A Homogeneity-based Zone Delineation Model for Land Use and Transportation Interaction Analysis: Investigating the Case of Light Rail Transit (LRT) Development in Kitchener – Waterloo

by

Pedram Fard

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Planning

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Matthias Sweet Associate Professor, Ryerson University
Supervisor(s)	Jeff Casello Professor, University of Waterloo
Internal Member	Pierre Filion Professor, University of Waterloo
Internal-external Member	Peter Johnson Associate Professor, University of Waterloo
Other Member(s)	Rob Feick Associate Professor, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the 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.

Statement of Contributions

Portions of this thesis have been previously published in a conference preceding and presented at conferences. The following provides a listing of the sections of the thesis that have been previously presented or published in whole or in part, and the associated citations. As the first author of these works, Pedram Fard conceptualized and initiated the studies, conducted the literature review, developed the quantitative models, performed the spatial and statistical data analyses, prepared tables and figures, and drafted and revised the submissions. Dr. Jeff Casello and Xiaomeng Xu (PhD Candidate), as the co-authors of these works provided constructive feedback on the models validation, interpretation of the results and their implications.

Content of Section 3.5.1 and Section 4.2.1:

> Fard, P., Casello, J. M., & Xu, X. M. (2021). An Approach to Measure Land Use Heterogeneity to Identify Homogeneous Urban Activity-Clusters. Presented at the Transportation Research Board (TRB) 100th Annual Meeting. Washington D.C. (Online), January, 2021.

Content of Section 3.5.3 and Section 4.2.3:

> Fard, P., and Casello, J. M. (2020). Identifying Homogeneous Urban Activity-Clusters: Application of Novel Spatial Aggregation Methods. Presented at the Municipal Property Assessment Corporation (MPAC) Annual Meeting. Ontario (Online), October, 2020.

Content of Section 3.4 and Section 3.5.3:

> Fard, P., and Casello, J. M. (2020). Implementation of an Enhanced Entropy Computation Algorithm for Measuring Urban Land Use Diversity. Presented at the ESRI Canada GIS in Education and Research Conference. Toronto, March, 2020.

Abstract

In an ever-increasingly urbanized world, planning policies bring direct and indirect societal and environmental impacts affecting quality of life for millions of people. Policy decisions are often complex, involving trade-offs between competing interests and high degrees of uncertainties. Quantitative methods have been used to understand the complexity of urban dynamics, to evaluate the alternative future scenarios and ultimately to help make more informed decisions. Despite the advantages these methods offer, they have been criticized for being ad-hoc, complicated and sensitive to the arbitrary choice of the indicators and the spatial scales of analysis.

In particular, transportation analysis and modeling often rely on pre-set structures of Traffic (or Transportation) Analysis Zones (TAZs) to conceptualize geographic space as it relates to urban activities and transportation flows. Theory suggests that appropriately created spatial structures for transportation analysis should represent areas with homogeneous characteristics in terms of land uses and activities. Reviewing literature indicates that conventional TAZs do not necessarily provide satisfactory levels of homogeneity due primarily to the insufficiency of density as the primary measure to create these zones and the arbitrary use of roadways in breaking the zones boundaries. As we move towards an era in which new mobility modes emerge and modern data sources open up great opportunities, it is necessary to rethink the way we conceptualize space within land use and transportation system interactions (LUTI) studies.

This research is motivated by the idea that land use diversity is equally important as densities (and other attributes) to define the spatial unit of analysis. The research aims to advance understanding of the impacts caused by the choice of analysis zones on the travel behavior and land use development analysis outcomes. This dissertation develops an enhanced measure of heterogeneity (i.e., land use diversity) and applies this measure to create a dynamic zonal structure through an iterative spatial aggregation method. This algorithm combines the input disaggregate zones that have similar diversity levels but also assembled from similar disaggregate land uses that make up their diversity.

The developed spatial models are examined and validated using a set of disaggregate land use, travel behavior and the building permits data from Waterloo Region in southern Ontario, Canada. This research examines the effects of land use heterogeneity and access to rapid transit on an ongoing urban dynamic in this fast-growing mid-size metropolitan region.

The first set of analyses explores the suitability of the proposed zonal structure – called Dynamic Activity Cluster Zones (DACZs) – compared to a commonly used pre-defined TAZ system and a graph-based spatial clustering model. The results indicate the advantages of the DACZ model in terms of concurrently creating more homogeneous zones with balanced size distribution. A sensitivity analysis is then performed to evaluate the robustness of the DACZ model in producing reliable zonal structures as a function of three parameters including aggregation heterogeneity threshold, levels of adjacency, and the original (input) spatial disaggregation. The results show that the model is effective in generating zones for which the size is defined as a function of homogeneity, as a result, these zones will generate more predictable outcomes in travel behavior modeling and analysis.

The second work investigates the regional daily travel behavior data aggregated and compared for both the DACZ and a conventional TAZ structure used in the regional planning called PLUM (an acronym for Population and Land Use Model). The comparisons reveal that the impacts of built environment homogeneity on travel behavior are more pronounced within DACZs, where the dynamic zones effectively capture variations of the active transportation and public transit mode shares. This analysis also uncovers a varying pattern of mode share and the average travel times across the built environment categories identified based on the population density and land use diversity levels; by increasing the levels of population density and land use diversity more trips are shown to be made by non-auto modes. This outcome supports the LUTI theories which contend that areas with diverse land uses and high population density are more conducive to active transportation and public transit trips.

The third investigation seeks to understand how the introduction of proposed and actual rapid transit investments are related to land use development trends. In a temporal analysis, the historical building permit data from 2000 to 2019 are analyzed focusing on two periods before and after the LRT project funding announcement (2010-2011). The adjusted permits construction values are calculated and compared across multiple scales including the study area, relative to the Regions' Central Transit Corridor (CTC) and within different heterogeneous built environment categories. The results identify areas that have disproportionately attracted more and higher valued developments, especially after announcement of the LRT project funding. The outcomes also confirm the role of higher levels of land use diversity and access to rapid transit on attracting greater scale of land use developments, while the density is found to have minimal association with this trend.

In summary, this study advances the research on land use and transportation system interactions by (i) articulating a novel spatial unit of analysis through developing and applying an enhanced homogeneity index and a spatial aggregation model, (ii) examining the associations between travel behavior patterns and heterogeneous built environment characteristics, (iii) providing insights on the development trends across Waterloo Region at multiple spatial-temporal scales that can be used in ongoing regional policy and planning evaluations, (iv) more generally facilitating the land use and transportation integration in planning and policy development through assessment and dissemination of a set of rigorous spatial modeling methods.

Acknowledgements

"If you want to go fast, go alone. If you want to go far, go together."
– African proverb

My PhD studies have been a long and rewarding journey where I have met many amazing people and brilliant minds, and made great experiences. I am incredibly grateful to those who have advised, encouraged, and supported me throughout my research and helped me thriving in this endeavor. I would like to take this opportunity to personally thank these individuals for their tremendous support.

First and foremost, I would like to express my utmost appreciation to my supervisor and friend Professor Jeff Casello. Thank you Jeff for your unwavering support and inspiration throughout this scholarly endeavor. Thank you for your always illuminating advice, dedicated mentorship and giving me the opportunities to engage in many impactful projects from advocating better transit for communities to delivering a more empowering education here at the University of Waterloo. It has been such a pleasure working with and learning from you, both personally and professionally. The many dinners and interesting conversations we had with you and Daria were joyful and are unforgettable to me and my family. We are forever grateful for your warm-hearted support.

I would also like to sincerely thank my dissertation committee members for their valuable guidance and support towards development and improvement of this research. Thank you Professor Filion for keeping me connected with the big picture. Your questions have always helped me to clarify my thoughts and directed me to better articulate my findings for delivering greater value to the planning community. Professor Johnson, thank you for your continued encouragement and suggestions on how to manage the methodological details of my work in light of the broader research questions. Thank you Professor Feick for being my committee member and providing your great insights on the avenues for future research. I am very fortunate that I had the opportunity to working with you all!

I wish to extend my appreciation to Professor Matthias Sweet from Ryerson University for reviewing my work as the external reader, and providing constructive feedback. I have enjoyed our dynamic discussions very much and am honored that you distinctively recognized and named my approach as the *"Fard Method"*. I would also like to express my gratitude to Professor Darren Scott from McMaster University for being the external examiner in my comprehensive examination and for providing valuable suggestions that have contributed to the development of this research. I have a special thank you to Professor Steven Roberts from Wilfrid Laurier University. Your valuable practical advice on spatial clustering and optimization models was highly instrumental in evaluating and validating my developed spatial methods. Thank you very much!

I have also been privileged to receive invaluable support and feedback at the genesis of this research. I am deeply thankful to Professor Dawn Parker and Professor Markus Moos for their thought provoking feedback that have greatly shaped my research foundation. Thank you Dawn for providing me the opportunity to collaborate with the Waterloo Institute of Complexity and Innovation (WICI) and the Urban Growth and Change Research Group.

I would like to acknowledge the financial support provided to this research by the Social Sciences and Humanities Research Council of Canada through the Partnership Development Grant awarded to Dr. Dawn Parker, and the Natural Sciences and Engineering Research Council through the Discovery Grant awarded to Dr. Jeff Casello. I was also honored to receive the President's Graduate Scholarship (PGS) during the early years of my PhD studies at University of Waterloo.

I owe a great debt of gratitude to Professor Eric Miller and the Data Management Group (DMG) team at University of Toronto and Uzo Russo from the Municipal Property Assessment Corporation (MPAC). This research would not have been possible without your generosity in providing the invaluable travel behavior and land use data. I am also very thankful to Reid Fulton and John Cicuttin from the Transit Development office at Grand River Transit, and Tristan Wilkin and Rehan Waheed from the Region of Waterloo's Planning Division for offering their valuable insights on the regional transit system evolution, travel behavior and land use dynamics in Waterloo Region.

Many thanks to Warren Davison (City of Waterloo), Markus Wieland (UW Geospatial Centre) and Scott MacFarlane (UW Mapping, Analysis & Design Lab) for sharing their GIS expertise and assembling the required municipal data sets. This research has enormously benefited from the practical programming advice generously offered by Jemmy Hu (Compute Canada), Hossein Hosseini (Esri Canada), Eddy Ionescu and the UW Data Science Club members. Also a big thank you to the good folks at the Portland R User Group and Toronto Data Workshop. Although we never actually met in person, your friendly and welcoming Meetup sessions were a great place for me to learn programming and troubleshoot my code at the pinnacle of pandemic. Thank you all very much!

I would also like to express my sincere thanks to the (current and former) colleagues and staff members across the campus community for helping me in many different ways. My deepest gratitude to Jane Russwurm at the UW Writing Centre for her compassionate guidance and mentorship above and beyond the writing sessions. My special thanks go to Sheree Solomon, Kelly Heald, Taylor Ertel, Alison Frayne, Lori McConnell and Laurie Larochelle for their keen diligence and prompt assistance in helping me navigating administrative procedures and logistical matters of my research. Many thanks to Dheana Ramsay and Maryam Latifpoor-Keparoutis from the Faculty's Office of Advancement for connecting me with the amazing community of UW alumni network.

Furthermore, I would like to thank my fellow students and dear friends at the School of Planning. To name a few, I wish to thank Rafael, Sara, Su, Nina, Matt, Robert, Xinyue, Yu, Fatima, Katherine, Abiral, Jon, Nick and Kourosh with whom I shared great times. My peer TAs at the Faculty of Environment, and Scott Anderson from the UW Centre for Teaching Excellence (CTE), I am thrilled that I was able to work with you during multiple offerings of the ENVS178 course. I was honored to receive the Excellence Award for Teaching Assistants for this class which was, in fact, recognition of our collaborative efforts in providing the best possible chance for our students to be successful in this fundamental course. Thank you all for your contributions and efforts.

A massive thank you to all my wonderful friends and colleagues at the Waterloo Public Transportation Initiative (WPTI) for the great accomplishments we have shared together along the way. Kevin, Kitty, Will, Adam, Jason, Ian, You-Jin, Veronica, Alie, Janelle, Jacob, Devin, Andrea, Binesh, Isha, Owen, Lindsay, Jeffrey and Aaron, you made my experience at Waterloo much richer and I look forward to your forthcoming contributions to a brighter transportation future for all.

I would also like to say a very special thank you to my collaborators and friends, Xiaomeng (Ming) Xu, Yue Zhao and Yixin (Candice) Chen who have been great companions for me and my family during the past few years. We are truly grateful to have you as our extended family!

Finally, I wish to express my profound gratitude to Professors Markus Stöckner, Christoph Hupfer and Jochen Eckart for providing such an impactful experience throughout iterations of the Hands-on Sustainable Mobility Workshop. I would like to extend my warmest thanks to Jule Merk and Julian Knust for graciously hosting us in beautiful Karlsruhe. Danke schön!

My PhD experience and those with whom I met and worked with enriched my life in meaningful ways both professionally and personally, and for that I am extremely grateful!

On a personal note, I would like to give a very special shout-out to my best friends from far and close, Abtin, Alex, Alireza, Roham, Ali, Bahman, Danon, Masoud, Naghmeh, Dena, Hossein, Ramin, Jeff and Sam. You have been a constant source of joy, laughter and encouragement over the many years. Thank you so much for cheering me up through the ups and downs of my PhD journey.

A heartfelt thank you to my loving family, my mother, my late father, my sister and my in-laws. You have cherished and supported me well before I began this endeavor. I am forever grateful for everything you have done for me to get here. Last but certainly not least, an immense thank you to my beautiful wife and daughter, Mengyi and Megan, you have brought the sunshine to my life and given me confidence when I most needed it. This work would not have been possible without your unconditional love and company, I wish to thank you with all my heart and dedicate this dissertation to you both.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Context

As we enter the third decade of 21st century, cities are increasingly confronted with challenges that drastically impact millions of their inhabitants. Disparities in access to jobs, affordable housing, education and health services are frequently observed in cities worldwide. These ever widening socio-economic inequalities, in addition to the adverse consequences of climate change and extreme events have necessitated and brought the notion of sustainable development to the forefront of urban planning agendas. More specifically, the United Nations' Sustainable Development Goals targeted equitable access to urban services and sustainable transport systems as universal resolutions to be achieved by 2030 (United Nations, 2015). The concept of sustainable development entails objectives of environmental protection, social equity and economic prosperity (Campell, 2016). In addition to their explicit implications in planning practices, these objectives are often conceptualized through two major strategies: Sustainable Urban Form and Sustainable Mobility.

Strategies to ensure sustainable urban form have their roots in substantive theory of planning, where appropriate shape and scale of developments are extensively discussed. The theory suggests that people and firms benefit from agglomeration and proximity advantages that cities and urban systems inherently provide. Historically, social interactions and business transactions could only become possible when different actors meet and interact locally. This was the driver for concentration of activities in densely developed form of historic cities from antiquity until the World War II era. However, while the concentration of activities benefits people and businesses, it comes with costs of land capitalization and higher land values in central and more accessible places (Harvey, 1985). Inspired by the historical instances of compact cities, to amplify their benefits and to curb their costs, sustainable urban form models advocate strategies to intensify developments and diversify activities within a network of connected communities.

Sustainable mobility, similarly, promotes integration of low-impact transportation modes - walking, cycling and public transit that aim to better connect those compact and diversified communities. It also provides measures to stimulate travel behavior changes towards lower degrees of car dependency (Banister, 2011a). The associated strategies often address major concerns of energy use and greenhouse gas emissions – higher land use densities and associated low-impact travel have been demonstrated to require lower energy uses per capita (Figure 1-1) – as well as infrastructure operation and maintenance costs (Kenworthy, 2007; Kenworthy et al., 2000). Provision and upgrading of public transit infrastructure as the backbone of sustainable mobility have been widely credited for encouraging social justice and better public health (Martens, 2016).

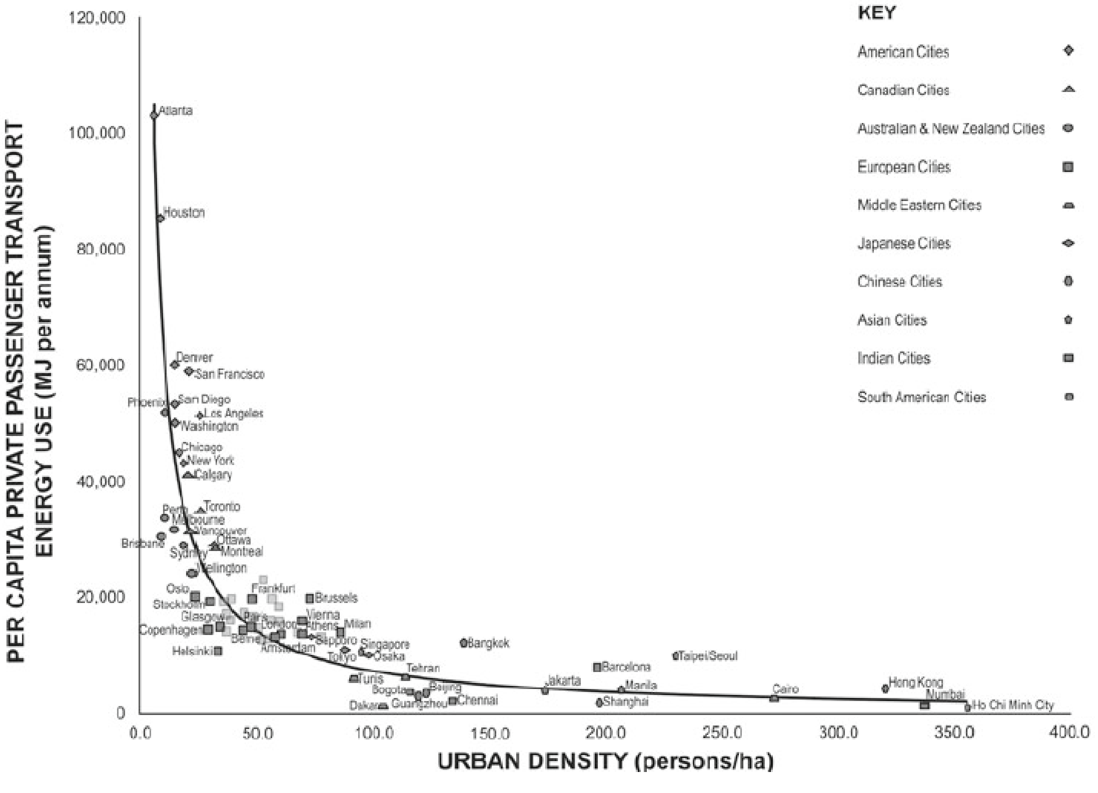


Figure 1-1 - Rapid decline in car use per capita with increasing urban density (Kenworthy et al., 2000)

Since sustainable mobility and urban form are interconnected, to achieve sustainable and economically viable urban development requires better integration of these two approaches. This is often referred to as integrated land use and transportation planning, which has been articulated under three broad concepts of New Urbanism, Smart Growth and Transit Oriented Development (TOD).

While integrated planning approaches have gained interest worldwide, the complex intertwined nature of land use and transportation system interactions (LUTI) need to be further investigated. Classical theories suggest that transportation network expansion in general, and transit infrastructure in particular, impact land use changes through two major mechanisms: reducing generalized transportation costs, and altering relative accessibility levels (Guiliano, 2004).

Both these factors significantly influence the location choices of key urban actors. Particularly, these influences are felt by: households and the people who live in them; firms who locate in metropolitan areas and hire employees; developers who supply space for activities in the area; and governments who influence the behavior of households, firms and developers.

The dynamics of interest in this research are land use and transportation system changes that take place in a complex urban context, in association with other established structures including political, economic, and planning systems and social norms.

Conventionally, governments are most often responsible for the supply and operation of the transportation system. The composition – the mix of modes available – as well as the system performance and costs are all directly controlled by them. Local governments also have an active role in establishing land use regulations and economic incentives that influence development.

While both transportation interventions and land use regulations are ideally to be implemented simultaneously, in practice, land use policy instruments are more widely adopted and zoning by far is the most common employed instruments among all (Qian, 2010). Land use instruments have more spatially concentrated impacts rather than desirable distributed effects. In contrast, transportation interventions potentially have wider and more socially inclusive and fair impacts (Lucas, 2012). In fact, temporal variance and the transportation changes precedence often gives an opportunity to planners and policy makers to leverage transportation interventions as a guiding force to shape desired land use related changes. However, when these changes materialized, they start to impact travel decisions and hence functionality of transportation systems (Figure 1-2).

Among many characteristics of land use, population and employment densities are typically considered the most significant characteristics that influence travel patterns at different spatial scales. Looking through the lens of integrated planning has led many municipalities to trying and achieve land use intensification, increasing the density and diversity of land uses such that more household activities can be accomplished with less (in distance and time) and less expensive travel. High levels of land use diversity provide multiple activity destinations within an easy access area which makes non-motorized transportation feasible. Moreover, higher levels of densities are necessary to support efficient public transit performance. Intensification also tends to promote the linking of multiple individual trips into a trip chain that again reduces per capita vehicle travel and energy consumption (Banister, 2011b).

Given this relationship, it is logical to try to embed a measure of the homogeneity (or heterogeneity) in studies of land use and transportation interactions. Hence, the focus of this research work is to consider another critical, but less well studied characteristic of built environment for land use and transportation system interactions analysis: land use diversity.

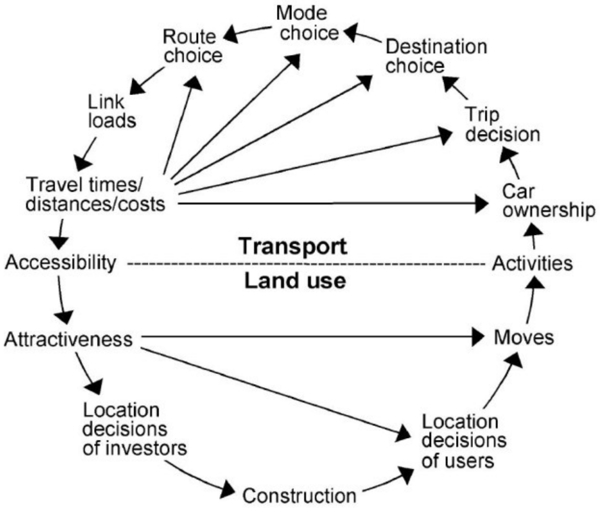


Figure 1-2 - The land use and transportation system interactions feedback cycle (Wegener, 2004)

1.1.1 Problem Statement and Motivation

Land use and transportation system interactions are complex. The absence of a unifying theory, coupled with a scarcity of consistent spatio-temporal data and limitations of analytical models have created challenges for understanding the extent and characteristics of these interactions.

In practice, planners have attempted to reverse a trend of providing transportation systems (often roadway expansions) in response to growing land use activity and now are trying to introduce high capacity, low impact modes in city centers to attract high density, high diversity land use development. Such strategies have been effective in some places but not in the others. In general, public transit systems are considered to be development catalysts and the literature offers evidence of land use impacts driven by transportation interventions (Cervero & Duncan, 2002). However, more recent studies have shown that to leverage the interrelationships between transportation system and land use development the context is decisive (Higgins, Ferguson, & Kanaroglou, 2014). Often the primary drivers for new land use developments or transportation systems interventions are external factors such as rapid demographic changes, economic cycles, emergence of new technologies, and shifts in social and environmental policies. Saturation in urbanization and transportation accessibility levels also impact land use and transportation system interactions beyond what intuitive theories offer (Kasraian, Maat, Stead, & van Wee, 2016). Furthermore, empirical studies that have examined the subject have often ended up with heterogeneous outcomes (Higgins & Kanaroglou, 2016).

Relevant to this research, a multi-decade longitudinal study on co-evolution of the transportation networks and land use in the Greater Toronto-Hamilton Area (GTHA) has found that the initial impacts of the transportation accessibility on urban development have increasingly decreased over time, while proximity to built up areas has had more lasting impacts on the land use development (Kasraian, Raghav, & Miller, 2020).

Current technologies and data availability make possible more robust studies of the interactions between land use and transportation. More specifically, emerging modern data acquisition and processing techniques along with rapid advancements of Geographic Information Systems (GIS) and High Performance Computing (HPC) offer the potential to overcome the barriers for understanding land use and transportation system interactions. Accordingly, this study utilizes data-driven spatial modeling methods to explore characteristics of these interactions in a fast growing metropolitan region in Ontario, Canada. In this interesting urban dynamic case, the Region of Waterloo, is employing an integrated (land use and transportation) approach to drive sustainable urban development in a dispersed metropolitan area. While the region is experiencing significant population and employment growth, new land use controls are incenting intensification around a newly constructed Light Rail Transit (LRT) system. Collectively, the introduction of LRT and expansion of the higher-order bus system, aim to transform mobility pattern across the region. The growth in the Region of Waterloo, and the related land use and transportation changes, provide a unique study area to evaluate land use and transportation system interactions in the presence of contemporary information sources.

From an analytical perspective, this research also investigates a longstanding practice applied in transportation analysis – the creation of a spatial (zonal) structure that is the basis for many evaluation tools, but more prominently estimates of trip activity. Traditional methods of creating zones attempt to generate analysis areas with relatively consistent total activities such as population or employment. Areas with higher densities are often represented with smaller zones, whereas lower density areas may be spatially larger. While this approach allows for consistency in important land use characteristics, it fails to consider the diversity of land uses contained in a proposed zone and, as a result, the zones created may not actually be homogeneous in the behaviors observed.

This issue of inadequate homogeneity of common transportation analysis zones combined with the arbitrary use of roadways in breaking the analysis zones boundaries are commonly recognized issues in transportation analysis. More specifically, zonal structures can be problematic, creating disproportionately negative impacts on the representation of non-auto mode trips and trips with purposes other than home-to-work. As we move towards an era in which active transportation options increases and new mobility modes emerge, it is necessary to rethink the way we conceptualize space within land use and transportation system interactions studies.

1.2 Research Objectives and Research Questions

There are two overarching objectives for this research. The first objective is to develop and evaluate a set of spatial models to examine the impact of data aggregation scale in land use and travel behavior interactions analysis. The second objective is to investigate the effects of land use heterogeneity and access to rapid transit on travel behavior and developments in metropolitan areas. This research aims to provide evidence-based information and analysis to answer the following questions:

1. What is the appropriate spatial scale to analyze land use and transportation system interactions and their impacts on travel behavior in a mid-size metropolitan region, based on observations from Waterloo Region? What are the benefits of using dynamically defined zones in travel behavior analysis?
2. How do built environment characteristics (e.g., density and diversity) and access to rapid transit influence travelers' behavior across a mid-sized region, based on observations from Waterloo Region?
3. How do proximity to LRT corridor and land use diversity levels impact the development dynamics across a mid-sized region, based on observations from Waterloo Region?

The responses to these questions provide not only methodological contributions to the literature. The answers generated through this research can help inform planning processes – improving the way in which critical infrastructure investments are conceived, designed, evaluated and built. Therefore, in tandem with conducting the case study analyses, a set of relatable results are selected and interpreted to convey the planning implications and knowledge that can arise from this research.

Nearly all planning tasks include a consideration of land uses; as a result, it is of high importance to be able to conduct these analyses appropriately and effectively using robust methods. One overarching contribution this research makes to planning is to assess how incorporating the diversity of land uses can create different outcomes compared to simply looking at density measures, and how that might influence the planning processes, particularly in transportation planning. The work done in this study is inherently linked to applied planning practices and outcomes from the Regional Municipality of Waterloo.

First, we present an example from the Region's Central Transit Corridor Monitoring Program, in which the Regional Municipality has implemented a set of quantitative indicators to evaluate the official development goals of *moving people more efficiently* and *building thriving communities*. In our analysis we assess the applicability of our proposed land use diversity measure to perform a planning assessment task, similar to the work conducted by the Region.

Second, we look into a recent major transportation project in the region, to demonstrate the strength of the concept of dynamic zone delineation in planning practice. We re-evaluate the work that was done on alignment selection for the recently developed regional LRT system, focusing on a section that currently serves the University of Waterloo Campus. We demonstrate how our proposed method might have influenced the planning outcomes if alignment selection were done using the enhanced diversity measurement in conjunction with dynamic zone delineation. We argue that our method could more convincingly identify areas with land uses that might be more supportive of an LRT system. This analysis may have pointed towards a different alignment in certain sections compared to the currently developed network.

Broadly, we propose and develop a set of spatial models and metrics to simultaneously synthesize heterogeneity of land use and travel behavior patterns, and contend that our methods can help to effectively identify the development gaps and potentials at the intersection of land use and transportation systems, at a very fine spatial scale. These abilities can enable planners to leverage existing potentials, and prioritize interventions where, for instance, the land use context is conducive for expansion of active transportation infrastructure. We conclude that the flexible dynamic zone delineation promises a methodological enhancement and a range of applications that are often not fulfilled in planning practice due to the limitations imposed by the conventional analysis zones such as the pre-defined TAZ structures. We elaborate on the new knowledge and original contribution of this dissertation under the Discussion and Contribution Sections.

1.3 The Conceptual Framework of Research

As theories suggests, travel behavior and location decisions are the two major foci that dictate land use and transportation system interactions (Figure 1-2). Empirical data gathered on both of these activities provide measurable outcomes that reflect dynamic choices and preferences of urban system actors. Figure 1-3 contains a graphical representation of how this study approaches the problem of land use and transportation system interactions. The left side of the diagram presents three critical areas of focus for which data have been gathered and relationships analyzed. The right side of the diagram concentrates on the appropriate spatial scale of analysis, comparing the proposed zonal structure (shown in red box) to those generated by modern state-of-the-art (graph-based) method as well as conventional analysis zones. Common to both sides of the diagram is the concept of heterogeneity, in terms of travel behavior as it relates to the built environment features and diversity of urban activities that should guide the creation of analysis areas.

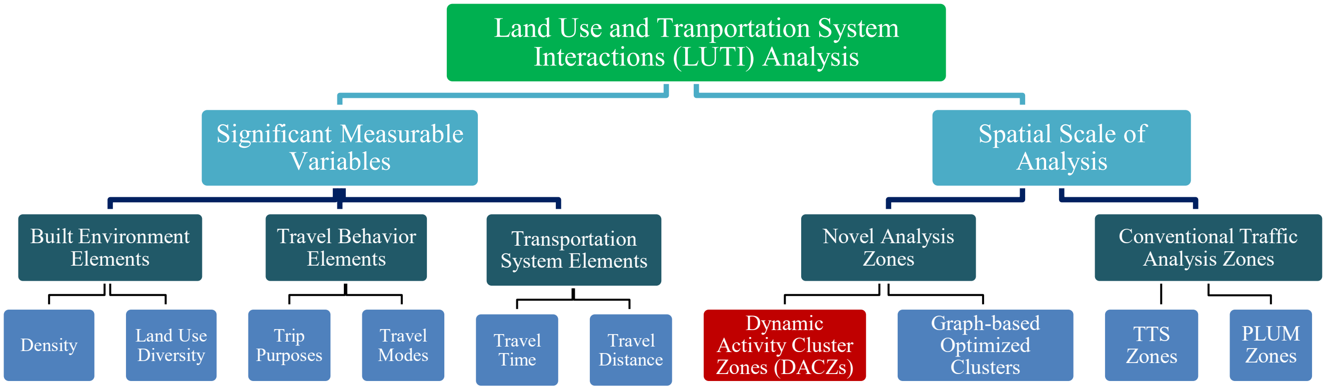


Figure 1-3 - The conceptual framework of research

Travel behavior and urban activities (land use) data in their simplest form consist of sets of attributes associated with some geocoded points in space (e.g., households location, trips origins and destinations). The first step in any land use and transportation system interactions analysis is to aggregate these point-based data into a kind of area-based zonal system. This process involves constructing a spatial abstraction that inherently introduces uncertainty in estimations (Kwan, 2012).

Looking through the lens of heterogeneity evaluation, this study develops a novel method to identify homogeneous zones and to minimize the uncertainty for travel behavior aggregation and built environment metrics measurement. The proposed method is implemented in a case study where sets of Dynamic Activity Cluster Zones (DACZs) are created for the cities of Waterloo and Kitchener. The DACZ system is then compared to the pre-defined conventional traffic analysis zones (TAZ) system. The research also implements a method that produces improved (i.e., minimum heterogeneity) spatial zones applying a graph-based spatial clustering method (also known as SKATER algorithm). This comparison aims to examine performance of these zone systems in producing more homogeneous zones in terms of land uses and variability of the zones sizes.

To answer the second question, regional travel behavior data derived from a regularly-conducted, widespread travel diary survey are aggregated at the DACZs level and contrasted with the pre-defined conventional TAZ called PLUM – used by the Regional Municipality of Waterloo for population, land use and transportation forecasts. Combining the land use data and the travel behavior data allows this study to synthesize built environment and travel behavior metrics in a cross-sectional analysis as laid out in Table 1-1.

Synthesis of Built Environment and Travel Behavior Metrics			Land Use and Population Density Heterogeneity Distribution			
			High-High	High-Low	Low-High	Low-Low
Travel Behavior Heterogeneity Distribution	Trip Purposes	Travel Modes	Each table cell represents statistical distribution of number of trips across travel time or distance bins			
	Home-based work	Auto				
		Public transit				
		Active modes …				
	Home-based non-work …	Auto				
		Public transit				
		Active modes …				

Table 1-1 - Conceptual framework of built environment and travel behavior metrics synthesis

In response to the third research questions, a spatial-temporal analysis is performed to explain dynamics of land development in Waterloo Region during a 20-year timespan, from a decade before announcement of the LRT project funding approval until a decade after that. This multi-level analysis aims to test the hypothesis that investment in the regional light rail rapid transit has stimulated more new developments to the Region's Central Transit Corridor (CTC). Historically, the CTC used to accommodate higher levels of population and jobs densities and is comprised of a relatively diverse urban environment. In recent years, areas within the CTC have also enjoyed access to high frequency transit and better active transportation connectivity compared to the rest of the region.

The historical records of building permits are used to assess whether the areas within the CTC have disproportionately attracted more and or larger developments (i.e., higher construction values per issued permit) compared to the rest of study area. Further, the analysis investigates the likely impacts of land use heterogeneity and population density levels on development activities trend.

1.4 The Analytical Framework of Research

The analytical framework of this research comprises a series of assessments to validate the developed spatial models, and examine the research questions in a case study in Waterloo Region. These analyses conducted to respond each question are organized under specific subsection within the Methods Chapter that logically linked to a subsection within the Results Chapter.

The work presented in Sections 3.5 and 4.2 aims to respond to the first research question. Here, we first develop an enhanced measure of entropy and use statistical metrics of variance to compare our measure with the conventional entropy index. We then apply the entropy measure to develop the Dynamic Activity Cluster Zones (DACZ) and evaluate the DACZ model for its robustness and effectiveness. Using a systematic sensitivity analysis, we assess the impacts of a range of heterogeneity thresholds and two levels of adjacency on the model outcome. We also provide detailed comparisons of the DACZ structure against the conventional TTS TAZs and the graph-based spatial clusters in terms of the overall homogeneity distribution and the zone size balance.

Section 3.6 (and 4.3) presents the analyses for answering the second research question that investigates specific instances of the general LUTI problem. Here we compare what we learned from two different zonal structures on travel behavior. We primarily look at the travel behavior distributions (mode share and proportions of trip purposes) within the DACZ and PLUM structures, for all zones and subset of classified zones. We then explore the travel time distributions for all zones within the DACZ system.

In Section 3.7 (and 4.4), our work addresses the third research question on investigating the likely effects of land use heterogeneity and access to rapid transit on historical trajectory of the land use developments across the region. In this temporal analysis we compare and interpret two metrics of the average adjusted construction values (AACV) and the total adjusted construction values (TACV) for subset of zones within the DACZ system and in an aggregate level.

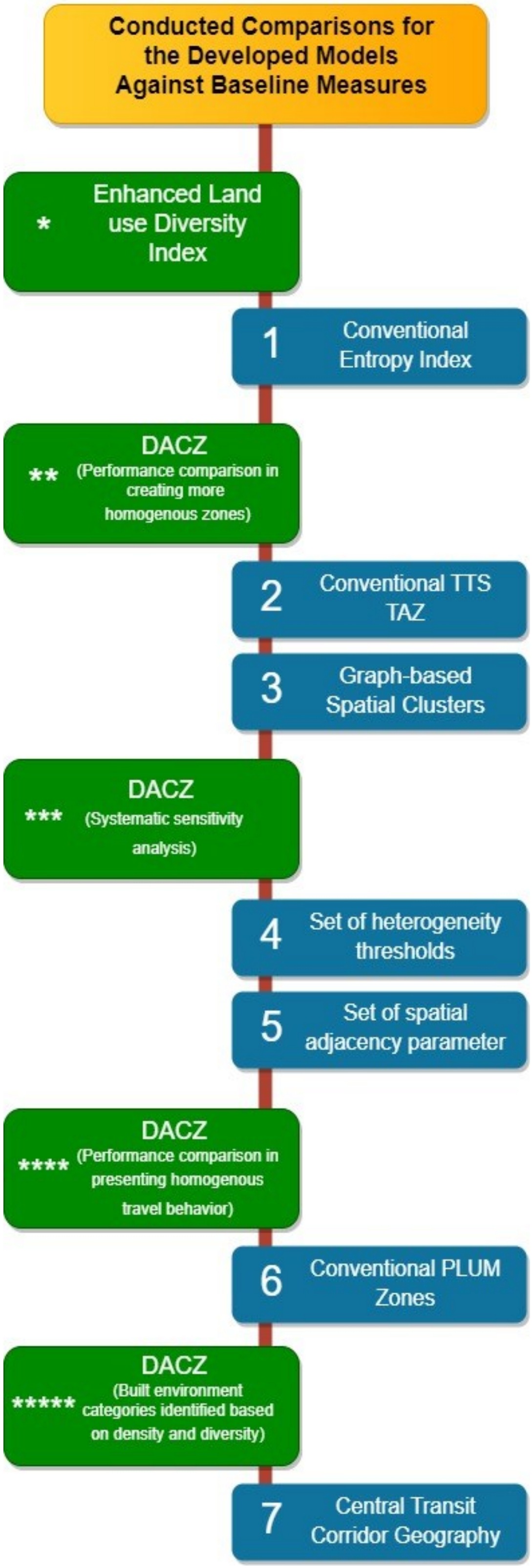


Figure 1-4 - Schematic diagram of the analytical framework of the research including the developed models and conducted evaluations

1.5 Thesis Outline

This thesis is divided into five chapters and includes a bibliography and appendices. Remainder of the thesis is organized as follows. Chapter 2 reviews the literature on the land use and transportation system interactions (LUTI), and presents implications of quantitative methods in LUTI analysis. It provides the contextual background for the research problem and sets out a theoretical framework for this research.

Chapter 3 presents the methods used and developed in this study. It begins by outlining the research data sources and introducing the Region of Waterloo and its planning context as the case study area. The chapter then portrays the proposed spatial heterogeneity measurement method (Enhanced Land Use Diversity Index) and describes the details of the developed spatial aggregation method that creates the Dynamic Activity Cluster Zones (DACZ), and the implemented graph-based regionalization model.

The research results are presented in Chapter 4. This includes the outcomes of the implemented spatial models, the sensitivity analysis and the complementary statistical analyses results. The theoretical, methodological and planning implications derived from this study are also discussed at the end of this chapter. Finally, Chapter 5 summarizes the research findings, outlines the contributions of this study, and concludes with recommendations for future research.

Chapter 2: Literature Review

2.1 Overview

Land use and transportation system interactions (LUTI) are one of the key concepts to understand complex urban dynamics. Extensive research has been conducted to describe the major components of these interactions manifested through location decisions and travel behavior. Furthermore, a variety of methods has been developed to quantify built environment characteristics and to measure transportation accessibility as drivers of location decisions and travel behavior. The focus of this research is on quantifying heterogeneity of the built environment, in terms of diversity of land uses and urban activities.

This study examines application of built environment heterogeneity measures to delineate dynamic analysis zones that is a novel approach in LUTI studies. The need for better conceptualization of built environment within land use and transportation analysis and modeling has been frequently discussed in the literature, and defining appropriate zonal structures remains an active area of research.

This review begins with a historical overview of the evolution of contemporary metropolitan regions in North America to illustrate trajectories of LUTI dynamics that have formed today's understanding and representation of urban structures. Seminal theories that explain these interactions are then synthesized, and empirical studies that have examined cases from North America are explored. Next, implications of land use and transportation integration for planning policy are discussed.

Finally, the role of quantitative methods and new data sources in LUTI analysis are reviewed and the significance of spatial representation and scale as crucial components of quantitative LUTI modeling methods are discussed. The chapter concludes by reflecting on implications of heterogeneity measurement for the analysis zones delineation, and presenting a set of conventional methods that are used in urban and transportation analysis and modeling.

2.2 Historical Background

Over the second half of the 20th century, extensive suburbanization of population and employment – the movement of households and jobs to areas outside of traditional urban cores – has dominantly shaped metropolitan landscapes across North America. Steady population and economic growth during the two decades after World War II stimulated great demand for housing development. The prevalence of anti-urban policies, and the availability of inexpensive land, along with rapid expansion of transportation networks and auto ownership, resulted in unprecedented outward metropolitan expansions (Southworth & Owens, 1993).

Through redistribution of population and jobs, this process profoundly changed the structure of metropolitan regions in North America, as conventional monocentric cities were replaced by a polycentric constructs (Cervero & Wu, 1997; Small & Song, 1994). This outward movement also led to pervasive central city and downtown decline in North American metropolitan areas (Filion, Hoernig, Bunting, & Sands, 2004).

Inspired by Borchert (1967), and Bunting and Filion (2010) this section is organized into to three subsections to describe distinct eras of metropolitan evolution in North American, including: rapid post-war expansion (1950 to 1975), deindustrialization (1975 to 1990), and post-industrialization (1990 to the present) periods.

2.2.1 From 1950 to 1975: Rapid Post-war Expansion

In the United States, where regional transportation investments had been historically perceived and promoted as generators of economic growth (Anas, Arnott, & Small, 1998), large-scale government investments were directed to expressway network expansion, while public transit had been receiving minimal proportions of investments (Filion, Bunting, McSpurren, & Tse, 2004). Thus, by the end of 1960s, US inter- and intra-metropolitan highways became extremely well-developed, covering entire states (including remote areas). These investments facilitated efficient transportation of resources and goods at regional and national scales, and provided better and more even access to labor markets and job opportunities across metropolitan regions (Kasarda, 1980). In this period, central areas in American urban centers began to decline.

In Canada, however, public investments made in transportation infrastructure development were divided more evenly between metropolitan public transit development and expressway expansions (although highways and roads, as in the US, received more funding). In turn, Canadian expressways were constructed and expanded on a much smaller scale as compared to the US, so jobs and urban activity continued to be concentrated in downtowns and other central urban areas with little perturbation (Filion, Bunting, et al., 2004; Hodge & Gordon, 2008).

Transportation network expansions in post-war North America provided an opportunity to reduce the importance of physical distance in the equilibrium of distance versus travel costs (Muller, 2004). Designated for high-speed traffic, expressways allowed longer distances to be traveled within a certain time budget compared to other conventional transportation modes of the time (e.g., streetcars and railways). The popularity and affordability of private car ownership - as a representation of consumer culture - occurred in tandem with these expressway network expansions, which soon allowed people to realize this potential (Bunting & Filion, 2010; Muller, 2004). The advent of low-cost transportation options diminished the conventional mobility barriers that historically had constrained the size and form of cities to public transit service areas and walking distances (Newman & Kenworthy, 1996). Now, residents were enabled to commute longer distances while maintaining similar levels of access to traditional urban cores. Concurrently, greater access to broader labor markets and customer bases allowed firms to relocate from downtowns (with the highest urban land prices) to more inexpensive land on the urban periphery (Filion & Bunting, 2010). As larger middle-class populations came to reside in suburbs, retail and commercial activities found their way into urban peripheries. By 1970 in major US metropolitan regions, suburbs hosted more jobs than traditional urban cores (Jackson, 1987).In both countries, the provision of functional road hierarchies, the separation of land uses, and the utilization of new building technologies characterized the planning objectives of this era (Fishman, 1987). Thus, favorable political and institutional conditions encouraged developers to take advantage of improved automobile accessibility at the metropolitan scale. Development of unprecedented large-scale suburban “new towns” (including various Levittowns in the US, and Don Mills north of Toronto) are the outcomes of these conditions (Harris, 2004; Muller, 2004). These trends continued throughout the post-war period and led to the widespread realization of car-oriented, low-density, single-use suburbs now referred to as suburban sprawl (Newman & Kenworthy, 1996).

Although initial attempts to “universalize suburbia” dated back to the 1920s and the initial auto industry “boom” as Fishman (1987) and Muller (2004) suggest, it was the simultaneous occurrence of favorable economic and political conditions, and conformity of interests among key actors (including all levels of governments and planning institutions, middle-class citizens and entrepreneurs) that helped massive peripheral developments to be realized in the post-war period.

2.2.2 From 1975 to 1990: Deindustrialization

In 1970s and 1980s, a series of significant changes in economic, political and social spheres marked a critical transition from the prior booming decades to a period of economic “slump” (Fishman, 1987). This era was marked by two major crises in North America: first, a long period of combined economic stagnation and inflation (known as “stagflation”) occurred in tandem with energy crises (Bunting & Filion, 2010). Second, national economies deindustrialized, and many jobs moved to countries with lower wage levels due to different structural economic conditions (Bunting & Filion, 2010). In the context of metropolitan development, the energy crises of 1970s profoundly impacted millions of suburban commuters, revealing the vulnerability of dominant car-dependent development to oil price shocks (Hodge & Gordon, 2008; Kenworthy, 2007).

This was a period when conservative leaders held federal office in the US (Reagan), UK (Thatcher) and Canada (Mulroney). The prevalent political conservatism coinciding with the consequences of aforementioned crises fueled support for the political ideology of neoliberalism, which aimed to restrict government and deregulate economic activities by the end of this era (Bunting & Filion, 2010). By limiting the mandate and capacity of public investment, this transformation had lasting impacts on governments' role in metropolitan dynamics, and heightened the role of the private sector in return (Grant & Filion, 2010). This also resulted in the advent of a new relationship between government and enterprises, which further formulated in form of public-private partnership initiatives and widely influenced urban transportation infrastructure investments in both Canada and the United States afterwards (Hodge & Gordon, 2008; Siemiatycki, 2006).

Despite the overall convergence, differences between the structure of local government in Canada and the United States had considerable impacts on metropolitan development trajectories in the two countries during this period. In the US, a more decentralized political system had left a great degree of autonomy and fiscal independence to metropolitan municipalities (Dilworth, 2009); in Canada, institutional arrangements were less well-suited to implementing explicit metropolitan-based policies (Donald, 2005). This encouraged the US metropolitan authorities to leverage the capacity of suburban land incentives to subsidize their regional growth to a larger extent than in Canada (Jonas, 1991).

