A DATA-DRIVEN SYSTEMATIC APPROACH TO IDENTIFY, CLASSIFY, AND ESTIMATE LONG-HAUL FREIGHT TRUCK PARKING SUPPLY

ERIK A NEVLAND

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

GRADUATE PROGRAM IN CIVIL ENGINEERING
YORK UNIVERSITY
TORONTO, ON

January 2020

ABSTRACT

Collisions involving trucks have increased fatality risks compared to passenger vehicles. Hours-of-Service (HOS) laws exist to reduce fatalities where truck driver fatigue is a contributing factor. Electronic logging devices (ELD) are being mandated to automatically track HOS and enforce compliance, creating a greater urgency for adequate truck parking. A lack of truck parking is often identified throughout North America; however, these studies are often limited to public rest areas despite evidence that drivers often utilize other types of parking.

To adequately compare truck parking supply and demand, an exhaustive truck parking classification scheme is developed based on important location attributes identified through extensive literature review. This scheme can be systematically implemented using available geospatial data. This data is then used to develop a truck parking supply model based on a negative binomial regression. The Region of Peel is used as the study area due to its considerably large freight industry.

DEDICATION

To my wife and better half,

Nadine Nevland,

for without her love, support, patience, and understanding

this work would not be possible.

ACKNOWLEDGEMENTS

I am extremely grateful for the instruction, guidance, encouragement, and cooperation of my supervisors during the two years of work on this research. I thank Dr. Peter Park for his years of support and for introducing me to transportation engineering in my undergraduate studies at the University of Saskatchewan. I appreciate the many opportunities that he has opened for me and am thankful that he extended the invitation for me to follow him to Ontario. I also thank Dr. Kevin Gingerich for taking me on as his first graduate student and for passing on his enthusiasm for goods movement. I appreciate the many times he worked with me late into the night, sometimes into the early morning, to help me with finishing submissions.

Special thanks and appreciation go to the Region of Peel, the City of Mississauga, the City of Brampton, the Town of Caledon, and the Smart Freight Centre for their data and support. I am also thankful for financial contributions from the Natural Sciences and Engineering Research Council (NSERC).

I am thankful to Dr. Sabbir Saiyed (Master's Thesis Supervisory Committee member), Roman Kuczynski, Sabrina Khan, Parshan Bahrami, and my other colleagues at the Region of Peel for their support in this research. I am proud to now be a part of the Region's Transportation System Planning section.

I would like to thank my colleagues in the Transportation Research Group at York University: Dr. Seun Daniel Oluwajana, Crystal Mingyue Wang, MD Tanvir Chowdhury, Ravichandra Rampure, Ahmad Mohammadi, Syed Ubaid Ali, Soha Saiyed, and Yashar Zarrin Zadeh. I am happy to call them my friends and very thankful for their collaboration and encouragement. The same regards go to my former colleague Justin Hart from the University of Saskatchewan. I appreciate the technical feedback from Dr. Mehdi Nourinejad, the internal member at my thesis defence. I am also thankful for the administrative support and guidance I received from Sindy Mahal, Gillian Moore, and Theresa Smellie.

I must acknowledge the opportunities I have received through the ITE York University Student Chapter. I am proud of the work we have done in forming this group and am confident that it will continue to inspire the next generation of transportation professionals. For this, I thank our many volunteers, the membership at large, and the Institute of Transportation Engineers.

Lastly, I appreciate the lifetime of support I have received from my friends and family. I appreciate everything my parents have done for me and am happy they always encouraged me to keep enthusiastically seeking answers. I would not have pursued this research if it were not for the encouragement of my wife, who thankfully has been incredibly supportive and patient during this endeavour.

Dominus dat sapientiam. Deo gratias.

TABLE OF CONTENTS

Abstract ... ii
Dedication ... iii
Acknowledgements ... iv
Table of Contents ... vi
List of Tables ... viii
List of Figures ... ix
List of Acronyms and Abbreviations ... x
I. Note on Formatting Conventions ... xi
Declarations ... xii
I. Co-Authorship Declaration ... xii
II. Declaration of Adopted Previous Works ... xiii
Chapter 1: Introduction ... 1
1.1 Research Motivation ... 1
1.2 Research Goal ... 3
1.3 Research Objectives ... 3
1.4 Research Scope ... 5
1.5 Thesis Structure ... 5
Chapter 2: Literature Review ... 7
2.1 Heavy Commercial Vehicle Parking ... 7
2.2 Collisions and Safety ... 15
2.3 Hours of Service ... 20
2.4 HCV Parking Modelling and Analysis ... 25
2.5 Important Characteristics of HCV Parking ... 28
2.6 Existing HCV Parking Classification Scheme ... 32
Chapter 3: Study Data ... 34
3.1 Case Study Area: Region of Peel ... 34
3.2 Data Sources ... 38
Chapter 4: Analysis Methods ... 47
4.1 Data Analysis Techniques ... 47
4.2 Regression Analysis ... 51
4.3 Goodness-of-Fit Tests ... 56
Chapter 5: Identification of HCV Parking Locations ... 60
5.1 Approach to Identifying HCV Parking ... 60
5.2 Identification of HCV Stop Events ... 61
5.3 Examining HCV Trips and Tours ... 62

5.4 Identification of HCV Rest Events....63
5.5 HCV Parking Locations....64
5.6 Verification of HCV Parking Location using Trip Purpose....67

Chapter 6: Development of an HCV Parking Classification Scheme....69
6.1 Classification Methodology....69
6.2 Identification of Key HCV Parking Characteristics....70
6.3 Analysis of HCV Parking Characteristic Combinations....71
6.4 Heuristic Stratification of HCV Parking Classifications....73

Chapter 7: Application of the HCV Parking Classification Scheme....76
7.1 Classification Approach for the Case Study Area....76
7.2 Results of Classification for the Case Study Area....83
7.3 Confirmation of Results....87

Chapter 8: Estimation of HCV Parking Supply....89
8.1 Modelling Approach....89
8.2 Scope for the HCV Parking Supply Model....90
8.3 Model Development....94
8.4 Candidate HCV Parking Supply Models....100
8.5 Selected HCV Parking Supply Model....108

Chapter 9: Conclusion....115
9.1 Summary....115
9.2 Contribution....116
9.3 Future Modelling Work....117
9.4 Future Planning Considerations....119
9.5 Concluding Statements....121

References....123

Appendices....136
Appendix A: History of Goods Movement....A-1
Appendix B: Details of HCV Stop Event Temporal Analysis....B-1
Appendix C: Details of Data Statistical Analysis....C-1
Appendix D: Numerical Goodness-of-Fit Methods....D-1
Appendix E: R Script for Data Analysis....E-1
Appendix F: Q-Q Plots....F-1
Appendix G: CURE Plots....G-1
Appendix H: HCV Parking Supply Spatial Autocorrelation....H-1

LIST OF TABLES

Table 2-1: Simplified Summary of Hours of Service Laws 21
Table 5-1: HCV GPS Ping Data Format 61
Table 5-2: Differences in HCV Rest Events by Scale 67
Table 6-1: Binary Conditions for Key HCV Parking Characteristics 71
Table 6-2: All Combinations of HCV Parking Characteristics 72
Table 6-3: HCV Parking Classification Scheme 74
Table 7-1: Indicating Variables for HCV Parking Location Characteristics 79
Table 7-2: Frequency Distribution of Property Parcels by Category 83
Table 8-1: Counted HCV Parking Spaces by Classification 93
Table 8-2: Selected Variables for Analysis 98
Table 8-3: Preliminary HCV Parking Models for Sample Size Determination 99
Table 8-4: Statistical Power Analysis and Sample Size Calculation 100
Table 8-5: Variable Combinations for Candidate HCV Parking Supply Models 101
Table 8-6: HCV Parking Supply Candidate Models 102
Table 8-7: Outliers Identified in Models 103
Table 8-8: HCV Parking Supply Candidate Models Goodness-of-Fit 105
Table 8-9: Zero-Prediction Accuracy of Zero-Inflated Models 107
Table 8-10: Total HCV Parking Spaces per TAZ in the Region of Peel 112

LIST OF FIGURES

Figure 1-1: Relationship between Research Objectives and Thesis Chapters 4
Figure 2-1: Vehicle Classification using FHWA 13-Category Scheme 9
Figure 2-2: WB-19 Design Vehicle Specifications 10
Figure 2-3: WB-19 Design Vehicle Turning Path 10
Figure 2-4: Number of Problematic HCV Parking Locations Reported in Canadian Provinces 13
Figure 2-5: Top Twenty Canadian Regions with Problematic Parking 14
Figure 2-6: Total Fatalities in Collisions involving HCVs 17
Figure 2-7: Fatalities in Collisions involving HCVs (per 100 Million Miles Driven) 18
Figure 2-8: HCV and Non-HCV Collision Statistics 19
Figure 2-9: Methods of Logging HOS 25
Figure 2-10: Atlanta Regional Commission Truck Parking Tiers 33
Figure 3-1: Region of Peel Case Study Area 35
Figure 3-2: 2019 HCV Traffic on Derry Road in the Region of Peel 36
Figure 3-3: Region of Peel Traffic Analysis Zones (TAZ) Map 38
Figure 3-4: Criteria for an Identified Stop Event 39
Figure 3-5: Interpretation of Entropy Index 41
Figure 4-1: Example of a Q-Q Plot 58
Figure 4-2: Example of a CURE Plot with Confidence Limits 59
Figure 5-1: Heatmap of Identified HCV Rest Events in the Region of Peel 64
Figure 5-2: GPS Data Showing Stop Events in Property Parcels 65
Figure 5-3: Heatmap of Identified HCV Parking Locations in the Region of Peel 66
Figure 7-1: HCV Parking Classification for the Region of Peel 80
Figure 7-2: Spatial Patterns of Parking Locations for Long-Haul HCV Trips 86
Figure 8-1: Visual Clues used to Manually Identify HCV Parking Spaces 91
Figure 8-2: HCV Parking Spaces Identified at a Warehouse Location 92
Figure 8-3: Histogram of Parcels Counted by Number of HCV Parking Spaces 94
Figure 8-4: Total HCV Parking Spaces per TAZ in the Region of Peel 113

LIST OF ACRONYMS AND ABBREVIATIONS

ATRI	American Transportation Research Institute
CAD	Computer-Aided Design
CAV	Connected and Autonomous Vehicles
CCMTA	Canadian Council of Motor Transport Administrators
CD	Census Divisions
CURE	Cumulative Residual
DOT	Department of Transportation
ELD	Electronic Logging Device
EPOI	Enhanced Points of Interest
FARS	Fatality Analysis Reporting System
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
GES	General Estimates System
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSA	General Services Administration
GSD	Ground Sample Distance
GTA	Greater Toronto Area
GTHA	Greater Toronto and Hamilton Area
HCV	Heavy Commercial Vehicle
HOS	Hours of Service
ICC	Interstate Commerce Commission
ITE	Institute of Transportation Engineers
ITS	Intelligent Transportation System
LISA	Local Indicators of Spatial Association
MAASTO	Mid America Association of State Transportation Officials
MAUP	Modifiable Area Unit Problem
MSA	Metropolitan Statistical Area

MTO	Ministry of Transportation Ontario
NAICS	North American Industry Classification System
NASEM	National Academies of Sciences, Engineering, and Medicine
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
OSHA	Occupational Safety and Health Administration
SIC	Standard Industrial Classification
SQL	Structured Query Language
STA	State Transportation Agency
TAZ	Traffic Analysis Zone
TDM	Transportation Demand Management
TPIMS	Truck Parking Information and Management System
TRI	Trucking Research Institute
TTS	Transportation Tomorrow Survey
V2I	Vehicle-to-Infrastructure

I. Note on Formatting Conventions

This thesis uses certain formatting conventions to improve clarity for the reader. All mathematical equations and variables, including those used in paragraphs, are formatted using italicized Cambria Math font (example: AIC and $P(y_i \mid \mu_i, \alpha)$). Equations are numbered sequentially throughout the thesis, including through the appendices. Coding functions and packages in R are formatted using Courier New font (example: `function()` and `Package`). Latin terms are formatted using italicised font (example: *a priori*, *post hoc*, and *et al.*). Lastly, the referencing system used in this thesis is the Harvard style, as outlined in "Cite Them Right: The Essential Referencing Guide" (Pears and Shields, 2016) and formatted using Mendeley citation manager (Mendeley Ltd., 2019).

DECLARATIONS

I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is the result of joint research, as follows: This thesis incorporates previous works, as described in the subsequent section, which were produced with and under the supervision of Dr. Peter Y. Park and Dr. Kevin Gingerich at York University. Segments of this research are located throughout this thesis. In all cases, the key ideas, primary contributions, data analysis and interpretation, and modelling efforts were performed by the author, and the contribution of co-authors was primarily conducted through the provision of supervision and final editing.

I am aware of the York University Faculty of Graduate Studies Policy on Intellectual Property for Graduate Programs and certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-authors to include the previously described material in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

Erik Nevland, EIT

II. Declaration of Adopted Previous Works

This thesis includes content adapted from the following conference proceedings and presentations:

Nevland, E. A., Gingerich, K. and Park, P. Y. (2019a) 'Development of a Classification Scheme and a Supply Model for Inter-Regional Truck Parking Facilities', in *98th Annual Meeting of the Transportation Research Board*. Washington, DC.

Nevland, E. A., Gingerich, K. and Park, P. Y. (2019b) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Canadian Institute of Transportation Engineers Annual Conference*. Ottawa, ON.

Nevland, E. A., Gingerich, K. and Park, P. Y. (2019c) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Joint Transportation Association of Canada / Intelligent Transportation Systems Canada Conference and Exhibition*. Halifax, NS.

Additionally, this thesis contains content adapted from a submitted journal article currently in press:

Nevland, E. A., Gingerich, K. and Park, P. Y. (2020) 'A Data-Driven Systematic Approach for Identifying and Classifying Long-Haul Truck Parking Locations', submitted for publication in *Transport Policy*. Washington, DC: Elsevier Ltd.

CHAPTER 1: INTRODUCTION

1.1 Research Motivation

According to the American Transportation Research Institute's (ATRI) most recent "Critical Issues in the Trucking Industry" survey, heavy commercial vehicle (HCV) parking and its associated issues are rated fifth highest among all freight industry concerns and rated third highest for commercial drivers (2019). The shortage of HCV parking is a serious freight transportation problem for three main reasons. Firstly, fatigued drivers may continue to drive until they find a safe and legal parking space. Reduced driver reaction times due to fatigue may contribute to an increase in HCV collisions in emergency situations (NASEM, 2016). Secondly, if additional driving is required to find a parking location, time is wasted and emissions are increased due to the additional fuel consumption (Kansas DOT and Turnpike Authority, 2016). Thirdly, drivers may eventually park in an unsafe and illegal location where the appropriate level of safety and security is not provided (Schmid *et al.*, 2018). HCVs can often be observed stopped on shoulders or highway ramps, leading to an unsafe road environment for both the driver and other roadway users (Boris and Johnson, 2015; Ioannou and de Almeida Araujo Vital, 2018).

Illegal parking is a particular concern from the perspective of driver health and safety. Boggs *et al*. (2019) showed a positive correlation between parking facilities reaching capacity and collisions on adjacent ramps subsequently used for overflow parking. When choosing to stop in an illegal parking location, drivers may not fully comprehend potential security threats such as vehicle thefts and other crimes (FHWA, 2015). A tragic incident occurred in 2009 when an HCV driver named Jason Rivenburg was killed while parked at an abandoned gas station in South Carolina (FHWA, 2015). Jason's Law was subsequently passed in the United States to address the lack of safe and legal parking spaces for HCVs. This law aims to increase the number of safe HCV parking spaces by constructing new public parking facilities and creating additional parking infrastructure at existing commercial truck stops and rest areas (USA: 112th Congress, 2012; FHWA, 2015). Jason's Law also mandated the United States Department of Transportation (US DOT) to perform nation-wide HCV parking surveys to evaluate the supply, demand, and quality of HCV parking around the country (FHWA, 2015).

Hours-of-Service (HOS) regulations and emerging requirements for electronic logging devices (ELDs) limit the extent of fatigued driving but also place a greater pressure on HCV parking (ATRI, 2018). The HOS regulations govern the duration of time that a driver can operate an HCV. The United States and Canada typically[1] limit daily driving to 11 and 13 hours, respectively (Government of Canada, 2009; FMCSA, 2016). While HOS regulations can result in safer roads, they may also create a greater challenge for HCV drivers who need to find convenient parking locations before exceeding their maximum driving hour limit. To ensure that HCV drivers are not falsifying or failing to complete their HOS logs, the United States mandated ELDs in 2017 and Canada is expected to mandate ELDs in 2021 (*Truck News*, 2017; Government of Canada, 2017; ATRI, 2018; Lamb, 2019). ELDs have already begun to impact HCV parking in the United States, especially for small businesses (Scott *et al.*, 2019). Parking shortages will likely be a growing concern due to stricter HOS enforcement from ELDs and continual HCV traffic growth along major highway routes in North America (NASEM, 2016).

The number of fatal HCV collisions has recently been on the rise (FMCSA, 2018; The Canadian Press, 2018) which may be in part caused by fatigue (NASEM, 2016) and insufficient HCV parking (Boggs *et al.*, 2019). The Atlanta Regional Commission has developed an HCV parking location classification schemes to help better understand the location and extent of existing HCV parking shortages as a matter of public safety (2018). Although this classification scheme is the best existing option, an improved quantitative HCV parking location classification scheme could be designed to be systematically implementable using commonly available geospatial data.

[1] HOS Laws in the United States and Canada have specific rules and exceptions relating to daily driving and working time (Government of Canada, 2009; Federal Motor Carrier Safety Administration, 2016).

In order to better understand these current-day concerns with HCV parking in North America, it is important to examine how they evolved over time. The movement of goods using HCVs first started in North America over 100 years ago. Since then, policy decisions and a lack of infrastructure development have increased the problems related to HCV parking over time along with the size and scope of the goods movement industry. By examining the evolution of this problem and the preliminary countermeasures that were put into place, it becomes easier to determine the causes and extent of the issues relating to HCV parking. A detailed history on this subject can be found in Appendix A: History of Goods Movement.

1.2 Research Goal

The aim of this research is the development of an approach to identify, classify, and estimate long-haul HCV parking supply to increase the accuracy of assessing HCV parking adequacy. GPS data is first processed to identify the locations where drivers are fulfilling their hours of service (HOS) requirements on long-haul trips. Then, an exhaustive HCV parking location classification scheme is developed based on extensive literature review. A systematic approach to apply this classification scheme using commonly available geospatial data is also developed. Lastly, a mathematical model is developed to estimate the HCV parking supply. The approaches developed as part of this research are applied to the case study area: the Region of Peel in Ontario, Canada.

1.3 Research Objectives

Literature on the topic of HCV parking has generally prioritized the modelling of demand over supply; however, comprehensive estimates of supplied HCV parking spaces are crucially needed to properly understand any existing HCV parking problems (Atlanta Regional Commission, 2018). HCV parking supply is often underestimated by jurisdictions due to the omission of valid parking spaces such as those belonging to private facilities. This oversight can result in inefficient and expensive infrastructure projects pursued based only on qualitative knowledge. If HCV parking deficiencies exist and are not properly understood or addressed, serious safety concerns can arise for both HCV drivers and the greater public at large. The four specific objectives of this research are:

1. Develop an approach to identify long-haul HCV rest stop parking locations;
2. Create a classification scheme for HCV parking locations based on an extensive literature review;
3. Devise a systematic approach to apply the above classification using geospatial data; and
4. Develop an HCV parking space count model to estimate total parking supply.

This research creates a set of tools for use by municipalities and transportation professionals to better determine the location and extent of any HCV parking deficiencies. This research will also address the existing gap in the literature related to HCV parking supply and classification. A diagram showing the relationship between the four research objectives is shown in Figure 1-1.

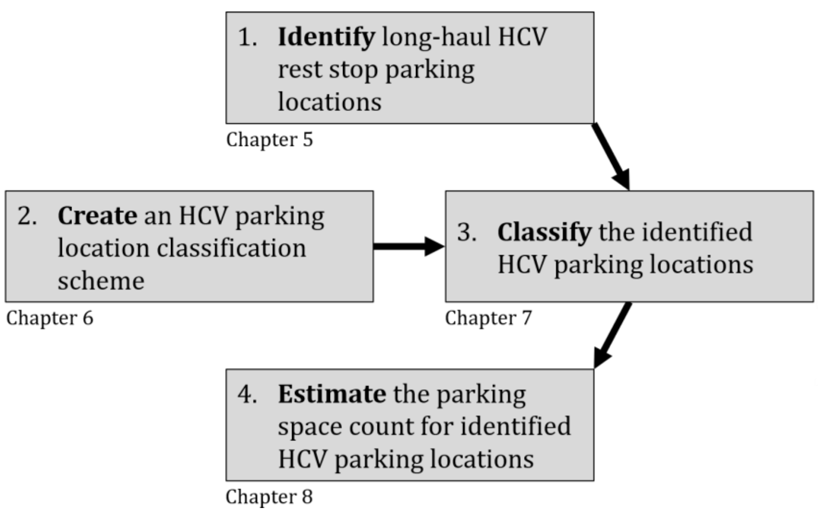


Figure 1-1: Relationship between Research Objectives and Thesis Chapters

1.4 Research Scope

The scope of this thesis is limited to HCV parking supply. HCV parking demand is not considered as part of this thesis, nor is the methodology for comparing HCV parking supply and demand. The total HCV parking supply is considered; therefore, spaces that are occupied are counted as part of the overall supply. A proper comparison between HCV parking supply and demand would require the derivation of parking utilization factors to account for HCV parking spaces that are blocked by improper parking, non-HCV vehicles, and or temporary storage. This comparison would also require an analysis of the temporal distribution of HCV parking demand; however, this is also outside the scope of this thesis.

The scope of this thesis is also limited to HCV parking related to rest stops used to comply with HOS laws on long-haul trips. This thesis does not examine short-haul trips and does not examine stops that were short in duration. Short stops would include those used to deliver or pick up goods where the duration of the stop is too short[2] to contribute to HOS-compliant off-duty time.

Lastly, this thesis does not examine the specific geometric requirements for parking HCVs. The scope includes the existence of HCV parking but not the adequacy of individual parking spaces. The turning radius and swept path of HCVs is not considered when examining existing HCV parking supply.

1.5 Thesis Structure

This thesis contains an additional eight chapters. In Chapter 2: Literature Review, the results of an extensive literature review are presented in order to provide context to the current issues surrounding HCV parking. In Chapter 3: Study Data, the data used in this thesis is discussed in detail and information is provided on the Region of Peel case study area located in Ontario, Canada. In

[2] Canadian HOS laws permit meeting the mandatory off-duty time in a sleeper birth in two periods as long as neither period is shorter than two hours (Government of Canada, 2009).

Chapter 4: Analysis Methods, the methods of analyzing data, building regression models, and checking model goodness-of-fit are described. In Chapter 5: Identification of HCV Parking Locations, GPS data is processed to identify the locations where HCVs are parking on long-haul trips to meet their HOS requirements. In Chapter 6: Development of an HCV Parking Classification Scheme, feasible combinations of key HCV parking location characteristics identified in the literature review are heuristically stratified to develop an exhaustive HCV parking location classification scheme. In Chapter 7: Application of the HCV Parking Classification Scheme, the proposed scheme is systematically applied using commonly available geospatial data to locations previously identified in the Region of Peel. In Chapter 8: Estimation of HCV Parking Supply, a negative binomial regression is estimated to model HCV parking supply using processed geospatial variables. In the final chapter, Chapter 9: Conclusion, the key findings of this thesis are summarized and recommendations for future technical and policy research are provided.

CHAPTER 2: LITERATURE REVIEW

This chapter presents the summary of a literature review performed for this thesis. Important topics discussed include the definition of HCV parking, driver safety, Hours of Service (HOS) laws, existing HCV parking analyses, key characteristics of HCV parking, and an existing HCV parking classification scheme.

2.1 Heavy Commercial Vehicle Parking

This thesis focuses on long-haul freight truck parking to meet the requirements of Hours of Service (HOS) laws. Due to the larger size and greater mass of HCVs, there are issues related to parking HCVs that are beyond that of parking passenger vehicles.

2.1.1 Definition

Freight truck parking can be defined by examining the definitions of both 'freight' and 'parking'. In this context, 'freight' is defined as "*goods in transit or being transported by rail, road, or sea*" (Oxford University Press, 2019a). This thesis focuses on those goods being transported by road in heavy commercial vehicles (HCVs), which are commonly referred to as trucks. 'Parking' is defined as "*space reserved or used for… [t]he placing or leaving of a vehicle or vehicles[, such as] a car park or other designated area, at the side of a road, etc.*" (Oxford University Press, 2019b). Taken together, 'freight truck parking' is defined in this thesis as space used for placing or leaving goods-moving HCVs. This thesis focuses on the policies and availability of the locations where these HCVs can or cannot legally park to rest during goods-moving long-haul trips.

2.1.2 Heavy Commercial Vehicles

A heavy commercial vehicle, or HCV, is a large goods-moving vehicle that is used for commercial purposes. These vehicles are distinct from passenger vehicles as their larger size and greater mass restrict the locations where they can park (Ogden, 1991).

One system for classifying vehicles is the FHWA 13-Category Vehicle Classification Scheme, which classifies vehicles based on axles, tires, and trailers (Hallenbeck *et al.*, 2014). A visual representation of this system is shown in Figure 2-1. Based on this vehicle classification system, HCVs can refer to vehicles Class 6 to 13; however, this thesis will focus on single trailer HCVs where the tractor and trailer can be separated. The project scope excludes single unit commercial vehicles (Class 6 and 7), often called delivery trucks, which do not always face the same parking limitations and are not often used for long-haul trips (Amer and Chow, 2017). The scope also excludes Long Combination Vehicles (LCVs) (Class 11 to 13) where the tractor is pulling multiple trailers; very different laws govern these vehicles and the parking limitations are much more severe (Ministry of Transportation Ontario, 2019; Parsons, 2019).

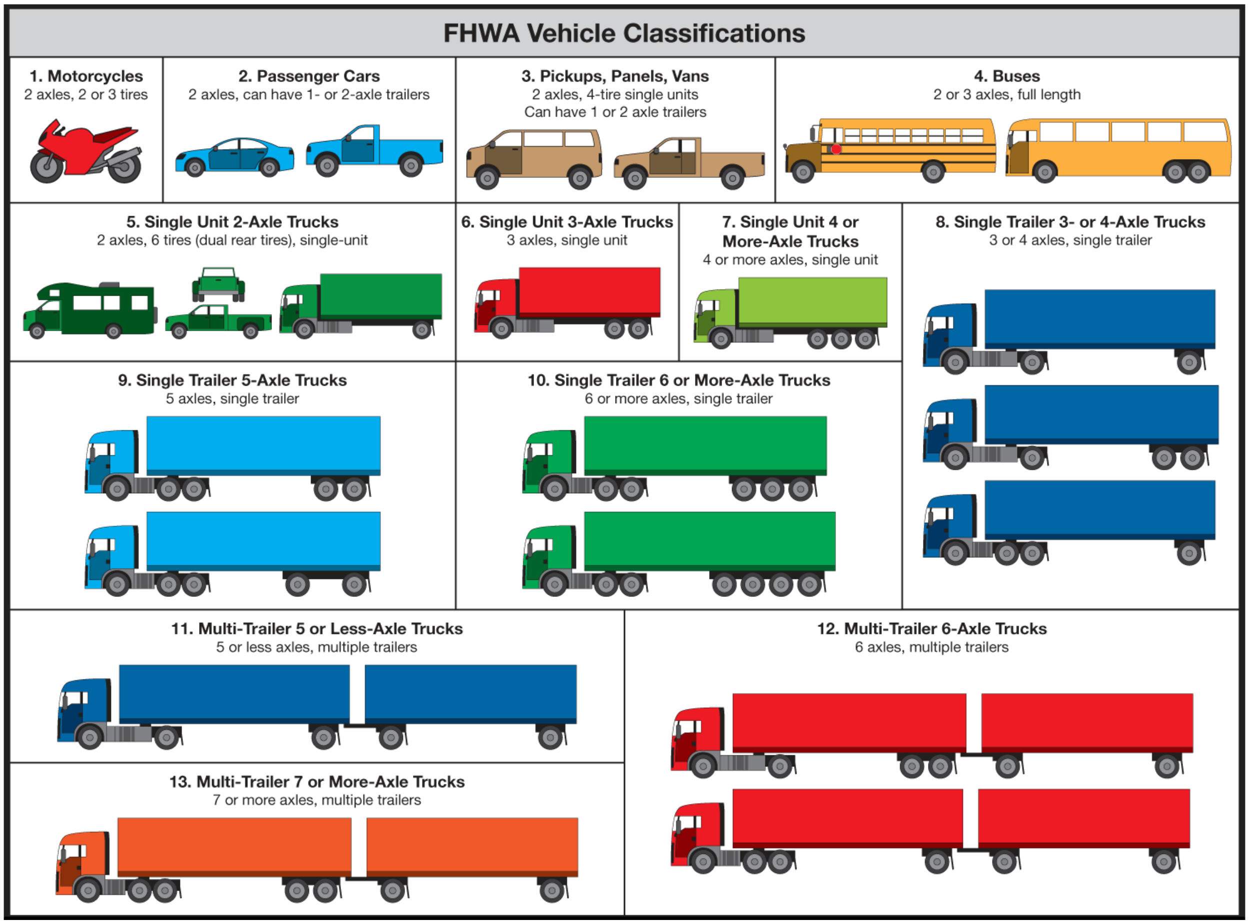


Figure 2-1: Vehicle Classification using FHWA 13-Category Scheme
(Texas DOT, 2012)

The most common heavy truck used in the North America is Class 9, which is part of the three classes this thesis focuses on (NASEM *et al.*, 2010; Hallenbeck *et al.*, 2014). Class 9 HCVs are tractor-trailer combinations with a total of five axles. Class 8 and 10 vehicles are the other single-trailer classifications and have similar turning radius limitations and overall dimensions. The dimensions of a typical Class 9 HCV, the WB-19 Design Vehicle, are found in Figure 2-2. The turning radius and swept path of this vehicle, dependent of the turn angle, is shown in Figure 2-3. The turning limitations of HCVs helps to define the locations in which they can maneuver and park (Harwood *et al.*, 2003). Large parcels of land are required to construct HCV parking lots, which is a key obstacle in the development of new HCV parking infrastructure (TRI *et al.*, 1996; Giuliano *et al.*, 2017).

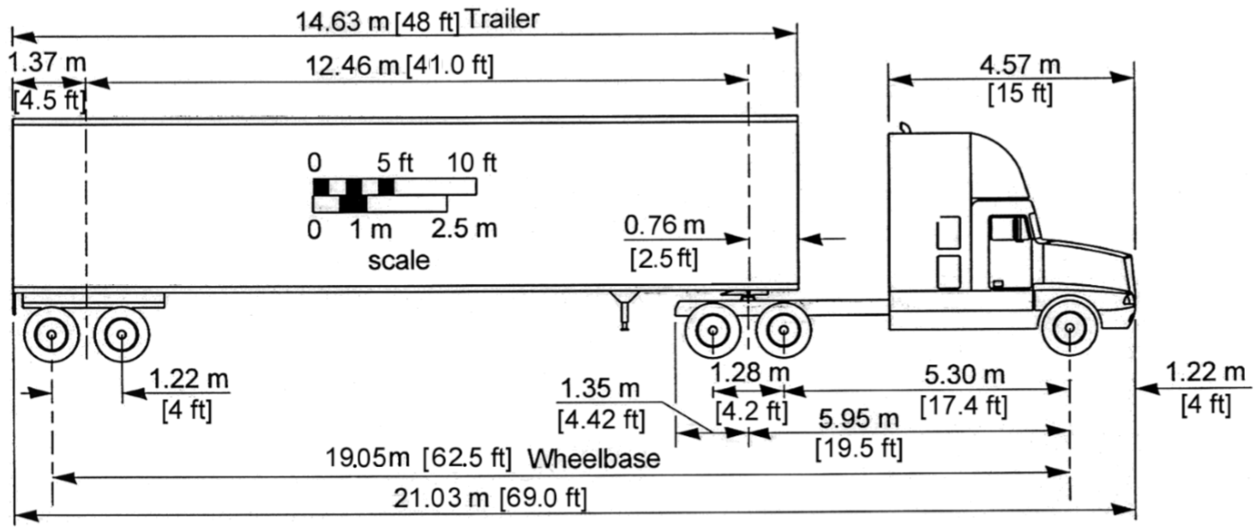


Figure 2-2: WB-19 Design Vehicle Specifications
(Harwood *et al.*, 2003)

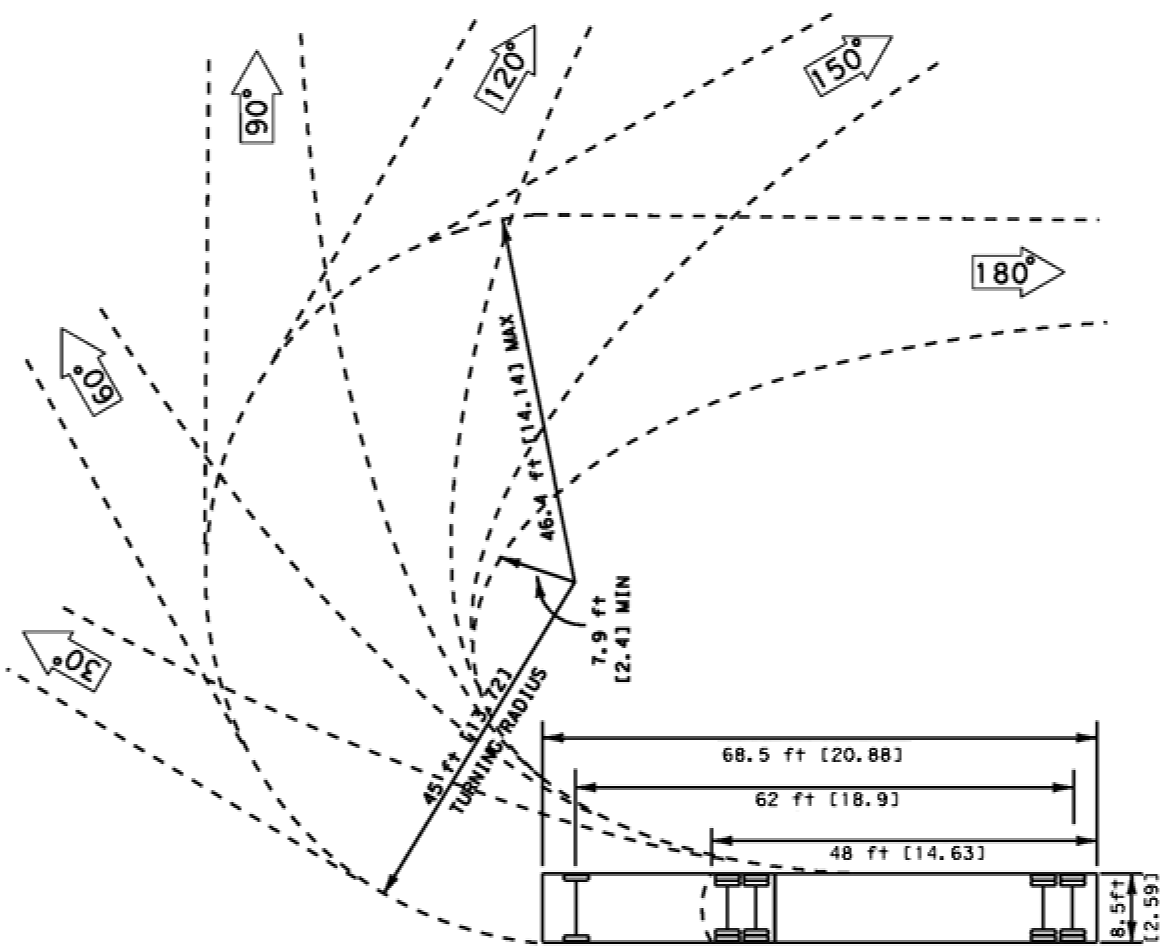


Figure 2-3: WB-19 Design Vehicle Turning Path
(Texas DOT, 2018)

2.1.3 Long-Haul Trips

There are two types of HCV trips that should be examined separately: short-haul trips and long-haul trips (Amer and Chow, 2017). Short-haul trips typically refer to urban deliveries that occur within a single city or region and typically arrive at their destination before a mandatory rest break is required (Fleger *et al.*, 2002; Gill and Macdonald, 2013). In contrast, long-haul trips typically occur between cities or regions and have a much greater likelihood of requiring a rest break to satisfy hours of service (HOS) laws (Fleger *et al.*, 2002; Gill and Macdonald, 2013). Average parking duration varies between short and long-haul trips, with drivers meeting rest requirements during long-haul trips in addition to eating, refuelling, performing maintenance, and using the restroom (Fleger *et al.*, 2002; Trépanier *et al.*, 2010; Kimley Horn, 2015).

Long-haul trips, or inter-regional trips, are often defined based on the trip distance with examples including a one-way distance of 400 km (Pécheux *et al.*, 2002) or 500 km (Ortuzar and Willumsen, 2011). Long-haul trips with a total distance travelled in excess of 900 km have been demonstrated to take substantially more time due to required rest breaks (Gingerich *et al.*, 2016a).

Short-term HCV parking, such as curbside spaces for urban deliveries (Amer and Chow, 2017; Mitman *et al.*, 2018; Schmid *et al.*, 2018), is predominantly concerned with timely access and close proximity to businesses for efficient goods transfer. This needs to be considered separately from long-term parking such as rest stops and overnight stops (Amer and Chow, 2017), where safety concerns and suitable amenities are paramount. Although this thesis focuses on long-term rest stops, readers who are interested in research on urban deliveries can find relevant studies conducted by Malik *et al.* (2017), Nourinejad and Roorda (2017), de Abreu e Silva and Alho (2017), Chow and Amer (2015), Yang *et al.* (2014), and Nourinejad *et al.* (2014).

2.1.4 Availability

A lack of HCV parking has been identified in locations all over North America. A 2015 FHWA national truck parking survey in the United States found this shortage of parking supply to be a national safety concern (2015). Studies from individual regions have found results that are consistent with the conclusion of this FHWA survey. A study of HCV parking on the I-5 interstate in California found that HCV parking was difficult to find and more than 70% of drivers have not been able to park at truck stops due to the facility being full (Martin and Shaheen, 2013). HCV drivers expressed an interest in a real-time parking availability application that would allow them to more easily find overnight parking. The Atlanta Regional Commission's Truck Parking Assessment Study (2018) found that there was a current lack of HCV parking that will worsen in the future due to significant ongoing growth. The study concluded that the necessary next steps are to add or expand HCV parking supply, develop HCV parking policies and partnerships, improve the sharing of HCV parking information, and monitor/integrate future technology developments into the freight industry. The study also called for a more detailed analysis of the current parking supply and the identification of existing brownfield sites that could be converted into legal and safe HCV parking.

In an analysis looking at trip generation in the province of Ontario, it was identified that the province as a whole, and especially the Greater Toronto Area (GTA) region, has experienced massive growth in the movement of freight. Sections of Highway 401 in the GTA experience an annual average daily truck traffic (AADTT) of 25,000 and Highway 403 in Mississauga experiences an AADTT of 30,000 (Ferguson *et al.*, 2014). Based on survey results, Ontario is considered to have the most problematic HCV parking in all of Canada, with the GTA area being considered the worst (Trépanier *et al.*, 2010). Figure 2-4 shows the Canadian provinces with the number of surveyed driver reports identifying problematic HCV parking locations in this study. Figure 2-5 identifies the top twenty worst regions for reports of problematic parking in Canada.

Cargo on any given truck in the GTA region is worth $500,000 on average and can be in the millions of dollars for high-end goods; therefore, any issues with freight transport can have large economic effects (Ferguson *et al.*, 2014). Problems associated with parking can delay shipments or, in a worst-case scenario, lead to an accident due to either fatigue or unsafe parking practices. Examining origin and destination information may help better understand the mechanics of freight movement throughout the province of Ontario (Ferguson *et al.*, 2014).

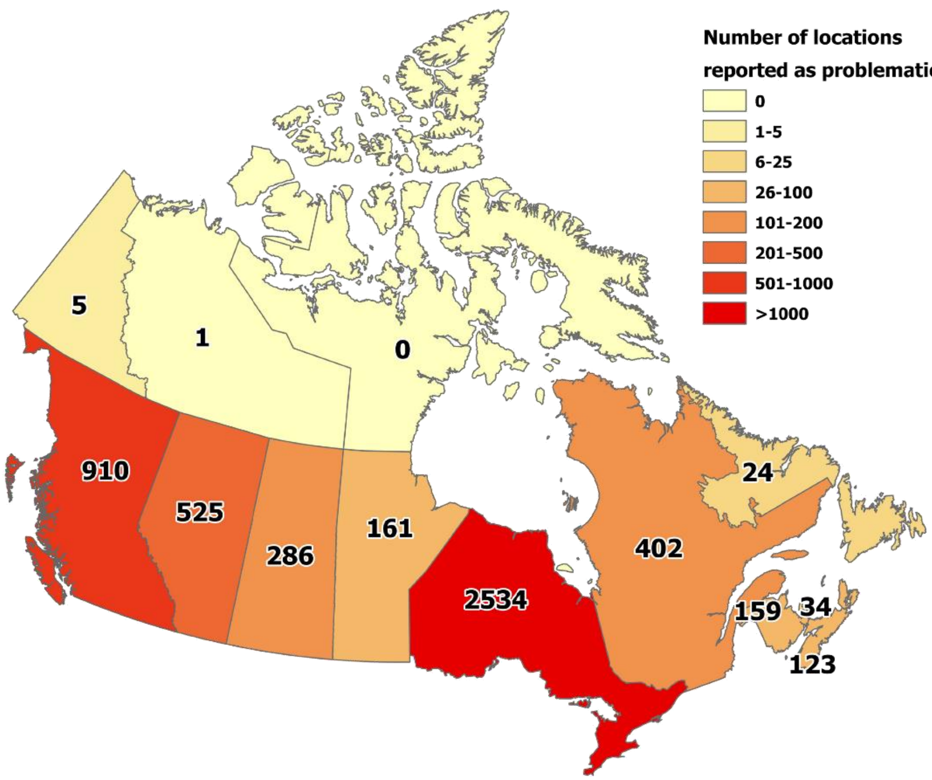


Figure 2-4: Number of Problematic HCV Parking Locations Reported in Canadian Provinces Adapted from Trépanier *et al.* (2010)

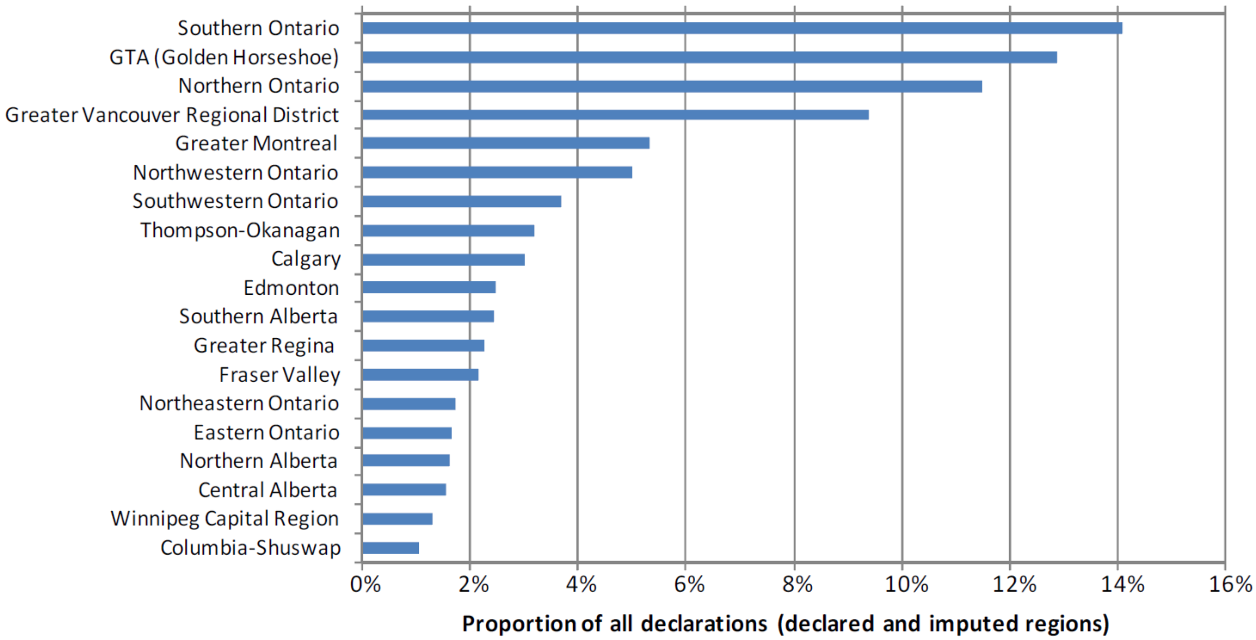


Figure 2-5: Top Twenty Canadian Regions with Problematic Parking
Adapted from Trépanier *et al.* (2010)

2.1.5 Passenger Vehicle Parking

Although HCV and passenger vehicle parking problems share some similarities, the two types of parking require a different management approach. The number of passenger vehicle parking spaces is often determined by legal requirements associated with different types of developed property (Shoup, 1995), but the demand for parking has risen to the point where it is no longer affordable to maintain established legal parking requirements especially in urbanized areas. Many municipalities have enacted transportation demand management (TDM) policies to encourage transit, active transportation, and/or carpooling (Shoup, 1995). TDM policies are enacted for many reasons, but one major benefit is a reduction in the need for passenger vehicle parking (Ison and Rye, 2008). Parking fees are another common tactic used to discourage passenger vehicle parking (Lehner and Peer, 2019; Yan *et al.*, 2019). TDM can be effective because it reduces the demand for passenger vehicle travel, but the demand for HCV parking may not be as easily manipulated through TDM policies due to the commercial nature of goods movement.

Passenger vehicle parking spaces in parking lots can often be counted using GIS (geographic information system) software and areal imagery (McKeown *et al.*, 1985; Wang and Hanson, 1998; Seo and Urmson, 2009; Kabak and Turgut, 2010). This approach often cannot be utilized for HCV vehicles since their parking spaces are not necessarily of uniform size, often do not have the same distinctly painted parking spot markers (parking lines), often have different surface coverage or conditions, and are often more *ad hoc* in nature, sharing their space with other equipment and containers. Demand for passenger vehicle parking can be calculated using the ITE's Parking Generation Manual (Hooper, 2019). This ITE manual outlines the various characteristics that are correlated with passenger vehicle parking demand.

2.2 Collisions and Safety

A key rationale for the analysis of HCV parking is the increased safety risks associated with HCVs. Compared to passenger vehicles, collisions involving HCVs introduce increased risk to the health and welfare of the general public due to the large mass of HCVs, especially when loaded. Collisions involving HCVs also have significant economic considerations due to the potential for costly infrastructure damage and the disruption to goods supply chains.

2.2.1 Fatigue

One contributing human factor to vehicle collisions is the impairing effects of fatigue (AASHTO, 2008). It is estimated that between 10 and 20% of the roughly 4,000 annual fatalities resulting from collisions involving trucks and buses in the United States may have involved fatigue (NASEM, 2016). This is likely due to the stress and lack of sleep associated with the lifestyle of commercial vehicle drivers. Fatigued driving is a substantial concern for governments and has resulted in the creation of Hours of Service (HOS) laws.

Fatigue itself is difficult to define and objectively measure. Police accident report-based databases, such as the National Highway Traffic Safety Administration's (NHTSA) General Estimates System (GES) and Fatality Analysis Reporting System (FARS), identify fatigue as a contributing factor in 1 to 2% of fatal collisions involving HCVs and buses. This statistic is likely a massive underestimate due to the sudden rush of adrenaline that occurs after an accident, which can make detecting previous fatigue impossible in most circumstances (NASEM, 2016). Another study from the National Transportation Safety Board (NTSB) found fatigue to be the most common cause of fatal-to-driver HCV collisions, contributing to 57 of the 182 collisions (31%) of those examined as part of the study (1990). Fatigue was found to be more common than medical conditions and impairment by alcohol or other drugs in contributing to these fatal-to-driver HCV collisions. Lastly, the study found an association between violation of hours of service (HOS) laws and drug usage, suggesting that drivers may attempt to use drugs to combat the symptoms of fatigue.

Driver fatigue has many causes, including extended wakefulness or acute sleep deprivation, poor sleep quality, lengthy driving or working times, chronic insufficient sleep, suboptimal work schedules, sleep disorders, and certain medications (NASEM, 2016). Drivers face increasing pressures to work despite being fatigued due to a decreasing workforce, increasing demand for next-day delivery, and financial pressures. Although research has shown a relationship between sleep deficiencies and decreased driver attention, it is still not fully understood how much sleep is required to eliminate fatigue concerns. This question is complicated due to the many factors that affect fatigue other than length of sleep. Overall, sleep deficits have been found to contribute to decreased alertness and decreased performance. This decease in performance can lead to driver error and inappropriate driving practices, which in turn can result in collisions (NASEM, 2016).

2.2.2 HCV Collisions and Fatalities

Data analysis has shown that the total number of fatalities involved in HCV collisions have been growing in recent years. As shown in Figure 2-6, the overall number of fatalities in collisions involving HCVs has been increasing between 2009 and 2016 in the United States (FMCSA, 2018). This trend is also seen in Ontario, Canada, where the number of fatal collisions involving HCVs reached a ten-year high in 2017 (The Canadian Press, 2018).

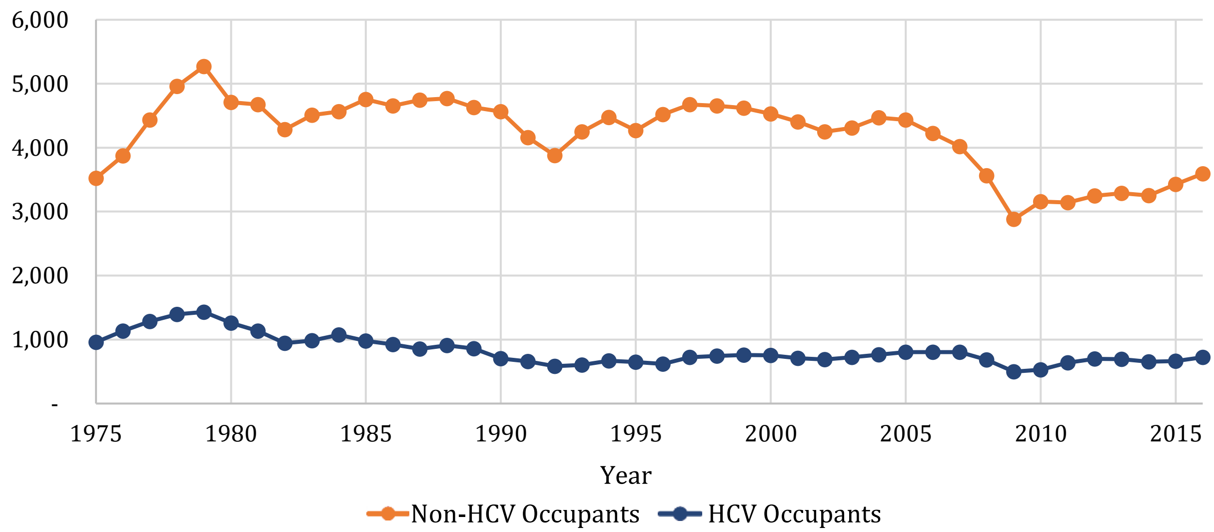


Figure 2-6: Total Fatalities in Collisions involving HCVs
Adapted from FMCSA (2018)

When the number of fatalities involving HCVs is examined per unit distance travelled, it can be shown that the overall rate has been decreased during the last four decades in the United States, as shown in Figure 2-7 (FMCSA, 2018). The massive increase in goods movement due to e-commerce has likely contributed towards the growing number of HCVs in the transportation network (Pettersson *et al.*, 2018), resulting in the overall rise in collisions despite decreasing rates per unit distance.

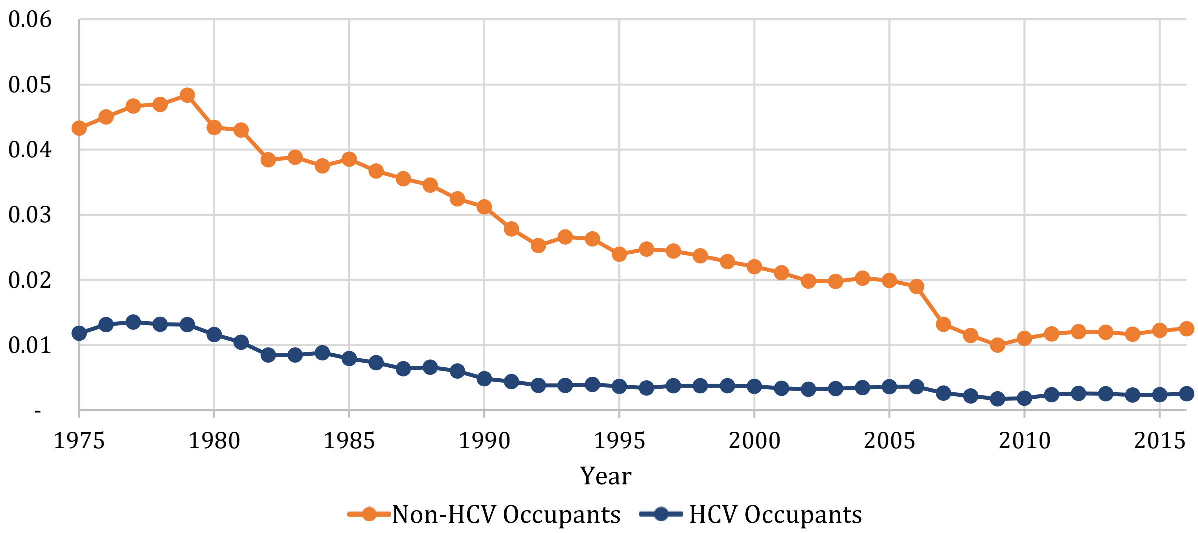
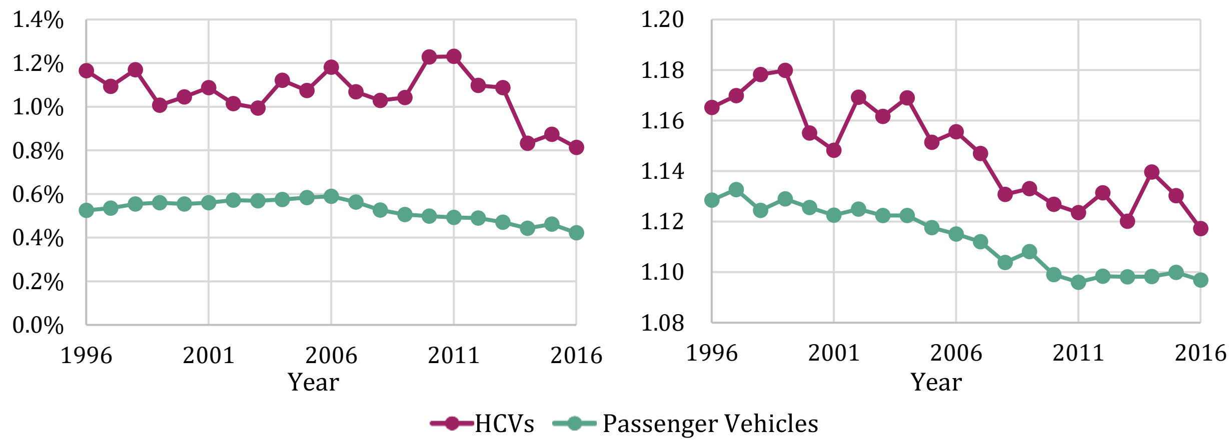


Figure 2-7: Fatalities in Collisions involving HCVs (per 100 Million Miles Driven)
Adapted from FMCSA (2018)

Figure 2-6 and Figure 2-7 show that in the United States, the majority of fatalities in HCV collisions are the occupants of non-HCV vehicles. Additionally, collisions involving HCVs are more likely to result in a fatality, as shown in Figure 2-8(a), and are more likely to have a higher number of fatalities compared to collisions that do not involve with HCVs, as shown in Figure 2-8(b) (FMCSA, 2018). These statistics show that although rates of HCV collisions are going down, collisions with these vehicles is still more likely to be fatal and will often result in a higher number of fatalities, with a disproportionate number of these fatalities being non-HCV occupants. Between 1975 and 2016, after normalizing for distance travelled, there have been 1.57 times the number of fatalities in collisions involving HCVs compared to those only involving passenger vehicles in the United States. HCV driver fatigue and parking issues likely contributed to some portion of this safety concern. These trends show that HCV safety is a concern for both HCV drivers and the general public.



