ASSESSING PAVEMENT CONDITIONS AND THEIR EFFECT ON ROAD SAFETY: ONTARIO EXPERIENCE

By

Iliya Nemtsov

Bachelor of Engineering, Ryerson University, Toronto, Ontario, Canada, 2016

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the program of

Civil Engineering

Toronto, Ontario, Canada, 2018

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

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

I understand that my thesis may be made electronically available to the public.

Assessing Pavement Conditions and Their Effect on Road Safety: Ontario Experience

By

Iliya Nemtsov

Master of Applied Science

Civil Engineering

Ryerson University

2018

Abstract

The increasing need to rebuild and repair Ontario highways has motivated this research aimed at maximizing the efficiency of pavement maintenance and design. The first of two complementary objectives were to evaluate the safety improvements of reduced pavement roughness on two-lane undivided Ontario highways using the Empirical Bayes and Cross-Sectional analysis methods. The second objective was to improve the prediction of pavement distress and surface roughness by examining the impact of local calibration of prediction models. The findings suggest that better pavement conditions can reduce the severity of fatal and injury collisions by as much as 12% in some cases and therefore that pavement maintenance decisions should incorporate road safety when assessing cost-life analysis. The results provide a basis for those decisions in that they can be used to estimate the safety effect of a specific improvement in roughness.

Acknowledgements

I would like to thank my parents for their continued encouragement and support during my graduate and undergraduate studies at Ryerson.

I would like to thank Dr. Yuan and Dr. Persaud for their knowledge and support to make all of this possible.

I would like to thank Warren Lee of Ministry of Transportation of Ontario and Craig Lyon for their advice and support throughout this work.

Lastly, I would like to acknowledge and thank the Ministry of Transportation of Ontario's Highway Infrastructure Innovation Funding Program for their funding and data support.

Table of Contents

List of Tables

List of Figures

List of Acronyms

a_{wz}	Ride Quality
AADT	Annual Average Daily Traffic
AC	Asphalt Concrete Pavement, Flexible Pavement
ARAN	Automatic Road Analyzer
β	Vector Regression Coefficients for Covariates
β_0	Intercept of the Generalized Linear Model
CMF	Crash Modification Factor
CMFunction	Crash Modification Function
CO	Composite Pavement, Semi-Rigid Pavement
CURE	Cumulative Residuals
CR	Crash Rate
DMI	Distress Manifestation Index
EB	Empirical Bayes
FC	Fatigue Cracking
FI	Fatal and Injury Crashes
GLM	Generalized Linear Model
GOF	Goodness-Of-Fit
GR	Gravel Road
HIIFP	Highway Infrastructure Innovation Funding Program
IRI	International Roughness Index
k	Over Dispersion Parameter
LHRS	Linear Highway Referencing System
MAD	Mean Absolute Deviation
ME	Mechanistic-Empirical
MEPDG	Mechanistic-Empirical Pavement Design Guide
MPB	Mean Predictor Bias
MRPNB	Multi Random Parameter Negative Binomial

MSE	Mean Squared Error
MSPE	Mean Squared Predictor Error
MTO	Ministry of Transportation of Ontario
MVAB	Motor Vehicles Accident Database Main Accident Record Table
NBR	Negative Binomial Regression
NCHRP	National Cooperative Highway Research Program
PC	Portland Cement Pavement, Rigid Pavement
PCI	Pavement Condition Index
PCR	Pavement Condition Rating
PDO	Property Damage Only
PSR	Present Serviceability Rating
RC	Reflection Cracking
RQI	Ride Quality Index
SAS©	Statistical Analyst System (Software)
SPF	Safety Performance Function
ST	Surface Treated Pavement
TC	Thermal Cracking
TOT	Total Collisions
TRB	Transportation Research Board
x_n	Independent Variable, or Covariate

Background and Motivation

Transport Canada's Canadian Motor Vehicle Traffic Collision Statistics for 2015 reported a 1.1% decrease (10,280 in 2015 to 10,397 in 2014) in severe injuries with a slight 0.3% increase (10,280 in 2015 to 10,397 in 2014) in total and fatal injuries (Transport Canada, 2015). Similarly, the Ontario Road Safety Annual Report reported the lowest fatality rate of 0.53 per 100,000 licenced drivers in 2014 (Ontario Ministry of Transportation, 2014b). However, an examination of the historical totals of fatal collisions for Canada and Ontario (Table 1) showed how moot road safety improvement has been since 2011.

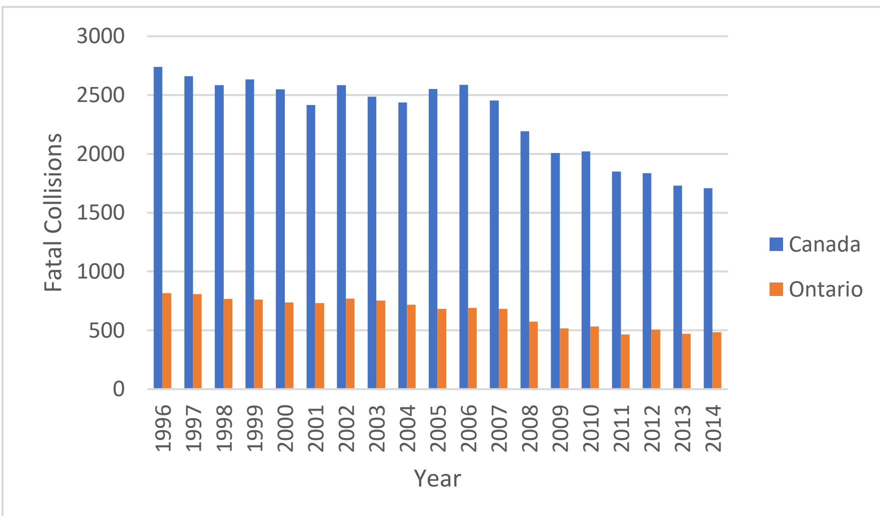


FIGURE 1. FATAL COLLISIONS (Ontario Ministry of Transportation, 2014b; Transport Canada, 2015)

There have been many breakthroughs in road safety research which have improved our understanding of how and why collisions happen and what we can do to avoid them. In a vacuum, identifying and addressing safety concerns by applying various safety treatments has been a universal standard. In reality, municipalities must choose optimal moments to invest in safety treatments or when to forgo them in favour of more urgent requirements such as road maintenance.

Road infrastructure is critical to ensuring the quality of life and prosperity of all Ontarians (Ontario Ministry of Infrastructure, 2017). Like most investments, it requires constant maintenance. The Canadian infrastructure report card (Project Steering Committee, 2016) states that without an increase in the current reinvestment rates, the condition of Canada's core infrastructure will gradually decline, resulting in substantial increases in deterioration and reinvestment costs. In addition, the delays in maintenance will increase driver discomfort, damage to the vehicles, and the chance for accidents (Chang et al., 2017).

One major element of pavement maintenance is addressing the surface roughness of the roads. Less road maintenance means greater driver discomfort. Research has primarily focused on the impact of pavement roughness on vehicle steering and braking capabilities. Nevertheless, research seldom looked at the relationship of road roughness and road safety. The quantification of road roughness on safety would be a first for Ontario and Canada as a whole (Tehrani, Falls, & Mesher, 2017).

It is important to note that simply maintaining the crumbling infrastructure is only a temporary solution. At some point in time, more roads will have to be built or replaced. This provides a chance for innovation and use of experience gained from previous research as a steppingstone for modernization.

One such example is when the MTO moved away from the uncertainty of empirical-based pavement design philosophy since the late 1990's and adopted the more reliable mechanistic-empirical (ME) methodology (Ontario Ministry of Transportation, 2014a). Similarly, MTO has dropped ME global calibration factors in favour of the more accurate local calibration factors.

The impact of the local calibration factors on highway pavement design (if any) has yet to be evaluated.

Objectives

This thesis is divided into two parts. The first part investigates the safety impact of road roughness of two-lane undivided highways and the second part is a case study in which the recently developed local calibration factors were applied to examine the significance they have on pavement structural design. Both sections address Ontario highways.

Part 1: Impact of the International Roughness Index on Safety

A literature review of recent research involving the correlation between road roughness and safety has yielded few results; of those found, most echoed the need to fill in this void in the research. Tehrani et al. (2017) estimated a set of possible Safety Performance Functions (SPF) to model the relationship between pavement conditions and collisions on Alberta highways. Until then, research on the road roughness and safety of the improvements has yet to be conducted in Canada. A search of the CMF Clearinghouse (University of North Carolina Highway Safety Research Center, n.d.), yielded only one result for rigid pavements. However, these findings are of little use for Ontario because only a small portion of highways have rigid pavements. On the other hand, flexible pavements accounted for 76% of all pavements in 2014 and 81% historically (1972 to 2014), examining this pavement type is justified.

The idea evaluating the impact of road roughness on safety was presented to the Ministry of Transportation of Ontario (MTO) in 2017 and was accepted as it fitted the goals set out by the Highway Infrastructure Innovation Program (HIIFP). Considering the possible depth of the project, this research was essential to develop CMFs and to lay a solid foundation for future research. Thus the principal objective of this part of the research was to estimate CMFs of increased IRI using a cross-sectional method and to evaluate the safety improvements following pavement maintenance activities using the Empirical Bayes (EB) before-after method. The idea is to use these results in planning pavement maintenance activities in concert with improving prediction of pavement performance, which is the objective of the second part of the research that is described next.

Part 2: Impact of the Local Calibration of Mechanistic-Empirical Pavement Design on Ontario Highways

Part 2 aimed to compare predicted performance of Ontario highways using locally calibrated design models versus the globally calibrated defaults in the AASHTOWare ME Pavement Design Software. In so doing, the sensitivity of calibration coefficients for the recently revised reflective cracking model was examined using level 3 input parameters.

Part 1: Impact of the International Roughness Index (IRI) on Safety

1. Introduction

Road safety research is a vast and critical area which has led to substantial discoveries that have improved our understanding of how and why collisions happen and what can be done to prevent them. It is essential to quantify the severity of the problem in a way that relates to real-world probability. The following sections discuss the fundamental concepts applied in this process.

1.1. Safety Performance Functions

Safety is defined as the expected number of crashes in a specified period of time. It is essential for engineers to be able to predict this number accurately. These estimates are obtained from Safety Performance Functions or SPFs. SPFs are causal mathematical relationship models between collisions in a span of time (usually a year) and various road characteristics such as section lengths, AADT, and shoulder and median widths. There are several types of collisions and levels of severity that need consideration. It is more common to develop a multitude of SPFs in order to capture the safety performance of the road segment.

The basic theory of SPFs stems from the concept of exposure. Roads with higher traffic volumes tend to experience more collisions than roads with lower traffic volumes (AASHTO, 2010b). Design and temporal characteristics need to be treated as variables as they also influence the collision rate. The SPF that summarizes this relationship is contextualized in the form of a multi-linear equation (Miaou, 1994):

$$Expected\ Num.of\ Collisions\ per\ Year = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (1)$$

Where

β_0 represents is the intercept

β_n represents are coefficients

x_n represents independent variables

The model is fit using regression analysis to identify the coefficients which indicate the strength of the relationship between the independent variable and collision frequency.

SPFs are commonly used in road network screening to identify locations with safety concerns. SPFs are more often used to study the effectiveness of safety treatments. As they allowed the observation of changes in the frequency of collisions for populations that were treated and those that were not. The difference between those frequencies is a possible indication of the average safety effect of the treatment (Hauer, 1997).

SPFs can be used in various analysis methods depending on what kind of data is available. If collision data is available before and after the implementation of the treatment(s), an Empirical Bayes(EB) before-after study can be conducted (Hauer, 1997). If collision data before the implementation of the treatment is unavailable, a cross-sectional analysis can be conducted instead. Both methods are discussed in more detail in Section 4.

1.2. Crash Modification Factors and Crash Modification Functions

Engineers use Crash Modification Factors (CMFs) or Crash Modification Functions (CMFunctions) when describing the effectiveness of safety treatments. A CMF is a decimal percent value used in conjunction with the predicted collision frequency to identify the overall safety impact of the treatment. For example, a CMF of 0.95 for a lane widening from 11 to 12 feet implies that a 5% reduction in collisions could be expected (University of North Carolina Highway Safety Research Center, n.d.). A CMFunction allows the CMF value to change based on variables that influence the treatment's effectiveness. Despite the fact that CMFunctions are being preferred, their development is often difficult and requires significantly more data to detect a difference in performance (Gross, Persaud, & Lyon, 2010).

Several methods have been developed to determine CMF's. The most common are:

- Cross-sectional studies
- Empirical Bayes before-after studies

Cross-sectional studies compare two sites that are nearly identical to one another. The only significant difference between them is the presence of the safety treatment at one of the sites and not the other. The difference in the crash frequencies at the two sites for the same period of time is inferred as the CMF. The empirical Bayes (EB) before-after study compares the collision frequencies of the treated sites to the frequencies of collisions for those sites in the event they were not treated. The EB before-after study is a more robust and preferred method of analysis given the availability of data (Gross et al., 2010). Both methods are discussed in more detail in Section 4.

1.3. Pavement Management

The objective of pavement management is to ensure that the road serviceability is maintained at a standard deemed necessary for users (Ontario Ministry of Transportation, 2013). As pavement ages, its overall condition deteriorates and the longer it remains in service, the more intensive the rehabilitation requirements are (Figure 2 left). The need for more intensive rehabilitation can be mitigated by performing minor, but more frequent maintenance (Figure 2 right). However, due to diminishing returns, the pavement would eventually need to be reconstructed. Agencies perform periodic pavement evaluations to determine the pavement conditions.

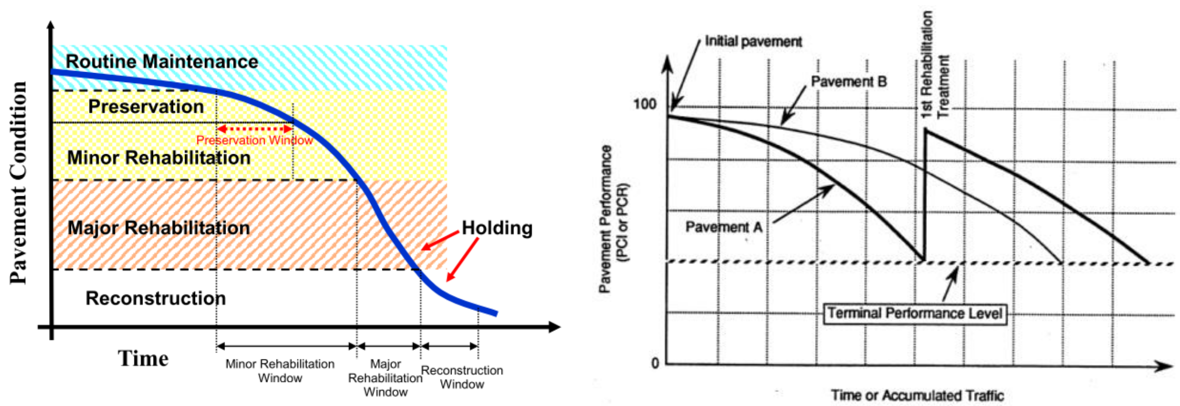


FIGURE 2. LEFT: PAVEMENT LIFE CYCLE RIGHT: PAVEMENT MAINTENANCE TO PERFORMANCE CYCLE (MTO, 2013)

The Ministry of Transportation of Ontario (MTO) has conducted routine pavement distress evaluations since the mid-1960's. The pavement is inspected to determine the following (Ontario Ministry of Transportation, 2013):

- How well the pavement served the travelling public
- Changes in pavement conditions over time
- When the pavement needs repair

- The extent of repair required

Initially, pavement distress evaluations were based on a subjective rating scale carried out by specially trained individuals who conducted road inspections and ride quality surveys to rank the road conditions using a descriptive table. The rating given was called the Pavement Condition Rating (PCR). The rating scale was problematic because results were dependent on the subjective opinion of the evaluator. Roguhness rating was especially challenging in the driving test as some people have a higher tolerance for rougher/bumpier roads.

PCR was changed in 1985 when more objective estimates became standard practice in pavement distress evaluations (Ontario Ministry of Transportation, 2013). The unified driving and inspection rating system were split into two parts: a driving test and a set of empirical calculations for the density of pavement distresses called the Distress Manifestation Index (DMI). Instead of using a subjective driving test, road roughness was now being determined from the vertical variation of suspension movement called the International Roughness Index (IRI). Together they formed the Pavement Condition Index (PCI) (Ontario Ministry of Transportation, 2013).

1.4. International Roughness Index (IRI)

Although MTO uses the IRI standard to calculate road roughness, there are several other ways to do so. The earliest type of measurement was created in the United States as a product of the AASHO road tests in the 1960's. The roughness of a road was represented as the Present Serviceability Rating (PSR), which ranged from zero (very poor) to five (excellent). The primary limitation of this method is that PSR is not scalable as it is based on the traveller's interpretation of ride comfort (National Research Council (U.S.). Highway Research Board., 1962).

The global standard was not established until the 1980's with the development of the International Roughness Index (IRI) during the International Road Roughness Experiment (IRRE) conducted by the World Bank. The criteria for development was to produce a unit that measures road roughness that is relevant, transportable, and stable over time. Transportable means that any road agency in the world should be able to obtain this measurement using the equipment they have on hand and the measurement is stable with time, valid for any road surface type and any range of pavement roughness (Sayers, Gillespie, & Queiroz, 1986).

2. Literature Review on Road Roughness and Safety

2.1. Road Roughness and IRI Thresholds

Driving is a very involved process that heavily relies on the driver's competence behind the wheel (AASHTO, 2010b). Drivers are also heavily influenced by their surroundings (Hauer, 2000), including the discomfort due to vibrations as a result of a rough road surface (Loprencipe & Zoccali, 2017). Road roughness has been linked to increase stress in vehicle structures, increase in the dynamic loads applied to pavement structures accelerating fatigue damage, and decrease in road vehicle-interactions such as steering and braking (Sayers & Karamihas, 1998). Canstisani and Loprencipe (2010) have found that drivers experience greater discomfort at higher IRI with faster driving speeds. The study only examined speeds up to 80 km/h. Múčka (2016) explained that IRI imposes an 80km/h limit because IRI was developed under this specification. The transferability of IRI remained intact as Loprencipe and Zoccalli (2017) determined that the Ride Number, Michigan Ride Quality Index (RQI), and Minnesota Ride Quality Index were comparable to the ride quality indices of the IRI.

The acceptable comfort threshold varies by region, and in Canada, the IRI thresholds have not been clearly defined (Jurgens & Chan, 2005; Tehrani, Falls, & Mesher, 2015; Transport Association of Canada, 2006). The most commonly used IRI threshold levels were presented by Jurgens and Chan (2005). These are defined as:

- GOOD (IRI < 1.5 meters per kilometre)
- FAIR (IRI 1.5-1.9 meter per kilometre)

- POOR (IRI > 1.9 meters per kilometre)

Tehrani et al., (2015) created a questionnaire for road users of Highway 2 in the province of Alberta. The primary objective was to define IRI thresholds based on driver perception of road roughness. The results indicated the following IRI comfort thresholds:

- VERY GOOD (IRI < 1.02 m/km)
- GOOD (1.02m/km<IRI<1.35m/km),
- FAIR (1.35m/km<IRI<1.6m/km),
- POOR (1.6m/km<IRI<1.85m/km)
- VERY POOR (IRI>1.85m/km)

It is now important to note that typically, every region will specify the level of "importance" of each highway. The Ministry of Transportation of Ontario Design Guidelines stipulates that there are four major divisions in highway classification: freeway, arterial, collector or local (Ontario Ministry of Transportation, 1985).

The second edition of the Pavement Design and Rehabilitation Manual (Ontario Ministry of Transportation, 2013) does not explicitly specify IRI thresholds by user perception, but rather by the aforementioned highway classification. This classification system starts with the most used, critical, and centralized routes used by the majority of the population and ends with the least frequently used routes that only serve smaller portions of the population. The MTO Pavement Design and Rehabilitation Manual provide the DMI and PCI thresholds which are the trigger values that indicate that the pavement needs to be treated and to be brought back up to user satisfaction. Since this research deals with flexible pavement, the trigger IRI values have been back-calculated for the flexible pavement in Table 1, using Equation 2 (Ontario Ministry of Transportation, 2013).

$$PCI = Max(0, Min(100,13.75 + 9 * DMI - 7.5 * IRI)) \quad (2)$$

TABLE 1. PERFORMANCE THRESHOLD PER FUNCTION CLASS

FUNCTION CLASS	DMI	PCI	IRI – BACK CALCULATION	TERMINAL IRI – ME DESIGN
FREEWAYS	7.3	65	1.93	1.9
ARTERIAL	7	55	2.9	2.3
COLLECTOR	6.8	50	3.33	2.7
LOCAL	6.8	45	3.99	3.3

However, the Interim Report for Ontario's Default Parameters for AASHTOWare Pavement ME Design (Ontario Ministry of Transportation, 2014a) specifies slightly smaller values for terminal IRI.