In the US, by the mid-1970s, the spatial structure of metropolitan regions had been mostly shaped in the dispersed form that is observed today, yet major dynamics were taking place within cities' social and economic structures (Filion, Bunting, et al., 2004). Reduced or deferred investments in infrastructure and social services had led to deteriorating quality of life in inner cities, where social conditions were further stressed by high rate of crime and unemployment. In turn, larger groups of middle-class households were leaving inner cities to settle down in growing and socio-economically homogeneous suburbs. By the mid-1970s, achieving the critical mass of population and jobs in peripheries, reinforcing decentralization dynamics have been triggered in most metropolitan areas (Jackson, 1987). Concentration of disadvantaged residents in inner cities, along with decaying physical conditions pushed out greater number of retail establishments, firms and businesses to relocate in suburban malls, and office parks, where new clusters of activities were forming (Hartshorn & Muller, 1989).

This relocation of residents and firms incrementally shaped self-sufficient subcenters of outlying areas, and polycentric metropolitan structures emerged. Studies of Chicago (McDonald, 1987;

McMillen & McDonald, 1998), Cleveland (Bogart & Ferry, 1999), Los Angeles (Giuliano, Redfearn, Agarwal, Li, & Zhuang, 2007; Giuliano & Small, 1991, 1993; Small & Song, 1994), San Francisco (Cervero & Wu, 1997; McMillen, 2001), Houston (Craig & Ng, 2001; McMillen, 2001), and the metropolitan areas of Dallas, New Orleans (McMillen, 2001), Atlanta, Baltimore-Washington, Boston, New York and Philadelphia (Casello, 2007; McMillen, 2003) identified the widespread emergence of suburban employment centers and polycentricity across the US metropolitan regions by the end of 1980s. However, despite increasing suburbanization Canadian cities have not shown significant subcenter formation (Bunting, Filion, & Priston, 2002; Burns & Marcy, 1979).

In Canada, however, metropolitan dynamics were somewhat different. Although auto-accessible suburban communities built after 1945 attracted a large proportion of middle-class residents and increasing number of activities (Hodge & Gordon, 2008), the rapid arrival of immigrants with strong preferences for public transit access and cheaper housing in older Canadian urban neighborhoods helped to maintain the livability of inner cities (Grant & Filion, 2010).

Planning for urban renewal and public housing projects also played a critical role in the coordination of transportation infrastructure investments towards a more concentrated land use pattern in Canada. In larger metropolitan regions, these investments mainly consisted of road widening and public transit expansions. However, as suburban commuters began to outnumber inner city dwellers, demand for auto access to downtowns increased as well. Demolition of existing neighborhoods for highway expansions along with excessive parking provision in downtowns were the planners' response to this dynamic, albeit to a greater degree in the United States than in Canada (Bunting & Filion, 2010; Yeates, 1990)

Nevertheless, Canadian downtowns evolved in response to these trends as well. The transformed landscapes of central business districts (CBDs) in major Canadian cities, highlighted by the construction of iconic office towers (Bunting & Filion, 2010) indicated that firms and businesses were beginning to take advantage of improved accessibility and infrastructure renewal attempts in core areas. Despite downtown improvements, in relative terms, suburban jobs scattered in peripheral locations with proper highway access increasingly represented larger proportion of metropolitan employments (Hartshorn & Muller, 1989).

Although Canadian metropolitan regions did not grow as large as their US counterparts (Hodge & Gordon, 2008), the morphology of development patterns shows the uniformity among the outer suburbs built in this period (Filion, Bunting, et al., 2004). These developments can mainly be characterized as low population density peripheries that encompassed lower order employment and activity subcenters. Despite the overall similarity, Canadian metropolitan areas did not yet evolve towards polycentricity to a comparable extent in the subsequent period (Bourne, 1989; Griffith, 1981).

In summary, this was a period of urban decay, as North American economies deindustrialized in the face of stagflation and energy crises. However, firms' movement to suburbs fueled the ever-spreading sprawl and shaped new polycentric metropolitan structures as we observe today.

2.2.3 From 1990 to Present: Post-industrialization

While 1975-1990 was marked by economic difficulties and a significant spatial reorganization of urban structures and actors, the period from the 1990s to the present has changed metropolitan regions even more drastically. This period has been marked by increasing socio-economic and environmental challenges associated with North America's globalized economy. This evolution, coupled with the neoliberal political climate that emerged in the previous period and the growing awareness of the adverse consequences of increasing suburbanization, all brought critiques of car-dependent development strategies. The result has been that the notion of sustainable development has risen to the forefront of urban studies and the planning agenda (Bourne, 1996; Campbell, 1996).

Beginning in the early 1990s, a paradigm shift towards sustainable development took place. In light of that, main objectives of environmental protection, social equity, and economic prosperity have been widely adopted within North America's public policy domain (Krueger & Gibbs, 2008). Though, critiques have been articulated that the early adoptions of sustainable development idea to a large extent were mixed of superficial or symbolic use of the concept rather than genuine integration (Berke

& Conroy, 2000). However, many metropolitan regions have made efforts to reverse their long-standing decentralization trend through re-urbanization strategies of land use intensification and reduced car dependency (Filion & Kramer, 2012), hence, great investments made in sustainable mobility provision in Canada and the US (Baum-Snow, Kahn, & Voith, 2005).

In terms of the urban structure in this period, the elements of monocentricity, polycentricity and dispersion have emerged across major Canadian regions, while the job centers played a dominant structural role in larger regions such as Montreal, Toronto, and Vancouver metropolitan areas (Sweet, Bullivant, & Kanaroglou, 2017).

2.2.4 Summary

Cities have historically emerged in central locations and benefited from economic advantages of agglomeration and proximity of activities. As cities expand and their populations grow, they increasingly rely on transportation systems to facilitate their functions in providing more diverse and complex social and economic opportunities.

Spatial structures of cities in North America are mainly the legacy of rapid expansion and the movement of households and jobs to areas outside of traditional urban cores after the Second World War period. We reviewed this history of metropolitan evolution to portray the role of transportation infrastructure investments in the emergence of metropolitan polycentricity. We emphasized on the impacts of automobile technology in shaping dispersed urban structure, the trend has profoundly formed a pattern of car-dependent travel behavior across metropolitan areas as we observe today.

We traced the roots of extensive suburbanization which has caused downtown decline and urban decay of the late 20th century. We provide this review to contextualize the investigation we have done in Waterloo Region, as a mid-size metropolitan area with representative auto-oriented characteristics. This study provides future research with a baseline expectation from applying our developed models on analyzing comparable case studies, where the confluence of factors is needed to be considered to interpret the quantitative outcomes.

2.3 Integrated Land Use and Transportation Systems

This section begins with an overview of seminal theories that explain land use and transportation system interactions. It then explores empirical research that have examined case studies from Canada and the United States, and finally discusses the implications of land use and transportation integration for planning policy.

2.3.1 Theoretical Constructs

A number of contemporary planning theories represent cities as complex systems (Batty, 2008a), where human and institutional actors continually interact within relatively stable structures, and steer dynamics of cities at various spatial and temporal scales (Wegener, 2014). This concept helps to explain metropolitan evolution through the investigation of key actors' location and in some instances travel decisions.

The key actors and their behaviors relevant to this study are households and the people who live in them; firms who locate in metropolitan areas and hire employees; developers who supply space for activities in the area; and governments who influence the behavior of households, firms and developers. The dynamics of interest in this research are land use and transportation system changes that take place in a complex urban context, in association with other established structures including political, economic, and planning systems and social norms.

Households and residents make their location choices based on many factors, including the composition of the household, income, transportation costs, and locations of jobs, goods and services. Firms choose to locate where there are a diverse workforce and a suitable customer base. Naturally, there is codependence between the decision-making of these actors (Pagliara & Wilson, 2010).

Developers also react to (and sometimes lead) movements in markets that influence location choices for households and firms. All of these actors have concerns about the transportation system that is available to them. Specifically, transportation "costs" (measured as time, convenience, reliability and actual expenses) directly impact households' budgets and firms' operating cost structures (McCann, 1998). These dynamics make certain locations more (and less) attractive to households and firms. At a higher scale, governments are most often responsible for the supply and operation of the transportation system. The composition – the mix of modes available – as well as the system performance and costs are all directly controlled by them. Local governments also have an active role in establishing land use regulations and economic incentives (Bemelmans-Videc, Rist, & Vedung, 1998).

Classical economic theories suggest that transportation network expansion in general, and transit infrastructure in particular, impact land use changes through two major mechanisms: reducing generalized transportation costs, and altering relative accessibility levels (Guiliano, 2004). Both these factors significantly influence the location choices of residents and firms (Anas et al., 1998).

By reducing generalized cost (e.g., travel time and monetary costs), residents are able to commute longer distances given their budget constraints. Hence, they can trade housing size and commuting distance, and likely reside farther from central areas (Handy, 2005). On the other hand, altering relative accessibility levels changes the utility of different locations for different activities (Hansen, 1959). This potentially enhances the appeal of development close to transit stops (Handy, 2005).

Firms' location decisions are considered to be more complex; given the diversity of factors involved in the firms' production or service delivery (i.e., the business inputs and outputs) their responses to accessibility changes are vary among different industries. While manufacturing firms choose to locate farther in peripheries, retailing, commercial and service firms may tend to relocate in more central places (Anas et al., 1998).

These conflicting effects of the transportation system on relocation choices have generated diverging forces encouraging decentralization in some ways and acting as a counterforce in others (Handy, 2005). This brings the urban theory full circle: people and firms benefit from agglomeration and proximity advantages that cities and urban systems inherently provide (Filion & Bunting, 2010) which, in turns, drives concentration of activities in densely developed form. However, while the concentration of activities benefits people, firms and businesses, it comes with the costs of land capitalization and higher land values in more accessible places, pushing developments further away from central places and stimulating emerging new growth centers (Alonso, 1964; Harvey, 1985).

Ultimately, from a planning perspective, it is useful to be able to predict, generally, the behaviors of actors in these complex systems, particularly in terms of how they may react to changes in system inputs. Obviously, the problem is difficult. In an increasingly diversified society, the stakeholders may be too diverse to be conceivably categorized into the three groups (Campbell, 2016) residents, firms and developers, assessed here. An additional way of trying to understand urban dynamics moves from the theoretical to applied – including the conduct of empirical studies.

2.3.2 Empirical Studies

As discussed, theories offer some insights that help to draw conclusions on land use and transportation system interactions. To complement that work, a variety of empirical case studies have been conducted to untangle these interactions, particularly providing evidence on the location decisions of key urban actors. This reviews empirical studies that have examined LUTI dynamics from cases in the US and Canada.

Knight and Trygg (1977) conducted an early review of the built environment impacts of transit system investments in North America. They concluded that in the presence of ubiquitous roads and expressways in contemporary metropolitan areas, the accessibility improvements driven by rapid transit investment tend to have marginal impacts on land use developments. The authors suggested that to induce and sustain desired land use developments, transit investments need to be accompanied by pro-development municipal policies. Moreover, a positive steady regional development trend and the availability of developable land and favorable physical site characteristics are required for the success of transit-driven development.

While investment in light rail transit (LRT) systems was gaining momentum in the 1980s, Cervero (1984) explored motivations for and land use impacts of 12 planned or in operation LRT systems across North America. He concluded that LRT systems (which feature fixed-alignment, permanent infrastructure) have inherent potential to trigger land use changes, promote compact development, result in property value growth, and stimulate redevelopment. The study indicated, however, that these potential land use effects can be realized only if other factors are already in place. A growing regional economy, supportive development policies (e.g., appropriate zoning and tax incentives), automobile discouragement strategies (e.g., parking restrictions), suitable physical conditions, and the availability of land adjacent to station areas were all identified as important factors to trigger land use changes in the presence of LRT. He also suggested the potential land use effects of LRT are most likely to be realized in downtowns and central areas.

Studying the Bay Area Rapid Transit (BART) system, one of the very first post-war urban rail mega projects, Cervero and Landis (1997), evaluated land use impacts of the system against the project objectives. While the project major motivation was defined as shaping “multi-centered settlement pattern.” (Cervero & Landis, 1997, p. 309), the authors stated that 20 years of system operation had produced only a moderate desirable impact on residential development in the area. In terms of redistributing employment, the system had encouraged employment growth in downtown San Francisco. However, with the exception of a few suburban stations, major growth took place within 800m from regional highway access points.

In a comprehensive review, Huang (1996) reported that rail transit planning in North America in many cases occurred without concurrent land use planning, these investments do not necessarily generate land use changes. Rather, infrastructure is planned to serve current populations, and their benefits may be limited to farebox revenues and property taxes.

In contrast to Huang’s (1996) findings, Cervero & Duncan (2002) reported that public transit systems are generally considered to be development catalysts, and their impacts are associated with increasing land values and higher rent premiums. Studying commercial properties in Santa Clara County, California, the authors reported a capitalization premium of 120% above the average regional property value growth rate for parcels in commercial business districts. These parcels were located within walking distance (800m) from commuter rail stations with upgraded express service. The authors also examined property values in the vicinity of new LRT stations and observed a modest yet significant increase of property value increases over 20% above the average regional growth rate.

Analyzing the trend of vacant land value changes in proximity to planned LRT stations , Knaap et al. (2001) reported significant land value growth in Washington County (Portland), Oregon. Similarly, studying land use impacts of LRT development in Minneapolis, Hurst and West (2014) observed significant land use and densities changes, and positive real estate market trend within the new LRT corridor, during six years of the system operation. Lee and Sener (2017) analyzed the land use impacts of LRT development in Houston – a metropolitan city without zoning – and found a spike in commercial development within new LRT corridors (800m buffer), whereas, only small changes were reported in high-density residential development near LRT stations.

Given the potential implications of property value changes in measuring impacts of transportation infrastructure investment, Higgins and Kanaroglou (2016) conducted an intensive review on four decades of land value uplift (LVU) studies in North America, and reported a significant heterogeneity in the studies' outcomes.

In a more recent study, a multi-decade longitudinal study on co-evolution of the transportation networks and land use in the Greater Toronto-Hamilton Area (GTHA) has found that the initial impacts of the transportation accessibility on urban development have increasingly decreased over time, while proximity to built up areas has had more lasting impacts on the land use development (Kasraian et al., 2020). Furthermore, as new mobility modes emerge studies have been done to forecast potential impacts of new mobility options for the LUTI dynamics. The results indicate that ridesharing and Automated Vehicles (AVs) at the forefront of changes are likely to increase vehicle miles traveled VMT and to reduce public transit modes share. The land use impacts of these new modes are expected to be significant as well. These impacts are likely to lead to more dispersed urban growth patterns in well-connected distant suburbs and rural settlements (Soteropoulos, Berger, & Ciari, 2019). Importantly, adoption of the shared mobility services over time has shown to be more influenced by demographic factors rather than the spatial land use patterns (Sweet & Scott, 2021).

2.3.3 Planning and Policy Implications

An increasing number of planning policies draw on the interrelated nature of land use and transportation systems. These policies aim to achieve sustainable urban form and sustainable mobility through better integration of land use and transportation planning. This approach is often referred to as integrated planning, which has been articulated under three broad concepts of New Urbanism, Smart Growth and Transit Oriented Development (TOD).

New Urbanism is primarily concerned with design elements and aesthetic improvements of neighborhoods, though, its ultimate goal is to recreate communities around convenient public transit and to mitigate urban sprawl (Handy, 2005). Similarly, Smart Growth as a government-driven strategy aims to direct growth into existing urban areas, and to improve the viability of public transit. Smart Growth promotes compact and vibrant urban environments that are served by sustainable transportation modes including active and public transit options (Duany, Speck, & Lydon, 2010).

In line with Smart Growth principles, Transit Oriented Development (TOD) suggests that urban developments be concentrated around higher-order (higher capacity, greater infrastructure investment) public transit nodes (e.g., BRT, LRT or Metro stations). This strategy helps to achieve improved levels of accessibility by encouraging higher densities and diverse choice of activities within a walkable and bike-friendly area. TOD promises to upgrade the quality of urban environments, and to utilize urban land and infrastructure in a more efficient manner (Evans, Pratt, Stryker, & Kuzmyak, 2007). Transit-induced densification has been shown to improve economic productivity as well (Chatman & Noland, 2013; Graham, 2007). However, studies suggest that higher density development may not significantly influence travel behaviors unless the level-of-service of the roads is reduced and parking requirements are eliminated (Boarnet, 2011; Chatman, 2008).

Depending on the political and institutional contexts, different levels of governments may use a range of measures to implement integrated planning policies including policy instruments and public investment. Policy instruments themselves comprise regulatory and economic measures (Bemelmans-

Videc et al., 1998). Table 2-1 summarizes the potential policy measures that are used to support integration of land use and transportation planning, with distinction of their implications for either land use or transportation.

Instrument type		Land use related	Transportation related
Policy instruments	Regulatory measures	Zoning (Stull, 1974) Urban growth boundaries (containment policies) (Dawkins & Nelson, 2002) Greenbelts and environmentally sensitive areas boundaries (Ding, Knaap, & Hopkins, 1999)	Mobilization of collective agency (Addie, 2013) Parking bylaws
	Economic measures	Tax incentives (Reese, Larnell, & Sands, 2009) Development charges Transfer of development rights (TDR) (Shahab, Clinch, & O'Neill, 2018) Non-financial compensation schemes (van der Veen, Spaans, & Janssen-Jansen, 2010)	Fuel pricing Public transit pricing
Public investments		Affordable housing development Social service provision Public land acquisition (Bengston, Fletcher, & Nelson, 2004)	Road network development Public transit development (Cervero & Duncan, 2002) Light Rail Transit Development (Cervero, 1984)

Table 2-1- Policy instruments and public interventions overview (Source: author)

While both transportation interventions and land use regulations are ideally to be implemented simultaneously, in practice, land use policy instruments are more diverse and widely adopted, ranging from zoning and development charges to greenbelts and urban growth boundaries. Instruments such as non-financial compensation and transfer of development rights (TDR) are often applied in practice to a lesser extent due to sophisticated regulatory requirements (Janssen-Jansen, Spaans, & van der Veen, 2009). On the other hand, zoning by far is the most common employed instruments among all (Qian, 2010). The land use instruments have more spatially concentrated impacts rather than desirable distributed effects. In contrast, transportation interventions potentially have wider and more socially inclusive and fair impacts (Lucas, 2012). In fact, temporal variance and the transportation changes precedence often gives an opportunity to planners and policy makers to leverage transportation interventions as a guiding force to shape desired land use related changes.

2.3.4 Summary

As discussed in earlier sections, land use and transportation system interactions are complex and to understand their roots and consequences it is necessary to attain multidimensional insights. As we approach our research questions, we conducted a series of reviews to provide this study with the required understandings on the broad context of LUTI problems. We explored a continuum of aspects from theoretical perspectives to empirical findings, and research methodologies.

Land use and transportation system interactions are often analyzed using economic theories. This ranges from more aggregate classical location choice theories to more recent behavioral approaches. Although, these theories offer some insights that help LUTI studies, these interactions are more complex and difficult to understand purely theoretically. Therefore, variety of empirical case studies have been conducted to untangle these interactions, providing evidence on the travel behavior and location decisions of key urban actors and their potential relationships.

However, reviewing empirical studies we found that a very few instances were able to comprehensively analyze the land use and transportation systems elements concurrently and with the balanced representations. Most studies had focused on only one or a few land use types. Moreover, while the auto and public transit modes have received the most attention in LUTI studies, other transportation modes (e.g., active and non-motorized modes) were often overlooked. As our review identified, these issues while prevalent, are potentially remediable. Since the need for more robust quantitative indicators and spatial models for improved LUTI analysis is evident through this review, we positioned our study to respond to this need through development of novel spatial modeling solutions.

2.4 Quantitative Models to Support Land Use and Transportation System Interactions Analysis

Land use and transport system interaction (LUTI) modeling framework has historically provided a broad methodological basis to understand the complex evolution of metropolitan areas at various spatial and temporal scales (Wegener, 1998). While this framework has been mainly informed by theories and quantitative methods from economics and geography, in recent days, it has increasingly adopted more interdisciplinary approaches such as complex systems (Batty, 2008b; Wegener, 2004), microsimulation and activity-based method (Acheampong & Silva, 2015; Miller, Douglas Hunt, Abraham, & Salvini, 2004).

To date, numerous operational LUTI models have been developed and employed in planning research and practice (Koomen & Borsboom-van Beurden, 2011). Hence, in the past three decades, an increasing number of research and review papers have been published to address theoretical and methodological advancements and challenges of these models (Acheampong & Silva, 2015; Hunt, Kriger, & Miller, 2005; Iacono, Levinson, & El-Geneidy, 2008; Nourian, Alipour, & Ache, 2021; Wegener, 1994, 2014).

LUTI modeling approaches are diverse in their conceptualization of spatial dimension, and on treatment of the key urban actors interacting within the dynamic urban environment. Depending on the models' construct and abstraction levels, these models range from those that treat the spatial dimension and key urban actors implicitly (such as statistical models) to explicitly (such as agent based models) (Wegener, 2014). This review includes those major LUTI modeling approaches that are built upon spatial conceptualization and take into account the key urban actors explicitly. According to these criteria statistical regression models, deterministic cellular automata (CA) and stochastic Markov chain models, as well as artificial neural networks (ANN) models that lack such explicit conceptualizations (van Schrojenstein Lantman, Verburg, Bregt, & Geertman, 2011) remain outside the scope of this study.

The review starts with an introduction of early economic location models, followed by an exploration of the progress made through incorporation of spatial interaction, complex systems and random utility theories. Then, agent-based and microsimulation models are discussed as disaggregated alternatives.

2.4.1 Location Choice and Spatial Interaction Models

Within the LUTI modeling framework, economic, and spatial interaction models were the first generation of quantitative methods built upon static location theories that aimed to describe the land use and activity locations in respect to transportation costs (Wegener, 2014). Although, these models could not capture the land use and transportation system interactions in their full depth, they have contributed to better understanding of the behavior of key actors (i.e., residents and firms) involved in those interactions (Pagliara & Wilson, 2010). The most influential examples of these models include the work of von Thunen (1826) on optimal allocation of production activities based on transportation costs, and Weber's (1929) least cost model of industrial location. In the 1960s, Alonso (1964) introduced a residential location model constructed based on bid rent theory. This model evolved to further incorporate more progressive residential location models (Anas, 1982; Mills, 1967; Muth, 1969). The successor variants of these models such as DRAM and EMPAL are still in use for policy analysis purposes at US metropolitan regions (Putman, 2010).

Two seminal discoveries helped link transportation and land use in quantitative models. First was the introduction of transportation accessibility as a representation of the intensity and opportunities for interaction (Hansen, 1959). Next were the concepts of gravity and entropy maximization to estimate spatial flows according to the distance and size (e.g., population) of given origins and destinations (Wilson, 1971). Together, these were major theoretical improvements that facilitated conceptualization of land use and transportation system dynamics at regional scale. Lowry’s (1964) Model of Metropolis was the first operational LUTI model of a kind, and initiated a string of comprehensive models. This type of models has been further expanded upon to include random utility theory and discrete choice models; these techniques are still used today (Engelberg, He, Le, & Zegras, 2021).

The introduction of the representation of actors as utility maximizing agents, modeled through random utility theory, opened up new opportunities to conceptualize fine scale decisions such as residential location choices (McFadden, 1978), and transportation mode choices (Ben-Akiva & Lerman, 1985). These innovations helped later modeling approaches better formulate their internal dynamics (Iacono et al., 2008).

2.4.2 Complex Systems Theory and System Dynamics Models

Parallel to development of economic based approaches, in the late 1960s, Forrester (1969) suggested that growth and evolution of cities can be better analyzed and understood through a system dynamics approach compared to the static view of the models that had been developed by that time. A system dynamics approach essentially aims to understand the complex and nonlinear systems' behaviors over time (e.g., land use change), through the actions of feedback loops and accumulations of stocks (e.g., population), considering the rate of flows (e.g., housing relocation or migration rate) (Forrester, 1969).

According to this perspective, land use and transportation systems are the two major subsystems of a city that continuously interact over a spectrum of spatial and temporal scales. The built environment and transportation networks experience the slowest interactions and have the lowest rate of changes while economic and demographic changes including peoples' and firms' relocations take place at intermediate speeds. Ultimately the movements of people and goods (mobility) take place in a quick and flexible manner, which tend to produce rapid impacts on other urban subsystems (Wegener, Gnad, & Vannahme, 1986).

The methodological simplicity and structural modularity of System Dynamics models motivated their application to study dynamic urban phenomena (Elsawah et al., 2017; Haghani, Lee, & Byun, 2003). The Metropolitan Activity Relocation Simulator (MARS), for example, is one the latest SD models that has been developed to assist long-term land use and transportation policy analysis in metropolitan areas (Pfaffenbichler, Emberger, & Shepherd, 2010).

2.4.3 Agent-Based Models (ABM)

In response to structural limitations of theory based models, Agent-Based Modeling (ABM) was introduced as an exploratory method to explicitly simulate individual decision making processes (Benenson & Torrens, 2004). This was a significant shift from a structural perspective to a behavioral representation of real world phenomena, that enabled researchers to replicate behaviors of complex and nonlinear systems (e.g., land use change) through bottom-up processes (Batty, Crooks, See, & Heppenstall, 2012).

ABMs' abilities to represent urban stakeholders and their decision-making processes at fine spatial and temporal scales offer unique advantages for modeling complex land use and transportation dynamics, through efficient simulation of travel behaviors, relocation choices, as well as land development decisions against various policy scenarios (Miller et al., 2004). However, it is a common challenge for ABMs to integrate feedback operations across various spatial scales, and also to properly aggregate the fragmented outcomes of the model processes (Parker, Berger, & Manson, 2001; Stanilov, 2012a). To alleviate these issues, coupling and integrating ABM with other modeling approaches, along with exploiting spatial dimension as a mediator to capture cross-scale feedbacks have been suggested as a methodological solution (Filatova, Verburg, Parker, & Stannard, 2013). The integration of ABMs in microsimulation models has already shown great potentials for land use and transportation system interaction studies (Miller et al., 2004; Wegener, 2014). Though, due to inconsistency of the modeled dynamics' scales, not all researchers agree on suitability of such integration (Batty, Crooks, et al., 2012).

2.4.4 Microsimulation models

As Miller (2003) argues, microsimulation is more of a modeling approach rather than a model per se. Urban microsimulation models attempt to integrate various disaggregate sub-models to benefit from their specific advantages in capturing processes at different spatial and temporal scales (Moekel, Spiekermann, Schürmann, & Wegener, 2003).

State of the art microsimulation LUTI models such as ILUTE and UrbanSim, feature explicit representations of key stakeholders including households, firms and developers that interact in housing and real estate markets. Firms' economic activity, as well as individual agents daily activities and their associated travel behaviors are simultaneously modeled (Miller et al., 2004; Waddell et al., 2003).

However, despite their great potential, there are theoretical and methodological challenges associated with urban microsimulation models that have yet to be addressed. Most importantly, the applicability of random utility theory as the core decision making concept for simulating major decisions such as residents and firms' relocation choices is questionable (Timmermans, 2003).

Moreover, though the reciprocal nature of land use and transportation system interactions, operational models often treat the transportation system (in terms of both physical infrastructure and mobility options) as fixed and an exogenous factor to the model (Miller, 2014). This restricts the abilities of the models to reflect emerging trends, for example shared and on-demand mobility services in an increasingly multimodal transportation landscape. Similarly, decisions (behaviors) of local governments and public agencies are usually considered as the given input to the model rather than the model-driven behaviors (Miller, 2004). Comprehensive microsimulation models have been also criticized for their high level of complexity, and extensive data requirements (Birkin & Wu, 2012).

Despite these issues, to date microsimulation models have provided reasonably accurate estimates of urban land use and transportation system dynamics, and the approach remains promising as a viable solution to inform major policy making processes (Birkin & Wu, 2012; Iacono et al., 2008).

2.4.5 Data Collection Methods to Support Spatial LUTI Modeling

The development of quantitative urban models as selective representations of real-world systems and processes, involves a high degree of abstraction and generalization (Burg & Stolk, 2004). These models need to adequately represent the urban demographic processes, the built environment, and functional urban infrastructure systems that carry flows of people, goods, and information (Batty, 2008b). They should also be able to consistently reproduce behaviors of the urban subsystems and actors within the scope of model applicability (Batty, 2012). Although theories play a major role to build a valid model, empirical data are essential to ensure the consistency of the models with reality as well (Voinov, 2008).

As Miller (2004) suggests, historical data from multiple time periods are essential elements for the purpose of LUTI model estimation – determining the model parameters statistically - as well as for the model validation - confirming conformity of the model outcomes versus the observed real-world phenomenon. He stresses that having a set of data gathered at three points in time with comparable specifications is necessary in this regard. The data from two time periods are used to establish the model construct (model estimation), and the third one for validation purpose.

Given the inherently broad scope of LUTI models, they often incorporate numerous variables from a range of domains. A generic LUTI model data inventory includes built environment and land use data, transportation network and travel diaries, demographic and employment records, market information and prices indicators from certain sectors (Haase & Schwarz, 2009; Waddell, Ulfarsson, Franklin, & Lobb, 2007). LUTI models are also developed at various scales. The scale in this context is defined not only in terms of the spatial and temporal dimensions, but also in respect to attributes of the modeled system, its associated agents and processes. Accordingly, the required data to support the modeling practice need to be collected with adequate granularity and specification (Miller, 2004).

For aggregate LUTI models, given their coarse spatial and temporal scales, the conventional data collection scheme could potentially support modelling efforts, if adequate historical data were available (Iacono et al., 2008). In such a scheme, aggregated population and employment data were collected through surveys or census every 5 or 10 years. Similarly, transportation activities were largely chronicled using the Census journey to work information or specifically designed Origins-Destination surveys. Further, built environment, and land use data were often provided by municipal authorities, and might be updated in annual basis.

However, Waddell (2011) argues that urban data sets have often been incomplete and subject to error. He estimates that roughly 75% of the effort and time involved in LUTI model development are still spent on data preparation and resolving data integration issues. Although these issues may partly be solved through the adoption of crowdsourcing and open data initiatives, their data accuracy yet to be addressed (Batty, Axhausen, et al., 2012).

For models designed with disaggregate approaches (e.g., agent-based and microsimulation) conventional data collection practices pose more significant challenges to satisfy the models' data requirements. Thus, in the past decade, large efforts have been made to collect and synthesize appropriate disaggregate data, especially on demographic and travel behavior aspects (Miller, 2014).

While disaggregate models tend to explicitly represent households or individuals and their heterogeneity of preferences, they need to be provided with demographic data at a micro scale (Miller, 2004). However, the census often doesn't provide data at such scales due to confidentiality concerns. To address this challenge, systematic population synthesis methods such as Iterative Proportional Fitting (IPF), and Combinatorial Optimization (CO) have been developed to generate sets of model agents that are statistically consistent with the observed demographic data (Pritchard & Miller, 2012; Ryan, Maoh, & Kanaroglou, 2009).

In terms of transportation system and travel behavior data, passive data collection methods for automobile systems are well established and widely used. For transit systems, modern data collection methods include Bluetooth detectors, along with Automatic Vehicle Location and Automatic Passenger Counting (AVL/APC) technologies. These approaches have far improved macro level transit performance data collection (Bachmann, Roorda, Abdulhai, & Moshiri, 2013). Miller (2014) emphasizes that unbiased estimates of travelers' behaviors can only be achieved from a model informed by disaggregate data at the individual trip maker scale. While collecting such data at large scale was technically infeasible before, rapid advancements of mobile devices and embedded sensors has diminished barriers to obtain those individual level data (Batty, Axhausen, et al., 2012).

The application of GPS loggers has shown great potential to identify travel modes and trip purposes in complex urban environments (Chen, Gong, Lawson, & Bialostozky, 2010). The widespread adoption of smartphones with embedded GPS units has advanced the data collection capabilities. Hence, a wide range of methods have been developed to extract more sophisticated travel behavior features (Wu, Yang, & Jing, 2016). These methods improved data collection opportunities particularly for sustainable transportation modes. This ranges from detection of transit trips (Nour, Hellinga, & Casello, 2016) and identifying transit route choices (Zahabi & Patterson, 2016), to tracing cycling activity and path choices (Casello & Usyukov, 2014) and capturing pedestrian tours (Xu, Casello, & Fard, 2021). However, Miller (2014) suggests that, as long as the planning and policy analysis needs to have better and more disaggregate models, demand for more disaggregate data tends to continue.

2.4.6 Spatial Representation in Quantitative Urban Models

Aggregate spatial modeling practices use either regular grid tessellation or irregular zonal representations to conceptualize the space within their constructs. Historically, transportation analysis and modeling rely on pre-defined spatial zones to convey variables that are being analyzed. This zonal representation helps to conceptualize geographic space as it relates to the built environment, on the ground activities and transportation flows. To define these zones, one can think of a continuum of spatial scales ranging from land parcels to city blocks, further to postal code boundaries and census tracts, up until the electoral districts and census divisions (Wolf, Knaap, & Rey, 2019).

However, depending on the scale and concept of space abstraction (Figure 2-1), both grid tessellation and aggregated zonal representation utilize the smallest unit of space (cell or zone) to portray only the dominant or aggregated characters of an area in real world geographic space (e.g., major land use type). Therefore, these types of representation overlook actual differences and attenuate heterogeneities at various levels (Koomen & Borsboom-van Beurden, 2011).

Moreover, the required spatial aggregation and disaggregation operations to transfer information between spatial scales inevitably involve information loss (Figure 2-1c. versus d.), which increase uncertainty and may result in random errors as well as systematic inconsistencies in spatial modeling practices (Jacobs-Crisioni, Rietveld, & Koomen, 2014; Järv, Tenkanen, & Toivonen, 2017).

Therefore, both concepts are subject to inherent issues from the uncertain geographic context problem (UGCoP) (Kwan, 2012) and the modifiable areal unit problem (MAUP) (Openshaw, 1984), to boundary pixel and sub-pixel issues (Fisher, 1997), spatial scale effect issue (Kok & Veldkamp, 2001), and aggregation bias or ecological fallacy (Wrigley, Holt, Steel, & Tranmer, 1996).

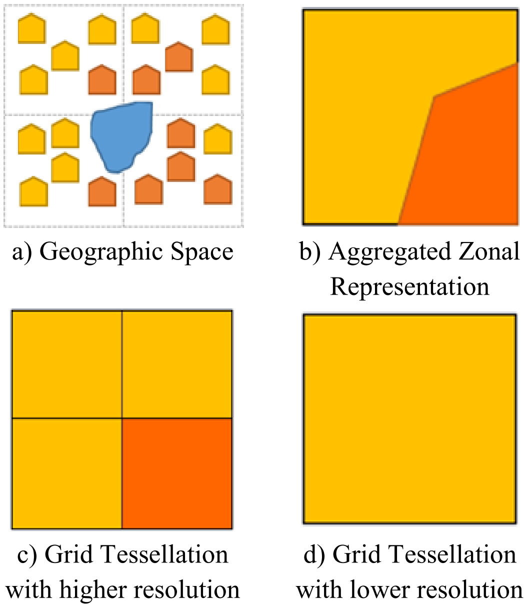


Figure 2-1 - Different approaches for spatial abstractions of geographic space

2.4.7 Summary

In recent years, the planning community worldwide has made efforts to reverse the challenges associated with car-dependent development strategies. They initiated numerous solutions including re-urbanization and TOD policies, which increasingly rely on public participation and evidence-based assessment of the development scenarios (e.g., What-If analysis). The final part of the literature review is focused on studying quantitative approaches that are used to inform planning and policy decisions, especially in the controversial area of large scale public infrastructure investment such as the Hamilton's revived LRT project or the reduced Hurontario LRT plan, both located within the GTHA.

Quantitative models are used as analytical tools to help evaluate envisioned development scenarios and often involve simulating key actors interacting within dynamic urban environment. These approaches are diverse in their conceptualization of space, and on treatment of the actors interacting within the models. This ranges from highly implicit to highly explicit alternatives.

Appropriate empirical data are central to support the modeling practices, in terms of both model construction and validation, as well as their forecasting application. The methodological advancements tend to make models more disaggregate, this helps them to better represent the real-world actors and their decision-making processes. Moreover, spatial representation and scale are crucial elements for reliable quantitative LUTI analysis.

To date, limitations of conventional data collection methods have been among major modeling challenges, thus further advancement of urban modeling especially in the context of land use and transport system interaction, is tied to progress of disaggregate data collection and spatial modeling methods. The recent decades have witnessed rapid advancements in digital sensors, and their integration with smartphones. While this trend has opened up opportunities for inexpensive large-scale urban data collection, the methods to efficiently analyze these data yet to be developed.

This research is a step towards better articulating spatial unit of analysis in LUTI studies by addressing and mitigating some of the significant methodological challenges including spatial scale effect issue and modifiable areal unit problem (MAUP).

2.5 Implications of Heterogeneity Measurement for Analysis Zones Delineation

The need for better conceptualization of the built environment within transportation analysis and modeling has been frequently discussed in the literature related to land use and transportation system interactions. Although the appropriate delineation of zonal boundaries still remains a major challenge in spatial analysis (Garreton & Sánchez, 2016; Wei, Rey, & Knaap, 2021), methods are available to quantify the corresponding effects and uncertainties (S.-I. Lee, Lee, Chun, & Griffith, 2019), and the adverse impacts of arbitrary choice of analysis zones and inappropriate spatial scale on the derived policy recommendations have been documented in previous urban studies (Gehrke & Clifton, 2016). However, addressing these issues has been largely overlooked in transportation analysis field (Pereira, Schwanen, & Banister, 2017).

The summary of practice, the known issues and their potential solutions have been reviewed in the works of Singleton and Clifton (2013), Gehrke and Clifton (2016) and Nowosad and Stepinski (2018). Ewing and Cervero (2010) have also offered a comprehensive framework constructed of seven dimensions whereby transportation and the built environment may interact. These dimensions – that referred to as “7-D’s” – include more agreed upon factors of density, diversity, and design, further extended by *destination accessibility*, *distance to transit, demand management* and *demographics*. However, the extent to which each of these factors contributes to the overall LUTI dynamics, and the potential collinearity among these variables have shown to be case-specific and dependent on the contextual environment (Handy, 2018).

Despite differences in the reviewed studies, incorporating land use diversity, enhancing spatial model constructs along with better aggregation methods have been central to the suggested solutions for appropriate LUTI investigation. Looking through the lens of activity-based analysis, land use diversity along with residential and employment densities are the most significant aspects of the built environment that influence travelers’ mode choice. As a result, we often observe contrasting travel patterns across different spatial scales (Ewing & Cervero, 2001; Gehrke & Clifton, 2014). This observation has led to many municipalities trying to achieve land use intensification, increasing the density and diversity of land uses such that more household activities can be accomplished with less (in distance and time) travel. Moreover, intensification also tends to promote the linking of multiple individual trips into a trip chain that again reduces per capita vehicle travel and energy consumption (Banister, 2011b). Given this relationship, it is logical to try to embed a measure of the homogeneity (or heterogeneity) in the decision making when creating zonal structures.

One such measure of heterogeneity is an Entropy index, a well-established means to quantify diversity as we perceive it within natural landscapes (Yoshida & Tanaka, 2005). This index was originally introduced in the context of information theory and widely adopted in scientific practices including urban and transportation modeling (Batty, 1974). We employ the concept of Entropy as it reveals heterogeneity levels and helps to create better transportation analysis zones.

In common practice, Traffic (or Transportation) Analysis Zones (TAZs) are devised to partition geographic space into contiguous non-overlapping polygons that cover the area of study. These polygons then are used to represent the attributes of these areas using required variables at an aggregate level. To ensure validity of the TAZs in representing the built environment and on the ground activities, there are common criteria to be considered. These include spatial contiguity and compactness of the zones, their alignment with other established boundaries such as municipal and political districts, and overall homogeneity of characteristics within the created zonal structures (You, Nedović-Budić, & Kim, 1997, 1998).

As Gehrke and Clifton (Gehrke & Clifton, 2014) and Clifton et al. (Clifton, Singleton, Muhs, & Schneider, 2016) suggest, it is crucial to find a right balance between tendencies of using more disaggregate scales versus simplifications that operational transportation models require. This leads to a practical question of what set of criteria should be used to determine the size and number of Traffic Analysis Zones. While in urban transportation studies there are motivations to identify a smaller number of larger areas that are homogeneous, it is of the same importance to recognize highly heterogeneous spaces where multiple activities take place within a small spatial area. As an example of the implications of inappropriate spatial scales, failing to consider highly heterogeneous areas may overlook areas that accommodate a large number of pedestrian tours, which, while important, are often not included in transportation modeling practices (Clifton et al., 2016).

Historically, classic urban models that built upon the concept of Central Business District (CBD) use employment density as the primary criterion to define their functional spatial clusters (Alonso, 1964; Mills, 1967). A next generation of studies that developed based on residential location models extends their criteria by incorporating population density along with employment densities (Anas, 1982; Muth, 1969). As computational power has risen, more recent works tend to employ iterative aggregation methods to identify analysis zones using more sophisticated parameters.

This study is inspired by the method suggested by Giuliano and Small (1991), and enhanced by Casello and Smith (2006) to construct urban activity centers – the zones that potentially produce and attract regionally significant number of trips. We also adopt an iterative spatial aggregation method developed by Casello and Fard (2017), and incorporate the land use diversity measurement method proposed by Yoshida and Tanaka (2005) to assess homogeneity of the aggregated zones.

2.6 Alternative Spatial Regionalization Methods to Define Analysis Zones

Recall that due to the spatial nature of land use and transportation system interactions, conceptualizing geographic space in form of zonal structures has been always a critical task in LUTI studies. In response, a range of solutions from the ad-hoc to more mathematical optimization approaches have been developed and used in urban studies (Wolf et al., 2019). An overview of conventional ad-hoc methods was presented in previous section. This section describes a set of more mathematical approaches that are used in the analysis zone definition practices known as spatial regionalization methods; "region" in this context implies a set of homogeneous and spatially contiguous geographic objects (Wei et al., 2021). In general, these solutions are classified under the quantitative clustering techniques with explicit or implicit spatial elements.

Spatial clustering techniques pertaining to this study are the ones that explicitly ensure the spatial contiguity objective, known as spatially constrained clustering algorithms. These methods include graph-based spatial partitioning method (also known as SKATER - an acronym for Spatial' K'luster Analysis by Tree Edge Removal) (Assunção, Neves, Câmara, & da Costa Freitas, 2006), the max-p-region method (Duque, Anselin, & Rey, 2012), the Automatic Zoning Procedure (AZP) (Openshaw, 1977) and the Evolutionary Multi-objective Optimization (EMO) method (Roberts, Hall, & Calamai, 2011). These methods basically aim for aggregating neighboring spatial objects (that are smaller in size) as the input to a model to create larger regions. These models are subject to one or more optimization criteria such as maintaining spatial contiguity, maximizing intra-zonal homogeneity (i.e., similarity of some attributes) and maximizing inter-zonal dissimilarity, maintaining a desirable compactness, and achieving a pre-defined number of clusters (Duque et al., 2012; Wei et al., 2021).

As the number of these criteria that are applied in a certain analysis increases, the complexity of the optimization methods increase considerably (Guo, 2008). Here, considerations need to be taken into account to ensure the computational feasibility of the optimization processes. For the methods that use rigorous mathematical optimization, the size of solution space to search for the optimal solution is exponentially scaled with respect to the degrees of adjacency and number of input zones. Moreover, depending on the selected optimization strategy, there is a potential outcome in which the model may get trapped into the sub-optimal solution and fail to identify the global optimal (i.e., best solution) (Aydin, Janikas, Assunção, & Lee, 2021). For the methods with a pre-defined number of clusters, determining this exact number is a subjective choice that can potentially impact reliability of the results. Accordingly, such algorithms are usually applied iteratively and are required to perform a large set of trial-and-error solutions before achieving satisfactory results. Hence, these clustering algorithms are often computationally intensive and demand processing resources beyond what ordinary workstation computers offer (Wei et al., 2021).

In recent years, spatial clustering and regionalization methods have been increasingly adopted in urban studies. These methods have shown effectiveness in a range of applications from travel behavior analysis (Guo, Zhu, Jin, Gao, & Andris, 2012; Kim & Yoon, 2021) to land market research (Helbich, Brunauer, Hagenauer, & Leitner, 2013) and land use and transportation systems interactions analysis (Casello & Fard, 2017; Zhang, Song, van Nes, He, & Yin, 2019; Zhong et al., 2015). As we approached our research questions, we reviewed those alternative spatial clustering methods that could potentially provide a baseline to assess our proposed model. We further implemented the graph-based partitioning method and compared its outcomes with the results from our dynamic zone creation model using the comparable input parameter (presented in detail in Sections 3.5.3 and 4.2.3).

2.7 Literature Review Summary

This chapter reviewed the literature on different aspects of land use and transportation system interactions (LUTI). To date extensive research has been conducted to describe and understand the major components of these interactions manifested through location decisions and travel behavior. Given the complex and intertwined nature of the LUTI and its elements, a multi-perspective approach was used to review the growing body of related research from historical, theoretical, empirical and policy perspectives. This work was then complemented with an overview of the quantitative methods that have been frequently used in LUTI research, and planning and policy development.

We started with a historical overview of the contemporary evolution of metropolitan regions in North America to reveal the roots of extensive suburbanization unfolded after the Second World War period. Our work portrayed the role of transportation infrastructure in the emergence of metropolitan polycentricity, and highlighted the impacts of automobile technology on rapid suburban expansions, prevalent central city decline, and dispersed metropolitan form as we observe today. Following the historical overview, seminal theories that explain land use and transportation system interactions were synthesized, which featured the importance of economic theories, from aggregate classical location choice theory to more recent behavioral approaches in shaping our understanding of urban dynamics and physical structure of metropolitan regions. The outcome indicated that although the current theories offer valuable insights on LUTI dynamics, these interactions are more complex and ambiguous to understand purely theoretically.

While the theoretical review revealed the lack of a unifying theory, our review of empirical studies suggested that evidence on the contributing factors, the direction and magnitude of the impacts within LUTI dynamics are mixed as well. This observation was further reinforced, when we explored the implications of these interactions for planning analysis and policy making. Our review presented integrated land use and transportation planning concepts such as Transit Oriented Development (TOD), and identified a spectrum of planning measures to shape desired integrated policy outcomes, which, zoning by far turned out to be the most commonly employed instruments.

This literature review concluded by presenting the role of quantitative methods and data sources in LUTI analysis, and showed the significance of spatial representation and scale as fundamental elements of the analytical methods. More elaborate description of the methods that are used to respond to our research questions is presented in the Methods Section. These reviews collectively highlighted the need for developing more robust quantitative indicators and spatial models to improve LUTI analysis and better inform planning practice and policy. This review also helped to contextualize our investigation of the ongoing urban dynamic in a fast-growing and car-oriented metropolitan area in Waterloo Region, Canada, with respect to the dominant urban development pattern in North America.

Chapter 3: Methods

3.1 Overview

Reviewing the literature and empirical studies on land use and transportation system interactions (LUTI) indicates a trend towards wider adoption of more disaggregate spatial models and analytical approaches in studies. In the LUTI field, where the theory offers only limited insight to draw definite conclusions and the empirical evidence is mixed, disaggregate data sources and analytical methods may hold promise for better understanding of ever-increasing urban dynamics. Recall that the key research questions that will be addressed in this study are:

- What is the appropriate spatial scale to analyze land use and transportation system interactions and their impacts on travel behavior in a mid-size metropolitan region?
- What are the benefits of using dynamically defined zones in travel behavior analysis?
- How do built environment characteristics (e.g., density and diversity) and access to rapid transit influence travelers' behavior across the Waterloo Region?
- How do proximity to LRT corridor and land use diversity levels impact land use development dynamics across the Waterloo Region?

