DEVELOPMENT OF SAFETY PERFORMANCE FUNCTIONS FOR HIGH-SPEED ROADWAYS IN SASKATCHEWAN

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Civil and Geological Engineering

University of Saskatchewan
Saskatoon

By

NADEEM ABBAS

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised the work presented in this thesis or, in their absence, by the Head of the Department or the Dean of the College through which my thesis work was done. It is understood that copying, publishing or using any portion of this thesis for financial gain is prohibited without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use of the material in my thesis.

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

Head of the Department of Civil and Geological Engineering

University of Saskatchewan

Engineering Building

57 Campus Drive

Saskatoon, Saskatchewan S7N 5A9

Canada

ABSTRACT

Saskatchewan experienced the highest rate of collision-related fatalities among all the Canadian provinces and territories in 2011. To implement measures to reduce collisions on roadways, highway safety professionals require a tool to quantify road safety. There are currently no safety performance functions (SPFs) for high-speed roadways in Saskatchewan, which limits the ability of highway safety professional to quantify and evaluate safety as part of the decision-making process. SPFs are an important component of the Highway Safety Manual's (HSM 2010) systematic safety management process, which evaluates the safety of roadways and the effectiveness of roadway safety improvements. SPFs are mathematical models that relate roadway features, such as traffic volume, geometry, and traffic control, to the observed collisions. They are validated by performing regression analyses. The objective of this study was to use local roadway features and collision data (2007 to 2011) to develop a set of local SPFs to predict the number of collisions on Saskatchewan high-speed roadways including interchanges. GIS maps were used to obtain roadway configuration data (basic freeway segments, ramps, weaving sections, etc.), which were correlated to existing collision data for those same roadways. The negative binomial (NB) distribution was used to derive the SPFs written in the programming language R. Statistical goodness of fit (GOF) tests were performed to identify the best-fitting SPFs, and the study produced 24 statistically significant models (eight roadway configurations and three levels of collision severity). The SPFs developed in this study can play a vital role in improving the planning and decision-making processes in roadway safety engineering and help to reduce the number of fatalities on roadways in Saskatchewan.

ACKNOWLEDGMENTS

My research has been an excellent learning experience for me. During the course of my research, I have met very talented and devoted people who have provided me full support and guidance. I would like to take this opportunity to thank my advisor Dr. Peter Y. Park for his professional advice and for giving me the chance to perform this valuable research. His dedication and support were vital. I would also like to thank my committee members Dr. Jian Peng, Dr. Haithem Soliman, Dr. Ryan Walker, Dr. Avi Akkerman, Jordan Parisien, and Brandon Harris for all their suggestions. Thank you, Dr. Tayyab I. Shah, for being available whenever I required help with GIS.

Finally, I would like to thank Saskatchewan Government Insurance (SGI) and the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this work. I also want to thank the various organizations who provided invaluable data, without which this research would not have been possible.

DEDICATION

This work is dedicated to my mother, Zareena Abbas, who passed away during the completion of this work. I will be forever grateful for the unconditional love, support, and guidance she gave me. Muhnjee Amma!

TABLE OF CONTENTS

PERMISSION TO USE i
ABSTRACT ii
ACKNOWLEDGMENTS iii
DEDICATION iv
TABLE OF CONTENTS v
LIST OF TABLES viii
LIST OF FIGURES x
LIST OF ABBREVIATIONS xi
CHAPTER 1: INTRODUCTION 1
1.1. Background ---- 1
1.1.1. Collisions History on Saskatchewan Roads 1
1.1.2. Measure of Road Safety 4
1.1.3. Rationale for the Development of Safety Performance Functions 7
1.1.4. Visualization of Collisions using ArcGIS 12
1.2. Study Objectives ---- 13
1.3. Study Scope ---- 14
1.4. Research Benefits ---- 15
1.5. Thesis Layout ---- 17
CHAPTER 2: LITERATURE REVIEW 18
2.1. Safety Performance Functions ---- 18
2.1.1. Literature Review on Safety Performance Functions 19
2.1.2. Development of Local SPFs for Saskatchewan 35
2.1.3. Selection of Best-Fitting Model 42
2.1.4. Validation of Models 44
2.2. Network Screening ---- 47
2.2.1. EPDO Average Collision Frequency with EB Adjustment 48
2.2.2. Excess Expected Average Collision Frequency with EB Adjustment 57
2.3. Chapter Summary ---- 60
CHAPTER 3: DATABASE INTEGRATION 61
3.1. Description of Datasets ---- 61
3.1.1. Spatial Datasets 61
3.1.2. Regina Spatial Dataset 62
3.1.3. Saskatoon Spatial Dataset 63
3.1.4. Saskatchewan Ministry of Highway and Infrastructure Spatial Datasets 64
3.1.5. Traffic Volume Datasets 65
3.1.6. Regina Traffic Volume Dataset 66
3.1.7. Saskatoon Traffic Volume Dataset 67
3.1.8. SMHI Traffic Volume Dataset 68

3.1.9. Collision Dataset 68
3.2. Segmentation Scheme -- 69
3.2.1. Basic Freeways inside Interchange System 70
3.2.2. Basic Freeways outside Interchange System 71
3.2.3. Off Ramps 71
3.2.4. On Ramps 71
3.2.5. Ramp Influence Areas 72
3.2.6. Weaving Sections 72
3.2.7. Signalized Ramp Terminals 72
3.2.8. Unsignalized Ramp Terminals 73
3.3. Development of Ready-to-Use, Integrated Database -- 74
3.3.1. Geospatial Referencing for Collisions 76
3.3.2. Roadway Reclassification 77
3.3.3. Managing Spatial Datasets 78
3.3.4. Development of Unique Location Identifiers 80
3.3.5. Plotting Collisions on the Base Layer of the Final Dataset Framework 82
3.3.6. Managing Traffic Volume Datasets 84
3.3.7. Managing Collision Dataset 86
3.3.8. Finalized Integrated Database 89
3.4. Chapter Summary -- 90

CHAPTER 4: DEVELOPMENT OF SAFETY PERFORMANCE FUNCTIONS 92

4.1 Developing Safety Performance Functions -- 92
4.1.1. SPFs for Basic Freeway Segments inside Interchange Systems 96
4.1.2. SPFs for Basic Freeway Segments outside Interchange Systems 99
4.1.3. SPFs for Off Ramps 101
4.1.4. SPFs for On Ramps 103
4.1.5. SPFs for Ramp Influence Areas 104
4.1.6. SPFs for Weaving Sections 107
4.1.7. SPFs for Signalized Ramp Terminals 109
4.1.8. SPFs for Unsignalized Ramp Terminals 110
4.2 Model Validation Results -- 113
4.3 Chapter Summary -- 116

CHAPTER 5: NETWORK SCREENING 117

5.1 Network Screening Methods -- 117
5.1.1. Network Screening for Basic Freeways inside Interchange System 119
5.1.2. Network Screening for Basic Freeways outside Interchange System 120
5.1.3. Network Screening for Off Ramps 121
5.1.4. Network Screening for On Ramps 122
5.1.5. Network Screening for Ramp Influence Areas 123
5.1.6. Network Screening for Weaving Sections 124
5.1.7. Network Screening for All High-Speed Roadway Segments (Excluding Ramp Terminals) 125
5.1.8. Network Screening for Ramp Terminals (signalized) 129
5.1.9. Network Screening for Ramp Terminals (unsignalized) 129
5.1.10. Network Screening for All Ramp Terminals 130
5.2 Chapter Summary -- 137

CHAPTER 6: SUMMARY AND CONCLUSION 138

6.1 Summary -- 138

6.1.1 Segmentation Scheme 139
6.1.2 Spatial Datasets 140
6.1.3 Traffic Volume Datasets 141
6.1.4 Collision Datasets 142
6.1.5 Development of SPFs 143
6.1.6 Validation of SPFs 144
6.1.7 Network Screening 144
6.1.8 GIS Maps for Top Ten Hotspots 145
6.2 Limitations ----- 145
6.3 Future Work and Recommendations ----- 146

REFERENCES 150

LIST OF APPENDICES 156

Appendix A- Study Area Map 157
Appendix B- Sample Calculation for Network Screening (Expected EPDO and Excess EPDO) 159
Appendix C- NCHRP Report 17-45 Segmentation Scheme 173
Appendix D- Sample of Final Database Used for Development of SPFs 175
Appendix E- Regression Results for All Candidate Models 179
Appendix F- Prediction Models 188
Appendix G- CURE Plots of All Candidate Models 198
Appendix H- Summary of Collision Records Excluded from Analysis 214

LIST OF TABLES

Table 1: Three-Year Collision Summary 1

Table 2: Three-Year Collision Summary by Road System 2

Table 3: Three-Year Summary of Losses on Provincial Highway System 2

Table 4: Three-year Average Societal Cost of Collisions on Provincial Highway System.............. 3

Table 5: NCHRP Report 17-45's Freeway Configurations with Available SPFs 20

Table 6: NCHRP Report 17-45's Ramp Segment Configurations with Available SPFs 20

Table 7: NCHRP Report 17-45's Ramp Terminal Configurations with Available SPFs 21

Table 8: Candidate Functional Forms for Basic Freeways Inside and Outside Interchange Systems 37

Table 9: Potential Functional Forms for On-Ramps and Off-Ramps 38

Table 10: Potential Functional Forms for Ramp Influence Areas 39

Table 11: Potential Functional Forms for Weaving Sections 40

Table 12: Potential Functional Forms for Ramp Terminals (Signalized and Unsignalized).......... 41

Table 13: Sample Information in Shapefile for Roadways in the City of Regina 63

Table 14: Sample Information in Shapefile for Roadways in the City of Saskatoon 64

Table 15: Sample Information in Shapefiles for Roadways in Saskatchewan Provided by the SMHI 65

Table 16: Sample of Traffic Volume Information from Saskatoon... 67

Table 17: Sample of Traffic Volume Information from SMHI. 68

Table 18: Sample Information of Raw Collision Data (HIGHSPEED_AC). 69

Table 19: Sample of Integrated Database used for SPF Development 75

Table 20: Sample of Final Database Compiled for the development of SPFs................................ 86

Table 21: Summary of Raw Collision Data 88

Table 22: Final Roadway Configurations and Number of Corresponding Collision Records 89

Table 23: Sample of Finalized Integrated Database for Different Roadway Types 90

Table 24: Number of Candidate Models and Number of Finalized Models.................................. 95

Table 25: Available Sites and Collisions for the Development of SPFs 96

Table 26: Candidate Models for Basic Freeway Segments inside Interchange Systems 97

Table 27: Regression Results for the Selected Models for Basic Freeway Segments inside Interchange Systems 97

Table 28: Candidate Models for Basic Freeway Segments outside Interchange Systems 99

Table 29: Regression Results of Selected Models for Basic Freeway Segments outside Interchange Systems 100
Table 30: Candidate Models for Off Ramps 101
Table 31: Regression Results for Selected Models for Off Ramps 102
Table 32: Candidate Models for On Ramps 103
Table 33: Regression Results of Selected Models for On Ramps 104
Table 34: Candidate Models for Ramp Influence Areas 105
Table 35: Regression Results of Selected Models for Ramp Influence Areas 106
Table 36: Candidate Models for Weaving Sections 107
Table 37: Regression Results of Selected Models for Weaving Sections 108
Table 38: Candidate Models for Signalized Ramp Terminals 109
Table 39: Regression Results of Selected Models for Signalized Ramp Terminals 110
Table 40: Candidate Models for Unsignalized Ramp Terminals 111
Table 41: Regression Results of Selected Models for Unsignalized Ramp Terminals 111
Table 42: Summary of Selected Functional Form for the Collision Prediction Models 113
Table 43: Validation of SPFs for 6 Classes of High-Speed Roadway Segments in Saskatchewan 114
Table 44: Number of Observations in Estimation and Validation Datasets 116
Table 45: Societal Costs of Collisions for the Year 2010 by SGI 118
Table 46: Basic Freeways inside Interchanged System Network Screening Results 120
Table 47: Basic Freeways outside Interchange System Network Screening Results 121
Table 48: Off Ramps Network Screening Results 122
Table 49: On Ramps Network Screening Results 123
Table 50: Ramp Influence Areas Network Screening Results 124
Table 51: Weaving Sections Network Screening Results 125
Table 52: All High-Speed Roadway Segments Network Screening Results 126
Table 53: Ramp Terminals (signalized) Network Screening Results 129
Table 54: Ramp Terminals (unsignalized) Network Screening Results 130
Table 55: All Ramp Terminals Network Screening Results 131

LIST OF FIGURES

Figure 1: Variation in Collision Frequency 6
Figure 2: Ramp Terminal Configurations (Transportation Research Board 26
Figure 3: Cumulative Residual Plot for a Good Collision Prediction Model 44
Figure 4: Cumulative Residual Plot for a Poor Collisions Prediction Model 44
Figure 5: Expected and Excess Collision Frequencies 58
Figure 6: SMHI Control Section Naming System 65
Figure 7: Sample of Traffic Volume Information from Regina 67
Figure 8: Classification System for Freeways 73
Figure 9: Classification System for Interchanges 73
Figure 10: SMHI's Control Sections with Small Line Segments 79
Figure 11: Examples of More Than One Location Identifier for One Roadway 81
Figure 12: Sample of the Basic Framework for the Final Integrated Dataset 82
Figure 13: Difference in Collision Location in Two Different Spatial Datasets (a: Regina Spatial Dataset, b: Basic Framework of New Integrated Spatial Dataset) 84
Figure 14: Description of Factors used in Network Screening Methods: (Left) EPDO with EB Adjustment Method and (right) Excess EPDO with EB Adjustment Method 119
Figure 15: Top Ten Locations (Segments) Ranked using Excess EPDO Method 127
Figure 16: Top Ten Locations (Segments) Ranked using Expected EPDO Method 128
Figure 17: Top Ten Ramp Terminal Locations using Excess EPDO Method 132
Figure 18: Top Ten Ramp Terminal Locations using Expected EPDO Method 133
Figure 19: Top Ten Segments and Ramp Terminals using Excess EPDO Method 135
Figure 20: Top Ten Segments and Ramp Terminals using Expected EPDO Method 136

LIST OF ABBREVIATIONS

$\widehat{y_\imath}$	= Predicted number of collisions at site I, using corresponding SPF
y_i	= Observed number of collision at site i
μ	= mean
a, b, c and d	= Regression coefficient
a, b1, b2,..,bn	= Regression coefficients
A2	= Three-leg ramp terminal at two-quadrant Parclo A
A4	= Four-leg ramp terminal at four-quadrant Parclo A
AADT	= Annual Average Daily Traffic Volume
$AADT_{en}$	= Average annual daily traffic (vehicles/day) on entering ramp
$AADT_{ex}$	= Average annual daily traffic (vehicles/day) on exiting ramp
$AADT_{in}$	=Average annual daily traffic (vehicles/day) for crossroad leg inside interchange system
$AADT_{m}$	= Mainline freeway annual average daily traffic volume (vehicle/day)
$AADT_{out}$	= Average annual daily traffic (vehicles/day) for crossroad leg outside interchange system
$AADT_{r}$	= Ramp annual average daily traffic volume (vehicle/day)
$AADT_{total}$	= Sum of approach volume from ramps and crossroad
$AADT_{xrd}$	= Average annual daily traffic (vehicles/day) for crossroad
AASHTO	= American Association of State Highway and Transportation Officials
AIC	= Akaike's Information Criterion
B2	= Three-leg ramp terminal at two-quadrant Parclo B
B4	= Four-leg ramp terminal at four-quadrant Parclo B
BIC	= Bayesian Information Criterion
CCPDO	= Collision cost for PDO collision severity
CCy	= Collision cost for each severity, y
C-D	= Collector and Distributor
$C_{fw,n,y,z}$	= Calibration factor developed for specific roadway configurations, such as freeway segment, with ″n″ number of lanes, ″y″ collision type (i.e., ′sv′ single vehicle, ′mv′ multiple vehicle, or ′a′ all type), and ″z″ severity type (i.e., ′fi′ fatal and injury, and or ′pdo′ property damage only)
CMF	= Collision Modification Factor

C_n	= Annual correction factor for year n
$C_{n(FI)}$	= Annual correction factor for fatal and injury collisions
$C_{n(FI)}$	= Annual correction factor for fatal and injury collisions
$C_{n(total)}$	= Annual correction factor for total collisions
$C_{n(total)}$	= Annual correction factor for total collisions
COS	= City of Saskatoon
Cr	= Calibration factor
CURE	= Cumulative residual plots
$D3_{en}$	=Three-leg ramp terminal with diagonal entrance ramp
$D3_{ex}$	=Three-leg ramp terminal with diagonal exit ramp
D4	= Four-leg ramp terminal with diagonal ramps
Diff	= Difference in design speeds of freeway and ramps
Dummy	= Variable zero for non-split ramp approach or for split ramp approach with separate right turn having yield control
EB	= Empirical Bayes
EPDO	= Equivalent Property Damage Only
Excess EPDO	= Excess equivalent property damage only
$Excess_n$	= Excess expected collisions for year n
$Excess_y$	= Excess expected collisions for year n
$f_{fatal\ (weight)}$	= EPDO fatal weight factor
FI	= Fatal and Injury Collisions
$f_{injury\ (weight)}$	= EPDO injury weight factor
$f_{y\ (weight)}$	= EPDO weighting factor based on collision severity y
GIS	= Geographical Information System
GOF	= Goodness of fit
HCM	= Highway Capacity Manual
HSM	= Highway Safety Manual
IHSDM	= Interactive Highway Safety Design Model
ITE	= Institute of Transportation Engineers
j	= Number of years in the study
k	= Dispersion parameter
K	= Overdispersion parameter from the appropriate SPF
L	= Length of the deceleration lane

Lx = Average length of roadway segment

MAD = Mean Observed Deviation

MPB = Mean Prediction Bias

MSE = Mean Squared Error

MSPE = Mean Squared Prediction Error

MV = Multiple Vehicle

$N_{(FI)}$ = Predicted fatal and injury collisions

$N_{(PDO)}$ = Predicted PDO collisions

$N_{(total)}$ = Predicted total collisions

$N'_{(FI)}$ = Fatal and injury component of the total collisions

$N'_{(PDO)}$ = PDO component of the total collisions

NB = Negative Binomial

NCHRP = National Cooperative Highway Research Program

$N_{expected,1(FI)}$ = EB-adjusted expected average FI collision frequency for year 1

$N_{expected,1(total)}$ = EB-adjusted expected average total collision frequency for year 1

$N_{expected,n}$ = EB-adjusted expected average collision frequency for year n

$N_{expected,n\ (EPDO)}$ = EPDO expected average collision frequency for year n

$N_{expected,n\ (FI)}$ = EB-adjusted expected average FI collision frequency for year n.

$N_{expected,n\ (PDO)}$ = EB-adjusted expected average PDO collision frequency for year n

$N_{expected,n\ (total)}$ = EB-adjusted expected average total collision frequency for final year n

$N_{expected,n,(FI)}$ = EB-adjusted expected average FI collision frequency for year n

$N_{expected,n,(total)}$ = EB-adjusted expected average total collision frequency for year n

$N_{observed,(I)}$ = Observed number of injury collisions

$N_{observed,(F)}$ = Observed number of fatal collisions

$N_{observed,(FI)}$ = Observed number of fatal and injury collisions

$N_{observed,n\ (FI)}$ = Observed number of fatal and injury collisions for year n

$N_{observed,n\ (total)}$ = Observed number of total collisions for year n

$N_{predicted,1\ (FI)}$ = Predicted number of fatal and injury collisions for year 1

$N_{predicted,1\ (total)}$ = Predicted number of total collisions for year 1

$N_{predicted,n}$ = SPF-predicted average collision frequency for year n

$N_{predicted,n\ (FI)}$ = Predicted number of fatal and injury collisions for year n

$N_{predicted,n\ (total)}$ = Predicted number of total collisions for year n

$N_{predicted,y}$ = Predicted average collision frequency for severity type y

N_{spf} = Predicted average frequency of collisions for roadway categories, damage only

$N_{spf,fw,n,y,z}$ = Predicted average collision frequency of a freeway segment with "n" number of lanes, "y" collision type (i.e., 'sv' single vehicle, 'mv' multiple vehicle, or 'a' all types), and severity type "z" (i.e., ′fi′ fatal and injury, and or ′pdo′ property damage only)

N_y = Uncalibrated predicted number of collisions for severity type y

OECD = Organization for Economic Co-operation and Development

p = Ramp position (right and left hand)

PDO = Property damage only collisions

P_F = Proportion of observed fatal collisions within FI collisions

P_I = Proportion of observed injury collisions within FI collisions

R^2_{FT} = Freeman Tukey R-Square

RIA = Ramp Influence Area

RID = Roadway Identification

RSMP = Roadway safety management process

RTM = Regression to Mean

SGI = Saskatchewan Government Insurance

SMHI = Saskatchewan Ministry of Highways and Infrastructure

S = Posted speed of the mainline freeway segment (km/h)

SPF = Safety Performance Functions

SV = Single Vehicle

TDM = Transportation Data Model

TRB = Transportation Research Board

type = Type of ramp (on-ramp or off-ramp)

$W_{(FI)}$ = Weight factor for fatal and injury collisions

$W_{(total)}$ = Weight factor for total collisions

$W_{EPDO,FI}$ = EPDO weight factor for fatal and injury collisions

W-Sec = Weaving Section

Wy = Empirical Bayes weight for severity y

X = Set of geometric variables characterizing roadway segment

X1, X2….Xn = Geometric and traffic features variables

y = observed value

ΣAAADT = Annual average daily traffic volume sum of freeway and ramps (vehicle/day)

$\Sigma RAADT$ = Annual average daily traffic volume sum of ramps (vehicle/day)

n = Locations count in the dataset

p = Degree of freedom

γ = gamma function

CHAPTER 1: INTRODUCTION

1.1. Background

1.1.1. Collisions History on Saskatchewan Roads

Transport Canada reported that Saskatchewan experienced the highest fatality rate associated with traffic collisions among all Canadian provinces and territories (i.e., 12.6 fatalities/100,000 population) and the third highest injury rate (i.e. 635.6 injuries per 100,000 population) (Canada, 2013). For instance, Saskatchewan Government Insurance (SGI) reported that Saskatchewan experienced a total of 31,741 collisions in 2013, which include 115 fatal collisions and 5,290 injury collisions (SGI, 2014). These statistics, compared to 2012, showed an increase of 5.02% in total collisions, and decreases of 24.3% and 4.9% in fatal and injury collisions, respectively. Although the number of fatal and injury collisions declined in 2013, PDO collisions rose by 7.4% from 2012 levels, and from 2011 to 2013, the total number of collision showed an increasing trend each year (Table 1).

Table 1: Three-Year Collision Summary ((SGI), 2011 to 2013)

Collision Type	Year			Change	Change
	2011	2012	2013	(2011 to 2012)	(2012 to 2013)
Property Damage Only	24,372	24,511	26,336	0.57%	7.45%
Injury	5,166	5,562	5,290	7.67%	-4.89%
Fatal	137	152	115	10.95%	-24.34%
Total	**29,675**	**30,225**	**31,741**	**1.85%**	**5.02%**

Looking at the total number of collisions that occurred only on provincial highway systems in Saskatchewan, an increasing trend can be observed between 2007 and 2013, except in 2012 when a decrease was observed. Table 2 shows a five-year summary of the total collisions from 2009 to 2013 for each type of road system, including provincial highways in Saskatchewan (SGI, 2014).

Table 2: Three-Year Collision Summary by Road System ((SGI), 2009 to 2013)

Road System	Year					
	2009	**2009***	**2010**	**2011**	**2012**	**2013**
Provincial Highways	13,107	6,923	7,008	7,621	6,867	7,737
Urban Roads	33,133	19,898	18,347	18,717	19,857	20,607
Rural Roads	6,178	2,950	2,579	2,368	2,374	2,565
Other Roads	1,811	834	864	945	1,127	832
Total	**54,229**	**30,605**	**28,798**	**29,651**	**30,225**	**31,741**

*A decreased reporting threshold in 2010 resulted in a significant change in the 2009 total number of collisions; the second column for 2009 shows collisions after the reduction in reporting threshold.

Table 3 shows a summary of losses (victims) caused by collision on provincial highways for the period of 2011 to 2013, during which Saskatchewan experienced an average of 106 fatalities, 1,854 injuries, and 6,080 PDO collisions each year.

Table 3: Three-Year Summary of Losses on Provincial Highway System ((SGI), 2011-2012)

Victims on Provincial Highway System	Year			Average Losses
	2011	**2012**	**2013**	
Fatalities	106	128	85	106
Injuries	1,880	1,749	1,934	1,854
Property Damage Only	6,280	5,588	6,373	6,080

Collisions result in various direct costs, such as the costs associated with property damage (e.g., vehicle repair costs, emergency response costs, etc.), medical expenses, travel expenses for appointments, and income replacement in cases when more than seven consecutive work days are missed (Flores et al., 2013). Collisions also result in indirect costs known as societal costs. This includes the cost that a community pays for preventing and reducing the risks associated with a collision involving serious injury and or fatality. Societal costs also include loss of productivity, pain, suffering and grief, loss of quality of life, the value of statistical life, etc. (Leur, 2010). The American Association of State Highway and Transportation Officials (AASHTO) provides estimated societal costs based on collision severity that can be used to monetize impacts of collisions on society and where available locally-developed societal costs can be used (AASHTO, 2010). Table 4 shows the societal costs estimated based on the locally developed SGI societal costs in 2010 dollars for a fatal, injury, and PDO collisions on provincial highways in Saskatchewan. The average annual cost is calculated by multiplying the average losses, given in Table 3, by the societal costs provided by SGI. The results show that Saskatchewan incurred a loss of $890.95 million in only one year from collisions on provincial highways. This cost represents the estimated societal costs from collisions per year based on SGI's collision data from 2011 to 2013.

Table 4: Three-year Average Societal Cost of Collisions on Provincial Highway System

Severity of Collision	**Societal Cost per Collision* (Million)**	**Average Losses**	**Annual Average Cost (Million)**
Fatal	$5.54	106	$589.09
Injury	$0.13	1,854	$241.06
PDO	$0.01	6,080	$60.80
Total loss due to victims on provincial highways			**$890.95**

*Societal cost per collision severity in 2010 dollars ((SGI), 2007-2011)

To improve safety on roads, it is imperative for road safety professionals to have a systematic, data-driven approach (e.g., six-step road safety management process) capable of quantifying roadway safety and identifying locations that require safety improvement to reduce the number of collisions to an acceptable limit. Without such an approach, limited resources are misused for the installation of safety improvement countermeasures on each road within a targeted network.

Although the agencies responsible for improving safety on Saskatchewan roadways have a systematic, data-driven approach for their urban road network in the City of Regina and the City of Saskatoon, none of the agencies, including Saskatchewan Government Insurance (SGI), has implemented such an approach for high-speed roadways in the province to estimate the expected number of collisions and identify problem areas for their decision-making process. The outcomes of this research (i.e., a set of safety performance function (SPFs) for high-speed roadways in Saskatchewan) will enable agencies in Saskatchewan to quantify safety for their roadway safety improvement projects. This study uses local data such as local traffic volumes, local collision data, the geometry of Saskatchewan highways, and other features to produce the best-fitting collision prediction models (i.e., SPFs). The developed SPFs can be employed in various engineering processes such as transportation planning, design, and operation projects and will improve the effectiveness of allocating scarce resources by directing them where they are most needed.

1.1.2. Measure of Road Safety

The 2010 Highway Safety Manual (HSM) introduced a six-step roadway safety management process (RSMP) through which targeted sites are evaluated for potential improvements. The RSMP is comprised of (1) Network Screening, (2) Diagnosis, (3) Select Countermeasures, (4) Economic Appraisal, (5) Prioritize Projects, and (6) Safety Effectiveness Evaluation.

The first step in the process (i.e., network screening) identifies locations by quantifying safety using SPFs within the targeted road network that have the potential and need for safety improvements. This step is important because it is neither practical nor feasible to perform an engineering evaluation for an entire road network and invest a significant amount of resources on data collection, analysis, and management. Locations identified in this step are then further investigated by following the remaining five steps of the RSMP. This first step identifies locations where safety improvements are needed, so that real safety benefits are achieved and incorrect screening and associated unnecessary budgetary expenditures can be avoided.

In Saskatchewan, transportation agencies such as the Saskatchewan Ministry of Highways and Infrastructure (SMHI) and the City of Saskatoon (COS) traditionally use the collision rate method and collision frequency method, respectively. Both these methods have proven limitations. For example, the collision rate method assumes a linear relation between collision frequency and traffic volume, a relationship that many researchers have proven is actually nonlinear. Furthermore, the collision rate method uses the AADT (exposure) in a denominator (i.e., by dividing the collision frequency by the exposure), which can result in locations with higher traffic volumes being underemphasized and sites with lower traffic volumes being overemphasized (Hamidi et al., 2010; Xiao Qin et al., 2005). In other words, sites with high collision rates do not necessarily have high collision frequencies, and sites with high collision frequencies might get overlooked when they have considerably high AADT.

On the other hand, the collision frequency method usually uses data for a short period (e.g., 2 to 3 years). This can neglect the long-term average collision frequency and ignores the RTM effect. The RTM effect is a statistical phenomenon present in collision records which says that a location observing a high number of collision in one year will show a decreased number of collisions in the next year (close to the long-term mean). Likewise, locations observing a low number of collisions in one year will show an increased number of collision the next year (close to the long-term mean)

Due to the randomness of collision occurrences, a short-term average collision frequency does not necessarily represent a real underlying trend and may produce misleading results about the safety performance of targeted locations, either overestimating or underestimating their safety. Another drawback of the collision frequency method is that it does not account for exposure (i.e., AADT). In this process, collisions are converted into equivalent property damage only (EPDO) collisions by assigning a weight to fatal and injury collisions. The weight is assigned based on the societal costs of a fatal or injury collisions equivalent to PDO collisions. As a result, locations with high rates of fatal or injury collisions are overemphasized in comparison to locations with higher rates of PDO collisions due to the assigned weight of fatal or injury collisions. Figure 1 shows the random error generated due to the short-term average of collision frequency, which varies from the long-term average and may not show the RTM effect.

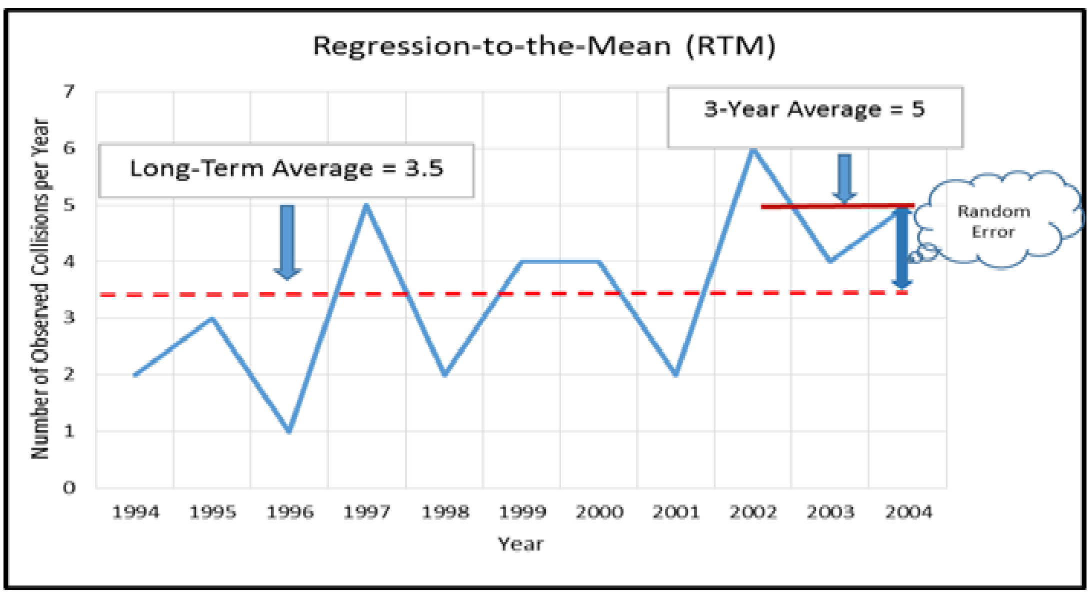


Figure 1: Variation in Collision Frequency (Susan Herbel, 2010).

The HSM provides a number of network screening measures, which include the expected collision frequency method. These new techniques duly account for the regression to the mean (RTM) effect. The expected collisions frequency method provides a stable list of locations that have a potential for safety improvements. The model for expected number of collisions addresses the biases of the collision frequency and/or collision rate methods by using safety performance functions. The expected number of collisions is considered as a long-term average number of collisions for a targeted location in a roadway network (e.g. freeway segment, speed-change lanes, ramps, ramp terminals, etc.). This network screening technique combines the actual number of collisions with the predicted number of collisions based on statistical models (i.e. SPFs) from representative locations. In other words, the expected number of collisions used to quantify the safety of a particular location is a combination of historical collision records and the predicted value obtained from a statistical model with a given AADT (Hauer, 1997; HSM, 2010). The historical collision data and the predicted number of collisions are combined using the Empirical Bayes (EB) method.

1.1.3. Rationale for the Development of Safety Performance Functions

To address safety concerns on Canadian roadways, including those in Saskatchewan, transportation professionals require a systematic process to quantify safety on roads. The AASHTO Highway Safety Manual (AASHTO, 2010) provides technical background and safety management methods for collisions on different types of roadways. The roadways types include 2-lane rural highway, multilane rural highways, and urban and suburban arterials. Facilities like freeways and interchanges are addressed by the National Cooperative Highway Research Program (NCHRP) Report 17-45. The report has been finalized and approved by AASHTO, and it will form the future chapters of HSM (Transportation Research Board (TRB), 2014).

The HSM and NCHRP report provide a number of SPFs for different roadway categories based on data from different states in the United States. These SPFs are intended for jurisdictions that do not have the resources or expertise to develop their own local SPFs. Therefore, these generic SPFs may not necessarily appropriately account for local conditions in jurisdictions outside the US. To address this issue, the HSM and NCHRP report provided calibration factors to be applied to the SPFs to account for local conditions, such as differences in collision frequency for different jurisdictions or different reporting thresholds in different jurisdictions (HSM, page C-18). Calibrated SPFs can somewhat better reflect local conditions and provide an expected number of collisions that is relatively more accurate (Lyon et al., 2005; Raghavan Srinivasan & Carter, 2011; Young & Park, 2013).

To verify the validity of generic SPFs and calibrated SPFs, various studies have been conducted, and these studies have documented a better fit to the local conditions for calibrated SPFs than for generic SPFs (Lu et al., 2012; Lyon et al., 2005; B. Srinivasan et al., 2011). Although the HSM and NCHRP 17-45 report provide calibration factors, they simultaneously strongly suggest that jurisdictions with the expertise and resources should develop their own set of SPFs based on their local data, which better reflect local conditions (AASHTO, 2010).

A study conducted by the Colorado Department of Transport (DOT) explored the potential benefits of developing of jurisdiction-specific SPFs (Colorado, 2009). In the study, the calibrated SPFs provided by the HSM were applied to signalized intersections and compared to local SPFs developed based on the state's local collision data and AADT. For this study, models were developed for ten interchange categories (Colorado, 2009). The study concluded that the local SPFs are more representative of local conditions and provide a more accurate expected collision frequency than the calibrated SPFs from the HSM.

Similar observations were reported by a study undertaken to examine the applicability of the Interactive Highway Safety Design Model (IHSDM) for a 2-lane rural road in New Brunswick (Marleau & Hildebrand, 2010). The IHSDM uses the same calibration procedure as provided in the HSM (2010). The study found that the calibrated SPFs overestimated collision frequencies. Lu et al. (2012) also found that the SPFs developed in Florida for road segments, intersections and ramps produce more accurate expected number of collisions than HSM's national default SPFs with local calibration factors. Garber and Rivera (2010), while checking the transferability of Minnesota SPFs to roadways in Virginia, concluded that SPFs developed for target intersections based on local data in Virginia were more representative of the observed collision data, even though the SPFs developed in Minnesota were modified based on the difference in roadside characteristics (i.e. topography and the difference in database; traffic, collision, and road characteristics). The authors, therefore, recommend using SPFs that are developed based on local data to account better for local conditions. Persaud et al. (2012) also concluded that local SPFs are superior to calibrated HSM models.

There is a need for the development of local SPFs for high-speed roadways in Saskatchewan, since the HSM provides a number of SPFs for such road facilities. One possible answer could be that the roadway classification needed to apply HSM SPFs is different from the roadway classification used in Saskatchewan. Furthermore, the application of HSM SPFs requires detailed information about roadway geometry, which is another barrier to applying the generic HSM SPFs. For example, the HSM's SPFs for speed change lanes requires detailed roadway geometric information (i.e., the length of a speed change lane from tapper point to gore point, etc.), which is not readily available.

Local high-speed SPFs are also required considering that urban (intercity) traffic characteristics are very different from high-speed (intracity) roadway characteristics. They have different traffic patterns, AADTs, operating speeds, traffic controls, conflict points, roadway geometries, collision configurations, etc. The operating speed on intercity roadways is usually less than the operating speed on high-speed roadways, whereas traffic volumes are higher on urban roadways, compared to high-speed roadways. Due to the high exposure (AADT) on the intercity roadways, the probability that a collision will occur higher, but the severity of such collisions is likely to be low due to the lower operating speed. On the other hand, AADT on a high-speed roadway is low, meaning that collision probability is lower, but, due to the high operating speed, the severity of a collision is likely to be high.

In a study based in Saskatchewan, Young and Park (2013) compared and analyzed the performance of SPFs for three categories of intersections in Regina: 3-leg unsignalized, 4-leg unsignalized, and 3- and 4-leg signalized. The comparison was made by using local data among uncalibrated HSM SPFs, calibrated HSM SPFs, and jurisdiction-specific SPFs. The study found that, as the AADT increases, the jurisdiction-specific SPFs outperformed both the uncalibrated and calibrated SPFs.

A review of the literature related to using the expected collision frequency method for quantifying safety on roadways in Saskatchewan reveals that local SPFs have been developed for intersections, rural roads, and urban roads. However, the same attention has not been paid to developing local SPFs for freeways and interchanges (i.e. ramps, speed-change lanes, etc.). The research presented herein bridges this gap in knowledge related to evaluating and quantifying safety for high-speed roadways and interchanges in Saskatchewan. The safety evaluation work in this study can be used for evaluation of geometric design alternatives, selection of countermeasures to improve the safety of high-speed roadways in Saskatchewan, and an evaluation of the safety once countermeasures are implemented.

The findings of this study can help roadway safety professionals make better-informed decisions supported by a scientific and data-driven approach, and allow them to communicate with the public and other stakeholders with confidence. The developed SPFs can be used to estimate an expected number of collisions per year for target locations (i.e., freeways inside interchange system, freeways outside interchange system, off ramps, on ramps, ramp influence areas, weaving sections, signalized ramp terminals, and unsignalized ramp terminals) for three severity types (i.e. total collisions, fatal-injury collisions, and PDO collisions.

The outcomes in the shape of collision prediction models for fatal and injury collisions, and not the fatality or injuries, will provide the basis for improved understanding of the underlying factors that contribute to the occurrence of collisions. It may be noted that the prediction models for fatalities or injuries will be dependant on the number of occupants in a vehicle. For example, a car involved in a fatal collision may have a smaller number of fatalities compared to a collision involving a bus with a large number of occupants. Since a number of occupants in a vehicle cannot be controlled or reduced, developing models for fatal and injury collisions instead of the numbers of fatalities or injuries appears more relevant. Reducing the number of fatal and injury collisions will in turn reduce the number of fatalities or injuries.

For high-speed roadways in Saskatchewan, there are certain technical challenges associated with compiling a database to be used in developing local SPFs. For example, most jurisdictions in Saskatchewan do not collect or maintain data solely for the purpose of safety analysis, and the quality and quantity of data available has an influence on the outcomes of the study. Missing annual traffic volumes and lacking information related to collision records may affect the expected number of collisions given by the locally developed SPFs. The 2010 Traffic Characteristics Report of Saskatoon mentions that the city has five permanent count stations that frequently record hourly traffic volumes and additional 546 short-term count stations that record traffic counts every three years (City of Saskatoon, 2010). Traffic volume data for certain minor legs of interchanges within the study area are missing.

A second technical challenge related to developing SPFs for high-speed roadways in Saskatchewan concerns the identification of the locations of collisions. The spatial datasets of the City of Saskatoon and the City of Regina use single point in the middle of a roadway segment as a location identifier, regardless of the direction of travel or location of collision along the roadway segment. Collisions occurring anywhere on the segment are therefore assigned to that location identifier, regardless of the actual location of collision along that segment. Likewise, there is only one point representing a ramp terminal, even if the terminal is split into more than one terminal point. Thus, any collision occurring at any of the terminal points of the ramp is assigned the single terminal point (i.e. location identifier) of the entire ramp terminal. The roadway classification provided in spatial datasets in Saskatchewan is different from the classification used for this research, and all the roadways used in this study therefore required reclassification. These technical challenges may represent a limitation in the development of very accurate SPFs for all configurations of high-speed roadways in Saskatchewan.

1.1.4. Visualization of Collisions using ArcGIS

The use of GIS technology is gaining popularity as an aid in decision-making in a variety of disciplines, including road safety (John M Bigham et al., 2009). The power of GIS to display and manage the complexity of a large amount of data has been helpful in making complex decisions, promoting safety, and improving the returns of public resource investments (Kim & Levine, 1996). The City of Regina, City of Saskatoon, and SMHI maintain that using GIS tools are advantageous in managing, displaying and analyzing large datasets for road networks within their respective jurisdiction (location identifiers, roadway classification, traffic controls, traffic directions, geometric design features, etc.). 80% of City of Saskatoon's municipal data can be spatially referenced (City of Saskatoon). Therefore, the SPFs developed in this research can be linked to the existing GIS datasets of all the authorities within the study area. The outcomes of the SPFs can be integrated into the spatial datasets to display expected collision frequencies based on collision severity and roadway configuration for all of the roadways within the scope of this research. This new GIS-based tool will help to easily identify or visualize problem areas on high-speed roadways in Saskatchewan.

1.2. Study Objectives

The goal of this research is to develop and propose a set of SPFs for high-speed roadways in Saskatchewan, including for provincial highways and expressways in Saskatoon and Regina.

To achieve this goal, the following four specific objectives were developed:

i. Develop an integrated and ready-to-use database from three different datasets (i.e. spatial dataset, traffic volume dataset, and historical collision records);

ii. Develop a set of SPFs for high-speed roadways that takes into account roadway configuration and collision severity, which can be used in various road safety engineering practices;

iii. Perform safety network screening based on the expected number of collisions produced by the developed SPFs; and

iv. Develop various GIS maps that highlight particularly unsafe road segments and/or locations using ArcGIS.

1.3. Study Scope

The scope of this study is focused on network screening and the development of local SPFs for high-speed roadways in Saskatchewan (i.e., the roadways within interchange systems and roadway segments approaching interchanges). The SPFs are developed for the following eight different roadway configurations:

- For freeways:
  - Freeways inside interchange area;
  - Freeways outside interchange area;
  - Ramp influence areas; and
  - Weaving sections.
- For interchanges:
  - Off ramps;
  - On ramps;
  - Signalized ramp terminals; and
  - Unsignalized ramp terminal.

The cross section of a high-speed roadway considered for this study is based on the SMHI's typical cross section. The cross-sectional elements of typical high-speed roadways in

Saskatchewan have a lane width of 3.7 m with a 3.0-m right shoulder width and 1.0-m left shoulder width. The cross slope of the traveled way and the left shoulder is 2%, and the cross slope of the right shoulder is 5%. Each traveled way has two lanes in each travel direction. The posted speeds for freeways, ramp influence areas, and weaving sections range from 80 to 110 km/hr, and approximately 79% of the posted speeds for these roadways vary between 90 and 110 km/hr. The ramps have one lane with a typical width of 5.0 m and posted speeds ranging from 30 to 100 km/hr, and more than 71% of the posted speeds on ramps vary between 50 and 80 km/hr.

The collisions used for analysis are only vehicle to vehicle collisions; collisions involving pedestrian, bicycles, animals, and trains are not included in this study. For the study, no new data was collected from a field. Only traffic and collision data, which were collected and maintained by governing agencies (e.g., the Ministry of Infrastructure and Transportation and SGI), are used. However, for some locations where traffic control information is lacking, Google Earth® or Google Maps® were used to collect the information, which was assumed to be the most up-to-date traffic control information.

The high-speed Saskatchewan roadways included in the scope of this study are (1) Highway 1, (2) Highway 11, (3) Highway 16, (4) Ring Road and Lewvan Drive in Regina, (5) Circle Drive in Saskatoon, and (6) Highways 2 and 3 in Prince Albert. The geographical scope of the study is presented on a map in Appendix A.

1.4. Research Benefits

Safety performance function's development in Saskatchewan have been mainly focused on intersections, arterial roads, rural roads and not a single study has been conducted for freeways and interchanges. In Saskatchewan, for instance, SPFs for intersections, arterial roads, and collector roads in urban settings have been developed for Saskatoon and Regina (Parisien, 2012; Young & Park, 2013), but no attempt has been made to develop SPFs for the high-speed roadways in the province. As a result, this study has generated new knowledge that will allow local road safety engineers in Saskatchewan develop a better understanding of road safety on local high-speed roadways. The following is a list of beneficial applications of the SPFs developed in this study:

- For new high-speed roadways and interchanges, the developed SPFs can estimate the level of safety of each alternative route option by providing an expected number of future collisions for each option;

- For existing highway facilities, the developed SPFs can allow engineers to perform network safety screenings that can determine the segments and interchanges that may require safety improvements; and

- The developed SPFs can help evaluate the performance of safety countermeasures by allowing roadway safety officials to compare the expected number of collisions before and after the implementation of potential countermeasures. Governing agencies can quantitatively estimate the effectiveness of proposed countermeasures using local SPFs.

Other benefits may include the formulation of more effective safety policies and regulations, as well as the development of more robust traffic enforcement plans. In addition, government agencies in Saskatchewan may use the developed SPFs to allocate their scarce resources better where they are expected to produce the highest safety benefits.

1.5. Thesis Layout

In the next chapter (Chapter 2), the scientific literature concerning methodologies for developing SPFs is reviewed and the most appropriate method for developing SPFs in this study is selected based on that literature review. Chapter 3 describes the local spatial, collision, and traffic volume datasets used to develop local SPFs for high-speed roadways in Saskatchewan. Chapter 3 also describes the steps required to compile an integrated database and the roadway segmentation process used for the development of those SPFs.

Chapter 4 presents the SPF development and validation process for high-speed roadways in Saskatchewan. In other words, the detailed procedures for developing and analyzing local SPFs are discussed.

The product of the developed SPFs (i.e., the expected number of collisions) are presented in Chapter 5 using GIS maps that show the results of network safety screening (i.e., collision-prone locations known as hotspots). Finally, Chapter 6 presents the limitations of this study, as well as its conclusions and recommendations for future work.

CHAPTER 2: LITERATURE REVIEW

The chapter describes development of safety performance functions for different roadway classifications and collision severities. Different functional forms used in the HSM and by past transportation safety researchers are described along with the data requirements. Different commonly used methods for developing best-fitting models and validating those models are also described. In the second part of this chapter, two network screening techniques, the EB-adjusted EPDO average collision frequency method and the excess expected average collision frequency method, are discussed.

2.1. Safety Performance Functions

Safety Performance Functions (SPFs) are mathematical models developed through regression analysis that relate roadway features such as traffic volume and geometric and traffic control features with observed collision frequencies (AASHTO, 2010). The regression analysis is often done using generalized regression techniques such as the negative binomial model. The network screening process of the RSMP uses SPFs to predict the expected number of collisions on roadways similar in type to the roadways for which they are developed.

SPFs can be developed for a number of roadway configurations (e.g., freeway segment, ramp segments, ramp terminals) traffic control features (e.g., signalized or unsignalized), levels of collision severity (e.g., fatal or injury collisions), and area types (e.g., urban or rural). A wide variety of inputs are used for SPF development, which include traffic volumes, posted speeds, number of lanes, lane widths, median widths, area type, etc. (Cafiso et al., 2010; Hauer & Bamfo, 1997; Hu & Donnell, 2010; Transportation Research Board (TRB), 2014). The input variables that are available in local databases have a vital impact on the outcomes of the SPF development. In other words, the quality and quantity of data has an effect on the prediction power of the model. Ideally, the jurisdiction intending to develop its own set of local SPFs should have access to all the necessary data to generate a high-quality set of SPFs. The availability of such detailed data, however, is not common in most cases since jurisdictions collect data for various purposes and not specifically for the development of SPFs.

Traffic volume (i.e., Annual Average Daily Traffic (AADT)) is the most frequently used input variable along with length when SPFs are developed specifically for a road segment (AASHTO, 2010; Chang, 2005; Lu et al., 2013b; Transportation Research Board (TRB), 2014). AADT values encompass various road features and are therefore widely used by transportation professionals as an essential input for different planning, design, and operation purposes. For example, AADT is used in geometric design to select the number of lanes to construct for a certain roadway, in structural design for determining the thicknesses of different pavement layers, and in the selection of traffic control devices/features. AADT can also reflect the roadway type since roads with small AADTs are likely to be smaller local roads and roads with higher AADTs are likely to be major arterial roads.

2.1.1. Literature Review on Safety Performance Functions

The NCHRP Report 17-45, provides a number of SPFs for basic freeways (i.e., through-lanes, and speed-change lanes) and interchange systems (i.e., ramps and ramp terminals). The types of roadway configurations for which SPFs are available in the report, based on AADT ranges, area types and number of lanes, are presented in Table 5 below.

Table 5: NCHRP Report 17-45's Freeway Configurations with Available SPFs (Transportation Research Board (TRB), 2014)

Roadway Configurations	Area Type	Number of Through Lanes	AADT Range (veh/day)
Freeway Segments and Speed-Change Lanes	Rural	4	0-73,000
		6	0-130,000
		8	0-190,000
	Urban	4	0-110,000
		6	0-180,000
		8	0-270,000
		10	0-310,000

Table 6 shows the list of ramp segment configurations for which SPFs are available in the NCHRP Report, according to area type and number of lanes in the ramps.

Table 6: NCHRP Report 17-45's Ramp Segment Configurations with Available SPFs (Transportation Research Board (TRB), 2014)

Roadway Configurations	Area Type	Number of Lanes	AADT Range (veh/day)
Ramp Segments	Rural	1	0-7,000
	Urban	1	0-18,000
		2	0-32,000

Table 7 shows the list of ramp terminal configurations for which of SPFs are available in the NCHRP report, in terms of traffic geometric design features and traffic control feature (stop control or Signal control) with crossroad and all ramps AADT:

Table 7: NCHRP Report 17-45's Ramp Terminal Configurations with Available SPFs (Transportation Research Board (TRB), 2014)

Ramp Terminal Configurations	Control Type	AADT Range Crossroad (veh/day)	AADT Range Total All Ramps (veh/day)
Three-leg Terminal with Diagonal Exit Ramp *(D3ex)*	Stop	0-22,000	0-8,000
	Signalized	0-34,000	0-16,000
Three-leg Terminal with Diagonal Entrance Ramp *(D3en)*	Stop	0-22,000	0-15,000
	Signalized	0-29,000	0-21,000
Four-leg Terminal with Diagonal Ramps *(D4)*	Stop	0-18,000	0-10,000
	Signalized	0-47,000	0-31,000
Four-leg Terminal at Four Quadrant Parclo A *(A4)*	Stop	0-21,000	0-12,000
	Signalized	0-71,000	0-30,000
Four-leg Terminal at Four Quadrant Parclo B *(B4)* Four	Stop	0-20,000	0-12,000
	Signalized	0-45,000	0-29,000
Three-leg Terminal at Two Quadrant Parclo A *(A2)*	Stop	0-17,000	0-12,000
	Signalized	0-46,000	0-25,000
Three-leg Terminal at Two Quadrant Parclo B *(B2)*	Stop	0-26,000	0-14,000
	Signalized	0-44,000	0-22,000

The predictive models which produce a value for the expected number of collisions (collisions/year) combine SPFs, collision modification factors (CMFs), and calibration factors for a given roadway facility. The predictive model for freeway segments is given by equations 2-1 to 2-5 (Transportation Research Board (TRB), 2014):

$$N_{p,fw,n,a,s} = N_{p,fw,n,mv,fi} + N_{p,fw,n,sv,fi} + N_{p,fw,n,mv,pdo} + N_{p,fw,n,sv,pdo} \quad \text{[Equation 2-1]}$$

$$\begin{aligned} N_{p,fw,n,mv,fi} = {} & C_{fw,n,mv,fi} \times N_{spf,fw,n,mv,fi} \\ & \times \left(CMF_{1,fw,n,mv,fi} \times \ldots \times CMF_{m,fw,n,mv,fi}\right) \\ & \times \left(CMF_{1,fw,n,a,fi} \times \ldots \times CMF_{m,fw,n,a,fi}\right) \end{aligned} \quad \text{[Equation 2-2]}$$

$$N_{p,fw,n,sv,fi} = C_{fw,n,sv,fi} \times N_{spf,fw,n,sv,fi} \times \left(CMF_{1,fw,n,sv,fi} \times \ldots \times CMF_{m,fw,n,sv,fi}\right) \times \left(CMF_{1,fw,n,a,fi} \times \ldots \times CMF_{m,fw,n,a,fi}\right)$$ [Equation 2-3]

$$N_{p,fw,n,mv,pdo} = C_{fw,n,mv,pdo} \times N_{spf,fw,n,v,pdo} \times \left(CMF_{1,fw,n,mv,pdo} \times \ldots \times CMF_{m,fw,n,mv,pdo}\right) \times \left(CMF_{1,fw,n,a,pdo} \times \ldots \times CMF_{m,fw,n,a,pdo}\right)$$ [Equation 2-4]

$$N_{p,fw,n,sv,pdo} = C_{fw,n,sv,pdo} \times N_{spf,fw,n,sv,pdo} \times \left(CMF_{1,fw,n,sv,pdo} \times \ldots \times CMF_{m,fw,n,sv,pdo}\right) \times \left(CMF_{1,fw,n,a,pdo} \times \ldots \times CMF_{m,fw,n,a,pdo}\right)$$ [Equation 2-5]

Where:

$N_{p,fw,n,y,z}$ = predicted average collision frequency of a freeway segment with "*n*" number of lanes, "*y*" collision type (i.e., '*sv*' single vehicle, '*mv*' multiple vehicle, or '*a*' all types), for severity type "z" (i.e., '*fi*' fatal and injury, and/or '*pdo*' property damage only);

$N_{spf,fw,n,y,z}$ = predicted average collision frequency of a targeted freeway segment with "*n*" number of lanes, "*y*" collision type (i.e., '*sv*' single vehicle, "*mv*" multiple vehicle, or '*a*' all type), and "z" severity type (i.e. '*fi*' fatal and injury, and/or '*pdo*' property damage only);

$CMF_{m,fw,n,y,z}$ = collision modification factor for geometric feature "*m*" for a freeway segment with "*n*" number of lanes, "*y*" collision type (i.e., '*sv*' single vehicle, '*mv*' multiple vehicle, or '*a*' all type), and "z" severity type (i.e., '*fi*' fatal and injury, and/or '*pdo*' property damage only); and

$C_{fw,n,y,z}$ = calibration factor developed for specific roadway configuration such as freeway segment in a specific area with "*n*" number of lanes, "*y*" collision type (i.e., '*sv*' single vehicle, '*mv*' multiple vehicle, or '*a*' all type), and "*z*" severity type (i.e., '*fi*' fatal and injury, and/or '*pdo*' property damage only).

