2014

Bayesian analysis of factors affecting crash frequency and severity during winter seasons in Iowa

Mohammad Saad Bin Shaheed
Iowa State University

Bayesian analysis of factors affecting crash frequency and severity during winter seasons in Iowa

by

Mohammad Saad B. Shaheed

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Konstantina "Nadia" Gkritza, Co-Major Professor
Shauna L. Hallmark, Co-Major Professor
Alicia L. Carriquiry
Jiangping Zhou
Jing Dong
Keith K. Knapp

Iowa State University

Ames, Iowa

2014

DEDICATION

To my beloved parents and wife

This Dissertation is dedicated to my beloved parents, Umme Salma and Ebney Hyder Mohammad Shaheed, and my wife, Tabassum Hayet, without whose caring support, inspiration, and motivation, it would have not been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

AWOS	Automated Weather Observing System
DOT	Department of Transportation
EB	Empirical Bayes
MCMC	Markov Chain Monte Carlo
PSI	Potential for Safety Improvement
RTM	Regression to Mean
RWIS	Roadway Weather Information System

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Drs. Konstantina Gkritza and Shauna L. Hallmark, and my committee members, Dr. Alicia Carriquiry, Dr. Keith K Knapp, Dr. Jing Dong, and Dr. Jiangping Zhou, for their guidance and support throughout the course of this research.

I would also like to express my special gratitude to Dr. Konstantina Gkritza for her valuable advice, guidance, and continuous support during the course of my doctoral program. Her continuous inspiration was the driving force for me to excel both technically and professionally. I would also like to extend my appreciation to Dr. Shauna L. Hallmark for her valuable advice and encouragement. I would like to thank Dr. Alicia Carriquiry, Professor at the Department of Statistics, for the statistical advice and help. My heartfelt gratitude to Zachary Hans, Research Engineer at the Institute for Transportation for providing me with the crash data needed for this research. I would also like to thank Inya Nlenanya and Anindya Das for helping me with the data processing. I would also like to extend my gratitude to my dearest friend in Ames, Md. Shaidur Rahman for his moral support during my Ph.D. years.

I would like to thank my family for their encouragement. This dissertation would not have been possible without the unconditional love and support from my siblings, Shammi Sharmin and Mohammad Noor Shaheed. Finally, I would like to express my deepest gratitude to my wife for her hours of patience, respect and love.

ABSTRACT

Traffic safety during winter seasons has been a serious concern in Iowa as hundreds of people are injured on Iowa's highways each winter. As the goal of the state transportation agency is to ensure the mobility of road users without compromising the safety during winter periods, it is important to understand the factors affecting winter-weather crash frequency and occupant injury risk through quantitative prediction models. It is of utmost importance to identify locations prone to winter-weather crashes to utilize the limited resources efficiently for improving safety during winter conditions. This research intended to develop a systematic prioritization technique to identify winter-weather crash hotspots by using Empirical Bayes technique that addresses the serious limitations of the traditional methods to screen road networks for identifying high crash locations. This research also addresses the issue of hierarchical structure in the crash data by developing quantitative models to predict occupant injury risk for crashes occurring during winter seasons to obtain unbiased and accurate estimation of the parameters for better management of road safety during winter seasons.

Along with developing site prioritization techniques for identifying roadway segments with potential for safety improvement through traditional statistical methods using raw crash data, Empirical Bayes technique is used to screen roadway segment through developing safety performance functions for winter-weather crashes. A novel approach is adopted to extract weather data from information reported by winter maintenance crew members to incorporate weather related factors in developing safety performance functions at network level for three roadway types in Iowa. Weather factors such as visibility, wind velocity, air temperature are found to have statistically significant effects on winter-weather crash frequency. The ranking of

roadway segments based on Potential for Safety Improvement (PSI) by employing Empirical Bayes technique differs from the ranking produced by simple crash frequency. Safety Performance Functions developed in this research can be used to produce ranking based on PSI by using crash observations made over a specific number of years for winter-weather crashes. Models predicting occupant injury risk with binomial logit formulation are developed considering the hierarchical structure of the crash data in a Bayesian framework in this research for weather-related crashes, non-weather related crashes, and all crashes occurring during the four winter seasons (2008/09 to 2011/12) in Iowa. These models are developed using disaggregate crash data with occupants nested within crashes. High values of between-crash variance for the three models underscore the justification of considering the hierarchical nature of the crash data due to the natural crash data collection process. Factors related to occupants (gender, seating position, trap status, ejection status, airbag deployment, safety equipment used) had statistically significant effects on occupant injury risk for all the models. Weather-related variables such as visibility and air temperature were found significant predictors of all crashes and weather-related crashes during the winter seasons. The variable representing road surface condition is also found to be a significant factor in all three models developed to predict occupant injury risk during the winter seasons.

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

Winter weather safety has been a serious concern for transportation agencies and road users in countries with severe winters such as the United States (U.S.), Canada, and Northern European countries. Transportation agencies spend significant amount of resources every year to keep highways and local roads clear of snow and ice to ensure the mobility of the road users without compromising on safety during the winter weather periods.

Literature (Nilsson and Obrenovic, 1998) shows that drivers are twice more likely to be involved in a crash in winter than summer for a given distance of travel. In the United Kingdom and the United States of America, weather related crashes account for 30% and 35% of total reported crashes (Andrew and Bared, 1998). According to Federal Highway Administration (FHWA) (2010), weather related crashes accounts for an average of 24% of total crashes resulting in about 7,400 fatalities and over 673,000 injuries annually. Thus it is important to implement effective winter weather road maintenance operations, such as plowing, salting, and sanding.

It has been estimated that the U.S. spends around $2 billion annually for winter weather maintenance and operations. This estimate does not include indirect costs such as damage occurring to environment especially in the form of salinization of land, ground and surface water, roadway infrastructure and vehicles due to salt use (Environment Canada, 2002). Given the limited resources available to agencies for winter weather operations, it is imperative to establish a prioritization methodology in terms of frequency (or occurrence) and severity of winter weather crashes to rank high risk road segments for improving safety to ensure an efficient allocation of resources. On the other hand, it is also critical to identify the causes for traffic crashes that occur during winter weather, and in particular, understand the factors affecting traffic crash frequency/severity during winter weather conditions. Identifying and quantifying the relationship between winter road safety and contributing factors can be achieved with the use of advanced statistical modeling for predicting the frequency of winter weather crashes or crash severity as a function of various significant factors (vehicles, roadway, environment, weather, occupants, maintenance and operations).

1.2 Research Motivation

The motivation of the current research derives from the fact that there are hundreds of people injured in winter-weather crashes on Iowa's highways each winter. In this research, winter weather-related crashes were defined as weather-related crashes occurring during the winter seasons with any of the followings reported for the crash event:

- Weather conditions: Sleet/hail/freezing rain or Snow or Blowing sand/soil/dirt/snow
- Surface conditions: Ice or Snow or Slush
- Vision obscurement: Blowing sand/soil/dirt/snow

About one-third of the total crashes and half of the rural interstate crashes were weather-related during the winter seasons of 2004/05 to 2011/12 in Iowa. Figure 1.1 shows the percentage of weather-related crashes during the four winter seasons (2008/09 to 2011/12) with respect to the total number of crashes during the same time period in Iowa. Though transportation agencies spend millions of dollars on proactive and reactive maintenance for ensuring the best possible pavement condition and visibility for traveling motorists, there is no systematic method to identify high risk winter weather-related crash locations in Iowa. Previous efforts to identify potentially problematic winter-related crash locations were based on using historical crash data during winter weather conditions. Traditional naïve statistical methods using raw crash data to identify crash hotspots have serious limitations, including the so called regression-to-mean (RTM) problem in highway safety. RTM refers to the tendency of the unusually high or low crashes in one time period to regress or return to the mean in subsequent time periods. For example, if an agency selects a location with an unusually high number of winter weather-related crashes over a short time period for targeted winter maintenance, it doesn’t necessarily mean that site has a higher than normal crash frequency. This will result in inefficient allocation of resources without the expected improvement in winter road safety with respect to the winter maintenance. Traditional methods to identify hotspots for improving winter road safety also do not consider incorporating weather information while selecting locations for targeted winter maintenance. At the same time, it is important to understand the factors contributing to or causing severe injuries to travelers during winter weather season to mitigate the causes of severe injury outcomes.

