2013-10-08

Modeling non-motorized travel demand at intersections based on traffic counts and GIS data in Calgary, Canada

Tabeshian, Maryam

UNIVERSITY OF CALGARY

Modeling non-motorized travel demand at intersections based on traffic counts and GIS data in Calgary, Canada

by

Maryam Tabeshian

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF CIVIL ENGINEERING

CALGARY, ALBERTA

SEPTEMBER, 2013

Abstract

In September 2009 the City of Calgary Council approved Plan It Calgary, which proposes policies that focus on the development of resilient neighborhoods through the intensification and diversification of urban activities around transit stations and routes. More intensive development and mixed land use encourage non-motorized trips and reinforce comfortable, safe and walkable streets. The development of high-density, mixed-use and transit- and pedestrian-oriented communities has the potential to generate trips with shorter destinations, which are expected to result in a higher share of active travel modes, such as biking and walking. Thus, there is a growing need to estimate the impact of land-use development scenarios and transportation policies on bicycle and pedestrian demand to predict future non-motorized trip volumes and adequately design the related infrastructure.

This study calibrates multiple linear and Poisson regression models to estimate non-motorized travel demand based on GIS, transportation data and road characteristics. The empirical models that have been developed in this research can be used to assess the impacts of urban design and built environments, such as developing high-density and mix-land-use areas, and building complete streets in the middle ring communities of the City of Calgary in influencing the demand for active travel modes. The developed models show the benefits of improved pedestrian infrastructure, such as improved network connectivity and increases in the length of pedestrian pathways, as well as the integration of transit and walking modes and transit and bicycle modes in encouraging more non-motorized travel demand. The method employed herein is a straightforward statistical analysis method, and the needed data are relatively easy to access.

Dedicated to:

The source of my happiness, my beloved husband Hadi, for his great support, encouragement, and undying love.

Acknowledgment

First and foremost, I would like to express my sincerest gratitude to my supervisor Dr. Lina Kattan who has supported me during my Master's study and research with her patience, motivation, and knowledge. One simply could not wish for a better or friendlier supervisor.

Many thanks go to the members of the dissertation committee, Dr. Andrew Hunter, Dr. Alex de Barros, and Dr. Dada for their time and useful comments on this thesis.

I am grateful for the financial support I received through the Geomatics for Informed Decisions Network (GEOIDE), the Neptis Foundation and Urban Alliance Professorship Research Funds.

My final, and most heartfelt, acknowledgment must go to my family: my parents, my brother and sister for supporting me with their unconditional love and for all their constant encouragement during my life.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Acronyms

Acronyms	**Definition**
CBD	Central Business District
CTP	Calgary Transportation Plan
GEOIDE	Geomatics for Informed Decisions
GIS	Geographic Information System
GOF	Goodness-of-Fit
LL	Log of the Likelihood
MDP	Municipal Development Plan
PYP	Plan Your Place
RMSD	Root Mean Square Deviation
RMSE	Root Mean Square Error
SSE	Sum of Square Errors
SSR	Regression Sum of Squares
SST	Total Sum of Squares
TAZ	Transportation Analysis Zone
TDM	Travel Demand Management
TOD	Transit Oriented Development
TSM	Transportation System Management

CHAPTER 1. INTRODUCTION

1.1. Background

In recent years, the overreliance on motorized transport modes such as private cars has brought about various side effects such as air pollution, traffic congestion, auto-oriented lifestyles, and safety problems especially for non-motorized modes. Different strategies have been implemented to overcome these problems including enhanced pedestrian and bicycle friendly designs, land use planning, Travel Demand Management (TDM), Transportation System Management (TSM), and improved transit systems (Kim, 2005).

Transportation planners and policy makers have become increasingly interested in improving non-motorized infrastructure and encouraging cycling and walking as active and eco-friendly means of transportation. The intent is the provision of safer roads and promotion of healthier lifestyles through the shaping of urban forms and innovative transportation solutions that bring about shorter destinations and walkable, pedestrian friendly designs.

In September 2009, the City of Calgary Council approved Plan It Calgary, which is an integrated Municipal Development Plan (MDP) and Calgary Transportation Plan (CTP). Plan It Calgary proposes policies that focus on the development of resilient neighborhoods through the intensification and diversification of urban activities around transit stations and routes. More intensive development and mixed land use encourage non-motorized trips and reinforce comfortable, safe and walkable streets (Plan It Calgary, 2009).

In a recent effort to promote sustainable communities, the PlanYourPlace (PYP) project, sponsored by GEOIDE (Geomatics for Informed Decisions) and supported by the Neptis Foundation, adopts a user-centered design approach that brings together planning tools founded on practical, academically sound principles that are designed to educate and stimulate interest in the development of future communities within the middle ring neighborhoods in the City of Calgary. The city's "middle ring" is comprised of neighbourhoods that were developed between the 1950s and 1970s. The dominant pattern of each of these neighbourhoods consists of a warped grid and crescent blocks organised around a central school and recreation field(s). Commercial development within the neighbourhood unit typically takes the form of auto-oriented strip malls with a large grocery store as an anchor store and large surface parking lots. In general, the PYP project's goal is to develop a process and guidelines that assist the City in transitioning these "middle ring" neighbourhoods into a sustainable future (Plan Your Place, 2011).

In particular, the intent of the PlanYourPlace project is to help change existing lifestyle choices and allow communities to switch to options that promote sustainable living where compact, mixed-use, pedestrian-friendly street network, and design are integrated to support walking, cycling, and high-quality transit. This form of development is commonly referred to as TOD (Transit Oriented Development).

In addition, the development of mixed-use communities and pedestrian friendly neighborhoods is expected to result in a higher share of active travel modes, such as biking and walking. Thus, there is a growing need to estimate the impact of development scenarios and transportation policies to properly estimate future non-motorized trip demand and to determine the important factors that can affect the frequency of walking and biking in Canadian cities and adequately design the related infrastructure. This is especially important for a highly auto-oriented city like Calgary. The ability to predict the future non-motorized volumes also helps practitioners to conduct non-motorized safety studies and public health studies.

1.2. Research Objective

The objectives of this study are twofold: 1) to examine the impact of important factors affecting the frequency of active modes; and 2) to estimate bicycle and pedestrian volumes at major intersections in the City of Calgary as affected by these elements. This research uses the sketch plan method and calibrates different regression models to estimate bicycle and pedestrian travel demand based on geographic information system (GIS), traffic data, and road characteristics. The outcomes from this study also identified the importance of land use characteristics, urban design, and transportation elements affecting non-motorized travel demand in the City of Calgary.

The required data — bicycle and pedestrian counts, GIS, transportation, and socioeconomic data — for developing the regression models were obtained from different sources including the City of Calgary, Spatial and Numeric Data Services at the University of Calgary, Calgary transit, census data, and other sources. Using these data and with the aid of ArcGIS software, 108 explanatory variables were defined and examined for calibrating the regression models.

Several regression models were developed to examine the relationships between non-motorized trip volumes (dependent variables) and built environments, land-use characteristics, socio-economic variables, and transportation services (independent variables). In this research multiple linear and Poisson regressions were used for calibrating non-motorized modes prediction models. The regression models developed in this research also identified several contributing factors that can affect frequency of walking and biking in Calgary. These models indicated that improved pedestrian and bicycle infrastructure, such as improved pedestrian network connectivity and pathway length, and improved transportation service integration, such as transit and bicycle integration, as well as safer routes for pedestrian and cyclists make a significant contribution to increasing non-motorized travel volumes.

Moreover, the validation process conducted for a sample of 18 intersections in Calgary showed that the bicycle and pedestrian prediction values estimated using linear and Poisson regression models have adequate predictability. Additionally, the predictability for both linear and Poisson regression models are almost at the same level.

The models developed in this research can be used to assess the qualitative impact of urban design, the built environment, non-motorized mode infrastructure and better integration of active travel modes with transit — all of which can influence the demand for active modes in the middle ring communities in the City of Calgary. Moreover, these models can be used to estimate present non-motorized travel demand at intersections where no count data are available. This method is a straightforward statistical analysis for practitioners and urban planners to estimate future pedestrian and cyclist volumes, assess their future travel needs and plan for adequate related facilities, such as walkways and bicycle infrastructure.

1.3. Research Contributions and Findings

According to the literature, non-motorized prediction models for most studies focus on American cities. Most of these models were developed to examine qualitatively the factors contributing to encouraging non-motorized trips. Therefore, these models do not have an acceptable goodness-of-fit and thus are not capable of predicting future pedestrian volumes with reasonably high accuracy. Despite these efforts, there is still a lack of an applicable model for the estimation of pedestrian and bicycle volumes in Canadian cities, especially for a highly auto-oriented city, such as Calgary.

This study adopts a sketch plan approach to model pedestrian and cyclist volumes in Calgary. The developed regression models in this thesis made the following contributions to:

i. *<u>Introducing the first bicycle regression model in Canada, and the first pedestrian regression model in Calgary:</u>*

The developed cyclist regression models in this research are the first bicycle prediction models calibrated for a Canadian city. The developed pedestrian regression models are the first pedestrian prediction models calibrated for the City of Calgary. It is to be noted that these prediction models were calibrated for P.M. peak and thus reflect both commute and non-commute trips.

ii. *<u>Having the ability to predict future bicycle/pedestrian travel demand:</u>*

The developed models have a relatively high goodness of fit, and all explanatory variables in these models are statistically significant. Therefore, these models can be used for the purpose of predicting bicycle and pedestrian travel demand. The developed model can also be used to examine the contributing land use and transportation elements that affect the frequency of walking and biking.

iii. *<u>Considering the effect of road narrowing and traffic calming technique on bicycle travel demand:</u>*

The variable "Lane", which is defined as the total number of street lanes reaching the intersection, is introduced in this research for the first time. The negative sign of the estimated parameter in this variable in the bicycle prediction model indicates that the cyclists are more willing to bike in narrower streets, which is attributed to safety problems that cyclists have on wide streets. This variable also highlights the important role that road narrowing plays as an efficient traffic calming technique as it forces motorists to reduce their speeds, thereby resulting in improved safety for cyclists and attracting more cyclists to use the infrastructure.

iv. *Introducing the variable "Street-Length" to consider the impact of connectivity on pedestrian travel demand:*

The variable Street-Length is another variable introduced in this research for the first time. This variable shows the centerline kilometers of streets in a specific buffer zone around the intersection and can be interpreted as connectivity. Increasing the number of streets with pedestrian sidewalks in a certain buffer area results in a denser street network in that area. This improves connectivity and is shown to increase pedestrian volumes.

1.4. Thesis Organization

The remaining chapters of the thesis are organized as follows. Chapter 2 presents a review of the literature on previous non-motorized estimation models. Chapter 3 consists of two sections. In the first section the data used in this research is described, including study area, pedestrian and bicycle count data, land-use variables, transportation services, and socio-economic characteristics. In the second section the conceptual framework and the methodology for developing the regression models are described. In Chapter 4 the results of the development of the linear and Poisson regression models for estimating active mode travel demand are presented. The over-dispersion test conducted to fulfill the Poisson distribution property that restricts the mean and variance to be equal is also described in this chapter. The model validation is also discussed at the end of Chapter 4. The final conclusions, research contributions, and recommendations for future work are included in Chapter 5.

CHAPTER 2. LITERATURE REVIEW

Chapter 2 presents a review of the literature of different methods to estimate non-motorized trip volumes. This chapter focuses mainly on previous literature related to non-motorized prediction models which were developed using sketch plan method.

2.1. Methods of Estimating Non-Motorized Travel Demand

In recent years, planning for active travel modes has become one of the important parts of transportation engineering and urban planning. Practitioners and policy makers need to have accurate information on non-motorized trip volumes for conducting public health studies (Cervero and Duncan, 2003), bicycle and pedestrian safety studies (Qin and Ivan, 2001; Miranda-Moreno et al., 2010), and feasibility studies for active modes infrastructure improvements (Lewis and Kirk, 1997; Ercolano et al., 1997). Various methods have been proposed for counting pedestrians, both manually and automatically (Pulugurtha and Repake, 2008; Schneider et al., 2009). However, to overcome the need of collecting data different kinds of methods have been proposed in order to predict non-motorized travel demand based on available data.

The *Guidebook on Methods to Estimate Non-Motorized Travel* (FHWA, 1999) describes five different methods for estimating non-motorized trip volumes. These methods include: aggregate behaviour studies; discrete choice models; regional travel models; comparison studies; and sketch plan methods. A summary of these methods is presented in Table 2.1. In this table the column Accuracy indicates accuracy level for one method in comparison with the other mentioned methods and Sensitivity to Design Factors demonstrates the ability of the methods to assess the effects of design factors on demand.

Table 2.1. Categorization of Non-Motorized Trips Estimation Method

Demand Estimation Method	Level of Study Area	Accuracy	Sensitivity to Design Factors
Aggregate Behavior Studies	Area/Regional Level	Very Low	Very Low
Discrete Choice Models	Facility Level Area/Regional Level	High	High
Regional Travel Model	Facility Level Area/Regional Level	Moderate	Moderate
Comparison Studies	Facility Level	Low	Low
Sketch Plan Method	Facility Level	Moderate	Moderate

Source: FHWA (1999)

The next sections provide a description of the details of the five different types of non-motorized trips estimation methods with special focus on the sketch plan methods.

2.1.1. Aggregate behavior studies

According to the *Guidebook on Methods to Estimate Non-Motorized Travel* (FHWA, 1999) Aggregate behavior studies involve "the development of models to predict mode split or other travel behavior characteristics at an area level, such as for residents of census tracts or metropolitan areas". In this method the prediction is based on socio-economic characteristics and other variables related to that area such as non-motorized infrastructure conditions and availability of other modes. Ashley and Banister (1989) and Nelson and Allen (1997) applied this method to predict the percentage of cyclist commuters in United Kingdom and United States areas, respectively.

In recent years, this method has not been very popular and is thus rarely adopted for predicting non-motorized travel demand. One of the disadvantages of this method is the large spatial unit of analysis needed, which is usually as big as a census tract or transportation analysis zone (TAZ). Another disadvantage of this method is that it is mainly focused on commuter trips and does not consider other types of travel, such as recreational and shopping trips. Moreover, in

Aggregate Behavior Studies the accuracy and the sensitivity to design factors[1] are very low, which can be considered as another shortcoming of this method.

2.1.2. Discrete choice models

The *Guidebook on Methods to Estimate Non-Motorized Travel* (FHWA, 1999) defines discrete choice models as follows: "A discrete choice model predicts a decision made by an individual (choice of mode, choice of route, etc.) as a function of any number of variables, including factors that describe a facility improvement or policy change". This method is very accurate and very sensitive to design factors. The method has been used in some studies in order to develop models for predicting bicycle and pedestrian mode choice. Wilbur Smith Associates (1996) developed discrete choice models in order to predict the effect of non-motorized improvements on transit access mode. Recently, Hunt and Abraham (2007) developed a discrete choice model to estimate bicycle demand in the City of Calgary.

Despite the advantages of discrete choice models, they require substantial effort for collecting the survey data and calibrating the discrete choice models. This method is thus costly and time consuming. In addition, it is not easy for practitioners as it needs individuals who are expert in discrete choice modeling techniques.

2.1.3. Regional travel models

According to the *Guidebook on Methods to Estimate Non-Motorized Travel* "regional travel models, commonly referred to as "four-step travel demand models", use data on existing and future population, employment, and transportation network characteristics, in conjunction with

---

[1] Sensitivity to design factors shows the ability of method to assess the effects of design factors on demand.

data on existing travel patterns and models of human behavior, to predict future travel patterns" (FHWA, 1999). This method comprises four steps:

- Trip generation
- Trip distribution
- Mode choice
- Route Assignment

The traditional four-step model is widely used to predict automobile and transit demand. The spatial unit of analysis developed for this method, which is mostly census tract or Transportation Analysis Zone, is more appropriate for automobile rather than bicycle and pedestrian use (Porter et al., 1999). However, in some studies, modified versions of this model were developed that were capable of taking pedestrian and bicycle demand into account (Purvis, 2003).

This method is very difficult to use by practitioners in comparison with other existing methods and requires extensive data. In addition, this method is mainly focused on commuter trips and do not consider other types of trips made for the sole purpose of recreation (FHWA, 1999).

2.1.4. Comparison studies

Comparison studies are one of the simplest methods for estimating bicycle and pedestrian volumes. According to this method, "non-motorized travel demand on a facility is predicted by comparing it to usage and to surrounding population and land use characteristics of other similar facilities" (FHWA, 1999). However, a comparison study usually provides only a very rough estimation of travel demand and, sometimes, it is very difficult to find facilities that are truly comparable (FHWA, 1999). Lewis and Kirk (1997), for example, estimated bicycle volumes on a proposed rail trail bikeway in a Massachusetts metropolitan area based on the bicycle counts from a similar bikeway in the Boston area.

2.1.5. Sketch plan methods

The sketch plan method is another technique for estimating bicycle and pedestrian volumes. According to the *Guidebook on Methods to Estimate Non-Motorized Travel (*FHWA, 1999**),** sketch plan methods were originally described as “methods [that] generally use bicycle and pedestrian counts and regression analysis to predict non-motorized trips volume as a function of adjacent land uses and indicators of transportation trip generation” (FHWA, 1999). This method has been used in many existing bicycle and pedestrian volume prediction models.

The studies done by Pushkarev and Zupan (1971) and Behnam and Patal (1977) were among the first efforts using the sketch plan method to estimate pedestrian trip volumes. They proposed regression models for forecasting pedestrian volume in central business districts (CBDs) based on land-use characteristics. Pushkarev and Zupan (1971) used aerial photography to develop a linear regression model to relate pedestrian volumes per block to commercial spaces, distance to transit stops, and amount of sidewalk. Behnam and Patal (1977) also used linear regression to estimate pedestrian volume per hour in Milwaukee (Wisconsin) based on land-use characteristics.

