MOUNTAIN-PLAINS CONSORTIUM

MPC 18-344 | M. Ahmed and R. Chalise

Calibration of the Highway Safety Manual's Safety Performance Functions for Rural Two-Lane Highways with Regional Considerations for the Rocky Mountains and Plain Regions

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Calibration of the Highway Safety Manual's Safety Performance Functions for Rural Two-Lane Highways with Regional Considerations for the Rocky Mountains and Plain Regions

Mohamed Ahmed, Ph.D., P.E.
University of Wyoming
Phone: (307) 766-5550
Email: mahmed@uwyo.edu

Rameshwor Chalise, M.Sc.
University of Wyoming
Email: rchalise@uwyo.edu

Department of Civil and Architectural Engineering
University of Wyoming
Laramie, WY 82071

March 2018

Acknowledgements

The funding for this study was provided by the USDOT FHWA to the Mountain Plains Consortium (MPC). Additional matching funds were provided by the Wyoming Department of Transportation (WYDOT). The data involved in this study were provided by the WYDOT.

Disclaimer

ABSTRACT

Over the last two decades, the use of advanced scientific and data-driven statistical methods has been continuously evolving to reduce the frequency and severity of crashes on our roadways. The American Association of State Highway and Transportation Officials (AASHTO) Highway Safety Manual (HSM) was first released in 2010. The manual is considered as a significant milestone in the advancement of the practices of roadway safety analyses. As a part of the 2010 HSM, safety performance functions (SPFs) are detailed as statistical models utilized to predict the expected number of crashes per year for a certain roadway facility as a function of traffic, roadway characteristics, and weather conditions. To better manage roadway safety, SPFs are used in: 1) assessing how safe or unsafe are specific roadway segments, intersections, special facilities, etc., and 2) evaluating how much safety has changed because of specific countermeasures or operational or design interventions. SPFs are commonly broken down into two main categories: simple and full. Simple SPFs typically utilize traffic volumes while full SPFs consider additional factors such as roadway geometry, driver characteristics, or weather conditions. Part C of the 2010 HSM, Predictive Methods, details the calibration procedure for jurisdiction specific SPFs. The primary limitation of the HSM Part C is that the developed SPFs for various road facilities were developed using data from only a few states in the United States, putting their validity and reliability into question as various states or regions are characterized by differing geographic features, weather conditions, and demographics. The primary purpose of this report is to examine the suitability and transferability of the 2010 HSM's SPFs to Wyoming-specific conditions.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The first edition of the American Association of State Highway and Transportation Officials (AASHTO) Highway Safety Manual (HSM) was published in 2010 by a considerable effort of the Transportation Research Board (TRB). The manual has already been proven to be a great asset and a major help for traffic safety researchers and practitioners as it has acted as a bridge between research and practice. It has helped transportation professionals make decisions to enhance traffic safety. Many of the states are already using the HSM along with the AASHTO design guidelines in order to quantify the safety of roadways and intersections. In fact, the HSM is the sole source for quantitative evaluation of roadway safety throughout the nation. The proper use of the HSM in planning, design, operation, and maintenance can bring a remarkable achievement in traffic safety.

Safety performance functions (SPFs) are regression models used to predict the expected number of crashes for a particular geographic space per unit time. The geographic space could be a roadway segment, an intersection, or a special facility. SPFs uses exposure factor e.g., traffic volume and sometimes other roadway characteristics, and weather parameters to predict the average crash frequency. The most common exposure factor used in the SPFs for roadway segments is average annual daily traffic (AADT).

SPFs are used to predict the average crash frequency for a particular location for a given site condition. The predicted crashes can be used with the Empirical Bayes Method to analyze the safety of the roadway system. In this method, crashes can be predicted using past crash experience (i.e., observed crashes) with their respective weight and the predicted crashes from the SPF. An accurate estimation of the safety performance of a roadway facility is crucial for the following safety studies:

1) Network Screening: SPFs can be used to determine the safety performance of a given location (segment or intersection) compared with the average safety performance of the locations with similar characteristics. This would be helpful in identifying the crash prone sites with potential for improvement in traffic safety in the road network.
2) The Safety Effectiveness of Different Countermeasures: After predicting average crash frequency, SPFs could be used alone or with crash history to estimate long-term crash frequency and the safety effectiveness of the different sites' treatments. This approach is useful to determine the most effective cost/safety treatment when multiple alternatives of countermeasures are compared. This approach not only helps to determine the best countermeasure alternative but also helps to quantify the benefits of each treatment approach.
3) Project Evaluation: After utilizing the SPF to evaluate the safety effectiveness of roadway improvements, further future planning, policy making, and programming decisions can be effectively made. SPFs are a critical component for before/after studies using the Empirical Bayes (EB) method, which combines the crash history for a given site with the predicted crashes from the SPF.

The HSM SPFs should be calibrated to account for the local factors and conditions that affect the safety of a roadway or network. The calibrated SPFs using the data from those locations give the predicted average number of crashes over a certain period of time to meet the facility local conditions. The Highway Safety Manual (HSM) provides the process to calibrate the provided SPFs. It is necessary to adjust the HSM SPFs when the conditions at the site of interest are different from the base site conditions mentioned in the HSM. Crash modification factors and calibration factors can be applied to adjust the HSM SPFs or new jurisdiction-specific SPFs can be developed if the site conditions and other additional factors are quite different as recommended in the literature.

1.1 Overview of the HSM

The Highway Safety Manual consists of four primary parts. Part A deals with fundamentals on traffic safety and human factors, and Part B deals with roadway safety and management process. There are nine chapters (Chapter 1 to Chapter 9) that comprise Part A and Part B in the HSM. The first three chapters deal with introduction, human factors, and fundamentals. These chapters highlight the use of the HSM in planning, design, operations, and maintenance activities. Chapter 3 in Part A provides background information needed to apply predictive methods, crash modification factors, and evaluation methods, which are later discussed in Parts B, C, and D of the HSM. Chapters 4–9 are concerned with crash monitoring in existing roadways and suggest steps to reduce crashes and their severities. Figure 1.1 shows the four primary parts of the 2010 HSM.

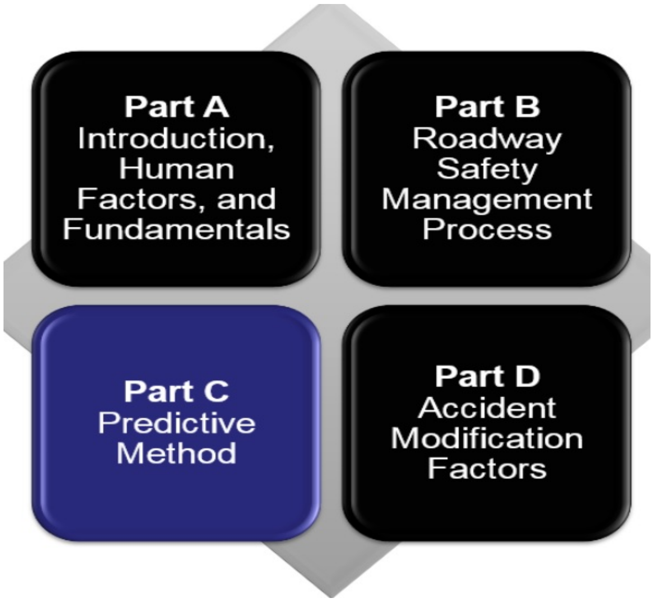

Figure 1.1 Parts of the 2010 Highway Safety Manual

1.2 HSM Predictive Methods

Crash prediction is important in planning, designing, and implementing countermeasures in order to enhance overall safety of the traffic system. Part C of the HSM deals with predictive methods (1). Predictive methods help to quantify the safety associated with roadway facilities such as roadway segments, intersections, and ramps. Expected number of crashes can be predicted using various different models. The expected number of predicted crashes depends on various factors such as traffic volume, weather conditions, and roadway geometry. Additionally, driver behavior highly influences crashes. However, such data are not always available. Naturalistic driving studies and driving simulator experiments may provide better insights on the impact of human factors on crashes. However, crash data can also provide very general information about driver behavior.

The predictive methods in Part C of the HSM are described in three chapters (Chapter 10–Chapter 12). Chapter 10 deals with rural two-lane two-way roads, Chapter 11 deals with rural multilane highways, and Chapter 12 deals with urban and suburban arterials. The primary focus of this study is the development of prediction models for rural two-way two-lane roadways.

2. METHODOLOGY

2.1 HSM SPFs

2.1.1 Safety Performance Functions

Safety performance functions are statistical models to estimate the expected number of crashes for a certain roadway facility. The SPFs are developed using data such as roadway geometry, AADT, or weather parameters. Each of those parameters has a different influence on crash occurrence and rate. Some of them heavily influence crash frequency and rates; however, some have only a minor influence. There are some factors that are a challenge to include in the safety predictive models. The changes in traffic laws, regulations, and policies can be difficult to account for. Other factors, such as economic status and fuel prices, may be represented by other confounding factors, such as traffic volume. For example, total fatalities in the United States were reduced significantly in 2008 because of an increase in fuel prices. The increase in fuel price decreases vehicle miles travelled (VMT), an exposure factor that can result in crash reduction. Factors like these cannot be explicitly incorporated in the predictive models estimation, although they seem to affect crash rates significantly. But the influence of these factors cannot be measured until they occur. Hence, only the measurable parameters are included in model development. The HSM states the only parameter that changes from each year is AADT, provided the geometry remains unchanged.

