1-1-2011

Calibration of highway safety manual predictive models and validation of safety performance funtions for canadian urban/suburban intersections

Shahzad Faisal
Ryerson University

CALIBRATION OF HIGHWAY SAFETY MANUAL PREDICTIVE MODELS AND VALIDATION OF SAFETY PERFORMANCE FUNCTIONS FOR CANADIAN URBAN/SUBURBAN INTERSECTIONS

By

Shahzad Faisal

Bachelor of Engineering, University of Engineering and Technology, Lahore, Pakistan, 1991

A Thesis

Presented to Ryerson University in partial fulfilment of the requirements of the degree of Master of Applied Science in the Program of Civil Engineering

Toronto, Ontario, Canada. 2011

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other Institutions or individuals for the purpose of scholarly research.

Signature

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other Institution or individuals for the purpose of scholarly research.

Signature

Shahzad Faisal

Department of Civil Engineering

Ryerson University

BORROWER'S PAGE

Ryerson University requires the signature of all persons using or photocopying this thesis. Please sign below and give your address and date.

Name	Signature	Address	Date

“CALIBRATION OF HIGHWAY SAFETY MANUAL PREDICTIVE MODELS AND VALIDATION OF SAFETY PERFORMANCE FUNCTIONS FOR CANADIAN URBAN/SUBURBAN INTERSECTIONS”

By

Shahzad Faisal

MASc, Civil Engineering, 2011, Department of Civil Engineering, Ryerson University

Abstract

In this research, the HSM predictive models for collisions on urban/suburban arterials are calibrated for collision data from the City of Toronto. It has been found that the use of calibration factors for applying HSM models to Toronto intersection data is not appropriate. New collision models are therefore developed by using local data. The HSM and Toronto models are then calibrated to City of Edmonton intersection collision data to determine whether it is better to calibrate HSM models for a Canadian jurisdiction or models from another Canadian jurisdiction. A related aspect of the research is the investigation of models for crash types. There is no safety performance function (SPF) available in the HSM to predict rear end collisions. Instead, rear end collisions are estimated as a proportion of predicted multivehicle collisions. To overcome this deficiency, Toronto data are used in the estimation of models for rear end collisions.

Acknowledgment

The author would like to express his deepest appreciation to his supervisor Dr. Bhagwant Persaud, for his invaluable advice, guidance, suggestions, patience and encouragement throughout the execution of this research program. His unfailing optimism and constant encouragement always prompted the author to overcome the difficulties in completing this research.

The author would like to express his sincere thanks to the City of Toronto, the City of Edmonton for providing data for this research.

The author is also thankful to the Civil Engineering Department in Ryerson University.

Finally, the author is grateful to his family, colleagues, and all friends for their support, and encouragement throughout this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

2U Two-lane undivided facilities

3SG Three Legged Signalized Intersection

3ST Three Legged Unsignalized Intersection

3T Three-Lane Facilities with centre two-way left-turn lanes

4SG Four Legged Signalized Intersection

4ST Four Legged Unsignalized Intersection

4U Four-lane undivided facilities

4D Four-lane divided facilities

5T Five-lane facilities with centre two-way left-turn lanes

AADT Annual Average Daily Traffic

CMF Collision Modification Factor

APM Accident Prediction Model

a,b,c,d,e Regression Coefficients

CURE Cumulative Residuals

Cr Calibration Factor

EB Empirical Bayes

F&I Fatal and Injury (Collisions)

GLM Generalized Linear Modeling

GOF Goodness-Of-Fit

HSM Highway Safety Manual

Pr Significance Level (p-value)

K Overdispersion Parameter

MAD Mean Absolute Deviation

MTO Ministry of Transportation, Ontario

n Number of years of collisions

NB Negative Binomial

N_{brmv} Multi Vehicle Collisions at road Predicted using SPFs

N_{brsv} Single Vehicle Collisions at road Predicted using SPFs

N_{drwy} Driveway Related Collisions Predicted

N_{pedr} Vehicle-Pedestrian Collisions Predicted

N_{biker} Vehicle-Bicycle Collisions Predicted

N_{bimv} Multi-Vehicle Collisions Predicted at Intersections

N_{bisv} Single-Vehicle Collisions Predicted at Intersections

N_{spf} Collisions Predicted by using SPFs

PDO Property Damage Only (collisions)

RTM Regression-To-Mean

SAS Statistical Analysis Software

SPF Safety Performance Function

1.0 Introduction

The Highway Safety Manual (HSM; AASHTO, 2010) documents state-of-the-art analytical tools for safety management processes and contains collision prediction methodologies to assess the safety of a design. The HSM is based on research which has involved more than seven National Cooperative Highway Research Program (NCHRP) projects since 2001 at a cost of $3.0 million. The scope of the HSM is to provide analytical tools and techniques to quantify the safety effects of decisions made in planning, design, operations, and maintenance.

Many jurisdictions have recognized the importance of assessment and implementation of the HSM predictive method, and have started research in this area. The research for this thesis also addresses this area, specifically the validity of the implementation of the HSM predictive method for Canada. The HSM provides important information and methodologies for practitioners to conduct highway safety analyses, including the prediction of expected collision frequencies for new and existing locations. The HSM facilitates evaluations of the safety impacts of alternate design scenarios by providing essential safety performance functions (SPFs) and collision modification factors (CMFs) for intersections and road segments.

The application of the HSM predictive method must first ensure that SPFs and CMFs recommended for use are compatible with the data available for specific jurisdictions. If they are not, then a recalibration of these tools is required. These needs define the scope and motivation of this research.

1.1 Scope

The scope of this research is limited to urban/suburban arterials. It is mainly focused on 3-legged and 4-legged signalized (3SG and 4SG) intersections. This research also includes 4-lane undivided road segments as shown below in Fig. 1.1

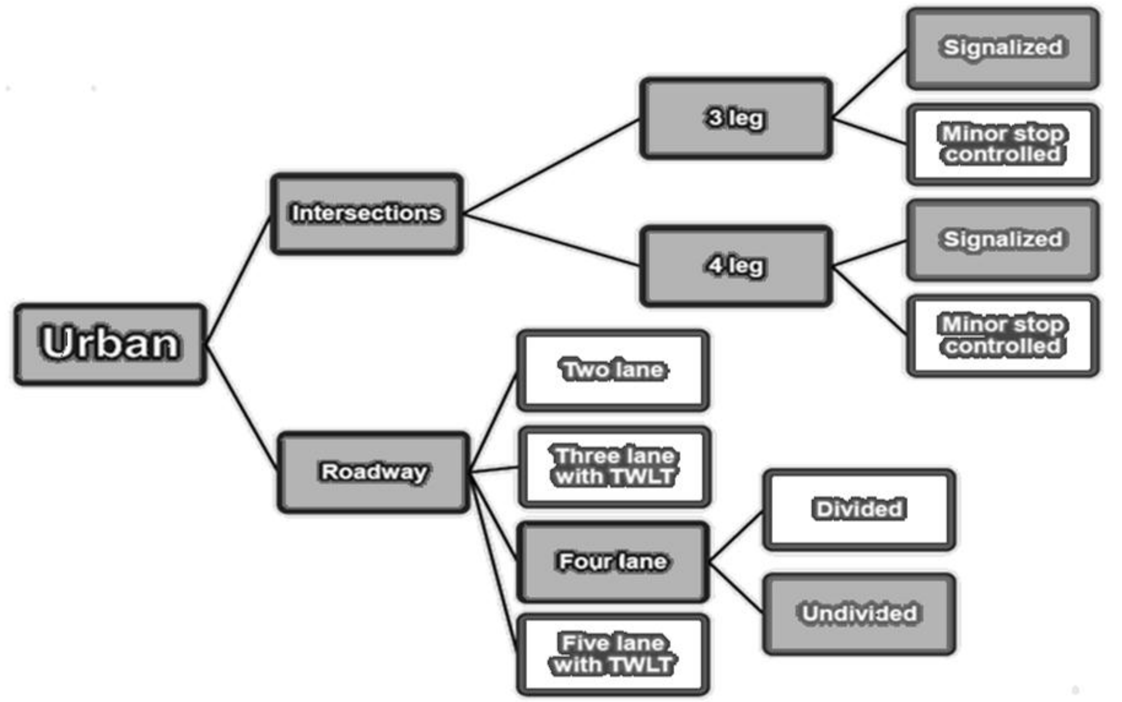


Figure 1.1: Division of urban/suburban facilities in the HSM

1.2 Research Objectives

The objective of this research is to evaluate the application of HSM predictive models for local conditions. The applicability of the HSM base models with Canadian collision data is not a given due to the many noticeable differences in collision data recording practices and driving behaviours between Canadian cities and the jurisdictions in which data were taken for developing the HSM models. In summary, the research objectives are:

- to calibrate HSM base models to local conditions, and

- to develop jurisdiction specific SPFs and compare them with the calibrated HSM models.

The purpose of base model calibration is to adjust the predictive base model to match the local conditions of different Canadian jurisdictions. The study investigates whether the methodology introduced by the HSM is practical, if the model predictions closely match the historical collision records, and whether the SPF variables are sensitive enough to be included in the collision prediction for local data. The scope of the application is limited to signalized intersections in urban/suburban areas.

1.3 Thesis Outline

This thesis is divided into eight chapters as follows.

Chapter 1- Introduction: introduces road safety techniques and SPFs, and outlines the objectives of the thesis.

Chapter 2- Review of the HSM: the HSM predictive method and prediction models are discussed.

Chapter 3-Literature Review: discusses the material related to collision prediction models, transferability of collision prediction models and statistical techniques used to analyse the collision data.

Chapter 4- Methodology: describes the procedure for calibration of collision prediction models and the application of goodness of fit (GOF) measures to assess the performance of collision prediction models.

Chapter 5- Application of Toronto Arterial Data to the HSM Predictive Method: describes the data requirements for the HSM predictive method, characteristics of available local data, and models used for collision predictions, and presents the calibration results.

Chapter 6- Statistical Analysis of Toronto Signalized Intersections: analysis is carried out to develop SPFs with data from local jurisdictions for use in the HSM predictive method.

Chapter 7- Transferability of HSM and Toronto Models to Edmonton Data: evaluates the application of the newly developed models for Toronto to the data of another Canadian jurisdiction (Edmonton) and compares them to the application of the HSM models and locally developed Edmonton models.

Chapter 8- Remarks and Conclusion: summarizes the results from the research, and provides recommendations for the application of the HSM predictive method in local Canadian jurisdictions.

2. Overview of Highway Safety Manual

The HSM has a science-based technical approach that takes the guesswork out of safety analysis. The HSM provides tools to conduct quantitative safety analyses, which allow for safety to be quantitatively evaluated alongside other transportation performance measures, such as traffic operations, environmental impacts, and construction costs.

The HSM provides the following tools:

• methods to develop an effective roadway safety management program and evaluation of its effects. A roadway safety management program is the overall process for identifying sites with potential for safety improvement, diagnosing problems for onsite conditions, evaluating conditions and identifying potential treatments at the sites, prioritizing and programming treatment types, and subsequently, evaluating the effectiveness of reducing crashes by the prescribed treatments,

• a predictive method to estimate crash frequency and severity. This method can be used to make informed decisions throughout the project development process, including: planning, design, operations, maintenance, and the roadway safety management process.

• a catalog of crash modification factors (CMFs) for a variety of geometric and operational treatment types, backed by robust scientific evidence. Several of the CMFs in the HSM have been developed by using high-quality before/after studies that account for regression to the mean.

2.1 Introduction to the HSM Predictive Method

The HSM predictive method provides a quantitative measure of the average expected crash frequency under both existing conditions and conditions which have not yet occurred. Part C in the HSM provides a predictive method to estimate the average expected crash frequency of a network, facility, or individual site, and introduces the concepts of SPFs and CMFs. For each facility type, prediction models for set base conditions are found. CMFs quantify the change in expected average crash frequency as a result of geometric or operational modifications to a site that differs from the set base conditions. The HSM predictive method can be applied to segments and intersections for the following facility types:

- rural two-lane, two-way roads,
- rural multilane highways, and
- urban and suburban arterials.

This research will focus on urban and suburban arterials.

2.2 Highway Safety Manual Models for Urban and Suburban Arterials

The predictive method addresses the following urban and suburban arterial facilities:

- two-lane undivided facilities (2U),
- four-lane undivided facilities (4U),
- four-lane divided facilities (4D),
- three-lane two-way left-turn lane facilities (3T), and
- five-lane two-way left-turn lane facilities (5T).

Separate models are provided to estimate intersection- and non-intersection-related crashes.

2.2.1 Segment Models

For segments, separate models are used to estimate the following types of collisions:

- multiple-vehicle non-driveway,
- single-vehicle,
- driveway-related,
- vehicle-pedestrian, and
- vehicle-bicycle.

The collision predictive models for roadway segments are as follows:

- multiple-vehicle non-driveway crashes

$$N_{brmv}=\exp\,(a+b\ln(AADT)+\ln(L)) \quad ....2\text{-}1$$

where

AADT= volume of annual average daily traffic on road segment

L=length of roadway segment (mi.)

a,b are the regression coefficients;

- single-vehicle crashes

$$N_{brsv}=\exp\,(a+b\ln(AADT)+\ln(L)) \quad ....2\text{-}2$$

where

AADT= volume of annual average daily traffic on road segment

L=length of roadway segment (mi.)

a,b are the regression coefficients; and

- multiple-vehicle driveway-related crashes

$$N_{brdwy}=\sum n_j N_j \left(\frac{AADT}{15000}\right)^t \quad ....2\text{-}3$$

where

N_j = number of driveway-related crashes per driveway per year for driveway type j

n_j = number of driveways within roadway segment of driveway type j including all driveways on both sides of the road

t = coefficient for traffic volume adjustment

The number of driveways is separately determined for each side of the road and then added together.

Driveway types are classified as:

- major commercial,
- minor commercial,
- major industrial/institutional,
- minor industrial/institutional,
- major residential,
- minor residential, and
- other.

Major driveways are those that serve sites with 50 or more parking spaces. Minor driveways are those that serve sites with less than 50 parking spaces. It is not intended that an exact count of the number of parking spaces be made for each site. Driveways can be readily classified as major or minor from a quick review of aerial photographs that show parking areas, or through user judgment based on the character of the establishment served by the driveway. Commercial driveways provide access to establishments that serve retail customers. Residential driveways serve single- and multiple-family dwellings. Industrial/institutional driveways serve factories, warehouses, schools, hospitals, churches, offices, public facilities, and other places of employment. Commercial sites with no restriction on access along an entire property frontage are generally counted as two driveways;

- vehicle-pedestrian crashes

$$N_{pedr} = N_{br}F_{pedr} \quad \text{.................}2\text{-}4$$

where

N_{br} = total number of crashes predicted excluding vehicle-pedestrian and vehicle-bicycle crashes

F_{pedr} = pedestrian crash adjustment factor; and

- vehicle-bicycle crashes

$$N_{biker} = N_{br}F_{biker} \quad \text{.............} \ 2\text{-}5$$

where

N_{br} = total number of crashes predicted excluding vehicle-pedestrian and vehicle-bicycle crashes

F_{biker} = bicycle crash adjustment factor.

CMFs are applied to adjust the HSM base model for local conditions. The CMFs for road segments are as follows:

- on-street parking,
- roadside fixed objects,
- median width,
- lighting, and
- automated speed enforcement.

2.2.2 Intersection Models

Separate models have been developed for the following types of intersections:

- 3-legged unsignalized (3ST),
- 4-legged unsignalized (4ST),
- 3-legged signalized (3SG), and
- 4-legged signalized (4SG).

For intersections, separate models are used to estimate the following types of collisions:

- multiple-vehicle,
- single-vehicle,
- vehicle-pedestrian, and
- vehicle-bicycle.

The collision predictive models for intersections are as follows:

- multiple-vehicle collisions

$$N_{bimv}=\exp\ (a+b\ln(AADT_{maj})+c\ln(AADT_{min})) \quad \text{............. 2-6}$$

where

$AADT_{maj}$=volume of annual average daily traffic on major roads

$AADT_{min}$=volume of annual average daily traffic on minor roads

a,b,c are the regression coefficients;

- single-vehicle collisions

$$N_{bisv}=\exp\ (a+b\ln(AADT_{maj})+c\ln(AADT_{min})) \quad \text{............. 2-7}$$

where

$AADT_{maj}$= volume of annual average daily traffic on major roads

$AADT_{min}$=volume of annual average daily traffic on minor roads

a,b,c are the regression coefficients;

- vehicle-pedestrian crashes

$N_{pedi} = \exp(a+bx\ln(AADT_{TOT}) + cx\ln(AADT_{min}/AADT_{maj}) + dx\ln(PedVol) + exn_{lanesx})$ 2-8

where

$AADT_{TOT}$ = sum of the AADTs on major and minor roads

n_{lanesx} = maximum number of lanes crossed by a pedestrian in any crossing manoeuvre at the intersection considering the presence of refuge islands; and

- vehicle-bicycle crashes

$$N_{bikei} = N_{bi}F_{bikei} \quad .... 2\text{-}9$$

where

N_{bi} = total number of crashes predicted excluding vehicle-pedestrian and vehicle-bicycle crashes

F_{bikei} = pedestrian crash adjustment factor.

CMFs available for intersections include:

- left-turn lanes,
- right-turn lanes,
- lighting,
- left-turn signal phasing,
- right-turn on red,
- red light cameras,
- bus stop locations,

- school locations, and
- alcohol sales establishments.

