeway	990
	Van4	CV Greenway at Rupert	Bicycle Path	543
	Van5	CV Greenway at Victoria	Bicycle Path	805
	Van6	Ontario at 11th St.	Bicycle Boulevard	788
Velo Quebec *Route Verte* (Locations across province of Quebec)	**VQ1**	Métabéchouan	Asphalt Bicycle Path	232
	VQ2	Duschesnay[2]	Gravel Bicycle Path	154
	VQ3	Québec[2]	Asphalt Bicycle Path	1015
	VQ4	Lennoxville[2]	Gravel Bicycle Path	207
	VQ5	Lévis[2]	Asphalt Bicycle Path	1034
	VQ6	Cabano[2]	Gravel Bicycle Path	144
	VQ7	St-Jean-sur-Richelieu[2]	Asphalt Bicycle Path	200
	VQ8	Longueuil[2]	Asphalt Bicycle Path	413
	VQ9	Cushing[2]	Asphalt Bicycle Path	112
	VQ10	Laval[2]	Asphalt Bicycle Path	590
	VQ11	Granby[2]	Asphalt Bicycle Path	267
	VQ12	Mont-Rolland[2]	Gravel Bicycle Path	359
	VQ13	Trois Rivières[2]	Asphalt Bicycle Path	542
	VQ14	Victoriaville[2]	Asphalt Bicycle Path	62
	VQ15	Gatineau[2]	Asphalt Bicycle Path	34
	VQ16	Blainville[2]	Asphalt Bicycle Path	135

[1]Average Daily Volume over study period (April – November)
[2]Locations for which geospatial data was available

2.3.4 Development of Factor Groups – Cluster Analysis

The k-means technique, which is an iterative clustering technique, was used to develop the factor groups, based on the hourly and daily expansion factors (**Equations 1-3**). The goal of this technique is to group n observations (each of which can be a multidimensional vector) into k clusters such that the within-cluster variance is minimized. The iterative procedure starts with initial mean values for each cluster, and then alternates between two steps. In the first step, each observation is assigned to the cluster such that the minimum within-cluster variance is produced. In the next step, the new means are calculated for each cluster. This procedure continues until no observations are moved in the first step.

The Calinski-Harabasz (CH) criterion was used to help guide the selection of the number of clusters, k. It can be represented as follows:

$$\frac{SS_{BC}}{SS_{WC}} x \frac{(n-k)}{(k-1)}$$

where SS_{BC} is the between-cluster variance, SS_{WC} is the within-cluster variance, and n is the number of observations. Large differences between cluster centers and low variation within clusters is desirable, so large CH values indicate favorable cluster outcomes. If a plot of CH values vs. k exhibits a "kink", then it may not necessarily be advantageous to choose one of the outcomes with similar CH values on the upper portion of the kink over another.

The *cluster kmeans* command in STATA was used to run this analysis, with the default dissimilarity measure, Euclidean distance. Because the final clusters can change depending on the initial mean values, random observations were chosen for cluster centers, and the random seed was changed multiple times to verify that major changes in the final outcome did not occur (particularly that the CH value did not change drastically).

2.3.5 Measurement of Land-use, Built Environment, and Demographic Characteristics

The unit of analysis was a buffer based on travel distance along the bike infrastructure network. The ArcMap Network Analyst extension was used to create buffers covering all points that could be reached along the bike network and within 2.5 km of each counter, which corresponds to 10 minutes of cycling at an average speed of roughly 15 km/h (El-Geneidy et al. 2007). An example of the buffer used for a typical counter location is presented in **Figure 4.**

After defining the unit of analysis, the ArcMap Intersect tool was used to determine the land-use compositions of the buffers surrounding each counter, represented simply by the proportion of each type of land-use contained in the buffer. In addition, population density surrounding each counter was calculated by first calculating the gross population density of each census tract. The tracts were intersected with the buffers, and a weighted average (based on the total area of each tract within the buffer) was used to estimate the buffer's density. The same method was used to calculate employee density. Finally, intersection density was estimated by dividing the total number of intersections (as determined by Network Analyst) by the buffer area.

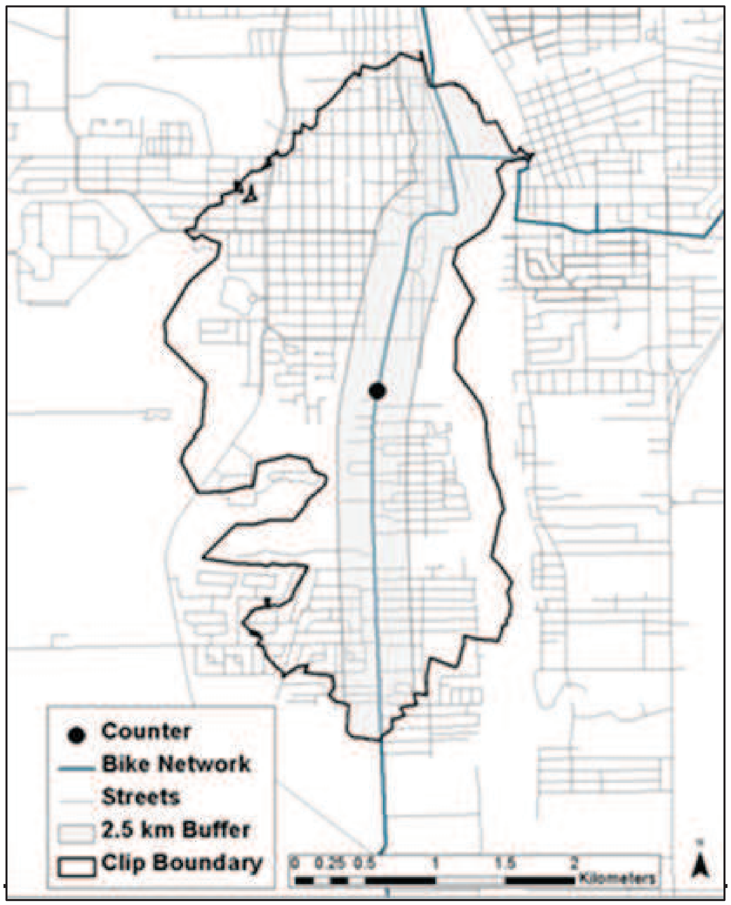


Figure 4. Example of Bike Network Buffer

2.4. RESULTS AND DISCUSSION

2.4.1. Cluster Analysis

The CH values indicated that, based on both the hourly and daily expansion factors, *k* equal to two, three and four would produce comparably distinct clusters; more than four clusters was inferior. The CH value for *k* equal to three was highest, *k* equal to four yielded two separate clusters with mixed patterns, and *k* equal to two is too few to be meaningful in this context. Therefore, it was decided that the final clusters would be based on *k* equal to three. Two of the clusters clearly contained the utilitarian and recreational sites, respectively, and the third cluster contained the mixed locations (**Figure 5**). This third cluster was split further with a separate k-means cluster analysis with *k* equal to two, based this time solely on hourly expansion factors.

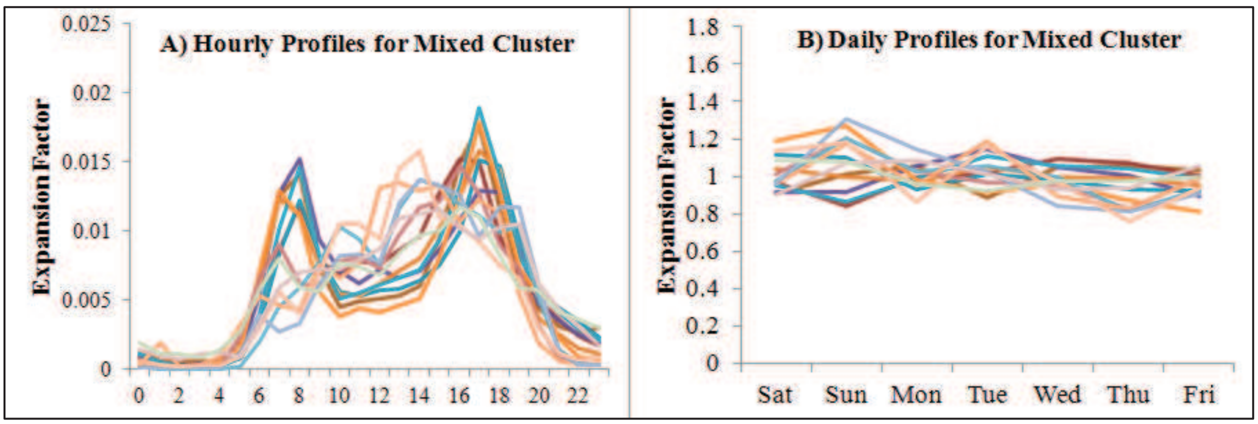
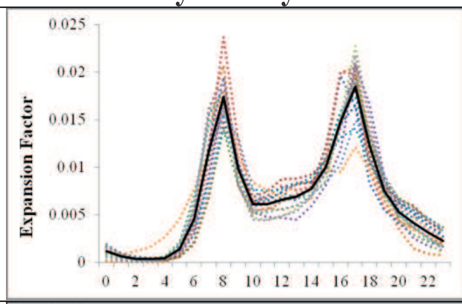
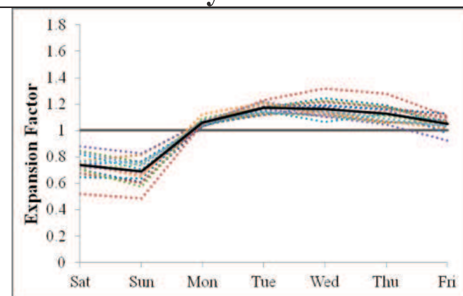
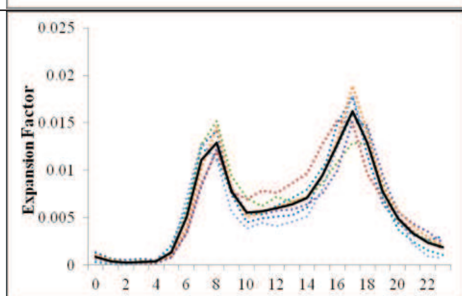
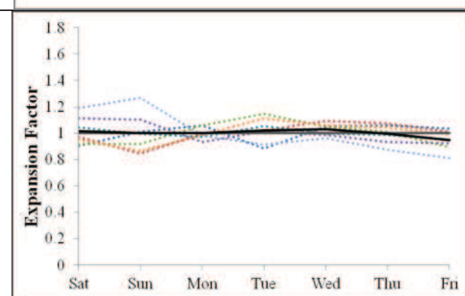
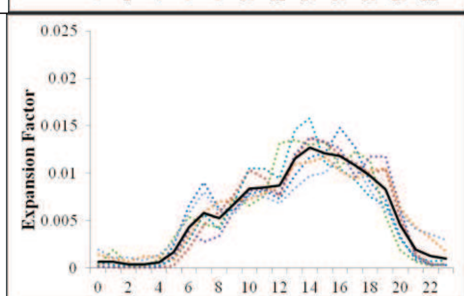
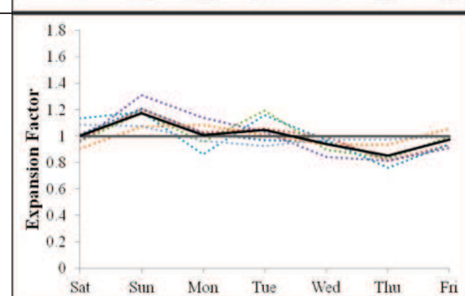
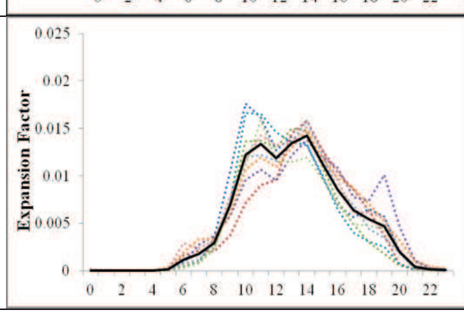
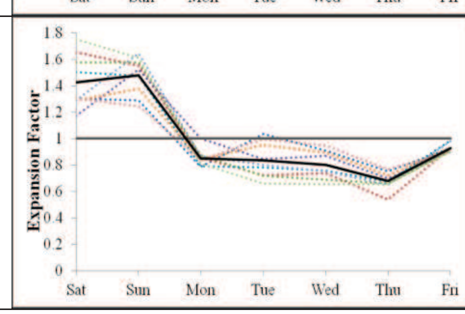


Figure 5. Hourly and Daily Profiles within Mixed Cluster

The resulting factor groups are very similar to those obtained in Miranda-Moreno et al. (2013), and the hourly and daily profiles within each cluster, in addition to the mean values (black lines) are presented in **Table 3**. The four factor groups are again labeled *primarily utilitarian, mixed-utilitarian, mixed-recreational* and *primarily recreational*. Again, the *primarily utilitarian* locations exhibit AM an d PM peaks on weekdays, and they experience the greatest use throughout the week. *Mixed-utilitarian* sites again experience AM and PM peaks, but their daily profiles are relatively flat; they are used equally as much on the weekends as on weekdays. Note that the maximum values of the AM and PM peaks are lower than those of the *primarily utilitarian* group. The *primarily recreational* sites exhibit only one peak, which appears to occur close to noon, but slightly before or after depending on the location. They experience greater use on the weekend than during the week. The *mixed-recreational* group also experiences one peak, which seems to occur in the PM. *Mixed-recreational* locations also exhibit greater ridership during the AM and usage appears to be distributed more thoroughly across the day. *Mixed-recreational* sites also have relatively consisted use throughout the week, but in this case appear to be slightly more skewed towards the weekend than the *mixed-utilitarian* sites are towards the weekdays.

Table 3. Hourly and Daily Profiles of Final Factor Groups

Group	Weekday Hourly Profiles	Daily Profiles
Primarily Utilitarian		
Mixed-Utilitarian		
Mixed-Recreational		
Primarily Recreational		

Hourly and daily patterns presented in **Table 3** appear relatively consistent across regions; i.e. primarily utilitarian locations in Vancouver exhibit very similar hourly profiles to those of primarily utilitarian facilities in Montreal, Ottawa, and Portland, for example. However, monthly patterns vary considerably across both classifications and regions (**Figure 6**). For example, when examining seasonal data from April through November, Vancouver's utilitarian facilities retain higher ridership in November than both Montreal and Ottawa, suggesting that because Vancouver has warmer winters, more of its utilitarian cyclists ride year-round. However, the Velo Quebec facilities, which share a climate similar to that of Montreal and Ottawa, retain far less ridership in the winter than both cities, presumably because recreational trips are more sensitive to cold weather.

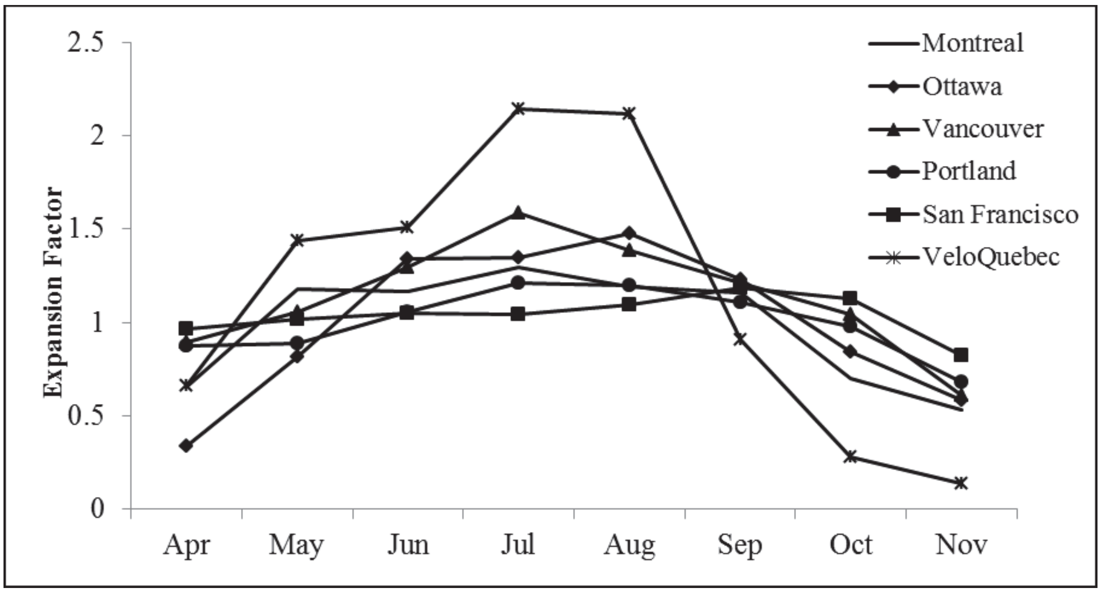


Figure 6. Monthly Expansion Factors

One counter location, SF8, was excluded from the final clusters for anomalous behavior. Though it exhibits AM and PM weekday peaks, indicative of *primarily utilitarian* or *mixed utilitarian* behavior, it has significantly higher volumes on the weekends than on weekdays; the WWI value for SF8 is 1.88, while the other *mixed-utilitarian* facilities have values much closer to 1.0. It is suspected that this anomalous behavior is due to the fact that SF8 is located on the border of Presidio National Park and is close to the Golden Gate Bridge. High tourist traffic, which would typically occur on the weekend (in addition to local recreational traffic) could explain why the daily profile is similar to that of a recreational site, while the hourly profile still exhibits utilitarian behavior. This highlights the need for more data in future analyses, and suggests that count locations near tourist sites may need particularly close attention when assigning factor groups and estimating AADB.

A summary of the locations within each group is presented in **Table 4**, along with the AADB values, the WWI and AMI traffic summary indexes presented in **Section 2.2.3**, and the groups to which each site was assigned in Miranda-Moreno et al. (2013). Overall, the groups are consistent with those obtained by Miranda-Moreno et al. (2013). One location which had previously been classified as *primarily utilitarian* was in this case classified as *mixed-utilitarian*, and four locations which had been classified as *mixed-recreational* were classified as *primarily recreational*. Again, as suggested by Miranda-Moreno et al. (2013), it seems quite possible that with more locations, factor groups would emerge that would split the current recreational group, producing more nuanced levels within the more recreational sites. Within this particular dataset, the AADB values decrease as sites become less utilitarian in nature.