As mentioned earlier, the prediction models (SPFs) included in the Highway Safety Manual were developed using data from different states. However, for better prediction, those models need to be calibrated for specific jurisdiction for more accurate crash prediction. The HSM SPF predicts the expected number of crashes for base conditions; and the base conditions for rural two-lane two-way roadways are as follows:

Lane width: 12 ft.

Shoulder Width: 6 feet

Shoulder Type: Paved

Roadside Hazard Rating: 3

Driveway Density (DD): 5 driveways per mile

Horizontal Curvature: None

Vertical Curve: None

Centerline Rumble Strips: None

Passing Lanes: None

Two-way left turn lanes: None

Lighting: None

Automated Speed Enforcement: None

Grade Level: 0%

For grade level, 0% grade is not accepted by many of the states due to drainage criteria. The HSM SPF uses 0% as base condition, but it needs to be adjusted for the actual grade.

Safety performance function for total crashes for two-lane two-way (TLTW) roadways meeting base conditions is shown in Equation 1.

Nspf = AADT x L x 365 x 10-6 x e-0.312 **Eq. 1**

Where:

Nspf = Predicted number of total crashes for each year

AADT = Average annual daily traffic for that specific roadway/segment

L = Length of roadway segment in miles

The relationship shows that number of expected crashes increases linearly with AADT and length of the roadway segment. The constants 365 and 10^{-6} are used to convert AADT into million vehicles miles travelled (MVMT).

Figure 2.1 shows the graphical representation of crash rates for rural two-lane two-way roadways for base conditions. Independent variables are AADT and length of roadway segments, and the dependent variable is total crashes. In the graphical representation, AADT is plotted along the X-axis, and crash rate (crashes per mile) is plotted along the Y-axis. The AASHTO HSM 2010 assumed a linear relationship between crashes per mile and AADT and developed a function for rural two-lane two-way roadways. The relationship is given by Equation 1. However, studies suggest that the crash frequency follows a Poisson distribution. The negative binomial model is an alternative to the Poisson model, which has been used extensively to deal with the over-dispersion in crash data (2) (3) (4) (5). The negative binomial model is preferred to other models since it captures the variability in crash data (6).

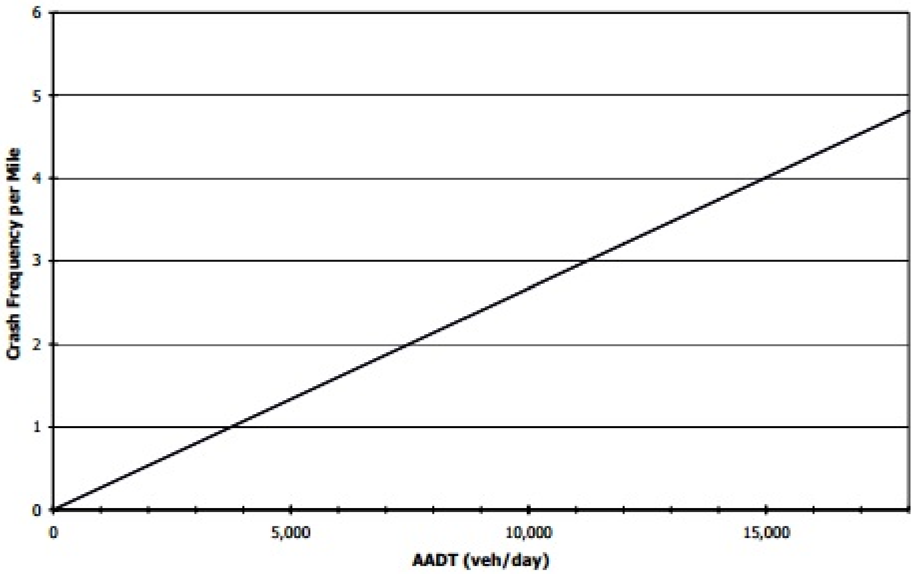

Figure 2.1 Crash Rate Based on AADT from 2010 HSM

2.1.2 Limitations of HSM SPFs

HSM SPFs were developed using the data collected from only six U.S. states. Figure 2.2 shows the states from which data were collected for the development of the SPFs provided in the HSM. In many ways, Wyoming potentially differs from these states (traffic composition, driver population, crash reporting threshold and forms, and weather conditions). Moreover, Wyoming faced an increase in energy-related activities over the last five years, which forms a shift in the traffic composition.

Figure 2.2 States from Which Data Were Collected for SPF Development in the 2010 HSM

According to the Wyoming Department of Transportation (WYDOT), in 2014 there were a total of 14,699 crashes, of which 6,215 crashes occurred in rural highways (of 131 fatal crashes, 104 occurred on rural highways). A better understanding of the factors that affect the number of crashes on rural roads will be an advantage for transportation officials to plan better to attain greater traffic safety in rural areas. A model developed using the crash data from rural highways will be able to predict the number of crashes for a specific segment of the roadway. It will also help to determine the hotspots, which are areas with high crash rates, and recommend suitable countermeasures to reduce the future number of crashes.

Traffic safety cannot be evaluated solely by observed crash frequencies and rates. Developed SPFs help to quantify the safety of the roadway system based on scientific approaches. They consider additional factors such as roadway geometry, weather conditions, and various additional factors that may affect the safety of the system.

While the effects of climate conditions may be accounted for by adjusting the SPFs using site-specific calibration factors, the effect of weather in the HSM first edition is not explicitly addressed. Through climate analysis, National Climatic Data Center (NCDC) scientists have identified nine climatically consistent regions within the contiguous United States, as shown in Figure 2.3. Although the development of the HSM SPFs depended on crash data collected from different states, these states are located in different climatic regions. Also, not all the climatic regions were considered when developing the HSM SPFs.

“Climate change will lead to fewer traffic accidents in West Midlands, UK,” stated a research study from the University of Gothenburg, Sweden (7). The results estimate climate change to decrease the number of days with temperature below zero degrees, which will also result in a reduction in the number of traffic crashes. U.S. Drought Monitor reported that 48 states were affected this year because of extreme heat this summer, and the odds of severe heat waves are increasing because of climate change according to various climate studies.

Ahmed et al. (2011) modeled crash occurrence in dry and snow seasons separately to account for the variability of climate conditions and found that crash risk could be significantly increased during snow seasons compared with dry seasons as a confounding effect between grades and pavement condition (8). Ahmed et al. (2012) concluded that crash frequency during snowy seasons could be approximately 82% higher than in dry seasons using crash data collected from the mountainous I-70 freeway in Colorado (9). Figure 2.3 from the NCDC show the different regions by precipitation and snowfall rates. Wyoming lies in the region where there is very low precipitation and high snowfall, which is different from the states from where data were collected. Hence, it is proposed that these regional differences should be considered in the process of developing SPFs.