2.2. Road Roughness Safety Performance Function Covariates

In general, the widely accepted Safety Performance Function (SPF) for a highway road segment is as follows (AASHTO, 2010b; Heydecker & Wu, 2001; Zeng, Fontaine, & Smith, 2014):

$$\frac{Collisions}{Year} = e^{\beta_0} AADT^{\beta_1} Length^{\beta_2} e^{\beta_n x_n} \quad (3)$$

Where:

β_0 = is the intercept

β_n=is the vector parameter(s)

x_n = is the value of the explanatory covariate(s)

n = the number of explanatory covariates

$AADT, Length$ represent the annual average daily traffic (AADT) and section length (respectively), which are considered key parameters for jurisdiction-based SPFs.

Since road roughness research is still in its infancy, the model structure has not yet been clearly defined. It is thus important to examine other possible SPF forms used to predict collisions.

The first model by Elghriany, Yi, Liu and Yu (2016) has abandoned the conventional model structure and opted to estimate the collision rate (CR):

$$CR = \frac{Crashes * 1,000,000}{AADT * 365 * Number\ of\ Years\ * Segment\ Length} \quad (4)$$

Crashes were filtered to use only pavement-roughness related to collisions in Ohio. The IRI and measured CR were used to fit linear ($CR = \beta_0 + \beta_1 IRI$), quadratic ($CR = \beta_0 + \beta_1 IRI + \beta_2 IRI^2$), exponential ($CR = \beta_0 e^{\beta_1 IRI}$), and power ($CR = \beta_0 * IRI^{\beta_1}$) models. The data sample only considered rigid highway pavements with an IRI greater than 1.5 m/km and speeds of 60mph to 65mph.

Abdel-Aty, Devarasetty and Pande (2009) developed SPFs that consider, the speed limit and the number of lanes to model total, rear-end and severity of collisions on multilane arterials in Florida. The SFP's were based on urban, suburban and rural land use. Abdel-Aty et al. (2009) did not incorporate section lengths into the dispersion parameter as a function as per Hauer (2001); instead, the data was split, based on a set of arbitrarily chosen segment lengths. The SPFs did not include any roughness index, despite the fact that they were conducting a before-after analysis of resurfacing projects.

In Tennessee, Chan, Huang, Yan and Richards (2010) separately modelled the rut depth, International Roughness Index (IRI) and Present Serviceability Index (PSI) as parameters, without the inclusion of section length, to predict the collision frequency per lane. Also, the time of day, weather, total, and peak/off-peak accidents were modelled. Tehrani et al., (2017) modelled the total collisions for two-way highways with dividers. The IRI and rut depth were used as parameters, but their values were taken as either an average or maximum. The weather, horizontal and vertical alignment, and surface conditions were the other model parameters.

Zeng, Fontaine, and Smith, (2014) modelled total, fatal, and injury crashes using shoulder size and lane width as parameters for rural two-lane undivided highways in the nine Virginia construction districts. Highway sections shorter than 0.1mi and longer than 10mi were excluded from the analysis. Lee, Nam and Abdel-Aty (2015) on the other hand used the Bayesian ordered logistics models for low, medium, and high-speed roads to model single and multiple vehicle collisions on Florida's highways. Their modelling choice was influenced by the pavement condition index (PCI). Since the PCI is ordinal, the comparison between different indexes could not be made directly, PCI was generalized to be modellable.

Lastly, Chen, Saeed, Alqadhi and Labi (2017) had created multivariate random parameter negative binomial models (MRPNB) to account for the unobserved effects across crash severity levels and existing heterogeneity across road segments. IRI was initially used as a parameter inside the MRPNB model. Chen et al. (2017) later work, used IRI outside the model according to the following thresholds: Excellent (60-100), Good-Fair (126-150), Fair (151-200), and Poor (≥ 200) (IRI in in/mi). The other parameters were average daily trucks, lane width, outside shoulder width, vertical curve grade and median width.

2.3. Safety Impact of Pavement Roughness

An initial investigation into the safety effects of pavement resurfacing projects were conducted in the late 1980's. A report by the Transportation Research Board (TRB) (1987) concluded that the effect of resurfacing projects was based on the reasoning for the pavement maintenance. If pavement maintenance was conducted because of structural concerns or poor riding quality, accidents were found to increase by an average of 2%. Of that 2%, there was a 10% reduction in wet pavement collisions offset by an equal increase in dry pavement collisions. When pavement resurfacing projects were selected to address the high number of wet pavement collisions it led to a decrease of 15-70%. Considering the full life of the project, the authors predicted it would average to a probable 20% reduction in wet pavement collisions overall. More so, the wet pavement collision treatment decreased total collisions by 5% and increased dry pavement collisions by 15%.

The National Cooperative Highway Research Program (NCHRP) project 17-9 investigated the safety effects of highway standards. The results indicated the existence of significant gaps in understanding the influence of design on safety. A follow-up project 17-9(2) explored the implications of pavement resurfacing. An empirical before-after study was conducted for two-lane rural and suburban highways in Washington, California, Minnesota, New York, and Illinois.

The objective was to see if resurfacing would make a substantial difference in safety when paired with minor and major safety improvements (Hughes, Prothe, McGee, & Hauer, 2001).

Huges et al. (2001) indicated that the procedure used in project 17-9(2) lacked the control over the regression-to-the-mean bias and other forms of biases. Furthermore, the improvements for Washington and Minnesota could not be directly compared due to differences in methodology for these States. As a result of these problems, results were inconclusive.

Zeng et al., (2014) conducted an Empirical Bayes before-after study of pavement rehabilitation for two-lane highways in Virginia. Their EB analysis revealed a 26% reduction in fatal injury crashes, but the overall crash frequency was not significantly affected by resurfacing. The results revealed no effect of pavement rehabilitation on reducing run-off-road collisions. Sideswipe collisions were reduced by 6.1% (10:1) and 5.3% (6:0) for fatal and injury collisions. Total nighttime collisions increased by 15% for the after a period. However, the number of fatal and injury collisions was reduced by 37.5%. The total number of wet pavement collisions did not indicate a notable change. However, 41.7% fewer fatal and injury crashes were observed.

Jaeyound, Nam, and Abdel-Aty (2015) found that multi-vehicle collisions are more likely to occur under poor pavement conditions (PCI 0.0-3.0) for all road speeds ($\leq$35 to $\geq$50 mph). Head-on, angle and turning collisions were 9%, 1%, and 2% more likely to happen (respectively). Conversely, rear-end collisions were 12% less likely to happen. The severity of single-vehicle collisions decreased on poorer pavement conditions for speeds $\leq$ 35mph but increased for speeds $\geq$50mph. The number of collisions related to hitting fixed objects or running into ditch/water hazards increased by 15% for poor pavement conditions, while the number of collisions related to guardrails and overturning was reduced.

Elghriany et al., (2016) came to similar conclusions for rigid pavements in Ohio, suggesting that roads are safest if the International Roughness Index (IRI) is kept close to 1.5 m/km. The results revealed 7.4% increase in the crash rate when the IRI is increased to 1.75 m/km, 231.8% increase in the crash rate when the IRI was increased to 2.250 m/km and 448.7% increase in the crash rate when the IRI was increased to 2.5 m/km.

Lastly, Chen at al., (2017; 2017) concluded that two-lane highways in poor condition have a lower frequency of no-injury crashes when compared to highways with four or six lanes. Their findings suggested that two-lane highway pavements in excellent condition are associated with a higher frequency of injury and no-injury collisions.

3. Database

This section provides information about the data sources, the creation of the primary database, and a summary of the analysis database.

3.1. Data Sources

3.1.1. Ontario Road Characteristic Data and AADT

The Ontario road characteristic data was provided by the Ministry of Transportation of Ontario (MTO) and extracted from the Integrated Highway Information System (IHIS) in 2011. The spreadsheet was organized using MTO's Linear Highway Referencing System (LHRS). This referencing system uses unique milepost identifiers to refer to the specific location within the road network. When LHRS is combined with a milepost offset, it is possible to reference any position along the road. Some sections can be referred by two different LHRS, where one LHRS is a subset of another. Since roads are continuous, the offset and lengths of LHRS reference can be referenced by the highway number instead.

For example, take the section with LHRS of 10008 which runs from the 1.163km mile point and ends at 2.208km mile point on Highway 1. In one data set, LHRS 10008 is referenced. In another set of data, the same section is not referenced, the closest LHRS is 10004 which runs from 0.23km to 5.658km on Highway 1. Thus, LHRS 10008 must be contained within LHRS as the highway number, and the beginning and end milepost do not exceed the LHRS 10004 limits. One way to work around this inconsistency is to use the highway number (Highway 1) instead of the LHRS. Since the mile makers are always increasing and do so exclusively for each highway, LHRS can be forgone when needed.

In addition to the various geometric features of every section of road, the data-set tracked pavement maintenance activities categorized as construction, reconstruction, resurfacing and widening.

The Annual Average Daily Traffic (AADT) volumes were also provided by the MTO as a shapefile. The shapefile was imported into the ArcGIS© software from which the AADT volumes were exported to a spreadsheet. The data was referenced by LHRS sections and included the following details:

- AADT volumes for the years 1988 to 2016
- Highway Number
- Offset
- Section Length

3.1.2. Collision Data

The raw collision data were provided by the MTO for years 2000 to 2013 in a spreadsheet format, accompanied by a PDF file that allowed the collision data to be decoded (Ontario Ministry of Transportation, 2004). Each accident was tied to a unique microfilm number, and the location of the collisions was referenced by LHRS and offset. Between 2010 and 2011, there was a change in the data entry procedures; the new procedure included vehicle specific information for each accident, unlike the previous years which only include the general information on the accident. As a consequence, the microfilm number identifiers were no longer unique. Because of the large data span, attention had to be paid when mixing old and new data formats, to ensure the proper removal of duplicate entries. In addition to the lateral location of the collision, the horizontal location was classified as being either on or off the highway. The MTO’s off-road collision classification refers to off-road collisions as any collision that has occurred in a location which is not a public road (Ontario Ministry of Transportation, 2004).

3.1.3. Pavement Management Data

The pavement performance database was provided by the Pavement and Foundation Section of the Ministry of Transportation of Ontario for this research in 2017. Two separate spreadsheets with performance data from 1971 to 2012 and 2013-2014 were merged into one dataset. Each row of data was referenced by LHRS, beginning mile, end mile, and the year. The pavement condition data included the Road Condition Index (RCI), the Damage Manifestation Index (DMI), and the Pavement Condition Index (PCI). Pavement roughness data included the International Roughness Index (IRI) and the Rut Depth. As discussed in Section 1.3, PCI is calculated from the DMI and IRI. IRI data were only recorded in the dataset following 1997. The five pavement types that appear in the spreadsheet file are:

- AC – Asphalt Concrete (Flexible pavement)
- CO – Composite (Semi-Rigid pavement)
- GR – Gravel Roads
- PC – Portland Cement (Rigid pavement)
- ST – Surface Treated

The pavement type was recorded in two columns, to indicate if there was a change in the pavement surface type between the beginning and end of the year. For example, a CO surface can become an AC surface later in its lifetime. Lastly, rehabilitation activity was recorded for the year in which it was conducted. The major categories of rehabilitation activities were: reconstruction, full depth reclamation, hot mix overlay, milling, rebuild, rehabilitation, and surface treatment. For more information on the type and expected service life of treatments, please refer to (Ontario Ministry of Transportation, 2013).

3.2. Database Development

The first objective of Part 1 was to develop a database that merges the road characteristics, AADT, accident, and pavement condition data for Ontario highways. The process flowchart is presented below.

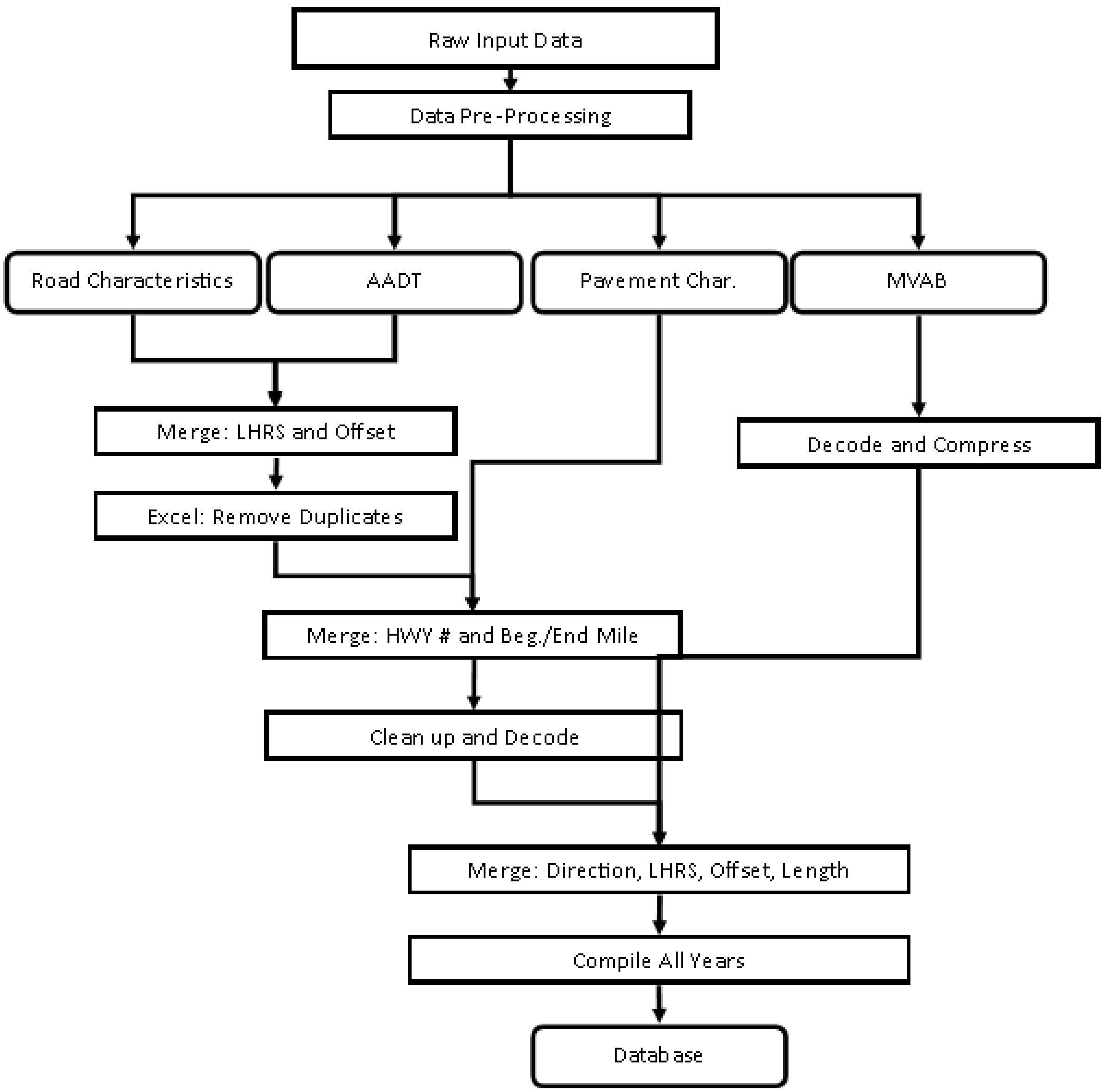


FIGURE 3. DATABASE PREPARATION FLOWCHART

Most of the process was done by using SAS© programing language. The general code is presented in Appendix A. Each step has been discussed in more details below.

Data Pre-Processing:

All four files came in their raw format as spreadsheet saved in the XLS or XLSX format. The following processing was done to each file in Excel as preparation for import to SAS.

Road Characteristics: Left as is.

Annual Average Daily Traffic (AADT): Amended to only include 'aadt' for years 2000-2013. The 'highway' and 'id' columns were dropped.

MVAB: There were two different data entry formats (new and old), it is sensible that the old data format dictated which fields could be used. All common columns were kept based on visual comparison. The duplicates of the years 2013 and 2012 were removed using the built-in Excel function, and the microfilm number was used to identify the duplicates. The reasoning is further explained in Section 3.1. The columns 'rdloc', 'rdsur1', 'rdsur2', and 'trafcon' values were a mix of string and numeric data. All those values were changed to numeric: 'rdloc' – B→10, 'rdsur1'/'rdsur2' - A→10, and 'trafcon'- A→10, and B→11. The MVAB kept the years as separate files.

Pavement Char.: This was a merge of the two files discussed in Section 3.3. The 2013 data had bridge deck information which was not included in the merge. The columns past 'friction' were not included. Data for the years 2000 to 2013 was included for flexible AC sections only. The 'Direction' column values included string entries indicating compass directions. They have been changed to the numeric equivalent specified in (Ontario Ministry of Transportation, 2004). Finally, the pavement file for years 2000 to 2013 was split into 14 files by year.

Excel: Remove Duplicates:

With the first merge complete the results were scrutinized. It was found that 'lhrs' and 'offset' do not match up causing multiple AADT values to be corresponding to some road segments. The road characteristics were exported to Excel and sorted by 'lhrs' and 'offset' (by road characteristics reference). A nested "IF" statement was then used to identify the duplicates based on the 'lhrs' and 'offset', and the overlapping AADT volumes were averaged. The processed file was then brought back into the SAS© environment.

Clean up and Decode:

Some inconsistencies had to be taken care of in preparation for the merge with the MAVB. 'routeaux' indicated auxiliary routes such as collectors and express lanes. These routes were removed since they are not the focus of this study. Sections that had a 'pqi', 'psi', 'pdi', 'iri' of zero, or 'surf_width' of blanks or zero were also removed.

The merge with the Pavement Char. added the second inconsistency. Unlike the last time, the sections had to be re-segmented as the 'activity' field could not be taken as an average. The 'from_d' and 'to_d' of pavement characteristics were compared to the 'begin_mile' and 'end_mile' of Pavement Char. to obtain all new section mile points, descriptions, and offsets.

Lastly, the terrain was decoded from a string to a number value. Since the Pavement Char. is direction dependent, in cases where the road had two directions (e.g. 1(North) and 2(South)), the AADT was divided in half. The direction was left as is if the direction was not specified/central (5) and 'ident' was calculated to serve as a unique identifier.

Decode and Compress:

At this step, the class of accident ('classac') was decoded. Since the study only considered highway accidents, any road locations ('rdloca') other than "on-highway" or "highway other" was not considered. Due to multiple observations for a single location, the accident counts had to be summed up.

Compile All Years:

The MVAB and Pavement Char. were initially separated by year to ease the merging of the previous steps. Since they all had the appropriate AADT and road characteristics appended, the table only needed to be updated to include them all in one file.

3.2.1. Analysis Database

The analysis database used for the Empirical Bayes and Cross-Sectional studies is a subset of the database discussed in Section 3.2. Many of the major requirements for the development of Safety Performance Functions were already met in the preceding section. This includes the calculation of the AADT volumes when direction dependent, taking the natural logarithm of AADT, length, and IRI.

Since the analysis only included two-lane undivided highways with a central pavement characteristic data ('direction' = 5), a query was made in SAS to select only those sections. Because of Excel import and export, some of the data were incorrectly categorized in a string format when it should have been numeric. This was then rectified in the code. The Total and Fatal

+ Injury calculations were then conducted. Lastly, all of the locations with a recorded International Roughness Index of zero were removed from the database since this was not considered a realistic number for the pavement type. IRI values closer to zero are only expected from highway runways (Sayers et al., 1986). The query used in SAS© to create the database is presented in Appendix B: SAS Code for Creating an Analysis Database .

