

Pedestrian Activity Model for prioritizing investment – A case
study of sidewalk snow clearing in the City of Waterloo

by

Bishoi Shinoda

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Masters of Environmental Studies

in

Planning

Waterloo, Ontario, Canada, 2018

©Bishoi Shinoda 2018

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Bishoi Shinoda

Abstract

This research emphasizes the needed integration for the empirical field of pedestrian modeling and the practical field of public works services. This is illustrated by focusing on the winter sidewalk maintenance delivery standards in Canadian municipalities, which often suffers from a mismatch with the spatial distribution of pedestrian activity. As a response, the first objective of this research is to predict the spatial distribution of pedestrian activity. This is done by reviewing five approaches to pedestrian modeling and demonstrating an understanding of the built environment and non-built environment variables that influence pedestrian demand. Based on common shortcomings to each approach, an analytical approach is proposed and used to construct a Pedestrian Activity Model (P.A.M.) predicting daily walking trip count per neighbourhood, with the City of Waterloo as the case study area.

Building on this, an analysis of the highest classes of the constructed P.A.M. is utilized to suggest a Pedestrian Priority Zone. This addresses the second research objective, which is to identify high foot traffic areas to construct a priority zone for delivering enhanced and efficient winter sidewalk maintenance.

Between the two tested regression types, the Spatial Error Regression (SER) is a better fit in capturing daily walking trips. Of the seven explanatory variables considered, only Transit Activity, Metric Reach (sidewalk connectivity), and Elementary and Secondary School Student Enrollment variables are significant under the SER model. As a result, the Pedestrian Activity Model is founded on the SER model and the 3 significant variables. The highest pedestrian activity class is found along University Avenue between University of Waterloo and Wilfrid Laurier University, while the second highest is found around the Uptown. A single Pedestrian Priority Zone is suggested based on amalgamating the three highest P.A.M. classes.

While the results are context-specific, the methodology is transferable. The process of constructing the predictive model can be used to validate other existing pedestrian models. Also, constructing a Pedestrian Activity Model could be an essential piece to decision making not just for enhancing public works services, but also for recommending new infrastructure connections, prioritizing streetscape enhancement projects, encouraging commercial and retail development, and boosting Real Estate market.

Acknowledgment

Dr. Clarence Woudsma, my supervisor, whose insights and support from the earliest stage to my research were essential in guiding my research interests, completing this thesis and broadening my planning knowledge, skills and experience.

Dr. Jean Audrey, my committee member, whose passion to teach and to see success in others has spirit lifted me during the darkest hours of this thesis and challenged my understanding of the material for a better grasp.

The School of Planning, including all academic and administrative staff (especially Prof. Jennifer Dean, Kelly Heald-Oliver, Taylor Ertel, and Tiffany Chen), without their support; I would not have been successful in fulfilling the degree requirements.

My father, who drove me to my committee meetings for the past year. My mother, who was worried if ate healthy as I stayed on-residence. My brother and his wife, who celebrated my success and constantly encouraged me. My grandma, who despite living in Egypt, always asked about me and prayed for my success. I am deeply grateful to all my family and for all the big and little things they did to support me.

My school colleagues, who shared strong support for each other and offered input and feedback supporting my academic growth.

My friends who motivated, and encouraged me as needed. My church family and friends, who prayed for me and encouraged my professional development.

Lastly but most importantly, I am grateful to my God, who is the reason for my success. I came to learn this in the past couple of years that I am extremely blessed with everyone and everything that came into my life in the particular order they happened in.

Table of Contents

AUTHOR’S DECLARATION	ii
Abstract	iii
Acknowledgment	iv
List of Figures	vii
List of Tables	ix
Chapter 1: Introduction	1
1.1. Research Justification	1
1.2. Winter Sidewalk Maintenance background	5
1.3. Research Goal and Objectives	8
1.4. Anticipated Research Contribution	9
1.5. Thesis Outline	10
Chapter 2: Literature Review	12
2.1. Active Transportation Planning	12
2.1.1. Common Active Transportation Modes	12
2.2. Winter Maintenance	14
2.2.1. Winter Road maintenance	14
2.2.2. Winter Sidewalk maintenance	17
2.3. Approaches for estimating pedestrian activity	19
2.3.1. Mode Choice (4SM)	20
2.3.2. Walkability Index	22
2.3.3. Space Syntax Model	25
2.3.4. Gravity Model	28
2.3.5. Microsimulation/Agent-based Model	30
Chapter 3: Methods	33

3.1. Overview of the Approach _____	33
3.2. Study Location _____	35
3.3. Research design _____	39
Constructing a Pedestrian Activity Model _____	39
3.3.1. Geographical Unit of Analysis _____	39
3.3.2. Built Environment and Non-built Environment Variables _____	42
3.3.3. Regression Analysis _____	63
3.3.4. Pedestrian Activity Model Formula _____	70
Pedestrian Priority Zone _____	71
Chapter 4: Findings and Discussion _____	72
4.1. Regression Models _____	72
4.2. Pedestrian Activity Model _____	89
4.3. Pedestrian Priority Zone Configuration _____	104
4.4. Research Implication _____	108
Chapter 5: Conclusion _____	110
5.1. Conclusion _____	110
5.2. Further Research _____	114
5.3. Recommendations _____	115
Bibliography _____	116
Appendices _____	127
Appendix A – Considered Variables – Not Included In Final Model _____	128
Appendix B – Single Variable Regressions’ Scatterplots _____	133

List of Figures

FIGURE 2.3.3.1	28
FIGURE 2.3.4.1	29
FIGURE 3.2.1	36
FIGURE 3.2.2	36
FIGURE 3.2.3	37
FIGURE 3.3.2.1	44
FIGURE 3.3.2.2	45
FIGURE 3.3.2.3	46
FIGURE 3.3.2.4	47
FIGURE 3.3.2.5	48
FIGURE 3.3.2.6	50
FIGURE 3.3.2.7	51
FIGURE 3.3.2.8	53
FIGURE 3.3.2.9	54
FIGURE 3.3.2.10	56
FIGURE 3.3.2.11	57
FIGURE 3.3.2.12	59
FIGURE 3.3.2.13	60
FIGURE 3.3.2.14	62
FIGURE 3.3.2.15	62
FIGURE 3.3.3.16	65
FIGURE 3.3.3.17	66
FIGURE 3.3.3.18	66
FIGURE 3.3.3.19	68
FIGURE 4.1.1	74
FIGURE 4.1.2	74
FIGURE 4.1.3	75
FIGURE 4.1.4	75
FIGURE 4.1.5	76
FIGURE 4.1.6	76

FIGURE 4.1.7	79
FIGURE 4.1.8	83
FIGURE 4.1.9	84
FIGURE 4.1.10	85
FIGURE 4.1.11	86
FIGURE 4.1.12	87
FIGURE 4.1.13	88
FIGURE 4.2.1	91
FIGURE 4.2.2	92
FIGURE 4.2.3	93
FIGURE 4.2.4	95
FIGURE 4.2.5	96
FIGURE 4.2.6	98
FIGURE 4.2.7	100
FIGURE 4.2.8	101
FIGURE 4.3.1	106
FIGURE 4.3.2	107
FIGURE B.1	133
FIGURE B.2	133
FIGURE B.3	134
FIGURE B.4	134
FIGURE B.5	135
FIGURE B.6	135
FIGURE B.7	136
FIGURE B.8	136
FIGURE B.9	137
FIGURE B.10	137

List of Tables

TABLE 1.2.1.....	7
TABLE 2.3.1.....	20
TABLE 2.3.2.....	32
TABLE 3.3.1.1.....	41
TABLE 4.3.1.....	103

Chapter 1: Introduction

1.1. Research Justification

Pedestrianism is not fully considered in infrastructure investment in North American cities.

Following the Second World War, population growth in these cities diverged away from city centers towards its outskirts. This marked the era of suburbanization and car-dependence that still persists till today. With car-dependence, pedestrian activity is discouraged by street design (Blomberg et. al., 2000), and built form. Many New Urbanism and 'Winter' City scholars and practitioners have spoken to the disconnect between the public realm (e.g., pedestrian environment) and private realm (e.g., building and site design, services' accessibility, and land use patterns) as one of the leading causes of low walking trips in urban and suburban communities (Trudeau, 2013; Coleman, 2010; Parolek et. al., 2008; Pressman, 1996; Hough Stansbury Woodland Ltd., 1990; Manty & Pressman, 1988).

In recent decades, there has been a growing interest in active transportation, especially in public health and transportation fields. While active transportation is an inclusive term, referring to walking, bicycling, skating, and other forms of non-motorized mobility, in this thesis active transportation solely refers to walking. In the public health field, researchers are concerned for the epidemic of obesity, Type II diabetes, and chronic heart diseases among urban populations due, in part, to an overall lack of physical activity (Public Health, City of Toronto, 2012; Sundquist et. al., 2011; Frank et. al., 2010; Lee and Moudon, 2004; Moudon and Lee, 2003). Walking has been proven to provide both a mean of transportation and a physical activity (Lee and Moudon, 2004). Many related public health studies try to better understand and measure the correlation between physical activity and walkability (City of Toronto, 2012; Sundquist et. al.,

2011; Frank et. al., 2010; Lee and Moudon, 2004; Moudon and Lee, 2003). Walkability is a measurement of walking potential based on the easiness of walking between places (City of Thunder Bay, 2017; Tsiompras and Photis, 2016; Region of Waterloo, 2014; Frank et. al., 2010).

From the perspective of urban planning, planning for active transportation has been and still is a big part of the growing effort to battle urban sprawl. These efforts could be classified into two approaches: bigger and beautified pedestrian scape, and improved integration with land use planning. These efforts are linked with contemporary planning movements, such as Smart Growth, Complete Streets, and New Urbanism (Hui et. al., 2018; Hong, 2016; Tracz, 2015; Trudeau, 2013; Knaap and Talen, 2005; Lund, 2003). Smart Growth focuses on integration with land use planning as it is known for infill and mixed-use development approach (Hong, 2016; Knaap and Talen, 2005). On the other hand, Complete Streets focus mainly on the pedestrian scape, emphasizing active transportation planning by re-designing the right of way (ROW) to accommodate all street users. A key signature for Complete Streets is often a very wide ROW to accommodate a space for each user – cars, buses, bicycles, and pedestrians (Hui et. al., 2018; Tracz, 2015). New Urbanism is somewhat of a hybrid of Smart Growth and Complete Streets as it advises both development design and the public realm (e.g., streetscape, and public space) to enhance pedestrian activity, creating livable neighbourhoods and cities (Hong, 2016; Trudeau, 2013; Knaap and Talen, 2005). As is true for many planning issues, there is no single solution nor perfect solution. It is our duty as community members to build on older solutions and approaches to improve or suggest new solutions improving our quality of life. As the world evolves, newer and improved solutions are needed, which is the case for active transportation planning too.

In the quest to enhance pedestrian activity, we must also consider the influence of climate on walking; whether heat, cold, rain, wind, or snow, people respond differently to different climate conditions. A study by Vanky et. al. (2017) based in the City of Boston, found a negative correlation between walking activity and various climate conditions and seasons, such as precipitation during spring and humidity during both Autumn and Spring. A Canadian case study found a negative correlation between snow and walking trips (Miranda-Moreno and Lahti, 2013). Similar findings on the effect of various weather elements and conditions on walking are documented in various studies around the globe (Hong, 2016; Shaaban and Muley, 2016; Böcker et. al., 2013).

This thesis is concerned with pedestrian activity and safety during the winter months. Issues such as snow accumulation, uneven snow clearing practices, pedestrian network disconnections, safety hazards, and greater trip durations often discourage or create barriers to winter walking activity (TriTag Transport Action Group, 2018; Vanky et. al., 2017; City of Toronto, 2016; Miranda-Moreno and Lahti, 2013; Li et. al., 2013). A recent study by the TriTag group based on the streets of Kitchener, found that 50% of people within a 50m-walk are likely to encounter an uncleared snow-packed patch of sidewalk during the winter months (Thompson, 2018; TriTag Transport Action Group, 2018). Some of these concerns can be traced back to inadequate and/or inconsistent winter sidewalk maintenance practices (e.g., snow shoveling, and salt application) (TriTag Transport Action Group, 2018; Li et. al., 2013). While city occupants and landlords are required under Ontario's Occupiers' Liability Act and Residential Tenancies Act to maintain hazard-free premises, which includes snow clearing (Preszler Injury Lawyers, n.d.), some municipalities have taken responsibility for snow clearing duties to improve pedestrian

accessibility and safety (Preszler Injury Lawyers, n.d.; City of Toronto, 2016). However, snow clearing practices vary across municipalities (see Table 1.2.1) and individual properties.

How could a municipality fund growing needs for a sidewalk snow clearing program?, Should all sidewalks be cleared or be prioritized? How would sidewalks be selected or prioritized for clearing, e.g., along bus routes, along only arterial roads, around malls, around hospitals, around co-op and social housing, around elementary schools, or based on mixed criteria? As discussed below and in the literature review, current snow clearing practices for sidewalk specific segments usually have simple justification. They rely on common-sense-based assumptions like downtown cores or all bus-routes and arterial roads all have similar pedestrian activity. This research challenges that status-quo approach by recommending sidewalk snow clearing prioritization in identified high foot traffic areas. From a planning and feasibility perspectives, it is important to invest where the most outcome is anticipated. In the thesis, this will be achieved by estimating walking trips, through the case study area, by testing, and using walking-associated explanatory variables (e.g., elementary school enrollment, land use mix, and sidewalk network). This is proof-of-concept research that links the field of active transportation planning (i.e., pedestrian activity modeling) and the field for delivering efficient public services (i.e., winter sidewalk maintenance) to improve public safety and improve winter-based walking activity convenience.

1.2. Winter Sidewalk Maintenance background

It was only recently that winter sidewalk maintenance was included in Ontario's minimum maintenance standards regulations (Ontario, 2018a; City of Toronto, 2014; Ontario, 2013).

Initially, some cities might have looked into the Accessibility for Ontarians with Disabilities Act (AODA) to verify any requirements for providing winter sidewalk maintenance, such as level of service, or minimum clearance width required, but the AODA also remains silent on the matter (Ontario, 2016; City of Toronto, 2014). The recent update to Ontario's Minimum Maintenance Standards was adopted in May 2018, and it updates the standards around the level of service and minimum clearance width for sidewalk snow clearing (Ontario, 2018a). Despite the newly added section 2.1, the newly announced standards are not mandatory but rather a set of guiding principles that is available to municipalities (Ontario, 2018b). In addition, subsection 44 (9) of the Municipal Act excuses municipalities from any liability due to personal injuries caused by snow or ice on sidewalks except in gross negligence scenarios (Ontario, 2018c). In addition, the newly adopted standards do not address prioritizing specific sidewalk segments by foot traffic volume unlike road snow clearing prioritization schemes that are based on road speed and vehicular traffic volume (Ontario, 2018a).

Prior to the latest version of Ontario's Minimum Maintenance Standards, local municipalities developed their own set of practices/guidelines for winter sidewalk maintenance. Most Canadian municipalities have a by-law addressing winter sidewalk maintenance, reducing liability due to related-personal injuries (e.g., slips and falls).

The case study location for this research is the City of Waterloo; therefore, it is important to understand the city's current practices and those of the adjacent municipality, the City of Kitchener. Both cities have assigned business and property owners the duty of clearing their

fronting sidewalks, while both municipalities will clear sidewalks adjacent to their public facilities and where no property owner is adjacent to the sidewalk (City of Kitchener, 2016; City of Waterloo, 2009). Also, both municipalities identified 24 hours after a snow event as the window for city residents, business owners, and its public works department to complete snow clearing activities (City of Kitchener, 2016; City of Waterloo, 2009). Following the updated standards' regulation, the City of Kitchener council approved 2 of 5 recommendations presented to the Community and Infrastructure Services Committee (Nielson, 2018; Pickel, 2018). The two approved proposals include proactive enforcement of the by-law to ensure timely response to storm events and snow accumulation, and funding for partnership programs to assist those in need (Nielson, 2018; Pickel, 2018).

Other municipalities have stepped up their winter sidewalk maintenance efforts by providing prioritized and customized services. Table 1.2.1 provides a summary of sidewalk snow clearing standards in major Canadian municipalities. These municipalities make decisions about their winter sidewalk maintenance practices (e.g. snow clearing, and ice prevention) using four considerations. The first consideration is the desired sidewalk surface condition, such as achieving bare pavement. The second and third are the triggering snow accumulation level for maintenance practices deployment and the period it takes to complete the first round of clean up after the end of each storm (see Table 1.2.1). The fourth consideration is prioritizing certain sidewalk networks for which there is earlier deployment of services and shorter completion time. According to the table below, all cities of Toronto, Ottawa, and Edmonton have a prioritized sidewalk system. However, they do not have a unified scheme for prioritizing specific sidewalk routes. Some prioritize sidewalks along arterial roads, in downtown areas, near transit stops and/or near school areas (City of Halifax, n.d.; City of Ottawa, n.d.). The common prioritization

regime often favours high car traffic areas, which may not align with the spatial distribution of pedestrian demand, especially in outer city edges. Only the City of Toronto identifies that it uses pedestrian volumes as an indicator of where sidewalks are prioritized (City of Toronto, 2013), but there is no record of the method or the data used. This is where the research in this thesis fits in, establishing a method to identify high pedestrian activity neighbourhoods as the basis to prioritize sidewalk snow clearing and other public works services.

Table 1.2.1: Sidewalk Snow Clearing Standards in Major Canadian Municipalities

City	Desired Pavement Condition by Sidewalk Class	Threshold for Plowing by Sidewalk Class	Completion Time for Plowing By Sidewalk Class	Comments
Toronto*	Safe & passable on all class; 90% salt / 10% sand blend is applied to provide traction and melt snow/ice All sidewalks cleared where mechanically possible.	2cm – high pedestrian volume 8cm – low pedestrian volume	13 hrs to complete one round on all road class (approx 4 hr high volume, 9 hr low volume)	Residents responsible for clearing snow on 1,100km of sidewalk under by-law Snow removal (haulage or on-site melting) is an unbudgeted activity.
Montréal	Safe & passable on all class; 50% salt / 50% crushed gravel is applied to provide traction as temperatures are typically too cold for salt All sidewalks are cleared.	2.5 cm for all sidewalk class	14 hrs Manual sidewalk clearing in downtown core	Sidewalk plowing supplemented by frequent snow removal on local roads in old city Snow removal budget is \$110M
Ottawa	Bare pavement when conditions allow on arterial, Safe & passable on local All sidewalks are cleared.	2.5 cm in downtown 5 cm for remaining sidewalks	4 hrs for areas with high concentrations (CBD/Malls/etc...) 12 hrs for most primary sidewalks 16 hrs for most residential sidewalks	Sidewalk plowing supplemented by frequent snow removal on local roads in old city Snow removal budget is \$8M
Winnipeg	Bare pavement in downtown core when conditions allow, Snow packed on remaining sidewalks All sidewalks are cleared.	5 cm for Priority 1 & 2 8 cm for residential sidewalks	36 hrs for most sidewalks 5 days for residential sidewalks	
Calgary	Safe & passable on all class Sidewalks are only cleared on major roads, collectors and bus routes where the sidewalk is not adjacent to private property.	Not specified	4 days after end of snowfall	The owner or occupant of a private piece of land shall remove snow and ice from adjacent sidewalks within 24hrs
Edmonton	Sidewalks are only cleared adjacent to City owned land. Sidewalks adjacent to private property are the responsibility of the property owner.	Not specified	24 hrs (only near transit facilities) 48 hrs (only near city-owned land)	Residents responsible for clearing snow on 3,562km of sidewalk under by-law Snow removal budget is \$5.4M
London	Snow packed on all class	8cm for all sidewalk class	24 hrs	

*Toronto standards represented are for 2015 - 2022

Winter Maintenance Levels of Service
For Bikeways, Windrows, Sidewalks & AODA - Supplementary Report

Source: City of Toronto. (2014). Staff Report – Confirmation of Levels of Service for Winter Maintenance of Bikeways, Windrow Opening, Sidewalk and AODA Compliance.

1.3. Research Goal and Objectives

The goal of this thesis is to develop an analytical approach to prioritize winter maintenance of sidewalks. The motivation is to enhance pedestrian winter mobility and mitigate associated safety hazards. The study is based on a case study methodology and deploys quantitative methods to achieve the research goal. The case study area is the City of Waterloo, which is located west of the Greater Toronto Area. The main form of quantitative analysis utilized is geospatial analysis using secondary data.

The first objective of this study is to predict the spatial distribution of pedestrian demand. Although we do not know the extent to which pedestrian activity is reduced in the case study area during winter, we assert that there very likely is a decline, based on findings for other locations (Vanky et. al., 2017; Miranda-Moreno and Lahti, 2013). While estimating pedestrian demand was initially inspired by walkability indices, the focus is instead on predicting actual trip counts spatially. This approach is dependent on available secondary data sources, their quality, and their correlation to the response variable (i.e., daily walking trips).

The second objective of this study is to suggest a Pedestrian Priority Zone. Based on the spatial distribution of predicted walking trips, neighbourhoods are categorized/classified according to their associated pedestrian activity level. An analysis of the highest pedestrian activity neighbourhood classes is used to suggest a configuration for a priority zone. The priority zone is the founding step to re-direct, focus, and prioritize sidewalk snow clearing.

1.4. Anticipated Research Contribution

The anticipated key contribution of this thesis is to provide a proof-of-concept approach for linking the field of active transportation planning, especially foot-traffic studies, and the field of winter sidewalk maintenance. The established link could potentially improve pedestrian mobility and safety through improved sidewalk snow clearance. This analysis does not include analysis of the temporal seasonality effects on pedestrian activity; rather the approach taken is to predict daily walking trips, which would then be used to prioritize sidewalk snow clearing routes. The research will highlight various data sources that are available to capture both actual pedestrian activity and its potential application at the street-level.

1.5. Thesis Outline

This thesis is divided into five chapters: Introduction, Literature Review, Methods, Findings and Discussion, and Conclusion. The Introduction Chapter introduces the research background, goal, and objectives.

Following the introduction, Chapter 2: Literature Review synthesizes academic and grey literature around the basics of active transportation planning, and the history for winter road and sidewalk maintenance. In addition, it explores five approaches for estimating or predicting pedestrian activity, which feeds into the Methods Chapter and the construction of the Pedestrian Activity Model. Each of the five approaches is examined in terms of commonality, construct, and pros and cons.

Chapter 3: The Methods chapter starts off with an outline of the approach plus an overview of the case study location. For the case study area, transportation statistics are shared to demonstrate the share of walking mode versus other modes. The research design comes in later with a focus on the components and the detailed approach for constructing the Pedestrian Activity Model. Prior to that is the presentation of the study's geographical unit of analysis. This chapter also outlines the regression analysis process, along with the shortcomings of using TTS reported walking trips as the response variable. Lastly, there is a brief description of an approach for identifying a Pedestrian Priority Zone.

Chapter 4: Findings and Discussion shares and interprets the findings from the regression analyses. Based on the findings, the Spatial Error regression model is used as the foundation for the predictive model: "Pedestrian Activity Model." Findings from the model are divided into generic findings and class-specific findings. The class-specific findings explain the local context behind the predicted walking trips using the significant variables and other features. Lastly, the

chapter reveals the analysis of the highest three classes of Pedestrian Activity Model, which are used to suggest a Pedestrian Priority Zone.

The Conclusion chapter re-caps the research goal, and objectives. Afterward, the chapter highlight the key findings of the thesis followed by the limitations for the Pedestrian Activity Model. Lastly, the chapter ends with suggesting potential improvements to this research, that were not carried out because of resource limitations, and with the next step to applying the Pedestrian Priority Zone to the field of winter sidewalk maintenance.

Chapter 2: Literature Review

2.1. Active Transportation Planning

As the term suggests, “active transportation” refers to a form of transportation that is also a physical activity. According to the Public Health Agency of Canada (2014), “Active transportation refers to any form of human-powered transportation.” Walking, bicycling, and others are the most common active transportation forms that appear in planning and public health literature (PHAC, 2014; Iacono, Krizek, and El-Geneidy, 2010; Lee, and Moudon, 2004; Cervero, and Kockelman, 1997). It was not until suburbanism revealed the full implications of a car-dependent culture and land use segregation, that active transportation gained grounds as a counter-movement in both transportation planning and public health fields. The growing interest has funded projects into understanding walking and biking activity as well as encouraging investments that improve active transportation participation and neighbourhoods’ quality of life.

2.1.1. Common Active Transportation Modes

Walking is the oldest mode of transportation dating back to the hunter-gatherer settlements. Despite walking being the oldest form of transportation, we do not have a full understanding of how to predict it. Unlike car drivers and cyclists, pedestrians are not confined by the edge of the sidewalk or by a boulevard. Pedestrians can cross middle of the road, walk over grass, or take a shortcut through a building or a trail through a woodlot. Pedestrian flow is more dynamic compared to motorized transportation; however, it is sensitive to environmental surroundings and affected by socio-economic status (Manaugh, and El-Geneidy, 2011). The set of variables contributing to a person’s decision to walk or not, is different than that to drive. Some scholars even found that trip purposes correspond differently to environmental variables (Manaugh, and El-Geneidy, 2011; Lee, 2004).

On the other hand, bicycles were more recently invented in the early 17th century. Most transportation agencies treat bicycles in a similar manner as automobiles in regard to traffic laws and designated infrastructure (e.g., roadways) (Ontario, 2014; California: 2011). Partly as a result, cyclists' movements are more predictable than pedestrian's and easier to model. The most redundant concern with cycling is safety. Municipalities are usually torn between expanding dedicated bicycling infrastructure (e.g., separated bike lanes) and balancing costs. From this point forward, the thesis will narrow focus on walking activity and how to approach pedestrian activity modeling.

2.2. Winter Maintenance

2.2.1. Winter Road maintenance

We have been systematically clearing snow from transportation infrastructure for nearly two centuries. Earliest record shows that railway companies used “horse-drawn plows to clear railways in 1831 and to clear city streets in 1862,” (Minsk, 1998; Sullivan, 1831). Mechanical plowing equipment was not developed till the late 19th century, to provide faster and more efficient way of clearing railways (Minsk, 1998). The first record of truck-mounted plows used for winter road maintenance was during the winter of 1920-21 in New York, USA (Minsk, 1970).

The growth of an integrated road network in the United States in the 20th century pushed further the development of snow removal technology (Minsk, 1998; Minsk, 1970). The second wave of winter maintenance advancement came in the late 20th century (Kuemmel, & National Cooperative Highway Research Program, 1994). National programs, like the Strategic Highway Research Program (SHRP), were initiated to study and report on the effectiveness and efficiency of various snow and ice control programs (Smithson, 2004; Kuemmel, & NCHRP, 1994). For example, SHRP reported on best practices such as Road Weather Information System (RWIS), which revolutionized winter road maintenance (Kuemmel, & NCHRP, 1994). Other research in the field of winter road maintenance also included examining the effectiveness of various deicing materials, such as salt (sodium chloride), and calcium chloride, in addition to the use of abrasive mixtures (e.g., a mix of sand and salt), for achieving bare pavement standards and improving driving conditions. Ontario, after testing the effectiveness of straight salt versus abrasive mix on achieving the intended bare pavement standard, expanded the use of straight salt to its winter maintenance practices of highways (Kuemmel, & NCHRP, 1994). Overall, the growing research

and the integration of technology since the late 1970s (Minsk, 1998) in the field of winter maintenance has advanced current practices' efficiency.

Winter maintenance practices can be classified into three categories: mechanical, chemical, and thermal (Minsk, 1998). The objective for the mechanical approach is to push, lift, or cast snow “sufficiently far ... to reduce the necessity for rehandling it” (Minsk, 1998). Snow is usually overthrown onto the boulevard or the roadside or hauled to designated storage areas.

Displacement plows are the most commonly used in winter road maintenance and are usually mounted to a truck's front and sides (TAC, 2013; Minsk, 1998). The design and development of Displacement plows were pioneered by the railway companies, as discussed above, to remove snow along the tracks (Minsk, 1998).

Unlike mechanical treatment, chemical treatments can be applied in both proactive and reactive scenarios. Salt has a limited effectiveness in lower temperatures, yet its low cost and its versatility make it the most common winter road maintenance chemical application (Nassiri et al., 2015; Minsk, 1998; Kuemmel, & NCHRP, 1994). Pre-wetting is an addition to direct salt spreading and is used to speed the effectiveness of salt application. Pre-wetting includes spreading a brine (usually water and salt mixture) or a liquid freeze-point depressant along with a solid salt application (TAC, 2013; Minsk, 1998).

The thermal application is the least used of the three winter maintenance practices categories. The use of heat to melt away snow, whether through built-in heating systems within roads and bridges, or hauling snow to melting stations, is too expensive, especially when considering the roadway dimensions (TAC, 2013; Minsk, 1998; Kuemmel, & NCHRP, 1994).

The degree to which each municipality uses each maintenance practices depends on different factors such as weather conditions, environmental restrictions, equipment availability, staff training, contractors and budgetary items. As a result, each municipality has a unique plan for battling snow and ice on its roadways. Upper-tier governments like the Province of Ontario set out minimum maintenance standards for municipalities to follow that include patrolling frequencies, snow accumulation limits, and maintenance time of completion restrictions; however, they do not set out the exact practice in achieving the standards (Government of Ontario, 2013).

News outlets and safety-related agencies remind us annually of dangerous winter driving conditions and accidents happening due to snowfall or icy roads. Road crashes have a financial toll of \$10 billion annually on the Canadian health care system, with weather-related collisions accounting for about \$1 billion (Andrey et. al., 2001). Several studies concluded that overall adverse winter weather conditions increase the risk of road accidents by poor road surface conditions while investing in winter road maintenance significantly reduces the risk by improving road surface conditions (Ye et. al., 2013; Usman et. al., 2010; Qiu and Nixon, 2008). Chemical application is a key player in improving safety and reducing the risk of road accidents. Salt application as well as sanding prevent the bond of snow to the road surface and increase the road friction. The deployment of chemical (e.g., salt) application is dependent on the four Rs: Right material, Right amount, Right place, and Right time (TAC, 2013).

One of the technological advancements in the field, which allows for more proactive and responsive winter road maintenance and therefore improved safety around winter driving, is Road Weather Information System (RWIS). RWIS stations provide real-time local information about road surface temperature, level of moisture, and presence of deicing chemicals using

various sensors (Nassiri et. al., 2015; TAC, 2013; Minsk, 1998). The local live data transmitted by the RWIS stations help the decision makers in regard to the four Rs for chemical application deployment.

Not all roads are cleared at once, nor to the same standards. Government bodies establish standards based on various considerations. Road authorities usually use “Average Daily Traffic” (ADT) as the key indicator of the heaviest traffic in prioritizing what roads receive treatment first and the desired level of service (Nassiri et. al., 2015; Kuemmel, & NCHRP, 1994). In addition to ADT, emergency routes, school areas, and major transit routes are factored into determining priority snow clearing routes (Nassiri et. al., 2015; Minsk, 1998). The Province of Ontario uses a combined approach of ADT and posted speed as the criteria for a road classification scheme (Ontario, 2013). The road classification scheme is integrated with the winter road maintenance deployment of the road-class-based level of service as illustrated in Ontario’s minimum maintenance standards regulation (Ontario, 2013).

2.2.2. Winter Sidewalk maintenance

Research and adaptation of best practices in the field of winter sidewalk maintenance are not as fast and as responsive as winter road maintenance (City of Toronto, 2014). As a result, winter sidewalk maintenance is limited mainly to manual shoveling, traditional mechanical plowing and basic applications of chemical treatment. The City of Toronto staff (2014) reported that sidewalk plowing equipment manufactures is slow in adopting the latest technologies from winter road maintenance field. Businesses and property owners are usually dependent on manual shoveling and personal snowblowers to clear snow, while municipalities usually use heavier mechanical equipment (e.g., Snow Plows). Some Canadian municipalities provide basic chemical treatments to sidewalks such as salting and sanding (City of Markham, n.d.; City of Kitchener, 2016; City of

Toronto, 2014). Private property owners are free to use the salt as they see fit on sidewalks, which raises concerns about over-salting and its environmental impact on vegetation and groundwater (Hosseini et. al., 2016; Fay and Shi, 2012).

In some rare instances, municipalities have adopted thermal application but it is usually limited to short sidewalk segments. Heating sidewalks to melt away snow and ice is impractical on large scale because of the overly large capital cost for heating system installment, while on the other hand, cost-effective alternatives include chemical and mechanical approach (Eugster, 2007). Examples of implemented thermally-heated sidewalks are reported in the Town of Nagaoka, Japan (Iwamoto et. al., 1998), and City of Aomori, Japan (Eugster, 2007), and is even being considered for Montreal, Canada (Lowrie, 2017).

2.3. Approaches for estimating pedestrian activity

The literature concerned with pedestrian activity measurement is rich and ever-evolving in incorporating new technological advancements. To gain a grasp of the various approaches and how they differ, I present Raford's and Ragland's (2006) classification of the three streams to estimating pedestrian activity. Under each of the streams are a number of approaches that represents that stream.

The first stream is the "Sketch Plan Model," which is usually adopted at the regional level to produce a rough estimate of pedestrian activity or its potential using simple planning guidelines (Raford and Ragland, 2006). The second stream is the "Network Analysis Model." It provides finer estimates of pedestrian activity than the former stream and is usually adopted at the city or neighbourhood level. The advantage of the second stream is its simplicity compared to the third stream, which has greater retention of details, often at street-level, than any of the first and second approaches. The third stream is computation heavy. It is called Microsimulation because it simulates the movement of individuals. Microsimulation was developed to map out individual movements in confined space like: airports, and malls (Raford, Ragland, 2006). It can also be used for small outdoor geographical areas (e.g., a single street or a small number of streets) due to the associated heavy computation (Omer and Kaplan, 2017). In total there are five approaches discussed below, that are classified under each of the three streams (see Table below). For each of the approaches reviewed below, I break it down into three sub-sections: commonality, construct, and pros and cons.

Table 2.3.1: Classification of Approaches to Estimate Pedestrian Activity (source: Raford, N., & Ragland, D. (2006). Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: The Case of Boston, Massachusetts.)

Approach	Sketch Plan Model	Network Analysis Model	Microsimulation
Tools and Measures	Mode Choice (4SM)	Gravity Model	Agent-based Model
	Walkability Index	Space Syntax Model	

2.3.1. Mode Choice (4SM)

Mode Choice is one of four stages in the travel demand forecasting model known as the Four-Step Model (4SM), which is commonly used by regional planning and transportation organizations in North American Cities (Clifton et. al., 2016; Davidson et. al., 2007; McNally, 2000). Transportation analysis has evolved tremendously in response to the post-war development and economic growth, during which, the 4SM (i.e., trip generation, trip distribution, mode split, route assignment) was developed (McNally, 2000). The 4SM is popular among large metropolitan cities, regional governmental bodies and state departments. In the US, transportation analysis and planning are conducted by Metropolitan Planning Organizations, Regional Planning Agencies, and States’ Department of Transportation (Davidson et. al., 2007).

Mode Choice is usually estimated using discrete choice models or nested logit models (Davidson et. al., 2007; McNally, 2000). It usually comes in the third step of the 4SM. The first step calculates trip generation, followed by trip distribution. The first step is dependent on survey data about the origins and destinations of trips. The second step, “trip distribution”, utilizes the “Gravity Model,” which will be discussed later, to distribute trips based on surrounding attractions and traffic impedance. The third step and the core of this sub-section is called the

“Mode Split” or “Mode Choice” because it identifies each transportation mode share for the study area. As mentioned above, there are two common methods to calculating the “Mode Split”; however I will only focus on the nested logit model because it is more commonly referenced in the literature (Davidson et. al., 2007; Jonnalagadda et. al., 2001; McNally, 2000). The logit analysis estimates trade-offs among variables for the transportation modes: examples include in-vehicle time, frequency of service, reliability, and crowdedness (Jonnalagadda et. al., 2001). Since these variables are car- and transit-oriented, Pedestrian Environmental Factors (PEFs) are incorporated to assist the model to predict pedestrian activity (Jonnalagadda et. al., 2001). PEFs include measures, like pedestrian network connectivity, perception of safety, urban vitality, and topological barriers. The required data sources are household surveys, land use database, and level of service (LOS) characteristics for each mode (Jonnalagadda et. al., 2001; McNally, 2000). Household surveys typically aim for a 5% sample target of households in the survey area (Data Management Group, 2014), and the level of service (LOS) is considered for all modes at every origin and destination (Jonnalagadda et. al., 2001). Estimating pedestrian activity levels is often limited to factors such as sufficient reporting of walking trips in the survey sample, and understanding of trip attractions for a pedestrian versus other users.

Despite advancements in estimating mode choice, the model output still fails to fully represent trips by non-auto modes, especially walking (Clifton et. al., 2016). The conventional four-step model (4SM) was adopted to estimate future travel demand based on large-scale infrastructure projects (McNally, 2000). The 4SM usually considers average vehicle occupancies as total person trips, which emphasizes the role of the automobile and by default de-emphasizing other modes. The success of this 4SM has been to data availability, which fueled the growing use of Household Surveys (Davidson et. al., 2007; McNally, 2000). Using the Household survey data

has a scale-related disadvantage when assessing pedestrian activity. Data are usually aggregated to geographical areas, suited for motorized modes, called Traffic Analysis Zones (TAZ) (Iacono et. al., 2010). TAZ unit areas are too big to capture the spatial movement of pedestrians (Eash, 1999 as cited in Iacono et. al., 2010). As a result, aggregating data into large areas (e.g., TAZ, Census Tracts) will most likely lead to lost travel demand sensitivity for low-share modes due to aggregation errors (Clifton et. al., 2016; McNally, 2000). Also, aggregated data could flatten the trip numbers across larger areas, which would make the activity looks low. Therefore, the suitability of large geographical units of analysis is questioned.