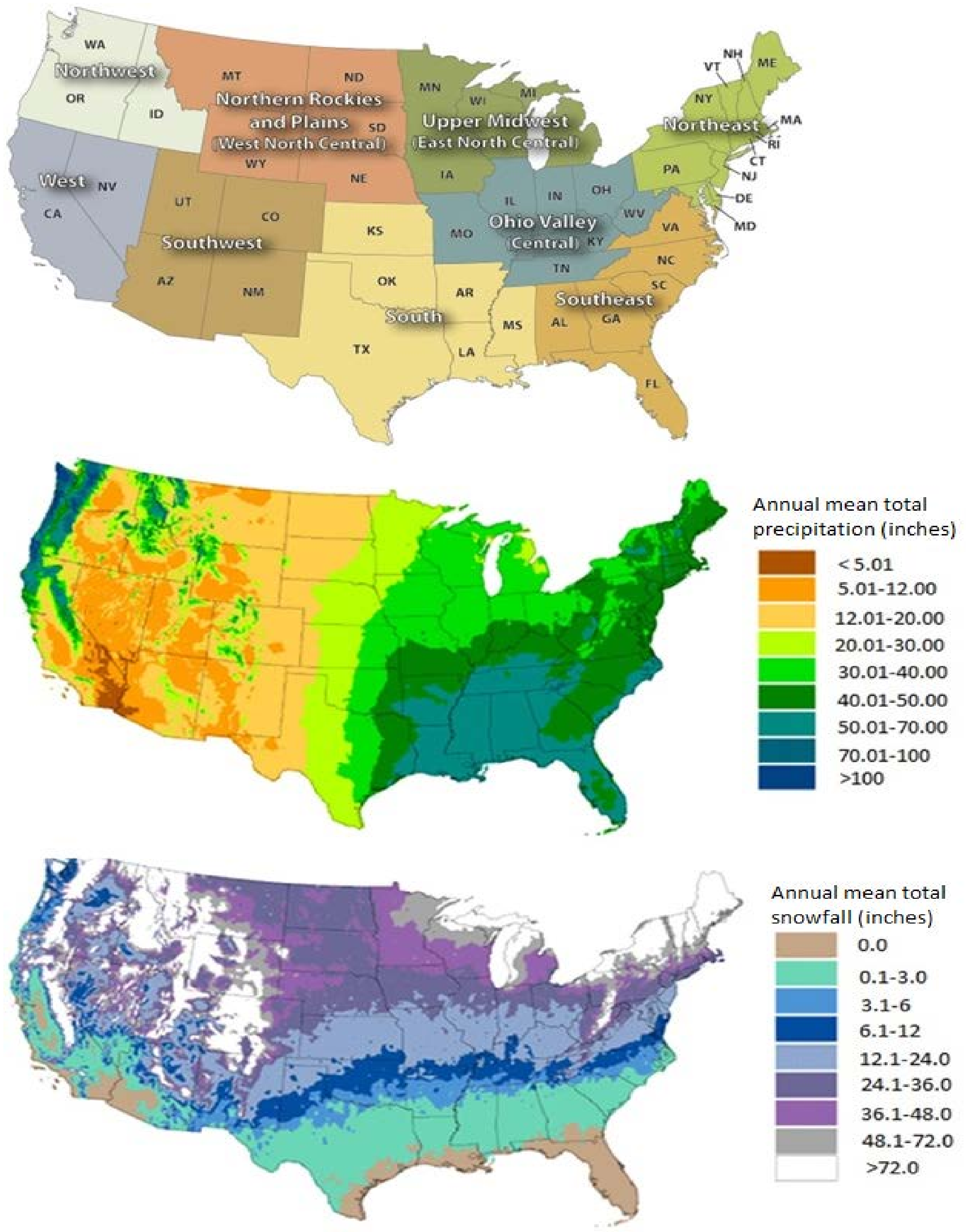

Figure 2.3 U.S. Climate Regions, Annual Mean Total Precipitation/Snow (NOAA)

The previously mentioned limitations may raise the need to have a specific SPF accounting for the extreme differences between Wyoming and the other states used to develop the HSM's SPFs.

The main goal of this research is to validate the adequacy and the transferability of the HSM SPFs to Wyoming. This would provide help in improvement of the implementation of the first edition of the AASHTO's Highway Safety Manual for the Rocky Mountains and Plain Regions. It would also be beneficial for better crash prediction capability to enhance the overall safety of the traffic system. The HSM safety management methodology includes economic evaluation of the expected crash outcomes of road improvement scenarios. There is a standard value for each different crash severity levels. If all of those factors are fully addressed, we can expect better crash predictions, which will provide further help for more accurate economic evaluations.

Different states have different crash cost estimates for different crash severity. A better evaluation of crash costs helps transportation officials invest in locations with higher crash frequency or severity where these investments will return the most of the capital in terms of crash reductions. The state of Wyoming has its own crash cost estimates based on crash severity levels, as shown in Table 2.1. The costs are based on base year 2012 and assume a 1.0107% cost escalation per year compounded from 2012 to 2014 (USDOT). The crash frequency data were provided by the Highway Safety Manual based on actual Wyoming counts.

Table 2.1 Crash Cost Estimates for Different Crash Severities

Crash Severity	**Societal Crash Cost Estimate**	**Frequency %**	**2012 Base Values from Highway Safety**	**HSM Crash Frequency for Comparison**
Fatal (K)	$9,299,000	0.80%	$9,103,485	1.3%
Disabling Injury (A)	$420,000	3.11%	$410,862	5.4%
Evident Injury (B)	$127,000	8.82%	$124,301	10.9%
Possible Injury (C)	$72,000	9.78%	$70,412	14.5%
Property Damage Only (PDO) (O)	$32,000	77.49%	$31,813	67.9%

2.2 Time Period for Analysis

A justifiable time period should be considered to get enough crash data. Crashes are known to be rare and random events. Hence, averaging the crash counts for several years helps in addressing the regression-to-the-mean phenomenon discussed in (10). In the case of rural highways with low traffic volume, generally a longer time period is needed because of the relatively small number of crashes over shorter periods of time. In order to get enough crash data to develop models, many studies have used multi-year crash data (Sacchi et al., 2012; Xie et al., 2011; Cafiso et al., 2010; Qin et al., 2005). For the development of jurisdiction-specific SPFs or for the calibration of the HSM SPFs, AASHTO 2010 recommends using a period that reflects the length of time for which the models will be used.

The likelihood of a crash occurrence is significantly affected by the short-term turbulence of traffic flow. Lee et al. (2003) claim that crash potential must be estimated on a real-time basis by monitoring the current traffic conditions (11). Wang et al. (2013) conducted a spatio-temporal analysis to explore the relationship between traffic congestion and crashes, and found that increased traffic congestion is associated with more injury crashes, which was just the opposite of their assumptions (12). Traditionally, aggregated variables have been used to include weather conditions in the model. However, aggregated variables might not be sufficient to capture the time-varying nature of driving environmental factors. This might result in significant loss of critical information on crash prediction. Chen et al. (2016) developed zero-inflated negative binomial models with site-specific random effects with unbalanced data panel data to analyze hourly crash frequency on highway segments, and found there is a unique significance of the real-time weather conditions, road surface conditions, and traffic data to predicted crash frequency (13).

2.3 Segmentation

Segmentation is an important process for developing SPFs. The roadways are divided into small segments and crashes are assigned to the corresponding segments. Development and application of SPFs depend upon the organization of data into distinct homogeneous entities, also known as segmentation, which depends on roadway geometry and traffic flow. However, Resende and Benekohal (1997) suggest that segmentation based on multiple variables may result in a lot of homogeneous segments, but with smaller lengths (14). The segment length affects the rate, and eventually the calibrated crash prediction models. If very short segments are formed, the lowered crash rates may result in inaccurate predictions. Some studies recommend using a segment length of 0.5 miles or longer. Since crashes are random and rare in nature, there might be a large number of segments with zero crashes (14). In order to avoid many segments with zero crashes, the segment length can be increased as for long as possible; however, unnecessary lengthening of segments might sacrifice homogeneity (15).

Various studies have already been done on segmentation. Miaou and Lum (1993) suggest that having very short segments can have a negative impact on estimation of coefficients for linear regression models, but not for Poisson regression models (2). They eliminated all roadway segments less than 0.05 miles in length. However, Ogle et al. (2011) indicate that segments with lengths less than 0.1 miles may cause uncertain results in crash analysis (16). If the length of homogeneous segments is less than 0.1 miles, they can either be removed (Miaou and Lum, 1993) or merged with adjacent segments with similar characteristics (Ahmed et al., 2011). Both of these approaches will eventually result in a fewer number of homogeneous roadway segments. Additionally, some studies suggest more localized studies that acknowledge the variability within a segment rather than averaging variables over different sub-segments to obtain a composite measure if the segments fail to be homogeneous with respect to some variables.

Berhanu (2004) defined roadway segments to be homogeneous on the basis of adjacent land use and cross-sectional characteristics. The variables for cross section were number of lanes, total road width, and median and shoulders widths (17). Cafiso (2011) used AADT as the only variable and segment length as the offset variable in order to identify the black spots on two-lane rural roads. It was also found that segment lengths should be related to AADT in order to obtain increased performance in identifying correct blackspots.

The HSM recommends developing the homogeneous segments based on AADT, number of lanes, curvature, presence of ramp at the interchange, lane width, outside and inside shoulder widths, median width, and clear zone width. The HSM recommends the length of a homogeneous segment to be at least 0.1 miles (1). If there are several segments with smaller segment lengths, different approaches can be taken for segmentation.

2.4 Crash Types and Severities

There are various factors that affect crash types and severity. Different weather conditions and roadway geometry, as well as driver behavior and performance, greatly define the type and severity of crashes. Crash injury severity is correlated with the type of the crash. For example, head-on collisions are more severe than that of rear-end collisions (18). Development of specific models based on crash types and severity helps planners and practitioners make effective decisions to increase the overall safety of the targeted roadway.

Roadway geometry greatly influences crash types. Khattak et al. (1998) also concluded that for adverse weather conditions, the severity of crashes increases for roadways with grades and curvatures (19). Persaud et al. (2004) concluded that rural highways generally do not feature physical barriers to separate opposing traffic flows, thereby increasing the chances of head-on and opposite sideswipe crashes. After installation of rumble strips, overall crashes and crash severity were found to decrease significantly. Head-on and opposite sideswipe crashes were found to be significantly reduced, and the number of crashes during nighttime was reduced to daytime levels (3). Many studies, such as Ahmed et al., (2011) have shown that the number and severity of crashes increase on downgrades and decrease on upgrades, especially on snowy, slushy, or wet roadway surfaces.

