2011-03-17

Calibration of the Highway Safety Manual Safety Performance Function and Development of Jurisdiction-Specific Models for Rural Two-Lane Two-Way Roads in Utah

Bradford Keith Brimley
Brigham Young University - Provo

Calibration of the Highway Safety Manual Safety Performance

Function and Development of Jurisdiction-Specific Models

for Rural Two-Lane Two-Way Roads in Utah

Bradford Brimley

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Mitsuru Saito, Chair
Grant G. Schultz
C. Shane Reese

Department of Civil and Environmental Engineering

Brigham Young University

April 2011

ABSTRACT

Calibration of the Highway Safety Manual Safety Performance
Function and Development of Jurisdiction-Specific Models
for Rural Two-Lane Two-Way Roads in Utah

Bradford Brimley
Department of Civil and Environmental Engineering, BYU
Master of Science

This thesis documents the results of the calibration of the Highway Safety Manual (HSM) safety performance function (SPF) for rural two-lane two-way roadway segments in Utah and the development of new SPFs using negative binomial and hierarchical Bayesian modeling techniques. SPFs estimate the safety of a roadway entity, such as a segment or intersection, in terms of number of crashes. The new SPFs were developed for comparison to the calibrated HSM SPF. This research was performed for the Utah Department of Transportation (UDOT).

The study area was the state of Utah. Crash data from 2005-2007 on 157 selected study segments provided a 3-year observed crash frequency to obtain a calibration factor for the HSM SPF and develop new SPFs. The calibration factor for the HSM SPF for rural two-lane two-way roads in Utah is 1.16. This indicates that the HSM underpredicts the number of crashes on rural two-lane two-way roads in Utah by sixteen percent.

The new SPFs were developed from the same data that were collected for the HSM calibration, with the addition of new data variables that were hypothesized to have a significant effect on crash frequencies. Negative binomial regression was used to develop four new SPFs, and one additional SPF was developed using hierarchical (or full) Bayesian techniques. The empirical Bayes (EB) method can be applied with each negative binomial SPF because the models include an overdispersion parameter used with the EB method. The hierarchical Bayesian technique is a newer, more mathematically-intense method that accounts for high levels of uncertainty often present in crash modeling. Because the hierarchical Bayesian SPF produces a density function of a predicted crash frequency, a comparison of this density function with an observed crash frequency can help identify segments with significant safety concerns.

Each SPF has its own strengths and weaknesses, which include its data requirements and predicting capability. This thesis recommends that UDOT use Equation 5-11 (a new negative binomial SPF) for predicting crashes, because it predicts crashes with reasonable accuracy while requiring much less data than other models. The hierarchical Bayesian process should be used for evaluating observed crash frequencies to identify segments that may benefit from roadway safety improvements.

Keywords: safety performance functions, Highway Safety Manual, crash modification factors, negative binomial, empirical Bayes, hierarchical Bayes, safety

ACKNOWLEDGMENTS

I would not have completed this work without the help and support of a number of people. Dr. Mitsuru Saito provided guidance and direction throughout the research and helped me during the writing process. Dr. Grant G. Schultz gave insightful suggestions, feedback, and encouragement. Dr. C. Shane Reese did what much of the engineering literature was incapable of doing—he explained Bayesian statistics at a level I could understand. Andrew Olsen spent a number of hours producing the negative binomial models and developing the code for the hierarchical Bayesian analyses.

The technical advisory committee at the Utah Department of Transportation provided funding and direction for this research. I express thanks specifically to Robert Hull, Scott Johnson, David Stevens, and Tim Taylor for participating in committee meetings to discuss the data and process for developing and using these models. It is reassuring to know how much concern they have for the public's safety—they recognize the need to reduce crashes on our roadways, and that the first step is to better understand the factors that are related to crashes.

I am grateful for my family that has encouraged me throughout my life and education. Finally, I express gratitude to my wife, Amanda, for providing more love, support, and patience than I deserve. This final product represents significant sacrifice from her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The Highway Safety Manual (HSM), published in 2010 by the American Association of State Highway and Transportation Officials (AASHTO), is designed as a resource for transportation professionals to make educated decisions that affect the safety of a roadway. AASHTO states that the HSM "assembles currently available information and methodologies on measuring, estimating and evaluating roadways in terms of crash frequency and crash severity" (AASHTO 2010, p. xxvii). Safety performance functions (SPFs) are some of those tools available in the HSM and are used to quantitatively measure the safety of a roadway in terms of number of crashes.

SPFs are crash prediction models (the term "model" is hereon used synonymously with SPF). They incorporate known information about a roadway entity into an equation that estimates the safety of the entity as a yearly crash frequency. The SPFs can thus be used in two ways. One is to predict the safety of a roadway for a future time period (such as after its construction or a realignment). The other is to identify locations (or sites) that experience extreme crash frequencies, where the observed frequency (the number of actual crashes) is much higher than the predicted value.

SPFs that accurately represent observed crash frequencies are valuable to state and local transportation agencies. The SPFs in the HSM have been developed through extensive research across the United States. However, AASHTO recognizes that there are many factors affecting safety, such as “driver populations, local roadway and roadside conditions, traffic composition, typical geometrics, and traffic control measures” (AASHTO 2010, p. xxviii). As these factors vary among jurisdictions, predicting crashes in one specific location is difficult when using a model developed from nationwide data. The HSM SPFs can thus be calibrated to account for these local factors that affect the safety of a roadway or network. The Utah Department of Transportation (UDOT) desired to calibrate the HSM SPF for rural two-lane two-way roads to fit local conditions for use in evaluating roadway safety. Also, because the HSM encourages developing new jurisdiction-specific SPFs when sufficient data exist and resources allow, new SPFs were also developed.

The following sections discuss the purpose and need for this research and the organization of the thesis.

1.1 Purpose and Need

The purpose of this project is to develop multiple SPFs for rural two-lane two-way roads in Utah by calibrating the HSM model and creating new jurisdiction-specific models using data from 2005-2007. It is anticipated that UDOT will be able to use one (or more) of these SPFs during the planning stage and in evaluating the safety of the roadways in its jurisdiction.

Rural roads play a major part in the nation’s highway system, especially in the western part of the country, where cities are less densely distributed and rural highways are used to provide access between urban areas. According to the National Highway Transportation Safety Administration (NHTSA), of the 33,808 crash-related fatalities in America during 2009, 19,259 (or 57 percent) occurred on rural roads (NHTSA 2010b). Understanding factors that affect the number of crashes on rural roads can help transportation planners make decisions that improve their safety. A model that accurately predicts the number of crashes on a roadway can help officials locate hot spots—locations with unusually high crash rates—and evaluate them to determine how to lower the number of observed crashes in future years.

In accordance with Title 23 of the United States Code, states are required to submit a yearly *5 Percent Report* that documents at least five percent of the locations in the state's highways with the most severe safety needs (USC 2006). SPFs can help determine the top five percent locations for crashes because they consider multiple factors that contribute to the safety of an entity and allow the user to compare roadways with different characteristics. A highway segment, for example, that experiences a crash frequency of six crashes in a single year, for which only two crashes are predicted, may be a candidate for a dangerous location on that year's *5 Percent Report.* At a minimum, the cause of the high crash frequency merits further investigation.

SPFs establish a fairer basis for evaluating the safety of roadway entities than singly using crash rates and frequencies, because they consider multiple factors that affect safety. They can be used for evaluating safety before a new roadway is constructed, predicting the change in safety as a result of an improvement, or comparing crash frequencies across a complete highway network.

1.2 Report Organization

Chapter 1 presented the general background for this research and its purpose and need. Chapter 2 provides background information on SPFs and crash modeling in general. Chapter 3 is a discussion of current findings in the literature. Chapter 4 presents the data collection process and model development methods. Chapter 5 contains the modeling results. Chapter 6 presents the conclusions of this research and recommendations for UDOT.

2 BACKGROUND INFORMATION

Safety has often been secondary to what some consider the more urgent concerns of our transportation systems: congestion, travel times, air pollution, and fuel consumption (Lord and Persaud 2004). In recent years, safety has received more attention as federal agencies have focused on reducing the yearly number of crashes and fatalities in America. Not only has there been a steady decline in the number of injuries from crashes since 1999 (NHTSA 2010a), but the number of fatalities in 2009 (33,808) was the lowest on record in the United States since 1950 (NHTSA 2010b).

Much of the decline in injuries and fatalities can likely be attributed to increased safety regulations (such as seat belt laws or manufacturing requirements) and efforts to make motorists aware of safety issues (such as cell phone use while driving). However, NHTSA attributes the noticeably sharp declines in the number of crashes and fatalities in 2008 and 2009 to economic changes, such as the rise in unemployment (NHTSA 2010a). These components of safety complicate forecasting crashes, because they are generally not noticed until after their occurrence. Despite these difficulties, there is still value in developing crash prediction models because other noticeable relationships can be found and measured.

This chapter introduces the reader to using the HSM prediction method; calibrating the HSM SPFs and developing jurisdiction-specific models; hierarchical Bayesian methods, which can be used for developing SPFs; data needs; and data biases that should be avoided. A summary of the chapter is also provided.

2.1 Using the HSM Predictive Method

The predictive method, contained in Chapters 10–12 (Part C) of the HSM, "provides a quantitative measure of expected average crash frequency" (AASHTO 2010, p. C-1) which may be used based on existing roadway conditions or for future conditions, such as a projected average annual daily traffic (AADT). Each chapter of Part C of the HSM focuses on a different classification of road. Chapter 10 is for rural two-lane two-way roads, Chapter 11 is for rural multilane highways, and Chapter 12 is for urban and suburban arterials. SPFs are provided for both roadway segments and intersections. This study focuses on rural two-lane two-way road segments as discussed in Chapter 10 of the HSM.

Chapters 1–9 of the HSM (Parts A and B) discuss, among other topics, the selection of countermeasures for reducing crashes, economic appraisals of the countermeasures, and prioritization of projects. When used with these chapters of the HSM, SPFs can help practitioners identify the locations that can benefit most from cost-effective improvements. For example, a hot spot (a site with an abnormally high crash frequency) may have a large number of run-off-road (ROR) crashes that can be reduced by increasing the shoulder width or installing shoulder rumble strips. By improving the geometry or surrounding conditions of a roadway, crashes may be reduced because the road is more forgiving to mistakes made by drivers. SPFs can help pinpoint locations where such changes may improve the safety of a roadway.

This following subsections discuss SPFs, crash modification factors (CMFs), and the Empirical Bayes (EB) method.

2.1.1 Safety Performance Functions (SPFs)

SPFs utilize known information about a roadway, such as geometry and AADT, to predict the number of crashes and their severity on a segment or at an intersection. As discussed in Chapter 1, some aspects of safety, such as policy regulations or economic changes, are very difficult to include in an SPF because they are difficult to define or often not measured until after their occurrence. The continual changes in the observed safety of roadways make it difficult to determine what variables should be used to predict the number of crashes at a given site. The only parameter that changes from year-to-year in the SPFs given in the HSM is AADT (unless the geometry changes from an improvement).

Chapter 10 in the HSM provides a crash prediction model for rural two-lane two-way road segments. The SPF was developed from studies that have involved a number of areas of the United States and may be calibrated to better predict the safety of a specific jurisdiction. Equation 2-1 is the SPF for rural two-lane roads meeting the base conditions as documented in the HSM (AASHTO 2010).

$$N_{spf} = AADT \times L \times 365 \times 10^{-6} \times e^{-0.312} \tag{2-1}$$

where, N_{spf} = predicted number of yearly crashes,

$AADT$ = average annual daily traffic, and

L = segment length (mi).

This model infers that crashes are directly proportional to the exposure in million vehicle miles traveled (MVMT)—note the coefficients 365 and 10^{-6} used to convert AADT and segment length to MVMT. Figure 2-1 shows the predicted crash rate (crashes per mile) for two-lane rural roads based on AADT. This model is used for segments meeting the base conditions and considers only exposure (AADT and segment length). The base conditions are defined in the HSM and presented in Table 2-1.

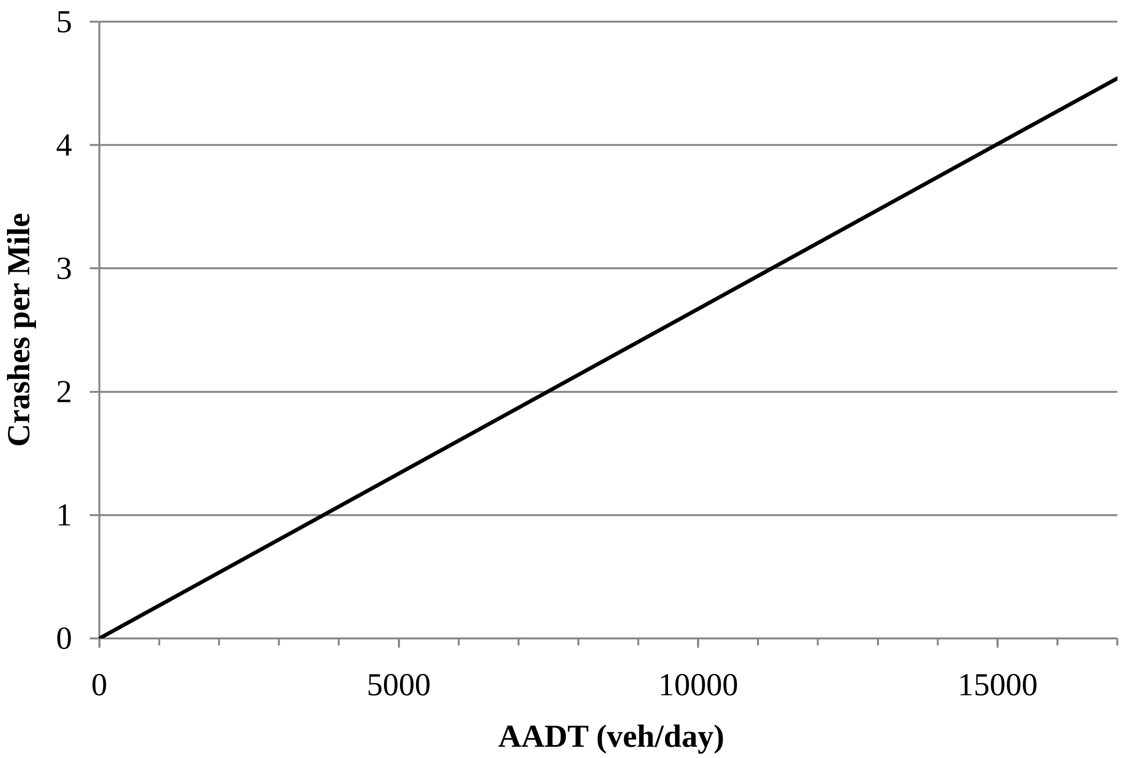

Figure 2-1: Predicted crash rate based on AADT (adapted from AASHTO 2010).

Table 2-1: Base Conditions for Rural Two-Lane Two-Way Roads (AASHTO 2010)

Lane Width: 12 ft	Vertical Curvature: None
Shoulder Width: 6 ft	Centerline Rumble Strips: None
Shoulder type: Paved	Passing Lanes: None
Roadside Hazard Rating: 3	Two-Way Left Turn Lanes: None
Driveway Density: 5 driveways per mile	Lighting: None
Horizontal Curvature: None	Automated Speed Enforcement: None

2.1.2 Crash Modification Factors (CMFs)

CMFs are used when the characteristics of a site deviate from the base conditions given in the HSM. CMFs are multiplied to the base prediction value (N_{spf}) which adjusts the base predicted crash frequency to that of the actual conditions. Equation 2-2 shows this relationship. A CMF greater than 1.0 indicates an increase in predicted crashes due to the non-base condition, while a CMF less than 1.0 represents a reduction in crashes.

$$N = N_{spf} \times CMF_1 \times CMF_2 \times \dots CMF_i \tag{2-2}$$

where, N = predicted number of crashes considering non-base conditions,

CMF_i = crash modification factor, and

N_{spf} = yearly number of predicted crashes from Equation 2-1.

The calibration procedure given in the HSM allows calibration of two groups of segments: those that conform strictly to base conditions, where each CMF = 1.0; and those with a variety of characteristics, whose CMFs are then included in the total prediction. This study focuses on the latter case and includes a calibration of segments that do not all conform to the HSM base conditions. This was done because it was practically impossible to find enough study segments that strictly conform to the base conditions in Table 2-1.

2.1.3 The Empirical Bayes (EB) Method

The EB method addresses two problems associated with crash prediction models: regression to the mean (RTM) and lack of data when few years of crash history are available (Hauer et al. 2002). RTM is a common bias when evaluating a network for sites with high crash frequencies because an area with a high number of crashes in one year may have a smaller, more typical, crash frequency the following year. The precision of a crash prediction model is low when it is based on a limited time period because it is unlikely that the small time period is representative of the true average crash frequency.

The EB method is used in recognition that the safety of a site is best estimated by considering both the number of observed crashes at the site and the number of crashes at sites with similar characteristics, as predicted by the SPF (Hauer et al. 2002). Use of the EB method produces the *expected* number of crashes at a site through a mathematical combination of the predicted and observed crash frequencies. Equation 2-3 is used to estimate the expected crash frequency. The weight, w, used in Equation 2-3, is calculated by combining the model's overdispersion parameter (φ) and N_{spf}, as shown in Equation 2-4.

$$N_{expected} = w \times N_{spf} + (1 - w) \times N_{observed} \tag{2-3}$$

where, $N_{expected}$ = expected number of crashes determined by the EB method,

w = weight determined by Equation 2-4, and

$N_{observed}$ = observed number of crashes at a site.

$$w = \frac{1}{1 + \varphi \times (N_{spf})} \tag{2-4}$$

where, φ = overdispersion parameter.