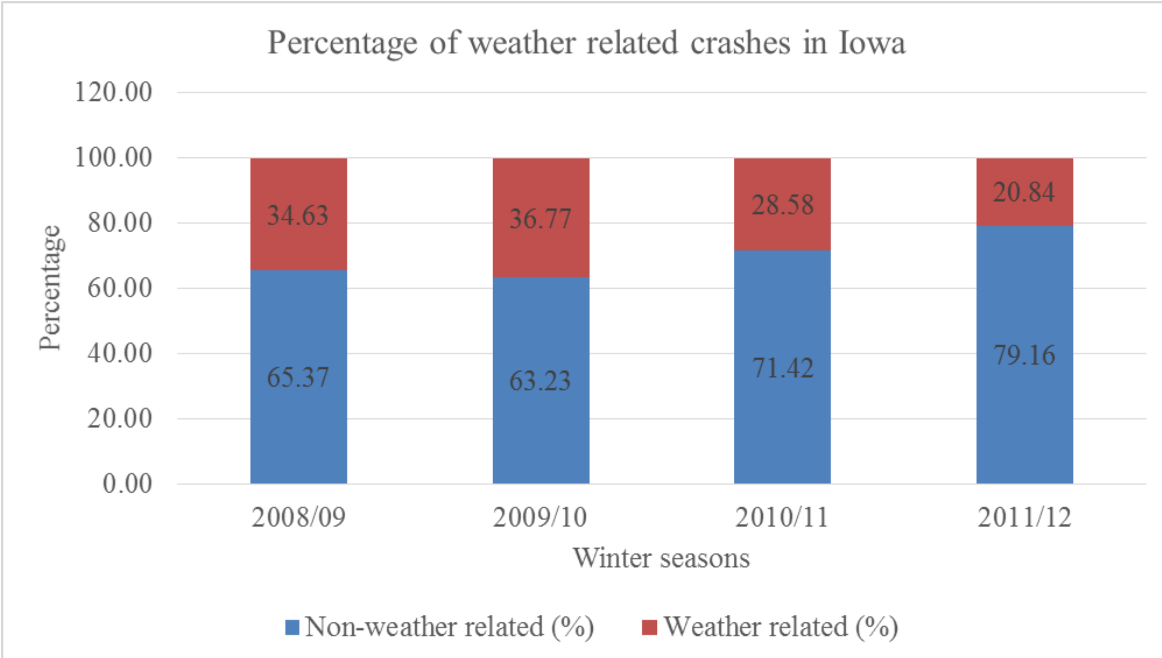


Figure 1.1 Percentage of weather-related crashes during four winter seasons (2008/09 to 2011/12)

1.2.2 Research Objectives

In recent years, techniques for screening road networks to identify high crash locations have become more sophisticated and require more data as inputs. Instead of relying on the traditional methods (crash frequency, rate, or severity) to identify candidate locations for safety improvements, this study aimed to utilize an Empirical-Bayes adjusted crash frequency to identify crash locations that have a potential for crash reduction during winter weather conditions. The Empirical-Bayes approach combines the expected number of crashes with the observed crash counts at a location to produce an improved estimate of the expected number of crashes. As crashes are random in nature, the Empirical-Bayes method takes into account the phenomenon of RTM. Extensive research has shown that the Empirical-Bayes approach is the most consistent and reliable method for identifying sites with promise (Cheng and Washington, 2005; Cheng and Washington, 2008; Hauer, 1996; Hauer et al., 2002).

Safety Performance Functions (SPF) need to be developed for applying Empirical-Bayes technique to screen road segments. SPFs are statistical models used to estimate the average crash frequency for a specific site type, based on traffic volume and roadway segment length. Typical safety performance functions (SPFs) have been developed to estimate crash frequency using site or roadway characteristics such as lane width and traffic volume expressed as annual average daily traffic (AADT). These SPFs did not incorporate additional independent variables as this process would be more complex and labor intensive. The proposed research will develop crash frequency models for three roadway types in Iowa to predict winter weather crashes with the crash frequency being a function of several factors related to winter weather conditions such as visibility, pavement temperature, air temperature, and wind speed.

The second objective of the proposed research was to develop statistical models to predict severity of winter weather crashes using crash data information along with winter weather related attributes such as wind speed, pavement temperature, and winter storm type. Traditional crash prediction models such as generalized linear regression models are not able to capture the multilevel structure of the traffic crash data. The underlying assumption for these types of traditional models is that the observations are sampled from a single homogenous population and each crash observation is an independent situation leading to independent residuals. However, this "independence" assumption may often not hold true since multilevel data structures exist extensively because of the crash data collection process. As traffic crash data are hierarchical in nature with possible correlation at the occupant or vehicle level, ignoring such within-class correlation might lead to the development of prediction models with biased parameter estimates and variables to be falsely significant. Because of the hierarchical nature of traffic crash data, it is reasonable to assume that the characteristics of the vehicle in which occupants are traveling affect the probability of casualties of occupants. In this case, casualties within the same vehicle would tend to be similar compared to casualties in different vehicles.

Similarly, vehicles/drivers involved in a multi-vehicle crash event might sustain damages/injuries with similar severity or vice versa. Whether positive or negative, ignoring such intra-class correlation in crash data is equivalent to ignoring the clustering nature of the data. When this correlation in the hierarchical crash data is ignored, the “independence” assumption which was mentioned previously fails to hold true. With crash prediction model being one of the most important techniques to investigate the relationship between crash occurrence/crash severity and risk factors associated with the various traffic entities, it is very important to obtain an unbiased and accurate estimation of parameter estimates to predict crash occurrence/crash severity for better management of road/winter weather safety. The use of hierarchical/multilevel models is one way of addressing the multilevel structure or the clustered nature of the crash data. Multilevel models have the ability to represent datasets with hierarchical or clustered nature. Multilevel models are likely to involve random effects defined over the clusters and possible correlation between different types of cluster groups. This research dedicates effort to develop a hierarchical discrete choice model in a Bayesian framework to consider the multilevel structure of the data.

1.3 Research Contribution

Although the effect of winter weather on road safety has been studied extensively by many researcher, studies related to development of a systematic prioritization technique for screening road segments for winter weather crashes are scarce in the literature. At the same time, studies addressing the hierarchical/clustered nature of the crash data in a Bayesian framework are also limited in the transportation literature. The contribution of this dissertation are as follows:

1. Development of statistical models relating winter crash frequency to weather-related information, roadway characteristics, and traffic volumes for three types of roadways
2. Development a systematic prioritization technique for screening road segments for winter weather crashes in an Empirical-Bayes framework using the crash frequency models.
3. Understanding the factors affecting winter weather crash-injury severity by developing severity models in a hierarchical Bayesian framework so as to consider the multilevel structure of the crash data on a micro scale (crash – vehicle/driver – occupant)

The following schematic shows the research framework.