Advancements to the 4SM included further details about each mode choice, which re-states the model holistic approach to transit analysis rather than this research's intention of examining just walking activity. Davidson et. al. (2007) speak about how the 4SM still produces reasonable predictions, which makes the switch to more accurate predictive models unnecessary for smaller regional planning organization. In terms of this research, the results obtained from the 4SM model could only be used as an overall validation or representation of pedestrian activity. However, the aggregated results would present an ecological fallacy issue when transferring the results to the disaggregate area (e.g., street level).

2.3.2. Walkability Index

Walkability refers to the ease of walking between places and is largely a function of physical proximity and the connectivity between origins and destinations (City of Thunder Bay, 2017; Tsiompras and Photis, 2016; Region of Waterloo, 2014; Frank et. al., 2010). Walkability indices, which are intended to score/quantify the level of walk potential of an area, are mainly developed and used by professionals and academics in both fields of transportation planning and public health. A well-known example of a walkability index is offered by Walkscore.com, a product

developed by a private company that provides scores for any address in the U.S., Canada, Australia, and New Zealand (Walkscore.com, n.d.-a). More locally, the Region of Waterloo offers a second example, with its own walkability index developed under the NEWPATH initiative in 2009 (Region of Waterloo, 2014). Walkability indices are used in various sectors: in the transportation planning sector to indicate levels of pedestrian activity potential (Raford and Ragland, 2006); in the public health sector to examine influences on healthy living and improves quality of life (Region of Waterloo, 2014); and by the real estate sector to highlight accessibility of a property in question via walking to surrounding services.

The easiest-to-interpret formulas for walkability are those based on simple sums. Because different variables have different units of measurement, it is common for indices to be based on the sum of z-scores (Frank et. al., 2010). There is no limit on the number of variables that a walkability index could include. Some indices have as few as four variables (Region of Waterloo, 2014; Frank et. al., 2010), and some include numerous variables (Cervero and Kockelman, 1997). The most commonly used factors are intersection density, net residential density, retail floor area ratio, proximity to basic land uses and land use mix (Tsiompras, and Photis, 2017; Region of Waterloo, 2014; Frank et. al., 2010; Cervero and Kockelman, 1997). These data come from different sources but they are mainly available through open public sources, which makes them accessible. Other academics proposed to utilize audit and assessment tools to gather primary data (e.g., field observations, focus groups, travel diaries, and interviews) (Clifton et. al., 2007; Day et. al., 2006; Dannenberg, 2005; Moudon and Lee, 2003). In the context of this thesis, gathering primary data is beyond the available resources: monetary, training and staff, as well as requires time.

The most basic form of a walkability equation assumes equal influence of variables, while the more advanced formulas use weighted variables as argued for by Tsiompras, and Photis (2017). The same authors also discredited population density as a significant factor to consider in a walkability index and recommend it to be combined with land use mix assuming highly mixed land use areas are associated with high population density areas also. Tsiomprass and Photis (2017) adopt Manaugh's and Kreider's (2013) land use interaction method, which provides an alternative representation of land use mix rather than the conventional entropy equation (Frank et. al., 2010; Cervero and Kockelman, 1997). Indeed, various approaches are taken to measuring or representing all of the important explanatory variables.

The walkability index's strength lies in its focused approach to examining potential pedestrian activity. However, its flexibility makes it prone to issues of multicollinearity and complexity of the construct. As discussed above, there is no cap on the number of variables/factors that can be incorporated into the walkability index. With multiple variables, there is a higher risk of multicollinearity, which duplicates measured influences resulting in a skewed index. Cervero and Kockelman (1997) used factor analysis to avoid multicollinearity between variables. On the other hand, Moudon and Lee (2003) warn of underestimating the power of a single variable, when trying to include too many variables of the same category. On the other hand, including just built form explanatory variables (e.g., land use mix, sidewalk availability, network connectivity, and retail floor area ration) would only examine the potential for pedestrian demand rather than actual demand. Demand representative variables include, but are not limited to, population density, employment density, bus boarding and alighting, and elementary school enrollment.

Another concern with using a walkability index to estimate pedestrian demand relates to the various geographic units associated with the various data sources. It is important to select compatible data sources based on their resolution and an appropriate geographical unit area for the study that is relevant to the spatial scale of pedestrian movement. Most of the data types mentioned above are typically found at a fine scale (e.g., street-level and block-level), which means a high-resolution analysis can be conducted to define walkability scores at the street-level with little to none lost information.

The resulting walkability score does not represent a specific trip frequency, but rather a relative score indicating potential pedestrian activity. The walkability index is relatively easy to calculate and to understand, which explains its popularity in walkability studies and crossover to other fields (e.g., real estate).

On the other hand, only few scholars and practitioners explore the option to validate their or others' versions of walkability indices (Duncan et. al., 2011; Manaugh, and El-Geneidy, 2011; Frank et. al., 2010). Convergent validity is the similarity between the tool's results and other tools' results which theoretically should be similar (Web Centre for Social Research Methods, 2006). Validity could be established for a walkability index by comparing its walkability scores with pedestrian count or reported pedestrian activity (assuming the validity of reporting). This process has two positive outcomes. Through a sensitivity analysis, validation can be used to fine tune the weights for the walkability score. Also, through validation, walkability scores can be substituted with corresponding trip counts for easier representation of pedestrian activity.

2.3.3. Space Syntax Model

The Space Syntax approach is the least famous of all five approaches discussed here to estimate/predict pedestrian activity. It predicts pedestrian flow through the analysis of network

connectivity. Raford (2010) does a superb job in examining why Space Syntax has not been adopted to the same scale as other transportation forecasting models in North American planning. In contrast, Space Syntax is widely adopted and common in the United Kingdom and other European countries as well as taught in many abroad universities (Raford, 2010).

Raford (2010) identified a total of eight challenges, to adopt Space Syntax in North American planning, through interviews with experts on both sides. Two key challenges stand out that push Space Syntax out of North America. First, space syntax was launched in North America, a decade after other transportation forecasting models, such as the Four-Step Model (Raford, 2010). The delayed exposure allowed industry standards to formulate around the older models. Adoption of space syntax revokes and challenges these standards. The second issue is more technical. Space Syntax has its own unique language and terminology as well as it requires advanced statistical expertise to interpret model outcomes (Raford, 2010). The added technical complexity of space syntax in comparison to conventional models has limited its adoption in the North American context, also considering incompatibility with existing industry standards.

Space syntax measures the degree of connectivity whether at the regional level (whole) or at the street-level (local) (Li et. al., 2017; Penn, 2003). In other words, Space Syntax is a configuration analysis of the street network (Omar & Kaplan, 2017). Connectivity is evaluated based on integration and choice measures. Integration is a measurement of closeness of each road segment to all other segments, while choice is a measurement of wholeness by counting how many times a segment is along the shortest path between every pair of road segment (Li et. al., 2017; Omar & Kaplan, 2017; Hajrasouliha & Yin, 2015).

Space syntax's advantage lies in its strong correlations with flow (Omar & Kaplan, 2017; Raford, 2010). Penn (2003) found in his study area that Space syntax is about 52% representative

of pedestrian flow ($R^2=0.527$). When combined with other explanatory variables such as distance to transit stops and tourist destinations, the R^2 jumps to be as much as 0.81 (81%) (Raford & Ragland, 2006). Also, the outcome can be calibrated with a sample pedestrian flow to produce an actual pedestrian count for the rest of the study region (Raford and Ragland, 2006). Another advantage is the emphasis on connectivity as the sole variable to account for most activity within the city or region.

Despite its strong correlation with pedestrian activity, space syntax has disadvantages. North American planning experts point out technical issues such as new and unique terminology, the absence of commercial software packages, and requisite advanced statistical expertise to interpret data as key barriers to adopting Space Syntax (Raford, 2010). Other social issues include late exposure to space syntax, its rejection of North American industry standards, and the immense efforts and hustle needed to back its adoption (Raford, 2010). In addition, from a pedestrian analysis perspective, the space syntax literature rarely mentioned how it accounts for streets with one-sided sidewalk or without any sidewalk, which has a greater impact on pedestrian flow in North America versus in European context (Li et. al., 2017; Omer & Kaplan, 2017; Hajrasouliha & Yin, 2015; Raford, 2010; Raford and Ragland, 2006; Penn, 2003). Omar and Kaplan (2017) also point to how Space Syntax predictive powers are lower in planned urban areas (e.g., most North American Cities, and most suburban communities) versus in self-organized urban growth (e.g., historical urban cores and old cities).

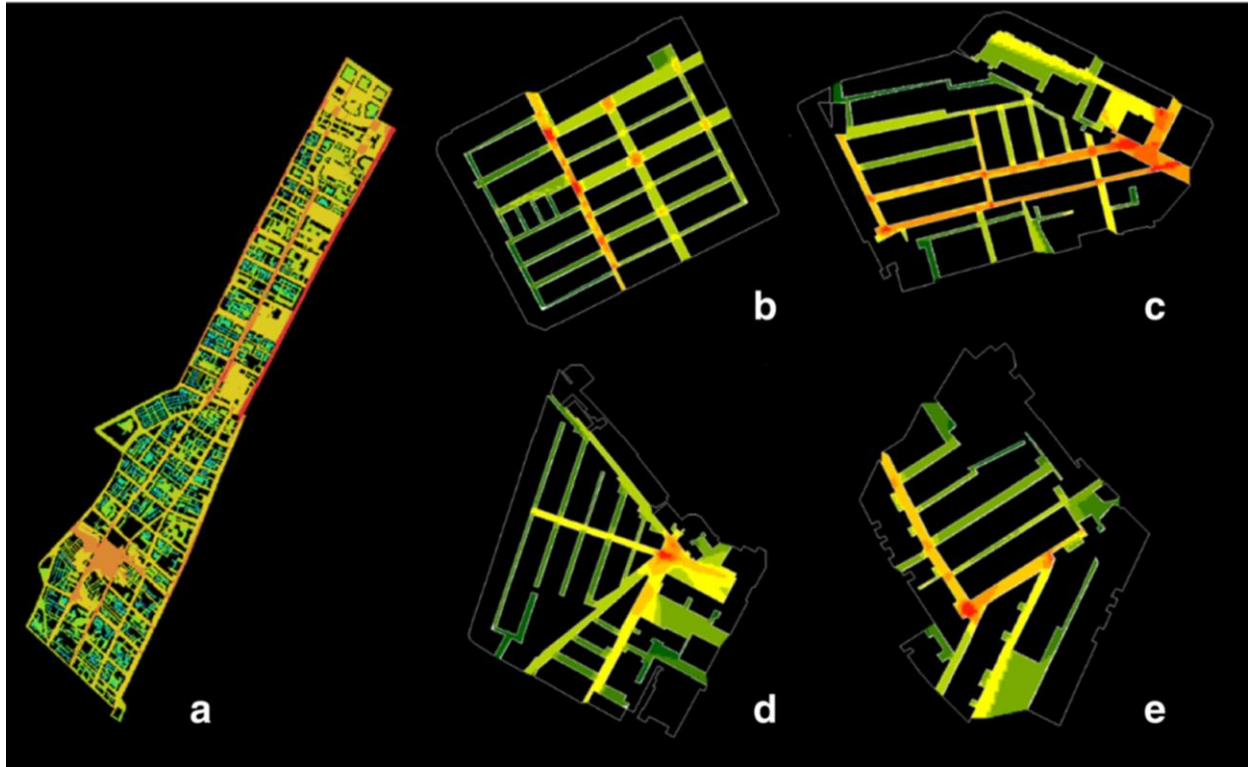


Figure 2.3.3.1: Example of Space Syntax analysis outcome (Source: Li, X., Lv, Z., Zheng, Z., Zhong, C., Hijazi, I., & Cheng, S. (2017). Assessment of lively street network based on geographic information system and space syntax.)

2.3.4. Gravity Model

The gravity-based measure is commonly used in accessibility models to determine ease of reaching destinations (Iacono et. al., 2010; El-Geneidy, and Levinson, 2006; Geurs and Van Wee, 2004; Rutherford, 1979). The gravity model is famous for being a part of the Four-Step Model and comes before the “Mode Split” step, as discussed previously. Gravity-based measures estimate trip distribution between origins and destinations based on their attractiveness and traffic impedances (e.g., travel time) (El-Geneidy, and Levinson, 2006; Luoma et. al., 1993; Rutherford, 1979). The model is based on land use and household survey data. Household data are usually found at the aggregate level (e.g., Census tract, or TAZ), which makes the conventional gravity model well suited to estimate motorized-based trips. Some researchers have

attempted to retrofit the model to examine pedestrian and cycling activity, but all agreed that there are challenges and limitations (Iacono et. al., 2010; Rutherford, 1979).

The gravity model is constructed similarly to Newton’s law of gravity as it is founded on the same concept: the attraction between two bodies is directly proportional to their mass (in this case, amount of attractions) (Rutherford, 1979). The gravity model trip distribution equation is highlighted in Figure 2.3.4.1. As discussed above, the model computes trip distribution using a friction factor, representation of attractions and traffic

$$T_{ijp} = P_{ip} A_{jp} F(t)_{ijp} / \sum_j A_{jp} F(t)_{ijp} \quad i, j = 1, 2, 3, \dots, n \quad (1)$$

where

- T_{ijp} = one-way trips from block i to block j for purpose p,
- P_{ip} = trips produced at block i for purpose p,
- A_{jp} = trips attracted to block j for purpose p, and
- $F(t)_{ijp}$ = friction factor based on the travel distance between block i and block j for purpose p (ordinarily travel time would be used but since the level of service for walking is nearly constant, it is easier computationally to substitute distance, which is then directly proportional to time).

Figure 2.3.4.1: Gravity Model Trip Distribution Equation (Rutherford, G. (1979). Use of the Gravity Model for Pedestrian Travel Distribution. Transportation Research Board. 728. 53-59.)

impedance (Rutherford, 1979). Accounting for the attractiveness to destinations is founded on Land-use data, which is usually coarse (Rutherford, 1979). In conventional auto-based models, the impedance factor is generally dependent on congestion levels and travel speed on road networks (Geurs and Van Wee, 2004). For non-motorized travel, the impedance factor is usually either travel time, distance, or cost (Iacono et. al., 2010; Rutherford, 1979).

The gravity model is limited by its scope and data sources in measuring non-motorized travel. The gravity model can use similar data sources as in the four-step model but it is then raising the same concerns. Eash (1999) as cited in Iacono et. al. (2010) points at how the model’s aggregated unit areas are poorly matched to the spatial scale of non-motorized movement. In other words, the geographical unit areas are too big to capture pedestrian trips between zones, which by default misses a considerable number of intrazonal trips. Rutherford (1979) used

pedestrian-specific surveys than only capture their movement. The advantage to these surveys is that data are collected at the disaggregate level and maintained at high-resolution to match the complex spatial movement of pedestrian activity. The downside of this approach is the dependency of the survey on available resources, such as training, surveyors, time limit, and monetary compensation. The alternative is publicly available aggregate data, which raises the issue of ecological fallacy when data are transferred to a disaggregate level.

2.3.5. Microsimulation/Agent-based Model

Agent-based modeling applications simulate ‘agents’ movement to replicate real-world pedestrian behavior to identify behavioral triggers and to assess an infrastructure’s level of service (e.g., sidewalk, hallway, intersection, or hall) (Chen, 2012; Torrens, 2012; Batty, 2001; Kerridge, Hine, & Wigan, 2001). In the literature, there are different ways to refer to agent-based modeling. It can be referred to as a multi-agent system (MAS), an agent-based simulation (ABS), or individual-based modeling (IBM) (Chen, 2012). Despite the varying definitions to what an ‘agent’ is, academics and researchers agree about two key defining properties: Autonomy, and Social ability (Chen, 2012). Autonomy is the agent’s ability to “carry out instructions and make decisions without direct interventions of others,” while social ability recognizes the community dynamics and how agents interact with each other in order to complete their task (Chen, 2012). Agent-based models (ABMs) application to the field of urban planning can help understand the impact of urban design on pedestrian flow, congestion, and social activity. Although ABMs can be applied at large scales like the city’s, its best-achieved predictions are at the local level (e.g., specific locations, building interiors, and intersections) (Kerridge, Hine, & Wigan, 2001).

Various agent-based models share the same concept of simulation a group of agents’ movements through a set of rules (Chen, 2012; Torrens, 2012; Batty, 2001; Kerridge, Hine, & Wigan, 2001).

A different set of rules lead to different discourses and versions of microsimulation models. The earliest form of ABM's rules were based on the fluid-flow law to guide agent movement (Torrens, 2012). More modern agent-based models like the PEDFLOW, divert away from fluid-flow rules towards incorporating rules that better mimics real-world pedestrian behaviour, such as static awareness, personal space measure, preferred walking speed ...etc. (Torrens, 2012; Kerridge, Hine, & Wigan, 2001). It is hard to discuss the construct of agent-based modeling as there are many discourses with varying contrasts so only the founding concept and area of differences are shared here.

The key advantage of using ABMs is the ability to conduct fine-scale analysis at the individual level. Also, agent-based modeling ability to measure the infrastructure's level of service capacity (Torrens, 2012).

On the other hand, ABMs are not intended to predict or estimate actual pedestrian demand, which is the intended outcome for this thesis. Level of service is more like potential sidewalk capacity assuming maximum flow, which is like assuming that the level of service is fully representative of actual pedestrian presence. Also, the ABMs' individual-analysis-level of detail is not necessary for this thesis as it will be discussed later how an aggregate level detail at the neighbourhood level is more sufficient for this study.

Table 2.3.5: Summary of Pros and Cons for Each of the Five Approaches for Estimating Pedestrian Activity

Approach for Estimating Pedestrian Activity	Pros	Cons
Mode Choice (4SM)	Industry standard, and easy to adopt and apply	Aggregate Datasets, and Large geographical units of analysis
Walkability Index	Simple construct and easy to interpret	Indicator value and lack of validation
Space Syntax Model	Focuses on connectivity, and strong correlation to pedestrian flow	Required extensive statistical knowledge and unfamiliar technical language
Gravity Model	Industry standard, and can stand on its own	Aggregate Datasets, and Large geographical units of analysis
Microsimulation	Street-level accuracy	Complex pedestrian behaviour and focus on Level of Service

Chapter 3: Methods

This thesis analysis bridges the fields of active transportation planning and winter maintenance of surface transportation systems. It focuses specifically on pedestrian presence and infrastructure, with the overall goal of contributing to the provision of an accessible and reliable sidewalk system during winter weather.

This chapter is organized into three main sections. The first provides an overview of the approach taken in the thesis. The second section introduces the study area, which is the City of Waterloo in southern Ontario. The third section elaborates on the research design, which includes the geographical unit of analysis for the study, the set of variables chosen for the construction of the Pedestrian Activity Model, and the steps taken to construct the model, like conducting various regression analyses, and lastly identifying the formula for the model. Lastly, the chapter explains briefly the approach taken to define a Pedestrian Priority Zone, where pedestrian-related public service (e.g., sidewalk snow clearing) should be prioritized, using the newly constructed model.

3.1. Overview of the Approach

Five approaches to estimating/predicting pedestrian activity were reviewed in the literature review chapter (i.e. Four-step Model, Walkability Index, Gravity Model, Space Syntax Model, and Agent-based Model). Each approach had its limitation and critique, in addition to also having advantages. The Four-step model (includes both the Mode Choice and the Gravity model approaches) is well established in traffic demand management and transportation planning; however, it is an aggregate model that was designed for an automobile dominant traffic. On the

other hand, Space Syntax model and Agent-based model are more modern and detail-oriented. The disadvantage to using Space Syntax modeling is its sole dependence on connectivity to explain pedestrian activity, while the disadvantage to using agent-based modeling is the complexity of replicating pedestrian behaviour and focus on measuring the infrastructure's level of service capacity. This leaves us off with one model to consider: Walkability index. The criticism for walkability index is its over-simplified formula, measurement of potential activity not actual plus its usual inclusion of just primarily built environment variables, and the lack of validation.

With the lack of validation in mind as well as the need for demand-representative variables, a new model is proposed. The new model is called "Pedestrian Activity Model" (P.A.M.). P.A.M. is inspired by the use of regression as a calibration and a validation tool and is also inclusive of both built environment variables, like that in walkability indices, and demand representative variables (e.g., population density, transit users). A multi-variable regression model would offer a better representation of reported pedestrian activity. In addition, the use of regression would calibrate the model to predict actual walking trips, rather than an indicator of pedestrian activity.

3.2. Study Location

For this thesis, I selected the City of Waterloo for the case study location. The City of Waterloo is located north of the City of Kitchener and, together with the City of Cambridge and four townships, comprises the Region of Waterloo (see Figure 3.2.1). Despite not being part of the Greater Toronto Area, the City of Waterloo is part of the Greater Golden Horseshoe as indicated in Ontario's Growth Plan (2017). The City is located 95 kilometers away from the City of Mississauga and 125 kilometers from the City of Toronto downtown core.

According to the 2016 census, the City of Waterloo's population is about 105,000 (Statistics Canada, 2017). According to the Transit Tomorrow Survey (TTS) for 2016, there are over 65,000 jobs in the City (Data Management Group, 2018a). The 2016 census shows that 22,000 of the city's residents commute to jobs within the city (Statistics Canada, 2017) and about 40,000 workers and employees commute in from outside. One-third of those who live and work within Waterloo hold an occupation in educational services (Statistics Canada, 2017).

Waterloo is home to three post-secondary institutions: University of Waterloo, Wilfrid Laurier University, and Conestoga College (see Figure). 32,000 full-time university students live in Waterloo, and another 10,000 commutes in for post-secondary education (Waterloo Town and Gown Committee, 2017). Another 3200 students attend classes in Conestoga College Waterloo Campus (Hicks, 2016). Together there are over 45,200 full-time post-secondary students, who roam the city from September to April of each year.

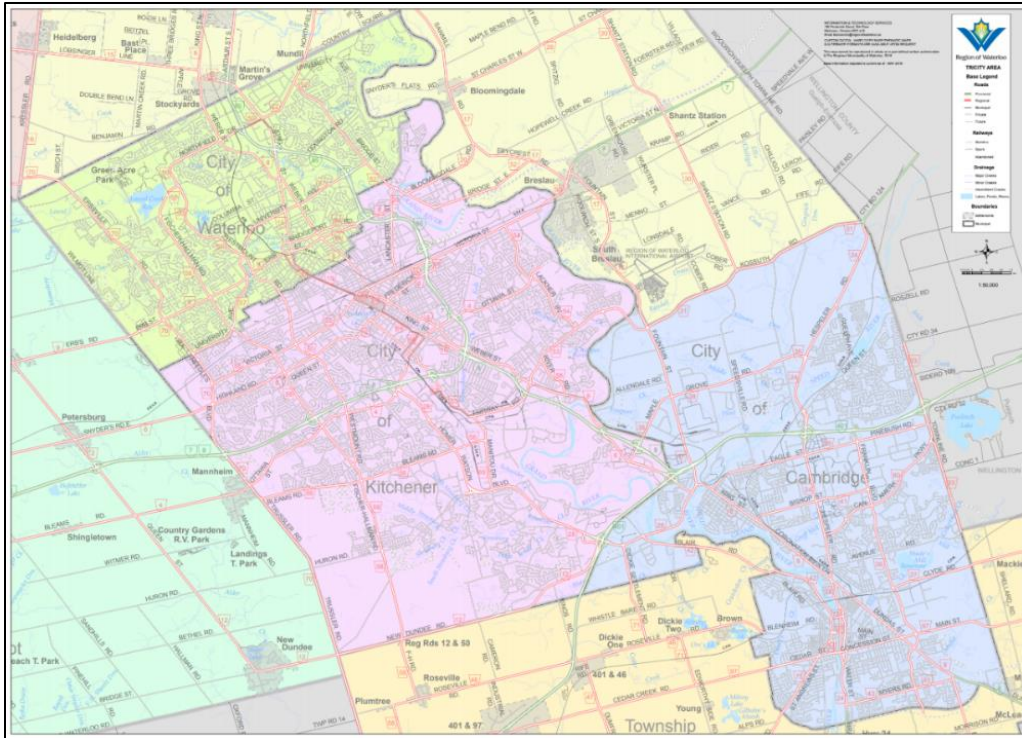


Figure 3.2.1: Tricity map (source: Region of Waterloo website, base maps)

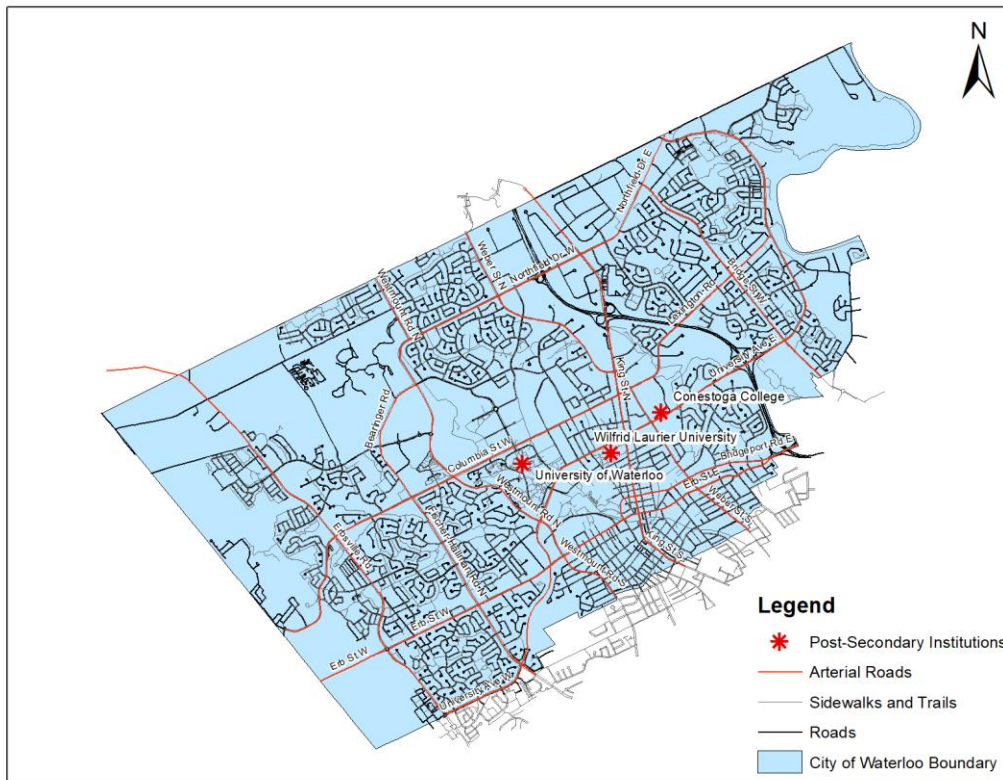


Figure 3.2.2: Post-secondary Institutions' locations, City of Waterloo (Source: City of Waterloo, Open Data Portal)

The City of Waterloo is predominantly a car-dependent environment with 83% of daily trips by household members being made by automobiles (either as a driver or a passenger) (see Figure 3.2.3) (Data Management Group, 2018a). Walking trips as reported by the 2016 TTS make up the second largest share of daily trips, accounting for 8%, followed by 6% public transit

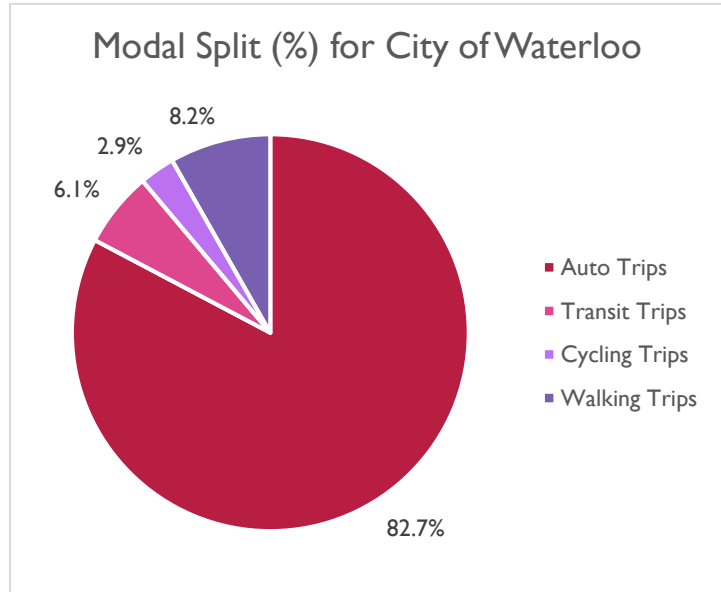


Figure 3.2.3: Model split of daily trips, City of Waterloo (Source: 2016 TTS)

trips. The remaining 3% is for cycling trips. The TTS only captures the primary mode of transportation for trips but, in reality, every trip includes a walking trip component. Public transportation users walk to and from stations and bus stops. Drivers and passengers walk to and from their cars whether in parking lots or on the street. Shopping trips are walking trips as customers walk from one store to another. Walking is common as a secondary mode of transportation, yet it is seldom reported. In addition to unrecognized walking trips, surveys, like the Transit Tomorrow Survey (TTS) and the Household Survey Census, focus on private households, in doing so, they under-count activity made by seniors and families living in public housing, and students living on-residence (Data Management Group, 2018b). The same document also reports under-representation for areas clustered with rental apartment buildings, as owners are unaware of their tenants' activity (Data Management Group, 2018b). On-campus residence for the both University of Waterloo and Wilfrid Laurier University provide approximately 25% of beds available to students in the City of Waterloo, while the majority of

the beds are located in multi-residential buildings (Waterloo Town and Gown Committee, 2017). Considering the limitations of the TTS survey and that the City of Waterloo is a university city, post-secondary student transportation activity is under-reported. The TTS would be better at explaining activity away from post-secondary student clusters.

3.3. Research design

Constructing a Pedestrian Activity Model

In the literature review chapter, five possible approaches were presented for estimating pedestrian activity. For this thesis, I chose a regression-based model to predict pedestrian activity spatially. The first sub-section is about choosing a suitable geographical unit of analysis for this study. The following sub-section examines walking-related built environment and demand representative variables examined in the regression analysis. The variables are examined in terms of commonality in previous studies, the data source and form, and the variable's measure and descriptive statistics. These variables were identified and chosen through an extensive review of the literature and considerations of data availability. Third sub-section explains the process of using a regression analysis to calibrate the Pedestrian Activity Model. Lastly, the final model formula's construct is revealed.

3.3.1. Geographical Unit of Analysis

Researchers and academics argue that traditional geographical units of analysis (e.g., TAZ) are not suitable to capture pedestrian activity because these areas were constructed to capture motorized forms of transportation in travel forecast models (Clifton et. al., 2016; Eash, 1999 as cited in Iacono et. al., 2010). An alternative, presented by Clifton et. al. (2016), was to cover the study area with grid cells, which had a pre-defined length such as 80m for a 1-minute walk. A counterpoint to using grid cells is their misalignment with the street network and constant overlap with buildings' footprint. This creates a challenge in calculating some of the variables such as Commercial Floor Area Ratio, and sidewalk network connectivity. Another challenge with using small grid cells as suggested by Clifton would be the immense required computational

power. A simple calculation to find out the number of cells in the City of Waterloo reveals there will be about 10 thousand cells in the overall grid, if the grid cell is 80m x 80m.

As a result, I re-consider conventional geographical units of analysis, for which I have 3 alternatives as highlighted in Table 3.3.1.1. The biggest unit area option is Traffic Analysis Zones, which has a mean area of 62.4 hectares and consists of 104 areas within the City of Waterloo and contains 3 areas that overlap with surrounding municipalities. All variables mentioned above are available at the disaggregate level, which makes using Traffic Analysis Zones optimal if considering all variables in the final model. In considering TAZ suitability to capture pedestrian movement, average TAZ area is a key decisive factor. A 5-minute walk is about 400m using 80m to represent a 1-minute walk standard (Clifton et. al., 2016). A 400m grid cell is 16 hectares in area. The average size TAZ is about 4 times the size of a 400m grid cell, which raises concern for TAZ's inability to capture intra-zonal pedestrian movement.

The second alternative is dissemination areas (DA). According to Statistics Canada (2015), a dissemination area is a "small area composed of one or more neighbouring dissemination blocks, with a population of 400 to 700 persons." DA is the smallest geographical area for all census data (Statistics Canada, 2015). With DAs, it is still possible to consider all variables, since they are either available at the DA level (i.e., Population and Employment densities) or at a finer scale (e.g., Transit Activity, Sidewalk Connectivity.). The average Dissemination area size is 42.5 hectares (Table 3.3.1.1), which is about 2.5 times the size of a 400m grid cell (16 hectares per cell). The concern remains for pedestrian flow misrepresentation with using big size geographical unit of analysis.

The last geographical unit of analysis to be considered is PLUM zones (PZ). The City of Waterloo has 336 PLUM zones plus 1 zone that overlaps with the City of Kitchener. The pros of using PLUM zones are their small average size (19.4 hectares), which is similar to that of a grid cell representing a 5-minute walk (16 hectares). PLUM zones were created in respect to municipal boundaries, traffic analysis zones, census tracts, and water and sewer service areas, which also aligns with roads and property lines, reducing the need to disaggregate data (GIS Region of Waterloo, n.d.). On the hand, the disadvantage to using PLUM zones is that not all variables are available at that fine scale. Variables such as population and employment densities are available at the dissemination area level. Incorporating aggregate-level variables would create a modifiable areal unit problem (MAUP). As discussed in Appendix A, both population and employment density variables have a strong correlation with other variables, such as transit and land use mix, which is considered for this study. Since transit and land use mix variables are considered, population and employment densities were no longer incorporated, to issues pertaining to multicollinearity and MAUP. By default, PLUM zones are adopted as the study's geographical unit of analysis.

Table 3.3.1.1: PLUM zones (PZ) versus Dissemination Areas (DA) versus Traffic Analysis Zones (TAZ) comparison

Category	PLUM Zones (PZ)	Dissemination Areas (DA)	Traffic Analysis Zones (TAZ)
Count	336 (conflict: 1)*	153	104 (conflicts= 3)*
Min. (ha)	0.32	6.19	4
Max. (ha)	233.33	532.17	368
Mean (ha)	19.4	42.5	62.41

*conflict areas where not counted and omitted due to overlap with adjacent municipalities

3.3.2. Built Environment and Non-built Environment Variables

Prior to exploring each variable, there are basic standards adopted throughout the analysis process. First, the baseline for the data is 2016 as it was a common year for updates across most of all data sources (e.g., census, and TTS). Second, going into the regression analysis, variables in the raw format were normalized to a percentage scale (0-100) using the following formula:

$$= \frac{X}{X \text{ max.}} \times 100$$

Land Use Mix – Entropy

Land use mix is a measure of diversity and it is the variation of land use types within geographical areas. The logic here is that with adequate land use mix, people are more motivated to walk for their everyday needs. In this thesis, I present two different measures for land use mix. The first is entropy and the second is interaction lines. Entropy equations have been commonly used in walkability studies (Ewing and Cervero, 2010; Frank et. al., 2010; Meghelal and Capp, 2011). Entropy measure considers both area per land use and the number of existing uses within each PLUM zone. A neighbourhood with just two big land use types will score less in the entropy equation versus a neighbourhood similar in size with five different land uses. From a review of previous studies, there was no standard to which land use types should be tested. In this study, I chose to evaluate the land use mix based on five land use types: residential, commercial, employment, institutional, and parks. Protected green lands and undeveloped lands are classified as open space and are not considered in the evaluated land use mix as suggested by Manaugh and Kreider (2013).

Land use data is available for the City of Waterloo based on their Official Plan. Data were retrieved from the Geospatial Centre at the University of Waterloo. Land use data is available as a polygon shapefile. The data attributes included polygons' area and land use types. In case a site contains multiple uses, the minor uses are reclassified under the site's primary use. Further data manipulation is required to isolate parks from open space zoning using parkland location and shapefile from the Region of Waterloo, which is also available through the Geospatial Centre.

The entropy equation used here is the same one showcased in the study by Brown et. al. (2009) about the mixed land use and walkability. The equation below has been adjusted to reflect the number of considered land use types and geographical unit of analysis (i.e., PLUM zones) for this study. The entropy equation is as follows:

$$Entropy = \frac{-A}{\ln(N)}$$

N – Total number of land uses considered in the analysis. In this research context, N = 5

$$A = \left(\frac{b1}{a} \times \ln \frac{b1}{a}\right) + \left(\frac{b2}{a} \times \ln \frac{b2}{a}\right) + \left(\frac{b3}{a} \times \ln \frac{b3}{a}\right) + \left(\frac{b4}{a} \times \ln \frac{b4}{a}\right) + \left(\frac{b5}{a} \times \ln \frac{b5}{a}\right)$$

Note: The terms in equation A is dependent on the number of available land use types in each geographical area of analysis

b1 – Residential Area

b2 – Commercial Area

b3 – Employment Area

b4 – Institutional Area

b5 – Parks Area

a – Sum of areas of all available land uses within each a PLUM zone

The entropy equation is a two-part equation and is dependent on both land use type availability and size. The raw results range from 0 to 1, with 0 indicating homogeneity and 1 indicating heterogeneity. The entropy equation structure is too complex to write on the GIS software (i.e., ArcMap). As a result, land use data were extracted then imported into Microsoft Excel for processing.