N_{spf} is used in both Equations 2-3 and 2-4 and is the number of predicted crashes. The overdispersion parameter (φ) is an element of the negative binomial distribution, which is indicative of the spread of the crash data. If the overdispersion parameter is large (indicating that the crash data are widely dispersed), the EB method places less weight on the SPF prediction.

2.2 Calibrating the HSM SPFs and Developing Jurisdiction-Specific SPFs

Driver behavior is difficult to predict or model. In fact, almost all crashes are a result of human errors caused by impairment, distractions, or aggressive behavior. Because these factors and their effects may be so diverse across different areas, calibrating models for local jurisdictions and development of new jurisdiction-specific SPFs are encouraged to produce the most accurate models that best predict safety (AASHTO 2010). This section discusses calibration of the HSM models and development of jurisdiction-specific SPFs.

2.2.1 HSM Model Calibration

Model calibration in the HSM is performed by applying a multiplicative factor to the given SPF so that the aggregate number of predicted crashes is equal to the aggregate number of observed crashes throughout a jurisdiction. A calibration factor allows the SPF to keep its original model form.

As discussed in the Appendix to Part C of the HSM, selected samples are used to find the calibration factor that will make the aggregate predicted crash frequency equal to the observed total in the jurisdiction. The HSM recommends using a minimum of 30-50 sites that are selected without regard to their crash frequencies (AASHTO 2010). Since the SPF calibration is for rural two-lane two-way roads in Utah, the local jurisdiction is the entire state. Equation 2-5 illustrates how to use the calibration factor:

$$N_{local} = C \times N_{spf} \tag{2-5}$$

where, N_{local} = total predicted crashes in a local jurisdiction, and

C = calibration factor.

The calibration factor is calculated by rearranging Equation 2-5 and substituting an observed crash frequency for N_{local} as recommended by the HSM (AASHTO 2010).

$$C = \frac{N_{observed}}{N_{spf}} \tag{2-6}$$

2.2.2 Development of Jurisdiction-Specific SPFs

When enough data are available, the HSM allows users to create jurisdiction-specific SPFs. It is recommended that SPFs be developed using negative binomial regression techniques to account for the dispersion present in crash data and estimate the overdispersion parameter. The SPFs must use exposure (AADT and segment length) and should be able to be converted to the HSM base conditions (if developed using non-base values) (AASHTO 2010).

2.3 Hierarchical (Full) Bayesian Methodology

In light of the number of unknown parameters affecting crash frequencies that are otherwise impossible to quantify, a full Bayesian approach is proposed as a method for estimating the true effect of parameters on the variable of interest. Bayesian methods are also valuable because they provide accurate estimates when few data are available, a circumstance commonly encountered with modeling crashes. This section is based on a thorough discussion of Bayesian statistics in Gelman et al. (2004).

Bayes' Theorem, the foundation for Bayesian applications, is outlined in Equation 2-7.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2-7}$$

where, *P(A|B)* = Conditional probability (or posterior) of *A*, given B,

P(B|A) = Conditional probability (or likelihood) of B, given A,

P(A) = Unconditional (or prior) probability of A, and

P(B) = Unconditional probability of B.

Bayesian models use the density function to estimate the effect of a given parameter on the model rather than a discrete coefficient. Use of the density function allows for greater understanding of the amount of uncertainty in the data (Lan et al. 2009). The density function for each parameter provides the likelihood that it has a certain predicting effect. In Equation 2-7, the conditional probability values may be replaced by density functions.

The unconditional prior distribution of *A* (simply referred to as a "prior") is information that is known about the parameter beforehand, or *a priori*. Priors are unconditional because they are not dependent upon other parameters, but represent the likelihood of a particular occurrence. For example, an informative prior may be the probability density function of AADT on rural two-lane two-way roads in Utah.

The application of Bayes' Theorem to a crash prediction model results in Equation 2-8. The conditional probabilities have been replaced by density functions. For this application of

Equation 2-7, the parameters θ and y are used in place of A and B, respectively, where θ represents a vector of the modeling parameters and y represents the known data.

$$f(\theta|y) = \frac{f(y|\theta)P(\theta)}{P(y)} \tag{2-8}$$

where, y = crash frequency, and

θ = vector for the unknown modeling parameters.

Because of the uncertainty associated with crash modeling, it is necessary to determine the probability of θ, given y (or $f(\theta|y)$), in order to later make inferences about the entire population. Thus $f(\theta|y)$ is represented as the posterior distribution ($\pi(\theta|y)$), shown in Equation 2-9. The prior distribution of θ is represented by $\pi(\theta)$.

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{P(y)} \tag{2-9}$$

where, $\pi(\theta|y)$ = posterior distribution of θ, determined from known crash data,

$f(y|\theta)$ = likelihood of y given θ, and

$\pi(\theta)$ = informational prior distribution of θ.

P(y), called the normalizing constant, is solved through integrating $f(y|\theta)\pi(\theta)d(\theta)$. Substituting this into Equation 2-9, the posterior distribution of θ is solved by Equation 2-10.

$$\pi(\theta|y) = \frac{f(\theta|y)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta} \tag{2-10}$$

The integral in the denominator makes Equation 2-10 a true density function. When multiple parameters are used for the vector θ, the denominator in Equation 2-10 must be integrated for each defined parameter. This is a very difficult equation to solve mathematically. Rather than performing the complex integration, Markov Chain Monte Carlo (MCMC) methods, specifically Metropolis-Hastings and Gibbs sampling, can be used to sample from the posterior distribution. Using thousands of samples, the true posterior distribution can be accurately estimated.

2.4 Data Needs

The HSM specifies what data are required for performing SPF calibration or developing new models. Data requirements for each model can be found in the HSM chapter for the respective model (Chapters 10–12). For developing new jurisdiction-specific SPFs in this study, variables other than those required by the HSM were included that were believed to contribute to crash frequencies. This section discusses the two types of data required for modeling crashes: crash data and facility data.

2.4.1 Crash Data

The prediction capability of a model is dependent upon the level of detail of the data. Regarding crash data, HSM model calibration only requires the total crash frequency for a period of one or more years (equal to the time period for which the calibration factor will be used). If desired, other crash data, such as crash severity, collision type, and contributing factors, can be included to evaluate these specific occurrences (AASHTO 2010). The KABCO scale is commonly used for defining severity levels: K-fatality, A-incapacitating injury, B-non-incapacitating injury, C-possible injury, O-no injury (NHTSA 2008).

2.4.2 Facility Data

Facility data include site characteristics such as roadway classification; geometric conditions (number of lanes, beginning and ending points of segments, presence of medians and rumble strips, lane and shoulder widths, curvature, traffic control devices, and lane configurations); and traffic volume data (including AADT for both major and minor streets if intersection crashes are being modeled). The HSM recognizes that some important data may not be readily available, such as radii of horizontal curves (AASHTO 2010). In this study, only tangent sections of roadways were included for both the calibration and development of new models.

2.5 Data Biases

The accuracy of an SPF is determined by the quality of its representative data. This section discusses the three sources of bias that can result in inaccurate models: inaccurate data, variation in crash reporting thresholds, and differences in crash reporting methods or definitions by agencies (AASHTO 2010).

2.5.1 Inaccurate Data

Data accuracy is crucial because the crashes represented by a particular model are only those that are reported as occurring within the boundaries of a specific time and place, regardless of the number of crashes that actually occurred. These crashes are the “observed” crashes. If crash locations (based on route and milepost) are recorded incorrectly by law enforcement, some crashes may not be included in the model, or some may be included that did not actually occur at the determined site.

2.5.2 Reporting Thresholds

The 2008 Model Minimum Uniform Crash Criteria (MMUCC) recommends that all crashes involving death, injury, or property damage valued at $1000 or greater be reported (NHTSA 2008). However, the MMUCC guidelines are not regulations and not every state uses the $1000 level. Each state has its own reporting threshold, and this amount may change with legislation. This should be considered when comparing crash rates between states. At the time this study was conducted, the crash reporting threshold for the state of Utah was $1000.

Property Damage Only (PDO) crashes are underreported because drivers, who are responsible for reporting crashes, tend to avoid reporting PDO crashes than face negative incidents on their driving record and accompanying insurance rate increases. Because PDO crashes are underreported, it is not uncommon for SPFs to only consider crashes with severity levels of K, A, B, or C since those crashes are reported more consistently (Fitzpatrick et al. 2008; Griffin et al. 1998; Park and Lord 2008). This study includes all levels of severity because the removal of PDO crashes from the dataset would reduce the number of observed crashes below a reasonable level for modeling.

2.5.3 Crash Reporting Methods and Definitions

Crash reporting methods may differ among the law enforcement agencies across the state. The MMUCC, developed by a committee of professionals from law enforcement agencies, safety and medical agencies, and governments, is an example of collaboration among agencies to develop standardized crash reporting methods. It provides definitions and guidelines for information that should be collected in a crash report to help maintain consistency between different law enforcement agencies. The third edition of the MMUCC, published in 2008, contains 107 data elements, 75 of which should be collected at the scene of the crash. The other elements may be obtained from crash scene information or data files maintained by the state (NHTSA 2008).

The following are some of the specific data that should be collected in crash reports according to MMUCC guidelines: date and time of crash, weather conditions, light conditions, type of intersection, make and model of motor vehicle, direction of travel before crash, total lanes in roadway, roadway alignment and grade, most harmful event for the motor vehicle, contributing circumstances, cargo body type, air bag deployment, gender and age of all persons involved, and alcohol test results (NHTSA 2008). Law enforcement should be responsible for the accuracy of these data, despite the level of detail required by so many data fields.

2.6 Summary

The HSM is a guide for engineers and agencies responsible for roadway safety. The SPF equations in Part C (Chapters 10–12) of the HSM can be used to predict the number of crashes of an entity, such as a road segment or intersection, for a given time period. Additionally, the EB method is a proven technique used to determine the *expected* crash frequency. This is accomplished through combining the predicted and observed crash frequencies using the defined weighting factor. When calibrated, the predictive capability of an SPF improves and can be used across an entire jurisdiction. If enough data are available, completely new jurisdiction-specific SPFs may be developed. These new models can take advantage of functional forms that can better fit local trends that are not shown in the HSM SPFs.

The hierarchical Bayesian approach better accounts for variability in the data by using density functions to estimate the posterior distribution of a parameter. A density function for crash frequencies can then be determined, which, when compared with the observed crash frequencies for individual road segments, can identify dangerous segments.

Accurate data collection is necessary for the models to represent actual conditions. If crashes are not reported correctly, the resulting model may underpredict or overpredict the crash frequencies of future time periods. If facility data are not collected correctly, the resulting model will not accurately represent contributing factors associated with crashes and crash predictions. Thresholds for reporting crashes, in addition to reporting methods, may vary by jurisdiction. It is important for agencies to maintain as much consistency as possible in reporting crashes so that correct comparisons and inferences can be made. The MMUCC provides guidelines on collecting facility and crash data to avoid possible biases caused by these issues.

3 LITERATURE REVIEW

This chapter is the culmination of a literature search related to crash predictions and modeling. A review of the following topics discussed in the literature is presented: SPFs, including independent variables and CMFs; crash types and severities; time periods for analyses; negative binomial models; empirical Bayesian methods, and hierarchical (full) Bayesian methods.

3.1 Safety Performance Functions (SPFs)

SPFs have long been developed for various types of entities, such as rural roads, urban arterials, intersections, freeways, and even freeway ramps (Zegeer et al. 1986; Vogt and Bared 1998; Lord and Persaud 2004; Lord and Bonneson 2005; Persaud and Lyon, Inc. and Felsburg Holt and Ullevig 2009). SPFs developed from different studies for similar entities may have the same model form, but will be different because of the characteristics of the local roads and population. This is the reason for recommending a calibration of the HSM, which has already been performed in a select few states (Banihashemi 2011; Sun et al. 2011; Xie et al. 2011).

This section discusses the selection of independent variables for crash predictions and the use of CMFs in new jurisdiction-specific models.

3.1.1 Variable Selection

Development of an SPF requires selecting the variables that will be included in the model. Mayora and Rubio (2003) discovered over 50 different roadway conditions that have some influence on crash rates. However, too many variables in a crash prediction model may result in overfitting the dependent variable in the dataset. Also, limited time and resources rarely allow an agency to incorporate so much data into a model. Ultimately, the best model considers independent variables that best predict crashes and has a reasonable functional form, showing a logical connection between the variables and results.

The HSM model for two-lane two-way rural roads considers exposure, cross section geometry (lane and shoulder widths), curvature, density of driveways, and roadside hazards, among other variables, to be measureable characteristics that affect safety (AASHTO 2010). Garber and Ehrhart (2000) examined models with the standard deviation of the speed of vehicles on a roadway, and noted that more crashes occur when there is a greater speed distribution. They summarized that crash rates were higher in the early-morning and late-day hours because there is greater variation in vehicle speeds at these times when fewer vehicles are on the road. Additionally, the percentage of multiple-vehicle crashes decreased as traffic volume decreased, while the percentage of single-vehicle crashes increased. Harwood et al. (2007) noted that wider lane widths provide a buffer against driver mistakes or inattentiveness and thus result in lower crash rates. The CMFs for lane width in the HSM reinforce this. However, Hauer (1999) cited research that contradicts these findings, stating that 12-ft lanes on rural roads are less safe than 11-ft lanes.

Mayora and Rubio (2003) presented the following factors as having high correlation with crash rates: access density, sight distance, speed limit, and proportions of no-passing zones.

Zegeer et al. (1986) included variables for flat or mountainous terrain in addition to other variables now suggested in the HSM for their SPFs. Another study found an increase in overall crashes with higher speed limits (Griffin et al. 1998). However, Vogt and Bared (1998) found no such relationship between crashes and speed. Contradictions among studies underscore the fact that the ability of each model to predict the safety of a roadway is dependent upon the accuracy of the data and the uniqueness of the population. Because crashes are a result of human errors more often than system inadequacies, modeling crashes using true causal factors is nearly impossible.

Mensah and Hauer (1998) confronted the problem of modeling crashes using actual causal factors. They discussed the issue with modeling crash frequencies based on average values, such as AADT, and asserted that an immediate value of traffic flow would be a better predictor of crashes than an average daily count. Crashes are, after all, discrete events and are the result of specific causes. Similar to modeling crash rates with the variation of vehicular speed as done by Garber and Ehrhart (2000), it may be useful to find a relationship between crashes and daily or hourly traffic flow variations. The number of trucks, daily commuters, or vacationers is always changing, as well as the demographics of the drivers (age groups, gender, class, etc.). Despite the understanding that these variables make some contribution to the safety of a roadway, AADT is used to avoid the impossible task of measuring all the changing components of traffic flow.

Part of the difficulty with selecting variables to model crashes is differentiating between correlation and causality. Correlation does not indicate that something is caused by a particular factor. For example, there is always a correlation between the total crashes on a roadway segment and its length, even though the segment length itself is not the cause of crashes. SPFs thus tend to be simplistic because they often contain predictive rather than actual causal factors (Lord and Persaud 2004). As mentioned previously in this section, Hauer (1999) noted that 12-ft lanes often experience higher crash rates than 11-ft lanes. This may happen because wider lanes encourage aggressive behavior, or it may be a product of higher design standards required of roads with higher traffic volumes. If either case is true, the lane width itself is not a causal factor, but can be used to predict crashes because of its correlation to observed crash rates.

After study variables have been selected and data collected, an investigation of the correlations among the variables is used to determine their independence. If one variable is significantly dependent upon another, both variables should not be included in a model. For example, the mean travel speed should not be included in a model that also uses the speed limit as an independent variable because they are likely to be highly correlated. Using both variables would provide little benefit.

3.1.2 Crash Modification Factors (CMFs)

Study methods for developing CMFs are classified in two ways: cross-sectional studies or before-after analyses (Gross et al. 2010). Variations exist for these two classifications (e.g., applying the EB Method to a before-after analysis), but the overall process can be defined as cross-sectional or before-after. Cross-sectional analyses compare the safety of multiple sites with different characteristics, and the resulting CMFs are often taken from the coefficients of the model's variables. Several studies have developed CMFs in this manner (Fitzpatrick et al. 2008; Gross et al. 2010). CMFs from before-after studies are preferred to those based on a single time period because there is a reference established for the actual change in safety (Gross et al. 2010).

Because a single value may not correctly represent the safety effect for sites with different characteristics, it can be advantageous to develop crash modification functions rather than single CMFs (Gross et al. 2010). For instance, the effect of a roadway's shoulder width may be dependent upon the AADT and lane width. Thus, a function of AADT and geometry may be used in place of a single CMF. Issues may also arise when multiple CMFs are used simultaneously. Fitzpatrick et al. (2008) and Gross et al. (2010) recognized that using CMFs that are not completely independent of each other may result in overestimating their combined benefit. When one CMF is applied with another, the estimated combined effect may likely be less than if the modifications had been made separately. The HSM also cautions against using a large number of CMFs for a single roadway entity (AASHTO 2010).

As CMFs may be used to quantify the safety of a roadway whose characteristics vary from base conditions, or to predict safety after an improvement has been made, Hauer (1997) argued that improvements, however, may change more than just the roadway geometry. They may affect driver behavior, or at least temporarily affect the driving task. This extra change complicates before-after studies and provides reason to use the EB method. Zegeer et al. (1986) suggested not using data on roadways with recent improvements until three years after any treatment.

When applying CMFs to a prediction model, it is important to consider the method by which they were derived, whether through a cross-sectional or before-after study. CMFs developed through a cross-sectional analysis, for example, may not accurately represent an expected change in safety for a proposed improvement that a before-after analysis may provide (Gross et al. 2010).

