Modeling Optimal Freight Logistics Facility Locations Based on The Clustering of Industries and Truck Trip Patterns in Ontario

Ayat Hussein
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
Hussein, Ayat, "Modeling Optimal Freight Logistics Facility Locations Based on The Clustering of Industries and Truck Trip Patterns in Ontario" (2020). Electronic Theses and Dissertations. 8369.
https://scholar.uwindsor.ca/etd/8369

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
MODELING OPTIMAL FREIGHT LOGISTICS FACILITY LOCATIONS BASED ON THE CLUSTERING OF INDUSTRIES AND TRUCK TRIP PATTERNS IN ONTARIO

By

Ayat Hussein

A Thesis
Submitted to the Faculty of Graduate Studies through the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science at the University of Windsor

Windsor, Ontario, Canada

2020

© 2020 Ayat Hussein
MODELING OPTIMAL FREIGHT LOGISTICS FACILITY LOCATIONS BASED ON THE CLUSTERING OF INDUSTRIES AND TRUCK TRIP PATTERNS IN ONTARIO

by

Ayat Hussein

APPROVED BY:

______________________________________________
J. Tofflemire
Department of Civil and Environmental Engineering

______________________________________________
W. Anderson
Department of Political Science

______________________________________________
C. Lee
Department of Civil and Environmental Engineering

______________________________________________
H. Maoh, Advisor
Department of Civil and Environmental Engineering

May 8th, 2020
DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
ABSTRACT

Freight transportation plays a major role in determining the economic health of a region. Efficient freight transportation systems are typically associated with reducing the cost of moving goods from and to logistics facilities. Understanding the clustering pattern of truck trip ends (i.e., productions and attractions) can help optimize the location of such facilities over space. This thesis explores freight activities, using data from a large sample of trucks owned by Canadian carriers for the month of September 2014, to propose ways to optimize the location of future logistics facilities in Ontario. Heat maps using the kernel density estimation method are generated to highlight the clustering of these trips by industry. Besides the exploratory work, multi-criteria decision analysis is performed to create a suitability surface to identify potential locations where new logistics facilities could be established. A total of 18 potential locations across Ontario are identified and used to execute a number of location-allocation scenarios. The ArcGIS 10.6 software and its extensions (namely Spatial Analyst and Network Analyst) are heavily utilized in the analysis to create the kernel density maps, suitability surface, and the Location-Allocation modeling work. The results indicate that Hamilton, Ontario would be the most optimal location for establishing a future logistics facility to complement the operations occurring in the Peel region. When factoring the Canada-US border, Windsor, Ontario can be considered the second most optimal location after Hamilton. The conducted analysis allows us to see the optimized locations for new logistics facilities to service local markets around Ontario as well as US markets serviced by the Ontario-US land border crossings.
DEDICATION

I dedicate this thesis to my beloved parents, Dr. Riyad and Suham Hussein, for always believing in me and pushing me to be a better version of myself.
ACKNOWLEDGEMENTS

First and foremost, I would like to express my deeply felt gratitude to Dr. Hanna Maoh for giving me the chance to start my MASc under his supervision. I am thankful for his constant guidance and encouragement. His passion in the field of transportation engineering and his hard-working attitude inspired me to do my best. Our enlightening conversations about transportation research helped me think more conceptually, critically and quantitatively for which I am grateful. In addition, I am also extremely thankful for the extra time Dr. Maoh spent to help me prepare and submit conference proceeding, presentations, and scholarship applications.

I would like to thank my committee members Dr. Chris Lee, Dr. William Anderson and Mr. John Tofflemire for taking the time to review my work and provide insightful comments which ultimately strengthened my thesis. Also, I would like to thank Dr. Ruparathna for chairing my MASc defense. I especially thank Dr. Shakil Khan for his expertise in managing the Traffic Research Lab, where I worked on my research, and for always helping me when I had troubles with my workstation.

Lastly, I must express my very profound appreciation to my parents and siblings for providing me with unfailing support and continuous encouragement. Without their love and support, this accomplishment would not have been possible. I am also thankful for my best friends Nariman Marqus, Shahad Bahi, Sally Joseph, and Shery Gergis for listening to my complaints and for their constant encouragement. Thank you!
TABLE OF CONTENTS

DECLARATION OF ORIGINALITY ............................................................... iii

ABSTRACT .............................................................................................. iv

DEDICATION ........................................................................................... v

ACKNOWLEDGEMENTS .......................................................................... vi

LIST OF TABLES ..................................................................................... x

LIST OF FIGURES .................................................................................. xi

Chapter 1: INTRODUCTION ..................................................................... 1
   1.1 Overview .......................................................................................... 1
   1.2 Research Objectives ......................................................................... 4
   1.3 Thesis Outline .................................................................................. 5

CHAPTER 2: LITERATURE REVIEW ......................................................... 6
   2.1 Analysis of Truck Movements ........................................................... 6
   2.2 Clustering Analysis ......................................................................... 9
      2.2.1 Logistics Clusters .................................................................... 9
      2.2.2 Trip Origin and Destination Clustering .................................. 11
      2.2.3 Types of Relationships among cluster members .................. 13
      2.2.4 Advantages of Logistics Clustering ....................................... 14
   2.3 Supply Chain and Logistics ............................................................. 15
   2.4 Location of Logistics Terminals ...................................................... 18
      2.4.1 Optimality of Locations ......................................................... 18
      2.4.2 Site Suitability ....................................................................... 19

CHAPTER 3: METHODS OF ANALYSIS .................................................. 21
   3.1 Study Area ..................................................................................... 21
   3.2 Data .............................................................................................. 23
      3.2.1 Truck GPS Trips ................................................................. 23
3.2.2 Population and Jobs ......................................................................................... 24
3.2.3 Commercial Vehicle Survey (CVS) Data ......................................................... 24
3.2.4 Geographic Information Systems (GIS) Data .................................................... 25
3.3 Data Exploration Techniques ............................................................................... 26
  3.3.1 Regression Analysis ......................................................................................... 27
  3.3.2 Kernel Density Analysis .................................................................................. 28
3.4 Network Modeling Techniques .......................................................................... 30
  3.4.1 Raster Data Analysis ...................................................................................... 32
  3.4.2 Multi-Criteria Decision Analysis (MCDA) ..................................................... 34
  3.4.3 Analytical Hierarchical Pairwise (AHP) .......................................................... 44
  3.4.4 Location-Allocation Modeling ...................................................................... 46

CHAPTER 4: RESULTS AND DISCUSSION ................................................................. 49
  4.1 Data Exploration ................................................................................................. 49
    4.1.1 Data Validation .............................................................................................. 49
    4.1.2 Trucking Trips Intensity ............................................................................... 52
    4.1.3 Firm Intensity ............................................................................................... 58
    4.1.4 Regression Analysis ...................................................................................... 61
    4.1.5 Border Trips .................................................................................................. 63
  4.2 Modeling Results ................................................................................................. 79
    4.2.1 Multi Criteria Decision Analysis (MCDA) Results: .................................... 79
    4.2.2 SAW Model Results .................................................................................... 82
    4.2.3 Location-Allocation Modeling Results ......................................................... 84

CHAPTER 5: CONCLUSION ......................................................................................... 93
  5.1 Clustering Patterns .............................................................................................. 94
  5.2 Logistics Facility Locations ............................................................................... 95
  5.3 Research Contributions ...................................................................................... 97
  5.4 Research Limitations and Future Research ....................................................... 98

REFERENCES ............................................................................................................. 99

APPENDICES ............................................................................................................ 105
  Appendix A: AHP Pairwise Evaluation ................................................................. 105
Appendix B: Truck Trip Productions Heat Maps per Industry ................................... 108
Appendix C: Truck Trip Attractions Heat Maps Per Industry ................................... 112
Appendix D: Rudimentary Location-Allocation Model .................................................. 114

VITA AUCTORIS ................................................................................................................ 115
# LIST OF TABLES

Table 3-1 GIS Data Layers .......................................................................................... 26
Table 3-2 Proposed MCDA Criteria ............................................................................. 36
Table 3-3 AHP Pairwise Comparison Matrix ............................................................... 45
Table 4-1 The Correlation Matrix Between Trip Productions and Trip Attractions Per Industry .................................................................................................................. 54
Table 4-2 Correlation Coefficients Between Truck Trip Generations and Jobs By Industry ......................................................................................................................... 60
Table 4-3 Top 10 Cities by Trip Productions ............................................................... 60
Table 4-4 Top 10 Cities by Jobs .................................................................................... 61
Table 4-5 Top 10 Cities by Firms .................................................................................. 61
Table 4-6 Parameter Estimates of Trip Generation Model ........................................... 63
Table 4-7 Trip Distances Along the Borders ................................................................. 64
Table 4-8 The Pair Wise Comparison Matrix of the Eight Considered Criteria .... 80
Table 4-9 Standardized Criteria Weight ....................................................................... 82
LIST OF FIGURES

Figure 3-1 Ontario Zones Presenting The Study Area Of The Thesis ............... 22

Figure 3-2 Kernel Density Estimation ................................................................. 29

Figure 3-3 3d Quartic Density Function ................................................................. 29

Figure 3-4 Network Analysis Process ................................................................... 31

Figure 3-5 Illustration Of Raster Data Representation Of Real-World Features .... 33

Figure 3-6 Raster Layer Expressing The Acceptable Locations Within A 10km Distance From The Ontario Major Airports ..................................................... 37

Figure 3-7 Raster Layer Presenting The Population Intensity In Ontario ............ 38

Figure 3-8 Raster Layer Expressing The Acceptable Locations Within A 2km Distance From The Ontario Rail Tracks .............................................................. 39

Figure 3-9 Raster Layer Showing The Available Locations Within A 2km Distance From The Full Ontario Network Including The Major Highways .............. 40

Figure 3-10 Raster Layer Illustrating All The Open Areas Available For The Construction Of A New Logistics Facility .............................................................. 41

Figure 3-11 Raster Layer Showcasing The Suitable Locations Within A 5km Boundary To The Existing Warehouses ................................................................. 42

Figure 3-12 Raster Layer Presenting The Agglomeration Economies Effects in Ontario ............................................................................................................... 43

Figure 3-13 Assigned Facility Using A P-Median Problem ................................. 47

Figure 4-1 Data Exploration Categories ................................................................ 49

Figure 4-2 Scatter Plot Of (a) Trip Productions And (b) Trip Attractions ........... 51

Figure 4-3 Truck Trip Productions Share By Industry ........................................... 54
Figure 4-4 Intensity Of Truck Trip Productions Related To (a) Manufacturing, (b) Transportation, (c) Retail And (d) Wholesale Industries........................................56
Figure 4-5 Intensity Of Truck Trip Attractions Related To (a) Manufacturing, (b) Transportation, (c) Retail And (d) Wholesale Industries........................................57
Figure 4-6 Intensity Of Firms Around Ontario.................................................................58
Figure 4-7 Scatter Plot Of Truck Trip Productions Versus Total Jobs..........................62
Figure 4-8 Frequency Of Trips Crossing The Ambassador Bridge From On To Us Based On (a) Kilometers And (b) Hours; And From Us To On Trips Based On (c) Kilometers And (d) Hours .................................................................67
Figure 4-9 Frequency Of Trips Crossing The Blue Water Bridge From On To Us Based On (a) Kilometers And (b) Hours; And From Us To On Trips Based On (c) Kilometers And (d) Hours .................................................................68
Figure 4-10 Frequency Of Trips Crossing The Peace Bridge From On To Us Based On (a) Kilometers And (b) Hours; And From Us To On Trips Based On (c) Kilometers And (d) Hours .................................................................69
Figure 4-11 Trips Within 650 Km From On To Us Across The (a) Ambassador Bridge, (b) Blue Water Bridge And (c) Peace Bridge .................................................72
Figure 4-12 Trips Greater Than 650 Km From On To Us Across The (a) Ambassador Bridge, (b) Blue Water Bridge And (c) Peace Bridge .........................73
Figure 4-13 Percentage Of Truck Trips From Peel To Mainland Us Crossing Via The (a) Ambassador Bridge, (b) Blue Water Bridge And (c) Peace Bridge...........76
Figure 4-14 Percentage Of Truck Trips From Certain Ontario Origins To Mainland Us Crossing Via The (a) Ambassador Bridge, (b) Blue Water Bridge And (c) Peace Bridge .................................................................................................................. 77

Figure 4-15 Suitability Surface Based On The Mcda ............................................................................. 83

Figure 4-16 Suitable Facility Locations With The Constraint Of Sufficient Open Land Size .................................................................................................................................................. 84

Figure 4-17 P-Median Model Results For The Placement Of One Facility Serving Ontario Demand Points .................................................................................................................................. 85

Figure 4-18 P-Median Model Results For The Placement Of Two Facilities Serving Ontario Demand Points .................................................................................................................................. 86

Figure 4-19 P-Median Model Results For The Placement Of Three Facilities Serving Ontario Demand Points .................................................................................................................................. 87

Figure 4-20 P-Median Model Results For The Placement Of Peel And Hamilton Logistics Facilities Serving Ontario Demand Points .................................................................................. 88

Figure 4-21 P-Median Model Results For The Placement Of One Facility Serving Us And Ontario Demand Points .................................................................................................................................. 90

Figure 4-22 P-Median Model Results For The Placement Of Two Facilities Serving Us And Ontario Demand Points .................................................................................................................................. 90

Figure 4-23 P-Median Model Results For The Placement Of Three Facilities Serving Us And Ontario Demand Points .................................................................................................................................. 91

Figure 4-24 Maximum Coverage Model Results For The Placement Of Three Facilities Serving All Of Ontario Demand Points With A 390 Min Cut-Off Radius .................................................................................................................................. 92
Chapter 1: INTRODUCTION

1.1 Overview

Freight transportation is a major driver of economic prosperity as it has a direct impact on the performance of many economic sectors (Mathew, 2009). In fact, using the Canadian Transportation Economic Account (CTEA) 2014, it has been calculated that the transportation sector contributes to 8% of the Canadian gross domestic product (GDP) (Transport Canada, 2019). Freight transportation has a major role in determining the cost of a product, it is typically the largest cost component of Supply Chain Management. Data from Logistics Management’s Annual Study of Logistics and Transportation Trends expresses that on average 10 to 11 percent of a company’s revenue is spent on transportation (Goodwill, 2017). Consequently, many companies across the food, commodity and consumer goods sectors struggle with defining their profit margins as transportation costs climb to nearly double the inflation rate (Johnson & Prentice, 2018).

Transportation is a complex domain having to adapt rapidly to changing political, social and economic conditions and trends. Accurate and efficient methods and tools are thus needed to enhance the planning and decision-making processes associated with freight logistics.

Logistics is generally the management of the movement of cargo between the location of origin to the location of the destination, meeting the requirements of customers and/or corporations. Although, logistics is sometimes thought of as the simple profession of “moving goods,” modern supply chain management consists of knowledgeable processes, sophisticated machinery, and advanced information and communication technologies. Logistics management is that part of supply chain management that plans, implements, and controls the efficient, effective, forward and reverse flow as well as storage of goods. To properly plan and design for optimized
freight movements a broader understanding of the geographic and infrastructure resources of a region becomes vitally important.

There are different levels of infrastructure for storing and loading cargo. Examples are distribution centres, warehouses, logistics hubs, and logistics centres. A warehouse is a platform used for storing cargo efficiently whereas a distribution centre is a logistics platform used to store and offer value-added services such as product mixing, order fulfillment, cross docking and packaging. A distribution centre stores products for a relatively lesser time compared to a warehouse although they have the capability of turning into a warehouse. Distribution centres are the centres of modern supply chains where its role is to efficiently meet customer’s requirements (Rodrigue, 2016). Similarly, a logistics hub is a central operating platform within a transportation network that serves to optimize transportation cost which houses the operations of unloading, scanning, sorting, labeling, bagging, and loading products. Finally, a logistics center is a designated location where all activities relating to transport, logistics and the distribution of goods for national and international transit, are carried out by various operators on a commercial basis (Higgins et al., 2012).