However, development of models for specific types and severities can be difficult because of the low number of crashes for a specific type and severity. The HSM suggests that researchers and practitioners multiply estimated total crashes by a percentage of jurisdiction-wide crashes of that type in order to predict the crashes by type. However, this approach might not be accurate since it assumes that the crash distribution by type is constant throughout the jurisdiction, which may not be valid. It is suggested to develop the model separately for different crash types rather than using the proportion approach because of the nonlinear relationships between crashes of different types and traffic flow.

2.5 Model Development

Many studies in the literature suggest developing jurisdiction-specific SPFs if enough data are available. Jurisdiction-specific SPFs are believed to predict the crashes more effectively compared with utilizing the HSM SPFs with adjustment using CMFs and calibration factors. Poisson and Negative Binomial (NB) models are widely used to calibrate the SPFs because the number of crashes on the sites is assumed to have a Poisson distribution, and crash variations across different sites are assumed to have a gamma distribution. The negative binomial distribution is also called as Poisson-gamma distribution since it encompasses characteristics of both the Poisson distribution (for crash frequency) and the gamma distribution (for variation of crashes across different sites). The negative binomial model is an alternative to the Poisson model. The values of the mean and variance coincide in the Poisson model. However, the NB model does not require the mean to be equal to the variance. The negative binomial model is preferred to other models since it allows for extra variability in the crash data. While developing jurisdiction-specific SPFs, use of the NB model is suggested in order to capture the dispersion of crash data. Dispersion of crash data is estimated by the over-dispersion parameter, which is a component of the NB model.

After segmentation, SPFs for different terrain and severity levels can be developed. Typically, a negative binomial model is used to develop SPFs. However, if the dataset consists of several zeros, a zero-inflated Poisson model can be used. Model diagnostics can also be performed to determine the goodness-of-fit of each model.

The negative binomial or Poisson-gamma regression model is the most widely used model for crash prediction. Crashes are random in nature, and the frequency of crashes often follows a gamma distribution where the variance of the crash counts exceeds the mean. The negative binomial model captures the randomness in crash frequencies and is expected to predict the expected number of crashes more efficiently. The expected number of crashes can be obtained from the equations below (2).

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}) \qquad \textbf{Eq. 2}$$

where:

λ is the expected number of crashes

β_0 is the intercept

x_{ji} is predictor variable *j* for observation *i*

β_j is population regression coefficient for predictor variable *j*

$$\text{Var} = \lambda_i + \alpha \lambda_i^2 \qquad \textbf{Eq. 3}$$

Where:

α is the overdispersion parameter

The Poisson regression model is a special case of NB model when α approaches to zero. The probability function of the NB distribution is given by the following as

$$P(Y = y) = \frac{\Gamma(y+1/\alpha)}{\Gamma(y+1)\Gamma(\frac{1}{\alpha})}\left[\frac{(\alpha\lambda)^y}{(1+\alpha\lambda)^{y+\frac{1}{\alpha}}}\right] \qquad \textbf{Eq. 4}$$

Where:

$\Gamma(x)$ = gamma function

y = number of crashes

α= overdispersion parameter

λ= mean of y

The probability function of the zero-inflated distribution is given by the following equation as

$$P(Y = y) = \begin{cases} w + (1-w)e^{-\lambda} & for\ y < 0 \\ \frac{(1-w)\lambda^y e^{-\lambda}}{y!} & for\ y > 0 \end{cases} \qquad \textbf{Eq. 5}$$

Simple SPFs include AADT as the only explanatory variable and the number of crashes as the response variable. The SPF functional form for a roadway segment is given by the following equation as

$N_{predicted} = e^{a} \times AADT^{b}$ **Eq. 6**

In order to find out the value of a and b, the equation can be rewritten as

$N_{predicted} = \exp(a + b \times \ln(AADT))$ **Eq. 7**

Where:

a and b are regression coefficients

$N_{predicted}$ is predicted number of crashes per mile per year

AADT is average annual daily traffic (vehicles per day)

Extra variation is accounted for by an overdispersion parameter in the negative binomial model. It indicates the statistical reliability of the SPF. The closer the value to zero, the more reliable is the SPF.

Default HSM SPFs can be multiplied by a calibration factor to obtain calibrated SPF, and the calibration factor can be obtained as follows:

$\text{Calibration factor (C)} = \frac{\sum_{All\ Sites} Observed\ Crashes}{\sum_{All\ Sites} Predicted\ Crashes}$ **Eq. 8**

The equation shows that the calibration factor can be obtained from the ratio of observed crashes to predicted crashes. However, the calibration factor is not required if jurisdiction-specific SPFs are calibrated.

3. DATA COLLECTION

The main objective of the data collection process is to select the roadway segments for calibration of Wyoming-specific SPFs. For calibration of simple SPFs, all the segments were selected for both rural two-lane two-way roadways. However, for calibration of full SPF, segments were selected in a random manner in order to, it is hoped, represent all parts of the state. Also, the segments were selected in a way to cover as many roadway characteristics and regions in Wyoming as possible. The HSM states that segments do not need to conform to base conditions for calibration and development of jurisdiction-specific SPFs (1).

The following sections explain data needs, data collection procedure for crash data, traffic data, roadway data, and weather data, and implications for data collection along with preparation of the dataset for analysis.

3.1 Data Needs

3.1.1 Crash Data

Crash frequency is the dependent variable in SPFs. The quality and detail of crash data are crucial for the development of accurate SPFs. Different severity levels and crash types can be used for modeling in order to obtain models for each crash type and severity level. However, if specific target crashes or severity levels are rare, combining multiple severity level techniques could be utilized. SPFs, if developed for different severity levels and crash types, can be helpful for applying effective countermeasures. The KABCO scale is commonly used to categorize different crash severity levels. KABCO defines crash severity levels in five categories.

i. K: Fatal injury
ii. A: Incapacitating injury
iii. B: Non-incapacitating injury
iv. C: Possible injury
v. O: No injury (property damage only)

For this study, models were developed for total crashes, property damage only (PDO), and fatal and injury (F+I) crashes, according to the HSM guidelines.

3.1.2 Facility Data

Characteristics and features of study sites are called facility data. The characteristics of sites are defined by roadway geometry and traffic volume. Roadway geometry typically includes length of segments, number of lanes, lane width, shoulder width, horizontal curvature, vertical grade, presence of rumble strips, traffic control devices, etc. Annual average daily traffic (AADT) and vehicle miles travelled (VMT) are the parameters commonly utilized to define traffic flow.

Weather conditions also can be treated as facility data since they define the characteristics of the site. Weather parameters are defined in the following subsection.

3.1.3 Weather Data

Weather conditions may influence crash frequency and crash rate. Weather parameters, if included in the model, can be helpful for more accurate crash prediction. Snowy and rainy conditions are the commonly used adverse conditions in modeling. However, visibility, fog, and other forms of precipitation also can be included in the model. In this study, only the snowy and rainy conditions are considered (9) (20) (5). The relationship between weather conditions and roadway geometry is important in order to suggest appropriate countermeasures to reduce the effects of adverse weather conditions.

3.2 Data Acquisition

For simple SPFs, all segments conforming to the jurisdiction were considered. AADT and crashes were incorporated with the corresponding segments accordingly. Segmentation was based on AADT for the development of simple SPFs.

For the development of full SPFs, the study segments were selected in a random way in order to represent all parts of the state. The number of crashes and traffic volumes were not considered to ensure all roadways of different ranges of AADT and crash history are covered in the data. Initially, the study segments were selected based on roadway geometry and roadway characteristics. Later, additional variables (AADT and weather parameters) were added to the data. WYDOT collects AADT for interstate highways, state highways, and local federally sponsored roadways, which can be obtained from the WYDOT database.

3.3 Data Preparation

Homogeneous segmentation approach was used for each corridor. Routes were selected from different parts of the state in order to capture the variation in land use, terrain, types of vehicles, and weather parameters within the state.

First, segmentation was done based on horizontal alignment and vertical grades. The degree of the curve was used to represent the horizontal alignment of segments. Vertical grades were represented as categorical variables. The vertical grade was divided into four categories, which are shown as follows:

i. Mild upgrade (0% to 3%)
ii. Steep upgrade (>3%)
iii. Mild downgrade ((-3) % to 0%)
iv. Steep downgrade (< (-3) %)

Using categorical variables for grades instead of continuous variables decreases the number of segments and will provide longer segments. This will help to reduce the number of segments with zero crashes. Also, it helps to eliminate the segments with a length less than 0.1 mile. In this study, the minimum length of segments of 0.1 mile was considered. Previous studies suggest that segments less than 0.1 mile in length should be removed or merged to an adjacent segment that has similar characteristics.

After the segmentation based on horizontal alignment and vertical grades was completed, other parameters were added for each segment. These parameters are AADT, the presence of rumble strips, the presence of passing lanes, shoulder width, average number of rainy days each year, and the average number of snowy days each year. Except for the presence of passing lanes and rumble strips, all variables are continuous. The presence of rumble strips and passing lanes were categorized as binary variables.