Equation 2-1 represents the sum of four collision frequencies: (1) fatal and injury "*fi*" collisions involving multiple vehicles "*mv*", (2) fatal and injury collisions "*fi*" involving a single vehicle "*sv*", (3) property damage only "*pdo*" collisions involving multiple vehicles "*mv*", and (4) property damage only "*pdo*" collisions involving a single vehicle.

The functional forms of SPFs for freeway segments, speed-change lanes (ramp entrances and exits) or collector-distributor roads (C-D) are given by Equation 2-6 (Transportation Research Board (TRB), 2014).

$$N_{spf} = L_x \times e^{a+b\times ln(c\times AADT)} \qquad \text{[Equation 2-6]}$$

Where:

N_{spf} = Predicted average frequency of collisions for freeways, speed change lanes, and collector-distributor roads;

L_x = Average length of freeway, ramp entrances or exits and speed change lanes;

AADT = Average annual daily traffic (vehicles/day) on a given segment (i.e., freeway, speed-change lane, or collector-distributor road); and

a, *b* = Regression coefficients.

The predicted numbers of collisions are categorized by the number of vehicles involved (multiple or single vehicle collisions), number of lanes (4-10 lanes), severity type (*fi* and *pdo*), and area type (urban and rural). Equation 2-6 assumes a linear relation with length and a nonlinear relationship with AADT.

The functional forms of SPFs for ramp segments are given by Equation 2-7 (Transportation Research Board (TRB), 2014).

$$N_{spf} = L \times e^{[a+b\times ln(c\times AADT)+d(c\times AADT)]} \quad \text{[Equation 2-7]}$$

Where:

N_{spf}	= Predicted average frequency of collisions for the ramp segment;
L	= Length of ramp segment;
$AADT$	= Average annual daily traffic (vehicles/day) on ramp segment; and
a, b, c and *d*	= Regression coefficients.

The predicted number of collisions are categorized into one-lane entrance/exit ramps or two-lane entrance/exit ramps, severity type (*fi* and *pdo*), and area type (urban and rural). Equation 2-7 assumes a linear relation with length and a nonlinear relationship with AADT.

The functional forms of SPFs for ramp terminals is given by Equation 2-8 (Transportation Research Board (TRB), 2014).

$$N_{spf} = e^{[a+b\times ln(c\times AADT_{xrd})+d\times ln(c\times AADT_{ext}+c\times AADT_{ent})]} \quad \text{[Equation 2-8]}$$

With:

$$AADT_{xrd} = 0.5 \times (AADT_{in} + AADT_{out}) \quad \text{[Equation 2-9]}$$

Where:

N_{spf} = Predicted average frequency of collisions for ramp terminal;

$AADT_{xrd}$ = Average annual daily traffic (vehicles/day) for crossroad;

$AADT_{in}$ = Average annual daily traffic (vehicles/day) for crossroad leg inside interchange system;

$AADT_{out}$ = Average annual daily traffic (vehicles/day) for crossroad leg outside interchange system;

$AADT_{ex}$ = Average annual daily traffic (vehicles/day) on exiting ramp;

$AADT_{en}$ = Average annual daily traffic (vehicles/day) on entering ramp; and

a, b, c and *d* = Regression coefficients.

The predicted number of collisions are categorized into six different types of ramp terminal configurations: (1) three-leg ramp terminal with diagonal exit or entrance ramp ($D3_{ex}$ and $D3_{en}$), (2) four-leg ramp terminal with diagonal ramps (D4), (3) four-leg ramp terminal at four-quadrant Parclo A (A4), (4) four-leg ramp terminal at four-quadrant Parclo B (B4), (5) three-leg ramp terminal at two quadrant Parclo A (A2), and (6) three-leg ramp terminal at two quadrant Parclo B (B2). Equation 2-8 assumes a nonlinear relation with AADT(s). Figure 2 shows ramp terminal configurations given in the NCHRP Report 17-45.

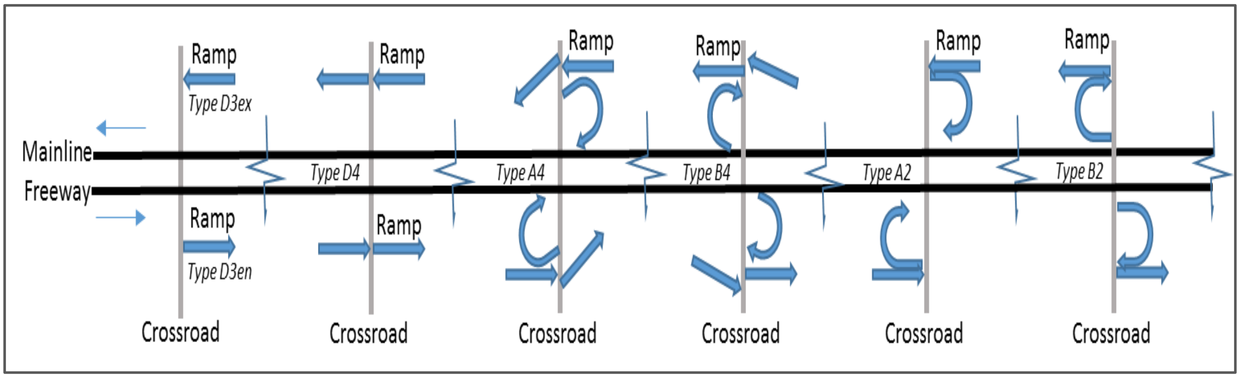


Figure 2: Ramp Terminal Configurations (Transportation Research Board (TRB), 2014)

The SPFs provided in the NCHRP report were developed using a large dataset from various states in the US. These SPFs were compared with local SPFs developed by researchers, including researchers in the US, and local SPFs generally provide a better fit to observed data in a given locality than the NCHRP SPFs.

In addition, collecting, analyzing and interpreting a large amount of data to develop SPFs for a relatively small jurisdiction like Saskatchewan is not practical, since the acquisition of such data requires considerable resources, time, and effort. For example, using Equation 2-6 requires a detailed measurement for the length of a speed-change lane beginning at the start of the taper of a speed-change lane to its gore point. Furthermore, accessing historic collision records for an adequate number of ramp terminals with similar characteristics (i.e., geometry and traffic control features), which is required when using Equation 2-8, represents another hurdle. In smaller jurisdictions, it may be necessary to develop local SPFs using datasets that are readily available.

Chen (2010) evaluated the safety performance of different freeway exit types (i.e., diverge areas and exit ramp sections) for a Florida highway system. A total of 792 sites were selected, comprising diverge areas, off ramps, and exit ramps. Data from 2004 to 2006 were used for the model development. The functional form developed is shown in Equation 2-10.

$$N_{spf} = L^{b1} \times AADT_m{}^{b2} \times AADT_r{}^{b3} \times e^{(a+b4X1+b5X2+\cdots+bnXn)} \quad \text{[Equation 2-10]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for different freeway exit types;

L = Length of the deceleration lane;

$AADT_m$ = Mainline freeway annual average daily traffic volume (vehicle/day);

$AADT_r$ = Ramp annual average daily traffic volume (vehicle/day);

X1, *X2*…. *Xn* = Geometric and traffic feature variables;

a, *b1*, *b2*,.., *bn* = Regression coefficients.

Moon and Hummer (2009) used three model types to develop safety performance functions for influence areas of ramps on mainline freeway sections: (1) generalized linear model with main effect variables only, (2) generalized linear model with main effect and interaction term variables, and (3) Hauer's method. These methods use linear and nonlinear relationships with dissimilar functional forms for each predictor. The model forms are given by Equations 2-11 to 2-13.

$$N_{spf} = e^{\alpha} \times (AADT_m)^{\beta 1} \times e^{\beta 2(p)} \times e^{\beta 3(type)} \times e^{\beta 4(diff)} \times (ADT_r)^{\beta 5} \quad \text{[Equation 2-11]}$$

$$N_{spf} = e^{\alpha} \times (AADT_m)^{\beta 1} \times e^{\beta 2 (AADT_m) \times (ADT_r)} \times e^{\beta 3 (p)} \times e^{\beta 4 (AADT_m) \times (ADT_r) \times diff} \times e^{\beta 5 (AADT_m) \times (ADT_r) \times type} \quad \text{[Equation 2-12]}$$

$$N_{spf} = \alpha * (AADT_m)^{\beta 1} \times e^{\beta 2 (ADT_r)} \times \beta 1 \times \beta 2 \quad \text{[Equation 2-13]}$$

Where:

N_{spf} = Estimated frequency of collisions per year on ramp influence areas on freeways;

$AADT_m$ = Mainline freeway annual average daily traffic volume (vehicle/day);

ADT_r = Ramp average daily traffic (vehicle/day);

p = Ramp position (right and left hand);

type = Type of ramp (on-ramp or off-ramp); and

diff = Difference in design speeds of freeway and ramps.

In the past, collision prediction models related collisions with the geometric features of a roadway (i.e., lane widths, shoulder widths, vertical and horizontal alignment features, etc.) and traffic characteristics such as speed limits and traffic volumes. These models were referred to as full models. Lu et al. (2012) suggested that having many independent variables in models may result in a possible correlation and, to avoid such a phenomenon, simpler models that only relate collision frequency with traffic volume may be used. These simpler models are gaining increasing acceptance due to their requirement for less data and because they are easier to use.

The results obtained using these simpler models are very much comparable to the results achieved from full models (Lu et al., 2012; Lu et al., 2013a).

Lu et al. (2012) developed simple SPFs in Florida by correlating collisions only with AADT. The simple models, developed for roadway segments and ramps are given by Equation 2-14.

$$N_{spf} = e^{\alpha} \times AADT^{\beta} \quad \text{[Equation 2-14]}$$

Where:

N_{spf}	= Estimated frequency of collisions per year for road segments/ramps;
AADT	= Annual average daily traffic volume (vehicle/day); and
α and β	= Regression coefficients.

Parajuli et al. (2006) developed SPFs for mainline freeway segments within the influence zones interchanges. The SPFs were developed for fatal, injury, and property damage only collisions (Parajuli et al., 2006). The functional form developed for mainline freeway segments and ramps is given by Equation 2-15.

$$N_{spf} = \propto \times (AADT)^{\beta} \times e^{\beta 2(L)} \quad \text{[Equation 2-15]}$$

Where:

N_{spf}	= Estimated frequency of collisions per year for mainline freeway/ramps;
AADT	= Annual average daily traffic volume (vehicle/day);

L = Length of the deceleration lane; and

α, β, and $\beta 2$ = Regression coefficients.

Raghavan Srinivasan and Carter (2011) developed SPFs for rural and urban freeway segments in North Carolina using negative binomial regression. The data used for model development were from 2004 to 2008. SPFs were developed for nine collision types including total collisions, fatal and injury collisions, PDO collisions, lane departure collisions, single- and multiple-vehicle collisions, etc. The functional form for the model is given by Equation 2-16.

$$N_{spf} = L \times e^{\alpha} \times (AADT)^{\beta} \qquad \text{[Equation 2-16]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for mainline freeway/ramps;

L = Length of the roadway segment;

$AADT$ = Annual average daily traffic volume (vehicle/day); and

α and β = Regression coefficients,

Kiattikomol et al. (2008) developed regression models for collision prediction based on data from North Carolina and Tennessee. The research focused on the development of collision prediction regression models using NB regression for interchanges and non-interchange segments of urban freeways. The models were developed for different severity types (i.e., fatal and injury, injury only, and PDO). Data for 276 segments from North Carolina and 381 segments from Tennessee, comprising interchange and non-interchange types, were used. Three-year data for years 2000 to 2002 were used, and different model forms were used. The finalized model is presented as Equation 2-17.

$$N_{spf} = \alpha \times L^{\beta 1} \times (AADT)^{\beta} \qquad \text{[Equation 2-17]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for interchange/non-interchange segments;

L = Length of the roadway segment;

$AADT$ = Annual average daily traffic volume (vehicle/day); and

α, $\beta 1$ and $\beta 2$ = Regression coefficients.

B. N. Persaud (1994), using data of Ministry of Transportation Ontario from the period 1987 to 1988, developed SPFs for rural highway segments, and similar data for 1989 was used for model validation. The functional form of the model is given by Equation 2-18.

$$N_{spf} = \alpha \times L \times (AADT)^{\beta} \qquad \text{[Equation 2-18]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for highway segments;

L = Length of the highway segment;

$AADT$ = Annual average daily traffic volume (vehicle/day); and

α and β = Regression coefficients.

Le and Porter (2012) performed a study on the safety implications of ramp spacing (the space between entering ramps and exiting ramps at interchanges) in California and Washington using data from 2006 to 2008. SPFs were developed for all severity types and for single- and multiple-vehicle collisions. The model form used is given by Equation 2-19.

$$N_{spf} = e^{((\beta \times X) + ln(L))} \quad \text{[Equation 2-19]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for roadway segments;

X = Set of geometric variables characterizing roadway segment;

L = Length of the roadway segment; and

β = Regression coefficients.

Lyon et al. (2011) developed SPFs for five categories of ramp terminals at diamond interchanges in Colorado. They used data for the period from 2000 to 2006 and selected sites that represent statewide geography and other features. SPFs were developed for total collisions and fatal and injury collisions. Equations 2-20 and 2-21 show the model forms developed for signalized ramp terminals and stop-controlled ramp terminals, respectively.

$$N_{spf} = \alpha \times (AADT_r)^{\beta 1} \times (AADT_{xrd})^{\beta 2} \times e^{\beta 3(dummy)} \quad \text{[Equation 2-20]}$$

$$N_{spf} = \alpha \times (AADT_{total})^{\beta 1} \times e^{\beta 2(dummy)} \quad \text{[Equation 2-21]}$$

Where:

N_{spf} = Estimated frequency of collisions per year for ramp terminal;

$AADT_r$ = Sum of approach volumes from ramps and service roads;

$AADT_{xrd}$ = Sum of approach volumes from both ends of the crossroad;

$AADT_{total}$ = Sum of approach volumes from ramps and crossroad;

dummy = Dummy variable that is equal to 0 for non-split approach ramps and equal to 1 for split approach ramps with separate right turn with yield control; and

α, $\beta 1$, $\beta 2$, and $\beta 3$ = Regression coefficients.

The equation 2-10 to 2-13 require detailed information about the geometric and traffic features roadways. Using these equations involves a considerable amount of efforts and time for the collection and interpretation of data. In contrast, Equations 2-14 to 2-19 are simple to use and do not require a large amount of data. Mainly used for roadway segments, only two input variables are utilized in these equations (AADT and length), which are often available, even in small jurisdictions. These equations assume a linear relation with one of the two input variables. On a larger scale, NCHRP Report 17-45 classifies six roadway categories for freeways and interchanges (1) Basic freeway segments, (2) speed-change lanes, (3) off ramp segments, (4) on ramp segments, (5) C-D roads, and (6) ramp terminals.

Roadway classification in Saskatchewan does not follow the classification given in the NCHRP report. Therefore, users may find it difficult to categories interchange segments in the manner required by the report. Also, compiling a collection of data for geometric features could be burdensome since it involves very detailed measurements (e.g., the distance between taper start point and gore point). For high-speed roadways Saskatchewan, it is practical to develop models based on local data that is organized according to the Saskatchewan roadway classification system and that is readily available in databases, in which an adequate number of reference location are made available. This data should be used to generate better collision prediction models.

To estimate an expected number of collisions at a particular site/location, a corresponding predictive method for that roadway type should be available. Individual estimates obtained from models are then combined to find the expected number of collisions for the roadway segment being investigated for safety concerns (i.e., an interchange) or for the entire network. The basic assumption when estimating an expected number of collisions is that, during that specific time of interest (years), roadway features such as geometry have not changed, and the traffic volumes are known. The expected number of collisions obtained from prediction models are summed, if necessary, and combined with the observed number of collisions for the same roadway type to get more reliable estimates of the expected average collision frequency. The Empirical Bayes (EB) method is used to combine the predicted number of collisions with the observed number of collisions to strike a balance between the estimated number of collisions provided by the prediction model and the observed number of collisions (Lu et al., 2012; Lyon et al., 2011; Parajuli et al., 2006; Parisien, 2012; Transportation Research Board (TRB), 2014; Young & Park, 2013). The detailed description of the EB method is provided below under the heading "2.3. Network Screening". Each SPF has a corresponding dispersion parameter "*k*" which provides an indication of the reliability of an SPF. Smaller *k* values are preferred over larger values. This overdispersion parameter is used in the EB method. (Lu et al., 2012; Parajuli et al., 2006; Parisien, 2012; Transportation Research Board (TRB), 2014; Young & Park, 2013).

2.1.2. Development of Local SPFs for Saskatchewan

For the development of SPFs for eight categories of freeways and interchanges in Saskatchewan, including ramp terminals, two distinct data sets were required; one for model development, and another for model validation. Owing to the small number of available data, the entire database was used for model development. Once the appropriate model was identified, the database was divided into two equally sized datasets for the generation of model parameters and for model validation, with the exception of data for weaving sections and signalized ramp terminals, which were too limited. The data used for model development and for model validation were randomly selected.

Researchers use various statistical tools to develop collision prediction models. Among those tools, the most commonly used are STATA, SAS, and R-Language. In this research, R-Language was selected as the tool to estimate model parameters. The generalized linear regression model, also known as the Negative Binomial (NB) model, was used to perform regression analyses. The NB model is widely accepted and commonly used for the development of SPFs, and it accounts for the overdispersion found in the collision data (Abdel-Aty et al., 2014; Abdel-Aty & Radwan, 2000; Chang, 2005; Lu et al., 2013b; B. Srinivasan et al., 2011; Sung, 2000; Young & Park, 2013). Collisions on roadways are random in nature, they fluctuate in number each year at a certain location and have non-negative integer values. It may appear that Poisson regression should be used for modeling collision predictions, but the Poisson model assumes that the variance in data is equal to the mean. In the case of collision data, this assumption does not hold. Strong evidence in support of using the Negative Binomial model, which accounts for the variance in collision data and overdispersion, is available (Abdel-Aty & Radwan, 2000; Hadi et al., 1995; Lu et al., 2013a; Young & Park, 2013).

The NB model is provided in Equation 2-22:

$$N = \frac{\gamma(y+k)}{\gamma(y+1)\gamma k} \times \frac{\left(\frac{1}{k}\mu\right)^{y}}{\left(1+\frac{1}{k}\mu\right)^{(y+k)}} \qquad \text{[Equation 2-22]}$$

Where:

k = dispersion parameter

μ = mean;

γ = gamma function; and

y = observed value.

Different functional forms were considered as potential SPFs for different configurations of high-speed roadways in Saskatchewan. A detailed discussion of roadway configuration is discussed in “3.2.1 Segmentation Scheme.” Parameters were estimated for each potential functional form that was considered for SPF development for each roadway configuration. Those potential functional forms are given in Table 8 to Tbale 12:

Table 8: Candidate Functional Forms for Basic Freeways Inside and Outside Interchange Systems

Potential Model Forms	
$\boldsymbol{N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 2-23]
$\boldsymbol{N = \alpha \times L^{\beta 1} \times \left(\frac{AADT}{1000}\right)^{\beta 2}}$	[Equation 2-24]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 2-25]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 2-26]

Where:

N = Estimated frequency of collisions per year for freeway segment (inside or outside interchange system);

L = Length of freeway segment (kilometer);

AADT = Annual average daily traffic volume of freeway (vehicle/day); and

α, β1, and *β2* = Regression coefficients.

Table 9: Potential Functional Forms for On-Ramps and Off-Ramps

Potential Model Forms	
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 2-27]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 2-28]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times L^{\beta 2}}$	[Equation 2-29]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times L}$	[Equation 2-30]

Where:

N = Estimated frequency of collisions per year for ramp segment (on ramp or off ramp);

L = Length of ramp segment (kilometer);

AADT = Annual average daily traffic volume of ramp (vehicle/day); and

α, β1, and *β2* = Regression coefficients.

Table 10: Potential Functional Forms for Ramp Influence Areas

Potential Model Forms	
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{AADT}{1000}\right)} \times exp^{\beta 3\left(\frac{\Sigma\, RAADT}{1000}\right)} \times S$	[Equation 2-31]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L \times S$	[Equation 2-32]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L$	[Equation 2-33]
$N = \alpha \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta} \times L$	[Equation 2-34]
$N = \alpha \times\ exp^{\left[\left(\beta 1\frac{AADT}{1000}\right) + (\beta 2 * S)\right]}$	[Equation 2-35]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$	[Equation 2-36]
$N = \alpha \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta}$	[Equation 2-37]
$N = \alpha \times\ exp^{\beta\left(\frac{\Sigma AAADT}{1000}\right)}$	[Equation 2-38]

Where:

N = Estimated frequency of collisions per year for ramp influence area;

L = Length of the ramp influence area (kilometer);

AADT = Annual average daily traffic volume of freeway segment (vehicle/day);

ΣRAADT = Annual average daily traffic volume sum of ramp AADTs (vehicle/day);

ΣAAADT = Annual average daily traffic volume sum of freeway and ramp AADTs (vehicle/day)

S = Posted speed of the mainline freeway segment (km/h); and

α, β1, β2 and *β3* = Regression coefficients.

Table 11: Potential Functional Forms for Weaving Sections

Potential Model Forms	
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{\Sigma\, RAADT}{1000}\right)} \times exp^{\beta 3 * S} \times \Sigma\, RAADT^{\beta 4}}$	[Equation 2-39]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{AADT}{1000}\right)} \times exp^{\beta 3\left(\frac{\Sigma\, RAADT}{1000}\right)} \times L \times S}$	[Equation 2-40]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{\Sigma\, RAADT}{1000}\right)} \times S}$	[Equation 2-41]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L}$	[Equation 2-42]
$\boldsymbol{N = \alpha \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta} \times L}$	[Equation 2-43]
$\boldsymbol{N = \alpha \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * s}}$	[Equation 2-44]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 2-45]
$\boldsymbol{N = \alpha \times \left(\frac{\Sigma AADT}{1000}\right)^{\beta}}$	[Equation 2-46]
$\boldsymbol{N = \alpha \times exp^{\left[\left(\beta 1 \frac{AADT}{1000}\right) + (\beta 2 * L)\right]}}$	[Equation 2-47]
$\boldsymbol{N = \alpha \times exp^{\left[\left(\beta 1 \frac{\Sigma AAADT}{1000}\right) + (\beta 2 * s)\right]}}$	[Equation 2-48]

Where:

N	= Estimated frequency of collisions per year for ramp influence area;
L	= Length of the weaving section (kilometer);
AADT	= Annual average daily traffic volume of freeway segment (vehicle/day);
ΣRAADT	= Annual average daily traffic volume sum of ramp AADTs (vehicle/day);
ΣAAADT	= Annual average daily traffic volume sum of freeway and ramp AADTs (vehicle/day);
S	= Posted speed of the mainline freeway segment (km/h); and
α, β1, β2, β4 and *β4*	= Regression coefficients.

Table 12: Potential Functional Forms for Ramp Terminals (Signalized and Unsignalized)

Potential Model Forms	
$$N = \alpha \times \left(\frac{\Sigma\ RAADT}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)^{\beta 2}$$	[Equation 2-49]
$$N = \alpha \times \left(\frac{AADT_{total}}{1000}\right)^{\beta}$$	[Equation 2-50]
$$N = \alpha \times \left(\frac{\Sigma\ RAADT}{1000}\right)^{\beta 1} \times\times \left(\frac{AADT_{total}}{1000}\right)^{\beta 2}$$	[Equation 2-51]
$$N = \alpha \times \left(\frac{AADT_{total}}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)$$	[Equation 2-52]

Where:

N = Estimated frequency of collisions per year for ramp terminal (signalized or unsignalized);

$\Sigma RAADT$ = Annual average daily traffic volume sum of ramp AADTs (vehicle/day);

$AADT_{xrd}$ = Annual average daily traffic volume of crossroad, average of both legs of crossroad (outside and inside interchange, in vehicle/day);

$AADT_{total}$ = Annual average daily traffic volume (sum of crossroad and entering and exiting ramp, in vehicle/day); and

α, $\beta 1$, and $\beta 2$ = Regression coefficients.

2.1.3. Selection of Best-Fitting Model

In the selection of potential prediction models, the linear, as well as nonlinear relation of length with predicted number collisions, was explored. For ramp terminals, AADT, representing traffic volumes for both major and minor roadways, was looked upon.

A number of methods were used in the selection of the most appropriate model forms for Saskatchewan high-speed roadways. These methods included calculation of the statistical significance (p-value) of predictors (Cafiso et al., 2010; Meng & Qu, 2012). Lower p-values represent a greater statistical significance, and acceptable p-values for collision prediction models have been taken up to 0.15 (Harwood et al., 2000). Akaike's Information Criterion (AIC) (Akaike, 1998; Cafiso et al., 2010; Meng & Qu, 2012), which provides a relative measure of goodness of fit (GOF) and lower AIC values indicate a better-fitting model. However, the AIC does not evaluate the performance of models. The Bayesian Information Criterion (BIC) (Haque et al., 2010; Young & Park, 2013) is similar to the AIC but also considers the number of data points in the model. Lower BIC values are preferred over higher values (Burnham & Anderson, 2004). Cumulative residual plots (CURE Plots) (Cafiso et al., 2013; Hauer, 2015; Manan et al., 2013) are horizontally plotted graphical representations of models that allow observed values and model predictions (i.e. cumulative residuals) to be compared. The CURE plots are a better means for measuring the performance of model than R^2 since the R^2 provides measure of overall performance of a model without providing details about the performance of model for different values of the input variable, whereas the CURE Plot shows how well the model is performing against each value of an input variable (Hauer, 2015). Good models have cumulative residuals oscillating around zero line (x-axis) and within ± two standard deviations, (Hauer, 2015; Hauer & Bamfo, 1997).

The AIC is calculated using following equation:

$$AIC = 2 \times (count\ of\ model\ parameters) - 2 \times \log(maximum\ likelyhood) \qquad \text{[Equation 2-53]}$$

The BIC is calculated using the following expression:

$$BIC = count\ of\ model\ parameters \times log(count\ of\ dataset\ observations) - 2 \times\ log(maximum\ likelyhood) \qquad \text{[Equation 2-54]}$$

A CURE plot representing a good model is shown in Figure 3. The green and red dotted lines represent deviations equal to two standard deviations in either direction, and the solid black line shows the cumulative residuals of the model. A CURE plot representing a poor model is similarly shown in Figure 4.

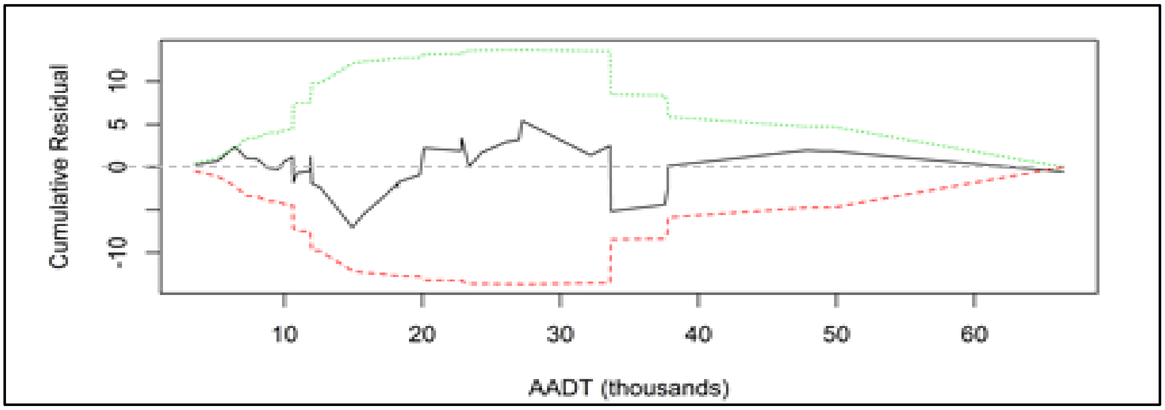


Figure 3: Cumulative Residual Plot for a Good Collision Prediction Model

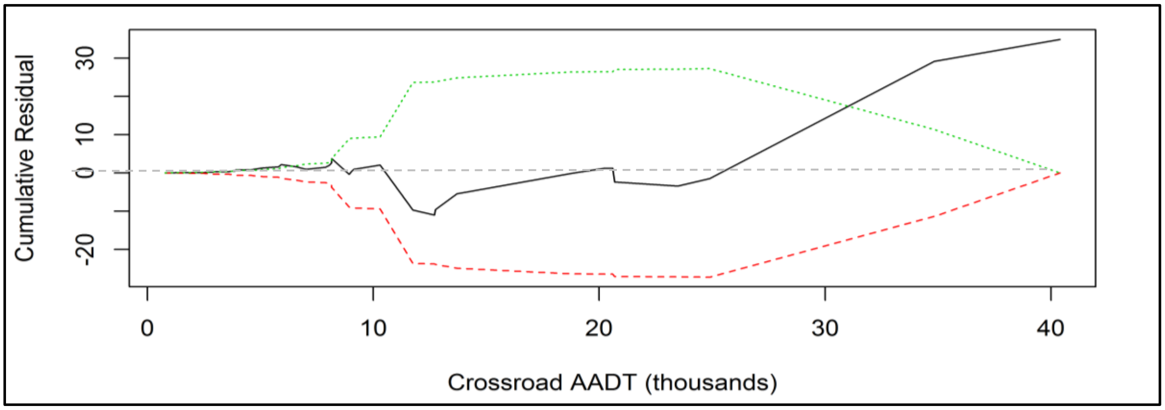


Figure 4: Cumulative Residual Plot for a Poor Collisions Prediction Model

2.1.4. Validation of Models

Literature review suggest using multiple statistical tests for validation of models (Cheng & Washington, 2005; Hadayeghi et al., 2006). The following validation methods were applied to the developed models for high-speed roadways in Saskatchewan.

The Mean Square Error (MSE), which measures the error associated with the models and is calculated by dividing the sum of the squared differences between the observed and a predicted number of collisions by the sample size, was determined for each of the developed models. Smaller MSE values reflect a model's better fit to the observed collisions.

Equation 2-55 shows the expression used to calculate MSE.

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-p)}$$ [Equation 2-55]

Where:

y = number of observed collisions;

$\hat{y}$ = number of predicted collisions;

n = sample size; and

p = number of parameters in the model.

The Mean Squared Prediction Error (MSPE) was also determined for each model using the validation dataset, and the results were compared with the MSE values. MSPE values that are greater than corresponding MSE values indicate over-fitting of the model and lower values depict that the model is under-fitting. However, similar values are the desired target. MSPE values closer or less than MSE values demonstrate that the model could be considered as transferable. MSPE equation is similar to the MSE with only one difference; that is the number of parameters are removed from MSE equation. Equation 2-56 shows the formula through which MSPE is calculated:

$$MSPE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n)}$$ [Equation 2-56]

Another measure to check the validity of models involves calculating the Mean Absolute Deviation (MAD). MAD shows the average magnitude of the variability in the prediction models. MAD values are always positive, and smaller values indicate less variability, and therefore a better fit. Calculating MAD values serves as another GOF test, which was used for each model using the validation dataset. The following equation is used to calculate MAD:

$$MAD = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{(n)} \qquad \text{[Equation 2-57]}$$

The Mean Prediction Bias (MPB) was also calculated for each model to check the magnitude and direction of the models' average biases. This calculation was again performed using the validation dataset. Smaller MBP values reflect a better prediction capability of a model. The following equation is used to calculate MBP:

$$MPB = \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)}{(n)} \qquad \text{[Equation 2-58]}$$

MAD and MPB are measures used to understand how accurately a model can predict the number of collisions.

The Freeman-Tukey R-Square (R^2_{FT}) test was suggested by Fridstrøm et al. (1995) to check the GOF. This test is similar to the R^2 test, and values closer to one are desired. The similar values of R^2_{FT} from development dataset and R^2_{FT} from validation datasets are an indication of model's better fit. The test is applied to both the dataset used to determine the parameters to be used in each model and the dataset used for validation. The following equations are used for this test:

$$R^2_{FT} = \frac{\sum_{i=1}^{n} (f_i - \bar{f})^2 - \sum_{i=1}^{n} \hat{e}_i^2}{\sum_{i=1}^{n} (f_i - \bar{f})^2} \qquad \text{[Equation 2-59]}$$

And:

$$f_i = \sqrt{y_i} + \sqrt{y_i + 1} \qquad \text{[Equation 2-59a]}$$

$$\bar{f} = \frac{\sum_{i=1}^{n}(f_i)}{n} \qquad \text{[Equation 2-59b]}$$

$$\bar{f} = \frac{\sum_{i=1}^{n}(f_i)}{n} \qquad \text{[Equation 2-59c]}$$

$$\hat{e} = f_i - \sqrt{4 \times \hat{y}_i}+1 \qquad \text{[Equation 2-59d]}$$

Where:

y_i = Observed number of collision at site *i*;

$\hat{y}_i$ = Predicted number of collisions at site *I*;

n = Number of similar locations in the dataset; and

p = Degrees of freedom.

2.2. Network Screening

The Highway Safety Manual (HSM) outlines 13 methods to perform network screening and identify sites that require safety improvements (AASHTO, 2010). These methods include determining the collision frequency, collision rate, equivalent property damage only (EPDO), excess equivalent property damage only (Excess EPDO), critical collision rate, etc. Each of the methods provided in the HSM has its strengths and weaknesses, and it is difficult to ascertain which one of is better than the others (HSM, 2010; Sun & Manthena, 2009). It is therefore desirable that more than one network screening method be used for the same site, and that sites identified as candidates for safety improvement interventions by each method be compared (Parisien, 2012). In this study, two methods are used for network screening purpose: (1)

Equivalent property damage only (EPDO) with EB adjustment and (2) Excess equivalent property damage only (Excess EPDO) with EB adjustment.

The methods used for the research are used in conjunction with the EB method to improve the estimates generated by SPFs. Using the EB method requires placing weight on predicted and observed numbers of collisions in order to reduce the RTM effect (Lyon et al., 2011; Parisien, 2012). The application of this procedure can only be made to the sites that have historic collision records and available SPFs, which suggest that it cannot be used for new design proposals (HSM, 2010). The EB weight factor is calculated for each location and the weight factor is dependent on the overdispersion parameter. Equation 2-60 is used to relate observed and predicted collisions.

$$W_y = \frac{1}{1+K \times (\sum_{study\ period} N_{predicted})} \qquad \text{[Equation 2-60]}$$

Where:

W_y = Empirical Bayes weight for severity, y;

K = Overdispersion parameter for the appropriate SPF; and

$N_{predicted,y}$ = Predicted average collision frequency for severity type y.

2.2.1. EPDO Average Collision Frequency with EB Adjustment

In this method, injury and fatal collisions are converted into equivalent property damage only (EPDO) collisions using the weight factor (separate from the weight factor described in above). The weight factors are calculated using the total societal cost of a PDO collision. This step accounts for the severity of the collision based on the societal collision cost for a particular collision severity relative to a PDO collision.

SPFs for targeted site are used to calculate predicted average collision frequencies for each severity type (fatal and injury collisions and PDO collision). The frequency of total collisions is calculated by adding the frequencies of collisions for each severity type. The SPFs are then calibrated to the local conditions using the most recent data for collision frequencies, traffic volumes, segment lengths, etc.

Step 1: Calculate the Predicted Average Collision Frequency using an Appropriate SPF

Step 1.1: Apply Developed SPF appropriate to the site

$$N_y = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \qquad \text{[Equation 2-61]}$$

Where:-

N_y = Uncalibrated predicted number of collisions for severity y;

$AADT$ = Annual average daily traffic volume; and

α and β = Regression coefficients.

Step 1.2: Calibrate SPFs

$$C_r = \frac{\sum_{all\ sites} observed\ collisions}{\sum_{all\ sites} predicted\ collisions} \qquad \text{[Equation 2-62]}$$

$$N_{predicted,y} = C_r \times N_y \qquad \text{[Equation 2-63]}$$

Where:

C_r = Calibration factor;

$N_{predicted,y}$ = Predicted number of collisions for severity y; and

N_y = Uncalibrated predicted number of collisions for severity y.

<u>Step 1.3: Apply Calibrated SPFs for Severity Type</u>

$$N_{(total)} = C_{r\,(total)} \times Ny_{(total)} \qquad \text{[Equation 2-64]}$$

$$N'_{(FI)} = C_{r\,(FI)} \times Ny_{(FI)} \qquad \text{[Equation 2-65]}$$

$$N'_{(PDO)} = C_{r\,(PDO)} \times Ny_{PDO} \qquad \text{[Equation 2-66]}$$

$$N_{(FI)} = N_{(total)} \times \left(\frac{N'_{(FI)}}{N'_{(FI)} + N'_{(PDO)}}\right) \qquad \text{[Equation 2-67]}$$

$$N_{(PDO)} = N_{(total)} - N_{(FI)} \qquad \text{[Equation 2-68]}$$

Where:

$N_{(total)}$ = Predicted total collisions;

$N'_{(FI)}$ = Calibrated fatal and injury collisions;

$N'_{(PDO)}$ = Calibrated PDO collisions;

$N_{(FI)}$ = Predicted fatal and injury collisions;

$N_{(PDO)}$ = Predicted PDO collisions; and

α *and* β = Regression coefficients.

Step 2: Calculate Annual Correction Factor *n*

Annual correction factors are applied to the SPFs to account for annual changes in traffic volumes.

$$C_{n\,(total)} = \frac{N_{predicted,n(total)}}{N_{predicted,1(total)}} \quad \text{and} \quad C_{n\,(fi)} = \frac{N_{predicted,n(FI)}}{N_{predicted,1(FI)}} \qquad \text{[Equation 2-69]}$$

$$f_{y(weight)} = \frac{CC_y}{CC_{pdo}} \qquad \text{[Equation 2-70]}$$

Where:

$C_{n(total)}$ = Annual correction factor for total collisions;

$C_{n(FI)}$ = Annual correction factor for fatal and injury collisions;

$N_{predicted,n(total)}$ = Predicted number of total collisions for year n;

$N_{predicted,1(total)}$ = Predicted number of total collisions for year 1;

$N_{predicted,n(FI)}$ = Predicted number of fatal and injury collisions for year n; and

$N_{predicted,1(FI)}$ = Predicted number of fatal and injury collisions for year 1.

Step 3: Calculate Weighted Adjustment

Calculate the EB weight factor for each location, where weight factor is dependent on overdispersion parameter, study period and predicted number of collisions and a decrease in any of these will result in decreased weight factor.

$$W_y = \frac{1}{1+K \times (\sum_{study\ period} N_{predicted})} \quad \text{[Equation 2-71]}$$

Where:

W_y = Empirical Bayes weight for severity y;

K = Overdispersion parameter from the appropriate SPF; and

$N_{predicted,y}$ = Predicted average collision frequency for severity type y.

Step 4: Calculate First Year EB-Adjusted Expected Average Collision Frequency

The EB-adjusted expected average collision frequency is calculated using the weight factor based on both the predicted and observed collisions, as described above. More emphesis will be placed on SPFs predicted collisions it the weight fatocr increase and vise versa.

$$N_{expected,n(total)} = W_{total} \times N_{predicted,n(total)} + (1 - W_{total}) \times \left(\frac{\sum_{n=1}^{j} N_{observed,n(total)}}{\sum_{n=1}^{j} C_{n(total)}} \right)$$

[Equation 2-72]

$$N_{expected,n(FI)} = W_{FI} \times N_{predicted,n(FI)} + (1 - W_{FI}) \times \left(\frac{\sum_{n=1}^{j} N_{observed,n(FI)}}{\sum_{n=1}^{j} C_{n(FI)}} \right)$$

[Equation 2-73]

Where:

$N_{expected,n,(total)}$ = EB-adjusted expected total average collision frequency for year n;

$N_{predicted,n(total)}$ = Calibrated predicted total average collision frequency from SPF;

$N_{observed,n(total)}$ = Observed number of total collisions for year n;

$w_{(total)}$ = EB-Weight factor for total collisions;

$C_{n(total)}$ = Annual correction factor for total collisions;

$N_{expected,n,(FI)}$ = EB-adjusted expected average FI collision frequency for year n;

$N_{predicted,n(FI)}$ = Calibrated predicted average FI collision frequency from SPF;

$N_{observed,n(FI)}$ = Observed number of FI collisions for year n;

$w_{(FI)}$ = EB-Weight factor for FI collisions;

$C_{n(FI)}$ = Annual correction factor for FI collisions; and

j = Number of years in the study.

Step 5: Calculate Five Year EB-Adjusted Average Collision Frequency

The ranking of location is based on the most recent year in the study period. The final year's expected collision frequency is calculated by multiplying the SPF-predicted collision frequency by the annual correction factor for the final year of the study period.

$$N_{expected,n(total)} = N_{expected,1(total)} \times C_{n(total)}$$ [Equation 2-74]

$$N_{expected,n(FI)} = N_{expected,1(FI)} \times C_{n(FI)}$$ [Equation 2-75]

$$N_{expected,n(PDO)} = N_{expected,n(total)} - N_{expected,n(FI)}$$ [Equation 2-76]

Where:

$N_{expected,n(total)}$ = EB-adjusted expected average total collision frequency for final year n;

$N_{expected,1(total)}$ = EB-adjusted expected average total collision frequency for year 1;

$N_{expected,n(FI)}$ = EB-adjusted expected average FI collision frequency for final year n;

$N_{expected,1(FI)}$ = EB-adjusted expected average FI collision frequency for year 1;

$N_{expected,n(PDO)}$ = EB-adjusted expected average PDO collision frequency for final year n; and

C_n = Annual correction factor for year n.

Step 6: Calculate Separate Weight Factors for Fatal Collisions and Injury Collisions

This step accounts for the severity of collisions based on the total societal costs, converted to EPDO, associated with each type of severity. This is performed separately for fatal collisions and injury collisions, resulting in an EPDO-weight for each type of collision.

$$f_{y(weight)} = \frac{CC_y}{CC_{pdo}} \qquad \text{[Equation 2-77]}$$

Where:

$f_{y(weight)}$ = EPDO weight factor based on collision severity y;

CCy = Cost of collision severity y; and

$CCPDO$ = Cost of PDO collision severity.

<u>Step 7: Calculate Proportion of Fatal and Injury Collisions</u>

Since fatal and injury collisions are considered in combination in SPFs, the proportion of fatal collisions and injury collisions is calculated using observed fatal and injury collisions.

$$P_F = \frac{\sum N_{observed,(F)}}{\sum N_{observed,(FI)}} \qquad \text{[Equation 2-78]}$$

$$P_I = \frac{\sum N_{observed,(I)}}{\sum N_{observed,(FI)}} \qquad \text{[Equation 2-79]}$$

Where:

P_F = Proportion of observed number of fatal collisions;

P_I = Proportion of observed number of injury collisions;

$N_{observed,(F)}$ = Observed number of fatal collisions;

$N_{observed\ ,(I)}$ = Observed number of injury collisions; and

$N_{observed,(FI)}$ = Observed number of fatal and injury collisions.

Step 8: Calculate Combined Weight of Fatal and Injury Collisions

The EPDO weight factor for fatal and injury collisions is obtained by summing the product of the proportion of fatal and injury collisions and their respective EPDO collision costs.

$$W_{EPDO,FI} = P_F \times f_{fatal(weight)} + P_I \times f_{injury(weight)} \quad \text{[Equation 2-80]}$$

Where:

$W_{EPDO,FI}$ = Combined EPDO weight factor for fatal and injury collisions;

$f_{injury(weight)}$ = EPDO weight factor injury collisions;

$f_{fatal(weight)}$ = EPDO weight factor for fatal collisions;

P_F = Proportion of observed number of fatal collisions; and,

P_I = Proportion of observed number of injury collisions.

Step 9: Calculate Final Year EPDO Expected Average Collision Frequency

The final year EPDO expected average collision frequency is calculated by summing the expected PDO collision frequency with the EPDO-weighted expected fatal and injury collisions.

$$N_{expected(EPDO)} = N_{expected,n,(PDO)} + W_{EPDO,FI} \times N_{expected,n(FI)} \qquad \text{[Equation 2-81]}$$

Where:

$N_{expected,n(EPDO)}$ = EPDO expected average collision frequency for year n;

$N_{expected,n(PDO)}$ = EB-adjusted expected average PDO collision frequency for year n;

$w_{EPDO,FI}$ = EPDO weight factor for fatal and injury collisions; and

$N_{expected,n(FI)}$ = EB-adjusted expected average fatal and injury collision frequency for year n.

<u>Step 10: Rank Sites Based on EB-Adjusted EPDO</u>

Based on the EPDO, sites are ranked from highest to lowest to identify locations that can benefit most from safety improvements.

2.2.2. Excess Expected Average Collision Frequency with EB Adjustment

The difference between the expected average collisions frequency and the EB-adjusted expected average collision frequency is referred to as the excess expected average collision frequency. The excess expected average collision frequency of the EB Adjustment is another tool used to rank roadway network locations described in the HSM.

The excess expected number of collisions highlights locations that exhibit a number of collisions that exceeds the number of collisions predicted by a developed model at a location with similar characteristics (AASHTO, 2010). This procedure starts when the predicted EB-adjusted collision frequencies are calculate using above described steps (i.e., follow steps 1-8 listed above). Figure 5 shows the logic behind the expected and excess collision frequencies. A positive difference between the EB-adjusted excess collision frequency and SPF-predicted collision frequency indicates that the site is likely to experience a greater number of collisions compared to other sites with similar characteristics (i.e. geometric features and traffic volumes).

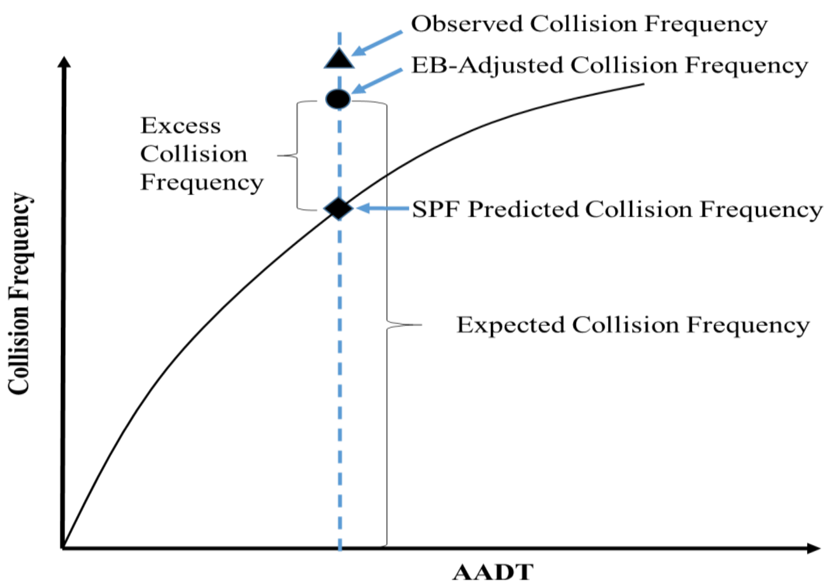


Figure 5: Expected and Excess Collision Frequencies

Excess collision frequency is calculated when the EB adjusted collision frequency is already calculated (i.e. step 1 to 8). The procedure for calculating the excess expected collision frequency is outlined below:

<u>Step 9: Calculate Excess Expected Average Collision Frequency</u>

$$Excess_y = \left(N_{expected,n(PDO)} - N_{prdcited,n(PDO)}\right) + \left(N_{expected,n(FI)} - N_{predicted,n(FI)}\right) \quad \text{[Equation 2-82]}$$

Where:

$Excess_n$ = Excess expected collisions for year n;

$N_{expected,n}$ = EB-adjusted expected average collision frequency for year n; and

$N_{predicted,n}$ = SPF-predicted average collision frequency for year n.

Step 10: Calculate Excess EPDO

Excess collisions are converted to EPDO to account for the severity of the collisions. The following equation is used to convert excess collisions into EPDO by applying a weight factor.

$$Excess_y = \left(N_{expected,n(PDO)} - N_{prdcited,n(PDO)}\right) + \left(N_{expected,n(FI)} - N_{predicted,n(FI)}\right) \times W_{EPDO.FI}$$

[Equation 2-83]

Where:

$Excess_y$ = Excess expected collisions for year n;

$N_{expected,n}$ = EB-adjusted expected average collision frequency for year n;

$N_{predicted,n}$ = SPF-predicted average collision frequency for year n; and

$w_{EPDO,FI}$ = EPDO weight factor for fatal and injury collisions.

Step 11: Rank Sites Based on EPDO including Excess EPDO

Based on excess EPDO, sites are ranked from highest to lowest to identify location in the roadway network that can benefit from safety improvements.

The sample calculations for both methods (i.e. excepted average collision frequency with EB adjustment and excess expected collision frequency) is presented in Appendix B.

2.3. Chapter Summary

A literature review was performed to examine which functional forms are commonly used for the development of local SPFs, in particular for freeways and interchanges. Special attention was paid to the roadway classifications systems used in the HSM and in other literature. Statistical models used for development of SPFs were examined, including the negative binomial (NB) model, also known as the Poisson-gamma model due to the overdispersion found in the collision datasets. Different commonly used GOF tests for model validation were also described.

The HSM's network screening methods were described out of which two network screening methods 1) EPDO average collision frequency with EB adjustment, and 2) excess expected average collision frequency were selected and their calculation procedure was detailed.

CHAPTER 3: DATABASE INTEGRATION

This chapter begins by describing the three following sets of data that are used in this study for the development of local SPFs for high-speed roadways in Saskatchewan: spatial dataset, traffic volume dataset, and collision dataset. The second part of this chapter discusses the roadway classification system used in this study for development of corresponding SPFs. Finally, the last part of this chapter discusses the how datasets in different formats from different sources were integrated into a ready-to-use database for development of SPFs in this study.

3.1. Description of Datasets

One of the objectives of this research was to transform the different sets of information that are required for the development of SPFs into an integrated database. For this purpose, the required data were collected and the following datasets were adapted for the compilation of an integrated dataset.

3.1.1. Spatial Datasets

The spatial datasets were received from the following three sources:

- City of Regina;
- City of Saskatoon; and
- Ministry of Highways and Infrastructure (SMHI).

These datasets contain spatial information about roadway networks in their respective jurisdictions. For example, the dataset from the City of Saskatoon only contains information about roadways within their administrative limits, as is the case for the spatial dataset from

Regina. The Saskatchewan Ministry of Highway and Infrastructure dataset contains information about provincial highways (City of Regina, 2013; City of Saskatoon, 2014; Saskatchewan Ministry of Highways and Infrastruture, 2014). Since this research is intended to cover high-speed roadways in the entire province of Saskatchewan, a separate spatial dataset containing spatial information for all roadways in the province was needed to serve as a basic framework for the integrated dataset describe above. For this purpose, another spatial dataset (SRN11) was provided by the SMHI and was used to develop the basic framework for the integrated dataset (Saskatchewan Ministry of Highways & Infrastructure, 20090101). The spatial datasets from the three sources above are discussed separately below.

3.1.2. Regina Spatial Dataset

The City of Regina provided a spatial dataset in shapefile format (City of Regina, 2013), which is the most widely used format in North America. This dataset format assigns a single line to a roadway segment that considers both directions combined as well as single point for intersections. This dataset contains very detailed information associated with each roadway. For example, each roadway had a location identifier written as a number under heading “KEYNUMBER”, a roadway classification under heading “ROAD_FUNC”, the name of the street under “STRT_NAME”, etc.

Table 13 shows a sample of the information provided in the spatial dataset from the City of Regina.

Table 13: Sample Information in Shapefile for Roadways in the City of Regina

Field	Description
KEYNUMBER	Unique Identifier
ROAD_ID	Unique Identifier
ROAD_FUNC	Road Classification (HWY, Expressway)
STRT_NAME	Posted Street Name

3.1.3. Saskatoon Spatial Dataset

The City of Saskatoon's GIS has a Transportation Data Model (TDM) in a shapefile format which is one of the most popular geospatial data formats in the region (i.e. shapefiles). Eighty percent of Saskatoon's municipal infrastructure can be georeferenced (i.e., roadways, water and sewer systems, etc.). TDMs are used to display various roadway information, such as roadway classification, traffic control features, structural and geometric roadway details, etc. (Saskatoon, 2005). The City of Saskatoon provided spatial datasets in the shapefile format for the roadways under their jurisdiction, which also contained very detailed information. This information includes location identification under the heading "UGRID", which represents a single point on the road segment or ramp terminal. In addition, the dataset also contains information about roadways classification (CETI_TYPE), traffic volumes, and posted street name (ONLINE_STR). Table 14 shows a sample of the information available in the spatial dataset for the City of Saskatoon.

Table 14: Sample Information in Shapefile for Roadways in the City of Saskatoon

Field	Description
ROAD_ID	Unique Identifier
ROAD_LENGT	Segment Length
ONLINE_STR	Posted Street Name
UGRID	Unique Identifier
CETI_TYPE	Roadway Classification

3.1.4. Saskatchewan Ministry of Highway and Infrastructure Spatial Datasets

The Saskatchewan Ministry of Highway and Infrastructure provided two spatial datasets, one for the roadways under their jurisdiction (Saskatchewan Ministry of Highways and Infrastruture, 2014) and another containing information for all roadways in the province, including roadways in cities (SRN11) (Saskatchewan Ministry of Highways & Infrastructure, 20090101).

Unlike the dataset from the City of Regine, the first SMHI dataset assigns a single line to each travel direction on a roadway. This dataset also contains very detailed information, including location identifiers under “ROADNAME” in a specific numerical format. The first three numbers in the location identifier represent the provincial highway number. The following two numbers represent the section of the highway. The two numbers after that represent the road type and the three letters at the end represent the roadway classification (divided or undivided, and travel direction). Figure 6 explains the naming system used by the SMHI for provincial highways.

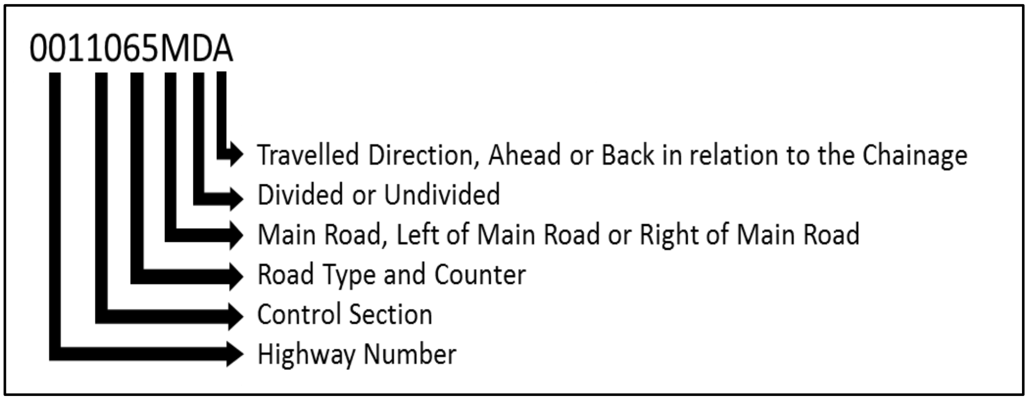


Figure 6: SMHI Control Section Naming System

Other information given in the SMHI datasets includes posted speed limit, community name, and route number. Table 15 provides a sample of the information available in both spatial datasets received from the SMHI (SMHI shapefile and SRN11 spatial dataset).

