QUANTIFYING THE ROAD SAFETY BENEFITS OF SUSTAINABLE TRANSPORTATION: TRANSIT

by

Ahsan Alam

B.Sc., Bangladesh University of Engineering & Technology, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The College of Graduate Studies
(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Okanagan)

December, 2010

ABSTRACT

As more and more people require transportation, not only must it be efficient and economically affordable, it must also increase safety and be environmentally friendly. High road traffic collisions have been recognized as a global major public health and safety problem. Present auto-dominant culture results in more traffic congestion, collisions, and environmental deterioration. So, there is a need to shift people from auto to a more sustainable transportation system that will promote reduced congestion levels, reduced environmental pollutions, improved road safety, and regional economic growth. This research quantified the road safety benefits of sustainable transportation in the form of increased transit services.

The objective of the research was: (1) to develop an auto-based transportation planning model of the Regional District of Central Okanagan (RDCO) for year 2006 and 2020; (2) to conduct road safety analysis due to future transit and road network improvements; and, (3) to identify collision prone locations of the region.

This research built AM period 4-step transportation planning model of the RDCO for the years 2006 and 2020. It proposed RDCO specific trip generation rate and a new auto mode share model, which can be used for future transportation modeling of the region. This research developed collision prediction models (CPMs) for the RDCO, which can be used to predict future AM period collisions. This research also built RDCO transportation planning model for 2020 having four sub-scenarios: (1) do-nothing; (2) only road improvements; (3) only transit improvements; and, (4) both transit and road improvements. It was found that transit improvements have the potential to significantly reduce urban and rural collisions. This research also suggests that construction of new roads in rural areas might result in collision increases. This is a very important consideration for transportation planners before constructing new roads in rural areas. This research also identified, ranked and analyzed collision prone locations (CPLs) which would help decision makers as they consider where to spend resources, targeting locations with the highest potential for safety improvements. It is believed that the results of this research would contribute significantly in future transportation planning and road safety evaluation of the region.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AADT = Annual Average Daily Traffic
Auto = Private Vehicles
BC = Province of British Columbia, Canada
B/C = Benefit-Cost Ratio
BCMoT = BC Ministry of Transportation
BH = Bus Headway
Black Spot = Hazardous location
BPR = Bureau of Public Roads
BStopD = Bus Stop Density
CMF = Collision Modification Factor
CPL = Collision Prone Locations
CPM = Collision Prediction Model
CRR = Collision Risk Ratio
DA = Dissemination Area
DUE = Deterministic User Equilibrium
EB = Empirical Bayes
GIS = Geographic Information System
GLM = Generalized Linear Model
GVRD = Greater Vancouver Regional District
INTD= intersection density
ITE = Institute of Transportation Engineers
KPR = Kelowna Pacific Railway
LOS = Level of Service
MLM = Multinomial Logit Model
MED = Million Entering Vehicles
MVC = Motor Vehicle Collisions
NB = Negative Binomial
NL = Nested Logit
OCP = Official Community Plan
O-D= Origin-Destination
PCR = Potential Collision Reduction
PDO = Property Damage Only
POPD = Population Density
PCU= Passenger Car Unit
Pop45_64 = Percentage of Population between Ages 45 and 64
RDCO = Regional District of Central Okanagan
RDNO = Regional District of North Okanagan
RSIP = Road Safety Improvement Programs
RTM = Regression-to-the-Mean
SAM3= Severe AM period predicted collisions for 3years
SD = Scaled Deviance
S-D = Socio-Demographic
SOE = System Optimum Equilibrium
SPED = Travel Speed
SRS = Sustainable Road Safety
TAM3= Total AM period predicted collisions for 3years
TAZ = Traffic Analysis Zone
TLKM = Total Road Lane Kilometres
TDM = Transportation Demand Management
VC= Volume to Capacity ratio (congestion)
VHH = Vehicles per Household
VKT = Vehicle Kilometres Travelled

ACKNOWLEDGEMENTS

It is my pleasure to thank the people who made this thesis possible.

First of all, I would like to offer my enduring gratitude to my supervisor, Dr. Gordon Lovegrove, P.Eng., MBA for his enthusiasm, inspiration, and supervision while working in his sustainable road safety (SRS) research group. He provided continuous encouragement, sound advice, and motivation throughout my research period.

I would like to thank the Regional District of Central Okanagan (RDCO), the City of Kelowna, BC Transit, and the Insurance Corporation of British Columbia (ICBC), for providing me their valuable data, without which this thesis would not have been completed.

My sincere thanks go to my family, friends, and colleagues for their spontaneous help, moral support, and friendship.

And above all, my thanks and love to my God for His wisdom and blessings in my life.

DEDICATION

To my parents

CHAPTER 1 INTRODUCTION

1.1 Background

Transport networks provide benefits to nations and individuals by facilitating access to activities such as jobs, economic markets, education, and health. As communities develops and grows, demand for transport increases which adds more vehicular trips to road network. As a result of increased traffic on the road network, more road collisions occur. All over the world high road collision frequencies have been recognized as a major public health and safety problem. The enormous social and economic cost of road collisions has become one of the greatest global challenges.

In Canada, demand for ground based transportation is increasing exponentially. But the supply of new facilities is not enough to meet the growing demand. Auto (i.e. private vehicles) is the dominant mode of transportation in Canada and this auto-dominant culture is resulting in more traffic congestion, environmental pollution, and road collisions. A safer, more efficient and more environmentally friendly transportation system is needed. This can be achieved by adopting different policies and infrastructure systems such as educating road users, enforcing strict laws, controlling land use and applying engineering strategies. All of these strategies support the movement towards a more sustainable transportation system.

Road transportation system has three components namely the driver, the vehicle and the road environment. Collision occurs from any of these components or a combination of them (Sayed et al., 1995). In North America the drivers' error is involved in 96% of road collisions, whereas the road environment and vehicle related components are involved in 30% and 10% of collisions respectively (Sayed et al., 1995). As driver error is the main cause behind road collisions, reduction of auto use may improve road safety. By developing Collision Prediction Models (CPMs), theoretically it has been found that reduced auto volume of the road network can result in reduced road collisions. So by shifting more people from auto to other mode such as transit, it may be possible to reduce road collisions. As transit is considered as a more sustainable transportation mode than auto, it is very important to find how increased transit use will reduce road collisions. This thesis aims to quantify the sustainable road safety benefits of increased transit use. The study area of the thesis is Regional District of Central Okanagan (RDCO), British Columbia, Canada.

In this introductory chapter, while section 1.1 highlights the importance of the research, section 1.2 discusses the motivation of the research by summarizing the global road safety context and the need for sustainable transportation. Section 1.3 introduces the geography, transportation network and road safety of the study area. Section 1.4 states the research objectives and Section 1.5 describes the structure of this thesis.

1.2 Motivation

1.2.1 Motivation-1: Road safety problem

Even though transportation has benefits, it also has some costs such as traffic congestion, environmental pollution, and road collisions. Road collisions have been shown to take lives, reduce productivity and hamper economic growth. So to improve road safety scenario, it is very important to understand the road safety problem. The following sub-sections will discuss the road safety problem in a global as well as local context.

1.2.1.1 Global picture of road safety

The problem of road collisions is being experienced by every country of the world. Each year over 1.3 million people die and 20-50 million people suffer from non-fatal injuries due to road collisions (WHO, 2009). To recognize the social burden resulting from road collisions, the World Health Organization (WHO) dedicated 'World Health Day 2004' to road safety to raise global awareness. The United Nations (UN) carried out 'Global Road Safety Week' in 2007 to improve road safety awareness of young people (UN, 2007). According to WHO (2009) report, throughout the world, road collision injuries are one of the top ten leading "causes of death". If new initiative are not taken then road collision will increase by 65% (Gasper, 2004) and the number of road collision fatalities will be doubled by 2020 (CGRF, 2009). Another report by WHO (2008) predicts that road collision injuries will become the 5^{th} leading causes of death by 2030.

The following Table 1.1 (WHO, 2004) predicts the current and predicted road safety problems in different regions of the world. The regions were described by the regional classifications of the World Bank and classified based on Gross National Income (GNI) of each country. According to the World Bank (2008), if GNI is less than $975 per capita then it is low-income country, if it is between \$976 and \$3,855 per capita then it is middle-income country and if it is greater than \$11,906 per capita then it is high-income country. The main assumptions of the prediction were: (a) present policies and actions in road safety will continue; and, (b) no additional road safety countermeasures will be put into place.

Table 1.1 Predicted road traffic fatalities by region (in thousands) (WHO, 2004)

Region	Number of Countries	Observed		Predicted		Change (%) between 2000-2020	Fatality Rate (deaths/ 100,000 persons)	
		1990	2000	2010	2020		2000	2020
East Asia and Pacific	15	112	188	278	337	79	10.9	16.8
East Europe and Central Asia	9	30	32	36	38	19	19	21.2
Latin America and Caribbean	31	90	122	154	180	48	26.1	31
Middle East and North Africa	13	41	56	73	94	68	19.2	22.3
South Asia	7	87	135	212	330	144	10.2	18.9
Sub-Saharan Africa	46	59	80	109	144	80	12.3	14.9
Sub-total	121	419	613	862	1124	83	13.3	19
High-Income Countries*	35	123	110	95	80	-27	11.8	7.8
Total	156	542	723	957	1204	67	13	17.4

* North America, Europe

The table shows that between 2000 and 2020, high-income countries will experience an average of 27% decrease in road collisions whereas low-income and middle-income countries will experience an average of 83% increase in road collisions. Even though for high income countries the decrease is 27%, still there will be 80,000 fatalities in 2020 which is a very high number. Among different regions South Asia will experience the largest growth in road traffic deaths, with a dramatic increase of 144%. A recent report of WHO (2009) indicates that more than 90% of road collision fatalities occur in low-income and middle-income countries, even though these countries have only 48% of the world's registered vehicles.

Figure 1.1 (WHO, 2004) shows that road collision injuries are responsible for 23% of all types of injury deaths. The number of premature deaths resulting from road collisions is higher than that of other causes. From a report of United Nations (ITC, 1999) it is found that around 70% of road collisions victim's age are within "working years" (e.g. age between 25 and 65). In 2004 road traffic injuries were within the top three causes of death of people aged between 5 and 44 years, and, by 2015 it will become the leading cause of premature deaths and disability for children aged 5 and above.

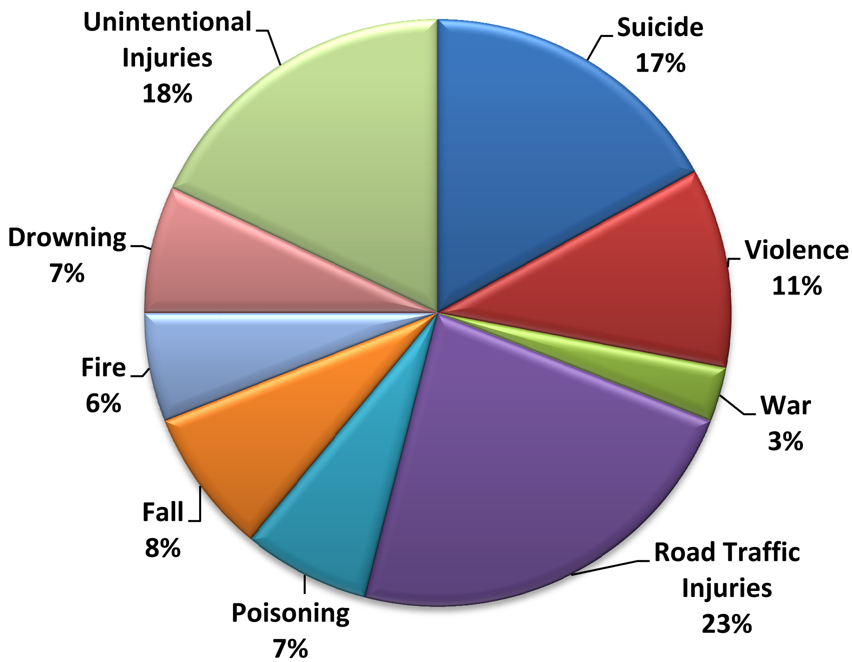


Figure 1.1 Distributions of global injury mortalities by cause (WHO, 2004)

The cost of road collision injuries and deaths is enormous. Globally, the economic cost of road injuries varies from 1% to 3% of Gross National Product (GNP), which equals US$ 518 billion annually. Especially in developing countries the cost is US$100 billion annually which is twice of the development assistance given to those developing countries (UN, 2003). If the current trends continue, by 2020 health departments of developing countries will be spending approximately 25% of their budgets on traffic collision caused casualties (ITC, 1999). So road safety is not only a global major public health issue but also an economic burden to society.

With a goal to improve global road safety, the United Nations (UN) has announced the period of 2011-2020 as the 'Decade of Action for Road Safety', The overall goal of the decade will be to halt or reverse the increasing global road fatalities trend by increasing activities at the national level. During the ‘Decade of Action’ period, up to 5 million lives is expected to be saved and 50 million injuries to be prevented which will be a reduction of 50% of predicted 1.9 million deaths globally by 2020 (UN, 2009). In 2003 the European Commission adopted ‘3rd European Action Program to Road Safety’ with a target to reduce the number of fatalities by 50% between 2001 and 2010 by reducing the number of fatalities from 54,000 to 27,000. Even though the target number of fatalities reduction by 2010 has not been completely met, significant progress has been observed such as the number of fatalities has fallen by more than 40% compared to 25% decrease between 1990-2000 (Europa, 2010). The European Commission also adopted plans for the next decade from 2010 to 2020, which targets reduction of 50% road fatalities on Europe's roads by setting higher vehicle safety standards, improving road user’s behaviour, and increasing the law enforcement (EC, 2010). Like other concerned countries, Canada has also undertaken road safety vision plans to reduce the number of road collision victims. The following sub-section will discuss about Canadian road safety.

1.2.1.2 Context of Canadian road safety

Road collision is a major problem in Canada. The following Figure 1.2 gives the Canadian road collision history which clearly shows that total number of fatalities is decreasing over time.

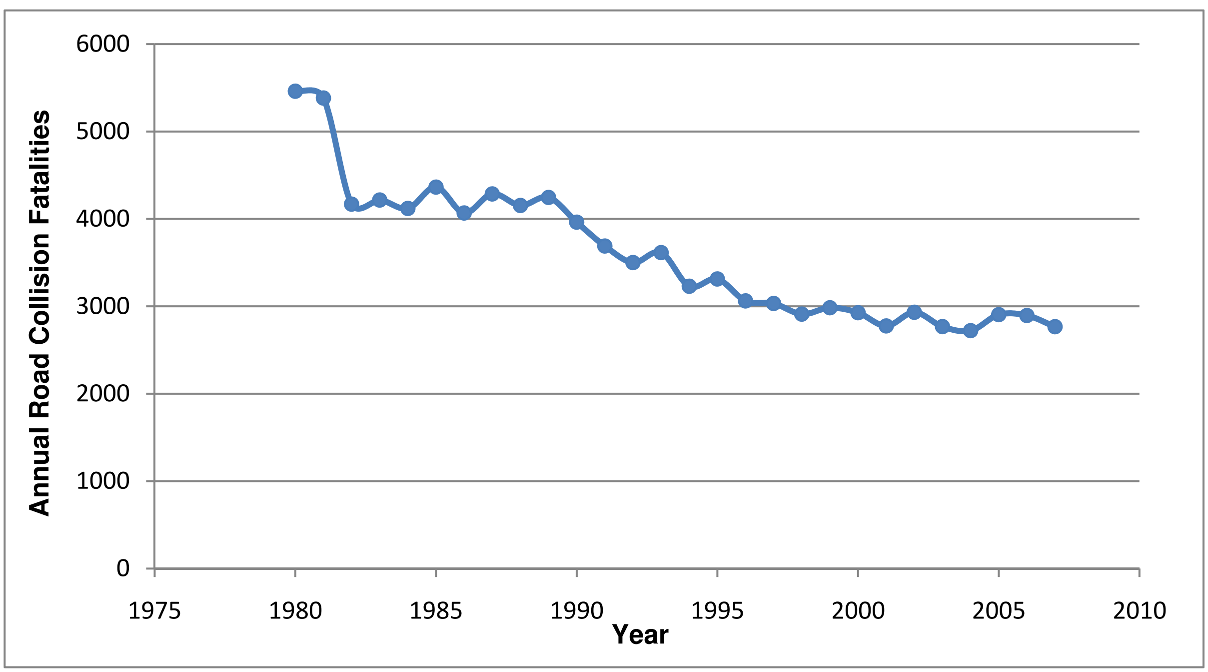


Figure 1.2 Canadian road collisions trend

In Canada each year, road collisions account for a large percentage of all accidental deaths (Transport Canada, 2004). From 2000 through 2004, 32% of injury deaths were the result of Motor Vehicle Collisions (MVCs) (Ramage-Morin, 2008). In 2001, there were 2,778 deaths and 24,403 injuries due to road collisions (Transport Canada, 2004). In 2006, it increased to 2,889 deaths and 199,337 injuries, or 9.1 deaths and 630 hospitalizations per 100,000 populations (Transport Canada, 2007).

Road collisions in Canada are also one of the leading causes of lost years of productive life (working years of a person's life). In the age group of 15 to 24 years, 70% of all accidental deaths were the result of road collision during 2000-2004 (Ramage-Morin, 2008). The Canadian Council of Motor Transport Administrators (CCMTA, 1998) reports that the average years of lost life due to a fatal collision is 40 years, compared to respiratory disease (9 years), circulatory disease (10 years), or tumors (15 years). In 2004 the direct economic cost (ignoring the indirect costs such as lost productivity, long term disability and pain sufferings) of road collisions to Canadians was estimated $25 billion annually which was about 2% of Canada's 2004 GDP, (Transport Canada 2004). Considering both direct and indirect costs the total collision cost in 2004 was $63 billion annually (Vodden et al., 2007) which was 4.88% of 2004 Canada's GDP.

With a vision of making Canada's roads the safest roads in the world, in 1996 Canadian Council of Motor Transport Administrators (CCMTA) initiated 'Road Safety Vision: 2001'. The main objectives of the vision were: (1) to raise public awareness; (2) to improve communication and collaboration between road safety agencies; (3) to enhance enforcement measures; and, (4) to improve national data quality and collection (CCMTA, 2007). 'Road Safety Vision 2001' was successful in terms of fatal and injury collision reduction as well as improvement of Canada's position in international road safety ranking (CCMTA, 2005). This success prompted the CCMTA to adopt 'Road Safety Vision: 2010' which would emphasize the importance of partnerships among agencies and the use of a wide variety of initiatives to reduce road collisions. The overall target of that vision was to reduce the number of road collision severities (e.g., fatalities or serious injuries) by 30 percent during the period of 2008 to 2010, compared to the period of 1996 to 2001 (CCMTA, 2005). The following Figure 1.3 has been prepared using CCMTA annual report (2005) and Transport Canada (2010) collision statistics to give a snapshot of Canada's road safety situation. This figure indicates the fatality trend line as well as the desired fatalities line for Vision 2010.

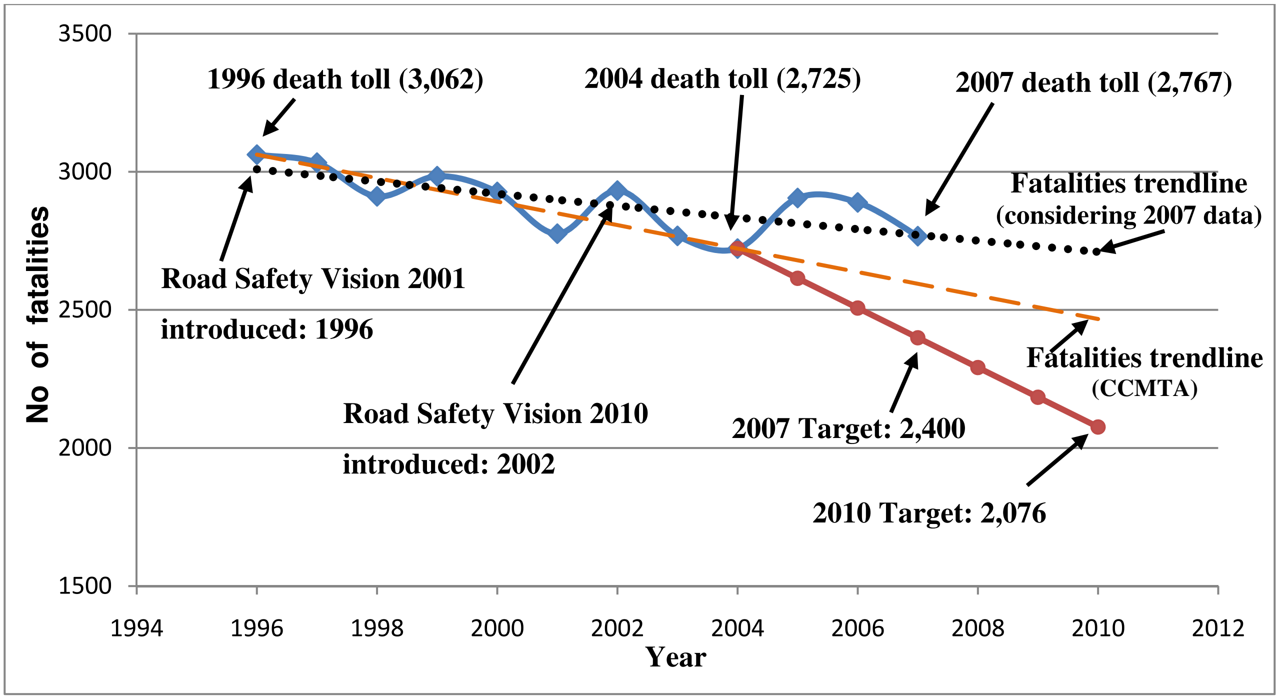


Figure 1.3 Canadian traffic fatality projections to 2010

It is important to mention here that the fatality trendline by CCMTA was projected in 2005 considering collisions data till 2004. But after observing 2005, 2006 and 2007 collisions data it can found that the actual fatality trendline is even higher than that of CCMTA. According to the CCMTA fatality trend line, in 2007 total number of expected road collision fatalities were 2,613, whereas according to the Vision: 2010 expected number of fatalities was 2,400. So, there is a gap of 213 fatalities. But if the actual trendline (considering 2007 collisions data) is considered, then it can be seen that the gap is even bigger. It clearly signifies the need of improving Canadian road safety condition.

1.2.2 Motivation-2: Sustainability

As communities develop, activities to meet their daily needs are increasing and making their ecological footprint on the earth. Ecological footprint is a measure of human demand on the Earth's ecosystems (Bonard et al., 2008). It is defined as ''total area of productive land and water required continuously to produce all the resources consumed and to assimilate all the wastes produced by a defined population, regardless of where that land is located" (Rees & Wackernagel, 1996). It aims to determine ''to what extent human load is within the present regenerative capacity of the biosphere or natural capital interest" (Haberl et al., 2001).

Human activities are responsible for creating an unbalance of the world. Concerning about present and future world's sustainability, in 1987 the United Nations published the report, "Our Common Future", which focused on the change of policies necessary for achieving sustainable development. This report defines sustainable development as "sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs." (UN, 1987). This report was the first report that discussed about different strategies and policies to achieve sustainable development. Hawken (2010) estimates that at present we are using 30% more of nature than what nature can regenerate. In other words, our present rates of natural resource extraction, consumption and waste require 1.3 times the carrying capacity of the world. A projection from the United Nations (UN) shows that by 2050, we will be using about twice the carrying capacity of the Earth (Hawken, 2010). Also 60% of the earth's protective ozone layer has been lost in last 50 years, and 30% of the earth's arable land has been lost in last 40 years (Hawken, 2010). Natural disasters, extreme climate changes, more droughts, more flooding, more variation in weather, wildfires, strong storms, spreading of tropical diseases etc. are the evidences of an unsustainable global environment. Global warming has become one of the greatest global threats as well. There are two debatable causes behind global warming; (1) natural cycle (Avery and Singer, 2007); and (2) green house gases. Different studies found that around 30% of green house gases are the contribution of transportation (OECD, 2001; USEPA, 2006).

Energy crises have also become a major global problem. In 2007, 82 percent of the world's total energy supply came from non-renewable sources: oil (34%), coal (27%) and natural gas (21%) (IEA, 2009a). If the current production rate remains constant, then the estimated oil reserve is enough to last another 40 years, natural gas 60 years and coal 155 years (Iwaro, 2010; IEA 2009b). From 2007 to 2035 worldwide the use of oil will increase 30%, while in the same period the use of oil in transportation sector will increase by 45% (USEIA, 2010). To meet the future demand huge volume of oil reserve is needed. But researchers have shown that oil production rate has reached or will reach to its peak between 2010 and 2020 (Edwards, 1997; Duncan &

Youngquist, 1999; Cavallo, 2002; Hirsch, 2005). So, it is very clear that in near future conventional oil production rate will enter a phase of permanent decline (Krumdieck et al., 2010). As oil production rates will decline and demand for oil will increase, the price of oil will increase (Pimenten, 2010). Thus, higher price of carbon-based fossil fuel will make private auto-based transportation less economically feasible. More sustainable alternative transport modes will become more attractive.

So there is a need to shift from private vehicles to mass transit and also from carbon-based fossil fuel to alternate fuel source. In 2007 the European Union (EU) targeted to increase the share of renewable energy and the target was "by 2020 renewable energy should account for 20% of the EU's final energy consumption". United States is also planning for renewable energy so that it can supply 10% of the nation's electricity by 2012 and 25% by 2020 (EERE, 2009). In Canada, even though some provinces have adopted renewable energy policies, but nationwide no initiatives have been taken yet.

1.2.2.1 Sustainable road safety

Clearly there is an urgent need for road collision reduction in a sustainable manner. Sustainable Road Safety (SRS) is a proactive approach to create a transportation system in which the road environment keeps pace with transportation demands and the probability of collision occurrence is limited to an acceptable level of risk by means of an inherently safe road environment (Van Schagen & Janssen, 2010; Croft 2005). SRS approach starts with human behaviour as driver error consist 96% of road collision (Sayed et al., 2005). Due to unpredictable human behaviour, road users are considered as the weakest link in the transport chain (Van Viliet & Schermers, 2000). But two other components behind road collisions (e.g. road environment and vehicles) are also important. That is why; SRS acknowledges that the safe transportation network cannot be achieved only relying on road users' behaviour, but also land use and transport infrastructure design, vehicle performance, and strict law enforcement. The SRS concept is based mainly on three main principles: (1) the functionality of roads, (2) the predictability of traffic situations, and, (3) the homogeneity of masses and speed (Van Schagen & Janssen, 2010, Engelsman & van Zyl, 2007). There are another two principles: (4) forgiving road environment and anticipation of road users' behaviour, and, (5) state awareness by the road user (SWOV 2006). Considering all of these principles, human behaviour errors and consequences of these errors are minimized so that a safer transportation system can be achieved.

Previously discussed sustainability problems (i.e. energy crises, environmental pollution) and road collision burden clearly promote the need of an auto-alternate transportation system which will be sustainable in nature. Sustainable transportation is such a system which focuses on a safe, efficient, economic, and environmentally friendly transportation system (Wiederkehr, 2004). It is concerned about finding a proper balance between environment, economic and society (Steg & Gifford, 2005; Ruckelhaus, 1989; Litman, 2009a). The European Union's Council of Ministers (2001) adopted the following definition of sustainable transportation:

- A sustainable transport system allows for the basic access and development needs of individuals, companies and societies in a manner consistent with human and ecosystem health. It also promotes equity within and between successive generations.

- It is affordable and efficient, offers a choice of different transport modes, supports a competitive economy, and facilitates balanced regional development.

- It limits emissions and waste to within the planet's ability to absorb them, consumes renewable resources at or below their rates of re-generation, and uses non-renewable resources at or below the rates of development of renewable substitutes while minimizing the impact on the use of land and the generation of noise.

In other words, sustainable transportation promotes reduced congestion levels, cleaner environments, improved road safety, and regional economic growth in such a way that it will have less harmful effect on the future of the earth. Basically, the concept of sustainable transportation is embedded in the broader concept of sustainable development and SRS. Sustainable transportation includes different modes such as walking, cycling, telecommuting, green vehicles, car-sharing and public transit (which can include heavy rail, trolley, express bus, conventional fixed-route bus, minibus, demand-responsive para-transit) (Litman and Burwell 2006). Among various modes, this thesis will focus on transit (e.g. bus) and transit oriented development to quantify the direct economic benefits of sustainable transportation in terms of road safety, by focusing on a study of the following geographical area.

1.3 Study area

1.3.1 Demographics and land use of RDCO

The study area of this thesis is the Regional District of Central Okanagan (RDCO) is situated along the shores of Okanagan Lake in the southern interior of British Columbia, Canada, as shown in the following Figure 1.4. The RDCO is comprised of the City of Kelowna, District of West Kelowna, District of Lake Country and District of Peachland.

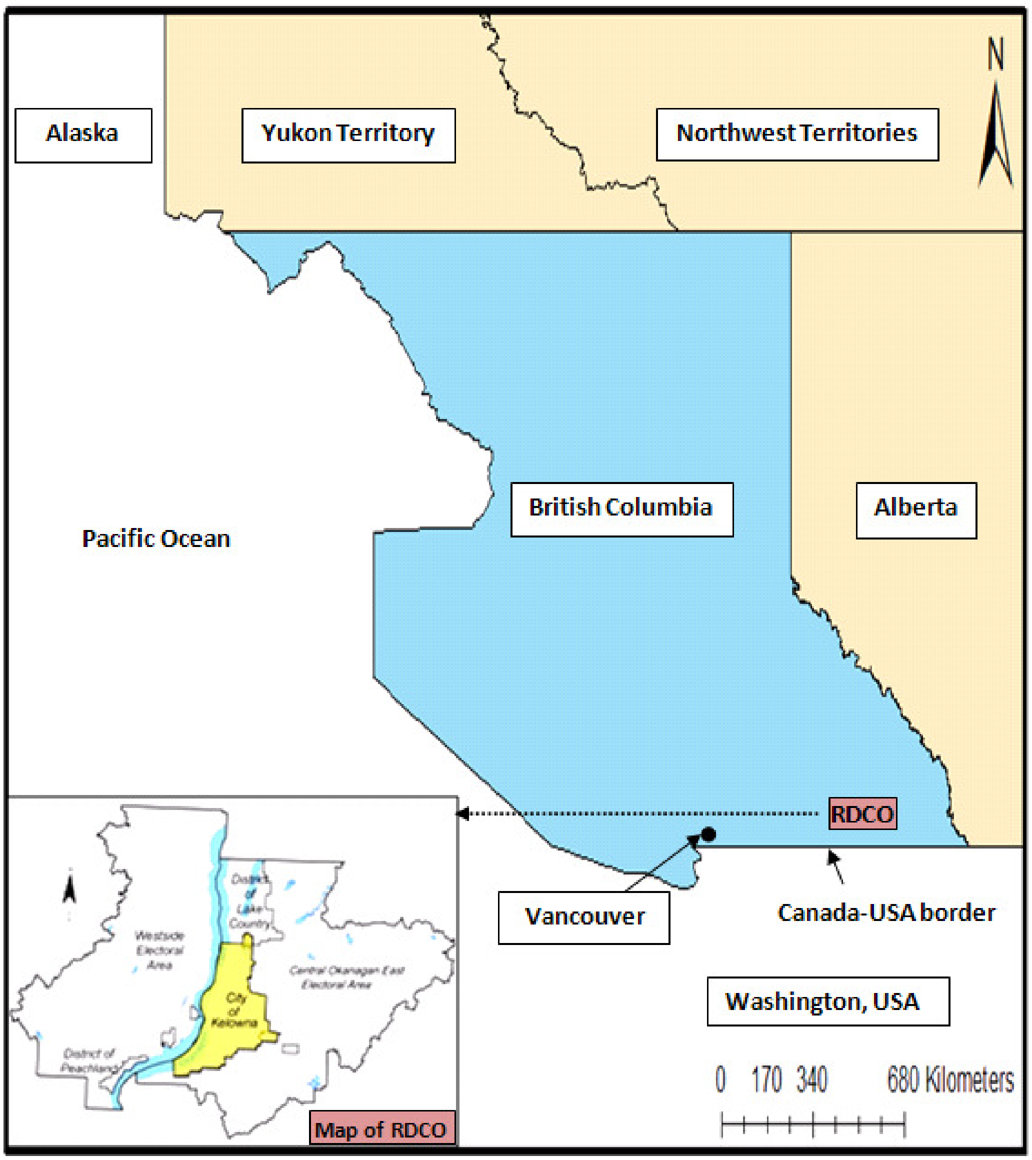


Figure 1.4 Geographic location of the Regional District of Central Okanagan

RDCO is one of the fastest growing regions in British Columbia, Canada. The region is expected to grow at approximately 2.31% per annum from 2006 population of 167,417 to some 221,589 persons in year 2020, a growth of 54,172 people in 14 years (BC STAT, 2010). Its economy is rooted in agriculture, forestry, manufacturing and tourism. Table 1.2 (BCSTAT, 2010; COEDC, 2009a) compares the historic and predicted population growth of those municipalities.

Table 1.2 Population statistics of the RDCO

Name	1996 census pop	2006 census pop	Annual change between 1996-2006 (%)	2006 pop. density per sq. km	Forecasted 2020 pop. (from OCP)	Annual change between 2006-2020 (%)
City of Kelowna ($211.69\ km^2$)	92,859	110,351	1.88	521.29	141,689	2.03
District of West Kelowna ($121.42\ km^2$)	n/a	28,793	n/a	237.14	34,447	1.40
District of Lake Country ($122.16\ km^2$)	9,330	9,790	0.49	80.14	15,243	3.98
District of Peachland ($15.98\ km^2$)	4,675	4,938	0.56	309.01	7,127	3.17
RDCO ($2,904.01\ km^2$)	141,628	167,417	1.82	57.65	221,589	2.31
City of Vancouver ($114.78\ km^2$)	536,516	599,780	1.18	5225.5	663,340	0.77
Greater Vancouver Regional District ($2877.36\ km^2$)	1,906,506	2,199,121	1.53	764.28	2,757,615	1.81
British Columbia ($924,815.43\ km^2$)	3,724,500	4,243,580	1.39	4.59	5,148,497	1.52

City of Kelowna

The city of Kelowna was incorporated in 1905. The 2006 population of Kelowna was 110,351 with a 1.88% annual growth over the 1996 population of 92,859 (BCSTAT, 2010). The projected 2020 population of the city is 141,689 (City of Kelowna, 2009), which indicates that from 2006-2020 the annual population growth rate of 2.2% will be higher than that of 1996-2006 period. It is the largest city in BC’s Okanagan valley and one of the fastest growing cities in BC. The city is famous for its natural beauty and it is a centre of global tourism. The city is the regional economic centre of the RDCO. Between 1995 and 2005 average household income increased almost 41% with an average income of $ 63, 932 in 2005, whereas between 1996 and 2006 total number of employed person increased 32.2% with 54,525 employed persons in 2006 (COEDC, 2009b). The city's vibrant retail and commercial base makes it the 2nd largest centre for business, commercial, retail, urban development, and health care services in BC next to Vancouver.

District of West Kelowna

The District of West Kelowna is incorporated in December 2007. It is the second largest municipality in the Central Okanagan region with a 2006 population of 28, 793. The forecasted 2020 population is 34,447 with 1.4% annual growth from 2006-2020. The district has a diverse economy, including agriculture, manufacturing, tourism, retail and construction. In 2005 average household income was $ 60, 639 and between 1996 and 2006 total number of employed persons increased 36.6% with 14,400 employed persons in 2006 (COEDC, 2009c). The largest segment of this labour force is in retail trade with 13.2% while 12.5% are involved in construction industry which is more than that of provincial 7.6% (District of West Kelowna, 2010).

District of Lake Country

District of Lake Country was incorporated in 1995. It is the northern-most community of the RDCO. At present the district is experiencing rapid population increase. Between 1996 and 2006 the observed annual population growth rate was 0.493% with 2006 population of 9,790. The forecasted 2020 population is 7,127 having a 3.98% annual growth from 2006 to 2020. Rapid population growth in this area has resulted in significantly increased residential and commercial activities. The main three industries that are employing most of the people in this area are manufacturing, retail trade, and construction. Between 1995 and 2005 average household income increased almost 47.3% with an average income of $ 68,882 in 2005; whereas between 1996 and 2006 total number of employed persons increased 22.95% with 5,250 employed persons in 2006 (COEDC, 2009d).

District of Peachland

This district was incorporated in 1909. Peachland is the southern-most community of the RDCO. The 2006 population of Peachland was 4,938 with a 0.56 % annual growth over the 1996 population of 4,675 (BCSTAT, 2010). The projected annual population growth from 2006 to 2020 is 3.17% which is much higher than the observed rate from 1996-2006. Between 1995 and 2005 average household income increased almost 17.9% with an average income of $55,143 in 2005; whereas between 1996 and 2006 total number of employed persons almost doubled with 2,155 employed person in 2006 (COEDC, 2009e). Over the past few years Peachland has undergone a major development program, including improvements to transportation facilities, municipal services and additional residential, commercial and retail facilities.

1.3.2 Transportation network in the region

The transportation network of this region consists of lakes and highways, railway corridors, and, an international airport as shown in Figure 1.5. Following sub-sections will describe each of the networks in the region:

Road Transportation

Highway 97 is Okanagan Valley's key north-south highway with connections to several east-west highways, including the Trans Canada Highway 1 and Highway 3, 3A, 33, 6 and 97C (Okanagan Connector). Highway 97 goes through each of the municipalities of RDCO. In the study area most of the case Highway 97 has four lanes except central part of the Kelowna where it has six lanes. In a hope to encourage more transit use and car pooling, from September, 2009 High Occupancy Vehicle (HOV) lane has been introduced in Highway 97 between Water Street, Kelowna to Highway 33, Kelowna. Highway 97 is the busiest corridor of the region. In 2009 Annual Average Daily Traffic (AADT) on this corridor was 27, 320 (BCMoT, 2010).

Auto is the predominant mode of transportation in the region with a 90% mode share for all kind of trip purposes. Transit, bicycle, walking mode share are 1.5%, 2.1% and 4.9% respectively (Winram, 2007). Greyhound Lines of Canada Ltd provides scheduled bus service to all Central Okanagan communities and to rest of Canada/US. There are some other services which operate within Okanagan valley. Public transit is not a popular mode in this region. For daily purposes, average travel time by transit is 31 minutes while by auto it is 15 minutes (Winram, 2007).

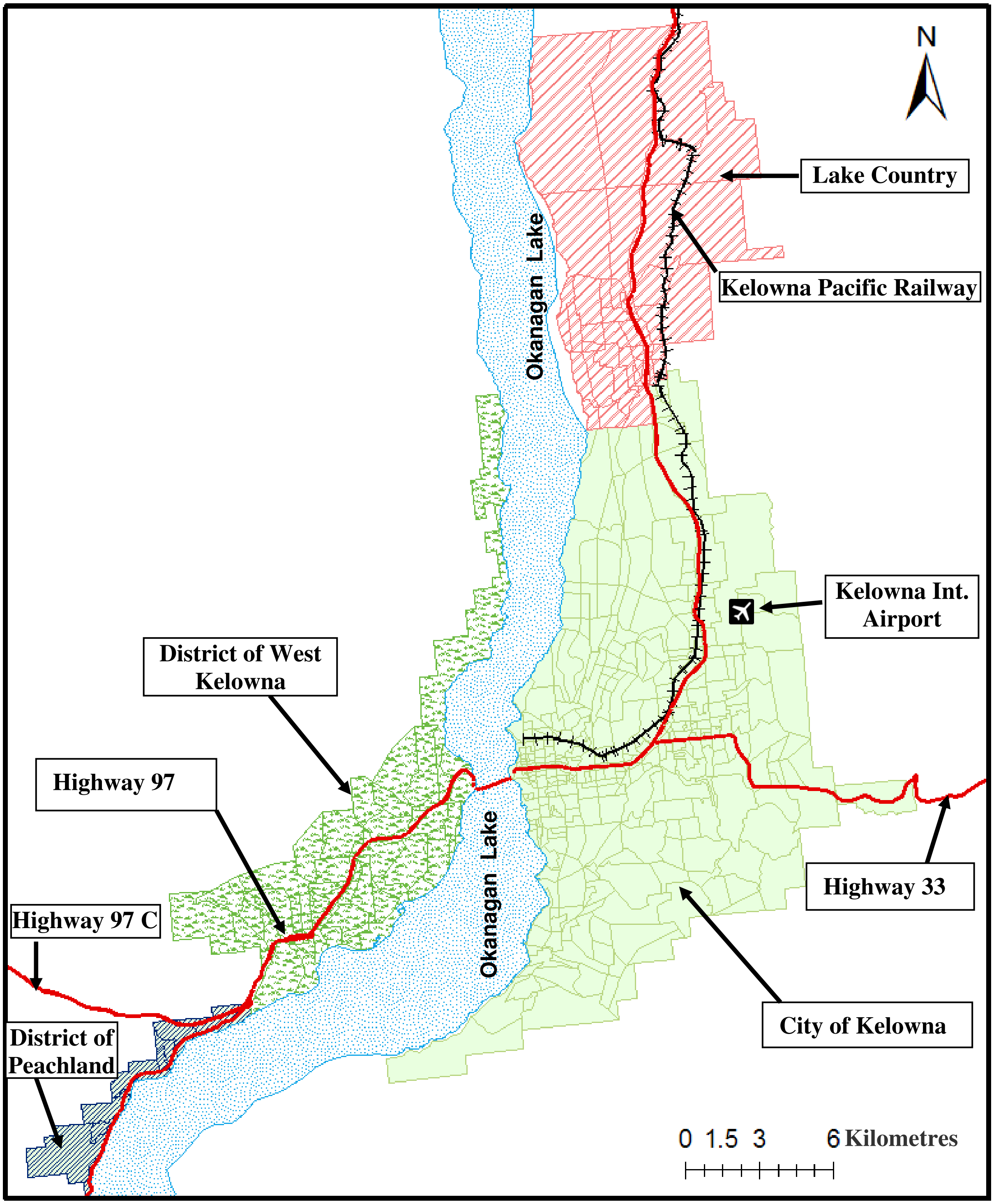


Figure 1.5 Transportation network of the Regional District of Central Okanagan

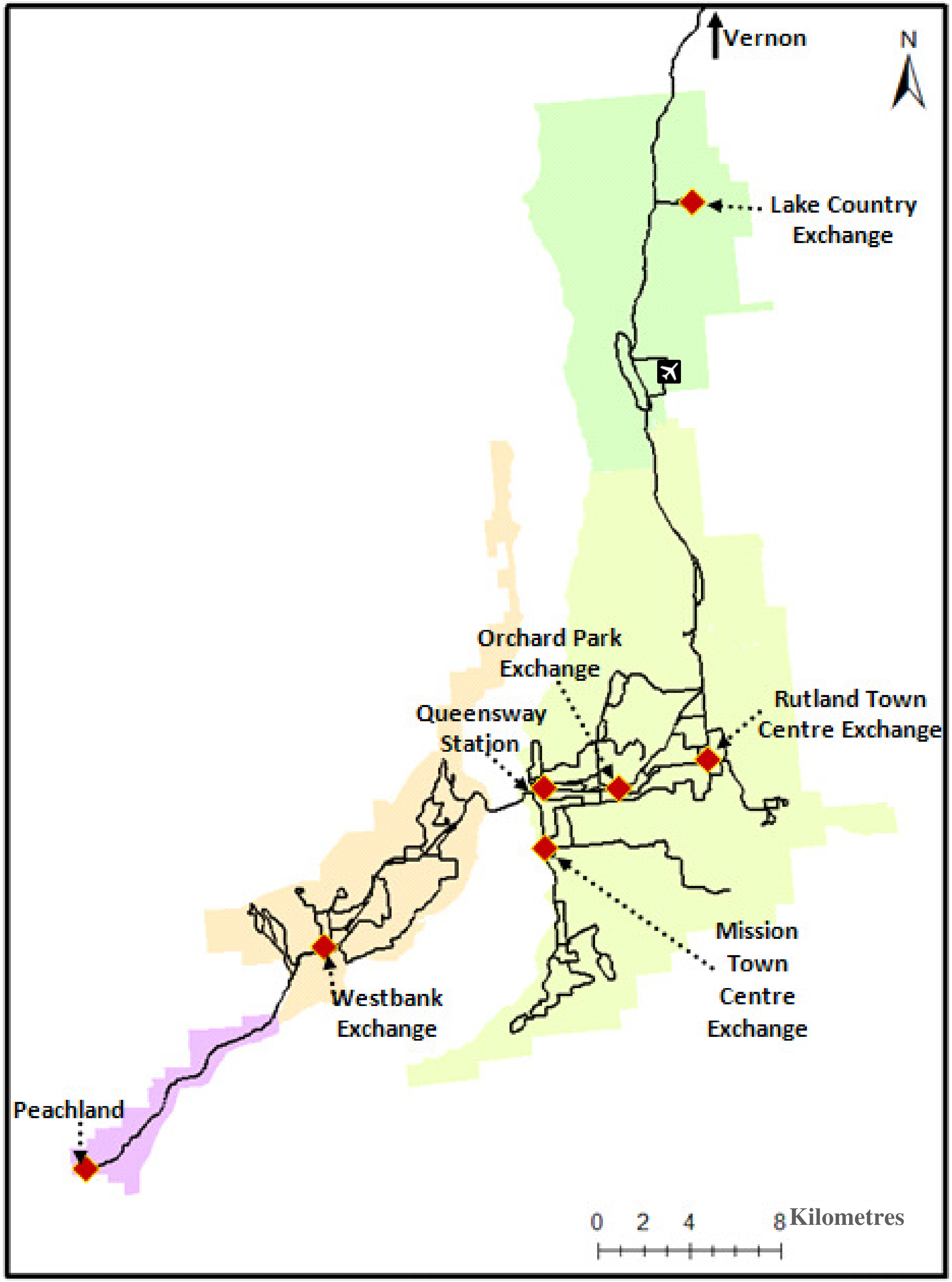


Figure 1.6 Transit route of the Kelowna Regional Transit Service

The RDCO transit service is operated by Kelowna Regional Transit System which started in 1977 and the first significant improvement of the system came in 1996 with the introduction of 15-minute peak hour commuter service buses on major routes. In 2005-06 total annual RDCO transit ridership reached over 3,000,000 trips. In 2007 average number of passengers per weekday was nearly 13,000 (BC Transit, 2007). A modest growth of 4 – 6% per annum was projected over 2020 (IBI Group, 2005). Figure 1.6 presents the transit route and important transit exchange within the region. At present 22 bus routes are operated in the region where 13 buses within Kelowna, 7 buses within West Kelowna, 1 bus within Lake Country and 1 bus within Peachland. In addition, Vernon Regional Transit System provides one bus service in northern RDCO which moves between Vernon and UBC-Okanagan.

Rail Transportation

Kelowna Pacific Railway (KPR) is the only railway in the study area and operates on a short haul rail line leased from Canadian National (CN) Rail. It consists of 167.7 km of mainline track, 10.2 km of associated sidings and spurs, and approximately 40.9 km of running rights on Canadian Pacific Railway (KPR 2010). It runs between Canadian National (CN) Railway's Kamloops yard and Kelowna, and, between Vernon and Lumby. At present this rail service is used only for freight (i.e. forest products, grain, industrial products etc.) transport. The freight service can handle approximately 16,000 carloads per year. Following Figure 1.7 shows the rail route of KPR.

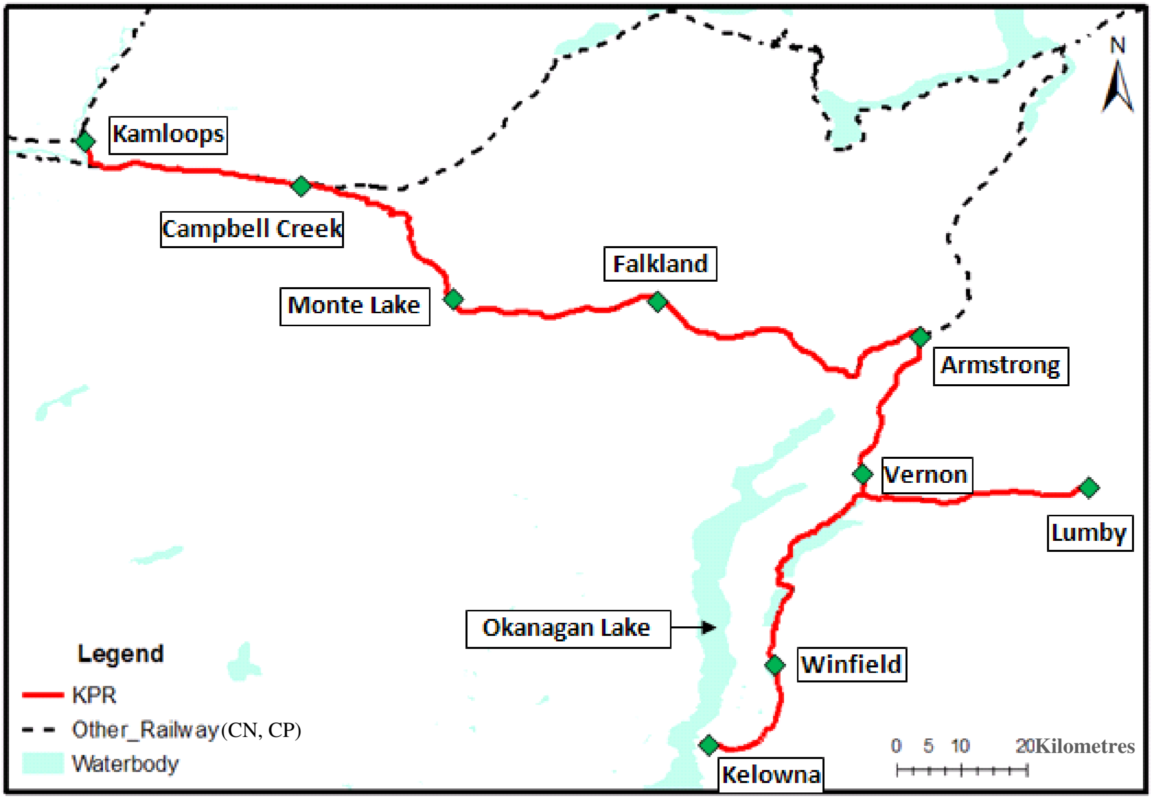


Figure 1.7 Kelowna Pacific Railway route

Air Transportation

Kelowna International Airport acts as a gateway into B.C.'s southern interior and facilitates the economic prosperity of the Okanagan Valley, as the 11st busiest airport in Canada. In 2009, the total number of trips was 1,367,631. Between 1996 and 2006, the number of annual trips increased from 0.16 million to 1.22 million indicating an annual growth of 10.33%. The total number of projected trips in 2020 will be 2.35 million, 2 million and 1.6 million for high, medium and low scenarios respectively. These projections translate to annual growth rate of 6.2% for high scenario and 4.57% for low scenario; which is higher than the annual growth rate of RDCO population for that period (City of Kelowna, 2010; InterVISTAS, 2007)

Water Transportation

The RDCO is blessed with many water bodies. Among them Okanagan Lake is the largest lake at 135 km in length, an average width of 4.5 km, and a surface area of 351 km². Historically, water transportation was a popular mode of travel, but no longer competitive with ground-based transportation because of longer travel time and inconvenience during winter season.

1.3.3 Road safety of the RDCO

Even though from 1996 to 2006 the RDCO population increased 18.2%, road traffic collision increased 87.8% in that period. In 1996 total number of collision was 2,067 and it almost doubled in 2006 with 3,882 collisions. Figure 1.8 shows the trend of road collision in RDCO. This figure indicates that the number of collision per thousand residents is increasing even though British Columbia (BC) is experiencing decreasing trend. That means each year the RDCO road users' lives are becoming less safe, which highlights the importance of road safety improvement of the region.

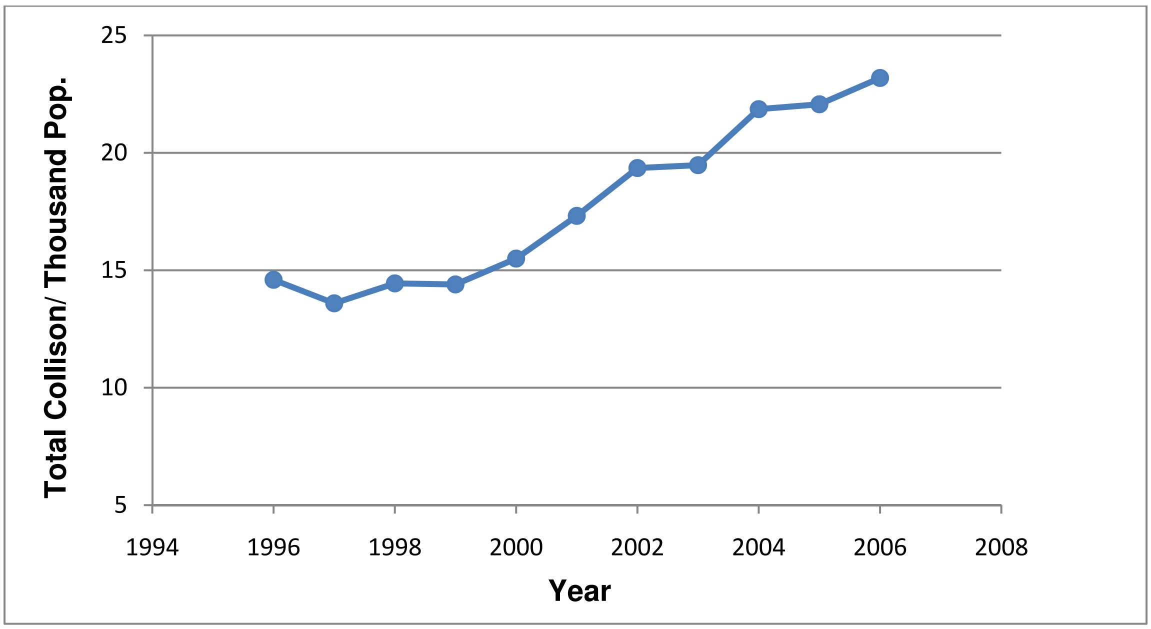


Figure 1.8 Total number of collisions per thousand populations in RDCO

1.4 Research objectives

The need for a safer, more energy efficient, and environmentally friendly transportation system promote the necessity of shifting from present auto-dominant transportation system to a more sustainable transportation system. Increased transit use has higher potential to achieve a sustainable transportation system. However reliable empirical results are needed for economic justification of major infra structure investment and policy changes. The primary focus of this research is to quantify the road safety benefits of increased transit by pursuing the following research objectives:

1. Develop an auto-based transportation planning model of the Regional District of Central Okanagan for (1) year 2006 and (2) year 2020.

2. Using the transportation planning models of 2006 and 2020, conduct an evaluation of RDCO road safety with the development of community based macro-level Collision Prediction Models (CPMs).

3. Using these models, road traffic collision analysis of the RDCO and identification of collision prone locations.

1.5 Thesis overview

The thesis is divided into six chapters. This first chapter discusses the topic of road safety, research motivation, the geography of the study area, and objectives of the research.

Chapter Two reviews the literature regarding traditional transportation modeling steps including trip generation, trip distribution, mode choice and trip assignment. It also reviews the safety literature, reactive and proactive engineering approaches, development and application of community-based macro-level CPMs.

Chapter Three describes the dataset and the methodology used to build transportation planning model of the region for (1) year 2006 (base year) and (2) year 2020. The chapter also describes the methodology used for the development and the use of macro level CPMs used in meeting the goals of this research.

Chapter Four describes the resulting AM period transportation planning model of the RDCO for years 2006 and 2020. It describes results for each step of the model and validates the results. It also describes four 2020 sub-scenarios, i.e. (1) do nothing (i.e. no road network and transit service improvement); (2) only road network improvements but no transit service improvement; (3) transit service improvements but no road network improvement; and, (4) improvement of both road network and transit services. This chapter also developed macro-level CPMs using 2006 data and predicts the number of AM period collisions for different sub-scenarios of 2020. Using the number of predicted collisions, road safety of each 2020 sub-scenario was evaluated.

Chapter Five discusses results of the RDCO road safety analysis. It provides collision density maps for different years and also identified collision prone zones using statistical techniques. It provides collision density map along Highway 97 and identifies top fifteen collision prone intersections in the RDCO. Later, this chapter analyses collision patterns to get better understanding of collisions and calculates the economic cost of RDCO collisions.

Finally Chapter Six presents research conclusions and recommends for future research work.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

The purpose of this literature review is to describe the theoretical foundations of transportation planning model and community-based macro-level CPMs. This chapter is divided into two main sections. The first section is about transportation planning model and the next section is about road safety improvement programs. Section 2.2 discusses about the importance and types of transportation planning model. It briefly reviews traditional four-step planning model by discussing each of the four steps of modeling. While Section 2.3 is about reactive road safety improvement programs and steps involved in reactive approach, Section 2.4 discusses about proactive road safety approach. Section 2.5 discusses about the development and application of community-based macro-level CPMs collision prediction models. Finally Section 2.6 summarizes the literature review.

2.2 Transportation planning

Transportation planning refers to practices for making transport policy, program and investment decisions. The primary purpose of the transportation planning process is to generate information useful to decision makers and is to find the consequences of transportation-related actions (Meyer & Miller, 1984). Traditionally, these models are used to forecast future traffic volume at a given point of time (Levinson, 1994). The main uses of transportation planning analysis are: (1) to improve and to modify existing facilities; (2) to design new facilities; (3) to predict future transportation network parameters such as traffic volumes, mode share percentages and link travel time; (4) to justify the performance of Transportation Demand Management (TDM) programs; and, (5) to observe the change in travel pattern due to change in land use.