Table 4. Summary of Final Factor Groups

Group	Facility	AADB	WWI	AMI	Prior Group[1]
Utilitarian	Mon1	2200	0.56	1.62	Utilitarian
	Mon2	4324	0.56	1.26	Utilitarian
	Mon3	3736	0.65	1.66	Utilitarian
	Mon4	3735	0.72	1.63	Utilitarian
	Mon5	1489	0.56	1.94	NA[2]
	Ott1	1637	0.67	2.52	Utilitarian
	Ott3	1390	0.43	2.38	Utilitarian
	Ott4	1250	0.69	2.59	NA[2]
	Port	4869	0.54	1.68	Utilitarian
	SF1	421	0.77	2.62	Utilitarian
	SF2	404	0.72	1.3	Utilitarian
	SF3	259	0.65	1.53	Utilitarian
	SF4	2475	0.71	1.77	Utilitarian
	Van3	990	0.61	2.05	Utilitarian
	Van4	543	0.53	2.93	Utilitarian
	Van5	805	0.56	2.97	Utilitarian
	Van6	788	0.72	2.17	Utilitarian
	Mean	**1842**	**0.63**	**2.04**	**---**
Mixed-Utilitarian	Mon6	1724	0.94	2.66	NA[2]
	Ott2	832	0.85	1.16	Mixed-Util.
	SF5	156	0.82	1.82	Utilitarian
	SF6	3452	1.08	1.62	Mixed-Util.
	SF7	188	1.01	1.7	Mixed-Util.
	Van1	3004	0.84	1.76	Mixed-Util.
	Van2	336	1.1	2.17	Mixed-Util.
	Mean	**1386**	**0.95**	**1.84**	**---**
Mixed-Recreational	VQ5	1034	1.25	0.48	Mixed-Rec.
	VQ7	200	0.96	0.71	Mixed-Rec.
	VQ8	413	1.14	0.88	Mixed-Rec.
	VQ10	590	1.2	0.85	Mixed-Rec.
	VQ13	542	1.08	0.63	Mixed-Rec.
	VQ14	62	1.2	0.53	Mixed-Rec.
	VQ15	34	1.2	0.44	Mixed-Rec.
	Mean	**411**	**1.14**	**0.65**	**---**
Recreational	VQ1	232	1.47	0.32	Mixed-Rec.
	VQ2	154	1.88	0.27	Recreational
	VQ3	1015	1.51	0.35	Mixed-Rec.
	VQ4	207	1.77	0.32	Recreational
	VQ6	144	1.34	0.32	Mixed-Rec.
	VQ9	112	2.26	0.26	Recreational
	VQ11	267	1.81	0.33	Recreational
	VQ12	359	2.13	0.24	Recreational
	VQ16	135	1.47	0.44	Mixed-Rec.
	Mean	**292**	**1.74**	**0.32**	**---**

[1]Classification obtained in Miranda-Moreno et al. (2013)
[2]Data was not incorporated in Miranda-Moreno et al. (2013)

2.4.2 Land-use, Built Environment, and Demographic Results

The average land use compositions for each factor group are presented in **Figure 7,** including 95% confidence intervals based on the t-distribution. Although classified as *utilitarian*, the Ott1 site, on the Ottawa River Pathway, exhibited drastically different land-use characteristics from the other *utilitarian* facilities. It was excluded from the average for the *utilitarian* group and is presented separately to facilitate discussion. Note that there was only one *mixed-utilitarian* counter with available geospatial data and therefore it has no confidence interval.

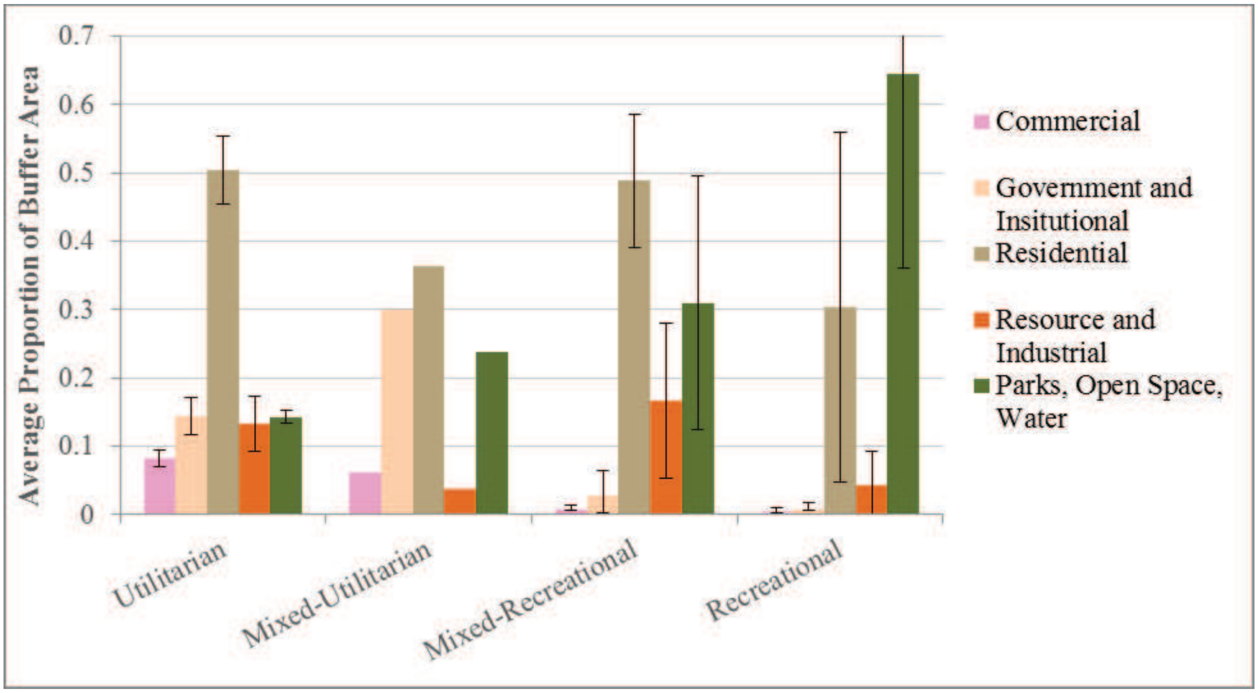


Figure 7. Average land use composition by factor group and for ORP counter

Recreational facilities are associated with the highest level of parks, open space, and water bodies, while *mixed-recreational* sites exhibit a much lower amount; *utilitarian* locations exhibit the least amount of parks, open space, and water bodies, with the *mixed-utilitarian* site in between. Conversely, *recreational* sites exhibit almost no commercial space, while *mixed-recreational* and *utilitarian* sites are associated with successively increasing levels of commercial land-use. This is intuitive as commercial land uses are associated with utilitarian trips like commutes to work or school and errands, and open space are not. *Utilitarian* and *mixed-recreational* sites are associated with comparable residential space, while recreational sites exhibit the least amount. *Utilitarian* sites in general exhibit a greater mix of land-uses. (**Figures 7 - 9**)

Average intersection, population, and employee density all increase from *recreational* to *utilitarian* locations (**Figure 10**).

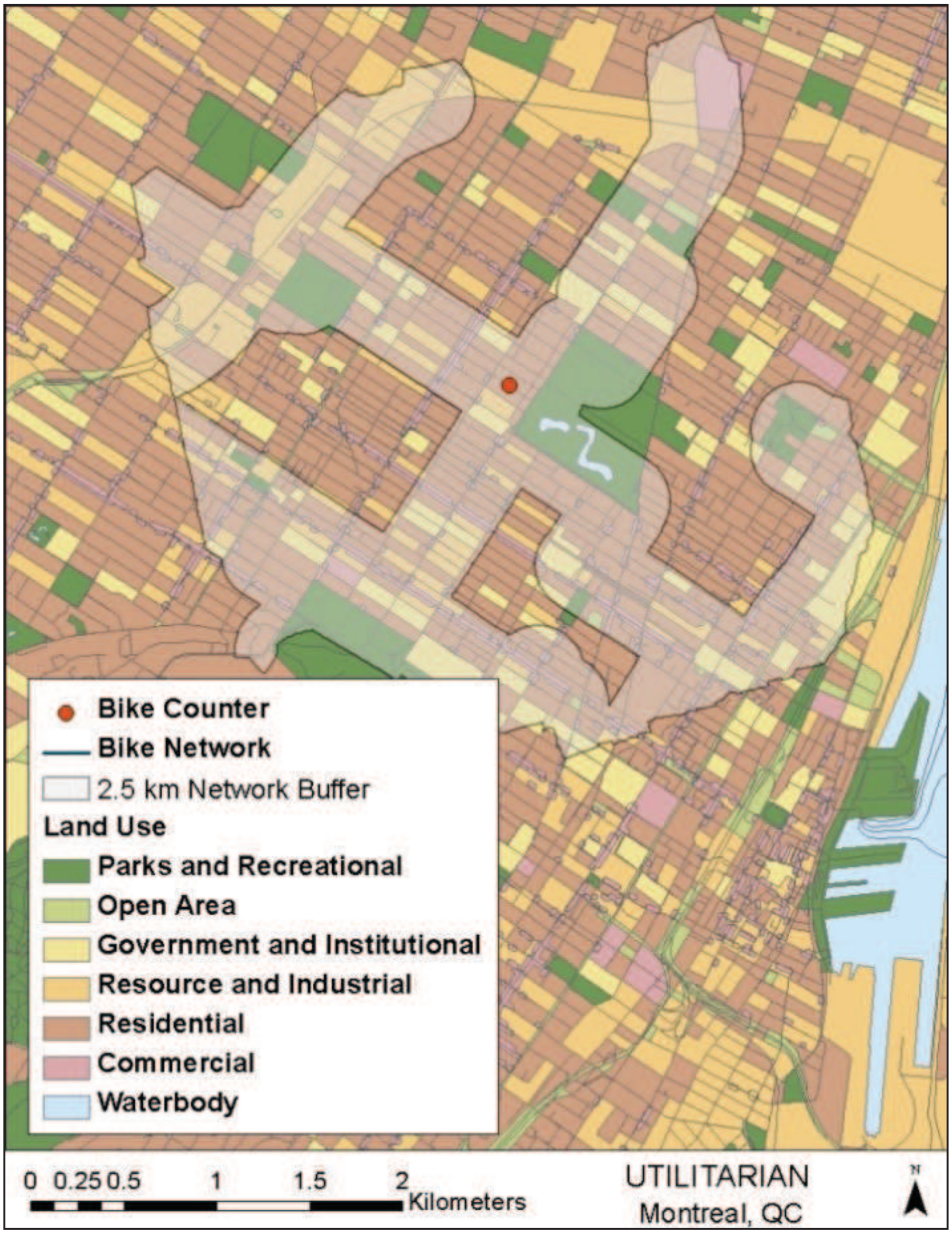


Figure 8. Typical Utilitarian Location

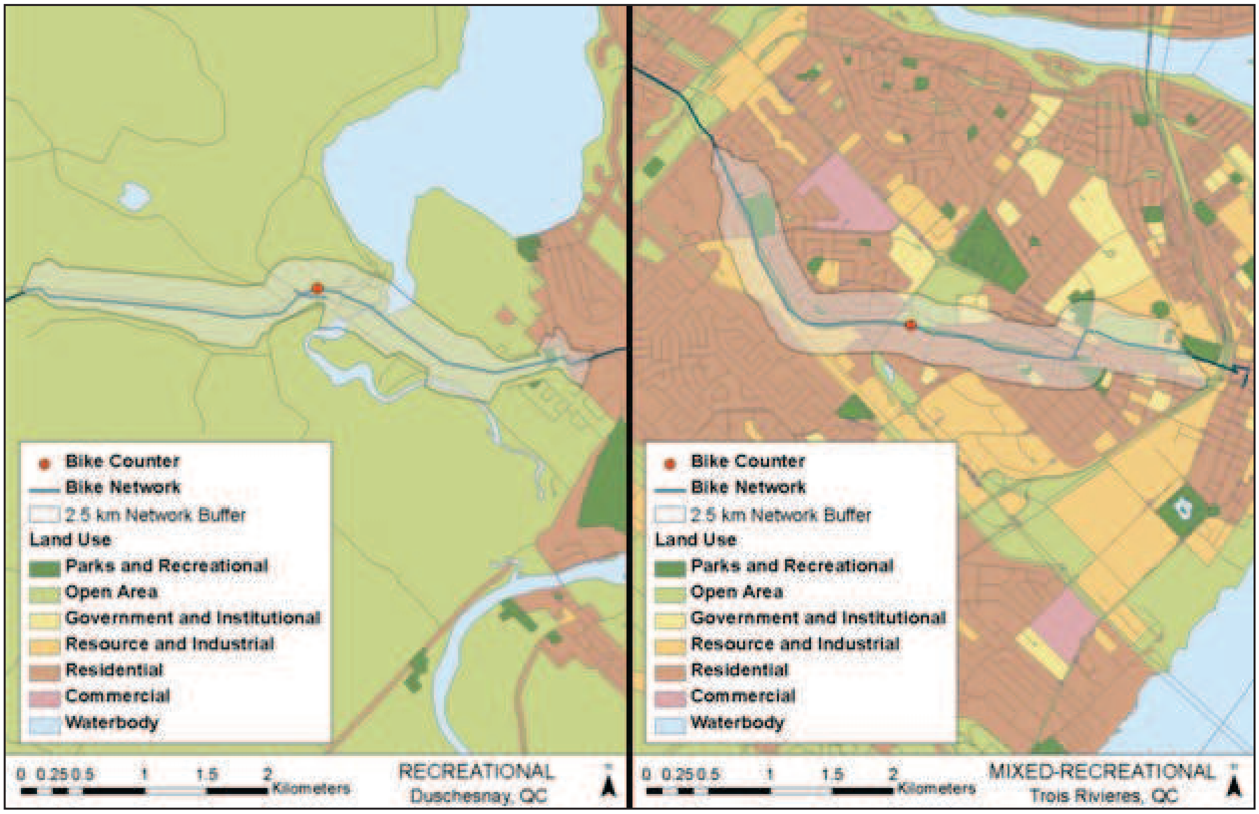


Figure 9. Typical recreational and mixed-recreational counter locations

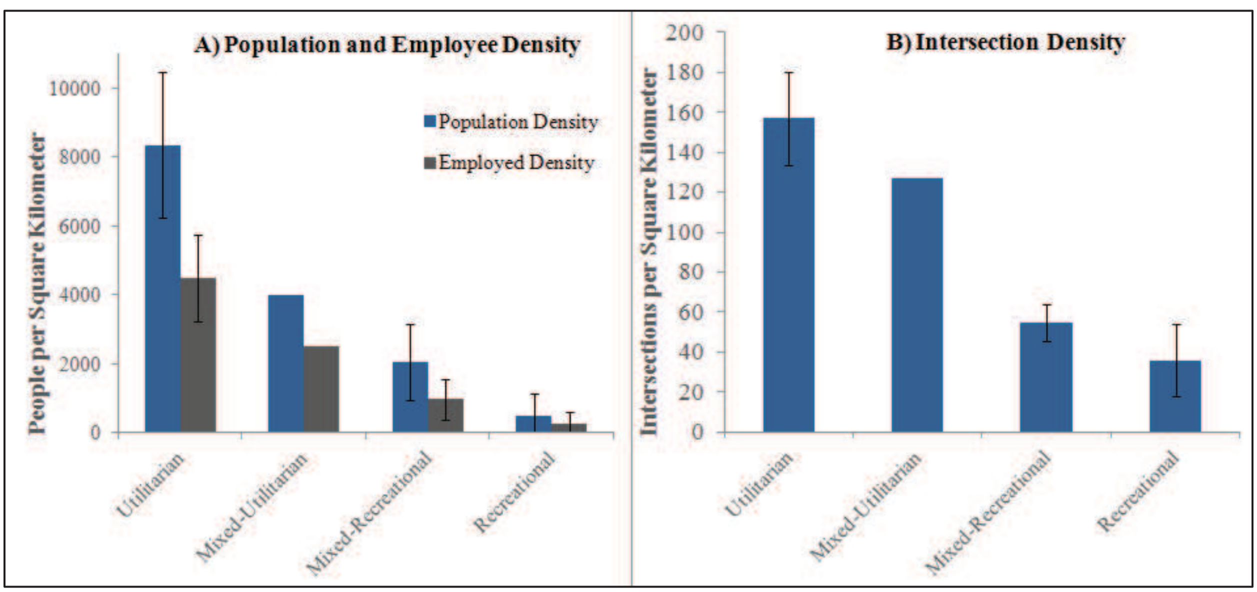


Figure 10. Intersection, population and employee density by factor group

The counters in Ottawa (**Figure 11)** exemplify the challenges with relating land-use to temporal traffic patterns. Note that only the buffers associated with Ott1 and Ott2 are shown here. First, Ott2 and Ott3 are less than a kilometer apart, yet Ott3 exhibits *primarily utilitarian* patterns while Ott2 exhibits *mixed-utilitarian* patterns. The *mixed-utilitarian* site, Ott2, exhibits high levels of park space, but also is associated with some commercial space (**Figure 7** and **Figure 11**); this likely explains the mixed nature of its use. However, Ott1 is located on the Ottawa River Pathway, and despite being associated with a relatively high level of open space and no commercial activity (**Figure 11),** which would suggest mixed or even recreational use, it is classified as *primarily utilitarian* by its temporal patterns. This highlights the complexity of urban travel behavior. Origins and destinations that result in utilitarian traffic patterns are served by the link which this counter is on, but they are too far to be captured by the buffer that was utilized. Much further work will be necessary to determine appropriate buffers or means of capturing such complex interactions between land-use, both near and far, and temporal patterns.

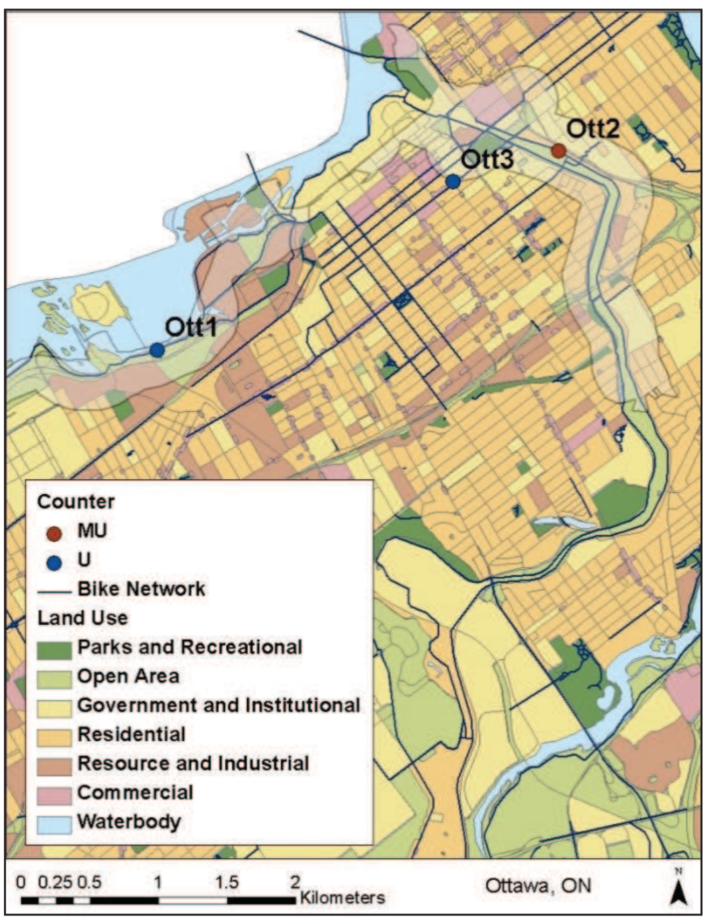


Figure 11. Ottawa Bicycle Counter Locations

2.5 CONCLUSIONS

The k-means clustering technique was used to demonstrate that a large set of long-term bicycle count locations can be classified into four factor groups – *primarily utilitarian, mixed-utilitarian, mixed-recreational* and *primarily recreational*. The groups are nearly identical to the groups defined by Miranda-Moreno et al. (2013).