During the past few years researchers tried to relate non-motorized travel demand into different variables such as: built environment characteristics (Dill, 2009; Guo et al., 2007; Haynes and Andrzejewski, 2010; Jones et al., 2010); infrastructure design characteristics

(Lindsey et al., 2007; Reynolds et al., 2007); socio-economics (Schneider et al., 2010; Miranda-Moreno and Fernandes, 2011); and weather and topography (Handy, 2005; Krizek, 2003; Cameron, 1976). One of the first attempts to study impact of neighborhood characteristics on travel demand was conducted by Levinson and Wynn (1963). This study demonstrated that high density neighborhoods have a lower frequency of private vehicle trips and higher frequency of the use of public transit and non-motorized trips in comparison with low density neighborhoods. In a recent study in Montreal, Canada, Miranda-Moreno and Fernandes (2011) used weather conditions and spatio-temporal patterns to calibrate a pedestrian prediction model. In this research linear and negative binomial regression were developed to predict aggregate 8-hour and disaggregate hourly pedestrian volumes. This study was one of the first attempts for developing pedestrian regression models in a Canadian city.

Estimating bicycle/pedestrian travel demand and determining important factors affecting non-motorized trip volumes have various applications in different research areas such as feasibility studies for active modes infrastructure improvements, non-motorized safety studies, and public health studies. The following examples briefly review the past research related to these applications:

- Ercolano et al. (1997) proposed a model to estimate pedestrian travel demand based on hourly vehicular volume, non-motorized mode share, and transit ridership in suburban areas. This model has been used to determine locations for pedestrian infrastructure improvements.

- Qin and Ivan (2001) developed a linear relationship between the natural log of weekly pedestrian counts, the built environment and road characteristics to estimate pedestrian volumes needed as input for pedestrian safety studies.

- Cervero and Duncan (2003) determined the effect of urban landscapes on walking and biking in San Francisco Bay Area in order to find important variables that have significant impacts on public health.

Some studies defined their independent variables, such as land-use characteristics, socio-economic variables and transportation services, using different scales of buffer zones around the studied intersections. Multiple-scale analyses showed that the best models were developed when independent variables were chosen from different scales of buffer zones (Schneider et al., 2010; Liu and Griswold, 2009; Miranda-Moreno and Fernandes, 2011).

Most of the studies on non-motorized prediction models have only focused on modeling pedestrian travel demand. Just a few of them tried to calibrate regression models to relate bicycle volumes into built environment characteristics. Haynes and Andrzejewski (2010) developed a linear regression model to estimate weekday afternoon peak hour bicycle travel demand in the City of Santa Monica. Their model showed that land use mix, PM bus frequency, and a bicycle network have a positive impact on bicycle volumes while having a young population under the age of 18 has a negative impact. Jones and Buckland (2010) used the sketch plan method to develop a linear regression model in order to estimate bicycle and pedestrian travel demand based on socio-demographic and physical factors in San Diego County, California. They also tried to use a natural logarithm of dependent variable to avoid predicting negative bicycle volumes. In another study, Griswold et al. (2011) tried to calibrate bicycle linear regression models during weekdays and weekends based on different independent variables such as: socio-economic, land-use, transportation system and intersection site characteristics in Alameda County, California. The models in this study showed that proximity to retail or a large university has a higher impact on bicycle volumes on weekdays than weekends. On the other hand the weekends models showed a lower impact for hilly terrain and a higher impact of bicycle facilities on bicycle trips in comparison with weekdays models.

Since non-motorized travel demand consists of non-negative integer values and is considered as count data, it can be best modeled by Poisson and negative binomial regressions (Washington et al., 2003; Davidson and MacKinnon, 2003). Kim (2011) estimated pedestrian volume using Poisson regression. Miranda-Moreno and Fernandes (2011) and Cao et al. (2006) developed negative binomial regression models for predicting pedestrian travel demand. In another study, Hankey et al. (2012) developed negative binomial regression models to estimate 12-hour bicycle and pedestrian count volumes. Table 2.2 shows a summary of the literature that adopted sketch plan methods.

Table 2.2. Features of Sketch Plan Method (source: Kim and Susilo, 2011)

Researchers	Level of study area	Observation frequency	Data requirements		Estimation technique
			Pedestrian volume	Land-use and socio-economic data	
Pushkarev and Zupan (1971)	Block (Midtown Manhattan)	Hourly	Pedestrian counts (aerial photography)	Square mile of office, retail, and restaurant space	Linear regression
Behnam and Patel (1977)	Block (CBD of Milwaukee, WI)	Hourly (extrapolated from 6-minute counts)	Pedestrian counts (real counts)	Commercial space, office space cultural and entertainment space, manufacturing space, residential space, parking space vacant space, storage and maintenance space	Linear regression
Davis et al. (1988)	Crosswalk level (Washington, DC)	5- to 10-minute time segments during peak hours	Pedestrian counts (real counts)	Vehicle traffic counts	Relationship between vehicle and pedestrian
Matlick (1997)	Corridor-level (Seattle, WA)	Daily	Transportation mode share information (Census)/National Personal Travel Survey (NPTS)	Housing types, density, persons per household unit, and hotels Retail, recreation, social facilities, schools, employment, and churches	Linear regression
Ercolano et al. (1997)	City level (Plattsburgh, NY)	Hourly (peak hour)	Vehicles per hour from traffic counts and mode share from Census	Vehicle traffic counts	Computation using spreadsheets
Targa and Clifton (2005)	City level (Baltimore City, MD)	One day	Number of walk trips from NHTS 2001	Car ownership in household, type of housing unit, household income, age, sex, driver status, educational status, attitudes/perceptions of pedestrians, household density, street connectivity, land-use diversity, proportion of commercial units	Poisson regression
Kim (2005)	Metropolitan Level (six counties and one city in Baltimore Metropolitan region)	One day	Number of walk trips from NHTS 2001	Age, driver status, education level, income, race, percentage of adult drivers in household, non-residential density (tract level), road density within ¼ mile, mixed land use (tract level)	Poisson regression
Cao et al. (2006)	Town level (six neighbourhoods in Austin, TX)	30 days	Number of pedestrians derived from a self-administered survey mailed in 1995	Major stores within walking distance, traffic volume, pedestrian connections, perception of stores, perception of walk advantage, perception of walk comfort, perception of traffic, miles to the nearest store, sex, age, worker status, presence of children, household income	Negative Binomial regression

Table 2.2. Features of Sketch Plan Method (Continued)

Researchers	Level of study area	Observation frequency	Data requirements		Estimation technique
			Pedestrian volume	Land use and socio-economic data	
Shay et al. (1985)	Town level (Southern Village in Chapel Hill, NC)	One day	Number of walking trips from travel diary	Sex, age, number of children, number of cars/household, number of licensed drivers per car, walking is enjoyable*, environmental protection is important*, value shops and services close by*, distance from home to activity centre *Scale based variables: 1 = strongly disagree to 5 = strongly agree	Negative binomial regression
Pulugurtha and Repaka (2008)	Intersection level (Charlotte, NC)	12 hours (extrapolated from hourly volumes)	Pedestrian counts (real counts)	Household units, population, total employment, urban residential area, neighborhood business, mixed land use, transit stops, speed limit, vehicular volume: All variables above are captured within ¼, ½, and 1 mile buffers	Linear regressions
Baran et al. (2008)	Town level (a New Urbanist community and conventional suburban neighbourhood)	One day	Number of walking trips from NHTS 2001: either leisure or utilitarian walk trips	Age, gender, household size, vehicles per household and respondent's occupational status, two space syntax variables (global integration, local integration, and control variable)	Poisson regression and negative binomial regression
Schneider et al. (2009)	Intersection level (Alameda County, CA)	Weekly (extrapolated from 2 hour volumes with distinction of weekdays and Sundays)	Pedestrian counts (real counts)	Total population, total employment, proportion of housing units (either vacant or rented), number of housing units (either vacant or rented), number of commercial properties, number of elementary/ middle/high schools and colleges, number of transit stations (bus, rail), sidewalk coverage, freeway presence, total street centreline distance, race (Caucasian), car ownership, income, age (categorical variable): All variables above are captured within both 1/10 and ¼ mile buffers Level of traffic, number of lanes, crosswalks, bicycle lanes, traffic signal, and curb radius	Linear regression

Source: Kim and Susilo (2011)

2.1.6. Space syntax method

The space syntax method can also be used to estimate non-motorized travel demand with the aid of regression models. In this model, street and pedestrian network characteristics that are considered independent variables include (Raford and Ragland, 2004):

a) connectivity: number of street segments that are directly connected to a given intersection;

b) mean depth: mean number of street segments between any node and any other node in the network;

c) visibility: view shed area from any node in the network;

d) relative asymmetry or integration: the required number of turns for traveling between two points in the network.

Raford and Ragland (2004) applied the space syntax method to the development of a regression model to estimate pedestrian travel demand in Oakland, California. McCahill and Garrick (2008) also used this method in order to predict morning peak hour bicycle volume in Cambridge, MA.

2.2. Summary

The above literature shows that the focus of most studies is the development of pedestrian estimation models in areas in the United States. Thus, these models may not be used for Canadian cities since the socio-cultural composition, urban form and mobility patterns are different in these two countries. For instance, Pucher and Buehler (2006) indicated that transit and active modes have higher modal share in urban Canada. Table 2.3 shows the modal share of the work trip in Canada and in the United States (Pucher and Buehler, 2006).

Table 2.3. Modal Share of the Work Trip in Canada and USA, 2000/2001

Transport Mode	United States (%)	Canada (%)
Auto	87.9	80.7
Transit	4.7	10.5
Bicycle	0.4	1.2
Walk	2.9	6.6
Other	4.1	1
Total	100	100

Source: Pucher and Buehler (2006)

In addition, most of the existing models do not have an acceptable goodness of fit, and some of the explanatory variables used in these models are not statistically significant. Therefore, most of the current pedestrian models are not capable of predicting pedestrian volumes with reasonably high accuracy. The application of these models is thus most often limited to identify the significant land-use variables affecting the frequency of walking (Hankey et al., 2012; Miranda-Moreno and Fernandes, 2011; Kim and Susilo, 2011; Jones et al., 2010).

CHAPTER 3. DESCRIPTION OF THE DATA AND METHODOLOGIES

Chapter 3 comprises two sections. The first section describes the data used in this research. It also includes a description of study area, the dataset, and the data processing effort in generating explanatory variables using ArcGIS software. The second section discusses the methodology used for developing the linear and Poisson regression models. The over-dispersion test is explained at the end of this section.

3.1. Data

This section describes the study area which consists of 34 major intersections in the city of Calgary, Alberta. It also describes the database used which includes pedestrian and bike count data, GIS data on land use, transportation services, employment, and socioeconomic characteristics.

3.1.1. Study Area

The City of Calgary is located on the Bow River in the southern portion of the province of Alberta, Canada. According to the 2011 census, Calgary's population was 1,096,833 making it the largest city in Alberta. Downtown Calgary is located in the centre of the City of Calgary and consists of five major neighborhoods: Eau Claire; the Downtown West End; the Downtown Commercial Core; Chinatown; and the Downtown East Village. Adjacent to, or directly radiating from the downtown, are the first of the inner-city communities, which, are also surrounded by relatively dense and established neighborhoods (The City of Calgary Land Use Planning and Policy, 2011).

Calgary's middle ring is comprised of approximately 80 neighborhoods that were developed between the 1950s and the 1970s. They now form a reasonably consistent band around the inner city grid neighborhoods, with the exception of the eastern part of the city, where a broad industrial corridor and the Calgary International Airport interrupt the residential fabric. The present dominant pattern of these middle ring neighborhoods consists of warped grid and crescent blocks organized around central schools and recreation fields. These middle ring communities are mostly auto-oriented, which is in conflict with the sustainability plans of the City of Calgary.

The public transportation system in Calgary, including buses and light rail, is provided by the City of Calgary's Calgary Transit. The transit system operates with more than 900 buses and about 200 light rail vehicles (Calgary Transit, 2012). As an alternative to the more than 290 km of shared bikeways on streets, Calgary has a network of multi-use (bicycle, walking, rollerblading, etc.) paths that span more than 700 km (Calgary Pathways and Bikeways, 2012).

3.1.2. Description of the Database

The objective of this research is the development of regression models to relate non-motorized travel demand to environmental and land-use characteristics, socio-economic variables and transportation services. This study focuses on pedestrian and bike count data gathered at 34 major intersections located in Calgary. These observed count data were correlated with several independent variables including: GIS, socio-economic and transportation data, as explained in this section.

3.1.2.1. Pedestrian and bike count data

Pedestrian and bicycle count data at 34 intersections located on major arterials were provided to this research by the City of Calgary. No intersection was selected from downtown area, since downtown urban form and land-use patterns are totally different from the rest of the city. The observations were 6-hour counts in 15 minute time intervals completed during three time intervals in a day: the 7:00–9:00 AM peak hour, the 11:00–13:00 noon peak hour and the 16:00–18:00 PM peak hour. Figure 3.1 shows the location of the 34 intersections in Calgary that were examined, and the descriptive statistics for pedestrian and bicycle counts are shown in Table 3.1.

These counts were taken in different years from 2007 to 2012 and in different months of the year from April to November. Vehicle counts were intersection turning movements, bikes were counted travelling in the same direction as vehicles, and pedestrians were counted as they crossed the leg of the intersection. It is to be noted that one of the major limitations of this data is that pedestrians who crossed more than one leg of the intersection were counted multiple times. Moreover, pedestrians crossing the streets far from the intersections and right-turning pedestrians on the sidewalk were not counted because they did not cross the roadway.

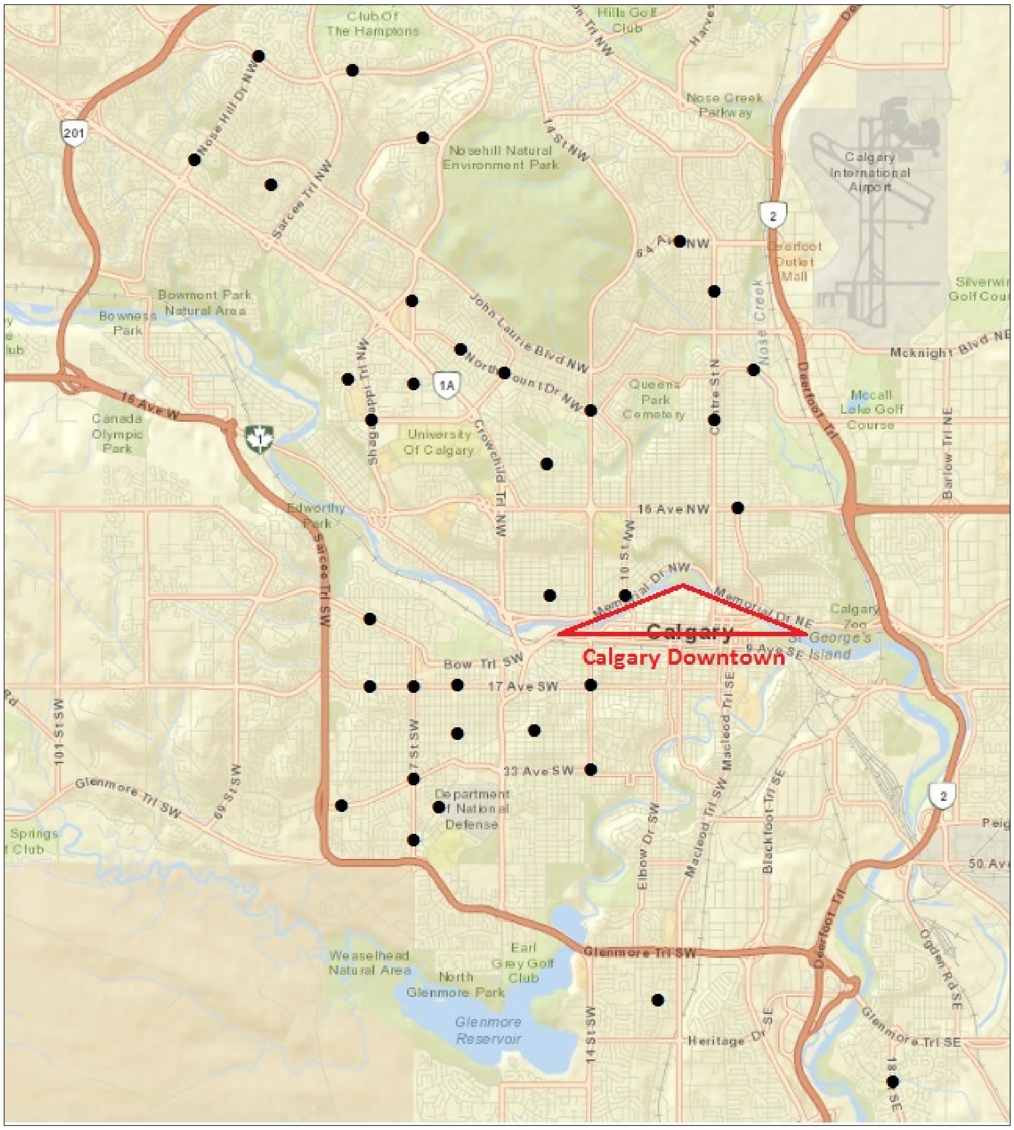

Figure 3.1. Location of the 34 examined intersections in Calgary used for calibrating the non-motorized regression models.

Table 3.1. Description of the Independent Variables

	Variable Name	Description	Number of intersections	Mean	Std. Dev.
Pedestrian Counts	Ped_Ave_AM_Peak	Average number of pedestrians crossing the intersection per an hour during AM peak	34	55.3	63.5
	Ped_Ave_Noon_Peak	Average number of pedestrians crossing the intersection per an hour during Noon peak	34	53.3	61.1
	Ped_Ave_PM_Peak	Average number of pedestrians crossing the intersection per an hour during PM peak	34	72.3	73.9
Bike Counts	Bike_Ave_AM_Peak	Average number of bikes crossing the intersection per an hour during AM peak	34	11.9	9.3
	Bike_Ave_Noon_Peak	Average number of bikes crossing the intersection per an hour during Noon peak	34	6.0	6.1
	Bike_Ave_PM_Peak	Average number of bikes crossing the intersection per an hour during PM peak	34	14.0	9.4

3.1.2.2. Independent (Explanatory) Variables

Twenty seven different independent variables were reviewed for developing the regression models. These explanatory variables belong to three different categories: socio-economic, population and employment characteristics, transportation characteristics and land use variables. These variables were extracted from various sources including GIS data, census data, Calgary Transit and City of Calgary websites. Descriptions of these variables are provided in Table 3.2.