Based on the objectives and scales of planning process, Litman (2009b) divided transportation planning in different categories which are as follows:

- *Traffic impact studies* evaluate traffic impacts for a particular development or project.
- *Local transport planning* develops neighbourhood and municipal transport plans.
- *Regional transportation planning* develops plans for a region.

- *State, provincial and national transportation planning* develops plans in a large scale for the province or the nation.
- *Strategic transportation plans* develop long-range plans for the future.
- *Transportation improvement plan* identifies projects that can be implemented in near future.
- *Corridor transportation plan* identifies projects to be implemented on a specific corridor.
- *Mode or area-specific transport plan* identifies various initiatives to improve the use of a particular mode (e.g. public transit, cycling, walking) or for a particular area.

Tennoy (2010) & Litman (2010) showed that the transportation system, travel behaviour, and land use are interconnected. Transportation demand is a sub-set of the Land Use (LU) planning process, because land use developments can significantly affect the spatial distribution of employment and population, and thus can affect the demand for transportation (Hasan & Hoque, 2008). Transportation modeling involves various steps. The following sections will discuss transportation planning model structure and steps involved in it.

2.2.1 Traditional transportation planning model

The four-step planning model is the most widely used model by transportation planning agencies because of its relatively understandable process and low economic cost (McNally 2000; Litman, 2009b). The total modeling process is divided into four steps, namely trip generation, trip distribution, mode choice, and trip assignment. At first the study area is divided into numerous aggregation units named Traffic Analysis Zones (TAZs). Generally TAZ boundaries are set by major roads or by zonal boundaries. Then total number of generated trips (i.e. produced and attracted) for each zone is calculated and these trips are distributed to different destinations. For each travel mode, trips are assigned to road network based on route specific generalized costs (i.e.travel time, travel distance, fuel cost, fare etc). Results from the final step are compared with actual observed road network volumes, and adjustments are made to calibrate the model until desired accuracy is achieved. The following Figure 2.1 illustrates the steps involved in planning process:

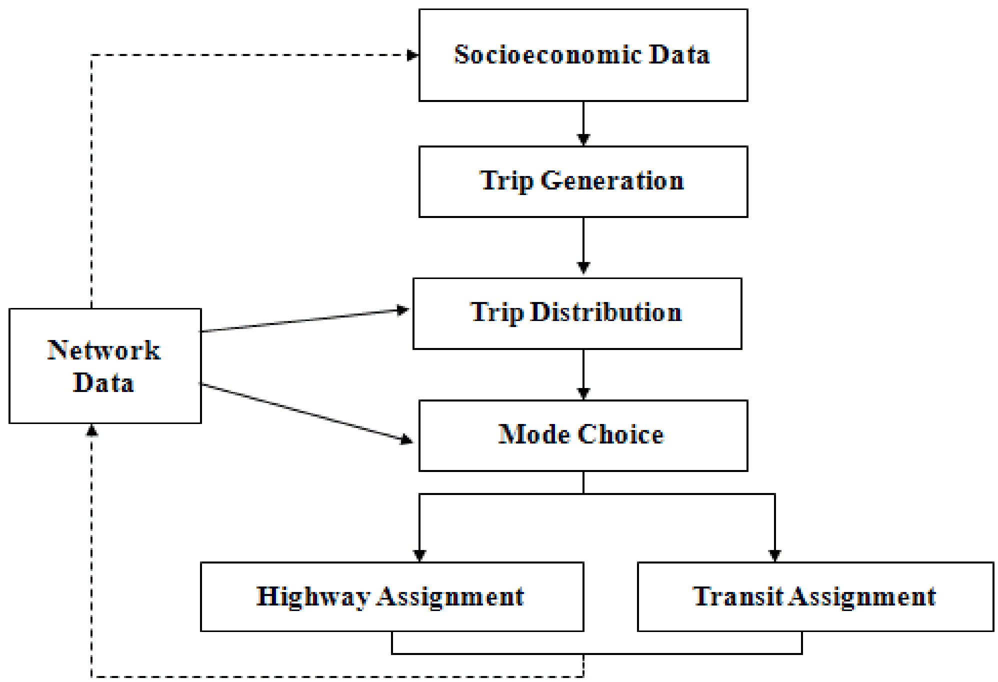


Figure 2.1 Four step transportation planning model

2.2.1.1 Trip generation

Trip Generation is the first step of the conventional four-step transportation planning process. It aims to predict the number of trips generated and attracted by each area unit (zone). In other words, this step answers the question of "how many trips produced or attracted to any zone". This step can calculate existing travel demand as well as forecast future travel demand.

Trips can be classified by purpose, by time of the day, and by person. Purpose based trips can be classified as trips for work, trips for education, trips for shopping, trips for recreation and other trips. Generally purpose based trips are classified as home-based work (HBW), home-based other (HBO) and non-home based (NHB). Trip purpose has the ability to affect travel behaviour of a person. But travel behaviour is highly influenced by the traveler's socio economic attributes such as income level, vehicle ownership, house hold size, age of traveler, family size and education level (Zhengbing et al., 2009). Buchanan (1963) and Handy (1993) also suggested that a person's travel is influenced by the character of the community he lives and also by the spatial structure of the region of which that community is a part. As the total number of generated trips is the sum of trip production and attraction, it is very important to understand variables that are responsible for producing and attracting trips. Meyer & Miller (2001) identified trip production variables as household income, vehicles per household, number of workers per household, residential density, and distance between trip origin and central business district (CBD); while trip attraction variables as employment levels, zonal business floor space, and accessibility to the work force. In addition to personal trips, freight trips are also important because of their significant contribution to congestion level. Freight trips are usually influenced by the number of employees, number of sales, and location of commercial firms (Mathew & Krishna Rao, 2007).

There are commonly three approaches used in the trip generation analysis: (1) regression analysis; (2) cross-classification analysis; and, (3) trip rate analysis. Regression technique is one of the popular methods for trip generation model (Ortuzar & Willumsen, 2001). It attempts to discover a relationship between the number of generated trips and the socioeconomic characteristics within the zone. It expresses the number of trips by a function containing explanatory variables. The most common form of trip generation model is a linear regression model as given in equation (2.1):

$$T_i = a_0 + a_1x_1 + a_2x_2 + \ldots\ldots + a_ix_i \tag{2.1}$$

where,

T_i= total number of generated trips;

a_i = the coefficient of regression equation; and,

x_i = explanatory variables.

There are some weaknesses associated with regression method. If the variable values are aggregated at zonal level, then lots of data (i.e. variation) can be lost during aggregation. Also if the trip generation rate and variables are not linearly related, then assumption of linearity would introduce some bias (Seo et al, 2000). Finally, to perform regression analysis it requires household survey and other data which may not be possible to collect due to some constrains.

Cross-classification technique is another popular method that considers trip making characteristics as a function of household attributes. It assumes similar trip making characteristics for households that are classified into same category (Anderson & Olander, 2002). This method works with the expected number of daily trips generated per household whose characteristics are categorized by household attributes (i.e. income per household, number of vehicle per household). For each category, the number of trips per household is tabulated in a table named 'cross-classification table'. Separate tables can be prepared for separate trip purposes. When each element (or cell) of the table is multiplied by the number of households of that category and these values are added, then total number of trips for that category is obtained. Adding the total number of trips of each category will give the total number of trips for the entire zone. For zone 'i', purpose 'p' and person type 'n', the total number of generated trips(T_{ip}) is (Daly, 1997):

$$T_{ip}=\sum_k h_{ki}.\pi_{knp} \tag{2.2}$$

where,

h_{ki} =the number of households of category 'k' in zone 'i';

π_{knp} = the average number of trips made by a household of that category by person type 'n' and purpose 'p'.

The average number of trips (π_{knp}) per category is estimated from a local home interview survey or is taken from standard trip-rate data. This method is a widely used trip generation model. However, this method has some weaknesses which are as follows: (Goodman & Kruskal, 1979; Stopher & McDonald, 1983; Chatterjee et al., 1977; Kitamura, 1981)

- No statistical test to evaluate the reasonableness of the observed trip rates;
- As the number of available sample for each cell is different, reliability of the individual cell value is not consistent. Even sometimes some cells do not have any values;
- To determine the best set of categories (i.e. household groups), there is no effective way unless an extensive trial-and-error procedure;
- The average value is used for trip rates and average value doesn't consider the variances among households; and,
- The estimated trip rates are reliable only when transport systems and land-use patterns have not undergone major changes (e.g. no new traffic analysis zone).

Third method for trip generation is the application of trip generation rate. Ortuzar & Willumsen (1990) suggested the use of trip rates instead of zonal total trips so that the effect of the zone size can be eliminated. This method can estimate the total number of trips for different land use and building types. Using attribute specific trip rate, it is possible to calculate total number of trips for an attribute, and adding trips for each attributes, total number of trips generated from an analysis zone can be calculated. The Institute of Transportation Engineers (ITE)'s trip generation informational report provides trip generation rates for different types of land uses, institutes and residential areas. These rates are developed based on the data obtained from the United States and Canada. It gives trip rates for different time periods (i.e. weekdays, weekends, peak hour) of a day and also the percentage of traffic to enter or exit a site. The limitation of rate method is it does not consider the location of the area, cost of transportation, and many other important factors. It is important to calibrate trip rates using locally observed data.

2.2.1.2 Trip distribution

Trip distribution is a process by which produced trips from one zone are distributed to other zones of the study area. This process determines how and where the produced trips will be distributed. Each zone is taken once at a time and a determination is made of how the produced trips will be attracted by other zones. The trip pattern within a study area is usually represented by means of a trip table which is called Origin-Destination (O-D) table. The distribution of trips is assumed to be dependent upon various factors such as availability of jobs, transportation facilities, travel time and zonal attraction.

Different methods have been used in trip distribution analysis. Among them, Fratar model is one of the oldest trip destination choice models. This method was presented by Thomas J. Fratar in 1954. By using growth rates the model extrapolates a base year trip table to obtain future traffic volume. Therefore, for each zone base year trip table and growth factor are needed. The method suffers from some limitations as it doesn’t take account the change of spatial accessibility due to increased supply or changes in travel patterns and congestion (Levinson & Kumar, 1995).

The Intervening opportunities model is another trip distribution model. Stouffer (1940) gave the basic hypothesis of the model as ‘for a particular trip purpose the probability of choosing a destination is directly proportional to the number of opportunities at the destination zone and inversely proportional to the number of intervening opportunities that are closer to the origin’. In this method, the attraction of a zone does not decrease continuously with distance, but by the number and attractiveness of intervening opportunities that is passed by traveler. The in-situ attractive properties of the destination are modeled as opportunities and the impedances are measured in terms of the number of opportunities in other zones that are close to the origin. Normally the application of the intervening opportunities model requires the destination zones to be specified in an order of decreasing accessibility from each origin (Muranyi and Miller, 1966; Ruiter, 1967; Wills, 1986). However, the model is not often used in practice, probably for the following reasons:

- The theoretical basis is not very clear to understand by practitioners;
- The idea of matrices with destinations ranked by accessibility is more difficult to handle in practice.

The Gravity Model is the most widely used trip distribution model to distribute trips between zones. This model is derived from Newton's Inter-Planetary model of attraction. The assumption of the model is all trips starting from a given zone are attracted to other zones in direct proportion to the attractiveness of the destination zone, and in inverse proportion to the travel impedance (i.e. operating cost, travel time, distance, parking fee, walk time, wait time etc.) between the zones (Sosslau et al., 1978). So the attractiveness and accessibility of zones are used as inputs, and, the output is an origin-destination table to show the distribution of trips from each zone. General form of the gravity model is given in equation (2.3):

$$T_{ij} = P_i * \frac{A_j F_{ij} K_{ij}}{\sum_{j=1}^{n} (A_j F_{ij} K_{ij})} \qquad (2.3)$$

where,

T_{ij}=number of trips from zone 'i' to zone 'j';

P_i=number of trip productions from zone 'i';

A_j=number of trip attractions in zone 'j';

F_{ij}= friction factor; and,

K_{ij}=adjustment factor.

K-factor is the reflection of socio-economic characteristics of each pair of zone. It is a specific zone-to-zone adjustment factor to allow the incorporation of the effect of socio-economic variables on travel patterns. It can reflect some conditions that are not captured by the model or difficult to formulate. The K-factor is the ratio between the observed and the estimated trips for any given pair of zones (Duffus et al., 1987). Stopher & Meyburg (1975) suggested the use of same K-value for future prediction unless K-factor is statistically related to the socio-economic characteristics of the zones. Later Duffus et al. (1987) suggested that the use of K-factors is not necessary in future predictions because K-factors can result in larger errors in prediction than without their use.

The friction factor (F) represents the spatial separation between zones. F is the reciprocal of impedances. The shape of the F-factor curve is usually an inverse exponential function (Duffas et al., 1987), which implies the increase of impedances between a pair of zones will decrease the probability of trip making between those zones. For different trip purposes the shape of the F-factor curve remains same, but for the same impedance value friction factor (F) changes based on the purpose of the trip. So it is important to calibrate the model to obtain purpose based friction factor equation. This can be done by making household travel survey and by traffic count. But this process is data extensive, time consuming, and expensive. Anderson et al. (2002) suggested that for a small area if the average trip length is similar for all trip purposes, then using the same value of friction factor will distribute trips in a similar fashion for all purposes.

Generally travel time is considered as the most important impedance in trip destination choice. However, inclusion of variables other than travel time can give better prediction of travel behaviour. This can be done by the use of discrete models. The role of discrete choice models in decision-making support has significantly grown, particularly in transportation, as it is critical to understand and forecast choice behaviour in a detailed way (Bierlaire, 2007). With the development of disaggregate approaches, discrete choice models have been used for travel demand estimation (Ben Akiva & Lerman, 1985; Bhat et al., 1998; Train, 1998). Discrete choice model assumes that the set of alternatives considered by the decision-maker, is finite and discrete (Levinson & Kumar, 1995). While the gravity model usually uses travel time as impedance, discrete models bring different discrete variables that represent a choice from a set of mutually exclusive choices. These discrete variables are expressed by a utility function. According to utility maximum theory, the traveler will always try to maximize the utility for his/her journey. Probability model for destination choice can be formulated using utilities. Allen (1984) used utilities of a discrete choice model to determine the impedance of trip distribution model. Ben-

Akiva & Lerman (1985) also developed combination of destination and mode choice utility based models for work and non-work trips.

2.2.1.3 Mode choice

Mode choice is the third step of traditional four-step transportation planning model. This step attempts to assign person-trips to various alternative modes (i.e. auto, transit, bike etc.) available in the study area. This step determined what mode will be used and to what extent. Mode choice depends on the behaviour of the traveler. There are two basic ways of modeling travelers' behaviour. One approach is called "aggregate approach" which directly models the mode share of all or a segment of travelers by choosing each alternative as a function of mode attribute characteristics and traveler's average socio-demographic attributes. The second approach is called "disaggregate approach" which recognizes travel behaviour as a result of numerous individual decisions and models individual choice responses as a function of available mode characteristics and socio-demographic attributes of each individual.

Koppelman & Bhat (2006) suggested that for the travel behaviour modeling of a group of individuals, the disaggregate approach has several important advantages over the aggregate approach. First, the disaggregate approach works with individual travel behaviour to explain the changes in travel behaviour as a result of changes in mode alternative attributes. Second, the disaggregate approach is likely to be more transferable to different time and different geographic context. Third, as discrete choice models are used to understand individual travel behaviour, it helps to change individual travel pattern in a proactive desired way by adopting mode attribute related strategies. Fourth, the disaggregate approach is more cost effective and efficient than the aggregate approach in terms of model reliability per unit cost of data collection, because aggregate approach loss variability of data during aggregation process and to obtain same level of precision aggregate approach require more data. Finally, if properly specified, disaggregate models can obtain un-biased parameter estimates than that of aggregate model.

Traveler's mode choice is related to socioeconomic variables and mode-related attributes. Arasan et al. (1996) identified three basic factors behind mode choice decision as: characteristics of the journey (e.g., length, time of day, and purpose), traveler's socioeconomic characteristics, and the transport system. Racca and Ratledge (2004) identified land use pattern as another important variables. However, travelers' mode choice not only depends on the socioeconomic structure and transport system but also on the traveler's perception of the mode (Gebeyehu & Takano, 2007).

In mode choice decision, utility is a very important parameter. Utility is an indicator of value to an individual. It is a function containing alternative attributes and individual characteristics that describes an individual's utility valuation for each alternative. Individual's choice behaviour is the outcome of a decision process in which individual tries to maximize his/her utility (McFadden, 1978; Ben-Akiva & Lerman, 1985). It is modeled by random variables, in order to account for many sources of uncertainty in the decision process.

Utility of each alternative is related to its attributes such as travel time, in-vehicle travel time, out-of-vehicle time, cost of travel, parking facility, walk distance and reliability on time arrival. Utility is also dependent on road users' characteristics such as income of traveler, traveler's sex and age, number of automobile in traveler's household and number of adults in the traveler's household. The utility value of a mode 'i' can be expressed as:

$$U_i = A + BX + CY + DZ + \ldots \qquad (2.4)$$

where, A, B, C, D are the coefficients and X, Y, Z are mode attributes.

The utility maximization theory states that a traveler will select a mode which will give him/her highest utility during decision making. That ensures the certainty of choosing highest utility based ranked mode under the observed choice conditions. The utility model that contains certain predictions of choices is called deterministic utility model. This model assumes that the analyst will understand individual's perception about different alternatives and the decision making process. But practically it is not possible to fully understand individual's perception and circumstances. Therefore, the deterministic model may lead to error in behaviour prediction.

Koppelman & Bhat (2006) identified three primary sources of error in using deterministic utility functions. First, traveler's perception about the alternative attribute may be incomplete or traveler may have incorrect information about some or all of the alternatives. Second, the analyst or observer may has incorrect understanding of the same attribute relative to the individuals and also incorrect or incomplete understanding of the utility function that is used to predict individual travel behaviour. Third, the analyst may not know or account specific circumstances of the individual's travel decision. Other than the above mentioned sources Manski (1977) identified other sources of error, such as measurement errors and proxy variables.

Therefore, mode choice models should recognize the analyst's lack of information about individual behaviour and alternative attributes. The analysts should consider the uncertainty associated with mode choice variables to achieve better result than predicting travel behaviour with certainty. To take account of uncertainty, utility can be considered in two components: (1) first component of the utility function is called deterministic or observable portion of utility which is observed by the analyst; and, (2) second component accounts the difference of unknown utility used by the individual and the analyst (Horowitz, 1991; Ben-Akiva & Bierlaire, 2003; Koppelman & Bhat, 2006). The second component is expressed by a random error term and the utility function becomes:

$$U_{it} = V_{it} + \varepsilon_{it} \tag{2.5}$$

where,

U_{it} is the true utility of the alternative 'i' to decision maker type 't';

V_{it} is the deterministic or observable portion of the utility estimated by the analyst; and,

ε_{it} is the random error part that represents uncertainty.

The error terms are unobserved, unmeasured and analyst has no prior information to model error term. A wide range of distributions could be used to represent the distribution of error terms over individuals and alternatives. The error terms for each individual can be assumed as a sum of small components which are missing from the deterministic utility function. Each of these components has relatively small impact on the value of each alternative and according to central limit theorem the sum of these small error components will be normally distributed (Koppelman & Bhat, 2006). This normal distribution assumption leads to the formulation of the Multinomial Probit (MNP) probabilistic choice model. This MNP model has the advantages of capturing all kind of correlations among alternatives (Ben-Akiva & Bierlaire, 2003) which makes it a flexible tool for demand modeling (Horowitz, 1991). However, due to high mathematical complexity of the MNP model, it becomes difficult to estimate, interpret and forecast. So in practice the application of MNP model has become very limited (Ben-Akiva & Bierlaire, 2003; Koppelman & Bhat, 2006).

The simplest and the most frequently used utility based model is Multinomial Logit Model (Horowitz, 1991; Ghareib, 1996; Thamizh Arasan et al., 1996; Ben-Akiva & Bierlaire, 2003). Koppelman & Bhat (2006) described the main assumptions of the Multinomial Logit Model (MLM) as: (1) the error components have an extreme-value (or Gumbel) distribution; (2) the error components are identically and uniformly distributed across attributes; and, (3) the error components are identically and uniformly distributed across individuals. Based on these assumptions a probability based model can be developed that will give the probability of choosing one mode over others. Following equation (2.6) is the general form of MLM which is given by (McFadden, 1974; Domencich & McFadden, 1975; Ben-Akiva & Lerman, 1985; Koppelman & Bhat, 2006). The probability of choosing an alternative 'i' (i=1,2,3…,j) from a set of 'j' alternatives is:

$$Pr(i) = \frac{e^{(V_i)}}{\sum_{i=1}^{j} e^{(V_i)}} \qquad (2.6)$$

where,

$Pr(i)$ is the probability of the decision-maker choosing alternative 'i'; and,

V_i is the systematic component of the utility of alternative 'i'.

Equation (2.6) implies that the probability of choosing an alternative will increase monotonically with the increase of the utility of that particular alternative and the probability will decrease with the increase of other alternative's utility. Following Figure 2.2 shows that if transit service improves, then the associated utility of transit will be higher and the percentage of transit ridership will increase. On the other hand, if the auto mode gives higher utility then people will use more autos and transit ridership will decrease.

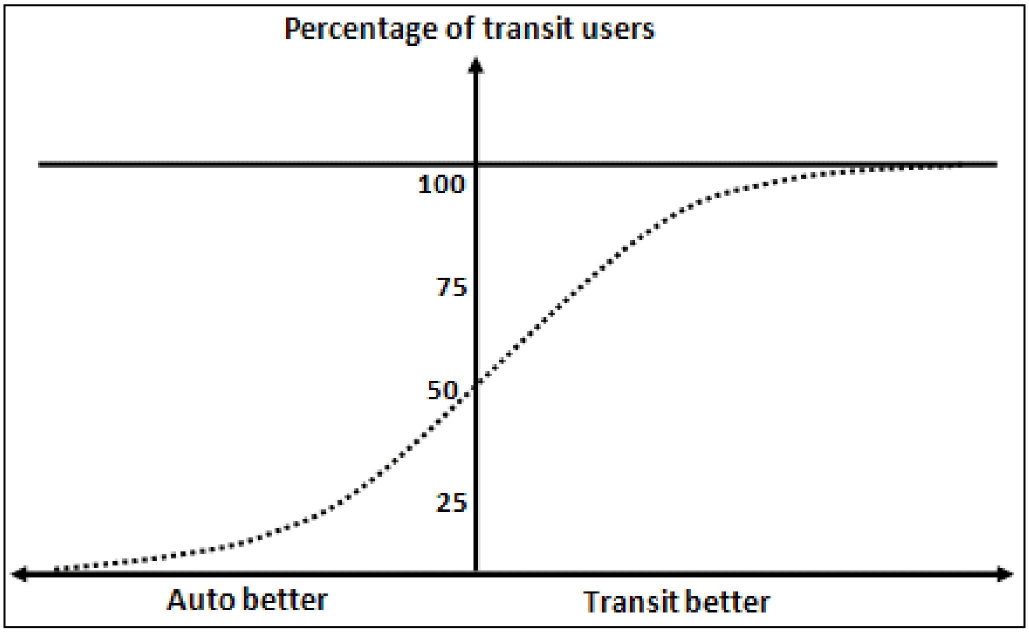


Figure 2.2 Multinomial Logit Model (MLM) curve

Koppelman and Bhat (2006) described the advantages of MLM as: (1) simple mathematical form; (2) ease of estimation and prediction; and, (3) ability to add or remove any alternatives from the model. However, the MNL model has been widely criticized for its Independence of Irrelevant Alternatives (IIA) property. Ben-Akiva & Benlaire (2003) defined 'IIA' property as "the ratio of the probabilities of any two alternatives is independent of the choice set". According to the property any change in an attribute of one alternative will have the same proportional impact on the probability of other alternatives. It assumes equal competition between all pairs of alternatives, but this assumption is inappropriate in many choice situations. For example, bus and light rail both are transit service and sometimes they have similar attributes such as same travel time, same fare structure, and same lack of privacy. According to the IIA property MLM doesn't consider these similarities and if these similarities are not considered it may give wrong prediction of travel behaviour. To overcome this limitation various models have been proposed so far and among them Nested Logit (NL) model (Williams, 1977; McFadden, 1978; Daly and Zachary, 1978; Gil-Moltó & Hole, 2003) is widely used.

The NL model retains most of the computational advantages of MLM except deviation of the IIA property (Borsch-Supan, 1987). The NL model divides alternatives of similar characteristics into subsets (nests) so that IIA assumption is valid within each nest but not for alternatives in other nests. So alternatives in a common nest will exhibit a higher degree of similarity and competitiveness compared to alternatives in other nests. Following Figure 2.3 shows a NL model structure where auto and public transit modes are divided into sub-nests. According to the figure, the competiveness between bus and light rail will be higher than other mode (i.e. auto, walk, and bicycle). Similarly for auto, drive alone and shared drive will exhibit more similarities and competiveness when auto is chosen as the mode of travel.

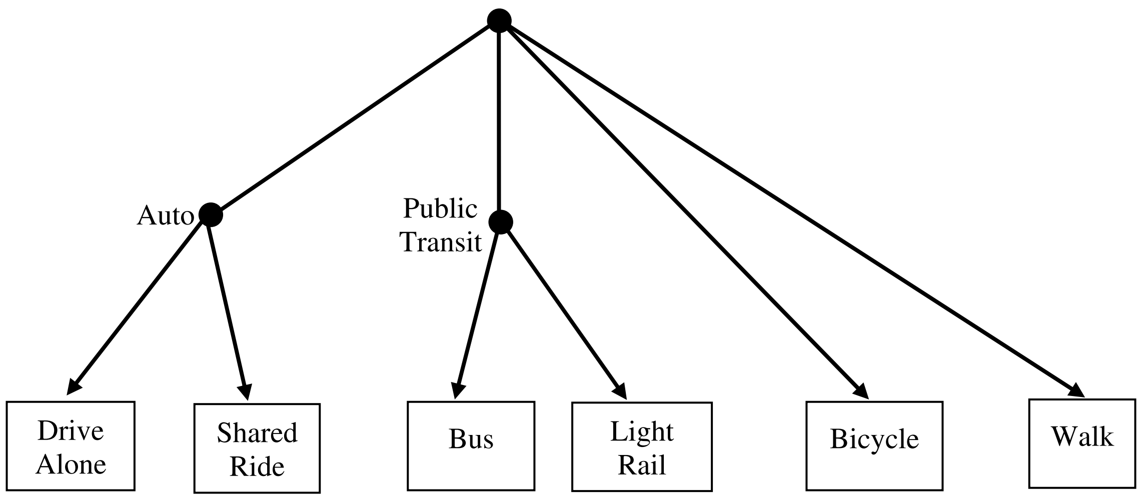


Figure 2.3 Nested Logit model structure

2.2.1.4 Traffic assignment

As the fourth and final step of the traditional transportation planning, traffic assignment is the process of how travel demand and transportation supply interact in transportation networks .This step assigns vehicular trips into road network and also determines which path will be used for a given set of origin-destination pair. For each mode it assigns trips based on the minimum generalized cost for the O-D pairs. The fundamental aim of the traffic assignment process is to reproduce the observed transportation system with traffic volume on roads. The major purposes of traffic assignment procedures are:

1. To estimate traffic volume on links of the network and to aggregate network measures.
2. To estimate inter zonal travel cost (i.e.travel time and distance, operating cost) between trip origins and destinations. These costs can be used in the trip distribution step to refine the model.
3. To analyze the travel pattern of each origin to destination (O-D) pair.

4. To identify heavily congested links to perform better traffic management.
5. To identify the routes that are used for a particular origin to destination (O-D) pair.
6. To analyze which O-D pairs use a particular link or path.
7. To obtain turning movements volume for the design of future junctions.

Different methods are used to assign vehicular trips on the road network. Among them All-or-Nothing is the oldest and simplest method. In this method all trips for a given O-D pair are assigned to a single path offering minimum impedance (i.e. cost, distance, time etc.) for the journey. This method is also called as Shortest Path method because all drivers assumed to use the shortest path for a given O-D pair. It assumes the capacity of links is unlimited and each link can carry any volume of traffic. This model is unrealistic because it assigns all trips in one single route regardless of the actual capacity of the link. As capacity and travel time on a congested route is not considered, travel time is a fixed input in the model and it does not vary depending on the congestion on links. However, this model can act as a building block for other types of improved assignment methods; because of its ability to identify driver's desired route for a given pair of O-D in the absence of congestion (Mathew, 2007).

Generally link travel time is a function of traffic volume and travel time increases with increased traffic volume. The relation between travel time and traffic volume is expressed by 'link cost' function. According to Suh et al. (1989), the link cost function used in traffic assignment is also named as capacity function, capacity restraint function, link capacity function, link performance function, and congestion function. Conventionally it is assumed that the link cost function increases monotonically with the increase of traffic volume. Different types of link cost functions have been proposed to calculate link travel time and delays due to congestion (Rudjito, 2006). The most widely used link cost function is 'BPR function' which is proposed by The Bureau of Public Roads (BPR), USA in 1964. Following equation (2.7) is an example of BPR function:

$$T = T_0[1 + \alpha(\frac{Q}{Q_{max}})^{\beta} \tag{2.7}$$

where,

T= actual travel time; T_0 = free flow travel time on uncongested link;

Q= actual traffic volume; Q_{max}= link capacity; and, α, β = constant

Generally the recommended α and β values are 0.15 and 4 respectively. The value of α=0.15 means that at capacity the travel time of the link will be 15 percent higher than the travel time at free flow condition. The parameter β determines the slope of the link cost function to represent level of congestion.

Considering the link capacity constrain, various types of traffic assignment techniques have been proposed such as incremental, user-equilibrium, system optimal and dynamic assignments. Incremental assignment is a process in which fractions of traffic volumes are assigned to network step by step. In each step, a fixed proportion of total trips are assigned based on all-or-nothing assignment. After assigning trips in each step, travel time for each link is recalculated using link cost function and the next increment of trips is assigned to a route that offers shortest travel time. This process continues until all the trips are assigned. Higher number of increments will increase the accuracy of the results. However, this method is influenced by the order in which volumes for O-D pairs are assigned. So there exists a possibility of bias in the results.

Selecting a route offering minimum cost is the normal tendency of drivers. This human tendency of minimizing cost was described by Wardrop (1952) in his first principle which is: "The travel time between a specified origin and destination on all used routes is same or less than the travel time on routes that would be experienced by a traveler on any unused route". This principle assumes that drivers will use only those routes which offer least cost to travelers and others routes of the network with higher cost will not be used. It gives the impression that drivers have a perfect knowledge of the travel costs of each route and driver can choose the best route to minimize his/her cost. This deterministic behavioural assumption leads to a model called 'Deterministic User Equilibrium' (DUE) model. However, it is not practical to completely determine driver's behaviour and this limitation of DUE model exhibits less accuracy to represent reality (Rudjito, 2006).

Wardrop's first principle was only for a single user class and it doesn't consider volume of traffic on other routes of the network. So Wardrop (1954) proposed an alternative way of assigning traffic to network, known as Wardrop's second principle which is: "At equilibrium condition, traffic in congested network should be arranged in such a way that the average travel time is minimum". According to this principle *drivers can cooperate with one another or can get* *network information in order to minimize total system travel time.* This assumptions lead to another assignment method called 'System Optimum Equilibrium' (SOE) assignment. However, the assumption of SOE is not realistic, but it can be useful to transport planners and engineers to minimize the total travel costs of the system (Mathew, 2007).

Again, deterministic driver's behaviour assumes that drivers will follow a route with minimum travel cost. But in reality this assumption is very strict in a sense that the driver's perception of travel cost for a particular route should not be the same for all drivers. Stochastic User Equilibrium (SUE) model assumes that driver's perceptions of costs of a particular route are not identical among all drivers. SUE model considers the variation of driver's evaluation of a route. So some drivers will consider one route as the cheapest route, whereas other drivers will choose another route as the cheapest. Thus, SUE model may select some routes as cheapest which are actually not cheapest (Van Vliet & Dow, 1979).

Previously discussed incremental, equilibrium, DUE, SUE models are the part of 'Static Assignment' technique. Static assignment is such a technique that assumes traffic system variables (i.e. traffic flows, travel times, costs) for an O-D matrix are constant over the analysis period. These static methods do not have the ability: (1) to clearly represent the variation of demand flows and network performances during congestion; and, (2) to reproduce some important dynamic phenomena (e.g. vehicle queue formation and dispersion) (VISUM, 2007). Even sometimes in large networks static assignment models overestimate link volumes and in that case Dynamic Traffic Assignment (DTA) can solve the problem (Merchant & Nemhauser, 1978; Carey, 1986).

The properties of dynamic traffic assignment have the ability to represent the actual travel behaviour, and hence it can improve the accuracy of the model results (Szeto & Lo, 2006). Szeto & Lo (2006) mentioned that the properties of DTA strongly depend on two components: (1) the travel choice principle; and, (2) the traffic-flow component. The travel choice principle models travelers' intensity to travel. If any travel is made then travel choice principle determines how to select modes, routes and departure time. Travel time is the important element of considerations in that process and Stochastic Dynamic User Optimal (SDUO) is used to make travel choice. Here SDUO is the dynamic extension of Wardrop's (1952) principle (Ran & Boyce, 1996). The traffic-flow component represents the propagation of traffic on a transport network and it gives the network performance in terms of travel time. Lo & Szeto (2002) shown that modeling this part could be cumbersome and rendering result formulation could be hard to solve. However, the uses of DTA models are still theoretical and they are not widely used for traffic assignment.

In general, trip based traditional transportation planning models are widely used due to its simplicity. But there are some limitations associated with the traditional modeling approach. The main sources of the limitations are given below: (Kitamura et al., 1995; & McNally & Rindt, 2008)

- The traditional four-step procedure does not represent the decision mechanisms underlying travel behaviour. It is very important to understand traveler's behaviour because even small changes can change individual travel pattern.
- The four-step procedure treats each trip as an independent entity for analysis, ignoring the spatial and temporal interrelationship between trips and activities.
- Inadequate specification of the interrelationships between travel and activity participation, including activity linkages and interpersonal constraints.
- It does not consider travel as a demand derived from activity participation decisions. People do not decide how many trips to make before deciding what to do, where to go, and how to get there
- It does not specify individual choice sets, when choices have to make among alternatives in a constrained environment.

2.2.2 Activity based model

As human activities are linked to each other, trips made to pursue different activities are also inter-related. So consideration of human behaviour has lead to another transportation planning technique which is called 'activity based modeling'. It starts with the recognition of human behaviour to understand why and how human activities are engaged over a time period. The lack of human behavioural consideration in traditional trip-based models has been criticized by several authors (Ben-Akiva, et al., 1998), who emphasize the importance of activity-based approach.

Jones et al. (1990) provide the definition of activity-based analysis as a "framework in which travel is analyzed as daily or multi-day patterns of behaviour, related to and derived from differences in life styles and activity participation among the population." Kitamura (1996) describes the activity-based approach as the only approach that can offer coherent frameworks for policy analysis and demand forecasting with different types of TDM programs to improve mobility and reduce environmental impact. A framework of the activity-based modeling approach can be found in Appendix A. The complex, data intensive and expensive nature of this approach hasn't made it widely used for transportation planning.

The activity based transportation planning approach has several advantages over traditional trip-based approach, as summarized below (Kitamura et. al., 1995; Jones et. al., 1990):

- It focuses on travel behaviour pattern and treats daily activity-travel pattern as a whole, not as discrete trips.
- It has the characteristics to incorporate various trips making constraints into model.
- It can address the issue of induced or suppressed demand by representing activity engagement human behaviour.
- It examines timing and duration of activities to predict travel behaviour along a time axis.
- It can evaluate the change in travel behaviour due to any TDM program. Therefore, it can act as a comprehensive evaluation tool to find the effect of transportation policy on daily travel pattern of people.
- It is flexible and versatile to be modified for any specific study objectives or to address various policy scenarios.
- It considered spatial, temporal and inter-personal constraints to predict better travel behaviour.

2.2.3 Transportation planning model software

Transportation planning models require huge amount of datasets and it is very hard to handle those spatial and non-spatial data. Planning for a large area involved lots of TAZs and associated data. So to do the transportation planning, several softwares are used, including:

Emme3

It is a very popular transportation planning modeling software. Over two decades, modellers and planners have relied on Emme and it has become one of the most sophisticated transportation planning model software (INRO, 2010). It has the ability to model private and public transit assignment, strategy analysis, and automation framework. It offers planners a complete and comprehensive set of tools for demand modelling, multimodal network modelling, and implementation of evaluation procedures. GIS integration and powerful 'Emme prompt' tool make it a very powerful modeling tool. It requires very good knowledge of script writing and sometimes it becomes very difficult to use for new practitioners.

TransCAD

TransCAD is the first Geographic Information System (GIS) designed transportation modeling software to store, display, manage, and analyze transportation data (Caliper Corporation, 2010). It combines GIS and transportation modeling capabilities in a single integrated platform. It can be used for all modes of transportation, at any scale or level of detail. It provides (1) powerful GIS engine; (2) tools for mapping, visualization, and analysis; and, (3) application modules for routing, travel demand forecasting, logistics, and territory management.

VISUM

VISUM is a comprehensive, flexible software system for transportation planning, travel demand modeling and network data management (PTV, 2010). It is used for metropolitan, regional, statewide and national planning applications. It can integrate all modes of transportation (i.e., car, car passenger, truck, bus, train, pedestrians and bicyclists) into one network model. It provides a variety of assignment procedures and 4-stage modeling components which include trip-end based as well as activity based approaches. It can be also integrated with GIS and other format dataset. This software is relative easy to learn for new practitioners.

2.3 Reactive road safety improvement programs

Worldwide road collisions have become a global major social and economic burden. Recognizing the fact, many road authorities and researchers have initiated various road safety improvement programs (RSIPs). While many RSIPs focus on education, enforcement, and engineering, transportation engineers focus mainly on engineering measures. The traditional road safety improvement program is based on reactive approach which addresses road safety in reaction to existing collision histories. The objective of reactive RSIPs is to identify and treat hazardous locations, based on the analysis of collision, traffic, and highway data (Sayed, 1998a; Khisty & Lall, 1998). Hazardous locations are defined as those locations where the collision rate is higher than the norm. Hazardous locations may consist of intersections or segments of roads/ highways. The main three steps in reactive RSIPs are (Sayed, 1998a):

1. Location identification: detecting hazardous locations (black spots);
2. Problem identification: diagnosing collisions and locations to find causes behind road collisions; and,
3. Solution identification: finding remedial measures to solve existing problem.

2.3.1 Black spot programs

Sayed et al. (1995) showed that 32% collisions in North America and the United Kingdom are due to road environment related factors. If collision prone locations (CPLs) can be identified, it is possible to reduce road environment related collisions by applying engineering measures. So it is very important to identify collision prone locations or black spots. A black spot is defined as any location that has a collision potential of significantly higher than the normal collision potential of a group of similar locations (Sayed, 1998a). Black spot programs include identification and ranking of hazardous locations. To ensure the proper use of limited resources spent only on true hazardous locations; it is important to have a sound procedure which will properly identify and rank black spots so that resources can be spent on those locations which have the highest potential for road safety improvement. In identifying black spots, collision potential of a location can be expressed by several measures, including: collision frequency, collision rate, collision severity, severity index, or a combination of them.

The collision frequency measure gives the number of collisions occurring at a location during a specific time period. To minimize the effects of random fluctuations, Zegeer (1982) recommended the use of an observation time period of between one and three years. McGuigan (1982) and Sayed (1998b) defined the limitations of this measure as it doesn't consider traffic volume or exposure level when comparing across different locations. As high traffic volume locations exhibit higher probability of collisions, frequency measure can lead to a bias by ranking high volume locations as more collision prone than low volume locations, even though low volume locations may exhibit higher collision potential. McGuigan (1982) recommended the use of traffic volume to identify black spots.

Collision rate is defined as collision frequency divided by some unit of exposure such as million-vehicle-kilometres (MVK) for sections, and million-entering-vehicles (MEV) for intersections. If any location experiences higher collision rate than a defined collision rate, then that location is identified as a black spot. Hauer (1995) and Hauer et al. (2002) identified collision rate as a traditional approach. However, this method has several problems. First, it identifies low volume locations as hazardous even though these location exhibit low collision frequency. Second, 'collisions per unit traffic volume' ratio assumes a linear relationship between traffic collisions and traffic volumes, which has been shown to be incorrect (Hauer, 1995). Third, for short road segment lengths, results can be biased (Zegeer, 1982; Nicholson, 1980).

Another measure to identify black spot is called as Collision Severity Index (CSI). It is defined as the weighted sum of fatal, injury, and Property Damage Only (PDO) collisions. Different weights are used by different agencies based on the cost and impact of collisions. In British Columbia (BC) some road authorities weight the severity of fatal collisions (F) as 100 times of PDO collisions and 10 times of injury collisions (I) (Sayed, 1998a), that gives:

$$CSI = 100F + 10I + PDO \qquad (2.8)$$

Many transportation agencies also use rate-quality control method to identify black spots. The method is based on the assumption that traffic collisions are rare events and the probability of their occurrence can be approximated by the Poisson distribution (Zegeer, 1977). This method compares the traffic collision rate for a particular intersection or roadway segment with a critical rate and if the observed collision rate is equal or higher than the critical rate, then that location is identified as a black spot. The critical collision rate is calculated by a function containing the average collision rate of similar locations and the vehicle exposure at the study location.

Among all of these measures, collision frequency and collision rate methods are widely used. Both of these methods have their limitations and there is a debate of choosing one method over another them. Zegeer & Dean (1977) suggested the use of both collision frequency and collision rate criteria in a two-step procedure to overcome their individual weaknesses in black spot identification. Commonly, collision frequency is used as the first step to identify black spots and collision rate is used to rank identified locations. However, whatever method is used, it requires some statistical techniques to identify black spots which is discussed in the following section.

2.3.2 Statistical techniques

Road collisions are not normal events, rather they are random events. Due to random nature of collision, it is not possible to calculate the true mean collision frequency of a location. But it can be estimated as an 'expected' value using reasonable assumptions and empirical techniques (Sayed, 1998a). Statistical techniques are used to eliminate randomness of observed collision frequency and to estimate the expected value of the mean collision frequency with some degree of accuracy.

The accuracy of an estimate is usually expressed in terms of standard deviation or variance of the estimate. Thus, calculation of variance of mean collision frequency is very important in road safety analysis. The value of variance depends on the distribution assumed for collision occurrence. As collisions are independent, random and non-negative events, Poisson distribution assumed to fit better because of the following properties: (Sayed, 1998a; Johnson, 2005):

- The number of events occurring in a particular time interval or specified region is independent of the number of collision that occurs in another time interval or region. (i.e. independent);
- The probability of occurring more than one event during a very short time interval or small region is negligible (i.e. rare events); and,

- The probability that a single event will occur during a very short time interval or in a very small region is proportional to the length of the time interval or size of the region.

One important statistical issue in road safety improvement programs evaluation is selection bias or Regression-to-the-Mean (RTM) bias (Hauer et al., 1988; Sayed & de Leur, 2001a). The RTM is the tendency of extreme events of a random variables to be followed by less extreme events if no change has occurred in the underlying causal mechanism (and vice versa) (Hauer et al., 1988; Sayed et al., 1997; Sayed & de Leur, 2001a). It means "highest goes lower and lowest goes higher". In road safety analysis, RTM bias also occurs when high collision frequency of a particular location is followed by a lower collision frequency, even though traffic condition remains same. Eventually, the collision frequency return to its long-term mean value as time goes by. As collision frequency fluctuates around its mean, RTM bias may occur when a site with high collision frequency is selected as black spot even. Because that high collision frequency is not the true mean of the site. This RTM bias may lead to a 'false' labelling of a site as hazardous. To overcome this problem, a statistical technique called Empirical Bayes (EB) method is used.

2.3.3 Empirical Bayes technique

The Empirical Bayes (EB) method is a statistical technique which is used in road safety analysis because of its ability to reduce RTM bias (Higle & Witkowski, 1988; Brude & Larsson, 1988; Hauer, 1992; Sayed et al., 1997; Sayed, 1998a). EB technique has the potential to identify black spots and also to measure the effectiveness of any countermeasure with higher degree of accuracy. The method not only considers the collision history of a particular location, but also considers the collision history of similar locations (the reference group). Thus, the EB approach combines two clues by means of a weighted average as given in the following equation (2.9) (Hauer et al., 2002a) as:

Estimate of expected collisions for a location = (weight) x collisions expected at similar locations

+ (1 – weight) x count of collisions at this location (2.9)

Here the value of 'weight' depends on the distribution assumed for observed and expected collisions. The value of 'weight' can be calculated using statistical methods which will be shown later. The main assumption of the EB technique is the Bayes theorem. According to the theorem, the conditional probability of the behaviour of an entity depends on two pieces of information: information of the entity itself and some engineering judgments on the underlying properties of the entity. Mathematically, the theorem can be stated as (Sayed, 1998a):

$$P(\emptyset|x) = \frac{P(x|\emptyset).P(\emptyset)}{\sum P(x|\emptyset).P(\emptyset)} \quad (2.10)$$

where,

Ø = a parameter;

P(Ø) = prior probability distribution of Ø events occurrence;

P(x|Ø) = observed distribution of the entity itself; and,

P(Ø|x) = the posterior distribution of Ø events expected to occur using prior and observed distribution.

The prior distribution if an entity is derived based on the properties and past experience of the entity (Sayed, 1998). The EB approach derives prior distribution based on the observed data of a reference group which consist of locations with similar properties. Higle & Witkowski (1988) set out the EB method to identify collision prone locations based on two assumptions. First assumption is the actual collision rate is treated as random variable and it follows Poisson distribution. So the observed probability distribution becomes:

$$P\{N_i = n|\overline{\lambda_\iota} = \lambda, V_i\} = \frac{(\lambda V_i)^n}{n!}.e^{-\lambda V_i} \quad (2.11)$$

where,

$\overline{\lambda_\iota}$= collision rate at location i;

N_i= number of collision at location 'i' during analysis period; and,

V_i= number of vehicle passing through the location 'i' for the analysis period.

The second assumption is the probability distribution of the 'reference group' collision rate, $f_R(\lambda)$, follows a gamma distribution with parameters α and β.

$$f_R(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\alpha} \quad (2.12)$$

So the first step is to estimate the value of α andβ. There are various methods (e.g. method of simple moment, method of moment estimates are available to calculate their values. Among them the method of simple moment (MSM) is the simplest method where α and β are chosen in such a way that the mean and variance associated with the gamma distribution are equal to the mean ($\bar{x}$) and variance (s^2) of the reference group. This can be done by making,

$$\bar{x} = \frac{\alpha}{\beta} \; and \; s^2 = \frac{\alpha}{\beta} \text{ , so that } \beta = \frac{\bar{x}}{s^2} \; and \; \alpha = \beta\bar{x} \quad (2.13)$$

The second step is to combine the 'reference group' probability distribution (the prior distribution) with the location specific collision observed rate to obtain the location specific probability density function (i.e. the posterior distribution), which is expressed as:

$$f_i(\lambda|N_{i,}V_i) = f(N_i|\lambda, V_i)f_R(\lambda) \quad (2.14)$$

The distribution of posterior distribution is also assumed to follow Gamma distribution and the value of α_i and β_i is calculated by

$$\alpha_i = \alpha + N_i \; and \; \beta_i = \beta + V_i \quad (2.15)$$

Thus, the resulting probability density function becomes

$$f(\lambda|N_i, V_i) = \frac{{\beta_i}^{\alpha}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta\lambda} \quad (2.16)$$

If the collision rate of the location, X_i, is greater than the observed mean of the reference group collision rate, X_R , then location 'i' is considered as collision prone. Mathematically, location 'i' is considered as collision prone if

$$P(\overline{\lambda_i} > X_R|N_i, V_i) > \delta \quad (2.17)$$

or if:

$$\left[1 - \int_0^{x_R} \frac{{\beta_i}^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i\lambda} d\lambda\right] > \delta \quad (2.18)$$

where,

δ= the desired level of confidence;

$X_R = \frac{\sum_{i=1}^{m} N_i}{\sum_{i=1}^{m} V_i}$; and,

N= the number of locations in the reference population.

Later Sayed & Abdelwahab (1997) recommended a small modification to Higle & Witkowski's method and the base of the modified method is collision correctability. This method focuses on those collisions which have correctable characteristics so that road safety agencies can focus on those collisions to reduce collisions. They introduced the concept of critical collision rate, λ_{c_i}, and formed the following equation:

$$\left[1-\int_0^{x_R}\frac{\beta_i^{\alpha+\lambda_{c_i}V_i}}{\Gamma\left(\alpha+\lambda_{c_i}V_i\right)}\lambda^{\alpha+\lambda_{c_i}V_i-1}e^{-\beta_i\lambda}d\lambda\right]>\delta \tag{2.19}$$

Solving equation (2.19) the value of critical rate (λ_{c_i}) is obtained. If ratio of the observed collision rate (λ_i) and the critical rate, (λ_{c_i}) is greater than one, then the location is considered as collision prone zone and vice versa.

Sayed, Navin & Abdelwahab (1997) proposed another modified black spot identification technique which is termed as 'countermeasure-based approach'. In the identification of black spots, this method does not follow traditional approach. Instead of looking for high collision frequency locations, this method identifies a location as collision prone if it exhibits well-defined collision patterns to be remedied by specific countermeasures. It assumes that a location having well defined collision pattern can be treated more effectively than a high collision frequency location with poorly defined collision patterns. This method starts with identifying collision patterns that can be targeted by countermeasures and later searching for locations that have an over representation of these patterns. Sayed, Navin & Abdelwahab (1997) defined the term 'over-representation' as a high likelihood of collision occurrence at a place.

Even though the EB method is widely used for road safety analysis, but the method described by Higle & Witkowski (1988) has two problems (Hauer, 1992; Sawalha, 2002). First, using a location's collision rate as a measure of its safety has been shown to have the potential for misleading results because of the non-linear relationship between traffic volume and road collisions (Hauer, 1995). Second, finding a suitable reference group is difficult to estimate the mean and variances of the prior distribution. To estimate mean and variance of the reference group, it requires a large sample set which is difficult to find. Without a suitable reference group reliable parameter estimation of the prior distribution is not possible. Also it is not possible to find reference locations which have exactly similar traits as the study location. The difficulty of finding reference locations increases with the increase of the number of traits. Great care is needed to select the members of the reference group and sometimes it is not possible to find exact reference group (Hauer, 1997). If the reference locations do not match with the study location then the result will not be reliable (Sawalha, 2002). This reference group selection related problem can be refined by the use of CPMs (Hauer, 1992; Sayed, 1998a).

2.3.4 Using collision prediction models

The problem of having a large reference group of similar characteristics can be overcome by multivariate regression technique (Hauer, 1992, 1995; Sawalha, 2002). The multivariate regression works with the development of CPMs. Even though the CPM variables are obtained from the reference group data, but the advantage is, the variables of the reference groups need not to be similar rather categorically similar (e.g. two-lane roads, signalized intersections, unsignalized intersections, etc.) (Hauer, 1992). So CPMs do not need a large reference groups with similar characteristics rather they express a relationship between the expected level of safety of an entity (the dependent variable) and a set of traits (explanatory variables). CPMs can be used to estimate the mean and variances of the imaginary reference group of each location (Hauer, 1992). Using the reference group data, CPMs are developed and location specific traits are entered into the CPM to predict the location specific mean of the prior distribution. Hauer et al. (1988) showed that the prior distribution follows gamma distribution with a mean E(Λ), and a variance:

$$Var\ E(\Lambda) = \frac{[E(\Lambda)]^2}{\kappa} \qquad (2.20)$$

The shape and scale parameters (i.e. α and β) can also be calculated as:

$$\alpha = \kappa \qquad (2.21)$$

$$\text{and } \beta = \frac{E(\Lambda)}{\kappa} \qquad (2.22)$$

Here ķ is a shape parameter which is derived during CPM development.

The next step is to use local collision data to refine the prior estimate of virtual reference group's data. Using Baye's theorem, this step combines location specific collision history and reference group's distribution to get a refined safety estimate of the location. The result of the refinement is a posterior distribution. The mean of the resulting Posterior distribution is termed as the empirical Bayes safety estimate for location 'i', (EB_i). The posterior distribution is also gamma distributed with shape and scale parameters as given below (Hauer et al 1988, 2002a; Kulmala 1995)

$$\alpha = \kappa + count \tag{2.23}$$

$$\beta = \frac{\kappa + E(\Lambda)}{E(\Lambda)} \tag{2.24}$$

The mean (EB_i), and the variance, Var (EB_i) of the posterior distribution is calculated as (Hauer, 1992):

$$EB_i = E(\Lambda | Y = count) = \frac{\alpha}{\beta} = \left[\frac{E(\Lambda_i)}{\kappa + E(\Lambda_i)}\right](\kappa + count) \tag{2.25}$$

$$Var\ (EB_i) = var(\Lambda | Y = count) = \frac{\alpha}{\beta^2} = \left[\frac{E(\Lambda_i)}{\kappa + E(\Lambda_i)}\right]^2 (\kappa + count) \tag{2.26}$$

Equation (2.26) can be written in the form of equation (2.9) (Sawalha & Sayed, 2005a):

$$EB_i = weight.E(\Lambda_i) + (1 - weight).count \tag{2.27}$$

where,

count= observed collision frequency of the study location;

weight= $\frac{\kappa}{\kappa + E(\Lambda_i)} = \frac{1}{1 + \frac{E(\Lambda_i)}{\kappa}} = \frac{1}{1 + \frac{Var(E(\Lambda_i))}{E(\Lambda_i)}}$; and,

Var $(E(\Lambda_i)) = \frac{E(\Lambda_i)^2}{\kappa}$ = the variance of the CPM estimate.

The value of weight increases with higher value of κ, as a result of that the EB estimate of the location will be closer to CPM prediction and vice versa. Hauer (1997) noted that if the predicted and observed times are not equal, then some modifications should be applied. He suggested the use of a ratio (r) between observed time and prediction time to. Thus, weight becomes:

$$weight = \frac{1}{1 + r.\frac{Var(E(\Lambda_i))}{E(\Lambda_i)}} \tag{2.28}$$

It has been shown that the EB method is effective in reducing RTM bias and for that reason the EB method is used for identification of collision prone locations.

2.3.5 Identification and ranking of collision prone locations using CPMs

The EB method can be used in identification of collision prone locations. The steps needed for CPL identification is set out by Sayed (1998a) and Sawalha & Sayed (1999). The identification can be done in three steps. In step one, with the development of CPM, the prior distribution of the imaginary reference population is defined with expected mean collision frequency and variance using equation 2.21 and 2.22. The mean of the prior distribution can be found from CPM prediction, $E(\Lambda)$ or the 50th percentile (P_{50}) of the Prior distribution's probability density function. In step two, the observed collision frequency (count) of the location is used, together with $E(\Lambda)$ and κ, to define gamma distributed posterior distribution. The shape and scale parameters of the distribution can be calculated using equation 2.23 and 2.24. In the last step the EB estimate of the location (EB_i) is calculated using 2.25 and the value of EB_i is compared to the regional average ($E(\Lambda)$ or P_{50}) calculated from the prior distribution of imaginary reference group. If for a significant probability, δ (usually not less than 0.95), the EB estimate of the location exceeds the average of the reference group then the location is considered as collision prone location (CPL). It is usually done by integrating the posterior probability density function in the range from *0* to $E(\Lambda)$, to determine with what level of probability EB_i and $E(\Lambda)$ values are different. Mathematically, the location is identified as collision prone if the following condition is met:

$$1-\int_0^{E(\Lambda)} f_{EB}(\lambda)d\lambda = \left[1-\int_0^{E(\Lambda)} \frac{\left[\frac{\kappa}{E(\Lambda)}+1\right]^{(\kappa+count)} \lambda^{(\kappa+count-1)} e^{-\left[\frac{\kappa}{E(\Lambda)}+1\right]\lambda}}{\Gamma(\kappa+count)} d\lambda\right] \geq \delta \qquad (2.29)$$

Figure 2.4 graphically shows the method of identifying CPL at a defined confidence level.