3.2.2. SFP Database Summary

The following table is a general summary of the data obtained from the query mentioned in Section 3.2.1. These data were used to develop samples and produce the Safety Performance Functions (SPF). Also, the same data were used to obtain the seasonal factors to calibrate the SPFs for the yearly variations. The data summary is presented in the tables below:

TABLE 2. SUMMARY STATISTICS FOR TWO-LANE UNDIVIDED HIGHWAYS

TWO-LANE UNDIVIDED HIGHWAY (N=1261)

DATA ITEM	Minimum	Maximum	Mean	Sum
AADT	0	32300	4736.28	-
SEGMENT LENGTH (KM)	0.01	46.9	5.78	7247.68
IRI	0.5	8.16	1.68	-
TOTAL CRASHES	0	46	3.96	67099
FATAL+INJURY	0	15	0.79	13423
PROPERTY DAMAGE ONLY	0	37	3.17	53676

TABLE 3. SUMMARY STATISTICS FOR TWO-LANE UNDIVIDED ARTERIAL HIGHWAYS

TWO-LANE UNDIVIDED ARTERIAL HIGHWAY (N=639)				
DATA ITEM	Minimum	Maximum	Mean	Sum
AADT	0	32300	5263.96	-
SEGMENT LENGTH	0.02	35.90	5.96	3791.87
IRI	0.5	5.07	1.58	-
TOTAL CRASHES	0	46	4.60	39830
FATAL+INJURY	0	15	0.94	8179
PROPERTY DAMAGE ONLY	0	37	3.66	31651

TABLE 4. SUMMARY STATISTICS FOR TWO-LANE UNDIVIDED COLLECTOR HIGHWAYS

TWO-LANE UNDIVIDED COLLECTOR HIGHWAY (N=506)				
DATA ITEM	Minimum	Maximum	Mean	Sum
AADT	0	28500	4791.71	-
SEGMENT LENGTH	0.01	46.9	5.78	2929.27
IRI	0.5	6.21	1.68	-
TOTAL CRASHES	0	34	3.65	25038
FATAL+INJURY	0	11	0.70	4806
PROPERTY DAMAGE ONLY	0	32	2.95	20232

The data used for these analyses covered 7200 km of the highways, 3800 km of which is arterials and 2900km of collectors. From this overview, the highways are kept in reasonable condition with an IRI around 1.5 to 1.6 on average. Arterial highways had more collisions than collector highways and all two-lane highways.

3.2.3. Empirical Bayes Dataset

In order to acquire before and after data for the Empirical Bayes (EB) analysis, the database had to be readjusted to only feature one highway section per row and arranged in a way that would allow for excel spreadsheet calculations. The primary challenge of EB analysis is the selection of before-after sections for evaluations. For this analysis, only sections with one pavement maintenance activity between 2003 to 2010 were selected. This was done for two reasons. One was to ensure that there was enough before and after data included in the comparison (3 years minimum in this case). The second reason was to ensure that only one improvement in IRI was observed because of the possibility that other possible design changes could have been implemented at the same time. The safety impacts of design changes are then thought to be minimized if only one pavement maintenance procedure is considered. The procedure in SAS© is presented in Appendix C: SAS Code for Creating an EB Analysis Database.

4. Methodology

This chapter describes details the methodologies used to determine the Safety Performance Functions (SPFs) and Crash Modification Factors (CMFs) for two-lane undivided highways. The methodologies detailed here are the Negative Binomial Regression (NBR) for SPFs and the Cross-Sectional and EB before-after analysis of IRI (Gross et al., 2010; Hauer, 1997).

4.1. Crash Prediction Models

Crash prediction models in the form of SFPs are used in EB and Cross-Sectional studies to estimate the number of collisions for a specific year (Gross et al., 2010). The SPFs are developed to relate the crash frequency to the characteristics of the site (Gross et al., 2010).

4.1.1. Distribution of Collisions

As discussed in Section 1, the original model format was thought to be multi-linear. Due to the tendency for multi-linear models to create negative values and shortcomings related to identifying factors that significantly affect the collision frequency the model was deemed to be inadequate (Jovanis & Chang, 1986).

The general consensus was that there is a higher chance of collisions at higher traffic volumes. If this positive linear relationship holds true, it will imply that the mean estimate would intersect with the mean of the data at all times; however, as Jovanis & Chang (1986) explain, as the traffic volume increases so does the variance of the collision frequency, violating the assumption made in linear regression. As Hauer (2015) points out, the shaping the model to the data is of primary concern; and parameter estimation is only secondary.

The Generalized Linear Model (GLM) with Poisson or Negative Binomial (NB) error structure is the most accepted formulation for developing SPFs (Hauer, 2015). The Poisson model has been studied at length and has been deemed inadequate. This was because observation data will be a subject of extended periods of zero observation becoming overdispersed and violating the Poisson assumption of variance equaling the mean (Abdel-Aty & Radwan, 2000; Hauer, 2001; Miaou, 1994; Tehrani et al., 2017). The Gaussian NB alternative was chosen by these researchers as it better fits the data better.

4.1.2. Generalized Linear Models

Since collisions should not be predicted as negative, the log-linear model has been assumed. The resulting GLM equation, which is also known as the SPF, involves a simple combination of variables as inputs. The general model structure is hence:

$$\mu = e^{\beta_0} e^{\beta_n x_n} \tag{5}$$

Negative Binomial Regression (NBR) has been the preferred method of identifying the β parameters of the covariates in the SPF equation. Although, NBR is more general than the Poisson method and requires more extensive computations. This method can be applied to many different combinations of variables of interest (Abdel-Aty & Radwan, 2000; Miaou, 1994). Several software packages can perform the required regression.

SAS© Enterprise Guide was used in this study (SAS, 2018). SAS© fits the GLM, utilizing the maximum likelihood estimation (MLE) of the parameter vector(s) in an interpretive process. The dispersion parameter (k) is also determined by the maximum likelihood estimate (SAS, 2018).

Lastly, the SPF must be calibrated for the yearly variation in observed collisions. The yearly variation is calculated by dividing the SPF prediction for each year by the observed number of collisions in the same year. The resulting factor can then be applied to the general SPF to estimate crashes for that year.

4.1.3. Dispersion Parameter (k)

In statistical regression, the modelling of k is done for two main reasons: 1) To indicate how dispersed the accident counts are around the mean of the estimated collision frequency 2) to serve as a weight factor in the estimation of the safety effect of treatment in the Empirical Bayes (EB) before-after study (Hauer, 2001).

The assumption that the value of k is not fixed was suggested by Heydecker and Wu (2001). Allowing k to vary according to the explanatory covariates would offer a better representation of the reality of the specific data set (Hauer, 2001). Miaou and Lord (2003) for example, found that the assumption of a fixed k underestimates collisions by 35% at individual intersections in Toronto.

Abdel-Aty et al. (2009) observed that SPF covariates varied considerably if the fitting was done using the whole data set rather than models estimated from arbitrarily chosen section lengths. The issues stem from the assumption that model covariates do no change when accident counts are assumed to be NB distributed. In doing so, the other estimates of the distribution only affect the precision to which the covariates are estimated (Hauer, 2001).

In general, the higher the accident counts for a road section, the more influence it should have on the estimate of the model parameters. When counts are assumed to be NB distributed with a fixed k, data points with a mean larger than the k will cause a significant increase in the variance of those accident counts. The increased variance of data with higher accident counts will have an incorrect effect for estimating the most likely value of the model parameters (Hauer, 2001).

This effect k has also extends into the EB methodology. Hauer (2001) has demonstrated that under the assumption of a constant k, comparing two roads, one with a length of one unit and another with n sections of various lengths that would add up to one unit, and have the same errors. This changed the log-likelihood function used to predict the covariates. If the k is now a function of segment length L and its coefficient β_k (common to all the population), then the log-likelihood is now (Hauer, 2015):

$$\begin{aligned} \ln[\quad |\mathcal{L}^*(\beta_0, \dots \beta_n \dots, \beta_k)] \\ = \sum_{i=1}^{n} \big[ln\Gamma(k_i + \beta_k L_i) - ln\Gamma(\beta_k L_i) + \beta_k L_i \ln(bL_i) \\ + k_i \ln\big(\hat{E}\{u_i\}\big) - (\beta_k L_i + k_i) \ln\big(\beta_k L_i + \hat{E}\{u_i\}\big)\big] \end{aligned} \quad (6)$$

4.2. Empirical Bayes Before-After Study

Crash modification factors show the changes in safety while filtering out the otherwise unobservable effects of factors that influence safety outside of the boundaries of the treatment (Gross et al., 2010). The most common and desirable method to obtain Crash Modification Factors (CMFs) is a before-after study. Three most common varieties are (Abdel-Aty et al., 2009):

- Naïve before-after
- Before-after with a comparison group

- Before-after with the Empirical Bayes approach

The effectiveness of the prediction depends on the ability to discern if there is an actual safety improvement or if there is a matter of regression-to-the-mean bias. In order to account for the regression-to-the-mean effect in estimating treatment effectiveness, it is best to compare the safety of the treated site to that for the same site had the treatment not been implemented. The Empirical Bayes (EB) approach accomplishes this (Abdel-Aty et al., 2009; Persaud & Lyon, 2007).

Hauer (1997) suggests that the following expression for the EB approach:

$$\lambda_i = C_{bi}\, a_i + (1 - a_i)K_i \tag{7}$$

Where

$\lambda_i =$ The expected number of crashes at site i in a given time period

$C_{bi} =$ The estimated number of crashes at site i in the time period as calculated by the SPF

$K_i =$ The observed number of crashes at site i before treatment in a given time period

$a_i =$ Weight factors used to combine the observed crashes (K_i) with the predicted crashes C_{bi} for the before period

The weight factor a is estimated from the mean and the variance of the SPF as follows:

$$a_i = \frac{1}{1 + \frac{k_i C_{bi}^2}{C_{bi}}} = \frac{1}{1 + k_i C_{bi}} \quad (8)$$

Where

k_i = The over dispersion parameter determined alongside the negative binomial regression of the SPF

For reasons mentioned in Section 4.1.3 the following equations are used to determine the dispersion parameter:

$$k_i = \mathrm{e}^{\beta_0} L_i^{\beta} \quad (9)$$

$$VAR(k_i) = (1 - a_i) k_i \quad (10)$$

Where

L_i = The length of the segment of segment i

Factor r_{ci} is used to determine the ratio between the collision volumes for the before and after periods.

$$r_{ci} = C_{ai}/C_{bi} \quad (11)$$

Where

C_{ai} = The SPF predicted crash frequency for site i over a given after time as calculated by the SPF.

The following equations are used to estimate the expected number of collisions in the after period had treatment not been implemented and its variance:

$$\pi_i = r_{ci} k_i \tag{12}$$

$$VAR(\pi_i) = r_{ci}^2 VAR(k_i) \tag{13}$$

The safety effect (θ, CMF) is then calculated as:

$$\theta = \frac{\frac{\lambda_{sum}}{\pi_{sum}}}{\left[1 + \frac{VAR(\pi_{sum})}{\pi_{sum}^2}\right]} \tag{14}$$

$$StDev(\theta) = \sqrt{\frac{\theta^2 \left(\frac{Var(\pi_{sum})}{\pi_{sum}^2} + \frac{Var(\lambda_{sum})}{\lambda_{sum}^2}\right)}{\left(1 + \frac{Var(\lambda_{sum})}{\lambda_{sum}^2}\right)^2}} \tag{15}$$

Where

λ_{sum} = Is the sum of all expected crases at all sites (0 to i)

π_{sum} = Is the sum of the converted before period crashes for all sites (0 to i)

4.2.1. EB Database Selection

Pavement maintenance is usually conducted when the pavement deteriorates below a set threshold, and there is enough funding and demand. In EB analysis, there must be a defined before and after period for comparison. In order to ensure that this is the case, data from sections that received maintenance in the first and last three years between 2000 and 2013 were excluded from the EB analysis. Furthermore, sections with two or more pavement maintenance activities were not considered in the EB analysis.

4.3. Cross-sectional Studies

One of the major shortcomings of Empirical Bayes before-after studies is that they cannot be used in cases where there is no record of before or after data and/or few treatment sites. In such cases, cross-sectional studies are the preferred method of determining the Crash Modification Factors (CMFs) (Gross et al., 2010). As the name implies, the study compares the collisions experienced at the treated site(s) to the collisions experienced at similar untreated sites in the same time frame.

It is necessary to ensure that the treated and non-treated sites are as similar as possible regarding characteristics that could impact collision frequency to yield reliable results. If this is done, it is possible to directly calculate the ratio between the treated and untreated collision frequencies, yielding the CMF. In practice, this is a difficult requirement to meet (Gross et al., 2010).

Since the Safety Performance Function is used to account for all variables that could affect safety, the CMF can be derived from the model covariates (Gross et al., 2010). To do so, the treated and untreated sites should be combined for the model development stage to generate a singular model which includes, among other variables that would affect safety, an indicator variable that would identify whether or not the site received any treatment or a continuous indicator variable indicating the magnitude of treatment (AASHTO, 2010b). The log-linear model with a Negative Binomial (NB) distribution is the suggested method of accounting for the yearly variation of crash frequency. Once the SPF is created, the CMF can then be determined using the exponent of the variable associated with the treatment, given that the model is log-linear (Carter, Sinivasan, Gross, & Council, 2012; Gross et al., 2010).

4.4. Goodness of Fit

Several metrics can be used to evaluate the goodness of fit of any statistical model. The most common ones used for Safety Performance Functions (SPFs) are summarized below.

4.4.1. Mean Prediction Bias (MPB)

The mean prediction bias is the mean of the difference between the predicted and observed values of interest (collision frequency). An MPB value of zero indicates that the model neither underestimates or overestimates the collision frequency. An MPB value greater than zero indicates that the model overestimates collision frequency, while an MPB value below 0 indicates that the model underestimates collision frequency (Qin, 2016). The general formula is:

$$MPB = \frac{\sum_{i=1}^{n}\left(\widehat{Y_i} - Y_i\right)}{n} \tag{16}$$

Where

n = Number of samples

$\widehat{Y_i}$ = Predicted value

Y_i = Observed value

4.4.2. Mean Absolute Deviation (MAD)

The mean absolute deviation is the mean of the absolute difference between the predicted and observed values of interest (collision frequency). Unlike MPB, MAD calculates the average magnitude of the variability between the predicted and observed values. Values closer to zero imply that there is little or no variability between the observed value and the model prediction (Qin, 2016). The general formula is:

$$MAD = \frac{\sum_{i=1}^{n} \left| \widehat{Y_i} - Y_i \right|}{n} \tag{17}$$

Where

n = Number of sample

$\widehat{Y_i}$ = Predicted value

Y_i = Observed value

4.4.3. Mean Squared Predictor Error (MSPE)

The mean squared prediction error is the squared difference between the predicted and observed values of interest (Qin, 2016). The general formula is:

$$MSPE = \frac{\sum_{i=1}^{n} \left(\widehat{Y_i} - Y_i \right)^2}{n_{val}} \tag{18}$$

Where

n_{val} = Validation sample size

$\widehat{Y_i} =$ Predicted value

$Y_i =$ Observed value

4.4.4. Mean Squared Error (MSE)

The means squared error is the squared difference between the predicted and observed values of interest divided by the sample size and the number of variables in the model (Qin, 2016).

$$MSE = \frac{\sum_{i=1}^{n}\left(\widehat{Y_i} - Y_i\right)^2}{n_{sample} - r} \tag{19}$$

Where

n_{sample} = Validation sample size

$r =$ Number of variables in the prediction model

$\widehat{Y_i} =$ Predicted value

$Y_i =$ Observed value

Similar MSE and MSPE values being close to each other indicates that the deterministic and stochastic components are stable between the comparison values (Begum, 2008).

4.4.5. Cumulative Residual Plots (CURE Plots)

Cumulative residual plots showed the cumulative difference between the observed and predicted values (residuals) plotted against one of the model covariates stacked in increasing order. Residuals plotted within the 95% confidence intervals lines indicate that the model is a good fit. Significant drops or increases of the residuals indicate outliers (Hauer, 2015; Qin, 2016).

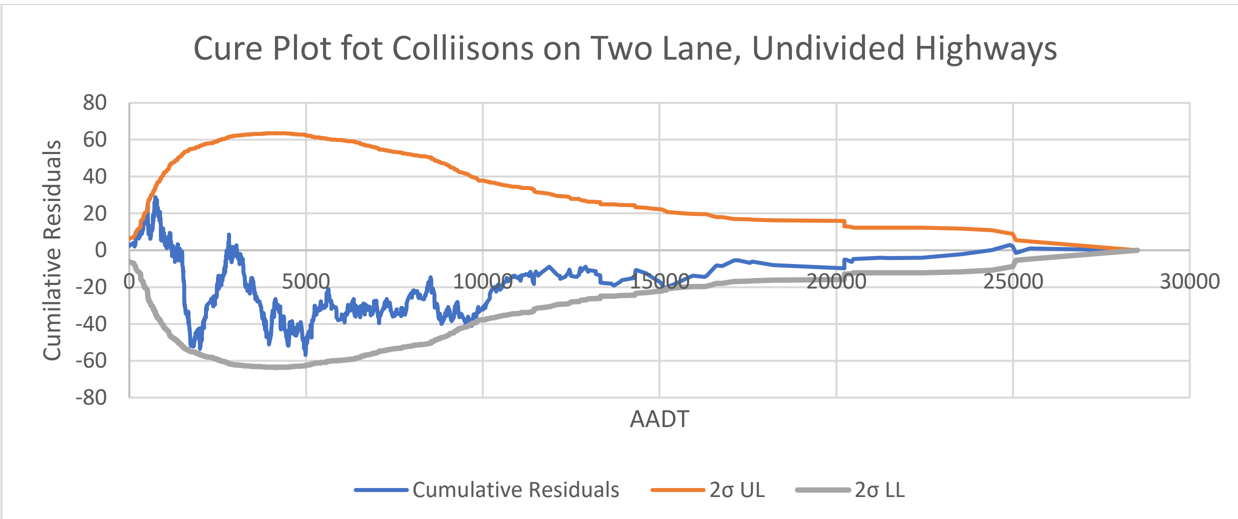


FIGURE 4. SAMPLE CURE PLOT

5. Results

5.1. Safety Performance Functions and Covariates

The following SPF model structure was used to predict the model covariates:

$$\frac{\boldsymbol{Collisions}}{\boldsymbol{Year}} = \boldsymbol{e^{\beta_0} AADT^{\beta_1} Length^{\beta_2} IRI^{\beta_3}} \tag{20}$$

$$\textbf{k} = \boldsymbol{e^{\beta_4} Length^{\beta_5}} \tag{21}$$

The factors for two-lane undivided roadways can be seen in Table 5. A significance level of 5% was used for each parameter of the collision prediction model.

TABLE 5. TWO-LANE UNDIVIDED COVARIATES

CRASH TYPE	TOTAL (TOT)		FATAL + INJURY (FI)		PROPERTY DAMAGE ONLY (PDO)	
PARAMETERS	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
$\boldsymbol{\beta_0}$	-4.962	0.072	-7.707	0.124	-4.938	0.078
$\boldsymbol{\beta_1}$	0.595	0.008	0.715	0.013	0.569	0.008
$\boldsymbol{\beta_2}$	0.866	0.007	0.924	0.013	0.853	0.008
$\boldsymbol{\beta_3}$	0.074	0.016	0.126	0.025	0.063	0.017
$\boldsymbol{\beta_4}$	-1.035	0.061	-1.237	0.195	-0.896	0.064
$\boldsymbol{\beta_5}$	-0.038	0.029	-0.005	0.089	-0.048	0.031

The factors for two-lane arterial roadways can be seen in Table 6. AADT and Length were significant for all severity levels. Only the Fatal + Injury and International Roughness Index parameters were statistically significant at the 5% level. The parameter insignificance may be due to the influence of Property Damage Only collisions, as IRI is insignificant in that case and must be significant enough to influence total collisions.

Table 6. Two-Lane Undivided, Arterial Covariates

CRASH TYPE	TOTAL (TOT)		FATAL + INJURY (FI)		PROPERTY DAMAGE ONLY (PDO)	
PARAMETERS	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
$\boldsymbol{\beta_0}$	-4.943	0.111	-7.723	0.188	-4.933	0.120
$\boldsymbol{\beta_1}$	0.606	0.012	0.726	0.020	0.582	0.013
$\boldsymbol{\beta_2}$	0.833	0.010	0.913	0.017	0.814	0.010
$\boldsymbol{\beta_3}$	0.014	0.021	0.069	0.033	-0.001	0.022
$\boldsymbol{\alpha_2}$	-1.081	0.082	-1.221	0.252	-0.877	0.083
$\boldsymbol{\beta_4}$	-0.111	0.040	-0.037	0.113	-0.161	0.041

The factors for collector two-lane undivided highways can be seen in Table 7. Each parameter for the collision prediction model was significant at the 5% level.