Table 3.2. Description of the Land-Use, Transportation, and Demographic Variables

Category	Variable Name	Description
Socio-economic	POP	Total population
	JOB	Total number of jobs
	POP_Under20	Total population under 18 years old
	POP_Over65	Total population over 65 years old
	POP_20_24	Total population between 20 years old and 24 years old
	Emp	Total employees
	Inc	Average family income
	Transit_Users	Total number of transit users
Land Use	School	Number of schools
	Community_Service	Total number of community service locations
	Institutional	Hectares of institutional space
	ParkRecreationEdu	Hectares of park, recreation and educational space
	Residential_Low	Hectares of low-density residential space
	Residential_Medium	Hectares of medium-density residential space
	Residential_High	Hectares of high-density residential space
	Dwell	Total number of dwellings
	Commercial	Hectares of commercial space
	Direct Control	Hectares of direct control space
	Commercial_Direct	Hectares of the sum of the commercial and direct control spaces
Transportation	Bus_Stop	Number of bus stops
	Bus_Route	Total kilometers of bus routes
	Bus_Frequency	Total vehicle Km of transit routes
	Bikeway	Total centerline kilometers of bikeways
	Pathway	Total centerline kilometers of pathways
	Street-Length	Total centerline kilometers of streets
	Lane	Number of street lanes reaching the intersection
	Bikeway_Lane	Number of bikeway lanes at the intersection

The data were provided to this research in the format of shapefile, which is a popular geospatial vector data format for geographic information system (GIS) software. The explanatory variables were defined using these shapefiles with the aid of ArcGIS software. In the following sections, the explanatory variables are described in detail:

a) Socio-economic variables

- Census data

  The census data on community level from 2006 to 2012 is available online on www.cityonline.calgary.ca website. Explanatory variables such as Number of Employees, Average Family Income, Number of Transit Users, Number of Dwelling Units, and Number of Young Population who are between 20 and 24 years of age are defined using this file (the City of Calgary, Census Data, 2006–2012).

- Plan It Calgary & Calgary Metropolitan Plan Scenario

  This shapefile was provided to the research by the City of Calgary via Spatial and Numeric Data Services at the University of Calgary. This file contains the number of jobs and population for the year 2006 and the forecasted data for the years 2014, 2019, 2024, 2029, 2034, 2039, and 2076. These data are provided on a Transportation Analysis Zone (TAZ) level. The explanatory variables including Population, Number of Jobs, Number of Seniors over 65, and Number of Young Population Under 20 are defined from this shapefile.

b) Land use variables

- Community Service

  This shapefile was obtained from the City of Calgary and was provided to this study via Spatial and Numeric Data Services at the University of Calgary. The file includes locations of different kinds of community services such as community centres, libraries, attractions, courts, hospitals, clinics, post-secondary schools, private schools, public schools, separate schools, visitor information, and social development centers in Calgary.

- School Location

  This shapefile consists of the location of all the schools, colleges, and universities in Calgary. The City of Calgary’s department of Transportation Planning provided this data.

- Land Use

  This shapefile shows the footprint of different types of land uses in Calgary. Therefore land use variables including commercial space, residential space, direct control space, industrial space, major infrastructure, parks, educational space, and recreational space were defined from this file. The land use shapefiles for 2007 and 2011 were obtained from the City of Calgary and were provided to this study via Spatial and Numeric Data Services at the University of Calgary. Since there were no data available to show the type of the buildings and the number of floors, only the footprint of different land use areas were considered for developing the regression models.

c) Transportation variables

- Bus Stops

  This shapefile consists of the location of all bus stops in Calgary. The City of Calgary provided this research with the data.

- Bus Routes

  This shapefile includes location of the bus routes in the year 2011. Using these files variables such as Bus_Route, and Bus_Frequency were defined. This file was also provided to this research by the Spatial and Numeric Data Services at the University of Calgary.

- Bikeways and Pathways

  These data are available online on www.cityonline.calgary.ca website in the shapefile format (the City of Calgary, Calgary Pathways and Bikeways, 2012).

- Street Network

  This shapefile consists of the whole street network in Calgary. Spatial and Numeric Data Services at the University of Calgary provided this research with this data. Variable Street-Length was defined using these data and with the aid of ArcGIS software as explained in the next section.

- Street Lane and Bikeway Lane

  These variables show the number of street lanes and the number of bikeway lanes reaching the intersection which were calculated using Google Maps satellite images.

3.1.3. ArcGIS 10.1 Software

In order to generate the explanatory variables all the shapefiles and intersections were imported into ArcGIS 10.1 software. Each intersection was geocoded and had unique X and Y coordinates. Four different buffer zones were defined around the intersections with radii of 0.1 miles (161 m), 0.25 miles (402 m), 0.50 miles (805 m), and 0.75 miles (1207 m) in order to do multiple scale analysis and to generate the explanatory variables in different scales (Figure 3.2). Hence a total of 108 independent variables were considered in the development of the regression models. It is to be noted that since the created buffer zones were small, each of them only included one intersection that was the main route for cyclists and pedestrians to move in the buffers.

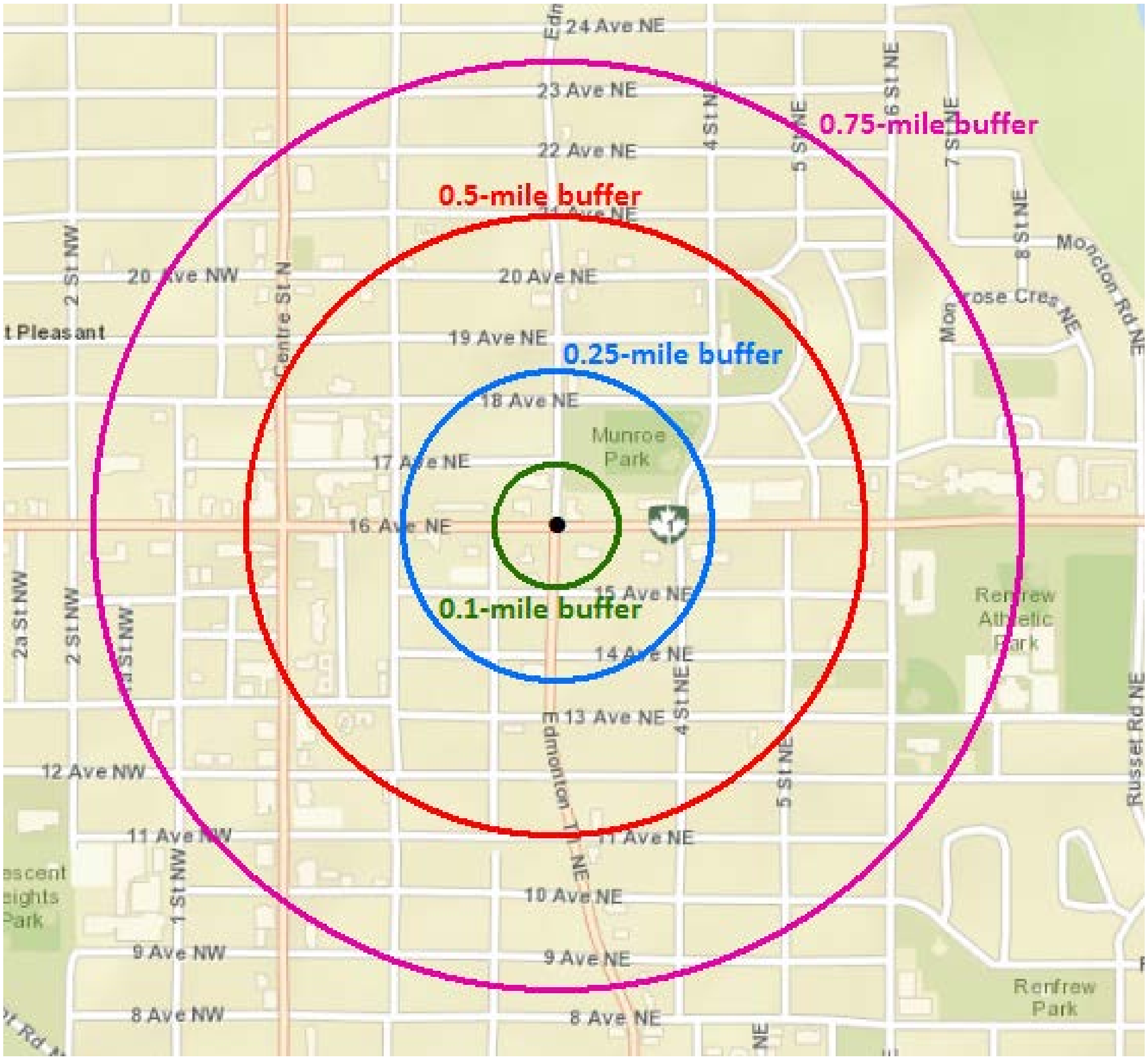

Figure 3.2. Schematic of the buffers around an intersection.

Table 3.3. Descriptive Statistics for Independent Variables

Category	Variable Name	0.1 mile (161 m)		0.25 mile (402 m)	
		Mean	Std. Dev.	Mean	Std. Dev.
Socio-economic	POP	95.0	53.8	597.1	325.0
	JOB	56.1	50.4	332.3	275.1
	POP_Under20	17.7	7.0	113.0	42.7
	POP_Over65	11.7	6.0	74.2	36.4
	POP_20_24	9.0	6.6	56.6	39.4
	Emp	48.9	22.3	305.7	133.6
	Inc	55439.3	24175.1	58758.2	26041.9
	Transit_Users	10.6	5.9	66.3	34.9
Land Use	School	0.0	0.0	0.3	0.4
	Community_Service	0.1	0.3	0.2	0.4
	Institutional	72.2	285.0	9718.2	19090.5
	ParkRecreationEdu	10120.1	10829.6	64429.6	54251.4
	Residential_Low	30997.4	26676.4	248215.4	145326.8
	Residential_Medium	9665.4	11176.3	66347.0	59251.6
	Residential_High	25.5	146.6	1791.8	8189.1
	Dwell	41.0	26.9	259.0	161.6
	Commercial	22141.7	21398.2	63406.4	72121.5
	Direct Control	5305.6	11385.8	35117.2	60723.3
	Commercial_Direct	27447.3	22641.0	98523.7	89753.5
Transportation	Bus_Stop	3.1	1.5	6.0	3.1
	Bus_Route	2.3	1.5	5.9	4.1
	Bus_Frequency	14.4	9.1	37.1	25.3
	Bikeway	176.0	204.3	529.7	516.8
	Pathway	0.1	0.1	0.3	0.4
	Street-Length	1.7	0.5	8.2	1.3
		Mean		Std. Dev.	
	Lane	1.0		1.3	
	Bikeway_Lane	16.0		5.1	

Table 3.4. Descriptive Statistics for Independent Variables (Continued)

Category	Variable Name	0.5 mile (805 m)		0.75 mile (1207 m)	
		Mean	Std. Dev.	Mean	Std. Dev.
Socio-economic	POP	2376.2	1173.7	5193.8	2221.7
	JOB	1282.0	1084.0	3012.0	3305.2
	POP_Under20	450.3	159.6	1004.9	317.1
	POP_Over65	298.4	136.3	653.5	260.1
	POP_20_24	224.9	140.7	486.5	266.0
	Emp	1206.6	472.3	2662.0	987.9
	Inc	60395.9	24129.1	59123.4	18479.2
	Transit_Users	263.9	124.1	584.8	269.7
Land Use	School	1.6	1.2	2.9	1.6
	Community_Service	0.5	0.7	1.2	1.1
	Institutional	55424.5	118313.9	148651.8	302249.8
	ParkRecreationEdu	251408.7	136544.0	618283.7	281210.9
	Residential_Low	1081166.7	499161.4	2417122.1	886650.3
	Residential_Medium	243065.7	210435.3	487529.7	360701.1
	Residential_High	10874.8	55845.6	28091.5	124581.9
	Dwell	1024.0	550.4	2268.0	1102.8
	Commercial	132913.1	152988.4	253775.3	249885.3
	Direct Control	145564.9	216836.4	308012.2	388740.5
	Commercial_Direct	278478.0	262610.0	561787.5	450017.3
Transportation	Bus_Stop	14.9	6.6	32.1	13.5
	Bus_Route	15.2	9.6	34.5	22.5
	Bus_Frequency	92.7	55.0	215.1	132.8
	Bikeway	1551.2	1426.5	3176.5	2159.0
	Pathway	1.1	1.2	2.6	2.7
	Street-Length	30.3	3.1	66.3	6.1
		Mean		Std. Dev.	
	Lane	1.0		1.3	
	Bikeway_Lane	16.0		5.1	

To generate required data in the buffer zones around the intersections census tract, transit, street network and land use data were intersected with the buffers and the data were compiled into excel tables. To generate some of the explanatory variables such as population, jobs, young population, senior population, and employment following formula was used:

$$P_i = \sum_j \frac{A_{j,i}}{A_j} \times P_j$$

where P_i is the population inside buffer zone *i*, P_j is the population of census block *j*, $A_{j,i}$ is the area of census tract inside buffer zone *i*, and A_j is the area of census block *j*.

The variable Average Family Income was generated in the buffer zones using following formula:

$$I_i = \frac{\sum_j \left(I_j \times A_{j,i}\right)}{\sum_j A_{j,i}}$$

where I_i is the average family income inside buffer zone *i*, I_j is the average family income of census block *j*, and $A_{j,i}$ is the area of census tract inside buffer zone *i*.

The variable Bus-Frequency was defined as the Vehicle.Km of transit routes which is expressed as Km of bus routes in a buffer zone multiplied by bus headways over peak duration of 2 hours:

*Bus_Frequency = Km of Bus Routes * (Bus Headways / 2 hr.)*

The bicycle and pedestrian count data provided were from 2007 to 2012. Therefore, in order to develop regression models the counts should have been related to their corresponding independent variables, which were in the same year as the count was. In some cases only the recent shapefiles (for the year 2011 or 2012) were available. All land use variables, except number of dwellings, and all transportation variables were collected in the year 2011 or 2012. Additionally, since the data were not available for previous years, the 2012 land use and transportation data were used for developing the pedestrian and bicycle regression models. This was justified by the fact that most recent development is taking place in the outskirts of the city with little development taking place around the intersections with observation counts.

The data in the 'Plan It Calgary & Calgary Metropolitan Plan Scenario' shapefile were available for the years 2006 and 2014; these data were used to determine the population for the years in between by applying linear interpolation.

The Census data were available for all the years from 2006 to 2012 thus the actual data were used in these years.

After generating the explanatory variables in the buffer zones around the intersections, all the independent variables and dependent variables were imported into SPSS 20.0 software for statistical analysis.

3.2. Description of the Regression Models

In this thesis, the non-motorized trips were calibrated using multiple linear and Poisson regression models. This section describes these models, their assumption and the goodness of fit measures. Then the over-dispersion test is described to fulfill the Poisson distribution property that restricts the mean and variance to be equal.

3.2.1. Linear Regression Theory

One of the applications of linear regression is developing predictive models based on an observed data set of X and Y values. These prediction models can be used to predict the value of Y if an additional value of X is given. In this study multiple linear regression models were calibrated to relate bicycle and pedestrian counts into built environment, transport, and demographic variables.

Linear regression is used in order to find a linear relationship between a dependent variable Y and one or more explanatory variables.

3.2.1.1. Simple linear regression

Simple linear regression consists of one independent variable, and multiple linear regression includes two or more independent variables. Simple linear regression model is expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$$

where Y_i is the dependent variable, X_{1i} is the independent variable, β_0 and β_1 are the constant terms, ε_i is the disturbance term, and $i = 1, 2, 3, \ldots, n$ corresponds to the observation.

3.2.1.2. Multiple linear regression

Multiple linear regression model is expressed in matrix notation and given by:

$$\boldsymbol{Y}_{n\times 1} = \boldsymbol{X}_{n\times p}\boldsymbol{\beta}_{p\times 1} + \boldsymbol{\varepsilon}_{n\times 1}$$

where n is the number of observations and p is the number of explanatory variables.

3.2.1.3. Linear regression goodness-of-fit measures

There are different measurements for assessing linear regression model GOF, including R-squared, adjusted R-squared (Davidson and MacKinnon, 2003). These measurements can be used in order to compare the results of the competing regression models in a single study, or to compare the regression models within several studies.

- R-squared (coefficient of determination)

  R-squared, coefficient of determination, is expressed as:

$$R^2 = \frac{(SST - SSE)}{SST} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$SSE = \sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2,\ SSR = \sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2,\ SST = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

where *SSE* is the Sum of Square Errors, *SSR* is the Regression Sum of Squares, and *SST* is the Total Sum of Squares. Y_i is the observed value, $\hat{Y}_i$ is the predicted value, and $\bar{Y}$ is the mean of the observed data.

The value of R-squared is between 1 and 0. $R^2 = 1$ indicates that fitted model is able to explain all the variation in the dependent variable. And $R^2 = 0$ indicates that the fitted model cannot define any linear relationship between dependent variable and explanatory variables.

- Adjusted R-squared

The adjusted R-squared is more appropriate for comparing models with different numbers of parameters. This measurement of GOF is obtained using following formula:

$$R^2_{adjusted} = 1 - \left(\frac{n-1}{n-p}\right)\left(\frac{SSE}{SST}\right)$$

where n is the number of observations and p is the number of parameters. $n - 1$ is the SSE degree of freedom and $n - p$ is the SST degree of freedom.

- T-statistic test

T-statistics is used to determine the statistically significant parameters entered in the regression model. When the sample size is larger than 30 the t-statistic is normally distributed.