(a) Percentage of Fatal Collisions **(b)** Fatalities per Fatal Collision

Figure 2-8: HCV and Non-HCV Collision Statistics
Adapted from FMCSA (2018)

2.2.3 Jason's Law

Jason's Law, passed in 2012 in the United States, aims to address the lack of safe and legal parking spaces for HCVs. Jason's Law is inspired by Jason Rivenburg, an American HCV driver who was murdered in his cab while he was parked at an abandoned gas station. Jason had not been able to find a safe and legal location to park, so the law named in his honour aims to ensure that other drivers do not face a similar situation. Jason's Law specifically aims to increase the amount of safe parking by constructing safe public HCV parking facilities and adapting existing facilities next to commercial truck stops and travel plazas (FHWA, 2015).

In order to properly assess the scope of the parking problem, Jason's Law called for a national Truck Parking Survey. The Jason's Law Truck Parking Survey found that HCV parking deficiencies were, in fact, an issue in the United States, especially around urban areas and along major highways. In the survey, over 90% of HCV drivers reported having issues finding safe and legal HCV parking during night hours. The report strongly encouraged state and regional governments to examine the issues related to HCV parking more thoroughly and outlined several areas that would benefit from government financial intervention (FHWA, 2015).

2.3 Hours of Service

Laws relating to Hours of Service (HOS) have a strong effect on HCV driver parking behavior. These laws have become stricter in recent years after the introduction of electronic logging devices (ELDs). In order to properly examine HCV parking concerns, it is important to first understand the context and purpose of these laws. The American Transportation Research Institute (ATRI) has recently rated HOS and the ELD mandate as the second and fourth highest concern for commercial drivers in the United States, respectively (2019).

2.3.1 Definition and Purpose

Hours of Service (HOS) refers to the total time that drivers of commercial vehicles either spend driving (daily driving time) or spend working (on-duty time). In Canada, limitations are put on both daily driving time and on-duty time through commercial vehicle driver hours of service regulations (CCMTA, 2010).

Canada and the United States each have their own strict HOS laws that govern the maximum duration of time that a driver can operate a commercial vehicle, as well as minimum off-duty times. In Canada, the HOS laws are contained within the National Safety Code for Motor Carriers Standard 9 (CCMTA, 2010). In the United States, the HOS laws are contained within U.S. Code: Title 49 - Transportation (FMCSA, 2016). In Canada and the United States, drivers are limited to driving 13 and 11 hours in one day or work shift, respectively. Additional rules regulate the length of break time, the duration of on-duty time, and the total duration of driving time in multiple day periods. A simplified summary of the HOS laws in Canada and the United States are shown in Table 2-1; however, it should be noted that additional laws can change the details of the requirements depending on the circumstance of the driver (CCMTA, 2010; FMCSA, 2016).

Table 2-1: Simplified Summary of Hours of Service Laws

Parameter	Canadian Freight Laws	American Freight Laws
Maximum Daily Driving Time	13 hours	11 hours
Daily Rest Break Requirements	10 hours (8 hours must be consecutive)	10 consecutive hours
Maximum Driving Time in a Cycle	70/120 hours	60/70 hours
Length of a Cycle	7/14 days	7/8 days
Off-Duty Time Required to Reset a Cycle	36/72 consecutive hours	34 consecutive hours

Note: These are general rules; there are particular circumstances described in the HOS laws that can affect the durations expressed in this table. (CCMTA, 2010; FMCSA, 2016)

The purpose of HOS laws are to ensure that driver faculties are not impaired due to tiredness which could jeopardize the safety of both the driver and the public (CCMTA, 2010). Despite this intended purpose, these strict requirements, in combination with a limited supply of parking spaces, often lead to drivers needing to choose between illegally parking or breaking HOS laws (Chatterjee and Wegmann, 2000). Drivers who violate HOS laws face fines, points against the driver and carrier safety record, and out-of-service declarations (MTO, 2009; Government of Canada, 2017; Scott *et al.*, 2019).

2.3.2 Stakeholders

HOS laws affect a large number of stakeholders. Arguably those that are the most affected by HOS laws and the issues relating to HCV parking availability are the HCV drivers themselves. Drivers face immense pressure to ensure that they deliver their goods on time; therefore, they often push the limits of HOS (FHWA, 2015). Lawmakers are passing laws that are viewed as increasingly inflexible without ensuring that adequate parking is available (ATRI, 2018). Trucking companies and owneroperators are now expected to deliver goods faster and in shorter time windows without additional cost to the consumer (Gill and Macdonald, 2013; Buldeo Rai *et al.*, 2019). The goods movement industry pays a significant amount of taxes and feels that more government investment in HCV parking infrastructure is warranted (Phelan *et al.*, 2016; Region of Peel, 2017). Additionally, labour unions involved in the goods movement industry advocate for shorter working shifts and more flexibility for factors outside of the driver's own control such as weather and parking shortages (FMCSA, 2000).

The government, at various levels, is also a large stakeholder whose mandate is affected by HOS laws and HCV parking concerns. National, state/provincial, and local governments all have the responsibility of balancing the needs of industry and the public at large. The construction of parking is often seen as an unnecessary and wasteful use of public resources; therefore, many elected officials are hesitant to endorse these activities (TRI, Apogee Research Inc. and Wilbur Smith Associates, 1996; Shoup, 1999). Planners who work for governments impose and face strict density targets and zoning laws which makes the construction of HCV parking facilities quite difficult (Shoup, 1999; Government of Ontario, 2019).

The general public is also a significant stakeholder in this issue. The public, as taxpayers, are typically opposed to having public funds used to finance projects that they perceive to only benefit the goods-movement industry (TRI, Apogee Research Inc. and Wilbur Smith Associates, 1996; Shoup, 1999). However, the general public is also the driving force behind HCV parking inadequacies as increasing demands for eCommerce goods and next day delivery are main contributors to the rising number of HCVs on the road (Pettersson *et al.*, 2018; Buldeo Rai *et al.*, 2019). Additionally, the general public is also the most affected by collisions involving HCVs, as over 80% of the victims of these incidents are not HCV drivers (FMCSA, 2018).

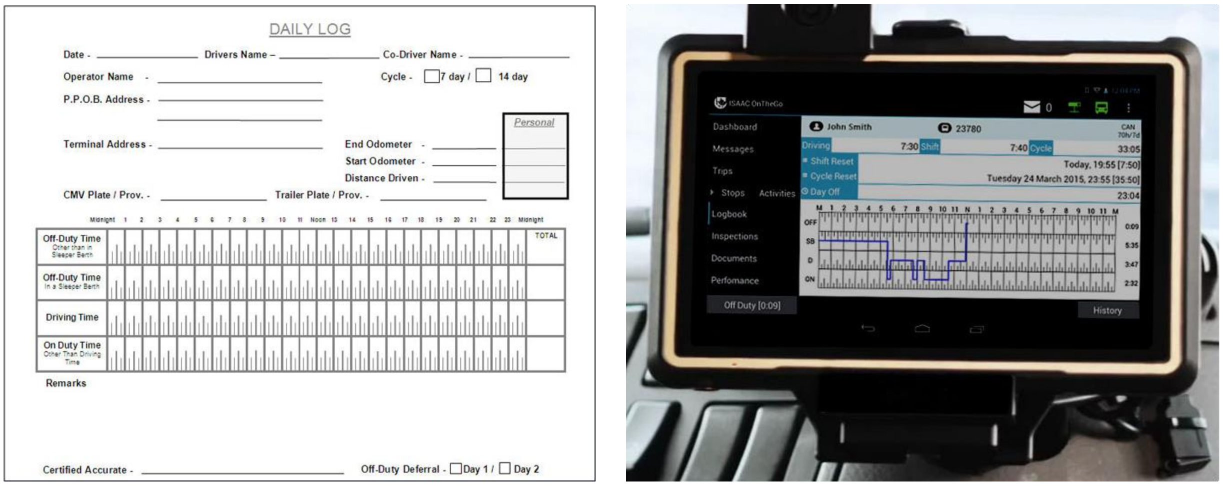
Other stakeholders affected by this complex issue include business owners, land use activists and environmentalists, the industries supported by goods movement, and the economy as a whole. Many business owners have conflicting needs and feel as though extensive government investment in one industry would be unfair to other industries. Land use activists and environmentalists have genuine concerns with the impacts that HCV parking facilities have on the land (Chatterjee and Wegmann, 2000; Davis *et al.*, 2010). Many industries are strongly dependent on goods movement and have an invested interest in ensuring that long-haul HCV trips can occur safely and efficiently. Lastly, the goods movement industry is a significant contributor to local and national economies, so it is essential that goods can be moved in an effective manner (Region of Peel, 2017; Government of Ontario, 2019).

2.3.3 Electronic Logging Devices

Recently, the American and Canadian governments began to examine the feasibility of using modern technology to automatically track hours of service instead of relying on drivers to complete manual HOS reports. Electronic logging devices (ELDs) are tools installed in an HCV that automatically record a driver's HOS (Government of Canada, 2017). ELDs can be installed in vehicles to ensure that hours of service are not forgotten, falsified, or erroneously completed. Drivers are less likely to violate HOS laws when they are electronically monitored, especially since ELDs are more difficult to falsify compared to traditional paper logbooks (Scott *et al.*, 2019). Examples of both the paper logs and an ELD interface are shown in Figure 2-9. The goal of ELDs is to increase HOS compliance and ultimately reduce HCV driver fatigue and collisions. This is achievable through the threat of automated violation detection and the penalties associated with noncompliance (Scott *et al.*, 2019).

The United States mandated ELDs in 2017 and Canada has announced it will follow suit in 2021 (*Truck News*, 2017; Government of Canada, 2017; ATRI, 2018; Lamb, 2019). This active approach to monitoring driving behaviour is often cited as a contributing factor to the increasing rates of illegal parking and the overall decreasing availability of parking supply. Many stakeholders in the trucking industry argue that this inflexible approach to monitoring HOS is unjust especially when considering the current truck parking shortages that are acknowledged at the highest government levels (*Truck News*, 2017). Some drivers and companies oppose the ELD mandate as there is no allowance for leniency as with paper logs. Those who oppose ELDs feel that there should be a 'common-sense grey area' that allows flexibility when drivers cannot find a safe parking space, or that ELDs should only be mandated for those with documented compliance issues (Dills, 2019).

Initial results from the United States ELD mandate have shown that driver compliance with HOS has increased (Scott *et al.*, 2019). One study examined the data from driver inspections and all federally recordable collisions during the first twenty months of the mandate. The increase in compliance was most commonly seen for drivers employed by small carriers; however, accident rates for these carriers also increased. This could be due, in part, to the increased frequency of speeding for these small carrier drivers as a response to the loss of productivity caused by the ELD mandate. Drivers for large carriers were less effected by the ELD mandate, as they were already more likely to be compliant with HOS laws (Scott *et al.*, 2019) or were already using this technology.



(a) Paper Log
(MTO, 2009)

(b) Electronic Logging Device
(Wikimedia Commons, 2015)

Figure 2-9: Methods of Logging HOS

2.4 HCV Parking Modelling and Analysis

Previous academic studies have examined HCV parking from both a demand and supply perspective. There has been significantly more research conducted on HCV parking demand, with mathematical models available to estimate demand in North America. HCV parking supply has garnered significantly less attention, with most estimates of supply developed using driver surveys.

2.4.1 Parking Demand

In 2002, the Federal Highway Administration (FHWA) commissioned the mathematical model (Pécheux *et al.*, 2002) that is frequently used to quantitatively estimate HCV parking demand on highways (Atlanta Regional Commission, 2018). This HCV parking demand model was developed for their 2002 national assessment of truck parking (Atlanta Regional Commission, 2018). The main formulas that make up this model are:

$$PHP_i = PPF_i \times THP_i \qquad (1)$$

$$THP_{LH} = (WPT/WDT) \times THT_{LH} + (D_{ST} \times THT_{LH})/60 \quad (2)$$

$$THP_{SH} = (D_{ST} \times THT_{SH})/60 \quad (3)$$

$$THT_i = P_i \times V_t \times TT \quad (4)$$

$$V_t = AADT \times P_t \times F_s \quad (5)$$

$$TT = L/S \quad (6)$$

where the subscripts SH and LH represent short-haul and long-haul, respectively, and i can represent either SH or LH; PHP is peak-hour parking demand (trucks/hour or spaces/hour); PPF is a peak-parking factor; THP is the daily truck-hours of parking demand (hours/day); WPT is weekly parking time; WDT is weekly driving time; THT is daily truck-hours of travel (hours/day); D_{ST} is the duration of short-term stops per hour traveled (min/hour); P_i is the proportion of total trucks that are either SH or LH; V_t is the seasonal peak daily truck volume (trucks/day); TT is the average truck travel time (hours/truck); $AADT$ is the annual average daily traffic (vehicles/day); P_t is the percent of total traffic that is trucks; F_s is a seasonal peaking factor (suggested 1.15); L is the analysis segment length (km); and S is the speed limit or average truck speed (km/h) (Pécheux *et al.*, 2002).

The results of this model are highly variable at the segment level but are more accurate when aggregated at the regional level (Pécheux *et al.*, 2002). It should also be noted that this model is used for general regions despite being derived from survey responses of drivers in the United States only; however, model calibration can increase the model accuracy in specific regions (Pécheux *et al.*, 2002).

Other studies on HCV parking demand include Morris *et al.* (2017), Boris and Brewster (2016), Haque *et al.* (2017), and Chatterjee and Wegmann (2000). There are two different types of HCV parking demand including longer rest stops (such as an overnight stay at a rest area) and short-term stops (such as urban deliveries). More information on the distinction between short and long-haul trips is provided in Section 2.1.3.

2.4.2 Parking Supply

Driver surveys are a typical source of information used to examine HCV parking supply and gain insight into the adequacy of existing parking locations and spaces. The surveys typically indicate that existing parking spaces are not sufficient. For instance, 92% of HCV drivers that travel through the Atlanta Region reported in one study that they require 30 minutes or longer to find a suitable parking space (Atlanta Regional Commission, 2018). HCV drivers also reportedly prefer rest areas with amenities such as access to fuel, food, internet, and restrooms (Trépanier *et al.*, 2010; Atlanta Regional Commission, 2018). Many studies have also suggested that HCV drivers could benefit from an advanced application that identifies and disseminates where HCV parking is available on a real-time basis (Trépanier *et al.*, 2010; Martin and Shaheen, 2013; Boris and Johnson, 2015; Atlanta Regional Commission, 2018). The SmartPark in-cab parking notification system in Minnesota is an example of such an application (Morris *et al.*, 2017). For more information about the results of these HCV parking supply studies, see Section 2.1.4.

HCV drivers have reported their long-haul parking preferences to include public rest areas, highway weigh stations, and commercial parking areas. The reported unfavourable locations include freeway shoulders, interchange ramps, and local streets since these locations lack security and amenities (FHWA, 2015). Other characteristics related to parking location preferences include an absence of HCV parking fees and the provision of exclusive parking spaces for HCVs (Pécheux *et al.*, 2002; Malik *et al.*, 2017).

The suggestion of a more thorough HCV parking supply count has existed in literature for quite some time. The Transportation Equity Act for the 21st Century (TEA-21), enacted in the United States in 1998, requires in Section 4027 that the FHWA obtain a count of all public and private sector HCV parking spaces and determine where any shortages may exist (USA: 105th Congress, 1998). The subsequent study released in 2002 analysed public parking facilities, rest areas, and truck stops and calculated an approximate ratio of the parking demand to a supply obtained from established surveys (Fleger *et al.*, 2002). While this was a useful step in examining the balance between parking demand and supply, it neglects a large percentage of viable, safe parking locations that exist in other areas. Currently, no mathematical models have been developed to estimate HCV parking supply.

2.5 Important Characteristics of HCV Parking

In order to properly develop a classification scheme for HCV parking locations, it is important to understand the key characteristics that can be used to describe these locations. Based on the extensive literature review conducted for this thesis, the most important aspects of HCV parking have been identified as: 1) legality, 2) accessibility, 3) ownership, 4) dedication to HCV parking, and 5) roadside parking. Each of these characteristics can be described as a binary variable, meaning that they can be represented as two possible states indicating true (1) or false (0) (Everitt and Skrondal, 2010).

2.5.1 Legality

Legal status is a very important characteristic of an HCV parking location (FHWA, 2015). The legality of a location in this context refers to authorized (legal) or unauthorized (illegal) HCV parking. Legal HCV parking locations do not have laws, bylaws, or policies prohibiting HCV parking; however, a legal parking location does not necessarily have to be accessible for all HCV drivers.

An unauthorized HCV parking location does not permit any HCV parking due to a law, bylaw, or policy which may or may not be related to the land use or the physical limitations of the location. For this thesis, an unauthorized parking location is more broadly defined as any location other than legal HCV parking areas. Unauthorized parking occurs periodically at vacant or abandoned lots, highway ramps, and other roadway segments where HCV parking is prohibited (Atlanta Regional Commission, 2018; Ioannou and de Almeida Araujo Vital, 2018). These locations are a concern due to increased criminal activities and/or traffic collisions associated with HCV parking (FHWA, 2015). This thesis classified all locations that did not legally allow HCV parking as unauthorized.

Chatterjee and Wegmann (2000) conducted an HCV parking study on Tennessee interstate highway corridors and reported that 468 (80%) of the HCVs were parked in legal designated spaces while 117 (20%) of the HCVs were parked illegally on the entrance and exit ramps. Boris and Brewster (2016) reported that 84% of American HCV drivers park in unauthorized locations at least once a week and almost 10% of drivers have admitted to daily violations.

While HCV drivers may prefer legal parking facilities that provide suitable amenities and legal piece of mind, unauthorized HCV parking may indicate that a driver could not find a legal parking space due to a lack of supply or insufficient knowledge of any available parking spaces nearby (Boris and Johnson, 2015). In addition, adverse weather and/or fatigue increase the likelihood of a driver stopping on a highway ramp (Anderson and Roll, 2018). Drivers reportedly prefer to stop as soon as possible once they become fatigued and view nearby ramps as the most convenient option for parking (Chatterjee and Wegmann, 2000). Despite the safety risks of parking on ramps, some enforcement agencies hesitate to ticket HCV drivers parking on ramps as stopping to rest can be regarded as less hazardous than fatigued driving (Kimley Horn, 2015).

2.5.2 Accessibility

The accessibility of an HCV parking location can be described as open or limited. Open access HCV parking locations are available to all HCV drivers while limited access HCV parking locations are accessible only to a specified group of HCV drivers. Limited access parking locations are often locations where goods are loaded and unloaded and typically available to a small number of carrier companies (Gingerich *et al.*, 2016b).

Many HCV parking locations are limited access. Private commercial businesses and industrial properties sometimes allow HCV parking to drivers who deliver or pick up goods from their facilities (Trépanier *et al.*, 2010; Corro *et al.*, 2019), but many locations restrict long-term parking for HCVs (Atlanta Regional Commission, 2018). This additional parking supply, although available only to certain drivers, reduces demand for HCV parking at open access HCV parking locations by increasing the supply located elsewhere.

2.5.3 Ownership

HCV parking sites may be publicly or privately held. Examples of publicly owned HCV parking facilities include welcome centers, weigh stations, HCV inspection locations, and some rest stops (FHWA, 2015). At some publicly owned locations, HCV parking availability is reduced due to spaces being used by non-commercial vehicles. For example, TRI *et al.* (1996) reported that non-commercial vehicles occupy 10 percent of HCV parking spaces in some designated rest areas.

Parking demand often exceeds parking supply at many public HCV parking locations, leading to vehicles parking in adjacent and potentially illegal locations. For example, the Virginia DOT found that over 80% of public HCV parking facilities in Virginia regularly experienced overflow parking demand on nearby highway ramps (Kimley Horn, 2015). HCV drivers interviewed by Boris and Brewster (2016) reasoned that the closure of existing public rest area facilities and the resulting increased distance to available parking facilities was a major concern.

In some cases, these problems could result from insufficient investments in HCV parking infrastructure capacity. Investment appears to vary widely by jurisdiction. The US DOT found that the ratio of private to public HCV parking spaces in the United States varied from 1.6 in Vermont to 39.9 in Oklahoma (FHWA, 2015).

2.5.4 Dedication to HCV Parking

The Atlanta Regional Commission's (2018) classification scheme included privately-owned parking locations. The Commission identified two groups: primary (dedicated) locations with space designed exclusively for HCV parking and secondary (non-dedicated) locations with spaces available for HCVs to park despite HCV parking not being the intended primary use of the space (Atlanta Regional Commission, 2018). Examples of the non-dedicated locations include shopping centres, hotels, and restaurants.

The locations offering non-dedicated HCV parking will naturally contain fewer amenities or considerations specifically for drivers, but they nonetheless provide additional capacity to a regional parking supply. An issue with previous classifications was the omission of these locations, thereby resulting in an underestimation of HCV parking.

2.5.5 Roadside Parking

Roadside HCV parking refers to parking occurring on the side of a road, such as a highway, local road, or ramp. Roadside parking often occurs along roadways with low traffic volumes such as dead-end roads, road shoulders, or in industrial areas. Such locations are not usually designed for long term HCV parking and HCV oil spills pose an environmental concern (Chatterjee and Wegmann, 2000; Boris and Brewster, 2018).

It is, however, important to acknowledge this HCV parking supply as many drivers rely on roadside parking to comply with hours of service laws, and roadside parking is especially important in areas such as northern Canada where dedicated parking may be infeasible. HCV drivers reported that northern Canada has substantially less parking facilities than other areas of the country (Phaneuf, 2008). Roadside parking is therefore the only parking option along many northern highways in

Canada, but the dangers are increased in the winter months due to snow and ice (Standing Committee on Transport and Government Operations, 2001).

Depending on the jurisdiction, location, and road type, roadside HCV parking can be authorized or unauthorized. Roadside HCV parking is often ignored when assessing supplied parking spaces despite contributing towards parking availability along certain roadways (Atlanta Regional Commission, 2018), but this should not be used as a justification for neglecting the creation of new dedicated parking spaces (Atlanta Regional Commission, 2018; Boris and Brewster, 2018).

2.6 Existing HCV Parking Classification Scheme

The most comprehensive HCV parking classification scheme available was developed by the Atlanta Regional Commission as part of their HCV parking assessment study (2018). They proposed the following categories: 1) publicly-controlled; 2a) privately-owned primary; 2b) privately-owned secondary; and 3) unauthorized. In this context, primary refers to land use purpose of built locations like truck stops, and secondary refers to locations like motels or shopping plazas where limited parking may occur, but the primary use is not HCV parking. This classification scheme is shown in Figure 2-10.

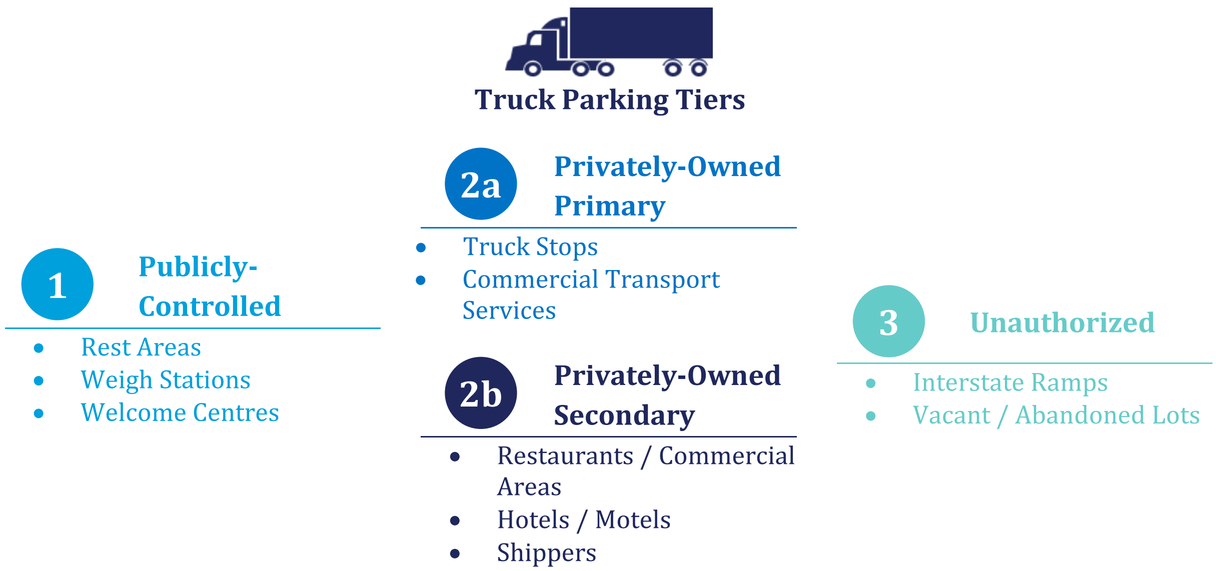


Figure 2-10: Atlanta Regional Commission Truck Parking Tiers
Adapted from Atlanta Regional Commission (2018)

Although the Atlanta Regional Commission's classification scheme considers legality, ownership, and land use purpose, the literature review of key HCV parking characteristics indicates that further improvements could create a more exhaustive classification. Another limitation of this classification scheme is that there is no quantitative method described to classify parking locations into one of these categories. The classification scheme that is developed as part of this thesis should be designed to be systematically implemented using geospatial data. The data used for this purpose is described in Chapter 3: Study Data.

CHAPTER 3: STUDY DATA

Geospatial datasets are required in order to identify where HCVs are parking, develop and apply an HCV parking classification scheme, and estimate HCV parking supply. This chapter discusses the case study area for this thesis as well as the various data sources used.

3.1 Case Study Area: Region of Peel

The case study area for this thesis has been selected as the Region of Peel in Ontario, Canada. The Region of Peel, shown in Figure 3-1, is located west of the City of Toronto and consists of three municipalities: the City of Mississauga, the City of Brampton, and the Town of Caledon. The Region of Peel was selected as the case study area due to the clustering of transport and warehousing activities within its boundaries (Ferguson *et al.*, 2014). In 2014, approximately $1.8 billion CAD worth of freight was transported through the Region of Peel daily; this contributed $49 billion CAD to the region’s economy and accounted for 43% of all employment in the region (Region of Peel, 2015, 2017). The Region of Peel has two major intermodal facilities, the Pearson International Airport and the Canadian National (CN) Brampton Intermodal Terminal. These are Canada’s largest air freight facility and largest intermodal rail terminal, respectively (David Kriger Consultants Inc. and CPCS, 1992; Region of Peel, 2017). This clustering of goods movement industries has resulted in roadways with high percentages of HCVs, including Derry Road as shown in Figure 3-2. The large number of HCVs corresponds to a large demand for HCV parking. The management of goods movement has become a priority for the Region of Peel, providing a unique opportunity in this thesis to examine an area that has both high numbers of HCV rest stop events and a large amount of background literature relating to goods movement.

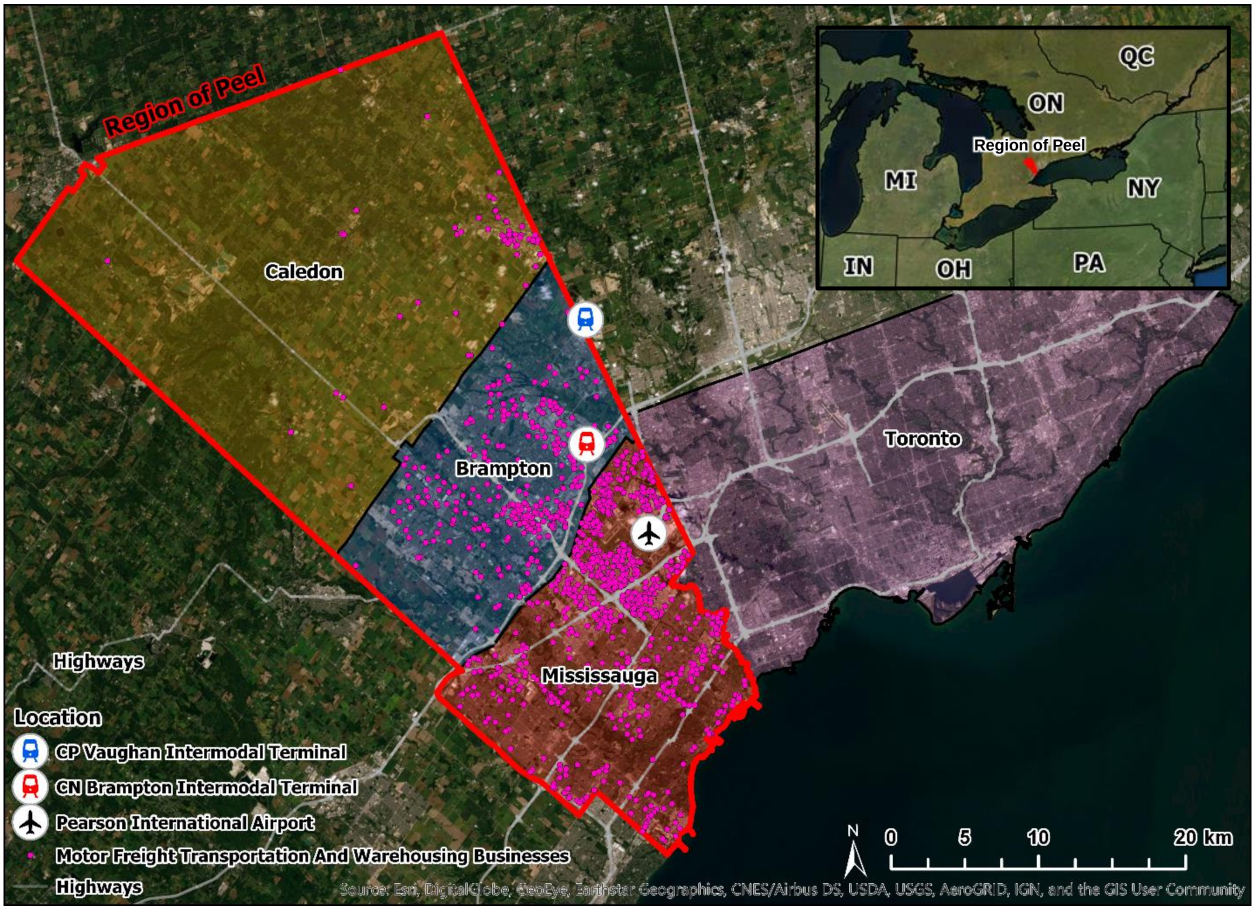


Figure 3-1: Region of Peel Case Study Area

The movement of goods are managed through various internal and external policies at the Region of Peel. The Peel Goods Movement Strategic Plan 2017-2021 document outlines the approach to managing the growing movement of goods within the region (Region of Peel, 2017). This five-year plan does not address HCV parking inadequacies directly but does highlight some concerns related to roadside HCV parking. The Region of Peel Goods Movement Long Term Plan discusses the lack of HCV parking supply as an emerging issue (Region of Peel, 2019a). Moreover, the Region of Peel sees the use of commercial connected and autonomous vehicles as a potential future solution to reducing HCV parking demand. The overall demand for goods movement is predicted to increase due to ecommerce, which is predicted to create a large need for logistics facilities and HCV parking supply. The Region of Peel's most recent Long Range Transportation Plan discusses intelligent transportation system (ITS) technologies that could be used for parking control and management; however, it does not consider ITS technologies in relation to freight (Region of Peel, 2019b). The previous Long Range Transportation Plan emphasized infrastructure improvements to benefit the large trucking industry; however, it also cited the challenges related to the construction, maintenance, and operation of infrastructure projects and discussed the possibility for partnerships with the private sector (Region of Peel, 2012a).

Figure 3-2: 2019 HCV Traffic on Derry Road in the Region of Peel
Photo courtesy of Ravichandra Rampure

Various road and land use classifications exist in the Region of Peel and can be integrated as part of this research project. The Strategic Goods Movement Network Study designated various regional roads as ‘primary truck routes’ and ‘connector truck routes’ (Region of Peel, 2013b). These routes are contrasted with roads that have HCV access restrictions put in place with municipal and regional bylaws, including both all day and partial day restrictions (Region of Peel, 2019d). Regional roads were further classified as part of the Region of Peel's Road Characterization Study, which established guidelines for idealized roadway cross sections and access control guidelines (Region of Peel, 2013a). Lastly, the region has established Generalized Land Use classifications (Region of Peel, 2014) as part of their Regional Official Plan (Region of Peel, 2016).

The Region of Peel has various avenues for goods movement research and advocacy. The Region of Peel Goods Movement Task Force was created in 2009 to promote and advocate for the efficient movement of goods in the region (Region of Peel, 2012b). Similarly, the Smart Freight Centre, a partnership with York University, McMaster University, and the University of Toronto, was established in 2018 to undertake evidence-based research, decision support, advocacy, training, and monitoring related to goods movement. The purpose of the Smart Freight Centre is to "improve the economic vibrancy of business, environmental sustainability, and quality of life for residents of the Greater Toronto and Hamilton Area [by coordinating] transportation infrastructure, land development, regulation, technology tools, and resources that improve goods movement activities" (Smart Freight Centre, 2019a). A study of HCV parking is mentioned as a future research area in the Smart Freight Centre Five-Year Work Plan (Smart Freight Centre, 2019b).

For the analysis in this thesis, the Region of Peel is divided into twenty-six traffic analysis zones (TAZ) based on the municipal wards used in the 2011 Transportation Tomorrow Survey (TTS) (Data Management Group, 2011). A map of these TAZ is shown in Figure 3-3. TAZ are numbered sequentially in each municipality: eleven in the City of Mississauga, ten in the City of Brampton, and five in the Town of Caledon.

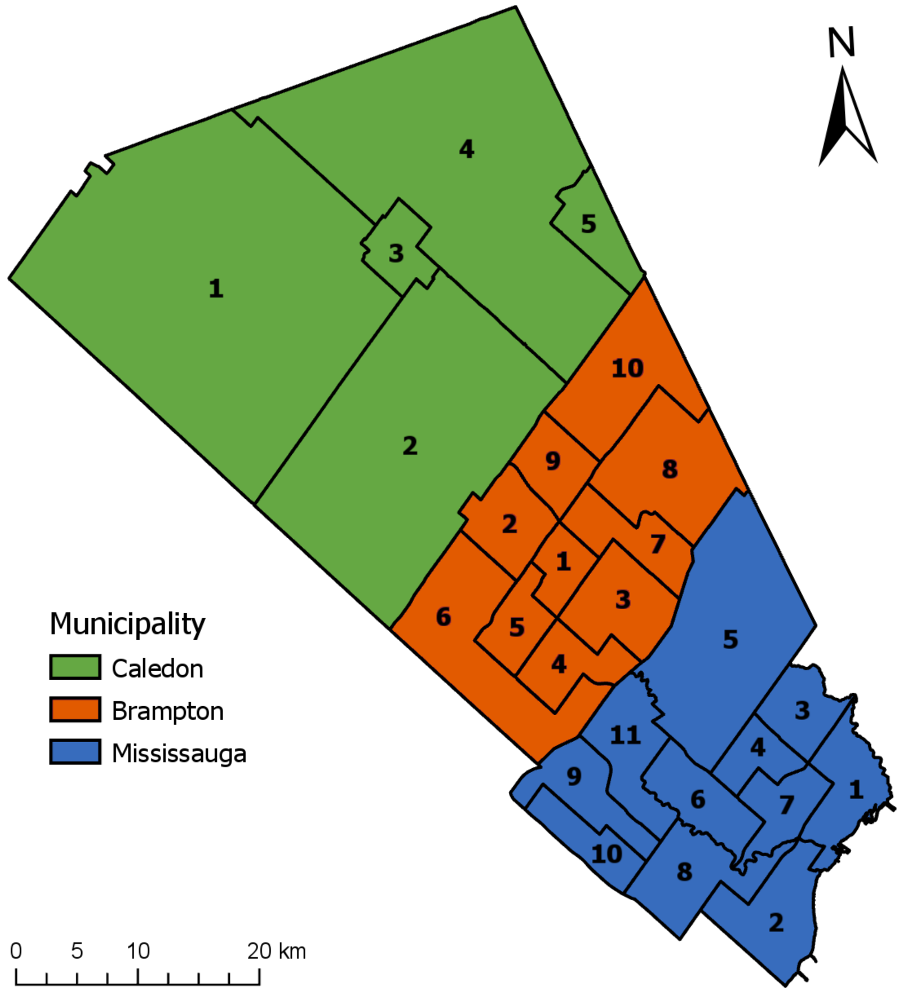


Figure 3-3: Region of Peel Traffic Analysis Zones (TAZ) Map
TAZ correspond to municipal wards as utilized in the 2011 Transportation Tomorrow Survey

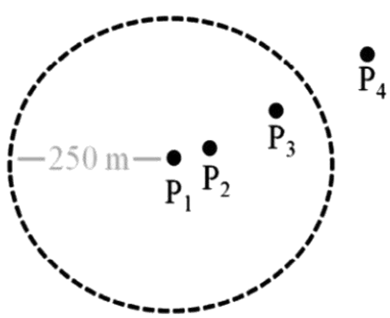
3.2 Data Sources

Various data sources were used in the analysis of this thesis. The key data source used to identify HCV parking locations is GPS pings observing HCV travel patterns in North America. Additional data sources included land use and transportation datasets, firmographic datasets, and various zoning schemes. Most datasets utilized either had a geospatial component or were geocoded to correspond with locations on the GIS mapping surface.

3.2.1 GPS Data of HCV Locations

The main dataset used in this thesis is a list of stop events processed from raw GPS data as described by Gingerich *et al.* (2016a, 2016b). The obtained data consisted of approximately 27 million HCV stop events from all over North America, with approximately 7.9 million occurring in the province of Ontario and nearly 1.9 million stop events in the Region of Peel. Approximately 3.3 million identified trips with origins and destinations occurring across North America were connected to the stop event data to provide additional context. This dataset was collected in the year 2014. The data contained unique vehicle identification numbers, time-stamps, and GPS coordinates of stop events.

The stop event data that was used in this thesis was previously identified from GPS pings originating from 28,487 HCVs belonging to 493 Canadian carriers. These HCVs had GPS data collected and stored within onboard data loggers at various intervals. Stop events were then identified using a Structured Query Language (SQL) database platform based on the vehicle staying within a 250 meter radius for at least 15 minutes, as illustrated in Figure 3-4. In this figure, Ping 3 would be identified as a stop event with a dwell time of 25 minutes (Gingerich *et al.*, 2016b).



Ping 1: $t = 0$ min, $d_w = 0$ min
Ping 2: $t = 5$ min, $d_w = 5$ min
Ping 3: $t = 25$ min, $d_w = 25$ min
Ping 4: $t = 30$ min, $d_w = 0$ min

where:
$t =$ cumulative time
$d_w =$ dwell time

Figure 3-4: Criteria for an Identified Stop Event (Gingerich *et al.*, 2016b)

GPS, or Global Positioning System, is a type of global navigation satellite system (GNSS) developed by the United States Air Force with the purpose of using the position of observed satellites to measure three-dimensional location, velocity, and precise time data in real-time (U.S. Air Force, 2015; Xu and

Xu, 2016). Data that is collected can be stored for further analysis at a later time, a capability has been readily adopted by the transportation industry since it directly or indirectly provides information that is valuable for determining policy and meeting planning needs (Zmud *et al.*, 2014).

The spatial accuracy of the GPS ping data used in this thesis was previously determined by calculating the perpendicular distance of a group of pings from two high-volume road segments. The nominal average perpendicular distance between the GPS pings and the road segments were 15 meters and 28 meters respectively, with 95th percentile distances of 90 meters and 145 meters, respectively (Gingerich *et al.*, 2016b). Factors that may lead to spatial inaccuracies of the GPS ping include the quality of the receiver, atmospheric conditions, signal blocking by obstructions, and multipath error (U.S. Air Force, 2017).

The purpose of stop events can be classified as primary or secondary. Primary stops occur when there is a transfer of goods between the HCV and the stop location (Gingerich *et al.*, 2016a, 2016b). These stops are different than secondary stops, which are defined as stops that meet driver's secondary concerns, such as refueling, eating, and using the restroom (Gingerich *et al.*, 2016a, 2016b).

For the HCV stop data provided, primary and secondary stops have previously been identified using the concept of entropy. An entropy index (EI) value was separately provided for each of the HCV stop events. Entropy is a principle that is used to "describe the state of order or chaos of a given system" (Gingerich *et al.*, 2016b). For this data, the EI value numerically expresses the variation of carrier fleets at a particular stop location. Expressed numerically, the EI of stop location q is:

$$EI_q = -\sum_{c=1}^{C}\left[\left(\frac{n_c}{N}\right)\ln\left(\frac{n_c}{N}\right)\right] \tag{7}$$

where n_c is the number of identified stop events for carrier c at location q, N is the total number of identified stop events at location q, and C is the total number of carriers at location q (Gingerich *et al.*, 2016b). An EI value of 0 relates to a location where stop events have been identified from only one carrier, with larger EI values representing locations with a more heterogeneous mixture of carriers (Gingerich *et al.*, 2016b). The general trend between heterogeneity and EI is shown in Figure 3-5. Stops with an entropy value larger than 2.75 are classified as secondary stops and stops with an entropy value smaller than 1.6 were classified as primary stops. Stops with an entropy value between 1.6 and 2.5 were classified as secondary stops unless there was a known firm within 200m of the stop location. This classification of stops as either primary or secondary can be useful in validating HCV rest events during long-haul trips as these are likely to be secondary stops.

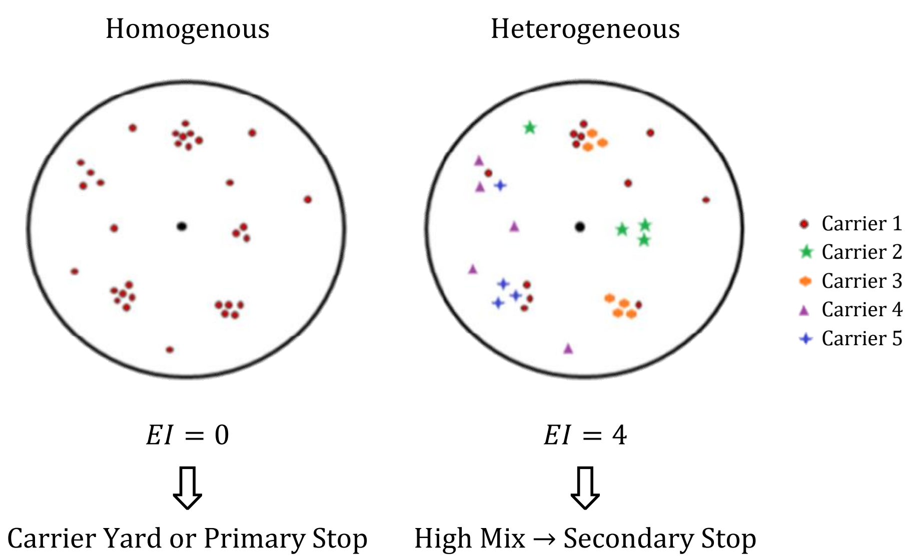


Figure 3-5: Interpretation of Entropy Index
(Gingerich *et al.*, 2016b)

Further information about the previously completed processing of the GPS dataset can be found in two reports prepared by Gingerich *et al.* (2016a, 2016b). Details on the post-processing of the GPS data for this thesis can be found in Section 5.2.

3.2.2 Land Use and Transportation

Data related to land use and transportation is necessary for the analyses outlined in Chapter 6: Development of an HCV Parking Classification Scheme, Chapter 7: Application of the HCV Parking Classification Scheme, and Chapter 8: Estimation of HCV Parking Supply.

Generalized Land Use data for the Region of Peel was obtained from the Peel Data Centre for the year of 2014 to match the time frame of the collected GPS data (Region of Peel, 2014). Generalized Land Use refers to a high-level land classification system used in the Region of Peel for land use and transportation planning, which includes categories such as Commercial / Industrial, Residential, and Rural / Agricultural. Generalized Land Use data was used as an ownership indicating variable in Section 6.2.1.

Road network elements are required to determine the location of roadways and ramps. The Ontario Road Network, which included information such as street names, road classifications (including ramps), and jurisdiction, was created by the Ontario Ministry of Natural Resources and obtained from Scholars Geoportal (Ontario Council of University Libraries, 2019). The road network was used as a roadside indicating variable in Section 6.2.1.

Data was used to identify the locations where HCVs legally could and could not park to satisfy HOS requirements. Information on HCV access and parking restrictions were taken from the City of Mississauga's 'The Traffic By-Law 555-00' (2000), the City of Brampton's 'Unauthorized Parking By-Law 104-2018' (2018) and 'Traffic and Parking By-Law 93-93' (2015), as well as the Town of Caledon's 'Parking on Private Property By-Law No. 2018-1' (2018), 'Traffic By-Law No. 2015-058' (2019), and 'Idling of Motor Vehicles By-Law No. 2014-078' (2014). This information was supplemented using parking violation ticket information obtained from Mississauga Open Data (City of Mississauga, 2019) and Brampton geoHUB (City of Brampton, 2019). The Mississauga parking ticket information was geocoded using ArcGIS Pro (Esri, 2018), the Brampton parking ticket data was already georeferenced, and no parking ticket information was able to be obtained for the Town of Caledon. The HCV parking violation information and parking restrictions outlined in the bylaws were used as legality indicating variable in Section 7.1.1.

Orthoimagery data consists of areal photography that is orthorectified and georeferenced, meaning the imagery has real-world coordinates and has been geometrically corrected to account for scale variations that result when taking the images. Ontario SPOT Pansharpened Orthoimagery data, produced by the Ontario Ministry of Natural Resources, was obtained from Scholars GeoPortal (Ontario Council of University Libraries, 2019). This raster data was captured using two satellites, SPOT 4 and SPOT 5, between 2005 and 2010. The data consisted of three multispectral bands at a 20 meter ground sample distance (GSD or pixel spacing) and a panchromatic band at 10 meter GSD. The Ontario Ministry of Natural Resources further processed this data, pan-sharpening and enhancing the images in Ontario to form a set of 10 metre multispectral (infra-red, red, and green) images which can be used to differentiate between ground coverings. This dataset was used in the supervised image classification in Section 8.3.2.

Default basemaps, as provided by ESRI in ArcGIS Pro, are used in several maps produced in this thesis to show locations in relation to satellite imagery (Esri, 2018). Google Maps 3D satellite imagery was also used to identify and count HCV parking spaces (Google, 2019) in Section 8.2.1. The 3D nature of this mapping application made it easier to identify HCV parking spaces since buildings or other objects occasionally obscure ground features at some angles.

3.2.3 Firmographic

Firmographic data relates to characteristics of businesses or firms similar to how demographic data relates to characteristics of populations and people (Van Wissen, 2000; Maoh and Kanaroglou, 2013).

Firmographics can be helpful in analyzing businesses as it can indicate correlations related to business types and attributes.

Business types and locations were acquired from the Enhanced Points of Interest (EPOI) dataset as produced by DMTI Spatial Inc. (2017). Each EPOI in the point shapefile includes four-digit Standard Industry Classification (SIC) information. SICs are standardized codes that represent various business types (OSHA, 2019). This dataset was available through Scholars GeoPortal (Ontario Council of University Libraries, 2019). In recent years, SIC has been superseded by the North American Industry Classification System (NAICS); however, this newer system was not used for the analysis in this thesis (Statistics Canada, 2017). SICs were used to identify indicating variables in Section 7.1.1 and as independent modelling variables in Chapter 8: Estimation of HCV Parking Supply.

Shapefiles containing the locations of gas stations and weigh stations were also obtained from Scholars Geoportal (Ontario Council of University Libraries, 2019). The coordinates of the Toronto Pearson International Airport, the CN Brampton Intermodal Terminal, and the CP Vaughan Intermodal Terminal were extracted from Google Maps (Google, 2019). These locations were used to develop indicating variables in Section 7.1.1.

Building footprint shapefiles were obtained for Mississauga, Brampton, and Caledon from Mississauga Open Data (City of Mississauga, 2019), Brampton geoHub (City of Brampton, 2019), and the Peel Data Centre (Region of Peel, 2019c), respectively. These shapefiles contained the outline of buildings on the ground derived from aerial imagery. These building footprint shapefiles were used in estimating the amount of paved area in property parcels in Section 8.3.2.

3.2.4 Zoning

Various systems of zoning were used for the analysis in this thesis. At the microscopic scale, individual land parcels were used to represent the locations where HCVs park. Information from the firms that exist on these land parcels can then be used to identify the corresponding HCV parking type in Chapter 7: Application of the HCV Parking Classification Scheme. Individual land parcels are identified by using the Ownership Parcel shapefile as defined by Teranet Enterprises Inc., the Ontario Municipal Property Assessment Corporation, and the Ontario Ministry of Natural Resources (Ontario Ministry of Natural Resources, 2010). Ownership Parcels are defined as plots of land that have been registered or deposited in Ontario's Land Registry or Land Titles Systems and cover most of Southern and Eastern Ontario. The Teranet parcels were obtained from Scholars GeoPortal (Ontario Council of University Libraries, 2019).

At the mesoscopic scale, the Region of Peel municipal wards, as utilized in the 2011 Transportation Tomorrow Survey (TTS), were chosen for HCV parking location count estimates developed in Chapter 8: Estimation of HCV Parking Supply (Data Management Group, 2011). For this thesis, the municipal wards used in the TTS were referred to as Traffic Analysis Zones (TAZ).

The TAZ shapefile was obtained from the Data Management Group at the University of Toronto (Data Management Group, 2011). The 2011 TTS was the most recently completed TTS at the initiation of this thesis and is used in the Region of Peel's current Travel Demand Forecasting model (IBI Group, 2016). These TAZ divide up the Region of Peel municipalities into wards of various sizes: five in the Town of Caledon, ten in the City of Brampton, and eleven in the City of Mississauga.

Municipality and regional boundaries, as created by DMTI Spatial Inc. (2017), were obtained from Scholars Geoportal (Ontario Council of University Libraries, 2019). These boundaries were used in various maps in this thesis.

Origins and destinations of HCV trips were identified using several separate jurisdictional levels (Gingerich *et al.*, 2016a). In Canada, census divisions (CD) as defined by Statistics Canada were selected as zonal boundaries (Government of Canada, 2019). In the United States, zonal boundaries were defined by metropolitan statistical areas (MSA); however, if an MSA did not exist for an area, the county boundaries were used instead (U.S. GSA, 2019).

CHAPTER 4: ANALYSIS METHODS

Analysis methods were required to examine the geospatial datasets to identify where HCVs are parking, develop and apply an HCV parking classification scheme, and estimate HCV parking supply. This chapter discusses various analysis techniques, modelling approaches, and goodness-of-fit measures.

4.1 Data Analysis Techniques

This thesis utilizes specialized software and various data analysis techniques for spatial and statistical analysis of the datasets. Spatial analysis is the examination of the patterns and correlations that exist in geospatial data in order to identify problems and connections that were previously not known (Esri, 2012). Spatial analysis is a necessary component for examining HCV parking locations. Statistical analysis techniques are necessary to ensure accurate results due to the complexity, diversity, and stochastic nature of transportation problems (Washington *et al.*, 2011).

4.1.1 Software Selection

Appropriate software must be selected in order to identify, classify, and estimate the HCV parking supply as part of this thesis. Various types of software exist and can be considered, including macroscopic traffic simulation software, microscopic traffic simulation software, computer-aided design (CAD) software, geographic information system (GIS) software, and statistical computing programming software.

Macroscopic traffic simulation software is used to model transportation systems based on the relationships between traffic flow, speed, and density. Sections are examined instead of modelling individual vehicles in the network (Alexiadis *et al.*, 2004). An example of a macroscopic traffic simulation software is Emme (INRO, 2017). This type of software could be used to estimate the movement of HCVs between zones in the network, which could assist in understanding HCV parking demand; however, this software is not appropriate for the objectives of this thesis.

Microscopic traffic simulation software is used to model individual vehicles in a transportation corridor or network using car-following and lane-changing theories (Alexiadis *et al.*, 2004). An example of a microscopic traffic simulation software is Vissim (PTV AG, 2018). This software could be used to simulate specific parking layouts, including roadside parking and its interaction with through traffic; however, this software is not appropriate for the objectives of this thesis.

