Validating Local Transit Accessibility Measures Using Transit Ridership

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Geography and Planning
University of Saskatchewan
Saskatoon

By

Sarah Bree

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Geography and Planning

117 Science Place
University of Saskatchewan
Saskatoon, SK Canada
S7N 5C8

ABSTRACT

Local transit accessibility measures are important tools used by planners to understand the effects of changes to public transit systems. Several local transit accessibility measures exist in the literature; however, it is not clear how these measures relate to public transit usage. Therefore, the aim of this study is to evaluate several transit accessibility measures that are commonly used in the literature by examining their association with ridership levels at the dissemination area level. The assessed transit accessibility measures ranged from a basic stop count, to gravity-based measures which use distance decay functions from a local household survey, and Walk Score's Transit Score. Using several land use and transit service datasets, including data collected from the fare box systems onboard the Saskatoon Transit buses, three types of model were tested. These models include ordinary least square models (OLS), spatial lag models (SLM), and spatial error models (SEM). The results from the models suggest that we can more closely predict actual public transit ridership when including a gravity-based accessibility measure in the model, while controlling for several household socioeconomic factors and built environment characteristics. In all cases, the measure that best fit the variation in ridership was the filtered frequency accessibility measure calculated using a 400 m network buffer and a distance decay function based on a Butterworth filter with a bandpass value of 250 m. This study offers transit planners and practitioners a better understanding of the performance of different transit local accessibility measures in relationship to actual transit ridership. Using the previous accessibility measure, the local accessibility of the proposed BRT system in Saskatoon was evaluated. The results showed an increase in accessibility in most of the city, even when the number of the stops was decreased from 1,443 to 994.

ACKNOWLEDGEMENTS

I would like to thank the following people all of whom made it possible for me to complete my master's degree! I would first like to thank my supervisor Dr. Ehab Diab, who continuously provided me with encouragement and was always willing and enthusiastic to assist me throughout the research project. Thank you Dr. Daniel Fuller and the INTERACT Team for your continued support and guidance. Thank you Dr. Ryan Walker, my committee chair, and Dr. Scott Bell for providing me with this opportunity. Thank you Phyllis for always being there to answer all of my questions, and rest of the Geography and Planning faculty for sharing their knowledge with me.

I would also like to give a special thank you to Dr. Bram Noble and Dr. Cherie Westbrook for all the extra support over the past two years. Thank you for believing in me when I began to question myself and for pushing me to keep going.

Thank you to the Canadian Institutes of Health Research (CIHR) funded *Environments and Health Signature Initiative Grant: Interventions, Research and Action in Cities Team* and the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding for this project. Thank you to Saskatoon Transit for providing the ridership data used in this research.

Finally, I would like to thank my family: Mom, Dad, Brandon, Lauren and Bruce, Papa and Great Gram, Great Grandma, Great Grandpa, Nicole and Michael. Without all of you pitching in to help, I would not have been able to achieve this goal.

I did it!

For Lauren

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

As the population of Saskatoon grows towards half a million, the city has outlined a plan for growth which places considerable focus on improving public transit. There are many benefits of a good public transportation system including: increased social equity, more sustainable urban layout, reduced travel times, and improved citizen health (Fayyaz, Liu, and Zhang 2017). Over the years, the City has spent time and money on its transit system. Despite constant efforts to improve the system, the city continues to experience a decline in transit ridership.

Saskatoon has planned a transit service network redesign. In 2026, the system will change from a spoke-and-hub network configuration to a grid network that is designed around a new BRT system. The system enhancements are estimated to cost between 90 and 150 million dollars (City of Saskatoon 2018). Experience with system changes from other cities has shown that transit users often prefer to walk farther than to wait longer for transit access (Moniruzzaman and Páez 2012). Planners have stated that the BRT enhancements are a move from a *coverage* based model to a *frequency* based model.

A similar transit enhancement approach taken by the City of Houston in 2015 (Natco 2018) saw a surge in ridership after such a plan implementation. There will be fewer stops but with higher frequency service. The proposed changes are primarily geographical in nature. Some areas of the city are expected to experience increased geographical accessibility, while other areas are expected to experience decreased geographical accessibility. The city is prepared to accept decreased transit route coverage in some areas in exchange for increased transit route coverage in other areas on the premise that improved service will prove the viability of the BRT enhanced system, justifying more enhancements in the future.

To successfully plan and implement such large scale infrastructure projects, costs and impacts must be estimated and reported to all stakeholders (Ding, Zhang, and Li 2018, Kim and Song 2018). Therefore, an appropriate local transit accessibility measure is needed to evaluate the changes in transit accessibility. Improving access to transit is an important issue that has been considered by many transit agencies and cities to increase ridership by attracting new passengers and retaining existing ones. *Local accessibility* to transit is usually measured using different tools (or measures), with the aim of understanding the effects of transit system upgrades and land use changes. Several accessibility measures exist in the literature, including stop count, coverage-based, and frequency-based measures. These measures aim to evaluate the offered transit service quality and can further be used in analyzing travel behaviour and transit ridership. Nevertheless, it is not clear how these measures are correlated with actual transit ridership. Therefore, this thesis examines the performance of several accessibility measures using Saskatoon as a case study, while controlling for the impacts of different influential socioeconomic and built environment factors. It applies the best preforming accessibility measure to the proposed BRT system to examine the relative impact it will have on local accessibility.

1.1 Scholarly and Societal Relevance

This thesis provides transportation planners, researchers, and engineers a methodological approach to measure and predict large scale ridership changes as the public transit system network is reconfigured. This resulting methodology can be used as a tool that can be adapted to analyze the accessibility of a transit system for any city. It also provides a practical contribution to society by allowing us to examine the relative impacts of changes in accessibility due to the implementation of the proposed BRT system in Saskatoon. Overall, this study offers a better understanding of the performance of different accessibility measures and the potential impacts of the planned Saskatoon BRT system on ridership.

1.1.1 Contributions to Knowledge

1. Demonstration of a method to analyze transit accessibility and identify the best method to explain ridership by comparing results from different methods to transit ridership.
2. Understanding how influencing variables affect accessibility changes and transit ridership in Saskatoon.
3. An estimate of the expected changes in public transit accessibility due to the planned BRT system implementation.

1.1.2 Contributions to Society

1. The estimates of accessibility may assist local planners with system design, and land use planning (i.e. transit corridors plans).
2. The model will provide the City of Saskatoon a better understanding of the factors that affect transit ridership within the city. This may be able to help them plan for and predict future ridership.

1.2 Purpose and Objectives

There are several methods used to measure different aspects of accessibility. However, methods for measuring large scale infrastructure changes remain limited and it is not clear how these measures are correlated with actual ridership and user behaviour (Kim and Song 2018, Ding, Zhang, and Li 2018). The purpose of this project is to study how the reconfiguration of urban transit routes and schedules impact transit ridership as users' access into the transit system changes. Specifically, this research:

1. Examines the performance of several transit accessibility measures by assessing their relationship with actual Saskatoon transit ridership. This will be done by:

    a. Identifying the most commonly used transit accessibility measures in the literature.

    b. Modeling ridership as a function of the proposed accessibility measure as well as other influencing factors, while testing different parameters for the accessibility measures such as distance decay functions.

2. Provides several recommendations on which accessibility measure should be used to estimate the potential changes in ridership in Saskatoon. This will help local authorities better understand the potential impacts of the planned BRT system on ridership.

3. Applies the accessibility measure that is deemed best to the proposed Saskatoon Transit System BRT configuration to estimate how accessibility will change as the routes are reconfigured.

2 LITERATURE REVIEW

2.1 BRT and Saskatoon Transit

To understand the impacts of the BRT system on ridership, it is important to first understand what a BRT system is, and how it differs from the current system. In addition, it is important to understand how the implementation of a BRT system has changed transit ridership in other cities. In general, a BRT (Bus Rapid Transit) system can be thought of as similar to a subway system, except with buses rather than subway trains. Like a subway system, the goal of a BRT system is to be fast and reliable. The Institute for Transportation and Development (ITD) has published a description of the characteristics that define a BRT system, called the *BRT Standard* (Policy 2016). These characteristics include: dedicated bus only median aligned transit ways, priority for buses at intersections, elevated loading platforms (i.e., stations) with prepaid fare zones, frequent service with up-to-date schedule information, and good branding. A scoring rubric for BRT systems is also defined. According to this rubric, BRT systems in Yichang China, Belo Horizonte Brazil, and Bogota Colombia meet the "gold standard" for BRT systems.

The Canadian Urban Transit Association (CUTA) describes BRT systems from a Canadian Perspective (Miller et al. 2018). CUTA defines BRT as a bus based rapid transit service that combines stations, vehicles, and running ways into a high quality, customer-focused service that is fast, reliable, comfortable and cost-efficient. They also list many of the features included in the ITD's BRT Standard as desirable components.

Saskatoon's proposed BRT system only incorporates a subset of the features a gold standard BRT system could have (City of Saskatoon 2018). Most of the system will utilize shared roadways and existing roadside bus stops. However, there will be a few blocks of dedicated transit ways and at least two "platform stations", but there are no plans for prepaid fare zones. Saskatoon's BRT system will be comprised of fewer stops than the current system, but the system overall will offer a higher frequency of service. Therefore, most of the proposed changes are both geographical and temporal in nature. Some areas of the city are expected to experience increased temporal accessibility, while other areas are expected experience decreased geographical accessibility.

2.2 The Benefits of BRT Systems from Other Cities

One of the first BRT systems in the world was implemented in Curitiba Brazil in 1972 (Goodman 2006). The system has proven to be very successful. The busses run at very high frequency, up to every 90 seconds. Prepaid fare zones keep the average stop duration to 15 seconds. It is estimated that 70% of the city's workers utilize the system for commuting, eliminating 27 million car trips per year. Per capita fuel use is 30% below the national average, resulting in some of the cleanest air in Brazil. Citizens spend approximately 10% of their income on transportation, well below the national average.

Following the success of Curitiba, the city of Bogotá Colombia implemented a BRT system in 2000 (Hidalgo 2015). It currently consists of 112 km of dedicated transit ways and serves two million passengers per day. Since its introduction, air pollution within Bogotá has decreased by 40%, and there has been a 79% reduction in collisions, a 75% reduction in injuries, and a 92% reduction in fatalities. The City of Yichang China implemented a BRT system in 2015 and increased public transit ridership by 20% (ITDP 2015). It includes 23 km of dedicated transit ways and 37 raised platform stations with prepaid fare zones. The system utilizes dedicated buses with doors on both sides.

The City of Houston Texas began planning a BRT system in 2012. Their goal was to increase transit ridership by increasing transit departure frequency on some routes while keeping costs unchanged for riders and taxpayers. The new BRT system provides 15 minute service along a few dozen routes. After the system was deployed in 2015, transit ridership immediately increased by 8% (Llamas 2015, Binkovitz 2016). The Houston project is considered a successful example of a BRT implementation.

2.3 Accessibility Overview

Accessibility is a term that is commonly used but often not properly understood, defined, or measured in a consistent manner (Geurs and van Wee 2004, Boisjoly and El-Geneidy 2017). Boisjoly and El-Geneidy (2017) studied the performance of various accessibility metrics. They state that accessibility has become a buzzword - a term thrown around by planners and decision makers as a key objective, goal or indicator in their development and growth plans. The authors also state that most of the plans they studied that referred to accessibility did not actually define the term (Boisjoly 2017).

One of the earliest definitions of transit accessibility is the "potential for opportunities for interaction" between users and the transit system (Hansen 1959). Other definitions include “the ease with which locations can be accessed” (Morris, Dumble, and Wigan 1979, Luo and Wang

2003), and, “the ease with which a user enters the system” (Langford, Fry, and Higgs 2012). Several researchers have examined the concept of accessibility as it is an essential component required to evaluate the performance of transport systems (El-Geneidy et al. 2015). The consensus among researchers is that measuring accessibility is as complicated as are the systems to which it applies.

2.3.1 Local Accessibility Measures

Local accessibility refers to what services residents can reach within their neighborhoods, e.g., local grocery or drugstores (Handy and Niemeier 1997). This study uses local accessibility to transit (i.e., access into the system) as suggested by Langford et al. 2012 to define accessibility. Transit services have spatial components (i.e., transit stops and stations) and temporal components (i.e., transit schedules). There are several methods that can be used to understand local accessibility to transit. In fact, several local accessibility measures were proposed and used in the literature to account for service quality at the aggregate level (Ibeas et al. 2012, Boisjoly and El-Geneidy 2017, Polzin, Pendyala, and Navari 2002).

The simplest way to compute local accessibility to transit is to count all the stops, routes, or departures within a geographical area (Thill and Kim 2005). However, Polzin et al. (2002) determined that using a very simplistic accessibility metric could overestimate model predictions.

Researchers have been examining more sophisticated local accessibility metrics. Ryus considered the “ease” with which pedestrians were able to reach a stop (i.e., sidewalks and paths) as their accessibility metric (Ryus et al. 2000). Gent computed an accessibility metric using the walking distance to a transit stop and frequency of service at that stop (Gent 2005). Ibeas used a count of the number of stops and the number of routes within a 400 m network buffer as their accessibility metric (Ibeas et al. 2012). Jacques and El-Geneidy used the number of stops in a census tract (Jaques 2014). Still, the number of stops is an important component in any accessibility metric. According to a recent survey among 343 practitioners around the world, most respondents indicated using the count of bus stops as an accessibility metric (Boisjoly and El-Geneidy 2017).

Other researchers have been investigating more articulated accessibility metrics. In 2003, Luo and Wang proposed an accessibility model to estimate geographical accessibility of medical services (Luo and Wang 2003). Their two-step model considered supply of surrounding services at a particular demand location, and the total demand on the services by surrounding locations. This model came to be known as the Two Step Floating Catchment Area model. In their initial model, they used a temporal buffer zone based on a thirty-minute travel distance, relative distance was not considered. For example, if a service was located 1 meter inside a demand location's catchment area, it was counted; but if a service was located 1 meter outside the catchment area buffer zone, it was not counted. However, it became apparent that distance matters. They conclude that in addition to local accessibility, regional accessibility must also be considered. That is even if a physician is located outside of the buffer zone in a neighbouring community they should also be considered to be accessible. They also suggest that socioeconomic factors such as income level and whether individuals own a personal vehicle should be considered as these factors will influence not only a user's accessibility to transit but if they require it.

In 2005, Thill and Kim studied public transit accessibility in Minneapolis–St. Paul (Thill and Kim 2005). A stepwise Poisson regression technique was used to model four types of trips: 1) home to work, 2) home to shopping, 3) home to school and 4) home to other locations. They concluded that no single metric appears suitable for measuring accessibility, as users' perception of transportation desirability depends on many factors. However, they did propose several options for geographical accessibility including a distance-to-service gravity function (i.e., a non-linear response to distance).

In 2009, McGrail and Humphreys examined the use of the 2SFCA model for estimating access to primary health care in rural Victoria, Australia (McGrail and Humphreys 2009). They proposed a distance decay weighting with an initial constant value of 1.0 (i.e., no decay) up to a certain breakpoint, followed by a linear decay to zero. They also examined the use of two different catchment buffer sizes. They looked at how a catchment buffer of 15 minutes would change the type of accessibility being defined. For example, using a catchment buffer of 15 minutes in rural Victoria would be considered local accessibility, whereas if the same catchment buffer of 15 minutes was used downtown Melbourne it would be considered regional accessibility. They explain that a single catchment buffer used for different areas will provide significantly different accessibility scores based on the type of region and the density of population and services it provides.

Additionally, in 2009, Luo and Qi proposed an enhanced 2SFCA (E2SFCA) model that applies a distance-decayed weighting to both steps of the original 2SFCA model (Luo and Qi 2009). They proposed discrete weightings that change in a stepwise fashion at certain defined distances. The step functions can be tailored to the service characteristics of the area served. For their study area, in northern Illinois, because of the low density of the area, a 1km quadrilateral grid was used to allocate population and housing counts from the 2000 census data. By implementing the enhancement of the stepwise distance decay weighting, they conclude that the results of using the E2SFCA reveal more of the accessibility details than 2SFCA. This is because the 2SFCA does not differentiate the spatial variation within each catchment.

In 2010, Dai examined the use of the 2SFCA algorithm for estimating access to health care in Detroit Michigan and proposed a non-linear distance decay factor based on a Gaussian distribution function (Dai 2010). He found that living in an area of poorer economic status with limited access to public transit limits the opportunities for education, employment, and health care access within a city. He also found that if they increased their buffer size to greater than 15 minutes, details about accessibility were lost as the results were over-smoothed because of the density and distribution of health care facilities in Detroit.

Later, Langford, Fry, and Higgs 2012 proposed a transit enhanced Two Step Floating Catchment Area algorithm for estimating geographical access into transit systems (Langford, Fry, and Higgs 2012). Their objective was to develop a transit-enhanced 2SFCA method for estimating local accessibility to a regional transit system. Previous to their study, there had been relatively little work on examining the potential utility of FCA techniques to public transportation. The transit enhancements applied by Langford et. al. to the 2SFCA algorithm include: 1) a Butterworth filter with a 250 m bandpass as a distance decay function between the demand location and the service location, and 2) a filter function such that only departures on a given route at the stop closest to the demand location are included. The method they produced is highly flexible and offers opportunities for further enhancement to include possible changes to the public transport network and schedules to explore spatio-temporal patterns of provision. In particular, they explain that there is potential for utilizing more intelligent demand-side and supply-side inputs into the model, and for the creation of more accurate temporal evaluations of accessibility.

Recently, Walk Score introduced *Transit Score* to quantify local accessibility to transit (Walk Score 2019). Transit Score is a filtered frequency model that uses departures per week as its service metric, while applying a weighted score for different transport modes (bus, rail, other). Scores are computed for points on a 153 m (500 ft) grid. The score considers distance to a stop, the routes it serves, and departure frequency. A custom distance decay function is used. The resulting scores are normalized to a scale from 0 to 100, with Manhattan scored at 100.

2.4 Modelling Ridership

A potential user's decision to use transit or not depends on many factors besides just access into the system (local accessibility). The importance of individual factors varies from user to user. There is no one formula that will determine if a potential user actually uses transit. Modelling ridership allows for consideration of the many factors that may not be considered in the local accessibility measure alone.

Moniruzzaman et al. modelled the Hamilton, Ontario transit system. They computed an "*accessibility **by** transit*" metric for use as a model input (Moniruzzaman and Páez 2012). The accessibility by transit metric computes a score for a *Dissemination Area* (DA) based on how many "destinations" can be reached from that DA by transit in a given time, and the desirability of those destinations. The factors that contribute to *accessibility* ***into*** *the transit system*, such as stop counts, route counts, and departure frequency are used directly as model inputs. Their results indicate that dwelling density is the only built environment factor (of those considered) that contributed to the model. They suggest investigating other built environment factors, such as number of traffic control features, amenities along the walking routes, even trees along the walking routes. Finally, they suggest a smaller scale (i.e., smaller than DAs) should be investigated as it could give a more accurate result. However, the smallest geographic area with a complete socioeconomic information in Canada is the dissemination area.