T-statistic is required to check if the entered explanatory variables are statistically significant with a significant percent of α. According to this test, the null (H_0) and alternative (H_A) hypotheses are defined as follows:

$$t - statistic = \frac{k_i}{st.error}$$

$$H_0: k_i = 0$$

$$H_A: k_i \neq 0$$

where k_i is the coefficient for independent variable X_i. If the t-statistic value is significant with a significant level of α, then the null hypotheses is rejected which means that the X_i variable is statistically significant.

3.2.1.4. Linear Regression Assumptions

- Linearity

Linearity assumption requires a linear relationship between dependent variable and explanatory variables. This can be checked by plotting model predicted values versus model residuals. The linearity assumption is satisfied when there is no curvilinear pattern in the plot.

- Uncorrelated disturbances

Uncorrelated disturbances require the observations not to be dependent across time, space, and individuals. For example, the correlation of disturbances across time can be checked by plotting disturbances versus time. When there is no trend in the plot the disturbances are not correlated.

- Normally distributed disturbance

This assumption needs the residuals to be normally distributed and to have the expected value of zero.

- Exogenous independent variables

This assumption requires explanatory variables to be exogenous, which means they should be defined by the factors that are outside the regression model.

3.2.2. Poisson Regression Theory

As bicycle and pedestrian volumes are count data in nature, they can be modeled using a Poisson regression model. In this model, the probability of intersection i having y_i pedestrians or bicycles per specific period (where y_i is a non-negative integer) is given by:

$$P(y_i) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}$$

where $P(y_i)$ is probability of intersection i having y_i pedestrians per specific period; and, λ_i is the Poisson parameter, which is equal to the expected number of pedestrians or bicycles per specific period.

For estimating Poisson regression, the Poisson parameter needs to be calculated as a function of independent variables. The relationship between the Poisson parameter and the variables is usually determined using a log-linear model:

$$\lambda_i = \exp(\beta X_i)$$

where X_i is the vector of independent (explanatory) variables, and β is the vector of estimable parameters. Estimating this model would be much easier using the log of the likelihood (*LL*) function:

$$LL(\beta) = \sum_{i=1}^{n} [-exp(\beta X_i) + y_i \beta X_i - \ln(y_i!)]$$

3.2.2.1. Poisson regression goodness-of-fit measures

Two measures for evaluating the goodness of fit for the Poisson model are ρ^2 and adjusted ρ^2:

$$\rho^2 = \left(1 - \frac{LL_{estimated\ model}}{LL_{base\ model}}\right)$$

$$\rho^2_{adj} = \left(1 - \frac{LL_{estimated\ model} - K}{LL_{base\ model}}\right)$$

where $LL_{estimated\ model}$ is the log-likelihood function of the estimated model, $LL_{base\ model}$ is the log-likelihood function of the base (constant only) model, and K is the number of estimated parameters in the model.

3.2.2.2. Assumption of Poisson Models and Over-Dispersion Test

For properly modeling the count data using Poisson regression, the assumption of equality between the mean and the variance needs to be satisfied, i.e., $E(y_i) = Var(y_i)$. Otherwise, the data are said to be over-dispersed if $E(y_i) < Var(y_i)$ or under-dispersed if $E(y_i) > Var(y_i)$. In cases where the mean does not equal the variance, it is more appropriate to model the count data using negative binomial regression.

Cameron and Trivedi (1990) provided an over-dispersion test. According to this test, the null (H_0) and alternative (H_A) hypotheses are defined as follows:

$$H_0\colon Var\ (y_i) = E(y_i)$$

$$H_A\colon Var\ (y_i) = E(y_i) + \alpha g\big(E(y_i)\big)$$

where α is the over-dispersion parameter, $E(y_i)$ is the predicted count, and $g\big(E(y_i)\big)$ is a function of the predicted count and equals:

$$g\big(E(y_i)\big) = E(y_i)$$

or

$$g\big(E(y_i)\big) = E(y_i)^2$$

This test is conducted by estimating a simple linear regression, where variable Z_i is regressed on W_i:

$$Z_i = bW_i$$

$$Z_i = \frac{\big(y_i - E(y_i)\big)^2 - y_i}{\sqrt{2}E(y_i)}$$

$$W_i = \frac{g\big(E(y_i)\big)}{\sqrt{2}}$$

According to this test, H_0 is rejected if parameter b in the regression model ($Z_i = bW_i$) is statistically significant in either case of $g\big(E(y_i)\big) = E(y_i)$ or $g\big(E(y_i)\big) = E(y_i)^2$. This means that the Poisson regression assumption is not satisfied, and the data would be modeled more appropriately using negative binomial regression.

CHAPTER 4. REGRESSION MODELS RESULTS AND VALIDATION

In the first two sections of this chapter the results of the linear and Poisson regression models are presented. Then the results of the over-dispersion test are discussed. In section 4, the results of the linear regression models are compared versus the results of the Poisson regression models. Finally, in the last section the developed prediction models are validated using another set of count data.

In this research, SPSS 20.0 was used to find the best multiple linear and Poisson regression models for predicting pedestrian and bicycle trip volumes. In the first step all the explanatory variables and count data were imported into the SPSS software.

To calibrate the regression models different pedestrian and bicycle count data were reviewed selected from different time intervals throughout the day. These count data are as follows:

- AM peak hour from 7:00 to 9:00.
- Noon peak hour from 11:00 to 13:00.
- PM peak hour from 16:00 to 18:00.
- 8-hour count done throughout the day: Sum of the AM peak, noon peak, and PM peak counts.
- Average number of counts during AM peak which is the average number of pedestrian/bicycle volumes per hour for the intervals 7:00 to 8:00 and 8:00 to 9:00.
- Average number of counts during noon peak which is the average number of pedestrian/bicycle volumes per hour for the intervals 11:00 to 12:00 and 12:00 to 13:00.

- Average number of counts during PM peak which is the average number of pedestrian/bicycle volumes per hour for the intervals 16:00 to 17:00 and 17:00 to 18:00.

Table 4.1 and 4.2 show different non-motorized count data and the corresponding descriptive statistics.

Table 4.1. Pedestrian Count Data Descriptive Statistics

Variable Name	Description	Number of intersections	Mean	Std. Dev.
Ped_7_9_AM_Peak	Number of pedestrians crossing the intersection during AM peak hour	34	110	128.9
Ped_Ave_AM_Peak	Average number of pedestrians crossing the intersection per an hour during AM peak	34	55.3	63.5
Ped_11_13_AM_Peak	Number of pedestrians crossing the intersection during noon peak hour	34	106.1	123.9
Ped_Ave_Noon_Peak	Average number of pedestrians crossing the intersection per an hour during noon peak	34	53.3	61.1
Ped_16_18_AM_Peak	Number of pedestrians crossing the intersection during PM peak hour	34	144.1	147.8
Ped_Ave_PM_Peak	Average number of pedestrians crossing the intersection per an hour during PM peak	34	72.3	73.9
Ped_Daily	Number of pedestrians crossing the intersection during 8 hours throughout the day	34	260.3	260.0

Table 4.2. Bicycle Count Data Descriptive Statistics

Variable Name	Description	Number of intersections	Mean	Std. Dev.
Bike_7_9_AM_Peak	Number of bicycles crossing the intersection during AM peak hour	34	23.3	18.8
Bike_Ave_AM_Peak	Average number of bicycles crossing the intersection per an hour during AM peak	34	11.9	9.4
Bike_11_13_AM_Peak	Number of bicycles crossing the intersection during noon peak hour	34	11.4	12.2
Bike_Ave_Noon_Peak	Average number of bicycles crossing the intersection per an hour during Noon peak	34	6.0	6.2
Bike_16_18_AM_Peak	Number of bicycles crossing the intersection during PM peak hour	34	28.0	18.6
Bike_Ave_PM_Peak	Average number of bicycles crossing the intersection per an hour during PM peak	34	14.3	9.4
Bike_Daily	Number of bicycles crossing the intersection during 8 hours throughout the day	34	50.4	45.8

Different linear regression models were developed and examined to identify the count data that can best be modeled using the available explanatory variables. According to the results in Tables A.1 to A.14 in Appendix A, the best regression models were developed when average number of pedestrians or bicycles during PM peak hour was selected as the dependent variable of the models. This may be explained by the fact that a significant share of walking and biking trips are usually conducted during PM peak and have recreational and shopping purposes. Most of these trips are done after the school or work hours, or on the way home from school or work. Therefore, it can be argued that PM counts can reflect better non-motorized trips such as shopping and recreational trips in addition to commute trips as compared to AM and noon counts. However AM and noon peak counts can still be used for estimating AM and noon pedestrian volumes.in addition, since the results of the regression models are average pedestrian volumes during PM peak hours, they should be adjusted using the peak hour factors to have the maximum number of pedestrians or cyclists for designing purposes.

Therefore, further statistical analysis was conducted using average number of bicycle and pedestrian volumes during PM peak to find the best Linear and Poisson regression models with the highest possible adjusted R-squared and adjusted ρ-squared, respectively.

4.1. Results of the Multiple Linear Regression Models

In order to find the best bicycle and pedestrian linear regression models the following method was applied:

a. Explanatory variables that had a high correlation with non-motorized volumes were selected.

b. Among the selected variables, those variables that had a high correlation with each other were removed.

c. Variables that were not statistically significant were removed from the model.

d. Finally, the model with the best overall fit and highest adjusted coefficient of determination ($R^2_{adjusted}$) was selected.

Tables 4.3 and 4.4 show the coefficients for the calibrated linear regression models for estimating pedestrian and bicycle travel demand. The regression models were both statistically significant with a confidence level of 99.9%. The t-statistics of all parameters were statistically significant and had the proper sign.

The adjusted R^2 for the pedestrian and bicycle prediction models were 0.921 and 0.900, respectively, which shows that the models fit the data and could provide a good estimation of non-motorized travel demand.

4.1.1. Linear Regression Model for Pedestrians

Table 4.3 shows the explanatory variables for calibration of the pedestrian linear regression model. The positive sign of the explanatory variables coefficients indicates that these variables had a positive effect on the pedestrian volumes at intersections.

The number behind the variable indicates the buffer zone to which the variable belongs. For example, School_0.50 indicates the total number of schools in the buffer zone of 0.50 miles around the intersection.

Table 4.3. Multiple Linear Regression Model for Pedestrians

Dependent Variable = Ped_Ave_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-213.7520	41.421	-5.161	0.000
Bus_Stop_0.10*	7.9988	3.351	2.387	0.025
Commercial_0.25	2.5400	0.730	3.500	0.002
Transit_User_0.50	0.1812	0.039	4.662	0.000
BusFrequency_0.75	0.0910	0.330	2.711	0.005
Street-Length_0.50	0.0040	0.001	2.957	0.007
Pathway_0.25	25.5860	11.552	2.215	0.036
School_0.50	15.4880	3.442	4.499	0.000
Job_0.75	0.0074	0.002	4.758	0.012
Overall Model				
R^2	0.940			
R^2_{adj}	0.921			
Sig.	0.000			

* The number behind the variable indicates the buffer zone to which the variable belongs.

According to this model, better transportation services, such as higher bus frequency and more bus stops, can help attract higher pedestrian volumes at intersections since walking is a vital part of public transportation. The positive sign of variables such as commercial space, number of jobs and number of schools indicates that having high-density, mixed-use communities can also increase pedestrian travel demand.

Street-Length_0.50 shows the centerline kilometers of streets in a 0.5-mile buffer zone. This variable can be interpreted as connectivity. Increasing the number of streets with pedestrian sidewalks in a certain buffer area results in a denser street network in that area, which improves connectivity. Better connectivity in an area can increase pedestrian volumes.

Moreover, the calibrated pedestrian model indicates that pedestrian facilities, such as separate pathways, can encourage people to walk more. This may be attributed to the fact that intensifying the pedestrian pathway network and building separate pedestrian pathways may result in improved pedestrian and cyclist safety, since these pathways are mostly located at a distance from streets and roads. This would result in a decrease in non-motorized exposure and, thus, increase the volume of the active travel modes.

The pedestrian prediction model also shows that the frequency of walking is higher among transit users since walking is a vital part of public transportation.

4.1.2. Linear Regression Model for Cyclists

Table 4.4. Multiple Linear Regression Model for Cyclists

Dependent Variable = Bike_Ave_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Institutional_0.50*	0.664	0.090	7.785	0.000
Residential_Low_0.10	2.250	0.360	6.207	0.000
Commercial_0.10	2.030	0.680	2.966	0.006
Lane	-0.504	0.158	-3.197	0.003
Bus_Stop_0.25	1.218	0.359	3.396	0.002
Overall Model				
R^2	0.915			
R^2_{adj}	0.900			
Sig.	0.000			

* The number behind the variable indicates the buffer zone to which the variable belongs.

The coefficients of the bicycle linear regression model are shown in Table 4.4. According to this table, land-use variables such as Institutional_0.50, Residential_Low_0.10, and Commercial_0.10 have positive signs and indicate the positive impact of high density and mixed land use on bicycle volumes. In this research different residential variables were used including low density, medium density, and high density residential. However, since the amount of high and medium density residential in Calgary is very low these variables are not included in the regression models. and it can be said that in this research the low density residential variable represents the general residential land use in Calgary.

The variable Lane is defined as the total number of street lanes reaching the intersection in all directions. The negative impact of this variable may be attributed to safety problems that cyclists have on wide streets. According to the *2009 University of Calgary Commuter Cyclist Survey Report* (Twaddle and Hall, 2009), the first barrier to the potential cyclist commuters was the lack of safe routes for cyclists. This report also mentioned that the three most requested on-route improvements by current and potential cyclists were the provisions of “more bicycle lanes on city roads, more pathways and more direct cycle routes” (Twaddle and Hall, 2009). Therefore, considering both the bicycle prediction model and the *2009 University of Calgary Commuter Cyclist Survey Report*, by allocating one lane of street lanes to cyclists, there will be an increase in bicycle trip volumes. These findings highlight the important role that road narrowing plays as an efficient traffic calming technique that forces motorists to reduce their speeds, thereby resulting in improved safety for cyclists and attracting more cyclists to use the infrastructure.

Using a bicycle in combination with transit may be a reason for the positive impact of the number of bus stops on bicycle demand. In Calgary, some bus routes are provided with bike racks (the City of Calgary, Calgary Pathways and Bikeways map, 2012). Cyclists are also able to use Calgary's light rail transit (C-Train) on weekends and during non-peak weekday hours (the City of Calgary, Calgary Pathways and Bikeways map, 2012). Another reason for the positive impact of the number of bus stops on bicycle volumes may be a greater willingness of transit users to use bicycles.

4.2. Results of the Poisson Regression Models

For developing the Poisson regression model, independent variables were selected according to the following four steps:

a. Variables with a low correlation ($|\rho|<0.2$) with the dependent variable were removed.

b. Independent variables that had a high correlation ($|\rho|<0.650$) with each other were removed.

c. Independent variables that were not statistically significant at a significance level of 0.05 were removed from the model.

d. Finally, the Poisson model with the highest adjusted ρ^2 value was selected.

Tables 4.5 and 4.6 show the coefficients and goodness of fit measures for the calibrated Poisson regression models for estimating pedestrian and bicycle travel demand. The adjusted ρ^2 value for the calibrated pedestrian and bicycle Poisson regression models were 0.790 and 0.321, respectively. Both models were statistically significant at a confidence level of 99.9%.

4.2.1. Poisson Regression Model for Pedestrians

All explanatory variables in the pedestrian Poisson regression model were positive and had a positive effect on pedestrian volumes. The variable Dwell_0.50 shows the number of dwelling units in a 0.50-mile buffer zone and indicates that an increased number of dwelling units and, thus, denser population results in an increase in pedestrian travel demand.

Other variables in this model have the same interpretations as those discussed for the pedestrian linear regression model.

Table 4.5. Poisson Regression Model for Pedestrians

Dependent Variable = Ped_Ave_PM_Peak			
Variable Name	Coefficient	Std. Error	Sig.
Bus_Stop_0.10*	0.0730	0.0220	0.001
StreetLength_0.50	0.0771	0.0023	0.000
BusFrequency_0.75	0.0017	0.0002	0.000
Dwell_0.50	0.0003	0.0000	0.000
Commercial_0.25	0.0498	0.0035	0.000
School_0.50	0.2100	0.0175	0.000
Pathway_0.25	0.2590	0.0708	0.000
Overall Model			
LL(β)	-203.128		
LL(0)	-999.394		
ρ^2	0.797		
ρ^2_{adj}	0.790		
Sig.	0.000		

* The number behind the variable indicates the buffer zone to which the variable belongs.

4.2.2. Poisson Regression Models for Cyclists

The explanatory variables in the bicycle Poisson regression model were exactly the same as those in the bicycle linear regression model and had the same interpretations.

Table 4.6. Poisson Regression Model for Cyclists

Dependent Variable = Bike_Ave_PM_Peak			
Variable Name	Coefficient	Std. Error	Sig.
Constant	1.5570	0.2518	0.000
Commercial_0.10*	0.1560	0.0389	0.000
Lane	-0.0200	0.0104	0.055
Residential_Low_0.10	0.1413	0.0258	0.000
Bus_Stop_0.25	0.0590	0.0186	0.002
Institutional_0.50	0.0341	0.0036	0.000
Overall Model			
LL(b)	-105.261		
LL(0)	-162.312		
ρ^2	0.351		
ρ^2_{adj}	0.321		
Sig.	0.000		

* The number behind the variable indicates the buffer zone to which the variable belongs.

4.3. Over-Dispersion Test

The theoretical background of the over-dispersion test was explained in Chapter 4. This section discusses how the over-dispersion test was conducted for bicycle and pedestrian Poisson regression models.

The over-dispersion test, as proposed by Cameron and Trivedi (1990), was conducted for the calibrated Poisson regression models. According to this test, both the pedestrian and bicycle Poisson regression models' p-values for coefficient b were not statistically significant, at a confidence level of 95%, under different cases of $g\big(E(y_i)\big) = E(y_i)$ and $g\big(E(y_i)\big) = E(y_i)^2$.

This shows that the data did not have any over-dispersion and that the assumption of Poisson regression was satisfied in both prediction models. Therefore, there is no need to model the data using negative binomial regression model and the data can be modeled properly with the aid of Poisson regression model.

