ABSTRACT

DEVELOPING CRASH MODIFICATION FACTORS FOR OPERATIONAL PARAMETERS ON URBAN FREEWAYS

by
Eugene Vida Maina

Studies have shown that, roadway safety has become an intensively investigated topic with the objective of improved understanding of the factors that cause crashes to occur. However, it has been shown that as traffic volumes continue to increase across the United States, 52% of drivers feel less safe on the roads today more than they did five years ago and that the American public feels that traffic safety is a serious problem that needs both the government and media to pay more attention to this issue.

In response to these public and driver grievances, State and National transportation agencies have been and continue to pursue and understand the causes and solutions that would significantly reduce roadway crash frequencies. At national level, through various and rigorous studies, the American Association of State Highway and Transportation Officials, AASHTO has published the Highway Safety Manual to quantify safety using predictive models and CMFs. Various efforts have been attempted at state level too, for example, Texas DOT has developed an Interim Roadway Safety Design Workbook that describes the relationship between various roadway elements and each element influences roadway safety.

In an effort to contribute towards understanding and resolving the factors that influence crash frequencies on roadways, through a thorough literature search. This study realizes that although there has been vast research in this area, no study has explicitly explained why there is variation in crash frequencies on roadways segments with similar physical/geometric features and annual average daily traffic (AADT). Studies suggest that these variations are due to volume changes throughout the day, an effect literature shows that can only be addressed by hourly volumes and not AADT.

Driven by these literature conclusions, this dissertation develops crash modification factors (CMFs) for urban freeways by considering level of service (LOS) deterioration due to change in hourly traffic volumes. Here, this study investigates LOS when it deteriorated from A to B, B to C, C to D, D to E and E to F using hourly volume and hourly crash data collected on urban freeway segments, specifically routes US 1, NJ 3 and NJ 21 in the State of New Jersey. Data were collected on 14 miles of urban freeway segments and 1344 hours of traffic volume count and crash data were analyzed for a period of four years, 2008, 2009, 2010, and 2011.

Results from this investigation, shows that operational elements have some influence on urban freeway safety. This dissertation shows that as LOS deteriorated from A to B, B to C, C to D, D to E and E to F, the estimated CMFs were 0.673, 1.110, 0.865, 1.452, and 0.370 respectively. These findings concur with those referred to in this dissertation's literature review findings, which showed that by adding capacity, that is, by reducing congestion initially results in safety improvement that diminishes as congestion increases.

DEVELOPING CRASH MODIFICATION FACTORS FOR OPERATIONAL PARAMETERS ON URBAN FREEWAYS

by
Eugene Vida Maina

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Transportation

Program in Transportation Engineering

May 2013

.

APPROVAL PAGE

DEVELOPING CRASH MODIFICATION FACTORS FOR OPERATIONAL PARAMETERS ON URBAN FREEWAYS

Eugene Vida Maina

Dr. Janice R. Daniel, Dissertation Advisor Date
Associate Professor of Civil and Environmental Engineering, NJIT

Dr. Athanassios Bladikas, Committee Member Date
Associate Professor, Mechanical and Industrial Engineering, NJIT

Dr. I Jy Chien, Committee Member Date
Professor of Civil and Environmental Engineering, NJIT

Dr. Rongfang Liu, Committee Member Date
Associate Professor of Civil and Environmental Engineering, NJIT

Dr. Jian Yang, Committee Member Date
Associate Professor of Management Science and Information Systems, Rutgers University

BIOGRAPHICAL SKETCH

Author: Eugene Vida Maina

Degree: Doctor of Philosophy

Date: May 2013

Undergraduate and Graduate Education:

- Doctor of Philosophy in Transportation Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2013
- Master of Science in Civil Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2008
- Bachelor of Science in Electronics Engineering Technology,
  University of North Texas, Denton, TX, 2005

Major: Electronics Engineering Technology

- Diploma in Building,
  The Kenya Polytechnic University College, Nairobi, Kenya, 2000

Major: Building Construction

Presentations and Publications:

Daniel, J. R. and E. Maina. "Relating Safety and Capacity on Urban Freeways". Procedia – Social and Behavioral Science. 6th International Symposium on Highway Capacity and Quality of service,. Stockholm, Sweden. 2011.

Maina, E. "Relating Safety and Capacity on Urban Freeways". Office of the Provost, New Jersey Institute of Technology. Dana Knox Student Research Showcase. Newark, NJ. 2011.

To My Father,
The Late Andrew Maina Andieri

In life, one has time, youth, and mind. Time will come and go no matter what, so will youth. One should enjoy these two responsibly - they have consequences. It gets tough as they both pass by - that is why the old days will always be the good old days. The mind is the most powerful of the three; the worst mistake would be to waste it. Instead, the mind should be kept young with time. How? Keep it thirsty for knowledge and happiness.

It is also a waste of life if these three are natured but not used to improve life for others, the environment, the future, and oneself.

Words of the late, Andrew Maina Andieri.

ACKNOWLEDGMENT

I lack the words to describe my appreciation to Dr. Janice Daniel, my dissertation advisor, friend, and mentor. Her sharp-witted mind, patience, persistence, thoroughness, and assistance have carried me through this enduring study and successful end. I cannot thank her enough. Together with Dr. Daniel, I would like to express my gratitude to my dissertation committee members: Dr. Athanassios Bladikas**,** Dr. I Jy Chien**,** Dr. Rongfang Liu, and Dr. Jian Yang for their insightful contribution. Special thanks to Dr. Jay Meegoda for his assistance and guidance throughout my graduate studies.

I am thankful and a beneficiary of the U.S. Army's Federal Tuition Assistance program, the U.S. Department of Veterans Affairs Montgomery GI Bill program, and the New Jersey Army National Guard's tuition waiver program. I am grateful for the New Jersey Department of Transportation's easy to access and detailed motor vehicle crash database archive. Without it, I would have not have had he required crash data to tackled this topic.

To end with, I hold my highest respect to my father, the late Andrew Maina Andieri, and my mother, Wilfrida Afandi Maina. They denied themselves everything for my Sisters (Caroline and Valentine), Brother (Cyprian) and I to have a decent education. I am thankful they for never gave up on me even when I seemed lost and undecided. I am pleased I listened to, and observed them. To my recently born Daughter, Efa Afandi Vida, I hope I will be to you what my parents were to me and that you stay healthy, be a law-abiding citizen, and prosper.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

**TABLE OF CONTENTS
(Continued)**

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Whenever a driver is on the road operating a vehicle, a risk of a crash occurring exists. The crash can be associated with human factors, roadway physical/geometrical features, weather conditions, environmental factors, operational elements, bad luck or a combination of several of these features.

Characteristically, all roads have some level of risk, but some roadway sites (segments or intersections) are contemplated to be more risky than others are. Some of those identified to be more hazardous possess similar physical or geometrical features but have varying crash frequencies. What could be the cause of these variation in crash frequencies even though the roadway sites have similar or same physical features? In the past, practitioners have considered geometrical elements and Annual Average Daily Traffic (AADT) to measure the absolute crash frequency at a roadway site to proclaim whether the site is a safety concern or not. This approach however tend not to address the variation in crash frequencies at these similar or same sites.

Roadway safety has become an intensively investigated topic with the objective of improved understanding of the factors that cause crashes to occur. If the factors that cause crashes are known, practitioners will be able to identify safety improvement countermeasures more efficiently and effectively. Earlier and recent studies have indicated that the relationship between crash frequency and traffic volume is nonlinear and depends on several variables.

Traffic can act exceptionally different depending on the roadway functional class, area type, physical features etc. Therefore, treating all sites the same does not reflect the influence operational conditions have on safety. This dissertation provides an empirical and unbiased methodology of measuring safety on urban freeway segments and as recommended by the HSM (2010) the measured safety was reported as CMFs.

1.2 Research Need

Most transportation studies address how geometrical elements influence roadway safety and not operational elements. The few that have, do acknowledge that operational elements have some influence on safety, but fail to quantify safety as those for geometrical elements have. For example, Kononov et al.'s, (2008) relationships between safety and congestion study indicated that even though practitioners believe that additional capacity afforded by additional lanes is associated with more safety, they do not specify how much safety and for what period of time the given freeway segment will maintain the safety conditions.

No published study, including the HSM (2010) has quantified safety by considering operational elements. There is need to explicitly understand how much safety and for what period of time or condition will a specified operational element affect a given urban freeway segment.

The reason that has made the use of operational elements data to quantify safety unpopular amongst transportation practitioners, is the hesitance by the HSM (2010) to use hourly traffic volumes instead of the traditional AADT. The analytical derivation and factors adopted by the HSM (2010) render the exercise also impractical. However, studies have shown that urban freeway segments with the same geometric conditions experience varying crash frequencies. For example, Qin et al. (2006) study on the relationship between crash occurrence and hourly traffic volumes shows that the expected crash count on two equal segments with the same AADT and physical characteristics varies according to the distribution of traffic volume through the time of the day. Variation in crash frequencies on similar segments is due to volume changes that affect roadway operations throughout the day (Abdel-Aty and Pande, 2007), an effect Hauer et al. (1996) states that only hourly volumes can accurately account for. As a result, this study focused on hourly traffic volumes instead of AADT to investigate operational influence on safety.

The HSM (2010) recommends that safety evaluation to be performed before implementing any given treatment. This exercise helps to predict the expected safety consequences of the suggested treatment if there are any. According to the HSM (2010), the safety consequence can be measured using either crash prediction models or CMFs or both. In accordance to the HSM (2010), this study developed CMFs to measure the safety effect of the suggested operational implementations. The HSM (2010) has effective CMFs for various types of treatments. However, none of the HSM (2010) CMFs is developed for roadway operational functions. It is for this reason that this study focuses on this new approach to quantify safety. The findings of this research will serve as a starting point to convince transportation researchers to further investigate this topic and eventually be included in the future version of the HSM.

No published study directly links the highway capacity manual, HCM (2010) to the HSM (2010). The HCM (2010) and the HSM (2010), two essential manuals referred to for concepts, guidelines and computational procedures for estimating operational elements and quantifying the safety effects of various engineering treatments proposed during roadway planning, design, operations or maintenance respectively are independent of each other. This study establishes a link between these two manuals showing dependency on each other when considering operational safety impacts.

1.3 Objectives of Study

This dissertation's findings will fill the gap regarding operational elements - specifically how they influence safety and eventually contribute towards the inclusion and improvement of the future versions of the HSM. Therefore, the seven main objectives for this research are:

1. Use hourly volumes to investigate the relationship between operational elements and safety.
2. Use traffic density and level of service (LOS) to investigate the relationship between operational elements and safety
3. Develop SPFs to determine predicted crash frequencies.
4. Develop Empirical Bayesian models to determine expected crashes.
5. Derive CMFs to quantify the impact operational elements have on safety under given conditions.

6. Use the findings of this study to show that operational elements have some influence on roadway safety.

7. Establish a link or relationship between the HCM and HSM.

This study has derived and presented a thorough procedure of developing CMFs and as a result, contributed to improving the quality of research and roadway safety design. In addition, a link between the HCM and HSM was established. Through the models developed and presented in this study, a better perception of operational elements and their influence on safety were revealed. The other intent here was also to encourage further investigations on this topic and eventual the inclusion in the future versions of HSM.

The objectives of this study were accomplished by attempting and completing the following tasks:

1.3.1 Literature Review

A detailed literature search on factors that influence roadway safety and development of CMFs was conducted. The objective of the literature search was to unearth information of previous findings to assist in obtaining the objectives of this study. The main sources this study used to obtain prior studies on this topic included *TRID, NJIT's Van Houten library, NJIT's National Center for Transportation and Industrial Productivity, Science Direct, Dissertation Abstracts Online, NJDOT* and *SCOPUS.*

1.3.2 Survey to Select Study Roadways

To achieve a nearly realistic and unbiased outcome, this dissertation studied roadways only classified as urban freeways only. This is because urban freeways have a high chance of experiencing all six LOSs changes in a twenty-four hour day period at a given sites and the drivers can be assumed to be local drivers and therefore the adjustment factor for presence of occasional or non-familiar drivers can be assumed to be 1.

A thorough survey was conducted on NJDOT's "*Straight Line Diagram, SLD*" database to identify roadway segments that were classified as urban freeways. Routes US 1, NJ 3 and NJ 21 were selected for investigation. Other site characteristics such as location, street name, speed limit (mi/hr.), number of lanes, median width (ft.), lane width (ft.) and shoulder width (ft.) were also collected from this database to assist in model development and CMF analysis.

1.3.3 Survey to Select Study Sites

Using the selected roadways, this study referred to NJDOT's "*Roadway Information and Traffic Counts*" database to select the study sites. Five one-mile (half-mile downstream and half-mile upstream) study sites were selected on each roadway at locations where NJDOT had placed volume count stations. Twenty-four hour annual average hourly traffic volumes for the years 2008, 2009, 2010 and 2011 were collected for analysis.

1.3.4 Determining the Level-of-Services, LOS

To determine the LOS for each of the 1,344 study hours LOSs, this study used the procedure given in the HCM (2010) finding the density first and then assigning LOS. To achieve this task, roadway characteristics data given *Section 1.3.2* and hourly traffic volumes given in *Section 1.3.3* were used to compute the density and assigning of the LOSs for each study hour.

1.3.5 Crash Data Collection

NJDOT's *"Crash Records"* database was the only source used to find the hourly crash frequencies for this study. This study then merged each hourly crash frequency value with its corresponding traffic volume, roadway characteristics, density and LOS and used these information to determine the variables included in the development of safety performance functions (SPFs).

1.3.6 Developing Safety Performance Functions and Crash Modification Factors

In accordance to the HSM (2010), the safety consequence was measured using CMFs, a function of observed crashes given in *Section 1.3.5*, predicted crashes determined from SPFs and expected crashes determined form Empirical Bayes, EB before-after studies. Therefore, the development of CMFs involved selecting roadway physical/geometrical and operational elements generally considered to be related to safety.

1.4 Definition of Important Acronyms Used

LOS: **Level-of-Service** is “a quality measure describing operational conditions within a traffic stream, generally in terms of such service measures as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience.” HCM (2010).

HCM **Highway Capacity Manual “**is a publication of the Transportation Research Board of the National Academies of Science in the United States. It contains concepts, guidelines, and computational procedures for computing the capacity and quality of service of various highway facilities, including freeways, highways, arterial roads, roundabouts, signalized and un-signalized intersections, rural highways, and the effects of mass transit, pedestrians, and bicycles on the performance of these systems.” HCM (2010).

SPF: **Safety Performance Functions** are statistical models relating crash frequencies to roadway and driver characteristics.

RTM: **Regression-to-Mean** is the tendency for the occurrence of crashes at a given location to fluctuate up and down, and to converge to a long-term average. Bahar (2010).

HSM: **Highway Safety Manual** is a "provides practitioners with information and tools to consider safety when making decisions related to design and operation of roadways. The HSM assists practitioners in selecting countermeasures and prioritizing projects, comparing alternatives, and quantifying and predicting the safety performance of roadway elements considered in planning, design, construction, maintenance, and operation." AASHTO (2010).

SLD: **Straight Line Diagram** is a diagram of a road where the road is shown as a straight line. Such diagrams are usually produced by a highway department, and display features along the road, including bridges and intersecting roads. Rows below the diagram show data about the road, usually including speed limit, number of lanes, bridge numbers, and historical data, among other data. Subway lines also frequently employ straight-line diagrams.

CMF: **Crash Modification Factors** is a multiplicative factor used to compute the expected number of crashes after implementing a given countermeasure at a specific roadway site. HSM (2010).

AADT: **Average Annual Daily Traffic** is the average 24-hour volume at a given location over a full 365/366 day year; the number of vehicles passing a site in a year divided by 365/366. HCM (2010).

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter summarizes a review of literature describing existing and proposed studies on (1) the methodology approach for estimating the operational performance for a basic freeway segment, (2) relationships between roadway geometric and operational elements; and safety and (3) methodologies for estimating crash modification factors (CMFs). This chapter is organized in eight Sections: Section 2.1 introduces the chapter. Section 2.2 discusses the role of the Highway Capacity Manual (HCM, 2010) and the Highway Safety Manual (HSM, 2010) in analyzing existing and proposed roadways in the US. Section 2.3 discusses the geometric elements considered by the Highway Safety Manual (HSM, 2010) in developing existing crash modification factors (CMFs). Section 2.4 discusses roadway operational elements and how they can affect safety. Section 2.5 discusses the Average Annual Daily Traffic (AADT) and how it affects safety. Section 2.6 discusses hourly volumes, how they affect roadway safety and why they may be used instead of AADT in developing the CMFs associated with freeway operation performance, Section 2.7 discusses the various methodologies used to develop CMFs and Section 2.8 summarizes the literature search chapter.

2.2 Analysis for a Basic Freeway Segment

This Section provides a review of the two relevant freeway segment analysis manuals used in the US: (1) The Highway Capacity Manual (HCM, 2010) published by Transportation Research Board (TRB) for preliminary roadway designs and (2) the Highway Safety Manual (HSM, 2010) published by American Association of State Highway and Transportation Officials (AASHTO). Though very effective, these two manuals are generally independent to each other. As demonstrated in Chapter 1 Section 1.2, this study intends to show that both these manuals can be referred to concurrently to develop CMFs to enable practitioners to measure the expected safety for both proposed and existing roadways.

2.2.1 Highway Capacity Manual (HCM, 2010)

This Section provides an overview of the HCM (2010). This is being provided because it will be used to provide the procedures and guidance practitioners follow to estimate the operational performance of a basic freeway segment. This manual provides concepts, guidelines and computational procedures for estimating maximum service flow rate (capacity) measured as passenger cars per hour per lane (pc/h/ln) and the level of service (LOS) for both uninterrupted and interrupted freeway Sections. The manual also incorporates the effects of mass transit, pedestrians and bicycles on the performance of roadway systems. The HCM (2010) also provides the guidelines to calculate the volume to capacity ratio (v/c ratio), a ratio of estimated or existing demand flow over the capacity of a given facility (Roess et al. 2011).

One of the freeway quality measures estimated in the HCM is LOS. As Section 2.4.2 of this chapter shows, LOS impacts safety, so does capacity which is the service flow rate at LOS "E". According to the HCM (2010), LOS is

> "*a quality of measure describing operational conditions within traffic stream, generally in terms of service measures such as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience*".

LOS is measured as a letter between A and F. A Section rated "A" gives the best operational quality and "F" the worst operational quality (HCM, 2010). The HCM (2010) indicates that the LOS of a basic freeway segment is generally determined in three steps: (1) Determining the free-flow speed, (2) Determining the demand flow rate under base conditions and (3) determining density, speed and LOS.

Another basic freeway quality measure is service flow rate, the maximum rate of flow that can reasonably be expected on a given lane or roadway under prevailing conditions while maintaining a particular LOS (HCM, 2010). The service flow rate at LOS "E" is the capacity for a basic freeway segment where capacity is the maximum number of vehicles that can be accommodated at the basic freeway segment (HCM, 2010). Service flow rate is calculated as follows.

$$SF_i = MSF_i \times N \times f_{HV} \times f_P \tag{2.1}$$

Where: SF_i = service flow rate for LOS "i," veh/h

MSF_i = maximum service flow rate for LOS 'i," pc/h/ln

N = number of lanes (in one direction) on the facility

f_{HV} = adjustment factor for presence of heavy vehicles as follows

$$\mathrm{f_{HV}} = \frac{1}{1 + \mathrm{P_T}(\mathrm{E_T} - 1) + \mathrm{P_R}(\mathrm{E_R} - 1)} \tag{2.2}$$

Where: P_T = proportion of trucks and buses in the traffic stream (given)

P_R = proportion of RV in the traffic stream

E_T = passenger-car equivalent for trucks and buses

E_R = passenger-car equivalent for RVs

f_P = adjustment factor for presence of occasional or non-familiar users of a facility

2.2.2 Highway Safety Manual (HSM, 2010)

The HSM (2010) provides factual information and tools for quantifying the safety effects of various engineering treatments proposed during roadway planning, design, operations or maintenance. The predictive methodologies of the highway Safety Manual express the safety performance of a roadway as the expected number of crashes. Safety treatments, and their impact on the expected number of crashes, are expressed as crash modification factors (CMFs). To develop CMFs, the HSM considers 33 roadway and traffic variables listed in Appendix A.

