Safety at Quebec's Roundabouts: Investigating Injuries and Accident Occurrence

Prepared by
Shaun Burns

Department of Civil Engineering and Applied Mechanics
McGill University, Montreal, Quebec

December 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Engineering

Contribution of co-Authors

A number of researchers have contributed to the work within this thesis. The second chapter is based on a paper written in collaboration with my supervisor, Dr. Luis Miranda-Moreno, Joshua Stipancic, a graduate student at McGill University, Dr. Nicolas Saunier, an associate professor at École Polytechnique de Montréal, and Dr. Karim Ismail, my co-supervisor from Carleton University. The paper is titled *An Accessible and Practical Geocoding Method for Traffic Collision Record Mapping: A Quebec Case Study* and will be published in the 2014 Transportation Research Record (Journal of the Transportation Research Board (TRB)). The third chapter is based on a paper submitted for presentation at the 2013 TRB Annual Meeting and was again written in collaboration with Dr. Luis Miranda-Moreno, Dr. Nicolas Saunier and Dr. Karim Ismail. The paper is titled *Crash Severity Analysis at Roundabouts: A case study in Quebec, Canada* and was presented at the Transportation Research Board's 92nd Annual Meeting. The final chapter will be considered for submission to an upcoming conference. In all cases I was the primary author of the papers.

Acknowledgments

I would like to thank the Québec Ministry of Transportation (MTQ) as well as Quebec's Automotive Insurance Board (SAAQ) for providing the accident data used in this paper. I would also like to thank the various engineers and municipal representatives who helped identify roundabout locations across the province.

I am very grateful to the Québec Fund for Research on Nature and Technology (FRQNT), the Québec Ministry of Transportation (MTQ), the Québec Fund for Health Research (FRQS), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Réseau de Recherche en Sécurité Routière (RRSR), the Centre interuniversitaire de recherche sur les reseaux d'entreprise, la logistique et le transport (CIRRELT) and McGill University for providing various levels of funding to both me personally and to the project as a whole. The financial support allowed for the purchase of necessary equipment and to allow me to focus solely on my research and not have to worry about living expenses.

I also owe thanks to a number of colleagues. From École Polytechnique de Montréal, Jean-Simon Bourdeau was of great help for having set up the inventory of roundabouts in Quebec, and Paul St-Aubin for his invaluable help during data collection and his computing knowledge during analysis. From McGill University, I owe thanks to Usamah Khan for helping during data collection, Taras Romancyshyn for his help in selecting control sites and Joshua Stipancic for being a great friend and allowing me to bounce ideas off of him.

I would also like to thank my co-supervisors, Doctor Luis Miranda-Moreno from McGill University and Doctor Karim Ismail from Carleton University. Finally, I would like to thank Doctor Nicolas Saunier from École Polytechnique de Montréal for his contributions. They supported and encouraged me throughout the various stages of my research. Without their knowledge and guidance I don't know where I would be today.

Abstract

Numerous studies both in the United States and Europe have investigated the roundabout as a means to improve the safety of intersections. The focus of these studies has been the investigation of the effectiveness of roundabouts as well as the contributing factors associated to crash risk looking at both the number and the severity of crashes. Despite the number of studies in this area, very little research exists in the Canadian or Quebec provincial context in which local factors could play in favour or against roundabouts. More importantly, the popularity of roundabouts in the Province of Quebec has increased significantly in the last years. Safety concerns associated with this type of intersection remain, with no clear answer.

That being said, an important amount of crash record data is not geocoded, which is essential for crash occurrence and severity analysis studies. This thesis attempts to investigate roundabout safety in the Province of Quebec, Canada by combining different sources of data and using various statistical approaches.

In order to accomplish this goal, the thesis is divided into three main objectives; 1) to develop a free and accessible method of geocoding crash record data, 2) to investigate the impact that roundabouts have on the safety of road users using historical crash record data and 3) to identify contributing factors associated to roundabout crash frequency and injury severity.

Among other results, a methodology for mapping crash data is proposed using a custom algorithm that calls on an online service such as the Google Maps application programming interface (API). It is found that a match rate of 78% can be achieved, which is comparable to commercially available geocoding options. With proper user revision, the results are sufficient for practical applications such as intersection safety analysis.

Through the application of an ordered logit model in order to identify factors that contribute to injury severity at roundabouts the research in the third chapter identified that factors such as a larger number of involved vehicles, accidents occurring within the intersection, vehicle rollovers, the involvement of buses, accidents occurring in the dark on unlit roads and snow conditions led to increased injury severity within roundabouts. Similarly, factors associated to accidents involving only cars, animal strikes and snow-covered roadways were found to be associated with less severe injuries.

Finally, from the results of Chapter 4, which deals with the safety effectiveness of roundabout, a number of observations were made. While found to be statistically insignificant, the observations suggest that smaller roundabouts, roundabouts in which only one type of user is present (i.e. local or highway driver behavior, not both) and roundabouts with a consistent speed on all of their approaches are observed to demonstrate a reduced crash frequency, as estimated through the before-after method with comparison group. One of the main weaknesses of this method is its inability to account for regression-to-the-mean effects; an idea that will be discussed throughout the work.

Résumé

De nombreuses études, tant aux États-Unis qu'en Europe, ont étudié le carrefour giratoire en tant que moyen d'améliorer la sécurité des intersections. L'objectif principal de ces études a été l'enquête de l'efficacité des carrefours giratoires, ainsi que les facteurs contributifs associés au risque de collision en regardant à la fois le nombre et la gravité des accidents. Malgré le nombre d'études dans ce domaine, très peu de recherches existent dans le contexte Canadien ou provinciale Québécois où les facteurs locaux pourraient jouer en faveur ou contre les carrefours giratoires. Plus important encore, la popularité des carrefours giratoires dans la Province de Québec a considérablement augmenté dans les dernières années. Les préoccupations liées aux problèmes de sécurité dans les carrefours giratoires restent, sans réponse claire.

Cela étant dit, une quantité importante de données d'accident ne sont pas géocodées, ce qui est essentiel pour les études d'occurrence d'accidents et les analyses de la gravité des blessures. Cette thèse tente d'enquêter la sécurité des carrefours giratoires dans la province de Québec, Canada, en combinant différentes sources de données et en utilisant diverses approches statistiques.

Pour atteindre cet objectif, la thèse est divisée en trois objectifs principaux; 1) développer une méthode gratuite et accessible de géocodage des données d'accident, 2) étudier l'impact que les carrefours giratoires ont sur la sécurité des usagers de la route à l'aide de données historiques d'accident, et 3) identifier les facteurs associée à la fréquence des accidents dans les carrefours giratoires et la gravité des blessures.

Parmi d'autres résultats, une méthodologie pour cartographier les données d'accident est proposé en utilisant un algorithme personnalisé qui fait appel à un service internet tel que l'interface de programmation d'application de Google Maps (API). On constate qu'un taux de correspondance de 78% peut être atteint, ce qui est comparable à des options commerciales. Avec la révision appropriée de l'utilisateur, les résultats sont suffisants pour des applications pratiques telles que l'analyse de la sécurité aux intersections.

Grâce à l'application d'un modèle logit ordonné afin d'identifier les facteurs qui contribuent à la gravité des blessures aux carrefours giratoires, la recherche dans le troisième chapitre a identifié que des facteurs comme un plus grand nombre de véhicules impliqués, les accidents survenus dans l'intersection, les renversements de véhicules, la participation des bus,

les accidents survenus dans le noir sur les routes non éclairées et les conditions de neige ont mené à une augmentation de la gravité des blessures dans les carrefours giratoires. De même, les facteurs associés à des accidents impliquant uniquement les voitures, les accidents avec des animaux et des routes enneigées ont été trouvés être associés à des blessures moins graves.

Enfin, à partir des résultats du chapitre 4, qui traite de l'efficacité de la sécurité aux carrefours giratoires, un nombre d'observations ont été faites. Bien que jugée statistiquement non significatif, les observations suggèrent que les petits carrefours giratoires, les carrefours giratoires dans lesquels un seul type d'usagé est présent (c'est-à-dire le comportement local ou autoroutier; mais pas les deux en même temps), et les carrefours giratoires avec une vitesse uniforme sur l'ensemble de ses approches tendent ont démontré une fréquence de collision réduite, telle qu'estimée par la méthodologie avant-après avec groupe de comparaison. L'une des principales faiblesses de cette méthode est qu'il ne prend pas en compte les effets de régression vers la moyenne; une idée qui sera discutée au long de ces travaux.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 CONTEXT AND MOTIVATION

In large parts due to the perceived safety and operational benefits associated with roundabouts from international studies, North American interest in their implementation has grown considerably in recent years (*1*) (*2*) (*3*). Although several guides have been created to help design roundabouts for the North American road environment such as the National Cooperative Highway Research Program's (NCHRP) *Report 572 - Roundabouts in the United States* and *Report 672 – Roundabouts: An Informational Guide* (*1*) (*2*), the Federal Highway Administration's *Roundabout Technical Summary* (*4*), and more locally the Quebec Ministry of Transportation's *Le carrefour giratoire: un mode de gestion différent* (*5*), no standard has yet to be uniformly adopted. As municipalities continue to make the decision to invest in new and converted roundabout intersections, it is increasingly important to study the actual safety and operational efficiency that is achieved due to the variations in geometry and signalization that exists across the road network.

Following the trend observed in North America, a growing number of municipalities in the Province of Quebec, Canada have shown interest in investing in roundabouts. Starting with the construction of the first roundabout in Canada in Ville Saint-Laurent, Québec, in 1998, it is estimated that the Province of Quebec now boasts over 150 roundabouts, of which 41 are found on the Quebec Ministry of Transportation (MTQ) road network according to recent communications with the ministry. With such a high total, Quebec is considered to have the highest number of roundabouts in Canada (*6*) (*7*). Because of this, the Quebec road network is well suited for an investigation into roundabout safety with regards to local design practices.

An extensive literature base exists in which European and Australian roundabouts have been widely studied and has generally come to agreeable results (*8*) (*9*) (*10*). Through the use of before-after safety studies it is shown that injury-accidents decrease in all cases after the construction of a roundabout. As mentioned, similar results are observed in a majority of the studies, although slight increases in accident rates are presented when targeting property-damage only and vulnerable-user crash categories. The exact causes of these observations are general not presented.

In this literature, it has been recognized that roundabout intersections are a valuable tool in the reduction of vehicle conflicts, and excessive speeding. However, environmental conditions, localized driver behavior and individual design elements can influence the level of safety that is achieved. Similarly, a number of recent American studies have been published, but to the author's knowledge very little has been studied in Canada and in particular the province of Quebec.

Consequently, there are concerns with respect to a roundabout's performance in terms of safety due to the particular environmental and driver behavioral conditions that exist within the province (*7*). Environment, weather and driving behavior differences across countries and societies lend support to the notion that local studies for safety are always beneficial for adapting designs and operating characteristics to the local conditions. A better understanding of roundabout safety as it pertains to the Quebec road network can help justify their continued implementation as well as help identify areas of strength and weakness in the local design culture.

Accordingly, this thesis focuses on the safety analysis of roundabouts in the province of Quebec. For this purpose, an approach based on historical crash records is implemented. In other words, this research studies roundabout safety performance and identifies contributing crash-risk factors and design elements through the analysis of historical accident data. Prior to the safety analysis, a methodology for the geocoding of each accident record is presented in order to enable the spatial location of each record within the road network. The safety effects of a group of features (i.e. of roundabouts) are then studied through the use of observational studies in the form of crash injury severity and frequency analyses (*11*).

It is important to mention that a surrogate safety approach can be implemented as a complementary component to investigate the safety of roundabouts, although this approach is out of the scope of this research. However, readers who wish to further explore the topic of surrogate safety for roundabouts in Quebec are invited to consult the work of St-Aubin et al. (*12*) (*13*).

For this research, ten years of crash records were made available through the provincial crash database of the *Ministère des Transports du Québec (MTQ)* and the *Société des Assurances Automobile du Québec (SAAQ)*. Geometry inventories and land use data was also required for the realization of this work.

The following section provides some general background, objectives and additional details on the data that was used.

1.2 BACKGROUND

1.2.1 Roundabout Design Theory

Roundabouts are a form of self-regulating traffic control achieved primarily through the layout of the intersection in terms of geometry and lane configuration, as well as through the design of driver sight lines using lighting and landscaping details (*14*). The primary goal of a roundabout design revolves around the management of vehicle speeds and the reduction of conflict points (*2*) (*5*) (*12*).

With the introduction of the yield-at-entry rule in the 1960s, it is important to distinguish between what is known as a modern roundabout and older traffic circles (*2*) (*5*). Traffic circles are generally larger, offer tangential approaches to the circulating roadway, allow for parking and can include signalization or otherwise not satisfy the yield-at-entry rule. Modern roundabouts, on the other hand, require any entering vehicle to yield to vehicles already on the circulating roadway, are unsignalized and provide perpendicular approaches to the circulating roadway. In both cases, North American traffic follows a counter-clockwise direction around the circulating roadway. FIGURE 1 illustrates the basic differences between the two concepts.

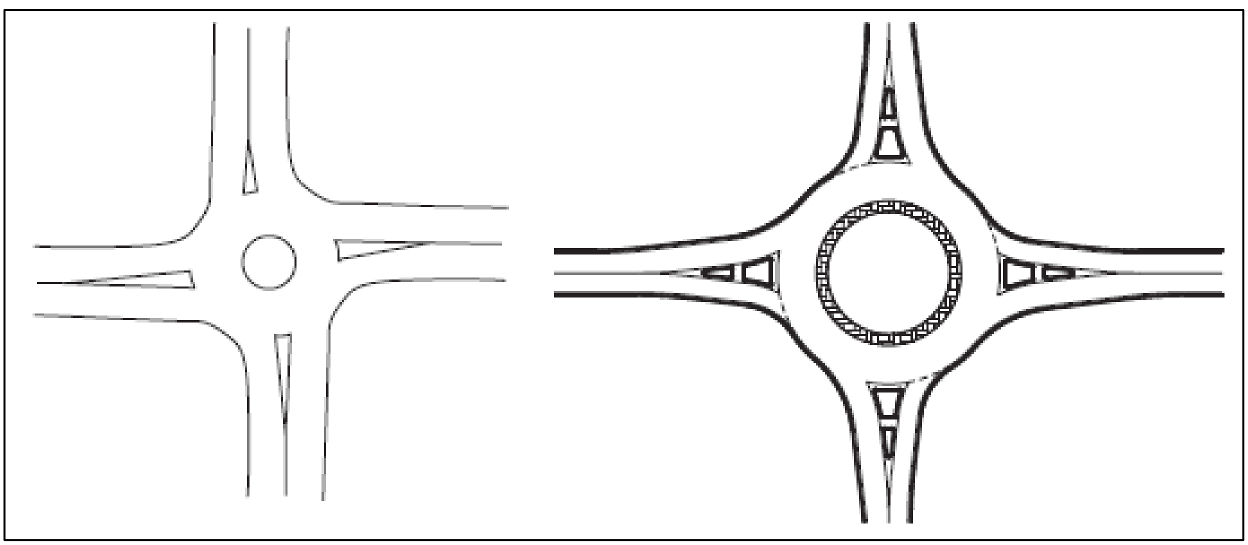

FIGURE 1: Comparison of a typical traffic circle (left) and a modern roundabout (right) (*5*).

Depending on the design guide that is consulted, additional requirements are quoted as being required for a circular intersection to fit the true definition of a modern roundabout.

Looking at the Quebec roundabout design guide and the prevailing American guide *Report 672 Roundabouts: An Informational Guide*, and presented in no particular order, these include (*5*):

- The existence of three or more approaches.
- Construction of splitter islands between entry and exit lanes.
- Trajectory deflection when entering the roundabout.
- The presence of a raised central island.
- Accessible pedestrian crossings.
- Construction of a truck apron (where a large design vehicle is expected).

With this in mind, the Quebec road network contains a number of circular intersections that do not fit the definition of a true modern roundabout. Looking at FIGURE 2 for example, two circular intersections that are not considered roundabouts can be seen. The intersection in Point-Claire, Quebec is a modified T-intersection in a residential sector that retains all-way stop signs on its approaches. In the case of the Dorval, Quebec intersection, this is an example of an older traffic circle in which the approaches are tangential to the circulating roadway and all approaches are signal-controlled.

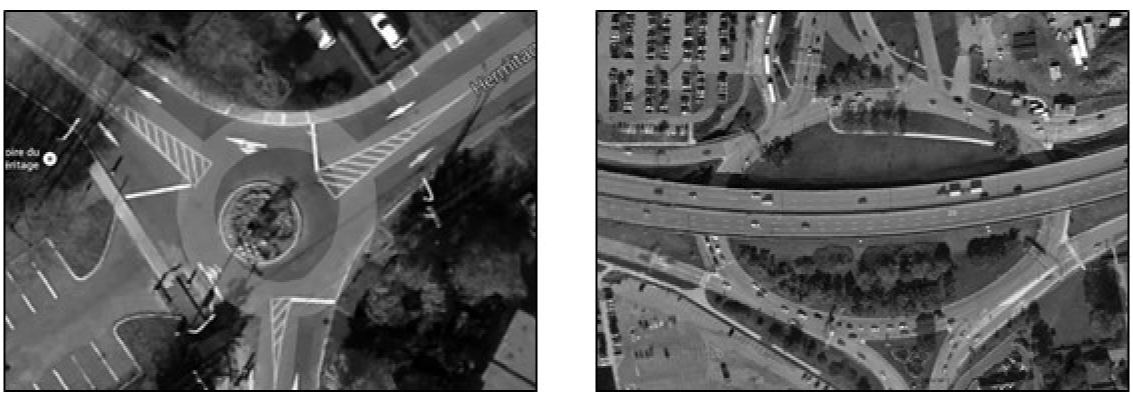

FIGURE 2: Circular intersection examples; Pointe-Claire, QC (left) and Dorval, QC (right) (*15*).

Other examples of circular intersections are considered to be roundabouts, although in the strictest sense they do not fulfill all of the criteria of a modern roundabout. Examples of this can be seen in FIGURE 3. In the case of the Bromont, Quebec roundabout example, the intersection only has two approaches. As mentioned above, the Quebec roundabout design guide defines a roundabout as having three or more approaches (*5*). In the case of the Trois-Rivières, Quebec example, the intersection presented is an early example of a traffic circle that was converted to operate as a modern roundabout. The intersection lacks many of the geometric characteristics of a traditional modern roundabout; namely, the existence of perpendicular approaches and an overall smaller footprint.

FIGURE 3: Roundabout-type intersections; Bromont, QC (left) and Trois-Rivières, QC (right) (*15*).

As will be discussed in further detail in the coming sections, a majority of the roundabouts being built across the territory of Quebec are of the single-lane variety with four approaches and integrated truck aprons to accommodate larger design vehicles such as delivery trucks, buses and emergency vehicles.

1.2.2 Roundabout Safety

An important trait of a roundabout intersection is the potential for improved safety. Although the benefit is typically more prominent for fatal and injury crashes, much of the international literature has shown that roundabouts are safer than stop and signal-controlled intersections. The reader is invited to consult (*2*) (*8*) (*16*) (*17*) for a sampling of research that reinforces this idea.

It is important to note that although these results are well documented in the international literature, it is difficult to infer the expected safety of roundabouts on the North American road network due to the relative infancy of their application as an intersection safety treatment and a population that has very little exposure on how to properly navigate them.

Persaud et al. (*17*) acknowledge the small sample sizes and limited results that have been extracted in American roundabout safety studies in the past, although growing interest in their use is quickly changing this weakness. As North American studies continue to grow in both quantity and quality, it will be possible to identify the true safety that is achieved at roundabouts considering differences in driver behaviour, environmental factors, and local design practices.

As the topic of roundabout safety benefits continues to be debated, it is increasingly important for researchers to ensure that their models account for weaknesses identified in past studies. These typically include an inability to account for regression-to-the-mean effects, loss of data through excessive aggregation, weak or incomplete datasets, and inappropriate resolution of recorded crash records, to name a few (*17*). Although some of these weaknesses can be controlled during the analysis, a certain responsibility must be taken by local transportation or government officials to put in place proper data collection procedures.

Before going any further with the topic of roundabout safety, a basic explanation of the guiding principles of a roundabout's safety performance is presented. As previously mentioned, the roundabout's safety record is attributed primarily to a smaller number of conflict points, as well as reduced vehicle speeds. These effects are based on the roundabout theory presented above and the geometry principles discussed in the following section (*5*).

Conflict points exist throughout the road network as a result of vehicle movements which have the possibility of occupying the same temporal and spatial planes. It is widely accepted that crash frequency at an intersection is related to both the number of these conflict points, as well as the relative magnitude of the conflicting vehicular flows.

As can be seen in FIGURE 4, a traditional 4-way intersection presents 32 vehicle-vehicle conflict points spread across its area. The roundabout, on the other hand, presents only 8 vehicle-vehicle conflict points, grouped in 4 distinct conflict zones. It is also important to note that the types of conflicts are also fundamentally different for roundabouts, which offers the potential for reduced crash severity. Conflict points relating to crossing paths and left-turn movements simply do not exist as a consequence of their design (*18*).

A similar principle is seen for vehicle-pedestrian conflicts, as presented in FIGURE 5 below. Not only do pedestrians have to be attentive to vehicles coming from multiple directions during green light phases, there is also the possibility of conflict points arising from red light running and right-turn-on-red movements. For roundabouts on the other hand, pedestrians need only cross one direction of traffic at a time at each approach.

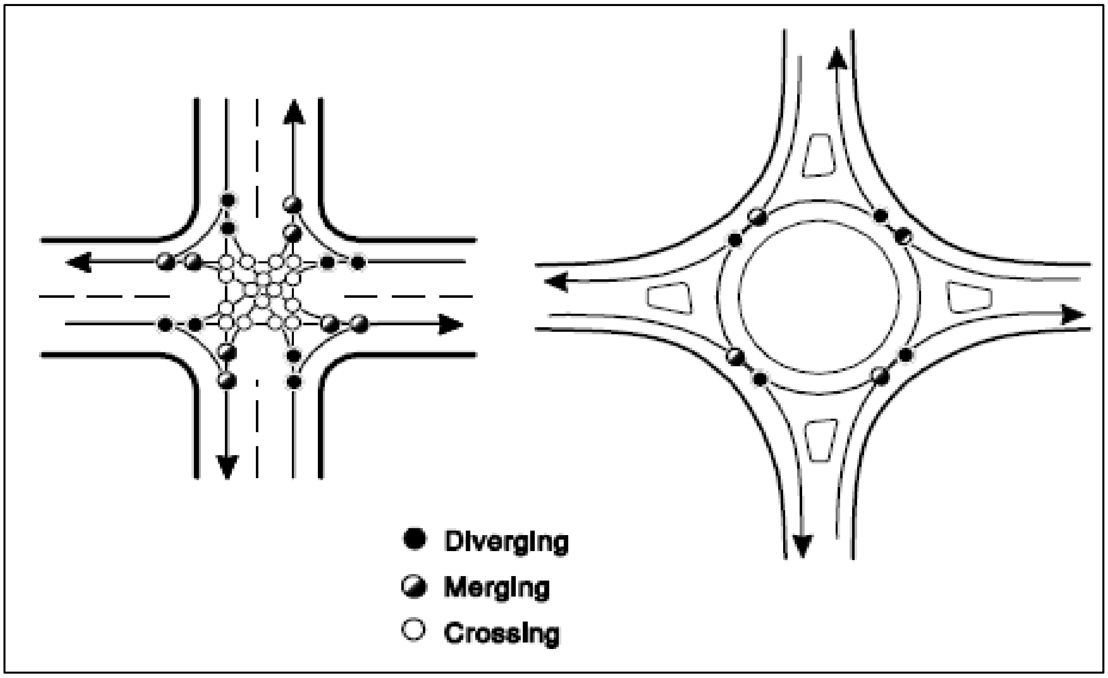

FIGURE 4: Vehicle conflict point for a traditional four-way intersection (left) and a roundabout (right) (*2*).

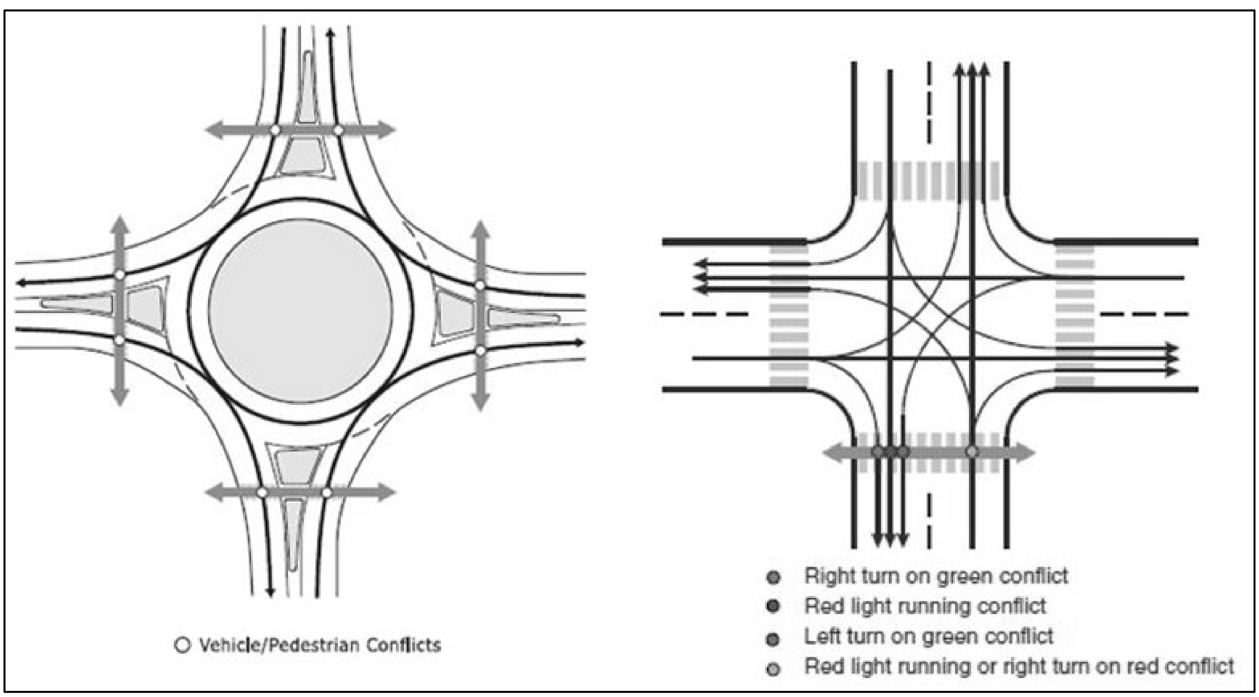

FIGURE 5: Comparison of pedestrian conflicts at roundabouts (left) and four-way intersections (right) (*2*).