Table 4.7 shows the results of the p-value for coefficient *b* in different cases, according to the over-dispersion test.

Table 4.7. Over-Dispersion Test

Poisson Regression Model	$g(E(y_i)) = E(y_i)$		$g(E(y_i)) = E(y_i)^2$	
	b	p-value	b	p-value
Pedestrian	0.023	0.236	0.000	0.814
Bicycle	-0.011	0.556	0.000	0.489

4.4. Comparing the results of linear and Poisson regression models

In this section the results of the linear regression models are compared with the results of the Poisson regression models.

4.4.1. Pedestrian prediction models

Figures 4.1 and 4.2 show observed values versus predicted values of pedestrian demand for linear and Poisson regression models, respectively. Points neighboring the 45 degree line indicate the lower error in perdicted values in comparison with the observed values. The Pearson correlations between observed and forecast pedestrian travel demand were 0.970 and 0.971 in linear and Poisson regression models respectively. The correlations values are very close to each other and both of them are very high.

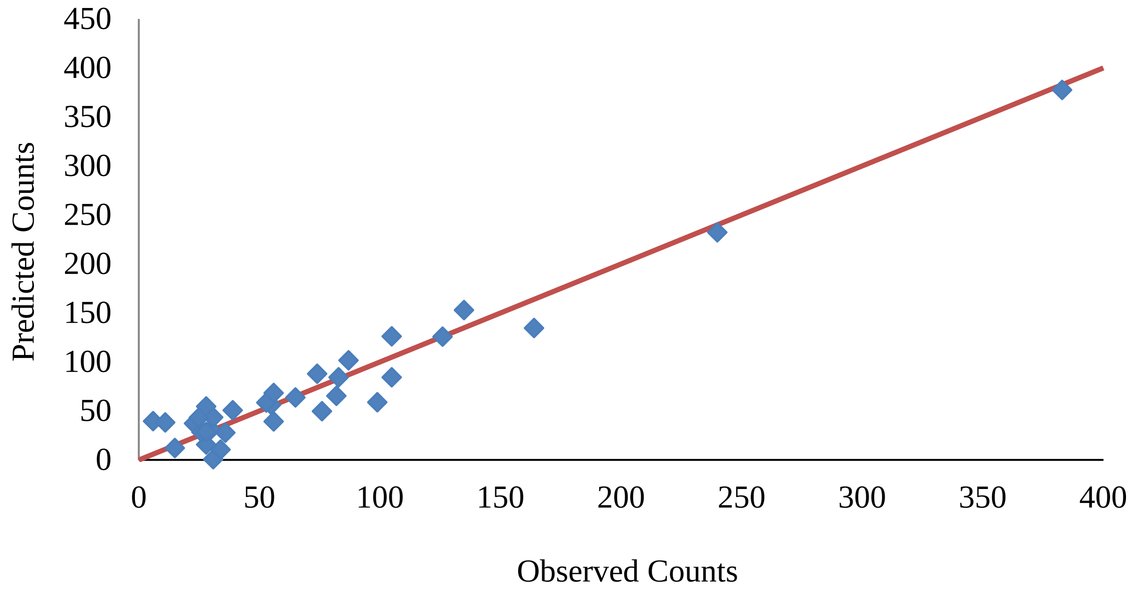

Figure 4.1. Correlation between observed and predicted pedestrian counts for linear regression model.

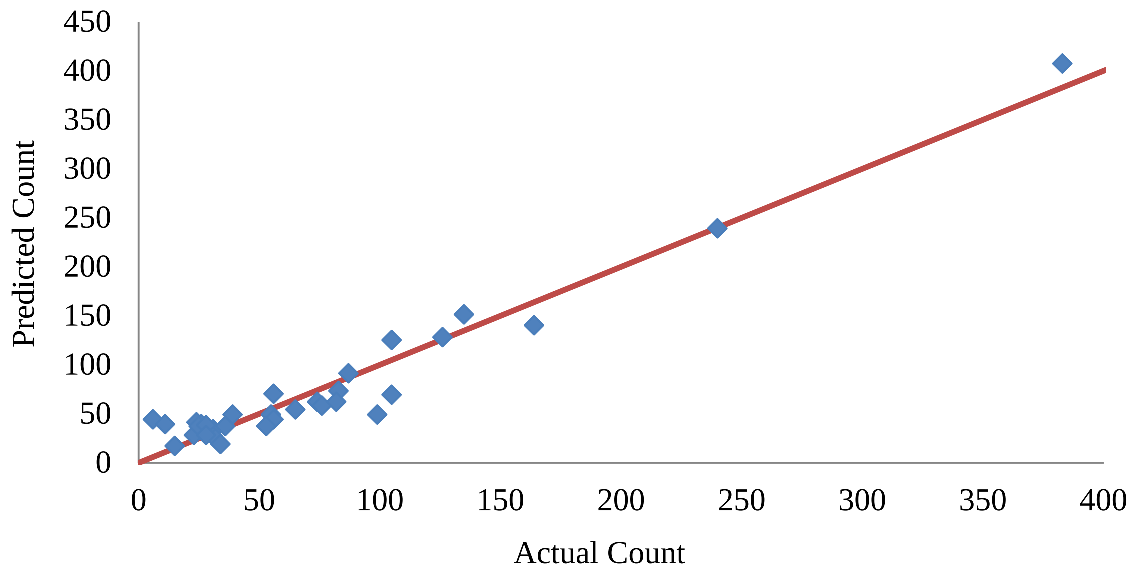

Figure 4.2. Correlation between observed and predicted pedestrian counts in Poisson regression model.

The linear and Poisson regression models residuals are demonstrated in Appendix B in Figures B.1 and B.2, respectively. Differences between observed and predicted values of pedestrian travel demand ranged between almost zero to 40 in linear regression model and zero to 50 in Poisson regression model.

Presented graphs in this section indicate that none of the multiple linear and Poisson regression models for estimating pedestrian travel demand is better or worse than the other one. Both models can be used for forecasting pedestrian volumes in Calgary.

4.4.2. Bicycle prediction models

Figures 4.3 and 4.4 show observed values versus predicted values of bicycle demand for linear and Poisson regression models, respectively. Points neighboring the 45 degree line indicate the lower error in perdicted values in comparison with the observed values. The Pearson correlations between observed and forecast pedestrian travel demand were 0.844 and 0.857 in linear and Poisson regression models respectively. The correlations values are very close to each other and both of them are high.

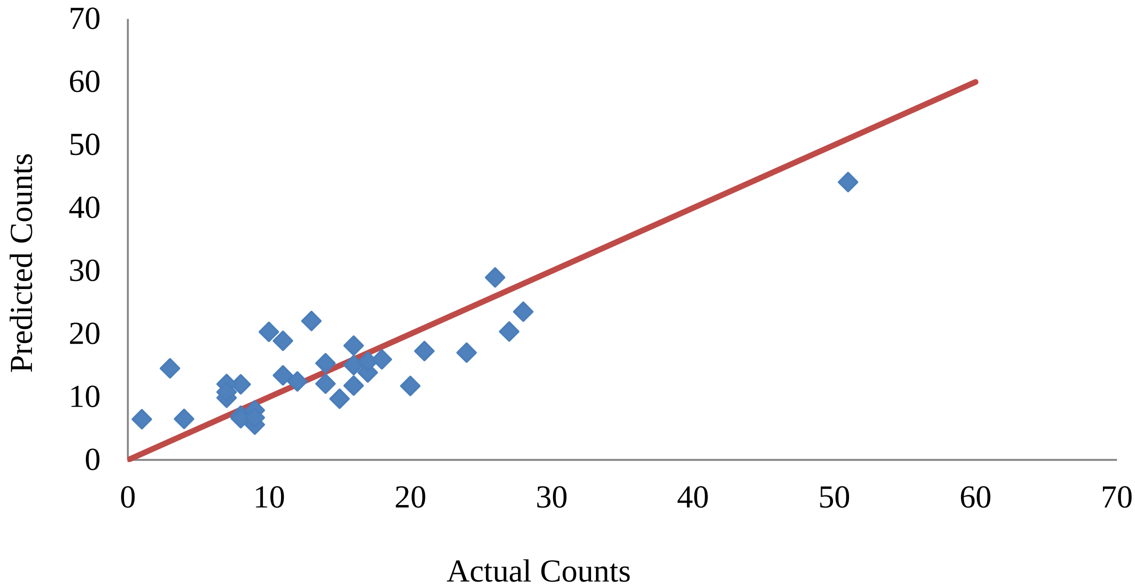

Figure 4.3. Correlation between observed and predicted bicycle counts in linear regression model.

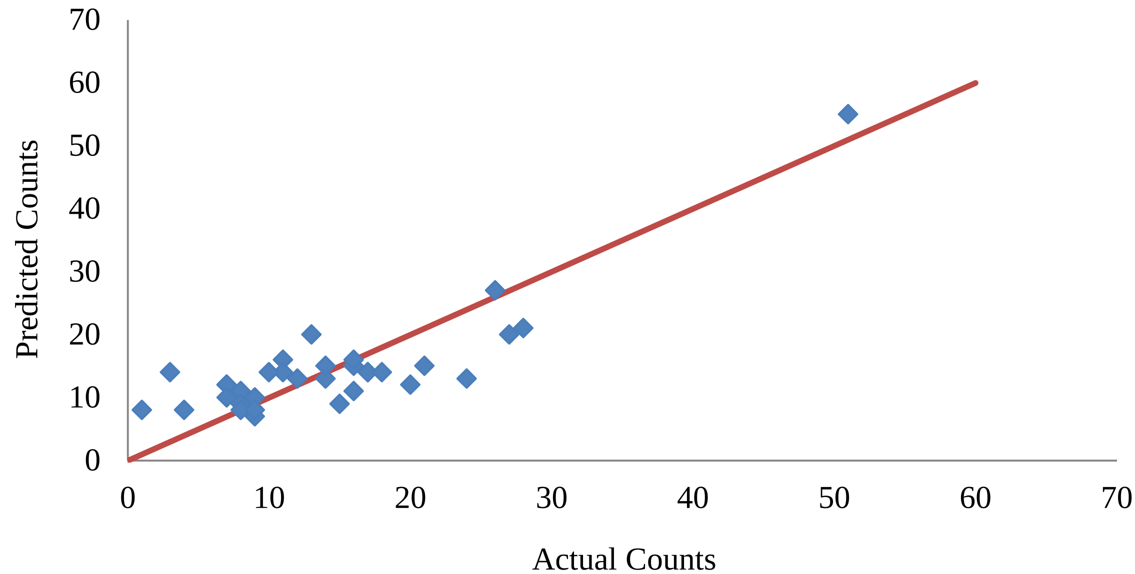

Figure 4.4. Correlation between observed and predicted bicycle counts in Poisson regression model.

The linear and Poisson regression models residuals are demonstrated in Appendix B in Figures B.3 and B.4, respectively. Differences between observed and predicted values of pedestrian travel demand ranged between almost zero to 11 in both linear and Poisson regression models.

Presented graphs in this section and in Appendix B indicate that none of the multiple linear and Poisson regression models for estimating bicycle travel demand is better or worse than the other one. As well, both models can be used for forecasting bicycle volumes in Calgary.

4.5. Model Validation

To validate prediction models developed in this research, a sample of 18 intersections was selected from Glamorgan community, a middle ring community in South West Calgary (Figure 4.5). The explanatory variables that are needed for estimating bicycle and pedestrian travel demand were generated for these intersections using GIS maps and with the aid of ArcGIS 10.1 software. The explanatory variables were then used in linear and Poisson regression models in order to predict bicycle and pedestrian volumes at the intersections.

Figure 4.5. Locations of sample intersections selected for model validation.

Figure 4.6 to 4.9 demonstrate predicted values versus observed values for bicycle and pedestrian volumes using different prediction models. Points neighboring the 45 degree line indicate the lower error in perdicted values in comparison with the observed values.

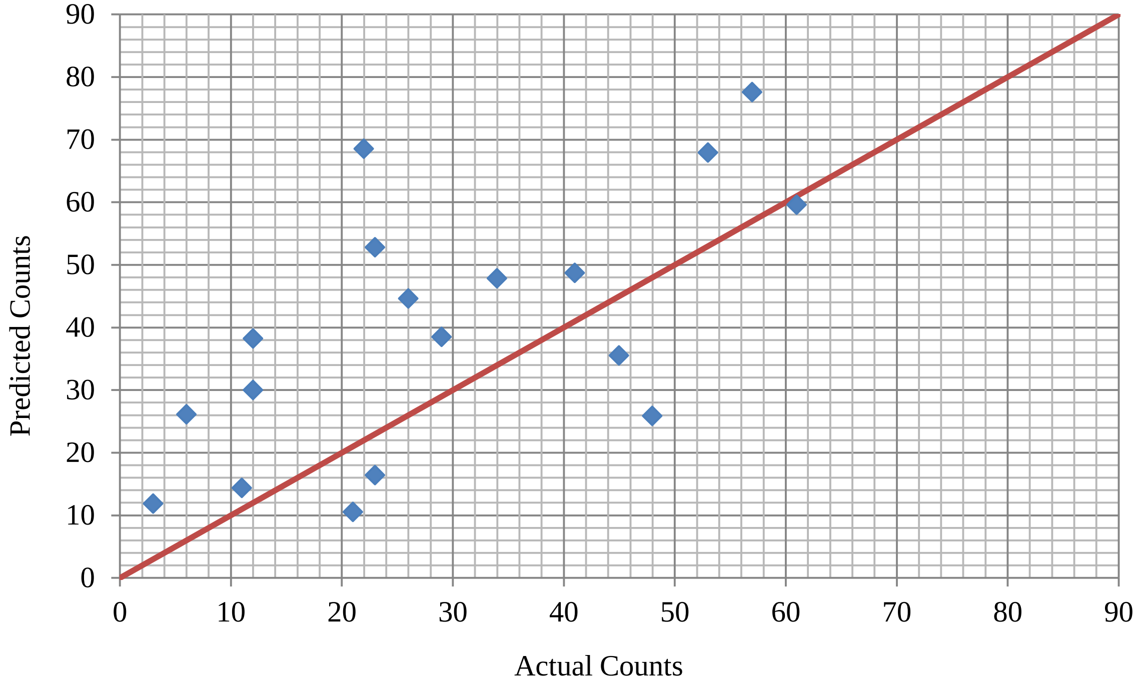

Figure 4.6. Observed values VS predicted values using pedestrian linear regression.

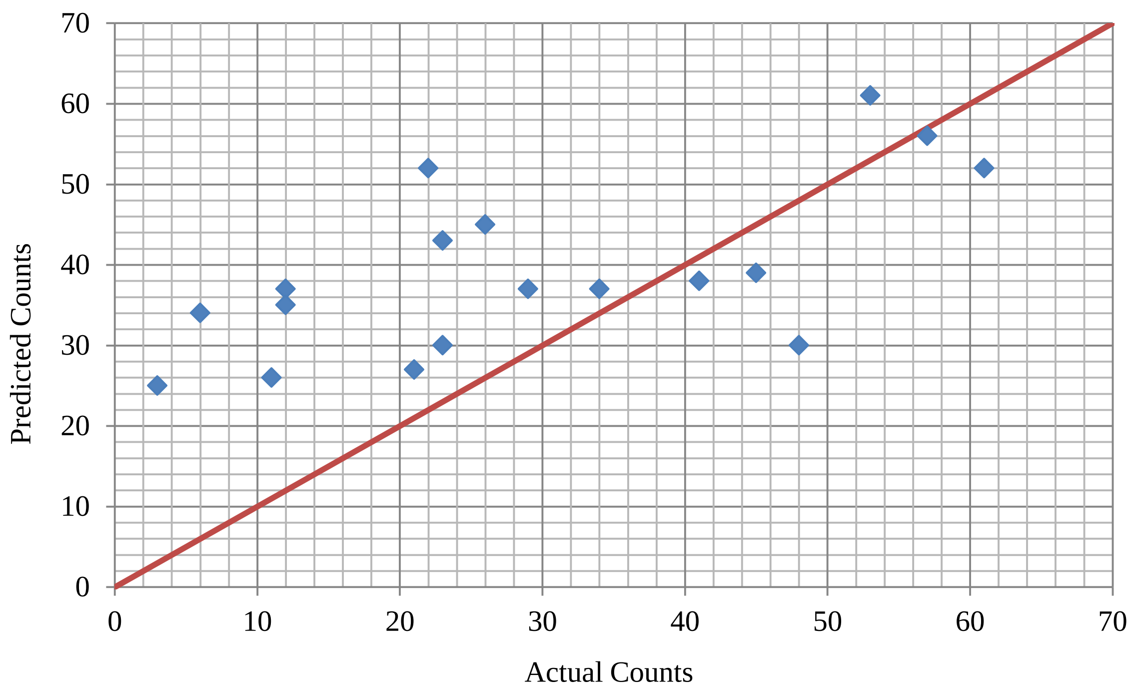

Figure 4.7. Observed values VS predicted values using pedestrian Poisson regression.