2.4.1 Factors Affecting Ridership

Several studies have focused on identifying the general factors affecting ridership at the aggregate level of city, neighborhood, and route. Different scenarios often require different measures. Factors such as land use and spatial distribution of jobs, shops, the transportation system itself, times at which travel opportunities are available, and the individual characteristics of users, such as their income and mobility needs must be considered (Miller et al. 2018).

Generally, these factors can be categorized as either internal or external (Taylor and Fink 2003). Internal factors include transit system attributes such as fare (cost), service frequency, network coverage, service reliability, safety, and comfort. Internal factors are normally within the control of the transit agency and have a direct impact on transit users' satisfaction, overall loyalty, and, therefore, travel behaviour (Zhao, Webb, and Shah 2014, van Lierop, Badami, and El-Geneidy 2018). More specifically, higher network coverage, more frequent service and better service performance (in terms of schedule adherence) are associated with higher ridership (Diab, Badami, and El-Geneidy 2015, Chapman et al. 2006).

External factors include population, household density, employment levels, income levels, and other socioeconomic and built environment factors (Thompson, Brown, and Bhattacharya 2012). Population and household density are identified as the main factors contributing to higher transit ridership levels (Guerra and Cervero 2011a). Lower income levels tend to increase transit ridership (Wang and Woo 2017). In addition, income is usually inversely related to other variables like unemployment rate (Miller et al. 2018). Built environment characteristics such as the presence of highways, road network distance, and mixed pattern road network are generally related to lower transit usage (Pasha et al. 2016). Higher Walk Score values are associated with areas that tend to have a greater mix of land use which has been found to have a significant and positive impact on transit usage (Miller et al. 2018). Other socioeconomic factors such as immigration rate and the percent of visible minorities also have a positive and significant impact on ridership (Miller et al. 2018, Kohn 2000).

2.5 Gaps in Literature

Over the years, accessibility has been extensively studied and has been frequently considered in transit system studies. The general consensus among scholars is that measuring accessibility is complicated as are the systems to which it applies. Methods used to measure or estimate accessibility have become increasingly sophisticated, but, there continues to be no “best” way to measure it (Thill and Kim 2005, Handy and Niemeier 1997, Miller et al. 2018). No previous study has provided a detailed understanding of the performance of different accessibility measures by assessing their relationship to actual ridership. Examining this relationship will help to determine which accessibly measure best represents transit ridership in Saskatoon. Further, by identifying the best fit measure, the potential impact of several policy interventions (e.g., BRT system) on ridership can be examined. In addition, this study considers spatial autocorrelation in the models assessed; the majority of the previous studies did not, which may have an impact of the quality of results.

3 METHODS

3.1 Study Context

Saskatoon is the largest city in the province of Saskatchewan, Canada, and in 2016 was home to approximately 280,000 people (Statistics Canada 2016b). The city has a dense urban core where active modes of transportation have been gaining in popularity. Automobile mode share remains high, at 89.5% (2016), particularly in the outer regions of the city (Statistics Canada 2016b). As the city's population grows towards 500,000, the City of Saskatoon has outlined an ambitious plan for growth, which places considerable focus on improving public transit (City of Saskatoon 2018). Over the years, the City of Saskatoon has spent time and money on its transit system. In September 2018, Saskatoon Transit, operated 41 regular bus routes serving 1465 stops with approximately 262,000 weekly departures (GTFS 2019). Figure 1 illustrates the stop and route locations of the Saskatoon Transit System as of September 2018. Figure 2 illustrates Saskatoon's proposed BRT system configuration as of April 2018. Figure 3 illustrates the three BRT lines planned for Saskatoon. The Red line runs from Blairmore in the west to Briarwood in the east, the Green line runs from Confederation Park in the west to Willowgrove in the east, and the Blue line runs from Stonebridge in the south to Lawson Heights in the north.

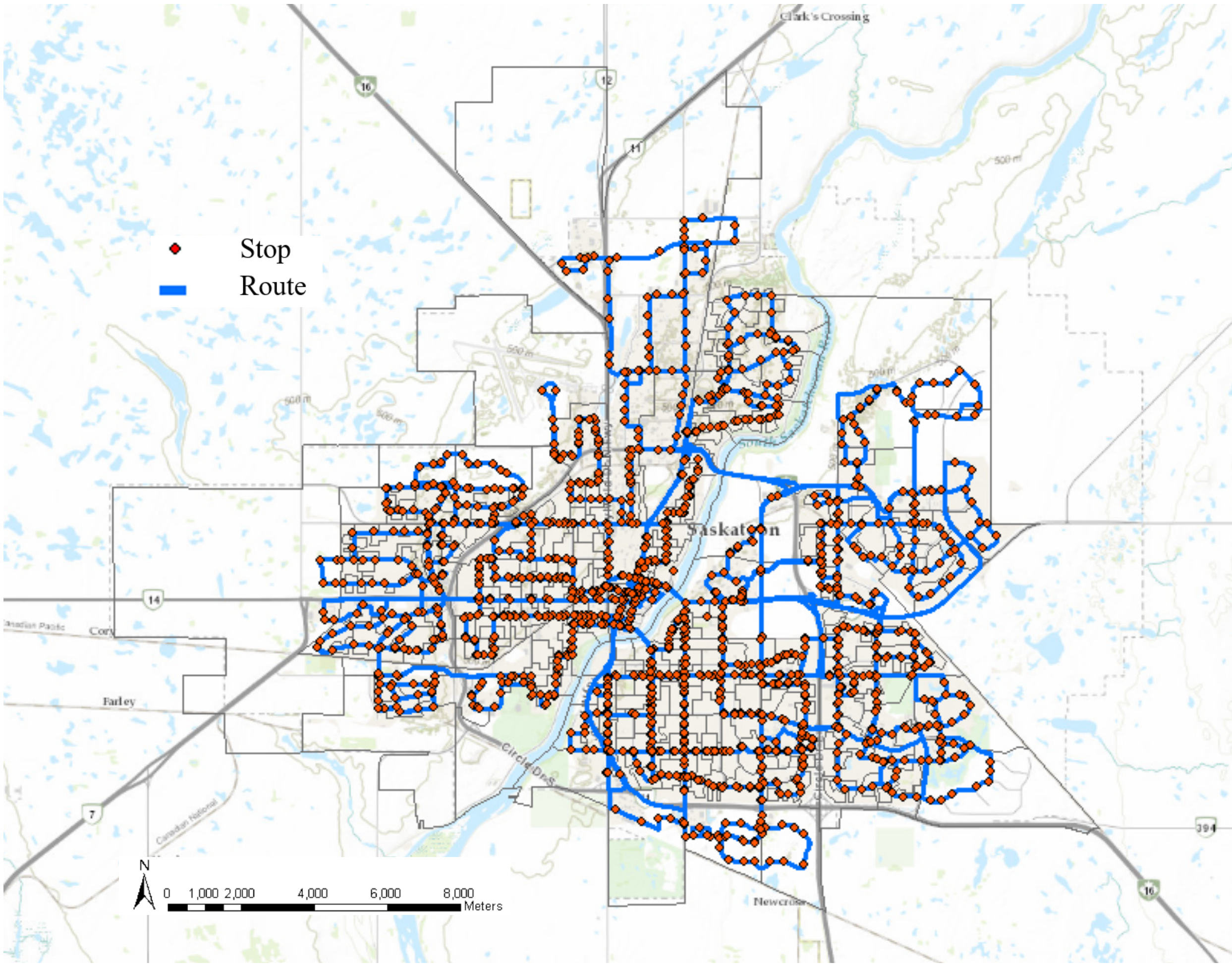


Figure 1: Saskatoon transit system configuration as of September 2018

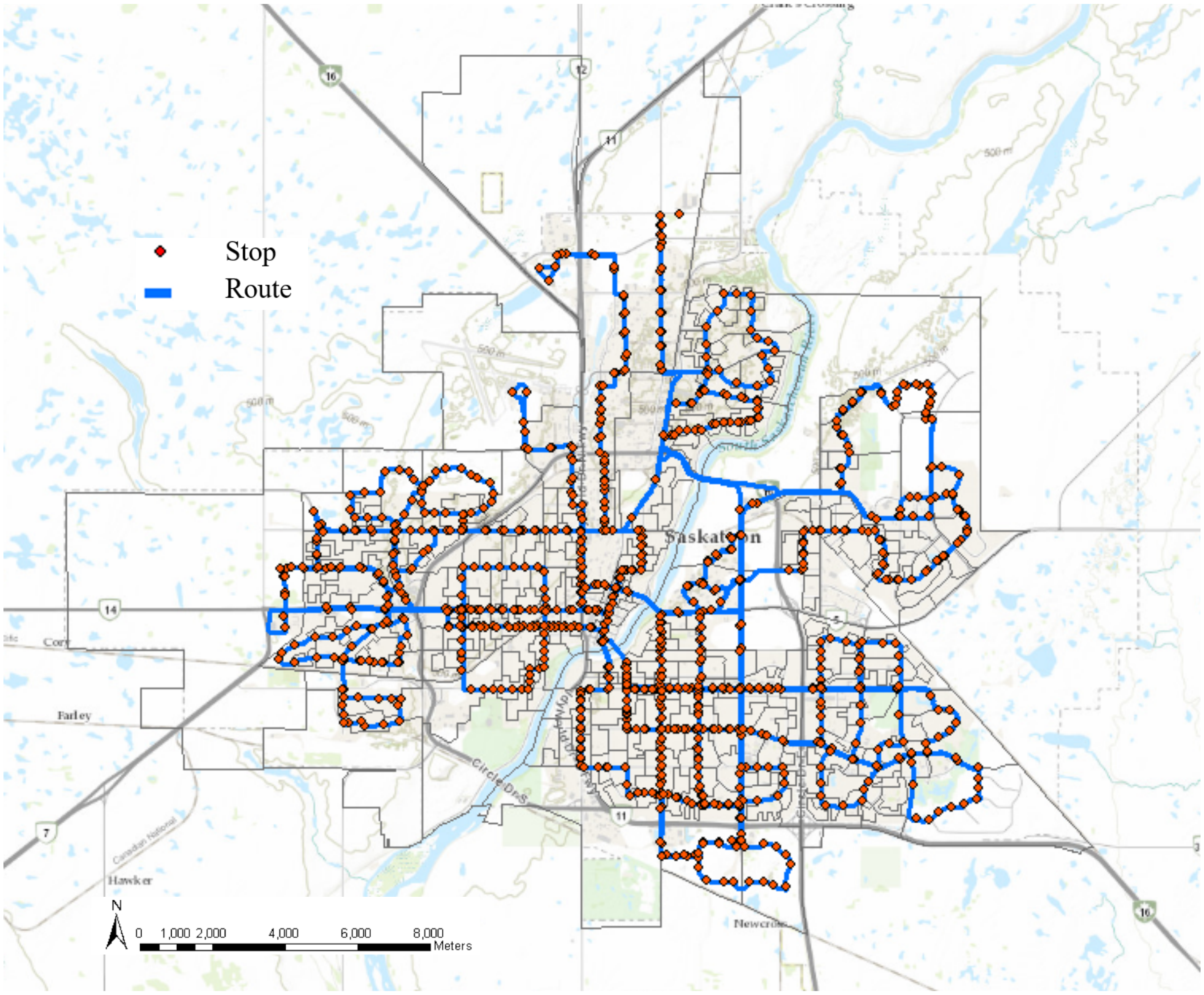


Figure 2: Saskatoon proposed BRT system configuration as of April 2018

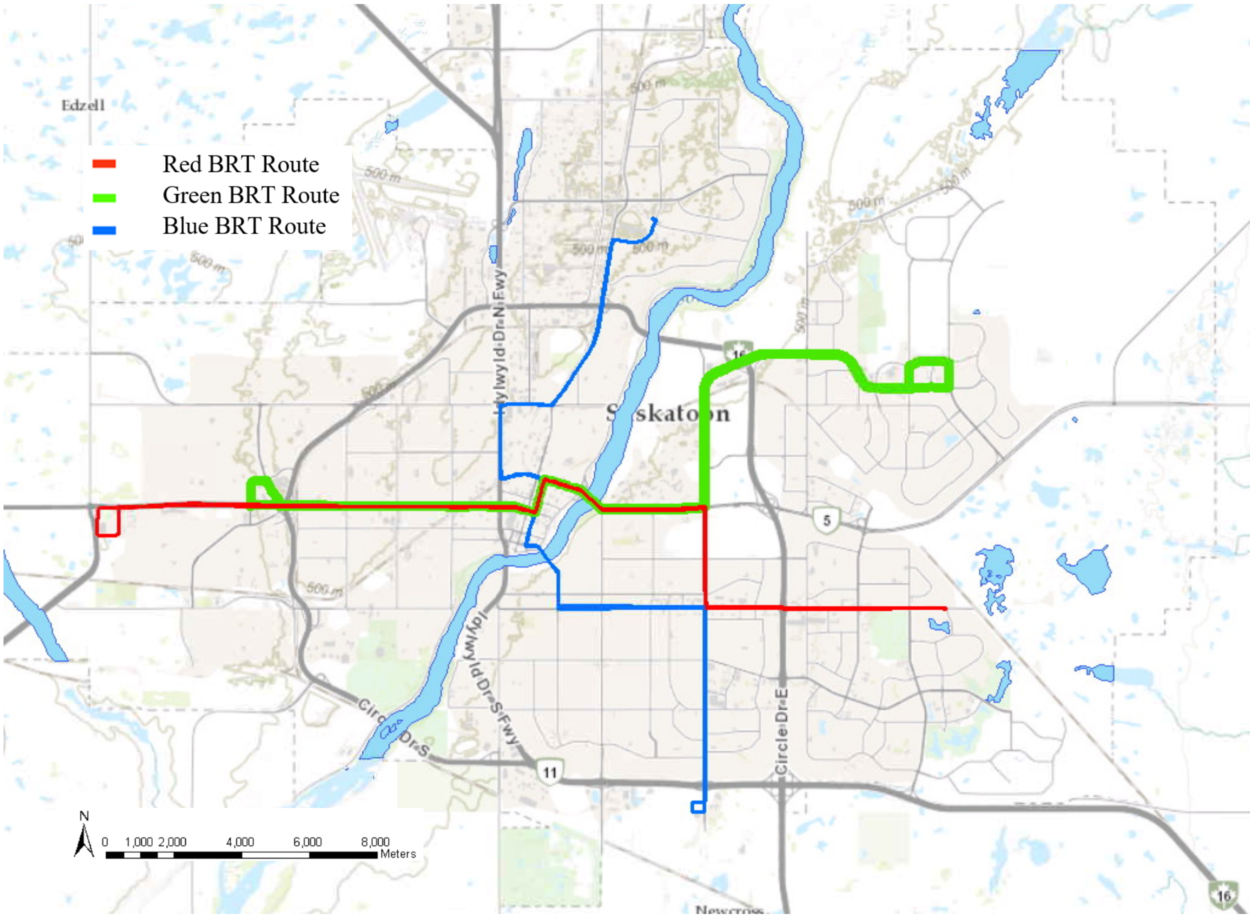


Figure 3: Proposed BRT routes as of April 2018

3.2 Data Collection

This research aims to evaluate several transit accessibility measures by examining their association with ridership levels at the dissemination area level. Several datasets were acquired to achieve this study goal. The first dataset is related to transit schedules and stop information in Saskatoon. This dataset, which is comprised of Saskatoon's transit system information as of September 2018, was obtained from the General Transit Feed Specification (GTFS) online database. This dataset includes the location of every stop, the routes served by each stop, and the departure direction and time of every weekly departure for each route at each stop. September was selected as it represents a typical month of transit usage in Saskatoon. During this period, school is in session, there are no major holidays, and the weather is typically temperate. Also, in Saskatoon major transit system changes can only occur on July 1 of a year because of a municipal bylaw. Therefore, the system will largely remain unchanged for the remainder of the year.

The second dataset is a Saskatoon Dissemination Area (DA) shapefile, which was extracted from the latest Statistics Canada Digital Boundary 2016 file (Statistics Canada 2016a). The third dataset contains socioeconomic data, which was also obtained from Statistics Canada (Statistics Canada 2016b). The fourth datasets were the road network and pedestrian sidewalk files obtained from the City of Saskatoon open data portal. The fifth dataset includes detailed ridership data at the stop level collocated from Saskatoon transit's fare box system for all the weekdays of September 2018. This comprehensive data set includes vehicle number, stop number, route number, passenger transit pass unique ID, fare type (i.e., pass, exact change, and transfer), and date and time of transaction. The sixth dataset is the 2013 Saskatoon Household Travel Survey (SHTS). SHTS is a disaggregate origin-destination survey that included 5% of the Saskatoon and region households, conducted in 2013. As with any household survey, this dataset included detailed data about home location, number of people per household, and trips made by individuals of each household. This dataset was used to observe and understand users walking patterns to transit including the average distance walked to reach a transit stop. Further, the data was used to calculate the distance decay function which was used in the accessibility measure computation.

The seventh and final dataset are the Walk Score and Transit Score values. These values were obtained from the Walk Score website using the APIs listed in Appendices A and B respectively. Walk Score computes their values based on a 152 m (500 ft) grid. Therefore, the Walk Score values obtained are the snapped latitude and longitude closest to the requested point. Once the Walk Score values and Transit Scores values were acquired for all 13,095 grid cell centroids, the scores for the DAs were computed by averaging the grid cell scores.

3.3 Transit Service Accessibility Measures Calculation

Accessibility to transit depends not only on the spatial location of the stop, but also on the level of provided service. This study examines several accessibility measure definitions. More specifically, local transit service accessibility, S_j, may be calculated in several ways. Consider the hypothetical scenario described in Table 1 in which three bus stops within a demand location's catchment area provide access to five routes.

Table 1: Transit service example

Route	**Stop 1** S_i ; 100 m	**Stop 2** S_i; 200 m	**Stop 3** S_i; 300 m
A	3/h	3/h	3/h
B		2/h	2/h
C			3/h
D	1/h	1/h	
E	1/h	1/h	

Based on the hypothetical example, five service measures may be defined:

1. **Stop count measure:** This measure counts the total number of stops within a demand location. In the hypothetical example, one stop can be reached within 190 m, and three stops are within 400 m.