3.2 Crash Types and Severities

Different roadway geometries affect different crash types. For example, ROR crashes are linked more closely to lane and shoulder widths than are rear end collisions (Zegeer et al. 1986).

Specific models by crash type allow researchers to discover links between certain geometric characteristics and their associated crash types. Such models that focus on a specific crash type can help policy makers and practitioners make educated decisions for road safety improvements that focus on specific types of crashes, especially those that are often fatal or cause severe injury.

Limiting models to specific crash types or severities can be difficult because the number of crashes may be too small to produce a reliable model (Jonsson et al. 2009). The HSM encourages users to predict crashes by type through multiplying the estimated total number of crashes by the percentage of jurisdiction-wide crashes of that type. This method may be inaccurate for certain applications, because it implies that the distribution of crash types is constant throughout a jurisdiction. Jonsson et al. (2009) noted that models by crash type produce a better fit than models with total crash estimates and proportions because of the nonlinear relationships between crashes by type and traffic flow.

3.3 Time Periods for Analysis

Multiple years of crash data are used for developing SPFs to compensate for low crash frequencies. This strategy is especially relevant for rural roads with low traffic volumes that often experience very few or no crashes in an average year. Because crashes are rare and random events, crash frequencies that fluctuate each year are best expressed as an average for a time period of multiple years. A multi-year time period takes advantage of the RTM phenomenon which is thoroughly discussed by Hauer (1997). For this reason, many studies that have developed SPFs used multi-year time periods (Cafiso et al. 2010; Fitzpatrick et al. 2008; Lord and Bonneson 2005; Lord and Bonneson 2007; Vogt and Bared 1998).

For calibration of the HSM SPFs or development of jurisdiction-specific SPFs, it is recommended to use a period that reflects the length of time for which the models will be used (AASHTO 2010).

3.4 Negative Binomial Models

Contemporary crash prediction models are most often developed using a negative binomial distribution. The negative binomial distribution is suited for modeling crashes because of the naturally high variability of crash frequencies whose variance is greater than the mean. Negative binomial models are also referred to as mixed Poisson-gamma models because crashes within a site fit a Poisson distribution, but the variation across multiple sites is Gamma distributed (Mitra and Washington 2007; Park and Lord 2008).

The gamma-distributed error in the negative binomial model is the source of the overdispersion parameter used in the EB method (Hauer et al. 2002). The overdispersion parameter (φ) is an indication of the precision of the model and the variability of the crash frequencies. A value of 1.0 is indicative of a fully Poisson-distributed model, while greater overdispersion values are indicative of more variability in the model.

3.5 Empirical Bayes Analyses

The EB method as discussed in Section 2.1.3 has been applied to road safety for a number of years, and received even more attention in 1997 with the publication of Hauer's *Observational Before-After Studies in Road Safety*. With the publication of the HSM, the EB method may be considered the standard in evaluating road safety. Many studies have used the EB method for various applications and there is a general concurrence regarding its usefulness in evaluating crashes and roadway safety (Elvik 2008; Persaud et al. 2010; Persaud and Lyon

2007). However, recent computing and software advances have made the more complex hierarchical (or full) Bayes technique a viable method for modeling crash frequencies (Lan et al. 2009).

3.6 Hierarchical Bayesian Analyses

As the full Bayes methodology is much more mathematically complex than negative binomial regression, its application has not yet received widespread use in safety analyses. There are some studies that propose using hierarchical Bayesian methods, which can be found in the literature (El-Basyouny and Sayed 2009; Lan et al. 2009; Olsen et al. 2011; Schultz et al. 2010; Schultz et al. 2011).

3.7 Summary

Modeling crash frequencies with SPFs has been a common technique for quantifying road safety for a number of years. By developing SPFs, researchers are able to discover variables that affect crashes and better understand how much they contribute to an entity’s safety (or lack thereof). However, care should be taken to not equate causation with correlation. Some variables may have predicting capability, but are not factors that necessarily cause crashes. There are multiple ways to develop CMFs, the most common of which is to use the variable coefficients provided in a statistical model. CMFs are valuable for predicting how a certain characteristic may affect the safety of a roadway when changed.

Negative binomial models are traditionally used to account for the dispersion present in crash data. The negative binomial distribution provides the overdispersion parameter, which assists in applying the EB method to estimate the expected crash frequency. The expected crash frequency (as developed by the EB method) is a combination of predicted and observed crash frequencies. Hierarchical Bayesian modeling, though more complex than the EB method, is a relatively new technique for producing accurate crash predictions.

4 DATA COLLECTION

The objective of the data collection process was to select study segments in a manner that was as random as possible and measure as many roadway characteristics that were hypothesized to have an effect on safety as was reasonable. Appendix A of Part C of the HSM notes that segments not conforming to the base conditions in the HSM can be used for SPF calibration and development of jurisdiction-specific SPFs (AASHTO 2010). Few rural roads in Utah conform to these strict conditions. In fact, only 14 of the 157 study segments in the final dataset meet the HSM base conditions.

This chapter discusses the data collection process, including the use of Roadview, Google Earth, UDOT traffic tables, the UDOT crash database, the UDOT 10 year construction database, and limiting the dataset to meet specific conditions; the raw data; correlation and independence of variables; data inconsistencies; limitations; and the modeling procedure. A summary of the chapter is also given.

4.1 Data Collection Process

The data collection process was structured such that study segments were selected without considering the number of crashes that occurred in a given year. The initial stages of data collection strictly involved only geometry and roadway characteristics. The latter portion involved adding AADT and crashes to the database and removing segments that had experienced construction during the study period or represented conditions not prevalent on Utah's rural two-lane two-way highways. A dataset of segments chosen without regard to crash frequencies or AADT (likely the most influential variable in causing crashes) was necessary to ensure that the data were comprised of samples that are as random as possible. UDOT annually collects AADT values for state highways and local federally sponsored roads, which are published and available to the public (UDOT 2011b). Because these data are easily accessible, all study segments are either a state or federal highway.

This section discusses the use of Roadview, Google Earth, UDOT traffic tables, the UDOT crash database, and a UDOT construction project database for selecting road segments and obtaining the necessary facility and crash data. The final subsection contains a discussion on the methodology for removing segments that did not conform to conditions prevalent on rural roads in Utah.

4.1.1 Roadview

The Roadview Explorer (UDOT 2010) provided photologs for selecting segments in the dataset. Figure 4-1 shows a photo from Roadview on Route 59. Note that the location is precise to 0.001 miles.

The selected segments were straight and homogeneous with no turning or passing lanes. The data collected from Roadview consisted of route number, starting milepost, ending milepost, type of striping used (double yellow line, solid yellow and dashed lines, or single dashed line), presence of centerline and shoulder rumble strips, speed limit, and number of driveways in each segment. The striping configuration was entered in the database as a 0, 1, or 2, designating the number of lanes with a passing opportunity, as shown in Figure 4-2.

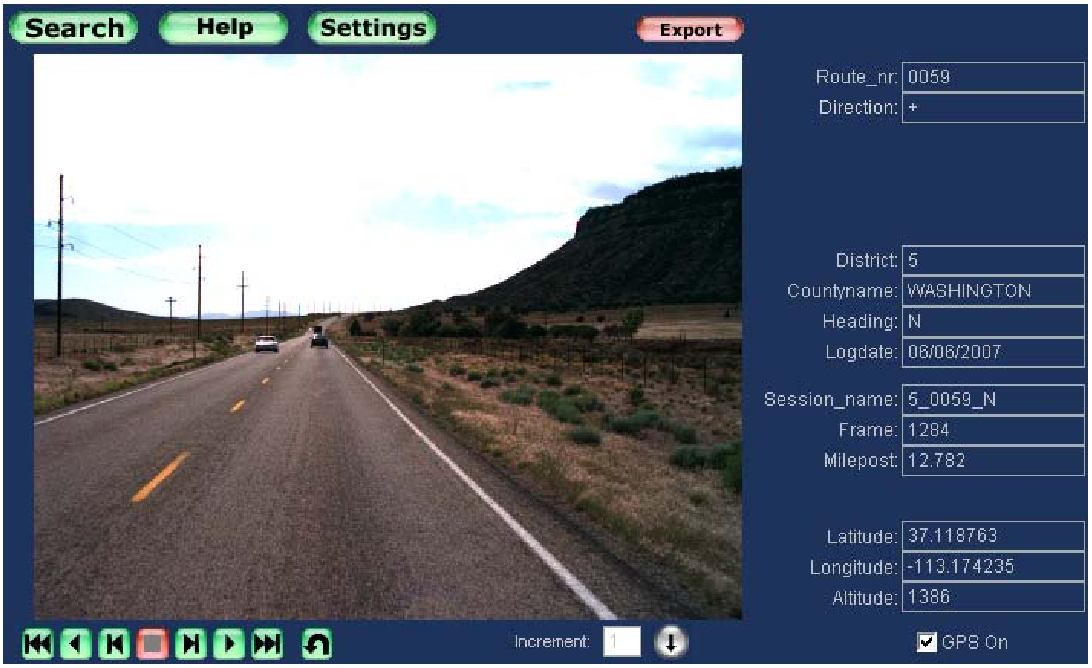

Figure 4-1: Screenshot from Roadview showing highway 59 at milepost 12.78.

The roadside hazard rating (RHR), expressed on a scale of 1-7, is a subjective rating and determined by multiple factors in the area surrounding the roadway (Harwood et al. 2000). Pictures from Roadview were initially used to rate the roadside based on sideslopes or obstacles that would severely affect a vehicle that leaves the roadway. However, there were no segments with dangerous sideslopes or obstacles in this study, because those characteristics are more often associated with winding roads in mountainous areas rather than straight, homogeneous roads. Assigning segments a subjective RHR value based on pictures provided in the HSM from Harwood et al. (2000) ceased after it was determined that there was little variability in RHRs that would result in any added benefit from studying the segments' roadside conditions.

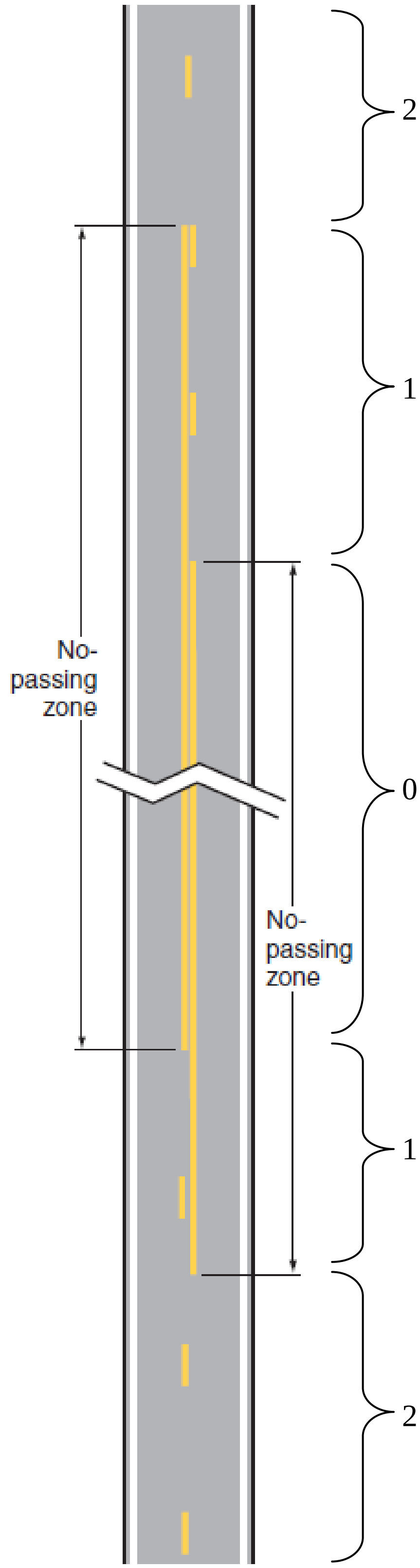

Figure 4-2: Data value entered to represent the lane striping on a road segment (adapted from figure 3B-1 in 2009 MUTCD, FHWA 2009).

Segments for the dataset were selected by the following procedure. This procedure can also be used or slightly modified to select segments for a different SPF model (such as multilane highway). Table 4-1 shows an example of the collected data.

1) Using the Roadview Explorer, select a route to examine two-lane two-way highways in rural areas.
2) Record the highway number and beginning milepost of a tangent section. Enter the speed limit, presence of shoulder and centerline rumble strips, and number of lanes with a passing opportunity (indicating roadway striping). Begin and end segments approximately 250 feet or 0.05 miles from curves, intersections, or locations where homogeneity is disrupted.
3) Visually inspect each segment to ensure that it is straight and homogeneous by clicking "play" and watching the streamlined video progress as if the viewer were the driver. Count the number of driveways or minor intersections.
4) Record the ending milepost of the homogeneous segment.

Table 4-1: Example of Data Collected from Roadview

Highway	Starting Milepost	Ending Milepost	Number of Driveways	Centerline Rumble Strip	Shoulder Rumble Strip	Passing Ability	Speed Limit
6	90.582	93.447	5	No	No	2	65
6	94.754	99.506	5	No	No	2	65
6	105.645	110.387	6	No	No	2	65
6	263.299	264.577	0	No	Yes	2	65
6	264.578	265.718	0	No	Yes	2	65
6	265.726	266.994	0	No	Yes	2	65
6	267.264	267.793	0	No	Yes	2	65

4.1.2 Google Earth

Roadview images were initially used to collect lane and shoulder width measurements. It was hypothesized that the lane width of a road could be determined by multiplying the on-screen lane widths from Roadview (generally 2-3 inches) by a factor that increased them to real-world values (mostly 11-12 ft). However, a comparison of in-field and on-screen measurements showed that this method provided inaccurate data.

Satellite imagery, provided by Google Earth (Google 2010), was a better method to measure lane and shoulder widths because removed the need to make site visits to each segment. Two measurements of the total cross-sectional width and the combined lane widths were made for each tangent section. Their values provided the average lane and shoulder widths for the model. Also, elevations at the endpoints of each segment were determined from Google Earth. The difference in two elevations was divided by the segment's length to obtain the average longitudinal grade.

UDOT has made available a layer in Google Earth that provides AADT measurements along Utah highways (UDOT 2011a). This layer gives values of AADT along defined segments of road, with each AADT segment as long as 10 or 20 miles. The data are identical to those contained in the Traffic on Utah Highways reports (UDOT 2011b). AADT data were recorded for each segment for all 3 study years (2005-2007). If a study segment ever contained two UDOT-defined AADT segments, the AADT from the longer portion of the two was used. Though not done for this study, Google Earth may be a useful application for measuring the radii of horizontal curves in the future.

4.1.3 UDOT Traffic Tables

The Truck Traffic on Utah Highways reports (UDOT 2011b) provided count data on trucks as a percentage of AADT. Values for single-unit trucks, combo-unit trucks, and a combined total truck percentage were used as variables in the database. Single-unit trucks are comprised of classes 4-7 of the Federal Highway Administration (FHWA) vehicle classification system. Combo-unit trucks are comprised of classes 8-13 (UDOT 2011b). Truck traffic through Utah is so high that some study segments have combo-unit trucks comprising more than 50 percent of AADT.

4.1.4 UDOT Crash Database

UDOT maintains a crash database with all crash-related data necessary to create useful SPFs. The database provided by UDOT was opened with Excel and delimited into fields. The crash data necessary for modeling was added to the study database using a macro that counts the number of crashes and severity within each defined segment. Other data relevant to future crash modeling activities (such as crash types, vehicle types, and ages of drivers) can be extracted from the crash records by modifying the macro. The macro used in this study can be found in the Excel workbook contained on the CD attached with this thesis.

The following statistics were recorded in the data table:

- Total number of crashes occurring on the segment,
- Number of crashes caused by wild animals,
- Number of crashes reported as occurring at a junction, and
- Number of crashes by each severity group using the KABCO scale.

As mentioned, more statistics can be added to the data table to include other characteristics about the crashes (such as other causes or interesting statistics). This information can be used to develop other SPFs that model more specific elements.

4.1.5 UDOT Construction Database

A database of all construction projects within Utah from 2000 to 2010 was used to verify that no construction had occurred on any study segments from 2004 to 2007. A buffer of one year before the study period with no construction was used rather than the three years suggested by Zegeer et al. (1986), because a longer buffer of no construction would have considerably reduced the amount of collected data.

4.1.6 Limiting the Dataset

The SPF for rural two-lane roads in the HSM is valid for highways with AADT up to 17,000 veh/day. The original 3-year dataset of 169 segments contained only a few segments with AADT greater than 10,000 veh/day. Initial modeling results showed that the most inaccurate predictions (based on residuals of observed and predicted crashes) were for segments with such extreme AADT values (even though they are allowed by the HSM model). Other poor predictions were for segments located near or within small towns. These were areas with speed limits of 45 mph or less.

Even though both of these situations are permitted in the HSM, segments that fit the criteria of characteristics with AADT values greater than 10,000 veh/day and speed limits less than 50 mph were removed from the dataset. This is considered data dredging (or data snooping) and is generally discouraged. The rationale to justify using this procedure is that those segments are not representative of the bulk of Utah’s rural two-lane highways and their extreme characteristics undermine the predicting capability of the SPFs. Eight segments were removed for having AADT values greater than 10,000 veh/day during at least one year. Four segments were removed for having a speed limit of 40 or 45 mph. Of the original 169 study segments, 157 remained.

The removal of segments from the model dataset limits the range of data that can be used for future predictions. The models are only capable of predicting crashes on segments whose characteristics fit within the range of the original data, which are discussed in section 4.2.