A CMF is a ratio between the number of crashes expected after a modification is implemented and the number of crashes if the change does not take place (HSM, 2010). Bahar (2010) describes a CMF as a multiplicative factor used to compute the expected number of crashes after a given countermeasure is implemented at a given roadway Section. Part D of the HSM provides six steps used to generate CMFs. These steps are a result of extensive literature review of published highway safety research studies spanning more than fifty years (Bahar, 2010). Bahar (2010) also states,

> "*evidence-based and rigorous review, supported by statistical evidence of the accuracy and validity of studies, was applied,*"

has resulted to adoption of the six steps by the HSM (2010) to determine CMFs. They are:

1. Determine the safety effect (CMF) of the implemented countermeasure as documented in a study publication;
2. Adjust the estimated CMF from step 1 to account for bias from either or both regression-to-mean (RTM) and changes in traffic volumes;
3. Determine the ideal standard error of the CMF
4. Adjust the standard error from step 3 by applying a given method of correction factor, MCF;
5. Adjusts the standard error to account for bias from either or both RTM and changes in traffic volumes;
6. Standardize or combine CMFs from similar roadway Sections.

Operational and safety elements are presently discussed in the HCM (2010) and HSM (2010). Both manuals are very important in analyzing proposed or existing roadways. However, both manuals are generally independent of each other and neither manual accounts for the impact operation performance has on safety. For example, the HCM (2010) provides the step to determine capacity, LOS and v/c ratio but does not show how each function impacts safety, as Section 2.4 of this dissertation will shows, these functions significantly affect safety. The HSM (2010) measures safety but only considers geometrical elements and AADT and not operation performance to generate CMFs. Although Foster et al. (2009) states that CMFs quantify the potential change in expected average crash frequency as a result of both geometric and operational modifications, the list of the 33 variables in Appendix A of this dissertation indicates that the HSM (2010) only considers geometric elements and AADT to develop CMFs. This study considered operational elements to developed CMFs based on the hourly traffic volumes on urban freeway segments with the intention that this topic will be further investigated by transportation practitioners and also be included in future versions of the HSM.

2.3 Geometric Elements Considered by the HSM to develop CMFs

Geometric design of roadways refers to portioning of the physical elements of a given roadway according to standards and constraints to provide smooth-flowing, crash-free facilities. As stated earlier, both the HCM and HSM evaluate the performance of a given roadway considering geometric elements and AADT. For safety performance assessment, the HSM guidelines take into account the site characteristics and traffic-volume variables shown in Tables 2.1 and 2.2 for uninterrupted roadway segments and interrupted segments respectively.

According to the HSM, the major characteristics are: length of roadway segment, number of through lanes, lane width, shoulder width, presence of median, median width and left turn lanes. This Section of the literature review discusses these elements and how they impact safety individually and if other elements, specifically operational elements were considered in the analysis.

Table 2.1 Segment Variables used in HSM Predictions

VARIABLES	**Rural Two-Lane, Two-Way Roads**	**Rural Multilane Highways**	**Urban/Suburban Arterials**
Area Type (rural/suburban/urban)	**X**	**X**	**X**
AADT	**X**	**X**	**X**
Length of roadway segment	**X**	**X**	**X**
Number of through lanes	**X**	**X**	**X**
Lane Width	**X**	**X**	
Shoulder width	**X**	**X**	
Shoulder type	**X**	**X**	
Presence of median (divided/undivided)		**X**	**X**
Median width		**X**	
Presence of concrete median barrier		**X**	
Presence of passing lane	**X**		
Presence of short four-lane Section	**X**		
Presence of two way left-turn lane	**X**		**X**
Driveway density	**X**		
Number of major commercial driveways			**X**
Number of minor commercial driveways			**X**
Number of major residential driveways			**X**
Number of minor residential driveways			**X**
Number of major industrial/institutional driveways			**X**
Number of minor industrial/institutional driveways			**X**
Number of other driveways	**X**		
Horizontal curve length	**X**		
Horizontal curve radius	**X**		
Horizontal curve super-elevation	**X**		
Presence of spiral transition	**X**		
Grade	**X**		
Roadside hazard rating	**X**		
Roadside slope		**X**	
Roadside fixed-object density			**X**
Roadside fixed-object offset			**X**
Percent of length with on-street parking			**X**
Type of on-street parking			**X**
Presence of lighting			**X**

Table 2.2 Intersection Variables used in HSM Predictions

VARIABLES	Rural Two-Lane, Two-Way Roads	Rural Multilane Highways	Urban/Suburban Arterials
Area type (rural/suburban/urban)	X	X	X
Major-road AADT	X	X	X
Minor-road AADT	X	X	X
Number of intersection legs	X	X	X
Type of intersection traffic control	X	X	X
Left-turn signal phasing (if signalized)			X
Presence of right turn on red (if signalized)			X
Presence of red-light cameras			X
Presence of median on major road		X	
Presence of major-road left-turn lanes(s)	X	X	X
Presence of major-road right-turn lane(s)	X	X	X
Presence of minor-road left-turn lanes(s)		X	
Presence of minor-road right-turn lane(s)		X	
Intersection skew angle	X	X	
Intersection sight distance	X	X	
Terrain (flat vs. level or rolling)		X	
Presence of lighting		X	X

2.3.1 Length of Roadway Segment

Anastasopoulos' et al. (2008) main objective was to determine the factors that influence the frequency and severity of accidents on homogeneous segments of Indiana's rural interstate highways to provide effective safety-related countermeasures, one of the factors studied was length of roadway segment. The study developed negative binomial regression models to analyze accident data collected on interstates I-64, I-65, I-70, I-74 and I-164 over a period of over 5 years. The accident data consisted of 322 homogeneous segments. The study found that crash frequencies increased as the segment lengths increased and conversely decreased when the segment length decreased. The study shows the expected trend when the segment length increases but does not indicate the expected crash frequency associated with given segment length values, also operation performance elements were not considered.

2.3.2 Shoulder Width

Strathman et al. (2001) investigated the statistical relationship between crash frequency and roadway design attributes on the Oregon state highway system. The study developed CRFs to analyze crash data obtained from ODOT's crash database and found that number of lanes, curve characteristics, vertical grade, surface type, median type, turning lanes, lane width and shoulder width were statistically related to crash activities. For shoulder width, the study shows that for every 1 foot of right shoulder width added to a freeway segment, the crash frequency decreased by a value of 0.04.

2.3.3 Number of Lanes

Practitioners generally believe that additional capacity afforded by additional lanes is associated with more safety, however, they do not specify how much safety, and for what period of time the given freeway segment will maintain the safety conditions are generally not considered (Kononov et al. 2008). A number of studies tend to state similar results as shown below.

Kononov et al. (2008) investigated the relationship between safety and number of lanes on urban freeways. The study developed safety performance functions (SPFs) using 5 years' of accident data from the states of California, Colorado and Texas. By comparing the slopes of the SPFs, the study showed that increasing the numbers of lanes on urban freeways, initially resulted in safety improvement that diminished as congestion increased.

Garber and Ehrhart's (2000) main objective was to develop mathematical relationships that describe the combined influence traffic volume and geometric characteristics have on crash occurrences. The study developed multivariate ratio of polynomials models to analyze freeway crash data obtained from Virginia's DOT data base and police accident reports from January 1993 to September 1995. The study concluded that the crash rates tended to increase as the standard deviations of speed increased. On the contrary, in the literature review performed by Garber and Ehrhart (2000), Lundy (1965) states that as the number of lanes increases, the crash rates decreases.

Milton and Mannering's (1998) intent was to develop a statistical model of accident frequency that could be used to isolate accident-prone Sections of highway. The study then developed negative binomial regression models and analyzed annual accident frequency data from Sections of principal arterials in Washington State for the years 1992 and 1993. In all, 31306 observations were used in the model. The study determined that more lanes tend to increase accident frequency.

These studies by Kononov et al. (2008), Garber and Ehrhart (2000), Milton, and Mannering (1998) indicate that as the number of lanes increase, the crash frequency also increases, but the severity specifically, fatalities, decrease. Controversially, Noland and Oh (2003) as well as Mussa and Chimba (2006) indicate that not only do the crash frequency increase when the number of lanes are increased on a given roadway, but the fatalities also increases.

Noland and Oh (2003) investigated whether various changes in road network infrastructure and geometric design could be associated with changes in road fatalities and reported accidents. The study developed negative binomial models to analyze data from the Highway Safety Information System (HSIS) for the State of Illinois and found that increases in number of lanes appears to be associated with both increased traffic-related accidents and fatalities.

Mussa and Chimba's (2006) objective was to build crash prediction models that would reveal significant variables that influence crash frequencies. The study involved developing Zero-Inflated negative binomial (ZINB) models to analyze crash data from Florida State Highway system and showed that non-limited access roadways with 6 or more lanes had both higher fatalities and crashes than 4-lane roadways.

2.3.4 Lane Width

Gross and Donnell's (2011) main objective was to compare the case-control and cross-Sectional methods to estimate measures of safety effectiveness using two independent datasets. The safety effects of various lane and shoulder widths were estimated and compared using cross-Sectional and case-control methodologies to estimate CMFs for fixed roadway lighting and the allocation of lane and shoulder widths in the States of Minnesota and Pennsylvania. Based on case-control method, the study indicated that the CMF for intersections with lighting was 0.886 and the CMF for the cross-sectional study was 0.881 for cross-sectional method.

2.3.5 Median Width (With Barrier)

Pande et al.'s (2010) study had two main objectives; (1) to outline some of the functions (i.e., geometric design and time of day) associated with crash frequency and (2) propose a classification tree based methodology of identifying traffic and highway design parameters are significantly associated with crashes on expressways/freeways. The study developed negative binomial regression models to analyze crash data on US Route 19 also known as SR 55 in Pasco County Florida and found that as the median width increased, the percentage of crashes decreased. The study investigated 18 ft., 24 ft. and 28 ft. median widths and found that they were associated with a reduction of 11.36%, 11.85%, and 9.70% crashes, respectively.

2.3.6 Median Width (No Barrier)

Bonneson et al.'s (2009) main objective was to investigate the relationship between various geometric design components and their corresponding safety effects. The findings of this investigation i.e., accident modification factors (AMFs) were to be adopted as safety design guidelines and evaluation tools by TxDOT designers in the planning and design stages of project development. The study used correlations to investigate the impact median width (no barriers) had on safety in the state of Texas. The study found that when the median width (no barrier) decreased from 64-ft. to 48-ft. the crash frequency increased by 4.1%. Therefore, it can be concluded that reduction in median width (no barrier) results in increased crashes. The study however did not investigate the operation conditions and how they influenced safety.

2.3.7 Conclusions

The studies sited in this Section show how geometric elements influences safety. However, with the exception of Strathman et al. (2001) and Gross et al. (2011), none of these studies measured the expected number of crashes upon implementing a given countermeasure. Table 2.3 shows the summary of geometric elements, methodology and whether or not operation performance was considered for each study. The table indicates that none of these studies considered operation performance as a variable in their study, a variable this study intends to investigate.

Table 2.3 Geometric Elements Literature Search Summary

AUTHOR(S)	**Geometric Element of Study**	**Methodology**	**Safety Quantified**
Anastasopoulos et al. (2007)	Segment Length	Negative Binomials	NO
Strathman et al. (2001)	Shoulder Length	CRFs	YES
Kononov et al. (2008)	Number of Lanes	SFPs (Safety Performance Functions)	NO
Milton and Mannering (1998)	Number of Lanes	Poisson Regression	NO
Garber and Ehrhart (2000)	Number of Lanes	Multivariate Ratio of Polynomials	NO
Noland and Oh (2003)	Number of Lanes	Negative Binomial Regression	NO
Mussa and Chimba (2006)	Number of Lanes	Zero-Inflated Negative Binomial (ZINB) Regression	NO
Gross and Donell (2011)	Lane Width	CMFs	YES
Pande et al. (2010)	Median Width (with Barrier)	Negative Binomial Regression	NO
Bonneson et al. (2009)	Median Width (no Barrier)	Correlations	NO

2.4 Operational Elements

Operation analyses define all traffic parameters, roadway parameters, and control conditions for an existing or proposed roadway segment (Roess et al. 2010). This analysis also facilitates the determining of expected LOS, capacity and v/c ratio of a given roadway segment. This Section shows the findings from previous studies on how roadway operation performance influences safety.

2.4.1 V/C ratio

Lord et al.'s (2004) main objective was to establish the relationship between crashes and various traffic flow characteristics in Quebec, Canada. The study used predictive models (i.e., functional forms) to evaluate 5 years of crash data and determined that although the effects of v/c ratios on safety have not been clearly established nor properly modeled, v/c ratio, along with other traffic flow characteristics have direct influence on the likelihood and severity of a crash. The study also concluded that crash risk and the number of crashes increases with higher vehicle density and v/c ratios.

2.4.2 Level of Service (LOS)

Lord et al. (2004) also investigated how LOS influenced roadway safety using the same data and predictive models from Section 2.4.1 above. The discussion of previous work prepared by Lord et al. (2004) discusses studies by Frantzeskakis and Iordanis (1987), Persaud, and Nguyen (2000). These studies examined the effects of LOS on safety and concluded that both crash frequencies and crash rates increased as the LOS decreased from LOS of "A" to LOS of "F."

2.4.3 Capacity

Kononov et al. (2008) investigated the effect capacity had on safety for urban freeways in the states of Colorado, California and Texas. Safety Performance Functions (SPFS) were developed based on crash data and determined that when capacity is increased, the number of crashes temporarily reduced but increased with congestion.

2.4.4 Density

Density is the number of vehicles occupying a given length of a roadway or lane, expressed as vehicles per mile or vehicles per mile per lane. Generally, density is difficult to measure due to an elevated vantage point from which the Section under study may be observed is required (Roess et al. 2011). However, density can be expressed in terms of speed and flow measurements as shown in this study.

Density is an important traffic stream measure because it is directly related to traffic demand. Because (1) Drivers choose speeds according to how close they are to other vehicles, the speed and density combine to give the observed rate of flow and (2) it also a measure of the nearness of other vehicles; this influences the freedom to maneuver and the psychological comfort of drivers.

2.4.5 Conclusions

The studies mentioned in this Section, show how safety is affected by the operational performance measures of a roadway including, v/c ratio, LOS, capacity and density. However, none of these studies measured the safety effect each studied function had on the crash frequency. To be specific, none of these studies has shown the expected number of crashes when each function is increased or decreased.

In order to determine the relationship between operational and safety considerations in geometric design improvements, Harwood (1995) states that

> "*It would be extremely valuable to know how safety varies with v/c ratio and what v/c ratios provide minimum accident rate. Only limited research has been conducted on the variation of safety with v/c [volume–capacity] ratio. More research of this type is needed, over a greater range of v/c ratios, to establish valid relationships between safety and traffic congestion to provide a basis for maximizing the safety benefits from operational improvement projects.*"

As a follow up to Harwood's (1995) statement above, this study investigates the safety effect of v/c ratio and density. The v/c ratios and densities in this study will be developed from hourly volumes and the resultant effects will be reported as CMFs. Knowledge of this relationship would help transportation practitioners precisely understand the safety implications for both projected traffic growth on existing highways and of highway improvements designed to increase capacity (Hall and Pendleton, 1990).

2.5 Average Annual Daily Traffic (AADT)

According to Section 2.2, AADT is the only operation performance function the HSM (2010) uses to measure the safety effect of a given roadway. The AADT is the average 24-hour volume at a given roadway segment over a full 365/366 day year; the number of vehicles passing a given site in a year divided by 365/366 days (HCM, 2010). In simpler terms, AADT is the average number of vehicles that pass a given roadway Section each day in a given year (Castro-Neto et al. 2009).

Garder (2006) investigated the segment characteristics and severity of head-on crashes on two-lane rural highways in the state of Maine. Probit regression models were developed to analyze crash data for the years 2000, 2001 and 2002 obtained from Maine's DOT data base. The study found that an AADT of less than 2000 veh/day was associated with 5.2% crashes and an AADT of more than 2000 veh/day was associated with 7.2% crashes. Contrary to all the previous studies in this Section, this study not only showed that crash frequencies increased with AADT, but also the severity increased with crash frequency.

State Department of Transportation and local transportation agencies have collected and predicted AADT for design, planning and administrative purposes (Seaver et al. 2000). However, for most cases, AADT does not represent the correct volume conditions at the time of crash (Castro-Neto et al. 2009). As a result, researchers are moving toward microscopic crash analysis which includes analysis of hourly crash data (Abdel-Aty and Pande, 2007).

Hourly volumes account for the uncertainty in the measurement of AADT values and incapability of AADT capturing accurate traffic flow variations (Abdel-Aty and Pande, 2007). Also, unlike AADT, hourly volumes show logical measures of congestion represented by v/c ratio and LOS (Frantzeskakis and Iordanis, 1987 and Persaud and Nguyen, 1998) along with distributional properties of variation in speed (Abdel-Aty et al.

2006). Based on these reviews, this study will consider hourly volumes and not AADT to develop CMFs to measure safety.

2.6 Hourly Volume

Hourly volume is the volume of traffic that traverses across a segment of a roadway in a given hour, expressed as vehicles per hour (veh/h) (Ivan et al. 2000). Unlike hourly volumes, the other measures applied to quantify the chances for crashes, such as AADT, VMT and NEV (Number of Entering Vehicles), do not consider the temporal traffic variation (Wang and Ivan, 2000). For example, AADT, VMT or NEV cannot accurately account for the distribution of weekday to weekend traffic volume that might vary from one location to another or from daytime to nighttime. This effect can be accurately for by hourly volumes (Hauer et al. 1996). As the following studies indicate, crashes at a given time should relate closely to the hourly traffic volume.

Qin et al. (2006) investigated the relationship between crash occurrence and hourly volume counts on rural two-lane highway segments in the States of Michigan and Connecticut. The study used a hierarchical Bayesian framework with Markov Chain Monte Carlo (MCMC) algorithms to estimate the posterior distributions for crash probabilities as a function of hourly volume and time of day. The study demonstrated that the expected crash count on two equal length segments with the same AADT and physical characteristics will vary according to the distribution of traffic volume through the time of the day.

Ceder and Livneh (1982) investigated the relationships between measures of accidents and traffic flow by considering hourly flow instead of ADT (Average Daily Traffic). The study developed power functions to analyze accident data collected on eight four-lane road Sections for a period of 8 years in England. The study indicated that there is a U-shaped relationship between hourly flow rate and the number of crashes. i.e., the number of crashes increased during day hours and decreased during night hours.

Persuad and Mucsi (1995) studied the relationship between hourly traffic volumes and crash frequency on two-lane rural roads in Ontario, Canada. The study developed negative binomial regression and empirical Bayesian (EB) models to analyze the crash data for different time periods (24 hr., day and night hours) and found that a convex relationship existed between hourly traffic volumes and crash frequency.