For the development of the HSM SPF, only the non-intersection crashes were taken into consideration. The crashes related to intersections and interchanges were removed from the dataset.

3.4 Data Utilized

Crash and traffic data from 2003 to 2013 were used to develop Wyoming-specific SPFs. Simple SPFs were developed using all the rural roadway segments. However, for calibration of full SPFs, segments were chosen randomly. Table 3.1 and Table 3.2 provide the summary of the data used for calibration of Wyoming-specific simple SPFs and Wyoming-specific full SPFs, respectively.

Table 3.1 Summary of Rural Roadway Segments Utilized in Simple SPFs

Category	**Terrain**	**Total Length of Segments (Miles)**	**Crash Data (2003-2013)**	
			Total Crashes	**Fatal and Injury Crashes**
Interstate Freeways	**Flat and Rolling**	1527.98	21447	5963
	Mountainous	102.44	2428	608
Two-Lane Two-Way	**Flat and Rolling**	4421.55	12371	4103
	Mountainous	513.02	2250	808

Table 3.2 Summary of Rural Roadway Segments Utilized in Full SPFs

Category	**Terrain**	**Total Length of Segments (Miles)**	**Number of Segments**	**Crash Data (2003-2013)**	
				Total Crashes	**Fatal and Injury Crashes**
Interstate Freeways	**Flat and Rolling**	355.82	40	5439	1425
	Mountainous	98.45	14	2391	596
Two-Lane Two-Way	**Flat and Rolling**	538.81	60	1556	489
	Mountainous	380.70	38	1435	471

Figure 3.1 provides the graphical representation of segment lengths and their frequencies used in development of full SPFs.

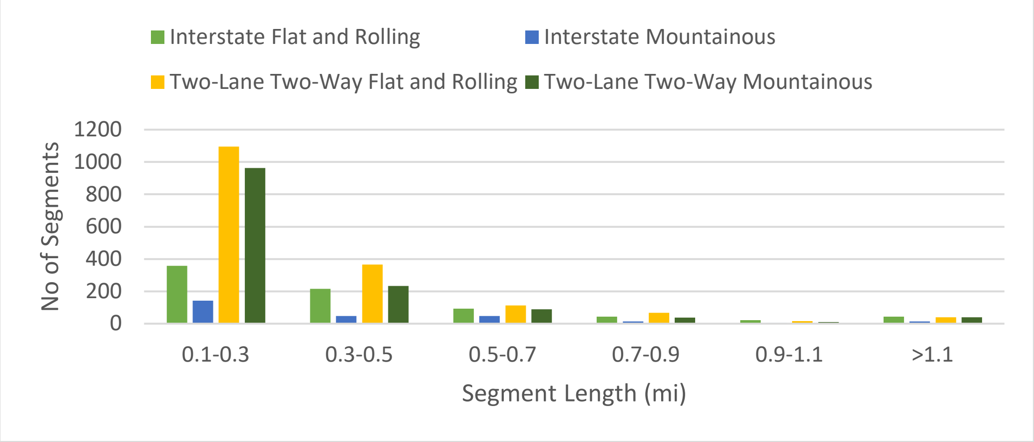

Figure 3.1 Segment Length Distribution

Table 3.3 displays and describes all of the variables considered in full SPF development.

Table 3.3 Description of Available Variables for Full SPF Development

Variables	Description	Interstate Flat and Rolling			
		Minimum	**Maximum**	**Average**	**Std. Deviation**
Response Variables					
Total Crashes	frequency of all crashes per segment	0	112	7	10.33
Fatal and Injury Crashes	frequency of F+I crashes per segment	0	27	2	2.77
Exposure Variables					
AADT	Average Annual Daily Traffic (veh/day)	1200	12171	4416	2160.58
VMT	Daily vehicle miles travelled	49.1	35106.38	2036.19	2898.27
Segment Length	length of road segment (mi)	0.1	5.85	0.46	0.55
Risk Factors					
Degree of Curve	Degree of horizontal curve	0	3.32	0.36	0.57
Grade	longitudinal grade, four categories: 0%-3%=1, >3%=2, 0%-(-3)%=3, <(-3)%=4	-5	5.49	-0.35	1.78
Inside Shoulder Width	Inside shoulder width (ft)	2	4	3.96	0.27
Outside Shoulder Width	Outside shoulder width (ft)	4	14	9.78	1.05
Median Width	Median width (ft)	32	148	97.51	36.16
No of Rainy Days	Average no of rainy days per year	44	112	80	23.17
No of Snowy Days	Average no of snowy days per year	11	58	33	13.91

3.5 Potential Data Issues

3.5.1 Traffic Data

Collecting an exact AADT for all roadway classification is not feasible since traffic data in Wyoming are collected for state and interstate highways only. Automatic traffic recorders are installed at various locations to record the traffic counts. Later, correction factors can be applied based on the number of access roads, residential areas, and other factors that cause divergence to estimate AADT at the desired segment of the highway. AADT might be missing and sometimes AADT for the desired location cannot be obtained. These implications may result in further elimination of roadways segments with no traffic data. One should choose highway segments only after the complete availability of all required data for the analysis.

3.5.2 Weather Data

Collecting precise weather data has always been an issue not only in traffic safety but also in other research works. It is even more difficult to collect weather data with the very sparse Road Weather Information System (RWIS) and weather stations at airports. Even though weather data are collected from weather stations, it is not always a true representation of the desired location since the weather stations can be far from the actual location. However, those data can be used assuming a minimum variance between real weather conditions and the weather data collected from the weather stations.

3.5.3 Crash Reporting

Errors may occur with crash reporting, the data entry process, or any other process related to storage of crash data. Sometimes police officers fail to fill out a certain section of the crash report, and sometimes those filled spaces may be omitted for various reasons. These inconsistencies in crashes cannot be avoided completely. A complete investigation of each crash during the study period is also not possible. Some of these mistakes are unavoidable; related personnel can only hope that they are not large enough to significantly affect the model results. Moreover, these types of missing or inaccurate data are random by nature and hence would not affect the accuracy of the calibrated models.

4. DATA ANALYSIS

4.1 Descriptive Analysis

Data were collected for rural two-lane two-way roadways and interstate freeways from 2003 to 2013. In this time, 21,423 crashes occurred on rural two-lane two-way roadways in Wyoming. The million vehicle miles travelled (MVMT) during the study duration for two-lane two-way roadways is approximately 131,778 MVMT, giving a crash rate of 0.16625 crashes per MVMT for Wyoming’s rural two-lane two-way roadways.

Table 4.1 displays the crash frequencies by weather condition, according to the crash report.

Table 4.1 Crash Distribution According to Weather Condition

Weather Conditions	**Two-lane two-way**	
	Crash Frequency	**Percentage (%)**
Blizzard	155	0.72
Blowing Dust or Sand or Dirt	20	0.09
Blowing Snow	186	0.87
Clear	17355	81.01
Cloudy or Overcast	494	2.31
Fog	165	0.77
Other	14	0.07
Raining	513	2.39
Severe Wind Only	335	1.56
Sleet or Hail or Freezing Rain	97	0.45
Smoke	6	0.03
Snowing	1897	8.85
Unknown	186	0.87

Table 4.2 displays the crash distribution throughout the state, according to the lighting conditions.

Table 4.2 Crash Distribution According to Lighting Condition

Lighting Conditions	**Two-lane two-way**	
	Crash Frequency	**Percentage (%)**
Darkness Lighted	377	1.76
Darkness Unlighted	8030	37.48
Dawn	1094	5.11
Daylight	10745	50.16
Dusk	1055	4.92
Other	2	0.01
Unknown	120	0.56

Table 4.3 displays the statewide crash frequencies for rural TWTL highways, separated by severity level.

Table 4.3 Crash Distribution According to Crash Severity

Crash Severity	**Two-lane two-way**	
	Crash Frequency	**Percentage (%)**
Fatal Injury	404	1.89
Incapacitating Injury	1410	6.58
No Injury	15680	73.19
Non- incapacitating Injury	2309	10.78
Possible Injury	1460	6.82
Unknown	160	0.75

4.2 Statistical Analysis

The datasets were prepared for rural TWTL highways in order to develop Wyoming-specific simple and full safety performance functions. The datasets were analyzed using SAS (statistical analysis software) version 9.4. For simple SPFs, all the roadway segments were considered and the AADT was the only explanatory variable. For full SPFs, all the variables mentioned in Table 3.3 were initially included to develop the negative binomial model. The variables with insignificant results were discarded and the model was rerun. For rural TWTL roadways, many segments had zero crashes. Hence, the zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) models were also developed for those roadways. A comparison of the negative binomial model and the ZIP model will be provided. Models were developed for each classification and terrain for summer and winter seasons and compared to discover the effects of different seasons in crashes.