Although these principles hold for single lane roundabouts, the presence of additional entry lanes and the need to provide wider circulating and exit roadways, multi-lane roundabouts present additional conflict points. These additional conflict points generally arise due to drivers failing to stay within their lane as well as improper lane use when exiting, as illustrated in FIGURE 6.

The additional conflict points present at multilane roundabouts are typically of the low-speed, side-swipe type and result in low severity crashes. As a consequence, although the number of conflicts increases for multilane roundabouts, the overall safety is still typically higher than alternative intersection designs due to the reduced conflict severity. A general rule in roundabout design is to use the smallest number of entry, circulating and exit lanes possible after taking capacity requirements into consideration, with the goal of maximizing the safety benefits of the design (*2*).

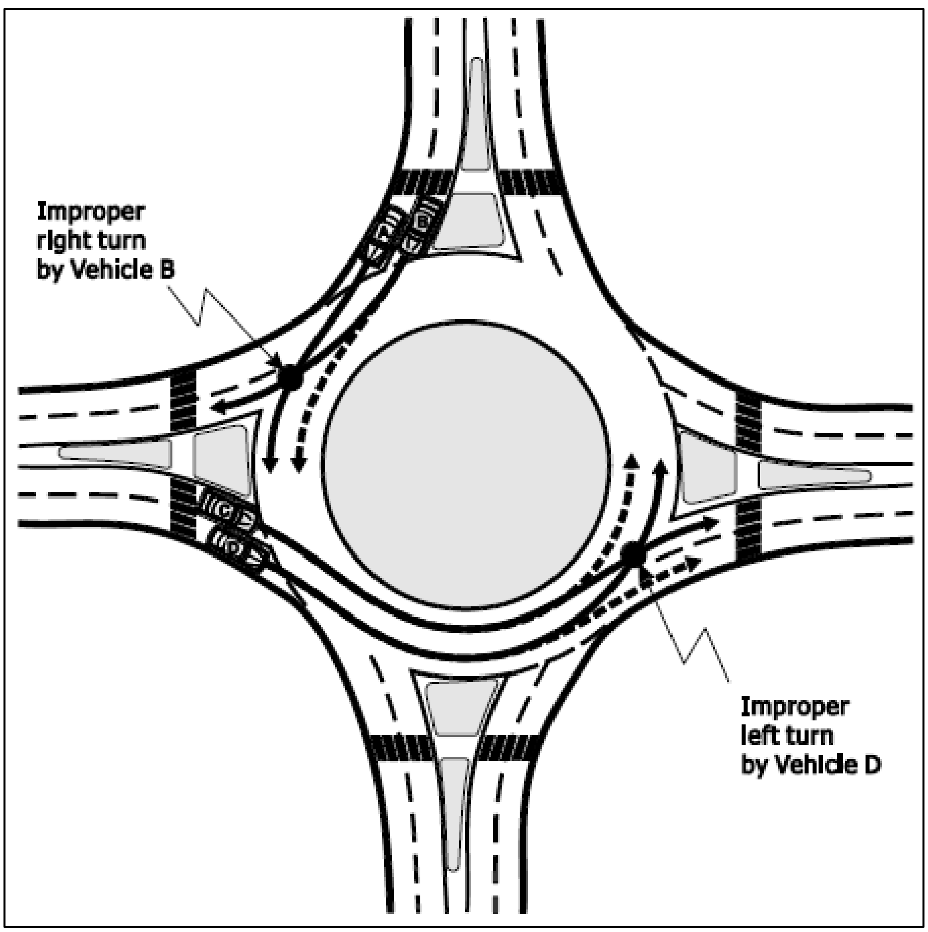

FIGURE 6: Lane conflicts at multi-lane roundabouts (*2*).

As with a majority of road infrastructure, the overall safety of a roundabout facility also varies according to the dimensions and characteristics of the intersection. One of the main decisions for designers to make is the size of the roundabout in question. A number of studies have found that greater safety benefits are achieved in smaller, single lane roundabouts (*5*) (*19*). This is quoted as resulting from the simplified operation of a roundabout with a single circulating lane where there are no lane change decisions to make (*2*).

With this in mind, designers must consider the type of users that can reasonably be expected to use the facility. Designs for areas with larger design vehicles (such as a large proportion of emergency vehicles, city buses or heavy industry vehicles) and a high pedestrian and cyclist volume will be considerably different from what can be expected at a rural roundabout with little pedestrian traffic.

In any roundabout design, an iterative approach must be taken in order to find a balance between available capacity and the expected safety of the facility (*2*). Varying implementation conditions will call for different features to be included in the roundabout design such as: addition or removal of lanes on individual approaches, use of slip lanes to alleviate increased right turn volumes, flare at entry to accommodate high entering volumes, etc. (*5*).

The geometric design of a roundabout also influences the reduction of vehicle speeds throughout the intersection, and is often quoted as being one of the most critical aspects of a safe roundabout design (*2*). As previous research has shown, an important link exists between practiced vehicle speeds and both crash occurrence and severity (*18*). In their research, Chen et al. (*20*) reinforced this idea by using Bayesian Poisson-Gamma and Zero-Inflated Poisson models to attempt to predict accident occurrence at roundabouts and found average approach speed to be the most significant predictor.

Fundamentally, a combination of design elements will influence the *deflection* (or the amount of deviation with respect to the straight path of a given vehicle measured from the approach), leading to the desired speed reduction. Theoretical speed reduction profiles for different entry speeds are presented in FIGURE 7. As can be seen, the average vehicle speed across a roundabout ranges from 25-40 km/hr, which allows drivers more time to react to potential conflicts compared to traditional intersection designs (10). The reduced speed differential between all road users also helps maintain low crash severity when crashes do occur, leading to fatalities and severe injuries being uncommon at roundabout intersections.

More detailed designs traits can also influence a driver's speed at a roundabout. A properly landscaped central island, for example, has been quoted as beneficial in this respect as it acts to break a driver's line of sight, leading to a reduction in speed. In rural environments, a series of successive, reversed curves is used on approaches with a high posted speed limit to attempt to begin slowing vehicles down to an acceptable entry speed for the roundabout. The general principle is illustrated in FIGURE 8.

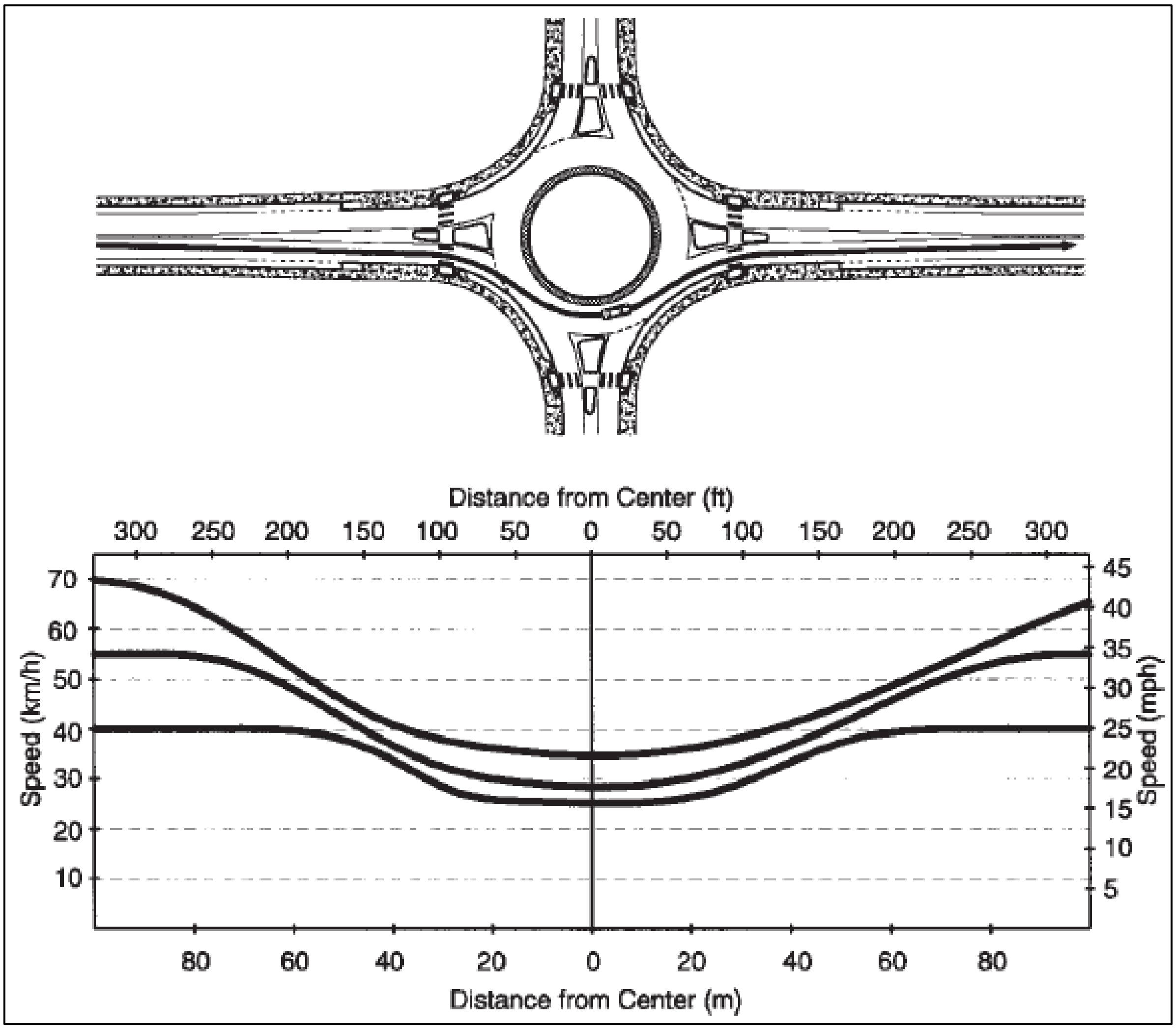

FIGURE 7: Theoretical speed reduction profiles for a given vehicle across a roundabout (3).

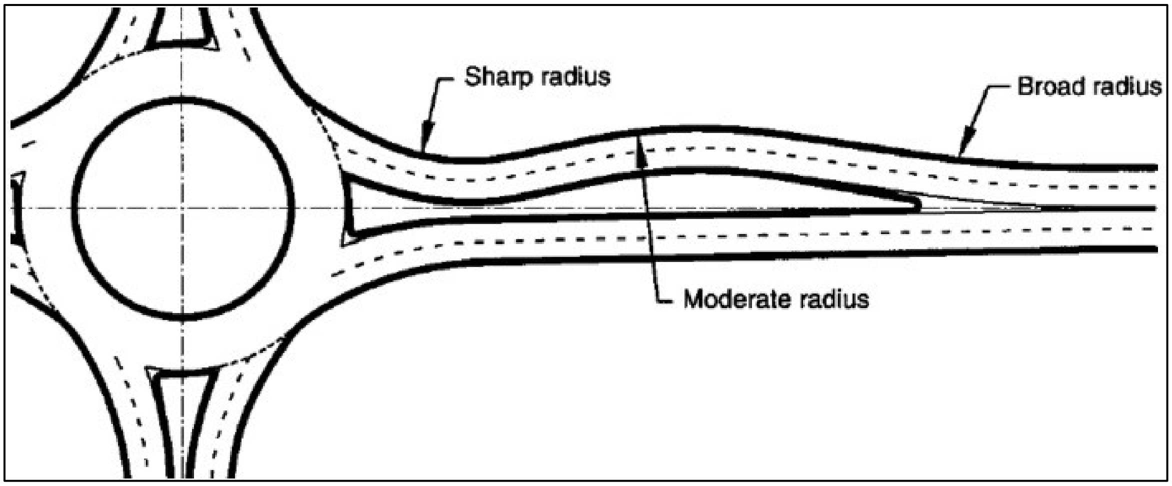

FIGURE 8: Successive reverse curves at high speed approaches (*14*).

It is interesting to note that although roundabouts are designed to reduce physical vehicular speeds, the reduced stopping and delays common to other forms of intersections can lead to improvements in the overall speed of travel.

1.3 OBJECTIVES AND THESIS STRUCTURE

This thesis aims to examine the safety of roundabout intersections in the province of Quebec using a historical crash-analysis approach. Through this research, an attempt was made to answer a number of general research questions, including: What are the data needs and limitations? What are the factors that influence crash severity and crash occurrence? Are roundabouts safer than traditional intersections? Which characteristics are the most appropriate? To answer these questions, the identification of the available data was the initial step.

As is often the case, important data limitations were quickly identified, such as the lack of geocoded crash data. In order to respond to the data limitations and the stated research questions, the objectives of this thesis are:

- To develop a free and accessible method of geocoding crash report data in preparation for safety analysis;
- To investigate the impact that roundabout intersections have on the safety of road users through the use of statistical methods in a before-after framework;
- To study and identify the contributing factors associated to roundabout crash frequency and injury severity;

- To recommend design practices that will help provide a greater level of safety at Quebec roundabouts.

For each of these objectives, an important amount of work was needed to integrate and process the various sources of data. The objectives are addressed throughout each of this thesis' five chapters. The first chapter provides an introduction to the topic of roundabout safety as well as the motivation behind the research project, the data that was used as well as the objectives that were investigated.

The second chapter presents the methodology for the geocoding of textual accident locations that was required in order to prepare the data for spatial analysis and allow for subsequent safety analysis. The development of a custom algorithm is presented, as well as the main results and a discussion of the expected level of accuracy that can be achieved.

The third chapter presents an injury severity analysis that was performed at the beginning of the research, in order to identify factors that influence the outcome of accidents occurring at roundabouts. Using the geocoded accident data, the fourth chapter presents a crash frequency analysis using a before-after framework in order to quantify the safety effect of roundabout construction on the Quebec road network.

The final chapter highlights the primary conclusions of the research and outlines directions for future work.

1.4 DATA

For this research, different types of data were required such as an inventory of existing roundabouts (including information on location, year of construction, operational jurisdiction and geometry), historical crash records and geographic data (including all files required for working with ArcGIS geographic information software).

This section provides the details of each data source.

1.4.1 Roundabout Inventory

According to an inventory compiled in 2011 by a research team at École Polytechnique de Montréal, there are approximately 147 roundabouts that have either been built, or are planned to be built in the Province of Quebec. Note that this inventory was assembled under the parent research project (Research on Roundabout Safety in Québec funded in Québec, Canada, within the Road Safety Research Program by FRQNT-MTQ-FRQS for the period 2011-2014), from which this thesis was created. Readers can consult (*21*) for a map and a full copy of the database.

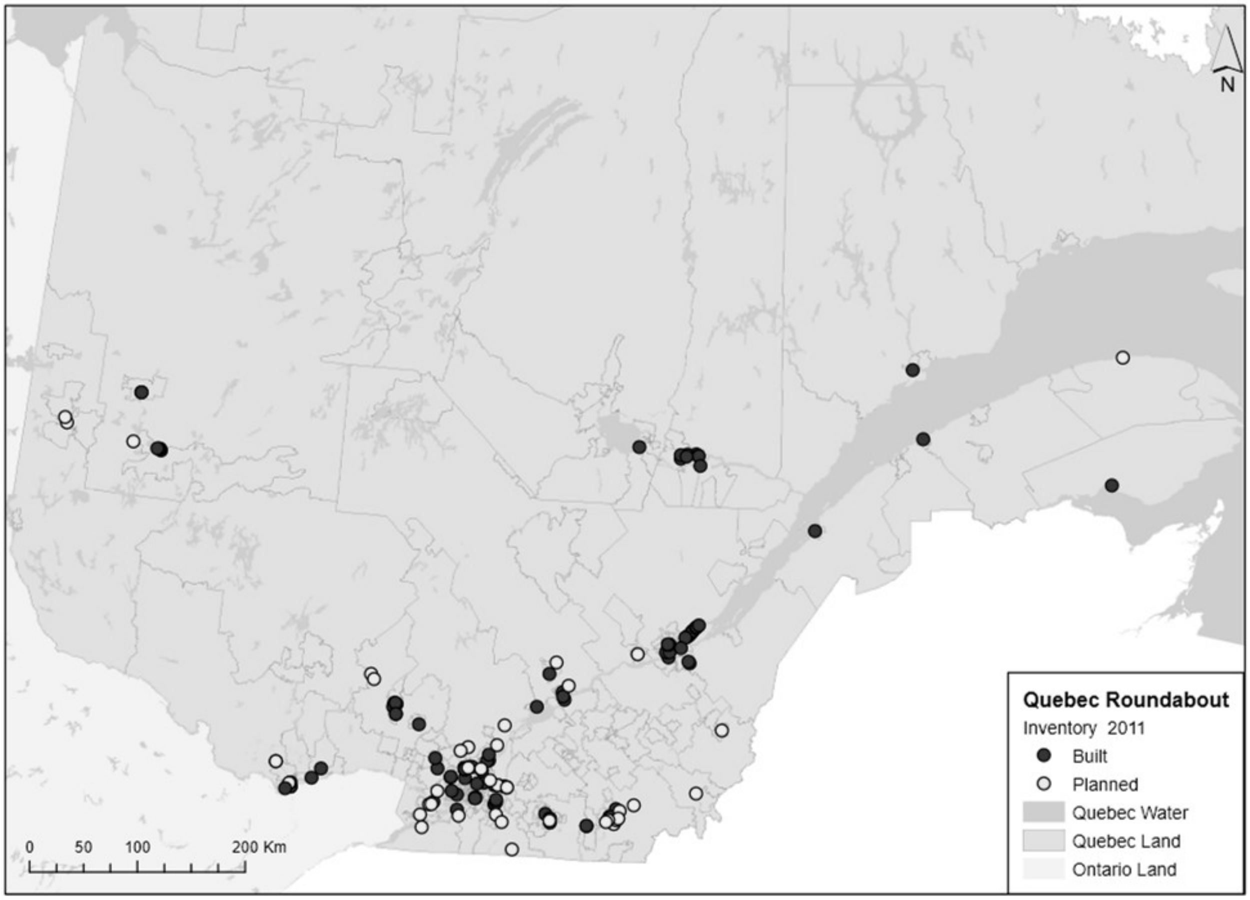

FIGURE 9: Inventory of built and planned roundabouts in Quebec (according to 2011 data) (22).

Of the 147 known roundabouts, 74 are under the jurisdiction of the provincial ministry of transportation (MTQ), with the rest being distributed across the municipal road network. FIGURE 9 presents a map illustrating the known roundabouts in Quebec. Furthermore, TABLE 1 presents a summary of the location of the built roundabouts with respect to the different jurisdictions of Quebec's Ministry of Transportation.

TABLE 1: Roundabout location summary.

Ministry of Transportation Region	Roundabout Frequency
Montérégie (Est-de-la-)	12
Montréal (Île-de-)	8
Laurentides-Lanaudière	9
Outaouais	6
Bas-Saint-Laurent-Gaspésie-Îles-de-la-Madeleine	2
Chaudière-Appalaches	4
Mauricie-Centre-du-Québec	5
Abitibi-Témiscamingue	6
Estrie	5
Côte-Nord	1
Saguenay-Lac-Saint-Jean-Chibougameau	8
Montérégie (Ouest-de-la-)	2
Capitale-Nationale	16
Laval-Mille-Îles	12
(Empty)	11
Total	**107**

A distribution of the construction dates for the 50 roundabout sites with known construction dates is presented in FIGURE 10. As can be seen, the construction of roundabouts in the province started with a small interest in 1998, with little growth over the next few years. As of 2003, however, interest in roundabout construction sees a considerable growth. Although variability exists over the years, interest in the construction of roundabouts clearly remains. According to the available database created in 2011, a further 20 roundabouts were in the planning and design stage for construction between the years 2012 and 2016.

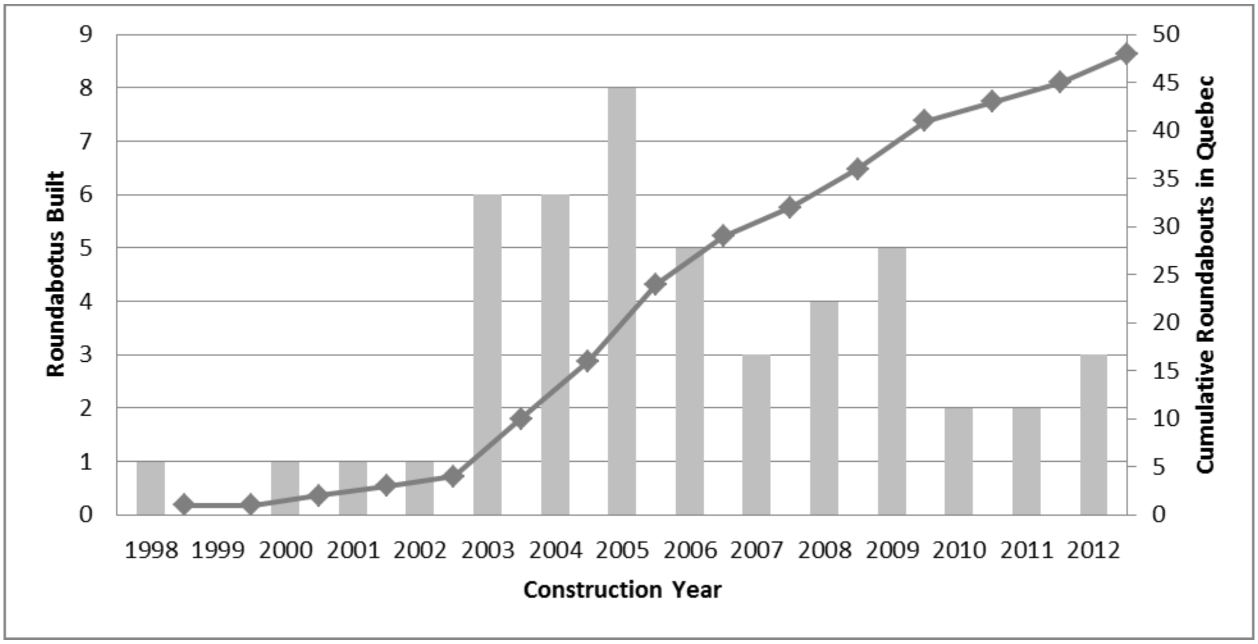

FIGURE 10: Distribution of Roundabout Construction Dates for 50 sites with known construction dates.

Using information gathered from design plans, available imagery and a select sampling of site visits for existing roundabouts, the assembled roundabout inventory contains fields relating to geometry, location and operational details. An overview of some of the available data fields for the 107 built roundabouts is presented in TABLE 2 below. As can be seen, considerable variability exists with respect to how and where roundabouts are built across the province. That being said, the most common roundabout characteristics observed throughout the dataset include four-legged roundabouts in residential neighborhoods with posted speed limits of 50 km\hr and 25 km\hr for the approach and circulating roadways, respectively. Furthermore, single-lane roundabouts with truck aprons around the central island to accommodate larger design vehicles are well represented in the dataset.

TABLE 2: Summary roundabout characteristics

Environment	Roundabout Frequency
Urban – Residential	36
Urban – Mixed	10
Urban – Commercial	7
Rural	25
Urban – Industrial	4
Suburban	9
(Empty)	16
Total	**96**

(A)

Number of Approaches	Roundabout Frequency
1	3
2	2
3	26
4	59
5	4
6	1
(Empty)	12
Total	**107**

(B)

Highest Posted Approach Speed	Roundabout Frequency
40	1
50	63
70	3
80	1
90	5
(Empty)	34
Total	**107**

(C)

Posted Speed (Circulating Roadway)	Roundabout Frequency
15	5
25	30
30	1
35	9
(Empty)	62
Total	**107**

(D)

Number of Circulating Lanes	Roundabout Frequency
1	75
2	18
Variable	2
(Empty)	12
Total	**107**

(E)

Truck Apron	Roundabout Frequency
Not Present	9
Present	77
(Empty)	21
Total	**107**

(F)

1.4.2 Geographic and Road Network Data

Due to the use of geospatial software such as ESRI's ArcGIS, a number of reference datasets were required in the form of Quebec's road network, municipal boundaries and water and land layers (for display purposes). These files were obtained from McGill University's *Transportation Research at McGill* (TRAM) research group (*22*).

When possible, roundabout site information was also verified either through site visits or through engineering plans provided by the various municipalities contacted at the beginning of the project.

Additionally, Google Maps © was extensively consulted for the collection of geometric and geographic data that was unavailable from other sources and would require too much travel for a site visit (*15*). Furthermore, the application provided the perfect platform for performing the comparison site selection described in Chapter 4.

The geographic data was collected in order to determine the type of setting in which a roundabout was constructed. Similar to a basic land use variable, the variable was considered to represent the following types of application:

- Roundabouts in a residential setting on local streets;
- Roundabouts located at the entrance to a municipality (usually on the boundary between local and regional roadways);
- Roundabouts located at the intersection of two or more regional highways;
- Roundabouts located at highway off-ramps and serving as a simple interchange;
- Roundabouts in a rural setting in which the surroundings are undeveloped;

More traditional land use categorizations such as urban, suburban, rural and industrial were also considered.