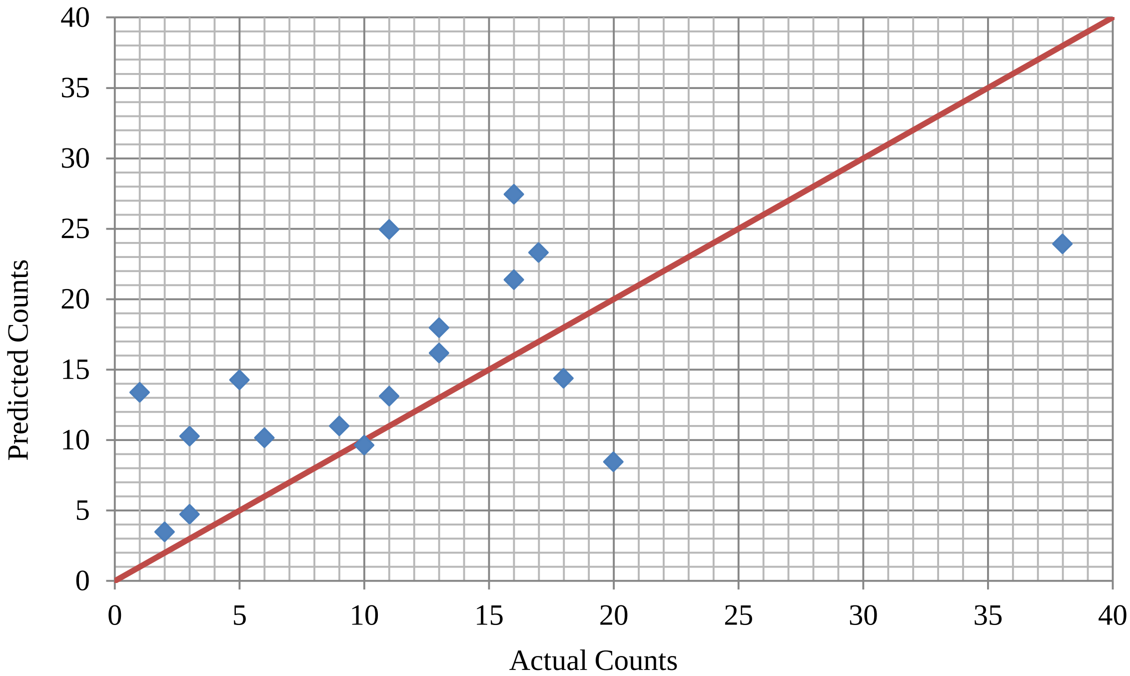

Figure 4.8. Predicted values VS observed values using bicycle linear regression.

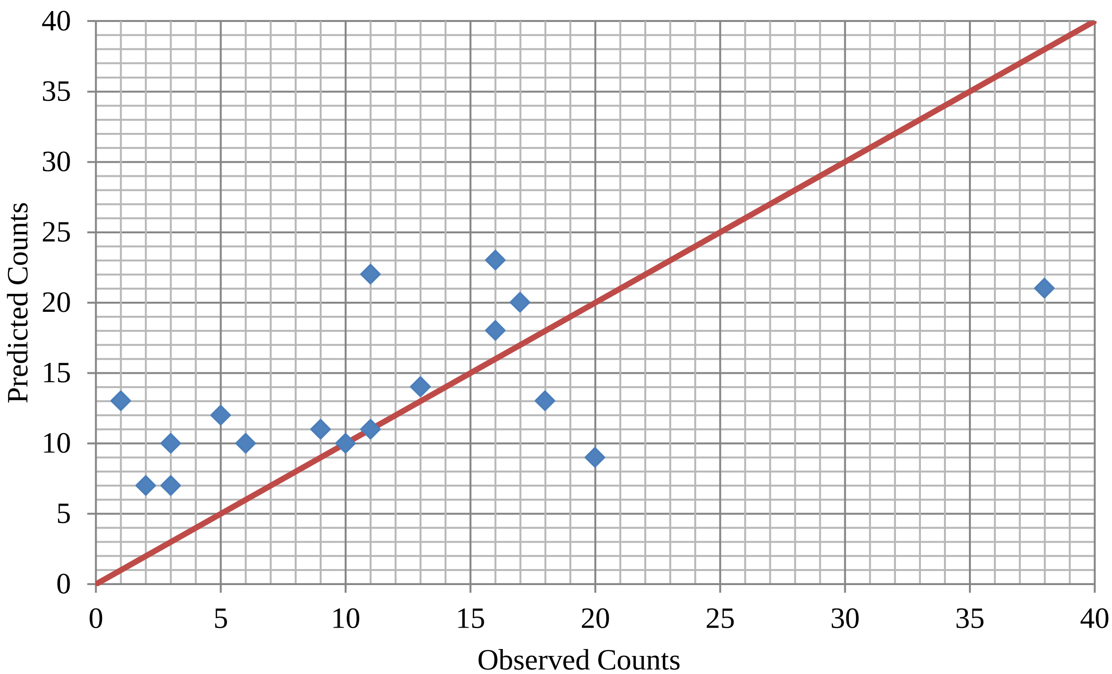

Figure 4.9. Predicted values VS observed values using bicycle Poisson regression.

The linear and Poisson regression models residuals for predicting pedestrian demand are demonstrated in Appendix C in Figures C.1 and C.2, respectively. Differences between observed and predicted values of pedestrian travel demand ranged between almost zero to 50 in linear regression model and zero to 30 in Poisson regression model.

The linear and Poisson regression models residuals for predicting bicycle demand are demonstrated in Appendix C in Figures C.3 and C.4, respectively. Differences between observed and predicted values of pedestrian travel demand ranged between almost zero to 15 in linear regression model and zero to 17 in Poisson regression model.

It can be seen that the majority of the predicted values are overestimated. This may be due to most of the observed intersections for developing the regression models were selected from the arterial and collector roadway intersections in Calgary while the intersections used for validating the models were mostly considered as minor intersections.

In addition to the figures discussed previously, two measures were used in order to determine the prediction accuracy: Pearson product-moment correlation coefficient and root mean square deviation.

4.5.1. Pearson product-moment correlation coefficient

This coefficient, r_{ij}, indicates the degree of linear association between two variables i and j, and ranges from +1, perfect positive correlation, to -1, perfect negative correlation. r_{ij} is given by:

$$r_{ij} = \frac{S_{ij}}{S_i S_j}$$

where S_{ij} is covariance between variables i and j. and S_i and S_j are the standard variations of variables i and j, respectively. This variable can show the predictability power of the developed models.

Table 4.8 shows the values of Pearson correlation between observed and predicted values for linear and Poisson regression models. According to this table there are high positive correlations between observed and predicted values for different bicycle and pedestrian prediction models.

Table 4.8. Pearson Correlations between Observed Values and Predicted Values

Observed Values	Bicycle Prediction Values		Pedestrian Prediction Values	
	Linear Regression	Poisson Regression	Linear Regression	Poisson Regression
Pedestrian	NA*	NA	0.634	0.633
Bicycle	0.586	0.592	NA	NA

* Not Applicable

4.5.2.Root mean square deviation

The root mean square deviation (RMSD) or root mean square error (RMSE) is usually used to measure the differences between observed values and predicted values. The RMSD is a good accuracy measure to compare forecasting errors of different models for a particular variable. The RMSD ranges from zero to infinity. And the lower values indicate less variance in residuals. RMSD is expressed as:

$$RMSD = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}}$$

where $\hat{y}_i$ is the predicted value and y_i is the actual value.

Table 4.9 shows the RMSE values for different prediction models for estimating bicycle and pedestrian volumes.

Table 4.9. RMSE Values for Different Prediction Models

Pedestrian Linear Model	Bike Linear Model	Pedestrian Poisson Model	Bike Poisson Model
19.182	7.8023	16.635	7.161

According to this table bicycle linear and Poisson regression models have similar prediction accuracy. This table also indicates that the pedestrian Poisson regression model has slightly better prediction accuracy than pedestrian linear regression model.

4.6. Summary

In this chapter the results of the developed linear and Poisson regression models for predicting bicycle/pedestrian trip volumes were presented.

According to the pedestrian models the following suggestions can be made in order to encourage more people to shift to walking mode:

- Enhancing transportation services with higher bus frequency and more bus stops. As well as encouraging more people to use public transit services.
- Developing high-density and mixed-use communities.
- Improving street network connectivity.
- Providing more extensive and safer pedestrian infrastructure such as separate pedestrian pathway network.

Considering the results from the bicycle regression models, the following recommendations can be implemented in order to encourage more people to use bicycles as their preferred travel mode:

- Applying traffic calming policies in order to increase cyclists' safety.
- Allocating separate on street lanes for bike users.
- Developing high-density and mixed-use communities.
- Improving transit and bike integration.

CHAPTER 5. SUMMARY AND DISCUSSIONS

This chapter presents the concluding comments of this thesis and suggests directions for future research. Overall conclusions and discussions are presented in section 5.1. Section 5.2 presents the author's perspective on the contributions of the research to the non-motorized demand modeling problem. Section 5.3 discusses the limitations and shortcomings in the research. Finally, section 5.4 describes the recommendations for future works.

5.1. Conclusions

The aim of this study was the development of prediction models for estimating non-motorized travel demand in the City of Calgary, Canada. For each pedestrian and bicycle demand, two empirical models were developed — one multiple linear regression model and one Poisson regression model — in order to relate pedestrian/bicycle counts to land-use variables, socio-economic characteristics and transportation services.

The required data — bicycle and pedestrian counts, GIS, transportation, and socioeconomic data — for developing the regression models were obtained from different sources including the City of Calgary, Spatial and Numeric Data Services at the University of Calgary, Calgary transit, census data, etc. Using these data and with the aid of ArcGIS software 108 explanatory variables were defined and reviewed for calibrating the regression models.

The developed pedestrian and bicycle linear regression models had adjusted R^2 values of 0.921 and 0.900, respectively; and, the models were statistically significant at a significance level of 0.01. The Poisson regression models for estimating bicycle and pedestrian travel demand had adjusted ρ-squared values of 0.790 and 0.321, respectively; and, the models were statistically significant at a significance level of 0.01. Although these values were not very high in comparison with the R-squared values, the results of the Poisson regression models were almost the same as the results of the linear regression models. This was shown by comparing the graphs demonstrated correlations between observed and predicted values for both linear and Poisson regression models, and also by examining the residual values for these models. The Pearson correlations between observed and forecast travel demand in linear and Poisson regression models were also close to each other. In addition, the over-dispersion test showed that the data can be modeled properly using Poisson regression model and there is no need to model them with aid of negative binomial regression model.

Moreover, the validation process conducted for a sample of 18 intersections in Calgary showed that the bicycle and pedestrian prediction values estimated using linear and Poisson regression models have adequate predictability. Additionally, the predictabilities of both linear and Poisson regression models are almost in the same level.

The models developed in this research can be used to assess the role of urban design and built environments, such as increases in population and employment density, mixed land use and building of complete streets in Calgary's middle ring communities, on the demand for active travel modes. Moreover, the model can be used to estimate the qualitative impact of development scenarios and transportation policies on pedestrian and bicycle demand. The developed prediction models in this research indicate that improved pedestrian and bicycle infrastructure, such as improved pedestrian network connectivity and pathway length, improved transportation service integration, such as transit and bicycle integration, and safer routes for pedestrian and cyclists make a significant contribution to increasing non-motorized travel volumes.

The calibrated empirical models may also be used to predict future non-motorized travel demand for improving non-motorized infrastructure in different parts of the city or for conducting bicycle/pedestrian safety and public health studies. Moreover, these models can be used to estimate non-motorized travel demand at intersections with no available counts. The method in this research is a straight-forward statistical analysis for practitioners, and the needed data is relatively easy to access.

5.2. Research Contributions

According to the literature, in the field of non-motorized prediction models the focus of most studies is the development of pedestrian estimation models in areas in the United States. Most of these models do not have an acceptable goodness of fit, and some of the explanatory variables used in these models are not statistically significant. Therefore, most of the current pedestrian models are not capable of predicting future pedestrian volumes with reasonably high accuracy. The application of these models is most often limited to identifying the land-use and transportation element variables affecting the frequency of walking. Furthermore, despite these efforts, there is still a lack of an applicable model for the estimation of pedestrian and bicycle volumes in Canadian cities, especially for a highly auto-oriented city, such as Calgary.

Thus, the developed regression models in this thesis made the following contributions:

i. *Introducing the first bicycle regression model in Canada, and first pedestrian regression model in Calgary:*

The developed cyclist regression models in this research are the first bicycle prediction models calibrated for a Canadian city. Additionally, the developed pedestrian regression models are the first pedestrian prediction models calibrated for the City of Calgary.

v. *Having the ability to predict future bicycle/pedestrian travel demand:*

This study adopts a sketch plan approach to model pedestrian and cyclist volumes in Calgary. The developed models have a relatively high goodness of fit, and all explanatory variables in these models are statistically significant. Therefore, these models can be used for the purpose of predicting bicycle and pedestrian travel demand as well as determining the variables that affect the frequency of walking and biking.

vi. *Considering the effect of road narrowing and traffic calming technique on bicycle travel demand:*

The variable Lane, which is defined as the total number of street lanes reaching the intersection, is introduced in this research for the first time. The negative impact of this variable on the bicycle trip volumes indicates that the cyclists are more willing to bike in narrower streets, which is attributed to safety problems that cyclists have on wide streets. This variable also highlights the important role that road narrowing plays as an efficient traffic calming technique that forces motorists to reduce their speeds, thereby resulting in improved safety for cyclists and attracting more cyclists to use the infrastructure.

vii. *Introducing the variable "Street-Length" to consider the impact of connectivity on pedestrian travel demand:*

The variable Street-Length is another variable introduced in this research for the first time. This variable shows the centerline kilometers of streets in a specific buffer zone around the intersection and can be interpreted as connectivity. Increasing the number of streets with pedestrian sidewalks in a certain buffer area results in a denser street network in that area, which improves connectivity; and, better connectivity in an area can increase pedestrian volumes.

5.3. Limitations

As is the case with any research, this research suffers from its limitations and shortcomings. The first limitation that can be named for this research is that the sample size of observed intersections is small since only 34 intersections used for the development of the regression models. With this small sample size of 34 intersections, some important variables that were not captured and were not included in the models, may have a significant relationship with non-motorized trips.

The second limitation is in the method used for counting pedestrians at the intersections. In pedestrian count data, pedestrians who crossed more than one leg of the intersection were counted multiple times. Moreover, pedestrians crossing the streets far from the intersections and right-turning pedestrians on the sidewalk were not counted because they did not cross the roadway.

The pedestrian and bicycle count data provided to this research were taken in different years from 2007 to 2012. This requires that the corresponding explanatory variables to be in the same year as the intersection counts are. However, some of the explanatory variables such as land use, pathways and bikeways, and number of schools are just available for the year 2012. This was justified by the fact that most recent urban development is taking place in the outskirts of the city with little development taking place around the intersections with observation counts.

Land use variables including commercial space, residential space, direct control space, industrial space, major infrastructure, parks, educational space, and recreational space are only the footprint of different types of land uses in Calgary. Since there were no data available about the building type and the number of floors for each building, these variables only indicate a rough estimation of different land use spaces in Calgary.

5.4. Future Works

The following future research can be considered imminent to the presented work and are suggested as:

- The number of observed intersections used for the development of the regression models should be extended to include more intersections in the City of Calgary.

- Due to the lack of data, some explanatory variables, such as: the number of private cars owned by families; weather conditions; sidewalk coverage; existence of bicycle racks; and the existence of vehicle parking and the number of different land use parcels (except number of dwelling units) were not examined in this research in the development of the regression models. These variables may be considered for calibrating the regression models in future work.

- New bicycle/pedestrian prediction models may be developed for downtown Calgary, as the downtown street network, land use, and travel patterns are completely different from the other parts of the city.

- The models developed in this research were calibrated for the City of Calgary. Different non-motorized prediction models can be developed for other Canadian cities or other countries.

- In this research a sample of 18 intersections were used for conducting model validation, in the future a larger sample may be used for validating the developed non-motorized prediction models.

- Using the developed models in this research, the estimated non-motorized travel demand may be predicted under different land use and transportation scenarios and the results may be simulated using micro-simulation software. Using this means the predicted traffic behaviour will be visually represented through animation to enable politicians and stakeholders to fully perceive the impacts and consequences of different proposed scenarios.

- For collecting non-motorized count data more accurately, smart phone applications can be used to collect information on non-motorized trips in a given buffer zone. This will alleviate the limitations of the data collection effort and avoid doubly counting the pedestrian and bike volume. It will also take into consideration the non-motorized trips activity in a larger buffer zone rather than only on major intersections.

REFERENCES

Ashley, C.A., and C. Banister. Cycling to Work from Wards in a Metropolitan Area. *Traffic Engineering and Control*, Vol. 30, 1989, pp. 297–302.

Baran P.K., D. Rodriguez, and A. Khattak. Space Syntax and Walking in a New Urbanist and Suburban Neighbourhoods. *Journal of Urban Design*, Vol. 13, 2008, pp. 5–28.

Behnam, J., and B.G. Patel. A Method for Estimating Pedestrian Volume in a Central Business District. *Transportation Research Record: Journal of the Transportation Research Board, No. 629*, Transportation Research Board of the National Academies, Washington, D.C., 1977, pp. 22–26.

Calgary Transit. http://www.calgarytransit.com/html/fleet_information.html. Accessed November 10, 2012.

Cameron, A.C., and P.K. Trivedi. Regression-based Tests for Overdispersion in the Poisson Model. *Journal of Econometrics*, Vol. 46 (3), 1990, pp. 347–364.

Cameron, R.M. Pedestrian Volume Characteristics. *Institute of Transportation Engineers Compendium of Technical Papers*, Vol. 47, 1977, pp. 36–37.

Cao, X., S.L. Handy, and P.L. Mokhtarian. The Influences of the Built Environment and Residential Self-selection on Pedestrian Behavior: Evidence from Austin, TX. *Transportation*, Vol. 33, 2006, pp. 1–20.

Cervero, R., and M. Duncan. Walking, Bicycling, and Urban Landscapes: Evidence from the San Francisco Bay Area. *American Journal of Public Health*, Vol. 93 (9), 2003, pp. 1478–1483.

Davidson, R., and J.G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2003.

Davis S.E., L.E. King, and H.D. Robertson. Predicting Pedestrian Crosswalk Volumes. *Transportation Research Record: Journal of the Transportation Research Board, No. 1168*, Transportation Research Board of the National Academies, Washington, D.C., 1988, pp. 22–26.

Dill, J. Bicycling for Transportation and Health: The Role of Infrastructure. *Journal of Public Health Policy*, Vol. 30 (Suppl. 1), 2009, pp. S95–S110.

Ercolano, J.M., J.S. Olson, and D.M. Spring. Sketch Plan Method for Estimating Pedestrian Traffic for Central Business District and Suburban Growth Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1578*, Transportation Research Board of the National Academies, Washington, D.C., 1997, pp. 38–47.

Federal Highway Administration. G*uidebook on Methods to Estimate Non-Motorized Travel: Supporting Documentation*. Publication FHWA-RD-98-166. FHWA, U.S. Department of Transportation, 1999.