Figure 3.3.2.1 shows the descriptive statistics for the entropy measure. The land use mix mean is 0.2 and the standard deviation is 0.13. The histogram shows that the distribution of land use mix is skewed towards the right side of the mean. The maximum entropy score calculated was 0.67, while the lowest was 0 representing no presence of any of the 5 land use types. There are more zones with lower than average land use mix compared to above average zones. According to Figure 3.3.2.2, the above average zones are scattered across the city, with no clear spatial pattern. A common criticism of entropy is its limitation to only measuring land use diversity from the perspective of just land use types and their size magnitude, with no regard to how well mixed are they within each analysis area (Manaugh and Kreider, 2013; Tsiompras and Photis, 2016).

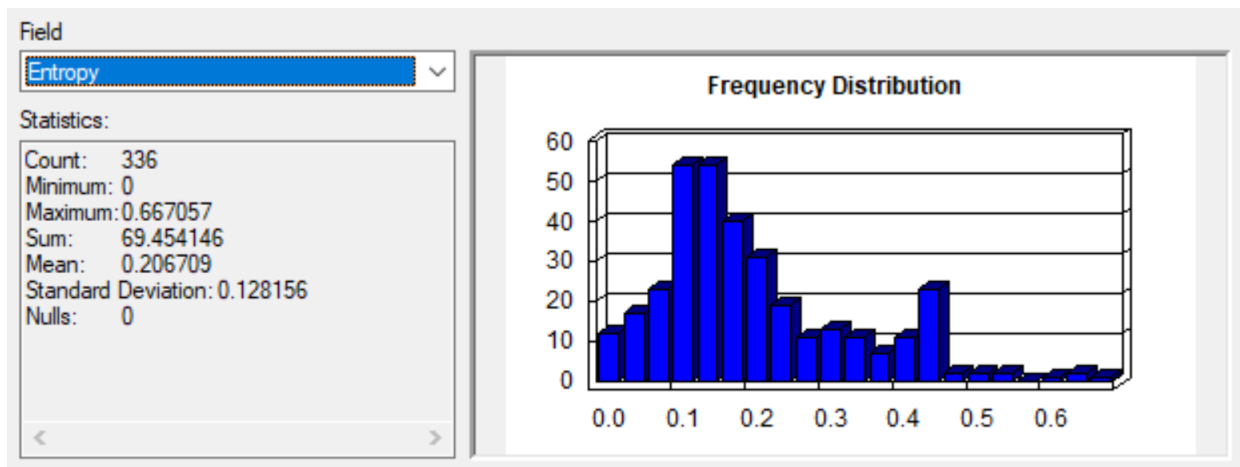


Figure 3.3.2.1: Land Use Entropy descriptive statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

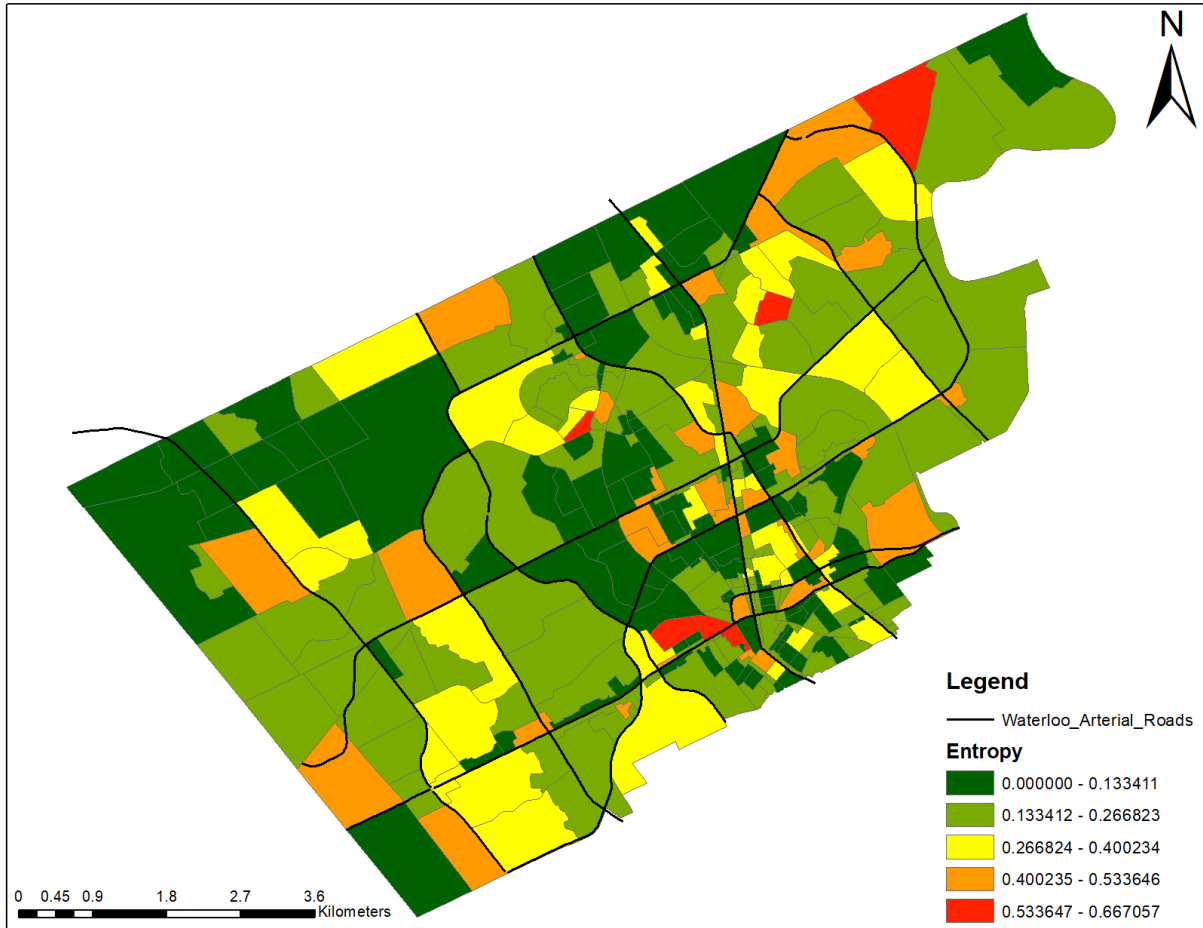


Figure 3.3.2.2: Land Use - Entropy's equal interval map by PLUM zones

Land Use – Interaction Lines

Interaction lines is another measure to evaluate the degree of variation among land uses. This method was proposed by Manaugh and Kreider (2013) in their article “What is mixed use?...” Interaction lines measure the distribution of land use mix by measuring the shared line between every two varying land use types. In other words, this measure accounts for the local land use mix. For example, all three figures in Figure 3.3.2.3 would score identically in entropy, despite variation in land use distribution and therefore diversity. A person who lives in the figure to the most left would walk further to reach the commercial area versus someone who lives in the figure to the most right.

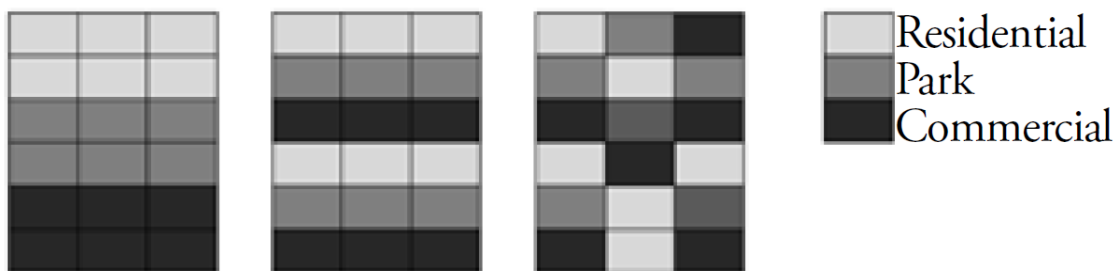


Figure 3.3.2.3: All three figures will have identical entropy score, while scoring differently on the interaction lines measure. The most-right figure would score the highest on interaction lines measure.

As discussed in the Entropy measure, land use data is a polygon feature shapefile. Land use data were converted to lines, using the “Polygon to Lines” tool in ArcMap. The new shapefile attributes include data such as land use types on the right and the left side of the interaction line. Using “Selection by Attributes”, the shared lines between every two varying land use types were selected and isolated. All shared lines are considered to have the same weight, without any preference for land use combinations. Since analysis areas vary in size from big around city skirts to small about the city core, a ratio of interaction lines to the area size is calculated. The

area is not based on the total PLUM zone size but is rather based on the sum of areas for all present five land uses (e.g., residential, commercial, employment, institutional, and parks). The interaction lines ratio is representative of the local land use mix per each PLUM zone.

The histogram below is heavily skewed towards the right with over 150 zones equal to zero. The maximum length on interaction lines per hectare is 270 m. High mixed-used distribution is predominately found in small PLUM zones as shown in Figure 3.3.2.5. Smaller PLUM zones tend to have a higher ratio due to its size. On the other hand, the shortest interaction line is 0 m. This can be explained by either single land use occupant of the zone or absence of any of the five land use types from the zone. Overall, there seem to be more zones with greater local land use mix distribution in the eastern half of the City (see Figure 3.3.2.5). Theoretically, Interaction lines along with Entropy should provide a balanced look at the impact of land use mix on walking activity.

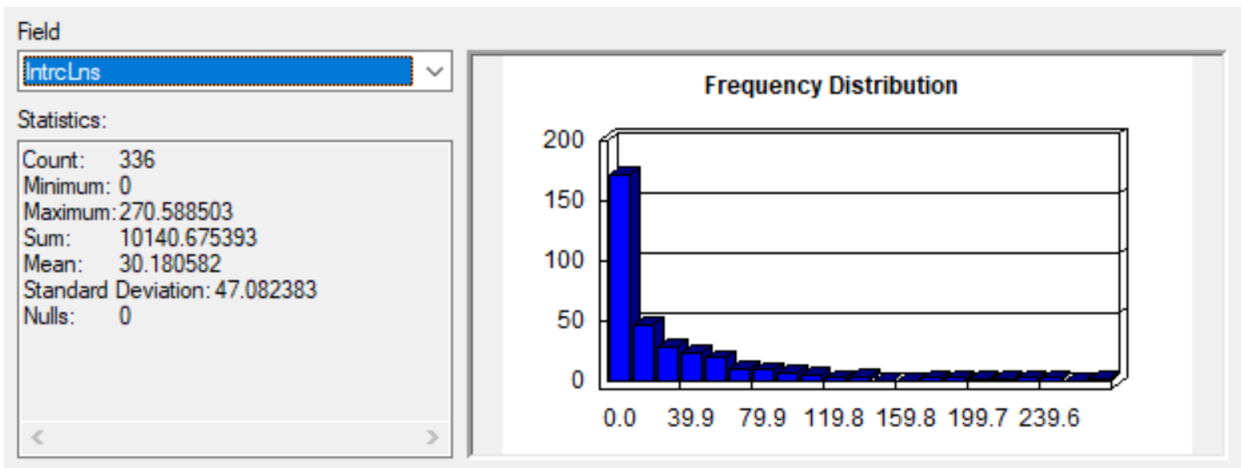


Figure 3.3.2.4: Land Use – Interaction Lines descriptive statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

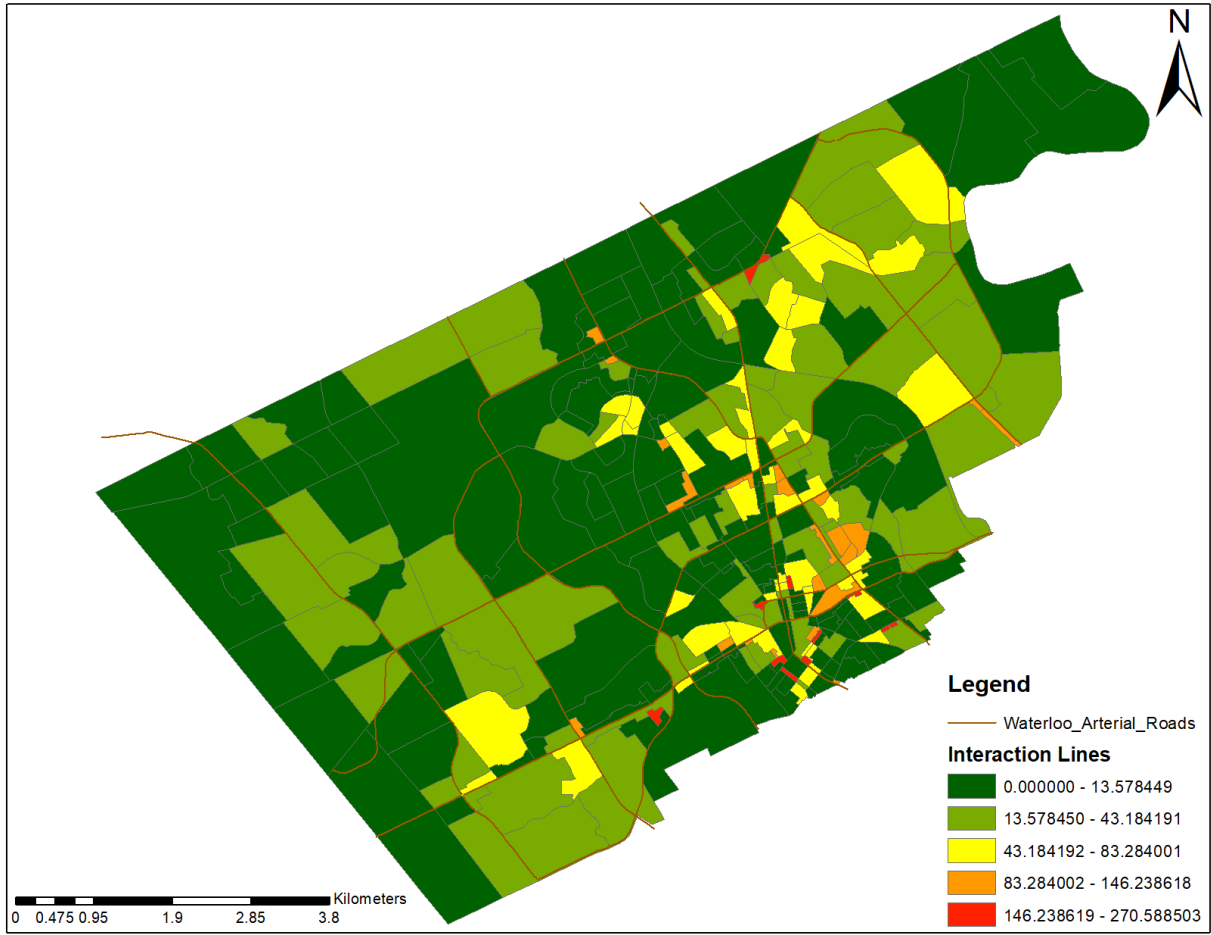


Figure 3.3.2.5: Land Use – Interaction lines natural breaks map by PLUM zones

Net Commercial Floor Area Ratio (FAR)

This measure is a modified edition of the conventional retail floor area ratio. Many academics and researchers found a correlation between walking and retail floor area ratio and incorporated it into walkability indices (Sundquist et. al., 2011; Frank et. al., 2010; Region of Waterloo, 2009; Leslie et. al., 2007). The floor area ratio (FAR) is structured to compute how much of a parcel is covered by the designated commercial/retail building. A pedestrian-oriented community will have a high floor area ratio indicating less space for cars and more space for walking activity. On the other hand, a car-oriented community will have a low floor area ratio due to large surface parking spaces.

For the City of Waterloo, the closest to retail-representative data is commercial zoning data, which includes but not limited to retail, and services (e.g., dry cleaners, and mechanics) (City of Waterloo, 2016). Data is available as polygon features for the City of Waterloo, representing zoning as in the Official Plan (OP), with attributes such as land use type and shape area size. Building footprints were available for the Region of Waterloo also as polygon features. Building footprint data did not have an attribute indicating its commercial use or not; therefore, building footprints overlapping OP's identified commercial zones were classified as commercial. The assumption used here is OP commercially designated lands capture all on-ground commercial uses.

The floor area ratio equation is a universal equation. It is the ratio of a building footprint to its parcel. The heaviest burden of computing the commercial floor area ratio is the data quality for property parcels. Another data quality challenge was when PLUM zone boundaries split big parcels into smaller portions as well as building footprints. Each overlapping case was visually

inspected and the building's area was allocated to the zone where the biggest portion of the building stands. In the case of a PLUM zone contained multiple commercial parcels, the mean was adopted as the final FAR value. Since PLUM zones have varying sizes, Floor Area Ratios (FAR) were weighted by the commercial area percentage of the total PLUM zone size. The outcome is classified as “net Commercial Floor Area Ratio.”

According to Figure 3.3.2.6, the distribution of net Commercial Floor Area Ratio is extremely skewed to the right. The FAR mean is 0.03, which is less than a tenth of the scale's maximum (0.33). This indicates the domination of car-oriented commercial space in the City of Waterloo. From Figure 3.3.2.7, I observe a spatial concentration of pedestrian-oriented commercial space (medium-high FAR) along the southern half of King St. by the city's core.

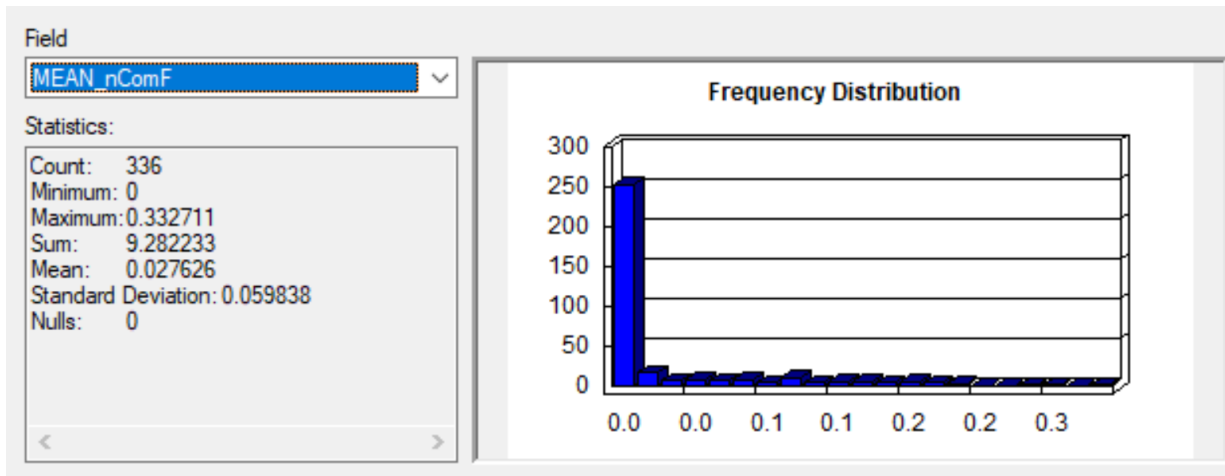


Figure 3.3.2.6: Net Commercial Floor Area Ratio descriptive statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

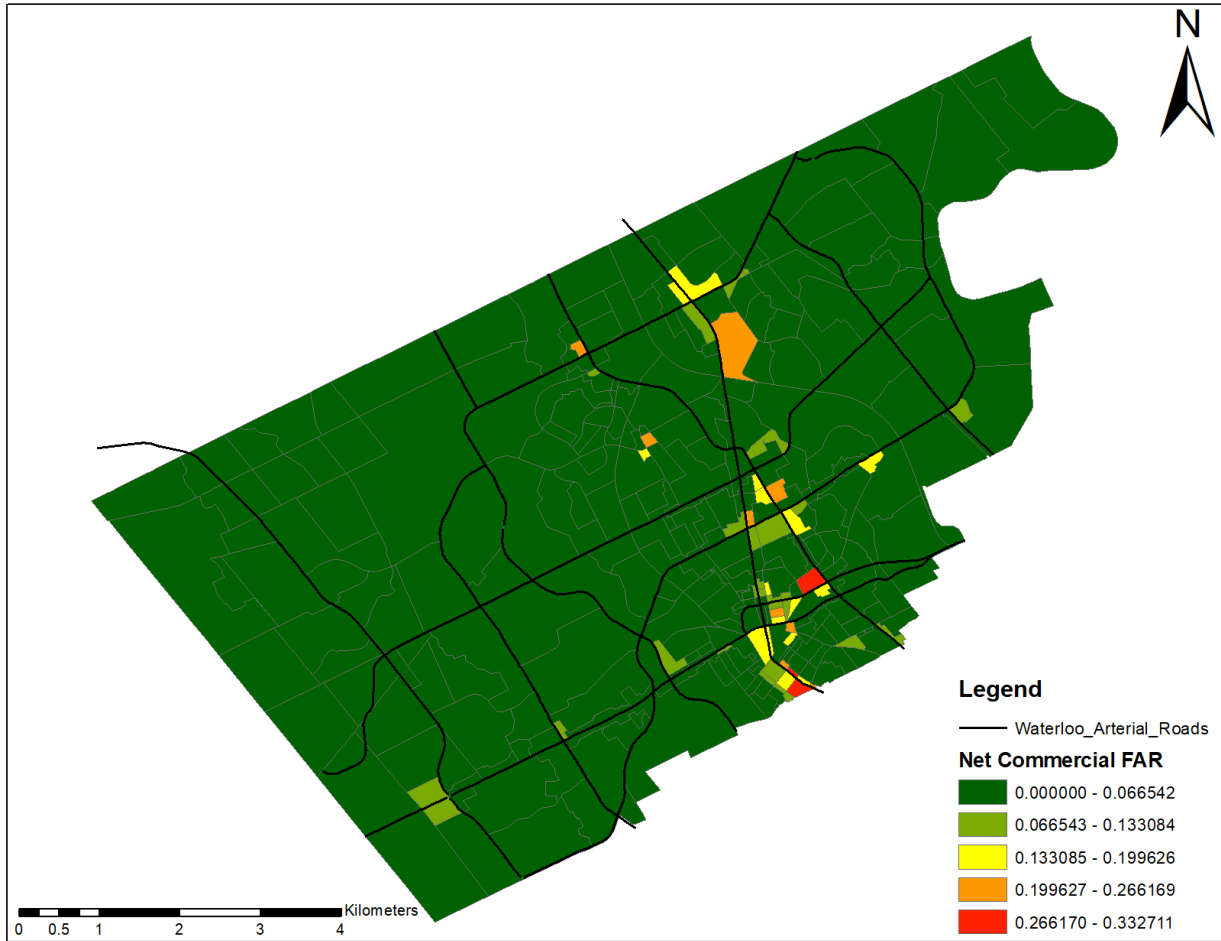


Figure 3.3.2.7: Net Commercial Floor Area Ratio equal interval map by PLUM zones

Connectivity – Metric Reach

Metric reach is an alternative to using intersection density for measuring network connectivity and coverage (see Appendix A for the Intersection Density measure). A recent study by Ellis et. al. (2016) looked at various measures for connectivity and found a significant correlation between walking and both intersection density and metric reach measures. Metric reach is the sum of sidewalk length in every direction if a person walked from the zone's centroid on the sidewalk in any direction for a set distance. The set distance in this study is 800 meters, while the industry's standard for a 10-minute walk is 800 m - 1000 m (Ellis et. al., 2016; Tsiompras and Photis, 2016; Frank et. al., 2010; Lee and Moudon, 2006). A 200-meter search tolerance was used to find the closest point to the zone's centroid along the sidewalk network. When adding the 200m search tolerance to the 800m maximum walking cut-off distance, the total does not exceed the 1 km threshold, which remains within the 10-minute walkable distance standard.

The active transportation infrastructure data is available through the Region of Waterloo and its municipalities' open data portal. The infrastructure data is a line feature making up a network of sidewalks, trails and pedestrian crossings. Trails were included as they often improve connectivity in neighbourhoods and not just provide recreational space. Despite data availability, there were multiple data layers and none contained the full network for the City of Waterloo. When trying the spatial join tool to merge all incomplete network layers, there were multiple cases of duplicate segments. The overlapping segments were not evenly spread nor consistent, which made it inapplicable to identify and eliminate. The remaining solution was to manually edit and trace missing links from other data sources into a single network layer. In addition, aerial photos and google maps were used to validate the various network links presence. Another challenge was deciding on road crossings locations particularly involving grass footpaths. Visual

clues, such as painted crossings, corner street’s concert pad orientation, and beaten crossings through grass boulevards, were used to determine road crossing locations. On the other hand, when it came to longer beaten pathways and trails through grass, they were not included as they are not official pedestrian infrastructure and are not subject to snow plowing. Also, despite how recent aerial photos are (two years old), grass and nature can reclaim unpaved pathways if unused.

According to the descriptive statistics below and measure’s histogram (Figure 3.3.2.8), there are over 30 PLUM zones that have a metric reach score of zero. Otherwise, the distribution of PLUM zones by metric reach is slightly skewed to the right towards the mean (15.18 km). The standard deviation is about 9.37 km. According to Figure 3.3.2.9, those zones with moderate-high metric reach score (23-39km) are clustered around Waterloo’s Uptown, especially along King St. and Erb St. In addition, there are some moderate Metric Reach zones scattered around the city peripheries, which can grow with further land development and network integration.

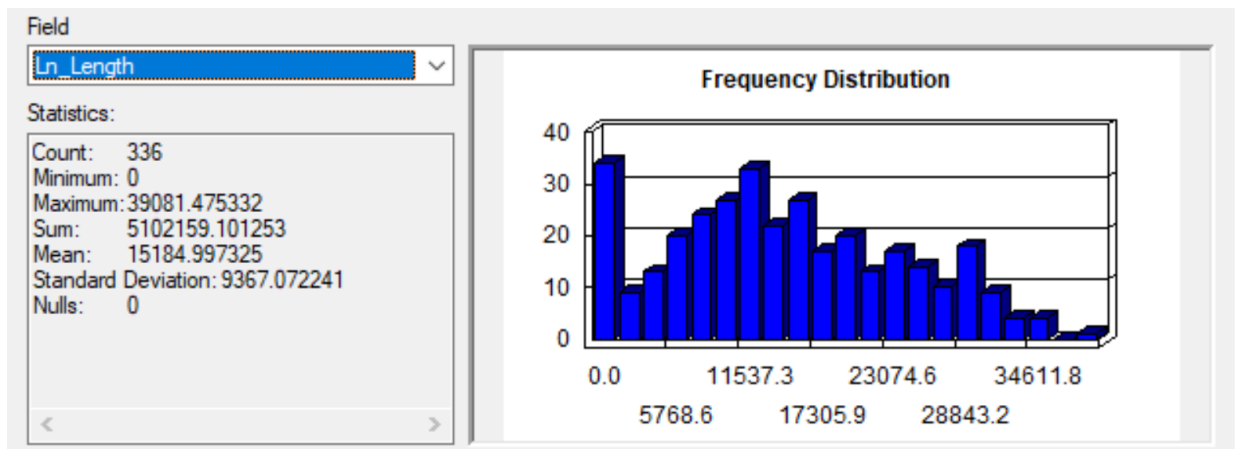


Figure 3.3.2.8: Connectivity – Metric Reach descriptive statistics (X-axis: variable’s value and Y-axis: PLUM zones’ count)

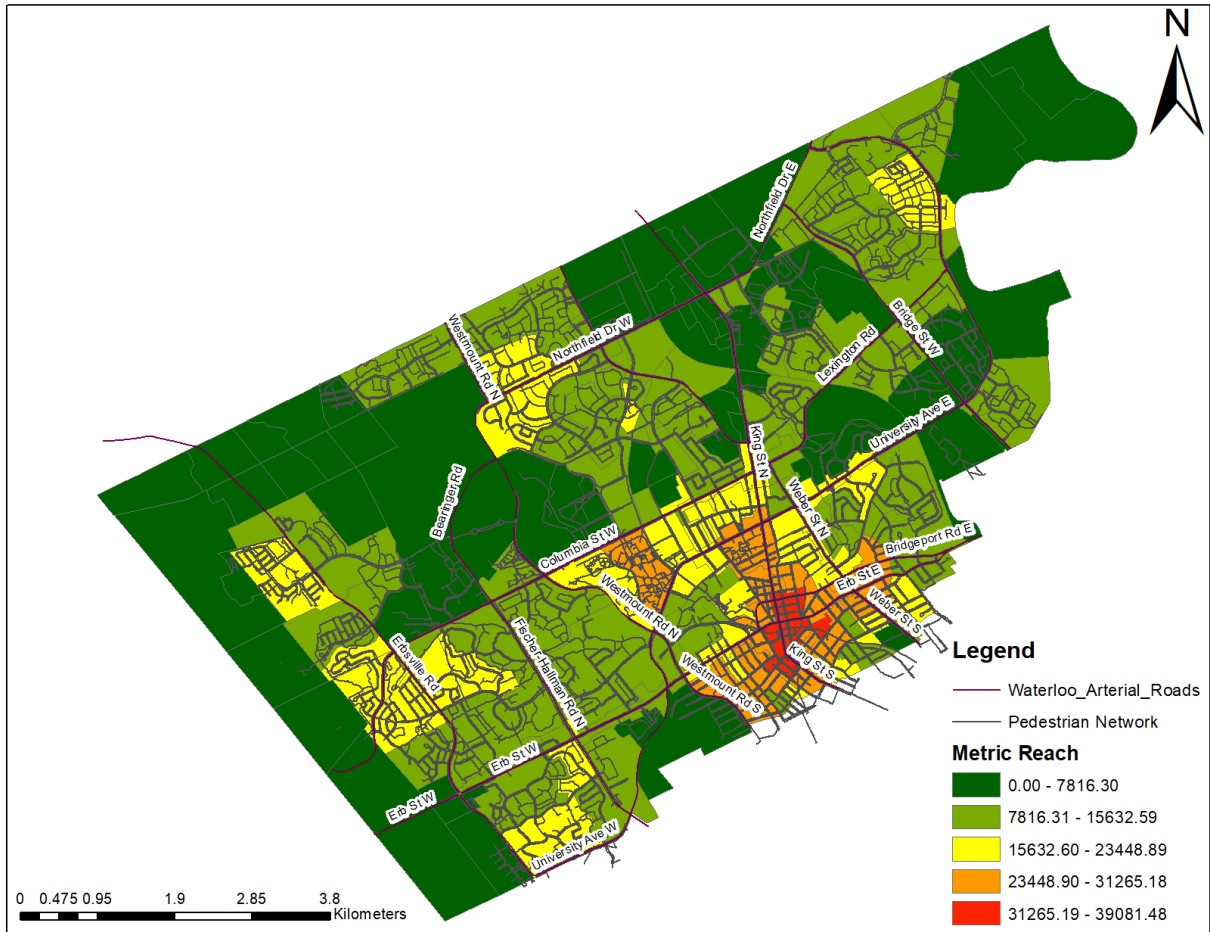


Figure 3.3.2.9: Connectivity – Metric Reach equal interval map by PLUM zones

Transit Activity

It was stressed that both population and employment density correlate to transit activity, which as well correlates to walking (see Appendix A) (Liu and Griswold, 2009; Ewing and Cervero, 2010; Tsiompras and Photis, 2016). Tsiompras and Photis (2016) found that proximity to transit would account for up to 20% of walkability's weight because of its multicollinearity with simulating pedestrian activity destinations, and population and employment density. In addition, according to the 2016 TTS, 50% of annual transit trips by the City of Waterloo permanent residents were accessed by walking. This further stress the importance of proximity to transit in the context of evaluating pedestrian activity.

Transit data were requested from the Region of Waterloo because the detailed 2016 data were not publicly available on the open data portal. The data acquired included point data for transit stops and attributes such as boarding and alighting activity per stop. Transit stops/stations are treated here as destinations and activity per stop was based on the sum of alighting and boarding to represent transit users passing through each stop. To account for accessibility to bus stops, buffer areas with a 400 m radius were built around each PLUM zone. Since all public transit in the City of Waterloo is yet serviced by buses only, 400 m buffer is the industry standard for typical walking distance to bus stops or stations (Oliver, 2014). Later, all activity within a PLUM zone and its 400m buffer zone were accounted towards that zone. Including the buffer zones ensure that intra-zonal walking trips to bus stops are captured.

According to Figure 3.3.2.10 below, over 100 out of 336 PLUM zones have almost zero transit activity. The Transit Activity histogram is extremely skewed towards the right. The highest activity per PLUM zone is about 23 thousand daily trips, while the average is 3.85 thousand

transit users. The standard deviation is approximately 4.5 thousand users. Figure 3.3.2.11 shows that post-secondary institutions are the key driver of transit activity as zones of moderate-high activity are found along University St. between the three institutions. There is a lighter activity presence extending along King St. south towards Uptown as well as by Conestoga Mall, which both serves as transit hubs.

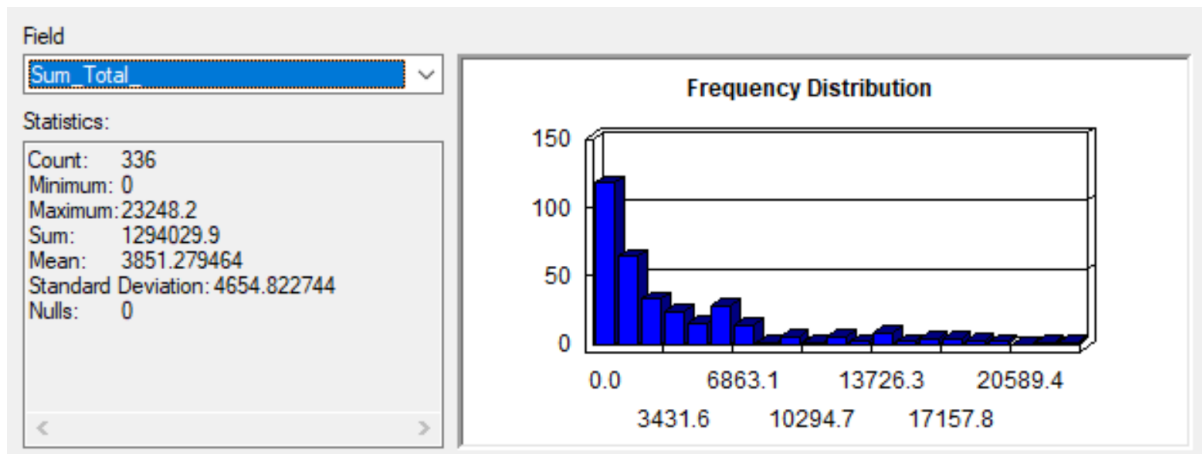


Figure 3.3.2.10: Transit Activity descriptive statistics(X-axis: variable's value and Y-axis: PLUM zones' count)

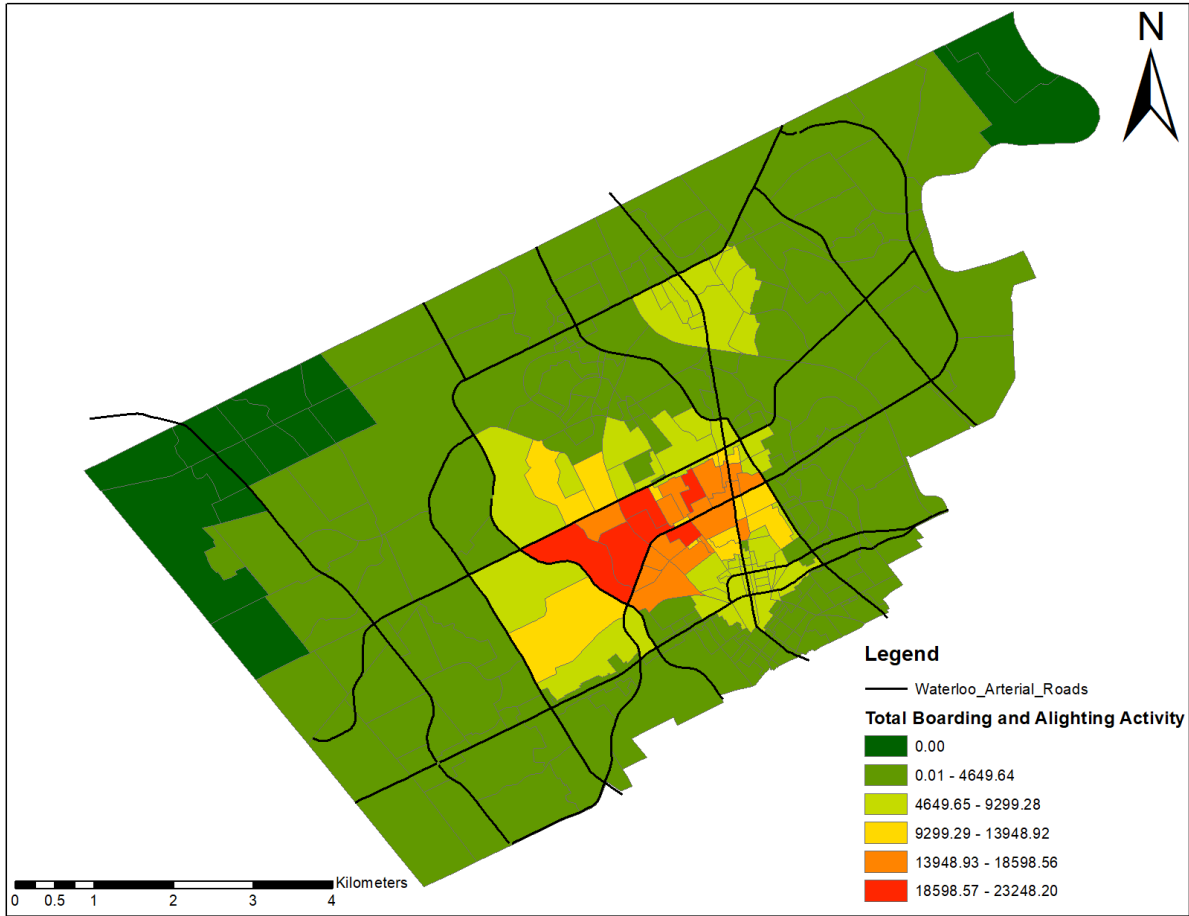


Figure 3.3.2.11: Transit Activity equal interval map by PLUM zones

Elementary and Secondary Schools' Student Enrollment

Unlike built environment variables that estimate pedestrian activity potential, enrollment data tend to shed light on actual pedestrian demand. Many school boards encourage walking to school and run multiple programs encouraging active transportation participation among kids to reduce the risk of chronic health complications (Student Transportation Services of York Region, 2018; Metrolinx, n.d.). In addition, academic authors, researchers, and municipalities have used some type of education-related variables to understand the spatial distribution of walking activity (Millward, Spinney, and Scott, 2013).