Computer-aided design (CAD) software are used in the transportation industry for create and examine models of physical infrastructure. AutoTURN is a CAD software used to examine vehicle swept path to ensure that adequate space is provided to permit HCV parking maneuvers (Transoft Solutions, 2015). This software can be used to analyse the geometry of HCV parking layouts; however, this software is not appropriate for the objectives of this thesis.

No existing commercial transportation software is specifically designed to identify, classify, and estimate the number of HCV parking locations in a region. Due to this, GIS software and statistical computing programming software were used in conjunction to accomplish the objectives of this thesis. ArcGIS Pro (Esri, 2018) and R (R Core Team, 2013) were selected as they can pass information to each other through the establishment of an R-ArcGIS bridge (Esri Canada, 2018). Additionally, EmEditor was used to view and format excessively large tabular datasets to ensure that they could be used in ArcGIS Pro and R (Emurasoft Inc., 2018).

The analysis for this thesis was possible due to software made available with academic licenses including ArcGIS Pro (Esri, 2018), EmEditor (Emurasoft Inc., 2018), and R (R Core Team, 2013).

4.1.2 Spatial Analysis

For the spatial aspects of the analysis, the Geographic Information System (GIS) used was ArcGIS Pro (Esri, 2018). A GIS is a computer software that connects geographic and descriptive information in a multi-layered virtual spatial environment (Esri, 2012). ArcGIS Pro is used to plot the GPS stop events to determine their locations in relation to land parcels and TAZ. ArcGIS Pro is capable of performing various types of spatial analysis techniques in a fast and efficient manner. The capabilities of ArcGIS Pro are greatly expanded when combined with R, a language and environment for statistical computing (R Core Team, 2013). ArcGIS Pro is capable of directly interfacing with R through an R-ArcGIS bridge, which allows each program to access, analyse, and modify datasets without the need for exporting or importing (Esri Canada, 2018). This bridge is created by accessing the R-ArcGIS Support Geoprocessing option in ArcGIS Pro after installing the `arcgisbinding` package in R.

Spatial autocorrelation exists in data with a spatial component when values are related in some way to their neighbours (Fischer and Wang, 2011). When data has significant spatial autocorrelation, it is important to incorporate spatial dependence and/or appropriate spatial variables within the model. Global indicators of spatial autocorrelation include Moran's I and Geary's c. Local indicators of spatial association (LISA) include local Moran's I, the Getis local statistic, and the Ord local statistic. The results of the HCV parking supply model were aggregated to the Traffic Analysis Zone (TAZ) and checked for spatial autocorrelation to examine the possibility of a spatial component. More information can be found in Appendix C.6: Spatial Autocorrelation.

The scope of the thesis covers HCV parking supply, which does not change over a short-term period across hours or days; therefore, an exhaustive temporal analysis is not included. Temporal analysis is necessary if the research extends into the realm of parking demand. Additional information related to the temporal distribution of HCV rest stop events from this thesis can be found in Appendix B: Details of HCV Stop Event Temporal Analysis.

4.1.3 Statistical Analysis

It is important to properly examine a dataset before using it for modelling purposes. Statistical models have various assumptions that must be met, so it is important to statistically analyze data before incorporating it into a model. This includes checking for data outliers, multicollinearity, heteroskedasticity, normal distribution, overdispersion, and spatial autocorrelation. These tests are used to develop the model in Chapter 8: Estimation of HCV Parking Supply.

Outliers are observations that appear to lie outside a large number of similar observations (Washington *et al.*, 2011). For the analysis in this thesis, outliers are removed using three statistical tests: Cook's distance, leverage, and Mahalanobis distance. If an observation is found to fail all three of these tests, it is omitted from the model. For the development of the HCV parking supply model, these tests are used to screen and remove the outliers from the parking count dataset. More information on the outlier removal process is found in Appendix C.1: Outliers.

Multicollinearity is the state when two or more independent variables are correlated with each other, which can effect the accuracy of a model (Washington *et al.*, 2011). The variance inflation factor and the multicollinearity condition number are two measurements that are used to identify multicollinearity for each independent variable and the overall model, respectively. For the development of the HCV parking supply model, these tests are used to ensure that the selected independent variables are truly independent. More information can be found in Appendix C.2: Multicollinearity.

Heteroskedasticity is the presence of non-consistent variance in one variable across the range of another (Washington *et al.*, 2011). Tests to identify heteroskedasticity include the Goldfeld–Quandt test, the Breusch-Pagan test, the Koenker-Bassett test, the Park test, and the White test. Each test identifies heteroskedasticity in its own way, and each test requires certain assumptions about the dataset to be met. For the development of the HCV parking supply model, these tests were used to ensure the validity of the regression modelling techniques and other statistical tests, as well as to obtain a better understanding of the variance. More information can be found in Appendix C.3: Heteroskedasticity.

Normality, or the normal distribution supposition, refers to the condition where a dataset or a model's error follows a normal distribution, which arises due to the central limit theorem (Washington *et al.*, 2011). Indicators of normal distribution include a skewness value of zero and a kurtosis value of three. Tests to determine if a dataset is normally distributed include the Jarque-Bera test, the Lilliefors test, and the Anderson Darling test. For the development of the HCV parking supply model, normality was examined to ensure the applicability of other statistical tests. More information can be found in Appendix C.4: Normality.

Overdispersion occurs in a dataset when the variance is greater than the mean and affects what type of model can be applied to a particular dataset (Washington *et al.*, 2011). Overdispersion is often examined when determining the appropriateness of the Poisson regression model and the negative binomial regression model, the former of which cannot handle overdispersion. The negative binomial regression is able to accommodate overdispersion using an overdispersion parameter. More information can be found in Appendix C.5: Overdispersion. A negative binomial regression is used to develop models in Chapter 8: Estimation of HCV Parking Supply.

4.2 Regression Analysis

Regression analysis is the process of developing mathematical expressions of relationships between variables with the purpose of prediction or explanation (Washington *et al.*, 2011). Various types of regressions can be used to investigate interdependence between variables with the purpose of developing a model (Washington *et al.*, 2011). It is important to ensure that models have the proper level of sensitivity, or power, to properly identify the variable correlation (Shmueli, 2011; Beck, 2013). Predictive and explanatory accuracy must be balanced based on the requirements of the model (Shmueli, 2011).

4.2.1 Selection of Model Type

There are various model types that can be considered for creating the HCV parking supply model; however, for this analysis only four are considered in detail: 1) a multiple linear regression, 2) a Poisson regression, 3) a negative binomial regression, and 4) a zero-inflated negative binomial regression. In Chapter 8: Estimation of HCV Parking Supply, only the latter two model types were examines; the former two could not be used due to limitations as discussed below.

The simplest model type often examined for multivariate modelling is a multiple linear regression model. Due to its simplicity, multiple linear regressions are extensively used to model a wide variety of relationships. A multiple linear regression model has the form:

$$y_i = a + b_i x_i + \varepsilon_i \tag{8}$$

where y_i is the dependent variable for each observation i, a is the model constant, b_i is the model parameters, x_i is the independent variables, and ε_i is the model error. This model is not appropriate for an HCV parking supply count model as it allows negative values to be predicted (Washington *et al.*, 2011). A multiple linear regression model can be developed in R using the `lm()` function (R Core Team, 2013).

A more complex model is the Poisson regression model, where the probability of observation i being value y_i is:

$$P(y_i \mid \mu_i) = \frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!} \tag{9}$$

where μ_i is the Poisson parameter, which is equal to the expected value for y_i. One common relationship between this parameter and the independent variables is:

$$\mu_i = \exp\left(x_i^T \beta + \varepsilon_i\right) \tag{10}$$

where x_i is a matrix of the independent variables and β is a vector of the model parameters. This shows that:

$$y_i \sim x_i^T \beta \tag{11}$$

The Poisson regression is a count model, meaning that it is designed to predict count data (non-negative integers). One limitation of the Poisson regression model is that it assumes no overdispersion of the dependent variable; therefore, the mean and the variance of this variable should be roughly equal (Washington *et al.*, 2011). A Poisson regression model can be developed in R using the `glm()` function with the argument `family=poisson()` (R Core Team, 2013); this function estimates regression coefficients using the maximum likelihood method (Zeileis *et al.*, 2008).

A more robust count model is the negative binomial regression, which is a more generalized version of the Poisson regression that features an overdispersion parameter to account for overdispersion in the dependent variable. For a negative binomial regression model, the probability of observation i being value y_i is:

$$P\left(y_i \,\middle|\, \mu_i, \alpha\right) = \frac{\Gamma\left(\frac{1}{\alpha} + y_i\right)}{\Gamma(y_i + 1)\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i} \tag{12}$$

where $\Gamma(\,)$ is the gamma function and α is the overdispersion parameter (Washington *et al.*, 2011). The calculation of this overdispersion parameter is discussed in Appendix C.5: Overdispersion. A negative binomial regression model can be developed in R using the `glm.nb()` function as part of the `MASS` package (Ripley, 2018); this function estimates regression coefficients using the maximum likelihood method (Zeileis *et al.*, 2008; Ripley, 2018).

In some datasets, observations can have a value of zero for two qualitatively different reasons: the count of zero and the inability to have a count (Washington *et al.*, 2011). This presence of two types of zero in the dependent variable can inflate the overdispersion of a model. To handle this issue, a zero-inflated modelling approach can be used to better process the zero values that occur due to the inability to have a count. In this regard, another model to consider is the zero-inflated negative binomial regression model. This approach applies a binomial choice model, such as a logit model, to the data to determine which observations should be automatically set to zero before performing a negative binomial regression. This model then takes the form:

$$y_i = \begin{cases} 0 \quad \text{with probability } p_i + (1-p_i)\left[\dfrac{\frac{1}{\alpha}}{\left(\frac{1}{\alpha}\right) + \mu_i}\right]^{\frac{1}{\alpha}} \\ y \quad \text{with probability } (1-p_i)\left[\dfrac{\Gamma\big((1/\alpha) + y_i\big)u_i^{1/\alpha}(1-u_i)^{y_i}}{\Gamma(1/\alpha)y_i!}\right], \; y = 1,\, 2,\, 3, \dots, \\ \qquad\qquad u_i = \dfrac{(1/\alpha)}{[(1/\alpha) + \mu_i]} \end{cases} \tag{13}$$

where p_i is the probability of being in the zero-state determined from the logit model, and y is the value that is determined from the negative binomial regression model. Zero-inflated models require two types of zeros to exist for the dependent variable; the approach is not designed for all datasets with many zero value observations (Washington *et al.*, 2011). A zero-inflated negative binomial regression model can be developed in R using the `zeroinfl()` function with the parameter `dist = "negbin"` as part of the `pscl` package (Jackman *et al.*, 2017); this function estimates regression coefficients using the maximum likelihood method and uses a logit model by default for the binomial choice model (Zeileis *et al.*, 2008; Jackman *et al.*, 2017).

In order to determine if a zero-inflated model is required, transportation engineering texts often recommend using a Vuong test (Washington *et al.*, 2011). Recent literature suggests that using a Vuong test to identify zero-inflation is incorrect (Wilson, 2015). Due to this, both the standard and zero-inflated negative binomial models should be considered if zero-inflation is suspected but not known, and goodness-of-fit tests should be used *post hoc* to identify the optimal model.

It can be noted that with sufficiently large means, both Poisson and negative binomial distributions begin to approximate a normal distribution (Montgomery and Runger, 2011). The use of a regression model with a normal distribution can be examined for regression models with sufficiently large means.

4.2.2 Statistical Power

It is important to ensure that an experiment has adequate power, or sensitivity, to detect the desired result. Statistical power $(1-\beta)$ is affected by the significance level (α), the effect size (Cohen's d), and the sample size (n) (Cohen, 1992). As the α and Cohen's d are set based on convention, only the sample size can be adjusted to ensure adequate power is achieved (Beck, 2013). It is recommended that the sample size be determined *a priori* to ensure that adequate power is obtained (Shmueli and Koppius, 2011). The R function `negbin.samp()` from the `RSPS` package was used to calculate the required sample size based on power (Bimali *et al.*, 2015).

4.2.3 Explanatory and Predictive Modelling

Statistical models are used in transportation planning and engineering for various purposes, mainly as a means of explaining some phenomenon or predicting new observations (Shmueli, 2011). Models with high explanatory power are often assumed to be highly predictive; however, this is not always the case (Shmueli, 2011).

Predictive modelling is often favoured in transportation engineering for many reasons. Transportation engineering datasets are often very large and complex with many relationships between variables that are difficult to infer. Predictive modelling allows engineers to develop models that can predict phenomenon with high accuracy despite not necessarily being explanatory. Additionally, predictive models help set measurable benchmarks and they have explicit recognition that not all relationships between variables are known or understood (Shmueli, 2011).

Statistical models will possess varying levels of explanatory and predictive accuracy. For the purpose of developing a model to estimate the number of truck parking spaces in a geographic area, it is most important that the predictive accuracy is maximized. The variables that are selected in the final model should provide logical results in terms of explanatory accuracy. Illogical relationships between the independent and dependent variables can highlight potentially erroneous results in the model. Various tests are performed on the data to explore for factors that can negatively affect the results (example: multicollinearity, heteroskedasticity, etc.); however, the explanatory nature of the model is used as a final check to ensure the results are logical (Shmueli, 2011). This approach is know as 'goodness-of-logic' (Miaou and Lord, 2003).

4.3 Goodness-of-Fit Tests

There are various ways to determine how predictive a model is. The most common approaches to testing goodness-of-fit are numerical methods; however, some graphical methods also exist.

In addition to examining the numerical and graphical goodness-of-fit tests, the residuals of the model can be examined for characteristics such as normality and spatial autocorrelation (Fischer and Wang, 2011; Washington *et al.*, 2011).

4.3.1 Numerical Goodness-of-Fit Methods

Multiple numerical methods for evaluating goodness-of-fit tests will be used in this thesis since there is no single test that can independently identify the most favourable model. Previous studies have stressed the importance of using multiple goodness-of-fit tests for determining best-fitting models since individual tests can provide inconsistent results (Lord and Park, 2008; Shmueli, 2011; Young and Park, 2013; Oluwajana, 2018). The numerical goodness-of-fit tests that are used to assess the accuracy of the final model include the Mean Squared Error (MSE), Mean Squared Predicted Error ($MSPE$), Mean Prediction Bias (MPB and MPB_{holdout}), Mean Absolute Deviation (MAD and MAD_{holdout}), Akaike's Information Criterion (AIC), corrected Akaike Information Criteria (AIC_c), and Bayesian Information Criterion (BIC). It should be noted that the MPB and MAD are calculated for both the data used to create the model and a holdout dataset. Also reported is the two-times log-likelihood at convergence value ($2LL(\theta)$) which is used to calculate the AIC, BIC, and AIC_c. A larger $2LL(\theta)$ value, meaning the negative value closest to zero, indicates a better model fit. More information on these numerical goodness-of-fit methods are found in Appendix D: Numerical Goodness-of-Fit Methods.

4.3.2 Graphical Goodness-of-Fit Methods

In addition to numeric methods for analysing goodness-of-fit, two graphical methods are used in this thesis. A commonly used graphical method is a Quantile-Quantile (Q-Q) plot, where the quantiles of the data residuals are plotted against the standard normal distribution quantiles or the predicted variables are plotted against the actual (Washington *et al.*, 2011; Aguinis *et al.*, 2013). A good model fit will have a nearly linear trend line. An example of a Q-Q plot is shown in Figure 4-1. A Q-Q plot can be generated in R using the functions `qqnorm()` for comparing data to a normal distribution and `qqplot()` for comparing two datasets (R Core Team, 2013).

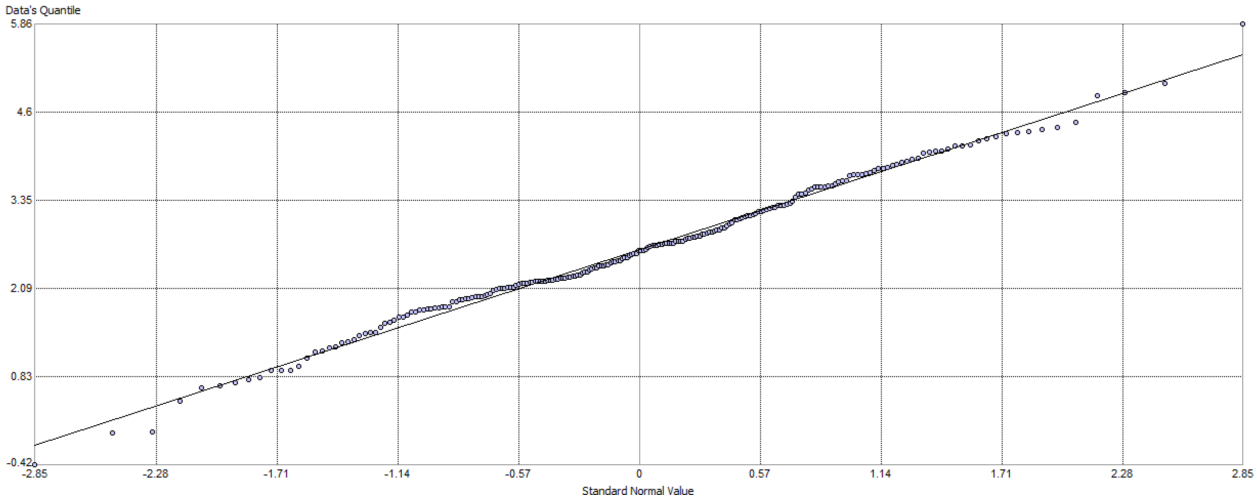


Figure 4-1: Example of a Q-Q Plot
(Gingerich, 2013)

Another useful graphical goodness-of-fit tool is the Cumulative Residual (CURE) plot (Hauer, 2015). The CURE plot examines cumulative residuals against an ordered variable of interest, referred to as an exposure variable. There are several benefits to this approach including being able to examine a clearer representation of the residuals and an indication of where the model is more and less predictive over the range of another variable (Hauer, 2015).

A good CURE plot will show the cumulative residual line staying within a certain range of confidence limits, typically set as two times the standard deviation of an unbiased random walk of the cumulative residual ($\pm 2\sigma_s'(i)$), which expressed mathematically is:

$$\pm 2\sigma_s'(i) = \pm 2\sqrt{\sum_{j=1}^{i}\left(y_j - \hat{y}_j\right)^2}\sqrt{\left(1 - \frac{\sum_{j=1}^{i}\left(y_j - \hat{y}_j\right)^2}{\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}\right)} \text{ for } i = 1,2,\dots,n \ ; \ j = 1,\dots,i \quad (14)$$

where i is the index location of the dependent variable sorted by the smallest to largest exposure variable, n is the total number of observations, y_j is the actual dependent value, and $\hat{y}_j$ is the predicted dependent value (Hauer, 2015). An example of a CURE plot is shown in Figure 4-2. CURE plots can be generated in Excel by ordering the residuals by the exposure variable and plotting the cumulative residuals along with the confidence limits.

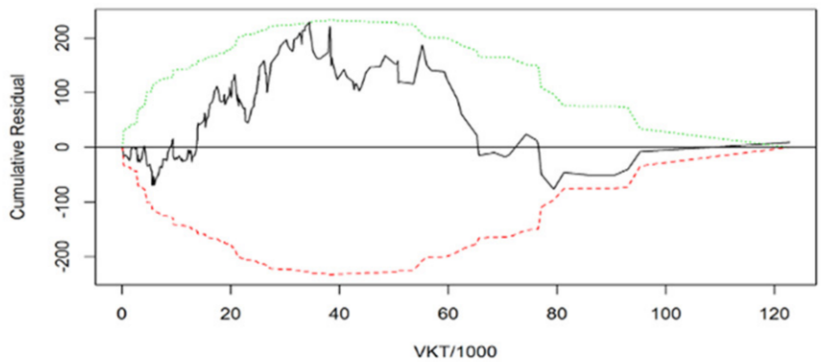


Figure 4-2: Example of a CURE Plot with Confidence Limits
(Takyi *et al.*, 2018)

A major benefit of CURE plots is the increase in clarity related to where a model is accurate within the range of an exposure. As the cumulative residual line climbs, the model is underpredicting the dependent variable. Vertical segments in the cumulative residual line can potentially indicate outliers in the dataset. If the cumulative residual line falls outside of the confidence limits further improvements may be required for the model; however, attempts to restrict the cumulative residual line by adding additional independent variables can lead to an overfitted model. CURE plots can be generated for each independent variable used in the model. (Hauer, 2015).

A model with sufficient goodness-of-fit would produce a CURE plot with a cumulative residual line that is close to zero for the full extent of the exposure variable, with roughly even amounts of the line above and below the zero point, and with a cumulative residual that converges to zero at the maximum value of the exposure variable (Hauer, 2015).

CHAPTER 5: IDENTIFICATION OF HCV PARKING LOCATIONS

It is important to identify the locations where HCV drivers park in a given area to meet their HOS requirements on long-haul trips in order to properly analyze the extent of any HCV parking deficiencies. To do this, sufficient analysis of geospatial data is required. This includes identifying HCV stop events, identifying HCV tours and trips, determining which HCV stop events are conducted for the purpose of meeting HOS requirements, and assigning these rest events to land parcels that can be identified as HCV parking locations.

5.1 Approach to Identifying HCV Parking

In order to properly classify HCV parking locations and estimate their capacity, it is first necessary to identify where HCV vehicles are parking during long-haul trips. This problem can be broken down into several key steps:

1. Identify the point locations of HCV stop events within the study area;
2. Identify the trips that corresponds with these HCV stop events;
3. Determine the HCV stop events of sufficient duration that occurred during a long-haul trip;
4. Identify the property parcels where these HCV rest events occurred; and
5. Verify the HCV parking locations using stop purpose.

This problem can be solved using geographic information systems (GIS) and data analysis. For the analysis, a combination of ArcGIS Pro (Esri, 2018) and R (R Core Team, 2013) is used by establishing an R-ArcGIS Bridge between the programs (Esri Canada, 2018). This allows for seamless integration between the two programs, with R able to access, analyse, and modify datasets that are within ArcGIS Pro without the need for exporting and importing data.

5.2 Identification of HCV Stop Events

The HCV data used for this thesis, as described in Section 3.2.1, consists of GPS pings originating from 28,487 HCVs belonging to 493 Canadian carriers. HCV stop events were previously obtained by Gingerich *et al.* (2016a, 2016b). This previous analysis identified stops from vehicles that stayed within a 250 meter radius for a time not less than 15 minutes. For each of these stops, spatial and temporal information was provided as shown in Table 5-1.

Table 5-1: HCV GPS Ping Data Format

Variable Name	Variable Meaning	Example Format
PointID	Numerical identifier of the GPS Ping at the end of an HCV's stop event	3067492536
CarrierID	A five-digit unique identifier for a carrier company	19999
TruckID	A unique numerical identifier for an HCV which is prefixed by the CarrierID	19999-130824
Latitude	The North-South component of the HCV's stop event geographic location	43.71947
Longitude	The East-West component of the HCV's stop event geographic location	-79.72179
DwellTimeEnd	The date and time at the end of the HCV's stop event	2014-08-27 6:33:39
DwellTimeMinutes	The duration of the stop event in minutes	402.9060

The HCV stop event dataset was imported into ArcGIS Pro for analysis. The HCV stop events that were located within the boundaries of the Region of Peel were isolated using the 'Select By Location' tool. There were nearly 1.9 million stop events identified within the Region of Peel of Peel in the year 2014; however, it should be noted that this dataset represents a sample of the total stop events that would have occurred. The HCV stop event dataset only represents Canadian drivers from 493 carriers, which may introduce some bias into the analysis. The potential bias was not considered pertinent to HCV parking supply locations; however, this would be a concern for estimating parking demand. These HCV stop events include stops for all purposes and various durations; therefore, further analysis is required to isolate stops corresponding to HCV driver stop events that meet HOS requirements.

5.3 Examining HCV Trips and Tours

To identify the HCV stop events corresponding to rest stops on long-haul trips, it is necessary to examine HCV tours. For this part of the analysis it becomes important to distinguish between trips and tours (Hunt and Stefan, 2007; Gingerich *et al.*, 2016b). A trip would correspond to an HCV's movement from an origin to a destination with a unique purpose, most likely to pick up and/or drop off goods. A tour is a series of one or more trips made sequentially by the same HCV with the origin of the first trip corresponding to the destination of the final trip. A tour could take several days and potentially require one or more rest events in order to comply with HOS laws. The HCV stop events of interest for this thesis must occur at some point in the middle of the tour instead of the terminus point at the start and end of the tour. The intermediate stops reflect events occurring to satisfy HOS requirements and therefore fall within the scope of this thesis.

The start and end points for trips within tours correspond with locations where goods are loaded or unloaded; these are primary stops (Gingerich *et al.*, 2016a). Also identified during these trips and tours were the intermediate secondary stops that occurred between the primary stop end points of the long-haul trips. Most rest events correspond with secondary stops; however, rest breaks sometimes occur immediately before or after a transfer of goods at the primary location (FHWA, 2015; Atlanta Regional Commission, 2018). A description of primary and secondary stops is found in Section 3.2.1.

5.4 Identification of HCV Rest Events

In order for an HCV stop event to be considered to be a rest event, they must be of sufficient duration of time and must occur during an identified tour.

The rest events were investigated in relation to hours of service (HOS) regulations. In Canada, long-haul HCV drivers have a mandatory daily break of eight consecutive hours off-duty. If a sleeper cab is attached to the power unit, the eight-hour consecutive time can be split into two periods with each period having at least two hours of consecutive rest (Government of Canada, 2009). Given these regulations, a two-hour threshold was used as the minimum dwell time required for a rest event occurring between tour ends. This threshold removed short stops, such as driver meals or HCV refueling, which are not the focus of this thesis. The upper threshold for rest events was set at fifteen hours, which is 1.5 times greater than the required 10 hours of off-duty time required for drivers every day (Government of Canada, 2009). This provides an allowance for drivers stopping longer then the minimum driver rest period but ignores very long rest periods that indicate an increased likelihood of other activities.

Using the R-ArcGIS bridge, a query was performed to identify the HCV stop events that occurred during an identified tour with a duration between two and fifteen hours. Using these thresholds, a sample of 14,766 rest events for long-haul HCV trips were identified in the Region of Peel for 2014. A heatmap showing the clustering of these rest events in the Region of Peel is shown in Figure 5-1. This figure shows the density of the total identified HCV rest events from the 2014 GPS dataset, representing a sample of the total number of HCV rest events that occurred in the region. The density of all HCV rest events, both those included and not included the sample group, are expected to show the same general distribution.

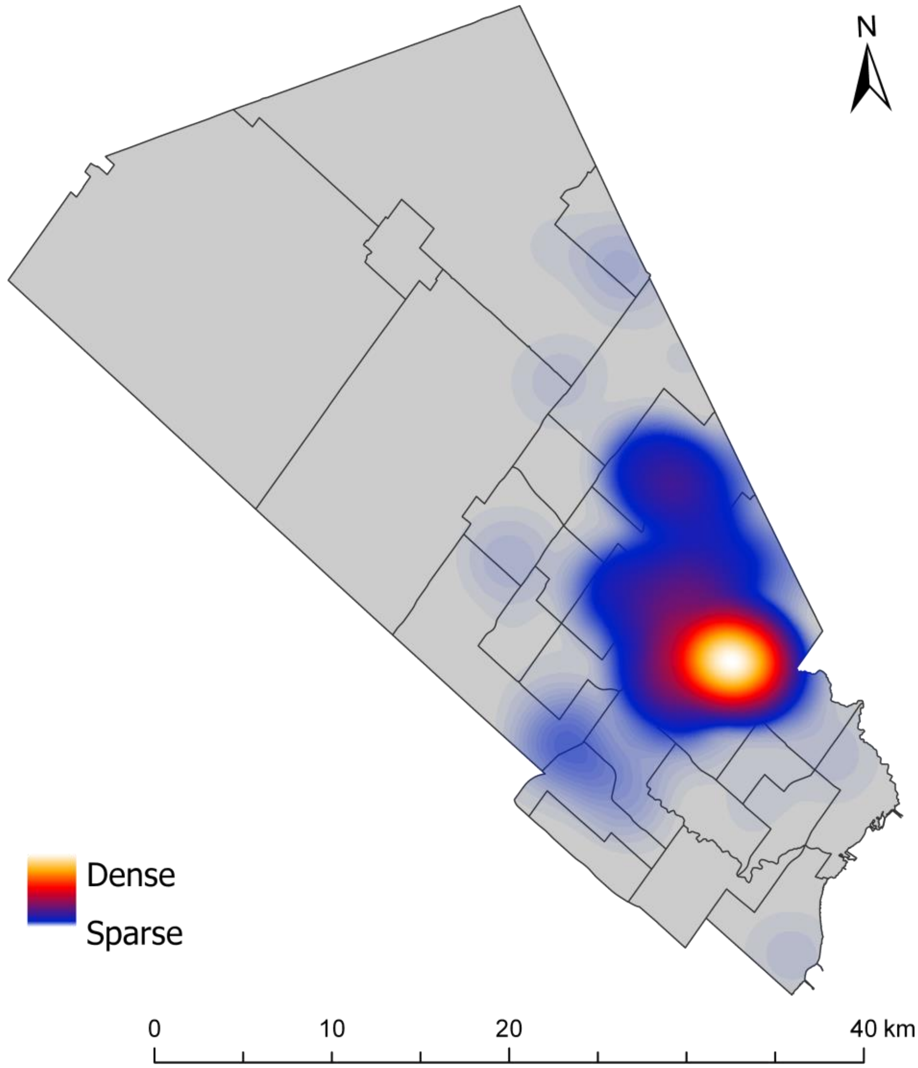


Figure 5-1: Heatmap of Identified HCV Rest Events in the Region of Peel

The clustering of HCV rest events corresponds with what is expected based on the literature review. Most rest events are occurring within close proximity of the Pearson International Airport, the largest freight hub in Canada. The locations where HCV drivers take rest events is discussed further in relation to land use and industrial areas at the end of Section 5.5.

5.5 HCV Parking Locations

In order to identify the common locations where these HCV rest events are taking place, a system for defining property locations was required. In this analysis, individual property parcels created by

Teranet Enterprises Inc. (Ontario Ministry of Natural Resources, 2010) are used to identify possible individual parking locations. Figure 5-2 shows an example of an industrial area with HCV rest events located on individual property parcels. The data shown is fictional, but the layout of stop events presented is typical of data from the Region of Peel.

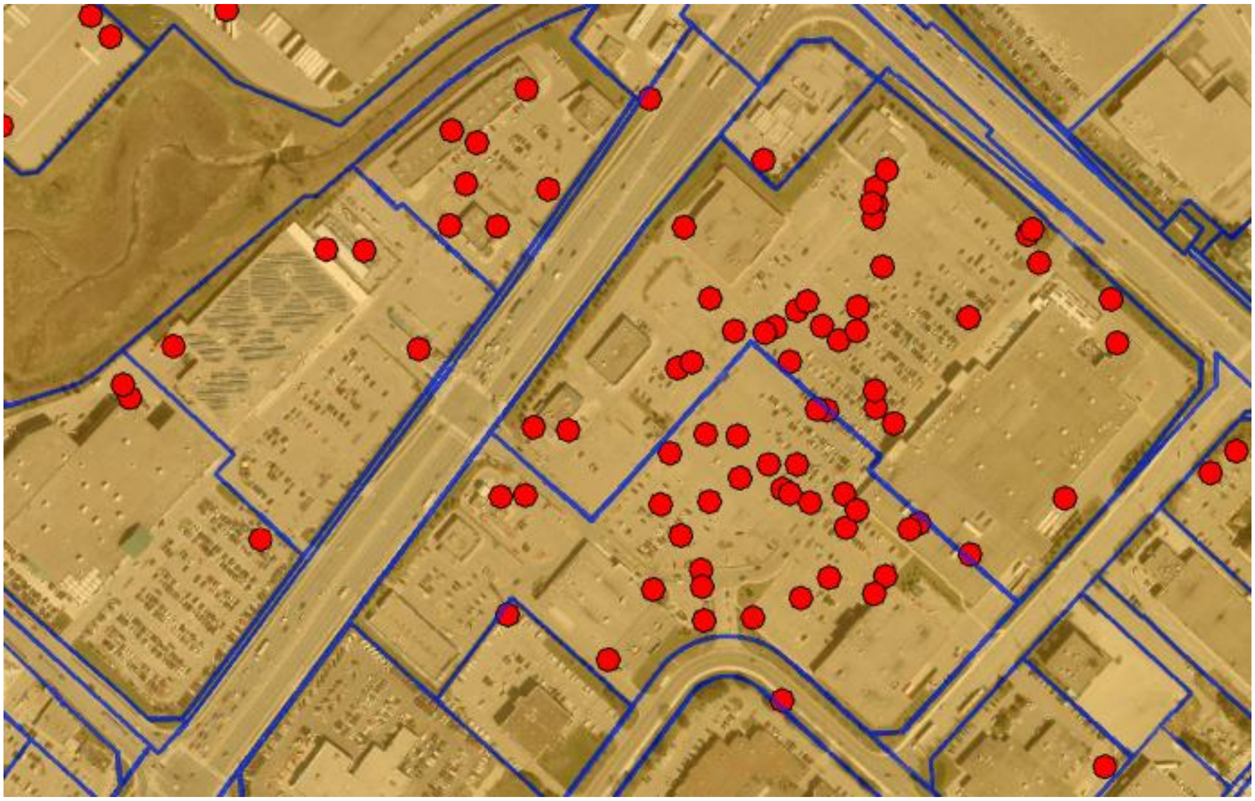


Figure 5-2: GPS Data Showing Stop Events in Property Parcels
Note: Due to confidentiality concerns, stop events shown are fictional and for illustration purposes only

The HCV rest events were assigned to the property parcels using a spatial join in ArcGIS Pro. The 14,766 rest events were located at 1,688 individual property parcels. These property parcels in the Region of Peel are the locations for long-haul HCV parking that are later classified and used to estimate HCV parking supply in this thesis. A heatmap showing the clustering of these HCV parking locations in the Region of Peel is shown in Figure 5-3.

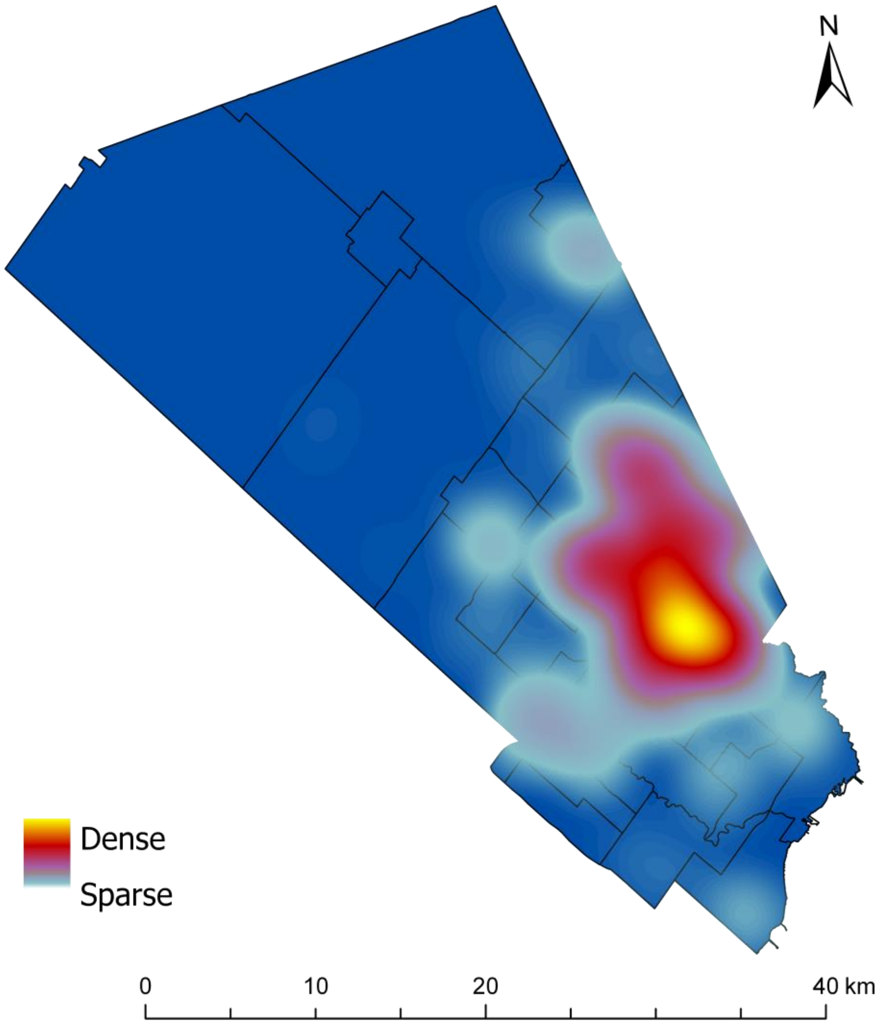


Figure 5-3: Heatmap of Identified HCV Parking Locations in the Region of Peel

The clustering of HCV parking locations in Figure 5-3 is similar to the clustering of HCV rest events shown in Figure 5-1; however, the cluster around Pearson International Airport is less concentrated and several other clusters are more visible. This change in distribution is expected, as less-utilized HCV parking locations will be more visible when the heatmap is not weighted based on the frequency of identified HCV rest events. One noteworthy additional cluster is located north of Pearson International Airport, corresponding with the Town of Bolton (Ward 5 in the Town of Caledon) commercial/industrial area, as identified in the Region of Peel Generalized Land Use classification (Region of Peel, 2014). Further examination of the Region of Peel Generalized Land Use map shows that each cluster visible in Figure 5-3 corresponds with a commercial/industrial land use area.

5.6 Verification of HCV Parking Location using Trip Purpose

This thesis focuses on long-haul HCV drivers and their stop events that meet HOS requirements. Primary stops should correspond with the start or end or long-haul trips, where secondary stops should correspond with stops in the middle of long-haul trips. The classification of HCV stop events as primary or secondary can be used to verify the identification of HCV Parking Locations. A more detailed description of primary and secondary stops is found in Section 3.2.1.

Rest events should, for the most part, correspond with secondary stops. Literature suggests that HCV drivers will sometimes take a rest break at a delivery point; however, this is often not allowed or is discouraged (FHWA, 2015; Atlanta Regional Commission, 2018). It is possible to look at the proportion of rest events that are primary and secondary in nature at various locations. An example of three locations with varying scale is shown in Table 5-2.

Table 5-2: Differences in HCV Rest Events by Scale

Location	Size (km^2)	Number of Identified Rest Events	Proportion of Primary Stops
Region of Peel	1,247	14,766	7.96%
City of Mississauga	292	11,351	7.22%
Mississauga Flying J	0.033	1,355	0.00%

The Region of Peel as a whole and its City of Mississauga each had between seven and eight percent of rest events corresponding with primary stops. This suggests that a small proportion of HCV rest events occur at locations where goods are either picked up or dropped off. In contrast is the Mississauga Flying J, a gas station near Pearson International Airport. This was the location with the largest number of identified rest events. None of these rest events were previously classified as primary stops, which is both an artifact of and is in alignment with the previously determined definition of a primary stop. As described in Section 3.2.1, primary stops were previously identified using an entropy index (EI) value that numerically expresses the variation of carrier fleets at a particular stop location; the Mississauga Flying J is a single location and therefore will have a constant entropy index (EI) value for all identified rest events.

Further verification of the HCV parking locations is performed as part of the application of the HCV parking location classification scheme to the Region of Peel, as discussed in Section 7.3.

CHAPTER 6: DEVELOPMENT OF AN HCV PARKING CLASSIFICATION SCHEME

This thesis proposes a new HCV parking location classification scheme based on the literature review outlined in Chapter 2: Literature Review. The primary criterion for the new scheme is that it should be exhaustive whereby all HCV parking location types are included. In addition, the scheme should be systematically implemented using readily available geospatial data and transferable to other regions in North America. A heuristic stratification is used to develop the proposed HCV parking location classification scheme using feasible combinations of key HCV parking location characteristics.

6.1 Classification Methodology

The proposed HCV parking location classification scheme was developed in three distinct steps:

1. Identify the most important HCV parking location characteristics based on a comprehensive literature review;
2. Analyze all possible combinations of the identified parking characteristics; and
3. Determine an exhaustive and easy to implement classification scheme using a heuristic stratification method.

The five key characteristics of HCV parking locations are combined into sixteen feasible combinations, which are then recombined into nine categories using a heuristic stratification approach. Both one-to-many and many-to-one relationships are used in the recombination of categories to ensure the classification scheme is easily interpretable. The final classifications are partitioned using the findings of the literature review.

The given approach has several key benefits. Firstly, the proposed scheme is designed to be general rather than specific to a jurisdiction. Secondly, each classification can be identified using discrete characteristics obtained from the literature review. This allows for a systematic classification of all HCV parking locations as long as data corresponding to the characteristics are available. Thirdly, the proposed scheme should be easily interpretable and implementable by various jurisdictions across North America. Overall, this approach will result in a classification scheme that is more rigorous than what currently exists while still easy for transportation specialists to understand and apply.

6.2 Identification of Key HCV Parking Characteristics

The first step in the creation of an HCV parking location classification scheme is to identify key HCV parking location characteristics. An extensive literature review was performed in order to identify these key characteristics. The results of this literature review are found in Section 2.5.

The most important characteristics identified from this literature review included 1) legality, 2) accessibility, 3) ownership, 4) dedication to HCV parking, and 5) roadside parking. Each characteristic is presented as a binary variable defined as the presence or absence of the given condition (Everitt and Skrondal, 2010). The meaning of each binary condition for these variables is shown in Table 6-1.

Table 6-1: Binary Conditions for Key HCV Parking Characteristics

Key HCV Parking Characteristic	Binary Condition	
	True	**False**
Legality	HCV parking is permitted	HCV parking is not permitted
Accessibility	HCV parking is available for all drivers	HCV parking is available only for some drivers
Ownership	HCV parking is publicly owned	HCV parking is privately owned
Dedication to HCV Parking	Location is dedicated to HCV parking	HCV parking is not a primary feature
Roadside Parking	HCV parking is on a roadside	HCV parking is not on a roadside

These key HCV parking location characteristics can be coded as binary variables in this thesis as they can only have one of two assigned values: true (1) or false (0) (Everitt and Skrondal, 2010).

6.3 Analysis of HCV Parking Characteristic Combinations

The second step in the establishment of an HCV parking location classification scheme is to examine all the possible combinations of key HCV parking location characteristics. These characteristics will provide insight into the feasible groupings of characteristics that can later be rearranged to match what is frequently discussed in literature.

Table 6-2 presents the twenty-four possible combinations of the five identified key HCV parking location characteristics. Accessibility was considered only for legal HCV parking locations since unauthorized parking locations would be inaccessible for public or private use. As a result, combinations that do not include legal parking have the accessibility characteristic designated as not applicable (N/A).

Table 6-2: All Combinations of HCV Parking Characteristics

ID	Characteristics					Examples Identified in the Literature Review
	Legal	Open	Public	Dedicated	Roadside	
1	Yes	Yes	Yes	Yes	Yes	*None (public roads are not dedicated to parking)*
2	Yes	Yes	Yes	Yes	No	Public rest areas
3	Yes	Yes	Yes	No	Yes	Road shoulder (in areas legally permitted)
4	Yes	Yes	Yes	No	No	Public weigh stations
5	Yes	Yes	No	Yes	Yes	*None (private roads are not dedicated to parking)*
6	Yes	Yes	No	Yes	No	Gas stations, private rest areas
7	Yes	Yes	No	No	Yes	Road shoulder (in areas legally permitted)
8	Yes	Yes	No	No	No	Private weigh stations
9	Yes	No	Yes	Yes	Yes	*None (public roads are not dedicated to parking)*
10	Yes	No	Yes	Yes	No	Public truck trailer parking locations
11	Yes	No	Yes	No	Yes	Publicly-owned limited access roads
12	Yes	No	Yes	No	No	Some government facilities
13	Yes	No	No	Yes	Yes	*None (private roads are not dedicated to parking)*
14	Yes	No	No	Yes	No	Private truck trailer parking locations
15	Yes	No	No	No	Yes	Private roads
16	Yes	No	No	No	No	Some logistics facilities, warehouses, and drop-off locations
17	No	N/A	Yes	Yes	Yes	*None (public roads are not dedicated to parking)*
18	No	N/A	Yes	Yes	No	*None (locations cannot be dedicated to unauthorized parking)*
19	No	N/A	Yes	No	Yes	Public roadways, ramps
20	No	N/A	Yes	No	No	Public parks
21	No	N/A	No	Yes	Yes	*None (private roads are not dedicated to parking)*
22	No	N/A	No	Yes	No	*None (locations cannot be dedicated to unauthorized parking)*
23	No	N/A	No	No	Yes	Private roads
24	No	N/A	No	No	No	Private businesses, passenger vehicle parking and abandoned lots

The combinations from Table 6-2 have been grouped into suitable categories that include both legal and illegal HCV parking locations for Step 3 of the classification development process. There are eight combinations of characteristics that were not found to be feasible and were subsequently disregarded from the final classification scheme. For example, combination 1 contains characteristics including public, dedicated to parking, and roadside, but this combination is not possible since public roads are primarily dedicated to vehicle mobility instead of parking. The eight infeasible categories are printed in grey italicized font in Table 6-2.

6.4 Heuristic Stratification of HCV Parking Classifications

The third and final step in the creation of an HCV parking location classification scheme is to regroup the sixteen feasible combinations of HCV parking characteristics, shown in Table 6-2, using a heuristic stratification based on HCV classifications identified in the literature. This is a necessary step to ensure that the HCV parking classification scheme is easily interpretable and implementable by various jurisdictions throughout North America. The sixteen feasible combinations of HCV parking characteristics were used to create the final categories seen in Table 6-3 using a many-to-many relational structure with both many-to-one and one-to-many relations.

The proposed classification scheme classifies HCV parking locations into nine categories: five authorized and four unauthorized HCV parking location categories. The five authorized HCV parking location categories are: ① Public Rest Areas and Gas Stations, ② Weigh Stations, ③ Open Access HCV Parking, ④ Limited Access HCV Parking, and ⑤ Authorized Roadside Parking. The four unauthorized HCV parking location categories are: ⑥ Unauthorized Roadside Parking, ⑦ Unauthorized Highway Ramp Parking, ⑧ Unauthorized Parking on Public Property, and ⑨ Unauthorized Parking on Private Property. Symbols have been developed to represent each of these categories and are used throughout the rest of this thesis.

Table 6-3: HCV Parking Classification Scheme

Category	Description	Table 6-2 IDs	Key Characteristics				
			Legal	Open	Public	Dedicated	Roadside
① Public Rest Areas and Gas Stations	Legal, publicly accessible locations such as rest areas, gas stations, and truck stops	2,6	Y	Y	Y‖N	Y	N
② Weigh Stations	Legal locations where parking is available adjacent to weigh station infrastructure	4,8	Y	Y	Y‖N	N	N
③ Open Access HCV Parking	Legal, publicly accessible business locations	4,6,8	Y	Y	Y‖N	Y‖N	N
④ Limited Access HCV Parking	Business locations that will typically restrict access to authorized vehicles only	10,12, 14,16	Y	N	Y‖N	Y‖N	N
⑤ Authorized Roadside Parking	Locations where roadside parking is considered legal	3,7, 11,15	Y	Y‖N	Y‖N	N	Y
⑥ Unauthorized Roadside Parking	Locations where roadside parking is not considered legal	19,23	N	N/A	Y‖N	N	Y
⑦ Unauthorized Highway Ramp Parking	Highway entrance and exit ramps where parking is usually illegal	19,23	N	N/A	Y‖N	N	Y
⑧ Unauthorized Parking on Public Property	Illegal, publicly owned locations for HCV parking such as parks, airports, and conservation areas	20	N	N/A	Y	N	N
⑨ Unauthorized Parking on Private Property	Illegal, privately owned locations where no HCVs are allowed to park for extended periods	24	N	N/A	N	N	N

Y = Yes; N = No; Y‖N = Yes or No; N/A = Not Applicable

A many-to-one relationship was utilized for combination ID 1 and ID 2 from Table 6-2 to be combined to create Category ① in the Table 6-3 parking scheme. Locations exhibiting the characteristics of these two combinations include public rest areas (ID 2) and gas stations or private rest areas (ID 6). These combinations were grouped together since no particular rationale could be found for HCV drivers to prefer public parking facilities more than private parking facilities or vice versa as long as the accessible amenities in the two types of parking facilities are comparable to each other.

A one-to-many relationship can be seen where Combination ID 19 has been split into two different categories including illegal roadside parking (Category ⑥) and illegal ramp parking (Category ⑦). The given combination was split into the two categories since they are often discussed separately in the literature (ATRI, 2018; Anderson and Roll, 2018). The other remaining relationships can be identified by connecting the ID in Table 6-3 back to the original combinations in Table 6-2.

The final classification in Table 6-3 includes five categories corresponding to legal parking locations. Public rest stops and gas stations are identified as Category ①, while weigh stations are assigned as Category ② as the latter category is often explicitly mentioned in existing literature (Martin and Shaheen, 2013; Pinjari *et al.*, 2014; FHWA, 2015; Atlanta Regional Commission, 2018). Category ③ and Category ④ pertain to open access HCV parking and limited access HCV parking. Accessibility is utilized as a defining characteristic for those categories to separate parking supply that accommodates all vehicles (open access) from restricted supply that is only available to specific users (limited access). Category ⑤ pertains to roadside locations where parking is authorized.

The remaining four categories in Table 6-3 correspond to unauthorized parking. This includes roadside parking where by-laws do not permit this activity, highway ramps, and public or private property used without permission. The unauthorized roadside parking is separated into roadway parking for Category ⑥ and ramp parking for Category ⑦. Finally, Category ⑧ and Category ⑨ are respectively organized as unauthorized parking occurring on private properties and public properties.

An application of this proposed classification scheme for the Region of Peel in Ontario, Canada is demonstrated in Chapter 7: Application of the HCV Parking Classification Scheme.

CHAPTER 7: APPLICATION OF THE HCV PARKING CLASSIFICATION SCHEME

This chapter provides an application of the proposed HCV parking location classification scheme, as discussed in Chapter 6: Development of an HCV Parking Classification Scheme. The classification scheme is applied to the Region of Peel in Ontario, Canada. The application of the HCV parking classification scheme found in this chapter is currently unique to the Region of Peel; however, the approach is transferrable to other regions since is systematically applied using commonly available geospatial data. This approach of using common data was used in order to make the classification scheme more rigorous and cost effective.

It should be noted that the HCV parking location classification scheme is only applied to the actively used locations identified in Chapter 5: Identification of HCV Parking Locations; therefore, the results do not represent an exhaustive dataset of all possible HCV parking locations, authorized or unauthorized. All locations that do not legally permit HCV parking could be considered as unauthorized; however, it is contended that such an exhaustive list would be less useful since many unused locations are likely undesirable or impractical for HCV parking. The same principal applies for authorized HCV parking, as any space that an HCV could physically park without legal repercussions could be considered an authorized parking space. The approach selected for this thesis is in line with the principle of diminishing returns, since the result of an approach considering every theoretical or conceivable HCV parking location would not be helpful from a policy perspective.

7.1 Classification Approach for the Case Study Area

The application of the HCV parking classification scheme, regardless of case study area, involves three key steps:

1. Identify variables from commonly available geospatial data that can indicate the absence or presence of each HCV parking location characteristic;
2. Convert the identified indicating variables to a binary format; and
3. Develop an algorithm that uses these binary variables to assign an HCV parking location classification to each identified parcel.

This approach identifies the binary state of each characteristic that applies to each location and assigns a classification based on the combination of characteristics. Since each classification is defined by the combination of binary characteristics, it is possible to classify each location if its characteristics are known.

The proposed HCV parking location classification scheme is applied to the Region of Peel in Ontario, Canada. This location was selected as the case study area due to the clustering of the transport and warehousing activities within its boundaries (Ferguson, Lavery and Higgins, 2014).

7.1.1 Identification of Key Characteristic Variables

In Section 2.5, five key HCV parking characteristics were identified: legality, accessibility, ownership, dedication to HCV parking, and roadside parking. These characteristics were used to develop the comprehensive classification scheme found in Chapter 6: Development of an HCV Parking Classification Scheme; however, as shown in Table 7-1 only four of the five key characteristics are required to classify parking locations in part due to a one-to-many relationship that was used in developing the scheme. The 'dedicated' key characteristic is not necessary a defining attribute of any of the nine proposed classifications; instead, the presence of a public rest area, gas station, or weigh station and the accessibility type are used to differentiate non-roadside locations where parking is permitted (Categories ① to ④). All other locations (Categories ⑤ to ⑨) are not dedicated to HCV parking either due to being roadside and/or not authorized for HCV parking. The HCV parking classification scheme can be found in Table 6-3.

In addition to variables indicating four of the five key characteristics, variables must be identified that indicate the presence of a gas station, a weigh station, and a highway ramp. The presence of a gas station, a weigh station, or highway ramp indicates that the HCV parking location should be classified as ① Public Rest Areas and Gas Stations, ② Weigh Stations, or ⑦ Unauthorized Highway Ramp Parking, respectively.

The seven attributes used to apply the HCV parking classification scheme to the Region of Peel are: legality, accessibility, ownership, roadside parking, gas station, weigh station, and highway ramp. The first four of these variables correspond with key HCV parking characteristics and are described in Section 2.5.

Table 7-1: Indicating Variables for HCV Parking Location Characteristics

HCV Parking Characteristic	Indicating Variable(s)	
	Attribute(s)	Data Source(s)
Legality	Land use, bylaws, and ticketing information	Region of Peel
Accessibility	Parcel contains SIC '42' firm[1]	DMTI Enhanced Points of Interest (EPOI) dataset
Ownership	Land use	Region of Peel
Roadside Parking	Parcel contains road segment	Ontario Road Network
Gas Station	Parcel contains SIC '5541' firm[2]	Enhanced Points of Interest (EPOI) dataset
Weigh Station	Parcel contains SIC '4785' firm[3]	Scholars Geoportal Dataset
Highway Ramp	Parcel contains highway ramp segment	Ontario Road Network

[1] SIC category 42 pertains to Motor Freight Transportation and Warehousing firms
[2] SIC category 5541 pertains to Gasoline Service Stations firms
[3] SIC category 4785 pertains to Fixed Facilities and Inspection and Weighing Services for Motor Vehicle Transportation firms

7.1.2 Classification Algorithm and Flow Chart

Figure 7-1 shows a flow chart outlining the algorithm used to allocate property parcels, as identified in Section 5.5, into one of the nine parking location categories outlined in Table 6-3. While the binary variable approach to classification can be applied to any jurisdiction, it may be necessary to consider local laws governing authorized HCV parking since bylaws will likely vary from region to region. Details on the creation of this algorithm follow the flow chart.