Federal Highway Administration. *Guidebook on Methods to Estimate Non-Motorized Travel Overview of Methods.* Publication FHWA-RD-98-165. FHWA, U.S. Department of Transportation, 1999.

Griswold, J.B., A. Medury, and R.J. Schneider. Pilot Models for Estimating Bicycle Intersection Volumes. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2247*, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 1–7.

Guo, J., C. Bhat, and R. Copperman. Effect of the Built Environment on Motorized and Non-motorized Trip Making: Substitutive, Complementary, or Synergistic? *Transportation Research Record: Journal of the Transportation Research Board, No*. 2010, 2007, pp. 1–11.

Handy, S. Critical Assessment of the Literature on the Relationships among Transportation, Land Use, and Physical Activity. *Transportation Research Board Special Report 282*, Available online: http://trb.org/downloads/sr282papers/sr282Handy.pdf , 2005.

Hankey, S., G. Lindsey, X. Wang, J. Borah, K. Hoff, B. Utecht, and Z. Xu. Estimating Use of Non-Motorized Infrastructure: Models of Bicycle and Pedestrian Traffic in Minneapolis, MN. *Landscape and Urban Planning*, Vol. 107 (3), 2012, pp. 307–316.

Haynes, M., and S. Andrzejewski. *GIS Based Bicycle & Pedestrian Demand Forecasting Techniques*. Travel Model Improvement Program Webinar. Available online: http://tmiponline.org/Clearinghouse/Items/20100429_-_GIS-based_Bicycle_and_Pedestrian_Demand_Forecasting_and_Traffic_Count_Programs.aspx. April 29, 2010.

Hunt, J.D., and J.E. Abraham. Influences on Bicycle Use. *Transportation: Planning, Policy, Research, Practice.* Vol. 34 (4), 2007, pp. 453–470.

Jones, M., S. Ryan, J. Donlon, L. Ledbetter, D.R. Ragland, and L. Arnold. *Seamless Travel: Measuring Bicycle and Pedestrian Activity in San Diego County and its Relationship to Land Use, Transportation, Safety, and Facility Type*. Caltrans Task Order 6117. California Department of Transportation, 2010.

Kim, N.S., and Y.O. Susilo. Comparison of Pedestrian Trip Generation Models. *Journal of Advanced Transportation*, Vol. 47 (4), 2011, pp. 399–412.

Kim, N.S., Trip Generation Model for Pedestrians Based on NHTS 2001. Master's Thesis, Department of Civil Engineering, University of Maryland, College Park, United States, 2005.

Krizek, K. Operationalizing Neighborhood Accessibility for Land Use — Travel Behavior Research and Regional Modeling. *Journal of Planning Education and Research*, Vol. 22 (3), 2003, pp. 270–287.

Levinson H.S., and F.H. Wynn. Effects of Density on Urban Transportation Requirements. *Highway Research Record, No. 2*, 1963, pp. 38–64.

Lewis, C.B., and J.E. Kirk. Central Massachusetts Rail Trail Feasibility Study. Central Transportation Planning Staff, Boston, MA, Massachusetts Bay Transportation Authority, 1997.

Lindsey, G., J. Wilson, E. Rubchinskaya, J. Yang, and Y. Han. Estimating Urban Trail Traffic: Methods for Existing and Proposed Trails. *Landscape and Urban Planning*, Vol. 81 (4), 2007, pp. 299–315.

Liu, X., and J. Griswold. Pedestrian Volume Modeling: Case Study of San Francisco. *Yearbook of the Association of Pacific Coast Geographers*, Vol. 71, 2009, pp. 164–181.

Matlick J.M. If We Build it, Will They Come? (Forecasting Pedestrian Use and Flows), 1996; 315–319.

McCahill, C., and N.W. Garrick. The Applicability of Space Syntax to Bicycle Facility Planning. *Transportation Research Record: Journal of the Transportation Research Board, No. 2074*, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 46–51.

Miranda-Moreno, L.F., and D. Fernandes. Modeling of Pedestrian Activity at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board, No. 2264*, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 74–82.

Miranda-Moreno, L.F., P. Morency, and A. El-Geneidy. How Does the Built Environment Influence Pedestrian Activity and Pedestrian Collisions at Intersections? *89th Annual Meeting of the Transportation Research Board*, Washington DC, 2010.

Nelson, A., and D. Allen. If You Build Them, Commuters Will Use Them. *Transportation Research Record: Journal of the Transportation Research Board, No. 1578*, Transportation Research Board of the National Academies, Washington, D.C., 1997, pp. 79–83.

Plan It Calgary. http://www.calgary.ca/PDA/LUPP/Pages/Municipal-Development-Plan/Plan-It-Calgary/Plan-It-Calgary.aspx?redirect=/planit, 2009.

Plan Your Place. *Plan Your Place A Geospatial Infrastructure for Sustainable Community Planning: Literature Review, 2011*. http://planyourplace.ca/publications/Literature%20Review_v_2.2.pdf. Accessed July 3, 2013.

Porter, C., J. Suhrbier, and W. L. Schwartz. Forecasting Bicycle and Pedestrian Travel: State of the Practice and Research Needs. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1674*, Transportation Research Board of the National Academies, Washington, D.C., 1999, pp. 94–101.

Pucher, J., and R. Buehler. Why Canadians Cycle More than Americans: A Comparative Analysis of Bicycling Trends and Policies. *Transportation Policy*, Vol. 13 (1), 2006, pp. 265–279.

Pulugurtha, S.S., and S.R. Repake. Assessment of Models to Measure Pedestrian Activity at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2073*, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 39–48.

Purvis, C. *Incorporating Effects of Smart Growth and TOD in San Francisco Bay Area Travel Demand Models: Current and Future Strategies*. Available online: http://www.mtc.ca.gov/maps_and_data/datamart/research/Incorporating_Smart_Growth_MTC_models.pdf, 2003.

Pushkarev, B., and J.M. Zupan. Pedestrian Travel Demand. *Highway Research Record, No. 355*, National Research Council, Washington, D.C., 1971, pp. 37–53.

Qin, X., and J.N. Ivan. Estimating Pedestrian Exposure Prediction Model in Rural Areas. *Transportation Research Record: Journal of the Transportation Research Board, No. 1773*, Transportation Research Board of the National Academies, Washington, D.C., 2001, pp. 89–96.

Raford, N., and D. Ragland. Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. *Transportation Research Record: Journal of the Transportation Research Board, No. 1878*, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 66–74.

Reynolds, K.D., J. Wolch, J. Byrne, C. P. Chou, G. Feng, and S. Weaver. Trail Characteristics as Correlates of Urban Trail Use. *American Journal of Health Promotion*, Vol. 21(4 Suppl.), 2007, pp. 335–345.

Schneider R., L. Arnold, and D. Ragland. A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. *Transportation Research Record: Journal of the Transportation Research Board, No. 2140*, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 13–26.

Schneider, R.J., L.S. Arnold, and D.R. Ragland. Methodology for Counting Pedestrians at Intersections: Use of Automated Counters to Extrapolate Weekly Volumes from Short Manual Counts. *Transportation Research Record: Journal of the Transportation Research Board, No. 2140*, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 1–12.

Shay E., Y. Fan, D. Rodriguez, and A. Khattak. Drive or walk?: Utilitarian Trips within a Neotraditional Neighborhood. *Transportation Research Record: Journal of the Transportation Research Board, No. 1985*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 154–161.

Targa F., and K Clifton. Built Environment and Trip Generation for Non-Motorized Travel. *Journal of Transportation and Statistics*, Special Edition (NHTS), 2004. Available online at: http://onlinepubs.trb.org/onlinepubs/archive/conferences/nhts/Clifton.pdf.

The City of Calgary Land Use Planning and Policy. *Beltline Area Redevelopment Plan*, 2011. http://web.archive.org/web/20110611041400/http://www.calgary.ca/DocGallery/BU/planning/pdf/centre_city/beltline/beltline_plan_one.pdf. Accessed January 28, 2013.

The City of Calgary, Calgary Pathways and Bikeways. https://cityonline.calgary.ca/Pages/Product.aspx?category=PDCTransportation&cat=CITYonlineDefault&id=PDC0-99999-99999-00508-P. Accessed November 10, 2012.

The City of Calgary, Census Data https://cityonline.calgary.ca/Pages/Category.aspx?cat=CITYonlineDefault&category=PDCCensusInformation&publicdata. Accessed July 6, 2012.

The City of Calgary, Pathways and Bikeways Map, Calgary, 2012. http://www.calgary.ca/CSPS/Parks/Documents/Pathways/path-bike-map.pdf. Accessed July 16, 2013.

Twaddle, H., and F. Hall. *2009 University of Calgary Commuter Cyclist Survey Report*. http://www.calgary.ca/Transportation/TP/Documents/cycling/2009_uofc_commuter_cyclist_surv_rep.pdf. Accessed July 16, 2013.

Washington, S.P., M.G. Karlaftis, and F.L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis.* Boca Raton, FL, Chapman & Hall/CRC, 2003.

Wilbur Smith Associates. *Non-motorized Access to Transit: Final Report*. Prepared for Regional Transportation Authority, Chicago, Ill., 1996.

APPENDIX A. PEDESTRIAN LINEAR REGRESSION MODELS

Tables A.1 to A.7 show pedestrian linear regression models developed using stepwise method:

Table A.1. Linear Regression Model for Morning Pedestrian Volumes

Dependent Variable = Ped_7_9_AM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-83.8650	32.630	-2.570	0.015
Commercial_0.50	0.0004	0.000	4.329	0.000
School_0.50	51.6180	11.600	4.450	0.000
Km_BusRoute_0.25	0.0090	0.004	2.604	0.014
Overall Model				
R^2	0.633			
R^2_{adj}	0.596			
Sig.	0.000			

Table A.2. Linear Regression Model for Average AM Peak Pedestrian Volumes

Dependent Variable = Ped_Ave_AM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-41.5670	16.311	-2.548	0.016
Commercial_0.50	0.0002	0.000	4.329	0.000
School_0.50	25.7960	5.798	4.449	0.000
Km_BusRoute_0.25	0.0050	0.002	2.597	0.014
Overall Model				
R^2	0.633			
R^2_{adj}	0.596			
Sig.	0.000			

Table A.3. Linear Regression Model for Noon Pedestrian Volumes

Dependent Variable = Ped_11_13_Noon_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-83.6670	35.130	-2.382	0.024
Commercial_0.1	0.0030	0.001	4.627	0.000
Transit_Users_0.75	0.1360	0.054	2.496	0.018
School_0.50	24.3610	11.282	2.159	0.039
Overall Model				
R^2	0.633			
R^2_{adj}	0.596			
Sig.	0.000			

Table A.4. Linear Regression Model for Average Noon Peak Pedestrian Volumes

Dependent Variable = Ped_Ave_Noon_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-41.7990	17.569	-2.379	0.024
Commercial_0.1	0.0020	0.000	4.624	0.000
Transit_Users_0.75	0.0680	0.027	2.503	0.018
School_0.50	12.2330	5.642	2.168	0.038
Overall Model				
R^2	0.633			
R^2_{adj}	0.597			
Sig.	0.000			

Table A.5. Linear Regression Model for Evening Pedestrian Volumes

Dependent Variable = Ped_16_18_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-164.9260	27.598	-5.976	0.000
JOB_0.75	0.0120	0.004	3.164	0.004
Commercial_0.1	0.0020	0.000	5.852	0.000
Transit_Users_0.75	0.1430	0.047	3.064	0.005
BusFrequency_0.75	0.0003	0.000	3.836	0.001
POP_P65_0.75	0.1170	0.045	2.619	0.014
Overall Model				
R^2	0.916			
R^2_{adj}	0.901			
Sig.	0.000			

Table A.6. Linear Regression Model for Average PM Peak Pedestrian Volumes

Dependent Variable = Ped_Ave_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-82.1450	13.781	-5.961	0.000
JOB_0.75	0.0060	0.002	3.162	0.004
Commercial_0.1	0.0010	0.000	5.850	0.000
Transit_Users_0.75	0.0710	0.023	3.064	0.005
BusFrequency_0.75	0.0001	0.000	3.850	0.001
POP_P65_0.75	0.0590	0.022	2.623	0.014
Overall Model				
R^2	0.916			
R^2_{adj}	0.902			
Sig.	0.000			

Table A.7. Linear Regression Model for Daily Pedestrian Volumes

Dependent Variable = Ped_Daily				
Variable Name	Coefficient	Std. Error	t	Sig.
(Constant)	-160.3560	64.279	-2.495	0.018
Km_BusRoute_0.75	0.0050	0.001	4.150	0.000
School_0.50	85.0970	22.732	3.744	0.001
Commercial_0.50	0.0010	0.000	4.359	0.000
Overall Model				
R^2	0.653			
R^2_{adj}	0.619			
Sig.	0.000			

Tables A.8 to A.14 show bicycle linear regression models developed using the stepwise method:

Table A.8. Linear Regression Model for Morning Bicycle Volumes

Dependent Variable = Bike_7_9_AM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	8.4210	5.014	1.679	0.103
Institutional_0.50	0.0001	0.000	3.338	0.002
POP_20_24_0.1	1.1770	0.423	2.781	0.009
Overall Model				
R^2	0.342			
R^2_{adj}	0.299			
Sig.	0.002			

Table A.9. Linear Regression Model for Average AM Peak Bicycle Volumes

Dependent Variable = Bike_Ave_AM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	4.5490	2.526	1.801	0.081
Institutional_0.50	3.80E-05	0.000	3.265	0.003
POP_20_24_0.1	0.5870	0.213	2.753	0.010
Overall Model				
R^2	0.334			
R^2_{adj}	0.291			
Sig.	0.002			

Table A.10. Linear Regression Model for Noon Bicycle Volumes

Dependent Variable = Bike_11_13_Noon_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-32.5060	4.728	-5.957	0.000
Institutional_0.75	3.99E-05	0.000	9.538	0.000
POP_20_24_0.1	0.6100	0.148	3.520	0.002
Bus_Stop_0.25	1.1820	0.381	3.645	0.001
Residential_Low_0.25	3.10E-05	0.000	1.858	0.074
Commercial_0.1	3.06E-04	0.000	3.294	0.003
Residential_Low_0.1	2.21E-04	0.000		
BusFrequency_0.75	1.91E-05	0.000	2.715	0.011
Overall Model				
R^2	0.856			
R^2_{adj}	0.817			
Sig.	0.000			

Table A.11. Linear Regression Model for Average Noon Peak Bicycle Volumes

Dependent Variable = Bike_Ave_Noon_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-11.6100	2.114	-5.492	0.000
Institutional_0.75	0.0000	0.000	9.119	0.000
POP_20_24_0.1	0.2790	0.083	3.355	0.002
Bus_Stop_0.25	0.7570	0.207	3.665	0.001
Residential_Low_0.1	0.0002	0.000	5.352	0.000
Commercial_0.1	0.0001	0.000	2.998	0.006
Overall Model				
R^2	0.805			
R^2_{adj}	0.771			
Sig.	0.000			

Table A.12. Linear Regression Model for Evening Bicycle Volumes

Dependent Variable = Bike_16_18_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Constant	-17.0950	7.384	-2.315	0.028
Institutional_0.50	0.0001	0.000	7.244	0.000
Bus_Stop_0.25	2.3260	0.749	3.104	0.004
Residential_Low_0.1	0.0010	0.000	4.860	0.000
Commercial_0.1	0.0003	0.000	2.218	0.035
Residential_High_0.25	0.0010	0.000	2.179	0.038
Overall Model				
R^2	0.715			
R^2_{adj}	0.664			
Sig.	0.000			

Table A.13. Linear Regression Model for Average PM Peak Bicycle Volumes

Dependent Variable = Bike_Ave_PM_Peak				
Variable Name	Coefficient	Std. Error	t	Sig.
Institutional_0.50	0.0001	0.000	6.581	0.000
Residential_High_0.25	0.0003	0.000	2.453	0.020
Bus_Stop_0.25	1.1100	0.192	5.768	0.000
Residential_Low_0.1	0.0001	0.000	4.928	0.000
Overall Model				
R^2	0.897			
R^2_{adj}	0.883			
Sig.	0.000			

Table A.14. Linear Regression Model for Daily Bicycle Volumes

Dependent Variable = Bike_Daily				
Variable Name	Coefficient	Std. Error	t	Sig.
Institutional_0.50	0.0003	0.000	5.883	0.000
Residential_Low_0.1	0.0010	0.000	4.884	0.000
Bus_Stop_0.25	6.6750	1.747	3.821	0.001
Lane	-1.9370	0.821	-2.359	0.025
Overall Model				
R^2	0.814			
R^2_{adj}	0.789			
Sig.	0.000			

APPENDIX B. RESIDUAL ANALYSIS FOR PREDICTION MODELS

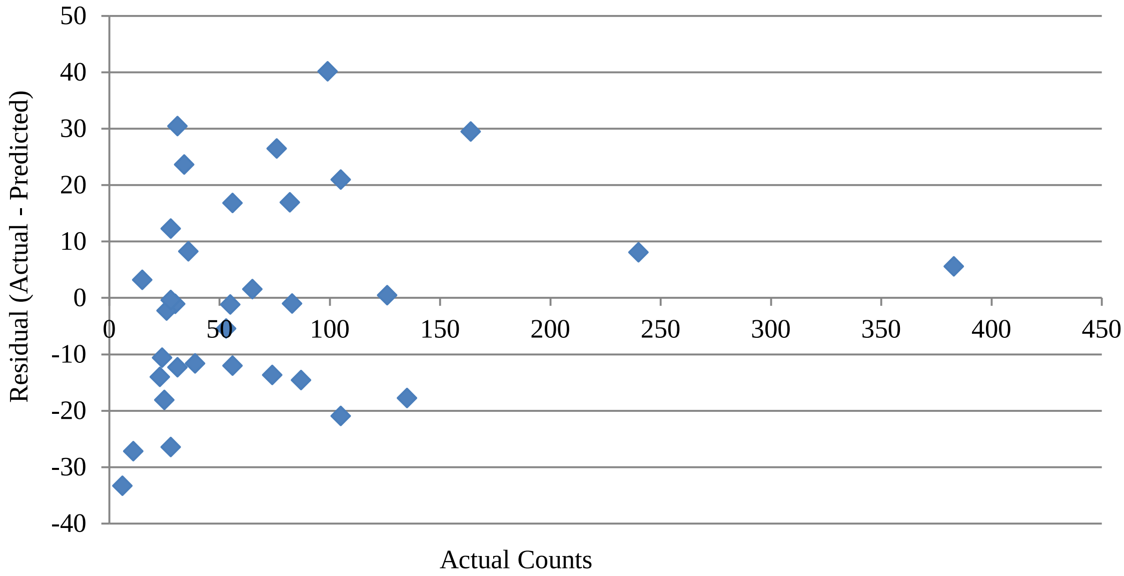

Figure B.1. Residual analysis for pedestrian linear regression model.