The data on school enrollment is available through an agreement with the University of Waterloo and the City of Waterloo. Otherwise, the enrollment data is not available through public open GIS portal. The data is available as a point feature layer, which includes attributes such as the school name, school board association, school class (i.e., elementary, and secondary), school addresses, and total student enrollment per school. Two elementary schools' enrollment data were missing; therefore, phone calls were made to the two schools' principal office to acquire an enrollment estimate for 2016/2017 school year.

Schools' catchment areas do not match the boundaries for PLUM zones. In addition, elementary schools and secondary schools each have their own set of catchment areas, which represented an issue trying to merge their data. In effort to resolve the spatial mismatch between the different boundaries, catchment areas were neglected and the enrollment data was assigned to the hosting PLUM zone. In case, where two schools lay in the same PLUM zone, the sum of both schools' enrollment data was passed down.

There are 34 schools, a total of both elementary and secondary schools, in the City of Waterloo. Together, they occupy 30 PLUM zones and provide education to almost 18,000 students. From Figure 3.3.2.13 below, all highest student enrollment is located on the city outskirts. According to Figure 3.3.2.12, the histogram is skewed to the right due to only a few PLUM zones that contain schools. There is a maximum of about 2 thousand students who arrive at the same PLUM zone daily. The average school enrollment is about 50 students and the standard deviation is about 200 (see Figure 3.3.2.12).

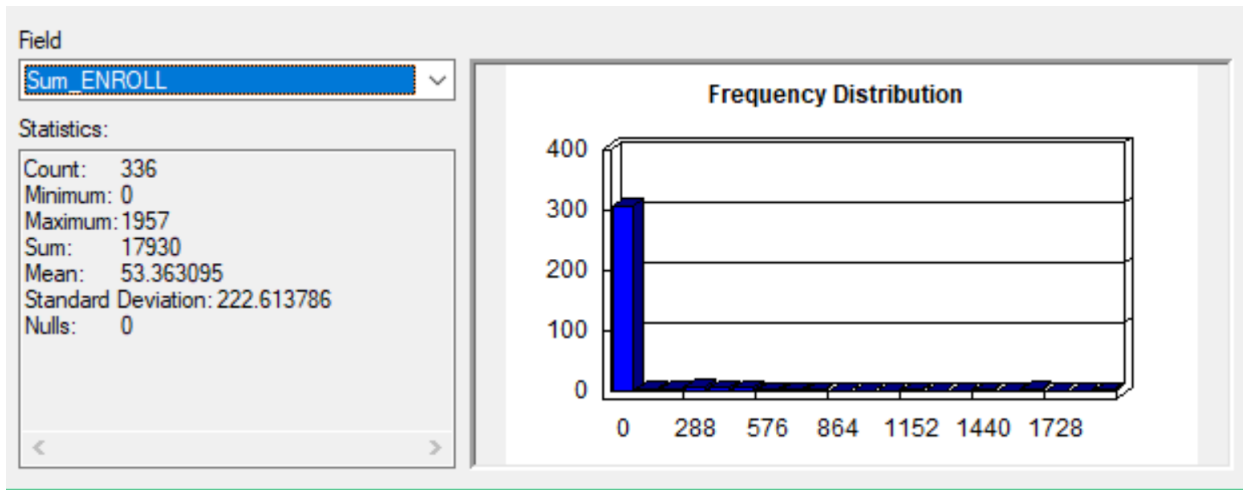


Figure 3.3.2.12: Elementary and Secondary Schools' Student Enrollment Descriptive Statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

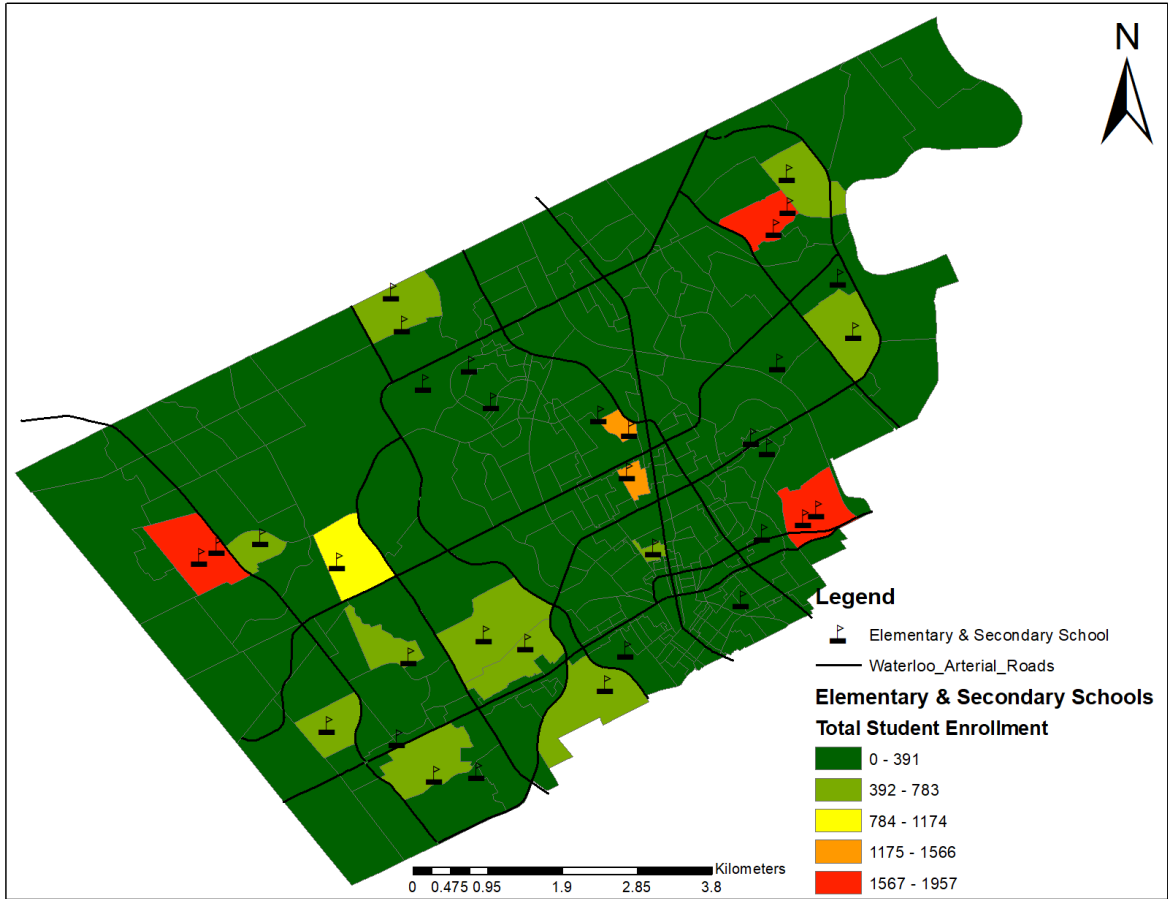


Figure 3.3.2.13: Elementary and Secondary Schools' Student Enrollment equal interval map by PLUM zones

Post-Secondary Institutional presence

Despite the potential multicollinearity between post-secondary institutions and transit activity, post-secondary institutional presence is included as a separate variable. Post-secondary institutions are not just educational hubs. They also attract land use diversity and employment. An optimal option is to use student registration data to measure the magnitude of activity on campus; however, such data is unavailable and would be hard to interpret since post-secondary students move across campus(es), which also intersects multiple PLUM zones.

With these difficulties mentioned above in mind, the intent is to use building permit data, which is accessible through the Geospatial Lab as a points feature class, to classify and isolate post-secondary institutional buildings that are used for educational purposes, then create a binary index (i.e., 0 or 1) representing post-secondary presence. Review of the data revealed that every building on post-secondary campuses does not necessarily have an available building permit data point. As long as there is one building permit point in the analysis area, that PLUM zone is assigned a value of 1. This overcomes the issue of incomplete building permit point data, while still representing post-secondary institutional presence.

In total, there are 10 PLUM zones, which include post-secondary institutional buildings. These zones represent the University of Waterloo campus, the Wilfrid Laurier University campus, and the Conestoga College campus. According to Figure 3.3.2.15 below, post-secondary institutions are located along University Street, with the University of Waterloo having the biggest campus of all three. When comparing Figure 3.3.2.15 with Figure 3.3.2.11, there is an association between transit activity and post-secondary institutional presence, which confirms the concern for multicollinearity.

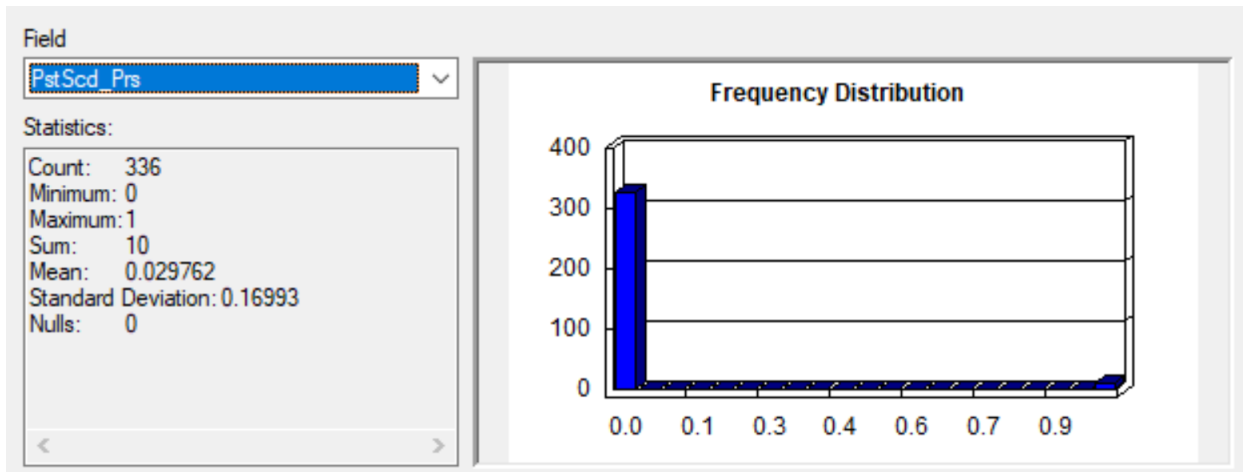


Figure 3.3.2.14: Post-Secondary Institutional Presence variable descriptive statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

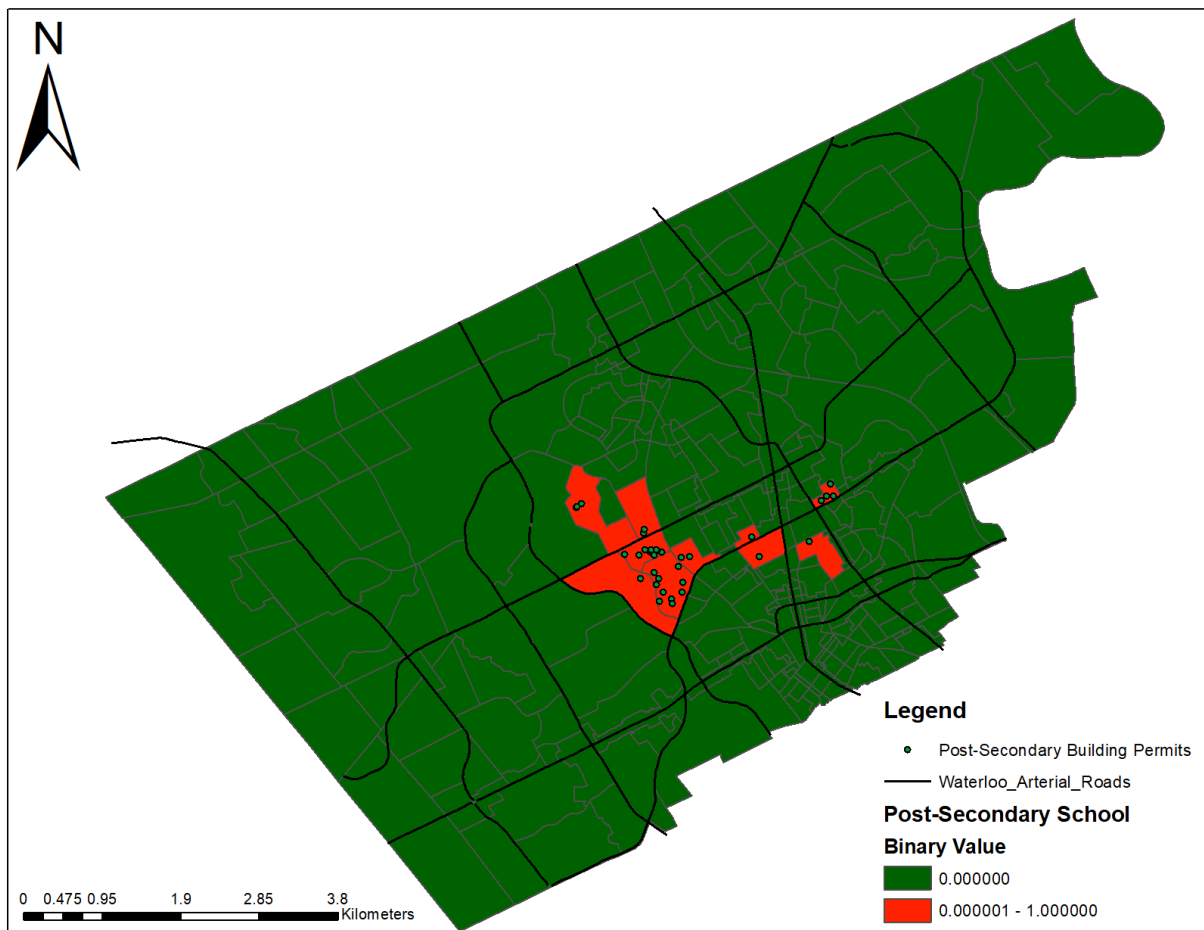


Figure 3.3.2.15: Post-Secondary Institutional Presence quantile map by PLUM zones

3.3.3. Regression Analysis

Regression analysis is used here to identify significant relationships between the explanatory variables and walking and to calibrate the model for accurate prediction. Since the first objective of this thesis is about predicting pedestrian activity, actual pedestrian trips are used to calibrate the model. Pedestrian data is available through the Transit Tomorrow Survey (TTS), which contain daily walking trips data based on reported trip activities and primary mode used (Data Management Group, 2018b).

Before explaining the process of conducting a regression analysis, it is important to understand the limitations of the TTS data pertaining to walking trip activity. According to the 2016 TTS: Data Expansion and Validation report, there are multiple concerns in regard to using TTS data, especially in university cities and for pedestrian representation (Data Management Group, 2018b). The first concern is that the TTS survey is based on mailing the survey to private households and depending on a single household member to remember and report trip activities for each other household member. This raises two alarms. The first is the exclusion of public housing residents and also students living in residence. Secondly, large households housing multiple students or being rented out are under-represented as landlords tend not to remember other members' trip activities or could care less for a transportation survey about the tenants' transit activities. These two alarms question the suitability of using TTS data in a university city like Waterloo. The second concern about the suitability of the TTS to capture pedestrian activity. In addition to previous concerns, the TTS survey only records primary trips' transit modes, which means the exclusion of secondary walking trips from cars to shopping strips, from store to store, from work to grocery stores to cars, and from schools to community centres to transit

stops. As a result, by default walking activity is under-represented through the TTS survey design.

Despite all the negatives mentioned, TTS is the only publicly available data source for consistent transportation activity records across multiple regions, cities, and towns in Ontario. Other data sources are available at the aggregate level, which are not suitable for the calibration process.

The optimal solution is to gather primary data through surveys or pedestrian count, but due to the study's limited resources, TTS pedestrian data are used, despite its limitations. Given that the study's contribution is primarily in developing the method as a proof of concept, data accuracy is less of a concern.

Daily walking trip count is available for TAZ in form of trip origins and destinations. It is recognized that not all trips originated from a TAZ will end up in the same TAZ and at the same time not all trips will end up in a different TAZ. To strike a balance, the maximum of either walking trips by origin or destination was used. Using the maximum captures the highest demand regardless of the trips' direction. The critique for using the maximum is the unknown traffic flow; however, it is not of concern for this study goal.

Figure 3.3.3.16 shows the spatial distribution of TTS reported walking trips. High foot traffic is located along King St. and along University Ave between the University of Waterloo and Wilfrid Laurier University. In addition, there is are a couple of isolated zones of high pedestrian activity located on the city peripheries. Review of the walking trips' histogram (Figure 3.3.3.17) reveals that the data is heavily skewed towards the right. More than 80% of the TAZs in the city have a reported maximum walking activity of 325 or less, which makes up about 10% of the maximum

recorded activity of 3248. On average 8% of the household daily trips are made by walking, while driving makes up 80% of the daily household trips (Data Management Group, 2018a).

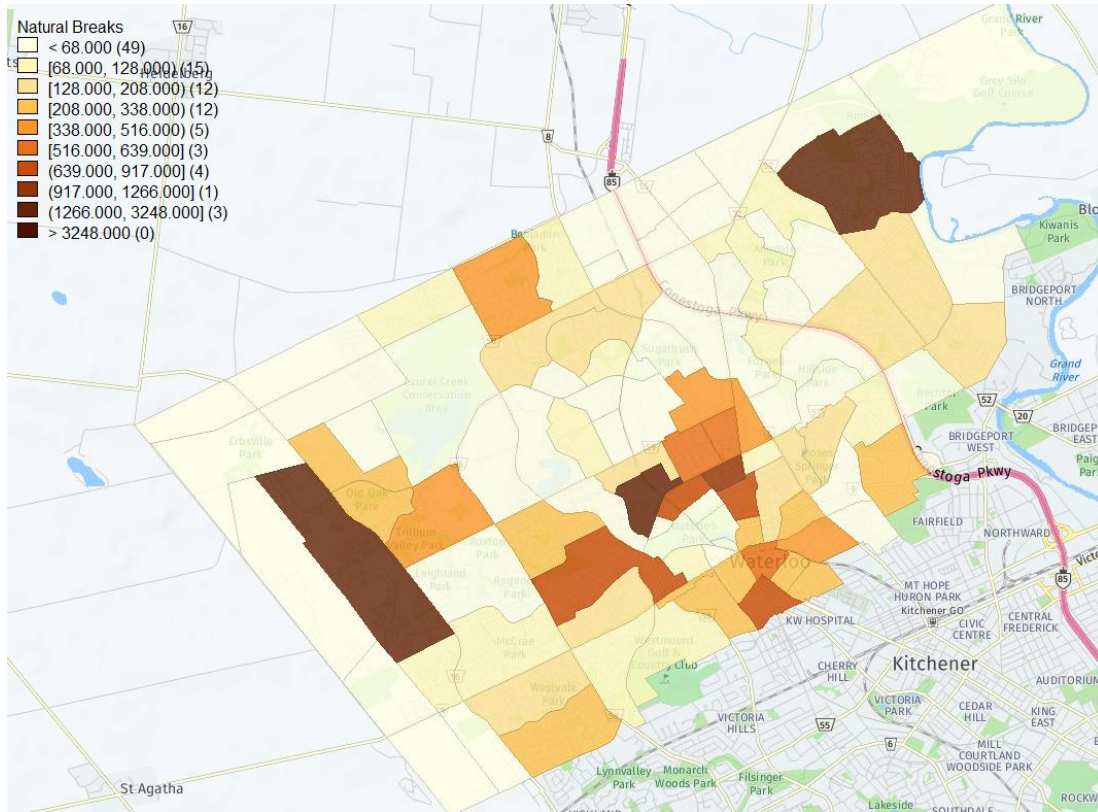


Figure 3.3.3.16: 2016 TTS Reported Walking Trips Distribution per TAZ Map

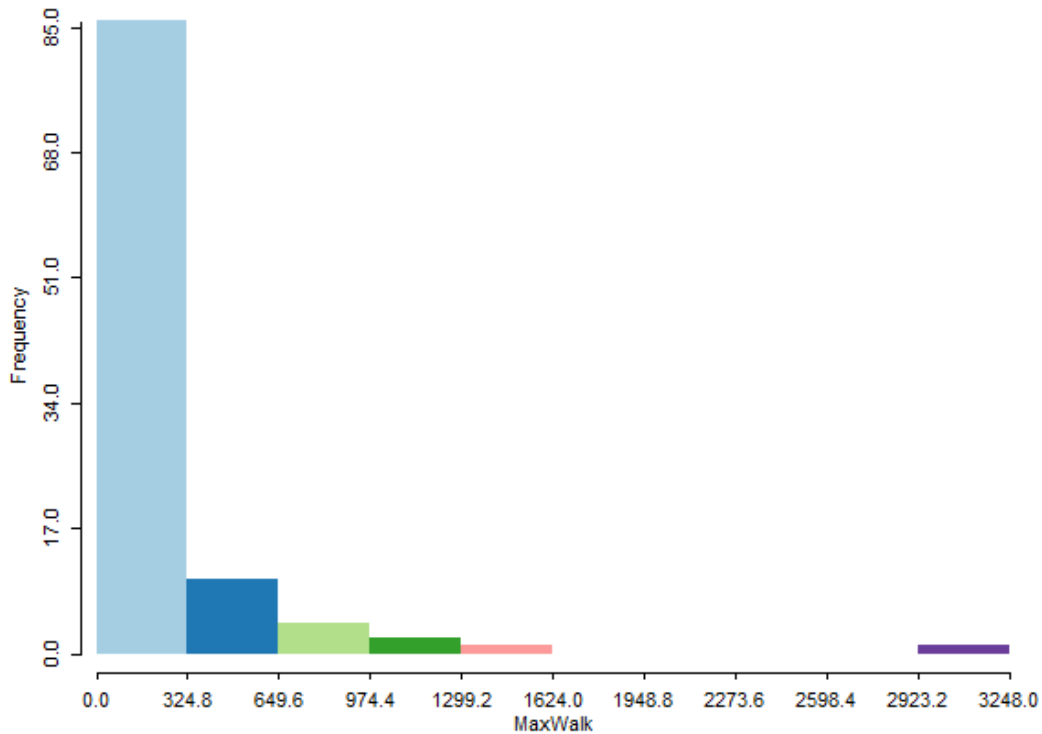


Figure 3.3.3.17: 2016 TTS Reported Walking Trips Histogram

TTS data is available for Traffic Analysis Zones (TAZ), while the study’s geographical unit of analysis is PLUM zones (PZ). To proceed with the regression analysis, there should be a unified geographical unit of analysis to avoid a Modifiable Areal Unit Problem (MAUP). Data at a coarse level, require disaggregation, which assumes that data is evenly distributed across an area. This assumption is often criticized for losing sensitivity, as discussed in the literature review.

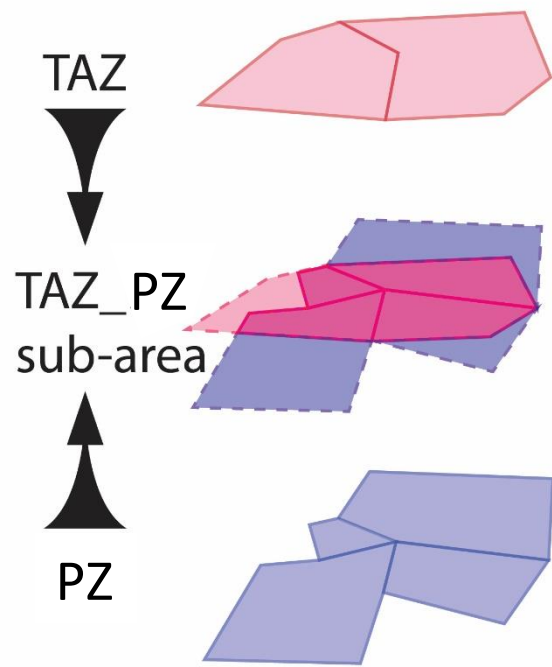


Figure 3.3.3.18: TAZ_PZ sub-area illustration

A simplification to data disaggregation is to conduct the analysis at the in-between scale, which in this study is identified as “TAZ_PZ sub-area” or “sub-area.” The assumption for evenly distributed data is retained to weight the land values based on the percentage of land making up the sub-area. For example, if a TAZ has 100 walking trips but only 25% of the TAZ make up the sub-area, then there are 25 walking trips in that sub-area. In total, there are 499 sub-area in the City of Waterloo, when intersecting TAZ and PZ layers.

Not all sub-areas are used for the regression analysis due to inconsistent data availability. Also, some sub-areas were too small to include as the TAZ and the PZ layers do not perfectly align. TTS walking trips data was not available for all TAZ in the City of Waterloo, which made it not available in all sub-areas. Other variables like net commercial floor area ratio, and Elementary and Secondary School Student Enrollment are not available per each sub_area. The criteria for choosing sample areas is based on specific data availability and area size. Sample areas were picked for being at least 16 hectares in size, which is the same size as a 400m grid cell representing 5-minute walk. Also, sample areas must have an associated TTS walking trip data, land use entropy, metric reach, and transit activity more than 0. In total, there are 69 sub-areas that match the sampling criteria (see Figure 3.3.3.19).

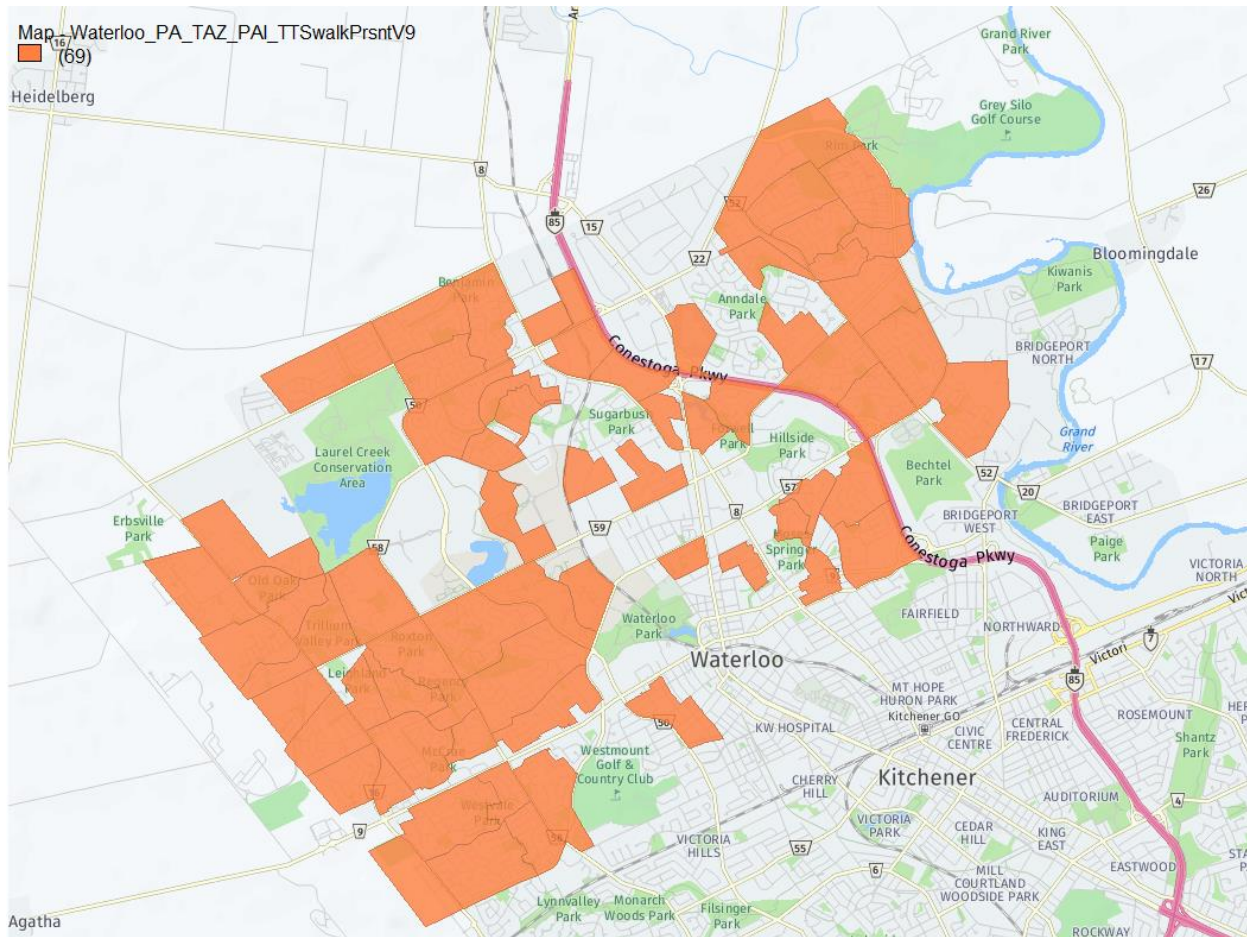


Figure 3.3.3.19: Map of sample TAZ_PZ sub-areas used in the regression analysis

The process for the regression analysis was inspired by Raford and Ragland (2006) when they used multiple linear regression to calibrate the Space Syntax model with a sample of pedestrian counts. In addition, other researchers in the field of walkability indices have used some sort of a correlation test to validate their walkability indices (Duncan et. al., 2011; Manaugh, and El-Geneidy, 2011; Frank et. al., 2010). Following similar steps, I employ two types of regression to reach optimal calibration of the Pedestrian Activity Model.

The first step taken was to test whether a raw data version or a transformed data version is better in explaining the correlation between the variables and walking trips. Transformation is usually used to adjust data distribution towards normality. Variables' data were first normalized to the

same scale using a percentage of the maximum, then transformed using natural log. Any zero value after the normalization process was replaced with a value of 1 because it is inapplicable to use natural log on zero values. Two sets of linear regression were conducted on each set of data versions. Overall, raw data had a better correlation than transformed data. Therefore, raw data was used in the next step of the regression analysis.

This step includes carrying both a multiple variable linear regression and a spatial error regression to find the best-fit regression model to represent the correlation between the considered explanatory variables and walking trips. The reason for considering a spatial error regression is for the typical spatial autocorrelation for foot traffic, meaning that the relationships between the considered variables and the response variable usually have an underlying spatial pattern. The regression tests were conducted using GeoDa, which can run both types of regression. Raw data were normalized for easier interpretation of variables' influence in relation to each other. All seven variables mentioned above were tested then the significant variables (>90%) were re-tested to finalize the Pedestrian Activity Model.

3.3.4. Pedestrian Activity Model Formula

As discussed in the literature review, some of the approaches to estimating pedestrian activity lacked calibration and validation to predict actual trip counts. Conventional walkability index formula is based on summing the variables per area. Advanced formulas included weighting these variables. Initially inspired by walkability indices, but with the ability to calibrate a predictive model using regression, a walkability index is no longer considered to capture pedestrian behaviour. The final Pedestrian Activity Model is based on the regression analysis results, which are revealed in the next chapter. The role of the regression analysis was not just to objectively find the optimal set of weights to maximize correlation to walking trips but to also identify significant variables. Below is the general formula for spatial error regression formula that the model is based on.

$$y_i = \hat{y}_i + u$$

y_i = observed value for area i

\hat{y}_i = predicted value for area i

u = residuals

$$y_i = \alpha + (\beta \times x_i)_n + u$$

$$u = \lambda W\epsilon + \epsilon_i$$

α = Y – intercept

$(\beta \times x_i)_n$ = explanatory variable term

β = regression coefficient

x_i = explanatory variable value for the area i

n = number of explanatory variables ($n= 1, \dots, 5$)

u = residuals/error term

$\lambda W\epsilon$ = spatial error = (lambda \times average of adjacent areas' residuals)

ϵ_i = unexplained error

Since the model is about predicting pedestrian activity for the rest of the study area, the error term was removed from the final formula. The new modal's final formula is:

$$\hat{y}_i = \alpha + (\beta \times x_i)_n$$

\hat{y}_i = predicted value for area i

α = Y – intercept

$(\beta \times x_i)_n$ = explanatory variable term

β = regression coefficient

x_i = explanatory variable value for the area i

n = number of explanatory variables ($n= 1, \dots, 5$)

Pedestrian Priority Zone

The purpose of this section is to illustrate briefly the process of determining Pedestrian Priority zones based on high foot traffic clustering. In ArcMap, from the symbology tab of the Pedestrian Activity Model layer properties, data were classified by Natural Breaks into 10 classes. Natural Breaks is useful in highlighting change according to the natural flow/distribution of the data. A statistical analysis of the gradual accumulation of the highest Natural Break classes is conducted to suggest a priority zone. The analysis looks at the pedestrian activity (predicted trips count), associated pedestrian infrastructure (length), and area coverage (area). The suggested number of priority zones is subject to change depending on the type of public investment, operation restrictions, and resource limitations.

Chapter 4: Findings and Discussion

Previous chapters established the conceptual framework and methods to predict pedestrian activity and to define the Pedestrian Priority Zone. This chapter's goal and structure are focused on sharing the findings for:

- The initial regression models (i.e., single variable linear regression, multiple linear regression, and spatial error regression),
- the Pedestrian Activity Model (includes a breakdown for some classes), and
- the Pedestrian Priority Zone configuration

4.1. Regression Models

Single Variable Linear Regressions

As illustrated in the Methods Chapter, all seven explanatory variables are not normally distributed; therefore, data transformation to restore normality was considered. Single variable linear regressions were carried out twice for each variable: once with raw data format, and again with data transformed using Natural Log. The reason is to find which data representation produces a higher correlation with the dependent variable (i.e., TTS walking trips). Figures 4.1.1 to 4.1.6 are scatterplots showing the linear regression best fit line and the model's R-square value for selected variables. The rest of the scatterplots are in Appendix B.

A major finding pertaining to some of the variables and their correlation to walking was their unexpected negative direction. According to the literature review, each of the variables considered in this study should have a positive correlation to walking trips. The variables with negative correlations were Land Use – Entropy (see Figure 4.1.1 and Figure 4.1.2 below), Land Use – Interaction Lines, and Net Commercial Floor Area Ratio (see Appendix B). One possible explanation behind the negative correlation is data quality issues (e.g., using OP land uses), as

discussed in the previous chapter and reinstated as a limitation of the model in the Conclusion Chapter.

In determining which data form is better, two additional linear regressions were carried out after summing all variables with respect to their correlation direction as revealed in the single variable linear regressions. Variables in the raw data format had varying data scales. Therefore, raw data were normalized to a percentage scale (0-100) based on the maximum value per each variable, then summed with respect to their associated correlation direction as found in the single variable regressions (Figure 4.1.5). In comparing Figure 4.1.5 and Figure 4.1.6, it can be seen that there is a stronger correlation between the normalized explanatory variables and walking activity ($R^2 = 0.406$) versus the transformed data ($R^2 = 0.255$). As of yet, these findings at their best (Figure 4.1.5) can only explain 40% of the variation in walking happening by zones in the City of Waterloo. It can be concluded that the normalized data, which substitute for the raw format, are better at explaining walking activity. However, these findings did not yet take into account weighting variables individually to maximize the correlation as well as eliminating insignificant variables, which is the next step in constructing the Pedestrian Activity Model.