2. **Coverage measure:** The service provided by each stop is determined by the number of different routes serving the stop. In the hypothetical example, Stop 1 provides access to three routes, Stop 2 provides access to four routes, and Stop 3 provides access to three routes. That is, $S_1 = 3$, $S_2 = 4$, *and* $S_3 = 3$. When using this measure to calculate accessibility, Stop 2 provides the most service.

3. **Frequency measure:** The service provided by each stop is determined by the number of departures over a time interval. In the example, $S_1 = 5$, $S_2 = 7$,*and* $S_3 = 8$. Stop 3 provides the most service when using this measure to calculate accessibility.

4. **Filtered coverage measure:** Routes are filtered out according to their distance from the demand location. Stop 1 provides the closest access to three routes, Stop 2 and Stop 3 each provide the closest access to one route. That is, $S_1 = 3$, $S_2 = 1$, $S_3 = 1$. When using this measure to calculate accessibility, Stop 1 provides the most service.

5. **Filtered frequency measure:** Departures on the same route from stops farther away from the demand location are not considered. Therefore, $S_1 = 5$, $S_2 = 2$, $S_3 = 3$. Stop 1 provides the most service when using this measure to calculate accessibility.

Each one of the model types described above involves service locations at different distances from the demand location. In each case, a weighting based on the distance between service location (bus stop) and demand location (i.e., a geographic area centroid) can be applied to the score contributed by each stop to calculate a gravity-based measure. Indeed, the smallest geographical sub-areas for which statistics such as population are available in Canada are Dissemination Areas (DAs). In total, 362 DAs are defined for Saskatoon in the 2016 census. All accessibility measures were computed using the bus schedules of the weekday morning peak period from 6 am to 9 am.

3.4 Dissemination Area Cropping and Partitioning

The 362 Dissemination Areas (DAs) defined by the 2016 census for the City of Saskatoon vary dramatically in size and population. While the majority of the DAs encompass developed areas of the city, some DAs around the city's periphery include vast swaths of undeveloped land. These enormous DAs have the potential to skew the results. Most DA polygons are roughly rectangular, but some have wildly irregular and concave shapes that cause the centroid to fall far outside the DA.

To address this issue, all DAs were partitioned into 100 m by 100 m square grid cells for further processing. Any grid cell farther than 1 km from a bus stop was discarded. Figure 4 shows the areas in blue that were clipped. The areas that were not considered are mainly undeveloped areas from DAs on the city's periphery, and some areas of undeveloped land within the city such as the University farmlands. An accessibility measure was then computed for each grid cell. The accessibility measure for every DA was computed by using the weighted average for all the cells that fall within the DA. Most cells within a DA will be 100 m by 100 m in size but some may be clipped by the DA boundary.

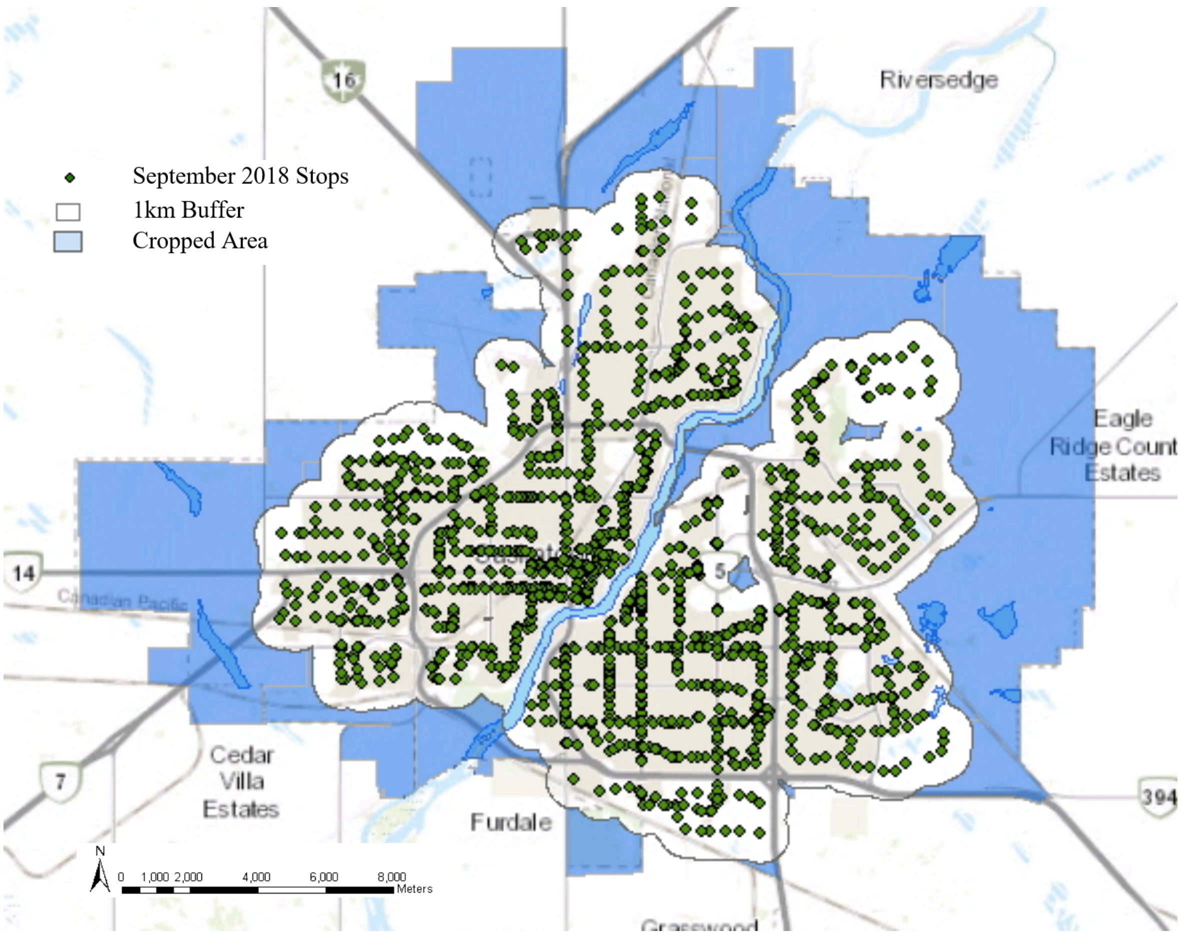


Figure 4: A 1 km buffer was placed around each September 2018 transit stop to crop large DAs

3.5 Processing the City of Saskatoon September 2018 Ridership Data

The goal of the data processing described in this section is to generate measures of transit usage per DA against which the computed accessibility metrics and other DA parameters can be correlated and used in a model to predict usage.

Since the definition of transit system accessibility used in this research is "ease of entry into the system", it is desirable to measure how many people are using the system. Other parameters such as trip length or number of transfers are not considered, although it is acknowledged they influence the decision of individual riders to use the system. However, due to data limitations, it was not possible to account these factors in this thesis.

Table 2: Data columns in ridership data

Column	Description
ID	Record Index - Not used in this analysis
Bus	Vehicle ID - Not used in this analysis
Date	Date of boarding
Time	Time of boarding
Route	Route ID
Stop	Stop ID
Pass	Pass ID
Employee	Employee ID (Driver?) - Not used in this analysis
Fare Type	E.g., Transfer, Cash Fare, Pass, etc.

The City of Saskatoon provided a PDF file containing a record of every boarding into the transit system for September 2018. The data categories recorded for each boarding into the system are listed in Table 2. The system boardings can be categorized into two types: entries, which are users embarking on new trips, and transfers, which are the continuation of an existing trip. Since transfers are not new entries into the system, they can be discarded. Since users often "tap" a card against a card reader to pay their fare, the term "tap" is used as a synonym for system entry. Entry taps or boardings were combined, and all transfers were discarded. A table which includes the number of all the different boarding types and their percent of the total taps can be found in Appendix C. Table 3 is a summary of the total entry taps and transfer taps.

Table 3: Boardings by type

Boarding Type	Counts	Percent of Boardings
Entry (Tap)	642,807	71.29%
Transfer	258,818	28.71%
All	901,652	100%

The data provided by the city contains a large number of taps with missing stop ID (the city is aware of this missing data). As well there are some taps with missing route ID. As the goal is to determine transit system usage per DA, taps with no stop ID cannot be located and must be discarded. In addition to taps with missing stop information, this stage of the filtering also discarded all taps belonging to ***Unknown*** routes. Many of the taps with missing route ID do have a valid stop ID. It is not desirable to include taps into the system from unknown routes (e.g. special event routes and high school routes) even if the stop ID is known. This is because ultimately, the goal of this research is to investigate the effect of the BRT changes on the existing standard transit system configuration. Therefore, taps belonging to special event routes and high school routes were discarded, resulting in 531,744 remaining taps for the next stage of filtering. A table with the number of good stops and bad stops per route can be found in Appendix D. Table 4 provides a summary of taps with valid stop IDs (good taps) and invalid stop IDs (bad taps) on known and unknown routes. Table 5 provides a summary of valid and invalid stop IDs for September 2018. A table containing the 20 stops with the highest valid tap count can be found in Appendix E.

Table 4: Summary of taps with stop IDs and without stop IDs on known and unknown routes

Routes	Good Taps	Bad Taps	Totals
Known	531,774 (82.72%)	90,751 (14.12%)	622,525 (96.85%)
Unknown	12,809 (2.00%)	7,473 (1.16%)	20,282 (3.15%)
All Routes	544,583 (84.72%)	98,224 (15.28%)	642,807 (100%)

Table 5: Summary of valid and invalid stop IDs for the September 2018 configuration

Stop ID Status	Tap Count	Percent
Valid	527,561	99.20%
Invalid	4,213	0.80%
Total	531,774	100%

3.5.1 Temporal Filtering by Day of Week

The 527,561 valid taps that were filtered from the tap data were then filtered by date. The results are listed in Table 6 and plotted in Figure 5.

Table 6: All September 2018 taps with valid stop ID and route ID

Date	Day	Taps	Date	Day	Taps
2018-09-01	Saturday	7,062	2018-09-16	Sunday	4,624
2018-09-02	Sunday	4,582	2018-09-17	Monday	24,677
2018-09-03	Monday	4,520	2018-09-18	Tuesday	24,939
2018-09-04	Tuesday	18,124	2018-09-19	Wednesday	25,488
2018-09-05	Wednesday	23,003	2018-09-20	Thursday	25,780
2018-09-06	Thursday	23,464	2018-09-21	Friday	22,518
2018-09-07	Friday	23,404	2018-09-22	Saturday	7,325
2018-09-08	Saturday	7,966	2018-09-23	Sunday	4,755
2018-09-09	Sunday	5,588	2018-09-24	Monday	24,632
2018-09-10	Monday	24,788	2018-09-25	Tuesday	24,999
2018-09-11	Tuesday	24,885	2018-09-26	Wednesday	25,719
2018-09-12	Wednesday	24,495	2018-09-27	Thursday	26,272
2018-09-13	Thursday	24,478	2018-09-28	Friday	24,222
2018-09-14	Friday	23,179	2018-09-29	Saturday	8,952
2018-09-15	Saturday	7,767	2018-09-30	Sunday	5,354
				Total Taps	527,561

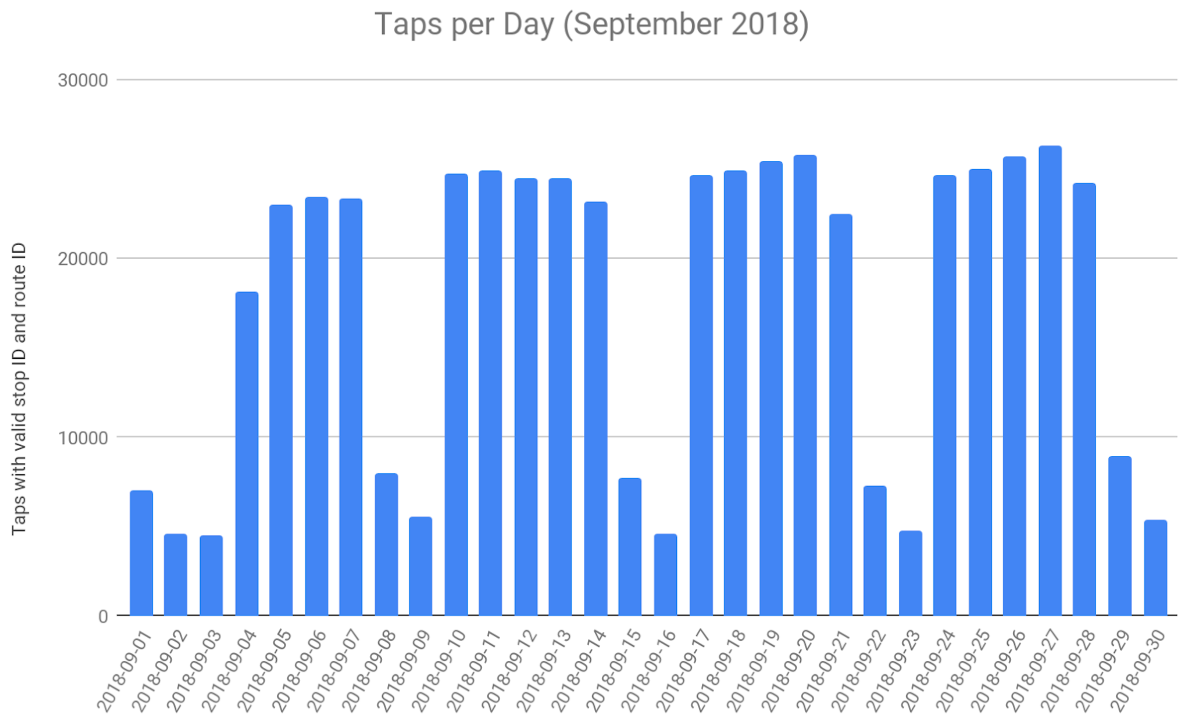


Figure 5: All September 2018 taps with valid stop ID and route ID

Monday September 3rd has a low tap count because it was the Labour Day holiday and the transit system was operating on a Sunday schedule. Counting Monday, September 3rd as a weekend day, the average number of taps on a weekend day was 6,226, and the average number of taps on a weekday was 24,161. This research concentrates on weekdays, because it is on those days that the planned BRT changes may be most impactful. Therefore, all weekend taps including Monday September 3rd were discarded, leaving 459,066 taps during the weekdays left for further processing.

Transit system accessibility may be good downtown or at the University, but it is accessibility at home that may be the determining factor in a decision to use transit. For this reason, it is desirable to map users to the DAs where they live. Therefore, all taps except those between

6 am and 9 am were discarded. This resulted in 110,894 taps during the morning period left for analysis. Table 7 shows the different types of valid taps observed during the morning hours of 6 am to 9 am on weekdays.

Table 7: Fare types for weekday taps between 6AM-9AM with valid stop ID and route ID

Fare Type	Taps	Percent of Total
Pass (multi ride card)	89,839	80.01 %
Rides (multi ride card)	11,535	10.40 %
Exact Fare	5,638	5.08 %
Override	2,200	1.98 %
Counter	834	0.75 %
Rides (disposable)	603	0.54 %
Cash Exceeded	218	0.19 %
Pass (disposable)	19	0.02 %
Override CNIB	7	< 0.01 %
ePurse (multi ride card)	1	< 0.01 %
Total	110,894	100 %

3.6 Ridership Level Calculation

When the taps per stop were joined to the DA, there were a large number of DAs with zero taps. This result was somewhat unexpected. However, most stops are located on major streets which also tend to be DA boundaries. In fact, most DAs only have stops on their boundaries. Depending on the accuracy of the stop locations in the GTFS database, and on the DA boundary locations in the Statistics Canada database, a stop could, by chance, fall into a DA on either side of the boundary. Because of this, there are many DAs containing no stops at all. Figure 6 illustrates an example of how the majority of transit stops fall along the DA boundaries in Saskatoon.

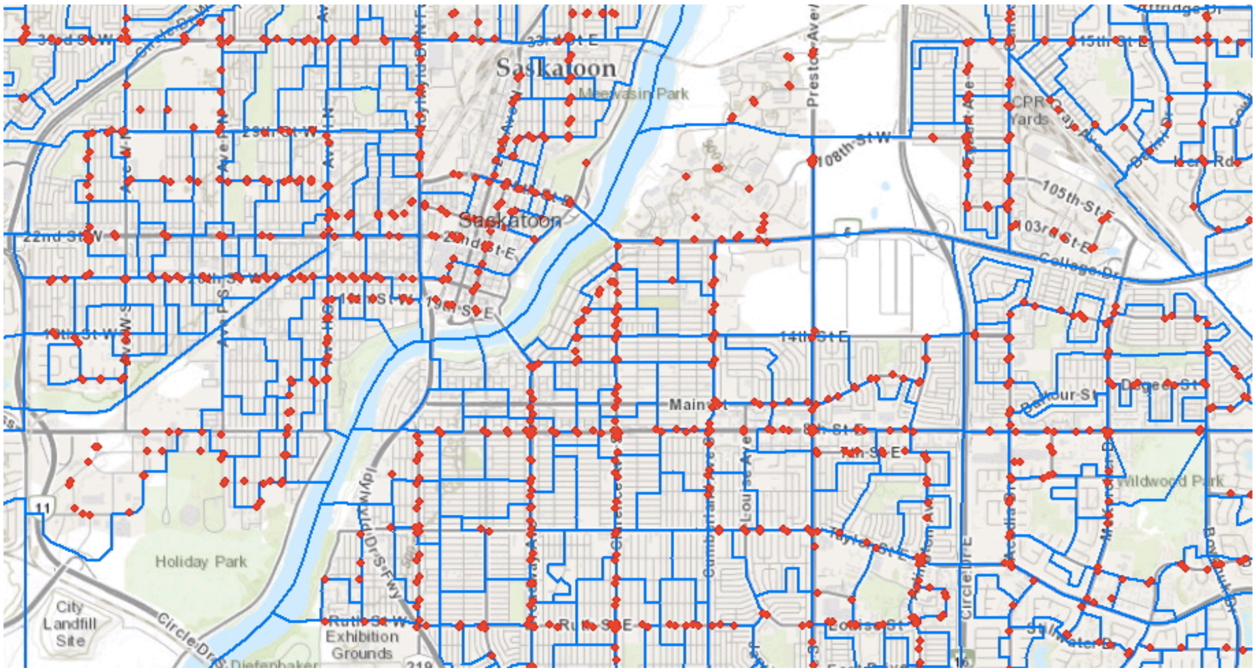


Figure 6: Many of the stops (red) fall along the DA boundaries (blue)

To solve this problem, a 400 m network buffer was created around every stop using the street network dataset. Intersections between the buffers and the DAs were computed. Each resulting intersection was weighted to give a geographic weight for number of transactions which occurred within the DA and summed to give the total ridership during the morning period for every DA and day. This was done to eliminate edge effect errors in estimating ridership at the DA level. Figure 7 and Figure 8 illustrate how the network constrained buffers look around the stops for the September 2018 transit route configuration and the proposed BRT configuration.