1.4.3 Accident records

As previously mentioned, the data used for the safety analysis presented in this thesis was provided from both Quebec's Ministry of Transportation (the *Ministère des Transports du Québec (MTQ))* and Quebec's Automotive Insurance Board (*Société de l'assurance automobile du Québec* (SAAQ)). It is based on a dataset containing all recorded motor-vehicle crashes on both the provincial and municipal road networks within the Province of Quebec for the period of

2000 to 2011. Initially, a total of 762,718 crash records, corresponding to crashes having occurred in municipalities that have roundabouts on their territory, were extracted. TABLE 3 presents a breakdown of the available crash records by municipality. From this subset, accidents occurring within proximity of a roundabout were identified for further analysis. The area covered by the available records accounts for 65 individual roundabouts.

Each record contains information on the time, date and location of the accident, as well as characteristics of the roadway, the environment, the vehicle and the driver. The injury severity levels of all individuals involved in a given accident are also included in the dataset.

One of the main limitations of the provided crash record database is that of the 762,718 available crash records, 295,619 (or 38.8% of the total records) lack the geo-localization that would enable spatial analysis. In other words, the location information provided with these records consists only of textual address fields, which cannot be directly analyzed by spatial methods.

With no indication as to why certain records contain geocoding information and others don't, the possibility of introducing important biases into the results is too great to ignore. In order to deal with this important issue, a methodology was developed for geocoding crash data based on online services such as the Google Maps Application Programming Interface (API) (*23*). This research is presented in the following chapter.

TABLE 3: Crash record summary by municipality.

Municipality	Total Records	Municipality	Total Records
Amos	6,641	Sainte-Agathe-des-Monts	4,888
Boucherville	8,049	Saint-Julie	5,498
Chambly	5,432	Saint-Gedeon	690
Gatineau	80,540	Saint-Henri	1,436
Granby	10,063	Saint-Irenee	227
Montreal	464,400	Shawinigan	20,127
Mont-Saint-Hilaire	4,221	Trois-Rivieres	51,471
Mont-Tremblant	6,771	Val-d'or	13,116
New Richmond	1,430	Vaudreuil-Dorion	11,540
Pointe-Lebel	400	Yamachiche (Louiseville)	1,137
Saguenay	64,641		
		Total	**762,718**

Chapter 2 Crash Data Mapping and Analysis

2.1 INTRODUCTION

Traffic crash records are essential input data in the road safety management process. Crash data is essential in different steps of traffic safety studies and programs, including network screening (hotspot identification), safety performance function development, before-after observational studies, among others (*24*). Traditionally, crash records have been the foundation of most road safety studies and of the development of road design guides and countermeasures. In most cases in transportation engineering, police reports are the main source of crash records. Other sources, such as injury data from ambulance and hospitalization reports, are less popular. The popularity of police reports in transportation engineering, particularly within North America, is partly due to the availability of a relatively large amount of information regarding the road environment, vehicles, passengers/drivers and weather. Records typically include a number of fields, the purpose of which is to capture the consequences of the crash (number and category of all injuries, assessment of damage, etc.), the characteristics of the crash (type of impact, number of vehicles involved, environmental and roadway conditions, etc.) as well as a police officer's professional opinion on the probable cause of a crash (*25*).

Despite its acceptability, crash data suffers from weaknesses including underreporting, localization errors, varying levels of detail, missing information, and misclassification. Another important issue is the lack of accuracy in the geographical location of each crash (X-Y geographic coordinates). Although text-based address fields are included on each report, the level of detail included can create substantial ambiguity (*26*). The inclusion of accurate spatial coordinates can have important implications in the results of a traffic safety study. The miss-location of crash records can lead to wrong conclusions as the diagnosis is largely based on the relationship between traffic crashes and road network characteristics (*25*).

Geocoding methods are employed in order to link text-based addresses to X-Y geographical coordinates. In general terms, a geocoding method is defined as the process of assigning geographic coordinates to an input feature (i.e. an address) (*26*). Although satisfactory results can be obtained under ideal conditions, geocoding of crash records can suffer from a number of issues. Incomplete address input data, as well as the inclusion of shorthand notation and spelling mistakes can impact the geocoding results. Another important issue is that commercially available geocoding programs tend to exhibit a lack of flexibility when presented with different data structures and output requirements (*26*) (*27*). Additionally, programs that allow for a high degree of matching accuracy often require a large cost, increased technical knowledge, and quality reference maps upon which to base the geocoding results. There is a need for flexible geocoding methods that provide a balance between implementation complexity and the acceptable level of accuracy (*27*).

This research proposes a simple and practical method based on online services such as the Google Maps Application Programming Interface (API) for geocoding crash data. The level of matching success as well as an indication of spatial accuracy achieved is evaluated as part of the objectives in this research. Additionally, the inherent limitations of this freely available web-based service will be discussed. In order to evaluate the accuracy of different geocoding systems and gain a better understanding of the results that can be obtained from custom API-based algorithms, a case study using the available Quebec crash data is used as an application environment. From here, the geocoding methodology is extended to the entire crash record database in preparation for the safety analysis presented in the following chapter.

2.2 LITERATURE REVIEW

As documented in the literature, there are three main methods for the localization of crashes, depending on the location information provided: link-node or address field, route-km point and the global positioning system (GPS)-based approach. In the link-node or address approach, crash location is identified using the distance from a node, with known points along the road being identified as nodes (e.g., intersections). In some cases, the address of the event is given (street number, name and municipality, etc.). In the route-km point approach, one makes use of unique route numbers and unique identifiers such as mile-markers that are assigned to a continuous section of road. This is typically used for mapping highway crashes. In the third case, the coordinates of each crash record are obtained directly from GPS units at the scene of a crash, reducing errors related to the spelling, description, and transcription of address descriptors and potentially increasing the location accuracy if used correctly (*28*). Despite the advantages of a GPS-based data collection method, many jurisdictions are still reporting crash data based on the first two methods. This approach has yet to be made available to all levels of first responders, and the availability of this data cannot be relied upon. Hence, the development of crash mapping tools has been highlighted in the literature, with the basic objective to assign a location (X-Y coordinates) to each crash report, as well as taking into account potential temporal variations in location names (*26*) (*27*) (*29*).

Due to its importance in fields such as public health, police crime tracking and traffic safety research, extensive literature exists that investigates the geocoding of spatial records (*30*) (*27*) (*31*). A prevalent conclusion of the existing literature is that irrespective of the method used, geocoding results are directly related to the quality and completeness of the address input (*26*) (*27*). A number of studies have shown that electronic field-based data collection and entry can both increase efficiency and accuracy of spatial matching, due to the reduction of transcription and typographical errors (*32*). Spatial accuracy can be further improved by the collection of postal codes on all crash reports, as postal codes are typically well known and less likely to promote spelling mistakes or colloquial descriptors of the location (*33*). Nevertheless, detailed records are of little help without an appropriate geocoding algorithm to properly interpret the input data and provide an output of the desired spatial coordinates.

The typical geocoding process involves three primary steps: data standardization, record matching, and event location (*25*) (*34*) (*27*). Data standardization is an important step to consider as real-world data is known to be imperfect, with incomplete fields, incorrect formatting, misspellings, use of shorthand notation and alternative place names often being quoted as problematic (*26*) (*35*) (*33*). Record matching is generally the most important step, and can lead to the greatest error. Goldberg et al. (*36*) caution that three different types of errors can be encountered, each one having different implications on the final results. Geocoding error can be considered to come from low spatial accuracy, false matches, or from the invalidity of assumptions made during the match (*36*).

Throughout the literature, two main categories of matching algorithms can be observed; deterministic and probabilistic matching (*33*). In general, both types of matching algorithms rely on the availability of appropriate reference tables used to match a similar address to the input field and return the linked geographical coordinates. Deterministic (or rule-based) matching can be difficult to set up, with adjustments to the rules often being required (*26*). The main weakness with this type of matching is that a binary output condition exists whereby a match is either successful, or the process fails. Probabilistic matching sorts potential matches from the reference table according to the degree of separation between the input address and the reference table.

Hence, the algorithm will return the most accurate match, and provide alternative matches according to a decreasing match score (*27*).

Throughout the literature, geocoding is undertaken through both commercially available software packages and custom algorithms. These options can vary greatly in price and quality (*28*). In terms of commercial options, many of the well-known GIS programs include geocoding tools or functionality (*28*) (*26*) (*29*). Online commercial geocoding services also exist, charging either a per-request fee or through purchase of a membership (*36*). More interesting to this research, however, are those services that offer access to an application programming interface, which has the potential to provide geocoding requests free (with some limitations), or with a small fee that provides for additional functionality (*23*) (*37*). Additionally, these services provide an alternative to the rigid data structure required by typical geocoding programs.

Evaluating the accuracy of a geocoding algorithm is often a difficult task. It is widely accepted that due to the sheer volume of records that typically require geocoding, perfection would be unattainable (*35*). Accuracy is not always the most important factor, as many applications do not require high accuracy to provide meaningful results (*34*) (*27*). In public health applications, it is sufficient to assign cancer occurrences to the census-tract level (*36*). Many studies quote the percentage of matched records as a measure of geocoding performance, although match rate is fundamentally different from the accuracy measure (*28*).

In general, the literature suggests that a match rate between 70% and 83% can be considered a good rate for address geocoding (*30*) (*26*) (*32*) (*38*). Bigham et al. (*31*) state that for intersection-coded collisions a success rate of 86% is acceptable. Finding similar results, although by a different method, Ratcliffe (*35*) proposed that a minimum acceptable geocoding rate of 85% was required for the mapped data to be representative of the final map if all records had been successfully geocoded. Attempting to improve the understanding of typical geocoding error, Zandbergen (*33*) provides an estimated range of geocoding positional error to be from 25 to 168 m. With this type of inaccuracy in mind, Levine et al (*29*) suggested that although records are often provided with a directional offset from an intersection, the difficulties in interpreting the offset along with the inherent inaccuracy of the geocoded estimation render the offset value ineffective. For this reason, they mapped all incidents to the nearest intersection.

2.3 METHODOLOGY

With the goal of investigating the use of online geocoding APIs, three main steps were followed. These steps include: i) the selection of an appropriate API, ii) the development of a custom algorithm, and iii) testing of the algorithm to establish the potential match rate and accuracy that could be expected from such a service. Details of each step are explained below.

2.3.1 API Selection

At the beginning of this work, a number of APIs were identified and considered for further investigation. These include services offered by companies such as Yahoo!, MapQuest and Google Maps, as well as from open source services such as GISgraphy (*http://www.gisgraphy.com/*) and Nominatim (*http://nominatim.openstreetmap.org/*) , which rely on OpenStreetMap geospatial data (*39*) (*40*).

A number of considerations needed to be taken into account, such as geographic coverage areas, level of detail and the reliability of finding an acceptable match given a certain quality of input data. Services that only covered the United States, for example, were not considered, as the primary dataset upon which this work is derived is from Canada. A number of tests were sent to the various API services for reliability assessment purposes. The comparison was simplified by an API comparison tool provided on the GISgraphy website. An example of the API comparison is presented in FIGURE 11**.** As can be seen, results vary across the different services, with some being completely unacceptable.

For this research, the Google Maps API was selected as the online geocoding service to be evaluated due to its consistency in returning addresses, and a seemingly higher accuracy (more details are presented in the results section). The API is also ideally suited for use with a mixed database of address description formats, as both link-node and address point geocoding localization methods are supported (*23*).

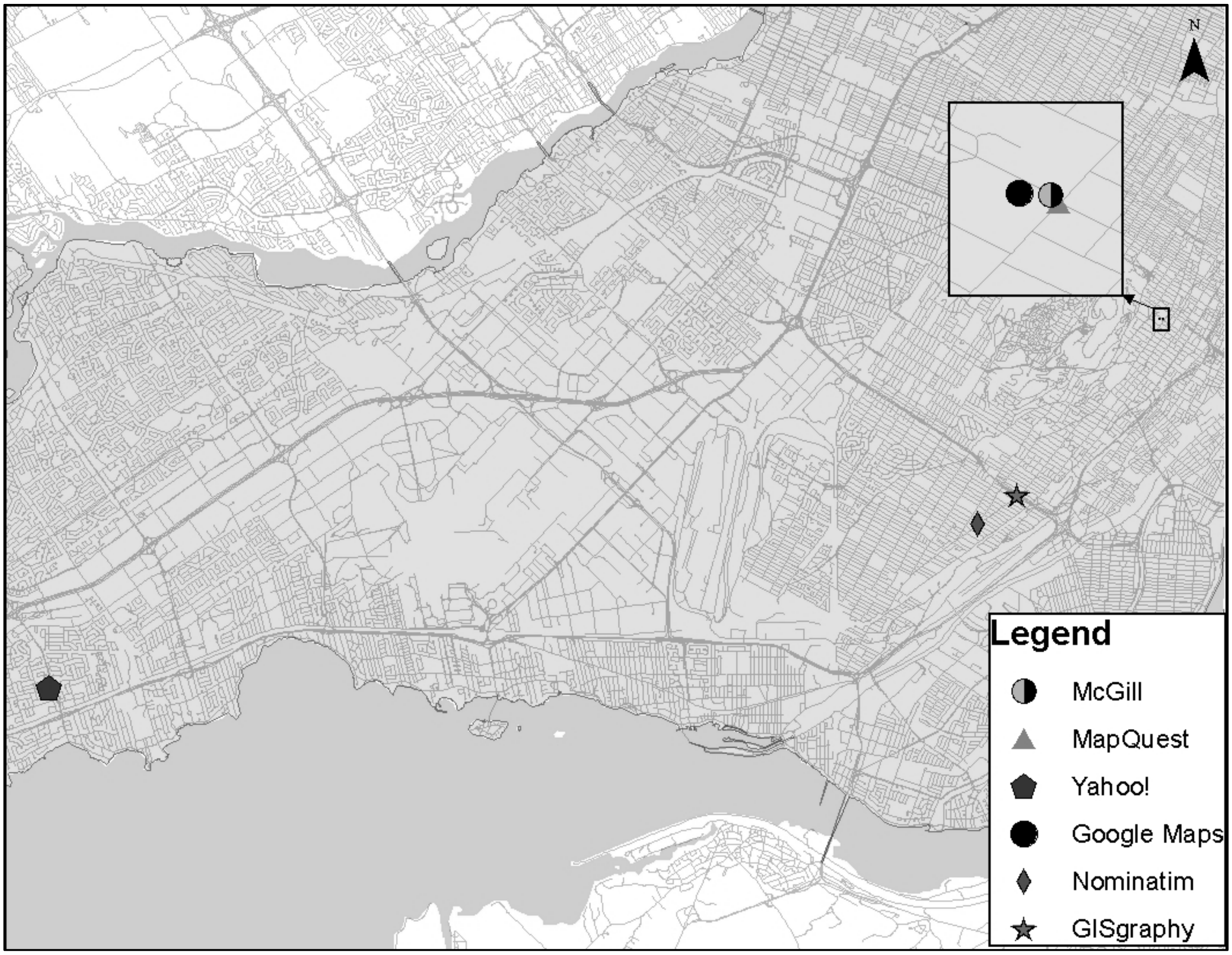

FIGURE 11: API comparison using McGill University's address (*39*).

2.3.2 Programming Environment

In order to evaluate the Google Maps API, a custom algorithm was created using the Python programming language. As can be seen in FIGURE 12, the algorithm is used to read a crash database file and interpret the supplied fields. The algorithm then attempts to clean the address fields by removing redundant information (if multiple fields contain the same information, for example), and special characters (such as capital letters, accented letters (in French) which may be misinterpreted, and punctuation). The algorithm can also be used to replace commonly misspelled words and typos, as setup by the user.

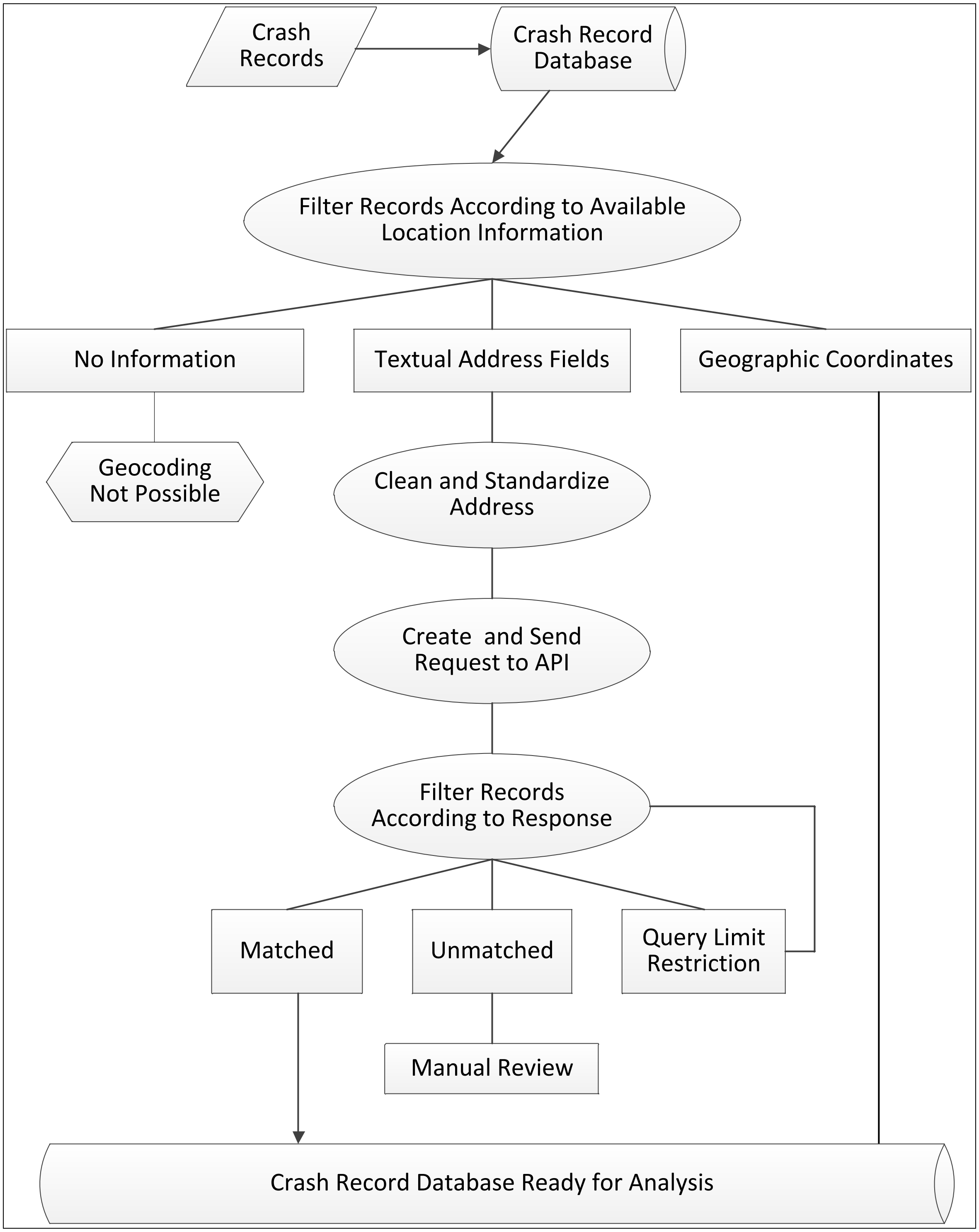

FIGURE 12: Flow chart illustrating the design of the Python algorithm.

Once this is done, the algorithm is set to call the Google Maps API using a Hypertext Transfer Protocol (HTTP) request message. An important benefit of using the API is that the address does not need to be parsed by the user. Being based on a probabilistic matching method, the API is able to parse the input, as well as infer the formatted address based on information within its proprietary reference tables. This server-side processing is beneficial as it considerably reduces the technical knowledge needed by the user.

The algorithm then collects the API's response, which can be in either *json* (JavaScript Object Notation) or XML formats (*23*). To help assess the accuracy of the returned location coordinates (i.e. latitude and longitude), the API includes a tag indicating the type of mapping accuracy that was successfully returned for a given match. The possible tags are as follows:

1. **Street_address:** Indicates that the result is a precise street address.
2. **Intersection:** Indicates that the result is at the intersection of two streets.
3. **Route:** Indicates that the result is a named street segment.
4. Political, country, administrative_area_level_1, administrative_area_level_2, administrative_area_level_3, locality, sublocality, neighborhood: Indicate that the result is within a political or civil entity (such as a municipality, province, etc.)
5. **Colloquial_area, premise, subpremise:** Indicates that the result is a named location, such as a well-known building.
6. **Postal_code:** Indicates that the result is a postal area.
7. Natural_feature, airport, park: Results are as indicated.
8. **Point_of_interest:** Indicates that the result is a local point of interest that does not fit in another category (*23*).

The API's responses are returned and ranked from the most to the least accurate match levels. For this paper, only the first three tags (i.e. *street_address, intersection* and *route*) were considered to be useful matches; all subsequent tag levels were considered to have not returned a match. Due to the API's usage restrictions, the algorithm is also programmed to identify when the maximum daily match limit has been met.

2.3.3 Analysis and Mapping

In order to evaluate the quality of the geocoding achieved by the Google Maps API implementation, two quality measures were analyzed. The first is the match rate. As mentioned above, a match is achieved if *street_address*, *intersection* or *route* was returned as a match indicator tag. The *route* tag is less accurate than the other two options due to the fact that it indicates that the record occurred somewhere along a street segment. Nevertheless, it is considered to be of sufficient accuracy for intersection safety analysis studies as it is still possible to conclude to which route the crash belongs (and to exclude it from analysis at intersections).

The second quality measure attempted to capture the level of accuracy that was provided by the geocoder. This was done by comparing the latitude and longitude provided by the Google Maps API geocoder to those that were already provided with some of the Quebec crash records. Although a comparison to the actual location as found on a map would be more representative of the true error, this would be impractical due to the number of records contained in the dataset. The results of the comparison with the previously geocoded records are presented below.

2.4 DATA

In order to evaluate the advantages and accuracy of the methodology, records from the municipality of Amos, Quebec were considered. This municipality was selected due to the large proportion of non-geocoded crash records, as well as its remote nature. The logic behind this was that if the API is capable of geocoding a smaller, remote municipality, it should be able to handle larger municipalities as well.

The total number of crash records for this municipality is 6641 records, with only 50% (or 3322) of records having been supplied with coordinate references. It is interesting to note that of the 3319 crash records that lack coordinates, only 22 records are from crashes occurring on roadways under provincial (MTQ) jurisdiction. This is most likely due to the addition of geographic coordinates at the time of digitization of the records.

Although a number of fields are contained in each record relevant to the crash, the primary focus in this study is on the location fields. These include:

- **ADR_NUMR_IMMBL:** Street number of a house/building near the crash site.

- **ADR_NOM_VOIE:** Street name on which the crash occurred.
- **VAL_NUMR_ROUTE:** Route number if applicable (such as a numbered highway, etc.).
- **NOM_VOIE_INTSC:** The name of a cross-street if the crash occurred at, or in proximity to an intersection.
- **VAL_AUTRE_IDENT_REPR:** Name of other identifying landmark if available.
- **VAL_DISTN_REPR:** Distance (in metres) to the intersection or landmark.
- **DES_TYPE_DIRCT:** Direction from the crash (if distance is not 0).

As with any form of real-world data, these fields are not always filled-in correctly. Because of this, it is possible to observe records with incomplete address information. Common issues include missing street numbers, partial street names, lack of a cross street or other landmark, among others. Each of these issues leads to a reduction in the address quality, and reduces the chances of accurate location information being returned by a geocoder.

2.5 RESULTS & DISCUSSION

2.5.1 API Selection

As previously mentioned, a number of APIs were considered for use in a custom algorithm. In order to determine which API was most likely to reliably return responses with an acceptable accuracy, the output was compared for a number of test cases. TABLE 4 presents a sample of these test cases, as well as an indicator to outline whether the coordinates are valid for the input location.

As can be seen, very different performance is obtained from the APIs. Surprisingly, Nominatim returned no results for any of the input addresses. It is possible that the API is not able to interpret the addresses that are being sent to it, or that its reference tables do not cover the region in question. Therefore, this API was dropped from consideration, as it would not allow for the geocoding of the available dataset. GISgraphy was similarly dropped due to the fact that the results returned were so inaccurate that they did not even fall in the municipality of interest. Both Yahoo! and MapQuest performed similarly, with half of the input addresses being returned successfully. The Google Maps API was ultimately selected for further analysis in this paper as it returned valid coordinates for all but the last test case. The last test case was handled in the same way by all of the proprietary source geocoders: due to the incomplete nature of the address, a guess was made as to the full address.

TABLE 4: Detailed API Comparison for (A) APIs with Proprietary Data and (B) Open Source Data.

(A) Proprietary Source

Address (in Amos, Quebec)	Google Maps			Yahoo!			MapQuest		
	Longitude	Latitude	Valid?	Longitude	Latitude	Valid?	Longitude	Latitude	Valid?
Des Metiers at Av Du Parc	-78.1229	48.5608	Yes	-78.1231	48.5607	Yes	-78.1231	48.5607	No
343 6e Rue Ouest	-78.1309	48.5693	Yes	-78.0121	48.6110	No	-78.1311	48.5686	Yes
94 Principale Sud at Du Metro	-78.1158	48.5697	Yes	-78.0121	48.6110	No	-78.1160	48.5736	No
4e Rue Est at Gravel	-78.1063	48.5650	Yes	-78.1065	48.5649	Yes	-78.1065	48.5649	Yes
82 1e	-78.1330	48.5731	Guess	-78.1133	48.5719	Guess	-78.1176	48.5719	Guess

(B) Open Source

Address (in Amos, Quebec)	GISgraphy			Nominatim		
	Longitude	Latitude	Valid?	Longitude	Latitude	Valid?
Des Metiers at Av Du Parc	-73.7058	45.5531	No	-	-	No
343 6e Rue Ouest	-73.8667	45.5480	No	-	-	No
94 Principale Sud at Du Metro	-73.3233	45.3214	No	-	-	No
4e Rue Est at Gravel	-73.6299	45.6001	No	-	-	No
82 1e	78.1064	48.5659	No	-	-	No

*Validity refers to whether the returned coordinates are within an acceptable distance from the true coordinates of the address.