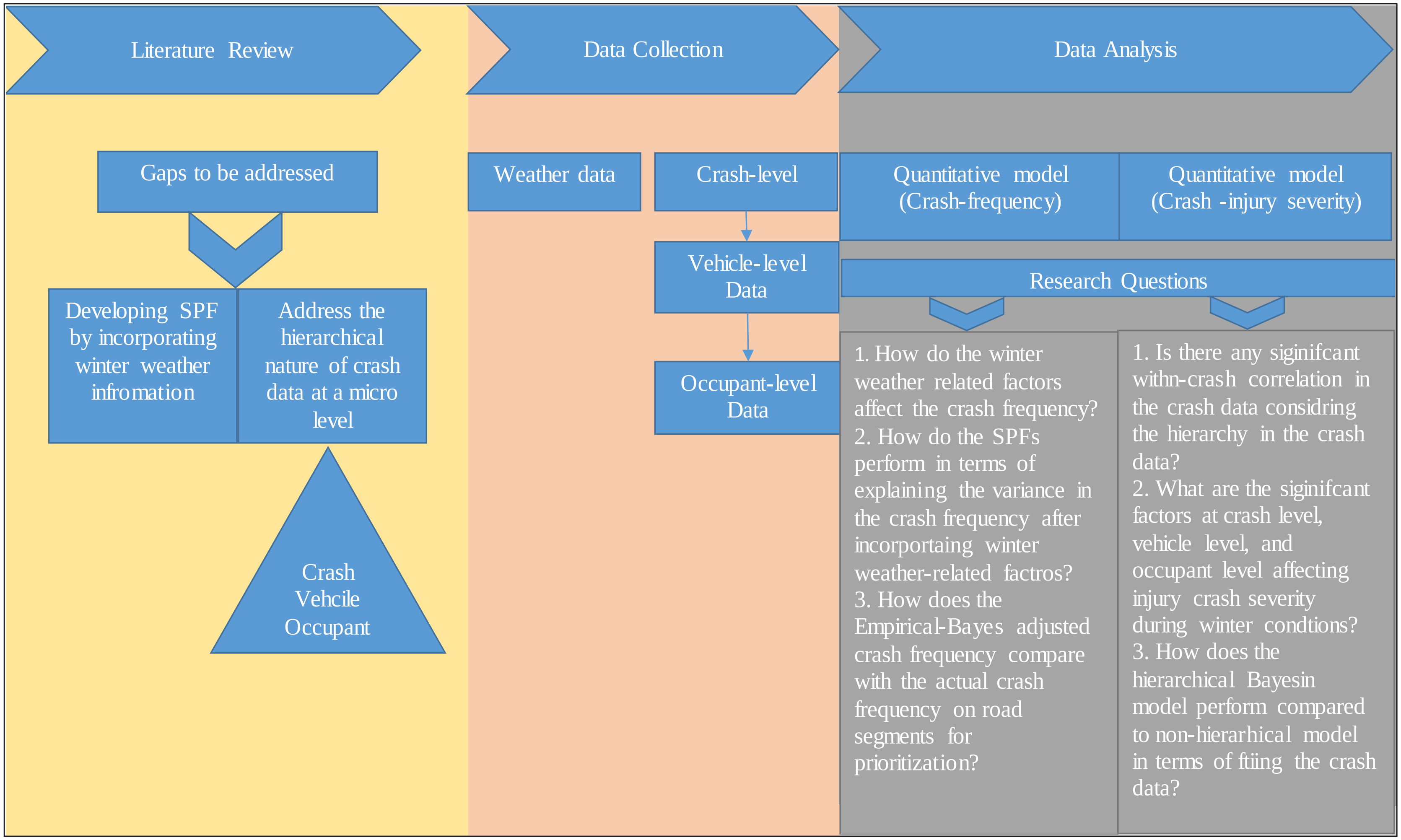


Figure 1.2 Research framework for the study

1.4 Organization of the Dissertation

This dissertation consists of seven chapters. Chapter 1 was the introduction to the problem. The remaining portion of the dissertation is organized as follows:

In Chapter 2, a literature review is presented in the area of factors affecting winter weather safety, crash frequency models, and multilevel modeling techniques affecting crash-injury severity.

Chapter 3 describes the data used for the analysis and data processing.

Chapter 4 describes the proposed methods used for achieving the research objectives.

Chapter 5 describes the development and calibration of safety performance functions used in the Empirical-Bayes analysis and the corresponding estimation results.

Chapter 6 describes the development of multilevel models for crash-injury severity analysis and the corresponding model estimation results.

Chapter 7 highlights the conclusions and main contributions of this research. A discussion of the research limitations and ideas for future research is also presented in this chapter.

CHAPTER 2

LITERATURE REVIEW

This chapter provides a thorough literature review of the factors affecting winter weather safety and methodologies that were used in previous studies to analyze crashes related to winter weather. This chapter also discusses the studies related to multilevel modeling techniques for crash injury severity analysis.

2.1 Factors Affecting Winter Weather Safety

2.1.1 Effect of Weather on Safety

From the literature, it has been found that some of the weather related factors affecting safety on the roadway include snow, rain, freezing rain, severe and major storms, temperature, visibility, and wind speed (Edwards 1998; Feng 2001; US Department of Commerce 2002; Strong et al 2010).

Andreescu and Frost (1998) conducted correlation and regression analysis of daily accidents with weather related variables (temperature, rain fall, and snowfall) using three years of crash data (1990-1992) from Montreal, Quebec. Differences in daily number of crashes and mean number of crashes over a week was used as the number of daily crashes for the three years of study period to reduce the variation in the number of crahses per day. The study results found that the number of crashes increased with increase in snowfall or rainfall intensity but no significant relationship with respect to temperature was found.

Aggregated data by intervals of six hours was used by Andrey et al. (2003) to analyze crash and precipitation data of six Canadian cities from 1995 to 1998 employing matching pair technique. Using this technique, the researchers compared crashes on periods of days under adverse weather conditions with crashes on periods of similar days under normal weather conditions. The results indicated 75% and 45% increase in frequency of overall collisions and injury severity collisions respectively due to precipitation. Although snowfall effects were more pronounced than rainfall for collisions, severity of the crashes were less in nature. Andrey and Knapper (2003) found that the crash risk associated with rainfalls is mainly due to visibility with crash rates dropping quickly near to normal once the rain stopped. The study results also revealed that high winds and fog are responsible for a small proportion of crashes. More recently, Andrey (2010) investigated the effects of weather on crash severities using data from 1984 to 2002 for 10 Canadian cities. Using a match paired technique, Andrey (2010) showed that the risk of minor injury crash increased by 74% and 89%, respectively due to rain fall and snowfall whereas the increase in major/fatal injury crash risk was 46% and 52%, respectively due to rainfall and snowfall.

Using 25 years of weather, traffic and crash data for the 48 US continental states, Eisenberg (2004) developed a set of state-level daily and monthly collision models that followed a negative binomial distribution. The estimated monthly models showed a reduction in fatalities and an increase in non-fatal crashes with snow precipitation. The estimated daily models showed a positive relationship between snow precipitation and the total number of crashes, and also revealed that fatalities increased with heavy precipitation. Eisenberg and Warner (2005) conducted an analysis using the same dataset to investigate the relationship between snowfall and crash rate and calibrated negative binomial models with number of crashes as the dependent variable and precipitation, traffic exposure and other factors as independent variables. The findings revealed that the number of non-fatal injury crashes and property damage crashes increased during the snowfall but the number of fatal crashes decreased.

Sherif (2005) attempted to establish a link between road surface temperature, surface moisture, and road safety using data for one winter season from the city of Ottawa, Ontario, Canada. A Pavement Moisture Risk Factor (PMRF) was developed using the ratio of crash rate on wet surface to that of on dry surface. The results from the study indicated that wet surfaces are found more hazardous when temperature ranges from +1 to -2 C. However, some of the major limitations of the study was large aggregation of crash and weather data at a high level, masking the variations within different types of highways, consideration of wet and icy surfaces to be equal in terms of their effect on safety.