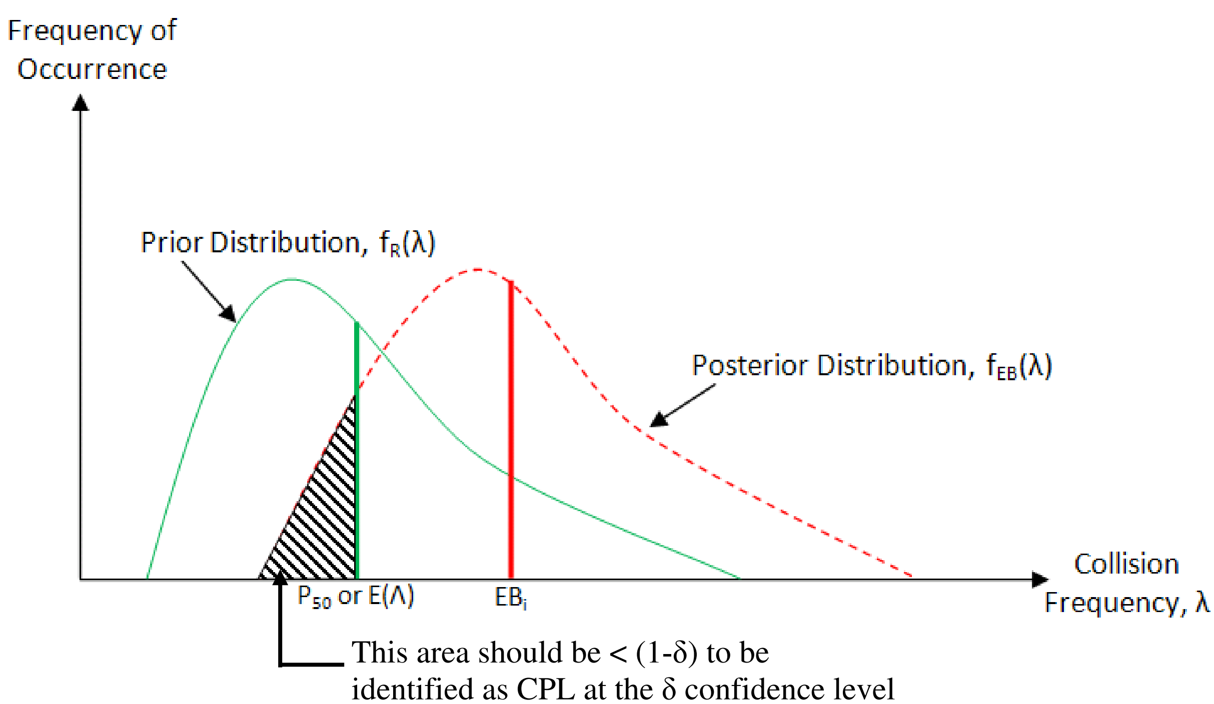


Figure 2.4 Empirical Bayes identification of collision prone location

After the identification of collision prone location, the next step is to rank those CPLs to ensure the locations in highest need of improvement are given more attention. Traditionally, ranking criteria use the difference between the observed and the expected collision frequency. This method is called Potential Collision Reduction (PCR) method.

$$PCR = EB - E(\Lambda) \tag{2.30}$$

PCR ranking method focuses on the highest potential for collision reduction. But this method has two weaknesses. The first weakness is the use of the mean value of expected collision, E (Λ). As mean implies that up to half of the 'similar' sites would have lower mean collision frequencies than the expected mean, the PCR value would be greater for lower collision frequency locations and these locations would be ranked higher. The second weakness of the PCR criterion is that it tends to ignore low collision frequency locations that have experienced significant collision increases. So another method named Collision Risk Ratio (CRR) is used which can address the weakness of PCR method by ranking the locations based on the ratio of observed to expected collisions.

$$CRR = \frac{EB}{E(\Lambda)} \tag{2.31}$$

A third method is proposed that uses collision modification factors (CMFs) to rank CPLs based on the anticipated safety benefits as a outcome of countermeasures (Hauer et al., 2002b):

$$\text{Anticipated safety benefit} = \text{observed (RTM corrected) collision frequency} \cdot CMF$$

$$= EB_i . CMF \tag{2.32}$$

This method also has some limitations as all of the identified black spots are not similar and they do not require similar types of countermeasure. Moreover, different countermeasures have differing CMFs. So, this criterion requires some prior knowledge of each site and its possible countermeasure (and CMF) to ensure an accurate ranking. But it is not possible to anticipate the possible countermeasure and associated CMF. That is why this method is not widely used; rather PCR and CRR methods are used. Instead of using PCR and CRR separately, by summing the ranks obtained from PCR and CRR, a combined ranking for each zone is prepared sometimes. The zone(s) with the higher ranking score are recommended for diagnosis and remedy.

2.3.6 Diagnosis and remedy of black spots

After the identification and ranking of black spots, the next step is to diagnose those locations to find the possible cause behind collisions. The diagnosis step is done in two steps. First, the collision history of the location is analyzed to identify the clusters of particular collision types. Second, physical and operational condition of the location is analyzed, including observations of driver characteristics and consultation with local road agencies. Based on the data gathered in two steps, predominant types of collisions and possible causes behind them are identified.

Once the causes have been identified, the next step is to identify possible countermeasures (remedies) to improve road safety condition of the location. Often, more than one possible countermeasure is identified for a location and economic analysis is conducted to find the feasibility of each alternative. The most popular economic analysis method involves a benefit-cost (B/C) analysis of each candidate countermeasure (Sayed, 1998a). While the costs are related to installation and operation of the countermeasure, benefits are measured in terms of reductions in: travel time, delay, collision frequency, and collision severity. If the B/C value is greater than one, then the benefit is greater than the associated cost, indicating the countermeasure is economically beneficial. If there is more than one countermeasure, the countermeasure having highest B/C is selected to maximize the economic benefits.

Reactive RSIP is the traditional approach that focuses on the reduction of collisions to improve safety. Even though the reactive approach has been proven to be very successful, it has some major drawbacks. This approach requires several years of collision data to identify black spots before the application of any countermeasure to improve road safety. As it takes lots of lives and damages properties before any action is taken, it makes a huge impact to society. Moreover, retrofitting countermeasures at identified black spots is usually costly. Therefore, road safety authorities and researchers have emphasized on pursuing more proactive engineering approaches to prevent black spots from occurring in the first place.

2.4 Proactive approach to road safety

It has been found that the driver, vehicle, and the road environment are the main three components of collision occurrence and collision can occur due one component or a combination of them. Sayed et al. (1995) found that 96% collisions are involved with driver's error. So it is very important to adopt such strategies that focus on minimizing driver's error. de Leur & Sayed (2003) identified the most effective and long lasting road safety engineering strategies as that proactively focus on engineering safer road environments to reduce exposure and collision risk of drivers and their passengers.

The proactive engineering approach to road safety improvement focuses on predicting and improving the safety of planned facilities rather than improving the safety of existing facilities (de Leur & Sayed, 2003). The aim of the approach is to evaluate safety throughout every stage of planning a new facility so that the safety risk can be minimized by precluding black spots from occurring at all. If road safety is explicitly addressed before a project is built, it reduces the number and cost of reactive safety countermeasures that have to be retrofitted into existing communities. There has been much work done on identification of the shortcomings of the reactive approach versus the need for a more proactive engineering approach to road safety (van Schagen & Janssen, 2000; de Leur & Sayed, 2003) and all of these studies suggest that adoption of proactive approach can reduce the social and economic burden of road collisions. When road safety is explicitly evaluated throughout the planning and design process, enhanced effectiveness and sustainability can occur before the drivers are exposed to the road system (Lovegrove, 2007).

Although there exist sufficient reliable empirical tools for traditional reactive approach, proactive approach is at a relatively early stage of their development and implementation (de Leur & Sayed, 2003; Lovegrove, 2007). de Leur and Sayed (2003) identified the main obstacle of using proactive approach as the lack of (1) opportunity within traditional planning process to proactively consider road safety; (2) proper methodology and reliable tools; and, (3) systematic approach and framework to evaluate road safety in a proactive manner. de Leur and Sayed (2003) highlighted the need for a guideline and framework to consider road safety throughout the design and planning process. They proposed a proactive road safety planning framework which is given below:

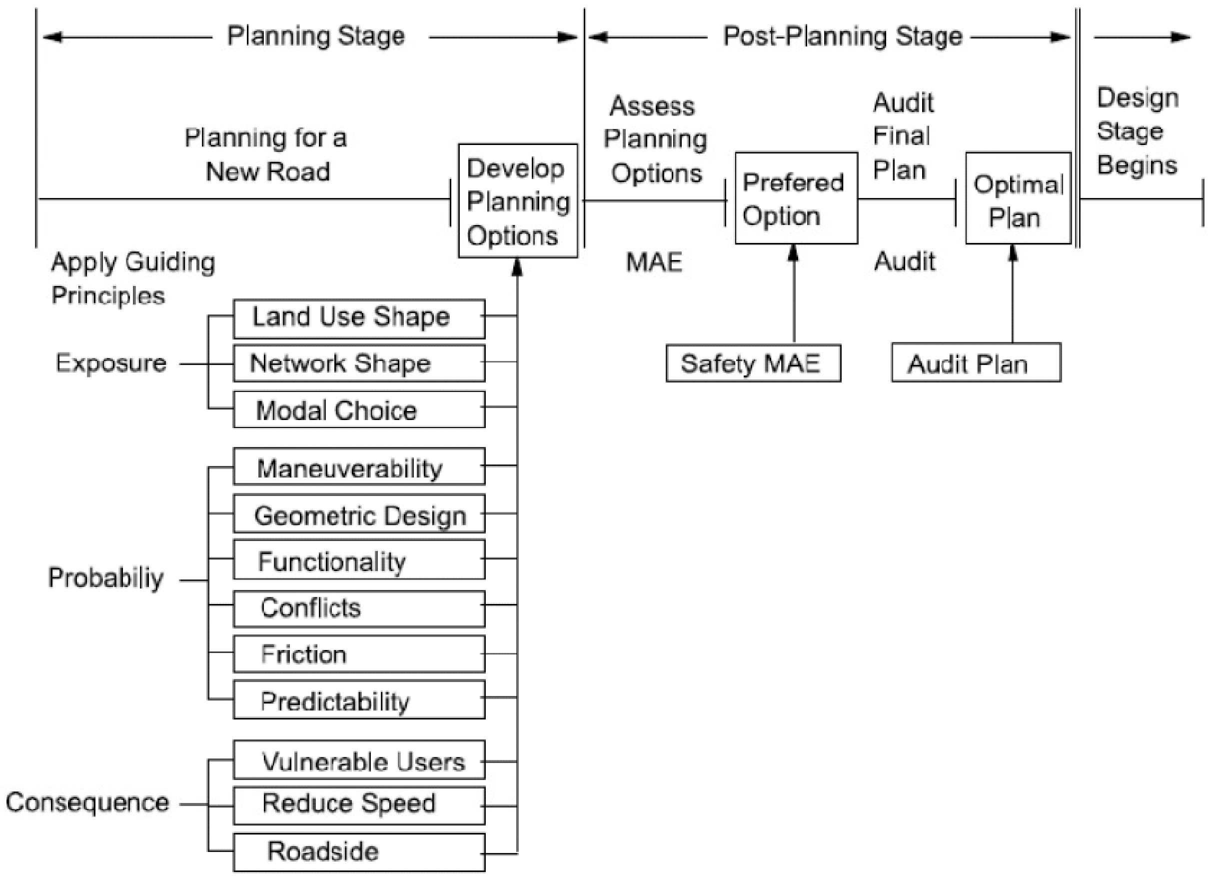


Figure 2.5 Proactive road safety planning framework (de Leur & Sayed, 2003)

The main goal of proactive approach is to consider road safety before the road users are exposed to the system. Different types of proactive techniques are being used. Sayed et al. (2010) summarizes the trend of different proactive engineering initiatives such as the use of safety conscious planning models, the explicit evaluation of safety in road design, and the automated safety analysis of video data. Following sub-sections will discuss two proactive empirical tools: (1) road safety audits; and, (2) combination of CPMs and regional transportation planning models.

2.4.1 Road safety audits

The internationally accepted definition of road safety audit is "Road safety audit is a formal and independent safety performance review of a road transportation project by an experienced team of safety specialists, addressing the safety of all road users" (Ho et al, 2011). Road Safety Audits (RSAs) can be applied either in reaction to an existing identified hazardous location, or proactively as a part of road planning and design exercise. Audits can be performed for a small project (i.e. single intersection) as well as for a large project (i.e. regional planning). RSA has the ability to provide an explicit, formalized safety evaluation of road projects of any size. RSA has been shown to be the most beneficial and easiest part to be integrated in an agency's existing safety program (Wilson, 2004). While the cost to conduct an RSA has been estimated at 0.2 – 0.5% of total project costs, the benefits are huge having a first year rate of return in the range of 120% to 146%, and a B/C ratio of 36:1 (ARRB, 1999; Jordan, 2001, Wilson, 2004). Austroads described an analysis of nine audit sites with 250 different design stages and the results showed that RSA associated benefit–cost ratios ranged from 3:1 to 242:1 while for existing roads, benefit–cost ratios ranged from 2:1 to 84:1 (Wilson, 2004). Moreover, RSAs are recommended as a part of all regional and community-wide planning exercises (Wegman, 1996; de Leur & Sayed, 2003; Hadayeghi et al., 2003).

2.4.2 Use of collision prediction models

CPMs can be used in proactive road safety programs. CPMs can be combined with transportation planning model to estimate the expected outcome of road safety of the plan. CPMs can be of two types: micro-level CPMs and macro-level CPMs. Micro-level CPMs works at a particular intersection or at a small segment of a road/highway while macro-level CPMs works at a larger scale such as zone, neighbourhood or community. Both of these models have been used as an empirical tool to do proactive, planning level analysis. Ho and Guarnasche (1998) combined micro-level CPMs with Greater Vancouver Regional District (GVRD) Emme/2 regional transportation planning model to evaluate road safety of the planning. They used traffic volume of major and minor roads as well as attributes of different facilities (e.g. types of intersections, number of lanes and road segment types) as input to the CPMs. The result of the analysis was not expected and they recommended to use refined CPMs for non-standard intersections and also refined forecasted traffic volume when comparing a base and future scenario. Later Lord & Persaud (2004) attempted to integrate refined micro-level CPMs with an Emme/2 transportation planning model of the Toronto area. The refined CPMs included different road and intersection configurations (e.g. 4 lane arterials, 3-way intersections), and collision categories (e.g. PDO, injury, and fatal). They concluded that use of micro-level CPMs in safety planning is cumbersome and the assumption of fixing the characteristics of all other intersection except the study intersection is not necessarily correct. These two studies highlighted the lack of micro-level CPMs to fill the gap between what is needed and what is available in terms of reliable proactive safety planning tools. These studies also suggested the use of macro-level CPMs as an empirical tool to pursue proactive road safety.

One of the main limitations of using micro-level CPMs in regional safety evaluation is to reliably predict traffic volumes on inner local roads of a neighbourhood. To overcome this problem Dutch researchers (Poppe, 1995; 1997a) proposed the use of neighbourhood traits such as population and area in lieu of traffic volumes on inner roads, which has lead to the development of community-based, macro-level CPMs. Hadayeghi et al. (2003) and Lardon de Guevara et al. (2004) developed macro-level CPMs that statistically significantly related collision frequency with population density, employment, intersection density, and road types. However, neither study attempted to use the CPMs in a road safety planning evaluation. In 2006, using a refined methodology, Lovegrove and Sayed (2006a) developed 47 macro-level CPMs for the Greater Vancouver Regional Districts (GVRD) using more than 200 input variables. Instead of looking at a specific micro-level location, they developed macro-level CPMs for relatively large areas (i.e. neighbourhoods). It allows the models to be applied at a larger scale which is more compatible with the planning process.

Lovegrove & Sayed (2006b) combined macro-level CPMs with road safety planning. They used CPMs to evaluate the road safety of four different neighbourhood road patterns (grid, cul-de-sac, fused and two other theoretical patterns), and of area-wide traffic calming techniques. Lovegrove & Sayed (2007) also demonstrated the application of macro-level CPMs to identify black spots and to recommend possible remedial countermeasures. Thus macro-level CPMs can be used to enhance traditional road safety methods, and to provide early-warning to road safety engineers.

Lovegrove et al. (2010) also demonstrated the use of the models to evaluate the road safety effects of a regional transportation plan. They used macro-level CPMs with a 3-year GVRD regional transportation improvement plan and proved that the new plan will reduce the expected collision frequency compared to a do-nothing scenario. They recommendations the practical use of macro-level CPMs as an improved empirical road safety evaluation tool for both existing and planned neighbourhoods.

It is important to mention that Khondaker et al. (2009) developed collision prediction models for the RDCO. But the data quality and results were not good because it used the output from an old transportation planning model. It also suffers from the problem of miscoding some important urban/rural areas. So the author successfully transferred macro level CPMs from GVRD (1996) to the City of Kelowna (2003) to predict number of collisions.

2.5 Development of community based macro-level collision prediction models

The need for macro-level collision prediction models in proactive road safety application have been demonstrated by several road safety researchers and agencies. Community-based, macro-level CPMs develop statistically significant relations between collision frequency and zonal (i.e. neighbourhood) traits. The main difference between micro-level and macro-level CPMs is instead of looking at a single intersection or a road segment, macro-level CPMs use multiple intersections and road segments across a zone. The advantage of macro-level CPM is it doesn't need accurate forecast of traffic volume of each intersection and road segment of the region.

Several researchers attempted to develop macro-level CPMs. Levine et al. (1995) developed a model that related roads collisions to zonal population, employment and road characteristics. But this model was based on linear regression which is not appropriate to represent the non-negative, non-normal, non-linear nature of collisions. Fotheringham (2000) used a Geographically Weighted Regression (GWR) model with some explanatory variables such as population and employment, to analyze the spatial variability of collisions between zones, with inconclusive results. Hadayeghi et al. (2003), following the same GWR methodology found similar inconclusive results. Using geo-statistical techniques, Kim & Yamashita (2002) tried to relate police collision data with land use categories and ended up without success. Even though geo-statistical techniques appear to hold promise, they are not successfully used in macro-level CPMs development.

Hadayeghi et al. (2003) developed a series of macro-level collision prediction models to predict the mean collision frequency for each zone by aggregating data across 463 traffic zones in Toronto, Canada. He used generalized linear modeling (GLIM) approach to predict collisions (fatal and non-fatal collisions) as a function of socio-demographic, traffic demand, and network data variables using Negative Binomial (NB) regression. Instead of using individual link or node data, he built macro-level CPMs using aggregated node and link data across each zone. Explanatory variables used in the research, included vehicle-kilometres travelled (VKT), arterial road lane-kms, number of households, area, posted speed, average zonal congestion, intersection density, total employed labour force, and total minor road kilometres. Several other possible explanatory variables were explored, including different employment sectors, land uses, neighbourhood geometry, driver age, gender, road conditions, collision reporting practices, and police enforcement levels. The following CPM in equation 2.33 was used to predict expected number of collisions.

$$E(\Lambda) = a_0 VKT^{b_0} e^{\sum b_1 x_1} \tag{2.33}$$

where,

E(Λ)=mean collision frequency;

VKT= zonal total forecasted VKT from Emme/2;

x_i= zonally aggregated explanatory variables (e.g. population, employment); and,

b_0,b_1, b_i = constants (model parameters)

Lardon de Guevara et al. (2004) created planning-level collision prediction models for Tucson, Arizona. Assuming NB error structure and using log-linear transformation technique, they develop macro-level CPMs assuming non-linear, exponential function which is as follows:

$$E(\Lambda) = e^{\sum b_i x_i} \tag{2.34}$$

where,

E(Λ)= collisions per two years;

x_i= independent variables; and,

b_i = constants (model parameters)

Hadayeghi et al. (2007) updated and improved the previous research (Hadayeghi et al., 2003) using collision, socioeconomic, demographic, road network, and traffic volumes data for the city of Toronto’s 481 traffic zones. Specific land uses, types of employment and the presence of transit facilities were new variables considered in that study. The model form was similar to equation (2.33) using NB error structure and generalized linear modeling (GLM) procedure.

The CPM should consider the effect of zero traffic on road. In case of equation (2.34) even if the traffic volume is zero, it may predict some collision to occur. Sawalha & Sayed (1999) noted that the model must yield logical results. Specifically, there must be zero risk of collision with zero exposure. Hauer et al. (1988) related collisions to exposure via the product of traffic flows raised to some power. In addition to traffic flow, Miaou (1996), and Sawalha & Sayed (1999) have shown that other variables such as geometric features of road, traffic controls system etc. can affect collision pattern. Based on empirical case studies, they suggest that the proper CPM form should consist of exposure measure(s) (raised to some power) multiplied by an exponential function incorporating other explanatory variables. Following a comprehensive literature review and data extraction process, Lovegrove (2007) developed and presented forty-seven macro-level collision prediction models, each significantly associated with one or more of twenty-nine variables identified using refined GLM methods. The general model form used is given in equation 2.35 which meets the principle 'zero-exposure = zero collision frequency'.

$$E(\Lambda) = a_0 Z^{a_1} e^{b_i x_i} \tag{2.35}$$

where,

E(Λ)= predicted collision frequency (over 3years);

VKT= external exposure variables (e.g. VKT, TLKM);

x_i= explanatory variables (e.g. population, employment); and,

a_0, a_1, a_i = constants (model parameters)

Lovegrove & Sayed (2006a) also developed community-based, macro-level CPMs for the Greater Vancouver Regional District (GVRD). They categorized all the input variables in four different groups and developed 16 different CPM groups which are shown in Table 2.1. The basis of different groups came from:

- Four themes of neighbourhood traits (traffic exposure, road network, socio-demographics, and TDM);
- Two classes of land use (rural or urban); and,
- Two sources of exposure data derivations (modeled or measured), whereas "modeled" exposure variables are output from transportation planning models and "measured" exposure variables are derived from geo-referenced mapping.

Table 2.1 Collision prediction model groups

Themes	Land Use	Derivation	Group #
Exposure	Urban	Modelled	1
		Measures	2
	Rural	Modelled	3
		Measures	4
Socio-Demographic	Urban	Modelled	5
		Measures	6
	Rural	Modelled	7
		Measures	8
Transportation Demand Management (TDM)	Urban	Modelled	9
		Measures	10
	Rural	Modelled	11
		Measures	12
Road Network	Urban	Modelled	13
		Measures	14
	Rural	Modelled	15
		Measures	16

Lovegrove & Sayed (2006a, 2007) also proposed guidelines to use macro level CPMs and these guidelines have been proven to have the potential to act as empirical tool to pursue proactive road safety. The following subsections will describe about regression technique and model formulation of macro-level CPMs.

2.5.1 Regression techniques

Two statistical modeling methods have been used mainly for estimating CPM parameters: conventional linear regression modeling, and, generalized linear regression modeling (GLM). Traditional approach uses linear regression, assuming a Normal (Gaussian) distribution error structure. However, it has been found that the normal distribution of error terms for modeling road collisions is not appropriate (Sayed & de Leur, 2001a). The non-normal error distributions are found to better describe the unexplained random variations in road collisions (Kulmala, 1995; Miaou, 1996). The GLM method predicts collisions that fit better with the observed collision data. GLM has the capability to be used for a wide range of probability distribution such as Normal, Poisson, Negative Binomial, Gamma error structure. More recently, Hadayeghi et al. (2010) investigated the local spatial variations in the relationship between the number of zonal collisions and potential transportation planning predictors through the development of Geographically Weighted Poisson Regression (GWPR). Hadayeghi et al. (2010) also showed the use of Full-Bayesian Semiparametric Additive (FBSA) and Geographically Weighted Poisson Regression (GWPR) methods can result in better prediction than GLM method.

2.5.1.1 The GLM Process

There are several GLM statistical software packages (e.g. GLIM4, SAS, GenStat) that can be chosen to do the regression analysis (Dobson, 1990). They can be used to model data that follow a wide range of probability distributions. The random nature of collision is explained by probability distribution of error term around the mean. Generally Poisson and Negative Binomial (NB) error structure are the most commonly used distribution in GLM. In Poisson structure the mean and the variance are assumed equal which makes the calculation easier. But the advantage itself is also a limitation of Poisson structure as most collision data are over-dispersed (the variance is greater than the mean) (Kulmala & Roine, 1988; Kulmala, 1995). Miaou & Lum (1993) identified there possible sources of over-dispersion in collision. They sources are: (1) CPMs cannot contain all the variables that explain collision occurrence; (2) there are uncertainties in vehicle exposure data and traffic variables; and, (3) collision data comes from a non-homogenous environment. For the over-dispersion characteristics of collisions, NB structure better represents the error structure than Poisson distribution.

The decision on whether to use Poisson or NB structure is based on a methodology proposed by Bonnenson & McCoy (1993). First, Poisson error structure is assumed and the assumption is verified by calculating a dispersion factor (σ_d). If σ_d exceeds the value of 1.0, then the data have greater dispersion and Poisson distribute assumption is incorrect. So NB error distribution should be used.

$$\sigma_d = \frac{Pearson\,\chi^2}{n-p} \tag{2.36}$$

where,

Pearson $\chi^2 = \sum_{i=1}^{n} \frac{[y_i - E(\Lambda_i)]^2}{Var\,(y_i)}$;

N = number of locations;

y_i = observed mean collision frequency at location ‘i’ over a specified time period;

Var (y_i) = variance of the observed mean collision frequency at location ‘i’;

$E(\Lambda_i)$ = expected mean collision frequency for location 'i' obtained from CPM; and,

P = number of CPM parameters.

Once an error distribution, model form, and link function is specified, the multivariate regression software can estimate the value of model parameters, and the value of the overdispersion or shape parameter, **κ,** using one of three methods: the Maximum Likelihood (MLE) method, the expected value of the χ^2 statistic method, and the mean Scaled Deviance method(Lawless, 1987). However, MLE method has been shown to provide better accuracy compared to other methods (Sawalha & Sayed, 2005a; Miaou, 1996). The next step is to select explanatory variables to be used in CPMs.

2.5.2 Selection of explanatory variables

There are different factors behind road collisions and it is important to identify those influencing factors so that road collisions can be predicted with higher accuracy. Sawalha & Sayed (2005a) recommended forward stepwise procedure to add independent variables in model form. In this procedure, variables are added to the model one by one and tested for significance, starting with the exposure variables. Each time a variable is added to the model, the addition of the variable is evaluated based on several criteria. First, its logic (i.e. +/- sign) should be assessed as to whether it meets with intuitive expectations or not. Second, the t-ratio of the added variable's estimated coefficient must be significant at the 95% confidence level. Last, the addition of the variable to the model should cause a significant drop in the CPM's Scaled Deviance (SD) at 95% confidence level. If the error structure is Poisson distributed, the SD is defined as:

$$SD = 2\sum_{i=1}^{n} y_i ln\left(\frac{y_i}{E(\Lambda_i)}\right) \tag{2.37}$$

and if the error structure follows a NB structure, SD is defined as (McCullagh and Nelder, 1989):

$$SD = 2\sum_{i=1}^{n}\left[y_i ln\left(\frac{y_i}{E(\Lambda_i)}\right) - (y_i + \kappa) ln\left(\frac{y_i+\kappa}{E(\Lambda_i)+\kappa}\right)\right] \tag{2.38}$$

If a drop of least $\chi^2_{0.05,1} = 3.84$ is observed then the added variable significantly enhances CPM predictive accuracy, and it is not correlated with other independent variables of the model (Sawalha & Sayed, 2005a). In the next step another variable is added to model and the criteria are evaluated for the newly added variables. Addition of variables continues unless all desired variables are tested.

Lovegrove & Sayed (2006a) identified the following variables that positively related to collision frequency:

- Exposure variables: vehicle kilometres travelled (VKT), total road lane kilometres (TLKM), and average congestion (VC).
- Socio-Demographic variables: job density (WKGD), population density (POPD), unemployment (UNEMP), residential unit density (NHD).
- TDM variables: shortcut capacity/attractiveness (SCC, SCVC), number of drivers (DRIVE), total commuters (TCM), total commuter density (TCD). SCC measures the ability of vehicles to shortcut on local roads through a particular zone, in vehicles per hour. SCVC is simply SCC multiplied by congestion level VC.
- Network variables: signal density (SIGD), intersection density per unit area (INTD), intersection density per lane-km (INTKD), arterial-local intersection percent (IALP), arterial lane kilometre percent (ALKP).

Lovegrove & Sayed (2006a) also identified variables that are negatively related to road collision frequency:

- Socio-Demographic variable: family size (FS).
- TDM variables: core size and percentage (CORE, CRP).
- Network variables: 3-way intersection percent (I3WP), local road lane-kilometre percent (LLKP), and Core area (CORE).

2.5.3 Goodness of fit

After the selection and addition of explanatory variables, it is necessary to test how well the model fits the observed data. So a goodness of fit test is necessary. Various researchers (Sawalha & Sayed,1999; Sayed & de Leur, 2001b; Sawalha & Sayed, 2005a) described the methodology for testing the goodness of fit for CPMs. One criteria is, the SD, Pearson χ^2 statistics value should be less than the χ^2 distribution value with (n - p - 1) degrees of freedom at 95% confidence level. Also a plot can be done between the Pearson Residuals (PR_i) and the predicted collisions for each location. PR is defined as the difference between the predicted and the observed collision frequency divided by the standard deviation (NAG, 1994).

$$PR_i = \frac{E(\Lambda_i) - y_i}{\sqrt{Var(y_i)}} \tag{2.39}$$

In a well-fit model PR_i values should cluster around zero over the full range of predictions (Bonnenson & McCoy, 1997). Another measure of model goodness of fit is to plot the average of squared residuals (SR) versus the predicted collision frequency. The average of squared residuals is defined as:

$$Average\ of\ SR = \frac{\sum_{i=1}^{n}(E(\Lambda) - y_i)^2}{n} \tag{2.40}$$

where n is the number of locations in each group (e.g. five locations at a time). This is typically done after ranking the locations in order of predicted collision frequency and by plotting the averages of predicted collisions (taken in groups) versus the averages of squared residuals (taken in groups) over the full set of observation (Sawalha & Sayed, 2001). For a well fit model, the points should cluster about the variance function line for a negative binomial error distribution. The variance function line is drawn using the following function:

$$Var(y_i) = E(y_i) + \frac{E(y_i)^2}{\kappa} \tag{2.41}$$

2.5.4 Outlier analysis

During the assessment of model goodness of fit, it may not meet the expectations. Lack of model fitness is often related to data quality problems. Outlier analysis is performed to identify and remove unusual or extreme observations that are not typical of the rest of the data. Normally outliers are the results of error in data collection and recording. Sayed & Rodriguez (1999) and Sawalha & Sayed (2005b) described the use of the Cook's Distance (CD) measure for outlier analysis.

$$CD_i = \frac{h_i}{p(1-h_i)}(r_i^{ps})^2 \tag{2.42}$$

where,

h_i= leverage value;

r_i^{ps}= is the standardized resudial of point 'i' = $\frac{PR_i}{\sqrt{1-h_i}}$; and,

p = number of parameters.

The higher value of CD_i indicates the stronger influence on the model. Sawalha & Sayed (2005b) suggested sorting the observations in descending order of CD and removing points with the largest CD value. After removal of each point, the model is re-run with κ fixed at its previous value to observe the change in goodness of fit. If the SD change is greater than ${\chi^2}_{0.05,1} = 3.84$, the GLM software is re-run to estimate all parameters, including a new κ and a new CD for each of the remaining point. Sawalha & Sayed (2005b) recommended monitoring the t-statistic of each variable to ensure the significance of the variables. This stepwise outlier analysis is repeated until the change in SD becomes less than 3.84.

2.5.5 Guideline for selecting appropriate model

Collision prediction models can be developed based on several criteria such as the type of collisions (e.g. fatal, injury), the time period (e.g. AM, PM, midday) and the traits of variables (exposure, S-D, network, TDM). For a particular objective it is very important to select the appropriate model to be used. A six-step selection process has been recommended by Lovegrove (2007) which is given below:

- The first step is to choose the type of CPM (i.e. micro- or macro-level) based on the scope of the safety evaluation.
- The second step is to select model by considering the safety application task (i.e. reactive or proactive safety analysis).
- The third step looks at the predominant type of land use in each neighbourhood under evaluation and select which land use type (i.e. urban or rural) will be used.
- The fourth step is involved in selecting trigger variables. Trigger variables are those variables which can significantly change the collision frequency and collision pattern.
- The fifth step identifies which type of collision (i.e. total, severe, AM, PM, nonrush, and/or pedestrian) are of interest in the safety evaluation.
- The sixth and final step is to check the source of data to use in the selected models. Adequate data ensures that the selected models will provide accurate and credible results.

2.5.6 CPM prediction

Community-based macro-level collision prediction models can predict the number of collisions based on some input variables. As it is a prediction, it could be associated with some prediction errors. CPM predictions are based on some random variables and the result can vary over a wide range. The variance of collisions prediction is described as (Lovegrove, 2007):

$$Var[E(\Lambda)] = \frac{[E(\Lambda)]^2}{\kappa} \quad (2.43)$$

where,

E(Λ)= predicted collision frequency (over 3years); and,

κ = shape parameter.

If the predicted collision frequency is used for comparing future scenarios, then only a relative comparison of the future scenarios is desired. To do that, CPMs are used to predict the future number of collisions for each TAZ and the number of collisions is typically summed across all zones of influence to provide an overall collision estimate for each scenario. The difference in the two sums between two scenarios is evaluated to find whether the difference is statistically significant or not for a desired level of confidence interval. The following equation (2.44) can be used for the evaluation purpose (Lovegrove, 2007):

$$\left| \frac{E(\Lambda_1)-E(\Lambda_2)}{\frac{\sqrt{Var[E(\Lambda)]}}{\sqrt{n}}} \right| \geq T\left[\frac{(1-\alpha)}{2}, n-1\right] \quad (2.44)$$

Where, n= the number of zones of influence used in deriving the sums;

$E(\Lambda_1)$ *and* $E(\Lambda_{2)}$= the sumes of zonal predictions for the scenarios being compared;

Var [$E(\Lambda)$] = the greater of Var [$E(\Lambda_1)$] or Var [$E(\Lambda_{2)}$];

α= the desired level of confidence typically 95%; and,

T= the T-statistics =1.96 @ 95% confidence for large n

2.6 Summary

Transportation planning produces useful information for decision makers so that it helps to make transport policy, program and investment decisions. Traditional transportation planning model includes four steps namely: trip generation, trip distribution, mode choice and trip assignment. While the first step determines the number of trips generated from an area, the second step determines the destination of each trip. In step three and four, trips are assigned to road network for each mode available in the planning context. Section 2.2 and 2.3 presented a comprehensive review of each step involved in transportation planning. The next Section 2.4 discussed about road safety programs.

Road collisions have become one of the greatest social and economical burdens to society. Each nation is suffering from road safety burden and to reduce road collision burden, road safety agencies have implemented various Road Safety Improvement Programs (RSIPs). The traditional RSIP is reactive approach which works in reaction to existing collision history. Identification and ranking of CPLs is an important part of reactive road safety programs. Diagnoses of collisions and recommendation of countermeasures are the next two steps when CPLs are identified. However, reactive approach requires a pre-existing collision history, treats a limited set of black spots based on available funding, and involves costly retrofits in existing communities. As a result, researchers and road safety agencies are perusing more proactive approach.

Proactive approach is an approach that focuses on predicting and improving the safety of a new planned facility so that the safety risk can be minimized before road users are exposed to it. Combination of collision prediction model (CPM) and regional planning model has proven to be an improved empirical tool to evaluate the road safety of a regional planning. Macro-level CPM can play a vital role in that safety evaluation process. This chapter discussed about the statistical issues, development, and application of CPM in proactive road safety program.

CHAPTER 3 DATABASE DESCRIPTION AND METHODOLOGY

3.1 Introduction

This chapter consists of four main sections. In section 3.2, the data sources of collisions, socio-demographic variables and road networks are given. The next section 3.3 describes the methodology of data extraction, including aggregation unit, aggregation approach and different variables definitions. In section 3.4, the four step transportation planning model development methodology is described, including method used and calibration process. Section 3.5 describes the methodology for collision prediction model development, including information on regression technique, model form, goodness of fit and outlier analysis. Section 3.6 describes the methodology for identifying collision prone locations. Finally Section 3.7 summarizes the data extraction and methodology of the research.

3.2 Data sources

The data used for this research came from several sources. The following sections will describe the data source for each type of variable.

3.2.1 Collision variables

Collisions data was obtained from Insurance Corporation of British Columbia (ICBC). ICBC handles most auto insurance collision claims in British Columbia (BC), and it is the central warehouse of collision database. They provided collisions data for the period of 2004 to 2006. The data is geo-coded and it includes details of each collision. The attribute file associated with ‘shapefile’ includes the severity of collisions as well as spatial and temporal description of the collisions. The availability of ICBC geo-coded claim data is a great advance in road safety analysis as it overcomes many of the traditional unreported, unattended, and/or incomplete (i.e. self-reported) collision data problems (Lovegrove, 2007). The given collision dataset defines three types of collisions: fatal, injured and PDO; whereas fatal collision as ‘F’; injury collision as ‘I’; some PDO collisions as ‘N’ (when property damage > CAD 1,000); and, some PDO collisions as ‘M’ (when property damage <= CAD 1,000). Figure 3.1 shows the spatial distribution of collisions in RDCO for year 2006. Figure 3.1 shows the total collision (i.e. includes all types of collisions) densities (collisions/Km^2) and Figure 3.2 shows severe collision (i.e. only fatal and injured collisions) densities in the region. Clearly for both types of collisions, collision density is higher in central Kelowna compared to other locations of the RDCO.

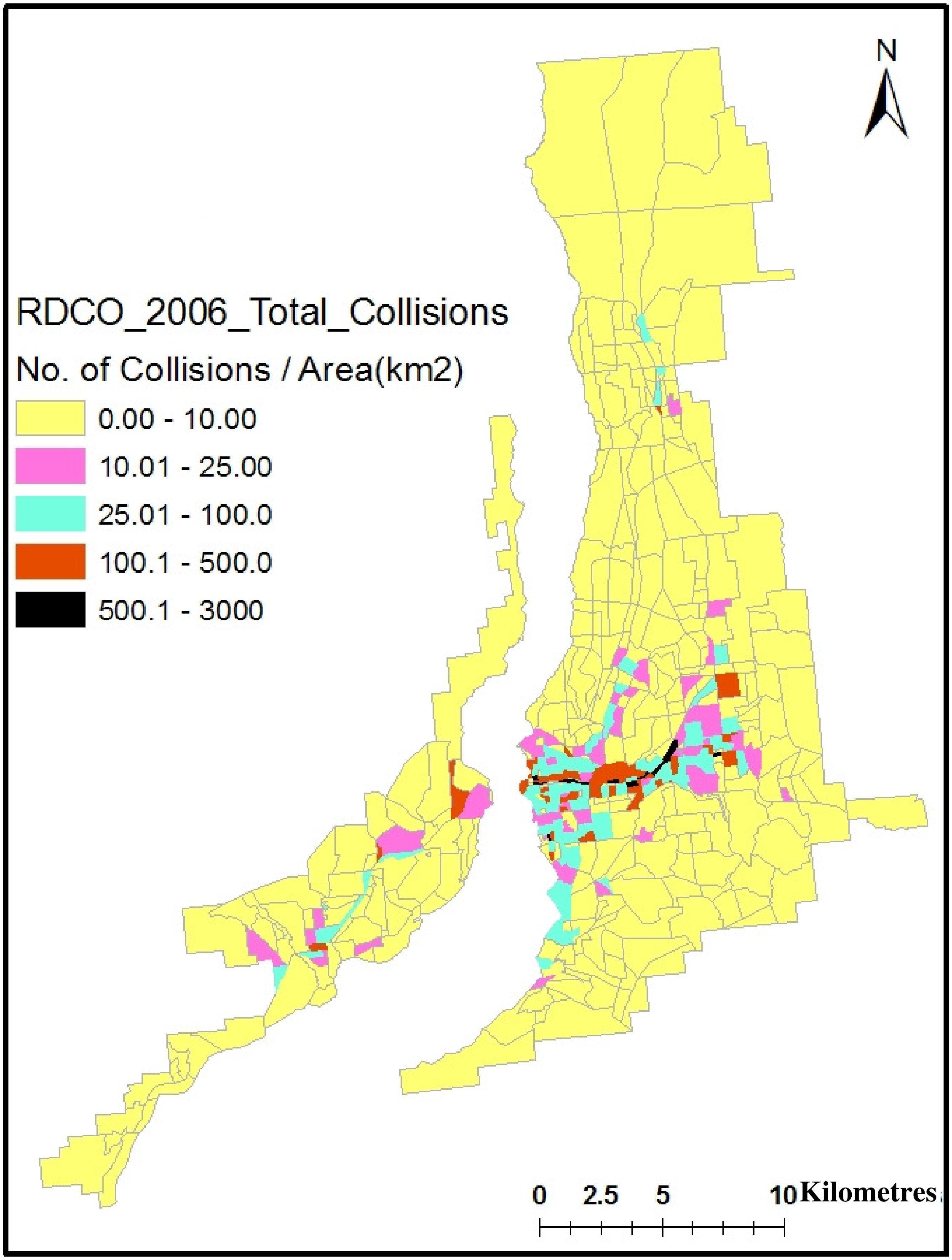


Figure 3.1 RDCO total collision densities

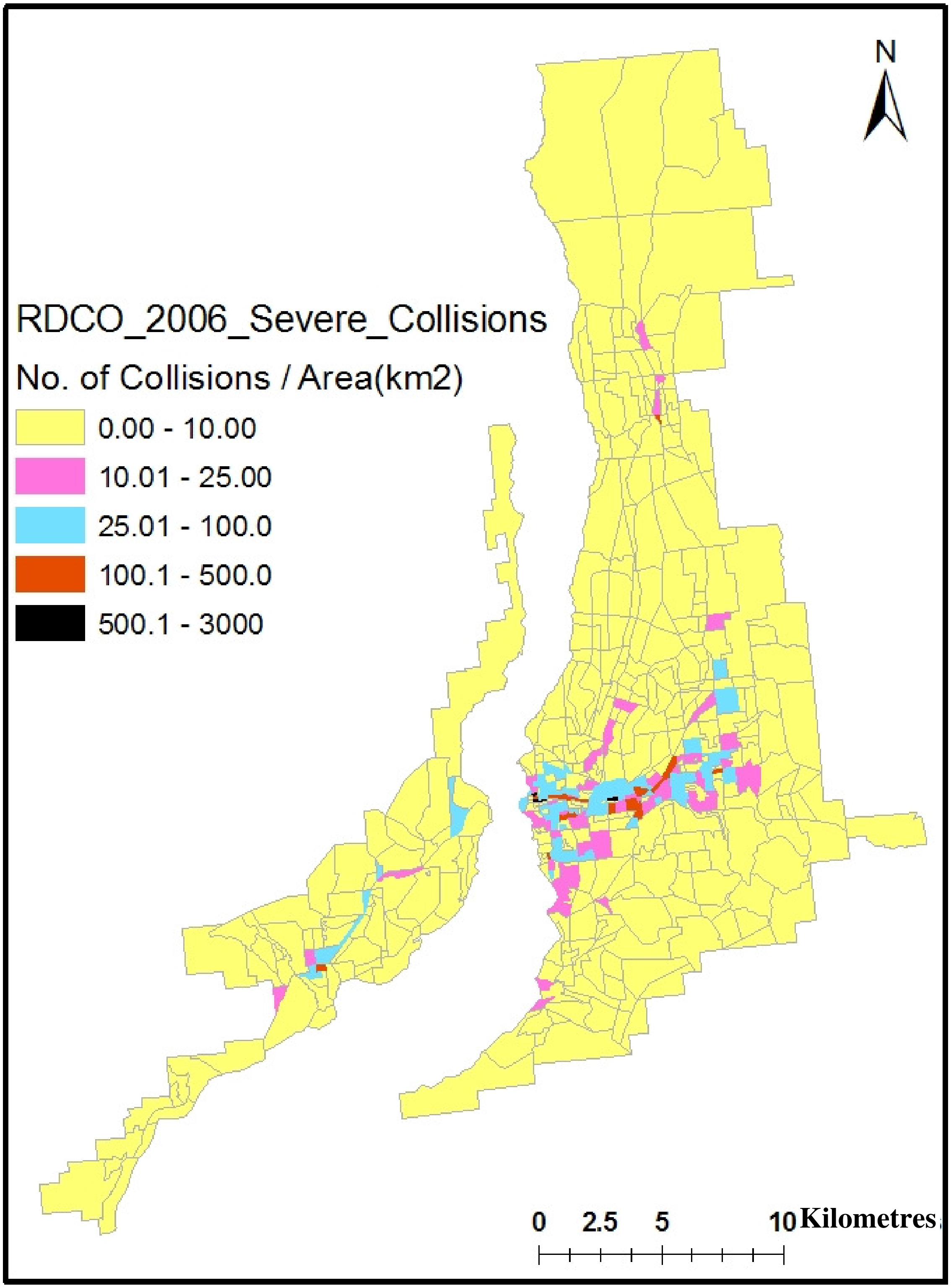


Figure 3.2 RDCO severe collision densities

3.2.2 Exposure variables

Exposure variables can be of two types: modeled data and measured data. The digital road network data was obtained from 'Geobase' (2010). The road network data contains several attributes such as length, number of lanes, posted speed, road class types etc. Traffic volume data on Highway 97 and Highway 33 was downloaded from the BC Ministry of Transportation (BCMoT, 2010) website. This highway data provides 24-hour traffic count for some selected intersections. During the period of the research, there was no reliable transportation planning model for RDCO which can give present and forecasted traffic volume on each link. So this study focused on building a four-step transportation planning model of the RDCO and the modeled exposure data (i.e. VKT, VC, SPED) for each link was obtained from the VISUM model. After that the link shapefile was imported to ArcGIS to calculate the value of each exposure variable. Zonal vehicle-kilometres-travelled (VKT) can be obtained by equation (3.1).

$$VKT_i = \sum_{i=1}^{n} Volume_i * Length_i \qquad (3.1)$$

where,

VKT_i= total VKT in TAZ_i;

$Volume_i$= traffic volume assigned to link 'i';

$Length_i$=length of link 'i' in TAZ_i; and,

n= number of links in the zone.

Other than VKT, speed and volume-capacity ratio (i.e. congestion) are also important exposure variables. Average zonal speed (SPED) is the average of link speed of a zone weighted by traffic volume. Average zonal congestion (VC) is the average of link volume/capacity ratio of a zone weighted by link length. Equation 3.2 was used to calculate average zonal congestion level. Measured exposure data (i.e. TLKM) was calculated manually from the digital road network.

$$VC_k = \frac{\sum_{i=1}^{n} VC_i * link\ length_i}{\sum_{i=1}^{n} link\ length_i} \qquad (3.2)$$

where,

VC_k= average zonal congestion level in TAZ 'k';

VC_i= congestion level at link 'i'; and,

$link\ length_i$=length of link 'i' in TAZ_k.

3.2.3 Network variables

In the study, all of the network variables are measured data. Some of these data were obtained digitally and some of these manually. The digital road network and road intersection map was obtained from 'Geobase'. Manual aggregations were performed to extract data on intersection, road geometry, and road-lane-kilometres. ArcGIS 9.2 software was used to calculate intersection density and total lane kilometre of different roads (i.e. highways, arterials, collectors and local roads). The geo-coded bus stop location of the RDCO transit service was provided by BC Transit (BC Transit, 2006). The digital bus route map was not available. So 2006 RDCO bus routes were created using 2006-bus schedule guide (City of Kelowna, 2006) and 2020 bus routes were prepared according to the Central Okanagan Smart Transit Plan (IBI Group, 2005).

3.2.4 Transportation demand management (TDM) variables

TDM variables were obtained from 2006 Census data. Several types of TDM variables were collected such as the number of commuters and mode split. Again the number of commuters (TCM) was broken down according to the choice of travel mode (i.e. drive, passenger, transit, bike, walk, taxi) which enabled to calculate the number of commuters by each mode. The traveller's trip characteristics were obtained from a RDCO household travel survey by Winram (2007). Among other CPM variables, CORE is an important TDM variable which is defined as the largest portion of the traffic zone, not bisected by major roads (van Minnen, 1999). Lovegrove (2007) identified two other TDM variables namely shortcut capacity (SCC) and shortcut attractiveness (SCVC). These variables provide a zonal descriptor of the neighbourhood's local road access network. Calculation of these variables needs huge time because it requires manual inspection through each TAZ. Also as this research is not going to develop TDM related CPMs, these variables were not calculated. The impact of these variables on results can be verified in future research.

3.2.5 Socio-demographic variables

The socio-demographic (S-D) variables were measured, and derived from Census Canada database. Databases of 1996, 2001 and 2006 were used to obtain data for each Dissemination Area (DA). Here, DA refers to a small area composed of one or more neighbouring blocks and it is the smallest standard geographic area for which all census data are disseminated (Census

Canada, 2006). The data includes various attributes such as population, average income, number of people per household and number of employed persons. The data for each DA was obtained in tabular form and the table was joined with the geo-coded map of DA so that information for each DA is available in 'shapefile' format. 'Shapefile' helps to do spatial analysis in ArcGIS software. The advantage of using ArcGIS is that it has the inherent ability to handle and manage bulk amount of data to do spatial analysis. This research used ArcGIS 9.2 software for all spatial analysis. As an example, population data was obtained in tabular form and using ArcGIS it was joined with DA map to calculate and display population density. The following Figure 3.3 shows the RDCO population distribution in terms of population density (POPD). The highest density zones are in red and the lowest density zones are in yellow color. Other density zones are in between. This figure clearly shows that the population density in 2006 was much higher in central Kelowna compared to other locations of the region.

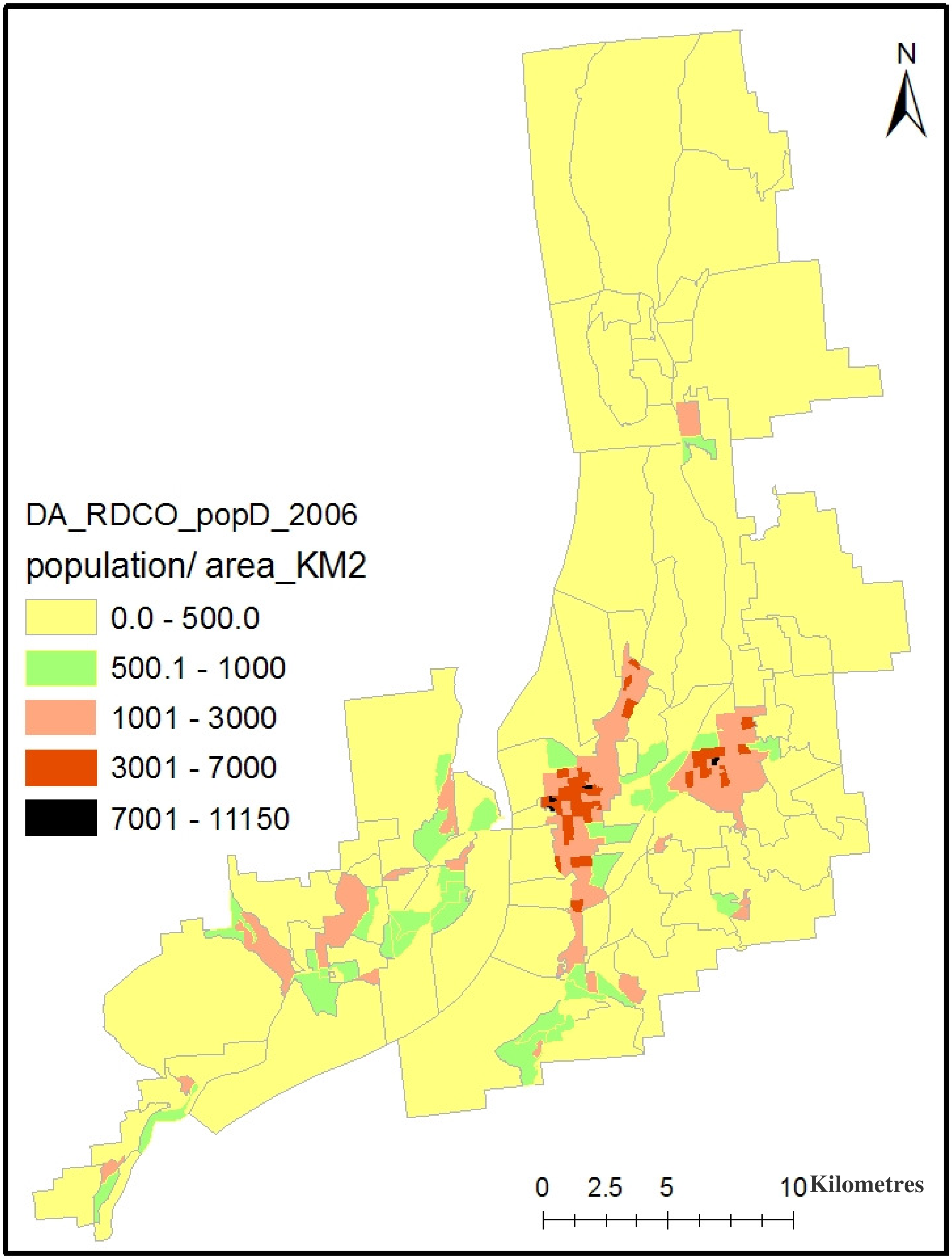


Figure 3.3 2006 population density of the RDCO

The number of students in education institutes (i.e. schools, colleges, universities) is also an important input variable for trip generation process. Schools include elementary schools, secondary schools and high schools. The location of 56 RDCO schools, 1 college and 1 university was obtained from CanMap® and the student number of each school for year 2004 - 2009 was obtained from BC Ministry of Education (2010) website.

3.2.6 Land use

Land use pattern is one of the main variables in transportation planning model as well as collision prediction model. RDCO land use map for year 2006 was obtained from CanMap®. While, the map includes different types of land use patterns, this research considered four different categories of land use types; they are: (1) residential area, (2) industrial area, (3) commercial area, and, (4) institutional area. The land use map for year 2020 was obtained from the Official Community Plan (OCP) of each of the municipalities in RDCO. Figure 3.4 shows future land use types of the region in 2020.

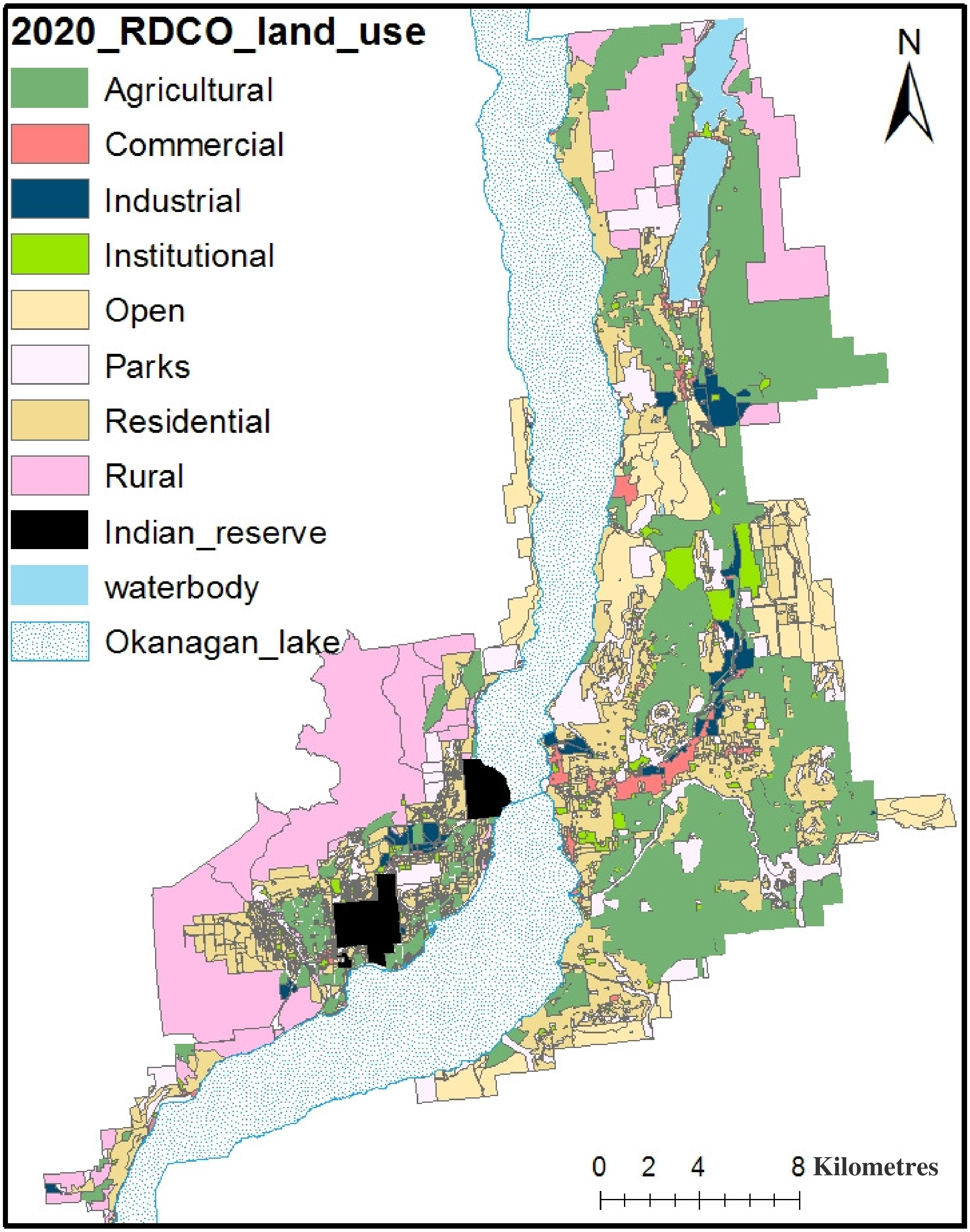


Figure 3.4 RDCO future land use in 2020

3.3 Data extraction methodology

Sufficient sound quality data is one of essential requirements of a reliable statistical analysis. Data should be collected from a reliable source. Sometimes data extraction and management is necessary to obtain data in useable format. A successful data extraction process increases the chance of obtaining a well-fitted statistical model. For this research, several criteria were considered during data selection and extraction process. These criteria were: (1) the variable data should be sufficient but not expensive to be collected; (2) the variable data should be collected and extracted in a relatively accurate and replicable way; (3) variable definitions should be reasonable and easy to understand; (4) the variables should be predictable to be used in future; and, (5) the variables should be relevant and practical for use. The next section discusses about aggregation unit, aggregation approach and data stratification process.

3.3.1 Aggregation

Sometimes aggregation is necessary when the unit of available data and the unit of analysis are different. For the research, most of the data was obtained from census Canada where the unit was DA. But for the analysis, Traffic Analysis Zone (TAZ) was chosen as analysis unit. So aggregation was necessary to obtain useable data for each TAZ.