2.5.2 Matching Proficiency

The algorithm output was obtained in a comma-separated file that could be analyzed in Microsoft Excel software to establish a preliminary match rate and accuracy estimation. For the Amos, Quebec crash records, it was found that of the 3319 records that lacked geographic coordinates, 2586 (or 78%) of the records were matched to either an *intersection* or *street_address* level. Assuming that the results are being used to perform an intersection safety analysis, it would also be possible to include the *route* results, as this would locate the records along a given route, indicating that they did not occur at an intersection. With this assumption, the match rate is found to increase to 85%. Adding these records to those that were previously geocoded by either the MTQ or the SAAQ, it is found that over 92% of all traffic crash records can be mapped for the municipality. From the results, it can be seen that using a custom algorithm to call upon an online geocoder service can provide a competitive match rate that falls within the accepted rate in the literature for commercially available systems. The main benefit of this method, however, is that it is not required to parse the input address information to match a specific format before passing the input to the geocoder. This information is automatically extracted by the geocoder, with a seemingly high level of confidence.

2.5.3 Accuracy Estimation

In terms of the accuracy estimation, results are less conclusive. As documented in the literature, measuring the accuracy is a difficult task to undertake, and often requires manual verification in order to obtain any level of confidence in the conclusion.

As previously mentioned, the results of the accuracy estimation were obtained by comparing the distance between the known coordinates and those provided by the algorithm. Looking at the raw results, a large discrepancy could be observed for a number of entries. A more comprehensive analysis found that the Google Maps API handled the geocoding of records identified by route number alone (and not the more common name of the road segment) very poorly. Removing these records from the estimation it was found that the average distance between the known and geocoded coordinates is 200 m. An interesting observation however, is that 54 % of the records have a distance between the two coordinate estimations of less than 30 m.

Initially, the average distance error between the previously geocoded coordinates and those obtained with the use of the custom algorithm seems relatively high at 200 m. Looking at the data, however, it can be seen that a wide range of estimations is obtained. Selecting a sample of records with a higher degree of match quality, however, yields an estimated average distance of only 22 m.

One consideration that was investigated in order to clarify this result is that the previously geocoded records were taken to be accurate representations of the crash location, although it is possible that they are in fact estimations in and of themselves. Looking at FIGURE 13, this hypothesis seems to be a possibility as neither the previously geocoded coordinates, nor those obtained from the algorithm are at the true location indicated in the crash record. A manual sampling of the results reveals that in fact, the records from the algorithm are more accurate than the previously known coordinates in many cases. From this, it can be concluded that estimating an accuracy measure by comparing the results to those previously geocoded may be flawed, and that the Google Maps API may in fact provide better estimates than originally thought, provided a high quality address record is available.

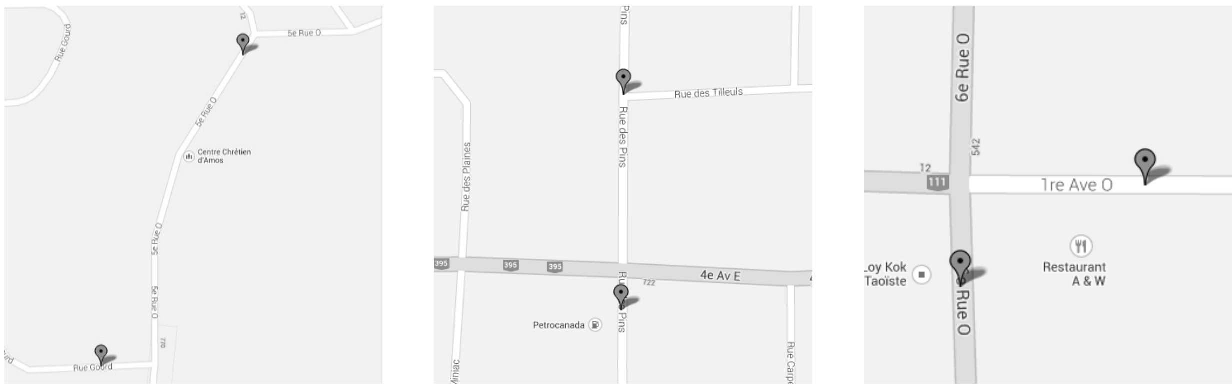

FIGURE 13: Examples of coordinate differences (*23*).

One of the main observations that should be taken away from the preliminary results, however, is that the Google Maps API tends to always provide a match for a given input, even if the match quality and accuracy are low. This may lie in the fact that the user has no direct control on the probabilistic matching limits, and thereby is forced to accept the result returned by the API. Because of this, revision of the resulting matches is suggested, as is caution in the use of the returned results. As manual revision is impractical for large datasets, an automated process should be investigated in order to improve the reliability of the geocoding.

2.5.4 Common Causes of Low-Quality Matching

Looking at the records with a high estimation of distance between the known and geocoded coordinates reveals that a majority of these records have some sort of ambiguity involved in their address fields. This ambiguity prevents the geocoder from returning high quality results. A summary of common shortcomings is presented in TABLE 5 below.

TABLE 5 Sample Address Problems Causing Low-Quality Matches.

Address	Problem
622 Des Javies, Amos, Quebec	Street does not exist in the municipality.
Ruelle Arriere Restaurant Succo, Amos, Quebec	This description is not recognized by the geocoder.
1132 RTE 111 E at 4e rue E, Amos, Quebec	The geocoder has difficulty identifying numbered roadways.
Taschereau at 10 Av E, Amos, Quebec	The geocoder interprets the "10" as a house number, and not the street number due to lacking formatting (i.e. 10e av.).
Av Authier O, Amos, Quebec	No street number is provided as reference on the street, so a general segment location of Avenue Authier is returned.
1e at 2e, Amos, Quebec	The street type is missing in both cases, leading to a guess on the location.
22 Principale, Amos, Quebec	No distinction between Principale North or South. Although a match is returned, the API provides a guess as to which street is meant.

Examples of the address record shortcomings include records located on streets with both "North" and "South" components with no distinction provided in the record, as well as records with addresses such as "1e at 2e". Without the inclusion of street types it is difficult for the geocoder to determine if this is a cross of First Street and Second Avenue, or a similar combination.

2.5.5 Extension of the Methodology to the Entire Roundabout Crash Record Database

Having investigated the use of a custom geocoding algorithm using a service such as the Google Maps API and proven the concept through a case study using a subsample of the crash records available in this research, it was possible to extend the proposed methodology to the entire crash record database as initially described.

It should be noted that in addition to the methodology presented above, an effort was made to manually screen the non-matched crash records for known roundabout intersections in order to be able to better analyze the safety at these types of facilities. To the order of magnitude of the database, this was not possible for all record types.

TABLE 6 presents a summary of the 762,718 records, broken down by region as well as an outline of the proportion of available geographical coordinates before and after having applied the methodology. Note that for clarity, only the matches with either an *intersection* or *street_address* level designation returned by the geocoder were included in this table.

TABLE 6: Results of Geocoding Algorithm Applied to full crash record database.

		Initially		After Geocoding	
Municipality	**Total Records**	**Records with Coordinates**	**% of Records**	**Records with Coordinates**	**% of Records**
Amos	6,641	3,322	50.0%	5908	89.0%
Boucherville	8,049	5,933	73.7%	6,673	82.9%
Chambly	5,432	4,471	82.3%	4,826	88.8%
Gatineau	80,540	64,859	80.5%	72,356	89.8%
Granby	10,063	8,167	81.2%	9,106	90.5%
Montreal	464,400	248,726	53.6%	45,8044	98.6%
Mont-Saint-Hilaire	4,221	3,391	80.3%	3,612	85.6%
Mont-Tremblant	6,771	5,245	77.5%	5,829	86.1%
New Richmond	1,430	1,286	89.9%	1,323	92.5%
Pointe-Lebel	400	305	76.3%	331	82.8%
Saguenay	64,641	48,957	75.7%	56,198	86.9%
Sainte-Agathe-des-Monts	4,888	3,782	77.4%	4,109	84.1%
Saint-Julie	5,498	4,588	83.4%	4,869	88.6%
Saint-Gedeon	690	541	78.4%	591	85.7%
Saint-Henri	1,436	1,224	85.2%	1,270	88.4%
Saint-Irenee	227	157	69.2%	175	77.1%
Shawinigan	20,127	10,493	52.1%	15,871	78.9%
Trois-Rivieres	51,471	32,969	64.1%	43,818	85.1%
Val-d'or	13,116	6,917	52.7%	10,142	77.3%
Vaudreuil-Dorion	11,540	10,810	93.7%	11,224	97.3%
Yamachiche (Louiseville)	1,137	956	84.1%	995	87.5%
Total	**762,718**	**467,099**	**61.2%**	**716,029**	**94.0%**

As can be seen, the crash record database was originally obtained with only 61.2% of records being associated with geographical coordinates. Applying the methodology resulted in a total of 94.0% of records having geographical coordinates, greatly increasing the records available for spatial analysis. As presented in the literature, this level of geocoding is considered to be quite high, and will help ensure that the safety analysis performed in this work is representative of the entire database.

2.5.6 Limitations

Among the limitations of the proposed algorithm, readers should take note that at present the APIs presented in this research are intended for the use of online application developers to include maps on their respective websites and/or mobile apps. Any use beyond this requires special permissions be obtained from the API owners. As such, this work remains largely a proof of concept, with the purpose of showing the potential applications of the technology that is currently on the market.

2.6 CONCLUSIONS

This research has explored the use of online geocoding services such as the Google Maps API as a simple and accessible tool for the geocoding of traffic crash records. Through the detailed analysis of a case study in the municipality of Amos, Quebec, it was found that at the strictest level, a match rate of 78% could be achieved through a custom algorithm. Relaxing of the matching conditions improved the match rate to 85%, although caution should be taken as not all applications can support the associated reduction in match reliability provided. These results are comparable to those obtained from commercially available geocoding options, although manual review indicated that a number of false matches occurred when incomplete input data was sent to the geocoding API.

Expanding the methodology to the entire crash record database available through this research, it was found that a total of 716,029 or 94.0% of the records could be associated with geographical coordinates, which is expected to allow for representative analysis to be explored in the safety analysis explored in the following chapters.

The research also explored the geocoder's spatial accuracy through the case study, although the results tend to vary substantially from record to record. Factors such as the completeness of the input address fields and the ability of the API to interpret the location description (in particular for route number addresses) have a large influence on this outcome.

It is suggested that with proper user revision, the results are sufficient for practical applications such as intersection safety analysis. The use of an intersection's area of influence should compensate for at least some of the observed inaccuracies, and allow for the crash records to be associated with either the respective intersection location, or street segment to which they belong.

Chapter 3 Injury Severity

In order to begin understanding the safety performance of Quebec roundabouts, the injury risk of accidents occurring at roundabouts was investigated. This topic is approached through the application of an ordered logit regression model. The research presented in this chapter was conducted in 2012 and was subsequently presented at the Transportation Research Board's 2013 Annual Meeting and published in the associated conference proceedings.

3.1 INTRODUCTION

The occurrence of road vehicle accidents is well documented with respect to the adverse economic and emotional effects they inflict on society (*41*). Because of this, traffic engineers and policy makers have shown considerable interest in finding ways to identify critical factors of both accident occurrence and their severity outcomes (*42*) (*43*). International experience with roundabouts tends to indicate improved safety and performance when compared to traditional sign- and signal-controlled intersections. The inherent safety benefits of a roundabout are often attributed to the fundamental design features which lead to lower travel speeds, the elimination of head-on and right-angle conflict areas and the need for pedestrians to only cross one vehicle movement at a time (*44*) (*18*) (*2*). However, largely due to the relative scarcity of roundabouts within the North American road network compared to typical intersection controls, few studies have investigated the severity of accidents occurring in or around roundabouts. Most of the previous work on roundabouts has been related to accident frequency or occurrence (*1*).

Due to the large number of roundabouts and particular weather and driver behavioral conditions that exist within the Province of Quebec, it is suggested that detailed safety studies be conducted in order to better understand which factors increase or decrease the severity of accidents (*7*). This could help identify countermeasures or actions that decrease accident consequences in roundabouts. Changes to the established design guides may be required in order to better suit local conditions.

In order to improve the safety of all road users within a roundabout, planners and engineers must be able to modify the physical environment in such a way as to reduce the dangers within each facility. Factors relating to vehicle, driver, roadway design and weather conditions need to be studied to better understand how each characteristic affects the injury severity outcome of a given roundabout. In this way, designers would be able to estimate the safety benefits of a specific design element change by holding all other factors constant.

This research aims to investigate a number of factors which may be associated with the injury severity outcomes of accidents occurring in and around roundabouts. Accordingly, the most significant factors which affect the severity of accident injury outcomes are identified using an ordered logit regression model. The studied variables pertain to roadway, environmental, vehicle and human behavior characteristics that are thought to impact injury outcomes (*45*) (*46*).

3.2 LITERATURE REVIEW

Over the last few decades, there has been extensive literature on injury severity analysis in road safety based on historical accident records. An important part of this literature deals with the identification of accident factors that are associated to injury severity outcomes at various levels. Among these factors, one can mention road geometry, signs and traffic control characteristics at the accident location, traffic conditions at the moment of the accident, and vehicle and passenger attributes including driver, weather, visibility, surface conditions, etc. As a unit of analysis, severity studies commonly use the individual (passenger and driver), vehicle or accident level. The analysis level selected is often dependent on data constraints (*47*). In this important literature, some studies have also looked at accident occurrence outcomes classified by injury severity types. Also, to model injury severity outcomes, many statistical methods have been proposed including traditional ordered response models to take into account the inherent ordering of the reported injury severity levels (*48*). Other models such as probit, multinomial, mixed logit and latent class model have been used (*49*) (*50*) (*51*). For a comprehensive literature review of the statistical methods used in this topic, one can refer to (*41*). Many injury severity studies have also identified the contributing factors of accidents at different location types such as rural roads, signalized urban intersections or special facilities such as highway railway crossings and freeway ramps (*42*) (*52*) (*53*).

In this rich literature dealing with a variety of issues, various reports have investigated the safety of roundabouts. In this type of intersection, a particular aspect that has attracted a lot of attention is the effectiveness of roundabouts when compared to either sign- or signal-controlled intersections. Many before-after studies have investigated this important question in Europe and North America. The effects of converting intersections to roundabouts have been documented in various literature-review studies. For instance, one can refer to the work of Elvik (2003) that carried on a meta-analysis of studies reported outside the United States. This study concluded that roundabouts are associated with a 30% to 50% reduction in the number of injury accidents and fatal accidents are reduced by 50% to 70% (*54*). The changes on property damage crashes are highly uncertain, and an increase often can occur in some conditions (e.g., three-leg intersections). An important NCHRP report, entitled *Roundabouts: An Informational Guide*, has also made an important effort to summarize this literature (*2*).

Despite this rich literature covering several issues in roundabout safety, very few studies have investigated which factors have the highest influence on the injury severity outcomes at roundabouts. Perhaps the only study looking at injury severities at the vehicle and crash level is the recent study of Daniels et al. (2011). They investigated the factors associated to severity of crashes or injury outcomes at roundabouts - using injury crash records on roundabouts in Flanders-Belgium. Logistic regression was used for this purpose to represent the two grouped outcomes (fatal and serious injuries as one category, and minor or property damage only in the second category). Among other results, they found that crash severity is strongly dependent of the involved types of road users. In particular, vulnerable users (pedestrians, bicyclists, moped riders and motorcyclists) have a higher probability of getting seriously injured in a roundabout crash. Bicyclists represent almost half of all those killed or seriously injured in multiple-vehicle collisions at the investigated roundabouts. As in other facilities, the effects of age, geometry and light conditions are less substantially correlated to the injury severity (*55*). No studies of this type have been reported in North America, in particular in Canada. This can be associated to the lack of data and spatial location of crashes. The research in this chapter aims at investigating potential factors that have a large influence on the injury severity outcomes of accidents at roundabouts. To the author's knowledge, this is the first severity study of this type in Canada and the Unites States.

3.3 DATA

The data used for this section of the research is comprised of a subset of the available accident records previously described. As this analysis was completed early in the research process, it does not consider the full extent of the database.

Each accident record provides a wide range of characteristics with the goal of capturing the exact circumstances of a given accident. From the information available in the accident reports, the variables considered to estimate the injury severity models are presented in TABLE 7. It should be noted that the table contains all tested variable formulations, and that some of the variables were tested independently from other variables that explain the same property.

Two subsets were identified within the data, and are distinguishable according to whether the accident record exists in the MTQ or SAAQ datasets. A degree of redundancy was observed between the two datasets. For this reason, only unique records were taken from the MTQ database to complement the SAAQ dataset.

According to data availability and various physical characteristics, 37 sites were considered in this study, as seen in FIGURE 14. Although a large amount of data was provided, the injury severity analysis in this chapter only considers accident records that occurred after the construction of a given roundabout.

From the remaining accident records, the reduced dataset was mined to extract the records that occurred within the roundabout's area of influence. The area of influence includes the area within the roundabout's boundaries, as well as a certain distance along the approaches, whereby drivers begin to respond to stimuli at the periphery of the intersection (*56*). The area of influence is difficult to define, and tends to vary for each intersection. For this research, the area of influence was taken to be the land contained within a 100 m radius from the center of the roundabout. This value falls within the range of influence area sizes of 15-150 m which can be found in the literature. The size of the influence area considers the size of the intersection as well as the posted speed limits (*56*).

Caution was taken when multiple intersections were present within the area defined to be a roundabout's area of influence. For these sites the data had to be further reduced to ensure that the accidents were not being influenced by an intersection other than the roundabout. This was done by considering the proximity to other intersections, the direction of travel and other characteristics of the accident.

TABLE 7: Variables Available for Analysis and their Legend.

Category	Variable	Legend
DEPENDENT	Crash Level	0=Property Damage Only, 1=Minor injuries, 2=Severe and Fatal Injuries
INDEPENDENT		
Season	Winter	0=Not Winter, 1=Winter (Dec-Jan-Feb-Mar)
Day of Week	Weekday	0=No, 1=Yes
	Friday	0=No, 1=Yes
	Weekend	0=No, 1=Yes
Time of Day	Day (6:00-18:00)	0=No, 1=Yes
	Evening (18:00-24:00)	0=No, 1=Yes
	Night (24:00-6:00)	0=No, 1=Yes
Lighting	Daylight	0=No, 1=Yes
	Nightlight	0=No, 1=Yes
	Dark (Night, no lights)	0=No, 1=Yes
	Twilight	0=No, 1=Yes
Accident Size	Number of Vehicles	Continuous
Vehicle Type	Truck	0=No, 1=Yes
	Bus	0=No, 1=Yes
Collision Type	Animal (Vehicles strikes animal)	0=No, 1=Yes
	Pole (Vehicle strikes utility, sign pole)	0=No, 1=Yes
	Vehicle (Vehicle strikes another vehicle)	0=No, 1=Yes
	Structure (Vehicle strikes building, bridge)	0=No, 1=Yes
	No Impact (Vehicle rolls over, falls in ditch)	0=No, 1=Yes
	Guardrail (Vehicle strikes a guardrail)	0=No, 1=Yes
Surface Condition	Dry	0=No, 1=Yes
	Wet	0=No, 1=Yes
	Snow_Ice	0=No, 1=Yes
	Bad Surface (Effects of both wet and snow)	0=No, 1=Yes
Weather	Clear	0=No, 1=Yes
	Rain_Fog	0=No, 1=Yes
	Snow	0=No, 1=Yes
Accident Type	Single Vehicle	0=No, 1=Yes
	Intersection (Occurred at an intersection)	0=No, 1=Yes
	Lane Change	0=No, 1=Yes
	Hit and Run	0=No, 1=Yes

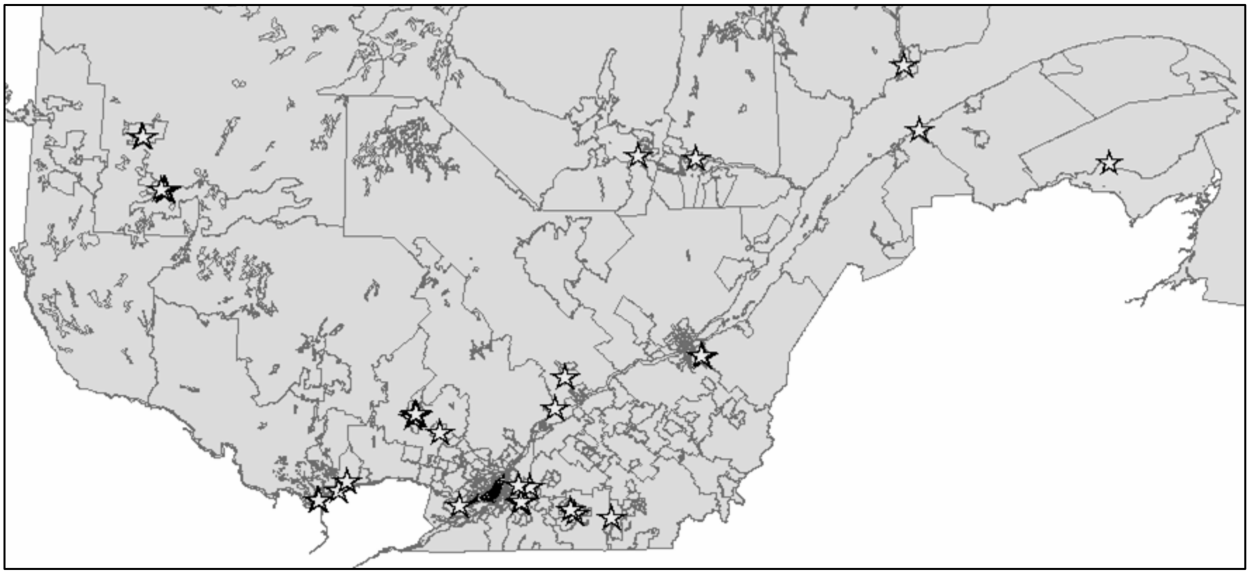

FIGURE 14: Study Site Locations.

From the initial datasets, a total of 1675 motor-vehicle accident records were found to satisfy the above-mentioned requirements; 1200 are from the SAAQ dataset and a further 475 are supplemented from the MTQ dataset. It is worth noting that each record contains the data of a given accident. In other words, the available records are presented in an aggregate accident-level format, which limits the availability of both individual and vehicle-based information. For the analysis of accident-level data, the injury severity was taken as the worst injury that occurred as a consequence of that accident (*57*).

TABLE 8 presents the distribution of crashes by severity level, as taken from the accident reports. Looking at this table, it can be seen that a majority of accidents (91.16 %) are of the property damage only type. This agrees with the findings of several studies in the United States which have found that accidents at roundabouts tend to be less severe than at typical intersections due to the exchange of the most severe crossing conflicts for less severe merging conflicts (*2*). It must be kept in mind that the distribution seen in TABLE 8 may be skewed due to the different reporting rates present across severity levels. Accidents causing only minor damage are much less likely to be reported than accidents that cause injuries. This variation is influenced by factors such as accident severity, time of day and number of users involved. Higher severity accidents tend to have higher reporting rates than damage-only accidents (*41*) (*57*).

TABLE 8: Frequency Distribution of Accident Severity Categories.

Crash Level	Frequency	Percent
Property Damage Only	1,527	91.2 %
Minor Injuries	141	8.4 %
Severe Injuries	5	0.3 %
Fatality	2	0.1 %
Total	1,675	100.0 %

Due to the small amount of observed severe and fatal injuries, as well as the similarity of their consequences, these two categories were combined together. According to Yang et al. (*58*), this merging will also help reduce the potential correlation effects between the closely related categories.

3.4 METHODOLOGY

As is evidenced through many of the most common classification methods, accident severity data displays an inherently ordinal nature in the form of no injury, minimal injury, minor injury, severe injury and fatal injury outcomes (*41*) (*58*) (*59*). Using the ordinal nature of the outcome variable is vital in the selection of an appropriate modeling approach (*41*).

Although numerous modeling methods are applicable to accident severity data, some of the better results have been obtained using extensions of basic multinomial logit modeling (*58*). One of the most fundamental concepts to remember when using these models is that the models are only applicable for the assumption that an accident has already occurred; these models do not attempt to predict accident occurrence, only their injury severity outcomes (*60*).

The model considered in this research is a single-level ordered logit model. The goal of this analysis is to predict the most severe injury level expected to result in a given accident based on the variables presented in TABLE 7. For this analysis, three injury severity outcomes were considered: Property Damage Only (PD), Minor Injury (MI) and Severe/Fatal injury (S/F).

An ordered response model is used in conjunction with a latent variable framework, and focuses on the principle that the choice process is based on a uni-dimensional index function. The general form of the model function is shown in Equation 1:

$$\boldsymbol{y^*}_{\boldsymbol{i}} = \boldsymbol{\beta X_i} + \boldsymbol{\varepsilon_i} \qquad \textbf{EQ. 1}$$

where

y_i^* = latent variable measuring injury risk of each accident i,

X_i = vector of observed explanatory variables measuring the characteristics of accident i,

β = vector of estimated parameters

ε_i = random error term.