Hermans et al. (2006a) investigated the effect of weather factors on road safety by using data collected in Netherlands in 2002 and considering a number of factors related to wind, temperature, precipitation, and visibility. The collected data included hourly data on cloudiness, precipitation duration, precipitation amount, relative humidity, presence of precipitation, presence of fog, snow, thunderstorm, black ice, hail, and visibility. The researchers estimated negative binomial models and found that duration of precipitation and wind gust speed were associated with higher crash frequency while the presence of light were associated with lower number of crashes. Hermans et al. (2006b) also analyzed frequency and severity of crashes based on monthly data collected from 1974 to 1999 in Belgium using a state space approach considering several weather variables. A state space approach is based on describing a time-varying process by a vector of quantities. Percent days with thunderstorm and precipitation were found to be positively associated with minor injury risk with statistical significance. Both minor and major fatal injury risk were found to be higher in days with precipitation and with increased sun light hours. On the other hand, risk for both types of injuries was found to be lower in days with freezing temperature.

Qiu and Nixon (2008) conducted a meta-analysis of past studies from 1967 to 2008 to illustrate weather related factors affecting road safety. According to that review, it was found that snow precipitation was likely to increase the total number of crashes by 73%, 85%, and 100% on average in USA, Canada, and UK, respectively. Rain was likely to increase the total number of crashes by 58%, 73%, and 24%. Injury crashes also followed same pattern. It is to be noted that the estimates considered in this meta-analysis from different studies were the gross averages in different countries in different time span. Many factors such as driving behavior, exposures, and maintenance operations attributed to the variations in the percentages mentioned above and thus the findings from this study cannot be generalized without considering specific traffic, maintenance, weather characteristics of a specific region.

2.1.2 Effect of Traffic Related Factors on Safety

Knapp and Smithson (2000) investigated the impact of winter storm events on traffic volumes. Sixty four winter storm events occurring between 1995 and 1997 on interstates in Iowa were considered that met certain traffic volume, storm duration and snowfall intensity criteria set by the researchers. Road Weather Information System (RWIS) data from seven sites near the interstates were used to collect roadway and weather condition data. Automatic Traffic Recorders (ATR) located near the RWIS were used to collect the hourly traffic volumes to approximate storm and non-storm event traffic volumes. In that study, multiple regression analysis was conducted to investigate the relationship between reduction in the percentage of traffic volumes during winter storm events, snowfall intensity, total snow fall, and other weather related variables. The percent reduction in traffic volume during a winter storm event was derived by calculating the percent reduction in volume from average traffic volume during a non-storm event. The analyses indicated that the percent reduction in traffic volume during winter storm events had a statistically significant and positive relationship with total snowfall and the square of maximum gust wind speed. Knapp et el. (2000) also studied the impact of winter storms on crash frequency and reduction in traffic volume using a standard Poisson regression count model as there was no evidence of presence of overdispersion in the crash data. Hourly data were collected for crashes, traffic volume and weather variables in Iowa for a 48 km long segment of the Interstate highway from 1995 to 1998 and identified 54 winter storm events based on freezing temperature, precipitation and non-dry pavement surface. The model results showed an increase in crash frequency with the increase in exposure (vehicles millions kilometers), snow storm duration, and snowfall intensity.

Knapp and Smithson (2001) investigated the change in vehicle speed during winter weather events involving the use of mobile video data collection equipment to collect traffic flow data (i.e., speed and volume), weather and road surface conditions during seven 1998-1999 winter weather events at an Interstate location in Iowa. The researchers discussed the effectiveness and some concerns related to using mobile video data collection equipment during winter weather. Exploratory data analysis revealed a 16 percent reduction in the average winter weather vehicle speed compared to the typical average speed at the same location during non-winter conditions. A 307 percent increase in the variability of vehicle speed during winter weather events was also found when compared to the typical speed variability. The multiple regression model developed as part of the study revealed that off-peak average winter weather vehicle speed would increase with the square of traffic volume, decrease with the decrease in visibility below 0.4 km and decrease when snow began to affect or cover the roadway lanes. That study assumed that traffic volume was a surrogate for the weather characteristics affecting variable speed and as such weather data were not collected during the winter weather events.

Padget et al. (2001) conducted a study to investigate winter weather speed variability in Sport Utility Vehicle (SUV), pick-up trucks and passenger cars. They collected and analyzed the speeds of SUVs, pickup trucks, and passenger cars on five different winter weather pavement surface conditions in Ames, Iowa. The analysis results revealed that all three types of vehicles had similar average speeds during the normal conditions with passenger cars having the highest average speed but this pattern reversed during the winter weather conditions with SUVs having the highest average speed and passenger cars having the lowest average speed. It was concluded that passenger cars generally traveled slower than SUVs during winter weather conditions but faster during normal conditions. It was found that there was a difference between the normal and the winter weather speed choice of SUV, passenger car and pickup truck drivers. However, the variability in the speed of SUV, pickup trucks and passenger cars increased during winter weather periods compared to those during normal conditions regardless of the time of day. Nighttime speeds for all three types of vehicles were found to be significantly less than daylight speeds. The analysis result also revealed that the average vehicle speed for all three types of vehicles decreased with poorer roadway surface conditions during the winter weather periods.

Lee and Ran (2004) developed a winter maintenance performance measure based on speed recovery durations during snow events using speed data collected from Automatic Traffic Recorders (ATR) and winter storm report data in Wisconsin. The authors defined speed recovery duration as the time between the stopping of the snow event and recovery of vehicle speed to normal. The speed recovery duration was proposed as a measure of winter maintenance performance measure in lieu of the total operational costs or salt usage. A regression model developed in the study showed that vehicle speed recovery duration to the normal condition was significantly associated with snow duration and maximum speed reduction during the snow storm. Lee et al. (2008) conducted a follow-up study involving a larger sample size to validate the findings of the Lee and Bran (2004) study. The follow-up study involved the investigation of vehicle speed changes during winter weather events with data extracted from Wisconsin winter maintenance logs. The study conducted a regression tree analysis with Speed Recovery Duration (SRD), which is the time required to regain the normal average speed from minimum speed during a winter storm event, as the dependent variable. SRD was found to be a promising factor to evaluate winter maintenance activities using vehicle speed data. According to the developed models, it was found that SRD would increase with the quick reduction of vehicle speed to the minimum speed during the winter storm events. A longer SRD would also be expected for increase in the percent of Maximum Speed Reduction (MSR). The study confirmed that vehicle speeds could be a good measure for indicating driving conditions during a winter weather event.

2.1.3 Effect of Winter Maintenance on Safety

Adams et al. (2006) developed regression tree models for estimating labor, equipment and material resources, cleanup cost and percent overtime cost associated with winter weather maintenance activities during storm events in Wisconsin. The researchers focused on estimating the required resources using regression tree models, which are independent from unit costs of labor, maintenance and equipment that change over time and vary from county to county. Models were developed for 72 counties in Wisconsin divided in 4 service groups depending on the percent highway coverage received by those counties during winter weather events. The regression models captured the effect of precipitation depth, storm duration, air and pavement temperature at the start of the storm, time of the day, and service level on resource requirements for winter maintenance. The analysis showed that temperature influenced labor and equipment requirements as well as materials usage for winter maintenance. This type of model is used by the state of Wisconsin for estimating resource requirements in case of an impending storm. These models are also applicable to different counties for estimating the resource requirements with varying unit labor, material and equipment costs.