These studies draw conflicting and consistent conclusions, showing that indeed a relationship exists between crash frequency and hourly volume even though the trend is still unknown.

2.7 Crash Modification Factors (CMFs)

CMFs quantify or measure the change in expected average crash frequencies or crash effects at a given location after implementing a particular treatment (also known as countermeasure, or intervention action, or alternative), design modification or change in operations (HSM, 2010). CMFs indicate that a change in either geometry or operation conditions could result in either an increase or a decrease in crashes (Mbatti, 2011). A CMF of 1.00 indicates no effect on safety, while a CMF of more than 1.00 shows that the treatment could result in a safety degradation and if the CMF is less than 1.00, a safety benefit is expected (Gross et al. 2010).

In CMF development, two main fundamental methods can be adopted: (1) Before-and-after studies and (2) observational cross-Section studies (Bahar, 2010). As the following sub Sections indicate, the availability of crash data to be analyzed determines which method to adopt. In addition, the amount of data determines the accuracy of the analysis, the larger the sample, the more reliable the results are (Stroud, 1995). There are other studies, however not commonly used. They are Case-control studies and Cohort Studies.

2.7.1 Before-and-After Study

Before-and-after studies involve assessing either the number of crashes or some other measure of risk before and after a given countermeasure is implemented on either a single or several sites (Gross et al. 2010). Generally, this analysis involves the comparison of a "treatment group" and a "control group": Sites in the "*treatment group*" have treatments implemented and number of crashes measured before and after the implementation occurs. Sites in the "*control group*" do not have treatments implemented and numbers of crashes are measured for the same before-and-after periods as for the "*treatment group*." (Lawson, 2011).

Generally, crash data for three years or more are used in before-and-after study evaluations. However, five years may give a more consistent picture of the long term risks at the study site (Lawson, 2011).

2.7.2 Cross Sectional Study

Cross Sectional studies are conducted in the event that a before-and-after study is impractical (Gross, 2006). This method is used when comparing the safety performance of a site with certain special features to another site without these special features (Connor et al. 2001). For example, this study could be used to measure the safety effect of a given countermeasure before it is implemented (Mbatti, 2011).

The National Cooperative Highway Research Program (NCHRP) Report 617 (Harkey, 2008), found this method to be appropriate while determining the safety effect for traffic engineering and intelligent transportation systems (ITS) improvements. In this study, the estimated safety effects of different geometric elements on roadways from 34 States were used to alter the geometric designs of roadways from all the 50 States.

HSM (2010) also indicates that the cross Sectional method may be appropriate when implementing a countermeasure(s) on a roadway which before crash information is missing or cannot be obtained. For example, Zhao et al. (2008) studied the safety effect of four-left side off ramps on freeways in Tampa Bay, Florida. In this study, the crash records of the four-left off ramps were used to also analyze four-right off ramps based on the geometric similarities between the two sets of ramps.

2.7.3 Developing of CMFs

Given the characteristics associated with crash data, practitioners have proposed substantial analysis tools, statistical methods and models for analyzing crash data. However, crash data have been known to have issues (Sando and Mohr, 2011) such as (1) over dispersion, (2) under-dispersion, (3) time-varying explanatory variables, (4) temporal and spatial correlation, (5) low sample mean and small sample size, (6) injury severity and crash type correlation, (7) under reporting, (8) omitted variables bias, (9) endogenous variables, (10) functional form and (11) fixed parameters (Lord and Mannering, 2010). Using such data could lead to erroneous results or conclusions if the wright statistical tool or method is not used (Lord and Mannering, 2010).

Several methods have been applied over the years to minimize crash data and methodological issues associated with crash frequency. The most common and recent methods are Safety Performance Functions (SPFs), Bayesian Models and the HSM (2010) statistical procedure of deriving CMFs.

2.7.4 Safety Performance Functions (SPFs)

A safety performance function is a mathematical relationship (model) between frequency of crashes by severity and the most significant causal factors of crashes for a specific type of road (Garber et al. 2010). The commonly used SPFs are developed using Poisson and Negative Binomial regression models (Stamatiadis et al. 2011). SPFs can be developed for either homogeneous or non-homogeneous segments. For homogeneous segments, SPFs use CMFs to estimate the safety effect of any operation or geometric variations (Mbatta, 2011).

Before a model is selected, several tests are conducted to establish its acceptability. These tests assist to verify the underlying distribution assumptions. As stated by Vogt and Bared (1997), the three most important tests for an acceptable model are: (1) the covariant for each estimated coefficient has to be statistically significant. This is so that one should be able to reject the null hypothesis that the coefficient is zero, (2) Engineering and instinctive judgments should be able to confirm the validity and practicability of the sign and rough importance of each estimated coefficient and (3) Goodness-of-fit measures and statistics such as R-Squared (the determinant coefficient), the deviance and Pearson chi-square should indicate that the variables have descriptive and predictive power. A value near 0.0 suggests a lack of correlation and a value of 1.0 suggests that the model estimates are in perfect agreement with the observed data (Famoye et al. 2001).

2.7.4.1 Poisson Regression Models

Because crash frequency data are non-negative integers, most studies have used Poisson regression models as a starting point in their safety effect analysis. In this type of model, the probability of a site entity (i) such as lane width or number of lanes having y_i crashes per given time period where y_i is a non-negative integer (Lord and Mannering, 2010) is given by equation 2.3 as shown below.

$$P_{(y_i)} = \frac{EXP(-\lambda_i)\left(\lambda_i{}^{yi}\right)}{yi\,!} \tag{2.3}$$

Where: $P_{\gamma i}$ = expected number of crashes per time period

λ_i = Poisson parameter for study site

y_i = crashes per time period

Although Poisson regression has been the starting point for crash data analysis, practitioners have found crash data characteristics that make Poisson regression problematic. In particular, Poisson Models cannot handle evaluating the crash counts from different sets of roads, or from the same set of roads but over different time periods, the distribution of the observed counts is often over-dispersed, in that the crash count variance is larger than the crash count mean (Lord et al. 2005). Poisson repression models are also adversely affected by low sample means and can produce biased results in small samples this is because Poisson models assume that the sample mean is equivalent to the sample variance.

2.7.4.2 Negative binomial Regression Models

Negative binomial models are introduced to overcome the problem of over-dispersion in Poisson regression models. The negative binomial model is derived by rewriting the Poisson parameter for each observation i as

$$P(y_i) = \left(\frac{\Gamma\left(y_i + \frac{1}{k}\right)}{y_i!\,\Gamma\left(\frac{1}{k}\right)}\right) * \left(\frac{k\mu_i}{1 + k\mu_i}\right)^{y_i} * \left(\frac{1}{1 + k\mu_i}\right)^{\frac{1}{k}} \tag{2.4}$$

Where: Γ = is the gamma function

μ = is the negative binomial distribution mean

k = is the dispersion parameter

The negative binomial is probably the most used crash frequency modeling methodology, however, it has limitations: (1) this model cannot handle under-dispersed data and (2) it has estimation problems when the data is characterized by low sample mean value and small sample sizes (Mbatti, 2011).

2.7.4.3 Zero Inflated Poisson and Zero Inflated Negative Binomial Models

One important characteristic about crash data is that some may contain large amounts of zeros and a long or heavy tail which results to highly dispersed data (Geedipally at al. 2011). This means that at the study site, the "no crashes observed" cases are so large that normal Poisson and negative binomial regression models cannot efficiently be used (Lord and Geedipally, 2011). To analyze such data, zero inflated Poisson and zero inflated negative binomial models are adopted, because they assume the mean can never be zero (Li et al. 2011).

2.7.5 Bayes Model

Another model used to predict the expected number of crashes at a given site is Empirical Bayes (EB). This model statistically predicts the number of crashes at a given site during the after period before treatment has been done (Qin et al. 2006). The EB before-after model is widely used because it has the ability to account for regression-to-mean (RTM) and traffic volume changes (Persuade et al. 2001) that are usually associated with crash data. EB models estimate estimates the expected number of crashes for a given time period as

$$m = w_1(x) + w_2(P) \tag{2.5}$$

Where: m = expected number of crashes

w_1 = Estimated mean

w_2 = Estimated variance

x = count of crashes for a given period of time

P = annual number of crashes

Where: $w_1 = {}^{P}/_{(K + nP)}$

$w_2 = {}^{K}/_{(K + nP)}$

$K = {}^{P^2}/_{[Var(P)]}$

n = number of time periods being evaluated

2.8 Crash Severity

Although several studies have been conducted to explore the factors that influence crash severity, few studies have explored the relationship between severity and roadway operation or level of service; no study has quantified this relationship. Two of the few studies that have addressed the influence of traffic volume and severity are Christoforou et al. (2010) and Martin (2002).

Christoforou et al. (2010) investigated the influence of speed and traffic volume on the injury level sustained by the vehicle occupants involved in road accidents on the A4-A86 motorway in Paris, France using Probit regression models. The study found that there is a significant relationship between severity and traffic volume and speed. Specifically, for lower traffic volumes, the probability of more severe accidents is significantly higher than for higher volumes. Christoforou et al. (2010) suggested that this finding verifies the assumption that under free flow roadway conditions drivers tend to travel at higher speeds and, therefore, increases the chance of higher severity levels. Christoforou et al.'s (2010) findings are similar to those of Martin (2002).

Martin's (2002) study investigated the relationship between crash severity and hourly traffic volumes based on two years of observations made on 2000 kilometers of French interurban motorways. Martin (2002) used regression models and showed that severity is greater at night and when hourly traffic is light.

2.9 Summary

Improving roadway safety is considered to be a very important task that saves both lives and available resources. To continue improving roadway safety, States and national transportation agencies have been and continue to research on the factors that contribute to roadway crashes.

As shown in this chapter, there has been vast research and findings on this topic and as a result, practitioners now have many resources available to help them understand safety improvements and decide which ones to implement. Among available resources, are CMFs. CMFs are effective in estimating the expected number of crashes after a given countermeasure is implemented. Previous studies mentioned in this chapter; indicate that generally only geometric elements and AADT have been considered in CMF development. However, as indicated by other studies, operational elements such as v/c ratio, capacity, density, and LOS do affect safety, but the relationships are not yet clear. Therefore, there is need to understand how operation performance of a roadway impacts safety and a result, this dissertation study will investigate the safety effect of v/c ratios and densities developed from hourly volumes.

CHAPTER 3

DATA ACQUISITION AND METHODOLOGY

As stated in Chapter 1, the main objective of this study was to establish the relationship between operational elements and safety on urban freeway segments using CMFs. Accordingly, the data collection process began by identifying urban freeways in the State of New Jersey. Data were collected on the same or similar segments under different operational conditions. This study focused only on data required for developing CMFs using the before-after method as described in the HSM (2010). The procedure of collecting and preparing the data for analysis is discussed in Section 3.1 and the methodologies used to analyze the prepared data are discussed in Section.

3.1 Data Collection Procedures

This study's first priority was to identify urban freeways because urban freeways are the only facilities that provide pure uninterrupted flow (Roess et al. 2011) and are characterized with base or ideal conditions that there are no heavy vehicles in traffic stream and the driver population is dominated by regular or familiar users of the facility (HCM, 2010). In addition, the measure of effectiveness used to define levels of service (LOS) is density (HCM, 2010).

In order to simplify the discussion on data collection procedures, the tasks involved were highlighted and presented in Figure 3.1. Sub-Sections 3.1.1 through 3.1.6 give details on each of the tasks attempted to prepare the data for analysis.

3.1.1 Identifying Urban Freeways

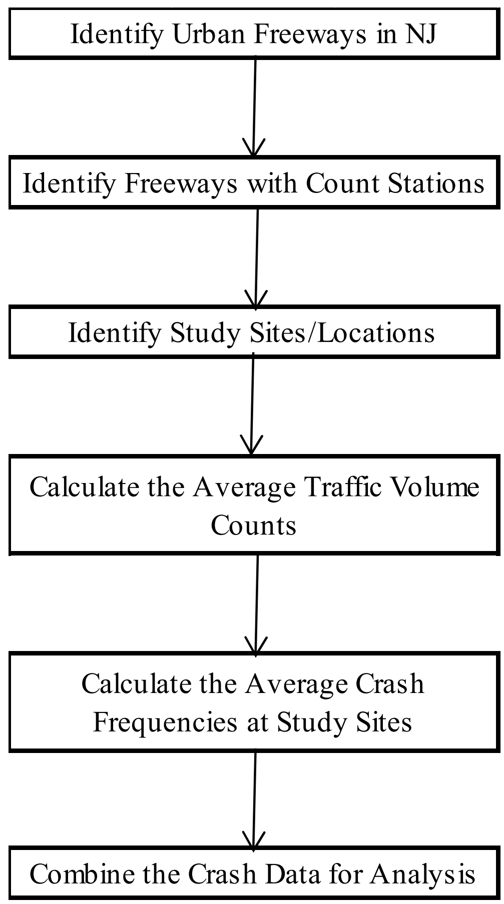

Figure 3.1 Tasks involved in collecting and preparation of crash data for Analysis

As Figure 3.1 shows, the first step of this procedure was to identify and select only urban freeways from all the roadways in the State of New Jersey. To attempt this task, NJDOT's straight-line roadway database was used to obtain information on all the roadways in the State. From the database, urban freeways were identified and selected. In the database, NJDOT has classified all the roadways by functional class with urban freeways are classified as functional class 12.

Those roadways not classified as urban freeways were not considered for this study. With this information, the next step was to identify the roadways that had count stations as shown in Figure 3.1 and discussed next in Sub-Section 3.1.2.

3.1.2 Identifying Urban Freeways with Traffic Volume Count Stations

Once all urban freeways were identified, this study referred to NJDOT's Roadway Information and Traffic Counts database to select the roadways that had at least four count stations so that this study would have at least four study sites on each roadway. Since each station has twenty-four hours of traffic volume, then there would be ninety-six total hours of study for each year on a given roadway. Accordingly, because four years of crash data were considered in this study, then the ninety-six hourly volumes for four years would result in three hundred and eighty-four hourly volume and hourly crashes for each roadway. The hourly volumes were the same for each year.

Three urban freeways, US Route 1, NJ Route 3, and NJ Route 21 had count stations that met these conditions. Those roadways that did not have at least four count stations were not considered for this study. The next step was to identify the specific study sites or locations as shown in Figure 3.1 and discussed next in Sub-Section 3.1.3.

3.1.3 Identifying the Study Sites

Still referring on NJDOT's Roadway Information and Traffic Counts database, the third step involved identifying and selecting roadways that had the count stations. The count stations had to be at least one mile apart because each segment of study was one mile long, therefore if they had to be less than a mile apart, some crash data would have been analyzed twice hence giving biased analysis results. Successive count stations had also to be not more than three miles away from each other to reduce chances of having significant change in traffic volumes and drivers. At the count stations, NJDOT collected hourly traffic volumes for an entire day for two or three days, the average traffic volume for each hour was then calculated for analysis.

The mileposts for the selected study sites for each freeway were summarized and presented in Table 3.1. This table also shows the study site's characteristics, that is the posted speed (mi/hr.), number of lanes, lane width (ft.), shoulder width (ft.), and median width (ft.). Five sites were selected on US 1 and NJ 21. Four sites were selected on NJ 3. The next step was to calculate the average hourly traffic volume count for each hour and year at all the study locations as shown in Figure 3.1 and discussed next in Sub-Section 3.1.4.

Table 3.1 Selected Study Site Characteristics

Roadway	Mile Post	Posted Spd. (mph)	Number of Lanes	Lane Width (ft.)	Shoulder Width (ft.)	Median Width (ft.)
US Route 1	46.00	50	3	8	6	8
	47.20	50	2	12	3	8
	48.20	50	2	12	8	4
	50.50	50	2	12	3	4
	52.29	45	2	12	0	4
NJ Route 3	0.80	55	3	12	12	26
	3.10	55	3	12	12	26
	5.50	55	3	12	12	8
	9.50	50	3	12	15	6
NJ Route 21	4.40	45	3	11	0	6
	5.00	50	3	12	12	6
	7.10	55	3	12	12	8
	9.70	55	3	12	12	8
	12.40	55	3	12	12	10

3.1.4 Calculating the Average Hourly Traffic Volumes

As earlier stated in Sub-Section 3.1.3, the traffic volumes are counted for about two or three days. In this step, to get the average traffic volume for each hour of the day, the total traffic volumes for each hour were divided by the number of days NJDOT collected the data. This exercise was done for all the study sites for each year of study. Having found the average traffic volumes for each hour, the next step as shown in Figure 3.1 was to get the total crash frequency for each hour of study.

3.1.5 Getting the Crash Frequency for Each Hour of Study

The total crash frequencies for each study year were collected from NJDOT's crash data database for all the three roadways of study. The total number of crashes for each hour was then determined. Thereafter the data shown in Table 3.1and the determined hourly crash frequencies were prepared for analysis as explained next in Sub-Section 3.1.6.

3.1.6 Preparing Collected Data for Analysis

This step involved four tasks: (1) assigning each crash frequency total to the respective hourly traffic volume; (2) determining the physical/geometric features; (3) calculating density; and (4) finally assigning each observation point or study hour with its respective LOS. Since each site had 96 hours of study and 14 sites were considered, a total of 1344 hours were analyzed by this study. The density calculation and assigning of LOS were in accordance to the procedures recommended by the HCM (2010). The final step for this task was to combine all the hours of data and results from all study sites with the same

LOS. This allowed the data to be analyzed according to their respective LOS for all the study sites using the methodologies explained in Section 3.2, which is discussed next.

3.2 Methodology

Crash data analysis is generally the most effective and frequently used resource for assessing the safety performance of any given freeway (Abdel-Aty and Pande, 2007). Crash occurrences can be viewed as a result of the interaction of several variables including road geometry, driver characteristics and driver behavior, operation conditions like operating speed, volume, and lastly environmental conditions (Christoforou et al. 2011). In the past, practitioners have analyzed the relationships between these variables and crash frequency to estimate the expected safety performance of a given freeway (Garber and Ehrhart, 2000) or to determine the expected severity levels of these crashes.

Crash data is analyzed to determine crash frequency or severity levels using different modeling techniques, ranging from conventional regression analysis (Garber and Ehrhart, 2000) to simulation models (Abdel-Aty and Pande, 2005; Abdelwahab and Abdel-Aty, 2002). In this study, the focus was on establishing the relationship between LOSs generated from hourly volumes, and crash frequency using both SPFs and Empirical Bayes before-after methodology to generate related CMFs. The subsequent Sub-Sections will outline the two commonly preferred SPFs and the role of Empirical Bayesian relationships adopted by this dissertation in an effort to show that operational elements have some influence on safety.

3.2.1 Safety Performance Functions (SPFs) Modeling

Crashes are intrinsic as vehicles traverse roadway segments or intersections (Tegge et al. 2010). Transportation practitioners rely on SPFs to predict expected crash frequencies for various physical, operational, human, or environmental factors. SPFs, also referred to as predictive models (Zhong et al. 2008), are used to show mathematical relationships between crash frequency and the most significant casual factor(s) (Garber et al. 2010). Therefore, developing germane SPFs is an important task that requires investigating and quantitatively identifying the relationships between safety and certain factors with the proper statistical models. So generally, there are two types of data required in developing SPFs, crash data and roadway attributes data (Zhong et al. 2008). These roadway attributes are used in SPF development used by the HSM (2010) are indicated in Appendix A as discussed in Chapter 2.