2.3 Overview of the Collision Prediction Algorithm

The HSM models are part of an algorithm that predicts the expected number of collisions for a site. The general form of the predictive model is as follows:

$$N_{predicted} = (N_{spf\,x} \times (CMF_{1x} \times CMF_{2x} \times \ldots\ldots\ldots\ldots \times CMF_{yx}) + N_{pedx} + N_{bikex}) \times C_x \quad \ldots\ldots\ldots\ldots 2\text{-}10$$

where

$N_{predicted}$ = predicted average crash frequency for a specific year for site type *x;*

$N_{spf\,x}$ = predicted average crash frequency determined for base conditions of the SPF developed for site type *x*;

N_{pedx} = predicted average number of vehicle-pedestrian collisions per year for site type *x*;

N_{bikex} = predicted average number of vehicle-bicycle collisions per year for site type *x*;

CMF_{yx} = collision modification factors specific to site type *x* and for specific geometric design and traffic control features *y*;

C_x = calibration factor to adjust SPF for local conditions for site type *x*.

This algorithm applies a base model and then refines the predictions of the base model by using CMFs. The base model predicts the expected number of collisions for sites that meet the base conditions. CMFs, which have been assembled by a team of experts and documented in the HSM, are then used to adjust the base model predictions to account for the effects of other variables that are subject to design decisions, i.e., conditions that are different from those of the base model.

3.0 Literature Review

3.1 Collision Modeling

The concept of modelling for collision predictions has been supported by recently carried out research, which have established the relation between collisions and traffic flow. Over the last few years, numerous road collision prediction models have been developed to investigate the effects of variables on collisions. A significant amount of research has described and explained the occurrence of road collisions. Some take into account the technical aspects of vehicles and infrastructures, while other viewpoints include psychological, behavioural and social-economic components. Milton et al. (2008) investigated the statistical properties of different regression models, and used Poisson and negative binomial (NB) regression models instead of a linear regression model to estimate the collision probability or frequency over a period of time.

The impacts of a variety of factors have also been investigated. For example, the relationship between collisions, weather and geometric elements was studied by Shankar et al. (1996). Independent variables such as the type and the quality of pavement, and the presence of park lanes and turn lanes were considered by Matthew et al. (2002). Collision prediction models typically use annual average daily traffic (AADT) as a significant variable to predict collisions (e.g., Persaud and Dzbik (1993) used a generalized linear model which shows a positive relationship between collision data and traffic flow). Collision risks for four lane freeways are found to be lower than freeways with more than four lanes under the same traffic volume.

The effect of the median width of four lane roads on collision rates was examined by Rodman et al. (1996) by using an NB distribution. Their study indicated that collision rate decreases with increasing median width, while wider medians are associated with a reduction in crossover collisions that involve head-on collisions between opposing traffic.

Lord and Persaud (2000) carried out one of the earliest studies which separately analyzed curves and tangents for road sections. The dependent variable used was collision frequency and the independent variables were traffic flow and road geometry. Their study indicated that collision frequency increases with AADT, section length (L) and curvature (1/R) for curves, and the number of collisions per year increases with AADT and the length of tangents. Abdel-Aty et al. (1998) used an NB distribution for collision frequency prediction. Collision frequency was estimated as a function of AADT; degree of horizontal curvature; L; width of lane, shoulders and median; and urban/rural designation. The model in this study also accounted for driver characteristics, including sex and age (young, middle aged and senior). The results showed that collision frequency increases with AADT, degree of horizontal curvature and L, and decreases with width of lane, shoulder and median.

Martin (2002) discussed the relationship between the collision rate and traffic volume per hour (VH). In this study, an NB distribution was used to represent the relationship between collision rate and VH. Higher collision rates were shown for both property damage-only and injury collisions when the VH was less than 400 V/h. The collision rate also rapidly decreased with increasing VH, passing through a minimum value of 1000 and 1500 v/h for two-lanes or three-lanes, respectively; after that, the collision rate gradually increased with an increase in traffic.

Golob and Recker (2003) used linear and non-linear multivariate statistical analyses to determine the factors that were related to the type of collision (such as traffic flow, weather, and lighting conditions). The approach of their study was to identify the most significant variables from the factors. Hauer (2004) developed a statistical road safety model by using NB distribution, in which the dependent variable was the number of collisions per year while the independent variables were traffic flow and geometric characteristics. In a later study, Hauer (2007) applied the above statistical model to estimate collision frequencies on undivided four lane urban roads. This model evaluated the number of collisions per year as a function of the following independent variables: AADT, percentage of trucks, degree and length of horizontal curve, grade of tangents, length of vertical curve, lane width, shoulder width and type, road side hazard rating, speed limit, access points (e.g. signalized intersections, stop-controlled intersections, commercial driveways and other driveways), and the presence and nature of both parking and two-way left-turn-lanes (TWLTL). The findings showed that the variables that had a significant effect are: AADT, the number of commercial driveways, and speed limit. Caliendo et al. (2007) also focused on variables that are related to traffic flow, geometric structure, pavement surface, and rainfall.

A model for injury and damage-only collisions at urban intersections in Canada was developed by Persaud et al. (2002), which described the relationship between collision risk and traffic attributes, including traffic volume. The time-series collision data from the study explicitly revealed temporal changes in safety conditions and enabled a comparison of the safety performance of junction types across different cities. Oh et al. (2003) compared the safety performance of single point and tight diamond intersections by using traffic flow information and collision data. They found no significant differences between the total numbers of collisions at the two types of intersections, but the differences in the frequency of injuries and fatalities are significant, with the single point intersection being apparently safer than the tight diamond intersection. Lyon et al. (2005), Wong et al. (2007), Schattler and Datta (2004), Kim et al.

(2005), Lum and Halim (2006), and Zhang and Prevedouros (2003) established SPFs and conducted specific collision prediction models, particularly for signalized intersections.

Lyon et al. (2005) described the development of SPFs for urban signalized intersections based on 5 years of collision data in Toronto. Separate SPFs were developed for 3SG and 4SG intersections, several impact types (rear-end, right angle, left turn, all combined) and severity levels, property damage only collisions (PDO), and fatal plus non-fatal injuries (F&I). They also found that the number of lanes on an approach is strongly correlated with the number of turning lanes.

3.2 Model Transferability

Different studies have been recently carried out to evaluate the validity and transferability of collision prediction models to other jurisdictions. We have reviewed these studies as follows.

3.2.1 Application of HSM Draft Chapter in Louisiana

Sun et al. (2006) applied the HSM calibration procedure to a road network in Louisiana state, which accounted for 13,400 miles of two-lane rural highways. The study investigated whether the methodology introduced by the HSM draft chapter is practical, whether the modelling results closely match historical collision records, and if the variables are sensitive enough to be included in the model in consideration of the data collection cost. The scope of the application was limited to segments of two-lane rural highways.

The calibration process did not follow the HSM recommended procedure because of the unavailability of data. The average predicted values were uniformly less than the average observed collision frequencies at all levels of AADT. The highest calibration parameter was 2.28

for an AADT less than 1,000 vehicles per day, and the lowest was 1.49 for an AADT greater than 10,000 vehicles per day.

3.2.2 Calibration of HSM Collision Prediction Model for Italian Secondary Road Networks

Martinelli et al. (2005) investigated how the calibration procedure of an HSM model can be applied to the road networks in the Italian province of Arezzo. This was carried out to evaluate the effective transferability of an HSM methodology to a region characterized by a different environment and different road characteristics, driver behaviour, and collision reporting systems in contrast to those on which the HSM model has been developed. The considered road network covered 1,300 km of two-lane rural highways, and a 3-year (2002–2004) collision database was used. There was a considerable difference between the road networks of Arezzo and Minnesota, on which the HSM model had been developed. When the HSM calibration procedure was applied to the two-lane highway road networks in Arezzo, it was noted that all the different possible calibration procedures resulted in an overestimation of the low collision sections and an underestimation of the high collision sections, which led to the conclusion that a constant calibration factor is not a realistic option for model transferability.

3.2.3 Calibration and Transferability of Collision Prediction Models for Urban Intersections

Persaud et al. (2002) discussed the transferability of collision prediction models to other jurisdictions. Toronto data were used to estimate models for three- and four-legged signalized and unsignalized intersections. Then, the performances of these models were compared with those of models from Vancouver and California, which were recalibrated for Toronto. The results were mixed, which suggest that a single calibration factor may be inappropriate and disaggregation by traffic volume might be preferable. The study concluded that the California models are not applicable for use with Toronto data as they generally predict more collisions with the latter when the minor AADT is low. A GOF test for a three legged intersection model with an overdispersion (k) of 2.34 in British Columbia indicated that the model fits the Toronto data better than the one for four-legged intersections in which k is 2.17.

3.2.4 Collision Prediction Models With and Without Trends - Application of Generalized Estimating Equation Procedure

Lord and Persaud (2000) discussed the application of a generalized estimating equation (GEE) procedure for calibrating collision prediction models when several years of data are available and when it is desirable to incorporate trends. The application was for a sample of 4SG intersections in Toronto by using data from 1990 to 1995. Generalized linear modeling (GLM) does not account for temporal correlation in the collision count data. The results from this study demonstrated that failure to account for temporal correlation does not affect the coefficient estimates, but considerably underestimates the variances of the estimates. This means that explanatory variables may be incorrectly attributed as significant if temporal correlation is not considered. The results also showed that accident prediction models (APMs) which incorporate time trends usually perform better than those that do not.

3.2.5 Pedestrian Collision Prediction Models for Urban Intersections

Lyon and Persaud (2002) developed aggregate collision prediction models for vehicle pedestrian collisions at urban intersections by using vehicle and pedestrian volumes as explanatory variables. One of the objectives was to test the transferability of the models, given that many jurisdictions do not have the resources or the data with which to calibrate them. Data were collected for 4SG intersections in the City of Hamilton for this purpose.The data collected from Hamilton for 4SG intersections were used to test the transferability of the Toronto models. The

Toronto model for 4SG intersections was modified for application to Hamilton by calculating a calibration factor that was added to the model as a multiplier. This calibration factor was calculated by dividing the summation of observed collisions over all intersections in the Hamilton data by the summation of predicted collisions obtained by applying the original Toronto model to the Hamilton data. In applying the Toronto model to the Hamilton data, the calibration factor, which was estimated at 1.17, was multiplied by the predictions from the Toronto model. The transferability of the 4SG intersection model to a different jurisdiction was tested and proved to be successful.

3.3 Statistical Approach for Modeling

Many models have been developed to predict collisions on rural and urban intersections. Three main statistical approaches are usually attempted by researchers to relate collisions to geometric and traffic related explanatory variables. These are:

- multiple linear regression,
- Poisson regression, and
- NB regression.

Multiple linear regressions, which are estimated by ordinary least squares, are the most straightforward approach. However, the model lacks distributional properties to adequately describe random, discrete, non-negative, and typically irregular vehicle collision events on the road. The Poisson model assumes that the ratio between the mean and variance is equal to 1. However, collision data are found to be significantly over dispersed relative to their mean. Therefore, the use of Poisson regression models may overestimate or underestimate the likelihood of vehicle collisions. A standard generalization of the Poisson model is NB distribution. NB regression provides a common tool to model cross-sectional count data, like collision frequencies at signalized intersections. Both Poisson and NB regressions are special cases of GLM.

3.3.1 Statistical Models of At-Grade Intersection Collisions

Harwood et al. (2000) developed NB regression models to fit collision data at three- and four-legged stop-controlled rural intersections and three-legged stop-controlled urban intersections. Lognormal regression models were found to be more appropriate for modelling collisions at four-legged stop-controlled and four-legged signalized urban intersections. The decision to use an NB or lognormal regression analysis was based on an evaluation of the collision frequency distribution for specific categories of intersections. Generally, multiple-vehicle collisions represent a large proportion of all collisions, except for those that occur at three-legged stop-controlled rural intersections. With large numbers of intersections that have no or low collision experience, the distribution tends to follow the shape of a Poisson distribution. When the number of intersections with no or low collision experience is relatively small, the distribution tends to follow the shape of a lognormal distribution. This is clearly seen in the case of four-legged signalized urban intersections and four-legged stop-controlled urban intersections. The NB model overcomes a limitation of the Poisson distribution in that the latter assumes that the mean equals the variance of the distribution, whereas in modelling collisions, the variance or dispersion of the data exceeds the estimated mean of the collision data distribution. The data are then said to be over dispersed, and the underlying assumption of the variance being equal to the mean for the Poisson distribution is violated. The NB, which is a discrete distribution, provides an alternative model to deal with k in the count data, such as collision frequencies. Statistical analysis software (SAS) provides a procedure for GLM (PROC GENMOD) that can be used to estimate regression coefficients by specifying the appropriate type of distribution.

3.3.2 A Review of Collision Prediction Models for Road Intersections

Collisions are random and discrete with nonnegative values. For this type of dependent variable prediction, as shown in Nambuusi et al. (2008), the Poisson regression model can be used when discrete response variables have counts as possible outcomes. There is no fixed upper limit for the outcome. Since the outcome must be a nonnegative integer, its distribution should cover a nonnegative range and the simplest of such a distribution is a Poisson distribution.

3.3.3 Methodology to Predict the Safety Performance of Urban/Suburban Arterials

Hauer et al. (2007) developed regression models to use as base models for the HSM methodology. Functional relationships between the expected number of roadways or intersection collisions and explanatory variables, including AADT and roadway segment or intersection characteristics, were developed by using a stepwise multiple regression approach, which assumed an NB error distribution of collision frequencies. The models were individually developed within each U.S. state and for a combination of states, and separately for each roadway, intersection and collision type.

To develop regression models from the data in individual states, SAS was used to estimate the model coefficients, k, the standard error of each estimate, and their significance levels (p-value). This procedure fits a generalized linear model to the data by a maximum likelihood estimation of the regression and k parameters; an NB error structure of collision counts is assumed. In developing regression models with the combined data from more than one state, a slightly different modelling approach was used. In this case, it was assumed that two states provide a random sample of data from the pool of all states. Thus, state was included in the NB regression as a random factor rather than a fixed factor like the other variables. This approach allows one to estimate the effect of each fixed effect while controlling for the random state effect. All other variable and model selection criteria were kept identical to those used in the modelling approach for the individual state. To develop models from combined state data, SAS was used to estimate the model coefficients "k" standard error of each estimate, and their significance levels (p-value), by assuming an NB distribution of collisions and a random error structure within and between states.

3.4 Summary

Different road collision prediction models have been developed to investigate the effects of different variables on collisions. The investigated impacts of a variety of factors have established that collisions and traffic flow have a significant relationship. In the HSM, separate SPFs are developed for 3SG and 4SG intersections. Several collision impact types and severity levels, PDO, and F&I are considered in collision modelling that have been recently carried out.

In recent research, HSM recommended procedures for collision prediction and model transferability are investigated for local jurisdictions. For example, an HSM calibration procedure has been applied to road networks in the Italian province of Arezzo to evaluate the effective transferability of this methodology to a region characterized by different environment and road characteristics, driver behaviour, and collision reporting systems from the characteristics on which an HSM model has been developed. When an HSM calibration procedure was applied to two-lane highway networks in Arezzo, it was noted that all of the different possible calibration procedures resulted in an overestimation of the low collision sections and an underestimation of the high collision sections, which led to the conclusion that a constant calibration coefficient is not a practical option for model transferability.

To date, there has been no research that evaluates the implementation of the HSM predictive methodology in Canada and the transferability and applicability of HSM signalized intersection models to local jurisdictions. Therefore, this research aims to fill these research gaps by evaluating the HSM predictive methodology for Toronto urban/suburban arterials.

4.0 Methodology

4.1 HSM Predictive Models

The HSM predictive method consists of:

- SPFs – regression equations used to estimate the predicted collision frequency at a site for a given "base condition",
- CMFs – used to adjust the "base condition" in an SPF to specific site characteristics, and
- Calibration factors (Cr) – adjust average collision frequencies calculated from an SPF to local site conditions.

There are two options in the development and use of the predictive method:

- calibrate HSM base models to local conditions (e.g. Toronto), and
- develop jurisdiction specific SPFs.

Some HSM users may prefer to develop SPFs with data from their own jurisdiction to use with the predictive method rather than calibrating the SPFs in the HSM. The HSM Appendix to Part C provides guidance on developing jurisdiction-specific SPFs that are suitable for use with the predictive method.

4.2 Procedure for Calibration

The accurate calibration of the predictive model is the first key step toward a successful HSM application. Model calibration accounts for differences caused by climate, animal population, driver population and trip purpose, collision reporting threshold, and collision investigation practices. It is not meant for geometric design factors that are justified by the safety prediction methodology. As presented in the HSM, the calibration procedure involves five steps:

- the identification of the facility type to be calibrated. The City of Toronto urban/suburban arterials and signalized intersection collision data are used for this research;
- the selection of sites for calibration. Approximately 1900 intersections were selected for this research, in which approximately 1700 are 4SG intersections and 200 are 3SG intersections in the City of Toronto;
- the obtaining of data for each facility type, i.e. collision data, major AADTs, minor AADTs, number of lanes and site characteristics;
- the application of an appropriate (Part C) predictive model to predict collisions as: $Npredicted = Nspf\ x \times (CMF1x \times CMF2x \times \ldots \times\ ) \times Cx$; and
- the computing of the calibration factors: separate calculations of Cx for roadway segments and intersections. A proportional factor that is applied to regional data is estimated as follows:

$$Cr(orCi) = \frac{\sum ObservedCrashes}{\sum \Pr edictedCrashes}$$

where

Cr = calibration factor for road segments

Ci = calibration factor for intersections.