Table 15: Sample Information in Shapefiles for Roadways in Saskatchewan Provided by the SMHI

1. SMHI Spatial Dataset

Field	Description
ROADNAME	Unique Identifier
PLACENAME	Community Name

2. SRN11 Spatial Dataset

Field	Description
POSTEDNAME	Posted Road Name
SPEEDLIMIT	Speed Limit
PLACENAME	Community Name
ROADCLASS	Road Classification
NBRLANES	Number of Lanes

3.1.5. Traffic Volume Datasets

Evaluating the safety performance of planned or existing roadways using collision prediction models essentially requires knowing the exposure (AADT) to estimate the probability of collisions (X Qin et al., 2004; Sayed & De Leur, 2008). For prediction models, exposure should ideally be expressed as yearly traffic volumes.

The traffic volume datasets were received from:

- City of Regina;
- City of Saskatoon; and
- Saskatchewan Ministry of Highway and Infrastructure.

All three datasets provided by the above jurisdictions were in different formats. The description of these datasets is given below.

3.1.6. Regina Traffic Volume Dataset

The City of Regina provided traffic volume information for the years 2007-2009 in more than one format that is pdf format and for the years 2010-2011 in Microsoft (MS) Excel spreadsheet (or txt) format (City of Regina, 2013). In the City of Regina dataset, traffic volume information for a road segment is combined for both travel directions. The traffic flow maps display AADT for freeways and interchanges in detail for ramps, crossroads, and weaving sections. However, at few locations, such as the interchange at Ring Road and Ross Avenue and the interchange at Ring Road and McDonald Street, the traffic volumes of the minor legs (i.e., ramps) of interchanges are missing. Figure 7 shows sample traffic volume information for interchange between Highway 1 and Highway 6 on the outskirts of Regina.

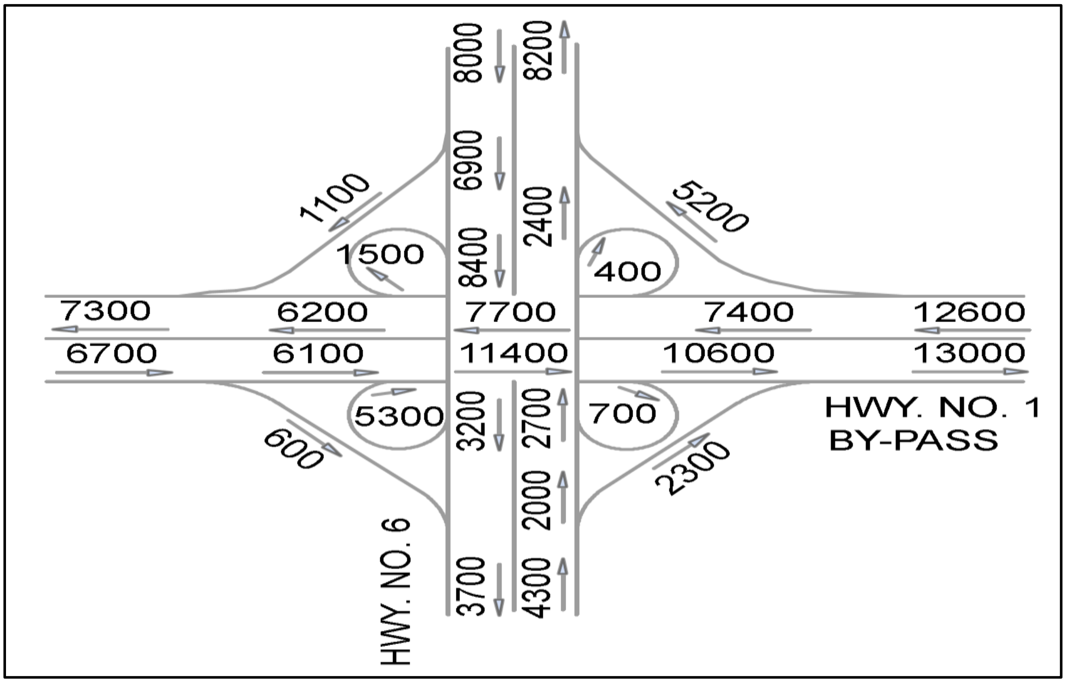


Figure 7: Sample of Traffic Volume Information from Regina

3.1.7. Saskatoon Traffic Volume Dataset

The City of Saskatoon (COS) provided traffic volume information combined for both travel direction in MS Excel spreadsheet format for 298 locations/roadways, with supporting station index maps in pdf format (City of Saskatoon, 2014). The spreadsheets provide a numerical identifier for each of the count stations under the heading "Stano". The dataset provided by the COS also has some missing AADT information. Table 16 provides a sample of the AADT data received from the COS.

Table 16: Sample of Traffic Volume Information from Saskatoon.

Stano	Location	AADT				
		2007	2008	2009	2010	2011
338	Taylor Street: Circle Drive - Acadia Drive			26400	30450	28750
365	Circle Drive: 22nd Street - Laurier Drive		24300			
376	Circle Drive: Attridge Drive - Preston Avenue	66900	66650			
377	Attridge Drive: Circle Drive - Central Avenue	37350				43250

3.1.8. SMHI Traffic Volume Dataset

The Saskatchewan Ministry of Highway and Infrastructure provided traffic volume information for 139 locations for the period from 2007-2012. The information was provided in spreadsheets with the location information under the heading "RD_ELEMENT". There were some locations with missing traffic volumes for minor legs of interchanges. The AADT information was supported appended pdf maps referred to under "APPENDIX LOCATION" for each roadway segment. Table 17 shows a sample of the AADT information provided by the SMHI.

Table 17: Sample of Traffic Volume Information from SMHI.

RD_ELEME	FR_KM	TO_KM	AT_KM	APPENDIX LOCATION	AADT12	AADT11	AADT10	AADT09	AADT08	AADT07	Location Descriptions
0011161LUB	0.00	0.14	1.02	C9	50	50	70	70	70	70	On Ramp from 0011160LUB to Highway No. 301 N.B.
0022060LUB	0.00	0.16	0.15	D9	70	70	50	50	50	50	Left Slip Off at Km 0.28 of S.B. Off Ramp to Highway No. 3 E.B.
0011160RUA	0.00	0.39	0.25	C8	70	70	60	60	60	60	W.B. Off Ramp at Knowles Overpass
0011860RUA	0.00	0.25			230	200	200	200	200	200	Off Ramp from W.B. to North Service/Frontage Road at Km 0.35

3.1.9. Collision Dataset

A raw collision dataset "HIGHSPEED_AC" was received from SGI in MS Excel format for the period from 2007-2012: The raw dataset contained collision records for a total of 10,187 collisions, of which 8,260 collisions occurred within the study period (2007-2011) both within and outside the study area. Each row in the raw dataset represents a collision with a unique case number written under the heading "CASENO". The location identifier for each collision is presented under "UGRID" or "CTRLSECT." The UGRID information applies to collisions that occurred in Regina, Saskatoon, and Prince Albert, whereas CTRLSECT information applies to collisions that occurred on provincial highways. The collision dataset also contains detailed information about collisions, such as the collision severity, year, and distinct road feature of collision site. Table 18 shows the description of the information provided in the raw collision dataset (HIGHSPEED_AC).

Table 18: Sample Information of Raw Collision Data (HIGHSPEED_AC).

Field	Description
CASENO	Unique Identifier for Each Collision Record
SEVERITY	Collision Severity; PDO, Injury, Fatal
UGRID	Urban Grid Code/Unique Location Identifier
CTRLSECT	Road Segment Unique Identifier
ATKM	Collision Location on a Particular CTRLSEC
ACCSITE	Distinct Road Feature of Collision Site
USTREET1	Roadway Name that the Collision Occurred
USTREET2	Intersections or Roadway Features or further description
ACCSITE	Intersections or Collision Related Distinct Roadway Features
CONFIG	Vehicle position to Roadway or Each Other; Turning Movements
Year	Year collision occurred

3.2. Segmentation Scheme

The roadway classification provided in spatial datasets was different from the classification required for development of safety performance functions. It is therefore imperative to adopt a unified segmentation system for this research (disintegration of roadway networks into different categories). Though, NCHRP Report 17-45 suggests classifying freeway segments into basic freeway segments and speed-change lanes, and classifying interchange segments into collector-distributor roads (C-D), ramp and ramps terminals. But to follow NCHRP Report’s classification system, a detailed data was required which was not available for Saskatchewan high -speed roadways. Further for the ramp terminals, NCHRP report identifies the six most common ramp terminal configurations which were not applicable to Saskatchewan’s data. To give an overview of the NCHRP classification system for interchange and freeway segments, exhibits are presented in Appendix C.

To devise a final roadway classification system, Highway Capacity Manual (HCM) was consulted in addition to NCHRP Report for the integrated dataset for the development of SPFs for high-speed roadways in Saskatchewan. The final classification system was devised by making best use of available data, NCHRP and HCM's approach, and engineering judgment, where needed. As a result, eight different roadway configurations were finalised for the final roadway classification system in this research. The eight final roadway were classified for which SPFs were developed.

To develop a roadway classification system for all the roadways in this study, the spatial dataset provided by the SMHI (SRN11) was used as the base layer for the compilation of the integrated data base with a common roadway classification system. This SRN11 was a starting point and modifications were made to this shapefile for the purpose of segmentation. Other data including AADT and collisions were also added to the base layer on the basis of the segmentation scheme. The same base layer was used for to display network screening results.

The eight final roadway segment classifications used for the development of local SPFs for high-speed roadways in Saskatchewan are described in detail below.

3.2.1. Basic Freeways inside Interchange System

A basic freeway inside an interchange system in this study is defined as a homogeneous segment having the same traffic and geometric features, (traffic volumes, key geometric design features, and traffic control features) along the whole length of the segment. At a diamond interchange, the basic freeway segment inside interchange system begins at the off ramp and ends at an on ramp. The start and end points are located where the centerlines of the ramps and the freeway intersect. A cloverleaf interchange contains two basic freeway segments inside the interchange system, as shown in Figure 9.

3.2.2. Basic Freeways outside Interchange System

A basic freeway segment outside the interchange system is a homogeneous segment between two interchanges or a homogenous segment when approaching an interchange when the segment is not between two interchanges. For basic freeway segments between two interchanges, the length of the particular segment depends on the distance between the end of the first interchange's ramp influence area and the start of the second interchange's ramp influence area. Freeway segments outside a single interchange (not between two interchanges) can be of any length depending on the homogeneity of the freeway characteristics, but that length ends at the ramp influence area reached on the approach to an interchange. Similarly, when departing from an interchange, the length starts from the end point of the interchange's ramp influence area until the end of the homogenous section.

3.2.3. Off Ramps

An off ramp refers to a ramp or loop leaving a freeway segment in the direction of travel, starting from the intersection point of the ramp's centerline and the centerline of the freeway, and ending at the intersection point of the ramp's centerline with the centerline of the crossroad.

3.2.4. On Ramps

An on ramp refers to a ramp or loop that merges with the mainline freeway in the direction of travel. An on ramp begins at the intersection point of the centerline of the ramp segment and the centerline of the crossroad, and ends at its merging point on the freeway (at the intersection point of the centerline of the freeway and the centerline of the ramp).

3.2.5. Ramp Influence Areas

A ramp influence area (RIA) is a stretch of freeway 450 m long, either before an off ramp when approaching an interchange or after an on ramp when departing from an interchange. However, the length may be shorter in some cases when the homogeneous freeway section has a length less than 450 m.

3.2.6. Weaving Sections

A weaving section (W-Sec) refers to a freeway segment between the on ramp and off ramp within a cloverleaf portion of an interchange (Figure 9), or to a freeway segment between two interchanges if the length of the section is less than 750 m. When the length of freeway segment between two interchanges is less than 750 m, it is solely referred to as a weaving section and will not have any ramp influence area or basic freeway segment outside interchanges. It can be understood as a segment of a freeway at which vehicles are both entering and exiting the freeway (weaving).

3.2.7. Signalized Ramp Terminals

A signalized ramp terminal is a terminal with signalized traffic control at the intersection of a ramp and a crossroad (three-leg ramp terminal), or at the intersection of two ramps and a crossroad (four-leg ramp terminal). Ramp terminal configurations may vary, but signalized ramp terminals are considered as one type of roadway segment, irrespective of the number of legs at the ramp terminal.

3.2.8. Unsignalized Ramp Terminals

An unsignalized ramp terminal is an unsignalized (stop control, yield control, or no control) terminal between ramp(s) and crossroad. As for signalized ramp terminals, no distinction is made between different unsignalized ramp terminal configurations.

Figures 8 and 9 show the system adopted in this study for classifying high-speed roadways (freeways and interchanges) in Saskatchewan.

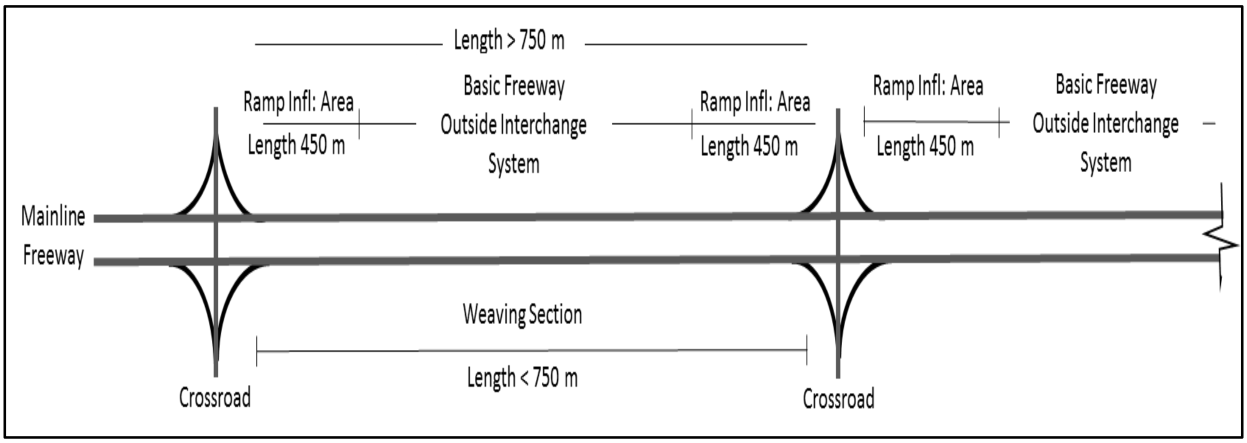


Figure 8: Classification System for Freeways

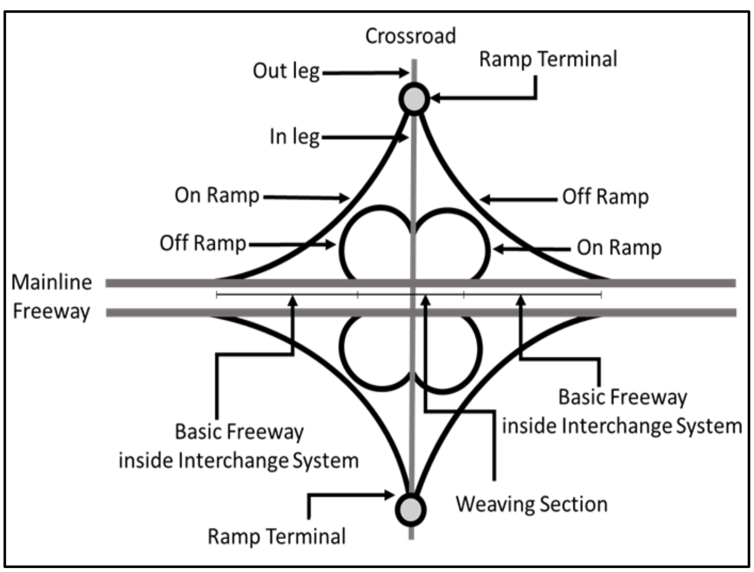


Figure 9: Classification System for Interchanges

Though the segmentation scheme followed a systematic approach where needed best engineering judgment was also used.

3.3. Development of Ready-to-Use, Integrated Database

The development of a ready-to-use database that integrates all of the different raw datasets is a crucial component of this study, and developing SPFs that take into account the broadest range of available information requires it. The finalized integrated dataset involved compiling four sets of data into one dataset: (1) roadway characteristics (from ArcGIS file format), (2) traffic volume (AADT) (from MS Excel spreadsheets and traffic flow maps in pdf format), (3) collision records (from MS Excel spreadsheets), and (4) traffic control information for ramp terminals (from MS Excel, pdf, and Google Maps® formats). A sample portion of the integrated dataset compiled for the generation of collision prediction models is shown in Table 19.

Table 19: Sample of Integrated Database used for SPF Development

a) Data Required for Basic Freeway Segment

RID	Site Type	Segment Length (m)	2009 Collisions			2010 Collisions			2011 Collisions			AADT		
			PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	2009	2010	2011
504960	BFW	807.1	2	1	0	1	1	0	4	0	0	20725	27071	27730
507163	BFW	705.6	2	0	0	0	0	0	0	0	0	6020	6020	5940
507172	BFW	783.0	5	0	0	5	1	0	3	0	0	17700	1700	17700
507204	BFW	871.4	18	0	0	19	4	0	21	1	0	27700	35167	41034
510390	BFW	901.8	8	3	0	11	2	0	14	2	0	29600	29600	29600

b) Data Required for Ramp Segment

RID	Site Type	Sub-Type	Segment Length (m)	2009 Collisions			2010 Collisions			2011 Collisions			AADT		
				PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	2009	2010	2011
597431	Off_Ramp	A	956.7	1	1	0	1	1	0	1	0	0	6000	9300	9900
788124	On_Ramp	B	267.2	1	0	0	0	0	0	0	0	0	3000	3000	3000
671335	Off_Ramp	C	476.3	1	1	0	1	0	0	1	0	0	8700	8700	8700
709527	On_Ramp	D	400.3	0	0	0	0	0	0	2	0	0	4600	4467	4334
788228	On_Ramp	B	942.1	1	0	0	1	0	0	0	0	0	3750	3750	3750

c) Data Required for Ramp Influence Area

RID	Site Type	Segment Length (m)	2009 Collisions			2010 Collisions			2011 Collisions			2009 AADT			2010 AADT			2011 AADT		
			PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	Freeway	Entrance Ramp	Exit Ramp	Freeway	Entrance Ramp	Exit Ramp	Freeway	Entrance Ramp	Exit Ramp
790718	Ramp_Inf_Area	450	18	2	0	14	0	1	9	0	0	66400	17850	16450	66150	17850	16450	65900	17850	16450
790720	Ramp_Inf_Area	450	3	0	0	1	1	0	0	0	0	48750	5750	5500	57050	5750	5500	65350	5750	5500
790665	Ramp_Inf_Area	450	0	0	0	0	1	0	1	1	0	34000	9200	11200	40300	8800	10500	46600	8400	9800
790668	Ramp_Inf_Area	450	0	0	0	1	1	0	1	1	0	18600	2600	1400	25667	2233	1367	32734	1866	1344
790673	Ramp_Inf_Area	450	0	0	0	1	0	0	2	0	0	34000	4600	3700	40300	6167	5633	46600	7734	7566

d) Data Required for Ramp Terminal

RID	Site Type	Control	2009 Collisions			2010 Collisions			2011 Collisions			2009 AADT				2010 AADT				2011 AADT			
			PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	In leg	Out Leg	Entrance Ramp(s)	Exit Ramp(s)	In leg	Out Leg	Entrance Ramp(s)	Exit Ramp(s)	In leg	Out Leg	Entrance Ramp(s)	Exit Ramp(s)
101	Terminal	Un_Sig	0	0	0	1	1	0	0	0	0	9300	16200	5200	1100	9300	16200	5200	1100	9300	16200	9300	1100
113	Terminal	Sig	1	0	0	4	1	0	1	1	0	10150	16400	5650	8800	16800	23800	4625	10000	22625	30900	6925	13600
114	Terminal	Sig	0	1	0	1	1	0	2	0	0	10150	10400	5300	2900	16800	16733	5100	3500	22625	23066	8250	6400
115	Terminal	Un_Sig	2	0	0	0	0	0	1	0	0	14700	10700	6600	4400	14700	10700	5900	4933	14700	10700	5200	5466
150	Terminal	Sig	5	1	0	4	5	0	6	0	0	35400	47800	8700	6100	35400	47800	8700	6100	35400	47800	8700	6100

However, the simplicity of Table 19 does not reflect the challenges faced while compiling the integrated database. A detailed discussion about managing and integrating all of the data from the different datasets is discussed in the following section.

3.3.1. Geospatial Referencing for Collisions

The spatial datasets from Regina, Saskatoon, and the SMHI were used to georeference collisions. Collisions in the Prince Albert region of Saskatchewan were plotted using the SRN11 shapefile, since a spatial dataset from Prince Albert is not available. Since each spatial dataset was specific to its respective jurisdiction (i.e. spatial dataset from Regina, Saskatoon, and SMHI), georeferencing all collisions using any one of the spatial datasets was not possible. Therefore, datasets were divided into four sub-datasets:

i. Regina Collisions;

ii. Saskatoon Collisions;

iii. Prince Albert Collisions; and

iv. Provincial Highway Collisions.

The collision datasets related to Regina and Saskatoon required bifurcation on the basis of segment related collisions and intersection related collisions. This bifurcation was necessitated due to the limitation that line feature file and point features file cannot be combined as one in ArcGIS. Therefore, segment and intersection related collisions were plotted by using bifurcated datasets on a respective line feature shapefile and point feature shape file. The collisions tables for Regina and Saskatoon were combined using ArcGIS with the spatial datasets using a common field (i.e. UGRID) present in collision dataset and spatial dataset. A spatial dataset for Regina required minor addition to match with location identifier field in the collision dataset. Regina shapefiles contained location information in a numerical format, whereas the collision dataset contained location information in an alphanumeric format with the letters “RE” in front of the location number (i.e. KEYNUMBER). Matching location information was achieved by adding a field into the spatial dataset and joining KEYNUMBER with the letters “RE” to produce a same field and using ArcGIS.

Prince Albert collisions were plotted with the help of the location descriptors in collision dataset found under headings USTREET 1, USTREET 2, and ACCSITE. These fields in the collision dataset provided the name of the roadway segment on which collision took place, the name of the intersecting road, distinct collision-related roadway feature information, etc. (SGI, 2007). The locations of collisions were also cross-checked using google maps® and the SMHI’s spatial dataset.

The collisions that occurred on provincial highway from the SMHI’s spatial dataset were plotted using the information given under the heading CTRLSECT in the dataset. The spatial information in the dataset is supported by another location descriptor under the heading “ATKM”. The ATKM provided an exact location of the collision (in kilometers) from the start of a particular control section (i.e., CTRLSECT). The collisions were plotted using the linear referencing tool of ArcGIS.

The X and Y coordinates for all collision records pertaining to Regina, Saskatoon, Prince Albert, and provincial highway were generated using ArcGIS so that they could be transferred to the base layer to form an integrated database.

3.3.2. Roadway Reclassification

All the spatial datasets had roadway classifications that were different from the classification system used for this research. For example, a freeway segment in a spatial dataset may be classified as a freeway or expressway, but according to the classification system for the integrated data, that segment could be classified as a freeway segment inside interchange system, a freeway outside interchange system, a ramp influence area, or a weaving section. Therefore, the roadway segments in the collision records were assigned their corresponding roadway classification according to the descriptions of the classification system for the integrated dataset as they are provided above.

Roadway classification for the segments of provincial highways on which collisions occurred was done by plotting collisions using X and Y coordinates in the integrated dataset, in which roadway classifications (i.e., freeway inside interchange system, off ramp, ramp terminal, etc.) were already defined. The “Locate Features Along Routes” linear referencing tool of ArcGIS was used for this purpose.

3.3.3. Managing Spatial Datasets

The length of a roadway segment is one of the important inputs for model development. In all spatial datasets, including SRN11, roadways (i.e. freeways, ramps, weaving sections and ramp influence areas) are broken down into smaller roadway segments with the same traffic volumes and similar geometric and traffic control features. Another important consideration for ramp segments with split legs was to determine which leg of the split ramp should be taken as the correct length for ramp segment. Figure 10 shows the small line segments on a legitimate control section on a provincial highway network. For example, CTRLSECT 0010970RDA is split into five small line segments, and a split ramp segment with CTRLSECT 0010160LUB contains two line segments having split legs at the terminal point. The smaller line segments in a control section are shown in different colours.

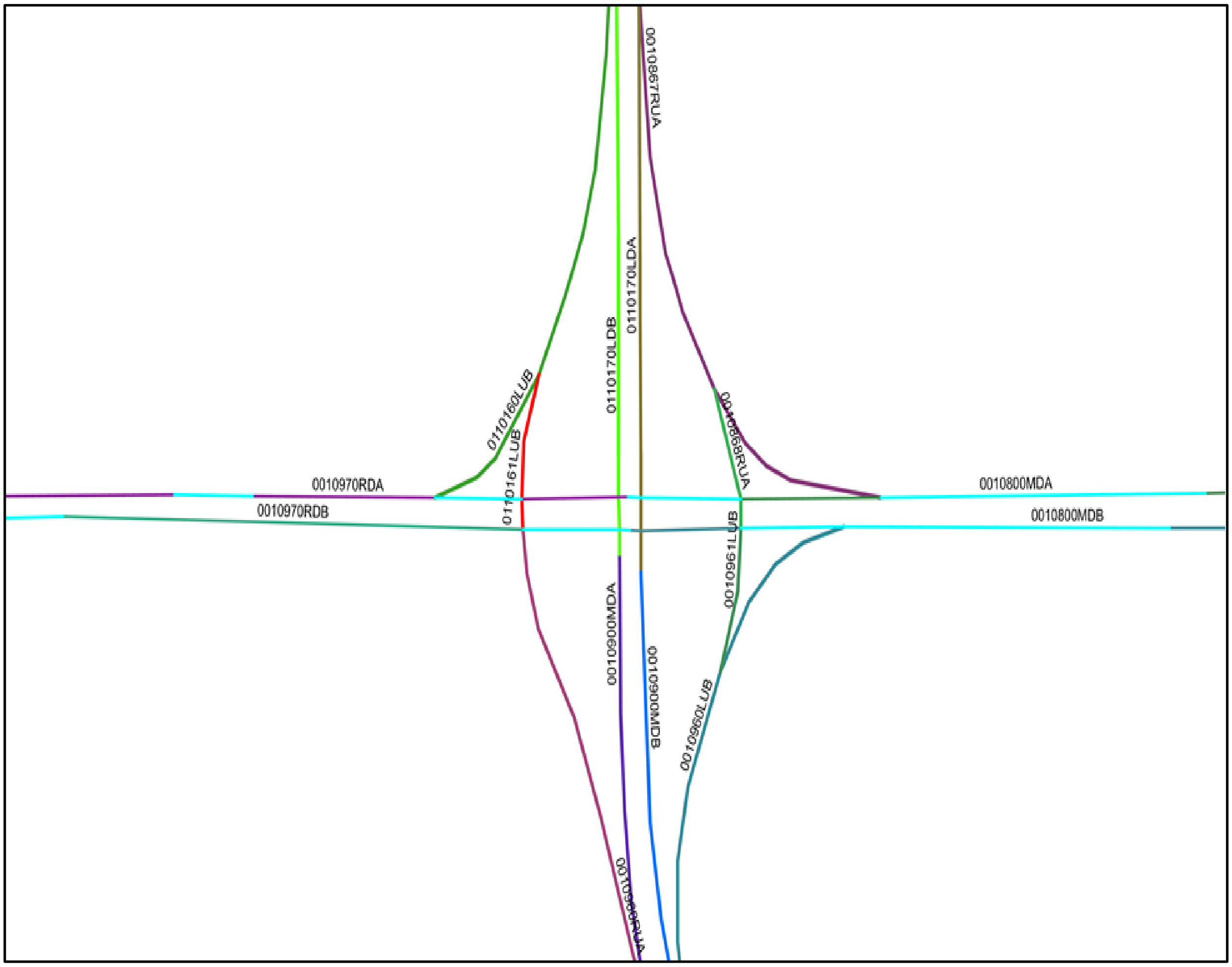


Figure 10: SMHI’s Control Sections with Small Line Segments

To overcome this issue, small line segments within legitimate road segments in the integrated dataset were carefully merged by using the merge and split feature in the editing tools of ArcGIS. Where needed, these merged lines were again split to develop roadway categories according to the classification system discussed above. To determine the length of ramps, the guidelines given in the NCHRP report were followed, which suggest that the length of the ramp should be taken from its start to the end of the split that has the highest AADT (Transportation Research Board (TRB), 2014).

To determine the lengths of basic freeway segments inside interchange systems, basic freeway segments outside interchange system, and weaving sections, the averages of lengths of both line segments in the SRN11 spatial dataset (one for each direction of travel for a roadway segment) were calculated. For ramp influence areas, a length of 450 m was chosen. Where the ramp merging and diverging points were staggered (i.e. not at the same location) and the length of influence area was less than 450 m then the average length of both line segments was taken as the length of ramp influence area.

3.3.4. Development of Unique Location Identifiers

The location identifiers given in the collision and spatial datasets were different for different jurisdictions. Therefore, there was a need to develop a unique yet similar location identifier for all the roadway configurations. All homogenous roadway segments, which for this research should be considered as a single roadway segments with a single location identifier, in each spatial dataset were comprised of more than one line having different location identifiers. For example, a basic freeway segment inside an interchange system, a weaving section, or a ramp segment, as it is defined for this research, comprised of more than one line segment having more than one UGRID or CTRLSECT. Therefore, small line segments needed to be combined into one homogenus seciton accoriding to the roadway calssification determined for this research, so that AADT, germoetric feature, and collision data could be assigen to that specific locations. The same situation was observed for ramp termianls even if the terminal points were more than one.

Figure 11(a) shows an interchange in Regina with four split ramps, with KEYNUMBER 703890 and 704450 (in red), each representing the location of a single ramp terminal point, even though each of the ramps is split into two terminal points. The other terminal points of the ramps are labelled by the number 4 (in red). Likewise, ramp segments in the same figure have more than one UGRID location identifier (i.e. RE999940 and RE999840) (in green).

Figure 11(b) shows an interchange on a provincial highway with control section information. It can be seen that the CTRLSECT 0011200MDB (line in red) is partly a basic freeway segment inside an interchange system and partly a weaving section. The same issue can be seen for CTRLSECT 0011300MDM (line in blue) in the same figure.

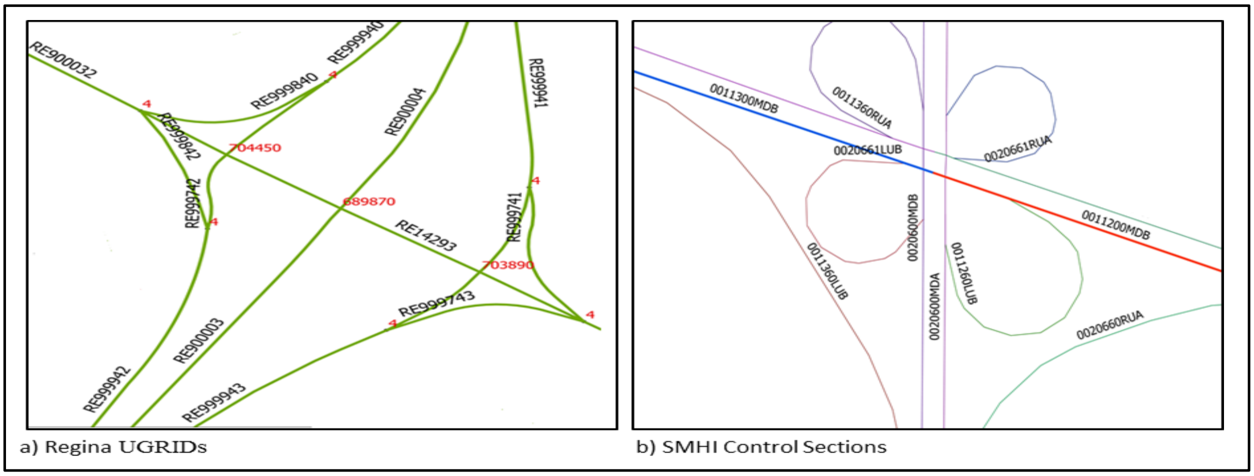


Figure 11: Examples of More Than One Location Identifier for One Roadway

To develop a new unique location identifier for all the roadways in the study area, the field "FID" in ArcGIS was used. ArcGIS automatically generates this field for all the features in a shapefile or when a new feature is added to the shapefile. However, this field is dynamic and changes whenever a new feature is added to the layer. To prevent this field from changing, a separate column was added to the attribute table of the shapefile where FID was copied. The new field, specifically for this research, was given the heading RID (roadway identification) and each roadway category was assigned its unique location identifier.

The SRN11 dataset did not contain ramp terminal points. Therefore, a point feature layer was created, and all the ramp terminal points were created manually using the "Create Feature" tool of ArcGIS. The FIDs of the manually drawn ramp terminal points were used to generated ramp terminal RIDs.

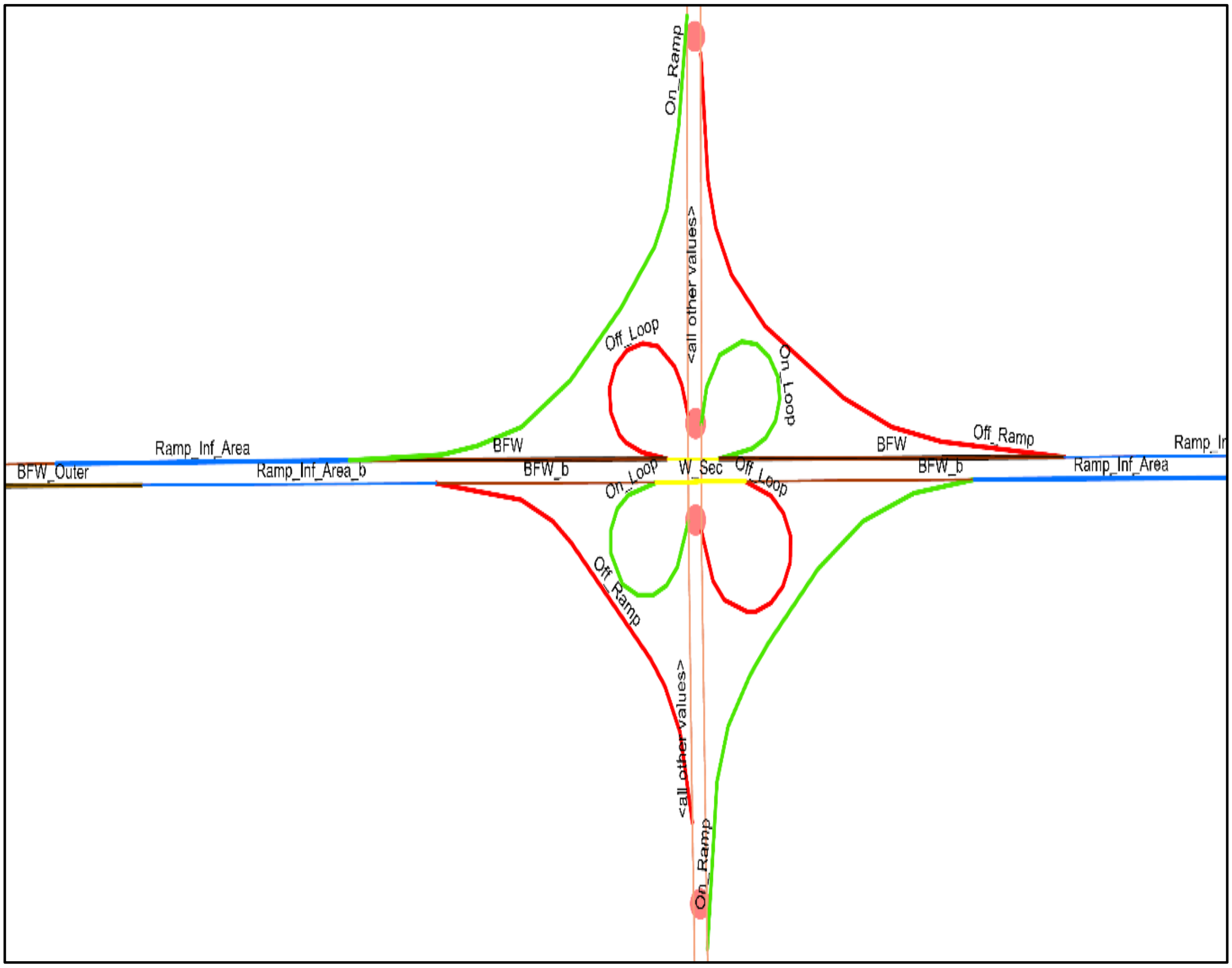
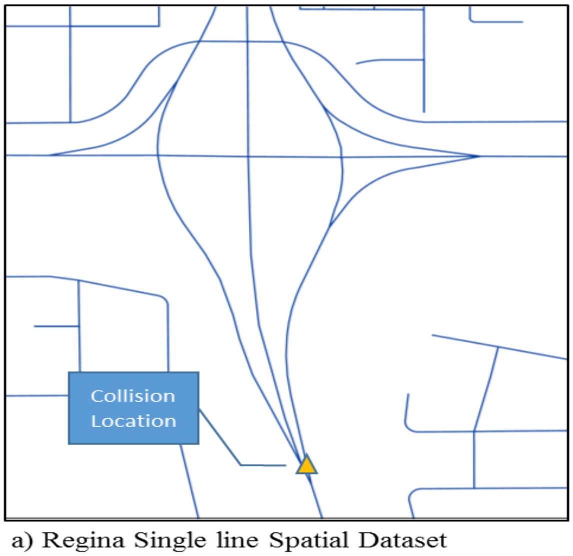


Figure 12: Sample of the Basic Framework for the Final Integrated Dataset

3.3.5. Plotting Collisions on the Base Layer of the Final Dataset Framework

The Regina and Saskatoon shapefiles assign one line to a road segment, representing both travel directions, whereas the SMHI's shapefiles assigned two lines to a road segment (one for each travel direction). The coordinates for collisions generated through these shapefiles needed to be integrated into the basic model framework developed using the SRN11 dataset, which also contains two lines for a road segment. Collisions and other essential information such as AADT, length, number of lanes, and posted speed, were assigned to only one line in the basic framework for the final integrated dataset while the second line for the opposite direction unchanged to display road network continuity and avoid the duplication of data. Figure 12 shows a sample of the base layer of the dataset framework in which thicker lines represent roadway segment to which all the data is attributed, and thinner line represent roadway segments used only for road network continuity. The roadway configuration labels ending with "_b" designate the thinner line to which no information is assigned. The circles represent ramp terminals. "X" and "Y" coordinates were used to plot collisions on the base layer of the integrated spatial dataset. Since the collisions and other information were to be assigned to only one line in the base layer, provincial highway collisions required manual shifting from a two-line representation to one line. Manually shifting Regina and Saskatoon collisions was also needed since their coordinates were generated from single-line shapefiles.

The difference in the georeferencing procedure for different jurisdictions also a required minor shifting of collision locations to their proper position. For example, if a collision occurred at the intersection of a freeway segment and ramp, its location may change to the nearest basic freeway segment inside an interchange system or to the ramp influence area when integrated into the integrated spatial dataset. Figures13 (a) and (b) show the same collision in two different spatial datasets (Regina and SRN11) at the same location (interchange at Victoria Ave and Road Rd Regina). The collision (yellow triangle) was initially plotted in the Regina shapefile but was transferred to the base layer of the developed integrated spatial dataset using its coordinates. It can be seen in Figure 13(b) that the location has changed from the intersection to the ramp influence area. Due to this minor shift in location, modifications to the location identifiers in the collision dataset were made when the location coordinates for collisions were generated.



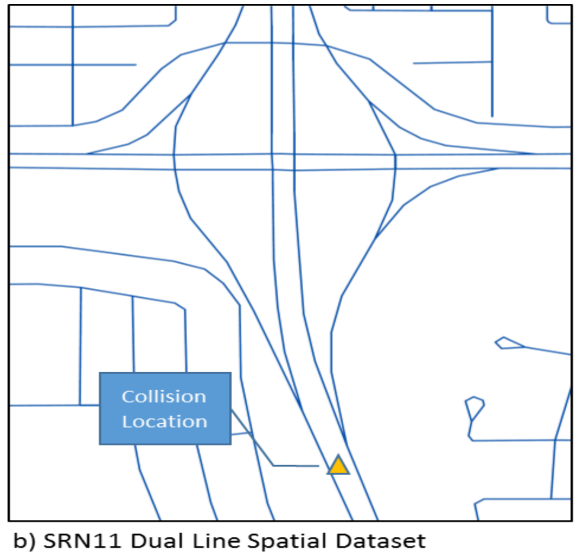


Figure 13: Difference in Collision Location in Two Different Spatial Datasets (a: Regina Spatial Dataset, b: Basic Framework of New Integrated Spatial Dataset)

3.3.6. Managing Traffic Volume Datasets

Managing the safety performance of planned or existing roadways using collisions prediction models requires known exposures (AADTs) (X Qin et al., 2004; Sayed & De Leur, 2008). Ideally, yearly traffic volumes are required for the development of SPFs. Unfortunately, most the jurisdictions like the City of Saskatoon or City of Regina do not count yearly traffic volumes every single year for many reasons. Missing yearly traffic volumes were estimated using HSM guidelines (AASHTO, 2010). For example, when two or more yearly AADTs were available, the missing year’s AADT was determined through interpolation or extrapolation, and when only a single year’s AADT was available, the same AADT was assumed for the entire study period. There were few locations in Saskatoon for which no AADT was provided in the traffic volume dataset. Therefore, the AADTs given in the Saskatoon spatial dataset were used for the entire study period (City of Saskatoon, 2014). The traffic volume data available for the City of Regina covers period from 2007-2009, so the traffic flow map of 2012 retrieved from City of Regina website was used to interpolate the missing traffic volumes for years 2010 and 2011 (City of Regina, 2014).

In Regina, some traffic volumes on minor legs (ramps) of interchanges were missing. These were estimated using available traffic volumes for the respective mainlines, crossroads and any other leg of the interchange for which data is available, with the basic assumption that traffic flow must be conserved (Sung, 2000). The relative traffic volumes of mainline freeways and crossroads on each side of an interchange were considered while estimating these missing volumes. Missing AADTs at ramp terminals and on weaving sections were calculated from the given mainline freeway traffic volumes and crossroad traffic volumes.

Since traffic volumes in Regina and Saskatoon could not be georeferenced because they are not associated with a spatial dataset, these traffic volumes were added manually to the base layer of the integrated spatial dataset by creating a separate field in the attribute table. All the manual entries were cross checked by comparing tables generated by the ArcGIS and the data provided by both the cities. Finally, all the information (i.e. AADT, segment length, the number of lanes, speed, etc.) in the base layer of the new integrated spatial dataset was combined with the collision dataset by developing queries in MS Access and using the Pivot Table function of MS Excel.

Finally, a database with a unique location identifier (RID) for each roadway segment classified using the same developed roadway configuration was developed. The integrated database's year wise subsets were developed to run regression analyses to develop SPFs. Table 20 shows a sample summary of the information present in the final integrated spatial database used for development of SPFs for high-speed roadways in Saskatchewan.

Table 20: Sample of Final Database Compiled for the development of SPFs

Year	RID	PDO	Injury	Fatal
2011	510390	14	2	0
2011	515656	10	2	0
2011	519979	6	4	1
2011	521246	0	2	1
2011	523505	2	1	0

3.3.7. Managing Collision Dataset

The raw collision dataset (HIGHSPEED_AC) for the study period (i.e. 2007-2011) includes collision records for roadways inside and outside of the scope of this study. The raw collision dataset, therefore, required some adjustments, such as separating out collisions that occurred in the study area, before using the data for development of an integrated database. The raw dataset was divided into sub-datasets based on the jurisdiction and location identifier. The sub-datasets for Regina and Saskatoon were then further divided on the basis of segment related and intersection related collisions. When these sub-datasets for Saskatoon were initially plotted, it was observed that the collision location (UGRID) was not in agreement with other location descriptors (USTREET1, USTREET2, COMMNAME, and ACCSITE), which created uncertainty. For example, a collision occurring near a ramp terminal might be assigned a UGRID for a ramp terminal, while the ACCSITE code of 12 suggests that the location of the collision was on a ramp segment. To remove any uncertainty/error in the data, a thorough examination of all the records of Saskatoon and Regina was carried out. The Regina and Saskatoon collisions were checked and validated by looking at the UGRID location identifiers of individual collision in the spatial datasets and comparing the described location with the collision dataset's location information (STREET1, STREET2, ACCSITE, CONFIG, and TAISACCDESC) (Saskatchewan Governmnet Insurance, 2013; SGI, 2007). Where there was disagreement, the final location was selected after carefully referring to all the location information available in the collision dataset.

As a result, a total of 499 collisions in Saskatoon required a shift in their UGRID location identifiers, whereas the UGRID location identifiers of collisions that occurred in Regine were in conformity with rest of the location descriptors.

Some of the provincial highway collision information had ATKM location references recorded as 999.99, which refers to an unknown location on a control section. There were a total of 714 collisions under this category that required thorough examination so that the greatest number of collision records could be retrieved and used in this study. The locations of only 14 collisions were identified using latitude/longitude and the position of 12 collisions were identified on the basis of roadway classification information (i.e. ramp), even though the ATKM location information was missing. As a result, a total of 26 collisions were retrieved.

Another 112 collisions on provincial highways had some location information, but the ATKM length exceeded the actual length of the segment and were therefore omitted. 82 other collision records had no corresponding CTRLSECT information in the spatial dataset and those collisions were also omitted. As a result, 894 collision records in total were unable to be integrated into the dataset for SPF development.

Table 21 shows a summary of all the collisions which sufficient location information both within and outside the study area.

Table 21: Summary of Raw Collision Data

Community Name	CTRLSECT	UGRID	Total Collisions
Saskatoon	371	2089	2,460
Regina	596	1690	2,286
Lumsden	288		288
Moose Jaw	280		280
North Battleford	277		277
Balgonie	236		236
Swift Current	184		184
Battleford	147		147
Prince Albert	1	113	114
Communities outside of study area	1,988		1,988
Grand Total Collisions			**8,260**

Since this research aims to develop local SPFs for collisions involving vehicles only, collisions involving trains, pedestrians, or animal were removed. These collisions were identified using the information provided in the datasets (ACCSITE, PEDACT, PEDMCF1, and TAISACCDESC).

Upon completing the roadway, classification and organization of the geometric features data, collision data, and traffic volume data, a total of 368 roadway segments were identified, which had the required information for the study. Table 22 summarizes the final number of roadways of different configurations with associated collision counts. These 368 roadway segments were used for the development of local SPFs for high-speed roadways in Saskatchewan.

Table 22: Final Roadway Configurations and Number of Corresponding Collision Records

Roadway Category	Number of Sites	Total Collisions
Basic Freeway inside Interchange System	39	672
Basic Freeway outside Interchange System	38	610
Off Ramp	82	176
On Ramp	84	122
Ramp Influence Area	43	328
Weaving Section	19	338
Terminal (signalised)	25	668
Terminal (unsignalized)	38	87
Grand Total	**368**	**3001**

3.3.8. Finalized Integrated Database

The three types datasets (i.e., collisions, traffic volumes, and spatial datasets) were successfully integrated into one new integrated spatial dataset with a single format. The integrated database was then used to group together the information related to each type of roadway configuration classifications in this study for the development and validation of SPFs.

Table 23 shows a sample of the final integrated database categorized into different high-speed roadway configurations in Saskatchewan for SPF development.

Table 23: Sample of Finalized Integrated Database for Different Roadway Types

a) Database for Basic Freeway inside Interchange System

RID	Roadway Configuration	Segment Length (km)	2009 Collisions			2010 Collisions			2011 Collisions				AADT	
			PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	2009	2010	2011
504960	BFW	0.8071	2	1	0	1	1	0	4	0	0	20725	27071	27730
507163	BFW	0.7056	2	0	0	0	0	0	0	0	0	6020	6020	5940
507172	BFW	0.783	5	0	0	5	1	0	3	0	0	17700	1700	17700
507204	BFW	0.8714	18	0	0	19	4	0	21	1	0	27700	35167	41034
510390	BFW	0.9018	8	3	0	11	2	0	14	2	0	29600	29600	29600

b) Dataset for Ramp Segments

RID	Roadway Configuration	Sub-Type	Segment Length (km)	2009 Collisions			2010 Collisions			2011 Collisions				AADT	
				PDO	Injury	Fatal	PDO	Injury	Fatal	PDO	Injury	Fatal	2009	2010	2011
597431	Off Ramp	A	0.9567	1	1	0	1	1	0	1	0	0	6000	9300	9900
788124	Off Ramp	B	0.2672	1	0	0	0	0	0	0	0	0	3000	3000	3000
671335	Off Ramp	C	0.4763	1	1	0	1	0	0	1	0	0	8700	8700	8700
709527	Off Ramp	D	0.4003	0	0	0	0	0	0	2	0	0	4600	4467	4334
788228	On Ramp	B	0.9421	1	0	0	1	0	0	0	0	0	3750	3750	3750

A larger sample of the final integrated database showing all the roadway configurations used for development of collision prediction models (SPFs) in this study is presented in Appendix D.

3.4. Chapter Summary

This chapter has described the various datasets (spatial, traffic volume, and collision datasets) and their formats as they were obtained form different jurisdictions in the study area. The roadway classification used in this study system that was designed using HSM and HCM guidelines was also described in this chapter. This chapter has discussed the development of the new integrated spatial dataset and the geospatial referencing of collisions used in that dataset, including the challenges faced and the modifications of the raw datasets required for the integration. The issues faced in compiling the integrated dataset were related to missing information, different dataset formats, transferring data from one dataset to the base layer of the integrated dataset, differences in the location identifiers used in the raw datasets, and the reclassification of roadways. A sample of the finalized integrated dataset that is used for the development of SPFs for high-speed roadways in Saskatchewan in this study was presented at the end of this chapter.

CHAPTER 4: DEVELOPMENT OF SAFETY PERFORMANCE FUNCTIONS

This chapter discusses the development of a set of SPFs using the negative binomial (NB) model and R-Language. Potential or candidate collision prediction models for a total of 8 different roadway classifications were developed. The number of locations (i.e. data points) used to generate the parameters in each model are presented for each roadway configuration. The criteria for the selection of the best models based on p-values, CURE Plots, AIC values, BIC values, and overdispersion parameters are presented along with a summary of the statistical results for each tested prediction model. The last part of this chapter discusses the validation of selected model for each roadway configuration using goodness-of-fit tests.

4.1 Developing Safety Performance Functions

The SPFs are developed for each different collision severities (i.e. total collisions, fatal and injury collisions (FI), and property damage only collisions (PDO)) for each roadway classification. The fatal and injury collisions were combined as one category (fatal and injury collisions) due to a limited number of fatal collisions in the collisions data. Regression analyses were performed, and local traffic volume data, geometric features, and traffic control features for the roadway segments were related to historical collisions records. The models are developed to predict the expected number of collisions, with the goal of identifying roadway configurations on which the numbers of expected collisions are the greatest. These new SPFs, developed specifically for high-speed roadways in Saskatchewan, can be used in the HSM's safety management process to evaluate the safety of a network of high-speed roadways in the province.

Collision data contains non-negative integers and naturally follow Poisson distributions. However, due to the randomness of collision occurrences and the overdispersion found in collision data, Poisson Gamma (a.k.a., Negative Binomial (NB) model) is a more suitable distribution to use for collision predictions modeling. The NB method is frequently used by transportation engineers and researchers because it has the capability of taking into account the RTM effect and the overdispersion found in the collision data due to random occurrences of collisions (Cheng & Washington, 2005; Elvik, 2007; Hauer et al., 2002; ITE, 2009; Persaud et al., 2001).

The integrated database for each type of roadway configuration was used to calculate the parameters in the respective candidate models and the candidate SPFs were derived based on suggested functional forms for SPFs in the literature, and were introduced into an R-Language program (a statistical computing program) along with the integrated database for the roadway configurations. The performance of each candidate model was checked by their p-values, AIC and BIC values, CURE Plots, and overdispersion parameters. Separate R programs according to the three severity levels were developed for this purpose. The R programs were also developed to perform intermediate tasks and generate input variables such as five-year average AADTs, the natural log of AADTs, combine five-year collision records according to the severity levels, etc. The R programs were written to regress segment- and terminal-related data in comma separated files (CSV). MS Access and Pivot Table were used to develop subsets of the integrated database to perform regression analysis. Each sub-database for each roadway segment class contains essential information, including five years of traffic volumes (AADT), collision records for 2007 to 2011, and geometric and traffic features such as segment lengths, a number of lanes, posted speeds, etc. The R programs were designed to perform intermediate tasks, such as to generate five-year AADT averages (divided by 1000 to produce coefficients that are easy to use), the sum of collision records, the natural log of input variables, CURE Plots.

The minimum sample size for the development of local SPFs has been debated in the literature. For example, the HSM indicates a desirable minimum sample size of 30–50 sites with more than 100 collisions per year for the calibration of SPFs (AASHTO, 2010). These guidelines for a minimum number of sites are based on best engineering judgment, but these guidelines are challenged by some researchers (Banihashemi, 2012). R. Srinivasan (2013) reported that the sample size needed for the development of SPFs should be substantially higher than the number suggested by the HSM and, in cases with less than 100 collisions per year, a larger group of sites should be identified so that the minimum number of collision per year is met (R Srinivasan et al., 2013). Ideally, there should also be sufficient data to be able to split the data into two portions: one to select a functional form and fix the parameters in the model, and another to ensure that the developed model is applicable to an entirely different set of data for similar roadways. However, in Saskatchewan, there are too few roadways of the same classification for this approach to be used. In order for local SPFs to be developed specifically for high-speed roadways in Saskatchewan, a different approach is required. Once a candidate model was selected based on the entire database corresponding to the roadway configuration, the same integrated database was then randomly divided into two sub-databases; one used for recalculating the model parameters (calibration) in the finalized functional forms, and the other for checking the model's transferability to other roadways with similar characteristics. The validation and GOF statistical methods used included MSE, MSPE, MPB, MAD, and Freeman-Tukey's R^2. The best collision prediction models will have lower statistical values (i.e. approaching zero), except for the Freeman-Tukey R^2 value, which will be closer to 1.

Table 24 below provides a summary of the candidate models used and the number of models finalized.

Table 24: Number of Candidate Models and Number of Finalized Models

Roadway Configuration	Candidate Models	Severity Types	Total	Finalized Models
Basic Freeway inside Interchange	4	3	12	3
Basic Freeway Outside Interchange	4	3	12	3
On Ramp	4	3	12	3
Off Ramp	4	3	12	3
Ramp Influence Area	8	3	24	3
Weaving Section	10	3	30	3
Ramp Terminal (signalized)	4	3	12	3
Ramp Terminal (unsignalized)	4	3	12	3
	Grand Total		**126**	**24**

The regression results for all the candidate models, including selected models, are ranked in Appendix E in terms of their correlation with the integrated database. The models in exhibits given at Appendix F showing all the candidate models for eight roadway configurations where dotted lines represent candidate models, and solid lines represent selected or finalized models. Appendix G shows the CURE Plots for each of the candidate models including selected models according to severity levels.