4.2 Raw Data

Figure 4-3 is a map of Utah that highlights the locations of the study segments. A portion of US-40 is expanded to show specific detail of how the study segments appear close-up. Although the selection process for this study was not completed with strictly random sampling techniques, this map shows that the study segments represent a variety of Utah's geography.

Figure 4-4 is a histogram of the segment length for all 157 study segments. Note that the segments are never less than 0.2 mi in length. The HSM recommends only using segments greater than 0.1 mi in order to reduce the work required by the data collection process (AASHTO 2010). Generally, the trend is a decrease in the number of segments as the segment length increases, with some inconsistencies. The mean segment length is 0.97 mi.

Of the 157 segments, 99 have lane widths greater than 12.0 ft and only 58 have shoulders at least 6.0 ft in width. The base conditions of the HSM SPF are 12.0-ft lane widths and 6.0-ft shoulder widths. Figure 4-5 shows a histogram of the segments' lane widths. Figure 4-6 is a histogram of shoulder widths for the study segments. The lane widths are grouped by 0.5-ft increments and the shoulder widths are grouped by 2-ft increments because the CMFs used in the

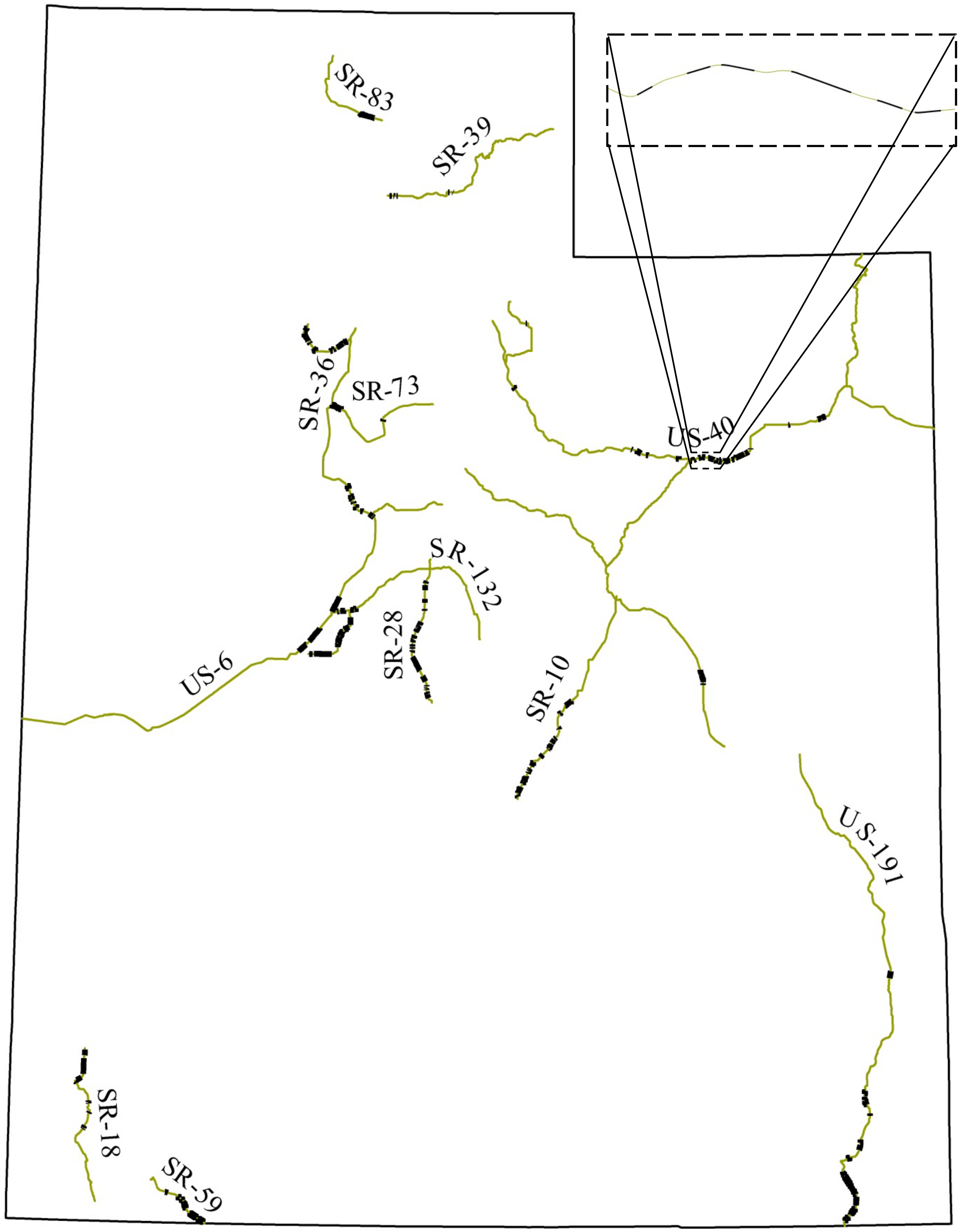

Figure 4-3: Map of selected study segments used for calibration and modeling.

HSM are given by these amounts. The mean lane and shoulder widths are 12.1 ft and 4.7 ft, respectively.

Figure 4-7 is a histogram of AADT values for the study segments. The AADT represented in the histogram is an average of the three AADT values for each segment during the study period. The mean value of AADT across the study segments is 2,787 veh/day. Because the new SPFs are developed based on a crash frequency for three years, the average value of AADT for three years is used.

Figures 4-8 and 4-9 are histograms of the average number of single-unit trucks and combo-unit trucks, respectively, for each segment as a percentage of AADT. The single- and combo-unit truck percentages are also averages for the three years.

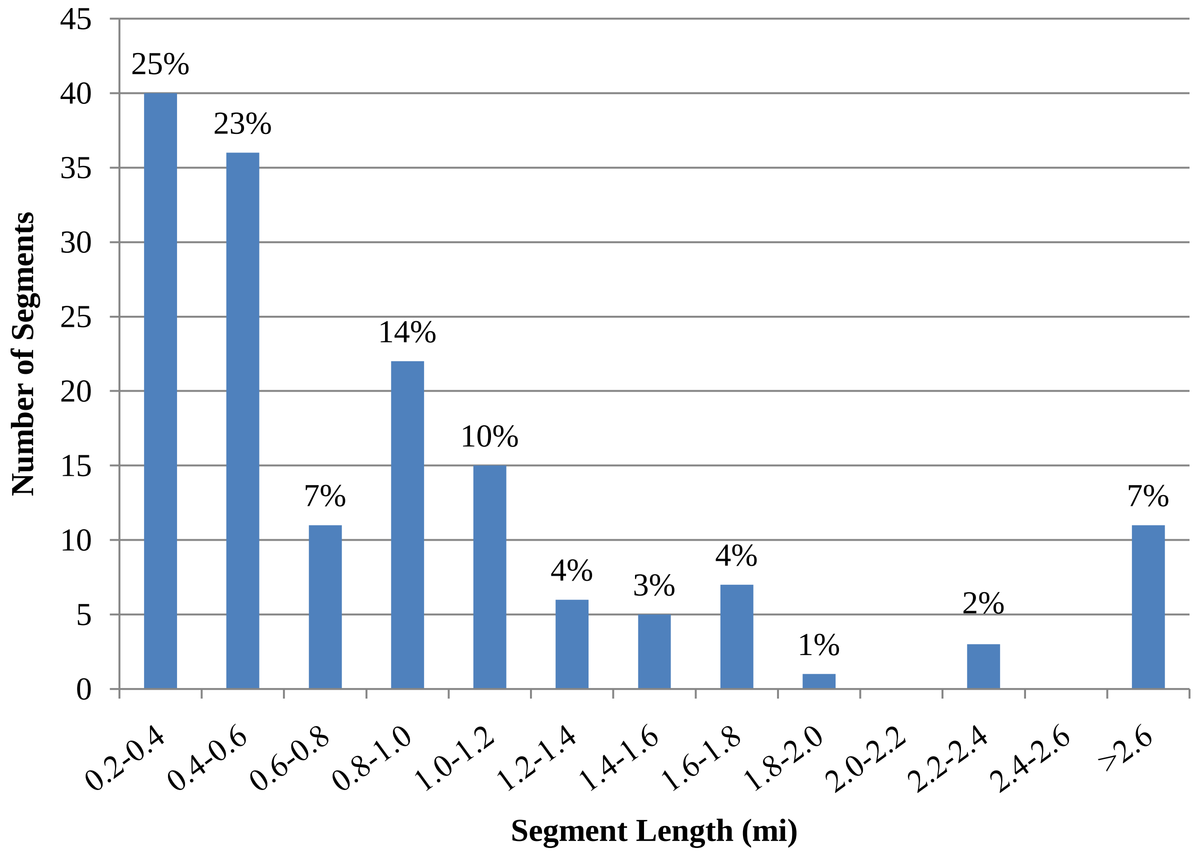

Figure 4-4: Histogram of segment length.

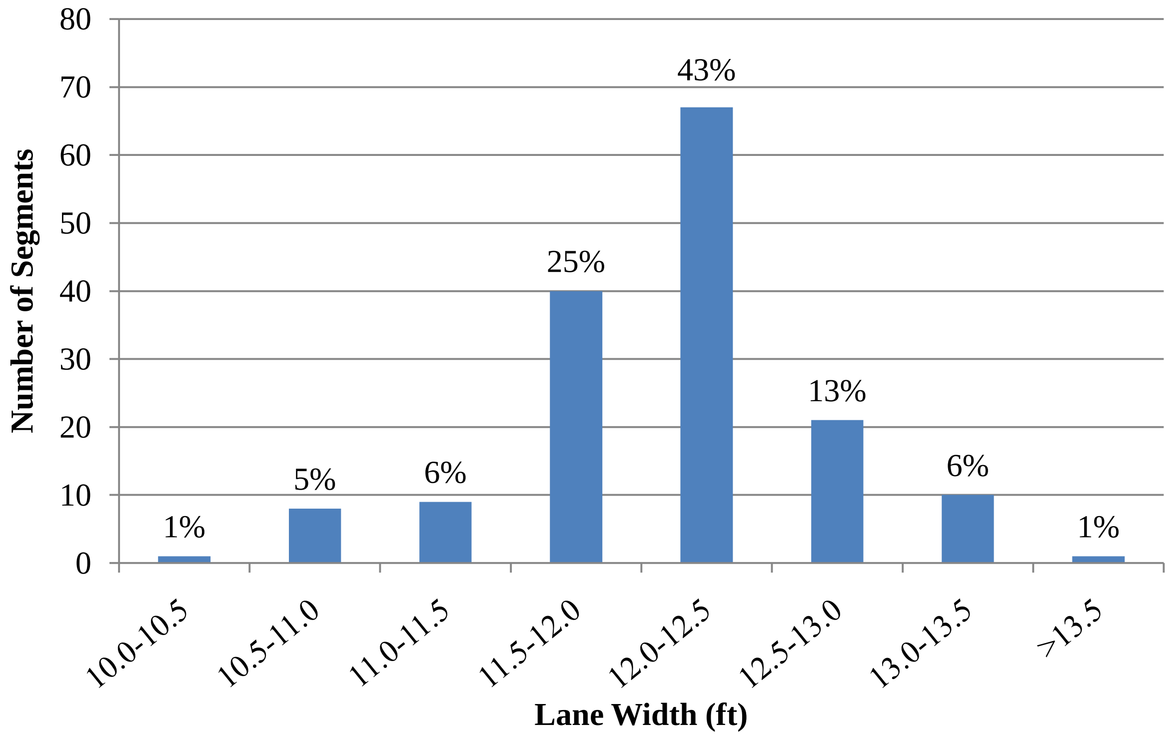

Figure 4-5: Histogram of lane widths for study segments.

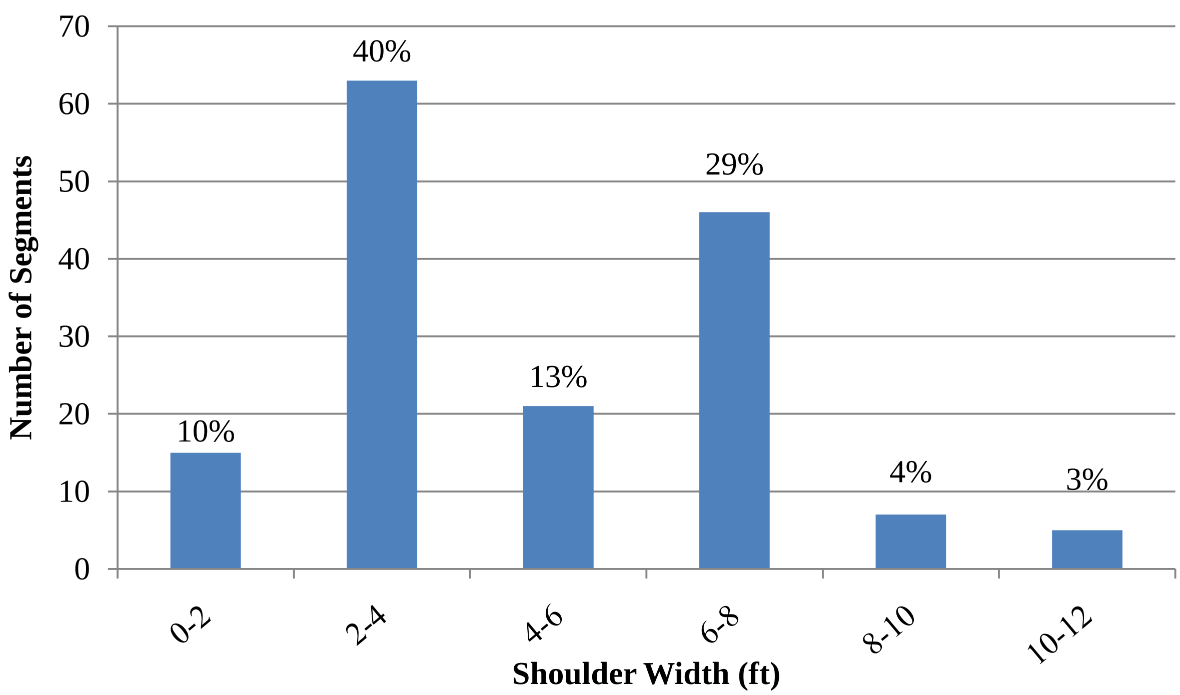

Figure 4-6: Histogram of shoulder widths for study segments.

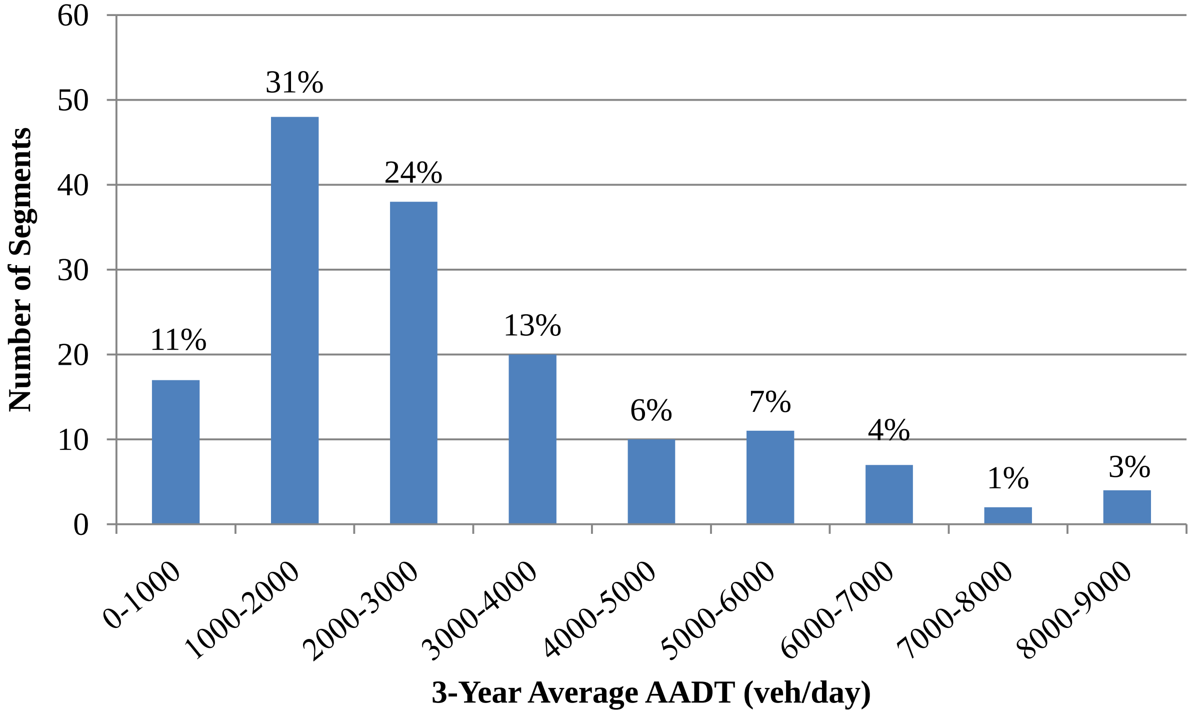

Figure 4-7: Histogram of average value of AADT over three years.