An interpretation of the research questions for the Methods section is described in this overview. New data sources have made individual level travel information and property level land use records accessible for large metropolitan areas that were challenging to obtain before. Although these fine scale data sets offer wider choices of scale for LUTI studies, they simultaneously raise important questions that need to be answered prior to the intended analysis:

- Is it methodologically feasible to take a disaggregate approach for an analysis?
- If aggregation is necessary what is the appropriate spatial scale to analyze land use and transportation system interactions in a certain urban area?

While disaggregate methods are acknowledged in the literature for their accuracy and transparency, they are computationally resource intensive and suffer from scalability issues in larger study areas. It is often more difficult to draw policy sensitive conclusions from the disaggregate analyses, and nonexpert audiences are more likely to get overwhelmed with the level of details and heterogeneity of their results. Hence, even a positive answer to the first question is not sufficient to warrant taking a disaggregate approach.

Aggregate approaches have been state-of-the-practice and used in urban studies for a long time (Páez & Scott, 2004) because these methods have many positive attributes. In the context of LUTI analysis, the results from the aggregate methods are easier to communicate and more straightforward to compare over time and across study locations. Moreover, when the available data sets are heterogeneous in their scales – similar to the case in this study – aggregation helps to bridge between multiple spatial scales (van Delden, van Vliet, Rutledge, & Kirkby, 2011).

Given the motivation to aggregate, we can now return to the second question asked above and attempt to create methods that help determine the appropriate scale of aggregation. The answer to the second question depends on the size and extent of the study area and the urban dynamics taking place within that region. Pragmatically, if an aggregate scale is chosen then the subsequent questions must be answered to determine the aggregation level:

- What are the variables of interest to investigate a certain research or policy question?
- What is the scale at which the required primary data were collected?
- How well do the aggregate analysis zones represent the required variables?
    - Are the zones adequately homogeneous with minimal variation in their attributes?
    - Does the zonal structure efficiently capture spatial heterogeneities across the study area?
- Does the change in scale of analysis significantly impact the results and their policy implications?
- Are there advantages in dynamically defining the analysis zonal structure compared to using pre-defined conventional analysis zones?

This research aims to examine the preceding questions in a case study to investigate specific instances of the general LUTI problem. More specifically, the research presented in this thesis examines and interprets the effects of land use heterogeneity and access to rapid transit on travel behavior in Waterloo Region. The chapter starts with an introduction to the Region of Waterloo as the broad study area and presents the Region's planning context. Then, research data sources and data specifications are discussed. This is followed by a description of the concepts and the methods utilized in this study.

The data-driven framework of this research comprises four major elements and their sub-components as follows:

1. Motivated by the idea that diversity is equally important as density (and other attributes) to the analysis of land use and transportation interactions, this research develops and implements a land use measurement method to quantify diversity as it is perceived within the urban transportation planning, research and policy analysis. This includes the development of an enhanced entropy measurement model that efficiently computes heterogeneity scores for a large number of spatial zones, consisting of any number of land use types. The utility of this model is twofold: to consistently measure land use diversity as a built environment characteristic that is central to the hypothesis testing in this research; and as an integral part of the iterative spatial aggregation model, to evaluate and ensure homogeneity of the derived dynamic analysis zones based on a certain heterogeneity level threshold (cap).
2. Based on the diversity metric created in step 1, the research develops and evaluates a novel spatial method to delineate analysis zonal structure for land use and transportation system interactions studies.
   A set of two distinct spatial regionalization models are implemented and examined. The first is a purpose-developed spatial aggregation model that uses an entropy-based heterogeneity score along with an adjacency measure as the aggregation criteria and generates dynamic zonal structures called Dynamic Activity Cluster Zones (DACZs). A sensitivity analysis is then conducted to evaluate how this model performs against variation in the major input parameters including heterogeneity threshold and the DACZs area size

cap. The second is a graph-based spatial partitioning model that is built upon graph theory. This method uses a spatial optimization approach to identify a predetermined number of homogeneous cluster of zones based on similarity of their enclosed land uses.

3. The DACZ method to create zonal structures developed in the previous step is then applied to a metropolitan case study area to assess its efficacy in generating more homogenous zones in terms of travel behavior. A specific DACZ structure created in step 2 along with a conventional zonal structure (i.e., PLUM zones) are used to classify the study area into a set of built environment categories based on the zones' heterogeneity score paired with their population density. A hierarchical aggregation model is then applied to summarize the travel behavior metrics distributions at a range of spatial scales ranging from disaggregate (i.e., the zone level), to intermediate (i.e., the built environment categories) and aggregate (i.e., the entire study area) levels.

   These hierarchical model outcomes provide the statistical basis for synthesizing and evaluating the likely impacts of heterogonous built environment characteristics on travel behavior patterns for different travel purposes.

4. This research examines the historical trend of development in a metropolitan region to assess whether proximity to high quality transit corridor and the built environment heterogeneity might influence land use development.

   Based on the DACZ zonal structure developed in step 2, a set of temporal analyses are performed at three spatial scales including intermediate (i.e., the built environment categories identified in step 3), semi-aggregate (i.e., the areas within or outside of the Regions' Central Transit Corridor), and aggregate (i.e., the entire study area) scales.

   These analyses aggregate and summarize the building permit records biannually to reveal the areas that might have disproportionately attracted more and or larger land use developments (i.e., higher construction values per issued building permits), potentially due to the benefits of heterogeneous built environment and access to rapid transit.

It is important to note that the terms *land use diversity, land use heterogeneity and entropy score* are used interchangeably throughout this study. Where a specific quantitative measure is expected, often the normalized entropy score is calculated and presented.

3.2 Study Area

3.2.1 Introduction

The Region of Waterloo is a fast-growing metropolitan region in southern Ontario, Canada. The region is an anchor of the High-Tech Toronto-Waterloo Corridor, extending for about 100 kilometers between the Region of Waterloo and the City of Toronto to the east. It is also part of the Greater Golden Horseshoe (GGH), an economic growth region officially designated under the Places to Grow Act (Government of Ontario, 2017) that has been a major contributor to the national economic growth. The GGH has a population of more than eight million and is estimated to attract an additional four and a half million residents and two million new jobs within two decades (MMAH, 2017).

The urbanized area of the region consists of three cities of Waterloo, Kitchener and Cambridge. With rapid demographic dynamics and population growth projected from 535,000 in 2016 to 729,000 by 2031 (RMOW, 2015) and 835,000 at 2041 (Government of Ontario, 2017), the Region has been one of the fastest growing communities not only in the province of Ontario, but also nationally across Canada. This growth is expected to bring more employment and consequently higher levels of travel demand to a predominantly car-oriented metropolitan area. In response, the Regional Municipality of Waterloo has been employing integrated land use and transportation tools – both policy and infrastructure investments – to drive more sustainable urban development and targeted higher levels of public transit and active transportation mode shares.

Historically, the region has been identified as a dispresed metropolitan area with polycentric urban structure. The multifunctional activity centers across the region consist of historical business districs of the Cities of Kitchener and Waterloo, namely Downtown Kitchener and Uptown Waterloo. In addition, at the northern and the southern corners of the region, two regional shopping malls and their surrounding employment and commercial districts have been major destinations for shopping, entertainment and services. The University of Waterloo and Wilfrid Laurier University with approximately 70,000 students, faculty and staff are also among the major acivity destinations within the region that generate considerable number of daily trips. A map of these activity centers along the Region's Central Transit Corridor (CTC) is depicted in Figure 3-1[1].

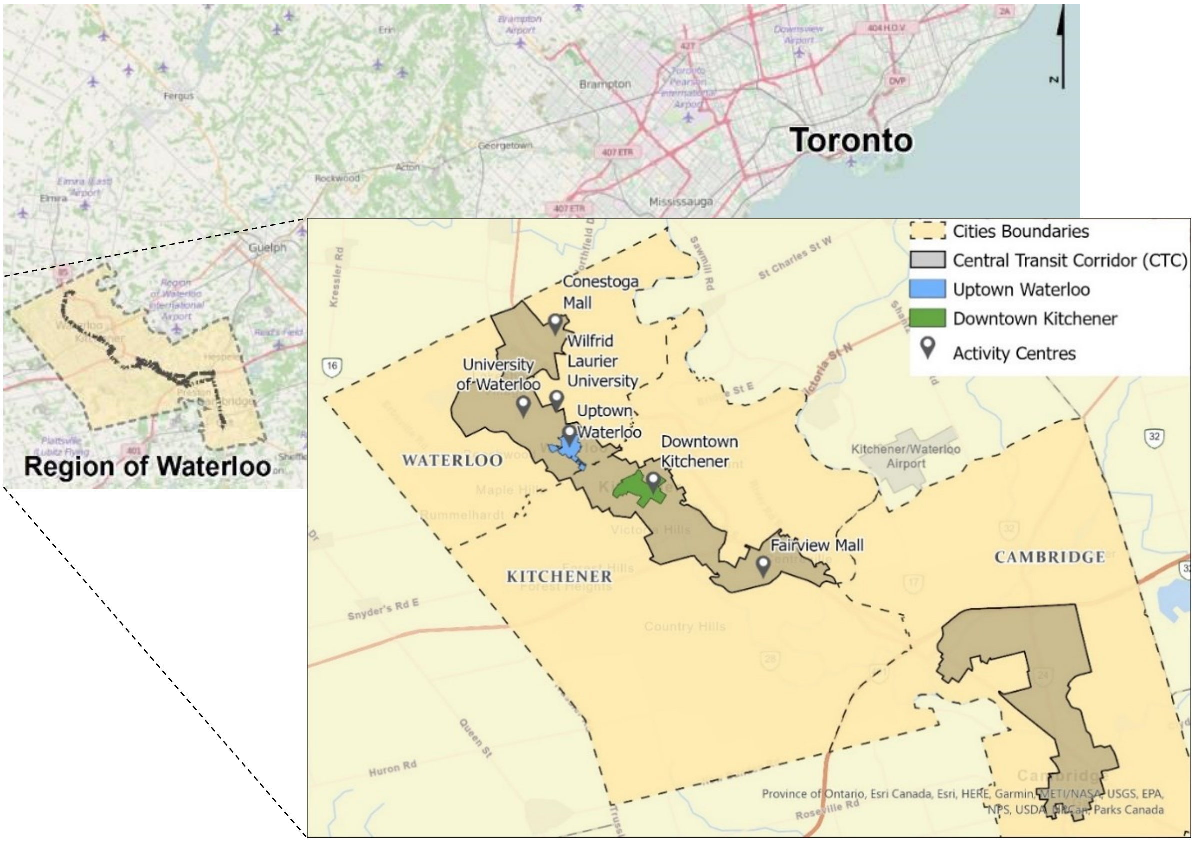


Figure 3-1 - Relative location of Region of Waterloo, the LRT alignment, the region's urban activity centers and the Central Transit Corridor (CTC)

In recent years the Region has invested in multiple major transit system upgrades including a set of higher order express bus transit routes branded as "iXpress", and the construction of 19 kilometers of Light Rail Transit (LRT) system, known as "ION". Operated as part of the Region's overall transit

---

[1] This map depicts an approximate representation of the CTC, delineated based on the boundaries of the Census Dissemination Blocks located within 800 meters or 10-minutes walking distance from the ION LRT stations.

system – Grand River Transit (GRT) – ION was planned as a two-phase project and received funding commitment from the Provincial and Federal governments in 2010. Construction of phase-one LRT between Conestoga Mall in Waterloo, and Fairview Park Mall in Kitchener began in 2014 and ended in 2018. ION passenger services started in June, 2019. This LRT segment connects the region's conventional activity centers and has been proposed to act as a catalyst for the Region's Central Transit Corridor (CTC) intensification along with creation of Transit Oriented Development (TOD) communities.

Phase-two of the project currently features an adapted Bus Rapid Transit (BRT) that extends the LRT system in the north to the Ainslie Street transit terminal in the Cambridge business district. This 17 kilometer segment will eventually be converted to LRT. The other pivotal point of the system is the forthcoming multi-modal hub called Central Station, located midway between Downtown Kitchener and Uptown Waterloo. Central Station integrates ION LRT with local GRT and inter-city buses, as well as the passenger train services including GO Rail (regional to the Greater Toronto Area) and VIA Rail (Canada's national railway system). This station will be a key node that bridges between the growing regional and inter-regional trips, and provides interchanges between all transportation modes (the ION system map is provided in Appendix A).

3.2.2 Planning Context of Region of Waterloo

In the early 2000s, the regional government consolidated its planning policies in a guiding development vision known as the Regional Growth Management Strategy (RGMS). This high level strategy aimed for regional re-urbanization by "Building Vibrant Urban Places", and "Providing Greater Transportation Choice" among its six major goals (RMOW, 2003). The strategy emphasized the role of public transit systems as a major measure to steer the future regional growth, and specifically supported the development of a Light Rail Transit (LRT) system within the Region's Central Transit Corridor (CTC). It also consisted of new land use controls including "a firm countryside line to limit urban sprawl" and encouraged intensification along the CTC to shift the trend of the Region’s outward growth. Following the adoption of this vision, the Region approved a new Regional Official Plan (ROP) in 2015 and an updated its Transportation Master Plan (TMP) called “Moving Forward” in 2018.

The ROP with a 20 year horizon elaborates on the RGMS re-urbanization agenda and introduces a set of policies grounded in principles of Transit Oriented Development (TOD). The highlights from the ROP include supporting mixed-use development with medium to high density, encouraging compact urban form, improving the built environment with street designs that promote multi-modal transportation, concentrating public services and institutional investments within comfortable walking distance from major transit stops, and more importantly directing new developments towards built up areas of the region.

The Moving Forward plan updates the previous TMP from 2010. It reflects on the addition of new transit infrastructure in the region and sets out policies on emerging technologies and evolving travel behavior patterns. In 2011, private car travel accounted for almost 90% of the trips within the region. The new plan puts forward an ambitious target for modal shares: shifting away 12% of the car mode share by 2041 with plans to accommodate those trips by public transit and active transportation modes. The TMP aims to achieve its goals through five strategies:

- Build a Transportation Network that Supports all Modes of Travel
- Promote a Healthy Community (by improving walking and cycling network)
- Develop a Frequent Transit Network
- Enhance Inter-Regional Connections
- Position the Network for New Mobility

In addition to these regional plans, multiple higher order provincial policies and regulations along with lower-tier municipality plans form the planning context in the Waterloo Region.

Planning in the Province of Ontario is structured in a multi-layer scheme corresponding to the Provincial and local levels of government. The Planning Act defines and governs the structure, and necessitates an interdependence of official plans with local levels of government (i.e., single-, upper-, and lower-tier municipalities). Within this hierarchy, Region of Waterloo must conform to provincial policies and objectives and, as an upper-tier municipality, potentially directs the lower-tier municipal plans.

Pertaining to this study, there are a number of plans and policies that provide directions on land use and transportation development in the region. The two major regulations with direct implications for the Region are the Provincial Policy Statement (PPS) (2014) under the Planning Act (Government of Ontario, 1990), and the Growth Plan for the Greater Golden Horseshoe (2017) under the Places to Grow Act (Government of Ontario, 2017).

The PPS lays out the primary policy foundation for land use planning and development across the province. The Region of Waterloo land use regulations reflected in the ROP must conform to the PPS "Building Strong and Healthy Communities" policies. These policies portray high level goals on economic development, land use management, housing, employment, infrastructure, energy conservation and climate change (Government of Ontario, 2014).

The Growth Plan, on the other hand, identifies more specific objectives for the GGH region. It provides guidance on the intended land use patterns and urban forms, prioritizing transit corridors and station areas. According to this plan, all municipalities within the GGH region must achieve a minimum of 40% intensification within their existing built up areas by 2041. Downtown Kitchener, and Uptown Waterloo are designated as urban growth centers and, as a result, must achieve or go beyond target gross density levels of 200 residents and jobs per hectare. In addition, the transit station areas that are served by ION LRT or bus rapid transit (BRT) are also considered as Major Transit Station Areas (MTSA), and required to maintain 160 residents and jobs per hectare combined.

Although not within the direct mandate of the Greater Toronto and Hamilton Area (GTHA), Waterloo Region also aligns its planning with the 2041 Regional Transportation Plan (RTP) for the Greater Toronto and Hamilton Area (2017). The RTP has identified the importance of the economic connections between the region and the GTHA and thus aims to strengthen this inter-regional engagement. The RTP main initiatives to this end are improving the (regional) GO Train frequencies and upgrading the (intercity) GO Bus services for additional origin-destination pairs. Widening the provincial highways and development of the High-Occupancy Vehicle (HOV) lanes to relieve the frequently congested sections of the Highways 401 are among the other objectives of this plan (RMOW, 2018).

Ultimately, in recent years across Ontario metropolitan regions, there has been increasing interest in development and expansion of light rail transit systems as a reliable transit mode. Waterloo Region's ION LRT, along with the under construction Eglinton Crosstown LRT in Toronto, the Hurontario LRT in Peel Region and the revived plan of Hamilton's LRT project are some of the successful examples that have been able to secure investment from Federal, Provincial and local governments.

As the official planning practices have increasingly recognized the need for more robust quantitative assessments in their procedures, the planning framework for the Region provides a rich environment for this study. For example, following the announcement of the LRT project, the Regional Municipality of Waterloo has initiated a monitoring project to assess the Region's progress towards the goals of the LRT project: higher density of development; increased transportation choice; greater utilization of sustainable transport modes. This research has contributed to the development of the baseline metrics for the monitoring report (Parkin et al., 2016). Moreover, by answering the proposed research questions, this study further addresses the need for improved analysis and modeling methods, and expands our understanding of the regional land use and transportation systems interactions at a very disaggregate level.

3.3 Data

This research relies upon multiple sets of (predominantly) spatial data collected from secondary data sources. These items range from publicly available data available on open portals, to institutionally administered data acquired through access agreements. This study also leverages valuable data sharing collaborations established between University of Waterloo and the Municipal Property Assessment Corporation (MPAC – the corporation tasked by the Province of Ontario to estimate and maintain a database of property values for all parcels in the Province) as well as the University of Toronto (where a research group administers travel surveys and forecasting for the Greater Golden Horseshoe). Development of these collaborations was essential for accessing the fine-scale land use and travel behavior data.

The inventory of collected research inputs include demographics, land uses, transportation networks, transit routes and travel behavior information. Integrating these thematically diverse data sets from multiple data sources involved issues of data type inconsistencies and conflicting geographic projection. To ensure consistency and accuracy of analyses, appropriate data preprocessing methods were utilized to address and resolve these errors.

Data processing, statistical analysis and visualizations generated for this study to a large extent were conducted using open source software and platforms including R programming language along with the Tidyverse data processing library. Network related data processing and travel time computations were done using ESRI ArcGIS platform. The following section elaborates on the different data sets and their specifications utilized in this study.

3.3.1 Property Level Land Use and Parcel Fabric Data[2]

Recall that one of the goals of this research is to develop and evaluate the value of alternative zonal structures that explicitly consider land use heterogeneity. In order to develop these very disaggregate

---

[2] Due to confidentiality terms of the research data access agreement, only aggregate level land use information is presented in this thesis.

estimates it was required to incorporate parcel level land use data into the proposed model. The property level land use data were available to this study for year 2019. This data set was provided by the Municipal Property Assessment Corporation (MPAC) covering an approximately 3km buffer around the LRT alignment within the cities of Waterloo and Kitchener. The MPAC land use dataset contains detailed land use classes, categorized for taxation and property assessment purposes.

MPAC's land use classifications are applied as per Provincial regulation (Ontario. Reg. 282/98) for the purpose of assessment and taxation (Government of Ontario, 2020). There are some peculiarities in these Regulations which can lead to differences between how properties end up being classified and general understanding of their land uses. The simplest example is vacant land, which defaults to a commercial classification unless "principally zoned" (an undefined term) for another use. Despite these irregularities, the dataset proves to be exceptionally useful to uncover diverse activities that take place within the urbanized area of the region.

This dataset consists of 150 land use classes within the cities of Waterloo and Kitchener. Using the unique Assessment Roll Numbers (ARN), land use records were integrated with the regional parcel fabric for the year 2018. The parcel fabric layer was made available for this research by the Geospatial Centre at University of Waterloo Library; the data were originally obtained from the Province of Ontario Land Registry Information System (POLARIS). POLARIS is administered by Teranet Company – the exclusive provider of land registry data in Province of Ontario.

The extensive classification system administered by MPAC consists of more than 300 land use classes and offers great detail on characteristics of a property (i.e., land parcel) as they are important for the financial institutions and municipalities for their property assessment and taxation operations. At the aggregate level, each property is categorized into one of the eight broad general categories of: Vacant, Farm, Residential, Commercial, Industrial, Institutional, Special & exempt and Government. These major eight "series" – as designated by MPAC – reflect the nature of activities in a land parcel given its economic function. In the disaggregate level, however, taking into account more specific business types, building structures and tenure types, each property is classified under one of the potential 300 predetermined classes. For instance, two similar properties that are classified as Residential at the aggregate level, are further specified according to their different ownership structures with complex designations: “Freehold townhouse/row house – more than two units in a row with separate ownership – property code 309”; and “Row housing, with three to six units under single ownership – property code 350” (MPAC, 2020). While this level of disaggregation may be useful for MPAC, for this research purpose over-segmentation of land uses would likely introduce unwanted complexities and biases into homogeneity measurement. In this instance, we assume that these similar residential classifications are likely to produce similar travel behaviors (such as trip generation) and, as such, can be aggregated without loss of fidelity.

More generally, to measure land use heterogeneity from a transportation planning perspective, the excessively complex MPAC classification system was remapped by assigning the observed land uses in the region into a more manageable set of classes. The process that was undertaken was to compare potentially similar MPAC classifications for possible aggregation based on the similarity of their travel behavior properties. To do so, MPAC classifications were compared based on their trip generation and attraction characteristics according to the ITE Trip Generation Manual (Institute of Transportation Engineers (ITE), 2017). Since 1960s, the ITE Trip Generation Manual has been the industry-standard handbook for estimating trip generation potentials for different land developments across North America. For the purpose of this research, a lookup table was developed to link each of the MPAC land use classes identified within the study area to a comparable reclassified land use class provided by the ITE Manual. The objective of this process was to reduce redundancies in a large set of land use categories without losing essential details needed for land use heterogeneity measurement.

This reclassification was inspired by the proposal for adjustment of the ITE Manual and a series of subsequent empirical studies (Clifton, Currans, & Muhs, 2015; Currans & Clifton, 2015, 2018). These studies suggest to consider a combination of characteristics for potential aggregation of land use categories for transportation analysis. The highlighted characteristics include temporal similarity of the activity related travels, dependence of trip-maker associated with the activity, size of the development, product base and institutional format of the activity, and ultimately social and recreational dimensions of activity located in a land parcel.

Historically, the provided methods for trip estimation rates in this Manual tend to over emphasize automobile trips and underrepresent non-motorized travel (Singleton and Clifton 2013). This issue has been recognized by ITE and in recent years efforts have been made to utilize expanded travel surveys to better account for multi-modal and non-automobile trips. As a result, an updated Supplement to the ITE Trip Generation Manual was released in 2020 with significantly expanded multi-modal information incorporated. In this study, the 10th edition of the Manual was used that provides information for 176 specific land use classes under 10 major categories.

As noted earlier, there are eight general MPAC property codes. Of them, the 400 series or the commercial properties had the highest number of observed classes, with 47 observed instances in the region. The original classes included various commercial, office, retail and entertainment uses; these MPAC classifications were mapped to 17 new reclassified categories. Similarly, the 300 series (residential properties) had 34 subclasses within the data set, which were reclassified into six categories from the ITE Manual including (1) Single-Family Detached and Low-Rise Multifamily Housing (combined), (2) Multifamily Mid and High-Rise Housing (combined), (3) Mid and High-Rise Residential with Commercial (combined), (4) Off-Campus Student Apartment, (5) Congregate Care Facility, (6) Recreational Homes. For industrial properties, 19 classes were reduced to just three new classes. The 800 series or governmental properties had a minimal presence within the study area with only four identified classes that were combined into one category. One additional class was preserved to address the missing land use records and inapplicable conversions.

The results of this aggregation include less complexity in land uses that will facilitate more straightforward zonal constructs while not generating significant loss in the representation of these land uses for transportation analysis.

3.3.2 Travel Behavior Data

One research objective of this study is to determine if zones created through the proposed method – identifying and merging areas with similar densities and heterogeneity of land uses – produces distinctly different transportation behaviors than zones established using existing methods. To this end, a data set containing disaggregate travel behavior was merged with the disparate spatial representations and metrics of travel propensities were calculated.

The Transportation Tomorrow Survey (TTS) is a household travel survey that is conducted by the University of Toronto's Data Management Group (DMG). The TTS covers selected municipalities with significant economic interconnections in Southern Ontario, including the Region of Waterloo and the Greater Toronto and Hamilton Area (GTHA). Although TTS collects detailed geocoded trip origins and destinations data, those data were not accessible due to confidentiality policies in place to ensure respondents' privacy. The DMG normally provides the aggregated travel behavior data at a scale known as Traffic Zone System (TTS TZS). These zones are conventionally defined to provide computational efficiency for transportation modeling and often represent relatively large districts within urban areas. The current TZS zones were last updated in 2006; the study area of this research is represented in the TZS as 192 zones.

However, this study aimed to analyze the travel behavior distributions at multiple spatial scales that were different from the conventional TAZs. Hence, it was required to process and aggregate the detailed geocoded trip origins and destinations data from the original survey. DMG helped with this data aggregation step as well as demographic data assignment based on the provided zonal structures.

The first incorporated structure was a set of modified TAZs that is used by the Region in their Population and Land Use Model (PLUM). The second structure was the novel Dynamic Activity Cluster Zones (DACZs) generated through homogeneity-based aggregation process.

For each zonal structure, a set of travel flow matrices were derived that included data columns corresponding to the trips origin and destination (O-D) unique zone IDs, total number of trips that share the same attributes, travel mode and purpose of trips.

3.3.3 Transportation Network and Transit Routes

To extend our travel behavior analysis further than the mode share comparisons, we proposed to conduct additional assessments on the distribution of the travel times for the trips within the region, given their travel mode. As we are interested to explore the built environment impacts on trips characteristics, we focused only on those trips that have their origins identified within the dynamic zones, for which are able to quantify the levels of heterogeneity and population density.

Despite the great insight the TTS data provide on travel behavior characteristics, its aggregated nature precludes the inclusion of explicit information on individual trips' travel time. So, in order to perform further travel behavior pattern analysis, it was necessary to explicitly estimate the travel time and distances based on the O-D pairs spatial information.

Travel time and distance estimation often relies on computation of a network-based shortest path using mode-specific routing algorithms. In recent years, an increasing number of out-of-the-box (online) services have started to provide such estimates for trip planning and research purposes. Google's Distance Matrix API and ESRI StreetMap Premium are among the well-known examples of such services. A recent review of these options indicates that given the different technologies and technical considerations, the results of computed travel times using different toolsets often differ considerably for the same O-D pair (Higgins, 2019). Moreover, due to the very large number of O-D pairs and multiple number of modes that pertain to this study, using online services could be a costly and less efficient choice compared to the development of an in-house transportation network model and travel query system. Hence, using the Network Analyst module from the ESRI ArcGIS platform, a digital network model was implemented and a flexible travel time and distance estimation toolbox was developed to perform distance and duration estimates for the transportation modes present in this study.

Basically, an appropriate transportation network model in a GIS format requires spatial features that resemble the geometry of the network linear elements and their connection points, along with the essential attributes such as traffic direction and the likely prohibitions associated with the transportation mode that is being modeled (e.g., prohibition of cycling on highways).

In this study, the basic geometric transportation network data and the required tabular attributes for developing walking, cycling and auto modes network were obtained from Ontario Open Data Portal, for the year 2019. Taking into account the potential data errors, the network spatial layer was first topologically “cleaned up” and the linear features (links) that represent traversable edges and their logical connections were converted into a Network Dataset – an industry standard format to store and retrieve transportation network data. In the absence of explicit cycling infrastructure and sidewalk network data, all the road segments except the designated highways were incorporated into the walking and cycling network.

Then, to increase the accuracy of the shortest path estimation for auto mode, the available tabular attributes including traffic direction, traffic restrictions, speed limit, and the pavement type were joined to the geometric elements of road segments and added to the model routing algorithm.

As the completed Network Dataset consists of hundreds of thousands road segments and the number of the potential O-D pair queries could be close to a million instances, it was necessary to improve the routing computational efficiency by preventing potential redundant path searches. This was done through spatial indexing of the road segments according to their functional hierarchy. Given the complexity of the road networks in a metropolitan area, a combination of the official road designations, speed limit and number of lanes was used to classify each road segments and then the spatial index of roads hierarchy implemented to the network model for Waterloo Region.

Having the first transportation network implemented with the capability of modeling walking routes, it became possible to develop an integrated multi-modal network supporting the combined walking and public transit modes queries (Figure 3-2). In this multi-modal network, route choices are determined by the ability of pedestrians to reach transit stops considering the availability of transit service within a time window. The transit routes, service schedules and frequencies for both the GRT bus network and the ION LRT were extracted from the General Transit Feed Specification (GTFS) data. This data set was provided by the Region of Waterloo Open Data Portal and reflected the latest, redesigned GRT network structure and the operating LRT service as of Fall 2019.

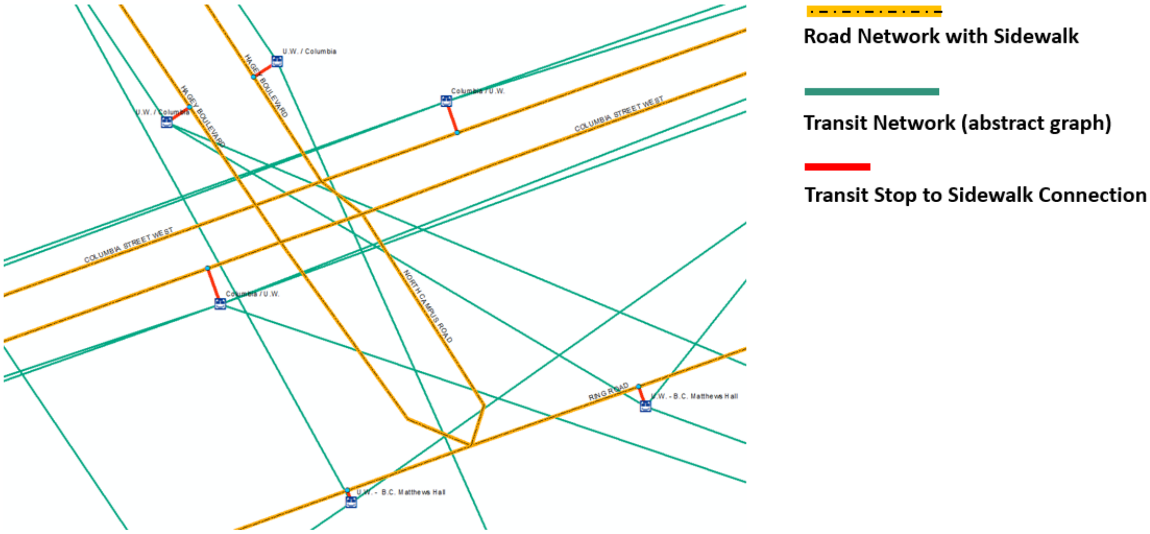
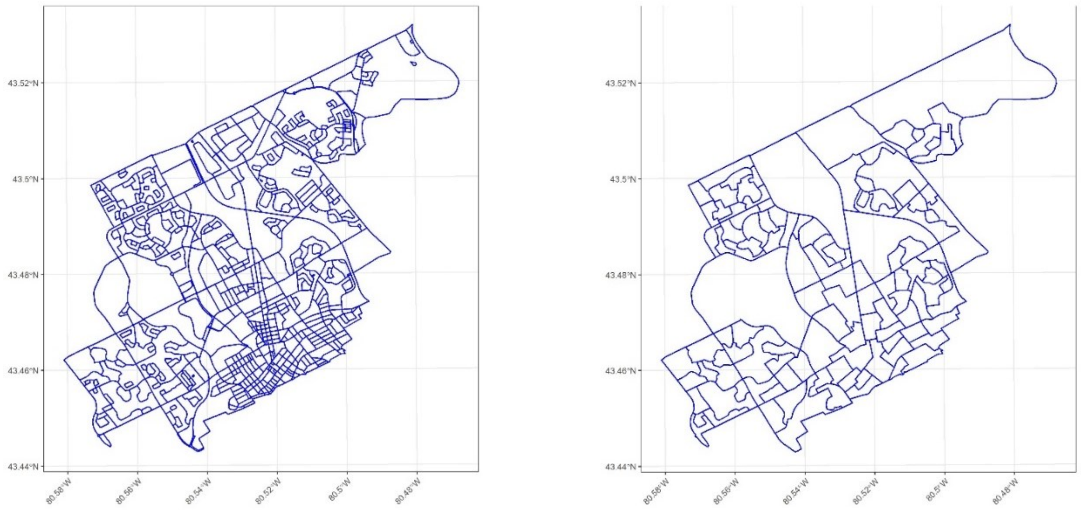


Figure 3-2 - Topological model of the multi-modal network for walking and transit

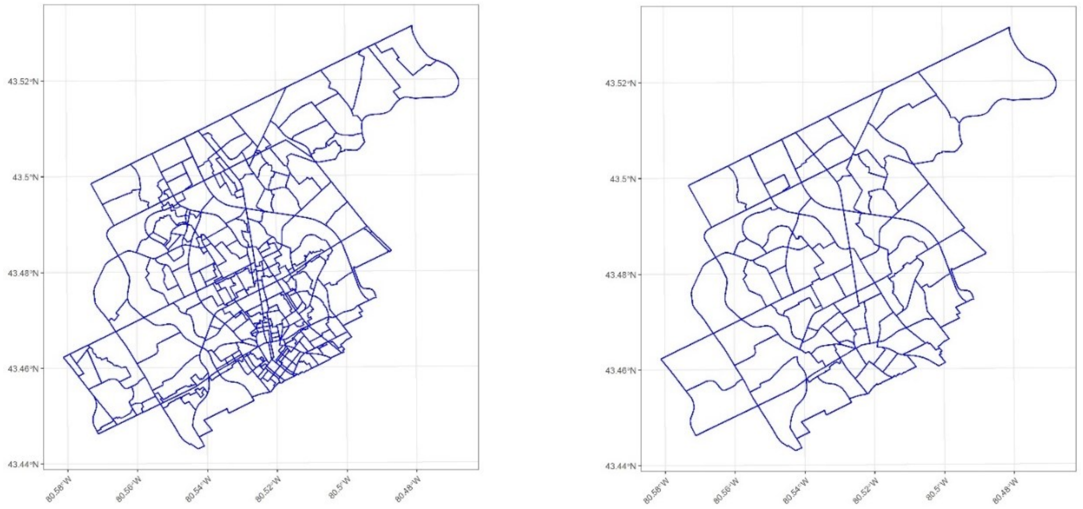
3.4 Analysis Zones Delineation for Land use and Transportation Studies

The pre-set structure of zones that are commonly used in transportation studies is referred to as Traffic (or Transportation) Analysis Zones (TAZs). TAZs often stand at the middle of the spatial continuum, larger than city blocks and smaller than census tracts depending on the scale of analysis (Figure 3-3).



(a) Census Dissemination Blocks
(559 zones)

(b) Census Dissemination Areas
(117 zones)



(c) Population and Land use Model Zones
(291 zones)

(d) Conventional TTS TAZs
(84 zones)

Figure 3-3 - Conventional Zonal Systems Used in Urban Studies (City of Waterloo Study Area)

While conventional TAZs may vary in their sizes and shapes, an effective zonal structure will result in land areas (zones) with a set of homogeneous underlying activities that for this analysis includes land use activities that generate and attract similar travel behavior patterns. While homogeneity may be accomplished by specifying very small areas, a large number of zones becomes computationally difficult. To balance the need for homogeneity and tractable numbers of zones, pre-defined TAZs are primarily defined based on a sort of density measure, where population and employment densities are most commonly used (You et al., 1997). Density measures require minimal data, are easy to compute and provide a relatively simple and intuitive basis for drawing zonal boundaries. However, this simplicity comes at the cost of effectiveness and applicability (Chang, Khatib, & Ou, 2002).

For instance, a high density residential area consisting of uniform land uses is more likely to generate and attract commuting trips that take place during the peak hours by auto. In contrast, a mixed use area with similar density levels tends to generate lower volume of outgoing trips. Since these diverse land uses provide a wider range of services and trip destinations locally, households are able to satisfy their daily activities through shorter trips, often with a greater proportion of trips made by walking or cycling. Density-based zonal definitions tend to overlook these differences and, as a result, produce similar size zones for both presented cases, despite different transportation activities that are likely to take place.

Although, density and diversity may have some characteristics in common, they impact travel behavior and the likelihood of travel by multiple modes differently. In the earlier example, within the mixed use area, the diverse land uses might produce a set of pedestrian tours over the mid-day period (off-peak hours) through which several households' activities could be accomplished. This wouldn't be possible in a homogeneous area of similar density.

In addition to the limitations specific to density-based methods, pre-defined zones and ad-hoc delineation approaches in general are subject to a set of issues arising from the inherent limitations of aggregation. Spatial aggregation as a data reduction technique inevitably involves information loss, increases uncertainty and introduces biases into analyses. The spatial scale effect issue and modifiable areal unit problem (MAUP) are among the well-known instances originating from these biases (Pereira, Banister, Schwanen, & Wessel, 2019).

Hence, these known issues need to be explicitly identified and treated to arrive at robust policy conclusions. This research explores multiple spatial aggregation and clustering strategies and particularly develops a method to effectively produce and evaluate zonal structures. This method facilitates understanding of inherent spatial scale issues and provides appropriate tool to address them in a scalable manner.

3.5 Dynamic Activity Cluster Zone (DACZs) Model Implementation

The work presented in this section addresses the first research question. This study aims to integrate a novel heterogeneity measurement and a spatial aggregation method to create new zonal boundaries over a large and heterogeneous urban area. This method helps differentiate areas with high concentration of diverse land uses from homogeneous areas with minimal variation in their activities. It is built upon a quantitative perspective of land use diversity as the degree to which various types of activities are closely located within a landscape. Land use diversity is being recognized as a significant zonal characteristic that influences travel patterns at different spatial scales (Gehrke & Clifton, 2014). High levels of diversity provide multiple activity destinations within an easy access area which increases accessibility, makes non-motorized transportation feasible, and supports efficient public transit performance.

3.5.1 Iterative Homogeneity-based Spatial Aggregation Approach[3]

Along with the novel heterogeneity computation, this research introduces a new zonal structure that is not pre-defined nor static, but rather created based on the heterogeneity of the parcels encountered over dynamic boundary formation. To help explain the process, consider the following examples.

---

[3] This section is based on the following peer-reviewed conference publication:
Fard, P., Casello, J. M., & Xu, X. M. (2021). An Approach to Measure Land Use Heterogeneity to Identify Homogeneous Urban Activity-Clusters. Transportation Research Board 100th Annual Meeting. January, 2021.

Suppose we wish to understand how local factors – like transportation infrastructure – influence transportation patterns at a neighborhood level. To do this, we might start with a single parcel in an urban core and review its attributes. We could then begin to gather the same data for adjacent properties and nearby properties. If we were to inventory these attributes, we will quickly identify a very high number of observed parcel types and purposes. This heterogeneity suggests that in order to model what's happening in this area – in terms of travel behavior, or accessibility – the analysis should take place over very small geographic areas, because increasing the zonal size would introduce too much complexity to represent behavior meaningfully in any estimates to be done.

Now consider that the starting point of analysis is at the center of a large, homogeneous suburban subdivision. We will likely have to evaluate a very large number of parcels at a significant distance from the starting point before reaching some threshold level of heterogeneity. This suggests that any modeling done on this area can involve a much larger spatial area without loss of representativeness. In the first example, a parcel say 200 meters from the starting point is likely to not be as relevant (again in terms of travel behavior) to the starting point, because of the heterogeneity of the land uses between them. In the second example, a single family detached house 1 km from the starting point may be completely relevant to the starting point, because of the homogeneity between them.

These two examples can be broadened to conceptualize the definition of analytically generated zonal structures called Dynamic Activity Cluster Zones (DACZs). Each activity cluster zone that is dynamically defined represents a set of land parcels that together satisfy a certain threshold of homogeneity. Once the aggregation exceeds an empirically derived threshold, the Activity Cluster gets bounded. Depending on the transportation mode that is being analyzed, DACZs are forged through spatial aggregation of land parcels located along a transportation network segment that serves these associated parcels. This approach helps to better reflect the impacts of parcel level heterogeneity and distinguish mode specific transportation accessibility in urban system analysis.

3.5.1.1 Enhanced Entropy Index as a Measure of Land Use Heterogeneity

Among the methods to evaluate heterogeneity, Entropy measurement provides a robust approach to capture diversity and to compare heterogeneity in a quantitative manner. Conceptually, an Entropy index can be perceived as a probabilistic measure that reflects the likelihood of finding multiple, different activities (in this case land use classes) within a given boundary of space. Entropy is an aggregate measure, meaning it reflects compound characteristics of a set of entities (land parcels) all together.

Entropy's mathematical representation is shown in **(Equation 1)**. The entropy score $(Entropy_i)$ is a unitless measure ranging from 0 to 1. In a land use diversity analysis, scores closer to 0 represent homogeneity and the value 1 indicates maximum heterogeneity of the zone that is being analyzed. In this definition, homogeneity means 1 minus the heterogeneity.

$$Entropy_i = \frac{-\sum_i \left(Q_{lu,i} \times \ln(Q_{lu,i})\right)}{\ln(n)} \quad while \quad Q_{lu,i} = \frac{S_{lu,i}}{S_i}$$

where:

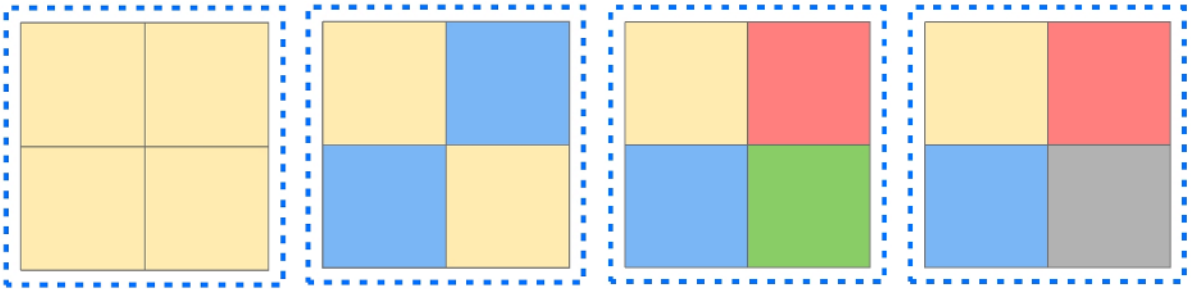
(n) refers to the number of land use classes observed within the zone (i),
(lu) iterates through observed land use classes within that zone,
(S_i) denotes total area of the zone (i),
$(S_{lu,i})$ represents the area of the specific land use class (lu) within the same zone,
(Q) is described as the proportional area measure of diversity in Entropy model.

In the proposed method, in addition to the conventional Entropy model parameters, a land use taxonomy reclassification scheme (lu') and a new coefficient (w) were introduced to better capture impacts of different land use types on diversity. Also, by employing a normalization strategy, the landscape-wide diversity levels are reflected on the local entropy score. The Enhanced Entropy model is formulated as follows (**Equation 2**):

$$Enhanced\ Entropy_i = \frac{-\sum_i \left(w_{lu'} \,.\, Q_{lu',i} \times \ln(Q_{lu',i})\right)}{\ln(\max(n))} \quad while \quad Q_{lu',i} = \frac{S_{lu',i}}{S_i}$$

To demonstrate how the Enhanced Entropy formulation in equation 2 works, a simple landscape is presented consisting of six different possible land use types (denoted by different colors). Within each zone (depicted as blue dashed squares), the land parcels are arranged as four evenly shaped cells. As Figure 3-4a. shows, four uniform cells – all parcels are suburban residential units – result in an Entropy score of zero, which is an indication of absolute homogeneity.

In Figure 3-4b., a relatively homogeneous zone including a balanced mix of two different land use types (e.g., residential and commercial) is illustrated. For this example, the model estimates the Entropy score as 0.5 because there are two sets of two homogeneous parcels, each occupying half the land area. Mathematically, in the Enhanced Entropy model denominator, a normalization is done that replaces the total number of land use classes observed within each zone (n) by the maximum number of land use classes observed within any given zone in the study area $(max(n))$. This change produces a significant difference from the conventional Entropy score, which would be equal to 1 in this case without the normalization applied.


(a) Entropy score = 0
maximum homogeneity

(b) Entropy score = 0.5
a relatively homogeneous zone consists of balanced land use mix

(c) Entropy score = 1
maximum heterogeneity

(d) Entropy score = 0.75
maximum heterogeneity decreases by presence of an activity averse land use

Figure 3-4 - Demonstration of computed Enhanced Entropy for a simple landscape

The third example (Figure 3-4c.) illustrates a zone in which every parcel represents a different type of land use. In this case both the conventional and Enhanced Entropy models produce scores equal to 1. This can be interpreted as the maximum heterogeneity within a zone. In the last example (Figure 3-4d.), there is a mixed use zone that consists of a land use that from a planning perspective adversely impacts the heterogeneity. One can imagine this as presence of a large parking lot within an urban core. The conventional Entropy model tends to overlook the differences of land use configuration presented in the case (c) and (d), and estimates Entropy score as 1. In contrast, this study's model differentiates the two cases through application of the coefficient ($w_{lu'}$). It returns the reduced value of 0.75 as the Entropy score for this zone (Figure 3-4d.).

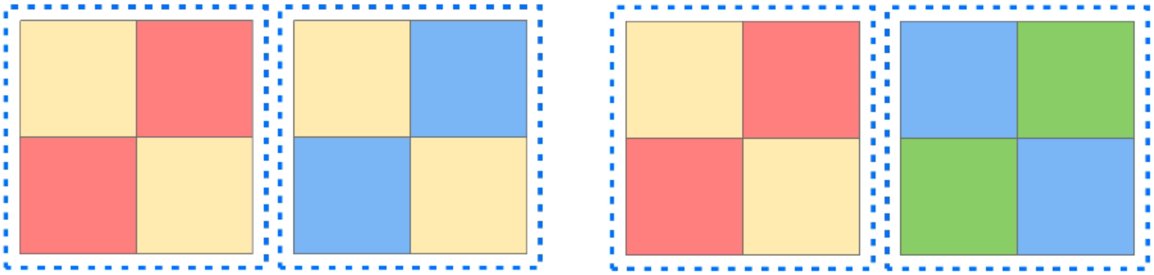
3.5.1.2 Dynamic Activity Cluster Zone (DACZs) Creation

The previous example demonstrates entropy calculations based on a total area that is delineated into four, equally-sized zones. The goal of creating a DACZ system is to determine an aggregation strategy that will combine subunits of land use together, thereby increasing the size of these areas, while maintaining a diversity (or entropy) measure that is lower than a threshold. The process is repeated across a region such that ultimately all possible aggregations that can be made are made, and a finite number of DACZs are created, all with entropy scores below the desired threshold.