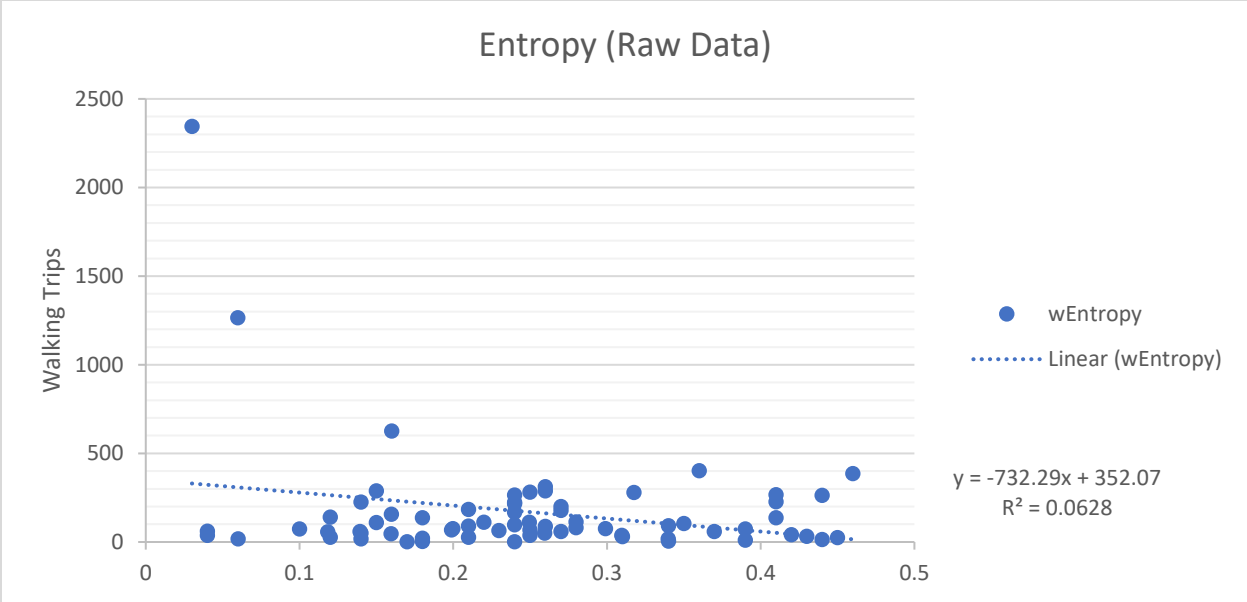


Figure 4.1.1: Land Use – Entropy (Raw Data version) Linear Regression

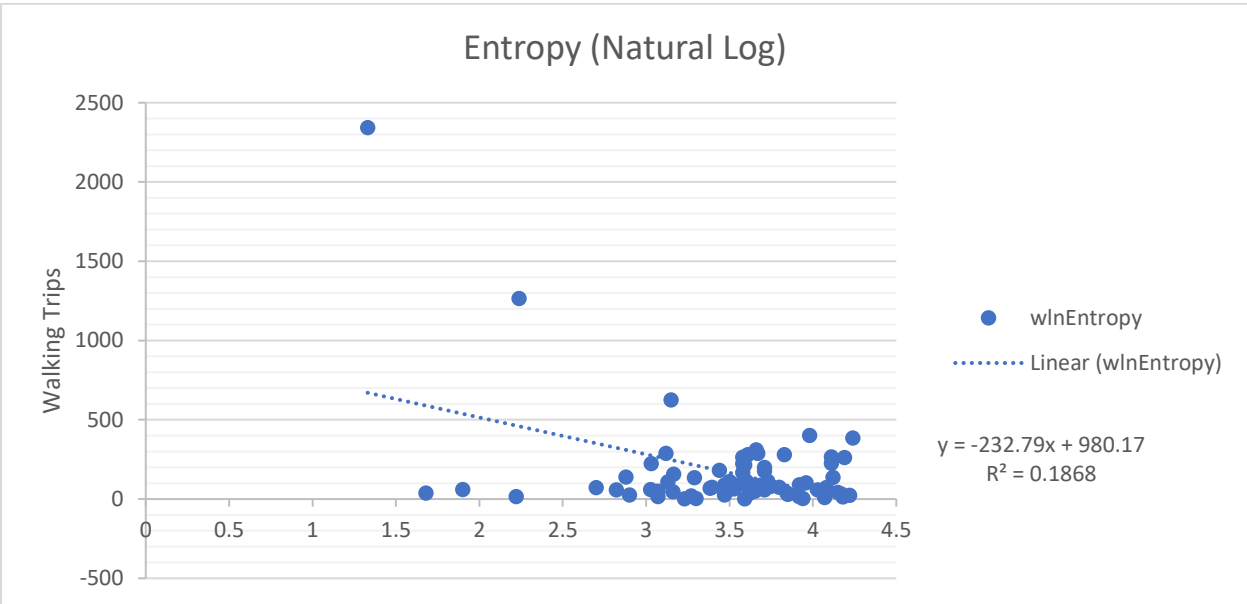


Figure 4.1.2: Land Use – Entropy (Natural Log version) Linear Regression

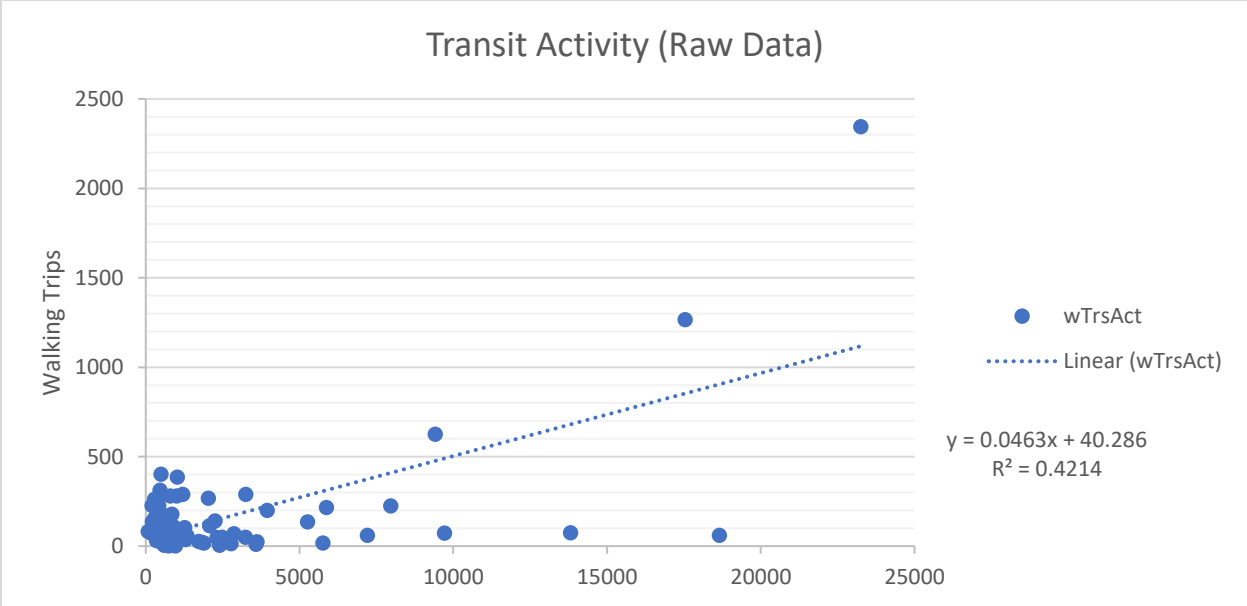


Figure 4.1.3: Transit Activity (Raw Data version) Linear Regression

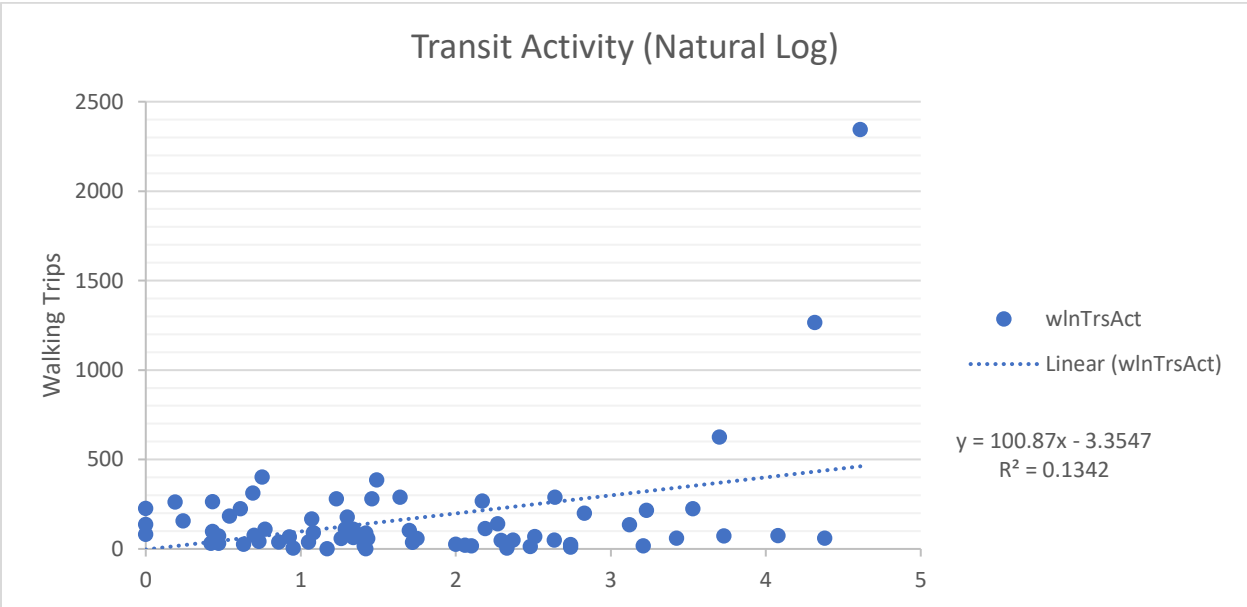


Figure 4.1.4: Transit Activity (Natural Log version) Linear Regression

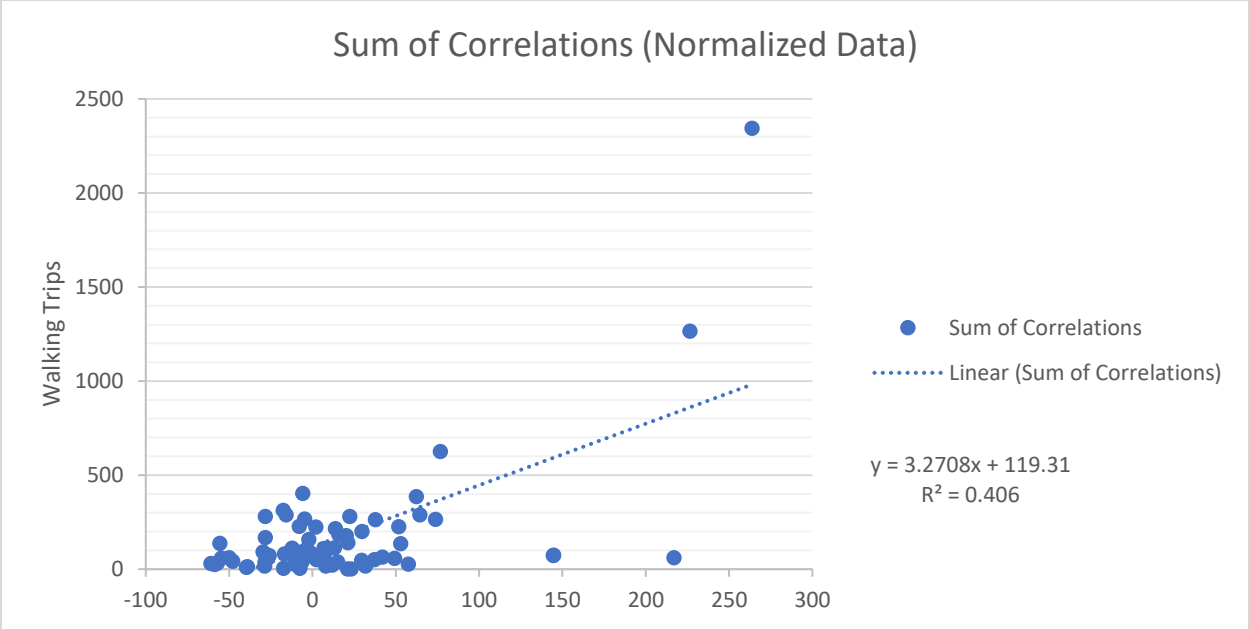


Figure 4.1.5: All Variables' Sum of Correlations (Raw Data version) Linear Regression

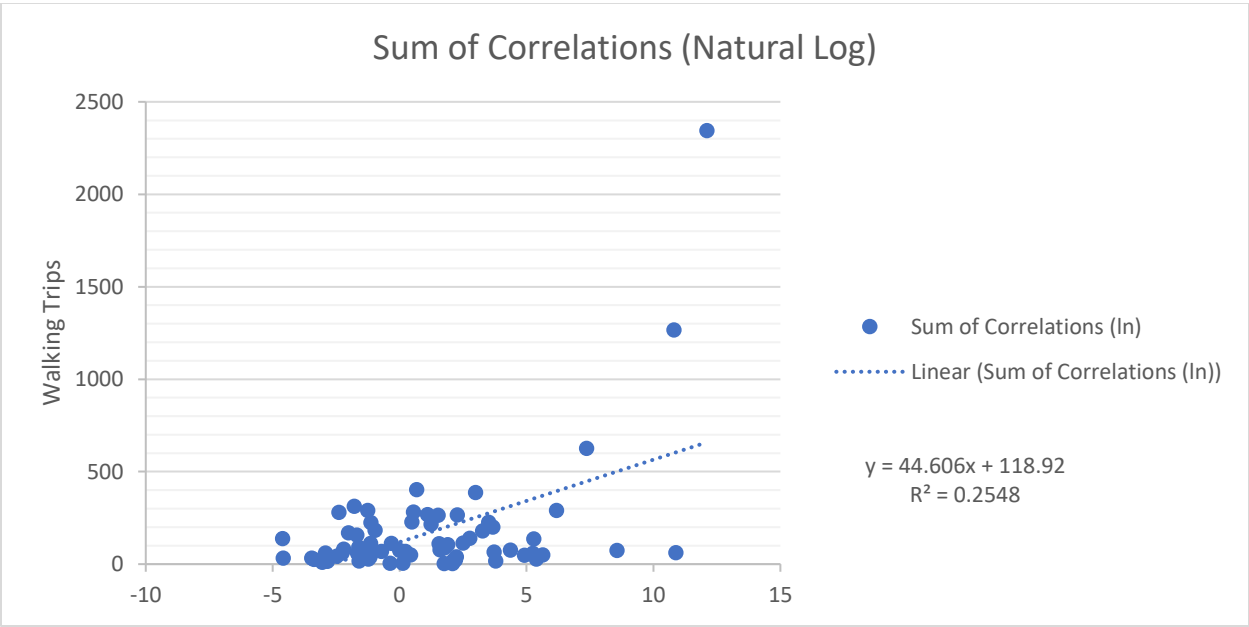


Figure 4.1.6: All Variables' Sum of Correlations (Natural Log version) Linear Regression

Ordinary Least Square Regression (OLSR)

The regression models referred to above were limited to a single variable for exploratory purposes. In considering multiple variable regression models, the work is completed using “GeoDa”. GeoDa is a Geographic Information System (GIS) analysis software that is used to run both a multiple variables linear regression as well as a spatial error regression.

The advantage of running a multiple variable Least Squares regression is the automated calibration of the coefficient for each variable as well as finding which of the variables is significant to the case study context. Another advantage to running a multiple variables linear regression is the freedom to use the explanatory variables in either raw form or normalized form; however, using normalized data has another advantage, which is the ability to interpret variables’ influence easily in relation to other variables with the same scale.

While there are different versions of multiple variable linear regression, Ordinary Least Square Regression (OLSR) is used due to (1) its availability on “GeoDa”, which can also run spatial-based regressions, and (2) the researcher’s familiarity with running and interpreting OLSR models.

Figure 4.1.7 shares the Ordinary Least Square Regression’s results, which is based on normalized data. The first thing to note is the R square value, which explains to what degree the model statistically explains variation in the the dependent variable (i.e., walking activity). The adjusted R square value is 0.48, which means the regression model using all variables explains 48% of the variability in pedestrian trips. This model has a confidence level of more than 95%. Another number to remember and keep in mind when comparing with other regression models is the Akaike Info Criterion (AIC), which also indicates how well the model represents the

dependent variable in comparison to other models. For this OLSR, the AiC number is 954.85.

The rule of thumb is the lower the AiC value the better representative is the model. In this case, the AiC value is ambiguous until compared to another model's AiC.

Besides the overall fit of the model, it is important to understand the significant components making up the regression model, their correlation direction, and their coefficients. As mentioned before, the variables considered in this study were expected to have a positive correlation with walking by various studies (Liu and Griswold, 2009; *Region of Waterloo, 2009*; Ewing and Cervero, 2010; Frank et. al., 2010; Meghelal and Capp, 2011; Sundquist et. al., 2011; Manaugh and Kreider, 2013; Millward, Spinney, and Scott, 2013; Ellis et. al., 2016; Tsiompras and Photis, 2016). According to the figure below, there are two variables - Post-Secondary School Presence, and Net Commercial Floor Area Ratio - that have negative correlations with walking. Despite having a relatively similar coefficients as other variables, these negatively correlated variables are insignificant (below the 90% confidence level).

Another major finding is that, of all the variables included in the model, Transit Activity has the most influence on pedestrian trip activity with a coefficient value of 14.44. The second most influential positively correlated variable is Metric Reach with a coefficient value of 4.95. In other words, with an increase of one unit in the explanatory variable Transit Activity increases walking trips three times that of Metric Reach. Both variables are significant at the 90% confidence level, while the remaining five variables are insignificant.

While linear regressions are common in active transportation studies, it comes up short in explaining the spatial component of the data. The correlations between the considered variables and pedestrian activity usually have an underlying spatial pattern. Also, linear regression

assumes that data is normally distributed, which is violated as per the Jarque-Bera test (p-value < 0.05) (see Figure 4.1.7). For these reasons, a spatial-based regression is tested.

REGRESSION

SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

```
Data set           : Waterloo_PA_TAZ_PAI_TTSwalkPrsntV6
Dependent Variable : wWalk_1 Number of Observations: 69
Mean dependent var : 171.116 Number of Variables : 8
S.D. dependent var : 318.485 Degrees of Freedom : 61

R-squared          : 0.531624 F-statistic           : 9.89105
Adjusted R-squared : 0.477876 Prob(F-statistic)      : 3.52514e-008
Sum squared residual: 3.27809e+006 Log likelihood       : -469.426
Sigma-square       : 53739.1 Akaike info criterion    : 954.851
S.E. of regression : 231.817 Schwarz criterion     : 972.724
Sigma-square ML    : 47508.5
S.E of regression ML: 217.964
```

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	-184.569	120.374	-1.53329	0.13037
wEntropy_1	0.487041	2.51277	0.193826	0.84696
PstScd_1	-3.25931	2.27334	-1.43371	0.15676
wEnS_Enr_1	2.12679	1.50903	1.40937	0.16380
wTrsAct_1	14.4386	3.31778	4.35187	0.00005
wMtcRch_1	4.94793	2.46985	2.00333	0.04959
wnComFAR_1	-4.99777	3.48263	-1.43505	0.15638
wIntrcLn_1	2.80683	5.97776	0.469545	0.64035

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 11.403450

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	235.9765	0.00000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	7	331.8169	0.00000
Koenker-Bassett test	7	60.8293	0.00000

===== END OF REPORT =====

Figure 4.1.7: OLSR results report – including all considered variables

Spatial Error Regression (SER)

For the spatial-based regression, a Spatial Error Regression (SER) is used because of its availability through “GeoDa” and also because of the researcher’s familiarity in running and interpreting SER models. Using “GeoDa” again ensures consistency in the result report format and easier interpretation. Evaluating whether the SER model or the OLSR model is better is dependent on the best-fit indicators (e.g., R^2 , and AiC). A better fit model would have a higher R square value while also scoring a lower Akaike info Criterion (AiC) value. Now, according to Figure 4.1.8 below, the SER model’s associated R^2 value is 0.57, and the AiC value is 951.76. On the other hand, the OLSR model’s associated R^2 is 0.53 and the AiC is 954.85. The numbers cited above indicate that the SER model is better at explaining the variation between the reported daily walking trips and the predicted values by about 4%.

Despite the fit differences between the two models, both models share similar results in terms of negatively correlated variables (i.e., Post-Secondary School Presence, and Net Commercial Floor Area Ratio), which are also insignificant variables. Also, Transit Activity and Metric Reach variables have the greatest pull among the remaining variables on pedestrian activity, with Transit Activity in the lead.

In considering the confidence level at 90% rather than the conventional 95% (see Figure 4.1.8), a third significant variable stands out, that is Elementary and Secondary School Student Enrollment. The Student Enrollment variable has a coefficient of 2.66, while the Metric Reach variable has a coefficient of 4.25 and Transit Activity’s coefficient is 11.53 (see Figure 4.1.8). In other words, one unit increase in Transit Activity equals four times increase in walking trips versus one unit increase in Student Enrollment and two times increase in walking trips versus one unit increase in Metric Reach.

For reasons like higher R squared value, and lower AiC value, the SER model is a better fit at representing the correlation between the explanatory variables and the pedestrian data. In addition, the SER model is the preferred option because it contains an additional significant variable bringing the model a step closer to understanding variables influencing walking in the City of Waterloo. Moving forward, the SER model is used to construct the Pedestrian Activity Model.

However, prior to constructing the Pedestrian Activity Model, the Spatial Error Regression (SER) model was re-run with only the three identified significant variables (i.e., Transit Activity, Metric Reach, and Elementary and Secondary School Student Enrollment). This process configures the coefficient values to maximize the correlation between the explanatory variables and the dependent variable.

The result of the new SER model is shown below in Figure 4.1.9, while Figure 4.1.11 shows the spatial distribution of the predicted values. The new SER model run has a slightly lower R square value (0.55) than the previous 0.57 R square value, but it eliminates insignificant variables. The new AiC value is smaller (946.34) compared to 951.29 for the former SER model (see Figure 4.1.9, and Figure 4.1.8).

The spatial distribution of the residuals shows positive residuals along the city outskirts plus at both post-secondary institutions' main campuses (Figure 4.1.12). This indicates that the predictive model under-represents walking trips at city outskirts and outlier high pedestrian activity areas. A histogram of the residuals shows slight skewness towards the right (see Figure 4.1.13).

While the new spatial error regression model at its best can only explain 55% of the variation in the walking activity, it is important to remember here the limitations for the TTS reported daily walking trips as well as the limitations associated with the sample area. The sample area underrepresents the downtown core, which might explain the unexpected insignificant correlations between some of the variables (e.g., Land Use Mix Entropy measure, Land Use Mix Interaction Lines measure) and walking activity.

```

REGRESSION
-----
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set           : Waterloo_PA_TAZ_PAITTSwalkPrsntV6
Spatial Weight     : Waterloo_PA_TAZ_PAITTSwalkPrsntV3
Dependent Variable : wWalk_1   Number of Observations: 69
Mean dependent var : 171.115942 Number of Variables   : 8
S.D. dependent var : 318.484707 Degrees of Freedom   : 61
Lag coeff. (Lambda) : -0.334103

R-squared          : 0.569548 R-squared (BUSE)      : -
Sq. Correlation    : -          Log likelihood          : -467.642371
Sigma-square       : 43661.9   Akaike info criterion : 951.285
S.E of regression  : 208.954   Schwarz criterion     : 969.158

-----
Variable           Coefficient      Std.Error      z-value      Probability
-----
CONSTANT           -149.993         101.002        -1.48505     0.13753
wTrsAct_1          11.525           2.73392         4.21557     0.00002
wMtcRch_1          4.25356          2.0666         2.05823     0.03957
wEntropy_1         0.625539         2.19949         0.284402    0.77610
PstScd_1           -1.53439         1.9267         -0.796384    0.42581
wEnS_Enr_1         2.65601          1.39767         1.90032     0.05739
wnComFAR_1         -4.23329         3.14249        -1.34711     0.17794
wIntrcLn_1         2.14568          5.07935         0.422432    0.67271
LAMBDA             -0.334103        0.144897        -2.3058     0.02112
-----

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST           DF      VALUE      PROB
Breusch-Pagan test      7      242.4255   0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : Waterloo_PA_TAZ_PAITTSwalkPrsntV3
TEST           DF      VALUE      PROB
Likelihood Ratio Test    1      3.5666    0.05895
===== END OF REPORT =====

```

Figure 4.1.8: SER results report – including all considered variables

```

REGRESSION
-----
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set           : Waterloo_PA_TAZ_PA_I_TTSwalkPrsntV9
Spatial Weight     : Waterloo_PA_TAZ_PA_I_TTSwalkPrsntV9
Dependent Variable : wWalk Number of Observations: 69
Mean dependent var : 171.115942 Number of Variables : 4
S.D. dependent var : 318.484707 Degrees of Freedom : 65
Lag coeff. (Lambda) : -0.365689

R-squared          : 0.552983 R-squared (BUSE) : -
Sq. Correlation    : - Log likelihood : -469.171256
Sigma-square       : 45342 Akaike info criterion : 946.343
S.E of regression  : 212.937 Schwarz criterion : 955.279

-----
Variable          Coefficient      Std.Error      z-value      Probability
-----
CONSTANT          -123.325         61.784         -1.99606     0.04593
nMtrRch           4.8325          1.96684        2.45699     0.01401
nTrsAct           8.87014         1.239          7.15912     0.00000
nEnS_Enrll       3.01644         1.33148        2.26548     0.02348
LAMBDA            -0.365689       0.143445      -2.54933     0.01079
-----

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST              DF      VALUE      PROB
Breusch-Pagan test 3      238.7847  0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : Waterloo_PA_TAZ_PA_I_TTSwalkPrsntV9
TEST              DF      VALUE      PROB
Likelihood Ratio Test 1      5.0160    0.02511
===== END OF REPORT =====

```

Figure 4.1.9: SER results report – including only significant variables

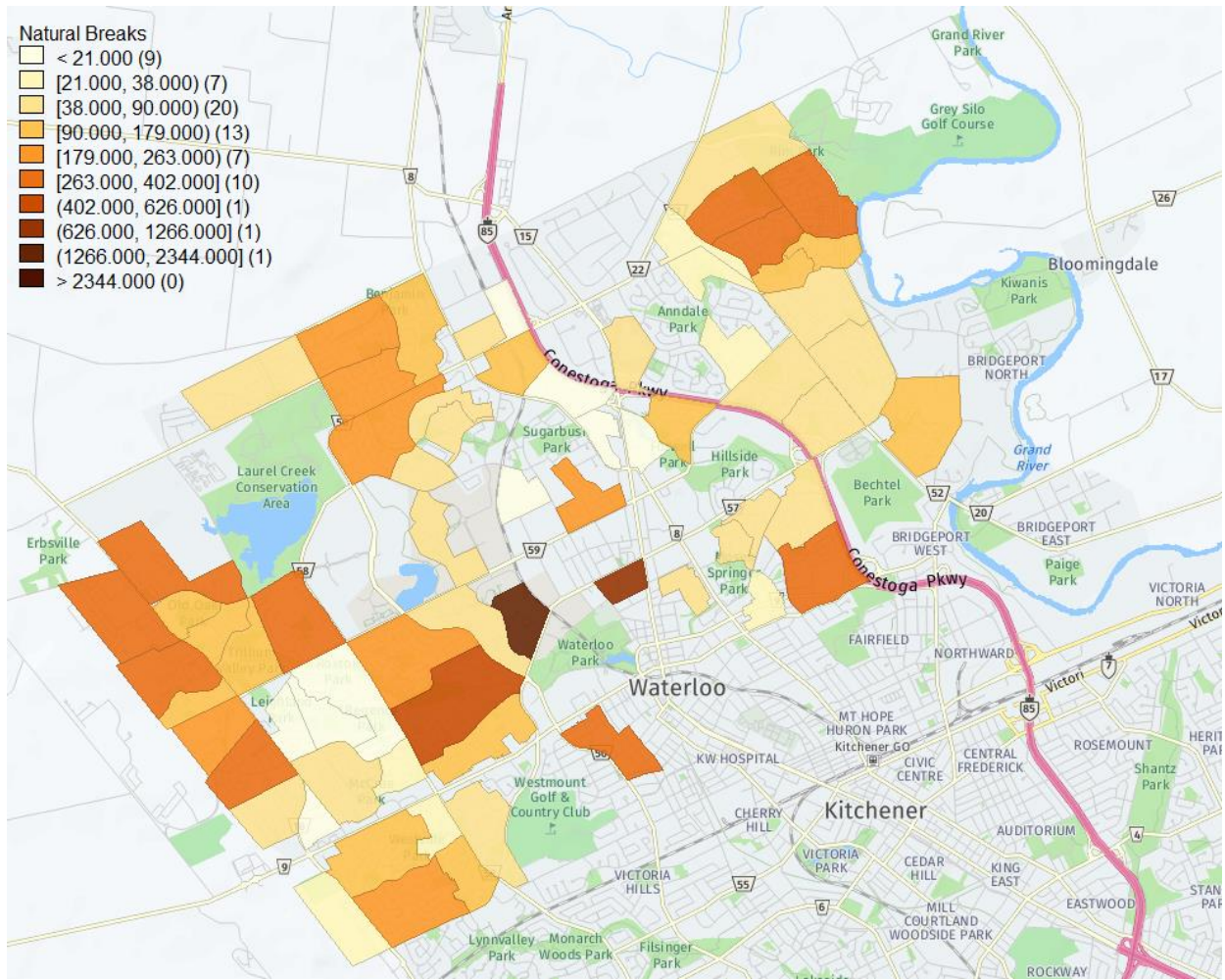


Figure 4.1.10: 2016 TTS reported daily walking trips per sample TAZ_PZ sub-areas

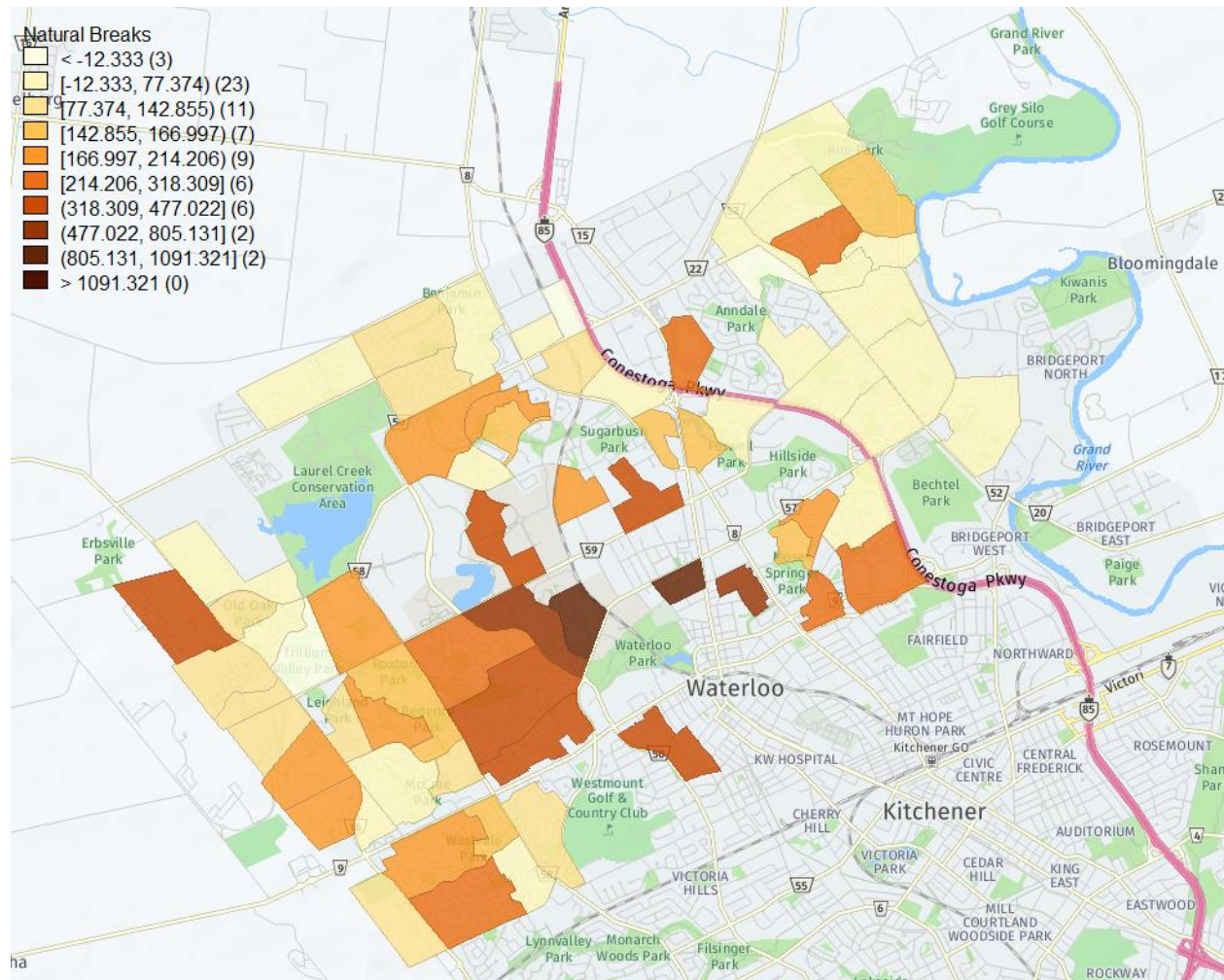


Figure 4.1.11: Spatial Error Regression's Predicted values (i.e., walking trips) map

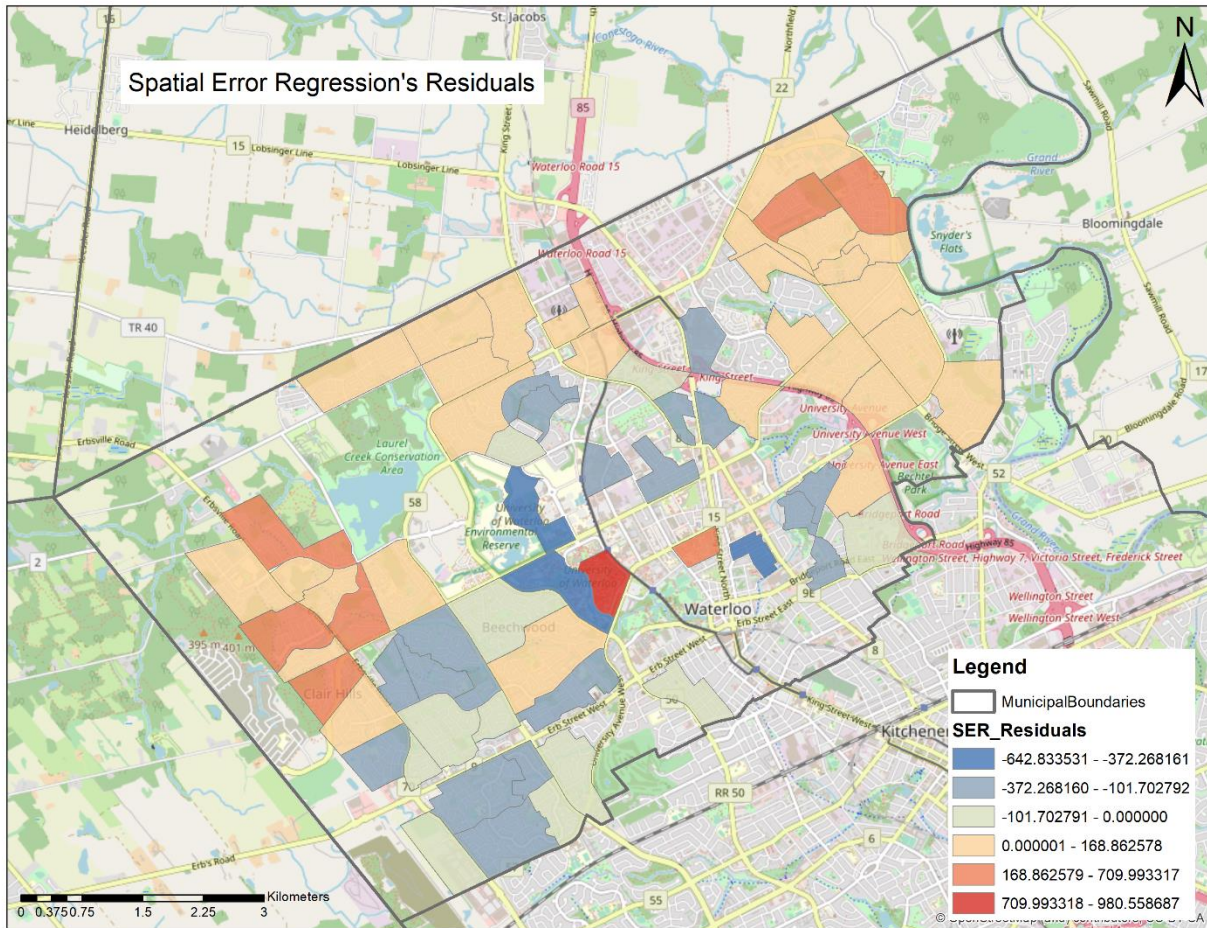


Figure 4.1.12: Spatial Error Regression's Residuals map

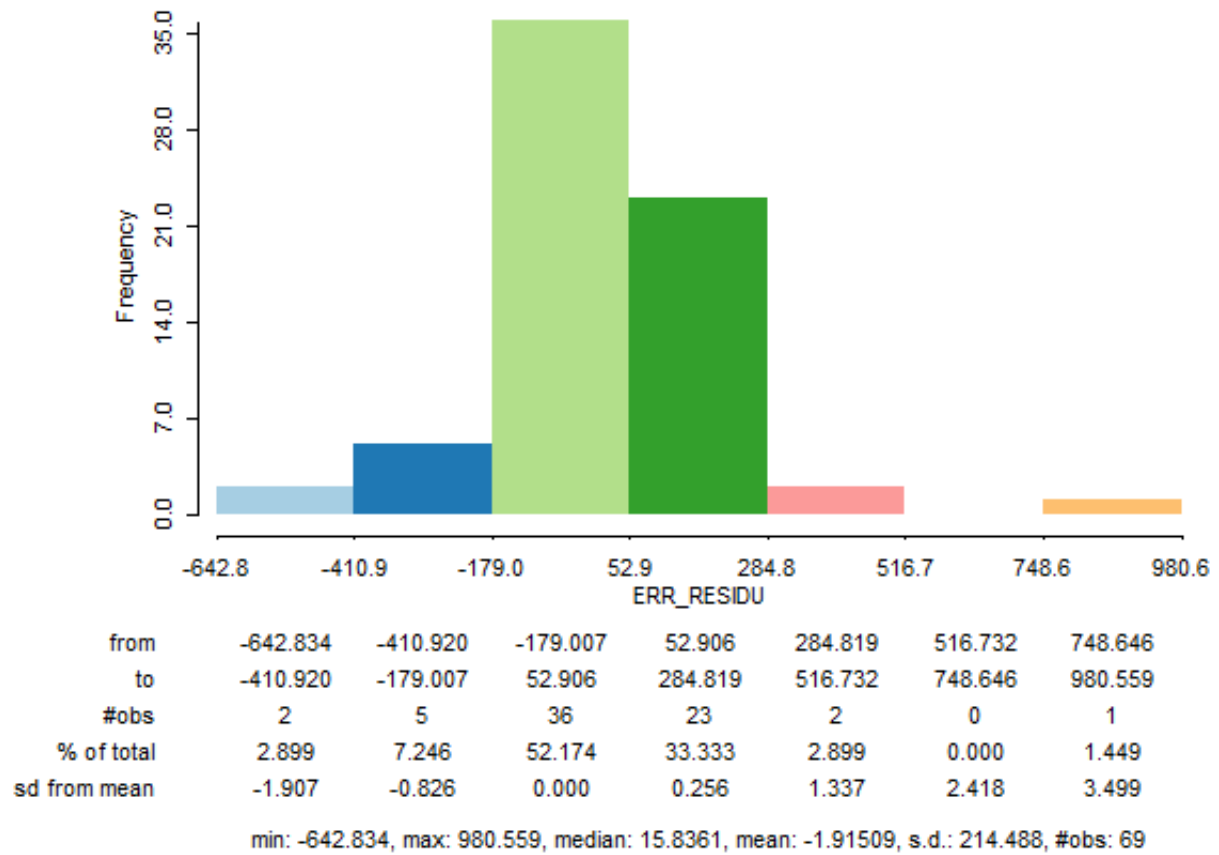


Figure 4.1.13: Spatial Error Regression's Residuals Histogram

4.2. Pedestrian Activity Model

A regular regression model is constructed as two parts. The first part is the predictive part of the model and the second part is the error term, which tells the difference between the predicted value and the observed value (i.e., TTS reported daily primary walking trips). Since the intent of the study is to predict pedestrian activity, the error term of the SER model is eliminated and the rest of the formula is adopted as the final formula to construct the Pedestrian Activity Model.