Relatively high levels of commercial land-use, low levels of open space, and increased population, employee and intersection density appear to be associated with a higher amount of utilitarian cycling at a given location. Conversely, low levels of commercial land-use, high levels of open space, and very low levels of population, employee, and intersection density are associated with a higher amount of recreational cycling at a given location. While these associations cannot be referred to as causal, and they are certainly not concrete rules, they generally support the intuition and educated guesses typically made by researchers and practitioners when assigning short-term count sites to long-term sites. However, even with extreme care, certain sites are liable to exhibit counter-intuitive patterns that will confound AADB estimates.

As exemplified by the Ottawa River Pathway counter, bicycle traffic patterns cannot always be explained by information gathered in the immediate vicinity of a counter. This highlights the complexity of the relationship between temporal traffic patterns and land-use. It also makes selecting an appropriate unit of analysis extremely difficult. While gravity-based measures may offer some help at overcoming the limitations of buffers, much more work and data is necessary to determine appropriate weights and metrics.

CHAPTER 3:
THE RELATIONSHIP BETWEEN BICYCLE COUNTS AND WEATHER

3.1. INTRODUCTION

Although several recent studies have been published on this topic (Lewin, 2011; Miranda-Moreno and Kho, 2012; Miranda-Moreno and Nosal, 2011; Rose et al., 2011; Thomas et al., 2013), there are several shortcomings in the literature. Most studies have used survey data, brief manual counts, or daily (aggregate) data, which cannot capture the effects of hourly (disaggregate) weather conditions, such as the effect of rain in the morning versus rain in the evening. Most past studies have either focused on one specific geographical area or a particular type of bicycle facility; few works have examined how the relationship between cycling and weather differs across cities or across different types of facilities. Very little research has examined the difference between the impact of weather on weekdays and the weekend.

This chapter will help address these shortcomings by using data from four North American cities – Montreal, Ottawa, Vancouver and Portland – and from locations along a bike network that spans Quebec, to develop hourly and daily cyclist ridership models. Separate models will be calibrated for utilitarian and recreational locations, as classified according to the factor groups presented in **Chapter 2**. Classifying locations in this way prior to modeling is advantageous because it is well established that facility locations within each group have the same temporal patterns. There is less assurance that count locations classified according to functional classification or practical experience actually exhibit the same temporal behavior.

While a number of researchers have pointed out that understanding the relationship between cycling and weather can help develop methods for adjusting short-term counts, few have developed models which are geared for that specific purpose. This chapter presents both absolute and relative models. Absolute models, in which bike counts are modeled a function of weather variables and temporal factors, might be used to estimate missing data or to identify long-term trends, as Thomas et al. (2013) demonstrated. Relative models, in which deviations in cycling counts from average cycling counts are modeled as a function of corresponding deviations in weather conditions, are more readily suited to adjusting short-term counts, as will be demonstrated in **Chapter 4**.

3.2 LITERATURE REVIEW

Though the relationship between cycling and weather has been studied as early as 1977 (Hanson and Hanson 1977), there has recently been an increase in research, and a wide range of data collection methods and statistical analyses have been utilized. Data collection methods include surveys, as well as manual and automatic count data. For instance, Nankervis (1999a; b) used survey data and counts of parked bicycles on a university campus; Hanson and Hanson (1977) used travel survey data to estimate the daily cycling modal share, and more recently, Lewin (2011), Rose et al. (2011), and Thomas et al (2013) used automatic long-term counts. Most studies incorporate data aggregated at the daily level, but more studies are emerging that use hourly cycle counts (Gallop and Tse 2011; Miranda-Moreno and Nosal 2011; Tin Tin et al. 2012). Regression models, with bicycle counts or the logarithm of bicycle counts modeled as a function of weather variables and various temporal factors, make up the bulk of the statistical analyses (Brandenburg et al. 2007; Lewin 2011; Miranda-Moreno and Nosal 2011; Nankervis 1999a; Nankervis 1999b; Rose et al. 2011; Thomas et al. 2013; Tin Tin et al 2012).

In the literature, the two main weather determinants are temperature and rain, but others such as humidity and wind speed have been identified. In general, cycle counts increase with temperature. Though varying specifications make it difficult to compare directly across studies, an increase in temperature of one degree Celsius is generally associated with an increase in cycle counts of less than five percent (Tin Tin 2012; Miranda-Moreno and Nosal 2011). Two studies found the square of temperature to be insignificant when entered into a model with temperature, suggesting that the effect of temperature is linear (Tin Tin et al. 2012; Rose et al. 2011). However, Lewin (2011) and Miranda-Moreno et al. (2011) used binary variables to show that extremely high temperatures are associated with a decrease in cycle counts, and Richardson (2000) observed a non-linear effect. Miranda-Moreno and Nosal (2011) found that increases in humidity are associated with decreases in cycling. Only one known study has utilized a thermal index, which simultaneously incorporates temperature and humidity, to describe the perception of weather by cyclists (Brandenberg et al. 2007). Other variables that have been examined include hours of sunshine and wind speed (Thomas et al. 2013), and cloud coverage (Hanson and Hanson, 1977).

Regression models generally incorporate precipitation as a continuous variable, as the duration of precipitation (in hours, for example), or simply as a binary variable denoting the presence of precipitation. Lewin (2011) found that rainfall in a day decreases the count by about 10% of the annual daily average, and Tin Tin et al. (2012) found that cycle counts decrease by 1.5% and 10.6 % per millimeter of rain in a day and hour, respectively. Gallop and Tse (2011) and Miranda-Moreno and Nosal (2011) found that rain in one of the previous three hours can have an effect on cycle counts in the current hour comparable to or greater than rain in the current hour. Miranda-Moreno and Nosal (2011) also found that rain in the morning can reduce cycling counts in the afternoon. These results suggest that the complex relationship between cyclist counts and precipitation may require more complex representations of precipitation.

These studies are based in European, Australian, and North American locations. Studies have typically only considered one region or city at a time, with a few exceptions, such as the work of Rose et al. (2011) that included data from Portland, Oregon, USA and Melbourne, Australia. This study presents an aggregate (daily) ridership model to study the effects of weather on bicyclist volumes. As one of the main results, it is found that cyclists in the two cities (Melbourne and Portland) exhibit different sensitivities to weather.

While some studies make no distinction, several have examined the effects of weather on utilitarian and recreational cycling separately (Brandenburg 2007; Hanson and Hanson 1977; Richardson 2000; Thomas et al. 2013). Brandenburg (2007) observed that commuting occurs more in cooler weather than recreational cycling, and that commuters are less sensitive to rain than recreational cyclists. Hanson and Hanson (1977) found that the effect of weather was greater on discretionary trips than trips reportedly for work. Richardson's (2000) study looked at the effects of weather on cycle trips in Australia using travel survey data, observing that cycling recreational users are more affected by extreme temperatures and rainfall than utilitarian users are. Thomas et al. (2013) also noted that different user groups respond differently to weather; recreational cycling is much more sensitive to weather than utilitarian cycling. In other words, discretionary trips can be more easily put off than trips to work. Though he did not look at utilitarian vs. recreational cyclists, Nankervis (1999a; 1999b) concluded that certain cyclist groups, like students, may respond to weather differently from others. Finally, Thomas et al. (2013) examined weekend and weekday cycling separately and concluded that cycling on utilitarian facilities is more sensitive to weather conditions on weekends.

3.3 METHODOLOGY

3.3.1 Study Areas

This study utilizes cyclist count data from 10 automatic counting stations in four North American cities – one in Portland, OR, USA; four in Vancouver, BC, Canada; four in Montreal, QC, Canada; and one in Ottawa, ON, Canada - and from 3 counting stations along the Green Route (GR) in the Province of Quebec. A brief description of each location is provided in **Table 5**. For the purposes of this study, cycle-tracks are on-street, but physically separated, bi-directional bicycle facilities; separated paths are off-road facilities that may or may not be shared with pedestrians; sidewalk bike facilities are sidewalks that permit use by cyclists, generally over bridges; and bicycle boulevards are quiet streets optimized for cycling with traffic calming and markings. All of the counter locations in the four cities exhibit *primarily utilitarian* traffic patterns and all of the GR counters exhibit *primarily recreational* patterns, as presented in **Chapter 2**.

Table 5. Summary of the urban and Green Route counting stations

City	Name	Location	Type	ADV * (Standard Deviation)
Montreal	Mon1	Maisonneuve Blvd.	Cycle Track	1896 (1024)
	Mon2	Maisonneuve Blvd.	Cycle Track	3575 (1627)
	Mon3	Brebeuf St.	Cycle Track	3456 (1626)
	Mon4	Berri St.	Cycle Track	3735 (1972)
Ottawa	Ott1	Ottawa R. Path	Separated Path	1721 (813)
	Ott2	Alexandra Bridge	Sidewalk Bike Facility	1108 (513)
Vancouver	Van1	Cambie St Bridge	Sidewalk Bike Facility	909 (451)
	Van2	CV Greenway at Rupert	Separated Path	525 (234)
	Van3	CV Greenway at Victoria	Separated Path	779 (334)
	Van4	Ontario &11th St.	Bicycle Boulevard	859 (349)
Portland	Port1	Hawthorne Bridge	Sidewalk Bike Facility	4367 (1677)
Velo Quebec	VQ1	Métabéchouan,QC	Separated Path	316 (286)
	VQ2	Duschesnay, QC	Separated Path	210 (191)
	VQ3	Quebec City, QC	Separated Path	943 (1013)

*Average seasonal daily volumes corresponding to the period April 01 – 30 November

3.3.2 Data

Bicycle Count Data

The bicycle count data was collected using the same equipment described in **Section 2.3.2**. In the context of this chapter, both Montreal and Vancouver have data available for 2008 - 2010, but Ottawa has data for only 2009 – 2010, and Portland and most Green Route counters have data for only 2010. Again, as explained in **Section 2.3.2**, only data from April – November (inclusive) was analyzed. The relationship between cycling and weather differs between the winter and summer, and therefore a separate analysis was conducted by Miranda-Moreno and Kho (2012).

Weather Data

Weather data, consisting of hourly values for temperature, relative humidity, and precipitation, were obtained from weather stations maintained by Environment Canada (EC), the *Ministère du Développement durable, de l'Environnement et des Parcs* (MDDEP), and the National Oceanic and Atmospheric Administration (NOAA). The bicycle count data locations in Montreal, Vancouver, Portland and Ottawa were all matched with data from four respective weather stations, while each "Route Verte" location was matched to a separate station. The weather stations are typically located within 2-6 km of the bicycle counters.

Data Processing

The bicycle and weather datasets were joined by date (and time, when applicable), and were screened for missing or erroneous data. Data were discarded on holidays, and when graphical inspection revealed irregular behavior characteristic of counter malfunction, such as days with zero counts. Also discarded were days with missing weather data, as they could not be used in the models. Even if a day was missing only one or two hours of weather or bike data, it was discarded in full, as observed maximum/minimum temperatures, precipitation totals, and so on could be erroneous. With the exception of the Montreal and the Green Route locations, which were missing roughly 15% and 13% of their data, respectively, all sites were missing less than 5% of their data. Only 2% of the Vancouver hourly weather dataset was missing, but missing values were interspersed throughout the dataset, rather than in one or two large, easily-excluded chunks. Therefore, if three or less values in a row for temperature or humidity were missing, the missing values were replaced by linear interpolation. If more than three were missing then that day's data were discarded. After interpolation, only 0.3% of the dataset was excluded due to missing weather data.

Due to a high presence of non-missing zeroes in the overnight hours, data from 20:00 - 06:00 were excluded from the analysis. Because the log of cycle counts was used as the dependent variable, remaining non-missing 0 counts were changed to one. For all of the urban locations, less than 1 percent of the data needed to be changed from 0 to 1. Due to lower demand, particularly in the colder months, 15% of the Green Route observations were changed

3.3.3 Counter classification according to temporal patterns

The bicycle locations were classified as either utilitarian or recreational based on the framework presented in **Chapter 2**. The advantage of classifying count locations in this manner, as opposed to relying upon functional classification, local experience, or simply whether a location is urban or suburban/rural, is the assurance that all of the count data locations included in each category exhibit the same temporal patterns. This increases the likelihood that those locations along bike facilities in each category are being used in a similar manner. As demonstrated in **Chapter 2**, both urban cycling facilities and facilities on the green route can exhibit mixed patterns, or characteristics of both utilitarian and recreational locations; including such locations in either of the utilitarian and recreational groups would be erroneous. Of those used in the weather analysis, all of the count data locations in Montreal, Ottawa, Vancouver and Portland exhibit utilitarian patterns and all of those along the green route exhibit recreational patterns.

3.3.4 Absolute Bike Count Modeling

The relationship between weather and both hourly and daily cyclist volumes was analyzed using log-linear regression models. Separate models were calibrated for weekdays and weekends, and counting stations from the same city were used to calibrate one general model per city. After verifying that they exhibited similar results when calibrated separately, all of the Green Route count data locations were used to calibrate one general model. All of the models have the following functional form:

$$\mathbf{ln}(N_{k,h,d,m,y}) = \alpha + \beta X_{h,d,m,y} + \upsilon W_{h,d,m,y} + \gamma D_{k,h,d,m,y} + \omega_{h,k,d,m,y} \qquad \textbf{(Equation 7)}$$

where:

$N_{k,h,d,m,y}$	= the number of bicycle counts on path *k,* in an hour *h,* during day of the week *d,* month *m,* and year *y*;
$X_{h,d,m,y}$	= a vector of hourly weather variables during the same date and time (*h, d, m and y*) as $N_{k,h,d,m,y}$; These include continuous variables, discrete variables, and non-linear transformations of weather variables;
$W_{h,d,m,y}$	= a vector of variables that incorporate weather conditions from hours other than hour *h,* such as a binaryvariable set to one if it has rained in the three hours prior to *h*;
$D_{k,h,d,m,y}$	= a vector of dummy variables for path *k,* hour period *h,* day of the week *d*, month *m*, and year *y*;
$\omega_{h,k,d,m,y}$	= either a random, independent error term, or when accounting for serial autocorrelation, equivalent to $\phi_1 e_{h-1,k,d,m,y} + \cdots + \phi_p e_{h-p,k,d,m,y} - \theta_1 z_{h-1,k,d,m,y} - \cdots - \theta_q z_{h-q,k,d,m,y}$, where $e_h$is the error in hour h and $z_h$is a white noise process, and ϕ and θ are parameters to be estimated from the data;
$\alpha, \beta, \upsilon, \gamma$	= vectors of parameters to be estimated from the data.

In the case of daily level models, the subscript h is removed from all variables above and in the error and white noise terms, ϕ and θ, respectively, h is replaced with *d*.

Both continuous and binary variables were used to incorporate weather conditions into the models. Weather variables that were available across all cities were tested, and those that had a significant effect on bicycle ridership were retained. Non-linear (such as second-order polynomial) effects of temperature and humidity were tested; for instance, both temperature and the square of temperature were entered into the models. Binary variables were used to relate precipitation from previous hours to a given hour (hourly models), or to specific in which portion of the day rainfall occurred (daily models). The Akaike information (AIC) criterion was used to evaluate different model specifications and aid with model selection.

Correlation between independent variables was evaluated. If two weather variables had a correlation coefficient greater than 0.5, whichever variable had a weaker correlation with cyclist ridership was excluded. For example, both Vancouver and Portland exhibited a strong negative correlation between temperature and relative humidity. Therefore, only the models for Ottawa and Montreal incorporate humidity.

In addition to the model coefficients, elasticities are computed in the following sections to facilitate comparisons between variables with different units. Elasticity is generally defined as proportional change in the dependent variable relative to proportional changes in a given independent variable. The combined elasticity of a continuous variable represented as a quadratic function is ($\beta_k \cdot \overline{X}_k + 2 \cdot \beta_{k+1} \cdot \overline{X}_k^2$), where $\overline{X}_k$ is the mean value of variable k and β_k is the estimated coefficient for variable k. This value is the percent change in the cyclist count for an hour (or day) corresponding to a 1% increase in the independent weather variable. If one of the terms, β_k or β_{k+1}, is insignificant, then it is excluded from the equation. The elasticity of a binary variable is [$\exp(\beta_k) - 1$], which is the relative change in the cyclist count expected when the conditions described by the variable are present.

3.3.5 Relative Bicycle Count Modeling

Linear regression models with ARMA errors were also calibrated to model deviations in cycling counts from average levels as a function of corresponding deviations in weather conditions. Relative models were calibrated using only counts aggregated at the daily level, and again, separate models were calibrated for weekdays and weekends. Note that Miranda-Moreno and Nosal (2011) developed a similar model, but deviations in cycling and weather were calculated relative to values averaged by month. It was found here that using averages calculated within a moving 21-day period produced a better fit. For a given weekday, the bicyclist count deviation is relative to the average of all weekday cyclist totals spanning from 10 days before and 10 days after. For a given Saturday or Sunday, the bicyclist deviation is relative to the average of all weekend cyclist totals within a period spanning from 10 days before to 10 days after. A twenty-one day period was used because it always includes 15 weekdays and 6 weekend days. Weather conditions deviations are relative to values averaged over the entire 21 day period. The model can be represented as follows:

$\Delta DB_{y,j} = (\beta * \Delta W_{y,j}) + (\alpha * P_{y,j}) + (\gamma * D_{y,j}) + \varepsilon_{y,j}$, where **(Equation 8)**

$\Delta DB_{y,j}$ $= \left[DV_{y,j} - \frac{\sum_{k=j-10}^{k=j+10} DB_k}{N}\right] \Big/ \left[\frac{\sum_{k=j-10}^{k=j+10} DB_k}{N}\right]$, the relative Daily Bicyclists deviation of day j in year y, where j ranges from 1 to the number of days in the year or cycling season. Depending on the model, only weekdays or weekends are included in the averages, and N is equal to 15 or 6, respectively.

$\Delta W_{y,j}$ $= \left[W_{y,j} - \frac{\sum_{k=j-10}^{k=j+10} W_k}{21}\right] \Big/ \left[\frac{\sum_{k=j-10}^{k=j+10} W_k}{21}\right]$, a vector of deviations in continuous weather variables (temperature, dewpoint, total precipitation, etc.) from their respective 21 day moving averages on day j in year y. (Note that although the normalized version is shown above, these variables may or may not be normalized)

$P_{y,j}$ = a vector of binary variables related to precipitation on day j in year y,

$D_{y,j}$ = a vector of binary variables related to temporal effects, like the day of week on which day j in year falls,

β, α, γ = vectors of coefficients to be estimated from the data, and

$\varepsilon_{y,j}$ = either a random, independent error term for day j in year y, or when accounting for serial autocorrelation, equivalent to $\phi_1 e_{j-1} + \cdots + \phi_p e_{j-p,} - \theta_1 z_{j-1} - \cdots - \theta_q z_{j-q}$, where e_j is the error on day j, z_j is a white noise process, and ϕ and θ are parameters to be estimated from the data.