3.3.1.1 Aggregation unit

For transportation planning model and collision prediction model development, TAZ was used as the aggregation unit. The whole RDCO region was divided into several TAZs. US Census Bureau (2001) defines TAZ as “TAZ is a special area delineated by state and/or local transportation officials for tabulating traffic-related data, especially journey-to-work and place-of-work statistics”. It is a unit of geography, most commonly used in conventional transportation planning models. Generally TAZ boundaries are set by major roads or by zone boundaries. The size of a TAZ is chosen in such a way that the socio-demographic characteristics (e.g. income, population density) among TAZs become homogenous. The TAZ map of the RDCO was obtained from the City of Kelowna (2003) in CAD file. ArcGIS 9.2 was used to convert the CAD file to ‘shapefile’. This map divided the whole region into 500 TAZs, whereas City of Kelowna had 378 zones, District of West Kelowna had 69 zones, District of Lake Country had 42 zones and District of Peachland had 11 zones. Most of the case, TAZ boundaries overlapped DA boundaries very closely and this reduced the possibility of aggregation bias. A map of the RDCO TAZ can be found in Appendix B.

3.3.1.2 Aggregation approach

Even though the TAZ boundaries are very close to DA boundaries, the boundaries do not always match. A TAZ can wholly or partially lie within a DA boundary. So it is possible to obtain one part of a TAZ in one DA and other part in another DA. To improve data quality and to reduce aggregation bias, it is very important to pay attention while aggregating data for TAZs. Some assumptions were considered during the aggregation process. One assumption was to consider the uniform distribution of DA attribute values across its entire area. Another assumption was to aggregate attribute values based on the common area between TAZ and DA. If any data was aggregated as a summation, the common area between TAZ and DA was used. As an example the following equation (3.3) was used to calculate the total number of population for TAZ 'i':

$$POP_{TAZ(i)} = \sum_{i=1}^{n} POP_{DA\ (i)} * \frac{common\ area\ between\ TAZ_i\ and\ DA_i}{area\ of\ DA_i} \qquad (3.3)$$

Another assumption about population aggregation was to consider 'zero' population if any TAZ was located in 'open area' land use category. Sometimes it was also necessary to calculate the average value of an attributes. In that case, instead of simply averaging DA values, weighted average was calculated based on the common area between TAZ and DA. As an example, for calculating the average value of household income (HHI) for a particular TAZ 'i', the following equation (3.4) was used:

$$HHI_{TAZ(i)} = \frac{\sum_{i=1}^{n}\left(HHI_{DA(i)} * common\ area\ between\ TAZ_i\ and\ DA_i\right)}{total\ area\ of\ TAZ\ (i)} \qquad (3.4)$$

This study also used land use map of the region. To obtain land use pattern of each TAZ, ArcGIS 9.2 was used. A built-in function called 'intersection' was used to obtain the land use pattern of each TAZ. It is important to mention here that the 2020 land use map of the District of West Kelowna didn't include two Indian reserve areas. The area of these Indian reserves is 7% of the total area of the West Kelowna district. This area was not too large to be ignored. So, for these two reserve areas, 2006 land use pattern was assumed.

To calculate the number of collisions inside of TAZ, 'spatial joint' function of ArcGIS was used. It enabled to calculate collision density and to identify zones with higher collision densities. Generally community-based macro-level CPMs development process ignores highways because of their limited access. Traffic flows in highways are not characterized by the traits of surrounding neighbourhoods (Hadayeghi et al., 2003). So, road collisions at highways are not significantly influenced by surrounding neighbourhood characteristics. But within RDCO, Highway 97 and Highway 33 do not have limited access everywhere. As it goes through the central part of Kelowna (especially Kelowna downtown), most of the cases it is easily accessible by local traffic. So, highway collisions are significantly related to local traffic volume and therefore, highway data was not excluded while developing CPMs for the RDCO.

3.3.1.3 Data stratification

To minimize aggregation bias and to obtain reliable results, stratification was performed for both independent and dependent (i.e. collisions) variables. For collisions, instead of using only one year data, three-year (2004 – 2006) data was used. The main reason of choosing three year period was to reduce the effect of random fluctuation of collisions. Another reason was of choosing 2006 collision data it would be consistent with the 2006 RDCO transportation planning model. Collisions can be of different types (i.e. fatal, severe and PDO collisions) and therefore, collisions were stratified into two groups: total collision (all types of collisions) and severe collisions (i.e. fatal and injury). Also collisions between 6AM and 9AM were selected to be used in CPMS, because the RDCO transportation planning model was built for AM period (6-9AM).

Stratification of explanatory variables was done in three levels according to the model development guidelines by Lovegrove, (2007). The first level included the four themes of: exposure, S-D, TDM, and network variables. The second level of stratification was based on exposure variables. Two types of exposure variables were used: modeled (i.e. VKT) and measured (i.e. TLKM). Measured data consist of observed data that can be measured from existing road network. On the other hand, modeled data consists of traffic volume, speed, and congestion (V/C) which are the output from transportation planning model. The third level of stratification was based on the type of land use. Literature suggests that the land use pattern can significantly explain collision patterns and there exists significant differences between urban and rural zones in terms of size, population, and employment. If the differences in land use pattern are not considered, it may introduce bias. Therefore, both types of land use, urban and rural, were considered. Zones were classified based on the 2006 population density (pop/km^2). If the population density to a zone was under 800 persons per square kilometre, then it was considered as a rural zone unless it was commercial or industrial area. Also if the population density was over 800 persons per square kilometre, then it was considered as urban zone. These population density criteria were also used in the Okanagan Valley Quest Model (Lovegrove and Stanos, 2006) which was the first RDCO regional integrated strategic growth plan for air, land, water and transport. Details description of this is beyond the scope of this research.

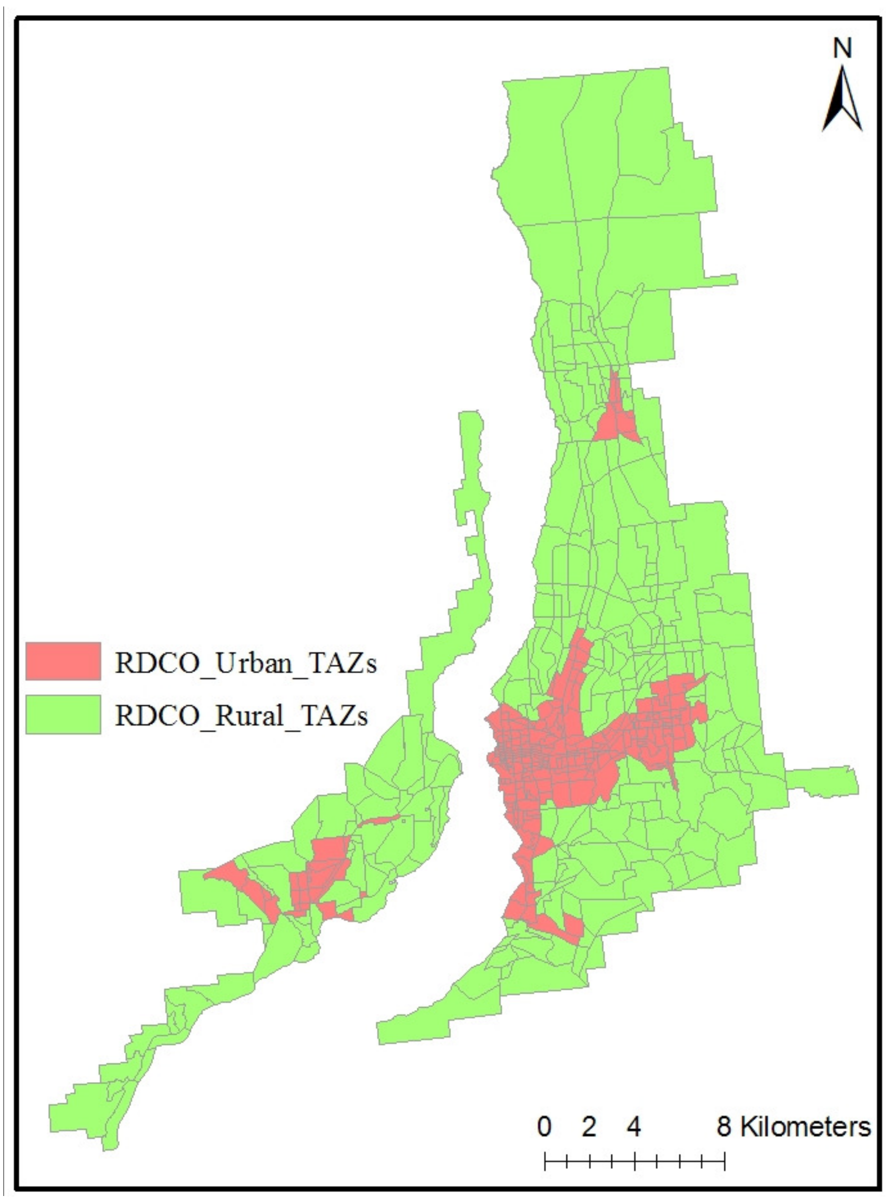


Figure 3.5 Urban and rural zones in RDCO

Table 3.1 lists the candidate variables with possible data source, year, abbreviation, extraction method (e.g. measured/modelled) and RDCO total value for each variable.

Table 3.1 Candidate variables for developing CPMs

Collisions	Symbol	Method	Source	Year	RDCO total	Zonal avg.	min	max	Std Dev
Total collisions over 3years	T3	Measured	ICBC	2004-2006	10871	22	0	371	44.88
Severe collisions over 3years	S3	Measured	ICBC	2004-2006	3033	6.06	0	183	19.18
Total AM period collisions over 3 years	TAM3	Measured	ICBC	2004-2006	1121	2.36	0	61	4.99
Severe AM period collisions over 3 years	SAM3	Measured	ICBC	2004-2006	413	0.87	0	27	2.26
Exposure									
Vehicle KMs Travelled	VKT	Modelled	VISUM model	2006	520950	1099.1	0	21160	10496
Average zonal congestion level	VC	Modelled	VISUM model	2006	n/a	0.1	0	0.78	0.135
Average zonal travel speed	SPED	Modelled	VISUM model	2006	n/a	50/61	0	87.41	14.6
Total lane KM	TLKM	Measured	Geobase	2006	3269.02	6.54	6.54	65.40	0.45
Zonal area (KM^2)	AR	Measured	City of Kelowna	2006	443.57	0.88	0.013	19.22	1.84
Socio- Demographics									
Urban zones	URB	Measured	Census	2006	242	n/a	n/a	n/a	n/a
Rural zones	RUR	Measured	Census	2006	258	n/a	n/a	n/a	n/a
Population	POP	Measured	Census	2006	167,417	334.84	0	24,306	2605
Population density	POPD	Measured	Census	2006	n/a	356.87	0	4861.25	448.1
Population between age 15 to 64 (%)	POP 15_64	Measured	Census	2006	n/a	63.25	0	79.76	11.35
Employment rate	EMPP	Measured	Census	2006	n/a	59.61	0	79.67	13.37
Unemployment rate (%)	UNEMPP	Measured	Census	2006	n/a	5.14	0	19.1	2.37
Average zonal family income	FS	Measured	Census	2006	n/a	2.65	0	3.3	0.43
Average children per family	CHILD	Measured	Census	2006	n/a	0.85	0	1.4	0.27
Average income	INCA	Measured	Census	2006	n/a	32476	0	62172	8430

Collisions	Symbol	Method	Source	Year	RDCO total	Zonal avg.	min	max	Std Dev
TDM									
Total commuters from each zone	TCM	Measured	Census	2006	537305	1074.61	105	3340	546.55
Commuter density	TCD	Measured	Census	2006	n/a	5481.48	25.82	67389.6	7943.1
No of commuters by biking (%)	BIKE	Measured	Census	2006	n/a	1.86	0	14.23	2.52
No of commuters by walking (%)	WALK	Measured	Census	2006	n/a	4.8	0	30.2	5.22
No of commuters driving (%)	DRIVE	Measured	Census	2006	n/a	80.5	10.45	97.01	10.87
No of commuters as car passengers (%)	PASS	Measured	Census	2006	n/a	7.47	0	19.02	3.03
No of commuters by bus transit (%)	BUS	Measured	Census	2006	n/a	2.35	0	19.02	2.52
Network									
No. of intersections	INT	Measured	Geobase	2006	6586	13	0	87	15
Intersection density (int/area_km2)	IND	Measured	Geobase	2006	n/a	18	0	296	38
No. of Arterial lane-KM	ALKM	Measured	Geobase	2006	337.49	0.706	0	11.96	1.568
No. of Arterial lane-KM/TLKM (%)	ALKP	Measured	Geobase	2006	n/a	12.24	0	100	22.69
No. of Collector lane-KM	CLKM	Measured	Geobase	2006	387.66	0.811	0	33.95	2.149
No. of Collector lane-KM/TLKM (%)	CLKP	Measured	Geobase	2006	n/a	12.37	0	100	20.64
No. of Local lane-KM	LLKM	Measured	Geobase	2006	2400.9	5.023	0	39.8	5.766
No. of Local lane-KM/TLKM (%)	LLKP	Measured	Geobase	2006	n/a	73.09	0	100	28.27

3.4 Methodology for 4-step transportation planning model

Transportation planning model is very important to provide adequate transport network information to decision makers. To do transportation planning, various transportation planning software (i.e.EMME3, VISUM, TransCAD) are available. Previous chapter discusses the features of each software. Considering level of efforts, availability, time consumption, and relative easiness, VISUM-11.0 software was selected to use to build the RDCO transportation planning model.

3.4.1 Building network

This research was intended to build transportation planning model for two scenarios: one was for 2006 scenario and another one was for 2020 scenario with four different sub-scenarios. The base year was selected as 2006 because of the availability of 2006 Census data. Also year 2020 was selected because of the availability of predicted 2020 official community plan (OCP) data. The basic steps involved in building and analysing networks were same for both scenario models. This section describes the basic assumptions and steps involved in the modeling process.

To build a road network in VISUM, four important elements are required; they are nodes, links, zones, and connectors. A node represents the end of a roadway or the intersecting point of roads. In VISUM a node is characterized by several attributes including through capacity, turning restrictions, and turning capacity. The capacity of a node can significantly influence the travel time of a trip. Canadian Capacity Guide (CCG, 2008) gives a list of saturation flow rates of signalized intersections for different regions in Canada. For BC it gives the flow rate for Victoria region and the recommended capacity in Victoria, BC is 1,735 Passenger Car Unit (PCU)/hr for sub-urban area, and 1,565 PCU/hr for downtown area. Also Highway Capacity Manual (HCM, 2000) recommends the saturation flow rate in the range of 1,700 to 2,200 PCU/hr. But saturation flow rate does not represent the actual capacity of an intersection. Because saturation flow rate gives the number of vehicles that pass an intersection if 100% time of a signal cycle is allocated to that approach. But practically this is not possible because other approaches also get green time to allow their vehicular movement. So capacity of a node is less than the saturation flow of that node. If the green time of a signal is 50% of the total cycle time, then capacity of that intersection will become 50% of the saturation flow. The GVRD EMME2 transportation planning model assumed node capacity of 400 PCU/hr/lane for stop sign controlled intersections and 1,600 PCU/hr/lane for free flow intersections. The RDCO digital road network does not include intersection capacity as well as the type of intersection control (i.e. signalized, stop sign). So the intersection capacity was assumed 1,000 PCU/ hr/lane. Future research is needed to identify intersection types as well as to adjust intersection capacity.

Links (i.e. roads) are one of the most important elements of road network. The main attributes of a link is characterized by the length, capacity (PCU/hr), free flow speed, number of lanes in each direction and travel time. The RDCO geo-coded road network contains information such as link length, number of lanes. But it doesn't include capacity of links, rather it includes road types (i.e. arterial, local etc.). So the capacity of each link was assumed based on the type of the road. Again GVRD transportation planning model was used to assume link capacity values. Following Table 3.2 gives the capacity assumed for each type of road. By multiplying the capacity and the number of lanes, total capacity (PCU/hr) of each link was calculated. Also as it was not possible to know the percentage of heavy vehicles on each link, the value of this was set to zero (i.e. default value). Future study could help better prediction of heavy vehicle volumes on roads.

Table 3.2 Assumed link capacity of each road class

Type of road	Link category	Capacity (PCU/hr/lane)
Alleyway / Lane	1	900
Local/Strata	1	900
Local/Street	1	900
Local/Unknown	1	900
Service lane	1	900
Resource/Recreation	2	500
Ramp	3	300
Collector	4	1300
Arterial	5	1500
Expressway/Highway	6	1900
Freeway	6	1900

Zones are the starting point and destination of trips. In VISUM every zone can be assigned a zone boundary which represents the spatial extension of the zone. Mathematically TAZs are reduced to a zone centroid. Centroid is assumed to be the generating point of all trips, meaning that all trips to/from a zone are assumed to produce from or attract to that point. As the location of centroid doesn't have any significant impact on the trip distribution and trip assignment results, for simplicity the centroid of residential area of a TAZ was assumed to be the centroid of that TAZ. This step was done using ArcGIS. The location of each centroid was examined carefully and if any centroid was found to be on water bodies, then the location of the centroid was moved manually to a new location that seemed to be the centroid of residential area. In the next step each centroid was linked to respective TAZ specific attributes. After that the centroid shapefile was imported to VISUM to represent 'zone'. Details of the TAZ attributes will be discussed later.

Connectors connect zones to road network. It connects zone and node so that mathematically the road users can exit and enter the zone. Each connector has two directions. The 'out' direction is used for distribution and 'in' direction is used to collect traffic from surrounding road network. This research used two-directional centroid connectors and most of the cases several connectors were used for each zone so that traffic can go to the nearest node depending on the direction of travel. It will distribute traffic in local streets which can give better neighbourhood specific traffic volume. The network elements of a typical transportation planning model can be found in Appendix G.

3.4.2 Trip generation

Trip generation is the step which calculates total number of trips produced and/or attracted from a zone. Previous chapter discussed three different methods for trip generation. Among them regression and cross classification method require trip generating characteristics of households which can be obtained by detailed traveller's surveys. But these surveys are expensive and time consuming. For the study area, no detailed survey and trip generation rate was available. Due to these constrains, it was decided to use ITE trip generation rate (8^{th} edition) to calculate the number of trips from each zone. The analysis period of this research is AM period (6am-9am) and ITE provides hourly AM period rates. As during the first hour of AM period (i.e., 6am-7am) traffic volume is not very high, it was assumed that the total traffic volume between 6am- and 9am would be 2.2 times than the hourly ITE AM period rates. Also it was decided to calibrate ITE trip rates based on the 2007 household survey data (Winram, 2007).

There are lots of variables which can affect the number of trips for a zone. Based on the data availability, following data sets were chosen to use for trip generation. Table 3.3 lists the trip generation variables along with ITE recommended trip generation rates (for 3hours). These trip rates were used for both 2006 and 2020 model.

Table 3.3 Trip generation variables and ITE recommended rates (ITE, 2008)

Variable	Unit	Trip production rate	Trip attraction rate
Residential area	Area (Acre)	1.768	0.312
Industrial area	Area (1000 ft^2)	0.528	1.232
Commercial area	Area (1000 ft^2)	1.320	1.320
Institutional area	Area (1000 ft^2)	0.312	1.248
No. of students in elementary school	No. of students	0.277	0.283
No. of students in secondary school	No. of students	0.268	0.656
No. of students in middle school	No. of students	0.446	0.545
No. of students in college	No. of students	0.132	0.308
No. of students in university	No. of students	0.132	0.308

It has been found that for most of the RDCO schools, the total number of students decreased from 2004 to 2009. Also from 1991 to 2006 the percentage of population under 19 years decreased from 25.3% to 22.4% and in 2020 this percentage will be 18.5% (City of Kelowna, 2009). This indicates that from 2006 to 2020, the percentage of population under 19 years will decreases by 3.9%. To calculate the number of school-going students in 2020, it was assumed that the student number would also decrease by 3.9% from 2006. The RDCO region has one college which is the Okanagan College, Kelowna campus. In 2006/07 it had total number of 4,067 student and the projected 2020 student number is 6,400 (RPG, 2007). The University of British Columbia Okanagan is the only university in the region. In 2006 the total number of students and employee was 5,000 and it is expected to have 8,500 by 2020 (PFS, 2010). The transportation planning model also allowed for external zones. The Regional District of North Okanagan (RDNO) was the North external zone and Penticton was the South external zone. This model also had provision for East and West external zones. But as the volume from/to East and West external zones were negligible, the volumes were set to zero in 2006 model. But the traffic volume for RDNO and Penticton was calculated using the 2007 household travel survey. It was found that during 6-9 AM 2,578 trips travelled from the RDNO to the RDCO and 1,998 trips travelled from RDCO to RDNO (Winram, 2007). As the 2007 household survey was held in April-May, 2007; it was decided to use 2007 data for the base year (i.e. 2006) VISUM model. The volume for Penticton was also calculated in the same way. The 2007 household travel survey data can be found in D.

3.4.3 Trip distribution

Trip distribution is the process which distributes trips to different destinations. This step determines the destination of each trip and produces an O-D table which shows the number of trips for each pair of zone. A comprehensive literature review was done in the previous chapter and it has been found that Gravity model is the most widely used trip distribution model. So this research also used the Gravity model. The form of the model is given in equation (2.3) which clearly shows that the possibility of making trips decreases with the increase of impedance. Travel time was used as impedance in the equation. The free flow travel time and journey travel time are calculated and stored in a 'skim matrix'. For each O-D pair, skim matrix stores different types of information like travel time, travel distance and cost.

Friction factor 'F' can be expressed by different functions. As the shape of the F-curve takes the shape of a negative exponential curve, the function assumed for F-curve was:

$$F_{ij}=e^{-ct} \tag{3.5}$$

where,

F_{ij} = friction factor for the pair 'i-j'; t= travel time in minutes; and, c is the coefficient.

The value of 'c' depends on the purpose of the trip and traveller's characteristics. The value of 'c' can be obtained by trial and error procedure followed by a calibration process. Unfortunately, there is no established friction factor curve for the study region. A traffic study was done on Okanagan Lake Bridge by Halcrow Group (2004), which used 'c' values ranging from 0.12 to 0.14. It is important to mention here that the study was done primarily for such traffic that would cross the proposed bridge. As a starting point of this research, initial 'c' value was considered 0.1 and the resultant trip distribution was observed and compared to 2007 distribution. The value of 'c' was increased gradually and after some trials; it was found that for a 'c' value of 0.2 the modeled distribution is close to the observed distribution. So, 'c' value was chosen as 0.2. Future research is needed to find more accurate 'c' value (i.e. friction factor curve) for the region. Also due to the unavailability of zonal employment data, it was not possible to calculate the value of K-factor. So the value of 'K' was assumed '1' for both 2006 and 2020 scenarios. Literature also suggests that the use of 'K' may introduce error in predicting future trip distribution (Duffus et al., 1987).

Using link length and free flow speed, the gravity model calculated the initial travel time for each O-D pair and it distributed trips based on the minimum travel time (in minutes). But as traffic volume increase, congestion and travel time also increase. As the new travel time can change trip distribution pattern, it is necessary to update O-D table based on the new travel time. So after performing the last step (i.e. trip assignment), new travel time was calculated and returned to the skim matrix so that refined travel time could improve trip distribution process. This process was continued until desired accuracy was achieved. It is also necessary to calibrate O-D table based on the observed data. But for the study region, there exists no O-D table to calibrate the 2006 VISUM model. So several auto trips were made within the study area and the observed travel time was used to verify the modelled travel time.

3.4.4 Mode choice

Mode choice is the third step where trips are assigned to a particular mode. This is the step that determines which mode will be chosen by a traveller for a particular trip. This step also calculates the percentage of mode share by travellers. Disaggregate mode choice models are the most widely used models to determine mode share. It calculates the utility of a mode based on the socio-economic condition and mode specific perceptions of the travellers. This information can be obtained by conducting a detailed travel survey at household level. Unfortunately, for the RDCO there was no mode related travel survey to allow calculating mode specific utility values at disaggregate level. Conducting a traveller's survey would be an important topic for future research. As it was not possible to obtain disaggregate level data, this research instead focused on building an aggregate based mode choice model. Census 2006 data were considered for the purpose. The following variables were considered for the mode choice regression purpose:

Table 3.4 Mode choice candidate variables

Variables	Symbol
No. of bus stops	BStop
Bus stops density	BStopD
Length of bus route (KM)	LBKM
Length of bus route/TLKM	LBKMM
Bus headway	BH
Walk-access time	WAT
Cost ratio (=transit fare/auto operating cost)	CostR
Average household income	HHI
Employed person (%)	EMP
Average vehicle per household	VHH
Population between 45-64 age (%)	POP45_64
Travel time ratio	TTR

Even though impedances (i.e. travel time) are very important for mode choice decisions, if there exist no bus route and/or bus stop in an area, people will not select bus as their travel mode. Therefore, this research considered transit related attributes such as the number of bus stops, bus headway, walk access time to bus stops. 'Walk-access time' indicates the time required to walk from the zonal centroid to the nearest bus stop. It was assumed (in minutes) manually by a careful observation of the location of zonal centroid and nearest bus stops.

Another important variable was 'cost ratio' which represents the ratio of transit fare and auto operating cost. For a 18,000 km per year driven car, the 2006 auto driving cost was $0.518 per km (for Cobalt LTZ car) and $ 0.652 per km (for Caravan car) (CAA, 2007). An average of $0.59 per km was used as the auto driving cost in 2006. Auto parking cost is another cost related variable. But in most of the RDCO, auto parking is almost free except downtown area. To be conservative, parking cost was not included in the 'cost ratio' variable. Assuming $2.00 transit fare and $0.59 per km auto operating cost, the cost ratio for 2006 was calculated as 3.89.

A household survey by Winram (2007) indicates that almost 40% of total and 65.8% of auto trips were made by travellers aged in the range of 45 to 65. So the percentage of population aged between 45 and 64 was also considered as a variable. Travel time is another important parameter for mode selection. Because people try to select such a mode which will minimize his/her travel time. In RDCO, the average travel time by bus in 2007 was 1.79 times more than that by driving auto. But this average value didn't represent the actual travel time ratio for each TAZ. Travel time ratio is dependent on various factors such as the O-D of the trip, availability of the preferred bus and transfer time needed to change bus. As for the region, there exist no travel time table for O-D pairs, 'travel time ratio' was not included in mode share equation. The bus headway for 2006 was obtained from Kelowna Regional Transit System and the expected headway in 2020 was obtained from Central Okanagan Smart Transit Plan (IBI Group, 2005). Other socio-demographic variables for 2006 were obtained from the 2006 Census Canada. The 2020 prediction of other variables was done only for those variables which were selected for the mode share equation and the predicted values are given in the next chapter.

'GenStat' statistical software was used to make mode share equation which will calculate the percentage of travellers using auto. DA was used as the unit of regression analysis and Census-2006 data was used for the regression analysis. At first, a linear regression model was attempted to build. Each variable was included in the model by following forward stepwise procedure. The goodness of fit and the t-statistics of each variable were assessed at 95% confidence level. If the modeled R^2 value was close to 1.0, then the model was considered as a well-fitted model. But if it was not close to 1.0, outliers were removed to improve the model fitness. Even after that if the R^2 value was not close to 1.00, then it was decided to use non-linear mode choice models. When a well-fitted model was developed, then the model form applied to estimate the percentage of auto user for each TAZ. The percentage of auto user was used to calculate the number of auto trips from each TAZ. It enabled to make an O-D table only for auto trips, which was applied into VISUM model to perform the next step which is trip assignment.

3.4.5 Trip assignment

Trip assignment is the last step of four-step transportation model that assigns trips to road network. If an assignment method is selected and the capacity constraints are defined, then VISUM can assign trips to the road network based on the minimum impedance criterion. Following BPR function was selected to consider link capacity constrain:

$$T_{cur} = T_0[1 + 0.15(\frac{Q}{Q_{max}})^4 \quad (3.6)$$

where,

T_{cur}= link travel time; and, $\frac{Q}{Q_{max}} = \frac{link\ traffic\ volume}{capacity\ of\ the\ link}$.

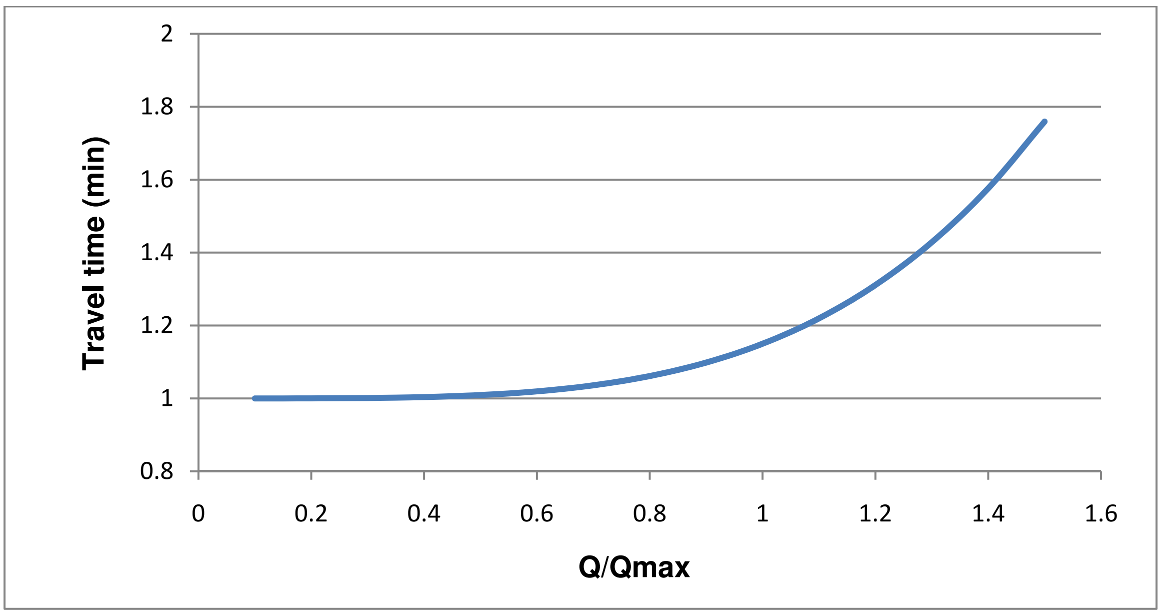


Figure 3.6 Relation between link travel time and congestion on link

Figure 3.6 shows the travel time on a link increases with the increase of Q/Q_{max}. It can be noted that the rate of travel time increase is relatively higher after the link volume exceeds the link capacity.

Node capacity and node type (i.e. signalized, unsignalized) are also important parameter to determine link travel time. The characteristics of signalized and non-signalized intersection are different. Presence of signalized intersections results in higher link travel time when compared to links having unsignalized intersections. So it is important to identify the type of intersection. But for the study area, types of all intersections were not known. Based on the observation, some major intersections were identified as signalized and different impedance function was assumed. Equation 3.7 represents node impedance function for unsignalized intersections, while equation 3.8 is for signalized intersections.

$$T_{cur}(\text{min}) = T_0 + 0.166 + 0.5 * \left(\frac{traffic\ volume}{node\ capacity}\right) + 0.25 * \left(\frac{traffic\ volume}{node\ capacity}\right)^2 \quad (3.7)$$

$$T_{cur}(\text{min}) = T_0 + 0.25 + 0.5 * \left(\frac{traffic\ volume}{node\ capacity}\right) + 0.25 * \left(\frac{traffic\ volume}{node\ capacity}\right)^2 \quad (3.8)$$

Quadratic form has the advantage of imposing some fixed delay in the equation. Signalized intersections increase delay due to traffic signal. Even stop sign controlled intersections also impose some delay because a vehicle will slow down or wait sometime at the intersection. So delay occurs at the intersection regardless of the type of intersections. As for the RDCO intersection types were not known, 0.166 minute (10 seconds) fixed delay for unsignalized intersection and 0.25 minute (15 seconds) fixed delay was imposed for signalized intersections.

Previous chapter discussed a lot of assignment techniques, where each technique has its own advantages and disadvantages. Among different techniques, 'user equilibrium assignment' technique was chosen due to its relatively better assignment result and less computational time requirement. After the trip assignment process, the traffic volume on each link was checked and if the difference between two successive iterations was more than +/- 10 autos, then a 'Go to the Operation' function was used. This function took the output from trip assignment process and feed it into 'skim matrix' to calculate new travel times. The new calculated travel time was used into the trip distribution step to refine O-D table. This process was continued until the difference of traffic volume on each link became less than 10 autos between two successive iterations. Finally the model was validated with the observed data obtained from BC Ministry of Transportation and City of Kelowna traffic count map. For each intersection, number of entering traffic volume was calculated and it was compared with observed data. If the percentage of error was within +/- 30%, then it was considered as acceptable. The screenline count was also compared, but in that case 10% error was considered to be acceptable (Lin & Navin, 1999). The overall transportation planning process is given in the following Figure 3.7:

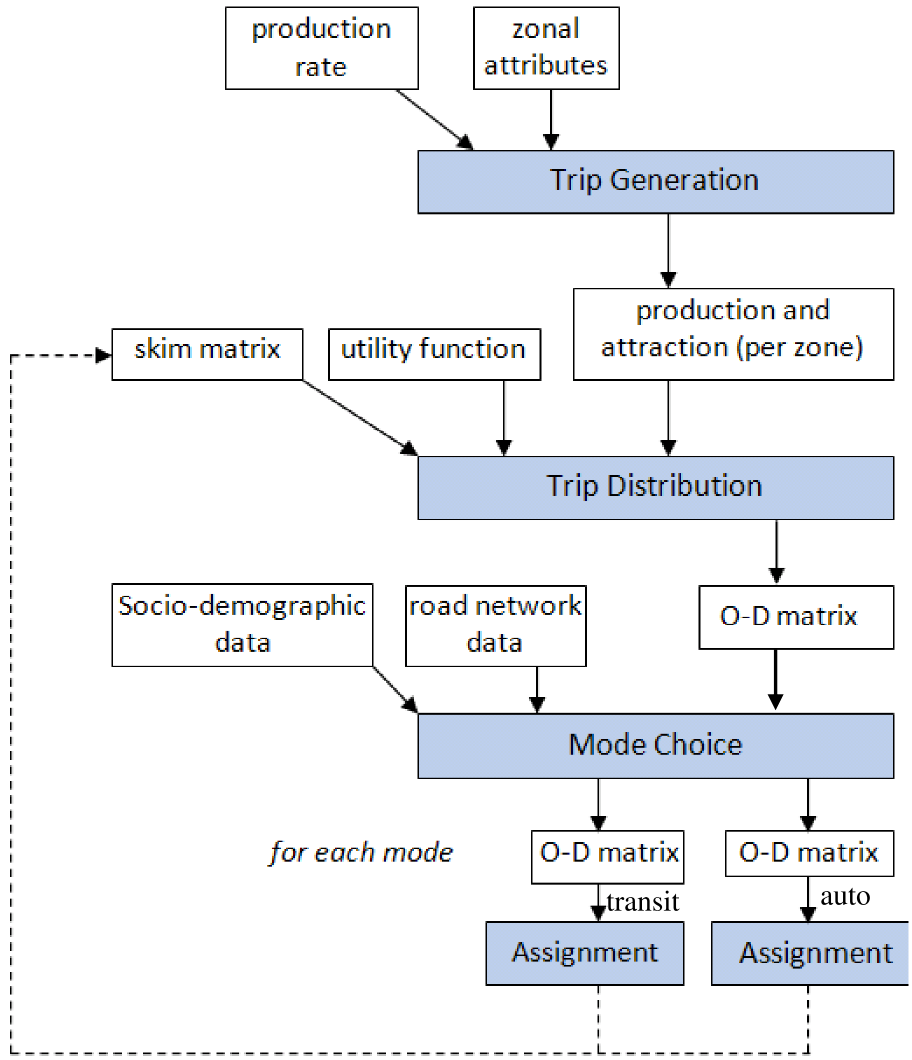


Figure 3.7 Steps involved in 4-step transportation planning model

3.5 Collision prediction model development methodology

The methodology for CPM development was based on the literature reviewed earlier. CPMs can be developed for four different traits given by Lovegrove (2007). One of the main objectives of the research was to compare the predicted collision frequency in 2020 for different scenarios. Based on the four-different traits, 16 different CPM groups can be developed as given in Table 2.1. But it would consume time and it would be beyond the main objective of the research. As exposure variables are the most influencing variable in CPMs, this research developed only exposure related CPMs. For each scenario following eight types of CPM was developed:

- Urban modeled total collision model;
- Urban modeled severe collision model;
- Urban measured total collision model;
- Urban measured severe collision model;
- Rural modeled total collision model;
- Rural modeled severe collision model;
- Rural measured total collision model; and,
- Rural measured severe collision model.

For developing CPMs, modeled outputs (i.e. VKT, VC, SPED) were obtained from the VISUM transportation planning model. Steps involved in calculating these variables are shown in Appendix C. Following sections describe the methodology and steps involved in the development of macro-level CPMs for the region.

3.5.1 Regression

It has been shown that Generalized Linear Regression Modeling (GLM) regression method has the advantages of overcoming limitations associated with the use of conventional linear regression in modeling traffic collisions (Hauer et al. 1988, Sawalha & Sayed, 2001). GLM method has an established methodology, while Geographically Weighted Poisson Regression (GWPR) and Full-Bayesian Semiparametric Additive (FBSA) methods are relatively new and complex. As the objective of this study is to compare the predicted collisions in 2020 and GLM is relatively easier to use, GLM technique was selected for the regression purpose. 'GenStat' statistical software was used for the development of collision prediction models. From a careful observation it was observed that the RDCO collisions were over-dispersed (variance is greater than the mean). It was also verified while developing collision prediction models. Based on the literature at first, the error structure was assumed to follow Poisson distribution and the value of σ_d was calculated. If the value became less than 1.00, then the model formulation was carried using Poisson distribution. Otherwise, if σ_d exceeded 1.00, then NB structure was used for the model development.

3.5.2 Form

Several types of CPM forms were discussed in the literature and after reviewing the literature, the CPM form suggested by Lovegrove (2007) was selected to use. The selected CPM form is given in equation (3.9). This equation contains an exposure term 'Z', which satisfies the 'zero risk' logic. Here, 'zero risk' means no road collision will occur if there is no traffic on roads. The value of 'Z' can be modeled (i.e. VKT) or measured (i.e. TLKM).

$$E(\Lambda) = a_0 Z^{a_1} e^{\sum b_i x_i} \tag{3.9}$$

where,

$E(\Lambda)$= predicted mean collision frequency;

a_0,a_1,b_i= GLM derived parameter estimates;

Z=exposure variable; and,

x_i=independent variables.

After log-linear transformation of equation (3.9), following equation (3.10) was obtained to use in GLM regression. Using the collision, exposure and explanatory variable data, the software can determine the constant term and parameter estimates of the equation (3.10).

$$Ln[E(\Lambda)] = Ln(a_0) + a_1 Ln(Z) + \sum_{i=1}^{n}(b_i . x_i) \tag{3.10}$$

3.5.3 Selection of explanatory variables

Explanatory variable selection followed a forward stepwise procedure, as described in Sawalha & Sayed (2005a). According to the procedure, each model was constructed by adding one variable at a time, and by testing the change in model fitness due to the added variable. The first variable tested in each model was the exposure variable (Z) due to its usual dominating prediction influence (Sawalha & Sayed, 2005a; Miaou, 1996). Additional candidate variables were then systematically added to the model. The decision to retain a variable in the model was based on four criteria. First, the logic (i.e. +/-) of the estimated parameter had to be associated intuitively with collisions. Second, the parameter estimate t-statistic had to be significant at the 95 percent confidence level (i.e. > 1.96). Third, the addition of the variable to the model should have caused a significant drop in the Scaled Deviance at 95 percent confidence level (i.e. SD> 3.84). Fourth, the variable had to show little or no correlation with any of the other independent variables. For example, transit mode split is highly correlated with auto mode split; hence, only one mode split variable (i.e. auto or transit) should be included in the model. Correlation between variables was checked by viewing correlation results in the software. Once explanatory variables were selected for a model, the next step was to find the model goodness of fit.

3.5.4 Goodness of fit

The method used to evaluate the model goodness of fit was described by several researchers. In the research, SD measures and Pearson χ^2 tests were performed to evaluate the goodness of fit. The equations of these two measures are repeated here for convenience:

$$\text{Pearson } \chi^2 = \sum_{i=1}^{n} \frac{[y_i - E(\Lambda_i)]^2}{Var\ (y_i)} \tag{3.11}$$

$$SD = 2\sum_{i=1}^{n} \left[y_i ln\left(\frac{y_i}{E(\Lambda_i)}\right) - (y_i + \kappa) ln\left(\frac{y_i + \kappa}{E(\Lambda_i) + \kappa}\right) \right] \tag{3.12}$$

If the value of these criteria were greater than the χ^2 distribution value with (n - p - 1) degrees of freedom at 95% confidence level, then the model was considered poorly fitted to observed data. So model refinement was done in the form of outlier analysis.

3.5.5 Model refinement

If any model failed to satisfy the goodness of fit test, then it was refined using outlier analysis. The refinement process was done based on the Cook's Distance (CD) technique described by Sawalha & Sayed (2005b). Using the equation (2.42) the CD value for each observed collision was calculated and ranked according to descending order. The point having highest CD was selected and removed from the model. Then the model was re-run without changing the previous value of **κ**. For 95% confidence level, if the new SD statistic value was less than 3.84 than the previous SD value, then the outlier was removed and the model was rerun to estimate a new **κ** and new estimates of CD. The process of removing outliers was continued until the resulting SD value satisfied the goodness of fit criteria. The t-statistics value of each variable was also checked to ensure the significance of each variable. If the resulting model satisfied all of the criteria, then it was considered to be a good fitted model to use. For each successfully fit model, no more than a few of the several hundred data points were found to be outliers (i.e. well under 5%).

3.6 Identification of collision prone locations

A black spot is a location which exhibits significantly higher collision potential when compared to average collision potential from a group of similar locations. To ensure that resources are spent only on locations with the highest potential for safety improvements, it is important to follow a sound procedure to properly identify and rank black spots for diagnosis and treatment. This research identified collision prone zones based on the literature described in Section 2.3.5. As urban and rural areas have significantly different collision potential, the identification process was done for both urban and rural areas using three years (2004-2006) collision data. Total number of collisions for each TAZ was calculated using ArcGIS 9.3. Also from the CPMs, predicted number of collisions, $E(\Lambda)$, for each TAZ was obtained. Using the observed and predicted collisions, the Empirical Bayes (EB) estimate of each location was calculated. Then EB of each TAZ was compared to the regional average, $E(\Lambda)$, which was obtained from CPMs. At 95% confidence interval, if the EB estimate of a location exceeded the reference group average, then the location was considered as collision prone zone (CPZ). Equation 2.29 was used for the identification purpose. After the identification of collision prone zones (CPZs), they were ranked by both Potential Collision Reduction (PCR) and Collision Risk Ratio (CRR) methods. Using the results from these two ranking methods, a combined ranking was prepared for both urban and rural areas.

A collision density map was also prepared for Highway 97. For that purpose, the entire length of the corridor was divided based on the TAZ boundary. 'Intersection' function in ArcGIS was used to the step. Later, the segment lengths were re-calculated so that it represents the actual divided segment length, not the previous undivided length. If any segment length was less than 1 km, then consecutive segments was merged together so that the length became more than 1 km. Later the number of collisions (2004-2006) was calculated for each segments and it was normalized by segment length to represent collision density.

In the 'intersection' shapefile, each intersection was represented by a single point. Road collisions do not occur just at the middle of an intersection, but also near the intersection point due to stop and deceleration of vehicles. So a 30 m radius of area was considered around each intersection and 'Buffer' function of ArcGIS was used to create circular areas. Then, the number of collisions inside of the circular area was calculated using 'Spatial Joint' and based on the collision frequency collision prone locations were ranked in a descending order.

3.7 Summary

This chapter is about the database description, data extraction and methodology used for building transportation planning model as well as for developing collision prediction model. Good, reliable source of data is the most important criteria for having a well fitted model. Several criteria was considered during data selection and extraction process, such as: (1) the variable data should be sufficient but not expensive to be collected; (2) the variable data should be collected and extracted in a relatively accurate and replicable way; (3) variable definitions should be reasonable and easy to understand; (4) the variables should be predictable to be used in future; and, (5) the variables should be relevant and practical for use. A sound procedure of data extraction and aggregation is also necessary. At first this chapter discussed about the data source of each kind of variables. ICBC, Census Canada 2006, Geosuite were the main data sources. The aggregation unit of analysis was TAZ and data were aggregated for each TAZ. Also data stratification was done based on the type of land use, time of collisions and severity of collisions. Table 3.1 was produced to list all the candidate variables for developing CPMs. This table gave the variable name, abbreviation, data source, year and statistics of each variable.

This chapter also discussed a detailed methodology used for building the RDCO transportation planning model for year 2006 and year 2020. VISUM-11.0 software was selected for modeling. Candidate variables for trip generation and mode share were listed in Table 3.3 and 3.4. Different assumptions and methods were also discussed for each of the steps involved in four-step transportation planning. Later this chapter described the methodology used for developing community based macro-level collision prediction models. It discussed about the regression techniques and model form. It also discussed the steps of selecting explanatory variables, assessing model goodness of fit and performing outlier analysis. The methodology defined in this chapter was applied to develop models which are presented in the next chapter.

CHAPTER 4 MODEL DEVELOPMENT RESULTS

4.1 Introduction

Based on the methodology outlined in Chapter Three, four step transportation planning model and CPMs were developed. In section 4.2, the results of the 2006 (base year) RDCO transportation planning model is presented. The next Section 4.3 discusses about the future vision of the Kelowna Regional Transit System and also about the future RDCO road network improvement plan. Section 4.4 represents the results of four different sub-scenarios of the 2020 transportation planning model. Section 4.5 discusses about the development of community-based macro-level CPMs for total collisions as well as for severe collisions by using the 2006 transportation planning model outputs. Section 4.6 predicts the number of collisions for each of the 2020 sub-scenarios and compares the road safety benefits of each scenario. Finally Section 4.7 summarizes the model development results.

4.2 The 2006 RDCO transportation planning model

This research focused on the development of a transportation planning model for the Regional District of Central Okanagan (RDCO). Year 2006 was considered as the base year for the model due to the availability of Census-2006 data that was used in the development, calibration, and validation processes. Also morning 'rush hours' (i.e. 6am -9am) was selected as the time period of the model, because demand load reaches its maximum in that period. Based on the method described in the previous chapter, the 2006 RDCO 4-step transportation planning model was developed and calibrated using VISSUM-11.0 software. The results of each step are presented in the following sub-sections.

4.2.1 Step 1 results: trip generation

Trip generation rates were followed from the 'Institute of Transportation Engineers' (ITE) recommended trip rates. Table 3.2 lists the names and associated trip generation rates of each of the trip generating variables. The variables values were aggregated for each TAZ by using ArcGIS-9.3. The values were entered into VISUM as 'zone attributes' and trip generation rates were entered into VISUM's trip generation module. As the VISUM model also allowed for external zones, external traffic volumes were also entered into the model. Based on the methodology, in 2006 North external zone produced 2,578 trips and attracted 1,998 trips. Also in

2006, South external zone produced 124 trips and attracted 1,281 trips. As the number of trips from/to West and East external zones was negligible, it was set to 'zero'. After performing the trip generation step, total number of generated (i.e. produced + attracted) trips from each zone was calculated. The total number of model-forecast trips produced from and attracted to each zone was compared with the 2007 inter-regional household travel survey data, which is summarized in Appendix D. It was found that the percent error in total production was 0.49% and percent error in total attraction was 20.8%. As the error in predicted attraction is higher than the observation, the attraction rates were readjusted based on the total number of attraction and percentage of error for each city. The average percentage of error was calculated weighted by the number of attractions and attraction rates were reduced accordingly for the next trial. After several re-adjustments the error in total attraction was reduced to 3.03%. As the error in prediction was very low, the trip generation results were considered acceptably calibrated. Table 4.1 represents the final adjusted trip generation rates and Table 4.2 gives the final predicted number of trips compared with the observed 2007 data.

Table 4.1 Final adjusted trip generation rates for the RDCO

Variable	Unit	Trip production rate	Trip attraction rate
Residential area	Area (Acre)	1.768	0.253
Industrial area	Area (1000 ft^2)	0.528	0.999
Commercial area	Area (1000 ft^2)	1.32	1.0713
Institutional area	Area (1000 ft^2)	0.312	1.013
No. of students in elementary school	No. of students	0.2772	0.23
No. of students in secondary school	No. of students	0.268	0.5324
No. of students in middle school	No. of students	0.4455	0.4419
No. of students in college	No. of students	0.132	0.2500
No. of students in university	No. of students	0.132	0.2500

Table 4.2 Trip generation results of the 2006 transportation planning model

Location	Type of trips	Modeled total trips (2006)	Observed total trips (2007)	Error (%)
RDCO	Production	95,566	95,097	0.49
	Attraction	97,473	94,519	3.03
City of Kelowna	Production	69,473	67,485	2.86
	Attraction	78,827	75,513	4.20
District of Lake Country	Production	8,357	8,164	2.31
	Attraction	4,894	5,067	-3.53
West side (West Kelowna & Peachland)	Production	15,034	19,448	-29.36
	Attraction	10,473	13,938	-33.09

From Table 4.2 it can be seen that the errors in calculating production and attraction are marginal except for the West side (West Kelowna & Peachland) region, these errors were expected as the District of West Kelowna was incorporated in 2007, and, between 2006 and 2007 it went through major developments in terms of employment, housing and other facilities. These improvements caused a significant increase in traffic between 2006 and 2007. It might be the reason of fewer trips in the 2006 VISUM model when compared to 2007-observed data. On the other hand, the City of Kelowna and the District of Lake Country are well-established municipalities and the resultant modeled trips are very close to the observed trips.

Finally, using the readjusted trip generation rates, total number of generated trips for each TAZ was calculated and following figures were prepared. Figure 4.1 shows the total number of generated trips across RDCO zones and Figure 4.2 shows the trip generation density (i.e. no of trips/area) across RDCO zones. From the figure it can be seen that trip generation density is focused in the town centres of Kelowna, Lake Country, and West Kelowna. These results were expected as most of these areas are dense residential/commercial urban areas and hence, produce more trips compared to rural areas.

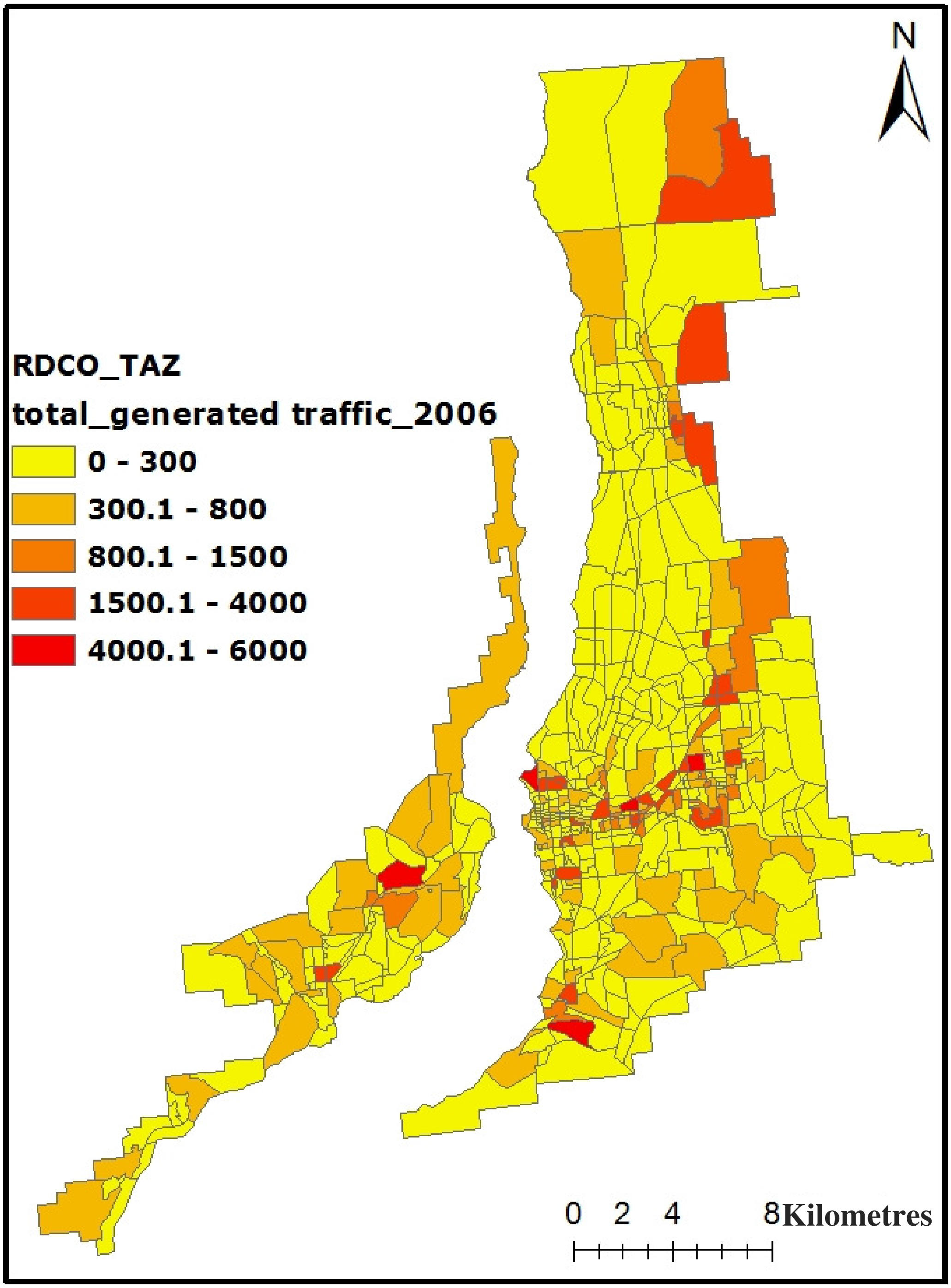


Figure 4.1 Total number of generated trips in 2006 across RDCO zones

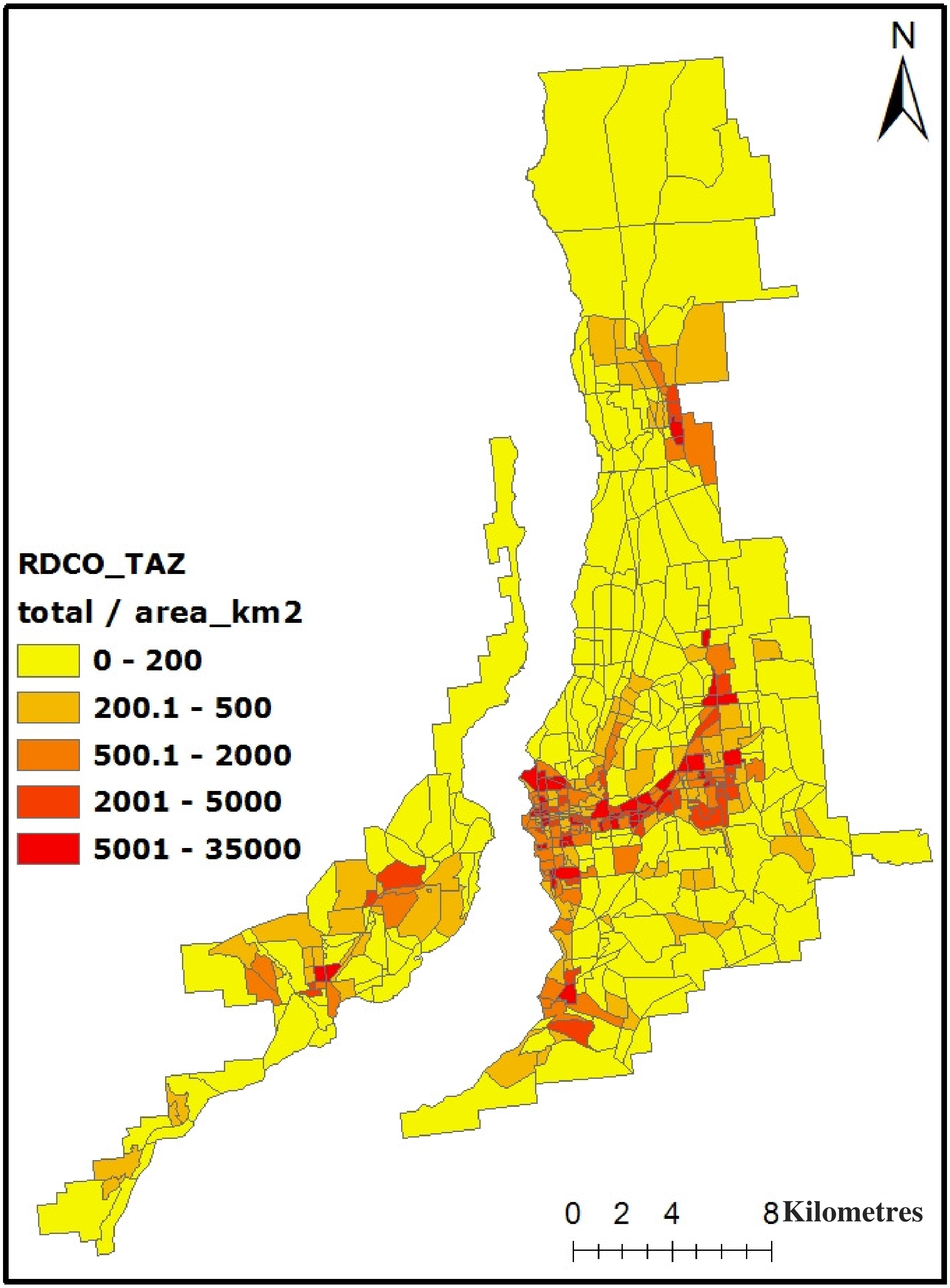


Figure 4.2 Trip generation density in 2006 across RDCO zones

4.2.2 Step 2 results: trip distribution

After calculating the total number of produced trips by each zone, the trip distribution step assigned the produced trips to various destination based on the Gravity model. Travel time was considered as travel impedance and the friction factor function was mentioned in the previous chapter. At the initial stage free flow travel time was used to distribute trips. Later capacity constraint function (i.e. BPR function) was used to update travel times by using the traffic volume on each link which was obtained after the trip assignment step. After several iterations, the final travel time for each trip was calculated and the average modeled travel time was found to be 15.99 minutes. It was very close to the observed average travel time of 16.8 minutes in 2007 AM periods (Winram, 2007). Following Figure 4.3 shows the travel time distribution and it can be seen that the travel time of almost 50% trips was in the range of 9 to 18 minutes.