As discussed, SPFs are statistical models used to relate crash frequencies to given roadway attributes. In this dissertation, crash frequencies are compared using regression analysis to determine which attribute(s) produce a significant cause-and-effect relationship (Tegge et al. 2010). Whether or not an attribute is significant, is based on a user-identified level-of-significance (α). The level of significance, measures the plausibility of the null hypothesis (Navidi, 2008), and usually ranges from 0.01 to 0.10 (Hauer, 1996). A smaller α shows, it is more difficult to declare an attribute significant. Crashes are rare but are a serious subject, therefore a larger α is usually adopted to include more attributes in the model (Tegge et al. 2010) and as a result the most suitable α selected is usually 0.10.

The predictor variables used in this research were selected using Pearson's correlation matrices at a correlation coefficient range of -0.03 and +0.03. This step was important because it prevented selecting predictor variables with multicollinearity. This is the situation where two or more predictor variables have strong correlation. In some multiple regression exercises where two or more predictor variables show strong correlation, the results may show some inconsistency. For example, the F-Test may show that the data fits well even though none of the predictor variables influences the dependent variable significantly (Kutner et al. 2004). In other words, multicollinearity misleadingly inflates the standard errors which in return makes some variables to be statistically insignificant while they actually are significant and vice versa.

A literature search conducted in this dissertation indicates that several methods have been used by previous studies to develop SPFs, with "*some more accurate than others*" according to Tegge et. al (2010). According to Harwood and Bauer (2000), the most accurate and commonly used SPF statistical models are lognormal and loglinear regression distributions.

3.2.1.1 Lognormal Regression Models

Lognormal regression models are usually used when the distribution of data is skewed (Tegge et al. 2010) because these models assume that the natural log of crash frequencies is normally distributed with the mean, μ and variance, σ^2. This model is usually effective when the data is naturally non-negative and the mean is relatively large. This type of distribution is preferred when analyzing crash data at signalized intersections with high volumes (Tegge et al. 2010). The model to estimate predicted crash frequency for the I^{th} intersection with q parameters, $X_{i1}, X_{i2}, X_{i3} \ldots X_{iq}$, takes on the following form:

$$\log(\mu) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{3.1}$$

Where β_0, β_1 … β_q are constant coefficients that need to be estimated. The assumption here is that the crash frequency is normally distributed and the linear regression coefficients are estimated using ordinary least-squared method (Harwood and Bauer, 2000).

3.2.1.2 Loglinear Regression Models

A loglinear regression model is a specific linear model where by the relationship between two or more variables is analyzed by taking the logarithm of the dependent variable (Tegge et al. 2010). There are several loglinear regression models but the two most accurate and common types are Poisson and negative binomial regression models (Tegge et al. 2010). These two main loglinear regression models are next discussed at length.

Poisson Regression Models

Poisson distribution is a distinct distribution that expresses the probability of a specific number of rare events occurring in a given amount of time, the events occur with a known probability and are independent of the previous event (Hogg and Tanis, 2001). The Poisson distribution is a limiting case of the binomial distribution because it assumes that the variance is equal or approximately equal to the mean. The probability of y_i events occurring at a certain time interval follows the form:

$$P_{(y_i)} = \frac{EXP(-\lambda_i)(\lambda_i{}^{yi})}{yi\,!} \tag{3.2}$$

Where: $P_{\gamma i}$ = predicted number of crashes per time period

λ_i = Poisson parameter for study site

y_i = crashes per time period

The linear model for the I^{th} roadway segment with q parameters $X_{i1}, X_{i2}, X_{i3} \ldots X_{iq}$, and regression coefficients β_0, $\beta_1 \ldots \beta_q$ takes on the following form:

$$\log(\mu) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{3.3}$$

The main difference between this model and the lognormal model is that here the model assumes a Poisson distribution.

Negative Binomial Regression Models

Like the Poisson distribution model, the negative binomial model describes the occurrence of random and rare events. The main difference between these two loglinear models is that unlike the Poisson distribution where it assumes that the variance is equal to the mean, the negative binomial distribution compensates for situations where the variance $(\mu + \mu^2/k)$ is larger than the mean, also referred to as overdispersion. The negative binomial model utilizes the following distribution function:

$$P(y_i) = \left(\frac{\Gamma\left(y_i + \frac{1}{k}\right)}{y_i!\,\Gamma\left(\frac{1}{k}\right)}\right) * \left(\frac{k\mu_i}{1 + k\mu_i}\right)^{y_i} * \left(\frac{1}{1 + k\mu_i}\right)^{\frac{1}{k}} \tag{3.4}$$

Where: Γ = is the gamma function

μ = is the negative binomial distribution mean

k = is the dispersion parameter

Here, as the overdispersion parameter nears zero, that is less variation, the distribution also nears a Poisson distribution (Tegge et al. 2010). The linear model for the I^{th} roadway segment with q parameters $X_{i1}, X_{i2}, X_{i3} \ldots X_{iq}$, and regression coefficients $\beta_0, \beta_1 \ldots \beta_q$ are takes on the following form:

$$\log(\mu) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{3.5}$$

The expression in *equation 3.5* above can be written to follow the form shown *in equation 3.6* below.

$$\mu_i = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq}) \tag{3.6}$$

The linear model assumes that the crash frequency follows a negative binomial distribution with parameters α and *k* as previously described (Harwood and Bauer, 2000). The form shown in *equation 3.6* can be used when the traffic volumes are constant or varying from the before to the after periods. In this study, the hourly traffic volume varied from the before to the after periods.

SPFs are used by most studies and the HSM (2010) to predict the crash frequency on a segment or at an intersection of a given roadway. However, SPFs do not account for regression-to-mean effect (RTM) associated with crash data. RTM is the tendency for crash data to regress back to the mean and therefore, a more involved analysis must be conducted to account for this effect to determine the actual or expected safety of a given location (Tegge et al. 2010).

To account for RTM, this dissertation adopted Empirical Bayesian relationships to compare both the observed crashes and the predicted crash frequencies found by the SPFs to estimate the expected crashes. EB relationships are discussed in the next Section.

3.2.2 Empirical Bayesian Relationship

As stated in the previous Section, SPFs were only part of the overall roadway safety evaluation process for this study. In addition, the observed crashes of a given segment or intersection need to be accounted for while determining the safety of a given roadway segment or intersection. However, observed crashes have occasionally been known to be misleading due to the regression-to-the-mean, RTM occurrence (Hauer, 2002). For example, a site may have high crashes at a given period and low crashes the next period without any safety implementations. Alternatively, a high-risk site may experience a period of randomly low crashes and therefore overlooked during safety evaluation. In observational studies,

> *"two methods, one simpler and one somewhat more complex, are preferred to derive a before-after study. The comparison group method being the simpler of the two, while EB method is more complex, but also more robust."* Gross et. al (2010).

EB models increase the precision of safety evaluation by estimating a weighted combination of the crash frequency expected in the before period without treatment and the observed crash frequency. The weights are determined as follows:

$$w = \frac{1}{1 + k\sum_{n=1}^{N} P_n} \tag{3.7}$$

Where: k = over dispersion parameter

P_n = predicted crash frequency for a given roadway in a period time n.

From the Empirical Bayes procedure, the weight factor is then applied to the predicted and observed number of crashes to determine the estimated number of crashes which is computed as follows:

$$N_{expected,T,B} = w.N_{predicted,T,B} + (1 - w)N_{observed,T,B} \tag{3.8}$$

Where: $N_{expected,T,B}$ = EB estimate of the expected crash frequency without treatment

$N_{predicted,T,B}$ = predicted crash frequency estimated by the SPF in the before period

$N_{observed,T,B}$ = observed crash frequency in the before period for the treatment group

Tegge et al. (2010) indicates that for this analysis, the more observations made, the smaller the weight factor, which makes the estimated crash frequency weighted more towards the observed crash frequency. This is consistent with the purpose of using Empirical Bayesian relationships that is to increase the precision by accounting for RTM. In other words as the number of observations increases, the RTM effect is not as severe. The RTM effect is illustrated in Figure 3.2.

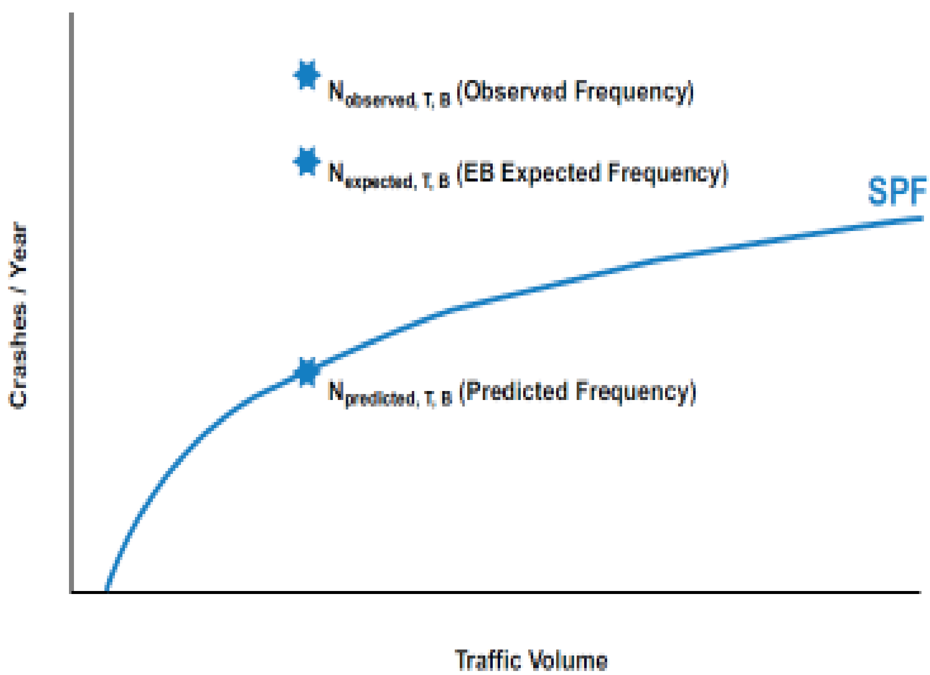

Figure 3.2 Illustration of the RTM effect on SPFs and how they are corrected using EB models.
Source: A Guide to Developing Crash Modification Factors by Gross et al. (2010)

Figure 3.2 shows how the SPF predicted crash frequency is weighted with the observed crash frequency to estimate Empirical Bayes expected crash frequency which fall in-between the observed and SPF predicted crash frequencies. The expected crash frequency found using Empirical Bayes relationships, is then used in the final analysis to develop the Crash Modification Factors, CMFs. The empirical Bayes estimate of the expected crash frequency without treatment, $N_{expected,T,B}$ is computed as shown in *equation 3.8*

3.2.3 Crash Modification Factors

The resultant $N_{expected,T,B}$ is then adjusted to account for the RTM effect. The adjusted value of the EB estimate $N_{expected,T,A}$, is the number of crashes in the after period without treatment and is estimated as follows:

$$N_{expected,T,A} = N_{expected,T,B} * \left(\frac{N_{predicted,T,A}}{N_{predicted,T,B}}\right) \tag{3.9}$$

Where: $N_{expected,T,A}$ = unadjusted EB estimate

$N_{predicted,T,B}$ = predicted number of crashes estimated by the SPF in the before period

$N_{predicted,T,A}$ = predicted number of crashes estimated by the SPF in the before period

The variance of $N_{expected,T,A}$ is estimated from the SPF predicted number of crashes in the before period as follows:

$$Var(N_{expected,T,A}) = N_{expected,T,A} * \left(\frac{N_{predicted,T,A}}{N_{predicted,T,B}}\right) * (1 - w) \tag{3.10}$$

The adjusted value of the EB estimate $N_{expected,T,A}$, is used to derive CMFs using the attributes shown in Table 3.1. CMFs and the variance are estimated as follows:

$$CMF = \frac{\left(\frac{N_{observed,T,A}}{N_{expected,T,A}}\right)}{1 + \left(\frac{Var}{N_{expected,T,A}{}^{2}}\right)} \tag{3.11}$$

$$Var\,(CMF) = \frac{CMF^2 * \left(\left(\frac{1}{N_{observed,T,A}}\right) + \left(\frac{Var}{N_{expected,T,A}{}^{2}}\right)\right)}{\left(1 + \frac{Var}{N_{expected,T,A}{}^{2}}\right)^2} \tag{3.12}$$

Where: $N_{observed,T,B}$ = observed crash frequency in the before period for the treatment group.

$N_{observed,T,A}$ = observed crash frequency in the after period for the treatment group.

$N_{predicted,T,B}$ = predicted crash frequency (sum of the SPF estimates) in the before period.

$N_{predicted,T,A}$ = predicted crash frequency (sum of the SPF estimates) in the after period.

Table 3.2 Summary of Notation for EB method

Time Period	Treatment Group (Observed Crashes)	SPF Prediction Crashes (SPF Developed from Reference Group)
Before	$N_{observed,T,B}$	$N_{predicted,T,B}$
After	$N_{observed,T,A}$	$N_{predicted,T,B}$

3.3 Example of the Empirical Bayes Method

Table 3.3 is similar to Table 3.2, which shows information to support calculations using the empirical Bayes method. This example uses data collected from this study when the LOS changed from A to B. The study found the weight was 0.025. The calculations of $N_{expected,T,A}$ and Var$(N_{expected,T,A})$ would be computed for each LOS and for each location individually and then summed to use in the estimation of the CMF and its standard error.

Table 3.3 Data for EB Before-After Study Example

Time Period	Treatment Group (Observed Crashes)	SPF Prediction Crashes (SPF Developed from Reference Group)
Before (LOS A)	408	Sum for 14 sites = 174.811
After (LOS B)	297	Sum for 14 sites = 191.368

The empirical Bayes estimate, $N_{expected,T,B}$, is calculated as:
$$\boldsymbol{N_{expected,T,B} = 0.025*174.811 + (1-0.025)*408 = 402.193}$$

The ratio of after period SPF estimates to before period SPF estimates is now:
$$\boldsymbol{N_{predicted,T,A} / N_{predicted,T,B} = 191.368/174.811 = 1.095}$$

The expected number of crashes in the after period in the treatment group that would have occurred without treatment ($N_{expected,T,A}$) is:
$N_{expected,T,A}$ = 402.193*1.095 = 440.286

The variance of $N_{expected,T,A}$ is estimated as:
Var ($N_{expected,T,A}$) = 440.286*1.095*(1-0.025) = 470.061

CMF = (297/440.286) / (1+ (470.061/440.286^2)) = <u>0.673</u>
Variance = (0.673^2((1/297) + (470.061/440.286^2)) / (1+470.061/440.286^2)^2)= 0.003

Taking the square root of the variance, the standard error of the CMF is 0.055.
The 95% confidence interval is 0.673 ± 1.96*0.0.55 = 0.565 to 0.781.

The example in Section 3.3 shows how the Empirical Bayes methodology is used to analyze crash data and the development of CMFs when the level of service deteriorated from LOS A to B. Using the same approach discussed in this Chapter, negative binomial, and EB models were used to determine the predicted and expected crash frequency values when the levels of service deteriorated from B to C, C to D, D to E, and E to F respectively. With all the observed, predicted, and expected crash frequencies for each LOS, CMFs were determined as discussed and presented in the Chapter 4 next.

CHAPTER 4

RESULTS AND DISCUSSION

Using the methodologies and procedures discussed in Chapter 3, the crash data discussed in Chapter 2 was analyzed to determine the relationship between crash frequency and operational elements, specifically when LOS deteriorated from A to B, B to C, C to D, D to E, and E to F. The main tasks in this Chapter were: (1) to determine variables to be used for analysis using; (2) use those variables to determine the predicted crashes, (3) use both the observed and predicted crash frequencies to determine expected crash frequencies and finally (4) used all observed, predicted and expected crash frequencies to determine the CMFs for each change in LOS. These tasks are further discussed in Sections 4.1, 4.2 and 4.3 of this Chapter.

4.1 Determining Predicted Crash Frequencies

Several statistical studies have shown that roadway variables and crash frequencies have a non-linear relationship, and therefore to show how the roadway variables influence crash frequency at a given site, SPFs were developed (Tegge et al. 2010). To determine which variables to consider for SPF development, this study used Pearson correlation matrices as discussed next in Sub-Section 4.1.1.

4.1.1 Determining Possible Models using Correlation Matrices

In this exercise, Pearson correlation matrices for all levels of service were developed using the SPSS 16.0 software. As discussed earlier, the purpose of this exercise was to identify the variables that were significant and uncorrelated and then to use them in the same SPF model since their resultant coefficients were unbiased.

Therefore, Pearson correlation matrices were determined to measure the strength of linear dependence between the individual variables. The Pearson correlation coefficient is usually denoted as r and is a value between +1 and -1. The lowest value that r can be is 0, this would show zero correlation or no relationship between the two given variables. The highest value that r can have is 1.00, this would show a perfect correlation or strong relationship between the two given variables and that is the two variables depend on each other. The values can either be positive or negative. A positive value indicates that an increase in one variable corresponds to an increase in the other variable. A negative value indicates that an increase in one variable corresponds to a decrease in the other variable. To select a model(s) with the same predictor variables and response variable at all levels of service, this study used data from all the levels of service in developing the correlation matrices. The Pearson correlation matrix for this study is presented in Table 4.1.

Table 4.1 Pearson Correlation Matrix

	Traffic Vol.	Posted Spd (mi/hr.)	Number of Lanes	Lane Width (ft)	Shoulder Width (ft)	Median Width (ft)	Density (veh/mi)
Traffic Vol.	1	0.211	0.401	-0.209	0.257	0.362	0.622
		0.000	0.000	0.000	0.000	0.000	0.000
Posted Spd (mi/hr.)	0.211	1	0.484	0.238	0.788	0.559	0.148
	0.000		0.000	0.000	0.000	0.000	0.000
Number of Lanes	0.401	0.484	1	-0.217	0.642	0.401	0.155
	0.000	0.000		0.000	0.000	0.000	0.000
Lane Width (ft)	-0.209	0.238	-0.217	1	0.257	0.090	-0.074
	0.000	0.000	0.000		0.000	0.001	0.007
Shoulder Width (ft)	0.257	0.788	0.642	0.257	1	0.390	0.116
	0.000	0.000	0.000	0.000		0.000	0.000
Median Width (ft)	0.362	0.559	0.401	0.090	0.390	1	0.250
	0.000	0.000	0.000	0.001	0.000		0.000
Density (veh/mi)	0.622	0.148	0.155	-0.074	0.116	0.250	1
	0.000	0.000	0.000	0.007	0.000	0.000	

The Pearson correlation matrix in Table 4.1 was used to select the model(s) to be used in SPF analysis. Model(s) selection followed the criteria that:

1. The predictor variables had to show no or weak correlation. The strength of relationship is classified by Choudhury (2009), Navidi (2008), and Kiemele et. al (2000) as presented in table 4.2.

Table 4.2 Classification of Correlation Strength

Value of *r*			Strength of Relationship
± 0.5	to	± 1.00	Strong
± 0.3	to	± 0.49	Moderate
± 0.1	to	± 0.29	weak
0.0	to	± 0.09	None or Very Weak

2. The predictor variables of the selected model(s) had to be statistically significant. This study used a 0.1 significance level.