As is the case with ordered logit models, all error terms are assumed to have a zero mean, and are assumed to be uncorrelated between observations (*61*). The goal of the model is to estimate the values of the unknown vector β. Standard regression techniques cannot be applied, however, as the dependent variable is unobservable. Within the dataset, however, we are provided with the observed variable Y_i, which is coded to represent the most severe injury outcome of a given accident (Y_i=0 if accident *i* results in property damage only; Y_i=1 if accident *i* results in minor injuries; Y_i=2 if accident *i* results in severe or fatal injuries). The relationship between the injury severity variable Y_i and the latent injury risk variable y_i^* is defined using threshold values as follows:

$$Y_i = \begin{cases} 0 & if -\infty \leq y_i^* \leq \psi_0 \\ 1 & if \quad \psi_0 \leq y_i^* \leq \psi_1 \\ 2 & if \quad \psi_1 \leq y_i^* \leq \infty \end{cases} \qquad \textbf{EQ. 2}$$

where the values $\psi_0$and $\psi_1$are unknown parameters that also need to be estimated. From here, the probability that injury *i* occurs during accident *j* is equal to the probability that injury risk y_i^* is found to be within a given set of thresholds. The probabilities are calculated as follows:

$$P(Y_i = 0) = P_{0j} = CDF(\psi_0 - \beta X_i)$$

$$\boldsymbol{P(Y_i = 1) = P_{1j} = CDF(\psi_1 - \beta X_i) - CDF(\psi_0 - \beta X_i)}$$

EQ. 3

$$P(Y_i = 2) = P_{2j} = 1 - CDF(\psi_1 - \beta X_i)$$

where CDF represents the cumulative distribution function of the random error term ε_i, and all other terms are as previously defined. The ordered logit model was applied to the dataset using Stata data analysis and statistical software (Version 10).

3.5 RESULTS

Multiple variable combinations were tested in order to obtain the best possible model. Subsequent models were compared using the likelihood-ratio test to ensure that any additional parameters provided sufficient explanatory power to the model. Furthermore, variables that fell outside of a 90 % confidence interval were not considered for the model. After numerous trials, the optimal model given the available data was obtained. In the analysis, correlation among variables was verified to avoid co-linearity. The results are presented in TABLE 9 below.

TABLE 9: Results of Ordered Logit Model

Independent Variable	Coefficient	t-ratio	Significance
Number of Vehicles	1.0838	4.18	0.000
Intersection	0.6750	2.96	0.003
Vehicle	-2.1676	-7.24	0.000
Animal	-2.2385	-2.11	0.035
No Impact	0.5806	2.00	0.046
Bus	1.4109	2.59	0.010
Dark	1.0914	2.61	0.009
Hit & Run	-0.5178	-1.83	0.067
Snow	0.8153	2.16	0.031
Snow & Ice	-0.6449	-2.05	0.040
Constant 1	2.9305		
Constant 2	6.1716		

As can be seen in TABLE 9, the variables retained in the final model were found to be statistically significant to a 90 % confidence interval. Parameters of interest during the model testing include the log likelihood value as well as the pseudo R^2 value. The goal for these parameters was to maximize the log likelihood value, and for the pseudo R^2 to be as close to 1 as possible. For the optimal model in this report a log likelihood value of -478.7 was obtained, as well as an R^2 value of 0.0943. Although these values could be improved, they are the best that were found given the data available at the time of the study.

3.5.1 Interpretation of Results

From the obtained model results, a number of interesting findings can be extrapolated. Both the sign and magnitude of the coefficients can be used to quantify the sensitivity of the latent injury severity measure with respect to the value of a given variable. A list and a brief explanation of the possible reasons for the coefficient signs are provided below. These explanations are provided in order to put the coefficient signs into context, and to indicate to the reader that the coefficients obtained from the model are reasonable. As the accidents were not directly observed for this study, other explanations are also possible.

A number of parameters within the model were found to have negative coefficients. Parameters with negative coefficients tend to decrease the injury severity measure. They are as follows:

- *Snow_Ice*: As evidenced in a number of studies, snow and ice-covered roads lead drivers to be more cautious, often cancelling the effects of the increased risk (*62*) (*63*).
- *Vehicle*: Due to the reduced severity of the conflicts present in or around roundabouts, two vehicles colliding have a smaller probability of severe injuries occurring (*2*).
- *Animal*: Due to the reduced speeds in or around roundabouts (20), it is unlikely that striking an animal would result in severe injuries.
- *Hit & Run*: Accidents in which a driver can leave the scene with a functional vehicle tend to indicate minor damage, and a lower risk of severe injuries.

A number of parameters were also found to have positive coefficients, which act to indicate factors present in accidents with increased injury severity outcomes. A list and a brief explanation of the possible reasons for the sign are provided below.

- *Number of Vehicles*: As evidenced by the data, crashes involving more vehicles tend to cause more severe injuries.
- *Snow*: During snow-storms the risk of injury can increase because of slippery conditions as well as a lack of visibility can exist. This should not be confused with the Snow_Ice variable (above), which indicates snow which is already on the roadway.
- *Intersection*: Although roundabouts have less conflict points than typical stop and signal-controlled intersections, conflicts still exist which could increase the risk of injury over the base case (i.e. on the approaches).
- *NoImpact*: Accidents such as rollovers and vehicles that fall into a culvert (without necessarily impacting another vehicle) tend to suffer more severe injuries.
- *Bus*: Due to their size and lack of seat belts, accidents involving buses have a greater risk of severe injuries.

- *Dark*: Accidents occurring on unlit roadways during the night tend to cause more severe injuries. This can be due to a lack of visibility.

The elasticity effects for the model are presented in TABLE 10. These effects are calculated by setting the dummy variables (those with possible values of either 0 or 1) to their default value of 0, whereas the continuous variable is set to its mean value. The elasticity for a given variable is then calculated by changing the desired dummy variable to a value of 1, or by adding an increment to the mean value of the continuous variable, with all other variables staying the same (*61*). Since the continuous variable deals with integer values only, the increment is taken to be a 1 unit increase.

TABLE 10: Probability Elasticity of the Final Model.

	Severity Level		
Independent Variable	$P(Y_i=0)$	$P(Y_i=1)$	$P(Y_i=2)$
Vehicle	0.392	-0.838	-0.884
Number of Vehicles	-0.383	0.761	1.856
Intersection	-0.235	0.478	0.931
No Impact	-0.200	0.410	0.762
Bus	-0.496	0.956	2.884
Dark	-0.386	0.766	1.877
Animal	0.397	-0.849	-0.892
Hit & Run	0.147	-0.311	-0.400
Snow	-0.286	0.578	1.210
Snow & Ice	0.178	-0.377	-0.471
Probability at mean value (%)	68.2	30.0	1.8

Looking at TABLE 10, these values provide interesting results inasmuch that researchers can predict which factors are most strongly related to injury severity. Furthermore, this is beneficial with respect to the allocation of limited accident and injury prevention resources. Using models such as the one presented above, researchers will be better able to determine the true effects of a proposed countermeasure by controlling for the effects of other significant factors.

As can be seen in TABLE 10, by far the largest elasticity can be found to occur with a change in value of the *bus* predictor variable. According to the calculations, an accident involving a bus reduces the likelihood of a no injury accident by 1.5 times, and increases the risk of a severe injury accident by as much as 3.88 times.

It can also be seen that the magnitude of the elasticity effects increases significantly for the *major injury* category. This is due to the small probabilities of an accident resulting in severe injuries, and the introduction of one of the presented variables leads to larger changes relative to the small probabilities. This makes sense as a majority of accidents fall into the *no injury* category under base conditions. This result further indicates that it would take a considerable change in the input variables for an accident to fall into the *major injury category*.

3.5.2 Limitations

Several potential accident contribution factors were not examined in this research due to a number of limitations imposed by the accident data. In an effort to limit the sharing of vehicle occupants' personal information (age, gender, location in vehicle, injury severity, etc.), certain areas of the accident reports were censored from the authors, effectively blocking access to any data at the occupant-level for analysis and therefore limiting the study to the collision-level for analysis.

As is typical with accident data records, the limited number of recorded events at the higher severity levels tends to reduce the power of the models to identify salient accident factors. Issues have also been identified relating to the quality of record-keeping within the province, as a number of fields on the accident reports are being left blank. Because of the lack of detailed accident records, several interesting variables had to be dropped from the analysis.

3.6 CONCLUSIONS

This research has identified and explored a number of factors and their effects on the injury severity levels of accidents that occur in or around roundabouts in the Province of Quebec. An ordered logit model was successfully estimated for accident data reported from 37 roundabouts within the road network.

It was found that the factors that significantly influence injury severity outcomes include: the season, the number and type of vehicles involved in the crash, the type of impact, the road surface conditions and the weather at the time of the accident, whether a hit and run was observed, the lighting conditions at the time of the accident, and also whether the accident occurred on an approach to the roundabout or within it.

One of the more interesting results from the model is the fact that vehicle occupants are 49.6 % more likely to suffer an injury during an accident involving a bus (given an accident occurring under base-value conditions). Furthermore, accidents involving multiple vehicles as well as those occurring in the dark are approximately 39 % more likely to experience injuries. These findings help reinforce the idea that roundabout designs need to consider the needs of the road users. For example, the geometry must be able to safely accommodate larger vehicles, such as buses; the lighting design should ensure that all areas are well lit for visibility purposes; and finally sight distances should be far enough so as to ensure that drivers can safely avoid obstacles or other vehicles on the road, but not so large as to encourage faster driving speeds (29).

The analysis of injury severity models can help improve safety at roundabouts by considering conditions unique to roundabouts in the Province of Quebec. The contributing factors identified in this study were based on the available data for Quebec roundabouts. A major limitation of this study was the limited information that could reliably be extracted from the dataset.

Chapter 4 Crash Frequency Analysis

This chapter investigates the safety effectiveness of roundabouts and explores the role of geometry through a traditional crash frequency analysis using the historical data treated in the previous chapters. This chapter presents this topic through a before-after analysis and a complimentary negative binomial regression applied in the role of an exploratory analysis to try and identify factors with an influence on crash occurrence at a sample of roundabouts.

4.1 INTRODUCTION

As has been referred to throughout this thesis, the primary objective of this research is the investigation of the safety performance of roundabouts in the Province of Quebec. A first attempt at investigating this topic was undertaken in Chapter 3 and dealt with an injury severity analysis. This work focused on looking at crashes which already occurred at roundabouts.

With the crash record mapping presented in Chapter 2 having increased the available crash record database from 467,099 to 716,029 (or 94.0% of the total records) of useable crash records, an attempt was made to further investigate roundabout safety in the form of a crash frequency analysis. This analysis was conducted in two main parts.

First, a before-after analysis is presented in order to investigate the crash trends at roundabouts as of their construction date through a comparison with a given set of comparison sites. These sites are taken from neighboring intersections of both the stop and signal-controlled varieties.

Second, a complementary analysis that applies a negative binomial regression to the crash records treated in the crash frequency analysis is presented, with the objective of identifying features that influence the number of crashes that occur at roundabouts in the Province of Quebec.

4.2 LITERATURE REVIEW

An overall conclusion of the international literature is that important safety benefits can be expected from the conversion of traditional intersections to modern roundabouts. That being said, some exceptions for pedestrian and bicycle crash risk exist. A number of studies have similarly illustrated these benefits for conversions from stop-controlled intersections, but caution that results for conversions from signalized intersections are less conclusive (*18*). Crash frequency modeling has been used to illustrate the safety performance of roundabout intersections, as well as to investigate the link between geometry, traffic conditions and crash occurrence (*55*). In some studies, the results have been further divided by type of crash (*2*).

Safety studies of roundabout conversions are particularly well documented in Europe due to their history of implementation. Elvik (*54*) summarizes 28 studies for a total of 113 effect estimates of roundabout safety. It was found that results were inconsistent across different types of roundabouts, although it was generally agreed that small roundabouts have better performance. A study by Daniels et al. (*64*) investigated 91 roundabouts for vulnerable road users and found that an increase in accidents tends to occur. The American National Cooperative Highway Research Program's (NCHRP) *Report 672* (*2*) summarizes crash reduction findings from both American and international studies. The results are presented in TABLE 11.

TABLE 11: Summary of International Roundabout Safety Studies (*2*).

Country	**Mean Reduction (%)**	
	All Crashes	**Injury Crashes**
Australia	41-61%	45-87%
France	-	57-78%
Germany	36%	-
Netherlands	47%	-
United Kingdom	-	25-39%
United States	35%	76%

As can be seen in the table, a conclusion that can be seen across a majority of studies is that a mean reduction in crashes at roundabouts is achieved. It is important to note that studies often quote the average safety effects of roundabouts. As is evidenced by the differing mean reduction values for all crashes and injury crashes in TABLE 11, the safety effects of roundabouts varies across different crash severity levels, as well as across the various types and configurations of roundabouts (*54*).

Before comparing the results between different safety studies, it is important to note that not all result formats are directly comparable. Studies focusing on individual facility types can be difficult to compare with studies focusing on the analysis of different injury types.

More and more North American studies of roundabout crash frequency are being documented, and note several safety benefits. For instance, Retting et al. (*44*) studied the benefits associated with roundabouts. Using a dataset with 17 converted roundabouts on high-speed rural roadways, this research conducted a crash frequency analysis. Among the main findings, they showed that the average crash frequency (before and after the transformation) was reduced by 38%, average injury crash rate was reduced by 76%, and fatal crashes were reduced by an estimated 90% (*44*). These findings are similar to those reported in the Highway Safety Manual (*65*), which presents a crash reduction of 48% for signalized intersections converted into roundabouts, and a reduction of up to 82% for injury crashes (*66*).

In general, the existing safety studies can be classified into two distinct classes: those that control for regression-to-the-mean effects[1] and those that don't. An example of an analysis approach that fails to correct for these effects is the traditional before-after study. A before-after study is used to compare crash counts in the before period with crash counts in the after period, which is referred to as the naïve before-after approach. The change in the crash count is considered to be the treatment effect (i.e. the effect of having built a roundabout). In addition to failing to control for regression-to-the-mean effects, this method suffers from a number of shortcomings, including failure to account for the effect of time, and other unknown causal factors.

A more appropriate method is the before-after studies with comparison groups, which can help address these shortcomings by capturing the effects of the unknown causal factors in the comparison group and allowing for a better prediction of the expected crash count in the after period. This method still fails to account for regression-to-the-mean effects however. Given this limitation, the Empirical Bayes (also referred to the EB method) is favored in safety analysis applications as it considers regression-to-the-mean (*11*). An illustration of the general principle of the comparison site and EB methods is presented in FIGURE 15. As can be seen, unlike the comparison-group approach which is based on direct comparison of observed crash counts, the EB method applies a statistical regression approach in order to determine the influence of individual factors on crash counts. In this way, the EB method controls for the randomness

---

[1] Regression-to-the-mean effects arise when sites with unusually high crash counts are targeted for treatment. The sites tend to observe a drop in crash counts in the period after application of the treatment, regressing to their long-term mean count. Regression-to-the-mean has the potential to affect the validity of many safety improvement studies if not properly controlled for.

inherent in crash count data by determining a relationship between counts and specific prediction factors (*11*).

Throughout the existing roundabout literature, the EB method is used extensively. A study in Michigan, United States, analyzed safety using both simple before-after and EB methodologies (*67*). More recently, a study by Gross et al. (*18*) presented a reduction in crash rates at roundabouts through a historical analysis using an EB methodology.

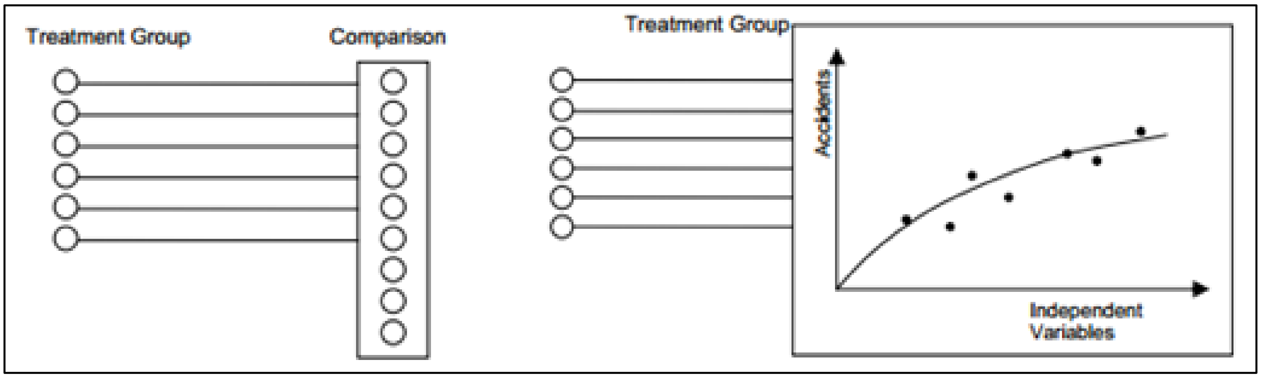

FIGURE 15: Illustration of before-after studies with comparison group (left) and EB method (right) (*11*).

Despite the importance of these previous efforts, few studies have investigated the influence of roundabout construction on safety in an environment with distinctive environmental and driver behavioral conditions as is found within the local and provincial road networks of the Province of Quebec, Canada. In fact, of the safety studies which mention roundabouts in Canada, a majority deal with the design considerations of roundabouts, as can be seen in (*68*) (*69*). Two studies based in the Province of Ontario, one dealing with the aspect of pedestrian safety in roundabouts by Henderson and Button (2013) and the other by Persaud et al (2010) dealing with the development of safety tools for roundabouts are part of the limited studies identified that deals with the observed safety performance of Canadian roundabouts (*70*).

In addition to the evaluation of the roundabout effectiveness, some research has been done to identify the contributing factors to crash frequency, although limited research applies directly to the Canadian road network. Keeping this in mind, a number of crash frequency regression models exist, and have been documented in numerous studies (*71*). As is accepted throughout the literature, crashes are both rare and random events and can exhibit considerable fluctuations in a given time period. Because of this, and their discrete, non-negative nature, typical regression modelling structures used in these studies include Poisson and Negative

Binomial regression (*72*). In this model, an assumption is made that negligible correlation exists with respect to the temporal variation between disaggregated crash observations (*73*).

Due to the weakness of Poisson regression in accounting for the overdispersion inherent in crash data, a Negative Binomial framework is considered to be better suited for the purposes of modelling the influence of various predictor factors on crash occurrence (*73*), as investigated in this work.

4.3 METHODOLOGY

The following section details the steps that were followed in order to investigate the crash frequency analysis on the Quebec crash record database. Although the use of the EB method that includes the development of safety performance functions is the most recommended approach in observational before-after studies, the available dataset suffered from some weaknesses that prevented their use (*74*). In particular, there was a lack of traffic exposure variables (both before and after construction) in the available data for this research. Without reliable traffic counts (from which ADDT[2] can be estimated), there is little basis for the prediction of crash rates. For this reason, this research focused on the use of a before-after study with comparison sites in order to try and account for time-trend effects throughout the study period or changes related to other external factors to the treatment (conversion of roundabouts). The author acknowledges the weakness associated with this methodology given the lack of traffic count data in the before and after period; therefore, the results should be interpreted accordingly.

4.3.1 Crash Data Preparation

Starting with the geocoded crash record database from Chapter 2, the next step was to filter the results to an acceptable level of accuracy. As discussed previously, an indication of the expected level of accuracy is provided by the online geocoding service. Following this step, a total of 716,029 records (or 94% of the total available records) were included in the geocoded database. According to a study by Ratcliffe (*35*), the effect of the loss of 6% of the available records is statistically insignificant and that the geocoded records should still be representative of the parent database.

---

[2] The AADT, or "average annual daily traffic" (*17*) is used in transportation engineering as a measure of traffic exposure on a given road facility.

In order to ensure that the most representative crash frequencies are investigated, the rejected crash records from the geocoding procedure were manually filtered in order to identify obvious roundabout crashes that were missed in the automated procedure. When records of this type were identified, they were added to the final database. Using a similar procedure, the database records were scanned for the inclusion of a field indicating the crash occurred outside of the road network, such as a parking lot. These records were removed from the dataset as they are not associated to the infrastructure being studied.

The next step of data preparation was to import the database into a geographic information system (such as ESRI's ArcGIS). This software package allowed for the spatial analysis of the crash reports in order to filter crashes that occurred within a roundabout's area of influence.

Initially, a maximum buffer with a radius of 100 m as measured from the center of each roundabout was applied. As roundabout infrastructure begins upstream of the circulating roadway on the approaches (such as reduction of speed limits, beginning of splitter islands, repeated reversed curves, etc.), the selected area of influence is larger than would typically be seen for a stop or signal-controlled intersection.

Accordingly, a custom buffer was developed for each roundabout. This new buffer considered the geometric design as well as the signalisation at the roundabouts to establish the most representative influence area, while maintaining the 100 m circular buffer as the maximum influence area of any given roundabout. On any given approach, the custom buffer is taken to be the smallest of the 100 m buffer or the furthest point along an approach in which roundabout infrastructure can still be found. This is done to distinguish between standard road infrastructure, and those aspects that are present because of the roundabout's existence. It is considered that any of these roundabout components could influence the observed safety at the intersection. An example of the custom buffer for the roundabouts in the municipality of Amos, Quebec is presented in FIGURE 16 below.

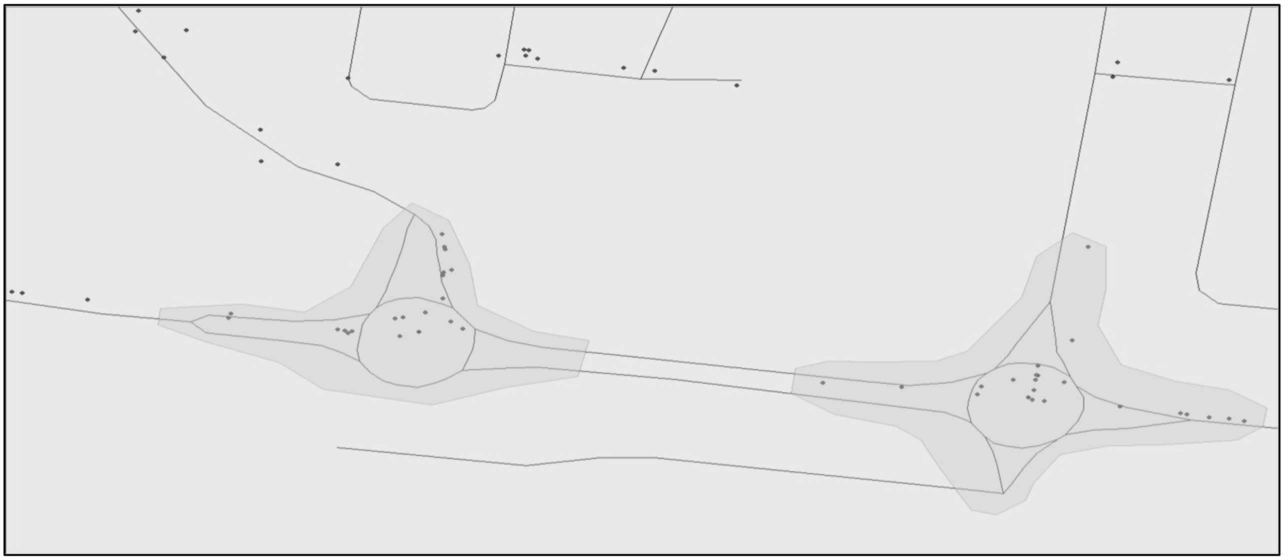

FIGURE 16: Example of custom buffers taking into account the intersection's area of influence.

Using these buffers as a reference, any crash record that falls within the buffer's area is logically considered to be within a roundabout's area of influence. Although this in no way suggests that these crashes occurred due to the existence of the roundabout, it leads to the possibility of such crashes having been influenced by the roundabout. Further analysis is required in order to hypothesize on what, if any, effects are due to the intersection being a roundabout instead of a more typical stop or signal-controlled variety.

Having isolated the crashes occurring at the 54 roundabouts for which adequate data was available, the next step in the methodology was to categorize the crash records according to whether they happened before, during, or after the construction of a given roundabout.

Due to uncertainties with respect to the duration of construction and the exact opening date of each roundabout, a range of +/– 6 months from the listed construction date was used. This range was selected in order to simplify the required assumptions for roundabouts for which only a construction year was provided. In these cases, the entire calendar year (from January 1st to December 31st of the same year) was considered to be the construction period. This period was selected in order to ensure that the data included in the analysis is taken from the sites after the opening of the roundabout to the public. As with the other sites, the roundabout facility is assumed to be in full operation at the end of the construction period.

In addition, the assumption of a range of dates for the construction period helped consider the adaptation period required for drivers to get used to a new facility; a period typically associated with higher crash frequency that could lead to false conclusions in a safety study. The reader is invited to explore a summary of the total crash frequency, construction date and earliest and latest available crash records presented in a table in Appendix B.

From here, it was necessary to further reduce the roundabouts considered for the before-after investigation since in a number of cases the crash data was not suitable for this type of analysis. Of note, sites with no available construction date, with crash records that do not occur in the time period of the roundabout's construction, or sites with either a before or after construction period of less than one year were dropped from the analysis as this information is required for further analysis. Due to the randomness and rarity of crash events, periods of less than a year are prone to large errors. A summary of the remaining sites is presented in TABLE 12.

4.3.2 Comparison Site Selection

As the name suggests, the central idea of a comparison site is to identify an untreated facility (for this study, an intersection that has not been converted to a roundabout) that approximates the features of a given roundabout. As such comparison sites are typically difficult to find and the available features may not correctly approximate those at the roundabout facility, a comparison group approach is used.

In this case, a group of comparison sites is selected to the best of the researcher's abilities within the same municipality as the roundabout. The goal here is that by averaging the crash rates throughout this comparison group, the effects of the different site facilities will compensate for each other, providing an indication of the trends affecting crash rates (such as environmental conditions, changing traffic patterns, etc.). Using the comparison groups as a basis for the analysis, it is possible to extrapolate the effects of a roundabout conversion on a specific site's safety.

The comparison site selection for this study was conducted by a research assistant from McGill University's Transportation Research Group. For the purpose of this study, the following factors were considered as a basis for site selection:

- Type of control
- Number of approaches
- Road classification
- Land use
- Physical characteristics (number of lanes, width, etc.)