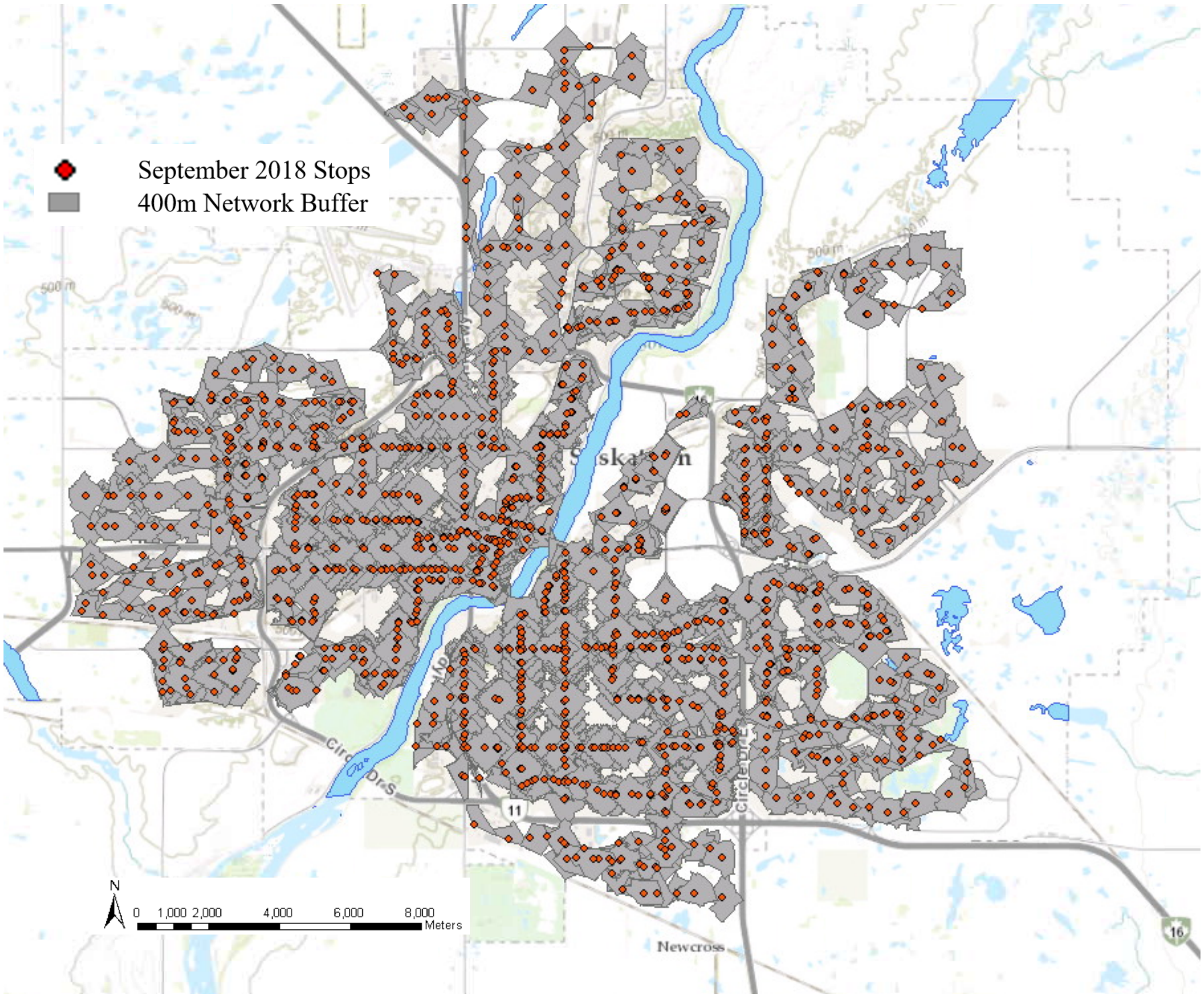


Figure 7: September 2018 stop locations with 400 m network buffers. There are 1,443 active stops.

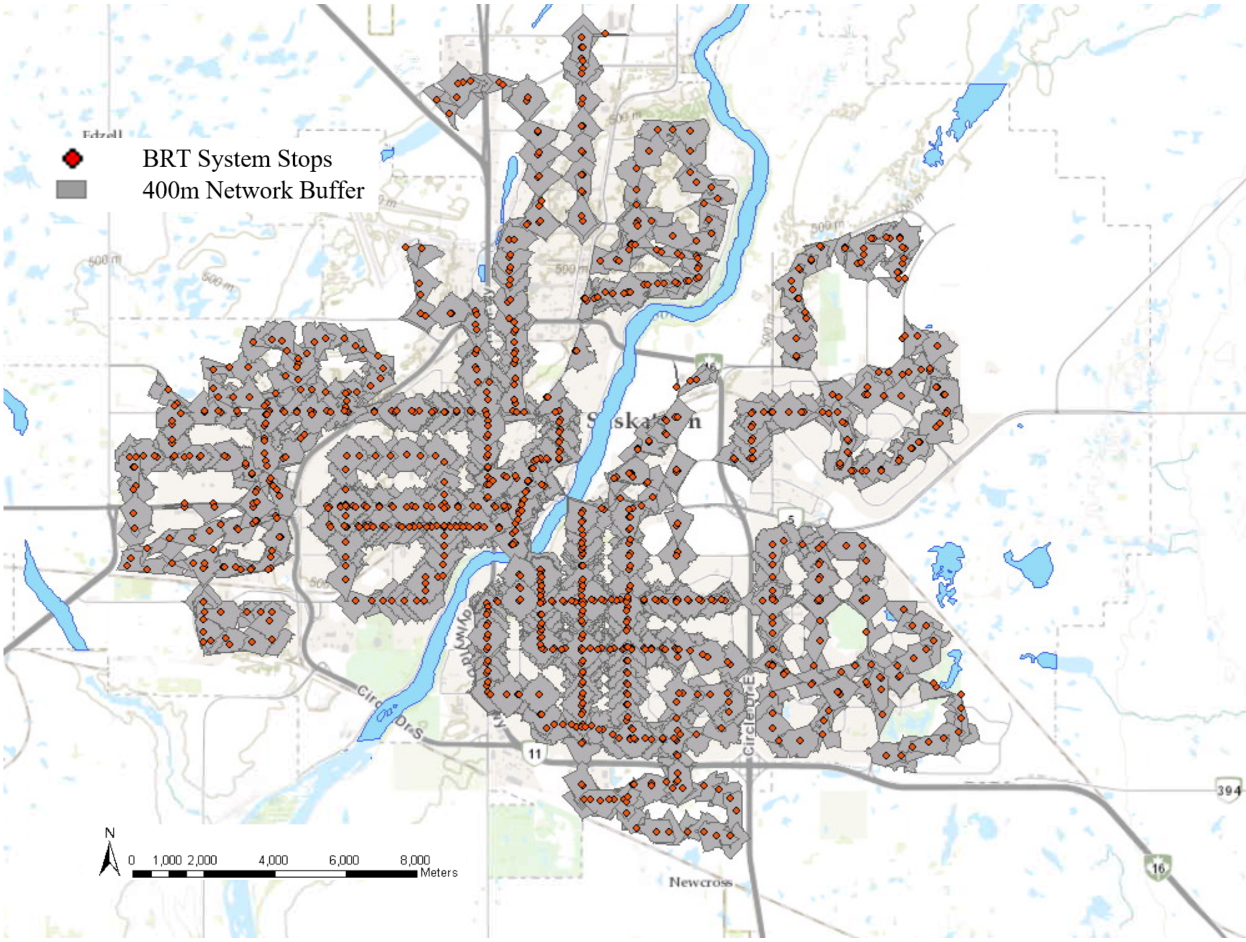


Figure 8: The proposed BRT configuration with 400 m network buffers around 994 stops

3.7 Model Development and Specification

Table 8 includes a list of variables that were prepared and tested for the inclusion in the final models. The table also presents the summary statistics of all the variables. After observing a skewed distribution of a number of key variables such as the dependent variable, number of taps per DA, all non-dummy variables were transformed into natural logarithmic form to approach normal distribution. The log-log model allows for the interpretation of model results in terms of elasticities, as has been done across the literature (Guerra and Cervero 2011b, Boisjoly et al. 2018, Taylor et al. 2009).

The external variables, as listed in Table 8, were prepared for the inclusion in the models. These variables were calculated at the DA level and include total population, occupied dwelling density, percent of aboriginal populations, average household size, percent of visible minorities, Walk Score, the percent of park and open space area, and road density per square kilometer. Walk Score is a measure of land use mix and is based on the number of services and amenities which can be accessed within an area, which should have a positive association with ridership. Total population occupied dwelling density and percent of visible minorities are also expected to have a strong positive association with ridership. In contrast, median income is expected to have a negative impact on transit service ridership. Additionally, several dummy variables were prepared for incorporation in the model. For example, a dummy variable for DAs with major transit stations was generated to control for other activities that are not residential, which can be found at these locations, such as the university campus main transit station. Another dummy variable was generated for DAs with high rise buildings, which is concentrated in a small area in the central business district (CBD).

A preliminary analysis was undertaken by exploring and comparing the general trends of ridership and other factors including the different accessibility measures. Having observed these relationships, an Ordinary Least Square (OLS) regression model (i.e., base model) was run first to explore the determinants of ridership at the DA level, while not including any of the developed accessibility measures in the models. An upward stepwise method for the dependent variable inclusion was used by removing insignificant variables once at a time. Therefore, different models were tested, while controlling for multicollinearity between variables, normality of residuals and the models' mean collinearity indicated by the variance inflation factor (VIF) indicator. Several external and dummy variables were found to be highly correlated or not significant, therefore, they were removed from the analysis. The final model achieved the highest R-squared value, while maintaining the maximum number of significant variables.

After developing the base OLS model, Moran's I test was computed to check for overall spatial autocorrelation of the OLS model's residuals. The results of the Moran's I test for the OLS model can be seen in Figure 9, which shows the spatial autocorrelation among the model residuals. Red DAs have spatial autocorrelation greater than the mean and blue DAs have spatial autocorrelation less than the mean. The test results were statistically significant ($p<0.01$) with a positive score, indicating that a spatial autocorrelation among the residuals exists and therefore are they not randomly distributed. The presence of spatial autocorrelation violates the OLS model requirement that all variables are independent. Therefore, a spatial lag model (SLM) and a spatial error model (SEM) were estimated to control for the spatial autocorrelation among residuals, while using the same set of variables.

The SLM model works by using a diffusion process in the dependent variable. It includes a spatially lagged dependent variable which is affected by the values of the dependent variables in adjacent locations. In contrast, the SEM model looks only at the spatial dependence in the errors of the observations and includes a spatial autoregressive error term that accounts for autocorrelation in the error with the weight's matrix. Spatial Error Models capture the influence of unmeasured independent variables. They evaluate how "clustering" of the dependent (output) variable can be accounted for using clusters of error terms (Mathew 2006).

With N observations and K independent variables, the SEM model is expressed as:

$$\boldsymbol{Y} = \boldsymbol{VC} + \boldsymbol{E}$$

where $\boldsymbol{Y}$ is an N x 1 vector of observations (dependent variable), $\boldsymbol{V}$ is and N x K matrix of observations and $\boldsymbol{C}$ is a K x 1 vector of regression coefficients, and $\boldsymbol{E}$ is an N x 1 vector of spatially autocorrelated error terms. $\boldsymbol{E}$ is given by:

$$\boldsymbol{E} = \lambda \boldsymbol{W}_E + \boldsymbol{u}$$

where λ is the autoregressive coefficient, $\boldsymbol{W}_E$ is a N x 1 vector of spatial lags for the errors, and $\boldsymbol{u}$ is another error term. A single prediction $\boldsymbol{y}_n$ (i.e., ridership in DA *n*) is therefore given by:

$$y_n = c_1 v_1 + c_2 v_2 + \ldots\ c_k v_k + e_n$$

where $\boldsymbol{c}_k$ is the coefficient for independent variable $\boldsymbol{v}_k$ and $\boldsymbol{e}_n$ is the spatially autocorrelated error term for observation $\boldsymbol{n}$.

Previous studies have examined several spatial weighting matrices and have found that the Queen's weighting approach at a similar spatial scale is superior (Ibeas et al. 2012). Therefore, in estimating the two models, a Queen-Contiguity matrix was applied, which incorporates DAs that share an edge or a vertex, while estimating the models' coefficients.

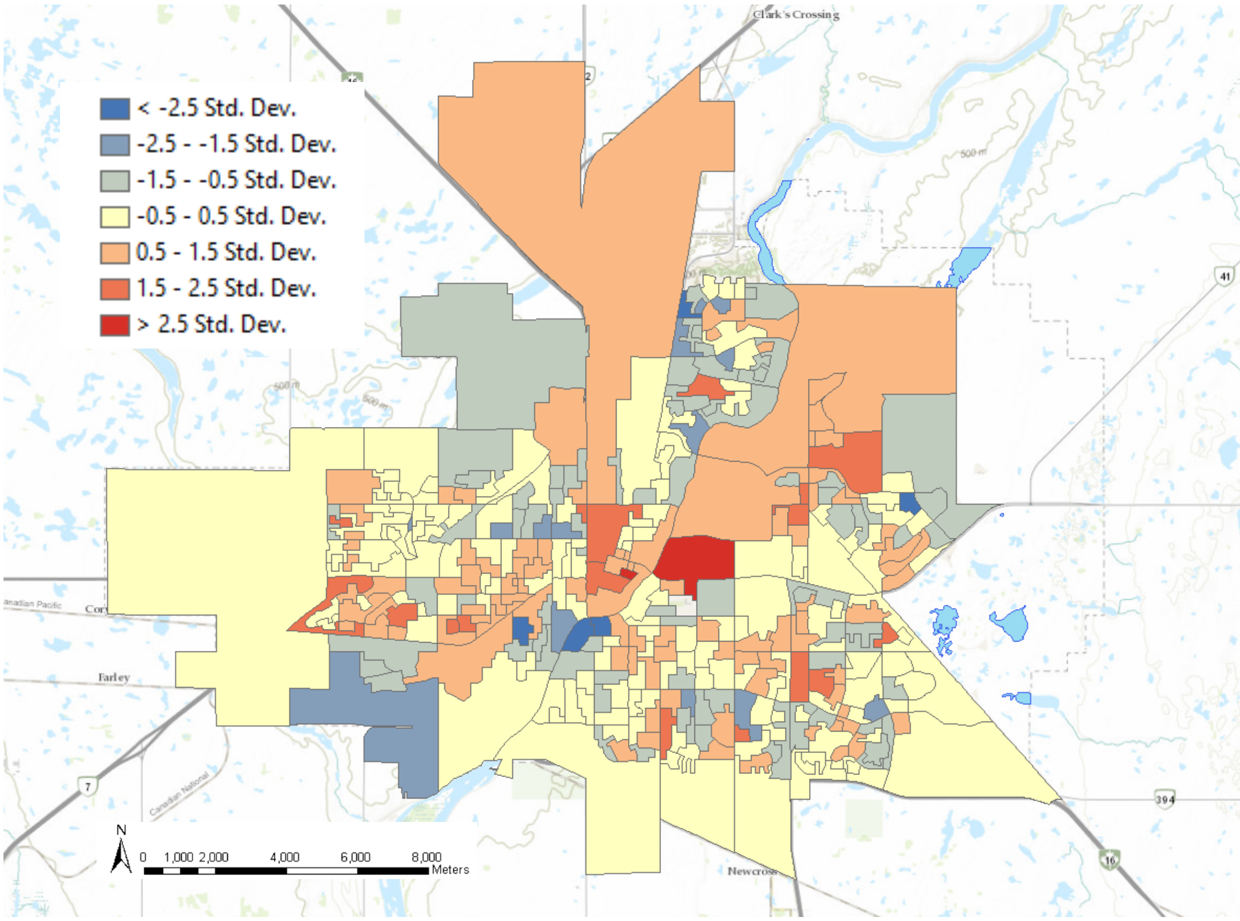


Figure 9: Spatial autocorrelation among the OLS model's residuals

After the development of the base models, the second phase of the research started by inputting the computed accessibility measures one at a time to the models, while comparing the fit of the different models in relation to the actual ridership. The best model was determined as the one having the lowest Akaike Information Criterion (AIC) value and the highest R2 and Log-Likelihood (LL) value. These values aid in model comparison when variables, like those in the base model, remain constant. After identifying the accessibility measure with the best fit, the final models were examined further.