4.3 Jurisdiction Specific SPFs

SPFs should be developed with a statistical technique, such as NB regression, which accounts for the k typically found in collision data, and quantifies the k parameters so that model predictions can be combined with observed collision frequency data by using an empirical Bayes (EB) method. Jurisdiction-specific SPFs should use the same base conditions as the corresponding

SPFs in Part C of the HSM or capable of being converted into such base conditions. Jurisdiction-specific SPFs should include the effects of the following traffic measures: volume of AADT for roadway segments, and volume of major and minor-road AADTs for intersections. Finally, jurisdiction-specific SPFs for any roadway segment and facility type should have a functional form in which the average predicted collision frequency is directly proportional to segment length.

4.3.1 Development of Jurisdiction Specific SPFs

The Statistical Analysis Software (SAS) package was used in this research to estimate the SPF coefficients by assuming an NB error distribution, which is now standard practice in developing regression models. The dispersion parameter, k, which relates the mean and variance of a regression estimate, is iteratively estimated from the model and the data. The value of k is such that a smaller value means a better model for a given set of data.

4.4 Goodness-of-Fit (GOF)

To examine how well a statistical model fits the data set, GOF measurements were examined. GOF measurements summarize the differences between the observed and predicted values from related SPFs. Several GOF measures are used to assess performance, including:

- "k", the value of the overdispersion parameter;
- the mean absolute deviation (MAD), average of the absolute value of observed minus predicted collision frequencies for each site;
- cumulative residual (CURE) plots; and

- comparisons of the ratio of observed to predicted values, summarized by the categories of variables of interest.

CURE plots are used to judge how well predictions fit the data over the full range of an independent variable. In this method, as documented by Hauer & Bamfo (1997), the cumulative residuals (difference between the observed and predicted collisions for each location) are plotted in increasing order for each covariate, e.g. AADT, separately. Also plotted are graphs with 95% confidence limits. If there is no bias in the model, the CURE plots should stay inside these limits.

5.0 Application of Toronto Urban/Suburban Arterial Data to the HSM Predictive Method

A database of urban arterial roads was made available by the City of Toronto for this study. This database is quite extensive and unique for a city jurisdiction.

5.1 Unsignalized Intersections

Toronto does not currently have an unsignalized intersection database. Collisions at unsignalized intersections are included in the current arterial segment data. It would be possible to create such a database by first observing which unsignalized locations have traffic counts (but not all will), then obtaining the geographic information system (GIS) coordinates to match to collisions (collisions within 25 m are assigned to intersections for a signal database). Information on turning lanes could be done in the field or by using Google maps. Lighting data could be collected in the field or assumptions can be made.

5.2 Signalized Intersections

A database of signalized intersections is available with many of the required variables. Missing variables of relevance to the HSM procedure include:

- lighting,
- left-turn phasing,
- right-turn-on-red,
- red-light cameras,
- presence of pedestrian refuge islands,

- bus stops within 1000 ft,
- schools within 1000 ft, and
- alcohol establishments within 1000 ft.

Note that bus stop, school and alcohol establishment CMFs only apply to vehicle-pedestrian collisions.

5.3 Segments

A database of arterial road segments exists in which each segment is defined as starting and ending at a signalized intersection or road end where applicable. Various variables are given an aggregated 'score' for the segment, e.g. for cases where on-street parking is not present on the total length of the segment. These variables include:

- median type,
- presence of TWLTL, and
- on-street parking.

The number of driveways is present in the data, but not broken down by type, as required by the HSM procedure for driveway-related collisions.

Missing variables of relevance to the HSM procedure include:

- median width,
- roadside fixed objects,
- lighting, and
- automated speed enforcement.

These missing variables are not considered in the collision prediction as CMFs as they are assumed to be a base condition equal to 1. The current Toronto segment data includes collisions at unsignalized intersections.

All the above-mentioned details are summarized in Table 5.1.

Table 5.1: Toronto Data Characteristics

Site Type	HSM Site Description	Variables Required for Applying SPFs	Variables Required for Applying CMFs	Model Output	Comments with regards to Toronto Inventory Data	Comments with regards to Toronto Collision Data
Arterial Segment	From intersection centre to the next intersection centre, or where there is a change in a homogeneous segment. Length is measured from the centre of the intersection.	Divided/ undivided/ TWLTL, AADT, Segment length, Number of driveways by type (major commercial, minor commercial, major industrial, minor industrial, major residential, minor residential, other), Posted speed	On-street parking, Roadside fixed objects, Median width, Lighting, Automated speed enforcement	Non-intersection-related collisions by collision type: multi-vehicle non-driveway; single-vehicle; driveway-related; vehicle-pedestrian; vehicle-bicycle	Median type is in a median score format, thus segments may have both divided and undivided parts. The same applies for TWLTLs and on-street parking. Does not have median width, roadside fixed objects, and lighting. Assume no automated speed enforcement. Have number of driveways, but not broken down by type.	Data presently assigned to segments include collisions at unsignalized intersections, excludes those within 20 m of the intersection. Collision data have intersection-related collisions and at/near private-driveways

Site Type	HSM Site Description	Variables Required for Applying SPFs	Variables Required for Applying CMFs	Model Output	Comments with regards to Toronto Inventory Data	Comments with regards to Toronto Collision Data
Signalized Intersections	Collisions at signalized intersections are those within the intersection limits and those on approaches that are related to the intersection.	Major road AADT, Minor road AADT, Pedestrian volume (numbers provided for rough estimates of pedestrian demand), Maximum number of traffic lanes crossed by a pedestrian in consideration of the presence of refuge islands	Left-turn lanes, Right-turn lanes, Lighting, Left-turn phasing, Right-turn-on-red, Red-light-cameras, Bus stops within a 1000’, Schools within a 1000’, Alcohol sales establishments within 1000’	Intersection-related collisions by collision type: multiple-vehicle; single-vehicle; vehicle-pedestrian; vehicle-bicycle	Does not have lighting, left-turn phasing, right-turn-on-red, red-light cameras, presence of pedestrian refuge islands, or bus stop , school and alcohol establishment information	In current database, includes those within 25 m of intersection.
Un-Signalized Intersection	Collisions at unsignalized intersections are those within the intersection limits and those on approaches that are related to the intersection.	Major road AADT, Minor road AADT	Left-turn lanes, Right-turn lanes, Lighting	Intersection-related collisions by collision type: multiple-vehicle; single-vehicle; vehicle-pedestrian; vehicle-bicycle	No unsignalized intersection database exists. This would need to be created. Requires latitude/longitude, AADTs, turn lanes and lighting information.	Unsignalized intersection collisions currently assigned to segments.

5.4 City of Toronto Signalized Intersections

Six years of collision data from 1999 to 2004 that pertain to approximately 1900 sites were used for calibration with the HSM predictive method in Part C of Chapter 12. This dataset is very rich and includes 3SG and 4SG intersections.

5.5 Summary Statistics

In the 5 year Toronto dataset, the collision types are angle, approach, rear end, side swipe, multi vehicle and single vehicle. In the 6 year dataset, multivehicle, single vehicle and pedestrian and bike collisions are available. Details are provided in Tables 5.2 and 5.3.

Table 5.2: Toronto Collision Data from 3-Legged Signalized Intersections

1999 to 2004 Toronto Collisions							
Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collisions	57.219	6	277	137	7839	100	6
F&I Collisions	15.285	0	67	137	2094	26.72	6
(PDO) Collisions	41.934	2	210	137	5744	73.28	6
Multi veh. Collisions	52.839	0	262	137	7238	92.34	6
Multi veh. F&I Collisions	12.328	0	53	137	1688	21.54	6
Multi veh. PDO Collisions	40.511	2	209	137	5550	70.79	6
Single veh. Collisions	1.204	0	8	137	164	2.10	6
Single veh. F&I Collisions	0.153	0	4	137	21	0.26	6
Single veh. PDO Collisions	1.05	0	8	137	143	1.83	6
Pedestrian Collisions	1.898	0	11	137	260	3.31	6
Bike Collisions	1.277	0	7	137	174	2.23	6

2000 to 2004 Toronto Collisions							
Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collisions	43.547	3	232	137	5965	100	5
Fatal Collisions	0.029	0	1	137	4	0.07	5
Injury Collisions	11.766	0	49	137	1612	27.01	5
F&I Collisions	11.795	0	49	137	1616	27.08	5
PDO Collisions	31.752	2	183	137	4350	72.91	5
Angle Collisions	8.314	0	52	137	1139	19.09	5
Approach Collisions	1.139	0	7	137	156	2.62	5
Rear End Collisions	12.898	0	85	137	1767	29.62	5
Side Swap Collisions	6.869	0	46	137	941	15.77	5

Table 5.3: Toronto Collision Data from 4-Legged Signalized Intersections

1999 to 2004 Toronto Collisions							
Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collisions	70.292	1	378	1691	118863	100	6
F&I Collisions	21.754	0	125	1691	36786	31.00	6
PDO Collisions	48.537	0	268	1691	82076	69.05	6
Multi veh. Collisions	64.997	0	370	1691	109909	92.46	6
Multi veh. F&I Collisions	17.691	0	120	1691	29915	25.16	6
Multi veh PDO Collisions	47.306	0	268	1691	79994	67.29	6
Single veh. Collisions	1.222	0	9	1691	2066	1.78	6
Single veh. F$I Collisions	0.243	0	4	1691	410	0.34	6
Single veh. PDO Collisions	0.979	0	8	1691	1655	1.39	6
Pedestrian Collisions	2.788	0	22	1691	4714	3.96	6
Bike Collisions	1.284	0	16	1691	2171	1.82	6

2000 to 2004 Toronto Collision Data							
Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collisions	55.247	1	313	1691	93422	100	5
Fatal Collisions	0.068	0	3	1691	114	0.12	5
Injury Collisions	17.054	0	100	1691	28953	30.86	5
F&I Collisions	17.122	0	100	1691	64467	31.00	5
PDO Collisions	38.124	0	229	1691	64467	69.00	5
Angle Collisions	8.266	0	60	1691	13977	14.96	5
Approach Collisions	1.445	0	15	1691	2443	2.61	5
Rear End Collisions	20.199	0	195	1691	34156	36.56	5
Side Swap Collisions	6.224	0	59	1691	10524	11.26	5

5.6 Prediction of Collisions by Using HSM Models

5.6.1 Toronto Intersection Collision Data

In the HSM, separate base models are used for collision prediction of 3SG and 4SG intersections, as there is no collision data for unsignalized intersections.

Two sets of signalized intersections are selected as follows:

- 3SG intersections, and
- 4SG intersections.

The HSM predictive models, described in Chapter 2, were applied to the data.

5.6.2 Collision Modification Factors for Intersections

The effects of individual geometric design and traffic control features of intersections are represented in the predictive models by CMFs. CMFS*1i* to CMFS*4i* are applied to multiple-vehicle and single-vehicle collisions at intersections, but not to vehicle-pedestrian and vehicle-bicycle collisions. CMFs that were used for intersection collision calibration include the following.

5.6.2.1 Left-Turn Lanes

The base condition is the absence of left turn lanes.

As left turn lane data are available, the CMFs from the HSM are used, as shown in Table 5.4.

Table 5.4: Collision Modification Factor (CMF1i) for Installation of Left Turn Lanes on Intersection Approaches

Intersection Type	Traffic control	Number of approaches with left turn lanes			
		One	Two	Three	Four
3-legged signalized intersections	Minor road STOP control	0.67	0.45	_	_
	Traffic signal	0.93	0.86	0.8	_
4-legged signalized intersections	Minor road STOP control	0.73	0.53	_	_
	Traffic signal	0.9	0.81	0.73	0.66

5.6.2.2 Right-Turn Lanes

The base condition is the absence of right turn lanes.

As right turn lane data are available, the CMFs from the HSM are used, as shown in Table 5.5.

Table 5.5: Collision Modification Factor (CMF3i) for Installation of Right Turn Lanes on Intersection Approaches

Intersection Type	Traffic Control	Number of approaches with right turn lanes			
		One	Two	Three	Four
3-legged signalized intersections	Minor road STOP control	0.86	0.74	_	_
	Traffic signal	0.96	0.92	_	_
4-legged signalized intersections	Minor road STOP control	0.86	0.74	_	_
	Traffic signal	0.96	0.92	0.88	0.85

5.7 Results

By using the HSM SPFs, collisions are predicted for each type of intersection and the calibration factors are calculated as follows.

5.7.1 Three Legged Signalized Intersections

Total 246 collisions are predicted for one year. Table 5.6 shows the calibration results. The total number of predicted collisions is about 18% of the observed number of collisions.

Table 5.6: Calibration Factors for Toronto 3-legged Signalized Intersections

	Pedestrian collision	Bike collision	Single vehicle collision	Multi vehicle collision	Total collision	Rear end collision
Observed	43	29	27	1206	1307	354
Predicted	5.88	2.61	20.64	215.35	246	628
Cr	7.4	11.17	1.33	5.60	5.34	0.56

In Table 5.6, the Cr value of single vehicle collisions is near 1, which shows that the HSM SPFs are a good fit for the prediction of single vehicle collisions in Toronto. The number of predicted collisions is much less for multi vehicle collisions, which shows the same behaviour as in the case of the total number of predicted collisions. In addition, the number of observed pedestrian and bike collisions are substantially higher in comparison to the predicted collisions. Rear end collisions, which are a component of multi vehicle collisions, are overpredicted i.e., the Cr value is less than 1.

5.7.2 Four Legged Signalized Intersections

In Table 5.7, the total number of predicted collisions is about 21% of the observed collisions. For the other collision types, the same trend is observed for the 4SG collisions as found in the 3SG collisions.

Table 5.7: Calibration Factors for Toronto 4-legged Signalized Intersections

	Pedestrian collision	Bike collision	Single Vehicle collision	Multi vehicle collision	Total collision	Rear End collision
Observed	785	362	344	18318	19810	6831
Predicted	61.37	29.26	279.80	3832.97	4203.43	9278
Cr	12.80	12.37	1.23	4.77	4.71	0.73

5.8 Summary of the Intersection Results

Without recalibration, the application of the HSM algorithm results in an underestimation of collision predictions for the signalized intersections in the Toronto data. The calibration factors to adjust the HSM models in local conditions of urban intersections are 4.7 for 4SG intersections, and 5.34 for 3SG intersections for the total number of observed collisions. The Cr value is very close to 1 for single vehicle collisions in 3SG and 4SG intersections. The rear end collisions are overestimated with a Cr value that is less than 1 for both 3SG and 4SG intersections.

5.9 Toronto Arterial Road Segment Collision Data

The Toronto arterial data consist of a variety of road segments and varying geometric and operational characteristics. The data includes 2412 segments with a total length of 103,977 m.

5.9.1 Types of Segments

In the Toronto arterial data, the road segment distribution is as follows:

- 2U =232 segments,
- 3T = 46 segments,
- 4U =1229 segments,
- 4D = 444 segments, and
- 5T = 48 segments.

Fifty percent of the data consist of 4U; therefore 4U segments are selected to apply the HSM predictive method.

5.9.2 CMFs for Segments

- CMFS1r is for on-street parking

The base condition is the absence of on-street parking on a roadway.

- CMFS2r is for roadside fixed objects

The base condition is the absence of roadside fixed objects on a roadway segment.

- CMFS3r is for median width

This is a CMF for median widths on divided roadway segments of urban and suburban arterials. The base condition for CMFs is a median width of 15 ft. The value of this CMF is 1.00 for undivided facilities.

- CMFS4r is for lighting

The base condition for lighting is the absence of roadway segment lighting.

- CMFS5r is for automated speed enforcement.

The base condition for automated speed enforcement is that it is absent.

5.10 Road Segments That Include Unsignalized Intersections

From the arterial collision data, the first subset selected is a 4-lane undivided major arterial road, Victoria Park Avenue, in the city of Toronto. This corridor also includes unsignalized intersections. This subset consists of 19 segments that range in length from 115 to 640 metres with an AADT range of 19,000 to 43,000 vehicles per day and posted speed of 60 km per hour.

A summary of the collisions and other road characteristics are given in Table 5.8 as follows.

Table 5.8: Summary of Urban 4-Lane Undivided Segments that Include Unsignalized Intersections

Variable	Min.	Max.	Mean	Frequency
Segment length (km)	0.115	0.664	0.351	16
AADT (Vehicles/day)	19,768	42,459	30,730	16
Driveway/km	5.14	121.5	40.45	16
All Severity Collisions/year	11	203	50.6	16
Fatal collisions in 5 years	0	1	0.063	16
Injury collisions in 5 years	6	72	16.75	16
PDO collisions in 5 years	4	131	33.82	16

The equations in the HSM predictive method are applied to predict collisions. Major AADTs and segment lengths are available in the Toronto collision data. Driveway related collisions, and major and minor industrial, institutional and residential driveways were counted by using Google map. In the Toronto data, major residential and commercial driveways are given, but the information is not very accurate. Therefore, to predict driveway related collisions, it was necessary to count all the driveways in the field as per the HSM given definitions on both sides of the road segments. The calibration factors calculated are shown in Table 5.9.

Table 5.9: Calibration Factors for Road Segments Including Unsignalized Intersections in the Toronto Data

Observed	810
Predicted	285
Cr	2.8

In Table 5.9, the number of predicted collisions is about 34% of the number of observed collisions. However, the total number of predicted collisions is only for road segments while the Toronto data also include collisions at unsignalized intersections. To handle this issue, we selected a subset of the arterials, which have zero access points; that is, there are no unsignalized intersections. This set of road segments only contains segment related collisions. The HSM predictive method was applied to these segments in subset 2.