4.2.1 Simple SPFs

4.2.1.1 General Simple SPFs

The following SPFs were developed for all crashes that occur on rural TWTL highways regardless of season or weather condition. For each developed SPF, the model coefficients and significance, as well as the dispersion, will be provided. Following this, the actual SPFs, separated into those for total crashes and those for F+I crashes, will be given. Note that cpm will designate the average crashes per mile. Table 4.4 and Table 4.5 detail the general simple SPFs.

Table 4.4 Coefficients for Simple SPFs

	Flat and Rolling				**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	0.0776	0.0505	-1.3	<.0001	0.3678	<.0001	-0.1318	0.2506
β	0.0003	<.0001	0.0004	<.0001	0.0001	0.038	0	0.6879
Dispersion	1.6274		0.392		1.0113		0.8652	

Table 4.5 Wyoming-Specific Simple SPFs

	Two-Lane Two-Way	
	Flat and Rolling	Mountainous
Total	cpm=exp(0.0776+0.0003AADT)	cpm=exp(0.3678+0.0001AADT)
F+I	cpm=exp(-1.3+0.0004AADT)	cpm=exp(-0.1318+0AADT)

Table 4.6 and Table 4.7 also detail the general simple SPFs, however, with a log transformation of traffic volumes (AADT).

Table 4.6 Coefficients for Simple SPFs (Log Transformation)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-4.4945	<.0001	-3.8935	<.0001	-1.3909	0.0141	-0.3549	0.624
β	1.6507	<.0001	1.2697	<.0001	0.6432	0.0008	0.0639	0.7928
Dispersion	1.4666		1.354		0.9947		0.8658	

Table 4.1 Wyoming-Specific Simple SPFs (Log Transformation)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-4.4945+1.6507AADT)	cpm=exp(-1.3909+0.6432AADT)
F+I	cpm=exp(-3.8935+1.2697AADT)	cpm=exp(-0.3549+0.0639AADT)

4.2.1.2 Seasonal Simple SPFs

Following the development of general simple SPFs, the distinctions made on a seasonal basis (summer and winter) are shown below. Table 4.8 and Table 4.9 detail the simple SPFs developed for summer conditions.

Table 4.8 Coefficients for Simple SPFs (Summer Season)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-3.9356	<.0001	-3.4073	<.0001	-0.0806	0.8991	0.4558	0.58
β	1.3395	<.0001	1.0413	<.0001	0.0745	0.7297	-0.2726	0.328
Dispersion	1.3769		1.1964		0.9585		0.8702	

Table 4.9 Wyoming-Specific Simple SPFs (Summer Season)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-3.9356+1.3395AADT)	cpm=exp(-0.0806+0.0745AADT)
F+I	cpm=exp(-3.4073+1.0413AADT)	cpm=exp(0.4558-0.2726AADT)

Table 4.10 and Table 4.11 detail the SPFs developed for the winter weather season.

Table 4.10 Coefficients for Simple SPFs (Winter Season)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-4.4074	<.0001	-3.9199	<.0001	-1.5055	0.0148	0.5019	0.6496
B	1.5155	<.0001	1.2161	<.0001	0.4896	0.016	-0.3738	0.2987
Dispersion	1.4016		1.2198		0.3809		0.3478	

Table 4.11 Wyoming-Specific Simple SPFs (Winter Season)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-4.4074+1.5155AADT)	cpm=exp(-1.5055+0.4896AADT)
F+I	cpm=exp(-3.9199+1.2161AADT)	cpm=exp(0.5019-0.3738AADT)

4.2.1.3 Simple SPFs by Weather Condition

The following SPFs were developed for the three primary weather conditions faced in Wyoming: clear, rainy, and snowy. Table 4.12 and Table 4.13 detail the developed SPFs for clear weather conditions.

Table 4.12 Coefficients for Simple SPFs (Clear Weather)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-4.2764	<.0001	-3.8453	<.0001	-0.5762	0.3542	0.5408	0.5048
B	1.5445	<.0001	1.2371	<.0001	0.2929	0.1649	-0.2557	0.3504
Dispersion	1.4791		1.3266		1.0253		0.9807	

Table 4.13 Wyoming-Specific Simple SPFs (Clear Weather)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-4.2764+1.5445AADT)	cpm=exp(-0.5762+0.2929AADT)
F+I	cpm=exp(-3.8453+1.2371AADT)	cpm=exp(0.5408-0.2557AADT)

Table 4.14 and Table 4.15 describe the developed SPFs for snowy weather conditions.

Table 4.14 Coefficients for Simple SPFs (Snowy Weather)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-3.8264	<.0001	-2.5394	0.0038	0.5707	0.645	0.4607	0.7966
B	1.1513	<.0001	0.6555	0.0159	-0.4009	0.3066	-0.47	0.4054
Dispersion	1.099		0.9133		0.2496		0	

Table 4.15 Wyoming-Specific Simple SPFs (Snowy Weather)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-3.8264+1.1513AADT)	cpm=exp(0.5707-0.4009AADT)
F+I	cpm=exp(-2.5394+0.6555AADT)	cpm=exp(0.4607-0.47AADT)

Table 4.16 and Table 4.17 detail the developed SPFs for rainy weather conditions.

Table 4.16 Coefficients for Simple SPFs (Rainy Weather)

Flat and Rolling					**Mountainous**			
	Total		Fatal+Injury		Total		Fatal+Injury	
	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq	Estimate	Pr>ChiSq
Intercept	-4.183	0.0179	-3.9675	0.0238	1.348	0.4378	0.9416	0.8044
β	1.1283	0.0382	1.1531	0.0383	-0.7779	0.1774	-0.792	0.5194
Dispersion	1.2401		1.7291		0		0	

Table 4.17 Wyoming Specific Simple SPFs (Rainy Weather)

	Two-Lane Two-Way Highways	
	Flat and Rolling	Mountainous
Total	cpm=exp(-4.183+1.1283AADT)	cpm=exp(1.348-0.7779AADT)
F+I	cpm=exp(-3.9675+1.1531AADT)	cpm=exp(0.9416-0.792AADT)

4.2.2 Analysis of Simple SPFs

Various Wyoming-specific simple SPFs were developed for rural two-lane two-way roadways. SPFs were developed for two different terrain types (flat and rolling and mountainous) as well as two severity levels (total crashes and fatal+injury crashes).

Wyoming-specific simple SPFs developed for rural two-lane two-way crashes indicate that there is an increase in total crashes with the increasing AADT. For both seasons (summer and winter) and weather conditions (clear, rainy, and snowy), there is a decrease in the number of total crashes with an increase in AADT. The models developed for rural mountainous two-lane two-way roadways are not significant at a 0.05 significance level. For mountainous two-lane two-way roadways, the segments have little variation in crash frequencies.

4.2.3 Full SPFs

In order to discover the impact of roadway geometry and weather conditions on crashes, full SPFs were developed. The roadway geometry variables included in full SPFs are lane width, shoulder width, horizontal curves, vertical grades, the manner of passing, the presence of passing lanes, and the presence of rumble strips.

The major weather-related variables are precipitation, snowfall, visibility, fog, and wind speed. Visibility is not included in the model since it is difficult to include visibility indices on an aggregate level. Hence, the only weather-related variables included were precipitation and snowfall. These variables were considered on an aggregate level as average number of rainy and snowy days per segment per year. Full SPFs for different terrain and roadways were developed, and coefficients are tabulated in Table 4.18 and Table 4.19.

Table 4.18 Wyoming-Specific Full SPFs for Flat and Rolling TWTL Roadways

Parameter	**Total Crashes**				**Fatal and Injury (F+I) Crashes**			
	Estimate	**Wald 95% Confidence Limits**		**Pr>ChiSq**	**Estimate**	**Wald 95% Confidence Limits**		**Pr>ChiSq**
Intercept	-4.4469	-4.9124	-3.9814	<.0001	-5.3682	-6.0994	-4.6371	<.0001
logVMT	1.9721	1.789	2.1553	<.0001	2.035	1.7571	2.3129	<.0001
Grade2	-0.1229	-0.3603	0.1144	0.3101	0.0529	-0.3153	0.4211	0.7783
Grade3	0.06	-0.0836	0.2035	0.4128	0.1792	-0.0423	0.4007	0.1129
Grade4	0.1285	-0.098	0.355	0.266	0.0482	-0.3267	0.4232	0.801
Degree of Curve	0.0241	-0.0045	0.0527	0.0992	0.0357	-0.0064	0.0778	0.0965
Shoulder Width	-0.098	-0.1307	-0.0653	<.0001	-0.1123	-0.1615	-0.0631	<.0001
Rumble Strips	-0.0484	-0.2236	0.1267	0.5879	-0.167	-0.4414	0.1074	0.2329
Passing Lane	-0.2436	-0.7124	0.2252	0.3085	-0.1458	-0.8481	0.5565	0.6841
No of Rainy Days	0.0032	-0.0007	0.0071	0.1084	-0.0052	-0.0122	0.0017	0.1404
No of Snowy Days	0.001	-0.0053	0.0074	0.7496	0.0111	0.0004	0.0219	0.0422
Dispersion	0.4954	0.3856	0.6366		0.5236	0.3013	0.9099	