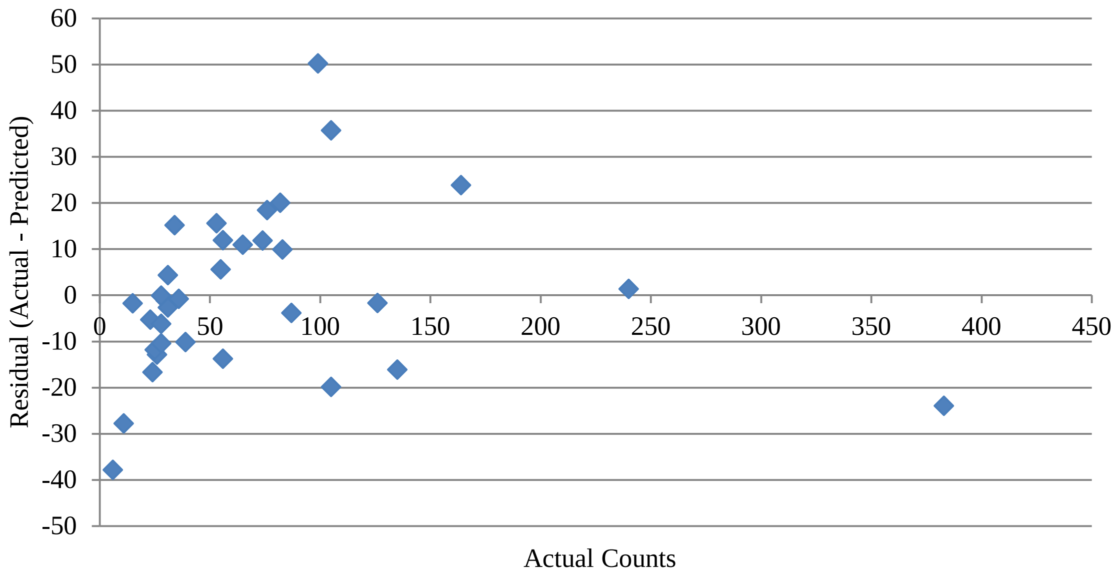

Figure B.2. Residual analysis for pedestrian Poisson regression model.

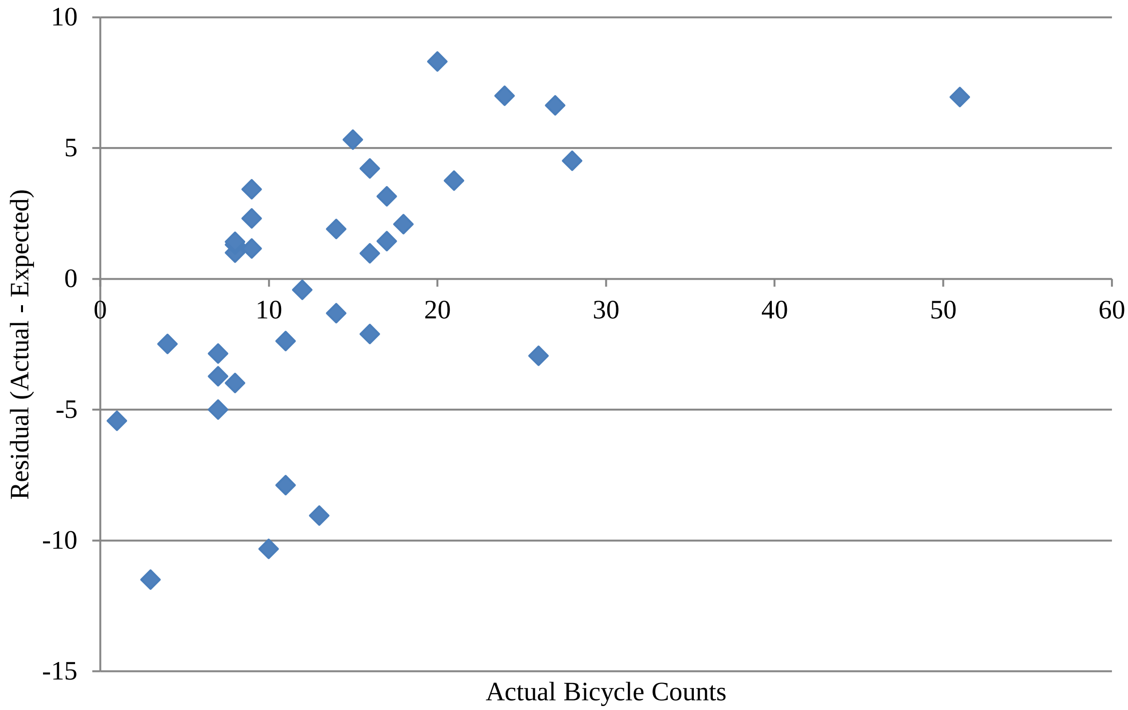

Figure B.3. Residual analysis for bicycle linear regression model.

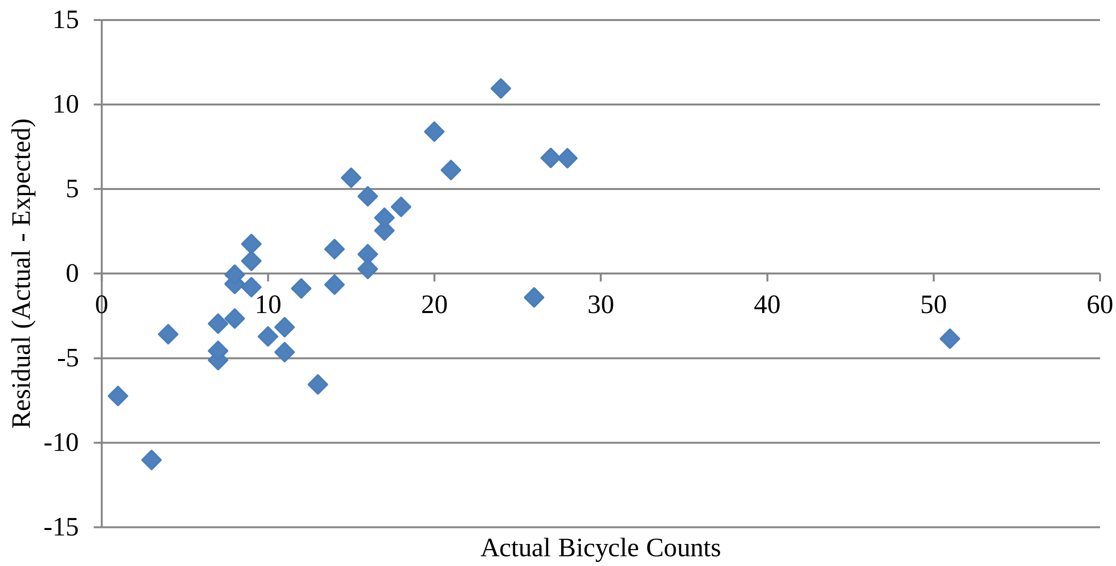

Figure B.4. Residual analysis for bicycle Poisson regression model.

APPENDIX C. RESIDUAL ANALYSIS FOR VALIDATING PREDICTION MODELS

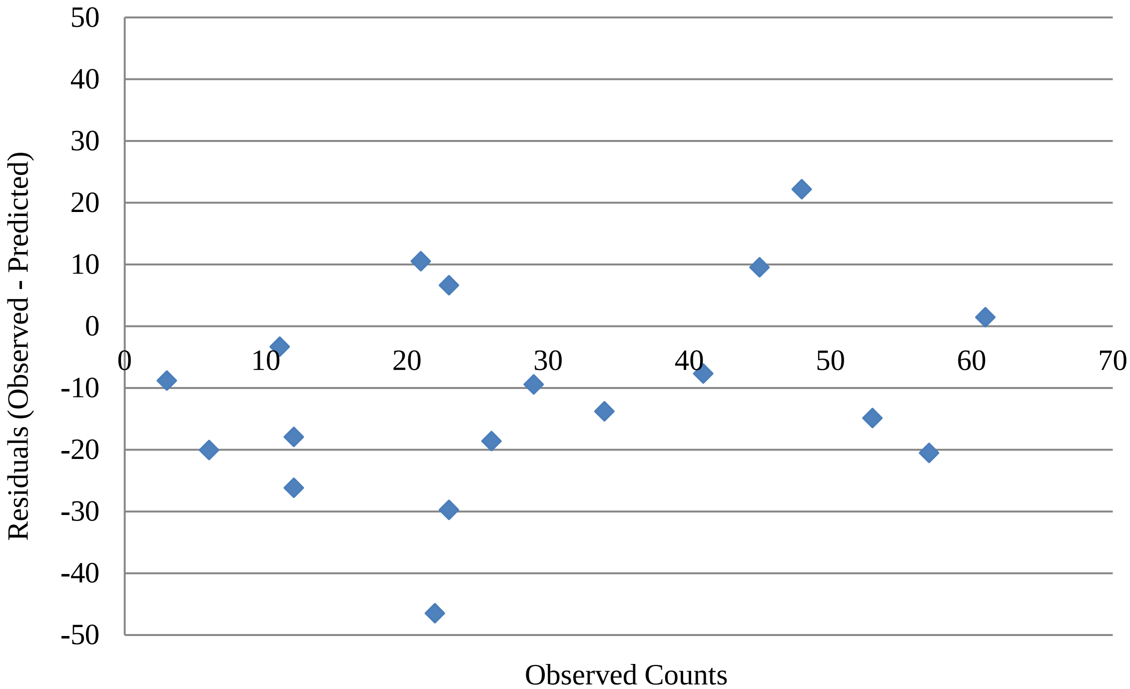

Figure C.1. Residual analysis for pedestrian linear regression model.

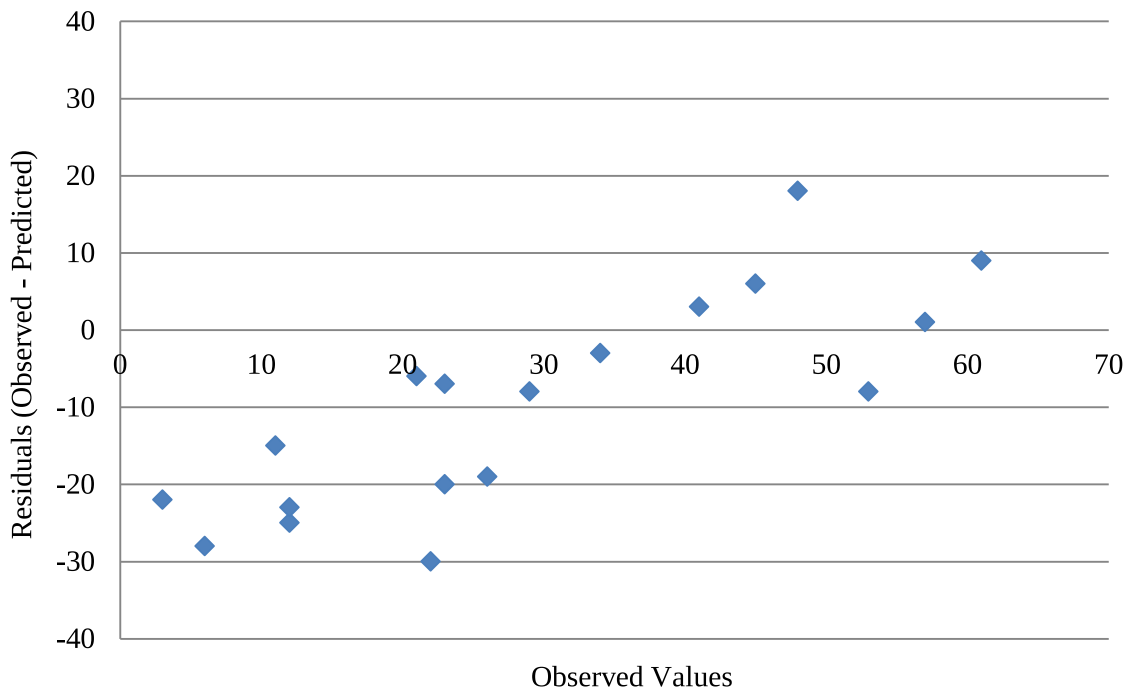

Figure C.2. Residual analysis for pedestrian Poisson regression model.

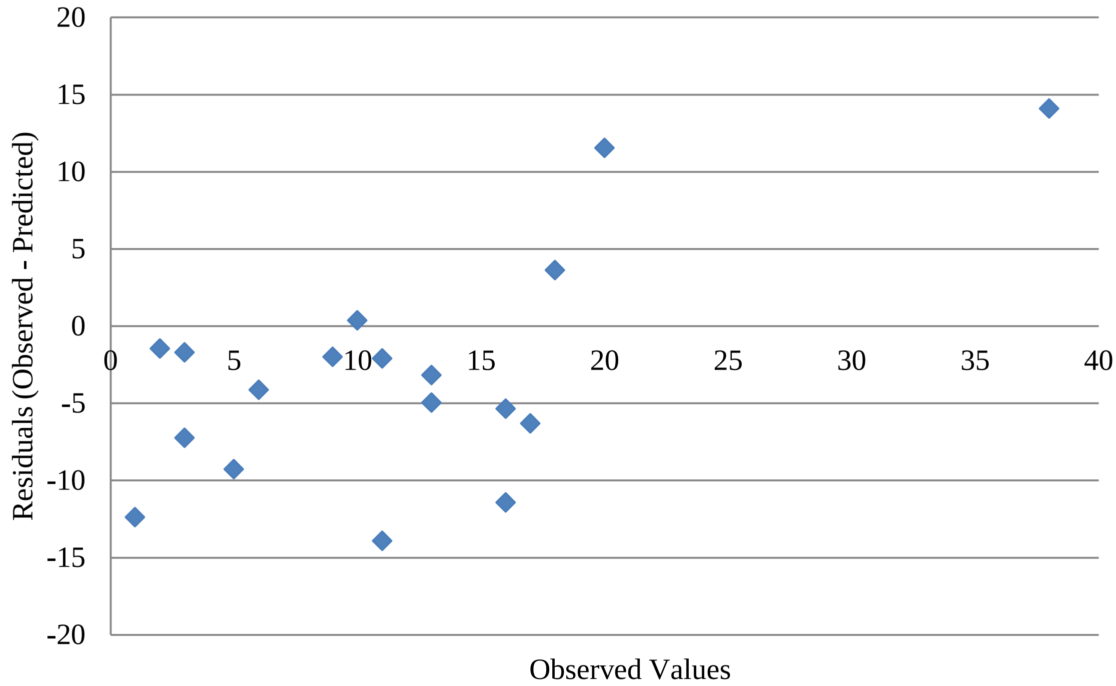

Figure C.3. Residual analysis for bicycle linear regression model.

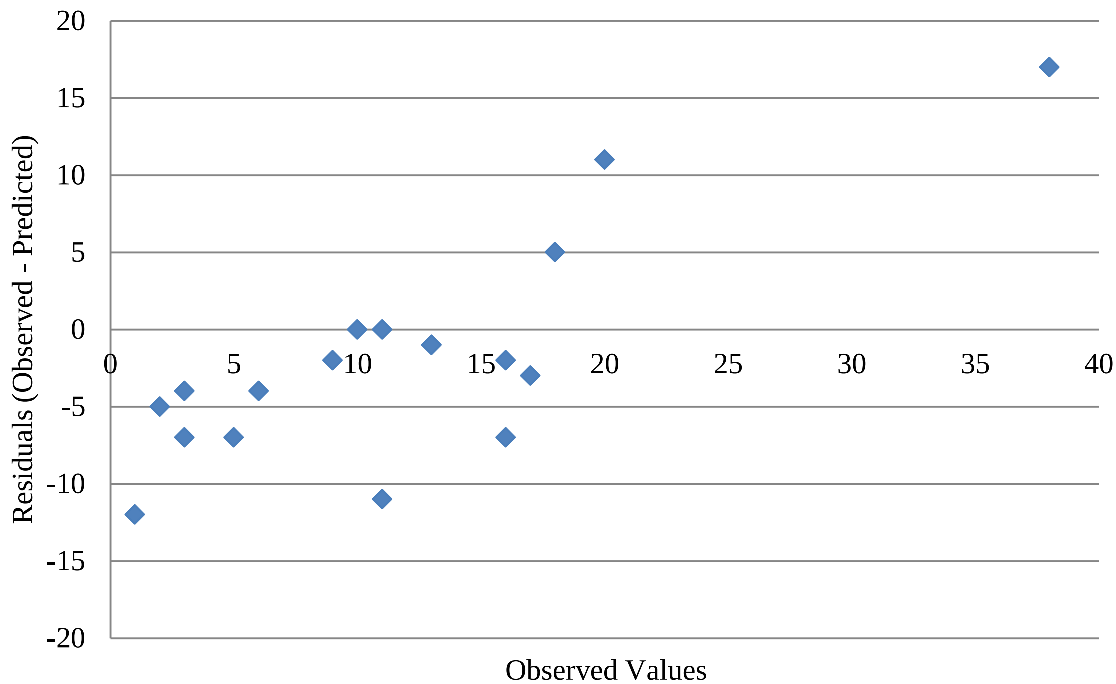

Figure C.4. Residual analysis for bicycle Poisson regression model.