3. The selected model(s) had to have traffic volume and density among the predictor variables.

Using the three criteria discussed, the following three models were selected for SPF analysis,:

Model I: Traffic Volume, Posted Speed, Lane Width, Density

Model II: Traffic Volume, Shoulder Width, Lane Width, Density

Model III: Traffic Volume, Posted speed, Lane Width, Number of Lanes, Shoulder Width, Median width, Density

SPFs for all the three models were developed and a goodness-of-fit test performed on all the models to measure how well each model explained the crash data. This task is discussed further next in Sub-Section 4.1.2.

4.1.2 Measuring Goodness-of-Fit for each Model

As is the case for all regression models, this study then tested the model(s) to determine the goodness-of-fit it was accepted for analysis. Generally, the goodness-of-fit test is performed using small-is-better criteria, the regression model with the smallest values is usually adopted over that with larger values. In this study negative binomial regression models were adopted since at all levels of service the data showed overdispersion – the dispersion coefficients were all larger than zero. A goodness-of-fit test was performed using deviance, Pearson chi-square and their respective degree of freedom to check how well the data fit the model. Before discussing the technical background for this criterion, it is important to first define deviance and the Pearson chi-square parameters.

Deviance is a measure of degree of fit defined as two times the difference of the log-likelihood for the maximum achievable model and the log-likelihood under the fitted model. The Pearson chi-square is a test that establishes whether or not an observed frequency distribution differs from a theoretical distribution. The Pearson chi-square is the squared difference between the observed and the predicted values divided by the variance of the predicted value summed over all observations in the model. Both Deviance and Pearson Chi-Square are calculated as shown in *Equations 4.1 and 4.2,*.respectively.

$$Deviance = \sum_{i=1}^{n} 2(y_i log\frac{y_i}{y_i} - (y_i - \hat{y}_i)) \tag{4.1}$$

$$Pearson\ Chi - Square = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{\hat{y}_i} \tag{4.2}$$

Where: $\hat{y}_i$ is the predicted value of y_i for both cases.

"The deviance has an approximate chi-square distribution with n-p degrees of freedom, where n is the number of observations, p is the number of predictor variables (including the intercept), and the expected value of a chi-square random variable is equal to the degrees of freedom" (SPSS 16.0 Brief Guide, 2007).

It is then accepted that a model fits the data well, when the ratio of the deviance to degree of freedom is approximately about one. A large ratio value may indicate model misspecification or an over-dispersed response variable; a less than one ratio also indicates model misspecification or an under-dispersed response variable (SPSS 16.0 Brief Guide, 2007). This study's goodness-of-fit values for each of the three models are presented in Tables 4.3a through 4.3c.

Table 4.3a Test for Goodness-of-Fit for Model I

LOS	Parameter	Value	df	Value/df
	Deviance	463.367	432	1.07
A	Pearson Chi-Square	458.562	432	1.06
	Sig.	0.000		
	Deviance	268.983	245	1.10
B	Pearson Chi-Square	276.635	245	1.13
	Sig.	0.001		
	Deviance	296.790	261	1.14
C	Pearson Chi-Square	259.209	261	0.99
	Sig.	0.000		
	Deviance	119.684	99	1.21
D	Pearson Chi-Square	100.920	99	1.02
	Sig.	0.290		
	Deviance	108.707	100	1.09
E	Pearson Chi-Square	101.093	100	1.01
	Sig.	0.000		
	Deviance	192.481	171	1.13
F	Pearson Chi-Square	179.641	171	1.05
	Sig.	0.000		

Table 4.3a presents the results of the goodness-of-fit test for Model I. At all levels of service, the *Value/df* is approximately one. This shows that the crash data fits well in this model. However, Model I was not selected for use as the model was not statistically significant for LOS D at significance level of 0.1.

Table 4.3b Test for Goodness-of-Fit for Model II

LOS	Parameter	Value	df	Value/df
A	Deviance	464.315	432	1.08
	Pearson Chi-Square	463.551	432	1.07
	Sig.	0.000		
B	Deviance	268.058	245	1.09
	Pearson Chi-Square	271.743	245	1.11
	Sig.	0.000		
C	Deviance	296.926	261	1.14
	Pearson Chi-Square	274.251	261	1.05
	Sig.	0.000		
D	Deviance	119.632	99	1.21
	Pearson Chi-Square	100.252	99	1.01
	Sig.	0.137		
E	Deviance	109.816	100	1.10
	Pearson Chi-Square	108.775	100	1.09
	Sig.	0.000		
F	Deviance	191.970	171	1.12
	Pearson Chi-Square	178.742	171	1.05
	Sig.	0.000		

Table 4.3b presents the results of the goodness-of-fit test for Model II. At all levels of service, the *Value/df* is approximately one. This shows that the crash data fits well in this model. However, Model II was not selected for use as the model was not statistically significant at LOS D at a significance level of 0.1.

Table 4.3c Test for Goodness-of-Fit for Model III

LOS	Parameter	Value	df	Value/df
A	Deviance	461.895	429	1.08
	Pearson Chi-Square	452.824	429	1.06
	Sig.	0.000		
B	Deviance	269.488	242	1.11
	Pearson Chi-Square	267.477	242	1.11
	Sig.	0.000		
C	Deviance	296.305	258	1.15
	Pearson Chi-Square	260.485	258	1.01
	Sig.	0.000		
D	Deviance	119.791	96	1.25
	Pearson Chi-Square	101.868	96	1.06
	Sig.	0.008		
E	Deviance	100.927	97	1.04
	Pearson Chi-Square	90.423	97	0.93
	Sig.	0.000		
F	Deviance	201.211	169	1.19
	Pearson Chi-Square	185.531	169	1.10
	Sig.	0.000		

Table 4.3c presents the results of the goodness-of-fit test for Model III. At all levels of service, the *Value/df* is approximately one. This shows that the crash data fits well in this model. This model was selected for use as the Model was statistically significant at a significance level 0.1 at all levels of service.

Tables 4.3a through 4.3c present results of the goodness-of-fit tests showing the values for the deviance and Pearson chi-square for all three models at all levels of services. According to the results, the *value/df* ratios for both deviance the and Pearson chi-square range between 0.93 and 1.25 for all models and all levels of services. Because they are close to one, as discussed in this Sub-Section, the model fits the data well for all the models. However, at a significance level of 0.1, only Model III was statistically significant at all levels of service and therefore was used in the development of SPFs discussed next in Sub-Section 4.1.3. The output of Models I, II, and III were determined using SPSS 16.0 and are presented in Appendix B of this dissertation.

4.1.3 Development of Safety Performance Functions

SPFs were used to predict the crash frequency for each LOS using Model III. The variables include in the model were crash frequency, lane width, posted speed limit, number of Lanes, shoulder width, median width, and density. Negative binomial models were used for the SPF development and the resultant coefficients for each variable presented in Tables 4.4a through 4.4f.

Table 4.4a Negative Binomial Parameter Estimates for LOS A

Parameter	Estimate	Std. Error	Sig.
(Intercept)	-1.949	2.044	0.014
Traffic Volume (Hourly)	0.001	0.001	0.058
Posted Speed Limit (m/h)	-0.070	0.030	0.020
Number of Lanes	0.865	0.301	0.004
Lane Width (ft.)	0.183	0.081	0.025
Shoulder Width (ft.)	-0.043	0.023	0.065
Median Width (ft.)	0.018	0.011	0.089
Density (veh/mi)	-0.092	0.121	0.049
Dispersion (*k*)	0.224	0.091	

Table 4.4b Negative Binomial Parameter Estimates for LOS B

Parameter	Estimate	Std. Error	Sig.
(Intercept)	-5.560	3.141	0.077
Traffic Volume (Hourly)	0.001	0.001	0.102
Posted Speed Limit (m/h)	0.061	0.036	0.086
Number of Lanes	-0.082	0.685	0.051
Lane Width (ft.)	0.225	0.127	0.075
Shoulder Width (ft.)	-0.093	0.029	0.001
Median Width (ft.)	-0.009	0.015	0.548
Density (veh/mi)	-0.147	0.137	0.028
Dispersion (*k*)	0.167	0.087	

Table 4.4c Negative Binomial Parameter Estimates for LOS C

Parameter	Estimate	Std. Error	Sig.
(Intercept)	-2.753	3.031	0.047
Traffic Volume (Hourly)	0.000	0.001	0.097
Posted Speed Limit (m/h)	-0.060	0.030	0.044
Number of Lanes	0.516	0.960	0.059
Lane Width (ft.)	0.307	0.106	0.004
Shoulder Width (ft.)	-0.031	0.021	0.014
Median Width (ft.)	-0.011	0.019	0.554
Density (veh/mi)	0.039	0.107	0.071
Dispersion (*k*)	0.204	0.074	

Table 4.4d Negative Binomial Parameter Estimates for LOS D

Parameter	Estimate	Std. Error	Sig.
(Intercept)	1.528	6.251	0.033
Traffic Volume (Hourly)	0.001	0.002	0.019
Posted Speed Limit (m/h)	-0.036	0.040	0.037
Number of Lanes	-2.051	2.765	0.058
Lane Width (ft.)	0.407	0.234	0.082
Shoulder Width (ft.)	-0.009	0.033	0.045
Median Width (ft.)	0.001	0.018	0.939
Density (veh/mi)	-0.115	0.171	0.053
Dispersion (*k*)	0.243	0.111	

Table 4.4e Negative Binomial Parameter Estimates for LOS E

Parameter	Estimate	Std. Error	Sig.
(Intercept)	-12.984	6.598	0.034
Traffic Volume (Hourly)	0.001	0.002	0.051
Posted Speed Limit (m/h)	0.198	0.048	0.000
Number of Lanes	-2.567	4.290	0.055
Lane Width (ft.)	0.431	0.192	0.025
Shoulder Width (ft.)	-0.108	0.096	0.025
Median Width (ft.)	-0.049	0.008	0.000
Density (veh/mi)	-0.001	0.156	0.009
Dispersion (*k*)	0.293	.	

Table 4.4f Negative Binomial Parameter Estimates for LOS F

Parameter	Estimate	Std. Error	Sig.
(Intercept)	-11.922	1.084	0.000
Traffic Volume (Hourly)	0.000	0.000	0.000
Posted Speed Limit (m/h)	0.366	0.057	0.000
Number of Lanes	-3.033	0.589	0.000
Lane Width (ft.)	0.177	0.082	0.032
Shoulder Width (ft.)	0.000	.	.
Median Width (ft.)	-0.043	0.005	0.000
Density (veh/mi)	-0.002	0.001	0.002
Dispersion (*k*)	0.099	0.034	

Tables 4.4a through 4.4f show the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for each level of service. Also shown are the dispersion estimates. The intercepts show the estimated number of crashes when all variables are held at zero.

At a significant level of 0.1, all the variables, except for median width, are significant and as a result, median width was not considered in the SPF development in this study. The dispersion coefficients for all LOSs are positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate for use.

The results show that at all levels of service the lane width is positive. The posted speed is positive at LOS B, E, and F and negative at LOS A, C, and D. The number of lanes is positive at all LOS except at LOS C and A. The shoulder width has a negative influence on crash frequency, however as the results show at LOS F, the shoulder width was set to zero because the parameter is redundant meaning at this LOS it is highly correlated with one of the predictor variables at this particular LOS. Density is negative at all LOS except at LOS C. A positive sign indicates that as these positive variables increase, the crash frequency also increases, consequently, a negative sign is an indication that as the variables increase, the crash frequency decreases. The individual SPSS output results are presented in Appendix B of this dissertation.

The coefficients found in Model III were substituted in equation 3.6 in Chapter 3 of this dissertation to determine the predicted crash frequency. However, as explained in Chapter 3, a more rigorous analysis was conducted to determine the safety of each LOS. To achieve this, both the predicted crash frequency found by the SPFs and the observed crash frequencies were combined using Empirical Bayesian relationships to find the expected crash frequencies. Empirical Bayes model was used to,

> *"More precisely estimate the number of crashes (denoted as $N_{expected,T,A}$ in the comparison group method) that would have occurred at an individual treated site in the after period had a treatment not been implemented. Similar to the comparison group method, the effect of the safety treatment is estimated by comparing the sum of the estimates of $N_{expected,T,A}$ for all*

treated sites with the number of crashes actually recorded after treatment." Gross et al. (2010).

The treatment in this case was the degrading of level of service from one level of service to the next level of service, i.e. from LOS A to LOS B. Specifically, the EB before-after model was used to account for the regression-to-mean (RTM) effect usually associated with the crash data. The procedure for calculating the expected crash frequencies is next discussed in Section 4.2.

4.2 Determining Expected Crash Frequencies

As stated in Section 4.1, SPFs are only part of the overall safety evaluation process in this study and the observed crashes need to be accounted for in determining safety. In this Section, the expected crash frequencies are determined using Empirical Bayesian method to increase the precision of safety estimation by accounting for the RTM bias. Substituting in equation 3.7, the weight factors were calculated using the overdispersion coefficients found in SPF modeling. The weight factors were then used to calculate the expected crash frequencies presented in Table 4.5.

Table 4.5 Expected Crash Frequency Estimates

LOS	*Actual Crashes*	*Predicted Crashes*	*Weight*	*Expected Crashes*
A	408	174.811	0.025	402.193
B	297	191.368	0.030	293.795
C	434	253.915	0.019	430.589
D	212	144.177	0.028	210.118
E	267	125.582	0.026	263.258
F	938	352.362	0.028	921.680

Table 4.5 presents the sum of all observed and predicted crash frequencies at all levels of service. The observed crash frequencies are the sum of the actual crashes observed at each hour for each LOS. Predicted crash frequencies are also the sum the SPF generated crash frequencies using *equation 3.6*. The weight factor was calculated using *equation 3.7* and the expected crash frequencies, $N_{expected,T,B}$ was calculated using *equation 3.8*.

Having found the expected crash frequencies for each level of service using the Empirical Bayes before-after models, the next procedure involved developing the CMFs for each deterioration in LOS as discussed next in Section 4.3.

4.3 Developing Crash Modification Factors

The final step in this study's data analysis involved estimating the CMF when the LOS deteriorated from A to B, B to C, C to D, D to E and E to F using the procedures outlined in Chapter 3 of this dissertation. Using the weight factor all the observed, predicted and expected crash frequencies found, CMFs were estimated and presented in Table 4.6.

Table 4.6 Crash Modification Factors

Parameter	***LOS A to LOS B***	***LOS B to LOS C***	***LOS C to LOS D***	***LOS D to LOS E***	***LOS E to LOS F***
$N_{expected,T,B}$ =	402.193	293.795	430.589	210.118	263.258
$N_{predicted,T,A}$ / $N_{predicted,T,B}$ =	1.095	1.327	0.568	0.871	1.053
$N_{expected,T,A}$ =	440.288	389.818	244.496	183.018	277.282
Var ($N_{expected,T,A}$) =	469.989	501.531	136.199	154.989	284.326
CMF =	**0.673**	**1.110**	**0.865**	**1.452**	**3.370**
Expected Effect on Safety =	**- 33%**	**11%**	**-13%**	**45%**	**237%**

Table 4.6 shows the results of the four steps described in Chapter 3 and presented in Sub-Section 3.3 to determine the CMFs when the LOS deteriorated on urban freeways. The four steps are finding the: (1) ratio of the predicted number of crashes during the after period to the predicted number of crashes during the before period ($N_{predicted,T,A}$ / $N_{predicted,T,B}$); (2) expected number of crashes in the after period in the treatment group that would have occurred without treatment ($N_{expected,T,A}$), (3) the variance of $N_{expected,T,A}$ and (4) finally estimation of the CMF.

The results in Table 4.6 can be interpreted as follows: When the LOS changed from A to B, the CMF was 0.67, a safety benefit of 33% or an expected reduction in crashes of thirty-three percent. When the LOS changed from B to C, the CMF was 1.11, a safety degradation of 11% or an expected increase in crashes of eleven percent. When the LOS changed from C to D, the CMF was 0.865, a safety benefit of 13% or an expected thirteen percent reduction in crashes. When the LOS changed from D to E, the CMF was 1.452, a safety degradation of 45% or an expected forty-five percent increase in crashes. Finally, when the LOS changed from E to F, the CMF was 3.37, a safety degradation of 237% or an expected increase in crashes of two hundred and thirty-seven percent.

Therefore, the largest reduction in crash frequency would be between LOS A and LOS B where the CMF is 0.673 with an expected reduction in crashes of thirty-three percent. The most hazardous change in LOS would be between LOS E to F where the CMF is 3.37 with an expected two hundred and thirty-even percent increase in crashes.

As the crash frequency increase when the LOS changed from E to F was significantly large, three individual study sites that experienced both levels of services of E and F were randomly selected to find out if the expected increase in crashes determined by the CMF was similar with the observed crash frequencies. The results of this task are tabled in Table 4.7.

Table 4.7 Observed Crash Averages for each LOS at Selected Study Sites

		ROUTE 1		ROUTE 3
Change in LOS	**CMF Findings (%)**	**MP 46 (%)**	**MP 52.29 (%)**	**MP 9.50 (%)**
A to B	-33	-40.00	-42.11	-70.00
B to C	11	150.00	400.00	185.71
C to D	-13	-25.00	-40.63	-61.54
D to E	45	533.33	-53.85	-50.00
E to F	237	231.25	242.86	150.00

The results for when the level of service changes from E to F presented in Table 4.7 show that on the average, at this LOS degradation, the number of crash frequencies tend to increase by 200%. This is similar to the findings in Table 4.6, which presents the expected number of crashes, CMFs and the percentage of expected effect on safety. However, the findings of B to C and D to E are different.

4.4 Severity Analysis

A similar investigation was conducted on crash severity to determine the CMFs for combined fatalities and injuries when LOS degraded from LOS A to LOS F. The first task involved determining the possible SPF models using Pearson's Correlation Matrices as discussed next in Sub-Section 4.4.1.

4.4.1 Determining the Possible Models

Table 4.8 Pearson's Correlation Matrix Table

	Traffic Volume	Posted Speed	Number of Lanes	Lane Width	Shoulder Width	Median Width	Density
Traffic Volume	1	0.219	0.407	-0.206	0.263	0.361	0.621
		0	0	0	0	0	0
Posted Speed	0.219	1	0.484	0.238	0.788	0.559	0.151
	0		0	0	0	0	0
Number of Lanes	0.407	0.484	1	-0.217	0.642	0.401	0.157
	0	0		0	0	0	0
Lane Width	-0.206	0.238	-0.217	1	0.257	0.09	-0.074
	0	0	0		0	0.001	0.007
Shoulder Width	0.263	0.788	0.642	0.257	1	0.39	0.118
	0	0	0	0		0	0
Median Width	0.361	0.559	0.401	0.09	0.39	1	0.25
	0	0	0	0.001	0		0
Density	0.621	0.151	0.157	-0.074	0.118	0.25	1
	0	0	0	0.007	0	0	

Referring to Table 4.8, the model with the best fit was selected using the discussed criteria that the predictor variables in the selected model were both uncorrelated and had a correlation coefficient that was within the range of -0.03 and +0.03. Two models were selected, they were:

Model I: Traffic Volume, Posted Speed, Lane Width, Density

Model II: Traffic Volume, Shoulder Width, Lane Width, Density

4.4.2 Measuring Goodness-of-Fit for each Model

Table 4.9a, shows the goodness-of-fit results for Model I. The *Value/df* for Deviance at LOS E, is much greater than one and therefore Model I was not selected as it is not fit for analysis. Crash frequency of fatalities and injuries for LOSs E and F were combined since they both have similar characteristics and both had very low frequencies. After combining LOSs E and F, the *Value/df* for Deviance reduced to approximately l. This indicates that at all LOSs A, B, C, D, and E & F, the data fits the model well.