In another study, Ye et al. (2009) investigated and evaluated the effect of weather information on winter weather maintenance cost. For this purpose, a general winter maintenance cost model was presented and neural networks and sensitivity analysis were used to identify key variables that had significant effect on cost. The analysis revealed that enhanced accuracy and frequent use of weather information could reduce winter maintenance cost significantly. The cost-benefit analysis conducted as part of the study revealed that weather information can be a promising way to improve winter maintenance and reduce agency costs.

Russ et al. (2008) conducted a study focused on addressing pretreatment protocol for winter maintenance of roadways in Ohio. The study was conducted in four parts consisting of conducting surveys of personnel in state departments of transportation and county manager in Ohio, conducting field durability studies of various applications of brine on Portland Cement Concrete and asphalt concrete pavements in Ohio, inspections of pretreatment during three winter seasons, and performing laboratory tests on Portland Cement Concrete (PCC) and Asphalt Concrete (AC) cores. Integration of the findings from these tasks resulted in a decision tree to aid in operational planning and pretreatment.

Blomqvist et al. (2011) combined an empirical model developed in Sweden with data on residual salt, road surface wetness, and traffic from 18 Danish field case studies to predict salt usage on road surface during winter weather. Results showed that the decay of residual salt could be modeled with traffic as an independent variable with a fair to quite good fit (with R^2 value ranging from 0.64 to 0.99). Road surface wetness was positively related to the rate of residual salt loss from the wheel tracks meaning wetter surface would expedite the salt leaving process from the wheel tracks. While passing only a couple of hundred vehicles on a wet road surface would result in almost no salt in the wheel track, it would take a couple thousand vehicles to pass on a moist road to achieve the same result.

2.2 Review of Past Methodologies for Modeling Winter Weather Crash Frequency

2.2.1 Winter Weather Crash Frequency Models

Usman et al. (2010) conducted a study to quantify the safety benefits of winter weather maintenance and operations employing event-based crash frequency models. Using crash and weather data from different sources in the province of Ontario, they developed event base models for predicting winter crashes controlling for visibility, Road Surface Condition index (RSI), traffic exposure, site specificity, and precipitation under snow storm events. The novelty of this research lied in introducing a Road Surface condition Index (RSI) which was assumed to reflect the maintenance operations during snow storm events. RSI was defined for major classes of road surface conditions having ordered categories in terms of the severity. RSI was introduced as a surrogate measure of the commonly used friction level and RSI was assumed to be similar to road surface friction values and varied from 0.1 (poorest, e. g., ice covered) to 1.0 (best, e. g., bare and dry). RSI was defined as a range of surface friction values assigned to different major classes of road surface conditions having based on the literature. Three types of modeling techniques were used to investigate the association of crash frequency during a snowstorm event with road surface conditions and the other controlling factors mentioned above. Results showed that the generalized negative binomial model offered the best fit for the data over the negative binomial and the zero inflated negative binomial models. The Road surface condition (RSI) index was found statistically related to influence a crash occurrence during a snow event.

Using disaggregate hourly data from the same winter snow storm events as in Usman et al. (2010) , Usman et al. (2011) developed a generalized negative binomial model for predicting winter crash frequency. This model was compared to the model calibrated from using aggregate event-based data to examine the impact of data aggregation (from event based data to hourly data) on modeling results. Results showed that data aggregation ignoring data correlation could result in loss of information and models with biased parameters. Some important factors turned out to be significant in the disaggregate model while it was insignificant in the event-based aggregate model. The same study also developed two Poisson-Lognormal (PLN) models using hourly winter crash data set with multi-level (event-hour structure) and single level data structure. The Multilevel data structure accounted for the within-event correlation of the observations at different hours. The single level and multilevel PLN models based on hourly data were very similar indicating that event-level correlation in the specific dataset used in this study was weak.

After establishing the effectiveness of calibrating a model with disaggregate dataset over aggregated data for predicting winter crash frequency, Usman et al. (2012) developed winter crash frequency models using disaggregated hourly dataset in a bid to investigate the link between winter road collision occurrence, weather, road surface conditions, traffic exposure, temporal trends, and site specific effects. Results showed that both the Generalized Negative Binomial (GNB) model and the Poisson-lognormal (PLN) model had a better fit when considering site-specific effects than without considering these effects. The PLN model considered the multi-level (event-hour level) structure of the data, while the GNB model was developed using the hourly data for winter crashes and the other factors mentioned above. The GNB model also has the ability to account for data heterogeneity through varying the overdispersion parameter. It was found that GNB provided a better goodness of fit compared to the PNL model suggesting that a single level model would be adequate without considering the multilevel structure of the data (event-hour hierarchy in this case). A multilevel model would not affect the significance of the variables considered in the single level model in this study.

Using the same dataset, Usman et al. (2012) developed crash-injury severity models to take into consideration the multilevel or hierarchical nature of crash data. They developed three types of models using the occupant-based data, vehicle based data, and collision or crash based data. The purpose was to consider the possible intra-class correlation at occupant or vehicle level observations. Multilevel multinomial logit, multilevel binary logit and multilevel ordered modeling structures were adopted to develop models using the winter crash data having occupant-vehicle-crash level hierarchy. The study compared these three alternative logistic models in a multilevel modelling framework. It was found that multilevel multinomial logit model had better fit to the occupant level and vehicle level data, while binary logit and ordered logit performed better for collision level data. Overall, multilevel multinomial logit models offered better predictions. It was found that aggregation of crash data at the collision level affected the parameters estimates significantly.

Qin et al. (2006) developed a negative binomial model for predicting crashes during winter storm events from 2000 to 2002 relating to winter storm severity in regard to duration and intensity, wind speed, deicing units used per lane mile, salt used per lane mile. The analysis was conducted for the Wisconsin State Trunk highway system. Results revealed that early deployment of winter maintenance operations could significantly reduce crash occurrence with the model showing a negative relationship between crash frequency and the time the crew spent out before the beginning of a storm. An inverse relationship between crash occurrence and the amount of deicing material used indicated a reduction in the number of crashes associated with the deployment of more deicing material. However, a positive relationship between salt units used and crash occurrence was also found; this was explained by the fact that there is a time lag between salting and snowplowing that can result in a slurry period during which the bare pavement might be slippery and more crashes could occur. Strom duration and wind speed were found to be positively associated with the crash frequency. Temporal distribution of the crashes during a snowstorm revealed that a large percentage of the crashes occurred during the initial stages of snowstorms. Though the temporal patterns for the percentage of crashes during snowstorms were similar for both state and local roads, a higher percentage of crashes occurred on local roads during the later stage of the snowstorm reflecting the different level of maintenance activities and usage of deicing materials.

2.2.2 Development of a Winter Severity Index

Nixon and Qiu (2005) developed a storm severity index using 252 winter storm events in Iowa. The storm severity index was proposed to provide a measure of the severity of any given storm based solely on a meteorological description of that storm. Storms were classified by six factors (storm type, in-storm road surface temperature, in-storm wind condition, early storm behavior, post storm temperature and post storm wind condition). A multiple regression model was estimated to produce a storm severity index between 0 and 1 with 0 indicating a very mild storm and 1 indicating a very severe storm. Winter maintenance personnel (maintenance garage supervisors) from Iowa DOT were asked to rank the severity of 10 representative storms (out of the 252 storm events considered for developing the multiple regression model) according to the level of difficulty that these events would pose to their maintenance activities. It was found that although there was general agreement between the supervisors' ranking and the initial severity index, there were areas of disagreement. The scores for the different factors considered in the regression model were adjusted according to the supervisors' ranking. This type of severity index for winter storm events can be helpful in assessing the performance of maintenance agencies as the severity of the storms they face can be quantified.