5.11 Road Segments That Exclude Unsignalized Intersections

The Toronto arterial data contain unsignalized intersections; therefore, to extract data which have only road-related collisions, road segments with zero access are selected. These selected segments have 4-lanes, are undivided and have no access road from start to end. These segments range in length from 50 to 616 metres. The posted speed varies from 50 km/h to 70 km/h. Most of them are in residential areas. A summary of the collisions and other characteristics are given in Table 5.10 as follows.

Table 5.10: Summary of Urban 4-Lane Undivided Segments in the Toronto Data

Variable	Min.	Max.	Mean	Frequency
Segment length (km)	0.049	0.616	336	46
AADT (Vehicles/day)	6,320	39,688	26,301	46
Driveway/km	0	180.86	23.75	46
All severity collisions/year	0	85	23	46
Fatal collisions in 5 years	0	1	0.044	46
Injury collisions in 5 years	0	30	7.11	46
PDO collisions in 5 years	0	55	15.84	46

The calibration factors are shown in Table 5.11.

Table 5.11: Calibration Factors for Road Segments in the Toronto Data

Observed	1035
Predicted	499
Cr	2.1

In Table 5.11, the Cr, which is equal to 2.1, shows that it is more accurate to predict collisions for road segments with zero access since unsignalized intersection collisions are included in the other segments.

5.12 Goodness-of-Fit (CURE Plots)

Figures 5.1 and 5.2 illustrate the CURE plots for AADT by using HSM SPFs for the total collisions in road segments.

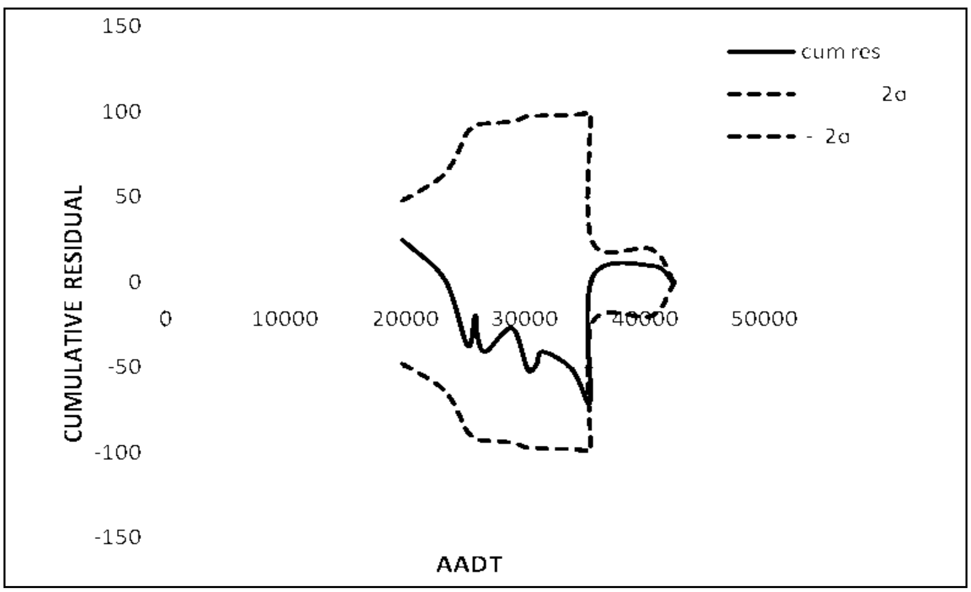


Figure 5.1: CURE Plot for Road Segments that include Unsignalized Intersections

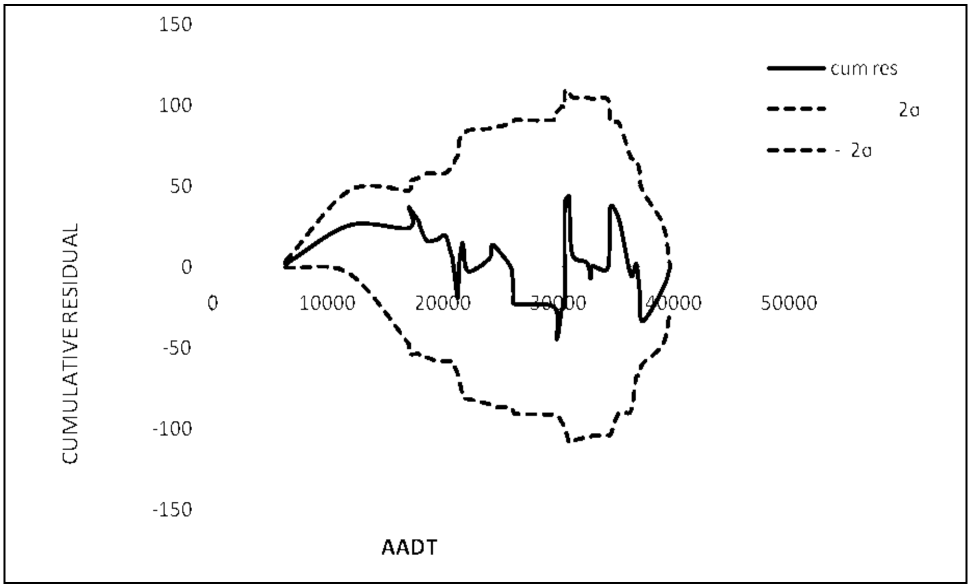


Figure 5.2: CURE Plot for Road Segments without Unsignalized Intersections

5.13 Summary

The CURE plots as shown in Figures 5.1 and 5.2 indicate that overall, the recalibrated models perform reasonably well and the data are a good fit. The Cr for road segments that do not include unsignalized intersections is smaller compared to the road segments that include unsignalized intersections. In both types of segments, the predicted number of collisions is underestimated compared to the observed number of collisions. This may be due to the differences in the collision reporting system between the City of Toronto and the city for which the HSM models are developed. Overall, the recalibration of the HSM algorithm for urban/suburban 4U segments is successful.

6.0 Statistical Analysis of Toronto Signalized Intersections

6.1 SPF Coefficients for Toronto Data by Using HSM Base Models

As per the HSM, users may develop SPFs with data from local jurisdictions for use in the predictive method rather than calibrating the SPFs. SPFs that are directly developed with local data may logically provide more reliable estimates than calibration. Therefore, jurisdictions that have the capability, and wish to develop their own models, are encouraged to do so. Guidance on the development of jurisdiction specific SPFs is presented in Section A.1.2 of the HSM.

6.1.1 Statistical Analysis

SPFs are developed by using a statistical technique called NB regression. The NB regression method accounts for k typically found in collision data. In doing so, it quantifies the k parameter so that the model predicted collisions can be combined with observed collision frequency by using the EB method for safety estimations as per Hauer (1997). SAS PROC GENMOD is used for statistical analysis.

6.2 Toronto vs. HSM SPF Coefficients

The HSM model equations as per Chapter 2, was used to estimate the coefficients for the Toronto collision data. The HSM base model coefficients and Toronto newly developed coefficients for the HSM base models in terms of multi vehicle, single vehicle, and pedestrian collisions are shown in Tables 6.1 to 6.6.

Table 6.1: Total Multivehicle Collision Coefficients for 3 –Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
HSM (Coefficients)	-12.13	1.11	0.26
Toronto (Coefficients)	-6.6024	0.6177	0.3874
Pr (Significance Level)	0.0315	0.0715	<.0001
K (Overdispersion)	0.2874		

In Table 6.1, the major AADT coefficient (b) has a Pr =0.07, which indicates a confidence level of 93%, and the minor AADT coefficient (c) has a Pr <0.0001, which is close to a 100% confidence level. An overdispersion value of k=0.28, shows that the fit to the data is good.

Table 6.2: Total Multi vehicle Collision Coefficients for 4-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
HSM (Coefficients)	-10.99	1.07	0.23
Toronto (Coefficients)	-7.4813	0.5661	0.5581
Pr (Significance Level)	<.0001	<.0001	<.0001
K (Overdispersion)	0.1992		

In Table 6.2, the coefficients for the independent variables, major and minor AADTs have a confidence level close to 100%. An overdispersion value of k=0.2 also shows that the fit to the data is good.

Table 6.3: Total Single Vehicle Collision Coefficients for 3-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
HSM (Coefficients)	-9.02	0.42	0.4
Toronto (Coefficients)	-0.2617	-0.5105	0.4288
Pr (Significance Level)	0.8936	0.0305	0.0003
K (Overdispersion)	0.4032		

In Table 6.3, the major AADT coefficient (b) has a Pr=0.031 which indicates a confidence level of 97%, and the minor AADT coefficient (c) has a Pr value near zero, which indicates a confidence level close to 100%. A low overdispersion value of k=0.4, shows that the data is a good fit, while a negative value for the major AADT coefficient illogically suggests that collisions will reduce with an increase in traffic. This trend is not realistic, therefore, this model is not recommended.

Table 6.4: Total Single Vehicle Collision Coefficients for 4-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
HSM (Coefficients)	-10.21	0.68	0.27
Toronto (Coefficients)	-6.8423	0.1377	0.487
Pr (Significance Level)	<.0001	0.0218	<.0001
K (Overdispersion)	0.2052		

In Table 6.4, the major AADT coefficient (b) has a Pr value of 0.0218 at a confidence level of 98%, and the minor AADT coefficient (c) has a value of Pr near zero, close to a 100% confidence level. A low overdispersion value of k=0.205, shows that the data is a good fit.

Table 6.5: Pedestrian Fatal and Injury Collision Coefficients for 3-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Total (b)	AADT Minor/ AADT Major (c)	Pedestrian Volume (d)	nMax-lanes (e)
HSM (Coefficients)	-6.6	0.05	0.24	0.41	0.09
Toronto (Coefficients)	-0.0419	-0.5397	0.0821	0.5424	0.0188
Pr (Significance Level)	0.9854	0.0229	0.4247	<.0001	0.8644
K (Overdispersion)	0.3158				

In Table 6.5, the independent variable, total AADT, has a Pr value of 0.0229 at a confidence level of 98%, while the minor/major AADT ratio, with a Pr value of 0.4247 at a confidence level of 58% is highly insignificant. Pedestrian volume has Pr <.0001; on the other hand, the maximum number of lanes with Pr= 0.8644 is not statistically significant. A low overdispersion value, k=0.3, shows that the data is a good fit, while a negative value for the total AADT coefficient shows that collisions will reduce with an increase in traffic. This trend is not realistic; therefore, the model is not recommended.

Regression analyses provide no assurance that the values of fitted regression coefficients represent cause-and-effect relationships of specific independent variables to safety. Thus, each regression result was reviewed to assess whether the results obtained were consistent with existing knowledge on the effect of that variable. Anomalous relationships, such as coefficients with an opposite sign to that which is expected, can result when variables in the model correlate with those that are not included in the model. Where such anomalous relationships were found, the variable in question was excluded from further modelling for the roadway or intersection type in question. This approach provided assurance that the resulting models were not only statistically significant, but also meaningful in engineering terms.

Table 6.6: Pedestrian Fatal and Injury Collision Coefficients for 4-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Total (b)	AADT Minor/AADT Major (c)	Pedestrian Volume (d)	nMaxlanes (e)
HSM (Coefficients)	-9.53	0.4	0.26	0.45	0.04
Toronto (Coefficients)	-4.8246	0.4718	0.4142	0	-0.0045
Pr (Significance Level)	<.0001	<.0001	<.0001	1	0.8817
K (Overdispersion)	0.5834				

In Table 6.6, the independent variables, total AADT and minor/major AADT, have a Pr <0.0001 at a confidence level near to 100%. The pedestrian volume coefficient has a Pr=1(0% confidence level) so it is highly insignificant. Also, the maximum number of lanes with Pr= 0.8817 is also highly significant. Therefore, the model is not recommended.

6.3 Multivehicle Collision Crash Type Distributions in HSM vs. Toronto SPFs

6.3.1 Rear End Collision Predictions

A distribution of multivehicle collisions in the HSM is given in Table 6.7. In this table, collisions are given as a fraction of the multivehicle collisions. For example, rear end collisions for 3SG intersections comprise 53% of F&I collisions, 51% of PDO collisions, and approximately the same percentage for 4SG intersections. Rear end collisions in the Toronto data for 3SG and 4SG intersections, as shown in Figures 6.1 and 6.2, make up 56% and 44%, respectively, of the total multi vehicle collisions. These collisions were analysed to determine suitable coefficients for application to the HSM base model so that the proportion of rear end collisions are dependent on major and minor traffic volumes.

Table 6.7: HSM Multivehicle Collision Distribution on 3-Legged Signalized Intersections and 4-Legged Signalized Intersections

	3-legged signalized		4-legged signalized	
Collision type	Fatal+Injury	PDO	Fatal+Injury	PDO
Rear end collisions	0.53	0.511	0.506	0.507
Head-on collisions	0.01	0.004	0.012	0.004
Angle collisions	0.387	0.3	0.421	0.305
Side (same direction) collisions	0.049	0.129	0.029	0.123
Side (opposite direction) collisions	0.004	0.004	0.003	0.005
Other collisions	0.02	0.052	0.029	0.056

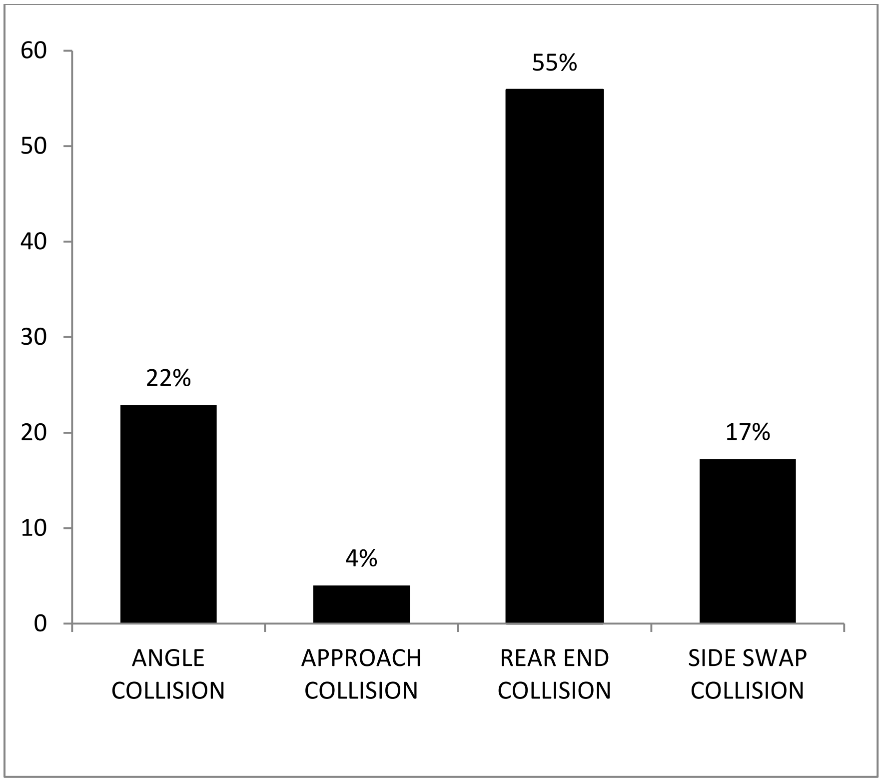


Figure 6.1: Toronto multivehicle collision distribution for 4-legged signalized intersections

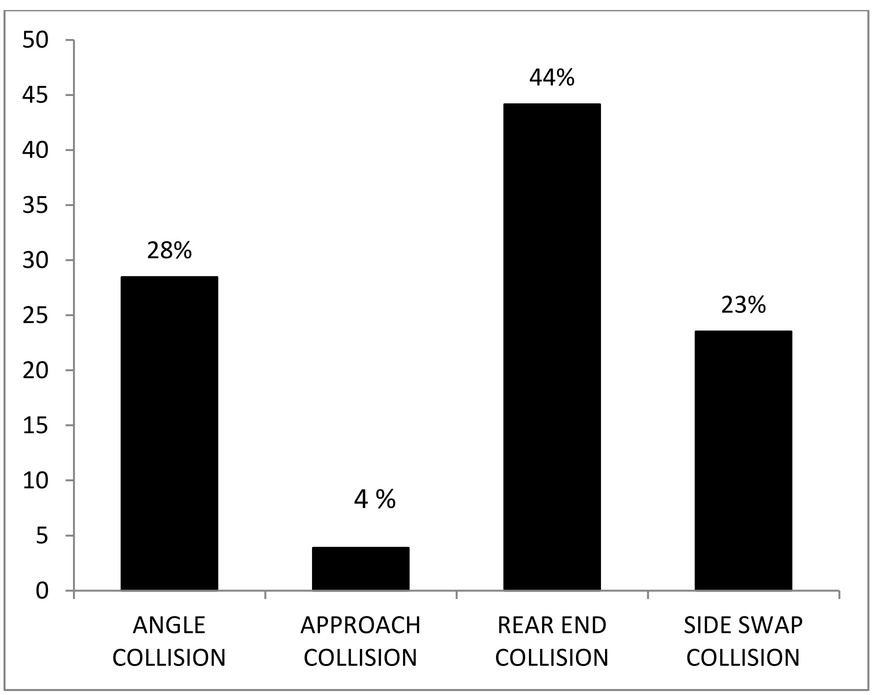


Figure 6.2: Toronto multivehicle collision distribution for 3-legged signalized intersections

6.3.2 Rear End Collision SPFs

The data for rear end collisions in Toronto for 3SG and 4SG intersections were used to calibrate HSM base models which took into consideration the major and minor AADTs as explanatory variables. The results are shown in Tables 6.8 and 6.9 below.