Table 7. Two-Lane Undivided, Collector covariates

CRASH TYPE	TOTAL (TOT)		FATAL + INJURY (FI)		PROPERTY DAMAGE ONLY (PDO)	
COEFFICIENT	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
$\boldsymbol{\alpha}$	-4.714	0.116	-7.543	0.197	-4.666	0.125
$\boldsymbol{\beta_1}$	0.549	0.012	0.686	0.020	0.518	0.013
$\boldsymbol{\beta_2}$	0.897	0.013	0.913	0.023	0.894	0.014
$\boldsymbol{\beta_3}$	0.200	0.026	0.232	0.041	0.196	0.029
$\boldsymbol{\alpha_2}$	-1.091	0.094	-1.294	0.334	-1.066	0.106
$\boldsymbol{\beta_4}$	0.088	0.045	0.027	0.158	0.137	0.050

Models for local two-lane highways were considered, but the sample was too small to draw any reliable conclusions.

5.2. The Goodness of Fit Measures

Table 8 presents the Goodness of Fit (GOF) statistics of the SPFs presented in Section 5.1.

For two-lane undivided highways (Table 8), the MPB shows that the SPF slightly underpredicts collision frequency. The MAD also shows instances of overprediction, but overall, the predictions of the model are close to real life observations. The MSPE and MSE show that there are instabilities between the stochastic and deterministic components within the comparison, especially the Total and PDO collisions.

TABLE 8. GOF: TWO-LANE UNDIVIDED N=16942

GOF MEASURE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
MPB	-0.006	-0.001	-0.002
MAD	2.222	0.706	1.927
MSPE	11.384	1.123	8.495
MSE	11.386	1.123	8.497

For two-lane undivided arterial highways (Table 9), the MPBs show that the SPFs slightly underpredict collision frequency. These slight differences are negligible in real life applications. PDO collisions were overpredicted by the model. The MAD also shows instances of overprediction, but overall, the predictions of the model are close to real life observations. The MSPE and MSE show that there are instabilities between the stochastic and deterministic components within the comparison, especially the TOT and PDO collisions.

TABLE 9. GOF: TWO-LANE UNDIVIDED, ARTERIAL, N=8653

GOF MEASURE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
MPB	-0.016	-0.002	0.060
MAD	2.381	0.795	2.055
MSPE	12.535	1.365	9.107
MSE	12.540	1.366	9.110

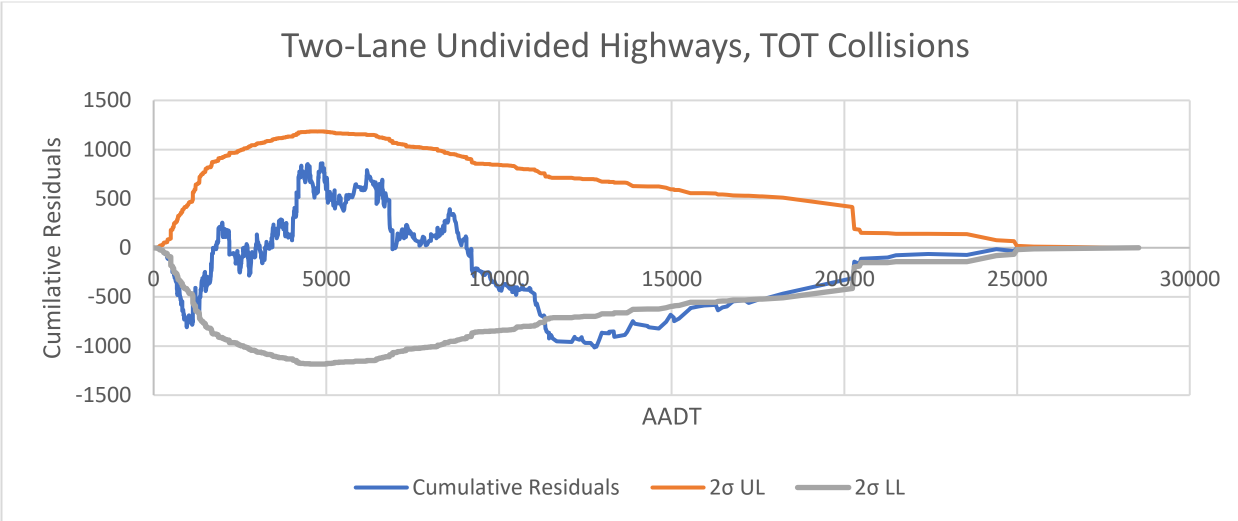
For two-lane undivided collector highways (Table 10), the MPB shows that the SPFs slightly under-predict collision frequency. This slight difference is negligible in real life applications. The MAD also shows instances of overprediction, but overall, the predictions of the model are close to real life observations. The MSPE and MSE show that there are instabilities between the predicted and observed collisions within the comparison, especially the TOT and PDO collisions.

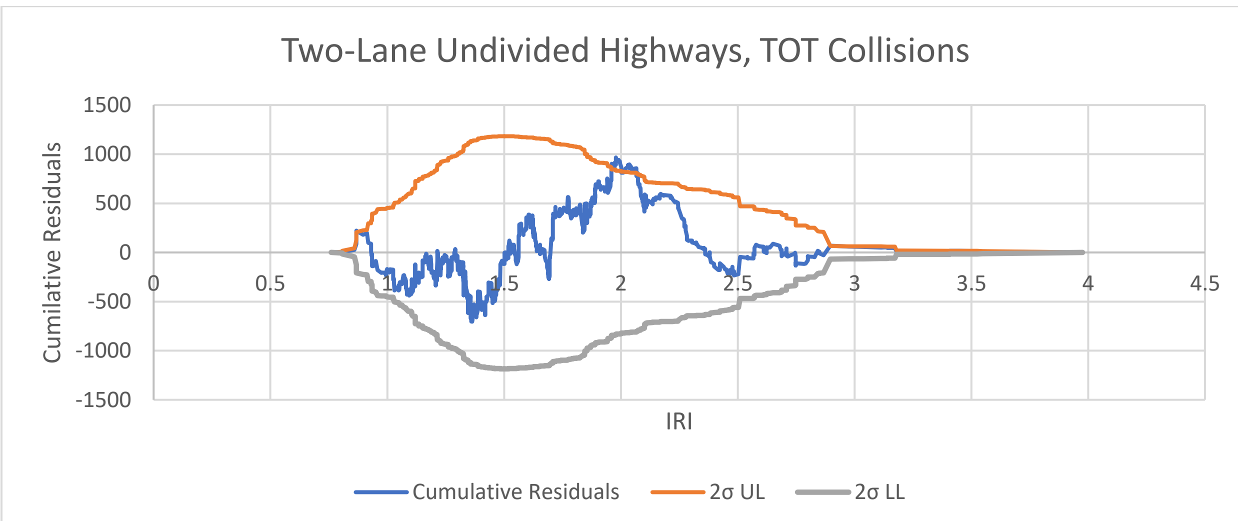
TABLE 10. GOF: TWO-LANE UNDIVIDED, COLLECTOR, N=6856

GOF MEASURE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
MPB	-0.003	-0.051	-0.003
MAD	2.225	0.661	1.947
MSPE	11.371	0.951	8.803
MSE	11.376	0.952	8.807

The CURE plots for total collisions on all two-lane undivided highways are shown in Figure 5. Overall, the AADT fit shows good oscillation but runs outside the 95% confidence boundary at the tail end of the observations. The sharper oscillations are the locations of the majority of the data, between 0 to 12500 AADT. Most of the model indicates a good fit, however, the observations become less dense past 12500 AADT.

For the IRI, the CURE plot indicates a very good fit with only a portion of the cumulative residuals outside the 95% limits around an IRI of 2. Although there are some outliers in the data, they are not significant enough to affect the overall performance of the model.



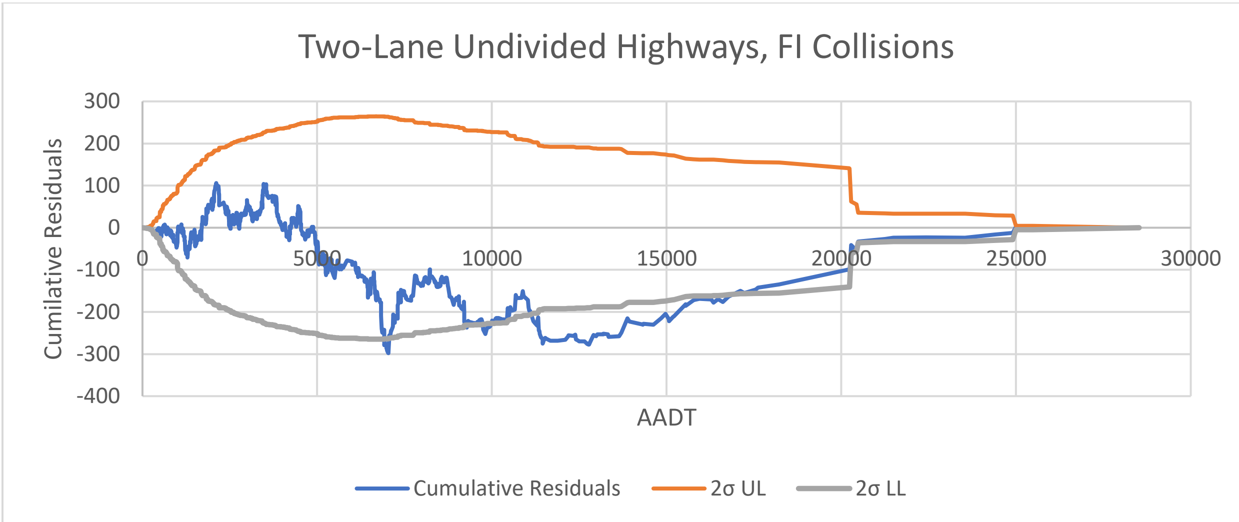


s

FIGURE 5. TWO-LANE UNDIVIDED HIGHWAYS, TOT COLLISIONS

The CURE plots for fatal and injury collisions on all two-lane undivided highways are shown in Figure 6. Overall, the AADT fit shows good oscillation with only some points outside the 95% confidence boundary between 12500 and 17500 AADT. A significant spike was observed at 4000 AADT, indicating a significant outlier.

For the IRI CURE plot, the model indicates a very good fit for the 95% limits. A significant spike in cumulative residuals was observed at an IRI of 1.5, this was possibly due to an outlier, causing the cumulative residuals to exceed the 95% confidence boundary.



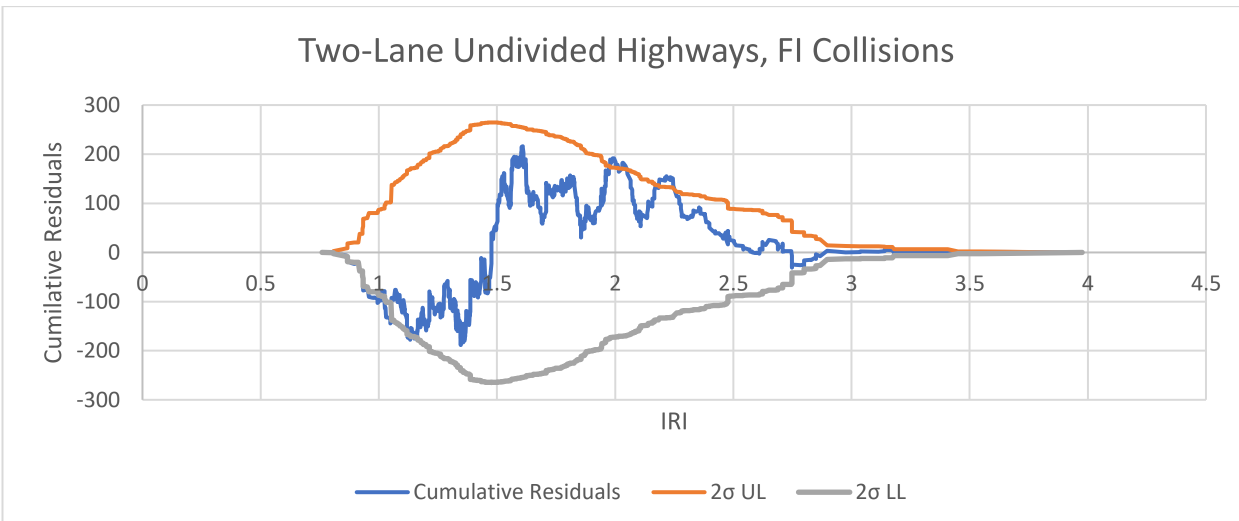
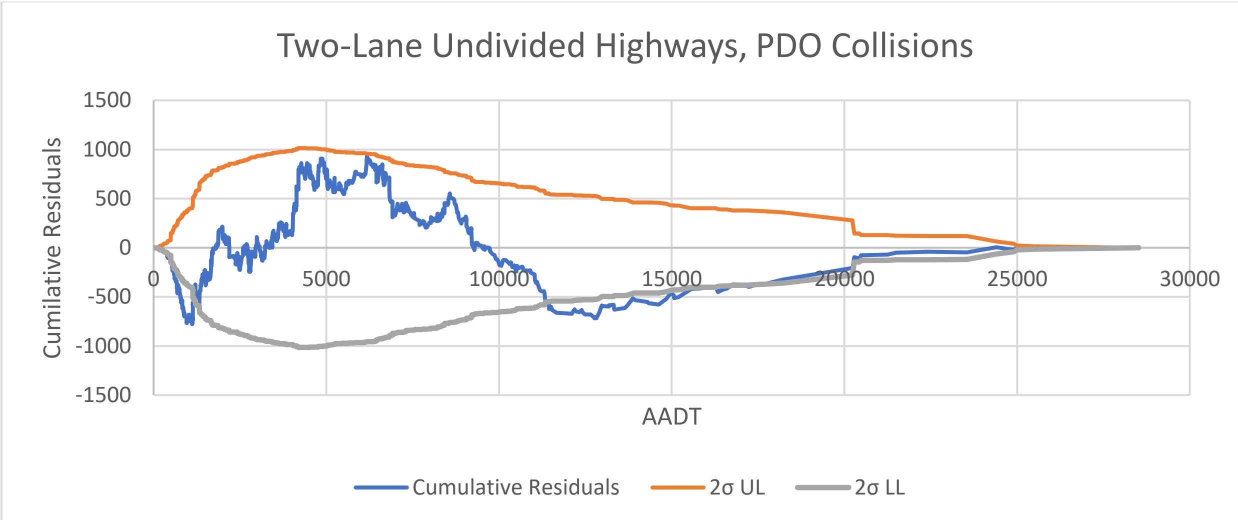


Figure 6. Two-Lane Undivided Highways, FI Collisions

The CURE plots for property damage collisions on all two-lane undivided highways are shown in Figure 7. Overall, the CURE plot for AADT shows a good fit for the 95% confidence intervals. The results reveal an outlier at the beginning of the AADT observations. There are two deviations from the confidence limits, one at 1000 AADT and another at 12500 to 17000 AADT range.

For the IRI CURE plot, the model shows a reasonably good fit, however, some of the predictions run outside the 95% confidence at IRI of 2. The oscillations of the cumulative residuals are excellent, and the only notable location for a possible outlier is at approximately 1.75 IRI.



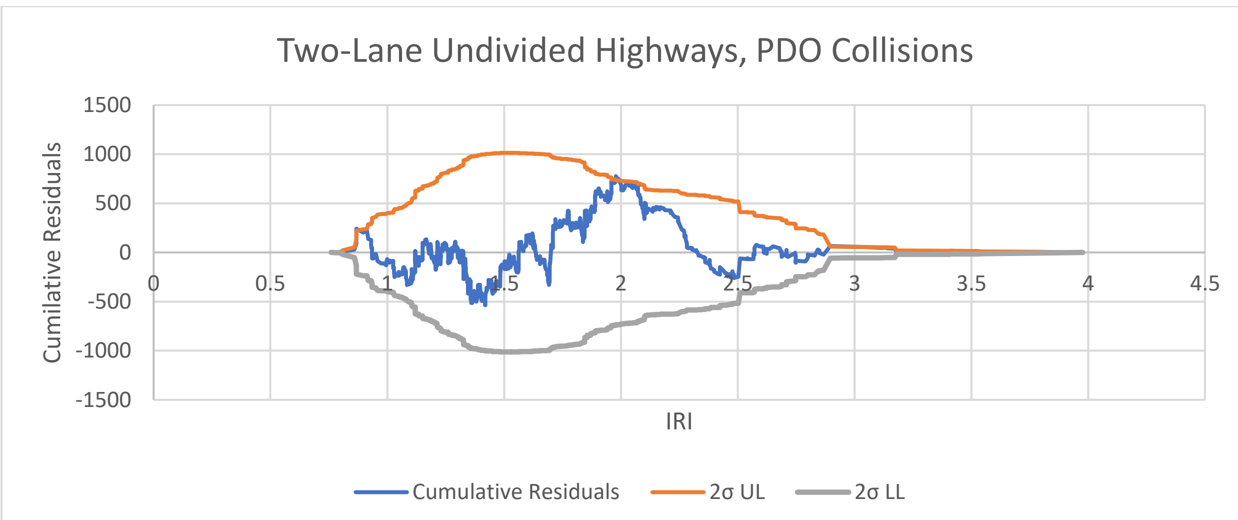
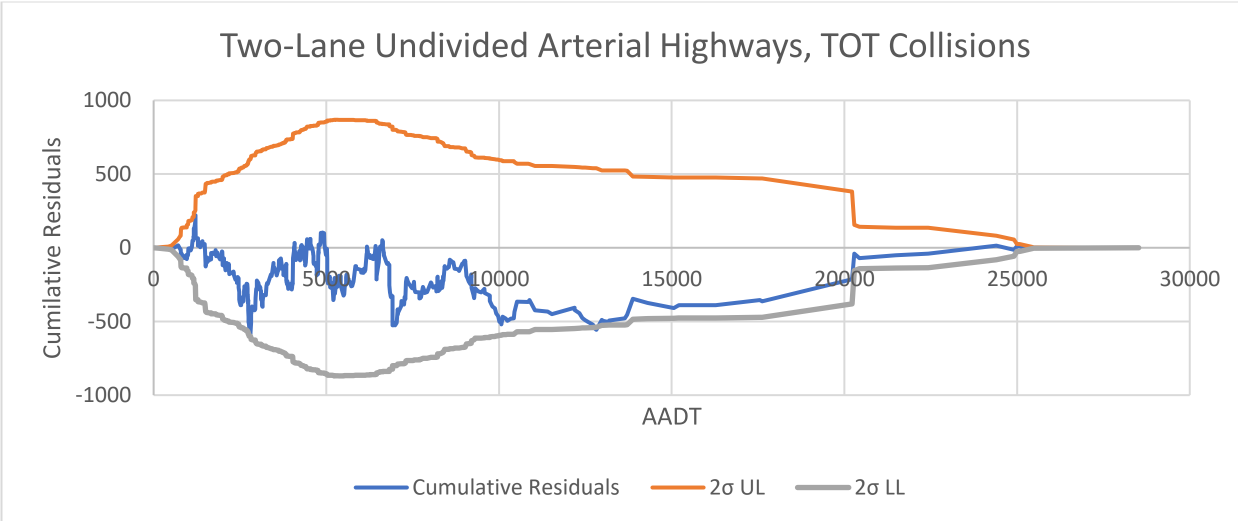


FIGURE 7. TWO-LANE UNDIVIDED HIGHWAYS, PDO COLLISIONS

The CURE plots for total collisions on arterial two-lane undivided highways are shown in Figure 8. The AADT cumulative residuals remained within the 95% confidence limits. The osculation between the observations is reasonable, with the except of a point around 2500 AADT where there is a significant drop almost outside of the 95% confidence limits.

For the IRI CURE plot residuals, the model shows an overall good fit, however, some of the predictions run outside the 95% confidence at an IRI of 2. It is possible that this is the result of the large outliers located at an IRI of 1.5 and an IRI of 1.2.