As in the case of the absolute models, multi-collinearity was checked to ensure that variables with correlation coefficients with absolute values greater than 0.5 were not included in the same model. The elasticity values for the continuous variables reported in the tables in the results sections are equal to ($\beta_k * 0.1$); for instance, for temperature, the elasticity value is equivalent to the relative change in cycling from the average corresponding to a 10% increase in temperature. The elasticity values reported for the binary variables are the coefficients themselves, and are equivalent to the relative change in cycling counts as a result of the presence of the conditions represented by the variable.

3.4 RESULTS

The results of the hourly and daily absolute models, as well as the relative model, are described in the following three subsections, respectively. Variables are significant at the five percent level if the t-statistic is great than 1.96. The model coefficients, t-statistic values, and elasticity values, are presented in each table. In order to conserve space, the fixed effects for month, day of the week, hour, and facility are not presented. To see an example of similar modeling results that include all parameters, refer to our previous work (Miranda-Moreno and Nosal 2011).

3.4.1 Hourly Absolute Model Results

The hourly modeling results for the utilitarian and recreational counting stations are presented in **Table 6**. To conserve space, only results for weekday models are shown in this subsection; weekend model results are shown for the daily level in the following subsections. With the exception of the second-order terms in the Ottawa and Green Route models, both the temperature and square of the temperature have a significant effect on hourly cycle counts. To illustrate this effect and the differences across the cities, the elasticity of ridership with respect to temperature is plotted in **Figure 12**. The effect is similar across the Montreal, Vancouver, and Portland models: elasticity increases until reaching a peak value, and stays positive over the full range of temperature values. Elasticity increases linearly with temperature for both Ottawa and the Green Route, as the squared terms for temperature are insignificant. Overall, the utilitarian locations exhibit lower elasticity than the Green Route locations; a 10% increase in temperature from the mean for each dataset results in an increase in ridership of 3.4% and 9.6% for the average utilitarian location and the Green Route locations, respectively.

Again, humidity was excluded from the Vancouver and Portland models due to the strong negative correlation between it and temperature. However, for the Montreal, Ottawa, and Green Route, as was done for temperature, variables for both humidity and the square of humidity were entered into the models. Both variables were significant for Montreal, but only first order term was significant in the Ottawa model, and only the second-order term was in the Green Route model. The elasticity values are plotted over a range of humidity in **Figure 12**. For all three locations, increases in humidity result in decreases in cycling counts at all or nearly all values of humidity. For the Montreal and Green Route models, the magnitude of this effect increases more rapidly at higher humidity. Again, utilitarian locations appear less sensitive; a 10% increase from the average humidity of roughly 65% results in an average decrease in cycling of 6.4% in Ottawa and Montreal, and of 16% at Green Route locations.

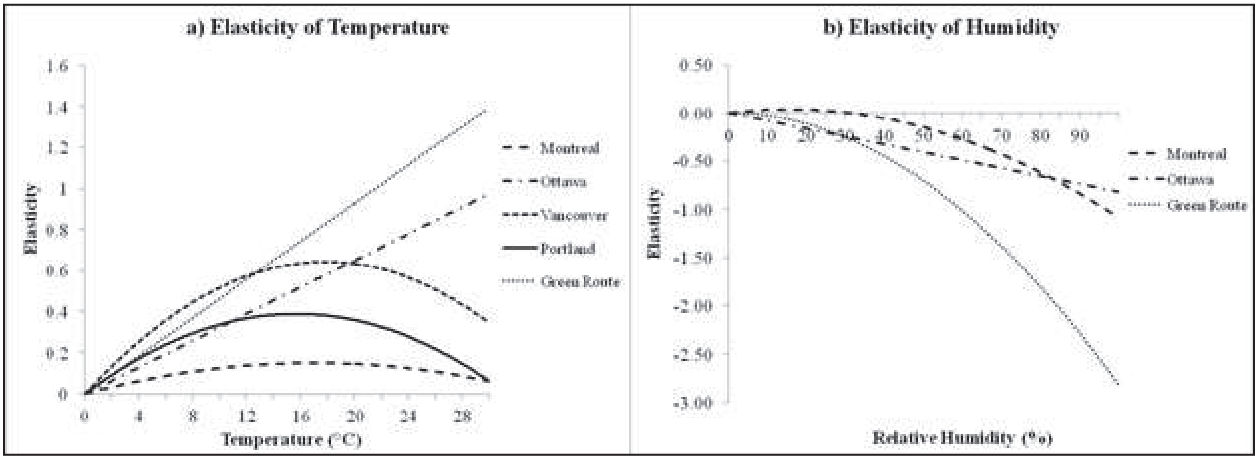


Figure 12.Elasticity of temperature and humidity - hourly models

After testing several different methods, it was determined that the best representation for direct precipitation was a three level factor, with separate binary variables corresponding to low (*rain1*), moderate (*rain2*), and heavy (*rain3*) precipitation. While the intervals for low, moderate and high correspond roughly to the National Weather Service's convention of .25 – 2.5 mm/hour, 2.5 – 7.6 mm/hour, and greater than 7.6 mm/hour, respectively, the intervals were tweaked to provide the best fit for each city. Intuitively, the presence of rain in a given hour has a negative impact on ridership that increases in magnitude with precipitation intensity. With the exception of Vancouver, the recreational locations are more sensitive to direct precipitation than the utilitarian ones.

Three other binary variables were entered into the models to account for lagged effects of precipitation. For a given hour, *RainPrev3Hrs* is set equal to one if it is not raining in the current hour and if it has rained in any of the previous three. It is significant in all models except for Portland's, and it results in a decrease in cycling that is comparable to direct precipitation. *AMrain* is set to one in the hours of 15:00-19:00 if it rained between 05:00-10:00, but did not rain at any other point in the day. It was significant in all of the utilitarian locations except Portland, and exhibited a negative effect on cycling, again comparable to the effect of direct precipitation. This suggests that, even if it does not rain in the afternoon, rain in the morning can result in lower cycling at utilitarian locations due to those who have switched modes or abandoned trips for the day. However, on the Green Route locations, *AMrain* has a positive effect, suggesting that rain in the morning shifts a higher concentration of recreational trips to the afternoon and evening. Finally, *PMrain* is set to one in the hours of 15:00-19:00 if it rains during that period, but at no other point in the day. It is significant in the Montreal and Ottawa models, and has a positive effect on cycling counts, which would serve to counteract the effect of the direct precipitation variables. This suggests that if it rains only in the afternoon or evening, a number of cyclists who would have otherwise changed modes are caught out with no choice but to cycle.

As noted earlier, the hourly models calibrated using OLS produced error series with significant autocorrelation. Incorporating ARMA error structures greatly reduced autocorrelation in the residuals errors. Autoregressive error terms up to the 11^{th} lag, first-order moving average terms, and seasonal autoregressive and moving average terms at the 15^{th} lag were tested for each model, and significant terms were retained (**Table 6**). However, significant correlation persisted at few lags. For example, the autocorrelation plots for the residual errors from both the OLS and regression with ARMA errors are presented in **Figure 13**. Previous research has demonstrated that complex ARIMA models are capable of producing hourly models with no correlation in the residuals errors (Gallop and Tse 2011). However, these models included lagged values of the dependent variable, cycle counts, as well as roughly 40 autoregressive error terms. Explaining variation in bicycle counts using lagged count values limits the practical applications of the model, and including a high number of lagged error terms increases the computational difficulty. Therefore, as an alternative solution to the correlation problem, and to verify the results obtained from the hourly data, daily models were also calibrated, the results of which are presented in the next section.

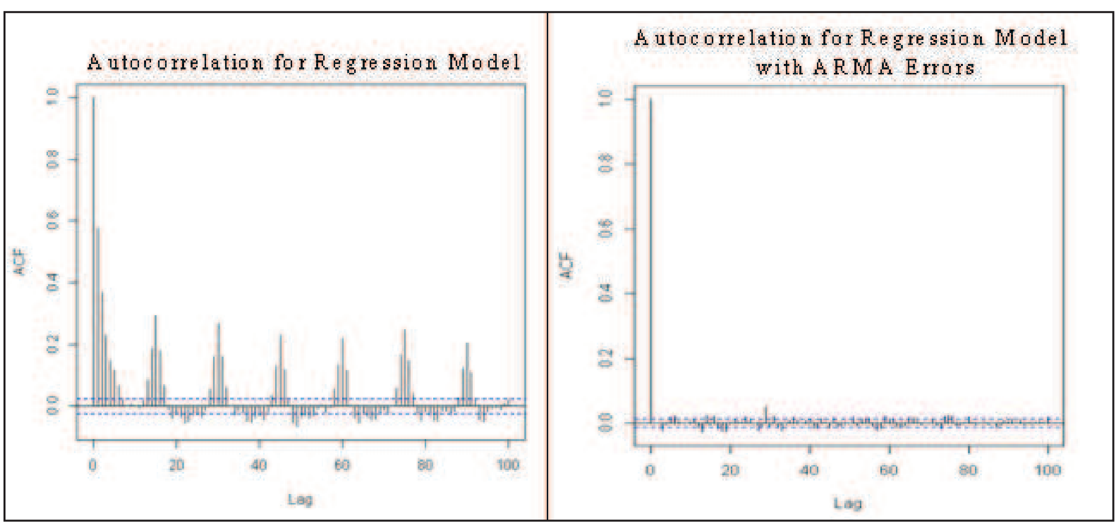


Figure 13. Autocorrelation plots for regression model with and without ARMA errors

Table 6. Hourly Modeling Results – Effects of Temperature, Humidity and Precipitation

	Montreal			Ottawa			Vancouver			Portland			Velo Quebec		
	Coef.	**t-stat**	**Elast.**	**Coef.**	**t-stat**	**Elast.**	**Coef.**	**t-stat**	**Elast.**	**Coef.**	**t-stat**	**Elast.**	**Coef.**	**t-stat**	**Elast.**
$Temp^2$	-0.00026	-6.7	0.12	-0.00027	-1.3*	0.35	-0.0010	-5.0	0.56	-0.00079	-5.1	0.32	0.00048	1.8*	0.96
Temp	0.018	13	---	0.033	4.3	---	0.072	13	---	0.050	8.3	---	0.047	5.1	---
$Humidity^2$	$-7.9*10^{-5}$	-18	-0.39	$-3.3*10^{-5}$	-1.3*	-0.89	---	---	---	---	---	---	$1.4*10^{-4}$	-4.2	-1.61
Humidity	0.0049	8.2	---	-0.0082	-2.2	---	---	---	---	---	---	---	-0.0044	-1.0*	---
Rain1	-0.040	-6.8	-0.039	-0.096	-3.0	-0.092	-0.28	-15	-0.24	-0.034	-3.3	-0.033	-0.19	-3.5	-0.18
Rain2	-0.046	-7.6	-0.045	-0.10	-2.5	-0.10	-0.26	-19	-0.23	-0.042	-2.7	-0.042	-0.24	-4.9	-0.21
Rain3	-0.071	-6.1	-0.068	-0.11	-1.8*	-0.11	-0.38	-15	-0.31	-0.081	-2.4	-0.077	-0.22	-4.6	-0.20
RainPrev3	-0.062	-12	-0.060	-0.12	-3.6	-0.12	-0.23	-14	-0.20	-0.014	-0.78	-0.013	-0.16	-3.7	-0.15
AMrain	-0.034	-3.8	-0.033	-0.090	-1.3*	-0.086	-0.087	-3.3	-0.083	0.0044	0.13*	0.0044	0.17	2.2	0.19
PMrain	0.018	2.3	0.018	0.16	3.3	0.17	0.040	1.5*	0.040	0.017	0.59*	0.017	0.077	1.2*	0.08
AR	1, 8, 9, 11, 15	---	---	1, 4, 15	---	---	1-3, 5-10, 15	---	---	1, 2, 4, 7, 9, 10, 15	---	---	1, 15	---	---
MA	1, 15	---	---	15	---	---	1, 15	---	---	1, 15	---	---	1, 15	---	---
AIC	-25839	---	---	3118	---	---	13608	---	---	-1927	---	---	12335	---	---
obs	20130	---	---	4110	---	---	18225	---	---	2550	---	---	6570	---	---

3.4.2 Daily Absolute Model Results

The coefficients for the weekend and weekday daily absolute models for both utilitarian and recreational locations are presented in **Table 7**. Residual errors from the daily OLS models exhibited far less autocorrelation than those from the hourly models; in general, fewer AR and MA terms were necessary to reduce autocorrelation, and in all cases it was possible to eliminate it fully.

The behavior of the elasticity values of temperature and humidity is similar in the weekday daily models to that of the hourly ones. The maximum daily temperature and the minimum daily humidity were found to provide the best fit in general. Again, the elasticity of temperature on weekdays increases to a maximum value for all locations (**Figure 14a**). However, in the daily models, all of elasticity values decrease more rapidly, crossing the y-axis between roughly 27°C and 30°C. Again, like in the hourly models, the elasticity values of humidity on weekdays are mostly negative (**Figure 14c**). However, Ottawa's elasticity plot is concave-up, crossing the y-axis at roughly 80% humidity. This could be due to the lower amount of data used in the daily models, or due to the fact that only roughly 10% of the minimum humidity values fall above 80% humidity. As with the hourly models, the magnitudes of the elasticity values are generally higher for the Green Route locations, suggesting that recreational locations are more sensitive to weather.

Slight modifications were made to adapt the precipitation variables to the daily models. The three discrete precipitation variables in **Table 7**, *Rain_hrs1 – Rain_hrs3*, relate the duration of rainfall between 06:00 and 20:00 to cycling counts, and correspond to 1 hour, 2-3 hours, and greater than 3 hours, respectively. In Montreal, Ottawa and Vancouver, the effect of rain on weekday cycle counts ranges from -13% to -47%, while the effect in Portland ranges only from -7% to -23%. The Green Route locations are more sensitive, with a reduction as high as 70% due to prolonged rain. In the daily models, *AMrain* is set to 1 if the day's rainfall occurred only between 05:00 and 10:00, and though it is only significant for Montreal and Ottawa, the signs and magnitudes are consistent with the hourly models. Finally, *PMrain* is set to 1 if the day's rainfall occurred only between 15:00 and 19:00, and again, the signs and magnitudes are consistent with the hourly models. *PMrain* in the Green Route models has a large magnitude relative to that of the utilitarian models. This is perhaps because a much larger proportion of cyclists rides in the afternoon at recreational locations and so, should it rain only in the afternoon, a large proportion would be caught out.

In general, cycle counts at utilitarian locations are more sensitive to weather on weekends than on weekday; the magnitudes of the elasticity of cycle counts with respect to temperature and humidity are greater **(Figure 14b and 14d)**. Furthermore, precipitation appears to have a larger negative effect on cycle counts on the weekends than during the week at utilitarian locations (**Table 7**). This is not true in a few cases, such as for *Rain_hrs1* for Montreal, but this may be due to the smaller amount of data available for the weekend models. The *AMrain* and *PMrain* variables are not significant in the weekend models. As for recreational locations, cycle counts exhibit similar elasticity values across weekends and weekdays. This is likely because recreational locations serve similar trip purposes on both weekends and weekdays, while utilitarian locations generally serve markedly different ones: trips for commuting and other obligatory purposes on weekdays, and trips for more recreational or leisurely purposes on weekends.

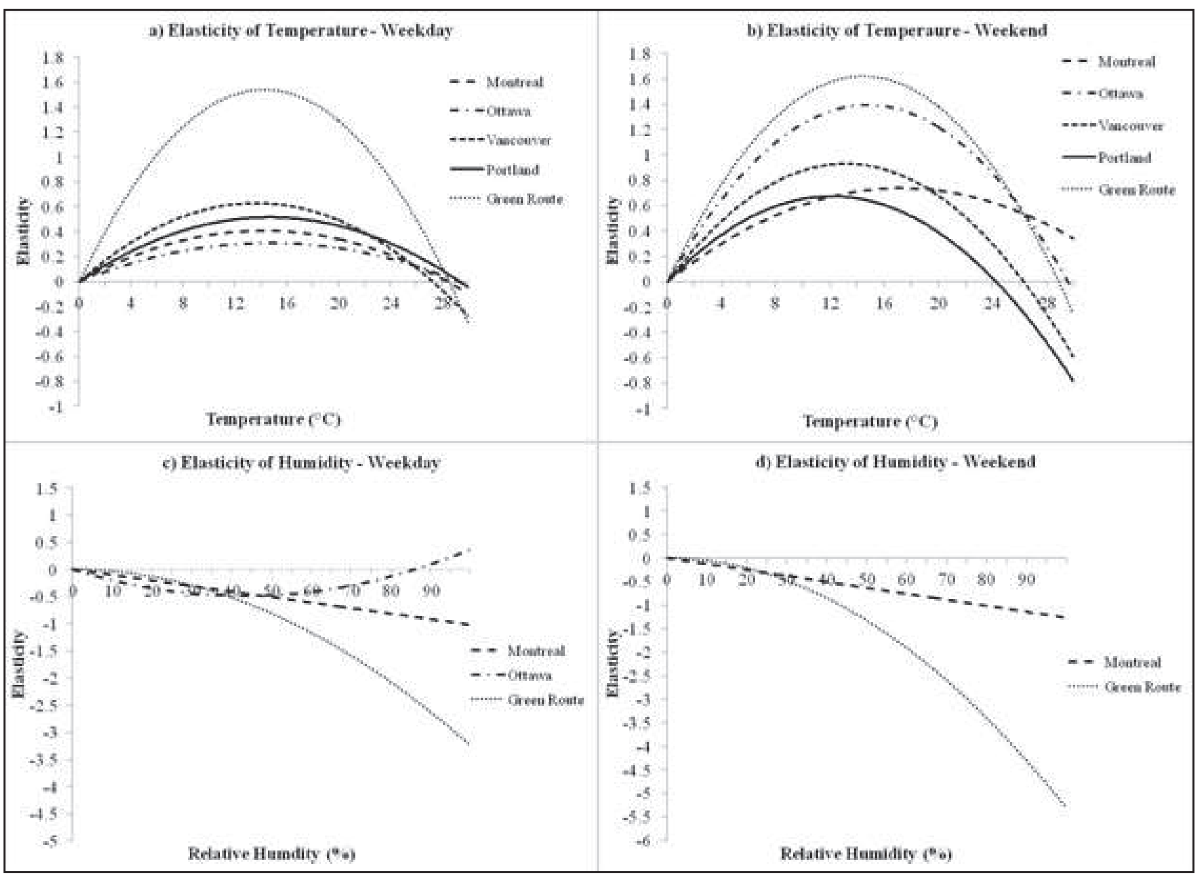


Figure 14. Elasticity of temperature and humidity – daily models

Table 7. Weekday and Weekend Daily-Level Model Results – Effects of Temperature, Humidity and Precipitation