The Pedestrian Activity Model formula is:

$$\hat{y}_i = (-123.325) + (8.87014 \times x_{Transit\ Activity}) + (4.8325 \times x_{Metric\ Reach}) \\ + (3.01644 \times x_{Elementary\ and\ Secondary\ School\ Student\ Enrollment})$$

\hat{y}_i = predicted daily walking trips for area i

Using the three significant variables' data to plug into the formula above, primary daily walking trips were predicted per each of the 336 PLUM zones within the City of Waterloo's boundary.

The predicted trips does not include secondary walking trips as inherited from the response variable (i.e. TTS reported primary daily walking trips).

Walking level and scale varies from a study to another and from location to another, which makes it hard to standardize its way of measurement. In respect to the local context and the findings of the P.A.M., natural breaks classification type is used to divide up the predicted walking trips' scale into ten classes based on the natural flow of the data. Ten classes is used rather than five to focus on the variation between the classes and try to understand the underlying context to each class. Figure 4.2.2 shows the Pedestrian Activity Model outcome map for the City of Waterloo, while Figure 4.2.1 shows the Natural Breaks classes distribution versus the model's data distribution. The findings are structured in this sub-section to first share general

findings pertaining to multiple zones, then share the breakdown of six zones representing six classes.

The first general observation is that some zones have negative trip counts. These zones are located along the city's north-east and north-west borders (Figure 4.2.2). The reason these zones have a negative trip counts is that because the model's y-intercept term is a negative value. To preserve the predictive model adopted from the SER model, the negative y-intercept is retained in the final formula as an indicator of the model's general over-prediction. It is also a reminder of the model's limitations and a chance for future improvements through better pedestrian trip data and expansion of the considered explanatory variables (e.g., vehicle ownership, employment density).

Another general observation of trip distribution shows that low pedestrian trip activity dominates the outer edges of the City of Waterloo in an upside-down "U shape". Adjacency to the City of Kitchener accounts for the moderate pedestrian activity along the City of Waterloo's south border.

On the other hand, the highest predicted walking trips are clustered along University Ave in close proximity to post-secondary institutions, while moderate-high pedestrian activity extends along King St. south of University Ave. towards Uptown. These findings support similar findings in the field around downtown cores and post-secondary institutions being key drivers for pedestrian movement.

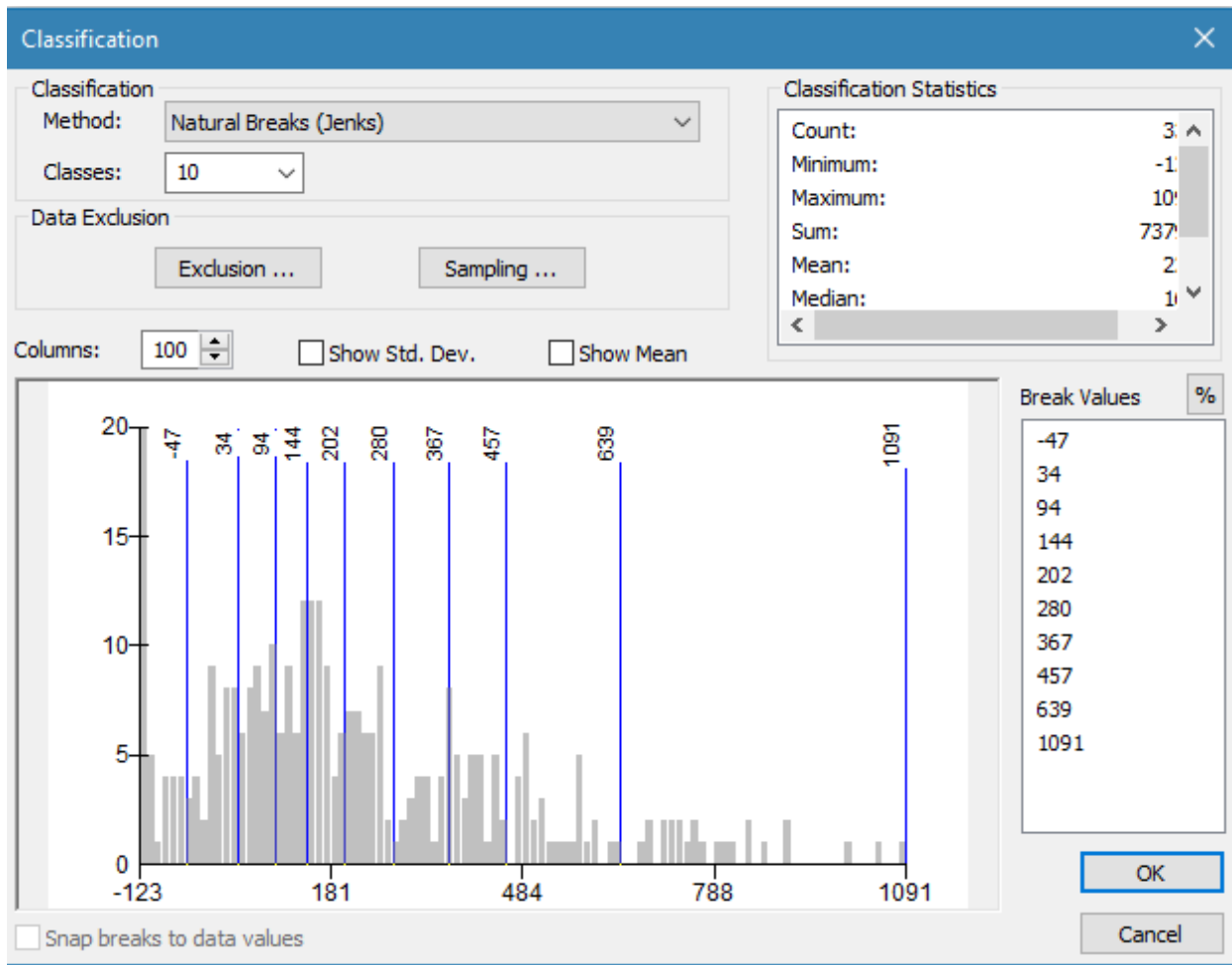


Figure 4.2.1: Pedestrian Activity Model's Natural Breaks classes versus data distribution

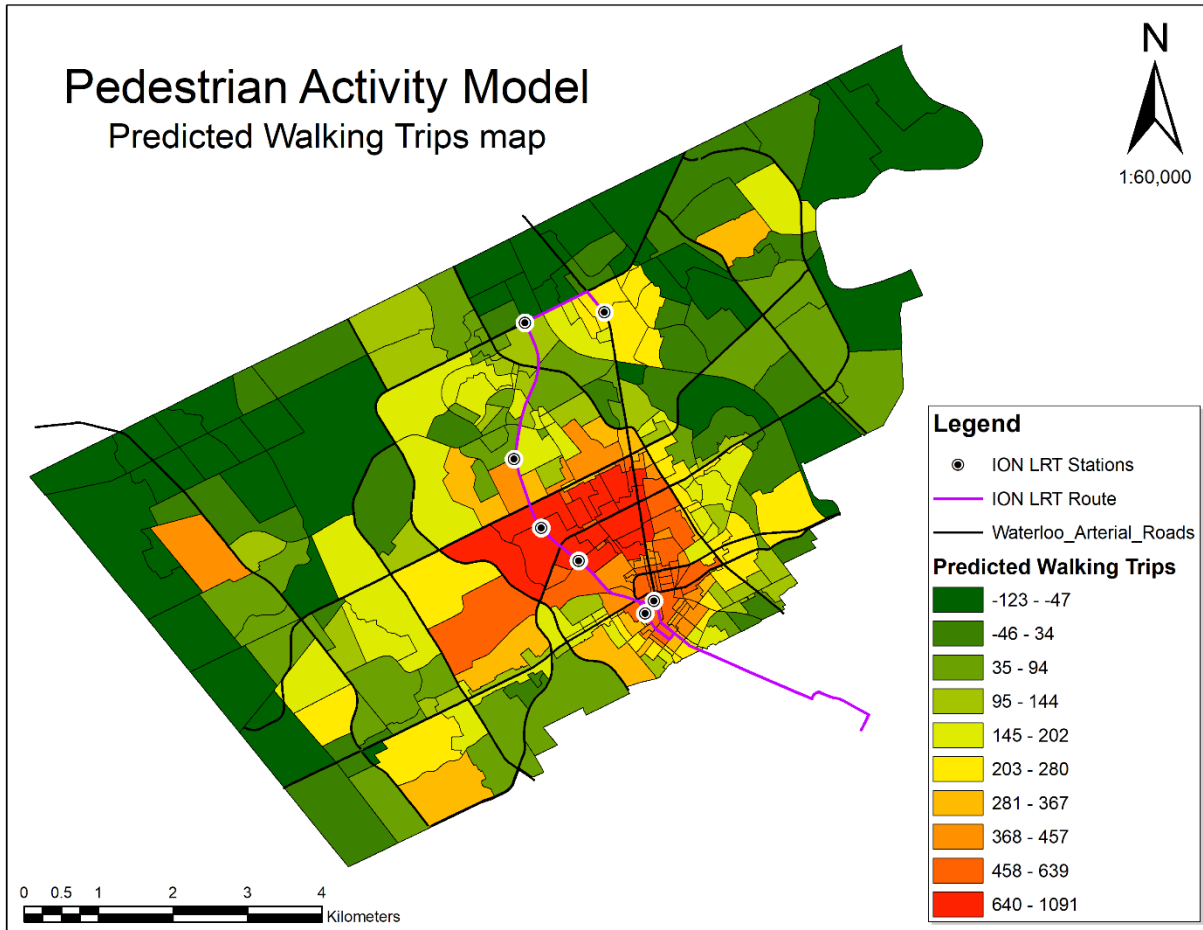


Figure 4.2.2: Pedestrian Activity Model Predicted Daily Walking Trips map

Pedestrian Activity Model’s predicted walking trips classification breakdown

This sub-section makes up the breakdown of the predicted walking trips classification by exploring six of the ten classes in Figure 4.2.2. The six classes showcase a wide variation from low predicted trip counts’ zone to high trip counts’ zones, not in that particular order. The classification of Pedestrian Activity Model predicted trips is based on natural breaks. Each of the six classes is explained by the varying role of each of the three significant variables. In addition, previously considered variables might be used to explain pedestrian activity in the zone, as well as other supplementary sources (e.g., policies, local programs, and planning documents), might be incorporated. Some of the classes have outliers , which are also explored below.

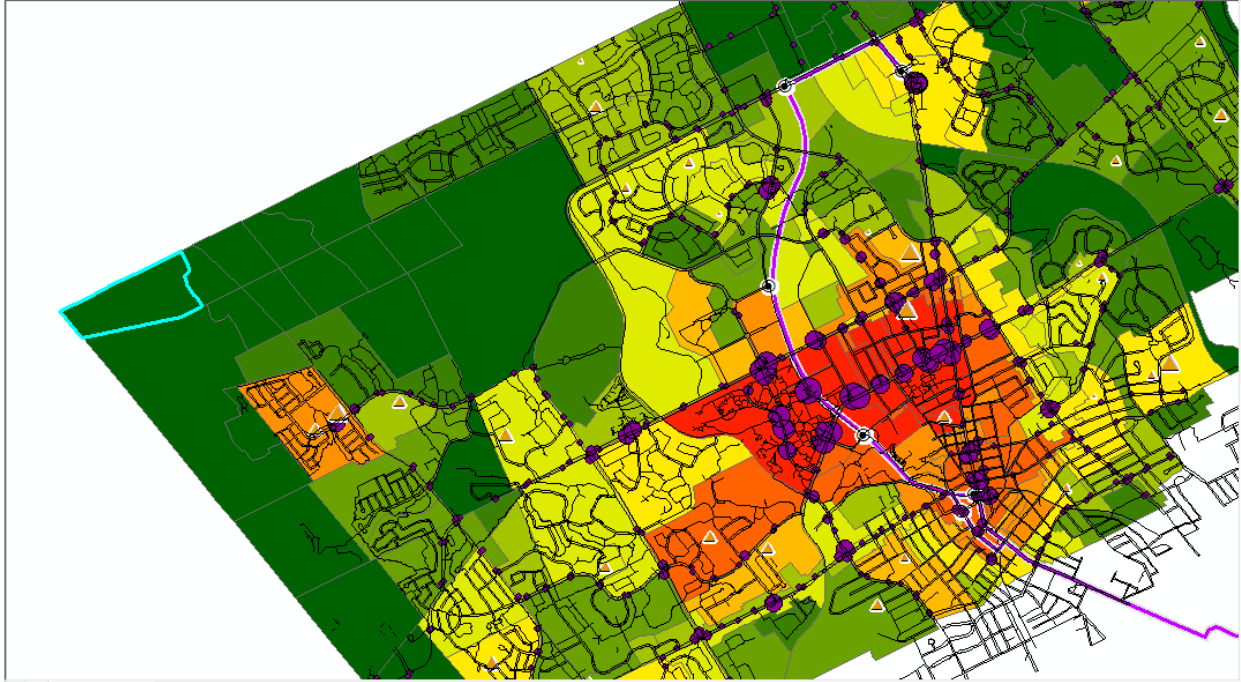


Figure 4.2.3: A PLUM zone (P.A.M. = -123) as Classified in the Lowest Predicted Trips class (-123 - -47)

Lowest Predicted Trips Class (-123 - -47)

If the y-intercept is removed, the zones in this class will have a minimum of zero trips. Logically, it is hard to say there are zero walking trips as people walk all the time as a secondary mean to get around or for recreational purposes. TTS reported daily walking trip counts capture only trips made solely by walking with no distinction between utilitarian or recreational purpose trips (Data Management Group, 2011). Since the model is built on the TTS daily walking trip data as the response variable, it inherits the same definition for predicted walking trips.

The highlighted zone, as in Figure 4.2.3, shows Pedestrian Activity Model (P.A.M.) trip counts of -123, which is the lowest predicted trips for the entire city. Figure 4.2.3 shows the spatial data for the three variables incorporated in the P.A.M. As observed, the selected zone is far off from any sidewalk network, transit stops, and schools, which means zero values per each of the variables. As a result, this zone scored the lowest predicted pedestrian trips. In addition, the land

use map as per the city's Official Plan marked these low pedestrian activity zones as rural areas, low-density residential, and open space (City of Waterloo, 2016). The land use findings are typical for an urban sprawl development, which carries association to the car-dependent environment.

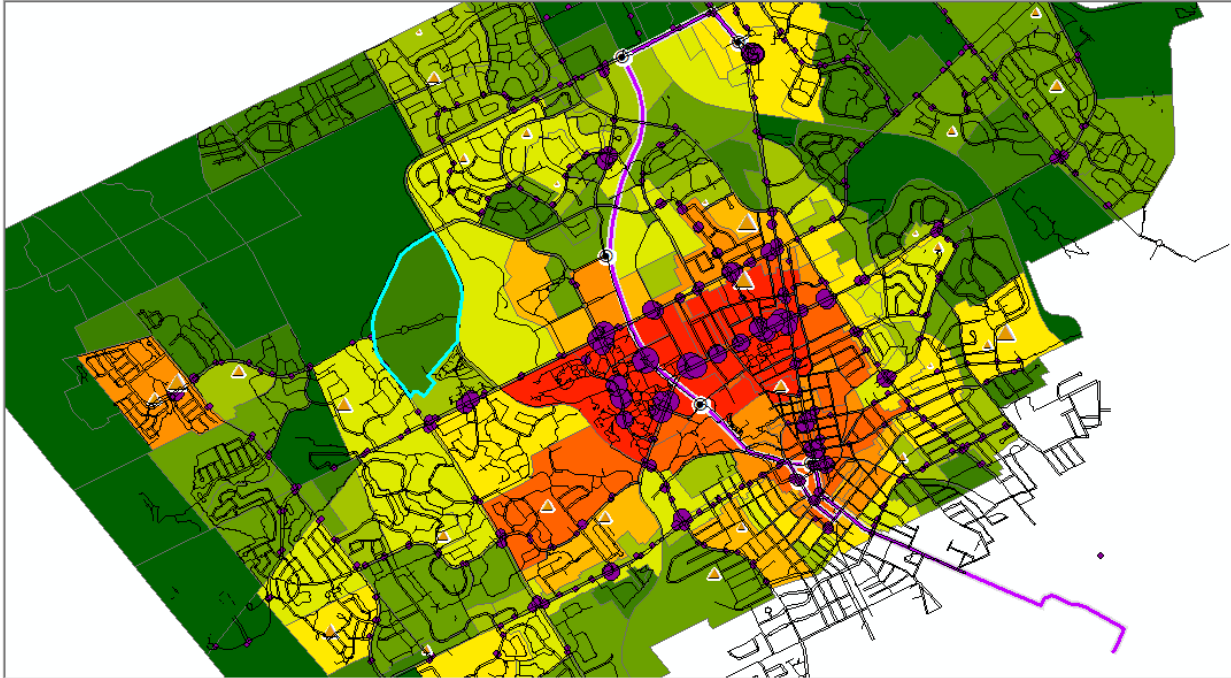


Figure 4.2.4: A PLUM zone (P.A.M. = -4) as Classified in the Second Lowest Predicted Trips class (-46 - 34)

Second Lowest Predicted Trips Class (-46 – 34)

Figure 4.2.4 shows a PLUM zone part of the second lowest predicted trips class ($-46 < \text{P.A.M.} < 34$). Overall, for the second lowest class, the observed trend is low sidewalk network presence and proximity to transit stops along one or two bus routes with an overall low activity. None of the second lowest class zones contain either an elementary or a secondary school. A statistical analysis reveals that the mean Transit activity for this class is 763 (3.28%) daily transit users, while a zone in the highest P.A.M. class has access as many as 23,000 (99.79%) daily transit users. As mentioned in the Methods Chapter, the Transit Activity variable is collinear to population and employment density; therefore, minimum access to transit users is a response to low population density and lack of employment opportunities. The average Metric Reach is 7.8 km, which is 20% of the maximum Metric Reach per zone (39 km).

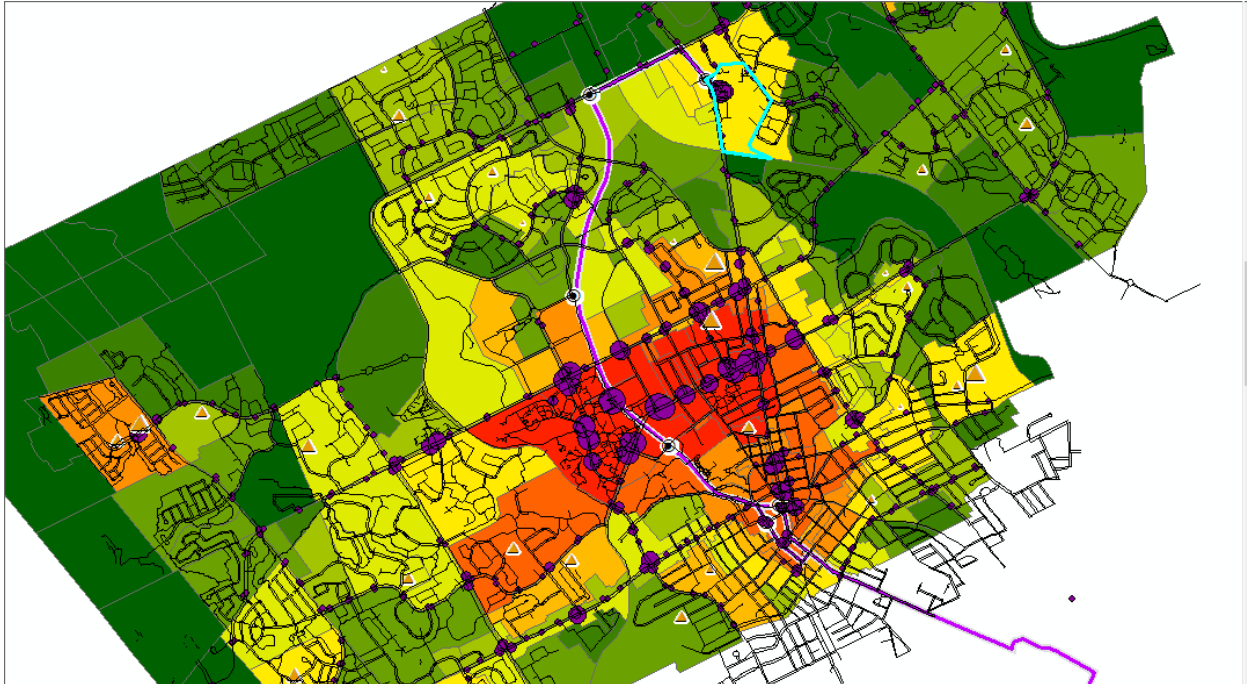


Figure 4.2.5: A PLUM zone (P.A.M. = 216) as Classified in the Moderate Predicted Trips class (203 - 280)

Moderate Predicted Trips Class (203 – 280)

While Transit Activity is the strongest of the three significant variables in influencing walking trips, it alone does not make an area vibrant with pedestrians. The proof is the Conestoga Mall zone, which, according to the map above (Figure 4.2.5), has 216 predicted daily trips and is part of the Moderate Predicted Trips Class (203 – 280). Conestoga Mall is considered a major commercial centre and a transit hub and even has its own ION station (LRT Line – not operational yet) (City of Waterloo, 2016; Region of Waterloo, 2011). Another finding particular to the Conestoga Mall is the low Metric Reach achievable (5 km) compared to the further reach of other lower-predicted trip class zones. Also, for a zone that is mainly commercially designated and strong transit presence (7,360 daily transit users), it has a low commercial floor area ratio (24%). This finding supports the claim the big box malls are car-oriented with low sidewalk presence and massive surface parking areas.

At this class, elementary and secondary schools start appearing in some of the zones. While the average student presence is about 4% (79 students); however, one of the zones contain both an elementary school (i.e., St. Agnes Catholic School) and a secondary school (i.e., Bluevale Collegiate), which combined have a high student enrollment (1648 students – 84%). Across all the zones within this class, Metric Reach is considered moderate with a mean of 46.96% (18.4 km), while the maximum recorded Metric Reach in the study is 39 km. On the other hand, the average transit users within and in proximity to this class is considerable low (14.05% = 3,267).

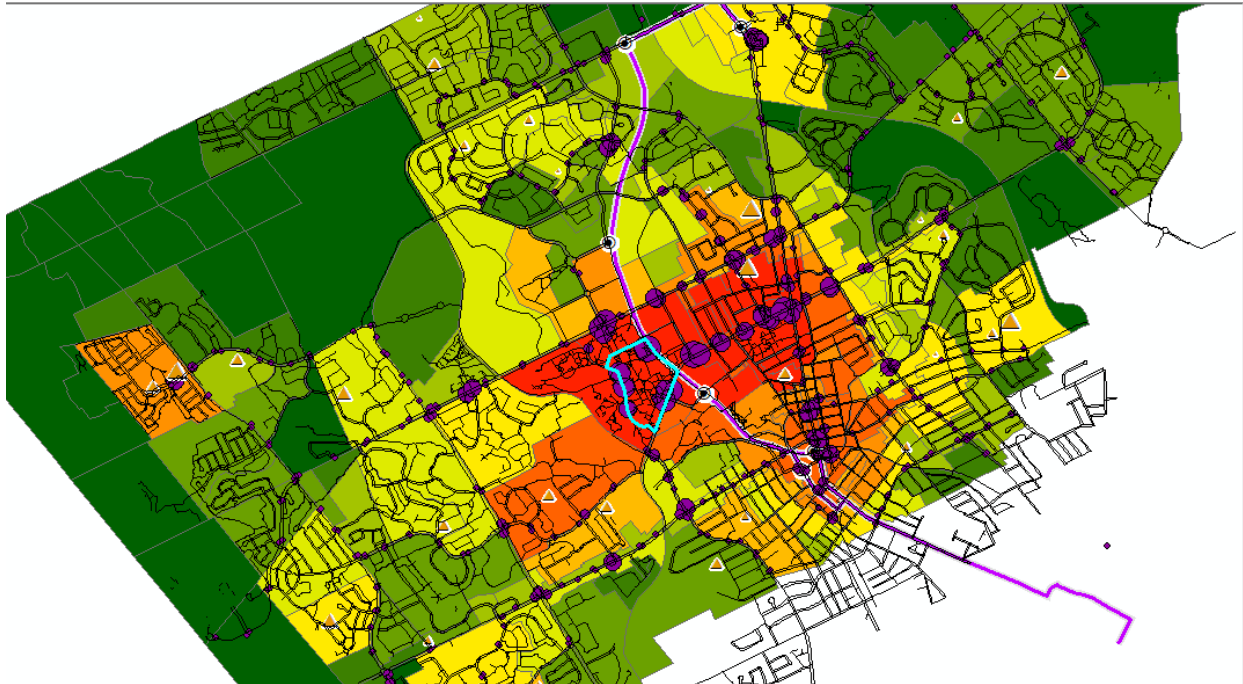


Figure 4.2.6: A PLUM zone (P.A.M. = 1046) as Classified in the Highest Predicted Trips class (640 - 1091)

Highest Predicted Trips Class (640 – 1091)

Zones part of highest predicted trips class are concentrated along University Avenue between the region's biggest post-secondary institutions (i.e., University of Waterloo and Wilfrid Laurier University) and have a predicted daily trip range of 640 to 1091. The key driver for pedestrian activity here is Transit Activity with a mean of 72% (16,739) and a maximum of 100% (23,249). It is important to note that the predicted daily walking trips are primary walking trips and does not include secondary walking trips to and from transit. P.A.M. inherits the same outcome as the response variable in the SER regression, which is primary daily walking trips. On the other hand, elementary and secondary school presence is low here but is offset by Transit Activity, which has slightly over two times bigger coefficient. An example is the University of Waterloo main campus, which does not have any elementary or secondary school institution but has access to an average of 23249 daily transit users and a moderate-high Metric Reach (67.8% = 26.5 km) (see

Figure 4.2.6). Another general finding pertaining to the class is the moderate Metric Reach (56.8% = 22.2 km). When the ION LRT line becomes operational, it is anticipated that some of the Transit Activity will shift, increasing pedestrian presence in zones along the ION's corridor.

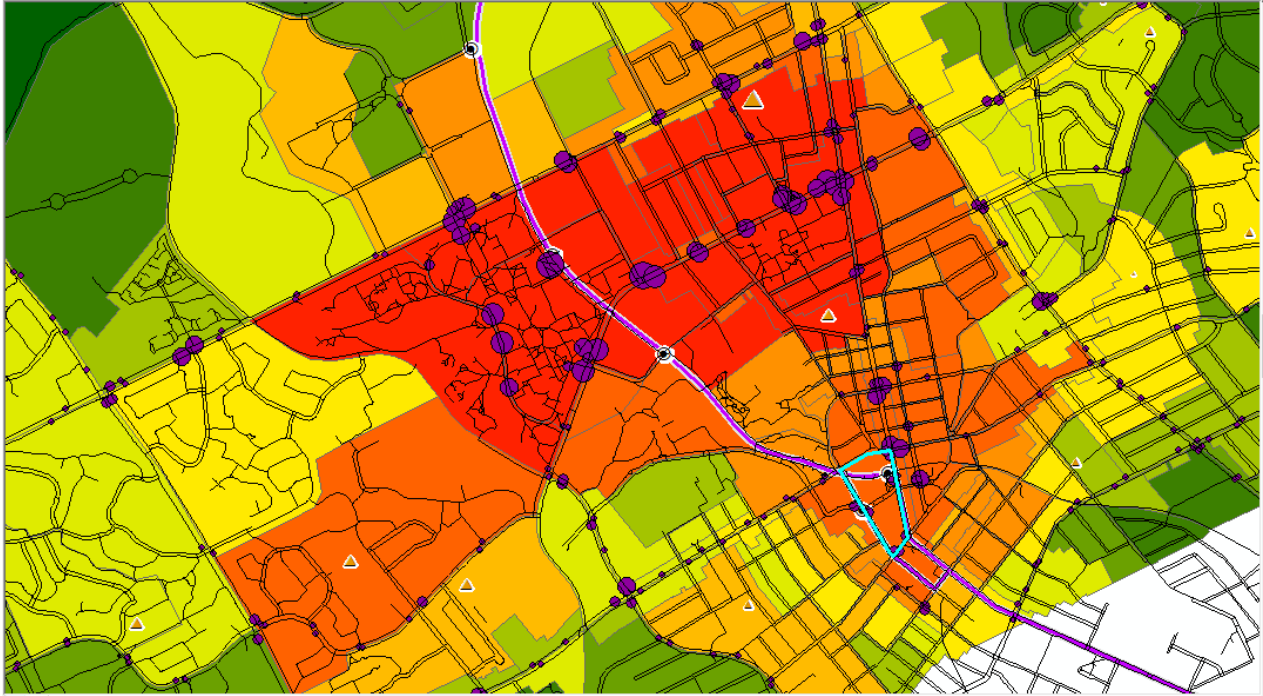


Figure 4.2.7: A PLUM zone (P.A.M. = 570) as Classified in the Second Highest Predicted Trips class (458 - 639)

Second Highest Predicted Trips Class (458 - 639)

While a few zones belonging to this class do not have a specific spatial distribution pattern, the remaining zones are clustered around the City of Waterloo’s Uptown. The selected zone above (Figure 4.2.7) is located at the heart of the Uptown. Zones located in and around the Uptown are different than other zones in terms of which of the variables is the key driver for predicting pedestrian activity. Figure 4.2.7 shows no elementary or secondary school presence in the Uptown and few bus stops with low-moderate activity. The selected zone has access to just 6,430 daily transit users. The key driver to pedestrian activity in the Uptown is predominantly Metric Reach. The grid street design, smaller blocks, and dual-sidewalk streets allow higher connectivity and accessibility for pedestrians moving in, out and about the Uptown. Zones belonging to this class and are located around the Uptown have a mean Metric Reach of 32.49 km and the selected zone has a Metric Reach of 36.2 km.

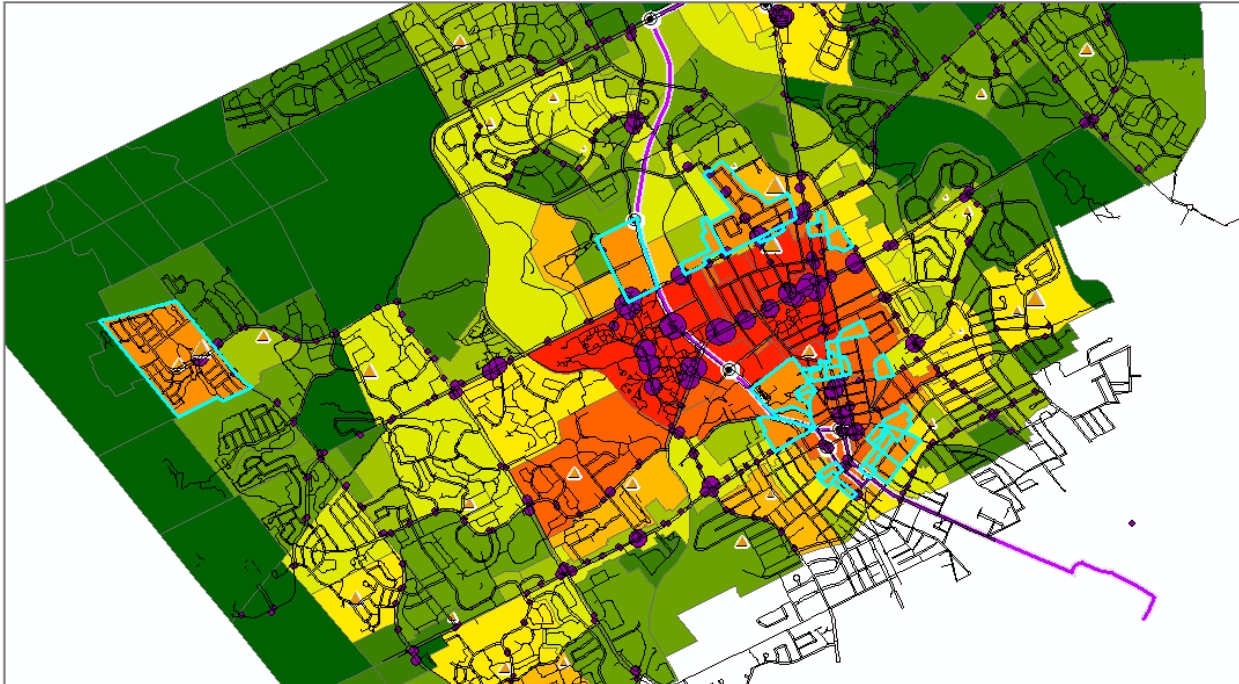


Figure 4.2.8: PLUM zones as Classified in the Third Highest Predicted Trips class (368 - 457)

Third Highest Predicted Trips Class (368 - 457)

Overall, there is no clear distinction for the make up of any specific zone(s) within this class.

Metric reach appears to be moderate-high with a mean of 26.24 km (67.14%). Transit activity is low with an average daily access to 5,106 transit users (21.96%). Only one zone has student enrollment and contains both an elementary and a secondary school.

On the other hand, this class as a whole plays two important roles in this Pedestrian Activity Index and identifying Pedestrian Priority Zones. The first role is completing the picture. The highest and second highest predicted trips classes have in-between spatial gaps and some zones do not align with the arterial street network. Including zones in this class allows the formation of one holistic area extending from the post-secondary institutions along University Avenue and down King St. to Uptown.

The second role is identifying zones with potential pedestrian activity growth. Located on the North East corner is a remote zone with a potential for pedestrian activity growth. This zone is home to both Abraham Erb Public School and Sir John A. Macdonald Secondary School. The combined student enrollment for that zone is 1957 students. Remote zones located on city peripheries tend to be car-centric; however, most students walk out of necessity. This emphasizes the role of schools in influencing walking activity in the neighbourhood. Pedestrian activity is anticipated to grow with further investment in transforming these zones to become pedestrian-focused.

Table 4.2.1: Analysis of the Pedestrian Activity Model's Natural Breaks highest 4 classes

Natural Breaks				
Highest Classes	Highest 1 Class	Highest 2 Classes	Highest 3 Classes	Highest 4 Classes
# of PLUM zones (336 Total)	24	53	85	105
Area (Hectares)	249.11	494.77	699.62	992.75
City of Waterloo Total PZs' areas (Hectares)	6514.81	6514.81	6514.81	6514.81
Area Percentage	3.82%	7.59%	10.74%	15.24%
Covered Accumulative Area Percentage	0.00%	198.62%	280.85%	398.52%
Class's Trips Maximum	1091			
Class's Trips Minimum	640	458	368	281
Actual Trips Minimum	676	475	371	294
Trips Total Per Selected Classes	19304	34822	47791	54422
Total Trips	73791			
Covered Accumulative Trips Percentage	26.16%	47.19%	64.77%	73.75%
Accumulated Change (%)	0	180.39%	247.57%	281.92%
Existing SW snow clearing Network (km)	143.8			
Total associated gross SW Network (km)	56.73	104.64	152.79	205.482
Private associated University SW Network (km)	25.35	25.35	25.46	25.94
Total associated net SW Network (km) (excluding UW & WL, & existing SW snow clearing network)	28.58	65.54	108.39	153.28
Total SW Network	716.674			
NET SW Network percentage of total network	3.99%	9.15%	15.12%	21.39%
Accumulated Change (%)	0	129.32%	279.25%	436.32%
Total New SW clearing Network (km)	172.38	209.34	252.19	297.08
Total New SW clearing Network (%)	24.053%	29.210%	35.189%	41.453%

4.3. Pedestrian Priority Zone Configuration

Modeling pedestrian activity is the foundation of many applications that require an understanding of the spatial distribution for pedestrian demand. For municipalities especially the public works department, pedestrian demand could be the next key criterion to prioritize active transportation-related projects, such as: extending the sidewalk network or providing sidewalk snow clearing or improving the streetscape. As mentioned in the Introduction Chapter, the research gap exists in bridging the fields of pedestrian modeling and providing sidewalk snow clearing. To address the existing gap, I propose identifying priority zone(s) based on analysis of the constructed Pedestrian Activity Model.

The existing sidewalk snow clearing network is about 143.8 kilometers spread across the City. As mentioned in the introduction chapter, the City of Waterloo clears snow off sidewalks adjacent to public properties (City of Waterloo, 2009). Suggesting a Pedestrian Priority Zone means expanding, not reconfiguring the existing sidewalk snow clearing network. The existing network still needs to be cleared by the City as it does not fall within other's jurisdiction.