Having a location that optimizes accessibility to a wide variety of suppliers and customer networks is a key aspect in increasing the efficiency of freight systems. As the saying in real estate circle goes, “the most important factors in determining value are location, location, and location” (Sheffi, 2012). Also, the idea of economic agglomerations, which is locating firms within close vicinity of one another, has cost saving impacts arising from urban agglomeration. The main advantages of agglomeration are the development of efficient and specified suppliers, knowledge sharing and spillover among the co-located firms, as well as the development of local labor pools with specialized skills (Duranton & Kerr, 2015).
Location optimization is a well-established area of research in both management science and transportation engineering. Computing advancements have led to the development of a variety of location optimization methods within the Geographic Information Systems (GIS) software. Example of the latter is the Network Analyst of the ArcGIS software, which can be used to solve a variety of location optimization problems while using the real road transportation network. Additionally, logistics clustering is an important phenomenon to consider when locating new facilities. Understanding the clustering patterns of truck trip ends is an important step for freight transportation planning and optimal warehouse location. Clustering is the classification of data into groups of similar objects allowing data to be represented in fewer categories, where the data loses some fine details in order to achieve simplification and highlight patterns (Sandag, 2018). Clustering is the current subject of research in several fields such as statistics, pattern recognition, and machine learning. Each group, called cluster, consists of objects that are similar between themselves and different to objects of other groups.

The use of clusters as a descriptive or exploratory tool for regional economic relationships provides a more affluent and meaningful representation of local industry drivers and regional dynamics. Companies tend to be geographically clustered which induces many economic advantages from their geographic proximity to each other (Maoh & Kanaroglou, 2007). Porter (1998) focuses on the advantages of logistics clusters suggesting that it increases the productivity of the co-located companies, increases the pace of innovation, and stimulates the formation of new businesses. Urban areas are obvious clusters of human activities indicating that it is of economic advantage to develop the many levels of infrastructure needed for enhanced economic performance. However, while logistics clusters are significant when locating new facilities, it is not the only aspect that should be considered when trying to determine the optimal location of
logistics facilities. Different criteria must also be considered and analyzed when making the decision.

Multi-Criteria Decision Analysis (MCDA) is an affective planning tool that assesses different criteria affecting the decision-making process (Masud & Ravindran, 2008). MCDA is used by humans every day to make personal decisions. It is also applied in the corporate world, by government agencies, and by medical centers among other types of industries. Multi-Criteria analysis was first introduced in the 1970s and many studies displayed the importance of such analysis. For the purpose of this thesis, MCDA will be utilized to propose possible logistics facility locations based on various attributes.

1.2 Research Objectives

The performance of any freight transportation system is affected by supply and demand of goods as well as the duration of shipments. Since the transportation of goods is a critical part of the entire supply-chain process, the location of distribution centers relative to demand points must be effectively analyzed. The focus of this thesis is on logistics clusters which are basically massive complexes or centres handling high volumes of freight activities. These could be distribution centers, warehouses, logistics companies, etc. Hence, the overall objective of this thesis is to explore the locational patterns of truck trips with respect to these clusters to understand the performance of freight transportation in Ontario. Such understanding will allow us to model the optimal location of new logistics facilities. To this end, the specific objectives of this thesis are as follows:

1) Advance knowledge on truck movement and the best practices for locating freight logistics facilities in Ontario
2) Analyze the clustering of truck trip generation intensities by industry

3) Identify potential locations for logistics facilities

4) Model the optimal location of logistics facilities

1.3 Thesis Outline

The rest of this thesis is organized into four chapters. Chapter 2 provides an overview of the most recent and relevant literature on the subject matter. Chapter 3 presents the study area, data and different analytical methods and techniques employed in this thesis to conduct the analysis. Chapter 4 is dedicated to present and discuss the results from the conducted research, while the last chapter is used to provide a conclusion and direction for future research.
CHAPTER 2: LITERATURE REVIEW

Freight transportation has become one of the most important activities in regional and urban planning due to the impact of freight transportation and distribution on roadways and the performance of the economy (Crainic et al., 1997). In fact, Urban freight transportation planning plays a major role for cities to reach sustainability (Sönke et al., 2007). Traditional urban planning focused on passenger transport and disregarded the impacts of freight transport on daily traffic. However, given the growth in freight transportation activities over the past 20 years, sustainable city planning had to consider truck movements on the road transportation network. The following chapter will highlight key aspects in the current knowledge pertaining to the analysis of freight movement.

2.1 Analysis of Truck Movements

In general, truck movements are more difficult to model than passenger trips since they tend to not follow daily standard routes (Muñuzuri et al., 2012). The freight movement literature suggests a variety of frameworks that were employed to classify and analyze truck movements based on different criteria. A review of different freight forecasting models has been provided by Chow et al. (2010) with association to their model development, data requirement, and the different objectives that may be achieved using each model. On the other hand, the study by Anand et al. (2012) follows a framework that classifies cargo movement models based on the objective of modeling, the stakeholders’ involvement, the descriptors for modeling purpose and the viewpoint for achieving the objective. Further, Zhou & Dai (2012) discuss freight models based on their techniques, procedures, major data requirements, and real-world application cases. Other reviews only focus on the geographical levels (i.e., national versus international) (De Jong et al., 2004; Yang et al., 2010), or describe models that were developed at the urban and metropolitan level.
only (Anand et al., 2012; Zhou & Dai, 2012; Regan & Garrido, 2001). Finally, an attempt to enhance the existing freight trip generation models was made by Holguín-Veras et al. (2013) in order to improve the predictive ability of these models and/or utilize them outside of the study area in which they were derived.

Many modeling efforts have been implemented in freight demand modeling, with the three most well known approaches being the Conventional Approaches, Supply-Chain Approaches and Simulation (Tour-Based) Approaches. The conventional approach usually starts with a defined model applied and adapted to accessible date (NCHRP 606, 2008). This approach is relatively straightforward, easy to perform, and inexpensive. The information needed to formulate the models is usually obtained from surveys or roadside counts. Usually, the surveyed collect information about the number of trips, the time of each trip, the origin and destination of each trip, the type of commodity or service shipped, the cost of the shipment/service, the weight of the shipment, the distance travelled, the mode choice, the route choice and other general information (NCHRP 606, 2008; Chow et al., 2010; Regan & Garrido, 2001; Gonzalez-Feliu & Routhier, 2012).

The supply-chain modelling approach attempts to clearly understand, predict, and describe the behavior of the different components of the supply chain within different scenarios; making it possible to analyze the various components that affect the supply-chain process including freight transportation and the various economic processes that are usually undertaken to fulfill an order. Since the supply-chain consists of all the parties involved in fulfilling a customer request directly and indirectly (Chopra & Meindl, 2016), supply-chain models have the ability of providing valued information about the system mechanics and its activities in order to clarify how changes in
external factors can affect the performance of logistics and transport system (Tavasszy et al., 1998).

Finally, the Simulation approach is used to deliver new vision for modeling the urban commercial vehicle movement. This approach highlights the importance of light truck movements and the occurrence of service trips as opposed to only goods deliveries within a city. It provides a thorough demonstration of truck movement tours without having to deal with shipment and the related complexities regarding the allocation of shipments to vehicle, conversion of commodity flows to shipment or routing as would usually done in the supply-chain approach (Hunt & Stefan, 2007). This modelling framework leads to a better planning of freight activities within a city since it permits “feedback effects” to efficiently evaluate response to changes in policies (Chow et al., 2010). The tour-based approach can be used to measure the impact of commercial vehicle movements with respect to the environment and offer measures to improve the performance of the system. The focus in this approach is on tours and not trips which provides more realism in the model and enables it to capture the true behavior of the agents (i.e., trucks) influencing the dynamics of the system.

The reviewed papers concluded that further research is required to effectively predict freight trip generations and produce accurate models. The importance of obtaining more data on freight flows and demand patterns is emphasized as a topic for future research. Thus, to better design and analyze freight transportation systems, detailed data must be collected to identify recurring patterns and to develop predictive models. Such data was derived through sample surveys of firms and trucks, which can end up being labour exhaustive and expensive (Hunt et al., 2006). However, new technology, such as Global Positioning Systems (GPS), is now used as an alternative and/or complimentary method to generate data on the movement of trucks. GPS data
loggers and transponders installed in moving trucks allow fleet managers to track the movement of vehicles in real-time. The collected GPS data have been used in recent years to gain a better understanding on the movement of trucks (Gingerich et al., 2016). In fact, GPS data have been used in transportation to also study the movement of people between home, work, and shopping trips, using different modes of travel such as car, train, truck, or walking (Gingerich et al., 2016).

In the context of freight, the biggest advantage of having GPS data is the large volume of information gathered on the movement of individual trucks. This raw data can be processed and transferred into multiple useful forms, such as origin-destination trips, truck routes, speed, truck tours, and bottlenecks (Bernardin, 2015). Many recent studies including but not limited to Ma et al. (2011), and Gingerich et al. (2016), used truck GPS data to create OD trips. For example, Ma et al. (2011) used origin-destination trips collected from GPS data between traffic analysis zones in Puget Sound, Washington, placing them in a custom software interface to provide users with freight movement patterns between the origins and destination pairs. More recently, Gingerich et al. (2016) utilized a large GPS data set that represented the movement of more than 60,000 Canadian trucks to study and analyze the movement patterns and behavior of these trucks within Canada and between Canada and the United States.

2.2 Clustering Analysis

2.2.1 Logistics Clusters

Logistics clusters are clusters of firms that provide logistics services. These firms include third-party logistics service providers, transportation carriers, warehousing companies and forwarders, as well as the logistics operations of industrial firms. A large body of literature since the 1980s has focused on understanding the clustering of firms at particular locations in urban
areas and devising techniques to locate firms based on criteria related to profit maximization. Sheffi (2012), explained the tendency of industries to be geographically “clustered” or gathered within a close proximity, although urban areas are the most obvious clusters of human activity. Such clusters necessitate the availability of certain geographical attributes such as a centered location and significant government investments in physical infrastructure.

A noticeable growth of population and businesses indicate an advantage for developing several levels of necessary infrastructure to improve economic performance. Rosenthal et al. (2003) agreed that economic clusters may give rise to having several hubs and becoming employment nodes in the city. A bigger economic force is created when firms cluster together, which is necessary for the livelihood of firms in urban areas (Maoh & Kanaroglou, 2007). In addition, the agglomeration of companies, firms, or corporate functions tends to draw economic advantages based on the geographic proximity to other firms or corporate functions within the same industry. This in fact is a phenomenon that was originally perceived and described by the British economist Marshall in his classic 1920 work ‘Principles of Economics’ (Sheffi, 2012). Thus, the development of industrial clusters or complexes indicates the presence of a positive area of collocation. Sheffi et al. (2010) suggested numerous methods to analyze “industrial complexes” or clusters. The authors note that such methods are critical to the understanding of the regional growth and developments. Similarly, Woudsma and Jakubicek (2019) attempted to examine the logistics clusters in urban Areas across North America. According to the authors, no consistency of facilities clustering trends are seen in the peri-urban areas as the usual case for residential developments. They found obvious occurring of clusters in some urban areas such as Los Angeles and Atlanta while other cities show a consistent change in pattern. The Study by Goodchild et al. (2009) explained the complexity of using metropolitan areas to serve the demand of customers
with the recent increase of using ‘mega’ distribution centres of more than 500,000 square feet. According to Sheffi (2012), clusters tend to grow due to “positive feedback” or “reciprocal reinforcement” forces. That is, as more companies of a certain type move in, more suppliers and customers move in, making the cluster even more attractive. Hence, large facilities are now being located selectively to assist large regions that would have previously been served through multiple local, or via smaller distribution centres.

2.2.2 Trip Origin and Destination Clustering

Origin-destination (OD) trip data are a special type of transportation data that consider origin and destination locations without paying attention to the actual routes used to make the trip. Typically, trip productions (based on the origin) and trip attractions (based on the destination) will exhibit some form of clustering since the firms producing and/or receiving these trips are also clustered. That is, agglomeration economies will give rise to trip clustering based on the origin or destination of these trips. Clustering methods are popular tools when it comes to pattern recognition in computer science (Bishop, 2006), although they have been used in transportation research in recent years. According to the literature, there are two basic types of clustering algorithms related to OD data: 1) spatial clustering of actual trajectories, and 2) point clustering of origin or destination points. The former is concerned with the matching and clustering of complex routes, while the latter is focused on the spatial clustering of trip ends.

According to Adrienko & Adrienko (2011), the spatial clustering of actual routes may be classified into two groups: partitioning and hierarchical clustering (Adrienko & Adrienko, 2011). On the other hand, He et al. (2018) noted that the point clustering method for OD data analyzes the traditional OD trip matrix by counting the number of trips produced from the origin points or attracted to the destination points in the study area. The scale of the origins and destinations can
be defined subjectively by users or derived from the data by using the density-based grouping method, which clusters geographically close points to each other (Adrienko & Adrienko, 2011). The study by Guo et al. (2012) presented an approach to group spatial points into clusters, derive statistical summaries, and visualize spatiotemporal mobility patterns. He et al. (2018) presented a new Simple Line Clustering Method (SLCM) designed to determine the fastest route for every OD trip within a certain radius. This simple line clustering method (SLCM) aggregates OD lines into small spatial clusters of sufficient size to reveal the spatial characteristics in terms of movements.

When dealing with truck flows, most studies in the literature used one of the following two methods (i) the Origin-Destination (OD) factoring approach, and (ii) direct facility flow approach. The OD factoring method uses existing freight flow data to predict future flow patterns by either using the entropy maximization mathematical programming process or the Fratar expansion technique (NCHRP 606, 2008; Stefan & Hunt, 2004). Entropy maximization conforms to the doubly constrained gravity model and is useful when noticeable changes to the transportation network is likely to occur in the future (Ortúzar & Willumsen, 2011). The Fratar expansion method, on the other hand, is an iterative proportional fitting technique applied to freight tables on the basis of the predicted production and attraction growth rates. Afterwards, the factored OD matrix is used as input to the traffic assignment model to estimate future flows. By comparison, the direct facility flow factor method can be used to predict future flows by considering the existing base year data. This approach relies on economic analysis and time series analysis to estimate the flows based on historical data and change in the level of economic activities. It could be used either to estimate future flow for a facility by applying factors that account for the diversion of flow from that facility to other routes or by applying growth factors to the flow on that facility (Yang et al., 2010; NCHRP 606, 2008).
2.2.3 Types of Relationships among cluster members

There are two main types of inter-firm relationships contributing to the success of clusters: Vertical and Horizontal Relationship. The two types will be explained in the following subsections.

2.2.3.1 Vertical Relationships

Vertical Relationships are the connections created amongst trading partners involving the achievement of business operations within the same production. These relationships are significant because the major share of value offered by enterprises to their customers, is usually obtained through the gaining of parts and services from their suppliers. It is trading partner agreements that governs the exchange for data, information, or items between parties. The sale side is responsible for interacting with suppliers and a variety of service providers, while the gaining or commercial enterprises are in charge of the interaction with material providers and parts suppliers. A good examples of vertical relationship clusters are those produced by a single “channel master” such as “Toyota city” (Sheffi, 2012; Porter, 1998). Managing such interactions between trading partners is of vital importance, especially since firms have been moving away from vertical integrations and progressively outsource various functions and stages of productions (Shain, 2009). Vertical relationships have been contributing to the growth in truck freight transportation in Canada and across North America.

2.2.3.2 Horizontal Relationships

Horizontal Relationships are links between firms of the same stage of productivity such as trades between automobile manufacturing plants in Detroit, Michigan, or even film studios in Hollywood, California. These firms would compete with each other yet cooperate along aspects that would benefit both parties. Such relationships can also exist between functions of different
industries since some agglomerations are not driven by supply but rather by demand. Demand driven agglomerations can be very advantageous to the customers since it would result in competitive prices, higher quality, and availability (Sheffi, 2012). In addition, horizontal attainment is a business strategy where one company takes over another that operates at the same level in an industry.

2.2.4 Advantages of Logistics Clustering

Transportation and warehousing are the core activities of logistics firms (Kasilingam, 1998). Given that such logistics activities do not depend on the specific characteristics of the good that is being handled inside the box (Sheffi, 2012), companies co-locating in proximity to each other usually experience operational advantages. Some of the noted advantages are tacit knowledge exchange, a collaborative environment, trust between cluster inhabitants, the availability of supply base, and the support for research and educational institutions (Sheffi, 2012). Tacit knowledge is the type of information that is difficult to transfer verbally or by writing but needs to be rather discussed over specifications with a supplier. Conducting this exchange of information within a cluster will result in an easier, faster, less expensive, and more effective transfer of information specifically when done face-to-face (Rodrigues-Posea & Crescenzi, 2008). Clustering also influence the cultural environment and legal management where common experiences allows people and organizations to develop a common trust amongst each other.