		Montreal			Ottawa			Vancouver			Portland			Green Route		
Weekday		Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.
	$Temp^2$	-0.0010	-10	0.25	-0.00072	-2.5	0.19	-0.0017	-8.5	0.48	-0.0012	-12	0.35	-0.0052	-10.4	0.99
	Temp	0.058	16	---	0.042	4.5	---	0.093	11	---	0.070	14	---	0.28	13.1	---
	$Humidity^2$	0.000021	1.3*	-0.39	0.00013	3.8	-0.40	---	---	---	---	---	---	-0.00026	-3.8	-1.17
	Humidity	-0.010	-5.7	---	-0.023	-5.4	---	---	---	---	---	---	---	0.0063	0.75*	---
	Rain_hrs1	-0.14	-8.1	-0.13	-0.18	-4.2	-0.16	-0.22	-8.9	-0.20	-0.076	-2.3	-0.073	-0.35	-3.4	-0.30
	Rain_hrs2	-0.32	-16	-0.27	-0.33	-6.8	-0.28	-0.37	-12	-0.31	-0.12	-3.7	-0.11	-0.74	-5.9	-0.53
	Rain_hrs3	-0.62	-29	-0.46	-0.54	-11	-0.42	-0.64	-27	-0.47	-0.26	-10	-0.23	-1.2	-12	-0.70
	AMrain	-0.096	-4.8	-0.092	-0.21	-3.9	-0.19	0.014	0.46*	0.014	-0.032	-0.9*	-0.032	0.21	1.7*	0.23
	PMrain	0.15	7.7	0.16	0.071	1.6*	0.073	0.061	2.0	0.063	0.035	1.3*	0.036	0.50	4.7	0.64
	AR	1-3	---	---	1	---	---	1	---	---	2	---	---	1	---	---
	MA	1	---	---	1	---	---	1, 2	---	---	---	---	---	1	---	---
	AIC	-810	---	---	979	---	---	-410	---	---	-266	---	---	609	---	---
	obs	1537	---	---	213	---	---	904	---	---	164	---	---	402	---	---
Weekend		Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.	Coef.	t-stat	Elast.
	$Temp^2$	-0.00124	-5.4	0.59	-0.0032	-4.6	0.89	-0.0027	-5.4	0.66	-0.0023	-4.6	0.21	-0.0047	-6.7	1.1
	Temp	0.085	10.8	---	0.19	6.8	---	0.14	7.6	0.00	0.11	6.0	---	0.26	8.6	---
	$Humidity^2$	0.000040	1.1*	-0.41	-0.00022	-1.5*	-0.88	---	---	---	---	---	---	-0.00016	-2.1	-.61
	Humidity	-0.013	-3.0	---	0.0071	0.61*	---	---	---	---	---	---	---	0.0065	0.47*	---
	Rain_hrs1	-0.13	-3.1	-0.12	-0.31	-2.6	-0.27	-0.19	-3.6	-0.17	-0.40	-4.4	-0.33	-0.45	-2.5	-0.36
	Rain_hrs2	-0.42	-8.6	-0.35	-0.29	-1.8*	-0.25	-0.51	-7.8	-0.40	-0.37	-3.8	-0.31	-0.41	-2.0	-0.34
	Rain_hrs3	-0.71	-14.9	-0.51	-1.0	-9.8	-0.65	-0.82	-19.2	-0.56	-0.54	-6.5	-0.42	-1.2	-5.8	-0.69
	AMrain	0.008	0.14*	0.008	0.11	0.65*	0.11	0.012	0.22*	0.012	0.13	1.2*	0.14	-0.19	-1.1*	-0.17
	PMrain	0.055	1.2*	0.057	0.14	1.0*	0.15	0.11	1.8*	0.12	0.14	1.7*	0.16	-0.098	-0.53*	-0.093
	AR	1,2,6,10,11	---	---	---	---	---	1	---	---	7	---	---	1	---	---
	MA	1	---	---	---	---	---	1	---	---	---	---	---	1	---	---
	AIC	339	---	---	53	---	---	78	---	---	-5.6	---	---	310	---	---
	obs	563	---	---	94	---	---	369	---	---	67	---	---	178	---	---

*not significant at 95% confidence level

3.4.3. Daily Relative Model Results

The coefficients for the weekday and weekend daily relative models are presented in **Table 8**. The differences across cities and between utilitarian and recreational locations are similar to those reported above for the hourly and daily absolute models; therefore, to conserve space, only results for the Montreal locations are presented in this section. Again, data from multiple locations were used to calibrate each model.

In general, as was the case in models reported previously, cyclist counts are more sensitive to weather conditions on the weekend; with the exception of *Reldel_hum_min*, the coefficients for significant variables have greater magnitudes on the weekend. The elasticity of *Reldel_temp_max*, which is the relative deviation in daily maximum temperature from the moving average, is positive. A ten percent increase in daily maximum temperature from the average daily maximum temperature results in 1.8% and 3% increases in cyclist counts above the average on weekdays and on the weekend, respectively. If the daily minimum relative humidity is 10% above average (*Reldel_hum_min*), the corresponding decrease in cyclist counts is roughly 3% from the average on both weekdays and the weekend. The magnitude of the reduction in cyclist counts from the average increases with the number of hours of rainfall in a day. On weekdays, *AMrain,* which is rain that occurs only in the morning and not during the rest of the day, further reduces the expected cyclist counts, and *PMrain*, which is rain that occurs only in the evening, counteracts the overall effect of rain. Both variables were not significant on the weekend.

Significant auto-correlation was found but was completely eliminated with the incorporation of an AR term at the firt and second lag for the weekend and weekday models, respectively. It is suspected that the second lag was significant for the weekend models because, given the structure of the weekend models, two steps behind a Saturday is the previous Saturday, for instance.

Table 8. Weekday and Weekend Relative Model Results for Montreal Counter Locations

	Weekday			Weekend		
	Coef.	**t-stat**	**Elast.**	**Coef.**	**t-stat**	**Elast.**
Reldel_temp_max	0.18	10	0.018	0.30	11	0.030
Reldel_hum_min	-0.32	-21	-0.032	-0.31	-9.2	-0.031
Rain_hrs1	-0.053	-3.1	-0.053	-0.032	-0.83*	-0.032
Rain_hrs2	-0.20	-11	-0.20	-0.28	-6.4	-0.28
Rain_hrs3	-0.31	-16	-0.31	-0.39	-12	-0.39
AMrain	-0.077	-3.9	-0.077	0.051	1.1*	0.051
PMrain	0.053	3.1	0.053	0.031	0.84*	0.031
AR	1	---	---	2	---	---
MA	---	---	---	---	---	---
AIC	-1392	---	---	-210	---	---
obs	1636	---	---	649	---	---

*not significant at 95% confidence level

3.5 DISCUSSION

This work confirms that temperature, humidity, and precipitation, in addition to temporal and location fixed effects, can be used to model hourly and daily bicycle counts. Results obtained at the two levels of data aggregation are consistent. Though two previous studies have found no or little evidence to support a non-linear effect of temperature (Rose et al. 2011; Tin Tin et al. 2012), it was found here that both temperature and the square of temperature were generally significant. In nearly all cases, this resulted in a positive relationship between cycling and temperature, the magnitude of which declines at higher temperatures. Humidity was also found to have a significant, non-linear effect on cycle counts for some locations. However, increases in humidity generally result in decreases in cycle counts, and the non-linear effect was less found less consistently than it was for temperature. Previously, only Gallop and Tse (2011) found a significant relationship between cycle counts and humidity.

Precipitation was found to have a negative effect on cycle counts that increases in magnitude with precipitation intensity. This is consistent with all prior literature. However, it was further demonstrated in this work that greater specificity with regards to when rainfall occurred over the course of the day can have a significant effect in models. If rainfall only occurs in the morning or evening, as opposed to throughout the course of day, cycle counts in a given hour or day can be affected differently. Furthermore, cycle counts in a given hour can be affected by rainfall that occurred in a previous hour, as was reported by Gallop and Tse (2012).

Utilitarian locations were shown to be less sensitive to weather conditions than recreational locations, as has been reported in prior research (Brandenburg 2007; Richardson 2000; Thomas et al; 2013). Furthermore, this research showed that not only do the magnitudes of precipitation variables differ between the two groups, but the dynamics of the effect can differ as well. For instance, rain in just the morning appears to have a negative effect on cycling in the afternoon at utilitarian locations, presumably because commuter cyclists have already switched modes for the day. However, rain in just the morning can increase cycling in the afternoon on recreational locations, presumably because cyclists delay their departure or exercise until later in the day. Recreational/leisure trips are less constrained by time.

For utilitarian locations, cycle counts on weekends were shown to be more sensitive to weather than cycle counts on weekdays, presumably because trips are more recreational or leisurely in nature on the weekend. This is consistent with the findings of Thomas et al. (2013).

This work confirms that the sensitivity of bicycle flows to weather conditions should be taken into account when collecting bicycle count data to estimate AADB. This is particularly true when data are collected over relatively short periods, such as for manual counts or brief automatic counter installations. The quality of these models and those presented by others suggests that methods may be developed to model weather and account for weather-related variation when estimating AADB.

In addition to the log-linear models, count data regression models were attempted with and without serial correlation. Time-series models for count data using more complex estimation methods (such as full Bayes and copula modeling) are promising; however, they use of more complex methods is beyond the scope of this paper.

3.6 CONCLUSION

This chapter utilized a rich hourly and daily bicycle count dataset, comprised of data from 10 counting stations, to investigate the effect of weather across bicycle facilities with different temporal patterns in four North American cities – Montreal, Portland, Ottawa and Vancouver - and on a recreational network across Quebec. This represents a considerable expansion of the evidence on the effects of weather on cycling counts in North America. Results were generally in accordance with prior research. However, this work observed non-linear effects of temperature and humidity while prior research was inconclusive or failed to identify such. Furthermore, this work identified the effects of more nuanced representations of precipitation, such as rain in the morning or afternoon, or rain in previous hours. Finally this work confirms limited research regarding the differences between weather's impact on weekday and weekend cycling.

Significant auto-correlation was identified in hourly cyclist count models, and was addressed using regression models with ARMA errors. Correlation in hourly models is a more serious problem than daily models, which had considerably less auto-correlation. However, consistent results were obtained in both models. The agreement between the hourly and daily models should serve as some validation for the hourly models.

The results of this research help with the understanding of how the effect of weather varies across cities and types of bicycle facilities. This work highlights and confirms the importance of understanding the characteristics of individual bicycle facilities during analysis, and will help to form methodologies to correct bicycle volumes and count data for weather. Relative models that were presented by Miranda-Moreno and Nosal (2011) were improved by utilizing a 21 day moving average as opposed to calculating averages by month. These models are well suited to adjusting short-term counts, as will be demonstrated in **Chapter 4**.

CHAPTER 4:

AADB ESTIMATION METHODS AND ACCURACY

4.1. INTRODUCTION

As noted in **Chapter 1** and highlighted in **Chapter 3**, weather has a significant impact on bicycle counts, and should be taken into account when collecting and adjusting short-term count to estimate AADB. This paper proposes two alternative AADB estimation methods that are designed to account for both weather-related bias and temporal variation. The first method utilizes a relative model, adapted from the model presented in **Chapter 3**, which relates deviations in daily cyclist counts from average daily counts to corresponding deviations in weather conditions. The second method is a disaggregated factor method, based on the individual daily cyclist totals from long-term counting sites.

The performance (accuracy) of the proposed methods, with respect to more traditional ones, is evaluated using data from a set of long-term counting sites in two Canadian cities, Montreal, Quebec and Ottawa, Ontario. The evaluation includes exploring how the location of the short-term count site, weather, time of the count, and duration of the count affect estimated AADB accuracy.

4.2. LITERATURE REVIEW

The most common incarnations of expansion factors today applied to bicycle data have been long-used to estimate annualized traffic for motor vehicles. Manuals like the Federal Highway Administration's Traffic Monitoring Guide (FHA 2001) and the Road Safety Manual (PIARC 2003) recommend that short-term counts be extrapolated by applying a daily and monthly expansion factor to a short-term count. The daily and monthly factors are typically equivalent to the average annual daily bicyclist count for a given day of the week or month, respectively, divided by the overall AADB. If the short-term count is less than 24-hours, an hourly expansion factor is required as well.

Three recent papers have provided the most thorough, if not only, in-depth analyses of the error associated with estimating AADB. Nordback et al. (2013) used a set of counters in Boulder, CO to test the standard expansion factor method. Focused primarily on the effect of the duration of the short-term count, they determined that at least one week of counts is optimal, and that estimates based on just one, two or three hours of data had average absolute error up to 58%. They also concluded that short-term data collected in the warmer months produced lower average error due to lower variability of daily counts. Esawey et al (2013) tested several different expansion factor methods, using data from Vancouver, British Colombia to estimate monthly average daily bicyclists. Rather than utilize the traditional method of producing daily factors by averaging over the course of the year, they produced daily factors for each month individually. They concluded that weekdays provided lower average estimation errors, and recommended against transferring expansion factors across years. They also accounted for weather by producing separate sets of expansion factors for wet and dry weather, finding that this method produced the lowest estimation error. Finally, Figliozzi et al. 2014 applied the expansion factor method to a counter in Portland, OR, and used regression analysis to model the AADB estimation errors as a function of weather conditions and day characteristics. They found that collecting short-term data on days with extreme weather and on holidays decreased the accuracy of the AADB estimates. They also demonstrated that using the regression equation to estimate the AADB estimation error and applying that error to the AADB estimate could improve accuracy. Both Figliozzi et al. (2014) and Nordback et al. (2013) demonstrated that using multiple days to estimate AADB improved accuracy.

4.3 METHODOLOGY

This section introduces the steps that were followed to evaluate the four proposed AADB estimation methods. The four methods are based on the scenario in which the traffic analyst has at least one site with one year or more of daily cyclist count data, and that she or he has one or more sites with at least one 24-hour short-term count (taken within the same year as the long-term data). The analyst would like to use the long-term daily count data to estimate AADB at the sites which have short-term counts. The short-term counts can come from manual data collection methods or temporary sensor installations, such as pneumatic tubes or infra-red sensors. In theory, the short-term count could be as brief as one hour and adjusted to reflect a 24-hour total. However, to simplify the scope of this paper, only methods beginning with a full 24-hour count were considered. To see a more thorough examination of methods based upon counts shorter than a full day, see Nordback et al (2013).

The four AADB estimation methods that were evaluated in this analysis are described briefly below:

- *Traditional Method:* expansion factors for each month and day of the week are computed over a whole year of data
- *Day by Month Method:* expansion factors for each day of the week are computed for each month separately
- *Weather Model Method:* a model that relates deviations from average cyclist counts to deviations from average weather conditions is used to adjust short-term counts
- *Disaggregate Factor Method*: an expansion factor is computed for each day of the year using the raw daily counts and the annual daily average.

Simulating the scenario described in the first paragraph of this section consisted of the following steps:

1. Long-term automatic counting stations in both Montreal and Ottawa were split into those that would represent long-term count sites and short-term count sites, dubbed throughout the rest of this text as long-term test sites and short-term test sites, respectively. Of eight total stations in Montreal, one served as a long-term test site; of five stations in Ottawa, two served as long-term test sites.

2. The long-term test sites were used to develop the frameworks for each of the four AADB estimation methods.
3. For each short-term test site, the four estimation methods were applied in turn to each individual day of count data to estimate AADB. AADB was estimated separately for each year of available data at a given short-term test site.
4. The estimated AADB values were compared to the observed AADB values to evaluate and compare the accuracy of the four estimation methods.

An overview of the study locations and data is presented first. Next, the development of the weather model, which will be used in one of the AADB estimation methods is discussed, followed by a detailed description of each of the four AADB estimation methods. Finally, the manner in which the accuracy of the estimation methods will be compared is discussed.

4.3.1. Study Locations

The locations of the long-term and short-term test sites are shown in the maps in **Figure 15**, and the nearest intersection, along with other brief summary information, is provided in **Table 9**. With the exception of M_S4, which is located on a grade-separated cycling facility, all of the counter locations in Montreal are on on-street cycling facilities. With the exception of M_S6, which is on a unidirectional bike lane, all are bidirectional and physically separated from traffic. In Ottawa, the Laurier counter locations are on paired unidirectional, on-street bike lanes and the Rideau Canal counters are on grade-separated, bidirectional pathways.

Montreal has seven short-term test sites and one long-term test site, which was selected because it had the longest contiguous period of reliable data. Ottawa has two short-term test sites on Laurier, which are associated with a long-term test site also on Laurier (all in red), and one short-term test site on the Rideau Canal path, paired with a long-term test site along the same path (in black) (**Table 9; Figure 15)**. Separate long-term tests sites were used because the difference between Laurier Avenue and the Rideau Canal path is stark, and it was assumed that most practitioners would act accordingly. Regardless, a detailed look at the consequences of long-term site selection is presented later on.

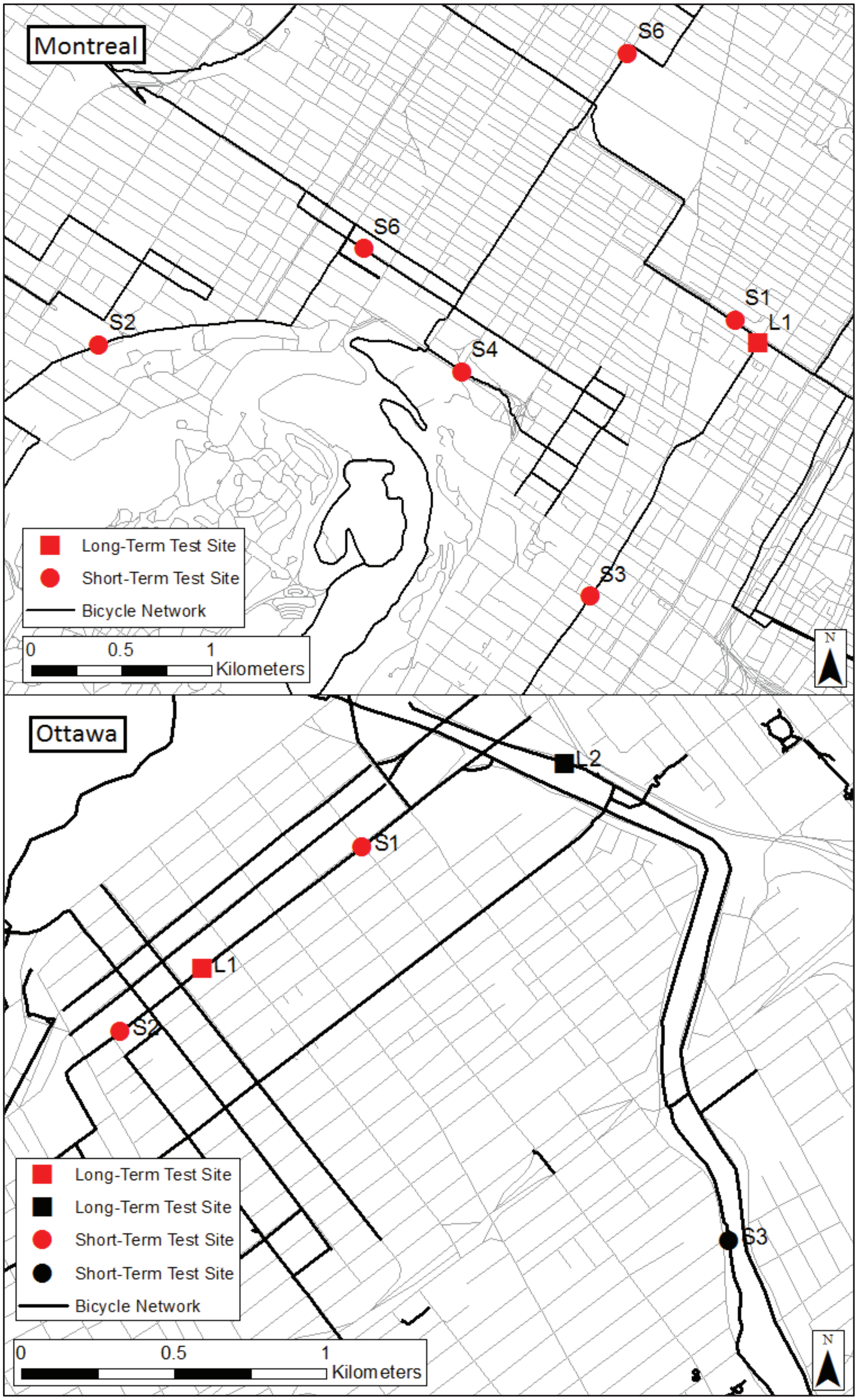


Figure 15. Montreal and Ottawa bicycle counter locations

4.3.2. Data

Bicycle Data

All of the bicycle data used in this study was obtained in the manner described in **Chapter 2**.