TABLE 12: Crash record summary for individual roundabout sites.

Site ID	Name	Municipality	Total Records	Records Before Construction	Records During Construction	Records After Construction
2	Frechette/Brassard	Chambly	58	15	6	37
3	Bourgogne/deSalaberry	Chambly	47	12	2	33
4	FerACheval/JulesChoquet	Sainte-Julie	20	12	1	7
5	R132/deMontarville/Rene-Levesque	Boucherville	17	4	1	12
6	A10/A15	Montréal	3	1	1	1
8	IleDesSoeurs/duGolf	Montréal	65	17	6	42
12	Ryan/Duplesis	Mont-Tremblant	55	10	3	42
15	R327/Ryan	Mont-Tremblant	107	57	14	36
16	Principale/Desjardins	Sainte-Agathe-des-Monts	100	46	11	43
20	R105/Montcalm	Gatineau	77	16	4	57
21	R138/Sherbrooke	Montréal	19	9	1	9
23	R218/R277	Saint-Henri	39	16	7	16
24	R132/Cyr	New Richmond	58	47	3	8
25	Station/delaBaie	Shawinigan	137	37	17	83
26	R138/PaysBrule	Louiseville	139	63	14	62
29	R148/desLaurentides	Gatineau	48	22	5	21
30	PierreL/Bruce	Granby	52	29	5	18
33	R111/Principale (R395)	Amos	63	13	7	43
34	R111/4e	Amos	63	18	9	36
35	R117/3e	Val-d'Or	57	20	7	30
36	R117/R397	Val-d'Or	169	53	15	101
37	R117/Hydro	Val-d'Or	57	20	4	33
38	R111/R117	Val-d'Or	109	48	7	54
41	R138/Granier	Pointe-Lebel	89	36	8	45
42	R170/deQueen	Saint-Gédéon	53	41	7	5
43	A40/CitédesJeunes	Vaudreuil-Dorion	23	1	2	20
44	R172/Roussel/du pont	Saguenay	87	81	2	4
45	R173/277	Saint-Henri	37	19	7	11
74	Talbot/Jacques-Cartier	Saguenay	83	27	10	46
75	duFoyer/Jacques-Cartier	Saguenay	27	1	4	22
86	desSousBois/delaForet	Mont-Tremblant	7	4	2	1
87	desSousBois	Mont-Tremblant	2	1	0	1
88	R138/Royale	Trois-Rivières	90	38	15	37
89	DesRecollets/Laviolette	Trois-Rivières	288	100	30	158
95	MarcA/Mousseau	Saguenay	16	14	1	1
96	SaintEmilie/SaintDenis	Saguenay	13	9	0	4
144	R362/Rang Terrebonne	Saint-Irenee	15	11	2	2
		Total	2389	968	240	1181

Although not explicitly considered to be a factor in site selection, comparison sites where the difference in the crash frequency in the before period varied by several orders of magnitude from that observed at roundabout sites, the comparison site was dropped from the analysis. This was done since a large difference in this value may indicate other incompatibilities between the two sites.

TABLE 13 illustrates a possible comparison site selected for the two roundabouts in Amos, Quebec. As can be seen, the sites approximate the features seen in the associated roundabout, but differences arising from the difficulty in exactly matching geometry, land use and traffic patterns exist. The difficulty in finding exact matches is amplified by the particular applications in which roundabouts are implemented. These include at entrances to a municipality, at highway interchanges, and at the crossing of two or more superior-class roadways. For this reason, similar sites are difficult to find in the smaller, more rural municipalities in which they are often constructed.

In order to populate the comparison group, a total of 10 comparison sites were initially selected. This was done so as to have a larger set of comparison sites in which to select the final comparison group and to improve the accuracy of the estimates. With the difficulty in finding appropriate comparison sites to populate the various comparison groups, however, a number of sites had to be dropped from the analysis as the analysis progressed. Reasons for the exclusion of sites include too small of a before or after period, an observed crash frequency that is several orders of magnitude larger than the other comparison sites, among other considerations. In all cases, a minimum of 2 comparison sites was used for each roundabout site, with a majority of the sites having between 4 and 6 comparison sites retained for the analysis.

TABLE 13: Example of comparison site selection for R111 and Rue Principale (top left) and R111 and 4e Rue (bottom left) for Amos, Quebec.

Roundabouts	**Comparison Sites**
	
	

Source: Personal Correspondence, 2013 **Source: Google Map**

4.3.3 Before-After Crash Data Analysis

The before-after analysis is applied by bringing together all of the work that was done throughout the course of this research. Using the before and after periods of the individual roundabout sites, the crash frequency at each comparison site is identified for identical periods. The intention of using this methodology is to estimate the expected change in crash frequency of an improvement while attempting to control for the effect of time on the crash record trends (*75*).

For this study, a correction factor is applied due to the varying time periods available throughout the data. From here, it is possible to compare the frequencies in the before and after periods for each treatment and comparison group.

As previously mentioned, the available data covers a period ranging from the year 2000 to 2011. Due to differences in construction dates, this leads to varying before and after time periods for the analysis. Although longer periods of data are generally preferred, results can still be extracted for sites with shorter time periods. As a minimum, a period of two years was considered in all cases to try and account for unusually high crash rates in any given year. TABLE 14 presents a summary of the before-after study period durations.

TABLE 14: Summary of before-after study periods.

Measure	Before Construction Period (years)	After Construction Period (years)
Mean	4.5	5.3
Standard Deviation	1.4	1.3
Minimum	2.0	3.0
Maximum	7.0	7.5

Although as previously mentioned no exposure data is available in this study, it is possible to hypothesize that a roundabout conversion is expected to serve at least the same level of traffic as was expected before the conversion, with a growth in traffic likely with the assumption that no major transportation network changes occur in the area during the study period. With this assumption, it is possible to comment on the crash frequency trends of the site under analysis.

Using the identified comparison groups and a before-after observational approach, the estimated number of crashes that would have occurred if no treatment were applied to a given intersection is defined as follows (*75*) (*76*):

$$\boldsymbol{\hat{\pi} = \hat{r}_c r_d K} \qquad \textbf{EQ. 4}$$

where:

$\hat{r}_c = \frac{N/M}{1+1/M}$ is the ratio of the number of crashes in the after to before periods for the comparison group;

$r_d = \frac{ta}{tb}$ is the ratio of the after to the before periods which is used to adjust the data for differences in the observation periods;

K, M = the number of crashes in the before period of a treated site and comparison group, respectively;

N = the number of crashes in the after period of a comparison group (L would be used in the case of a treated site).

In other words, the expected number of crashes at a given site had a roundabout not been constructed (and assuming constant traffic flow) is predicted by multiplying the observed number of crashes in the before period by the ratio of the number of crashes in the after period to the number of crashes in the before period for the comparison sites. The estimated difference in the number of crashes ($\hat{\delta}$) due to the construction of a roundabout (i.e. the treatment) at a given site is then calculated as follows:

$$\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\pi}} - \hat{\boldsymbol{\lambda}} \qquad \textbf{EQ. 5}$$

where: $\hat{\lambda}$ is the number of crashes after the application of the treatment.

In this form, if EQ.5 produces a positive value, there is an estimated reduction in the number of crashes in the after period; otherwise, there is an estimated increase in the number of crashes. Similarly, we can define $\hat{\theta}$ as the index of effectiveness, which is given by the formula:

$$\hat{\boldsymbol{\theta}} = \frac{\hat{\boldsymbol{\lambda}}}{\hat{\boldsymbol{\pi}}} \qquad \textbf{EQ. 6}$$

The index of effectiveness is the ratio of what the observed safety is with a treatment (or roundabout in this case), to the estimate of what safety could have been had the treatment not been applied. For a treatment to be considered effective, a value of $\hat{\theta}$ <1 is required. From here, it is possible to calculate the percentage reduction in the expected crash frequency, which is defined as (*76*):

$$\%\, Reduction = 100 * (1 - \theta) \qquad \textbf{EQ. 7}$$

As illustrated by Hauer (1997), however, the θ parameter is a biased estimate, which can be corrected by a simple adjustment as shown in Equation 8 (*76*):

$$\hat{\boldsymbol{\theta}} = \frac{\boldsymbol{\lambda}/\boldsymbol{\pi}}{\left(\mathbf{1} + \frac{\boldsymbol{Var}(\hat{\boldsymbol{\pi}})}{\boldsymbol{\pi}^{\mathbf{2}}}\right)} \qquad \textbf{EQ. 8}$$

Accordingly, the variance of π and θ can be calculated as follows:

$$\boldsymbol{Var}\left(\hat{\boldsymbol{\theta}}\right) = \boldsymbol{\theta}^{\mathbf{2}}\left(\frac{\left(\frac{\boldsymbol{Var}(\hat{\boldsymbol{\lambda}})}{\boldsymbol{\lambda}^{\mathbf{2}}} + \frac{\boldsymbol{Var}(\hat{\boldsymbol{\pi}})}{\boldsymbol{\pi}^{\mathbf{2}}}\right)}{\left(\mathbf{1} + \frac{\boldsymbol{Var}(\hat{\boldsymbol{\pi}})}{\boldsymbol{\pi}^{\mathbf{2}}}\right)^{\mathbf{2}}}\right) \qquad \textbf{EQ. 9}$$

$$\boldsymbol{Var(\widehat{\pi}) = \widehat{\pi}^2\left(\frac{1}{K}+\frac{1}{M}+\frac{1}{N}+var(\omega)\right)} \quad \textbf{EQ. 10}$$

Where K, M and N are as previously defined. The value var(ω) is a modification factor used to account for non-ideal comparison groups. Whereas an ideal comparison group will have yearly crash trends which are identical to what is observed in the treatment group, this is not always the case. That being said, it is accepted that the value of the modification factor is generally small. Gross et al. (2010) recommend that the variance can be estimated without the modification factor, with the acknowledgement that the estimate is conservatively low as a consequence (*77*). Consequently, a simplified formula is presented in Equation 11:

$$\boldsymbol{Var(\widehat{\pi}) = \widehat{\pi}^2\left(\frac{1}{K}+\frac{1}{M}+\frac{1}{N}\right)} \quad \textbf{EQ. 11}$$

In order to aid in the interpretation of the results, safety effectiveness calculations are aggregated to both a cluster and overall roundabout effect level. These calculations are performed according to the estimation of pooled observations as presented by Hauer (1997) (*76*).

4.3.4 Negative Binomial Modelling of Design Parameter Influence

In order to try and gain a better understanding of the contributing factors of safety at Quebec roundabouts, an exploratory analysis is performed to identify factors linked to crash occurrence. The variables are limited to those representing type, setting and basic geometric factors of the various roundabouts, as presented in TABLE 15. The studied variables are limited in part by the availability of the information in the dataset, as well as by those considered in other studies such as (*20*), (*78*), (*18*), among others. Both continuous and dummy variables are considered. Note that more detailed variables were possible, but according to the available sample size and the lack of vehicular flow data, the added level of detail was considered to be unjustified, and the general categories listed below were employed.

The dependent variable in this case is taken to be the observed number of crashes that occurred at a given roundabout, with the scope limited to crashes occurring in the period after a roundabout's construction. Unlike the before-after analysis, this model does not consider crash records that occurred in the before construction period. Because of this, sites that were previously dropped from the scope of the analysis due to missing or inappropriate before period data could be brought back into the analysis, resulting in a larger sample of roundabouts.

TABLE 15: List of variable categories considered for Negative Binomial regression.

Dummy Variables	Legend
Central Island Landscaping	0 = Landscaping, 1 = Built/Structures
Median Island Type	0 = Concrete, 1 = Painted
Land Use	0 = Natural, 1 = Built
Presence of a Truck Apron	0 = No, 1 = Yes
Type of Roundabout (Modern or converted)	0 = Modern, 1 = Converted
Presence of Pedestrian or cyclist facilities	0 = No, 1 = Yes
Presence of slip lanes	0 = No, 1 = Yes
Proximity of accesses	0 = No, 1 = Yes
Clustering of Roundabouts in a municipality	0 = No, 1 = Yes
Presence of a stop area on an approach	0 = No, 1 = Yes
Presence of a pedestrian traffic signal	0 = No, 1 = Yes
Location of roundabout on the MTQ network	0 = No, 1 = Yes
Approach Spacing	0 = 90 degrees, 1 = Other
Approach Configuration (tangential or perpendicular)	0 = Perpendicular, 1 = Tangential
Use of successive curves along approaches	0 = No, 1 = Yes
Continuous Variables	
Posted Speed Limit	Speed in km/hr
Circle Speed Limit	Speed in km/hr
Closest intersect along an approach	Distance in meters
Roundabout Diameter	Diameter in meters
Number of approach lanes	unit
Number of circulating lanes	unit

As previously mentioned, a Negative Binomial (NB) model is used in order to try and identify a set of predictor variables as a complement to the crash frequency analysis described above.

The general form of the Poisson model of which the NB model is as shown in Equation 10:

$$\boldsymbol{\mu_i = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_n X_{in})} \quad \textbf{EQ. 12}$$

Where:

$X_{i1}, X_{i2}, \ldots, X_{in}$ represent the values of the factors considered in the model;

$\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ represent a vector of parameters to be estimated from the regression (*79*).

With this formulation, it is assumed that the number of crashes at a given site i, Y_i, follows a Poisson distribution with a mean μ_i. In the NB model formulation, however, a Gamma random effect is introduced to μ_i which accounts for the variance and mean not being equal.

Consequently, the mean number of crashes is considered to be randomly distributed and can be represented by μ_{ε_i} as illustrated in Equation 11:

$$\mu_{\varepsilon_i} = \mu e^{\varepsilon_i} \qquad \textbf{EQ. 13}$$

Where:

μ_i is as previously described;

e^{ε_i} represents a multiplicative random effect.

Taking this into account and considering that the number of crashes at a given site *i*, Y_i, remains Poisson distributed, the probability function can be parameterized as follows (*79*):

$$\mathrm{P}(Y_i = y_i; \alpha) = \frac{\Gamma(y_i+\alpha^{-1})}{y_i!\Gamma(\alpha^{-1})}\left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i} \qquad \textbf{EQ. 14}$$

Where:

α is a parameter representing the overdispersion inherent in negative binomial regression.

It is interesting to note that as α approaches 0, the negative binomial regression is simply a Poisson regression. The negative binomial regression was applied to the crash record dataset using Stata data analysis and statistical software (*80*). Due to the different construction dates of the various roundabouts, the periods of available data vary for each site. This must be controlled for in the model, as a greater exposure period exists in some cases, leading to the possibility of more crashes occurring.

Unlike other statistical analysis packages, Stata allows users to directly control for this bias through the inclusion of an *offset* option in its regression parameters. An inherent risk in this approach is that the exposure has nonlinear effects on crash frequency. For the purposes of this analysis, these effects are assumed to be negligible.

The *offset* option is applied to the logarithmic value of the exposure variable (for the roundabout crash analysis, this exposure is the period of available crash records in the period after roundabout construction). The general form of the negative binomial regression implemented with the *offset* variable in the Stata software can be illustrated as follows:

$$nbreg\ dependant\_var\ var_1\ var_2 \ldots var_N, offset(\ln(exposure))$$

Unlike a typical model variable, the coefficient of the *offset* variable is not determined by the model estimation, but is instead fixed to the value of '1' by Stata. This is done as the variable's main purpose is to directly control for the effects of a longer exposure (or analysis) period for various observations.

4.4 RESULTS

As described in the methodology section, it is important to keep in mind that the research presented in this chapter applies a before-after analysis with comparison groups, which suffers primarily from its lack of consideration for regression-to-the-mean effects. The results presented below were observed throughout the data, however without consideration of the regression-to-the-mean effects and other factors such as traffic exposure (which is assumed to be constant in this research due to the lack of this information in the dataset), the changes cannot be definitively attributed to the conversion of the intersections into roundabouts. Because of this, the reader is cautioned when drawing conclusions from these results.

That being said, without talking about a definitive measure of safety, this research serves as an attempt to lay the groundwork for future studies and to identify patterns that merit future analysis given the possibility of obtaining richer datasets.

The results of the crash frequency analysis were divided into two main sections. They are discussed below.

4.4.1 Before-After Analysis

After having processed the data and removed any sites that did not have sufficient information to properly apply the before-after analysis, a total of 25 roundabout sites remained. Due to similarities in the design and implementation of the roundabouts, a grouping of the sites was performed in order to try and isolate general trends in their crash frequencies. The retained clusters are presented in TABLE 16 below.

TABLE 16: Cluster descriptions.

Cluster Number	Description
1	Roundabouts located at highway interchanges.
2	Smaller, single lane roundabouts typically located in residential sectors.
3	Larger, multi-lane roundabouts on local roads.
4	A mix of roundabout types and lane configurations on main roads and regional highways. High speeds.
5	Roundabouts converted from traffic circles; large diameter, multilane and tangential approaches.
6	Roundabouts located at the entrance to a municipality or subdivision.

Although manually selected through direct observation of the individual design characteristics, these cluster groupings demonstrate agreement with those identified by St-Aubin et al. using k-means clustering for a similar dataset (*81*).

The results of the before-after analysis after having applied the above groupings to the 25 studied roundabouts are summarized in TABLE 17 below. Readers who are interested in a more detailed overview of the crash record data used to perform the before-after analysis with comparison sites are invited to consult Appendix C.

Taking a general look at the results in TABLE 17, it can be seen that the crash experience varies substantially from site to site. An interesting observation is that of the 25 total sites, 13 roundabouts (or 52.0 % of the sites analyzed) are observed through the θ parameter to experience an increase in the number of crashes that occurred in the after period with respect to the expected level obtained through the comparison analysis. Specifically, two clusters demonstrate on average some safety benefits (clusters 2 and 4); however, the rest of the clusters show the opposite – on average, safety deterioration is observed. It is important to highlight that the statistical confidence intervals related to these estimations are very large. In a majority of cases, the lower limit of a 95% confidence interval lies below 1.0, and the upper limit lies about 1.0. Since the value of 1.0 is within the confidence interval, it is not possible to eliminate the possibility that the safety effect can take a value of 1.0, which indicates no observable effects occurred.

Considering the unbiased estimator of θ and the associated 95% confidence interval presented in TABLE 17, it is not possible to exclude the possibility that roundabouts have no effect on intersection safety (i.e. crash frequency statistically remains the same in the before and after periods). This statistical insignificance can be associated to different factors such as underreporting, mis-location of crashes, regression-to-the-mean effects (as previously explained), and the small sample sizes used in the analysis.

Looking at the trends across the identified clusters, however, observations can still be made. Although some exceptions do exist, the roundabouts assigned to a given cluster exhibit similar safety trends. For example, all of the roundabouts assigned to cluster number 4 experience a reduction in the crash frequency observed in the period after their construction, using the comparison group as a point of reference.

Despite the variability of the results across sites and clusters, it is possible to extrapolate interesting observations from these results. As can be seen in clusters 1 and 6, roundabouts in locations at which different types of drivers are forced to interact such at highway interchanges, or roundabouts located at the entrance to a municipality with a considerable speed differential and polarized flows experience increased crash frequency. Similarly, as was discussed in the literature (*5*) (*19*) (*2*), cluster 3 reinforces the notion that multi-lane roundabouts are typically less safe than other design options, as can be seen by the apparent increase in the crash frequency of all three analyzed sites.

Finally, both of the roundabouts from the municipality of Trois-Rivières assigned to cluster number 5 experience an increase in the crash frequency observed in the period after their conversion to modern roundabouts. Whereas a study by Gates & Maki (2000) explored the topic of converting old traffic circles to modern roundabouts and found that good performance can be obtained by such conversions, the statistical significance of the result from cluster number 5 seems to indicate the opposite. A possible explanation, as illustrated by Gates & Maki, is that adequate measures need to be taken to address the geometric differences between traffic circles and modern roundabouts (*82*). In the case of the Trois-Rivières roundabouts, the geometry retains many of the traits of the old traffic circle design. With the cluster sample of only two sites, however, it is difficult to conclude on the effectiveness of this type of facility.

TABLE 17: Cluster grouping of before-after analysis results.

Cluster	Municipality	ID	Number of Crashes in the After Period (λ)	Expected Number of Crashes in After Period if Roundabout not Constructed (π)	Change in Expected Crashes in the After Period ($\hat{\delta}$)	θ	Var(θ)	95 % Confidence Interval	
								Lower Bound	Upper Bound
1	Boucherville	5	12	11.3	-0.7	0.583	0.093	-0.014	2.357
	Louiseville	26	62	31.5	-30.5	1.772	0.323	0.659	2.851
	Overall Effect					1.547	0.251	0.566	2.711
2	Chambly	3	33	34.6	1.6	0.758	0.104	0.124	2.388
	Saguenay	74	46	46.1	0.1	0.930	0.071	0.408	2.297
	Mont-Tremblant	86	1	13.1	12.1	0.052	0.002	-0.032	2.005
	Saguenay	96	4	9.7	5.7	0.333	0.035	-0.035	2.183
	Overall Effect					0.770	0.035	0.406	2.180
3	Chambly	2	37	33.0	-4.0	0.867	0.144	0.124	2.483
	Montréal	8	42	35.0	-7.0	1.072	0.132	0.360	2.455
	Gatineau	20	57	36.7	-20.3	1.295	0.253	0.309	2.716
	Overall Effect					1.217	0.097	0.608	2.368
4	Sainte-Julie	4	7	9.2	2.2	0.718	0.091	0.125	2.354
	Mont-Tremblant	12	42	48.2	6.2	0.600	0.081	0.041	2.326
	Mont-Tremblant	15	36	101.0	65.0	0.340	0.008	0.166	2.057
	Shawinigan	25	83	69.1	-13.9	1.122	0.090	0.535	2.350
	Gatineau	29	21	29.4	8.4	0.674	0.043	0.266	2.212
	Val-d'Or	36	101	169.4	68.4	0.545	0.026	0.230	2.147
	Val-d'Or	37	33	38.0	5.0	0.781	0.070	0.263	2.294
	Val-d'Or	38	54	65.4	11.4	0.799	0.031	0.454	2.167
	Overall Effect					0.699	0.010	0.508	2.067
5	Trois-Rivières	88	37	32.4	-4.6	1.113	0.061	0.629	2.268
	Trois-Rivières	89	158	81.0	-77.0	1.932	0.058	1.458	2.260
	Overall Effect					1.707	0.035	1.342	2.181
6	Saint-Henri	45	11	9.0	-2.0	1.189	0.162	0.399	2.526
	Montréal	21	9	9.3	0.3	0.887	0.133	0.172	2.458
	Saint-Henri	23	16	11.1	-4.9	1.368	0.201	0.489	2.610
	Amos	33	43	24.7	-18.3	1.446	0.330	0.321	2.864
	Amos	34	36	34.9	-1.1	0.899	0.108	0.256	2.396
	Val-d'Or	35	30	27.2	-2.8	1.031	0.097	0.420	2.369
	Overall Effect					1.216	0.049	0.784	2.229

Conversely, as can be seen in clusters 2 and 4, roundabouts serving mainly a single type of traffic (either local or regional in nature) with little or no change in speed limits in the areas around the roundabouts seem to exhibit on average a reduction in crash frequency in the after construction period when compared to the predicted value.

This result for clusters 2 and 4 is as expected, as the existing literature illustrates many examples of roundabouts constructed in urban and rural settings with increased safety performance being observed (*2*) (*18*).

The above observations notwithstanding, the large confidence intervals calculated for the data demonstrate that the safety effectiveness calculations do not have enough statistical power to draw solid conclusions. The addition of more sites for each cluster (an increase of sample size) would be necessary in order to improve the reliability of these results.

To measure the general effectiveness of roundabout intersections, it is possible to consider a pooled dataset. The larger sample size of the pooled dataset reduces the variance associated with the individual observations, increasing the likelihood of identifying statistically significant effects. The results of the pooled analysis of the available dataset are presented in TABLE 18 below. From the totals of the presented data, $\hat{\boldsymbol{\theta}}$ (or the treatment effect variable) can be calculated to be 1.01 with a variance of 0.008 (as calculated by Eq. 7 & 8 above with consideration of the differences between the before and after periods using the r_d parameter). Applying a 95% confidence interval, θ is found to lie between 0.8 and 1.2. As before, unfortunately this value is still statistically insignificant, as it is impossible to exclude a value of 1.0 from θ.

TABLE 18: Pooled Site Analysis Data.