Logistics clustering also includes the benefits of transportation cost reductions, increased levels of customer service, increased services, resource sharing, and higher levels of employment (Sheffi, 2012, Rivera, et al., 2016, Van Den Heuvel et al., 2012, Bowen, 2008). Van Den Heuvel et al. (2012) reported that firms that collaborate tend to send part of their freight in the trucks of
colleague firms. Schuldt and Werner (2007) also mentioned that under high levels of communication and strong inter-company linkages, shipping companies will share spaces in containers and will ultimately cooperate to design more efficient shipping routes.

Logistics parks further facilitate the benefits of clustering because firms are part of a larger institution (the park) that help strengthen inter- and intra-companies’ links (Battezzati & Magnani, 2000). In fact, logistics clusters can end up serving as an infrastructure to other industries that may need specified logistics competencies and can end up replacing manufacturing jobs. These clusters have the ability to serve multiple industries making them less vulnerable to the changes of any particular industry (Sheffi, 2012). Additionally, Rosenthal and Strange (2003) argued that the advantages arising from co-locating within an industry decrease rapidly over the first few miles of distance between firms. Thus, logistics parks, closed and delimited agglomerative environments, offer higher benefits than open clusters where firms locate at somewhat greater distances.

2.3 Supply Chain and Logistics

Freight transportation is the main element supporting supply chains, global commodity, complexes and functional integrated networks of production. It also supports trade and service activities that cover all stages of production from the transformation of raw materials to market distribution and after-market services (Leinbach & Capineri, 2007). Hence, logistics planning is essential to reduce extra expenses, avoid damaging goods and/or missed deadlines. Logistics planning should be a priority for any business that relies on shipped goods. This includes providing ways to reduce transportation costs and finding an optimal location to ensure timely planning (Adam, 2017). According to Ruriani (2014), choosing optimal warehouse locations based on transportation efficiency is even more important than leasing rates or tax incentives. According to
this study, centralizing the transportation locations is profitable, where a single focused location is beneficial analytically, and strategically.

Knowing the important decision-making factors enables companies to select their optimal distribution structure including Distribution Centre location(s) (Alexander et al., 2018). This is important because a good structure ensures high customer service levels and reduced logistics costs. A good structure also helps companies to adapt to rapid changes in customer preference (Lotze-Campen et al., 2008). Alexander et al. (2018) suggested that three main factors are dominant when companies decide on their distribution structures: Supply Chain management, Transportation, and Geography. Multiple distribution centres are required in the case of high volume and spatially dispersed products ensuring a reduced transport cost and a higher level of service. However, another study by Chopra (2016) found that multiple distributions centres may result in increased logistics costs. The author developed a distribution network design framework based on product characteristics, but also on network requirements such as response time and return ability. Meixell et al. (2005) pointed that a few supply chain models have a comprehensive approach that includes outsourcing and supply integration after reviewing global supply chain design models.

According to recent studies, warehousing has experienced a major change in recent years. National and global level trends play a vital role on the current characteristics of warehousing and logistics facilities (Woudsma & Jakubicek, 2019). Global trade in consumed goods has shifted to the use of containers from the original methods of sea transport. This has changed the way regions compete with each other for logistics activity and the selection of potential of ports for handling these shipped containers (Rodrigue et al., 2015). Containerization opened new opportunities for inland facilities that may have previously been tied to a dock in the ocean port. Although, some
storage facilities are unable to move from port to inland. According to Rodrigue et al. (2015), containerized goods are different in terms of being able to use intermodal rail services to move goods inland easily without transhipment at the port.

With the rise of e-commerce and the execution centres supporting it, the idea of becoming closer to the consumer has become very important. Thus, in order to locate new facilities or expand an existing one factors such as availability of land, local regulations and operational characteristics of the cities comprising the urban region become critical. Woudsma & Jakubicek (2019) explored the logistics operation in Canada, noting that these complex supply chain considerations are undertaken by companies doing business throughout the country. This is reflected through the constantly increasing amount of operative logistics and warehousing facilities (Woudsma & Jakubicek, 2019). By comparison, in the UK, it is observed that having a logistics facility located at a port has an advantage because the container has to be trans-loaded in all cases. It is also beneficial to deconsolidate the container quickly and close to the port (Mangan et al., 2011). While reasons exist to have goods processed at ports, the rise of inland ports has been well documented in the literature. Growth in intermodal transport is driving local authorities to focus on logistics firms as a strategy for economic development and focus on railways as a method of increasing traffic to their intermodal terminals. The growth in intermodal traffic and subsequent growth in inland centres as hubs is also noted in Bowen (2008), who wrote about the Ohio valley and Chicago.

According to Woudsma & Jakubicek (2019), Canada is a large country with extensive geography to be overcome by shippers. As such, choosing a location for a large or ‘mega’ distribution centre becomes a source of competition amongst region. In recent years, the growth in intermodal transportation appears to be pushing many urban regions to develop “logistics
parks”, centered around airports or newly constructed intermodal terminals, with a recognition that these areas need to have a number of supporting services to be successful and attract shipment values. In fact, the competitiveness of regions in global trade is contingent in part of their innate geographical advantages over others, but the geographic importance of regions does not guarantee their desirability as a region for logistics activity. The study by Woudsma & Jakubicek (2019) further stated that “Metropolitan regions can influence the amount of land available for logistics related activities”, which may have the effect of reducing the importance of these regions in the global logistics landscape.

2.4 Location of Logistics Terminals

2.4.1 Optimality of Locations

Identifying the best site for a particular facility is not an easy task since planners must consider economics and demographic factors, while addressing political realities. According to Aykin (1995) different methods were implemented to find the optimal location of facilities (e.g. stores, warehouses, etc.). Over the past twenty-five years, many of these methods have been implemented into computer software to solve practical problems. An example of such software is ArcGIS, a mainstream Geographic Information Systems (GIS) software. The advantage of using ArcGIS is the ability to solve a variety of network optimization problems including the shortest path problems, travel salesman problems, vehicle routing problems and location-allocation problems (see for example: Aaron et al., 2007). Boliang (2016) emphasized the importance of determining the location and size of distribution centers which translates into solving “a location-allocation” problem. He further discusses that “a reasonable location and appropriate size is a guarantee for considerable profits”. Besides solving network optimization problems, ArcGIS can
also be utilized to perform Multi-Criteria Decision Analysis (MCDA) for site suitability modeling (Rosenthal & Strange, 2003).

The locations of warehousing facilities are based on a set of complex considerations and decision-making processes. Jakubicek & Woudsma, (2011) stated that facilities are required to be well connected to international flows of goods for the global linkages of supply chains. Although different factors affect the location choices on a local level, these are similar with global flows but grounded in local realities (Jakubicek & Woudsma, 2011). The supply chain research has focused on a broad set of factors including logistics costs, service level and their trade-offs (Alexander et al., 2018). When modeling the optimal location of facilities or distribution centres, discrete facility location models (e.g., P-median model) can be used. The model assumes a finite set of potential locations (i.e., p facilities) where the optimal locations are chosen by minimizing the total logistics costs (Melo et al., 2007; Alexander et al., 2018).

2.4.2 Site Suitability

Selecting the optimal location for warehouses and distribution centers is an essential step in determining the efficiency of an infrastructure. As noted above, location-allocation models can be used to determine optimal locations based on the availability of potential sites. Thus, to solve the problem, a list of potential sites will be required as input to the problem. Obviously, not any location can be considered as a potential location and as such a methodical approach much be followed to determine the list of potential locations. A well-established approach for coming up with a list of potential candidates or alternatives is the Multi-Criteria Decision Analysis (MCDA). According to Marsh et al. (2016), the MCDA method can be used to support the decision-making process by evaluating a set of alternatives for conflicting criteria and objectives. More specifically, MCDA is a methodology that uses various conflicting criteria to derive a set of alternatives for
evaluation. This is done by normalizing the values of the various criteria on a similar scale and then combining them into one overall evaluation (Belton and Stewart, 2002). Each criterion considered in MCDA has a relative weight within the final evaluation to reflect its relative importance within the decision context. Some objectives include maximizing attendance, which will locate optimal stations from a group of potential locations such that more individuals or fleets are served, and minimizing travel time or distance, which aims to minimize travel time or distance between the supply and target consumers (Greene & Hensher, 2003). The subject of the evaluation is scored according to how it performs within each criterion. The aggregation of the scores received for each criterion multiplied by their relative weights results in a composite score. The scores of each alternative option can be compared to improve the objectivity, transparency and consistency of decision making (Inotai et al., 2018).
CHAPTER 3: METHODS OF ANALYSIS

3.1 Study Area

Ontario’s economy succeeds through its exclusive combination of resources, manufacturing expertise and exports, (Government of Ontario, 2016). Due to its economy and trading patterns, Ontario attracts and produces a large volume of freight movements through trucks. In fact, Natural Resources Canada Office of Energy Efficiency relays that Ontario has the highest number of heavy-duty trucks in the country (2009). In addition, most of Canada’s exports to the United States originates from Ontario as it is home to fourteen Canada-U.S. road border crossings. Also, the four hundred series highway availability in Ontario plays a major role in determining the locations of freight facilities as it has a large role with the distribution of freight in Ontario (Ghamrawi, 2018).

In recent years, immense developments of truck producers have occurred in Ontario, with vast residential labor force growth. Specifically, in relevant goods movement sectors, directed by the four hundred series of highways and managed by zoning and planning regulations (Ferguson et al., 2012). Many of these highway sections are among Ontario's most congested roads. Consequently, the province of Ontario was selected as the study area for this thesis due to its extensive freight activity and boarder relations where it’s strategic location of boarding the United States makes it an international gateway for people and commerce. Ontario occupies approximately 1.076 million square kilometers of Canada’s land housing more than 38% of the Canadian population. In order to further analyze the truck movements originating from Ontario and determine the location of truck trip origins and destinations, Ontario was divided into zones based on its census divisions as shown in Figure 3.1.
Figure 3-1 Ontario zones presenting the study area of the thesis
3.2 Data

Multiple datasets have been used to come up with the analysis of an optimized warehouse location which is the main objective of this thesis. The following sub-sections of the chapter will explain in detail the data used in the research and their various sources.

3.2.1 Truck GPS Trips

The main data used for analysis in this thesis represents truck trips carried by Canadian carriers for the month of September 2014. These trips were analyzed and generated from a large sample of GPS pings, which depicted the movement of 22,865 trucks owned by 449 Canadian carriers. These trucks operate across Canada and the United States. The granular GPS data was acquired by Transport Canada and lent to the Cross-Border Institute at the University of Windsor for analysis (Gingerich et al., 2016). When the granular GPS data were processed, Canada and the US were divided into zones to determine the location of truck trip origins and destinations as shown in Figure 3.1. On the Canadian side, the census division delineations were used to represent the zones. While, on the U.S. side, the delineations of Metropolitan Statistical Areas (MSA) and counties were used. In total, there was 293 census divisions and 3163 MSAs and Counties. The processed data resulted in the origin and destination information indicating the zones where a truck trip started and where it ended. The trip records also included the exact longitude and latitude coordinates of the trip starts and ends. These coordinates were used to assign the trip ends to the nearest business establishment location. The industry for which the nearest business belongs to was then assigned to the trip. Further, the international trips between Canada and the U.S. were identified by the crossing that the trucks used to move between the two countries. More specifically, trips moving through the Ambassador Bridge, Blue Water Bridge and the Peace Bridge were identified.
3.2.2 Population and Jobs

Population and job information at the census division level were used in the analysis. This data was obtained for the year of 2016 from the Canadian census developed by statistics Canada. According to the 2016 Statistics of Canada Ontario’s population is 13.45 million people. Most of which reside in the greater Toronto Hamilton region (GHTA) which is more than 39% of total Ontario Population. In fact, 20 percent of the Ontario Population reside in Toronto and percent of which reside in Peel. In addition, Ontario houses more than 86.2 thousand jobs in different sectors (Statistics Canada, 2016). Although each industry offers a certain amount of positions with the transportation section being on the lead; 28% pf the jobs available in Ontario are related to the transportation industry. While Manufacturing, Retail and Wholesale sectors also play a big role on the economy of Ontario with each sector holding about 21%, 15%, and 8% of the total jobs in Ontario respectively.

3.2.3 Commercial Vehicle Survey (CVS) Data

Another dataset that was used in the analysis was based on the commercial vehicle survey (CVS) of the Ontario Ministry of Transportation (MTO). The Commercial Vehicle Survey (CVS) is a survey of truck drivers conducted by the Ontario Ministry of transportation to gather detailed information about the movement of trucks and cargo on major roads and highways across the province. The survey collects information about trip, vehicle, and commodity from arbitrary selected trucks on the Provincial Highways, international border crossings, some municipal roads, as well as some special generators (GIS, 2018). Trip and commodity details, and vehicle characteristics, including dimensions and weights are collected from the selected trucks from across the province at more than 200 directional survey sites. The survey entails collecting traffic count data for a number of locations in ON, and identifying vehicles by class whether it’s a
passenger vehicle, a bus, a single unit truck, or a Multi-unit truck. The 2012 CVS database used for this thesis contains data from forty-five thousand interviews conducted between 2010 and 2014. The ministry applies a set of procedures to process the data and generate origin - destination information (O-D matrix) at the census division level. According to 2012 data, the total number of truck trips was around 130 Thousand trucks traveling around Ontario on a daily basis.

The O-D CVS data can be used to identify the volume of goods moving in Ontario by trucks. The average daily value of goods from the 2012 CVS is derived from the commodity data collected by the survey and attached to the road network based on commodity and trip origin and destination as well as routing information (Ashrafi et al, 2017). Although not used in this thesis, the data can also be applied to measure the impact of these trucks on the highway by calculating the average daily Equivalent Single Axle Load (ESAL). This may further help pavement engineers design roads to accommodate the impact of truck traffic. Another application of the CVS survey is to determine the traffic volumes at data collection sites where average hourly traffic counts data by vehicle class is collected. In addition, such data can be used to identify the Origin and Destination of the different commodities. Finally, this survey data can also be used to calculate how commodities are moving in the province by weight where daily commodity weights in metric tons can be derived from the weight of the major commodities carried by the surveyed truck. In this thesis the CVS data will be used to validate the GPS trip data based on trip productions and attractions.

3.2.4 Geographic Information Systems (GIS) Data

A Geographic Information System is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data. For the purpose of this thesis a number of ArcGIS layers will be created for use as input in the analysis. As will be discussed in the
following chapters, these layers will be used to represent the different criteria needed for the location-allocation modeling. Table 3.1 lists the considered spatial layers (variables) along with their description and sources. Further details highlighting the rational for using these spatial layers will be provided in the Multi-Criteria evaluation section.

**Table 3.1 GIS Data Layers**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land Use</td>
<td>A shape file showcasing weather an area is residential, industrial, or an open space ready for occupation.</td>
<td>DMTI Incorporation</td>
</tr>
<tr>
<td>Existing warehouses</td>
<td>A shape file of all the existing warehouses locations in Ontario</td>
<td>EPOI from DMTI Incorporation</td>
</tr>
<tr>
<td>Population</td>
<td>A shape file containing the population of each census division</td>
<td>Statistics Canada 2016</td>
</tr>
<tr>
<td>Ports</td>
<td>A shape file of the locations of all the major airports for freight transportation across Ontario</td>
<td>DMTI</td>
</tr>
<tr>
<td>Rail</td>
<td>A shape file showcasing the railway tracks available for freight transportation across Ontario</td>
<td>DMTI</td>
</tr>
<tr>
<td>Network</td>
<td>A shape file showcasing the Highway tracks available for freight transportation across Ontario</td>
<td>Route logistics data from DMTI</td>
</tr>
<tr>
<td>Labor Force</td>
<td>A shape file containing the number of jobs available within each industry for very census division.</td>
<td>Stats Canada</td>
</tr>
</tbody>
</table>

**3.3 Data Exploration Techniques**

It is critical to explore the data used in order to be able to better understand the data trends and verify the results which is the initial stage of any data analysis. The following sections will thoroughly explain the different techniques used to explore the data, identify patterns, and evaluate the most optimized potential locations.
3.3.1 Regression Analysis

Regression analysis is a reliable data exploration technique for estimating the relationships between an outcome variable and features on a specific topic of interest. The most common form of regression analysis is linear regression, in which a researcher finds a linear function that most closely fits the data according to a specific mathematical criterion. The process of performing a regression allows the analyst to confidently determine which factors matter the most, which factors can be ignored, and how these factors influence each other. Regression analysis also helps determine trends found within the data based on the data variability. The analysis of every regression model requires the analyst to determine the dependent variable \((Y_i)\), that is hypothesized to be influenced by one or several independent variables \((X_{1i}, X_{2i}, ..., X_{ki})\). Regression analysis is used to provide a mathematical equation that can be used to make predictions. The equation will take the following form:

\[
Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \epsilon_i
\]

Where \(\beta\)’s are parameters to be estimated based on the observed data. The procedure for estimating the linear equation is based on minimizing the error \((\epsilon_i)\) in the data, which is done by reducing the distance between each observed data point \((Y_i)\) and the estimated point \((\hat{Y}_i)\) on the regression line (i.e., \(\epsilon_i = Y_i - \hat{Y}_i\)). The estimation of the \(\beta\) parameters will be associated with p-values to determine if these parameters are statistically significant or not (i.e. whether they are different from zero or not). A significant parameter is one that will impact the dependent variable \((Y_i)\) and will be associated with a very small p-value. Regression analysis also provides an R-Square value which measures how well the model fits the data. R-square is the ratio of the amount of variability captured by the model to the total variability in the data. As such, R-square values close to 1 would suggest that the model was able to capture a high amount of the variability in the
data, whereas, an R-square close to 0 would suggest a weak model that fails to capture any variability in the observed data. In this thesis, the Regression Tool of Microsoft Excel 365 will be used to explore the data.