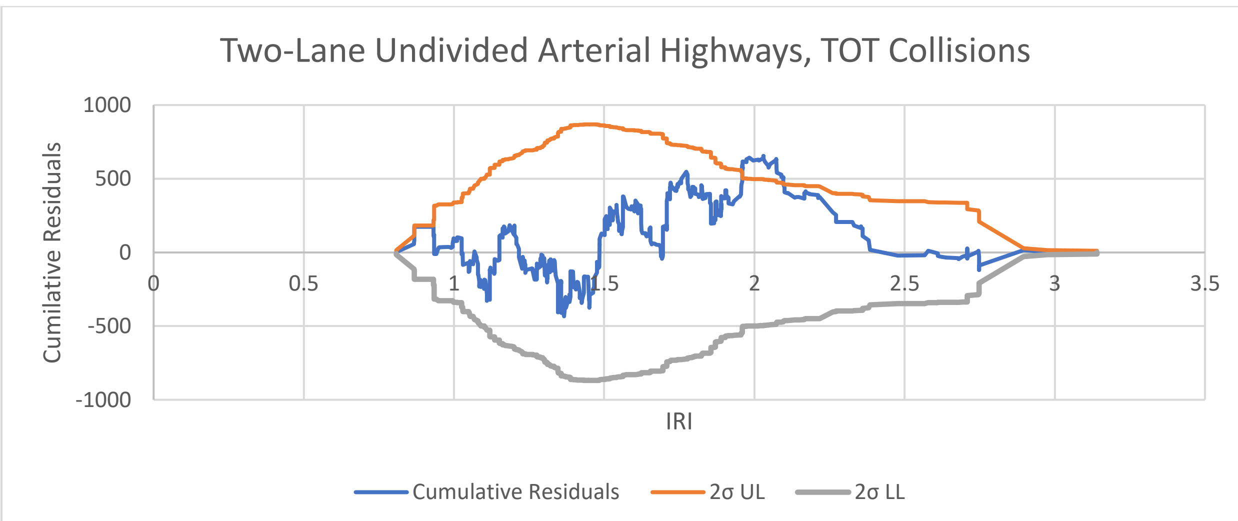
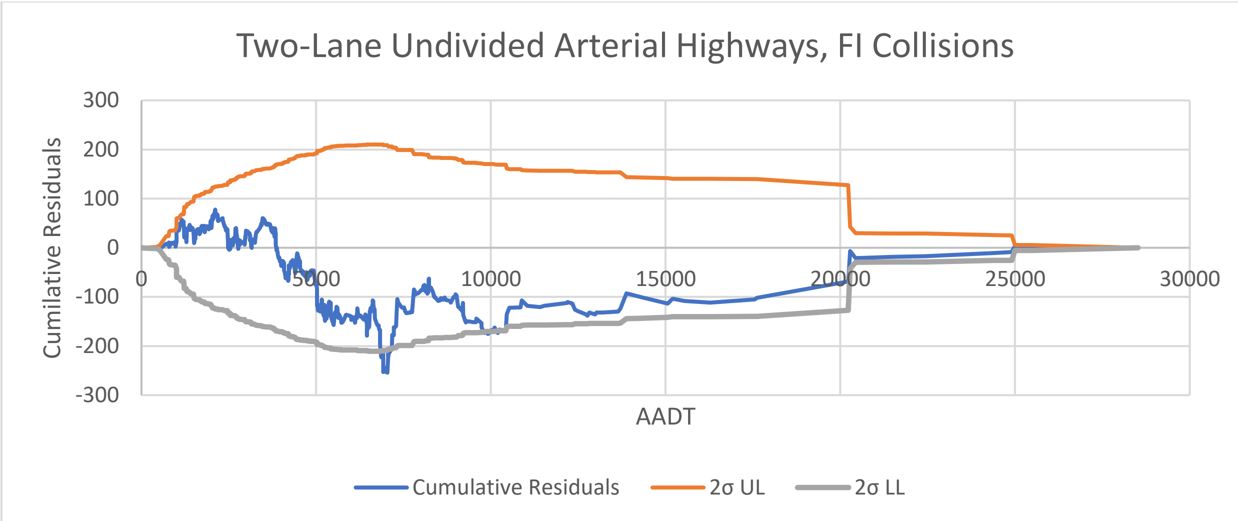


FIGURE 8. TWO-LANE UNDIVIDED ARTERIAL HIGHWAYS, TOT COLLISIONS

The CURE plots for fatal and injury collisions on arterial two-lane undivided highways are shown in Figure 9. The AADT CURE plot indicates a very good fit, with most of the residuals staying inside the 95% confidence interval boundaries. At around 7000 AADT there is a large drop in residuals, indicating the location of a potential outlier. The outlier also causes the model to fall outside of the 95% confidence interval for the same portion.

For the IRI CURE plot, the model shows a near perfect fit. Outliers exist around 1.5 to 1.6 IRI, but their effects were not significant enough to push the cumulative residuals outside of the 95% significance threshold.



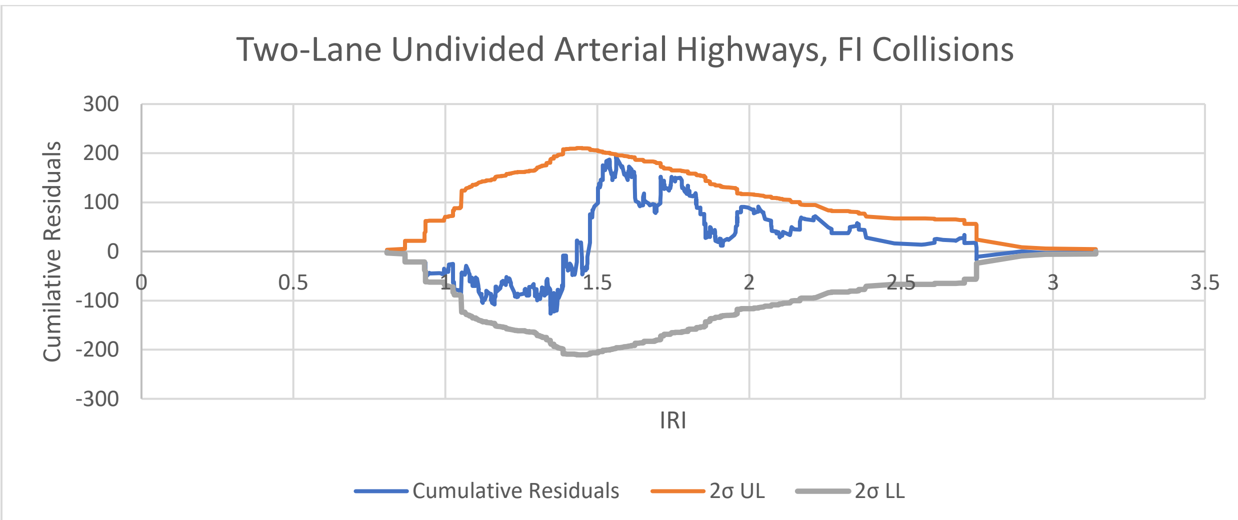
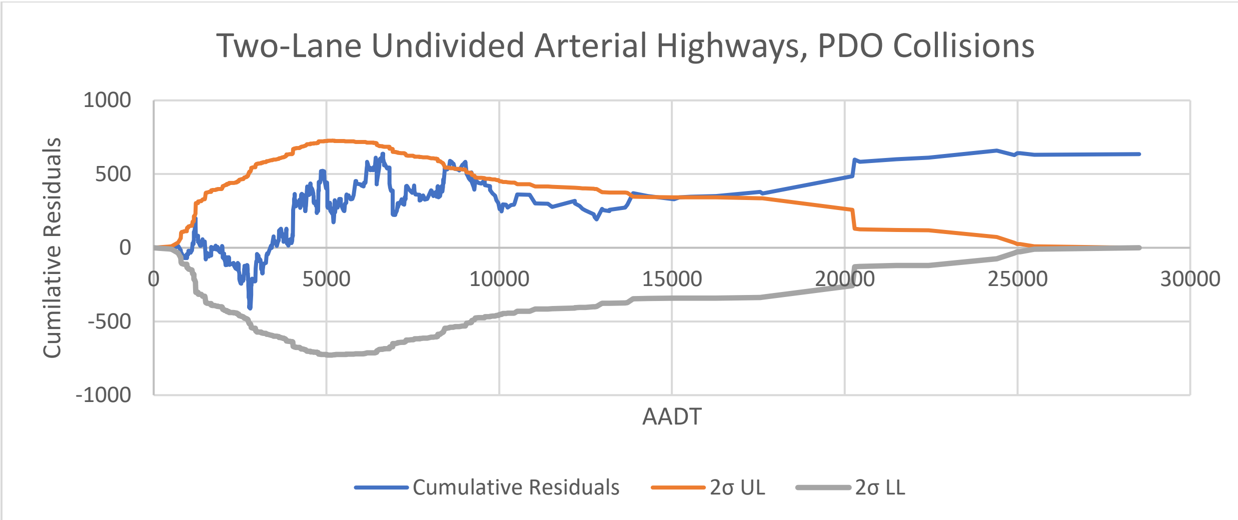


FIGURE 9. TWO-LANE UNDIVIDED ARTERIAL HIGHWAYS, FI COLLISIONS

The CURE plots for property damage collisions on arterial two-lane undivided highways are shown in Figure 10. The AADT CURE plot indicates a severe overestimation in the data set that is causing the results to be well outside of the 95% confidence limits for AADT higher than 16000. The overestimation is mostly due to the abnormal result of the IRI covariate, which unlike other cases is negative.

For the IRI CURE plot, the model shows a good fit for IRI up to 1.7. However, the cumulative residuals exceed the 95% confidence after 1.7 IRI. Once again this is the possible influence of the negative covariate of IRI.



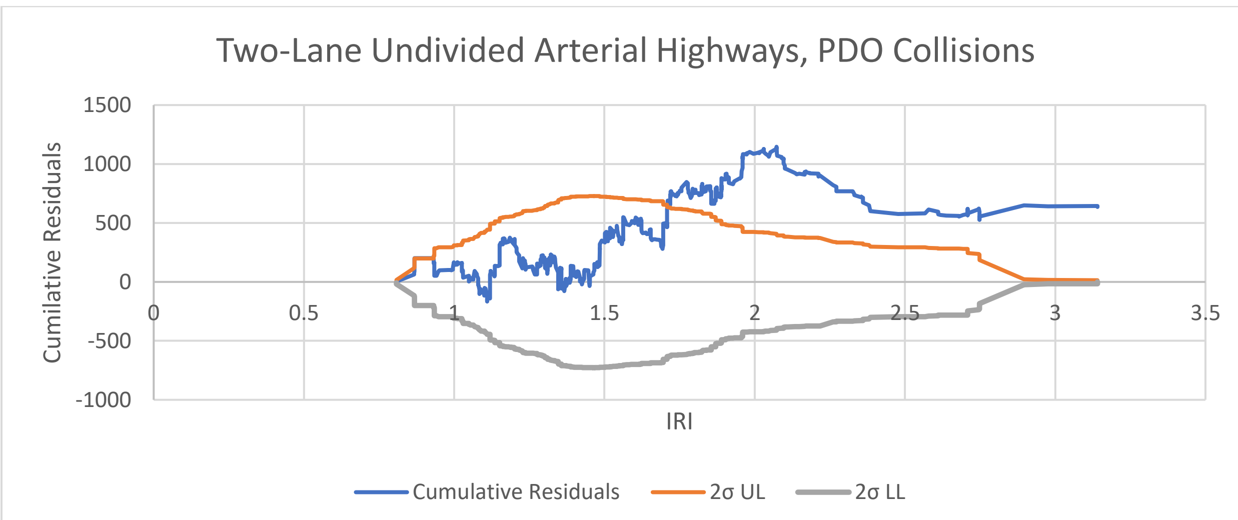


FIGURE 10. TWO-LANE UNDIVIDED ARTERIAL HIGHWAYS, PDO COLLISIONS

The CURE plots for total collisions on two-lane undivided collector highways are shown in Figure 11. The AADT CURE plot shows a significant fit. As seen in Figure 5 and Figure 7 the outliers responsible for this result are possibly located in this dataset. The significance was also greatly influenced by the outlier located around 2000 AADT, which ensured that the predictions remained within the 95% confidence interval up until 12000 to 15000 AADT where they briefly strayed outside of the 95% confidence interval.

For the IRI CURE plot, the model shows a near perfect fit. The oscillations and the lack of significant outliers that impact the cumulative residuals ensured that the model remained significant for all the IRI values.

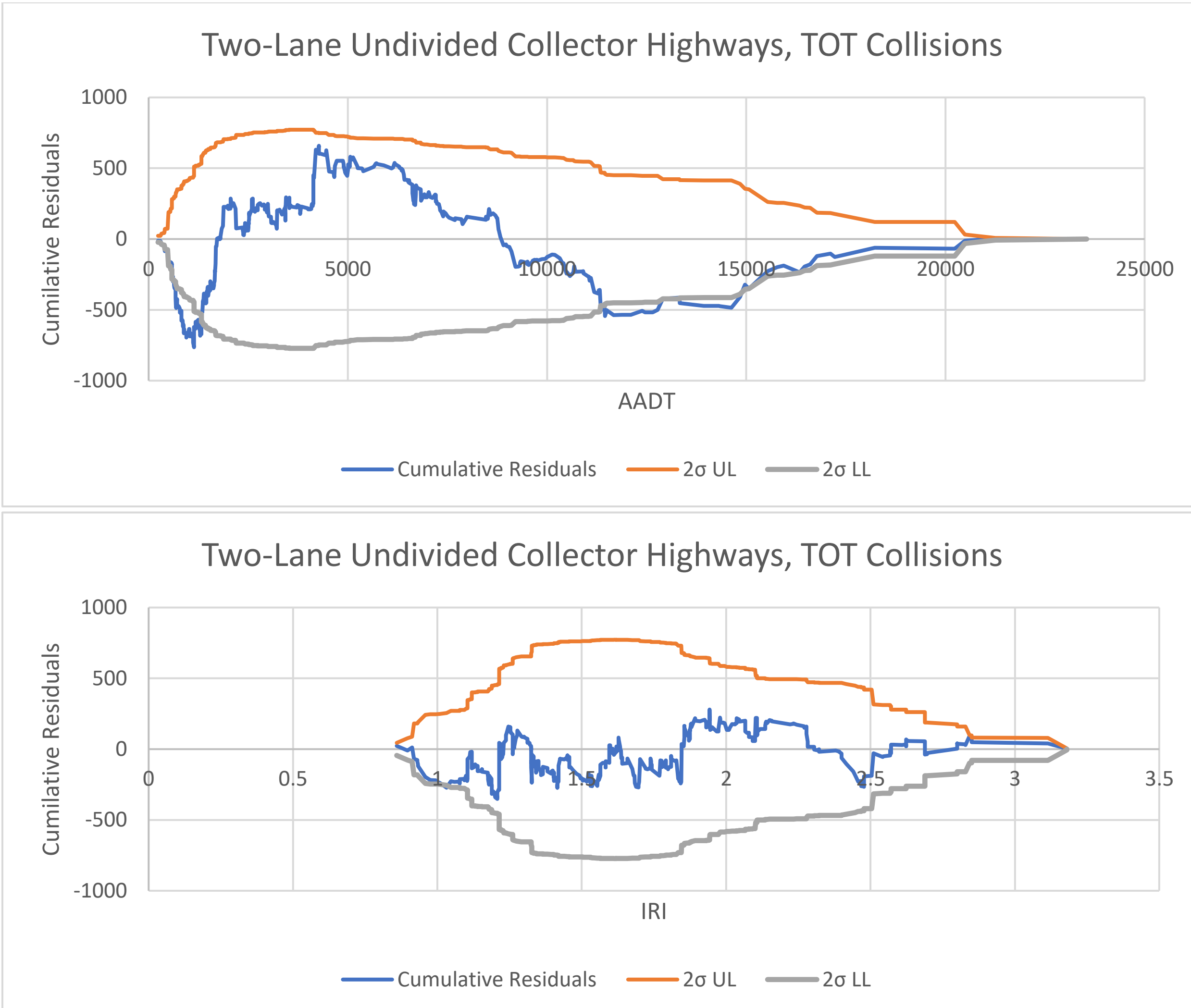


FIGURE 11. TWO-LANE UNDIVIDED COLLECTOR HIGHWAYS, TOT COLLISIONS

The CURE plots for fatal and injury collisions on two-lane undivided collector highways are shown in Figure 12. The AADT CURE plot shows a good fit for the 95% confidence intervals, however, there is still some deviation from the confidence limits around the 11000 to 15000 AADT range. The deviation is probably due to a trend of multiple outliers as far back as the 6000 to the 7500 AADT range, followed by another one at 11000 AADT.

For the IRI CURE plot, the model shows a near perfect fit. Between 0.8 and 1.2 IRI, the cumulative residuals fall in and out of the 95% confidence interval. The large outlier at 1.2 IRI ensured that the cumulative residuals remained significant for the rest of the IRI range.

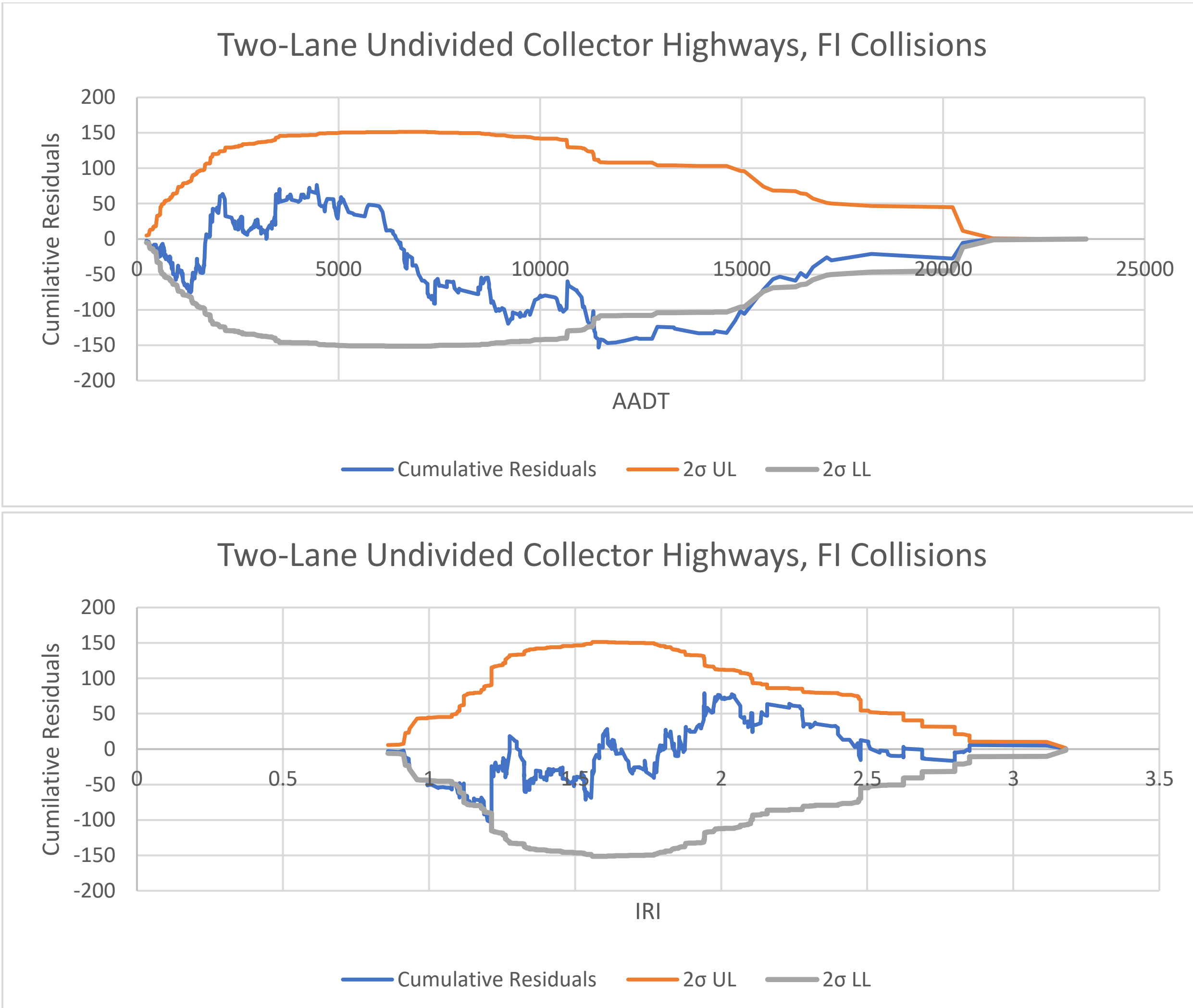
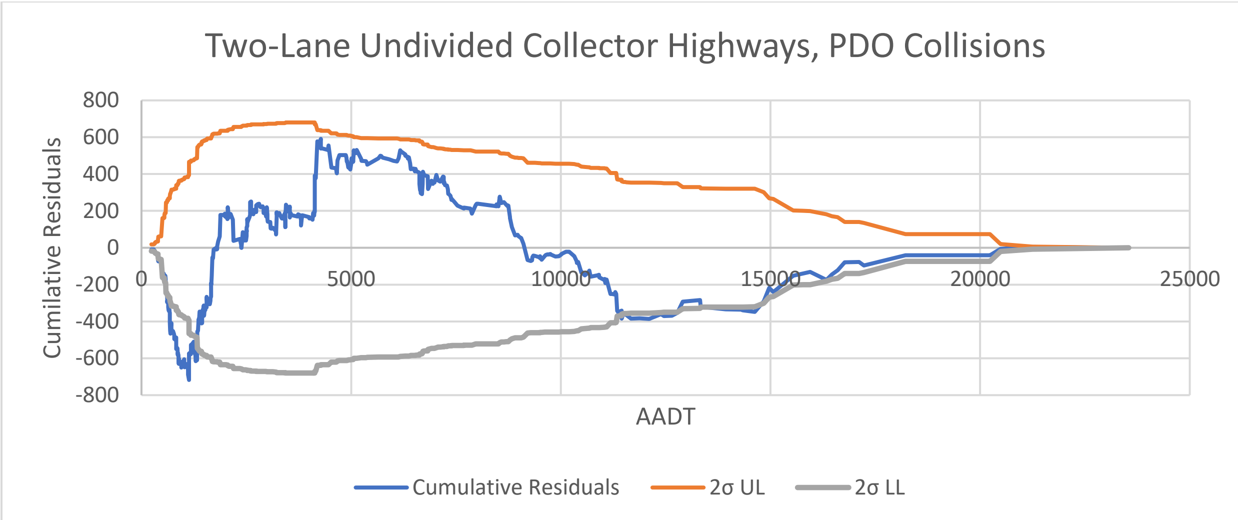


FIGURE 12. TWO-LANE UNDIVIDED COLLECTOR HIGHWAYS, FI COLLISIONS

The CURE plots for property damage collisions on two-lane undivided collector highways are shown in Figure 13. The AADT CURE plot suggests a good fit for the range of observations, however, significant outliers in the 0 to 3000 AADT range bring the residuals outside of the 95% confidence interval and then back. In the 12000 to 15000 AADT range, the cumulative residuals lie on the lower boundary of statistical significance.