Table 8: Variables considered in Ridership Models

Variable Name	Description	Mean	Std.Dev.
Taps per day (ln)	Number of taps per DA per day during the morning period between 6:00 am and 9:00 am (Dependent Variable).	3.4	1.12
External Variables			
Park percent (ln)	The percent of park area located within each DA	-3.1	4.5
Roads per km^2 (ln)	The total length of roads per square kilometer located within each DA	3.1	0.5
Aboriginal percent (ln)	Total percentage of self-identifying Aboriginal people living within each DA	2.2	0.9
Average household size (ln)	The average number of people living within a house for each DA	0.9	0.2
Area (ln)	The total area of each DA	12.2	0.9
Occupied dwellings (ln)	Total number of occupied dwellings within each DA	5.4	0.6
Occupied dwelling density (ln)	The total number of occupied dwellings in each DA divided by the DA area	7.0	0.8
Visible minorities (ln)	The percentage of the population in each DA that is a visible minority	2.6	1.0
Renter percent (ln)	The percentage of people who rent their home within each DA	2.9	1.1
Total population (ln)	The total population in each DA	6.3	0.6
Population density (ln)	The total population in each DA divided by the DA area	7.9	0.8
Owner percent (ln)	The percentage of people who own their homes within each DA	3.9	0.8
Median income (ln)	The household median income for each DA	11.2	0.9
Married percent (ln)	The percentage of the population that is married in each DA	3.8	0.4
Employment percent (ln)	The percentage of people within each DA who are employed	4.1	0.4
Post-secondary education (ln)	The percentage of people within each DA who have a post-secondary education. This includes trades, certificates diplomas, bachelor’s degrees and higher	3.7	0.4
No high school education (ln)	The percentage of the population who did not complete high school within each DA	2.4	0.5

Variable Name	Description	Mean	Std.Dev.
Major stations	A dummy variable. DAs with major stations are assigned 1 and DAs without major stations are assigned 0	0.1	0.15
High-rise	A dummy variable. DAs with residential high-rises are assigned 1 and DAs without residential high-rises are assigned 0	0.1	0.28
Walk Score (ln)	The average Walk Score value for each DA	3.7	0.5
Transit Accessibility Measures			
Transit score (ln)	The average transit score calculated for each DA	3.8	0.2
Stop Count 400 (ln)	A stop count measure calculated using a 400m network buffer	1.7	0.6
Stop Count 400 / 250 (ln)	A stop count measure calculated using a 400m network buffer and a Butterworth filter distance decay with a bandpass value=250m	1.0	0.6
Stop Count 532 (ln)	A stop count measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS)	2.1	0.6
Stop Count 532 / -0.003 (ln)	A stop count measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS) and an exponential distance decay of -0.003 calculated from SHTS.	0.8	0.6
Coverage 400 (ln)	A coverage measure calculated using a 400m network buffer	2.7	0.8
Coverage 400 / 250 (ln)	A coverage measure calculated using a 400m network buffer and a Butterworth filter distance decay with a bandpass value=250m	2.0	0.8
Coverage 532 (ln)	A coverage measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS)	3.1	0.8
Coverage 532 / -0.003 (ln)	A coverage measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS) and an exponential distance decay of -0.003 calculated from SHTS	2.1	0.8
Frequency 400 (ln)	A frequency measure calculated using a 400m network buffer	2.8	0.8
Frequency 400 / 250 (ln)	A frequency measure calculated using a 400m network buffer and a Butterworth filter distance decay with a bandpass value=250m	2.1	0.8
Frequency 523 (ln)	A frequency measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS)	3.2	0.8
Frequency 523 / -0.003 (ln)	A frequency measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS) and an exponential distance decay of -0.003 calculated from SHTS	2.1	0.8

Variable Name	Description	Mean	Std.Dev.
Filtered Coverage 400 (ln)	A filtered coverage measure calculated using a 400m network buffer	1.4	0.6
Filtered Coverage 400 / 250 (ln)	A filtered coverage measure calculated using a 400m network buffer and a Butterworth filter distance decay with a bandpass value=250m	1.1	0.6
Filtered Coverage 532 (ln)	A filtered coverage measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS)	1.6	0.6
Filtered Coverage 532 / -0.003 (ln)	A filtered coverage measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS) and an exponential distance decay of -0.003 calculated from SHTS	0.9	0.6
Filtered Frequency 400 (ln)	A filtered frequency measure calculated using a 400m network buffer	2.0	0.7
Filtered Frequency 400 / 250 (ln)	A filtered frequency measure calculated using a 400m network buffer and a Butterworth filter distance decay with a bandpass value=250m	1.5	0.7
Filtered Frequency 532 (ln)	A filtered coverage measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS)	2.2	0.7
Filtered Frequency 532 / -0.003 (ln)	A filtered frequency measure calculated using a 532m network buffer (85^{th} percentile of walking distance to transit calculated from SHTS) and an exponential distance decay of -0.003 calculated from SHTS	1.4	1.7

4 RESULTS

4.1 Correlation with Actual Ridership

Transit ridership depends on factors that are not related to transit service quality, such as residents' trip purpose, distance, and socioeconomic status. However, it should be noted that there is a relationship between transit service supply and demand. While the demand for transit is influenced by its supply, transit supply itself is adjusted by transit agencies in response to demand changes over time to provide an efficient service. In other words, demand and transit accessibility measures should be correlated.

A Pearson's correlation coefficient was computed to understand the relationship between ridership and each accessibility measure in Saskatoon at the DA level. The total number of bus pass entry taps between 6 am and 9 am is illustrated in Figure 10. The values in the Figure have been normalized by the population of each DA. Z-scores were calculated for a selection of the accessibility measures identified in the literature review to allow for visual comparison. Figure 11 illustrates the results of each of the five types of accessibility measures calculated with a 400m network buffer and a distance decay using a Butterworth filter with a 250 m bandpass value. The Figure also illustrates the correlation results for Walk Score's Transit Score and the Enhanced Two Step Floating Catchment Area (E2SFCA) measure as suggested by Langford et al. Table 9 lists the results of the correlation tests. As shown in the table, the best correlation (Pearson's r value of 0.590) was obtained by the ***Filtered Frequency 400 / 250*** measure. The poorest correlation (Pearson's r value of 0.036) was from the E2SFCA measure.

The low Pearson's correlation coefficient for the E2SFCA measure can be explained as follows. Langford et al. looked at regional accessibility to transit in a rural location. Because of the lower population density, in their measure population was used as a demand rather than a supply. That is, a higher population reduces available transit supply much like it can fill a physician's practice. However, unless the busses are consistently full, the population does not reduce the supply of the service. That is why the E2SFCA method performed poorly in Saskatoon. It should be noted that the best preforming measure, the ***Filtered Frequency 400 / 250***, is *identical* to the E2SFCA measure *except* for the final step where the service is divided by the population. Based on these results, the E2SFCA measure was removed from the analysis prior to model development.

Table 9: Accessibility measure correlation with transit ridership

Accessibility Measure	Pearson's *r*
E2SFCA	0.036
Transit Score	0.313
Stop count 400 / 250	0.423
Coverage 400 / 250	0.534
Frequency 400 / 250	0.551
Filtered Coverage 400 / 250	0.552
Filtered Frequency 400 / 250	0.590

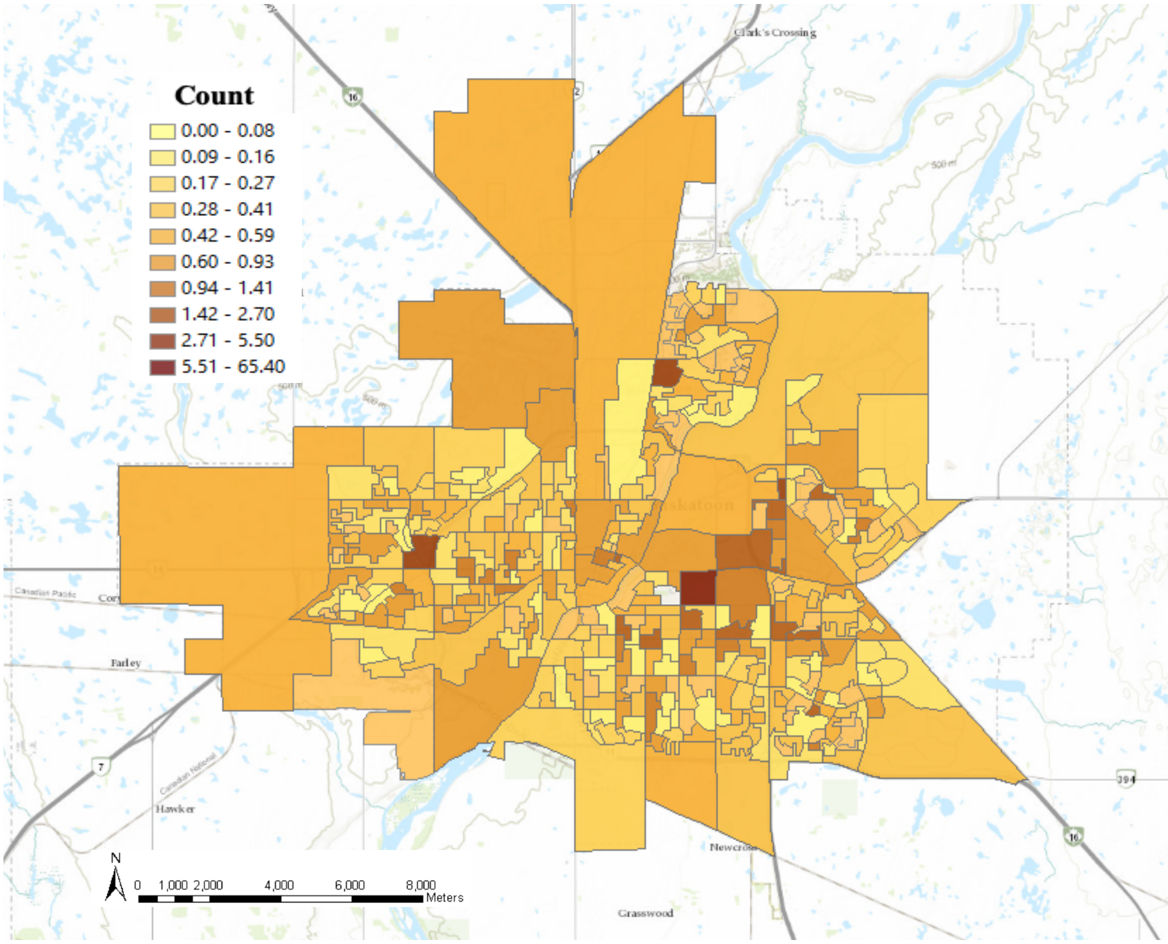


Figure 10: Entry Taps per DA between 6 am and 9 am normalized by the population

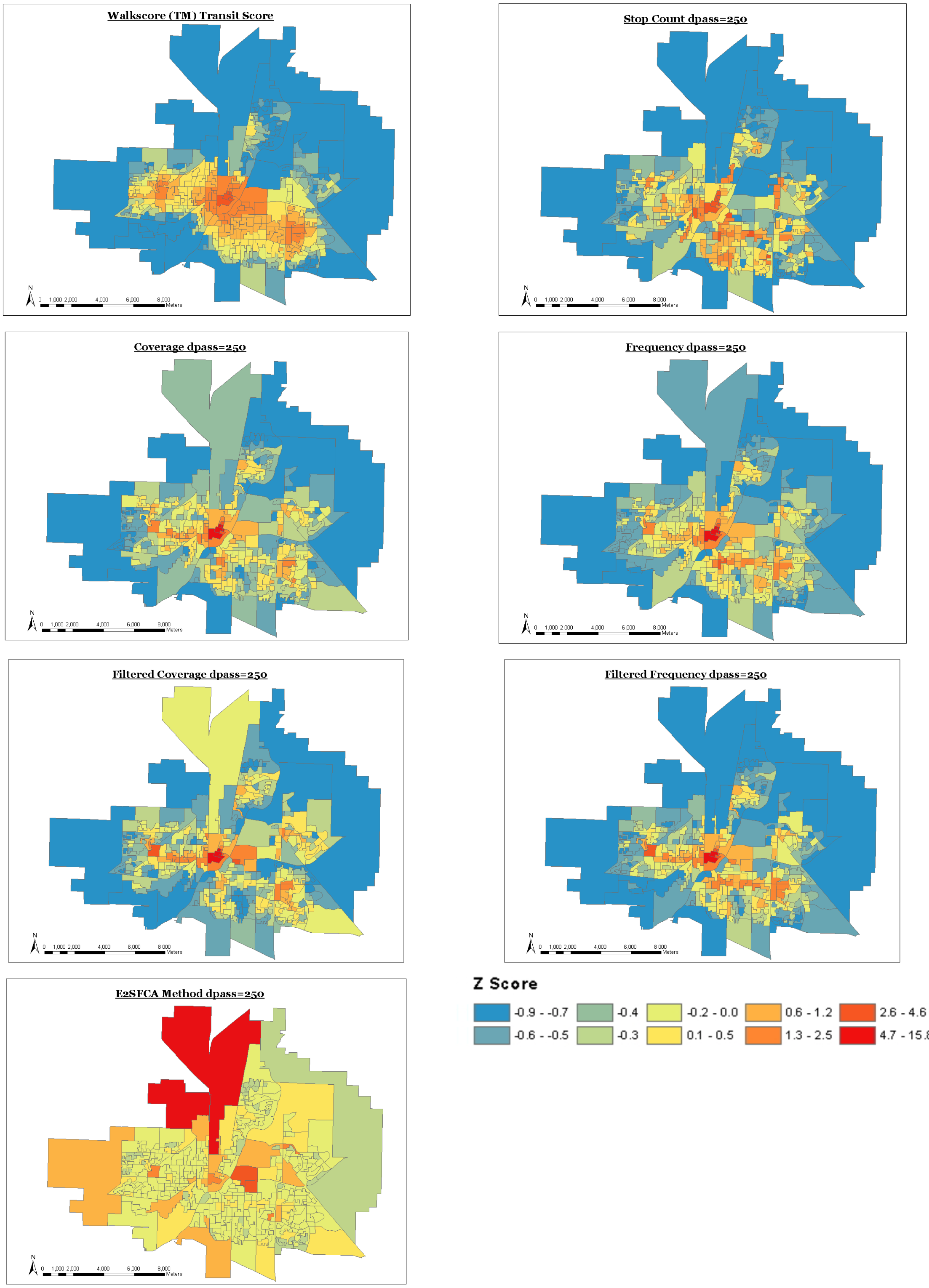


Figure 11: Accessibility measures calculated with a 400 m network buffer and a distance decay function using a Butterworth filter with a 250 m bandpass value

4.2 Base Models

Table 10.A lists the results of the ordinary least squares (OLS) base model, Table 10.B lists the results of the spatial lag (SLM) base model, and Table 10.C lists the results of the spatial error (SEM) base model. The Z-value and the statistical significance are reported in the tables along with the independent variable coefficients. The models are based on 360 observations (DAs) due to the removal of two outliers, which were DAs with zero population.

As shown by Table 10.A, the base OLS model contains 360 records and explains 65% of the variance in the log of the total transit usage. This proportion of explained variance is comparable to other ridership prediction models presented in the literature (Taylor et al. 2009). Further, in the base OLS model, all the variables show the expected sign and direction as predicted by theory. For example, for every 1% increase in Walk Score, a 0.9% increase in ridership is predicted. This is expected as Walk Score measures land use mix, and it is based on the number of local services and amenities that can be reached within a walking distance. In other words, more walkable areas will enjoy higher ridership levels.

In the OLS model, the size of dissemination areas is positively associated with higher ridership. This is because in Saskatoon there has been considerable growth of population in large DAs at the edges of the city. Dwelling density has a positive and statistically significant association with ridership. A 1% increase in dwelling density is associated with a 0.7% increase in transit ridership, while keeping all other variables constant at their mean values. In contrast, a 1% increase in the household median income is associated with a 1.4% decrease in transit ridership. These findings are reasonable and previously reported in the literature (Miller et al. 2018).

Other socioeconomic factors were also important determinants of transit ridership. More specifically, the percent of visible minorities and percent of people with post-secondary education is positively associated with transit usage. As shown, a 1% percent increase in visible minorities and a 1% increase in people with post-secondary education is associated with 0.1% and 0.7% increases in ridership, respectively. This is expected since visible minorities in Canada are usually less able to afford personal automobiles. For percent of people with post-secondary education, the University of Saskatchewan is a major education and employment center relative to the size of the city, and it is responsible for a significant percentage of transit users in Saskatoon. In fact, the University offers discount transit passes for its faculty and staff, and all undergraduate and graduate students are required to purchase bus passes as part of their tuition.

The model also predicts that transit usage increases with the increase in average household size. In other words, a 1% increase in the average household size is associated with a 2.1% increase in transit ridership. This is reasonable within the context of Saskatoon, where the transit ridership is mainly dominated by younger populations and students. Finally, DAs with a major transit station have a positive association with transit ridership, increasing predicted ridership by 2%.

Table 10.B lists the results of the base spatial lag model (SLM). As shown in the table, all of the variables' signs and magnitudes closely follow the OLS model described previously. However, the SLM coefficients cannot be interpreted directly, due to the existence of the "Weighting Taps" variable in the model, which accounts for the spatial correlation between DAs. Weighting Taps shows a statistically significant positive association with ridership, illustrating the

need for using such a model. Similarly, all the SEM model variables' signs and magnitudes listed in Table 10.C follow the OLS model closely with a few exceptions. However, the spatial error coefficient, LAMBDA, shows a statistically significant positive association with ridership.

Table 10: Base models

	A. OLS - Ordinary Least Square		B. SLM - Spatial Lag Model		C. SEM - Spatial Error Model	
	Coeff.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat
Constant	-7.41	-3.28	-9.87	-4.77	-7.22	-3.30
Walk Score (ln)	0.89	9.05	0.71	7.57	0.91	7.10
Area (ln)	1.12	15.6	1.10	16.7	1.07	17.2
Dwelling Density	0.67	7.42	0.65	7.94	0.45	5.50
Average Household Size (ln)	2.14	5.99	1.93	5.92	1.54	4.55
Transfer Stations	1.96	7.61	1.75	7.47	1.41	6.47
Median Income (ln)	-1.42	-7.03	-1.16	-6.24	-1.09	-5.88
Visible Minorities (ln)	0.09	2.46	0.07	1.96	0.04*	1.11
Post-Secondary Education (ln)	0.74	3.51	0.60	3.12	0.42	2.16
Weighting Taps (ln)			0.40	7.74		
LAMBDA					0.61	10.85
Adjusted R2	0.65					
AIC	732.5		685.2		680.0	
LL	-357.2		-332.6		-331.0	

Number of records = 360
Most Values are significant at 99%, * Significant at 90%

Table 11 lists the fit of ridership prediction models with and without the transit accessibility measures included. The accessibility measures were input into to the OLS, SLM and SEM base models one at a time. Because the independent control variables remained the same in each of the base models, the performance of the models could be compared using the AIC and LL values respectively. In total, 21 local transit accessibility measures were tested. As shown in the table, the base models that did not include an accessibility measure had the poorest fit to the data.

Models with the ***Filtered Frequency 400 / 250*** accessibility measure included performed the best. Recall this accessibility measure was calculated using a filtered frequency method with a 400 m network buffer and a distance decay based on a Butterworth filter with a bandpass value of 250 m. Other models that performed well were developed using the ***Frequency 400 / 250***, ***Filtered Frequency 532 / -0.003,*** and the ***Filtered Coverage 400 / 250*** measures. In contrast, according to the SLM and SEM models, the worst performing models were developed using Walk Score's ***Transit Score***, Filtered ***Coverage 532****,* and stop count measures including ***Stop Count 532***.

The 532 m network buffer was calculated using the 85th percentile of walking distance to transit obtained from SHTS. Overall, the models computed using gravity-based measures (distance decay) performed better than models that did not. Additionally, the results show some variation between measures in terms of performance based on the type of modelling technique (OLS vs. SLM vs. SEM), with only a few exceptions.