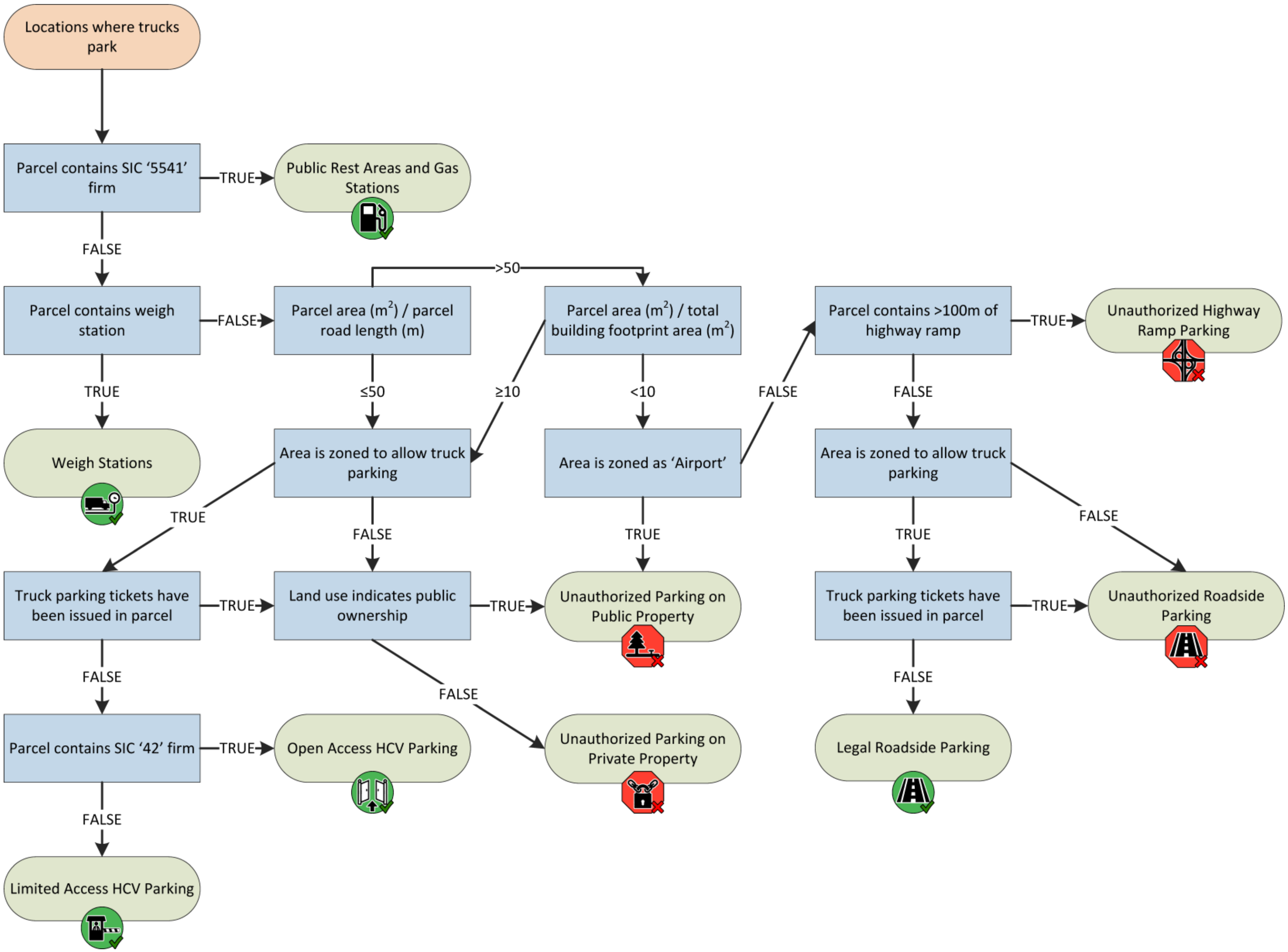


Figure 7-1: HCV Parking Classification for the Region of Peel

The 1,688 property parcels in the Region of Peel identified as HCV rest event locations were classified into the categories described in Table 6-3 using the flow chart in Figure 7-1. ① Public Rest Areas and Gas Stations were identified using the business type of the property parcel using the corresponding four-digit standard industrial classification (SIC) of ‘5541’. There was only one location classified as ② Weigh Stations in the Region of Peel but there were no rest events identified at this location using GPS data.

③ Open Access HCV Parking and ④ Limited Access HCV Parking both pertain to locations consisting of privately-owned businesses, but it was difficult to systematically identify the type of access available to a given business as this information is not publicly available. It was assumed that ③ Open Access HCV Parking applied to transportation logistics and warehousing facilities based on parcels–containing a business with the two-digit SIC for Motor Freight Transportation and Warehousing (‘42’); while ④ Limited Access HCV Parking applied to other businesses. The SIC 42 group contains three subcategories: Trucking and Courier Services, except Air (‘421’), Public Warehousing and Storage (‘422’), and Terminal and Joint Terminal Maintenance (‘423’) (OSHA 2019). This definition of the SIC 42 group and its three subcategories indicates that the presence of one of these firms will work sufficiently to identify ③ Open Access HCV Parking; however, this broad assumption may not hold for all establishments and should be examined in the future.

The ⑤ Authorized Roadside Parking category required several criteria for identification. The property parcel had to contain part of the identified roadway network; however, calibration was required since some non-roadway property parcels contained a small part of the road network. To calibrate this classification process, a sample of 30 locations that intersected the road network were investigated manually. Based on this analysis, a maximum ratio of 50 square meters of property parcel area for each meter of roadway length was established. If this ratio was exceeded, the property parcel was not considered a roadway location and was allocated to a different category. Based on additional analysis, property parcels were also not identified as roadways if more than ten percent of the property parcel area was covered by buildings or if the area was zoned as an airport.

Another condition for ⑤ Authorized Roadside Parking considered the legal status arising from land use zoning categories based on municipal by-laws. Locations were considered legal if a property parcel was zoned as commercial, industrial, rural, undeveloped, or estate land uses and does not experience parking violations. The parking infraction information obtained from Mississauga and Brampton included the time, location, and type of infractions as recorded by parking enforcement officials. Caledon violations were not available for this thesis but could be added in the future. Examples of violations included 'parking heavy vehicles in prohibited areas' and 'parking heavy vehicle in residential zone'.

All roadside parking that were found not to permit legal HCV parking were classified as either ⑥ Unauthorized Roadside Parking or ⑦ Unauthorized Highway Ramp Parking. The later category was distinguished from the former using the roadway data in the shapefile supplied by the Region of Peel. Property parcels containing a highway ramp were classified into this category. A minimum length of 100 meters was found to be useful for avoiding false positives.

⑧ Unauthorized Parking on Public Property and ⑨ Unauthorized Parking on Private Property were considered in terms of land use. Public ownership was assumed to apply to conservation areas, provincial parks, undeveloped settlements, government land, institutions, landfills, or airports. Other land uses were assumed to be privately owned. These locations are determined to be unauthorized for HCV parking either due to the identification of parking violations or if zoning does not permit for HCV parking. Most land types were not zoned to permit HCV parking. Only land use types of employment, rural, undeveloped settlements, and estates were identified to be zoned to authorize HCV parking.

7.2 Results of Classification for the Case Study Area

Table 7-2 shows the results of the classification of the 1,688 property parcels with HCV rest stop events. This table shows the frequency of the five categories of authorized property parcels and the four categories of unauthorized property parcels. The results identify HCV parking locations from a very comprehensive list of classifications which represents an exhaustive list for the case study area. Past HCV parking studies have not included all the categories of HCV parking found and have often ignored a large number of authorized HCV parking spaces.

Table 7-2: Frequency Distribution of Property Parcels by Category

Category	Property Parcels
① Public Rest Areas and Gas Stations	20
② Weigh Stations	1
③ Open Access HCV Parking	227
④ Limited Access HCV Parking	978
⑤ Authorized Roadside Parking	275
Total Authorized Property Parcels	***1,481***
⑥ Unauthorized Roadside Parking	85
⑦ Unauthorized Highway Ramp Parking	36
⑧ Unauthorized Parking on Public Property	19
⑨ Unauthorized Parking on Private Property	47
Total Unauthorized Property Parcels	***187***
Grand Total	**1,688**

Most identified property parcels in Peel were authorized (1,481) whereas there were significantly less identified unauthorized locations (187). It should be noted that these results count the number of total identified locations and not the number of HCV parking spaces at these locations. Most authorized property parcels (978 out of 1,481; 66%) provided limited access parking at private firms that are not transportation/warehousing firms, which have often been overlooked in many previous parking supply studies.

Figure 7-2 shows the spatial distribution of each HCV parking location category. The locations are aggregated and displayed by municipal ward. This figure shows the areas with HCV parking locations most widely used on long-haul trips. Each category needs to be inspected individually as the scale varies from category to category.

Figure 7-2 shows that most parking locations, both authorized and unauthorized, are clustered in the north-east portion of Mississauga, shown to be Mississauga TAZ 5 in Figure 3-3. When these results are compared to Figure 3-1, the north-east portion of Mississauga corresponds to many motor freight transportation and warehousing firms clustered around Pearson International Airport, as is expected for one of the biggest freight trip attractors/generators in the region (Fleger *et al.*, 2002; Atlanta Regional Commission, 2018). Additional locations with significant numbers of authorized HCV parking locations are Brampton TAZ 3 and TAZ 8, which corresponds with what is known as the CN Intermodal Area in the Region of Peel's Strategic Goods Movement Study (Region of Peel, 2013b). A large portion of this land is also classified as commercial / industrial in the Region of Peel's Generalized Land Use map (Region of Peel, 2014).

The unauthorized parking locations can provide insight into where additional legal HCV parking and/or enforcement may be required. Frequent unauthorized HCV parking at a location may indicate that drivers are unable to find suitable parking nearby. Alternatively, it is possible that drivers have become complacent and do not expect parking regulations to be enforced. Reviewing citation records, surveying drivers, and examining the distance to other legal parking locations may assist in determining the causes of unauthorized parking. Unauthorized HCV parking, and in particular ⑨ Unauthorized Parking on Private Property, is distributed fairly uniformly throughout the entire region in both industrial and residential areas. It is worth nothing that this uniform distribution is fairly low in terms of absolute value, with six or less locations identified per TAZ.

Each of the three local municipalities can be examined individually. Within the City of Mississauga, the largest share of authorized parking properties is located within TAZ 5. This is the location of Pearson International Airport and is additionally located near the CN Brampton Intermodal Terminal. There are also significant authorized parking parcels identified in TAZ 2, which contains a large industrial presence. In the City of Brampton, as mentioned previously a substantial share of authorized parking parcels has been identified along the commercial and industrial areas in TAZ 3, 7, and 8. TAZ 8 in Brampton contains Canada's highest volume intermodal rail terminal. As the least populous municipality in the Region of Peel, the Town of Caledon contains the lowest total of authorized parking locations. The largest cluster of legal parking parcels for this municipality were identified in TAZ 5 where the community of Bolton is located.

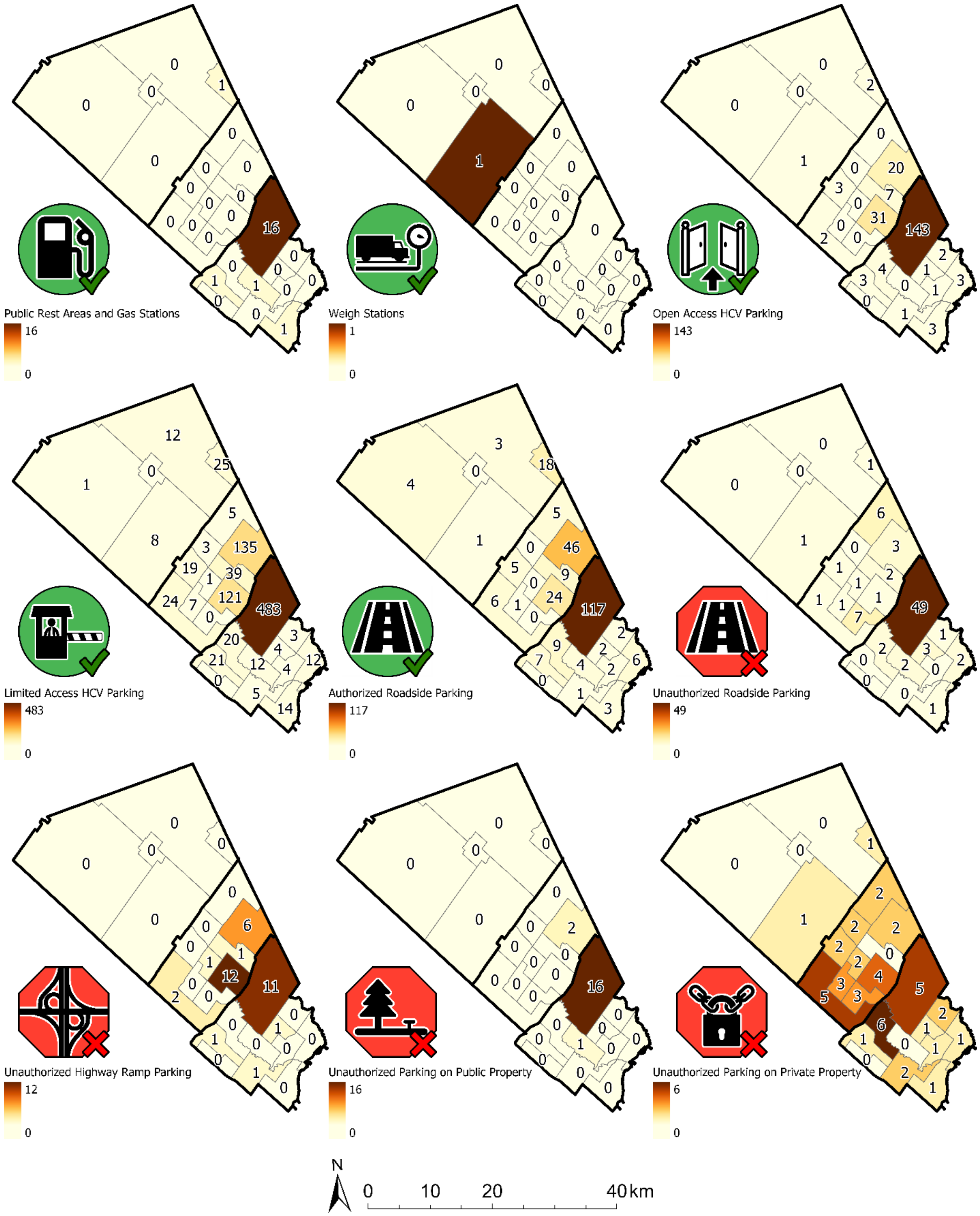


Figure 7-2: Spatial Patterns of Parking Locations for Long-Haul HCV Trips

Note: Map areas are subdivided by TAZ, numeric labels represent counts of classified HCV parking locations

7.3 Confirmation of Results

A total of one-hundred-seventy-one parcels were used to verify the application of the HCV parking classification scheme for the Region of Peel.

A sample of ninety-five parcels was selected from the original 1,688 parcels to confirm that the authorized non-roadside parking location categories in Table 1 were assigned appropriately. Additionally, the number of HCV parking spaces was manually counted to be used in the development of an HCV parking supply model. This sample included twenty ① Public Rest Areas and Gas Stations, eighteen ③ Open Access HCV Parking, and fifty-seven ④ Limited Access HCV Parking locations. In each case, the category and number of visible parking spaces was confirmed to the extent possible using the location data and satellite imagery provided by Google Maps (Google, 2019), as well as the geospatial data used to classify the locations. The sample size of ninety-five parcels was originally selected as a preliminary estimate of the number of HCV parking locations; however, the sample size was later increased to 150 HCV parking locations to obtain adequate statistical power for the development of the HCV parking supply model. Details of this analysis are given in Chapter 8: Estimation of HCV Parking Supply.

The other six categories of HCV parking locations were verified without counting any parking spaces. This is mainly due to the fact that no dedicated parking spaces exist at any of these locations. The single ② Weigh Stations location was examined and confirmed to be categorized correctly. An additional fifteen parcels were selected from each of the remaining categories: ⑤ Authorized Roadside Parking, ⑥ Unauthorized Roadside Parking, ⑦ Unauthorized Highway Ramp Parking, ⑧ Unauthorized Parking on Public Property, and ⑨ Unauthorized Parking on Private Property. These seventy-five total parcels were all determined to be properly categorized to the extent possible using the location data and satellite imagery provided by Google Maps (Google, 2019), as well as the geospatial data used to classify the locations.

These results were obtained for the third iteration of the HCV parking classification algorithm for the Region of Peel as outlined in Figure 7-1. Preliminary versions of the flow chart did not include the thresholds for road and ramp length in parcels. The addition and calibration of these thresholds, along with the final order of the steps and identification of the indicator variables, allowed for the systematic application of the HCV parking classification scheme at a high level of accuracy.

CHAPTER 8: ESTIMATION OF HCV PARKING SUPPLY

This chapter provides a process for developing an HCV parking supply model using the Region of Peel as a case study area. HCV parking supply refers to the total number of designated parking spaces available that can be used to meet HCV parking demand. Although the supply model developed here is specific for the Region of Peel, a similar approach could be used to develop a model for another jurisdiction. It is possible that other jurisdictions would have different variables that correlate with the number of parking spaces and that variables have a different relationship than they do in the Region of Peel. The final HCV parking supply model is presented here along with an estimate of the number of HCV parking spaces in the parking locations identified in the Region of Peel.

8.1 Modelling Approach

In order to ensure the accuracy, reproducibility, and validity of a truck parking supply model, it is important to determine how the model will be developed *a priori*. The approach that was used to develop the HCV parking supply model had the following steps:

1. Develop a procedure for processing the data to be used in the model;
2. Select the appropriate variables to be used in the model;
3. Determine the appropriate model sample size;
4. Create several HCV parking models using different variables and model types; and
5. Select the best model using various goodness-of-fit methods.

The data used, processing techniques, model types, and goodness-of-fit methods to be used are described in Chapter 3: Study Data.

An approach was developed using R to process the data and report only the information pertinent to the current analysis stage (Washington *et al.*, 2011). R is a useful tool as predefined functions can be used in combination with custom functions and custom script to ensure that only the necessary information is reported. A custom script was written in R to perform a data analysis in order to identify outliers and analyse the data structure. This script analyses the data to identify outliers, multicollinearity, heteroskedasticity, normality, and overdispersion using a total of fourteen different statistical tests. Additional information related to this data analysis is found in Appendix C: Details of Data Statistical Analysis. The R script written to perform this analysis is found in Appendix E: R Script for Data Analysis.

8.2 Scope for the HCV Parking Supply Model

The HCV parking supply model developed for this thesis focuses on HCV parking spaces that can be utilized for HOS-related rest events on long-haul trips. An HCV parking space is defined based on HCVs with a single trailer, corresponding to Classes 8 through 10 in the FHWA 13-Category Scheme (Hallenbeck *et al.*, 2014). More information on the design vehicle used for this thesis is found in Section 2.1.2.

The HCV parking space model developed as part of this thesis is only applicable for the Region of Peel. The estimates of HCV parking spaces are based on the parcels previously identified as part of Chapter 5: Identification of HCV Parking Locations and classified in Chapter 7: Application of the HCV Parking Classification Scheme.

8.2.1 Defining an HCV Parking Space

A heuristic approach was used to count the HCV parking spaces that were used in the development of the supply model. Google Maps 3D satellite imagery was used to identify and count the HCV parking spaces (Google, 2019). Figure 8-1 showcases the various visual clues that were used to identify the HCV parking spaces in the satellite imagery.

Figure 8-1: Visual Clues used to Manually Identify HCV Parking Spaces
Adapted from Google (2019)

As part of the verification process for the HCV parking location classification, the HCV parking spaces were counted using the visual clues represented in Figure 8-1 for a total of ninety-five parcels, including twenty ① Public Rest Areas and Gas Stations, eighteen ③ Open Access HCV Parking, and fifty-seven ④ Limited Access HCV Parking locations. After the results of the statistical power analysis, as described in Section 4.2.2, the HCV parking spaces in an additional fifty-five parcels were counted, including an additional eleven ③ Open Access HCV Parking and forty-four ④ Limited Access HCV Parking locations. An example of a parcel with HCV parking spaces identified using this approach is shown in Figure 8-2. It should be noted that the turning radius of HCVs was not used to verify the validity of parking spaces.

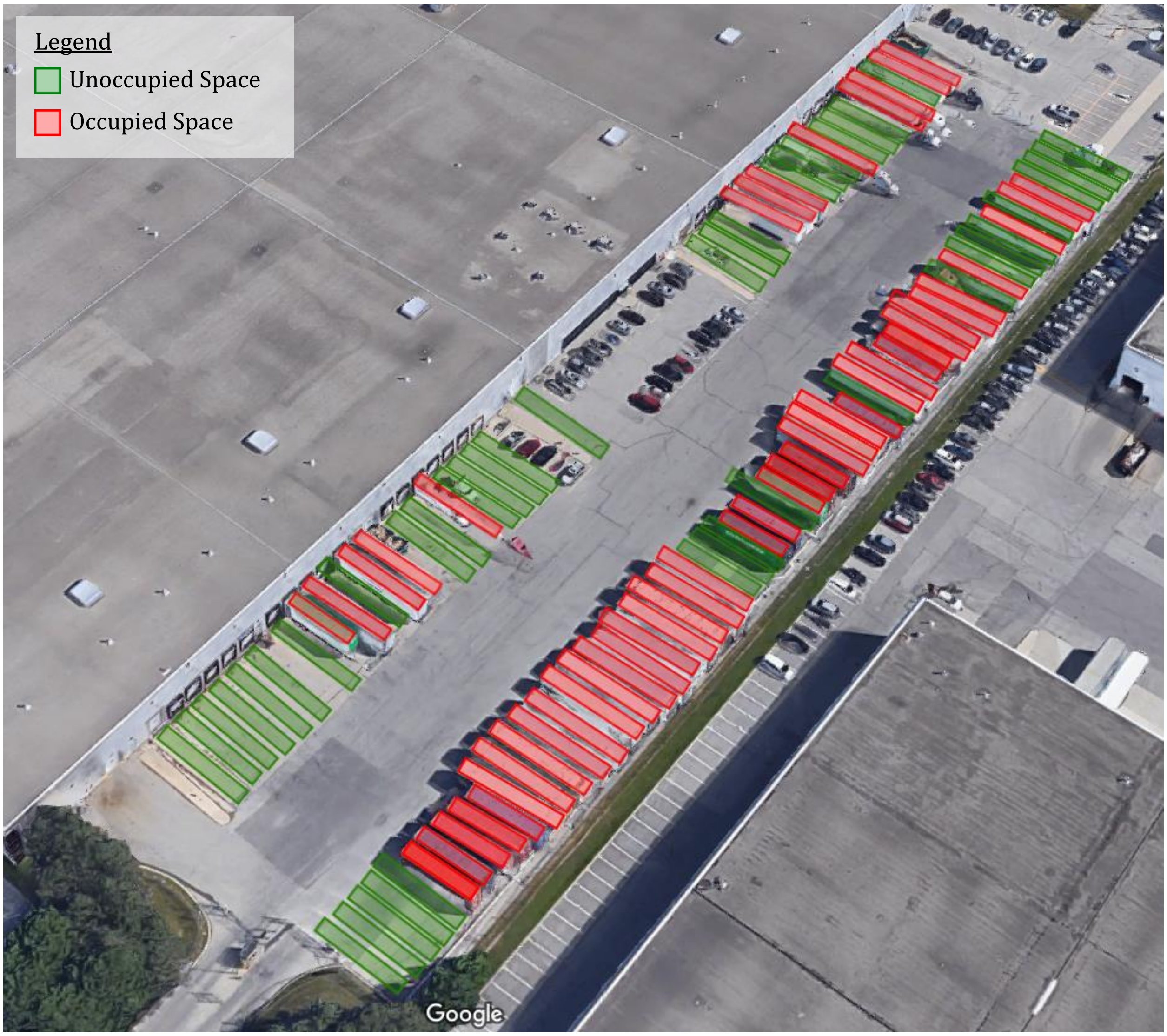


Figure 8-2: HCV Parking Spaces Identified at a Warehouse Location
Adapted from Google (2019)

As computer vision, machine learning, and image processing become more advanced and as more variables become available, it may become possible to automatically count the HCV parking spaces at these locations. Despite it being possible to use GIS software to identify passenger vehicle parking lots (Zhang and Couloigner, 2006; Davis *et al.*, 2010) and count their parking spaces (McKeown *et al.*, 1985; Wang and Hanson, 1998; Seo and Urmson, 2009; Kabak and Turgut, 2010), this is currently not feasible for HCVs due in part to the differences in the consistency of parking space size and markings, as described in Section 2.1.5.

The minimum, average, and maximum number of HCV parking spaces for each HCV parking location classification examined is shown in Table 8-1. These values show that ① Public Rest Areas and Gas Stations are the locations with the smallest number of HCV parking spaces, ③ Open Access HCV Parking are the location with the largest number, and ④ Limited Access HCV Parking locations have a number of HCV parking spaces between the other two. This result is expected, as many ① Public Rest Areas and Gas Stations are smaller in size with less designated HCV parking, many ③ Open Access HCV Parking locations are large warehouses with a large amount of HCV parking, and ④ Limited Access HCV Parking locations tend to be smaller than their open access counterparts. Limited access locations likely serve a smaller number of HCVs compared to open access locations since they serve a limited fleet of drivers from select companies.

Table 8-1: Counted HCV Parking Spaces by Classification

Classification	Quantity Counted	Number of HCV Parking Spaces		
		Minimum	Average	Maximum
① Public Rest Areas and Gas Stations	20	0	14	80
③ Open Access HCV Parking	29	0	122	763
④ Limited Access HCV Parking	101	0	27	408
All	**150**	**0**	**43**	**763**

The number of HCV parking spaces showed a positive skew for the frequency of locations based on the number of parking spaces available, as shown in Figure 8-3 for the sample of 150 HCV parking locations. This distribution also shows a large proportion of locations exhibiting zero designated HCV parking spaces available. Potential outliers in this dataset included six locations with more than 200 HCV parking spaces. These six locations all correspond with large logistics facilities with a large quantity of HCV parking spaces.

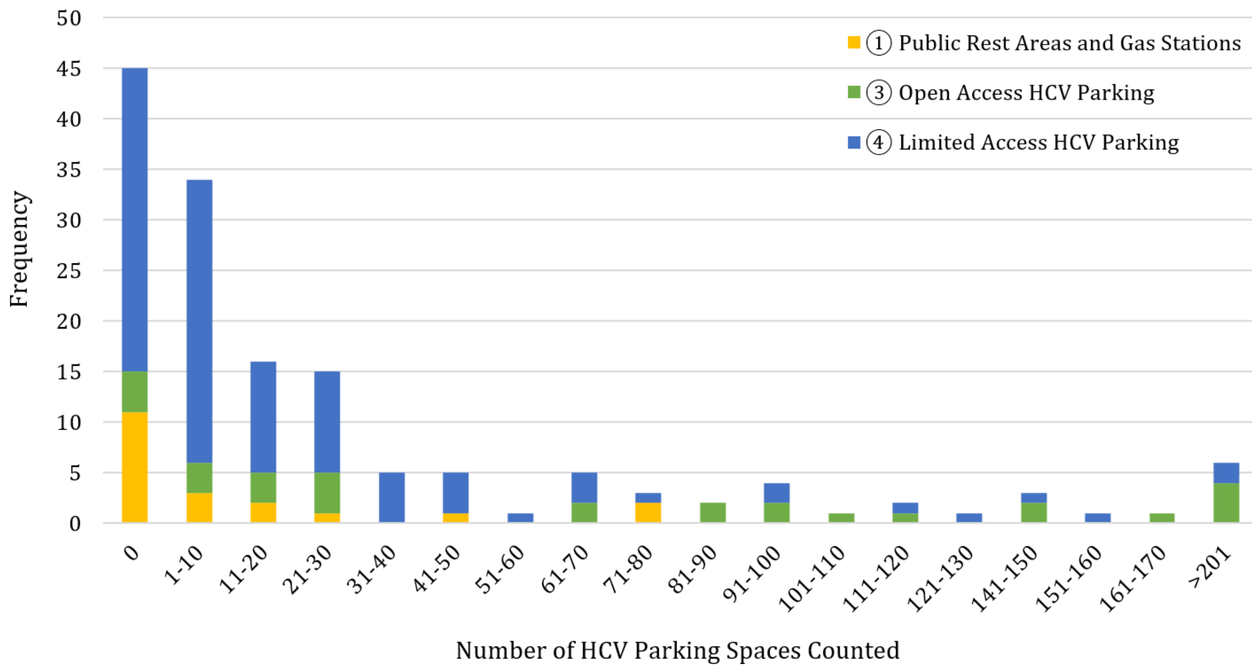


Figure 8-3: Histogram of Parcels Counted by Number of HCV Parking Spaces

8.3 Model Development

The number of HCV parking spaces is utilized as the dependent variable in the models developed for this thesis. A preliminary sample of ninety-five parcels, as identified in Section 7.3, was used to assess the potential variables for the HCV parking supply model.

The distribution and characteristics of the dependent variable influences the modelling type that can be adopted. For example, a multiple linear regression was rejected as a suitable technique in this thesis due to the inability of the model to avoid negative predicted values. In addition, a Poisson regression model was also rejected due to overdispersion: the preliminary sample's dependant variable mean of 49.5 is significantly lower than its variance of 14,489.

Six candidate HCV parking supply models were developed, and the best-fit model was later selected as the final model. Three combinations of variables were used with two separate model types: a negative binomial regression model and a zero-inflated negative binomial regression model.

Typically a Vuong test is used to determine if a zero-inflated model is necessary (Washington *et al.*, 2011); however, a recent study suggests that this is not a correct approach (Wilson, 2015). Due to this, both model types were used, and the final model was selected using both numeric and graphical goodness-of-fit indicators. These goodness-of-fit methods are described in Section 4.3.

8.3.1 Data Processing Procedure

A data processing approach was developed in advance of the model creation to minimize any bias. An algorithm was developed using R language to identify and eliminate outliers, and then test for multicollinearity, heteroskedasticity, normality, and overdispersion. This data processing algorithm is not specifically designed for the datasets used in this thesis; therefore, it is a tool that can also be used for other projects.

One key consideration in the algorithm was the criteria for identifying and eliminating outlier data records. The removal of data records as outliers should not be done *ad hoc* whenever it benefits the accuracy of a model but should be justified using statistical tests (Washington *et al.*, 2011). For this thesis, it was determined that records would be classified as outliers if they fail all three of the Cook's D, leverage, and Mahalanobis distance tests, as discussed in Appendix C: Details of Data Statistical Analysis. The results of these tests will vary depending on the combination of independent variables selected; therefore, the data analysis algorithm will need to be used independently for each model.

Multicollinearity between independent variables should be avoided as this can lead to model inaccuracy and interpretability issues (Washington *et al.*, 2011). The combination of variables selected should not have multicollinearity. The multicollinearity condition number will be used to identify multicollinearity between the independent variables as a whole and variance inflation factors (VIF) will be used to identify specific variables that have issues with multicollinearity. The application of these approaches is discussed in Appendix C: Details of Data Statistical Analysis.

Normality and overdispersion are also examined in the algorithm to identify if it is possible to use a more simplistic model type, such as a multiple linear regression model or a Poisson regression model. The R code for this algorithm is provided in Appendix E: R Script for Data Analysis.

8.3.2 Variable Selection

Several variables were selected as appropriate to use in the HCV parking supply model, including land use and transportation, firmographic, and zoning variables.

It is expected that land use variables related to parcel geometry will have substantial influence on predictive capabilities of the models, with larger land parcels leading to additional parking spaces. Two variables are proposed here but are highly correlated and therefore introduced in separate models. The first of these variables is the perimeter length of a given parcel. This variable is considered simple to introduce in the model and subsequently provides higher transferability for future study areas.

The second geometric variable created for the model represents the total pavement surface area available in each land parcel. Since no dataset initially existed to provide this information, it was created by taking the total parcel area and subtracting the building footprint and any green space or water features identified on the parcel. The proportion of green space and water coverage were calculated from a supervised image classification model created using ArcMap software with three categories including pavement/building footprints, greenery, and water coverage. The supervised image classification training was performed on sixty-five locations with twenty meter by twenty meter raster cells obtained from Landsat satellite imagery. An analysis with a hold-out sample of ten parcels found that the categories assigned to raster cells were 100% correct for nine out of ten locations. The last location in the testing had one raster cell incorrectly identified. Based on this performance, the image classification was considered suitable as part of the calculation for the total pavement surface area in each property parcel.

The other variables that were selected included a binary variable identifying rural parcels identified from Generalized Land Use classifications (Region of Peel, 2014), two HCV parking location classifications identified in Chapter 7: Application of the HCV Parking Classification Scheme, and a binary variable identifying parcels with Miscellaneous Retail firms as identified from the Enhanced Points of Interest (EPOI) dataset. Other industry types were considered; however, lower correlation values were observed, and multicollinearity existed between various firm types. A complete list of variables included in the models developed in this analysis are provided in Table 8-2.

Additional variables were tested during the development of the preliminary models including other HCV parking location classifications, the industry code for other types of firms based on the standard industrial classification (SIC) system, and parcels with low observed demand (occasional use) based on the rest events observed from GPS data. It should be noted that the latter is not included in the model due to low statistical significance and due to it is not part of commonly available geospatial dataset. Binary or continuous variables representing proximity to major goods movement facilities, such as airports or intermodal facilities, could also be considered in addition to land use, transportation, firmographic, and zoning variables.

Table 8-2: Selected Variables for Analysis

Variable	Description
Perimeter_i	Length of perimeter for parcel i (10^{-3} m)
Rural_i	1 if parcel i is rural, 0 otherwise
$\text{Perimeter}_i \times \text{Rural}_i$	Interaction term for Perimeter_i and Rural_i
$\text{Area}_{\text{paved},i}$	Area of pavement in parcel i (10^{-6} m^2)
$\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$	Interaction term for Perimeter_i and $\text{Area}_{\text{paved},i}$
$\text{Class}_{\text{open},i}$	1 if parcel i is classified as ③ Open Access HCV Parking, 0 otherwise
$\text{Class}_{\text{rest},i}$	1 if parcel i is classified as ① Public Rest Areas and Gas Stations, 0 otherwise
$\text{IND}_{59,i}$	1 if parcel i contains a Miscellaneous Retail (SIC59) firm, 0 otherwise

A priori expectations were determined for the selected variables to assist with goodness-of-logic checks. Larger locations are expected to have a larger number of HCV parking spaces, as are locations with a greater amount of paved area. Rural and miscellaneous retail locations are expected to have a lower number of HCV parking spaces, and locations classified as ① Public Rest Areas and Gas Stations or ③ Open Access HCV Parking are expected to have a higher number of HCV parking spaces. Based on these expectations, characteristics that are predicted to have a larger number of HCV parking spaces are expected to have a positive coefficient and characteristics that are predicted to have a lower number of HCV parking spaces are expected to have a negative coefficient.

8.3.3 Sample Size Determination

In order to determine the necessary sample size required to achieve adequate power, the `negbin.samp()` function from the `RSPS` package in R was used (Bimali *et al.*, 2015). Two preliminary HCV parking models were developed using the ninety-five parcels that were used to confirm the authorized non-roadside parking classification in Section 7.3. Unauthorized and roadside parking spaces were not included in this model since an exact count does not exist: any location can be used as an unauthorized parking space and roadside HCV parking spaces are not clearly defined. The number of HCV parking spaces were manually counted for these locations. A pair of negative binomial regression models were established using this preliminary data with the purpose of determining the sample size required for adequate statistical power. The creation of these preliminary negative binomial regression models is required by the `negbin.samp()` function, since the model's overdispersion parameter is required. The first preliminary model featured Perimeter_i as the key geographic variable and the second featured $\text{Area}_{\text{paved},i}$. Additional independent variables were selected by identifying the largest subset of variables that did not have multicollinearity. The preliminary HCV parking supply models are shown in Table 8-3.

Table 8-3: Preliminary HCV Parking Models for Sample Size Determination

Variable	Preliminary Model 1		Preliminary Model 2	
	β	**p**	**β**	**p**
Intercept	2.15	0.00***	3.09	0.00***
Perimeter_i	1.13	0.00***		
Rural_i			−2.60	0.01**
$\text{Perimeter}_i \times \text{Rural}_i$	−2.43	0.00***		
$\text{Area}_{\text{paved},i}$			8.39	0.00***
$\text{Class}_{\text{open},i}$	1.27	0.00***	1.60	0.00***
$\text{Class}_{\text{rest},i}$	−0.44	0.27	−1.77	0.08*
$\text{IND}_{59,i}$	−1.79	0.00***	−2.02	0.00***
AIC	756		767	
$2LL(\theta)$	−742		−753	

(***), (**), (*) represent statistical significance to 99%, 95%, or 90% respectively

The sample size calculation using the `negbin.samp()` function was performed using an $\alpha = 0.05$ and a desired power of $(1 - \beta) = 0.8$ for two levels of effect size: $d = 0.5$ for a medium effect size and $d = 0.65$ for a medium-large effect size (Cohen, 1992). The selected desired power, $(1 - \beta) = 0.8$, was set based on convention proposed for general use (Cohen, 1992). The mean HCV parking space count ($\bar{y}$) was determined to be 49.49 and the results of Preliminary Model 1 were used calculate the overdispersion parameter ($1/\theta$) as 2.256. The seed value and number of simulations were set to their default values, and the significant digits was set as three. The inputs and results of this sample size calculation are shown in Table 8-4.

Table 8-4: Statistical Power Analysis and Sample Size Calculation

Input Variables		Output	Effect Size	
			0.5	**0.65**
Power, $(1-\beta)$	0.8	Actual Power, $(1-\beta)$	0.809	0.805
Mean, $\bar{y}$	49.49	Required Sample Size, n	40	111
Overdispersion, $1/\theta$	2.256	Standard Error	0.012	0.013
Type I Error, α	0.05			
Seed	20			
Number of Simulations	1000			
Significant Digits	3			

Based on this result of this statistical power analysis, a minimum sample size of 40 was required for a medium effect size ($d = 0.5$) and a minimum sample size of 111 was required for a medium-large sample size ($d = 0.65$). A large effect size ($d = 0.8$) would require a minimum sample size of 370 which is not feasible for this study. To ensure a large-enough sample size was used, the HCV parking spaces were counted for a total sample of 150 parcels: 130 parcels to build the model and 20 parcels to act as a hold-out data set. The `negbin.samp()` function was also used for a *post hoc* power check after the final model was selected to ensure adequate power was maintained.

8.4 Candidate HCV Parking Supply Models

Six candidate models were established using the HCV parking space count dataset ($n = 130$) for comparison using various goodness-of-fit methods. The three negative binomial regression models and the three zero-inflated negative binomial regression models were compared. The regression coefficients were estimated using the maximum likelihood method for each of these model types. A logit model was used as the binomial choice model for the zero-inflated negative binomial regression.

8.4.1 Models

The candidate models were created using 130 of the 150 sample locations, with the same 130 locations used in each model. Three combinations of variables were used for the two model types: negative binomial regression and zero-inflated negative binomial regression. The first two variable combinations were selected to be the same as the preliminary model, and the third combination was established as the largest combination of variables that did not have issues with multicollinearity (except with interaction terms) as identified with the data analysis algorithm. The combinations of variables used in the models are shown in Table 8-5. Models 1, 2, and 3 are negative binomial regression models and models 4, 5, and 6 are zero-inflated negative binomial regression models. For the zero-inflated negative binomial regression models, predictions were automatically set to zero if the logit model estimated a greater than 50% chance of being a zero.

Table 8-5: Variable Combinations for Candidate HCV Parking Supply Models

Variable	Models 1 & 4	Models 2 & 5	Models 3 & 6
Perimeter_i	Y	N	Y
Rural_i	N	Y	Y
$\text{Perimeter}_i \times \text{Rural}_i$	Y	N	N
$\text{Area}_{\text{paved},i}$	N	Y	Y
$\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$	N	N	Y
$\text{Class}_{\text{open},i}$	Y	Y	Y
$\text{Class}_{\text{rest},i}$	Y	Y	N
$\text{IND}_{59,i}$	Y	Y	Y

Y = Yes; N = No

The dataset with 130 records was first processed with the data analysis algorithm to identify and remove outliers, diagnose multicollinearity, and check for normality and overdispersion. The remaining dataset was then used to create the candidate models using R. The final results are shown in Table 8-6.

Table 8-6: HCV Parking Supply Candidate Models

	Negative Binomial						Zero-Inflated Negative Binomial					
Variable	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**		**Model 6**	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**
Intercept	1.49	<0.001***	2.27	<0.001***	0.89	0.014**	2.07	<0.001***	2.76	<0.001***	1.77	<0.001***
Perimeter_i	1.96	<0.001***			2.19	<0.001***	1.56	<0.001***			1.62	<0.001***
Rural_i			1.48	0.051**	−0.35	0.659			2.25	<0.001***	0.87	<0.001***
$\text{Perimeter}_i \times \text{Rural}_i$	−2.63	0.001***					−1.91	<0.001***				
$\text{Area}_{\text{paved},i}$			53.04	<0.001***	82.90	0.005***			38.25	<0.001***	40.32	<0.001***
$\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$					−49.46	0.025**					−20.67	<0.001***
$\text{Class}_{\text{open},i}$	1.02	0.006***	0.57	0.139	0.60	0.111	0.97	<0.001***	0.54	<0.001***	0.67	<0.001***
$\text{Class}_{\text{rest},i}$	−0.12	0.781	−0.44	0.322			0.50	<0.001***	0.10	0.184		
$\text{IND}_{59,i}$	−2.13	0.007***	−2.17	0.006**	−2.04	0.009***	−1.80	<0.001***	−2.17	<0.001***	−1.98	<0.001***
Zero-Inflated Variables												
Intercept							0.14	0.664	0.32	0.923	0.11	0.744
$\text{Area}_{\text{paved},i}$							−163.07	0.003***	−124.72	0.005***	−156.99	0.004***
$\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$							68.40	0.051*	69.31	0.024*	37.76	0.055*
Data												
Removed Outliers	3		1		2		3		1		2	
n	127		129		128		127		129		128	
k	5		5		6		5 + 2		5 + 2		6 + 2	
df	121		123		121		118		120		118	

(***), (**), (*) represent statistical significance to 99%, 95%, or 90% respectively

The candidate models each removed between one and three outliers. When all variables identified in Table 8-5 are considered, the outlier removal process described in Section 4.1.3 identifies a fourth outlier, Location #139. The locations that were determined to be outliers are listed in Table 8-7.

Table 8-7: Outliers Identified in Models

Location	Parking Count	Paved Area (m^2)	Total Area (m^2)	Perimeter (m)	Class	Outlier in Model					
						1	2	3	4	5	6
#83	750 (2nd)	83,600 (4th)	149,400 (8th)	2,900 (3rd)	3 Open Access HCV Parking	Y	N	Y	Y	N	Y
#109	408 (3rd)	12,900 (51st)	164,000 (6th)	1,800 (8th)	4 Limited Access HCV Parking	Y	N	N	Y	N	N
#125	47 (31st)	862,800 (1st)	1,742,700 (1st)	5,600 (1st)	1 Public Rest Areas and Gas Stations	Y	Y	Y	Y	Y	Y
#139	0	340 (143rd)	408,700 (2nd)	2,600 (4th)	4 Limited Access HCV Parking	N	N	N	N	N	N

Y = Yes;   N = No

Examining the geometric variables for these HCV parking locations provide some insight into why these locations were determined to be outliers. Locations #83 and #109 have the second and third most HCV parking spaces, and Location #125 has the largest paved area, total area, and perimeter. These extremes likely contributed to their removal as outliers in the various models. It should be noted that only Location #125 was removed from all six models, which is likely due to it being geographically large with comparatively few HCV parking spaces. Location #139 was not removed in any individual model; however, it is identified as an outlier if all variables used to create the candidate models are considered. This location was a unique outlier as there were zero identified HCV parking spaces despite having the second largest total area and the fourth largest perimeter; very little paved area was identified at this rural location. Geographically these locations correspond with rural locations in the Town of Caledon or employment areas in the City of Mississauga.

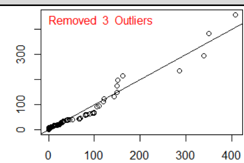
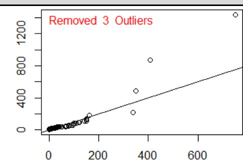
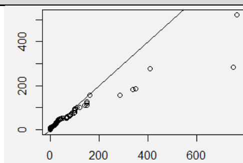
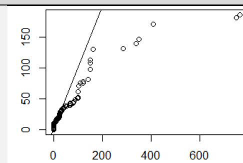
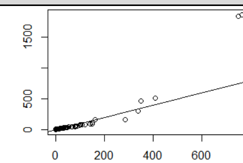
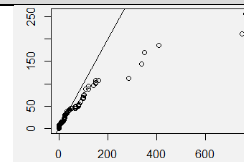
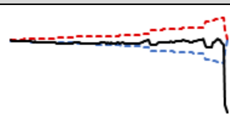
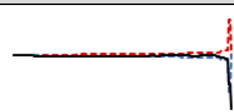
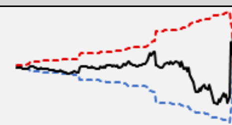
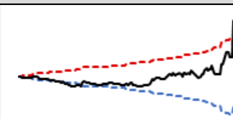
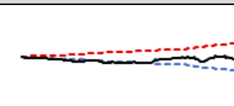
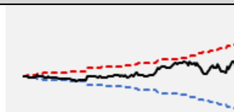
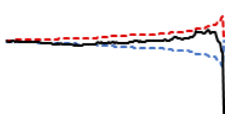
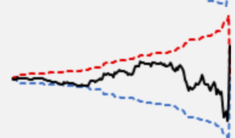
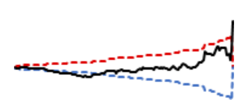
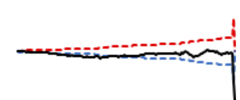
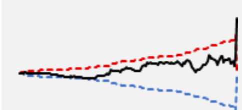
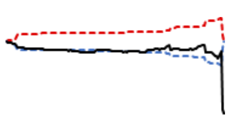
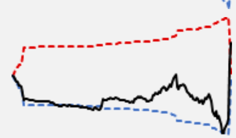
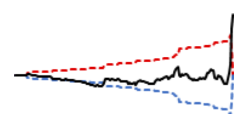
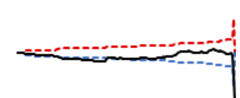
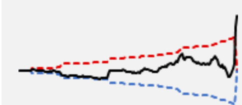
Although outliers were not included in the model development, they are important to examine when assessing model goodness-of-fit. Outliers are not screened from the holdout dataset in order to assess the overall goodness-of-fit of the model. Additionally, outliers are included in the graphical goodness-of-fit plots shown in Table 8-8.

8.4.2 Goodness-of-Fit Comparison

Various goodness-of-fit techniques, as described in Section 4.3, were used to determine the most appropriate HCV parking supply model. To do this, the negative binomial regression models were first compared separately from the zero-inflated negative binomial regression models, and the best model from each was compared to identify the best overall model. The various goodness-of-fit techniques, both numerical and graphical, are shown in Table 8-8. The numerical goodness-of-fit tests, as described in Section 4.3.1, were performed using R. The graphical goodness-of-fit tests, as described in Section 4.3.2, were also assessed: Q-Q plots were generated using R and CURE plots were generated using Excel.

The graphical goodness-of-fit plots can also be found in Appendix F: Q-Q Plots and Appendix G: CURE Plots. The Q-Q Plots are based on all observations ($n = 150$), both the model record set and holdout dataset. An exception to this is Model 1 and Model 2, which omit three extreme outliers to ensure the figure is interpretable. The CURE plots are based on all observations except the four outliers identified in Table 8-7. The square roots of $\text{Area}_{\text{total},i}$ and $\text{Area}_{\text{paved},i}$ were used as CURE plot exposure variables since the area variables grow at a faster rate than the distance variable, Perimeter_i. This also allowed for better visualization of the CURE plot at smaller values of the exposure variable. The $\text{Area}_{\text{total},i}$ variable was used as an exposure variable despite not being included in the model as it provides a good indication of how the accuracy of the model changes as HCV parking locations get larger.

Table 8-8: HCV Parking Supply Candidate Models Goodness-of-Fit

	Negative Binomial			Zero-Inflated Negative Binomial		
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Numerical Goodness of Fit Metrics						
$2LL(\theta)$	**–942**	–982	–957	**–1254**	–1497	–1307
MSE	3314	7961	**2480**	**1468**	2035	1528
$MSPE$	194	721	**94**	75	148	**69**
MPB	6.05	9.38	**2.63**	**0.26**	–1.95	–1.65
MPB_{holdout}	158	703	**–30**	–65	127	**–58**
MAD	28	34	**27**	**22**	26	**22**
MAD_{holdout}	194	721	**94**	75	148	**69**
AIC	**956**	996	973	**2526**	3012	2634
AIC_c	**956**	996	974	**2527**	3012	2635
BIC	**976**	1016	996	-	-	-
Q-Q Plots						
x: Actual y: Predicted						
CURE Plots						
$\sqrt{\text{Area}_{\text{total},i}}$						
Perimeter_i						
$\sqrt{\text{Area}_{\text{paved},i}}$						

Grey highlighted cell = optimal value (separate cells are highlighted each for the two model types)

As described in Section 4.3.2, CURE plots examine the cumulative residuals of the dependent variable ordered and plotted against an exposure variable. Sharp movements either up or down on the plot may only indicate that a particular outlier has not been modeled properly. Additionally, a less accurate fit at the extreme values of the exposure variable may indicate that the model is being extrapolated to observations beyond that which it was calibrated using. The exposure variables selected for this thesis are all geometric variables; therefore, only the large values of these exposure variables would be considered extreme values. The exception to this could be the square root of $\text{Area}_{\text{paved},i}$, where parcels some parcels had very little or even no paved area. The presence of some parcels without paved area could explain some of the sharp movements near the origin of these plots. Additionally, Table 8-7 shows that observations with comparatively large values of these geometric variables tend to be identified as outliers using appropriate statistical tests.

The three negative binomial regression models (Models 1, 2, and 3) were examined to determine which had the best overall fit, as shown in Table 8-8. The goodness-of-fit tests based on $2LL(\theta)$, AIC, AIC_c, and BIC, indicated that Model 1 has the best fit. The other goodness-of-fit tests, including those based on the holdout dataset and the graphical methods, indicated that Model 3 is the best fit. The Q-Q plot for Model 3 shows less deviation from the line compared too the other two models, which each had three predicted values that were too large to fit on the Q-Q plot. The CURE plots for Model 3 also show a better trend, keep within the confidence limits better, and tended closer to zero at the largest value of the exposure variables. The CURE plots show that most observations were predicted well with some exceptions at extreme values of the exposure variables. Based on these results, Model 3 was determined to be the most appropriate negative binomial regression model.

The three zero-inflated negative binomial regression models (Models 4, 5, and 6) were examined to determine which had the best overall fit, as shown in Table 8-8. Numerical goodness-of-fit tests based on the model data, MSE, MPB, and those based on $2LL(\theta)$, the AIC and AIC_c values, indicated that

Model 4 has the best fit. The numerical goodness-of-fit tests based on the holdout dataset, $MSPE$, MPB_{holdout}, and MAD_{holdout}, indicated that Model 6 is the best fit. The MAD value was the same for Model 4 and 6. The Q-Q plot for Model 6 performed better than those for the other two models. The CURE plots for Model 6 also show a better trend, keep within the confidence limits better, and tended closer to zero at the largest value of the exposure variables.

The accuracy of the zero-prediction of the zero-inflated negative binomial models was compared to identify the model that best identifies zero value observations. Table 8-9 lists the correct identification of zeros, the false positive identifications of zeros (type I error), and the false negative predictions of zeros where a zero is not identified in the model (type II error). Type I errors occur when the binomial choice model predicts a zero value when the number of HCV parking spaces for an observation is more than zero. Type II errors occur when the binomial choice model fails to predict a zero value when the number of HCV parking spaces for an observation is actually zero. It should be noted that a zero-inflated negative binomial model identifies two types of zeros: the count of zero and the inability to have a count. The type II error provided in Table 8-9 does not account for this, as the model predicted very low (close to zero) HCV parking supply for some of these observations. It can be noted that Model 6 performed the best in terms of zero prediction.

Table 8-9: Zero-Prediction Accuracy of Zero-Inflated Models

Model	Zero Prediction		
	Correct	Type I Error	Type II Error
4	8	6	37
5	6	2	39
6	8	4	37

Based on the goodness-of-fit results and the zero-prediction accuracy, Model 6 was determined to be the most appropriate zero-inflated negative binomial regression model.

Model 3 and Model 6 are both based on the same variable combinations. Model 3 provides a simpler approach, has much better values for the tests based on $2LL(\theta)$, the AIC and AIC_c values, and comparable values for all other numeric tests. The Q-Q and CURE plots for these two models are very similar, meaning that the models are predicting the number of HCV parking spaces quite similarly. Model 6 has the added benefit that it identifies several of the locations to have zero HCV parking spaces; however, there is no definitive evidence that the HCV parking supply count is truly zero-inflated with two types of zeros. Based on this information, Model 3 was selected as the final HCV parking supply model for this thesis.

8.5 Selected HCV Parking Supply Model

The final selected model for this thesis is Model 3. This model is based on a negative binomial regression model and included the following independent variables: Perimeter_i, Rural_i, $\text{Area}_{\text{paved},i}$, $\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$, $\text{Class}_{\text{open},i}$, and $\text{IND}_{59,i}$. The coefficients for this model are shown in Table 8-6.

The *post hoc* power check was performed for this final selected model using the `negbin.samp()` function, an average HCV parking count ($\bar{y}$) of 36.68, and an overdispersion parameter ($1/\theta$) of 2.474. An $\alpha = 0.05$ and a desired power of $(1 - \beta) = 0.8$ yielded a minimum sample size of 47 for a medium effect size ($d = 0.5$) and a minimum sample size of 121 for a medium-large sample size ($d = 0.65$). The seed value and number of simulations were set to their default values, and the significant digits was set as three.