Table 25 shows the eight roadways configurations for high-speed roadways in Saskatchewan with the number of data points available for regression (number of sites) and the number of collisions (total collisions) for each roadway configuration. The total number of collisions used for development of SPFs was 3001.

Table 25: Available Sites and Collisions for the Development of SPFs

Roadway Category	Number of Sites	Total Collisions
Basic Freeway inside Interchange System	39	672
Basic Freeway outside Interchange System	38	610
Off Ramp	82	176
On Ramp	84	122
Ramp Influence Area	43	328
Weaving Section	19	338
Terminal (signalised)	25	668
Terminal (unsignalized)	38	87
Grand Total	**368**	**3001**

4.1.1. SPFs for Basic Freeway Segments inside Interchange Systems

A total of 39 roadway segments classified as basic freeway segments inside interchange systems for which sufficient data are available were identified. Of these 39 segments, four are located in rural areas, and the remaining 35 are located in urban areas. Due to the small number roadway segments in this categories, segments in urban and rural areas were combined, and the SPFs were developed using entire dataset for this class of roadway segment for each of the three severity types.

Four candidate models were tested to find a best-fitting model. Table 26 shows the model forms used for this classification of high-speed roadways in Saskatchewan.

Table 26: Candidate Models for Basic Freeway Segments inside Interchange Systems

Candidate Models	
$\boldsymbol{N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-1]
$\boldsymbol{N = \alpha \times L^{\beta 1} \times \left(\frac{AADT}{1000}\right)^{\beta 2}}$	[Equation 4-2]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 4-3]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-4]

The results of regression analyses for basic freeway segments inside interchange systems clearly highlighted the models to be selected for the three severity types. The results of the regression analyses of the selected models are shown in Table 27. The selection was based on the resulting p-values, CURE Plots, lower AIC values, lower BIC values, and smaller overdispersion parameters.

Table 27: Regression Results for the Selected Models for Basic Freeway Segments inside Interchange Systems

Site	Severity	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC
BFW	Total	0.019	0.009	0.998	0.022	2.590	0.004	2.242	277.650	285.056
	FI	0.089	0.055	2.702	0.000	0.894	0.038	1.637	176.610	184.012
	PDO	0.017	0.007	1.064	0.015	2.200	0.015	2.232	260.340	267.743

The calculated p-values indicate that Equation 4-3 in Table 26 was the best-fitting model for the total or PDO collisions, and Equation 4-2 was the best-fitting model for FI collisions. The selected model for total collisions had the second lowest AIC value (277.650), BIC value (285.056), and overdispersion parameter (2.242). Due to a very small difference between first and second lowest values for all three of these measures, the selection of the best-fitting model was based on the CURE Plots. For PDO collisions, the difference between the AIC, BIC, and overdispersion parameter values for all the candidate models was small, and the selection of the best-fitting model was based on the CURE Plot with ± 2 standard deviations. For FI collisions, the selection was based on the lowest AIC value (260.340), BIC value (267.743), and overdispersion parameter (2.232), as well as the best-fitting CURE Plot. All the selected models had p-values within the 99.9% confidence interval. The finalized models for basic freeway segments inside interchange systems for high-speed roadways in Saskatchewan are given below.

The selected functional form for total and PDO collisions is given by equation 4-3:

$$\boldsymbol{N_{total\ or\ PDO} = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}} \qquad \text{[Equation 4-3]}$$

The selected functional form for FI collisions is given by equation 4-2:

$$\boldsymbol{N_{FI} = \alpha \times L^{\beta 1} \times \left(\frac{AADT}{1000}\right)^{\beta 2}} \qquad \text{[Equation 4-2]}$$

Where;

N	= predicted number of collisions on a basic freeway segment;
L	= average length (km) of the basic freeway segment inside an interchange system for both travel directions;
AADT	= annual average daily traffic, combined for both directions; and
α, β1, and β2	= regression coefficients.

4.1.2. SPFs for Basic Freeway Segments outside Interchange Systems

For this type of roadway, 38 roadway segments with all the essential information available were identified, of which 12 were in rural areas and 26 were in urban areas. As for basic freeway segments inside interchange systems, the SPFs were developed for one set of basic freeway segments outside interchange systems, regardless of their urban/rural setting. Due to the small number of roadway segments in this classification, the entire dataset was used for the development of SPFs and to determine the most suitable model from within candidate models.

Four candidate models were tested to find the best-fitting model for each of the three severity types. Table 28 shows the model forms used as candidate models for the development of SPFs for basic freeway segments outside interchange systems.

Table 28: Candidate Models for Basic Freeway Segments outside Interchange Systems

Candidate Models	
$\boldsymbol{N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-5]
$\boldsymbol{N = \alpha \times L^{\beta 1} \times \left(\frac{AADT}{1000}\right)^{\beta 2}}$	[Equation 4-6]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 4-7]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-8]

The results of the regression analyses of the selected models are shown in Table 29. The selection was based on the resulting p-values, CURE Plots, lower AIC values, lower BIC values, and smaller overdispersion parameters.

Table 29: Regression Results of Selected Models for Basic Freeway Segments outside Interchange Systems

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
BFWO	Total	0.055	0.000	1.442	0.000	1.271	275.564	281.114
	FI	0.137	0.018	0.742	0.019	1.208	204.774	210.325
	PDO	0.044	0.000	1.378	0.000	1.377	246.518	252.068

The calculated p-values indicated that Model 4-5 in Table 28 was best for the total or PDO collisions. The model for total collisions had the lowest AIC value (275.564), lowest BIC value (281.114) and smallest overdispersion parameter (1.271) among all candidate models. The model for PDO collision also showed a similar trend. Model 4-8 was the best fit for FI collisions, and the AIC values, BIC values, and overdispersion parameters of the candidate models varied only slightly. Therefore the selection was based on the performance of CURE Plots. The p-values of the selected models were within the 99.9% confidence interval. The finalized models for basic freeway segments outside interchange systems of high-speed roadways in Saskatchewan are given below.

The selected functional form for total and PDO collisions is given by equation 4-5.

$$\boldsymbol{N_{total\ or\ PDO} = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}} \qquad \text{[Equation 4-5]}$$

The selected functional form for FI collisions is given by equation 4-8:

$$\boldsymbol{N_{FI} = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}} \qquad \text{[Equation 4-8]}$$

Where;

N = predicted number of collisions;

L	= average length (km) of basic freeway segment inside interchange system for both travel directions;
AADT	= annual average daily traffic, combined for both directions; and
α and β	= regression coefficients.

4.1.3. SPFs for Off Ramps

There were 82 off ramps (combined for ramps and loops) for which the essential data was available. Of those 82 sites, 9 sites were in the rural areas and 73 were in urban areas. These 82 off ramps were considered together for the development of the SPFs for this type of roadway.

Four candidate models were tested to find the best-fitting models for each of the three severity types. Table 30 shows the model forms used as candidate models for the development of SPFs for off ramps of high-speed roadways in Saskatchewan.

Table 30: Candidate Models for Off Ramps

Candidate Models	
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 4-9]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-10]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times L^{\beta 2}}$	[Equation 4-11]
$\boldsymbol{N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-12]

The results of the regression analyses of the selected models are shown in Table 31. The selection was based on the resulting p-values, CURE Plots, lower AIC viand BIC values, and smaller overdispersion parameters.

Table 31: Regression Results for Selected Models for Off Ramps

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
off Ramp	Total	0.230	0.000	0.913	0.000	1.229	281.292	286.842
	FI	0.069	0.000	0.575	0.005	0.066	124.508	130.059
	PDO	0.162	0.000	1.011	0.000	1.393	254.516	260.067

The Model 4-12 in Table 30 showed a better fit in terms of lowest AIC and BIC values, and better fitting CURE Plot, whereas, the overdispersion parameters of all the candidate models did not show much variation. The calculated p-values of the selected model were significant and within the 99.9% confidence interval. The finalized model for off ramps of high-speed roadways in Saskatchewan is given by equation 4-12.

$$\boldsymbol{N_{total,FI,PDO}} = \boldsymbol{\alpha} \times \boldsymbol{L} \times \left(\frac{\boldsymbol{AADT}}{\boldsymbol{1000}}\right)^{\boldsymbol{\beta}} \qquad \text{Equation 4-12}$$

Where;

N = predicted number of collisions;

L = average length (km) of off ramp;

AADT = annual average daily traffic; and

α, and β = regression coefficients.

4.1.4. SPFs for On Ramps

There were 84 on ramps (combined for ramps and loops) for which all the essential information was available; 11 in rural areas and 73 in urban areas. Therefore, all on ramps were considered together for the development of SPFs for on ramps.

Four candidate models were used to find the best fitting models for all three severity types. Table 32 shows the model forms used as candidate models for the development of SPFs for on ramps of the high-speed roadways in Saskatchewan.

Table 32: Candidate Models for On Ramps

Candidate Models	
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * L}}$	[Equation 4-13]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}}$	[Equation 4-14]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times L^{\beta 2}}$	[Equation 4-15]
$\boldsymbol{N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L}$	[Equation 4-16]

The results of the regression analyses of the selected models are shown in Table 33. The selection was based on the resulting p-values, CURE Plots, AIC values, BIC values, and overdispersion parameters.

Table 33: Regression Results of Selected Models for On Ramps

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
On Ramp	Total	0.111	0.000	0.742	0.001	3.134	239.355	244.906
	FI	0.028	0.000	0.747	0.011	2.146	124.178	129.728
	PDO	0.082	0.000	0.753	0.001	3.049	211.293	216.843

The calculated p-values indicated that Model 4-14 in Table 32 was best for all collision severities, with a 99.9% confidence interval. It was observed that the length did not appear to be a significant predictor in the province where the length of ramps does not vary significantly. The finalized model for on ramps of high-speed roadways in Saskatchewan is given by equation 4-14.

$$\boldsymbol{N_{total,FI,PDO}} = \boldsymbol{\alpha} \times \left(\frac{\boldsymbol{AADT}}{\boldsymbol{1000}}\right)^{\boldsymbol{\beta}} \qquad \text{[Equation 4-14]}$$

Where;

N = predicted number of collisions;

AADT = annual average daily traffic; and

α, and β = regression coefficients.

4.1.5. SPFs for Ramp Influence Areas

There were 43 ramp influence areas (RIA) for which all the essential information needed for development of SPFs was available; Due to a lower number of data points, the SPFs were developed combined for rural and urban settings. Eight candidate models were tested to find the best-fitting models for all three severities. Table 34 shows the model forms used as candidate models for the development of SPFs for ramp influence areas of the high-speed roadways in Saskatchewan.

Table 34: Candidate Models for Ramp Influence Areas

Candidate Models	
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{AADT}{1000}\right)} \times exp^{\beta 3\left(\frac{\Sigma\, RAADT}{1000}\right)} \times S$	[Equation 4-17]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L \times S$	[Equation 4-18]
$N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}$	[Equation 4-19]
$N = \alpha \times L \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta}$	[Equation 4-20]
$N = \alpha \times exp^{\left[\left(\beta 1 \frac{AADT}{1000}\right) + (\beta 2 * S)\right]}$	[Equation 4-21]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$	[Equation 4-22]
$N = \alpha \times \left(\frac{\Sigma AAADT}{1000}\right)^{\beta}$	[Equation 4-16]
$N = \alpha \times exp^{\beta\left(\frac{\Sigma AAADT}{1000}\right)}$	[Equation 4-23]

The results of the regression analyses of the selected models are shown in Table 35. The selection was based on the resulting p-values, CURE Plots, lower AIC and BIC values, and smaller overdispersion parameters.

Table 35: Regression Results of Selected Models for Ramp Influence Areas

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
Ramp Influence Area (RIA)	Total	0.057	0.000	1.073	0.000	1.245	251.966	257.517
	FI	0.104	0.008	0.699	0.013	0.871	154.081	159.632
	PDO	0.033	0.000	1.149	0.000	1.385	226.506	232.056

The calculated p-values indicated that Model 4-22 in Table 34 was the best for the total or PDO collisions, and Model 4-19 was best for FI collisions. All p-values for predictors remained within a 99.9% confidence interval. The p-values for Model 4-17 and 4-21 for total and PDO collisions were not significant (>0.05). The AIC values, BIC values, and overdispersion parameters for the remaining models were within a narrow range. Therefore, the selection of the best-fitting models was based on the CURE Plot. The finalized models for ramp influence areas of high-speed roadways in Saskatchewan are given below by equations 4-22 (total and PDO collisions) and 4-19 (FI collisions).

$$\boldsymbol{N_{total\ or\ PDO} = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}} \qquad \text{[Equation 4-22]}$$

$$\boldsymbol{N_{FI} = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}} \qquad \text{[Equation 4-19]}$$

Where;

N = predicted number of collisions;

AADT = annual average daily traffic;

L = average length (km) of ramp influence area for both travel directions; and

α and, β = regression coefficients.

4.1.6. SPFs for Weaving Sections

For Weaving Sections (W-Sec), all 19 segments were located in urban areas. The essential information was available for all weaving sections, which were used in the development of SPFs.

Ten candidate models were tested to find the best-fitting models for all three collision severities. Table 36 shows the model forms used as candidate models for the development of SPFs for weaving sections of high-speed roadways in Saskatchewan.

Table 36: Candidate Models for Weaving Sections

Candidate Models	
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{\Sigma\, RAADT}{1000}\right)} \times exp^{\beta 3 * S} \times \Sigma\, RAADT^{\beta 4}$	[Equation 4-25]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{AADT}{1000}\right)} \times exp^{\beta 3\left(\frac{\Sigma\, RAADT}{1000}\right)} \times L \times S$	[Equation 4-26]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2\left(\frac{\Sigma\, RAADT}{1000}\right)} \times S$	[Equation 4-27]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta} \times L$	[Equation 4-28]
$N = \alpha \times \left(\frac{\Sigma AADT}{1000}\right)^{\beta} \times L$	[Equation 4-29]
$N = \alpha \times \left(\frac{\Sigma AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 * S}$	[Equation 4-30]
$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$	[Equation 4-31]
$N = \alpha \times \left(\frac{\Sigma AADT}{1000}\right)^{\beta}$	[Equation 4-32]
$N = \alpha \times exp^{\left[\left(\beta 1 \frac{AADT}{1000}\right) + (\beta 2 * L)\right]}$	[Equation 4-33]
$N = \alpha \times exp^{\left[\left(\beta 1 \frac{\Sigma AADT}{1000}\right) + (\beta 2 * S)\right]}$	[Equation 4-34]

The results of the regression analyses of the selected models are shown in Table 37. The selection was based on the resulting p-values, CURE Plots, lower AIC values, lower BIC values, and smaller overdispersion parameters.

Table 37: Regression Results of Selected Models for Weaving Sections

Site	Severity	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC
Weaving Section	Total	0.001	0.000	1.082	0.000	0.047	0.005	0.426	142.945	150.345
	FI	0.000	0.000	1.038	0.002	0.073	0.001	0.321	91.441	98.841
	PDO	0.001	0.001	1.085	0.000	0.038	0.024	0.452	134.703	142.104

The calculated p-values for Model 4-25 to 4-29 were found to be insignificant and greater than 0.05. Among the remaining models, the derived p-values indicated that Model 4-30 in Table 36 was the best prediction model for all collision severities, with a 99.9% confidence interval. The AIC, BIC, and overdispersion parameter values were similar, and therefore the selection of the prediction model was based on the Cure Plot. The finalized model for on ramps of high-speed roadways in Saskatchewan is given by equation 4-30.

$$\boldsymbol{N_{total,FI,PDO} = \alpha \times \left(\frac{\Sigma AATD}{1000}\right)^{\beta 1} \times exp^{\beta 2 * S}} \qquad \text{Equation 4-30}$$

Where;

N = predicted number of collisions;

ΣAADT = sum of annual average daily traffic for weaving section and entering and exiting ramps;

S = posted speed on mainline freeway; and

α, β1, and *β2* = regression coefficients.

4.1.7. SPFs for Signalized Ramp Terminals

There were 25 signalized Ramp Terminals (RTsig) for which the essential information was available. All signalized ramp terminals were located in urban areas. Therefore, SPFs could not be developed for rural area type. The entire dataset for signalized ramp terminals was used in the development of SPFs.

Four candidate models were tested to find the best-fitting models for all collision severities. Table 38 shows the model forms used as candidate models for the development of SPFs for signalized ramp terminal (RTsig) on high-speed roadways in Saskatchewan.

Table 38: Candidate Models for Signalized Ramp Terminals

Candidate Models	
$N = \alpha \times \left(\frac{\Sigma\, RAADT}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)^{\beta 2}$	[Equation 4-35]
$N = \alpha \times \left(\frac{AAADT_{total}}{1000}\right)^{\beta}$	[Equation 4-36]
$N = \alpha \times \left(\frac{\Sigma\, RAADT}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{total}}{1000}\right)^{\beta 2}$	[Equation 4-37]
$N = \alpha \times \left(\frac{AADT_{total}}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)$	[Equation 4-38]

The results of the regression analyses of the selected models are shown in Table 39. The selection was based on the resulting p-values, CURE Plots, lower AIC and BIC values, and smaller overdispersion parameters.

Table 39: Regression Results of Selected Models for Signalized Ramp Terminals

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
Ramp Terminal (Sig)	Total	0.000	0.001	2.454	0.000	1.923	194.266	199.817
	FI	0.000	0.000	2.430	0.000	1.157	138.489	144.040
	PDO	0.000	0.002	2.405	0.000	1.980	179.566	185.117

The calculated p-values indicated that Model 4-36 in Table 38 was best for all collision severities, with a 99.9% confidence interval. The AIC and BIC values for the selected model were lowest among all the candidate models. The overdispersion parameters for all the models were similar. Therefore, based on the AIC, BIC, and CURE Plots, Model 4-36 was found to be the best-fitting model for all severity types. The finalized model for on ramps of high-speed roadways in Saskatchewan is given by equation 4-36.

$$\boldsymbol{N_{total,FI,PDO} = \alpha \times \left(\frac{AADTtotal}{1000}\right)^{\beta}} \qquad \text{Equation 4-36}$$

Where;

N = predicted number of collisions;

$AADT_{total}$ = sum of AADTs of approaching and leaving traffic of the crossroad and the entering and exiting ramps; and

α and β = regression coefficients.

4.1.8. SPFs for Unsignalized Ramp Terminals

There were 38 unsignalized ramp terminals (RTunsig) for which the essential information was available, 4 in rural areas and 34 in urban areas. The SPFs were developed for combined urban and rural unsignalized ramp terminals. The entire dataset for signalized ramp terminals was used in the development of SPFs.

Four candidate models were tested to find the best-fitting models for all collision severities. Table 40 shows the model forms used as candidate models for the development of SPFs for unsignalized ramp terminals of the high-speed roadways in Saskatchewan.

Table 40: Candidate Models for Unsignalized Ramp Terminals

Candidate Models	
$N = \alpha \times \left(\frac{\Sigma\ RAADT}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)^{\beta 2}$	[Equation 4-39]
$N = \alpha \times \left(\frac{AADT_{total}}{1000}\right)^{\beta}$	[Equation 4-40]
$N = \alpha \times \left(\frac{\Sigma AADT}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{total}}{1000}\right)^{\beta 2}$	[Equation 4-41]
$N = \alpha \times \left(\frac{AAADT_{total}}{1000}\right)^{\beta 1} \times \left(\frac{AADT_{xrd}}{1000}\right)$	[Equation 4-42]

The results of the regression analyses of the selected models are shown in Table 41. The selection was based on the resulting p-values, CURE Plots, lower AIC values and BIC values, and smaller overdispersion parameters among all the competing models.

Table 41: Regression Results of Selected Models for Unsignalized Ramp Terminals

Site	Severity	α	p-value	β	p-value	Overdispersion Parameter	AIC	BIC
Ramp Terminal (Unsig)	Total	0.000	0.000	2.030	0.000	2.950	113.980	119.527
	FI	0.001	0.001	1.408	0.017	3.745	69.199	74.749
	PDO	0.000	0.000	2.421	0.000	2.882	97.428	102.978

The calculated p-values indicated that Model 4-40 in Table 40 was best for all collision severities, with a 99.9% confidence interval. The AIC and BIC values for the selected model were the second lowest, but close to lowest values. The overdispersion parameters for all the models were similar. Therefore the selection was primarily made on the basis of the CURE Plot, and secondarily on the basis of AIC and BIC values.

The finalized model for on ramps of high-speed roadways in Saskatchewan is given by equation 4-40.

$$\boldsymbol{N_{total,FI,PDO} = \alpha \times \left(\frac{AADTtotal}{1000}\right)^{\beta}} \qquad \text{Equation 4-40}$$

Where;

N = predicted number of collisions;

$AADT_{total}$ = sum of AADTs of approaching and leaving traffic of the crossroad and the entering and exiting ramps; and

α and β = regression coefficients.

The selected functional forms for the SPFs for all of the roadway segment classifications are summarized in Table 42 according to collision severity.

Table 42: Summary of Selected Functional Form for the Collision Prediction Models

Roadway Configuration	Total Collisions	Fatal-Injury Collisions	PDO Collisions
Basic Freeway inside I/C System	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 \times L}$	$N = \alpha \times L^{\beta 1} \times \left(\frac{AADT}{1000}\right)^{\beta 2}$	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta 1} \times exp^{\beta 2 \times L}$
Basic Freeway outside I/C System	$N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}$	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$	$N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}$
Off Ramp	$N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}$		
On Ramp	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$		
Weaving Section	$N = \alpha \times \left(\frac{AADTtotal}{1000}\right)^{\beta 1} \times exp^{\beta 2 \times S}$		
Ramp Influence Area	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$	$N = \alpha \times L \times \left(\frac{AADT}{1000}\right)^{\beta}$	$N = \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$
Ramp Term (sig)	$N = \alpha \times \left(\frac{AADTtotal}{1000}\right)^{\beta}$		
Ramp Term (unsig)	$N = \alpha \times \left(\frac{AADTtotal}{1000}\right)^{\beta}$		

4.2 Model Validation Results

The validation results for the SPFs for 6 out of 8 roadway segment classes (basic freeway segment inside interchange system, basic freeway segment outside interchange system, off ramp, on ramp, ramp influence area, and unsignalized ramp terminal) are given in Table 43. However, the validation tests were not performed for weaving sections and signalized ramp terminals since sufficient segments of these types with adequate data were not available. It is assumed that the validation results for these two roadway segment classes are similar to the other six roadway segment classes.

Table 43: Validation of SPFs for 6 Classes of High-Speed Roadway Segments in Saskatchewan

Estimation Dataset (50 %)				Validation Dataset (50%)			
Roadway Configuration	**Severity Type**	**MSE**	**R^2_{ft}**	**MSPE**	**MPB**	**MAD**	**R^2_{ft}**
Basic Freeway inside Interchange System	Total	872.534	0.140	374.668	4.954	12.700	-0.279
	FI	111.836	-0.105	31.641	1.559	3.178	-0.593
	PDO	540.353	0.122	239.145	3.944	10.508	-0.281
Basic Freeway outside Interchange System	Total	421.111	-0.165	269.702	-3.247	11.221	0.378
	FI	22.256	0.162	28.002	-0.775	4.187	0.048
	PDO	197.300	-0.075	159.385	-2.528	8.091	0.389
Off Ramp	Total	4.725	0.507	16.442	1.160	2.382	-0.257
	FI	0.294	0.527	0.724	0.243	0.616	-0.390
	PDO	4.404	0.434	13.481	1.027	2.154	-0.270
On Ramp	Total	12.229	0.369	4.850	0.590	1.527	-0.616
	FI	0.719	0.073	0.516	0.070	0.466	-0.053
	PDO	8.316	0.127	3.021	0.506	1.138	-0.746
Ramp Influence Area	Total	18.835	-0.055	329.670	-6.507	9.152	0.036
	FI	8.418	-2.326	8.594	1.192	2.433	-0.054
	PDO	12.696	-0.045	233.014	-4.836	7.270	0.102
Ramp Terminal (unsignalized)	Total	55.493	0.082	9.290	-0.515	1.057	-2.470
	FI	3.125	0.078	1.861	-0.332	0.478	-10.121
	PDO	35.502	0.039	2.994	-0.200	0.620	-1.704

It can be seen in the given table that the models for FI and PDO collisions for basic freeways inside interchange systems, FI collisions on off ramps, FI and PDO collisions on on ramps, FI collisions on ramp influence areas, and FI collisions on ramp terminal (unsignalized) have MSE and MSPE values that are similar in magnitude. These similar values indicate a high level of transferability for these models. For the remaining models, the MSPE values are lower than the MSE values, except for total and PDO collisions on off ramps and ramp influence areas, which is a strong indication of under-fitting of collision prediction models and that over-fitting is unlikely. The higher MSPE values could be attributed to the low number of available sites or the high variation in the collision dataset.

The MPB measure provided the information about the magnitude and direction of average model bias as compared to the validation dataset. The MPB values vary between -6.507 and 4.954 for total collisions on basic freeway inside interchange system and total collisions on ramp influence area respectively. The prediction models for off ramp, on ramp and ramp terminal signalised can be termed as better prediction models for observed data based on low MPB values compared to the rest of the models having MPB values between (+) 1.16 and (-) 0.515. The MPB values of remaining models are also reasonably smaller it can, therefore, be considered that the SPFs are transferable to the locations having similar characteristics.

The MAD statistics provide the information about the average deviation of predicted number of collisions from an observed number of collisions. The MAD values for both configurations of freeways and the ramp influence area are higher than rest of the roadway categories. This phenomenon is understandable since higher traffic volumes on these freeway segments mean high exposure thus more possibilities of collisions and more variation in the collision data.

The MPB values provide an indication of better prediction models for off ramp, on the ramp and ramp terminal signalized SPFs whereas, the MSPE and MSE values show variations for these roadway categories. It is, therefore, a better practice to apply more than one statistical test for checking the transferability of models.

Negative values of R^2ft for validation dataset are a common occurrence when negative binomial models are used since R^2 measures were developed for linear modeling. The higher R^2 values for estimation dataset compared to validation dataset is probably an indication of the presence of a relatively small number of sites in the validation dataset.

Table 44 shows the number of observations in estimation dataset and validation datasets.

Table 44: Number of Observations in Estimation and Validation Datasets

Roadway Configuration	**Number of Observations**	
	Estimation Dataset	**Validation Dataset**
Basic Freeway inside Interchange System	20	19
Basic Freeway outside Interchange System	19	19
On Ramp	42	42
Off Ramp	41	41
Ramp Influence Area	22	21
Ramp Terminal (unsignalized)	19	19

4.3 Chapter Summary

This chapter presented the description of the development of the SPFs for high-speed roadways in Saskatchewan using the integrated database. The entire database for each roadway configuration was used for development purpose of models. For basic freeway inside interchange system, basic freeway outside interchange system, on the ramp, off ramp, ramp terminal signalized, and ramp terminal unsignalized were used. However ramp influence area, and weaving section total eight and ten candidate models were used respectively. The best-fitting model was selected using statistical tests such as significant p-values, AIC, BIC, CURE Plots, and overdispersion parameters.

For validation database was bifurcated by randomly selecting into two groups. Fifty percent of data was used for generating parameters of selected SPFs and remaining half data was used for validation purpose. The validation and transferability tests were performed using mean square error, mean square prediction error, mean prediction bias, mean absolute deviation, and Freeman Tukey R-Squared value. The validation result was presented at the end of the section.

CHAPTER 5: NETWORK SCREENING

This chapter describes the safety performance screening of a network of high-speed roadways in Saskatchewan using the developed SPFs for eight different roadway classifications. The chapter begins by describing the two chosen methods for network screening (i.e., excess EPDO method with EB adjustment and the expected EPDO method). This chapter also presents the ranking of the ten roadway segments and terminals of each configuration that are most likely to benefit the most from safety improvement projects, according to each screening method. The results are presented in tables supported by six GIS maps showing the top ten locations in the roadway network.

5.1 Network Screening Methods

The Highway Safety Manual's safety management process allows transportation engineers to screen roadway networks and identify sites which are most in need of safety improvements (AASHTO, 2010). In this management process, roadway segments are ranked according to their safety performance, based on the expected number of collisions on those roadway segments. The top ten sites appearing in the ranking list have the greatest number of expected collisions and are referred to as hotspots. Hotspots are obvious targets for safety improvements. The network screening results can be used to aid in the decision-making process so that resources are allocated to sites that are most in need of safety improvement projects, therefore maximizing the benefit of resource investments into safety improvements.

The SPFs developed for this research were used for screening the safety of the network of the high-speed roadways in Saskatchewan. Two network screening methods were used: (1) Excess Expected Average Collision Frequency with EB Adjustment, and (2) EPDO Average

Collision Frequency with EB Adjustment. The results of the network screening were incorporated into ArcGIS to visualize the magnitudes of collision frequencies in order to easily identify hotspots. Separate maps were developed for roadway segments and ramp terminals. Separate maps that combine segments and ramps were also prepared to show both types of roadways and to have a holistic view of the performance of high-speed roadway networks in Saskatchewan.

The EPDO with EB adjustment network screening method involves the conversion of fatal and injury collisions into equivalent property damage only collisions by using weight factors. Weight factors are determined based on the societal cost of collisions. The societal costs of collisions of each severity were provided for this study by Saskatchewan Government Insurance (SGI). The societal costs of each type of collision are presented in Table 45.

Table 45: Societal Costs of Collisions for the Year 2010 by SGI

Severity Type	Societal Costs
Fatal	$ 5,543,800.00
Injury	$ 134,600.00
PDO	$ 10,900

The Excess EPDO with EB adjustment method calculates the difference between EB-adjusted estimates and the predicted number of expected collisions. A positive difference shows that the targeted location is performing poorly compared to the expected number of collisions, while a negative difference shows that the site is performing better compared to the expected number of collisions. Figure 14 displays the difference between the two network screening methods used in this study.

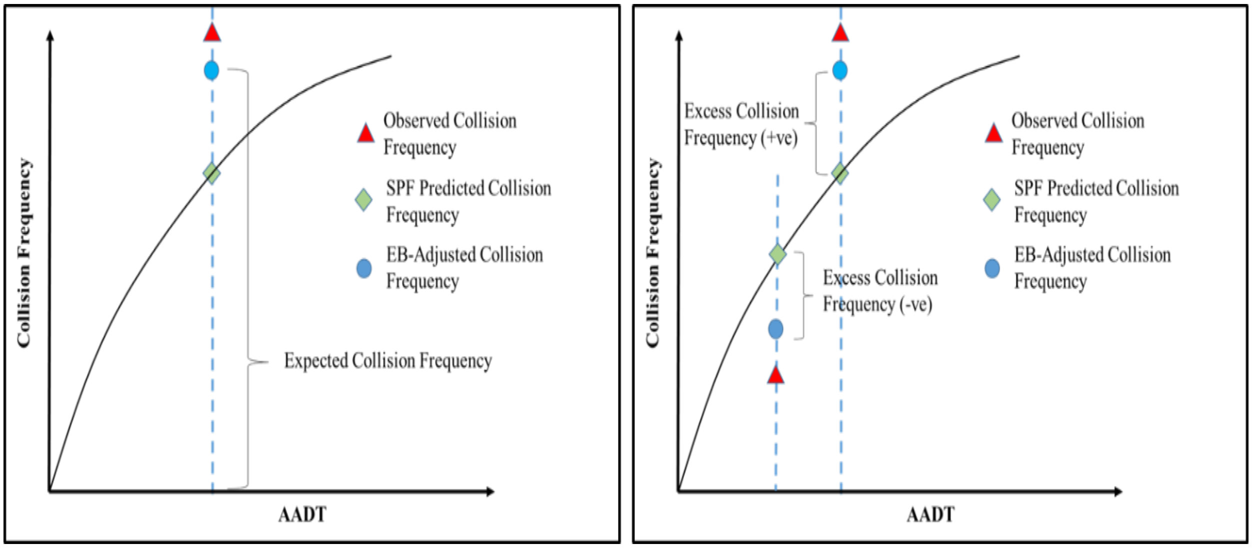


Figure 14: Description of Factors used in Network Screening Methods: (Left) EPDO with EB Adjustment Method and (right) Excess EPDO with EB Adjustment Method

5.1.1. Network Screening for Basic Freeways inside Interchange System

The two network safety screening methods were used to identify which of the 38 basic freeway segments in interchange systems in this study are expected to experience the greatest number collisions. Table 46 lists the ten freeway segments that have the greatest number of expected collisions based on the screening methods described above. Both screening methods identified the same locations ten locations in a slightly different order, and the top three hotspots identified (all in Regina) by each screening method were the same (all in Regina): (1) segment RID 516099 on Highway 1, (2) segment RID 507204 on Ring Road, and (3) segment RID 523015 on Highway 1. The positive excess EPDO values show that all ten hotspots are performing poorer than expected.

Table 46: Basic Freeways inside Interchanged System Network Screening Results

RID	Roadway	Place	EPDO		Rank	
			Excess	Expected	Excess	Expected
516099	Highway 1	Regina	90.139	118.011	1	1
507204	Ring Rd	Regina	43.899	81.778	2	2
523015	Highway 1	Regina	43.813	64.758	3	3
520489	Highway 16	Saskatoon	27.431	46.222	4	7
510390	Ring Rd	Regina	23.605	53.263	5	4
516650	Ring Rd	Regina	22.286	51.906	6	5
521246	Highway 11	Regina	15.692	32.972	7	9
520557	Highway 11	Saskatoon	15.421	42.647	8	8
521199	Highway 11	Lumsden	13.741	50.736	9	6
790679	Ring Rd	Regina	13.482	22.506	10	10

5.1.2. Network Screening for Basic Freeways outside Interchange System

Hotspots classified as basic freeway segments outside interchange systems are listed in Table 47. Of the 39 basic freeway segments outside interchange systems in this study, seven segments appeared among the top ten hotspots of each screening method. Two locations, one on Highway 11 in Saskatoon and one on Highway 1 in Regina were ranked in the same position by both network screening methods. However, the negative excess EPDO for segment RID 522978 on Highway 1 in Regina suggests that the location is actually performing better than indicated by the expected EPDO.

Table 47: Basic Freeways outside Interchange System Network Screening Results

RID	Roadway	Place	EPDO		Rank	
			Excess	Expected	Excess	Expected
520552	Highway 11	Saskatoon	102.187	805.653	1	1
519979	Highway 11	Saskatoon	69.284	461.978	2	5
516813	Highway 11	Lumsden	52.943	423.106	3	6
790661	Highway 1	Regina	47.441	389.608	4	4
790666	Highway 1	Regina	41.536	484.818	5	3
521202	Highway 11	Lumsden	40.846	325.855	6	8
517825	Highway 16	Saskatoon	35.349	682.566	7	2
500031	Highway 1	Balgonie	33.165	288.023	8	11
518809	Highway 11	Lumsden	26.35	242.913	9	14
515656	Highway 1	Moose Jaw	25.162	303.168	10	12
515698	Highway 1	Regina	23.237	246.647	11	9
520655	Highway 16	Saskatoon	11.233	296.022	15	10
522978	Highway 1	Regina	3.499	300.403	17	7

5.1.3. Network Screening for Off Ramps

Hotspots classified as off ramps are listed in Table 48. Segment RID 517449 on Highway 11 in Saskatoon appeared at the top of the list for both network screening methods. Segment RID 788231 on Highway 11 in Saskatoon was also identified as a hotspot by both screening methods. However, the remaining eight top ten hotspots identified by each network screening method are entirely different. Seven out of those eight hotspots identified by the expected EPDO with EB adjustment method have negative excess EPDO values, which suggests that these locations are actually performing better in terms of excess EPDO.

Table 48: Off Ramps Network Screening Results

RID	Name*	Place	EPDO		EPDO Rank	
			Excess	Expected	Excess	Expected
517449	A	Saskatoon	3.892	27.762	1	1
788231	C	Saskatoon	1.705	10.780	2	6
657702	A	Saskatoon	1.533	6.129	3	14
787898	C	Saskatoon	1.353	5.005	4	20
653455	C	Regina	1.168	5.791	5	17
787669	A	Saskatoon	0.773	3.575	6	28
671335	C	Saskatoon	0.663	6.028	7	15
788787	C	Saskatoon	0.545	2.678	8	44
780864	C	Saskatoon	0.507	6.284	9	11
788571	a	Saskatoon	0.446	4.076	10	24
597431	A	Regina	0.438	12.077	11	4
788117	C	Saskatoon	-0.095	6.517	50	10
520436	B	N. Battleford	-0.128	8.818	56	8
513644	A	Regina	-0.442	14.404	74	2
788570	A	Saskatoon	-0.684	10.583	78	7
788574	c	Saskatoon	-1.062	11.171	80	5
506092	A	Regina	-1.574	7.394	81	9
519975	A	Saskatoon	-3.958	12.600	82	3

* The capital letters represent ramps, and small letters represent loops in an interchange. Starting in a travel direction from East to West or from South to North in anticlockwise fashion.

5.1.4. Network Screening for On Ramps

Hotspots classified as on ramps are listed in Table 49. For this particular roadway segment class, 84 locations were analyzed. Nine of the top 10 identified hotspots appear on the lists produced using both screening methods, and the top four identified hotspots were the same for both methods. One on ramp, RID 789034 on Highway 16 in Saskatoon, has been identified by the top ten locations when using the expected EPDO method.

Table 49: On Ramps Network Screening Results

RID	Name*	Place	EPDO		EPDO Rank	
			Excess	Expected	Excess	Expected
693837	B	Saskatoon	5.452	7.927	1	2
771544	D	Saskatoon	4.884	6.552	2	3
788575	D	Saskatoon	4.346	6.354	3	4
788573	B	Saskatoon	4.285	7.970	4	1
788230	D	Saskatoon	2.813	3.714	5	6
517702	B	Saskatoon	2.322	4.289	6	5
598644	A	Saskatoon	2.257	2.989	7	10
790742	B	Saskatoon	2.183	3.242	8	9
691122	D	Regina	1.505	3.447	9	7
788228	B	Saskatoon	1.041	2.205	10	12
789034	B	Saskatoon	0.737	3.279	13	8

* The capital letters represent ramps and small letters represent loops in an interchange. Starting in travel direction from East to West or from South to North in anticlockwise fashion.

5.1.5. Network Screening for Ramp Influence Areas

Hotspots classified as ramp influence areas are listed in Table 50. Seven of the hotspots identified by the two screening methods appear in the top 10 hotspots on both lists, and the two screening methods identify the same top 6 hotspots. Three ramp influence areas, two on Highway 1 and one on Highway 16, have been identified in the top ten ranking for expected EPDO, but negative excess EPDO values suggest that these locations are actually performing better than indicated by the expected EPDO.

Table 50: Ramp Influence Areas Network Screening Results

RID	Roadway	Place	EPDO		EPDO Rank	
			Excess	Expected	Excess	Expected
505591	Highway 1	Regina	52.092	71.610	1	1
521425	Highway 1	Regina	16.895	27.946	2	4
790680	Ring Rd	Regina	13.808	32.224	3	3
790718	Highway 16	Saskatoon	11.956	38.515	4	2
790735	Highway 11	Saskatoon	11.740	23.634	5	5
790707	Highway 11	Lumsden	11.727	19.573	6	6
790694	Highway 1	Moose Jaw	7.475	13.860	7	11
520435	Highway 16	N. Battleford	4.204	14.189	8	10
518346	Highway 16	N. Battleford	3.655	13.640	9	12
790713	Highway 1	Swift Current	2.731	8.193	10	19
790668	Highway 1	Regina	-0.823	14.632	21	8
790665	Highway 1	Regina	-2.296	18.026	29	7
517733	Highway 16	Saskatoon	-7.412	14.519	37	9

5.1.6. Network Screening for Weaving Sections

Hotspots classified as weaving sections are presented in Table 51 Lists the top ten locations that have been identified for the 4-leg unsignalized intersections. For this particular group, 19 locations were analyzed, and six of these locations appear in the top ten hotspots identified by both screening methods. Segment RID 510667 (weaving section) on the Ring Road in Regina is the site which has the highest potential for safety improvement m according to both excess EPDO and expected EPDO. Four weaving section hotspots according to the expected EPDO method have a negative excess EPDO. The negative excess EPDOs suggest that these weaving sections are actually performing better than expected.

Table 51: Weaving Sections Network Screening Results

RID	Roadway	Place Name	EPDO		EPDO Rank	
			Excess	Expected	Excess	Expected
510667	Ring Rd	Regina	25.325	51.955	1	1
517796	Highway 16	Saskatoon	9.115	19.945	2	6
516608	Ring Rd	Regina	8.469	40.160	3	2
507076	Highway 1	Regina	4.068	30.157	4	3
517798	Highway 16	Saskatoon	2.539	9.898	5	10
500126	Ring Rd	Regina	2.495	28.772	6	4
517734	Highway 16	Saskatoon	1.684	8.145	7	11
520692	Highway 16	Saskatoon	0.795	7.373	8	12
516185	Highway 1	Moose Jaw	0.406	3.012	9	18
662585	Ring Rd	Regina	-0.041	0.227	10	19
519626	Highway 1	Regina	-0.434	10.765	12	9
790688	Highway 11	Regina	-7.358	17.481	16	7
501074	Ring Rd	Regina	-7.962	21.932	17	5
790749	Highway 16	Saskatoon	-21.379	14.264	19	8

5.1.7. Network Screening for All High-Speed Roadway Segments (Excluding Ramp Terminals)

The results of the expected EPDO and excess EPDO screening methods for each type of roadway segment classification (excluding ramp terminals) were compiled to identify the most dangerous hotspots in the network of high-speed roadways in Saskatchewan, regardless of their classification. The screening results for a total of 305 roadway segments were compiled for this purpose. The most dangerous hotspots on high-speed roadways in Saskatchewan according to each of the network screening methods, independent of roadway classification, are presented in Table 52. Eight hotspots were ranked within the top 10 most dangerous hotspots by both screening methods. Segment RID 520552 on Highway 11 in Saskatoon was identified as the single most dangerous hotspot in the studied high-speed roadway network by both roadway safety screening methods. The results of the province-wide network screening for high-speed roadways show that according to the excess EPDO screening method, the most dangerous hotspots are all located on basic freeway segments outside interchange systems, with the exception of three basic freeway segments inside interchange systems and one ramp influence area. Table 52 shows the compiled results of both network screening methods.

Table 52: All High-Speed Roadway Segments Network Screening Results

RID	Roadway Configuration	Roadway	Place	EPDO		Rank	
				Excess	Expected	Excess	Expected
520552	Basic Freeway Outside	Highway 11	Saskatoon	102.187	161.131	1	1
516099	Basic Freeway Inside	Highway 1	Regina	90.139	118.011	2	4
519979	Basic Freeway Outside	Highway 11	Saskatoon	69.284	101.073	3	6
516813	Basic Freeway Outside	Highway 11	Lumsden	52.943	96.107	4	7
505591	Ramp Influence Area	Highway 1	Regina	52.092	71.610	5	11
790661	Basic Freeway Inside	Highway 1	Regina	47.441	111.228	6	5
507204	Basic Freeway Outside	Ring Rd	Regina	43.899	81.778	7	9
523015	Basic Freeway Inside	Highway 1	Regina	43.813	64.758	8	15
790666	Basic Freeway Outside	Highway 1	Regina	41.536	119.767	9	3
521202	Basic Freeway Outside	Highway 11	Lumsden	40.846	74.623	10	10
517825	Basic Freeway Outside	Highway 16	Saskatoon	35.349	135.491	11	2
522978	Basic Freeway Outside	Highway 1	Regina	3.499	83.177	46	8

The network screening results compiled for all roadway segments of high-speed roadways in Saskatchewan were plotted on the following GIS maps (Figures 15 and 16) to show the locations of the locations of the most dangerous hotspots according to each network screening method. Figure 15 shows the top ten hotspots on the basis of Excess EPDO, and Figure 16 shows the top ten hotspots on the basis of Expected EPDO.

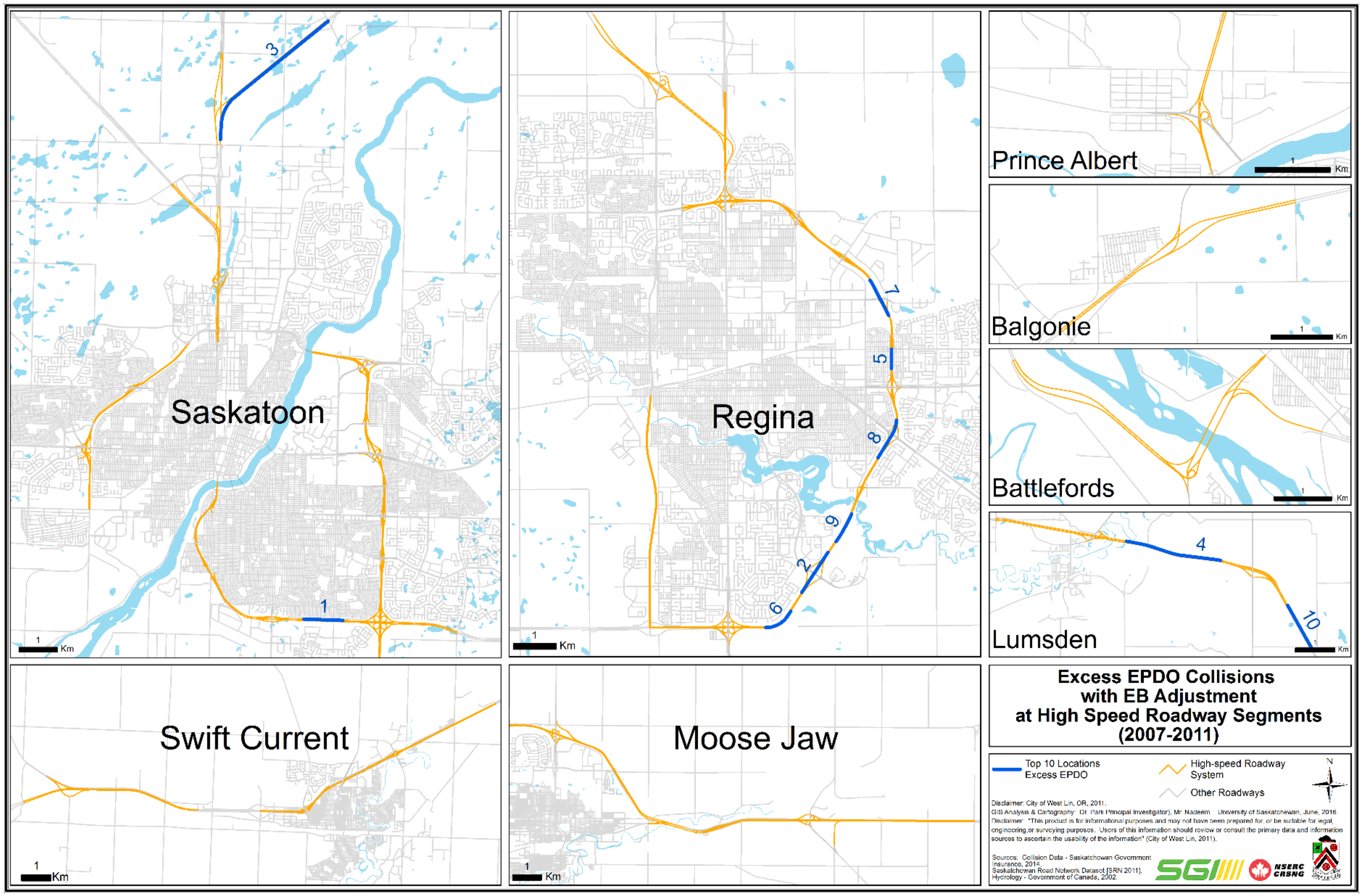


Figure 15: Top Ten Locations (Segments) Ranked using Excess EPDO Method

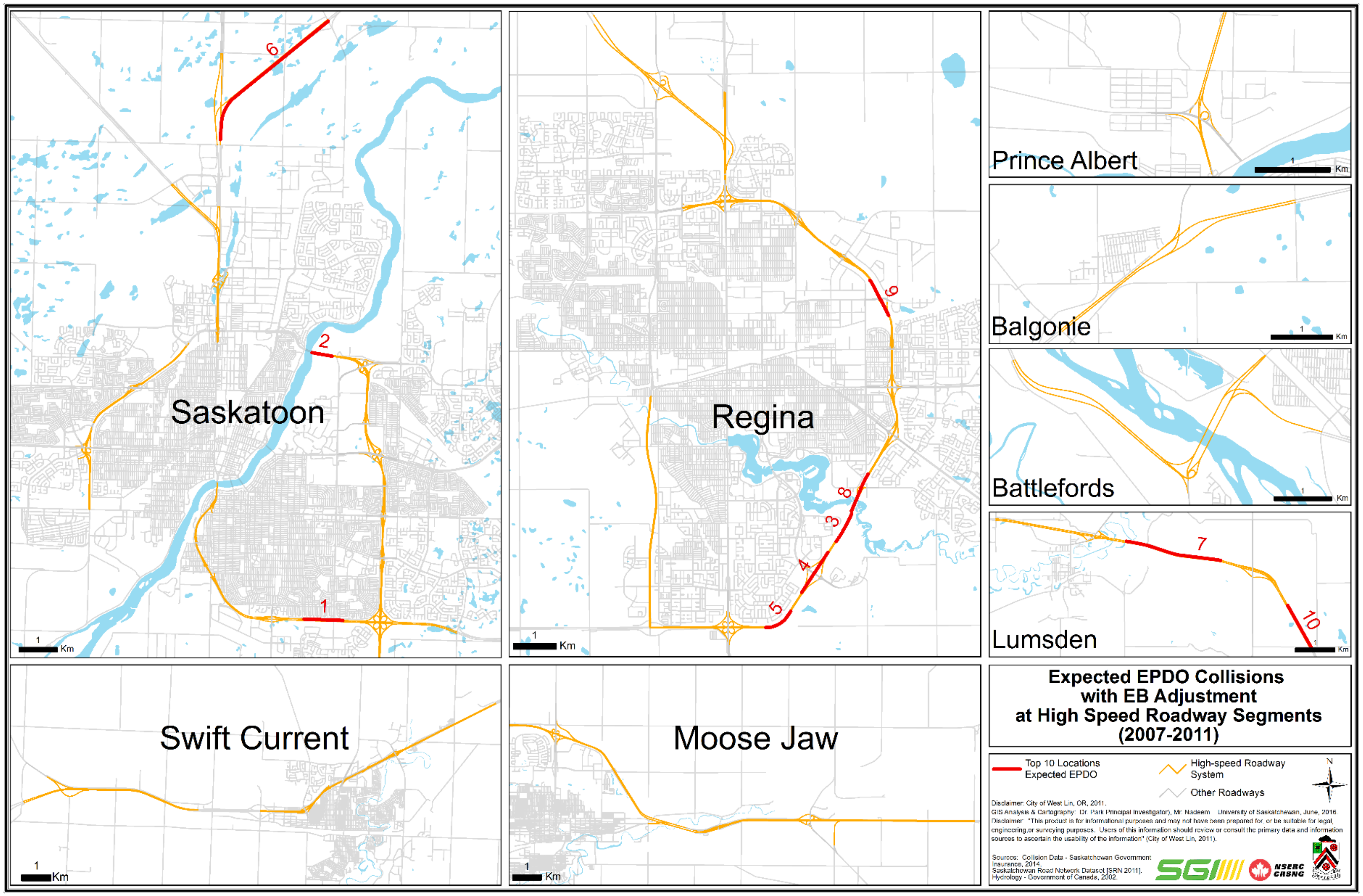


Figure 16: Top Ten Locations (Segments) Ranked using Expected EPDO Method

5.1.8. Network Screening for Ramp Terminals (signalized)

There were 25 signalized ramp terminals in the studied roadway network. The hotspots classified as signalized ramp terminals identified by each screening method are listed in Table 53. Seven sites appear on both lists, and the topmost location is the same for both network screening methods.

Table 53: Ramp Terminals (signalized) Network Screening Results

RID	Mainline	Crossroad	Direction	Place	EPDO		EPDO Rank	
					Excess	Expected	Excess	Expected
138	Highway 16	8th St E	E	Saskatoon	140.153	183.149	1	1
136	Highway 16	Taylor St E	E	Saskatoon	39.735	54.566	2	3
142	Highway 16	College Dr	E	Saskatoon	28.565	62.544	3	2
137	Highway 16	Taylor St E	W	Saskatoon	11.653	22.979	4	7
106	Highway 1	Arcola Ave E	SE	Regina	7.570	28.194	5	5
104	Highway 1	Wascana Pkwy	SE	Regina	3.843	6.620	6	17
110	Ring Rd	Dewney Av E	E	Regina	3.450	14.853	7	13
114	Ring Rd	McDonald St	SE	Regina	3.432	22.849	8	8
139	Highway 16	8th St E	W	Saskatoon	1.034	41.342	9	4
109	Ring Rd	Dewney Av E	W	Regina	0.593	8.944	10	15
108	Highway 1	Victoria Ave E-Ring Road	W	Regina	-18.727	18.556	22	9
107	Highway 1	Victoria Ave E-Ring Road	E	Regina	-39.874	16.593	23	10
150	Highway 16	Highway 11-Highway-12 (circle/Idywld)	W	Saskatoon	-40.845	23.412	24	6

5.1.9. Network Screening for Ramp Terminals (unsignalized)

Table 54 lists the hotspots classified as unsignalized ramp terminals identified by the network screening results. Of the 38 unsignalized ramp terminals studied, seven terminals appear on both lists produced by the network screening methods. Both screening methods identified the same top three most dangerous unsignalized ramp terminals. Three road segments, (1) terminal RID 156 (Highway 1/Highway 39), (2) terminal RID 155 (Highway 1/Highway 39), and (3) terminal

RID 130 (Highway 11/Ruth street), were identified as hotspots by the expected EPDO method, but had negative excess EPDO, suggesting that these ramp terminals are actually performing better than expected. Likewise, terminals RID 126, RID115, and RID 147 are identified as hotspots in terms of expected EPDO but have negative excess EPDO values that also suggest that these three locations are performing better than expected.

Table 54: Ramp Terminals (unsignalized) Network Screening Results

RID	Mainline	Crossroad	Direction	Place Name	EPDO		EPDO Rank	
					Excess	Expected	Excess	Expected
124	Ring Rd	Albert St	S	Regina	30.002	40.373	1	1
118	Ring Rd	Argyle St N	N	Regina	28.830	33.364	2	2
125	Highway 16	Attridge Dr-Preston Ave N	W	Saskatoon	14.133	30.910	3	3
112	Ring Rd	Rose Av E	W	Regina	4.245	8.701	4	5
132	Highway 11	Ruth Street - Adelaide St E&W	W	Saskatoon	2.093	3.147	5	7
157	Highway 1	Main St N	N	Moose Jaw	1.561	2.163	6	10
101	Highway 1	Highway 6 (Canam HW-Albert St)	N	Regina	1.170	7.585	7	6
156	Highway 1	Highway 39	N	Moose Jaw	-0.003	0.084	8	38
155	Highway 1	Highway 39	S	Moose Jaw	-0.016	0.185	9	37
130	Highway 11	Ruth Street - Adelaide St E/W	W	Saskatoon	-0.029	0.215	10	36
126	Highway 16	Attridge Dr-Preston Ave N	E	Saskatoon	-2.498	22.414	29	4
115	Ring Rd	Winnipeg St	N	Regina	-4.174	2.283	33	8
147	Circle Dr	22nd St W	E	Saskatoon	-10.561	2.264	37	9

5.1.10. Network Screening for All Ramp Terminals

The network screening results for all the ramp terminals, regardless of the traffic control features/signalization, are presented in Table 55. Of the 63 terminals available for analysis, eight locations are identified among the top ten most dangerous ramp terminals by both screening methods. Terminal RID 138 (signalized ramp terminal) at the intersection of Highway 16 and 8th Street E is identified as the dangerous ramp terminal in the network by both screening methods.