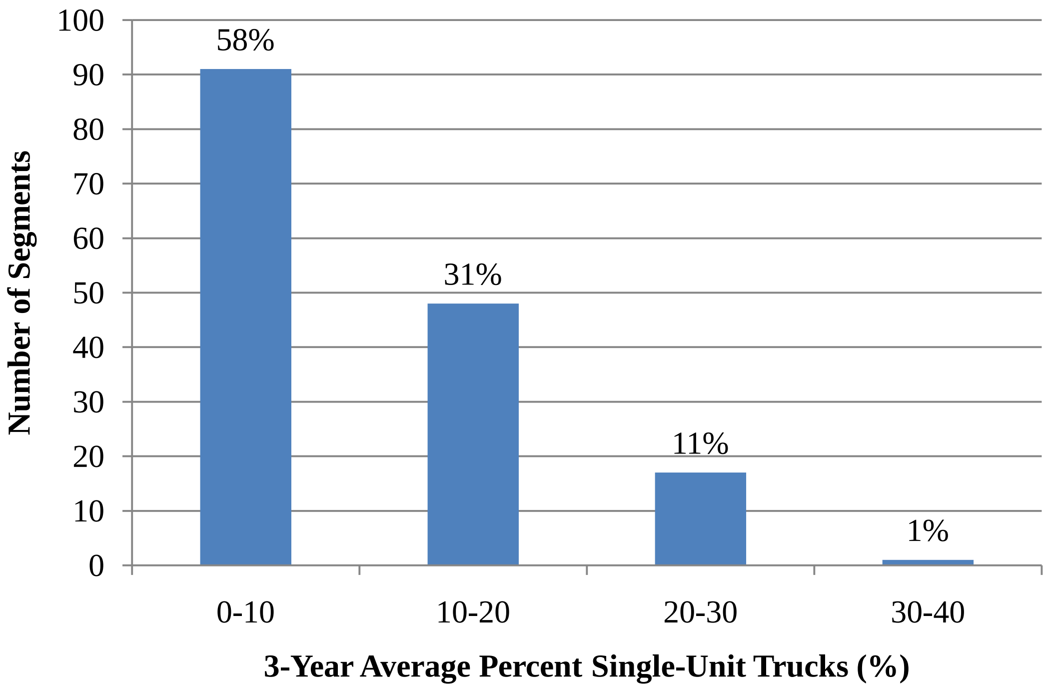

Figure 4-8: Three year average of single-unit trucks as percent of AADT.

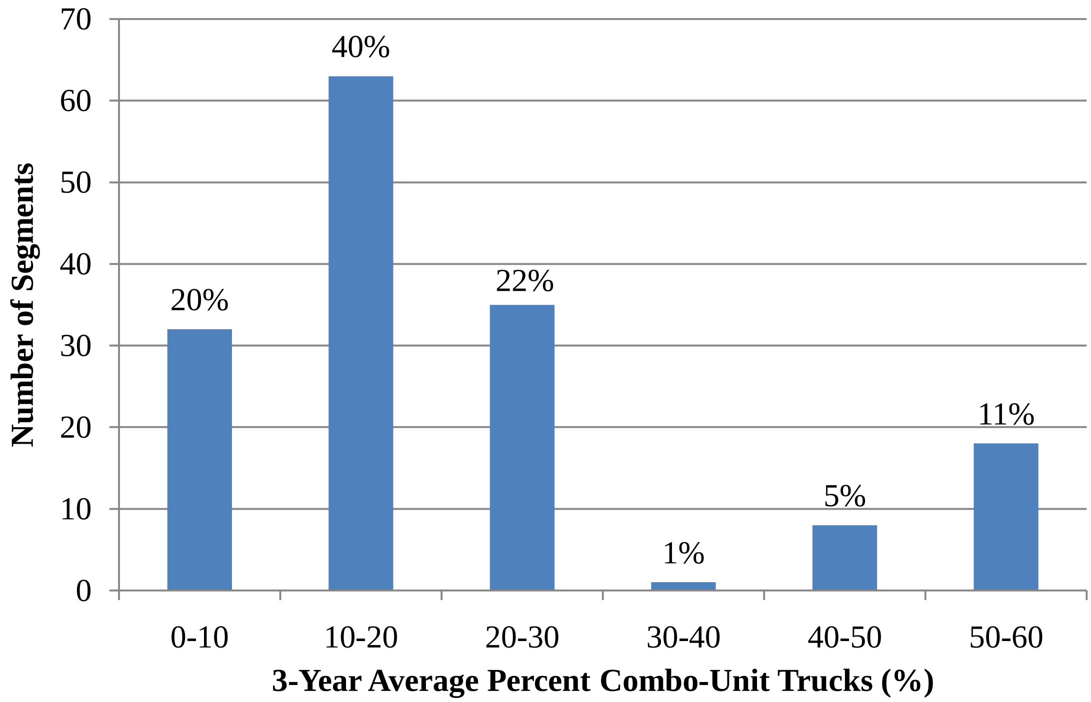

Figure 4-9: Three year average of combo-unit trucks as percent of AADT.

Figure 4-10 is a histogram of the number of crashes on each segment, displayed as individual values for each year. Segments are thus represented as three data points (157 segments × 3 years = 471 total segments). This histogram shows that, generally, it is more common for a study segment to have no crashes in a given year than to experience any crashes. On average, 53 percent of segments in a given year did not experience any crashes (249 of 471).

Figure 4-11 is a histogram of the 3-year combined crash frequency experienced by the 157 segments. Forty-five segments (29 percent) never experienced a single crash during the 2005-2007 study period.

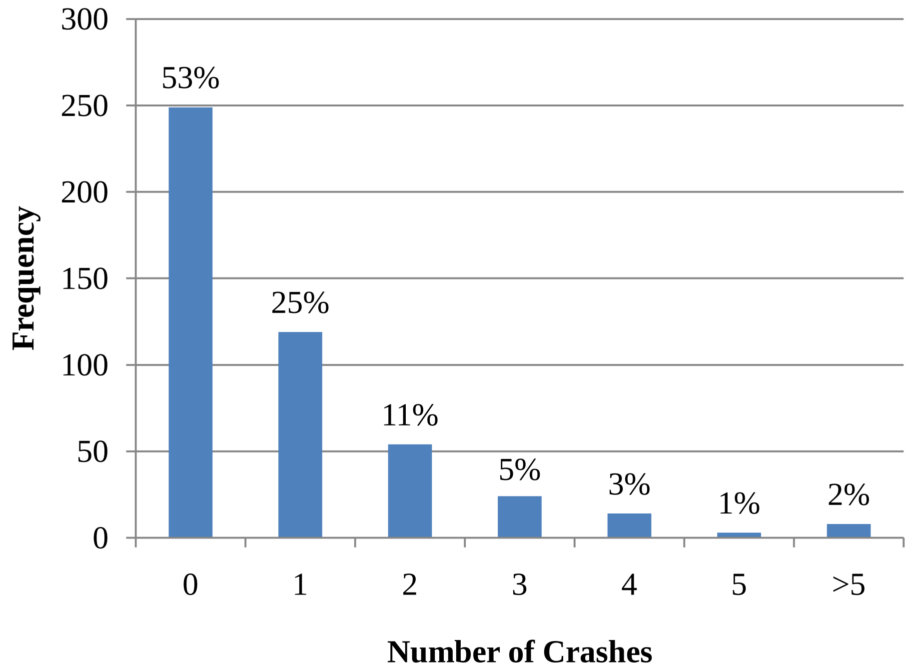

Figure 4-10: Histogram of crash counts on each segment for individual years.

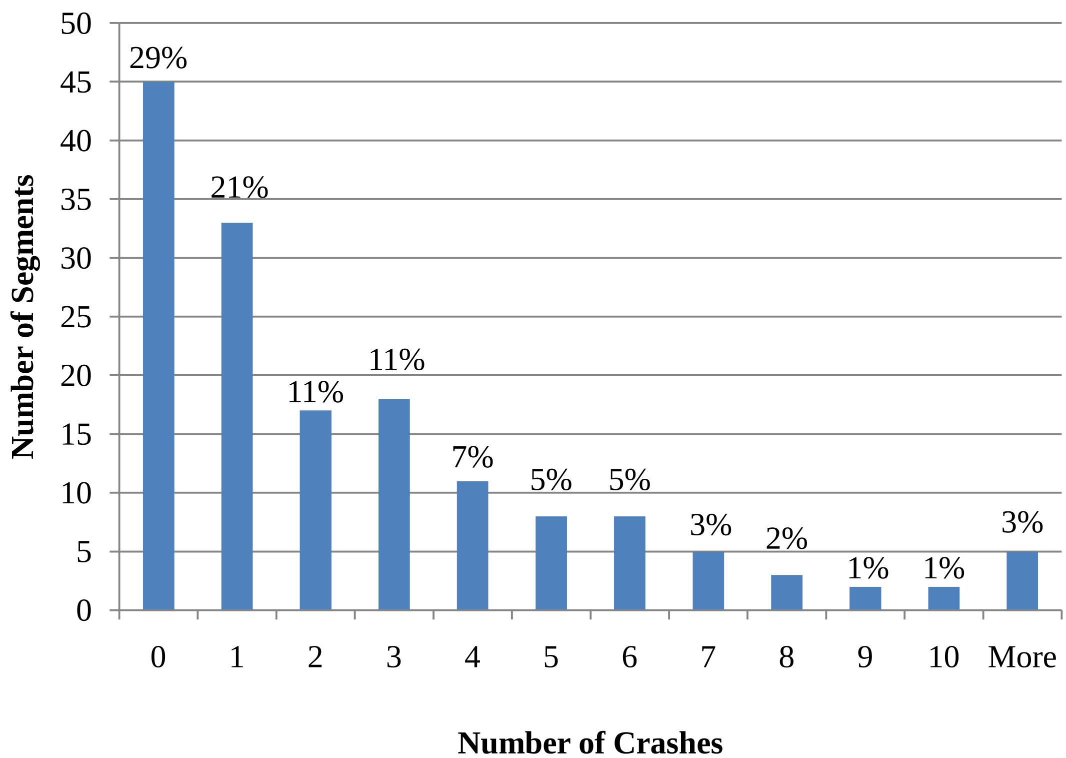

Figure 4-11: Histogram of total crash frequency for each study segment (three years).

4.3 Correlation and Independence of Variables

It is necessary to examine the correlation among variables in order to ensure independence. High correlation indicates that one variable is dependent upon another. When correlation is present, only one of the two correlated variables should be used in the model.

Table 4-2 shows the correlation among the data variables. Some correlation is expected and understandable. For instance, the number of driveways on a segment is correlated with segment length (longer segments will contain more driveways). Naturally, the number of driveways is also correlated with the driveway density, so it is important to ensure that only one of those variables is ever present in a model. There is negative correlation (though not remarkable) between speed limit and driveway density. Residences (and thus driveways) are more common in areas with a 55 mph speed limit than areas with a 65 mph speed limit. The highest correlation is between total truck percentage and combo-truck percentage. Again, none of the models should use both of these variables.

It is important to note that there is a distinction between correlation and causality. Correlation indicates that one variable is dependent upon another, not that it causes something to occur. This study evaluates the correlation between crashes and predictive factors and is not intended to declare causality. For example, there may be a correlation between speed limits and crashes, but that does not indicate that a speed limit causes crashes.

4.4 Data Inconsistencies

As discussed in Section 2.5, model accuracy is dependent upon the accuracy of the encoded crash data and collected facility data. A very thorough review of the crashes used in this study revealed some impossible crash attributes. Three specific examples are given. First, one

Table 4-2: Correlation Values of Investigated Variables

	Segment Length	Grade	Number of Driveways	Driveway Density	Shoulder Rumble Strip	Passing Ability	Speed Limit	Average Lane Width	Average Shoulder Width	AADT	Single Truck Percentage	Combo Truck Percentage	Total Truck Percentage	Total Crashes
Segment Length	1													
Grade	-0.284	1												
Number of Driveways	0.444	-0.183	1											
Driveway Density	-0.105	-0.015	0.638	1										
Shoulder Rumble Strip	-0.052	-0.054	-0.029	0.03	1									
Passing Ability	0.288	-0.123	0.107	-0.028	-0.058	1								
Speed Limit	0.138	0.018	-0.347	-0.555	0.016	-0.029	1							
Average Lane Width	0.032	-0.097	-0.135	-0.156	0.141	0.058	0.234	1						
Average Shoulder Width	0.08	-0.017	-0.047	-0.032	0.333	0.083	0.168	0.280	1					
AADT	-0.03	-0.214	0.132	0.196	0.359	0.155	-0.219	0.239	0.225	1				
Single Truck Percentage	0.135	-0.062	0.21	0.264	0.460	-0.069	-0.101	-0.095	0.273	0.214	1			
Combo Truck Percentage	-0.197	0.13	-0.29	-0.272	-0.016	-0.132	0.281	-0.008	-0.077	-0.172	-0.337	1		
Total Truck Percentage	-0.157	0.112	-0.224	-0.185	0.165	-0.165	0.261	-0.044	0.028	-0.095	0.035	0.929	1	
Total Crashes	0.538	-0.232	0.346	0.05	0.2	0.238	0.095	0.113	0.068	0.466	0.260	-0.249	-0.159	1

Correlation Value	0.40-0.49	0.50-0.59	0.60+

crash is coded as occurring at an off-ramp, but the crash record location is a solitary, open stretch of road. Second, another crash is coded as occurring at a bridge (either overpass or underpass), but the given location has no bridge nearby. Finally, another crash is coded as occurring at a 4-leg intersection, where the real location contains a T-intersection.

A complete investigation of each crash during a 3 year period is unreasonable. Therefore, there should be a reasonable expectation for accuracy from the law enforcement agents and those responsible for data entry to alleviate any concern over the data. Because mistakes are possible by any party involved, one can only hope that they are not large enough to significantly affect model results.

4.5 Limitations

The current data collection process requires intense efforts to find appropriate segments and record their attributes. For future calibration or model development, the same segments can be used, meaning the only necessary future data collection is to record the AADT values and crash data. This is an immense reduction in the required amount of work but is only possible if the segments remain unchanged (i.e., there have been no improvements or realignments to the road since the previous calibration).

Lane and shoulder widths and elevation were recorded using Google Earth satellite imagery. Measurements of the lane and shoulder widths were measured by zooming in on the Google Earth image and measuring the visible cross-section of the roadway. The accuracy depended on the pixilated image on-screen and the quality of the images, which were collected by satellite and airplanes for state GPS surveys. Blurry pixels on the screen could affect the measurement by as much as 0.5 ft. One must also be careful to not over-interpret the value for the longitudinal grade. The grade measurement used in this study was the average rate of elevation change between the two endpoints, and does not provide insight to the many grade variations at specific locations throughout a segment.

There are clearly more variables to road safety than just those gathered in this study. More data can be collected that help identify the features that affect safety and can be used to predict the number of crashes that may occur in a given period of time. The specific data mentioned above and given in the HSM are good starting points for modeling crash rates. Other possible data that can improve the model may include type of terrain, population of the nearest city, distance from the nearest city, quality of the pavement, or average age of registered drivers. Ultimately, the ability of the model to predict crashes is dependent upon the accuracy and amount of the data used in the model, which is determined by those performing the analysis under the constraints of time and resources.

4.6 Modeling Procedure

In addition to the HSM SPF calibration, six on negative binomial models and one hierarchical Bayesian model were investigated. Of the six negative binomial models, only four were fully developed into SPFs because the changes used in the remaining two models failed to produce significant improvements. The four developed negative binomial models were based on two model types performed at two levels of significance. The first model type incorporated the original data. The second used a natural log transformation of AADT. Each type was first developed at a lower level of confidence (75 percent) and then at a higher level of confidence (95 percent). The third model type that was investigated used rounded lane and shoulder widths. This was done to produce a SPF that uses CMFs similar to the HSM, whose CMFs are selected from whole integers for these measurements. Thus, a 10.6-ft lane would have the same effect in the model as an 11.4-ft lane. Because this did not improve upon the ability to predict crashes and did not provide significant parameters at the selected confidence levels, the model was not fully developed and the results are only briefly discussed in Chapter 5.

The negative binomial models were developed with the statistical software *SAS* using a backward stepwise technique (SAS 2011). Backward stepwise regression is performed by simultaneously introducing all potential variables to the model and, one by one, removing the least significant variable until only significant ones remain (based on a predefined confidence level). This study examined models at 75 and 95 percent confidence levels. For a 75 percent confidence level, the P-value of each parameter must be less than 0.25. For a 95 percent confidence level, the P-value must be less than 0.05.

The statistical software package *R* was used for developing the hierarchical Bayesian model (R 2011). Only one model was developed using hierarchical Bayesian techniques. The chosen model used the following variables: segment length, natural log of AADT, speed limit, and combo truck percentage. The variables selected for the hierarchical Bayesian model are intentionally the same as the variables contained in the final negative binomial model, because their significance in a model had already been established. Because the hierarchical Bayesian model uses distributions for parameters and provides a predictive distribution for a crash frequency, rather than single point estimates and predictions, it is difficult to make direct comparisons between the negative binomial and hierarchical Bayesian models. In Chapter 5, the hierarchical Bayesian model is shown to be effective for identifying hot spots by comparing the observed crash frequency of a roadway segment with the predicted distribution.

4.7 Summary

The modeling results are presented and discussed in Chapter 5. Figure 4-12 presents the components of the database that served as input for the SPF calibration and new model development.