Table 6.8: Rear End Collision Coefficients for 3-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-11.8262	1.152	0.2388
Pr (Significance Level)	<0.0001	0.0003	0.0003
K (Overdispersion)	0.137		

In Table 6.8, the coefficients for major and minor AADTs are statistically significant for rear end collisions with a confidence level near 100% and a low overdispersion value of k= 0.137, which indicates a good fit for rear end collision predictions when calibrating HSM base models to Toronto data.

Table 6.9: Rear End Collision Coefficients for 4-Legged Signalized Intersections in Toronto

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-10.786	0.8195	0.531
Pr (Significance Level)	<.0001	<.0001	<.0001
K (Overdispersion)	0.1874		

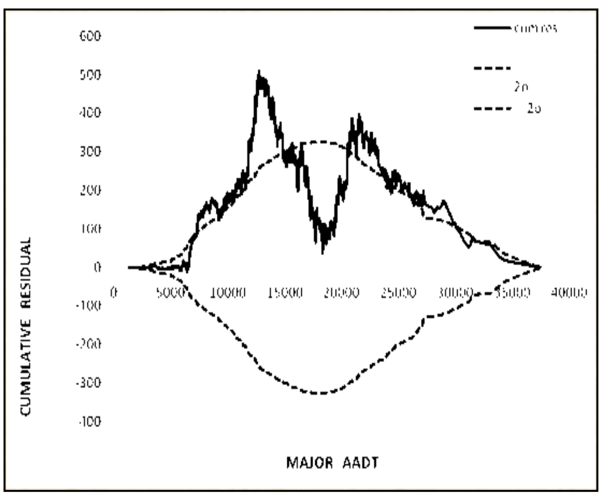
In Table 6.9, the coefficients for the major and minor AADTs are statistically significant for rear end collisions with a confidence level near 100% and a low overdispersion value of k=0.187, which indicates a good fit.

6.4 Goodness-of-Fit

6.4.1 CURE Plots

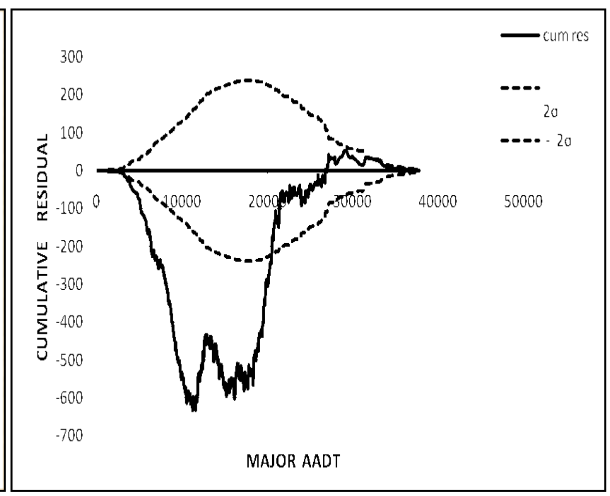
The new coefficients derived from an analysis of the Toronto collision data were used to predict collisions. How well these observed and predicted collisions fit were analysed by using CURE plots. The CURE plots for minor and major AADTs by using Toronto and HSM SPFs are shown below in Figures 6.3 to 6.10.

HSM Model

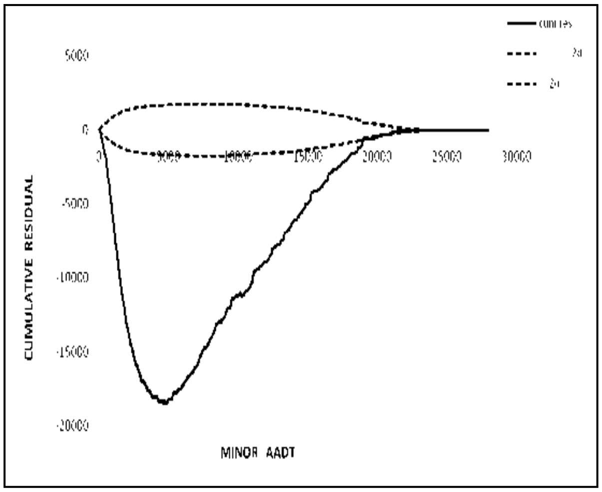


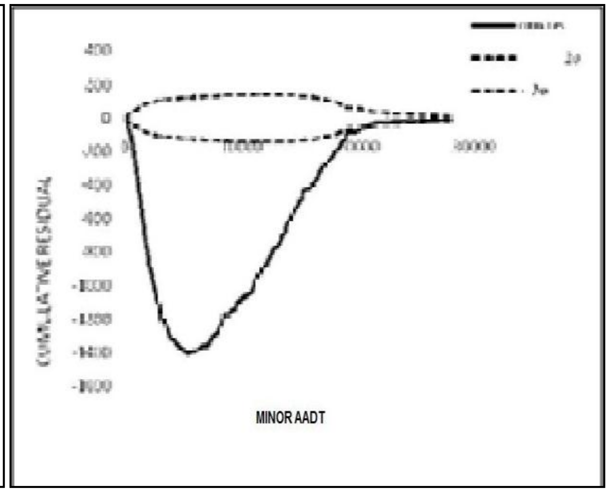
6.3

Toronto Model



6.4

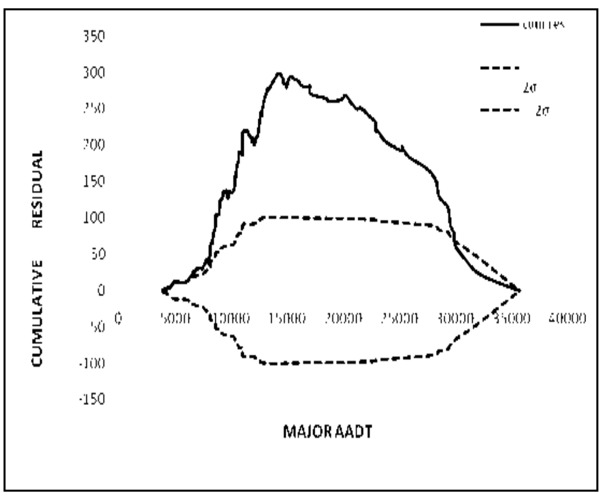


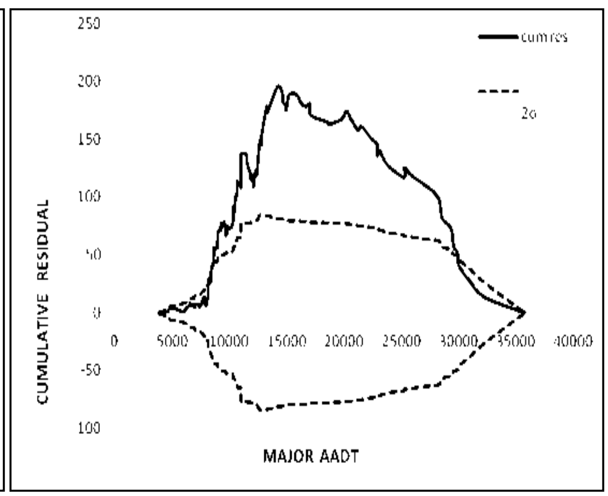


6.5 6.6

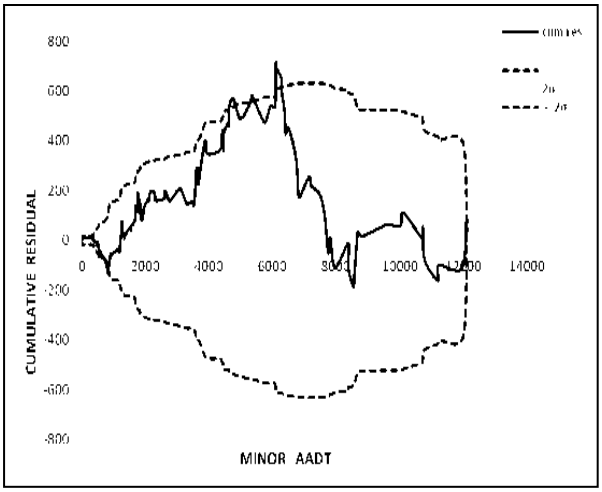
Figures 6.3 to 6.6: Multivehicle Collision CURE Plots for 4-legged Signalized Intersections based on Toronto Data

HSM Model Toronto Model

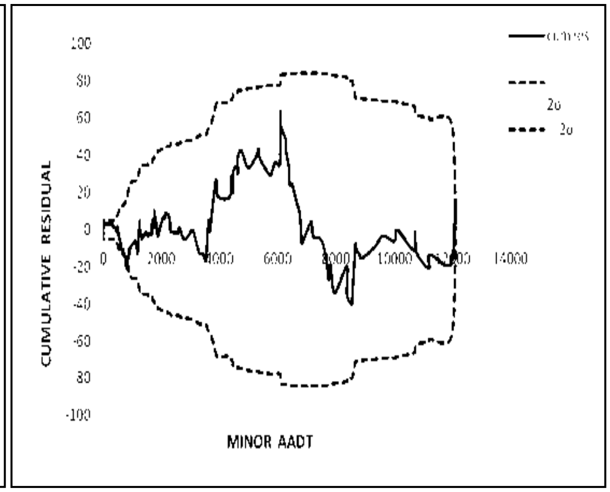




6.7 6.8



6.9



6.10

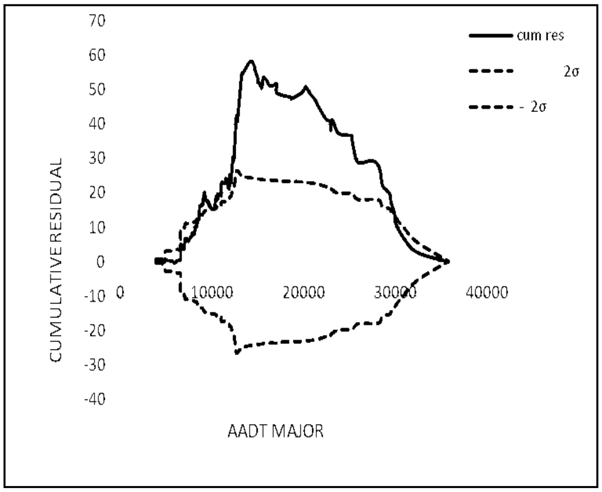
Figure 6.7 to 6.10: Multivehicle Collision CURE Plots for 3-legged Signalized Intersections based on Toronto Data

For major and minor AADTs, the total number of collisions tends to be underpredicted for all ranges of AADTs. The CURE plot strays outside the two standard deviation boundaries, which confirms that there is bias in the model predictions for both 3SG and 4SG intersections. For multivehicle collisions at 3SG intersections, the CURE plot of major AADTs for the total number of collisions strays outside the two standard error boundaries for all ranges of major AADTs, which shows that there is bias in the model predictions. In terms of minor AADT for the total number of collisions, the CURE plot just slightly strays outside the two standard deviation boundaries, which confirms that there is only a slight bias in the model predictions.

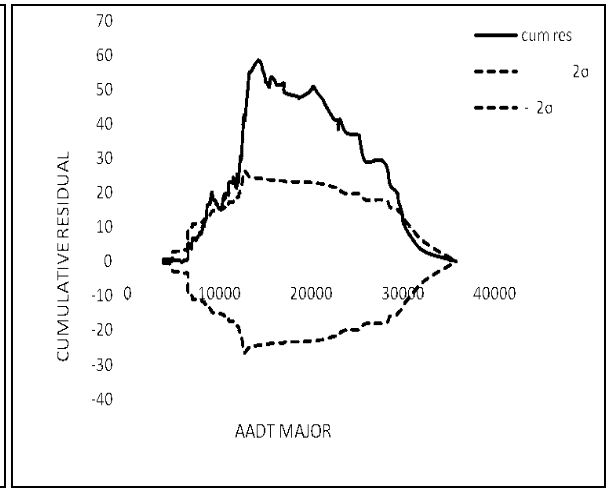
CURE plots for rear end collisions at 3SG and 4SG intersections are shown below in Figures 6.11 to 6.18.

HSM Model

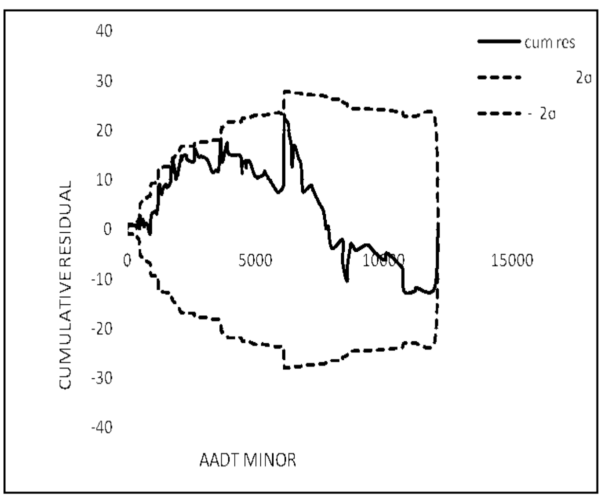
Toronto Model



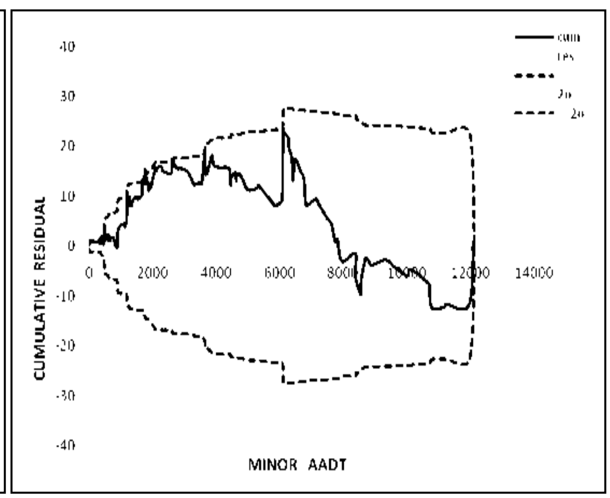
6.11



6.12



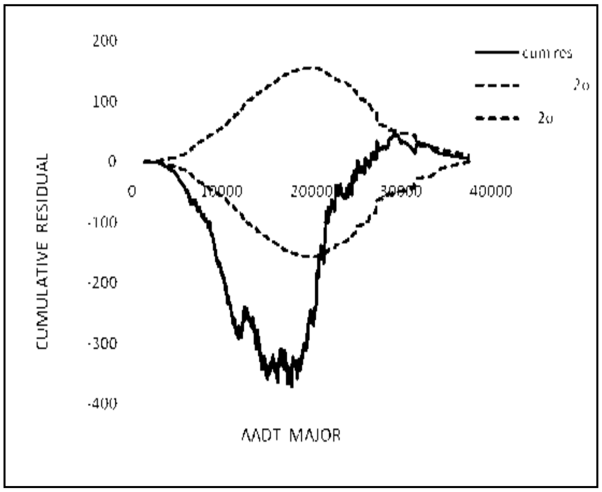
6.13



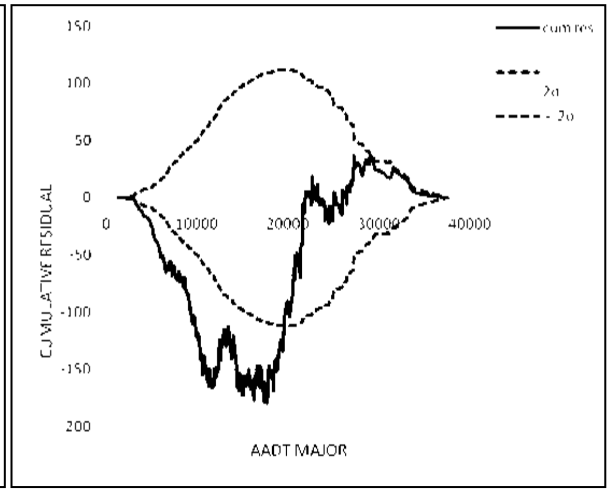
6.14

Figures 6.11 to 6.14: Rear End Collision CURE Plots for 3-legged Signalized Intersections based on Toronto Data

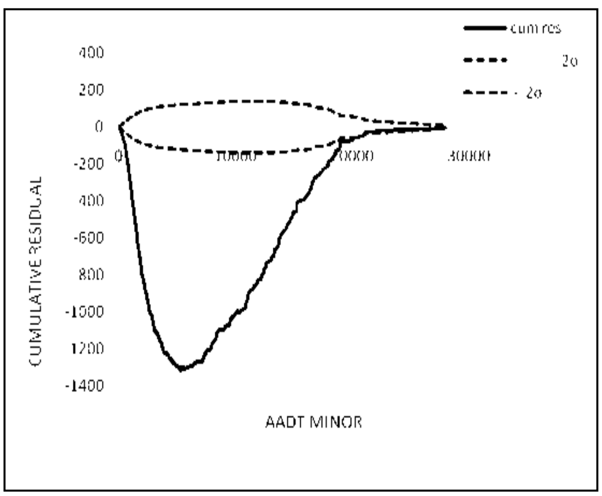
HSM Model | Toronto Model



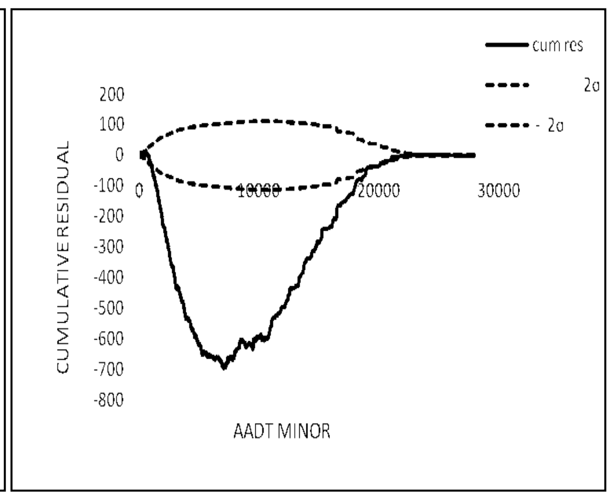
6.15



6.16



6.17



6.18

Figures 6.15 to 6.18: Rear End Collision CURE Plots for 4-legged Signalized Intersections based on Toronto Data

For major AADTs, all ranges of the CURE plot stray outside the two standard deviation boundaries, which confirm that there is bias in the model predictions for both 3SG and 4SG intersections. For 4SG intersections, the CURE plot of minor AADTs for the total number of collisions stays inside the two standard error boundaries for all ranges, which shows that there is no bias in the model predictions. In terms of minor AADTs for 3SG intersections, the CURE plot strays outside the two standard deviation boundaries for all ranges, which confirms that there is bias in the model predictions.