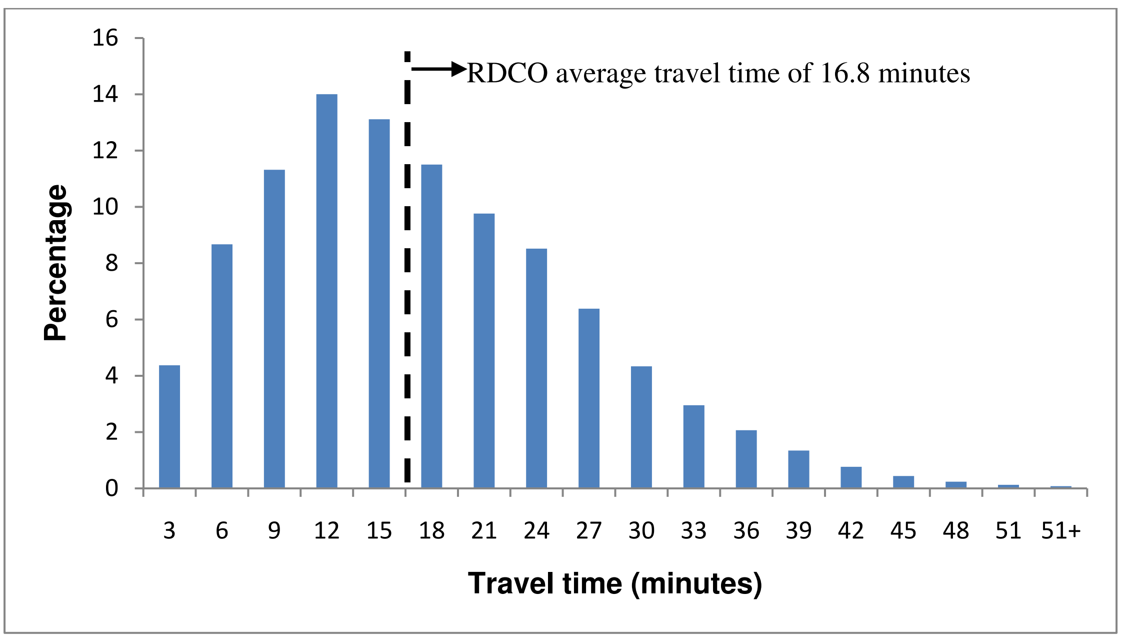


Figure 4.3 Distribution of travel time from the 2006 RDCO VISUM model

The distribution of travel distance was also obtained from the VISUM software, which is shown in the following Figure 4.4. It can be seen that travel distances of almost 50% trips were in the range of 6 to 12 km. Assuming 50 kmph travel speed and average travel time of 15.99 minutes, average travel distance becomes 13.2 km and from the VISUM model the average travel distance was calculated as 14.04km, which is close to the observed 13.2 km.

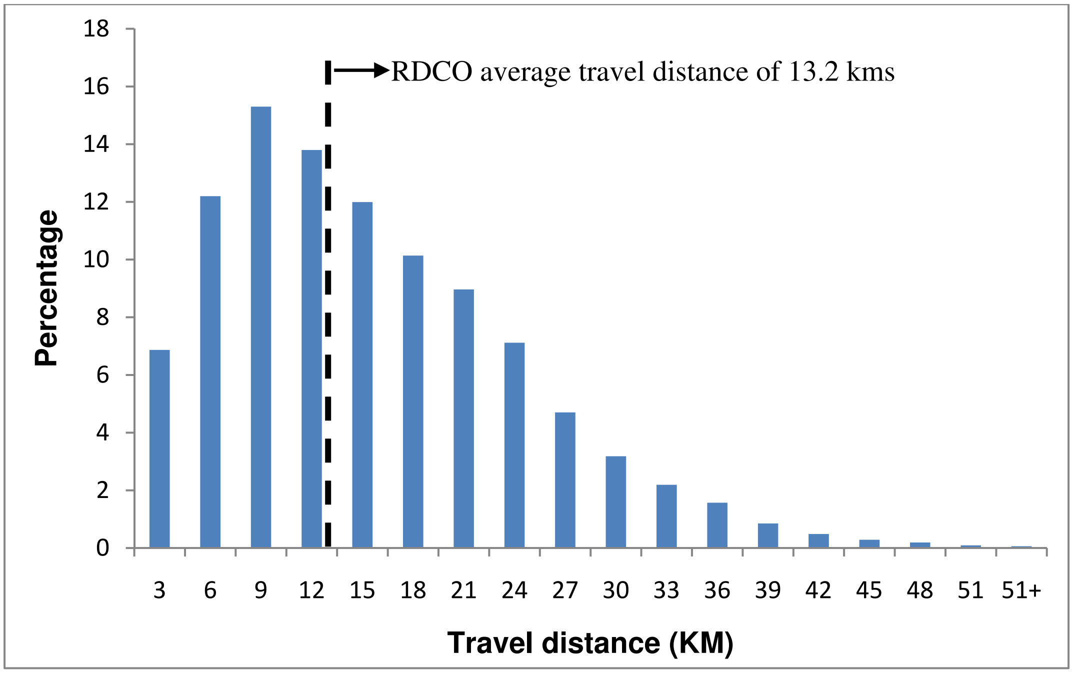


Figure 4.4 Distribution of travel distance from the 2006 RDCO VISUM model

To calibrate the trip distribution model, it is very important to know the observed traffic volume between each pair of origin and destination, which can be obtained from O-D table. But for the study area, there existed no O-D table and hence, it was not possible to calibrate the model using observed data. In lieu of data, was decided to compare modeled travel time with observed travel time for a sample of predefined origins and destinations. Several trips (in Fall, 2010) were made inside the RDCO region to observe travel time. Table 4.3 compares the modeled and observed travel time for these selected routes. The routes that were used for the purpose can be found in Appendix E. Each trip started and ended at a zonal centroid as per VISUM modelled network. For easy recognition of the zone location, nearest road intersection is mentioned in the Table 4.2 and the zone numbers are given in parenthesis. The survey trips were made in the AM periods (6-9 AM) and in the early part of AM period the traffic volume are usually less and vehicles can move at a higher speed compared to the speed during midday. That is why the observed travel times in 2010 were very close to the modeled travel time of 2006. Also it can be seen that for route #5, the observed travel time was higher than the modeled travel time. The reason behind higher travel time could be the significant increase of Highway#97 traffic volume between 2006 and 2010.

Table 4.3 Comparison of observed and modeled travel time

Route No.	From (zone)	To (zone)	Modeled Travel time (min) (2006)	Observed Travel time (min) (2010)	Error (%)
1	Pandosy St./ Cadder Ave (274)	Hwy 33/Hollywood Rd (493)	10min 9 sec	10 min 50sec	-6.73
2	Hwy 33/Hollywood Rd (493)	Highway 97/Abott St. (459)	10 min 32 sec	10 min 40sec	-1.27
3	Highway97/Pandosy St. (460)	Keith Rd/ Lakeshore Rd (309)	10 min 50 sec	10 min 30 sec	3.08
4	Gordon Dr/ McClure Rd (350)	Cadder Ave/Long St (277)	11min 8 sec	10 min 57 sec	-7.34
5	Cadder Ave/Long St (277)	Highway97/Lodge Rd (414)	23 min 48 sec	26 min 20sec	-10.6
6	Highway97/Glenmore Rd (412)	UBC Okanagan (40)	23 min 31 sec	24 min 2 sec	-2.2

4.2.3 Step 3 results: mode choice

Mode choice selects a mode that will be used for a particular trip. It is related to individual traveller's characteristics and mode specific attributes. Due to a lack of disaggregate level data, it was decided to build aggregate based mode choice model for auto. Mode choice variable values were obtained from the 2006 Census data and the average value of each variable was calculated for each DA. Initially linear regression analysis was performed to calculate the percentage of auto user for each DA. But the R^2 value of the linear model was found to be very low. So it was decided to build a non-linear regression model to calculate percentage of auto user. From the literature (McFadden, 1974; Domencich & McFadden, 1975; Ben-Akiva & Lerman, 1985; Koppelman & Bhat, 2006) it was observed that the utility for each mode is exponentially related to mode specific attributes. So in this research, a non-linear exponential form was considered.

In the previous chapter Table 3.4 listed all the candidate variables for mode choice modeling. While selecting the candidate variables, three criteria were considered: (1) the logic (i.e. +/-) of the estimated variable had to be associated intuitively with auto usage; (2) the t-statistics of each variable should be greater than 1.96 at 95% confident limit; and, (3) the variables should not be correlated. A generalized regression model was used to develop the mode choice model and after the regression analysis following model was obtained:

$$Auto\ user\ (\%) = 33.13.e^{-0.000925*BStopD+0.002349*BH+0.3815*VHH+0.0024*Pop45\_64} \quad (4.1)$$

where,

Auto user (%) = percentage of auto user from a zone;
BStopD = bus stop density (i.e. no. of bus stops/ TAZ area in KM^2)
BH = bus headways (in minutes);
VHH = vehicles per household; and,
Pop45_64 = percentage of population between ages 45 and 64.

For degrees of freedom (df) of 230, the Pearson chi-square (λ^2) of the model was 205.66 which is less than ${\lambda^2}_{0.05,230} = 266.38$. This implies the model fits the observed data. The details of the mode choice model can be found in Appendix F.

In the next step the mode choice model was applied to calculate the percentage of auto user in each TAZ by using TAZ specific mode choice attributes. The average RDCO auto user was calculated as 86.1%, which means on an average 86.1% of total RDCO trips will be made by auto. This modeled value of 86.1% is very close to the observed 87.3% auto user in 2007, which indicates sufficient predictability of the 2006 mode choice model. Next, using the percentage auto user, the total number of auto trips produced from each TAZ was calculated, in order to build an O-D auto trip table to perform trip assignment.

4.2.4 Step 4 results: trip assignment

The last step of four-step transportation planning model, trip assignment, assigns traffic to the road network. As the purpose of this research was to build an auto-based transportation model, only auto trips were assigned to the road network to obtain auto volumes on each road. Transit assignment and development of a transit model could be the topic for future research. For the assignment purpose, ‘Equilibrium Assignment’ method was selected and also to obtain better results ‘Go to Operation’ function was used.

After performing the trip assignment, the traffic volume on road links were compared to observed data obtained from the BC Ministry of Transportation and the City of Kelowna. Table 4.4 compares traffic volume at some signalized intersections as well as compares the screenline counts. It can be seen that modeled traffic volume was less than the observed volume. The reason may be the presence of additional through-traffic volume on Highways not accounted for in the model. Although the model considered RDNO and Penticton as the external zones, other traffic (passenger cars and freights truck) does come from other external regions and use the RDCO highway facilities to go to US/Canada border. This unaccounted traffic volume on highways could be the reason of higher observed volume. But as the percentage error for intersections was less than +/-30% (Lin & Navin, 1999), the trip assignment results are acceptable. This table also compares total traffic coming from/to the Lake country and it can be seen that the observed and the model values were very close. Finally it compares the traffic coming from/to the Westside. It can be seen that that the percentage error was little bit higher. For Westside, the observed traffic volumes were collected in 2007; and, as between 2006 and 2007 Westside regions went through major developments the difference in Westside traffic prediction is considerable. Also after combining the total modeled and observed volume, it was found that total error was 12.83% which can be considered as acceptable. A screenshot of the assigned auto-network can be found in Appendix G.

Table 4.4 Comparison of modeled and observed traffic volume

Type	Location	Modeled	Observed	Diff.	Error (%)
Intersections	Highway 97 & University way	5828	6683	-855	-14.67
	Highway 97 & Cooper Rd	7592	8079	-487	-6.41
	Highway 97 & Spall Rd	8913	10244	-1331	-14.93
	Highway 33& Gerstmer Rd	4708	4957	-249	-5.29
	Highway 33 & Rutland Rd	3694	4250	-556	-15.05
Screenline	From Lake Country to Kelowna (Southbound traffic)	3520	3706	-186	-5.28
	From Kelowna to Lake Country (Northbound traffic)	2009	2278	-269	-13.39
	Westbound trips across the bridge	1909	2194	-285	-14.93
	Eastbound trips across the bridge	5690	7101	-1411	-24.8
	Total	43863	49492	-5629	-12.83

After the auto assignment, it was possible to calculate total auto volume, VKT, VC, travel speed, and journey speed on each link. Among these attributes vehicle kilometres travelled (VKT), volume-capacity ratio (VC) and travel speed (SPED) are the most important exposure variables. For the development of CPMs it is very important to calculate total VKT for each TAZ. VKT for each link can be obtained directly from VISUM, but it doesn't aggregate total VKT inside of each TAZ. So, it was done by manually by ArcGIS 9.3. If a link goes through two TAZs and if the original length of the link is used to calculate VKT for each TAZ, then it overestimates VKT because the length used for calculating VKT does not represent the actual length of the link inside of the TAZ, rather it uses the original length of the link. So dividing a link based on the zonal boundary is very important. 'Intersection' function of ArcGIS was used to divide links crossing TAZ boundary. Next, the lengths of new link segments were calculated in ArcGIS. Finally, equation 3.1 was used to calculate VKT and the total VKT in 2006 was calculated as 520,950.

Another aggregation bias arises when an attribute value is averaged across the entire zone. The average VC and travel speed (SPED) of each TAZ was calculated using four different methods; which are: (1) simple average; (2) weighted by VKT; (3) weighted by traffic volume; and, (4) weighted by link length. It was found that the values obtained weighted by VKT and traffic volume were very close and the values weighted by length were always less than the values weighted by VKT and volume. It was important to avoid correlation between variables while calculating weighted average. As VC is the ratio between volume and capacity, average value of VC was calculated weighted by link length; and, as travel speed (SPED) is the ratio between distance and time, average value of SPED was calculated weighted by traffic volume. Following Figure 4.5 shows the values of VC in different RDCO zones and compares VC to give an idea about the congestion level indifferent part of the RDCO. It can be seen that VC is relatively higher in central Kelowna and also in some parts of Lake Country. As Central Kelowna is the main activity centre of the RDCO, higher VC is intuitive. For the Lake Country, the reason could be the number of fewer roads and higher traffic volume which made the value of VC higher.

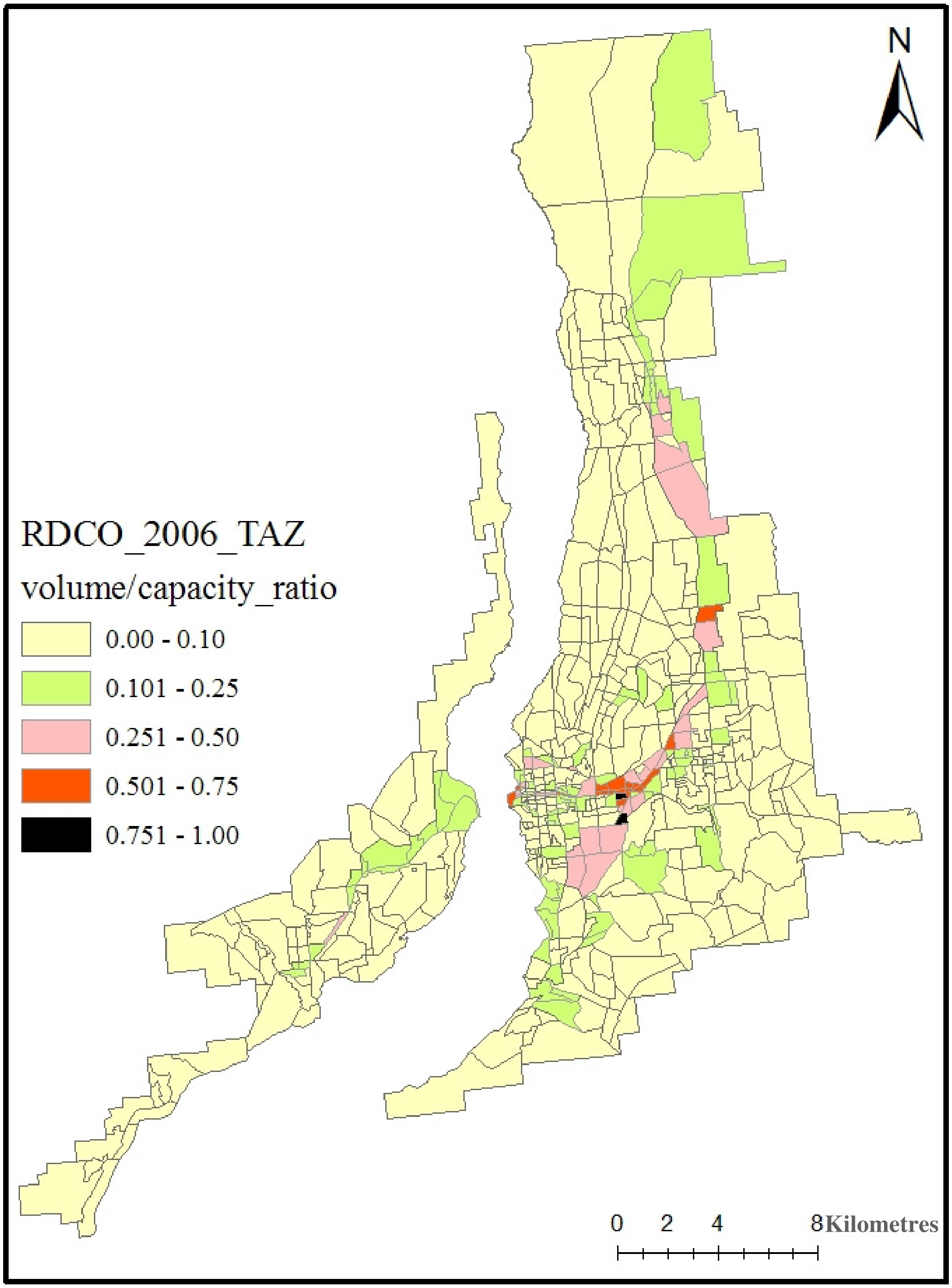


Figure 4.5 Volume-capacity (VC) ratio in 2006

4.3 Transit service and road network improvements in 2020

One of the main objectives of this research is to build RDCO transportation planning models for year 2020. Even though the research is focused to quantify the road safety benefits of increased transit usage, it is also very important to consider future road network improvements, which may add new roads to the existing road network and thus increase VKT and/or decrease VC. So this research considered both of transit and road improvements while building the 2020 RDCO transportation planning model. But the improvements were considered in different combinations, which were necessary because of the uncertainty involved in improving both road network and transit service at the same time. The uncertainty arises due to the need of huge financial support and policy changes. Four different sub-scenarios were considered in 2020 model. The sub-scenarios are: (1) do-nothing (no road network and no transit service improvement) scenario; (2) scenario with only road network improvement but no transit service improvements; (3) scenario with transit service improvements but no road network improvement; and, (4) scenario with improvements of both road network and transit service.

4.3.1 Kelowna Regional Transit System

RDCO transit service is operated by the Kelowna Regional Transit System which started in 1977. This transit service operates in each of the RDCO municipalities. It provides fixed route transit services and Flexi-route 'HandyDART' services for special needs passengers. Transit ridership in the region has increased significantly since 1977. The following Figure 4.6 shows the trend in transit ridership in the region. The ridership was stable from 1979 to 1990, even though from 1981 to 1987 overall ridership declined. During 1987 to 2000 due to increase in transit services, transit ridership increased significantly. In 2007 total annual ridership was 3,000,000 passengers.

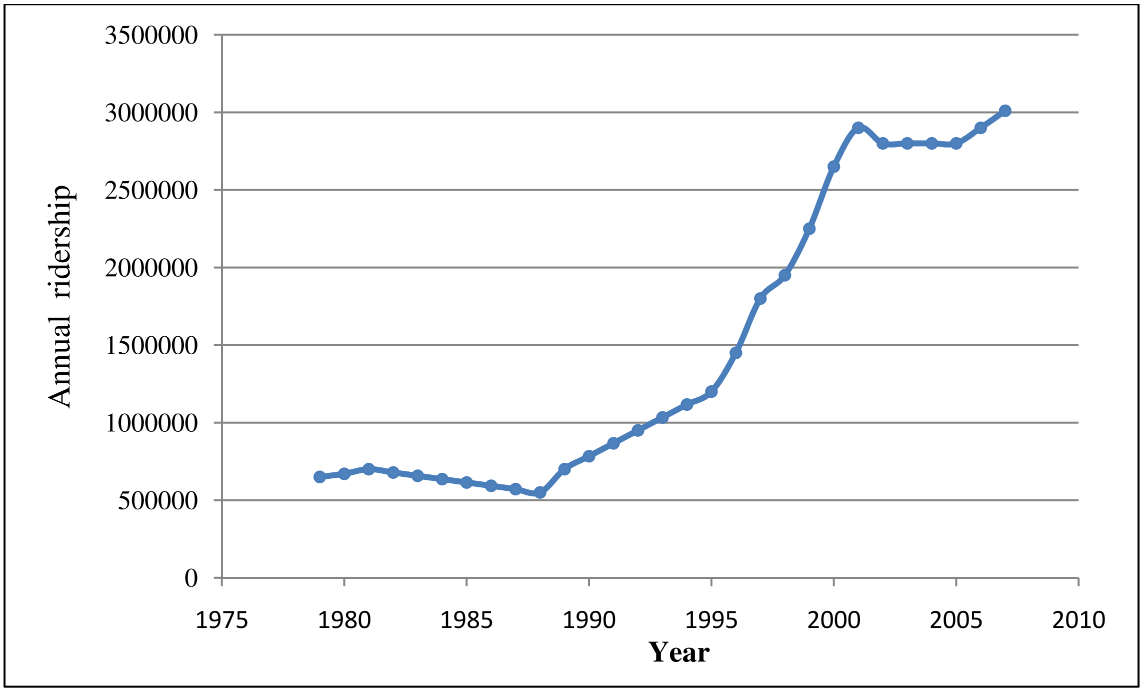


Figure 4.6 The Kelowna Regional Transit System ridership trend (IBI Group, 2005)

At present, 22 bus routes are operated in the region. Some of the buses carry higher passengers and some of them carry lower. In 2006, 33.4% of the RDCO area was covered by transit service when a 10 minutes walking distance (@ 1 m/sec walking speed) is considered. It was also found that transit covered 89% of urban areas and 24% of rural areas as well. The area coverage by transit service is given in Appendix I. Table 4.5 compares weekday transit ridership for different buses and also for different years. From 2006 to 2008, total transit ridership increased about 26%. One of the main reasons was the introduction of the 97-Express bus and the U-Pass program for University of British Columbia Okanagan (UBCO) students. A typical 2006 weekday bus ridership count is given in Appendix J.

Table 4.5 Comparison of weekday transit ridership in the RDCO

Bus route no.	Bus route name	October, 2005	October, 2006	November, 2007	October, 2008	Growth (%) between 2005-2008
1	Lakeshore	2,036	2,056	2,241	2,468	21.22
2	North End Shuttle	143	153	172	224	57
3	Dilworth	122	99	119	146	20
7	Glenmore	919	993	985	1,260	37.11
8	University	3,687	3,091	2,688	2,913	-20.99
9	Shopper Shuttle		675	134	141	n/a
10	North Rutland	2,780	2,979	2,821	2,710	-3
11	South Rutland	1,553	1,419	2,169	2,277	47
12	McCulloch	135	102	113	110	-18.52
14	Black Mountain		47	72	110	n/a
15	Crawford		24	29	31	n/a
16	Southwest Mission		90	102	136	n/a
20	Lakeview	369	308	389	426	15
21	Glenrosa	1,368	1,367	1,570	1,610	17.69
22	Peachland		84	28	21	n/a
23	Lake Country	354	448	421	576	62.71
24	Shannon Lake	483	459	518	590	22
25	East Boundary			85	99	n/a
27	Horizon	61		54	42	-31.15
28	Smith Creek	55	61	36	41	-25.46
29	Bear Creek		59	20	15	n/a
97	97 Express			1,905	2,709	n/a
	School Special	310	237	234	253	-18
	Total	15,326	15,055	16,903	18,925	23.48

In 2005 the Central Okanagan Smart Transit Plan was prepared, which reviewed the existing transit services and also proposed major transit improvements by 2020. Some of the major proposed transit improvements are given below:

- Introduction of queue jump lane to reduce travel time.
- Left turn modification to reduce delays at some specified intersections.
- Traffic signal priority to reduce delay at some selected Highway 97 intersections.
- Bus stop inventory of the bus stops used by the Kelowna Regional Transit System.

- Improvements of the access and egress facilities at the Orchard Park Mall bus exchange.
- Long term vision with different types of bus services, such as: (1) high frequency rapid transit with limited stops; (2) high frequency major town centre connector service; (3) medium frequency feeder service and, (4) low frequency community service.

The 2005 transit master plan also proposed bus headways for different buses. The following Table 4.6 (source: IBI group, 2005) summarizes AM period bus headways for year 2005 as well as for year 2020.

Table 4.6 Proposed transit headways in 2020 (IBI group, 2005)

Bus route no.	Route name	Type	Headway 2005 (AM period)	Headway 2009-21 (AM period)
1	Lakeshore	Connector	15	10
2	North End Shuttle	Local	30	10
3	Dilworth	Local	30	30
7	Glenmore	Local	30	20
8	University	Connector	15	20
9C	Central Rutland/ Black Mt.	Community	30	10
10	North Rutland	Connector	15	
10E	University BRT	BRT	30	10
11	South Rutland	Community	15	10
12	McCulloch	Community	30	20
19	Rutland Night	Local	Replaced by Community Service	Replaced by Community Service
20	Lakeshore	Local	30	30
21E	Westside BRT	BRT		10
21C	Glenrosa Shuttle	Community	20	10
22	Peachland	Community	45	30
23	Lake Country/ University	Community	40	30
24	Shannon Lake	Local	30	30
26	Westside Night	Local		Replaced by Community Service
27	Horizon	Community	30	30
28	Smith Creek	Community	30	30
29	Oakaview/Crawford	Community		30
30	McKenzie/Ellison	Community		30
31	Clifton/ Glenmore Valley	Community		30
32	McDougall	Community		30

In 2007 BC Transit reviewed the RDCO transit service and proposed some additional new improvements. The major improvements of the reviewed plan are listed below:

4.3.1.1 Service proposals for the City of Kelowna

- Additional UBC-Okanagan express trips
- Introductory bus rapid transit (BRT): downtown Kelowna to UBC-Okanagan
- Reallocation to introduce shopper shuttle#9
- Reallocation to combine South Rutland#11 and Central Rutland#9
- Service reliability improvements of Glenmore # 7
- Introduction of 15-minute commuter service to Glenmore area
- Route adjustments of McCulloch# 12 and North Rutland #10
- Additional Sunday service on University#8
- Additional HandyDART van service (shared with regional district)
- Increased evening, weekend and holiday services
- Introduction of Ellison / McKenzie community bus (shared with regional district)
- Community bus extension to Pandosy town centre
- Extension of Dilworth#3 route to Magic Estates

4.3.1.2 Proposals for the District of Lake Country

- Increased evening, Sunday and holiday service
- Introduction of HandyDART service
- Additional weekday 15-minute commuter service
- Introduction of Oyama / Copper Hill bus route

4.3.1.3 Service Proposals for the West side

- Additional morning Westside UBC-O express
- Additional morning Peachland commuter service
- Expansion of the existing Horizon#27 and Smith Creek#28
- Proposed East Boundary#25 service
- Proposed extension to Bear Creek / Westside rd.
- 15-minute commuter service for Lakeview #20 and Shannon Lake #24

- Re-routing of Glenrosa#21 via Mt. Boucherie complex
- Westside and Peachland extended evening, weekend and holiday service
- Additional HandyDART service
- Bus rapid transit (BRT) phase III: Westbank to downtown Kelowna

Based on the changes in 2020 transit routes, following Figure 4.7 was prepared which shows the 2006 transit routes as well as highlights those routes which will be re-routed or introduced. A map was also produced to show transit coverage map and it was found that for 10 minutes walking @ 1m/sec speed, transit routes in 2020 will cover 46% of the total RDCO area. Following Table 4.7 gives an idea about transit service changes between 2006 and 2020. Some of the values are from year 2003 as shown in the table.

Table 4.7 Transit service comparison between 2006 and 2020*

Parameters	2006	2020
Service hours	112,000 (1)	N/A
Route km	282 (1)	427 (1)
Bus stops	802 (1)	1281 (3)
Average headways	24.7 (1)	17.7 (1)
Average travel time	28.8 (2)	N/A
No. of bus	37 (1)	N/A
Area coverage for 10 minutes walking distance	33.4% of RDCO (3)	38.5% of RDCO (3)
No. of bus stops/ route km	2.84 (1)	3.0 (3)
Passenger/year	2,750,000 (1)	5,087,500

* Sources: (1) BC Transit (2) 2007 HH Survey for RDNO, RDCO (3) Estimated (AA)

4.3.2 Future road improvements

It is very important to forecast future travel demand and to plan necessary improvements to accommodate future demand. In case of the RDCO, each municipality have reviewed future transportation needs and planned for necessary changes. Details of the road improvement plans can be found in the Official Community Plan (OCP) of each municipality. Primarily future road improvements include construction of new roads and addition of new road lanes (i.e. widening of existing roads). For this research, future transportation network map was obtained from each city and Figure 4.8 was produced to highlight changes in future roads. The roads in black color represent future roads which will be constructed or modified in 2020. The roads in red color are the existing roads in 2006. It can be seen that lots of road improvements are planned for 2020 and this research will quantify how these road improvements will affect the road safety.

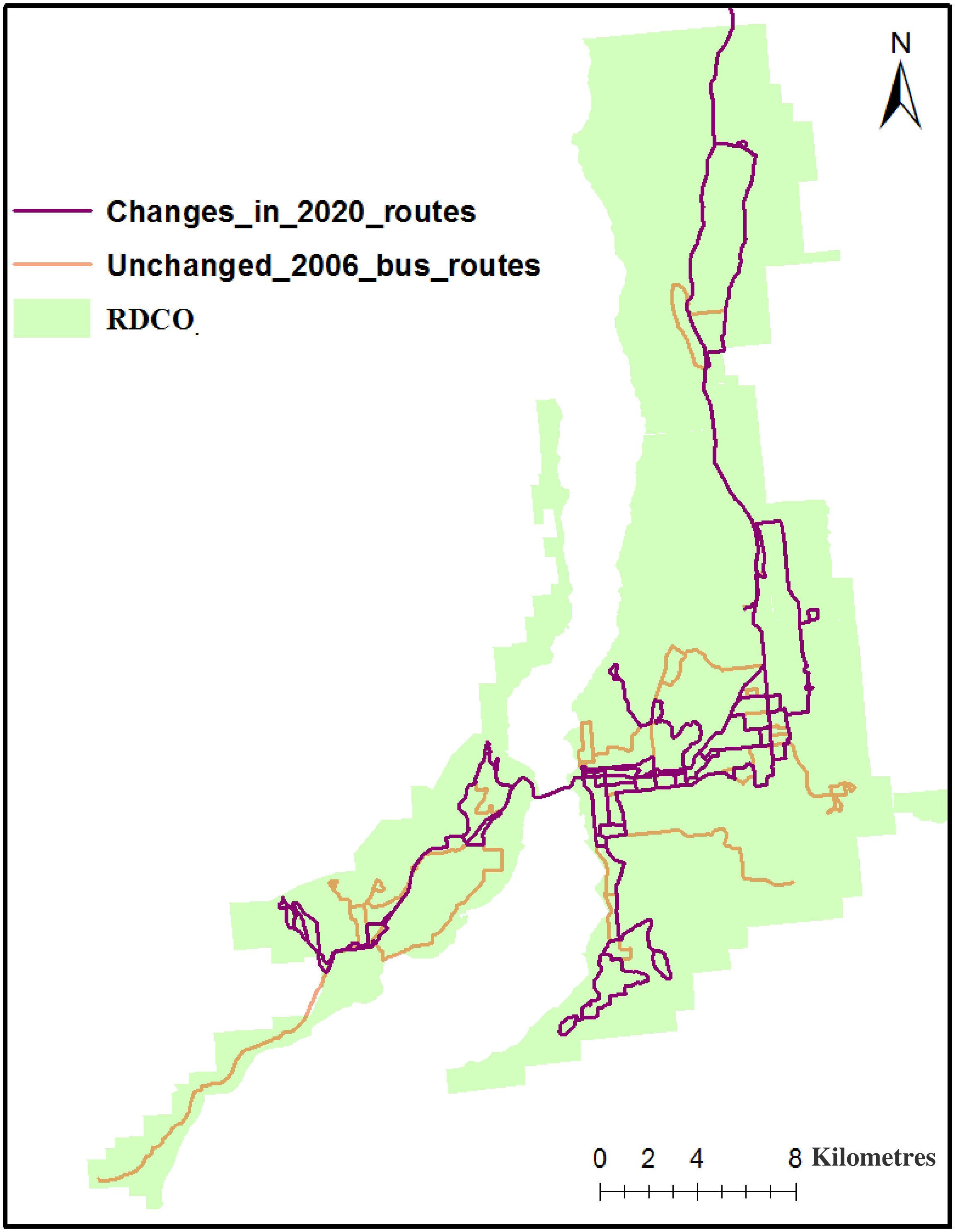


Figure 4.7 Changes in 2020 transit routes

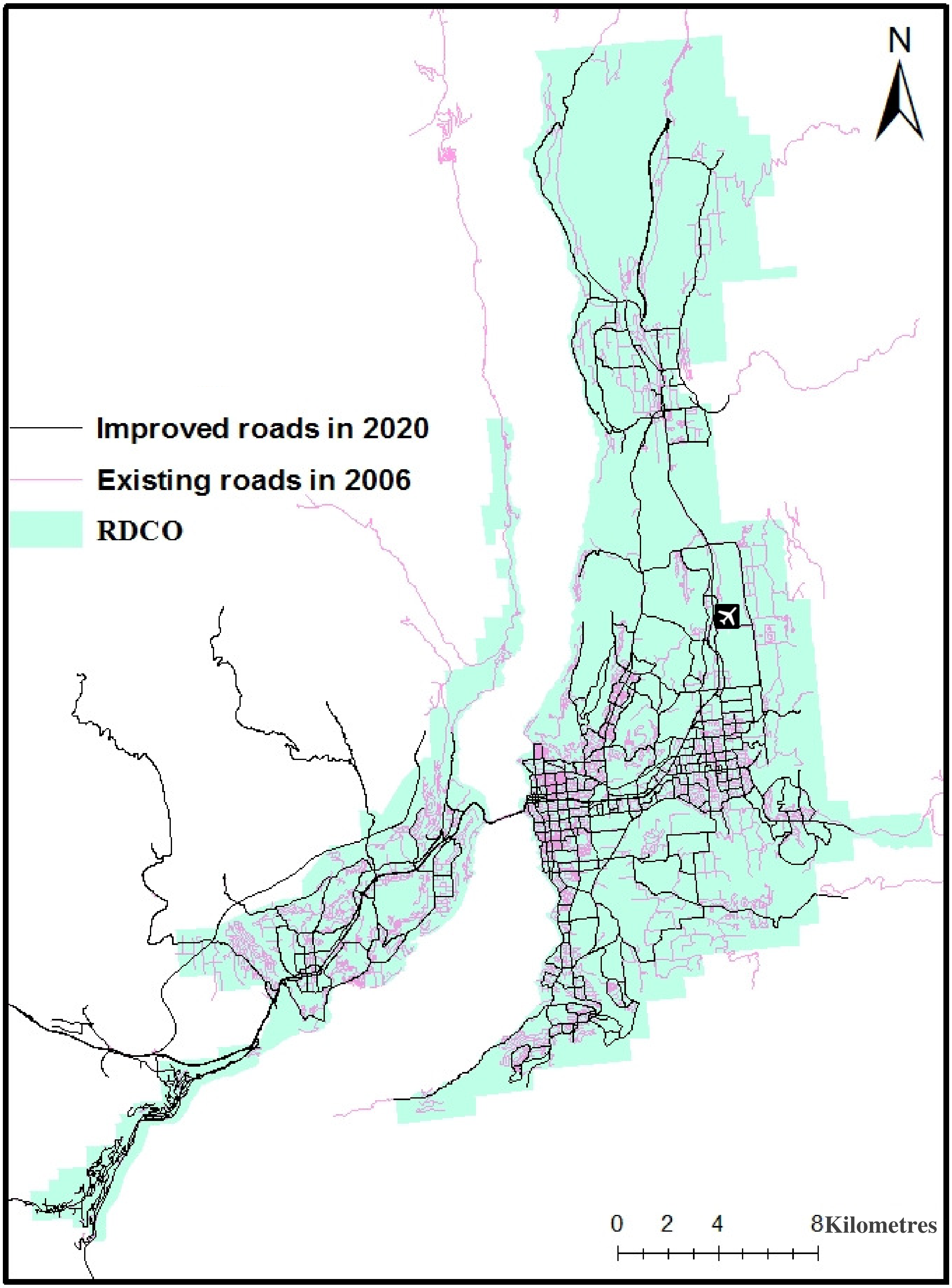


Figure 4.8 Road networks of the RDCO in 2020

4.4 The 2020 RDCO transportation planning model

As discussed earlier, this research considered both of transit and road improvements while building the 2020 RDCO transportation planning model. Four different sub-scenarios were considered to develop 2020 transportation planning model. As each sub-scenario planning model involved the same steps, the results are discussed below by describing the results of each step:

4.4.1 Trip generation results

For trip generation process, the 2006 variables and their trip generation rate were used. Based on the 2020 OCP of each city, the land-use related trip generation variables (i.e. residential area, commercial area etc.) were calculated for each TAZ. As the number of school going children was also in 2006 model, it was necessary to predict the number of school-going students in year 2020. For each of the RCDO school (i.e. school district #23), a careful observation was made and it was observed that the number of students decreased from 2004 to 2006. From the population projection made by the City of Kelowna (City of Kelowna, 2009) it was found that the percentage of population between ages 0 to 19 would decrease by 3.9 percent from 2006 to 2020. So, the number of students in 2020 was predicted by applying 4% decrease (compared to 2006) in student numbers.

As the 2006 base model includes two external zones (i.e. RDNO and Penticton), prediction of 2020 traffic volume coming from these two external zones were also important. The Okanagan Valley Transportation Plan (OVTP) gives the inter-regional daily traffic volume for year 1994 and year 2020. The traffic volume from OVTP can be found in Appendix K. The total number of daily trips coming from/to these external zones was obtained from the OVTP model. But as the VISUM model was for AM period, it was necessary to calculate AM period traffic volume for these external zones. Assuming a linear traffic volume growth, the daily traffic volume in 2007 was calculated from the OVTP model and it was compared with the AM period traffic volume found in 2007 Okanagan household survey. For Penticton it was found that in AM period, 1.9% of daily trips came from Penticton to the RDCO and 17.09% of daily trips went from the RDCO to Penticton. In case of RDNO in the AM period, 41.7% of daily trips came from RDNO to the RDCO and 27.6% went from the RDCO to RDNO. Using the percent of AM period traffic, total number of 2020 external Am period trips was calculated. Finally using the trip generation rate, total number of trips generated from each zone was calculated. Total number of produced trip was calculated as 215,522 and attracted trips as 313,628. Following Figure 4.9 shows the trip generation densities in 2020. The darker color represents higher trip density and lighter color represents lower density. It can be seen that in 2020 the town centres of each municipality will generate higher number of trips compared to other part of the municipality.

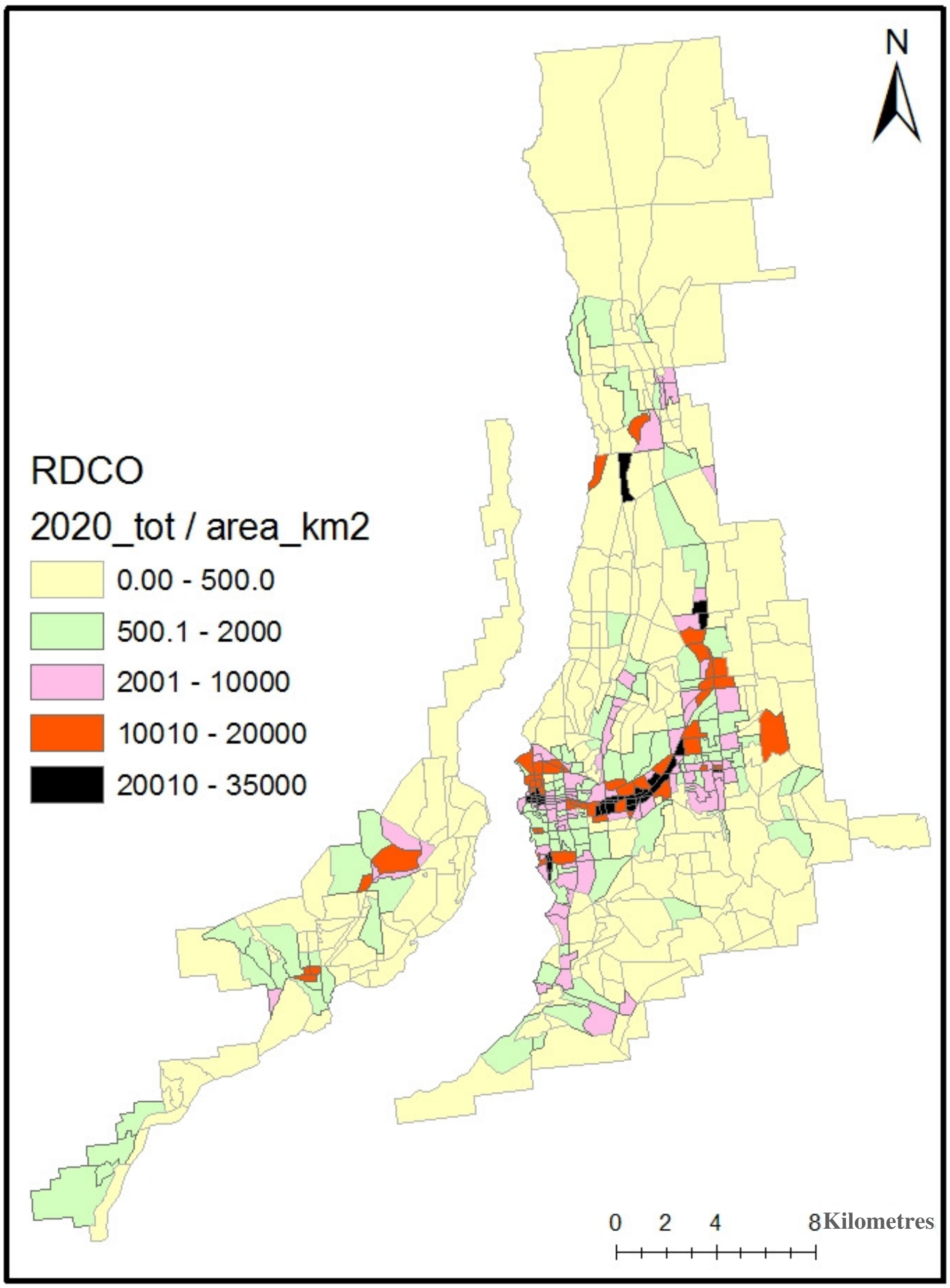


Figure 4.9 Trip generation density in 2020

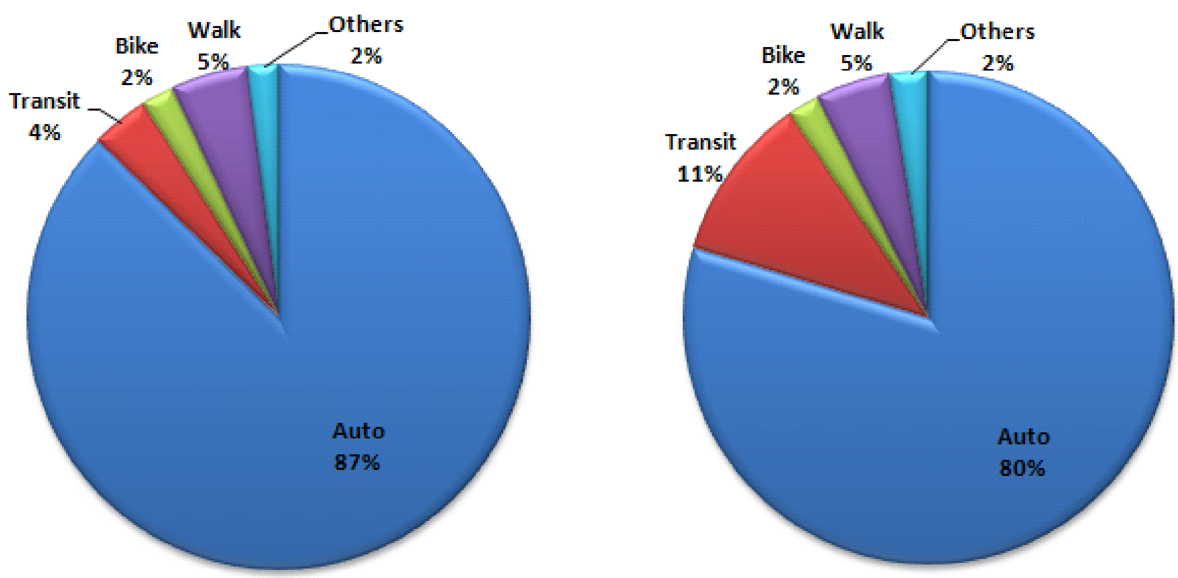
4.4.2 Trip distribution results

In the next step, trips produced from each TAZ were distributed to other zones by Gravity model. 'Travel time' was considered as the impedance and the same friction factor function was used as 2006 model. After several refinements, an O-D table was prepared to show number of trips between each pair of zones.

4.4.3 Mode choice results

To predict the percentage of auto user in each TAZ, mode choice equation 4.1 was used. The equation contains four different variables, namely (1) bus stop density (BStopD); (2) bus headway (BH); (3) vehicle per household (VHH); and, (4) percentage of population ages between 45 and 64 (Pop45_64). For do-nothing scenario, even though there will be no transportation and road network improvements, the age distribution of the population will change. From a population projection (City of Kelowna, 2009) it was found that for the RDCO 'Pop45-64' would decrease by 1.2% and for the City of Kelowna it would decrease by 1.8% from 2006 to 2020. So for each TAZ, the 2020 'Pop45_64' was calculated using the above mentioned percentage reduction. These predicted 'Pop45-64' values were used for each of the four 2020 scenarios.

Except Pop45_64 variable, for Scenario#1 (i.e. do-nothing) and Scenario#2 (only road improvements), mode choice variable values were kept same as year 2006 and the auto user percentage for each TAZ was calculated using equation 4.1. The average auto user percentage for both scenarios was calculated as 86.8%. But Scenario-3 and Scenario-4 consider transit improvement and hence, the values of mode choice variables were different than Scenario-1 & 2. Transit improvements can change the value of bus stop density, bus headway, and vehicle ownership. The bus headway was obtained from the proposed 2020 transit master plan which is listed in Table 4.5. For each TAZ the average transit headway was calculated and for the whole RDCO average headway in 2020 was calculated as 17.8 minutes compared to 2006 average headway of 24.7 minutes. The 2020 transit master plan doesn't give any detailed plan about bus stop improvements. It discusses about the inventory and improvement of bus stops within the region. So it was not possible to calculate the exact bus stop density in 2020. An increase of 2% annually was assumed to predict 2020 bus stop density. Also for the proposed new bus routes, bus stops were considered 500 m apart for each direction. Transit improvement has the potential to reduce average auto usage and vehicle ownership by 5% to 10% (Kim & Kim, 2004; Litman, 2007; Litman, 2010). Again, auto operating cost and auto parking cost are the two important mode choice variables which were not included in the 2006 mode share equation. But auto operating cost will increase due to higher fossil fuel price. Also in future the availability of free parking area in the RDCO may become less and people may have to spend more money for auto parking. These two cost issues may discourage people to use fewer autos. Considering all the facts, it was assumed that due to transit improvements vehicle ownership would decrease by 8% over 14 years (i.e. 2006-2020); which means 0.57% annual decrease. Finally using the equation 4.1, the percentage of auto user for each TAZ was calculated and the average auto user for the RDCO was calculated as 79.6%. It can be seen that due to transit improvements, auto share reduced from 86.8% to 79.6%. If the percentages of other mode users remain constant, then transit improvement will increase by 7.2%. Following Figure 4.10 shows the mode share for different modes in 2006 and 2020.



(a) 2006 mode share (b) 2020 mode share

Figure 4.10 Mode share percentages for different modes in the RDCO

Also in case of external zones, it was assumed that in 2020 there will be no transit service between the RDCO and Penticton; and all trips coming from/to the Penticton would be made by auto. In case of Vernon, it was found that in 2009 AM period, by using transit (bus#90) an average of 70 persons made their trips from the RDNO to the RDCO; which means 2.6% of total AM trips were made by transit. As the demand for this route is increasing, it was assumed that in 2020 one extra bus would be added to the existing transit service and 100 persons would come from the RDNO to the RDCO. These 100 trips will be 2.95% of the total trips (= 3,390 trips) in 2020. So it was assumed that 97.05% of total trips coming from the RDNO to the RDCO would be made by auto.

4.4.4 Trip assignment results

As Scenario-1 & 3 do not include any road improvements, the 2006 road network was used for their assignment process. But both of the Scenario#2 and Scenario# 4 include road improvements. Using the 2020 road network plan for each city, new improved RDCO road network map was built and used for the assignment purpose of Scenario-2 & 4. After assigning auto trips to road network, different exposure variables such as traffic volume, VKT, VC, SPED was calculated. Following Table 4.8 summarizes the exposure variables for each of the sub-scenarios in 2020. Also a sample distribution of VC can be found in Appendix L.

Table 4.8 Comparison of the RDCO transportation planning model outputs

Scenario	**Total VKT** (x 1000)	**Total TLKM**	**Avg. INTD**	**Avg. urban VC/TAZ**	**Avg. rural VC/TAZ**	**% of Auto user/TAZ**
2006 Scenario	520.95	3,269.02	38	0.139	0.07	86.76
2020 Do-Nothing (Scenario #1)	1,385.57	3,269.02	38	0.285	0.228	86.76
2020 Road Impv. (Scenario #2)	1,422.21	5,850.24	40	0.225	0.219	86.76
2020 Transit Impv. (Scenario #3)	1,270.16	3,269.02	38	0.261	0.209	79.56
2020 Road+ Transit Impv. (Scenario #4)	1,306.94	5,850.24	40	0.121	0.114	79.56

It can be seen that Scenario #2 has the highest VKT while, Scenario#3 has the lowest. Scenario#2 is about road improvements which will add new roads to the existing road network. As a result of that length of total road will increase and consequently VKT will increase.

Scenario# 3 has no road improvements and hence, there is no possibility of increasing VKT. But Scenario# 3 includes transit improvement which can reduce the total number of auto trips and hence, can reduce VKT. On the other hand, Scenario# 4 includes both road improvements and transit improvements and that is why the VKT in this scenario is in between the VKT of Scenario# 2 and 3. The next parameter TLKM represents the length of road lanes and for road improvement scenarios it has the highest value of 5,850.25 km. In case of VC, it can be seen that the average value of VC is very low. The reason of having low VC could be due to aggregation error. More research is needed to refine the value of VC. But from the table it can be seen that VC will decrease due to road improvement as well as for transit improvement. The highest reduction (almost 50%) of VC will occur when both of road and transit improvements will occur at the same period. Road improvements will increase the road capacity and on the other hand transit improvement will decrease the auto volume. For their combined effect the reduction of VC will be higher compared to other scenarios. It can be also noted that in Scenarios#2, for urban areas road improvement will reduce higher VC than that of rural areas. In urban areas traffic volume on roads is higher and addition of a new lane can significantly reduce VC while in case of rural area it will have less reduction due to less traffic volume in rural areas. If we look only at transit improvement, it can be seen that transit improvement alone can reduce VKT by 8.3% and VC by 2%. The reduction of VKT and VC will reduce road congestion and can decrease the travel time of road users. It also has the potential to reduce road collisions which will be quantified by developing community-based macro-level collision prediction models.

4.5 Development of community-based macro-level collision prediction models

Based on the methodology described in Chapter three, AM period (i.e. 6am-9am) CPMs were developed by using 2004-06 collisions data and 2006 VISUM model outputs. The correlations between CPM input variables were checked and they were not found to be correlated. Also it was found that collision data followed Negative Binomial distribution. As the main objective of the research is to compare the road safety condition of 2020 sub-scenarios, it was decided to develop only exposure based CPMs. A total of 15 CPMs where developed. Following Table 4.9 and Table 4.10 lists all CPMs along with their goodness of fit and t-statistics of variables.

Table 4.9 2006 urban collision prediction models for the RDCO

Model No.	Model Form	κ	DoF	Pearson χ^2	SD	$\chi^2_{0.05, dof}$	t-Stat
	Urban CPMs (total collisions)						
01	$TAM3 = 0.02213.VKT^{0.7289}$	0.909	235	198.11	202.00	271.76	Const.: -8.67 VKT: 11.19
02	$TAM3 = 0.01535.VKT^{0.724}.e^{0.00562INTD}$	0.874	234	177.15	189.88	270.68	Const.: -9.10 VKT: 11.14 INTD: 2.94
03	$TAM3 = 1.01.TLKM^{0.5904}$	1.09	223	256.88	253.04	258.84	Const.: 0.07 TLKM: 6.24
	Urban CPMs (severe collisions)						
04	$SAM3 = 0.004214.VKT^{0.8272}$	0.885	235	218.67	197.47	271.76	Const.: -9.07 VKT: 9.59
05	$SAM3 = 0.002419.VKT^{0.8341}.e^{0.00699\,INTD}$	0.848	234	200.55	185.38	270.68	Const.: -9.38 VKT: 9.57 INTD: 3.13
06	$SAM3 = 0.3749.TLKM^{0.597}$	0.985	226	251.66	221.41	262.07	Const.: -5.14 TLKM: 5.27

Table 4.10 2006 rural collision prediction models for the RDCO

Model No.	Model Form	κ	DoF	Pearson χ^2	SD	$\chi^2_{0.05,\ dof}$	t-Stat
	Rural CPMs (total collisions)						
07	$TAM3 = 0.07401.VKT^{0.4488}$	0.954	210	216.35	196.54	244.81	Const.: -6.62 VKT: 7.65
08	$TAM3 = 0.06001.VKT^{0.4313}.e^{0.0163\ INTD}$	0.925	209	211.89	187.53	243.73	Const.: -6.96 VKT: 7.45 INTD: 2.99
09	$TAM3 = 0.2612.TLKM^{0.725}$	1.00	206	216.84	206.40	240.23	Const.: -5.03 TLKM: 6.23
10	$TAM3 = 0.2139.TLKM^{0.636}.e^{0.02379\ INTD}$	1.03	208	197.61	216.87	242.65	Const.: -5.4 TLKM: 5.48 INTD: 4.34
	Rural CPMs (severe collisions)						
11	$SAM3 = 0.002383.VKT^{0.7949}$	0.676	212	196.96	124.60	246.97	Const.: -9.98 VKT: 9.66
12	$SAM3 = 0.005483.VKT^{0.957}.e^{-0.0343SPED}$	0.662	210	178.73	118.59	244.81	Const.: -8.22 VKT: 8.96 SPED:-3.28
13	$SAM3 = 0.002251.VKT^{0.763}.e^{0.01406\ INTD}$	0.666	211	196.67	120.85	245.89	Const.: - 10.03 VKT: 9.46 INTD: 2.23
14	$SAM3 = 0.1171.TLKM^{0.591}$	0.732	211	192.15	141.46	245.89	Const.: -6.94 TLKM: 4.47
15	$SAM3 = 0.09.TLKM^{0.59}.e^{0.01375\ INTD}$	0.721	209	185.92	137.93	243.93	Const.: -7.14 TLKM: 4.42 INTD:2.21

For total AM collisions (TAM3), it can be seen that VKT has the major influence on collision predictions. Increase in VKT will result in higher collisions. The co-efficient of VKT was between 0.5 and 1 which was also found in previous researches (Tanner, 1953; Poppe, 1997a, 1997b; Lovegrove, 2007). During the development of CPMs it was found that VC was positively correlated with total collisions. This result is intuitive because if the congestion level increases, the probability of collision occurrence also increases. But the t-statistics of VC was low (i.e. <1.96) and hence, VC was not included in the developed CPMs. The lower value of VC could be the reason for being insignificant. Typically VC is a very important CPM parameter and more research should be done to address this issue.