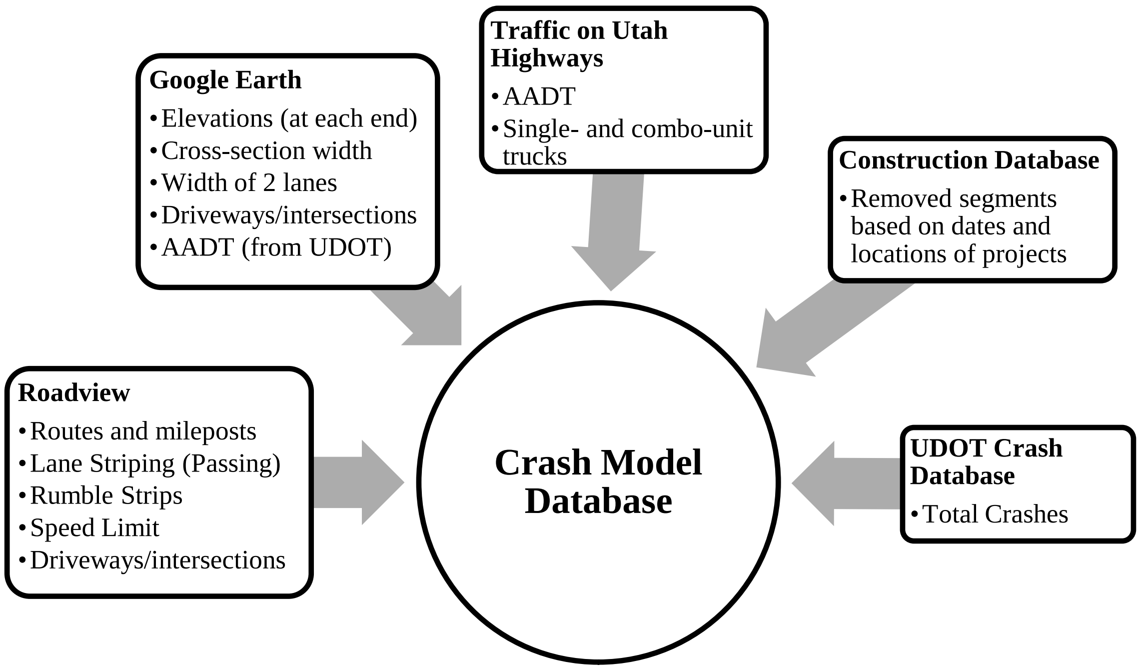

Figure 4-12: Origins of data used for the HSM SPF calibration and new SPFs.

The process of collecting data was completed with the intent of calibrating the HSM SPF and developing new SPFs. More data were collected than required by the HSM to discover which variables significantly affect crashes in a new model. Two modifications to the original data were made to examine any potential benefit for a model. These modifications were a natural log transformation of AADT and a dichotomization of lane and shoulder widths to whole and even numbered values, respectively. However, the dichotomization did not produce significant results. The two other negative binomial models were developed at 75 and 95 percent confidence levels, which resulted in four separate models. Additionally, a hierarchical Bayesian model was developed. This model incorporated the same variables used in the negative binomial model with the natural log of AADT at a 95 percent confidence level.

5 RESULTS

Development of the HSM calibration factor and jurisdiction-specific models was accomplished as discussed in Sections 2.2.1, 2.3, and 4.6. This chapter provides the modeling results. First, the result of the calibration of the HSM SPF for two-lane two-way roads is presented. Next, results of the four negative binomial SPFs and the attempts to develop two additional negative binomial models are discussed. Then, selection of a negative binomial SPF is presented. Finally, results of the hierarchical Bayesian modeling are given, followed by a proposed application of the Bayesian model for determining unsafe segments. A summary of the modeling results is also provided.

5.1 HSM Model Calibration

The HSM SPF for base conditions on rural two-lane two-way roads is reprinted in Equation 5-1 (simplified from Equation 2-1).

$$N_{spf} = AADT \times L \times 2.672 \times 10^{-6} \tag{5-1}$$

There were 426 reported crashes on the 157 segments for 2005-2007. The HSM predicts 368 total crashes for these three years using all applicable CMFs. Using Equation 2-5 (from the calibration methodology presented in the HSM), the calibration factor was found to be 1.16 (1.16

= 426/368). Equation 5-2 gives the Utah-calibrated HSM SPF for two-lane two-way rural highway segments using the calibration factor.

$$N_{local} = 1.16 \times AADT \times L \times 365 \times 10^{-6} \times e^{-0.312} \qquad (5\text{-}2)$$

Equation 5-2 is simplified to Equation 5-3.

$$N_{local} = AADT \times L \times 3.09 \times 10^{-6} \qquad (5\text{-}3)$$

As Equation 5-3 is the calibrated SPF from the HSM, CMFs should still be applied as directed by the HSM (AASHTO 2010).

5.2 Negative Binomial Model Form

The form of the negative binomial model used for developing jurisdiction-specific SPFs is shown in Equation 5-4.

$$\ln(N) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \qquad (5\text{-}4)$$

where, N = number of crashes (predicted or observed),

β_0 = intercept,

β_i = coefficient for variable x_i,

x_i = independent variable, and

n = number of independent variables.

A rearrangement of Equation 5-4 can directly predict the number of crashes for a given year as illustrated in Equation 5-5.

$$N = e^{\beta_0 + \sum_{i=1}^{n} \beta_i x_i} \tag{5-5}$$

Equation 5-5 may be expressed as outlined in Equation 5-6. The SPFs that follow in Section 5.3 are generally written with this form.

$$N = \exp\left[\beta_0 + \sum_{i=1}^{n} \beta_i x_i\right] \tag{5-6}$$

In Equation 5-6, each coefficient is multiplied with its respective variable and added together, as shown in the brackets. This sum is then exponentiated to determine the number of predicted crashes on a road segment. With this model form, coefficients less than zero show a reducing effect on crash frequencies, while coefficients greater than zero show an increasing effect on crash frequencies.

5.3 Utah-Specific Negative Binomial Models

This section presents the negative binomial models that were investigated as discussed in Chapter 4.6. First, two conventional models (with no data changes) are presented at 75 and 95 percent confidence levels. Next, two models are given that use the natural log of AADT at 75 and 95 percent confidence levels. Finally, an attempt at modeling crashes using dichotomized values of lane and shoulder widths is reviewed. This SPF was not fully developed because the parameters of interest were not significant enough to warrant inclusion.

Bayesian information criterion (BIC) is used for selecting a new SPF. When using BIC for comparing and selecting models, the model with the smallest BIC value should be selected. Note: there is no relationship between BIC and the hierarchical Bayesian methodology. The BIC value is determined by Equation 5-7 (Ramsy and Schafer 2002).

$$BIC = n \times \ln(RSS) + p \times \ln(n) \tag{5-7}$$

where,	BIC	=	Bayesian information criterion,
	n	=	number of observations,
	RSS	=	sum of squared residuals, and
	p	=	number of independent variables.

According to Equation 5-7, the value of BIC increases with higher squared residuals (RSS) and number of independent variables (p). Thus, BIC emphasizes the importance of parsimony, which addresses the concern that complex models may overfit the data and have poor predictive capability.

5.3.1 Conventional Negative Binomial Models

The conventional models were formed with no data modifications. Table 5-1 shows the estimates and P-values from the first conventional model at a 75 percent confidence level (P-value < 0.25). In this model, the least significant variables are Driveway Density (P-value = 0.2139) and No Shoulder Rumble Strips (P-value = 0.1120).

Table 5-1: Conventional Model at a 75 Percent Confidence Level

Parameter		Estimate	P-Value
Intercept		-7.488	0.0019
AADT		0.0002	<0.0001
Segment Length		0.429	<0.0001
Driveway Density		0.0286	0.2139
Passing	Prohibited	-1.596	0.0312
	1-Lane	-0.128	0.488
No Shoulder Rumble Strip		-0.268	0.1120
Combo Truck Percentage		-0.0219	0.0003
Speed Limit		0.1036	0.0009

The written model is shown in Equation 5-8.

$$
\begin{aligned}
N = \exp[&-7.488 + (0.0002)(AADT) + (0.429)(L) + (0.0286)(DD) \\
&- 1.596(\mathit{No\ Passing}) - 0.128(\mathit{1\text{-}Lane}) - (0.268)(\mathit{No\ SRS}) \\
&- (0.0219)(CT) + (0.1036)(\mathit{Speed})]
\end{aligned}
\quad (5\text{-}8)
$$

where, *DD* = driveway density (driveways/mi),

No Passing = 1 if passing is prohibited, 0 if permitted for one or two lanes,

1-Lane = 1 if passing is permitted for one lane, 0 if otherwise,

No SRS = lack of shoulder rumble strip (1 if not present, 0 if present),

CT = percentage of combo-unit trucks (%), and

Speed = speed limit (mph).

The overdispersion parameter for this model is 1.20. As discussed in Section 2.1.3, a high overdispersion parameter is indicative of great variation in the data and results in less weight on the SPF prediction when using the EB Method. The BIC value is 607.4.

In the first conventional model (Equation 5-8), two-lane passing ability and the presence of a shoulder rumble strip are automatically incorporated into the intercept of the model. If passing is prohibited or allowed for one lane, the variables *No Passing* or *1-Lane* should be 1.0, respectively. When passing is allowed for two lanes, both variables should be 0. The negative coefficients for *No Passing* or *1-Lane* indicate that crashes likely decrease on segments where only one lane is allowed to pass or passing is completely restricted. When there is no shoulder rumble strip, *No SRS* has a value of 1.0. The negative coefficient for *No SRS* indicates that fewer crashes are predicted on segments without a shoulder rumble strip.

Even though the P-value for 1-lane passing (0.488) is higher than the confidence level allows (0.25), this variable is still used in the SPF because the P-value for *No Passing* is significant at the 75 percent confidence level (0.0312). It is interesting to note that this model shows combo-unit trucks contributing to a decrease in crashes.

Driveway density and shoulder rumble strip are the only parameters that needed to be removed for the model to reach 95 percent confidence (P-value<0.05). Table 5-2 shows the estimates and P-values from the model with only parameters significant at the 95 percent confidence level.

The written SPF for this model is given in Equation 5-9.

$$N = \exp[-7.116 + (0.0003)(AADT) + (0.423)(L) - 1.506(\textit{No Passing}) - 0.0812(\textit{1-Lane}) - (0.0219)(CT) + (0.0938)(\textit{Speed})] \quad (5\text{-}9)$$

Table 5-2: Conventional Model at a 95 Percent Confidence Level

Parameter		Estimate	P-Value
Intercept		-7.116	0.0017
AADT		0.0003	<0.0001
Segment Length		0.423	<0.0001
Passing	Prohibited	-1.506	0.0427
	1 Lane	-0.0812	0.6603
Combo Truck Percentage		-0.0219	0.0003
Speed Limit		0.0938	0.0017

The overdispersion parameter for Equation 5-9 is 1.24. The BIC value is 601.5. Similar to Equation 5-8, the parameter for 1-lane passing is not significant but still included in the model because prohibited passing is significant. Again, the coefficients for *No Passing* and *1-Lane* indicate a decrease in crash frequency for segments where passing ability is restricted. Combo-unit trucks are linked with a decrease in crash frequencies, while higher speed limits are associated with higher crash frequencies.

5.3.2 Log Transformation of AADT

It was hypothesized that AADT has a different effect on crash frequencies than that shown by the form of the previous two SPFs, where AADT is part of the exponent. To experiment with a different model form that may also reveal different relationships between the variables and crash frequencies, the natural log of AADT was used as a variable in place of AADT. Additionally, the log-transformation reduced the skew in the AADT data. Figure 5-1 shows the original distribution of the average AADT value for three years (reprinted from Figure 4-7). Figure 5-2 is a histogram of the natural log of the AADT averages.

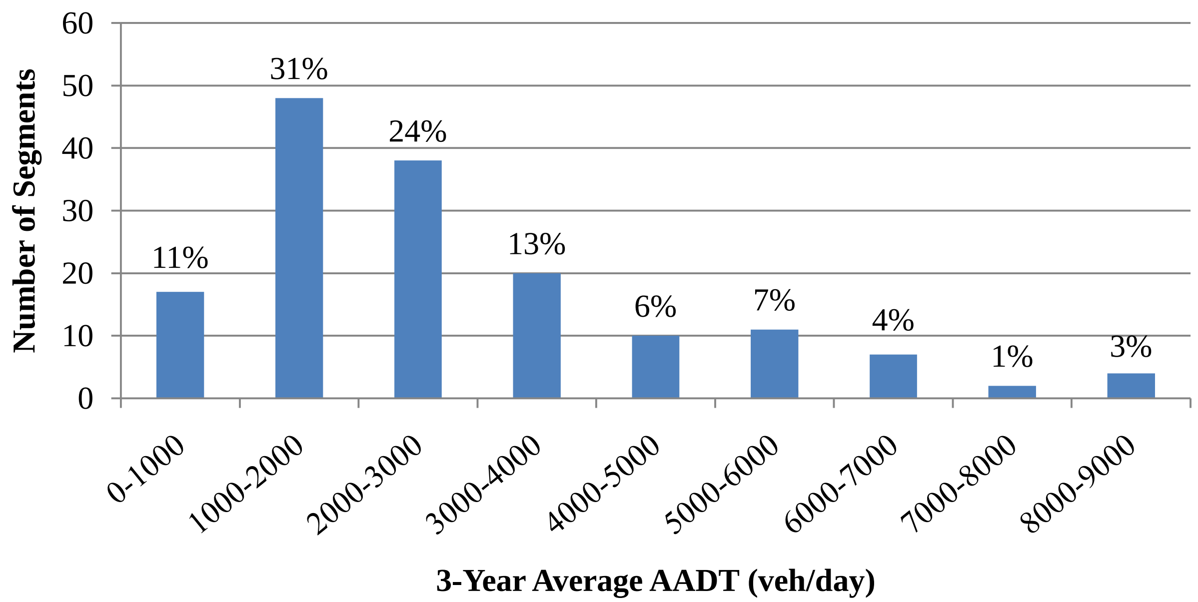

Figure 5-1: Histogram of average value of AADT over three years.

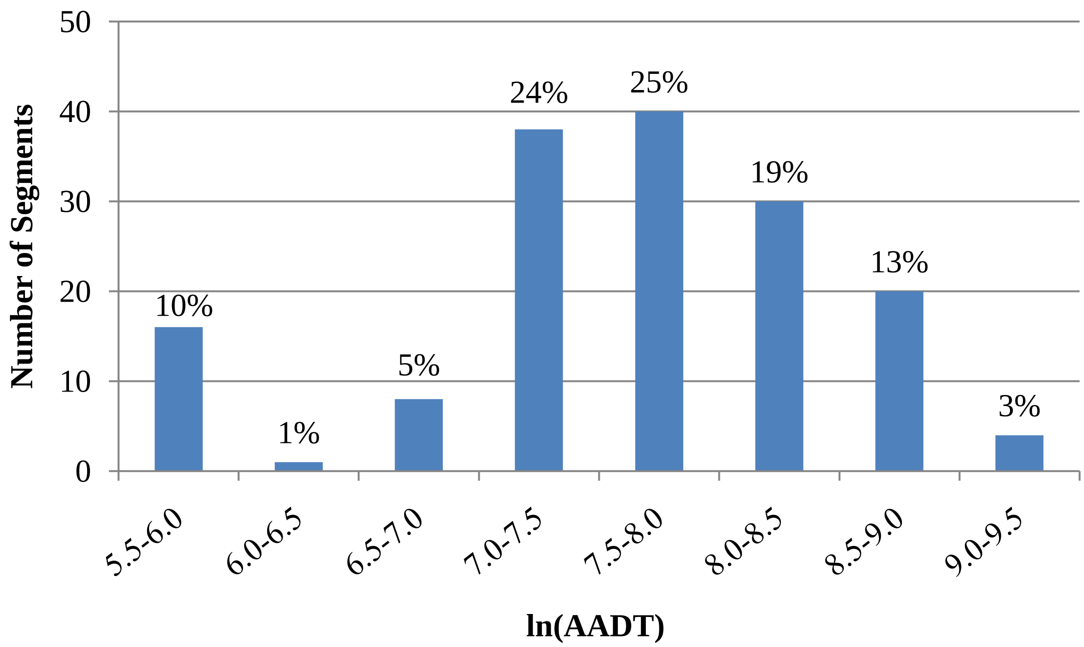

Figure 5-2: Histogram of the natural log of AADT.

Table 5.3 shows the parameters, estimates, and P-values of the model using the natural log of AADT at the 75 percent confidence level. The written SPF from this model is given in Equation 5-10.

Table 5-3: Model Using ln(AADT) at a 75 Percent Confidence Level

Parameter		Estimate	P-Value
Intercept		-12.11	<0.0001
Segment Length		0.442	<0.0001
ln(AADT)		0.753	<0.0001
Average Shoulder Width		-0.0498	0.0489
Passing	Prohibited	-1.222	0.0985
	1 Lane	-0.116	0.510
Driveway Density		0.0277	0.184
No Shoulder Rumble Strip		-0.346	0.0281
Combo Truck Percentage		-0.0257	<0.0001
Speed Limit		0.101	0.0006

$$N = AADT^{0.753} \times \exp[-12.11 + (0.442)(L) + (-0.0498)(SW) - 1.222(\mathit{No\ Passing}) - 0.116(\mathit{1\text{-}Lane}) + (0.0277)(DD) - 0.346(\mathit{No\ SRS}) - (0.0257)(CT) + (0.101)(\mathit{Speed})] \quad (5\text{-}10)$$

where, SW = shoulder width (ft).

The overdispersion parameter is 1.14. The BIC value is 596.7. This model generally shows no striking differences from the previous two conventional models, except that shoulder width is now a significant parameter. The coefficient for shoulder width shows an inverse relationship to crashes, which indicates that fewer crashes could be expected with wider shoulders. Like the previous models, passing restrictions, combo-unit trucks, and the absence of a shoulder rumble strip are all associated with a decrease in crashes.

The SPF based on the model using the natural log of AADT with a 95 percent confidence level is given in Table 5-4 and Equation 5-11.

Table 5-4: Model Using ln(AADT) at a 95 Percent Confidence Level

Parameter	Estimate	P-Value
Intercept	-12.06	<0.0001
Segment Length	0.450	<0.0001
ln(AADT)	0.840	<0.0001
Combo Truck Percentage	-0.0271	<0.0001
Speed Limit	0.0824	0.0040

$$N = AADT^{0.840} + \exp[-12.06 + (0.450)(L) - (0.0271)(CT) + (0.0824)(Speed)] \tag{5-11}$$

The overdispersion parameter for this model is 1.19. The BIC value is 583.7. This model has very few parameters, compared to the conventional models, but each variable is very significant. The BIC value for this model is quite smaller than those of the previous models, likely because this model contains fewer parameters than the other models. As seen previously, combo-unit trucks are associated with a decrease in crash frequencies and the speed limit is associated with an increase in crash frequencies.