3.3.2 Kernel Density Analysis

In order to gain a better understanding of the available data and be able to visualize the trucking trip patterns, a kernel density analysis was conducted. Kernel density estimation is another very useful statistical tool or technique used to create smooth curves or surfaces given a set of data. It is a popular technique utilized for spatial data analysis, done through the use of counts per unit area. Originally, kernel estimations were developed to attain an even estimate of univariate and multivariate probability density from an experiential sample of observations (Delmelle, 2009). Estimating the intensity of a spatial point pattern is similar to estimating a bivariate probability density where bivariate kernel estimations can be easily adapted to provide an estimate of the intensity as shown in Figure 3.2. The purpose of kernel estimation is to create a continuous surface known as a density surface. Kernel maps are created by directly using a point pattern or by using an attribute of the point pattern. If the point pattern is used directly, the kernel intensity \( \lambda_\tau \) is calculated as follows:

\[
\lambda_\tau = \sum_{i=1}^{n} \frac{1}{\tau^2} k \left( \frac{s - s_i}{\tau} \right)
\]

Where \( s \) is a location in the study area \( R \), \( s_1, s_2, \ldots, s_n \) are the location of the events, \( \tau \) is a search radius and \( k () \) is the kernel function. On the other hand, if the point pattern will be based on the attributes \( y_i \) of the point pattern, then the kernel intensity is calculated as follows:

\[
\lambda_\tau = \sum_{i=1}^{n} \frac{1}{\tau^2} k \left( \frac{s - s_i}{\tau} \right) y_i
\]
Various types of densities could be used to represent the kernel function including uniform, triangular, quartic or Gaussian. In practice, the following quartic function (see also Figure 3.3), is the most frequently used to calculate the kernel density:

\[ k = \frac{3}{\pi} \left( 1 - \left( \frac{s - s_i}{\tau} \right)^2 \right) \]
Three main parameters will affect the resulting kernel density surface: cell size, bandwidth or search radius $\tau$, and type of calculation. Given that the output is a raster file, the cell size will determine the coarseness of the resulting density surface. A larger cell size will result in a coarse surface whereas a smaller cell size will result in a smoother surface. The bandwidth $\tau$, which is the area around each cell, will also affect the calculation. A small radius $\tau$ will restrict density patterns to the immediate area of the point event whereas a larger radius $\tau$ will allow the density patterns to become generalized. Typically, $\tau$ is determined by trial and error. The third parameter is the type of calculation used in interpolating the density surface (i.e. the kernel function type).

Kernel mapping, also known as heat maps, has the advantage of providing a powerful visualization tool to identify the areas of highest density or “hot spots” using a gradient color. For the purpose of this thesis, we use kernel estimations to visualize the clustering of the point pattern produced by trucking trip ends based on the exact longitude and latitude of each trip to create the density surfaces. These kernels will then be used to visualize and analyze the areas of the highest amount of trucking trip activities for each of the twelve industries to showcase potential optimized locations for the new warehouses. The twelve industries being agriculture, communication, construction, finance, manufacturing, mining, wholesale, public administration, retail, services, transportation, and non-classifiable. The kernel maps will also be used to visualize the clustering of the existing warehouses to further determine the most economically active zones.

3.4 Network Modeling Techniques

Network analysis will be employed to determine the optimal location of potential distribution centers. The Network Analyst extension of ArcGIS will be used to examine the effectiveness of the locations of each distribution center. For the purpose of this thesis, the geo-
spatial analysis process shown in Figure 3.4 will be utilized. The starting point of the analysis is the *Site Suitability Modeling*. The purpose of this step is to create a suitability map to identify a number of potential locations where distribution centers could be established in the study area. The next step in the analysis is the development of a *Network Data Model* based on an adequate network database. The latter in its simplest format must consist of a number of line segments (Arcs and links) that are interconnected to each other via nodes, where network topology is maintained. Topological relationships between the various links of the network are critical for network data modeling since these relationships define the connectivity of the network. Besides the topology, the network should be associated with proper attributes including link length and travel speed to determine the travel time (i.e., impedance) on the network. Our network database is derived from the 2016 DMTI Inc. Route Logistics Spatial database and focus on major roads, highways and freeways in the province of Ontario. The impedance used in the analysis is the free-flow travel time (in minutes) on the network.

*Figure 3-4 Network analysis process*
As shown in Figure 3.4, Site Suitability and Network Data Modeling will then be employed to perform *Location-Allocation Modeling* in ArcGIS. The objective of the latter is to determine the optimal location from a set of potential locations that will be derived from the *Site Suitability Modeling* stage. To do so, GIS raster analysis will be performed in ArcGIS to create the data layers representing the different criteria needed in the analysis. Here, Multi-Criteria Decision Analysis (MCDA) will be used to generate a suitability map depicting the potential sites. Next, the network data model created from the *Network Data Modeling* stage will be employed to represent the transportation network used by trucks. In a nutshell, we will be using the following geo-spatial analysis tools in GIS: raster data analysis, weighted sum MCDA calculations, and finally the location-allocation network analysis. The following sub-sections will provide a detailed description of each of the different components that will be provided and utilized to conduct the analysis.

### 3.4.1 Raster Data Analysis

Raster data models are used to represent both fields and objects. Raster data analysis requires using one or more grid (raster) layers as input to create a specific output. It simply uses cellular organization to divide space in a series of grid units, each unit being similar in size to all others (also See Figure 3.5) although, output is not always dependent on simple additions of input layers or criteria. The resolution of the Raster is based on the pixel size which is the grid cell or picture element, defining the level of spatial detail in ground units. A small grid cell dimension indicates fine resolution and therefore large storage space. Raster data analysis is based on operations applied to the cells forming the raster. Such analysis not only can be performed at the level of individual cells, but also can be applied to a group of cells depending on the type of required analysis. In raster GIS analysis several models may need to be created and executed in a particular order.
For the purpose of this thesis, several raster models were created, representing various criterion (e.g. highways, rail, existing warehouses, land use and origin destination clusters, etc.). A coding scheme was than developed for the different raster layers in order to make the calculation for a site suitability map that will provide the potential location of logistics hubs in the study area. The coding will insure a consistent cell size in all the generated layers. When generating the raster layers, the coding can follow either a Boolean coding approach or a continuous but normalized approach. In the Boolean approach, raster cells will be coded as 1 if they meet the criteria, otherwise they will be assigned a value of 0 (Eastman et al., 1995). In the normalized approach, values in the raster cells are by convention continuous (e.g. distance to an existing major market). However, these values must be normalized on a scale of 0 to 1. To do so, the following formula could be used (Carver, 1991):

\[ n_{v_c} = \frac{v_c - v_{\text{min}}}{v_{\text{max}} - v_{\text{min}}} \]

Where \( n_{v_c} \) is the normalized value for grid cell \( c \), \( v_c \) is the actual value of raster cell \( c \), \( v_{\text{max}} \) is the maximum value among all raster cells in the layer being normalized while \( v_{\text{min}} \) is the minimum value among all raster cells in that layer. Once all raster layers are created, the Multi-

---

*Figure 3-5 Illustration of raster data representation of real-world features (Source: ESRI)*
Criteria Decision Analysis (MCDA) approach will be applied to create a site suitability map. The latter will provide the potential locations of logistics facilities based on the chosen criteria and their weights (based on ranking of importance) as will be discussed in the next sub-section.

3.4.2. Multi- Criteria Decision Analysis (MCDA)

Multi-Criteria Decision Analysis deals with the evaluation of a set of alternatives that are generated from a set of conflicting criteria to meet a specific objective. When making comprehensive or important decisions, multiple criteria and levels of scale need to be accounted for (Carver, 1991). Comparing conflicting sets of criteria, such as quality and costs, can sometimes lead to confusion and lack of clarity when each criterion is treated on its own. Fortunately, the MCDA has the advantage of combining the individual criteria to come up with a meaning score for different alternatives for evaluation. Additionally, the application of MCDA in GIS provides a very powerful tool for decision making processes. GIS can be used to visualize the various alternatives spatially in which a decision rule can be modeled more adequately. For this thesis, MCDA will be applied with the help of ArcGIS to determine potential logistics facility locations. Several key attributes/ criteria will be considered and then weighted relative to their influence on the process. These criteria will be coded as separate GIS raster layers in which the raster cell values will be coded using the Boolean scheme or will be normalized on a scale of 1 to 1 if the original values are continuous.

To create the suitability raster surface, one of two traditional approaches can be used: 1) Additive Boolean and 2) Simple Additive Weighting (SAW) methods. The Additive Boolean approach can be applied when all the raster layers are coded using the Boolean scheme. Ideally, the raster layers are grouped logically in which a raster grid cell in the output raster layer will be set to 1 if the raster grid cell of all input criteria has a value of 1, 0 otherwise. The output suitability
surface will have grid cells with values equal to 1 and 0. Here, all the alternative locations (i.e., raster cells) within the output raster layers will have equal chance of being a potential site without distinction. By Comparison, the SAW method (also referred to as weighted linear combination (WLC) or scoring methods) distinguish between the various potential alternatives in the output raster layer. In the SAW method, not all criteria are coded following a Boolean scheme but some are rather standardized as described in the previous section. Unlike the Additive Boolean approach, SAW assigns weights to each grid layer before calculating the suitability map. The suitability raster layer is then created as follows (Eastman et al., 1995):

\[ S_c = \sum_{r=1}^{R} w_r n v_{c,r} \]

Where \( S_c \) is the suitability of raster cell \( c \), \( w_r \) is the weight raster layer \( r \), and \( n v_{c,r} \) is the normalized score of raster cell \( c \) that belongs to raster layer \( r \). We contend that the SAW Method is more robust compared to the Additive Boolean method and as such it will be used in this thesis.

As shown in Table 3.2, eight distinct criteria are proposed to come up with the potential locations for new logistics facilities in Ontario. The Analyzed criteria include Access to Airports, Population, Land Size, Proximity to Rail, Proximity to Highways, Connectivity, Land Use, Existing Warehouses and Agglomeration Economies Effects. Many of the GIS layers were created using the Buffer Geospatial tool of ArcGIS and then converted from Vector to Raster layers. Also, all raster layers were generated at a raster spatial resolution of 2Km raster cells. Further, any potential site for a logistics facility must be large enough to accommodate freight activities. Therefore, a constraint is applied in which the final list of potential sites must have a land size area of at least 1000 acres. The text that follows provides an explanation of how the GIS raster layers are created for use as input to the MCDA model.
Table 3-2 Proposed MCDA Criteria

<table>
<thead>
<tr>
<th>Analysis Criterion</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Access to Airports</td>
<td>In order to ensure efficiency, the new warehouse should be close to airports for easier access to cargo and freight transportation sights.</td>
</tr>
<tr>
<td>2. Population</td>
<td>The use of cargo increases as the population size increases in an area and so it is more efficient to design for new warehouses in highly populated areas.</td>
</tr>
<tr>
<td>3. Proximity to rail</td>
<td>Freight transportation activities need to access to rail to ensure goods are transported efficiently and smoothly. Logistics facilities will benefit from being in close proximity to rail network.</td>
</tr>
<tr>
<td>4. Proximity to highways</td>
<td>Freight transportation in Canada is mostly done through trucks and as such, logistics facilities must be well-connected with markers and suppliers. Therefore, they must be in close proximity to major highways.</td>
</tr>
<tr>
<td>5. Transportation Connectivity</td>
<td>The idea of using the road network by cover we ensure that we are covering all access points on the highway and also knowing that we are in proximity to the highway which increases the chance of being in proximity to available land</td>
</tr>
<tr>
<td>6. Land use</td>
<td>The way land is used determines whether it is appropriate or inappropriate to place a new infrastructure. For the purpose of this thesis, it is essential to find an open area that is able to hold a larger size infrastructure such as a warehouse.</td>
</tr>
<tr>
<td>7. Existing warehouses</td>
<td>In order to reduce cost and to ensure the efficiency of the new location it would be advantageous to place the new warehouse near the already existing ones.</td>
</tr>
<tr>
<td>8. Agglomeration Economies Effects</td>
<td>A high labor force usually indicates a high access to businesses and companies requiring goods to be transported to and from the different locations which is why it is efficient to locate warehouses near such locations.</td>
</tr>
</tbody>
</table>

3.4.2.1 Access to Airports

Freight transportation activities are usually intermodal by nature and relies on different modes such as trucks, aircrafts, and rail. For this reason, the first criterion accounted for in this
analysis is access to airports. A lot of the cargo being transferred to other countries, provinces, or even to farther cities are most likely transferred using air. Therefore, it is of great benefit to place logistics facilities within a close proximity to those airports to insure a smooth inter-modal operation and to reduce transportation costs. For the purpose of this analysis, the 12 major airports in Ontario were considered. A maximum of 10km distance between the logistics facility locations and these airports was deemed adequate. The data was obtained from DMTI Inc. and a 10km buffer was then created around those airports. A Boolean coding scheme was used where 1 is assigned to all locations falling within the 10km buffer and 0 otherwise, as shown in Figure 3.6.

![Raster layer expressing the acceptable locations within a 10km distance from the Ontario major airports](image)

**Figure 3-6** Raster layer expressing the acceptable locations within a 10km distance from the Ontario major airports

### 3.4.2.2 Population Intensity

The second criterion considered in the MCDA model is the population intensity of a region. Population is the main drive for demand and so it is key to place logistics facilities at the locations
of highest populations and therefore highest demand. Availability of population resembles access to labor force, something that is beneficial for logistics facility operations. The population information of each region in Ontario was acquired from statistics Canada and then plotted in ArcGIS to come up with a normalized raster layer using a scale from 0 to 1. Here, 1 indicates a high population intensity and 0 low intensity, as shown in Figure 3.7.

![Figure 3-7 Raster layer presenting the population intensity in Ontario](image)

3.4.2.3 Proximity to Rail, Highways and Connectivity

Proximity to rail and major highways is believed to be very important criteria when locating logistics facilities. A lot of cargo is transferred through trains especially larger bulky commodities and other finished products such as cars and vans. Therefore, being in proximity to rail is of great importance for logistics facilities to insure efficient inter-modality operations. Also, large volumes of goods are shipped daily by trucks. These trucks rely heavily on highways to move shipment in
a timely and efficient manner. In the MCDA work, GIS layers depicting locations that are within
2km from rail tracks and major highways were generated and assigned a value of 1, otherwise the
layer is assigned a value of 0. These layers are shown in Figures 3.8 and 3.9. Further, the full
Ontario network was also acquired to represent the connectivity criterion. Connectivity determines
the ease of transporting cargo and ensuring there are sufficient roads connecting between the
logistics facilities and markets or suppliers. A 2km distance between the facility and the road
network was also found sufficient for the intended purpose.