8.5.1 Examination of Result

The selected HCV parking supply model can be inspected to determine the explanatory nature of the model. All of the variables included in this model are statistically significant to 95% except for the Rural_i and $\text{Class}_{\text{open},i}$ variables.

This model has two main geometric variables: Perimeter_i ($\beta = 2.19$; $p < 0.001$) and $\text{Area}_{\text{paved},i}$. ($\beta = 82.90$; $p = 0.005$). Both of these variables have a positive coefficient and a statistically significant p-value; therefore, HCV parking locations that are larger and that have more paved area are more likely to have a larger number of HCV parking spaces. The interaction between these variables, $\text{Perimeter}_i \times \text{Area}_{\text{paved},i}$ ($\beta =-49.46$; $p = 0.025$) indicates that an increase in both of these variables does not provide the same increase in HCV parking spaces as an increase in only one of these variables. This interaction term results in a different effective parameter for each of these variables depending on the value of the other. As the value of Perimeter_i increases, the effect that $\text{Area}_{\text{paved},i}$ has on the number of HCV parking spaces decreases. As the value of Perimeter_i decreases, the effect that $\text{Area}_{\text{paved},i}$ has on the number of HCV parking spaces increases. This means that for narrow parcels (larger Perimeter_i) the effect of $\text{Area}_{\text{paved},i}$ on HCV parking space count is decreased, meaning that larger amounts of paved area do not add as many HCV parking spaces in narrow parcels as in wide parcels. This could be due to the narrow width of the parcel minimizing the ability of HCVs to maneuver around the parcel and therefore reducing the efficiency of HCV parking space placement.

The $\text{Class}_{\text{open},i}$ ($p = 0.139$) variable is not statistically significant, meaning that there is no statistically significant effect on HCV parking supply in this model if a location is classified as ③ Open Access HCV Parking. The $\text{IND}_{59,i}$ ($\beta =-2.01$; $p = 0.009$) variable is statistically significant with a negative coefficient, meaning that parcels that contain a Miscellaneous Retail (SIC59) firm have less

HCV parking supply. This relationship is expected given the larger priority that retail establishments will generally give to passenger vehicles.

Rural_i ($p = 0.659$) is not statistically significant in the final selected model; however, its parameter is positive and significant in other models indicating that rural locations have a larger number of spaces than urban locations. This could be due to the value of land in urban areas being higher and therefore less abundant or available for parking spaces. The number of HCV parking spaces in the model only includes dedicated parking spaces while rural properties tend to have open areas. This reduces the number of visible parking spaces available at these locations even though there is open space that could be used as temporary parking. This means that Rural_i would likely be more significant and more positive if unmarked parking spaces could be identified in rural areas. It is worth noting that in Model 1 the interaction variable $\text{Perimeter}_i \times \text{Rural}_i$ ($\beta_{\text{Model1}} = -49.46$; $p = 0.001$) has a statistically significant negative coefficient, meaning that the number of HCV parking spaces increases slower in rural locations with an increase in the size of the HCV parking location when compared to urban locations.

Although not part of the selected model, in Preliminary Model 2 the $\text{Class}_{\text{rest},i}$ ($\beta_{\text{PrelimModel2}} = -1.77$; $p = 0.08$) variable has a somewhat statistically significant coefficient. This may indicate that locations classified as ① Public Rest Areas and Gas Stations have less HCV parking spaces than other locations; however, this value is not significant in other models. This finding does correspond with evidence found during the validation process that many gas stations in the Region of Peel do not provide a large amount of dedicated truck parking.

The predictability of Model 3, the final selected model, can be judged by examining the numerical goodness-of-fit metrics. The MAD and MAD_{holdout} values are 27 and 94, meaning that the average absolute prediction error is 27 and 94 HCV parking spaces for the model record set ($n = 130$) and holdout dataset ($n = 20$), respectively. The MPB and MPB_{holdout} values are 2.63 and -30, meaning that the total prediction error in these models is tending to zero (but more strongly in the model record set). Lastly, the square root of the MSE and $MSPE$ are approximately 50 and 10, again showing that the magnitude of the errors is still significant when additional weight is given to large deviations and less weight is given to smaller deviations.

To check the suitability of aggregating the results of Model 3 to the TAZ level, the errors of all observations ($n = 150$) are checked for spatial autocorrelation. This calculation was performed using GeoDa (Anselin, 2017) with the spatial weight matrix being manually set to one for observations in the same TAZ and zero for observations in different TAZ. The global Moran's I was found to not be statistically significant with a pseudo p-value of 0.170 after 999 permutations. Several points on a Moran scatter plot exhibited positive spatial autocorrelation; however, this was likely caused by outlier observations.

The insignificant global Moran's I suggests that the model error within in each TAZ is randomly distributed; therefore, aggregating to the zonal level is expected to decrease the total zonal error by balancing individual negative and positive errors. After aggregation to the TAZ level, the model can be considered as a suitably predictive estimate of HCV parking supply for high-level planning and policy purposes.

8.5.2 Region of Peel HCV Parking Supply

To determine the HCV parking supply of the Region of Peel, the selected HCV parking model is applied to all the long-haul HCV parking locations identified in the region. After applying the model, the HCV parking supply from all locations are aggregated to the TAZ level using the R-ArcGIS bridge. This approximation of the HCV parking supply for each TAZ can be used for policy planning at the Regional level. Due to the large variance of the number of HCV parking spaces, the HCV parking supply estimates are not intended to be used at the level of individual property parcels. If the number of HCV spaces is required for one specific location, it is recommended that a manual count be undertaken.

The number of HCV parking spaces in the Region of Peel is shown in Table 8-10. There were no HCV parking spaces identified in City of Mississauga TAZ 10, City of Brampton TAZ 4, or Town of Caledon TAZ 3. It should be noted that no authorized non-roadside HCV parking locations were identified in these locations; therefore, the values for these TAZ were expected to be zero regardless of the model results.

Table 8-10: Total HCV Parking Spaces per TAZ in the Region of Peel

City of Mississauga		**City of Brampton**		**Town of Caledon**	
TAZ	**Estimated Supply**	**TAZ**	**Estimated Supply**	**TAZ**	**Estimated Supply**
1	736	1	16	1	523
2	1,242	2	792	2	216
3	187	3	3,931	3	0
4	165	4	0	4	7,423
5	19,588	5	195	5	886
6	259	6	939		
7	90	7	973		
8	136	8	10,067		
9	449	9	37		
10	0	10	645		
11	1,206				
Total	**24,058**	**Total**	**17,594**	**Total**	**9,049**
Region of Peel Total: 50,701					

The number of HCV parking spaces is shown graphically in Figure 8-4. It should be noted that large TAZ, such as those in the Town of Caledon, may suffer from the modifiable area unit problem (MAUP) which occurs when point data is aggregated into zones (Wong, 2009). MAUP refers to the inconsistency of the results depending on the size and position of the zoning selected. Figure 8-4 gives the appearance that the HCV parking spaces are distributed throughout the large zones; however, they are primarily clustered in relatively small industrial areas. Although it is not possible to perfectly overcome the MAUP, this observation could show that the municipal wards are not the most appropriate TAZ for the Town of Caledon.

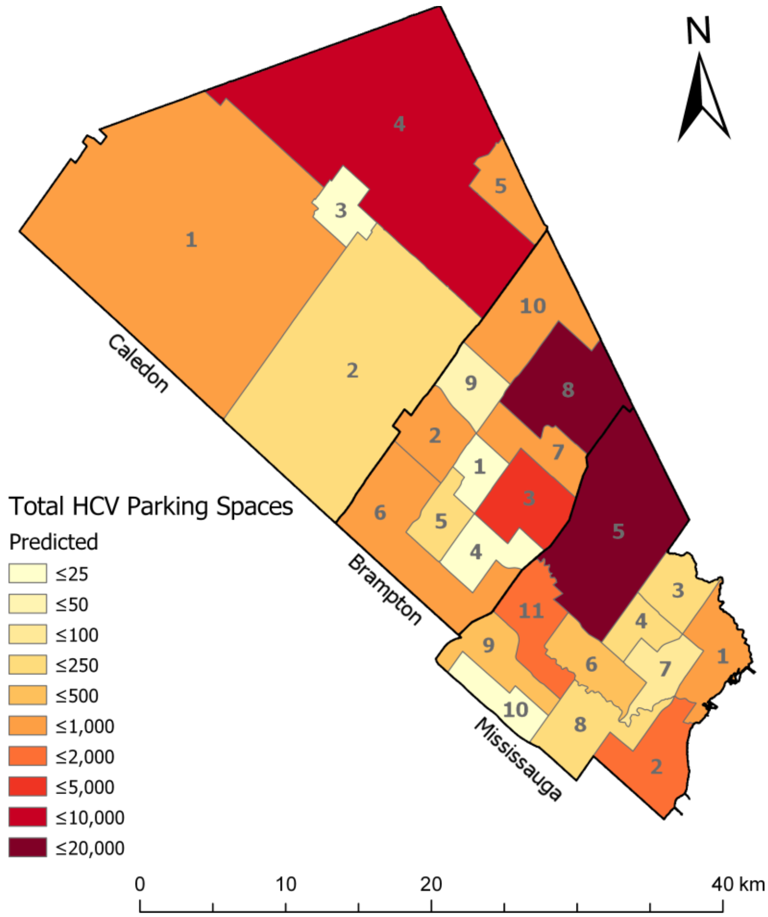


Figure 8-4: Total HCV Parking Spaces per TAZ in the Region of Peel
Note: Labels represent the TAZ number for each municipal ward in the Region of Peel

As expected, the majority of the estimated HCV parking supply is near the Pearson International Airport. There is also a significant HCV parking spaces around other industrial / commercial areas in the region, including around the commercial/industrial area of the Town of Bolton, as identified in the Region of Peel Generalized Land Use classification scheme (Region of Peel, 2014).

It should be noted that this estimate of HCV parking supply is limited to non-roadside HCV parking related to rest stops used to comply with HOS laws on long-haul trips. Additionally, this estimated supply cannot be compared at a one-to-one basis with an estimate of HCV parking demand. Parking utilization factors would need to be derived to account for HCV parking spaces that are inaccessible due to factors like improper parking, usage by non-HCV vehicles, and spaces being used for alternate purposes such as storage.

The estimated HCV parking supply in the Region of Peel was assessed for spatial autocorrelation. It was determined that there was no statistically significant spatial autocorrelation related to HCV parking supply with the exception of Brampton TAZ 7, which indicated a small amount of statistically significant spatial autocorrelation. This means that the distribution of HCV parking spaces is essentially random when aggregated up to the TAZ level. This is not unexpected, as the TAZ selected for this thesis are large in size and encompass fairly uniform land uses. Commercial and residential TAZ are distributed throughout the map without any apparent order. This also means that the number of HCV parking spaces in one TAZ are not significantly affected by the neighbouring TAZ. More details can be found in Appendix H: HCV Parking Supply Spatial Autocorrelation.

CHAPTER 9: CONCLUSION

9.1 Summary

The key objective of this research is the development of tools to identify, classify, and estimate long-haul HCV parking supply in order to more accurately assess the location and extent of any HCV parking deficiencies. These tools are of value to municipalities and transportation professionals as there are large economic and safety concerns related to HCV parking and driver fatigue.

An exhaustive literature review was conducted to better understand the issues surrounding HCV parking. The results of this literature review informed the analysis conducted for this thesis. Key HCV parking locations characteristics were found to be 1) legality, 2) accessibility, 3) ownership, 4) dedication to HCV parking, and 5) roadside parking.

The approach to identify HCV parking locations was developed by analyzing GIS pings from HCV vehicles using the R-ArcGIS bridge. Long-haul HCV rest events that correspond to HOS requirements were identified by identifying HCV stop events that lasted between two and fifteen hours in duration and occurred during a long-haul trip. This approach was applied to the Region of Peel to identify 1,688 individual property parcels that corresponded with HCV rest events.

An exhaustive HCV parking location classification scheme was developed using a heuristic stratification of key HCV parking location characteristics. A total of nine categories were established. The five authorized HCV parking location categories are: ① Public Rest Areas and Gas Stations, ② Weigh Stations, ③ Open Access HCV Parking, ④ Limited Access HCV Parking, and ⑤ Authorized Roadside Parking. The four unauthorized HCV parking location categories are: ⑥ Unauthorized Roadside Parking, ⑦ Unauthorized Highway Ramp Parking, ⑧ Unauthorized Parking on Public Property, and ⑨ Unauthorized Parking on Private Property.

An approach to systematically apply the HCV parking location classification scheme was developed for the Region of Peel. An algorithm using variables selected from commonly available geospatial data and implemented using an R-ArcGIS bridge.

A Region of Peel HCV parking supply model was developed using a negative binomial regression with variables identified from commonly available geospatial data in order to understand the factors that give rise to legal HCV parking spaces on individual properties. Appropriate *a priori* data analysis was performed to minimize any bias in the model. Six candidate models were developed: three negative binomial regressions and three zero-inflated negative binomial regressions. Using a combination of numerical and graphical goodness-of-fit metrics, the final model was selected and further examined to better understand the its explanatory and predictive accuracy. The results of the HCV parking supply model are intended to be aggregated up to the TAZ level. The final HCV parking supply model was applied to the whole Region of Peel, which predicted a total of 50,701 HCV parking spaces. Variables that were found to have a positive relationship with the number of HCV parking spaces include the perimeter length of the property and the paved area of the property. The presence of a miscellaneous retail industry firm was found to have a negative relationship with the number of HCV parking spaces.

The findings of this research project and the various tools developed can be used by municipalities and transportation professionals to better determine the location and extent of any HCV parking deficiencies by comparing the estimated HCV parking supply to a value of HCV parking demand.

9.2 Contribution

This research adds to the existing transportation literature by creating a classification scheme that can be applied to locations of long-haul HCV parking supply, along with an approach to estimate the number of HCV parking spaces in a Traffic Analysis Zone (TAZ). Previous parking studies have focused on the most visible locations for parking, such as dedicated rest areas, while giving little regard to additional legal parking. For example, ① Public Rest Areas and Gas Stations and ② Weigh Stations are commonly analyzed categories, but these two categories comprised only 21 (1.4%) of the 1,481 legal parking locations identified in the case study area, as shown in Table 7-2. Other parking locations are available to HCV drivers, even if they are considered less desirable but functionally adequate. This research contends that neglecting all viable HCV parking locations can lead to an underestimate of the available HCV parking supply.

The creation of an HCV parking supply model is an important development as no comparable tools are currently available that accomplish this task. An accurate estimate of HCV parking supply is a necessary element to understand the current parking needs of HCV drivers on long-haul trips. An analysis of current or future potential parking shortages can be examined when comparing an estimate of HCV parking supply with a suitable model of HCV parking demand. This is an important area for transportation policy due to the safety and environmental implications of inadequate HCV parking capacity.

9.3 Future Modelling Work

Future study can examine additional data sources to more-accurately classify HCV parking locations. While the proposed classification scheme provides a practical categorization, further sub-categories or tiers could be created. For example, ① Public Rest Areas and Gas Stations could be separated into two sub-categories; however, this would require additional efforts to properly define them. Additionally, the differentiation between ③ Open Access HCV Parking and ④ Limited Access HCV Parking currently based on the industry classification of the firms within the property parcel. Parcels containing firms with a two-digit standard industry classification (SIC) for Motor Freight Transportation and Warehousing ('42') were classified as open access and all others were classified as limited access. Another option that could be explored to determine the accessibility of a location is the entropy approach as described in Gingerich *et al.* (2016b). Entropy in this analysis is used as a measurement of the variability of carriers at an identified HCV parking location; locations with a greater entropy index indicate more open access. Although this technique may provide a more accurate prediction of accessibility, the data required for this analysis is not commonly available to jurisdictions at this time. Future data availability may make this a more feasible approach for classification.

HCV parking supply model developed as part of this thesis is intended to be aggregated from the property parcel level to the TAZ level due to the accuracy of the model at predicting HCV parking supply at a single property parcel. A future HCV parking supply model can focus on an approach to develop a model that is appropriately accurate to report the HCV parking space estimates at the property parcel level. Computer vision, machine learning, and image processing could be integrated into an HCV parking supply model as the analysis techniques become more advanced and as more data becomes available. Further analysis can also work to increase statistical significance of other variables in the HCV parking supply model.

The HCV parking supply model in this thesis is limited to non-roadside HCV parking related to rest events used to comply with HOS laws on long-haul trips. This estimated supply cannot be compared at a one-to-one basis with an estimate of HCV parking demand, as the supply represents all existing HCV parking spaces regardless if they can be used by an HCV driver. Parking utilization factors would need to be derived to account for HCV parking spaces that are inaccessible due to factors like improper parking, usage by non-HCV vehicles, and spaces being used for alternate purposes such as storage.

Lastly, more advanced approaches to estimating HCV parking demand could be developed using a data-driven systematic approach similar to that taken for HCV parking supply in this thesis. A more accurate estimate of HCV parking demand would provide for a better estimate of the extent and location of any HCV parking deficiencies when compared to HCV parking supply estimated using the tools developed in this thesis.

9.4 Future Planning Considerations

The assessment and management of HCV parking supply effects the end users of the transportation infrastructure in a jurisdiction. Policy areas relating to land use, HOS, and development influence existing and new parking supply. Various technological advancements can also be leveraged to increase the efficiency of the existing parking supply.

It is important that jurisdictions continue to improve access to geospatial data through open data portals. The input variables for this thesis were obtained from open data sources. As data availability increases, the accuracy of similar HCV parking models will continue to improve. This will allow for more efficient selection of new HCV parking locations to ensure that resources are used in a cost and safety-efficient manner. In addition to selecting efficient locations, jurisdictions should reconsider their current zoning restrictions related to HCV parking. Policy changes also have the potential to be very effective in alleviating HCV parking deficiencies, as they could allow for existing infrastructure to be used for HCV parking with minimal to no renovation. Jurisdictions can also examine technological advancements that can reduce fatigue-related collisions instead of relying on HOS laws as the primary defence against fatigued drivers.

One large obstacle in the creation of new dedicated HCV parking is the significant amount of land required, as land is usually scarce and expensive in urban areas (TRI *et al.*, 1996; Giuliano *et al.*, 2017). This land requirement is due in part to the large turning radius and swept path of an HCV when parking. Land zoning policies and urban densification targets often prohibit or discourage the construction of suitable HCV parking lots (Meyer, 1999; Giuliano *et al.*, 2017; Government of Ontario, 2019). Regulations may cap the maximum number of non-roadside parking spaces provided, and may not consider parking requirements related to HCV trips (Jaller *et al.*, 2013). This trend becomes evident in the Region of Peel when examining the distribution of authorized and unauthorized HCV parking locations, as shown in Figure 7-2. Most authorized parking locations are clustered near the Pearson International Airport (City of Mississauga TAZ 5) where industrial land has historically been more cost effective; in contrast, unauthorized parking locations are distributed throughout the region in most TAZ regardless of land costs.

It is important that jurisdictions examine shifting demand and the inter-regional nature of goods movement. Areas that have a deficiency of HCV parking spaces may shift the excess demand to areas that have a surplus of parking supply (Pécheux *et al.*, 2002). An area that identifies a deficiency of HCV parking may find that a deficiency still exists after they add HCV parking supply if neighbouring areas have persisting HCV parking deficiencies. HCV drivers have been known to travel 30 minutes or longer to find suitable parking (Atlanta Regional Commission, 2018), so neighbouring jurisdictions must work together to ensure that adequate HCV parking supply is available.

Large-scale innovation within the goods movement industry is expected to result from the development of new technologies and the adaptation of existing technologies (Cassetta *et al.*, 2017). Connected and autonomous vehicle (CAV) technology will have a substantial impact on HCV parking supply and demand in the future; however, it is difficult to predict the nature and timeframe of these changes (Phelan *et al.*, 2016). Less time at HCV parking facilities would be required if drivers can reduce their level of attention during vehicle operations (Short and Murray, 2016). As levels of automation increase, parking requirements could become similar to that experienced by current day team drivers. Team driving is when two people take turns to virtually eliminate the need to stop for HOS requirements (Goel *et al.*, 2019). Truck Parking Information and Management Systems (TPIMS), which include detection systems, data management, and communications technology, are being used in coordination with Vehicle-to-Infrastructure (V2I) systems to increase the efficiency of existing HCV parking (Perry *et al.*, 2017). These types of technologies should be considered in future studies.

The Region of Peel is like many jurisdictions in that they do not directly address the issues related to freight parking in their main policy documents. The existence of these policies and their contents show that the region takes the safety and economic effects of the industry seriously and is working to implement meaningful change. The tools and findings of this thesis can be used when updating the Region of Peel's strategic plans related to goods movement in order to properly plan for long-haul HCV parking. It is important for the Region of Peel to examine these issues as HCV safety effects the welfare of all residents who interact with the transportation network.

9.5 Concluding Statements

Both the United States and Canada consider ELDs to be an effective way of enforcing HOS and reducing fatigue-related collisions; however, this approach only works if enough HCV parking is available. One way to make HOS and ELD laws less frustrating for drivers is to increase the supply of HCV parking locations. As new parking locations can often be cost or space prohibitive, expanding or adapting existing areas such as weigh stations to accommodate HCV parking can provide additional capacity in areas where it is needed. Furthermore, vacant urban property parcels can be adapted to provide additional HCV parking locations (Perry *et al.*, 2017). Non-vacant urban property parcels, such as shopping malls or big-box stores, where excess parking capacity in daytime hours could potentially be used or adapted to accommodate HCV parking at night when the locations are empty and the demand for HCV parking is high.

In order to implement any of these potential solutions and before identifying areas for potential HCV parking expansion, it is vitally important to first examine if and where any HCV parking deficiencies exist. This exercise requires an accurate understanding of current HCV parking supply and a detailed understanding of the current and future parking demands. Examining the locations where illegal HCV parking occurs can be valuable for identifying areas that may require additional parking supply. The HCV parking location classification scheme and supply model developed in this thesis are tools that can be used to estimate current HCV parking supply, identify the locations where HCVs are illegally parking, and identify where current HCV parking supply deficiencies may exist. The adequacy of HCV parking is both a financial and safety concern for the goods movement industry and for the general public at large.

REFERENCES

Abdi, H., and Molin, P. (2007) 'Lilliefors/Van Soest's Test of Normality', in Salkind, N. J. (ed.) *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: SAGE Publications, Inc. Available at: http://www.utd.edu/~herve/Abdi-Lillie2007-pretty.pdf.

de Abreu e Silva, J., and Alho, A. R. (2017) 'Using Structural Equations Modeling to Explore Perceived Urban Freight Deliveries Parking Issues', *Transportation Research Part A: Policy and Practice*, 102, pp. 18–32. doi: 10.1016/j.tra.2016.08.022.

Adams, B. T., Srivastava, P., Wang, B. X., and Ogard, L. (2009) *Low Cost Strategies For Short Term Parking On Interstate Highways Of The MVFC*. Madison, WI: University of Wisconsin National Center for Freight & Infrastructure Research & Education.

Aguinis, H., Gottfredson, R. K., and Joo, H. (2013) 'Best-Practice Recommendations for Defining, Identifying, and Handling Outliers', *Organizational Research Methods*, 16(2), pp. 270–301. doi: 10.1177/1094428112470848.

Alexiadis, V., Jeannotte, K., and Chandra, A. (2004) 'Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer', I(July), p. 34.

Alin, A. (2010) 'Multicollinearity', *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), pp. 370–374. doi: 10.1002/wics.84.

Amer, A., and Chow, J. Y. J. (2017) 'A Downtown On-Street Parking Model with Urban Truck Delivery Behavior', *Transportation Research Part A: Policy and Practice*. Elsevier Ltd, 102, pp. 51–67. doi: 10.1016/j.tra.2016.08.013.

American Association of State Highway and Transportation Officials (2008) *Highway Safety Manual*.

American Transportation Research Institute (2018) *Critical Issues in the Trucking Industry – 2018*. Arlington, VA. Available at: http://canue.ca/wp-content/uploads/2018/03/Data_Dictionary_CanMap_Content_Suite_v2016_3.pdf.

American Transportation Research Institute (2019) *Critical Issues in the Trucking Industry – 2019*. Arlington, VA. Available at: https://truckingresearch.org/wp-content/uploads/2019/10/ATRI-Top-Industry-Issues-2019-FINAL.pdf.

Anderson, J. C., and Roll, J. (2018) 'Understanding Probable Reasons for Freeway Ramp and Shoulder Parking by Truck Drivers: An Emerging Safety Issue to Oregon Highway Users', in *Transportation Research Board Annual Meeting - Online Archived Content*. 18-05304, pp. 1–9.

Andersson, S. (2015) *Package 'het.test'*. Available at: https://cran.r-project.org/web/packages/het.test/het.test.pdf%0D.

Anselin, L. (2017) 'GeoDa'. Chicago, Il: University of Chicago. Available at: https://geodacenter.github.io/.

Atlanta Regional Commission (2018) *Atlanta Regional Truck Parking Assessment Study*. Atlanta, GA.

Beck, T. W. (2013) 'The Importance of A Priori Sample Size Estimation in Strength and Conditioning Research', *Journal of Strength and Conditioning Research*, 27(8), pp. 2323–2337.

Bimali, A. M., Usset, J., and Fridley, B. L. (2015) *Package 'RSPS'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/RSPS/RSPS.pdf.

Boggs, A., Hezaveh, A. M., and Cherry, C. R. (2019) 'Shortage of Commercial Vehicle Parking and Truck-Related Interstate Ramp Crashes in Tennessee', in *98th Annual Meeting of the Transportation Research Board*. Washington DC.

Boris, C., and Brewster, R. (2018) 'A Comparative Analysis of Truck Parking Travel Diary Data', in *Transportation Research Board Annual Meeting - Online Archived Content*. 18-00538.

Boris, C., and Brewster, R. M. (2016) *Managing Critical Truck Parking Case Study – Real World Insights from Truck Parking Diaries*. Arlington, VA: American Transportation Research Institute.

Boris, C., and Johnson, M. A. (2015) *Managing Critical Truck Parking Tech Memo #1: Commercial Driver Perspectives on Truck Parking*. Arlington, VA: American Transportation Research Institute.

Breusch, T. S., and Pagan, A. R. (1979) 'A Simple Test for Heteroscedasticity and Random Coefficient Variation', *Econometrica*, 47(5), pp. 1287–1294.

Buldeo Rai, H., Verlinde, S., and Macharis, C. (2019) 'The "Next Day, Free Delivery" Myth Unravelled: Possibilities for Sustainable Last Mile Transport in an Omnichannel Environment', *International Journal of Retail and Distribution Management*, 47(1), pp. 39–54. doi: 10.1108/IJRDM-06-2018-0104.

Burks, S. V., Belzer, M., Kwan, Q., Pratt, S., and Shackelford, S. (2010) *Trucking 101: An Industry Primer*. Washington, DC, DC: Transportation Research Board.

Burnham, K. P., and Anderson, D. R. (2002) *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. 2nd edn. New York, NY: Springer. doi: 10.1017/CBO9780511802461.005.

Canadian Council of Motor Transport Administrators (2010) *National Safety Code for Motor Carriers Standard 9: Hours of Service*.

Cassetta, E., Marra, A., Pozzi, C., and Antonelli, P. (2017) 'Emerging Technological Trajectories and New Mobility Solutions. A Large-Scale Investigation on Transport-Related Innovative Start-Ups and Implications for Policy', *Transportation Research Part A: Policy and Practice*. Elsevier, 106(March), pp. 1–11. doi: 10.1016/j.tra.2017.09.009.

Chatterjee, A., and Wegmann, F. J. (2000) 'Overnight Truck Parking Along Tennessee's Interstate Highways and Rest Areas', *Transportation Research Record: Journal of the Transportation Research Board*, 1734, pp. 64–68. doi: 10.3141/1734-10.

Chow, J. Y. J., and Amer, A. (2015) 'A Downtown On-Street Parking Model with Urban Truck Delivery Effects: A Case Study of Toronto's Financial District'.

City of Brampton (2015) *Traffic and Parking By-Law 93-93*.

City of Brampton (2018) *Unauthorized Parking By-Law 104-2018*.

City of Brampton (2019) *Brampton geoHUB*. Available at: http://geohub.brampton.ca/ (Accessed: 1 April 2019).

City of Mississauga (2000) *The Traffic By-Law 555-00*.

City of Mississauga (2019) *Mississauga Open Data*. Available at: http://data.mississauga.ca/ (Accessed: 1 April 2019).

Cohen, J. (1992) 'A Power Primer', *Psychological Bulletin*, 112(1), pp. 155–159. doi: 10.1037/0033-2909.112.1.155.

Conover, W. J. (1972) 'A Kolmogorov Goodness-of-Fit Test for Discontinuous Distributions', *Journal of the American Statistical Association*, 67(339), pp. 591–596. doi: 10.1007/sl0551-008-9945-3.

Corro, K. D., Akter, T., and Hernandez, S. (2019) 'Comparison of Overnight Truck Parking Counts to GPS Derived Counts for Truck Parking Facility Utilization Analysis', in *98th Annual Meeting of the Transportation Research Board*. Washington, DC.

Dallal, G. E., and Wilkinson, L. (1986) 'An Analytic Approximation to the Distribution of Lilliefors's Test Statistic for Normality', *The American Statistician*, 40(4), pp. 294–296.

Data Management Group (2011) *2011 Transportation Tomorrow Survey: Regional Municipality of Peel Summary By Wards*. Toronto, ON: Data Management Group, Department of Civil Engineering, University of Toronto. Available at: http://dmg.utoronto.ca/pdf/tts/2011/travel_summaries_by_ward/peel_wards11.pdf.

David Kriger Consultants Inc., and CPCS (1992) *Regional Transportation Plan Legislative Review Backgrounder: Urban Goods Movement*. Metrolinx, Government of Ontario.

Davis, A. Y., Pijanowski, B. C., Robinson, K. D., and Kidwell, P. B. (2010) 'Estimating Parking Lot Footprints in the Upper Great Lakes Region of the USA', *Landscape and Urban Planning*, 96(2), pp. 68–77. doi: 10.1016/j.landurbplan.2010.02.004.

Dills, T. (2019) 'ELD Mandate Meets Reality', *Overdrive Online*, 31 January. Available at: https://www.overdriveonline.com/eld-mandate-meets-reality/.

Dissanayake, A., and Wijekoon, P. (2016) *Package 'lrmest'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/lrmest/lrmest.pdf.

DMTI Spatial Inc. (2017) *CanMap Content Suite Data Dictionary v2017.3*. Richmond Hill, ON. Available at: https://library.carleton.ca/sites/default/files/find/gis/metadata/UserManual_CanMap_Content_Suite_v2017_3.pdf.

Durdle, K. (2006) 'Understanding the New Federal Hours Of Service Regulations'. Available at: http://www.transportation.alberta.ca/Content/docType276/Production/New_HOS.pdf.

Emurasoft Inc. (2018) 'EmEditor'. Available at: https://www.emeditor.com/.

Esri (2012) *What is GIS?* Redlands, CA. Available at: http://www.esri.com/library/bestpractices/what-is-gis.pdf.

Esri (2018) 'ArcGIS Pro'.

Esri Canada (2018) *Getting Started with the R-ArcGIS Bridge*. Toronto, ON. Available at: https://esricanada-ce.github.io/r-arcgis-tutorials/1-Getting-Started.pdf.

Everitt, B. S., and Skrondal, A. (2010) *The Cambridge Dictionary of Statistics*. 4th edn. New York, NY.

Federal Highway Administration (2015) *Jason's Law Truck Parking Survey Results and Comparative Analysis*. Washington, DC: US Department of Transportation. Available at: http://www.ops.fhwa.dot.gov/freight/infrastructure/truck_parking/jasons_law/truckparkingsurvey/jasons_law.pdf.

Federal Motor Carrier Safety Administration (2000) *Federal Register Volume 65, Number 85 (Tuesday, May 2, 2000)*. Washington, DC, DC: US Department of Transportation. Available at: https://www.govinfo.gov/content/pkg/FR-2000-05-02/html/00-10703.htm.

Federal Motor Carrier Safety Administration (2016) *U.S. Code Title 49: Transportation - Subtitle B - Chapter III - Subchapter B - Part 395: Hours of Service of Drivers*. United States of America: US Department of Transportation. Available at: http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=1&SID=4bf49f997b04dcacdfbd637db9aa5839&ty=HTML&h=L&mc=true&n=pt21.2.101&r=PART#se21.2.101_11.

Federal Motor Carrier Safety Administration (2018) *Large Truck and Bus Crash Facts 2016*. Washington, DC: US Department of Transportation. doi: FMCSA-RRA-16-001.

Ferguson, M. R., Lavery, T. A., and Higgins, C. D. (2014) *Truck Freight Generators & Attractors in the Province of Ontario*. Hamilton, Ontario: McMaster Institute for Transportation and Logistics.

Fischer, M. M., and Wang, J. (2011) *Spatial Data Analysis: Models, Methods and Techniques*. Heidelberg: Springer. doi: 10.1007/978-3-642-21720-3.

Fleger, S. A., Haas, R. P., Trombly, J. W., Cross III, R. H., Noltenius, J. E., Pécheux, K. K., and Chen, K. J. (2002) *Study of Adequacy of Commercial Truck Parking Facilities – Technical Report*. Washington, DC: Science Applications International Corporation, McLean, VA and Federal Highway Administration, US Department of Transportation.

Forsberg, L. (2014) 'Jarque-Bera Test of Normality'. Uppsala, Sweden: Uppsala University.

Gill, V., and Macdonald, A. (2013) *Understanding the Truck Driver Supply and Demand Gap and Its Implications for the Canadian Economy*. The Conference Board of Canada.

Gingerich, K. (2013) *Modelling Non-Residential Real Estate Prices and Land Use Development in Windsor with Potential Impacts from the Windsor-Essex Parkway*. University of Windsor.

Gingerich, K., Maoh, H., and Anderson, W. (2016a) 'Characterization of International Origin–Destination Truck Movements Across Two Major U.S.–Canadian Border Crossings', *Transportation Research Record: Journal of the Transportation Research Board*, 2547, pp. 1–10. doi: 10.3141/2547-01.

Gingerich, K., Maoh, H., and Anderson, W. (2016b) 'Classifying the Purpose of Stopped Truck Events: An Application of Entropy to GPS Data', *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 64, pp. 17–27. doi: 10.1016/j.trc.2016.01.002.

Gingerich, K., and Nevland, E. A. (2019) 'Effectiveness of Student-Led Co-Curricular Software Learning in STEM Education', *Applications of LinkedIn Learning in Ontario's Post-Secondary Institutions*. Available at: http://ite.club.yorku.ca/files/2019/08/Gingerich-Nevland-2019-Effectiveness-of-Student-Led-Co-Curricular-Software-Learning-in-STEM-Education.pdf.

Giuliano, G., Showalter, C., Yuan, Q., and Zhang, R. (2017) *Managing the Impacts of Freight in California*. Los Angeles, CA: National Center for Sustainable Transportation, University of Southern California.

Goel, A., Vidal, T., and Kok, A. L. (2019) *To Team Up or Not – Single Versus Team Driving in European Road Freight Transport*. Rio de Janeiro, Brazil: Pontifical Catholic University of Rio de Janeiro. Available at: https://w1.cirrelt.ca/~vidalt/papers/Team-vs-Single.pdf.

Google (2019) 'Google Maps'. Available at: https://www.google.ca/maps/.

Government of Canada (2009) *Commercial Vehicle Drivers Hours of Service Regulation (SOR/2005-313)*. Canada: Minister of Justice.

Government of Canada (2017) 'Regulations Amending the Commercial Vehicle Drivers Hours of Service Regulations (Electronic Logging Devices and Other Amendments)', *Canada Gazette*. Ottawa, 151(50). Available at: http://gazette.gc.ca/rp-pr/p1/2017/2017-12-16/html/reg5-eng.html.

Government of Canada (2019) *Open Government Potal*. Available at: https://open.canada.ca/en/open-data.

Government of Ontario (2019) *A Place to Grow: Growth Plan for the Greater Golden Horseshoe*. Toronto, ON.

Great Western Transportation (2019) *The History of Semi Trailer Trucks*. Available at: https://www.gwtrans.com/the-history-of-semi-trailer-trucks/ (Accessed: 30 August 2019).

Gross, J., and Ligges, U. (2015) *Package 'nortest'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/nortest/nortest.pdf.

Hallenbeck, M. E., Selezneva, O. I., and Quinley, R. (2014) *Verification, Refinement, and Applicability of Long-Term Pavement Performance Vehicle Classification Rules*. Elkridge, MD.

Hamner, B., Frasco, M., and LeDell, E. (2018) *Package 'Metrics'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/Metrics/Metrics.pdf.

Haque, K., Mishra, S., Paleti, R., Golias, M. M., Sarker, A. A., and Pujat, K. (2017) 'Truck Parking Utilization Analysis Using GPS Data', *Journal of Transportation Engineering*, 143(9). doi: 10.1061/JTEPBS.0000073.

Harwood, D. W., Torbic, D. J., Richard, K. R., Glauz, W. D., and Elefteriadou, L. (2003) *National Cooperative Highway Research Program Report 505: Review of Truck Characteristics as Factors in Roadway Design*. Washington, DC: Transportation Research Board. Available at: http://onlinemanuals.txdot.gov/txdotmanuals/tri/tri.pdf.

Hauer, E. (2015) *The Art of Regression Modeling in Road Safety*. New York, NY: Springer. doi: 10.1007/978-3-319-12529-9.

Historica Canada (2014) *Trucking Industry*, *The Canadian Encyclopedia*. Available at: https://www.thecanadianencyclopedia.ca/en/article/trucking-industry# (Accessed: 1 April 2014).

Hooper, K. G. (2019) *Parking Generation Manual*. 5th edn. Washington, DC: Institute of Transportation Engineers.

Horseless Age Company (1912) 'General Motors Truck Company Advertisement', *The Horseless Age*. New York, NY, July 10.

Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., and Mitchell, D. (2018) *Package 'lmtest'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/lmtest/lmtest.pdf.

Hunt, J. D., and Stefan, K. J. (2007) 'Tour-Based Microsimulation of Urban Commercial Movements', *Transportation Research Part B: Methodological*, 41(9), pp. 981–1013. doi: 10.1016/j.trb.2007.04.009.

IBI Group (2016) *Travel Demand Forecasting Model Update: 2011 AM and PM Peak Period Models (Final Report)*. Brampton, ON: Region of Peel.

INRO (2017) 'Emme'.

Ioannou, P., and de Almeida Araujo Vital, F. (2018) *Intelligent Parking Assist for Trucks with Prediction*. Davis, CA: National Center for Sustainable Transportation Research, University of California.

Ison, S., and Rye, T. (2008) 'Introduction: TDM Measures and their Implementation', in *The Implementation and Effectiveness of Transport Demand Management Measures: An International Perspective*.

ITE York University Student Chapter (2019) *ITE York University Student Chapter: York University's Community of Transportation Professionals*. Available at: http://ite.club.yorku.ca/ (Accessed: 22 March 2018).

Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., and Meers, Z. (2017) *Package 'pscl'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/pscl/pscl.pdf.

Jaller, M., Holguín-Veras, J., and Hodge, S. (2013) 'Parking in the City: Challenges for Freight Traffic', *Transportation Research Record: Journal of the Transportation Research Board*, 2379, pp. 46–56.

Kabak, M. O., and Turgut, O. (2010) 'Parking Spot Detection from Aerial Images', *Seven*, (5229284), pp. 1–3.

Kansas Department of Transportation, and Kansas Turnpike Authority (2016) *Kansas Statewide Freight Network Truck Parking Plan*. Topeka, KA.

Kimley Horn (2015) *Virginia Truck Parking Study*. Richmond, VA: Virginia Department of Transportation. Available at: http://www.virginiadot.org/projects/resources/VirginiaTruckParkingStudy_FinalReport_July2015.pdf.

Koenker, R. (1981) 'A Note on Studentizing a Test for Heteroscedasticity', *Journal of Econometrics*, 17(1), pp. 107–112. doi: 10.1016/0304-4076(81)90062-2.

Lamb, E. (2019) 'Canada Announces Final ELD Mandate to Take Effect in 2021', *Transport Topics*, 13 June.

Lehner, S., and Peer, S. (2019) 'The Price Elasticity of Parking: A Meta-Analysis', *Transportation Research Part A: Policy and Practice*, 121(January), pp. 177–191. doi: 10.1016/j.tra.2019.01.014.

Lopes, S. B., Brondino, N. C. M., and Rodrigues da Silva, A. N. (2007) 'Exploratory and Confirmatory Spatial Data Analysis Tools in Transport Demand Modeling', in *10th International Conference on Computers in Urban Planning and Urban Management*. Iguassu Falls, Brazil.

Lord, D., and Park, P. Y. (2008) 'Investigating the Effects of the Fixed and Varying Dispersion Parameters of Poisson-Gamma Models on Empirical Bayes Estimates', *Accident Analysis and Prevention*, 40(4), pp. 1441–1457. doi: 10.1016/j.aap.2008.03.014.

Maesschalck, R. De, Jouan-Rimbaud, D., and Massart, D. L. (2000) 'The Mahalanobis Distance', *Chemometrics and Intelligent Laboratory Systems*, 50(1), pp. 1–18.

Malik, L., Sánchez-Díaz, I., Tiwari, G., and Woxenius, J. (2017) 'Urban Freight-Parking Practices: The Cases of Gothenburg (Sweden) and Delhi (India)', *Research in Transportation Business and Management*. Elsevier, 24(May), pp. 37–48. doi: 10.1016/j.rtbm.2017.05.002.

Maoh, H., and Kanaroglou, P. (2013) 'Modelling Firm Failure: Towards Building a Firmographic Microsimulation Model', in Pagliara, F., de Bok, M., Simmonds, D., and Wilson, A. (eds) *Employment Location in Cities and Regions*. Berlin, Germany: Springer, pp. 243–261. doi: 10.1007/978-3-642-31779-8.

Martin, E. W., and Shaheen, S. A. (2013) 'Truck Parking and Traffic on I-5 in California: Analysis of a Clipboard Survey and Annual Average Daily Traffic Data', in *Transportation Research Board Annual Meeting - Online Archived Content*, pp. 1–15.

McKeown, D. M., Harvey, W. A., and McDErmott, J. (1985) 'Rule-Based Interpretation of Aerial Imagery', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(5), pp. 570–585. doi: 10.1109/TPAMI.1985.4767704.

Mendeley Ltd. (2019) 'Mendeley Desktop'. Amsterdam, Netherlands: Elsevier.

Meyer, M. D. (1999) 'Demand Management as an Element of Transportation Policy: Using Carrots and Sticks to Influence Travel Behavior', 33, pp. 575–599.

Miaou, S. P., and Lord, D. (2003) 'Modeling Traffic Crash-Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes Versus Empirical Bayes Methods', *Transportation Research Record*, (1840), pp. 31–40. doi: 10.3141/1840-04.

Ministry of Transportation Ontario (2009) *Commercial Vehicle Operators' Safety Manual*. Available at: http://www.mto.gov.on.ca/english/trucks/commercial-vehicle-operators-safety-manual/module-9.shtml.

Ministry of Transportation Ontario (2019) *Long Combination Vehicle (LCV) Program Conditions*. Available at: http://www.mto.gov.on.ca/english/trucks/long-combination-vehicles-program-conditions.shtml (Accessed: 23 September 2019).

Mitman, M. F., Davis, S., Armet, I. B., and Knopf, E. (2018) *Curbside Management Practitioners Guide*. Washington, DC: Institute of Transportation Engineers. Available at: https://www.ite.org/pub/?id=C75A6B8B-E210-5EB3-F4A6-A2FDDA8AE4AA.

Montgomery, D. C., and Runger, G. C. (2011) *Applied Statistics and Probability for Engineers*. 5th edn. New York, NY: John Wiley & Sons Inc.

Morris, T., Murray, D., Fender, K., Weber, A., Morellas, V., Cook, D., and Panaikolopoulos, N. (2017) *A Comprehensive System for Assessing Truck Parking Availability*. Arlington, VA: American Transportation Research Institute. Available at: http://www.cts.umn.edu/publications/researchreports.

National Academies of Sciences, Engineering, and Medicine (2016) *Commercial Motor Vehicle Driver Fatigue, Long-Term Health, and Highway Safety: Research Needs*. Washington, DC: The National Academies Press. doi: 10.17226/21921.

National Academies of Sciences, Engineering, and Medicine, Woodrooffe, J., Sweatman, P., Middleton, D., James, R., and Billing, J. R. (2010) *Review of Canadian Experience with the Regulation of Large Commercial Motor Vehicles*. Washington, DC: The National Academies Press. doi: 10.17226/14458.

National Transportation Safety Board (1990) *Fatigue, Alcohol, Other Drugs, and Medical Factors in Fatal-to-the-Driver Heavy Truck Crashes, Safety Study*. Washington, DC: NTSB/SS-90/01.

Nevland, E. A., Gingerich, K., and Park, P. Y. (2019a) 'Development of a Classification Scheme and a Supply Model for Inter-Regional Truck Parking Facilities', in *98th Annual Meeting of the Transportation Research Board*. Washington, DC.

Nevland, E. A., Gingerich, K., and Park, P. Y. (2019b) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Canadian Institute of Transportation Engineers Annual Conference*. Ottawa, ON.

Nevland, E. A., Gingerich, K., and Park, P. Y. (2019c) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Joint Transportation Association of Canada / Intelligent Transportation Systems Canada Conference and Exhibition*. Halifax, NS.

Nevland, E. A., Park, P. Y., Hanson, R., Klashinsky, R., and Der, T. (2015) 'Utilizing Weigh-In-Motion for Integrated Average Speed and Weight Enforcement', in *2015 Transportation Association of Canada Conference & Exhibition*. Charlottetown, PE.

Nourinejad, M., and Roorda, M. J. (2017) 'Parking Enforcement Policies for Commercial Vehicles', *Transportation Research Part A: Policy and Practice*. Elsevier Ltd, 102, pp. 33–50. doi: 10.1016/j.tra.2016.04.007.

Nourinejad, M., Wenneman, A., Habib, K. N., and Roorda, M. J. (2014) 'Truck Parking in Urban Areas: Application of Choice Modelling within Traffic Microsimulation', *Transportation Research Part A: Policy and Practice*, 64, pp. 54–64. doi: 10.1016/j.tra.2014.03.006.

Occupational Safety and Health Administration (2019) *SIC Division Structure*. Washington, DC: United States Department of Labor. Available at: https://www.osha.gov/pls/imis/sic_manual.html.

Ogden, K. W. (1991) 'Truck Movement and Access in Urban Areas', *Journal of Transportation Engineering*, 117(1), pp. 71–90.

Oluwajana, S. D. (2018) *Development of Hotzone Identification Models for Simultaneous Crime and Collision Reduction*. York University.

Ontario Council of University Libraries (2019) *Scholars GeoPortal*. Available at: http://geo1.scholarsportal.info.ezproxy.library.yorku.ca/.

Ontario Ministry of Natural Resources (2010) 'Ontario Parcel'. Available at: http://thelakelands.ca/wp-content/uploads/2014/07/OntarioParcel.pdf.

Ortuzar, J. de D., and Willumsen, L. G. (2011) *Modelling Transport*. 4th edn. Chichester, West Sussex, England: John Wiley & Sons, Ltd.

Oxford University Press (2019a) 'freight, n.', *Oxford English Dictionary Online*. Available at: https://www-oed-com.ezproxy.library.yorku.ca/view/Entry/74453 (Accessed: 21 September 2019).

Oxford University Press (2019b) ‘parking, n.’, *Oxford English Dictionary Online*. Available at: https://www-oed-com.ezproxy.library.yorku.ca/view/Entry/137961 (Accessed: 21 September 2019).

Paradis, E. (2005) *R for Beginners*. France: Institut des Sciences de l'´Evolution. doi: 10.1016/S0140-6736(09)61155-7.

Parsons (2019) *Long Combination Vehicle (LCV) Usage Study, Region of Peel*. Brampton, ON.

Pears, R., and Shields, G. (2016) *Cite Them Right: The Essential Referencing Guide*. 10th edn. London, UK: Palgrave.

Pécheux, K. K., Chen, K. J., Farbry, J., and Fleger, S. A. (2002) ‘Model Development For National Assessment of Commercial Vehicle Parking’, p. 46.

Perry, E. B., Oberhart, E., McFerren, R., Marion, S., Adams, T., and Zietlow, B. (2017) *An Evaluation of Vacant Urban Land for Truck Parking*. Madison, WI: Mid-America Freight Coalition. Available at: http://midamericafreight.org/wp-content/uploads/2017/03/MAFC-White-Paper_Truck-Parking-1.pdf.

Pettersson, F., Winslott Hiselius, L., and Koglin, T. (2018) ‘E-commerce and Urban Planning – Comparing Knowledge Claims in Research and Planning Practice’, *Urban, Planning and Transport Research*. Routledge, 6(1), pp. 1–21. doi: 10.1080/21650020.2018.1428114.

Phaneuf, I. (2008) ‘Where To Park’, *Truck News*, November.

Phelan, T., Mantero, V., Purdy, J., and Kearney, T. (2016) *National Coalition on Truck Parking Activity Report*. Washington, DC: Federal Highway Administration, US Department of Transportation.

Pinjari, A. R., Zanjani, A. B., Thakur, A., Irmania, A. N., Kamali, M., Short, J., Pierce, D., and Park, L. (2014) *Using Truck Fleet Data in Combination with Other Data Sources for Freight Modeling and Planning*. Tampa, FL: University of South Florida.

PTV AG (2018) *PTV VISSIM 6 User Manual*. Karlsruhe, Germany.

R Core Team (2013) ‘R: A Language and Environment for Statistical Computing’. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.r-project.org/.

Region of Peel (2012a) *Peel Region Long Range Transportation Plan*. Brampton, ON. Available at: https://www.peelregion.ca/planning/residents/transportation/long-range.htm.

Region of Peel (2012b) *Peel Region Moving into Action: Goods Movement Strategic Plan 2012 to 2016*. Brampton, ON. Available at: https://www.tc.gc.ca/eng/ctareview2014/pdf/Peel - Goods Movement StrategicPlan-2012-2016.pdf.

Region of Peel (2013a) *Region of Peel's Road Characterization Study*. Brampton, ON.

Region of Peel (2013b) *Strategic Goods Movement Network Study Technical Report*. Brampton, ON. Available at: https://www.peelregion.ca/pw/transportation/goodsmovement/pdf/peel-final-technical-report.pdf.

Region of Peel (2014) ‘Generalized Land Use 2014 Map’. Brampton, ON. Available at: http://opendata.peelregion.ca/media/33414/generalizedlanduse2014_pdf.pdf.

Region of Peel (2015) *Region of Peel Goods Movement Economic Impact Analysis Study*. Brampton, ON.

Region of Peel (2016) *Official Plan*. Brampton, ON.

Region of Peel (2017) *Peel Region Goods Movement Strategic Plan 2017-2021*. Brampton, ON. doi: 10.1002/ejoc.201200111.

Region of Peel (2019a) *Goods Movement Long Term Plan*. Brampton, ON: WSP.

Region of Peel (2019b) *Let's Move Peel: Long Range Transportation Plan 2019*. Brampton, ON.

Region of Peel (2019c) *Peel Data Centre*. Available at: http://opendata.peelregion.ca/ (Accessed: 1 April 2019).

Region of Peel (2019d) 'Truck Restrictions on Local Roads Intersecting Regional Roads (August 2019)'. Brampton, ON.

Ripley, W. N. V. and B. D. (2018) *Package 'MASS'*. R Foundation for Statistical Computing. doi: 10.1111/j.1467-9876.2010.00729.x>.Maintainer.

Saal, T. F., and Golias, B. J. (1997) *Famous but Forgotten: The Story of Alexander Winton, Automotive Pioneer and Industrialist*. Golias Pub.

Savitsky, T. (2018) *Package 'growfunctions'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/growfunctions/growfunctions.pdf.

Schmid, J., Wang, X. (Cara), and Conway, A. (2018) 'Commercial Vehicle Parking Duration in New York City and Its Implications for Planning', *Transportation Research Part A: Policy and Practice*. Elsevier, 116(November 2016), pp. 580–590. doi: 10.1016/j.tra.2018.06.018.

Scott, A., Balthrop, A., and Miller, J. W. (2019) *Did the Electronic Logging Device Mandate Reduce Accidents?* Available at: https://ssrn.com/abstract=3314308.

Seo, Y.-W., and Urmson, C. (2009) *A Hierarchical Image Analysis for Extracting Parking Lot Structures from Aerial Images*. Pittsburgh, Pennsylvania: Robotics Institute, Carnegie Mellon University.

Shmueli, G. (2011) 'To Explain or to Predict?', 25(3), pp. 289–310. doi: 10.1214/10-STS330.

Shmueli, G., and Koppius, O. R. (2011) 'Predictive Analytics in Information Systems Research', *MIS Quarterly*, 35(3), pp. 553–572.

Short, J., and Murray, D. (2016) *Identifying Autonomous Vechile Technology Impacts on the Trucking Industry*. doi: 10.1016/j.ijhydene.2008.11.062.

Shoup, D. C. (1995) 'An Opportunity to Reduce Minimum Parking Requirements', *Journal of the American Planning Association*, 61(1), pp. 14–28. doi: 10.1080/01944369508975616.

Shoup, D. C. (1999) 'The Trouble with Minimum Parking Requirements', *Transportation Research Part A: Policy and Practice*, 33(7–8), pp. 549–574. doi: 10.1016/S0965-8564(99)00007-5.

Smart Freight Centre (2019a) 'Smart Freight Centre'. McMaster University, University of Toronto, York University, and Region of Peel. Available at: http://smartfreightcentre.ca/.

Smart Freight Centre (2019b) *Smart Freight Centre Five Year Plan (2019-2023)*. McMaster University, University of Toronto, York University, and Region of Peel.

Standing Committee on Transport and Government Operations (2001) *TRGO Committee Meeting - May 3, 2001: Evidence*. Ottawa, ON: Government of Canada.

Statistics Canada (2017) *North American Industry Classification System (NAICS) Canada*. Ottawa, ON. Available at: https://www150.statcan.gc.ca/n1/en/pub/12-501-x/12-501-x2016001-eng.pdf.

Stephens, M. A. (1974) 'EDF Statistics for Goodness of Fit and Some Comparisons', *Journal of the American Statistical Association*, 69(347), pp. 730–737.

Stieve, T. (2012) *Moran's I and Spatial Regression*. Available at: https://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjR67-G2djdAhUB5IMKHVasCBcQFjAAegQIBRAC&url=http%3A%2F%2Fsites.tufts.edu%2Fgis%2Ffiles%2F2013%2F11%2FMorans-I-and-Spatial-Regression.docx&usg=AOvVaw12ilC_zqrjRh23-kBMk2b2.