The variables used in this fourth SPF contain data that UDOT regularly maintains. AADT values and the percentage of combo-unit trucks are collected yearly and are available Online (UDOT 2011b), and speed limits on road segments can be extracted from roadway shapefiles available through the Utah GIS Portal (Utah AGRC 2011).

5.3.3 Dichotomized Values of Lane Width and Shoulder Width

The HSM provides CMFs for lane and shoulder widths based on a whole integer for the measured variable. There are CMFs for lane widths of 9, 10, 11, and 12 ft and shoulder widths of 0, 2, 4, 6, and 8 ft. To produce an SPF that has parameters used like the CMFs in the HSM, lane and shoulder widths were dichotomized and grouped similar to those in the HSM by rounding measured lane widths to the nearest whole number (between 10 and 13 ft) and shoulder widths to the nearest even number (between 0 and 8 ft). These attempts proved to not produce statistically-reliable results. Not only were most of the grouped parameters not significant in the model, but there was no logical progression of coefficients from the lower to upper bounds of the variables (e.g., the coefficient for a 4-ft shoulder may have been negative, while the coefficients for the 2- and 6-ft shoulders may have been positive). It was determined that there was no need to further pursue such models or report specific results.

5.4 Model Selection

A comparison of the BIC values of the new negative binomial SPFs is the basis for recommending that UDOT implement a particular SPF for future use. Table 5-5 presents the new models with their respective BIC values. Of the new SPFs shown in Section 5.3, it is recommended that the fourth SPF (from Equation 5-11) be used because it has the smallest BIC value.

Table 5-5: Model Selection Using BIC

SPF	BIC
Conventional, 75% Confidence Level (Equation 5-8)	607.4
Conventional, 95% Confidence Level (Equation 5-9)	601.5
ln(AADT), 75% Confidence Level (Equation 5-10)	596.7
ln(AADT), 95% Confidence Level (Equation 5-11)	583.7

5.5 Hierarchical Bayesian Model

The hierarchical Bayesian model was developed using the variables that were significant in the fourth SPF (the model using the natural log of AADT with a 95 percent confidence level). The variables selected for the Bayesian model were deliberately chosen from those used in Equation 5-11 because their significance in a negative binomial framework had been established and Equation 5-11 is the least complex of the four developed models. The objective of developing this model is to demonstrate its effectiveness in accounting for uncertainty in modeling crash predictions and determining the specific locations with the greatest safety concerns.

The posterior distributions of the model parameters are shown in Figure 5-3. As the hierarchical Bayesian technique does not produce an exact estimate of a parameter, but rather a distribution, the significance of a parameter can be obtained by observing its posterior prediction density function. A parameter whose distribution is centered at 0.0 may be considered insignificant.

Figure 5-3 shows the density functions of the posterior distributions of the four variables (lane width, AADT, combo-unit truck percentage, and speed limit) and the intercept. Because the posterior distribution is a density function, the area under each curve is equal to 1.0.

5.6 Determining Unsafe Sites Using the Hierarchical Bayes Model

The posterior distribution of the expected crash frequency is the foundation for determining dangerous roadway segments, where the observed crash frequency exceeds the expected number of crashes, like a hot spot. The likelihood of an observed crash frequency occurring on a segment can be ascertained by comparing the observed crash frequency with the predictive density function, determined by the roadway characteristics and posterior distributions of the parameters. For a hot spot, the observed crash frequency will be far to the right of the density function.

Figures 5–4 through 5–8 show the distributions of expected crash frequencies for the five segments with the least likely observed crash frequencies above the predicted distribution. The single vertical line in each figure represents the actual crash frequency for the 3-year period from 2005-2007. A location considered dangerous, such as these five, will have an observed crash frequency to the right of the distribution. For example, the observed crash frequency in Figure 5-4 is eight crashes. This is far outside the predicted distribution that shows that a much smaller crash frequency should have occurred.

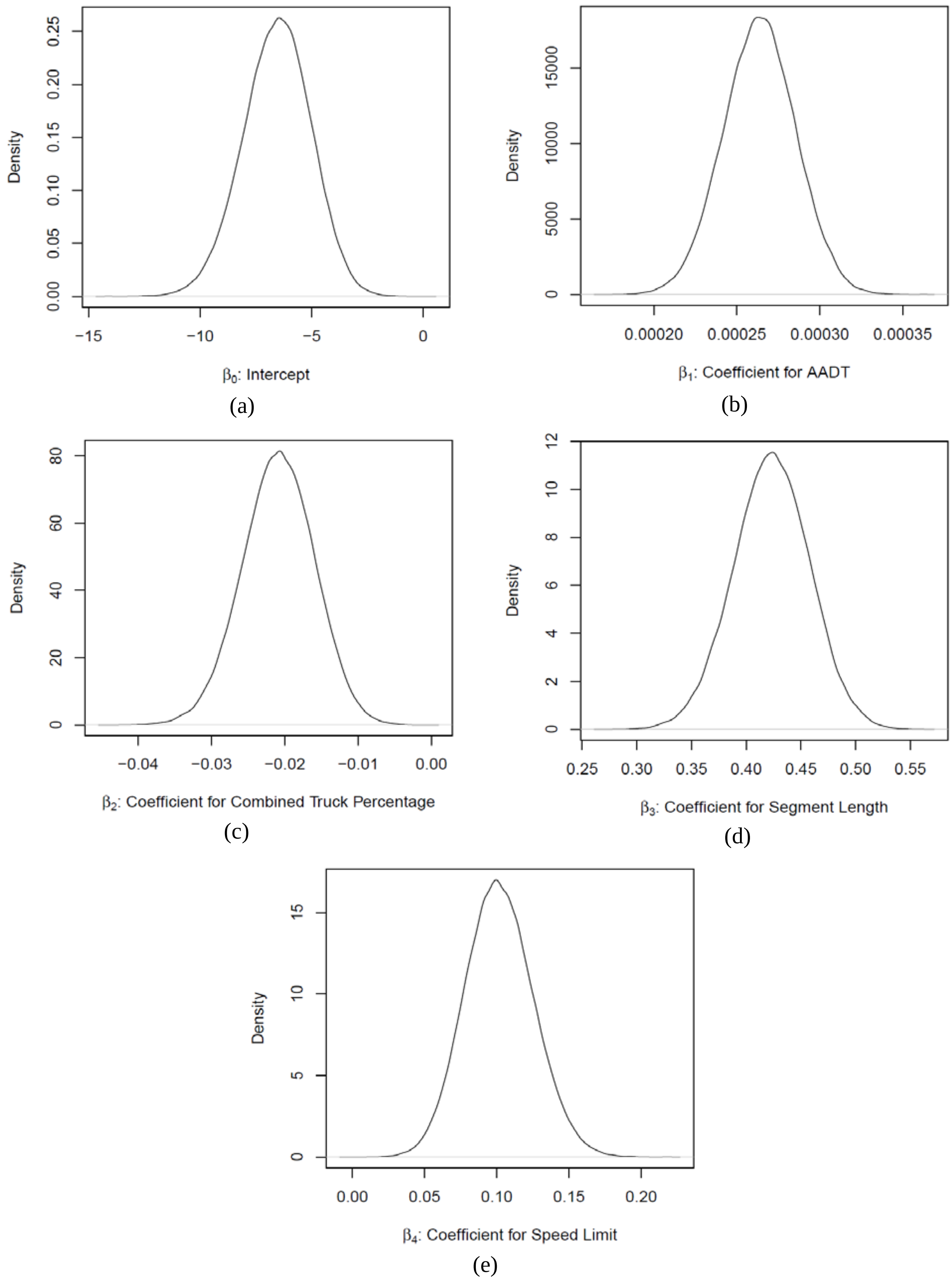

Figure 5-3: Posterior distributions for the hierarchical Bayesian model coefficients.

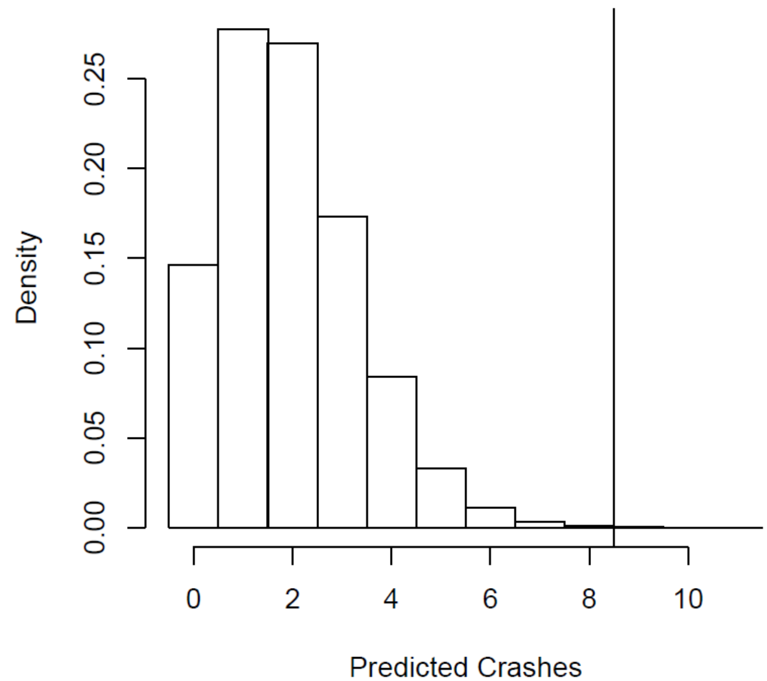

Figure 5-4: Comparison of the observed crash frequency and predicted distribution on segment 133 (Route 132 between mileposts 5.5 and 6.6).

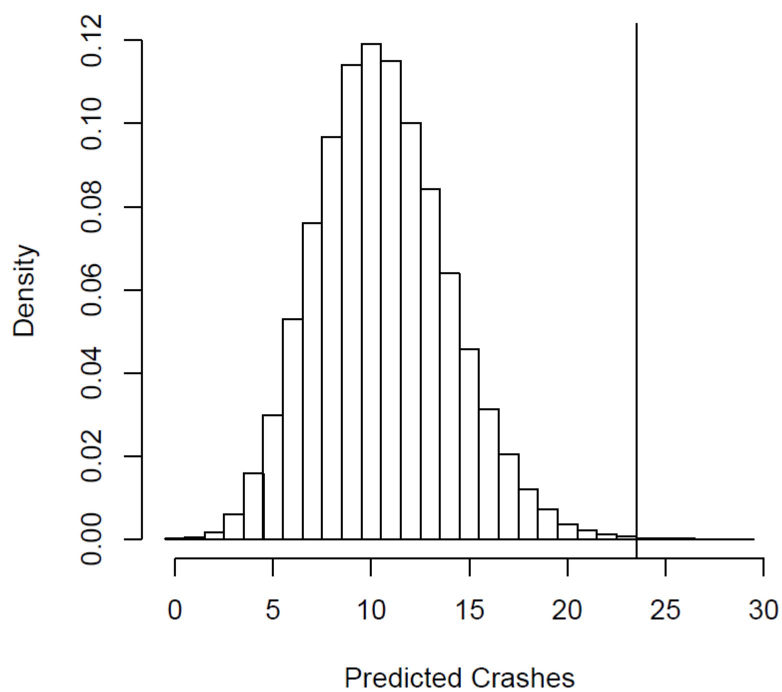

Figure 5-5: Comparison of the observed crash frequency and predicted distribution on segment 102 (Route 40 between mileposts 100.8 and 103.4).

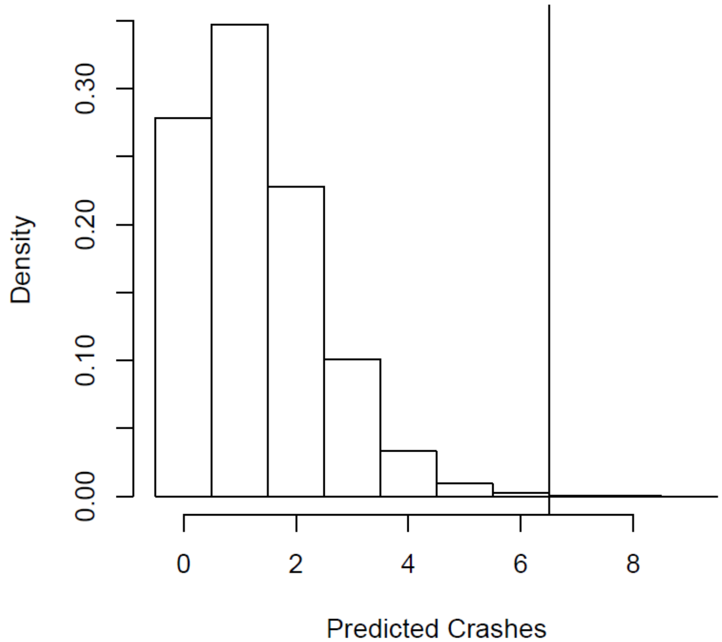

Figure 5-6: Comparison of the observed crash frequency and predicted distribution on segment 30 (Route 10 between mileposts 18.1 and 19.2).

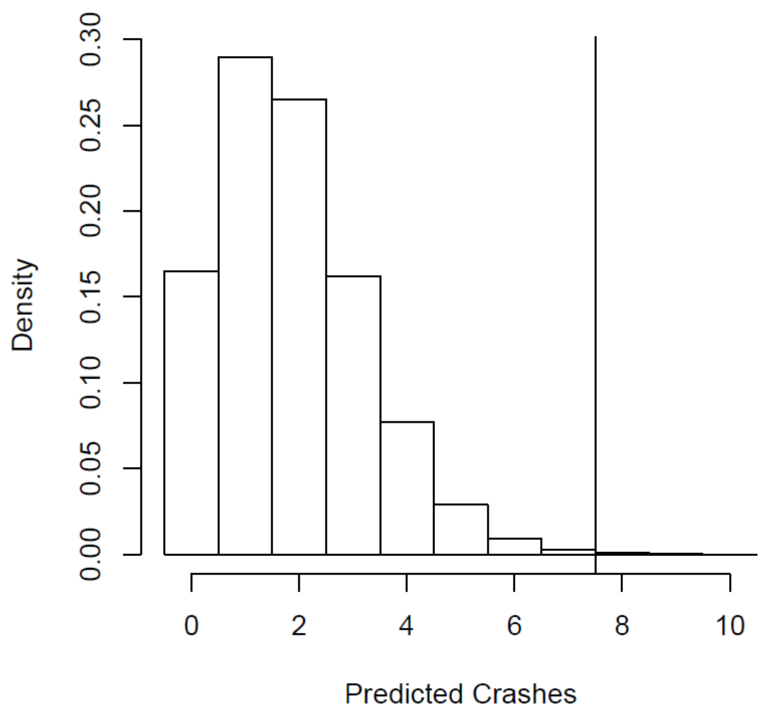

Figure 5-7: Comparison of the observed crash frequency and predicted distribution on segment 89 (Route 40 between mileposts 21.0 and 21.6).

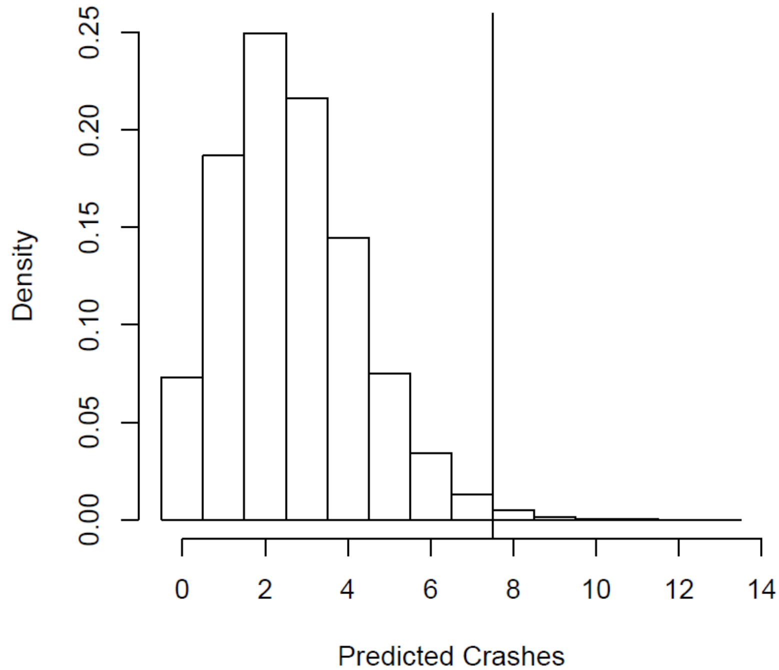

Figure 5-8: Comparison of the observed crash frequency and predicted distribution on segment 35 (Route 10 between mileposts 31.8 and 32.2).

5.7 Summary

The calibration factor of the HSM SPF for rural two-lane two-way roads in Utah was found to be 1.16. This indicates that the HSM underpredicts the number of crashes occurring on Utah's roads by 16 percent. Then, jurisdiction-specific SPFs were developed as an attempt to find a better prediction model to recommend for use by UDOT.

Not one SPF developed in this study includes every variable examined by the data collection process. Each variable makes a unique contribution to the SPF in which it is significant. It is worthwhile to compare the contributions of the variables in each model to find general consistencies (or inconsistencies) among the factors that contribute to a crash prediction.