Table 4.19 Wyoming-Specific Full SPFs for Mountainous TWTL Roadways

Parameter	Total Crashes				Fatal and Injury (F+I) Crashes			
	Estimate	Wald 95% Confidence Limits		Pr>ChiSq	Estimate	Wald 95% Confidence Limits		Pr>ChiSq
Intercept	-4.3099	-4.8622	-3.7577	<.0001	-5.8456	-6.7118	-4.9793	<.0001
logVMT	1.9004	1.7033	2.0975	<.0001	1.8223	1.5364	2.1082	<.0001
Grade2	-0.1256	-0.3866	0.1354	0.3457	-0.1304	-0.5083	0.2474	0.4987
Grade3	0.1267	-0.045	0.2985	0.148	0.1469	-0.0946	0.3884	0.2331
Grade4	0.1821	-0.0634	0.4276	0.146	0.1729	-0.1756	0.5215	0.3308
Degree of Curve	0.0522	0.0224	0.0819	0.0006	0.0675	0.0269	0.1082	0.0011
Shoulder Width	-0.0092	-0.0447	0.0264	0.6141	0.0215	-0.0286	0.0716	0.3999
Rumble Strips	-0.1009	-0.3847	0.183	0.4862	0.0754	-0.3095	0.4602	0.7011
Passing Lane	-1.5715	-3.1797	0.0368	0.0555	-21.8188	-63259.7	63216.09	0.9995
No of Rainy Days	-0.0055	-0.01	-0.0011	0.0155	0.0006	-0.006	0.0072	0.8599
No of Snowy Days	0.0115	0.0054	0.0175	0.0002	0.0078	-0.0008	0.0163	0.0763
Dispersion	0.6823	0.5441	0.8556		0.6248	0.3942	0.9903	

Wyoming-specific full SPFs were developed for different classifications and terrain. Negative binomial models were fit for each case using SAS 9.4 software. The variables included in the model were logarithmic transformation of vehicle miles travelled (VMT), grade, degree of curvature, shoulder width (two-lane two-way roadways only), inside and outside shoulder width and median width (for interstate freeways only), rumble strips, passing lanes/climbing lanes, number of rainy days, and number of snowy days. VMT was used in the model instead of AADT in order to represent the real exposure of traffic and the length of the roadway segment.

Some of the variables in the models were statistically insignificant at a 0.05 significance level. Hence, the models were rerun removing one insignificant variable at a time, and only statistically significant variables are included in the model.

The following tables will detail the final results (only significant variables) for the full SPFs.

Table 4.20 NB Model for Flat and Rolling TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-4.4121	0.2175	-4.8384	-3.9857	411.36	<.0001
logVMT	1	1.9614	0.0895	1.7860	2.1368	480.37	<.0001
Degree of Curve	1	0.0241	0.0146	-0.0045	0.0527	2.72	0.0989
Shoulder Width	1	-0.1015	0.0141	-0.1291	-0.0738	51.80	<.0001
No of Rainy Days	1	0.0037	0.0011	0.0017	0.0058	12.71	0.0004
Dispersion	1	0.4985	0.0636	0.3882	0.6402		

Table 4.21 NB Model for Mountainous TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-4.3194	0.2815	-4.8711	-3.7677	235.45	<.0001
logVMT		1	1.8760	0.0950	1.6898	2.0621	390.17	<.0001
Grade	2	1	-0.1179	0.1329	-0.3784	0.1425	0.79	0.3748
Grade	3	1	0.1307	0.0875	-0.0407	0.3021	2.23	0.1351
Grade	4	1	0.1856	0.1251	-0.0596	0.4308	2.20	0.1379
Degree of Curve		1	0.0509	0.0151	0.0212	0.0805	11.31	0.0008
Passing Lane		1	-1.5647	0.8213	-3.1745	0.0450	3.63	0.0568
No of Rainy Days		1	-0.0055	0.0023	-0.0100	-0.0010	5.76	0.0164
No of Snowy Days		1	0.0116	0.0031	0.0055	0.0176	14.12	0.0002
Dispersion		1	0.6824	0.0789	0.5441	0.8559		

Table 4.22 NB Model for Summer Flat and Rolling TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-4.8732	0.2687	-5.3999	-4.3465	328.84	<.0001
logVMT		1	1.8391	0.1123	1.6191	2.0591	268.33	<.0001
Grade	2	1	0.0208	0.1583	-0.2894	0.3310	0.02	0.8954
Grade	3	1	0.1926	0.0958	0.0049	0.3803	4.04	0.0443
Grade	4	1	0.1317	0.1563	-0.1745	0.4380	0.71	0.3992
Shoulder Width		1	-0.0962	0.0179	-0.1313	-0.0612	28.96	<.0001
No of Snowy Days		1	0.0068	0.0021	0.0026	0.0110	10.15	0.0014
Dispersion		1	0.5332	0.1054	0.3619	0.7856		

Table 4.23 NB Model for Winter Flat and Rolling TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-4.9952	0.2898	-5.5632	-4.4272	297.11	<.0001
logVMT		1	1.9800	0.1151	1.7544	2.2056	295.85	<.0001
Grade	2	1	-0.2337	0.1593	-0.5460	0.0785	2.15	0.1423
Grade	3	1	-0.0643	0.0933	-0.2472	0.1187	0.47	0.4911
Grade	4	1	0.1645	0.1444	-0.1185	0.4475	1.30	0.2545
Degree of Curve		1	0.0385	0.0176	0.0040	0.0731	4.77	0.0289
Shoulder Width		1	-0.0984	0.0182	-0.1341	-0.0628	29.24	<.0001
No of Rainy Days		1	0.0027	0.0013	0.0001	0.0053	4.08	0.0433
Dispersion		1	0.6114	0.1096	0.4303	0.8687		

Table 4.24 NB Model for Summer Mountainous TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.7946	0.2907	-5.3643	-4.2249	272.09	<.0001
logVMT	1	1.7009	0.1127	1.4801	1.9218	227.92	<.0001
Degree of Curve	1	0.0361	0.0198	-0.0027	0.0748	3.33	0.0681
No of Snowy Days	1	0.0043	0.0019	0.0006	0.0081	5.06	0.0244
Dispersion	1	0.6847	0.1247	0.4792	0.9784		

Table 4.25 NB Model for Winter Mountainous TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-5.3827	0.3824	-6.1322	-4.6332	198.12	<.0001
logVMT		1	2.0668	0.1277	1.8166	2.3171	262.06	<.0001
Grade	2	1	-0.1099	0.1770	-0.4568	0.2370	0.39	0.5346
Grade	3	1	0.2909	0.1130	0.0694	0.5123	6.63	0.0100
Grade	4	1	0.0715	0.1691	-0.2600	0.4029	0.18	0.6726
Degree of Curve		1	0.0665	0.0198	0.0277	0.1052	11.27	0.0008
No of Rainy Days		1	-0.0090	0.0034	-0.0156	-0.0024	7.07	0.0078
No of Snowy Days		1	0.0167	0.0044	0.0080	0.0253	14.25	0.0002
Dispersion		1	0.9314	0.1338	0.7029	1.2342		

Table 4.26 ZIP Model for Flat and Rolling TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-4.0513	0.1913	-4.4262	-3.6765	448.61	<.0001
logVMT	1	1.8751	0.0703	1.7374	2.0128	712.32	<.0001
Degree of Curve	1	0.0313	0.0135	0.0048	0.0577	5.36	0.0207
Shoulder Width	1	-0.0851	0.0114	-0.1073	-0.0628	56.10	<.0001
No of Rainy Days	1	0.0029	0.0009	0.0012	0.0046	11.36	0.0007
Scale	0	1.0000	0.0000	1.0000	1.0000		

Table 4.27 ZIP Model for Mountainous TWTL Roadways (Total Crashes)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-3.8146	0.2374	-4.2799	-3.3494	258.24	<.0001
logVMT		1	1.7300	0.0723	1.5882	1.8717	572.04	<.0001
Grade	2	1	0.0020	0.1083	-0.2102	0.2141	0.00	0.9856
Grade	3	1	0.2181	0.0640	0.0927	0.3435	11.62	0.0007
Grade	4	1	0.2328	0.0980	0.0407	0.4248	5.64	0.0175
Degree of Curve		1	0.0599	0.0129	0.0347	0.0851	21.66	<.0001
Passing Lane		1	-1.6083	0.7196	-3.0186	-0.1979	5.00	0.0254
No of Rainy Days		1	-0.0056	0.0020	-0.0096	-0.0016	7.48	0.0062
No of Snowy Days		1	0.0126	0.0026	0.0075	0.0176	23.39	<.0001
Scale		0	1.0000	0.0000	1.0000	1.0000		

For all of the models developed, there is an increase in total crashes and F+I crashes with the increase in VMT, which conforms to the results obtained from simple SPFs (in simple SPFs, AADT is used as exposure factor). It indicates that the greater the number of vehicles, the higher the crash frequency.