Table 4.2.1 shows the statistical analysis of various priority zones based on cumulative combinations of the highest 4 classes of the Pedestrian Activity Model as shown in Figure 4.3.1 and Figure 4.3.2. The highest class of pedestrian activity is made up of 24 PLUM zones containing both University of Waterloo's main campus and Wilfrid Laurier University and properties in between along University Avenue. While these 24 PLUM zones make up less than 5% of the City of Waterloo's total area, they host approximately 25% of the predicted daily walking trips. Associated public sidewalks and pedestrian crossing make 4% of the total city's public sidewalks and pedestrian crossings. If this class is considered as a priority zone, it would

imply increasing the existing sidewalk snow clearing network (143.8 km) to be 172.4 km and ensure safer winter pedestrian travel for 25% of the predicted daily trips across the city.

Further statistical analysis of the all 3 highest classes combined provides better accommodation to pedestrian demand. The associated area for the all 3 classes combined makes up only 11% of the City of Waterloo's total. Although the associated trip activity is as much as 65% of total estimated daily walking trips for the city, this much activity takes place only on 15% of the total length of public sidewalks and pedestrian crossings. While the highest class of pedestrian activity only captures zones around post-secondary institutions, the second class extends south towards the Uptown but does not necessarily captures all PLUM zones in proximity to Uptown. The third class of pedestrian activity fills in the gaps left by the first and second classes adding to the captured walking trips as much as 18% (see Table 4.2.1 and Figure 4.3.1). In addition, the third class captured a remote neighbourhood with potential pedestrian activity growth as discussed above.

When considering priority sidewalk snow clearing, it is important to allocated plows and other winter maintenance to zones with the most impact to justify the cost-benefit analysis case. For this reason, it is recommended to combine and use all 3 highest classes to form a single Pedestrian Priority Zone to deliver demand-based sidewalk snow clearing services. The priority zone can be adjusted to meet budget and operation restrictions as seen fit by the public works staff.

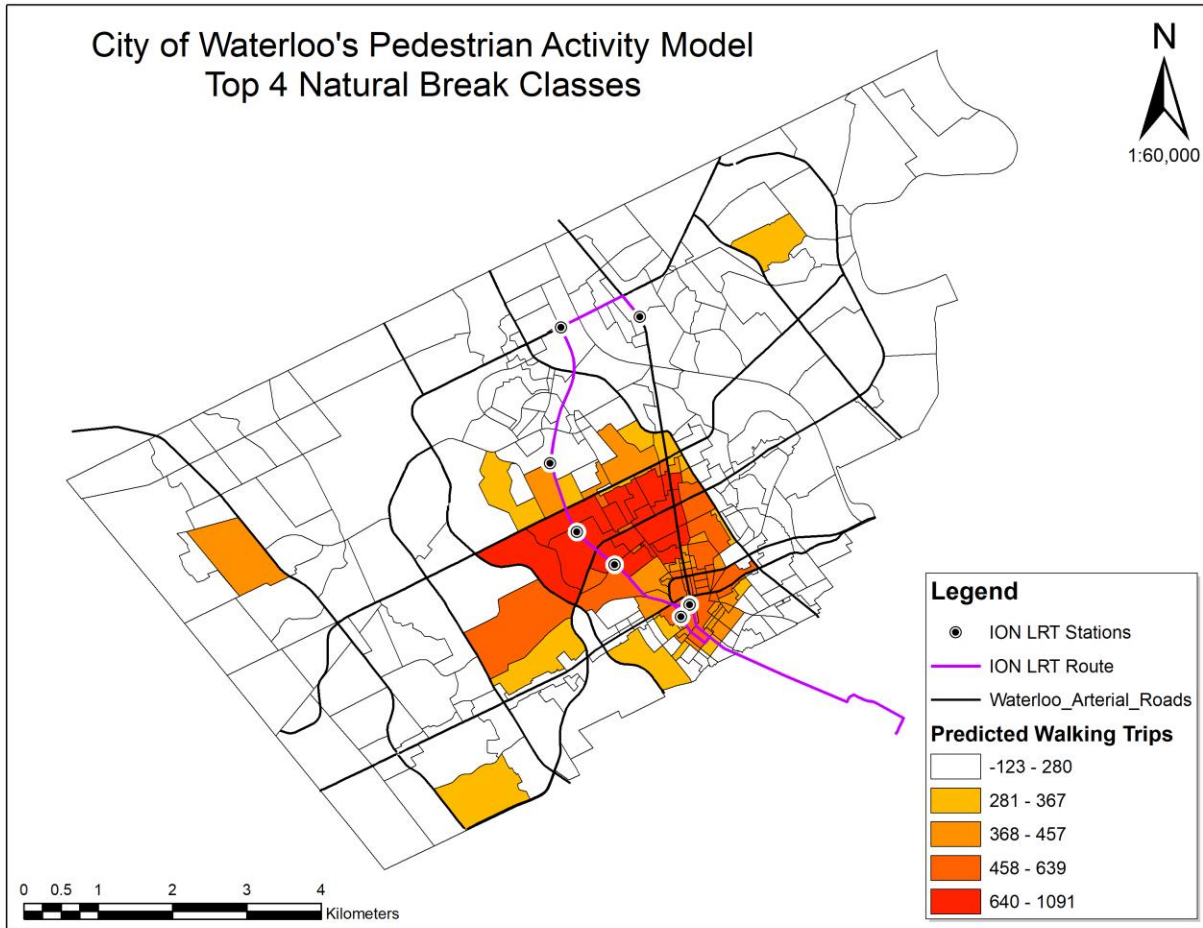


Figure 4.3.1: City of Waterloo's Pedestrian Activity Model's Highest 4 Natural Break Classes

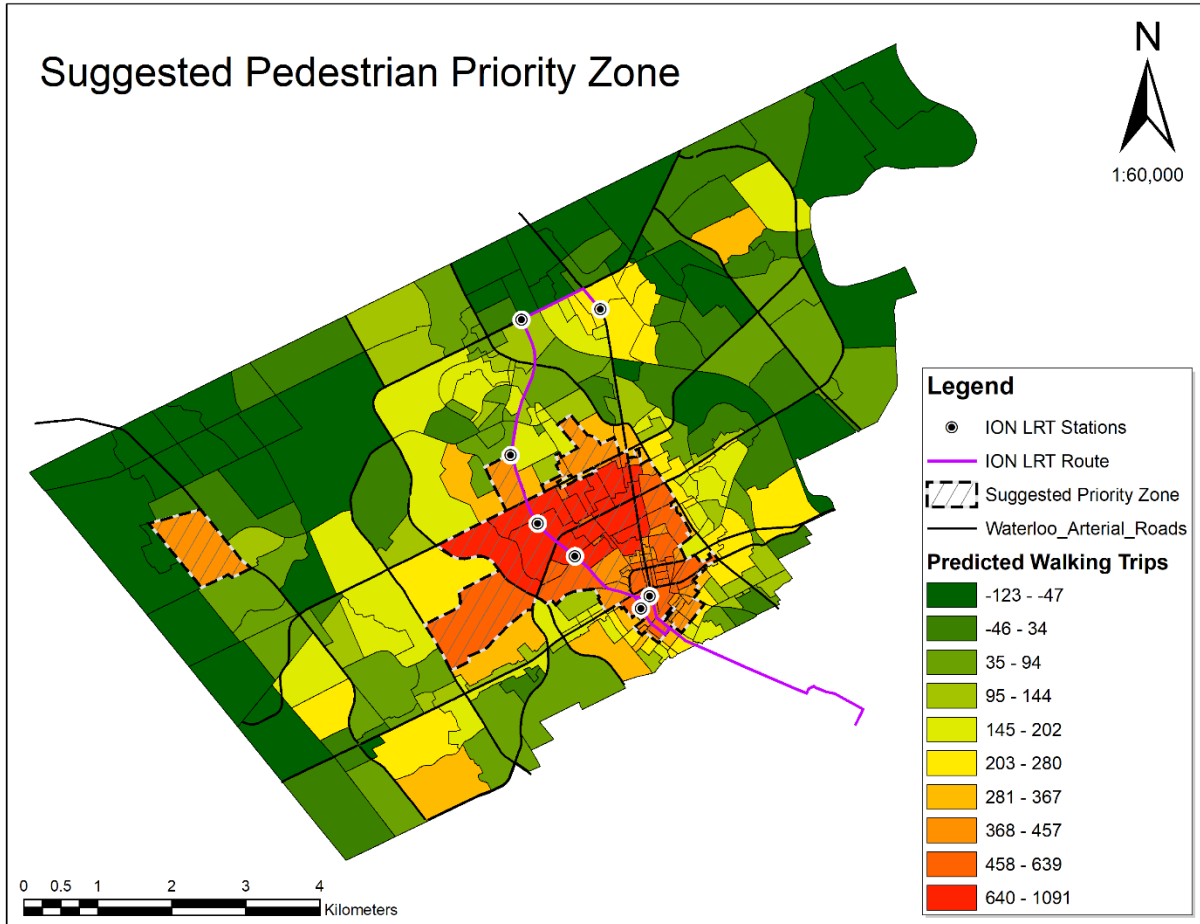


Figure 4.3.2: City of Waterloo's Suggested Pedestrian Priority Zone

4.4. Research Implication

While the findings have lots of empirical implications, it also has theoretical implications.

Starting with the land use mix's entropy and interaction lines measures, which according to the Spatial Error Regression, are insignificant. Could that mean that land use mix is not truly representative of pedestrian activity and that pedestrian activity is not influenced by mixed use development? There are two counterpoints to this notion. The first pertains to the data quality of the land use data. Land use data in this study were adopted from the Official Plan, which aggregated minor uses under the major use of the property. As a result, the land use data is coarse and does not provide a detailed view of the land use mix. The second counterpoint is that there is the possibility that having Transit Activity in the regression is causing land use mix measures to be insignificant. Transit Activity is demand representative variable that is collinear to population and employment density (Cervero and Kockelman, 1997; Ewing and Cervero, 2001; Liu and Griswold, 2009; Ewing and Cervero, 2010; Meghelal and Capp, 2011; Tsiompras and Photis, 2016). These same variables are commonly collinear with land use mix. As a result, Transit Activity is considered to be collinear to land use mix measures. Multicollinearity could be the reason for why land use mix measures were insignificant in the regression model.

Another implication to this research pertains to the local winter cycles. Depending on the location of the application, micro climate might differ. Municipalities located in the southern parts of the Province of Ontario are prone to numerous freeze and thaw cycles during the winter months. As a result, these municipalities tend to suffer from ice formation rather than snow accumulation on sidewalks that still hinder winter mobility and increase slips and falls risk. The implications for these different local winter cycles is to adjust the winter sidewalk maintenance applications to focus more on delivering an appropriate treatment plan (e.g., chemical treatment

to prevent ice formation). In this research, the term sidewalk snow clearing is used because of its commonality and it is not just exclusive to just mechanical treatment but it is meant to represent all appropriate winter sidewalk maintenance treatments. It is the responsibility of the local municipality to determine and deploy the appropriate treatment depending on their local winter cycle.

Chapter 5: Conclusion

The goal of this research is to develop an analytical approach to prioritize winter maintenance of sidewalks. This is manifested through two research objectives. The first objective is to predict the spatial distribution of pedestrian activity. The second objective is to identify a Pedestrian Priority Zone to potentially re-direct and focus pedestrian-related investment and services like sidewalk snow clearing.

5.1. Conclusion

This study concludes through the development of a Pedestrian Activity Model (P.A.M.) predicting daily walking trips to be used as foundation for prioritizing pedestrian-related investment. The strength in using the P.A.M. is in its simple construct and outcome, and in addressing the shortcomings identified, in the Chapter 2: Literature Review, for the approaches for estimating pedestrian activity. The shortcomings include:

- Mismatched geographical unit of analysis to represent pedestrian activity
- Aggregate data to accommodate analysis for other transportation modes
- Too many explanatory variables or just a single explanatory variable included
- Included insignificant explanatory variables
- Unweighted explanatory variables
- Lack of approach validation
- Required statistical expertise to interpret data and outcome

The Pedestrian Activity Model addresses these shortcomings in one way or another as highlighted in both Chapter 3: Methods, and Chapter 4: Findings and Discussion. Based on the

regression analysis, the Spatial Error Regression (SER) model has a better fit than the Ordinary Least Squares Regression, explaining additional 3% of walking activity. Of the 7 considered variables, only three – Transit Activity, Metric Reach, and Elementary and Secondary School Student Enrollment – have significant correlation (90% confidence) to walking activity. Transit Activity has the most explanatory powers in all three variables. One unit increase in the explanatory variable Transit Activity equals two times increase in walking trips versus one unit increase in the explanatory variable Metric Reach or equals approximately two and a half times increase in walking trips versus one unit increase in the explanatory variable Student Enrollment.

The Pedestrian Activity Model (P.A.M.) is constructed based on the SER's predictive model. Predicted daily walking trips are classified using natural breaks into ten classes to showcase the variation and the spatial distribution of pedestrian activity. The highest predicted walking trip class highlights PLUM zones along University Avenue between University of Waterloo and Wilfrid Laurier University. The second highest predicted walking trip class encompasses Uptown.

The existing sidewalk snow clearing network for the City of Waterloo is 143.8 km, which includes sidewalks adjacent to public properties. Using P.A.M., sidewalk snow clearing network should expand in high foot traffic areas. The criteria for analyzing highest predicted walking trips classes, is to suggest minimal sidewalk network addition for the maximum inclusion of daily walking trips. The highest three predicted walking trips classes make up the suggested Pedestrian Activity Zone. The new suggested zone encompasses approximately 110 km of sidewalk and pedestrian crossings and facilitates 65% of predicted daily walking trips. The new sidewalk snow clearing network is 250 km long.

Despite P.A.M.'s many advantage points, it has challenges and limitations too. The first challenge pertains to some of the variables initially considered. As discussed in Chapter 3: Methods under the Geographical Unit of Analysis sub-section, some variables – Population Density, and Employment Density – were not available at the lowest disaggregate level. Transforming these variables would cause a modifiable areal unit problem (MAUP). In addition, the literature review revealed that these same variables tend to be collinear to other consider factors like Transit Activity (see Appendix A).

Another initially considered variable was Intersection Density, however Metric Reach provides the opportunity to measure both sidewalk connectivity and presence in each area. A recent study Ellis et. al. (2016) found that Intersection Density and Metric Reach have strong correlation to walking activity. As a result, in this thesis Metric Reach substitutes Intersection Density in measuring the pedestrian infrastructure's (i.e. sidewalks) connectivity.

The second challenge pertains to the response data - TTS reported daily walking trips. The TTS reporting on walking activity is limited on trips sole made by walking and is unavailable for about one third of the Traffic Analysis Zones (TAZ) in the City of Waterloo (29 out of 104). In addition, the TTS survey design is biased towards private low rise and low-density households (Data Management Group, 2018b). By default, this under-represents apartment building tenants. The survey also depends on a single household member or the landlord to remember and report trip activity for all household members, which poses a challenge in capturing trip activity by big family households and multiple tenant rentals. This represents a challenge especially in a university community, such as the City of Waterloo.

The third challenge pertains to the sample area used in all regression models to find the best fit model and to configure the Pedestrian Activity Model. The size criteria used to define the sample

area excludes downtown PLUM zones because of their typical smaller area sizes. This potentially explains underrepresentation of pedestrian activity in the predictive model.

5.2. Further Research

Building on the model's limitations, this sub-section is dedicated for steps to improve the P.A.M.'s predictive powers. Starting with the reference data used to calibrate the model, the TTS data was criticized for many shortcomings through survey design and underrepresenting pedestrian activity as a part of other modal travel and transportation choices made by apartment tenants and students. The desired alternative is to gather primary data via a new survey, or conduct counts of passing pedestrians in sample areas and intersections, or a mix of both data collection methods. An improved reference data would improve P.A.M.'s predictive accuracy.

The second technical obstacle to overcome is the land use data quality. Land use measures were found insignificant in the conducted regression analyses. Data quality is part of that problem. Regression analysis tried finding correlation between reported walking trips versus coarse land uses. The Official Plan groups minor land uses under each site's primary use, which underrepresents the City's true land use mix. Detailed land use data is essential not just to calculate land use diversity but to also measure accurately commercial/retail floor area ratio.

5.3. Recommendations

The next step following suggestion of the Pedestrian Priority Zone is realizing the P.A.M.'s potential application in winter sidewalk maintenance. This would be manifested through transforming the priority zone into a priority plowing route, and differentiating the status quo level of service (LOS) from that associated with the priority plowing route. Under the new Ontario Minimum Standards regulations, the required level of service (LOS) for winter sidewalk maintenance are set to the minimum, which most municipalities' status quo winter sidewalk maintenance LOS already adheres to. This process, for configuring a second level of service, includes determining an appropriate snow and ice clearing method, frequency for snow clearing and patrolling, service triggers, and time of completion. Each municipality has unique context and limitations (e.g., staff, or budget) which mean that each would adopt a priority-based winter sidewalk maintenance differently.

Bibliography

1. Batty, M. (2001). Agent-Based Pedestrian Modeling. *Environment and Planning B: Planning and Design*, 28(3), 321-326.
2. Blomberg, J., Jordan, G., Killingsworth, R., and Konheim, C. (2000). Pedestrian Transportation: A Look Forward. *Transportation Research Board (TRB)*. Retrieved from: <https://trid.trb.org/view/639798>
3. Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: A literature review. *Transport Reviews*, 33(1), 71-91. 10.1080/01441647.2012.747114 Retrieved from http://resolver.scholarsportal.info/resolve/01441647/v33i0001/71_ioewoibipalr
4. Brown, Yamada, Smith, Zick, Kowaleski-Jones, & Fan. (2009). Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity. *Health and Place*, 15(4), 1130-1141.
5. Cervero R., and K. Kockelman. (1997). Travel demand and the 3 Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment* 2 (3): 199-219.
6. Chen, L. (2012). Agent-based modeling in urban and architectural research: A brief literature review. *Frontiers of Architectural Research*, 1(2), 166-177.
7. City of Halifax. (n.d.). Winter Operations. Retrieved on March, 2017. Retrieved from: <http://www.halifax.ca/snow/winterops.php>
8. City of Kitchener. (September, 2016). Staff Report – Winter Sidewalk Maintenance. Retrieved from: <https://lf.kitchener.ca/WebLinkExt/DocView.aspx?id=1481953&searchid=5538e315-1d20-43bf-83a3-6459809e04dc&dbid=0&cr=1>
9. City of Markham. (n.d.). How Markham Manages Snow Removal. *City of Markham*. Retrieved from: <https://www.markham.ca/wps/portal/Markham/Residents/RoadsParking/RoadSidewalkMaintenance/WinterRoadServices/howmarkhammanagessnowremoval/>
10. City of Ottawa. (n.d.) Winter Maintenance. Retrieved on March, 2017. Retrieved from: <http://ottawa.ca/en/residents/transportation-and-parking/road-and-sidewalk-maintenance/winter-maintenance#snow-removal-and-disposal>

11. City of Thunder Bay. (2017). Walkability and Pedestrian Safety in Thunder Bay. *City of Thunder Bay*. Retrieved from:
<http://www.tbdhu.com/sites/default/files/files/resource/2017-11/Walkability%20and%20Pedestrian%20Safety%20in%20Thunder%20Bay.pdf>
12. City of Toronto. (2016). Preventing Injuries from Wintertime Slips and Falls in Toronto. *City of Toronto: Staff Report*. Retrieved from:
<https://www.toronto.ca/legdocs/mmis/2016/hl/bgrd/backgroundfile-97431.pdf>
13. City of Toronto. (April 2014). Staff Report – Confirmation of Levels of Service for Winter Maintenance of Bikeways, Windrow Opening, Sidewalk and AODA Compliance.
14. City of Waterloo. (2016). Official Plan. City of Waterloo. Retrieved from:
<https://www.waterloo.ca/en/government/officialplan.asp>
15. City of Waterloo. (2009). By-Law No. 9-156: A By-Law to Regulate the Removal of Snow and Ice from Sidewalks and Roofs. Retrieved from:
http://www.waterloo.ca/uploads/94/2009_156_Removal_Snow_Ice_Sidewalks_Roofs_Bylaw.pdf
16. Clifton, Singleton, Muhs, & Schneider. (2016). Representing pedestrian activity in travel demand models: Framework and application. *Journal of Transport Geography*, 52, 111-122.
17. Clifton, Livi Smith, & Rodriguez. (2007). The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning*, 80(1), 95-110.
18. Coleman, P. (Jan-Feb 2010). The Livable Winter Community. *The Review* 83 (1).
19. Dannenberg, A., Cramer, T., & Gibson, C. (2005). Assessing the Walkability of the Workplace: A New Audit Tool. *American Journal of Health Promotion*, 20(1), 39-44.
20. Data Management Group. (2018a). Transportation Tomorrow Survey 2016: Data Retrieval. *University of Toronto*. Retrieved from:
<https://dmg.utoronto.ca/idrs/drsQuery/tts>
21. Data Management Group. (2018b). TTS 2016: Data Expansion and Validation. *University of Toronto*. Retrieved from:
http://dmg.utoronto.ca/pdf/tts/2016/2016TTS_DataExpansion.pdf

22. Data Management Group. (2011). Transportation Tomorrow Survey 2011: Design and Conduct of the Survey. *University of Toronto*. Retrieved from: <http://dmg.utoronto.ca/pdf/tts/2011/conduct2011.pdf>
23. Davidson, Donnelly, Vovsha, Freedman, Ruegg, Hicks, . . . Picado. (2007). Synthesis of first practices and operational research approaches in activity-based travel demand modeling. *Transportation Research Part A*, 41(5), 464-488.
24. Day, Boarnet, Alfonzo, & Forsyth. (2006). The Irvine–Minnesota Inventory to Measure Built Environments: Development. *American Journal of Preventive Medicine*, 30(2), 144-152.
25. Duncan, Aldstadt, Whalen, Melly, & Gortmaker. (2011). Validation of Walk Score® for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas. *International Journal of Environmental Research and Public Health*, 8(11), 4160-79.
26. El-Geneidy, A., and Levinson, D. (2006). Access to Destinations: Development of Accessibility Measures. *Minnesota Department of Transportation*. Retrieved from: <https://www.lrrb.org/pdf/200616.pdf>
27. Ellis, G., Hunter, R., Tully, M., Donnelly, M., Kelleher, L., & Kee, F. (2016). Connectivity and physical activity: Using footpath networks to measure the walkability of built environments. *Environment and Planning B: Planning and Design*, 43(1), 130-151.
28. Eugster, W. (2007). Road and Bridge Heating Using Geothermal Energy – Overview and Examples. European Geothermal Congress. Retrieved from: <https://pangea.stanford.edu/ERE/pdf/IGAstandard/EGC/2007/147.pdf>
29. Ewing, Reid, & Cervero, Robert. (2010). Travel and the built environment: A meta-analysis.(Report). *Journal of the American Planning Association*, 76(3), 265-294.
30. Ewing, & Cervero. (2001). Travel and the built environment - A synthesis. *Land Development And Public Involvement In Transportation*, (1780), 87-114.
31. Fay, L., & Shi, X. (2012). Environmental Impacts of Chemicals for Snow and Ice Control: State of the Knowledge. *Water, Air, & Soil Pollution*, 223(5), 2751-2770.
32. Frank, L., Sallis, J., Saelens, B., Leary, L., Cain, K., Conway, T., & Hess, P. (2010). The development of a walkability index: Application to the Neighborhood Quality of Life Study. *British Journal of Sports Medicine*, 44(13), 924-33.

33. Geurs, & Van Wee. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, 12(2), 127-140.
34. GIS Region of Waterloo. (n.d.). Layer: PLUM Zones. Retrieved from: <https://gis.region.waterloo.on.ca/arcgis/rest/services/Walkability/MapServer/7>
35. Kerridge, Hine, & Wigan. (2001). Agent-based modelling of pedestrian movements: The questions that need to be asked and answered. *Environment And Planning B-Planning & Design*, 28(3), 327-341.
36. Knaap, G., & Talen, E. (2005). New Urbanism and Smart Growth: A Few Words from the Academy. *International Regional Science Review*, 28(2), 107-118.
37. Kuemmel, D., & National Cooperative Highway Research Program (NCHRP). (1994). *Managing roadway snow and ice control operations* (Synthesis of highway practice. 207). Washington, D.C.: National Academy Press.
38. Hajrasouliha, A., & Yin, L. (2015). The impact of street network connectivity on pedestrian volume. *Urban Studies*, 52(13), 2483-2497.
39. Hicks, J. (2016). "Conestoga to double size of Waterloo campus with \$43.5-million makeover." *The Record*. Retrieved from: <https://www.therecord.com/news-story/7027578-conestoga-to-double-size-of-waterloo-campus-with-43-5-million-makeover/>
40. Hosseini, F., Hossain, S., & Fu, L. (2017). Bio-based materials for improving winter pavement friction. *Canadian Journal of Civil Engineering*, 44(2), 99-105.
41. Hough Stansbury Woodland Ltd, Ontario. Ministry of Municipal Affairs, Sault Ste. Marie, Ontario Professional Planners Institute. Northern District, & Winter Cities Association. (1990). *Winter cities design manual*. Sault Ste-Marie: Ministry of Municipal Affairs: Northern District of Ontario Professional Planners Institute.
42. Hui, N., Saxe, S., Roorda, M., Hess, P., & Miller, E. J. (2018). Measuring the completeness of complete streets. *Transport Reviews*, 38(1), 73-95. 10.1080/01441647.2017.1299815 Retrieved from http://resolver.scholarsportal.info/resolve/01441647/v38i0001/73_mtcocs
43. Iacono, Krizek, & El-Geneidy. (2010). Measuring non-motorized accessibility: Issues, alternatives, and execution. *Journal of Transport Geography*, 18(1), 133-140.

44. Iwamoto, K. Nagasaka, S. Hamada, Y. Nakamura, M. Ochifuji K. and Nagano, K. (1998). Prospects of snow melting systems (SMS) using underground thermal energy storage (UTES) in Japan. In -Proceedings of the Second Stockton International Geothermal Conference 物pp. 243-251.
45. Jonnalagadda, N., Freedman, J., Davidson, W., & Hunt, J. (2001). Development of Microsimulation Activity-Based Model for San Francisco: Destination and Mode Choice Models. *Transportation Research Record: Journal of the Transportation Research Board*, 1777, 25-35.
46. Lee, & Moudon. (2006). The 3Ds + R: Quantifying land use and urban form correlates of walking. *Transportation Research Part D*, 11(3), 204-215.
47. Lee, C., & Moudon, A. (2004a). Activity -friendly Communities: Correlates of Transportation or Recreation Walking, and Correlates of Physical Activity for Different Sub-populations, *ProQuest Dissertations and Theses*.
48. Lee, C., & Moudon, A. (2004b). Physical Activity and Environment Research in the Health Field: Implications for Urban and Transportation Planning Practice and Research. *Journal of Planning Literature*, 19(2), 147-181.
49. Leslie, Coffee, Frank, Owen, Bauman, & Hugo. (2007). Walkability of local communities: Using geographic information systems to objectively assess relevant environmental attributes. *Health and Place*, 13(1), 111-122.
50. Li, X., Lv, Z., Zheng, Z., Zhong, C., Hijazi, I., & Cheng, S. (2017). Assessment of lively street network based on geographic information system and space syntax. *Multimedia Tools and Applications*, 76(17), 17801-17819.
51. Li, Y., Hsu, J., & Fernie, G. (2013). Aging and the use of pedestrian facilities in Winter— The need for improved design and better technology. *Journal of Urban Health*, 90(4), 602-617. doi:10.1007/s11524-012-9779-2
52. Lowrie, M. (2017). Montreal moves forward on heated sidewalk project. *CBC News*. Retrieved from: <http://www.cbc.ca/news/canada/montreal/montreal-moves-forward-on-heated-sidewalk-project-1.3956703>
53. Lund, Hollie. (2003). Testing the claims of new urbanism: Local access, pedestrian travel, and neighboring behaviors. *Journal of the American Planning Association*, 69(4), 414.

54. Luoma, Mikkonen, & Palomäki. (1993). The threshold gravity model and transport geography: How transport development influences the distance-decay parameter of the gravity model. *Journal of Transport Geography*, 1(4), 240-247.
55. Maghelal, P. K., & Capp, C. J. (2011). Walkability: A review of existing pedestrian indices. *URISA Journal*, 23(2), 5-19.
56. Manaugh, & El-Geneidy. (2011). Validating walkability indices: How do different households respond to the walkability of their neighborhood? *Transportation Research Part D*, 16(4), 309-315.
57. Manaugh, K., & Kreider, T. (2013). What is mixed use? Presenting an interaction method for measuring land use mix. *Journal of Transport and Land Use*, 6(1), 63-72.
58. Mänty, J., & Pressman, N. (1988). *Cities designed for winter*. Helsinki: Building Book.
59. Metrolinx. (n.d.). Active and Sustainable School Travel. Metrolinx. Retrieved from: http://www.metrolinx.com/en/projectsandprograms/schooltravel/school_travel.aspx
60. McNally, M. (2000). The Four Step Model. *University of California*. Retrieved from: <https://escholarship.org/uc/item/7j0003j0#main>
61. Millward, Spinney, & Scott. (2013). Active-transport walking behavior: Destinations, durations, distances. *Journal of Transport Geography*, 28, 101-110.
62. Minsk, L. (1998). *Snow and ice control manual for transportation facilities*. New York: McGraw-Hill.
63. Minsk, L. (1970). A Short History of Man's Attempts to Move through Snow. Special Report 115. *Highway Research Board*. pp. 1-7.
64. Miranda-Morena, L., and Lahti, A. (2013). Temporal trends and the effect of weather on pedestrian volumes: A case study of Montreal, Canada. *Transportation Research Part D*. (22). 54-59.
65. Moudon, A., & Lee, C. (2003). Walking and Bicycling: An Evaluation of Environmental Audit Instruments. *American Journal of Health Promotion*, 18(1), 21-37.
66. Nassiri, S., Bayat, A., & Salimi, S. (2015). Survey of Practice and Literature Review on Municipal Road Winter Maintenance in Canada. *Journal Of Cold Regions Engineering*, 29(3).

67. Neilson, K. (2018). "Kitchener council chooses to defer large portion winter sidewalk clearing review." *Global News*. Retrieved from:
<https://globalnews.ca/news/4297578/kitchener-city-council-sidewalks-snow/>
68. Oliver, S. (2014). Exploring the Impact of the Pedestrian Environment on Public Transportation: A Case Study of Waterloo Region. (Master's Thesis). *University of Waterloo, Waterloo, Ontario*.
69. Omer, & Kaplan. (2017). Using space syntax and agent-based approaches for modeling pedestrian volume at the urban scale. *Computers, Environment and Urban Systems*, 64, 57-67.
70. Ontario (Government of Ontario). (2018a). Reg. 239/02: Minimum Maintenance Standards for Municipal Highways. *Government of Ontario*. Retrieved from:
<https://www.ontario.ca/laws/regulation/020239>
71. Ontario (Government of Ontario). (2018b). Reg. 366/18: Minimum Maintenance Standards for Municipal Highways. *Government of Ontario*. Retrieved from:
<https://www.ontario.ca/laws/regulation/r18366>
72. Ontario (Government of Ontario). (2018c). Municipal Act, 2001. *Government of Ontario*. Retrieved from: <https://www.ontario.ca/laws/statute/01m25#BK58>
73. Ontario (Government of Ontario). (2017). The Growth Plan for the Greater Golden Horseshoe. *Government of Ontario*. Retrieved from:
http://placestogrow.ca/index.php?option=com_content&task=view&id=430&Itemid=14
74. Government of Ontario (Ontario) (2016). Accessibility for Ontarians with Disabilities Act. Retrieved from: <https://www.ontario.ca/laws/statute/05a11#BK9>
75. Ontario (Government of Ontario). (2014). Ontario Traffic Manual: Book 7. Retrieved from: https://ontario-traffic-council.s3.amazonaws.com/uploads/2018/04/cona-OTM_Book_7_-_Temporary_Conditions_Jan_20141.pdf
76. Ontario (Government of Ontario). (2013). Reg. 239/02: Minimum Maintenance Standards for Municipal Highways. *Government of Ontario*. Retrieved from:
<https://www.ontario.ca/laws/regulation/020239/v5>
77. Parolek, D., Crawford, K., & Parolek, Karen. (2008). Form-based codes: A guide for planners, urban designers, municipalities, and developers. Hoboken, N.J.: J. Wiley & Sons.

78. Penn, A. (2003). Space Syntax And Spatial Cognition: Or Why the Axial Line? *Environment and Behavior*, 35(1), 30-65.
79. Perrier, Langevin, & Campbell. (2007). A survey of models and algorithms for winter road maintenance. Part IV: Vehicle routing and fleet sizing for plowing and snow disposal. *Computers and Operations Research*, 34(1), 258-294.
80. Pickel, J. (2018). "Kitchener City Council defers bulk of sidewalk snow clearing pilot project." *KitchenerToday.com*. Retrieved from: <https://www.kitchenertoday.com/local-news/kitchener-city-council-defers-bulk-of-sidewalk-snow-clearing-pilot-project-966331>
81. Pressman, N. (1996). Sustainable winter cities: Future directions for planning, policy and design. *Atmospheric Environment*, 30(3), 521-529.
82. Preszler Injury Lawyers. (n.d.). What is the Law in Ontario for Clearing Snow and Ice from Your Premises? *Blog*. Retrieved on: January, 2018. Retrieved from: <https://www.preszlerlaw.com/blog/what-is-the-law-in-ontario-for-clearing-snow-and-ice-from-your-premises/>
83. Public Health Agency of Canada (PHAC). (2014). What is Active Transportation. *Public Health Agency of Canada*. Retrieved on December, 2016. Retrieved from: <http://www.phac-aspc.gc.ca/hp-ps/hl-mvs/pa-ap/at-ta-eng.php>
84. Public Health, City of Toronto. (2012a). Road to Health: Improving Walking and Cycling in Toronto. *City of Toronto*. Retrieved from: <https://www.toronto.ca/legdocs/mmis/2012/hl/bgrd/backgroundfile-46520.pdf>
85. Public Health, City of Toronto. (2012b). The Walkable City: Neighbourhood Design and Preferences, Travel Choices and Health. *City of Toronto*. Retrieved from: https://www1.toronto.ca/city_of_toronto/toronto_public_health/healthy_public_policy/hphe/files/pdf/walkable_city.pdf
86. Qiu, L., & Nixon, W. (2008). Effects of Adverse Weather on Traffic Crashes: Systematic Review and Meta-Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 2055(2055), 139-146.
87. Raford, N. (2010). "Social and Technical Challenges to the Adoption of Space Syntax Methodologies as a Planning Support System (PSS) in American Urban Design." *Journal of Space Syntax*. 1(1). ISSN: 2044-7507. Retrieved from: <http://joss.bartlett.ucl.ac.uk/journal/index.php/joss/article/view/28>

88. Raford, N., & Ragland, D. (2006). Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: The Case of Boston, Massachusetts. IDEAS Working Paper Series from RePEc, IDEAS Working Paper Series from RePEc, 2006.
89. Region of Waterloo. (2014). Staff Report: NEWPATH Research Project. *Region of Waterloo*. Retrieved from:
http://www.regionofwaterloo.ca/en/gettingAround/resources/P-14-021_PH-14-006__NEWPATH_RESEARCH_PROJECT.pdf
90. Region of Waterloo. (2009). Development of Walkability Index for Waterloo Region. *Region of Waterloo*. Retrieved from:
https://www.publichealthontario.ca/en/LearningAndDevelopment/Events/Documents/8%20DEVELOPMENT_OF_A_WALKABILITY_SURFACE_FOR_WATERLOO_REGION.pdf
91. Rutherford, G. (1979). Use of the Gravity Model for Pedestrian Travel Distribution. *Transportation Research Board*. 728. 53-59.
92. Shaaban, K. & Muley, D. (2016). Investigation of Weather Impacts on Pedestrian Volumes. *Transportation Research Procedia*, 14, 115-122.
93. Smithson, L. (2004). Implementing Snow and Ice Control Research. *Sixth International Symposium on Snow Removal and Ice Control Technology*. ISSN 0097-8515.
94. State of California (California). (2011). Article 4: Operation of Bicycles, Division 11: Rules of the road. *Vehicle Code*. Retrieved from:
http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=VEH&division=11.&title=&part=&chapter=1.&article=4.
95. Statistics Canada. (2018a). Labour Force Survey (LFS). Ottawa. Retrieved from:
<http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3701>
96. Statistics Canada. (2018b). Labour Force Survey documentation. Ottawa. Retrieved from:
http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvDocument&Item_Id=213696&InstaId=793002
97. Statistics Canada. 2017. Waterloo, CY, Ontario and Waterloo, RM, Ontario. Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Retrieved from: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E>

98. Statistics Canada. (2015). Census Dictionary: Dissemination Area (DA). *Statistics Canada*. Retrieved from: <http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo021-eng.cfm>
99. Student Transportation Services of York Region. (2018). Active School Travel Programs. Student Transportation Services of York Region. Retrieved from: <http://net.schoolbuscity.com/active-school-travel-programs>
100. Sullivan, J. (1831). Method of clearing the Baltimore rail-way of snow during the late winter. *American Journal of Science*. (20). (no. 24 of Miscellanies).
101. Sundquist, Eriksson, Kawakami, Skog, Ohlsson, & Arvidsson. (2011). Neighborhood walkability, physical activity, and walking behavior: The Swedish Neighborhood and Physical Activity (SNAP) study. *Social Science & Medicine*, 72(8), 1266-1273.
102. Thompson, C. (2018). "Kitchener plows ahead with sidewalk shovelling study." *The Record.com*. Retrieved from: <https://www.therecord.com/news-story/8088425-kitchener-plows-ahead-with-sidewalk-shovelling-study/>
103. Torrens, P. (2012). Moving Agent Pedestrians Through Space and Time. *Annals of the Association of American Geographers*, 102(1), 35-66.
104. Tracz, A. T. (2015). *Complete streets design elements and their impact on travel behavior: Learning from the bagby street reconstruction in Houston, TX* (Order No. 1594367). Available from ProQuest Dissertations & Theses Global. (1707657552). Retrieved from <http://search.proquest.com.proxy.lib.uwaterloo.ca/docview/1707657552?accountid=14906>
105. Transportation Association of Canada (TAC). (2013). Winter Maintenance Equipment and Technologies. Syntheses of Best Practices-Road Salt Management. Retrieved from: <http://www.tac-atc.ca/sites/tac-atc.ca/files/site/doc/resources/roadsalt-9.pdf>
106. TriTag Transport Action Group. (2018). Winter Sidewalk Study. *TriTag*. Retrieved from: <http://www.tritag.ca/static/uploads/TriTAG-Winter-sidewalk-report.pdf>
107. Trudeau, D. (2013), New Urbanism as Sustainable Development?. *Geography Compass*, 7: 435–448. doi:10.1111/gec3.12042

108. Tsiompras, & Photis. (2017). What matters when it comes to "Walk and the city"? Defining a weighted GIS-based walkability index. *Transportation Research Procedia*, 24, 523-530.
109. Tuner, S., Hottenstein, A., and Shunk, G. (1997). Bicycle and Pedestrian Travel Demand Forecasting: Literature Review. *Texas Transportation Institute and Texas Department of Transportation*.
110. Usman, Fu, & Miranda-Moreno. (2010). Quantifying safety benefit of winter road maintenance: Accident frequency modeling. *Accident Analysis and Prevention*, 42(6), 1878-1887.
111. Vanky, A., Verma, S., Courtney, T., Santi, P., and Ratti, C. (2017). Effect of weather on pedestrian trip count and duration: City-scale evaluations using mobile phone application data. *Preventive Medicine Reports*. (8). 30-37.
112. Walkscore.com. (n.d.-a). Research and Data Services. *Walkscore.com*. Retrieved from: <https://www.walkscore.com/professional/research.php>
113. Walkscore.com. (n.d.-b). Walk Score Methodology. *Walkscore.com*. Retrieved from: <https://www.walkscore.com/methodology.shtml>
114. Waterloo Town and Gown Committee. (2017). Waterloo Student Housing. *City of Waterloo*. Retrieved from: <https://www.waterloo.ca/en/contentresources/resources/government/CTTEE2017-004-Town-and-Gown-WSA-Update.pdf>
115. Web Centre for Social Research Methods. (2006). Measurement Validity Types.
116. Ye, Z., Veneziano, D., & Shi, X. (2013). Estimating Statewide Benefits of Winter Maintenance Operations. *Transportation Research Record*, (2329), 17-23.