One ramp terminal (RID 150) is identified as a hotspot in terms of expected EPDO, but it has a negative excess EPDO value, suggesting that this terminal is actually performing better than expected.

Table 55: All Ramp Terminals Network Screening Results

RID	Mainline	Crossroad	Direction	Place Name	EPDO		EPDO Rank	
					Excess	Expected	Excess	Expected
138*	Highway 16	8th St E	E	Saskatoon	140.153	183.149	1	1
136*	Highway 16	Taylor St E	E	Saskatoon	39.735	54.566	2	3
124	Ring Rd	Albert St	S	Regina	30.002	40.373	3	5
118	Ring Rd	Argyle St N	N	Regina	28.830	33.364	4	6
142*	Highway 16	College Dr	E	Saskatoon	28.565	62.544	5	2
125	Highway 16	Attridge Dr-Preston Ave N	W	Saskatoon	14.133	30.910	6	7
137*	Highway 16	Taylor St E	W	Saskatoon	11.653	22.979	7	10
106*	Highway 1	Arcola Ave E	SE	Regina	7.570	28.194	8	8
112	Ring Rd	Rose Av E	W	Regina	4.245	8.701	9	20
104*	Highway 1	Wascana Pkwy	SE	Regina	3.843	6.620	10	23
139*	Highway 16	8th St E	W	Saskatoon	1.034	41.342	16	4
150*	Highway 12 (Idywld Dr)	Circle Dr	W	Saskatoon	-40.845	23.412	62	9

* Signalized Ramp Terminals

The network screening results of all ramp terminals in the network of high-speed roadways in Saskatchewan are plotted on GIS maps to show the top ten hotspots on the basis of Excess EPDO (Figure 17) and Expected EPDO (Figure 18).

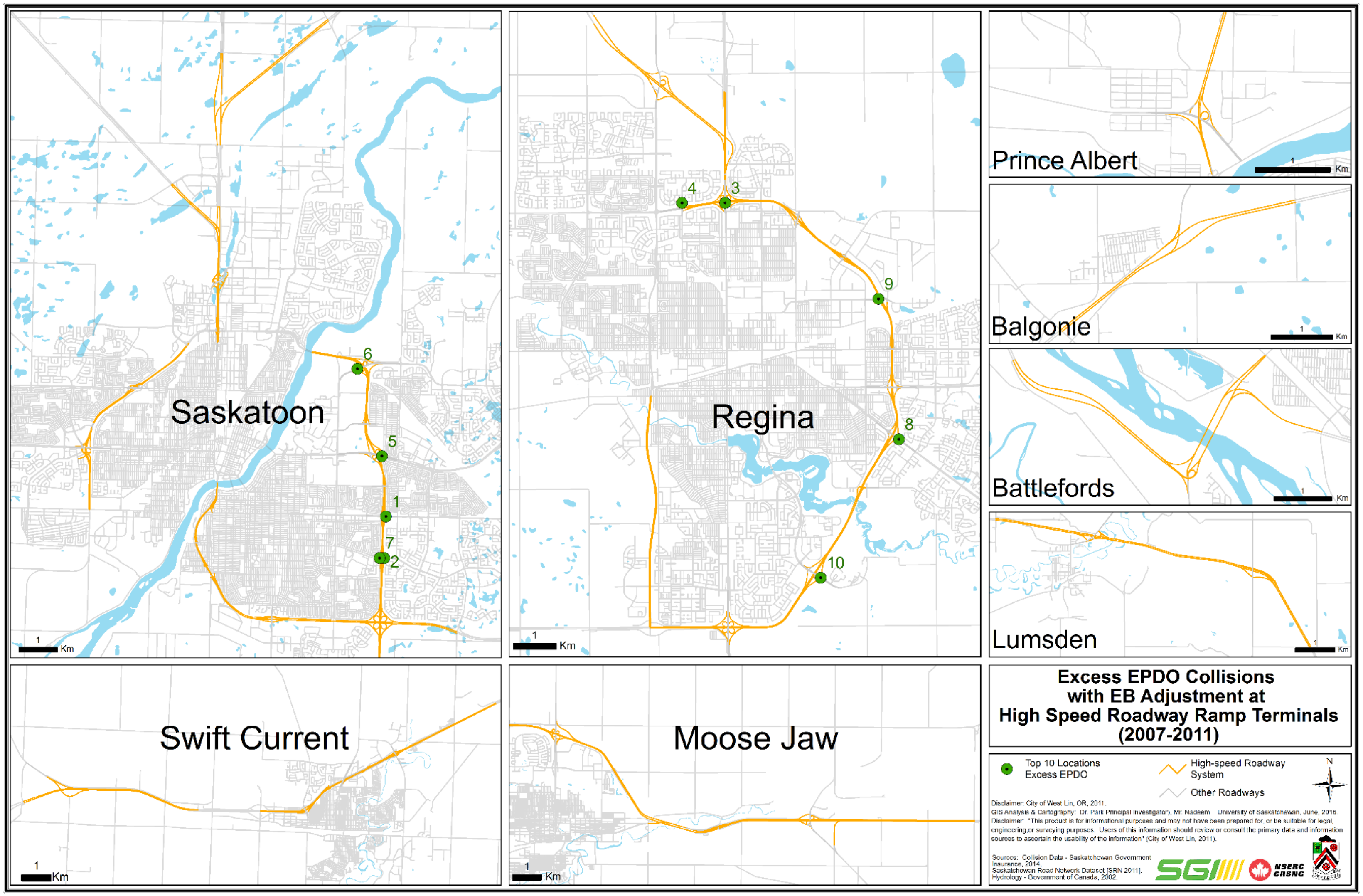


Figure 17: Top Ten Ramp Terminal Locations using Excess EPDO Method

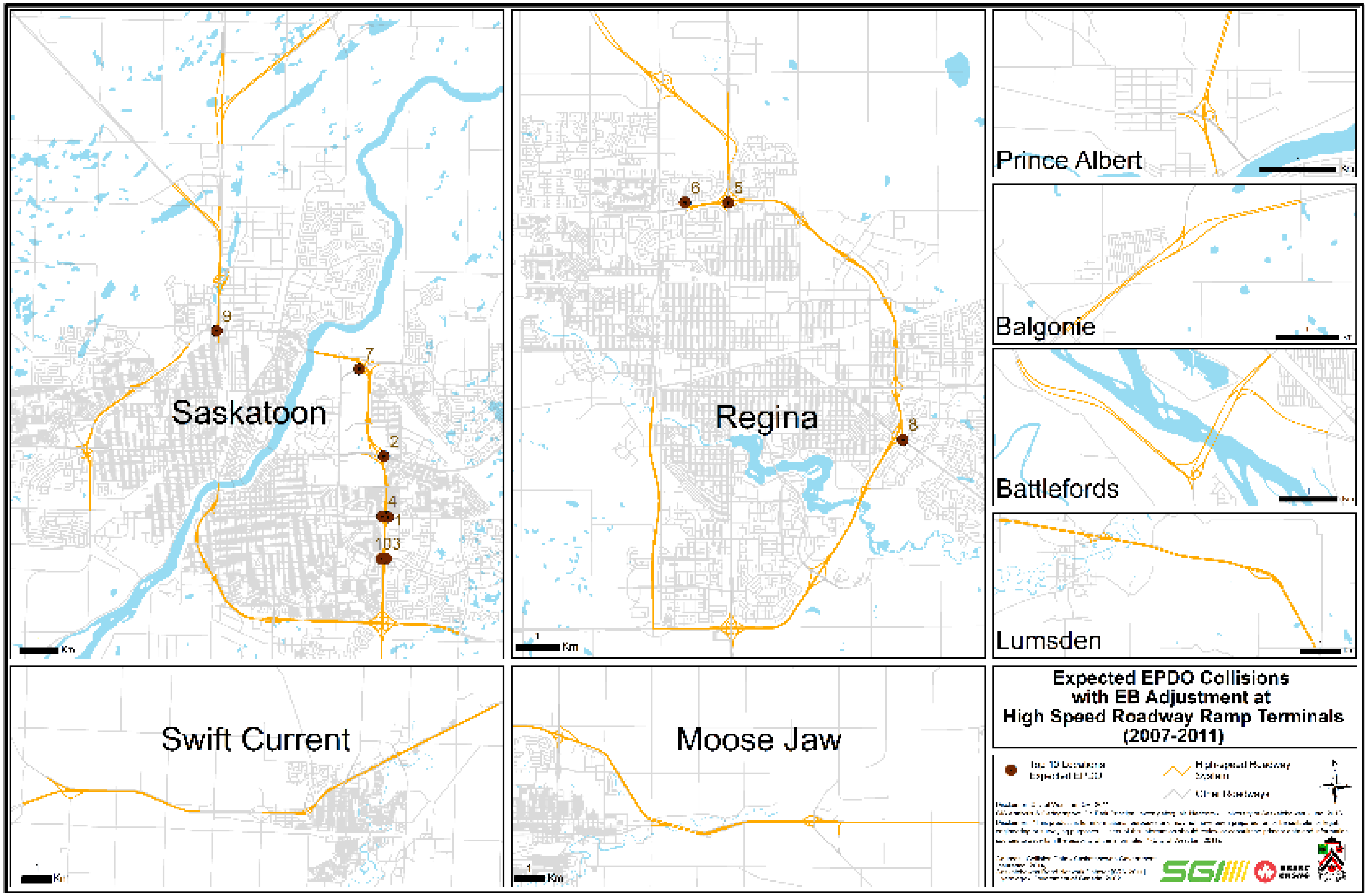


Figure 18: Top Ten Ramp Terminal Locations using Expected EPDO Method

The combined network screening results for all roadway segments and ramp terminals within the studied network of high-speed roadways in Saskatchewan are plotted on the GIS maps below. The GIS maps show the locations of the top ten hotspots on roadway segments and the top ten hotspots are ramp terminals, ranked on the basis of Excess EPDO (Figure 19) and on the basis of Expected EPDO (Figure 20).

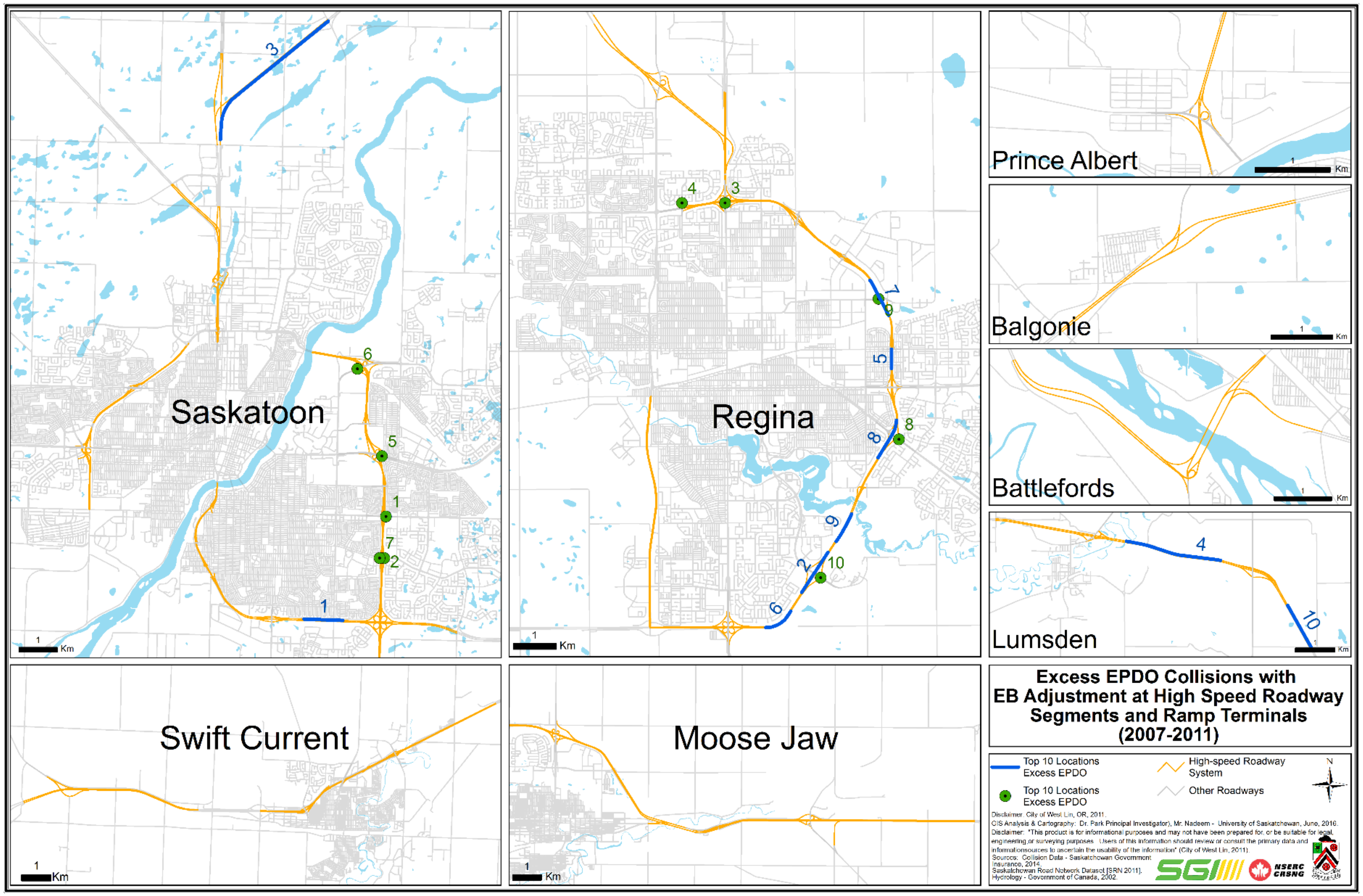


Figure 19: Top Ten Segments and Ramp Terminals using Excess EPDO Method

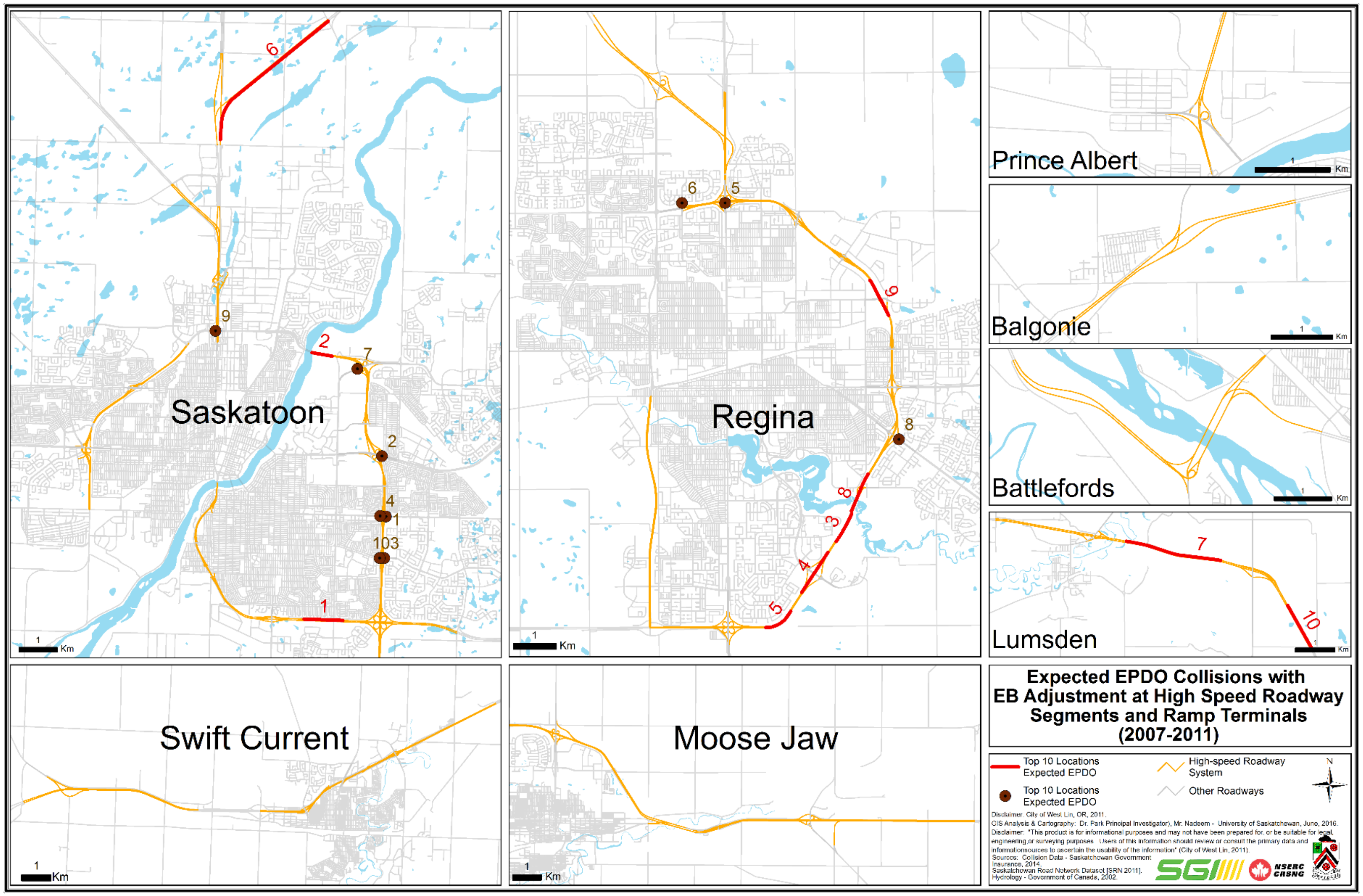


Figure 20: Top Ten Segments and Ramp Terminals using Expected EPDO Method

5.2 Chapter Summary

This chapter presented a summary of the results of the network screening performed for each of the eight roadway classifications used in this study. The two different screening methods used to rank the hotspots identified based on the expected EPDO with EB adjustment and on the excess EPDO with EB adjustment. The screening results showing the top ten most dangerous hotspots based on expected EPDO and excess EPDO were presented in tables for each of the eight roadway classifications. The results of the network safety screening were also combined for all roadway segment classifications (excluding terminals) and presented in a table and on GIS maps. The RIDs generated specifically for this research were discussed; however, end users may generate a different roadway identification system that meets their requirements using the details provided in Section 3.2, Segmentation Scheme. The results of the network screening for both types of ramp terminals (signalized and unsignalized) were also combined and presented in a table and GIS maps. Finally, the results of screening the entire high-speed roadway network in this study are summarized in two GIS maps (Figures 20 and 21) showing the locations of the roadway segment hotspots and the ramp terminal hotspots, each map showing the results of one of the network safety screening methods used.

CHAPTER 6: SUMMARY AND CONCLUSION

6.1 Summary

This study focused on the development of safety performance functions for high-speed roadways in Saskatchewan. The collisions considered for development of SPFs were the only vehicle to vehicle collisions. The study period for this research was selected from 2007-2011, based on the most recent available data for high-speed roadways including interchanges on Highway 1, Highway 11, and Highway 16. The primary goal of this research was to develop SPFs for different roadway configurations of high-speed roadways in Saskatchewan. The goal was achieved by developing an integrated database using historic collision records, annual average daily traffic volumes, and roadway geometric and traffic feature information. The SPFs were developed using Negative Binomial distribution, and the regression was performed using R-Language. Network screening was conducted using Excess EPDO method and Expected EPDO method with EB adjustment. The SPFs were used for the first step of HSM's road safety management processes "Network Screening." In total, eight roadway configuration were finalized for development of SPFs in this research 1) Basic freeway inside interchange system, 2) Basic freeway outside interchange system, 3) Off ramp, 4) On ramp, 5) Ramp influence area, 6) Weaving section, 7) Ramp terminal signalized, and 8) Ramp terminal unsignalized.

Total 126 functional forms were identified through literature review which were used to select the most appropriate functional form for each of the roadway configuration and for three collisions severity levels. As a result of regression analysis total 24 models. In all the finalized models, AADT and length were found as significant predictors except for weaving section where speed also appeared as a significant predictor. Different goodness-of-fit tests were run to check the transferability of models to the roadway with similar characteristics.

The developed SPFs were used in the first step of HSM safety management process to perform network screening using two of the thirteen network screening methods given in HSM, 1) Expected EPDO average collision frequency with EB adjustment method, and 2) Excess EPDO average collision frequency with EB adjustment method. As a result of network screening, top ten sites (hotspots) belonging to each roadway configuration were ranked from high to low score. The results of both network screening methods, 80% of the site appeared in top ten ranks, which indicates that these sites are strong candidates requiring safety improvement interventions. Most of the hotspots related to roadway segments were identified on the Ring Road Regina whereas, ramp terminals in Regina and Saskatoon showed an equal number of sites appearing in both screening methods.

An additional aspect of this research defined a methodology through which different spatial and nonspatial datasets can be joined to develop an integrated database that can be used for multipurpose research and not only for highway safety analysis. The outcomes of this research will enable highway safety professionals to quantify safety on the high-speed roadways in Saskatchewan, which can aid decision-making process and support selection of safety improvement projects. Additionally, the approach used to develop the integrated database using the information from different jurisdictions in different formats can also provide some insight to the researchers who intend to develop similar data for their analysis.

6.1.1 Segmentation Scheme

The development of roadway configurations required a disintegration of roadway facilities into different roadways types (i.e. segmentation). The segmentation was performed in light of the guideline provided in NCHRP Report 17-45 and consulting Highway Capacity Manual and where needed best engineering judgment was used. Segmentation resulted in following eight roadway configurations for Saskatchewan's data.

1. Basic Freeway inside Interchange System;
2. Basic Freeway outside Interchange System;
3. Off Ramp;
4. On Ramp;
5. Ramp Influence Area;
6. Weaving Section;
7. Ramp Terminal Signalized; and
8. Ramp Terminal Unsignalized.

6.1.2 Spatial Datasets

Mainly, study area comprised of roadways falling in urban areas of Regina, Saskatoon, Moosejaw, Swift current, Battlefords, and Prince Albert under the administrative control of Regina, Saskatoon, and Saskatchewan Ministry of Highway and Infrastructure. There was no single spatial dataset which had the roadways information of study area where necessary information required for SPFs development could be joined. Therefore, collision records were segregated on jurisdiction basis and were then joined with respective spatial dataset using matching location identifiers found in collisions and spatial datasets. The collisions on the provincial highway though required a linear referencing technique for joining collisions with the

SMHI spatial dataset. Once all the study area related collisions were plotted, their corresponding coordinates were extracted using ArcGIS. During this exercise, all the study area related collisions were manually assigned roadway type information through creating a separate field in the GIS database.

A SRN11 dataset (shapefile), provided by SMHI that contained all the roadway information of the study area, was modified by performing segmentation and plotting collisions. The modified shapefile (base layer) was also added with AADT, speed, average length, traffic control features information to develop the integrated database.

6.1.3 Traffic Volume Datasets

The development of SPFs, for estimation of predicted number of collisions, ideally requires yearly traffic volume information, which most of the jurisdictions do not collect on a yearly basis for all the roadways in their network owing to a variety of reasons. To estimate the missing annual traffic volumes, HSM's approach that is interpolation and extrapolation were adopted. The missing traffic volumes for minor legs (i.e. ramps) of some of the interchanges in the study area were also estimated based on the assumption that whatever traffic leaves the freeway will return to the freeway from opposite direction. While estimating, due consideration was given to the freeway and crossroad traffic volumes either approaching or leaving the interchange.

The segmentation was done using ArcGIS, and small line segments present in a legitimate roadway type were merged into one segment and then hose segments, where needed, were split according to the nomenclature finalized for segmentation. A geometric variable such as length was calculated using tools given in ArcGIS, and other information such as AADT was manually added to the base layer. Collisions were added to the base layer by using generated coordinates and where required careful manual shifting of collisions was made to position collisions at their respective location based on the roadway type information added to the collision dataset. The essential data (i.e. AADT, length of segment, number of lanes, speed, historical collision records, etc.) were extracted from the base layer. The extracted data from attribute tables of the spatial dataset was joined by producing summaries using Pivot Table tool of MS Excel and writing a number of queries into MS Access. The finalized database was then split into different sub-databases according to the roadway types so as to perform regression analysis and find the most suitable collision prediction model based on the local data. It may be kept in mind that each manipulation though required, was based on some assumptions thus creating uncertainty in final outcomes. The roadways where AADT was not available, were not included into the analysis.

6.1.4 Collision Datasets

The raw collision dataset was provided by Saskatchewan Government Insurance (SGI). The dataset required separation into three main categories based on the location identifiers, (i.e. UGRID for Regina, UGRID for Saskatoon, and CTRLSECT joined with ATKM for Provincial Highways). The collision data related to the Saskatoon showed some discrepancy in terms of collision locations while initial examination and it was observed that some location identifiers (i.e. UGRID) did not correspond to other location descriptions given in collision dataset. Therefore, a thorough examination was made for each of the collision records of Saskatoon by checking location in relation to other location descriptors provided in the collision dataset. As a result, a significant number of the collision were reassigned with the corrected location identifier (i.e. UGRID) to improve positional accuracy. The collision record of provincial highway also showed a significant number of collisions that could not be plotted using given information. A major chunk of these type of collisions with no specific location information (i.e. ATKM). Therefore, these collisions could not be plotted on the spatial dataset. Additionally, there were some other collision records where either the given ATKM length exceeded the legitimate length of a segment or the control section information was not available in SMHI's spatial dataset. Hence, collisions falling in those categories were excluded from the analysis.

6.1.5 Development of SPFs

The SPFs provided in the HSM may not necessarily translate Saskatchewan's local conditions for the reason that those SPFs were developed using data from different states in the US. It was found during the literature review that the local models developed for different jurisdictions in the US provided a better fit to the local data when compared to HSM's models; even though HSM's models were developed based on US data. Strong evidence is present in the literature that supports the development of local SPFs to better predict expected number of collisions based on local data. The use of Poisson-gamma model a.k.a. Negative Binomial model remained the preferred choice of highway safety researchers for developing collisions prediction models. Through literature review, a total of 126 potential model forms were identified for high-speed roadways in Saskatchewan. These models were used to find the most appropriate models for the roadway configurations and collision severity types adopted for this research. Out of those 126 candidate models, 25 models, three for each of the eight roadway configuration were finalized base on the significance (p-value), AIC, BIC, CURE Plot, and overdispersion parameter values. To select a best-fitting model, these values were considered.

6.1.6 Validation of SPFs

The finalized models were then checked for their transferability. The various goodness of fit tests were performed to check the validity of models. Initially, entire data pertaining to respective roadway configuration was used to find the most appropriate/good fitting model, and then the parameters of finalized models were re-estimates by using 50% of randomly selected data. The remaining 50% dataset was used for validation purpose. Validation was done using statistical techniques such as Mean Square Error (MSE), Mean Square Prediction Error (MSPE), Mean Absolute Deviation (MAD), Mean Prediction Bias (MPB), and Freeman-Tukey R2. The developed models showed reasonably good fit to the high-speed roadway data of Saskatchewan. However, due to a small number of observations and possible overdispersion in the collision database, these models should be used carefully with the best engineering judgment.

6.1.7 Network Screening

HSM provides thirteen different network screening methods each having its strengths and weaknesses. None of the methods has been proven as the best method. Therefore, this research adopted two methods 1) Expected EPDO average collision frequency with EB adjustment method, and 2) Excess EPDO average collision frequency with EB adjustment method. The results of each of the method provided a list of the site based on their safety performance. Both lists ranked hotspots which are likely to benefit from safety improvement projects. The Excess EPDO method identified roadways which are poorly performing with respect to the roadways having similar characteristics whereas, the expected EPDO method identified locations with a high frequency of collisions. Both lists ranked locations from high to low sores and selecting top ten location from each of the lists can identify hotspots which are in urgent need of safety improvements.

6.1.8 GIS Maps for Top Ten Hotspots

For roadway segments and rap terminals, a number of GIS maps using ArcGIS were developed to show top ten locations resulting from Excess EPDO and expected EPDO methods. It was observed that most of the location appearing in the top ten list of both network screening methods were identical with a slight difference in their rankings in each table. The ranking tables were joined with the base layer, and GIS maps were developed to show network screening results.

GIS maps were prepared for roadway segments and ramp terminals using results from Excess EPDO and expected EPDO methods. Another set of GIS maps was produced combined for roadway segments and ramp terminals on the basis of Excess EPDO and expected EPDO. By reviewing these GIS maps, particularly combined with all segments and all terminals, visual analysis of the entire road network under study is made easy. This could help safety professionals in identifying hotspots on roadways/interchanges which then be taken up for application of remaining steps of the HSM's safety management process. Hence, identification of hotspots is made much easier to select locations where safety improvement projects can be initiated.

6.2 Limitations

The SPFs are entirely dependent on the quality of data, and diligence and scrutiny were used in the development of the integrated database. It is imperative for the users of these models to understand the limitations of these SPFs and for users to be cautious while using these models to assess the safety of their road networks. For example, interpolation or extrapolation of traffic volumes to obtain yearly AADT, estimation missing traffic volumes on minor legs of interchange, and reassigning of UGRIDs to the Saskatoon collisions, may have induced more uncertainty into the models. Also, the limited number of observations available for the development of these SPFs may have affected their quality. Although the minimum required number of observations to develop good models is discussed in the literature, there is no clear limit defined. It is therefore advised that users understand the limitations of the developed SPFs for this research and use sound engineering judgment while selecting locations for safety improvements.

While selecting appropriate countermeasure to reduce collisions on a targeted roadway, it should also be kept in mind that the selected countermeasure may reduce collision on targeted location but at the same time may give rise to a bunch of other problems related to operations and maintenance. It is also a possibility that the applied countermeasure may shift the problem from the location being improved to the nearby segments, it is, therefore recommended that while evaluating safety complete interchange facility may be examined for potential improvements.

The results obtained by using these models may not be considered as absolute, and this fact is also acknowledged by HSM (HSM 2010).

6.3 Future Work and Recommendations

It is observed through this research that the agencies responsible for highway safety in Saskatchewan are maintaining their data in different formats. When any sort of data analysis is required involving data from more than one agency, such as this study, a considerable amount of resources and time is required. The agencies in Saskatchewan therefore, require an effective communication and a combined effort to develop a unified database of good quality that could be used for jurisdiction level as well as provincial level analysis. In other words, the centralized database can significantly increase the capability of analysis involving multiple jurisdictions and save a considerable time and resources. The consolidated/centralized database can be used for various analysis and not limited only to the development of SPFs or performing network screening. Other uses of the unified database could be related to another engineering, and managerial filed such planning, designing, asset management, etc.

The challenges faced during development of an integrated database utilizing datasets in different formats from various agencies for current research may serve as an example to show that why such integrated data in an amalgamated format is needed.

For studies like this, positional accuracy of collision is of vital importance, it is therefore, recommended that improvements may be made to the data recording system whereby the location coordinates of a collision could be recorded. This can be done either at the time of data entry by onsite collision reporting personnel or by the data entry personnel while updating collision records into the database. The coordinate information will overcome the issue of missing location information of collision. As a result, a maximum number of collisions and their locations will be available for analysis . Missing ATKM information, excess lengths of control sections, nonavailability of control sections in the spatial dataset, and single-point information (i.e. UGRID) representing collision location irrespective of actual location can all be taken as a good examples of missing collision location information.

The AADT data used in this research was combined for both directions. It is recommended that the AADT for roadways including interchanges may be collected for both travel directions. This will result in the availability of a larger number of data points or site for analysis thus increasing the estimation power of collision prediction models.

It is also recommended that the fresh SPFs should be developed for high-speed roadways in Saskatchewan when the AADT for both travel direction is available and meanwhile some new interchanges will also be added to the network. For example, the interchange on Circle Drive South and Preston Avenue South is developed after 2011 and was not included in this research. The new SPFs will also translate changes caused by improved vehicles, changed driver behaviors, improved geometry, improved traffic control features, changed vehicular patterns, etc. Additionally, as more year's data and sites will be available the prediction capability of models will defiantly increase.

The SPFs were developed for the freeways in the vicinity of interchanges and the roadways in the interchange (such as ramps, ramp terminals, etc.). It is, therefore, recommended that the SPFs may also be developed for the high-speed roadways outside of the interchange influence zone. This recommendation is based on the observation that high-speed roadways in Saskatchewan have varying geometric features such as divided or undivided highways, varying median widths, varying guard rails locations, varying electric poles locations, etc. The SPFs developed through this research were based on the high-speed roadways in interchange influence zones having almost similar geometric and environmental conditions. It is therefore unclear whether these SPFs can account for the variation found in the geometric features of all high-speed roadways outside of interchange influence zones or outside of urban high-speed roadways while estimating the expected number of collisions for the rest of the high-speed roadways in the province. The development of dedicated SPFs for high-speed roadways (i.e. freeways/expressways/ highways) outside of interchange infaunae zone will improve the ability of highway safety professionals to quantify safety for these type of roadways in Saskatchewan with even more confidence.

Saskatchewan has distinct weather conditions during winter and summer seasons. It is, therefore recommended that a separate set of SPFs be developed for winter and summer seasons. This will increase the capacity of transportation safety professionals to better understand the influence of seasonal variation in climatic conditions on traffic collisions. Additionally, the development of a set of SPFs is also recommended for different types of vehicles. For example, a collision involving heavy vehicles will have a different collision severity compared to a collision involving light vehicles. Collision severities can also be different if a collision occurred on a high-speed urban roadway where operating speeds are relatively low and traffic volumes are high, compared to intracity high-speed roadways with higher operating speeds and lower traffic volumes. Similar sets of SPFs can be developed for daytime and nighttime to identify the effect of roadway lighting and visibility on the safety of roadways.

Finally, it is strongly recommended that a study be conducted to evaluate the capability of the SPFs developed through this research to quantify the safety performance of high-speed roadways in Saskatchewan, in comparison with the generic SPFs provided in HSM.

REFERENCES

Saskatchewan Government Insurance (SGI). Traffic collision Statistics Annual Report. Saskatchewan Government Insurance (SGI). Regina, 2007-2011.

AASHTO. (2010). *Highway Safety Manual-1st Edition*: American Association of State Highway and Transportation Officials.

Mohamed A Abdel-Aty, PE Chris Lee, P Eng, Juneyoung Park, Jung-Han Wang, Muamer Abuzwidah, & Saif Al-Arifi. (2014). VALIDATION AND APPLICATION OF HIGHWAY SAFETY MANUAL (PART D) IN FLORIDA.

Mohamed A Abdel-Aty, & A Essam Radwan. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention, 32*(5), 633-642.

Hirotugu Akaike. (1998). A Bayesian analysis of the minimum AIC procedure *Selected Papers of Hirotugu Akaike* (pp. 275-280): Springer.

Mohammad Banihashemi. (2012). *Highway safety manual, calibration dataset sensitivity analysis.* Paper presented at the Proceedings from the 2012 TRB 91st Annual Meeting. Washington, DC January.

Kenneth P Burnham, & David R Anderson. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.

Salvatore Cafiso, Carmelo D'Agostino, & Bhagwant Persaud. (2013). *Investigating the influence of segmentation in estimating safety performance functions for roadway sections.* Paper presented at the 92th Annual Meeting of the Transportation Research Board, Washington, DC.

Salvatore Cafiso, Alessandro Di Graziano, Giacomo Di Silvestro, Grazia La Cava, & Bhagwant Persaud. (2010). Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables. *Accident Analysis & Prevention*, *42*(4), 1072-1079.

Transport Canada. (2013). Canadian Motor Vehicle Traffic Collision Statistics 2013. Transport Canada website: Transport Canada.

Li-Yen Chang. (2005). Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. *Safety science, 43*(8), 541-557.

Hongyun Chen. (2010). Comparison of safety performance by design types at freeway diverge areas and exit ramp sections.

Wen Cheng, & Simon P Washington. (2005). Experimental evaluation of hotspot identification methods. *Accident Analysis & Prevention, 37*(5), 870-881.

City of Regina. (2013). *Street Basemap; Traffic volume 2009-2011*.

City of Regina. (2014). 2012 Traffic Flow Map. In T. C. o. Regina (Ed.).

City of Saskatoon. Corporate Geographic Information Services Branch. Retrieved June. 14, 2014, from http://www.saskatoon.ca/DEPARTMENTS/Corporate%20Services/Corporate%20Information%20Services/GIS/Pages/default.aspx

City of Saskatoon. (2010). Traffic Characteristics Report. Saskatoon.

City of Saskatoon. (2014). *Street Basemap of Saskatoon; AADT for 2009-2011* [Shapefile].

DOT Colorado. (2009). Safety Performance Functions for Intersections. Denver, Colorado: Colorado Department of Transportation DTD Applied Research and Innovation Branch.

Rune Elvik. (2007). *State-of-the-art approaches to road accident black spot management and safety analysis of road networks*: Transportøkonomisk institutt.

Mariniel Flores, Peter Y Park, Angela Gardiner, & Justine Nyen. (2013). *Development of a Traffic Safety Action Plan for the City of Saskatoon.* Paper presented at the 2013 CONFERENCE AND EXHIBITION OF THE TRANSPORTATION ASSOCIATION OF CANADA-TRANSPORTATION: BETTER-FASTER-SAFER.

Lasse Fridstrøm, Jan Ifver, Siv Ingebrigtsen, Risto Kulmala, & Lars Krogsgård Thomsen. (1995). Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis & Prevention, 27*(1), 1-20.

Nicholas J Garber, & Griselle Rivera. (2010). Safety Performance Functions for Intersections on Highways Maintained by the Virginia Department of Transportation.

Alireza Hadayeghi, Amer S Shalaby, Bhagwant N Persaud, & Carl Cheung. (2006). Temporal transferability and updating of zonal level accident prediction models. *Accident Analysis & Prevention, 38*(3), 579-589.

Mohammed A Hadi, Jacob Aruldhas, Lee-Fang Chow, & Joseph A Wattleworth. (1995). Estimating safety effects of cross-section design for various highway types using negative binomial regression. *Transportation Research Record, 1500*, 169.

Ajmal Hamidi, Michael D Fontaine, & Michael J Demetsky. (2010). A Planning-level Methodology for Identifying High-crash Sections of Virginia's Primary System: Virginia Transportation Research Council.

Md Mazharul Haque, Hoong Chor Chin, & Helai Huang. (2010). Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections. *Accident Analysis & Prevention, 42*(1), 203-212.

Douglas W Harwood, FM Council, E Hauer, WE Hughes, & A Vogt. (2000). Prediction of the expected safety performance of rural two-lane highways.

Ezra Hauer. (1997). *Observational Before/After Studies in Road Safety. Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*.

Ezra Hauer. (2015). *The Art of Regression Modeling in Road Safety*: Springer.

Ezra Hauer, & Joseph Bamfo. (1997). *Two tools for finding what function links the dependent variable to the explanatory variables.* Paper presented at the Proceedings of the ICTCT 1997 Conference.

Ezra Hauer, Douglas W Harwood, Forrest M Council, & Michael S Griffith. (2002). Estimating safety by the empirical Bayes method: a tutorial. *Transportation Research Record: Journal of the Transportation Research Board, 1784*(1), 126-131.

HSM. (2010, Aug 06, 2013). An Introduction to Highway Safety Manual,. from http://www.highwaysafetymanual.org/Documents/HSMP-1.pdf

Wen Hu, & Eric T Donnell. (2010). Median barrier crash severity: Some new insights. *Accident Analysis & Prevention, 42*(6), 1697-1704.

ITE. (2009). BEFORE-AND-AFTER STUDY TECHNICAL BRIEF (T. S. Council, Trans.). Washington, DC 20005 USA: Institute of Transportation Engineers.

John M Bigham, Thomas M Rice, Swati Pande, Junhak Lee, Shin Hyoung Park, Nicolas Gutierrez, & David R Ragland. (2009). Geocoding police collision report data from California: a comprehensive approach. *International Journal of Health Geographics*.

Vasin Kiattikomol, Arun Chatterjee, Joseph E Hummer, & Mary Sue Younger. (2008). Planning level regression models for prediction of crashes on interchange and noninterchange segments of urban freeways. *Journal of Transportation Engineering, 134*(3), 111-117.

Karl Kim, & Ned Levine. (1996). Using GIS to improve highway safety. *Computers, environment and urban systems, 20*(4), 289-302.

Thanh Le, & Richard Porter. (2012). Safety evaluation of geometric design criteria for spacing of entrance-exit ramp sequence and use of auxiliary lanes. *Transportation Research Record: Journal of the Transportation Research Board*(2309), 12-20.

Paul de Leur. (2010). Collision Cost Study: CRISP.

Jinyan Lu, Albert Gan, Kirolos Haleem, Priyanka Alluri, & Kaiyu Liu. (2012). *Comparing Locally Calibrated and SafetyAnalyst-Default Safety Performance Functions for Florida's Urban Freeways.* Paper presented at the Transportation Research Board 91st Annual Meeting.

Jinyan Lu, Kirolos Haleem, Priyanka Alluri, & Albert Gan. (2013a). Full versus simple safety performance functions: comparison based on urban four-lane freeway interchange influence areas in Florida. *Transportation Research Record: Journal of the Transportation Research Board*(2398), 83-92.

Jinyan Lu, Kirolos M Haleem, Priyanka Alluri, & Albert Gan. (2013b). *Full versus Simple Safety Performance Functions: A Comparison Based on Urban Four-Lane Freeway Interchange Influence Areas in Florida.* Paper presented at the Transportation Research Board 92nd Annual Meeting.

Craig Lyon, Anwar Haq, Bhagwant Persaud, & Steven T Kodama. (2005). Safety performance functions for signalized intersections in large urban areas: Development and application to evaluation of left-turn priority treatment. *Transportation Research Record: Journal of the Transportation Research Board, 1908*(1), 165-171.

Craig Lyon, Bhagwant Persaud, & Jeremy Hahn. (2011). Safety Performance Functions for Ramp Terminals at Diamond Interchanges.

Muhammad Marizwan Abdul Manan, Thomas Jonsson, & András Várhelyi. (2013). Development of a safety performance function for motorcycle accident fatalities on Malaysian primary roads. *Safety science, 60*, 13-20.

Matthew Marleau, & Eric Hildebrand. (2010). *Collision Prediction for Two Lane Rural Roads Using IHSDM: A Canadian Experience.* Paper presented at the Proceedings of the 20th Canadian Multidisciplinary Road Safety Conference, Niagara Falls, Ontario.

Qiang Meng, & Xiaobo Qu. (2012). Estimation of rear-end vehicle crash frequencies in urban road tunnels. *Accident Analysis & Prevention, 48*, 254-263.

Jae-Pil Moon, & Joseph E Hummer. (2009). Development of safety prediction models for influence areas of ramps in freeways. *Journal of Transportation Safety & Security, 1*(1), 1-17.

Bishnu Parajuli, B Persaud, Craig Lyon, & Janice Munro. (2006). *Safety performance assessment of freeway interchanges, ramps, and ramp terminals.* Paper presented at the 2006 ANNUAL CONFERENCE AND EXHIBITION OF THE TRANSPORTATION ASSOCIATION OF CANADA: TRANSPORTATION WITHOUT BOUNDARIES; HELD 17-20 SEPTEMBER, CHARLOTTETOWN, PRINCE EDWARD ISLAND.

Jordan Parisien. (2012). Development of safety performance functions and a GIS based spatial analysis of collision data for the City of Saskatoon.

Bhagwant N Persaud. (1994). Accident prediction models for rural roads. *Canadian Journal of Civil Engineering, 21*(4), 547-554.

Bhagwant N Persaud, Richard A Retting, Per E Garder, & Dominique Lord. (2001). Safety effect of roundabout conversions in the united states: Empirical bayes observational before-after study. *Transportation Research Record: Journal of the Transportation Research Board, 1751*(1), 1-8.

X Qin, JN Ivan, & N Ravishanker. (2004). SELECTING EXPOSURE MEASURES IN CRASH RATE PREDICTION FOR TWO-LANE HIGHWAY SEGMENTS. *Accident Analysis & Prevention, 36*(2).

Xiao Qin, John N Ivan, Nalini Ravishanker, & Junfeng Liu. (2005). Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov Chain Monte Carlo modeling. *Journal of Transportation Engineering, 131*(5), 345-351.

Saskatchewan Governmnet Insurance. (2013). 2011 Saskatchewan Traffic Accidents Facts. Available from SGI Retrieved July 14, 2014, from Saskatchewan Government Insurance https://www.sgi.sk.ca/about/publications/collisionstats/2011.html

Saskatchewan Ministry of Highways & Infrastructure. (20090101). Saskatchewan Updates Road Network [SRN11] (mapDigital). Retrieved September. 23, 2013, from Saskatchewan Ministry of Highways & Infrastructure ftp://portaldata:freedata@ftp.isc.ca/PackagedData/Sask_Highways/SURN11.zip

Saskatchewan Ministry of Highways and Infrastruture. (2014). *SaskHighway Network; AADT 2009-2011*.

City of Saskatoon. Transportation Data Model (V1.23.02). City of Saskatoon Corporate Information Services Branch. Saskatoon, Saskatchewan, 2005.

Tarek Sayed, & Paul De Leur. (2008). *Collision prediction models for British Columbia*.

SGI. (Ed.) (2007) Saskatchewan Traffic Accident Information System (TAIS) Data Dictionary

Regina, Saskatchewan: Saskatchewan Government Insurance.

SGI. (2014). 2013 Saskatchewan Traffic Accident Facts: Saskatchewan Government Insurance.

Bhagwant Srinivasan, Craig Lyon, Bhagwant Persaud, Carol Martell, & Jongdae Baek. (2011). *Methods for Identifying High Collision Concentration Locations (HCCL) for Potential Safety Improvements: Phase II, Evaluation of Alternative Methods for Identifying HCCL*: California Department of Transportation.

R Srinivasan, D Carter, & K Bauer. (2013). Safety Performance Function Decision Guide: SPF Calibration vs SPF Development: Report FHWASA-14-004. Federal Highway Administration, Washington, USA.

Raghavan Srinivasan, & Daniel Lance Carter. (2011). Development of Safety Performance Functions for North Carolina: North Carolina Department of Transportation, Research and Analysis Group.

Xiaoduan Sun, & Ram Manthena. (2009). *Applying HSM Network Screening Models to Louisiana Rural Two-Lane Highway Network.* Paper presented at the Logistics@ sThe Emerging Frontiers of Transportation and Development in China.

Nakmoon Sung. (2000). *Evaluation of safety at freeway interchanges.* Michigan State University. Department of Civil and Environmental Engineering.

Lorrie Laing Susan Herbel, Colleen McGovern. Highway Safety Improvement Program (HSIP) Manual. U.S.DOT. Washington DC, 2010.

Transportation Research Board (TRB). (2014). NCHRP 17-45 [Completed], (Vol. 2014).

Jason Young, & Peter Y Park. (2013). Benefits of small municipalities using jurisdiction-specific safety performance functions rather than the Highway Safety Manual's calibrated or uncalibrated safety performance functions. *Canadian Journal of Civil Engineering*, *40*(6), 517-527.

LIST OF APPENDICES

Appendix A- Study Area Map

Appendix B- Sample Calculation for Network Screening (Expected EPDO and Excess EPDO)

Appendix C- NCHRP Report 17-45 Segmentation Scheme

Appendix D- Sample of Final Database Used for Development of SPFs

Appendix E- Regression Results for All Candidate Models

Appendix F- Prediction Models

Appendix G- CURE Plots of All Candidate Models

Appendix H- Summary of Collision Records Excluded from Analysis

Appendix A

Study Area Map

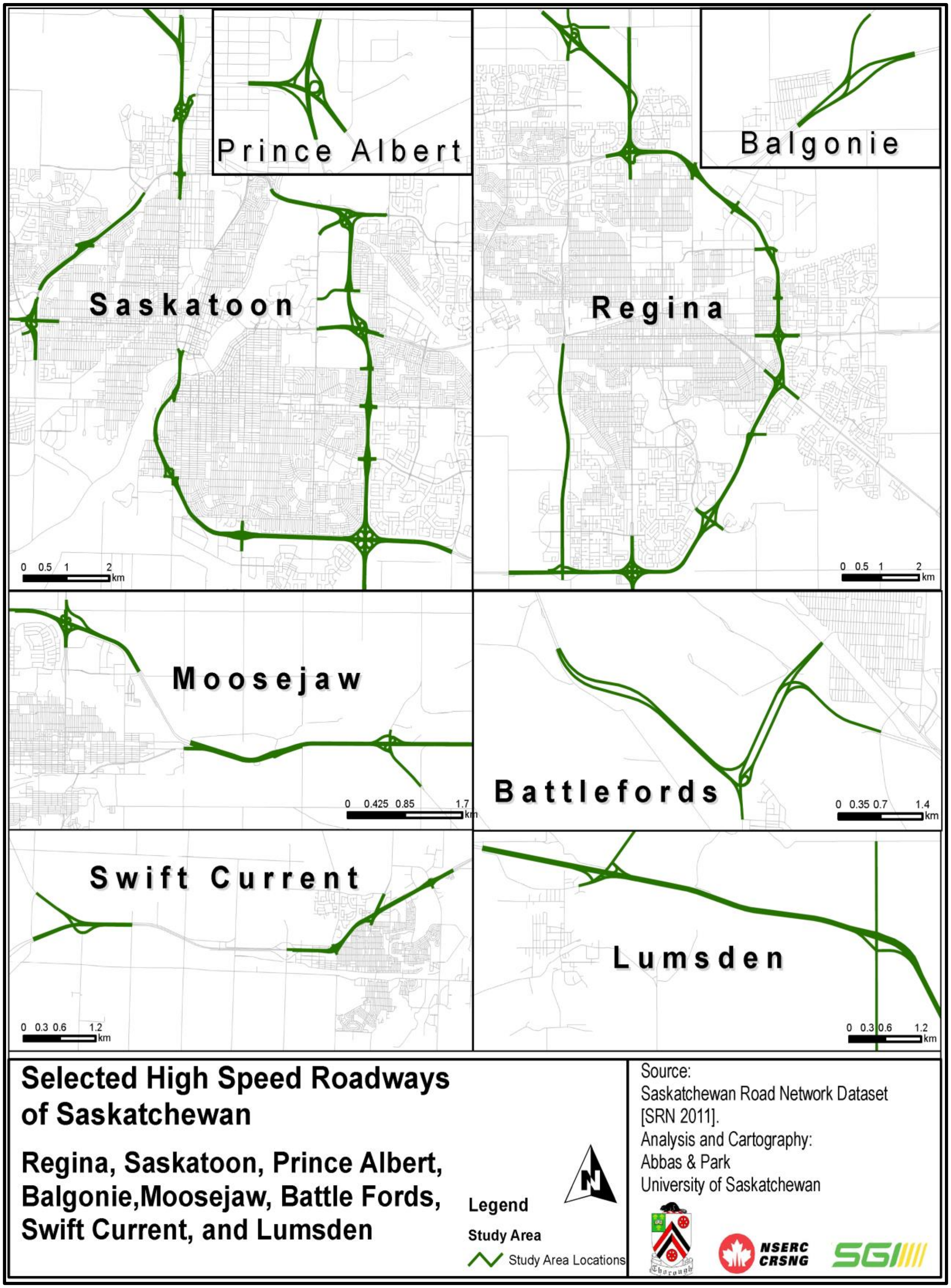


Figure A1: Study Area Map

Appendix B

Sample Calculation for Network Screening

(Expected EPDO and Excess EPDO)

SAMPLE CALCULATION FOR NETWORK SCREENING METHODS

EPDO Average Collision Frequency with EB Adjustment

A sample calculation is provided for an on-ramp segment. The selected roadway is **RID 585514** an on ramp at an interchange on Ring Road and Winnipeg Street, Regina

Step 1: Calculate the Predicted Average Collision Frequency using an Appropriate SPF

Step 1.1: Calibrate SPF

$$C_r = \frac{\sum_{all\ sites} observed\ collisions}{\sum_{all\ sites} predicted\ collisions}$$

$$N_{predicted,y} = C_r \times N_y$$

Where:

Cr = Calibration factor;

Npredicted,y = Predicted number of collisions for severity, y; and,

Ny = Uncalibrated predicted number of collisions for severity, y.

$C_{r\ (total)} = \frac{122}{118.133}$; $C_{r\ (total)} = 1.033$

$C_{r\ (FI)} = \frac{30}{29.730}$; $C_{r\ (FI)} = 1.009$

$C_{r\ (PDO)} = \frac{92}{88.613}$; $C_{r\ (PDO)} = 1.038$

Step 1.2: Determine and Apply the SPF for the site Type (on ramp)

$$N_{(total)} = C_{r\ (total)} \times \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$$

$$N'_{(FI)} = C_{r\ (FI)} \times \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$$

$$N'_{(PDO)} = C_{r\ (PDO)} \times \alpha \times \left(\frac{AADT}{1000}\right)^{\beta}$$

$$N_{(FI)} = N_{(total)} \times \left(\frac{N'_{(FI)}}{N'_{(FI)} + N'_{(PDO)}}\right)$$

$$N_{(PDO)} = N_{(total)} - N_{(FI)}$$

Where:

$N_{(total)}$ = Predicted total collisions;

$N'(FI)$ = Fatal and injury component of the total collisions;

$N'(PDO)$ = PDO component of the total collisions;

$N_{(FI)}$ = Predicted fatal and injury collisions;

$N_{(PDO)}$ = Predicted PDO collisions; and,

α *and* β= regression coefficients.

$$N_{(total)} = 1.033 \times 1.11e^{-01} \times \left(\frac{4400}{1000}\right)^{0.742} \qquad ; \qquad N_{(total)} = \mathbf{0.345}$$

$$N'_{(FI)} = 1.009 \times 2.78e^{-02} \times \left(\frac{4400}{1000}\right)^{0.747} \quad ; \quad N'_{(FI)} = 0.085$$

$$N'_{(PDO)} = 1.038 \times 8.21e^{-02} \times \left(\frac{4400}{1000}\right)^{0.753} \quad ; \quad N'_{(PDO)} = 0.260$$

$$N_{(FI)} = 0.345 \times \left(\frac{0.085}{0.085+0.260}\right) \quad ; \quad N_{(FI)} = \mathbf{0.085}$$

$$N_{(PDO)} = 0.345 - 0.085 \quad ; \quad N_{(PDO)} = \mathbf{0.260}$$

Step 2: Calculate Annual Correction Factor n

Annual correction factors are then applied to the SPFs to account for annual changes in traffic volumes, since traffic volumes may change form one year to other.

$$C_{n\,(total)} = \frac{N_{predicted,n(total)}}{N_{predicted,1(total)}} \quad \text{and} \quad C_{n\,(fi)} = \frac{N_{predicted,n(FI)}}{N_{predicted,1(FI)}}$$

Where:

$C_{n(total)}$ = Annual correction factor for total collisions;

$C_{n(FI)}$ = Annual correction factor for fatal and injury collisions;

$N_{predicted,n(total)}$ = Predicted number of total collisions for year, n;

$N_{predicted,1(total)}$ = Predicted number of total collisions for year 1;

$N_{predicted,n(FI)}$ = Predicted number of fatal and injury collisions for year, n; and,

$N_{predicted,1(FI)}$ = Predicted number of fatal and injury collisions for year 1.

$$C_{5\,(total)} = \frac{0.406}{0.345} \quad ; \quad C_{5\,(total)} = 1.18$$

$$C_{5\,(FI)} = \frac{0.100}{0.085} \quad ; \quad C_{5\,(FI)} = 1.18$$

Note: Calculate the annual correction factors for all the years in a study period and sum them ($\sum_{n=1}^{j} C_{n(total\ and\ FI)}$).