![Image of map showing locations within 2km from Ontario rail tracks]

*Figure 3-8 Raster layer expressing the acceptable locations within a 2km distance from the Ontario rail tracks*
Land use type is one of the criteria considered for modeling the potential location of logistics facilities. It is of great importance to place new logistics facilities at an open area away from residential regions for safety, zoning laws, heavy traffic and possible pollution footprints. Hence, it is most advantageous to locate new potential logistics facilities near industrial regions or at an open area. In order to calculate the raster layer for the open and industrial areas in Ontario, a land use shape file was acquired from Scholars Geoportal where industrial and open areas were extracted to allow for a clear identification of available land in the Ontario region. The raster layer is then calculated using Boolean coding schemes for further Suitability map calculations where 1 identifies the available locations in industrial or open areas and 0 identifies the locations that do not qualify for that criteria (See Figure 3.10).
3.4.2.5 Existing Warehouses

Proximity to the existing warehouse facilities was one of the other criteria used for the MCDA model. Knowing a facility already exists at a specific location is a great indication of suitability for a logistics facility. This is also related to the concept of economic agglomerations where placing specialized firms next to each other allows for a higher demand in that region and thus a higher profit and reduction in transportation cost. An area where a warehouse already exists would suggest that it is already zoned for logistics facility operations. It is worth noting that the existing warehouses were mostly clustered around the regions where the intensity of truck trip activities were the highest. Therefore, locating a new logistics facility within a close proximity to existing warehouses would be advantageous.

Figure 3-10 Raster layer illustrating all the open areas available for the construction of a new logistics facility
A shape file of all the existing warehouses across Ontario was acquired from Enhanced Points of Interest (EPOI) database maintained by the DMTI Inc. It was decided that a 5Km distance around the existing warehouses was sufficient for the establishment of new potential logistics facilities. As such, a 5Km buffer was created around the warehouses and a Boolean raster layer was then created where location falling within the buffer were assigned a value of 1 and locations outside the buffer were assigned a value of 0. The raster layer of the warehouse suitability is shown in Figure 3.11.

**Figure 3-11** Raster layer showcasing the suitable locations within a 5km boundary to the existing warehouses

### 3.4.2.6 Agglomeration Economies Effects

Another criterion considered for locating new potential logistics facilities is the agglomeration economies effects (i.e., job intensity) of every region. Ideally, zones with the highest intensity of jobs are a better option for housing logistics facilities. This is explained by
agglomerations of economic activities where firms co-locate near each other to achieve cost reduction. As more firms in related fields of businesses cluster together, their costs of production may decline significantly and even when competing firms in the same sector cluster, the outcomes can still be advantageous because the cluster attracts more suppliers and customers. Hence, placing the new logistics facility in a region where a lot of industrial activities are happening will ultimately result in a higher profit. This may also result in reduced shipping costs where carriers could place shipments from various firms on the same truck. The jobs of each zone were acquired from statistics Canada where a shape file was then created to better visualize the intensity of the jobs in each zone. This was then converted to a raster layer using a normalized scale of 0 to 1 to for further use in the suitability map calculations. The normalized raster layer representing the agglomeration effects is shown in Figure 3.12.

*Figure 3-12* Raster layer presenting the agglomeration economies effects in Ontario
3.4.3 Analytical Hierarchal Pairwise (AHP)

Decision maker can assign weights of each criteria (i.e. raster layer) in the MCDA based on its relative importance. There are several methods for deriving criteria weight values including ranking method, rating methods, Trade-off analysis methods, and Analytical Hierarchal Pairwise (AHP) Comparison methods. In the ranking method, the criteria are ranked according to a certain order. Once rankings are established the weights for each criterion can be calculated using the following equation:

\[ w_j = \left(1/r_j\right)/\left(\sum_k 1/r_k\right) \]

Where \( r_j \) is the rank of a given criterion. This method has the advantage of simplicity making it very practical when the number of criteria is less than 10. However, it lacks a theoretical foundation, so the AHP method can offer a better alternative.

The analytic hierarchy process (AHP) is a designed method for establishing and analyzing complex decisions, representing the most accurate approach for quantifying the weights of criteria. Instead of coming up with a “correct” decision, AHP helps find a decision that best suits the goal of the study (Saaty, 1990). This method offers a comprehensive framework for constructing a decision problem by relating the various criteria affecting the overall goal of the problem to each other. More specifically, AHP is utilized to estimate the relative magnitudes of factors through pair-wise comparisons. Usually, a survey is conducted to ask a certain population to participate in the pair-wise comparisons. Each of the respondents has to compare the relative importance between any two criteria. The designed questionnaire uses a Likert scale (-9 to 1 to 9) to estimate the importance of each criterion relative to the other criteria. For a total of \( n \) criteria, there will be \( C(n, 2) \) pairwise combinations, where:

\[ C(n, 2) = \frac{n!}{2! (n - 2)!} \]
Table 3-3 AHP Pairwise Comparison Matrix

<table>
<thead>
<tr>
<th></th>
<th>Criteria 1</th>
<th>Criteria 2</th>
<th>...</th>
<th>Criteria i</th>
<th>...</th>
<th>Criteria j</th>
<th>Criteria n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria 1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria 2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria i</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td>k</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Criteria j</td>
<td></td>
<td></td>
<td></td>
<td>1/k</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Criteria n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The pairwise comparisons can be summarized in a matrix that reflects the importance of one criterion to another, as shown in Table 3.3. In the matrix, if Criterion i is k times more important than Criterion j, then Criterion j is 1/k times more important than criterion i. This treatment allows the AHP approach to structure a hierarchy for the criteria affecting the problem. The actual measurements or subjective opinions provided in the above matrix can be used to calculate the criteria weights and Consistency Index (C.I).

Once the pair wise comparison matrix is created, a normalized pair wise matrix is to be calculated by dividing each importance value by the sum of all the values. For example, the normalized weight of criteria i compared to criteria j it would 1/k divided by the sum of all the values. The average of the normalized scores for each criterion are than calculated to obtain the criteria weights. Finally, the consistency index is to be calculated by the following formula,

\[
CI = \frac{\gamma_{max} - n}{n - 1}
\]

45
Where CI is the consistency index, $\gamma_{\text{max}}$ is the average of ration for all elements, n is the number of compared elements. Appendix A presents the pairwise evaluations that were introduced to come up with ratios for the pair wise comparisons between every criterion as compared to the others. A calculation of the weighting and consistency ration of each criterion were then calculated.

3.4.4 Location-Allocation Modeling

Location-Allocation Modeling in the Network Analyst of ArcGIS will be utilized to determine the optimal location of a specific number of logistics hubs based on the suitability map results. The location-allocation problem can be classified as either private or public sector problems. The private sector problem’s main concern is profit and ways to maximize revenue, whereas public sector problems are more concerned about Safety/Equity. The problem we are dealing with could be classified as a public sector problem. A number of well-established location-allocation models have been developed over the last few decades. These models are used to serve specified objectives. Among the most widely used models is the P-median problem model. Other models that have also been used and incorporated in commercial software like ArcGIS include the P-Median with Maximum Distance Constraint, Minimize Total Powered Distance model, Attendance Maximizing Problem model, Maximal Covering Location model, and the Maximal Covering Location Constrained model. All of these methods can be thought of as extensions of the P-median Problem model but impose various constraints depending on the problem being analyzed. These models can be used to determine the location of a given number of facilities, $P$, in order to minimize total distance traveled subject to various constrains. For instance, the P-median problem model is based on two constraints: (1) every demand node travels to its closest facilities, and (2) only a certain number of facilities, $P$, will be located. The objective function of the P-median problem can be formulated as follows (Daskin & Maass, 2015):
\[
\begin{align*}
\min_{\{x_{ij}\}} & \quad \sum_{i=1}^{n} \sum_{j=1}^{m} w_i c_{ij} x_{ij} \\
\text{Subject to:} & \\
\sum_{j=1}^{m} x_{ij} &= 1 & i = 1, \ldots, n \\
x_{ij} - x_{lj} &\geq 0, & i = 1, \ldots, n; j = 1, \ldots, m; i \neq j \\
\sum_{j=1}^{m} x_{jj} &= p
\end{align*}
\]

Where \( i \) is the demand points (given \( n \) demands points) and \( j \) is the supply point (\( m \) supply points), \( w_i \) is the amount of demand at node \( i \), \( c_{ij} \) is the shortest path travel cost from node \( i \) to node \( j \), and \( p \) is number of facilities. In the above model, \( x_{ij} = 1 \) if demand location \( i \) is assigned to facility site \( j \), \( 0 \) otherwise. Similarly, the value of \( x_{jj} = 1 \) if the facility opened at site \( j \), \( 0 \) otherwise. The nature of the solution of the P-median model is that facilities tend to be located at the weighted centre where most demand points are, as shown in Figure 3.13.

*Figure 3-13 Assigned facility using a P-median problem*
For the purpose of this thesis P-median model was utilized with the centroids of the census Divisions being considered as facility locations and the population of these centroids are the demand. For this problem type, facilities are located such that the sum of all weighted costs between demand points and solution facilities is minimized. It is traditionally used to locate warehouses, because it can reduce the overall transportation costs of delivering goods to outlets. Since Minimize Impedance reduces the overall distance the public needs to travel to reach the chosen facilities, the minimize impedance problem without an impedance cut-off is ordinarily regarded as more equitable than other problem types for locating some public-sector facilities such as libraries, regional airports, museums, department of motor vehicles offices, and health clinics. For this thesis, the p-median problem was set to solve for an optimal location to serve all demand points with the lowest cost.

For further analysis, the Maximal Covering Location Problem type from the location allocation model was also utilized to locate the number of facilities required in order to serve all demand points with a designated impedance cut-off. The maximum coverage problem locates facilities such that as many demand points as possible are allocated to solution facilities within the impedance cut-off. Maximize Coverage chooses facilities such that as much demand as possible is covered by the impedance cut-off of facilities. For this thesis, a cut-off time of 6.5 hours was used to test how many facilities will be needed to cover all 49 demand points within the Canadian consecutive driving time limit. The problem is set so that any demand point outside all the facilities' impedance cut-offs is not allocated, whereas a demand point inside the impedance cut-off of one facility has all its demand weight allocated to that facility.
CHAPTER 4: RESULTS AND DISCUSSION

4.1 Data Exploration

The analysis for modeling the most optimal locations for new logistics hubs around Ontario was done through multiple stages, using the trip data generated from the truck GPS records. As a first step, the trip data were validated. Next, the locations associated with the productions and attractions of the trips were explored. The exploration then examined the clustering of firms across Ontario to identify potential associations between the location of industries and freight activities. Once this is done, simple linear regression analysis was performed to confirm the association from steps 2 and 3 with the help of inferential statistics. Finally, the trucking trips along the three Canada-US borders (i.e. the Ambassador Bridge, the Peace bridge, and the Blue Water bridge) were explored. The five steps of data exploration are highlighted in Figure 4.1.

Figure 4-1 Data Exploration Categories

4.1.1 Data Validation

In order to examine the validity of the trips generated from the acquired GPS data, trip productions and attractions at the census division level were compared to the trip productions and
attractions obtained from the CVS data. The latter represents the most credible source of truck freight information for the province of Ontario. Pearson correlation was calculated for the productions and attractions, respectively. Trip productions of both data sets were highly correlated with a coefficient of approximately 95%. Likewise, trip attractions were also greatly correlated with a coefficient of 94%. A scatter plot of both data sets is presented in Figure 4.2. The correlation results indicate that the GPS data is a very good representation of the trucking trips occurring around Ontario as shown in the scatter plots presenting the productions and attractions of both data sets.

It is worth noting that the GPS trips represent the activities associated with the movement of a sample of trucks in Ontario over a course of one month. By comparison, the CVS trips represent the activities associated with the movement of all trucks in Ontario over a course of 24 hours. Despite the temporal difference, the productions and attractions from the GPS trips are well in line with those given in the CVS data. According to the CVS data, a total of 102,221 truck trips were generated on a typical weekday in Ontario. Whereas, in the GPS trips, the total number of produced and attracted trips from the sample trucks during the month of September 2014 was 86,207 and 85,126, respectively. This translates to a ratio of approximately 1.2 between the CVS 24-hour trips and the GPS one-month trips.
Figure 4-2 Scatter plot of (a) trip productions and (b) trip attractions
4.1.2 Trucking Trips Intensity

The sample trip productions and attractions from the GPS data depicted the movement of more than 22 thousand trucks that belonged to about 450 Canadian carriers. In order to organize these truck trips and better understand the trucking trip activities, the locations associated with the productions and attractions of the trips were explored. The ArcGIS mapping software was used by locating each trip origin and destination on the Ontario map based on its longitudinal and latitudinal coordinates. This was done to identify the zones related to the start and end of each truck trip activity. Once the truck trips were plotted on the map, it was necessary to calculate the intensity of the trips in each zone to determine the most active locations. Kernel density via the Spatial Analyst extension of ArcGIS was used to create heat maps of those trip ends by type of industry to visualize the intensity of the activities around Ontario. The heat maps in this section will showcase the most intense regions of Ontario in terms of trucking trips with correlation to the industry type being served by these trips. For brevity, only the industries with the most trip activities will be discussed in this chapter although the heat maps of the trip activities related to all 12 industries will be presented in Appendix B and C. The industries of most significance to Ontario are transportation, manufacturing, wholesale and retail. These industries represent more than 70% of the total trucking trip generations around Ontario.

The pie chart in Figure 4.3 shows the percentage share of the trip generations related to each industry. The percentages of the trip attractions are very similar to the ones acquired for trip productions as both production and attraction trips are highly correlated. Table 4.1 portrays the correlation between the trip productions and Trip attractions related to each industry; it shows a correlation of more than 95% for most industries indicating that a very similar number of trips are being produced and attracted to the various census divisions of Ontario.
<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation</td>
<td>86207</td>
<td>100%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>23902</td>
<td>28%</td>
</tr>
<tr>
<td>Retail</td>
<td>13107</td>
<td>15%</td>
</tr>
<tr>
<td>Services</td>
<td>13043</td>
<td>15%</td>
</tr>
<tr>
<td>Wholesale</td>
<td>7259</td>
<td>8%</td>
</tr>
<tr>
<td>Construction</td>
<td>5491</td>
<td>6%</td>
</tr>
<tr>
<td>Nonclassifiable</td>
<td>2589</td>
<td>3%</td>
</tr>
<tr>
<td>Finance</td>
<td>927</td>
<td>1%</td>
</tr>
<tr>
<td>Agriculture</td>
<td>587</td>
<td>1%</td>
</tr>
<tr>
<td>Public Administration</td>
<td>576</td>
<td>1%</td>
</tr>
<tr>
<td>Communication</td>
<td>566</td>
<td>1%</td>
</tr>
</tbody>
</table>
Figure 4-3 Truck trip productions share by industry

Table 4-1 The correlation matrix between trip productions and trip attractions per industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>0.97</td>
</tr>
<tr>
<td>Communication</td>
<td>0.95</td>
</tr>
<tr>
<td>Construction</td>
<td>0.99</td>
</tr>
<tr>
<td>Finance</td>
<td>0.94</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.94</td>
</tr>
<tr>
<td>Mining</td>
<td>0.97</td>
</tr>
<tr>
<td>Non-classifiable</td>
<td>0.82</td>
</tr>
<tr>
<td>Public Administration</td>
<td>0.93</td>
</tr>
<tr>
<td>Retail</td>
<td>0.85</td>
</tr>
<tr>
<td>Services</td>
<td>0.98</td>
</tr>
<tr>
<td>Transportation</td>
<td>1.00</td>
</tr>
<tr>
<td>Wholesale</td>
<td>0.99</td>
</tr>
</tbody>
</table>
4.1.2.1 Intensity of Truck Trip Productions and Attractions

According to the kernel density maps shown in Figure 4.4 a constant high intensity of trip productions occurs in the Greater Toronto and Hamilton Area (GTHA). The intensities observed in the GTHA is likely associated with the high presence of firms in these regions. The intensity of firms and its connection to the intensity of trip activities will be discussed in the next section. As can be seen in Figure 4.4a, the intensity of truck trip generations for the manufacturing sector is concentrated at the peel region and the cities surrounding it. By comparison, little to no activities seem to occur in the rest of Ontario. A similar pattern is seen for the transportation sector (Figure 4.4b), although a high density of truck trip activities is noticed to extend from the Toronto-York region all the way to London and Oxford. In fact, a similar pattern is noticed for the retail and wholesale sectors, the other two most important industries in Ontario. The GTHA also contains the highest intensity of trucking trips, although some noticeably high trip activities are also noticed in Hastings indicating that a good amount of wholesale and retails related trips are being produced and distributed from Hastings as well.