This Enhanced Entropy computation model was implemented as a function within R programming language. This function further is being used at multiple steps through the spatial aggregation method. This Entropy computation module is a crucial element of dynamic zonal system creation method, especially where potential aggregation of a base zone with neighboring zones takes place. The following examples demonstrate how the diversity-based spatial aggregation is performed. Again a simple landscape of neighboring zones is presented to evaluate potential homogeneity-based aggregation.

Beginning with an analysis of Figure 3-5a., there are two adjacent zones each made up of four parcels. The left zone is made up of two yellow parcels and two red parcels. The right zone is made up of two yellow and two blue parcels. Both the left and right zones have an Entropy index of 0.5. Aggregating these two zones will produce an output of four yellow, two red and two blue parcels all of equal area. The entropy score for this merged zone is 0.75. In Figure 3-5b., there is a similar, but not equivalent set of parcels. The left zone is again comprised of two yellow and two red parcels. But the left zone has two green and two blue parcels. While each zone has the same initial entropy score as Figure 3-5a., the aggregation of these zones generates four sets of two parcels, each with different functions. As a result, the combined Entropy index is 1. Hence, when trying to aggregate zones to create clusters one cannot just aggregate zones with similar Entropy scores. **Only zones that have similar Entropies but also assembled from similar disaggregate land uses that make up their Entropy can be aggregated**. This means each and every diversity-based potential aggregation requires to be evaluated against a new inventory of land uses within the potential cluster of combined zones. This task becomes challenging from methodological perspective, especially when processing urban landscapes consisting thousands of base zones and much larger number of adjacent neighbors.

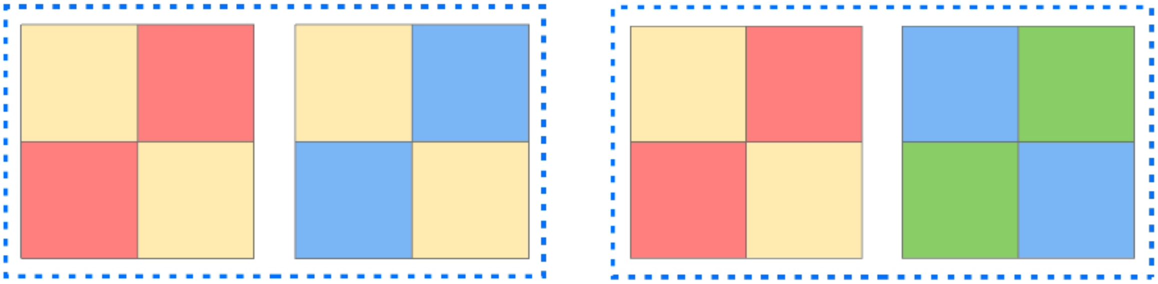
The steps required to perform such evaluation in a modular spatial model were implemented in an automated GIS model, presented in the following section.



(a) Entropy score per each base zone = 0.5

(b) Entropy score per each base zone = 0.5

These are relatively homogeneous individual zones consist of balanced mixed of two land use types.



(c) Aggregate entropy score = 0.75
Aggregation of the two base zones create a cluster that would still be relatively homogeneous.

(d) Aggregate entropy score = 1
Aggregation of the two base zones create too much heterogeneity that exceeds the acceptable threshold.

Figure 3-5 - Demonstration of computed Enhanced Entropy for different aggregation scenarios

3.5.1.3 Implementation of the DACZ Creation Model Algorithm

Dynamic zone creation is an iterative process where each and every base zone along with its adjacent zones are evaluated against a homogeneity criterion. The homogeneity threshold is defined through an exploratory sensitivity analysis that ensures the intended level of diversity within derived Activity Clusters.

With any kind of zonal creation one has to have some sort of base level of spatial disaggregation to start with. This could be all the way down to the parcel level, however, depending on the landscape configuration, distribution of parcel size and the road network patterns, reliance on the parcel level spatial aggregation is prone to forming irregular shape clusters and causing spatial gaps. This study proposes an intermediate spatial unit called block segment to initiate the spatial aggregation process. Block segments are defined based on Voronoi polygons (Boots, Okabe, & Sugihara, 1992) structure and help to form and maintain coherent set of parcels. An example of such a set of polygons is shown in Figure 3-6.

The Dynamic Activity Cluster Zone (DACZ) creation method performs a multi criteria assessment over block segments and incrementally identifies Activity Clusters using following heuristic:

I. Generating building blocks of the Dynamic Activity Cluster Zones (DACZs)
    1) Generate a hypothetical spatial grid over the study area to construct the base units for spatial aggregation. The potential alternative approaches here are:
        a. Regular spatial tessellation (e.g., fishnet or hexagon grid);
        b. Voronoi tessellation based on the location of the transportation network nodes. In this method, for every transportation network node (i.e., road junction) an encapsulating polygon is created where the edges of this polygon are halfway distance between nearest nodes. Conceptually, these are shown as the dashed lines in Figure 3-6. An actual example of the generation of intersection nodes and the creation of polygons is shown in Figure 3-7a. and b. respectively.

This study uses Voronoi tessellation approach since it better fits the complexity of urban parcel fabric.

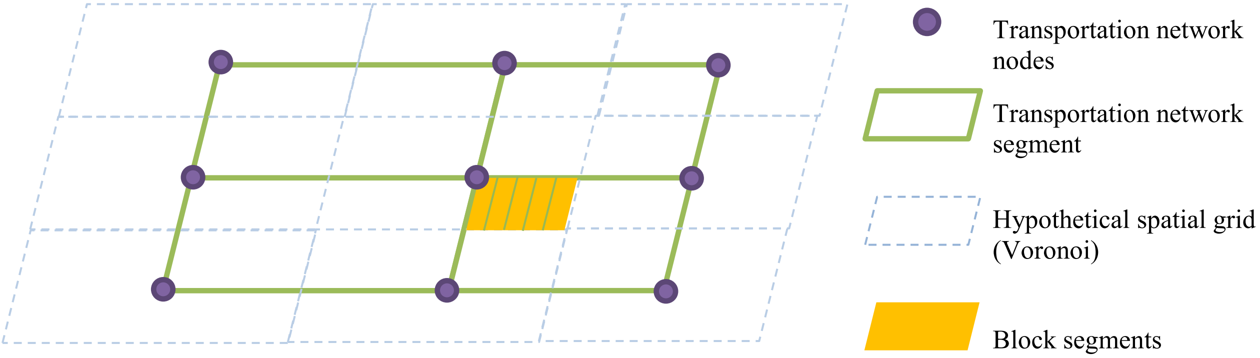


Figure 3-6 - Conceptual illustration of the hypothetical spatial grid and its association with the transportation network

2) Overlay the hypothetical spatial grid with underlying city blocks. Divide the city blocks into block segments containing several adjacent parcels (Figure 3-7c.).
3) Measure heterogeneity for each block segment created in step 2 using the enhanced entropy index.
4) Sort and rank the generated block segments based on their score of enhanced entropy index from step 3 (Figure 3-7d.).

II. Aggregating building blocks and forming Dynamic Activity Cluster Zones (DACZs)

5) Identify the block segment with the lowest entropy score as the core element for the DACZs creation.
6) Find all adjacent block segments where adjacency is defined as:
    a. sharing a border of any length;
    b. not spatially separated from the initial block segment by a physical or natural feature (e.g., river, cliff, highway);
    c. not previously assigned to another DACZs.
7) Add an adjacent block segment to the DACZs if:
    a. the land use diversity as measured by the combined entropy index remains lower than a defined threshold;
    b. the size of the DACZs does not exceed a certain threshold as an absolute value or relative to the landscape (e.g., 5 times larger than the largest base zone within study area).
8) Repeat step 7 until all adjacent block segments have been evaluated. The qualified block segments are added to the core unit and form a Dynamic Activity Cluster Zone (DACZs).

9) Return to step 6 to evaluate next degree of adjacency if it has been chosen by the user to do so.
10) Return to step 5 until all the block segments have been assessed.

The result of this heuristic approach is a set of aggregated zones with a composite entropy index that is lower than the user defined threshold. Since this approach generates zones for which the size is defined as a function of homogeneity, the expectation is that these zones will generate more predictable outcomes of future analysis.

(a) Extract roads' intersections

(b) Generate Voronoi polygons based on the extracted intersections

(c) Create block segments by overlaying city blocks geometry and Voronoi polygons

(d) Compute heterogeneity score and rank each and every block segment

Figure 3-7 - Generating building blocks of the Dynamic Activity Cluster Zones (DACZs)

3.5.2 Spatial Aggregation Model Sensitivity Analysis

Following the implementation of the spatial aggregation model, the next step is to evaluate the sensitivity of the model in terms of the number and structure of the output zones as a function of the input parameters. The essential parameters to assess the DACZ creation model are the aggregation homogeneity threshold (evaluated at step 7a), and the DACZs area size cap (processed at step 7b). Moreover, to explore the impact of different base zonal structure on the outcome DACZs, both analytically created Voronoi polygons and block segments (the model default input) were tested as the base layer for the aggregation model.

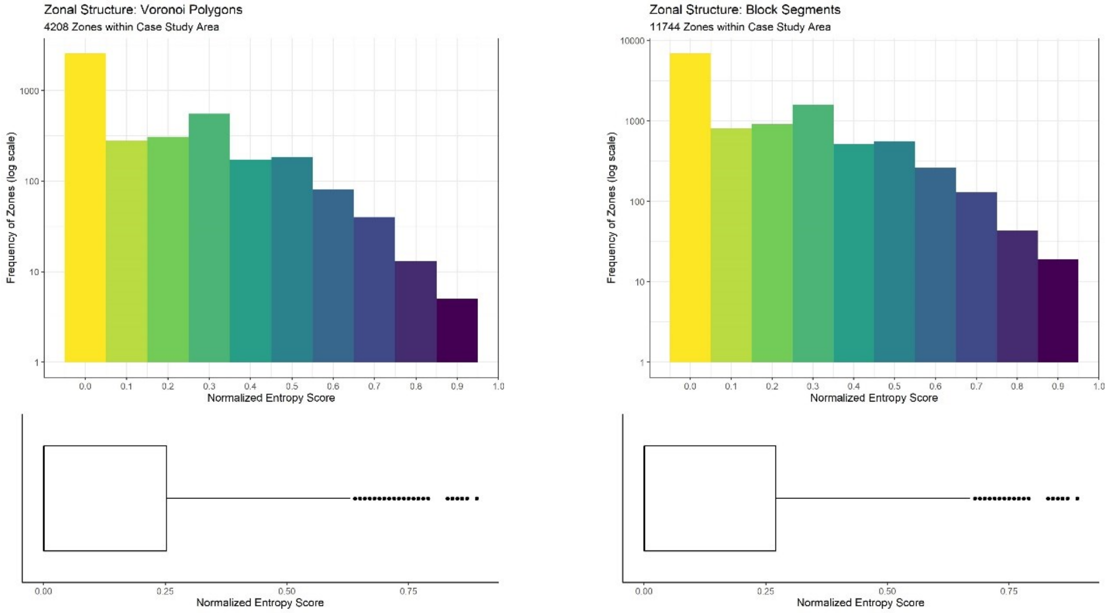
The required sets of parameters' values for this sensitivity analysis were systematically determined, taking into account theoretical context of spatial dependence and a balanced range of parameter variability within the study area. The following section describes the value selection for sensitivity analysis parameters and the aggregate results. The detailed outcomes and interpretations of the sensitivity analysis will be presented in the Results Chapter, under Section 4.2.2.

3.5.2.1 Heterogeneity Threshold

Recall that the land use heterogeneity level is the cornerstone of DACZ model. Hence, a wide range of variation was considered to examine this parameter based on the observed statistical distribution of heterogeneity levels. While the heterogeneity score itself is a continuous variable, it was primarily sequenced by 10 percentiles ranges and a set of six heterogeneity thresholds were derived at 45, 55, 65, 75, 85 and 95 percentiles of the normalized entropy score distribution for both input zonal structures (Table 3-1).

	Average	*Standard deviation*	*Max*	*45th percentile*	*55th percentile*	*65th percentile*	*75th percentile*	*85th percentile*	*95th percentile*
Voronoi	0.12	0.18	0.89	0	0.01	0.12	0.252	0.33	0.51
Block segments	0.13	0.19	0.89	0	0	0.14	0.27	0.33	0.52

Table 3-1 - Summary statistics of the observed heterogeneity scores for the two disaggregate zonal structures at regional scale



(a) Distribution of heterogeneity scores for voronoi polygons (in log scale)

(b) Distribution of heterogeneity scores for block segments (in log scale)

Figure 3-8 - Statistical distributions of the observed heterogeneity scores for the two disaggregate zonal structures at regional scale

Since the block segments are derived from the Voronoi polygons, their overall entropy score distribution patterns are very similar (Figure 3-8). However, as the size and shape of the zones vary for each structure (Appendix B), so that the land uses enclosed within them and their exact entropy score distributions vary too. These subtle differences are depicted in Table 3-1. This table presents the summary statistics of the observed normalized entropy (i.e., heterogeneity) scores for the Voronoi polygons and the block segments structures respectively.

As this summary highlights, for both zonal structures the median entropy score is equal to zero. Median – or the 50th percentile – is a pivotal point of a statistical distribution where half of the observations (i.e., zones) have the lower values (i.e., entropy scores), and the other half have the higher values than this point.

Considering that the theoretical range of the entropy score is from 0 to 1, having a median value of zero indicates a highly skewed distribution across the region. This means a disproportionate number of zones (i.e., at least half of the zones within each structure) are completely homogeneous, or in other words, they consist of only one land use type. In contrast, no instances of perfectly heterogeneous zone were identified in the region, and only a handful of zones revealed high levels of heterogeneity with the entropy score above 0.8. These are the zones mainly located in the region's historical activity centers in Uptown Waterloo and Downtown Kitchener, where multiple diverse activities positioned adjacent to each other.

Given such a skewed distribution, it is important to interpret the applied percentile thresholds by looking at their corresponding on-the-ground land use context. As the summary statistics table implies, the 45th percentile threshold associates with the lowest possible entropy score of zero. By applying this threshold, the DACZ model only allows the aggregation of those zones that consist of a same single land use type. This restriction maintains the heterogeneity levels of the created DACZs at the absolute minimum level too.

Interestingly, the second threshold at 55th percentile does not show a traceable impact on the associated entropy score for the block segments. It maintains the score at the absolute minimum level of zero. Nevertheless, for the Voronoi structure, this percentile threshold raises the corresponding entropy score slightly higher to the 0.01 level. This insignificant heterogeneity level can be interpreted as a zone consisting of a uniform land use pattern (e.g., a residential suburb) with a very few parcels (i.e., less than 1% of the zone's area) with different land use types.

Moving to the upper percentiles within the distribution, however, the entropy scores start to reveal some higher levels of heterogeneity. The 65th and 75th percentiles reflect sequential rises in their entropy score levels to 0.1 and 0.2 levels respectively. These entropy scores still represent homogenous areas like an inner suburb residential zone with presence of a few non-residential activities. These levels of heterogeneity thresholds open up the opportunity for some neighboring zones with a few dissimilar land uses to get merged.

Shifting further to the 85th percentile, the 0.3 entropy score shows a moderate level of heterogeneity that can be interpreted as a mixed-use zone where the single land use type dominance is waning.

Ultimately, the 95th percentile threshold shows a sharp upward rise to the 0.5 entropy score levels where the zones may include a balanced mixed of different land use types (e.g., residential, commercial and retail) that occupy relatively equal proportions of the zone's area.

Figure 4-4 presents the results of applying this set of percentile thresholds and their corresponding entropy scores on both zonal structures.

3.5.2.2 Dynamic Zones Area Size Threshold

Although the heterogeneity score threshold is a robust criterion to define the homogeneous dynamic zones, depending on the land use configurations "on the ground" there is a potential outcome in which the model fails to define zones with appropriate sizes. For example, in suburban fringes where uniform residential land use areas extend for several kilometers, it is likely for the iterative aggregation model to keep adding neighboring blocks until a very large encompassing zone is created while the heterogeneity levels still remain within the acceptable threshold. Such a large zone is likely not to be useful for some transportation analysis such as route assignment.

As we approached this problem in our model, we created a heuristic to get the right balance between the zone size and the number of disaggregate neighboring (input) zones to be merged for creating a dynamic zone (output). Following the steps visually presented at Section 3.5.1.3, for each starting disaggregate zone we envisioned two sets of neighboring zones to be evaluated for potential aggregations. We start with those immediately adjacent zones that shared a boundary with the central zone. If they could be merged by satisfying the aggregation criteria, we then look into the next level of (distant) neighboring zones, which where adjacent to the immediate neighbors. We repeat this aggregation evaluation and adding more neighbors until the combined entropy index and combined area of the zones remain lower than a defined threshold.

To determine the right upper limit for the size of a dynamic zone, we used a hexagonal grid analogy that helps manifest a well-shaped hypothetical aggregate zone based on a number of individual disaggregate zones. In practice, we implemented the DACZs area size cap parameter – as a multiplier of median input zones size – to prevent creating gigantic zones. This parameter was also conceived to implicitly set out an upper bound on the approximate number of individual zones that could be part of a dynamic zone.

This study utilizes zonal representation (as opposed to the rigid grid representation) as it better fits complex land use configurations within urban environment (Benenson & Torrens, 2004). While the zonal approach comes with the advantage of flexibility, the approximation of distance and neighborhood definition within zonal models are computationally demanding tasks (Moreno, Wang, & Marceau, 2009; Vliet, White, & Dragicevic, 2009; Wei et al., 2021). This applies to the iterative aggregation model we developed to create dynamic zones, where the computational requirements (i.e., memory and processing cycles) grow exponentially with the number of input zones and the intended degrees of (higher-order) adjacencies. To tackle this computational challenge, a multi-level dynamic method was taken to intuitively balance the model flexibility and its computational performance.

At the first level, the model identifies immediate adjacent zones for all input objects and stores that information in a mutable table; this precomputation saves a considerable amount of time required for the future spatial adjacency searches. Then, during each iteration (technically, through steps 6 to 8 of the iterative spatial aggregation model) if an aggregation happens, the model updates the adjacency table to reflect the new immediate neighbors of the generated zone.

This method leverage the Chessboard distance metric to determine higher degrees of adjacency. This metric is adopted from the regular grid tessellation approaches, where every neighboring zone (L1) that shares a boundary or corner point is considered one unit away the central zone (L0) . The regular grid abstractions offer computational efficiency for spatial adjacency and neighborhood definition (Al-Ogaibi, Sharieh, & Bremananth, 2014), and provide methodological robustness for complex spatial models (Stanilov, 2012b; Wegener, 2001). The detailed information on the choice of model distance metric is provided in Appendix C that provides future research with potential pathways to define distance metric assumptions. Following the expansion of the model spatial adjacency search through this multi-level heuristic, it was then necessary to set an upper bound for the zones size, and consequently to construct each dynamic zone as a unified whole. These solutions are again aimed to balance the size and number of the output dynamic zones within a computationally tractable level. The suitability of this methodological integration was evaluated through testing two threshold values for the area size cap (multiplier) corresponding to one and two degrees of spatial adjacency.

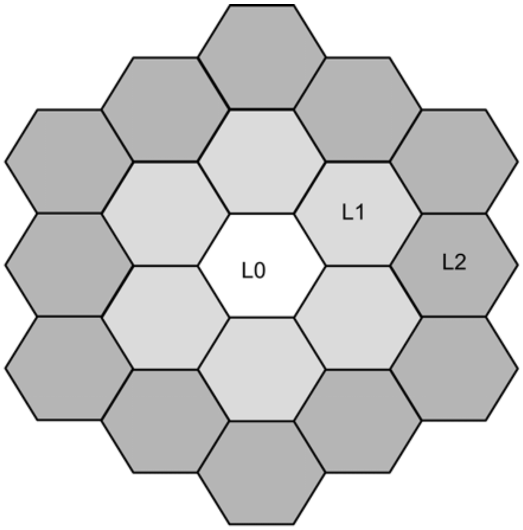


Figure 3-9 - Spatial adjacency levels in a hexagonal grid (Al-Ogaibi et a. 2014)

A value of 7 was selected as the first threshold to be tested as it implies one degree of adjacency within an evenly partitioned hexagonal grid. In this hypothetical zone system, a well-shaped dynamic zone could be comprised of a central cell (L0) with its six surrounding adjacent cells (L1). The size of this dynamic zone would be equal to 7 times of the input cells elements area.

Similarly, a value of 19 was used to approximate two degrees of adjacency from a given zone (Figure 3-9), where the size of the resultant dynamic zone could be as large as 19 times of the input cell elements area.

The choice of these values intended to approximate the area size cap in the irregular zonal structures of this study, using a hexagonal grid analogy. The hexagon is the optimal geometry to model neighboring zones adjacencies assuming the space is partitioned evenly (Birch, Oom, & Beecham, 2007). Moreover, connectivity of the zones in such a grid system is more efficiently computed through a discrete distance metric between the two zones, where the hexagonal tessellation minimizes distortion of the discrete distance metric compared to the continuous Euclidean distance (Ortigoza, 2015), as illustrated in Figure 3-9. This heuristic turned out to be computationally very efficient to create spatially compact and well sized zones considering the regional zones size distribution.

3.5.3 Graph-Based Partitioning Approach

To compare the proposed homogeneity-based aggregation model of this study, a graph-based spatial partitioning model was implemented to provide a baseline measure for performance evaluations. This comparative analysis was required to examine if zones created through the proposed method generated significantly more homogeneous analysis zones than the zones produced by the established methods. From a set of identified state-of-the-art contiguity based clustering approaches, the graph-based partitioning method was selected as it is known for computational efficiency and its methodological robustness (Aydin et al., 2021).

The graph-based partitioning is an outstanding technique to tackle the spatial optimization problem through a heuristic approach that is distinctly different from its counterpart methods. As its name implies, this method uses top-down partitioning of space, instead of bottom-up aggregation of zones. In fact, it embeds the required spatial contiguity criterion into the model construct and reduces the computational processing burden of adjacency search and evaluation for potential neighboring zones.

This heuristic is applied through conversion of an area-based spatial structure into a contiguity-based dual graph (Figure 3-10a.). This graph structure provides a simpler mathematical representation of the spatial zones that otherwise must be presented by means of complex objects along with their topological relationships. The graph model represents each input zone – regardless of its shape and size – as a node within a graph structure. It then constructs potential edges between those nodes if their corresponding zones share a border of any length. The method uses an indirect strategy to maximize intra-zonal homogeneity, considering a set of attributes of the connected nodes (i.e., their enclosed land uses in this study). First, if an edge exists between two specific nodes (i.e., having two neighboring zones) the algorithm calculates a distance metric between the nodes taking into account dissimilarity of their attributes (Figure 3-10b.). This distance measure is computed for each pair of directly connected graph nodes – meaning for all the neighboring zones in the original structure.

For example, when two zones are being evaluated to merge into a cluster, they have to have similar land use compositions. In a multivariate model, each land use class is considered an independent attribute that contributes to the total distance between the two zones. When land use classes within each zone considerably differ from a neighboring zone (i.e., connected node in the dual graph), the computed distance between them would be much larger indicating higher land use dissimilarity. It is important to note that this computed distance measure (i.e., dissimilarity measure) does not reflect the actual geographic distance metric per se.

The Minkowski distance metric is often used for purpose of dissimilarity measurement in the graph-based partitioning method, since it flexibly accommodates multivariate model that is necessary for calculating dissimilarity of multiple attributes (Seber, 2009). The mathematical representation of Minkowski distance between a given pair of nodes X, and Y is presented in (Equation 3):

$$X = (x_1, x_2, \dots, x_n) \; and \; Y = (y_1, y_2, \dots, y_n) \; \in \mathbb{R}^n$$

$$D(X,Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}$$

where:

(x_i) denotes i^{th} attribute from the vector of attributes X

(y_i) denotes i^{th} attribute from the vector of attributes Y

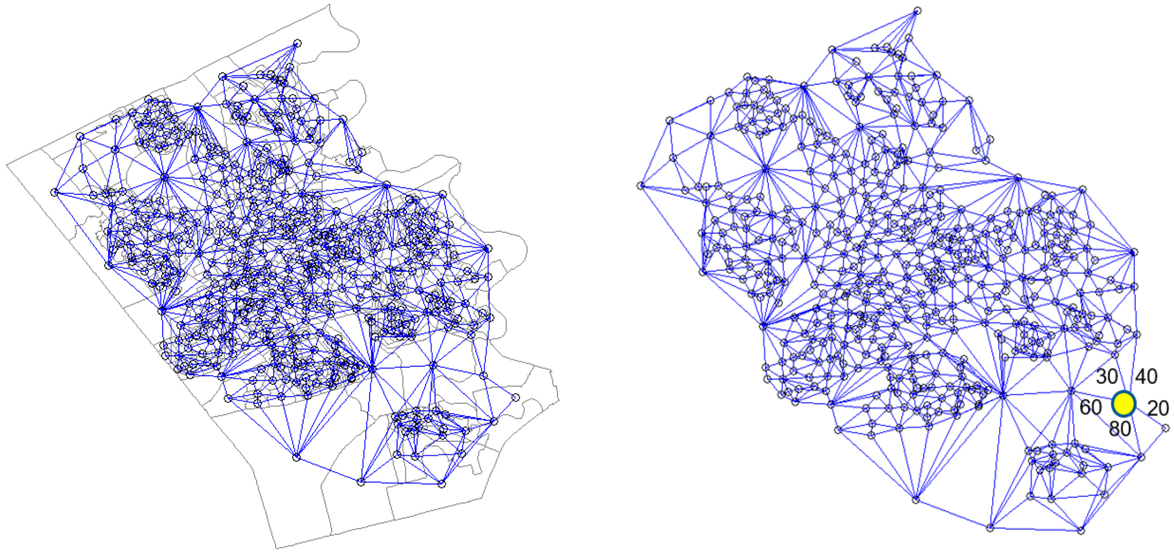
(p) refers to the order of distance measure. Minkowski distance is often applied with *p* being 2 corresponding to the Euclidean distance, or 1 resembling the Manhattan distance.

In the next step, the model starts with eliminating dissimilar neighbors to limit the optimization search space. This process is accomplished by extracting a particular sub-graph called minimum spanning tree (MST). The MST is a subset of the dual graph that connects all the nodes with minimum possible total distance of the edges. The MST is a sub-graph without any cycles meaning that for a graph with *n* nodes (i.e., *n* zones) it has *n*-1 edges. This reduced number of the optimization elements significantly decrease the model complexity and improves the computational efficiency of this method (Assunção et al., 2006). Eventually, the algorithm starts pruning the MST graph with the goal of identifying a predetermined number of clusters as subsets of the minimum spanning tree (Maravalle & Simeone, 1995).

Using the values derived from the sensitivity analysis in previous section, a set of desired number of clusters (i.e., equivalent to the number of produced DACZ) were applied to identify homogeneous clusters of block segments dual graph within Waterloo Region. Technically, a sequential series of steps were implemented to perform the graph-based partitioning as follows:

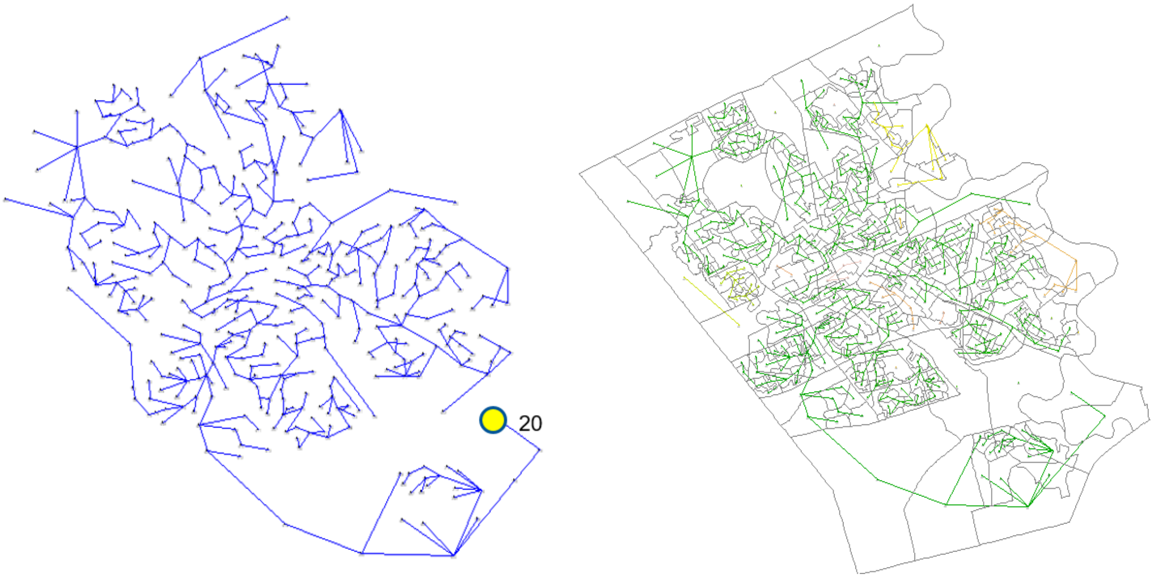
1) Computing surface area of all land use classes observed within each block segment
2) Constructing the dual graph based on the block segments (Figure 3-10a.)
3) Standardizing the land uses surface area
4) Transforming the standardized values into the dual graph nodes
5) Computing the pairwise land use dissimilarity between the connected nodes and assign the derived values as the distance attribute of the connecting edge (Figure 3-10b.)
6) Carrying out the dual graph pruning to extract the minimum spanning tree (MST) (Figure 3-10c.)
7) Defining a set of values derived from the sensitivity analysis as the desired number of clusters
8) Performing the SKATER algorithm to identify clusters of the nodes with homogeneous land uses
9) Reflecting back the algorithm results into the original zonal structure (Figure 3-10d.)
10) Reconstructing the area-based objects to represent the homogeneous activity clusters

Figure 3-10 shows the major steps in the process of graph-based partitioning in the study area. For the purpose of map clarity, the Census Dissemination Area (DA) scale was used as the base zonal structure in this visualization.



(a) Constructing a dual graph based on the base zonal structure

(b) Computing the pairwise dissimilarity distances between the connected nodes (zones)



(c) Extracting the minimum spanning tree (MST)

(d) Identifying homogeneous clusters using SKATER algorithm

Figure 3-10 - Identifying homogeneous land use clusters using graph-based method

3.6 Statistical Analysis of Travel Behavior Heterogeneity

One of the empirical objectives of this study is to evaluate usefulness of the homogeneity based analysis zone delineation for transportation related analysis and modeling. This section presents the analyses for answering the second research question which investigates how the built environment characteristics (e.g., density and diversity) and access to rapid transit can influence travelers' behavior. In earlier sections, we developed a dynamic zonal structure through bottom up aggregation of block segments. We then aggregated certain travel behavior attributes at dynamic zones and a conventional TAZ structure.

In this section we compare what we learned from these two different zonal structures on travel behavior distribution across the cities of Waterloo and Kitchener. This work helps to determine whether or not these zonal structures provide zones for which the travel behavior is more (or less) heterogeneous. We primarily look at the travel behavior distributions (mode share) within the DACZ and PLUM structures, for all zones and subset of classified zones. We then expand the analysis to explore the travel time distributions for all zones, this time only within the DACZ system.

The first zonal structure used in the comparison is a set of Dynamic Activity Cluster Zones (DACZs) created through homogeneity based spatial aggregation. The second zone system is a pre-defined conventional traffic analysis zones (TAZ) that is used by the Waterloo Region Municipality in their Population and Land Use Model called PLUM.

To evaluate the suitability of these zonal structures, detailed geocoded travel data from the latest regional travel behavior survey (TTS 2016) were aggregated for each zone system. The TTS comprehensive set of travel attributes include the trips' origin and destination, travel mode and the purpose of trip. The TTS travel data are collected during an entire workday period and while the trip purposes are broken down into different categories (e.g., home-based work (HBW), home-based school (HBS)), each trip count includes a return trip specified by either ends of that trip. For instance, a HBW trip includes the trips from home to work and from work to home. The trip purposes presented within TTS data set and used in this research are:

a) Home-based work trips (HBW)
b) Home-based school (HBS)
c) Home-based discretionary (HBD)
d) Non-Home based (NHB)

TTS also provides 12 categories of primary mode of travel. In this study, multiple sub-categories were combined together and five major transportation modes defined as follows:

a) Auto as driver
b) Auto as passenger (including trips by taxi and paid rideshare)
c) Public transit (including bus rides, school bus trips and rail-based trips)
d) Walking
e) Cycling and other modes

Once the trip data for the intended travel modes and trip purposes have been processed and reclassified, a multi-level approach was employed to identify travel behavior heterogeneity across the study area. This framework consisted of three spatial scales to allow drawing a broad perspective on the likely impacts of the region's built environment heterogeneity on its travel behavior patterns. For each designated scale the transportation mode share and the average travel time were computed for different trip purposes. One additional analysis was performed for all the trips regardless of their specific purpose. The selected metrics were calculated for both zone systems incorporating the trips originating within cities of Waterloo and Kitchener.

Within this multi-level framework, the most disaggregate scale consisted of individual analysis zones corresponding to the DACZs (N=1228) and the PLUM (N=692) zones. In contrast, the most aggregate scale encompassed entire study area as a whole (N=1). At the middle of this spatial spectrum, an intermediate spatial scale was analytically defined by taking into account the population density and land use diversity levels of the zones identified earlier (N=9).

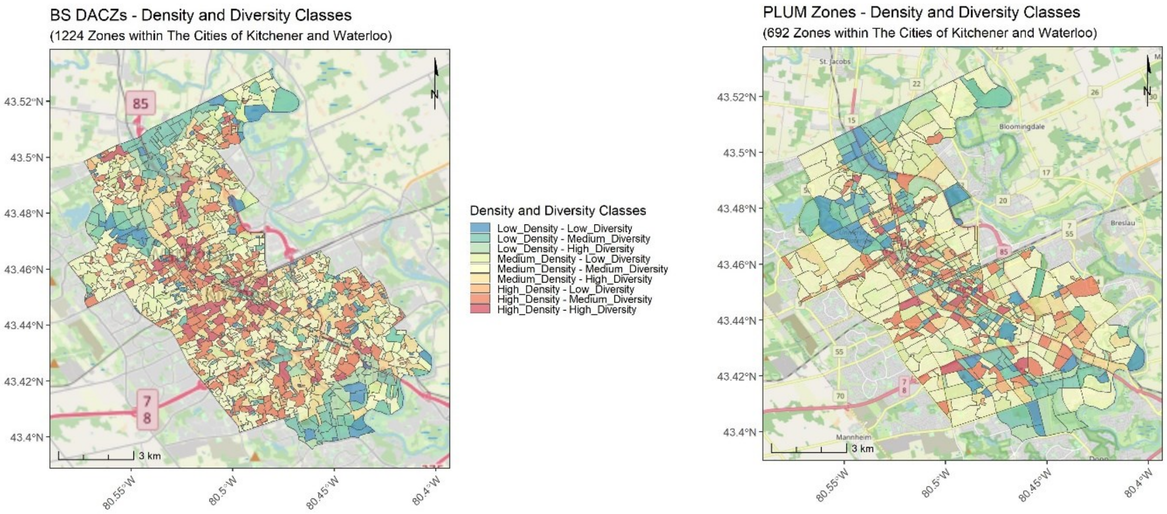
Devising this scale was motivated by theories that suggest travel behavior patterns are determined by the built environment characteristics along with access to transportation infrastructure. For instance, an area with diverse land uses and high levels of population density is believed to be more conducive to active transportation and public transit trips compared to a dispersed uniform subdivision. The subsequent section describes details of the steps taken to define the build environment categories as a complementary spatial scale for LUTI analyses.

3.6.1 Built Environment Classification Based on the Combined Population Density and Land Use Diversity Levels

In this step, land use diversity levels for the zones were computed using the enhanced entropy measurement method introduced in Section 3.6.1. Then, this heterogeneity score was paired with the population density of the zones and used to classify each zone into a certain built environment category.

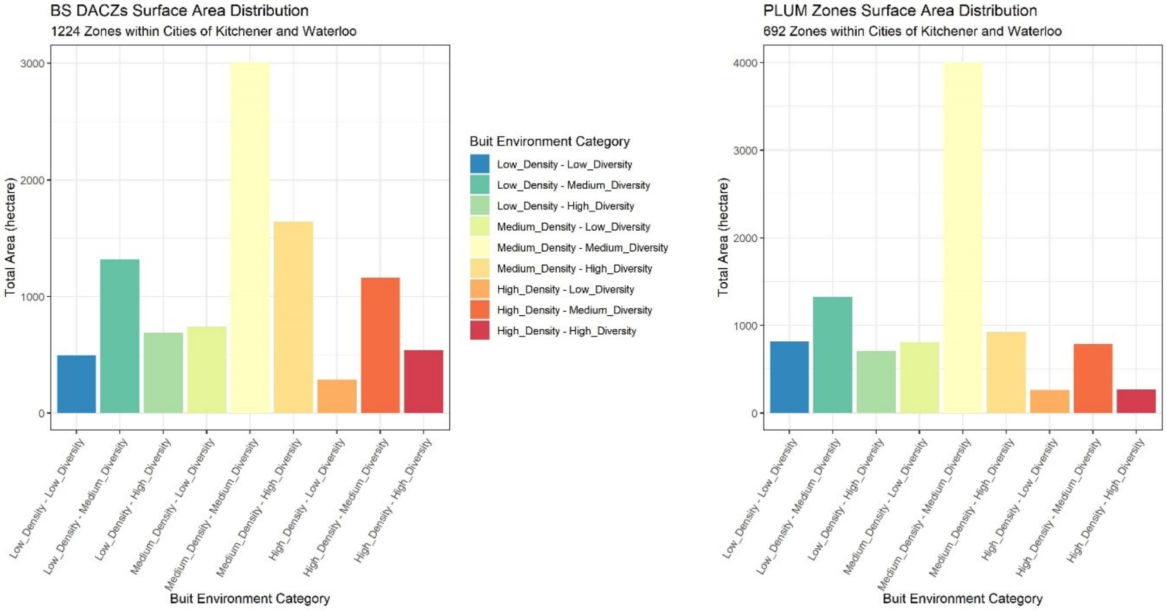
Since the land use diversity and the population density levels have to be interpreted relative to the regional levels, using the statistical distribution three distinct levels of population density and entropy score were identified for these variables namely *High, Medium* and *Low* categories. The specified High- categories are associated with the top 25 percentile of the high levels, and the Low- categories related to the bottom 25 percentile of the low levels of density and diversity distributions.

To represent the full range of built environment heterogeneity within the study area, the three classes of population density and three classes of land use diversity were intersected and a set of nine classes combined were formed. Figure 3-11a. and b. show the spatial distribution of the consolidated built environment categories for the DACZs and the PLUM TAZs accordingly. A blue-to-red color gradient is used to illustrate the levels of density and diversity for the zones, where at one end of the gradient, areas in red corresponding to the “High Density – High Diversity” class and on the other end, areas in blue showing the "Low Density – Low Diversity" category. Figure 3-11c. and d. also depict the statistical distributions of the total surface areas (in hectare) identified under each category.



(a) Spatial distribution of the DACZ zones by their built environment heterogeneity

(b) Spatial distribution of the PLUM zones by their built environment heterogeneity



(c) Statistical distribution of the total surface area in each built environment category (DACZs)

(d) Statistical distribution of the total surface area in each built environment category (PLUM zones)

Figure 3-11 - Spatial Distribution of Zones by Their Built Environment Heterogeneity Levels, (These categories and their characteristics will be discussed in more detail in the next chapter.)

3.7 Spatial-Temporal Analysis of Regional Land Use Development

The final objective of this study is to provide a historical perspective on the spatial dynamics of land use developments in Waterloo Region. In respond to the third research questions, this analysis aims to test the hypothesis that investment on the regional light rail rapid transit has stimulated more new developments to the Region's Central Transit Corridor (CTC). It examines the historical records of building permits from a decade before announcement of the LRT project funding until a decade after, to assess whether the areas within the CTC have disproportionately attracted more and or larger developments (i.e., higher construction values per issued permit) compared to the rest of study area. Further, the analysis investigates the likely impacts of land use heterogeneity and population density levels on development activities trend.

From the building permits dataset provided by Cities of Waterloo and Kitchener, the construction value variable is used as a proxy to quantify land use development activities. Since the original dataset reflects the nominal value of the issued building permits, a time-varying price index was used to adjust the permit value figures taking into account the inflation rates during the study period. Statistics Canada provides two sets of price indices to deflate nominal prices in the housing and construction sectors based on a fixed dollar value. The New Housing Price Index (NHPI) is a monthly index to measure the changes in the selling prices of new residential houses of all new residential unit types (i.e., single homes, semi-detached homes and townhomes) constructed and listed for sale or sold in Canada (Statistics Canada, 2021b). The Building Construction Price Index (BCPI) is a quarterly index that reflects changes in the prices for constructing a range of commercial, institutional, industrial and residential buildings in major census metropolitan areas in Canada (Statistics Canada, 2021a). Although BCPI could be a more relevant measure to adjust the building permits construction values over time, no consistent time series of the BCPI is available for the temporal period of this study due to the methodological changes happened in calculation of this index.

Hence, to ensure comparability of the building permit values over time, the monthly measure of New Housing Price Index ($NHPI_t$) – with the land value excluded from the prices – was applied to adjust the nominal prices from observations for any given time (denoted by $Nominal\ Value_t$ in Equation 4). For every issued building permit, the inflation adjusted construction value was then calculated based on the dollar value at the end of study period (December 2019). The applied index was derived from the metropolitan areas of the Province of Ontario with the reference period being December 2016 (the latest released) for which the NHPI equals 100.

Equation 4:

$$Inflation\ Adjusted\ Value_{2019-12} = \frac{Nominal\ Value_t}{NHPI_t} \times NHPI_{2019-12}$$

where:

$Subscript_t$ denotes the time period corresponding to the date (Year-Month) of issued permits

In the next step, using the DACZ zonal structure developed and evaluated in Section 3.9.1, a set of three aggregation model was developed to summarize temporal records of the issued building permits at biannual intervals. The biannual temporal scale was chosen to smooth out the fluctuation observed at the annual and seasonal resolutions within the original dataset. This approach helps to reveal the overarching patterns of the regional land use developments from 2000 to 2019, specifically with distinction between the two periods *before* and *after* the LRT project funding announcement (2010-2011), and during the construction period of the phase-one of the LRT line that began in 2014 and ended in 2018.

The first two models were intended to summarize the building permit records by the permit type at highly aggregate spatial scales. Model 1 accumulated the issued building permits within the study area all together. Model 2, in contrast, accounted for the geography of the Central Transit Corridor (CTC) and aggregated the building permit records for the areas *within* the CTC and *outside* of the CTC independently.

Model 3 was built upon an intermediate spatial scale using the built environment categories developed earlier by accounting for the latest land use diversity levels (for year 2019) and the population density of the zones (for year 2016). This model was used to sum up the construction values over the nine heterogeneous built environment categories regardless of the permit types. Ultimately, the revealed patterns of land use developments were analyzed to unravel the areas that might have disproportionately attracted more and or greater land use developments (i.e., higher construction values) potentially due to the benefits of heterogeneous built environment and access to rapid transit.

3.8 Methods Summary

This chapter outlined the methods that were used in this study, and demonstrated the spatial models developed and examined to answer the research questions. It started with introduction of the conventional entropy index as a land use diversity measurement method. Then, an enhanced land use diversity index was proposed and implemented to improve upon the conventional measure. Using this enhanced diversity metric, an iterative spatial aggregation model was developed to delineate Dynamic Activity Cluster Zones (DACZ) – as a novel analysis zonal structure for land use and transportation system interactions studies. This model's behavior in creating homogeneous DACZs was then evaluated through a systematic sensitivity analysis, taking into account variability of the spatial adjacency parameter and a range of heterogeneity threshold.

Next, the spatial aggregation method to create zonal structure was applied to a mid-size metropolitan case study area and evaluated against multiple conventional zonal structures. To demonstrate the model ability to represent more homogenous zones, a specific set of DACZ was tested compared to the TTS traffic analysis zones (TAZs). This evaluation was further extended by implementing and testing a graph-based spatial partitioning method as a modern state-of-the-art approach for creating spatial clusters (i.e., zones).

Furthermore, the same set of DACZ was examined to assess the model efficacy in generating more homogenous zones in terms of travel behavior compared to the PLUM zones. This is the analysis zonal structure for the Waterloo Region's Population and Land Use Model. A multi-level approach was employed and statistical distributions of the mode share metrics were calculated and compared for both zonal structures at multiple spatial scales. As a prerequisite to this assessment, a set of heterogeneous built environment categories was identified for both the DACZ and the PLUM zones by taking into account the land use diversity and the population density levels of their zones.

Lastly, the historical trend of development in Waterloo Region was evaluated for a decade before and after the announcement of the LRT project funding approval in 2010. Based on the DACZ zonal structure developed earlier, the temporal building permits data from 2000 to 2019 were aggregated at biannual intervals and their construction values were adjusted using time-varying price index for Province of Ontario. Then, a set of temporal analyses were performed at three spatial scales to reveal the areas that might have disproportionately attracted more and or larger land use developments (i.e., higher construction values per issued building permits), potentially due to the benefits of heterogeneous built environment and or proximity to high quality transit corridor within the CTC. The corresponding scales range from the intermediate (i.e., the built environment categories) to semi-aggregate (i.e., the areas within or outside of the Regions' Central Transit Corridor), and aggregate scales (i.e., the entire study area).

Chapter 4: Results and Discussion

4.1 Overview

Following the presentation of the methods in Chapter 3, this section discusses the results of the spatial models' implementation and complementary statistical analyses that were performed to answer the key research questions. As presented in earlier sections, a set of four major interconnected analyses were conducted corresponding to the overarching objectives and specific research questions of this study. This series of analyses initiated with development of a novel zonal structure through a spatial aggregation approach taking into account homogeneity of the land uses. This new zonal system was then applied in a case study in Waterloo Region and it was compared against a set of conventional zonal systems.

The first two analyses aimed to respond to the questions of the appropriate spatial scale for investigating land use and transportation system interactions, and to reveal the potential advantages of using dynamically defined zones in travel behavior studies. Hence, in an exploratory work, using the proposed spatial aggregation method a Dynamic Activity Cluster Zones (DACZ) structure was created and examined over a suburban subdivision in the City of Waterloo that has been recently connected to the newly developed regional LRT system. This DACZ structure was then compared to the pre-defined conventional traffic analysis zones (TAZ) in terms of effectiveness in representing more homogeneous zones, and in delineating spatial boundaries. Section 4.2 describes details of this analysis further.

Next, by performing a systematic sensitivity analysis the potential impacts of the spatial model parameters were tested across multiple zonal structures created for a broader area covering cities of Kitchener and Waterloo. This analysis resulted in 20 DACZ structures each of them comprised of different number of zones with varying performance in producing homogeneous zones. To further examine the DACZ approach in parallel with the conventional spatial regionalization methods, a state-of-the-art graph-based spatial partitioning method was applied and a new set of zonal structures created, each of them having the exact same number of the zones as their DACZ counterparts. Section 4.3 elaborates on the results of the DACZ model sensitivity analysis, and Section 4.4 presents the performance comparison between the results of the DACZ method and the graph-based partitioning approach in creating zonal structures.