Strohl, D. (2017) 'Automotive Hall of Fame to Induct Tractor-Trailer Pioneer Fruehauf', *Hemmings Daily*. Available at: https://www.hemmings.com/blog/2017/03/29/automotive-hall-of-fame-to-induct-tractor-trailer-pioneer-fruehauf/.

Takyi, E. A., Oluwajana, S. D., and Park, P. Y. (2018) 'Development of Macro-Level Crime and Collision Prediction Models to Support Data Driven Approach to Crime and Traffic Safety (DDACTS)', *Transportation Research Record: Journal of the Transportation Research Board*, 985, pp. 1–15. doi: 10.1177/0361198118777356.

Texas Department of Transportation (2012) *Traffic Recorder Instruction Manual*. Austin, TX. Available at: http://onlinemanuals.txdot.gov/txdotmanuals/tri/tri.pdf.

Texas Department of Transportation (2018) *Roadway Design Manual*. Austin, TX. Available at: http://onlinemanuals.txdot.gov/txdotmanuals/rdw/rdw.pdf.

The Canadian Press (2018) 'Fatal Collisions Involving Transport Trucks Hit 10-Year High in Ontario: Police', *CTV News*, 9 April. Available at: https://www.ctvnews.ca/canada/fatal-collisions-involving-transport-trucks-hit-10-year-high-in-ontario-police-1.3877833.

The Gale Group Inc. (2003) 'Trucking Industry', *Dictionary of American History*. 3rd edn. Charles Scribner's Sons.

Tobler, W. R. (2018) 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography*, 46, pp. 234–240. Available at: https://www.jstor.org/stable/143141.

Town of Caledon (2014) *Idling of Motor Vehicles By-Law No. 2014-078*.

Town of Caledon (2018) *Parking on Private Property By-Law No. 2018-1*.

Town of Caledon (2019) *Traffic By-Law No. 2015-058*.

Transoft Solutions (2015) 'AutoTURN: Advanced Vehicle Simulations'. Available at: https://www.cadac.com/media/3085/productblad_autoturn.pdf.

Trapletti, A., Hornik, K., and LeBaron, B. (2018) *Package 'tseries'*. R Foundation for Statistical Computing. Available at: https://cran.r-project.org/web/packages/tseries/tseries.pdf.

Trépanier, M., Morency, C., Demers, J., and Gaziello, J. (2010) *Analysis of the Commercial Vehicle Driver Survey Results (Rest Area Research Project—Phase I)*. Polytechnique Montréal.

Truck News (2017) 'Canada to Require ELDs by 2020', 18 December.

Trucking Research Institute, Apogee Research Inc., and Wilbur Smith Associates (1996) *Commercial Driver Rest and Parking Requirements: Making Space for Safety Final Report*. Washington, DC: Federal Highway Administration, US Department of Transportation. doi: 10.3141/1595-08.

U.S. Air Force (2015) *Global Positioning System*. Available at: http://www.af.mil/About-Us/Fact-Sheets/Display/Article/104610/global-positioning-system/ (Accessed: 25 January 2018).

U.S. Air Force (2017) *GPS Accuracy*. Available at: https://www.gps.gov/systems/gps/performance/accuracy/ (Accessed: 29 January 2018).

U.S. General Services Administration (2019) *Data.gov*. Available at: https://www.data.gov/.

UCLA: Statistical Consulting Group (2011) *FAQ: What Are Pseudo R-Squareds?* Available at: https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/ (Accessed: 30 November 2018).

United States of America: 105th Congress (1998) *Transportation Equity Act for the 21st Century*. United States of America. Available at: https://www.fhwa.dot.gov/tea21/pl105178.pdf.

United States of America: 112th Congress (2012) *H. R. 3235 - 'Jason's Law'*. United States of America: House of Representatives.

Wang, M., Nevland, E. A., and Park, P. Y. (2019) 'Average Speed Enforcement for Heavy Vehicles by Weight Classification: Simulation Results for the Laidlaw-Golden Freight Corridor', in *2019 Canadian Institute of Transportation Engineers Annual Conference*. Ottawa, ON.

Wang, X., and Hanson, A. R. (1998) 'Parking Lot Analysis and Visualization from Aerial Images', *WACV*, pp. 36–41.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L. (2011) *Statistical and Econometric Methods for Transportation Data Analysis*. 2nd edn. Boca Raton, FL: Taylor and Francis Group, LLC.

Washington, S., Persaud, B., Lyon, C., and Oh, J. (2005) *Validation of Accident Models for Intersections*. Washington, DC: Federal Highway Administration, US Department of Transportation.

White, H. (1980) 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity', *Econometrica*, 48(4), pp. 817–838. doi: 10.2307/2223855.

Wikimedia Commons (2015) *ISAAC InControl - Electronic On-Board Recorder (EOBR)*. Available at: https://commons.wikimedia.org/wiki/File:ISAAC_InControl_-_Electronic_on-board_recorder_(EOBR)_-_ISAAC_Instruments_inc.jpg (Accessed: 23 September 2019).

Wilson, P. (2015) 'The Misuse of the Vuong Test for Non-Nested Models to Test for Zero-Inflation', *Economics Letters*. Elsevier B.V., 127, pp. 51–53. doi: 10.1016/j.econlet.2014.12.029.

Van Wissen, L. (2000) 'A Micro-Simulation Model of Firms: Applications of Concepts of the Demography of the Firm', *Papers in Regional Science*, 79(2), pp. 111–134. doi: 10.1007/s101100050039.

Witten, G., Witten, D., Hastie, T., and Tibshirani, R. (2013) *An Introduction to Statistical Learning with Applications in R*, *Current Medicinal Chemistry*. New York: Springer. doi: 10.1007/978-1-4614-7138-7.

Wong, D. W. S. (2009) 'Modifiable Area Unit Problem', in Kitchin, R., and Thrift, N. (eds) *International Encyclopedia of Human Geography*. 1st edn. Elsevier.

Xu, G., and Xu, Y. (2016) *GPS: Theory, Algorithms and Applications*. 3rd edn. Springer. doi: 10.1007/978-3-662-50367-6.

Yager, T. (2009) 'Hours of Service: Background Information'. Washington, D.C.: Federal Motor Carrier Safety Administration, US Department of Transportation. doi: 10.1136/bmj.2.5411.758-a.

Yan, X., Levine, J., and Marans, R. (2019) 'The effectiveness of parking policies to reduce parking demand pressure and car use', *Transport Policy*. Washington, DC: Elsevier Ltd, 73(May 2018), pp. 41–50. doi: 10.1016/j.tranpol.2018.10.009.

Yang, X., Sun, Z., Ban, X., and Holguín-Veras, J. (2014) 'Urban Freight Delivery Stop Identification with GPS Data', *Transportation Research Record: Journal of the Transportation Research Board*, 2411, pp. 55–61. doi: 10.3141/2411-07.

Young, J., and Park, P. Y. (2013) 'Benefits of Small Municipalities using Jurisdiction-Specific Safety Performance Functions rather than the Highway Safety Manual's Calibrated or Uncalibrated Safety Performance Functions', *Canadian Journal of Civil Engineering*, 40(6), pp. 517–527. doi: 10.1139/cjce-2012-0501.

Zeileis, A., Kleiber, C., and Jackman, S. (2008) 'Regression Models for Count Data in *R*', *Journal of Statistical Software*, 27(8). doi: 10.18637/jss.v027.i08.

Zhang, Q., and Couloigner, I. (2006) 'Benefit of the Angular Texture Signature for the Separation of Parking Lots and Roads on High Resolution Multi-Spectral Imagery', *Pattern Recognition Letters*, 27(9), pp. 937–946. doi: 10.1016/j.patrec.2005.12.003.

Zmud, J., Lawson, C. T., and Pisarski, A. (2014) *Making Trucks Count: Innovative Strategies for Obtaining Comprehensive Truck Activity Data*. Washington, DC: Transportation Research Board. doi: 10.17226/22327.

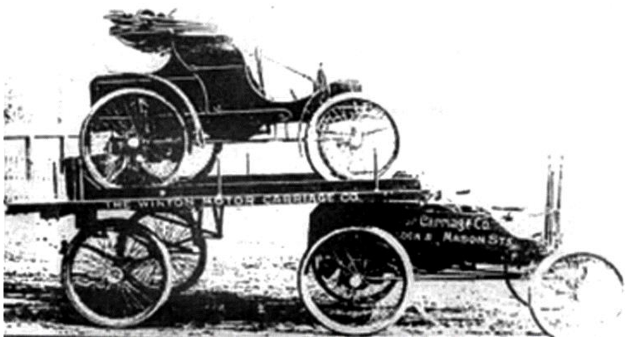
APPENDICES

APPENDIX A: HISTORY OF GOODS MOVEMENT

The modern trucking industry in North America is generally accepted to have evolved out of the horse-drawn carts of the late nineteenth century. Rail was the key mode for long-haul goods movement; however, this limited origins and destinations as locations adjacent to rail tracks. Horse-drawn carts helped solve this last-mile delivery problem, as these carts were capable of delivering goods to more secluded areas (Burks *et al.*, 2010).

In the late 1890s, motor vehicles began to be developed and used throughout North America. These vehicles primarily moved people, but some were designed as motor wagons to assist with the movement of goods. Between 1897 and 1900, several commercial motor wagons were introduced to the market. These primitive vehicles were designed to function similarly to a horse-drawn carriage and used electric engines which were favoured over nascent gasoline and steam engines. These "horseless carriages" were limited to short-hauls on paved urban roads due to driving-range, infrastructure, freight capacity, and reliability constraints. In part due to these restrictions, these early motorized vehicles were mostly used as a novelty during the first decade of the 1900s (The Gale Group Inc., 2003).

During this early period of commercial motor wagons, the first semi-truck was invented in 1898 by Alexander Winton to haul automobiles, as shown in Figure A-1(a). This vehicle was a modified short-wheeled touring automobile with a cart attached to its rear end (Saal and Golias, 1997). Over the next decade, technological improvements began to make the development of commercial motor wagons profitable (The Gale Group Inc., 2003). Initially, the most popular design of commercial motor wagons were single-unit trucks like that shown in Figure A-1(b); however, in 1912 August Charles Fruehauf built what he called the semi-trailer, as shown in Figure A-1(c). This semi-trailer, initially designed to haul a boat but later adapted to haul other goods such as lumber, was attached to a modified Ford Model T (Strohl, 2017; Great Western Transportation, 2019). The unique aspect of this design was that the hitch was positioned above the rear axle instead of behind it, which greatly increased the loading capacity of the trailer (Strohl, 2017). After this invention, the semi-trailer industry began to take off (Great Western Transportation, 2019).



(a) Alexander Winton's 1898 Semi-Truck
(Saal and Golias, 1997)

(b) General Motors 1912 Model H 3½ ton
(Horseless Age Company, 1912)

(c) August Fruehauf's 1914 Semi-Trailer on a Modified Ford Model T
(Strohl, 2017)

Figure A-1: Early Heavy Commercial Vehicles

Due to the advancements in combustion engines and improvement in the design of trucks, first and last-mile freight movement began to be moved increasingly by motor vehicle. The first World War solidified North America's reliance on the combustion engine. Twenty-five thousand trucks were produced in 1914, and by 1920 there were over 1.1 million registered trucks in the United States. Solid tires restricted speeds to 25 km/h, but the invent of air-filled tires that could handle heavier loads increased this speed dramatically. The development of interstate roadways and better engines in the 1920s allowed for the possibility of long-haul trips. This gave truck drivers the ability to travel further and for longer periods of time, which led to the first concerns with long-haul truck driver fatigue (The Gale Group Inc., 2003).

The US Interstate Commerce Commission (ICC), a regulatory agency in the United States, began examining an increasing number of collisions on the road involving freight vehicles. After several proposals and hearings, in 1937 the ICC adopted the first hours-of-service (HOS) rules in North America. Technology at this time allowed trucks to travel as fast as 65 km/h with a maximum daily distance of 400 km (The Gale Group Inc., 2003). The ICC concluded that drivers required approximately eight hours of sleep every twenty-four hours, so a law was passed to limit driver on-duty time to fifteen hours out of every twenty-four. Additionally, only twelve of these fifteen hours could be used for functions such as loading, unloading, driving, or reporting, with the intent that drivers be given three hours each day for meals and breaks (DOT, 2000).

These laws were quickly challenged by interest groups, including various representatives of organized labour, who argued for HOS limits of eight hours per day and forty-eight hours per week. The ICC conceded to the interest groups that there was no quantitative information available at the time that confirmed an exact amount of sleep required for drivers; however, they also expressed that in their opinion the labour group’s suggestions were not reasonable. In 1938, the ICC changed the HOS laws to limit driving to ten hours in twenty-four but only required eight hours between shifts. In addition to this, the daily on-duty limit was removed which essentially allowed drivers to work for as long as sixteen hours in a twenty-four-hour period as long as a weekly maximum of sixty hours was not exceeded. Additionally, an extra two hours of driving time was permitted in the case of unfavourable weather conditions impacting the scheduled route (DOT, 2000).

In Canada, HOS laws were not differentiated from general labour requirements until 1987 when they were moved from the Canada Labour Code to the Motor Vehicle Transportation Act. The federal government first began to regulate trucking in 1954 with the Motor Vehicle Transport Act, where it delegated the majority of the regulatory responsibility to the provinces (Historica Canada, 2014). Many of the provinces did not begin to specifically regulate HOS separate from general labour laws until the late 1980s when Standard 9 of the National Safety Code was established and required minimum restrictions to be followed. Standard 9 outlined a thirteen-hour limit to driving in twenty-four hours and a minimum of eight-hours of off-duty time between, with an exception allowing for a single four-hour rest period a week (Standing Committee on Transport and Government Operations, 2001).

As more research began to be undertaken in relation to HOS in North America, it quickly became clear that the existing rules had many issues. In the mid-1990s, the House Service Committee of the Canadian Council of Motor Transport Administrators (CCMTA) began to review the current rules relating to HOS with the intent of revising Standard 9 to be a more scientifically rigorous document to guide provincial regulators. In 1995, the ICC Termination Act required the US government to address its current HOS laws. The Canadian and US governments agreed to collaborate on their research as they both shared the goal of creating more effective, harmonized HOS laws (Durdle, 2006; Yager, 2009).

At the same time as HOS laws were being re-examined in North America, HCV drivers first began to report large-scale truck parking deficiencies. A report from the US Federal Highway Administration (FHWA) in 1996 found that HCV parking shortages were present along the US highway system. This was said to be due to significant growth in goods being moved by HCV and increased demand for parking relating to strict HOS enforcement. The report concluded that although the failure of meeting HCV driver's parking needs could pose a significant risk to the public, it is expected that significant private investment in expanding HCV parking facilities could improve the current shortfall in parking (TRI, Apogee Research Inc. and Wilbur Smith Associates, 1996). This attitude of inaction was the general approach of most jurisdictions, as HCV parking would not begin to be recognized as a nationwide concern for nearly two decades (United States of America: 112th Congress, 2012).

After nearly a decade of collaborative study, the US and Canadian governments both reformed their HOS laws. The United States adopted new HOS laws in 2003 restricting HCV drivers to eleven hours behind the wheel, fourteen total hours of on-duty time (not extendable with external circumstances), and a minimum of ten hours off duty time (Yager, 2009). In Canada, proposed regulations were published in 2003 and were eventually implemented in 2007. Although the Canadian laws became stricter, they were still less restrictive than the US laws. Canadian drivers were permitted thirteen hours of driving and fourteen total hours on duty time followed by ten hours of off-duty time which could be broken into two parts (Durdle, 2006). Although the two countries both adopted different regulations, they were both consistent in making restrictions tougher and reducing loopholes in their laws. This approach led to an increased pressure on HCV drivers to stay within the law as little to no leeway was provided for drivers.

Once consequence of strictly enforced HOS laws is that HCV drivers are often forced to choose between parking illegally or driving longer than the law legally permits if they cannot easily find suitable parking. Further complicating this matter is the fact that an American study found that 63% of HCV drivers had no awareness of nearby truck stops (Adams *et al.*, 2009). It was only after the death of Jason Rivenberg and the passing of Jason's Law in the United States that opinions began to change on the importance of adequate HCV parking.

Summaries of the changes to HOS laws over time in both Canada and the United States are shown in Table A-1 and Table A-2, respectively.

Table A-1: HOS Laws in Canada over Time

Year	Driving Limit	Total Work Limit	Off-Duty Requirement	7/8/14 Day Limits
Prior to 1987	HOS governed by various labour laws			
1987	13 hours	none	8 hours	60/70/120 hours
2007	13 hours	14 hours	10 hours	70/-/120 hours
2021	Expected the same as before, with mandatory ELD usage			

(Standing Committee on Transport and Government Operations, 2001; Durdle, 2006; Historica Canada, 2014; Government of Canada, 2017; Lamb, 2019)

Table A-2: HOS Laws in the United States over Time

Year	Driving Limit	Duty Time	Off-Duty Requirement	Time Limit Extensions	7/8 Day Limits
1937	10 hours	24 hours	8 hours	Yes	60/70 hours
1962	10 hours	15 hours	8 hours	Yes	60/70 hours
2003	11 hours	14 hours	10 hours	No	60/70 hours
2017	Same as 2003, with mandatory ELD usage				

(DOT, 2000, FMCSA, 2016; The Gale Group Inc., 2003; Yager, 2009)

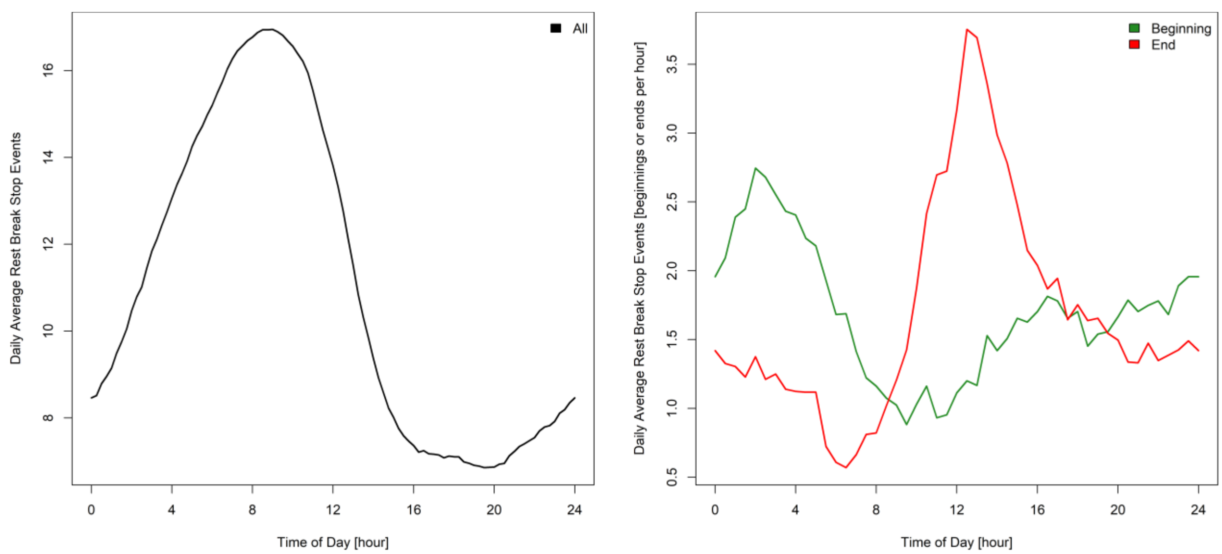
APPENDIX B: DETAILS OF HCV STOP EVENT TEMPORAL ANALYSIS

HCV parking demand, which leads to spaces being occupied, varies over time both hourly, daily, and seasonally. Additional factors that may yield in temporal changes include weather, events such as sales or natural disasters that change the demand for HCV trips, and the economy. In order to properly understand HCV parking demand, it is important to identify the temporal distribution and peak demand time of parked vehicles. The peak demand period corresponds to the time where driver rest time tends to be highest (FHWA, 2015). It is commonly understood that the peak demand period occurs sometime during the nighttime (FHWA, 2015). The peak demand period is important in determining parking adequacy as it corresponds to the time when parking supply is most likely to be fully utilized and capacity shortfalls may lead to illegal parking in areas like freeway ramps and shoulders (FHWA, 2015; Anderson and Roll, 2018). The Atlanta Regional Truck Parking Assessment Study found that peak demand often occurs during either the evening (7:00 PM to midnight) or the early morning (midnight to 5:00 AM) (Atlanta Regional Commission, 2018).

The following temporal distributions have been developed by examining the 14,766 stop events for long-haul HCV trips identified within the Region of Peel for the year 2014. The temporal datum of the results has not been verified; this means that although the shape of the distributions is likely correct that the position of the distributions may need to be shifted. Further verification of the time stamps associated with the GPS pings is required before comparing HCV parking supply and demand. The figures were developed in R using the `plot()` function (R Core Team, 2013).

Figure B-1 (a) shows the average number of identified long-haul rest events that were occurring at any time on an average day in the Region of Peel in 2014. This distribution shows that the examined sample has a peak long-haul parking demand at approximately 9:00 AM and the lowest demand occurs around 8:00 PM. This does not align with what is identified in the literature, which often notes that peak HCV parking demand occurs in the nighttime between 7:00 PM and 5:00 AM (FHWA, 2015; Atlanta Regional Commission, 2018). This could indicate that the time stamps associated with the GPS pings may need to be corrected before further HCV parking demand analysis.

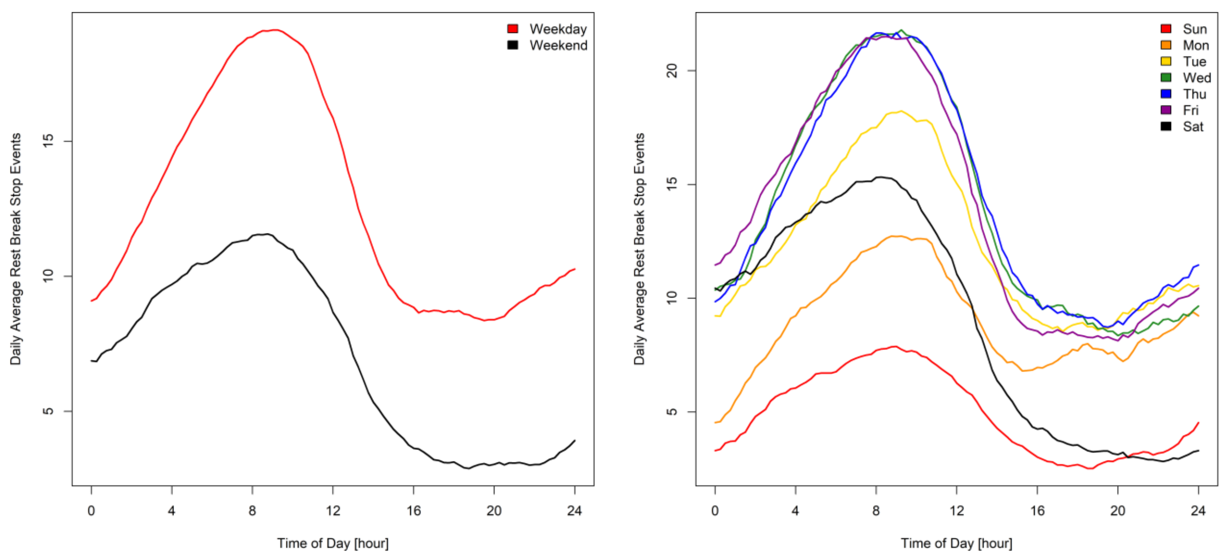
Figure B-1 (b) shows the distributions of when HCV drivers start and end their rest events. This figure shows that drivers most often start their rest events around 2:00 AM and end their rest events around 12:00 PM (noon), noting that these times may require correction.



(a) Total Stop Events Occurring **(b)** Stop Event Beginning and End Times

Figure B-1: Long-Haul HCV Stop Events in the Region of Peel

Figure B-2 shows the distribution of HCV rest events occurring during different days of the week. Assuming that the time stamps do not require adjustment greater than ±8 hours, these figures indicate that there is a greater demand for HCV parking during the weekdays when compared to the weekends. In fact, it appears as though the greatest demand for HCV parking is seen on Wednesdays, Thursdays, and Fridays, and that Saturday experiences a greater peak demand than Monday. All five weekdays experience a similar minimum demand, which is larger than the minimum demand experienced on the weekends.



(a) Weekday vs Weekend **(b)** Each Day Separately

Figure B-2: Day of Week Variations of Long-Haul HCV Stop Events in the Region of Peel

Figure B-3 shows the seasonal variation of HCV parking demand in the Region of Peel. This figure shows that the highest peak demand is seen in the summer (July to September) and the minimum is experienced in the winter (January to March). Spring (April to June) and fall (October to December) experience similar demand that is much closer to winter than to summer.

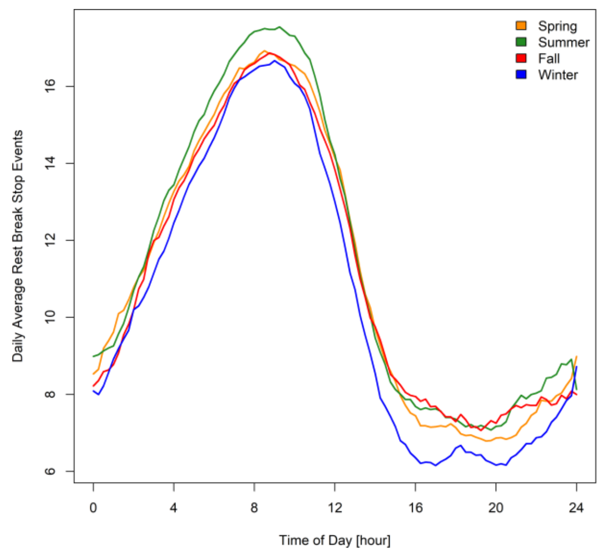


Figure B-3: Seasonal Variations of Long-Haul HCV Stop Events in the Region of Peel

Figure B-4 shows the demand for HCV parking separated by the length of the rest event. It can be seen that short and long rest events have very different peak demand times. Without correcting for any issues related to timestamps, it appears as though the highest demand for short rest events (two to five hours) occurs in the evening hours, the highest demand for longer length rest events occurs in highest demand for longer rest events (greater than six hours) occurs in the morning hours. After rest stops exceed nine hours in duration, the net demand begins to decrease and the difference between peak period and non-peak period demand diminishes. For this thesis, rest events were defined to take between two and fifteen hours, which is the minimum duration that can count towards HOS requirements and 1.5 times the required off-duty time (ten hours) for HOS laws in Canada (Government of Canada, 2009).

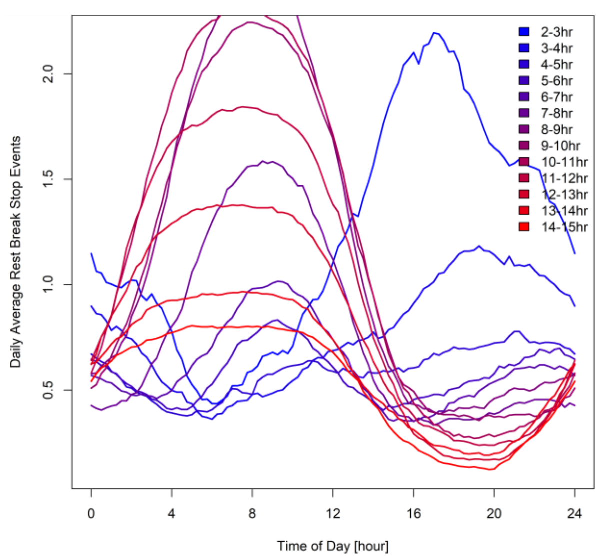


Figure B-4: Stop Length Variation of Long-Haul HCV Stop Events in the Region of Peel

APPENDIX C: DETAILS OF DATA STATISTICAL ANALYSIS

Appendix C.1: Outliers

Outliers are defined as observations that 'appear' to lie outside a large number of similar observations. Outliers should be removed if they result from erroneous data collection or analysis as this will negatively affect a developed model. However, outliers can also naturally arise in a dataset and removing these values can constrict the accuracy of a model to predict certain cases. Whenever possible, it is important to consider these values in some way, be it modeled separately or through a detailed justification of their omission (Washington *et al.*, 2011).

Several techniques for identifying outliers include calculating and examining the Cook's D, leverage, and Mahalanobis distance values for each observation (Aguinis *et al.*, 2013). These techniques are definitely not exhaustive; however, these three tests in combination will be used to identify outliers in a dataset.

The first technique considered in identifying outliers is calculating Cook's Distance: a measurement of a data point's influence on the regression coefficients as a whole (Aguinis *et al.*, 2013). Cook's distance can be found using the formula:

$$D_i = \frac{\sum_{j=1}^{n}\left(y_j - \hat{y}_{j(i)}\right)^2}{k(MSE)} \tag{15}$$

where n is the number of observations, y_j is the dependent variable, $\hat{y}_{j(i)}$ is the predicted dependent variable from a multiple linear regression, k is the number of independent variables, and MSE is the mean square error, as outlined in equation (36) (Washington *et al.*, 2011). A cutoff value for Cook's D can then be calculated from an F distribution with $\alpha = 0.50$ and $df = (k + 1, n + 1)$ as outlined by Aguinis, Gottfredson, and Joo (2013); if a Cook's D is larger than the F test value then the observation can be considered an outlier (Aguinis *et al.*, 2013). In R, the Cook's D values can be calculated using the `cooks.distance()` function (R Core Team, 2013).

The second method used to identify outliers is calculating the leverage of an observation: the measurement of how much that observation is an outlier when observed in the space of predictors (Aguinis *et al.*, 2013). This leverage is the diagonal elements of the hat matrix:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \text{ so that } \widehat{\mathbf{Y}} = \mathbf{HY} \tag{16}$$

where $\mathbf{H}$ is the hat matrix, $\mathbf{X}$ is a matrix of the independent variable observations, $\mathbf{Y}$ is a vector of the dependent variable values, and $\widehat{\mathbf{Y}}$ is a vector of the predicted values for the dependent variable from a multiple linear regression (Washington *et al.*, 2011). This means the individual leverage values are:

$$h_{ii} = \mathbf{X}_i^T(\mathbf{X}^T\mathbf{X})\mathbf{X}_i \tag{17}$$

where h_{ii} is the leverage value for observation i and $\mathbf{X}_i$ is the vector that corresponds to the ith row of $\mathbf{X}$ (Washington *et al.*, 2011). For large samples, outliers can then be identified as having a leverage value greater than $2(k+1)/n$ where n is the number of observations and k is the number of independent variables (Aguinis *et al.*, 2013). In R, the Leverage values can be calculated using the `hat()` function (R Core Team, 2013).

The final method used to identify outliers is the Mahalanobis distance which is defined as the distance between a data point and the centroid of the remaining data points (Aguinis *et al.*, 2013). The Mahalanobis distance is calculated similarly to Euclidian distance except with a consideration of a variance-covariance matrix. The formula for Mahalanobis distance is:

$$MD_i = \sqrt{(\boldsymbol{x}_i - \overline{\boldsymbol{x}})\mathbf{C}_x^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^T} \tag{18}$$

where $\boldsymbol{x}_i$ is a vector of observations for each independent variable, $\overline{\boldsymbol{x}}$ is a vector of the average for each independent variable, and $\mathbf{C}_x$ is the variance-covariance matrix calculated as:

$$\mathbf{C}_x = \frac{1}{n-1}\mathbf{X}_c^T\mathbf{X}_c \tag{19}$$

where n is the number of observations and $\mathbf{X}_c$ is the mean-centred matrix for the independent variable $(\mathbf{X} - \overline{\mathbf{X}})$ (Maesschalck *et al.*, 2000). A cutoff value for outliers in large sized sample can be set to the chi-squared distribution (χ^2) value with $\alpha = 0.05$ and $df = k$ where k is the number of independent variables (Aguinis *et al.*, 2013). In R, the Mahalanobis distance values can be calculated using the `mahalanobis()` function (R Core Team, 2013).

As previously mentioned, error outliers are one additional type of outlier that should be removed. Error outliers are observations that result from inaccuracies or mistakes. It is important to examine a dataset to ensure that all values occur within the limits of possibility and are entered correctly. No *a priori* theory is required to justify removing an error outlier (Aguinis *et al.*, 2013).

Appendix C.2: Multicollinearity

Multicollinearity is the state when two or more independent variables are correlated with each other. Multicollinearity among independent variables in a regression can cause issues with the accuracy of the model. Multicollinearity can lead to high standard errors of the estimated regression coefficients. In addition to this, multicollinearity makes the interpretation of regression coefficients counterintuitive as some of the explained variance from one independent variable is being explained by a coefficient for another correlated independent variable. This makes it impossible to accurately determine the effects of changing the value of one independent variable by one unit. Multicollinearity can also lead to large variability in estimated regression coefficients depending on the dataset used to obtain the model. In order to minimize the multicollinearity in a model it is important to properly design a model *a priori* (Washington *et al.*, 2011).

One measure of the total multicollinearity in a model is the multicollinearity condition number. The multicollinearity condition number determines multicollinearity by examining the eigenvalues of a matrix derived from the selected independent variables. It is suggested that multicollinearity will exist for each small eigenvalue of this matrix. The multicollinearity condition number is calculated as:

$$\kappa = \sqrt{\frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}} \tag{20}$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalues of $\mathbf{X}'\mathbf{X}$ where $\mathbf{X}$ is the matrix of the independent variables. This matrix, $\mathbf{X}$, comes from the multiple linear regression formula:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon \tag{21}$$

where $\mathbf{X}$ has dimensions $n \times (k + 1)$ where n is the number of observations and k is the number of independent variables. The first column of $\mathbf{X}$ consists of a vector of ones to account for a constant in the model (Alin, 2010).

In interpreting the multicollinearity condition number, a value of $\kappa < 10$ is generally considered to indicate no multicollinearity exists in the model. If $\kappa > 30$, it is said to indicate multicollinearity and $\kappa > 100$ is said to indicate extreme multicollinearity. In R, the multicollinearity condition number can be calculated using the `checkm()` function as a part of the `lrmest` package (Dissanayake and Wijekoon, 2016).

One approach often used to determine where multicollinearity exists in a set of independent variables is to develop a matrix of correlation values (Witten *et al.*, 2013). The Pearson product-moment correlation parameter, or simply the correlation parameter, can be calculated for a sample as:

$$r = \frac{COV(x_1, x_2)}{s_{x_1} s_{x_2}} \tag{22}$$

where $COV(x_1, x_2)$ is the sample covariance between the variables x_1 and x_2. s_{x_1} and s_{x_2} are the sample standard deviations of x_1 and x_2, respectively (Washington *et al.*, 2011). The sample covariance between x_1 and x_2 can be calculated as:

$$COV(x_1, x_2) = \frac{\sum_{i=1}^{n} (x_{1,i} - \bar{x}_1)(x_{2,i} - \bar{x}_2)}{n - 1} \tag{23}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means for x_1 and x_2, respectively, and n is the sample size (Washington *et al.*, 2011). The Pearson correlation parameter can be calculated using `cor()` function in R (Paradis, 2005).

The Pearson correlation parameter has limited usefulness as three or more variables may be highly correlated despite no two pairs having a noteworthy correlation value. A more robust test to identify the individual variables that are correlated is to calculate the variance inflation factor (VIF) for each of the independent variables (Witten *et al.*, 2013). The VIF is calculated as:

$$VIF_i = \frac{1}{1 - R_i^2} \quad \text{for } i = 1, 2, \ldots, k \tag{24}$$

where i is the variable the VIF is being calculated for, k is the total number of independent variables, and R_i^2 is the coefficient of multiple determination of x_i (independent variable i) on the remaining explanatory variables (Alin, 2010). An explanation of the coefficient of multiple determination can be found with equation (41). Each independent variable's VIF measures the increase in the standard errors of the linear regression coefficients due to the multicollinearity and relative to the variance that would be present without multicollinearity (Alin, 2010). In R, the VIF values can be calculated using the `checkm()` function as a part of the `lrmest` package (Dissanayake and Wijekoon, 2016).

Appendix C.3: Heteroskedasticity

Heteroskedasticity, also spelled heteroscedasticity, is defined as non-consistent variance in one variable across the range of another. The opposite of heteroskedasticity is homoskedasticity. In a linear regression model, the values of regression parameters do not require the variables be homoskedastic; however, the variance of these parameters will change across the range of another variable. This means that the model is no longer the best linear unbiased estimate (BLUE) due to the heteroskedasticity (Washington *et al.*, 2011).

One test used to detect heteroskedasticity is the Goldfeld–Quandt test. This test examines the null hypothesis of homoskedasticity and an alternative hypothesis that the observation variance has a linear relationship with the square value of the observation values ($\sigma_i^2 = cx_i^2$). This is examined using two OLS regressions: (1) using data hypothesised to have low-variance errors and (2) using data hypothesised to have high-variance errors. Typically, any data with an average variance (the middle 20%) is excluded from this test. The test statistic for the Goldfeld–Quandt test is:

$$\frac{SSE_{\text{low variance}}}{SSE_{\text{high variance}}} \approx F\left(\alpha; df = \left(\frac{n-d-2k}{2}, \frac{n-d-2k}{2}\right)\right) \tag{25}$$

where SSE is the sum of square error, $F(\,)$ is an F statistic test, n is the total sample size, d is the number of observations excluded, and k is the number of independent variables in the model. Downsides to the Goldfeld–Quandt test is that it requires ordering the observations by their variance and it requires normally distributed errors (Washington *et al.*, 2011). The Goldfeld–Quandt test can be performed in R using the `gqtest()` function as part of the `lmtest` package (Hothorn *et al.*, 2018).

Another test used to detect heteroskedasticity is the Breusch-Pagan test (Washington *et al.*, 2011). This test examines the variance of residuals from an ordinary least squares (OLS) regression to see if they are correlated with the independent variables; if so, heteroskedasticity is present. This approach is based on the Lagrangian multiplier test (Breusch and Pagan, 1979). The Breusch-Pagan test does not require the observations to be ordered; however, it does assume that the errors of the OLS regression are normally distributed (Washington *et al.*, 2011). The Breush-Pagan test can be performed in R using the `bptest()` function as part of the `lmtest` package (Hothorn *et al.*, 2018).

The Koenker-Bassett test is a variant of the Breusch-Pagan test that eliminates the need for normal distribution of the errors (Koenker, 1981). In this test, the studentized version of the test statistic will be used which in part removes the requirement for the data to be normally distributed (Koenker, 1981; Hothorn *et al.*, 2018). The Koenker-Bassett test can be performed in R using the `bptest()` function as part of the `lmtest` package with the argument `studentize = TRUE` (Hothorn *et al.*, 2018).

The Park test is another method to detect heteroskedasticity in data. The Park test examines the same circumstance as the Breusch-Pagan test: an OLS model developed from heteroskedastic data should have errors that are related to one or more independent variables. The Park test examines the null hypothesis of homoskedasticity by performing an OLS regression on the data and then setting up a second regression in the form:

$$\ln\left(\varepsilon_i^2\right) = \beta_0 + \beta \ln(x_i) + u_i \tag{26}$$

where ε_i is the error from the original regression, u_i is the disturbance term (the error in the second regression), and the β is the regression parameter in the second regression. The null hypothesis supposes that $\beta = 0$. If there is a statistically significant relationship between $\ln\left(\varepsilon_i^2\right)$ and $\ln(x_i)$, the null hypothesis can be rejected and heteroskedasticity can be assumed; however, this test does assume that the errors of the OLS regression are normally distributed (Washington *et al.*, 2011). Although there is no known direct function in R to perform the Park test, code has been written using the `lm()` function and a simple logarithmic transformation (Witten *et al.*, 2013).

The White test uses a covariance matrix estimator to determine if there is heteroskedasticity in a model without knowledge of the exact form of the heteroskedasticity (White, 1980). This test does not require normality of error (Washington *et al.*, 2011). The test statistic for this approach is:

$$nR^2 \approx \chi^2\left(\alpha, df = \frac{k(k+1)}{2}\right) \tag{27}$$

where n is the sample size, R^2 is the constant-adjusted squared multiple correlation coefficient from the model, $\chi^2(\,)$ is the chi-squared distribution, and k is the number of independent variables in the model (plus 1 if a constant is included). The null hypothesis for this text is homoskedasticity (White, 1980). The White test can be performed in R using the `whites.htest()` function as part of the `het.test` package (Andersson, 2015).

Appendix C.4: Normality

Many tests and analyses require either the data or at least the errors to be normally distributed. Normally distributed data follows the distribution that arises due to the central limit theorem. A probability density function of the normal distribution is written as:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{28}$$

where μ is the distribution mean and σ^2 is the distribution variance (Montgomery and Runger, 2011). Two important characteristics of a normal distribution include symmetry and a kurtosis value of three (Washington *et al.*, 2011). Skewness is a measure of asymmetry and a symmetric distribution will have a skewness value of zero. The skewness of a sample is calculated as:

$$\hat{S} = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^3}{\hat{\sigma}_y^3} \tag{29}$$

where n is the number of observations, y_i is the values of the observations in the sample, $\bar{y}$ is the average value of y_i in the sample, and $\hat{\sigma}_y$ is the standard deviation of y_i in the sample (Forsberg, 2014). A visual aid of skewness, including the relationship between mean, median, and mode in symmetric, right-skewed, and left-skewed distributions, can be found in Figure C-1.

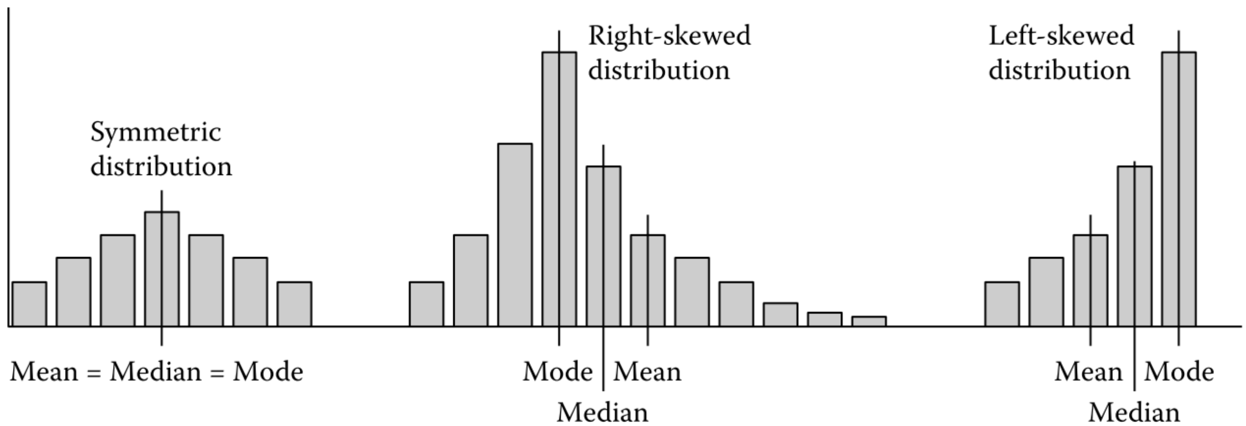


Figure C-1: Skewness of a Distribution
(Washington *et al.*, 2011)

Kurtosis is a measure of the 'flatness' or 'peakedness' of a distribution. A normal distribution is said to be mesokurtic, which means that it has a kurtosis value of three (Forsberg, 2014). Distributions with a kurtosis value larger or smaller than three are said to be leptokurtic or platykurtic, respectively, as seen in Figure C-2 (Washington *et al.*, 2011). The value of kurtosis can be calculated as:

$$\widehat{\mathrm{K}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^4}{\left(\hat{\sigma}_y^2\right)^2} \tag{30}$$

where n is the number of observations, y_i is the values of the observations in the sample, $\bar{y}$ is the average value of y_i in the sample, and $\hat{\sigma}_y$ is the variance of y_i in the sample (Forsberg, 2014). It should be noted that some literature discusses excess kurtosis, which is equal to the kurtosis value minus three; a normal distribution has an excess kurtosis value of zero (Forsberg, 2014).

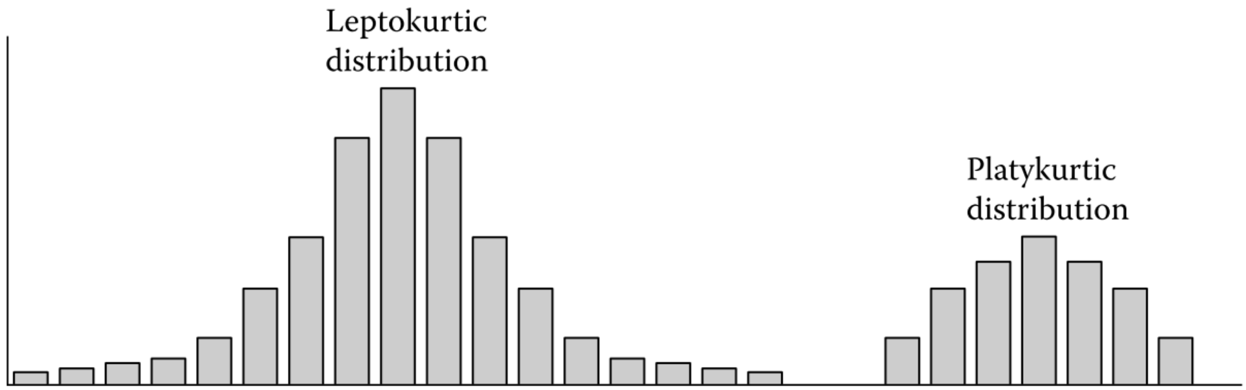


Figure C-2: Kurtosis of a Distribution
(Washington *et al.*, 2011)

The Jarque-Bera test is one method to determine if a dataset is normally distributed. This test uses the sample size, skewness, and kurtosis to test the null hypothesis that the data is normally distributed. The test statistics for the Jarque-Bera test is:

$$JB = n\left[\frac{\hat{S}^2}{6} + \frac{\left(\hat{K} - 3\right)^2}{24}\right] \tag{31}$$

where n is the sample size, $\hat{S}$ is the sample skewness, and $\hat{K}$ is the sample kurtosis. It should be noted that for a perfectly normal distribution, $JB = 0$ (Forsberg, 2014). This test statistic can be compared to the critical value of a chi-squared (χ^2) distribution with $df = 2$ given the preselected level of significance; if $JB > \chi^2(\alpha, df = 2)$, then the null is rejected and the sample can be assumed to not be normally distributed (Forsberg, 2014). The Jarque-Bera test can be performed in R using the `jarque.bera.test()` function as part of the `tseries` package (Trapletti *et al.*, 2018).

The Lilliefors test is a special case of the Kolmogorov-Smirnov test that is used to determine if a sample has been drawn from a normally distributed population. The Kolmogorov-Smirnov test can be used to determine if a sample follows a particular reference distribution (Conover, 1972). The Lilliefors test uses this same approach while setting the reference distribution to a normal distribution. This test is often used to determine if the error of a model is normally distributed. For this case, the null hypothesis for this test is normal distribution of the error (Abdi and Molin, 2007). The test statistic for this test is the maximum absolute difference between the observed distribution and the cumulative normal distribution function. This is compared to the p-value calculated from the Dallal-Wilkinson formula (Dallal and Wilkinson, 1986) if $p < 0.1$, or from a modified statistic derived from Stephens (1974) if $p > 0.1$. The Lilliefors test can be performed in R using the `lillie.test()` function as part of the `nortest` package (Gross and Ligges, 2015).

The Anderson Darling test is another test for normality in a dataset with the test statistic:

$$W = \frac{1}{12n} + \sum_{i=1}^{n} \left(\Phi([x_i - \bar{x}]/s) - \frac{2i-1}{2n} \right)^2 \tag{32}$$

where n is the sample size, $\Phi(\,)$ is the cumulative normal distribution function, x_i is the value of observation i, and $\bar{x}$ and s are the average value and standard deviation of x_i. Similar to the Lilliefors test, this modified test statistic can be compared to a modified statistic outlined by Stephens (1974). The Anderson-Darling test can be performed in R using the `ad.test()` function as part of the `nortest` package (Gross and Ligges, 2015).

Appendix C.5: Overdispersion

One consideration when selecting a model type is the relationship between the mean and the variance. If the variance is greater than the mean, the data is said to be overdispersed. The reverse of this, when the mean is larger than the variance, is underdispersion (Fischer and Wang, 2011). For transportation engineering data, overdispersion is much more common than underdispersion (Washington *et al.*, 2011).

One measurement of overdispersion, the overdispersion parameter, is used in the Negative binomial model (Washington *et al.*, 2011). The overdispersion parameter, α, is the value that satisfies the following equation:

$$\mathrm{E}[y_i] + \alpha \mathrm{E}[y_i]^2 = \mathrm{VAR}(y_i) \tag{33}$$

where $\mathrm{E}[y_i]$ is the mean of the dependent variable y_i and $\mathrm{VAR}(y_i)$ is its variance (Washington *et al.*, 2011). It should be noted that the negative binomial function in R, `glm.nb()`, reports θ as an alternative overdispersion parameter. This value of θ is simply the inverse value of α (Zeileis *et al.*, 2008).

Appendix C.6: Spatial Autocorrelation

The concept of spatial autocorrelation is a result of Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things" (Tobler, 2018). Spatial autocorrelation refers to a condition where values are related in some way, positively or negatively, to their neighbours. Positive spatial autocorrelation exists when values that are close to each other have similar values: the data can be said to be clustered. The opposite, negative spatial autocorrelation, exists when close observations are more dissimilar than further values: the data is said to be dispersed (Fischer and Wang, 2011; Stieve, 2012).

There are various methods for defining what areas are close to each other (are neighbours). Most measurements of spatial autocorrelation depend on a spatial weight matrix, W_{ij}, that is derived based on some measurement of distance between zones i and j (Fischer and Wang, 2011). Some approaches to deriving W_{ij} include contiguity or Euclidean distance between centroids. Distance weights can involve some sort of a decay function or can apply even weights to the k-nearest neighbours. Contiguity can use either a rook or queen rule, with the former not including areas that only share a corner as opposed to a length of border (Stieve, 2012).

Moran's I is one commonly used metric for spatial autocorrelation. Moran's I indicates how similar (or dissimilar) values are from their neighbours with a range of values from -1 to 1. Values close to either of these extremes indicate spatial autocorrelation where a value of 0 indicates a lack of spatial autocorrelation (Lopes *et al.*, 2007). The formula for calculating Moran's I is:

$$I = \frac{n}{W_o} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{34}$$

where n is the number of areas in the sample, i and j are any two zones, x_i is the observation of a variable in zone i, W_{ij} is the spatial weight matrix (where all $W_{ii} = 0$), and W_o is a normalizing factor equal to:

$$W_o = \sum_{i=1}^{n} \sum_{j \neq i}^{n} W_{ij} \tag{35}$$

(Fischer and Wang, 2011). A simplified visual representation of the Moran's I values can be seen in Figure C-3.

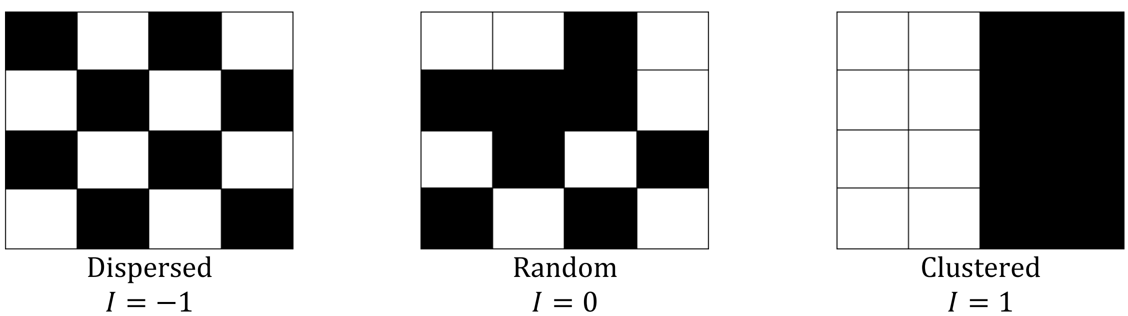


Figure C-3: Visual Interpretation of Moran's I
(Stieve, 2012, p. 10)

Most software will calculate a pseudo-p value to approximate the statistical significance of the Moran's I. This is performed by iterating the random assignment of values to the areas and calculating the distribution of the Moran's I for each of the iterations (Fischer and Wang, 2011; Stieve, 2012).

Local Indicators of Spatial Association (LISA) are local statistics relating to spatial autocorrelation in contrast to the Moran's I global statistic. LISA statistics break down global statistics in order to show the contribution of different areas (Lopes *et al.*, 2007; Fischer and Wang, 2011). One LISA statistic, a local Moran's I value, can be computed as:

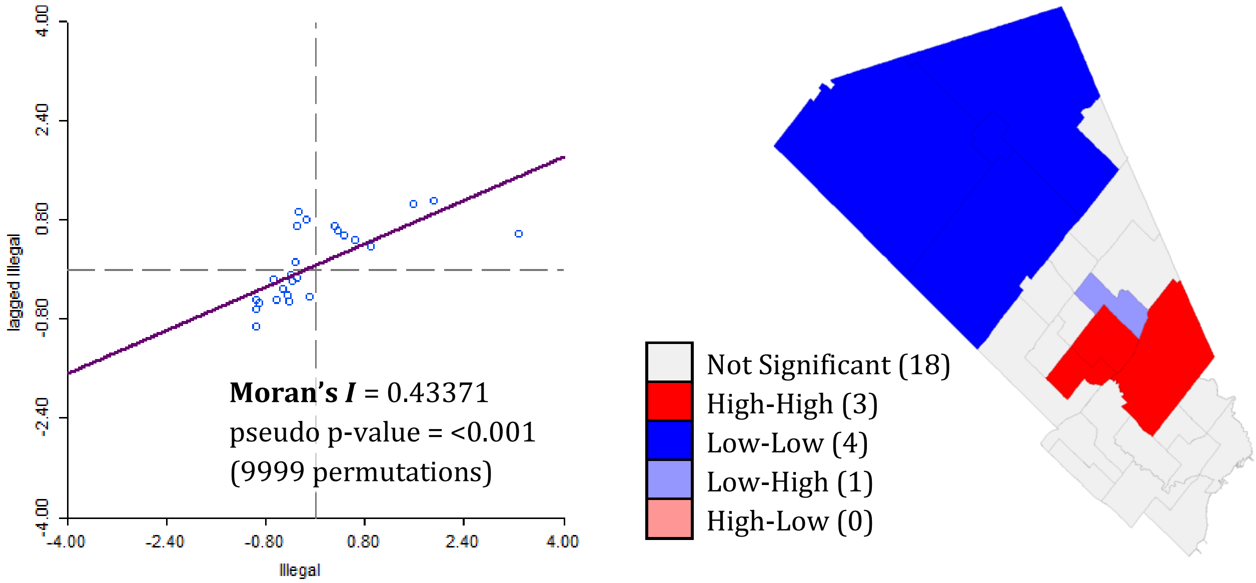
$$I_i = (x_i - \bar{x}) \sum_{j \in J_i}^{n} W_{ij} \left(x_j - \bar{x}\right)^2 \tag{36}$$

where J_i is the neighbourhood set of area i and other variables are the same as the global Moran's I (Fischer and Wang, 2011). Another LISA statistic, the global spatial variable, examines the relationship between an area's value, the area's neighbour's average value, and the global average. The classification of this statistic is shown in Table C-1 with the individual global spatial variables plotted in Figure C-4(a) and shown graphically in Figure C-4(b) (Lopes *et al.*, 2007).