				Treatment Sites			Comparison Sites							
Municipality	Site ID	Before Years	After Years	Before Crashes	Adjusted Before Crashes*	After Crashes	Before Crashes	Adjusted Before Crashes*	After Crashes	R_d	λ	R_c	π	Var(π)
					K	L		M	N					
Chambly	2	3.0	7.0	15	35.0	37	35.0	81.7	78.0	2.3	37.0	0.9	33.0	318.5
Chambly	3	4.0	6.0	12	18.0	33	17.0	25.5	51.0	1.5	33.0	1.9	34.6	308.8
Sainte-Julie	4	7.0	3.0	12	5.1	7	34.0	14.6	28.0	0.4	7.0	1.8	9.2	4.7
Boucherville	5	2.0	5.0	4	10.0	12	24.0	60.0	69.0	2.5	12.0	1.1	11.3	104.9
Montréal	8	4.0	6.0	17	25.5	42	81.0	121.5	168.0	1.5	42.0	1.4	35.0	146.9
Mont-Tremblant	12	2.5	7.5	10	30.0	42	32.0	96.0	156.0	3.0	42.0	1.6	48.2	1050.9
Mont-Tremblant	15	4.5	5.5	57	69.7	36	81.0	99.0	145.0	1.2	36.0	1.5	101.0	477.9
Gatineau	20	3.5	6.5	16	29.7	57	40.0	74.3	93.0	1.9	57.0	1.2	36.7	268.9
Montréal	21	6.0	4.0	9	6.0	9	70.0	46.7	74.0	0.7	9.0	1.6	9.3	7.8
Saint-Henri	23	6.0	4.0	16	10.7	16	81.0	54.0	57.0	0.7	16.0	1.0	11.1	7.1
Shawinigan	25	3.5	6.5	37	68.7	83	190.0	352.9	356.0	1.9	83.0	1.0	69.1	332.9
Louiseville	26	5.5	5.5	63	63.0	62	31.0	31.0	16.0	1.0	62.0	0.5	31.5	109.8
Gatineau	29	5.0	5.0	22	22.0	21	124.0	124.0	167.0	1.0	21.0	1.3	29.4	51.4
Amos	33	3.5	6.5	13	24.1	43	58.0	107.7	111.0	1.9	43.0	1.0	24.7	125.1
Amos	34	3.5	6.5	18	33.4	36	81.0	150.4	158.0	1.9	36.0	1.0	34.9	180.0
Val-d'Or	35	5.0	5.0	20	20.0	30	80.0	80.0	110.0	1.0	30.0	1.4	27.2	52.8
Val-d'Or	36	3.0	7.0	53	123.7	101	80.0	186.7	257.0	2.3	101.0	1.4	169.4	2706.8
Val-d'Or	37	4.0	6.0	20	30.0	33	74.0	111.0	142.0	1.5	33.0	1.3	38.0	160.8
Val-d'Or	38	5.0	5.0	48	48.0	54	137.0	137.0	188.0	1.0	54.0	1.4	65.4	143.0
Vaudreuil-Dorion	43	6.5	3.5	1	0.5	20	76.0	40.9	74.0	0.5	20.0	1.8	1.0	0.5
Saint-Henri	45	7.0	3.0	19	8.1	11	93.0	39.9	45.0	0.4	11.0	1.1	9.0	2.5
Saguenay	74	4.0	6.0	27	40.5	46	165.0	247.5	283.0	1.5	46.0	1.1	46.1	154.4
Mont-Tremblant	86	4.0	4.0	4	4.0	1	6.0	6.0	23.0	1.0	1.0	3.3	13.1	79.5
Trois-Rivières	88	6.0	4.0	38	25.3	37	166.0	110.7	143.0	0.7	37.0	1.3	32.4	26.0
Trois-Rivières	89	6.0	4.0	100	66.7	158	412.0	274.7	335.0	0.7	158.0	1.2	81.0	63.1
Saguenay	96	4.0	5.0	9	11.3	4	27.0	33.8	30.0	1.3	4.0	0.9	9.7	22.4
Total					**829.1**	**1031.0**					**1031.0**		**1011.4**	**6907.3**

Pooled Results	θ	Var(θ)	95% Confidence interval	
			Lower Bound	Upper Bound
	1.012	0.00781	0.8	1.2

*The adjusted before crash counts are obtained by applying the R_d factor in order to correct for differences in the observation length between the before and after periods.

Even if the pooled θ was found to be statistically significant, the validity of the result could be questioned due to the assumption of the safety effect of roundabouts being the same in all applications. From the observations made in TABLE 17, there is evidence of the safety of roundabouts varying according to their settings and configurations. Although safety studies presenting aggregated safety results for roundabouts are plentiful in the literature, studies considering the types of roundabouts and their settings have been rarely reported in the literature; a fact that limits the possibility of comparing these results across other locations or jurisdictions.

To summarise, evidence can be seen throughout the before-after analysis that similarly to the findings of Elvik (2003) (*54*) as well as the crash modification factors provided in the Highway Safety Manual (*65*), there is evidence that the safety performance of roundabouts on the Quebec road network is inconsistent, and largely dependent on the type of roundabout and the setting that it is constructed in. This is in contrast to what has generally been found in other North American roundabout safety studies such as (*2*) (*44*), which demonstrate safety benefits for all roundabouts.

It is observed that small roundabouts, and roundabouts constructed in an environment serving a single type of traffic tend to demonstrate safety improvements after their construction. On the other hand, larger multi-lane roundabouts and roundabouts used in special applications such as in highway interchanges and at the entrance to a municipality or residential development tend to demonstrate reduced safety with respect to an observed increase in crash frequency after their construction. Further research is necessary to validate these observations, however, as the results obtained through this methodology have been found to be statistically insignificant.

4.4.2 Negative Binomial Modelling Results

Keeping in mind that the available dataset did not include any vehicular flow data, the results for the few variables that were found to be significant in the exploratory analysis of the roundabout data are presented in TABLE 19 below. While trying to obtain the strongest possible regression model for the dataset, as previously listed in TABLE 15, variables relating to land use, roundabout geometry and design features were all tested. For the purpose of clarity, the variable *Approach Configuration* was a dummy variable created to indicate whether the roundabout's had tangential or perpendicular approaches with respect to the circulation roadway. The variable *Presence of Stop Area* is a dummy variable created to indicate if a stop area was located on any of the approaches, such as bus stops, passenger drop-off areas, etc.

Furthermore, it is possible to demonstrate that at a fundamental level, some variables (such as the number of lanes, for example) show strong correlation with vehicle count data, and can be considered as proxies for the missing variables.

TABLE 19: Negative Binomial regression results.

Log-likelihood	-168.01				
Variable	**Category**	**Coefficient**	**Standard Error**	**Significance**	**Elasticity**
Presence of Truck Apron	Dummy*	2.194	1.231	0.075	0.889
Maximum Posted Speed on Approach	Continuous	0.021	0.008	0.011	1.188
Number of Circulating Lanes	Continuous	0.606	0.284	0.033	0.454
Approach Configuration	Dummy*	2.467	1.350	0.068	0.915
Presence of Stop Areas	Dummy*	1.012	0.329	0.002	0.636
Intercept	-	-2.535	1.291	0.049	-
After Construction Period Duration	Continuous	Exposure	-	-	-
ln(alpha)	-	-0.783	0.237	-	-
alpha	-	0.457	0.108	-	-

*Dummy variables can take the value of 0 or 1.

Performing a negative binomial regression against the number of accidents occurring at the 36 roundabouts for which crash data was available in the after construction period, it is observed that at the 90% significance level, the retained explanatory variables are all related to the design of a given site.

Looking at the retained variables, there appears to be further evidence that larger roundabouts suffer from more crashes, and hence reduced safety compared to their smaller counterparts. This is inferred by the *Presence of Truck Apron, Maximum Posted Speed on Approach* and *Number of Circulating Lanes* variables which are typically associated with larger roundabouts. A possible explanation for the reason these variables were found to be significant is because of the fact that factors associated with larger roundabouts are correlated to higher traffic count values. Without having the data to validate this, however, it is also possible that confounding factors exist.

In addition to the variables discussed in the previous paragraph, it can also be seen in TABLE 19 that roundabouts with tangential approaches (as often found in older traffic circles or conversions), and roundabouts with stopping areas built into their approaches both show the same trend. In the case of the tangential approach variable, this is often attributed to the lack of proper deflection that is achieved with this configuration (*82*). This result agrees with the observed results from cluster 5 in the before-after analysis presented above, where the two traffic circle conversions in Trois-Rivières, Québec are found to have reduced safety performance. It should again be noted, however, that a small sample makes it difficult to conclude on the actual effectiveness and its cause. The positive sign on the coefficient of all of the variables indicates than in each case, an increase in the variable's value will result in a larger prediction of the crash frequency.

Furthermore, the regression modeling provides a likelihood-ratio test which provides strong evidence that the null hypothesis of $\alpha = 0$ should be rejected. This validates the use of the negative binomial regresion model over standard Poisson as the data exhibits overdispersion.

Similarly to the results of the before-after analysis presented previously, the evidence suggests that on average, smaller, single-lane roundabouts are safer than their larger counterparts. However, statistically, the results are not conclusive given the large confidence intervals obtained in the before-after analysis.

4.5 CONCLUSIONS

This chapter proposes a methodology to investigate roundabout safety through a crash frequency analysis. Two complimentary approaches were considered. The first was a before-after analysis with comparison group in order to look at the effectiveness of roundabouts in the province, and the second was through an attempt at identifying crash prediction variables through the application of a negative binomial modeling approach.

The two procedures were investigated through a subsample of the 54 roundabouts for which crash records were available for study. Although further investigations using the stronger Empirical Bayes method are suggested before drawing final conclusions, a number of interesting observations can be made. From the available evidence, it can be seen that in its current form, the crash frequency analysis of a sample of Quebec roundabouts suggests that safety varies across sites, although similarities can be observed between sites of a same subgrouping.

While based on the confidence intervals the results are not statistically significant, observations suggest that small roundabouts, and roundabouts constructed in an environment serving a single type of traffic tend to demonstrate, on average, a safety improvement after their construction. Similarly, larger multi-lane roundabouts and roundabouts used in special applications such as in highway interchanges and at the entrance to a municipality or residential development show on average an increase in crash frequency.

As mentioned previously, caution must be used when interpreting these results, as a number of unobserved factors exist. These include regression-to-the-mean effects, as well as factors such as the unavailable traffic flow variables. As is recognized throughout much of the literature, a link exists between traffic flow and crash frequency (*76*). Facilities with large traffic flows have greater chances for crashes to occur. With this in mind, it is not possible to isolate the effect of increased traffic flow on the safety performance of large roundabouts.

Furthermore, a limitation of the study is tied to the fact that the crash frequencies were not classified by type. Because of this, it is impossible to observe whether a specific type of crash exhibited different trends than the crash totals as a whole. Finally, future work should consider the possibility of an adaptation period affecting the safety trends observed in the data. This extension of the work would require a richer dataset, however, with a greater time period of data in the after construction period.

Due to the large variance related to a site-by-site analysis approach, the data was pooled to investigate the overall safety performance of roundabouts. Using this approach, a treatment effect (θ) of 1.01 with a variance of 0.008 was obtained. Applying a 95% confidence interval, θ was found to lie between 0.8 and 1.2. Although the variance was reduced, the results of the overall effect were also found to be statistically insignificant. Furthermore, a weakness inherent overall effectiveness methodology is that the general effect of roundabout conversions is the same across all applications, which contradicts the trends observed in the first part of the analysis – showing important variability across clusters (subgroups of treated sites).

The negative binomial regression identified a number of statistically significant variables related to the prediction of roundabout crash occurrence, with evidence of factors related to larger roundabouts causing increases in the predicted crash frequencies. These findings are in line with both North American and international literature on roundabout safety, which often links this trend to the increased conflict points created by lane changes on the approaches and circulating roadway (*2*) (*54*).

As safety can be considered as the long term average of a site's crash frequencies and a number of complex relationships govern their occurrence, it is suggested that authorities better prepare themselves for long-term safety audits by improving their data collecting practices and ensuring that they are followed. As is seen in this research, the level of analysis that is possible is highly dependent on the quality of the data that is made available to researchers.

Keeping in mind that a definitive conclusion on the safety effect of roundabouts in the province of Quebec is impossible due to the statistical insignificance of the results, there is evidence that smaller roundabouts, roundabouts constructed in an environment serving the same type of road user and roundabouts with little differential in their approach speeds tend to demonstrate safety improvements after their construction. This observation is in agreement with the conclusions of previous research (*2*) (*54*).

Although the analysis as a whole could provide greater analytical power given access to vehicular flows in combination with more detailed design variables, the analysis provides a good basis for future work. Due to the complexity of the relationships between the road network and crash occurrence, it is suggested that future work should investigate roundabout safety through a stronger model such as the Empirical Bayes method in order to control for regression-to-the-mean effects, among others. Similarly, future work applying a disaggregate approach in order to capture the influence of individual design elements and even the role of driver behavior at individual roundabouts is warranted.

Finally, this work considered all crashes together, with no distinction for injury severity. As was demonstrated in Chapter 3, injury severity is reduced in roundabouts and a link between crash outcomes and their occurrence should be pursued in future work.

Chapter 5 Final Conclusions and Future Work

This research has proposed a methodology for mapping and analyzing crash data for evaluating the safety performance of roundabouts in the province of Quebec, Canada. The proposed methodologies and empirical evidence are summarized as follows.

Chapter 2 provides a tool based on online geocoding services such as the Google Maps API for the geocoding of crash records that are provided without geographical coordinates. Through the detailed analysis of a case study in a municipality in Quebec, Canada, it was found that at the strictest level, a match rate of 78% could be achieved through a custom algorithm. This research demonstrated that with some precautions, a custom-designed algorithm can provide matching rates similar to those that can be found on the commercial market, and with an accuracy level that is acceptable for most applications. Expanding the methodology to the entire crash record database available as a part of this research project, it was found that a total of 716,029 or 94.0% of the records could be associated with geographical coordinates. The research in Chapter 2 also explored the geocoder's spatial accuracy through the case study. It was found that although the results tend to vary substantially from record to record, factors such as the completeness of the input address fields and the ability of the API to interpret the location description (in particular for route number addresses) have a large influence on this outcome. It is suggested that with proper user revision, the results are sufficient for practical applications such as intersection safety analysis.

Chapter 3 investigates the factors linked to injury severity at Quebec roundabouts. Through the use of an ordered logit model, a number of factors that significantly influence injury severity outcomes at roundabouts were found to include: the season, the number and type of vehicles involved in the crash, the type of impact, the road surface conditions and the weather at the time of the accident, whether a hit and run was observed, the lighting conditions at the time of the accident, and also whether the accident occurred on an approach to the roundabout or within it. These findings help reinforce the idea that roundabout designs need to consider the needs of the road user first and foremost, as the injury severity that occurs during a given crash event is dependent on a number factors.

Chapter 4 investigates roundabout safety through the use of a crash frequency analysis, and was accomplished with mixed results. A before-after study with control groups was implemented due to the absence of vehicular count data. A complementary negative binomial regression was also performed in order to investigate factors affecting roundabout crash occurrence.

Given the large variability of the results (potentially associated with small sample sizes), the findings are limited due to the large confidence intervals. With the available dataset, only inconclusive evidence could be extracted, as complex relationships exist in the factors affecting crash occurrence and could not be isolated. Nevertheless a number of interesting observations were made. From the available evidence, it can be seen that in its current form, the crash frequency analysis of a sample of Quebec roundabouts suggests that safety varies across subgroups or clusters of sites, although similarities can be observed between sites of a same subgrouping. Keeping in mind the statistical limitations of the results (with large confidence intervals), there is evidence that smaller roundabouts, roundabouts constructed in an environment serving a single type of traffic and roundabouts with little differential in their approach speed limits tend to demonstrate safety improvements after their construction. Although the analysis as a whole could provide greater analytical power given access to vehicular flows in combination with more detailed design variables, the analysis provides a good basis for future work. Similarly, the proposed methodologies remain valid.

Although the expected results may not be immediately applicable in practice, the research presented as part of this thesis will help align future efforts to answer the questions posed at the beginning of this work. Of note, providing that proper licensing is obtained for the use of the online API services, future work should focus on a more detailed investigation using alternative methods for the accuracy estimation associated with online geocoding services and their implementation through a custom algorithm.

With regards to Chapter 3 and 4, future work is dependent on the eventual availability of a more complete dataset with a larger set of treated sites. With this, more robust modelling frameworks (e.g., empirical and full Bayes approaches) can then be considered in order to account for the weaknesses identified throughout this work, and to validate the observations presented throughout. Enabling stronger statistical models will also help strengthen the conclusions of this thesis.

More specifically, future work for the injury severity analysis presented in Chapter 3 should introduce regular stop and signal-controlled intersections into the analysis to be able to compare how the influence of the various factors changes with respect to injury severity.

Similarly, future work for Chapter 4 should investigate the effectiveness of roundabout safety by type and severity of crash in order to get a better idea on how crash occurrence and injury severity are linked at roundabouts. Given the availability of traffic flow data, future work should also aim to develop safety performance functions for various treatment configurations and geometric factors. Finally, the author proposes that a study based on the detailed investigation of individual sites be considered due to the observed variability in roundabout crash frequency performance presented in Chapter 4. This methodology would help identify the influence of individual design elements on the safety effectiveness of roundabouts.

References

1. Transportation Research Board. Roundabouts in the United States. NCHRP Report 572. Transportation Research Board of the National Academies, Washington, D.C., 2007.

2. Transportation Research Board. Roundabouts: An Informational Guide. NCHRP Report 672. Transportation Research Board of the National Academies, Washington, D.C., 2010.

3. Pellecuer, L., and M. St-Jacques. Dernières avancées sur les carrefours giratoires. *Revue Canadien de Génie Civil*, Vol. 35, 2008, pp. 542-553.

4. Federal Highway Administration. Roundabouts: Technical Summary. U.S. Department of Transportation, Technical Summary 2010.

5. Ministère des Transports du Québec. Le carrefour giratoire: un mode de gestion différent. Les Publications du Québec, Québec, 2002.

6. Filion, N. Roundabouts: When going around gets you somewhere. CAA-Quebec, 2011. http://www.caaquebec.com/DocumentLibrary/UploadedContents/RadFiles/Salle_de_presse/Roundabouts-Spring-2011.pdf.

7. Beaupré, A. Évaluation et recommandations pour le développement des carrefours à sens giratoire au Québec. École de technologie supérieure, Montreal, Thesis 2011. http://espace.etsmtl.ca/876/.

8. Flannery, A., and L. Elefteriadou. A Review of Roundabout Safety Performance in the United States. in *Enhancing Transportation Safety in the 21st Century ITE International Conference*, Washington, D.C., 1999, p. 12. http://www.ite.org/traffic/documents/CCA99A33.pdf.

9. Stone, J., K. Chae, and S. Pillalamarri. The effects of roundabouts on pedestrian safety. The Southeastern Transportation Center, Knoxville, Tennessee, 2002.

10. Flannery, A., and T. K. Datta. Modern Roundabouts and Traffic Crash Experience in United States. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1553, no. 1, January 1996, pp. 103-109. http://trb.metapress.com/content/5373766055T27053.

11. Institute of Transportation Engineers. Before-and-After Study Technical Brief. Transportation Safety Council, Washington, D.C., 2009.

12. St-Aubin, P., N. Saunier, L. F. Miranda-Moreno, and K. Ismail. Use of Computer Vision Data for Detailed Driver Behavior Analysis and Trajectory Interpretation at Roundabouts. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2389, 2013, pp. 65-77.

13. St-Aubin, P., N. Saunier, and L. F. Miranda-Moreno. Road User Collision Prediction Using

Motion Patterns Applied to Surrogate Safety Analysis. in *Transportation Research Board 93rd Annual Meeting*, Washington, D.C., 2014, p. 13p.

14. Ritchie, S., and M. Lenters. High Speed Approaches at Roundabouts. *Transportation Research Circular*, 2005, p. 56.

15. Google Inc. Google Maps. https://www.google.ca/maps. Accessed October 2014.

16. Kim, S., and J. Choi. Safety Analysis of Roundabout Designs based on Geometric and Speed Characteristics. *LSCE Journal of Civil Engineering*, Vol. 17, no. 6, 2013, pp. 1146-1454.

17. Persaud, B. N., R. A. Retting , P. E. Garder, and Lord. Observational before-after study of the safety effect of U.S. roundabout conversions using the empirical Bayes method. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1751, 2001.

18. Gross, F., C. Lyon, B. Persaud, and R. Srinivasan. Safety effectiveness of converting signalized intersections to roundabouts. *Accident Analysis and Prevention*, April 2012. http://www.sciencedirect.com/science/article/pii/S0001457512001455.

19. Pellecuer, L. Étude de faisabilité de l'implantation des carrefours giratoires au Québec. École de Technologie Supérieure, Montréal, Québec, Thesis 2003.

20. Chen, Y., B. Persaud, E. Sacchi, and Bassani. Investigation of models for relating roundabout safety to predicted speed. *Accident Analysis & Prevention*, Vol. 50, January 2013, pp. 196-203.

21. Saunier, N. Roundabout Safety in Québec. December 21, 2012. http://giratoires.confins.net/. Accessed October 20, 2014.

22. Transportation Research at McGill. http://tram.mcgill.ca/. Accessed October 25, 2014.

23. Google, Inc. The Google Geocoding API. *Google Developers*, 2013. https://developers.google.com/maps/documentation/geocoding/. Accessed July 25, 2013.

24. National Research Council (US). Transportation Research Board. Task Force on Development of the Highway Safety Manual and American Association of State Highway and Transportation Officials. *Highway Safety Manual*. AASHTO, Washington D.C., 2010.

25. Chen, J. Black Spot Determination of Traffic Accident Locations and Its Spatial Association Characteristic Analysis Based on GIS. *Journal of Geographic Information System*, Vol. 4, 2012, pp. 608-617.

26. Christen, P., T. Churches, and A. Willmore. A Probabilistic Geocoding System based on a National Address File. in *The Australasian Data Mining Conference*, Cairns, Australia, 2004.

27. Davis Jr., C. A., F. T. Fonseca, and K. A.d.V. Borges. A flexible Addressing System for Approximate Geocoding. in *Proceedings of the Fifth Brazilian Symposium on geoInformatics*, Sao Paulo, Brazil, 2003.

28. Goldberg, D. A. A Geocoding Best Practices Guide. University of Southern California, Guide, 2008.

29. Levine, N., and K. E. Kim. The Location of Motor Vehicle Crashes In Honolulu: A Methodology For Geocoding Intersections. *Computers, Envrionment and Urban Systems*, Vol. 22, no. 6, 1998, pp. 557-576.

30. Velavan, K. Developing Tools and Data Model for Managing and Analysing Traffic Accident. University of Texas at Dallas, Dallas, Texas, Thesis 2006.

31. Bigham, J. M., T. M. Rice, S. Pande, J. Lee, S. H. Park, N. Gutierrez, and D. R. Ragland. Geocoding Police Collision Report Data from California: A Comprehensive Approach. *International Journal of Health Geographics*, Vol. 8, no. 72, December 2009.

32. Cherry, E., R. Floyd, T. Graves, S. Martin, and D. Ward. Crash Data Collection and Analysis System. ARCADIS G&M of North Carolina, Inc. , Pheonix, Arizona, Final Report FHWA-AZ-06-537, 2006.

33. Zandbergen, P. A. A Comparison of Address Point, Parcel and Street Geocoding Techniques. *Computers, Environment and Urban Systems* , Vol. 32, 2008, pp. 214-232.

34. Davis Jr., C. A., and F. T. Fonseca. Assessing the Certainty of Locations Produced by an Address Geocoding System. *Geoinformatica*, Vol. 11, 2007, pp. 103-129.

35. Ratcliffe, J. H. Geocoding Crime and a First Estimate of a Minimum Acceptable Hit Rate. *Juornal of Geographical Information Science*, Vol. 18, no. 1, January-February 2004, pp. 61-72.

36. Goldberg, D. W., J. P. Wilson, and C. A. Knoblock. From Text to Geographic Coordinates: The Current State of Geocoding. *URISA Journal*, Vol. 19, no. 1, 2007, pp. 33-46.

37. Yahoo!, Inc.. Yahoo! Maps Web Service. 2013. http://developer.yahoo.com/maps/. Accessed July 25, 2013.

38. Qin, X., S. Parker, Y. Liu, A. J. Grettinger, and S. Forde. Intelligent Geocoding System to Locate Traffic Crashes. *Accident Analysis and Prevention*, Vol. 50, 2013, pp. 1034-1041.

39. GISgraphy. *GISgraphy Results Comparator*, 2012. http://www.gisgraphy.com/compare/. Accessed July 28, 2013.

40. OpenStreetMap. OpenStreetMap. 2013. http://www.openstreetmap.org/. Accessed July 28, 2013.

41. Savolainen, P. T., F. L. Mannering, D. Lord, and M. A. Quddus. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention*, Vol. 43, no. 5, 2011, pp. 1666-1676.

42. Abdel-Aty, M. Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of Safety Research*, Vol. 34, no. 5, 2003, pp. 597-603.

43. Al-Ghamdi, A. S. Using logistic regression to estimate the influence of accident factors on accident severity. *Accident Analysis and Prevention*, Vol. 34, no. 6, 2002, pp. 729-741.

44. Retting, R. A., B. N. Persaud, P. E. Garder, and D. Lord. Crash and injury reduction following installation of roundabouts in the United States. *American Journal of Public Health*, Vol. 91, no. 4, April 2001, pp. 628-631.

45. Angel, A., and M. Hickman. Analysis of the Factors Affecting the Severity of Two-Vehicle Crashes. *INGENIERÍA & DESARROLLO*, Vol. 24, 2008. http://ciruelo.uninorte.edu.co/pdf/ingenieria_desarrollo/24/11_Analysis%20of%20the%20factors.pdf.

46. Lee, J., and F. Mannering. Analysis of Roadside Accident Frequency and Severity and Roadside Safety Management. Seattle, Research Project T9903,. http://www.wsdot.wa.gov/research/reports/fullreports/475.1.pdf.

47. Ye, F., and D. Lord. Comparing Three Commonly Used Crash Severity Models on Sample Size Requirements: Multinomial Logit, Ordered Probit and Mixed Logit Models. in *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., 2010, p. 20.

48. Kockelman, K. M., and Y.-J. Kweon. Driver injury severity: an application of ordered probit models. *Accident Analysis & Prevention*, Vol. 34, no. 3, May 2002, pp. 313-321.

49. Ye, F., and D. Lord. Investigating the Effects of Underreporting of Crash Data on Three Commonly Used Traffic Crash Severity Models: Multinomial Logit, Ordered Probit and Mixed Logit Models. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2241, 2011, pp. 51-58.

50. Eluru, N., M. Bagheri, L. Miranda-Moreno, and L. Fu. A latent class modelling approach for identifying vehicle driver injury severity factors at highway-railway crossings. *Accident Analysis & Prevention*, Vol. 47, no. 1, 2012, pp. 119-127.

51. Eluru, N., R. Paleti, R. M. Pendyala, and C. R. Bhat. Modelling Injury Severity of Multiple Occupants of Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2165, 2010, pp. 1-11.

52. Huang, H., H. C. Chin, and M. M. Haque. Severity of driver injury and vehicle damage in traffic crashes at intersections: A Bayesian hierarchical analysis. *Accident Analysis & Prevention*, Vol. 40, no. 1, 2008, pp. 45-54.