<u>Step 3: Calculate Weighted Adjustment</u>

The EB weight factor is then required to be calculated for each location. The weight factor is dependent on the overdispersion parameter, study period, and the predicted collisions obtained as a result of calibrated SPFs. An increase in any if this variable will cause a decrease in the weight factor.

$$W_y = \frac{1}{1 + k \times \left(\sum_{study\ period} N_{predicted}\right)}$$

Where:

W_y = Empirical Bayes weight for severity, y;

K = Overdispersion parameter from the appropriate SPF; and,

$N_{predicted,y}$ = Predicted average collision frequency for severity type, y.

$$W_{total} = \frac{1}{1+3.134 \times (0.345+0.345+0.345+0.376+0.406)} \quad ; \quad W_{total} = 0.510$$

$$W_{FI} = \frac{1}{1+2.146 \times (0.085+0.085+0.085+0.092+0.100)} \quad ; \quad W_{FI} = 0.149$$

Step 4: Calculate First Year EB-Adjusted Expected Average Collision Frequency

The EB-adjusted expected average collision frequency is obtained by applying the weight factor to the predicted and observed collision frequencies. As the weight factor increase, more emphasis will be placed on the predicted collisions frequency obtained from the SPFs. Based on the definition of the weight factor, shorter study periods and fewer predicted number of collisions.

$$N_{expected,n(total)} = W_{total} \times N_{predicted,n(total)} + (1 - W_{total}) \times \left(\frac{\sum_{n=1}^{j} N_{observed,n(total)}}{\sum_{n=1}^{j} C_{n(total)}} \right)$$

$$N_{expected,n(FI)} = W_{FI} \times N_{predicted,n(FI)} + (1 - W_{FI}) \times \left(\frac{\sum_{n=1}^{j} N_{observed,n(FI)}}{\sum_{n=1}^{j} C_{n(FI)}} \right)$$

Where:

Nexpected,n,(total)	= EB-adjusted expected total average collision frequency for year, n;
Npredicted,n(total)	= Calibrated predicted total average collision frequency from SPF;
Nobserved,n(total)	= Observed number of total collisions for year, n;
w(total)	= Weight factor for total collisions;
Cn(total)	= Annual correction factor for total collisions;
Nexpected,n,(FI)	= EB-adjusted expected fatal and injury average collision frequency for year, n;

$N_{predicted,n(FI)}$ = Calibrated predicted fatal and injury average collision frequency from SPF;

$N_{observed,n(FI)}$ = Observed number of fatal and injury collisions for year, n;

$w_{(FI)}$ = Weight factor for fatal and injury collisions;

$C_{n(FI)}$ = Annual correction factor for fatal and injury collisions; and,

j = Number of years in the study.

$$N_{expected,1(total)} = 0.149 \times 0.345 + (1 - 0.149) \times \left(\frac{2}{1.00+1.00+1.00+1.09+1.18}\right)$$

$$N_{expected,1(total)} = 0.37$$

$$N_{expected,1(FI)} = 0.510 \times 0.085 + (1 - 0.510) \times \left(\frac{0}{1.00 + 1.00 + 1.00 + 1.09 + 1.18}\right)$$

$$N_{expected,1(FI)} = 0.04$$

Step 5: Calculate Final Year EB-Adjusted Average Collision Frequency

The ranking of location is based on the most recent year of the study period. The final year's expected collision frequency is calculated by multiplying the SPF predicted collision frequency by the annual correction factor for the final year of the study period.

$$N_{expected,n(total)} = N_{expected,1(total)} \times C_{n(total)}$$

$$N_{expected,n(FI)} = N_{expected,1(FI)} \times C_{n(FI)}$$

$$N_{expected,n(PDO)} = N_{expected,n(total)} - N_{expected,n(FI)}$$

Where:

Nexpected,n(total)	= EB-adjusted expected total average collision frequency for final year, n;
Nexpected,1(total)	= EB-adjusted expected total average collision frequency for year 1;
Nexpected,n(FI)	= EB-adjusted expected fatal and injury average collision frequency for final year, n;
Nexpected,1(FI)	= EB-adjusted expected fatal and injury average collision frequency for year 1;
Nexpected,n(PDO)	= EB-adjusted expected PDO average collision frequency for final year, n; and,
Cn	= Annual correction factor for year, n

$N_{expected,5(total)} = 0.37 \times 1.18$; $N_{expected,5(total)} = 0.441$

$N_{expected,5(FI)} = 0.04 \times 1.18$; $N_{expected,5(FI)} = 0.047$

$N_{expected,5(PDO)} = 0.441 – 0.047$; $N_{expected,n(PDO)} = 0.394$

**small discrepancies in the numbers are due to rounding*

Step 6: Calculate Weighting Factors for Collision Severity

This step accounts for the severity of the collision based on the societal collision cost for a particular collision severity relative to a PDO collision.

$$f_{y(weight)} = \frac{CC_y}{CC_{pdo}}$$

Where:

$f_{y(weight)}$ = EPDO weighting factor based on collision severity, y;

CC_y = Collision cost for each severity, y; and,

CC_{PDO} = Collision cost for PDO collision severity.

This study used the societal costs provided by SGI for year 2010:

Severity	**Societal Cost**
Fatal	$5,543,800.00
Injury	$134,600.00
PDO	$10,900

$$f_{fatal(weight)} = \frac{\$\ 5{,}543{,}800}{10{,}900} \quad ; \quad f_{fatal(weight)} = 508.606$$

$$f_{injury(weight)} = \frac{134{,}600}{10{,}900} \quad ; \quad f_{yinjur(weight)} = 12.349$$

$$f_{PDO(weight)} = \frac{10{,}900}{10{,}900} \quad ; \quad f_{PDO(weight)} = 1.000$$

Step 7: Calculate the Proportion of Fatal and Injury Collisions

Since the predicted fatal and injury collisions are combined into a single SPF, the weight factor that is applied to calculate an EPDO score must be relative to the proportion of the observed fatal and injury collisions.

$$P_F = \frac{\sum N_{observed,(F)}}{\sum N_{observed,(FI)}}$$

$$P_I = \frac{\sum N_{observed,(I)}}{\sum N_{observed,(FI)}}$$

Where:

P_F	= Proportion of observed number of fatal collisions out of FI collisions;
P_I	= Proportion of observed number of injury collisions out of FI collisions;
$N_{observed,(F)}$	= Observed number of fatal collisions;
$N_{observed,(I)}$	= Observed number of injury collisions; and,
$N_{observed,(FI)}$	= Observed number of fatal and injury collisions.

$P_F = \frac{0}{30}$; $P_F = 0.00$

$P_I = \frac{30}{30}$; $P_I = 1.00$

Step 8: Calculate the Weight of Fatal and Injury Collisions

The EPDO weight factor for fatal/injury collisions is obtained by summing the product of the proportion of fatal and injury collisions with their respective EPDO collision cost.

$$W_{EPDO,FI} = P_F \times f_{fatal(weight)} + P_I \times f_{injury(weight)}$$

Where:

WEPDO, $_{FI}$ = EPDO weight factor for fatal and injury collisions;

$f_{injuryn}(weight)$ = EPDO injury weight factor;

$f_{fatal}\,(weight)$ = EPDO fatal weight factor;

PF = Proportion of observed number of fatal collisions out of FI collisions; and,

PI = Proportion of observed number of injury collisions out of FI collisions.

$W_{EPDO,FI} = 0 \times 508.606 + 1 \times 12.349$; $W_{EPDO,FI} = 12.349$

**small discrepancies in the numbers are due to rounding.*

<u>Step 9: Calculate the Final Year EPDO Expected Average Collision Frequency</u>

The final year EPDO expected average collision frequency is calculated by summing the expected PDO collision frequency with the EPDO weighted, expected fatal/injury collisions.

$$N_{expected(EPDO)} = N_{expected,n,(PDO)} + W_{EPDO,FI} \times N_{expected,n(FI)}$$

Where:

$N_{expected,n(EPDO)}$= EPDO expected average collision frequency for year, n;

$N_{expected,n(PDO)}$ = EB-adjusted expected PDO average collision frequency for year, n;

$w_{EPDO,FI}$ = EPDO weight factor for fatal and injury collisions; and,

$N_{expected,n(FI)}$ = EB-adjusted expected fatal and injury average collision frequency for year, n.

$N_{expected(EPDO)} = 0.400 + 12.349 \times 0.047$; $N_{expected(EPDO)} = 0.980$

**small discrepancies in the numbers are due to rounding.*

Step 10: Rank Sites by EB-Adjusted EPDO Score

Sites are then ranked from highest to lowest EPDO score to identify the locations from most likely to least likely to benefit from a safety improvement.

Excess Expected Average Collision Frequency with EB Adjustment

The Excess Expected Average Collision Frequency with EB Adjustment is used to rank locations based on the difference between estimates provided by the SPFs and the EB-adjusted estimates.

This procedure is intended to identify those locations sites which experience more collisions than expected for other locations with similar characteristics (AASHTO, 2010). To calculate the excess collision frequency, the EB adjusted collision frequency must be calculated as described previously (i.e., follow steps 1-8 listed above). The procedure for calculating the excess collision frequency is listed below.

A sample calculation is provided for an on-ramp segment. The selected roadway is **RID 585514** (same as previous) an on ramp at an interchange on Ring Road and Winnipeg Street, Regina.

Step 9: Calculate the Excess Expected Average Collision Frequency

The excess collision frequency is the difference between the EB-adjusted collision frequency and the predicted collision frequency obtained from the SPF. A positive excess collision frequency indicates that a location is not performing as well as other locations with similar traffic volumes and geometric characteristics.

$$Excess_y = \left(N_{expected,n(PDO)} - N_{prdcited,n(PDO)}\right) + \left(N_{expected,n(FI)} - N_{predicted,n(FI)}\right)$$

Where:

$Excess_n$ = Excess expected collisions for year, n;

$N_{expected,n}$ = EB-adjusted expected average collision frequency for year, n; and,

$N_{predicted,n}$ = SPF predicted average collision frequency for year, n.

$Excess_5 = (0.394 - 0.306) + (0.047 - 0.100);\quad Excess_y = .035$

Step 10: Calculate EPDO Excess

The excess collisions can be converted into EPDO scores in order to account for the severity of the collisions. This is accomplished through applying a weighting factor.

$$Excess_y = \left(N_{expected,n(PDO)} - N_{prdcited,n(PDO)}\right) + \left(N_{expected,n(FI)} - N_{predicted,n(FI)}\right) \times W_{EPDO.FI}$$

Where:

$Excess_y$ = Excess expected collisions for year, n;

$N_{expected,n}$ = EB-adjusted expected average collision frequency for year, n; and,

$N_{predicted,n}$ = SPF predicted average collision frequency for year, n; and,

$w_{EPDO,FI}$ = EPDO weight factor for fatal and injury collisions.

$Excess_5 = (0.394 - 0.306) + (0.047 - 0.100) \times 12.349 \quad ; \quad Excess_5 = -0.567$

**small discrepancies in the numbers are due to rounding.*

Step 11: Rank Sites by Excess EB-Adjusted EPDO Score

Sites are then ranked from highest to lowest excess EPDO score to identify the locations from most likely to least likely to benefit from a safety improvement

Appendix C

NCHRP Report 17-45 Segmentation Scheme

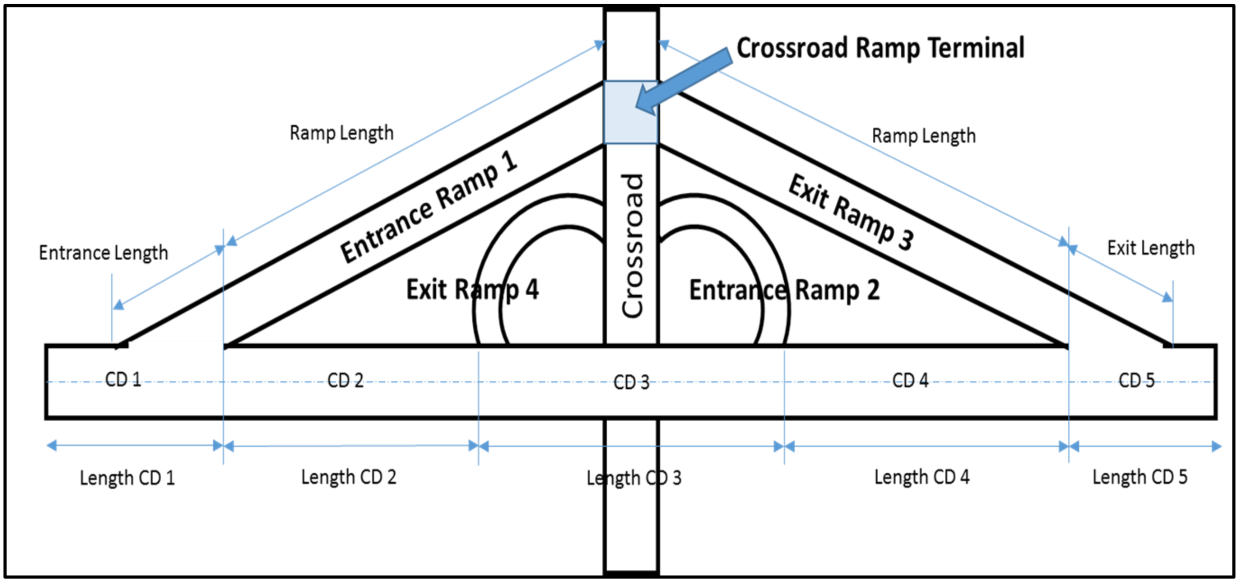


Figure C1: Segmentation Scheme for Interchanges

(One Side of Interchange with Collector Distributor (C-D) road)

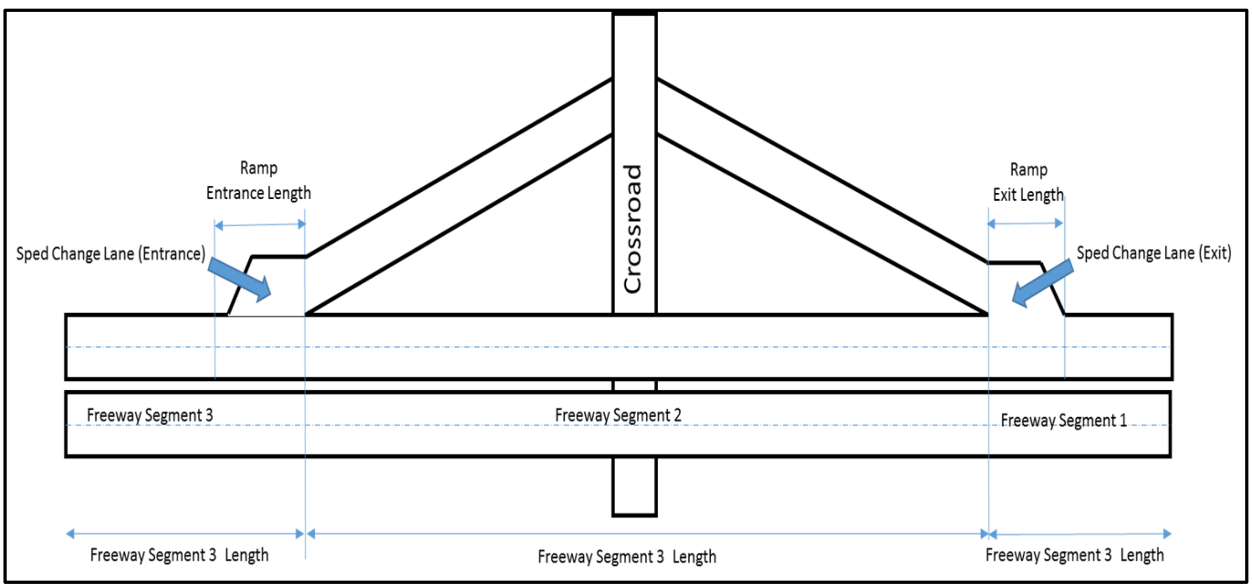


Figure C2: Segmentation Scheme for Freeway Segments and Speed Change Lanes

Appendix D

Sample of Final Database Used for Development of SPFs

Table D1: Sample of Finalized Database Used for Basic Freeway inside Interchange System

RID	MODEL	AvLength_z	AADT07	AADT08	AADT09	AADT10	AADT11	FAT7	FAT8	FAT9	FAT10	FAT11	INJ7	INJ8	INJ9	INJ10	INJ11	PDO7	PDO8	PDO9	PDO10	PDO11	AADT	length
517420	DEV	709.231	3480	5020	5020	5020	4660	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	4.64	0.709
521199	DEV	1294.275	10760	11160	11160	13080	13800	0	0	1	0	0	5	1	1	2	1	1	4	3	8	0	11.992	1.294
516336	DEV	536.676	6970	7138	7418	7343	7303	0	0	0	0	0	2	0	0	0	1	1	2	3	3	0	7.2344	0.537
523505	DEV	1429.637	9620	9910	10088	10292	10356	0	0	0	0	0	7	0	0	1	1	1	1	0	1	2	10.0532	1.430
516012	DEV	925.010	10540	10540	10540	10540	10520	0	0	0	0	0	0	0	1	1	0	5	3	1	0	0	10.536	0.925
516047	DEV	706.479	4890	4910	4910	5220	5220	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.03	0.706
523449	DEV	745.921	12440	12440	12440	12440	12440	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12.44	0.746

Table D2: Sample of Finalized Database Used for Basic Freeway outside Interchange System

RID	MODEL	AvLength_z	AADT07	AADT08	AADT09	AADT10	AADT11	FAT7	FAT8	FAT9	FAT10	FAT11	INJ7	INJ8	INJ9	INJ10	INJ11	PDO7	PDO8	PDO9	PDO10	PDO11	AADT	length
522205	DEV	2168.076	4680	4680	4680	4680	4680	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.68	2.168
517017	DEV	2917.134	4780	5440	5440	540	4720	0	0	0	0	0	0	0	0	0	0	2	1	0	5	1	4.184	2.917
519137	DEV	1192.256	5180	5830	5470	5640	5640	0	0	1	0	0	0	0	1	0	0	1	2	1	1	0	5.552	1.192
516078	DEV	1240.053	5700	5820	5820	6140	6140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.924	1.240
519684	DEV	3518.798	6230	6230	6280	6660	6840	0	0	0	0	0	0	1	1	0	0	2	4	3	1	3	6.448	3.519
501488	DEV	2443.482	6240	6400	6860	6900	6900	0	0	1	0	0	0	0	1	0	0	1	2	2	2	1	6.66	2.443
519979	DEV	4205.973	6380	6380	8300	8360	8360	0	0	0	0	1	2	1	2	4	4	2	5	2	8	6	7.556	4.206

Table D3: Sample of Finalized Database Used for Off Ramp

RID	MODEL	AADT07	AADT08	AADT09	AADT10	AADT11	FAT7	FAT8	FAT9	FAT10	FAT11	INJ7	INJ8	INJ9	INJ10	INJ11	PDO7	PDO8	PDO9	PDO10	PDO11	AADT	length
500471	DEV	20	22	25	31	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.026	0.605
519664	DEV	60	60	60	60	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.062	0.387
508152	DEV	120	1270	1270	1270	1270	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.040	0.282
520451	DEV	190	240	240	240	240	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.230	0.423
522839	DEV	200	206	206	244	244	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.220	0.375
518710	DEV	230	230	230	230	390	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.262	0.394
512773	DEV	230	230	230	230	230	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.230	0.340

Table D4: Sample of Finalized Database Used for On Ramp

RID	MODEL	AADT07	AADT08	AADT09	AADT10	AADT11	FAT7	FAT8	FAT9	FAT10	FAT11	INJ7	INJ8	INJ9	INJ10	INJ11	PDO7	PDO8	PDO9	PDO10	PDO11	AADT	length
521724	DEV	20	22	25	31	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.026	0.370
521726	DEV	35	35	35	35	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.035	1.131
506725	DEV	60	60	60	60	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.062	0.450
500373	DEV	143	143	115	72	107	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.116	0.623
518825	DEV	190	240	240	240	240	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.230	1.186
522844	DEV	200	206	206	244	244	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.220	0.213
702633	DEV	227	227	227	227	227	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.227	0.528

Table D5: Sample of Finalized Database Used for Ramp Influence Area

RID	MODEL	AADT07	AADT08	AADT09	AADT10	AADT11	FAT7	FAT8	FAT9	FAT10	FAT11	INJ7	INJ8	INJ9	INJ10	INJ11	PDO7	PDO8	PDO9	PDO10	PDO11	AADT	length	SPEEDLIMIT
519776	DEV	3860	3470	3470	3470	3470	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.548	0.450	100
790729	DEV	4780	5440	5440	5440	4720	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	5.164	0.450	110
790701	DEV	5180	5830	5470	5640	5640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.552	0.450	100
516079	DEV	5700	5820	5820	6140	6140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.924	0.450	100
522204	DEV	6230	6230	6280	6660	6840	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	6.448	0.450	100
790708	DEV	6970	7138	7418	7343	7303	0	0	0	0	0	0	0	2	0	0	0	2	0	0	0	7.234	0.450	110
790710	DEV	8020	8020	8020	8020	8020	0	0	0	0	0	0	0	1	0	0	3	2	1	1	0	8.020	0.450	100

Table D6: Sample of Finalized Database Used for Weaving Section

RID	MODEL	speed	length	AADT	ΣAAADT	ΣARAADT	TOT	FI	PDO
500126	DEV	100	0.274	38.820	62.560	23.740	32	9	23
500596	DEV	100	0.538	8.386	12.437	4.051	2	1	1
501074	DEV	100	0.722	43.080	65.260	22.180	24	5	19
507076	DEV	100	0.411	27.840	50.660	22.820	26	8	18
507213	DEV	100	0.388	9.740	14.730	4.990	3	2	1
510667	DEV	100	0.415	37.080	54.495	17.415	46	17	29
516185	DEV	100	0.085	5.819	7.619	1.800	6	1	5

Table D7: Sample of Finalized Database Used for Ramp Terminal (Signalized)

RID	MODEL	Control	AADTxrd	AADTramps	AADTtotal	AADTxrdsum	TOT	FI	PDO
159	DEV	Sig	6.700	2.250	15.650	13.400	1	0	1
160	DEV	Sig	7.320	2.400	17.040	14.640	0	0	0
113	DEV	Sig	17.468	15.200	50.135	34.935	8	2	6
114	DEV	Sig	14.247	9.190	37.685	28.495	11	3	8
136	DEV	Sig	20.385	10.850	51.620	40.770	77	17	60
137	DEV	Sig	16.945	11.450	45.340	33.890	19	8	11
103	DEV	Sig	16.900	13.160	46.960	33.800	7	3	4

Table D8: Sample of Finalized Database Used for Ramp Terminal (Unsignalized)

RID	MODEL	Control	AADTxrd	AADTramps	AADTtotal	AADTxrdsum	TOT	FI	PDO
101	DEV	Un_Sig	12.750	7.120	32.620	25.500	2	1	1
102	DEV	Un_Sig	6.600	2.900	16.100	13.200	1	0	1
111	DEV	Un_Sig	9.145	3.075	21.365	18.290	0	0	0
112	DEV	Un_Sig	8.935	4.440	22.310	17.870	6	1	5
115	DEV	Un_Sig	12.700	10.900	36.300	25.400	5	0	5
116	DEV	Un_Sig	13.700	10.960	38.360	27.400	0	0	0
118	DEV	Un_Sig	11.750	3.860	27.360	23.500	19	6	13

Appendix E

Regression Results for All Candidate Models

Table E1: Regression Results of Candidate Models for Basic Freeways inside Interchanged System

Roadway Configuration	Severity Type	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
Basic Freeway inside Interchange System	Total Collisions	3	0.019	0.009	0.998	0.022	2.590	0.004	2.242	277.650	285.056	1
		1	0.320	0.000	0.885	0.000	NA	NA	0.000	958.470	964.019	2
		2	0.304	0.336	2.106	0.001	0.955	0.025	2.151	276.280	283.682	3
		4	0.129	0.128	1.114	0.016	NA	NA	2.620	280.930	286.481	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	FI Collisions	2	0.089	0.055	2.702	0.000	0.894	0.038	1.637	176.610	184.012	1
		3	0.002	0.000	0.936	0.033	3.590	0.000	1.664	177.170	184.570	2
		1	0.106	0.083	0.732	0.099	NA	NA	2.045	180.320	185.869	3
		4	0.130	0.149	0.627	0.196	NA	NA	2.660	187.040	192.591	4
	Severity	**Model* No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	3	0.017	0.007	1.064	0.015	2.200	0.015	2.232	260.340	267.743	1
		4	0.080	0.047	1.221	0.008	NA	NA	2.483	261.830	267.383	2
		1	0.112	0.084	1.122	0.010	NA	NA	2.217	258.100	263.647	3
		2	0.187	0.179	1.874	0.003	1.018	0.018	2.137	258.910	266.306	4

Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E2: Regression Results of Candidate Models for Basic Freeways Outside Interchanged System

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
Basic Freeway outside Interchange System	Total Collisions	1	0.055	0.000	1.442	0.000	NA	NA	1.271	275.564	281.114	1
		2	0.047	0.003	1.072	0.000	1.501	0.000	1.269	277.511	284.912	2
		3	0.032	0.004	1.333	0.001	0.660	0.002	1.332	279.086	286.487	3
		4	0.658	0.651	0.593	0.092	0.000	0.000	1.701	284.932	290.482	4
	Severity	**Model* No.**	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	**Rank**
	FI Collisions	4	0.137	0.018	0.742	0.019	0.000	0.000	1.208	204.774	210.325	1
		3	0.016	0.000	1.284	0.000	0.462	0.021	0.970	201.659	209.059	2
		2	0.019	0.000	0.810	0.005	1.439	0.000	0.921	200.202	207.603	3
		1	0.012	0.000	1.603	0.000	NA	NA	0.942	198.600	204.150	4
	Severity	**Model* No.**	α	p-value	**β1**	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	**Rank**
	PDO Collisions	1	0.044	0.000	1.378	0.000	NA	NA	1.377	246.518	252.068	1
		2	0.028	0.001	1.197	0.000	1.537	0.000	1.362	248.171	255.572	2
		3	0.016	0.001	1.373	0.001	0.761	0.001	1.427	249.582	256.982	3
		4	0.527	0.517	0.537	0.153	0.000	0.000	1.919	256.392	261.943	4

Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E3: Regression Results of Candidate Models for Off Ramp

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
Off Ramp	Total Collisions	4	0.230	0.000	0.913	0.000	NA	NA	1.229	281.292	286.842	1
		1	0.072	0.000	0.922	0.000	0.850	0.039	1.172	282.218	289.618	2
		3	0.184	0.000	0.929	0.000	0.684	0.039	1.188	282.337	289.738	3
		2	0.113	0.000	1.005	0.000	NA	NA	1.391	285.241	290.792	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	FI Collisions	4	0.069	0.000	0.575	0.005	NA	NA	0.066	124.508	130.059	1
		1	0.017	0.000	0.563	0.009	1.249	0.000	0.000	124.609	132.010	2
		3	0.079	0.000	0.539	0.009	1.222	0.000	0.000	126.027	133.427	3
		2	0.033	0.000	0.751	0.002	NA	NA	0.741	134.937	140.487	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	4	0.162	0.000	1.011	0.000	NA	NA	1.393	254.516	260.067	1
		1	0.054	0.000	1.018	0.000	0.756	0.087	1.312	255.233	262.634	2
		3	0.124	0.000	1.025	0.000	0.611	0.088	1.328	255.308	262.708	3
		2	0.080	0.000	1.088	0.000	NA	NA	1.484	256.647	262.197	4

Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E4: Regression Results of Candidate Models for On Ramp

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
On Ramp	Total Collisions	2	0.111	0.000	0.742	0.001	NA	NA	3.134	239.355	244.906	1
		4	0.218	0.000	0.787	0.000	NA	NA	3.198	239.348	244.898	2
		3	0.153	0.000	0.756	0.001	0.481	0.324	3.062	240.371	247.771	3
		1	0.078	0.000	0.746	0.001	0.615	0.375	3.062	240.521	247.921	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	I Collisions	2	0.028	0.000	0.747	0.011	NA	NA	2.146	124.178	129.728	1
		4	0.056	0.000	0.727	0.011	NA	NA	2.217	125.090	130.641	2
		1	0.020	0.000	0.718	0.012	0.607	0.408	1.953	125.420	132.820	3
		3	0.037	0.000	0.729	0.011	0.375	0.488	2.012	125.632	133.033	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	2	0.082	0.000	0.753	0.001	NA	NA	3.049	211.293	216.843	1
		4	0.156	0.000	0.817	0.001	NA	NA	3.133	210.860	216.411	2
		3	0.116	0.000	0.780	0.001	0.540	0.279	2.994	212.171	219.572	3
		1	0.056	0.000	0.767	0.001	0.655	0.353	2.985	212.420	219.821	4

Note: The Column labelled "Model No." Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E5: Regression Results of Candidate Models for Ramp Influence Area

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	β3	p-value	Overdispersion Parameter	AIC	BIC	Rank
Ramp Influence Area	Total Collisions	6	0.057	0.000	1.073	0.000	NA	NA	NA	NA	1.245	251.966	257.517	1
		7	0.053	0.000	1.004	0.000	NA	NA	NA	NA	1.178	250.224	255.775	2
		4	0.121	0.005	1.004	0.000	NA	NA	NA	NA	1.043	246.134	251.685	3
		2	0.001	0.000	1.173	0.000	NA	NA	NA	NA	0.938	246.615	252.166	4
		3	0.129	0.008	1.073	0.000	NA	NA	NA	NA	1.111	248.045	253.595	5
		8	0.470	0.014	0.033	0.000	NA	NA	NA	NA	1.206	251.329	256.879	6
		1	0.000	0.000	1.418	0.050	-0.074	0.072	0.134	0.003	0.938	247.615	256.866	7
		5	0.000	0.004	1.454	0.000	0.039	0.052	NA	NA	1.142	250.948	258.349	8
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**β3**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	FI Collisions	3	0.104	0.008	0.699	0.013	NA	NA	NA	NA	0.871	154.081	159.632	1
		2	0.001	0.000	0.794	0.004	NA	NA	NA	NA	0.790	152.217	157.768	2
		4	0.102	0.006	0.649	0.011	NA	NA	NA	NA	0.835	153.493	159.043	3
		1	0.000	0.000	1.873	0.026	-0.101	0.025	0.111	0.019	0.665	153.777	163.028	4
		5	0.000	0.001	1.197	0.001	0.051	0.016	NA	NA	0.746	153.864	161.264	5
		7	0.044	0.000	0.654	0.014	NA	NA	NA	NA	0.966	156.692	162.242	6
		6	0.045	0.000	0.705	0.016	NA	NA	NA	NA	1.004	157.252	162.803	7
		8	0.199	0.000	0.019	0.039	NA	NA	NA	NA	1.027	158.272	163.822	8
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**β3**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	6	0.033	0.000	1.149	0.000	NA	NA	NA	NA	1.385	226.506	232.056	1
		7	0.029	0.000	1.094	0.000	NA	NA	NA	NA	1.287	244.353	229.903	2
		4	0.066	0.001	1.095	0.000	NA	NA	NA	NA	1.155	220.892	226.442	3
		2	0.001	0.000	1.250	0.000	NA	NA	NA	NA	1.217	222.276	227.826	4
		3	0.075	0.002	1.150	0.000	NA	NA	NA	NA	1.252	223.208	228.758	5
		8	0.297	0.000	0.037	0.000	NA	NA	NA	NA	1.287	224.635	230.185	6
		1	0.000	0.000	1.161	0.134	-0.060	0.166	0.142	0.003	1.024	222.242	231.493	7
		5	0.001	0.010	1.465	0.000	0.033	0.131	NA	NA	1.311	226.763	234.163	8

Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E6: Regression Results of Candidate Models for Weaving Section

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	β3	p-value	β4	p-value	Overdispersion Parameter	AIC	BIC	Rank
Weaving Section	Total Collisions	6	0.001	0.000	1.082	0.000	0.047	0.005	NA	NA	NA	NA	0.426	142.945	150.345	1
		9	0.003	0.004	0.054	0.000	0.057	0.003	NA	NA	NA	NA	0.430	142.667	150.068	2
		10	0.009	0.009	0.030	0.000	0.047	0.008	NA	NA	NA	NA	0.525	145.802	153.202	3
		7	0.090	0.017	1.093	0.000	NA	NA	NA	NA	NA	NA	0.569	146.326	151.876	4
		8	0.081	0.027	0.984	0.001	NA	NA	NA	NA	NA	NA	0.603	147.316	152.866	5
		1	0.000	0.000	1.899	0.014	-0.018	0.578	0.062	0.002	-0.357	0.676	0.321	142.543	153.644	6
		3	0.001	0.000	1.252	0.001	-0.012	0.578	NA	NA	NA	NA	0.460	145.002	152.403	7
		4	2.477	0.436	0.460	0.191	NA	NA	NA	NA	NA	NA	0.836	152.594	158.145	8
		5	3.337	0.352	0.325	0.343	NA	NA	NA	NA	NA	NA	0.870	153.216	158.766	9
		2	0.004	0.051	1.469	0.310	-0.025	0.696	-0.039	0.142	0.000	0.000	0.691	153.367	162.617	10
	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	β3	p-value	β4	p-value	Overdispersion Parameter	AIC	BIC	Rank
	FI Collisions	6	0.000	0.000	1.038	0.002	0.073	0.001	NA	NA	NA	NA	0.321	91.441	98.841	1
		9	0.000	0.000	0.053	0.001	0.089	0.000	NA	NA	NA	NA	0.279	90.763	98.163	2
		10	0.000	0.001	0.026	0.004	0.074	0.003	NA	NA	NA	NA	0.408	93.631	101.032	3
		7	0.035	0.011	0.953	0.015	NA	NA	NA	NA	NA	NA	0.664	98.879	104.429	4
		8	0.027	0.014	0.904	0.018	NA	NA	NA	NA	NA	NA	0.667	99.027	104.577	5
		1	0.000	0.000	1.664	0.064	-0.030	0.433	0.092	0.000	0.020	0.985	0.171	91.963	103.064	6
		2	0.000	0.003	3.112	0.068	-0.099	0.157	-0.026	0.285	0.000	0.000	0.391	98.016	107.266	7
		4	0.440	0.531	0.515	0.183	NA	NA	NA	NA	NA	NA	0.617	98.177	103.727	8
		3	0.000	0.000	1.011	0.042	-0.006	0.829	NA	NA	NA	NA	0.541	98.198	105.598	9
		5	0.463	0.593	0.438	0.241	NA	NA	NA	NA	NA	NA	0.631	98.514	104.065	10
	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	β3	p-value	β4	p-value	Overdispersion Parameter	AIC	BIC	Rank
	PDO Collisions	6	0.001	0.001	1.085	0.000	0.038	0.024	NA	NA	NA	NA	0.452	134.703	142.104	1
		9	0.006	0.010	0.053	0.000	0.047	0.014	NA	NA	NA	NA	0.461	134.514	141.915	2
		7	0.060	0.006	1.126	0.000	NA	NA	NA	NA	NA	NA	0.526	135.618	141.169	3
		8	0.055	0.013	1.010	0.001	NA	NA	NA	NA	NA	NA	0.568	136.818	142.369	4
		10	0.017	0.023	0.029	0.001	0.037	0.036	NA	NA	NA	NA	0.560	137.403	144.803	5
		1	0.000	0.001	1.928	0.016	-0.015	0.658	0.053	0.011	-0.415	0.645	0.342	134.538	145.639	6
		3	0.000	0.000	1.296	0.001	-0.013	0.536	NA	NA	NA	NA	0.429	134.678	142.079	7
		4	1.866	0.610	0.469	0.204	NA	NA	NA	NA	NA	NA	0.890	143.649	149.199	8
		5	2.635	0.475	0.321	0.371	NA	NA	NA	NA	NA	NA	0.930	144.282	149.832	9
		2	0.003	0.057	1.436	0.349	-0.022	0.749	-0.041	0.140	0.000	0.000	0.749	144.774	154.025	10

Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

Table E7: Regression Results of Candidate Models for Ramp Terminal (Signalized)

Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
Ramp Terminal (Signalized)	Total Collisions	2	0.000	0.001	2.454	0.000	NA	NA	1.923	194.266	199.817	1
		4	0.001	0.007	1.334	0.040	NA	NA	1.961	194.751	200.302	2
		3	0.000	0.004	0.495	0.583	2.107	0.041	1.905	195.960	203.360	3
		1	0.003	0.003	0.935	0.215	1.612	0.047	1.916	196.097	203.498	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	FI Collisions	2	0.000	0.000	2.430	0.000	NA	NA	1.157	138.489	144.040	1
		4	0.000	0.001	1.325	0.031	NA	NA	1.166	138.881	144.431	2
		3	0.000	0.000	0.303	0.707	2.221	0.013	1.163	140.351	147.752	3
		1	0.001	0.000	0.743	0.282	1.719	0.015	1.166	140.484	147.884	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	2	0.000	0.002	2.405	0.000	NA	NA	1.980	179.566	185.117	1
		4	0.001	0.007	1.292	0.054	NA	NA	2.012	179.993	185.544	2
		3	0.000	0.005	0.455	0.623	2.080	0.049	1.969	181.325	188.725	3
		1	0.003	0.003	0.891	0.251	1.592	0.056	1.976	181.443	188.844	4

Note: The Column labelled "Model No." Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

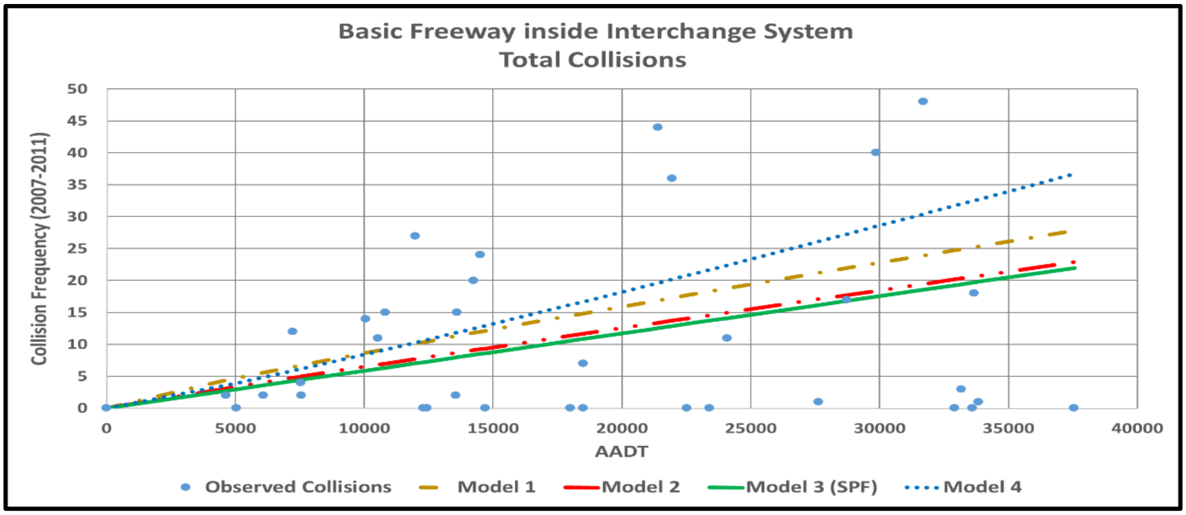
Table E8: Regression Results of Candidate Models for Ramp Terminal (Unsignalized)

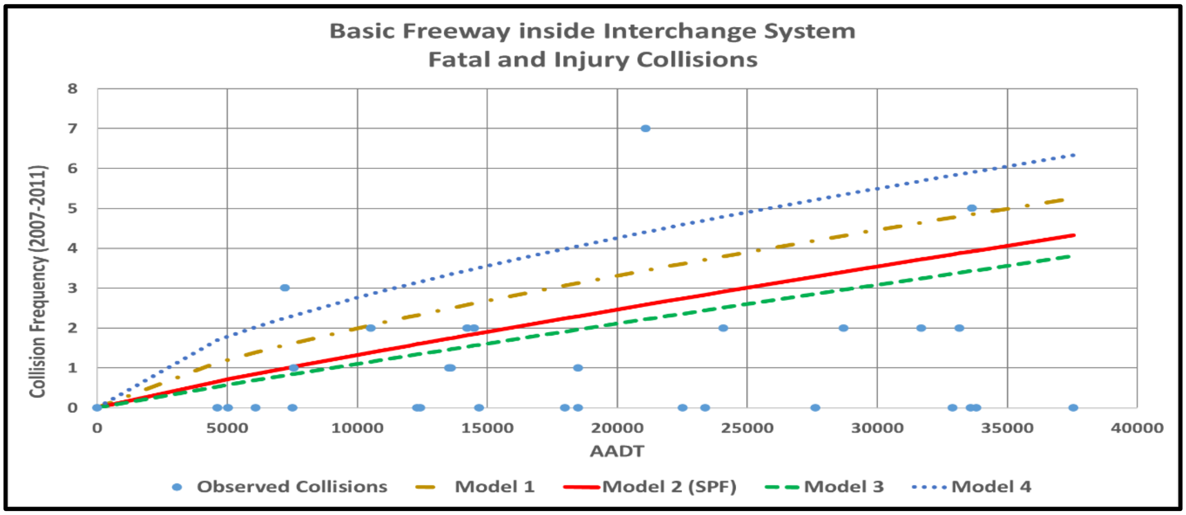
Roadway Configuration	Severity	Model No.	α	p-value	β1	p-value	β2	p-value	Overdispersion Parameter	AIC	BIC	Rank
Ramp Terminal (Unsinalized)	Total Collisions	2	0.000	0.000	2.030	0.000	NA	NA	2.950	113.980	119.527	1
		4	0.001	0.000	1.058	0.037	NA	NA	2.976	113.950	119.505	2
		1	0.004	0.000	0.497	0.263	1.568	0.003	2.770	115.400	122.801	3
		3	0.001	0.000	0.260	0.589	1.760	0.003	2.849	115.840	123.238	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	FI Collisions	2	0.001	0.001	1.408	0.017	NA	NA	3.745	69.199	74.749	1
		4	0.003	0.003	0.419	0.485	NA	NA	3.774	69.107	74.658	2
		1	0.005	0.000	0.307	0.566	1.128	0.073	3.509	70.930	78.331	3
		3	0.001	0.001	0.146	0.802	1.258	0.073	3.636	71.171	78.572	4
	Severity	**Model No.**	**α**	**p-value**	**β1**	**p-value**	**β2**	**p-value**	**Overdispersion Parameter**	**AIC**	**BIC**	**Rank**
	PDO Collisions	2	0.000	0.000	2.421	0.000	NA	NA	2.882	97.428	102.978	1
		4	0.000	0.000	1.473	0.011	NA	NA	2.941	97.512	103.063	2
		1	0.001	0.000	0.793	0.085	1.747	0.003	2.597	98.155	105.556	3
		3	0.000	0.000	0.574	0.247	1.871	0.003	2.667	98.729	106.130	4

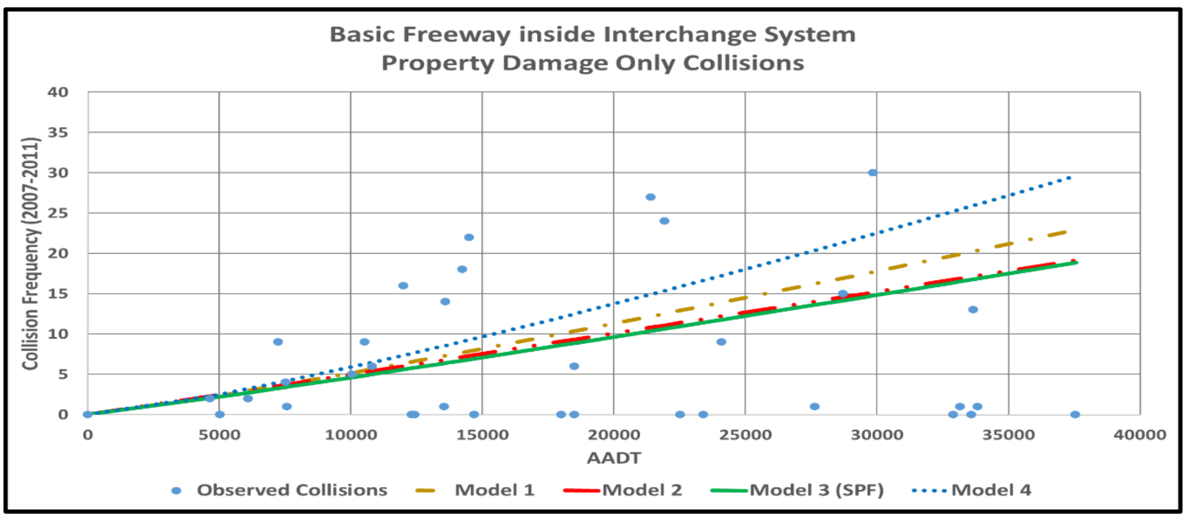
Note: The Column labelled “Model No.” Shows Models in the Order of their Appearance in the Respective Tables of Candidate Models

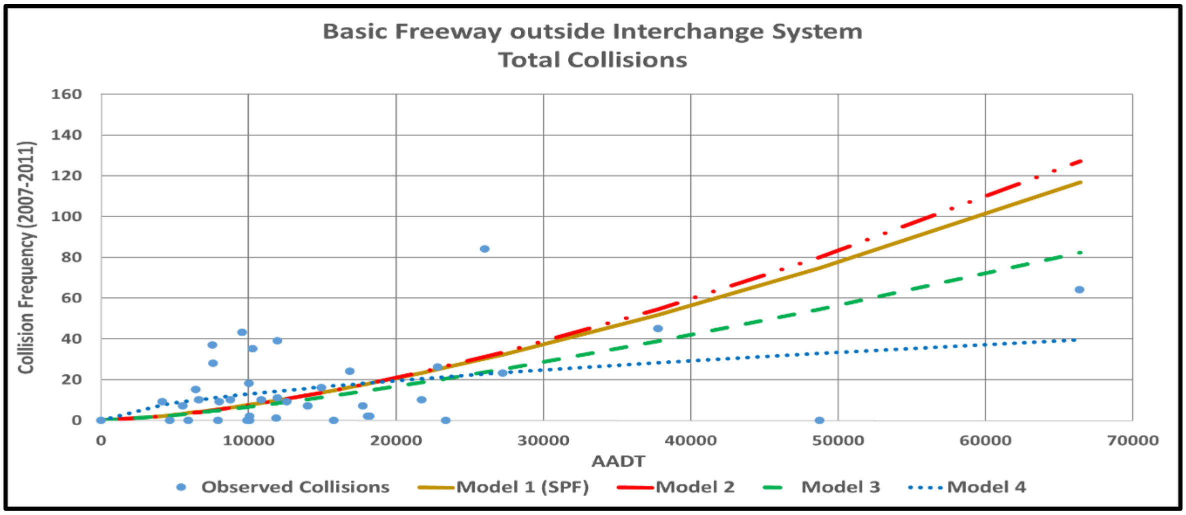
Appendix F

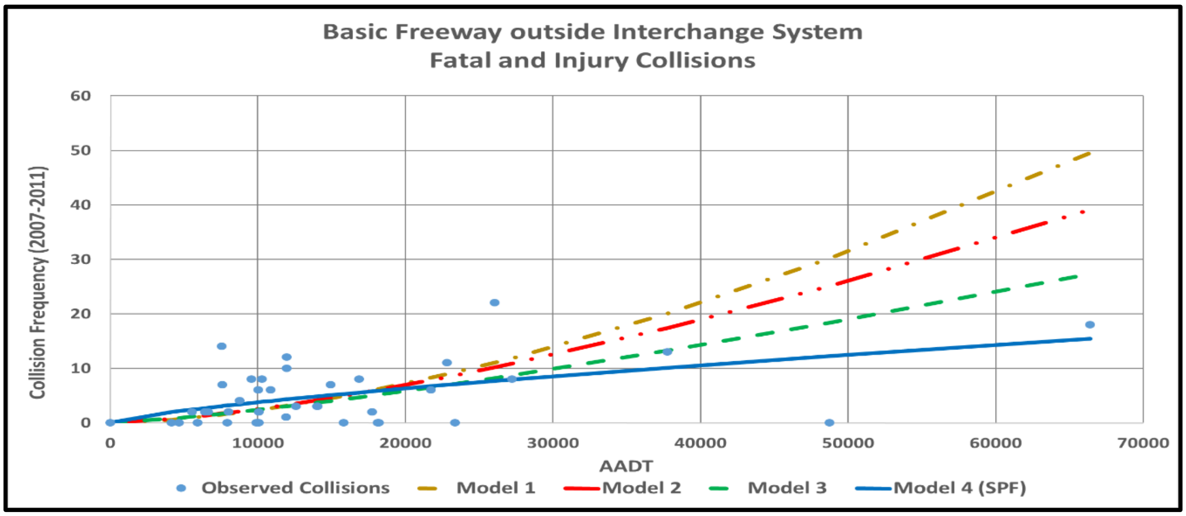
Candidate Models

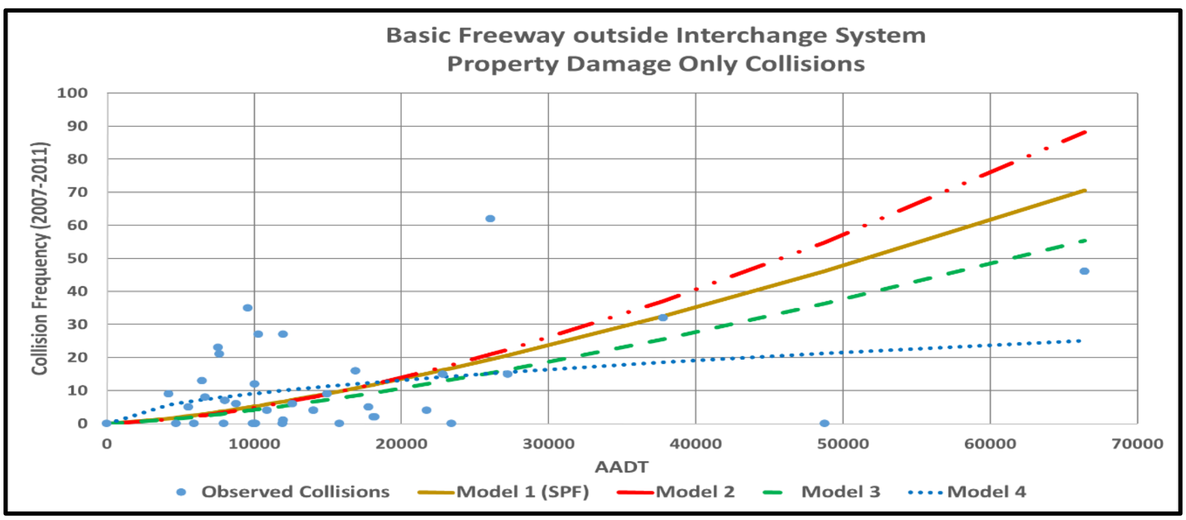

Basic Freeway inside Interchange System
Total Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 1
Model 2
Model 3 (SPF)
Model 4

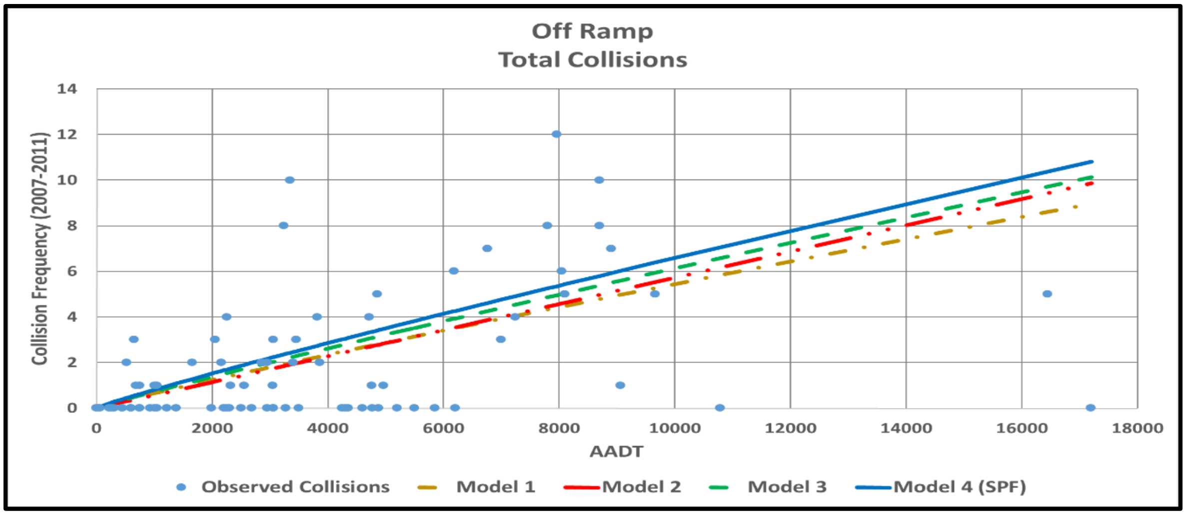

Basic Freeway inside Interchange System
Fatal and Injury Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 1
Model 2 (SPF)
Model 3
Model 4

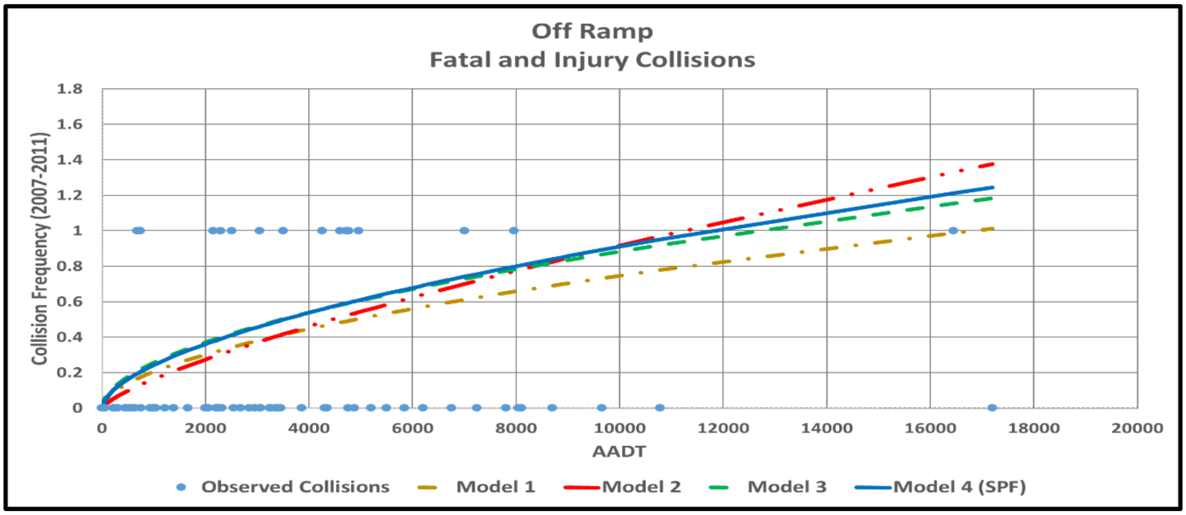

Basic Freeway inside Interchange System
Property Damage Only Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 1
Model 2
Model 3 (SPF)
Model 4