Weather Data

Weather data was obtained in the manner described in **Chapter 3**.

Data Processing

Because some of the bicycle counters are located on facilities that are not maintained in the winter, their count data becomes unreliable in the colder months. Therefore, to be consistent across as study locations, data from December through March were excluded from the analysis. The AADB values utilized throughout this study effectively average seasonal daily values. This is however still a useful metric for bicycle studies, and the methods presented here could easily be extended to full years if data are available. Furthermore, holidays were removed, resulting in a loss of roughly 2% of the data. The irregularity of traffic on holidays makes them difficult to include in the calibration and application of the weather models, and it was decided that they could be removed without significantly affecting AADB estimates. Finally, the datasets were combed thoroughly to identify missing data, which can be caused by counter malfunction, construction detours, and so on. This resulted in the loss of another 2.7% and 4% of daily observations in Montreal and Ottawa, respectively.

If more than a few days were missing over the course of a season, the entire year was discarded, as the observed AADB could not reliably be obtained. Because the number of missing days in each season was small, missing data were not estimated. The available years of data for each site are provided in **Table 9**. Again, for each short-term site, each year was treated separately, resulting in 19 test years. AADB was estimated for each year four ways, resulting in 76 different estimates. For reference, the observed AADB values computed over each site's full dataset is also provided in **Table 9**.

Table 9. Short-Term and Long-Term Test Sites

Type	Name*	Location	Years with Data	AADB
Long-Term	M_L1	Maisonneuve at Berri	2008-2012	4429
Short-Term	M_S1	Berri at Maisonneuve	2008 - 2010, 2012	3390
Short-Term	M_S2	Cote St. Catherine at Mceachran	2012	1662
Short-Term	M_S3	Maisonneuve at Peel	2008, 2010-2012	2176
Short-Term	M_S4	Parc at Duluth	2011, 2012	2420
Short-Term	M_S5	Rachel at Papineau	2012	3838
Short-Term	M_S6	St. Urbain at Mt. Royal	2008 - 2010	1917
Long-Term	O_L1	Laurier at Lyon	2012	1015
Short-Term	O_S1	Laurier at Metcalfe	2012	1437
Short-Term	O_S2	Laurier at Bay	2012	418
Long-Term	O_L2	Rideau Canal Western Pathway at First	2012	1210
Short-Term	O_S3	Rideau Canal Eastern Pathway by Laurier	2012	1181

*In each name, L and S correspond to long-term and short-term test sites, respectively

4.3.3. Weather Model Formulation

As noted earlier, in addition to temporal factors, weather can have a significant effect on bicycle traffic volumes, which can in turn have a large effect on AADB estimates. In an attempt to account for that effect, a relative model was developed which relates deviations in daily cyclist totals from the average daily total to respective deviations in daily weather conditions from average conditions. If a researcher knew that weather conditions on the day of a given short-term count were better or worse for cycling than average, she or he could use this model to adjust their short term count accordingly. This model will be incorporated into one of the tested AADB estimation methods presented in the following subsection. The model can be represented as follows:

$\Delta DB_{y,j} = \left(\beta * \Delta W_{y,j}\right) + \left(\alpha * W_{y,j}\right) + \left(\gamma * FE_{y,j}\right) + \varepsilon_{y,j}$, where **(Equation 9)**

$\Delta DB_{y,j}$ $= \left[DV_{y,j} - \frac{\sum_{k=j-10}^{k=j+10} DB_k}{21}\right] \Big/ \left[\frac{\sum_{k=j-10}^{k=j+10} DB_k}{21}\right]$, the relative Daily Bicyclists deviation of day *j* in year *y*, from a 21 day moving average of Daily Bicyclist totals, where *j* ranges from 1 to the number of days in the year or cycling season,

$\Delta W_{y,j}$ = $\left[W_{y,j} - \frac{\sum_{k=j-10}^{k=j+10} W_k}{21}\right] \Big/ \left[\frac{\sum_{k=j-10}^{k=j+10} W_k}{21}\right]$, a vector of deviations in continuous weather variables (temperature, dewpoint, total precipitation, etc.) from their respective 21 day moving averages on day j in year y. (Note that although the normalized version is shown above, these variables may or may not be normalized)

$W_{y,j}$ = a vector of continuous weather conditions on day j in year y,

$D_{y,j}$ = a vector of binary variables related to temporal effects, like the day of week on which day j falls,

β, α, γ = vectors of coefficients to be estimated from the data, and

$\varepsilon_{y,j}$ = a random, independent error term for day j in year y.

Linear regression was used to calibrate the model coefficients. Multi-collinearity was checked to ensure that variables with correlation coefficients with absolute values greater than 0.5 were not included in the same model.

This model differs slightly from the one presented in **Chapter 3**. Rather than calculate separate averages for weekdays and weekends within the 21-day moving period, the average across the entire 21 days was used. However, because the response of cycle counts to weather conditions varies between weekdays and weekends, the model coefficients were calibrated using only weekdays. Therefore, since the average values were computed using all days, the model relates deviations in weekday cycle counts from the overall average to deviations in weather conditions. This makes it easier to relate short-term counts taken on weekdays to the overall AADB, which includes weekends.

4.3.4. AADB Estimation Methods

The first two methods are based on those described in the Federal Highway Administration's Traffic Monitoring Guide (FHA 2001), and account for temporal variation only. The third and fourth methods are similar, but attempt to control for both temporal and weather-related variation. Note that although the AADB values reflect bicyclist counts on all days, only weekdays were used to estimate AADB.

Traditional Method

This method accounts for daily and seasonal variation in traffic volumes with individual factors for each day of the week, averaged over the whole season or year, and for each month. It has been widely used to annualize both motor vehicle and bicycle and pedestrian traffic counts.

$\widehat{AADB}_{i,y,j} = SDB_{i,yj,d,m} * {}^{1}/_{DF_d} * {}^{1}/_{MF_m}$, where **(Equation 10)**

$\widehat{AADB}_{i,j}$ = the estimated AADB for short-term site i, and year *y,* based on the short-term count taken on day *j,* which ranges from 1 to the number of days in the cycling season or year,

$SDB_{i,yj,d,m}$ = the observed Short-term Daily Bicyclists at short-term site i on day j in year *y,* which falls on day of the week d in month *m,*

DF_d = the Day-of-the-week Factor for day of the week *d*.

MF_m = the Month Factor for month *m.*

Both DF_d and MF_m are calculated using data from a long-term test site. In this case, the DF_d is the ratio of the average daily total cyclists on a given day of the week, *d,* averaged over the entire season or year, divided by the overall AADB. MF_m is the ratio of the average daily total cyclists in month *m,* divided by the overall AADB. Both DF_d and MF_m were calculated separately for each year.

Day by Month Method

This method is similar to the traditional method, but rather than account for daily and seasonal variation with separate factors, they are accounted for by computing the DF_d separately for each month. For instance, for an 8 month cycling season there would be 56 total factors.

$\widehat{AADB}_{i,y,j} = SDB_{i,y,j,d,m} * {}^{1}/_{DF_{d,m}}$, where **(Equation 11)**

$\widehat{AADB}_{i,y,j}$ = the estimated AADB for short-term site i, and year *y,* based on the short-term count taken on day *j,* which ranges from 1 to the number of days in the cycling season or year,

$SDB_{i,y,j,d,m}$ = the observed Short-term Daily Bicyclists at short-term site i on day j in year *y,*, which falls on day of the week d in month *m,*

$DF_{d,m}$ = the Day-by-month Factor for day of the week d and month *m.*

The $DF_{d,m}$, again calculated using data from a long-term test site, is the average daily total cyclists for each day of the week, *d,* within each month, m, divided by the overall AADB. $DF_{d,m}$ was calculated separately for each year.

Weather Model Method

This method attempts to account for the effect of weather on daily cyclist counts and subsequent AADB estimations by using the expected cyclist count deviation, obtained from the model described in **Section 4.3.3**, to adjust the observed short-term count. The method is executed in two steps: first, the short-term count is adjusted based on the predicted deviation from the 21-day moving average due to weather; second, the weather-adjusted count is temporally adjusted to reflect how the 21-day average varies from the AADB. The first step can be summarized as follows:

$\widehat{MADB}_{i,y,j} = SDB_{i,y,j,d,m}/(1 + \widehat{\Delta DB}_j)$, where **(Equation 12)**

$\widehat{MADB}_{i,y,j}$ = the estimated Moving Average Daily Bicyclists for short-term site *i*, and year *y,* centered at day *j*, which ranges from 1 to the number of days in the cycling season or year,

$SDB_{i,y,j,d,m}$ = the observed Short-term Daily Bicyclists at short-term site *i* in year *y,* on day j, which falls on day of the week *d* in month *m.*

$\widehat{\Delta DB}_{y,j}$ = the expected deviation in daily bicyclists on day *j* in year *y,* based on the weather conditions on day *j* and obtained from **Equation 9**, after calibrating the model with bicycle data from a long-term site.

For example, if the weather on day *j* was particularly well-suited to cycling, $\widehat{\Delta DB}_j$ will be positive, and the short-term daily bicyclists count, $SDB_{i,d,m}$, will be adjusted downward. The second step can be represented as follows:

$\widehat{AADB}_{i,y,j} = \widehat{MADB}_{i,y,j} * {}^{1}/_{MAF_j}$, where **(Equation 13)**

$\widehat{AADB}_{i,y,j}$ = the estimated AADB for short-term site *i* and year *y*, based on the short-term count taken on day *j,* which ranges from 1 to the number of days in the cycling season or year,

$\widehat{MADB}_{i,y,j}$ = the estimated Moving Average Daily Bicyclists for short-term site *i* and year *y*, centered at day *j*, as estimated using **Equation 12**.

MAF_j = $\frac{\sum_{k=j-10}^{k=j+10} DB_k}{21} / AADB$, the Moving Average Factor, centered at day *j* and calculated using data from a long-term site.

The coefficients of the weather model were estimated using data from the long-term sites. For Montreal, all 5 years of data were used to calibrate one model. If a contiguous section of data was missing, then a section spanning from ten days before to ten days after the missing data was excluded. If a single day was missing, then only twenty days were used to calculate the moving average, when applicable.

Disaggregate Factor Method

The disaggregate factor method is perhaps the simplest. For a long-term test site, each daily bicyclist total is divided by the overall AADB. Essentially, an expansion factor is created for each day of the year. It is expected that, as long as the long-term and short-term test sites experience the same weather, this method will account for deviations in weather conditions, and temporal factors like day of the week and month. It can be represented as follows:

$\widehat{AADB}_{i,y,j} = SDB_{i,y,j} * {}^{1}/_{DF_{y,j}}$, where **(Equation 14)**

$\widehat{AADB}_{i,y,j}$ = the estimated AADB for short-term site *i* and year *y*, based on the short-term count taken on day *j,* which ranges from 1 to the number of days in the cycling season or year,

$SDB_{i,y,j}$ = the observed Short-term Daily Bicyclists at short-term site *i,* on day *j* in year *y*,

$DF_{y,j}$ = $DB_{y,j}/AADB_y$, the Disaggregate Factor for day *j* in year *y*, where $DB_{y,j}$ and

$AADB_y$ are the total cyclists on day j in year y and the AADB, respectively, for the long-term count site. Again, j ranges from 1 to the number of days in the cycling season.

4.3.5. Evaluation of Accuracy

Each day of available count data was used to estimate AADB for a given short-term test site and year. Therefore, for each AADB estimation method, each day's estimate was compared to the observed AADB using the absolute percent error:

$$\left|Error_{i,y,j}\right| = 100\% * \left|\widehat{AADB}_{i,y,j} - AADB_{i,y}\right| / AADB_{i,y}, \text{ where} \quad \textbf{(Equation 15)}$$

$\left|Error_{i,y,j}\right|$ = absolute percent error for short-term site *i,* based on the AADB estimated on day *j* in year *y,* and calculated for each estimation method,

$\widehat{AADB}_{i,y,j}$ = the estimated AADB for short-term site *i* and year *y*, based on the short-term count taken on day *j,* which ranges from 1 to the number of days in the cycling season or year,

$AADB_{i,y}$ = the observed AADB for site *i* and year *y.*

In the results section, unless otherwise noted, the average absolute percent errors (AAPE) (averaged across all sites and years) are used to compare the accuracy of the different methods.

4.4. RESULTS AND DISCUSSION

The results of the weather model calibration are first discussed briefly, followed by the results and discussion regarding the different AADB estimation methods.

4.4.1. Weather Model

The coefficients of the weather model are presented in **Table 10**, along with corresponding p-values and a description of each variable. All of the results related to the signs and magnitudes of the estimated coefficients are in accordance with previous research.

It was found that positive deviations in temperature from the average were statistically significantly associated with increases in cyclist counts. However, this effect is tempered when the temperature is above twenty and deviations from the average temperature were positive; when it is already hot, increases in temperature make cycling less appealing.

For incorporating the effects of humidity on cycling, it was found that the relative deviation in maximum daily dew point depression explained a greater amount of the variance than relative humidity. Dew point depression is the difference between the air temperature and the dew point temperature, the temperature at which water vapor will condense into a liquid. The larger the dewpoint depression, the less humid the air feels. Increases in dewpoint depression from the average were associated with increases in cyclist counts.

Precipitation was entered into the model as continuous variable. Though precipitation decreases cyclist counts, a non-linear effect was observed: the magnitude of its negative effect increases less rapidly at higher levels of precipitation. To the average cyclist, the difference between no rain and light rain is greater than the difference between moderate rain and heavy rain.

In addition to the weather-related variables, fixed effects for Tuesday, Wednesday and Thursday were significant, meaning that average ridership on those days varies with respect to Monday. A fixed effect for Friday was found to be insignificant. Finally, a constant was significant and had a positive magnitude. This reflects the fact that the dependent variable in this model is the deviation in daily cyclist counts from the overall average count (calculated using all days of the week), but the model was calibrated using only weekdays. Counts at locations used to calibrate this model are generally higher during the week than on the weekend.

Regression models with ARMA errors were tested, but because the neither the coefficient values nor their significance changed meaningfully, simple OLS models were used in this case.

Table 10. Weather Model Coefficients - Montreal

Category	Variable	Description	Coefficient	P-Value
ΔW_i	del_temp_max	Deviation of maximum daily temperature from average	0.027	0.000
	del_dpd_max	Deviation of maximum dew point depression (temperature minus dew point temperature) from average	0.023	0.000
W_i	temp_max _o20_pdev	Equal to maximum daily temperature when maximum temperature is above 20°C and del_temp_max is positive (equal to 0 otherwise).	-0.0036	0.000
	$(\text{total\_precipitation})^2$	The square of total daily precipitation (in mm^2).	0.000065	0.000
	total_precipitation	Total daily precipitation (in mm).	-0.018	0.000
FE_d	fmon (reference)	---	---	---
	ftue	Equal to 1 if Tuesday, 0 otherwise.	0.12	0.000
	fwed	Equal to 1 if Wednesday, 0 otherwise.	0.11	0.000
	fthu	Equal to 1 if Thursday, 0 otherwise.	0.11	0.000
	constant	---	0.17	0.000
R^2 = 0.61				

4.4.2. AADB Estimation

With the exception of M_S6, the disaggregate method produced the lowest AAPE for all sites, followed by the weather method; the traditional and day by month methods performed comparably, and produced the least accurate estimates with the highest AAPE values (**Figure 16**). The magnitudes of the average absolute errors for the traditional method, which is the most comparable method, are in accordance with those obtained in other analyses (Nordback et al. 2013; Esawey et al 2013; Figliozzi et al. 2014). The rest of the results section is broken up to examine more specific factors that affect AAPE.

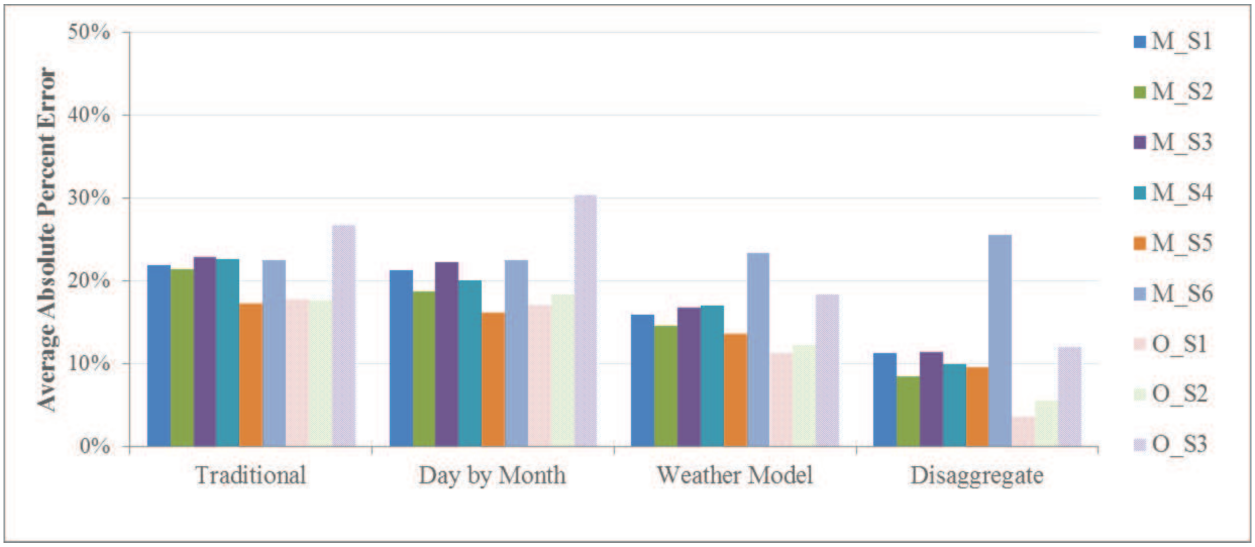


Figure 16. AAPE by Estimation method and Short-Term Count Location

Location of Short-Term Site and Long-Term Site Selection

Although an in-depth analysis of the contextual factors related to each short-term site was beyond the scope of this study, some conclusions can be drawn from a basic examination of the AAPE by short-term site locations. For some sites, such as M_S6, it was not possible to produce a reasonably accurate estimate with any method (**Figure 16**). An examination of the average daily cyclists by month for the Montreal locations with data available in 2008 reveals why this is so (**Figure 17**). While the monthly traffic profiles for the other two short-term sites with data, M_S1 and M_S3, closely match that of the long-term site (M_L1), M_S6 has a very different ridership pattern. It is not immediately clear why M_S6 exhibits such a markedly different pattern. Counter error is a possible explanation, but there were no known issues with the equipment. Regardless, this highlights the need to determine before estimating AADB whether the long-term and short-term sites have compatible traffic patterns. This in practice can be a difficult task if the analyst is not familiar with the traffic dynamics in the different corridors of the network. How to determine whether this is the case will require much further research.