It was also found that road collision is negatively associated with travel speed on the link. Intuitively, increased travel speeds should be associated with increased collision frequency and severity, but the opposite situation may arise due to various reasons. Lovegrove (2007) identified three possible causes which are: (1) the existence of correlation between speed and other un-identified variable that makes the result counter-intuitive; (2) the stratification of variables is not correct and is needed to be further investigated; and, (3) as the operating speed is the model output, error in predicting individual link's speed may cause the error. From observation it has been found that in the RDCO, most of the collisions occurred near the intersection and that is why the operating speed on links may not be intuitively correlated with collisions. To observe the effect of intersections in collision prediction, intersection density (INTD) was used in the CPM development process. It was found that intersection density is significantly related to total collisions as well as to severe collisions. The TLKM variable was also positively related to the number of collisions. Increase in TLKM will increase the exposure level of road users and it would increase the possibility of collisions.

4.6 Quantifying the road safety benefits among 2020 sub-scenarios

Community-based macro-level collision prediction models are very useful proactive empirical tool which relate number of collisions with neighbourhood traits. CPMs also have the capability to predict collisions at different scenarios. So, 2006 CPMs were used to predict the number of collisions in 2020. For urban total AM collisions, model # 02 (includes VKT & INTD) and for urban severe collision, model # 06 (includes VKT and INTD) were selected to predict 2020 urban collisions. For rural total collisions, model # 10 (includes VKT and INTD) and for rural severe collisions, model # 15 (includes VKT and INTD) were selected. Using these selected models, the number of total collisions (i.e. TAM3) as well as severe collisions (i.e. SAM3) was calculated for each scenario. Also the significance of the difference between sub-scenarios was calculated using equation 2.44. The results of the predicted road collisions are given in Figure 4.11 and Figure 4.12.

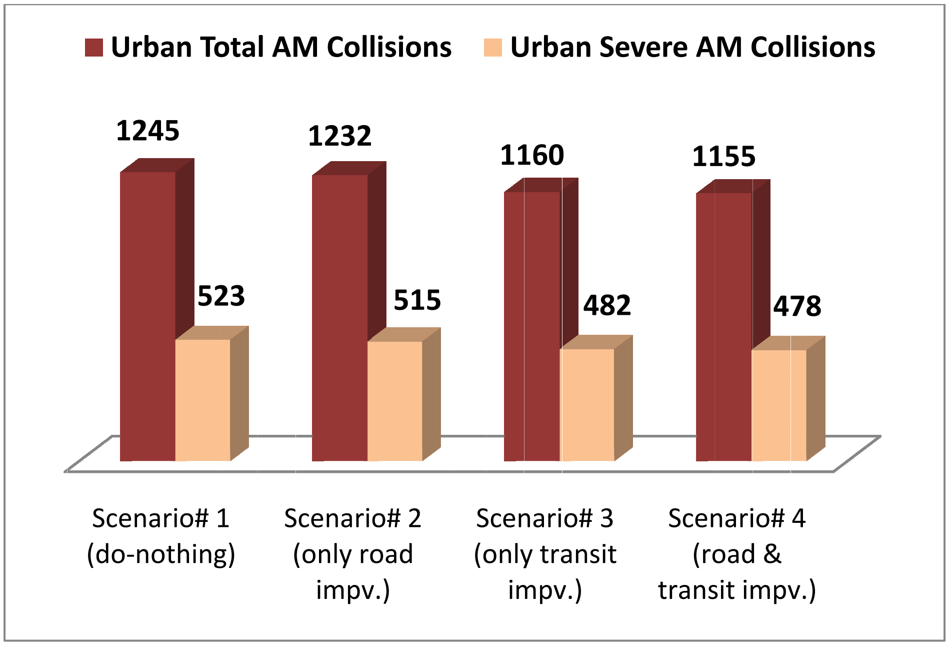


Figure 4.11 Predicted 3 years road collisions in RDCO urban areas in 2020

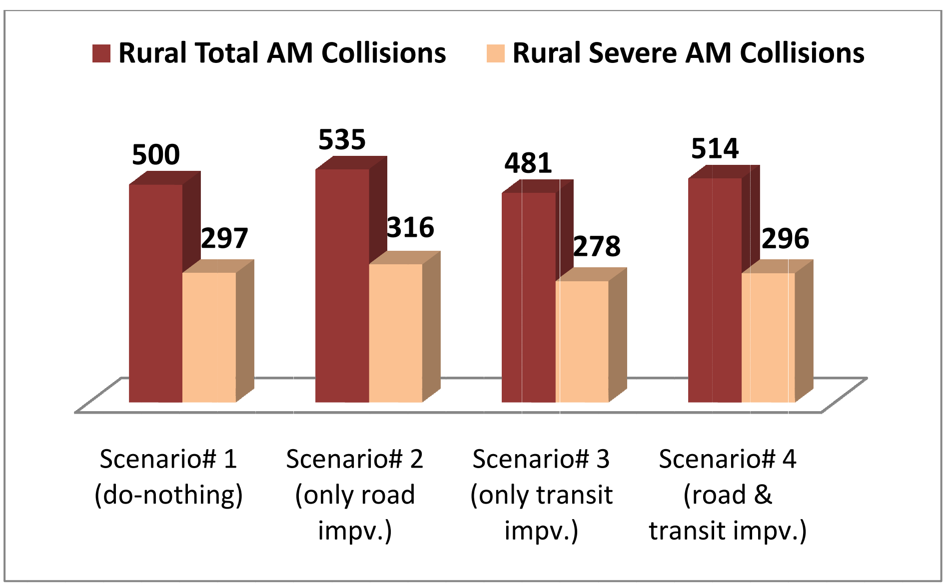


Figure 4.12 Predicted 3 years road collisions in RDCO rural areas in 2020

From Figure 4.11 it can be seen that in 2020 for 'do-nothing' scenario total number of urban collisions would be 1,245 and for road improvements scenario it would be 1,232 collisions. Due to road improvements the reduction is only 1%. From the transportation model it has been found that due to road improvements, VKT haven't changed drastically and for that reason the reduction of collisions is very few. In case of transit improvement, it could decrease total collisions by 6.8%. But if both transit and road improvements are considered together then total collisions reduction is only 7.2%. These results clearly suggests that road improvements will not have significant impact in reducing collisions. But because of reduced VKT transit improvements have the capability to reduce collisions. It can be also observed that transit improvements could reduce 7.8% of urban severe collisions.

In case of rural collisions, do-nothing scenario will experience 500 total collisions and 297 severe collisions. Scenario#2 (i.e. road improvements) will experience 535 total collisions and 316 severe collisions. It is interesting to see that due to road improvements, rural areas will experience an increase in total and severe collisions. As road improvement adds new roads to the network, it increases VKT and INTD. As VKT increase, people will be more exposed and hence, there will be more possibility of collisions. Another possible explanation for increased rural VKT due to road improvements is that these future road improvements may have the unintended impact of shifting commuting, through traffic out of urban areas (hence reduced VKT) and through rural areas that lie on the desire line between adjacent urban areas. So, for rural areas instead of reducing collisions, road improvements would increase the number of collisions. This result could be very important to transportation planner during planning of new road facilities in rural areas. On the other hand, transit improvements can decrease rural total collisions by 3.8% and rural severe collisions by 6.4%. These reductions are lower when compared to urban area. The reason for this is, transit service does not cover all rural areas and it will not significantly reduce the number of auto users where transit service is not available. So the reduction of VKT and the number of collisions in rural areas is not as much as for urban areas.

Due to road improvements, intersection density will not increase significantly. Also the affect of road improvements (includes increase in intersection density) are insignificant in reducing collisions. So another comparison was made to quantify road safety benefits by considering

CPMs having VKT only. Table 4.11 gives the number of predicted collisions for each of the 2020 sub-scenario.

Table 4.11 Predicted 2020 road collisions (considering CPMs with VKT only)

		Scenario# 1 (do-nothing)	Scenario# 2 (road impv.)	Scenario# 3 (transit impv.)	Scenario# 4 (road & transit impv.)
Urban	Total Coll./3yrs	1,280	1,251	1,193	1,171
	Severe Coll./3yrs	536	522	496	484
Rural	Total Coll./3yrs	532	562	510	539
	Severe Coll./3yrs	330	347	308	324

Again from the table it can be found that road improvements would increase rural collisions by 6%. On the other hand, transit improvement could reduce urban total collisions by 6.8% and rural total collisions by 4%. In both cases, it is clear that transit improvements have the potential to reduce total and severe collisions. On the other hand, road improvements will not significantly increase road safety; instead it would increase the number of rural collisions.

4.7 Summary

Based on the methodology, this chapter developed transportation planning model for the RDCO as well as collision prediction models. Due to the availability of Census-2006 data, year 2006 was used as the base year for the RDCO transportation planning model. The total number of trips generated from each zone was calculated using the ITE trip generation rates and it was calibrated by comparing with the 2007 observed data. In the next step trips were distributed to various destinations and it was found that average travel time in the RDCO was 15.99 minutes and average trip distance was 14.04 km. A travel time survey was conducted within the RDCO boundary and it was compared with the modeled travel time. The error in modeled travel time was found to be very small. This research also predicted a mode share equation which will calculate the percentage of auto user in each TAZ. The average auto user percentage in the RDCO was calculated as 86.14% which was very close to the observed 87.3%. In the trip assignment step, the trips were assigned to the road network and the number of traffic volume in various intersections and screenlines were compared to observe the error in prediction. Finally from the auto network various network related variables such as VKT, VC, travel speed, travel time were calculated.

This research also built transportation planning model for four different 2020 sub-scenarios. They are: (1) do nothing scenario; (2) scenario with only road improvements scenario; (3) scenario with only transit improvements scenario; and, (4) scenario with both transit and road improvements. Among these four scenarios, Scenarios #3 and #4 includes transit improvements which was associated with increased bus stop density, decreased bus headway and decreased vehicle per household. It was found that due to transit improvements, auto user percentage would drop by 7% from 87% to 80%. Table 4.7 listed different modeled network outputs and compared the changes in VKT, VC and TLKM for each scenario. It was found that Scenario # 4 has the lowest VKT and VC compared to other scenarios, because transit improvement will decrease auto volume on roads and road improvement will increase the road capacity. So due to the combined effect, Scenario # 4 can achieve 50% less VC than Scenario#1.

The research also developed collision prediction models based on 2006 VISUM model outputs. It has been found that collisions are significantly related to VKT, TLKM, and intersection density. The coefficient of CPM parameters varied depending on the types of collisions and land use (i.e. urban/rural). A total of 15 CPMs were developed for 2006 and four types (i.e. urban total, urban severe, rural total and rural severe) of CPMs were selected to predict 2020 collisions. It was found that instead of reducing collisions road improvements would decrease the number of urban collisions and increase the number of collisions in rural areas. The possible reason for higher rural collisions is that due to road improvements commuting, through traffic will use new rural roads that lie adjacent to urban areas and hence, urban VKT would decrease and rural VKT would increase. On the other hand, transit improvement can reduce urban total collisions by 6-7% and urban severe collisions by 4-8%. It can also reduce rural collisions by 4%. So it can be concluded that for both urban and rural areas, transit improvement has the potential to make roads safer.

CHAPTER 5 ROAD TRAFFIC COLLISION ANALYSIS OF THE REGIONAL DISTRICT OF CENTRAL OKANAGAN

5.1 Introduction

While previous chapters discussed the global road safety context, developed RDCO transportation planning model, and identified transit related road safety benefits, this chapter analyses the RDCO road safety situation. Specially, this chapter identifies collision prone locations (i.e. zones, highways and intersections) and finds collision patterns in the RDCO. Section 5.2 discusses the road safety context of British Columbia (BC) as well as of RDCO. Section 5.3 identifies not only CPZs but also ranks those CPZs to pursue priority-based countermeasures. Section 5.4 prepares a risk map along Highway 97 and Section 5.5 identifies collision prone road intersections based on their collision frequency. Section 5.6 estimates the economic cost of road collision and finally Section 5.7 summarizes all of the results.

5.2 Road safety context

Between 1996 and 2006, the population growth rate in BC was 1.39% annually. In that period number of licensed vehicles increased 1.9% annually and the number of licensed driver increased 1.63% annually (ICBC 2000, 2006). Also traffic collisions increased at an annual rate of 0.57% between 1997 and 2006. Figure 5.1 compares the fatality rate in BC with other Canadian provinces for 2007. It can be seen that in 2007, BC had a fatality rate of 1.58 per 10,000 vehicles which was higher than the Canadian average of 1.38. This figure also compares the 2007 fatalities with the average fatalities during 1994 to 1998. It is clear that all of the provinces experienced less collisions in 2007 compared to 1994-1998. BC has also experienced a reduction of 0.42 fatalities per 10,000 vehicles.

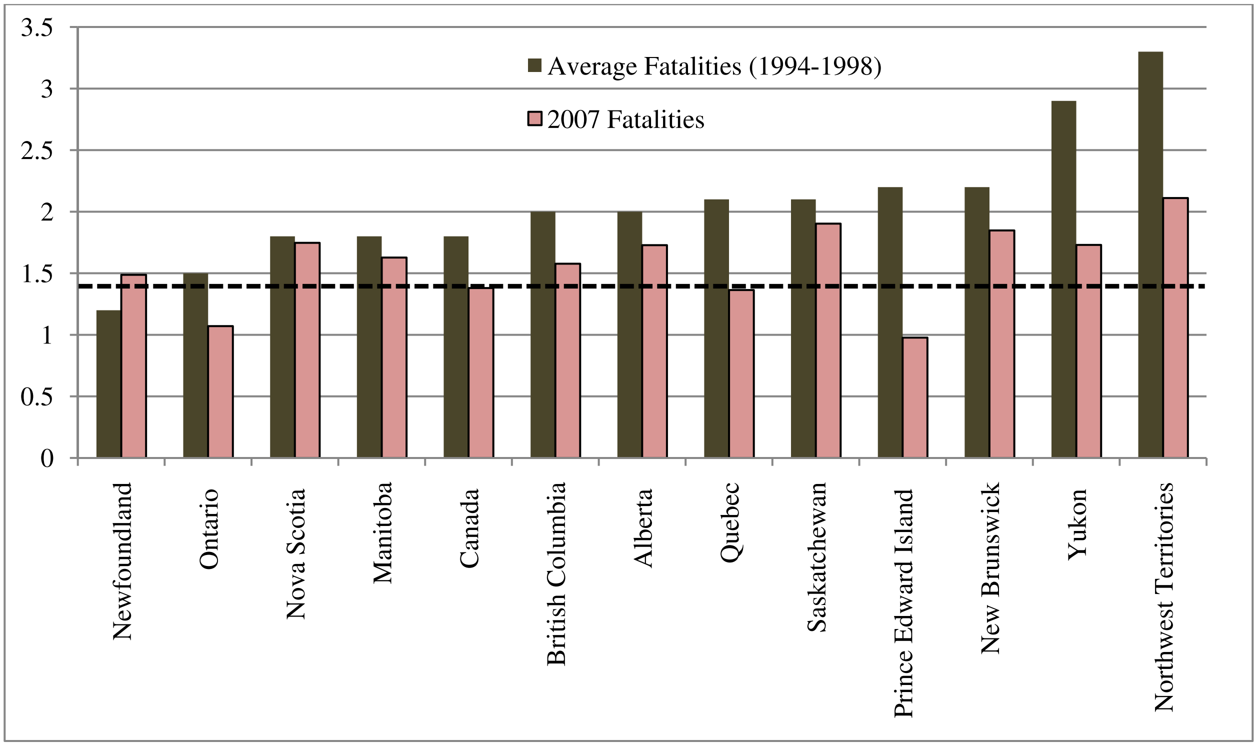


Figure 5.1 Fatality rate per 10,000 vehicles in different provinces of Canada

In BC, for last 20 years the total number of road collision victims has decreased significantly. In 1988, the total number of victims was 44,776 and in 2007 it was 26,414, indicating a 2.16% annual reduction. Figure 5.2 shows the recent fatality trend in BC. There was a sudden drop in the number of victims between 1995 and 1997. The reason was due to a major collision reporting policy change in 1996 when police in some municipalities began to attend only collisions of a more serious nature. So property damage only (PDO) and personal injury collisions have been under-reported since 1996. This affect of the policy change can be observed in the figure. By comparing the number of collisions before and after 1996, it can be seen that difference between the average numbers of collisions was around 48,000 collisions. That means almost 50% collisions were under-reported due to policy change, or in other words, after 1996 the total number of collisions (reported + unreported) would be double if no policy change was made. Considering this fact, it may be concluded that the decrease in road collisions was small even though it followed a decreasing trend. More details of the collision history in BC can be found in Appendix M.

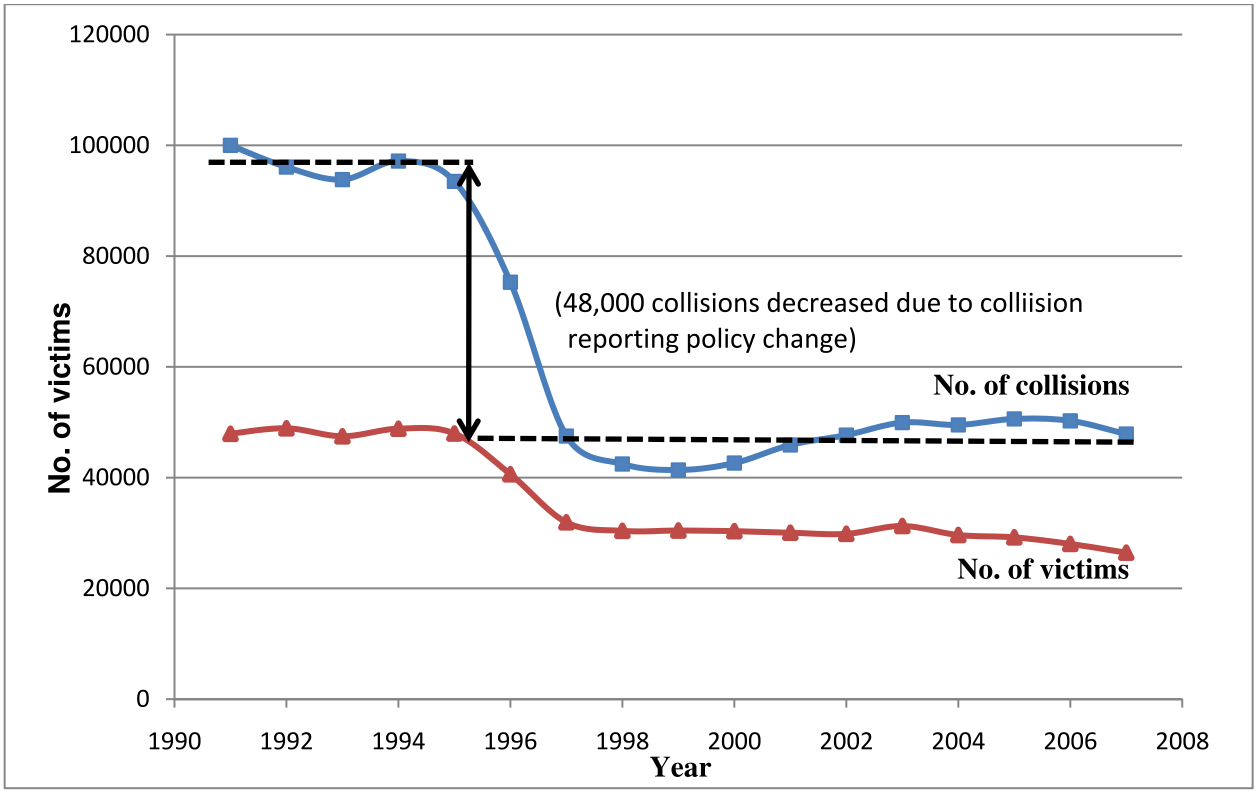


Figure 5.2 Number of victims due to road collisions in British Columbia

The Insurance Corporation of British Columbia (ICBC, 2007) reported that in BC a total of 47,870 road collisions occurred in 2007, resulting in 17,914 injuries and 372 fatalities. In other words, each day, on average, 1.14 fatalities and 68.7 injuries occurred. To improve Canadian road safety, Canadian Council of Motor Transport Administrators (CCMTA) initiated "Vision 2010" which set a collisions reduction target for BC. The fatality trend in BC and the target line of Vision 2010 (CCMTA 2007, ICBC 2007) are given in Figure 5.3. The figure indicates that the fatality trend line in BC was always higher than the target Vision 2010 line. In 2007 there was a difference of 100 fatalities between these two lines. This clearly indicates that the target has not been met and road safety improvements are needed to improve the existing situation.

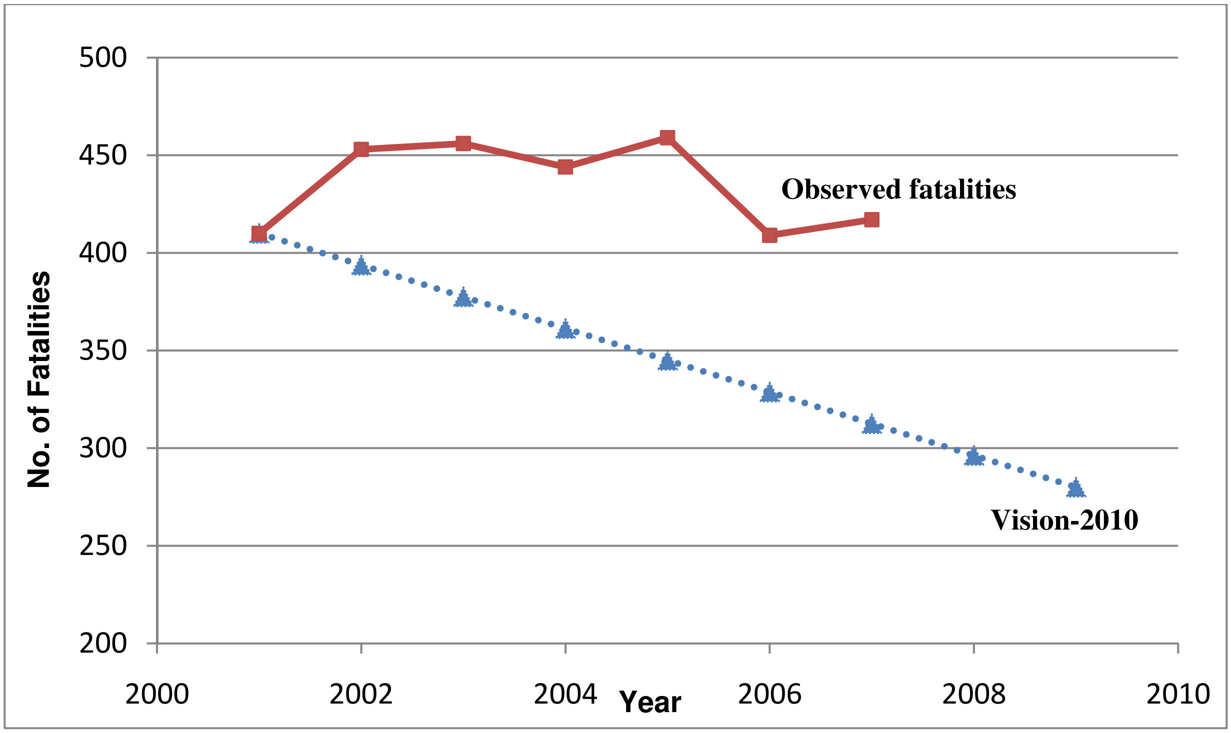


Figure 5.3 Fatality statistics in British Columbia

In the RDCO, between 1996 and 2006, the annual population growth rate was 1.82%. Rapid population growth produced more vehicular trips. As a result of more vehicles, more road collisions took place. In 1996, the total number of collisions was 2,067 and it almost doubled in 2006 to 3,882 collisions. Figure 5.4 shows the collision trend in the RDCO. It can be observed that the number of total and severe collisions increased significantly between 1996 and 2006, but the rate of increase was greater for total collisions than severe collisions. Figure 5.5 compares the number of collisions per 1,000 people between BC and RDCO. The increase of collision rate in the RDCO was much higher than that for BC. This clearly shows the road safety condition in the RDCO was not improving.

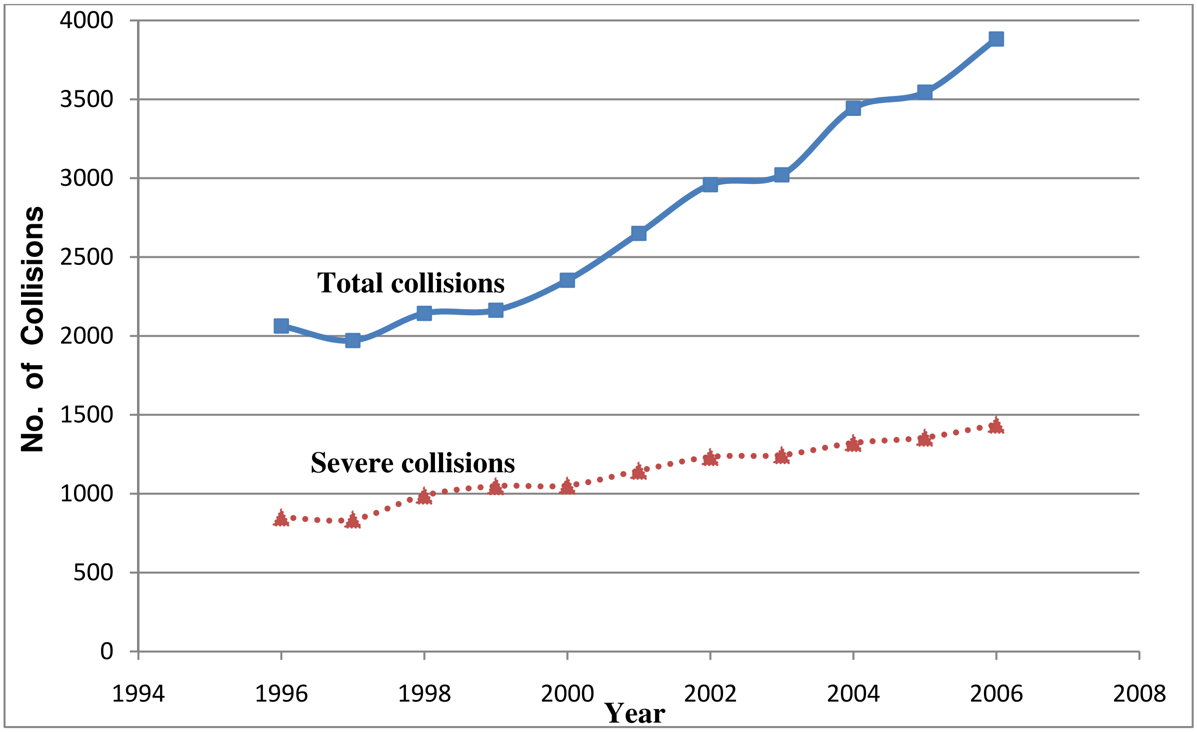


Figure 5.4 Total number of road collisions in RDCO

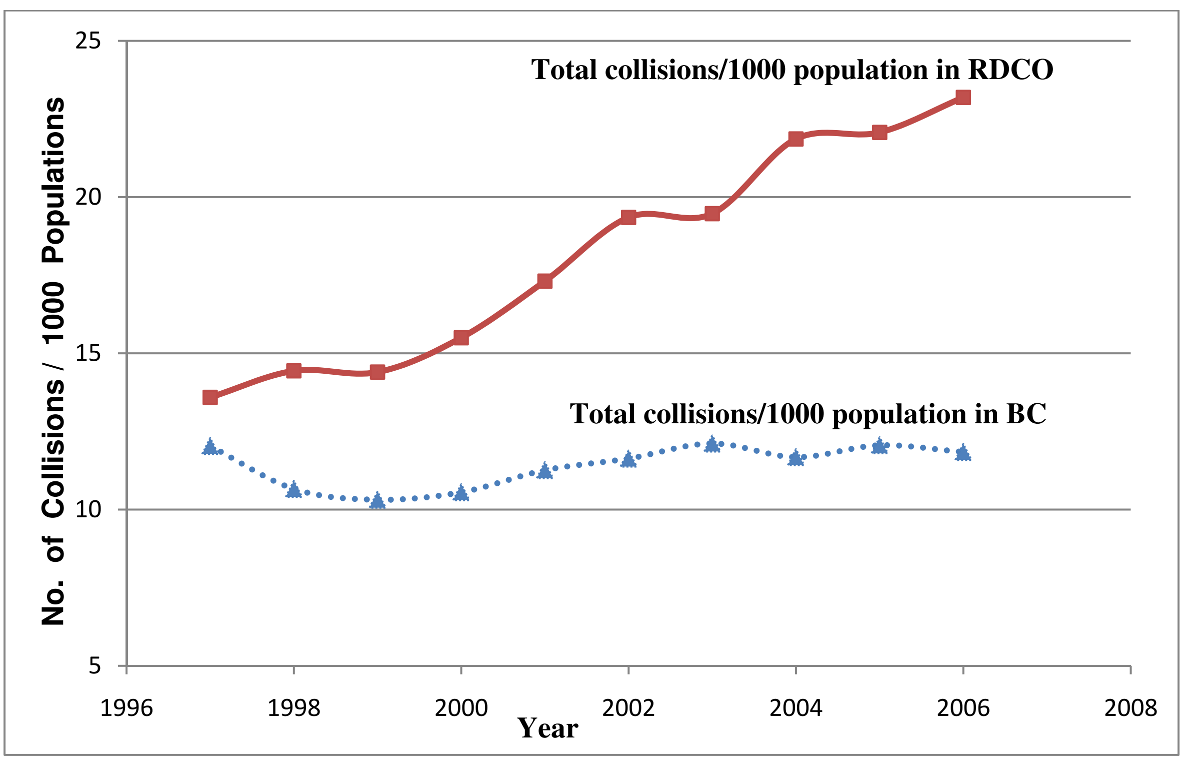


Figure 5.5 Comparison of road collisions trend between BC and the RDCO

5.3 Collision prone zones in the RDCO

It is clear that the road safety condition in the RDCO was not improving, rather it was experiencing increasing rate of collisions. So, it is important to know which locations experienced higher collision frequency than the average frequency. Collision density maps can be very useful to identify locations with higher frequency. The number of collisions inside each zone was calculated and was normalized with respective to the TAZ area. As collisions are rare events and collision frequency fluctuates around its long-term mean, three years of collision data were used for each case. Using 1996-1998 collision data, one map was produced and the other map was produced using 2004-2006 collisions data. After preparing the collisions density maps, it was found that most of the higher collision density areas were located around the central (i.e. busiest part) of Kelowna, which is shown in Figure 5.6 and 5.7. Both maps were scaled in the same way so that they could be compared. By comparing the maps, it appears that each zone experienced more collisions in 2004-2006 compared to 1996-1998.

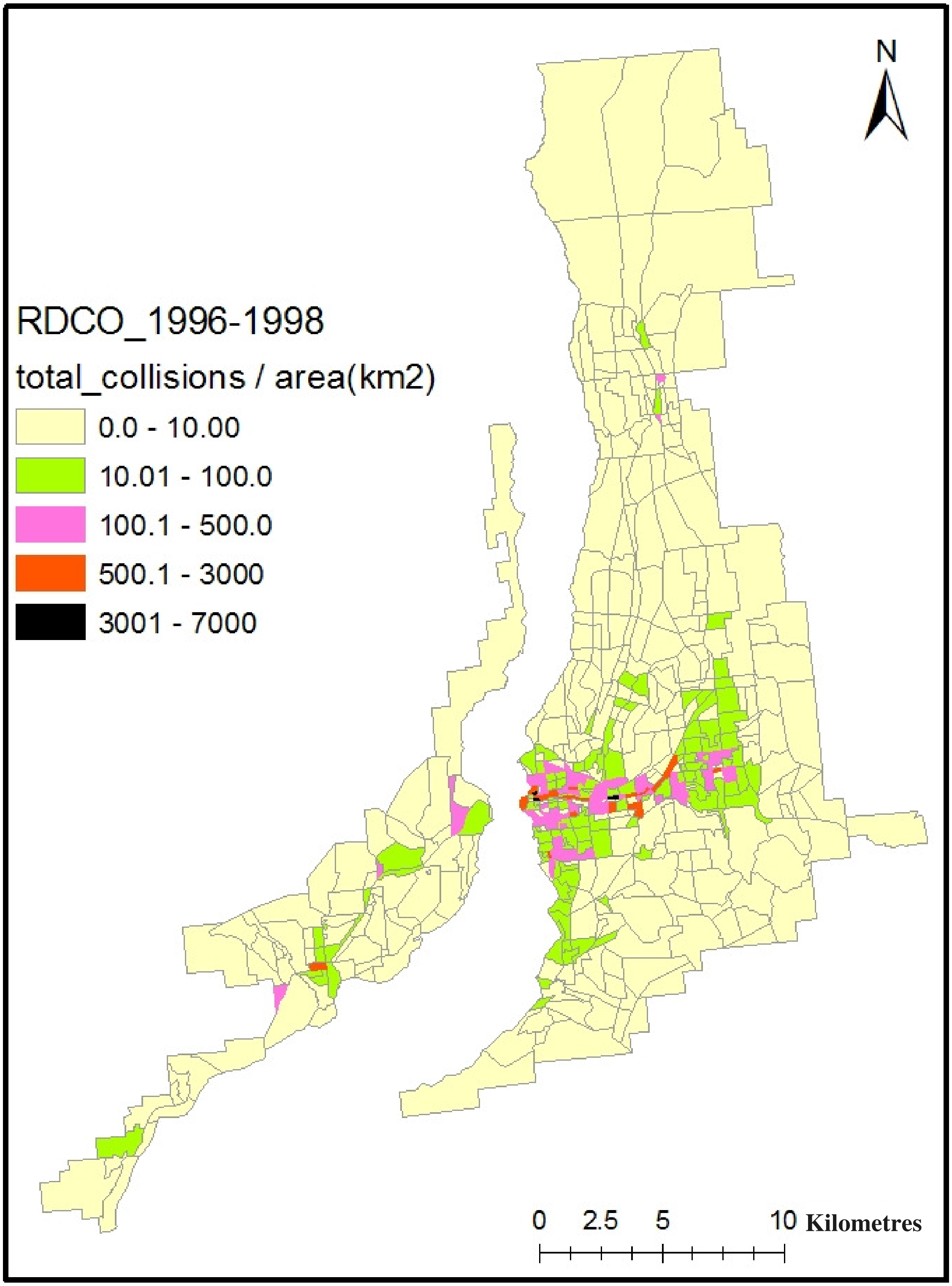


Figure 5.6 Collision density map of the RDCO for year 1996-98

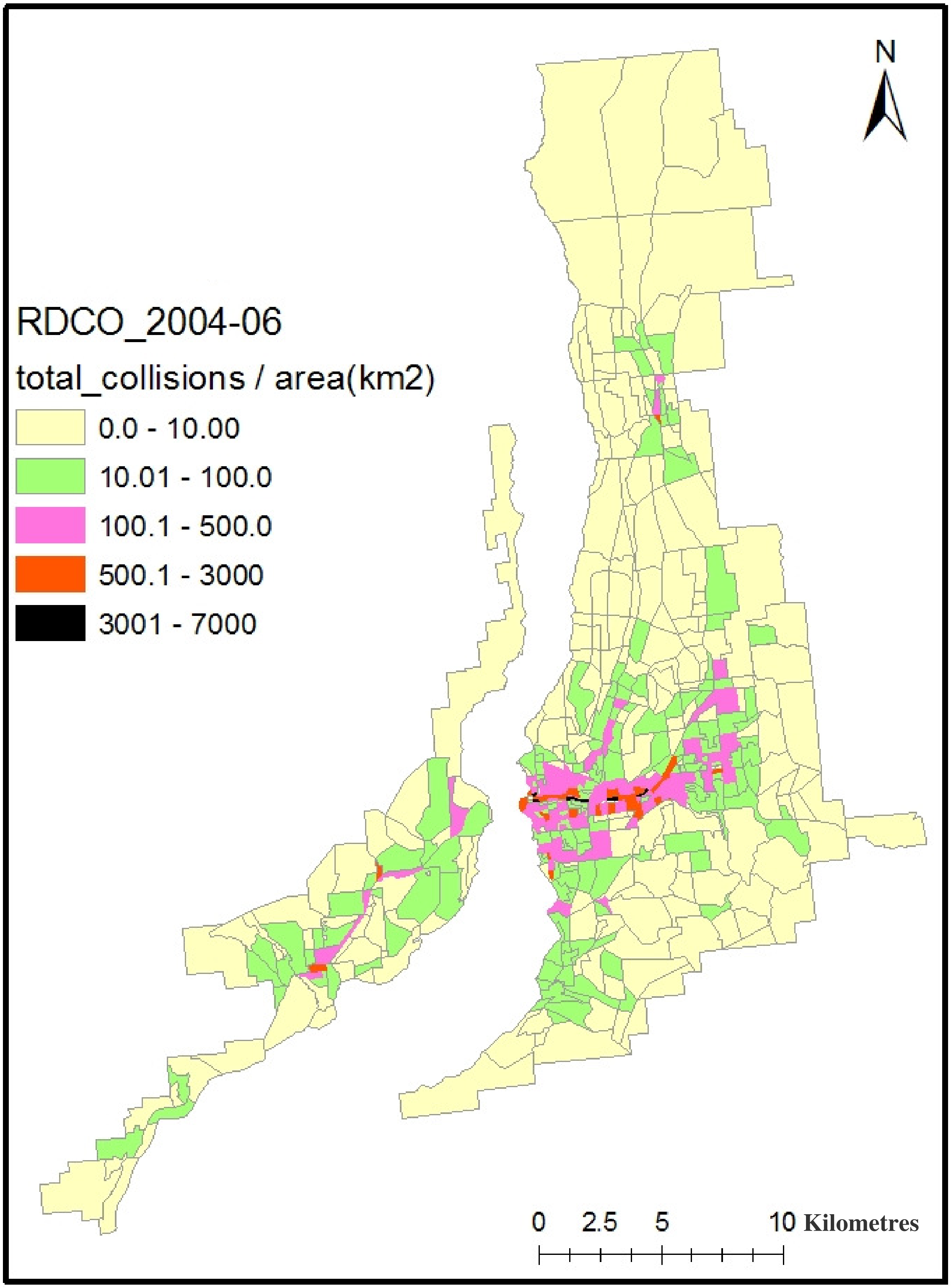


Figure 5.7 Collision density map of the RDCO for year 2004-06

5.3.1 Identification of collision prone zones in 2006

A black spot is a location which exhibits significantly higher collision potential when compared to average collision potential from a group of similar locations. To ensure that resources are spent only on locations with the highest potential for safety improvements, it is important to follow a sound procedure to properly identify and rank black spots for diagnosis and treatment. Identification of collision prone zones is a very important part of road safety analysis because it will identify zones that are experiencing higher road collisions frequency.

Identification of collision prone zones was done by using 2004-2006 collision data. It was done based on the method described in Section 3.6. The number of predicted collisions, E (Λ), was obtained using such CPMs that considers VKT and INTD. Using the observed and predicted collision data, Empirical Bayes (EB) estimate of each zone was calculated and using equation 2.29, a zone was identified as CPZ or non-CPZ. Equation 2.29 is repeated here for convenience.

$$1-\int_0^{E(\Lambda)} f_{EB}(\lambda)d\lambda = \left[1-\int_0^{E(\Lambda)} \frac{\left[\frac{\kappa}{E(\Lambda)}+1\right]^{(\kappa+count)} \lambda^{(\kappa+count-1)} e^{-\left[\frac{\kappa}{E(\Lambda)}+1\right]\lambda}}{\Gamma(\kappa+count)} d\lambda\right] \geq \delta\ (=0.95) \qquad (5.1)$$

As urban and rural areas have different characteristics in terms of predicting collisions, CPZs were identified separately for urban and rural areas. Twenty five urban CPZs were identified out of 242 urban TAZs and nine rural CPZs were identified out of 258 TAZs. Figure 5.8 shows the location of CPZs. It can be seen that most of the urban CPZs were around downtown Kelowna. Higher number of CPZs in downtown Kelowna is intuitive because this is the centre of activities that generated high volume of traffic. Also, most of the CPZs were found to be located along Highway 97. The reason could be the higher traffic volume on Highway 97.

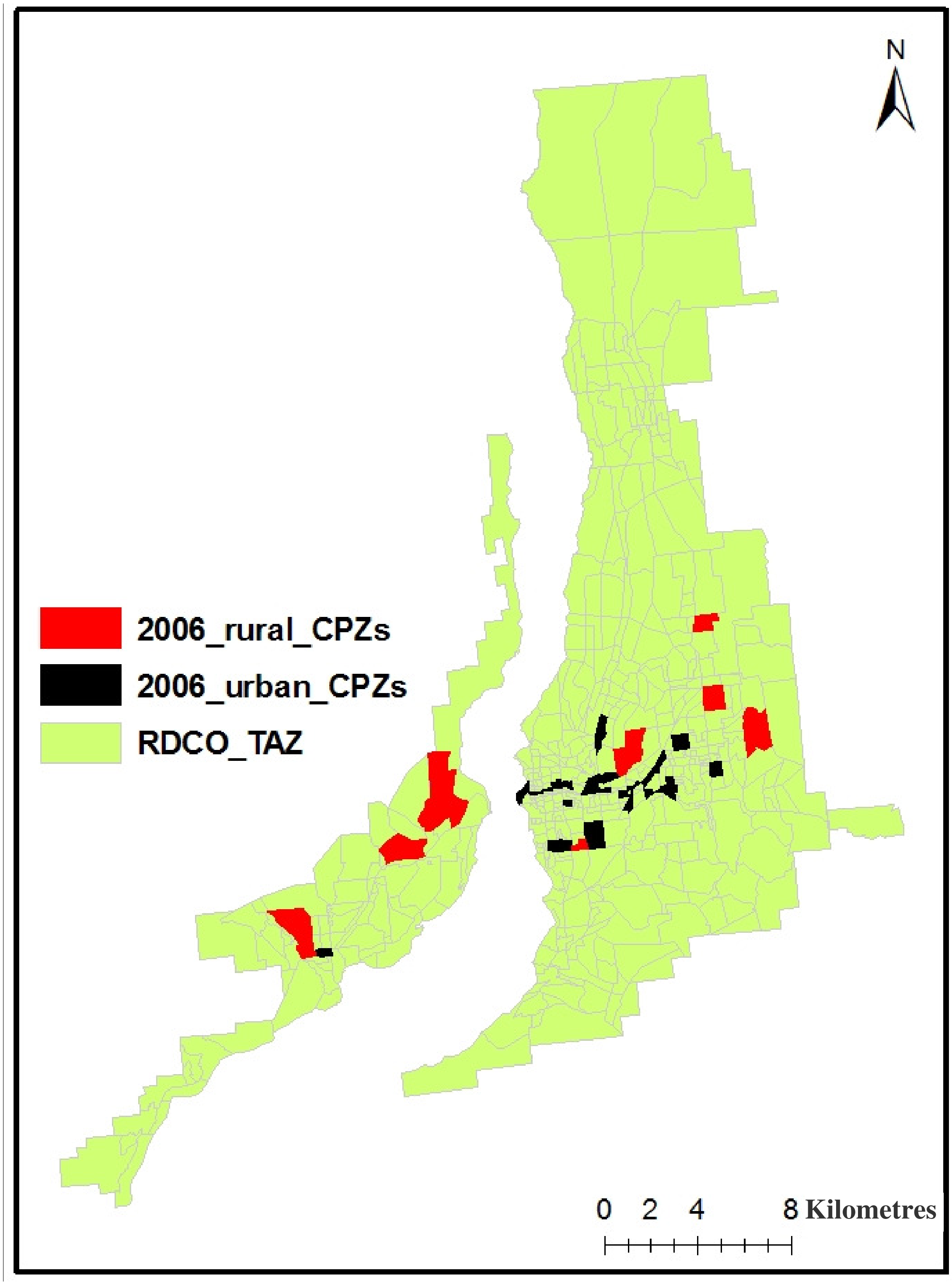


Figure 5.8 Collision prone zones in the RDCO in 2006

5.3.2 Ranking of 2006 CPZs

After identifying the CPZs, the next important step was to rank CPZs to take countermeasures on a priority basis. Based on the literature review, ranking was done for both of potential collision reduction (PCR) and collision risk ratio (CRR) method. At first, each CPZ was ranked for both methods and then the rankings of each CPZ were summed. A combined ranking was obtained where zone having lowest ranking number was ranked as the most collision prone zones.

Table 5.1 Ranking of 2006 urban CPZs

Urban Area					
Rank	Zone ID	observed coll./3yrs	modeled coll./3yrs	EB estimate / 3yrs	observed coll. density (3yrs coll./area_KM^2)
1	171	31	7.9	28.7	173.18
2	99	18	4.9	16.0	40.09
3	168	15	4.00	13.0	180.72
4	247	16	4.8	14.3	18.48
4	179	24	9.3	22.7	357.41
6	55	18	7.0	16.8	39.22
7	347	8	1.1	4.9	102.56
8	173	9	2.4	7.2	30.2
9	231	7	1.3	4.8	152.17
10	155	30	14.8	29.2	175.54

Table 5.2 Ranking of 2006 rural CPZs

Rural Area					
Rank	Zone ID	observed coll./3yrs	modeled coll./3yrs	EB estimate / 3yrs	observed coll. density (3yrs. coll./area_KM2)
1	395	61	6.1	53.8	66.09
1	34	38	2.6	28.8	43.33
3	19	12	2.2	9.1	21.78
4	249	13	3.3	10.9	56.52
5	385	10	2.2	7.7	6.31
6	3	6	0.8	3.2	3.3
7	397	6	1.6	4.4	2.04
8	366	5	1.3	3.5	2.44
8	201	5	1.4	3.5	3.11

Table 5.1 and 5.2 ranks the top ten CPZs for both urban and rural areas. This table also gives the observed collision density for each CPZ. In urban areas, collision densities were higher compared to rural areas. It can be observed that high collision density TAZs are not always ranked higher, rather opportunity to reduce collisions has more importance in the ranking. The potential for significant collisions reduction is important, because significant number of collisions can be reduced by applying countermeasures in high ranked CPZs.

5.3.3 Collision prone zones in 2020

This section made an attempt to identify potential collision prone zones in 2020. The total number of 2020 collisions was predicted for do-nothing scenario and calculated using the CPMs having VKT and INTD. As for year 2020 it is not possible to get observed data, the RDCO collisions trend was considered to extrapolate the number of possible collisions in 2020. It was found that if the current trend continues, then by 2020 the total number of collisions would be 1.6 times of 2006 collisions. So it was assumed that in 2020 each TAZ would experience 1.6 times of 2006 collisions. Assuming this, the number of 2020 collisions for each TAZ was calculated and considered as 'observed' collisions to calculate TAZ specific EB estimate. Later based on the methodology, zones were identified as CPZ or not. It was found that there would be thirty two urban CPZs and fifteen rural CPZs. Figure 5.9 shows the urban and rural CPZs in 2020. It can be observed that the number of CPZs would increase between 2006 and 2020, and hence, more road safety improvements should be taken to reduce the number of possible 2020 CPZs.

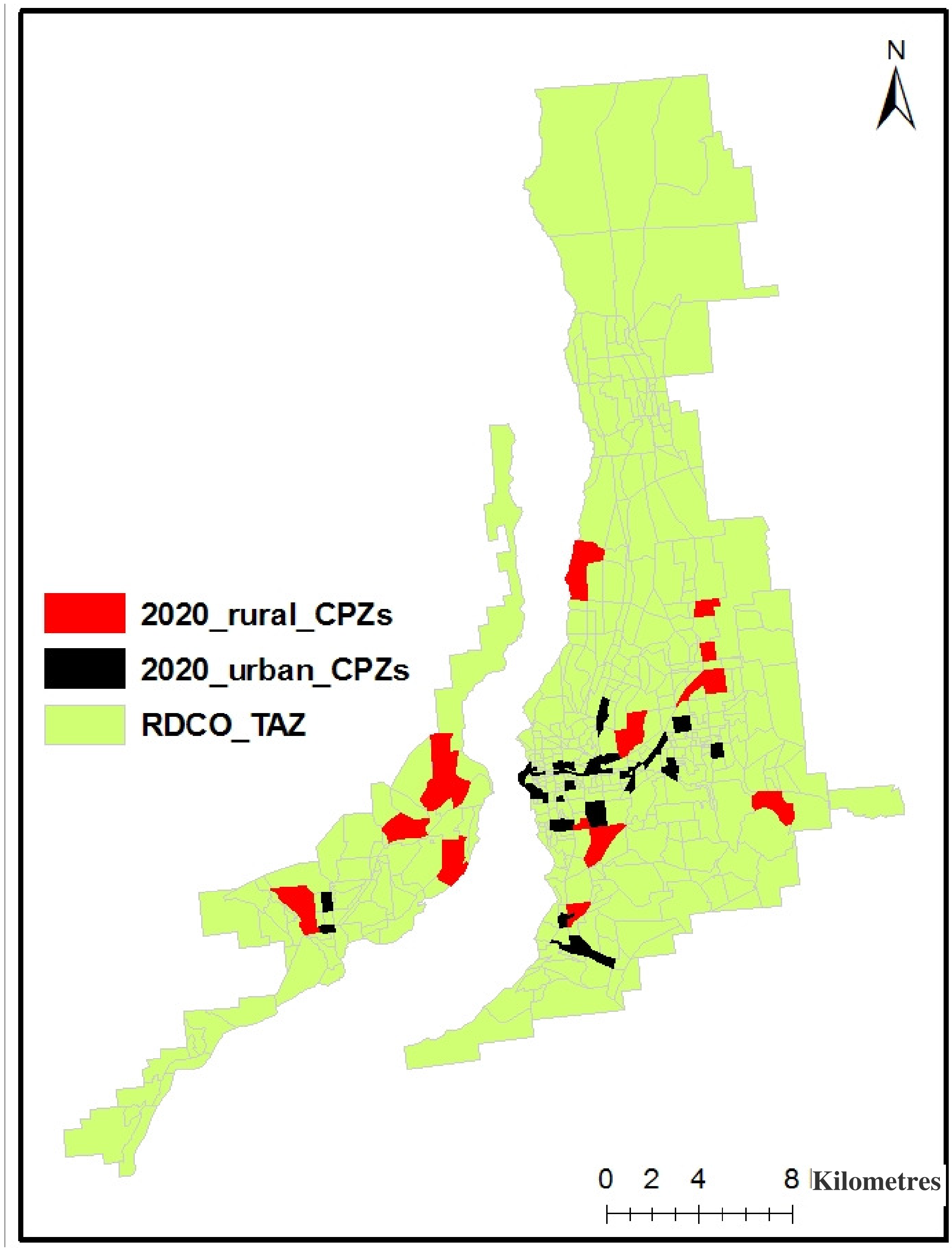


Figure 5.9 Collision prone zones in the RDCO in 2020

5.4 Collision density map along Highway 97

From the above result, it is evident that most of the CPZs were located along Highway 97. The Highway 97 corridor is the busiest transportation corridor in the region. This corridor is used by both local and through traffic. Between 1997 and 2001, the total number of collisions on

Highway 97 was 3,088, while between 2002 and 2006 it increased to 4,780 collisions; indicating 1.5 times increase of collision frequency. So it is important to understand the safety along Highway 97 and a risk map can be very useful to compare collision risks at different segments along the highway.

Within the RDCO boundary, the total length of Highway 97 is about 66 km. Based on the methodology, the entire length of the corridor was divided and shorter (<1 km) segments were merged together so that the minimum length of each segment became more than 1 km. Also using three years of collision data (i.e. 2004-2006), the collision density for each segment was calculated. After preparing the collision density map, it was found that most of the collision prone segments (CPSs) were in central Kelowna. Figure 5.10 shows the collision density map focusing on the City of Kelowna. Segments with high collision density are marked in dark color, while segments with lower collision density are marked in light color. It was found that most collision prone segments were located near Orchard Park mall in Kelowna. This is intuitive because Orchard Park mall is a centre of activities that produces higher traffic and hence, it has higher probability of collisions. From the collision data, it was found that 56% of Highway 97 collisions occurred between Abbott St. and McCurdy Rd. This corridor consists of 32 intersections within a length of 7.5 kms (i.e. 4.27 intersections/km). The presence of a high number of intersections could be the reason for high collision occurrence. So an intersection analysis was performed, which is described in the next section.

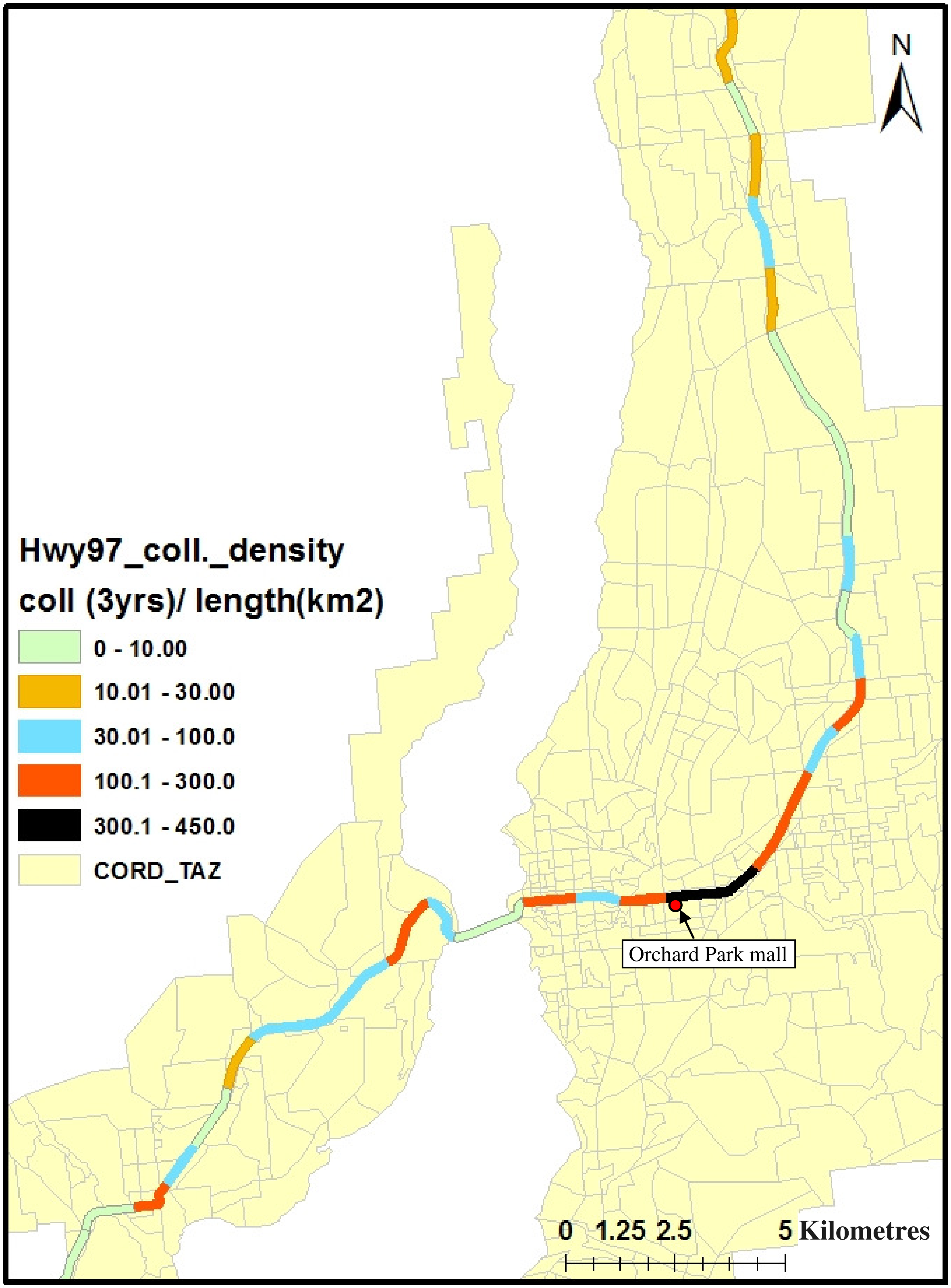


Figure 5.10 Collision density map of Highway 97 (near Kelowna downtown)

5.5 Intersection analysis

This section identified collision prone intersections based on the collision frequency of each intersection. Collision data for three years (i.e. 2004-2006) were used for the analysis. A 30 m buffer area was created for each intersection and the number of collisions inside of each circular area was calculated. Worst fifteen collision prone intersections were selected and the average annual collisions density (i.e. collisions/ AADT) for each intersection was calculated. Based on the collision density, intersections were ranked as shown in Table 5.3. Not surprisingly it was found that 12 intersections were at Highway 97. This table also shows the total number of traffic volume entering each intersection. The entering traffic volume at these intersections is very high compared to other intersections in the region. High traffic volume at these intersections was not only responsible for higher collisions, but also for lowering the level of service (LOS). From a study of the City of Kelowna (2007) it was found that in 2007 LOS of these intersections was between D and F (where LOS-A gives the least and LOS-F gives the maximum average delay per vehicles) and these intersections would experience LOS-F by 2020 if no improvements were taken.