Table 4.9a Goodness-of-Fit for Model I

LOS	Parameter	Value	df	Value/df
A	Deviance	291.094	426	0.683
	Pearson Chi-Square	443.919	426	1.042
	Sig	0.001		
B	Deviance	199.079	249	0.8
	Pearson Chi-Square	267.09	249	1.073
	Sig	0.127		
C	Deviance	223.261	263	0.849
	Pearson Chi-Square	262.051	263	0.996
	Sig	0.18		
D	Deviance	87.987	99	0.889
	Pearson Chi-Square	117.037	99	1.182
	Sig	0.539		
E	Deviance	10326.96	89	116.033
	Pearson Chi-Square	4.45E+33	89	5.00E+31
	Sig	0		
F	Deviance	203.206	182	1.117
	Pearson Chi-Square	195.712	182	1.075
	Sig	0		
E & F	Deviance	283.616	277	1.024
	Pearson Chi-Square	317.108	277	1.145
	Sig	0		

Table 4.9b, shows the goodness-of-fit results for Model II. The *Value/df* for Deviance at LOS E, is much greater than one and therefore this model was not included as fit for analysis. Fatalities and injuries for LOSs E and F were combined since they both have similar characteristics and both had very low severity counts. After combining fatalities and injuries for LOSs E and F, the *Value/df* for Deviance reduced to approximately l. The *Value/df* for Pearson Chi-Square at LOS E & F was 1.203. As a result, Model II was not considered for severities SPF analysis.

Table 4.9b Goodness-of-Fit for Model II

LOS	Parameter	Value	df	Value/df
	Deviance	291.174	426	0.684
A	Pearson Chi-Square	448.923	426	1.054
	Sig	0.001		
	Deviance	199.199	249	0.8
B	Pearson Chi-Square	264.002	249	1.06
	Sig	0.122		
	Deviance	223.54	263	0.85
C	Pearson Chi-Square	261.656	263	0.995
	Sig	0.192		
	Deviance	87.535	99	0.884
D	Pearson Chi-Square	102.783	99	1.038
	Sig	0.048		
	Deviance	11570.142	89	130.002
E	Pearson Chi-Square	2.43E+37	89	2.73E+35
	Sig	0.002		
	Deviance	203.132	182	1.116
F	Pearson Chi-Square	203.01	182	1.115
	Sig	0		
	Deviance	281.946	277	1.018
E & F	Pearson Chi-Square	3.33E+02	277	1.203
	Sig	0		

Even though the *P-Values* for both Models I and II are more than 0.1, Model I was used based on the goodness-of-fit test results. The SPF analysis for Model I is presented next. The original SPSS outputs are presented in Appendix C of this dissertation. The next task-involved development of SFPs as discussed in Sub-Section 4.4.3 next.

4.4.3 Development of Safety Performance Functions

SPFs were used to predict the crash frequency by severity for each LOS based on the goodness-of-fit test, Model I was selected for the development of SPFs. This model's variables were crash frequency by severity, which is the dependent variable, traffic volumes, lane width, posted speed limit, and density. Negative binomial models were used for the SPF development and the resultant coefficients for each variable presented in Tables 4.10a through 4.10e.

Table 4.10a Crash Frequency of Fatalities and Injuries Parameter Estimates for LOS A

Parameter	B	Std. Error	Sig.
(Intercept)	5.317	1.9006	0.005
Traffic Volume (Hourly)	0.002	0.0007	0.002
Posted Speed Limit (m/h)	-0.088	0.0354	0.013
Lane Width (ft.)	-0.083	0.121	0.495
Density (veh/mi)	-0.433	0.1248	0.001
Dispersion (*k*)	2.964	0.492	

Table 4.10a shows the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for LOS A. Also

shown are the dispersion estimates. The intercept in Table 4.10a shows the negative binomial regression estimated coefficients when all variables are held at zero.

At a significance level of 0.1, all the variables except for lane width are significant and as a result, lane width was not considered in the SPF development for LOS A. The dispersion coefficient is positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate.

Table 4.10b Crash Frequency of Fatalities and Injuries Parameter Estimates for LOS B

Parameter	B	Std. Error	Sig.
(Intercept)	4.193	2.4619	0.089
Traffic Volume (Hourly)	0	0.0003	0.447
Posted Speed Limit (m/h)	-0.003	0.0416	0.937
Lane Width (ft.)	-0.266	0.1542	0.084
Density (veh/mi)	-0.137	0.0844	0.106
Dispersion (*k*)	1.948	0.4358	

Table 4.10b shows the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for LOS B. Also shown are the dispersion estimates. The intercept in Table 4.10b shows the negative binomial regression estimated coefficients when all variables are held at zero.

At a significance level of 0.1, all the variables except for traffic volume and posted speed are significant and as a result, were not considered in the SPF development for LOS B. The dispersion coefficient is positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate.

Table 4.10c Crash Frequency of Fatalities and Injuries Parameter Estimates for LOS_C

Parameter	B	Std. Error	Sig.
(Intercept)	0.751	2.2226	0.735
Traffic Volume (Hourly)	0	0.0002	0.219
Posted Speed Limit (m/h)	-0.08	0.0332	0.016
Lane Width (ft.)	0.26	0.1776	0.143
Density (veh/mi)	-0.055	0.0612	0.373
Dispersion (*k*)	1.543	0.3642	

Table 4.10c shows the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for LOS C. Also shown are the dispersion estimates. The intercept in Table 4.10c shows the negative binomial regression estimated coefficients when all variables are held at zero.

At a significance level of 0.1, all the variables except for traffic volume and density are significant and as a result, were not considered in the SPF development for LOS C. The dispersion coefficient is positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate.

Table 4.10d Crash Frequency of Fatalities and Injuries Parameter Estimates for LOS D

Parameter	B	Std. Error	Sig.
(Intercept)	-1.18	4.374	0.787
Traffic Volume (Hourly)	0	0.0003	0.176
Posted Speed Limit (m/h)	-0.096	0.0589	0.101
Lane Width (ft.)	0.394	0.3982	0.322
Density (veh/mi)	-0.033	0.0796	0.676
Dispersion (*k*)	1.53	0.5369	

Table 4.10d shows the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for LOS D. Also shown are the dispersion estimates. The intercept in Table 4.10d shows the negative binomial regression estimated coefficients when all variables are held at zero.

At a significance level of 0.1, all the variables except for lane width and density are significant and as a result, were not considered in the SPF development for LOS D. The dispersion coefficient is positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate.

Table 4.10e Crash Frequency of Fatalities and Injuries Parameter Estimates for LOS E & F

Parameter	B	Std. Error	Sig.
(Intercept)	-4.341	1.5927	0
Traffic Volume (Hourly)	0	0.0001	0
Posted Speed Limit (m/h)	0.048	0.0401	0.229
Lane Width (ft.)	0.396	0.0866	0
Density (veh/mi)	-0.002	0.0011	0.081
Dispersion (*k*)	0.786	0.1614	

Table 4.10e shows the negative binomial parameter estimates, standard errors, and statistical significance of the intercept and the predictor variables for LOS E & F. Also shown are the dispersion estimates. The intercept in Table 4.10e shows the negative binomial regression estimated coefficients when all variables are held at zero.

At a significance level of 0.1, all the variables except for posted speed are significant and as a result, was not considered in the SPF development for LOS E & F.

The dispersion coefficient is positive and greater than zero, suggesting over-dispersion and therefore the negative binomial model was appropriate.

The coefficients found in Model I were used for calculating the crash frequency of fatalities and injuries. However, as explained in Chapter 3, to account for the RTM effect, a more rigorous analysis was conducted to determine the expected number of crashes by severity of each LOS. To achieve this, both the predicted severities found using the SPFs in Sub-Section 4.4.3 and the observed severities were combined using Empirical Bayesian relationships to find the expected severities. Empirical Bayes model was used to, estimate the number of fatalities and injuries that would have occurred had at the study site in the after period had a treatment not been implemented (Gross et al., 2010). The treatment in this case was the progressive degrading of levels of services from LOS A to LOS E &F. Specifically, the EB before-after model was used to account for the regression-to-mean (RTM) effect usually associated with the crash data. The procedure for calculating the expected crash frequencies is next discussed in Sub-Section 4.4.4.

4.4.4 Determining Expected Crash Frequencies

As stated in Section 4.1, SPFs are only part of the overall safety evaluation process in this study and the observed severities need to be accounted for in determining expected severity. In this Sub-Section, the expected number of crashes by severity was determined using Empirical Bayesian method to increase the precision of safety estimation by accounting for the RTM bias. Using the overdispersion coefficients found during SPF modeling, the expected number of crashes by severity was estimated and presented in Table 4.11.

Table 4.11 Estimates of Expected Crash Frequency by Severity

LOS	*Actual Fatalities & Injuries*	*Predicted Fatalities & Injuries*	*Weight Factor*	*Expected Fatalities & Injuries*
A	266	246.667	2.964	208.697
B	167	106.003	1.948	48.177
C	180	103.535	1.543	62.014
D	75	32.519	1.530	10.004
E & F	389	311.847	0.786	328.358

Table 4.11 presents the sum of all observed and predicted fatalities and injuries at all levels of services. The observed crash frequency by severity is the sum of the actual fatalities and injuries observed at each hour for each LOS. Predicted fatalities and injuries are the sum the SPF generated crash frequency by severity using *equation 3.6*. The weight factor was calculated using *equation 3.7* and the expected crash frequency by severity was calculated using *equation 3.8*.

Having found the expected number of fatalities and injuries for each level of service using the Empirical Bayes before-after models, the next procedure involved developing the CMFs for each deterioration in LOS as discussed next in Section 4.4.5.

4.4.5 Developing Crash Modification Factors

The final step in data analysis involved estimating the CMF when the LOS deteriorated from A to B, B to C, C to D, and D to E & F using the procedures outlined in Chapter 3 of this dissertation. Using the weight factor all the observed, predicted, and expected crash frequency by severity found, CMFs were estimated and presented in Table 4.12.

Table 4.12 Crash Frequency by Severity Crash Modification Factors

Parameter	***LOS A to LOS B***	***LOS B to LOS C***	***LOS C to LOS D***	***LOS D to LOS E F***
$N_{expected,T,B}$ =	208.697	48.177	62.014	10.004
$N_{predicted,T,A} / N_{predicted,T,B}$ =	0.430	0.977	0.314	9.590
$N_{expected,T,A}$ =	89.685	47.056	19.478	95.934
$Var\ (N_{expected,T,A})$ =	-75.695	-43.570	-3.322	-487.589
CMF =	**1.880**	**3.902**	**3.885**	**4.282**
Expected Effect on Severity =	**88%**	**290%**	**288%**	**328%**

Table 4.12 shows the results of the four steps described in Chapter 3 and presented in Sub-Section 3.3 to determine the CMFs when the LOS deteriorated on urban freeways. The results presented in Table 4.12 can be interpreted as follows: When the LOS changed from A to B, the CMF was 1.88, a severity degradation of 88% or an increase in the number of fatalities and injuries by severity eighty-eight percent. When the LOS changed from B to C, the CMF was 3.90, a severity degradation of 290% or an increase in the number of fatalities and injuries by two hundred and ninety percent. When the LOS changed from C to D, the CMF was 3.89, a severity degradation of 288% or an increase in the number of fatalities and injuries by two hundred and eighty-eight percent increase in severity, finally, when the LOS changed from D to E 7 F, the CMF was 4.28, a severity degradation of 328% or an increase in the number of fatalities and injuries by three hundred and twenty-eight percent increase in severity.

The results from the severity analysis suggest that as the LOS degraded from LOS A to B, B to C, C to D and D to E & F, severity degraded too. These findings are not similar to those of Christoforou et al. (2010) and Martin (2002) discussed in the literature review of this dissertation. According to Christoforou et al. (2010) and Martin (2002), the number of severities should decrease as the traffic volumes increase and speed reduces.

4.5 Conclusions

This Chapter applied the crash data collected on the selected urban freeways to the methodology developed. The first task involved determining the predicted crash frequencies and severities using SPF analysis. Here, three models were determined using Pearson's correlation matrices and the model(s) that fitted the data well was/were selected by applying the goodness-of-fit tests. To be specific, model III and model I were selected for crash frequency and severity respectively.

As discussed in Chapter 2, of all crash affecting variables, traffic volume has the most influence and was accounted for during SPF analysis, especially if it is not constant. However, SPFs fail to account for the regression-to-mean (RTM) effect on the predicted crash frequencies and therefore, using a weight factor and Empirical Bayes the RTM effect was corrected and the Expected crash frequencies and severities calculated. Finally, the CMFs for each LOS deterioration, beginning from LOS A to LOS F was calculated, presented, and discussed. Further interpretations, recommendations, and conclusions of this results and research were then presented in Chapter 5 next.

CHAPTER 5

INTERPRETATIONS, CONCLUSIONS and RECOMMENDATIONS

5.1 Interpretations

5.1.1 Crash Frequencies

The main objective of this study was to show whether freeway operational elements influenced crash frequencies. If so, then this study would recommend that operational elements, specifically level of service to be considered in the future versions of the Highway Safety Manuals. The *effect on safety* findings in Table 4.6 in Chapter 4 were graphically presented in Figure 5.1 to show the expected percentage trend in crash frequencies when the LOS deteriorated.

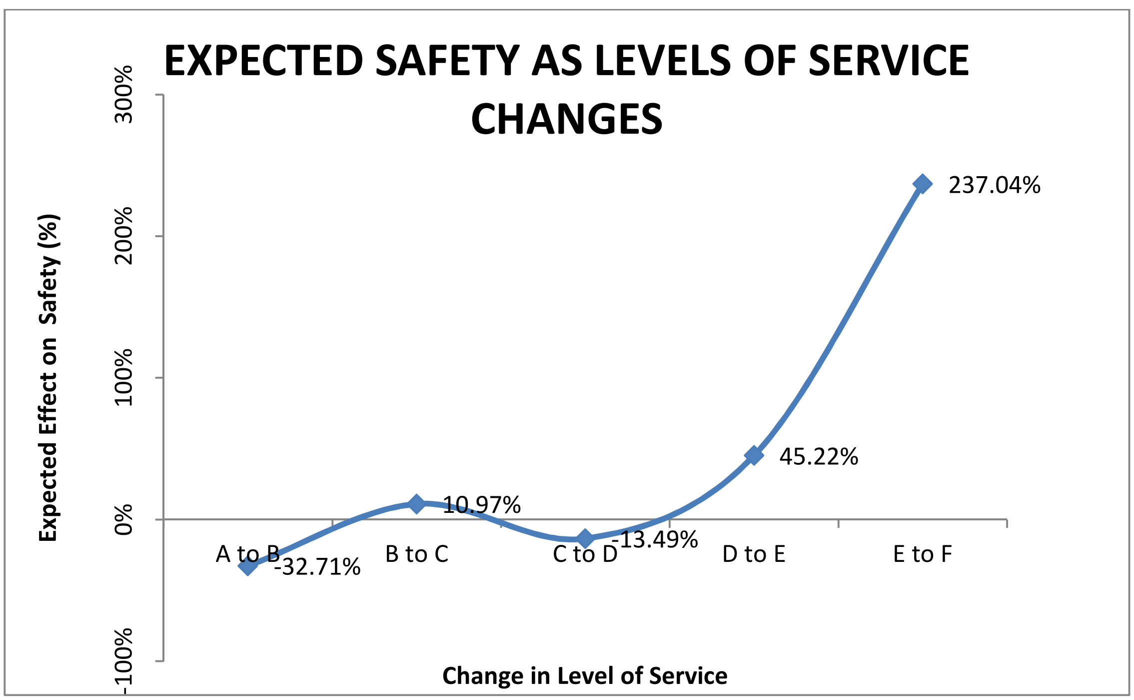

Figure 5.1 Crash frequency trend as the levels of services deteriorated.

Figure 5.1 shows that there is a relationship between operational elements and crash frequencies. The curve in the figure indicates that on urban freeways, as the level of services deteriorate from A to B, B to C, C to D, D to E and E to F, the influence is almost sinusoidal. That is as the LOS degrades from A to B the crash frequency percentages reduces, the crash percentages then increase as the LOS degrades from B to C and then decreases when the LOS degrades from C to D. The crash frequency percentages then reduces as LOS changes from D to E and increases significantly when LOS changes from E to F.

The findings of this study seem to concur with the few other studies that have acknowledged that operational elements have some influence on safety. For example, (1) Kononov et al.'s, (2008) relationships between safety. Kononov et al. states that

> *"Relating safety to the degree of congestion suggests that safety deteriorates with the degradation in the quality of service expressed through the level of service. Practitioners generally believe that additional capacity afforded by additional lanes is associated with more safety. How much safety and for what time period are generally not considered. Comparison of SPFs of multilane freeways suggests that adding lanes may initially result in a temporary safety improvement that disappears as congestion increases."*

And (2) in his discussion of previous work, Lord et al. (2004) discusses studies by Frantzeskakis and Iordanis (1987), Persaud, and Nguyen (2000) and concluded that crash frequencies increased as the LOS decreased from LOS of "A" to LOS of "F." However, this study is unique in that it uses CMFs as directed by the HSM (2010) to quantify safety when the operational conditions deteriorate on urban freeways.

5.1.2 Crash Severities

This study also investigates whether operational elements influenced the number of fatalities and injuries specifically when the LOS deteriorated from A to B, B to C, C to D, and D to E & F. If so, then this study would recommend that operational elements, specifically level of service to be considered in the future versions of the Highway Safety Manuals. The *effect on safety* findings in Table 4.12 in Chapter 4 are graphically presented in Figure 5.2 to show the expected percentage change in fatalities and injuries when the LOS deteriorated.

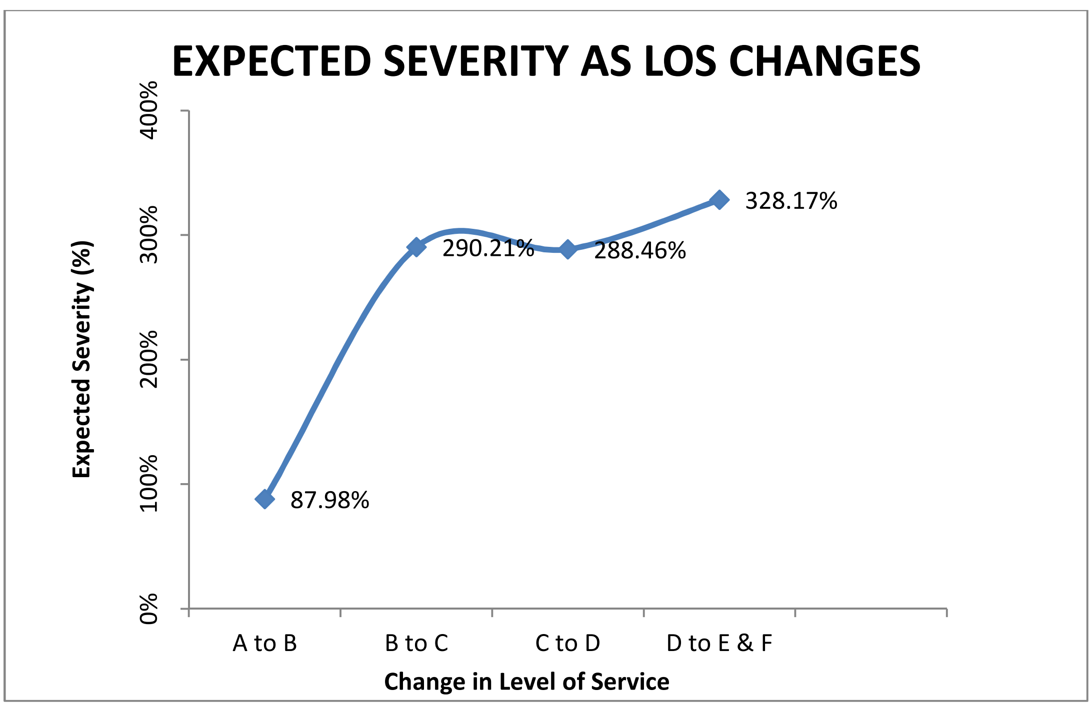

Figure 5.2 Crash severity trend as the levels of services deteriorated.