Also, more crashes (both total crashes and fatal+injury crashes) can be expected with the increase in the degree of curve and number of snowy days. A decrease in total crashes is observed with the increase in the number of rainy days, except for two-lane two-way flat and rolling roadways.

For flat and rolling two-lane two-way roadways, the upgrade was found safer compared with downgrade for both seasons. Upgrade for winter was found safer with respect to summer. For winter seasons, the steep downgrade was found to be riskier than the moderate downgrade. A moderate downgrade for summer was riskier, and steep a downgrade was safer. Horizontal curvature was found riskier during the winter seasons. For both seasons, there was a reduction in crashes with an increase in shoulder width. The relationships of the presence of rumble strips and the presence of passing lanes to total crashes were found statistically insignificant at a 0.05 significance level. Shoulder width was more effective during winter seasons compared with the summer seasons.

For mountainous two-lane two-way roadways, there was no statistical relationship between grade and total crashes for the summer season. Downgrades were found riskier for winter seasons. For both seasons, an increase in crashes was observed with the increase in the degree of curve. For both seasons, no statistical relationship was found between total crashes to passing lanes, rumble strips, and shoulder width.

4.2.3.1 Comparison of NB and ZIP Models

Since two separate model types were utilized, out of necessity due to crash infrequency, comparisons of NB and ZIP models were made. These are shown in Table 4.28.

Table 4.28 Comparison of NB and ZIP Models

	Negative Binomial Distribution			**Zero Inflated Poisson Distribution**		
Criterion	DF	Value	Value/DF	DF	Value	Value/DF
Deviance	1686	1606.6584	0.9529		4094.3392	
Scaled Deviance	1686	1606.6584	0.9529		4094.3392	
Pearson Chi-Square	1686	1838.3471	1.0904	1686	2181.3413	1.2946
Scaled Pearson X2	1686	1838.3471	1.0904	1686	2181.3413	1.2946
Log Likelihood		-1117.9699			-1165.5510	
Full Log Likelihood		-1999.5885			-2047.1696	
AIC (smaller is better)		4011.1771			4106.3392	
AICC (smaller is better)		4011.2270			4106.3891	
BIC (smaller is better)		4043.7755			4138.9376	

For both terrain types, the NB model was found to be the better fit. The parameter “Value/DF” for Pearson Chi-Square and Scaled Pearson $\chi 2$ is a measure that defines the deviation of predicted values from observed values and is better if close to 1. For both terrains, the measures are closer to 1 in the NB model. Also, the measures like “Log Likelihood” and “Full Log Likelihood” are higher for NB models where a higher value of likelihood indicates a better model. The values of AIC, AICC, and BIC are all lower for the NB models compared with ZIP models. All these goodness-of-fit measures indicate that the NB model is preferable to the ZIP model.

5. CONCLUSIONS

The calibration of the Highway Safety Manual (HSM) Safety Performance Functions (SPFs) to local conditions is a necessary step for a better estimation of the expected safety performance of various roadway facilities. The HSM SPFs were developed using data from only a few states that have different climate conditions, roadway geometry, and vehicle compositions. Although the HSM already has SPFs and it can be calibrated for local conditions, jurisdiction-specific SPFs are considered better for crash prediction. SPFs have two main applications: crash prediction and identification of high crash locations. The main objective of this research is to investigate whether the calibrated HSM SPFs, Wyoming-specific simple SPFs, or the Wyoming-specific full is the best for more accurate crash prediction.

Data were collected from various sources and Wyoming-specific SPFs were developed using crash data from 2003 to 2013 for both total and F+I crashes. SPFs were developed for rural two-lane two-way roadways for flat and rolling and mountainous terrains. First, simple SPFs were developed. Wyoming-specific simple SPFs were developed for different seasons (summer and winter) and different weather conditions (clear, snowy, and rainy conditions). Wyoming-specific full SPFs were later calibrated for the same roadway classification, terrain, and seasons and can be compared to simple SPFs. Within SPF development, segmentation is an important step in data preparation since it determines the goodness-of-fit of the model. A homogeneous segmentation approach was used for all the models.

The negative binomial model was used to develop Wyoming-specific SPFs. However, for rural two-lane two-way roadways, alternative models were also developed, since many segments had zero crashes. Goodness-of-fit measures were applied for NB and ZIP models in order to determine the best fitting model. The NB models were found to fit better for the crash data. Most of the variables were statistically significant; however, some coefficients had higher p-values, but were still retained in the model to quantify their effects on crash prediction. The overdispersion parameter is used to account for additional variability in the crash data. In addition, different goodness-of-fit measures were used to compare the reliability of NB models developed for rural two-lane two-way roadways.

A two-lane two-way roadway segment was selected randomly to compare the Wyoming-specific SPF to the calibrated HSM SPF adjusted with CMFs and a calibration factor. A paired t-test was performed to check the statistical difference between the expected number of crashes obtained from using calibrated HSM SPFs versus Wyoming-specific simple and full SPFs. It was found that the calibrated HSM SPF under-predicted the number of total crashes on the 16-mile roadway segments. Hence, it can be concluded that the calibrated HSM SPFs might not be the most adequate to predict crash frequencies in Wyoming. For the same segment, the Wyoming-specific simple SPF over-estimated the number of crashes. However, the Wyoming-specific full SPF was statistically more accurate in predicting the number of total crashes. Hence, the Wyoming-specific full SPF is recommended for estimating the expected number of crashes.

6. REFERENCES

1. Highway Safety Manual. s.l. : American Association of State Highway and Transportation Officials, 2010.

2. Miaou, S, and H Lun. 1994. "The Relationship Between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions" Accident Analysis & Prevention.

3. Persaud, B, C Lyon, and T Nguyen. 1999. "Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement" Transportation Research Record: Journal of the Transportation Research Board.

4. Hauer, E, F Harwood, F Council, and M Griffith. 2002. "Estimating Safety by the Empirical Bayes Method: A Tutorial" Transportation Research Record: Journal of the Transportation Research Board.

5. Shankar, V, F Mannering, and W Barfield. 1995. "Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies" Accident Analysis & Prevention, 3 ed.

6. Chin, H, and M Quddus. 2003. "Applying the Random Effect Negative Binomial Model to Examine Traffic Accident Occurrence at Signalized Intersections" Accident Analysis & Prevention.

7. Climate Change will lead to fewer traffic accidents, Research suggests. Gothenburg, University of. 2010, Science Daily.

8. Ahmed, M, H Huang, M Abdel-Aty, and B Guevara. 2011. "Exploring a Bayesian Hierarchical Approach for Developing Safety Performance Functions for a Mountainous Freeway" Accident Analysis & Prevention.

9. Ahmed, Mohamed, M Abdel-Aty, and R Yu. 2012. "Assessment of Interaction of Crash Occurrence, Mountainous Freeway Geometry, Real-Time Weather, and Traffic Data" Transportation Research Record: Journal of the Transportation Research Board.

10. Hauer, E. 1997. Observational Before–After Studies in Road Safety. London: Pergamon Publication.

11. Lee, C, B Hellinga, and F Saccomanno. 2003. "Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic" Transportation Research Record: Journal of the Transportation Research Board.

12. Wang, C, M Quddus, and S Ison. 2013. "A Spatio-Temporal Analysis of the Impact of Congestion on Traffic Safety on Major Roads in the UK" Transportmetrica A: Transport Science.

13. Chen, F, S Chen, and X Ma. 2016. "Crash Frequency Prediction Using Real-Time Environmental and Traffic Data and Unbalanced Panel Data Models" International Journal of Environmental Research and Public Health.

14. Resende, P TV, and R F Benekohal. 1997. "Effects of Roadway Section Length on Accident Modeling." In Traffic Congestion and Traffic Safety in the 21st Century: Challenges, Innovations, and Opportunities.

15. Cafiso, S, A Di Graziano, G Di Silvestro, G La Cava, and B Persaud. 2010. "Development of Comprehesive Accident Models for Two-Lane Rural Highways Using Exposure, Geometry, Consistency, and Context Variables" Accident Analysis & Prevention.

16. Ogle, J H, P Alluri, and W A Sarasua. 2011. "Model Minimum Uniform Crash Criteria and Minimum Inventory Roadway Elements: Role of Segmentation in Safety Analysis" Transportation Research Board 90th Annual Meeting.

17. Berhanu, G. 2004. "Models relating to traffic safety with road environment and traffic flows on arterial roads in Addis Ababa" Accident Analysis & Prevention.

18. Gårder, P. 2006. "Segment characteristics and severity of head-on crashes on two-lane rural highways in Maine" Accident Analysis & Prevention.

19. Khattak, A, P Kantor, and F Council. 1998. "Role of Adverse Weather in Key Crash Types on Limited-Access Roadways: Implications for Advanced Weather Systems." Transportation Research Record: Journal of the Transportation Research Board.

20. Edwards, J. 1998. "The Relationship Between Road Accident Severity and Recorded Weather" Journal of Safety Research.