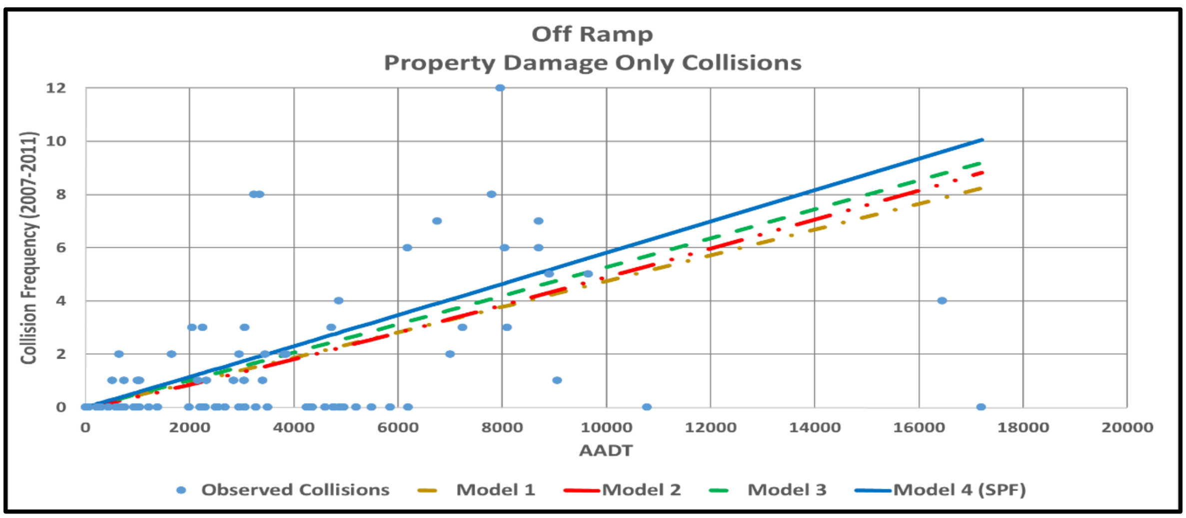

Basic Freeway outside Interchange System
Total Collisions
Collision Frequency (2007-2011)
AADT
0
20
40
60
80
100
120
140
160
0
10000
20000
30000
40000
50000
60000
70000
Observed Collisions
Model 1 (SPF)
Model 2
Model 3
Model 4

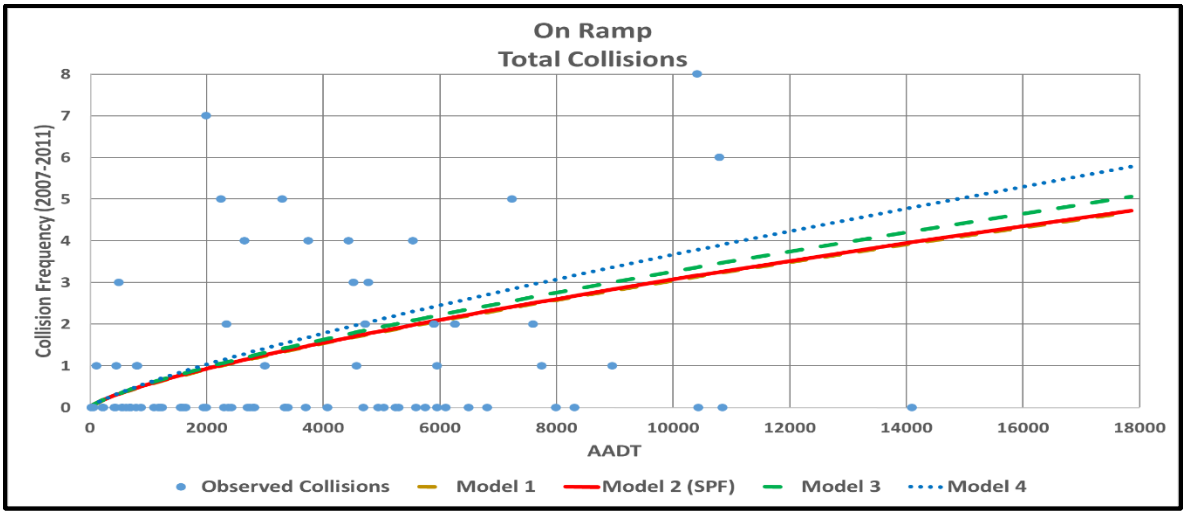

Basic Freeway outside Interchange System
Fatal and Injury Collisions
Collision Frequency (2007-2011)
AADT
0
10
20
30
40
50
60
0
10000
20000
30000
40000
50000
60000
70000
Observed Collisions
Model 1
Model 2
Model 3
Model 4 (SPF)


Basic Freeway outside Interchange System
Property Damage Only Collisions
Collision Frequency (2007-2011)
AADT
0
10
20
30
40
50
60
70
80
90
100
0
10000
20000
30000
40000
50000
60000
70000
Observed Collisions
Model 1 (SPF)
Model 2
Model 3
Model 4

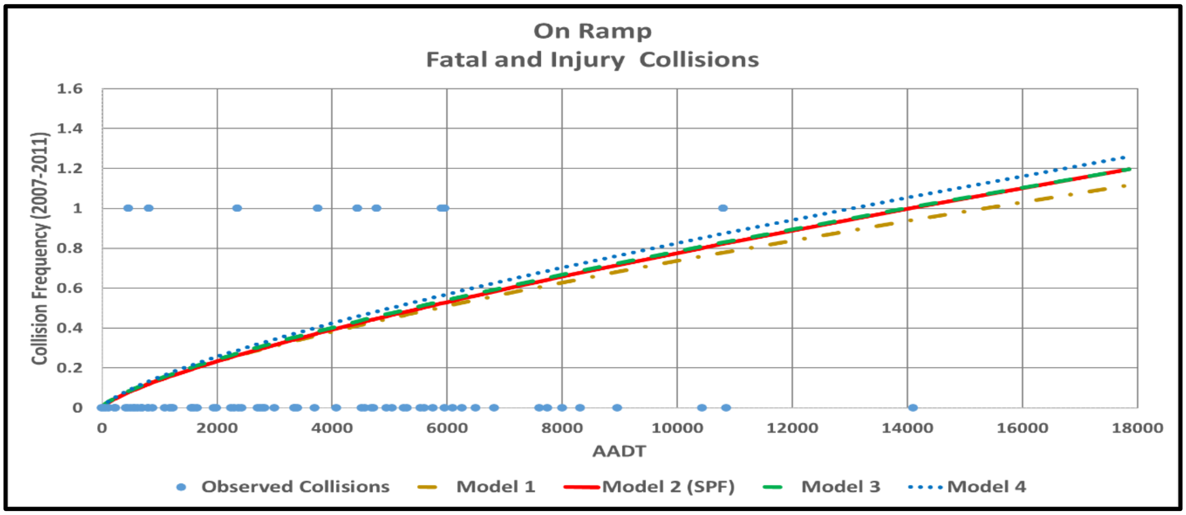

Off Ramp
Total Collisions
Collision Frequency (2007-2011)
0
2
4
6
8
10
12
14
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
AADT
Observed Collisions
Model 1
Model 2
Model 3
Model 4 (SPF)


Off Ramp
Fatal and Injury Collisions
Collision Frequency (2007-2011)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000
AADT
Observed Collisions
Model 1
Model 2
Model 3
Model 4 (SPF)

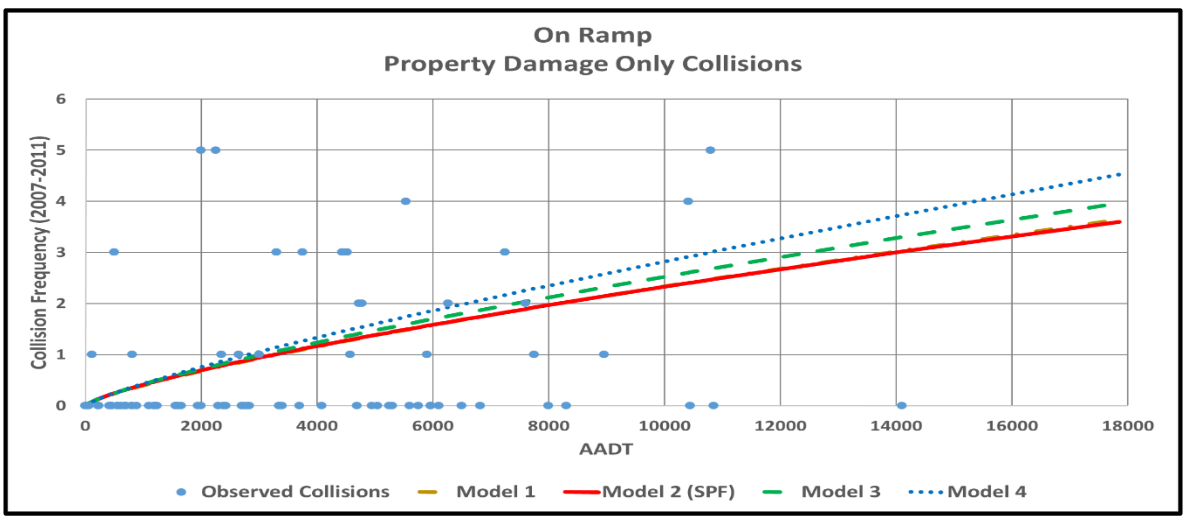

Off Ramp
Property Damage Only Collisions
Collision Frequency (2007-2011)
0
2
4
6
8
10
12
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000
AADT
Observed Collisions
Model 1
Model 2
Model 3
Model 4 (SPF)



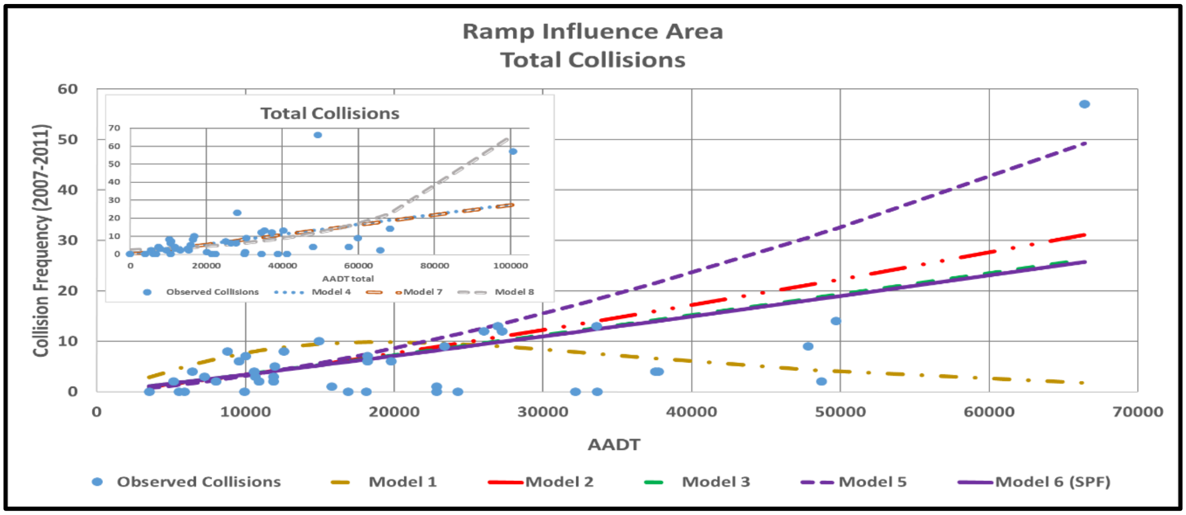
On Ramp
Total Collisions
Collision Frequency (2007-2011)
8
7
6
5
4
3
2
1
0
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
AADT
Observed Collisions
Model 1
Model 2 (SPF)
Model 3
Model 4



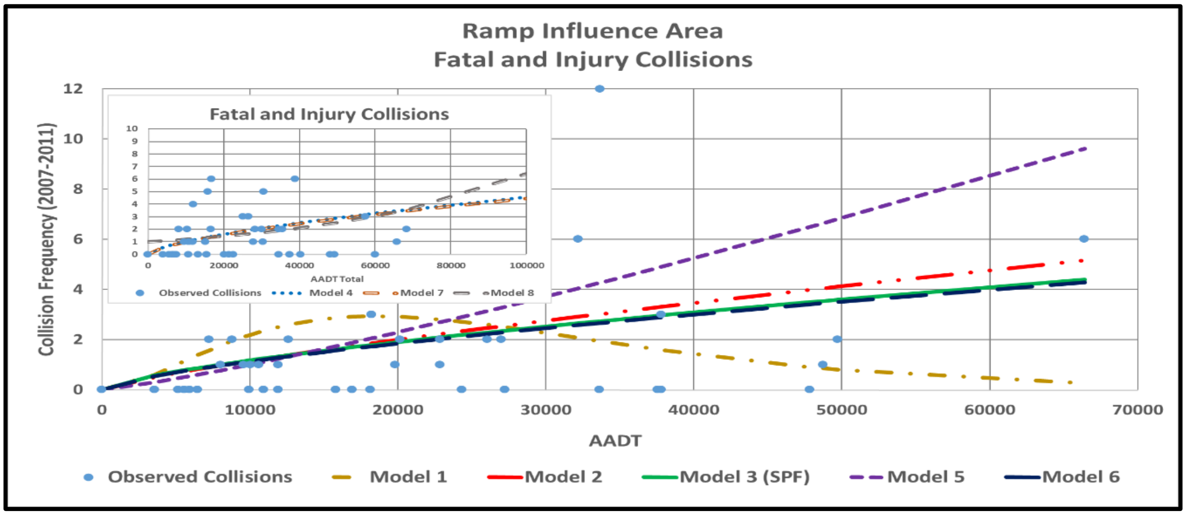
On Ramp
Fatal and Injury Collisions
Collision Frequency (2007-2011)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
AADT
Observed Collisions
Model 1
Model 2 (SPF)
Model 3
Model 4



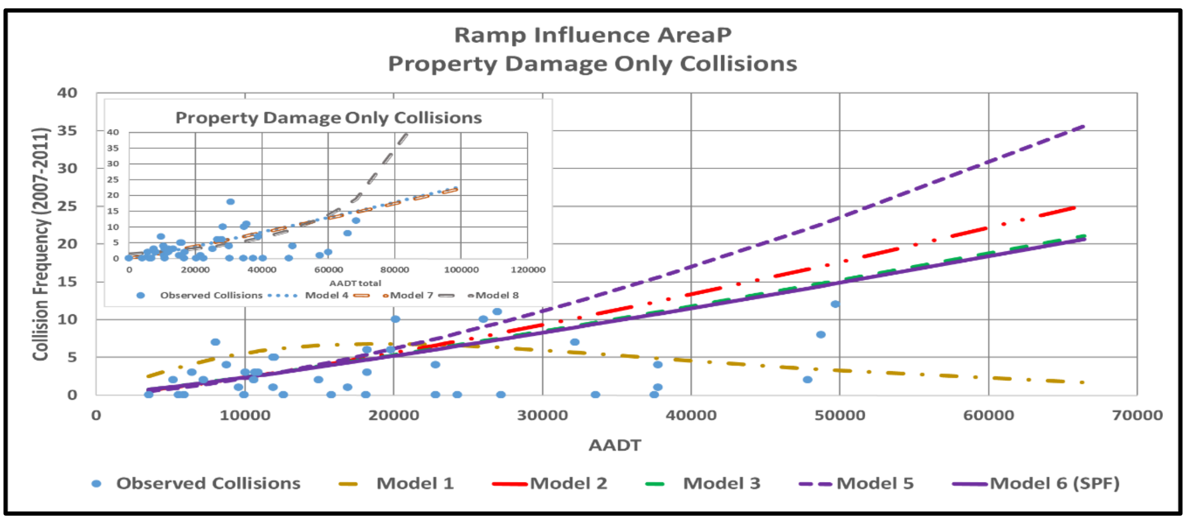
On Ramp
Property Damage Only Collisions
Collision Frequency (2007-2011)
6
5
4
3
2
1
0
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
AADT
Observed Collisions
Model 1
Model 2 (SPF)
Model 3
Model 4



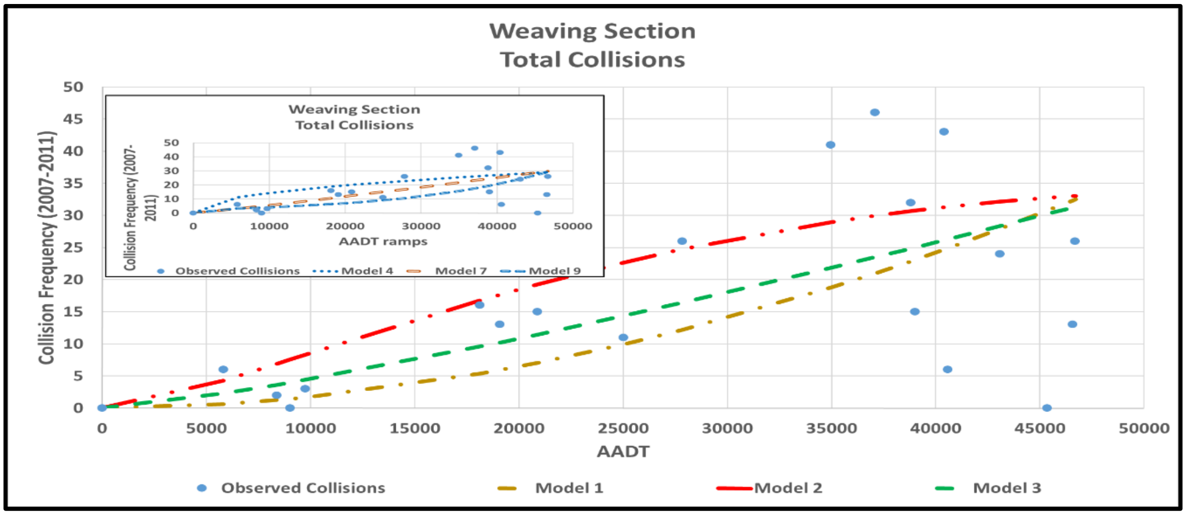
Ramp Influence Area
Total Collisions
Total Collisions
Collision Frequency (2007-2011)
AADT
AADT total
Observed Collisions
Model 1
Model 2
Model 3
Model 5
Model 6 (SPF)
Model 4
Model 7
Model 8

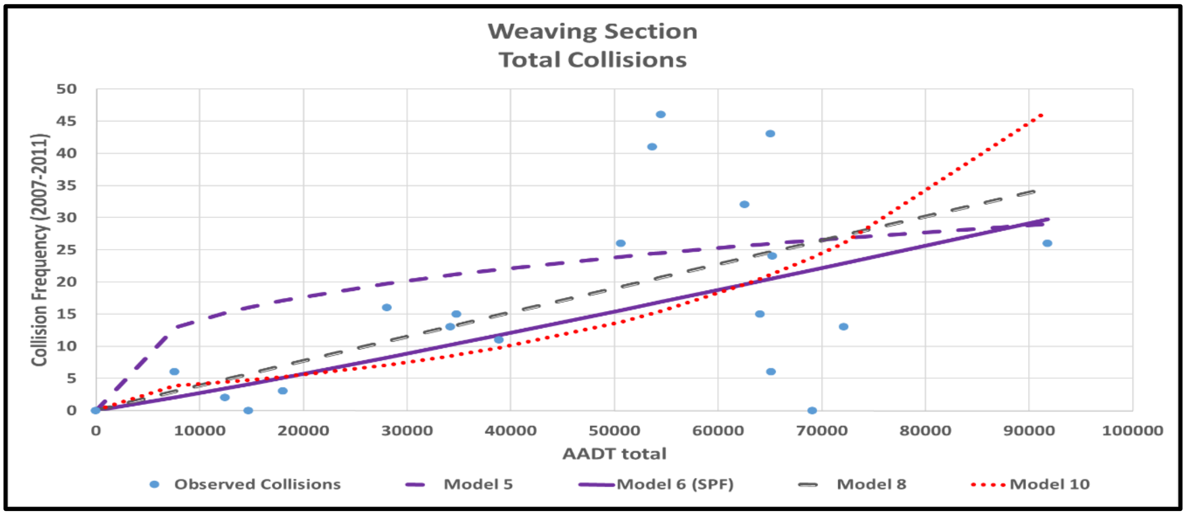


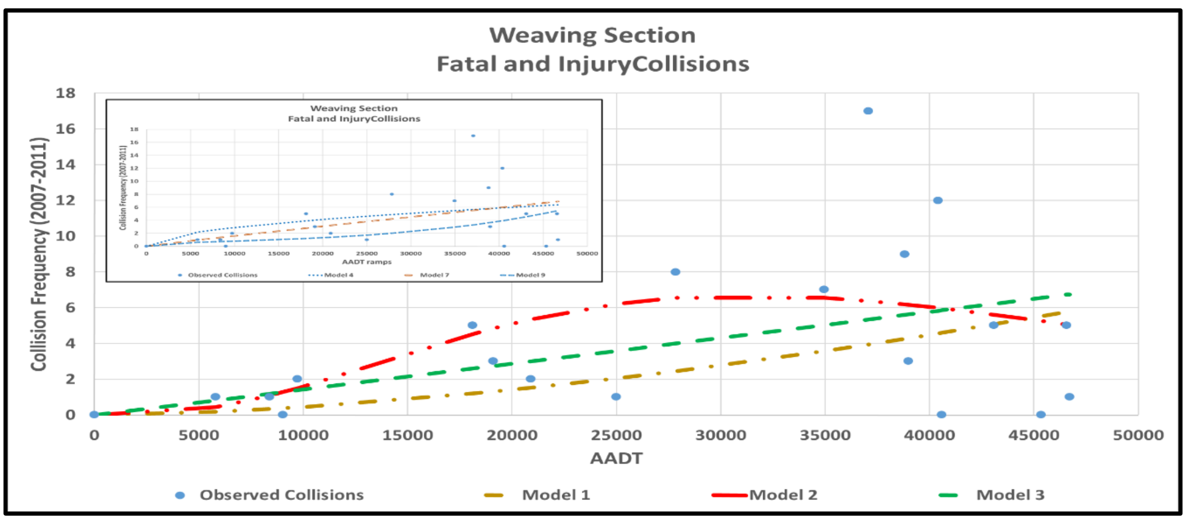
Ramp Influence Area
Fatal and Injury Collisions
Fatal and Injury Collisions
Collision Frequency (2007-2011)
AADT
AADT Total
Observed Collisions
Model 1
Model 2
Model 3 (SPF)
Model 5
Model 6
Model 4
Model 7
Model 8

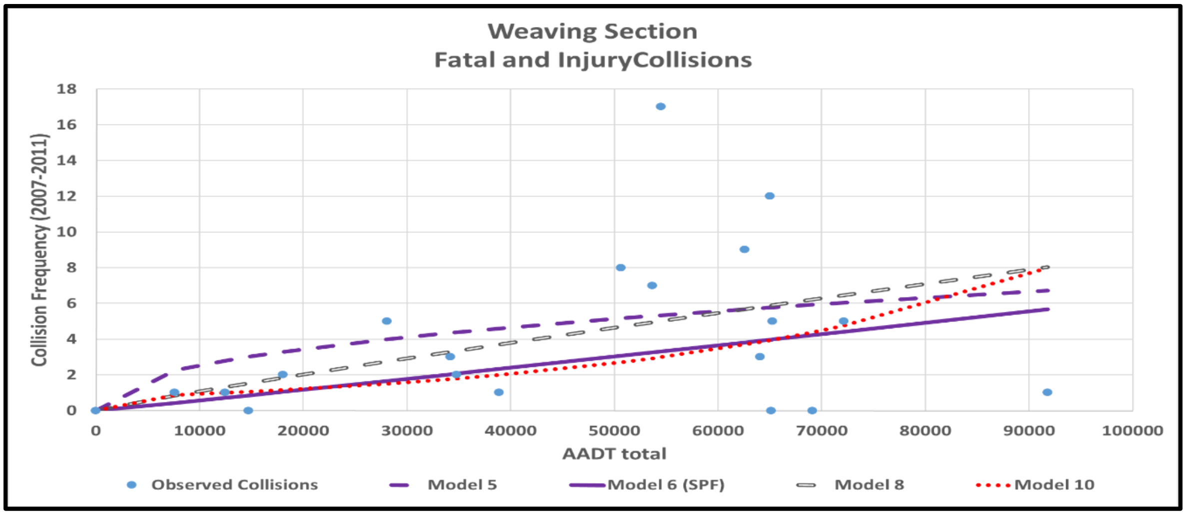


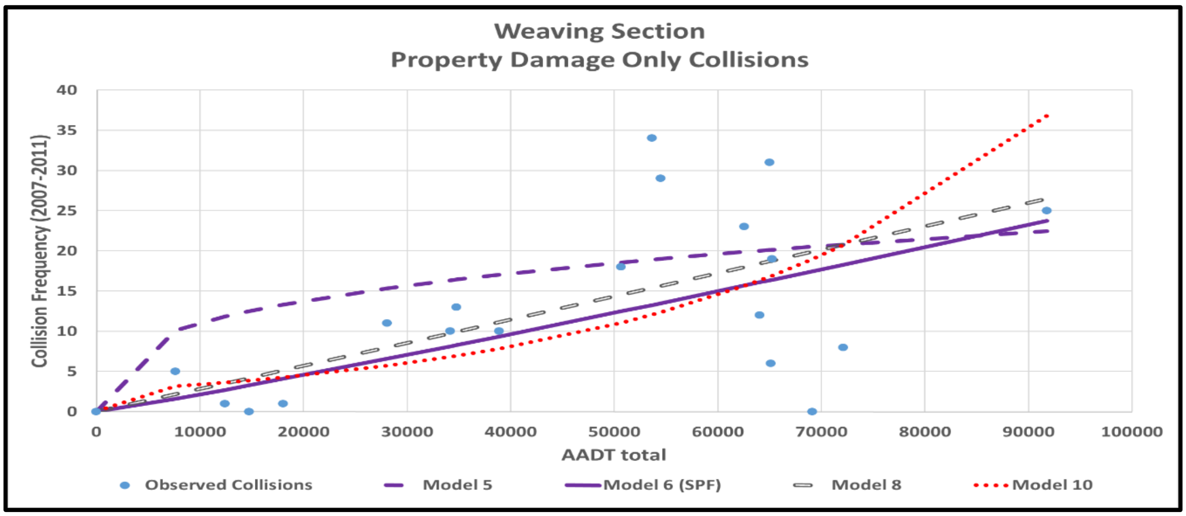
Ramp Influence AreaP
Property Damage Only Collisions
Property Damage Only Collisions
Collision Frequency (2007-2011)
AADT
AADT total
Observed Collisions
Model 1
Model 2
Model 3
Model 5
Model 6 (SPF)
Model 4
Model 7
Model 8


Weaving Section
Total Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 1
Model 2
Model 3
Weaving Section
Total Collisions
Collision Frequency (2007-2011)
AADT ramps
Observed Collisions
Model 4
Model 7
Model 9

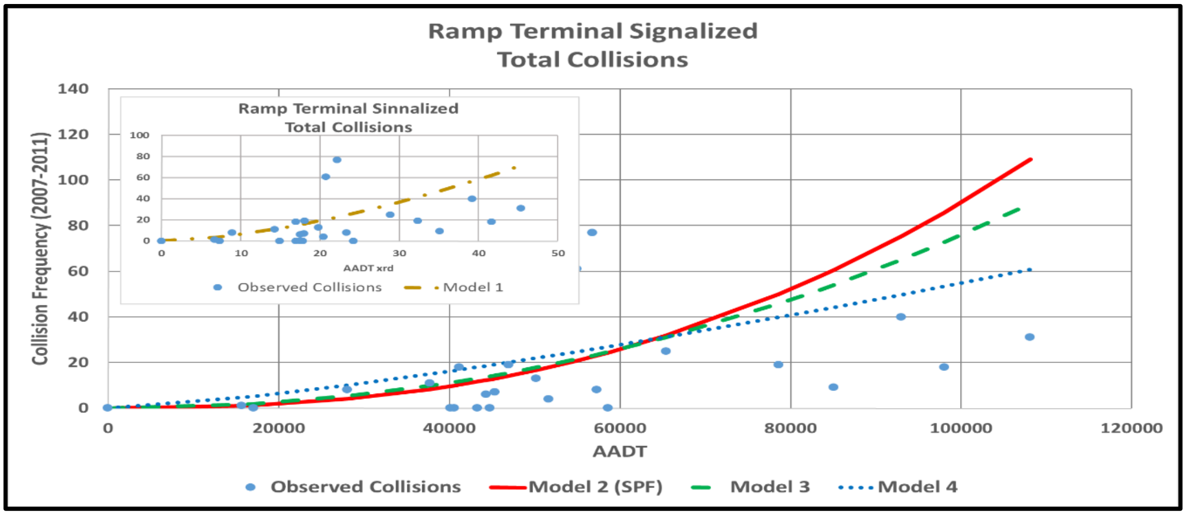

Weaving Section
Total Collisions
Collision Frequency (2007-2011)
AADT total
Observed Collisions
Model 5
Model 6 (SPF)
Model 8
Model 10


Weaving Section
Fatal and InjuryCollisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 1
Model 2
Model 3

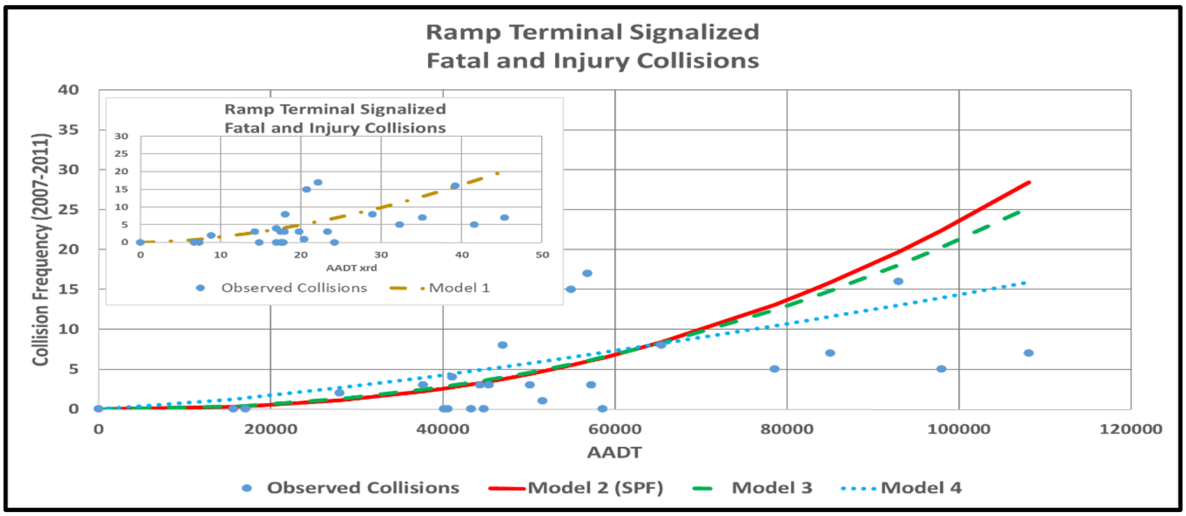

Weaving Section
Fatal and InjuryCollisions
Collision Frequency (2007-2011)
0
2
4
6
8
10
12
14
16
18
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
AADT total
Observed Collisions
Model 5
Model 6 (SPF)
Model 8
Model 10


Weaving Section
Property Damage Only Collisions
Collision Frequency (2007-2011)
0
5
10
15
20
25
30
35
40
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
AADT total
Observed Collisions
Model 5
Model 6 (SPF)
Model 8
Model 10

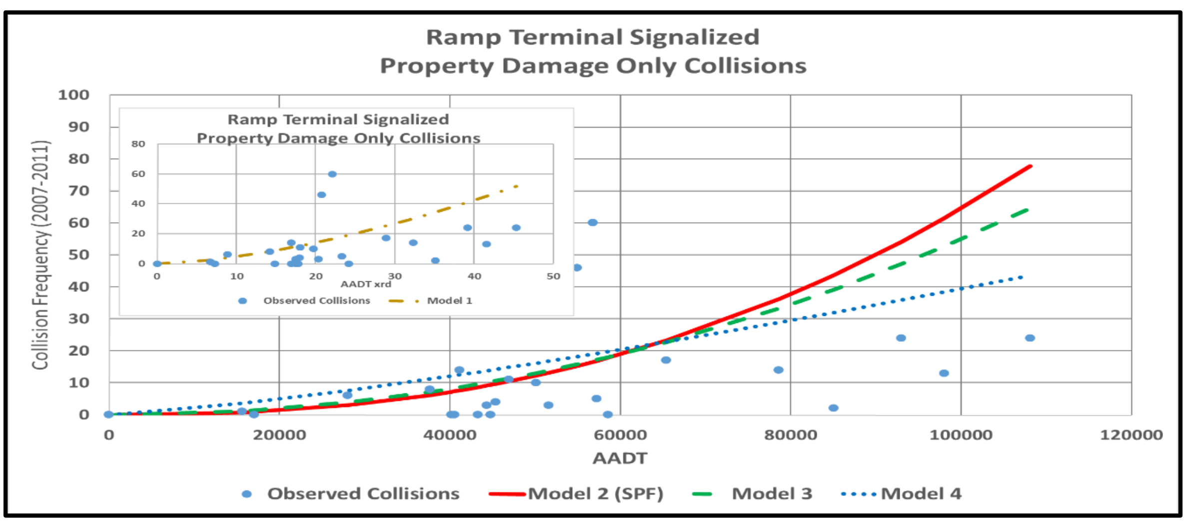

Weaving Section
Property Damage Only Collisions
Collision Frequency (2007-2011)
0
5
10
15
20
25
30
35
40
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
AADT
Observed Collisions
Model 1
Model 2
Model 3
Weaving Section
Property Damage Only Collisions
AADT ramps
Observed Collisions
Model 4
Model 7
Model 9



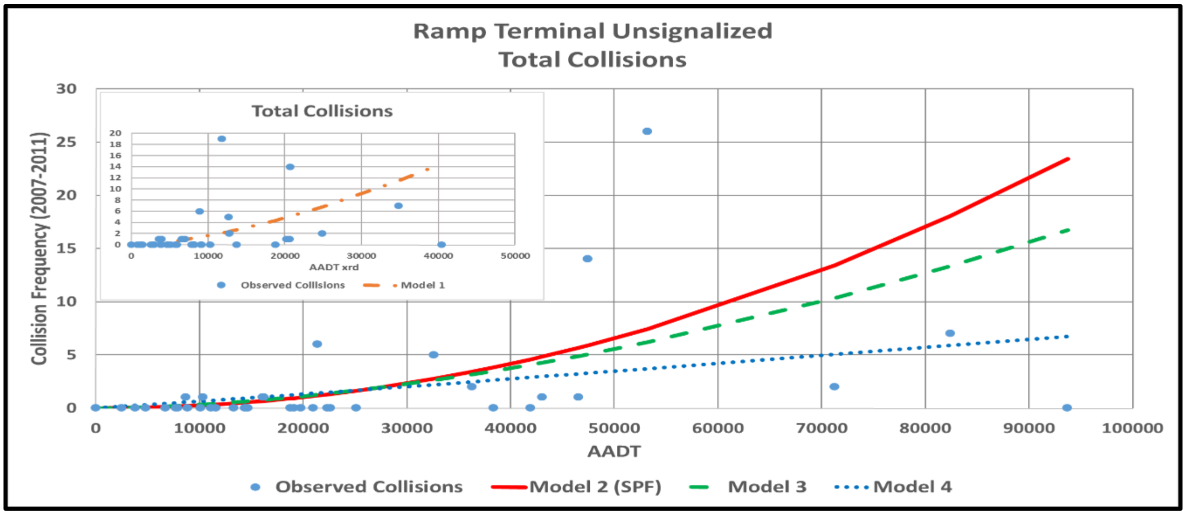
Ramp Terminal Signalized
Total Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4
Ramp Terminal Sinnalized
Total Collisions
AADT xrd
Observed Collisions
Model 1

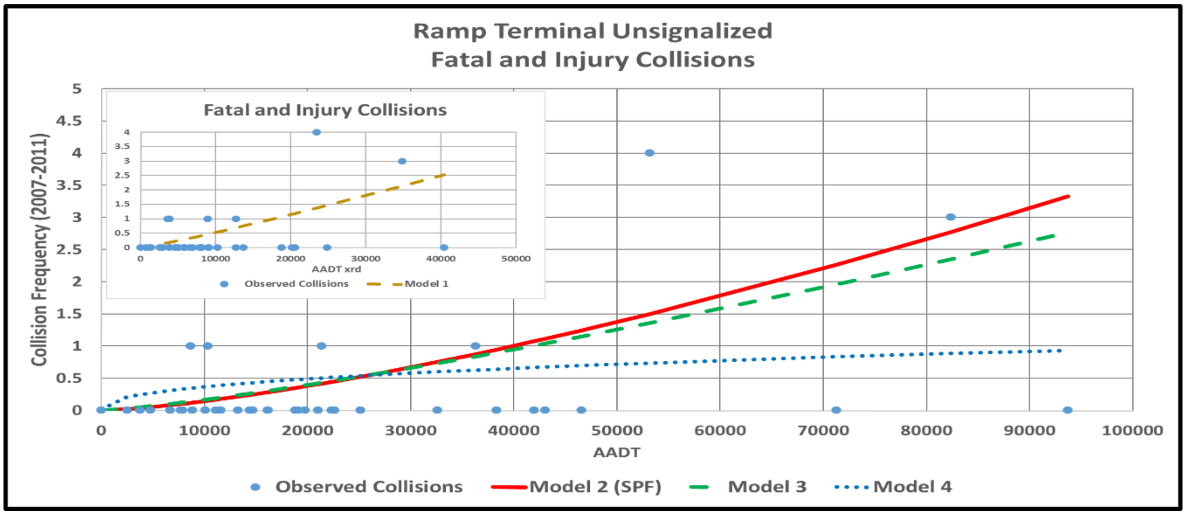


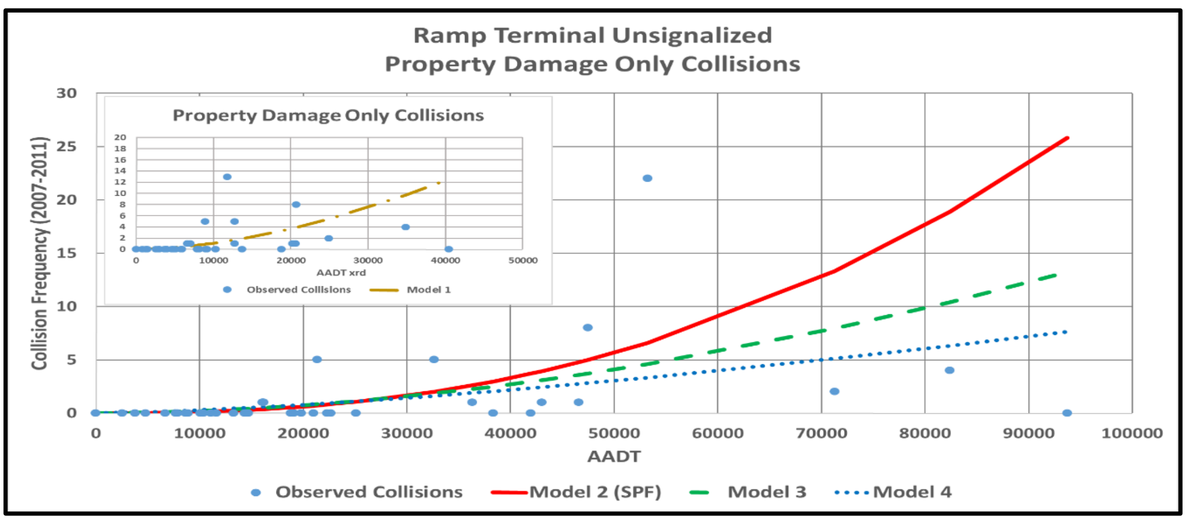
Ramp Terminal Signalized
Fatal and Injury Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4
Ramp Terminal Signalized
Fatal and Injury Collisions
AADT xrd
Observed Collisions
Model 1



Ramp Terminal Signalized
Property Damage Only Collisions
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4
Ramp Terminal Signalized
Property Damage Only Collisions
AADT xrd
Observed Collisions
Model 1


Ramp Terminal Unsignalized
Total Collisions
Total Collisions
AADT xrd
Observed Collisions
Model 1
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4


Ramp Terminal Unsignalized
Fatal and Injury Collisions
Fatal and Injury Collisions
AADT xrd
Observed Collisions
Model 1
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4

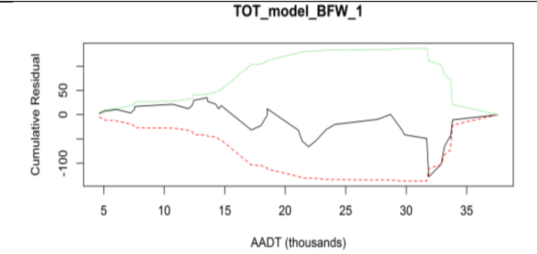

Ramp Terminal Unsignalized
Property Damage Only Collisions
Property Damage Only Collisions
AADT xrd
Observed Collisions
Model 1
Collision Frequency (2007-2011)
AADT
Observed Collisions
Model 2 (SPF)
Model 3
Model 4

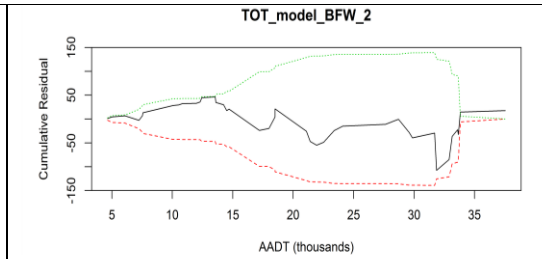
<u>Appendix G</u>

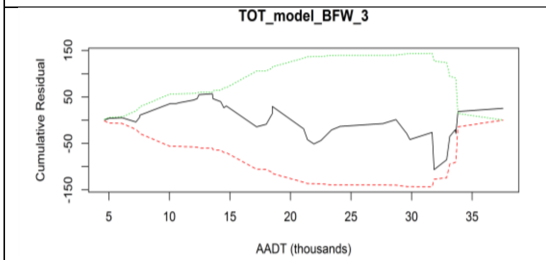
CURE Plots of All Candidate Models

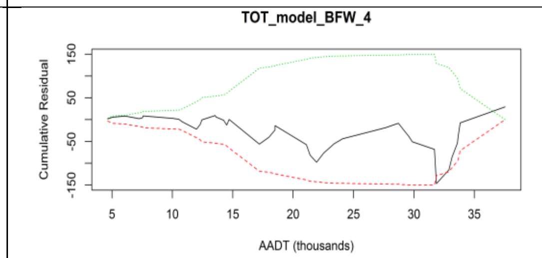
Basic Freeway inside Interchanged System

Total Collisions

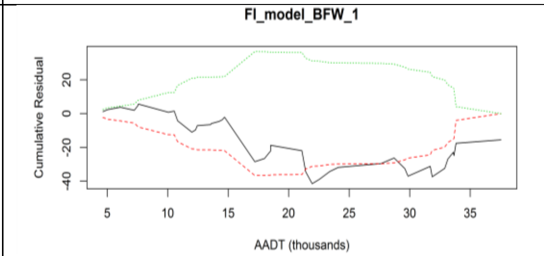


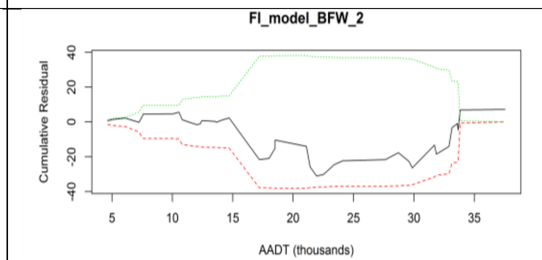


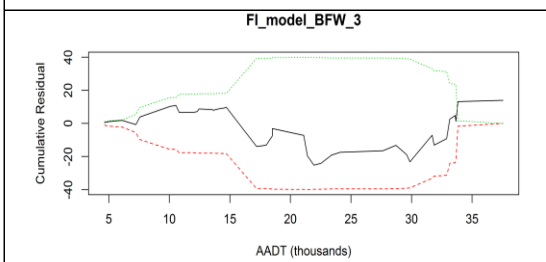


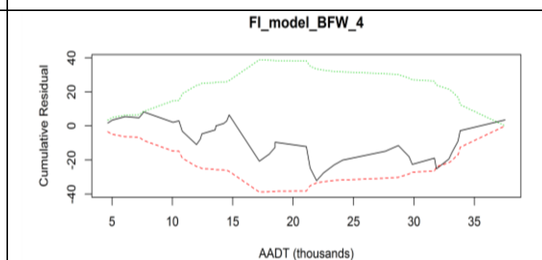


Fatal and Injury (FI) Collisions

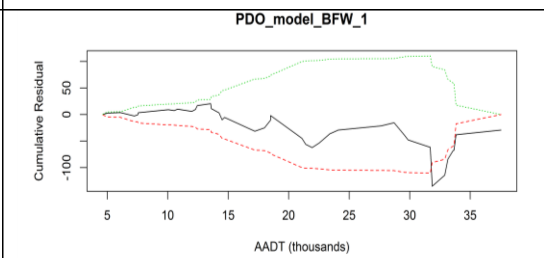


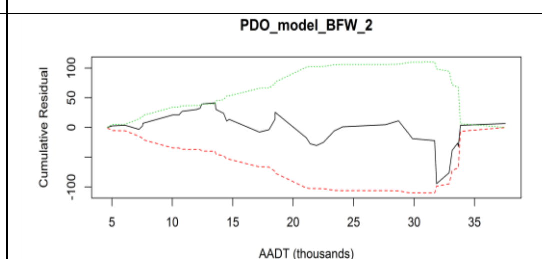


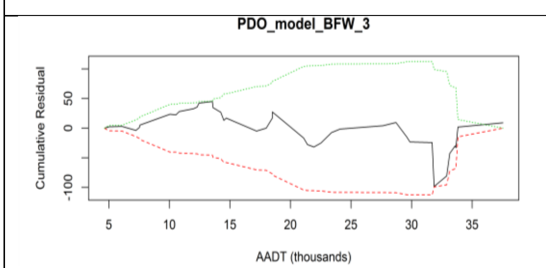


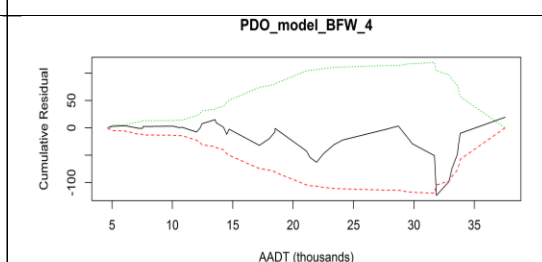


Property Damage Only (PDO) Collisions

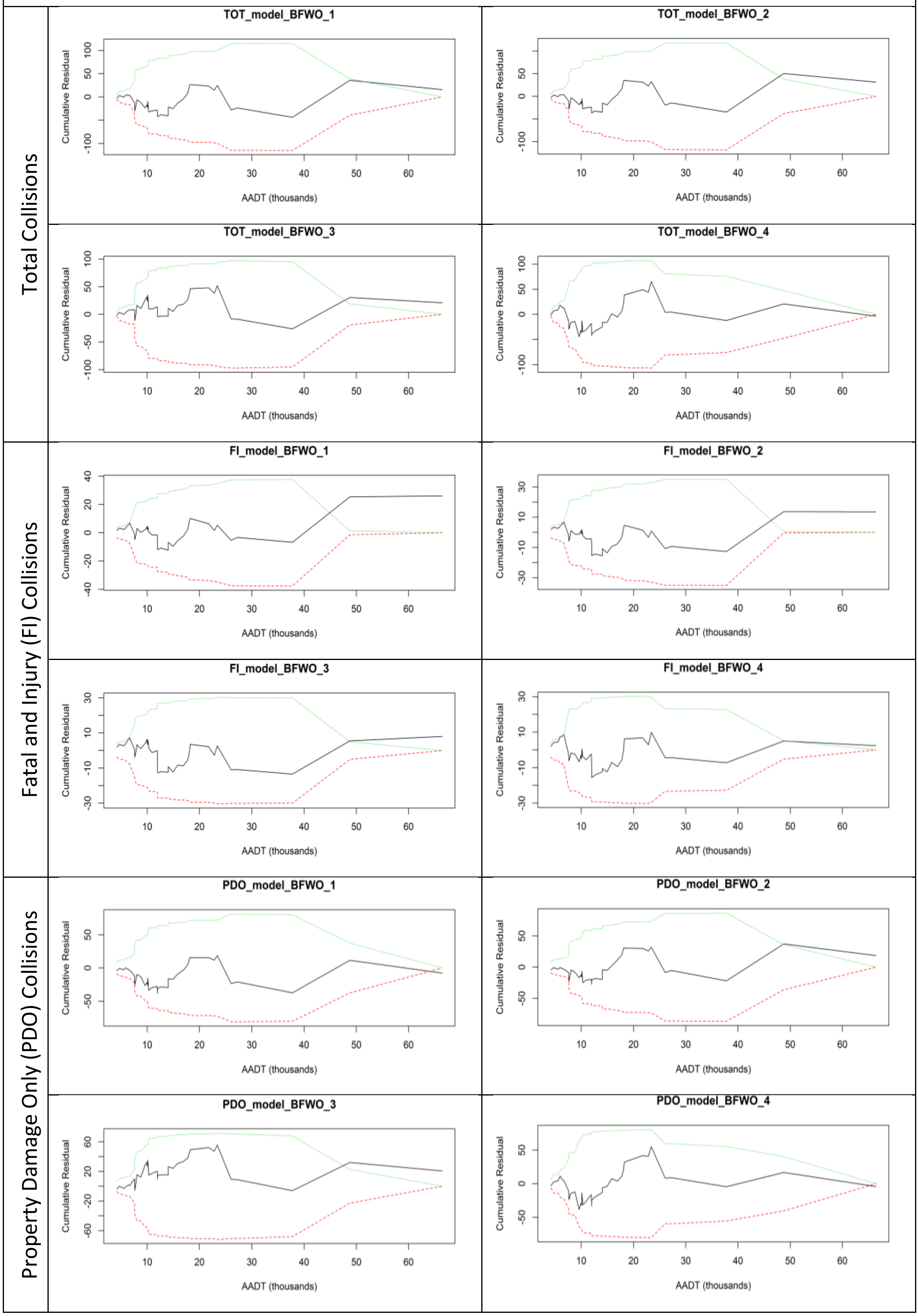








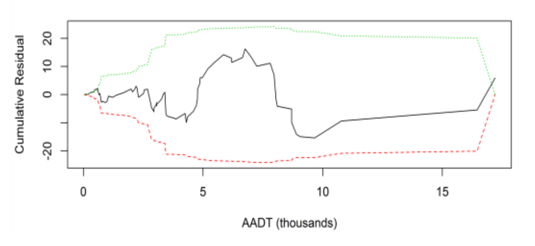
Basic Freeway Outside Interchanged System