The next analysis was geared towards the application of generated DACZ in a regional travel behavior study. It was intended to highlight the advantages of dynamically defined zones in travel behavior analysis by exploring how built environment characteristics (e.g., density and diversity) and access to rapid transit may influence travelers' behavior across Waterloo Region. This analysis was developed using the PLUM zones (N=692) as the baseline zonal structure in tandem with a previously produced DACZ, comprised of relatively close number of zones (N=1228). This DACZ structure was created based on a heterogeneity threshold level that allowed the creation of moderate mixed-use zones over the block segment polygons. Section 4.5 demonstrates the results of this analysis, where the latest available regional daily travel behavior data from TTS 2016 were processed taking into account different travel modes and trip purposes across the region.

The final analysis sought to explain the trend and pattern of land development in the region as a function of land use homogeneity levels and access to rapid transit. A set of temporal analyses were performed and the trend of changes in the building permits' construction values were presented at multiple spatial scales. The average and the total permits construction values were calculated to highlight differences in development activities across heterogeneous built environment categories and relative to the Regions' Central Transit Corridor. These biannual trends, covering a 20-year period, were then interpreted to identify the areas that might have disproportionately attracted more and or greater land use developments between 2000 and 2019. The results of this work are discussed in Section 4.4.

4.2 Dynamic Activity Cluster Zone (DACZs) Model Assessment

This section presents the outcomes of the DACZ model application as the proof of concept for the homogeneity-based iterative spatial aggregation method. This investigation comprises a series analyses including development and statistical assessment of an enhanced measure of entropy, application of this measure to create the Dynamic Activity Cluster Zones (DACZ) and evaluation of them in a systematic sensitivity analysis. The section concludes by reporting the comparisons of the DACZ structure against the conventional TTS TAZs and the graph-based spatial clusters in terms of the overall homogeneity distribution and the zone size balance.

4.2.1 DACZ and the Conventional Traffic Analysis Zones (TAZs) Comparison[4]

In order to illustrate the improvements upon conventional methods, a set of conventional TAZs and the dynamically defined zones were employed and evaluated within a diverse metropolitan district. The results revealed more effective performance of the proposed method in forming of homogeneous zonal structures. And, as a result, improved ability to forecast transportation behavior.

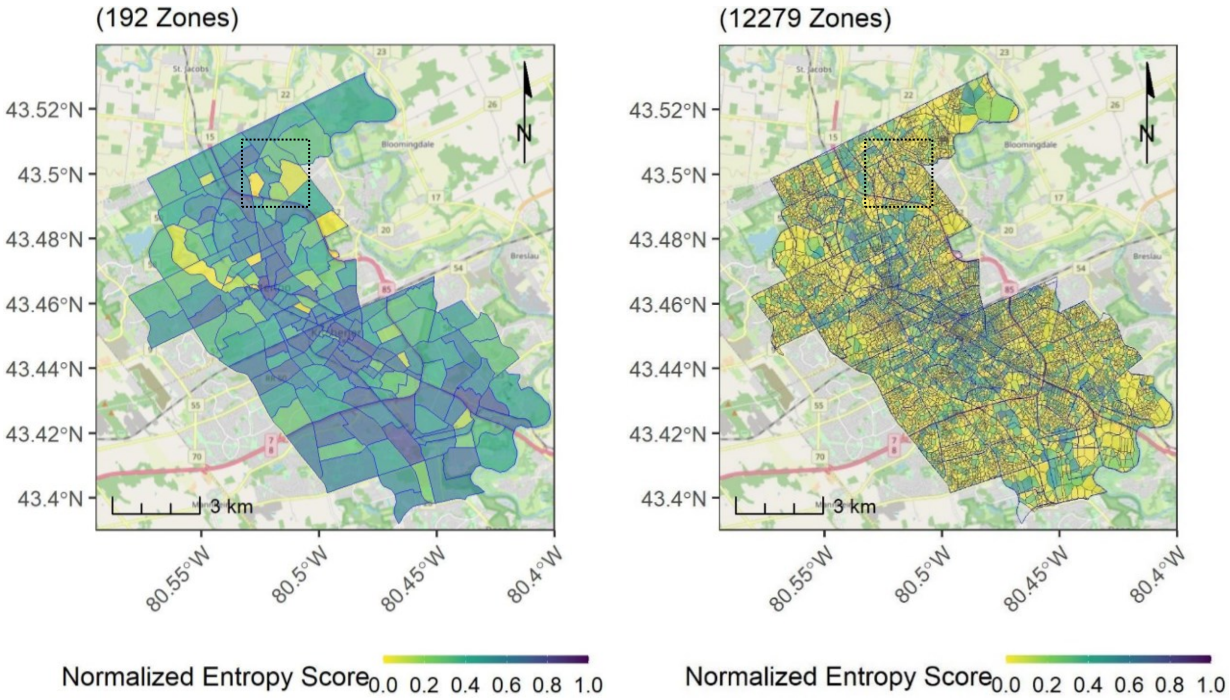
In the first step, the Enhanced Entropy model was applied to calculate heterogeneity score for each block segment as demonstrated in section 3.6. Figure 4-1a. and b. show the spatial distribution of heterogeneity index computed for both the block segments and conventional TAZs.

Within the cities of Kitchener and Waterloo, in the area for which this study had access to land use information, 53,000 parcels were analyzed and the Enhanced Entropy scores were computed for 192 conventional TAZs, along with 12,000 analytically generated block segments.

---

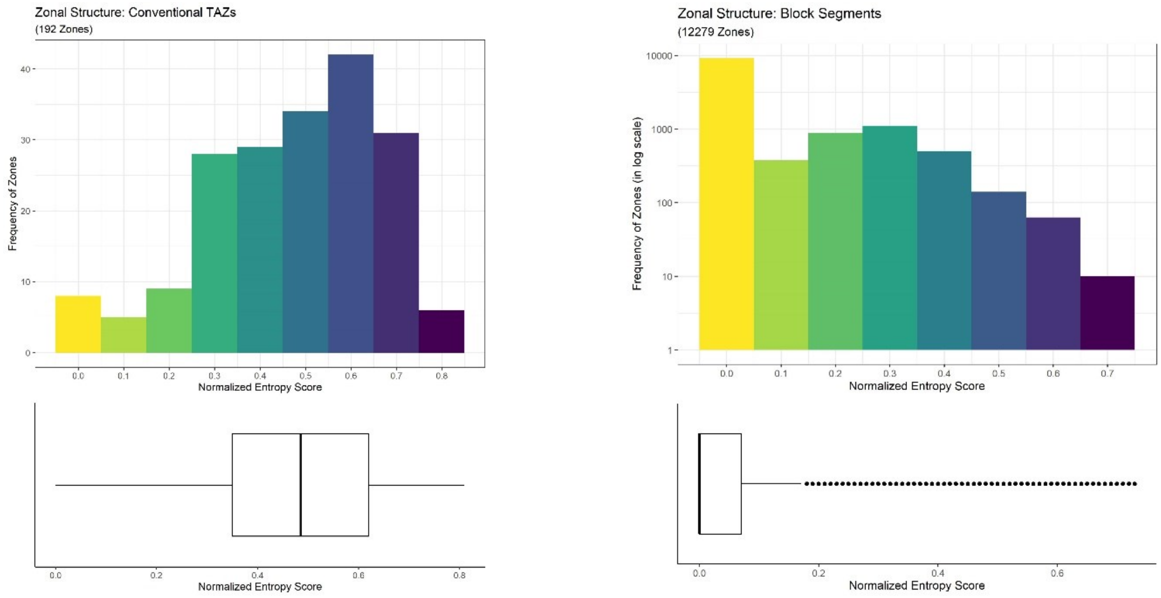
[4] This section is based on the following peer-reviewed conference publication:
Fard, P., Casello, J. M., & Xu, X. M. (2021). An Approach to Measure Land Use Heterogeneity to Identify Homogeneous Urban Activity-Clusters. Transportation Research Board 100th Annual Meeting. January, 2021.

It is important to note that the transportation network data that was used to generate block segments for this analysis was from the year 2018. Also, the land use reclassification scheme applied in this analysis is different from the rest of this thesis in that it consisted of 30 classes across the Waterloo Region, while only 16 of these classes were observed within the suburban case study area.



(a) Heterogeneity score within conventional TAZs

(b) Heterogeneity score within block segments



(c) Distribution of heterogeneity scores for conventional TAZs (in absolute scale)

(d) Distribution of heterogeneity scores for block segments (in log scale)

Figure 4-1 - Spatial and statistical distributions of heterogeneity index at regional scale for the tested zonal structures

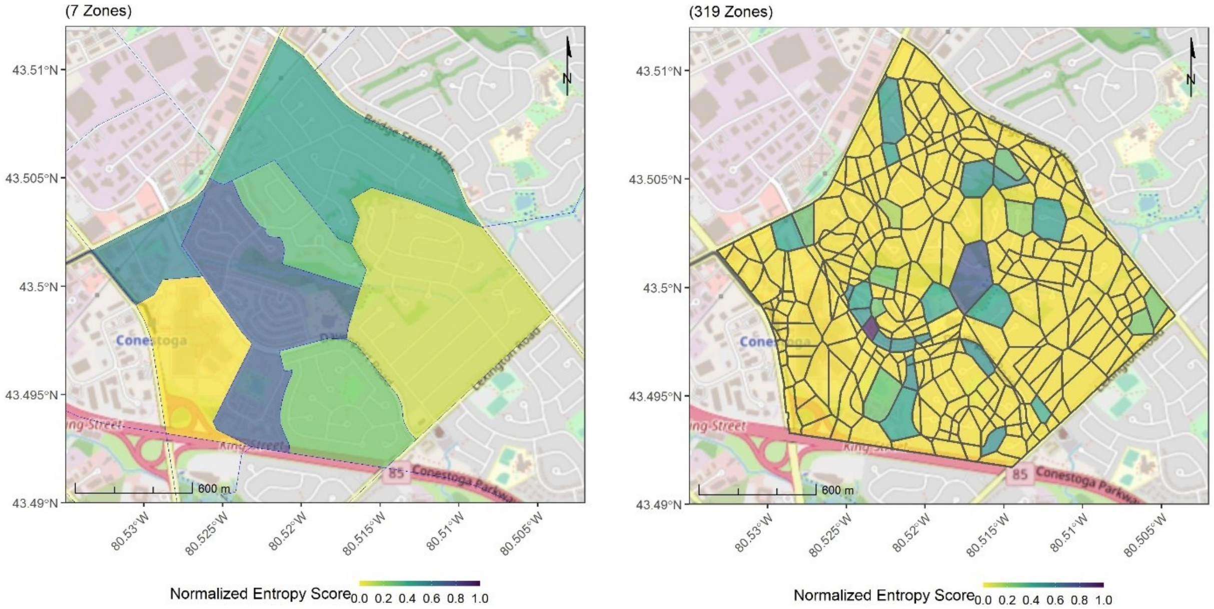
These numbers provide an outlook on the computational resources required by the Dynamic Activity Cluster Zones (DACZ) creation model. The number of the initial block segments sets the base number for the combinatorial possibilities of aggregation, as a result, the computation scale that the spatial aggregation model should be able to handle. The number of TAZs indicates the practically desired number of zones, upon which the current regional transportation model is built.

Table 4-1 summarizes descriptive statistics for the conventional and Enhanced Entropy scores calculated for the TAZs, and the block segment zonal structures. The figures calculated for both regional scale (Figure 4-1a. and b.), and at the case study area level (Figure 4-2a. and b.).

Spatial Scale and Extent on Analysis	*Number of Zones*	*Average Conventional Entropy Score*	*Standard Deviation of Conventional Entropy Score*	*Average Normalized Entropy Score*	*Standard Deviation of Normalized Entropy Score*
All the TAZs within cities of Waterloo and Kitchener	n=192	0.638	0.207	0.473	0.188
TAZs within Suburban Subdivision	n=7	0.474	0.318	0.337	0.258
Block Segments within cities of Waterloo and Kitchener	n=12279	0.197	0.354	0.070	0.134
Block Segments within Suburban Subdivision	n=319	0.077	0.246	0.041	0.135

Table 4-1 - Descriptive Statistics of the Entropy Scores Computed for the Conventional TAZs and Block Segments

As the statistical distributions indicate, the conventional TAZs have a wider range of Entropy score and hence each TAZ represents a more heterogeneous set of land uses compared to block segments. This is also evident from Figure 4-1c. and d. where the absence of yellow bar on the left graph, compared to the prevalence of yellow bar on the right indicates more intense heterogeneity of the TAZs. In contrast, within the block segments zonal structure, there are larger number of zones that have lower Entropy scores. These relatively small and homogeneous zones are building blocks for creating larger homogeneous activity clusters.



(a) Heterogeneity score within conventional TAZs (b) Heterogeneity score within block segments

Figure 4-2 - Comparison of heterogeneity score at the subdivision scale

In the next step, the Dynamic Activity Cluster Zone (DACZ) structure was created by aggregating those block segment that are not only similar in their lower levels of entropy but also have consistent types of land uses. To achieve a reasonably homogeneous set of DACZs, the potential heterogeneity levels for the aggregation process were explored and a set of thresholds was applied based on the statistical distribution of heterogeneity scores across the entire study area. As Figure 4-1c. and d. highlight, the heterogeneity score of block segments has a skewed distribution with a disproportionate number of segments with lower entropy values, peaking at zero. This means at the regional scale there are large number of block segments that are perfectly homogeneous. However, on the right side of distribution there is a smaller number of heterogeneous zones mainly in urban cores and likely within the Region's Central Transit Corridor (CTC) – the highest density and most diverse corridor in the region now served by Light Rail Transit. With a maximum heterogeneity score of 0.7, there is no perfectly heterogeneous zone found in the region. This is partly due to specific local land use and parcel fabric configurations, and also as a result of exclusion of certain land uses from entropy score computation. This distribution confirms the visual impression that the heterogeneity index map at the regional scale (Figure 4-1b.) and also the case study area map (Figure 4-2b.) convey.

Considering this empirical distribution, three exploratory scenarios were designed to perform the Dynamic Activity Cluster Zones creation. For Scenarios A, B and C, the heterogeneity percentile thresholds were established at 75, 85 and 95 percentile of the regional distribution, respectively. Table 4-2 presents descriptive statistics for the derived DACZs under each exploratory scenario.

Figure 4-3a. and b. show the steps taken to create DACZs. Figure 4-3c. presents the raw output of spatial aggregation model on block segments using heterogeneity threshold from Scenario C. Figure 4-3d. reflects the same output while incorporates geometries from the underlying parcel fabric. The differences in the total number of zones in these two figures and the observed spatial gaps within Figure 4-3d. originated from mismatches between the land use records and the parcels fabric data.

A comparison of Figure 4-3c. and d. with Figure 4-2a. illustrates qualitatively the improvement to be expected from applying the proposed algorithm. The larger areas of paler color (that is, areas of higher homogeneity) illustrate the improvement. A comparison of the average Normalized Entropy Scores for Scenarios A, B, and C (shown in Table 4-2, column 5, rows 1-3) with the average Normalized Entropy Score for the current TAZs (shown in Table 4-1, column 5, row 2) illustrates the same improvement in quantitative terms.

Exploratory Scenarios for DACZs created using alternative threshold from heterogeneity score distribution	*Percentile Threshold*	*Heterogeneity Score Corresponding to the Percentile Threshold*	*Number of Clusters (with valid land use data)*	*Average Normalized Entropy Score*	*Standard Deviation of Normalized Entropy Score*
Exploratory Scenario A	75th	0.07	n=20	0.194	0.191
Exploratory Scenario B	85th	0.25	n=19	0.201	0.220
Exploratory Scenario C	95th	0.39	n=18	0.245	0.232

Table 4-2 - Descriptive Statistics of the Entropy Scores Computed for Dynamic Activity Cluster Zones (DACZs)

(a) Heterogeneity score within block segments

(b) Block segments ranked based on their normalized heterogeneity score

(c) Dynamic Activity Cluster Zones created based on analytically generated block segments

(d) Dynamic Activity Cluster Zones reflecting parcel boundaries (only for parcels with valid land use data)

Figure 4-3 - Heterogeneity score at block segments and their derived DACZs (Scenario C)

Overall, the Dynamic Activity Cluster Zones represent sets of more homogenous land which likely generate more consistent travel behavior. In contrast, the way that the conventional TAZ was built here was more arbitrary in terms of breaking boundaries (presented as the blue line on the map). For instance, in this subdivision there is no reason for the central area of the subdivision to be broken up to multiple TAZs because land uses are sufficiently homogeneous that the travel behavior of the people who live there is likely to be very consistent. Moreover, in the southern part of subdivision, higher levels of heterogeneity can be observed that necessitates creating multiple zones, where the conventional TAZ fails to address that need.

4.2.2 DACZ Model Sensitivity Analysis

This section demonstrates the outcomes of the DACZ creation model sensitivity analysis including the number of output zones, their average land use heterogeneity levels and their size distributions as a function of the input parameters. Since the set of test input parameters included five distinct aggregation thresholds applied over two levels of spatial adjacency, the sensitivity analysis resulted in 10 particular DACZ structures for each input zonal layer (i.e., Voronoi polygons and block segments).

Figure 4-4 depicts the aggregate results of sensitivity analysis in terms of the number of generated dynamic zones (y-axis), given a set of aggregation heterogeneity threshold percentiles (x-axis). The different levels of spatial adjacency (i.e., multiplier parameters) are denoted by line color and different input base zonal structures denoted by line types. In this graph, by increasing the heterogeneity threshold levels, a decreasing trend of the number of generated DACZs is evident.

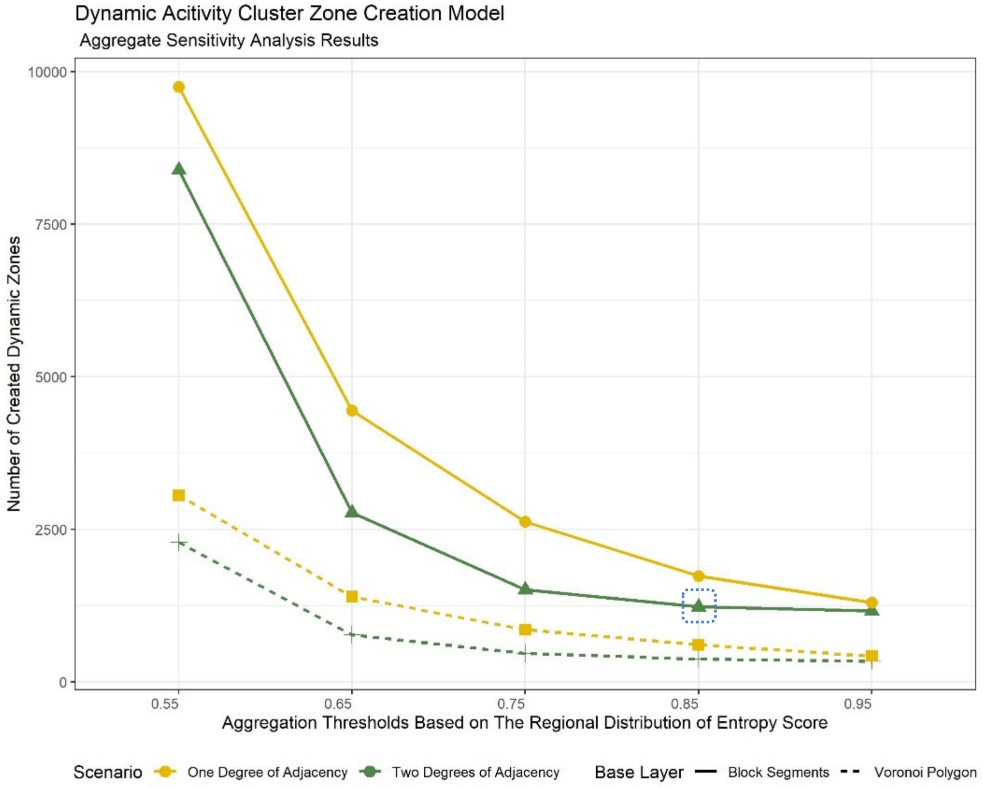


Figure 4-4 - DACZ model sensitivity analysis results (number of created zones)

On the left side of the chart, the DACZs have to have very low entropy scores, indicating very low heterogeneity. Here the model produces larger numbers of more homogeneous zones. Moving to the right, the higher threshold values allow the zone to be more heterogeneous, so it produces fewer zones that are larger in size.

The blue box in this graph highlights the point with the preferred level of disaggregation, where applying a reasonable level of heterogeneity threshold and using the block segments as the input to the model, the DACZ algorithm produces a similar number of zones (N:1228) compared to the baseline PLUM zones (N:692). As depicted in Table 3-1, within the regional distribution of the heterogeneity scores, the 85th percentile corresponds to an entropy score of 0.33 which can be interpreted as moderate levels of land use heterogeneity where mixed-use areas start to emerge. Using this threshold allows the model to form simple mixed-use dynamic zones, but prevents aggregation of very diverse mixed-use areas like urban core activity center zones. This DACZ layer will be further used for the built environment and travel behavior analysis in Sections 4.3 and 4.4.

Figure 4-5 provides a closer look at the heterogeneity score distribution for the generated zonal structures. It depicts the average normalized entropy score within DACZ structures (y-axis), plotted against the same set of aggregation heterogeneity threshold percentiles (x-axis) as presented in previous figure. This plot aims to examine the performance of the iterative aggregation model in producing homogeneous zones.

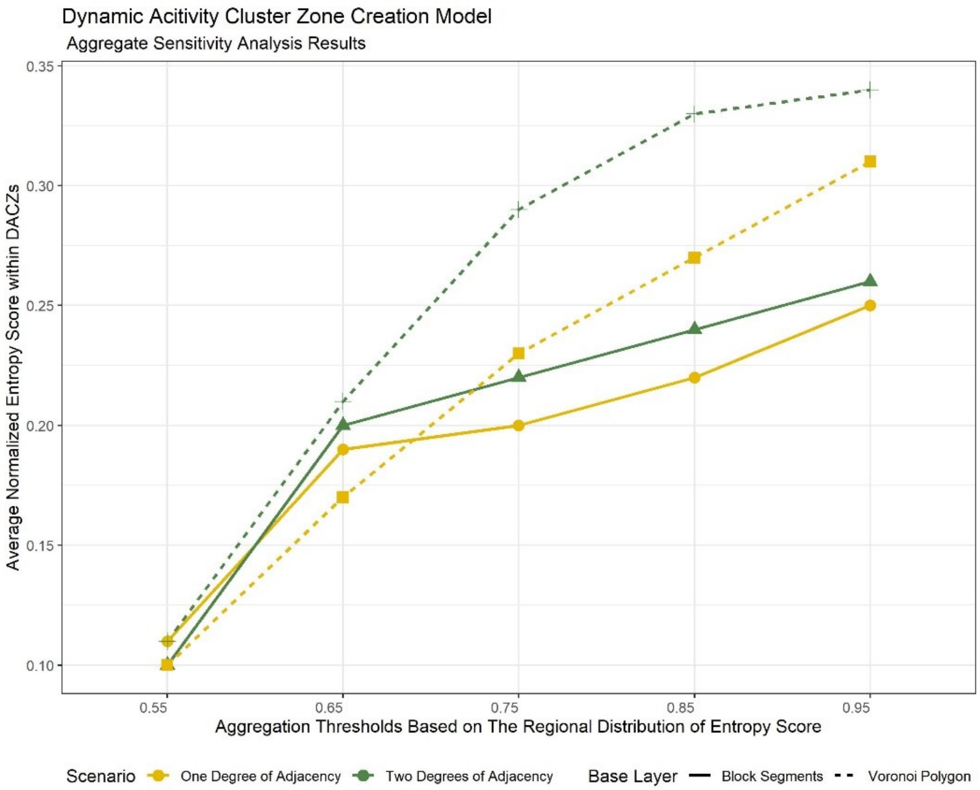
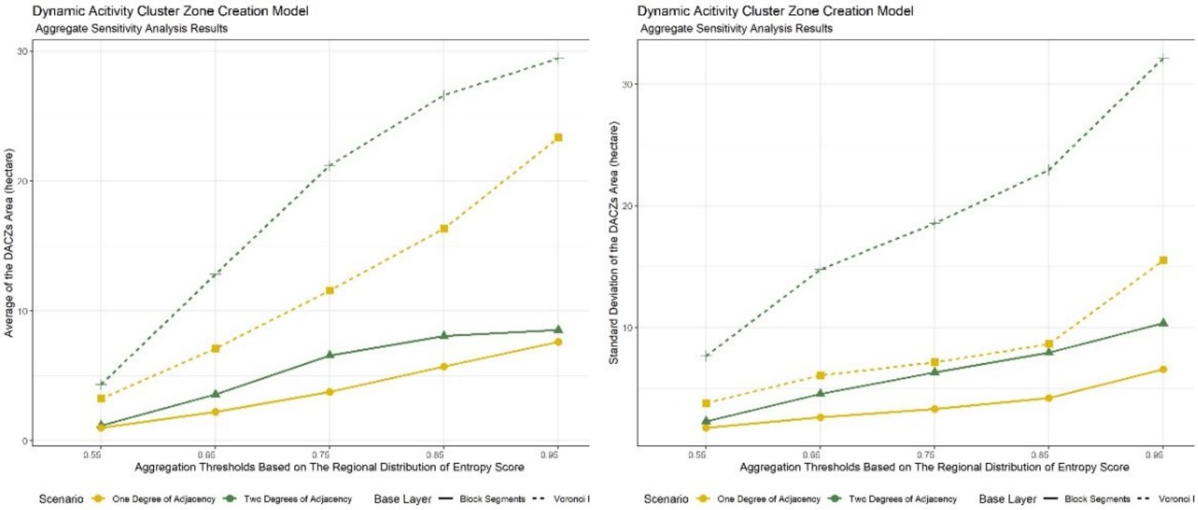


Figure 4-5 - DACZ model sensitivity analysis results (average normalized entropy score)

As the pattern of changes indicates, for all the applied aggregation scenarios, increasing the heterogeneity threshold levels produces a consistent upward trend in the average normalized entropy score. However, for the two different base zone systems tested in this analysis, a slight difference in their score gradient is detectable. While the gradient for the Voronoi polygons rises steadily, for the block segments the gradient tends to flatten off starting from the 65th percentile threshold. This model outcome can be attributed to the overall smaller size of block segments compared to the Voronoi polygons which has led to producing smaller more homogeneous zones.

Figure 4-6a. and b. illustrate the average and standard deviation of the zones size for the DACZ structures (y-axis) plotted against the aggregation heterogeneity threshold percentiles (x-axis). The average surface area graph (Figure 4-6a.) shows an overall increase in the zones size in response to shifting the heterogeneity threshold to higher levels. Also, the DACZs derived from the Voronoi polygons constantly have larger average sizes compared to the block segments, which matches prior expectations given the larger average size of the base Voronoi polygons. Moreover, impacts of the different area cap parameters (i.e., multiplier indicating different levels of spatial adjacency) can be clearly identified, where applying two degrees of spatial adjacency has perpetually resulted in larger zones compared with those DACZs derived from applying one degree of spatial adjacency.

However, as Figure b. illustrates, for the different base structures the zone size distributions are indicative of a diverging pattern. While the DACZ structures developed based on the block segments have average sizes ranging from 1 to 9 hectares and standard deviation between 2 to 10 hectares, corresponding figures for the Voronoi polygons show a wider variability ranging from 5 to 30 hectares for the average size and from 4 to 32 hectares for the standard deviation respectively.



(a) Average of zones area (b) Standard deviation of zones area

Figure 4-6 - DACZ model sensitivity analysis results (average and std. dev. of zones area)

This investigation demonstrates consistency and reliability of the DACZ model in producing output zonal structure as a (predictable) function of aggregation heterogeneity threshold, different levels of adjacency, and the original spatial disaggregation methods (i.e., Voronoi vs block segments). The results prove the robustness of DACZ model as initially proposed, and provide future studies with information on how the outcomes might change if we make different initial assumptions.

4.2.3 DACZ Model and the Graph-based Spatial Partitioning Method Comparison

As discussed in section 3.8, among a broad range of spatial regionalization methods, the graph-based spatial partitioning model was implemented to provide the baseline measure for evaluation of the DACZ model performance. The graph-based model offers a unique feature in that it partitions the geographic space into a number of homogeneous regions, where this number is an exogenous parameter to the model. This feature was essential for a pairwise analysis of a set of previously created DACZ structures against a comparable set of structures generated by a well-established method. Figure 4-7 shows the results of this pairwise comparison in terms of the average heterogeneity levels of the produced zones (y-axis) for both DACZ and graph-based methods given certain numbers of zones (x-axis). These numbers are endogenous derivatives of the DACZ model and exogenous parameter to the graph-based model. In this figure, the yellow line indicates values related to the DACZ structures and the green dotted line presents values for the graph-based clusters. As the graph shows, for the DACZ model growing the number of the zones from 336 to 3056 decreases the average normalized entropy score steadily from 0.35 to nearly 0.10.

This trend indicates that the number of generated zones is inversely correlated with the average normalized entropy score, meaning that the DACZ model is able to flexibly produce larger numbers of more homogeneous zones or alternatively generates fewer numbers of heterogeneous zones. This outcome in fact proves the intended model behavior in generating zones for which the size is defined as a function of homogeneity and, as a result, these zones will generate more predictable outcomes in travel behavior modeling and analysis.

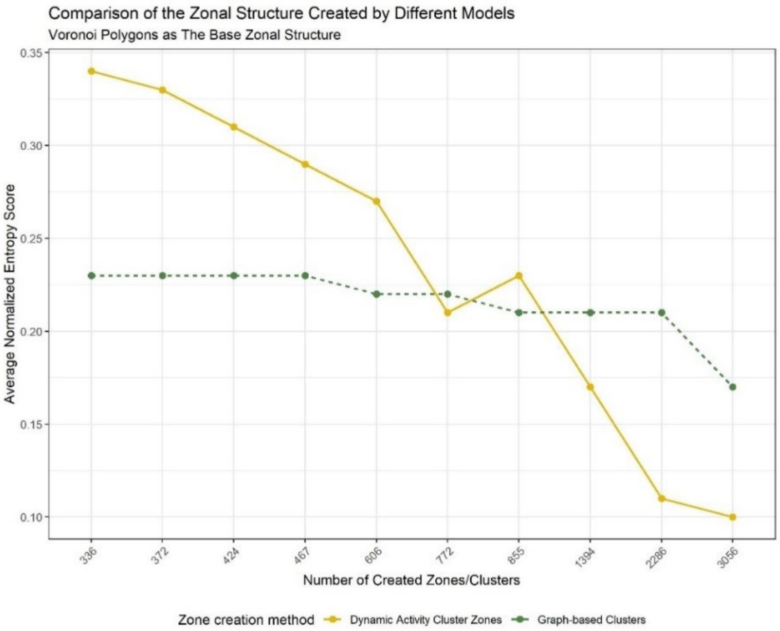
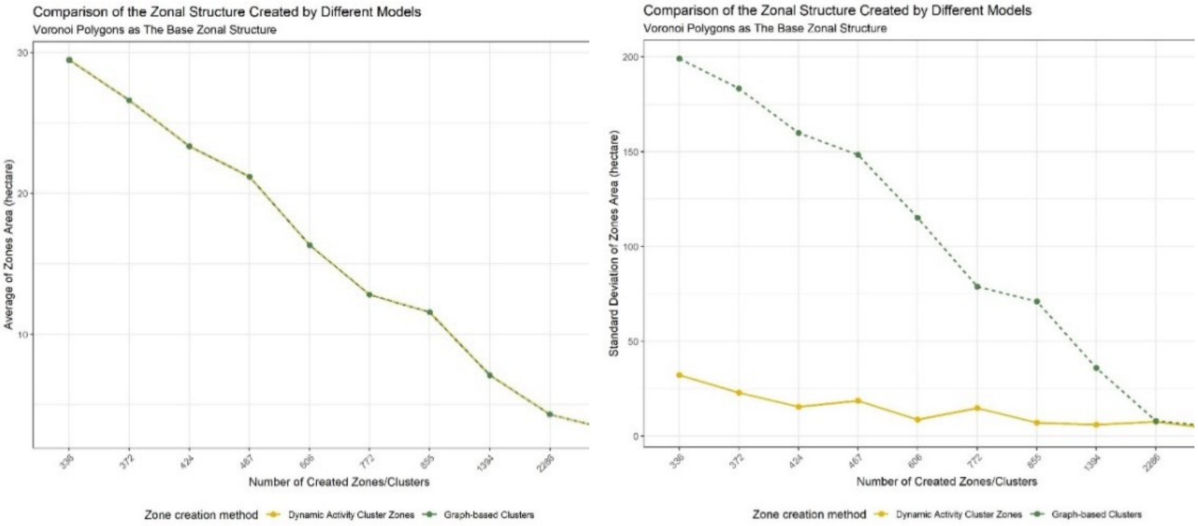


Figure 4-7 - DACZ and graph-based methods results comparison (average entropy score)

In contrast, as the green dotted line shows, for the graph-based model applied over the same study area having the exact same number of zones, the average normalized entropy scores are almost unchanged throughout the observed range. This can be indicative of significantly different spatial arrangement of the zones within graph-based structures compared to their DACZ counterparts.

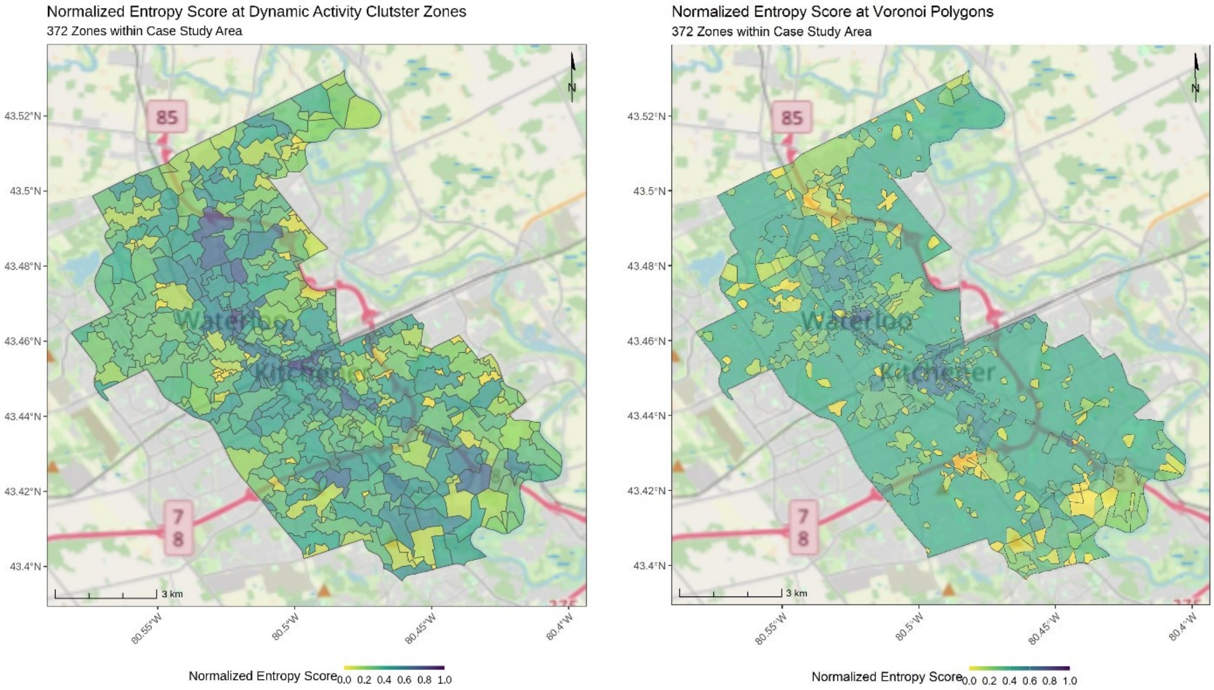
Figure 4-8a. and b. provide some insight on these differences by depicting the average and standard deviation of the zones area (y-axis) plotted against the set of zone numbers (x-axis). Figure a. shows the average size of the zones ranges from 5 to 30 hectares and since both models are compared against the same set of zone numbers no differences exist between them. However, the standard deviation of the zones size reveals an interesting pattern where the DACZ structures show minimal variation in their area; the zone sizes are relatively constantly around or below 30 hectares, while the dual-graph clusters suffer from a large variability in the clusters size ranging from 4 to almost 200 hectares. Looking at the spatial distributions of these two zonal structures in Figure 4-9 better highlights the considerable differences in zones formation between the two systems, where the DACZ model consistently demonstrates more effective aggregation.



(a) Average of zones area

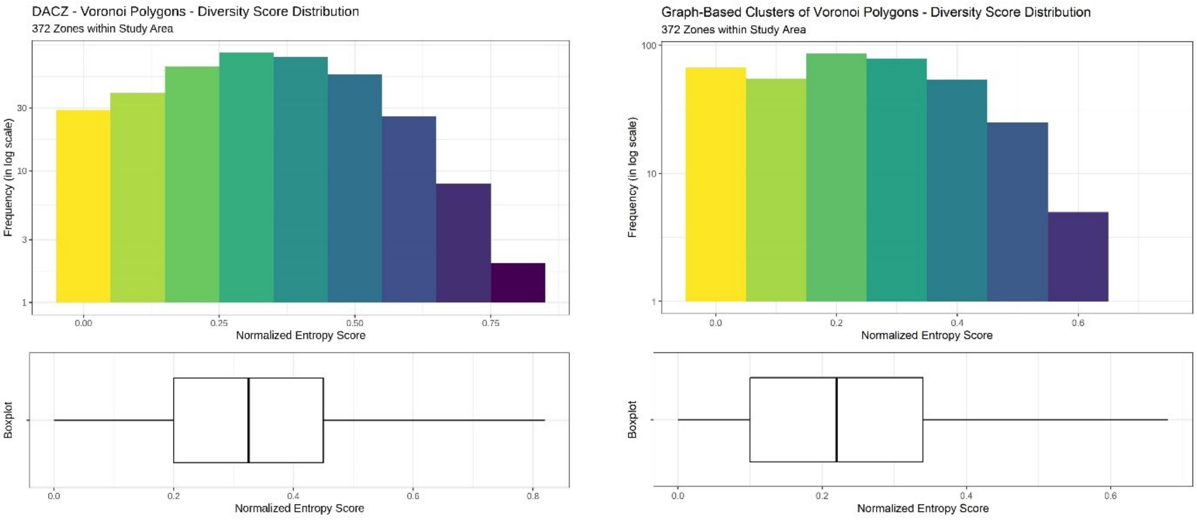
(b) Standard deviation of zones area

Figure 4-8 - DACZ and graph-based methods results comparison (avg. and std. dev. of areas)



(a) Heterogeneity score within DACZ

(b) Heterogeneity score within graph-based clusters



(c) Distribution of heterogeneity scores for dynamic zones (in log scale)

(d) Distribution of heterogeneity scores for graph-based clusters (in log scale)

Figure 4-9 - Spatial and statistical distributions of the generated zones and heterogeneity scores for DACZ and graph-based structures (aggregation threshold: 85th, two degrees of adjacency)

4.3 Travel Behavior Heterogeneity Distributions

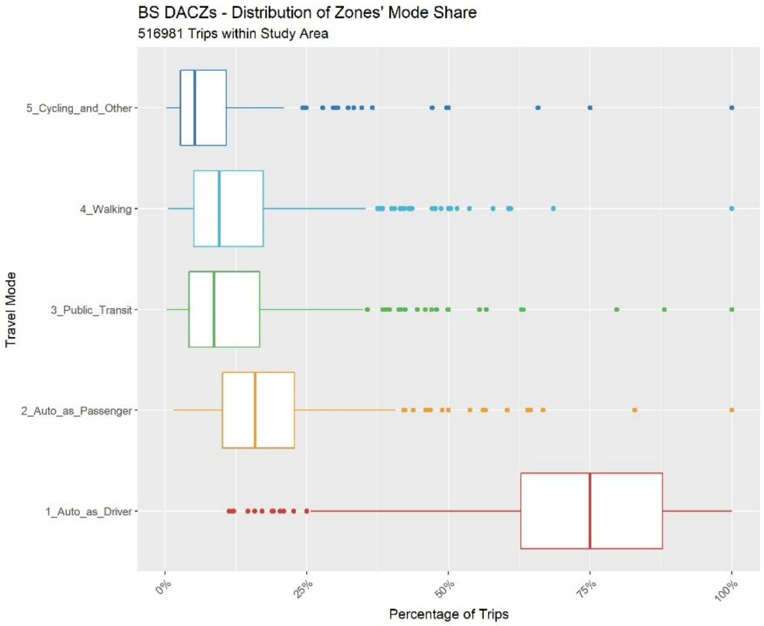
The subsequent sections describe the results of travel behavior analyses at three spatial scales for the observed travel data from TTS 2016. At each level a set of travel behavior metrics are computed for both DACZs and the conventional PLUM TAZs and the results are compared.

4.3.1 Mode Share Distribution Comparison between DACZs versus TAZs

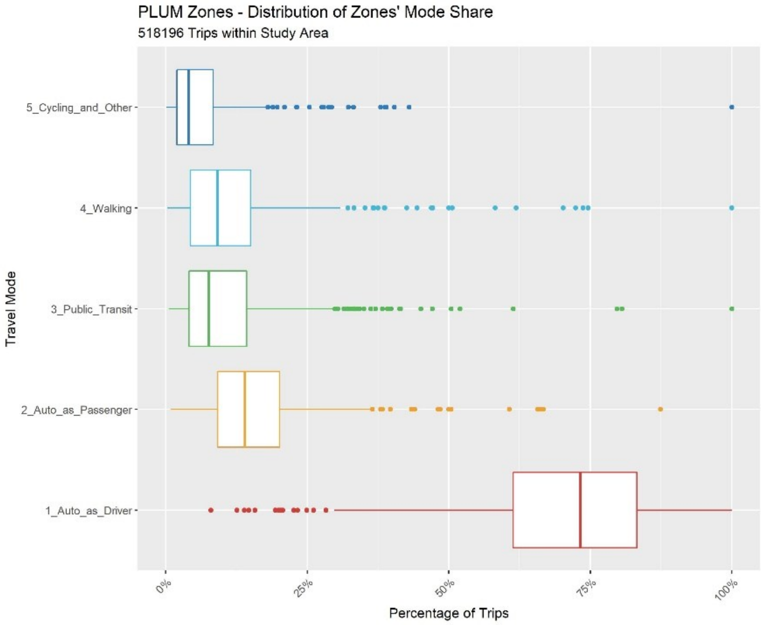
As a first comparison, regional travel mode shares were computed in terms of the overall proportions of trips made by different transportation modes. Figure A--5-7a. illustrates this aggregate mode share that is dominated by auto trips (82%) with most trips made by drivers (68%) followed by trips made as passengers (14%). In third place, public transit made up less than one tenth of the trips (8.3%) and active transportation modes combined come close to one tenth (9.2%) of total trips. There is a minor difference in the total number of the trips between the two zonal structures that results in a slightly higher number of trips (0.2%) for the PLUM zones. This is due to the inconsistency of study area boundary which was originally defined based on the Census Tracts.

While it makes sense for both zonal systems to reveal the same mode share levels at aggregate scale, as the DACZs and the PLUM zones differ in their average zone sizes and homogeneity levels, there is an expectation to observe different distributions of mode shares within each system. In line with this assumption, the zonal structure with greater intra-zonal homogeneity – that means greater inter-zonal heterogeneity – should show more consistent travel behavior at zonal level and higher variance in the regional distribution of travel behavior. To assess this expected difference, the distribution of mode shares were calculated for all zones within both systems. Figure 4-10a. and b. show the results of mode share distributions for the two zone systems.

Consistent with the aggregate regional figures, the zone level mode shares also illustrate the dominance of auto trips, where only quarter of the zones have auto mode share below 60%.



(a) Distribution of zones by their travel mode share for DACZs



(b) Distribution of zones by their travel mode share for PLUM Zones

Figure 4-10 - Statistical Distribution of Analysis Zones by Their Mode Share

The most visible characteristic of these mode share distributions for both zone systems is the left-skewed distribution of auto mode (as driver) in contrast to the right-skewed distributions of all other modes. As Figure 4-10a. and b. indicate, for almost half of the zones (i.e., the distribution median) auto mode accounts for at least 75% of their total trips. Moreover, there are only a very few zones where the auto mode share (as driver) falls below 25%. Combining all the auto trips (either as driver or passenger) under one category would result in an even more skewed distribution where auto mode share would rarely drop under 40% across the study area.

The results also imply that at regional level both the DACZs and the PLUM zones perform comparably in representing travel behavior patterns. This contradicted the expectation that the DACZs would reveal substantially different behaviors in terms of mode shares. Hence, in the next step the analysis expanded to incorporate the built environment heterogeneity as a factor that might influence the zonal systems suitability in revealing travel behavior heterogeneity.

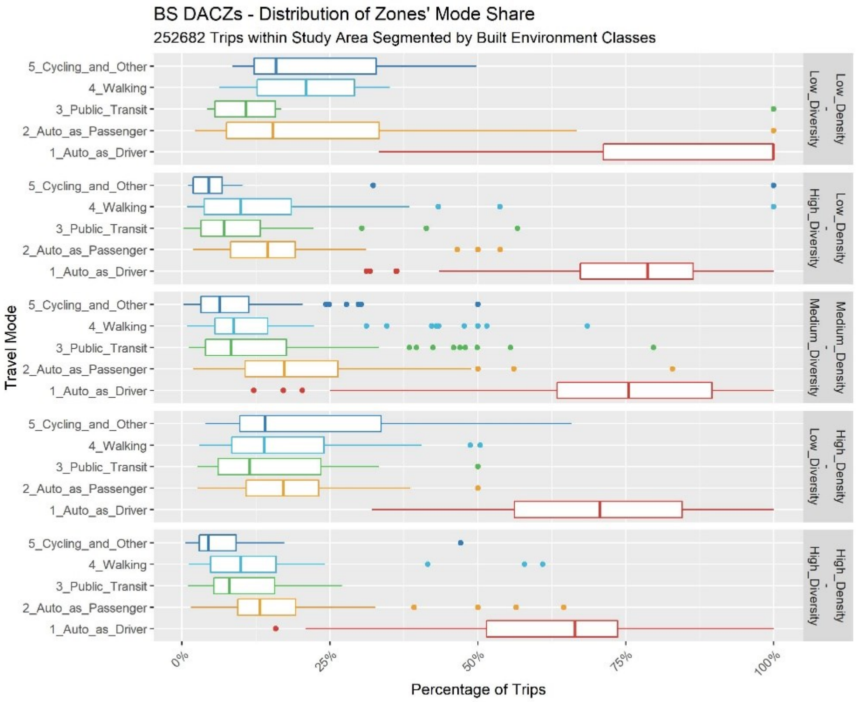
4.3.2 Comparison among Built Environment Classes Categorized According to their Combined Density and Land Use Diversity Levels

The expectations in this analysis were that creating zones with greater homogeneity will capture activities for which transportation behavior will also be more homogeneous. The previous section shows that at the macro level, that assumption did not hold true. In this section, the analysis further compares the performance of zones belonging to one of the nine categories described in Section 3.6.1 – high, medium and low density and diversity.