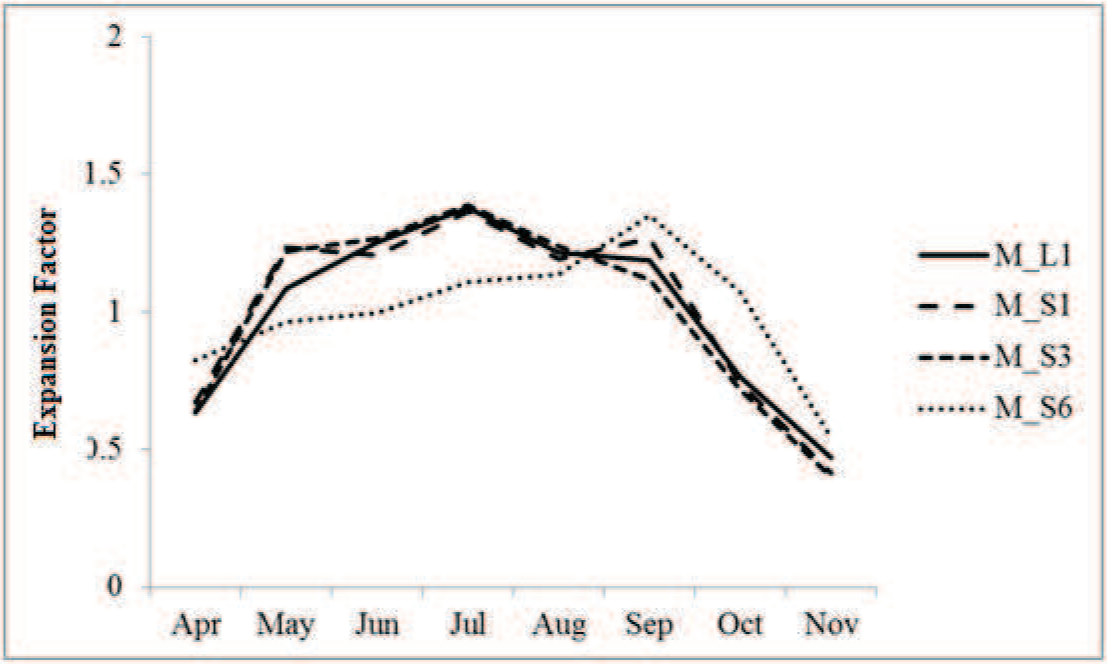


Figure 17. Monthly Expansion Factors at Montreal Locations in 2008

Ottawa presents another example of the effect that long-term site selection can have on the accuracy of the estimated AADB. One of the sites with the lowest AAPE values is O_S1, which is on the same corridor as and within one kilometer of its associated long-term site, O_L1 (**Table 9; Figure 18**). Both O_S1 and O_L1 are *primarily utilitarian,* and the AAPE value obtained when the disaggregate method is used to estimate the AADB of O_S1 with O_L1 is just 6%. However, when the data from the long-term test site on the Rideau Canal, O_L2, was used to estimate AADB of O_S1, the AAPE increased to 33% (**Figure 18**). O_L2 is *mixed-utilitarian*, and a plot of the day-of-week expansion factors reveals how this discrepancy increases the error (**Figure 19)**. The two sites located on Laurier, O_L1 and O_S1, exhibit nearly identical daily distributions, but O_L2 is markedly different – counts are more uniform through the week, rather than being skewed towards workdays. Being grade separated and along the river, it is used more recreationally than the facility on the city street. This again highlights the importance of matching sort-term sites to the appropriate permanent counting stations, in particular when they are not in the same corridor.

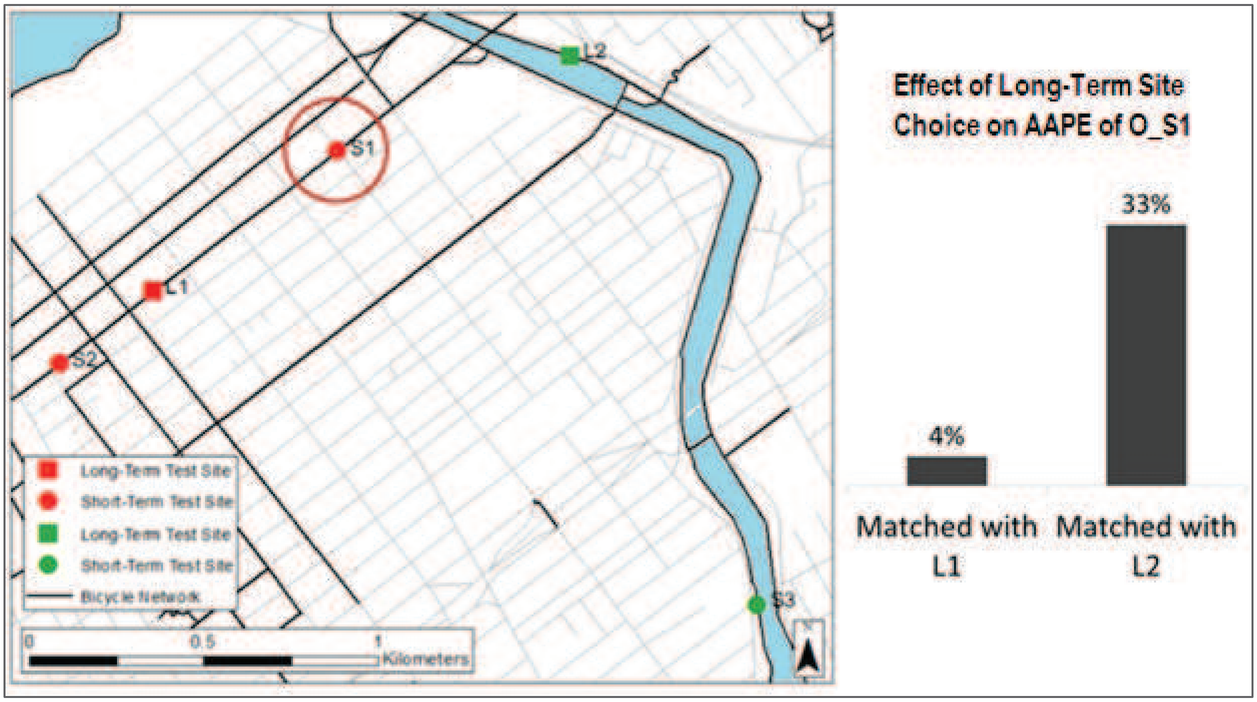


Figure 18. Effect of Long-Term Site Location on AAPE of O_S1

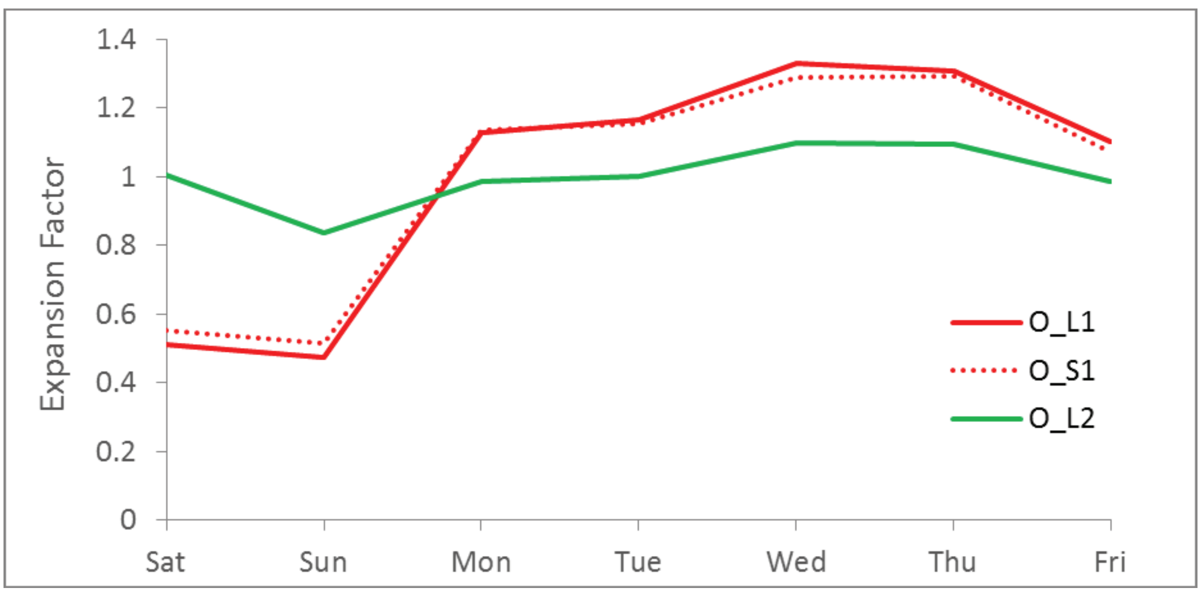


Figure 19. Day of Week Expansion Factors for Select Ottawa Locations

Weather Conditions

For all four estimation methods, when temperatures are warmer, estimates are more accurate (**Figure 20a**). However, the difference between estimates obtained during colder periods and warmer periods is more pronounced for the traditional and day by month methods; the traditional method produces errors that are roughly four times lower when the temperature is above 30 than when it is less than zero.

Estimates are more accurate when short-term counts are performed in the absence of rain (**Figure 20b**). However, for the disaggregate method, the difference in accuracy is less pronounced between dry and wet weather. This makes this method particularly attractive, as it suggests that, for instance, even days on which it rained during a pneumatic tube installation could be reliably used for AADB estimation. Perhaps surprisingly, the weather method produces a relatively large difference in accuracy between wet and dry days. This suggests that further work is needed to accurately model how precipitation affects cyclist counts. This reinforces the recommendation that data collection campaigns for short-term counts ought to take place in good weather conditions and not during winter.

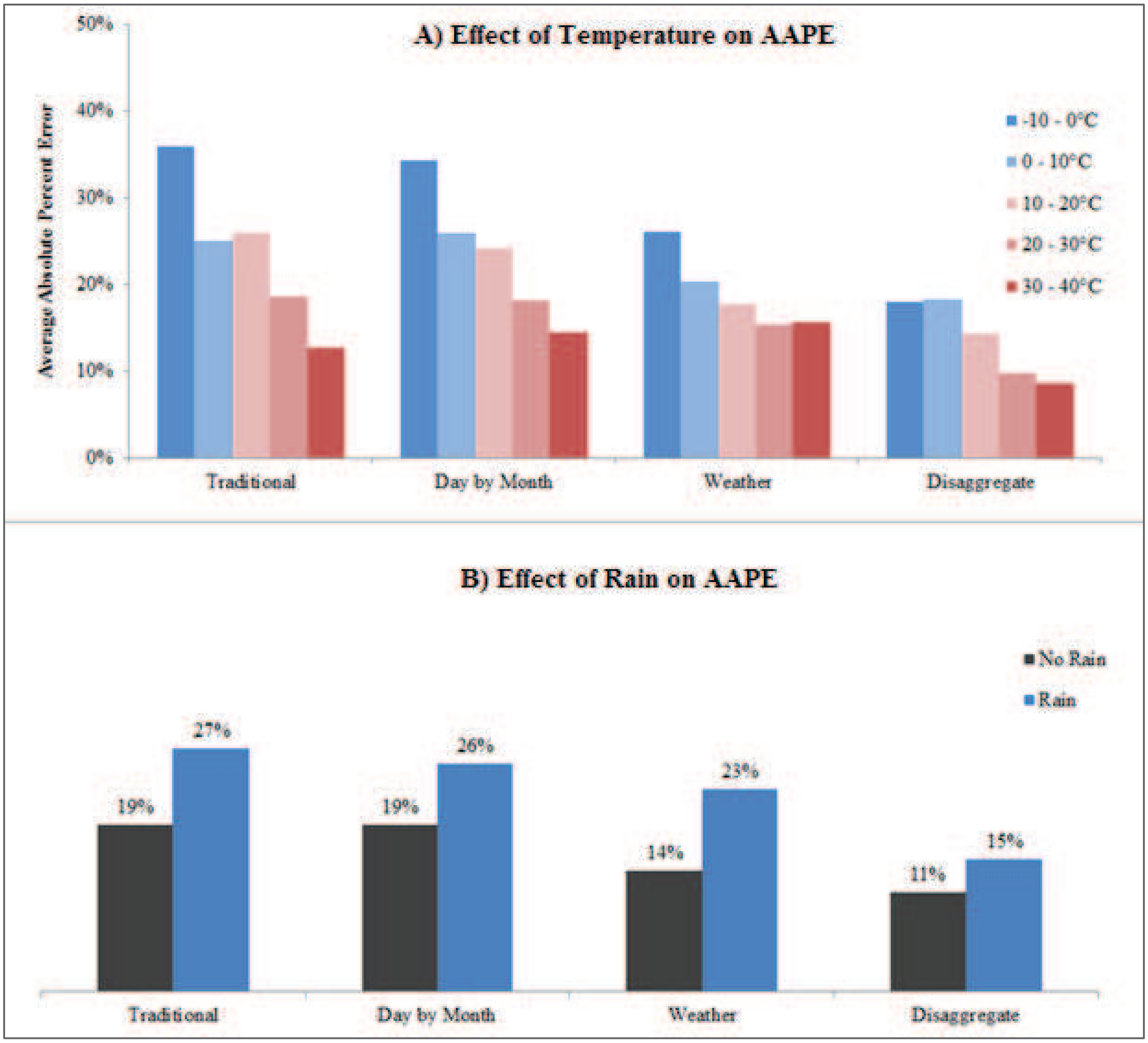


Figure 20. Effect of Weather Conditions on AAPE by AADB Estimation Method

Time of Temporary Count

In general, there is little variation in accuracy across days of the week (**Figure 21a**). For the traditional, day-by-month, and weather methods, it appears that Thursdays may be the best day on which to collect a short-term count. However, this may be specific to this set of counters. Furthermore, as suggested earlier in this suggest, it is clear that more accurate AADB estimates are produced in the warmer months (**Figure 21b**). It appears that, in this case, short-term counts taken in August produce the lowest AAPE. This is in accordance with prior work (4,5) and should serve as a clear guideline for when is best to collect short-term data.

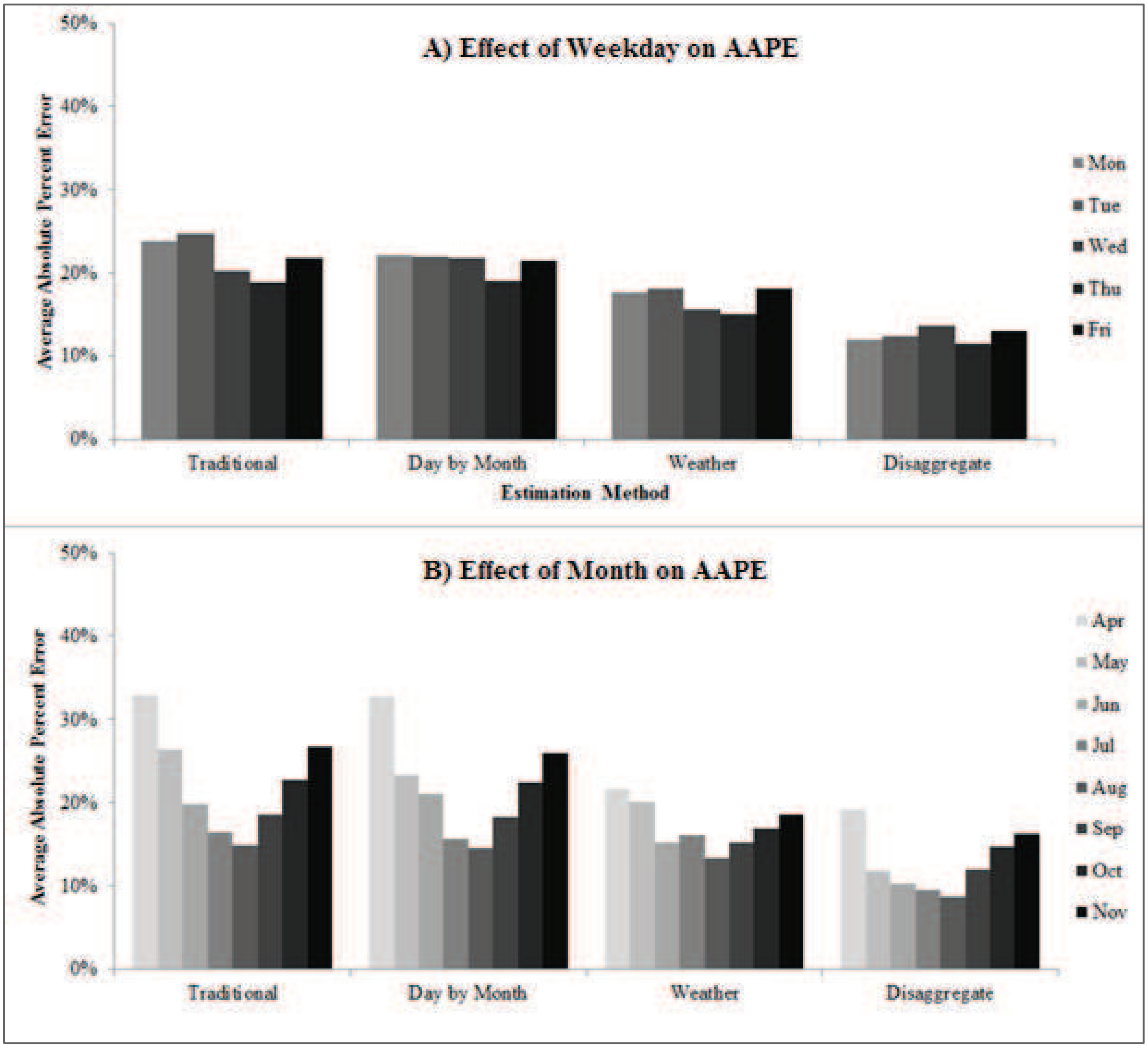


Figure 21. Effect of Time of Short-Term Count on AAPE by AADB Estimation Method

Duration of Count

In order to test the effect of the duration of the short-term count on AAPE, the AADB estimates obtained on contiguous weekdays were averaged, and the number of days was increased from 1 day up to 30. For the first three methods, large gains in accuracy can be obtained by increasing the duration of the short-term count (**Figure 22**); increasing the count duration from one day to five can decrease the average error by roughly one-third. For the traditional and day by month methods, the error associated with a month of counts is roughly half that of a one-day count, and for the weather method, it is roughly 60%. While gains can be made for the disaggregate method, they are less pronounced; after 5 days and 30 days, the AAPE is only 15% and 34% lower than the AAPE associated with a one-day count, respectively.