Table C-1: Global Spatial Variable

Value	Neighbour's Average	Quadrant on Moran's Scatterplot	Statistic
Higher than average	Higher than average	Q1	High-High
Lower than average	Lower than average	Q2	Low-Low
Lower than average	Higher than average	Q3	Low-High
Higher than average	Lower than average	Q4	High-Low

It should be noted that the slope of the best fit line in the Moran's Scatterplot, shown in Figure C-4(a), is the global value of the Moran's I. Moran's scatterplot plots the difference in a value compared to the global average and the neighbours average value compared to the global average (Lopes *et al.*, 2007).



(a) **(b)**

Figure C-4: Spatial Autocorrelation Metrics
(a) Moran's Scatterplot; (b) LISA: Individual Global Spatial Values
Created using GeoDa (Anselin, 2017) based on the Region of Peel Illegal Parking Parcel Density

An alternative that is sometimes used instead of Moran's I is Geary's c, which mathematically uses the squared differences instead of cross-products to measure value association (Fischer and Wang, 2011). Geary's c is calculated as:

$$c = \frac{(n-1)}{2W_o} \frac{\sum_{i=1}^{n} \sum_{i=1}^{n} W_{ij} \left(z_j - \bar{z}\right)^2}{\sum_{i=1}^{n} (z_i - \bar{z})^2} \tag{37}$$

where n is the number of areas in the sample, i and j are any two zones, z_i is the observation of a variable in zone i, W_{ij} is a spatial weight matrix (where all $W_{ii} = 0$), and W_o is a normalizing factor calculated in the same manner as for Moran's I (Fischer and Wang, 2011). The local Geary's c is calculated as:

$$c_i = \sum_{j \in J_i}^{n} W_{ij} \left(x_i - x_j\right)^2 \tag{38}$$

where the variables can be interpreted in the same was as for the local Moran's I (Fischer and Wang, 2011).

Two other LISA statistics that can be computed are the Getis local statistic (G_i) and the Ord local statistic (G_i^*) which both measure the local association of the neighbour values to the global values, with neighbour values determined by some previously determined distance. The Getis local statistic is calculated as:

$$G_i(d) = \frac{\sum_{j \neq i}^{n} W_{ij}(d) x_j}{\sum_{j \neq i}^{n} x_j} \tag{39}$$

and the Ord local statistic as:

$$G_i^*(d) = \frac{\sum_{j=1}^{n} W_{ij}(d) x_j}{\sum_{j=1}^{n} x_j} \tag{40}$$

where $W_{ij}(d)$ is the weight matrix as a function of distance d. The main difference between the two is that the Getis local statistic is exclusive of i (areas do not include their own value in the calculation) where the Ord local statistic is not (Fischer and Wang, 2011).

APPENDIX D: NUMERICAL GOODNESS-OF-FIT METHODS

There are many numerical methods for evaluating goodness-of-fit. One classic goodness-of-fit statistic for linear models is the coefficient of (multiple) determination, R^2, which is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{41}$$

where y_i is the actual dependent value, $\hat{y}_i$ is the predicted dependent value, and $\bar{y}$ is the mean of y_i (Washington *et al.*, 2011). In R, the easiest way to determine the R^2 is to look at the `r.squared` attribute of the `summary()` of the `lm()` function (R Core Team, 2013).

An R^2 that is adjusted to account for more than one independent variable in the model is:

$$R^2_{\text{adj}} = 1 - \left(\frac{n-1}{n-k}\right)\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{42}$$

where n is the number of observations and k is the number of parameters (Washington *et al.*, 2011). In R, the easiest way to determine the R^2_{adj} is to look at the `adj.r.squared` attribute of the `summary()` of the `lm()` function (R Core Team, 2013). There are also various pseudo-R^2 metrics for non-linear models that can provide a comparable goodness-of-fit value (UCLA: Statistical Consulting Group, 2011).

Mean Squared Error (MSE) is an unbiased estimate of the variance in a regression model and is calculated as:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-k} \tag{43}$$

where y_i is each actual value, $\hat{y}_i$ is the predicted values, n is the sample size, and k is the number of parameters (Washington *et al.*, 2011). The MSE can be found in R using the `mse()` function as part of the `Metrics` package (Hamner *et al.*, 2018). For examining data from another source like a validation or holdout dataset, the Mean Squared Predicted Error ($MSPE$) test is often used. The $MSPE$ is calculated as:

$$MSPE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n_2} \tag{44}$$

where y_i is the actual value from the external data set, $\hat{y}_i$ is the predicted value, and n_2 is the number of observations in the sample set. The $MSPE$ can be found in R using the `MSPE()` function as part of the `growfunctions` package (Savitsky, 2018). It can be noted that both the MSE and $MSPE$ square the difference between the predicted and actual data ensuring that all values are positive. This results in a positive sum with the smallest value indicating the best fit (Washington *et al.*, 2005).

One test that does not square the difference is the Mean Prediction Bias (MPB). The MPB is calculated as:

$$MPB = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \tag{45}$$

where y_i is each actual value, $\hat{y}_i$ is the predicted values, and n is the sample size. Since positive and negative values will cancel each other out, the MPB value that is the closest to zero (either positive or negative) will indicate the best fit (Washington *et al.*, 2005). MPB can be calculated for either the data used to build the model or a holdout dataset. Although there is no quick R function commonly available for this metric, it can be calculated using the `average()` function of the difference between the predicted and actual values (R Core Team, 2013).

A test similar to the MPB is the Mean Absolute Deviation (MAD); however, this test examines the absolute error in the model. The MAD is calculated as:

$$MAD = \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{n} \tag{46}$$

where y_i is each actual value, $\hat{y}_i$ is the predicted values, and n is the sample size. Positive and negative errors will not cancel each other out, so the smallest MAD value will indicate the best fit (Washington *et al.*, 2005). MAD can be calculated for either the data used to build the model or a holdout dataset. Although there is no quick R function commonly available for this metric, it can be calculated using the `average()` function of the absolute difference between the predicted and actual values (R Core Team, 2013).

More complex tests for determining model goodness-of-fit include the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both of these functions rely on the log-likelihood at convergence, $LL(\theta)$. AIC is calculated as:

$$AIC = 2k - 2LL(\theta) \tag{47}$$

(Washington *et al.*, 2011) and BIC as:

$$BIC = \ln(n)\, k - 2LL(\theta) \tag{48}$$

where k is the number of parameters and n is the sample size (Washington *et al.*, 2011). The smallest value of AIC and BIC for a series of models indicate the best fit (Washington *et al.*, 2011). These values can be calculated in R using the `AIC()` and `BIC()` functions (R Core Team, 2013). The $LL(\theta)$ is the maximized log of the likelihood function, which for a negative binomial regression is represented by:

$$\begin{aligned} LL(\theta) \ = \sum_{i=1}^{n} y_i \ln\left(\frac{\alpha \exp\left(x_i^T\beta\right)}{1+\alpha\exp\left(x_i^T\beta\right)}\right) - \frac{1}{\alpha}\ln\left(1+\alpha\exp\left(x_i^T\beta\right)\right) + \ln\Gamma\left(\frac{1}{\alpha}+y_i\right) - \ln\Gamma(y_i+1) \\ - \ln\Gamma\left(\frac{1}{\alpha}\right) \end{aligned} \tag{49}$$

where y_i is the dependent variable for each observation i, n is the number of observations, α is the overdispersion parameter, x_i is a matrix of the independent variables, β is a vector of the model parameters, and $\ln\Gamma(\,)$ is the log gamma function (Washington *et al.*, 2011; Takyi *et al.*, 2018).

A variant of the AIC goodness-of-fit test is the corrected Akaike Information Criteria (AIC_c). The corrected version of this test adjusts for any bias that may be present in the model due to a small sample size. AIC_c is calculated as:

$$AIC_c = AIC + \frac{2k^2 + 2k}{n - k - 1} = 2k - 2LL(\theta) + \frac{2k^2 + 2k}{n - k - 1} \tag{50}$$

where k is the number of parameters, $LL(\theta)$ is the maximized log of the likelihood function, and n is the sample size (Burnham and Anderson, 2002). It should be noted that:

$$\lim_{n \to \infty} AIC_c = AIC \tag{51}$$

since an infinitely-large sample size does not require adjustment for small sample size bias (Burnham and Anderson, 2002). AIC_c can be calculated in R by adding the small sample size bias term to the result of the `AIC()` function (R Core Team, 2013).

APPENDIX E: R SCRIPT FOR DATA ANALYSIS

Appendix E.1: Script

```
# R Language: Data Analysis Script
# By: Erik Nevland
# Prepared (in part) for DCAD 7060 Final Project
# Current as of November 28, 2018

# Must pre-load Data (SSC) and file path location (FileFolderLocation)

# Add Packages

library(MASS)
library(pscl)
library(lrmest)
library(lmtest)
library(tseries)
library(het.test)
library(nortest)

##############################################################################
##############################################################################

# Setup data

DV <- c("Parking_Count")
IVs <- c("PER_1000","Cl_Pub","Cl_Gas","IND59","PER_1000_RURAL") #Previous
Model #1
Alpha = 0.05
RemoveOutliers <- FALSE
ZeroInflate <- FALSE

Formula <- formula(paste(DV,"~",paste(IVs, collapse="+"),sep=""))
Data <- SSC
ZFormula <- formula(paste(DV,"~",paste(IVs, collapse="+")," | ",paste(ZIVs,
collapse="+"),sep=""))

EndMessage <- paste("Data Analysis Results",sep="")

##############################################################################
##############################################################################

# Outliers

Total <- "Total"
Outlier <- setNames(data.frame(matrix(ncol = 4, nrow = length(Data$PIN))),
c("CooksD", "Leverage", "Mahalanobis", "Total"))

#CooksD
```

```
Z1=6  #(Predictors + 1)
Z2=90       #(Observations - Predictors -1)
Outlier[,1] <- qf(.5,df1=Z1,df2=Z2) < cooks.distance(lm(Formula,data=Data))

#Leverage
Outlier[,2] <- hat(model.matrix(Formula,data=Data)) >
(2*(length(IVs)))/length(Data$PIN)

#Mahalanobis
Outlier[,3] <- qchisq(.95, df=length(IVs)) < mahalanobis(Data[, (names(Data)
%in% c(DV,IVs))],center=FALSE,cov(Data[, (names(Data) %in% c(DV,IVs))]))

#Total
Outlier$Total <- (Outlier$CooksD + Outlier$Leverage + Outlier$Mahalanobis)
for (i in (1:length(Outlier$Total))) {
      if (Outlier$Total[i] > 1) {
            Outlier$Total[i] <- TRUE } else
            Outlier$Total[i] <- FALSE }
Data <- data.frame(Data,Outlier)
names(Data)[names(Data) == 'Total'] <- 'Outlier'

EndMessage <- paste(EndMessage,"\n","  Outliers Identified","\n","    -
Cook's D: ",sum(Outlier$CooksD),"/",length(Data$PIN),"\n","    - Leverage:
",sum(Outlier$Leverage),"/",length(Data$PIN),"\n","    - Mahalanobis:
",sum(Outlier$Mahalanobis),"/",length(Data$PIN),"\n","    - Total Removed:
",sum(Outlier$Total),"/",length(Data$PIN)," (failed 2 or more tests)",sep="")

if (RemoveOutliers == TRUE) {RemovedData <- Data[(Data$Outlier==1),]}
if (RemoveOutliers == TRUE) {Data <- Data[!(Data$Outlier==1),]}

##########################################################################
#####01000101##01110000##01110011##01110100##01100101##01101001##01101110####

# Multicollinearity TestTypes       #checkm() is part of lrmest package

ConditionNumber <- checkm(Formula,data=Data)$`*****Condition number*****`
if (ConditionNumber < 30) {ConditionNumberResult <- "good"}
if (ConditionNumber >= 30 & ConditionNumber < 100) {ConditionNumberResult <-
"multicollinearity present"}
if (ConditionNumber >= 100) {ConditionNumberResult <- "extreme
multicollinearity"}

# VIF

VIF <- checkm(Formula,data=Data)$`*****VIF*****`

EndMessage <- paste(EndMessage,"\n","  Multicollinearity Tests","\n","    -
Condition Number: ",round(ConditionNumber, digits=3),"
(",ConditionNumberResult,")","\n","    - VIF:  ",sep="")

VIFResults <- ""
VIFCheck <- ""
```

```
for (i in (1:length(colnames(VIF)))) {
      if (i!=1) {VIFResults <- paste(VIFResults,"             ",sep="")}
      VIFResults <- paste(VIFResults,colnames(VIF)[i],":
",round(VIF[i],3),sep="")
      if (VIF[i] < 10) {VIFCheck <- "good"} else {VIFCheck <-
"Multicollinearity Present"}
      VIFResults <- paste(VIFResults," (",VIFCheck,")","\n",sep="")}

EndMessage <- paste(EndMessage,VIFResults,sep="")

#############################################################################
#####01100100##01101001##01100100##01101110##00100111##01110100#############

# Heteroskedasticity                #Remedy with Box-cox transformation

#Goldfeld-Quandt        #gqtest() is part of lmtest package

GQResult <- gqtest(Formula,data=Data)
GQSignificant <- ""
if (GQResult$`p.value` > Alpha) {GQSignificant <- "good"} else {GQSignificant
<- "heteroskedasticity present"}

EndMessage <- paste(EndMessage,"  Heteroskedasticity","\n","    - Goldfeld-
Quandt: ",round(GQResult$`statistic`,3),"
(p=",formatC(GQResult$`p.value`,format = "e", digits = 2),",
",substr(as.character(gqtest(Formula,data=Data)[2]),3,nchar(as.character(gqte
st(Formula,data=Data)[2]))-1),"; ",GQSignificant,")",sep="")

#Breusch-Pagan          #bptext() is part of lmtest package #Apollo #Franklin

BPResult <- bptest(Formula,data=Data, studentize = FALSE)
BPSignificant <- ""
if (BPResult$`p.value` > Alpha) {BPSignificant <- "good"} else {BPSignificant
<- "heteroskedasticity present"}

EndMessage <- paste(EndMessage,"\n","    - Breusch-Pagan:
",round(BPResult$`statistic`,3)," (p=",formatC(BPResult$`p.value`,format =
"e", digits = 2),", df=",BPResult$`parameter`,"; ",BPSignificant,")",sep="")

#Koenker-Bassett
KBResult <- bptest(Formula,data=Data, studentize = TRUE)
KBSignificant <- ""
if (KBResult$`p.value` > Alpha) {KBSignificant <- "good"} else {KBSignificant
<- "heteroskedasticity present"}

EndMessage <- paste(EndMessage,"\n","    - Koenker-Bassett:
",round(KBResult$`statistic`,3)," (p=",formatC(KBResult$`p.value`,format =
"e", digits = 2),", df=",KBResult$`parameter`,"; ",KBSignificant,")",sep="")

#Park
```

```
Park1 <- lm(Formula,data=Data)
Formula2 <-
formula(paste("log((Park1$`residuals`)^2)~Park1$`model`$",paste(IVs,
collapse="+Park1$`model`$"),sep=""))

ParkResult <- lm(Formula2)
ParkP <- summary(ParkResult)$coefficients[,4]
ParkText <- paste("\n","    - Park:  ",sep="")

for (i in (2:(length(IVs)+1))) {
      if (!(i==2)) {ParkText <- paste(ParkText,"             ",sep="")
      }
            ParkText <- paste(ParkText,IVs[i-1],":
p=",formatC(ParkP[i],format = "e", digits = 2)," ",sep="")
      if (ParkP[i] > Alpha) {
            ParkText <- paste(ParkText,"(good)","\n",sep="")
      } else {
            ParkText <- paste(ParkText,"(heteroskedasticity
present)","\n",sep="")
      }
}

EndMessage <- paste(EndMessage,ParkText,sep="")


#White            #whites.htest() is part of het.test package

WResult <- whites.htest(VAR(Data[, (names(Data) %in% c(DV,IVs))],p=1))
WSignificant <- ""
if (WResult$`p.value` > Alpha) {WSignificant <- "good"} else {WSignificant <-
"heteroskedasticity present"}

EndMessage <- paste(EndMessage,"    - White: ",round(WResult$`statistic`,3),"
(p=",formatC(WResult$`p.value`,format = "e", digits = 2),",
df=",WResult$`degrees`,"; ",WSignificant,")",sep="")

############################################################################
#####01101011##01101001##01101100##01101100##################################

#Normal Distribution Supposition

# Jarque-Bera
JBResult <- jarque.bera.test(Data[, (names(Data) %in% c(DV))])
JBResultText <- ""
if (JBResult$`p.value`>Alpha) {JBResultText <- "normal"} else {JBResultText
<- "not normal"}
EndMessage <- paste(EndMessage,"\n","  Normal Distribution
Supposition","\n","    - Jarque-Bera:  ",DV,":
X^2=",round(JBResult$`statistic`,0)," (df=",JBResult$`parameter`,",
p=",formatC(JBResult$`p.value`,format = "e", digits = 2),";
",JBResultText,")",sep="")
```

```
for (i in (1:length(IVs))) {
      JBResult <- jarque.bera.test(Data[, (names(Data) %in% c(IVs[i]))])
      if (JBResult$`p.value`>Alpha) {JBResultText <- "normal"} else
{JBResultText <- "not normal"}
      EndMessage <- paste(EndMessage,"\n","                    ",IVs[i],":
X^2=",round(JBResult$`statistic`,0)," (df=",JBResult$`parameter`,",
p=",formatC(JBResult$`p.value`,format = "e", digits = 2),";
",JBResultText,")",sep="")
}

#Lilliefors

LResult <- lillie.test(Data[, (names(Data) %in% c(DV))])
LResultText <- ""
if (LResult$`p.value`>Alpha) {LResultText <- "normal"} else {LResultText <-
"not normal"}
EndMessage <- paste(EndMessage,"\n","    - Lilliefors:  ",DV,":
",round(LResult$`statistic`,3)," (p=",formatC(LResult$`p.value`,format = "e",
digits = 2),"; ",LResultText,")",sep="")

for (i in (1:length(IVs))) {
      LResult <- lillie.test(Data[, (names(Data) %in% c(IVs[i]))])
      if (LResult$`p.value`>Alpha) {LResultText <- "normal"} else
{LResultText <- "not normal"}
      EndMessage <- paste(EndMessage,"\n","                   ",IVs[i],":
",round(LResult$`statistic`,3)," (p=",formatC(LResult$`p.value`,format = "e",
digits = 2),"; ",LResultText,")",sep="")
}

#Andersson Darling

ADResult <- ad.test(Data[, (names(Data) %in% c(DV))])
ADResultText <- ""
if (ADResult$`p.value`>Alpha) {ADResultText <- "normal"} else {ADResultText
<- "not normal"}
EndMessage <- paste(EndMessage,"\n","    - Andersson Darling:  ",DV,":
",round(ADResult$`statistic`,2)," (p=",formatC(ADResult$`p.value`,format =
"e", digits = 2),"; ",ADResultText,")",sep="")

for (i in (1:length(IVs))) {
      ADResult <- ad.test(Data[, (names(Data) %in% c(IVs[i]))])
      if (ADResult$`p.value`>Alpha) {ADResultText <- "normal"} else
{ADResultText <- "not normal"}
      EndMessage <- paste(EndMessage,"\n","
",IVs[i],": ",round(ADResult$`statistic`,2),"
(p=",formatC(ADResult$`p.value`,format = "e", digits = 2),";
",ADResultText,")",sep="")
}

################################################################################
#####01101000##01101001##01101101##01110011##01100101##01101100##01100110####
```

```
#OverDispersion

OverDispersionParameter <- 1/(glm.nb(Formula,data=Data)$`theta`)
EndMessage <- paste(EndMessage,"\n","  Overdispersion","\n","    -
Overdispersion Parameter: ",round(OverDispersionParameter,3),"
(SE=",round(glm.nb(Formula,data=Data)$`SE.theta`,3),")",sep="")

#############################################################################
#############################################################################


EndMessage <- paste(EndMessage,"\n","  Spatial Autocorrelation","\n","    -
Please import data into GeoDa for Moran's I and LISA","\n","If results are
favorable, please proceed to power check, modeling, and goodness-of-fit
analysis. Thank you.",sep="")

# Save Data

if (TRUE == FALSE) {
      write.csv(Data, row.names = FALSE,
            (file =
paste(FileFolderLocation,"ShapefileTables/DCAD_Var1.csv",sep="")))
      }

#############################################################################
#############################################################################

RanBefore <- "Yes"

cat(EndMessage)
```

Appendix E.2: Output

```
Data Analysis Results
  Outliers Identified
    - Cook's D: 2/130
    - Leverage: 17/130
    - Mahalanobis: 26/130
    - Total Removed: 2/130 (failed 3 or more tests)
  Multicollinearity Tests
    - Condition Number: 6.762 (good)
    - VIF:  PavedArea_10E6: 8.185 (good)
            Perimeter_1000: 2.565 (good)
            ClPublic: 1.089 (good)
            IND59: 1.016 (good)
            IsRural: 1.168 (good)
            PavedXPerim: 10.325 (Multicollinearity Present)
  Heteroskedasticity
    - Goldfeld-Quandt: 2.95 (p=3.54e-05, df1 = 57, df2 = 57;
heteroskedasticity present)
    - Breusch-Pagan: 353.101 (p=3.33e-73, df=6; heteroskedasticity present)
    - Koenker-Bassett: 42.875 (p=1.23e-07, df=6; heteroskedasticity present)
    - Park:  PavedArea_10E6: p=2.92e-07 (heteroskedasticity present)
             Perimeter_1000: p=3.59e-04 (heteroskedasticity present)
             ClPubic: p=5.18e-04 (heteroskedasticity present)
             IND59: p=7.16e-01 (good)
             IsRural: p=2.31e-05 (heteroskedasticity present)
             PavedXPerim: p=2.83e-05 (heteroskedasticity present)
  Normal Distribution Supposition
    - Jarque-Bera:  ParkingCount: X^2=2265 (df=2, p=0.00e+00; not normal)
                    PavedArea_10E6: X^2=131 (df=2, p=0.00e+00; not normal)
                    Perimeter_1000: X^2=79 (df=2, p=0.00e+00; not normal)
                    ClPublic: X^2=56 (df=2, p=5.67e-13; not normal)
                    IND59: X^2=2755 (df=2, p=0.00e+00; not normal)
                    IsRural: X^2=2755 (df=2, p=0.00e+00; not normal)
                    PavedXPerim: X^2=574 (df=2, p=0.00e+00; not normal)
    - Lilliefors:  ParkingCount: 0.297 (p=1.98e-31; not normal)
                   PavedArea_10E6: 0.224 (p=2.50e-17; not normal)
                   Perimeter_1000: 0.146 (p=4.95e-07; not normal)
                   ClPublic: 0.496 (p=1.45e-91; not normal)
                   IND59: 0.541 (p=3.63e-109; not normal)
                   IsRural: 0.541 (p=3.63e-109; not normal)
                   PavedXPerim: 0.269 (p=1.77e-25; not normal)
    - Andersson Darling:  ParkingCount: 16.46 (p=3.70e-24; not normal)
                          PavedArea_10E6: 8.88 (p=1.11e-21; not normal)
                          Perimeter_1000: 4.99 (p=2.14e-12; not normal)
                          ClPublic: 34.89 (p=3.70e-24; not normal)
                          IND59: 47.01 (p=3.70e-24; not normal)
                          IsRural: 47.01 (p=3.70e-24; not normal)
                          PavedXPerim: 16.19 (p=3.70e-24; not normal)
  Overdispersion
    - Overdispersion Parameter: 2.474 (SE=0.056)
  Spatial Autocorrelation
    - Please import data into GeoDa for Moran’s I and LISA
If results are favorable, please proceed to power check, modeling, and
goodness-of-fit analysis. Thank you.>.
```

APPENDIX F: Q-Q PLOTS

Outliers removed in Figure F-1 and F-2 for display purposes only. See Section 8.4.2.

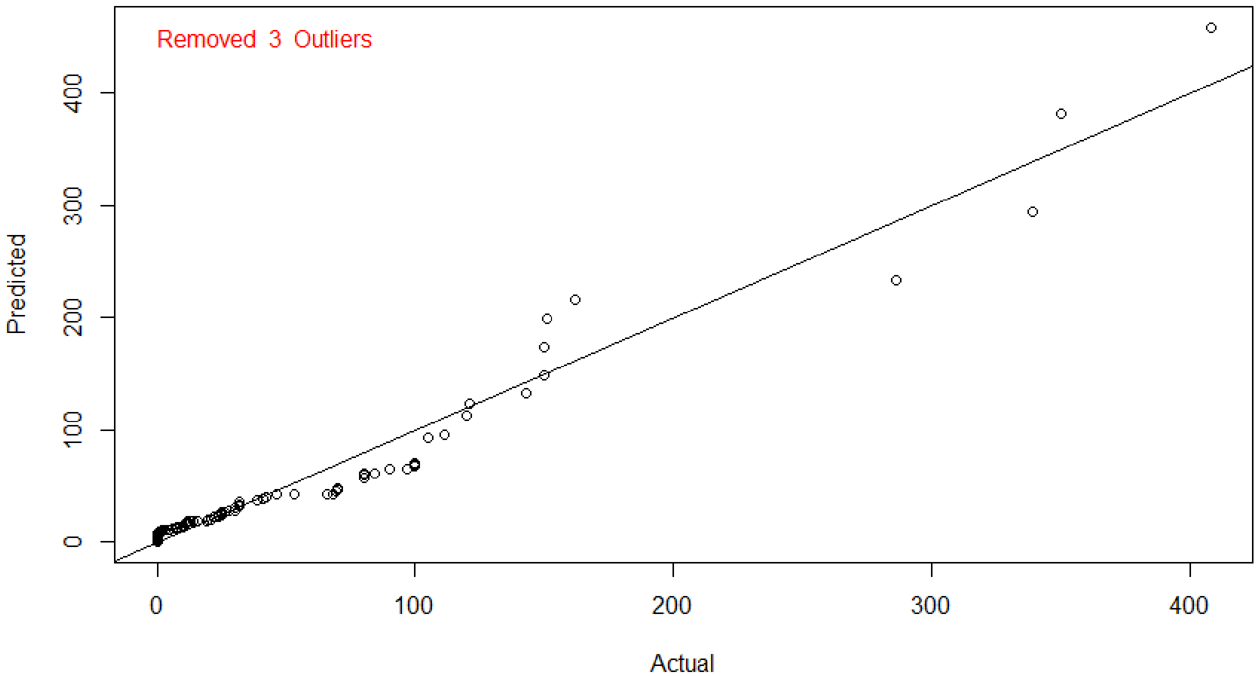


Figure F-1: Model 1 (Negative Binomial #1) Q-Q Plot

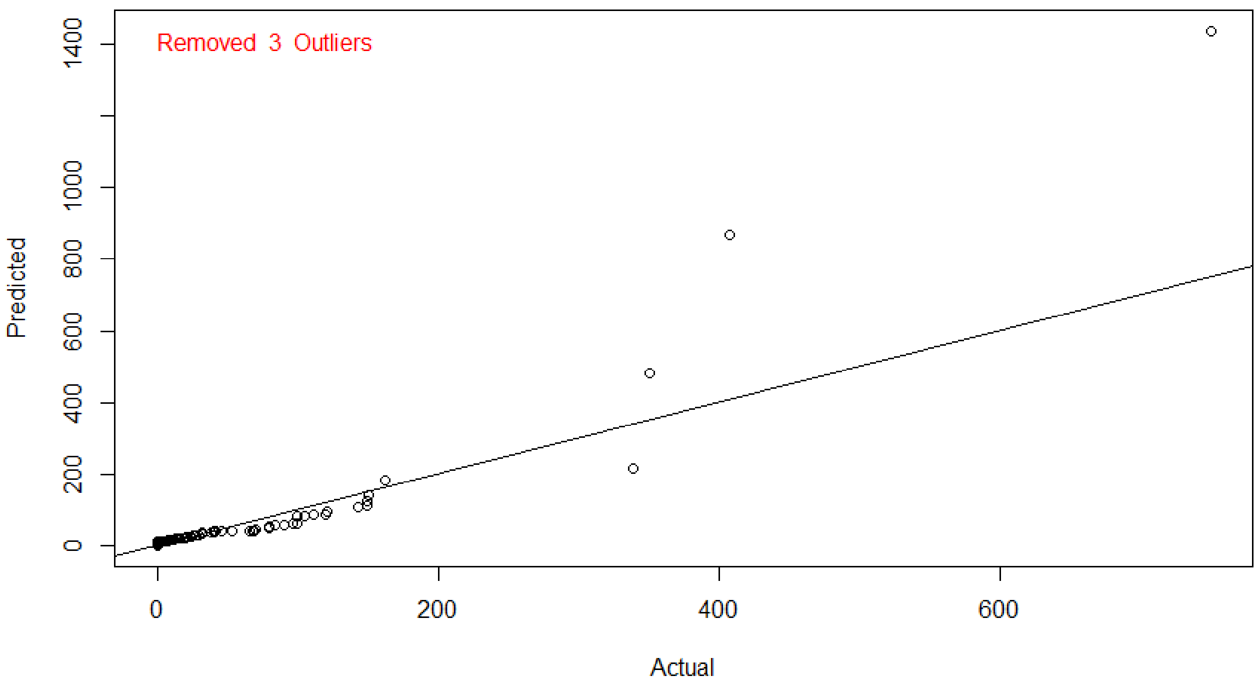


Figure F-2: Model 2 (Negative Binomial #2) Q-Q Plot

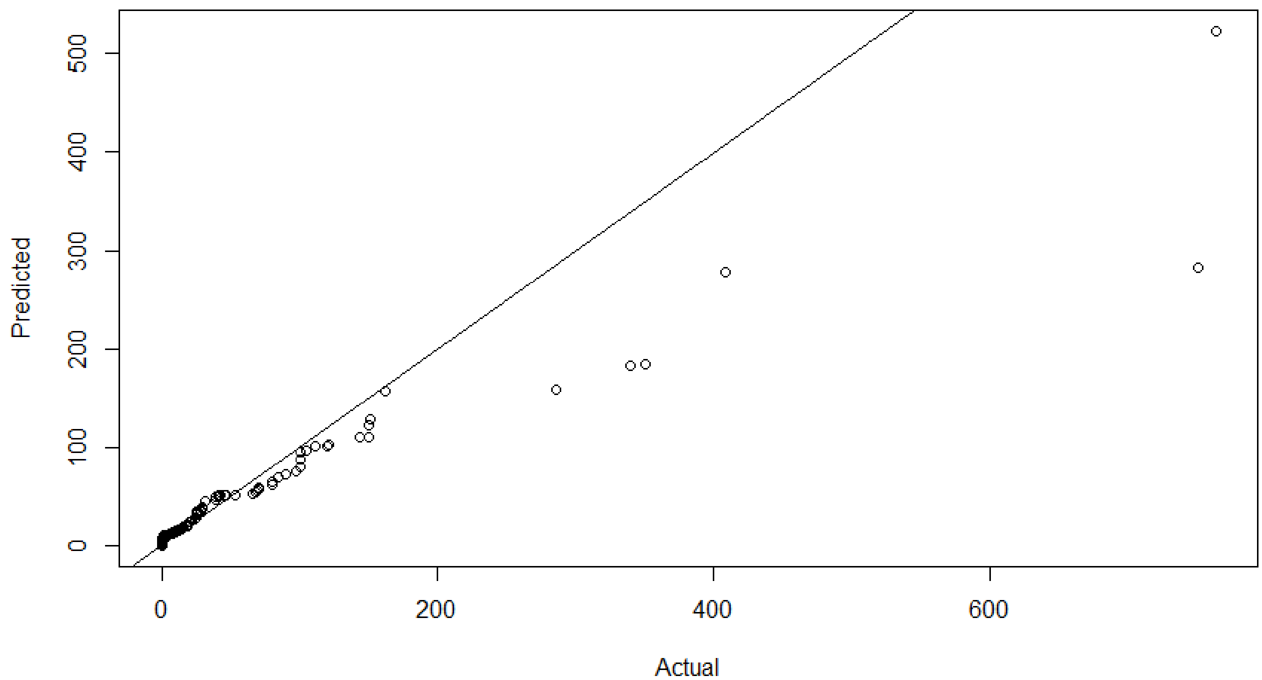


Figure F-3: Model 3 (Negative Binomial #3) Q-Q Plot

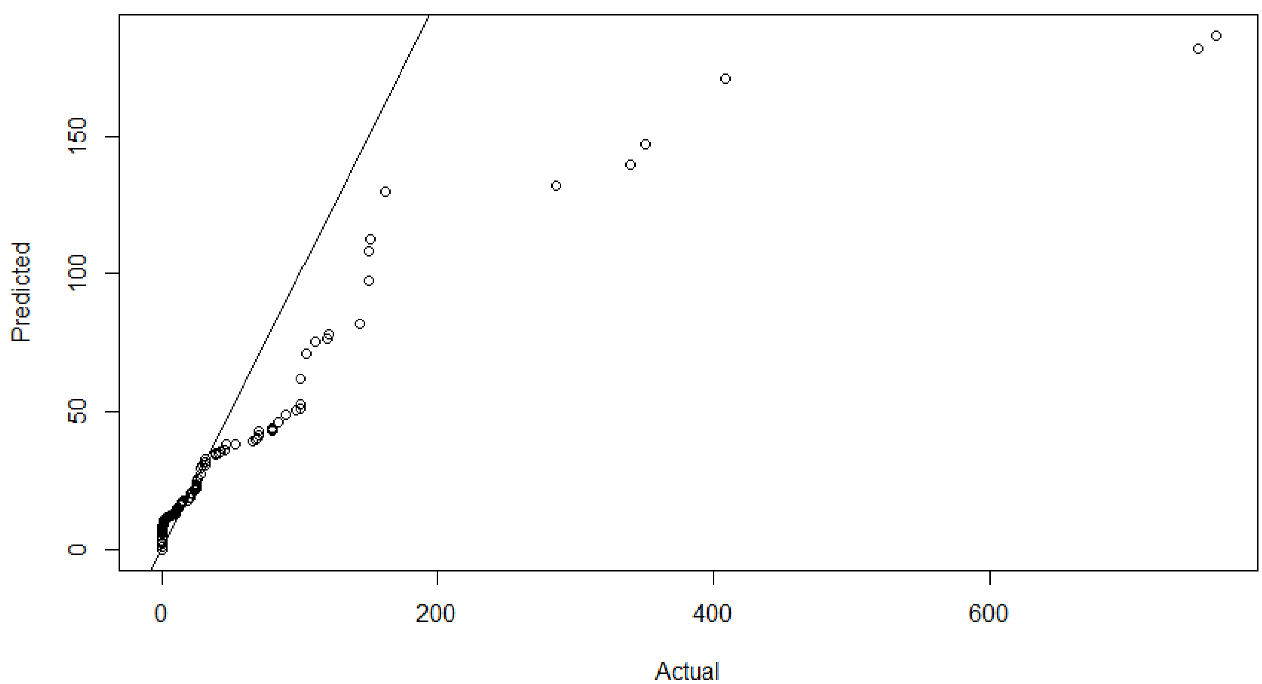


Figure F-4: Model 4 (Zero-Inflated Negative Binomial #1) Q-Q Plot

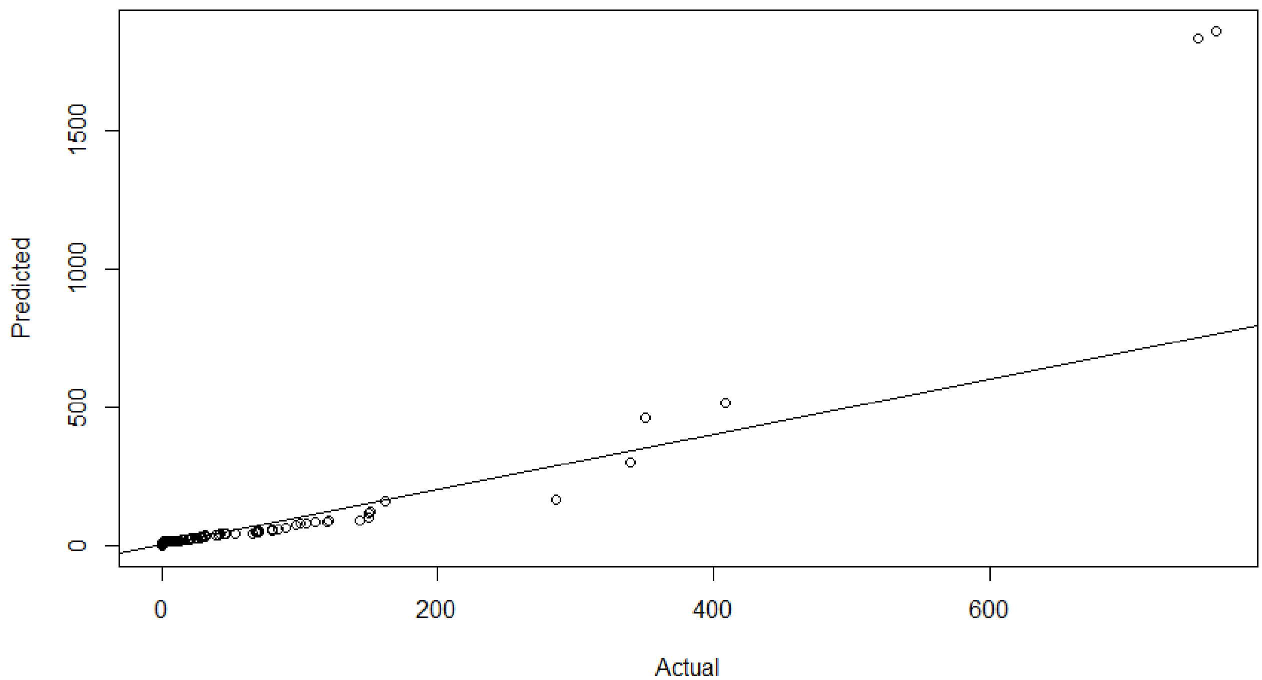


Figure F-5: Model 5 (Zero-Inflated Negative Binomial #2) Q-Q Plot

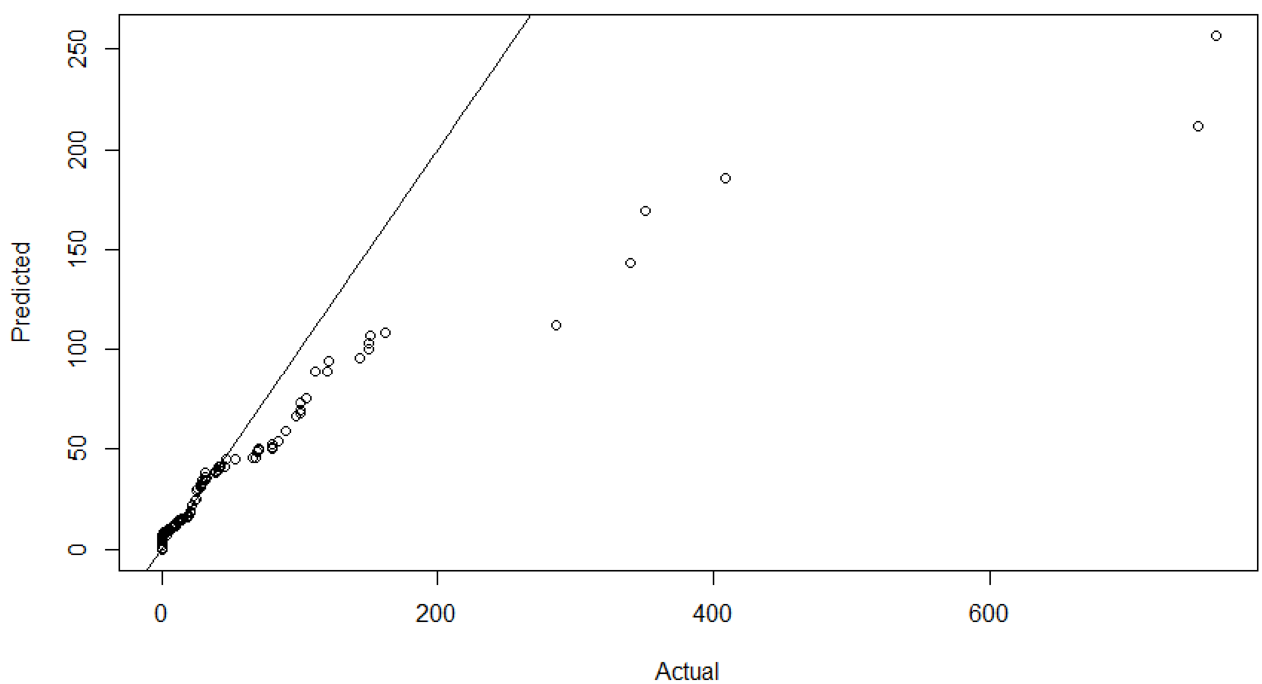


Figure F-6: Model 6 (Zero-Inflated Negative Binomial #3) Q-Q Plot

APPENDIX G: CURE PLOTS

Appendix G.1: Model 1 (Negative Binomial #1)

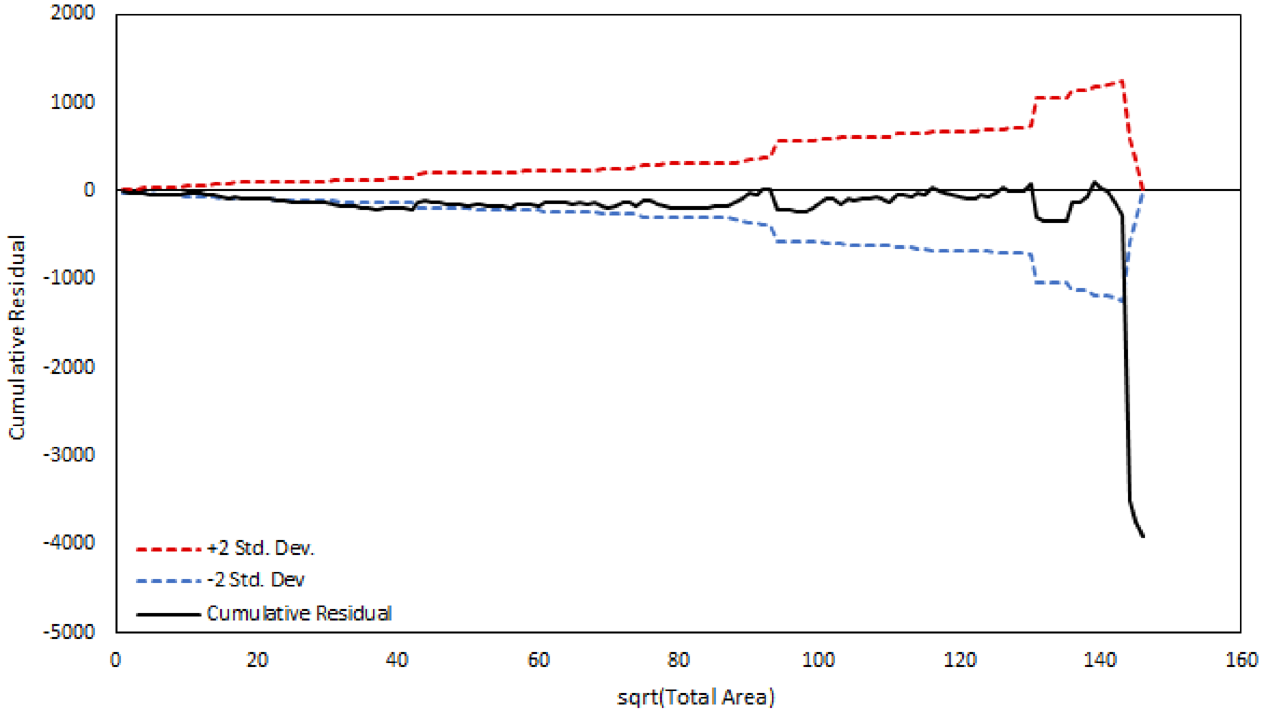


Figure G-1: Model 1 CURE Plot – Total Area

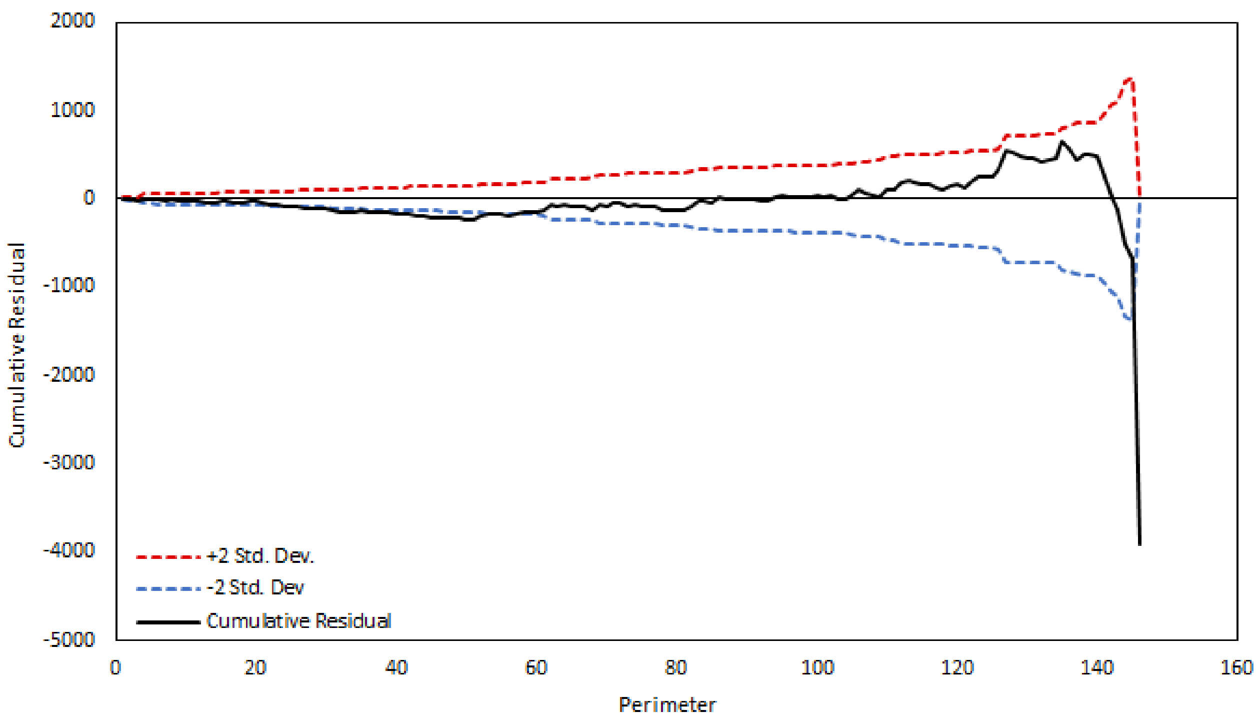


Figure G-2: Model 1 CURE Plot – Perimeter

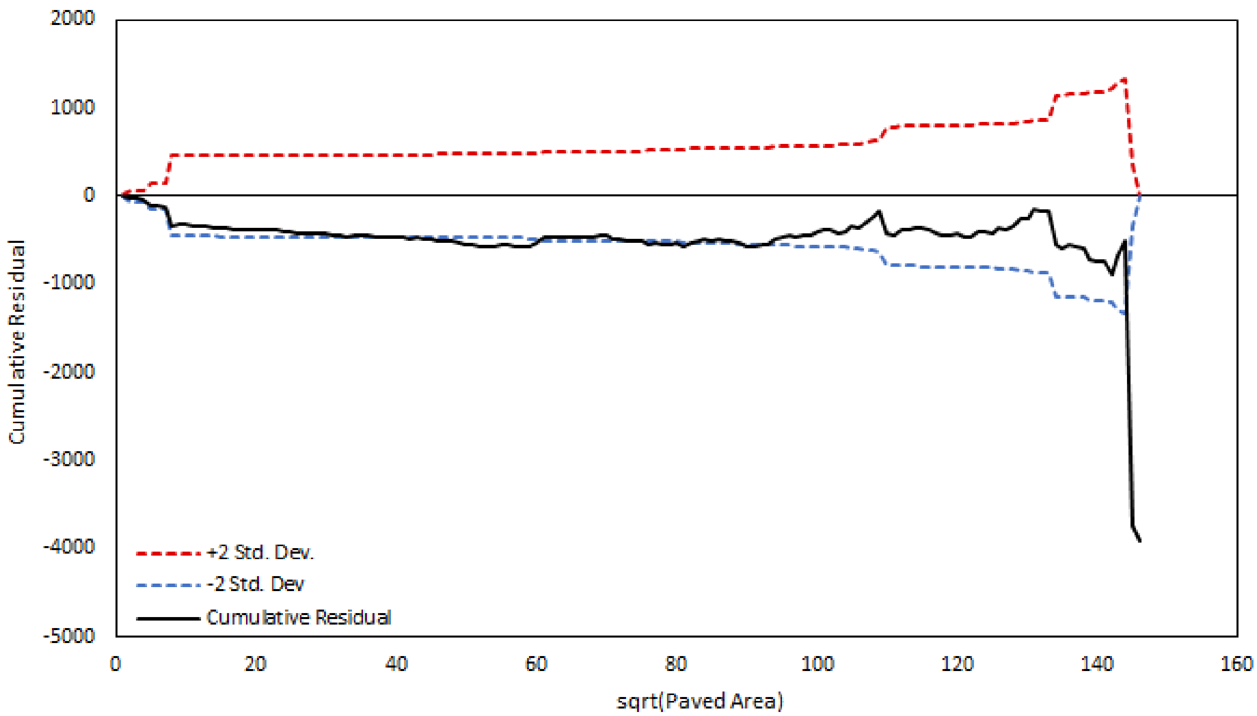


Figure G-3: Model 1 CURE Plot – Paved Area

Appendix G.2: Model 2 (Negative Binomial #2)