53. Chen, H., L. Cao, and D. B. Logan. Analysis of Risk Factors Affecting the Severity of Intersection Crashes by Logistic Regression. *Traffic Injury Prevention*, Vol. 13, no. 3, 2012, pp. 300-307.

54. Elvik, R. Effects on Road Safety of Converting Intersections to Roundabouts: Review of Evidence from Non-U.S. Studies. *Transportation Research Record: Journal of the Transportation Research Board*, 2003, pp. 1-10.

55. Daniels, S., T. Brijs, E. Nuyts, and G. Wets. Extended prediction models for crashes at roundabouts. *Safety Science*, Vol. 49, no. 2, 2011, pp. 198-207.

56. Wang, X., M. A. Abdel-Aty, A. Nevarez, and J. B. Santos. Investigation of Safety Influence Area for Four-Legged Signalized Intersections: Nationwide Survey and Empirical Inquiry. in *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2083, Washington, D.C., 2008, pp. 86-95.

57. Daniels, S., T. Brijs, E. Nuyts, and G. Wets. Externality of risk and crash severity at roundabouts. *Accident Analysis and Prevention*, Vol. 42, no. 6, November 2010, pp. 1966-1973.

58. Usman, T., L. Miranda-Moreno, and L. Fu. Analysis of Injury Severity Outcomes of Highway Winter Crashes: Multilevel Modeling Approach. in *91st Annual Meeting of the Transportation Research Board*, Washinton, D.C., 2012, p. 20.

59. Yang, Z., L. I. Zhibin, L.I. U. Pan, and Z.H. A. Liteng. Exploring contributing factors to crash injury severity at freeway diverge areas using ordered probit model. *Procedia Engineering*, Vol. 21, no. 0, 2011, pp. 178-185.

60. Shankar, V., F. Mannering, and W. Barfield. Statistical analysis of accident severity on rural freeways. *Accident Analysis and Prevention*, Vol. 28, no. 3, 1996, pp. 391-401.

61. Zahabi, S. A.H., J. Strauss, K. Manaugh, and L. F. Miranda-Moreno. Estimating Potential Effect of Speed Limits, Built Environment, and Other Factors on Severity of Pedestrian and Cyclist Injuries in Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2247, no. 1, 2011, pp. 81-90.

62. Edwards, J. B. The Relationship Between Road Accident Severity and Recorded Weather. *Journal of Safety Research*, Vol. 29, no. 4, 1998, pp. 249-262.

63. Fridstrom, L., J. Ifver, S. Ingebrigsten, R. Kulmala, and L. K. Thomsen. Measuring the Contribution of Randomness, Exposure, Weather, and Daulight to the Cariation in Road Accident Counts. *Accident Analysis and Prevention*, Vol. 27, no. 1, 1995, pp. 1-20.

64. Daniels, S., E. Nuyts, and G. Wets. The effects of roundabouts on traffic safety for bicyclists: An observational study. *Accident Analysis and Prevention*, Vol. 40, no. 2, 2008, pp. 518-526.

65. American Association of State Highway Transportation Officials (AASHTO). *Highway Safety Manual*, 1st ed. AASHTO, Washington, D.C., United States of America, 2010.

66. James, B., Y. Chen, and B. Persaud. Assessment of the Crash Modification Factors in the Highway Safety Manual for use in Canada. in *Annual Conference of the Transportation Association of Canada*, Halifax, Nova Scotia, 2010, p. 17.

67. McIntosh, K., C. Redinger, and J. Bagdade. Evaluating the Performance and Safety Effectiveness of Roundabouts. Opus International Consultants Inc., West Bloomfield, Michigan, Research Report OR 09083, 2011.

68. Canadian Institute of Transportation Engineers. Roundabouts in Canada: A Primer for Decision-Makers. Canadian Institute of Transportation Engineers, 2013. http://www.cite7.org/tlc/documents/Roundabouts_in_Canada_2013.pdf.

69. Lenters, M. Roundabout Planning and Design for Efficiency & Safety Case Study: Wilson Street/Meadowbrook Drive/Hamilton Drive--City of Hamilton. in *The Transportation Factor 2003. Annual Conference and Exhibition of the Transportation Association of Canada. (Congres et Exposition Annuels de l'Association des transport du Canada)*, St. John's, Newfoundland and Labrador , 2003, p. 21.

70. Henderson, R., and N. Button. Pedestrian Safety in Roundabouts. in *2013 Conference of the Transportation Association of Canada*, Winnipeg, Manitoba, 2013, p. 9. http://conf.tac-atc.ca/english/annualconference/tac2013/session9/button.pdf.

71. El-Basyouny, K., and T. Sayed. Comparison of Two Negative Binomial Regression Techniques in Developing Accident Prediction Models. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1950, 2006, pp. 9–16.

72. Hadayeghi, A., A. S. Shalaby, and B. N. Persaud. Macro-Level Accident Prediction Models for Evaluating the Safety of Urban Transportation Systems. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1840, no. 1, January 2003, pp. 87-95.

73. Usman, T., L. Fu, and L. F. Miranda-Moreno. Accident Prediction Models for Winter Road Safety. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2237, no. 1, 2011, pp. 144-151.

74. Persaud and Lyon Inc.. Safety Performance Functions for Intersections. Colorado Department of Transportation , Denver, Colorado, Report No. 2009-10 , 2009.

75. Harwood, D., K. Bauer, D. Potts, D. Torbic, K. Richard, E. Kohlman Rabbani, E. Hauer, and L. Elefteriadou. Safety Effectiveness of Intersection Left- and Right-Turn Lanes. Federal Highway Administration, McLean, Virginia, FHWA-RD-02-089, 2002.

76. Hauer, E. *Observational Before-After Studies in Road Safety*, 2nd ed. Elsevier Science,

Kidlington, Oxford, United Kingdom, 1997.

77. Gross, F., B. Persaud, and C. Lyon. A Guide to Developing Quality Crash Modification Factors. Federal Highway Administration, Washington, D.C., FHWA-SA-10-032, 2010.

78. Persaud, B., C. Lyon, and Y. Chen. Tools for Estimating the Safety and Operational Impacts of Roundabouts. Ryerson University, Ottawa, Ontario, Report 2010.

79. Erdman, D., L. Jackson, and A. Sinko. Zero-Inflated Poisson and Zero-Inflated Negative Binomial Models Using the COUNTREG Procedure. in *SAS Global Forum 2008*, Cary, North Carolina, 2008, pp. 1-11. http://www2.sas.com/proceedings/forum2008/322-2008.pdf.

80. StataCorp LP. STATA Data Analysis and Statistical Software. 2007.

81. St-Aubin, P., N. Saunier, and L. F. Miranda-Moreno. Large-Scale Microscopic Traffic Behaviour and Safety Analysis of Québec Roundabout Design. in *TRB Annual Meeting Compendium of Papers*, Washington, D.C., 2015.

82. Gates, T. J., and R. E. Maki. Converting Old Traffic Circles to Modern Roundabouts: Michigan State University Case Study. in *Proceedings of ITE 2000 Annual Meeting and Exhibit*, Nashville, Tennessee , 2000, p. 23.

83. Guo, T., Y. Liu, and W. Deng. A Disaggregate Speed Consistency Measure for the Safety Evaluation of Freeway Diverging Area. in *3rd International Conference on Road Safety and Simulation*, Indianapolis, 2011.

Appendix A: General Roundabout Design Features

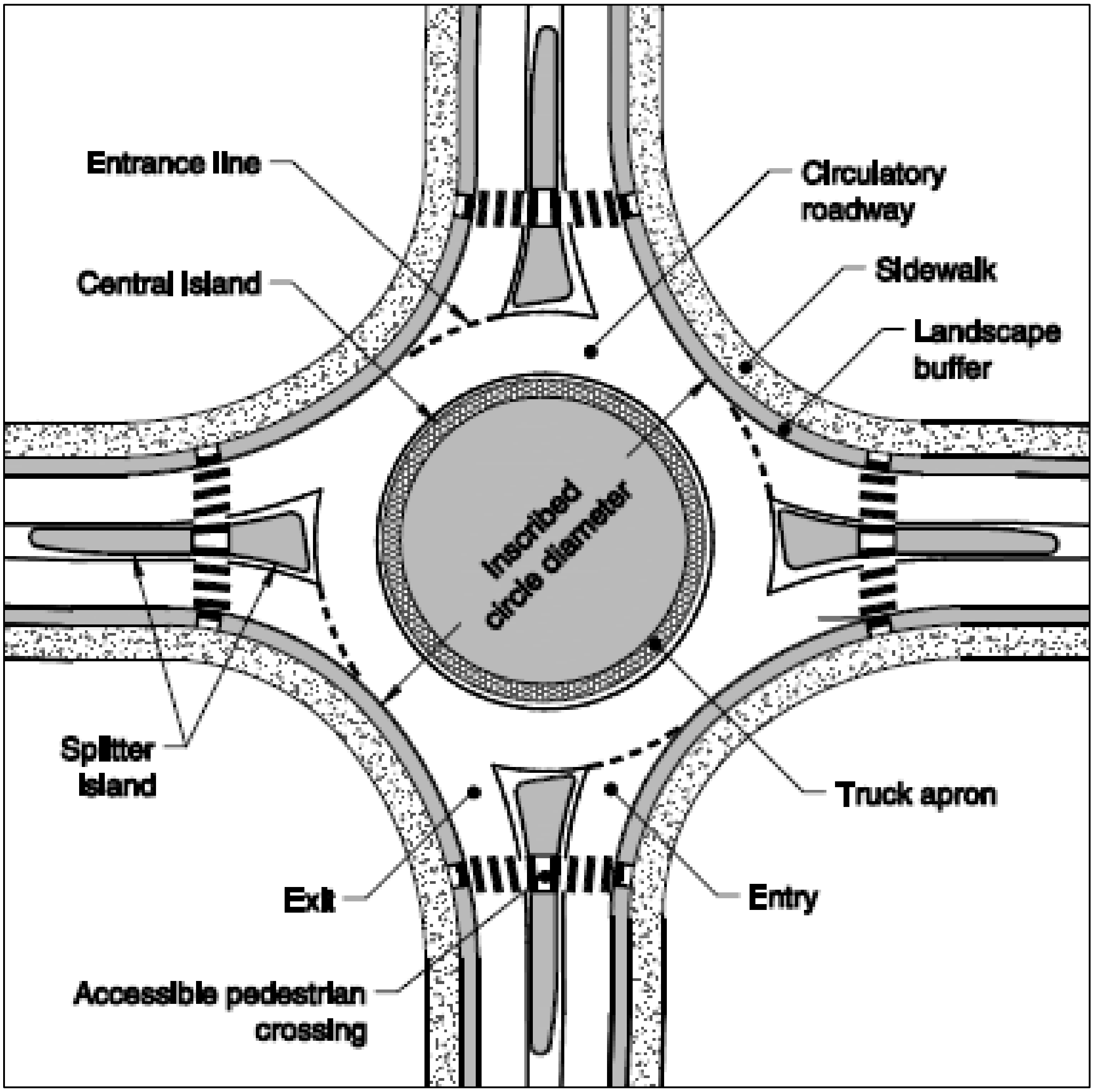

FIGURE A-1: General Roundabout Design Features (*2*).

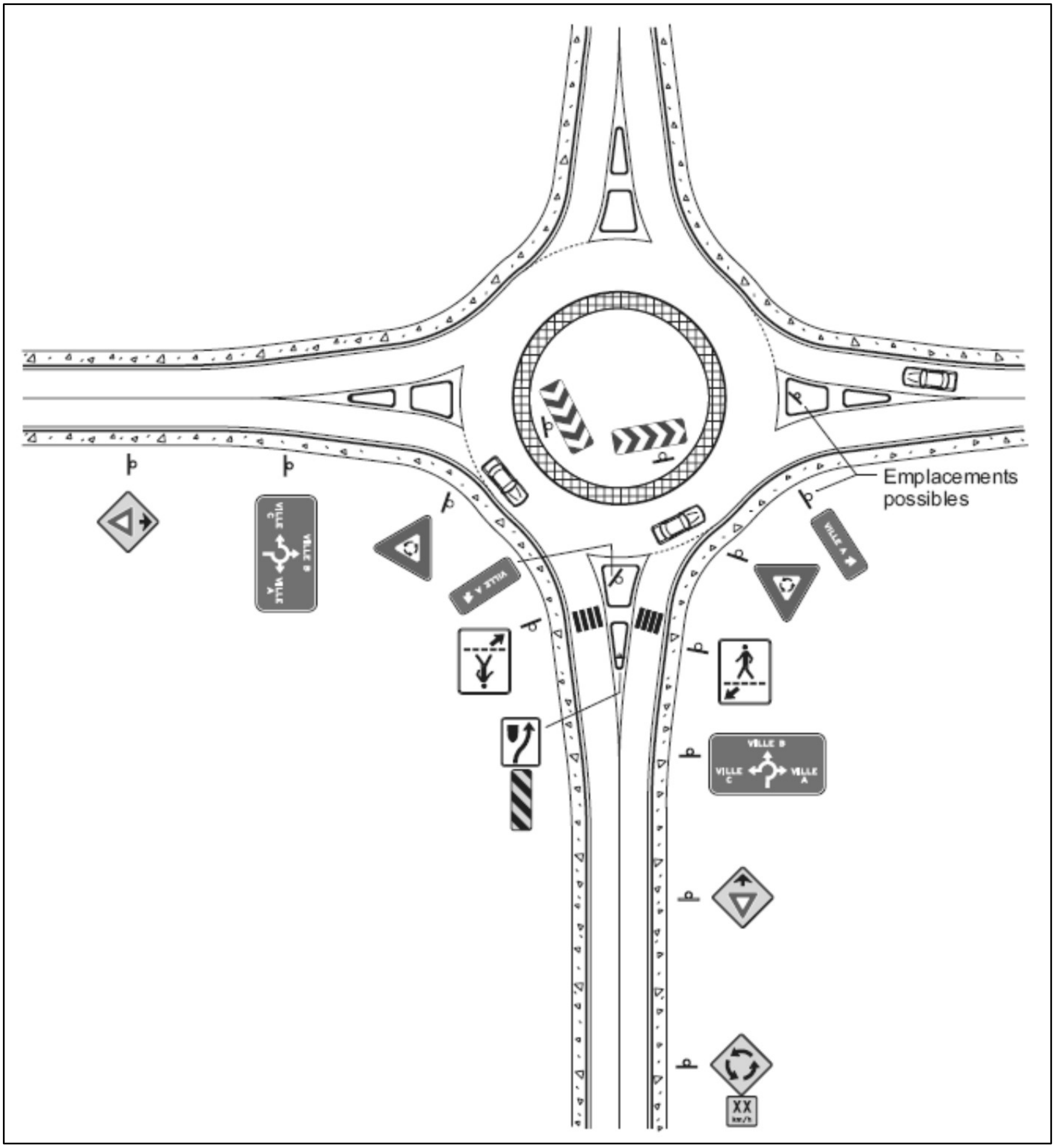

FIGURE A-2: Standard roundabout signalisation as illustrated in the Quebec design guide (*5*)**.**

Appendix B: Summary of Roundabout Crash Records

Site ID	Municipality	Total Accidents	Before Accidents	During Accidents	After Accidents	Construction Date	Earliest Record	Latest Record	Analysis?
1	Chambly	53	0	4	49	2003	28/10/2002	11/10/2011	No
2	Chambly	58	15	6	37	2003	19/01/2000	20/11/2011	
3	Chambly	47	12	2	33	2004	13/08/2000	28/07/2011	
4	Sainte-Julie	20	12	1	7	2007	16/01/2000	11/03/2011	
5	Boucherville	17	4	1	12	2005	18/09/2002	08/11/2011	
6	Montréal	3	1	1	1	2008	11/12/2007	22/06/2011	
7	Montréal	1	0	0	1	2008	29/12/2010	29/12/2010	No
8	Montréal	65	17	6	42	2004	10/01/2000	14/09/2011	
9	Montréal	10	0	0	10	2003	17/11/2004	01/05/2011	No
10	Mont-Saint-Hilaire	0	0	0	0	2003	-	-	No
12	Mont-Tremblant	55	10	3	42	01/11/2002	12/02/2000	13/07/2011	
13	Mont-Tremblant	0	0	0	0	2006	-	-	No
14	Mont-Tremblant	17	0	0	17	2006	08/03/2008	06/11/2011	No
15	Mont-Tremblant	107	57	14	36	01/12/2004	16/01/2000	13/09/2011	
16	Sainte-Agathe-des-Monts	100	46	11	43	01/11/2008	20/02/2000	02/09/2011	
17	Gatineau	90	0	0	90	01/12/2005	03/10/2006	20/11/2011	No
18	Gatineau	99	0	0	99	01/12/2005	23/08/2006	23/11/2011	No
19	Gatineau	380	0	0	380	01/11/2004	04/07/2005	11/11/2011	No
20	Gatineau	77	16	4	57	01/11/2003	01/09/2000	13/06/2011	
21	Montréal	19	9	1	9	01/08/2006	08/01/2000	22/11/2011	
23	Saint-Henri	39	16	7	16	01/10/2006	15/03/2000	29/09/2011	
24	New Richmond	58	47	3	8	01/11/2009	21/01/2000	21/10/2011	
25	Shawinigan	137	37	17	83	01/12/2003	10/01/2000	10/10/2011	
26	Louiseville	139	63	14	62	01/11/2004	13/07/2000	18/08/2011	
27	Montréal	15	0	0	15	1998	18/01/2002	26/09/2009	No
29	Gatineau	48	22	5	21	01/08/2005	25/01/2000	01/08/2011	

Appendix B: Summary of Roundabout Crash Records (Continued)

30	Granby	52	29	5	18	01/11/2009	31/01/2000	11/11/2011	
33	Amos	63	13	7	43	01/11/2003	11/04/2000	15/12/2010	
34	Amos	63	18	9	36	01/11/2003	21/07/2000	28/11/2011	
35	Val-d'Or	57	20	7	30	24/06/2005	10/03/2000	13/10/2011	
36	Val-d'Or	169	53	15	101	01/10/2003	08/02/2000	19/09/2011	
37	Val-d'Or	57	20	4	33	01/10/2004	21/04/2000	30/07/2011	
38	Val-d'Or	109	48	7	54	01/10/2005	28/01/2000	22/10/2011	
41	Pointe-Lebel	89	36	8	45	01/07/2005	03/02/2000	15/11/2011	
42	Saint-Gédéon	53	41	7	5	01/11/2009	24/05/2000	09/08/2011	
43	Vaudreuil-Dorion	23	1	2	20	01/11/2006	11/07/2003	19/10/2011	
44	Saguenay	87	81	2	4	2010	01/02/2000	30/08/2011	
45	Saint-Henri	37	19	7	11	01/09/2007	11/01/2000	01/12/2010	
57	Montréal	7	0	2	5	2003	07/08/2003	12/05/2011	No
61	Saguenay	184	-	-	-	**	-	-	No
62	Saguenay	4	-	-	-	**	-	-	No
65	Montréal	8	0	3	5	2007	31/07/2007	29/06/2010	No
69	Trois-Rivières	4	-	--	-	**	-	-	No
74	Saguenay	83	27	10	46	01/06/2004	14/01/2000	15/10/2011	
75	Saguenay	27	1	4	22	2002	21/06/2001	16/09/2011	
86	Mont-Tremblant	7	4	2	1	2006	31/01/2002	19/12/2009	
87	Mont-Tremblant	2	1	0	1	2006	05/06/2003	27/12/2007	
88	Trois-Rivières	90	38	15	37	01/06/2006	14/05/2000	06/11/2011	
89	Trois-Rivières	288	100	30	158	01/06/2006	04/02/2000	23/11/2011	
95	Saguenay	16	14	1	1	01/04/2009	01/02/2000	30/04/2011	
96	Saguenay	13	9	0	4	01/06/2005	09/04/2000	01/12/2009	
104	Mont-Tremblant	86*	-	-	-	01/11/2012*	25/07/2000	07/08/2011	No
143	Mont-Tremblant	34*	-	-	-	01/11/2012*	25/09/2002	10/05/2011	No
144	Saint-Irénée	15	11	2	2	2010	15/07/2001	10/10/2011	

*** = Roundabout built outside of data range** **** = No construction date available**

Appendix C: Summary of Before-After Analysis Data

Municipality	ID	Total Crashes	Before Period (years)	After Period (years)	Number of Crashes Before Construction	Number of Crashes After Construction	Site ID	Total Crashes	Number of Crashes Before Construction	Number of Crashes After Construction
Chambly	2	58	3.0	7.0	15	37	2.1	17	3	14
							2.2	30	3	27
							2.3	56	23	33
							2.4	10	6	4
Chambly	3	47	4.0	6.0	12	33	3.1	30	7	23
							3.2	13	4	9
							3.3	17	5	12
							3.4	8	1	7
Sainte-Julie	4	20	7.0	3.0	12	7	4.1	5	3	2
							4.2	13	6	7
							4.3	9	5	4
							4.4	12	7	5
							4.5	5	1	4
							4.6	18	12	6
Bouchervi lle	5	17	2.0	5.0	4	12	5.1	48	9	39
							5.2	8	2	6
							5.3	14	3	11
							5.4	23	10	13
Montréal	8	65	4.0	6.0	17	42	8.1	42	10	32
							8.2	21	8	13
							8.3	52	19	33
							8.4	71	21	50
							8.5	63	23	40
Mont-Tremblant	12	55	2.5	7.5	10	42	12.1	23	3	20
							12.2	40	9	31
							12.3	15	6	9
							12.4	42	3	39
							12.5	68	11	57

Appendix C: Summary of Before-After Analysis Data (Continued)

Municipality	ID	Total Crashes	Before Period (years)	After Period (years)	Number of Crashes Before Construction	Number of Crashes After Construction	Site ID	Total Crashes	Number of Crashes Before Construction	Number of Crashes After Construction
Gatineau	20	3.5	6.5	77	16	57	20.1	82	21	61
							20.2	19	5	14
							20.3	17	9	8
							20.4	15	5	10
Montréal	21	6.0	4.0	19	9	9	21.1	32	14	18
							21.2	41	21	20
							21.3	23	14	9
							21.4	48	21	27
Saint-Henri	23	6.0	4.0	39	16	16	23.1	10	5	5
							23.2	14	6	8
							23.3	18	11	7
							23.4	10	7	3
							23.5	30	17	13
							23.6	21	12	9
							23.7	35	23	12
Shawinigan	25	3.5	6.5	137	37	83	25.1	145	42	103
							25.2	94	31	63
							25.3	141	52	89
							25.4	86	35	51
							25.5	80	30	50
Louiseville	26	5.5	5.5	139	63	62	26.1	16	9	7
							26.2	31	22	9
Gatineau	29	5.0	5.0	48	22	21	29.1	19	8	11
							29.2	31	13	18
							29.3	26	12	14
							29.4	59	25	34
							29.5	63	26	37
							29.6	93	40	53
Amos	33	3.5	6.5	63	13	43	33.1	61	19	42
							33.2	45	15	30
							33.3	21	6	15
							33.4	17	11	6
							33.5	25	7	18

Appendix C: Summary of Before-After Analysis Data (Continued)

Municipality	ID	Total Crashes	Before Period (years)	After Period (years)	Number of Crashes Before Construction	Number of Crashes After Construction	Site ID	Total Crashes	Number of Crashes Before Construction	Number of Crashes After Construction
Amos	34	3.5	6.5	63	18	36	34.1	45	15	30
							34.2	152	48	104
							34.3	17	11	6
							34.4	25	7	18
Val-d'Or	35	5.0	5.0	57	20	30	35.1	21	11	10
							35.2	59	23	36
							35.3	39	17	22
							35.4	25	9	16
							35.5	46	20	26
Val-d'Or	36	3.0	7.0	169	53	101	36.1	132	25	107
							36.2	66	15	51
							36.3	40	15	25
							36.4	99	25	74
Val-d'Or	37	4.0	6.0	57	20	33	37.1	23	9	14
							37.2	26	9	17
							37.3	61	18	43
							37.4	40	17	23
							37.5	23	7	16
							37.6	43	14	29
Val-d'Or	38	5.0	5.0	109	48	54	38.1	126	52	74
							38.2	61	26	35
							38.3	40	19	21
							38.4	98	40	58
Saint-Henri	45	7.0	3.0	37	19	11	45.1	10	5	5
							45.2	15	9	6
							45.3	19	12	7
							45.4	12	9	3
							45.5	27	18	9
							45.6	18	13	5
							45.7	37	27	10

Appendix C: Summary of Before-After Analysis Data (Continued)

Municipality	ID	Total Crashes	Before Period (years)	After Period (years)	Number of Crashes Before Construction	Number of Crashes After Construction	Site ID	Total Crashes	Number of Crashes Before Construction	Number of Crashes After Construction
Saguenay	74	4.0	6.0	83	27	46	74.1	100	31	69
							74.2	57	19	38
							74.3	59	22	37
							74.4	75	27	48
							74.5	73	30	43
							74.6	84	36	48
Mont-Tremblant	86	4.0	4.0	7	4	1	86.1	9	2	7
							86.2	20	4	16
Trois-Rivières	88	6.0	4.0	90	38	37	88.1	28	10	18
							88.2	27	15	12
							88.3	66	42	24
							88.4	53	31	22
							88.5	54	15	39
							88.6	81	53	28
Trois-Rivières	89	6.0	4.0	288	100	158	89.1	60	30	30
							89.2	305	163	142
							89.3	100	63	37
							89.4	67	29	38
							89.5	215	127	88
Saguenay	96	4.0	5.0	13	9	4	96.1	11	3	8
							96.2	9	6	3
							96.3	9	6	3
							96.4	8	4	4
							96.5	14	4	10
							96.6	6	4	2
Mont-Tremblant	15	4.5	5.5	107	57	36	15.1	21	6	15
							15.2	47	22	25
							15.3	15	7	8
							15.4	30	2	28
							15.5	41	8	33
							15.6	72	36	36