Table 11: Overall fit of models with and without transit accessibility measures included

	A. OLS - Ordinary Least Squares			B. SLM - Spatial Lag Model		C. SEM - Spatial Error Model	
	Adjusted R2	**AIC**	**LL**	**AIC**	**LL**	**AIC**	**LL**
Base model	**0.650**	**732.5**	**-357.2**	**685.2**	**-332.6**	**680.0**	**-331.0**
Models with transit accessibility measures							
Stop Count 400 (ln)	0.694	685.3	-332.7	646.7	-312.4	594.3	-287.1
Stop Count 400 / 250 (ln)	0.675	666.9	-323.5	626.2	-302.1	569.2	-274.6
Stop Count 532 (ln)	0.682	702.9	-341.4	660.9	-319.5	624.9	-302.5
Stop Count 532/-0.003 (ln)	0.711	680.1	-330.05	637.2	-307.58	583.8	-281.9
Coverage 400 (ln)	0.730	640.2	-310.1	609.9	-293.9	539.5	-259.7
Coverage 400/250 (ln)	0.746	618.7	-299.4	585.8	-281.9	511.1	-215.6
Coverage 532 (ln)	0.713	663.5	-321.6	630.9	-304.5	586.74	-283.4
Coverage 532/-0.003 (ln)	0.734	641.7	-310.7	608.7	-293.4	545.85	-262.9
Frequency 400 (ln)	0.745	618.8	-299.4	590.3	-284.1	528.6	-254.3
Frequency 400/250 (ln)	0.762	594.2	-287.1	560.6	-269.3	497.0	-238.3
Frequency 532 (ln)	0.731	648.0	-314.0	616.4	-297.2	573.2	-276.6
Frequency 532/-0.003 (ln)	0.744	623.7	301.9	591.7	-284.8	530.9	-255.5
Filtered Coverage 400 (ln)	0.710	669.0	-324.5	633.7	-305.9	603.2	-291.6
Filtered Coverage 400/250 (ln)	0.752	613.0	-296.5	578.7	-278.3	508.0	-244.2
Filtered Coverage 532 (ln)	0.698	690.9	-335.4	654.2	-316.1	640.9	-310.5
Filtered Coverage 532/-0.003 (ln)	0.736	624.9	-311.5	609.4	-293.7	559.0	-269.5
Filtered Frequency 400 (ln)	0.748	635.1	-307.5	609.4	-293.7	589.9	-284.9
Filtered Frequency 400/250 (ln)	0.794	541.2	-299.5	515.4	-246.7	476.6	-228.3
Filtered Frequency 532 (ln)	0.714	663.3	-321.64	653.4	-306.7	627.5	-303.8
Filtered Frequency 532/-0.003 (ln)	0.756	609.64	-294.9	585.1	-281.5	552.4	-266.2
Transit score (ln)	0.669	713.4	-346.7	678.0	-328.0	671.7	-325.8

Number of records = 360, all the models includes the variables used in the base model (Table 10)
The best performing models are highlighted in **gray.**

Using the previous SEM model coefficients listed in Table 10, Figure 12 illustrates the log of the predicted values of ridership in terms of the number of taps per DA for a few selected models. It also shows the log of the actual ridership in terms of the number of actual taps (the dependent variable used in the model). With the addition of the accessibility measures, the model robustness, based on the Lagrange Multiplier Test, was significant only for the SEM model. Therefore, the SEM model was chosen as the recommend modelling approach.

In Figure 12, the DAs are grouped into 10 percentile ranges (low to high ridership) and plotted. The Y-axis value is the mean of the natural log of the observed ridership for the DAs in the percentile group. This was done to facilitate the comparison between accessibility measures for each percentile. In the Figure, the green line represents the actual ridership levels, while the orange line shows the base model's predicted ridership. As shown by the Figure, the base model's ridership predictions are furthest from the actual ridership values. The pink line shows the predicted values using the ***Filtered Frequency 400 / 250*** accessibility measure. This model performs the best when predicting ridership in low and high ridership DAs.

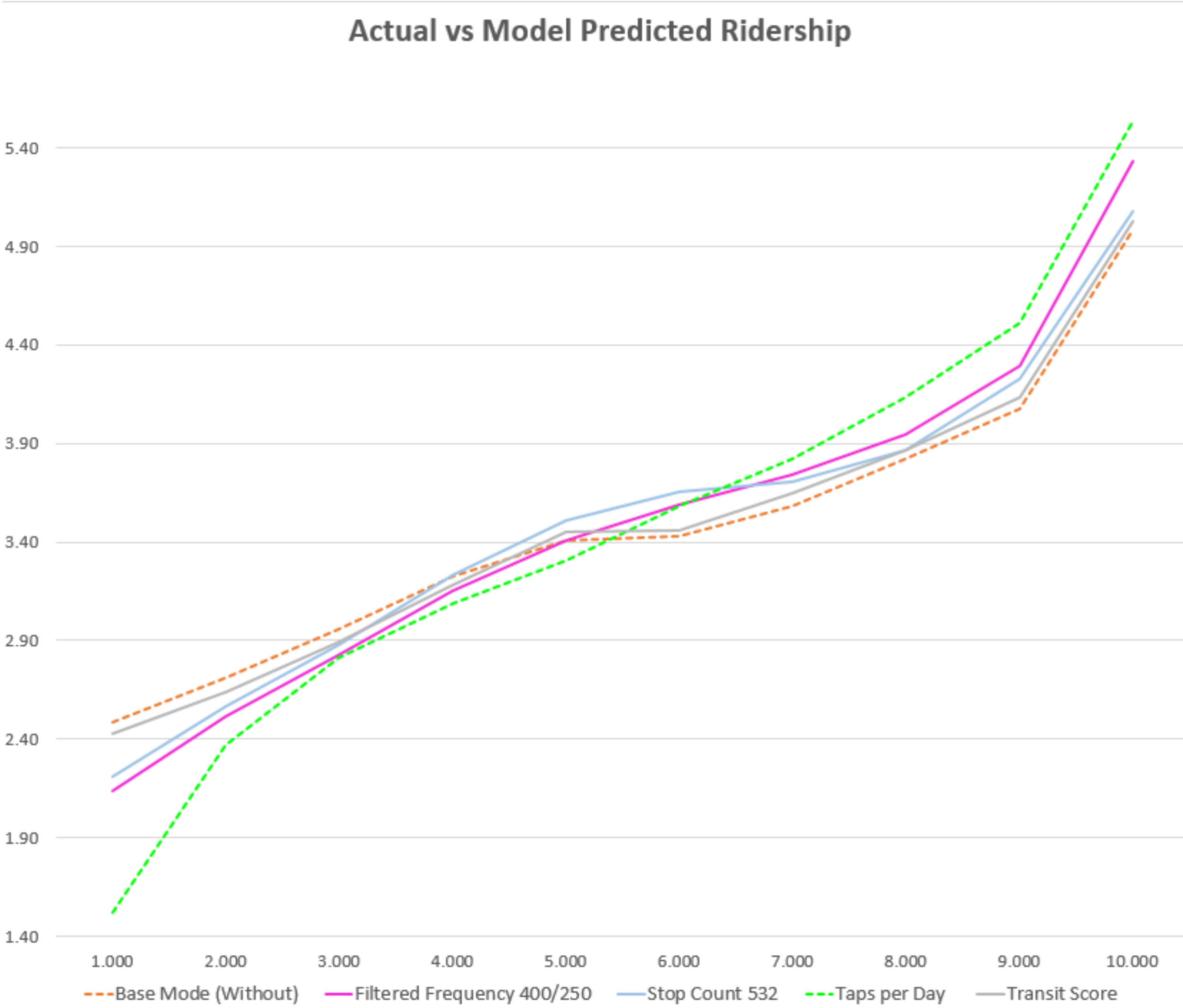


Figure 12: Actual ridership in terms of the log of number of taps per DA and the log of the predicted values of ridership for a few selected SEM models

Table 12 lists the best performing models' results with the ***Filtered Frequency 400 / 250*** accessibility measure. As listed in the Table, most of the coefficients' directions follow those of the first set of models in Table 10, with a few exceptions in terms of magnitude and significance. Figures 13 to 15 illustrate spatial autocorrelation among the models' residuals. Figure 13 shows the Local Indicator of Spatial Association (LISA) significance map for spatial autocorrelation for the OLS, SLM, and SEM models with the ***Filtered Frequency 400 / 250*** accessibility measure included. LISA helps to identify areas with significant spatial autocorrelation. In the figure, the grey colour shows areas with insignificant spatial autocorrelation, while the green highlights areas with significant spatial autocorrelation. Figure 14 shows the type of association and clustering for the significant locations. More specifically, the Figure illustrates areas with positive local spatial autocorrelation as high-high (red) or low-low (blue). It also shows areas with negative local spatial autocorrelation as high-low (light red) and low-high (light blue). Figure 15 shows scatterplots of the relationship between spatially lagged residuals (Y-axis) and the original residuals (X-axis) and the local Moran's I value for the models.

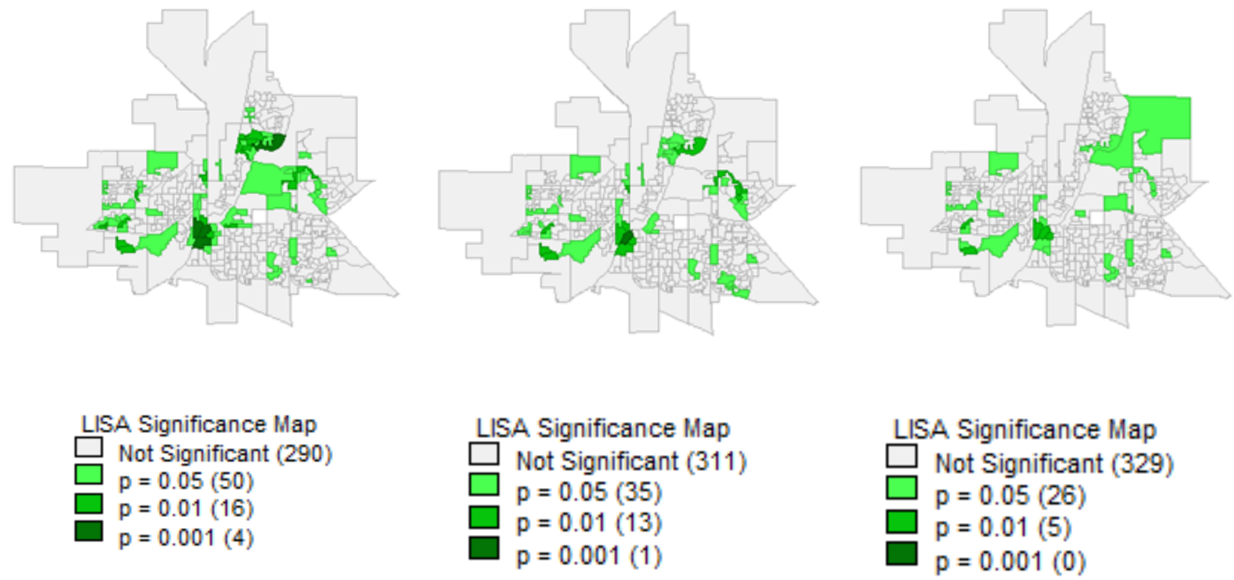
As shown by the Figures, the SEM model produces a better estimation of ridership with fewer areas of spatial autocorrelation in the error term. The LISA significance and cluster maps in Figure 13 and Figure 14 illustrate that there are fewer significant clusters produced by the SEM model compared to the other two models. For example, the number of significant clusters declined from 4 to 0 at the 99.9% significance level by switching from the OLS to the SEM model (Figure 13). For the locations with significant spatial autocorrelation, Figure 14 shows a mix of positive and negative associations between DAs and their neighbours. In addition, the cluster map shown in Figure 14 illustrates that there are fewer significant clusters for the SEM model compared to the other models. Finally, the local Moran's I scatterplot shown in Figure 15 shows an almost horizontal line with no spatial relationship, between ridership predicted at different DAs and their neighbours, in terms of the error. The global Moran's I test was also not significant ($p>0.05$) for the SEM model, indicating that this model with ***Filtered Frequency 400 / 250*** measure captured all the spatial variation while predicting ridership. This indicates that an SEM model with a ***Filtered Frequency 400 / 250*** accessibility measure included provides a more accurate estimation of ridership while reducing the impacts of spatial errors. Further, this will help transit authorities to understand the relative impact of local accessibility changes in the transit system.

Table 12: Best fit model with *Filtered Frequency 400 / 250*

	A. OLS - Ordinary Least Square		**B. SLM - Spatial Lag Model**		**C. SEM - Spatial Error Model**	
	Coeffi.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat
Constant	-12.69	-7.21	-13.78	-8.26	-13.09	-7.78
Walk Score (ln)	0.26	3.09	0.20	2.52	0.25	2.38
Area (ln)	1.08	19.60	1.07	20.46	1.05	22.29
Dwelling Density	1.67	6.07	0.44	6.55	0.30	4.74
Average Household Size (ln)	1.67	6.07	1.57	6.06	0.98	3.80
Transfer Stations	1.17	5.78	1.11	5.79	0.83	5.01
Median Income (ln)	-0.69	-4.31	-0.59	-3.88	-0.43	-2.96
Visible Minorities (ln)	0.05 *	1.67	0.04 **	1.32	0.01 **	0.43
Post-Secondary Education (ln)	0.29 *	1.77	0.24 **	1.56	0.10 **	0.68
Filtered Frequency 400 / 250 (ln)	0.95	15.78	0.87	14.29	1.02	16.64
Weighting Taps (ln)			0.25	5.47		
LAMBDA					0.62	11.14
Adjusted R2	0.79					
AIC	541.12		515.35		467.65	
LL	-260.55		-246.68		-228.32	

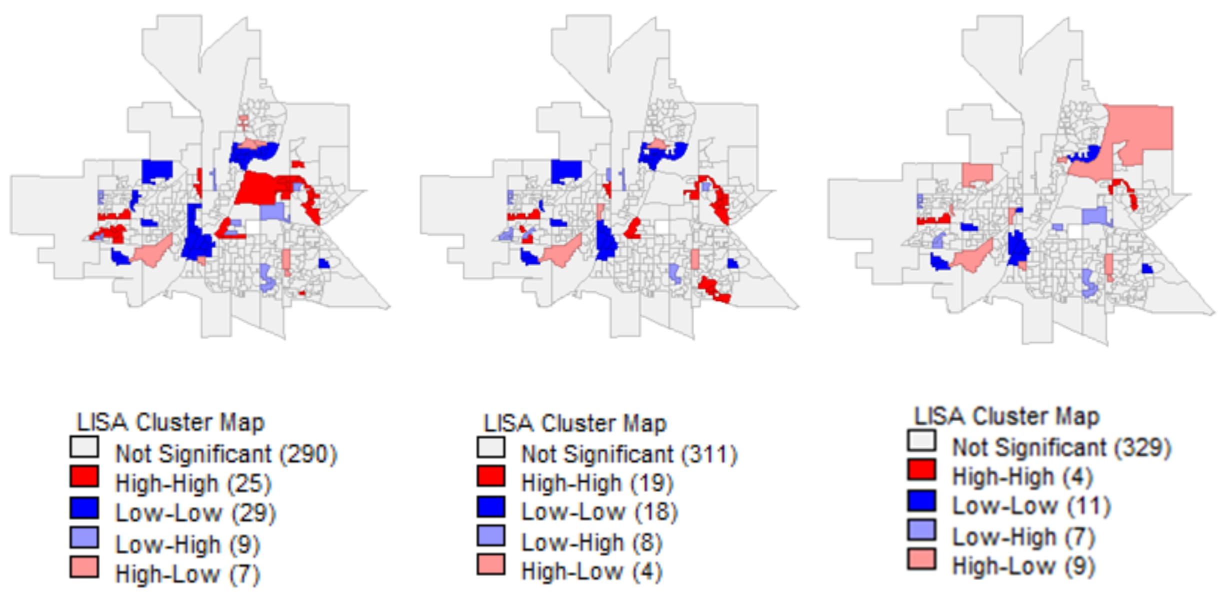
Number of records = 360

Most values are significant at 99% , * Significant at 90% ,** Not Significant



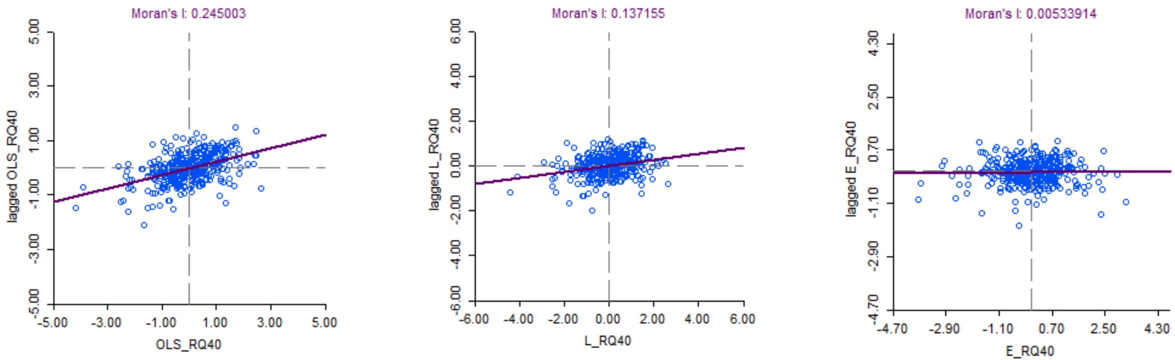
A. B. C.

Figure 13: Significance map for spatial autocorrelation for A. OLS, B. SLM, and C. SEM models with *Filtered Frequency 400 / 250*



A. B. C.

Figure 14: LISA cluster maps for the A. OLS, B. SLM, and C. SEM models with *Filtered Frequency 400 / 250*



A. B. C.

Figure 15: Local Moran's I scatter plot for the A. OLS, B. SLM, and C. SEM models with *Filtered Frequency 400/250*

5 EFFECT OF THE PROPOSED BRT SYSTEM

5.1 BRT Accessibility Measures

Accessibility measures for the proposed BRT transit system configuration were computed using the ***Filtered Frequency 400 / 250*** measure identified previously as producing the best SEM ridership model. Figure 16 illustrates the difference between the September 2018 accessibility measures and the BRT accessibility measure. The difference was computed by subtracting the September 2018 accessibility measures from the BRT accessibility measures. A positive result indicates increased accessibility, a negative result indicates decreased accessibility, and a result of zero indicates no change. Red indicates increased accessibility; blue indicates decreased accessibility. Figures 17 and 18 respectively are accessibility heatmaps of the September 2018, and the proposed BRT system.