6.4.2 Goodness-of-Fit Tests for Re-Calibration Based on the Toronto Data

The MAD and k parameter results are shown in Table 6.10 as follows.

Table 6.10:" MAD" and "k" Tests for Re-Calibration Based on Toronto Data

	Type of Facility	Type of Collision	Type of Severity	Observed Collisions	Predicted Collisions	Cr	MAD[(a)]	k[(b)]
HSM SPF (Toronto SPF)	3SG	Multi	Total	7239	1292 (8172)	5.60	6.03 (5.38)	0.56 (0.43)
HSM SPF (Toronto SPF)	3SG	Multi	Rear end	1767	3140 (1836)	0.56	1.36 (1.41)	0.56 (0.39)
HSM SPF (Toronto SPF)	4SG	Multi	Total	109910	22998 (90615)	4.78	5.72 (4.10)	0.44 (0.26)
HSM SPF (Toronto SPF)	4SG	Multi	Rear end	34156	46391 (25176)	0.74	2.36 (1.92)	0.65 (0.39)
HSM SPF	4U Segment	Total	Total	810	341	2.37	3.39	1.12
HSM SPF	4U excluding unsignalized intersections	Total	Total	1035	499	2.07	2.24	0.51

Note: (a) "MAD" is based on one year of collisions

(b) "k" is recalibrated overdispersion parameter

A Valid "MAD" value should be at a similar level of the observed average collision frequency and a smaller "MAD" value indicates a better fit. The "k" value is a relative measure to compare two different models. The model which shows a smaller "k" is a better fit. By this standard, the test shows that the Toronto models "MAD" and "k" values are better than those of the HSM models. Overall, both models are a good fit, with the "MAD" and "k" values acceptable.

6.5 Comparison of Observed and Predicted Collisions vs. Total Entering AADTs

Calibration factors for different ranges of AADT are shown in Tables 6.11 to 6.14 as follows.

Table 6.11: Multivehicle Collision Cr vs. Total Entering AADTs for 4-legged Signalized Intersections in Toronto

AADT	Intersections	Observed/Predicted	
		Calibrated HSM Model	Toronto Model
0 to 15000	673	0.85	1.03
15000 to 30000	867	0.93	1.18
>30000	150	1.45	1.52

In Table 6.11, the Cr value for the HSM calibrated model is the best for the medium ranges of total entering AADTs. For the Toronto model, the Cr value is best for the lower ranges of total entering AADTs for multivehicle collisions at 4SG intersections.

Table 6.12: Rear End Collision Cr vs. Total Entering AADTs for 4-Legged Signalized Intersections in Toronto

AADT	Intersections	Observed/Predicted	
		Calibrated HSM Model	Toronto Model
0 to 15000	673	0.50	1.02
15000 to 30000	867	0.66	1.27
>30000	150	1.30	1.92

In Table 6.12, the Cr value for the HSM calibrated model is the best for the higher range of total entering AADTs. For the Toronto model, the Cr value is good for the lower range of total entering AADTs for rear end collisions at 4SG intersections.

Table 6.13: Multivehicle Collision Cr vs. Total Entering AADTs for 3-Legged Signalized Intersections in Toronto

AADT	Intersections	Observed/Predicted	
		Calibrated HSM Model	Toronto Model
0 to 15000	74	1.36	1.02
15000 to 30000	46	1.19	1.02
>30000	17	0.48	0.52

In Table 6.13, the Cr values for the HSM calibrated model and Toronto model are the best for the medium ranges of total entering AADTs. For the Toronto model, the Cr value is also best for the lower ranges of entering AADTs for multivehicle collisions at 3SG intersections.

Table 6.14: Rear End Collision Cr vs. Total Entering AADTs for 3-Legged Signalized Intersections in Toronto

AADT	Intersections	Observed/Predicted	
		Calibrated HSM Model	Toronto Model
0 to 15000	74	0.67	1.16
15000 to 30000	46	0.64	1.04
>30000	17	0.37	0.63

In Table 6.14, the Cr value for the HSM calibrated model is the best for the lower ranges of total entering AADTs. For the Toronto model, the Cr value is also good for the same range of total entering AADTs for rear end collisions at 3SG intersections.

6.6 Comparison of Observed and Predicted Multivehicle Collisions by Number of Left Turn and Right Turn Approaches

The calibration factors for different numbers of approaches with left turn lanes and right turn lanes are shown in Tables 6.15 and 6.16 as follows.

Table 6.15: Multivehicle Collision Cr vs. Total Left Turn and Right Turn Approaches for 4-legged Signalized Intersections in Toronto

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted			
			Multi vehicle		Rear end	
			Toronto Model	Calibrated HSM Model	Toronto Model	Calibrated HSM Model
0	0	343	1.01	0.84	0.99	0.50
1	0	50	1.18	1.10	1.17	0.68
0	1	159	1.07	0.70	1.13	0.49
2	0	10	0.78	0.79	0.89	0.53
1	1	75	0.94	0.65	1.12	0.50
0	2	268	1.10	0.84	1.16	0.58
2	1	13	0.92	0.69	1.4	0.72
1	2	121	1.08	0.79	1.22	0.40
0	3	61	1.09	0.79	1.15	0.56
2	2	87	1.07	0.82	1.14	0.41
1	3	66	1.07	0.76	1.17	0.58
0	4	109	1.39	1.33	1.52	0.96
2	3	40	1.24	0.93	1.39	0.73

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted			
			Multi vehicle		Rear end	
			Toronto Model	Calibrated HSM Model	Toronto Model	Calibrated HSM Model
3	2	14	0.98	0.79	1.04	0.55
1	4	81	1.51	1.45	1.72	1.11
4	2	3	0.98	0.88	0.62	0.36
3	3	11	1.68	1.36	1.86	1.05
2	4	82	1.53	1.49	1.85	1.24
4	3	4	1.33	1.16	1.78	1.05
3	4	37	1.79	1.91	2.31	1.67
4	4	42	1.82	2.00	2.29	1.68

In Table 6.15, the Cr values of multi vehicle collisions in the HSM calibrated model and Toronto model show an increasing trend when the number of approaches with left turn lanes are 3 or more and the total number of approaches with left turn lanes and right turn lanes are 5 or more. This can be one reason why the CMF used is not effective, since the number of approaches increases up to a certain number. In addition, it may be due to the increasing number of observed collisions for these types of sites. This means that the CMFs for a larger number of approaches are not as precise as expected. The same trend is observed for the Cr of rear end collisions in the Toronto model. The Cr value of the HSM calibrated model is better for the base condition for multi vehicle and rear end collisions. The best Cr value is observed in multivehicle collisions for the HSM calibrated model, and intersections that have 3 approaches with left turn lanes and 2 approaches with right turn lanes. In the Toronto model, the best Cr value is observed for the base condition for multivehicle and rear end collisions. In the rear end collisions for the HSM calibrated model, the best Cr value is 0.96, which is for intersections that have 4 approaches with left turn lanes.

Table 6.16: Multivehicle Collision Cr vs. Total Left Turn and Right Turn Approaches for 3-Legged Signalized Intersections in Toronto

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted			
			Multi vehicle		Rear end	
			Toronto Model	Calibrated HSM Model	Toronto Model	Calibrated HSM Model
0	0	45	1.03	1.35	0.97	0.56
0	1	20	1.32	1.65	1.16	0.67
1	1	27	0.88	0.99	1.07	0.62
0	2	10	0.53	0.50	0.71	0.42
1	2	16	0.53	0.55	0.67	0.39

In Table 6.16, the Cr values in multivehicle and rear end collisions for the HSM calibrated model and Toronto model show a decreasing (except for the one approach with a left turn lane) trend as the number of approaches with left turn lanes and right turn lanes increase. This trend shows that the number of observed collisions decreases as the number of approaches increases.

6.7 Summary

The analysis result shows that multi vehicle collision model regression coefficients for 3SG and 4SG intersections are statistically significant. Single vehicle collisions model regression coefficients for 4SG intersections are statistically significant while for 3SG intersections, the regression coefficients are statistically insignificant. In the case of rear end collisions, statistically significant results are obtained for both 3SG and 4SG intersections.

7.0 Transferability of HSM and Toronto Models to Edmonton Data

The collision data from the city of Edmonton is used to check the transferability of Toronto's newly developed model for other local jurisdictions. Also, the Toronto models are compared with the HSM models by applying Edmonton data to both models.

7.1 Edmonton Data

The Edmonton data shows that for both 3SG and 4SG intersections, 70% are PDO collisions and 30% are F&I collisions.

7.1.1 Summary Statistics

Tables 7.1 and 7.2 list the 3SG and 4SG collision data in Edmonton, and Figures 7.1 to 7.2 show the collision distributions.

Table 7.1: 3-Legged Signalized Intersection Collision Data in Edmonton

Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collisions	46.94	0	194	88	4130	100	6
Total Injury Collisions	14.05	0	71.1	88	1236	29.93	6
Total PDO Collisions	32.9	0	140	88	2895	70.08	6
Right Angle Collisions	2.6	0	56	88	228	5.53	6
Rear End Collisions	7.48	0	52	88	658	15.93	6
Followed Too Close Collisions	28	0	153	88	2464	59.65	6
Failed to Observe Signal Collisions	5.57	0	61	88	490	11.86	6
Total Multi Vehicle Collisions	43.7	0	260	88	3845	93.09	6
Total Rear End	35.48	0	205	88	3122	75.58	6

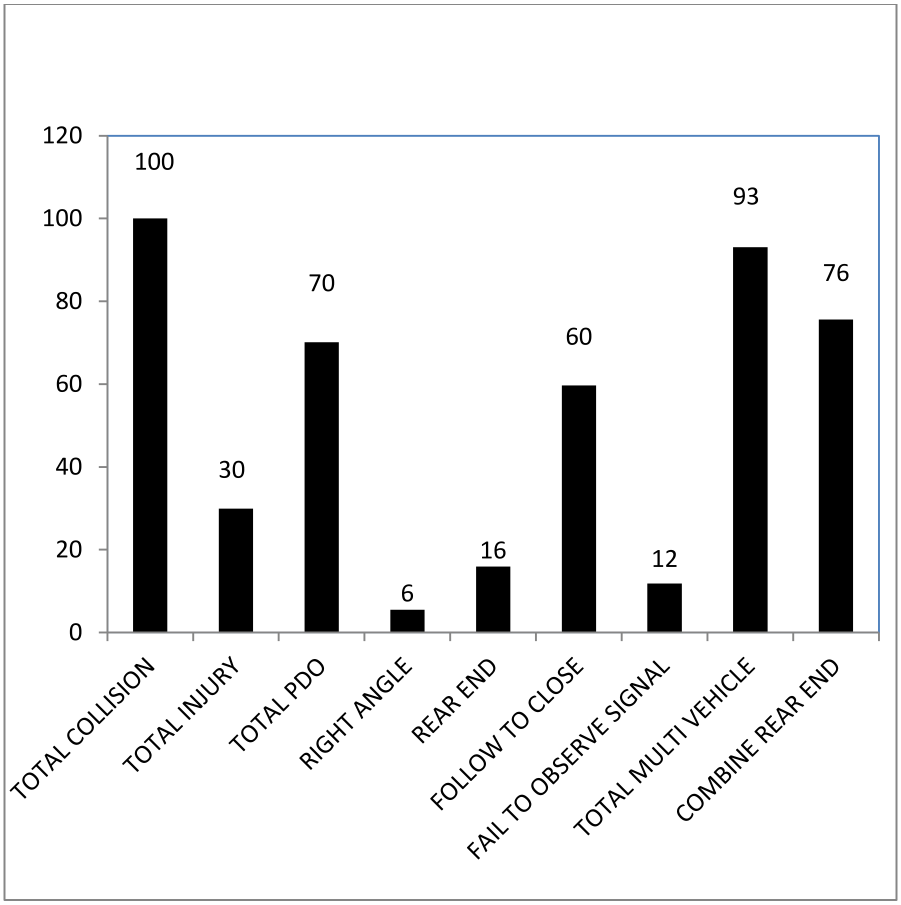


Figure 7.1: % distribution of collisions on 3-legged signalized intersections in Edmonton

Table 7.2: 4-Legged Signalized Intersection Collision Data in Edmonton

Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Total Collision	81.3	0	499	515	41869	100	6
Total Injury Collisions	24.26	0	154	515	12493	29.84	6
Total PDO Collisions	57	0	363	515	29355	70.16	6
Right Angle Collisions	9.27	0	154	515	4774	11.40	6
Rear End Collisions	11.7	0	114	515	6025	14.39	6
Followed Too Close	41.96	0	423	515	21609	51.61	6

Collision Type	Mean	Minimum	Maximum	Sites	Total	%	Years
Collisions							
Failed to Observe Signal Collisions	12.5	0	157	515	6437	15.37	6
Total Multi Vehicle Collisions	75.55	0	555	515	38908	92.92	6
Combine Rear End Collisions	53.66	0	537	515	27634	66.00	6

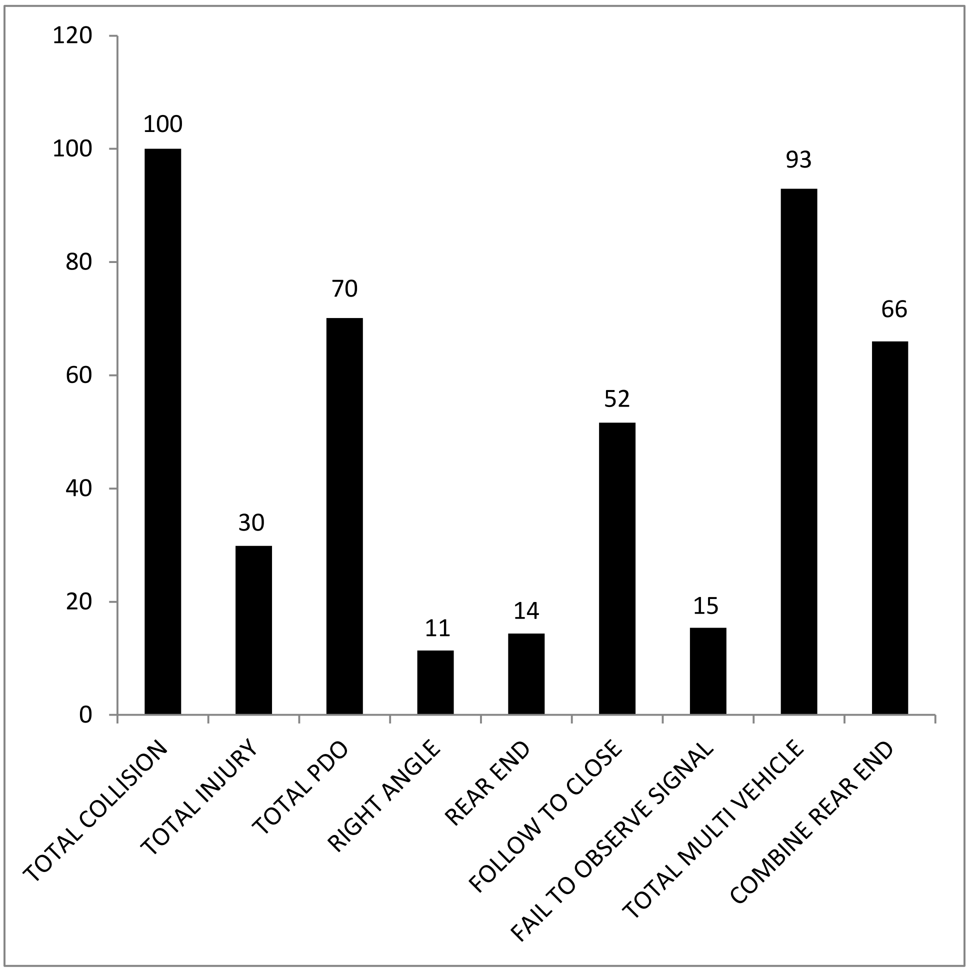


Figure 7.2: % distribution of collisions on 4-legged signalized intersections in Edmonton

7.2 Edmonton SPF Coefficients

The Edmonton data was analysed by using HSM base model equations, and the newly developed coefficients are shown in Tables 7.3 to 7.6.