Figure 4-11 shows the mode share distributions results considering the built environment heterogeneity categories. Because the expectation is that the proposed zonal structure will perform better at the extremes – i.e., at high and low densities and diversities – the presentation of results is simplified to exclude combinations of high (or low) and medium levels of density and diversity from these graphs. The "Medium Density – Medium Diversity" class was preserved as it is the largest identified built environment category that accounts for more than half of the analysis zones (Figure 3-11).

Looking at different built environment classes in Figure 4-11, a constant slight shift in the mode share distributions is recognizable. Increasing population density and land use diversity, generally more trips were made using non-auto modes. This overarching pattern can be quantitatively described by decreasing medians of the auto mode share as we move from “Low Density – Low Diversity” to “High Density – High Diversity” zones. It is also evident that the high levels of land use diversity are potentially a better predictor of the lower auto mode share compared to the high levels of density. In fact, finding high density zones that have low diversity is likely to produce errant results in terms of expected travel behavior. Overall, the observed trend reinforces the land use and transportation system interactions theories which contend that areas with diverse land uses and high levels of population density are more conducive to active transportation and public transit trips compared to more dispersed uniform areas.

However, while the two zonal systems reveal similar pattern, they tend to unfold it differently. For the zones within the “High Density – High Diversity” category both zonal systems show identical shifts in decreasing auto mode share. Within this category, only a quarter of the zones (i.e., the distribution upper quartile) have auto mode shares that account for 75% or higher proportions of their trips. For the mixed categories of “Low Density – High Diversity” and “High Density – Low Diversity”, although similar patterns are visible, variations of the active transportation and public transit mode shares are more pronounced within the DACZ zonal structure compared to the conventional PLUM zones. Ultimately, the forceful car dependence in the “Low Density – Low Diversity” areas is better reflected by the DACZ system, where the auto mode share reaches 100% for almost half of the zones.



(a) Distribution of zones by their travel mode share for DACZs

(b) Distribution of zones by their travel mode share for PLUM Zones

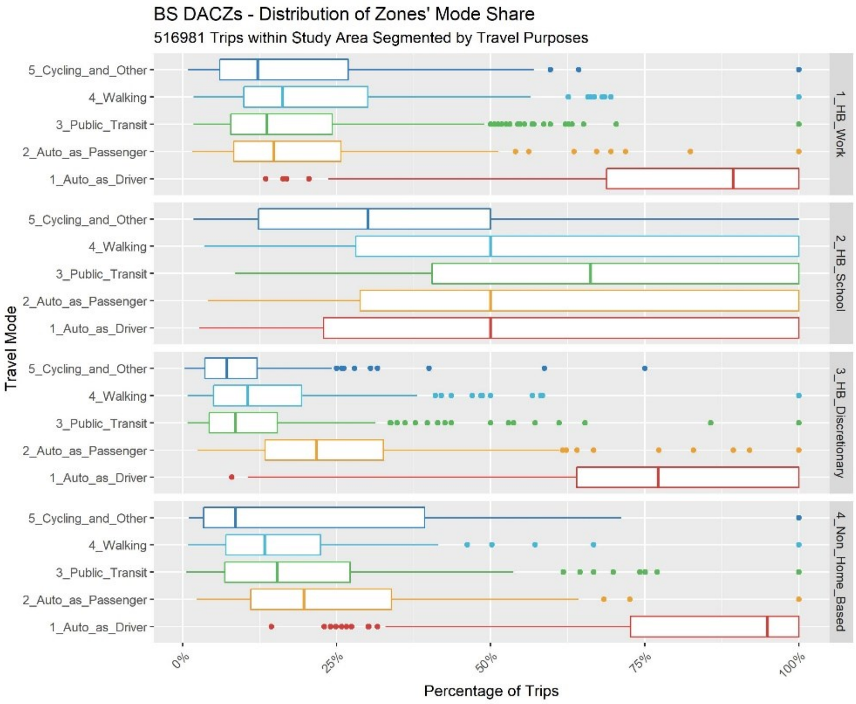
Figure 4-11 - Distribution of Analysis Zones by Their Mode Share Segmented by the Built Environment Heterogeneity (Selected Categories)

4.3.3 Mode Share Distribution by Trip Purpose (DACZs)

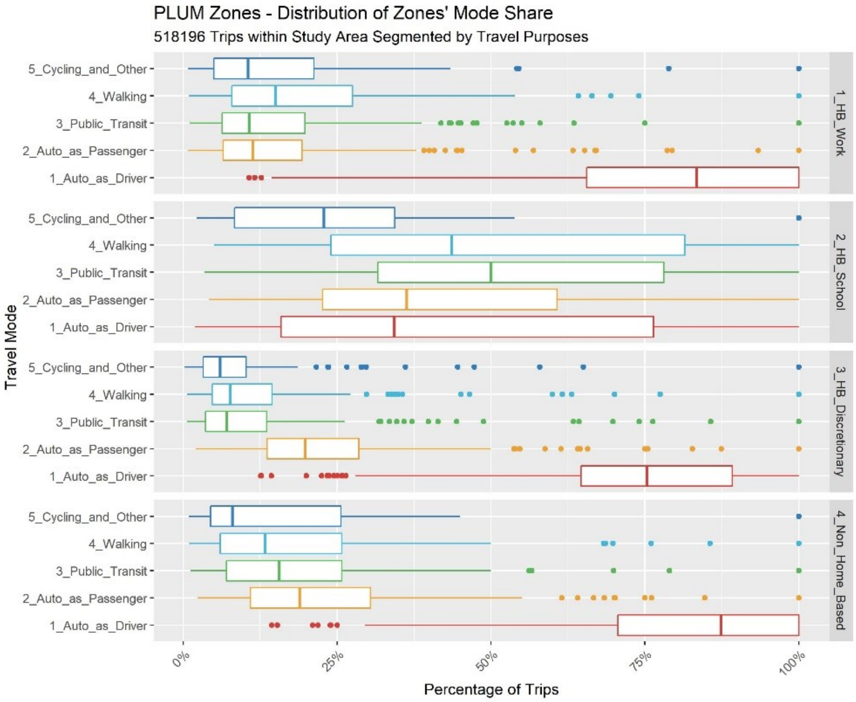
The aggregate mode share analysis indicated that the generated trips within the study area are dominated by auto trips across the heterogonous built environment categories from the "Low-Density Low-Diversity" to "High-Density High-Diversity" zones. Further analysis was performed to examine whether this pattern persists for different trip purposes. In this section, a comparison of mode shares is made not based on diversity or density of land uses, but rather based on trip purposes within the DACZ and PLUM structures. As illustrated in Figure 4-12a. and b. home-based school trips (HBS) are the only category of trips that shows a different and more evenly distributed mode share between auto, public transit and active transportation modes. A potential reason for this distribution can be associated with the specific characteristics of the population group that makes the home-based school trips (HBS). Individuals in this group are more likely in younger age bracket, which restricts their access to driving license and auto vehicles – major factors that can considerably impact their mode choice. However, it is important to note that from a total of more than a half million generated daily trips in the study area, 30.6% are home-based work trips (HBW), 40.8% home-based discretionary trips (HBD), 17.9% non-home based (NHB) trips and just a small fraction (around 10.6%) are home-based school trips (HBS).

The results again indicate that both the DACZs and the PLUM zones perform comparably in representing travel behavior patterns, and especially the DACZs have not demonstrated substantially different results. Based on the lesson learned from previous section, the expectations in this analysis were that the DACZ structure with greater intra-zonal homogeneity or interchangeably greater inter-zonal heterogeneity should show higher variance in the regional distribution of travel behavior. This assumption held true for most of the travel modes and trip purposes, but it was not fulfilled for the home-based work trips (HBW).

The detailed statistics of trip purpose distribution for all generated trips aggregated within the study area, and with distinction by travel mode are provided in Appendix D.



(a) Distributions for DACZs



(b) Distributions for PLUM Zones

Figure 4-12 - Distribution of Analysis Zones by Their Mode Share Segmented by Trip Purposes

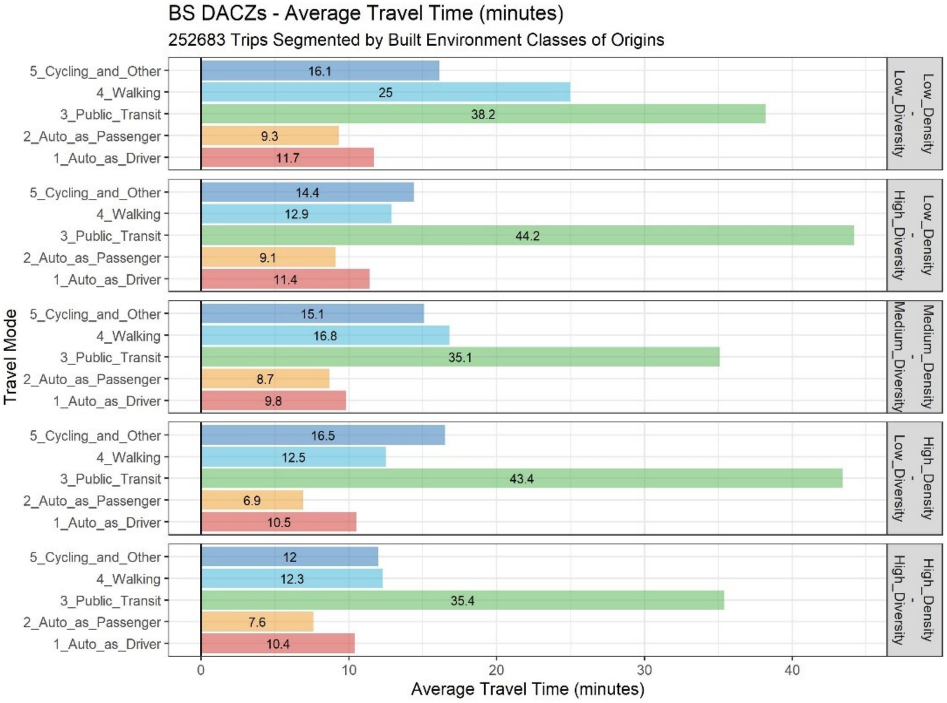
4.3.4 Travel Time Distribution by Built Environment Classes

One way that planners evaluate travel behavior patterns at the macro level is to analyze travel time statistics. In particular, the average travel time is a metric that reflects the overall transportation system performance and imply the levels of access to spatially distributed activities. To complement our analyses, we conducted an additional assessment on the distribution of the travel times for the trips originated within the cities of Waterloo and Kitchener.

The expectations in this investigation were that the built environment heterogeneity and density levels impact the trips characteristics considerably. We focused only on those trips that have their origins identified within the dynamic zones, for which we were able to compute the heterogeneity and density levels. Again, we simplified the results to exclude combinations of high (or low) and medium levels of density and diversity from the presented graphs as described in Section 4.3.2.

Figure 4-13a. shows the average travel time (ATT) distributions in minutes, given the built environment heterogeneity categories for the trips origins. Figure 4-13b. represents the differences of the ATTs for each mode compared to its regional level, computed as percentages. Looking at disparate built environment classes, a consistent pattern of travel times is recognizable where the travel modes determine the average trip duration. Across the categories, transit trips show the longest ATT, ranging from 35 in “High Density – High Diversity” to 45 minutes in the mixed categories. In contrast, auto trips, either as driver or passenger have the shortest ATT between 10 to 12 minutes, where the longer trips associated with the “Low Density” zones. Similarly, walking and cycling show a variation among classes in which the longer trips are more pronounced for the “Low Density – Low Diversity” zones.

Interestingly, these findings show a high degree of agreement with the expected travel behavior. They also provide some insights on the observed imbalanced mode share in earlier analysis. For example an explanation for the lower walking mode share, particularly in areas with lower levels of density and diversity, can be justified with the longer average walking time and the extra burdens which potentially preclude larger proportion of trips to be made by walking.



(a) Average travel time (ATT) in minutes

(b) Difference of average travel time for each mode compared to the regional average

Figure 4-13 - Distribution of Average Travel Times (ATT) in minutes for Generated Trips, Segmented by the Built Environment Heterogeneity (Selected Categories)

4.4 Spatial-Temporal Patterns of Regional Land Use Development

Recall that a research question to be answered in this thesis is an exploration of how development patterns in a metropolitan region are (or are not) correlated spatially with the presence (or absence) of land use diversity. Similarly, the research seeks to understand how the introduction of proposed and actual rapid transit investments are correlated to development. To address these questions, this section presents the results of an analysis of the historical building permit data in the Region of Waterloo over a 20-year period, from 2000 to 2019, within and outside the Region's Central Transit Corridor (CTC).

The original building permits dataset consisted of a very wide range of construction values ranging from zero to tens of millions CAD$, a monetary lower bound of 20,000 CAD$ (after adjusting for the inflation rate) was applied to remove those records associated with minor building modifications. To further "clean" the data, erroneous observations and duplicate records were detected manually and removed. This data preprocessing resulted in a set of 19,946 items for the entire period of study. In terms of the building permit types, after resolving the classification inconsistencies, the final dataset comprised seven different permit types within the cities of Kitchener and Waterloo, including: Residential (52.3%), Multi-unit Residential (12.6%), Mixed-use (0.6%), Commercial (1.9%), Institutional (1.2%), Industrial (1.5%), Other Non-residential (29.9%) permit types.

Figure 4-14 depicts the spatial distribution of issued building permits cumulatively. It presents all available data points covering the cities of Waterloo, Kitchener and Cambridge. The color gradient in this map represents the intensity of the developments through time. The relative intensity measure was computed using kernel density method considering both the construction value and the proximity of surrounding issued building permits in each location. In this map areas in brighter colors show greater concentration of higher-valued constructions. Therefore, the aggregate spatial pattern is indicative of development agglomeration along the CTC, especially to the north of the Uptown Waterloo. Following sections investigate the temporal patterns of these developments across multiple spatial scales including the entire study area (within the cities of Kitchener and Waterloo), relative to the CTC and within different heterogeneous built environment categories identified in Section 3.6.1.

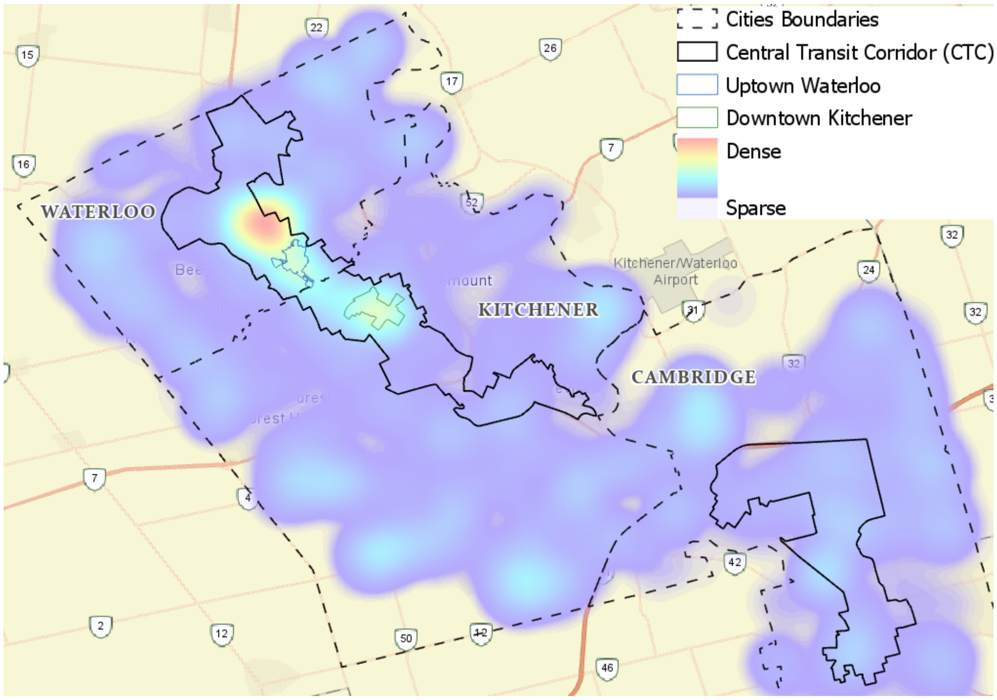


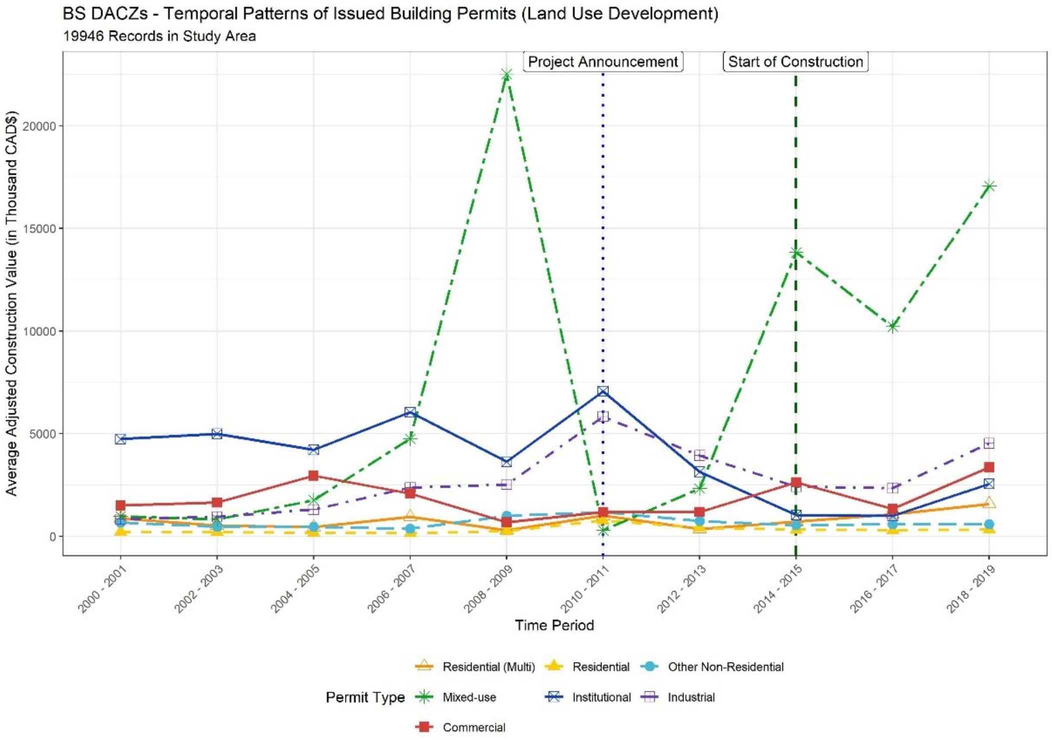
Figure 4-14 - Spatial Distribution of the Cumulative Issued Building Permits (from 2000 to 2019)

4.4.1 Temporal Trends for the Different Building Permit Types

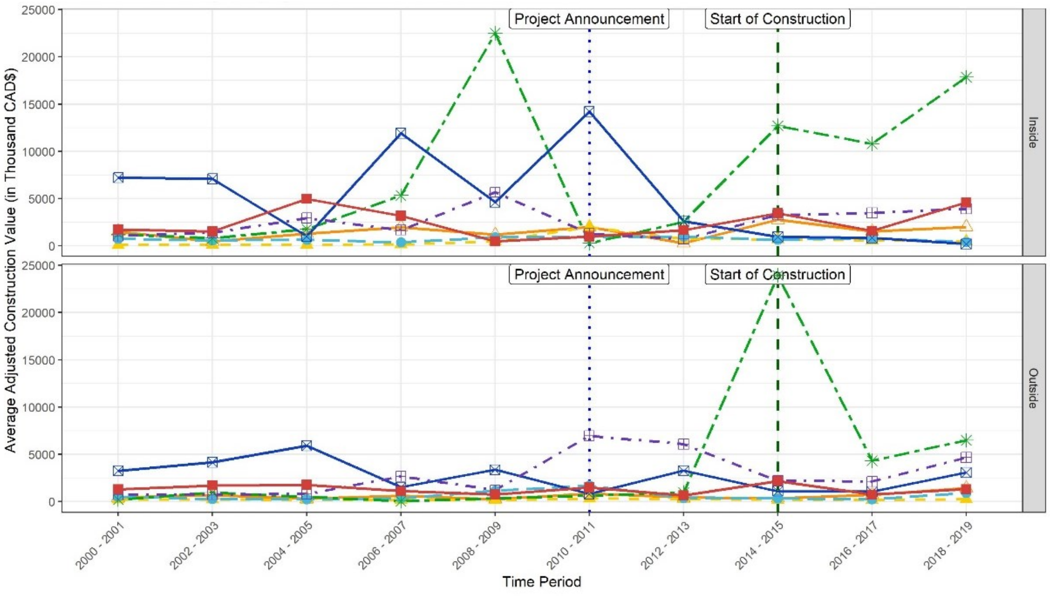
Figure 4-15 depicts the temporal changes in the average adjusted construction values (y-axis) for different types of building permits – denoted by different line types and colors.

Figure 4-15a. provides an aggregate overview in which for every point in time (x-axis) the regional statistical average of the permits' adjusted construction value (in thousand CAD$) are independently calculated for each building permit type (e.g., residential or commercial). Figure 4-15b. offers a more disaggregate view of the same data by incorporating geography of the Region's Central Transit Corridor (CTC). This figure consists of two panels; the top one shows the average values calculated for the areas within the CTC, and the bottom panel demonstrates the average values for the areas outside of the CTC. Both panels use the same scale range on the vertical axis.

In a similarly organized visualization, Figure 4-16a. and b. illustrate the temporal trend of the total construction values (in million CAD$) for the entire study area, and relative to the CTC geography, respectively.



(a) Average Adjusted Construction Value (AACV) for the entire study area



(b) Average Adjusted Construction Value (AACV) relative to the CTC geography

Figure 4-15 - Temporal Distribution of Building Permits Average Construction Value 2000-2019 (Segmented by the building permit type)

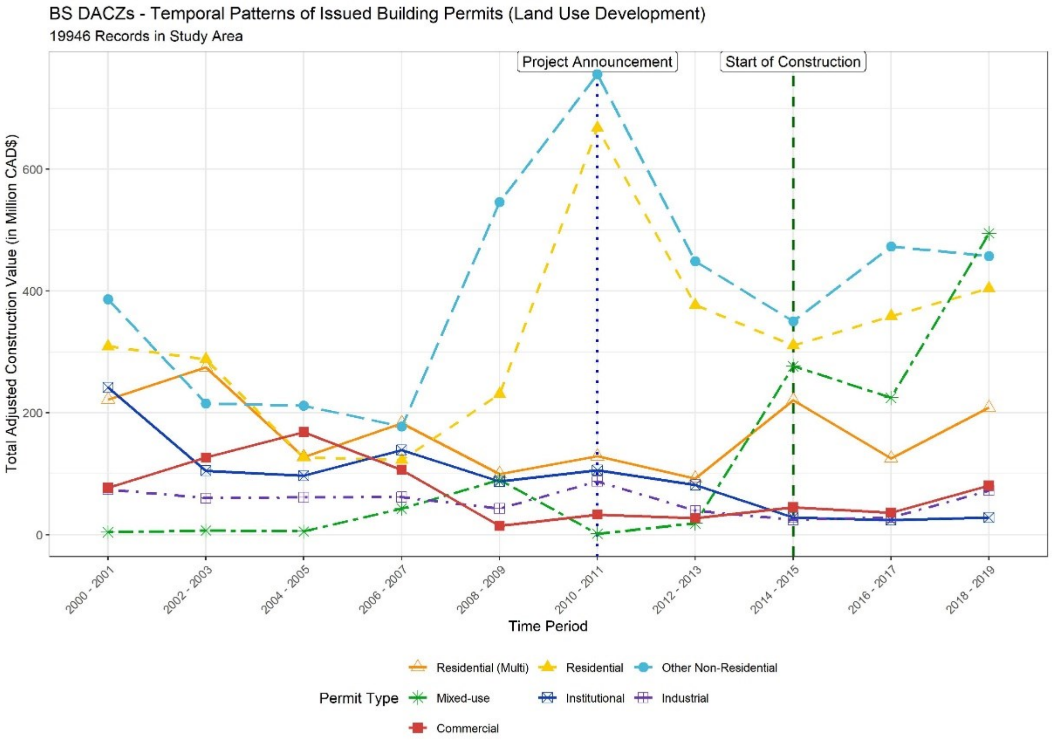
A primary observation from Figure 4-15a. is that most of the permit types – except the institutional (the dark blue line) and mixed-use buildings (the dashed green line) – display a relatively stable range of average adjusted construction values (AACV) through time. Residential permits (specifically single unit buildings) exhibit the lowest AACV among all the permit types, while the highest AACV at each time point, has interchanged between the mixed-use and institutional building permits during the study period. However, a closer look at individual time points is indicative of a (vertically) scattered pattern of the AAVC during each biannual period. This cross-sectional variation in AAVC reveals the differences between the monetary scales of developments for different permit types.

For instance while the differences between the residential AACV (the yellow line) and other non-residential AACV (the light blue line) is almost constantly negligible, its differences with the mixed-use permits AACV (the dashed green line) reaches 20-fold in 2008-2009 and abruptly vanishes in the subsequent period. Although the differences between the permits' average construction values are not unexpected given the nature of various land uses and their construction specifications, the divergent pattern of the AACVs towards the end of study period warrants further exploration.

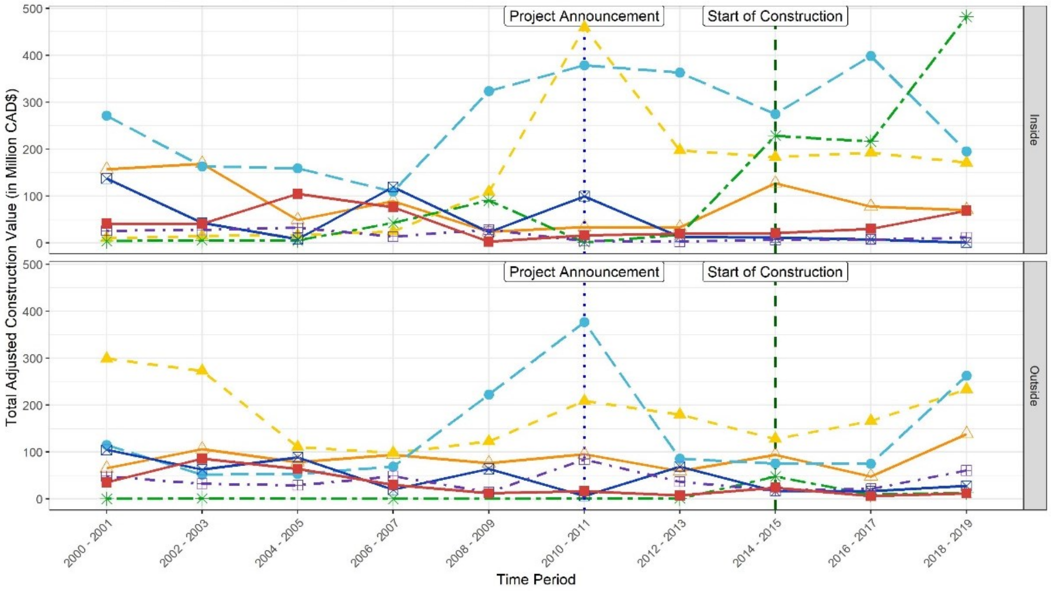
As Figure 4-15b. shows, for almost all permit types, areas within the CTC have been the major driver for pushing average construction values higher. This pattern has only one exception in that the industrial permit type (the dashed purple line) displays higher AACV outside of the CTC for the second half of the study period. This might be due to the Region's re-urbanization policies and its zoning regulation that was put in place in earlier periods. Theoretically, this trend also makes sense given that the industrial development requires larger land parcels on average, while in fact the larger and cheaper land within the CTC has been becoming scarce when competing uses such as residential and commercial developments have tendency to pay higher prices for the premium central locations within the CTC.

In addition to the average construction values trends, the total adjusted construction values (TACV) metric is expected to provide further insights to reveal the development patterns in the region. Figure 4-16a. and b., present the temporal trend of total adjusted construction values (y-axis) for different permit types. In these graphs for every point in time (x-axis) the TACVs are accumulated and presented in million CAD$.

As compared to the AACVs trends, the TACVs patterns are more scattered and appeared in different order, in terms of the relative magnitude of the TACVs for different permit types. Here, single unit residential (the dashed yellow line) and other non-residential (the dashed blue line) categories made up the largest proportions of the issued permits for the most part. They experienced sharp increases prior the announcement of the LRT project and continued to attract great amount of development afterwards. This pattern is replicated at both scales: at the entire study area and within and outside the CTC geography. However, the mixed-use permit type (the dashed green line) shows an interesting trajectory that was mainly driven by development within the CTC area. The regional mixed-use developments started with negligible values in the early 2000s. It continued at modest levels, almost lower than all other permit types until 2013. In 2014, concurrent with the construction of the LRT project, a significant increase in the TACV for mixed-use permits increased their share to the third largest development type in the region. Despite a slight decrease in the associated TACV in the subsequent period, the strong tendency for new mixed-use development particularly within the CTC has made it the largest category of issued permit type in 2019.



(a) Total Adjusted Construction Value (TACV) for the entire study area



(b) Total Adjusted Construction Value (TACV) relative to the CTC geography

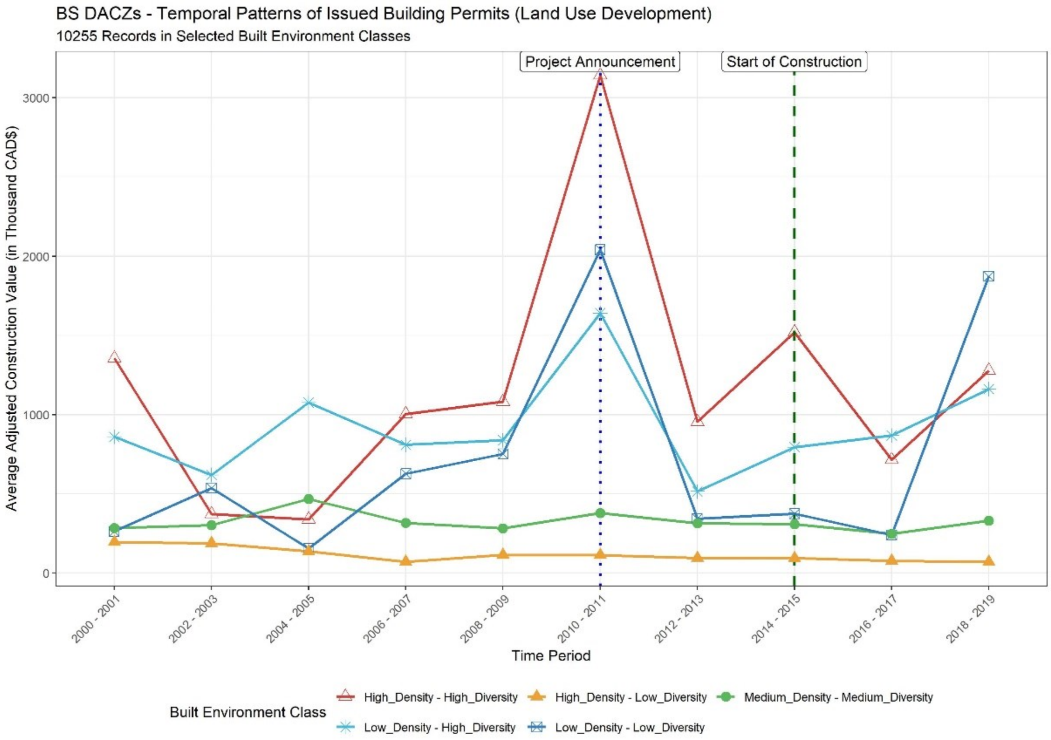
Figure 4-16 - Temporal Distribution of Building Permits Total Construction Values 2000-2019 (Segmented by the building permit type)

4.4.2 Temporal Trends within the Different Built Environment Categories

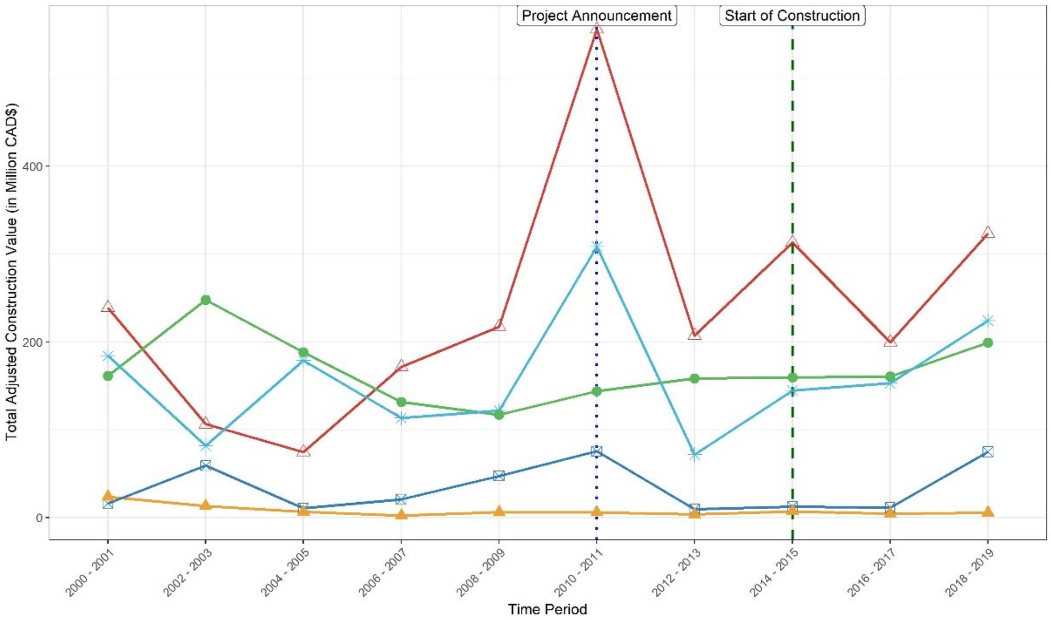
This section assesses building permit data from a different perspective. Figure 4-17a. and b. illustrate the trend of changes in the average and the total adjusted construction values of the issued permits (y-axis) respectively. These summary metrics were calculated based on the building permits records grouped by their location within the heterogeneous built environment categories identified earlier, from the “Low-Density Low-Diversity” to “High-Density High-Diversity” zones. Recall that the built environment groups presented in this section primarily defined based on the upper and lower bound (i.e., 25th percentiles) of the density and diversity levels distributions. It is also important to note that no distinction was made between the types of permits in this analysis.

As with the previous analysis, this review concentrates on extreme observations of land uses; as such, among the nine identified built environment categories, combined high / low and medium results are not presented in these graphs. This focus on certain segments of the heterogeneous built environment reduces the number of building permit records to 10,255 points, located within the five major categories. These categories are denoted by different line colors in the following figures.

Looking at the trends of changes in the average adjusted construction values (AACV) in Figure 4-17a., a few patterns are evident. The first is the sharp spike in AACVs in certain built environment categories coinciding with announcement of the LRT project in 2010-2011. This significant rise in AACVs is not observed for the “High Density – Low Diversity” and “Medium Density – Medium Diversity” areas. These two categories possess the lowest AACVs with relatively stable value around 300,000 CAD$ throughout the study period. In contrast, those areas identified as “High Diversity” have predominantly enjoyed the higher levels of AACV, and constantly stayed among the top 3 highest average permits values. However, the trend of AACVs changes indicates a noticeable decline following their peak in 2010-2011.



(a) Average Adjusted Construction Values (AACV) within selected built environment categories



(b) Total Adjusted Construction Values (TACV) within selected built environment categories

Figure 4-17 - Temporal Distribution of Building Permits Construction Values 2000-2019 (Aggregated by the selected built environment categories)

Since the average values presented here might be impacted by the presence of outliers, and the fact that an uneven number of observations within certain built environment groups could bias the AACVs, the next analysis was performed with total adjusted construction values (TACV), for the heterogeneous built environment categories.

Figure 4-17b. depicts the results of this analysis where, interestingly, similar patterns unfold. A prevailing higher TACVs are more pronounced for the “High Density – High Diversity” areas throughout the study period, except between 2002 and 2005. These areas also accommodate the development projects with the highest overall TACV, based on the permits that were issued during 2010-2011.

More importantly, it is evident that areas classified as “Low Diversity” have experienced the lowest total value of the issued permits through time. The two isolated lines (in yellow and dark blue) at the bottom of this figure highlight the significant differences in magnitudes of developments taking place within areas that have uniform land uses compared to those with more diverse built environment (the red and light blue lines). These results also indicate that within the studied heterogeneous metropolitan region, high density has not been sufficient, nor necessary for attracting greater scale of land use developments.

4.5 Discussion

This research has developed a homogeneity-based spatial aggregation method to create zonal structures for land use and transport system interactions (LUTI) studies. This spatial modeling approach was applied to a mid-size metropolitan region to identify and analyze areas with similar travel behavior profiles based on the composition of land uses within those areas.

The expectation was that one would observe different performance in representing the built environment and travel behavior distributions, between the proposed dynamic structure (i.e., DACZ structure) and the conventional TTS TAZ and PLUM zones. In line with this initial assumption, the DACZ structure with greater intra-zonal homogeneity – meaning greater inter-zonal heterogeneity – was anticipated to show more consistent built environment and travel behavior characteristics at zonal level, and higher variance in their regional distributions respectively.

This study benefited from access to fine-scale land use and travel behavior data for the urbanized areas of the Region of Waterloo, covering an approximately three kilometer buffer around the LRT alignment within the cities of Waterloo and Kitchener. This multidimensional data repository enabled this research to examine the performance of the implemented DACZ model in a range of spatial scales. In addition to that, the DACZ model was used to analyze historical trends of land use development illustrating the applicability of the DACZ for additional analyses, not directly related to transportation. The temporal assessment of the DACZ model in terms of how do land use changes get operationalized and reconciled through multiple time periods, remains to be explored in future research.

This dissertation sought to advance understanding of the impacts caused by the choice of analysis zones on the travel behavior and land use development analysis outcomes. It enhanced the methods for land use homogeneity measurement and spatial aggregation modeling and introduced a dynamic zonal structure with a range of applications in planning research and practice, particularly in transportation planning. The findings of this study offer implications for land use and transportation system interactions theories, and contribute to informing more integrated land use and transportation policies.

The theoretical, methodological and planning implications derived from this study are discussed in the subsequent sections.

4.5.1 Theoretical Implications

This research built upon the concepts that suggest appropriate spatial structures for LUTI analysis should represent homogeneous land use areas, in addition to the more established criteria such as densities and size attributes. In particular, the DACZ model which was developed and evaluated in this study enhances previous methods suggested by Casello and Smith (2006) to construct urban activity centers, and extends the iterative spatial aggregation method developed by Casello and Fard (2017) to identify the mega zones that potentially produce and attract regionally significant number of trips.

The application of the DACZ model to the Region of Waterloo revealed some interesting outcomes that reinforce the land use and transport system interactions theories to a large extent. While the multilevel travel behavior analysis showed that auto trips are dominant all over the region, the zone level mode shares displayed a trend in which increasing the levels of population density and land use diversity generate greater proportions of trips made using non-auto modes. This observation is consistent with the theories which contend that areas with diverse land uses and high levels of population density are more conducive to active transportation and public transit trips.

Moreover, accounting for travel purposes in the region, the auto mode dominance was also observed for different travel purposes, though the home-based school (HBS) trips showed a different and more evenly distributed mode share between auto, public transit and active transportation modes. The specific characteristics of the population group that makes HBS trips perceived to be the potential reason for this distribution. Individuals in this group are likely in the younger age bracket that restricts their access to driving license and vehicles – major factors that considerably impact their mode choice.

This work also uncovers questions on how the co-location (e.g., within the central districts) or isolation (e.g., within satellite areas) of the zones with similar levels of land use diversity and density might influence transportation function and travel behavior for different areas across a metropolitan region. These outcomes emphasize the importance of demographic, socioeconomic and locational characteristics as determinant factors in LUTI dynamics that should be addressed in future studies.

In terms of the land use development trends, the historical building permit data during the two periods before and after the LRT project funding announcement were analyzed at biannual intervals. The temporal analysis of changes revealed the patterns confirming the role of land use diversity and access to rapid transit on the magnitude of developments. The results indicated that for almost all permit types, areas within the Regions' Central Transit Corridor have been the major driver for pushing the adjusted average construction values higher through time. Likewise, higher levels of land use diversity were found to be associated with disproportionately more frequent and higher-valued land use developments (i.e., higher construction values) among different built environment categories.

In summary, findings from both the land use and travel behavior analyses suggest that the high levels of land use diversity are potentially a better predictor of the lower auto mode share compared to the high levels of density. Similarly, higher land use diversity is found to better explain the likelihood of attracting greater scale of developments, while the higher density levels shown less explanatory power. However, despite the identified spatial variations and temporal trends over the selected set of land use and transportation variables, the results of this research suggest that across Waterloo Region, levels of land use diversity and population density are likely not yet high enough to produce significant changes in travel behavior patterns, and to form highly intensified urban districts.

4.5.2 Methodological Implications

The research compared the dynamic zones generated and zones resulting from two conventional structures (i.e., TTS TAZ and PLUM zones) as well as a state-of-the-art graph-based partitioning.

The results suggest that the homogeneity-based spatial aggregation approach (i.e., DACZ model) provides a potentially better solution to define aerial units for LUTI analysis and modeling, as it often outperforms the conventional methods. The results also confirm the significance of the built environment homogeneity consideration for creating the analysis zones. While the literature review revealed that homogeneity criterion is often overlooked in delineating analysis zones in urban and transportation studies, this research was able to quantitatively demonstrate the problematic results of such practices. In addition, the conducted comparisons revealed that the commonly used density criterion can be an insufficient metric to use as the primary measure for defining zonal structure.

At an aggregate level, the sensitivity analysis presented with regard to the DACZ model indicates that the number of generated dynamic zones is inversely correlated with the average normalized entropy score, meaning that the DACZ model is able to flexibly produce larger numbers of more homogeneous zones or, alternatively, generate fewer more heterogeneous zones. This outcome proves the DACZ model's main advantage in generating zones for which the size is defined as a function of homogeneity and, as a result, these zones will generate more predictable outcomes in travel behavior modeling and analysis.

The application of the DACZ model to Waterloo Region also demonstrated the model’s effectiveness for application to travel behavior and land use development studies. In this specific example, an enhanced land use diversity measurement approach was implemented to compute heterogeneity scores for block segments within a metropolitan area. These scores were then evaluated against a set of heterogeneity thresholds under different spatial aggregation scenarios. The model created a relatively similar number of dynamic zones to the number of conventional TAZs within the same area. However, both qualitative and quantitative comparisons between these two zonal structures highlighted instances where the conventional TAZs perform less effectively in representing homogeneous land use areas. In all the presented scenarios, considerable homogeneity improvements were observed; more specifically, the average heterogeneity index values within dynamic zones were typically 50 percent lower than their levels within the conventional zones.

In addition to creating more suitable zonal structures for transportation analysis, the homogeneity-based approach offers an advantage in that it avoids relying on existing transportation infrastructure as boundaries. As Jung and Casello (2019) pointed out, the use of roadways as boundaries often causes problems when modeling travelers’ behavior from opposite sides of a same street. It even becomes more problematic when, for example, attempting to designate origin zones to travelers from different corners of an intersection. The DACZ model has addressed these issues by encapsulating roadway intersections and their associated parcels as major elements of zonal definition. Importantly, this research suggests that using entropy-based heterogeneity measurement to ensure homogeneity of the derived zonal structure is a substantial improvement upon the conventional TAZs creation methods.

In tandem with the conventional zones, the homogeneity-based spatial aggregation method also proved to be more effective in producing dynamic zonal structures compared to the modern graph-based spatial partitioning method in terms of variability of the zones sizes.

In summary, despite the limitations of the DACZ model most notably the exclusion of the design elements, the model shown to be effective in generating zones for which the size is defined as a function of homogeneity, an ability that can benefit transportation analysis and modeling in better representing spatial heterogeneity, especially compared to the conventional zones examined in this study. Finally, while the broad range of conducted evaluations have emphasized the robustness of the DACZ method for LUTI analysis, the modest variability of the observed characteristics in Waterloo Region can be attributed to the Region's particular urban context including moderate land use diversity and density distributions.

4.5.3 Planning Implications

By conducting a case study at the Region of Waterloo, this dissertation provides the planning practice with new insights on travel behavior distributions and development trends in a mid-size metropolitan region, which has been experiencing dynamic land use and transportation changes.

The Region's investment in expanding its rapid transit network, specifically the development of the LRT system, had as an objective *moving people more efficiently* and *building thriving communities*. As part of the official planning evaluation practice, the Region initiated the Central Transit Corridor Monitoring Program in 2015 to quantitatively assess the ongoing progress towards the foreseen objectives. The program comprised a set of 16 indicators to be measured and reported at annual intervals, primarily to provide *“a lens for monitoring the achievement of the goals of growth, density and vibrancy in the corridor”* (Parkin et al., 2016). Pertaining to this study, the Region has recognized the importance of land use diversity indicator, and introduced a measure which reflects the number of land uses that are found in the CTC as a percentage of all land uses observed within the region.

We argue that the methods developed in our study to measure land use diversity and to define the spatial zonal structure would better serve the monitoring program in Waterloo Region, and the planning practice in general. The Region’s implemented index is in fact a simplified mixed-use metric that was used as a proxy measure to capture land use diversity levels. Although this simple metric might be easier to communicate, it does not provide a consistent and generalizable method to convey the intended diversity concept, and to compare the diversity levels over time and across different locations. Similarly, the Region’s approach in computing a single index value for the entire CTC geography is evidently insufficient to describe the land use heterogeneity observed across the CTC.

In addition, one way that integrated planning practice evaluates land use and transportation system interactions is to correlate the land use characteristics with travel behavior patterns – access modes and the possibility of trip chaining and pedestrian tours in localized areas – for which the DACZ model would yield better results as it generally produces more homogeneous zones with balanced size and number, and hence reveals the likely correlations between the intended variables more effectively.

Moreover, as the Region is proactively working to implement re-urbanization and Transit Oriented Development (TOD) strategies – to maximize the benefits derived from investment in the LRT project – the DACZ model and our multi-level analytical approach can better inform planning policies and help achieving development goals in a variety of ways. For example, the re-urbanization objectives are often manifested through zoning plans, where the complex and inherently context-dependent zone delineation task can pose critical challenges to planning practice. We assert that the DACZ model can be used to facilitate the creation of the zoning boundaries under different alternative development scenarios and at a very disaggregate scale. Its fine-tuned zonal structure can help to make more informed zoning decisions with respect to the favorable land use diversity levels that best serve the development objectives, similar to the ones that have been envisioned in the Greater Golden Horseshoe Growth Plan and the Waterloo Region's new Regional Official Plan.