For IRI CURE plot, the overall fit is near perfect as most of the cumulative residuals remains inside the 95% confidence intervals with no indication of any significant outliers.



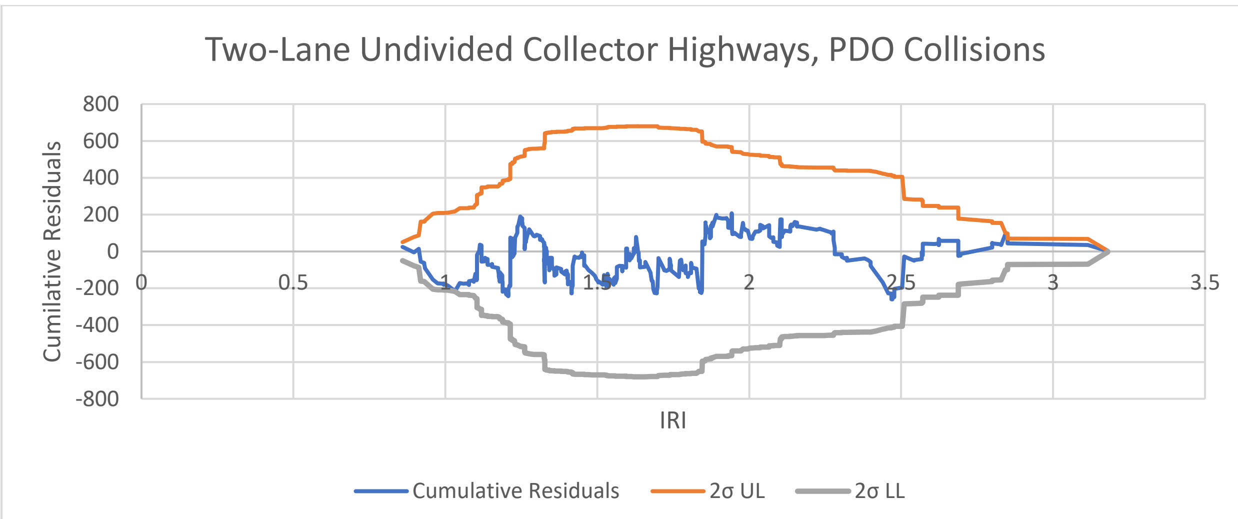


FIGURE 13. TWO-LANE UNDIVIDED COLLECTOR HIGHWAYS, PDO COLLISIONS

5.3. Empirical Bayes (EB) Before-After Analysis

Before the EB analysis can be accomplished, the Safety Performance Functions have to be adjusted for yearly variations in the collision frequency. The coefficients are calculated by dividing the yearly observed collisions by the yearly predicted collisions from the SPFs. The coefficient used are as follows for each year, highway function class and crash severity:

TABLE 11. SPF YEARLY CALIBRATION FACTORS

	ALL			**ARTERIAL**			**COLLECTOR**		
YEAR	TOT	FI	PDO	TOT	FI	PDO	TOT	FI	PDO
2000	0.946	1.128	0.900	0.926	1.092	0.900	0.959	1.128	0.920
2001	0.942	1.084	0.906	0.935	1.078	0.906	0.950	1.089	0.917
2002	1.081	1.179	1.055	1.037	1.106	1.055	1.139	1.277	1.106
2003	1.143	1.201	1.127	1.125	1.195	1.127	1.173	1.208	1.164
2004	1.121	1.151	1.112	1.138	1.156	1.112	1.095	1.149	1.082
2005	1.110	1.175	1.092	1.084	1.128	1.092	1.158	1.268	1.131
2006	1.052	1.080	1.043	1.044	1.036	1.043	1.063	1.115	1.050
2007	1.093	1.063	1.099	1.102	1.090	1.099	1.092	1.012	1.111
2008	1.037	0.961	1.055	1.034	0.947	1.055	1.035	0.977	1.048
2009	0.956	0.892	0.971	0.962	0.927	0.971	0.948	0.836	0.975
2010	0.880	0.809	0.896	0.897	0.838	0.896	0.852	0.760	0.874
2011	0.625	0.591	0.633	0.656	0.639	0.633	0.582	0.524	0.596
2012	0.988	0.857	1.020	1.016	0.884	1.020	0.951	0.840	0.978
2013	1.055	0.914	1.089	1.085	0.953	1.089	1.022	0.884	1.055

The EB study results in Table 12 shows that for two-lane undivided highways, performing pavement maintenance activities have the potential to reduce Fatal + Injury collisions by 12%. These results are statistically significant at the 5% significance level. Total and Property Damage Only collisions are unaffected. Zeng et al., (2014) also observed a decrease in crash severity with no effect on overall collision frequency.

TABLE 12. EB RESULTS: TWO-LANE UNDIVIDED HIGHWAYS

COLLISION TYPE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
OBSERVED “AFTER” COLLISIONS	8433	1735	6698
EXPECTED “AFTER” COLLISIONS	8489.50	1977.35	6869.19
CMF	0.99	0.88	0.97
SD	0.01	0.02	0.01
CMF 95% LOWER BOUND	0.96	0.83	0.94
CMF 95% UPPER BOUND	1.02	0.92	1

For arterials highways (Table.13), the EB results indicate more substantial safety effect of pavement maintenance. The results revealed a 6.5% reduction in TOT collisions, a 12% decrease in FI collisions and a 9% decrease in PDO collision reduction. All these effects are statistically significant at the 5% significance level. This improvement is consistent with the general understanding of driver’s expectation; since these roads are only one step away from being classified as freeways, driver’s expectancies are higher, and thus they expect better surface conditions on these roads.

TABLE 13. EB RESULTS: TWO-LANE UNDIVIDED ARTERIAL HIGHWAYS

COLLISION TYPE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
OBSERVED “AFTER” COLLISIONS	5394	1167	4227
EXPECTED “AFTER” COLLISIONS	5762.80	1319.04	4621.86
CMF	0.94	0.88	0.91
SD	0.02	0.03	0.02
CMF 95% LOWER BOUND	0.90	0.82	0.88
CMF 95% UPPER BOUND	0.97	0.95	0.95

For collector highways (Table 14), the results revealed an 11% increase in overall TOT collisions, a 10% increase in PDO collisions, and an 11% reduction in FI collisions. All these effects are statistically significant at the 5% significance level. The results indicate that for collector highways, pavement maintenance reduces the severity of collisions while increasing TOT and PDO collisions. On such roads, better than expected conditions with more lenient design can be dangerous. Better surface roughness conditions allow for greater control of the vehicle by the driver but do so in a way that emboldens the driver, possible causing them to perform more aggressive maneuvers.

TABLE 14. EB RESULTS: TWO-LANE UNDIVIDED COLLECTOR HIGHWAYS

COLLISION TYPE	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
OBSERVED “AFTER” COLLISIONS	2862	545	2317
EXPECTED “AFTER” COLLISIONS	2585.47	611.61	2094.06
CMF	1.11	0.89	1.10
SD	0.03	0.05	0.03
CMF 95% LOWER BOUND	1.05	0.80	1.04
CMF 95% UPPER BOUND	1.16	0.98	1.17

5.4.Cross-Sectional Analysis

As discussed in Section 4.3, the CMF is taken as the natural exponent of the unit change in the magnitude of treatment. In this case, treatment was the IRI. These CMFs are shown in Table 15.

TABLE 15. CMFS FOR A UNIT INCREASE IN IRI FROM THE CROSS-SECTIONAL ANALYSIS: TWO-LANE UNDIVIDED HIGHWAYS

FUNCTION CLASS:	ALL			ARTERIAL			COLLECTOR		
SEVERITY:	TOT	FI	PDO	TOT	FI	PDO	TOT	FI	PDO
$\boldsymbol{\beta_3}$	0.074	0.126	0.063	0.014	0.069	-0.001	0.200	0.232	0.196
CS CMF	1.08	1.13	1.06	1.01	1.07	1	1.22	1.26	1.22

The CS CMFs findings for all two-lane undivided highways show a 8%/13%/6% increase in Total, Fatal + Injury and Property Damage Only collisions, respectively for a unit increase in IRI. FI collisions are the most sensitive to change in IRI. TOT and PDO collisions are somewhat sensitive. The CS findings for arterial two-lane undivided highways show a 1%/7%/0% increase in TOT, FI, and PDO collisions, respectively for a unit increase in IRI, while the CS findings for collector two-lane undivided highways show a 22%/26%/22% increase in TOT, FI, and PDO collisions, respectively. In sum, collector highways seem to be the most affected by the increase in IRI with all collision severities increased by at least 22%. Also, for every function class the FI collisions were the most affected by the increase in IRI.

The unique aspect of the cross-sectional analysis is that the safety effect can be inferred from any change in IRI. The Empirical Bayes analysis looked at the change in the number of collisions before and after road maintenance has been done. It is possible to calculate the average IRI before and after road maintenance from the EB dataset. With the appropriate β_3 parameter the average CMF of IRI can be calculated before and after road maintenance. Then the before-after ratio of the CMF is the CMF of conducting road maintenance. Table 16 shows the CMFs for pavement treatments conducted on all two-lane undivided highways using CS analysis, while Table 17 shows the CS CMF results for arterial highways and Table 18 shows the CS CMF results for collector highways.

TABLE 16. CROSS-SECTIONAL BEFORE-AFTER ANALYSIS: TWO-LANE UNDIVIDED HIGHWAYS, ALL CLASSES COMBINED

	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
$\boldsymbol{\beta_3}$	0.074	0.126	0.063
BEFORE AVG. IRI/YEAR	2.33	2.33	2.33
AFTER AVG. IRI/YEAR	1.14	1.14	1.14
$(\boldsymbol{Before\ Avg.IRI})^{\beta_3}$	1.06	1.11	1.05
$(\boldsymbol{After\ Avg.IRI})^{\beta_3}$	1.01	1.02	1.01
CMF $= \frac{\boldsymbol{After}}{\boldsymbol{Before}}$	0.95	0.91	0.96

TABLE 17. CROSS-SECTIONAL BEFORE-AFTER ANALYSIS: TWO-LANE UNDIVIDED HIGHWAYS, ARTERIAL

	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
$\boldsymbol{\beta_3}$	0.014	0.069	-0.001
BEFORE AVG. IRI/YEAR	3.93	0.76	3.14
AFTER AVG. IRI/YEAR	3.95	0.78	3.13
$(\boldsymbol{Before\ IRI})^{\beta_3}$	1.01	1.05	1.00
$(\boldsymbol{After\ IRI})^{\beta_3}$	1.02	1.08	1.00
CMF	0.99	0.96	1.00

TABLE 18. CROSS-SECTIONAL BEFORE-AFTER ANALYSIS: TWO-LANE UNDIVIDED HIGHWAYS, COLLECTOR

	TOTAL (TOT)	FATAL + INJURY (FI)	PROPERTY DAMAGE ONLY (PDO)
$\boldsymbol{\beta_3}$	0.200	0.232	0.196
BEFORE AVG. IRI/YEAR	2.49	2.49	2.49
AFTER AVG. IRI/YEAR	1.12	1.12	1.12
$(\boldsymbol{Before\ IRI})^{\boldsymbol{\beta_3}}$	1.20	1.24	1.20
$(\boldsymbol{After\ IRI})^{\boldsymbol{\beta_3}}$	1.02	1.03	1.02
CMF	0.85	0.83	0.86

The previous section had developed Empirical Bayes CMFs for pavement maintenance. In contrast, the present section had developed CMFs for pavement maintenance using CS analysis. This presents a rare opportunity to compare both sets of CMFs in one study. The CMFs for pavement maintenance from Empirical Bayes and cross-sectional analyses are summarized in Table 19.

There is a definite similarity in comparing the CMFs for all two-lane undivided highways. As previously concluded, the EB CMF for FI collisions was the only statistically significant result. Overall EB and CS CMFs are similar with CS CMF underpredicting the effectiveness of the treatment by 3% relative to that for the EB.

The same underprediction of treatment effectiveness was seen for arterial highways. The TOT EB CMF predicted a 5% larger improvement in safety over its CS counterpart, while the CS CMF for PDO collisions showed no improvement in contrast to the 9% improvement indicated by the EB CMF.

TABLE 19. COMPARISON OF CRASH MODIFICATION FACTOR RESULTS FROM EMPIRICAL BAYES AND CROSS-SECTIONAL ANALYSIS

FUNCTION CLASS:	**ALL**			**ARTERIAL**			**COLLECTOR**		
SEVERITY:	TOT	FI	PDO	TOT	FI	PDO	TOT	FI	PDO
EB CMF	-	0.88	-	0.94	0.88	0.91	1.11	0.89	1.10
CS CMF	0.95	0.91	0.96	0.99	0.96	1.00	0.85	0.83	0.86
Δ	-	-0.03	-	-0.05	-0.08	-0.09	0.25	0.06	0.24

Lastly, for collector highways the CS CMF indicates a reduction in collisions while the EB CMF indicates the opposite effect. On the other hand, both the EB and CS were predicting a decrease in FI collisions with the CS CMF indicating a 6% larger reduction in FI collisions than the EB CMF.

The underprediction of the collision-reduction CMFs from the CS analysis is consistent with Hauer (2015) where it is suggested that relying on the fitted parameter model (SPF) to estimate the safety in the unit change would most likely be an underestimation of the safety effect (CMF).

6. Conclusions from Part 1

The findings of this portion of thesis generally indicate a reduction in the severity and, with notable exceptions, the frequency of collisions after pavement maintenance. The Empirical Bayes results showed that treated two-lane undivided highways in Ontario had a 12% reduction in fatal + injury collisions, while arterials experienced the most benefit with crash reductions around 6%-10% at all severity levels. Lastly, collector roads had the most interesting results, with an increase in total and property damage only collisions accompanied by an equal decrease in fatal and severe collisions. Cross-sectional results confirmed that the reduction (improvement) of IRI could have a pronounced effect in reducing collisions severity that can be mathematically related to a change in IRI in estimating the potential safety effect of contemplated IRI improvement.

In retrospect, the SPFs used for EB and CS results can include other parameters that will influence the outcome but picking and choosing parameters to keep and drop is something that requires more work.

Control of the vehicle is an essential element in safe driving supplemented by other forms of built-in safety along the road that would help drivers perceive and avoid danger. In road design conventions less important roadways have a lower building standard with possibly inadequate levels of built-in safety. Without that extra safety net protecting the driver, collision avoidance will boil down to driver experience and how good the vehicle interacts with the pavement surface.

Part Two: Impact of the Local Calibration of Mechanistic-Empirical Pavement Design on Ontario Highways

7. Introduction

Besides road safety, our understanding of pavement mechanics has also improved dramatically. Many large-scale road tests were conducted between 1950 and 1980, but the AASHO Road Test was the most significant of them all.

At that point of time, the principles of pavement structural design have been primarily based on empirical understanding. Despite the excellent performance of the roads it was difficult to say with certainty if the design was built as economically as possible for the purpose it served. Also, it is hard to determine how the growing demand and heavier axle loads would affect the pavement's life expectancy (National Research Council (U.S.) Highway Research Board., 1962).

The AASHO road tests were not intended for innovation but rather to see how pavement performance was influenced by the structural design choice (National Research Council (U.S.) Highway Research Board., 1962). Empirical equations were adopted with the publishing of the Interim Guide for the Design of Rigid and Flexible Pavement in 1961. Subsequence studies since then improved upon those models with the release of the AASHTO Guide for the Design of Pavement Structures in 1993.

Due to the increasing demands for mechanically based pavement design procedures (AASHTO, 2008), NCHRP projects 1-37A and, 1-40 were conducted to lay down the framework for the development of what is now known as The AASHTO Mechanistic-Empirical Pavement Design Guide (MEPDG) (AASHTO, 2008).

MEPDG represents a paradigm shift in pavement design. Unlike its predecessors, it establishes a direct relationship between pavement distress and various design inputs through mechanical analyses and empirical relation models. The complex and thorough design methodology has been packed in a user-friendly working platform now called the AASHTOWare Pavement ME software (previously the DARWin software).

The Ministry of Transportation of Ontario had been using the DARWin design software with a special version of AASHTO Guide for the Design of Pavement Structures improved for Ontario highways since the early 1990's. With the development of the MEPDG Manual of Practice issued by AASHTO in 2008 and the launch of the AASHTOWare MEPDG Design software in 2011, MTO has been working towards adopting it for Ontario highways (Ontario Ministry of Transportation, 2014a).

7.1. Empirical Models for Local Calibration

Although the AAHTOWare MEPDG Design software offers a user-friendly design approach, it leaves something to be desired in the modelling of the local calibration parameters. The software offers six sets of empirical models that can be subjected to local calibration. Models displayed in Table 20 show their respective local calibration coefficients in boldface. Over the past several years, Ryerson research team has been calibrating the coefficients to fit Ontario’s local materials, climatic and traffic conditions, practice in pavement design, maintenance, and rehabilitation. A summary of the local calibration work is presented in the next section.