Table 7.3: Coefficients for Total Multivehicle Collisions on 3-Legged Signalized Intersections in Edmonton

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-10.7896	0.8718	0.4289
Pr (Significance Level)	0.231	0.3535	0.0074
K (Overdispersion)			

Table 7.4: Coefficients for Total Multivehicle Collisions on 4-Legged Signalized Intersections in Edmonton

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-10.7281	0.7634	0.5872
Pr (Significance Level)	0.0005	0.0125	0.0014
K (Overdispersion)	0.6244		

Table 7.5: Coefficients for Rear End Collisions on 3-Legged Signalized Intersections in Edmonton

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-18.5626	1.5174	0.5378
Pr (Significance Level)	0.0211	0.0609	0.0008
K (Overdispersion)	1.9462		

Table 7.6: Coefficients for Rear End Collisions on 4-Legged Signalized Intersections in Edmonton

	Intercept (a)	AADT Major (b)	AADT Minor (c)
Regression Coefficient	-15.173	1.2699	0.4474
Pr (Significance Level)	0.0001	0.0003	0.0342
K (Overdispersion)	0.7376		

7.3 Comparison of Observed and Predicted Collisions in Edmonton data by using Toronto, Edmonton and HSM Model Coefficients

The calibration factors for multi vehicle and rear end collisions for Edmonton data are shown in Tables 7.7 and Table 7.8 as follows.

Table 7.7: Comparison of Observed and Predicted Multivehicle Collisions in Edmonton

	Edmonton Model	HSM Model	Toronto Model
3SG	1.70	2.34	0.49
4SG	1.49	2.36	0.52

In Table 7.7, the predicted number of multi vehicle collisions at 3SG and 4SG intersections show similar behaviour for the Edmonton and HSM models. A Cr greater than 1 represents an underprediction of collisions. For the Toronto model, the prediction is a larger number than the observed number of collisions which is the reverse for the Edmonton and HSM models in which the Cr is less than 1. For the Toronto model, this is consistent with the HSM model prediction for the Toronto data, in which the Cr is greater than 1, which indicates that the number of observed collisions in Toronto is greater than the number of collisions predicted by the HSM.

Table 7.8: Comparison of Observed and Predicted Rear End Collisions in Edmonton

	Edmonton Model	HSM Model	Toronto Model
3SG	1.64	1.56	1.09
4SG	1.87	1.16	0.89

In Table 7.8, rear end collisions in the Toronto model have the best Cr value for 3SG and 4SG intersections, while the Edmonton and HSM models show underestimation of collisions. In addition, the HSM model shows better Cr values for all types of collisions and intersections in the base condition. In comparing the Cr for all collisions, the Cr value for the base condition is much better, i.e., closer to 1.

7.3.1 Comparison of Observed and Predicted Multivehicle Collisions for Total Entering AADT Ranges

Calibration factors for different ranges of total entering AADTs are shown in Tables 7.9 to 7.12.

Table 7.9: Multivehicle Collisions Cr vs. Total Entering AADTs for 4-Legged Signalized Intersections in Edmonton

AADT	Intersection.	Observed/Predicted		
		Calibrated HSM Model	Calibrated Toronto Model	Edmonton Model
0 to 15000	40	0.70	0.63	1.09
15000 to 30000	164	0.82	0.88	1.33
>30000	55	1.45	1.29	1.83

In Table 7.9, as the AADT increases, the Cr value also increases for all models in terms of multivehicle collisions at 4SG intersections. This shows that the observed number of multivehicle collisions has increased, but the predicted number of collisions does not increase at the same rate. The best results for the Toronto and HSM models are for an AADT range of 15,000 to 30,000 vehicles per day. The Edmonton model shows a good Cr value for a lower AADT range of 0 to 15,000 vehicles per day.

Table 7.10: Rear End Collisions Cr vs. Total Entering AADTs for 4-Legged Signalized Intersections in Edmonton

AADT	Intersections	Observed/Predicted		
		Calibrated HSM Model	Calibrated Toronto Model	Edmonton Model
0 to 15000	40	0.64	0.57	1.35
15000 to 30000	164	0.89	0.82	1.52
>30000	55	1.87	1.39	2.52

In Table 7.10, as the AADT increases, the Cr value also increases for all models with respect to rear end collisions at 4SG intersections. This shows that the observed number of multivehicle collisions has increased, but the predicted number of collisions does not increase at the same rate. The best results for the Toronto and HSM models are for an AADT range of 15,000 to 30,000 vehicles per day, while the Edmonton model has a good Cr value at a lower AADT range of 0 to 15,000 vehicles per day.

Table 7.11: Multivehicle Collisions Cr vs. Total Entering AADTs for 3-Legged Signalized Intersections in Edmonton

AADT	Intersections	Observed/Predicted		
		Calibrated HSM Model	Calibrated Toronto Model	Edmonton Model
0 to 15000	10	1.05	0.76	1.77
15000 to 30000	41	0.95	0.81	1.54
30000 to 45000	24	1.12	1.16	1.92
>45000	13	0.91	1.10	1.61

In Table 7.11, as the AADT increases, the Cr values also increase for all models with respect to multivehicle collisions at 3SG intersections, except for Edmonton which has a slight reduction in the AADT range of 15,000 to 30,000 vehicles per day. Furthermore, when the AADT increases above 45,000 vehicles per day, the Cr value again tends to be reduced in all models. This shows that the observed number of multivehicle collisions increases up to a certain range of AADT and then decreases for higher ranges of AADTs. The Toronto and HSM models show the best predictions for AADTs in the range of 15,000 to 30,000 vehicles per day.

Table 7.12: Rear End Collisions Cr vs. Total Entering AADTs for 3-Legged Signalized Intersections in Edmonton

AADT	Intersections	Observed/Predicted		
		Calibrated HSM Model	Calibrated Toronto Model	Edmonton Model
0 to 15000	10	1.29	0.84	3.13
15000 to 30000	41	1.36	0.88	1.99
30000 to 45000	24	1.74	1.11	1.86
>45000	13	1.59	1.00	1.25

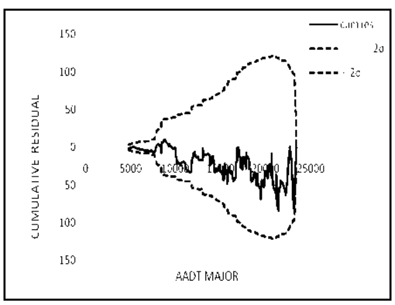
In Table 7.12, as the AADT increases, the Cr value also increases for all models in terms of rear end collisions at 3SG intersections, except for Edmonton at an AADT range of 0 to 15,000 vehicles per day. When the AADT increases above 45,000 vehicles per day, the Cr value again tends to be reduced in all the models. This shows that the predicted number of rear end collisions is decreased up to a certain AADT range and then again increases for higher ranges of AADTs.

7.4 Goodness-of-Fit

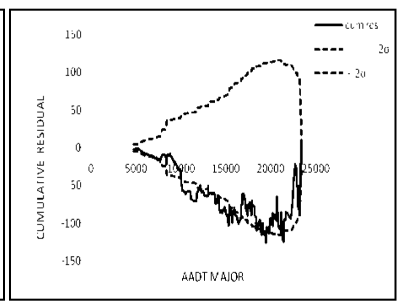
7.4.1 CURE Plots

CURE plots were drawn for the explanatory variables, i.e. major and minor AADTs, to determine how well the data fit the model used for the prediction of collisions in 3SG and 4SG intersections. Separate CURE plots were used for multivehicle and rear end collisions, and are shown in Figures 7.3 to 7.26.

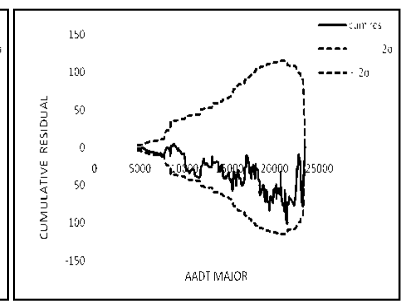
HSM Model | Toronto Model | Edmonton Model



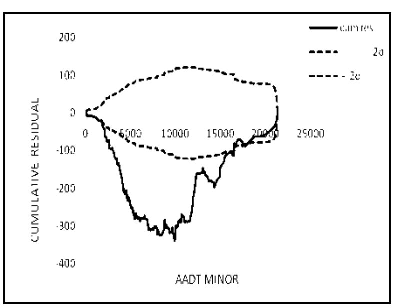
7.3



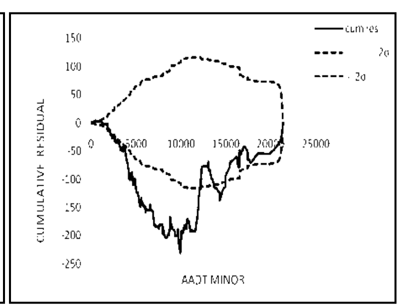
7.4



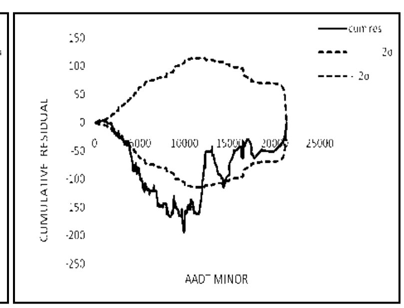
7.5



7.6



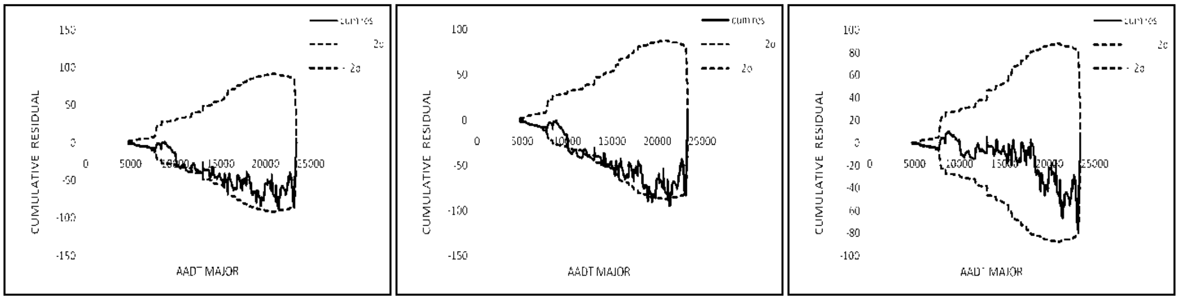
7.7



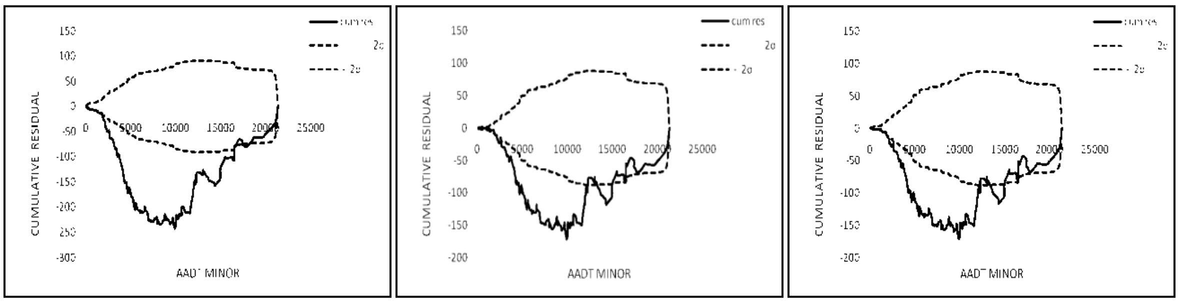
7.8

Figures 7.3 to 7.8: CURE Plots of Multivehicle Collisions at 4-legged Signalized Intersections in Edmonton

HSM Model Toronto Model Edmonton Model



7.9 7.10 7.11

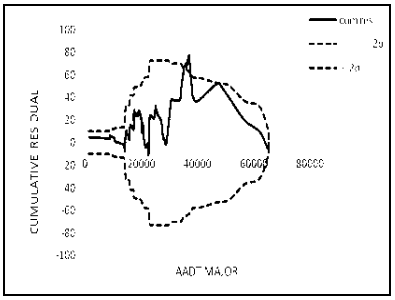


7.12 7.13 7.14

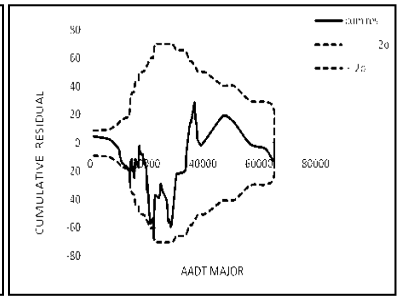
Figures 7.9 to 7.14: CURE Plots of Rear End Collisions at 4-legged Signalized Intersections in Edmonton

The CURE plots of multivehicle collisions at 4SG intersections that are shown in Figures 7.3 to 7.8 for the HSM and Edmonton models show a good fit with the major AADT. The CURE plots of rear end collisions as shown in Figures 7.9 to 7.14 for the Edmonton data fit well while Toronto and HSM models are biased for some of the AADT ranges. The CURE plots of minor AADT for all models are biased and also stray outside the standard deviation limits for most of the AADT ranges.

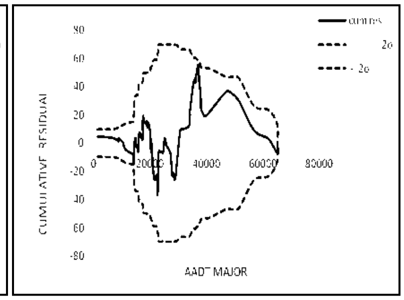
HSM Model	Toronto Model	Edmonton Model



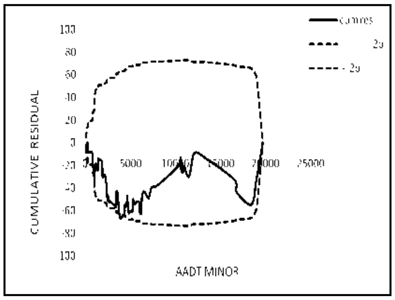
7.15



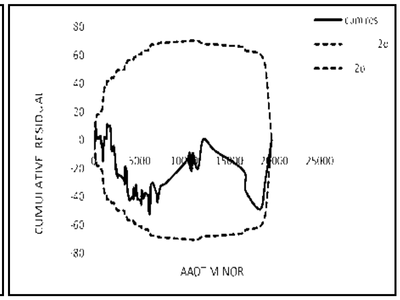
7.16



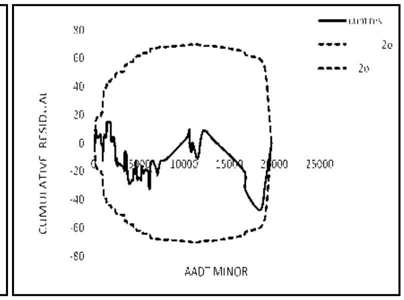
7.17



7.18



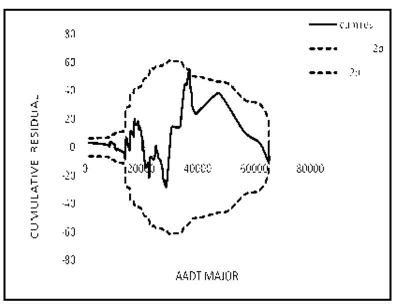
7.19



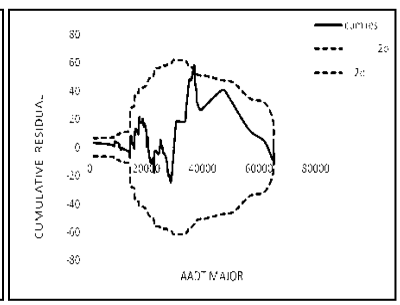
7.20

Figures 7.15 to 7.20: CURE Plots of Multivehicle Collisions on 3-legged Signalized Intersections in Edmonton

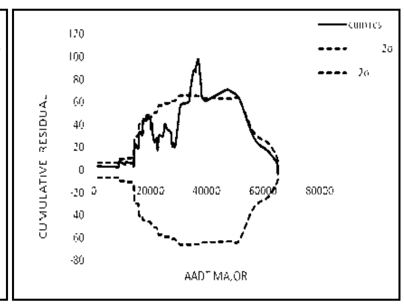
HSM Model	Toronto Model	Edmonton Model



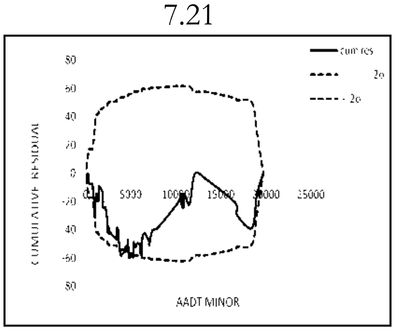
7.21



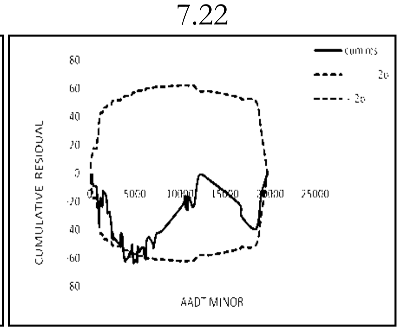
7.22



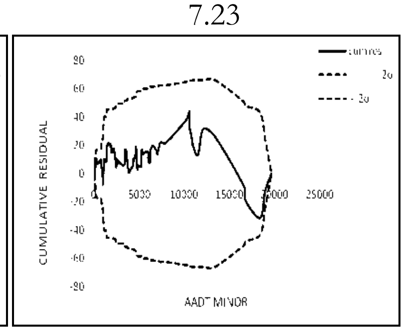
7.23



7.24



7.25



7.26

Figures 7.21 to 7.26: CURE Plots of Rear End Collisions on 3-legged Signalized Intersections in Edmonton

The CURE plot results for the 3SG intersections with the Edmonton data in all three models show a good fit.

7.4.2 Goodness-of-Fit Tests for Re-Calibrations Based on the Edmonton Data

The MAD and k parameter results for all three models are shown in Table 7.13.