2.2.3 Comparison of Crash Injuries during Winter and Non-Winter Events

Khattak and Knapp (2001) conducted a study to compare the winter snow event crash injury and non-injury crash rates with comparable winter non-snow event crash rates on selected Interstate highway locations in Iowa. Winter snow events were defined based on the definition in Knapp et al. (2000) and the same dataset and location was used for this study. They also compared the crash injury occurrence during winter snow event periods and comparable winter non-snow events along with the assessment of the impact of snow event elements on snow event crashes using binary logit models. Comparable non-snow periods were identified and extracted for the same hours on the same weekdays within the same month of the winter snow events. Results revealed significant increase in injury and non-injury crash rates during winter snow events than those rates during comparable winter non-snow events. However, the modeling results indicated that crashes during snow events involved fewer injuries than crashes during comparable non-snow periods. It was also revealed that snow event elements such as higher wind gust speed tended to result in more injury crashes during snow events, while higher snowfall intensity resulted in crashes involving fewer injuries during snow events.

2.3 Multilevel Modeling Techniques for Crash Severity Analysis

Multilevel techniques for modeling injury severity of individual occupants take into account crash hierarchy (crash-vehicle-occupant). The hierarchy can be expanded to include geographical elements such as road segments or sites, regions, and so on.

Jones and Jorgensen (2003) and Lenguerrand et al. (2006) were among the first studies to recognize the need to account for the crash level hierarchy (crash-vehicle-occupant) in crash data when modeling crash severity. Jones and Jorgensen (2003) first attempted to consider the natural crash hierarchy in disaggregate crash data using a crash dataset in Norway. A total of 16,332 crashes along Norwegian roads spanning from 1985 to 1996 were obtained. The probability of injury (fatal or serious) severity of occupants was modeled with a binomial model. The hierarchy specified for this study was occupants nested within crashes and crashes nested within municipalities. The predictor variables included variables related to crash characteristics (light condition, road type etc.) and occupant characteristics (gender, age etc.). The random variations at the crash level and municipal level were significant. The Intra-class Correlation Coefficient (ICC) revealed that the largest proportion of variation in the injury severity outcomes was attributed to the lowest level of the crash hierarchy (occupants with 83%), while 16% of the variation was attributed to the crash level and 1% was attributed to the municipal level.

However, this study did not provide any comparison between single-level and multilevel models in terms of parameter significance and model fit as a single-level model was not computed. Jones and Jorgensen (2003) ignored the intermediate “vehicle” level as majority of the vehicles in the crash data considered for the analysis included only a single occupant sustaining serious or fatal injury leaving little information to differentiate between vehicles and occupants. This can be attributed to a fundamental characteristic of crash data but not to the particularity of the data analyzed by Jones and Jorgensen. While the number of crashes can be large, the number of cars per crash and of individuals or occupants is typically very low.

Valnar (2005) illustrated the advantage of multilevel modeling compared to statistical techniques that ignore hierarchies, based on two empirical traffic safety examples. The study showed two important consequences of ignoring the hierarchical structure in the data. The first consequence is the underestimation of standard errors, which was illustrated with data from an observational study on seatbelt behavior. It was found that two factors (passenger: a dummy variable indicating whether the observed subject was a front seat passenger or a driver, and weekend night: a dummy variable to indicate the time span of a crash) were significant at five percent level in a single-level model while those were insignificant in a two-level model. The second consequence, related to contextual information, was illustrated with data from a roadside survey on drink driving. The second consequence relating to contextual information is illustrated with the frog pond theory in Hox, 2002. For example, in traffic safety, this theory is applied in the form of the effect an explanatory variable (for example, willingness to take risk) might have on the dependent variable (for example, choice of speed by drivers). Speed choice may depend on the average speed of other drivers at a particular location. Of particular interest for investigating the second consequence was the relationship between gender, traffic counts (total number of vehicles driving by the road site during police check) and the odds of drivers exceeding the legal limit of Blood Alcohol Concentration (BAC). The drivers were nested within road sites and a multilevel model was fitted. A significant relationship between gender and the outcome variable was a nice illustration of the frog pond theory in this case. Although an insignificant cross-level interaction was found, a significant cross-level interaction would mean a varying influence of individuals' gender on odds for drunk driving with different values of traffic counts at different sites.

Lenguerrand et al. (2006) proposed a binomial model to model the probability of vehicle occupants accounting for the hierarchical structure of the crash data with three levels: crash, car/vehicle, and occupant. They used crash data from French road injury crash census for a four year period from 1996 to 2000 and tested three different modeling techniques with logistic models, generalized estimating equations (GEE), and multilevel logistic models. It was revealed that multilevel models yielded better results compared to the other two modeling techniques reinforcing the importance of accounting for the hierarchical nature of the crash data. One important observation from this study was that the variance of vehicle random effect was falsely estimated to be zero for 36% of the cases. These incorrect estimates were attributed to the small number of observations per vehicle and per crash.

Kim et al. (2007) used a sample of 548 crashes collected from 91 two-lane intersections to model the probability of occurrence of five types of crashes using binomial multilevel models with crashes clustered into intersections. For each crash type, a separate model was developed. It was found that the random variation of the intercept across intersections was significant except for one crash type (head-on). It meant that the average probability for these types of crashes to occur varied significantly from intersection to intersection. The modeling results showed that multilevel models provide similar results compared to the traditional models except for one crash type (sideswipe opposite direction).

Helai et al. (2008) used a binomial multilevel model to predict the severity level of driver injury and vehicle damage in traffic crashes based on a total of 4,095 crashes occurring at signalized intersections that involved 7,084 driver-vehicle units. A driver-vehicle unit was defined by both the vehicle and the driving person involved. A binary dependent variable was defined by combining the driver injury severity and vehicle damage severity for the vehicle-driver units involved in crashes. The authors compared the results of a traditional binomial regression model with the multilevel binomial model and found that the ICC (Intra-class Correlation Coefficient) was 28.9%. It meant that 28.9% of the variation in the probability for driver-vehicles units to have experienced severe damage resulted from between crash variance or within crash correlation. The comparison with the classical model also revealed a better model fit for the multilevel model formulation.

Yannis et al. (2010) modeled the probability for each individual occupant in a vehicle to sustain different levels of injuries using injury severity levels as a multinomial response. They used a dataset containing 1,300 crashes that involved 3,500 occupants. The results revealed no random variation at the vehicle or crash level for the probability of injuries sustained by the occupants.

Dupont et al. (2010) modeled the probability of a fatality for each individual occupant using a set of fatal car-car crashes. They considered the country, crash, and vehicle hierarchy for the multilevel modeling. The results indicated no significant random variation at the higher level. Comparison with a single-level model formulation revealed similar values and significance for the coefficients of the variables considered for that modeling purpose.

2.4 Summary

This chapter provided a comprehensive review of the literature on the impact of weather, traffic, and maintenance related attributes on winter weather safety. Past studies related to modeling winter weather crash frequencies using both aggregate and disaggregate level crash data were described in this chapter. Though most of those studies investigated the effect of weather, traffic and maintenance parameters on road safety, the development of a site prioritization technique for improving winter weather safety using available crash data and maintenance crew reported weather data was scarce in the literature.

This chapter also reviewed studies that employed a multilevel modeling technique for analyzing crash data at a disaggregate level. It was evident that when the size of the crash sample was small, previous researchers (Lenguerrand et al., 2006; Jones and Jorgensen, 2003) ignored the vehicle level when analyzing disaggregate crash data and directly modeled the nesting of individuals within crashes. Alternatively, Helai et al. (2008) considered vehicle-driver entity as the unit of observation and ignored the nesting of occupants within vehicles.

The next chapter presents the data sources, description of the data used for the analysis, and data processing steps.