Table 5.3 Top fifteen collision prone intersections in the RDCO

Rank	Location	No. of collisions/3yrs	Traffic volume (entering)	avg. annual coll. rate (coll. /AADT)
1	Highway 97 & Westside Dr	178	52,496 (2006)	1.13
2	Dilworth Dr & Springfield Rd	224	114,207 (2006)	0.65
3	Highway 97 & Abbott St.	216	110,121 (2005)	0.65
4	Highway 97 & Banks Rd.	183	100,827 (2006)	0.60
5	Highway 97 & Dilworth Dr	256	148,678 (2006)	0.57
6	Highway 97 & Cooper Rd.	222	143,374 (2006)	0.52
7	Highway 97 & Gordon Dr	218	144,000 (2004)	0.50

Rank	Location	No. of collisions/3yrs	Traffic volume (entering)	avg. annual coll. rate (coll. /AADT)
8	Highway 97 & Spall Rd.	236	166,251 (2006)	0.47
9	Bernard Ave. & Spall Rd.	119	84,630 (2005)	0.47
10	Highway 97 & Leckie Rd	145	112,815 (2006)	0.43
11	Highway 97 & Sexsmith Rd.	175	150,406 (2006)	0.39
12	Highway 97 & McCurdy Rd	116	100,071 (2006)	0.39
13	Highway 97 & Leathead Rd	136	119,506 (2006)	0.38
14	Benvoulin Rd. & KLO Rd.	109	97,132 (2006)	0.37
15	Highway 97 & Highway 33	168	354,440 (2006)	0.16

5.6 Collision pattern

One of the important steps of road safety improvement programs (RSIPs) is to understand collision patterns (i.e. collision types, time of occurrence, severity etc.). By understanding the actual collision pattern, selected countermeasures will be more effective in reducing collisions. The 3-year collision pattern (2004-2006) was analyzed and Figure 5.11 was produced. It can be seen that 80% of the total collisions occur on weekdays and most collision occur on Friday (18%). Also this figure indicates that 30% of the collisions occur between 3pm - 6pm followed by 23% of the collisions between 12pm - 3pm. More than 50% of the collisions occur in the afternoon period (12pm - 6pm). Also 10% of the collisions occur in the morning rush hours (6am - 9am) compared to 30% collision in afternoon rush hours (3pm - 6pm). Later, collisions were examined based on the type of area (i.e. urban or rural). It was found that 14% of the rural collisions occurred from 6am to 9am and 44% occurred from 12pm to 6pm. In urban areas, 9% of the collisions occurred from 6am to 9am and 56% occurred from 12pm to 6pm. The reason for fewer morning collisions in urban zones could be presence of less social and shopping trips at this time. On the other hand, collisions between 12pm – 6pm are relatively higher in urban areas due to higher social and shopping trips. Also due to higher intersection density in urban areas, it was found that 48% of urban collisions were 'Rear End'; whereas 31% of rural collisions were 'Rear End'. As traffic volume is positively related with collision occurrence, traffic volume composition at the identified collision prone intersections was also analyzed and it was found that 15% of daily traffic was between 6am – 9am and 10% of collisions were occurred at that period. On the other hand, 45% of daily traffic was between 12pm – 6pm and 57% of collisions occurred in that period. It means that between 12pm – 6pm the occurrence of collisions is relatively higher than other periods of the day.

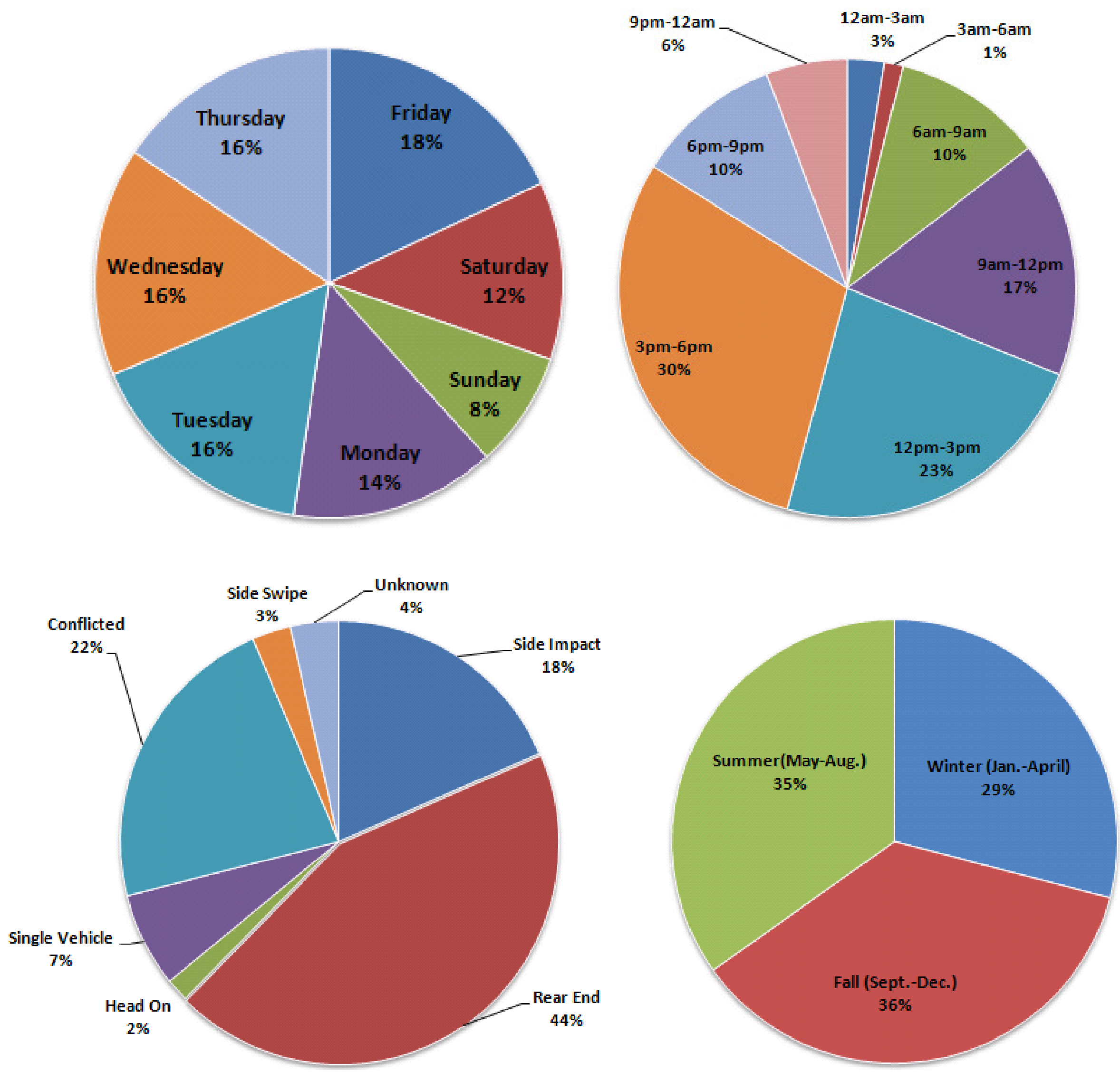


Figure 5.11 Collision patterns for year 2004-2006

Also by looking at the month, it was found that the collision occurrence is almost same in each month. Three different seasons were considered: (1) fall (i.e. September-December); winter (i.e. January-April); and, (3) summer (April-August). It was found that 36% of total collisions occurred in fall, whereas 29% in winter. In summer, it was 35%. As the RDCO is the centre of tourism activities in summer, tourists coming to the RDCO significantly increases vehicular trips. The number of students driving generally falls in the summer and increases in the fall. In winter, people make less trips and hence, there is less probability of road collisions.

After analyzing the collision types, it was found that the predominant type of collision was 'rear end' collisions, which was 44% of total collisions. 'side impact' collisions were 18% and 'conflicted' collisions were 22%. Here it is necessary to mention that by definition 'conflicted' collisions include those collisions which didn't have clear description to identify actual collision type. Usually 'rear end' type collisions occur mostly at intersections because of the presence of traffic signals and stop signs. For the RDCO it was observed that most of collision occurred at the intersections and also it can be realized from the CPMs where intersection density (INTD) was a significant collision variable. More attention should be given to reduce intersection related collisions. Signal co-ordination, changes in signal timing, addition of new turning lanes, construction of roundabouts etc. could be effective countermeasures to improve intersection safety. Also, as Kelowna downtown has high intersection density, construction of a bypass to avoid Highway 97 traffic could be a possible solution. More research is needed to find appropriate countermeasures.

5.7 Economic cost of collisions

The direct and indirect economic costs of road collisions are very high. Globally, the economic cost of road injuries varies from 1% to 3% of Gross National Product (GNP), and in Canada the total collision cost in 2004 was $63 billion which was 4.88% of Canada's 2004 GDP (UN, 2003; Vodden et al., 2007). Due to high collision cost of collisions, it is very important to find associated collision costs so that the economic benefits of any road safety improvement measure can be evaluated.

Collision cost has several components (i.e. victim related cost, property damage, administrative cost). An example of different types of collision components and associated costs are given in Appendix N. Based on the geography and socio-economic condition, collision cost varies from place to place. For BC, Table 5.4 lists three different methods of cost estimate (Lovegrove, 2007). The Human Capital Cost (HCC) method is based on the estimated value of an individual person's contribution to society in terms of direct impact on GNP (Miller, 1993). On the other hand, the Willingness-to-Pay (WTP) method is based on the willingness of the road users to pay for a safe journey (Miller, 1998). WTP is a measure of how much a society or an individual is willing to pay to avoid death/injury resulting from a road collision (Jones-Lee et.al., 1985; Mishan and Schelling, 1971). These two methods are based on economic assessments that involve a tradeoff between risk and economic resources. In contrast, the Insurance Claims (IC) method does not consider social costs, only the average actual auto insurance payment is considered for the estimation (ICBC, 2004).

Table 5.4 Typical road collision costs ($CDN 2004) in BC for a single collision (Lovegrove, 2007)

Collision Severity	Human Capital Cost	Willingness to Pay	Insurance Claims
Fatal Collision	CDN $ 1,000,000	CDN $ 4,200,000	CDN $ 281,000
Injury Collision	CDN $ 35,000	CDN $ 100,000	CDN $ 44,000
Property Damage Only	CDN $ 5,000	CDN $ 6,000	CDN $ 4,500

This table clearly indicates that the economic cost of collision varies depending on the type of estimating method. The WTP method estimates the highest and the IC method estimates the lowest cost. Based on 2007 collision data (ICBC, 2007) and the method used, it was estimated that in 2007 the collision cost in BC varied from $1 billion to $3.6 billion (in $CDN 2004), or in other words $230 to $850 annually per BC resident. Figure 5.12 shows collision cost (all values are in 2004 Canadian dollars) of the RDCO for 1996-2006. It clearly shows that the collision cost increased between 1996 and 2006. According to Insurance Claims (IC) method, between 1996 to 2001 collision costs increased by 41% and between 2002 to 2006 it increased by 22%. But according to WTP method it increased by 15% during 1996-2001 and 33% during 2002-2006. A sample calculation can be found in Appendix O.

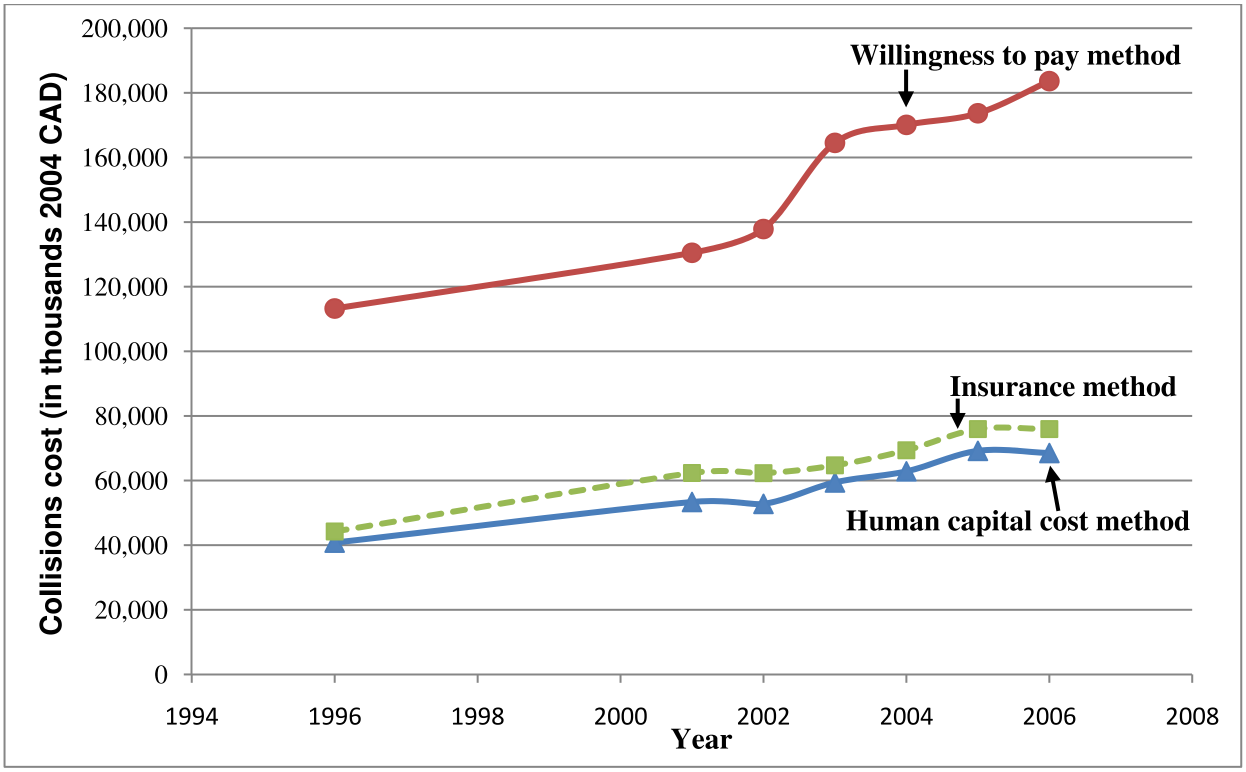


Figure 5.12 Economic cost of collisions (in thousands 2004 CAD) in the RDCO

Nevertheless, whatever methods are used, it is evident that the collision costs are increasing, and these costs are very significant when compared to average income of people. In 2006, the total RDCO population was 162,276 and the average individual income was $32,030. So, according to WTP method the cost of collision in 2006 was $1,097 per person which was 3.42% of the annual income of a person living in the RDCO. In the words, in 2006 the collision cost was 2.55% of 2006 GDP of BC. To give a better idea about these collision costs, they can be compared with transit fares. In 2006, the monthly transit pass was $47 and the average collision cost per person was $1,097. It implies that monthly collision cost of $91 is twice the cost of monthly transit pass. Road safety collisions have become not only a social burden but also an economic burden to society.

5.8 Summary

The primary objective of this chapter was to identify collision prone locations and to find collision patterns. It was found that the number of collisions in the RDCO had an increasing trend while in BC it had a decreasing trend. In 1996, the total number of collision in RDCO was 2,067 and it almost doubled in 2006 to 3,882 collisions.

To pursue road safety improvement, it is important to identify and rank collision prone zones so that priority-based countermeasures can be taken. To identify CPZs, three years of data (2004-2006) were used. A total of twenty five urban CPZs and nine rural CPZs were identified. Most of the TAZs were in central Kelowna and most of them were along Highway 97. After preparing a collision density map along Highway 97 it was found that the most collision prone highway segments were in the centre (i.e. busiest part) of Kelowna, especially near Orchard Park mall. An intersection analysis found that most of the high ranked collision prone intersections were situated at Highway 97. Also after identifying 2020 CPZs it was found that new zones would become CPZs and hence, more road safety improvements should be taken to preclude them from becoming collision prone.

After analyzing the collision patterns, it was found that 30% of collisions occurred between 3 pm and 6 pm followed by 23% collisions between 12 pm and 3 pm. More than 50% of collisions occurred between 12 pm and 6 pm. It was also found that ’rear end’ collisions was the predominant type of collisions, which was 44% of total collisions. As ‘rear end’ collisions are very common at intersections, more research and investigation are needed to improve intersection safety.

Finally, the economic cost of collisions in the RDCO was calculated by three different methods. The Willingness to Pay (WTP) method estimated collision cost of $183,628,000 in 2006, while Insurance Claims method estimated $75,929,000. Also based on the WTP method, the average collision cost in 2006 was calculated as 3.42% of the annual income of a person living in the RDCO. The results indicate the need for more attention to collision prone locations so that by reducing the number of road collisions a safer, economic and efficient road network can be obtained.

CHAPTER 6 CONCLUSIONS, CONTRIBUTIONS & FUTURE RESEARCH

6.1 Summary & conclusions

Road traffic collisions have been recognized as a global major public health and safety problem. It has become one of the top ten leading 'causes of death'. The enormous social and economic burden of road collisions was one of the motivations of this research. Another motivation came from the sustainability concern. The present automobile or vehicle-based transportation system is responsible for more congestion, environmental pollution, and road collisions. Energy crises have become a threatening issue to the earth and in future, increasing trend of fossil-fuel will make transportation more expensive. All of these issues emphasize the need of shifting towards an auto-alternative sustainable transportation system which will reduce environment pollutions, improve road safety, and enhance regional economic growth in such a way that it will have less harmful effect on the future of the earth.

This thesis intended to quantify the road safety benefits of sustainable transportation in the form of increased transit service. The study area of the research was the Regional District of Central Okanagan (RDCO) which is situated along the shores of Okanagan Lake in the southern interior of British Columbia, Canada. The first objective of the research was to develop an auto-based transportation planning model of the RDCO for year 2006 and year 2020. The second objective was to conduct road safety analysis due to future transit and road network improvements, and the third objective was to identify collision prone locations (i.e. zones, highway segments and intersections) of the region.

This research built a 4-step transportation planning model of the RDCO for year 2006. The model was built for the morning rush hours (6am-9am). Using the Institute of Transportation Engineers (ITE) trip generation rates total number of generated trips for each Traffic Analysis Zone (TAZ) was calculated. In the next step, based on the Gravity model, the trips were distributed to different zones and an origin-destination (O-D) table was prepared. This thesis also developed a new mode share model for the RDCO, which can calculate the percentage of auto (i.e. private vehicle) user in each TAZ. Using the auto-mode share equation, auto-trips were assigned to the network to get traffic volume on links. Finally using the model output (i.e. VKT, VC, SPED) and measured variables (i.e. TLKM, INTD) different types of community based macro-level collision prediction models were developed for both urban and rural areas. It was found that the collision number was significantly related to vehicle kilometres traveled (VKT), volume capacity ratio (VC), intersection density (INTD) and total lane kilometre TLKM). Based on available data, these well-fitted Collision Prediction Models (CPMs) have the capability to predict the number of collisions at community-level.

The 2020 transportation planning model also allows for four different sub-scenarios: (1) do-nothing; (2) only road improvements; (3) only transit improvements; and, (4) both transit and road improvements. For each sub-scenario, separate RDCO models were created and different outputs (i.e. VKT, TLKM, VC etc.) were extracted. It was found that transit improvements can reduce auto (i.e. private vehicles) mode share from 86.7% (do-nothing) to 79.6% (transit improvement). The number of RDCO road collisions was predicted for each scenario with the help of CPMs. The difference between sub-scenarios results were statistically significant and it was found that road improvements and transit improvements have the ability to reduce urban total collisions by 1-2% and 7-8% respectively. Transit improvement was also found to be effective in reducing rural road collisions by 4-8%. But it was observed that instead of reducing collisions, road improvements would increase rural collisions. Due to road improvements some commuting, through traffic could use rural roads instead of urban roads and hence, it would increase the number of rural collisions. It suggests that unless justified, improving (i.e. construction of new roads) road network in rural areas will not increase road safety. This result can be very useful to transportation planners before implementing any road improvements in rural areas.

Using CPMs and a standard Empirical Bayes (EB) statistical technique, this thesis identified and ranked collision prone zones CPZs for both urban and rural areas, many of which were adjacent to Highway 97. These would help to adopt priority-based countermeasures to reduce collision frequency in these CPZs. A collision density analysis along Highway 97 identified collision prone intersections in the RDCO. After analyzing the collisions, it was found that more than 50% collisions occurred between 12pm and 6pm and 44% of total collisions were ‘rear end’ type collisions, suggesting the need for intersection safety improvements. Finally, this research calculated the economic cost of road collisions in the RDCO. It was found that in 2006 average collision cost to a person was 3.42% of his/her annual income. This emphasizes the huge economic burden of road collision to society.

6.2 Research contributions

The following items represent the main contributions of this research:

A. Development of RDCO transportation planning model for 2006 and 2020

Transportation planning models can generate information useful to decision makers to access the consequences of transportation-related actions. One of the main contributions of the research was to develop 4-step transportation planning models for morning (i.e. 6am – 9am) traffic in the RDCO. One model was for year 2006 model and the other for 2020. Different types of land-use variables were used to calculate the total number of trips from each zone. An auto origin-destination (O-D) table was prepared to show the distribution of trips. A mode choice model was developed to predict the number of auto users from each TAZ and finally, after assigning trips traffic volume on each link was found. The 2006 base year model was calibrated with measured (i.e. counted) data and this base year model can be used for any future planning related works.

For the 2020 RDCO transportation planning model, future land use, transit improvements, and road network improvement plans were used. It was realized that other than transit, land use and road improvements have profound impact on road safety. Future land use was used to predict future traffic volume and future road network was used in different combinations with transit. Four different 2020 sub-scenario models were built, which were: (1) do-nothing (no transit and road improvements); (2) only road improvements but no transit improvements; (3) only transit improvements but no road improvements; and, (4) both transit and road improvements. Even though the first two steps (i.e. trip generation and trip distribution) were same for each sub-scenario, mode choice and trip assignment results were different. Each scenario output would help to realize different situations in 2020. These models would help to predict future traffic volume on the road network and these would help decision makers plan for the future. The impact of any land-use strategy can also be evaluated with the help of these planning models.

1. **Adjustment of trip generation rates for the RDCO**

For the development of 2006 transportation planning model, ITE trip generation rates were used to calculate total number of trips produced by and attracted to each zone. After that total number of trips for each municipality was calculated and compared with the observed 2007 household survey data. By comparing, trip generation rates were re-adjusted in such a way that errors in model prediction became less. After a few iterations, the trip generation rates were re-adjusted and new trip generation rates for the RDCO was obtained. These RDCO specific trip generation rates can be used for any future trip generation predictions.

2. **Development of the first mode choice model for the RDCO**

Creation of a mode choice model is the third step of traditional 4-step transportation model. It calculates how many trips will be made by each available transportation mode in the study area. Unfortunately, there exists no mode choice model for the region. So, this research developed a new mode choice model for auto, that includes different input parameters (i.e. bus headway, bus stop density, vehicle per household and percentage of population between age 45-64) to calculate the percentage of auto user in each zone. The input data were obtained from the available Census-2006 data and the predicted auto user percentage was very close to the observed data. This mode choice model can be very useful for future mode shift analysis because it can predict the changes in auto user due to any changes in mode choice variables. Beside auto, this research also calculated other mode choice variables that can be used for developing other mode choice equations.

B. Development of collision prediction models for the RDCO

Collision prediction models are very powerful proactive road safety tool that can be used to predict the number of collisions based on independent variables. This research developed several AM period CPMs for both urban and rural areas of the RDCO. These CPMs can explicitly evaluate the road safety of any proposed road improvements. Even though this research developed only exposure related CPMs, it also calculated the values of other CPM variables (i.e. TDM, S-D, network) which can be used to develop different forms of CPMs.

C. Evaluation of road safety for future road network and transit service improvement

This research built four different sub-scenarios for 2020 including transit and road network improvements. For each sub-scenario, the total number of expected collisions was predicted with the help of collision prediction models for both urban and rural areas. The road safety condition of each sub-scenario was compared with other sub-scenarios. This research suggests that transit improvement could reduce both urban and rural collisions. But instead of reducing collisions, road improvement would increase rural collisions. This result could be very useful to transportation planners before constructing any new roads in rural areas.

D. Black spot analysis

Black spot analysis is very important to properly identify and rank black spots for diagnosis and treatment. Using CPMs and EB technique, this research identified collision prone zones for both urban and rural areas. It also identified future zones that will exhibit higher collision potential than the average if no transit and road improvements will be done. This identification would help to focus on existing and potential CPZs to reduce collision frequency. This research also prepared a collision density map along Highway 97 and identified collision prone intersections in the RDCO. The collision pattern was also analyzed to understand and implement appropriate countermeasures. This research has found that collision frequencies are highest at road intersections and higher intersection density in the Highway 97 corridor made it one of the most collision prone locations in the region. Identification of collision prone locations opens the opportunity to do some future research to improve road safety at intersections.

E. Development of RDCO database for future research

A big contribution of the research was to build an updated and expanded RDCO database for socio-demographic and transportation variables. The development of a transportation planning model requires huge spatial datasets, which are not easy to handle and manage. Different types of data were obtained from different sources in different formats. This research converted them in useable format and aggregated them for each TAZ. Some of them were used in the research and some of them can be used for future research. The construction of the RDCO database will make future research more efficient and less time consuming.

6.3 Recommendations for future research

During the research, several limitations and the need for future improvements were identified. Therefore, this section presents some of the recommendations for the future research to enhance and strengthen the methodologies and the outcome of the transportation planning model as well collision prediction model:

- Travellers make trip decisions based on different factors such as trip purpose, travel time, travel cost etc. Among these factors, trip purpose is the most important factor to influence trip making decisions. Usually if the trip is work related, travellers do not hesitate to make relatively long distance trips. But in case of social trips, the probability of making a trip decreases with the increase of travel cost (i.e. time, distance). That is why different trip purposes have different friction factors and it is very important to know the purpose of the trips. But due to a lack of local trip purpose survey, this research considered only one general trip purpose. Identifying different trip purposes and conducting a trip survey to develop purpose based friction factor equation would be a very important topic for future research.

- This research developed a mode share equation for auto. But due to lack of data, it was developed based on an aggregated approach. It has been proven that disaggregate level mode choice models have several advantages over aggregate models and thus, disaggregate mode choice models have been widely used throughout the world. Disaggregate models require a detailed traveller survey to understand travellers' mode specific perceptions. A detailed traveller survey can be undertaken to develop disaggregate level mode choice model. To be conservative, this research didn't include the effect of auto operating cost and auto parking cost. But in future the price of these two variables will increase and it will discourage auto travel. Another important variable was 'journey time' which was also not used in the mode share equation. Inclusion of these variables can enhance the mode choice equation to predict better mode splits. A detailed household survey can be done to obtain disaggregate level data.

- For transportation planning model calibration, it is very important to have observed data which can be compared with modeled output. For the RDCO, there exist no origin-destination (O-D) trip tables to show the observed number of trips for each O-D pair.

Development of an O-D table can be very useful to calibrate the model. This can be done by counting traffic, household and roadside survey.

- A good quality digital road network is the corner stone of building a transportation planning model. It should contain sufficient road network information such as free flow speed, number of lanes, link length, link and node capacity, type of nodes etc. For this research the digital road network didn't include the capacity of links, rather it included link types. So based on the link type, capacity of each link was defined manually. But similar types of links had different capacity in different locations based on the geometry and surrounding environments. In some cases, it was found that the digital road network map had more lanes than the existing network. So, a careful observation was made to correct the capacity, speed limit, and numbers of lanes. Further investigation is required to increase the model efficiency. Also there is a need of digital signalized intersection map which could be used to define the types (i.e. signalized, non-signalized) and capacity of intersections. This can be done by collecting better data from the respective authority and also by manual inspection.
- In the RDCO transportation planning model, four external zones were considered. But external traffic volume was only used for the Vernon and Penticton external zones. For East and West zones, traffic volume was set to zero. More research is needed to estimate future traffic volume for these zones. This research also assumed that by using transit, 100 trips will come from RDNO to RDCO and no trip will come from Penticton to the RDCO. The validity of the assumption also needs to be verified. This can be done by collecting the future transportation plan from the associated authority, if available.
- This research has found that increasing transit services in rural areas would result in a reduction of rural collisions. From road safety perspective, transit should be improved (i.e. increased bus frequency, new routes) in rural areas. Again improvements of transit services will result in reduced green house gas emission, noise pollution, and traffic congestion. Meanwhile, in rural areas transit ridership is relatively low and transit improvement will cost a significant amount of money. A detailed social cost-benefit analysis (SCBA) could be a future topic for improving transit in rural areas.
- This research has built auto-based transportation planning models. It is recommended that a transit model be constructed to predict transit volume on links, transit travel time, transit

ridership, and transit demand. Collecting transit ridership data for each bus stop would help to predict transit ridership profile along transit routes. It would also help to improve the existing transit system. The use of a railway can be a part of future RDCO transit system and inclusion of railway transit system in the Okanagan Valley may result in a better transportation network in terms of safety, efficiency, and environmental impact. The cost and benefits associated with a future Okanagan valley railway service could be a very important topic for future research.

- The collision prone zones, Highway 97 collision density map and collision prone intersections were identified in this research. It also identified potential collision prone zones in 2020. More research is needed to determine the actual cause for collision prone locations and to take necessary countermeasures to improve existing road safety. It is also important to take necessary actions to preclude other locations from becoming collision prone.
- The RDCO VISUM transportation planning model calculates volume capacity (VC) ratio for each link. But it was observed that the average zonal volume capacity ratio was very low. Aggregation error might be the problem of having lower values of VC. It was also found that VC is not statistically significant in developing collision prediction models. Lower value of VC could be the problem. So, more research is needed to clarify this VC problem.
- Development of collision prediction models requires two time consuming steps of analysis. In the first step, data are extracted and manipulated that requires manual and computational methods, several software, and several disparate datasets. In the second step, data are entered into regression software, and an iterative procedure is followed to develop the CPMs. This includes selection of error distribution structure (i.e. Poisson or Negative Binomial), stepwise explanatory variable selection, goodness of fit test and refinement of model by outlier analysis. These steps are very time consuming and an automated procedure should be developed to perform these steps.
- Different types of data were obtained from different sources and aggregated to calculate TAZ specific values. But spatial aggregation is related to a bias called Modifiable Areal Unit Problem (MAUP). This problem can significantly change the aggregated value of the variables as well as the parameter estimates of CPM variables. Giving more attention to address the effect of MAUP could be an important topic for future research. Also instead of

using GLM method, other statistical regression methods (i.e. GWPR, FBSA etc.) could be used for developing collision prediction models. These new methods might be useful to increase the accuracy of predictions.

REFERENCES

Allen, B. (1984). Trip distribution using composite impedance. *Transportation Research Record*, 944, 118-127.

Anderson, M. D., and Olander, J. P. (2002). Evaluation of two trip generation techniques for small area travel models. *Journal of Urban Planning and Development*, 128(2), 77-88.

Arasan, V. T., Rangaraju, V. R., Krishna Rao, and K. V. (1996). Trip characteristics of travelers without vehicles. *Journal of Transportation Engineering*, 122(1), 76–81.

ARRB Transport Research. (1999). Evaluation of road safety audits, progress report 1: literature review. *Draft report*, ARRB, Australia.

Asensio, J. (2002). Transport mode choice by commuters to Barcelona's CBD. *Urban Studies*, 39(10), 1881–1895.

Austroads. (2001). R*oad safety audit.* 2nd Edition, Sydney, Australia.

Avery, D.T. and Singer, S. F. (2007). *Unstoppable global warming: every 1,500 years*. Rowman & Littlefield Publishers, Inc., ISBN-978-0742551176

BC Ministry of Education. (2010). *School reports*. Retrieved January 18, 2010 from http://www.bced.gov.bc.ca/reporting/school_data_summary.php

BC Ministry of Transportation (BCMoT). (2010). *Highway data.* Retrieved March 10, 2010 from http://www.th.gov.bc.ca/trafficData/index.htm

BC Ministry of Transportation (BCMoT). (2010). *Annual day of week summary for 2009*. BC Ministry of Transportation and Infrastructure, Retrieved September 14, 2010 from http://www.th.gov.bc.ca/trafficdata/tradas/reports/AllYears/2009/01/AV02/AV02%20%20Sit e % 20Gormans%20Mill%20-%20P-25-11NS%20-%20C%20on%2001-01-2009.pdf

BC Transit. (2006). *Geo-coded bus stop files*. Planning division, Victoria, BC, Canada

BC Transit. (March 2007). *Kelowna Regional Transit: System review and detailed service pla*n. Retrieved January 5, 2010 from http://lakecountry.civicweb.net/Documents /DocumentList.aspx? ID=240

BC STATS. (2010). *Regional population estimates and projections.* Retrieved January 6, 2010 from http://www.bcstats.gov.bc.ca/data/pop/pop/dynamic/PopulationStatistics/SelectRegion Type.asp?category=Census

Ben-Akiva, M.E., and Lerman, S.R. (1985). D*iscrete choice analysis: theory and application to travel demand.* The MIT Press, Cambridge, MA.

Ben-Akiva, M., and Lerman, S. (1985). D*iscrete choice analysis*, MIT Press, Cambridge.

Ben-Akiva, M, Bowman, J., Ramming, S., and Walker, J. (1998). Behavioral realism in urban transportation planning models. *In Transportation Models in the Policy-Making Process: A Symposium in Memory of Greig Harvey*, Asilomar Conference Center, California.

Ben-Akiva, M. and Bierlaire, M. (2003). Discrete choice models with applications to departure time and route choice. *Handbook of Transportation Science*, International Series in operations Research & Management Science, 2003, Volume 56, Part 2, 7-37, DOI: 10.1007/0-306-48058-1_2.

Bhat, C.R., Govindarajan, A. and Pulugurta, V. (1998). Disaggregate attraction-end choice modeling: formulation and empirical analysis. *Transportation Research Record* 1645, 60-68.

Bonard, A.,Lafragette, A. and Dridi-Dastrevigne, E. (2008). *Sustainable development assessment tools dedicated to territorial collectivities.* 6th annual international conference of Territorial Intelligence, Besançon, France

Bonneson, J. A., and McCoy, P.T. (1993). Estimation of safety at two-way stop-controlled intersections on rural highways. *Transportation Research Record*, 1401, 83-89.

Börsch-Supan, A. (1987). *Econometric analysis of discrete choice*. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, Germany, 202- 211.

Bierlare, M. (2007). *Recent methodological developments in discrete choice models*. Report TRANSP-OR 070415 Transport and Mobility Laboratory, School of Architecture, Ecole Polytechnique Federale de Lausanne, Switzerland.

Brude, U., and Larsson J. (1988). The use of prediction models for eliminating effects due to regression-to-the-mean in road accident data. Accident Analysis and Prevention, 20(4), 299-310.

Buchanan, C. (1963). Traffic in towns; a study of the long term problems of traffic in urban areas. Reports of the steering group and working group appointed by the Minister of Transport, Ministry of Transport, London, Great Britain.

Bureau of Public Roads. (1964). Traffic assignment manual. Urban Planning Division, US Department of Commerce, Washington DC.

Caliper Corporation (2010). *TransCAD transportation planning software overview*. Retrieved November 21, 2010 from http://www.caliper.com/tcovu.htm

Canadian Automobile Association (CAA). (2007). *Driving costs.* Retrieved June 18, 2010 from http://www.ama.ab.ca/images/images_pdf/2007-04-27DrivingCostsBrochure2007.pdf

Canadian Capacity Guide (CCG). (2008). *Canadian capacity guide for signalized intersections*. The Institution of Transportation Engineers, Third edition, 3.38-3.29, Retrieved January 6, 2010 from http://www.cite7.org/resources/documents/ CanadianCapacityGuide2008.pdf

Canadian Council of Motor Transport Administrators (CCMTA). (1998). Road safety vision: making Canada's roads the safest in the world. *Annual Report*, Transport Canada, Ottawa, Ontario.

Canadian Council of Motor Transport Administrator (CCMTA), (2005). *Annual report 2005 on road safety vision 2010.* Retrieved January 09, 2010 from http://www.ccmta.ca/english/pdf/rsv_report_05_e.pdf

Canadian Council of Motor Transport Administrator (CCMTA), (2007). *Road safety vision 2010: mid-term review*. Retrieved January 09, 2010 from http://www.ccmta.ca/english/pdf/rsv2010_midtermreport_final.pdf

Carey, M. (1986). A constraint qualification for a dynamic traffic assignment model. *Transportation Science*,20, 55–58.

Cavallo, A. J. (2002). Predicting the peak in world oil production, *Natural Resources Research*, 11[3], 187–195.

Census Canada. (2006). *Boundary files, reference guide*, *Census year 2006.* -Third edition, Catalogue no. 92-160-GIE, Statistics Canada, Retrieved January 08, 2010 from http://abacus.library.ubc.ca.ezproxy.library.ubc.ca/bitstream/10573/41796/1/92-160-GIE200600 3.pdf

Central Okanagan Economic Development Commission (COEDC). (2009a). *2009 economic profile regional district of the central Okanagan.* Retrieved February 12, 2010 from http://www.investkelowna.com/documents/2009RegionalProfile-IntrotoCO.pdf

Central Okanagan Economic Development Commission (COEDC). (2009b). *City of Kelowna: 2009 demographic profile.* Retrieved August 26, 2010 from http://www.regionaldistrict.com/docs/edc/KelProfile09.pdf

Central Okanagan Economic Development Commission (COEDC). (2009c). *District of West Kelowna: 2009 demographic profile*. Retrieved August 26, 2010 from http://www.regionaldistrict.com/docs/edc/DWKProfile09.pdf

Central Okanagan Economic Development Commission (COEDC). (2009d). *District of Lake Country: 2009 demographic profile*. Retrieved August 26, 2010 from http://www.regionaldistrict.com/docs/edc/DLCProfile09.pdf

Central Okanagan Economic Development Commission (COEDC). (2009e). *District of Peachland: 2009 demographic profile*. Retrieved August 26, 2010 from http://www.regionaldistrict.com/docs/edc/DPProfile09.pdf

Chatterjee, A., Khasnabis, L. and Slade, L. J. (1977). Household stratification models for travel estimation. *Journal of Transportation Engineering*, ASCE, 103(1), 199-213.

City of Kelowna. (2003). TModel 2 transportation planning model: TAZ map", Transportation planning office, Kelowna, Canada.

City of Kelowna. (2006). *Route and schedule information.* Regional services department, Kelowna, BC, Canada.

City of Kelowna. (2007). *Corridor Study: Central Okanagan Multi-Modal Corridor*.

City of Kelowna. (2009). *Official Community Plan: Population and housing projections discussion paper.* Retrieved July 4, 2010 from http://www.kelowna2030.ca/spaw2/uploads/files/Population%20and%20Housing%20Projections%20Discussion%20Paper.pdf

City of Kelowna. (2010). *Facts and statistic.*, Retrieved August 26, 2010 from http://www.city.kelowna.bc.ca/ CM/Page182.aspx

Commission for Global Road Safety (CGRF). (2009). *Make roads safe.* ISBN-13: 978-0-9553198-0-8, Retrieved August 05, 2010 from http://www.unece.org/trans/roadsafe/docs/CGRS_Report.pdf

Croft, P. (2005). *Sustainable road safety engineering*. European Transport Conference, France, Strasbourg, Retrieved August 25, 2010 from www.etcproceedings.org/paper/download/ 235

Daly A.J., and Zachary, S. (1978). *Improved multiple choice models*. Determinants of Travel Choice, Hensher and M.Q. Dalvi Editors, Saxon House Studies, Farnborough, UK, 335–357.

Daly, A. (1997). Improved methods for trip generation. Vol. P415 edn, PTRC, ed., Proceedings of the 25th European Transport Conference, London. 207-222.

de Leon, M. R. M., Cal, P. C. and Sigua, R.G. (2005). Estimation of socio-economic cost of road accidents in metro manila, Journal of the Eastern Asia Society for Transportation Studies, 6, 3183 - 3198, Retrieved May 25, 2010 from http://www.easts.info/on-line/journal_06/3183.pdf

de Leur, P., and Sayed, T. (2003). A framework to proactively consider road safety within the road planning process. Canadian Journal of Civil Engineering, 30(4), 711-719.

District of West Kelowna. (2010). Economic Landscape of West Kelowna. Retrieved July 5, 2010 from http://www.districtofwestkelowna.ca/index.aspx?page=39

Domencich, T. A. and McFadden, D. (1975). Urban travel demand: a behavioral analysis. North-Holland/ American Elsevier, New York.

Duffus, L. N., Alfa, A.S., and Soliman, A.H. (1987). The reliability of using the gravity model for forecasting trip distribution. *Transportation*, 14(3),Publisher: Martinus Nijhoff Publishers, ISSN: 0049-4488

Duncan, R. C., and Youngquist, W. (1999). Encircling the peak of world oil production. *Natural Resources Research*, 8[3], Retrieved September 06, 2010 from http://www.npg.org/forum_series/iforums/worldoil.html

Edwards, J. D. (1997). *Crude Oil and alternative energy production forecasts for the twenty-first century: the end of the hydrocarbon era.* AAPG Bulletin, 81[8],1292–1305.

Energy Efficiency and Renewable Energy (EERE). (2009). President Obama calls for greater use of renewable energy. *U.S. Department of Energy*. Retrieved August 20, 2010 from http://apps1.eere.energy.gov/news/news_detail.cfm/news_id=12194

Engelsman, J.C. and van Zyl, N. (2007). *Sustainable safety in the Netherlands and the applicability in South Africa.* 26th Annual Southern African Transport Conference-2007, SATC, ISBN: 192001702X

European Commision (EC). (2010). *Towards a European road safety area: policy orientations on road safety 2011-2020.* Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions, Brussels, Retrieved September 06, 2010 from http://ec.europa.eu/transport/road_safety/pdf /com_20072010_en.pdf

Europa. (2010). *Road safety programme 2011-2020: detailed measures*. Retrieved September 06, 2010 from http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/ 10/343 &format=HTML&aged=0&language=EN

European Union's Council of Ministers. (April 2001). *Transport / Telecommunications*. 2340 council meeting, Luxembourg, Retrieved January 7, 2010 from http://corporate.skynet.be/sustainablefreight/trans-counci-conclusion-05-04-01.htm

Gaspers, K. (2004). *On the road to danger: who call traffic injuries a 'global public health problem*. Safety + Health, June, National Safety Council, Itasca, Illinois, USA.

Gebeyehu, M., and Takano, S. (2007). Diagnostic evaluation of public transportation mode choice in Addis Ababa. *Journal of Public Transportation*, 10(4), 27-50.

Geobase. (2010). *National road network.* Retrieved at July 20, 2009 from http://geobase.ca/geobase/en/search.do?produit=nrn&language=en

Ghareib, A. H. (1996). Evaluation of logit and probit models in mode choice situation. *Journal of Transportation Engineering*, 122(4), 282–290.

Goodman, L.A., and Kruskal, W.H. (1979). *Measures of association for cross classifications*. Spring Series in Statistics, Springer-Verlag, pp.32.

Haberl, H., Erb, K.-H. and Krausmann, F. (2001). How to calculate and interpret ecological footprints for long periods of time: the case of austria 1926–1995. *Ecological Economic,s* 38, 25–45.

Hadayeghi, A., Shalaby, A., and Persaud, B. N. (2003). *Macro-Level accident prediction models for evaluating the safety of urban transportation systems*. Presented at Transportation Research Board Annual Meeting, TRB, Washington, D.C.

Hadayeghi, A., Shalaby, A., and Persaud, B. (2009). Development of planning level transportation safety models using full Bayesian semiparametric additive techniques. *Journal of Transportation Safety & Security,* 2(1),45 – 68.

Hadayeghi, A., Shalaby, A. S., and Persaud, B. (2010). Development of planning level transportation safety tools using geographically weighted Poisson regression. *Accident Analysis and Prevention*, 42(2), 676–688.

Halcrow Group Inc. (2004). *Okanagan lake bridge comprehensive traffic report*. Ministry of Transportation, BC.

Handy, S. (1993). Regional versus local accessibility: implications for non work travel. *Transportation Research Record*. 1400, 58–66.

Hasan. S. and Hoque. M. S. (2008). *A simplified travel demand modeling framework: in the context of a developing country city.* Cooperation for Urban Mobility in the Developing World XIII conference, Vietnam.

Hauer, E., Ng, J.C.N., and Lovell, J. (1988). Estimation of Safety at Signalized Intersections. *Transportation Research Record,* 1185, 48-61.

Hauer, E. (1992). Empirical Bayes approach to the estimation of "unsafety": the multivariate regression method. *Accident Analysis & Prevention*, 24(5), 457-477.

Hauer, E. (1995). On exposure and accident rate. *Traffic Engineering and Control*, 36(3),134-138.

Hauer, E. (1997). *Observational before-after studies in road safety – estimating the effect of highway and traffic engineering measures on road safety.* Elsevier Science Incorporated, Tarrytown, NY, USA.

Hauer, E., Harwood, D.W., Council, F.M., and Griffith, M.S. (2002a). Estimating safety by the Empirical Bayes method: a tutorial. *Transportation Research Record,* 1784, 126- 131.

Hauer, E, Kononov, J., Allery, B., and Griffith, M.S. (2002b). Screening the road network for sites with promise. *Transportation Research Record,* 1784, 27-32.

Hawken, P. (February 2010). *Sustainable world sourcebook: critical issues, viable solutions, resources for actio*n. New Society, ISBN 9780615309804

Highway Capacity Manual. (2000), *Transportation Research Board*, Washington, DC.

Higle, J.L., and Witkowski, J.M. (1988). Bayesian identification of hazardous locations. *Transportation Research Record*, 1185, 24-36.

Hirsch, R. L. (2005). The inevitable peaking of world oil production. *The Atlantic Council of United State*, XVI[3,] Retrieved September 15, 2010 from http://newwebsite.pagnet.org/documents /GTSEP/HirschPOintro.pdf

Ho, G., and Guarnaschelli, M. (1998). *Developing a road safety module for the regional transportation model, technical memorandum one: framework*. ICBC, December, Vancouver, Canada.

Ho, G., Zein, S., and de Leur, P. (2001). *The Canadian road safety audit guide.* Transportation Association of Canada, Ottawa, ON, Canada.

Horowitz, J. L. (1991). Reconsidering the multinomial probit model. *Transportation Research*, Part B: Methodological, 25(6), 433-438.

IBI Group. (April 2005). *Central Okanagan smart transit plan*. Retrieved June 10, 2009 from http://www.kelowna.ca/citypage/docs/pdfs/Transportation%20Division/Smart%20Transit%20Plan%20-%20Executive%20Summary.pdf

Institute of Transportation Engineers (ITE). (2008). Trip generation, 8th edition, ISBN:978-1-933452-43-2

Insurance Corporation of British Columbia. (2000) *Traffic collision statistics, 2000*. Retrieved June December 20, 2009 from http://www.icbc.com/road-safety/safety-research/collisions-2000.pdf

Insurance Corporation of British Columbia. (2004). *Traffic collision statistics: police-attended injury and fatal collisions*. British Columbia, Canada.

Insurance Corporation of British Columbia. (2006). *Traffic collision statistics, 2006*. Retrieved December 20, 2009 from http://www.icbc.com/road-safety/safety-research/collisions-2006.pdf

Insurance Corporation of British Columbia. (2007) *Traffic collision statistics, 2007*. Retrieved December 20, 2009 from http://www.icbc.com/road-safety/safety-research/traffic-coll-stats-2007.pdf

International Energy Association (IEA). (2009a). *Key world energy statistics. 2009.* Retrieved July 5, 2010 from http://www.iea.org/textbase/nppdf/free/2009/key_stats_2009.pdf

International Energy Association (IEA). (2009b). *World energy outlook: exclusive summery*. Retrieved July 10, 2010 from http://www.worldenergyoutlook.org /docs/ weo2009/ WEO2009_es_english.pdf

INRO. (2010). *Transport modelling with Emme-INRO*. Retrieved November 21, 2010 from http://www.inro.ca/en/products/emme/modelling.php

Inland Transport Committee (ITC). (1999). *Consideration of the possible approaches to the work on a United Nations long-term strategy on road safety*. Economic and Social Council, United Nations.

Iwaro, J. and Mwasha, A. (2010). Towards energy sustainability in the world: the implications of energy subsidy for developing countries. *International Journal of Energy and Environment*, 1[4], 705-714.

Jones-Lee, M.W., Hammerton, M. and Philips, P. (1985). The value of safety: results of a national sample survey. *The Economic Journal*, 95(377), 49–72.

Jones, P., F. Koppelman , and J.P. Orfueil. (1990). *Activity analysis: state-of-the-art and future directions*. In P. Jones (ed.) Developments in Dynamic and Activity-Based Approaches to Travel Analysis, Gower Publishing, Aldershot

Jordan, P. (2001). *Road safety audit – low costs, big benefits*. Editor F.P.D. Navin, Proceedings from the Fourth International Conference on Accident Investigation, Reconstruction, Interpretation and the Law, hosted by the Department of Civil Engineering, University of British Columbia, Vancouver, Canada, 179-183.

Kelowna Pacific Railway (KPR). (2010). *Short line Overview.* Retrieved May 6, 2010 from http://cnplus.cn.ca/IT/Shortlines/SL_Static.nsf/shortlines/52BA5D16C4C3C7D2862568EB00645FE4?opendocument

Khisty, C. J., and Lall, B. K. (1998). *Transportation engineering: an introduction*. 2nd Edition, Prentice Hall, Upper Saddle River, NJ, USA, 664-696.

Khondaker, B., Sayed, T. and Lovegrove, G.R. (2009). Transferability of community-based macro-level collision prediction models between different time-space regions for use in road safety planning applications. *Journal of Transportation Engineering*, 136(10), 871-880.

Kim, H.S. and Kim, E. (2004). Effects of public transit on automobile ownership and use in households of the USA. *Journal of the Applied Regional Science,* 16[3].

Kitamura, R. (1981). A stratification analysis of taste variations in work trip mode choice. *Transportation Resea*rch, 15A(6), 473-485.

Kitamura, R., Pendyala, R.M., Pas, E.I. and Reddy, P. (1995). *Application of AMOS, an activity-based TCM evaluation tool, to the Washington, D.C., metropolitan area*. 23rd European Transport Forum: Proceedings of Seminar E Transportation Planning Methods, PTRC Education and Research Services, Ltd., London, 177-190.

Kitamura,R. (1996). *Applications of models of activity behavior for activity based demand forecasting*. pp. 119-124, Texas Transportation Institute: Activity-based travel forecasting conference, Texas.

Koppelman, F. S. and Bhat, C. (2006), *A self instructing course in mode choice modeling: multinomial and nested logit models*. Prepared For U.S. Department of Transportation Federal Transit Administration.

Kulmala, R., and Roine, M. (1988). *Accident Prediction models for two-lane roads in Finland.* Conference on traffic safety theory and research methods proceedings, Session 4: Statistical analysis and models, Amsterdam: SWOV, 90-103.

Kulmala, R. (1995). *Safety at rural three- and four-arm junctions: development and application of accident prediction models.* Dissertation for the degree of Doctor of Technology, Technical Research Centre of Finland,VTT Publication 233, Espoo, Finland.

Ladron de Guevara, F., Washington, S. P., and Oh, J. (2004). Forecasting crashes at the planning level: a simultaneous negative binomial crash model applied in Tucson, Arizona. *Transportation Research Record*, 191-199, ISSN: 0361-1981.

Levinson, D. M. (1994). An evolutionary transportation planning model: structure and application. *Transportation Research Board*, 1493, 64-73.

Levinson, D.M., Kumar, A. (1995). Multi-modal trip distribution: structure and application. *Transportation Research Record,*-1466, 124-131.

Lin, F. C. and Navin, F. P. D. (1999). *Errors in transportation planning*. Civil Engineering Department, University of British Columbia, Vancouver, Canada.

Litman, T. (2004). *Evaluating household savings from high quality public transit service*. Victoria Transport Policy Institute.

Litman, T. and Burwell, D. (2006). Issues in sustainable transportation. *International Journal of Global Environmental Issues*, 6[4], 331-347.

Litman, T. (2007). *Win-win transportation solutions*. Victoria transport policy institute

Litman, T. (2009a). *Sustainable transportation indicators: a recommended research program for developing sustainable transportation indicators and data*. Presented at the 88th annual meeting of Transportation Research Board, TRB, Washington, DC, USA.

Litman, T. (2009b). *Introduction to multi-modal transportation planning*. Victoria transport policy institute, Victoria, Canada.

Litman. T. (2010). *Evaluating transportation land use impacts*. Victoria transport policy institute, Retrieved September 05, 2010 from http://www.vtpi.org/landuse.pdf

Lo, H.K. and Szeto, W.Y. (2002). A Cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research Part B*, 36, 421- 443.

Lord, D.e, and Persaud, B. N. (2004). Estimating the safety performance of urban road transportation networks. *Accident Analysis & Prevention*, 36, 609- 620.

Lovegrove, G., and Stanos, J., (2006). *Okanagan Valley Quest Model Transportation Database.* Technical report.

Lovegrove, G. R., and Sayed, T. (2006a). Macro-Level collision prediction models for evaluating neighbourhood traffic safety. *Canadian Journal of Civil Engineering*, 33(5),609–621.

Lovegrove, G.R. and Sayed, T. (2006b). Using macro-level collision prediction models in road safety planning applications. *Transportation Research Record,* 1950, 73 - 82.

Lovegrove, G. R. (2007). *Road safety planning: new tools for sustainable road safety and community development.* VDM Verlag Dr. Muller, Saarbrucken.

Lovegrove, G., and Sayed, T. (2007). Using macro-level collision prediction models to enhance traditional reactive road safety improvement programs. *Transportation Research Record*, 65-73, ISSN: 0361-198.

Lovegrove, G., Lim, C., and Sayed, T. (2010). Community-Based, macrolevel collision prediction model use with a regional transportation plan. *Journal of Transportation Engineering*, 136(2), 120-128.

Manski, C. (1977). *The structure of random utility models*. Theory and Decision 8, pp. 229-254.

Mark Bradley Research & Consulting (MBRC) (2009). *Strategy for activity-based travel demand model development with travel survey.* SCAG PROJECT 09-012, Southern California Association of Governments. 24.

Mathew, T. V, Krishna Rao, and K. V. (2007). *Chapter 7: trip generation.* Introduction to Transportation Engineering, Department of Civil Engineering, Indian Institute of Technology, Mumbai, India.

Mathew, T. V. (2007). *Chapter 10: traffic assignment.* Introduction to Transportation Engineering, Department of Civil Engineering, Indian Institute of Technology, Mumbai, India

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behaviour*. In P. Zarembka (Ed.), Frontiers in Econometrics. Academic Press, New York.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record*, 672, 72-77.

McFadden, D., and Train, K. (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15, 447-470.

McGuigan, D.R.D. (1982). *Non-Junction Accident rates and their use in 'black-spot' identification*. Traffic Engineering & Control, 60-65.

McNally, M. G. (2000). The four step model. Handbook of Transport Modelling, 1st Ed., D. A. Hensher and K. J. Button, eds., Elsevier Science, Kidlington, Oxford, U.K.

McNally, M.G., and Rindt, C. R.(2008). The activity-based approach. in: D.A. Hensher and Button, K.J. (eds.) Handbook of Transport Modelling, 2nd edition, Elsevier, Amsterdam, 55 – 73.

Merchant, D. K., and Nemhauser, G. L. (1978), A model and an algorithm for the dynamic traffic assignment problem. *Transportation Science*, 12, 183–199.

Meyer, M. D., and Miller, E. J. (1984). *Urban transportation planning: a decision-oriented approach.* McGraw-Hill, New York.

Meyer, M. D., and Miller, E. J. (2001). Urban transportation planning: a decision-oriented approach. 2nd Ed., McGraw-Hill, New York, ISBN-13: 978-0072423327

Miaou, S., and Lum, H. (1993). Modeling vehicle accident and highway geometric design relationships. *Accidents Analysis and Prevention*, 25(6), 689-709.

Miller, T. R. (1993). *Crash Costs for British Columbia. economic analysis working paper prepared for planning services branch*. B.C. Ministry of Transportation & Highways, Victoria, Canada.

Miller, T. R. (1998) *Benefit-Cost analysis: past and future direction.* Proceedings of a Conference on Highway safety: at the crossroads. March. Sponsored by the highway division of the American Society of Civil Engineering, Texas.

Mishan, E.J. (1971). Evaluation of life and limb: a theoretical approach. *The Journal of Political Economy*, 1971, 79 (4), 687-705.

Muranyi ,T. C. and Miller, J. D. (1966*). New applications of the opportunity model of traffic planning and design.* Chicago Area Transportation Study, 130 North Franklin Street, Chicago.

Nicholson, A.J. (1980). *Identification of hazardous locations*. PRU Newsletter No. 66, Ministry of Transport, New Zealand.

Numerical Algorithms Group (NAG). (1994). *The GLIM system, release 4 manual.* The Royal Statistical Society.

Organisation for Economic Co-Operation and Development (OECD). (2001). *Strategies to reduce greenhouse gas emissions from road transport: analytical methods*. Retrieved March 20, 2010 from http://www.internationaltransportforum.org/Pub/pdf /02GreenhouseE.pdf

Ortúzar, J. de D. (1983). Nested logit models for mixed-mode travel in urban corridors. *Transportation Research*, 17A(4), 283–299.

Oruzar, J. de D. and Willumsen, L.G. (1990). *Modelling Transport.* Chichester: Jihn Wiley & Sons Ltd. Xvi-376, ISBN 0 471 92629 9

Ortuzar, J., and Willumsen, L. G. (2001). *Modelling transport*, Wiley, West Sussex, U.K.

Parry, I. W.H. (2004). Comparing alternative policies to reduce traffic accidents, *Journal of Urban Economics*, 54[2], 346-368, Table 2.

Pimentel, D. (2008). Chapter 5: Pick oil, investments and the economy in an uncertain future. *Biofuels*, Solar and Wind as Renewable Energy Systems, Springer, ISBN: 978-1-4020-8653-3

Phillips Farevaag Smallenberg (PFS). (2009). *University of British Columbia Okanagan Master Plan*.

Poppe, F. (1995). Risk figures in the traffic and transport evaluation module (EVV): A contribution *to the definition study traffic safety in EVV*. Rep. No. R-95-21, Dutch Road Safety Research Institute (SWOV), Leidschendam, The Netherlands.