TABLE 20. EMPIRICAL MODELS FOR LOCAL CALIBRATION

EMPIRICAL MODELS	EXPRESSION	CALIBRATED COEFFICIENTS	SUMMARY STATISTICS
RUTTING	$\frac{\varepsilon_{p,AC}}{\varepsilon_{r,AC}} = k_z \boldsymbol{\beta_{AC}} 10^{-3.3541} T^{1.5606\boldsymbol{\beta_T}} N^{0.4791\boldsymbol{\beta_N}}$ $\frac{\varepsilon_{p,GB}}{\varepsilon_{r,GB}} = 2.03\boldsymbol{\beta_{BG}}\phi(N,\alpha)$ $\frac{\varepsilon_{p,SG}}{\varepsilon_{r,SG}} = 1.35\boldsymbol{\beta_{SG}}\phi(N,\alpha)$	$\beta_{AC} = 1.7692$ $\beta_T = 1.0$; $\beta_N = 0.6262$ $\beta_{GB} = 0.0968$ $\beta_{SG} = 0.2787$	Bias = 0 SD = 1.0 mm
FATIGUE CRACKING	$N_f = 0.007566 C_V C_H \boldsymbol{\beta_f} \varepsilon_t^{-3.9492\boldsymbol{\beta_\varepsilon}} E^{-1.281\boldsymbol{\beta_E}}$ Bottom-Up: $\mathrm{FC_{bt}}(t) = \frac{100\%}{1+\exp(\boldsymbol{C_1}C_1' - \boldsymbol{C_2}C_2' \log_{10} 100D)}$	All β's remains 1.0 $C_1 = 0.5236$; $C_2 = 0.1404$	Bias = 0; SD = 6.14%
	Top Down: $\mathrm{FC_{top}}(t) = \frac{10560}{1+\exp(\boldsymbol{C_3}C_3' - \boldsymbol{C_4}C_4' \log_{10} 100D)}$	Tried, but not calibrated	NA
THERMAL CRACKING	$A = k_t \boldsymbol{\beta_{t1}} 10^{4.389-2.52\log_{10}(E_{AC}\sigma_m n)}$ $\text{Thermal Cracking(TC)} = \boldsymbol{\beta_{t2}}\Phi\left[\frac{1}{\sigma_d}\log\left(\frac{C_d}{h_{AC}}\right)\right]$	Tried, but not calibrated	NA
REFLECTION CRACKING	$\Delta C_{bend} = \boldsymbol{k_1}\sum AK_b^n\, dN$ $\Delta C_{shear} = \boldsymbol{k_2}\sum AK_s^n\, dN$ $\Delta C_{thermal} = \boldsymbol{k_3}\sum AK_t^n\, dN$ $D_T = \boldsymbol{C_1}\Delta C_{bend} + \boldsymbol{C_2}\Delta C_{shear} + \boldsymbol{C_3}\Delta C_{thermal}$ $RC = \frac{\alpha}{\boldsymbol{C_4}+\exp(\boldsymbol{C_5}\log_{10} D_T)} \times 100\%$	Sensitivity analysis done, not calibrated	NA
IRI	$\mathrm{IRI} = \mathrm{IRI_0} + \boldsymbol{C_1}\mathrm{RD} + \boldsymbol{C_2}\mathrm{FC} + \boldsymbol{C_3}\mathrm{TC} + \boldsymbol{C_4}\mathrm{SF}$ Where SF is site factor	$C_1 = 55.096$ (c.f. 40.0) $C_2 = 1.088$ (c.f. 0.400) $C_3 = 0.008$ (global) $C_4 = 0.015$ (global)	Bias = 0; SD = 0.30 m/km

Of the six empirical models, only the rutting, bottom-up fatigue cracking, and IRI have been currently calibrated for use on Ontario highways. The local calibration results were are described and discussed in details by Yuan and Lee (2017). Trial calibration was attempted for top-down fatigue cracking, and thermal cracking models but not reflective cracking. There are current discussions about revising those models, once the new models are finalized, databases such as the one developed by (Yuan & Lee, 2017) can be used to find the parameters.

7.2. Summary of Previous Work

Although the mechanistic concept offers a more realistic methodology for predicting pavement models, there is much work associated with addressing the margin of error (AASHTO, 2010a). Just like Safety Performance Functions (SPF), more confidence is instilled when the predicted pavement designs performance mirrors what is happening in the field (closing the error gap between theory and reality).

The MEPDG design software comes preinstalled with "global" calibration parameters for the distress models. Municipalities with little to no funding can therefore still obtain acceptable results without large investment required for local calibration efforts. However, just like SPFs, the performance cannot be accurately predicted unless modelling/calibration procedures are taken to account for the local nuances specific to the region. As expected, preliminary studies applying the "global" model parameters for Ontario's pavement did not accurately predict the distress and performance observed on Ontario highways (Yuan & Lee, 2017).

To address this problem, MTO had commissioned Ryerson University for three major projects since 2010 under the MTO Highway Infrastructure Innovation Funding Program (HIIFP). The first project focused on the development of local calibration databases needed for analyses and preliminary calibration. The second project expanded the calibration database to include pavement sections and focused mainly on the calibration of the rutting models. The third and last project continued focus on the local calibration of the cracking model and the international roughness index (IRI) using the more accurate performance data collected by the new ARAN 9000 system.

The significant findings and calibration results are presented in Yuan & Lee (2017), and the significant changes in database development are discussed in Jannat (2012). Intermediate local calibration results for the rutting, fatigue cracking and IRI models are reported in (Gautam, Yuan, Lee, & Li, 2016; Jannat, Yuan, & Shehata, 2016; Waseem & Yuan, 2013). The reflective cracking model for rehabilitated pavements was not investigated, partly because the model was recently modified by the model developer. Meanwhile, it is also interesting for one to understand how much the local calibration may bring to practical pavement design.

To bridge the knowledge gap, the primary objectives of this part of the study were to perform a sensitivity analysis of the recently revised reflective cracking model in MEPDG and to examine the impact of the locally calibrated distress models on pavement design for Ontario highways.

8. Sensitivity Analysis of the Reflective Cracking Model

Reflective Cracking (RC) is defined as the cracking of the resurfacing/overlay layer due to underlying cracks or joints. As the underlying layer moves, the cracks propagate upward into the new layer (Zhou et al., 2010). The original model used to predict RC in MEPDG was mainly empirical, based solely on principle of bottom-up fatigue cracking. Miner's cumulative damage theory was replaced with the Paris-Erdogan crack propagation theory in version 2.2.0 of the AASHTOWare software.

The Paris-Erdogan theory states that all types of cracks may be reflected and propagated to the surface of the overlay. The total cumulative critical response parameter (D_T) is therefore estimated (as shown in Table 20), alongside the transfer function. The model itself has eight calibration coefficients: five C's and three k's. The software allows for the changes for all 16 coefficients, as there are eight coefficients for the fatigue cracking model and another eight for the thermal cracking model. The three k values function in the exact same way (and thus carry the same influences) as C_1, C_2 and C_3 for both the fatigue and thermal models. The only relatively independent parameters that can be used for sensitivity analysis are would be the five C's of both models.

8.1.1. Summary of Trial Sections

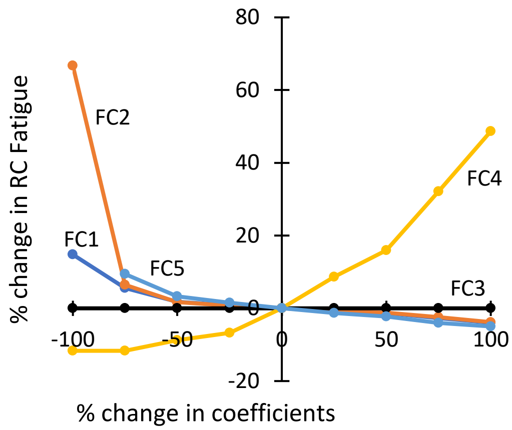
When conducting the sensitivity analysis, the current input level possible for MTO is level 3. Under these conditions, only the transfer function and a portion of the critical response parameter (D_t) can be changed in the AASHTOWare MEPDG Software. To perform the sensitivity analysis, five AC-over-AC sections were selected. Trial sections were selected with parameters diversity in layer combinations, layer thickness, AADT, functional class and location within Ontario. The selected sections are summarized below:

TABLE 21. TRIAL SECTIONS FOR SENSITIVITY ANALYSIS OF REFLECTIVE CRACKING MODEL

SECTION	114	158	477	976	1260
BEGIN LHRS	13640+2.1	14200+0.300	20940+2.1	46969+0	48250+2.7
HWY #	6	7	17	400	403
FUNCTION CLASS	Arterial	Arterial	Arterial	Freeway	Freeway
LOCATION	West	Eastern	North-Eastern	North-Eastern	Central
YEAR	2008	2012	2004	2008	2013
AADT	638	1280	680	2032	16817
LAYER 1(MM)	SP 12.5 FC1(40)	SP 12.5 FC2 (40)	SP 12.5(40)	SP 12.5 FC2(40)	SP 12.5 FC2(40)
LAYER 2(MM)	SP 19 (50)	SP 19 (100)	SP 19 (50)	SP 19 (50)	SP 19 (40)
LAYER 3(MM)	SP 19 (50)	-	CIR (75)	CIR (100)	HBD (200)
LAYER 4(MM)	-	Pulverized Layer (250)	-	-	-
LAYER 5(MM)	Granular A (225)	Old Granular Subbase (450)	HL-4(200)	HL-4(40)	Old Granular Base (300)
LAYER 6(MM)	-	-	Granular A (640)	Granular A (420)	(Old Granular Subbase (300)
SUBGRADE (RESIL. MOD.)	SM (35)	SM (50)	SM-Bedrock (50)	CL-ML (35)	CL (35)

8.1.2. Procedure and Results of the Sensitivity Analysis

To perform the safety analysis, the initial damage to the sections was to set to a degree of accuracy of level 3. The axle load distributions for the year 2015 (the year the database was created) were also imported for all section from the Ministry of Transportation of Ontario's iCorridor website. The preceding equation that calculates the bend, shear and thermal response parameters were taken as a black box, and the behaviour was observed by only manipulating C one through five by changing the base value in the range of -100% to +100% of themselves. The average change in the reflection cracking value was plotted against the percentage change of the coefficients. The results are plotted in Figure 14. The letter "F" in front of the C's represents the coefficients for the fatigue model, while the letter "T" represents the coefficients for the thermal model.



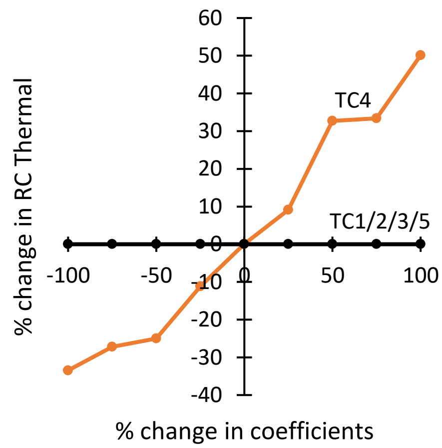


FIGURE 14 SENSITIVITY OF THE LOCAL CALIBRATION COEFFICIENTS OF THE REFLECTION CRACKING (RC) MODELS

The reflection fatigue cracking model indicates that FC2 and FC4 are the most sensitive coefficients and display opposite trends of one another. Overall, the relative percent change in RC fatigue is asymmetric and non-linear. Since FC3 is the weight factor for the fatigue cracking in the D_T model the sensitivity of zero is expected.

The thermal reflection cracking model revealed a completely different pattern than its counterpart. TC4 was the only coefficient that exhibited any sensitivity. However, changes in thermal reflective cracking have been observed with larger changes in TC5. The variation of TC4 among the five sections was negligible.

9. Impact of Local Calibration on Pavement Design

It has already been explored and Yuan and Lee (2017) that, in the current state, the use of the global calibration factors for the Ontario region results in MEPDG designs with overprediction of rutting, underprediction of fatigue, and alligator cracking. It is important to not only produce significant calibration parameters but also to explore what these calibration results tell us. The seven years of work that have been directed into researching these local calibration factors must not stop at just producing significant calibration parameters, but to explore what results the calibrations give us.

9.1.Summary of Trail Sections

Sections that represent various AADTT, function classes, layer thicknesses, subgrade combinations, Ontario regions, and a healthy mix of two rehabilitated (AC-over-AC) sections and three new or reconstructed segments were selected to capture the big picture of the effect that local calibration has on pavement design. The sections are summarized below:

TABLE 22. TRIAL SECTIONS FOR EXAMINING THE IMPACT OF LOCAL CALIBRATION

SECTION	114	698	835	1217	1260
BEGIN LHRS	13640+2.1	29590+3.900	39119+0.0	48140+2.000	48250+2.7
HWY #	6	41	93	402	403
FUNCTION CLASS	Arterial	Arterial	Collector	Freeway	Freeway
LOCATION	West	Eastern	Central	West	Central
YEAR	2008	2007	2004	2003	2013
AADT	638	180	632	6400	16817
LAYER 1(MM)	SP 12.5 FC1(40)	SP 12.5 (40)	SP 12.5 FC1 (40)	SP 12.5 FC2(40)	SP 12.5 FC2(40)
LAYER 2(MM)	SP 19 (50)	SP 19 (50)	SP 19 (100)	SP 19 (120)	SP 19 (40)
LAYER 3(MM)	SP 19 (50)	-	-	SP 25 (180)	HBD (200)
LAYER 4(MM)	-	Pulverized Layer (200)	Granular A (150)	-	-
LAYER 5(MM)	Granular A (225)	Granular A (190)	Granular B1 (450)	Granular A (550)	Old Granular Base (300)
LAYER 6(MM)	-	Old Granular Subbase (150)	Granular A (250)	Granular B1 (375)	Old Granular Subbase (300)
SUBGRADE (RESIL. MOD.)	SM (35)	SM (80)	CL (20)	CL-ML (27)	CL (35)

9.2.Procedure and Results of Calibration Impacts

With the sections chosen and imported into the AASHTOWare MEDPG software, it was time to define the threshold values for IRI, total rut depth, AC rut depth, and allegation cracking. All of which have been taken from the Ontario Default Parameter Guide for MEPDGD (Ontario Ministry of Transportation, 2014a). It is important to note that all five sections were designed using the Ontario's Pavement design and Rehabilitation Manual (Ontario Ministry of Transportation, 2013) and are deemed to satisfy the design requirements set out by the province.

The upper limit for the design life of reconstructed sections (15 to 18 years) and overlay sections (8 to 10 years) were used in this study.

For each trial section, two design iterations were carried out: one using the global calibration models and the other one using the locally calibrated models as specified by Yuan and Lee (2017) (the values are specified in Table 22).

The only design criteria evaluated were IRI, total rut depth, AC rut depth, and alligator fatigue cracking for reasons covered earlier in Section 7.1. In either iteration, when the design criterion was not met, the structural layer (usually the second layer of AC from the top) was increased until each criterion was satisfied. For section 1217, two structural layers can potentially be increased. The increase of the structural layer (red) required to satisfy each criterion has been recorded is presented in Table 23.

Table 23. Impact on MEPDG with and with MEPDG

SECTION	114		698		835		1217		1260	
FUNCTION	Arterial		Arterial		Collector		Freeway		Freeway	
CONSTRUCTION	Overlay		Recon		New		New		Overlay	
STRUCTURE	40 SP12.5FC1 **50 SP19** 50 old SP19 225 Gran A SM Soil		40 SP12.5 **50 SP19** 200 Pulverized 190 Gran A 150 Gran B SM soil		40 SP12.5 FC1 **100 SP19** 150 Gran A 450 Gran B1 CL Soil		40 SP12.5FC2 **120 SP19** **180 SP25** 550 Gran A CL-ML soil		40 SP12.5FC2 **40 SP19** 200 old HDB 300 old Gran A 300 old Gran B Cl soil	
TARGET RELIABILITY	85%		85%		75%		95%		95%	
DESIGN LIFE (YEARS)	10		18		18		18		10	
GLOBAL (G)/CALIBRATED (C)	G	C	G	C	G	C	G	C	G	C
	Increase in AC Layer Thickness (mm) for different distress modes									
IRI	50	0	40	0	0	0	>500	>500	>500	>500
AC BOTTOM-UP CRACKING	0	0	0	0	0	0	0	0	0	0
DEFORMATION- AC ONLY	0	0	0	0	0	0	>500	0	>500	0
DEFORMATION - TOTAL	50	0	20	0	100	0	100	0	300	0

The results revealed an overall benefit, as the sections using the global factors required thicker layers to meet the design requirements. Freeway sections 1217 and 1260 revealed the most significant overall benefit from local calibration, although IRI did not reach a satisfactory level to specify thickness. This can be partly attributed to the partially calibrated IRI model.

Arterial section 114 revealed moderate benefit to IRI and total deformation performance, while the similarly designed arterial section 698 benefited the least from a globally calibrated model. The probable cause is because theTable 22 AADT of section 114 is about five times greater than section 698. Similarly, the two freeway sections share the same observation as section 1260 has about 2.5 times the AADT as section 1217 it is seen that a thicker overall pavement is required to prevent failure due to total deformation. Lastly, sections 114 and 835, although different in many ways, shared similar AADT volumes and both required a 100mm increase if the pavement thickness increase to meet all the distress requirements.

These samples illustrate the impact of local calibration on the predicted performance of the roadway. A lack of local calibration of the designs leads to an unnecessary increase in the capital spent on structural Asphalt Cement. In essence, local calibration helps designers make more informed decisions while reducing unnecessary costs.

10. Conclusions from Part 2

In summary, the sensitivity analysis of the reflective cracking model revealed that factors FC2 and FC4 were the most influential out of the fatigue cracking criteria. The sensitivity analysis of the thermal cracking model revealed that only FC4 coefficient showed any reaction to the change of calibration coefficients. The change in question was however negligible compared to the level of response of the other five parameters. As previously mentioned, this analysis was conducted using level 3 inputs. When more detailed data becomes available for level 1 and 2 inputs, more analysis could be conducted for the local calibration of factors.

The analysis of the impact of local calibration on design indicated that the prediction model yielded better performance results than global values. The IRI model still needs to be fully calibrated as the results indicate highways with high AADT filed to meet the design standard. With new distress models and improvement in data quality steps already being taken that would soon allow us to develop the final set of local calibration factors.

Summary, Recommendation and Next Steps

With the findings from both parts of the research presented it is now appropriate to bridge the gap of understanding between these seemingly two different research fields (road safety and pavement design). It was found that poor quality pavements do indeed have a negative impact on road safety.

The idea drawn out from the results is that pavement management should not be based on economic decisions alone. It was proven that better pavements improve road safety. In the literature review, it was made clear that the pavement maintenance thresholds did not coincide with driver's preference for road roughness regarding IRI. The MTO, for example, has thresholds for roughness that are two times worse than what a driver would even rate as inadequate. MTO and other pavement agencies should consider revising their thresholds and risk assessment to incorporate road safety considerations. Although some roads can be potentially left to degrade further as a cost-cutting measure, it is obvious from the findings that they represent a real hazard to the users.

The study presented here is a new beginning to bringing pavement management and road safety management closer together. Like any other research, this work has two main limitations that should be addressed in future studies. First, is to expand the model parameters of the Safety Performance Functions to improve their predictive capabilities. Second, is to start considering other roads function classes such as freeways.

Lastly, when the MEPDG calibration is finished, exploration should be undertaken to see if its calculated IRI can be incorporated into the models presented in Part 1. Finally, the safety performance degradation along the pavement life can be an interesting future study area as once this degradation curve is established, more safety-conscious pavement management decisions can be made.