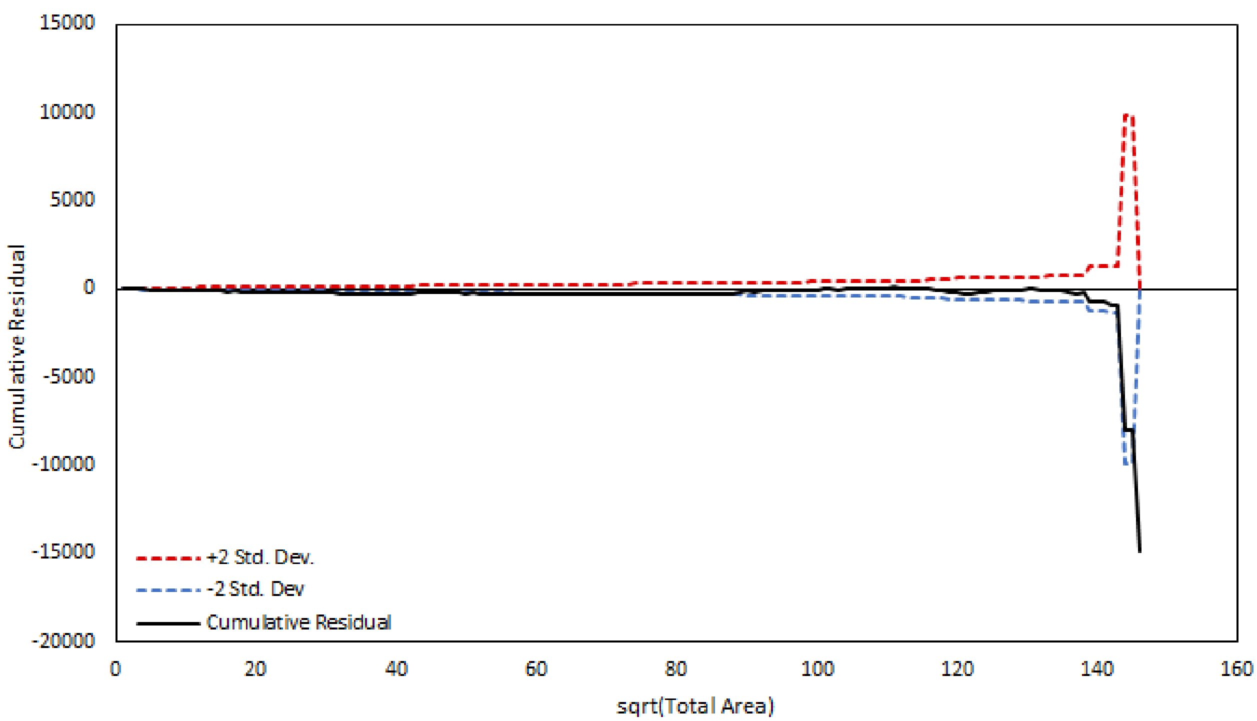


Figure G-4: Model 2 CURE Plot – Total Area

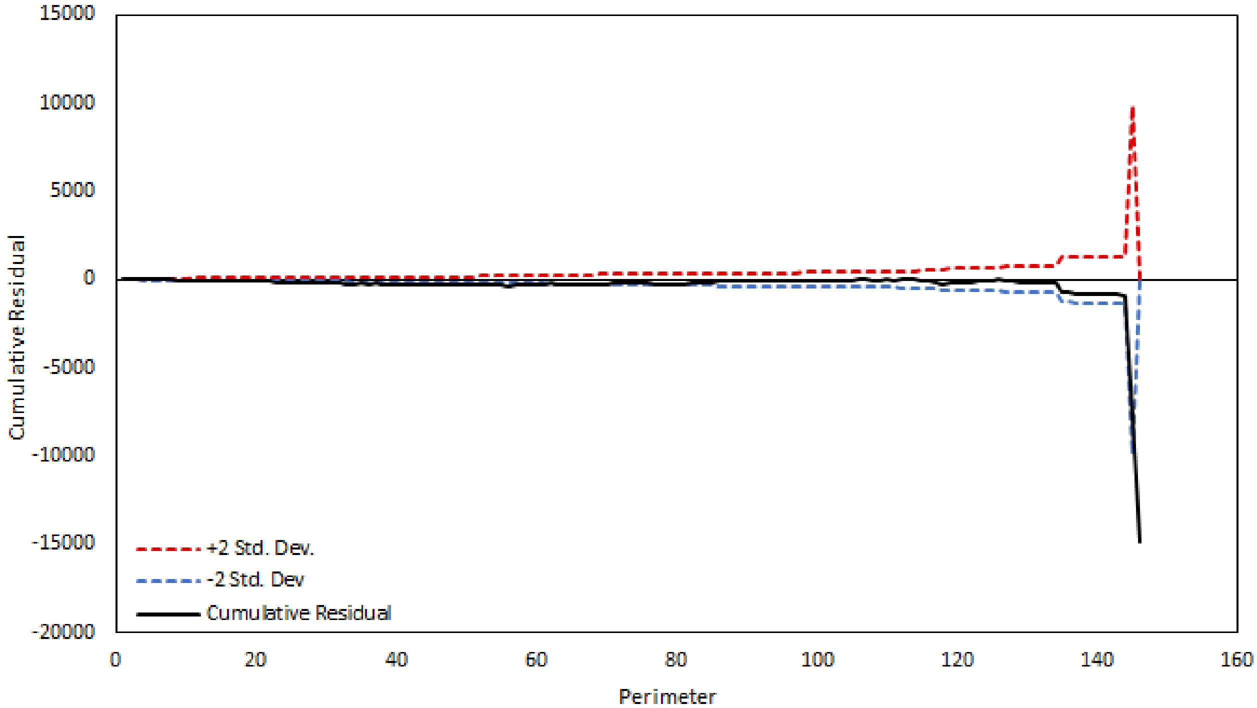


Figure G-5: Model 2 CURE Plot – Perimeter

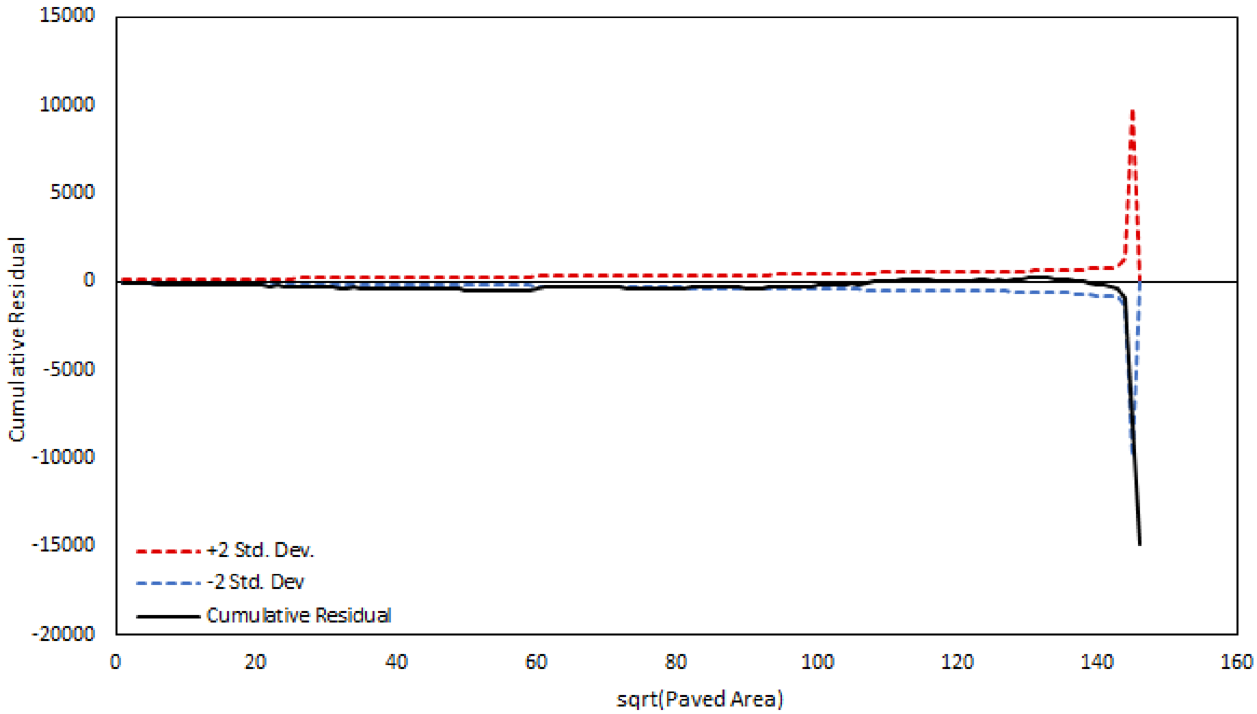


Figure G-6: Model 2 CURE Plot – Paved Area

Appendix G.3: Model 3 (Negative Binomial #3)

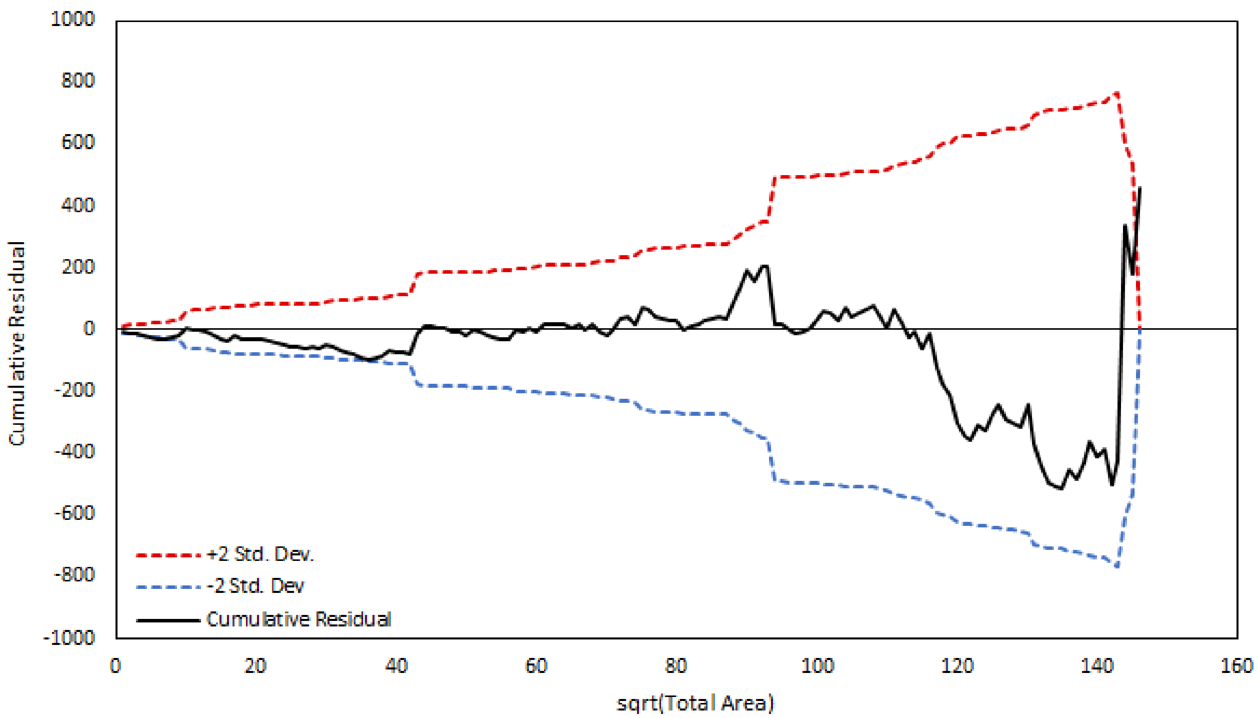


Figure G-7: Model 3 CURE Plot – Total Area

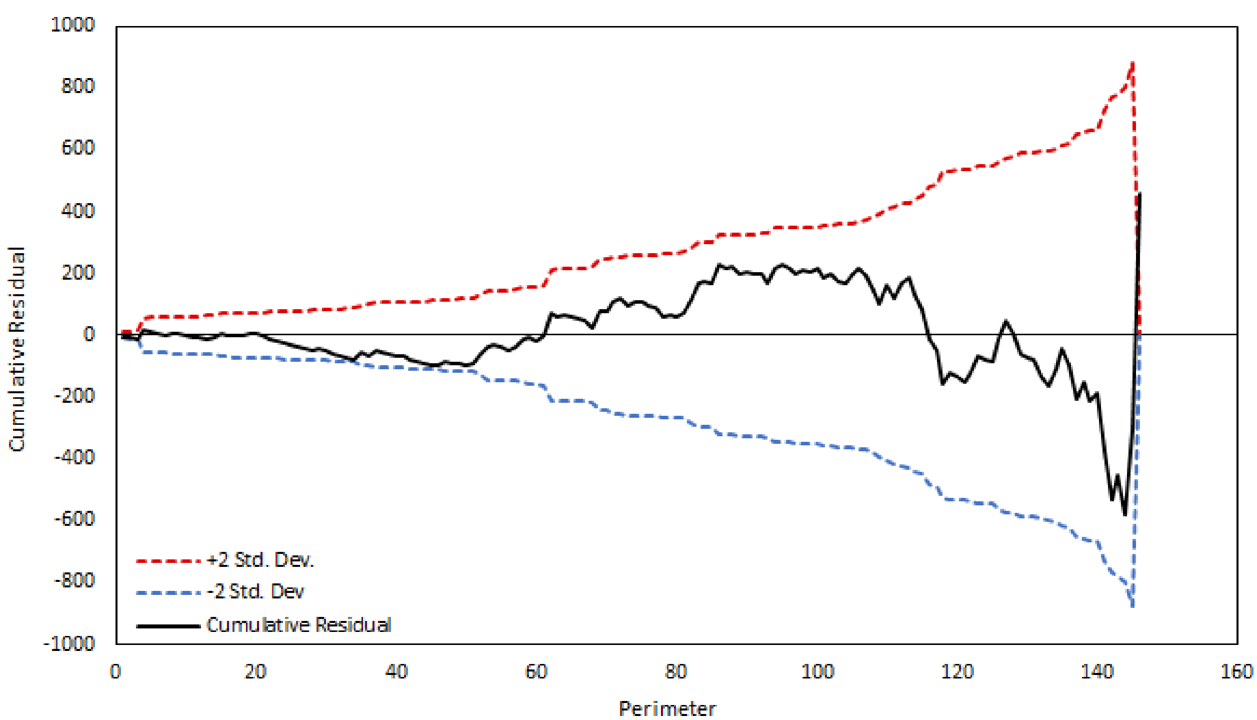


Figure G-8: Model 3 CURE Plot – Perimeter

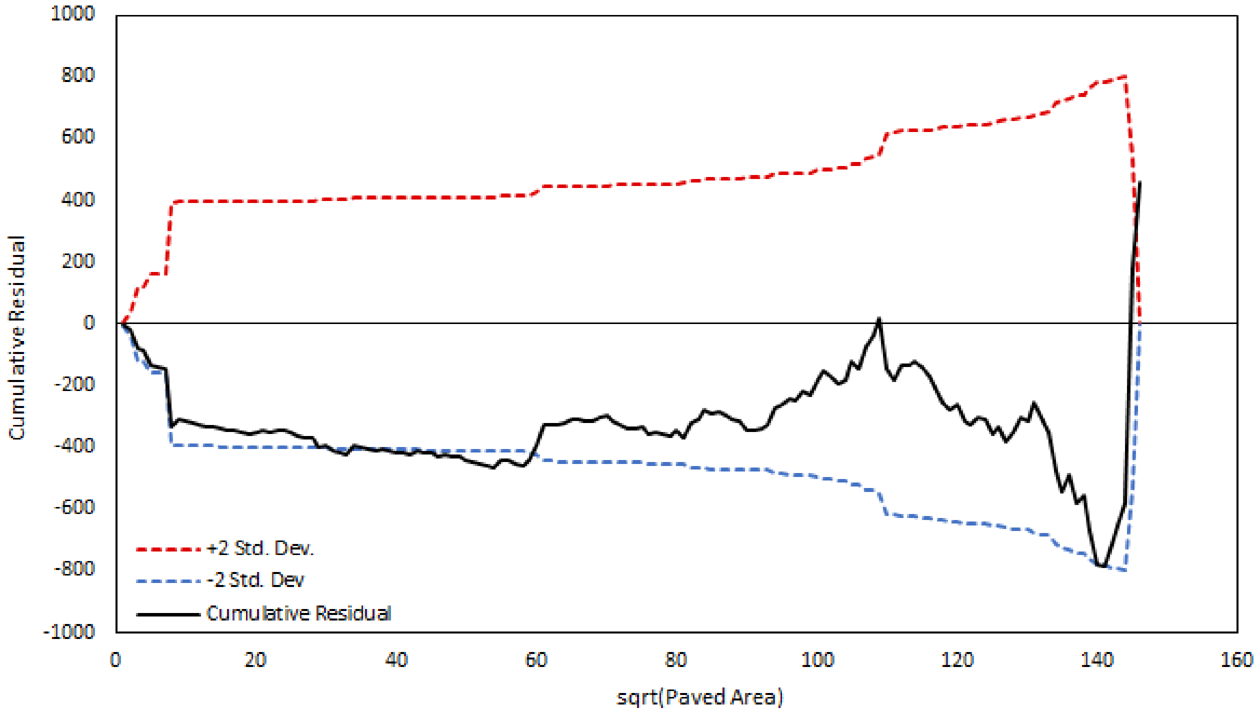


Figure G-9: Model 3 CURE Plot – Paved Area

Appendix G.4: Model 4 (Zero-Inflated Negative Binomial #1)

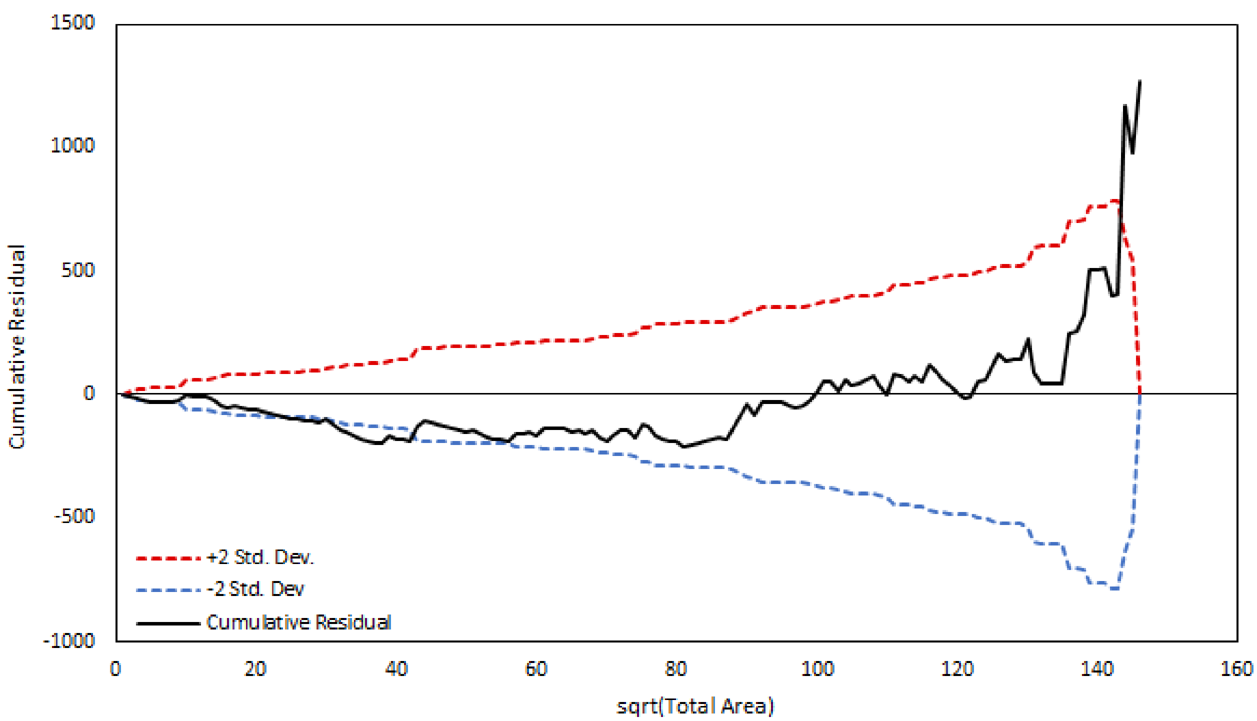


Figure G-10: Model 4 CURE Plot – Total Area

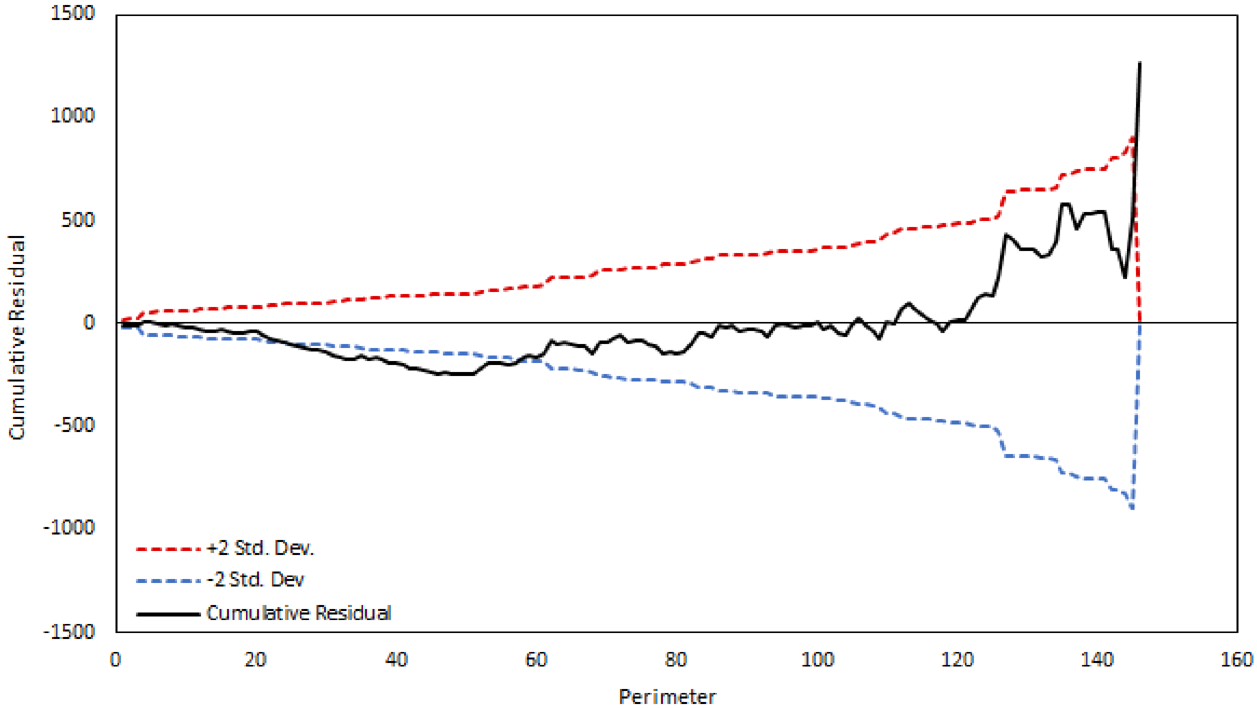


Figure G-11: Model 4 CURE Plot – Perimeter

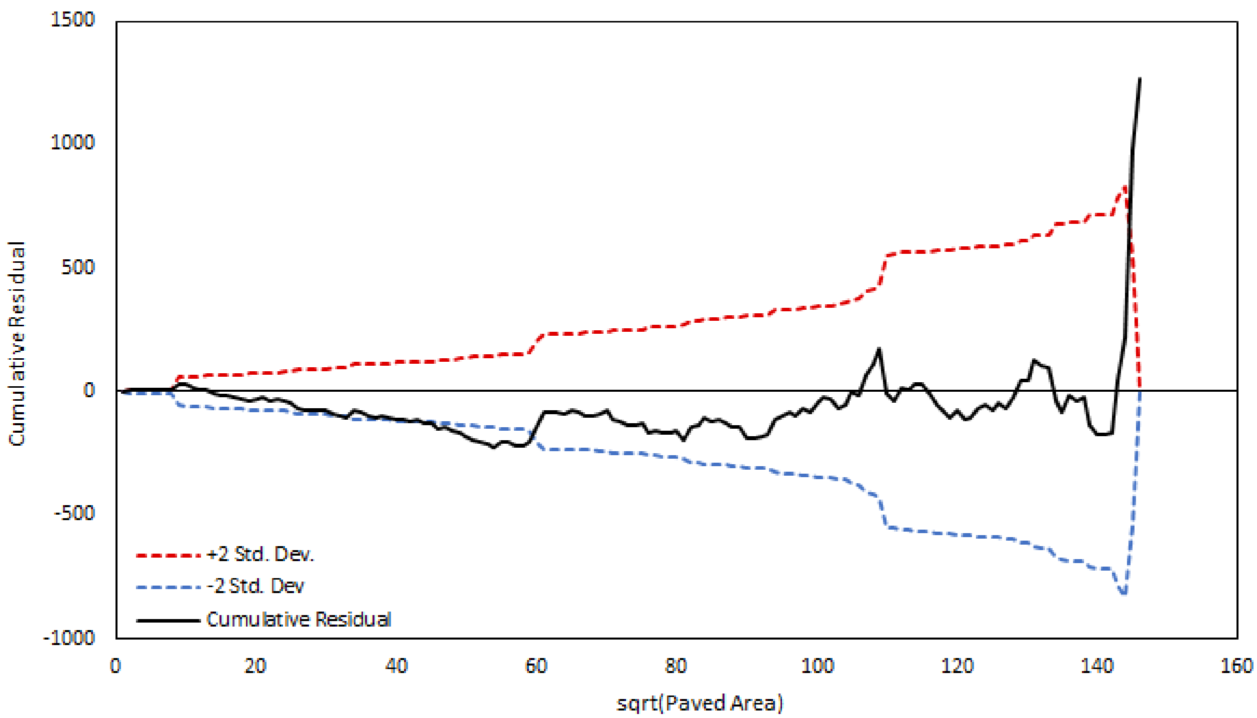


Figure G-12: Model 4 CURE Plot – Paved Area

Appendix G.5: Model 5 (Zero-Inflated Negative Binomial #2)

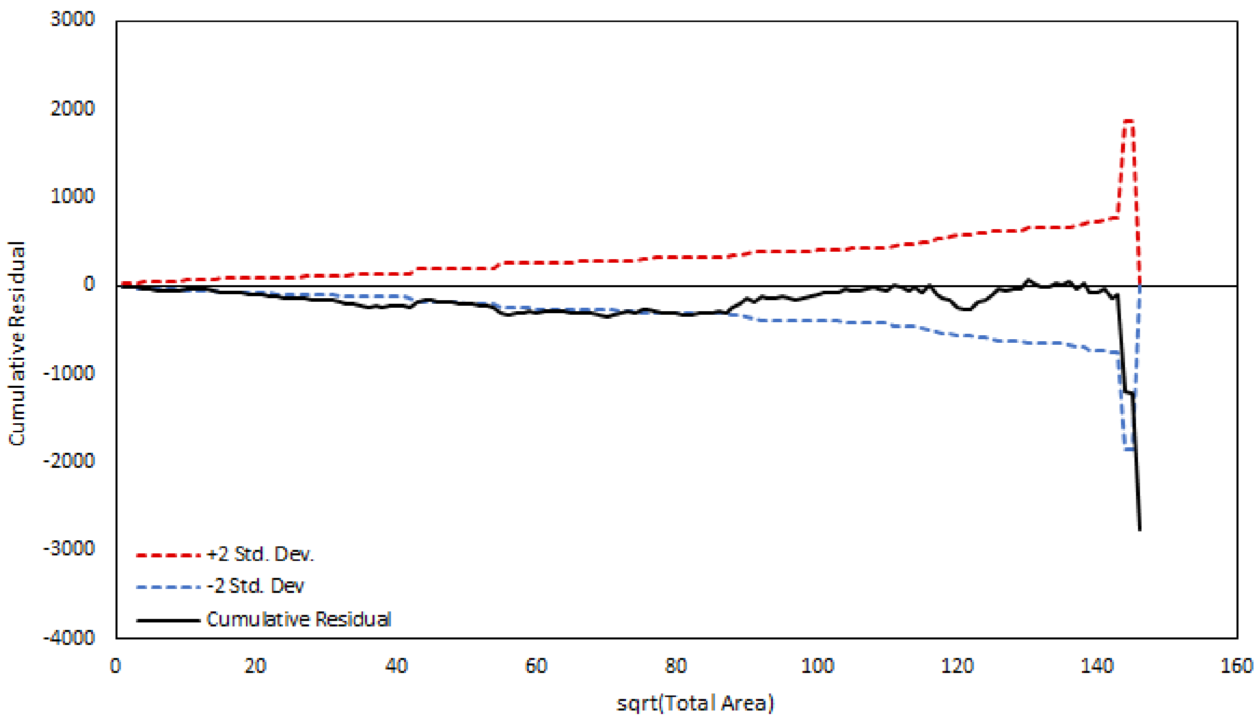


Figure G-13: Model 5 CURE Plot – Total Area

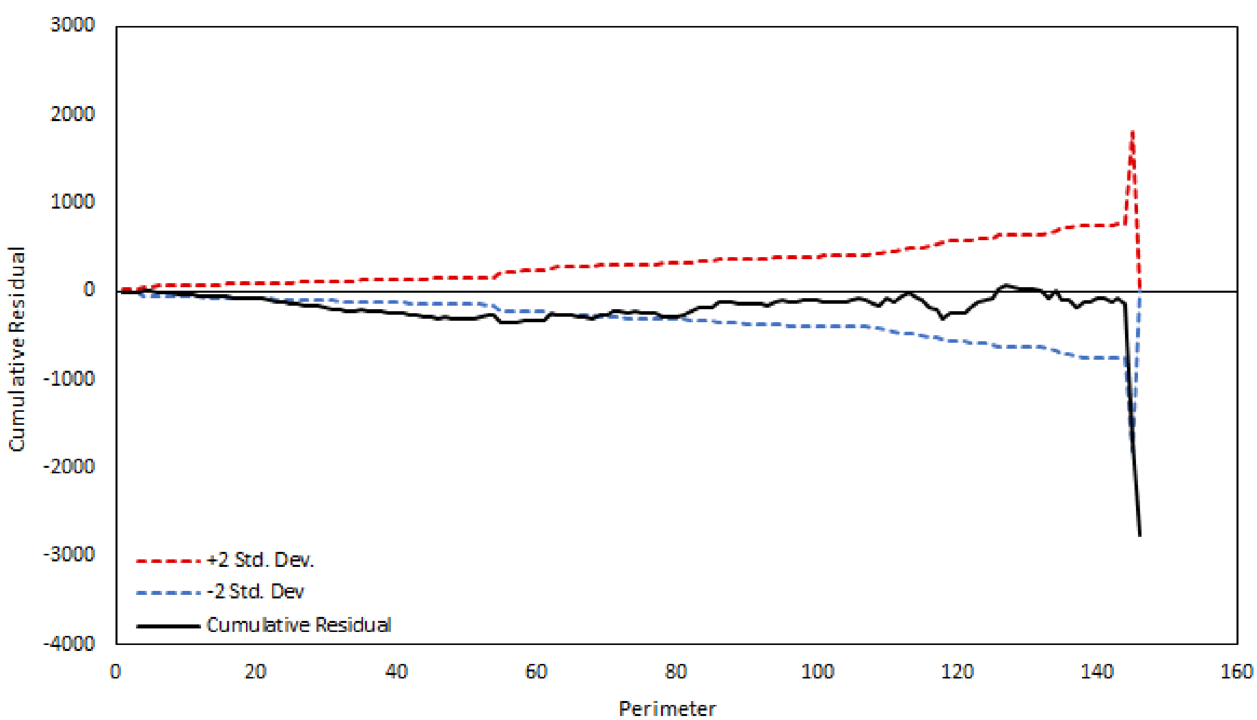


Figure G-14: Model 5 CURE Plot – Perimeter

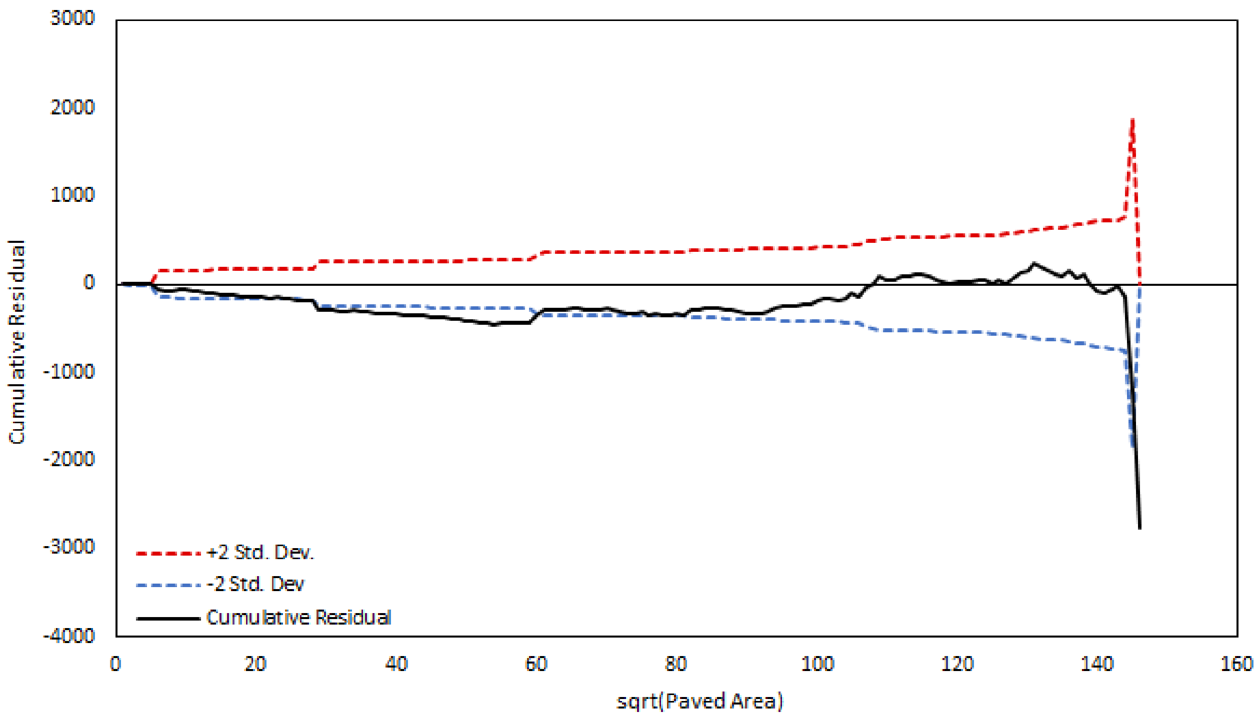


Figure G-15: Model 5 CURE Plot – Paved Area

Appendix G.6: Model 6 (Zero-Inflated Negative Binomial #3)

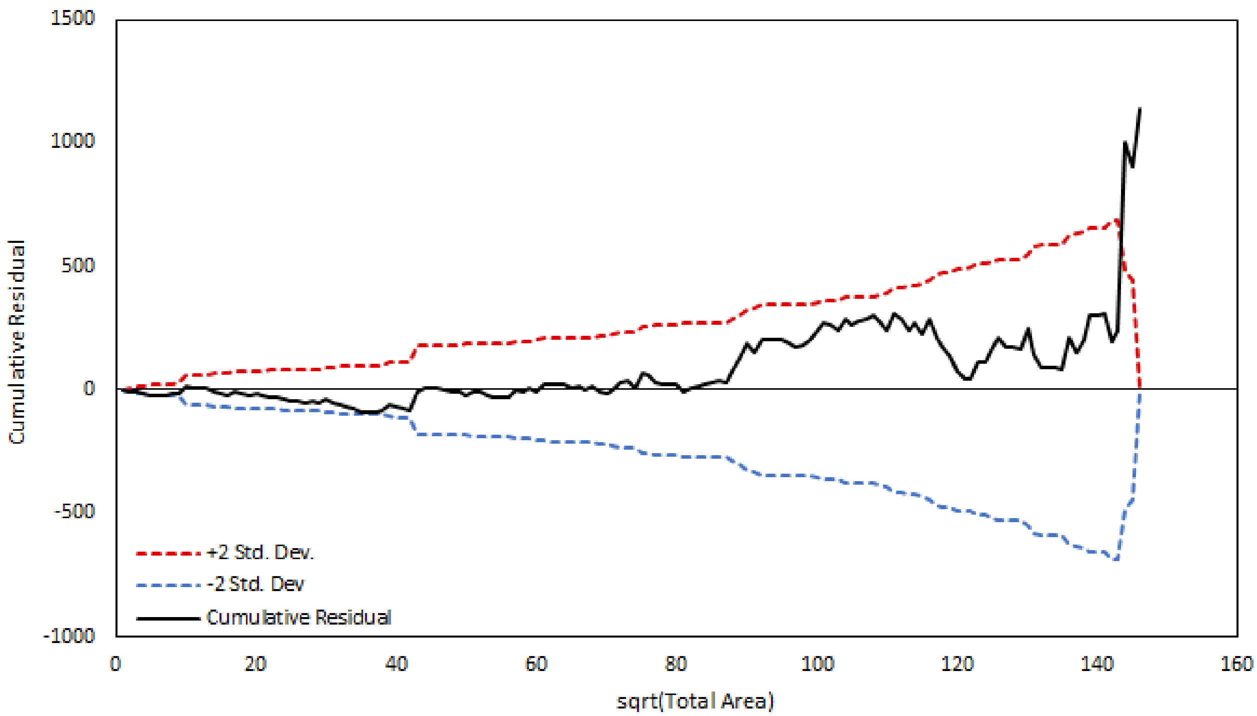


Figure G-16: Model 6 CURE Plot – Total Area

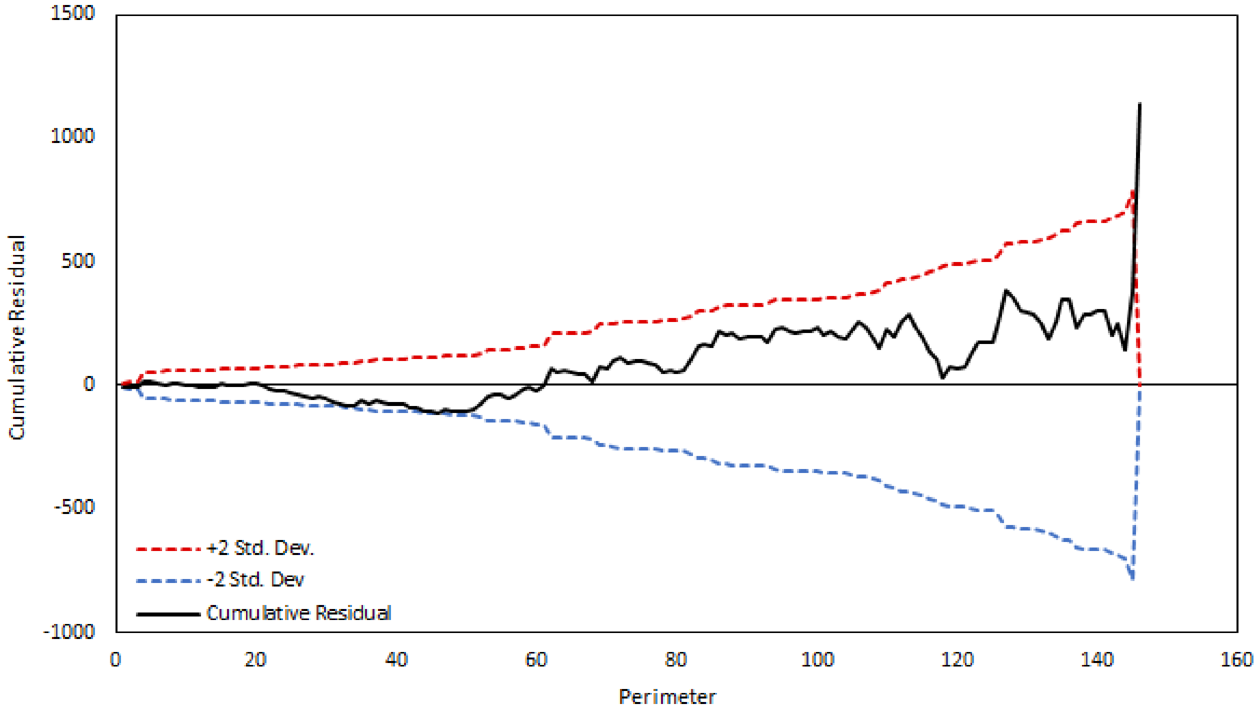


Figure G-17: Model 6 CURE Plot – Perimeter

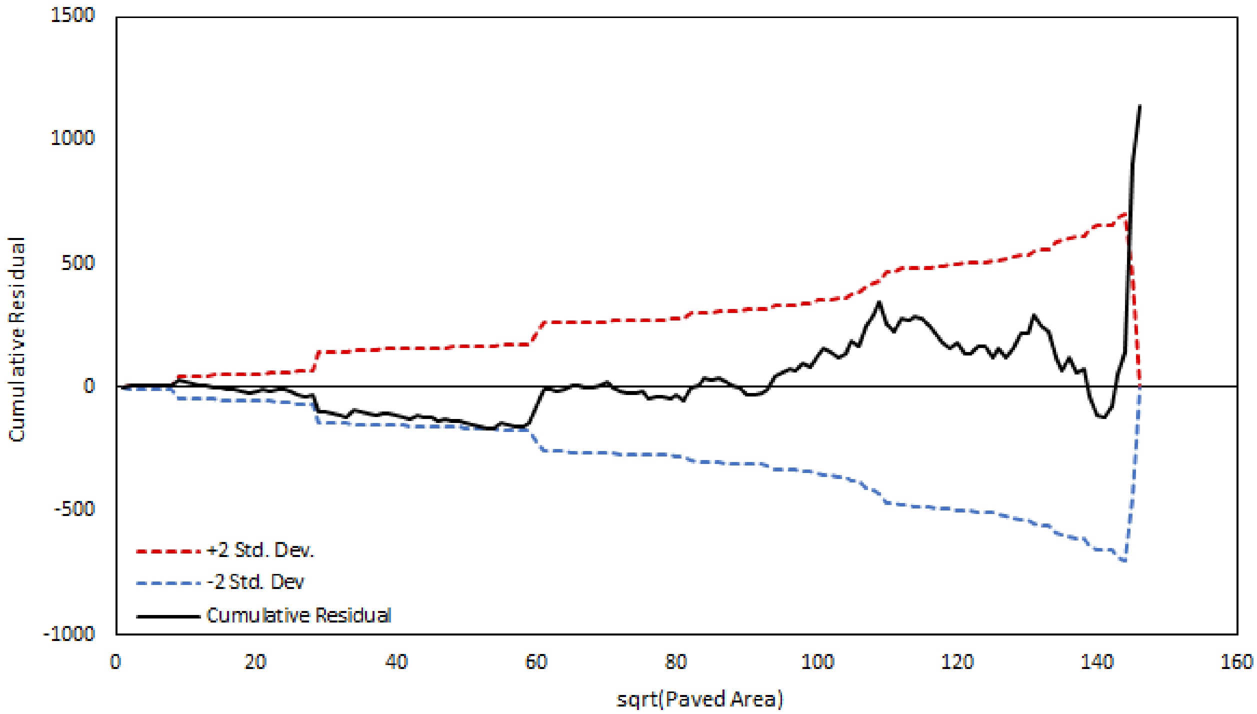


Figure G-18: Model 6 CURE Plot – Paved Area

APPENDIX H: HCV PARKING SUPPLY SPATIAL AUTOCORRELATION

Various tests of global and local spatial autocorrelation were performed on the results of the HCV parking supply estimation. With the exception of Brampton TAZ 7, which indicated a small amount of statistically significant spatial autocorrelation, no spatial autocorrelation was identified in the estimated HCV parking demand. Spatial Autocorrelation was assessed using the software GeoDa (Anselin, 2017).

The global Moran's *I* value was not found to be statistically significant, with a pseudo p-value of 0.189 after 999 iterations. This is shown in Figure H-1.

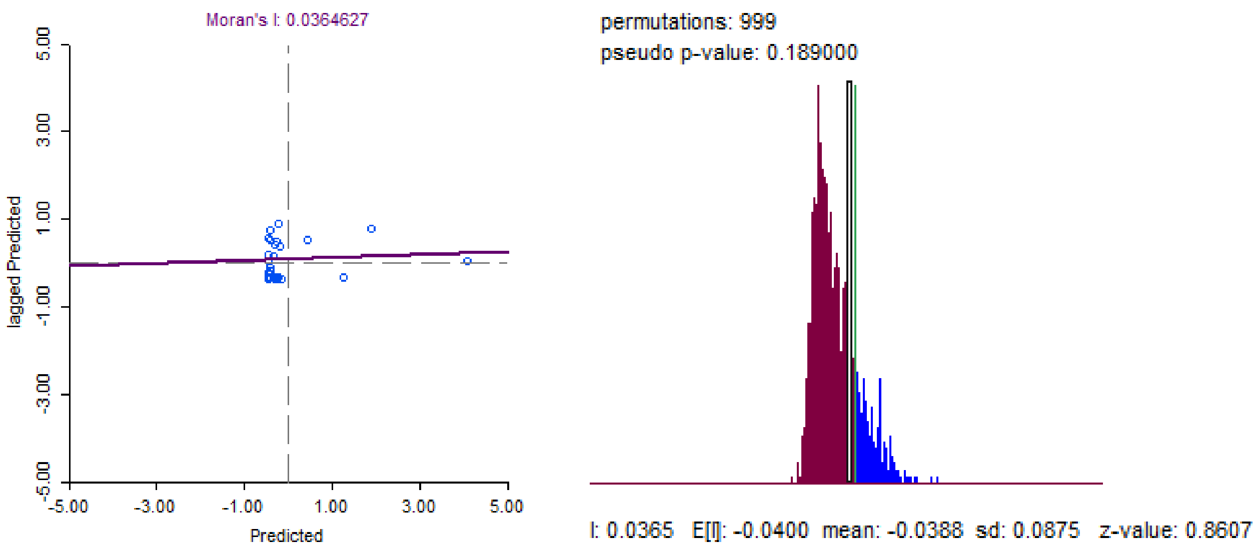


Figure H-1: Global Moran's *I* for HCV Parking Supply Estimates

Several Local Indicators of Spatial Association (LISA) were examined. For the local Moran's *I*, only a Low-High relationship was identified for Brampton TAZ 7 with a p=0.05. This means that this location has a lower than average number of HCV parking spaces with neighbours that have a higher than average number of HCV parking spaces. This is shown in Figure H-2.

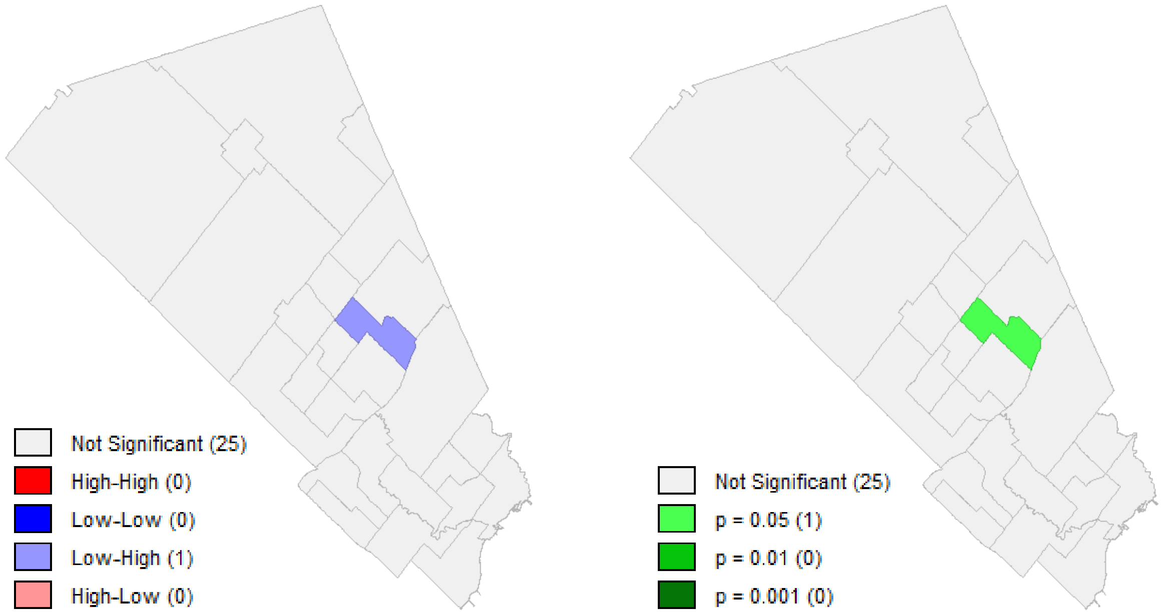


Figure H-2: Local Moran's *I* for HCV Parking Supply Estimates

For the Getis local statistic (G_i), a High relationship was identified for Brampton TAZ 7 with a p=0.05. This means that this TAZ has a high local association with its neighbours compared to the global values. This is shown in Figure H-3.

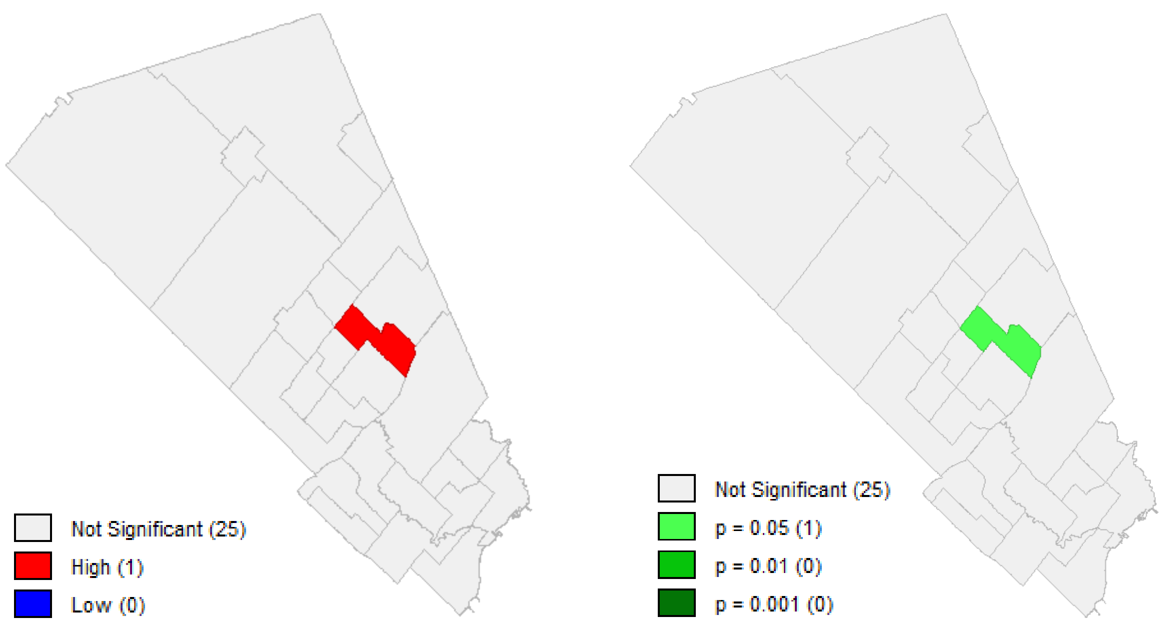


Figure H-3: Getis Local Statistic for HCV Parking Supply Estimates

For the Geary's *c* LISA, a negative relationship was identified for Brampton TAZ 7 with a p=0.05. This means that this TAZ has a statistically significantly lower HCV parking supply compared to its neighbours. This is shown in Figure H-4.

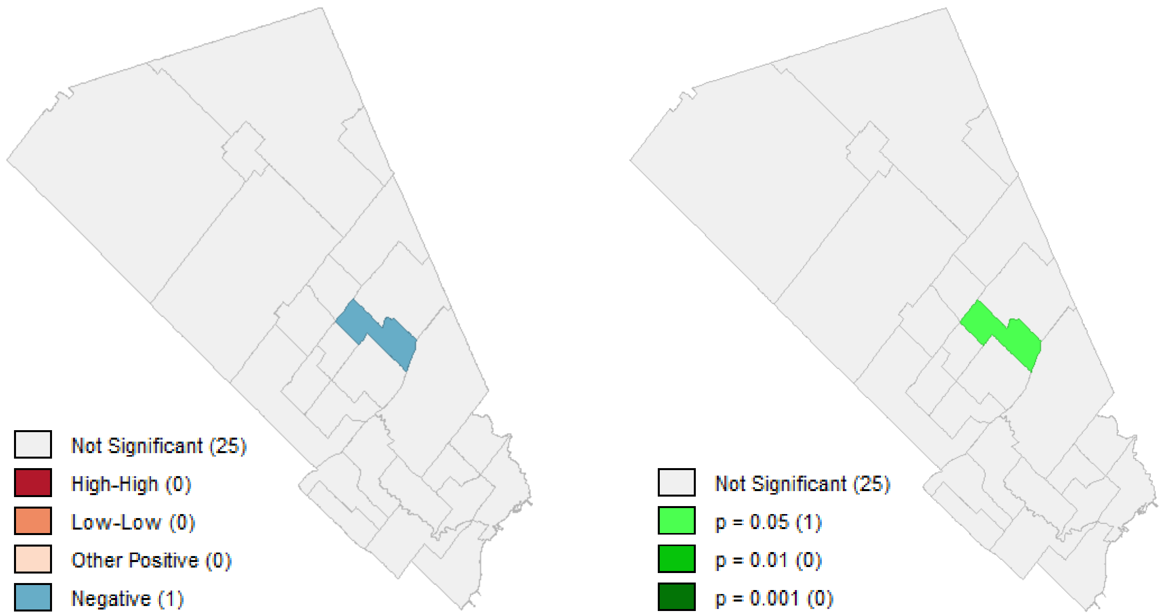


Figure H-4: Geary's *c* for HCV Parking Supply Estimates

The weight matrix used for this analysis followed the Queen rule with an order of contiguity of one. The weight matrix connectivity graph, connecting neighbours, is shown in Figure H-5.

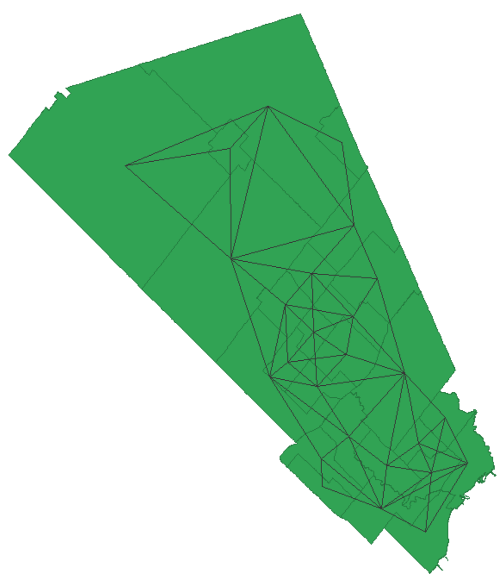


Figure H-5: Weight Matrix Connectivity Graph

VITA AUCTORIS

Name:	Erik Alexander Nevland, EIT
Place of Birth:	Prince Albert, Saskatchewan
Year of Birth:	1992
Education:	Certificate in Professional Communication (2017) University of Saskatchewan: Graham School of Professional Development Bachelor of Science in Engineering (2018) University of Saskatchewan: College of Engineering Master of Applied Science in Civil Engineering (2020) York University: Lassonde School of Engineering
Presentations:	Nevland, E. A., Park, P. Y., Hanson, R., Klashinsky, R. and Der, T. (2015) 'Utilizing Weigh-In-Motion for Integrated Average Speed and Weight Enforcement', in *2015 Transportation Association of Canada Conference & Exhibition*. Charlottetown, PE. Nevland, E. A., Gingerich, K. and Park, P. Y. (2019) 'Development of a Classification Scheme and a Supply Model for Inter-Regional Truck Parking Facilities', in *98th Annual Meeting of the Transportation Research Board*. Washington, DC. Nevland, E. A., Gingerich, K. and Park, P. Y. (2019) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Canadian Institute of Transportation Engineers Annual Conference*. Ottawa, ON. Wang, M., Nevland, E. A. and Park, P. Y. (2019) 'Average Speed Enforcement for Heavy Vehicles by Weight Classification: Simulation Results for the Laidlaw-Golden Freight Corridor', in *2019 Canadian Institute of Transportation Engineers Annual Conference*. Ottawa, ON. Nevland, E. A., Gingerich, K. and Park, P. Y. (2019) 'Parking Classification and Supply Modelling for Inter-Regional Truck Trips', in *2019 Joint Transportation Association of Canada / Intelligent Transportation Systems Canada Conference and Exhibition*. Halifax, NS. [*Presented by Gingerich and Park*]
Publications:	Gingerich, K. and Nevland, E. A. (2019) 'Effectiveness of Student-Led Co-Curricular Software Learning in STEM Education', *Applications of LinkedIn Learning in Ontario's Post-Secondary Institutions*.
Employment:	Transportation Planner (2019-Present) Region of Peel: Transportation System Planning
Volunteerism:	Founding President (2018-19) ITE York University Student Chapter

Committee Member-At-Large (2019-20)
Young Professionals in Transportation Toronto Chapter

Communications Coordinator (2019-20)
ITE Toronto Section

Vice President Education (2020)
Region of Peel Toastmasters Club

Awards:

Outstanding Leadership Award (2020)
ITE York University Student Chapter

11th Annual Joint CITE Section Student Presentation Competition Winner (Graduate Student Category) (2019)
ITE Toronto, Hamilton, and Southwest Ontario Sections

York University President's Sustainability Leadership Award (2019)
York University President's Sustainability Council
awarded to the ITE York University Student Chapter during presidency

CITE Student Chapter Delta Activity Award (2019)
Canadian Institute of Transportation Engineers
awarded to the ITE York University Student Chapter during presidency