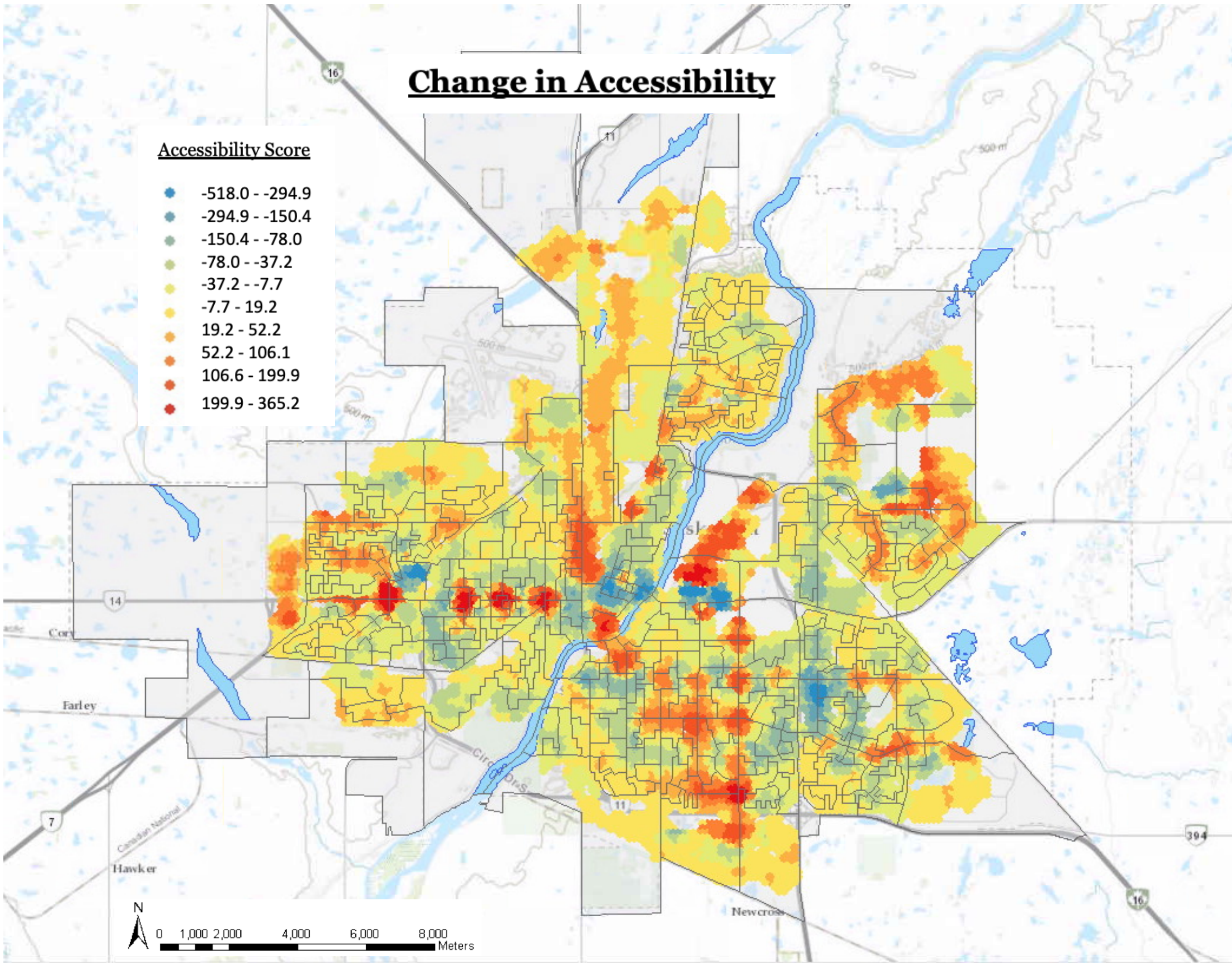


Figure 16: Change in the *Filtered Frequency 400 / 250* accessibility measure between the September 2018 transit system configuration and the proposed BRT system

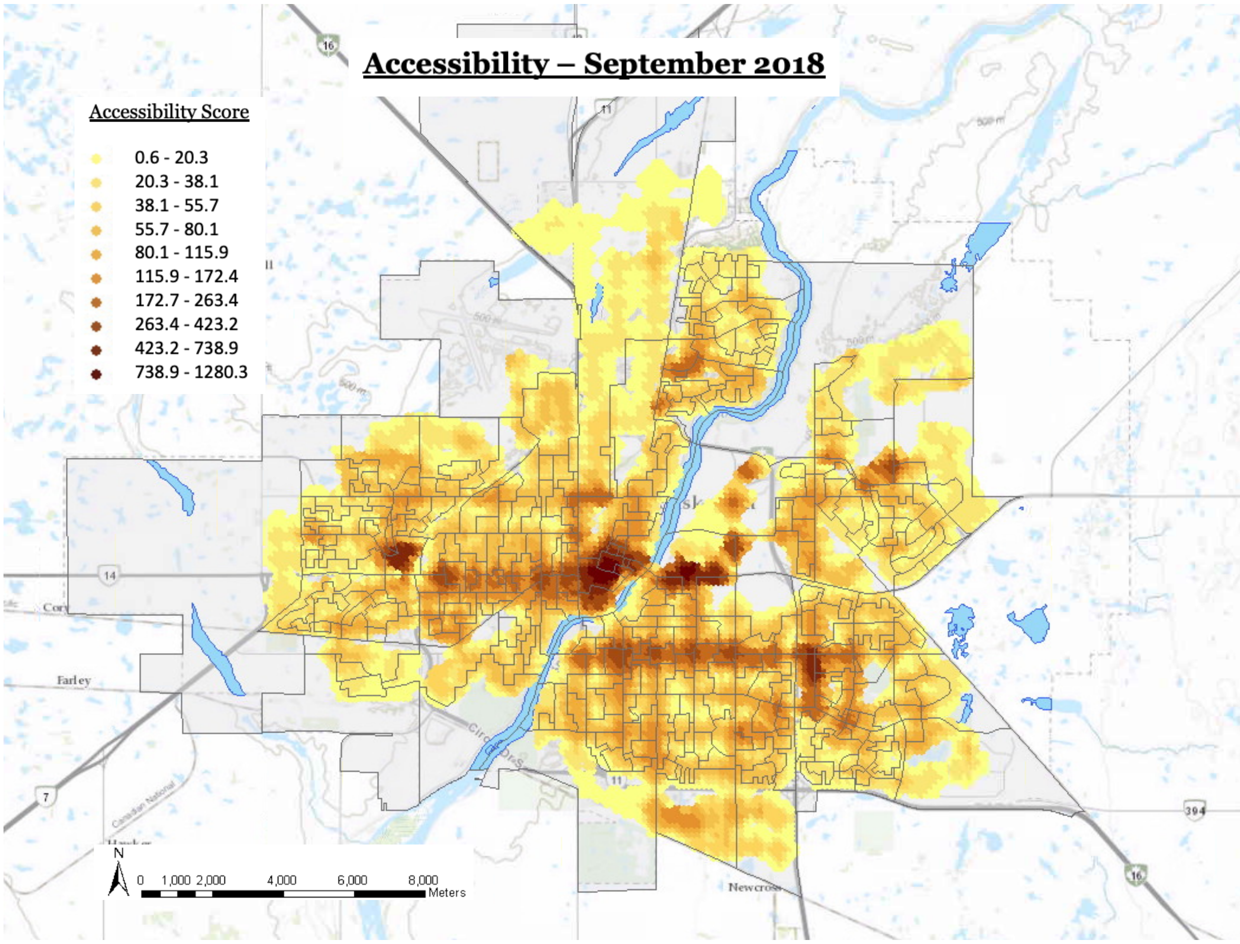


Figure 17: Heatmap of the *Filtered Frequency 400 / 250* accessibility measure for the September 2018 configuration

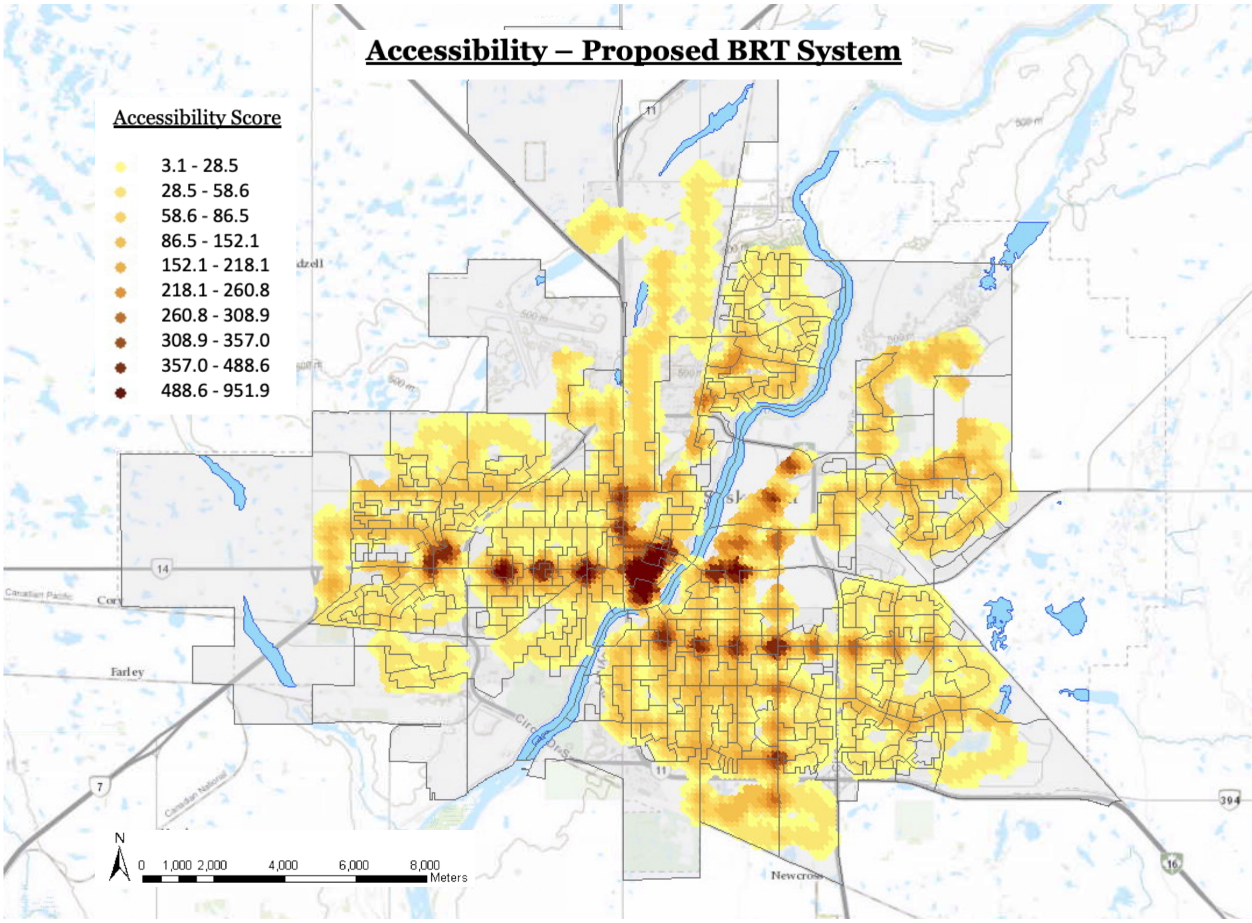


Figure 18: Heatmap of the *Filtered Frequency 400 / 250* accessibility measure for the proposed BRT system

The most obvious feature of this result is the increased accessibility around the proposed BRT stop locations. In fact, one can visualize the BRT routes on this heatmap. There is also increased accessibility along Taylor Street (southeast), the University Campus (excluding Place Riel), Idlywyld Drive (north central), and in some of the newer neighbourhoods on the city's periphery.

There is decreased accessibility at some of the current transfer and terminal locations, including Place Riel, the Downtown Bus Mall, Confederation Mall (west) and Circle Park Mall (east). There is also decreased accessibility along Central Avenue in Sutherland, and in some of the older central neighbourhoods.

Figure 16 illustrates relative changes in accessibility, not absolute values. Areas with reduced accessibility that are illustrated in blue in Figure 16, may still in fact have very good accessibility. These accessibility changes are consistent with the City of Saskatoon's stated strategy of increasing service is some areas at the expense of others in order to demonstrate the viability of BRT service.

6 DISCUSSION

This project examined a wide array of accessibility measures, with the aim of understanding how they perform in comparison to actual ridership. It also explored different types of statistical models to achieve the study goal, while minimizing estimation errors. Of those considered, the best performing measure was a filtered frequency algorithm with 400 m network constrained buffers and a Butterworth distance decay factor with a bandpass value of 250 m. In fact, this measure is identical to Langford's transit enhanced 2SFCA minus the step in which the service is divided by the population.

The resulting model did have difficulty predicting taps in DAs with the highest and lowest ridership. These DAs are those with very large or very small numbers of taps given their population. Examples of DAs with very large numbers of taps include those containing Place Riel and the Downtown Bus Mall. For example, Place Riel accounts for 17.3% of all September 2018 taps, yet the population of the DA that contains it is very small as it is comprised mostly of the U of S campus. The two most under predicted DAs were Place Riel and the Downtown Bus Mall, which are both typical morning destinations. The two most over predicted DAs were the Confederation Terminal and the Circle Park Terminal which are also both typical morning origin points. A table of the most over predicted DAs can be found in Appendix F.

6.1 Limitations

As of this writing, the BRT data acquired in kind from the City of Saskatoon is just a proposal and may be incomplete; many of the feeder routes have only a few stops assigned. The

City is expected to finalize the proposed BRT system in the near future. With the current BRT proposal, assumptions regarding stop locations can be made to compute accessibility measures. In addition, the most recent census data from Statistics Canada that breaks down population counts throughout the city based on dissemination areas was collected in 2016. Finally, the Saskatoon travel data from the SHTS was acquired in 2013.

6.2 Future Work

This study provides vital information for transit planners and decision makers to assess the impacts of transit system upgrades and land use changes to justify changes and costs to citizens. Important tools such as local transit accessibility measures allow planners to build more precise ridership models, which can be used to predict future ridership changes.

The results of this study only apply to system boardings using bus passes during the morning rush hour between 6 am and 9 am. This time was chosen on the assumption that riders would enter the system close to where they live, thus increasing the significance of DA population in the analysis. However, it is not known if the predicted changes in ridership extend to other times and methods of fare payment. Future work could include models for other times. For example, the afternoon rush hour, weekend taps, daily taps, annual taps, etc.

The SEM model used to predict ridership is composed of external socioeconomic variables such as median income, education levels, and built environment factors such as DA population and area. The factors used were chosen because they are typically used in these types of models.

Future work could include comparing model results from a more exhaustive combination of model external variables in order to predict transit ridership for the BRT system.

Saskatoon's DAs vary greatly by size, population, and other factors. In this work, two DAs with almost zero population were omitted from the analysis. These were the DA containing the Fieldhouse, Griffiths Stadium, and the University residences (DA ID 47110107), and the DA containing the new neighbourhood of Brighton (DA ID 47110689). However, there are several extreme outlier DAs including 47110106 that contains Place Riel (which alone accounted for 17.3% of all observed September 2018 taps), and the DAs that border the Downtown Bus Mall. Although these DAs were designated as transfer points in this work, future work could explore the handling of outlier DAs.

The network constrained buffers and distances used in this analysis were generated using a City of Saskatoon road network shapefile from 2012 which was the most recent available. Future work could include utilizing a more up-to-date road network. Additionally, paths and alleys could be incorporated into the network analysis.

The ***Filtered Frequency 400 / 250*** algorithm used to generate the accessibility measures utilized 400 m network constrained buffers and a Butterworth distance decay factor with a bandpass value of 250 m. These are the same parameters that Langford used with a transit-enhanced two step floating catchment area algorithm. Several other parameters were investigated, including 532 m network constrained buffers and an exponential distance decay function. Future work could include a more exhaustive examination of various values for these parameters.

The ***Filtered Frequency 400 / 250*** algorithm used to compute the accessibility measure is a measure of the opportunity to enter the system. That is, more frequent departures on more routes at closer stops all increase the opportunity. But there are many other factors that influence a potential rider's decision to utilize transit. For example, what is the total trip time? Are transfers required? Are all departures equal? Does a departure only serve one subsequent stop before turning around? Future work could include exploration of various algorithm enhancements to more accurately model ridership.

Finally, it would be very interesting to test the accuracy of the predicted increase in ridership over time. The techniques described in this work could be used to model ridership before and after the actual BRT rollout. Then, the predictions could be compared to actual observed changes in ridership.

7 CONCLUSION

This study examined a wide array of accessibility measures with the aim to understand how the models in which they were included performed in comparison to actual ridership data. Local transit accessibility measures were calculated using GTFS data for the 6 am to 9 am period of weekdays in September 2018. A detailed ridership data set was also obtained from Saskatoon Transit's farebox system to estimate ridership during the same time period. Three types of models were assessed to achieve the study goal. These models included ordinary least square models (OLS), spatial lag models (SLM), and spatial error models (SEM).

The models' results suggest that using accessibility measures with a distance decay function provides a better prediction of transit service ridership, while controlling for several household socioeconomic factors and built environment characteristics. The measure that best explained the variation in ridership across the three types of models used in this study was the ***Filtered Frequency 400 / 250*** accessibility measure that is based on 400 m network buffer and a Butterworth filter distance decay function with a bandpass value of 250 m. A similar filtered frequency accessibility measure using a distance decay function derived from the 2013 Saskatoon Household Travel Survey (SHTS) also performed well. However, this performance was hindered by the fact that the survey was conducted more than 5 years ago, and students were overrepresented in the survey. Therefore, by using a similar methodological approach as other transit agencies and researchers we can explore the performance of different local transit accessibility measures at different locations using different distance decay functions computed from more recent local household surveys.

The worst performing measure included a stop count measure that used a 400 m network buffer. In addition, Transit Score, which is commonly used in the literature, performed poorly in comparison to the other accessibility measures. However, one of the advantages of using Transit Score that was not captured in this study is that it incorporates a weighting for different modes of transit. For example, Transit Score weights heavy rail and light rail services higher than bus services. Using Saskatoon Transit as a case study, which operates only buses, precluded the examination of the performance of this measure in a multimodal transit system environment. Nevertheless, this study provides a methodological approach to incorporate factors such as local transit accessibility measures to allow planners to build more precise ridership models, which can be used to predict future ridership changes.

The relative effect of changes in local transit accessibility resulting from the planned BRT system configuration was calculated. Despite the number of active stops declining from 1,443 to 994, this study predicts there will be increased accessibility around the proposed BRT stop locations. It also predicts increased accessibility along Taylor Street (southeast), Idlywyld Drive (north central), and in some of the newer neighbourhoods on the city's periphery. It also predicts decreased accessibility at some of the current transfer and terminal locations. It further predicts decreased accessibility along Central Avenue in Sutherland, and in some of the older central neighbourhoods. These results are in keeping with the City of Saskatoon's stated goals and expectations.