The DACZ model is also instrumental to identify transit alignments and potential station areas especially for TOD projects, where the actual levels of land use diversity is perceived as a determinant factor in locating major transit stations (Singh, Fard, Zuidgeest, Brussel, & Maarseveen, 2014). For instance, in a conceivable scenario, if the DACZ model would have been used in the planning processes of the LRT project in Waterloo Region, a different choice of alignment and station locations could have been made to better connect the hotspots of population density and activity clusters across the region – aside from land acquisition considerations. Figure 4-18 shows the spatial distribution of the areas with combined high density and high diversity levels adjacent to the currently implemented LRT alignment (the green line) and around an alternative corridor (illustrated with the dashed lines). As the map indicates, the DACZ model has revealed multiple activity clusters across the King Street and University Avenue within the City of Waterloo, which had not been considered when the Region was making their choice about the LRT alignment. Presumably, if the Region had this land use diversity data as input to the alignment selection processes it might have resulted in a different corridor, where multiple stations could have been located within proximity of the identified high density and high diversity DACZs (highlighted with gray polygons). This alternative alignment could constructively support the intended TOD goals as both land use diversity and higher density developments coincided within the proposed corridor.

From the land use development perspective, findings of this study also highlight the importance of higher levels of land use diversity and access to rapid transit on larger magnitude of developments. These observations have important implications to inform planning policy on leveraging land use diversity in attracting desirable investments towards re-urbanization objective. The application of the DACZ model can help to accurately identify areas with the right levels of land use diversity. These areas are often smaller in size and tend to be smoothed out within the boundaries of generally larger conventional zones.

In conclusion, this research was a step towards better articulating land use diversity measure, and establishing a novel spatial unit of analysis in urban studies. The findings of this study reinforce the importance of diversity measure in land use and travel behavior interactions analyses, complementing the density and other built environment characteristics that are more commonly used in planning practice. The methods developed in this dissertation have shown to be effective to inform evidence-based policy development and planning project assessments, particularly in transportation planning.

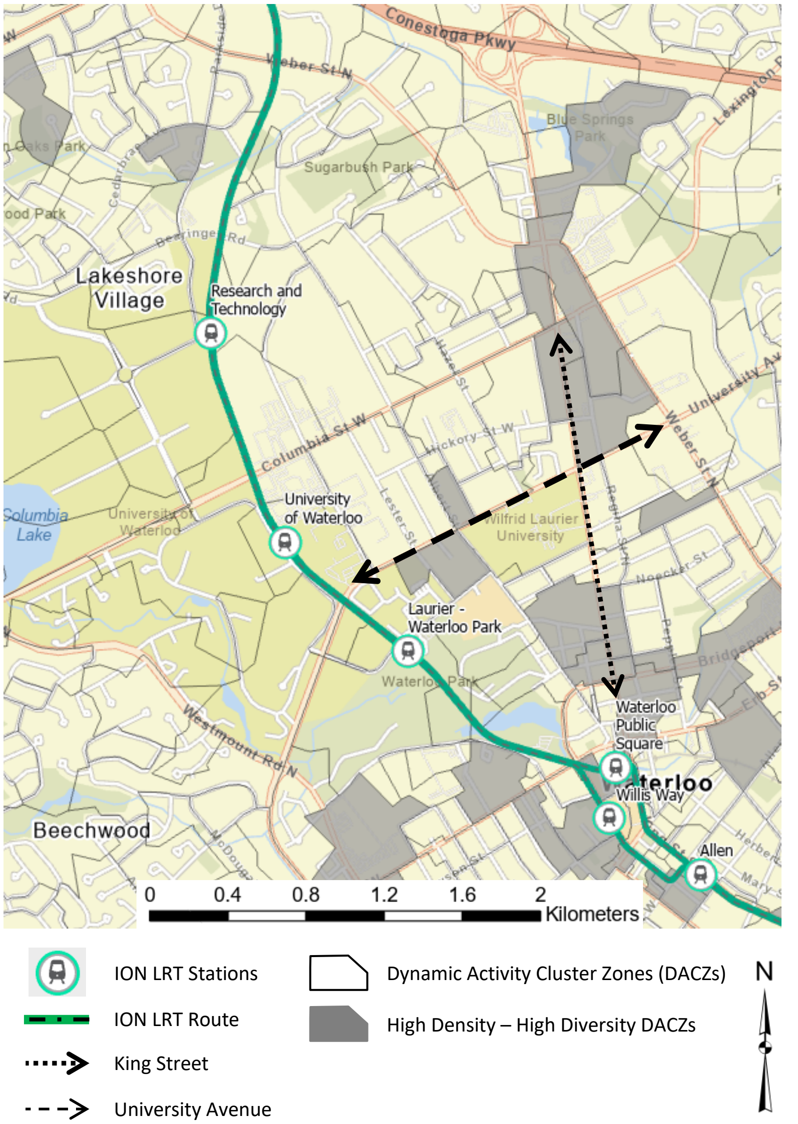


Figure 4-18 - Diagram of the Alternative LRT Corridor in Relation to the Currently Implemented ION LRT Alignment within the City of Waterloo

Chapter 5: Conclusion

This research examined and interpreted the effects of land use heterogeneity and access to rapid transit on an ongoing urban dynamic in Waterloo Region in southern Ontario, Canada. We initially established two overarching research objectives to pursue in this dissertation. The first objective was to develop and evaluate a set of spatial models to examine the impacts of data aggregation scale in land use and transportation system interactions analysis. The second objective was to investigate the effects of land use heterogeneity and access to rapid transit on travel behavior and developments.

We developed, applied and validated a novel spatial aggregation method to create dynamic analysis zones for land use and transportation system interactions studies. This aggregation method embeds an enhanced measure of heterogeneity to ensure (certain level of) land use homogeneity as the determinant factor for spatial delineation of the derived dynamic zones. This heterogeneity index was originally developed to quantify land use diversity as it is perceived within urban transportation planning and policy analysis.

The proposed spatial aggregation method and its associated quantitative models were tested using a set of land use, travel behavior and the building permits data from the Region of Waterloo, a mid-size metropolitan area with a growing economy and population in Ontario, Canada. The following section summaries the key findings from this research.

5.1 Key Findings

This dissertation consists of a series of interconnected analyses to answer the proposed research questions summarized as follows:

1. What is the appropriate spatial scale to analyze land use and transportation system interactions and their impacts on travel behavior in a mid-size metropolitan region, based on observations from Waterloo Region? What are the benefits of using dynamically defined zones in travel behavior analysis?

Sections 3.5 and 4.2 present the methods and outcomes of the implemented models to respond to the primary question. The first analysis explored the suitability of dynamic zones in representing more homogeneous land use areas (see Sections 3.5.1 & 4.2.1). By applying the enhanced diversity measurement, a set of Dynamic Activity Cluster Zones (DACZ) structure was created and examined over a suburban subdivision in the City of Waterloo. This DACZ system was then compared to the commonly used pre-defined conventional traffic analysis zones (TAZ) called TTS TZS, which has been the standard for transportation analysis across the Greater Toronto and Hamilton Area (GTHA).

The results indicated that the dynamic zones performed more effectively in creating homogeneous zones, and in delineating spatial boundaries where the higher levels of heterogeneity dictate dividing up an area. In contrast, the way that the conventional TAZ was built in the studied area was more arbitrary in terms of breaking boundaries. The TAZ also fails to address the need for creating multiple zones where the higher levels of heterogeneity necessitates to do so. Moreover, the TAZ tends to over subdivide areas that are sufficiently homogeneous and likely to exhibit very consistent travel behavior.

The second analysis was a systematic assessment of the DACZ model performance in producing a variety of zones taking into account variability of the spatial adjacency parameter and a range of heterogeneity thresholds. The sensitivity analysis demonstrated the robust behavior of the DACZ model in generating zones for which the size is defined as a function of homogeneity (see Sections 3.5.2 & 4.2.2). This was an essential characteristic to be proven to ensure that the derived dynamic zones are able to generate predictable outcomes in travel behavior modeling and analysis.

This evaluation was further extended in the third analysis, by implementing and testing a graph-based spatial partitioning method that is commonly used in a wide range of studies for creating homogeneous aerial units (see Sections 3.5.3 & 4.2.3). The graph-based spatial partitioning method was shown to be highly monotonic in producing zones with minimum variation in their heterogeneity distribution while illustrating a significant variance in zone sizes. In contrast, the DACZ method was found to be more effective in flexibly producing larger numbers of smaller and more homogeneous zones or, alternatively, generating fewer but larger and more heterogeneous zones.

2. How do built environment characteristics (e.g., density and diversity) and access to rapid transit influence travelers' behavior across a mid-sized region, based on observations from Waterloo Region?

Sections 3.6 and 4.3 present the work to respond to the second research question. The first analysis was a comparison of the DACZ and a conventional TAZ (called PLUM), in a regional travel behavior study. The latest available regional daily travel behavior data from TTS 2016 were aggregated for both zonal structures, taking into account different travel modes, trip purposes and built environment characteristics (e.g., density and diversity) across Waterloo Region (see Section 3.6.1).

For both zonal systems, each zone was classified into a non-overlapping built environment category given the population density of the zone paired with the land use diversity levels – computed using the enhanced entropy measurement method. To represent the full range of built environment heterogeneity, three classes of population density and three classes of land use diversity were statistically identified and intersected that together formed a set of nine combined classes from the "High Density – High Diversity" to the "Low Density – Low Diversity" category.

Both zonal systems revealed similar distributions of mode shares, in which, while the auto trips were dominant, by increasing the levels of population density and land use diversity more trips were made using non-auto modes. Despite the two systems' similarity in representing aggregate travel behavior, DACZ zonal structure was found to be more effective in capturing variations of the active transportation and public transit mode shares. The results of the mode share distributions comparison also showed that the potential impacts of the built environment homogeneity on travel behavior were more pronounced within dynamic zones compared to the conventional TAZs.

Importantly, it was evident that the high levels of land use diversity is potentially a better predictor of the lower auto mode share compared to the high levels of density. In fact, finding high density zones that have low diversity is likely to produce errant results in terms of expected travel behavior.

When considering different trip purposes, the mode share analysis reinforced previous findings in that the DACZ system showed higher variance in the regional distributions of travel mode shares for most of the travel modes and trip purposes. However, no significant differences was observed for the home-based work (HBW) trips.

In an extended analysis, using the DACZ system and the multi-modal transportation network developed in this research, we computed and compared the average travel times (ATT) for the trips originated within the cities of Waterloo and Kitchener. Assessing the distribution of ATTs across the built environment classes revealed an identifiable variation between different categories in which the longer trips often were associated with the “Low Density” zones. Compared to the mode share evaluation, the findings of the travel time analysis showed a higher degree of agreement with the expected travel behavior hypothesized in the literature.

We conclude that the DACZ structure is able to represent more consistent travel behavior at zonal level, which results in higher variance in the regional distribution of travel behavior. This ability arises from the greater intra-zonal homogeneity that results in larger inter-zonal heterogeneity at macro level.

3. How do proximity to LRT corridor and land use diversity levels impact the development dynamics across a mid-sized region, based on observations from Waterloo Region?

Sections 3.7 and 4.4 address the third research question that seeks to understand how the introduction of proposed and actual rapid transit investments are correlated to development. After adjusting the building permits construction values based on the historical inflation rates, a temporal investigation was conducted to analyze the trend of changes for different permit types, across three spatial scales. The historical building permit data were analyzed at biannual intervals from 2000 to 2019. Special distinction was made to highlight the two periods before and after the LRT project funding announcement (2010-2011), and during the construction period of the phase-one of the LRT line that began in 2014 and ended in 2018. The average and the total permits construction values were calculated across multiple scales including the entire study area within the cities of Kitchener and

Waterloo, relative to the Regions' Central Transit Corridor (CTC) and within different heterogeneous built environment categories.

Multiple trends were identified during the studied 20-year period and the results were indicative of areas that have disproportionately attracted more and higher valued land use developments during certain periods, especially after announcement of the LRT project funding. This analysis also revealed some distinguished patterns confirming the role of higher levels of land use diversity (heterogeneity) and access to rapid transit on larger magnitude of developments. Areas within the CTC have been the major driver for the higher average construction values through time, for almost all permit types. Contrary to the intuitive expectations, high levels of density have shown little association with attracting greater scale of land use developments.

In conclusion, synthesizing multi-scale travel behavior and built environment analysis indicates that within the cities of Kitchener and Waterloo, while the trends of land use development have been highly related with the access to the light rail transit corridor, the levels of land use diversity and population density, these characteristics are likely not yet high enough to produce significant changes in travel behavior patterns.

5.2 Contributions

This research was motivated by the idea that land use diversity is equally important as densities (and other attributes) to the analysis of land use and transportation system interactions. It sought to advance understanding of the impacts caused by the choice of analysis zones on the travel behavior and land use development analysis outcomes. By fulfilling the overarching objectives established at the outset of this study, this dissertation makes four methodological and empirical contributions to land use and transportation systems interactions studies.

First, it developed and evaluated an enhanced land use diversity measurement method to quantify diversity as it is perceived within urban land use and transportation planning. This method extends the conventional entropy measurement by normalizing it by the maximum number of land use classes observed within any given zone in a landscape (study area). This modification to the model construct produces a significantly more generalizable index to consistently quantify land use diversity as a built environment characteristic. Land use diversity has increasingly become a standard indicator for built environment assessment in urban and transportation research and policy analysis. Our developed method along with the implemented computational script provide future research with an effective toolset to efficiently compute heterogeneity scores for a large number of spatial zones, consisting of any number of land use types.

Second, this study applied the enhanced diversity measure to create a set of Dynamic Activity Cluster Zones (DACZ) for conducting further LUTI analysis. As a novel spatial structure, DACZ helps to aggregate travel behavior and land use data in analytically defined homogeneous zones. This process starts with generating disaggregate zones to be evaluated and combined to form the dynamic zones. By developing a new and flexible disaggregate spatial unit called block segment, this dissertation addresses the methodological limitations of the previous studies that confined to the use of roadways as boundaries for delineating analysis zones. By providing an intermediate unit, the block segment facilitates the transformation and integration of data points collected at road intersections and city blocks edges that are often tend to be randomly assigned and cause biased spatial estimates.

Introduction of the DACZ as a homogeneity-based and data-driven spatial unit, enhances upon the pre-defined analysis zones (e.g., TAZs and Zip code areas) in a simple but effective way. The DACZ creation model only aggregates input zones (block segments) that have similar diversity levels but also assembled from similar disaggregate land uses that make up their diversity. This capability offers an important advantage in generating zones for which the size is defined as a function of homogeneity and, as a result, these zones will generate more predictable outcomes in travel behavior modeling and analysis. In addition, by better capturing the complexity of urban road networks and parcel fabric, the DACZ method decreases uncertainties involved in mapping travel behavior data such as the GPS tracking and AVL/APC data points into aerial units. These improvements enable researchers to further evaluate and mitigate some of the frequently discussed biases originating from the spatial scale issue, the boundary effect and modifiable areal unit problem (MAUP). Technically, our enhanced entropy measurement model contributes to the vector based landscape metrics analysis, which is a developing area of research within the spatial modeling domain. By implementing a multi-level dynamic adjacency search strategy, we overcame the challenges arising from higher-order adjacency identification. This approach enabled us to expand the spatial window of analysis for the aggregation algorithm beyond the immediate neighboring zones – a common limitation seen in previous studies.

In addition to enhancing methods on defining the spatial unit of analysis, this study also contributes to empirical integrated land use and transportation planning and policy evaluation, especially for mid-sized cities and regions.

First, this research used a dynamic zonal structure developed and examined in earlier steps to investigate the travel behavior patterns across the Waterloo Region. The Region is employing an integrated planning approach to drive sustainable urban development and transportation in a dispersed metropolitan area. This study evaluated the degree to which the mode share of trips and their average travel times were influenced by land use heterogeneity patterns. This work found that although the auto trips are dominant in the study area, by increasing the levels of population density and land use diversity more trips are made using non-auto modes. This outcome supports the LUTI theories which contend that areas with diverse land uses and high levels of population density are more conducive to active transportation and public transit trips compared to more dispersed and uniform areas.

As the DACZ structure possesses greater intra-zonal homogeneity and hence greater inter-zonal heterogeneity, it tends to represent more consistent travel behavior at the zonal level, which in turn shows higher variances in the overall distribution of travel behavior. Hence, in the case of Waterloo Region our method shown slightly more effectiveness in capturing variations of the active transportation and public transit mode shares in regard to the built environment heterogeneity.

These specifications highlight the potential value of the dynamic zone delineation for studies of non-auto transportation modes, where the finer scale and more homogeneous structure of these analytically derived zones would allow for better modeling of pedestrian and cyclist trips. These trips are often shorter and take place within boundaries of a conventional analysis zone that increases the risk of their underestimations. As non-motorized transportation modes are increasingly attracting attention in planning and livable community development, this work can potentially benefit those studies in better understanding of the travel behavior patterns associated with non-auto transportation modes through improved spatial representation.

Second, this study sought to understand how the introduction of proposed and actual rapid transit investments are correlated to land use development in a fast growing metropolitan area. It investigated the effects of land use heterogeneity and access to rapid transit on temporal patterns of building permits issued over the past 20 years. Findings of this analysis highlight the importance of higher levels of land use diversity (heterogeneity) and access to rapid transit on larger magnitude of developments. From the temporal perspective, the results showed that the period after announcement of the LRT project funding witnessed disproportionately higher valued developments attracted to the region. Moreover, for almost all identified building permit types, areas within the Regions' Central Transit Corridor (CTC) have been the major driver for the higher average construction values through time. These observations have important implications for planning policy to better inform re-urbanization and Transit Oriented Development (TOD) strategies on leveraging land use diversity in attracting desirable investments. Applied to planning practice, more informed zoning decisions can be made with respect to the favorable land use diversity levels that best serve the development objectives.

In summary, this research was a step towards better articulating spatial unit of analysis in urban studies, thereby easing the disconnect in land use and transportation integration in planning and policy development. Our methodological and empirical contributions are anticipated to be of interest to the urban and transportation researchers and planners as well as for evidence-based policy development and evaluation purposes.

5.3 Limitations and Future Research

Building upon the concept of land use and transportation system interactions (LUTI), this dissertation contributed to the development of analytical methods, and extending our understanding of the complex interactions taking place within urban dynamics. However, as with all research, our results should be interpreted in the light of its assumptions and limitations.

First, this study investigated the observed heterogeneity of travel behavior and land use development trends by focusing on the levels of land use diversity, population density and the access to the transit corridor in a mid-size metropolitan region. These characteristics, although essential, do not provide a complete picture of the processes taking place within LUTI dynamics. From a theoretical perspective, there is a wider set of factors is involved in these dynamics that are difficult to measure. This, for example, includes the perception of safety or the environmental design that are challenging to quantify. In particular, in the case of Waterloo Region, future studies should address the impacts of the built environment design elements such as road network pattern, walkability, parcel divisions and zoning specifications on both travel behavior and land use development.

Second, in this study, no specific distinction was made in conducting analyses considering the socioeconomic or demographic context. However, as it was revealed in our analysis of mode share distributions by trip purposes, the demographic (and likely socioeconomic) characteristics are determinant factors in regard to LUTI dynamics, especially on travel mode choice decisions. Also, as a study with emphasis on improving aggregation methods, this research has not explored micro level travel preferences or individual land development decisions. These processes are involved with complex interactions between individuals and interplay between personal preferences and the wider context of urban environment, which need to be addressed in future studies. Moreover, as literature suggests, negative externalities of the transportation infrastructure investments and the TOD strategies, most notably including gentrification, NIMBYism and increased crime are often overlooked in the re-urbanization projects and policy analysis. Hence, future research may benefit from incorporating these outcomes, likely through employing a cost-benefit analysis framework.

Third, this research demonstrated that both the block segment as the basis for zone delineation, and the Dynamic Activity Cluster Zones (DACZ) as the spatial unit for LUTI analysis are promising concepts to overcome the limitations of the conventional roadway-based and density-based zones. However, assembling the population, household, and employment counts for the resulting zones shown to pose implementation challenges. While these data are often collected at a very disaggregate level, but are disseminated at an aggregate spatial scale corresponding to the administrative boundaries such as the census tracts or Zip code areas. Future research should investigate potential solutions to address this challenge systematically. Moreover, despite the advantages the DACZ model provides, it comes with the cost of greater computational requirements. An intuitive approach to decrease the computational intensity of the dynamic zone creation can be using the regular spatial tessellation (e.g., fishnet or hexagonal grid) instead of the block segments, especially for greater study areas that may cover both urbanized and non-urbanized environments with fragmented patches.

Fourth, the DACZ creation model relies on the disaggregate land use and transportation network data to produce zones that are suitable for LUTI analysis and modeling. While these data sets are often accessible to the municipalities and government agencies, however, accessing these still remains a challenge for academic research, especially for large scale studies (Johnson & Sieber, 2012). For example, due to the data access limitations, this research had to focus only on the central areas of the cities of Kitchener and Waterloo for most parts of the analyses. This limited spatial extent has likely precluded us from capturing the potentially greater variance in travel behavior and development patterns. This issue warrants further studies to cover the entire Region to draw a more complete conclusion. Hence, future work may benefit from assessing and incorporating publicly accessible land use data from open access sources such as the OpenStreetMap (OSM).

Further studies may also improve generalizability of the DACZ method by adopting land use reclassification strategy provided by the Land-Based Classification Standards (LBCS). LBCS as a hierarchical and multidimensional land use classification system (American Planning Association, 2000), is being increasingly used as a standard reference in planning studies in North America.

Fifth, this research conducted a sensitivity analysis to examine reliability of the DACZ model in producing zonal structure as a (predictable) function of three parameters including aggregation heterogeneity threshold, levels of adjacency, and the original spatial disaggregation methods (i.e., Voronoi vs. block segments). While the results showed the robustness of this model, future studies are suggested to extend this evaluation by testing a set of alternative assumptions such as higher order spatial adjacency and different distance metrics that can impact neighbor definition for disaggregate zones. These parameters are anticipated to change the model outcomes in creating zones with varied compactness and size distribution. Moreover, since both the DACZ model and the graph-based clustering were computationally intensive algorithms and required excessive processing time on standard workstations, we implemented and run our work on the Graham Cluster, a High Performance Computing platform provided by Compute Canada consortia. Although using vectorization capability of R programming language enabled the models to perform reasonably efficiently, future work may wish to adopt more advanced parallel computing strategies for scaling up these models.

Finally, there are potential opportunities to further extend this research through methodological and empirical studies. Most notably, application of the enhanced entropy index and the DACZ structure can significantly contribute to the statistical association analysis (e.g., regression) and causal inference modeling in LUTI studies. The enhanced index provides a rigorously quantified measure of built environment heterogeneity, and the DACZ system facilitates better spatial representation of variables, both are critical to the robustness of the models. Similarly, the methods developed through this research can enhance the travel demand models. The ongoing developments of the next generation models aim to respond to the long standing call for improving theoretical conceptualizations and model implementations (Miller, 2020) which are both anticipated to benefit from improvements the DACZ method offers. Ultimately, as this study was conducted and covered the period prior to the operation of the ION LRT in Waterloo Region, a great addition to this research, would be to perform a follow up study in the mid-term to examine the land use and travel behavior consequences that will be realized after the operation of the LRT in its full capacity.

References

Acheampong, R. A., & Silva, E. (2015). Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land Use; Vol 8, No 3 (2015)*. Retrieved from https://www.jtlu.org/index.php/jtlu/article/view/806

Addie, J.-P. D. (2013). Metropolitics in Motion: The Dynamics of Transportation and State Reterritorialization in the Chicago and Toronto City-Regions. *Urban Geography*, *34*(2), 188–217. https://doi.org/10.1080/02723638.2013.778651

Al-Ogaibi, A. A., Sharieh, A., & Bremananth, R. (2014). A Robust Method for Finding Nearest-Neighbor using Hexagon Cells. *International Journal of Mathematical and Computational Sciences*, *8*(1), 38–45.

Alonso, W. (1964). *Location and Land Use*. Harvard University Press, Cambridge.

American Planning Association. (2000). *Land-Based Classification Standards (LBCS)*. Retrieved from https://www.planning.org/lbcs/standards/

Anas, A. (1982). *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Policy Analysis with Discrete Choice Models.*

Anas, A., Arnott, R., & Small, K. A. (1998). Urban Spatial Structure. *Journal of Economic Literature*, *36*(3), 1426–1464. Retrieved from http://www.jstor.org/stable/2564805

Assunção, R. M., Neves, M. C., Câmara, G., & da Costa Freitas, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, *20*(7), 797–811.

Aydin, O., Janikas, M. V, Assunção, R. M., & Lee, T.-H. (2021). A quantitative comparison of regionalization methods. *International Journal of Geographical Information Science*, 1–29. https://doi.org/10.1080/13658816.2021.1905819

Bachmann, C., Roorda, M. J., Abdulhai, B., & Moshiri, B. (2013). Fusing a bluetooth traffic monitoring system with loop detector data for improved freeway traffic speed estimation. *Journal of Intelligent Transportation Systems*, *17*(2), 152–164.

Banister, D. (2011a). Cities, mobility and climate change. *Journal of Transport Geography*, *19*(6), 1538–1546. https://doi.org/10.1016/j.jtrangeo.2011.03.009

Banister, D. (2011b). The trilogy of distance, speed and time. *Journal of Transport Geography*, *19*(4), 950–959. https://doi.org/10.1016/j.jtrangeo.2010.12.004

Batty, M. (1974). Spatial Entropy. *Geographical Analysis*, *6*(1), 1–31. https://doi.org/10.1111/j.1538-4632.1974.tb01014.x

Batty, M. (2008a). *Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies.*

Batty, M. (2008b). Fifty years of urban modeling: Macro-statics to micro-dynamics. In *The dynamics of complex urban systems* (pp. 1–20). Springer.

Batty, M. (2012). A Generic Framework for Computational Spatial Modelling. In *Agent-Based Models of Geographical Systems* (pp. 19–50). https://doi.org/10.1007/978-90-481-8927-4_2

Batty, M., Axhausen, K., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., … Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, *214*(1), 481–518. https://doi.org/10.1140/epjst/e2012-01703-3

Batty, M., Crooks, A. T., See, L. M., & Heppenstall, A. J. (2012). Perspectives on Agent-Based Models and Geographical Systems. In *Agent-Based Models of Geographical Systems* (pp. 1–15). https://doi.org/10.1007/978-90-481-8927-4_1

Baum-Snow, N., Kahn, M. E., & Voith, R. (2005). Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000 [with Comment]. *Brookings-Wharton Papers on Urban Affairs*, 147–206. Retrieved from http://www.jstor.org/stable/25067419

Bemelmans-Videc, M.-L., Rist, R. C., & Vedung, E. O. (1998). *Carrots, sticks, and sermons: Policy instruments and their evaluation* (Vol. 1). Transaction Publishers.

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand* (Vol. 9). MIT press.

Benenson, I., & Torrens, P. M. (2004). *Geosimulation: Automata-based modeling of urban phenomena*. John Wiley & Sons.

Bengston, D. N., Fletcher, J. O., & Nelson, K. C. (2004). Public policies for managing urban growth and protecting open space: policy instruments and lessons learned in the United States. *Landscape and Urban Planning*, *69*(2–3), 271–286. https://doi.org/10.1016/j.landurbplan.2003.08.007

Berke, P. R., & Conroy, M. M. (2000). Are We Planning for Sustainable Development? *Journal of the American Planning Association*, *66*(1), 21–33. https://doi.org/10.1080/01944360008976081

Birch, C. P. D., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, *206*(3), 347–359. https://doi.org/https://doi.org/10.1016/j.ecolmodel.2007.03.041

Birkin, M., & Wu, B. (2012). A Review of Microsimulation and Hybrid Agent-Based Approaches. In *Agent-Based Models of Geographical Systems* (pp. 51–68). https://doi.org/10.1007/978-90-481-8927-4_3

Boarnet, M. G. (2011). A Broader Context for Land Use and Travel Behavior, and a Research Agenda. *Journal of the American Planning Association*, *77*(3), 197–213. https://doi.org/10.1080/01944363.2011.593483

Bogart, W. T., & Ferry, W. C. (1999). Employment centres in greater Cleveland: Evidence of evolution in a formerly monocentric city. *Urban Studies*, *36*(12), 2099–2110.

Boots, B., Okabe, A., & Sugihara, K. (1992). *Spatial tessellations*. John Wiley and Sons Ltd, England.

Boots, B., Okabe, A., & Thomas, R. W. (2003). *Modelling geographical systems : statistical and computational applications*.

Borchert, J. R. (1967). American Metropolitan Evolution. *Geographical Review*, *57*(3), 301–332. https://doi.org/10.2307/212637

Bourne, L. S. (1989). Are new urban forms emerging? Empirical tests for Canadian urban areas. *Canadian Geographer / Le Géographe Canadien*, *33*(4), 312–328. https://doi.org/10.1111/j.1541-0064.1989.tb00918.x

Bourne, L. S. (1996). REURBANIZATION, UNEVEN URBAN DEVELOPMENT, AND THE DEBATE ON NEW URBAN FORMS. *Urban Geography*, *17*(8), 690–713. https://doi.org/10.2747/0272-3638.17.8.690

Bunting, T., & Filion, P. (2010). Epochs of Canadian Urban Development. In T. Bunting, P. Filion, & R. Walker (Eds.), *Canadian cities in transition : perspectives for an urban age* (Fourth). Oxford University Press.

Bunting, T., Filion, P., & Priston, H. (2002). Density Gradients in Canadian Metropolitan Regions, 1971-96: Differential Patterns of Central Area and Suburban Growth and Change. *Urban Studies*, *39*(13), 2531–2552. https://doi.org/10.1080/0042098022000027095

Burg, L. van den, & Stolk, E. (2004). *Urban Analysis Guidebook: typomorphology*. Delft University of Technology, Faculty of Architecture.

Burns, J. R., & Marcy, W. M. (1979). Causality: Its characterization in system dynamics and KSIM models of socioeconomic systems. *Technological Forecasting and Social Change*, *14*(4), 387–398. https://doi.org/10.1016/0040-1625(79)90036-2

Campbell, S. (1996). Green Cities, Growing Cities, Just Cities?: Urban Planning and the Contradictions of Sustainable Development. *Journal of the American Planning Association*, *62*(3), 296–312. https://doi.org/10.1080/01944369608975696

Campbell, S. (2016). The Planner's Triangle Revisited: Sustainability and the Evolution of a Planning Ideal That Can't Stand Still. *Journal of the American Planning Association*, *82*(4), 388–397. https://doi.org/10.1080/01944363.2016.1214080

Casello, J. M. (2007). Transit competitiveness in polycentric metropolitan regions. *Transportation Research Part A: Policy and Practice*, *41*(1), 19–40. https://doi.org/http://dx.doi.org/10.1016/j.tra.2006.05.002

Casello, J. M., & Fard, P. (2017). Automated Tool for Geographic Information Systems That Supports Transit Network Design by Identifying Urban Activity Centers. *Transportation Research Record: Journal of the Transportation Research Board*, *2651*, 12–21. https://doi.org/10.3141/2651-02

Casello, J. M., & Smith, T. (2006). Transportation Activity Centers for Urban Transportation Analysis. *Journal of Urban Planning and Development*, *132*(4), 247–257. https://doi.org/10.1061/(ASCE)0733-9488(2006)132:4(247)

Casello, J. M., & Usyukov, V. (2014). Modeling Cyclists' Route Choice Based on GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, (2430), 155–161.

Cervero. (1984). Journal report: light rail transit and urban development. *Journal of the American Planning Association*, *50*(2), 133–147.

Cervero, & Duncan. (2002). Transit's value-added effects: light and commuter rail services and commercial land values. *Transportation Research Record: Journal of the Transportation Research Board*, *1805*(1), 8–15.

Cervero, R., & Landis, J. (1997). Twenty years of the Bay Area Rapid Transit system: Land use and development impacts. *Transportation Research Part A: Policy and Practice*, *31*(4), 309–333. https://doi.org/http://dx.doi.org/10.1016/S0965-8564(96)00027-4

Cervero, R., & Wu, K.-L. (1997). Polycentrism, Commuting, and Residential Location in the San Francisco Bay Area. *Environment and Planning A*, *29*(5), 865–886. https://doi.org/10.1068/a290865

Chang, K.-T., Khatib, Z., & Ou, Y. (2002). Effects of Zoning Structure and Network Detail on Traffic Demand Modeling. *Environment and Planning B: Planning and Design* , *29*(1), 37–52. https://doi.org/10.1068/b2742

Chatman, D. G. (2008). Deconstructing development density: Quality, quantity and price effects on household non-work travel. *Transportation Research Part A: Policy and Practice*, *42*(7), 1008–1030. https://doi.org/https://doi.org/10.1016/j.tra.2008.02.003

Chatman, D. G., & Noland, R. B. (2013). Transit Service, Physical Agglomeration and Productivity in US Metropolitan Areas. *Urban Studies*, *51*(5), 917–937. https://doi.org/10.1177/0042098013494426

Chen, C., Gong, H., Lawson, C., & Bialostozky, E. (2010). Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the New York City case study. *Transportation Research Part A: Policy and Practice*, *44*(10), 830–840.

Clifton, K. J., Currans, K. M., & Muhs, C. D. (2015). Adjusting ITE's Trip Generation Handbook for urban context. *Journal of Transport and Land Use*, *8*(1 SE-), 5–29. https://doi.org/10.5198/jtlu.2015.378

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016). Development of destination choice models for pedestrian travel. *Transportation Research Part A: Policy and Practice*, *94*, 255–265.

Craig, S. G., & Ng, P. T. (2001). Using Quantile Smoothing Splines to Identify Employment Subcenters in a Multicentric Urban Area. *Journal of Urban Economics*, *49*(1), 100–120. https://doi.org/10.1006/juec.2000.2186

Currans, K. M., & Clifton, K. J. (2015). Using household travel surveys to adjust ITE trip generation rates. *Journal of Transport and Land Use*, *8*(1 SE-), 85–119. https://doi.org/10.5198/jtlu.2015.470

Currans, K. M., & Clifton, K. J. (2018). Exploring ITE's Trip Generation Manual: Assessing age of data and land-use taxonomy in vehicle trip generation for transportation impact analyses. *Transportation Research Part A: Policy and Practice*, *118*, 387–398. https://doi.org/https://doi.org/10.1016/j.tra.2018.09.007

Dawkins, C. J., & Nelson, A. C. (2002). Urban containment policies and housing prices: an international comparison with implications for future research. *Land Use Policy*, *19*(1), 1–12. https://doi.org/10.1016/S0264-8377(01)00038-2

Dilworth, R. (2009). *The city in American political development*. Routledge.

Ding, C., Knaap, G. J., & Hopkins, L. D. (1999). Managing urban growth with urban growth boundaries: A theoretical analysis. *Journal of Urban Economics*, *46*(1), 53–68.

Donald, B. (2005). The politics of local economic development in Canada's city-regions: New dependencies, new deals, and a new politics of scale. *Space and Polity*, *9*(3), 261–281. https://doi.org/10.1080/13562570500509935

Duany, A., Speck, J., & Lydon, M. (2010). *The smart growth manual*. McGraw-Hill.

Duque, J. C., Anselin, L., & Rey, S. J. (2012). The max-p-regions problem. *Journal of Regional Science*, *52*(3), 397–419.

Elsawah, S., Pierce, S. A., Hamilton, S. H., van Delden, H., Haase, D., Elmahdi, A., & Jakeman, A. J. (2017). An overview of the system dynamics process for integrated modelling of socio-ecological systems: Lessons on good modelling practice from five case studies. *Environmental Modelling & Software*, *93*, 127–145. https://doi.org/https://doi.org/10.1016/j.envsoft.2017.03.001

Engelberg, D., He, H., Le, D.-T., & Zegras, P. C. (2021). *Chapter 21 - Accessibility, land use models, and modeling* (C. Mulley & J. D. B. T.-U. F. and A. Nelson, Eds.). https://doi.org/https://doi.org/10.1016/B978-0-12-819822-3.00019-5

Evans, Pratt, Stryker, & Kuzmyak. (2007). Transit Oriented Development. In *Transit Cooperative Research Program (TCRP). Traveler Response to Transportation System Changes*.

Ewing, R., & Cervero, R. (2001). Travel and The Built Environment: A Synthesis. *Transportation Research Record: Journal of the Transportation Research Board*, *1780*(1), 87–114.

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Filatova, T., Verburg, P. H., Parker, D. C., & Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: Challenges and prospects. *Environmental Modelling & Software*, *45*, 1–7. https://doi.org/10.1016/j.envsoft.2013.03.017

Filion, P., & Bunting, T. (2010). Fundamentals of Cities. In T. Bunting, P. Filion, & R. Walker (Eds.), *Canadian cities in transition : perspectives for an urban age* (Fourth). Oxford University Press.

Filion, P., Bunting, T., McSpurren, K., & Tse, A. (2004). Canada-U.S. Metropolitan Density Patterns: Zonal Convergence and Divergence1. *Urban Geography*, *25*(1), 42–65. https://doi.org/10.2747/0272-3638.25.1.42

Filion, P., Hoernig, H., Bunting, T., & Sands, G. (2004). The Successful Few: Healthy Downtowns of Small Metropolitan Regions. *Journal of the American Planning Association*, *70*(3), 328–343. https://doi.org/10.1080/01944360408976382

Filion, P., & Kramer, A. (2012). Transformative Metropolitan Development Models in Large Canadian Urban Areas: The Predominance of Nodes. *Urban Studies*, *49*(10), 2237–2264. https://doi.org/10.1177/0042098011423565

Fisher, P. (1997). The pixel: a snare and a delusion. *International Journal of Remote Sensing*, *18*(3), 679–685.

Fishman, R. (1987). Beyond Suburbia: The Rise of the Technoburb. In *Bourgeois Utopias: The Rise and Fall of Suburbia*. Basic books.

Forrester, J. W. (1969). *Urban Dynamics*. Pegasus Communications.

Garreton, M., & Sánchez, R. (2016). Identifying an optimal analysis level in multiscalar regionalization: A study case of social distress in greater Santiago. *Computers, Environment and Urban Systems*, *56*, 14–24.

Gatrell, A. C. (1983). *Distance and Space: A Geographical Perspective*. Clarendon Press.

Gehrke, S. R., & Clifton, K. J. (2014). Operationalizing land use diversity at varying geographic scales and its connection to mode choice: Evidence from Portland, Oregon. *Transportation Research Record*, *2453*(1), 128–136.

Gehrke, S. R., & Clifton, K. J. (2016). Toward a spatial-temporal measure of land-use mix. *Journal of Transport and Land Use*, *9*(1), 171–186.

Giuliano, G., Redfearn, C., Agarwal, A., Li, C., & Zhuang, D. (2007). Employment Concentrations in Los Angeles, 1980–2000. *Environment and Planning A*, *39*(12), 2935–2957. https://doi.org/10.1068/a393

Giuliano, G., & Small, K. A. (1991). Subcenters in the Los Angeles region. *Regional Science and Urban Economics*, *21*(2), 163–182.

Giuliano, G., & Small, K. A. (1993). Is the Journey to Work Explained by Urban Structure? *Urban Studies*, *30*(9), 1485–1500. https://doi.org/10.1080/00420989320081461

Government of Ontario. *Province of Ontario Planning Act.* , (1990).

Government of Ontario. *Provincial Policy Statement (PPS).* , (2014).

Government of Ontario. *Places to Grow Act.* , (2017).

Government of Ontario. *Ontario Regulation 282/98 under Assessment Act, R.S.O. 1990, c. A.31.* , (2020).

Graham, D. J. (2007). Agglomeration, productivity and transport investment. *Journal of Transport Economics and Policy (JTEP)*, *41*(3), 317–343.

Grant, J. L., & Filion, P. (2010). Emerging Urban Forms in the Canadian Cities (T. Bunting, P. Filion, & R. Walker, Eds.). *Canadian Cities in Transition : Perspectives for an Urban Age*, Fourth. Oxford University Press.

Griffith, D. A. (1981). Evaluating the transformation from a monocentric to a polycentric city. *The Professional Geographer*, *33*(2), 189–196. https://doi.org/10.1111/j.0033-0124.1981.00189.x

Guiliano, G. (2004). Land Use Impacts of Transportation Investments - Highway and Transit. In Susan Hanson and Genevieve Giuliano (Ed.), *The Geography of Urban Transportation* (3rd ed., pp. 237–273).

Guo, D. (2008). Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science*, *22*(7), 801–823. https://doi.org/10.1080/13658810701674970

Guo, D., Zhu, X., Jin, H., Gao, P., & Andris, C. (2012). Discovering Spatial Patterns in Origin-Destination Mobility Data. *Transactions in GIS*, *16*(3), 411–429. https://doi.org/https://doi.org/10.1111/j.1467-9671.2012.01344.x

Haase, D., & Schwarz, N. (2009). Simulation models on human-nature interactions in urban landscapes ea review including system dynamics, cellular automata and agent-based approaches. *Living Reviews in Landscape Research*, *3*(2).

Haghani, A., Lee, S. Y., & Byun, J. H. (2003). A system dynamics approach to land use/transportation system performance modeling Part I: Methodology. *Journal of Advanced Transportation*, *37*(1), 1–41. https://doi.org/10.1002/atr.5670370102

Handy, S. (2005). Smart Growth and the Transportation-Land Use Connection: What Does the Research Tell Us? *International Regional Science Review*, *28*(2), 146–167. https://doi.org/10.1177/0160017604273626

Handy, S. (2018). Enough with the "D's" Already—Let's Get Back to "A." *Transfers Magazine*. Retrieved from https://trid.trb.org/view/1709460

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73–76. https://doi.org/10.1080/01944365908978307

Harris, R. (2004). Suburbanization and the Employment Linkage. In *Building Work And Home. Manufacturing Suburbs* (pp. 221–236). Retrieved from http://www.jstor.org/stable/j.ctt14bs74p.14

Hartshorn, T. A., & Muller, P. O. (1989). SUBURBAN DOWNTOWNS AND THE TRANSFORMATION OF METROPOLITAN ATLANTA'S BUSINESS LANDSCAPE. *Urban Geography*, *10*(4), 375–395. https://doi.org/10.2747/0272-3638.10.4.375

Helbich, M., Brunauer, W., Hagenauer, J., & Leitner, M. (2013). Data-Driven Regionalization of Housing Markets. *Annals of the Association of American Geographers*, *103*(4), 871–889. https://doi.org/10.1080/00045608.2012.707587

Higgins, C. D. (2019). Accessibility toolbox for R and ArcGIS. *Findings*, 8416.

Higgins, C. D., Ferguson, M., & Kanaroglou, P. S. (2014). Light Rail and Land Use Change: Rail Transit's Role in Reshaping and Revitalizing Cities. *Journal of Public Transportation*, *17*(2), 5.

Higgins, C. D., & Kanaroglou, P. S. (2016). Forty years of modelling rapid transit's land value uplift in North America: moving beyond the tip of the iceberg. *Transport Reviews*, 1–25. https://doi.org/10.1080/01441647.2016.1174748

Hodge, G., & Gordon, D. L. A. (2008). *Planning Canadian communities : an introduction to the principles, practice and participants*. Nelson Thomson.

Huang, H. (1996). The Land-Use Impacts of Urban Rail Transit Systems. *CPL Bibliography*, *11*(1), 17–30. https://doi.org/10.1177/088541229601100103

Hunt, J. D., Kriger, D. S., & Miller, E. J. (2005). Current operational urban land-use–transport modelling frameworks: A review. *Transport Reviews*, *25*(3), 329–376. https://doi.org/10.1080/0144164052000336470

Hurst, N. B., & West, S. E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics*, *46*, 57–72. https://doi.org/10.1016/j.regsciurbeco.2014.02.002

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of Transportation and Land Use Change: A Guide to the Territory. *Journal of Planning Literature*, *22*(4), 323–340. https://doi.org/10.1177/0885412207314010

Institute of Transportation Engineers (ITE). (2017). *Trip Generation Manual, 10th Edition, Volume 2 — Data Plots*. Retrieved from https://itetripgen.org/index.html

Jackson, K. T. (1987). The Drive-in Culture of Contemporary America. In *Crabgrass Frontier: The Suburbanization of the United States*. Oxford University Press.

Jacobs-Crisioni, C., Rietveld, P., & Koomen, E. (2014). The impact of spatial aggregation on urban development analyses. *Applied Geography*, *47*, 46–56. https://doi.org/10.1016/j.apgeog.2013.11.014

Janssen-Jansen, L., Spaans, M., & van der Veen, M. (2009). New instruments in spatial planning. *An International Perspective on Non-Financial Compensation., Amsterdam*.

Järv, O., Tenkanen, H., & Toivonen, T. (2017). Enhancing spatial accuracy of mobile phone data using multi-temporal dasymetric interpolation. *International Journal of Geographical Information Science*, *31*(8), 1630–1651. https://doi.org/10.1080/13658816.2017.1287369

Johnson, P. A., & Sieber, R. E. (2012). Motivations driving government adoption of the Geoweb. *GeoJournal*, *77*(5), 667–680. https://doi.org/10.1007/s10708-011-9416-8

Jonas, A. E. G. (1991). URBAN GROWTH COALITIONS AND URBAN DEVELOPMENT POLICY: POSTWAR GROWTH AND THE POLITICS OF ANNEXATION IN METROPOLITAN COLUMBUS. *Urban Geography*, *12*(3), 197–225. https://doi.org/10.2747/0272-3638.12.3.197

Jung, Y.-J., & Casello, J. M. (2019). GIS-based transit trip allocation methods converting stop-level boarding and alighting trips into TAZ trips. *Transportation Planning and Technology*, *42*(8), 848–867. https://doi.org/10.1080/03081060.2019.1675321

Kasarda, J. D. (1980). The implications of contemporary redistribution trends for national urban policy. *Social Science Quarterly*, *61*(3/4), 373–400. Retrieved from http://www.jstor.org/stable/42860760

Kasraian, D., Maat, K., Stead, D., & van Wee, B. (2016). Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies. *Transport Reviews*, 1–21.

Kasraian, D., Raghav, S., & Miller, E. J. (2020). A multi-decade longitudinal analysis of transportation and land use co-evolution in the Greater Toronto-Hamilton Area. *Journal of Transport Geography*, *84*, 102696. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102696

Kenworthy, J. (2007). Urban Planning and Transport Paradigm Shifts for Cities of the Post-Petroleum Age. *Journal of Urban Technology*, *14*(2), 47–70. https://doi.org/10.1080/10630730701531708

Kenworthy, J., Laube, F., Newman, P., Barter, P., Raad, T., Poboon, C., & Guia Jr, B. (2000). *An international sourcebook of automobile dependence in cities 1960-1990*.

Kim, N., & Yoon, Y. (2021). Regionalization for urban air mobility application with analyses of 3D urban space and geodemography in San Francisco and New York. *Procedia Computer Science*, *184*, 388–395. https://doi.org/https://doi.org/10.1016/j.procs.2021.03.049

Knaap, G. J., Ding, C., & Hopkins, L. D. (2001). Do Plans Matter?: The Effects of Light Rail Plans on Land Values in Station Areas . *Journal of Planning Education and Research* , *21*(1), 32–39. https://doi.org/10.1177/0739456X0102100103

Knight, R. L., & Trygg, L. L. (1977). Evidence of land use impacts of rapid transit systems. *Transportation*, *6*(3), 231–247. https://doi.org/10.1007/BF00177453

Kok, K., & Veldkamp, A. (2001). Evaluating impact of spatial scales on land use pattern analysis in Central America. *Agriculture, Ecosystems & Environment*, *85*(1–3), 205–221. https://doi.org/10.1016/S0167-8809(01)00185-2

Koomen, E., & Borsboom-van Beurden, J. (2011). *Land-use modelling in planning practice* (Vol. 101). Springer Science & Business Media.