CHAPTER 3

DATA

This chapter describes the various data sources, data processing steps, and description of the data used for the development of the SPFs for winter weather crashes and crash injury severity models. The data collection spanned from the 2008-2009 winter season to the 2011-2012 winter season. The Iowa DOT defines the winter season as the period from October 15th to April 15th. Crash data, roadway information for different types of roadways, and weather related information were collected for these four winter seasons using a variety of data sources.

3. 1 Data Sources

Three types of data were used in this study: crash data, roadway information and traffic data, and weather related data. These data were gathered from different sources and integrated to compile a suitable dataset for achieving the research objectives.

3.1.1 Crash Data

The Iowa DOT maintains a database of all crashes that have occurred along Iowa's roadways. Information on weather related crashes occurring on Iowa's roadways for the winter seasons from 2008-2009 to 2011-2012 were collected. For developing safety performance functions, it was important to obtain crashes occurring on specific segments of roadway within a certain period. Weather related crashes occurring on one-mile road segments for three types of roadways during the four winter seasons were collected. The crash database contains detailed information related to individuals as well as vehicles involved in the crashes. Each crash was already assigned to a unique crash key. Data on the crash occurrence time and space were also collected for data aggregation over time and space. Crash data were spatially joined with the roadway and traffic volume data based on the spatial location of the crashes occurred along different types of roadways. Information on date and time of crash occurrences was necessary to integrate the crash data and weather data to represent the weather conditions during the crash occurrences.

3.1.2 Weather Data

A novel approach was taken in this research to obtain the weather related information for the weather related crashes considered during the four winter seasons in this research. Weather related information for the crashes on one-mile road segments were collected from the nearest cost center maintained by the Iowa DOT. Cost centers are maintenance garages maintained by Iowa DOT containing maintenance materials, equipment and dispatch maintenance crews for winter weather maintenance performance. A cost center is represented by an accounting code for a specific place of work. So in this context, each cost center represents a field maintenance garage. Each maintenance garage belongs to a district (there are six DOT districts in Iowa). Each cost center has its own jurisdiction of roadways on which the maintenance crews associated with that particular cost center perform the maintenance activities during winter weather. There are several cost centers associated with each district. The size of a cost center varies and the area of the cost center is determined by the resources at the cost center and the operational needs of the roads in the area. The boundaries of the cost centers do not have any association with political boundaries of county lines or city limits. Figure 3.1 shows the cost centers (orange triangles) and the roads they are responsible for (colored lines) Maintenance crew report the weather related information such as air temperature, pavement temperature, wind speed, visibility, precipitation type, and precipitation amount while performing the maintenance activities according to information provided by Iowa DOT. The cost center information was used instead of the Road Weather Information System (RWIS) because of the authenticity and the consistency of the crew reported weather information. A Road Weather Information System (RWIS) consists of sensor stations in the field, a communication system for data transfer, and central system to collect data from the sensors. RWIS data in Iowa contain information about air temperature, pavement temperature, and wind speed recorded by the RWIS stations near roadways. RWIS stations record data every 10 minutes. An example of sample data from RWIS is given in Appendix A. The Iowa DOT officials confirmed that the crew reported weather information is representative to that of the RWIS information. As shown in the figure 3.2, the colored lines represent the segments covered by each cost center and the green circles represent the RWIS sites.

3.1.3 Roadway and Traffic Volume Data

Roadway geometry data and traffic volume data for each road segment to be considered for developing safety performance functions for the three types of roadways were collected from the Iowa DOT Office of Transportation Data, Division of Planning and Programming maintained through Geographic Information Management System (GIMS). Several roadway geometry related attributes such as surface width, lane width, number of lanes, shoulder width, shoulder type and others are reported in the GIMS database along with the Annual Average Daily Traffic (AADT) for specific segment of roadways. Roadway geometry and traffic volume data were collected for one-mile road segments for the four winter seasons. Each one-mile road segment was assigned to a unique identification named as ROWID in the GIMS data. It is to be noted that real time traffic volume collected from the Automatic Traffic recorder (ATR) could have been used to incorporate in the SPFs. It is natural that traffic volumes differ from seasons to seasons and considering the real time traffic volume would represent the exact traffic conditions during crash occurrences for the winter seasons. The problem associated with collecting ATR reported traffic volume was related to the format in which volume data are stored by Iowa DOT. Iowa DOT used to utilize a different format for collecting and storing ATR reported traffic data prior to 2011. Iowa DOT introduced a more efficient format to collect and store the volume data starting from 2011. Thus, it was difficult to integrate the ATR reported volume data from the two different format. A seasonal factor could have been applied to the AADT to represent the traffic volume during the winter seasons. This is discussed as a limitations at the end of the dissertation. The same problem was also faced for developing the severity models.