Figure 5.2 shows that there is a relationship between operational elements and crash severities. The curve in the figure indicates that on urban freeways, as the level of services deteriorate from A to B, B to C, C to D, and D to E & F, the number of fatalities and injuries is expected to increase. That is as the LOS degrades from A to B the fatal and injury percentages increases, the crash percentages continue to increase as the LOS degrades from B to C and then decreases when the LOS degrades from C to D. The percentage change in fatalities and injuries then increases as LOS changes from D to E & F.

5.2 Conclusions

The main objective of this research was to determine if operational elements, specifically if levels of service influenced crash frequencies on urban freeways. If a relationship is established, it is recommended that AASHTO include these findings in the future versions of the Highway Safety Manual. The results presented in Table 4.6 and Figure 5.1 show that indeed due to changes in Level of Service, operational elements have some influence on crash frequencies.

This study used hourly volumes in the investigation to calculate the density and assigning of levels of service for each hour. This research used the Highway Capacity Manual (HCM, 2010) in both calculating density and assigning the level of service. Thereafter, the CMFs due to change in level of service were estimated using the procedures recommended by the Highway Safety Manual (HCM, 2010). The reference to these manuals in this study, showed that both manuals can be used together in designing and construction of both existing and proposed urban freeways.

As indicated in the literature review of this study, studies by Frantzeskakis and Iordanis (1987), Persaud, and Nguyen (2000) which examined the effects of LOS on safety and concluded that both crash frequencies and crash rates increased as the LOS degraded from LOS of "A" through LOS of "F." However, this was not exactly the same as the findings of this study. Explanation and recommendations on the results are discussed next in Section 5.3.

5.3 Recommendations

The main objective of transportation practitioners is to design and maintain roadways that are safe as possible. Using the results found by this study, to meet these condition on urban freeways, the levels of service should be maintained between levels of service A through C. The level of service change between B and C being more safer of the two. This could be due to the higher speeds and the free flow conditions associated with LOS A and B.

Levels of service between A and B on urban freeways were also found to be safe however due to congestion and reduced speeds, the travels conditions are not free and therefore not recommended. This condition avoided by directing traffic to other roadways and increasing the number of lanes. According to the results found by this study, the most hazardous level of service on urban freeways is between LOS D and E.

Not only do the crash frequencies increase, the operational conditions deteriorated as well.

The results in this study are not conclusive however, should provide a basis to be used to influence more research on this topic so that there is through and better understanding on how operational elements precisely influence crash frequencies. Thereafter, more constraints should be included to the variables such as by crash type, road surface type, and surface condition. Not only should the constraints be included, but also different methodology approach should be encourage to reduce the errors associated with crash data to find a more accurate solution as possible.

APPENDIX A: VARIABLES USED IN HSM (2010) SAFETY PREDICTIONS

1. Area Type (rural/suburban/urban)
2. AADT
3. Length of roadway segment
4. Number of through lanes
5. Lane Width
6. Shoulder width
7. Shoulder type
8. Presence of median (divided/undivided)
9. Median width
10. Presence of concrete median barrier
11. Presence of passing lane
12. Presence of short four-lane Section
13. Presence of two way left-turn lane
14. Driveway density
15. Number of major commercial driveways
16. Number of minor commercial driveways
17. Number of major residential driveways
18. Number of minor residential driveways
19. Number of major industrial/institutional driveways
20. Number of minor industrial/institutional driveways
21. Number of other driveways
22. Horizontal curve length

23. Horizontal curve radius
24. Horizontal curve super-elevation
25. Presence of spiral transition
26. Grade
27. Roadside hazard rating
28. Roadside slope
29. Roadside fixed-object density
30. Roadside fixed-object offset
31. Percent of length with on-street parking
32. Type of on-street parking
33. Presence of lighting

APPENDIX B

Crash Frequency Negative Binomial SPSS 16.0 Output Estimates

MODEL I

LOS A Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	3.138	1.0049	1.168	5.107	9.751	1	0.002
vol	0.002	0.0003	0.001	0.00	32.433	1	0.000
psd	-0.082	0.0183	-0.117	-0.046	19.773	1	0.000
LW	0.089	0.072	-0.052	0.23	1.526	1	0.217
D	-0.364	0.0638	-0.489	-0.239	32.61	1	0.000
(Neg. Bin.)	0.276	0.0972	0.138	0.55			
Dependent Variable: crash							
Model: (Intercept), vol, psd, LW, D							

LOS B Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-0.841	1.3695	-3.525	1.843	0.377	1	0.539
vol	0.001	0.0002	0	0.001	13.17	1	0
psd	-0.024	0.0226	-0.068	0.02	1.144	1	0.285
LW	0.106	0.1079	-0.106	0.317	0.958	1	0.328
D	-0.051	0.0496	-0.148	0.047	1.038	1	0.308
(Neg. Bin.)	0.222	0.0946	0.096	0.512			
Dependent Variable: crash							
Model: (Intercept), vol, psd, LW, D							

LOS C Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	1.358	1.1677	-0.93	3.647	1.353	1	0.245
vol	0	0.0001	6.95E-05	0	6.801	1	0.009
psd	-0.093	0.0176	-0.127	-0.058	27.619	1	0.000
LW	0.236	0.0902	0.059	0.413	6.845	1	0
D	-0.003	0.03	-0.062	0.056	0.012	1	0.913
(Neg. Bin.)	0.211	0.0747	0.105	0.422			
Dependent Variable: crash							
Model: (Intercept), vol, psd, LW, D							

LOS D Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-2.697	2.3159	-7.236	1.842	1.356	1	0.244
vol	7.37E-05	0.0001	0	0	0.255	1	0.613
psd	-0.035	0.0297	-0.093	0.023	1.375	1	0.241
LW	0.375	0.1955	-0.008	0.758	3.684	1	0.055
D	0.012	0.0376	-0.061	0.086	0.106	1	0.745
(Neg. Bin.)	0.248	0.1117	0.103	0.6			
Dependent Variable: crash							
Model: (Intercept), vol, psd, LW, D							

LOS E Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-11.776	1.8344	-15.371	-8.18	41.209	1	0.000
vol	0	0.0002	0	7.66E-05	2.579	1	0.108
psd	0.124	0.0453	0.035	0.213	7.476	1	0.006
LW	0.243	0.0934	0.06	0.427	6.79	1	0.009
D	0.133	0.0266	0.081	0.185	24.902	1	0.000
(Neg. Bin.)	0.131	0.0681	0.047	0.363			
Dependent Variable: crash							
Model: (Intercept), vol, psd, LW, D							

LOS F Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-8.431	1.1351	-10.655	-6.206	55.166	1	0.000
vol	0	0.0001	8.80E-05	0	8.032	1	0.005
psd	0.049	0.0349	-0.019	0.117	1.976	1	0.160
LW	0.458	0.0656	0.33	0.587	48.761	1	0.000
D	-0.001	0.0007	-0.002	0	2.088	1	0.148
(Neg. Bin.)	0.303	0.0582	0.208	0.442			

Dependent Variable: crash
Model: (Intercept), vol, psd, LW, D

Negative Binomial SPSS 16.0 Output Estimates for MODEL II

LOS A Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	-0.831	0.8654	-2.527	0.865	0.922	1	0.337
vol	0.002	0.0004	0.002	0.003	32.871	1	0
SW	-0.062	0.015	-0.091	-0.032	16.827	1	0
LW	0.111	0.0757	-0.038	0.259	2.133	1	0.144
D	-0.429	0.0748	-0.576	-0.282	32.854	1	0
(Neg. Bin.)	0.282	0.098	0.142	0.557			

Dependent Variable: crash
Model: (Intercept), vol, SW, LW, D

LOS B Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	-3.254	1.4852	-6.165	-0.343	4.801	1	0.028
vol	0.001	0.0002	0.001	0.001	21.696	1	0
SW	-0.057	0.0187	-0.094	-0.021	9.425	1	0.002
LW	0.249	0.1266	0.001	0.497	3.881	1	0.049
D	-0.119	0.0544	-0.225	-0.012	4.745	1	0.029
(Neg. Bin.)	0.188	0.0895	0.074	0.478			

Dependent Variable: crash
Model: (Intercept), vol, SW, LW, D

LOS C Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	-3.671	1.341	-6.299	-1.042	7.493	1	0.006
vol	0	0.0001	0	0.001	9.515	1	0.002
SW	-0.067	0.0144	-0.095	-0.039	21.581	1	0
LW	0.299	0.1029	0.097	0.501	8.448	1	0.004
D	-0.016	0.0316	-0.078	0.046	0.268	1	0.604
(Neg. Bin.)	0.23	0.077	0.119	0.443			

Dependent Variable: crash
Model: (Intercept), vol, SW, LW, D

LOS D Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-4.369	3.0338	-10.315	1.577	2.074	1	0.15
vol	9.80E-05	0.0002	0	0	0.232	1	0.63
SW	-0.022	0.0287	-0.079	0.034	0.615	1	0.433
LW	0.384	0.2279	-0.063	0.831	2.838	1	0.092
D	0.008	0.0397	-0.07	0.086	0.041	1	0.84
(Neg. Bin.)	0.255	0.1127	0.107	0.607			
Dependent Variable: crash							
Model: (Intercept), vol, SW, LW, D							

LOS E Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-13.336	2.6758	-18.58	-8.091	24.838	1	0
vol	0.001	0.0005	0	0.002	5.281	1	0.022
SW	-0.194	0.0929	-0.376	-0.012	4.373	1	0.037
LW	0.682	0.1648	0.36	1.005	17.157	1	0
D	0.019	0.047	-0.073	0.111	0.168	1	0.682
(Neg. Bin.)	0.142	0.0707	0.054	0.377			
Dependent Variable: crash							
Model: (Intercept), vol, SW, LW, D							

LOS F Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Square**	**df**	**Sig.**
(Intercept)	-8.998	1.0841	-11.123	-6.873	68.891	1	0
vol	0.001	9.58E-05	0	0.001	35.05	1	0
SW	-0.064	0.0275	-0.118	-0.01	5.424	1	0.02
LW	0.614	0.067	0.482	0.745	83.775	1	0
D	-0.002	0.0007	-0.004	0	11.446	1	0.001
(Neg. Bin.)	0.024	0.057	0.201	0.43			
Dependent Variable: crash							
Model: (Intercept), vol, SW, LW, D							

Negative Binomial SPSS 16.0 Output Estimates for MODEL III

LOS A Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	-1.949	2.0437	-4.893	1.621	0.134	1	0.014
vol	0.001	0.0007	0	0.002	0.551	1	0.058
psd	-0.07	0.03	-0.129	-0.011	5.396	1	0.02
N	0.865	0.3008	0.276	1.455	8.278	1	0.004
LW	0.183	0.0814	0.023	0.342	5.035	1	0.025
SW	-0.043	0.0233	-0.089	0.003	3.413	1	0.065
MW	0.018	0.0106	-0.003	0.039	2.895	1	0.089
D	-0.092	0.1214	-0.33	0.146	0.574	1	0.049
(Neg. Bi.)	0.224	0.0908	0.101	0.496			

Dependent Variable: crash
Model: (Intercept), vol, psd, N, LW, SW, MW, D

LOS B Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	-5.56	3.1408	-11.716	0.596	3.134	1	0.077
vol	0.001	0.0007	0	0.003	2.667	1	0.102
psd	0.061	0.0358	-0.009	0.131	2.942	1	0.086
N	-0.082	0.6846	-1.424	1.26	0.014	1	0.051
LW	0.225	0.1265	-0.023	0.473	3.161	1	0.075
SW	-0.093	0.0289	-0.149	-0.036	10.301	1	0.001
MW	-0.009	0.0146	-0.037	0.02	0.361	1	0.548
D	-0.147	0.137	-0.416	0.121	1.156	1	0.028
(Neg. Bi.)	0.167	0.0869	0.061	0.463			

Dependent Variable: crash
Model: (Intercept), vol, psd, N, LW, SW, MW, D

LOS C Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	-2.753	3.0305	-6.539	1.033	0.505	1	0.047
vol	2.58E-05	0.0007	-0.001	0.001	0.001	1	0.097
psd	-0.06	0.0297	-0.118	-0.002	4.057	1	0.044
N	0.516	0.9603	-1.366	2.398	0.289	1	0.059
LW	0.307	0.1064	0.098	0.515	8.318	1	0.004
SW	-0.031	0.021	-0.072	0.01	2.147	1	0.014
MW	-0.011	0.0189	-0.048	0.026	0.351	1	0.554
D	0.039	0.1065	-0.17	0.248	0.135	1	0.071
(Neg. Bi.)	0.204	0.0737	0.1	0.414			
Dependent Variable: crash							
Model: (Intercept), vol, psd, N, LW, SW, MW, D							

LOS D Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	1.528	6.2511	-9.196	12.252	0.007	1	0.033
vol	0.001	0.0016	-0.002	0.004	0.652	1	0.019
psd	-0.036	0.0401	-0.114	0.043	0.788	1	0.037
N	-2.051	2.7645	-7.469	3.367	0.55	1	0.058
LW	0.407	0.2337	-0.052	0.865	3.025	1	0.082
SW	-0.009	0.0331	-0.073	0.056	0.068	1	0.045
MW	0.001	0.0182	-0.034	0.037	0.006	1	0.939
D	-0.115	0.171	-0.45	0.221	0.449	1	0.053
(Neg. Bi.)	0.243	0.1109	0.099	0.594			
Dependent Variable: crash							
Model: (Intercept), vol, psd, N, LW, SW, MW, D							

LOS E Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	-12.984	6.5978	-25.916	-0.053	4.492	1	0.034
vol	0.001	0.0021	-0.003	0.006	0.434	1	0.051
psd	0.198	0.0478	0.105	0.292	17.216	1	0
N	-2.567	4.2896	-10.975	5.84	0.358	1	0.055
LW	0.431	0.192	0.054	0.807	5.031	1	0.025
SW	-0.108	0.0955	-0.295	0.079	1.274	1	0.025
MW	-0.049	0.008	-0.064	-0.033	37.38	1	0
D	-0.001	0.1563	-0.308	0.305	0	1	0.009
(Neg. Bi.)	0.293	.	.	.			

Dependent Variable: crash
Model: (Intercept), vol, psd, N, LW, SW, MW, D

LOS F Parameter Estimates							
			95% Wald Confidence Interval		Hypothesis Test		
Parameter	**B**	**Std. Error**	**Lower**	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**
(Intercept)	-11.922	1.0838	-14.047	-9.798	146.574	1	0
vol	0	8.11E-05	0	0.001	19.929	1	0
psd	0.366	0.057	0.254	0.477	41.115	1	0
N	-3.033	0.5887	-4.187	-1.879	26.539	1	0
LW	0.177	0.0823	0.015	0.338	4.612	1	0.032
SW	0	.	.	.	.	.	.
MW	-0.043	0.0052	-0.053	-0.033	68.85	1	0
D	-0.002	0.0006	-0.003	0	9.592	1	0.002
(Neg. Bi.)	0.099	0.0339	0.05	0.194			

Dependent Variable: crash
Model: (Intercept), vol, psd, N, LW, SW, MW, D

APPENDIX C:

Crash Frequency of Fatalities and Injuries SPSS 16.0 Output Estimates

MODEL I

LOS A Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	5.317	1.9006	1.592	9.042	7.827	1	0.005
vol	0.002	0.0007	0.001	0.003	9.511	1	0.002
psd	-0.088	0.0354	-0.158	-0.019	6.229	1	0.013
lw	-0.083	0.121	-0.32	0.154	0.466	1	0.495
D	-0.433	0.1248	-0.677	-0.188	12.039	1	0.001
Dispersion	2.964	0.492	2.141	4.104			

LOS B Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	4.193	2.4619	-0.632	9.018	2.901	1	0.089
vol	0	0.0003	0	0.001	0.579	1	0.447
psd	-0.003	0.0416	-0.085	0.078	0.006	1	0.937
lw	-0.266	0.1542	-0.568	0.036	2.981	1	0.084
D	-0.137	0.0844	-0.302	0.029	2.613	1	0.106
Dispersion	1.948	0.4358	1.256	3.02			

LOS C Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	0.751	2.2226	-3.605	5.108	0.114	1	0.735
vol	0	0.0002	0	0.001	1.51	1	0.219
psd	-0.08	0.0332	-0.145	-0.015	5.845	1	0.016
lw	0.26	0.1776	-0.088	0.608	2.147	1	0.143
D	-0.055	0.0612	-0.174	0.065	0.793	1	0.373
Dispersion	1.543	0.3642	0.971	2.45			

LOS D Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-1.18	4.374	-9.753	7.393	0.073	1	0.787
vol	0	0.0003	0	0.001	1.829	1	0.176
psd	-0.096	0.0589	-0.212	0.019	2.684	1	0.101
lw	0.394	0.3982	-0.386	1.175	0.981	1	0.322
D	-0.033	0.0796	-0.189	0.123	0.175	1	0.676
Dispersion	1.53	0.5369	0.769	3.043			

LOS E Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-84.21	5.4439	-94.879	-73.54	239.277	1	0
vol	0	0.0006	-0.002	0	2.443	1	0.118
psd	0.269	0.1174	0.039	0.499	5.254	1	0.022
lw	0	.	.	.	.	.	0
D	0.079	0.0753	-0.069	0.226	1.095	1	0.295
Dispersion	1.624	0.5605	0.826	3.194			

LOS F Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-9.157	1.7462	-12.58	-5.735	27.501	1	0
vol	0	0.0001	7.61E-05	0.001	6.318	1	0.012
psd	0.063	0.0497	-0.034	0.161	1.627	1	0.202
lw	0.326	0.0931	0.143	0.508	12.24	1	0
D	-0.002	0.0011	-0.004	0	3.002	1	0.083
Dispersion	0.541	0.1413	0.324	0.903			

LOS's E & F Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-4.341	1.5927	-12.873	-6.63	37.486	1	0
vol	0	0.0001	0	0.001	15.647	1	0
psd	0.048	0.0401	-0.03	0.127	1.446	1	0.229
lw	0.396	0.0866	0.226	0.566	20.921	1	0
D	-0.002	0.0011	-0.004	0	3.044	1	0.081
Dispersion	0.786	0.1614	0.525	1.175			

Negative Binomial SPSS 16.0 Output Estimates for MODEL II

LOS A Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	0.946	1.4232	-1.843	3.736	0.442	1	0.506
vol	0.003	0.0008	0.001	0.004	10.886	1	0.001
sw	-0.075	0.0293	-0.132	-0.017	6.528	1	0.011
lw	-0.049	0.1246	-0.293	0.195	0.155	1	0.693
D	-0.518	0.1433	-0.799	-0.237	13.085	1	0
Dispersion	2.954	0.4907	2.134	4.091			

LOS B Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	4.229	2.037	0.237	8.222	4.311	1	0.038
vol	0	0.0004	0	0.001	0.155	1	0.694
sw	0.011	0.0342	-0.056	0.078	0.097	1	0.756
lw	-0.295	0.1632	-0.615	0.025	3.266	1	0.071
D	-0.116	0.0955	-0.304	0.071	1.485	1	0.223
Dispersion	1.943	0.4353	1.252	3.014			

LOS C Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-3.79	2.5268	-8.742	1.163	2.249	1	0.134
vol	0	0.0002	-7.19E-05	0.001	2.791	1	0.095
sw	-0.063	0.0265	-0.115	-0.011	5.594	1	0.018
lw	0.334	0.1986	-0.055	0.723	2.826	1	0.093
D	-0.076	0.0636	-0.2	0.049	1.416	1	0.234
Dispersion	1.541	0.3648	0.969	2.451			

LOS D Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-26.315	19.227	-63.999	11.37	1.873	1	0.171
vol	0.002	0.0008	-8.59E-05	0.003	3.446	1	0.063
sw	-0.237	0.1259	-0.484	0.01	3.545	1	0.06
lw	2.087	1.6089	-1.066	5.241	1.683	1	0.194
D	-0.149	0.1217	-0.387	0.09	1.494	1	0.222
Dispersion	1.28	0.4815	0.613	2.676			

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-82.688	2.9782	-88.525	-76.851	770.868	1	0
vol	0.002	0.0014	-0.001	0.004	1.495	1	0.221
sw	-0.286	0.2544	-0.784	0.213	1.262	1	0.261
lw	0	.	.	.	.	.	0
D	-0.126	0.1457	-0.411	0.16	0.744	1	0.388
Dispersion	1.834	0.6137	0.952	3.534			

LOS F Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-8.682	1.5087	-11.639	-5.725	33.112	1	0
vol	0.001	0.0001	0	0.001	18.705	1	0
sw	-0.048	0.0394	-0.125	0.03	1.467	1	0.226
lw	0.477	0.0954	0.29	0.663	24.971	1	0
D	-0.003	0.0011	-0.005	0	7.533	1	0.006
Dispersion	0.543	0.142	0.326	0.907			

LOS E &F Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-9.739	1.2849	-12.257	-7.22	57.443	1	0
vol	0.001	0.0001	0	0.001	35.23	1	0
sw	-0.066	0.0321	-0.129	-0.003	4.245	1	0.039
lw	0.56	0.0893	0.385	0.735	39.292	1	0
D	-0.003	0.0012	-0.006	-0.001	8.44	1	0.004
Dispersion	0.778	0.1601	0.52	1.164			

REFERENCES

"*Analysis of Accident Statistics*," Permit Writers Workshop. http://kiewit.oregonstate.edu/pdf/pw_accid.pdf (accessed January 2011).