Krueger, R., & Gibbs, D. (2008). 'Third Wave' Sustainability? Smart Growth and Regional Development in the USA. *Regional Studies*, *42*(9), 1263–1274. https://doi.org/10.1080/00343400801968403

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, *102*(5), 958–968.

Lee, R., & Sener, I. N. (2017). The effect of light rail transit on land-use development in a city without zoning. *Journal of Transport and Land Use; Forthcoming Papers*. Retrieved from https://www.jtlu.org/index.php/jtlu/article/view/926/877

Lee, S.-I., Lee, M., Chun, Y., & Griffith, D. A. (2019). Uncertainty in the effects of the modifiable areal unit problem under different levels of spatial autocorrelation: a simulation study. *International Journal of Geographical Information Science*, *33*(6), 1135–1154. https://doi.org/10.1080/13658816.2018.1542699

Lowry, I. S. (1964). A Model of Metropolis. In *Rand Corporation, Santa Monica, CA*.

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy*, *20*, 105–113. https://doi.org/10.1016/j.tranpol.2012.01.013

Maravalle, M., & Simeone, B. (1995). A spanning tree heuristic for regional clustering. *Communications in Statistics - Theory and Methods*, *24*(3), 625–639. https://doi.org/10.1080/03610929508831512

Martens, K. (2016). *Transport justice: Designing fair transportation systems*. Routledge.

McCann, P. (1998). The Location of the Firm in Theory. In *The Economics of Industrial Location: A Logistics-Costs Approach* (pp. 17–47). https://doi.org/10.1007/978-3-662-03702-7_3

McDonald, J. F. (1987). The identification of urban employment subcenters. *Journal of Urban Economics*, *21*(2), 242–258. https://doi.org/http://dx.doi.org/10.1016/0094-1190(87)90017-9

McFadden, D. (1978). *Modelling the choice of residential location*. Institute of Transportation Studies, University of California.

McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics*, *50*(3), 448–473. https://doi.org/10.1006/juec.2001.2228

McMillen, D. P. (2003). Identifying Sub-centres Using Contiguity Matrices. *Urban Studies*, *40*(1), 57–69. https://doi.org/10.1080/00420980220080161

McMillen, D. P., & McDonald, J. F. (1998). Suburban Subcenters and Employment Density in Metropolitan Chicago. *Journal of Urban Economics*, *43*(2), 157–180. https://doi.org/http://dx.doi.org/10.1006/juec.1997.2038

Miller, E. J. (2003). Microsimulation. In K. Goulias & B. Raton (Eds.), *Transportation Systems Planning Methods and Applications*. CRC Press.

Miller, E. J. (2004). Integrated land use/transport model requirements. In *Handbook of transport geography and spatial systems* (pp. 147–165). Emerald Group Publishing Limited.

Miller, E. J. (2014). Transportation Models. In *Contributions to Economic Analysis*: *Vol. 293*.

Handbook of Microsimulation Modelling (pp. 13–385). https://doi.org/doi:10.1108/S0573-855520140000293012

Miller, E. J. (2020). *Chapter 3 - Travel demand models, the next generation: Boldly going where no-one has gone before* (K. G. Goulias & A. W. B. T.-M. the T. B. G. Davis, Eds.). https://doi.org/https://doi.org/10.1016/B978-0-12-817340-4.00003-6

Miller, E. J., Douglas Hunt, J., Abraham, J. E., & Salvini, P. A. (2004). Microsimulating urban systems. *Computers, Environment and Urban Systems*, *28*(1–2), 9–44. https://doi.org/10.1016/S0198-9715(02)00044-3

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 197–210.

MMAH, M. of M. A. and H. *Growth Plan for the Greater Golden Horseshoe.* , (2017).

Moekel, R., Spiekermann, K., Schürmann, C., & Wegener, M. (2003). Microsimulation of Land Use. *International Journal of Urban Sciences*, *7*(1), 14–31. https://doi.org/10.1080/12265934.2003.9693520

Moore, A. (2002). *The case for approximate Distance Transforms*. https://doi.org/10.1.1.98.5245

Moreno, N., Wang, F., & Marceau, D. J. (2009). Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model. *Computers, Environment and Urban Systems*, *33*(1), 44–54. https://doi.org/10.1016/j.compenvurbsys.2008.09.008

MPAC, M. P. A. C. *Ontario Designated Property Codes.* , (2020).

Muller, P. O. (2004). Transportation and urban form: stages in the spatial evolution of the American metropolis. In S. Hanson & G. Giuliano (Eds.), *The Geography of Urban Transportation* (Third).

Muth, R. F. (1969). *Cities and housing, the University of Chicago press, Chicago*.

Newman, P. W. G., & Kenworthy, J. R. (1996). The land use—transport connection. *Land Use Policy*, *13*(1), 1–22. https://doi.org/http://dx.doi.org/10.1016/0264-8377(95)00027-5

Nour, A., Hellinga, B., & Casello, J. (2016). Classification of automobile and transit trips from Smartphone data: Enhancing accuracy using spatial statistics and GIS. *Journal of Transport Geography*, *51*, 36–44.

Nourian, F., Alipour, M., & Ache, P. (2021). Model-theory interaction in urban planning: A critical review. *Planning Theory*, 14730952211026692. https://doi.org/10.1177/14730952211026693

Nowosad, J., & Stepinski, T. F. (2018). Spatial association between regionalizations using the information-theoretical V-measure. *International Journal of Geographical Information Science*, *32*(12), 2386–2401. https://doi.org/10.1080/13658816.2018.1511794

Openshaw, S. (1977). A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling. *Transactions of the Institute of British Geographers*, 459–472.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem. Concepts and Techniques in Modern Geography*. Geo Books.

Ortigoza, G. M. (2015). Unstructured triangular cellular automata for modeling geographic spread. *Applied Mathematics and Computation*, *258*, 520–536. https://doi.org/10.1016/j.amc.2015.01.116

Páez, A., & Scott, D. M. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, *61*(1), 53–67. https://doi.org/10.1007/s10708-005-0877-5

Pagliara, F., & Wilson, A. (2010). *The State-of-the-Art in Building Residential Location Models*. https://doi.org/10.1007/978-3-642-12788-5_1

Parker, D. C., Berger, T., & Manson, S. M. (2001). Agent-based models of land-use and land-cover change: report and review. *International Workshop, October 4-7, 2001, Irvine, California, USA*.

Parkin, M., Parker, D., Casello, J., Moos, M., Jin, X., Babin, R., … Yeung, K. (2016). *Central Transit Corridor Monitoring Program: Baseline Monitoring Report*. Region of Waterloo.

Pereira, R. H. M., Banister, D., Schwanen, T., & Wessel, N. (2019). Distributional effects of transport policies on inequalities in access to opportunities in Rio de Janeiro. *Journal of Transport and Land Use*, *12*(1 SE-). https://doi.org/10.5198/jtlu.2019.1523

Pereira, R. H. M., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews*, *37*(2), 170–191. https://doi.org/10.1080/01441647.2016.1257660

Pfaffenbichler, P., Emberger, G., & Shepherd, S. (2010). A system dynamics approach to land use transport interaction modelling: the strategic model MARS and its application. *System Dynamics Review*, *26*(3), 262–282. https://doi.org/10.1002/sdr.451

Pritchard, D. R., & Miller, E. J. (2012). Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. *Transportation*, *39*(3), 685–704. https://doi.org/10.1007/s11116-011-9367-4

Putman, S. H. (2010). *DRAM Residential Location and Land Use Model: 40 Years of Development and Application*. https://doi.org/10.1007/978-3-642-12788-5_3

Qian, Z. (2010). Without zoning: Urban development and land use controls in Houston. *Cities*, *27*(1), 31–41. https://doi.org/10.1016/j.cities.2009.11.006

Reese, L. A., Larnell, T. B., & Sands, G. (2009). Patterns of Tax Abatement Policy: Lessons From the Outliers? *The American Review of Public Administration*.

RMOW, R. M. of W. (2003). *Regional Growth Management Strategy (RGMS)*. Retrieved from https://www.regionofwaterloo.ca/en/resources/RegionalGrowthManagementStrategy.pdf

RMOW, R. M. of W. *Regional Official Plan (ROP)*. , (2015).

RMOW, R. M. of W. *Moving Forward 2018 Transportation Master Plan*. , (2018).

Roberts, S. A., Hall, G. B., & Calamai, P. H. (2011). Evolutionary Multi-objective Optimization for landscape system design. *Journal of Geographical Systems*, *13*(3), 299–326. https://doi.org/10.1007/s10109-010-0136-2

Ryan, J., Maoh, H., & Kanaroglou, P. (2009). Population Synthesis: Comparing the Major Techniques Using a Small, Complete Population of Firms. *Geographical Analysis*, *41*(2), 181–203. https://doi.org/10.1111/j.1538-4632.2009.00750.x

Seber, G. A. (2009). Multivariate observations, vol. 252. *Hoboken, New Jersey, USA: John Wiley & Sons*.

Shahab, S., Clinch, J. P., & O'Neill, E. (2018). Estimates of Transaction Costs in Transfer of

Development Rights Programs. *Journal of the American Planning Association*, *84*(1), 61–75. https://doi.org/10.1080/01944363.2017.1406816

Siemiatycki, M. (2006). Implications of Private-Public Partnerships on the Development of Urban Public Transit Infrastructure. *Journal of Planning Education and Research*, *26*(2), 137–151. https://doi.org/10.1177/0739456X06291390

Singh, Y. J., Fard, P., Zuidgeest, M., Brussel, M., & Maarseveen, M. van. (2014). Measuring transit oriented development: a spatial multi criteria assessment approach for the City Region Arnhem and Nijmegen. *Journal of Transport Geography*, *35*(0), 130–143. https://doi.org/10.1016/j.jtrangeo.2014.01.014

Singleton, P. A., & Clifton, K. J. (2013). Pedestrians in regional travel demand forecasting models: State-of-the-practice. *92nd Annual Meeting of the Transportation Research Board, Washington, DC*, 13–4857.

Small, K. A., & Song, S. (1994). Population and employment densities: structure and change. *Journal of Urban Economics*, *36*, 292–313.

Soteropoulos, A., Berger, M., & Ciari, F. (2019). Impacts of automated vehicles on travel behaviour and land use: an international review of modelling studies. *Transport Reviews*, *39*(1), 29–49. https://doi.org/10.1080/01441647.2018.1523253

Southworth, M., & Owens, P. M. (1993). The Evolving Metropolis: Studies of Community, Neighborhood, and Street Form at the Urban Edge. *Journal of the American Planning Association*, *59*(3), 271–287. https://doi.org/10.1080/01944369308975880

Stanilov, K. (2012a). Space in Agent-Based Models. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-Based Models of Geographical Systems*. https://doi.org/10.1007/978-90-481-8927-4

Stanilov, K. (2012b). Space in Agent-Based Models. In *Agent-Based Models of Geographical Systems* (pp. 253–269). https://doi.org/10.1007/978-90-481-8927-4_13

Statistics Canada. (2021a). *The Building Construction Price Index (BCPI)*. Retrieved from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2317

Statistics Canada. (2021b). *The New Housing Price Index (NHPI)*. Retrieved from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2310

Stull, W. J. (1974). Land Use and Zoning in an Urban Economy. *The American Economic Review*, *64*(3), 337–347. Retrieved from http://www.jstor.org/stable/1808886

Sweet, M. N., Bullivant, B., & Kanaroglou, P. S. (2017). Are major canadian city-regions monocentric, polycentric, or dispersed? *Urban Geography*, *38*(3), 445–471. https://doi.org/10.1080/02723638.2016.1200279

Sweet, M. N., & Scott, D. M. (2021). Shared mobility adoption from 2016 to 2018 in the Greater Toronto and Hamilton Area: Demographic or geographic diffusion? *Journal of Transport Geography*, *96*, 103197. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2021.103197

Timmermans, H. (2003). The saga of integrated land use-transport modeling: how many more dreams before we wake up? *Keynote Paper, Moving through Nets: The Physical and Social Dimension of Travel, 10th International Conference on Travel Behaviour Research, Lucerna, Www. Ivt.*

Baug. Ethz. Ch/Allgemein/Pdf/Timmermans. Pdf.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234-240 CR-Copyright © 1970 Clark Universi. https://doi.org/10.2307/143141

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. In *General Assembley 70 session*.

van Delden, H., van Vliet, J., Rutledge, D. T., & Kirkby, M. J. (2011). Comparison of scale and scaling issues in integrated land-use models for policy support. *Agriculture, Ecosystems & Environment*, *142*(1–2), 18–28. https://doi.org/10.1016/j.agee.2011.03.005

van der Veen, M., Spaans, M., & Janssen-Jansen, L. (2010). Using compensation instruments as a vehicle to improve spatial planning: Challenges and opportunities. *Land Use Policy*, *27*(4), 1010–1017. https://doi.org/https://doi.org/10.1016/j.landusepol.2010.01.003

van Schrojenstein Lantman, J., Verburg, P. H., Bregt, A., & Geertman, S. (2011). *Core Principles and Concepts in Land-Use Modelling: A Literature Review*. https://doi.org/10.1007/978-94-007-1822-7_3

Vliet, J. van, White, R., & Dragicevic, S. (2009). Modeling urban growth using a variable grid cellular automaton. *Computers, Environment and Urban Systems*, *33*(1), 35–43. https://doi.org/10.1016/j.compenvurbsys.2008.06.006

Voinov, A. (2008). Sensitivity, Calibration, Validation, Verification. In *Encyclopedia of Ecology, Volumes 1-5*. Elsevier.

von Thünen, J. H. (1826). Der isolierte Staat in Beziehung auf Nationalökonomie und Landwirtschaft. *Gustav Fischer, Stuttgart (Reprinted 1966)*.

Waddell, P. (2011). Integrated Land Use and Transportation Planning and Modelling: Addressing Challenges in Research and Practice. *Transport Reviews*, *31*(2), 209–229. https://doi.org/10.1080/01441647.2010.525671

Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M., & Ulfarsson, G. (2003). Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Networks and Spatial Economics*, *3*(1), 43–67. https://doi.org/10.1023/A:1022049000877

Waddell, P., Ulfarsson, G. F., Franklin, J. P., & Lobb, J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice*, *41*(5), 382–410.

Weber, A., & Friedrich, C. J. (1929). *Alfred Weber's theory of the location of industries.* Chicago, Ill.: The University of Chicago Press.

Wegener, M. (1994). Operational Urban Models State of the Art. *Journal of the American Planning Association*, *60*(1), 17–29. https://doi.org/10.1080/01944369408975547

Wegener, M. (1998). *Applied Models of Urban Land Use, Transport and Environment: State of the Art and Future Developments*. https://doi.org/10.1007/978-3-642-72242-4_14

Wegener, M. (2001). New spatial planning models. *International Journal of Applied Earth Observation and Geoinformation*, *3*(3), 224–237. https://doi.org/10.1016/S0303-2434(01)85030-3

Wegener, M. (2004). Overview Of Land-Use Transport Models. In *Transport Geography and Spatial Systems*. Pergamon/Elsevier Science,.

Wegener, M. (2014). Land-Use Transport Interaction Models. In *Handbook of Regional Science*. Springer.

Wegener, M., Gnad, F., & Vannahme, M. (1986). *The time scale of urban change*. IRPUD.

Wei, R., Rey, S., & Knaap, E. (2021). Efficient regionalization for spatially explicit neighborhood delineation. *International Journal of Geographical Information Science*, *35*(1), 135–151.

Wilson, A. G. (1971). A Family of Spatial Interaction Models, and Associated Developments. *Environment and Planning A*, *3*(1), 1–32. https://doi.org/10.1068/a030001

Wolf, L. J., Knaap, E., & Rey, S. (2019). Geosilhouettes: Geographical measures of cluster fit. *Environment and Planning B: Urban Analytics and City Science*, *48*(3), 521–539. https://doi.org/10.1177/2399808319875752

Wrigley, N., Holt, T., Steel, D., & Tranmer, M. (1996). Analysing, modelling, and resolving the ecological fallacy. *Spatial Analysis: Modelling in a GIS Environment*, 23–40.

Wu, L., Yang, B., & Jing, P. (2016). *Travel mode detection based on GPS raw data collected by smartphones: a systematic review of the existing methodologies*.

Xu, X. M., Casello, J. M., & Fard, P. (2021). *Capturing Pedestrian Tours and Activities through Smartphone Data*.

Yeates, M. (1990). *The North American city*. Harper & Row.

Yoshida, T., & Tanaka, K. (2005). Land-use diversity index: a new means of detecting diversity at landscape level. *Landscape and Ecological Engineering*, *1*(2), 201–206.

You, J., Nedović-Budić, Z., & Kim, T. J. (1997). A GIS-based traffic analysis zone design: technique. *Transportation Planning and Technology*, *21*(1–2), 45–68. https://doi.org/10.1080/03081069708717601

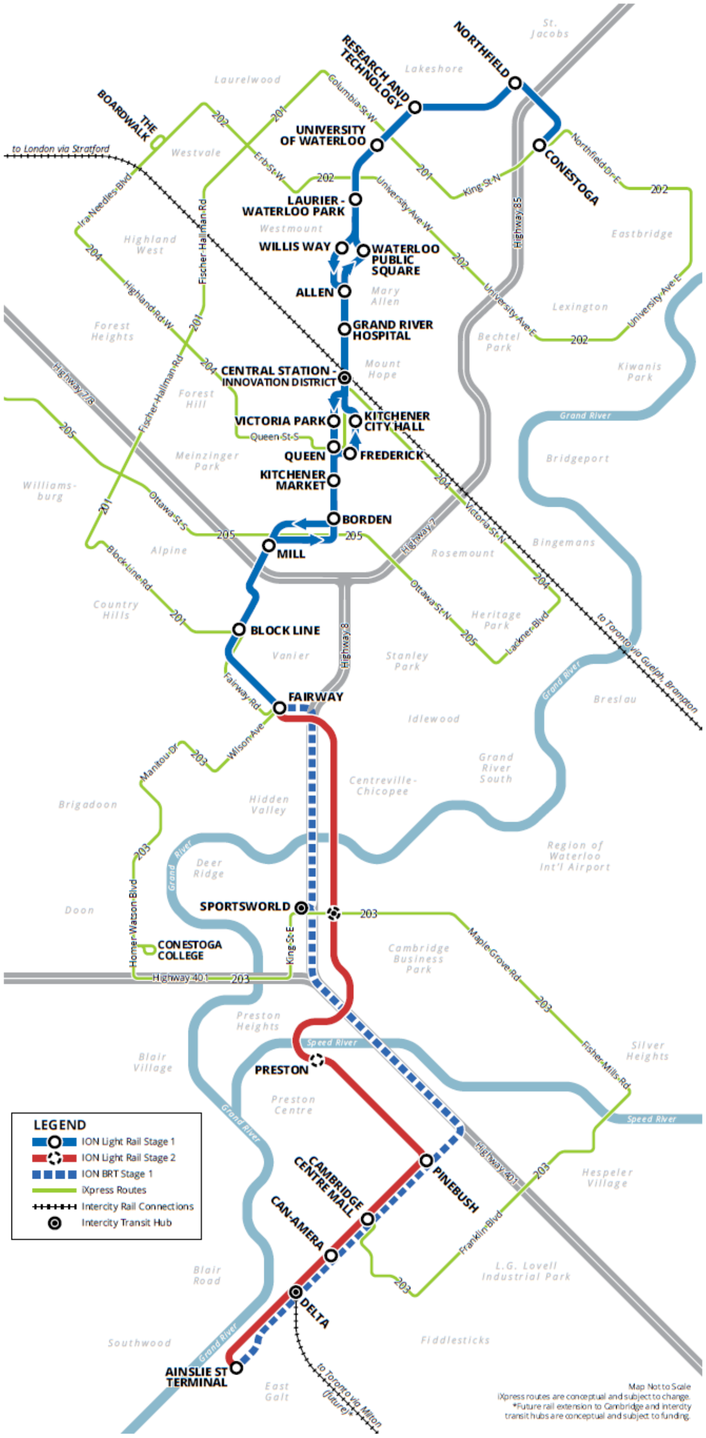
You, J., Nedović-Budić, Z., & Kim, T. J. (1998). A GIS-based traffic analysis zone design: implementation and evaluation. *Transportation Planning and Technology*, *21*(1–2), 69–91. https://doi.org/10.1080/03081069708717602

Zahabi, S. A. H., & Patterson, Z. (2016). *Towards Transit Trip Itinerary Inference from Smartphone Data: A Case Study from Montreal, Canada*.

Zhang, Y., Song, R., van Nes, R., He, S., & Yin, W. (2019). Identifying Urban Structure Based on Transit-Oriented Development. *Sustainability* , Vol. 11. https://doi.org/10.3390/su11247241

Zhong, C., Schläpfer, M., Müller Arisona, S., Batty, M., Ratti, C., & Schmitt, G. (2015). Revealing centrality in the spatial structure of cities from human activity patterns. *Urban Studies*, *54*(2), 437–455. https://doi.org/10.1177/0042098015601599

Appendix A



Map Not to Scale
iXpress routes are conceptual and subject to change.
*Future rail extension to Cambridge and intercity transit hubs are conceptual and subject to funding.

Figure A--5-1 - ION Rapid Transit System Route Map

Appendix B

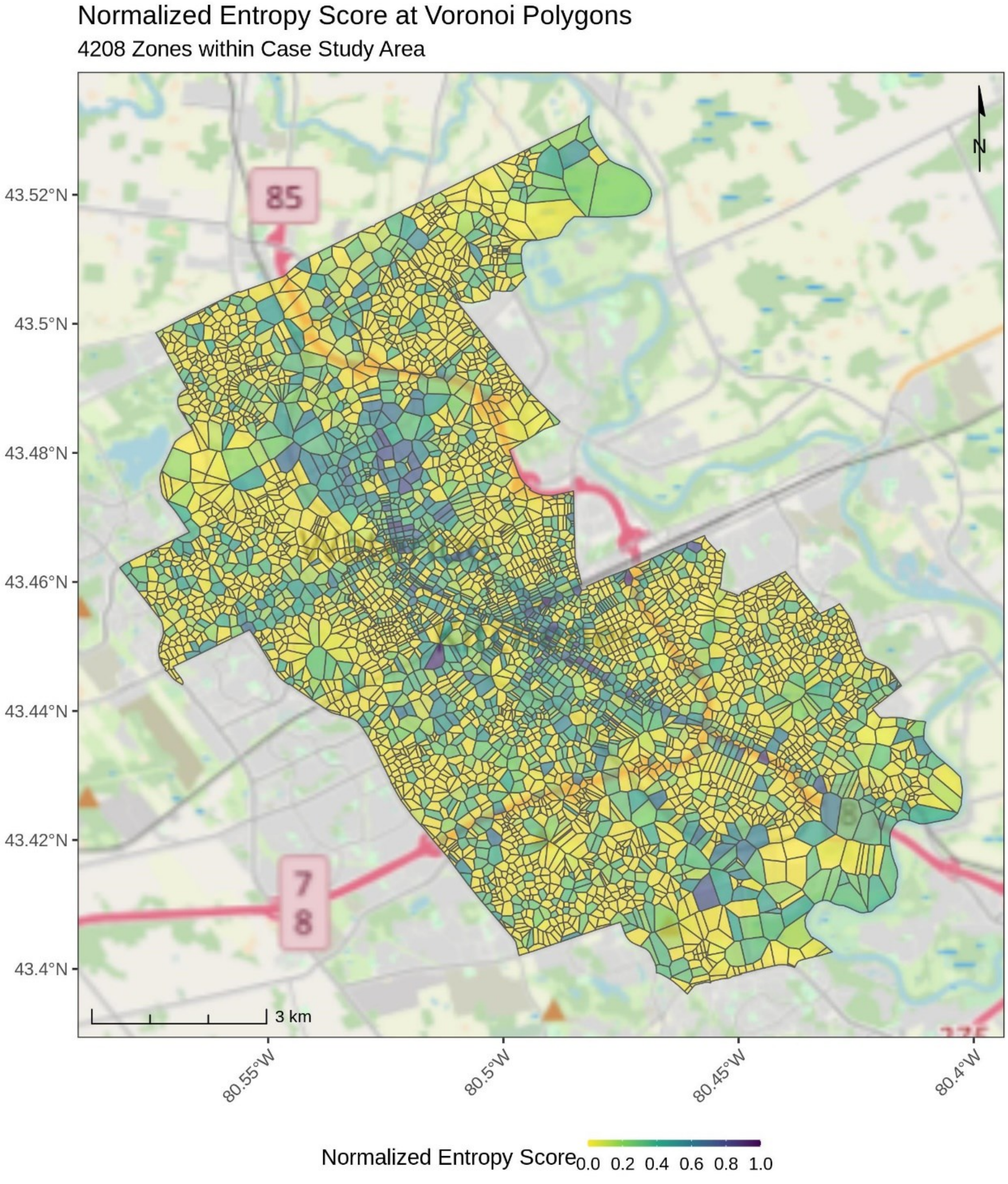


Figure A--5-2 - Spatial Distribution of Normalized Entropy Score for the Voronoi polygons

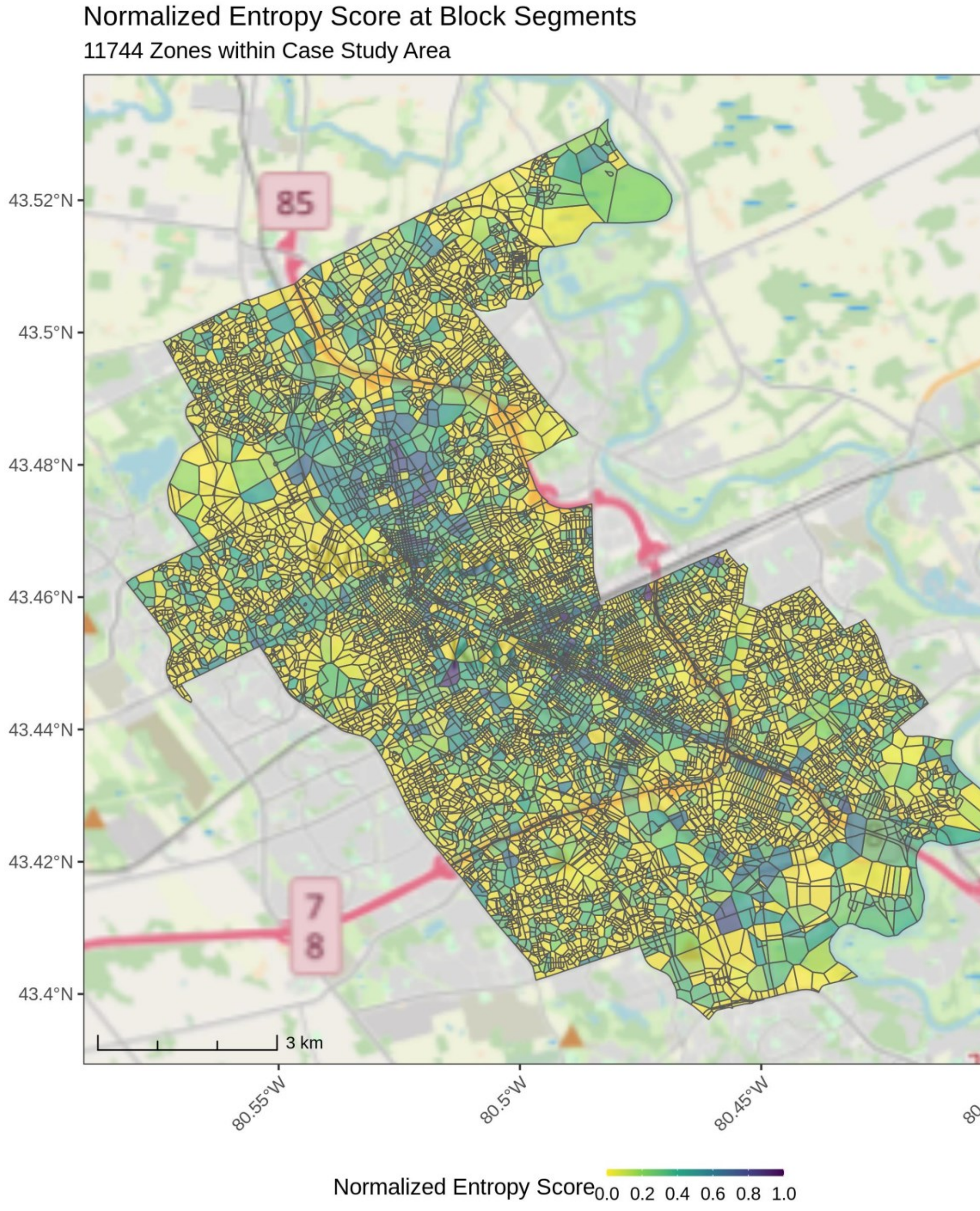


Figure A--5-3 - Spatial Distribution of Normalized Entropy Score for the Block Segments

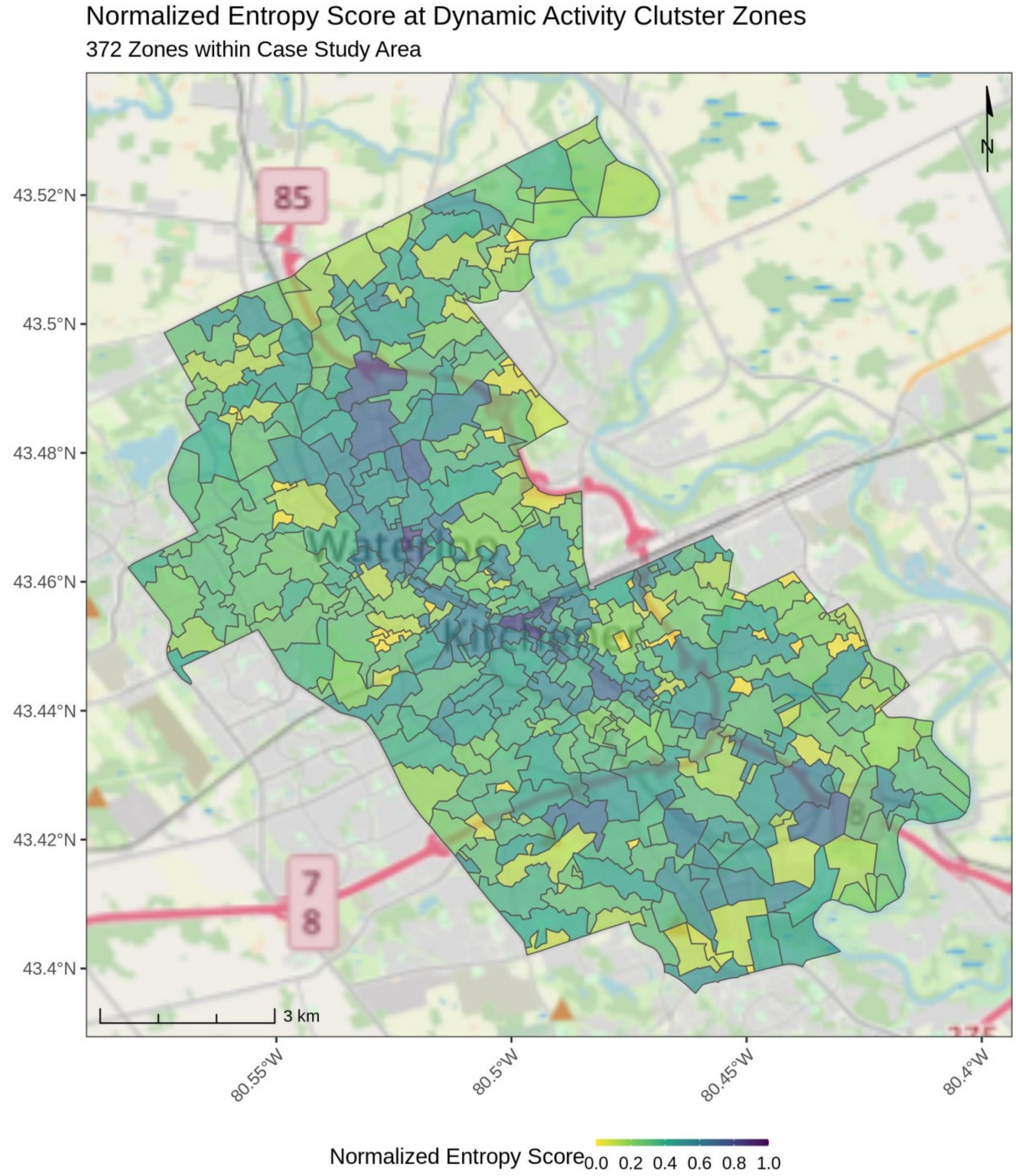


Figure A--5-4 - DACZ created based on the Voronoi polygons

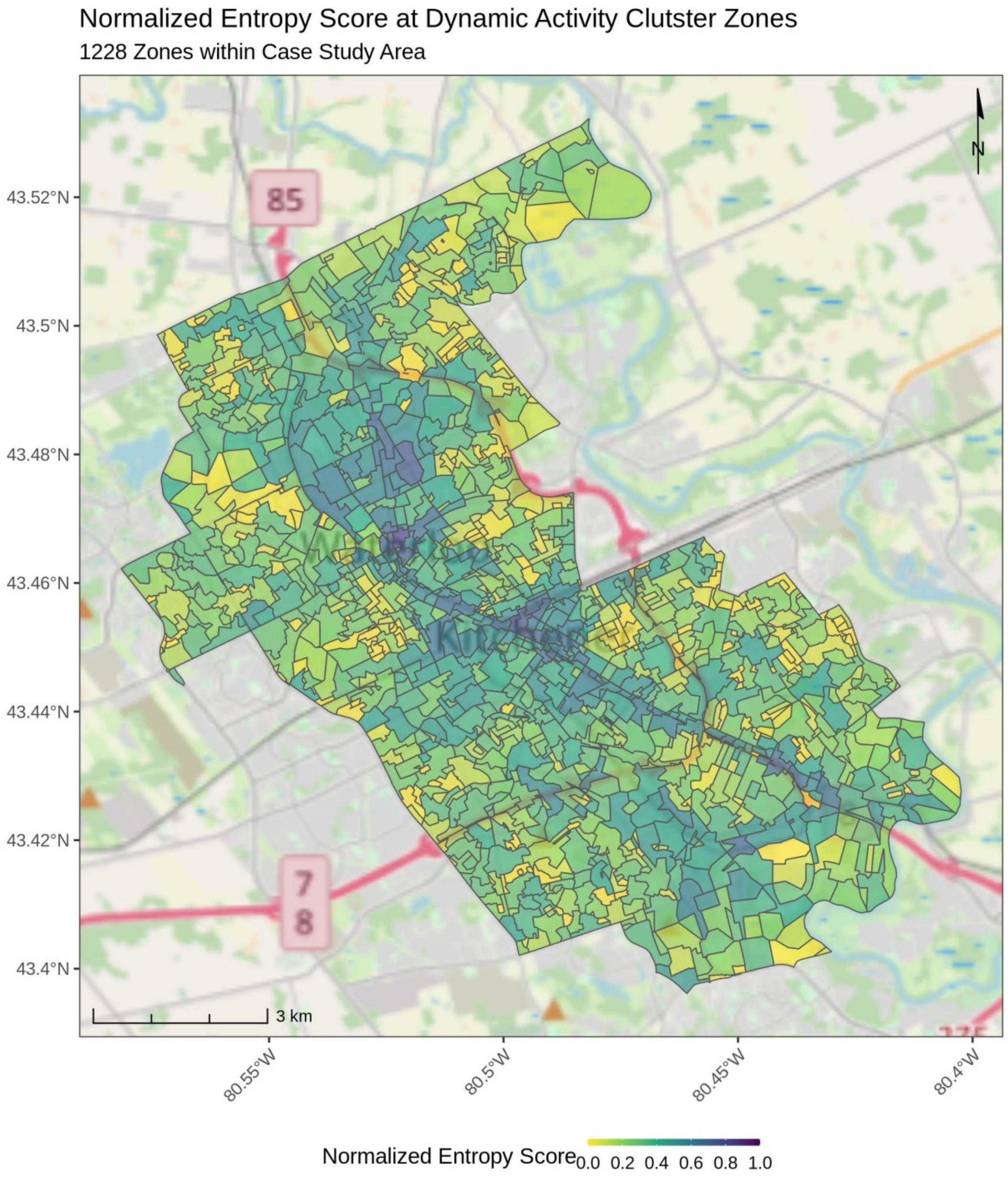


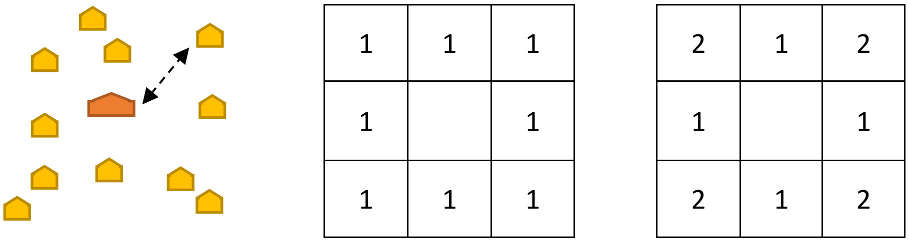
Figure A--5-5 - DACZ created based on the Block Segment polygons

Appendix C

Neighborhood Definition and Distance Metrics Relationship

The notion of spatial adjacency and neighborhood definition are the central concepts for quantitative spatial modeling. They determine the way a spatial model abstracts real world phenomena and captures the potential interactions of spatial elements within a landscape. Neighborhood definition in spatial models relies on the First Law of Geography that states “Everything is related to everything else, but near things are more related than distant things” (Tobler, 1970). This theory is built on the notion of distance as it applies to the geographic space. Euclidean Distance – “defined as the length of a straight line linking the fixed geographical positions of two objects” (Gatrell, 1983) – is the most commonly used measure for representing *continuous* geographic distance as it intuitively perceived (Figure A-6a.). However, within the context of aggregate spatial modeling, alternative distance metrics need to be used to reflect the *discrete* length unit corresponding to the spatial arrangement of entities (i.e., zones) (Boots, Okabe, & Thomas, 2003). Chessboard Distance (Figure A-6b.) and Manhattan Distance (Figure A-6c.) are among the well-established discrete distance metrics that are applied within different grid based spatial models (Moore, 2002). These distance measures rely on the topological relationships between two objects and take into account the number of shared corner points and the common edges of entities (i.e., zones or grid cells) to calculate distances.

The DACZ model developed in this study uses the Chessboard distance metric to determine higher degree of spatial adjacency.



a) Euclidean Distance

b) Chessboard Distance

1	1	1
1		1
1	1	1

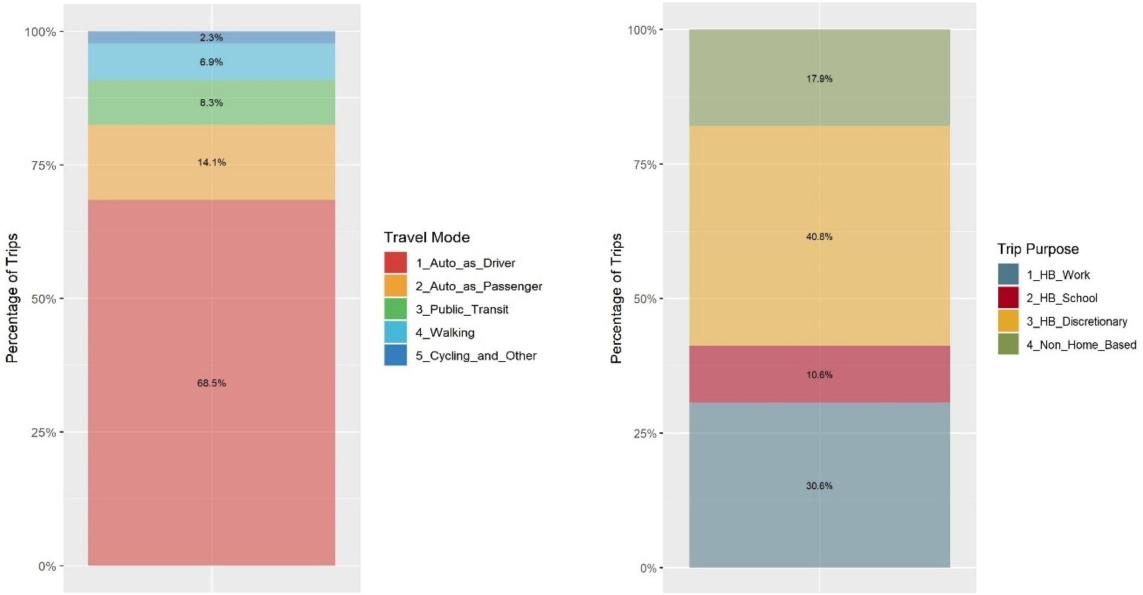
c) Manhattan Distance

2	1	2
1		1
2	1	2

Figure A--5-6 - Different distance metrics and their spatial propagation

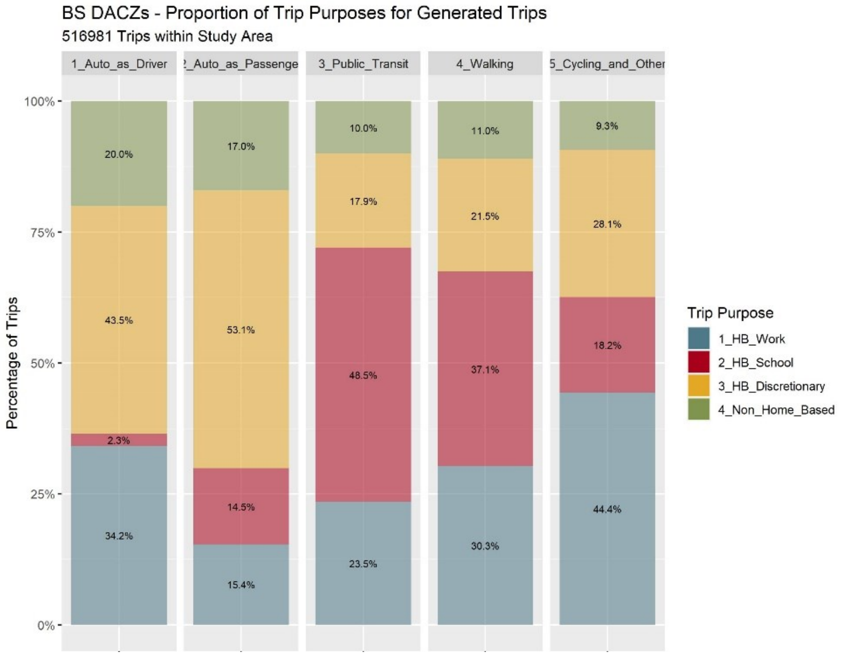
Appendix D

Travel Behavior Heterogeneity Distributions



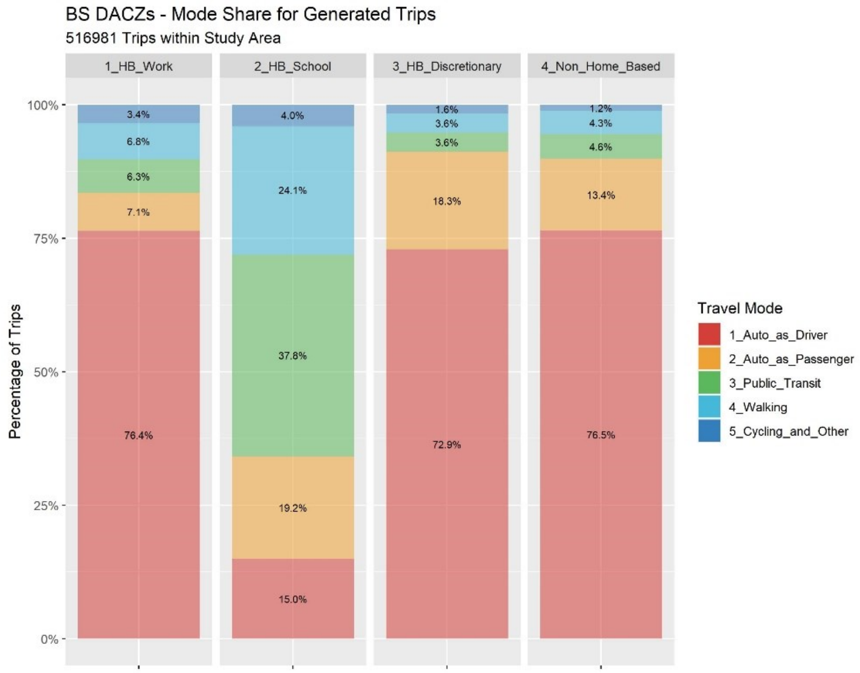
(a) Mode share for generated trips

(b) Proportion of trip purposes for generated trips

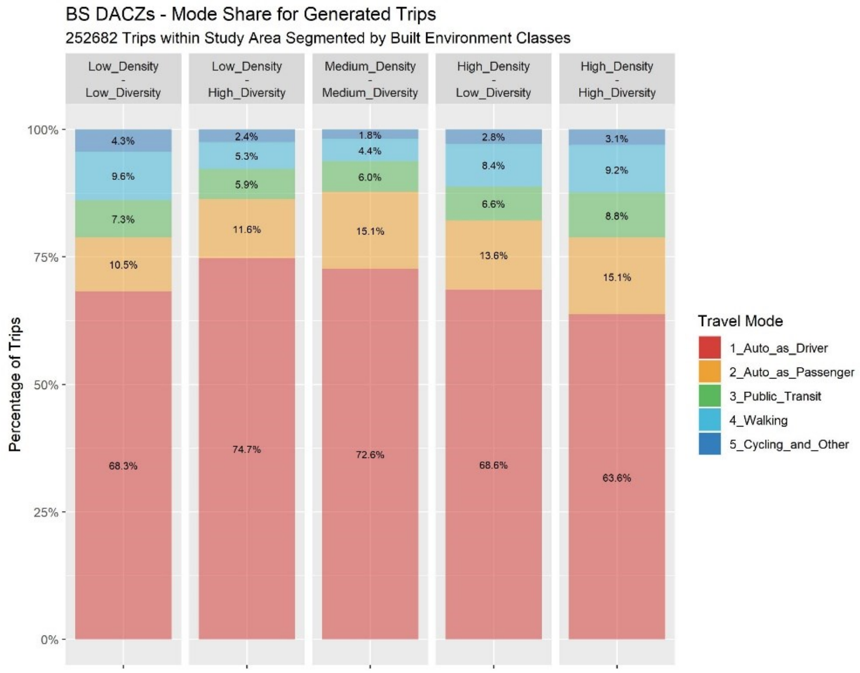


(c) Aggregate trip purpose statistics segmented by travel mode

Figure A--5-7 - Aggregate Mode Share and Trip Purposes Statistics within the Study Area



(a) Aggregate mode share statistics segmented by trip purposes



(b) Aggregate mode share statistics segmented by built environment categories

Figure A--5-8 - Aggregate Mode Share within the Study Area