The truck trip attractions of the four major industries (i.e., manufacturing, retail, transportation, and wholesale) exhibit very similar trends to the truck trip production activities of the same industries as can be seen in Figures 4.5. The GTHA continues to have the highest intensity of trip attractions in almost all industries. A clear high intensity of trucking trip attractions can also be observed in Hastings especially with trips related to the wholesale, manufacturing and retailed sectors.
Figure 4-4 Intensity of truck trip productions related to (a) manufacturing, (b) transportation, (c) retail and (d) wholesale industries.
Figure 4-5 Intensity of truck trip attractions related to (a) manufacturing, (b) transportation, (c) retail and (d) wholesale industries
4.1.3 Firm Intensity

The truck trip intensity patterns highlighted in the previous sections can be explained by the distribution of firms and/or jobs throughout Ontario. The intensity of firms in Ontario were also explored to further understand the relationship and visualize where most of Ontario firms exist. There is a clear relationship between the intensity of the truck trip generations and the intensity of the existing firms offering the products transferred by these trips. As shown in Figure 4.6 most of Ontario’s firms are in fact located within the GTHA which explains the high intensity of truck trip activities (i.e. truck trip productions and attractions) in these regions as discussed in the previous section. Nevertheless, firms exist all over the province of Ontario with a considerably high intensity in Ottawa as well.

Figure 4-6 Intensity of firms around Ontario
Additionally, firm intensity is not only a good representation for truck trips but also a great representation of the jobs available around Ontario. The concentration of firms around Ontario and the jobs available are very highly correlated when compared at the census division level. In fact, the correlation coefficient of firms and jobs in the construction industry is 0.98, similarly for the wholesale industry the correlation coefficient was 0.97. Additionally, the firms in the transportation and manufacturing industries were highly correlated with their jobs as each sector exhibits a correlation coefficient of 0.95. This will in turn justify the clustering of the trucking trip patterns as the jobs available are also very highly correlated to the generation and attraction of truck trips shown in Table 4.2. Such relation will be explored with the help of regression analysis in the next section.

Moreover, the top ten most active cities were explored in terms of jobs, truck trips, and firm intensities to attain a deeper understanding of the relationship between the trips and jobs as well as the trips and firm intensity. It was found that the top ten cities of all factors are very similar as shown in Tables 4.3 to 4.5. In fact, the top 10 cities in terms of trip productions were all amongst the highest cities by jobs and Firms. However, Essex which is the ninth highest city in terms of trips, was not part of the top 10 in terms of jobs and firms. Looking further indicated that this region was the twelfth highest city by jobs and firms. This indicates that there is a high correlation between truck trip and the amount of jobs available in a specific census Division. This also verifies the high correlation found between the jobs and firms available at the census division level.
### Table 4-2 Correlation coefficients between truck trip generations and jobs by industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>0.20</td>
</tr>
<tr>
<td>Communication</td>
<td>0.62</td>
</tr>
<tr>
<td>Construction</td>
<td>0.79</td>
</tr>
<tr>
<td>Finance</td>
<td>0.55</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.82</td>
</tr>
<tr>
<td>Mining</td>
<td>0.91</td>
</tr>
<tr>
<td>Non-classifiable</td>
<td>0.60</td>
</tr>
<tr>
<td>Public Administration</td>
<td>0.62</td>
</tr>
<tr>
<td>Retail</td>
<td>0.82</td>
</tr>
<tr>
<td>Services</td>
<td>0.76</td>
</tr>
<tr>
<td>Transportation</td>
<td>0.44</td>
</tr>
<tr>
<td>Wholesale</td>
<td>0.05</td>
</tr>
</tbody>
</table>

### Table 4-3 Top 10 cities by trip productions

<table>
<thead>
<tr>
<th>CD name</th>
<th>ZONEID</th>
<th>Total Trip Productions</th>
<th>Total Jobs</th>
<th>Total Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peel</td>
<td>3,521</td>
<td>19,531</td>
<td>730,865</td>
<td>51,293</td>
</tr>
<tr>
<td>Toronto</td>
<td>3,520</td>
<td>7,212</td>
<td>1,437,555</td>
<td>151,573</td>
</tr>
<tr>
<td>Waterloo</td>
<td>3,530</td>
<td>5,823</td>
<td>291,060</td>
<td>24,552</td>
</tr>
<tr>
<td>York</td>
<td>3,519</td>
<td>4,851</td>
<td>590,640</td>
<td>47,350</td>
</tr>
<tr>
<td>Halton</td>
<td>3,524</td>
<td>4,788</td>
<td>297,755</td>
<td>22,524</td>
</tr>
<tr>
<td>Hastings</td>
<td>3,512</td>
<td>4,377</td>
<td>63,920</td>
<td>6,586</td>
</tr>
<tr>
<td>Middlesex</td>
<td>3,539</td>
<td>4,062</td>
<td>233,845</td>
<td>22,941</td>
</tr>
<tr>
<td>Durham</td>
<td>3,518</td>
<td>3,241</td>
<td>343,745</td>
<td>22,386</td>
</tr>
<tr>
<td>Essex</td>
<td>3,537</td>
<td>2,954</td>
<td>189,675</td>
<td>17,916</td>
</tr>
<tr>
<td>Hamilton</td>
<td>3,525</td>
<td>2,830</td>
<td>271,990</td>
<td>22,605</td>
</tr>
</tbody>
</table>
Table 4-4 Top 10 cities by jobs

<table>
<thead>
<tr>
<th>CD name</th>
<th>ZONEID</th>
<th>Total Jobs</th>
<th>Total Trip Productions</th>
<th>Total Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>3,520</td>
<td>1,437,555</td>
<td>7,212</td>
<td>151,573</td>
</tr>
<tr>
<td>Peel</td>
<td>3,521</td>
<td>730,865</td>
<td>19,531</td>
<td>51,293</td>
</tr>
<tr>
<td>York</td>
<td>3,519</td>
<td>590,640</td>
<td>4,851</td>
<td>47,350</td>
</tr>
<tr>
<td>Ottawa</td>
<td>3,506</td>
<td>501,085</td>
<td>2,195</td>
<td>39,401</td>
</tr>
<tr>
<td>Durham</td>
<td>3,518</td>
<td>343,745</td>
<td>3,241</td>
<td>22,386</td>
</tr>
<tr>
<td>Halton</td>
<td>3,524</td>
<td>297,755</td>
<td>4,788</td>
<td>22,524</td>
</tr>
<tr>
<td>Waterloo</td>
<td>3,530</td>
<td>291,060</td>
<td>5,823</td>
<td>24,552</td>
</tr>
<tr>
<td>Hamilton</td>
<td>3,525</td>
<td>271,990</td>
<td>2,830</td>
<td>22,605</td>
</tr>
<tr>
<td>Simcoe</td>
<td>3,543</td>
<td>251,950</td>
<td>2,530</td>
<td>21,602</td>
</tr>
<tr>
<td>Middlesex</td>
<td>3,539</td>
<td>233,845</td>
<td>4,062</td>
<td>22,941</td>
</tr>
</tbody>
</table>

Table 4-5 Top 10 cities by firms

<table>
<thead>
<tr>
<th>CD name</th>
<th>ZONEID</th>
<th>Total Trip Productions</th>
<th>Total Jobs</th>
<th>Total Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto</td>
<td>3,520</td>
<td>7,212</td>
<td>1,437,555</td>
<td>151,573</td>
</tr>
<tr>
<td>Peel</td>
<td>3,521</td>
<td>19,531</td>
<td>730,865</td>
<td>51,293</td>
</tr>
<tr>
<td>York</td>
<td>3,519</td>
<td>4,851</td>
<td>590,640</td>
<td>47,350</td>
</tr>
<tr>
<td>Ottawa</td>
<td>3,506</td>
<td>2,195</td>
<td>501,085</td>
<td>39,401</td>
</tr>
<tr>
<td>Waterloo</td>
<td>3,530</td>
<td>5,823</td>
<td>291,060</td>
<td>24,552</td>
</tr>
<tr>
<td>Middlesex</td>
<td>3,539</td>
<td>4,062</td>
<td>233,845</td>
<td>22,941</td>
</tr>
<tr>
<td>Hamilton</td>
<td>3,525</td>
<td>2,830</td>
<td>271,990</td>
<td>22,605</td>
</tr>
<tr>
<td>Halton</td>
<td>3,524</td>
<td>4,788</td>
<td>297,755</td>
<td>22,524</td>
</tr>
<tr>
<td>Durham</td>
<td>3,518</td>
<td>3,241</td>
<td>343,745</td>
<td>22,386</td>
</tr>
<tr>
<td>Simcoe</td>
<td>3,543</td>
<td>2,530</td>
<td>251,950</td>
<td>21,602</td>
</tr>
</tbody>
</table>

4.1.4 Regression Analysis

To further enforce the relationship between trip generation and jobs, a linear regression model was estimated by regressing trucking trip generation at the census division level against the
jobs available in each census division. The bi-variate relationship was also explored using scatterplots, and correlation factors. A scatterplot is one of the methods used to explore the pattern in the data by highlighting the nature of the relationship between two variables. It is also useful to explore the strength of the relationship and to determine the presence of outliers in the trend. The scatter plot in Figure 4.7 shows a clear positive relationship between the jobs and trip productions. Interestingly, the scatterplot suggests that regions of Peel, Hastings, Toronto, York and Ottawa are Ontario’s biggest outliers. In the case of Peel and Hastings, the number of generated truck trips is disproportional to the number of jobs that exists in that region. That is, despite the fact that Peel and Hastings have less jobs relative to other regions, they produce much more trips. On the other hand, Toronto, York and Ottawa house more jobs but do not produce as many truck trips. This goes along to show that the most active cities are the outliers having the highest trips and/or jobs.

![Figure 4-7 Scatter plot of truck trip productions versus total jobs](image)

*Figure 4-7 Scatter plot of truck trip productions versus total jobs*

Given the outliers and the linear pattern observed in Figure 4.14, a multivariate regression model is estimated. The results shown in Table 4.3 confirms the positive relationship between the
truck trips and jobs in Ontario. More specifically, a larger amount of trips is associated with more jobs. Further, the parameters pertaining to the regional dummies meet our a priori expectation in terms of their signs. More specifically, Peel and Hastings are associated with positive parameters while Toronto, York and Ottawa are associated with negative parameters. The estimated regression equation is significant with an \((F (6, 42) = 192.44 \text{ and } p\text{-value } =0.0000)\), with an \(R^2\) of 0.96. The \(R^2\) value of the model explains the percentage of variability in the observed variable which is basically the coefficient of determination that measures how well the model fits the data.

**Table 4-6 Parameter estimates of trip generation model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Beta</th>
<th>t-stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-25.90</td>
<td>-0.20</td>
</tr>
<tr>
<td>Total Jobs</td>
<td>0.0135</td>
<td>12.91</td>
</tr>
<tr>
<td>Toronto Dummy (1 if Census Division is Toronto, 0 Otherwise)</td>
<td>-12150.17</td>
<td>-7.84</td>
</tr>
<tr>
<td>York Dummy (1 if Census Division is York, 0 Otherwise)</td>
<td>-3088.97</td>
<td>-3.76</td>
</tr>
<tr>
<td>Ottawa Dummy (1 if Census Division is Ottawa, 0 Otherwise)</td>
<td>-4537.15</td>
<td>-5.93</td>
</tr>
<tr>
<td>Peel Dummy (1 if Census Division is Peel, 0 Otherwise)</td>
<td>9699.84</td>
<td>10.50</td>
</tr>
<tr>
<td>Hastings Dummy (1 if Census Division is Hastings, 0 Otherwise)</td>
<td>3540.82</td>
<td>5.63</td>
</tr>
<tr>
<td>Number of Obs.</td>
<td></td>
<td>49</td>
</tr>
<tr>
<td>Adj. (R^2)</td>
<td></td>
<td>0.96</td>
</tr>
</tbody>
</table>

### 4.1.5 Border Trips

To model optimized logistics hub locations, it is important to explore where the Ontario truck trips are going to and where they are coming from. It turns out that only 57% of the trips stay within the Ontario border while more than 25% of the truck trips originating in Ontario cross the three major Canada-US border points. The major border points considered in this thesis are the Ambassador, Blue Water, and Peace bridges. Table 4.7 gives a detailed breakdown of the percentage of trips moving along the three bridges. It is worth noting that there is a limit of allowed
consecutive driving hours for truck drivers. For instance, in Canada, a truck driver is not allowed
to drive more than 13 consecutive hours or spend more than 14 hours of on-duty time. Assuming
an average speed of a 100km/hr, a truck driver would be able to drive a maximum of thirteen
thousand kilometers per day. Keeping that in mind, a full 2-way trip would translate into a
maximum distance of 650 km in each direction. The amount of trips within these limits are also
highlighted in Table 4.7. As can be seen, the majority of trips (i.e., 66%) between Ontario and the
US via the Ambassador Bridge are long trips with a distance greater than 650 km. This trend is
opposite in the case of the border related trips using the Peace Bridge and Blue Water Bridge where
the majority are short trips especially in the case of the Peace Bridge.

Table 4.7 Trip distances along the borders

<table>
<thead>
<tr>
<th>Trip Type</th>
<th>Total Number of Trips</th>
<th>Distance ≤ 650 km</th>
<th>Distance &gt; 650 km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of Trips</td>
<td>Percentage of Trips</td>
</tr>
<tr>
<td>ON-ON</td>
<td>48,987</td>
<td>48,356</td>
<td>99%</td>
</tr>
<tr>
<td>Ambassador</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bridge</td>
<td>ON-US</td>
<td>6,478</td>
<td>2,176</td>
</tr>
<tr>
<td></td>
<td>US-ON</td>
<td>6,011</td>
<td>2,015</td>
</tr>
<tr>
<td>Peace Bridge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ON-US</td>
<td>2,084</td>
<td>1,441</td>
</tr>
<tr>
<td></td>
<td>US-ON</td>
<td>1,208</td>
<td>955</td>
</tr>
<tr>
<td>Blue Water</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bridge</td>
<td>ON-US</td>
<td>2,223</td>
<td>1,311</td>
</tr>
<tr>
<td></td>
<td>US-ON</td>
<td>3,131</td>
<td>1,612</td>
</tr>
</tbody>
</table>
4.1.5.1 Frequency of Trips across the Borders

As can be seen in Table 4.7, a substantial amount of the trips crossing the borders are beyond the consecutive driving time limits, indicating that those trips cannot be made in a one-day trip. To better comprehend the intensity of trips moving through the borders and have an understanding of the nature of distance with respect to the current logistics hub locations, the frequency of these trips was graphed using histograms. Histograms are a plotting method that shows the frequency distribution of a variable with respect to its class interval. In this case the variable is the truck trips with respect to distance. The histograms in Figure 4.8 – 4.10 characterize border trips between Ontario and the U.S along the three bridges based on the kilometers traveled (i.e., distance) and hours of driving.

Figures 4.8 (a-d) portray the frequency of the trips along the Ambassador Bridge. These histograms further prove that more than 66% of the ON-US trips using the Ambassador Bridge are long trips that take longer than 650 km (i.e. 6.5 Hours). In fact, there is a high frequency of trips above 27 hours (i.e. 2700 km) in one direction indicating that some trips may take more than 2 days to deliver when following the consecutive driving time limits. Whereas some trips take up to 82 hours (i.e. 8200 km) equating to more than 6 days which further demonstrates the need to place logistics facilities at a more optimal location to reduce driving distance/ time. A similar pattern is seen for trips being attracted from US to ON across the same bridge.

Secondly, the frequency histograms of truck trips going through the Blue Water Bridge were explored and presented in the Figures 4.9 (a – d). This shows that the majority of the trips are short trips of less than 650 km in distance. Although, a big fraction of the trips are long trips with a driving distance of more than 650 k-m taking more than 6.5 hours of driving in one direction. In fact, a frequency of 200 truck trips takes between 17-25 hours of driving in one direction from
ON to the US to deliver its good, equating to around 1.5 to 2 days to deliver the intended goods and double it to go back. A similar pattern is noticed for the attracted trips from US to ON although, a higher frequency of long trips is noticed with a frequency of 100 truck trips taking between 70 to 79 hours of driving in one direction to come to ON from the US through the same bridge.