Table 7.13: "MAD" and "k" tests for re calibration based on Edmonton data

Theme	Type of Facility	Type of Collision	Type of Severity	Observed Collisions	Predicted Collisions	Cr	MAD[a]	k[b]
HSM SPF Toronto SPF (Edmonton SPF)	3SG	Multi	Total	640	273 1303 (377)	2.34 0.49	5.37 5.38 (5.12)	1.67 1.42 (1.99)
HSM SPF Toronto SPF (Edmonton SPF)	3SG	Multi	Rear end	521	333 478 (317)	1.56 1.09	4.61 4.53 (6.55)	1.71 1.74 (3.04)
HSM SPF Toronto SPF (Edmonton SPF)	4SG	Multi	Total	2033	862 3948 (1362)	2.36 0.52	4.93 5.40 (4.88)	0.85 1.11 (1.37)
HSM SPF Toronto SPF (Edmonton SPF)	4SG	Multi	Rear end	1199	1030 1341 (642)	1.16 0.89	3.56 3.77 (3.23)	1.12 1.35 (1.68)

Note: (a) – "MAD" is mean absolute deviation
(b) – "k" is the overdispersion parameter

In Table 7.13, the "MAD" results from the GOF tests are acceptable for all of the models. Only for the rear end collisions on 3SG intersections in Edmonton in which the "MAD" value is slightly higher but still less than the average frequency of collisions observed in Edmonton. The value of the "k" parameter is relatively good for multivehicle collisions at 4SG intersections for all three models. The Cr value for the rear end collisions in Toronto is the best fit for 3SG and 4SG intersections. These results show that the HSM models are applicable for predicting collisions from local Edmonton data. However, the local models can also be applied to predict collisions for another local jurisdiction. Moreover, the Toronto rear end model can be used for other Canadian intersections rather that the HSM approach which uses a proportion of the predicted multivehicle collisions.

7.5 Comparison of Observed and Predicted Multivehicle Collisions by Number of Approaches with Left and Right Turn Lanes

The effect of the different number of approaches with left turn lanes and right turn lanes on the Cr value for multivehicle collisions is shown in Table 7.14 for 4SG intersections.

Table 7.14: Multivehicle Collision Cr vs. Total Left Turn and Right Turn Approaches at 4-legged Signalized Intersections in Edmonton

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
0	0	82	0.96	0.63	0.63
1	0	3	1.27	0.86	0.98
0	1	8	1.06	0.69	0.65
2	0	4	1.15	0.73	0.86
1	1	9	2.25	1.54	1.41
0	2	36	1.15	0.78	0.75
2	1	3	1.47	1.02	1.18
1	2	14	1.36	0.93	0.91

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
0	3	2	1.57	0.93	0.86
2	2	17	1.21	0.80	0.75
1	3	5	2.30	1.56	1.76
0	4	10	1.09	0.72	0.72
2	3	5	2.77	1.87	1.87
1	4	9	1.68	1.15	1.14
4	2	9	2.42	1.68	1.61
3	3	3	1.22	0.83	0.72
2	4	11	1.65	1.14	1.23
4	3	6	2.66	1.82	1.88
4	4	21	3.51	2.42	2.55

In Table 7.14, the Cr values of multivehicle collisions for the HSM calibrated model and Toronto model show an increasing trend when the number of approaches with left turn lanes are 3 or more, and total number of approaches with left turn lanes and right turn lanes are 5 or more. This means that the CMF for the number of approaches with turn lanes may not be applicable for larger values of this variable. The same trend is observed for rear end collision Cr in the Toronto model.

The HSM models are well behaved for the base condition Cr in multivehicle and rear end collisions. The best Cr value is observed in the case of multivehicle collisions for the HSM calibrated model with a value of 0.93 for intersections that have 3 approaches with left turn lanes and 1 with right turn lanes. In the Toronto model, the best Cr value is for the base condition.

In the case of the HSM model, the best Cr value is 0.96 for sites where there are 4 approaches with left turn lanes and the total number of approaches is 4. For multivehicle collisions at 4SG intersections, as the number of approaches increase, the Cr value also increases. In some cases, especially when the number of approaches increases to more than 4, the Cr value increases more than the normal value which can be due to two reasons. First, the CMFs used are not applicable for cases in which the number of approaches is more than 4 and secondly, the observed number of collisions on higher side.

The effect of the different number of approaches with left turn lanes and right turn lanes for predicted and observed rear end collisions are shown in Table 7.15 for 4SG intersections.

Table 7.15: Rear end Collision Cr vs. Total Left Turn and Right Turn Approaches at 4-Legged Signalized Intersections in Edmonton

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
0	0	82	1.06	0.56	0.64
1	0	3	1.13	0.58	0.76
0	1	8	0.75	0.39	0.42
2	0	4	1.66	0.79	1.01
1	1	9	3.03	1.73	1.91
0	2	36	1.17	0.64	0.72
2	1	3	1.74	0.92	1.26
1	2	14	1.64	0.89	1.03
0	3	2	1.37	0.64	0.62
2	2	17	1.42	0.76	0.83

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
1	3	5	2.64	1.38	1.79
0	4	10	1.02	0.54	0.63
2	3	5	3.58	1.97	2.31
1	4	9	2.1	1.16	1.36
4	2	9	3.23	1.83	2.11
3	3	3	1.20	0.68	0.71
2	4	11	2.56	1.37	1.75
4	3	6	3.48	1.90	2.32
4	4	21	5.59	3.03	3.77

In Table 7.15, as the number of approaches increases, the Cr value also increases in terms of rear end collisions at 4SG intersections. This is the expected trend as the CMFs reduce the predicted collisions and it is also expected that the observed number of collisions should be reduced. However, in some cases, especially where the number of approaches increases to more than 4, the Cr value increases more than the normal value, which can be due to two reasons. First, the CMFs used are not affected in the case where the number of approaches is more than 4. Secondly, the observed collisions are slightly on the higher side for 4SG intersections.

The effect of the different number of approaches with left turn lanes and right turn lanes for predicted and observed multivehicle collisions is shown in Table 7.16 for 3SG intersections.

Table 7.16: Multivehicle Collision Cr vs. Total Left and Right Turn Approaches at 3-Legged Signalized Intersections for Edmonton

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
0	0	25	1.13	0.62	0.72
1	0	2	0.81	0.91	1.27
0	1	11	0.86	0.49	0.44
1	1	19	1.63	0.95	0.95
0	2	2	2.83	1.58	1.55
2	1	4	1.44	1.02	0.81
1	2	3	1.26	0.64	0.84
2	2	23	2.50	1.51	1.47

In Table 7.16, the Cr value is reasonable for 3SG intersections except when there are two or more approaches with left turn lanes and right turn lanes.

The effect of the different number of approaches with left turn lanes and right turn lanes for predicted and observed rear end collisions is shown in Table 7.17 for 3SG intersections.

Table 7.17: Rear end Collision Cr vs. Total Left and Right Turn Approaches at 3-Legged Signalized Intersections for Edmonton

Approaches with Right Turn Lanes	Approaches with Left Turn Lanes	Intersections	Observed/Predicted		
			Edmonton Model	Calibrated Toronto Model	Calibrated HSM Model
0	0	25	1.1	0.57	0.87
1	0	2	0.65	0.66	0.99
0	1	11	0.92	0.41	0.64
1	1	19	1.66	0.94	1.45
0	2	2	2.55	1.15	1.81
2	1	4	1.08	0.92	1.45
1	2	3	1.78	0.80	1.22
2	2	23	2.63	1.64	2.57

In Table 7.17, the Cr value for the HSM model is very close to 1, i.e. 0.99, in the case where there is one approach with right turn lanes at a 3SG intersection. The Cr value for the Toronto model is very close to 1. Also, for the Edmonton model, the Cr has a value of 0.92, which is very close to 1, in the case where there is one approach with left turn lanes.

7.6 Chapter Summary

When Toronto SPFs for multi vehicle and rear end collisions are used to predict collisions in Edmonton with six years of collision data, the Cr values for multi vehicle collisions at 3SG and 4SG intersections are 0.76 and 0.5, respectively. For rear end collisions, the Cr values are 1.5 and 0.75, respectively. This shows that the Cr value decreases in 4SG intersections as compared to 3SG intersections. These values also show that the Toronto SPF coefficients can be useful for the prediction of collisions in the city of Edmonton. Overall, the CURE plot shows that the data is a good fit with some bias in the lower AADT ranges. Only the cumulated residuals for minor AADT in rear end collisions at 4SG intersections stray outside the two standard deviation boundaries, showing that the data is not a good fit.

8.0 Remarks and Conclusions

8.1 Remarks

The following observations can be made for this research work.

1. The calibration factors for different ranges of AADT are found to vary. Therefore, a single factor as recommended in the HSM is not good enough for application to a single site. A single calibration factor is satisfactory if we are predicting for a set of intersections that represent all ranges of AADT values.

2. A set of data should have the same range of AADTs for calibration as the SPFs are analysed for such ranges.

3. Temporal effects may change the values of the calibration factors over time as the characteristics of sites change and collisions are affected by such changes. No quantified consideration is given for the temporal effect on the prediction of collisions. It is only recommended that the calibration process be repeated every three years.

4. For both Toronto and Edmonton data, the SPFs do not fit well with a low range of AADTs and low collision intersections. In addition, for high ranges of AADTs and intersections with high collisions, the calibration factors perform better than for the low AADT intersections.

5. In terms of the CURE plots, the residuals might stay inside the upper and lower deviation boundaries, but the Cr value may not be realistic or near 1.00, and vice versa. For a good SPF, a reasonable calibration factor that is near 1 and a CURE plot within standard deviation boundaries is preferred.

6. The analysis shows that a variable can be statistically significant, but the impact patterns are not necessarily rational. In some cases of AADTs, negative coefficient values are found. The

negative value means that with an increase in the AADT, collisions will be reduced, which is not realistic. Therefore, merely a statistically significant result does not mean that the models should include that variable and that the coefficients are acceptable for use.

7. In intersections, the percentage of single vehicle collisions is substantially smaller than multi vehicle collisions. On the other hand, rear end collisions make up 25% of the total number of collisions at intersections. Therefore, rear end collisions must be separately accounted for and deserve independent analyses. The Toronto SPF for rear end collisions was calibrated for the Edmonton data with satisfactory results.

8. For multi vehicle, single vehicle, pedestrian and bike collisions, the calibration factors are separately assessed. For each type of collision, the values of the calibration factors are different. Therefore, a single calibration factor for the total number of collisions is not a good estimate to predict collisions. Consequently, a single model is required in which collision severities should be given proportionate weight to find a single calibration factor.

9. For road segments, the HSM base model results are a good fit and the values of the calibration factors are significant. The only problem is in data collection for segments. It is somewhat difficult as many cities have data that include unsignalized intersection collisions. Moreover, it is difficult to differentiate the road segment data for urban/suburban roads. Also, for driveway related crashes, data are not available as the collision reporting system in Toronto does not have separate records for driveway related collisions. To locate driveway types and numbers on a set of road segments can be difficult and time consuming. Therefore, a simple base model is required for inclusion in the HSM to predict driveway related collisions.

10. For multi vehicle collisions at signalized intersections, the HSM base models with new statistically significant coefficients are satisfactory. For single vehicle collisions, the HSM

coefficients or new Toronto coefficients can both be used. For pedestrian collisions, the HSM coefficients give a very high Cr value, and the Toronto data analysis results are also not significant.There is the need to develop a new model for pedesterian collision prediction in local jurisdictions.

8.2 Conclusions

The HSM models are applicable for predictions with local data and behave equally well as compared to locally developed models with other local data. Also, the HSM and locally developed models behave better for base conditions. For medium ranges of AADT, all models give a calibration factor near 1 as compared to the low ranges or higher ranges of AADT, which reflect overpredictions and underpredictions.

A locally developed rear end collision model shows good results and should be used for predictions with local data. The GOF test results for all models demonstrate a reasonably good fit. It is practical to directly use the HSM model when local data are not available in good quantity and quality. Moreover, local SPFs should be developed if data are available and the new SPFs are within the scope of the HSM predictive method.

8.3 Further Study

From the result attained in this research, there are some areas that could benefit from further study. There is the need to develop a collision prediction model which works equally well for all ranges of AADT. In the case of a driveway related multi vehicle collision model for road segments, the base collision prediction model of the HSM requires overly extensive efforts to collect data for the prediction of collisions. There is the need to develop a simple-to-apply model to locally predict driveway related collision.

Bibliography

- Abdel-Aty, M. A., Chen, C. L. and Schott, J. R. (1998).An Assessment of the Effect of Driver age on Traffic Accident Involvement Using Log-Linear Models. Accident Analysis and Prevention, vol. 30, no. 6, pp. 851–861.

- Bauer, K. M. and Harwood, D. W. (2000). Statistical Models of At Grade Intersection Accidents—Addendum Publication, FHWA- Rd- 99-094, Transportation Research Board, Washington, D. C.

- Caliendo, C., Guida,M. and Parisi, A. (2007).A Crash Prediction Model for Multilane Roads. Accident Analysis and Prevention, vol. 39, pp 657–670.

- Dell'Acqua, G. and Russo, F. (2003).Accident Prediction Models for Road Networks. Department of Transportation Engineering, University of Naples, Italy.

- Golob, T. and Recker, W. (2003).Relationships among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions. Journal of Transportation Engineering, vol. 129, pp 342-353.

- Hauer, E. (1997). Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Pergamon Press, Elsevier Science Ltd, Oxford, England.

- Hauer, E and Bamfo, J. (1997).Two Tools for Finding What Function Links the Dependent Variable to the Explanatory Variables. ICTCT, Conference. Lund, Sweden.

- Hauer, E. (2004).Statistical Road Safety Modelling. Transport Board Research, record no. 1897. Transportation Research Board, Washington. D.C.

- Harwood, D. W. and Hauer, E. (2007).Methodology to Predict the Safety Performance of Urban and Suburban Arterials. Final Report for NCHRP Project 17-26. Transport Research Board.

- Highway Safety Manual. (2010).American Association of State Highway and Transportation Officials.

- Karlaftis,M. G. and Golias, I.(2002).Effects of Road Geometry and Traffic Volumes on Rural Roadway Accident Rates. Accident Analysis and Prevention, vol. 34 (3): pp 357-65.

- Lyon, C. and Persaud, B. (2002).Pedestrian Accident Prediction Models for Urban Intersection. Transportation Research Board, record no. 1818, pp 102-107.

- Lyon, C., Haq,A., Persaud, B. and Kodama, T. (2005).Safety Performance Functions for Signalized Intersections in Large Urban Areas Development and Application to Evaluation of Left Turn Priority Treatment. Transportation Research Board, record no. 1908, pp 165-171.

- Lord, D. and Persaud, B. (2000).Accident Prediction Models With and Without Trend Application of the Generalized Estimating Equations Procedure. Transportation Research Board, record no. 1717, pp 102-108.

- Martinelli, F., Torre,F. and Vadi, P.(2005).Calibration of the Highway Safety Manual's Accident Prediction Model for Italian Secondary Road Network. Transportation Research Board, record no. 2103, pp 1-9.

- Martin, J. (2002).Relationship between Crash Rate and Hourly Traffic Flow on Interurban Motorways. Accident Analysis and Prevention, vol. 34, issue no. 5, pp 619-629.

- Milton, J.C., Shankar, V.N. and Mannering, F.L. (2008).Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis. Analysis and Prevention, vol. 40, issue no. 1, pp 260-266

- Nambuusi,B. B., Brijs,T. and Hermans, E. (2008). A review of Accident prediction models for Road Intersections. Infrastructuur en ruimte, UHasselt, PHL, VITO, VUB, UGen, RA-MOW-2008-004

- Oh, J., Lyon, C., Washington, S., Persaud, B. and Bared, J. (2003). Validation of FHWA Crash Models for Rural Intersections. Transportation Research Board, record no. 1840, pp 41-49.

- Persaud, B. and Dzbik,L. (1993). Accident Prediction Models for Freeways. Transportation Research Board, record no. 1401, pp. 55-60.

- Persaud,B.(2001).Statistical Methods in Highway Safety Analysis, National Cooperative Highway Research Program Synthesis 294. Transportation Research Board, Washington. D.C.

- Persaud, B., Lord, D. and Palmisano, J. (2002).Calibration and Transferability of Accident Prediction Models for Urban Intersections. Transportation Research Board, record no. 1784, pp 57-64.
- Rodman, D.L., Knuiman, M.W. and Ryan, G. A.(1996).An Evaluation of Road Crash Injury Severity Measures .Accident Analysis and Prevention, vol. 28, issue no. 2, pp. 163–170.
- Shankar, V., Mannering,F. and Barfield, W.(1996).Statistical Analysis of Accident Severity on Rural Freeways. Accident Analysis and Prevention, vol. 3, pp 391-401.

- Sun, X., Li, Y., Magri,D. and Shirazi, H.(2006).Application of Highway Safety Manual Draft Chapter Louisiana Experience. Transportation Research Board, record no. 1950, pp 55-64.

- Vodden, K., Smith, D., Eaton, F. and Mayhew, D. (2007).Analysis and Estimation of the Social Cost of Motor Vehicle Accidents in Ontario. Transport Canada.

- Wong, S.C., Szea, N. N. and Lia, Y .C. (2007).Contributory Factors to Traffic Crashes at Signalized Intersections in Hong Kong. Accident Analysis and Prevention, volume 39, issue 6, pp 1107–1113.

- Washington, S.,Leonard, J. and others. (2001).Scientific Approaches to Transportation Research. Volumes 1 and 2 Prepared for National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.