Appendices

Appendix A – Considered Variables – Not Included in Final Model	128
Appendix B – Single Variable Regressions’ Scatterplots	133

Appendix A – Considered Variables – Not Included in Final Model

Net Population Density

Population density is among the most popular variables in walking and walkability studies (Cervero and Kockelman, 1997; Ewing and Cervero, 2001; Ewing and Cervero, 2010; Meghela and Capp, 2011; Tsiompras and Photis, 2016). Walkscore.com is a widely-known walkability index that uses population density to indicate the pedestrian friendliness of any property (Walkscore.com, n.d.-b). Population density is sometimes measured as gross density (population divided by area) and other times as net population density (ratio of population to a residential area).

The main data source in Canada for the population is the 2016 Census. The lowest disaggregate level, for which the population data is publicly available, is dissemination areas (DA). The data is in a polygon form representing dissemination areas. The land use data is available through the Geospatial Centre at the University of Waterloo for the City of Waterloo based on their Official Plan. It is also a polygon form data, with an attribute specifying the land use type (e.g., residential, commercial, employment, institutional,...etc.). Using residential land use area, net population density was calculated using the following formula:

$$= \frac{\textit{Population per DA}}{\textit{Residential Area per DA}}$$

As shown in the histogram below, the net population density is slightly skewed to the left as indicated by its tail. The net population density mean is 50 people/hectare. The maximum density is 162.97 persons per hectare, which is found at the northeast corner of King St. and University St. On the other hand, low residential density could be as low as 12.35 persons/hectare plus a no-

residential area bound by King St. and Weber St. north of the city. The low mean value can be explained by the common housing norm in Waterloo, which is single-detached, and also the Census survey is a private household-based survey, which explains low residential density DAs containing on-campus residences and student housing.

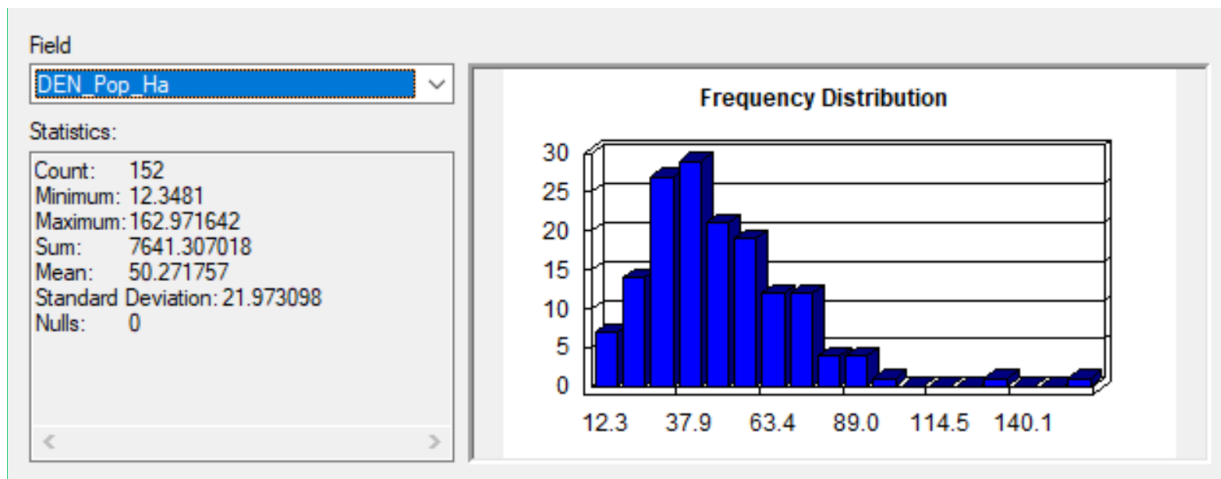


Figure A.1: Net Population Density's descriptive statistics (X-axis: variable's value and Y-axis: PLUM zones' count)

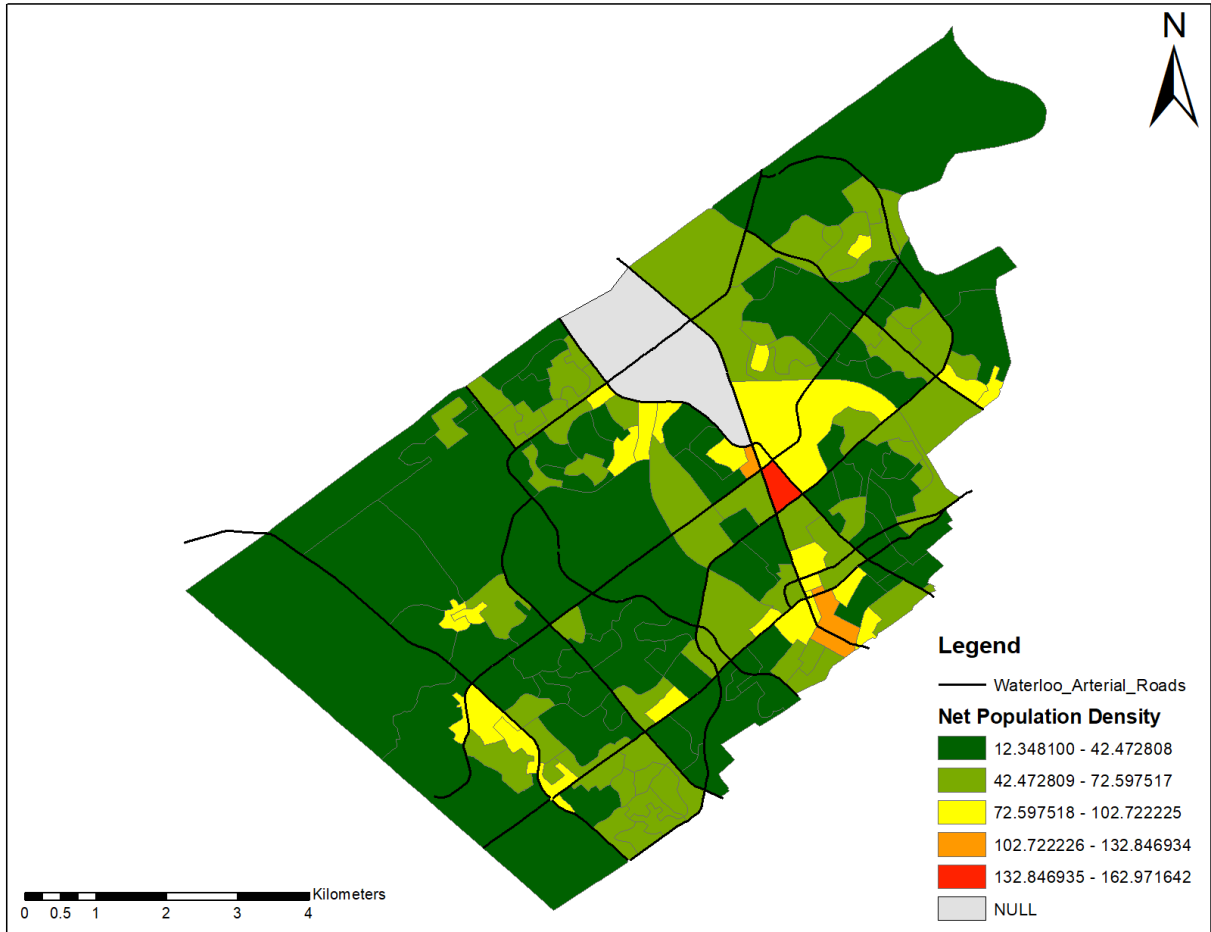


Figure A.2: Net Population Density's equal interval map by dissemination area (DA)

Employment Density

Similar to net population density, net employment density is common in walkability indices because of its correlation with walking (*Tsiompras and Photis, 2016*; Ewing and Cervero, 2010; Liu and Griswold, 2009). Employment density is an indicator of pedestrian demand as workers walk to surrounding commercial and retail areas during their lunch hour or to the transit stop at the end of the workday.

The 2016 Census only contains data such as how much of the labour force resides in each area and not actually employment location. For this type of data, the Labour Force Survey is the key source for employment data in Canada but publicly available data is only found at the aggregate level (e.g., Census Metropolitan Areas and Economic Regions) (Statistics Canada, 2018a; Statistics Canada, 2018b). Disaggregate level employment data is made available through the Transit Tomorrow Survey (TTS), which is available at the Traffic Analysis Zones (TAZ). As discussed in the Methods Chapter, TAZ was ruled out as an appropriate geographical unit of analysis, which had the implication of not including the employment density variable.

Unlike net population density, calculating net employment density is more complex. Employment overlaps with various land use types (e.g., mixed-use, residential, employment, commercial, and industrial lands). In residential and mixed-use areas, employees could be working from home, which adds to the complexity of whether residential areas should be included in the net employment density formula. In addition to the employment data's overlap across land uses and its complexity to calculate, the data is not commonly found at a suitable disaggregate level.

Connectivity – Intersection Density

Connectivity is a design feature that describes how accessible is a neighbourhood through sidewalk presence or density. Intersection density is a famous measure that falls under the connectivity umbrella and correlates with walking (Ellis et. al., 2016; Ewing and Cervero, 2010; Frank et. al., 2010). Intersection density is the number of 3-leg or more intersections per square kilometer (Ellis et. al., 2016). Previous studies have used the street network for calculating intersection density, assuming pedestrians walk along every route even where sidewalks are absent.

The active transportation infrastructure data is available through the Region of Waterloo and its municipalities' open data portal. The infrastructure data is a line-based network of sidewalks and trails. Despite data availability, there were multiple data layers and none contained the full sidewalk and trails network in the City of Waterloo. When I choose the spatial join tool to merge all incomplete network layer, there were multiple cases of replica segments, which resulted in the over-count of intersection density. The overlapping segments were not evenly spread, which made it inapplicable to use a constant error coefficient.

Descriptive statistics is not available due to the data's bad quality. Since this measure only addressed connectivity, while alternative measures (i.e., Metric Reach) account for both connectivity and sidewalk availability, the Intersection Density measure was not considered in the final model.

Appendix B – Single Variable Regressions' Scatterplots

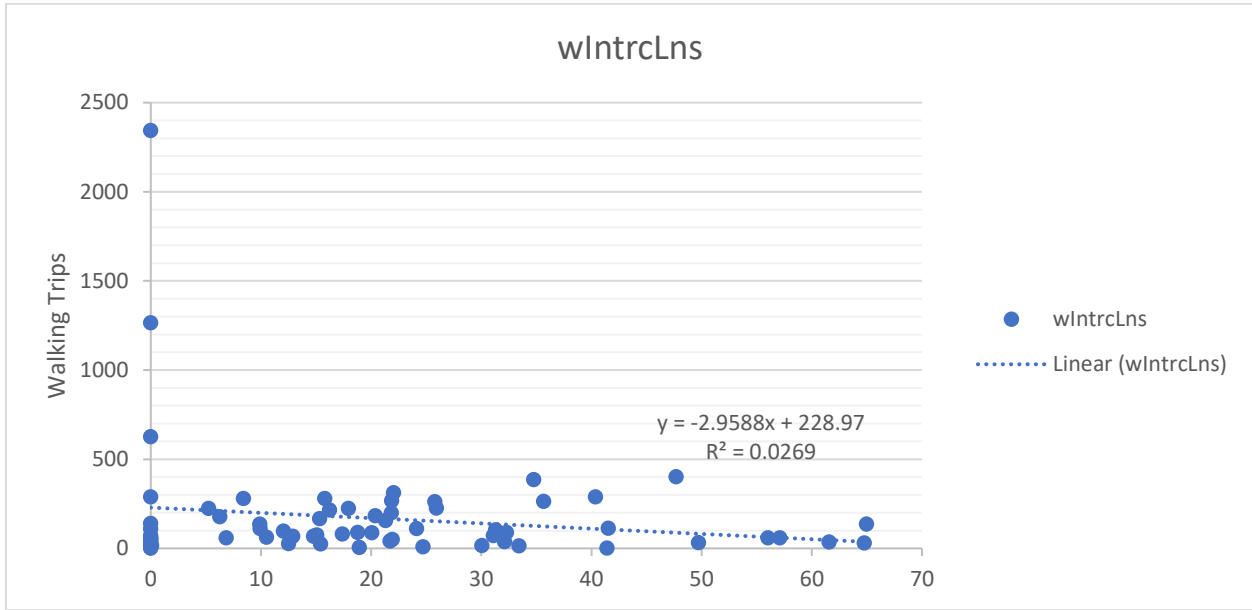


Figure B.1: Interaction Lines (Raw Data version) Linear Regression

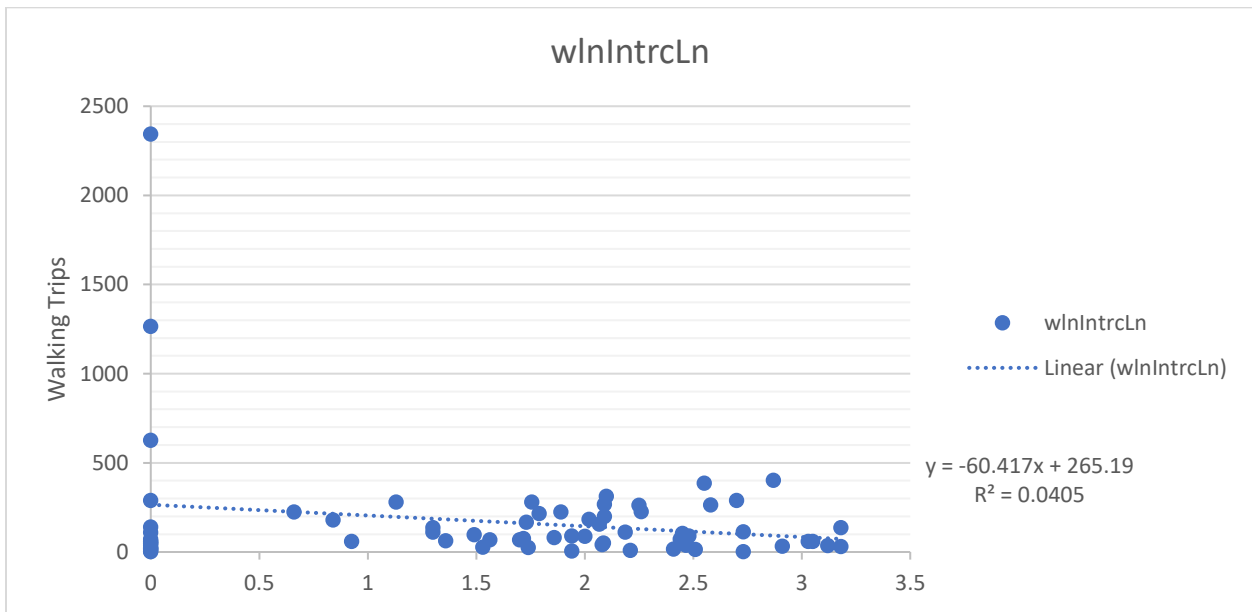


Figure B.2: Interaction Lines (Natural Log version) Linear Regression

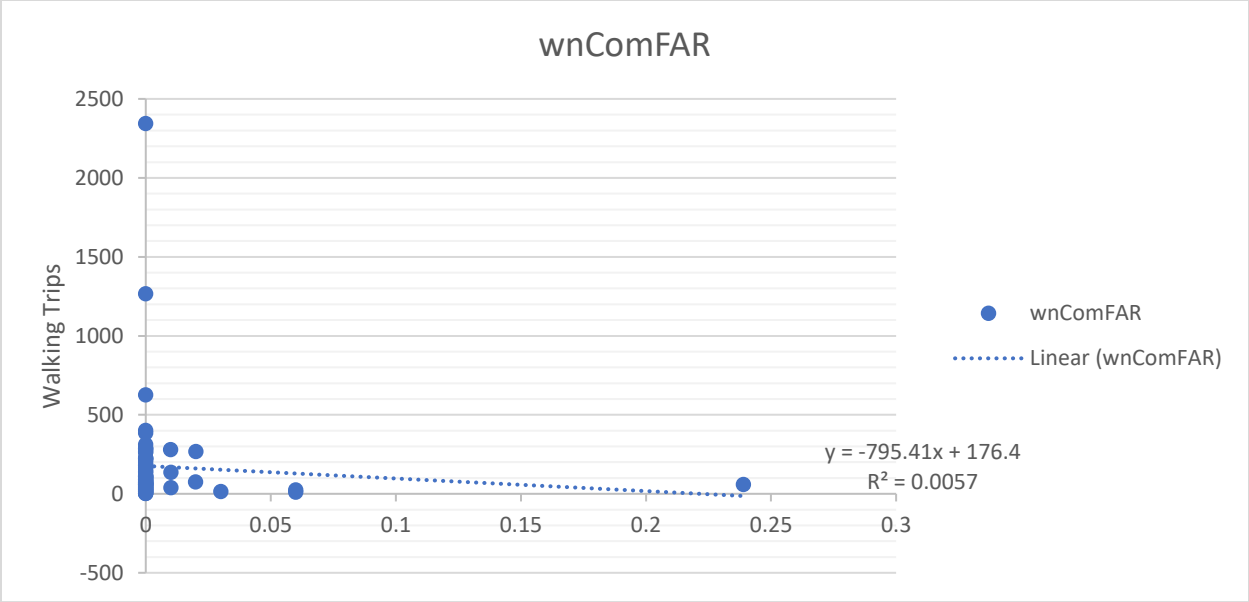


Figure B.3: net Commercial Floor Area Ratio (Raw Data Version) Linear Regression

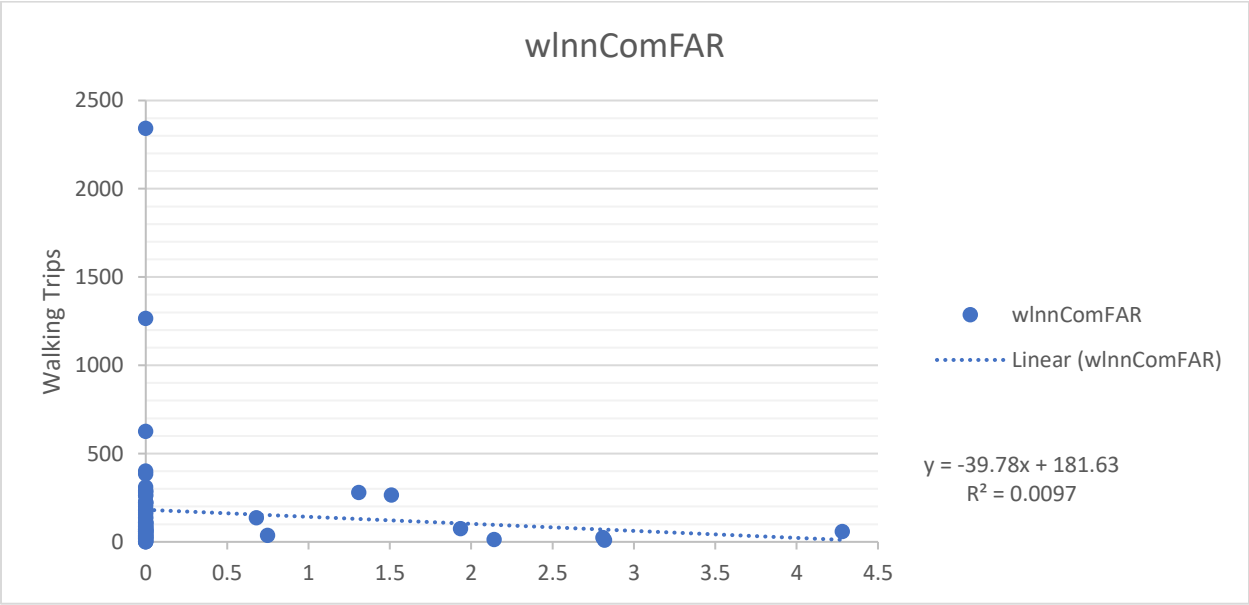


Figure 0.4: net Commercial Floor Area Ratio (Natural Log Version) Linear Regression

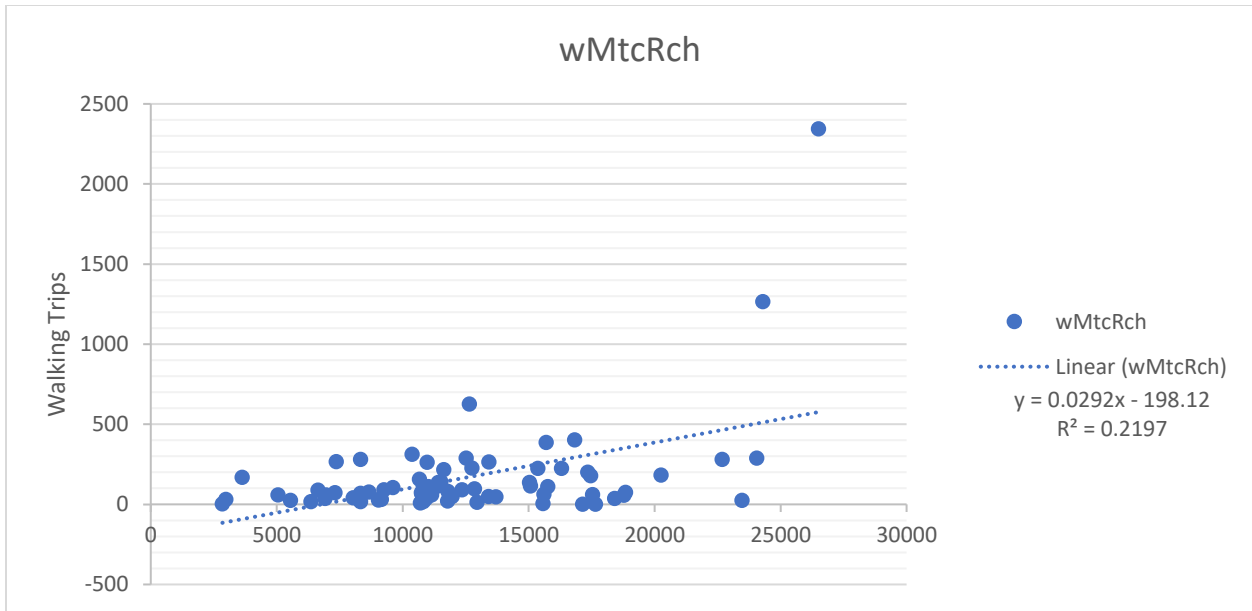


Figure B.5: Metric Reach (Raw Data Version) Linear Regression

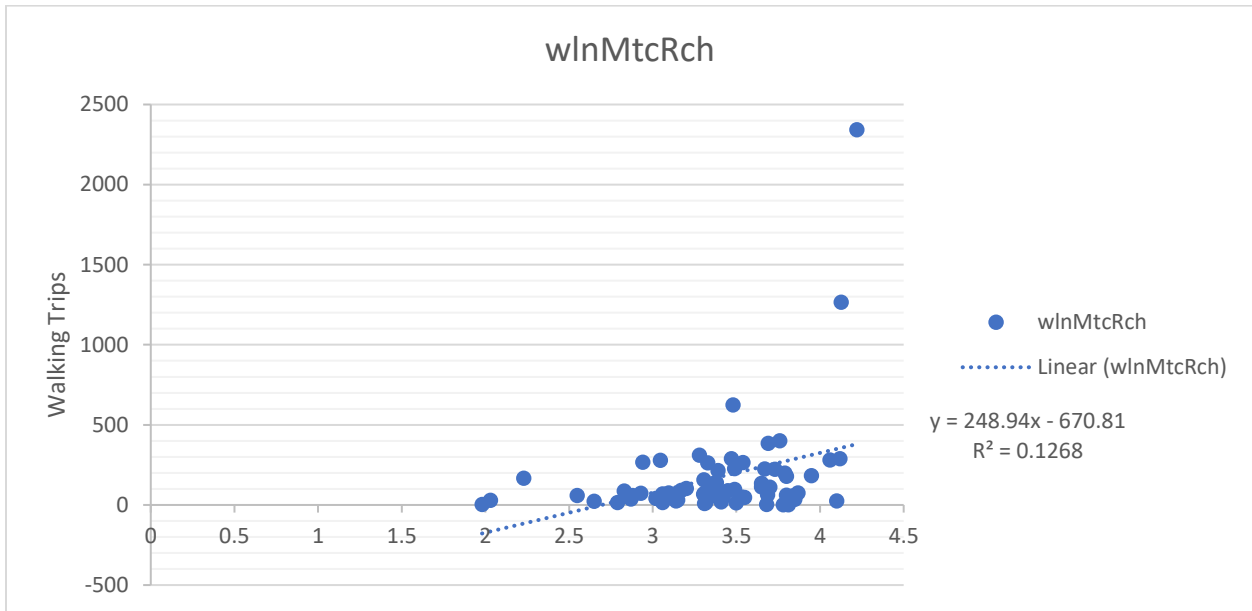


Figure B.6: Metric Reach (Natural Log Version) Linear Regression

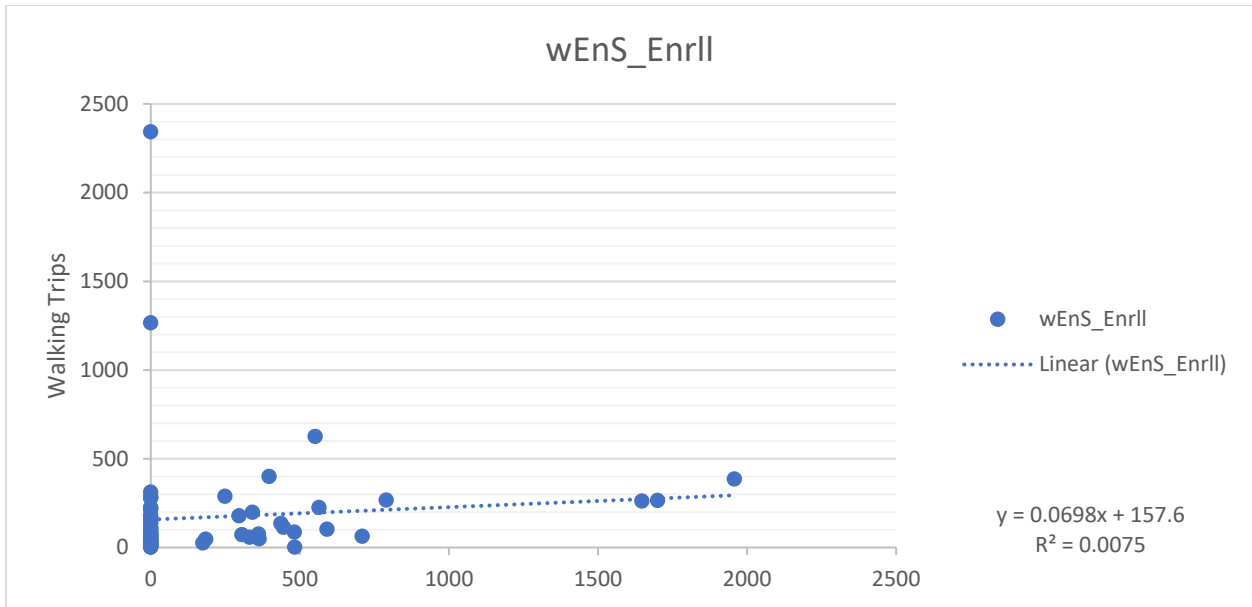


Figure B.7: Elementary and Secondary School Student Enrollment (Raw Data Version) Linear Regression

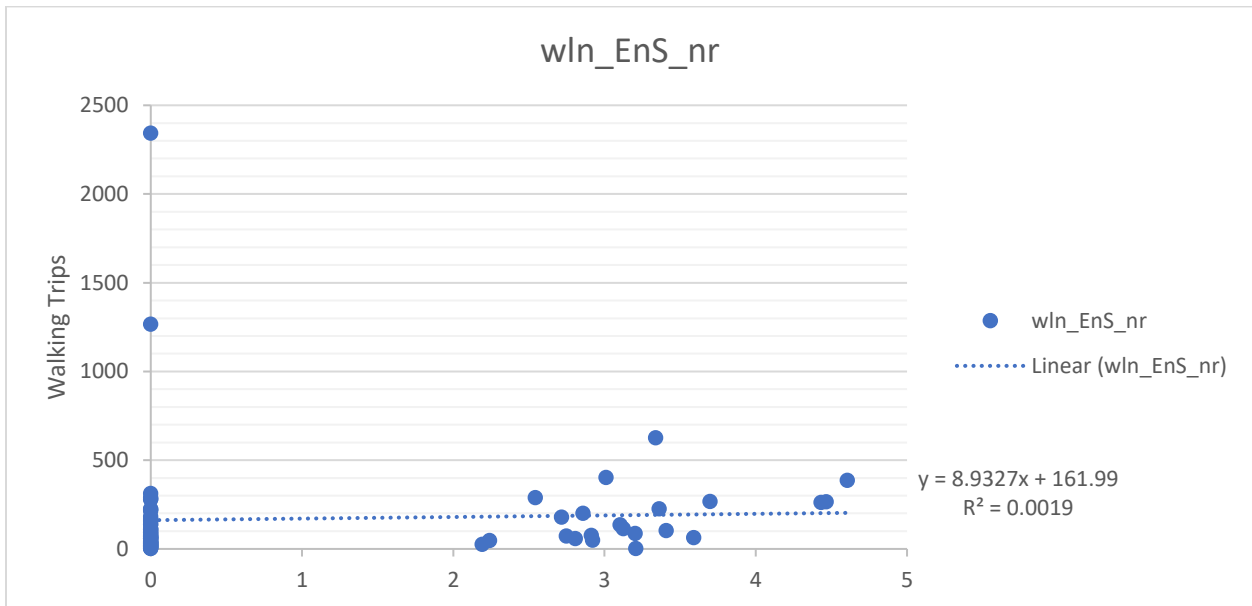


Figure B.8: Elementary and Secondary School Student Enrollment (Natural Log Version) Linear Regression

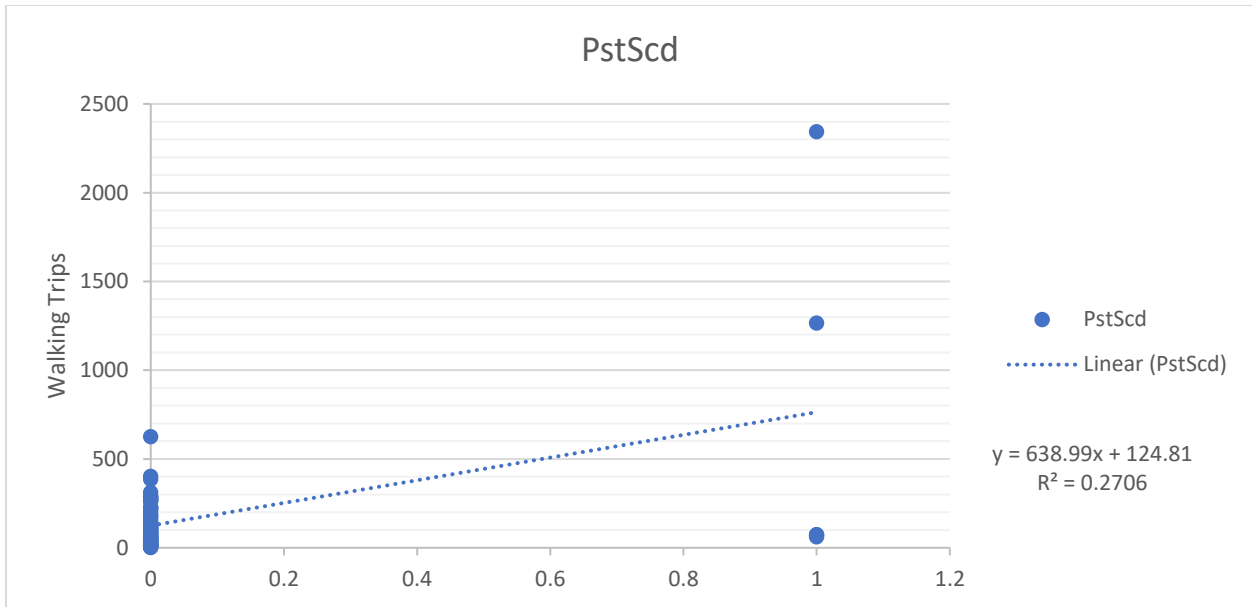


Figure B.9: Post-Secondary Presence (Raw Data Version) Linear Regression

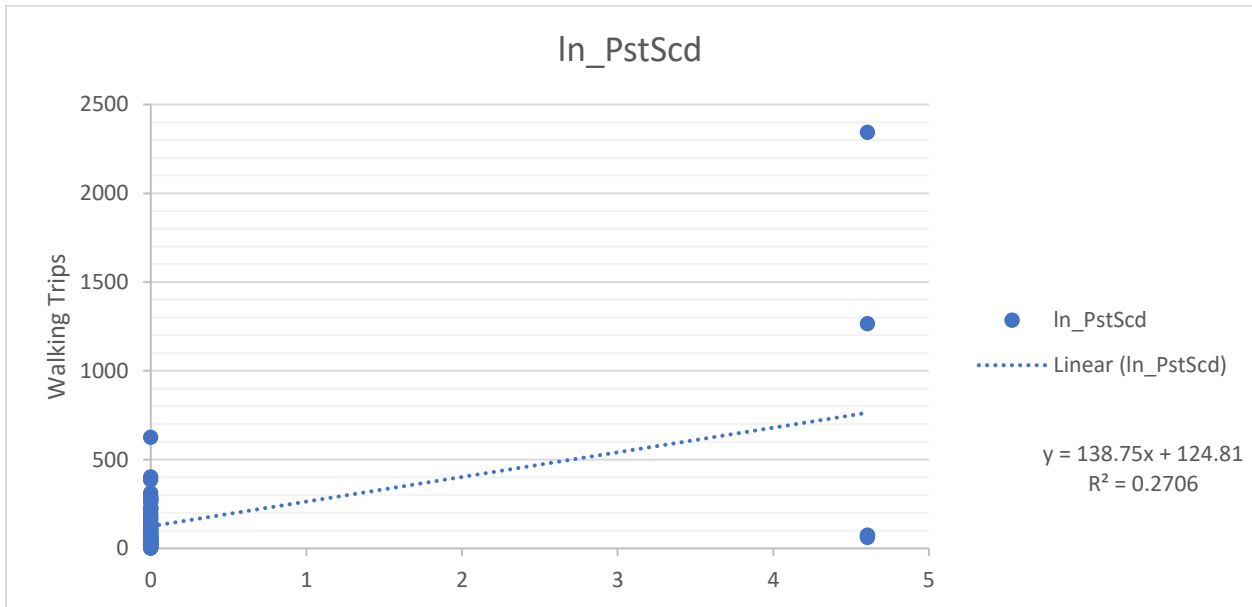


Figure B.10: Post-Secondary Presence (Natural Log Version) Linear Regression