Finally, the frequency histograms of truck trips going from Ontario to the United States and vice versa using the Peace Bridge were explored and presented in Figures 4.10 (a – d). A noticeably smaller ratio of long trips is notices through this Bridge where a very high frequency of trips shown for trips less than 643 km in one direction. This could be due to the fact that the Bridge is placed within a considerably close distance to the census division with the highest intensity of firms and logistics facilities. Nevertheless, a considerable high frequency of trips exceeds the 6.5 hours of consecutive driving limit in one direction and so an optimized logistics facility location is still of benefit. This will be further discussed in the next section and the locations of where these trips are going to will be further explored.
Figure 4-8 Frequency of trips crossing the Ambassador Bridge from ON to US based on (a) kilometers and (b) hours; and from US to ON trips based on (c) kilometers and (d) hours
Figure 4-9 Frequency of trips crossing the Blue Water Bridge from ON to US based on (a) kilometers and (b) hours; and from US to ON trips based on (c) kilometers and (d) hours.
Figure 4-10 Frequency of trips crossing the Peace Bridge from ON to US based on (a) kilometers and (b) hours; and from US to ON trips based on (c) kilometers and (d) hours
4.1.5.2 Cross Border Truck Trip Intensities

According to the above calculations, there is a high number of trips transferring cargo across the Canada-US border. To obtain a complete picture of the problem at hand, it is necessary to see where these trips are originating from to further understand the locations of the logistics hubs around Ontario. Figure 4.11 portrays the intensity of the trips going from Ontario across the three bridges. These are trips that falls within the 650km consecutive driving Canadian limit; that is, trips with a driving time of approximately 6.5 hours or less. The intensity maps for each Bridge show that the most active zones associated with cross-border trips are in fact the ones located within a close proximity to the census division housing the bridge. For example, Windsor is transferring most of the cargo carried by the trips moving through the Ambassador Bridge, while Peel transfers the most of the cargo moving across the Peace Bridge, and Sarnia produces most of the truck trips going through the Blue Water Bridge. The highlighted patterns indicate that there might be potential to exploit locations in close vicinity to the border to establish logistics facilities (e.g., warehouses) or industries that have trade relationships with the US. That could translate into time and monetary savings since it will allow cross-border trips to occur in one day.

Furthermore, cross-border truck trips that are more than 650km were explored. The intensities of these trips were calculated and are shown in Figure 4.12. A clear pattern is observed where Peel is the census division where most trips originate from regardless of which Bridge was used. This is an expected result as previous calculations showed that most existing firms are located within that area. However, this may cause a truck driver to take more than one day to make the trip to abide by the 13-consecutive hours of driving time limit. This in turn raises the question about the potential benefits of building new logistics facilities closer to the border and if that will help optimize freight movement between Ontario and the US. Obviously, locations in proximity
to the border might not be adequate but the benefits of trying to shift the origin of the trips from Peel to a significantly less congested region is a worthwhile endeavor that needs further exploration. The modeling work presented in the next section of this chapter tries to shed some light on this matter.
Figure 4-11 Trips within 650 km from ON to US across the (a) Ambassador Bridge, (b) Blue Water Bridge and (c) Peace Bridge
Figure 4-12 Trips greater than 650 km from ON to US across the (a) Ambassador Bridge, (b) Blue Water Bridge and (c) Peace Bridge
4.1.5.3 Cross Border Truck Trip Destinations

It has been observed that Peel is the most active zone in Ontario in terms of cross border freight movements especially for trip lengths that are more than 650km. It is beneficial to explore where these trips are ending at to have a better visualization of the effects of the location of logistics facilities. The percentage of truck trips going to Mainland US were calculated and mapped to show the trip destinations of trucks originating from Peel and traveling through the three bridges. Looking at the Choropleth maps presented in Figure 4.13, a pattern can be noticed where the US map is divided into three sections with each Bridge being in charge of serving a specific part of the US, mainly areas in close proximity to the Bridge itself. For example, more than 30% of the trips passing through the Blue Water Bridge end at the Chicago and Indianapolis metropolitan statistical areas which are the closest in distance to the Blue Water Bridge. By comparison, most trips passing through the Peace Bridge end at the east side of the US with more than 15% of those trips ending at Boston specifically. Additionally, Since the Ambassador Bridge is close in distance to the Blue Water Bridge a noticeably similar pattern between the trips crossing the two bridges is noticed although the trips passing through the Ambassador Bridge are covering a wider range of central US regions. More than 20% of these trips are ending at Chicago and Indianapolis but a high percentage is also ending at Dallas and Laredo which are more in the Central-Southern part of the US. This further highlights the importance of optimizing the location of facilities serving US markets in which truck trips crossing the Ontario-US border may make a full trip in one-day without exceeding the consecutive driving time limits. Such initiative to optimize the location of border-related logistics facilities may lead to more trading opportunities among the two countries allowing for a larger trip distance range in a more optimized shipping time.
The trips originating from other active regions aside from Peel were also explored to gain a better understanding of where these trips end at when crossing the border. These active zones were calculated in the previous section where they had the highest truck trip production intensities crossing the three bridges as shown in Figures 4.12. Accordingly, the percentage of border related trips ending at Mainland US from highly active Ontario regions, excluding Peel, are portrayed in Figures 4.14. A more constant trip distribution is noticed in this case where almost all of the US regions are being served. Although a noticeably higher concentration is seen within the areas closest to the bridge being used to transfer the goods. Trucks passing through the Blue Water Bridge are mostly serving central US, whereas the trucks passing through the Peace Bridge are concentrating more on the East side of the US. Finally, truck trips crossing Ambassador Bridge tend to be spread throughout mainland US.
Figure 4-13 Percentage of truck trips from Peel to Mainland US crossing via the (a) Ambassador Bridge, (b) Blue Water Bridge and (c) Peace Bridge
From Waterloo, Wellington, Halton, Toronto, and York

From Waterloo, Middlesex, Toronto, York, Halton, Huron, and Lambton

From Waterloo, Niagara, Toronto, Wellington, Hamilton, Brant, York, and Durham

Figure 4-14 Percentage of truck trips from certain Ontario origins to Mainland US crossing via the (a) Ambassador Bridge, (b) Blue Water Bridge and (c) Peace Bridge
4.2 Modeling Results

4.2.1 Multi Criteria Decision Analysis (MCDA) Results:

Multicriteria decision analysis (MCDA) was implemented for this thesis analysis in order to show, order, prioritize, and identify the potential alternative sites that could be used to model the most optimal locations for new logistics facilities. In order to find the most optimized location for a logistics facility, a set of criteria were considered. According to the literature, the key factors determining logistics facility locations are mostly related to transportation, geography, and supply chain management (SCM). SCM and transportation focus on distribution center location selection from a viewpoint of service level and logistics cost factors. Hence, criteria such as access to airports, connectivity, proximity to rail and highways, and proximity to existing warehouses were considered in this analysis. Additionally, the geographic literature focuses on spatial distribution of logistics facility locations that are mostly explained by location factors such as land use, population, and agglomeration economies effects. As such, the latter location factors were also included in the MCDA analysis. The weights of each criterion play a major role in MCDA models since they represent the relative importance of the utilized criteria. Several different methods are developed to determine the weights although for the purpose of this thesis the analytic hierarchy process (AHP) was used to calculate the weighted sum of each criterion.

The relative importance of each criterion was evaluated to come up with a decision that best suits the goal of this analysis. First, a hierarchy for the criteria is constructed by coming up with a pair-wise matrix comparing each criterion to others using a Likert scale from -9 to 1 to 9. This was done using a questionnaire comparing the importance of every criterion with respect to the others (Attached in Appendix A). The resulted pairwise comparison matrix for the eight criteria is shown in Table 4.8.
Table 4-8 The Pair wise comparison matrix of the eight considered criteria

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1.00</td>
<td>5.00</td>
<td>0.20</td>
<td>5.00</td>
<td>0.33</td>
<td>0.33</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>C2</td>
<td>0.20</td>
<td>1.00</td>
<td>0.11</td>
<td>0.33</td>
<td>0.11</td>
<td>0.11</td>
<td>0.20</td>
<td>0.33</td>
</tr>
<tr>
<td>C3</td>
<td>5.00</td>
<td>9.00</td>
<td>1.00</td>
<td>9.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>C4</td>
<td>0.20</td>
<td>3.00</td>
<td>0.11</td>
<td>1.00</td>
<td>0.11</td>
<td>0.11</td>
<td>0.33</td>
<td>3.00</td>
</tr>
<tr>
<td>C5</td>
<td>3.00</td>
<td>9.00</td>
<td>1.00</td>
<td>9.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>C6</td>
<td>3.00</td>
<td>9.00</td>
<td>1.00</td>
<td>9.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>C7</td>
<td>0.33</td>
<td>5.00</td>
<td>0.11</td>
<td>3.00</td>
<td>0.11</td>
<td>0.11</td>
<td>1.00</td>
<td>3.00</td>
</tr>
<tr>
<td>C8</td>
<td>0.33</td>
<td>3.00</td>
<td>0.11</td>
<td>0.33</td>
<td>0.11</td>
<td>0.11</td>
<td>0.33</td>
<td>1.00</td>
</tr>
</tbody>
</table>

C1: Access to airports; C2: Population; C3: Connectivity; C4: Land use; C5: Proximity to rail; C6: Proximity to highways; C7: Existing warehouses; and C8: Agglomeration economies effects.

These assessments highlighted in Table 4.8 were then converted to numerical values that can be processed and compared over the entire range of the problem. A normalized pairwise matrix was then calculated in order to measure weighting and consistency ratios. Finally, the average of the normalized scores for each criterion was calculated to obtain the criteria weights. Before using the calculated weights, it is important to check their consistency to ensure the pairwise process was logical and was not part of a random evaluation. Here, a consistency ratio (CR) is calculated using the following equation:

$$CR = \frac{CI}{RI}$$

Where the term CI is calculated as follows:

$$CI = \frac{\gamma_{max} - n}{n - 1}$$
In the above equations, $CI$ is the consistency index, $\gamma_{max}$ is the average of ratio for all elements, $n$ is the number of compared elements, and $RI$ is the consistency index of a randomly generated pairwise matrix. If the pairwise comparison was logical and rational, then the consistency ratio $CR$ should be less than 0.1 to insure the validity of the obtained weights. The 0.1 threshold suggests that if the hierarchy established is logical then the $CI$ value should differ from the $RI$ value since the latter is the outcome of a random pairwise evaluation process. When the number of elements $n$ is 8, the $RI$ value that should be used is 1.45. The $CR$ value for the above criteria weights was calculated to be 0.0223 which is less than the standard threshold of 0.1, indicating that the criteria weights are logical and may be used for the analysis. Once these values were assigned and validated the suitability map was calculated using the Spatial Analysis extension of the ArcGIS software.

Table 4.9 presents the standardized weight of each criterion which also allows us to rank the importance of the utilized criteria. It is observed that proximity to major highways and proximity to rail are the most important criteria when locating logistics facilities with a respective weight value of 19% for each. Trucks rely heavily on highways to transfer cargo to ensure timely and fast deliveries. Whereas transportation connectivity is the second most important with a weight of 18% to ensure there are sufficient roads connecting between the logistics facilities where cargo originate from and shipped by truck to a particular destination location.

Land use is the fourth most important criterion when it comes to locating logistics facilities with a calculated criteria weight of 9%. It is of great importance to place new logistics facilities at an open area away from residential regions for safety, zoning laws, heavy traffic and possible pollution footprints. Also, the fifth criterion accounted for in this analysis is access to airports. A lot of the cargo being transferred to other countries, provinces, or even to farther cities are most
likely transferred using air. The next most important criterion considered for locating new potential logistics facilities is the agglomeration economies effects (i.e., job intensity) of every region. Ideally, zones with the highest intensity of jobs are a better option for housing logistics facilities. Finally, proximity to the existing warehouse facilities is the last criteria to be considered for the MCDA model because an area where a warehouse already exists would suggest that it is already zoned for logistics facility operations.

Table 4-9 Standardized criteria weight

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Criterion Weight (%)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to airports</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Population</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Connectivity</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Land use</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Proximity to rail</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Proximity to highways</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Existing warehouses</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Agglomeration economies effects</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total Weight</strong></td>
<td><strong>100</strong></td>
<td></td>
</tr>
</tbody>
</table>

4.2.2 SAW Model Results

Once the weights of the MCDA have been found and verified, the simple additive weighting (SAW) method is applied to come up with a suitability surface. SAW is performed using the Raster Calculator of the Spatial Analyst extension of ArcGIS. The output is a surface that combines the various criteria and associated weights to depict the potential locations that could be used to establish a new logistics facility. The suitability surface of all the possible locations for the new logistics facility around Ontario is shown in Figure 4.15. According to the generated map, locations with the highest level of suitability (i.e. suitability level of 90 to 100 percent) are the key cities of Ontario. This is not surprising since most economic activities are taking place in these urban areas. These locations are referred to as primary locations but since urban centres are not tailored to house logistics facilities, we turned our attention to what we refer to as secondary
locations were the suitability level is between 80 to 90 percent. These locations are of paramount importance because they are not exactly in the urban area but are close enough to these urban centres.

Locations associated with the secondary suitability level scores resulted in approximately 78 candidate locations that are scattered around Ontario. These locations were filtered to only include sites that have a minimum area of 1000 acres. That is, a sufficient logistics facility should have an area of no less than 1000 acres in size. Thus, a land size constraint was introduced which resulted in approximately 18 candidate locations as shown in Figure 4.16. It can be observed that some of the 78 highly suitable locations such as Peel were not chosen because they did not meet the land size constraint.

Figure 4-14 Suitability surface based on the MCDA

Locations associated with the secondary suitability level scores resulted in approximately 78 candidate locations that are scattered around Ontario. These locations were filtered to only include sites that have a minimum area of 1000 acres. That is, a sufficient logistics facility should have an area of no less than 1000 acres in size. Thus, a land size constraint was introduced which resulted in approximately 18 candidate locations as shown in Figure 4.16. It can be observed that some of the 78 highly suitable locations such as Peel were not chosen because they did not meet the land size constraint.
4.2.3 Location-Allocation Modeling Results

4.2.3.1 Location-allocation Model serving all the Ontario Demand Point

In order to find the most optimized location for the new potential logistics facility or facilities, the location-allocation modeling tool from the Network Analyst extension of ArcGIS was applied. The placement of logistics facilities is considered as a private sector problem where the main objective is to minimize cost. Therefore, the p-median problem was then used to determine the location of a given number of facilities, \( P \), ensuring a minimized total travel time. The eighteen candidate locations found from the suitability map were used as the candidate logistics facility locations whereas the population of each census division was considered as the demand points. The application of the p-median problem located the logistics facility/facilities such that the total sum of the weighted impedance is minimized when serving all demand points.
within Ontario. Three scenarios were tested: 1) establishing one logistics facility, 2) establishing two logistics facilities and 3) establishing three logistics facilities.

Under the first scenario, the location-allocation model selects a location in Hamilton, Ontario as the most optimal for placing one new logistics facility as shown in Figure 4.17. The chosen location is able to serve all the demand points around Ontario with minimal cost and also has sufficient open space to place a logistics facility that is at least 1,000 acres. According to Ferguson et al. (2012), Hamilton has a suitable geographical location, a busy port, an international airport, good highway and rail access, and an educated labor force. The study by Ferguson et al. (2012) also proposed the development of Hamilton as a gateway for goods movement in, out, and through the area which further supports our finding by placing the new logistics facility in Hamilton. The obtain result indicate that Hamilton could be the next most active city after the Peel region in terms of regional freight activities.

Figure 4-16 P-median model results for the placement of one facility serving Ontario demand points
Under Scenarios 2, the location-allocation model selected a site in Ottawa besides Hamilton, as shown in Figure 4.18. The achieved results are justified by the clustering of freight activities in these regions and the high demand in both the Hamilton and Ottawa regions. Further, as shown in Figure 4.19, the location-allocation model under Scenario 3 selected Hamilton, Ottawa and Thunder Bay as the three optimal sites for establishing three logistics facilities.

When checking the cost in terms of travel time for each scenario, the results suggest that having one facility (i.e., Scenario 1) is more optimal than having two facilities (i.e., Scenario 2). The total time of serving all demand points from a logistics facility from Hamilton under Scenario 1 is 165.2 hours. By comparison, the overall cost under Scenario 3 is 124.63 hours, a reduction of 24.56% from Scenario 1. However, adding a third facility in Thunder Bay to service Northern Ontario might not be economical given the small demand in that part of the province.