"Iowa DOT High Crash Locations Ranking Procedure." IOWA DOT. http://www.ctre.iastate.edu/research/hcl/documents/ranking.htm (accessed January 2011).

AAA Foundation for Traffic Safety. *2010 Traffic Safety Culture Index.* Washington, DC. AAA Foundation for Traffic safety. 2010.

ASHTO, 2001. A Policy on Geometric Design of Highways and Streets, Washington , DC.

AASHTO, 2010. An Introduction to the Highway Safety Manual, American Association of State Highway and Transportation Officials, Washington, DC. http://www.highwaysafetymanual.org (accessed December 2010)

AASTO, 2010. Highway Safety Manual, American Association of State Highway and Transportation Officials, Washington, DC. http://www.highwaysafetymanual.org/Documents/HSMP-1.pdf (accessed February 2011)

Abdel-Aty, M., Dilmore, J. and A. Dhindsa. Evaluation of variable speed limits for real-time freeway safety improvement. *Accident Analysis and prevention*. Vol. 38, No.3, 2005, pp. 335-345.

Abdel-Aty, M.A., and A.E. Radwan. Modeling Traffic Accident Occurrence and Involvement. *Accident Analysis and Prevention,* Vol. 32, No. 5, 2000, pp. 633-642.

Addel-Aty, M. A. and R. Pemmanabonia. Calibrating a real-Time Traffic Crash-Prediction Model Using Archived Weather and ITS Traffic Data. *IEEE Transactions on Intelligent Transportation Systems*. Vol. 7, 2006, pp. 167-174.

Agent, K. R., Stamatiadis, N. and S. Jones. *Development of Accident Reduction Factors*. KTC-96-13. KTC, Kentucky Transportation Cabinet, 1996.

Bahar, G. *Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the Highway Safety Manual*. TRB E-C142, Transportation Research Board, Washington, DC, 2010.

Bonneson, J. A. and M. P. Pratt. *Roadway Safety Design workbook*. Publication FHWA/TX-09/0-4703-P2. FHWA, Texas Department of Transportation, 2009.

Bonneson, J., Zimmerman, K., and K. Fitzpatrick. *Interim Roadway Safety Design workbook*. Publication FHWA/TX-06/0-4703-P4. FHWA, Texas Department of Transportation, 2006.

Chen, H., Liu, P., Lu, J. J. and B. Behzadi. Evaluating the Safety Impacts of the Number and Arrangements of Lanes on Freeway Exit Ramps: *Accident Analysis and Prevention*. Vol. 41, No. 6, 2009, pp. 543-551.

Christoforou, Z. Cohen, S. and M. G. Karlaftis. Vehicle Occupant Injury Severity on Highways: An Empirical Investigation. *Accident Analysis and Prevention*. Vol. 42, No. 6, 2010, pp. 1606 – 1620.

Council, F. M., Reurings, M., Srinivasan, R., Masten, S. and D. Carter. *Development of a speeding-Related Crash Typology*. FHWA-HRT-10-024, Federal Highway Administration, McLean, VA. 2010.

Council, F.M., and J.R. Stewart. *Safety Effects of the Conversion of Rural Two-Lane to Four-Lane Roadways Based on Cross-Sectional Models*. FHWA-990327. FHWA, U.S. Department of Transportation, Washington DC. 2000.

Cruzado, I. and E. T. Donell. Factors Affecting Driver Speed Choice along Two-lane Rural Highway Transition Zones. *Journal of Transportation Engineering,* Vol 136, 2010, No. 8.

Daniel, J. R. and E. Maina. Relating Safety and Capacity on Urban Freeways. *Procedia - Social and Behavioral Science*. 6th International Symposium on Highway Capacity and Quality of Service, Stockholm, Sweden. 2010.

Dissnanayake, S. and L. Liu. *Speed Limit-Related Issues on Gravel Roads*. Kansas Department of Transportation, K-TRAN: KSU-06-5, Kansas Department of Transportation, Topeka, KS. 2009.

FHWA. *Crash Modification Factors Clearinghouse*. FHWA-SA-10-008, U.S Department of Transportation, Federal Highway Administration, Washington, DC, 2010.
http://www.cmfclearinghouse.org/ (accessed August, 2011).

FHWA. *The U.S. Department of Transportation Rural Safety Initiative*. U.S Department of Transportation, Federal Highway Administration, Washington, DC, 2008. http://www.dot.gov/affairs/ruralsafety/ruralsafetyinitiativeplan.htm (accessed December, 2009).

FHWA. *Transportation Planer's Safety Desk Reference*. FHWA-HEP-10-001, Department of Federal Highway Administration Office of safety, Washington, DC, 2010.

Fitzpatrick, K., Zimmerman, K., Bligh, R., Chrysler, S. and B. Blaschke. *Criteria for High Design Speed Facilities*. FHWA/TX-07/0-5544-1. FHWA, Texas Department of Transportation, Austin, TX., 2007.

FMCSA, 2006. Large Truck Crash Overview. Analysis Division Federal Motor Carrier Safety Administration U.S. Department of Transportation. Washington, DC.

Garber, N., and L. A. Hoel. *Traffic & Highway Engineering,* 4th ed. Cenage Learning, Toronto, Canada, 2009.

Garber, N. J., Haas, P. R. and Gosse. Development of Safety Performance Functions for Two-Lane Roads Maintained by the Virginia Department of Transportation. Final Report VTRC 10-R25. *Virginia Transportation Research Council*. Virginia Department of Transportation, Richmond, VA. 2010.

Garber, N.J. *The Effect of Speed, Flow, and Geometric Characteristics on Crash Rates for Different Types of Virginia Highways.* Virginia Transportation Research Council, Richmond, VA. 2000.

Gross, F. Persaud, B. and C. Lyon. *A Guide to Developing Quality Crash Modification Factors*. FHWA-SA-10-032, Federal Highway Administration, Washington, DC, 2010.

Gross, F., Jovanis, P. P. and K. Eccles. Safety Effectiveness of Lane and Shoulder Width Combinations on Rural, Two-Lane, Undivided Roads. In *Transportation Research Record*: *Journal of the Transportation Research Board*, No. 2103, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 42–49.

Gross, F., Persaud, B. and C. Lyon. *A Guide to Developing Quality Crash Modification Factors*. FHWA-SA-10-032. FHWA. US Department of Transportation, 2010.

Hallmark, S., McDonald, T. and R. Sperry. *Evaluation of Rumble Stripes on Low-Volume Rural Roads in Iowa-Phase I*. IHRB Project TR-577. IHRB, Iowa Department of Transportation, Ames, IA., 2009.

Hauer, E., Bonneson, J. A., Council, F., Srinivasan, R. and C. Zegeer. Crash Modification Factors - Foundational Issues. *Transportation Research Record: Journal of the Transportation Research Board, No. 2279,* Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 67–74. DOI: 10.3141/2279-08.

Ivan, J. N., Garrick, N. W. and G. Hanson. *Designing Roads that Guide Drivers to Choose Safer Speeds*. Connecticut Department of Transportation , JHR 09-321, Connecticut Department of Transportation, Rocky Hill, CT, 2009.

Ivan, J. N., Wang, C. and Bernardo, N. R. Explaining Two-Lane Highway Crash Rates Usingand Use and Hourly Exposure. *Accident Analysis and Prevention,* Vol. 32, No. 5, 2006, pp. 787-795.

Jones, J. C., Knopp, M. C., Fitzpatrick, K., Doctor, M. A., Howard, C. E., Laragan, G. M., Rosenow, J. A., Struve, B. A., Thrasher, B. A. and E. G. Young. *Freeway Geometric Design for Active Traffic Management in Europe*. FHWA-PL-11-004. American Trade Initiatives. Alexandria, VA. 2011.

Jones, K., Yunk, K. and D. Carter. *The CMF Clearinghouse: A Handy Safety Tool*. FHWA-HRT-11-001. FHWA. US Department of Transportation, 2010.

Jurewicz, C. and B. Thompson. Crash risk estimation and assessment tool. *ARRB Group Limited*.Victoria, Australia. 2010.

Kononov, J., Bailey, B., and B. K. Allery. Relationships between Safety and Both Congestion and Number of Lanes on Urban Freeways. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2083, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 26–39.

Kononov, J., Bailey, B., and B. K. Allery: Exploratory Analysis of Relationship between the Number of Lanes and Safety on Urban Freeways. Presented at 87th Annual meeting of the Transportation Research Board, Washington, D.C., 2008.

Korkut, M. *Freeway Crash and Lane Compliance under Truck Lane Restriction and Differential Speed Limit*. Master's Thesis. Louisiana State University, Department of Civil Engineering. Baton Rouge. 2009.

Kuhn, B., Balke, K., Chaudhary, N., Jasek, D., Karkee, G., Obeng-Boampong, K., Shelton, J.and S. Venglar. *Managed Lanes Strategies Feasible for Freeway Ramp Applications*. FHWA/TX-08/0-5284-2. Texas Department of Transportation,. Austin, TX. 2008.

Kutner, M. H., Nachtsheim, C. J. and J. Neter. *Applied Linear Regression Models*. Fourth Edition. McGraw Hill Companies, Inc., New York, 2004.

Lee, C., Hellinga, B. and F. Saccomanno. Evaluation of variable speed limits to improve traffic safety. *Accident Analysis and prevention*. Vol. 14, Part C, 2006, pp. 213-228.

Lee, J. and B. Park. Lane Flow Distribution on Basic Segments of Freeways under Different Traffic Conditions. Presented at 89th Annual meeting of the Transportation Research Board, Washington, D.C., 2010.

Liu, P., Chen, H., Lu, J. J., and Cao, B. How Lane Arrangements on Freeway Mainlines and Ramps Affect Safety of Freeways with Closely Spaced Entrance and Exit Ramps. *Journal of Transportation Engineering*, Vol. 136, 2009, No. 7.

Lu, J., Haleem, K., Gan, A. and W. Wu. Application of Fisher's Clustering Method to Improve Empirical Bayes Estimation for the Identification of High Crash Locations. Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.

Martin, J. L. Relationship between Crash Rate and Hourly Traffic flow on Interurban Motorways. *Accident Analysis and Prevention,* Vol. 34, No. 5, 2001, pp. 619-629.

Mbatta, G. Developing Crash Modification Factors for Urban Highway with Substandard Wide Curb Lane. *Ph. D. Dissertation*. Florida State University. Tallahassee, FL. 2011.

Milton, J., and F. Mannering. The Relationship among Highway Geometries, Traffic-Related Elements and Motor-Vehicle Accident Frequencies, *Transportation*, Vol. 25, No. 4, 1998, pp. 395-413.

Mulinazzi, T. E., Schrock, D. S., Rescot, R., Jasrotia, R., Hovey, K., Li, Y. and J. Pohlman. *Mitigating Wind Induced Truck Crashes*. 25-1121-0001-263. MATC TRB RIP No. 18467, Mid-America Transportation Center, U.S. Department of Transportation, Lincoln, Nebraska, 2009.

Mussa, R. and D. Chimba. Analysis of Crashes Occurring on Florida six-lane Roadways. *Transport Research Laboratory*: Advances in Transportation Studies., Vol. 8, No.B, 2006, pp. 43-56.

Mussa, R. and D. Chimba. Analysis of Crashes Occurring on Florida Six-Lane Roadways. *Advances in Transportation Studies*. Vol. 8, 2006, pp. 43-56.

Navidi, W. *Statistics for Engineers and Scientists*. Second Edition. McGraw Hill Companies, Inc., New York, 2008.

NCHRP, 2008. Report 617. Accident Modification Factors for Traffic Engineering and ITS Improvements, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

Nise, N. S. *Control Systems Engineering*. Fourth Edition. John Wiley & Sons, Inc., Hoboken, NJ, 2004.

Noland, R.B, and L. Oh. The Effect of Infrastructure and Demographic Change on Traffic-Related Fatalities and Crashes: A Case Study of Illinois County-Level Data. *Accident Analysis and Prevention*, Vol. 35, No. 4, 2004, pp. 525-532.

Pande, A., Abdel-Aty M. and Das, A. A Classification Tree Based Modeling Approach for Segment Related Crashes on Multilane Highways. *Accident Analysis and Prevention*. Vol 38 391-398. 2000.

Park, B., Fitzpatrick, K., and D. Lord: Evaluating the Effects of Freeway Design Elements on Safety. Presented at 89th Annual meeting of the Transportation Research Boar, Washington, D.C., 2010.

Park, B., Fitzpatrick, K., and D. Lord: Evaluation of the Safety Effects of Lane Arrangements on Freeway Segments with Closely Spaced Entrance and Exit Ramps. Presented at 89th Annual meeting of the Transportation Research Boar, Washington, D.C., 2010.

Poole, R. W. *When Should we Provide Separate Auto and Truck Roadways?* Discussion Paper No. 2009-24, International Transport Forum, Reason Foundation, Los Angeles, 2009.

Qin, X., Ivan, J. N. and N. Ravishanker. Selecting Exposure Measures in Crash Rate Prediction for Two-Lane Highway Segments. *Accident Analysis and Prevention,* Vol. 36, No. 3, 2004, pp. 183-191.

Qin, X., Ivan, J. N., Ravishanker, N., Liu, J. and Tepas D. Bayesian Estimation of Hourly Exposure Functions by Crash Type and Time of Day. *Accident Analysis and Prevention,* Vol. 38, No. 5, 2000, pp. 1071-1080.

Roess, R. P., Prassas E. S., and W. R. McShane. *Traffic Engineering*. Pearson Prentice Hall Inc., Upper Saddle River, NJ, 2004.

Sarvi, M. Measuring and Modeling the Dynamic of Heavy commercial Vehicles and their Interactions with other Vehicles: AHB40 Highway Capacity and Quality of Service. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.

Savolainen, P. T., and Ghosh, I. *Using Detector Data to Identify and Examine Crashes and Incidents on Freeways*. 051WY02. USDOT. USDOT Region V Regional Transportation Center Final Report. Detroit, MI, 2010.

Sisiopiku, V. P., Sullivan, A. and G. Fadel. Implementing Active Traffic Management Strategies in the U.S. In UTCA, Report No. 08206, Transportation Research Board of the National Academies, Washington, D.C., 2009.

Stamatiadis, N., Pigman, J., Sacksteder, J., Ruff, W. and D. Lord. *Impact of Shoulder Width and Median Width on Safety*. NCHRP Report 633. National Cooperative Highway Research Program. Transportation Research Board, Washington, D.C., 2009.

Stavrinos, D., Jones, J.L., Impact of Distracted Driving on Safety and Traffic Flow, *Accident Analysis and Prevention* (2010), doi:10.1016/j.aap.2013.02.003

Strathman, J. G., K. J. Dueker, J. Zhang, and T. Williams. *Analysis of Design Attributes and Crashes on the Oregon Highway System. FHWA-OR-RD-02-01. Portland State University, Portland, OR, 2001.*

Tegge, R. A., Jo, J. and Y. Ouyang. Development and Application of Safety Performance Functions for Illinois. Research Report ICT-10-066. *Illinois Center for Transportation*. Illinois Department of Transportation, Springfield, IL. 2010.

Torbic, D. J., Harwood, D. W., Gilmore, D. K. and K. R. Richard. *Interchange Safety Analysis Tool (ISAT): User Manual*. FHWA-HRT-07-045. FHWA, U.S. Department of Transportation, 2007.

TRB, 2010. Highway Capacity Manual, Transportation Research Board, Washington, DC.

Vogt, A., and J. G. Bared. Accident Models for Two-Lane Segments and Intersections. *Transportation Research Record1653*. Washington, DC, 1998.

Waller, S. T., Ng, M. W., Ferguson, E., Nezamuddin, N. and D. Sun. *Speed Harmonization and Peak-period Shoulder Use to Manage Urban Freeway Congestion*. FHWA/TX-10/0-5913-1. FHWA. Texas Department of Transportation, Austin, TX., 2009.

Washburn, S. S., Yin, Y., Modi, V. and A. Kulshrestha. *Investigation of Freeway Capacity: a) Effect of Auxiliary Lanes on Freeways Segments Volume Throughput and b) Freeway Segment Capacity Estimation for Florida Freeways.*TRC-FDOT-73157-2010, Florida Department of Transportation, 2010.

Wiegmann, D. A., von Thaden, T. L. and A. M. Gibbons. A Review of Safety Culture Theory and its Potential Application to Traffic Safety. *Foundation for Traffic Safety*. American Automobile Association. Heathrow, FL, 2007.

Xu, T., Sun, X., Zhang, G. and Y. Chen: Impacts of Large Vehicles on Freeway Speed and Safety. Presented at 88th Annual meeting of the Transportation Research Board, Washington, D.C., 2009.

Yanmaz-Tuzel, O. and K. Ozbay. A Comparative Full Bayesian before-and-after Analysis and Application to Urban Road Safety Countermeasures in New Jersey. *Accident Analysis and Prevention,* Vol. 42, No. 6, 2010, pp. 2099-2107.