The following are general observations about the variables used in this study.

- Exposure (both AADT and segment length) was always significantly associated with higher crash frequencies.

- The longitudinal grade was never a significant variable in any model. This may be a result of the majority of segments being relatively flat. Although there were some non-flat segments in the database, the study methodology involved finding straight, homogeneous segments, which generally excluded many areas with a noticeable grade.
- Driveway density was associated with a higher number of crashes. Even though this study excluded crashes at intersections (coded as "junctions" in the crash database), the models show that the presence of driveways has a positive correlation with crash frequencies.
- The absence of shoulder rumble strips consistently had a negative correlation with crash frequencies. One hypothesis may explain this relationship: rumble strips are already placed at locations with higher crash frequencies and have previously been labeled as "unsafe." Thus, when no shoulder rumble strips are present, fewer crashes are predicted. Another hypothesis is that shoulder rumble strips, which are used to reduce ROR crashes, may indeed increase crash frequencies because drivers make overcorrecting maneuvers after inadvertently driving on a rumble strip. The benefit of rumble strips is that they reduce severe and fatal crashes, but other, less severe, crash types may increase.
- Speed limit was significant in every model and always had a positive coefficient, showing an increasing relationship with crashes.
- Lane width was never a significant factor in any model. As discussed in Section 3.1.1, there is a disagreement in the literature regarding the effect of lane width on crashes. The HSM, however, claims that wider lanes result in fewer crashes (AASHTO 2010).
- Shoulder width rarely had a significant effect on crash frequency. The model in which shoulder width was significant indicated a decreasing effect on crashes. This is in

harmony with the HSM research that claims wider shoulders are related to lower crash frequencies (AASHTO 2010).

- Segments with limited or restricted passing opportunities (one-lane passing or prohibited) had a decrease in the number of crashes compared to segments where passing is allowed for two lanes (based on the model coefficients). It is possible that drivers are more aggressive on segments where passing is permitted.
- The percentage of single-unit trucks was never significant in any model.
- The percentage of combo-unit trucks was a significant variable in every model, with a decreasing effect on crashes. One hypothesis for this relationship is that professionally trained drivers operate combo-unit trucks, and tend to be less aggressive and have more consistent and defensive driving habits than other drivers. When a large portion of AADT is comprised of these vehicles, there is a noticeable reduction in dangerous situations that may be more-often caused by other vehicles.

The hierarchical Bayesian model is particularly useful for determining unsafe segments. Its strength is in its ability to produce the distribution of a variable's coefficient, rather than just a point estimate, and thus predict a crash frequency with the density function. Unsafe locations be determined by comparing the observed crash frequency with the likelihood of that frequency occurring, given the established parameters. This may especially be applicable in determining the top five percent of unsafe roadways.

6 CONCLUSION

The purpose of this study was to calibrate the HSM SPF for rural two-lane two-way roads to represent conditions in Utah and develop jurisdiction-specific SPFs that may be used in place of the HSM SPF. This thesis presents the procedure and results of the calibration and new model development. The Utah-specific SPFs were developed from the same dataset used for calibration of the HSM SPF with some additional variables that were hypothesized to affect crash frequencies or at least have predicting ability. The Utah-specific models were developed through negative binomial regression and can be used with the EB method. Additionally, a hierarchical Bayesian model was produced to show its effectiveness in evaluating crash frequencies.

The results of this study show that reasonable crash predictions can be made using the simpler models that require less data. Also, some of the variables that the HSM claims make important contributions to the safety of a roadway were not found to be significant in most of the jurisdiction-specific models. The following sections discuss in detail the conclusions of this study, recommendations for UDOT, and needs for further research.

6.1 Conclusions

Calibration of the HSM SPF for two-lane two-way road segments was performed using a dataset of 157 roadway segments in Utah. The calibration factor obtained (1.16) shows that there are more crashes occurring on Utah’s rural two-lane two-way roads than are predicted by the

HSM. Each of the new negative binomial models uses a unique set of variables that predict crash frequencies differently. The selection of a particular model for predicting crashes should be based on its predicting capability and the availability of data and resources.

Because there are 12 conditions for which CMFs can be applied to the HSM SPF for rural two-lane two-way roads (Table 2-1), the new jurisdiction-specific SPFs require fewer data variables than the HSM model. These simpler models that require less data may be adequate for predicting crashes. As the HSM emphasizes the use of the EB method to determine an expected crash frequency, an overdispersion parameter is provided with each of the new SPFs for use with the EB method.

The application of hierarchical Bayesian modeling is likely to become more widespread as computing capabilities increase and its effectiveness in safety evaluations continues to be proven. It is worthwhile to explore the use of this modeling technique in the future. In this study, the predictive distribution of the expected crash frequency developed from the posterior distributions of the model parameters was used to determine the road segments that had excessively high crash frequencies. This approach has promising application for identifying hot spots that could benefit from roadway improvements.

6.2 Recommendations to UDOT

The results of this study indicate that the relationships between crashes and roadway characteristics in Utah are different than those presented in the HSM. Selecting a specific model to express these relationships is dependent upon data availability and model accuracy. Regarding crash modeling and safety evaluations in Utah, the following recommendations are given to UDOT:

- Equation 5-11 should be used for predicting crashes on two-lane two-way rural roadway segments in Utah. This SPF is repeated in Equation 6-1 and should be used to avoid the intense process of collecting geometric data required by the HSM.

$$N = AADT^{0.840} + \exp[-12.06 + (0.450)(L) - (0.0271)(CT) + (0.0824)(Speed)] \quad (6\text{-}1)$$

where,

$AADT$ = average annual daily traffic,

L = segment length (mi),

CT = percentage of combo-unit trucks (%), and

$Speed$ = speed limit (mph).

- The calibration factor of the HSM SPF for two-lane two-way rural roads is 1.16. This SPF should only be used to predict crashes if UDOT elects to use the HSM SPF. The CMFs given in the HSM must be used with this model.
- SPFs for other classifications of roads (e.g., multilane highways, rural interstates, or urban arterials) should be developed for both road segments and intersections to initiate a comprehensive program for evaluating safety in Utah. This may include calibrating the SPFs in the HSM.
- As radii for horizontal curves and specific elevation data become available, SPFs should be redeveloped for rural two-lane two-way roads that consider horizontal and vertical curvature.

- Hierarchical Bayesian techniques should be used to identify locations on the state highway network that require safety-related attention. This may be in conjunction with the annua*l 5 Percent Report* required by federal law (USC 2006). A yearly evaluation that incorporates a previous 3-year period will help UDOT be aware of the locations that consistently experience unusually high crash frequencies.

6.3 Further Research

As more attention is given to evaluating crashes on roadways in Utah, especially in light of the UDOT Zero Fatalities campaign, there will be increased value in discovering the relationship between crashes and possible causal or predictive factors. As this study used a limited sample of rural two-lane two-way roads, other SPFs derived from Utah data may result in relationships between such factors and crashes that were not found in this study.

The process of developing SPFs or calibrating the HSM SPF models can be immensely improved through incorporating geographic information system (GIS) technology. A GIS can contain the database and scripts used for developing prediction models, and then visually display locations with dangerously high crash frequencies. One study presented at the Transportation Research Board Annual Meeting has successfully applied GIS with the HSM prediction model for rural two-lane two-way roads (Wellner and Qin 2011).

Again, the level of detail for a successful GIS application depends upon the availability of data and resources. It is unlikely for UDOT to have the time and money to develop a new database with characteristics as specific as lane widths, driveways, and striping that designates passing capabilities. However, there are data on AADT, percentages of trucks, and speed limits that have already been collected or are collected on a yearly basis. These data may be sufficient for developing a GIS-based model that can be used to evaluate observed crash frequencies throughout the state on a yearly or multi-year basis.

It was discovered that between 25 and 30 percent of all yearly reported crashes occurring on the segments used in this study from 2005-2007 were caused by wild animals. Future studies may include a variable indicative of the habitats or locales conducive to living conditions for wild animals. This would likely improve upon the predictive ability of the SPF. Another approach would be to remove from the collected data all crashes caused by wild animals and note any changes in the effects on variables used in the models.

As there are multiple functional classes of roads that have unique relationships between safety and the factors that cause crashes (whether considered in this study or not), there are many opportunities to develop jurisdiction-specific SPFs or calibrate the HSM SPFs. It is recommended that UDOT conduct research on the remaining HSM models to provide consistent metrics to safety evaluations.

The SPFs developed in this study do not consider any interacting effects of the variables. It is assumed, for example, that the effect on safety of shoulder rumble strips is completely independent of that of shoulder width. A shoulder rumble strip, however, may be more effective in the presence of a narrow shoulder than a wide shoulder. Further modeling that considers the multiplicative and interactive effects of the model variables could discover these relationships and result in a better understanding of these components that affect safety.

Ongoing safety-related research is critical to maintain safe roads as the components of our roadway system and its users change. Future research into evaluating crashes can lead to more advances in mitigating causes of high crash frequencies. Higher standards can be set as agencies begin to understand the relationship between the factors discussed in this project or others not yet considered and safety.

REFERENCES

American Association of State Highway and Transportation Officials (AASHTO). (2010). *Highway Safety Manual, Volume 2*. Washington, DC.

Banihashemi, M. (2011). "Highway Safety Manual, new model parameters vs. calibration of crash prediction models." *Compendium of the 90th Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, DC.

Cafiso, S., Di Graziano, A., Di Silvestro, G., La Cava, G., and Persaud, B. (2010). "Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables." *Accident Analysis & Prevention*, 42(4), 1072-1079.

El-Basyouny, K., and Sayed, T. (2009). "Accident prediction models with random corridor parameters." *Accident Analysis & Prevention*, 41(5), 1118-1123.

Elvik, R. (2008). "The predictive validity of empirical Bayes estimates of road safety." *Accident Analysis & Prevention,* 40(6), 1964-1969.

Federal Highway Administration (FHWA). (2009). *Manual on Uniform Traffic Control Devices*. U.S. Department of Transportation, Washington, DC.

Fitzpatrick, K., Lord, D., and Park, B. (2008). "Accident modification factors for medians on freeways and multilane rural highways in Texas." *Transportation Research Record: Journal of the Transportation Research Board.* 2083, Transportation Research Board, Washington, DC, 62-71.

Garber, N., and Ehrhart, A. (2000). "Effect of speed, flow, and geometric characteristics on crash frequency for two-lane highways." *Transportation Research Record: Journal of the Transportation Research Board.* 1717, Transportation Research Board, Washington, DC, 76-83.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis*, CRC press, Boca Raton, FL.

Google Earth (Google). (2010). <http://www.google.com/earth/index.html> (October 15, 2010).

Griffin, L., III, Pendleton, O., and Morris, D. (1998). An evaluation of the safety consequences of raising the speed limit on Texas highways to 70 miles per hour., Texas Transportation Institute, The Texas A&M University System.

Gross, F., Persaud, B., and Lyon, C. (2010). A Guide to Developing Quality Crash Modification Factors, FHWA-SA-10-032. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Harwood, D., Bauer, K., Richard, K., Gilmore, D., Graham, J., Potts, I., Torbic, D., and Hauer, E. (2007). Methodology to predict the safety performance of urban and suburban arterials, NCHRP 17-26. Transportation Research Board, Washington, DC.

Harwood, D., Council, F., Hauer, E., Hughes, W.and Vogt, A. (2000). Prediction of the Expected Safety Performance of Rural Two-Lane Highways, FHWA-RD-99-207. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

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

Hauer, E. (1999). Safety in geometric design standards. University of Toronto, Toronto.

Hauer, E., Harwood, D., Council, F., and Griffith, M. (2002). “Estimating safety by the empirical Bayes method: a tutorial.” *Transportation Research Record: Journal of the Transportation Research Board.* 1784, Transportation Research Board, Washington, DC, 126-131.

Jonsson, T., Lyon, C., Ivan, J., Washington, S., Van Schalkwyk, I., and Lord, D. (2009). “Differences in the performance of safety performance functions estimated for total crash count and for crash count by crash type.” *Transportation Research Record: Journal of the Transportation Research Board.* 2102, Transportation Research Board, Washington, DC, 115-123.

Lan, B., Persaud, B., Lyon, C., and Bhim, R. (2009). “Validation of a full Bayes methodology for observational before-after road safety studies and application to evaluation of rural signal conversions.” *Accident Analysis & Prevention*, 41(3), 574-580.

Lord, D., and Bonneson, J. (2005). “Calibration of predictive models for estimating safety of ramp design configurations.” *Transportation Research Record: Journal of the Transportation Research Board.* 1908, Transportation Research Board, Washington, DC, 88-95.

Lord, D., and Bonneson, J. (2007). “Development of accident modification factors for rural frontage road segments in Texas.” *Transportation Research Record: Journal of the Transportation Research Board.* 2023, Transportation Research Board, Washington, DC, 20-27.

Lord, D., and Persaud, B. (2004). “Estimating the safety performance of urban road transportation networks.” *Accident Analysis & Prevention*, 36(4), 609-620.

Mayora, J., and Rubio, R. (2003). “Relevant variables for crash-rate prediction on Spain’s two-lane rural roads.” *Compendium of the 82nd Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, D.C.

Mensah, A., and Hauer, E. (1998). “Two problems of averaging arising in the estimation of the relationship between accidents and traffic flow.” *Transportation Research Record: Journal of the Transportation Research Board.* 1635, Transportation Research Board, Washington, DC, 37-43.

Mitra, S., and Washington, S. (2007). “On the nature of over-dispersion in motor vehicle crash prediction models.” *Accident Analysis & Prevention*, 39(3), 459-468.

National Highway Traffic Safety Administration (NHTSA). (2008). Model Minimum Uniform Crash Criteria. <www.mmucc.us> (January 16, 2011).

National Highway Traffic Safety Administration (NHTSA). (2010a). An Analysis of the Significant Decline in Motor Vehicle Traffic Fatalities in 2008, DOT HS 811-346. National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington, DC.

National Highway Traffic Safety Administration (NHTSA). (2010b). Highlights of 2009 Motor Vehicle crashes, DOT HS 811-363. National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington, DC.

Olsen, A. N., Schultz, G. G., Thurgood, D. J. and Reese, C. S. (2011). “Hierarchical Bayesian modeling for before and after studies.” *Compendium of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.

Park, B., and Lord, D. (2008). “Adjustment for maximum likelihood estimate of negative binomial dispersion parameter.” *Transportation Research Record: Journal of the Transportation Research Board*. 2061, Transportation Research Board, Washington, DC, 9-19.

Persaud and Lyon, Inc., and Felsburg Holt and Ullevig. (2009). Safety Performance Functions for Intersections, CDOT-2009-10. Colorado Department of Transportation, Denver, CO.

Persaud, B., Lan, B., Lyon, C., and Bhim, R. (2010). “Comparison of empirical Bayes and full Bayes approaches for before-after road safety evaluations.” *Accident Analysis & Prevention*, 42(1), 38-43.

Persaud, B., and 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.

The R Project for Statistical Computing (R). (2011). <http://www.r-project.org> (February 18, 2011).

Ramsey, F., and Schafer, D. (2002). *The Statistical Sleuth*. Duxbury, Pacific Grove, CA.

SAS Institute, Inc. (SAS). (2011). SAS® 9.2. SAS, Cary, NC.

Schultz, G. G., Thurgood, D. J., Olsen, A. N., and Reese, C.S. (2010). “Transportation safety data and analysis, volume 1: analyzing the effectiveness of safety measures using Bayesian methods, UT-10.12a, Utah Department of Transportation Research Division, Salt Lake City, UT.

Schultz, G. G., Thurgood, D. J., Olsen, A. N., and Reese, C. S. (2011). “Analyzing raised median safety impacts using Bayesian methods.” *Compendium of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.

Sun, X., Magri, D., Shirazi, H. H., and Gillella, S. (2011). “Application of the Highway Safety Manual: Louisiana experience with rural multilane highways.” *Compendium of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.

United States Code (USC). (2006). Highway Safety Improvement Program, Title 23, Section 148(c)(1)(D).

Utah Automated Geographic Reference Center (Utah AGRC). (2011). “Utah GIS Portal.” <http://www.gis.utah.gov/sgid-vector-download/utah-sgid-vector-gis-data-layers-by-category#TRANSPORTATION> (March 8, 2011).

Utah Department of Transportation (UDOT). (2010). “Utah Department of Transportation “Roadview Explorer” Website.” <http://www.roadview.udot.utah.gov> (February 17, 2011).

Utah Department of Transportation (UDOT). (2011a). Traffic Maps (AADT). < http://www.udot.utah.gov/main/f?p=100:pg:0::::V,T:,2256> (February 18, 2011).

Utah Department of Transportation (UDOT). (2011b). Traffic on Utah Highways (AADT). <http://www.udot.utah.gov/main/f?p=100:pg:0::::T,V:529,> (February 18, 2011).

Vogt, A., and Bared, J. (1998). Accident Prediction Models for Two-Lane Rural Roads: Segments and Intersections, FHWA-RD-98-133. Federal Highway Administration, Washington, DC.

Wellner, A., and Qin, X. (2011). "GIS-based highway safety metrics implementation and evaluation." *Compendium of the Transportation Research Board 90th Annual* Meeting, Transportation Research Board, Washington, DC.

Xie, F., Gladhill, K., Dixon, K. K., and Monsere, C. M. (2011). "Calibrating the Highway Safety Manual predictive models for Oregon state Highways." *Compendium of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington, DC.

Zegeer, C. V., Hummer, J., Reinfurt, D., Herf, L., and Hunter, W. (1986). "Safety Effects of Cross-Section Design for Two-Lane Roads." *Transportation Research Record: Journal of the Transportation Research Board*. 1195, Transportation Research Board, 20-32.