Figure 4-17 P-median model results for the placement of two facilities serving Ontario demand points
A fourth Scenario was also tested to examine what would happen if we assumed a chosen logistics facility is to be added to Peel\textsuperscript{1}, which is currently the gateway for most trade around Ontario. Therefore, the location-allocation model was applied in the presence of Peel as an existing facility in the model. The results show that the presence of a facility in Peel denies the establishment of a new facility in Hamilton, particularly because Peel would be able to supply all demand point within the vicinity. Ottawa continues to be the next chosen facility and results in decreased travel time if the second logistics facility is added, as shown in Figure 4.20. These results further prove that the location of potential logistics facilities will shift depending on the demand

\textsuperscript{1} A rudimentary location-allocation model was tested using the centroids of the census divisions of Ontario as potential locations of logistics facilities and census division population as the demand. The model selected the Peel region as the optional location as shown in Figure D-1 in Appendix D.
of the population. More demand means increased benefits of placing more potential facilities given that other vital criteria mentioned in previous sections are accounted for.

Figure 4-19 P-median model results for the placement of Peel and Hamilton logistics facilities serving Ontario demand points

4.2.3.2 Location-allocation model serving Ontario and US demand points

As previously mentioned, about 25% of Canadian cargo originating in Ontario are transferred to the US across the three major Ontario-US land borders. Hence, it is important to model the optimal logistics facility location that is also able to serve US demand points. For this, three new scenarios with 1, 2, and 3 facilities were tested but this time US demand points were introduced in the model for the three main land borders. The model results are presented in Figures 4.21 - 4.23. The demand \( D_c \) for US markets via land crossing \( c \) was estimated relative to the Ontario population \( D \) by adjusting the latter based on the ratio of observed truck trips \( T_c \) that used land crossing \( c \) in relation to the observed truck trips \( T \) that took place within Ontario, that is:
\[ D_c = \frac{T_c}{T} \cdot D \]

The results in Figure 4.21 further prove that Hamilton is an optimal location that is capable of serving all demand point including US demand. When adding one more facility in the second scenario, we see Windsor emerging as the second most optimal location. Interestingly, the Windsor location seems to be serving the demand associated with both the Ambassador Bridge and Blue Water Bridge, as shown by the spider lines in Figure 4.22. As was discussed in earlier sections, more than 66% of the trips crossing the Ambassador Bridge exceed the executive driving time limit causing trucks to take more than one day to complete their trips. However, placing a new logistics facility in Windsor will be advantageous since it would help reduce the distance of trucking trips, which may allow more trips to meet executive time limits. Placing a new logistics facility in the Windsor region will also reduce driving times crossing the Blue Water Bridge while Hamilton facility continues to serve freight demand crossing the peace bridge. Finally, when running the model to select three facilities, Ottawa continues to be the third most optimal location serving the eastern side of Ontario, as shown in Figure 4.23.
Figure 4-20 P-median model results for the placement of one facility serving US and Ontario demand points

Figure 4-21 P-median model results for the placement of two facilities serving US and Ontario demand points
4.2.3.4 Maximum Coverage Problem

For further analysis, the maximum coverage problem was tested to account for the optimized number of facilities that is able to serve all demand points within a cutoff impedance. The cutoff impedance was set to be 390 Minutes (6.5 hours) which is the allowable consecutive one way driving time limit. The results shown in Figure 4.22 show that the optimal number of facilities needed in Ontario to be able to serve all demand point within the specified time limit is in fact 3. Although, the maximum coverage problem mainly focusses on serving a maximum number of demand point the results also show that 2 of the chosen facilities are placed in Hamilton and Thunder Bay which is very similar to our previous results from the p-median problem although the Ottawa location is shifted a little to be able to serve more demand within the cutoff impedance.
Figure 4-23 Maximum Coverage model results for the placement of three facilities serving all of Ontario demand points with a 390 Min cut-off radius
CHAPTER 5: CONCLUSION

This thesis advanced knowledge on truck movement and the best practices for locating freight logistics facilities in Ontario. It did so by exploring truck freight transportation activities in Ontario and by modeling the optimal location of freight logistics facilities in this Canadian province. Four data sources were employed in the analysis: 1) Truck GPS data, 2) Population and jobs of Ontario, 3) Commercial Vehicle Survey (CVS) data, and 4) Geographic Information Systems (GIS) data.

Using the truck GPS data enabled us to explore the clustering of freight movement activities around Ontario and to use the outcome to engage in regional network modeling. To date, most of the truck production activities that supply the various markets in the province and also move between the Canada-US borders have mostly relied on Peel’s logistics facilities. Peel is currently the gateway for most trade around Ontario and this is causing increased traffic congestion in that region. Traffic congestion increases transportation costs especially for longer distance trips within the province and across the border.

The analysis in this thesis was done by exploring Ontario truck trip productions and attractions carried by Canadian carriers for the month of September 2014. The first objective of the analysis was to explore the locational patterns of truck trip productions and attractions with respect to truck trip clusters for a better understanding of the performance of freight transportation system in Ontario. The second objective was to apply the extracted information and advanced geospatial methods to model an optimized potential location for new logistics facilities in Ontario.
5.1 Clustering Patterns

Clustering patterns of freight movements are critical in the modeling process of potential logistics facilities. In this analysis, the clustering of truck trips and firms for Ontario Census Divisions are analyzed with the use of truck GPS trips data. The intensity of truck trip productions and attractions in Ontario were calculated using the Kernel density calculation via the Spatial Analyst extension of the ArcGIS mapping tool. In addition, the intensity of the existing warehouse facilities in Ontario were explored to provide a clear explanation of the clustering patterns of the truck trips. Finally, the intensity of truck trips crossing the Canada-US border were also accounted for and explored to see where most of these truck trips start and end.

The results of the exploration analysis show a constant high intensity of trip productions and attractions occurs in the Greater Toronto and Hamilton Area (GTHA) with a noticeable high amount of trip activities occurring in Hastings. This was explained by the results of the firms’ allocations around Ontario. A clear relationship between the intensity of the truck trip activities and the intensity of the existing firms exists with most of Ontario’s firms being located within the GTHA as well. The analysis confirmed that firm intensity is not only a good representation for truck trips but also a great representation of the jobs available around the region. The concentration of firms around Ontario and the jobs available are very highly correlated with correlation coefficients of firms and jobs being between 0.95 to 0.98 depending on the industry type. This also justified the clustering of the trucking trip patterns as the jobs available turn to be very highly correlated to the truck trips as shown in Table 4.2 in chapter 4.

In order to model for an optimal logistics facility location, truck trip activities across the Canada-US border were also analyzed as more than 25% of the truck trips originating in Ontario cross the three major Canada-US land border points (namely Ambassador, Blue Water, and Peace
Bridges). The resulted intensity maps demonstrate that the most active zones associated with these trips are the ones located within a close proximity to the census division housing the land border especially for trips that are less than 650km in distance. A limit of 650km consecutive driving limit for one way is implemented by the Canadian government and so a longer distance might result in a multi-day trip. However, trips that exceed the 650km limit agree with the high clustering of previous results as they also tend to be mostly located at the GTHA regardless of the Ontario-US Bridge being used.

5.2 Logistics Facility Locations

Potential logistics facility locations are modeled using Multi Criteria Decision Analysis (MCDA). Several criteria were considered in order to guarantee identifying suitable sites that could be used to establish future logistics facilities. The relative importance of each criterion was evaluated using the analytic hierarchy process to come up with a decision that best suits the goal of this analysis. The results indicate that the proximity to highways, rail and having a strong road network connectivity close to the logistics facility are the most vital criteria. Access to airports and type of land uses are ranked as the second most important criteria followed by the effects of agglomeration economies. Furthermore, locations within close proximity of the already existing warehouses and locations that are highly populated tend to be important. Using these weight criteria, the simple additive weighting (SAW) method was applied to come up with a suitability surface. The latter was used to select a list of 18 candidate locations after using a land size constraint in which a logistics facility needs a minimum of 1000 acres of land. Finally, the location-
allocation modeling tool from the Network Analyst extension of ArcGIS was applied to find the potential optimal logistics facility/facilities.

The optimal logistics facility location analysis was done by testing two cases with 3 scenarios for each. The first set of scenarios were considered to place 1, 2, or 3 logistics facilities to serve the demand of the Ontario census divisions. The second set of scenarios considers placing 1, 2, or 3 logistics facilities to serve Ontario and US demand points. For the first set of scenarios, the location-allocation model selects a location in Hamilton, Ontario as the most optimal for placing one new logistics facility, Ottawa besides Hamilton for placing two new logistics facilities, and Thunder Bay along Hamilton and Ottawa for establishing three logistics facilities. The achieved results are justified by the clustering of freight activities in these regions and the high demand in both the Hamilton and Ottawa regions. Secondly, when considering the US demand alongside the Ontario demand, the results further prove that Hamilton is an optimal location that is capable of serving all demand point including US demand. When adding one more facility, we see Windsor emerging as the second most optimal location. Interestingly, the Windsor location seems to be serving the demand associated with both the Ambassador Bridge and Blue Water Bridge. Finally, when running the model to select three facilities, Ottawa continues to be the third most optimal location serving the eastern side of Ontario. The obtained results indicate that Hamilton could be the next most active city after the Peel region in terms of regional freight activities given its central location with respect to southern Ontario and key US markets. The results highlight the role of Hamilton as a gateway city. A gateway is identified as a city, or some transport and logistics oriented area in a city, that is associated with goods movement in, out, and through the area (Ferguson et al., 2012). Based on the analysis, Hamilton has a suitable geographical location, an international airport, good highway and railway access, high cluster of
truck trips and high labor force. These are all factors that help make a city a gateway along with the fact that Hamilton was chosen as the city to hold the most optimal potential logistics facility location.

5.3 Research Contributions

The analysis presented here offers a pioneering effort to address an important gap in the current transportation knowledge in terms of optimized logistics facility locations in an economically vibrant province like Ontario. The research in this thesis will contribute to the area of logistics freight transportation. Specifically, it introduces novel application of well-established methods to the fields of freight movement by analyzing one of a kind dataset on trucks that did not exist in the past. The performance of any freight transportation system is affected by supply and demand of goods as well as the duration of shipments. Since the transportation of goods is a critical part of the entire supply-chain process, the location of logistics facilities, whether distribution centers or major hubs, relative to demand points must be effectively analyzed. The contributions of this thesis are as follows:

1. The analysis provided evidence-based insights about the clustering patterns of trucking trip production and attraction patterns.

2. This is the first study to look at the factors affecting the optimal location of logistics facilities around Ontario in recent years. The site suitability map modeling and location-allocation analysis will produce new knowledge that can (a) help regions to make informed decisions about their future plans when it comes to investing in freight infrastructure projects, and (b) assist companies to identify the best locations to introduce new logistics facilities to ensure efficiency and optimized revenue.
5.4 Research Limitations and Future Research

The overall location-allocation results can be improved with access to and the use of additional information. Detailed data reflecting truck trip activities in the GTHA region will improve the accuracy of the locations for the proposed facilities. Also, considering existing zoning bylaws and any safety requirements for site selection can also ensure that the most suitable sites can eventually be chosen for future logistics facilities. A limitation of the current MCDA is not examining all potential criteria of interest. Costs of property acquisition or leasing will vary with the different census divisions. Factoring these costs into the analysis will introduce an additional criterion to the suitability analysis, which will add to the value of the results. Including more impedance in the analysis such as the cost of developing a new facility and comparing it to the cost of driving longer distances would also be advantageous. Also, it would be useful to include other type of major transportation ports such as multi-modal rail yards and sea ports to the analysis. In addition, adding traffic congestions to the analysis will improve the results. Nevertheless, these limitations could be rectified in future research. Future developments of this research could aim to investigate more criteria to insure the consideration of all factors affecting logistics facility development.

Additionally, future work should examine the costs and benefits of all sustainability pillars at a cost per weight of cargo transferred or per truckload. Another consideration in future research would be to examine the role of interacting with other provinces especially the province of Quebec. This province is home for Montreal, the 2nd largest market in Canada. Given the importance of Montreal and its proximity to the Ontario border, a number of large giants like Walmart and Shoppers Drug Mart have established their distribution centres in Cornwall, Ontario, which is in close proximity to the border between Ontario and Quebec.
REFERENCES


Border Crossings. Transportation Research Record: Journal of the Transportation Research Board. 2547. 1-10. 10.3141/2547-01.


APPENDICES

Appendix A: AHP Pairwise Evaluation

<table>
<thead>
<tr>
<th></th>
<th>1.00</th>
<th>3.00</th>
<th>0.33</th>
<th>5.00</th>
<th>0.20</th>
<th>7.00</th>
<th>0.14</th>
<th>9.00</th>
<th>0.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative to Access to Ports, Size of Land where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, population of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, connectivity of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, use of Land where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, proximity to rail of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, proximity to Highways of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, Proximity to existing warehouses where new warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, O-D Linkages where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Access to Ports, Access to labor force where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to size of land, population of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to size of land, connectivity of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to size of land, use of Land where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Relative to size of land, proximity to rail of where warehouse will be located is</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Relative to size of land, proximity to Highways of where warehouse will be located is

<p>| Relative to size of land, Proximity to existing warehouses where new warehouse will be located is | ✓ |
| Relative to size of land, O-D Linkages where warehouse will be located is | ✓ |
| Relative to size of land, Access to labor force where warehouse will be located is | ✓ |
| Relative to Population, connectivity of where warehouse will be located is | ✓ |
| Relative to Population, use of Land where warehouse will be located is | ✓ |
| Relative to Population, proximity to rail of where warehouse will be located is | ✓ |
| Relative to Population, proximity to Highways of where warehouse will be located is | ✓ |
| Relative to Population, Proximity to existing warehouses where new warehouse will be located is | ✓ |
| Relative to Population, O-D Linkages where warehouse will be located is | ✓ |
| Relative to Population, Access to labor force where warehouse will be located is | ✓ |
| Relative to Connectivity, use of Land where warehouse will be located is | ✓ |
| Relative to Connectivity, proximity to rail of where warehouse will be located is | ✓ |
| Relative to Connectivity, proximity to Highways of where warehouse will be located is | ✓ |
| Relative to Connectivity, Proximity to existing warehouses where new warehouse will be located is | ✓ |
| Relative to Connectivity, O-D Linkages where warehouse will be located is | ✓ |
| Relative to Connectivity, Access to labor force where warehouse will be located is | ✓ |
| Relative to Land Use, proximity to rail of where warehouse will be located is | ✓ |</p>
<table>
<thead>
<tr>
<th>Relative to Land Use, proximity to Highways of where warehouse will be located is</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative to Land Use, Proximity to existing warehouses where new warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Land Use, O-D Linkages where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Land Use, Access to labor force where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Rail, proximity to Highways of where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Rail, Proximity to existing warehouses where new warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Rail, O-D Linkages where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Rail, Access to labor force where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Highways, Proximity to existing warehouses where new warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Highways, O-D Linkages where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to Highways, Access to labor force where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to existing Warehouses, O-D Linkages where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to Proximity to existing Warehouses, Access to labor force where warehouse will be located is</td>
<td>✓</td>
</tr>
<tr>
<td>Relative to O-D Linkages, Access to labor force where warehouse will be located is</td>
<td>✓</td>
</tr>
</tbody>
</table>
Appendix B: Truck Trip Productions Heat Maps per Industry

Figure B-1 Intensity of truck trip productions for the Agriculture Industry

Figure B-2 Intensity of truck trip productions for the communication industry
Figure B-3 Intensity of truck trip productions for the construction industry

Figure B-4 Intensity of truck trip productions for the finance industry
**Figure B-5** Intensity of truck trip productions for the mining industry

**Figure B-6** Intensity of truck trip productions for the public administration industry
Figure B-7 Intensity of truck trip productions for the services industry

Figure B-8 Intensity of truck trip productions for the non-classifiable industry
Appendix C: Truck Trip Attractions Heat Maps Per Industry

Figure C-1 Intensity of Truck Trip Attractions for the Finance Industry

Figure C-2 Intensity of truck trip attractions for the mining industry
Figure C-3 Intensity of truck trip attractions for the public administration industry

Figure C-4 Intensity of Truck trip attractions for the non-classifiable industry
Appendix D: Rudimentary Location-Allocation Model

Figure D-1 Rudimentary Location–Allocation Model
NAME: Ayat Hussein

PLACE OF BIRTH: Windsor, ON

YEAR OF BIRTH: 1995

EDUCATION: Vincent Massey Secondary School, Windsor, ON, 2013

University of Windsor, B.Sc., Windsor, ON, 2017

University of Windsor, M.Sc., Windsor, ON, 2020