Poppe, F. (1997a). *Traffic models: inner areas and road dangers*. Rep. No. R-97-10, Dutch Road Safety Research Institute (SWOV), Leidschendam, The Netherlands.

Poppe, F. (1997b). *Cost-benefit analysis of a sustainably safe road traffic system*. SWOV Report D-97-17, Leidschendam, Netherlands.

PTV traffic mobility logistic (2010). *PTV AG: VISUM*. Retrieved November 21, 2010 from http://www.ptvag.com /software/transportation-planning-traffic-engineering/software-system-solutions /visum/

Racca, D. P., and Edward C. R. (2004). *Project report for factors that affect and/or can alter mode choice.* Delaware Transport Institute and State of Delaware Department of Transport.

Ramage-Morin, P. L. (2008). *Motor vehicle accident deaths, 1979 to 2004*. Component of Statistics Canada Catalogue no. 82-003-X Health Reports, Vol. 19, no. 3, Statistics Canada. Retrieved August 27, 2010 from http://www.statcan.gc.ca/pub/82-003-x/2008003/article/10648-eng.htm

Ran, B., and Boyce, D. (1996). *Modeling dynamic transportation networks. an intelligent transportation system oriented approach*. Second Revised Edition. Springer-Verlag, Heidelberg.

Rees, W.E. and Wackernagel, M. (1996). Urban ecological footprints: why cities cannot be sustainable – and why they are the key to sustainability. *Environmental Impact Assessment Review,* 16, 223–248.

Resource Planning Group (RPG) INC. (2007). *Okanagan College master plan*, Executive Summary, Retrieved July 13, 2010 from http://www.okanagan.bc.ca/Assets/Departments+%28Administration%29/Legal+Affairs/Campus+Plan+Executive+Summary.pdf

Ruckelhaus, W.D. (1989). *Toward a sustainable world*. Scientific American, September, 114–120.

Rudjito, D. (2006). *The application of fuzzy logic to traffic assignment in developing countries*. Doctoral thesis to Department of Civil and Environmental Engineering, Imperial College of Science Technology and Medicine London

Ruiter, E. R. (1967). Toward a better understanding of the intervening opportunities model. *Transportation Research,* 1, 47-51

Sawalha, Z., and Sayed, T. (1999). *Accident prediction and safety planning models for urban arterial roadwa*ys. Report for the Insurance Corporation of British Columbia by the Department of Civil Engineering, University of British Columbia, Vancouver, Canada.

Sawalha, Z.A. (2002). *Traffic accident modeling: statistical issues and safety applications*. Dissertation for the degree of Doctor of Philosophy, Faculty of Graduate Studies, Department of Civil Engineering, University of British Columbia, April, Vancouver, Canada.

Sawalha, Z., and Sayed, T. (2005a). Traffic accident modeling: some statistical issues, Canadian *Journal of Civil Engineering*, 33(9), 1115–1124.

Sawaiha, Z., and Sayed. T. (2005b). Transferability of accident prediction models. *Safety Science*, 44(3), 209-219,

Sayed, T, Abdelwahab, W., and Navin, F. (1995). Identifying accident-prone locations using fuzzy pattern recognition. *Journal of Transportation Engineering*, ASCE, 121(4), 352-358.

Sayed, T., & Abdelwahab, W. (1997). Using accident correctability to identify accident-prone locations. *Journal of Transportation Engineering*, 123(2), ASCE, 107-113.

Sayed, T., Navin, F., and Abdelwahab, W. (1997*).* A countermeasure-based approach for identifying and treating accident prone locations*. Canadian Journal of Civil Engineering*, 24, CSCE, Canada, 683-691.

Sayed, T. (1998a). *New and essential tools, part 1: selecting sites for treatment*. Proceedings from Seminar on New Tools for Traffic Safety, Section 3, hosted by the Institute of Transportation Engineers, August 9th, ITE, Toronto, Canada, 1-13.

Sayed,T. (1998b). *Road safety improvement programs*. ICBC Consultant Training, 6-14.

Sayed, T., and de Leur, P. (2001a). *Program evaluation report: road improvement program*. Prepared for the Insurance Corporation of British Columbia, ICBC, North Vancouver, Canada.

Sayed, T., Saunier, N., Lovegrove, G. and de Leur, P. (2010). *Advances in proactive road safety planning*. 20th Canadian Multidisciplinary Road Safety Conference, Niagara Falls, Ontario.

Seo, J., Ishiguro, K., and Inamura, H. (2000). An overview of trip generation model and future perspective: focus on the freigth transportation, *DJTSGKHKG journal*, 1999, 438-439.

Sosslau, A. B., Hassam, A. B., Carter, M. M., and Wickstrom, G. V. (1978). Quick response urban travel estimation techniques and transferable parameters: user's guide. *Transportation Research Board*, 187.

Steg, L. and Gifford, R. (2005). Sustainable transportation and quality of life. *Journal of Transport Geography*, 13. 59-69.

Stopter, P.R., and Meyburg, A. H. (1975). *Urban transportation modelling and planning*. Lexington Books, D. C. Heath and Co., Lexington, MA.

Stouffer, S. A. (1940). Intervening opportunities: a theory relating to mobility and distance. *American Sociological Review*, 5[6], 845–867. doi:10.2307/2084520.

Stopher, P. R., and McDonald, K. G. (1983). Trip generation by cross-classification: an alternative methodology. *Transportation Research Record*, 944, pp. 84-91

Suh, S., Park, C.H., and Kim, T.J. (1989). A highway capacity function in Korea: measurement and calibration. *Transportation Resea*rch, 24A, 177 - 86.

Susan Krumdieck, S., Page, S. and Dantas, A. (2010). Urban form and long-term fuel supply decline: A method to investigate the peak oil risks to essential activities. *Transportation Research Part A: Policy and Practice,* 44[5], 306-322.

SWOV (Institute for Road Safety Research, The Netherlands). (2006). *Advancing Sustainable Safety*. National Road Safety Exploration for 2005-2020, Retrieved August 05, 2010 from http://www.swov.nl/rapport/DMDV/Advancing_Sustainable_Safety_brief.pdf

Szeto,W. Y., and Lo, H. K. (2006). Dynamic traffic assignment: properties and extensions. *Transportmetrica Journal*, 2(1), 31-52.

Tanner, J.C. (1953). Accidents at rural 3-way junctions, *Journal of Institution of Highway Engineers*, 2[1], 56-67.

Tennoy, A. (2010). Why we fail to reduce urban road traffic volumes: does it matter how planners frame the problem. *Transport Policy* ,17(4), 216-223.

Train, K. (1998). Recreation demand models with taste differences over people. *Land Economics,* 74, 230-240.

Transport Canada. (2004). *Health Canada. 2004, road safety in Canada – an overview*. ISBN 0-662-36440-6, Ottawa, Canada.

Transport Canada. (2010). *Canadian motor vehicle collision statistics: 2007.* Retrieved January 06, 2010 from http://www.tc.gc.ca/eng/roadsafety/tp-tp3322-2007-1039.htm

United Nations (UN) (1987). *Our common future.* Chapter 2: Towards Sustainable Development, A/42/427, UN documents.

United Nations (UN). (2003). Global road safety crisis: report of the secretary general. *General Assembly,* Retrieved August 23, 2010 from http://www.searo.who.int/LinkFiles /whd04_documents_un_secretary_general_report_on_orad_traffic_injuries.pdf

United Nations (UN). (2007). *UN global road safety week*. Retrieved August 15, 2010 from http://www.un.org/events/roadsafety/

United Nations (UN). (2009). *UN decade of action for road safety 2011-2020.* Retrieved August 05, 2010 from http://www.makeroadssafe.org/publications/Documents /decade_is_action_booklet.pdf

U.S. Energy Information Administration (USEIA). (2010). *International energy outlook – 2010.* U.S. Department of Energy, Retrieved September 14, 2010 from http://www.eia.doe.gov/oiaf/ieo/pdf/0484%282010%29.pdf

United State Environment protection Agency (USEPA). (2006). *Greenhouse gas emissions from the U.S. transportation sector 1990–2003.* Retrieved March 20, 2010 from http://www.epa.gov/oms/climate/420r06003.pdf

van Minnen, J. (1999). *The suitable size of residential areas: a theoretical study with testing to practical experiences*. SWOV Report R-99-25, Leidschendam, Netherlands.

Van Schagen, I. and Janssen, T. (2000). Managing transport risks – sustainable safety in the netherlands. *Journal of IATSS Research*, ISSN:0386-1112, 24[2], 18-27.

van Schagen, I. and Theo, J. (2000). Managing road transport risks: sustainable safety in the Netherlands. *Risk Management in Transport*, IATSS Research, 24(2), 18-27.

van Vliet, D. and Dow, P.C. (1979). Capacity restrained road assignment 1, the convergence of stochastic methods. *Traffic Engineering and Control*, 20, 296 - 299.

VISUM. (2007). *VISUM 10, User Manual*, PTV Vision.

Vodden, K., Smith, D., Eaton, F. and Mayhew, D. (2007). *Analysis and estimation of the social cost of motor vehicle collisions in Ontario*. Transport Canada, Retrieved July 05, 2010 from http://www.tc.gc.ca/media/documents/roadsafety/tp14800e.pdf

Wardrop, J.G. (1952). Some theoretical aspects of road traffic research. *Institution of Civil Engineer*, Part II, Vol. [36], 325 - 62.

Wegman, F. (1996). *Sustainable safety in the Netherlands* SWOV Institute for Road Safety Research, Leidschendam, The Netherlands.

Wiederkehr, P.,Gilbert, R.,Crist, P. and Caïd, N. (2004). Environmentally sustainable transport (EST): concept, goal, and strategy – The OECD's EST PROJECT. *European Journal of Transport and Infrastructure Research* , 4[1], 11-25.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning*, Part A, 9, 285-344.

Wills, M. J. (1986). A flexible gravity-opportunities model for trip distribution. *Transportation Research*, 20B, 89-111.

Wilson, E.M. (2004). NCHRP synthesis of highway practice 336: road safety audits. *Transportation Research Board*, National Research Council, Washington, D.C.

Winram, J. (2007). *2007 North and Central Okanagan household travel survey*. The City of Vernon & the City of Kelowna, BC.

World Bank. (2010). *Country classifications.* Retrieved July 01, 2010 from http://data.worldbank.org/about/country-classifications

World Health Organization (WHO). (2004). *World Report on Road Traffic Injury Prevention: Summary.* Geneva, Retrieved June 5, 2009 from http://www.who.int/violence_injury_prevention/publications/ road_traffic/world_report/en/

World Health Organization (WHO). (2008). *World health statistics: 2008*. Retrieved February 1, 2010 from http://www.who.int/whosis/whostat/EN_WHS08_Full.pdf

World Health Organization (WHO). (2009). *Global Status report on road safety: time for action*. Retrieved December 22, 2009 from www.who.int/violence_injury_prevention/road_safety_status/2009

Zegeer, C.V., and Dean, R.C. (1977). Identification of hazardous locations on city streets. Traffic Quarterly 31, ENO Foundation for Transportation, Connecticut, USA, 549-570.

Zegeer, C. V. (1982). *Highway Accident Analysis Systems.* NCHRP, 91, Transportation Research Board, Washington, D.C.

Zhengbing, H., Shoufeng, M. and Xiting, T. (2009). Empirical study on the influence of learning ability to individual travel behaviour. *Journal of Transportation Systems Engineering And Information Technology*, 9(2), 75-80.

APPENDICES

APPENDIX A Example of an Activity Based Modeling Framework

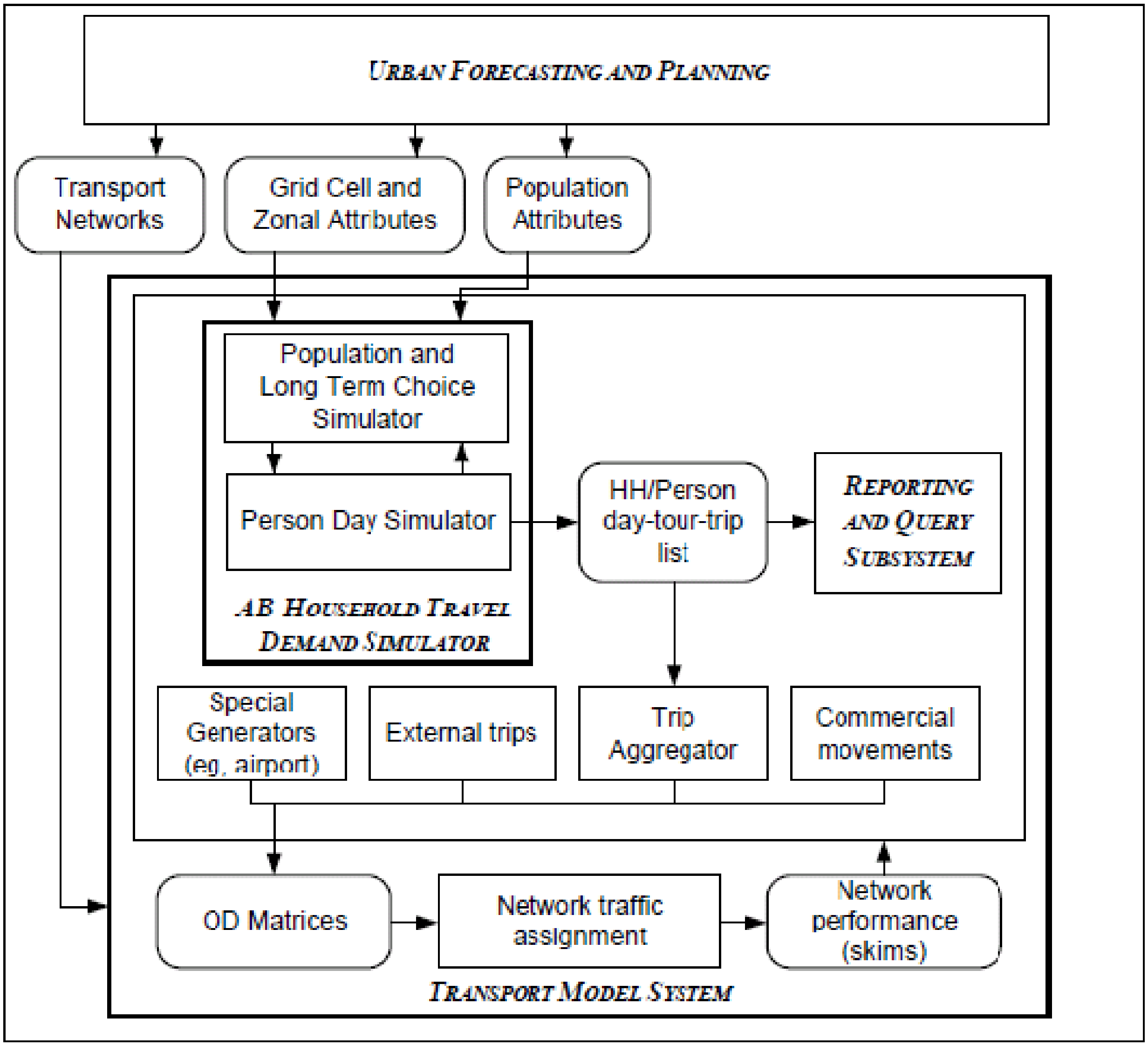


Figure A.1 Example of an activity based modeling framework (MBRC, 2009)

APPENDIX B Traffic Analysis Zones (TAZs) of the RDCO

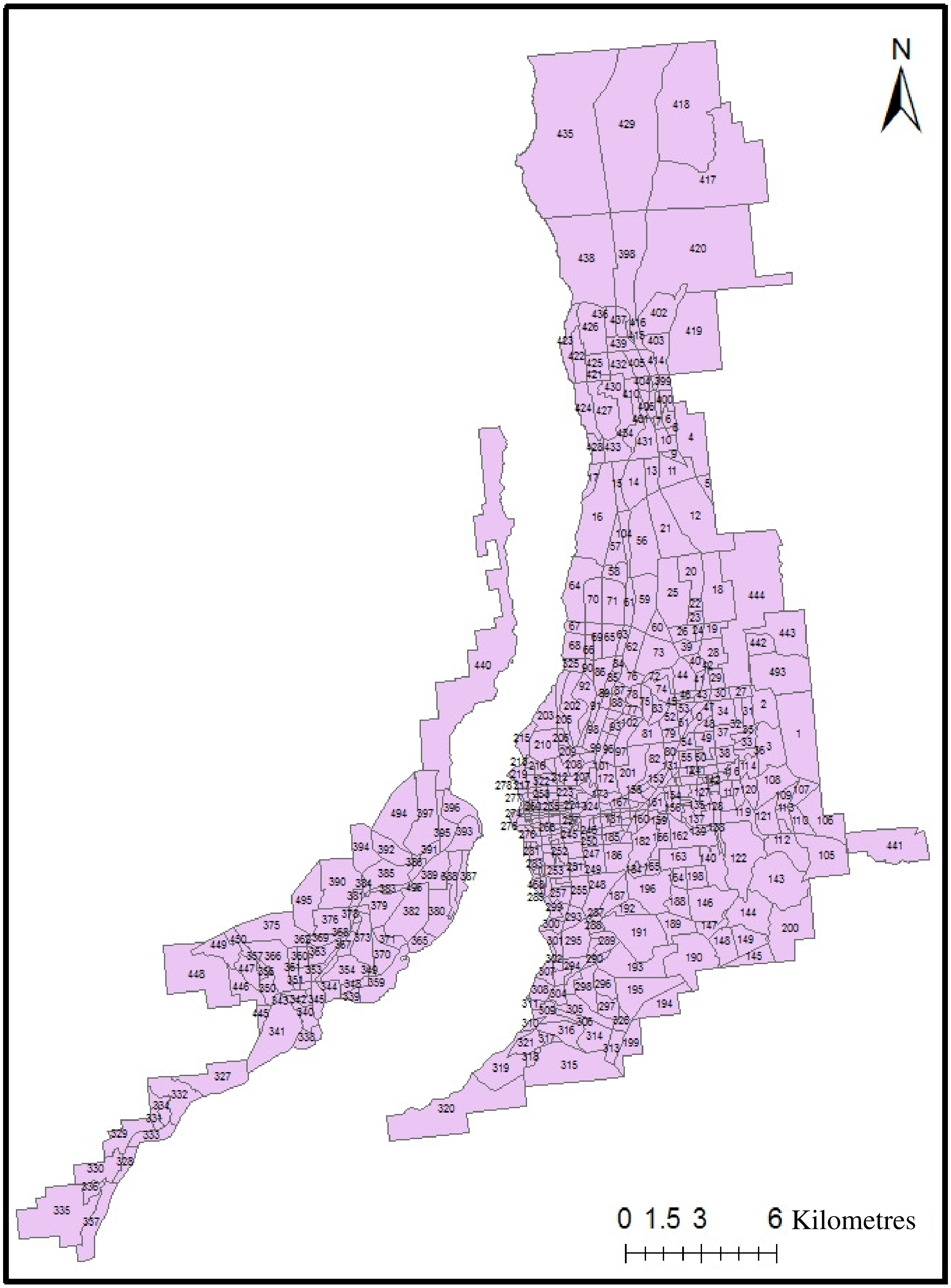
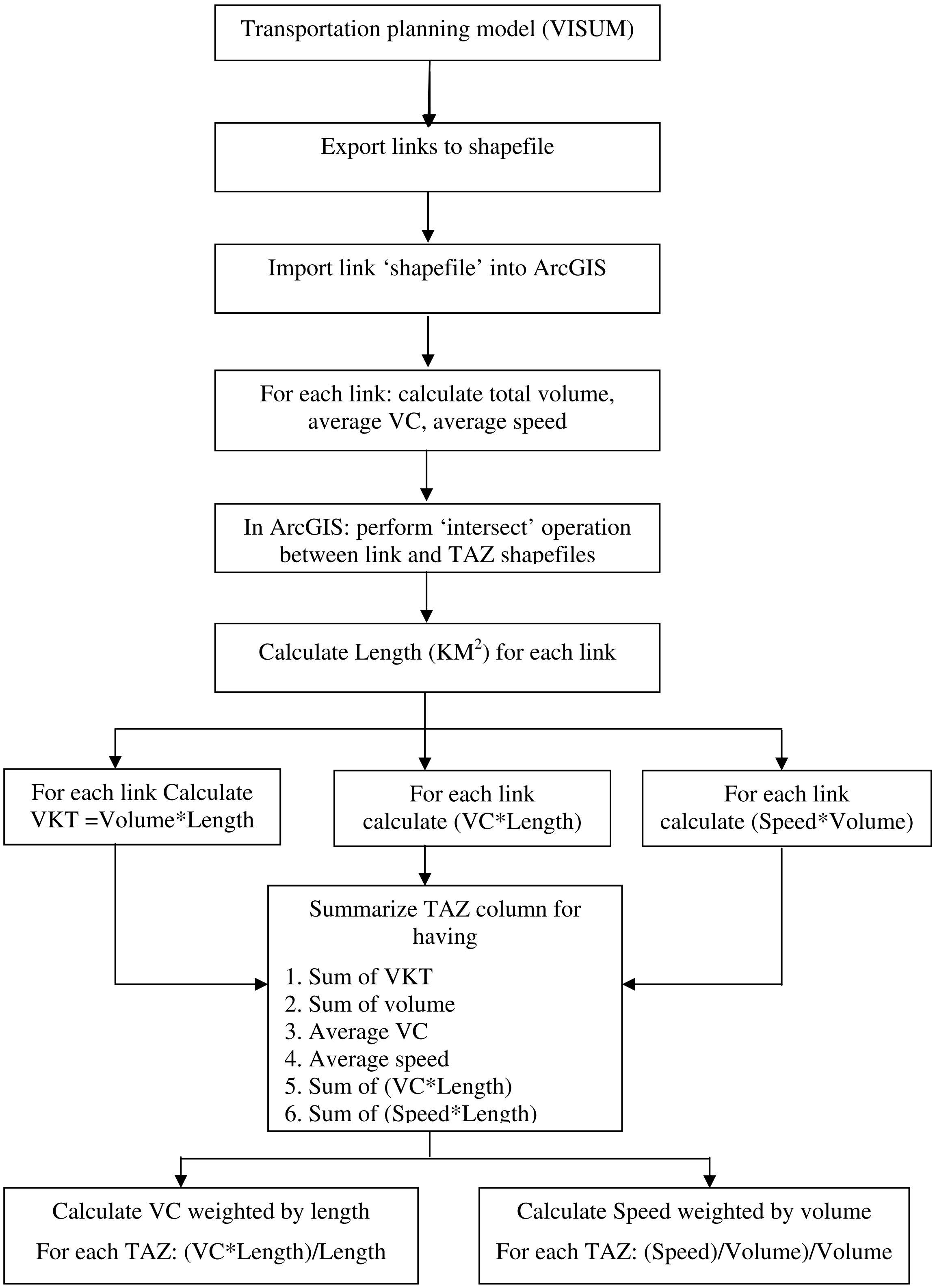


Figure B.1 TAZ map of the RDCO

Table B.1 RDCO traffic analysis zone statistics

	City of Kelowna	District of West Kelowna	District of Lake Country	District of Peachland	RDCO
No. of urban zone	211	25	6	0	242
No. of rural zone	166	45	36	11	258
Total area (km^2)	237.71	77.97	110.75	17.15	443.58
Average area (km^2)	0.63	1.11	2.64	1.56	0.89
Minimum area (km^2)	0.01	0.02	0.05	0.31	0.01
Maximum area (km^2)	9.77	15.33	19.22	5.27	19.22
Standard deviation	0.98	1.96	4.57	1.34	1.84

APPENDIX C Steps Involved in the Calculation of Modeled Variables



APPENDIX D Number of AM Period Trips (2007 Okanagan Household Survey)

Table D.1 Number of AM period trips (2007 Okanagan household survey)

Trip Destination	Trip Origin								
	RDCO			RDCO					
	City of Vernon	CS, Lumby (east)	Spal/Arm/End (north)	Lake Country (north & east)	West side (WB & PL) (west)	Central Kelowna	Sub. Kelowna	**Out of region**	**total**
RDNO	19,644	7,491	11,578	731	395	261	611	24	40,735
City of Vernon	16,409	3,197	4,034	674	272	261	427	24	25,298
CS, Lumby (east)	1,168	3,886	387	37	25	0	157	0	5,660
Spal/Arm/End (north)	2,067	408	7,157	20	98	0	27	0	9,777
RDCO	1,573	458	546	7,377	18,376	22,066	43,999	124	94,519
Lake Country (north & east)	185	18	22	3,847	66	230	699	0	5,067
West side (WB & PL) (west)	91	0	72	75	11,670	1,058	898	74	13,938
Central Kelowna	820	343	236	1,690	5,017	14,603	20,228	22	42,959
Sub. Kelowna	477	97	216	1,765	1,623	6,175	22,174	27	32,554
Out of region	192	18	589	56	677	113	435	22	2,102
Total	21,409	79,67	12,713	8,164	19,448	22,440	45,045	170	13,7356

APPENDIX E 2010 Field Trip Routes

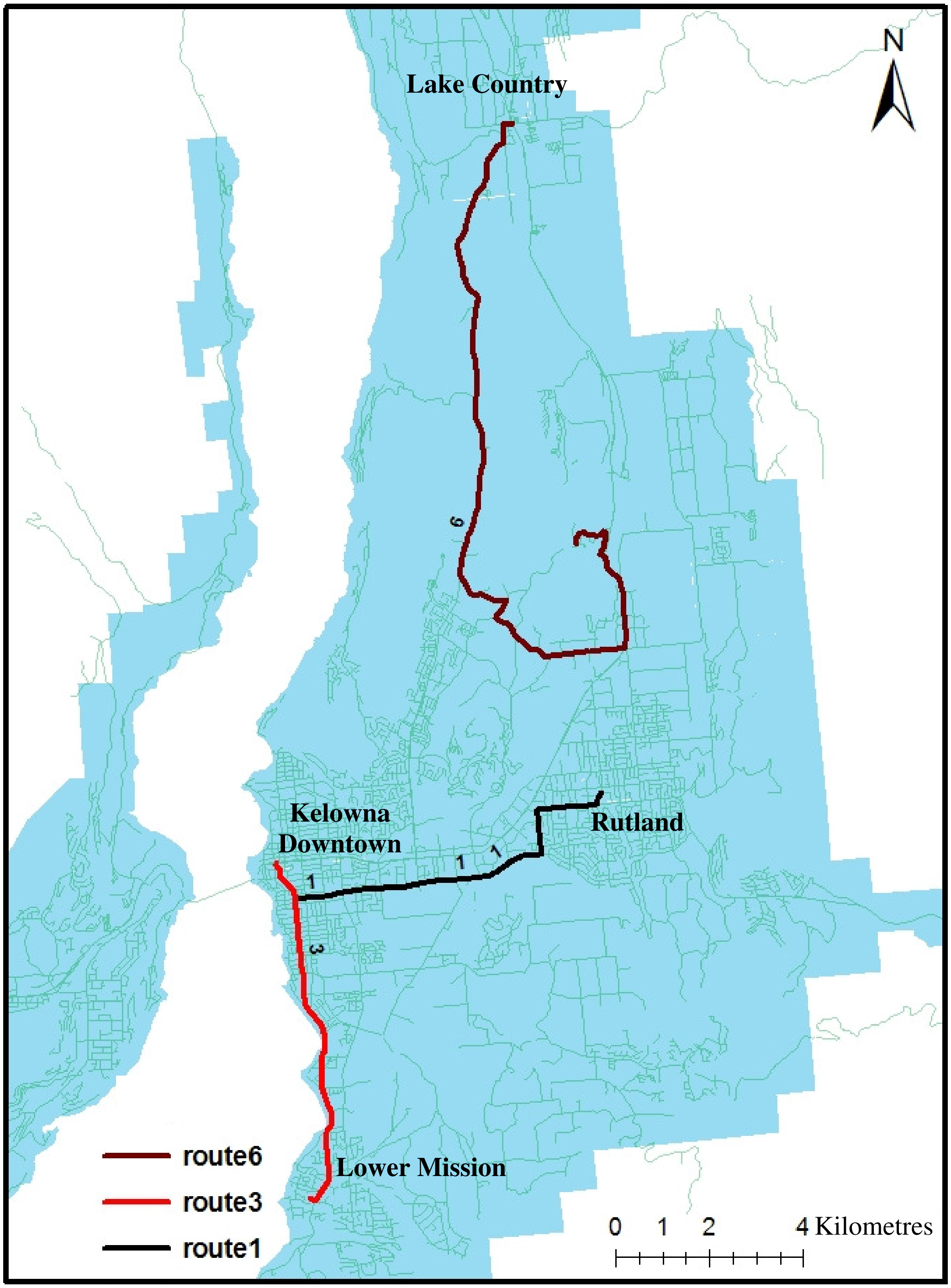


Figure E.1 2020 field trip routes # 1, 3 & 6

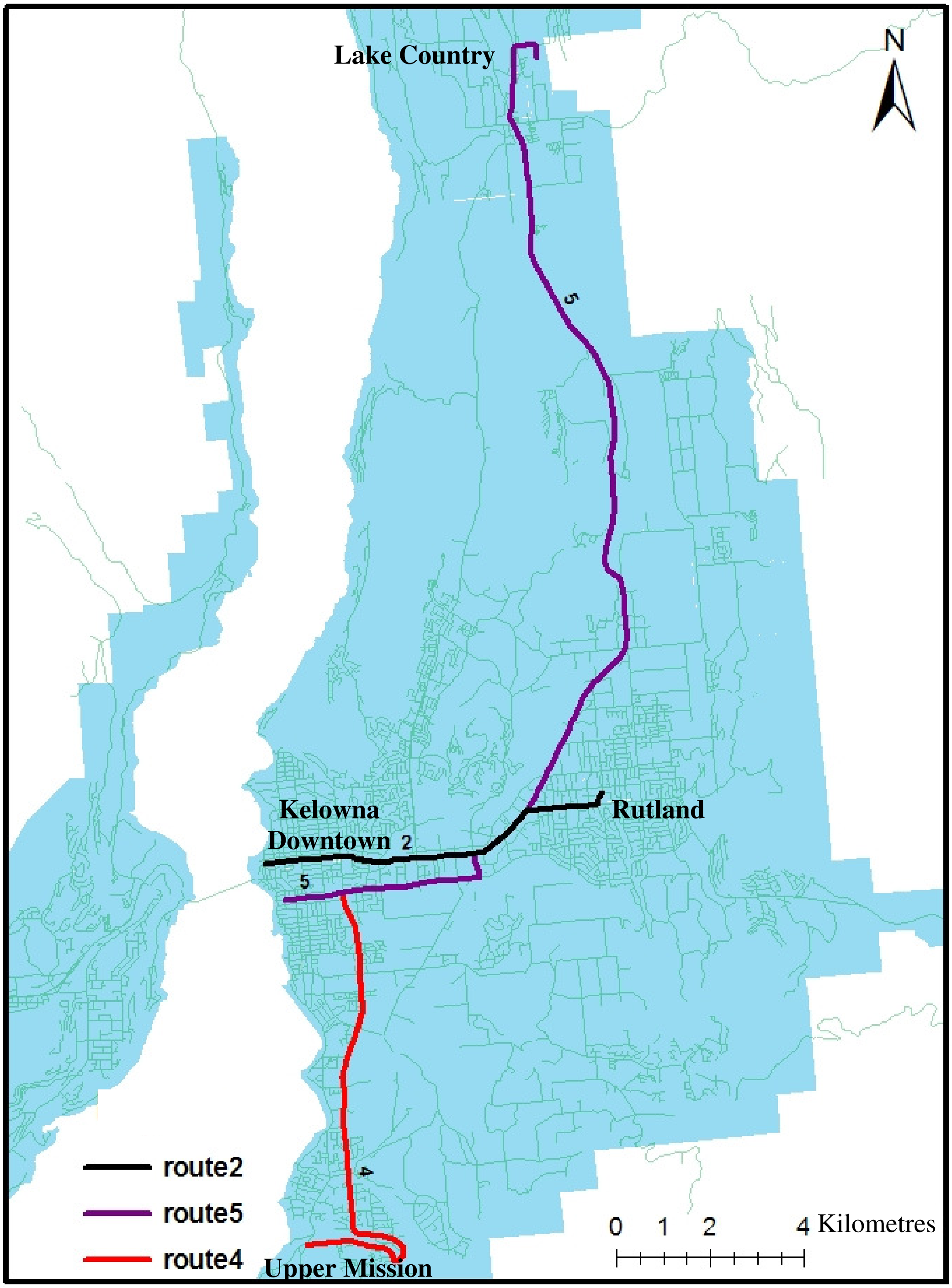


Figure E.2 2020 field trip routes # 2, 4 & 5

APPENDIX F 2006 Mode Share Equation Results

$$Auto\ user\ (\%) = 33.13.e^{-0.000925*BStopD+0.002349*BH+0.3815*VHH+0.0024*Pop45\_64}$$

where,

Auto user (%) = percentage of auto user from a zone;
BStopD = bus stop density (i.e. no. of bus stops/ area in KM^2)
BH = bus headways (in minutes);
VHH = vehicles per household; and,
Pop45_64 = percentage of population between age 45 and 64.

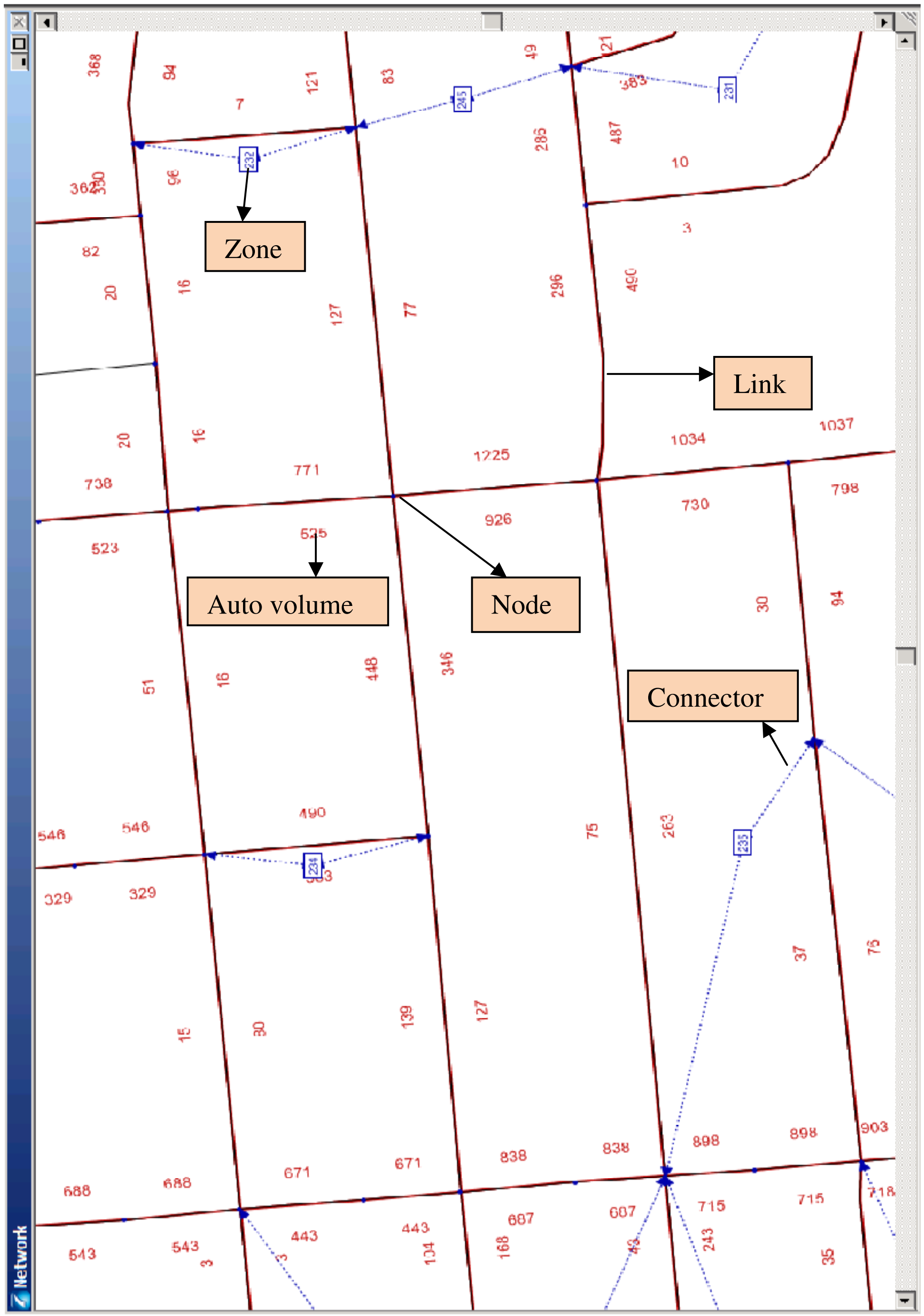
<u>Mode share equation parameter estimates:</u>

Parameter	estimate	s.e.	t(232)	t pr.	antilog of estimate
Constant	3.507	0.125	28.16	<.001	33.34
Pct45_64age	0.00459	0.00216	2.13	0.034	1.005
BH	0.002699	0.000662	4.07	<.001	1.003
BStopD	-0.000937	0.000388	-2.41	0.017	0.9991
VHH	0.3752	0.0523	7.1	<.001	1.455

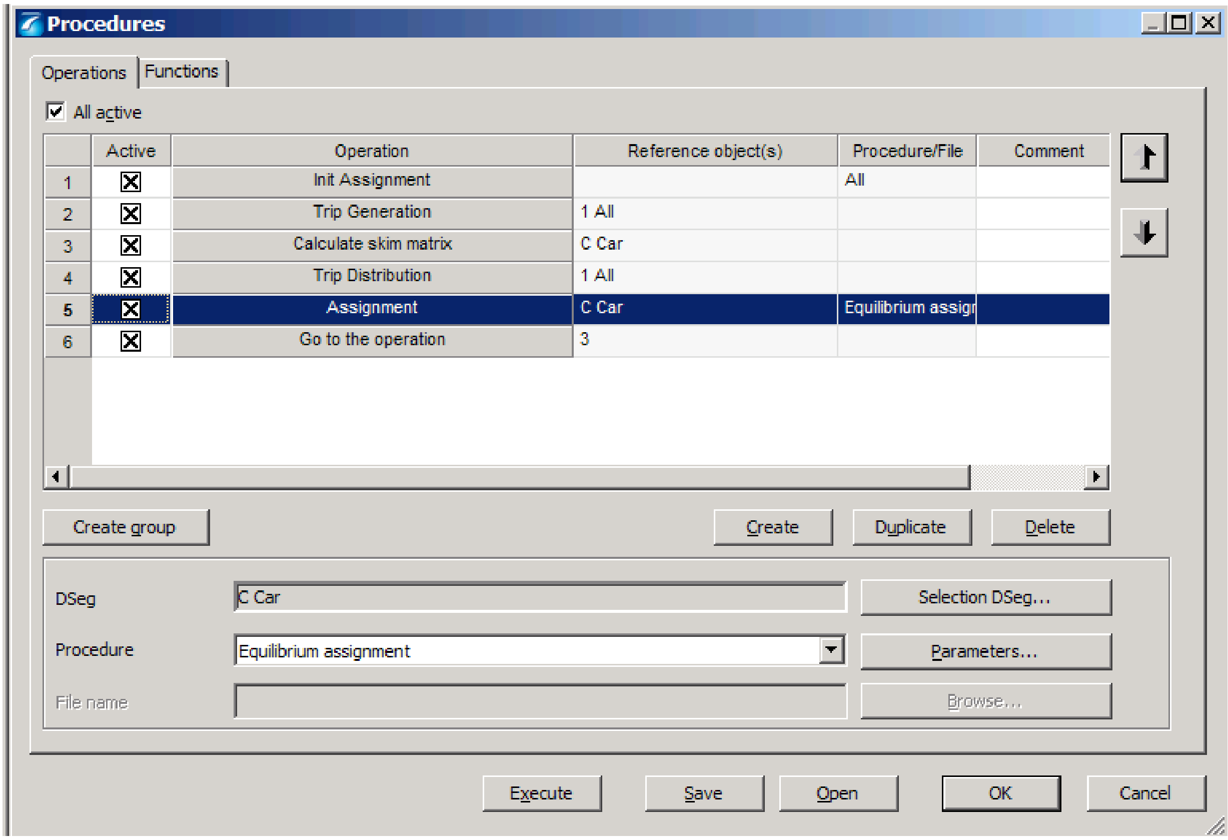
Table F.1 Correlations between parameter estimates

Parameter	Constant	Pop_45_64	BH	BStopD	VHH
Constant	1	-0.215	-0.7	-0.343	-0.961
Pct45_64age	-0.215	1	-0.281	0.117	-0.007
BH	-0.215	-0.281	1	0.276	-0.012
BStopD	-0.343	0.117	0.276	1	0.239
VHH	-0.961	-0.007	-0.012	0.239	1

APPENDIX G VISUM Screenshot of the 2006 RDCO Transportation Planning Model



APPENDIX H Sample Procedure of Transportation Modeling in VISUM



APPENDIX I Transit Coverage Map in 2006

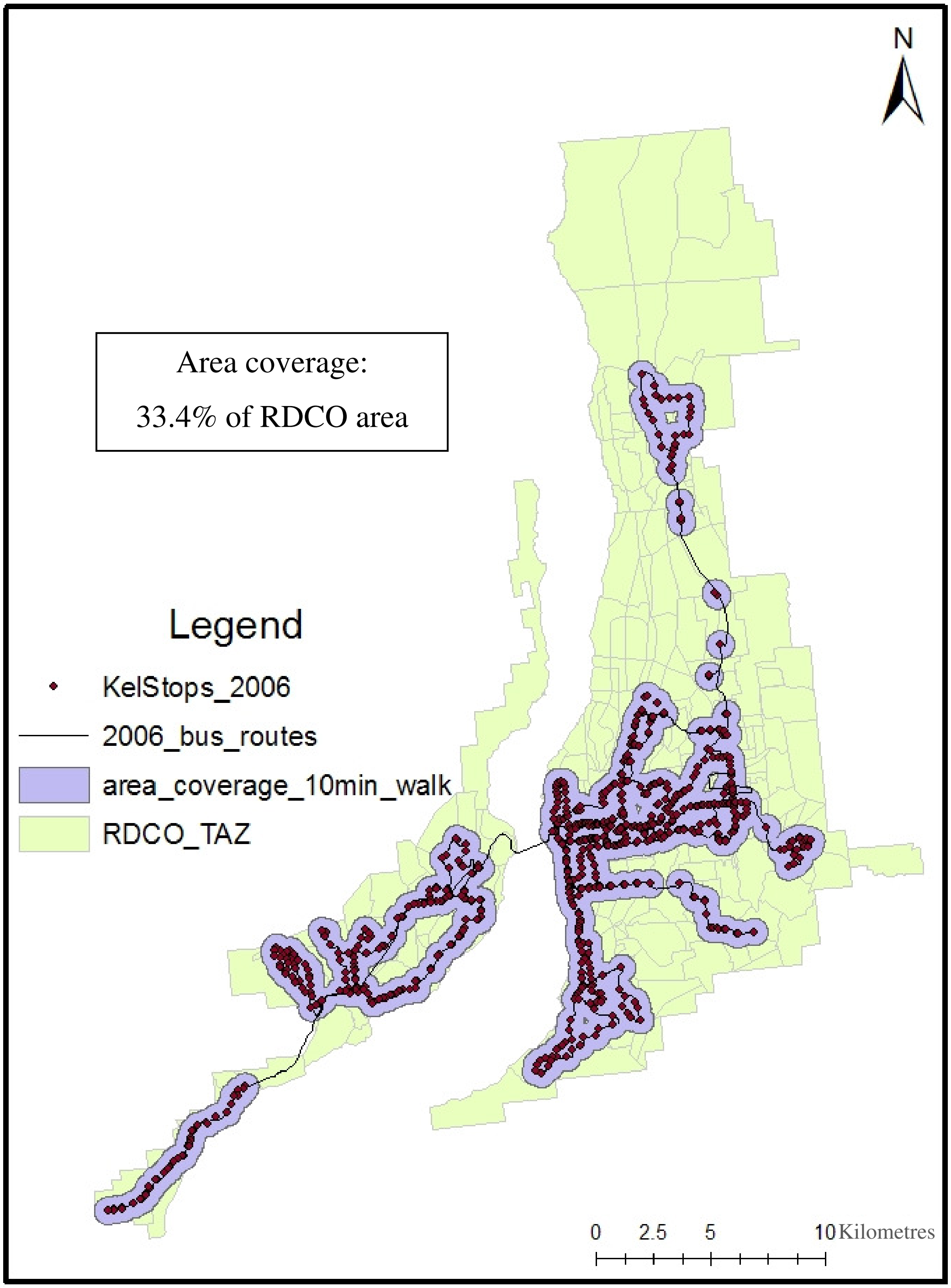


Figure I.1 Transit coverage map in 2006

- for 10 minutes walking distance @ walking speed of 1m/sec

APPENDIX J Weekday Transit Ridership in 2006

Table J.1 Weekday RDCO transit ridership in 2006

		Number of passengers			
Bus No	Bus name	AM Peak (6am-9am)	PM Peak (3pm-6pm)	Daily (24 hours)	AM period share (%)
1	Lakeshore	548	644	2056	26.65
2	North End Shuttle	29	45	153	18.95
3	Dilworth	38	24	99	38.38
7	Glenmore	286	270	993	28.80
8	University	710	992	3091	22.97
9	Shopper Shuttle	190	235	675	28.15
10	North Rutland	657	857	2979	22.05
11	South Rutland	341	499	1419	24.03
12	McCulloch	31	37	102	30.39
14	Black Mountain	15	15	47	31.91
15	Crawford	5	10	24	20.83
16	Southwest Mission	35	40	90	38.89
19	S. Rutland Night	0	25	224	0
20	Lakeview	88	114	308	28.57
21	Glenrosa	426	445	1367	31.16
22	Peachland	28	23	84	33.33
23	Lake Country	142	129	448	31.7
24	Shannon Lake	126	180	459	27.45
26	Westbank Night	0	0	78	0
27	Horizon	17	31	61	27.87
28	Smith Creek	21	20	59	35.59
99	School Special	0	237	237	0
	Total	3734	4872	15055	24.8

APPENDIX K Inter-Regional Trips from the Okanagan Valley Transportation Plan (OVTP, 1995)

Table K.1 Intercity trips between Kelowna and Vernon in 1994

From/To	Osoyoos	Penticton & Summerland	Peachland & Westbank	Kelowna	Vernon
Osoyoos	1,427	2,387	81	289	10
Penticton & Summerland	2,686	1,046	1,112	3,119	58
Peachland & Westbank	75	1,631	1,227	14,230	477
Kelowna	171	3,076	14,283	6,431	4,268
Vernon	17	92	892	3,338	1,352

Table K.2 Intercity trips between Kelowna and Vernon in 2020

From/To	Osoyoos	Penticton & Summerland	Peachland & Westbank	Kelowna	Vernon
Osoyoos	2,012	3,882	158	500	11
Penticton & Summerland	4,368	1,696	2,309	6,164	70
Peachland & Westbank	133	3,713	2,911	29,482	923
Kelowna	292	6,370	29,631	15,133	8,670
Vernon	20	108	1,762	6,364	3,363

APPENDIX L Distribution of VC for Urban Zones in 2020 Scenario#1 (do-nothing)

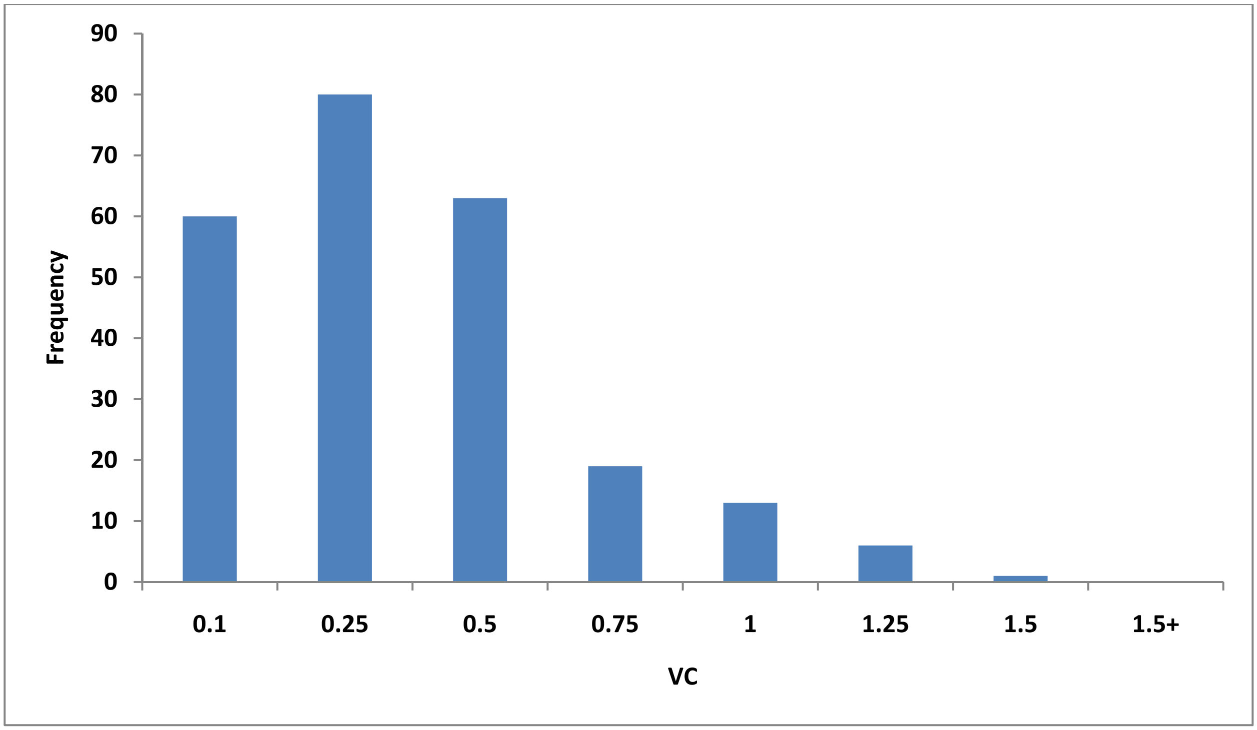


Figure L.1 Distribution of VC for urban zones in 2020 Scenario#1 (do-nothing)

	Value of VC
Average	0.285
Standard deviation	0.267
Maximum	1.261
Minimum	0.000
No. of zones	242

APPENDIX M Road Collision History of British Columbia, Canada

Table M.1 Road collision history of British Columbia, Canada (ICBC, 2007)

	No. of collisions				No. of victims		
Year	PDO	Injury	Fatal	Total	Injured	Killed	Total
1988	101277	30935	535	132747	44181	615	44776
1989	111300	39062	491	150853	47471	587	48058
1990	122300	35073	567	157940	50599	654	51223
1991	66563	32073	463	99099	47383	537	47920
1992	62331	39328	419	102078	48435	473	48908
1993	60984	32293	412	93689	46952	512	47464
1994	63362	33337	458	97157	48299	534	48833
1995	60398	32681	411	93490	47474	493	47967
1996	47783	27145	357	75285	40201	407	40608
1997	26094	21077	341	47512	31542	391	31933
1998	22117	19170	378	41665	29995	433	30388
1999	20998	20002	380	41380	30030	419	30449
2000	22255	20016	361	42632	29927	404	30331
2001	25320	20223	354	45897	29657	398	30055
2002	27161	20114	399	47674	29361	453	29854
2003	28599	20940	400	49939	30810	456	31266
2004	28811	20300	409	49520	29194	444	29638
2005	29822	20375	404	50601	28759	459	29218
2006	30003	19883	366	50252	27622	409	28031
2007	18683	18683	372	37738	25997	417	26454

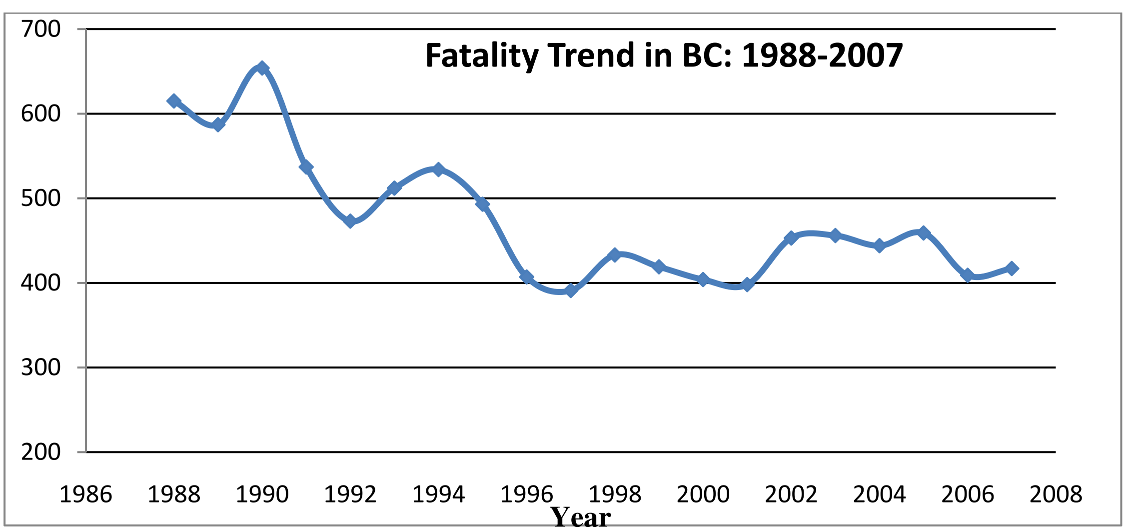


Figure M.1 Road collision trend in British Columbia, Canada (ICBC, 2007)

APPENDIX N Collision Cost Components

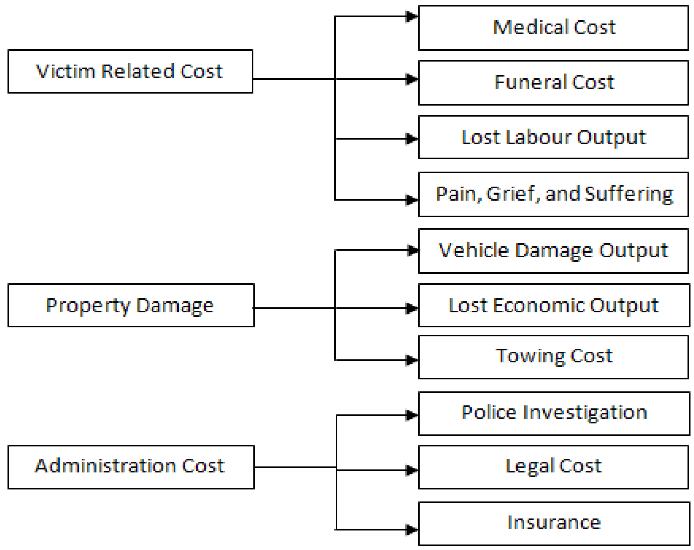


Figure N.1 Collision cost components (De Leon et al., 2005)

Table N.1 Road collision cost components (Parry, 2004)

Crash cost by severity (costs are in 2004 U.S. dollars)					
Crash Cost Components	**Fatal Injury**	**Disabling Injury**	**Evident Injury**	**Possible Injury**	**Property Damage Only**
Medical	22,095	19,471	5,175	3,485	140
Household Productivity	0	6,944	1,854	1,244	85
Lost Wages	0	25,014	6,239	4,160	155
Legal Costs	102,138	5,167	1,101	681	15
Insurance Administration	37,120	5,999	1,776	1,181	152
Property Damage	10,273	4,357	3,824	3,413	1,642
Police & Fire Services	833	175	112	90	31
Travel Delay	5,247	885	797	785	696
Employer Costs	0	1,679	665	461	67
Total, Excluding Quality of Life Costs	186,480	69,479	21,543	15,500	2,983
Quality of Life Costs (pain, grief etc.)	3,000,000	83,239	19,560	10,725	464
Total, Including Quality of Life Costs	3,186,408	152,718	41,103	26,225	3,447

* "Quality of Life Costs" represents the value of non-monetary costs such as pain, grief, and reduced enjoyment due to deaths and injuries.

APPENDIX O History and Economic Cost of RDCO Collisions

Table O.1 RDCO collision history

Year	Total Collisions	Fatal Collisions	Population	Total Collisions/ 1000 population
1996	2,067	5	141,628	14.59
1997	1,977	4	145,463	13.59
1998	2,148	3	148,762	14.44
1999	2,164	1	150,305	14.4
2000	2,357	7	152,098	15.5
2001	2,669	0	154,156	17.31
2002	3,013	1	155,692	19.35
2003	3,082	8	158,264	19.47
2004	3,483	6	159,333	21.86
2005	3,600	4	163,142	22.07
2006	3,882	6	167,417	23.19

source: BC Transit and BC Stat

Table O.2 Sample RDCO collision cost estimation for year 2006

Type of collision	No. of Collisions	Each Collision Cost (IC)	Total Cost (x1000) (IC)	Each Collision Cost (WTP)	Total Cost (x1000) (WTP)	Each Collision Cost (HCC)	Total Cost (x1000) (HCC)
PDO	2,438	4,500	10,971	6,000	14.628	5,000	12,190
Injury	1,438	44,000	63,272	100,000	143.800	35,000	50,330
Fatal	6	281,000	1,686	4,200,000	25.200	1,000,000	6,000
Total	3,882		75,929		183,628		68,520

- IC= Insurance Claims method
- WTP= Willingness to Pay method
- HCC= Human Capital Cost method