In addition to recommending an accessibility measure to best predict transit ridership in Saskatoon, this study provides City officials with a methodology that can be consistently used over time to predict and compare changes in accessibility and ridership as the transit system is reconfigured. The model suggested in this study can be easily adapted to include more recent transit ridership data like that data included in the household travel survey once it becomes available. In addition, the model can also be adapted to calculate accessibility during different times of the day or months of the year. This will allow planners and decision makers to re-evaluate routes to increase ridership over time. Furthermore, the city will be able to compare their plan for growth to the accessibility heatmaps to evaluate how many people will have local access to the planned transit corridors and transit hubs. The heatmaps generated in this study can be used as visual aids to explain the changes in transit accessibility to stakeholders.

Finally, this research allows the City of Saskatoon to investigate how ridership will change as the transit system is reconfigured. Using the best fit model, they can investigate external variables in areas with higher ridership and compare them to areas with lower ridership. It is unlikely the City will be able to plan and implement a system that will serve and be attractive to all citizens equally. However, with this model, City officials can better understand the factors that influence change in accessibility and will, therefore, be able to explain and justify the system enhancements to citizens.

References

Binkovitz, L. 2016. "How Houston (?!) became a model for city planning." *Houston Cronicle*. Accessed Feb 2019. https://www.houstonchronicle.com/local/gray-matters/article/How-Houston-became-a-model-for-city-planning-6893430.php.

Boisjoly, G., and Ahmed E. 2017. "Measuring performance: accessibility metrics in metropolitan regions around the world.".

Boisjoly, G., and A. El-Geneidy. 2017. "The insider: A planners' perspective on accessibility." *Journal of Transport Geography* 64:33-43. doi: https://doi.org/10.1016/j.jtrangeo.2017.08.006.

Boisjoly, G., E. Grisé, M. Maguire, M. Veillette, R. Deboosere, E. Berrebi, and A. El-Geneidy. 2018. "Invest in the ride: A longitudinal analysis of the determinants of public transport ridership in 25 North America cities." The 97th Annual Meeting of the Transportation Research Board, Washington D.C., USA.

Chapman, B, H. Iseki, B. Taylor, and M. Miller. 2006. The effects of out-of-vehicle time on travel behavior: Implications for transit transfers. California: California Department of Transportation.

City of Saskatoon, and Robert Dudiak. 2018. "Transit plan.". City of Saskatoon, accessed July 24th https://www.saskatoon.ca/engage/transit-plan.

Dai, Dajun. 2010. "Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit." *Health & Place* 16 (5):1038-1052. doi: https://doi.org/10.1016/j.healthplace.2010.06.012.

Diab, E., M. Badami, and A. El-Geneidy. 2015. "Bus transit service reliability and improvement strategies: Integrating the perspectives of passengers and transit agencies in North America." *Transport Reviews* 35 (3):292-328.

Ding, Jishiyu, Yi Zhang, and Li Li. 2018. Accessibility measure of bus transit networks. *IET Intelligent Transport Systems* 12 (7): 682-688.

El-Geneidy, A., Buliung, R., Diab, E., Lierop, D., Langlois, M., and Legrain, A. 2015. "Non-stop equity: assessing daily intersections between transit accessibility and social disparity across the Greater Toronto and Hamilton Area (GTHA)." *Environment and Planning B: Planning and Design* 43 (3):540-560. doi: 10.1177/0265813515617659.

Fayyaz S., Kiavash, S., Xiaoyue C., and Guohui Z. 2017. "An efficient general transit feed specification (GTFS) enabled algorithm for dynamic transit accessibility analysis." *PLOS ONE* 12 (10):e0185333. doi: 10.1371/journal.pone.0185333.

Gent, C., Symonds, G. 2005. Advances in public transport accessibility assessments for development control - a proposed methodology. In *Transportation Research Labratory*, edited by Transport Research Board.

Geurs, T., and Wee, B. 2004. "Accessibility evaluation of land-use and transport strategies: review and research directions." *Journal of Transport Geography* 12 (2):127-140. doi: https://doi.org/10.1016/j.jtrangeo.2003.10.005.

Goodman, J., Laube, M. , and Schwenk, J. . 2006. "Curitiba bus system Is model for rapid transit." *Race, Poverty, and Environment*.

GTFS. 2019. Saskatoon Transit GTFS.

Guerra, E., and R. Cervero. 2011a. "Cost of a Ride: The effects of densities on fixed-guideway transit ridership and costs." *Journal of the American Planning Association* 77 (3):267-290. doi: 10.1080/01944363.2011.589767.

Guerra, E., and R. Cervero. 2011b. "Cost of a Ride:The effects of densities on fixed-guideway transit ridership and costs." *Journal of the American Planning Association* 77 (3):267-290. doi: 10.1080/01944363.2011.589767.

Handy, S. L., and D. A. Niemeier. 1997. "Measuring accessibility: an exploration of issues and alternatives." *Environment and Planning A: Economy and Space* 29 (7):1175-1194. doi: 10.1068/a291175.

Hansen, Walter G. 1959. "How accessibility shapes land use." *Journal of the American Institute of Planners* 25 (2):73-76. doi: 10.1080/01944365908978307.

Hidalgo, D. 2015. "After 15 years of moving people, here's what Bogotá's BRT should do next." In *TheCityFix*, edited by Ross Centre.

Ibeas, Á., Cordera, R., Dell'Olio, L., Coppola, P., and Dominguez, A. 2012. "Modelling transport and real-estate values interactions in urban systems." *Journal of Transport Geography* 24:370-382. doi: https://doi.org/10.1016/j.jtrangeo.2012.04.012.

ITDP. 2015. "New 'Gold-Standard' BRT in Yichang, China leads the way for mid-sized chinese cities." Institute for Transportation and Development Policy, accessed feb. https://www.itdp.org/2015/07/15/new-gold-standard-brt-in-yichang-china-leads-the-way-for-mid-sized-chinese-cities/.

Jaques, C., El-Geneidy, A. 2014. "Does travel behavior matter in defining urban form? A quantitative analysis characterizing distinct areas within a region." *Journal of Transport and Land Use* 7 (1):1-14. doi: 10.5198/jtlu.v7i1.377.

Hyun, K., and Song, Y. 2018. "An integrated measure of accessibility and reliability of mass transit systems." *Transportation* 45 (4):1075-1100. doi: 10.1007/s11116-018-9866-7.

Kohn, H. 2000. "Factors affecting urban transit ridership." Statistics Canada, Transportation Division.

Langford, M., Fry, R., and Higgs, G. 2012. "Measuring transit system accessibility using a modified two-step floating catchment technique." *International Journal of Geographical Information Science* 26 (2):193-214. doi: 10.1080/13658816.2011.574140.

Llamas, J. 2015. "Houston's groundbreaking new transit system." *Strong Towns*. https://www.strongtowns.org/journal/2016/1/25/texas-transportation-funding-its-not-all-bad-news.

Wei, L., and Qi, W. 2009. "An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians." *Health & Place* 15 (4):1100-1107. doi: https://doi.org/10.1016/j.healthplace.2009.06.002.

Wei,L., and Wang, F. 2003. "Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region." *Environment and Planning B: Planning and Design* 30 (6):865-884. doi: 10.1068/b29120.

Mathew, S. 2006. GeoDA and spatial regression modelling. Pennsylvania University Population Research Centre.

McGrail, Matthew R., and John S. Humphreys. 2009. "Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method." *Applied Geography* 29 (4):533-541. doi: https://doi.org/10.1016/j.apgeog.2008.12.003.

Miller, E., Shalaby, A., Diab, E., and Kasraian, D. 2018. Canadian transit ridership trends study. Toronto, Ontario: Candian Urban Transit Association (CUTA).

Moniruzzaman, M., and Páez, A. 2012. "Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters." *Journal of Transport Geography* 24:198-205. doi: https://doi.org/10.1016/j.jtrangeo.2012.02.006.

Morris, M., Dumble, P., and Wigan, M. 1979. "Accessibility indicators for transport planning." *Transportation Research Part A: General* 13 (2):91-109. doi: https://doi.org/10.1016/0191-2607(79)90012-8.

NATCO 2018. "Case study: metro bus network redesign, Houston.", accessed July 23rd. https://nacto.org/case-study/metro-bus-network-redesign-houston/.

Pasha, M., Rifaat, S., Tay, R., and De Barros, A. 2016. "Effects of street pattern, traffic, road infrastructure, socioeconomic and demographic characteristics on public transit ridership." *KSCE Journal of Civil Engineering* 20 (3):1017-1022.

Policy, Institute for Transportation and Development. 2016. "The BRT Standard." accessed Aug 26th 2019. https://www.itdp.org/library/standards-and-guides/the-bus-rapid-transit-standard/.

Polzin, S., Ram P., and Sachin, N. 2002. "Development of time-of-day–based transit accessibility analysis tool." *Transportation Research Record* 1799 (1):35-41. doi: 10.3141/1799-05.

Ryus, P., Ausman, J., Teaf, D., Cooper, M., and Knoblauch, M. 2000. "Development of Florida's transit level-of-service indicator." *Transportation Research Record* 1731 (1):123-129. doi: 10.3141/1731-15.

Statistics Canada. 2016a. "2016 Census - Boundary files." accessed July 20. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm.

Statistics Canada. 2016b. 2016 Census of Canada. Ottawa, Canada.

Taylor, B., and Fink, C. 2003. Los Angeles, California.

Taylor, B., Miller, D., Iseki, H., and Fink, C. 2009. "Nature and/or nurture? analyzing the determinants of transit ridership across US urbanized areas." *Transportation Research Part A: Policy and Practice* 43 (1):60-77.

Thill, J., and Marim, K. 2005. "Trip making, induced travel demand, and accessibility." *Journal of Geographical Systems* 7 (2):229-248. doi: 10.1007/s10109-005-0158-3.

Thompson, G., Brown, J., and Bhattacharya, T. 2012. "What really matters for increasing transit ridership: understanding the determinants of transit ridership demand in Broward County, Florida." *Urban Studies* 49 (15):3327-3345. doi: 10.1177/0042098012443864.

Van Lierop, D., Badami, M., and El-Geneidy, A. 2018. "What influences satisfaction and loyalty in public transit? A critical view of the literature." *Transport Reviews* 38 (1):52-72.

Walk Score. 2019. "Walk Score ", accessed July 20. https://www.walkscore.com/.

Wang, K., and Woo, M. 2017. "The relationship between transit rich neighborhoods and transit ridership: evidence from the decentralization of poverty." *Applied Geography* 86:183-196. doi: https://doi.org/10.1016/j.apgeog.2017.07.004.

Zhao, J., Webb, V., and Shah, P. 2014. Customer loyalty differences between captive and choice riders.

Appendix A

The Walk Score values were obtained over several days using the following API (Application Programming Interface):

```
http://api.walkscore.com/score?format=json&lat=<YOUR-LAT>
&lon=<YOUR_LON>&transit=1&bike=1&wsapikey=<YOUR-WSAPIKEY>
```

where <YOUR-LAT> and <YOUR-LON> are the latitude and longitude of the grid cell centroid, and <YOUR-WSAPIKEY> is the API key. The results are returned in JSON format as shown in the following example response:


```
{
  "status": 1,
  "walkscore": 98,
  "description": "Walker's Paradise",
  "updated": "2019-02-17 04:40:31.218250",
  "logo_url": "https://cdn.walk.sc/images/api-logo.png",
  "snapped_lat": -52.456356,
  "snapped_lon": 106.593884,
  "transit" : None,
  "bike" : None,
}
```


Note that for Saskatoon, the Walk Store API does not return any transit or bike scores, although the Transit Score can be obtained using the separate Transit Score API.

Appendix B

The Transit Score values were obtained over several days using the following API (Application Programming Interface):

```
http://transit.walkscore.com/transit/score?format=json&
city=Saskatoon&country=CA&lat=<YOUR-LAT>&
lon=<YOUR-LON>&wsapikey=<YOUR-WSAPIKEY>
```

where `<YOUR-LAT>` and `<YOUR-LON>` are the latitude and longitude of the grid cell centroid, and `<YOUR-WSAPIKEY>` is the API key. The API requires the city and country arguments even though the latitude and longitude are specified. The results are returned in JSON format as shown in the following example response:


```
{
  "description": "Some Transit",
  "help_link": "https://www.redfin.com/how-walk-score-works",
  "logo_url": "https://cdn.walk.sc/images/transit-score-logo.png",
  "summary": "6 nearby routes: 6 bus, 0 rail, 0 other",
  "transit_score": 35
}
```

Appendix C

Boardings by Fare Type

Fare Type	Boardings	Percent of Total
Pass (multi ride card)	499192	55.36%
Transfer - Pass (multi ride card)	202392	22.44%
Rides (multi ride card)	59878	6.64%
Exact Fare	43419	4.81%
Transfer - Rides (multi ride card)	27802	3.08%
Transfer - Ticket Accepted	25875	2.86%
Override	23285	2.58%
Counter	8675	0.96%
Rides (disposable)	5239	0.58%
Transfer - Rides (disposable)	2436	0.27%
Cash Exceeded	2372	0.26%
Pass (disposable)	313	0.03%
Override CNIB	307	0.03%
Transfer - Pass (disposable)	242	0.02%
ePurse (multi ride card)	127	0.01%
Transfer - ePurse (multi ride card)	71	< 0.01%
Total Boardings	901625	100%

Appendix D

Count of “good” stops and “bad” stops per route

Route ID	Route Name	Taps with Stop ID	Taps missing Stop ID	% Missing
6	Broadway / Market Mall	39412	6474	14.11 %
8	8th Street / City Centre	35047	3442	8.94 %
17	Stonebridge / University	32487	4019	11.01 %
65	University / Kensington	25481	1098	4.13 %
60	University / Confederation	25023	1798	6.7 %
13	Lawson Heights / Broadway	21564	4924	18.59 %
26	Forest Grove / University	19225	3025	13.6 %
82	Main Street / Centre Mall	18863	2973	13.62 %
5	City Centre/ McCormack	18253	3433	15.83 %
50	Lakeview / University	18198	2973	14.04 %
55	Lakeridge/ University	18035	2323	11.41 %
7	Dundonald / City Centre	17653	1396	7.33 %
18	College Park / University	17459	2012	10.33 %
2	Meadowgreen / City Centre	16262	5550	25.44 %
81	Cumberland / Centre Mall	15647	2562	14.07 %
27	Silverspring / University	15519	928	5.64 %
44	Willowgrove / City Centre	12838	3502	21.43 %
19	City Centre / Centre Mall	12503	1797	12.57 %
61	University/ Blairmore	12213	2222	15.39 %
9	Riversdale / City Centre	11902	1180	9.02 %
45	Arbor Creek / City Centre	11494	1801	13.55 %
35	Silverwood / City Centre	10953	1857	14.5 %
63	University / Hampton Village	10358	1581	13.24 %
4	City Centre	10261	2796	21.41 %
12	River Heights / City Centre	9934	1124	10.16 %
10	Meadowgreen / City Centre	9872	1638	14.23 %
43	Evergreen / City Centre	9419	962	9.27 %
11	Airport / City Centre	8759	289	3.19 %
30	Lawson Heights / City Centre	8202	1991	19.53 %
1	City Centre / Exhibition	7658	2574	25.16 %

83	Centre Mall / Stonebridge	5688	3139	35.56 %
62	University / Montgomery	5251	551	9.5 %
14	North Industrial / City Centre	5057	1177	18.88 %
86	Rosewood / Centre Mall	4734	1924	28.9 %
3	City Centre / Hudson Bay Park	4533	4267	48.49 %
84	Briarwood / Centre Mall	2398	904	27.38 %
22	Confederation / City Centre	1820	1443	44.22 %
0	UNKNOWN	1803	4870	72.98 %
506	UNKNOWN	1670	208	11.08 %
21	University	1582	1204	43.22 %
518	UNKNOWN	1164	109	8.56 %
336	UNKNOWN	961	33	3.32 %
315	UNKNOWN	768	89	10.39 %
555	UNKNOWN	753	0	0 %
338	UNKNOWN	554	66	10.65 %
305	UNKNOWN	554	0	0 %
352	UNKNOWN	548	109	16.59 %
316	UNKNOWN	526	74	12.33 %
311	UNKNOWN	467	100	17.64 %
581	UNKNOWN	434	48	9.96 %
517	UNKNOWN	393	0	0 %
325	UNKNOWN	354	92	20.63 %
502	UNKNOWN	258	23	8.19 %
309	UNKNOWN	244	1002	80.42 %
102	UNKNOWN	244	68	21.79 %
332	UNKNOWN	240	0	0 %
358	UNKNOWN	197	59	23.05 %
15	UNKNOWN	182	131	41.85 %
333	UNKNOWN	156	125	44.48 %
20	South Industrial / City Centre	141	105	42.68 %
582	UNKNOWN	124	0	0 %
307	UNKNOWN	105	4	3.67 %
101	UNKNOWN	61	96	61.15 %
25	Sasktel Centre	50	1381	96.51 %
550	UNKNOWN	27	1	3.57 %
808	Field House / City Centre	21	412	95.15 %
9999	UNKNOWN	19	78	80.41 %

514	Sasktel Centre / North Ind	5	0	0 %
999	UNKNOWN	3	87	96.67 %
34	UNKNOWN	0	1	100 %

Appendix E

Top 20 stops by tap count for September 2018

Stop ID	Stop Name	Tap Count
5904	Place Riel / Terminal E&S	69301
5903	Place Riel / Terminal W&N	22830
5899	Downtown Terminal North	7372
5912	Confederation Terminal	7171
5900	Downtown Terminal South	6483
5910	Centre Mall Terminal I/B	6255
5901	Downtown Terminal West	6196
4174	Superstore	5511
3077	1st Avenue / 21st Street	5386
5906	Market Mall Terminal I/B	5370
5897	3rd Avenue / 23rd Street	5318
5413	Downtown Terminal West 1	4872
5908	Lawson Terminal I/B	4817
4315	Idylwyld / 33rd Street	4409
5909	Centre Mall Terminal O/B	4154
5902	Downtown Terminal East	4104
5465	Bowlt / Shaw Center	4076
3993	Preston / Preston Crossing	4022
5556	8th Street / JYSK Store	3642
3173	25th Street / 5th Avenue	3436

Appendix F

Over predicted dissemination areas

DA ID	DA Description	Taps Observed Sept. 2018	Taps Predicted Sept. 2018	Difference
47110365	Confederation Park Terminal	308	414	106
47110451	Circle Park Terminal	151	206	55
47110311	U of S fields (east of Circle Dr)	50	85	35
47110570	S 22nd St. - Ave. H to Ave N	31	58	27
47110069	S 22nd St. - Ave. E to Ave. H	37	60	23
47110067	S 22nd St. - Idlwyld to Ave. E	45	67	22
47110412	N 22nd St. - Idlwyld to Ave. E	18	38	20
47110397	Airport	13	31	18
47110540	Gordie Howe Sports / Landfill	18	35	17
47110669	S.E. Circle Dr	10	26	16