As more days of short-term data are included, the AAPE values improve in a non-linear manner. For the first three methods and the disaggregate method, roughly 75% and 60%, respectively, of the improvement to be had by adding more data has occurred by the addition of the 10^{th} day. Furthermore, the AAPE values across the four methods converge as more data is added. After 10 days of short-term data collection, the weather method and the disaggregate method produce the same AAPE, and after 20 days, all are essentially the same.

The results presented in **Figure 22** highlight the potential advantage of the disaggregate method. On average, ten days of short-term data collection are necessary for the traditional and day-by-month methods to produce estimates which are comparable in error to one day of the disaggregate method, and 5 days of data collection are required for the weather method to do so.

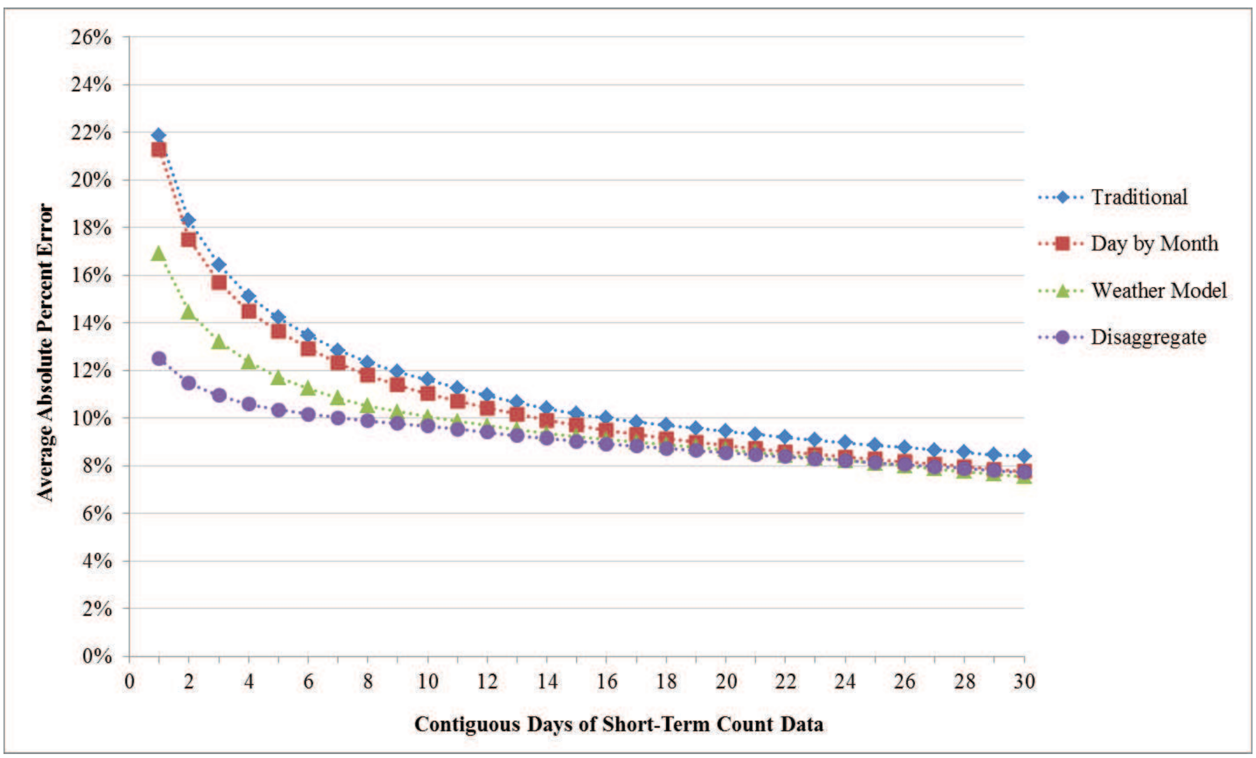


Figure 22. Effect of Duration of Short-Term Count on AAPE by AADB Estimation Method

4.5. CONCLUSION

This chapter evaluates the performance of four methods to estimate AADB from short-term counts, including two novel methods. A set of long-term count sites in two Canadian cities, Montreal and Ottawa, were divided into those that simulated long-term and short-term count data sites. Using data from long-term test sites, the four methods were applied to data from the short-term test sites to estimate AADB. The accuracy of the four methods was evaluated based on the average absolute relative error between the estimated and observed AADB values.

In general, it was found that the disaggregate method performs better than the other three methods, particularly when compared to the traditional and day-by-month methods. The weather adjustment method was the second best option, performing in some cases as well as or better than the disaggregate method. It was observed that the selection of the long-term location is critical; lowest error is obtained when the traffic patterns at the long and short term sites match well. This could be even more important than the selection of the factoring method.

The effect of weather conditions, as well as the time and duration of the short-term count was also evaluated, and it was found that greater accuracy can be obtained by considering these factors when planning a short-term count. Short-term data collected on dry days in warmer periods, particularly in the month of August, produced the lowest error for this set of sites. Collecting data on Thursday also appears to improve accuracy slightly. Furthermore, increasing the number of days of short-term data reduces error considerably, albeit in a non-linear fashion. After around 10 days of data collection, further gains in accuracy are marginal.

The weather and disaggregate methods are advantageous in that they produce more accurate AADB estimates. Because of their ability to account for weather conditions, it appears that less short-term data is needed to obtain accuracy comparable to more traditional methods. However the data needs of the weather method, and the fact that both methods are only applicable to short terms counts collected in the same year as the long-term counts, reduces their utility.

CHAPTER 5:

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

5.1 FINAL COMMENTS

With the role of bicycling in urban transportation expanding, more cities are investing in infrastructure and programs, and more cities are collecting bicycle count data. Bicycling-related research is increasing as well to help understand and enhance the role of bicycling in an urban transportation system. Accurate, non-biased measures of traffic flow are essential for many tasks in the planning, design and operation stages; AADB is a common, important metric with a wide range of research and practical applications. However, conventional methods for estimating AADT were developed largely for use with motor vehicle counts, and they tend to produce large errors when applied to bicycle counts. There are two primary sources of AADB estimation error. First is the failure match short-term data collection sites to appropriate long-term data collection sites with similar temporal patterns. Second is the failure of conventional AADB estimation methods to account for weather conditions experienced during the short-term count.

To address the first source of error, in **Chapter 2** forty-one long-term bicycle counter locations were grouped into four factor groups – *primarily utilitarian, mixed-utilitarian, mixed-recreational,* and *primarily-recreational* – using the k-means clustering technique. These groups confirmed those developed by Miranda-Moreno et al. (2013), and demonstrate that a large proportion of short-term count locations could likely be matched with long-term sites that have similar traffic patterns, improving the quality of AADB estimates. Also presented in **Chapter 2** was a preliminary land-use analysis to examine whether land-use data could be associated with the different factor groups. This analysis demonstrated that engineering judgement and intuition is largely correct when classifying count locations; for instance, *primarily recreational* sites are associated with the highest proportion of open space and park space in the vicinity of the bicycle facility or road section under analysis. However, strong caveats were discovered in that the land-use surrounding certain sites may suggest a particular type of traffic pattern, but a different pattern is observed due to particularities of the site, such as tourism activity or reasons which are not always immediately apparent.

The second source of error, failure to account for weather conditions, was addressed in **Chapter 3**. It was demonstrated that weather conditions – namely temperature, humidity, and precipitation – have a significant effect on hourly and daily bicycle counts. Non-linear effects were observed for temperature and humidity, and it was demonstrated that precipitation can affect the magnitude of bicycle counts throughout the day, beyond the hours in which it actually rained. These effects vary across the different factor groups. For instance, *primarily utilitarian* facilities are generally less sensitive to weather conditions than *primarily recreational* facilities. Furthermore, it was demonstrated that the dynamics of the effect of weather on bicycle counts are consistent across cities, but the magnitude of that effect can vary. This suggests that when accounting for weather, models and applied methods will likely need to be specifically developed for the given city and/or factor group. Finally, in addition to absolute models, a relative model was developed that related deviations in daily bicycle counts from a 21-day moving average to corresponding deviations in weather conditions. Previous models had used averages grouped by month, which produced inferior results. Regression with ARMA errors was tested to account for serial correlation of errors.

Chapter 4 compared the performance of four different AADB estimation methods, including two novel methods that account for weather conditions, using a simulation study format and data from different facilities. It was found that the two methods that accounted for weather conditions during the short-term count produced lower errors on average. One of the proposed methods, the disaggregate method, is capable of producing AADB estimates with very low error, even with relatively short data collection periods. In addition, it was demonstrated that failure to consider the appropriate factor group, or to match a short-term site with an appropriate long-term site, can drastically increase error.

5.2 FUTURE RESEARCH

This work paves the way for considerable future research. The cluster analysis presented in **Chapter 2** will be updated with more bicycle facilities from different regions and more years of data. This should give greater confidence in the derived factor groups and may even result in more nuanced groups. In addition, much more research will be necessary to identify a reliable optimization method for matching short-term sites to appropriate long-term sites. The land-use analysis in this work serves only as a beginning; more work will be necessary to identify proper units of analysis and metrics to determine how land-use and other factors interact and affect short and long-term variability of bicycle ridership.

With regards to **Chapter 3**, future work will include developing further improvements to the relative weather model. This includes finding potentially more appropriate means of incorporating weather data into the model – such as thermal indexes, better representations of precipitation – and calibrating models for a wider array of locations and using more data. From the statistical point of view, more sophisticated models for count data can be attempted to deal with missing counts and serial correlation in the model error terms.

Finally, with regards to the analysis presented in **Chapter 4**, future work will include testing the AADB estimation methods with short-term counts less than a full day long. This will reflect the manner in which much of the bicycle count data is collected. In addition, more locations and more years of data will be included to help increase the understanding of what factors affect the accuracy of AADB estimated. Extensions of the weather model method will be developed, such as the ability to use short-term counts from different years, and the standardization of data from different years for comparative or post-project evaluation studies.

REFERENCES

Allen, D. Patrick, Nagui Rouphail, Joseph Hummer, and Joseph Milazzo II. 1998. "Operational Analysis of Uninterrupted Bicycle Facilities." *Transportation Research Record: Journal of the Transportation Research Board* 1636: 29–36.

Brandenberg, C., Matzarakis, A., Arnberger, A. 2007. Weather and cycling – a first approach to the effects of weather on cycling. Meteorological Applications, 14, pp. 61-67.

Brown, B. Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight and obesity. Vol. 15, No. 4, 2009, pp. 1130–1141.

City of Portland Bureau of Transportation. "2011 Bicycle Counts Report." 2012. Portland, OR.

El-Geneidy, A.; K.J. Krizek; M.J. Iacono. "Predicting Bicycle Travel Speeds Along Different Facilities Using GPS Data: a Proof-of-Concept Model". Washington, DC: Transportation Research Board. 2007

Esawey, M. El, C. Lim, T. Sayed, and A. I. Mosa. Development of Daily Adjustment Factors for Bicycle Traffic. *Journal of Transportation …*, No. August, 2013, pp. 859–871.

Ewing, R., and R. Cervero. Travel and the Built Environment. *Journal of the American Planning Association*, Vol. 76, No. 3, Jun. 2010, pp. 265–294.

Federal Highway Administration. Office of Highway Policy. Traffic Monitoring Guide. 2001, US DOT: Washington, DC.

Figliozzi, M., Johnson, P., Monsere, C., Nordback, K., A Methodology to Characterize Ideal Short-term Counting Conditions and Improve AADT Estimation Accuracy, Forthcoming 2014 Journal of Transportation Engineering.

Gallop, C., Tse, C., 2011. A Seasonal Autoregressive Model of Vancouver Bicycle Traffic Using Weather Variables. Presented at the Transportation Research Board Conference. Washington, D.C.

Hanson, S., Hanson, P., 1977. Evaluating the impact of weather on bicycle use. Transportationu Research Record: Journal of the Transportation Research Board, No. 629, 43–48

Hernandez, H. *TEMPORAL VARIATIONS FOR MONITORING TRAFFIC IN URBAN AREAS*. University of Manitoba, 2012.

Hess, Paul Mitchell; Moudon, Anne Vernez; Logsdon, M. G. Measuriing Land Use Patterns for Transportation Research. *Transportation Research Record*: *Journal of the Transportation Research Board*, Vol. 1780, 2001, pp. 17-24.

Jones, Michael. 2009. "National Bicycle & Pedestrian Documentation Project." In *Transortation Research Board Conference*. Washington, DC.

Lewin, A., 2011. Temporal and Weather Impacts on Bicycle Volumes. Transportation Research Record: Journal of the Transportation Research Board, No. 2536, 18p

Li, M.-T., F. Zhao, and Y. Wu. Application of Regression Analysis for Identifying Factors That Affect Seasonal Traffic Fluctuations in Southeast Florida. *Transportation Research Record*, Vol. 1870, No. 1, Jan. 2004, pp. 153–161.

Lusk A., Morency, P., Miranda-Moreno, L., Walter W. (2013) "Bicycle Guidelines and Crash Rates on Cycle Tracks in the United States" *American Journal of Public Health*, (in press).

Miranda-Moreno, L. and T. Nosal. Weather or not to cycle; whether or not cyclist ridership has grown: a look at weather's impact on cycling facilities and temporal trends in an urban environment. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 2300, Transportation Research Board of the National Academies, Washington, D.C., 2011

Miranda-Moreno, L. F., C. Kho, 2012. Winter Cycling in North American Cities: Climate and Roadway Surface Conditions. In: The 91st Transportation Research Board annual meeting, Washington D.C.

Miranda-Moreno, L.; T. Nosal; R. Schneider; F. Proulx; "Classification of Bicycle Traffic Patterns in five North American Cities". *The 93rd Annual Transportation Research Board*

Conference. Washington, D.C. January 15, 2013. *Forthcoming in the Transportation Research Record.*

Nankervis, M., 1999. The effect of weather and climate on bicycle commuting. Transport Research Part A, 33, pp. 417–431.

Nankervis, M., 1999. The effects of weather and climate on urban bicycle commuters' decision to ride. A pilot study. Road and Transportation Research, 8, pp. 85-97.

Nicholls, S. "Measuring the accessibility and equity of public parks; a case study using GIS". Managing Leisure. Volume 6. 2001. Pp. 201-219.

Nordback, K., Piatowski, D., Janson, B., Marshall, W., Krizek, K., Main, D., 2011. Using Inductive Loops to Count Bicycles in Mixed Traffic. Journal of Transportation of the Institute of Transportation Engineers, 2, 39-57

Nordback, K., W.E. Marshall, B.N. Janson, E. Stolz. Estimating Annual Average Daily Bicyclists: Error and Accuracy. Presented at 92nd Annual Meeting of the Transportation Research Board,Washington, D.C., 2013.

Parker, Kathryn M., Jeanette Gustat, and Janet C Rice. 2011. "Installation of Bicycle Lanes and Increased Ridership in an Urban, Mixed-Income Setting in New Orleans, Louisiana." *Journal of Physical Activity & Health* 8 (January): S98–S102.

PIARC Technical Commitee on Road Safety. (2003). Traffic Count. In *Road Safety Manual.* Route2Market: 2003; pp 510-524

PIARC Technical Commitee on Road Safety. (2003). Traffic Count. *Road Safety Manual.* Route2Market: 2003; pp 510-524

Pucher, J., and Beuhler, R., 2011. Analysis of Bicycling Trends and Policies in Large North American Cities: Lessons for New York. Washington, DC: Prepared for the Research and Innovative Technology Administration, United States Department of Transportation

Pucher, John, Jennifer Dill, and Susan Handy. 2010. "Infrastructure, Programs, and Policies to Increase Bicycling: An International Review." *Preventive Medicine* 50 (2010) (January): S106–25

Richardson, A.J., 2000. Seasonal and Weather Impacts on Urban Cycling Trips. Victoria: TUTI Report for The Urban Transport Institute

Robichaud, K. and M. Gordon (2003). "Assessment of Data-Collection Techniques Transportation Research for Highway Record: Journal of Agencies." The Transportation Research Board, No. 1855, Transportation Research Board of the National Academies, Washington, D.C., pp. 129–135.

Robitaille, E. *Geographical Indicators of the Built Environment and Services Environment Influencing Physical Activity, Diet, and Body Weight*. Institut national de santé publique du Québec. 2009.

Rose, G., Ahmed, A. F., Figliozzi, M., Jakob, C., 2011. Quantifying and Comparing Effects of Weather on Bicycle Demand in Melbourne, Australia, and Portland, Oregon. Transportation Research Record: Journal of the Transportation Research Board, No. 3205

Sharma, S.C., Gulati, B.M., and S.N. Rizak. (1996). "Statewide Traffic Volume Studies and Precision of AADT Estimates." Journal of Transportation Engineering, 122 (6),pp. 430-439.

Song, Y.; D. Rodriguez. The measurement of the level of mixed land uses: a synthetic approach. Carolina Transportation Program White Paper Series. 2004

Sprinkle Consulting, Editors. Pedestrian and Bicycle Data Collection, I. AMEC E&I and I. 2011, USDOT: Washington, DC

Strauss, J., Miranda-Moreno, L. Morency M. (2013) "Cyclist Activity and Injury Risk Analysis at Signalized Intersections: A Bayesian Modeling" *Accident Analysis and Prevention (in press).*

Thomas, T., Jaarsma, R., Tutert, B., 2013. Exploring temporal fluctuations of daily cycling demand on Dutch cycle paths: the influence of weather on cycling. Transportation 40, 1–22

Tin Tin, S., A. Woodward, E. Robinson, and S. Ameratunga. Temporal, seasonal and weather effects on cycle volume: an ecological study. *Environmental health : a global access science source*, Vol. 11, No. 1, Jan. 2012, p. 12.

Traffic Monitoring Guide, F.H.A. Office of Highway Policy Information, Editor. 2001, US DOT: Washington, DC.

Yang, S., C. Lu, F. Zhao, R. Reel, and J. D. O'Hara. Estimation for Seasonal Factors of Similarity-Based Traffic for Urban Roads in Florida. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2121, No. -1, Dec. 2009, pp. 74–80.

Zhang, X.; H. Lu; J.B. Holt. "Modeling spatial accessibility to parks: a national study". The International Journal of Health and Geographics. Volume 10. 2011