Appendix A: SAS Code for Creating a Main Database

The following code was used to make the ROADCHARAADT:

```
DATA aadt1; set aadt;
drop highway;
drop id;
run;
proc sort data=aadt1;
by LHRS1 OFFSET1;
run;

proc sort  data=roadchar;
by LHRS OFFSET;
run;
proc sql;
    create table roadcharaadt as
    select *
    from  roadchar road, aadt aadt1
where road.LHRS = aadt.LHRS1 and
((aadt.OFFSET1 >= road.OFFSET and aadt.OFFSET1 < (road.OFFSET + road.LENGTH)) or
((aadt.OFFSET1 + aadt.LENGTH1)  <= (road.OFFSET + road.LENGTH) and (aadt.OFFSET1 +
aadt.LENGTH1) > road.OFFSET) or
(aadt.OFFSET1 < road.OFFSET and (aadt.OFFSET1 + aadt.LENGTH1) > (road.OFFSET +
road.LENGTH)));
quit;

proc sort data=roadcharaadt;
by LHRS OFFSET;
run;
```

Where XX was replaced by 00-13:

```
proc sql;
    create table rap20XXf as
    select *
    from  roadcharaadt3 rc, pavement20XX pave

where (rc.HWY = pave.routenum and rc.from_d >= pave.begin_mile and  rc.to_d <
pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d > pave.begin_mile and  rc.to_d <= pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d = pave.begin_mile and  rc.to_d = pave.end_mile)
```

```
or (rc.HWY = pave.routenum and rc.from_d < pave.begin_mile and rc.to_d > pave.begin_mile
and  rc.to_d < pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d > pave.begin_mile and rc.from_d <
pave.begin_mile and  rc.to_d > pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d < pave.begin_mile and  rc.to_d = pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d > pave.begin_mile and  rc.to_d = pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d = pave.begin_mile and  rc.to_d < pave.end_mile)
or (rc.HWY = pave.routenum and rc.from_d = pave.begin_mile and  rc.to_d > pave.end_mile)
;
quit;

data rap20XXfm (drop= end_desc to_desc begin_desc from_desc routeaux over_sur_typeg
surface contract_num begin_refp end_refp id length dir secnum routetype begin_mile
end_mile to_lhrs from_lhrs pavetypc aadta pct_trucka pavetype fric routenum rci dmi pci sai
AADT00 AADT01 AADT02 AADT03 AADT04 AADT05 AADT06 AADT07 AADT08 AADT09 AADT10
AADT11 AADT12 AADT13 AADT14);
set rap20XXf;

if routeaux = 'A' then delete;
if routeaux = 'B' then delete;
if pqi = 0 and psi = 0 and pdi = 0 and iri = 0 then delete;
if surf_width ='' or surf_width = 0 then delete;



offset3 = substr(begin_refp, 7, 5);
offset2 = input(offset3, 8.);
drop offset3;

if from_d<begin_mile then fromn_d = begin_mile;
if from_d>=begin_mile then fromn_d = from_d;
if to_d>end_mile then ton_d = end_mile;
if to_d<=end_mile then ton_d = to_d;

if from_d<begin_mile then from_descn = begin_desc;
if from_d>=begin_mile then from_descn = from_desc;
if to_d>end_mile then to_descn = end_desc;
if to_d<=end_mile then to_descn = to_desc;

if from_d<begin_mile then offsetn = offset2;
if from_d>=begin_mile then offsetn = offset;
if to_d>end_mile then offsetn = offset2;
if to_d<=end_mile then offsetn = offset;
drop offset2;
drop offset;
```

```
if terrain = 'FLAT' then do;
terrain = '1';
end;
if terrain = 'ROLLING' then do;
terrain = '2';
end;
if direction^=5 then do;
aadt=AADTXX/2;
end;
if direction=5 then do;
aadt=AADTXX;
end;
lengthn = ton_d-fromn_d;
logaadt = log(aadt);
loglength = log(lengthn);
lane_width = surf_width/num_lanes;
ident=(fromn_d + ton_d + direction)/100 + lhrs;
run;

proc sort;
by LHRS offsetn;
run;

data mvab20XXp; set mvab20XXs;
if RDLOCA= 10 or RDLOCA= 1;
if CLASAC = 1 then F=1; else F=0;
if CLASAC = 2 then I=1; else I=0;
if CLASAC = 3 then PD=1; else PD=0;
if CLASAC = 4 then NRpt=1; else NRpt=0;
if CLASAC = 5 then OCl=1; else OCl=0;
if INTIMP = 0 then IIU=1; else IIU=0;
if INTIMP = 1 then App=1; else App=0;
if INTIMP = 2 then Ang=1; else Ang=0;
if INTIMP = 3 then RE=1; else RE=0;
if INTIMP = 4 then SS=1; else SS=0;
if INTIMP = 5 then TM=1; else TM=0;
if INTIMP = 6 then SMVU=1; else SMVU=0;
if INTIMP = 7 then SMVO=1; else SMVO=0;
if INTIMP = 8 then NApl=1; else NApl=0;
if INTIMP = 9 then OII=1; else OII=0;
if RDSUR1 = 1 then DRY=1; else DRY=0;
if RDSUR1 = 2 then WET=1; else WET=0;
if RDSUR1 = 3 then LS=1; else LS=0;
if RDSUR1=  4 then SLU=1; else SLU=0;
```

```
if RDSUR1 = 5 then PKS=1; else PKS=0;
if RDSUR1 = 6 then ICE=1; else ICE=0;
if RDSUR1 = 7 then MUD=1; else MUD=0;
if RDSUR1 = 8 then LSG=1; else LSG=0;
if RDSUR1 = 9 then SPL=1; else SPL=0;
drop RDSUR1;
drop intimp;
drop clasac;
run;

proc sort data=mvab20XXp  out=mvab20XXps;
by reference_no offset DIRTRA;
run;

PROC MEANS data=MVAB20XXPs NOPRINT; by REFERENCE_NO OFFSET DIRTRA; VAR F I NRpt
PD OCl IIU App Ang RE SS TM SMVU SMVO NApl OII DRY WET LS SLU PKS ICE MUD LSG SPL;

OUTPUT OUT=MVAB20XXsum sum= F I NRpt PD OCl IIU App Ang RE SS TM SMVU SMVO NApl
OII DRY WET LS SLU PKS ICE MUD LSG SPL;

data mvab20XXsum; set mvab20XXsum;
drop _TYPE_;
drop _freq_;
run;

proc sql;
create table db20XXint as
select *
from rap20XXfm rap, mvab20XXsum mvab
where mvab.REFERENCE_NO = rap.LHRS and (rap.direction = mvab.dirtra or rap.direction=5 or
mvab.dirtra=9) and (rap.offsetn <= mvab.offset < (rap.offsetn + rap.lengthn));
quit;

proc sort;
by ident;
run;

proc means data=db20XXint noprint; by ident; VAR F I NRpt PD OCl IIU App Ang RE SS TM SMVU
SMVO NApl OII DRY WET LS SLU PKS ICE MUD LSG SPL;
OUTPUT OUT=db20XXsum sum= F I NRpt PD OCl IIU App Ang RE SS TM SMVU SMVO NApl OII
DRY WET LS SLU PKS ICE MUD LSG SPL;
run;

proc sort data=rap20XXfm out=rap20XXfms;
by ident;
```

```
run;

data db20XX;
update rap20XXfms db20XXsum;
by ident;
run;
data db (drop= from_d to_d _type_ _freq_);
merge db2000 db2001 db2002 db2003 db2004 db2005 db2006 db2007 db2008 db2009 db2010
db2011 db2012 db2013;
by ident year;
run;
```

Appendix B: SAS Code for Creating an Analysis Database

```
data db (drop= from_d to_d _type_ _freq_);
merge db2000 db2001 db2002 db2003 db2004 db2005 db2006 db2007 db2008 db2009 db2010
db2011 db2012 db2013;
by ident year;
run;

data db2;
set db;
array F2 _numeric_;
do over F2;
if F=. then F=0;
end;
array I2 _numeric_;
do over I2;
if I=. then I=0;
end;
array PD2 _numeric_;
do over PD2;
if PD=. then PD=0;
end;
TOT = F + I + PD;
FI = F + I;
run;

data dir5lane2;
set dir5;
if num_lanes ^=2 then delete;
if divided ^= 'NO' then delete;
logaadtn = input(logaadt,8.);
drop logaadt;
speed = input(posted_speed,8.);
drop posted_speed;
terrainn = input(terrain,8.);
drop terrain;
shldrw = input(shld_width,8.);
drop shld_width;
if iri = 0 then delete;
logiri = log(iri);
run;
```

Appendix C: SAS Code for Creating an EB Analysis Database

```
data YYYint;
set dir5lane2;
if FUNC_CLSS ^= 'ZZZ' then delete;/*Omit if no function class*/
if year = 2000 then do;
AADT_00 = AADT;
iri_00 = iri;
TOT_00 = TOT;
FI_00 = FI;
PD_00 = PD;
end;

if year = 2001 then do;
AADT_01 = AADT;
iri_01 = iri;
TOT_01 = TOT;
FI_01 = FI;
PD_01 = PD;
end;

if year = 2002 then do;
AADT_02 = AADT;
iri_02 = iri;
TOT_02 = TOT;
FI_02 = FI;
PD_02 = PD;
end;

if year = 2003 then do;
AADT_03 = AADT;
iri_03 = iri;
TOT_03 = TOT;
FI_03 = FI;
PD_03 = PD;
end;

if year = 2004 then do;
AADT_04 = AADT;
iri_04 = iri;
TOT_04 = TOT;
FI_04 = FI;
```

```
PD_04 = PD;
end;

if year = 2005 then do;
AADT_05 = AADT;
iri_05 = iri;
TOT_05 = TOT;
FI_05 = FI;
PD_05 = PD;
end;

if year = 2006 then do;
AADT_06 = AADT;
iri_06 = iri;
TOT_06 = TOT;
FI_06 = FI;
PD_06 = PD;
end;

if year = 2007 then do;
AADT_07 = AADT;
iri_07 = iri;
TOT_07 = TOT;
FI_07 = FI;
PD_07 = PD;
end;

if year = 2008 then do;
AADT_08 = AADT;
iri_08 = iri;
TOT_08 = TOT;
FI_08 = FI;
PD_08 = PD;
end;

if year = 2009 then do;
AADT_09 = AADT;
iri_09 = iri;
TOT_09 = TOT;
FI_09 = FI;
PD_09 = PD;
end;

if year = 2010 then do;
```

```
AADT_10 = AADT;
iri_10 = iri;
TOT_10 = TOT;
FI_10 = FI;
PD_10 = PD;
end;
if year = 2011 then do;
AADT_11 = AADT;
iri_11 = iri;
TOT_11 = TOT;
FI_11 = FI;
PD_11 = PD;
end;

if year = 2012 then do;
AADT_12 = AADT;
iri_12 = iri;
TOT_12 = TOT;
FI_12 = FI;
PD_12 = PD;
end;

if year = 2013 then do;
AADT_13 = AADT;
iri_13 = iri;
TOT_13 = TOT;
FI_13 = FI;
PD_13 = PD;
end;

run;

proc sort;
by ident lengthn;
run;
proc means data=YYYint noprint; by ident lengthn; var AADT_00 iri_00 TOT_00 FI_00 PD_00
aadt_01 iri_01 tot_01 fi_01 pd_01 aadt_02 iri_02 tot_02 fi_02 pd_02 aadt_03 iri_03 tot_03
fi_03 pd_03 aadt_04 iri_04 tot_04 fi_04 pd_04 aadt_05 iri_05 tot_05 fi_05 pd_05 aadt_06
iri_06 tot_06 fi_06 pd_06 aadt_07 iri_07 tot_07 fi_07 pd_07 aadt_08 iri_08 tot_08 fi_08 pd_08
aadt_09 iri_08 tot_08 fi_08 pd_08 aadt_09 iri_09 tot_09 fi_09 pd_09 aadt_10 iri_10 tot_10
fi_10 pd_10 aadt_11 iri_11 tot_11 fi_11 pd_11 aadt_12 iri_12 tot_12 fi_12 pd_12 aadt_13
iri_13 tot_13 fi_13 pd_13;
output out=YYYsum sum=;
run;
```

References

AASHTO. (2008). *Mechanistic-Empirical Pavement Design Guide: A Manual of Practice* (Interim ed). Washington, D.C.

AASHTO. (2010a). *Guide for the Local Calibration of the Mechanical-Empirical Pavement Design Guide*. Washington DC.

AASHTO. (2010b). *Highway Safety Manual* (Vol. 1st). Washington, D.C.: American Association of State Highway and Transportation Officials.

Abdel-Aty, M., Devarasetty, P. C., & Pande, A. (2009). Safety evaluation of multilane arterials in Florida. *Accident Analysis & Prevention*, *41*(4), 777–788. https://doi.org/10.1016/j.aap.2009.03.015

Abdel-Aty, M., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, *32*(5), 633–642. https://doi.org/10.1016/S0001-4575(99)00094-9

Begum, S. M. M. A. (2008). *Investigation of model calibration issues in the safety performance assessment of Ontario highways*. Ryerson University.

Cantisani, G., & Loprencipe, G. (2010). Road Roughness and Whole Body Vibration: Evaluation Tools and Comfort Limits. *Journal of Transportation Engineering*, *136*(9), 818–826. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000143

Carter, D., Sinivasan, R., Gross, F., & Council, F. (2012). *NCHRP 20-7(314) Final Report - Recommended Protocols for Developing Crash Modification Factors*. Washington, D.C. Retrieved from http://www.cmfclearinghouse.org/collateral/CMF_Protocols.pdf

Chan, C. Y., Huang, B., Yan, X., & Richards, S. (2010). Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS). *Journal of Advanced Transportation*, *44*(3), 150–161. https://doi.org/10.1002/atr.129

Chang, C. M., Nazarian, S., Vavrova, M., Yapp, M. T., Pierce, L. M., Robert, W., & Smith, R. E. (2017). *Consequences of Delayed Maintenance of Highway Assets*. Washington, D.C.: Transportation Research Board. https://doi.org/10.17226/24933

Chen, S., Saeed, T. U., Alqadhi, S. D., & Labi, S. (2017). Safety impacts of pavement surface roughness at two-lane and multi-lane highways: accounting for heterogeneity and seemingly unrelated correlation across crash severities. *Transportmetrica A: Transport Science*, *0*(0), 1–16. https://doi.org/10.1080/23249935.2017.1378281

Chen, S., Saeed, T. U., & Labi, S. (2017). Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. *Analytic Methods in Accident Research*, *16*, 75–89. https://doi.org/10.1016/j.amar.2017.09.001

Elghriany, A., Yi, P., Liu, P., & Yu, Q. (2016). Investigation of the effect of pavement roughness on crash rates for rigid pavement. *Journal of Transportation Safety & Security*, *8*(2), 164–176. https://doi.org/10.1080/19439962.2015.1025458

Gautam, G. P., Yuan, X.-X., Lee, W., & Li, N. (2016). Local Calibration of the MEPDG Rutting Models for Ontario's Flexible Roads: Recent Findings. Washington, D.C.

Gross, F., Persaud, B., & Lyon, C. (2010). *A Guide to Developing Quality Crash Modification Factors*. Retrieved from https://safety.fhwa.dot.gov/tools/crf/resources/fhwasa10032/fhwasa10032.pdf

Hauer, E. (1997). *Observational before-after studies in road safety : estimating the effect of highway and traffic engineering measures on road safety*. Oxford: Pergamon.

Hauer, E. (2000). *Lane width and safety. Literature*. Retrieved from https://www.researchgate.net/publication/261363624_Lane_width_and_safety_Literature

Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis & Prevention*, *33*(6), 799–808. https://doi.org/10.1016/S0001-4575(00)00094-4

Hauer, E. (2015). *The Art of Regression Modeling in Road Safety*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-12529-9

Heydecker, B. G., & Wu, J. (2001). Identification of sites for road accident remedial work by Bayesian statistical methods: an example of uncertain inference. *Advances in Engineering Software*, *32*(10–11), 859–869. https://doi.org/10.1016/S0965-9978(01)00037-0

Hughes, W., Prothe, L., McGee, H., & Hauer, E. (2001). Impact of Resurfacing Projects With an Without Additional Safety Improvements. *National Cooperative Highway Research Program*, (255), 1–10. Retrieved from http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rrd_255.pdf

Jannat, G. (2012). *Database Development for Ontario's Local Calibration of Mechanistic - Empirical Pavement Design Guide (MEPDG) Distress Models*. Ryerson University.

Jannat, G., Yuan, X.-X., & Shehata, M. (2016). Development of regression equations for local calibration of rutting and IRI as predicted by the MEPDG models for flexible pavements using Ontario's long-term PMS data. *International Journal of Pavement Engineering*, *17*(2), 166–175. https://doi.org/10.1080/10298436.2014.973024

Jovanis, P. P., & Chang, H. (1986). Modeling the Relationship of Accidents to Miles Traveled. *In Transportation Research Record, No. 1068, TRB, National Research Council*, *1068*, 42–51. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.462.5206

Jurgens, R., & Chan, J. (2005). Highway Performance Measures for Business Plans in Alberta. In *Annual Conference of the Transportation Association of Canada*. Calgary, Alberta.

Lee, J., Nam, B., & Abdel-Aty, M. (2015). Effects of Pavement Surface Conditions on Traffic Crash Severity. *Journal of Transportation Engineering*, *141*(10), 4015020.

https://doi.org/10.1061/(ASCE)TE.1943-5436.0000785

Loprencipe, G., & Zoccali, P. (2017). Ride Quality Due to Road Surface Irregularities: Comparison of Different Methods Applied on a Set of Real Road Profiles. *Coatings*, *7*(5), 59. https://doi.org/10.3390/coatings7050059

Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, *26*(4), 471–482. https://doi.org/10.1016/0001-4575(94)90038-8

Miaou, S.-P., & Lord, D. (2003). Modelling Traffic Crash-Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes Versus Empirical Bayes Methods. *Transportation Research Record: Journal of the Transportation Research Board*, *1840*(1), 31–40. https://doi.org/10.3141/1840-04

Múčka, P. (2016). Current approaches to quantify the longitudinal road roughness. *International Journal of Pavement Engineering*, *17*(8), 659–679. https://doi.org/10.1080/10298436.2015.1011782

National Research Council (U.S.) Highway Research Board. (1962). *The AASHO Road Test, Report 7*. *Special Report 61A - The History and Description of the Project, Publication No. 1061*. Washington, D.C.

Ontario Ministry of Infrastructure. (2017). *Build ON: 2017 Infrastructure Update*. Retrieved from https://www.ontario.ca/page/buildon-2017-infrastructure-update

Ontario Ministry of Transportation. (1985). *Geometric Design Standards For Ontario Highways*. Downsview, On: Queen's Printer for Ontario.

Ontario Ministry of Transportation. (2004). *AIS MS Access Query User Guide*.

Ontario Ministry of Transportation. (2013). *Pavement Design and Rehabilitation Manual*. *Materials Engineering and Research Office* (Second Edi). Downsview, On: Materials Engineering and Research Office. Retrieved from http://www.bv.transports.gouv.qc.ca/mono/1165561.pdf

Ontario Ministry of Transportation. (2014a). *Ontario's Default Parameters for AASHTOWare Pavement ME Design – Interim Report 2014*. Downsview, On. Retrieved from https://www.library.mto.gov.on.ca/SydneyPLUS/Sydney/ViewRecord.aspx?template=Books&record=da29aab0-7c84-4ff6-ad4b-5dff77b3c5a7&lang=en-US

Ontario Ministry of Transportation. (2014b). *Ontario Road Safety Report 2014*. Toronto, On.

Persaud, B., & Lyon, C. (2007). Empirical Bayes before–after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention*, *39*(3), 546–555. https://doi.org/10.1016/j.aap.2006.09.009

Project Steering Committee. (2016). *Canadian Infrastructure Report Card*. https://doi.org/10.17226/12014

Qin, L. (2016). *SAFETY EVALUATION OF FREEWAY SPEED-CHANGE LANES BASED ON CRASHES AND SIMULATED CONFLICTS*. Ryerson University.

SAS. (2018). SAS Customer Support. Retrieved March 23, 2018, from https://support.sas.com/en/support-home.html

Sayers, M. W., & Karamihas, S. M. (1998). *The Little Book of Profiling*. *The Regent of the University of Michigan*. Michigan. Retrieved from http://www.umtri.umich.edu/content/LittleBook98R.pdf

Sayers, Gillespie, T. D., & Queiroz, C. A. V. (1986). *The International Road Roughness Experiment*. Washington, D.C.: The International Bank for Reconstruction and Development/THE WORLD BANK.

Tehrani, S. S., Falls, L. C., & Mesher, D. (2015). Road users' perception of roughness and the corresponding IRI threshold values. *Canadian Journal of Civil Engineering*, *42*(4), 233–240. https://doi.org/10.1139/cjce-2014-0344

Tehrani, S. S., Falls, L. C., & Mesher, D. (2017). Effects of pavement condition on roadway safety in the province of Alberta. *Journal of Transportation Safety & Security*, *9*(3), 259–272. https://doi.org/10.1080/19439962.2016.1194352

Transport Association of Canada. (2006). *Performance Measures for Road Networks: A Survey of Canadian Use*.

Transport Canada. (2015). *Canadian Motor Vehicle Traffic Collision Statistics 2011*. Retrieved from http://www.tc.gc.ca/media/documents/roadsafety/TrafficCollisionStatisitcs_2011.pdf

Transportation Research Board. (1987). *Relationship Between Safety and Key Highway Features. a Synthesis of Prior Research*. *State-of-the-Art Report*. Retrieved from http://onlinepubs.trb.org/Onlinepubs/state-of-the-art/6/6.pdf%0Ahttps://trid.trb.org/view/282039

University of North Carolina Highway Safety Research Center. (n.d.). CMF Clearing House. Retrieved November 15, 2017, from http://www.cmfclearinghouse.org/faqs.cfm

Waseem, A., & Yuan, X.-X. (2013). Longitudinal Local Calibration of MEPDG Permanent Deformation Models for Reconstructed Flexible Pavements Using PMS Data. *International Journal of Pavement Research and Technology*, *6*, 304–312. https://doi.org/10.6135/ijprt.org.tw/2013.6(4).304

Yuan, X., & Lee, W. (2017). *Ontario's Local Calibration of the MEPDG Distress and Performance Models for Flexible Roads : a summary*. Retrieved from http://www.tac-atc.ca/en/conference/papers/ontarios-local-calibration-mepdg-distress-and-performance-models-flexible-roads-summary

Zeng, H., Fontaine, M., & Smith, B. (2014). Estimation of the Safety Effect of Pavement Condition on Rural, Two-Lane Highways. *Transportation Research Record: Journal of the*

Transportation Research Board, *2435*, 45–52. https://doi.org/10.3141/2435-06

Zhou, F., Lytton, R. L., Hu, S., Luo, R., Tsai, F.-L., & Lee, S. I. (2010). *Models for Predicting Reflection Cracking of Hot-Mix Asphalt Overlays*. Washington, D.C.: National Academies Press. https://doi.org/10.17226/14410