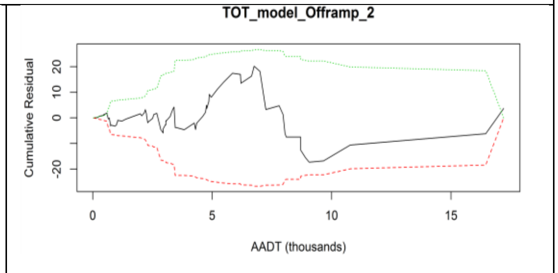
Off Ramp

Total Collisions

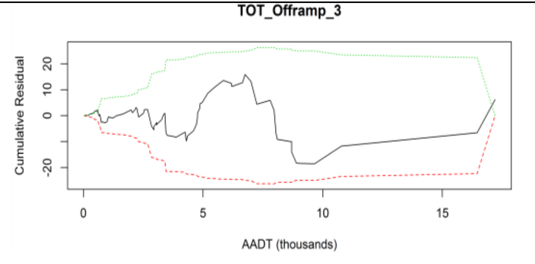
TOT_model_Offramp_1



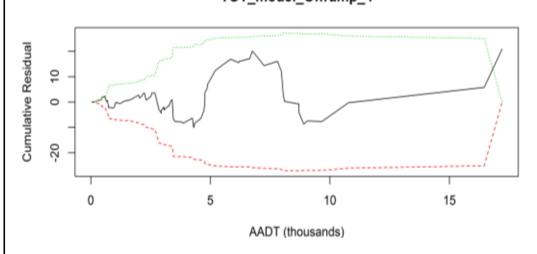
TOT_model_Offramp_2



TOT_Offramp_3

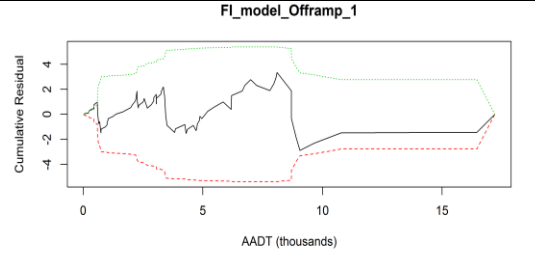


TOT_model_Offramp_4

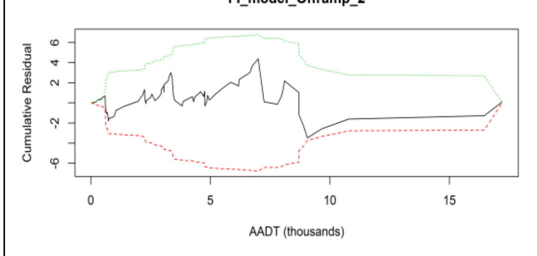


Fatal and Injury (FI) Collisions

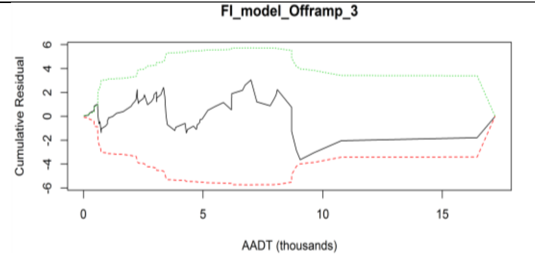
FI_model_Offramp_1



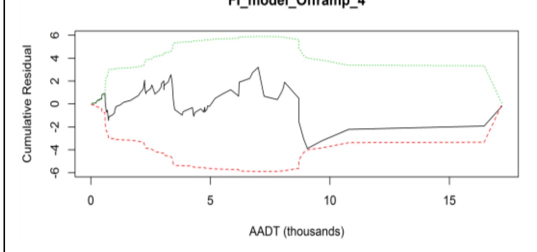
FI_model_Offramp_2



FI_model_Offramp_3

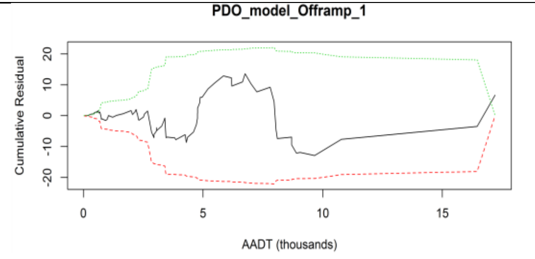


FI_model_Offramp_4

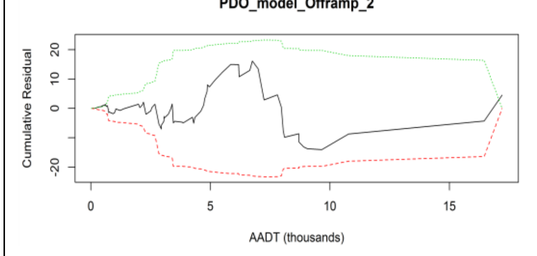


Property Damage Only (PDO) Collisions

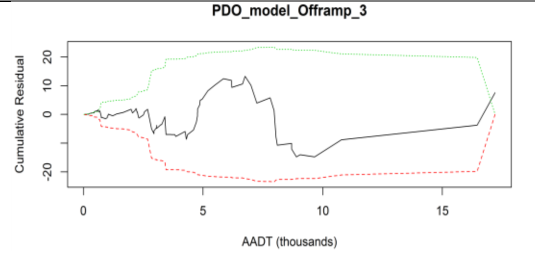
PDO_model_Offramp_1



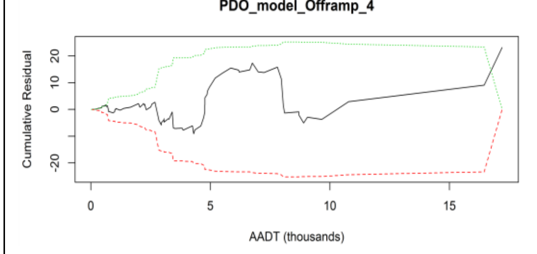
PDO_model_Offramp_2



PDO_model_Offramp_3

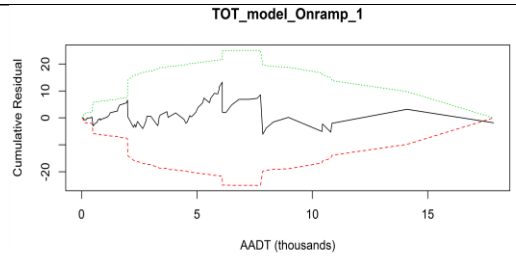


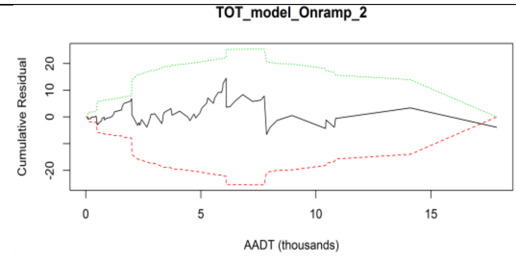
PDO_model_Offramp_4

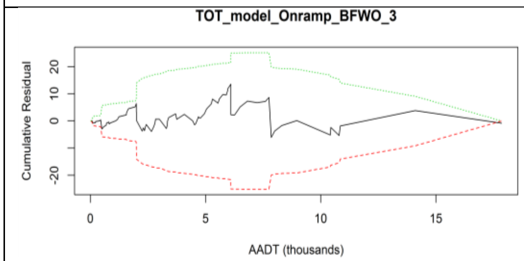


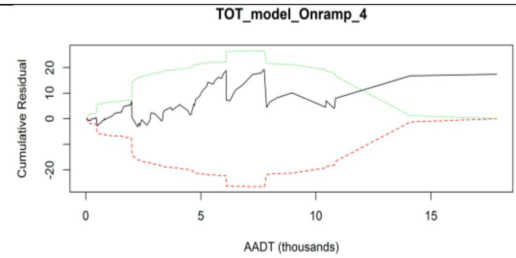
On Ramp

Total Collisions

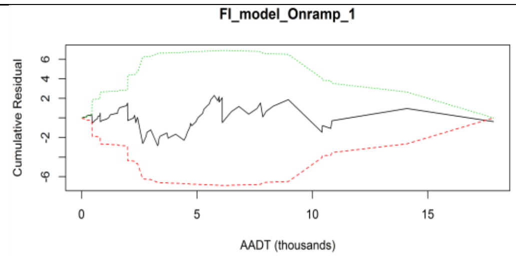


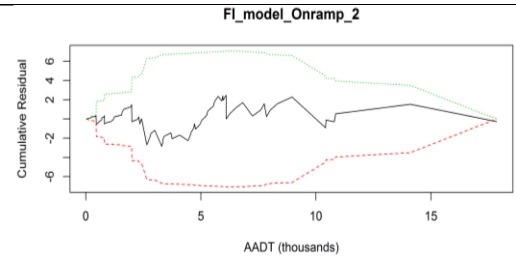


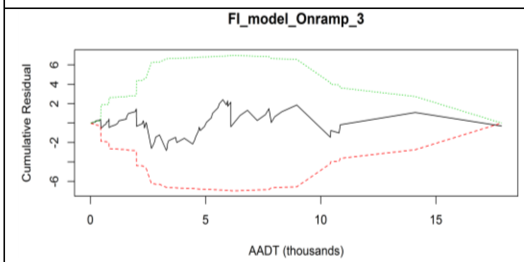


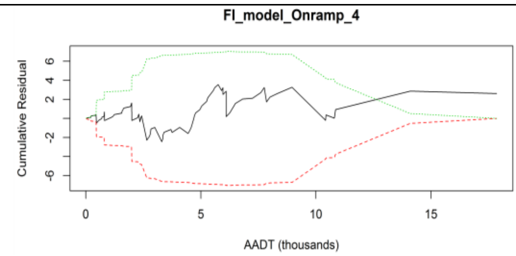


Fatal and Injury (FI) Collisions

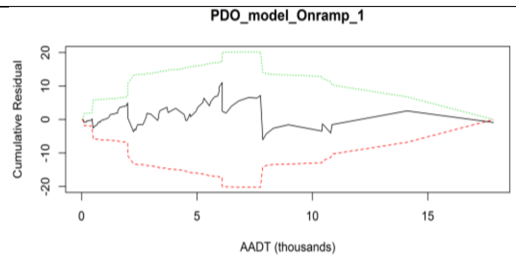


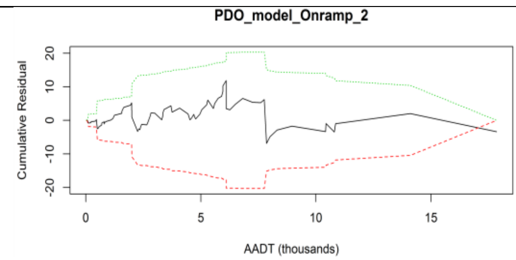


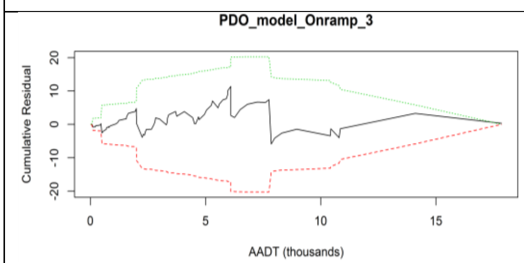


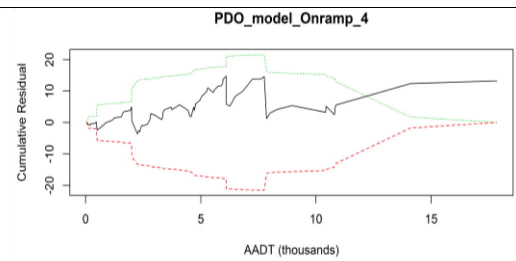


Property Damage Only (PDO) Collisions

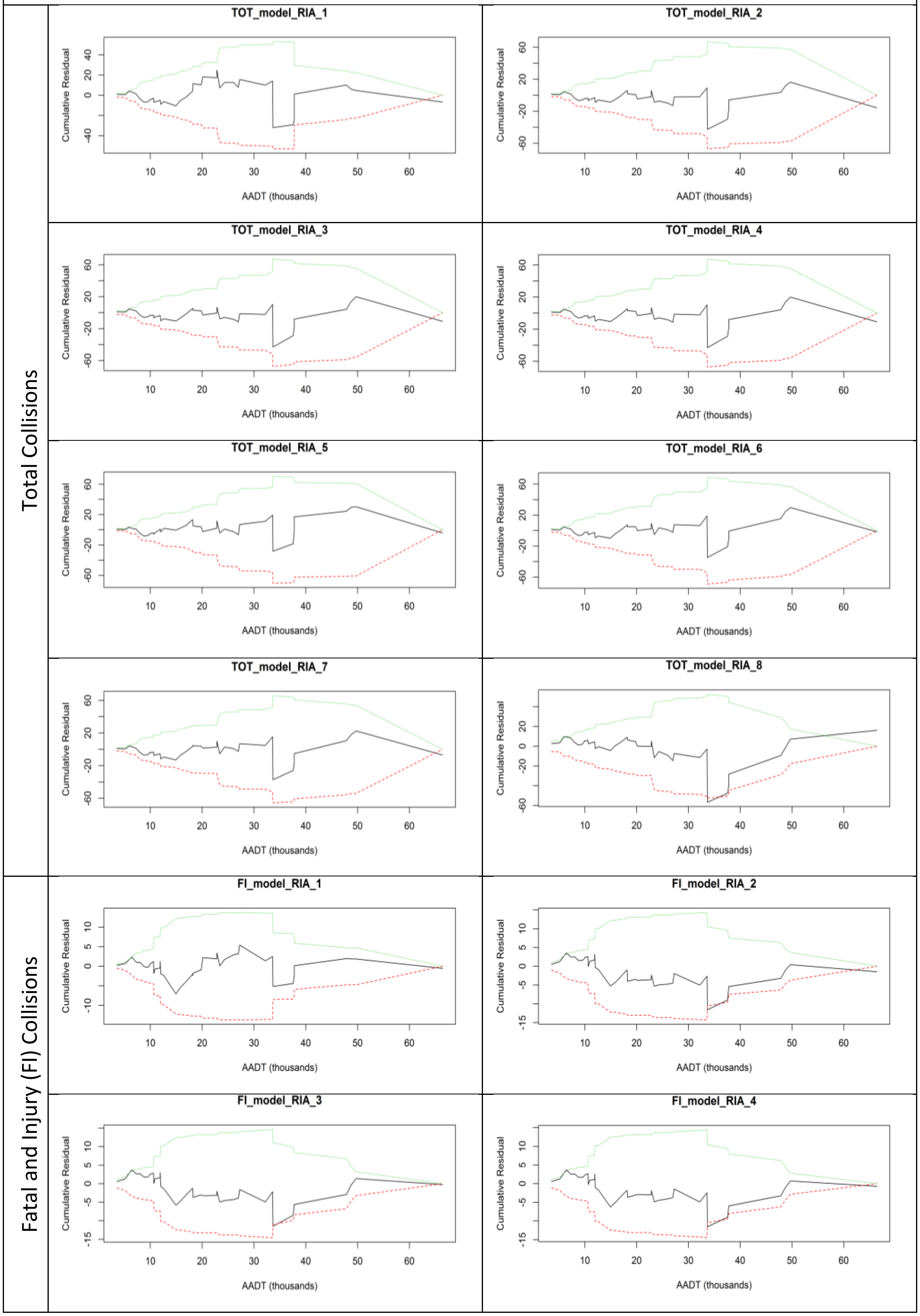




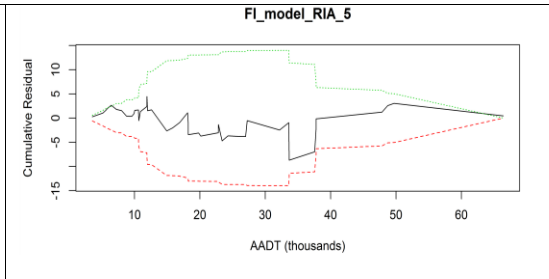


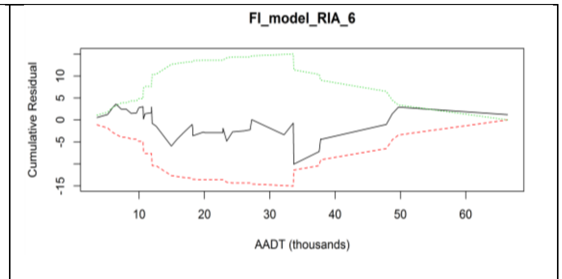


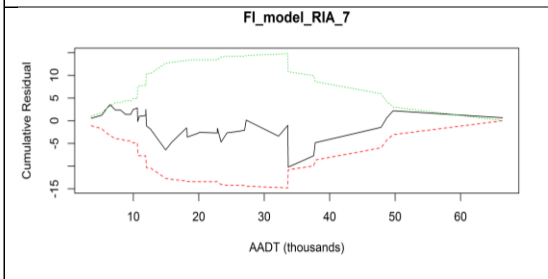
Ramp Influence Area

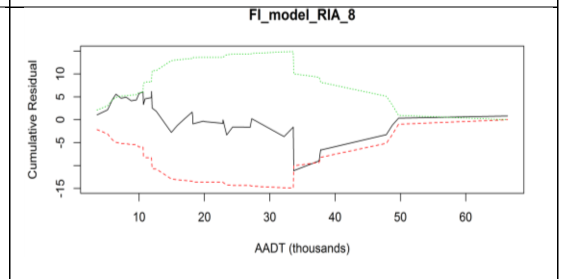


Fatal and Injury (FI) Collisions

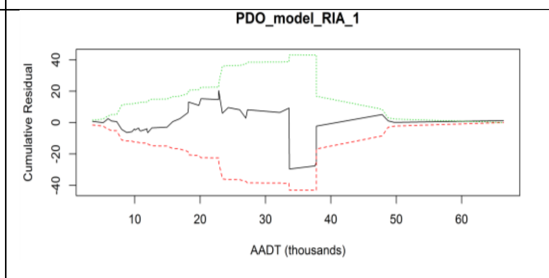


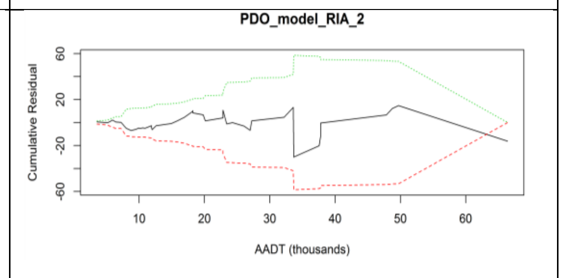


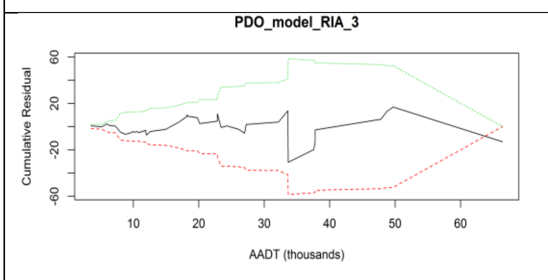


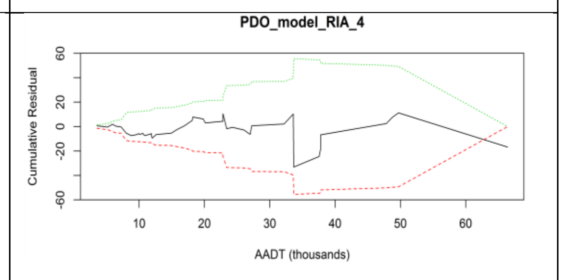


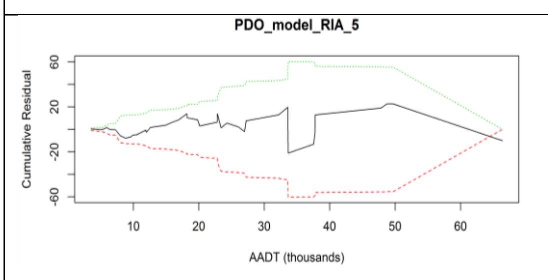
Property Damage Only (PDO) collisions

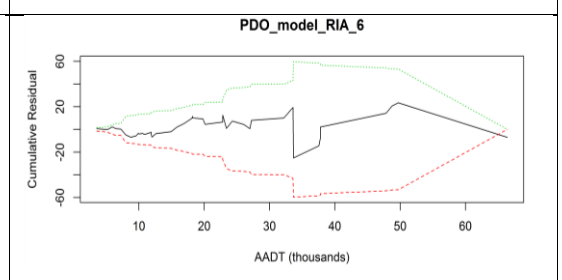


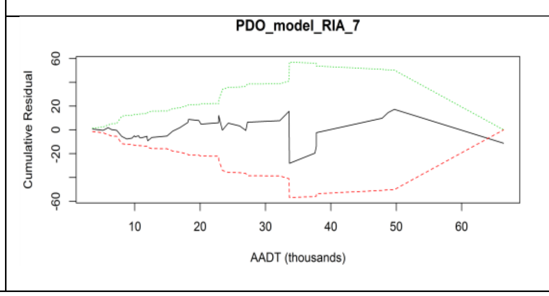


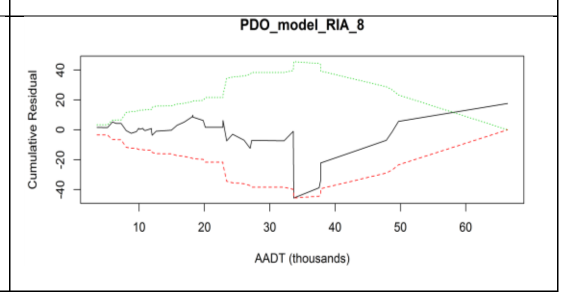






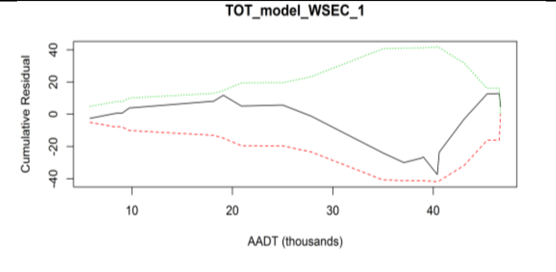


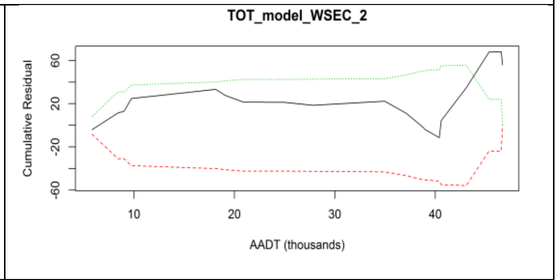


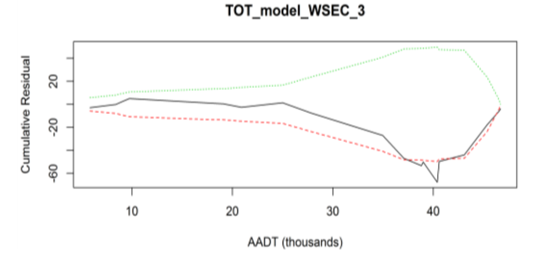


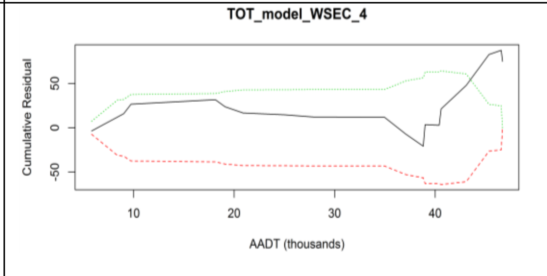
Weaving Section

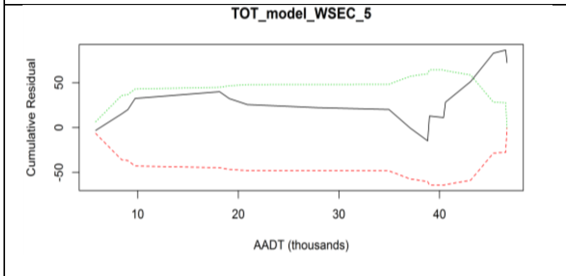
Total Collisions

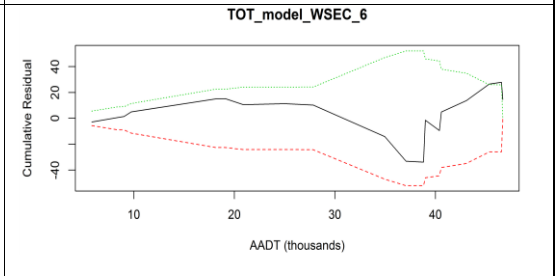


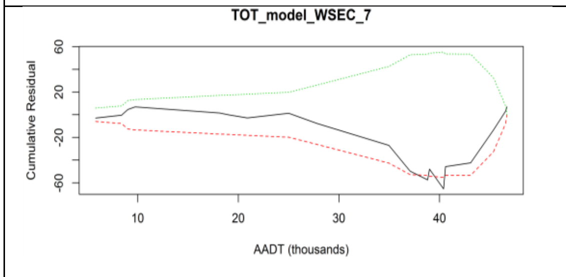


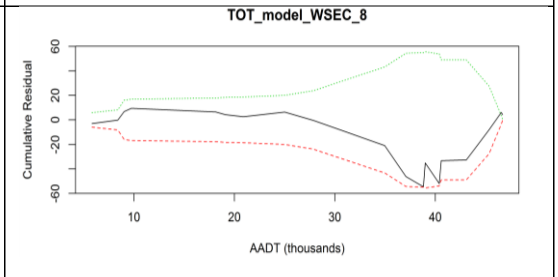


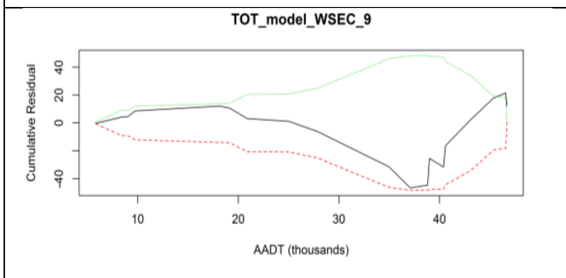


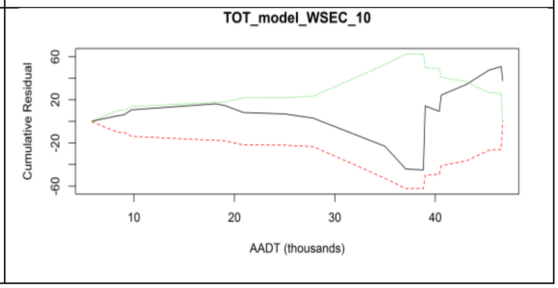






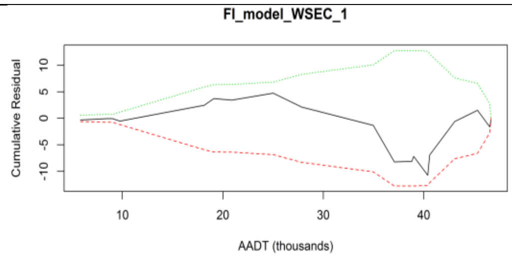


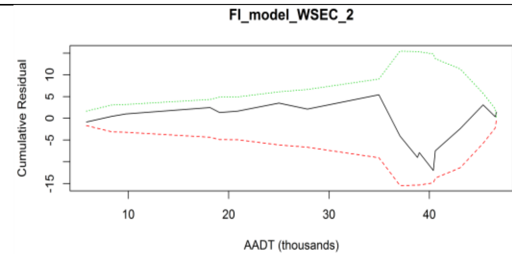


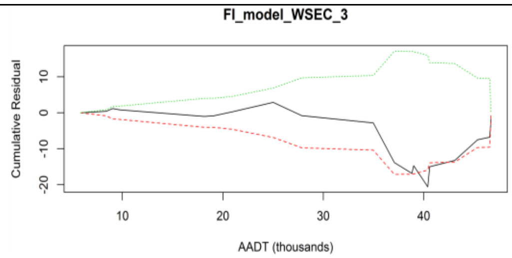


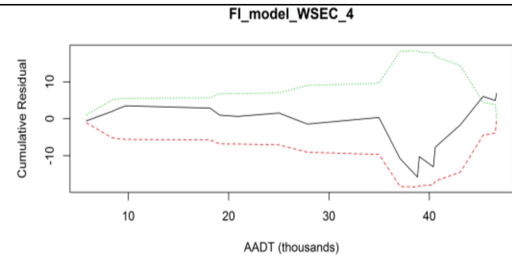
Weaving Section

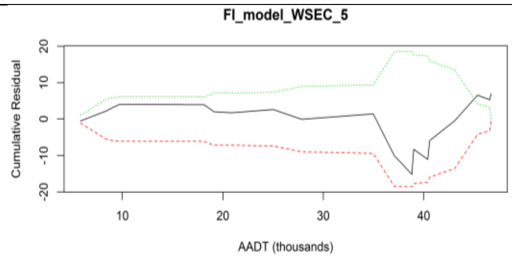
Fatal and Injury (FI) Collisions

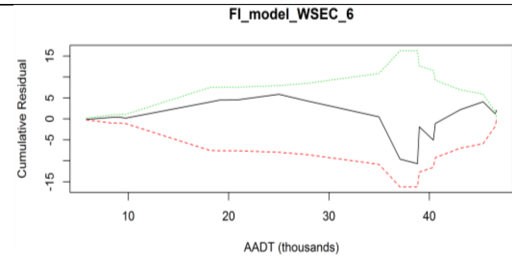


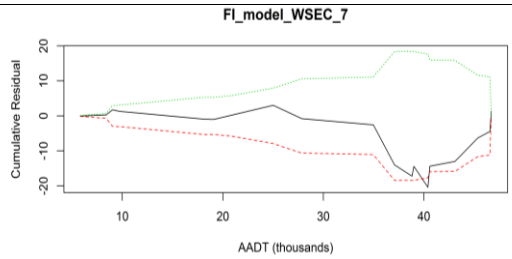


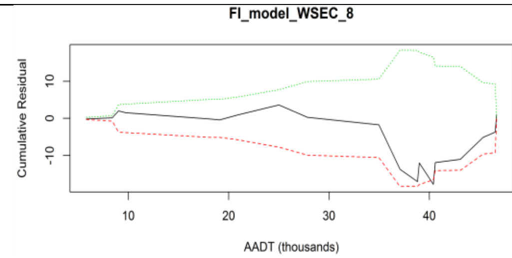


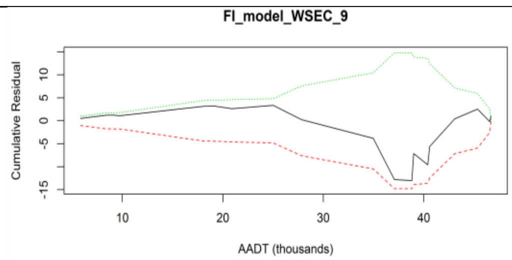


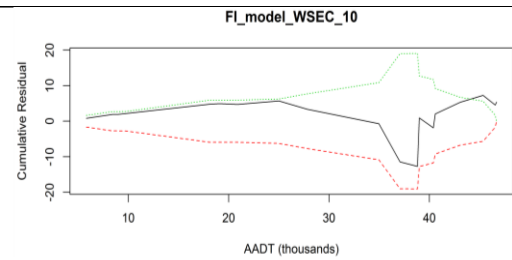






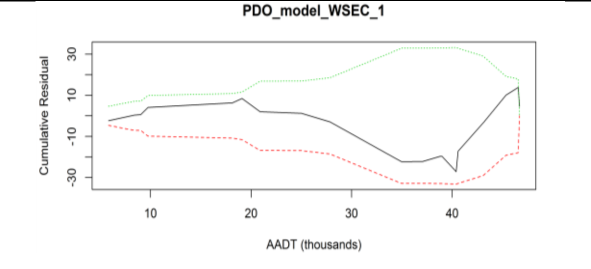


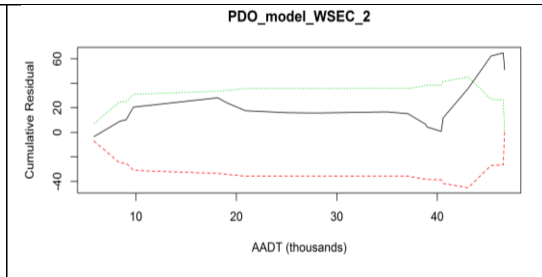


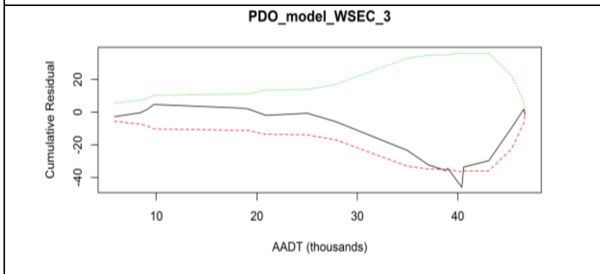


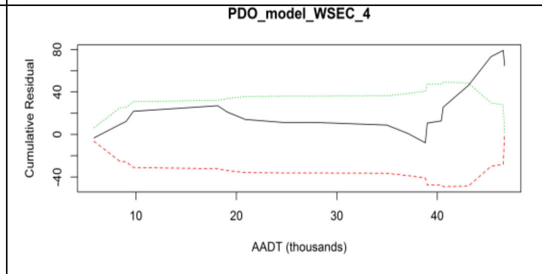
Weaving Section

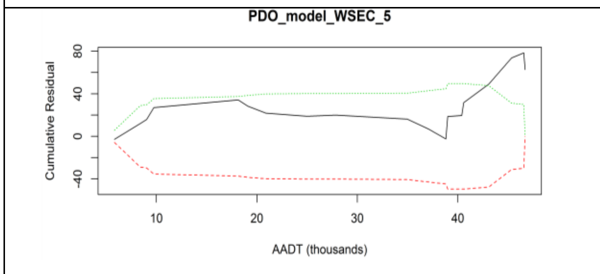
Property Damage Only (PDO) Collisions

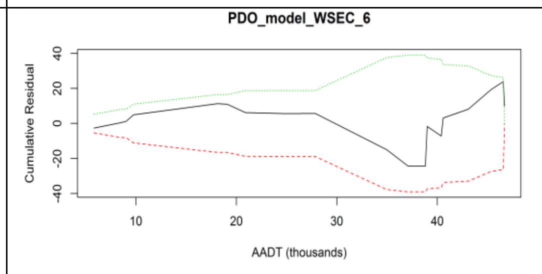


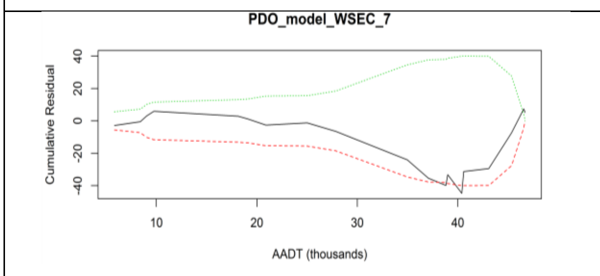


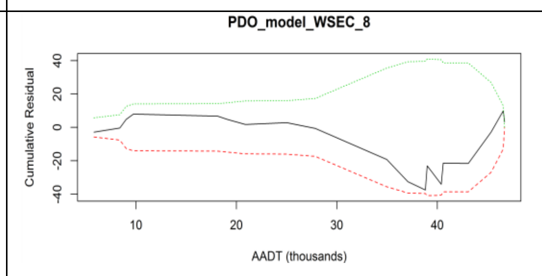


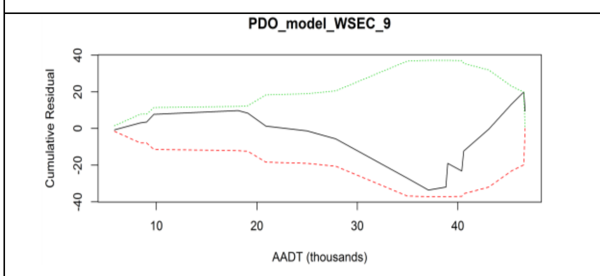


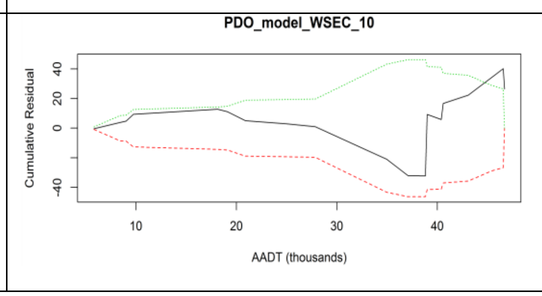






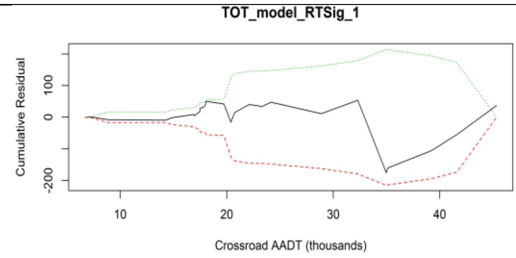


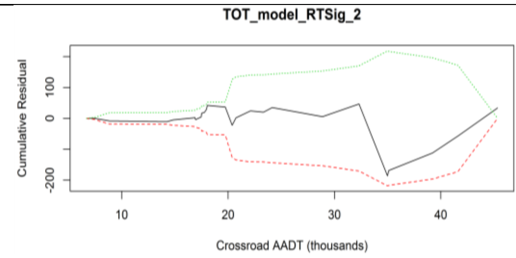


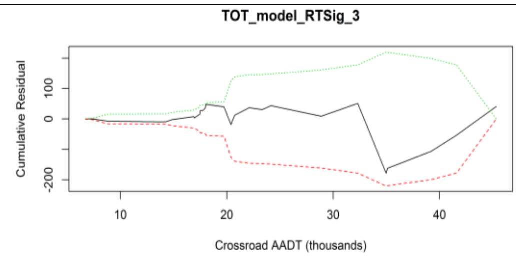


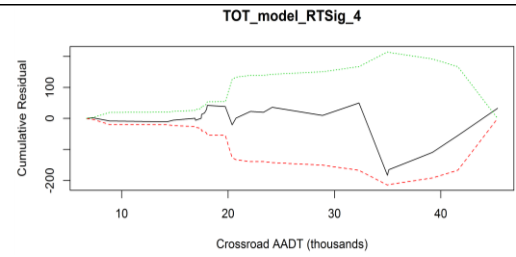
Ramp Terminal (signalized)

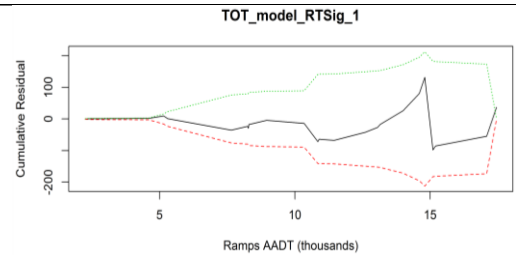
Total Collisions

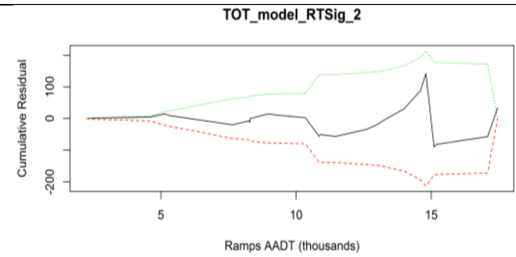


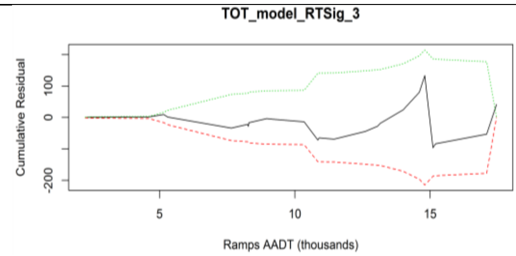


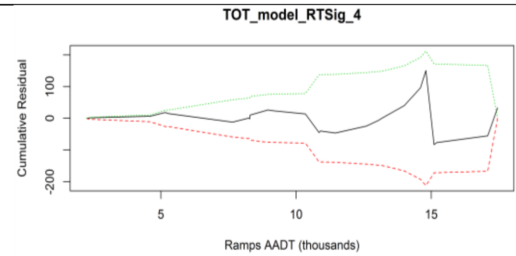






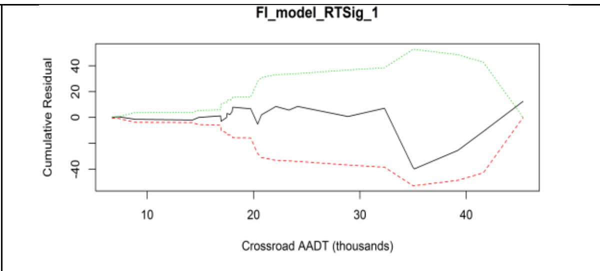


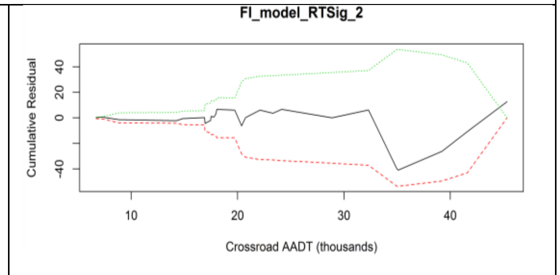


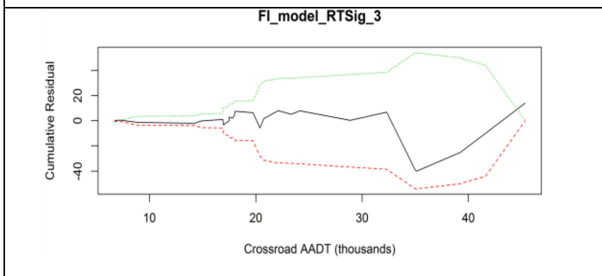


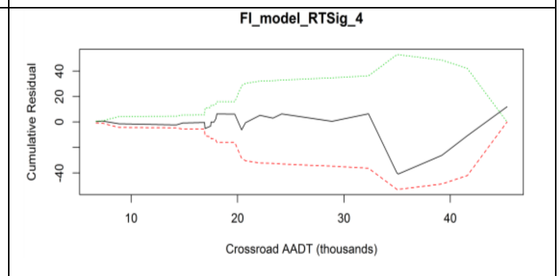
Ramp Terminal (signalized)

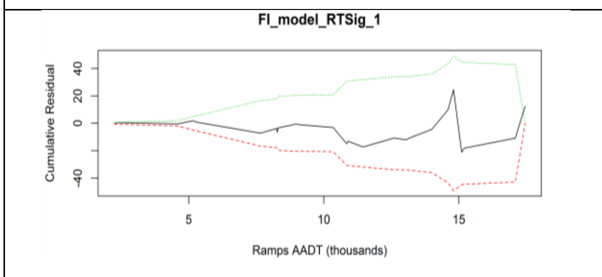
Fatal and Injury (FI) Collisions

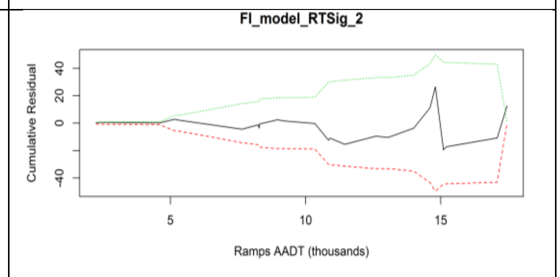


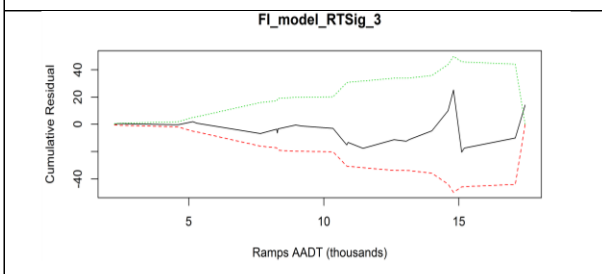


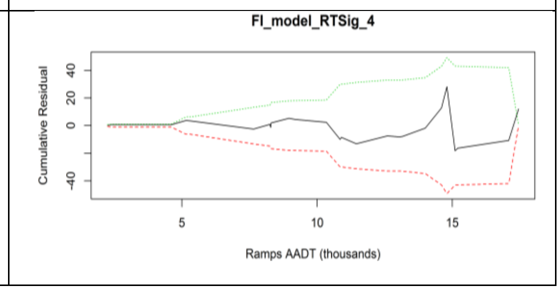






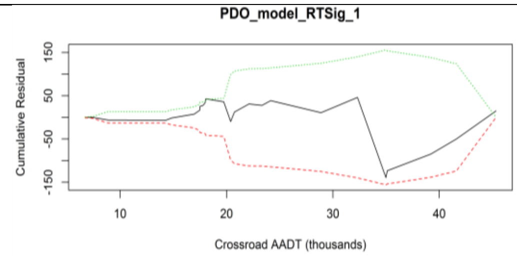


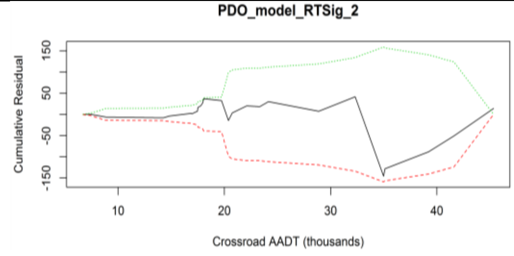


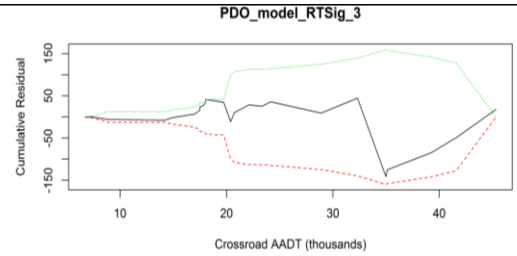


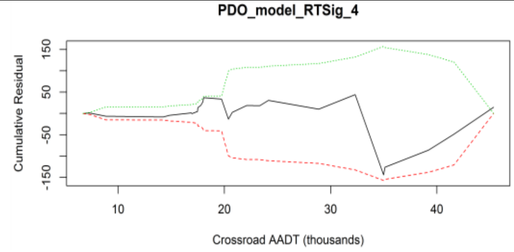
Ramp Terminal (signalized)

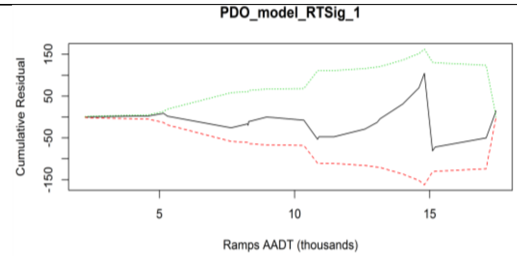
Property Damage Only (PDO) Collisions

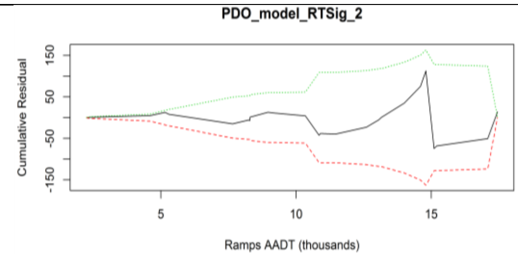


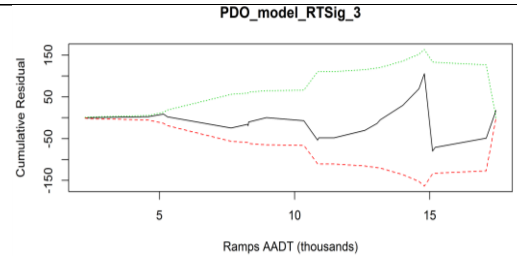


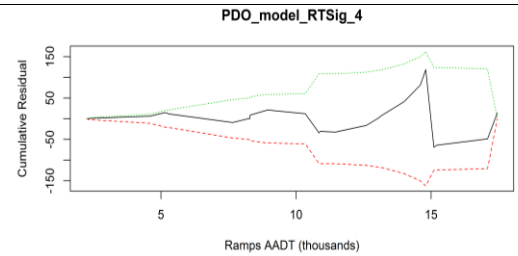






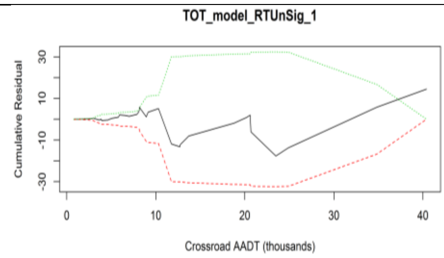


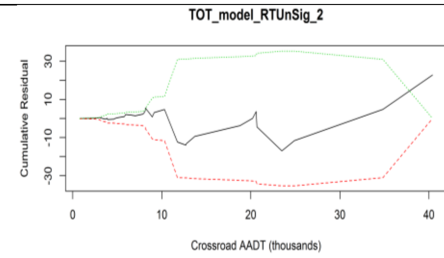


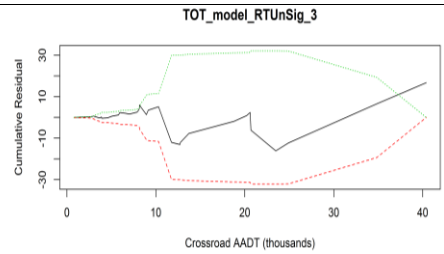


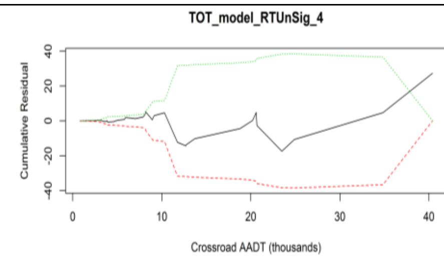
Ramp Terminal (unsignalized)

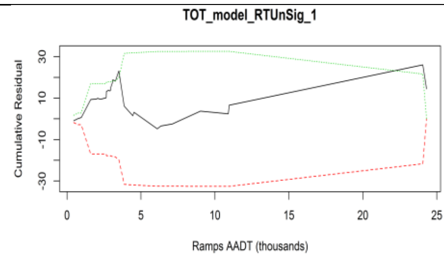
Total Collisions

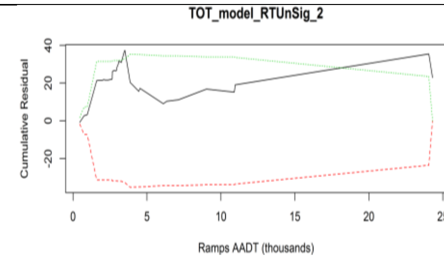


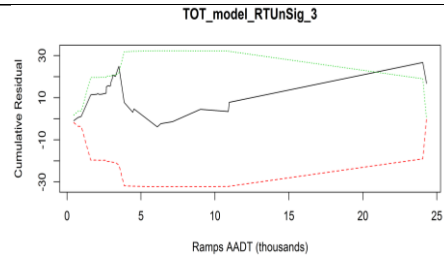


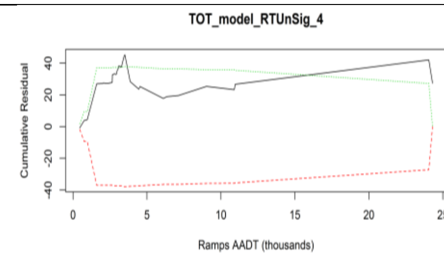






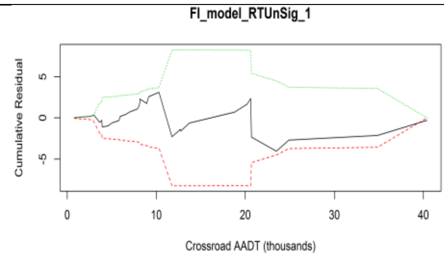


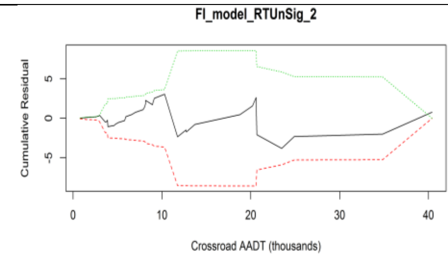


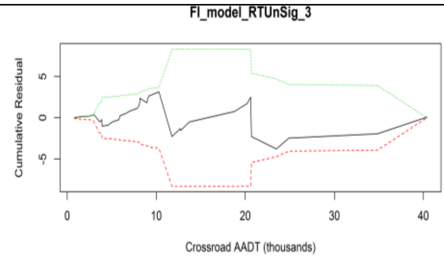


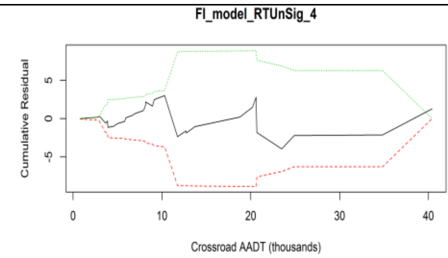
Ramp Terminal (unsignalized)

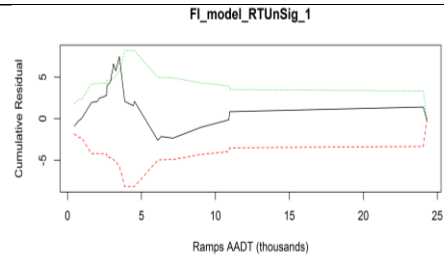
Fatal and Injury (FI) Collisions

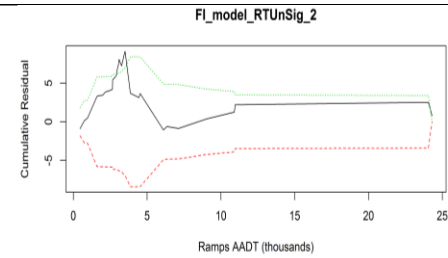


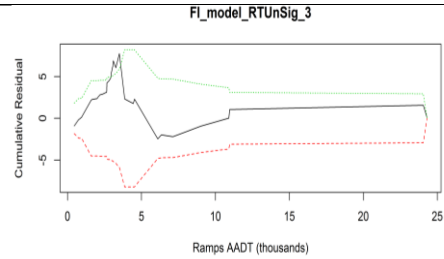


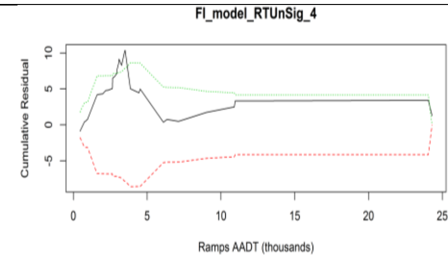






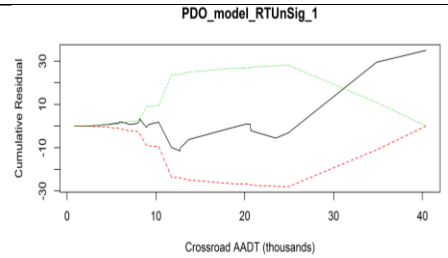


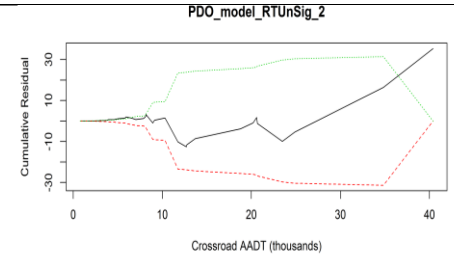


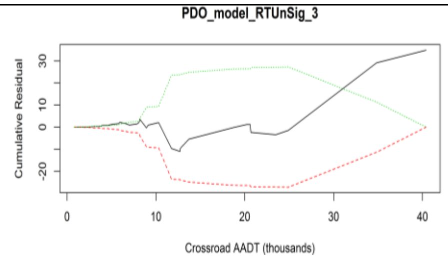


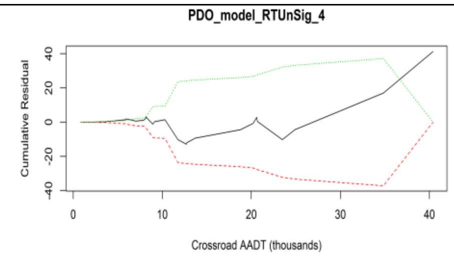
Ramp Terminal (unsignalized)

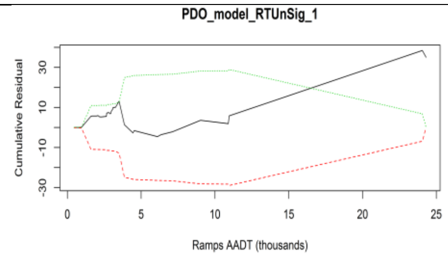
Property Damage Only (PDO) Collisions

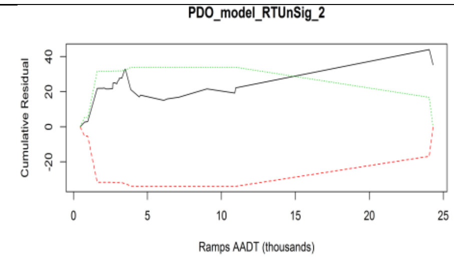


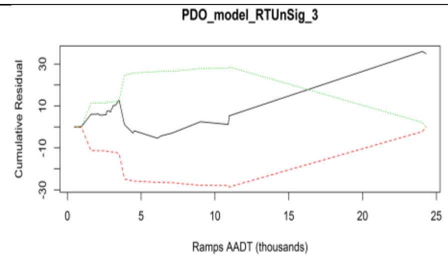


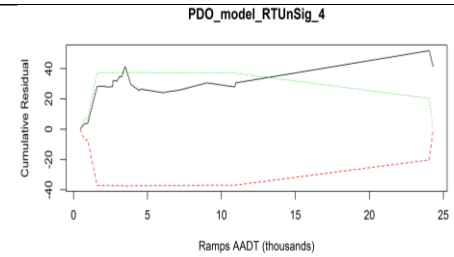












Appendix H

Sites not Included in Analysis

Due to Non-Availability of Traffic Volumes

Table H1: Details of Location Not Included in to Analysis

Community	Mainline Highway	Crossroad	Roadway Category	Number of Site
Lumsden	Highway 11	Before Lumsden	Off Ramp	2
			On Ramp	2
			Ramp Influence Area	2
			Terminal (unsignalized)	2
Moose Jaw	Highway 1	Before Moose Jaw	Terminal (unsignalized)	2
Regina	Ring Rd	Albert St	On Ramp	1
Saskatoon	Highway 11	Clarence Ave S	Off Ramp	2
			On Ramp	2
			Ramp Influence Area	2
			Terminal (signalized)	2
		Ruth Street -Adelaide St E & W	Terminal (unsignalized)	1
Swift Current	Highway 1	11th Ave NW	Terminal (unsignalized)	2
		Central Ave N	Terminal (unsignalized)	4
			Total Site Excluded	**26**