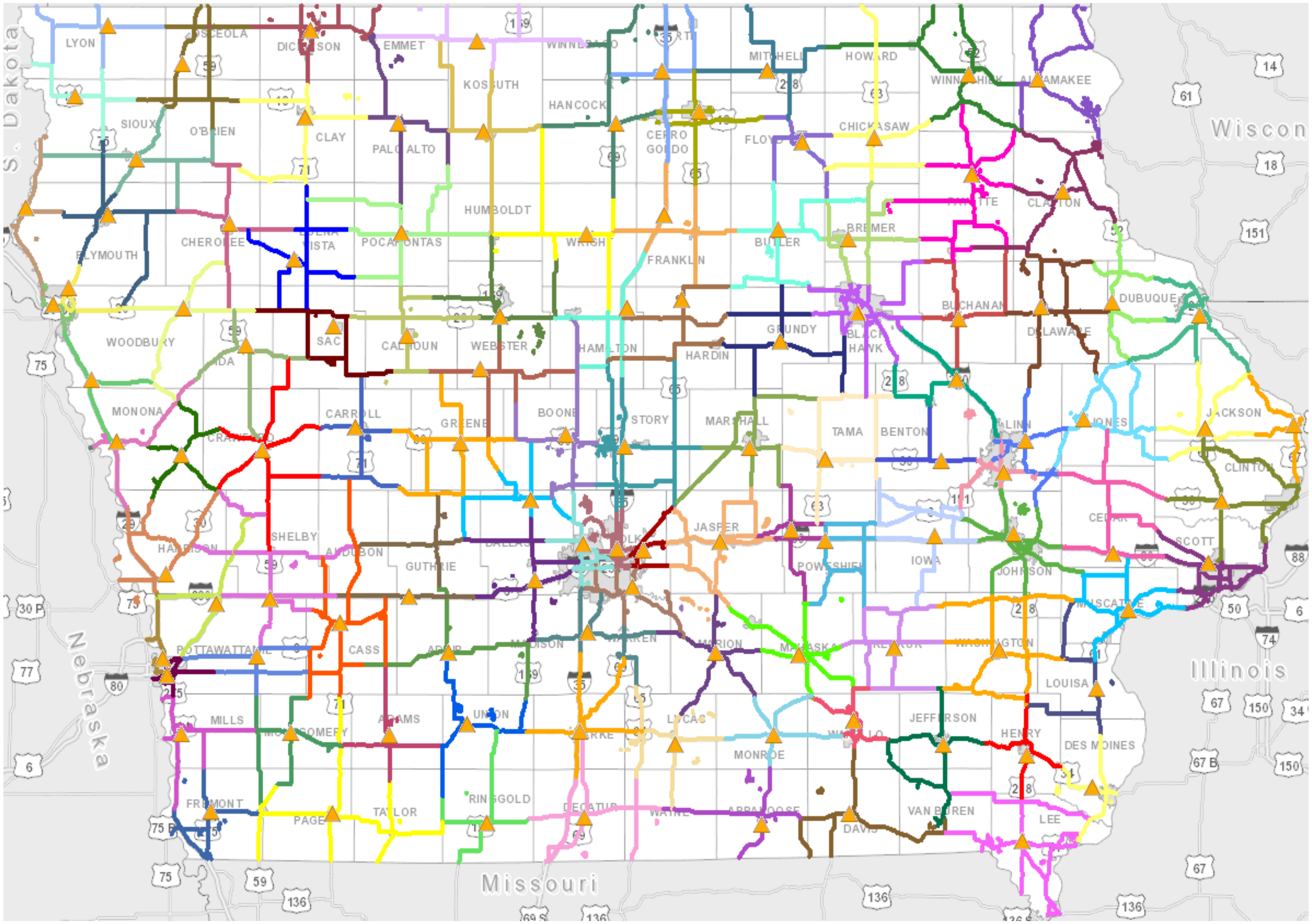


Figure 3.1 Location of cost centers and associated roads

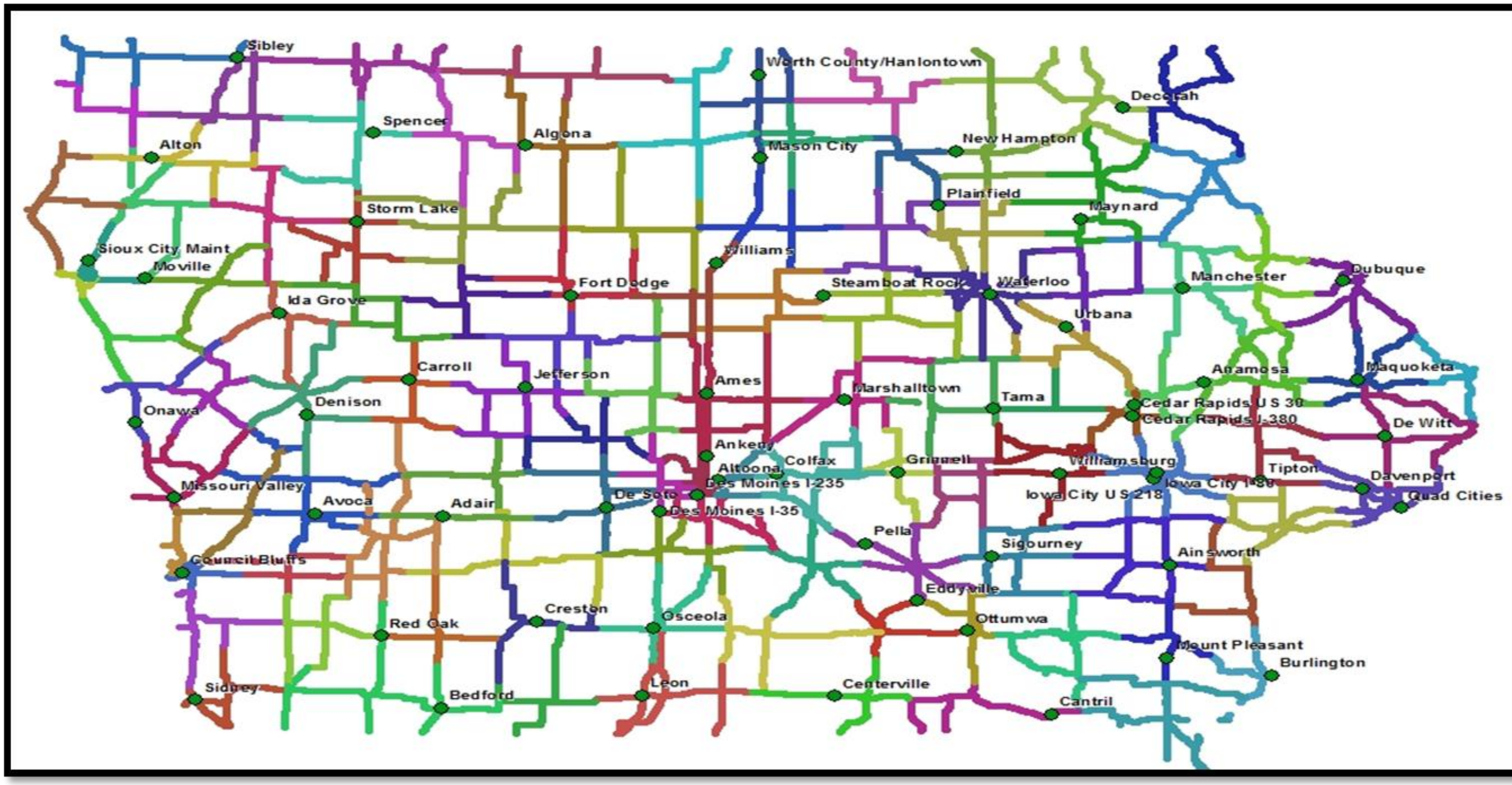


Figure 3.2 Locations of RWIS sites and roadway jurisdictions associated with cost centers

3.2 Data Processing for Developing Safety Performance Functions

SPFs have been typically developed to estimate crash frequency using site or roadway characteristics such as lane width and traffic volume expressed as annual average daily traffic (AADT). Incorporating weather related attributes that correspond to the specific crash situations in SPFs is more complex and labor intensive. Integrating weather data with the crash data is herein critical as the proposed research aims to develop safety performance functions for three types of roadways in Iowa to predict winter weather crashes as a function of variables related to winter weather conditions such as visibility, pavement temperature, air temperature, and wind speed.

Figure 3.5 shows the steps involved for data processing and integration of the different data sources collected. The challenge for processing the data was to integrate the crew reported weather information with the crashes occurring on the one-mile road segments for different roadway types. The crashes on different types of roadways were assigned to appropriate one-mile road segments matching their spatial locations in ArcGIS. Each crash was spatially joined with the nearest cost center through ArcGIS. Crew reported weather information for each crash was integrated with each crash based on the date and time of the crashes. Multiple crew reports were obtained for quite a few crashes as there were multiple crews reporting weather information on the same day those crashes occurred. One crew report was kept for each crash based on two criteria. Initially, precipitation intensity was considered to screen out crew reports. The crew reported precipitation intensity having higher priority was kept from multiple reports having other types of precipitation intensities. Table 3.1 shows the priorities for precipitation intensities according to Iowa DOT. If precipitation intensity was the same for a crash, precipitation duration was considered. The crew reported precipitation duration having the highest value was kept in this case.

Once all these steps were completed, 13,859 winter weather crashes that occurred during the four winter seasons were found to be associated with crew reported weather information out of the 15,096 winter weather crashes on the three types of roadways. It was also important to consider the difference in time of crash occurrence and time the crew members departed roadways. The weather condition during crashes would not be representative if the time difference between crew departure and time of crash occurrence is large. It was found that more than 80% of the 13, 859 crashes occurred before the crew departed. For the rest of the crashes, more than 50% occurred within 4 hours of crew departure and 19% within 4-6 hours. Not a large number of crashes occurred after crew departure with a time interval of 6 hours or more. Those crashes were still included in the analysis by assuming that the weather conditions were the same during the crashes as reported by crew members. Figure 3.3 and 3.4 shows distribution of crash percentage based on crew departure and time difference for crashes occurred after crew departure. These crashes were assigned to the one mile road segments according to the roadway types of interstate/freeway, multilane, and two-lane roadway. The last step involved the integration of this dataset with the GIMS to assign the geometric characteristics on each one-mile road segment.

Table 3.1 Priority ranking of precipitation intensity

Precipitation Intensity	Rank According to Priority
Freezing rain	1
Sleet	2
Snow	3
Blowing snow	4
Refreeze	5
Road frost	6
Rain	7
Fog	8
Bridge frost	9

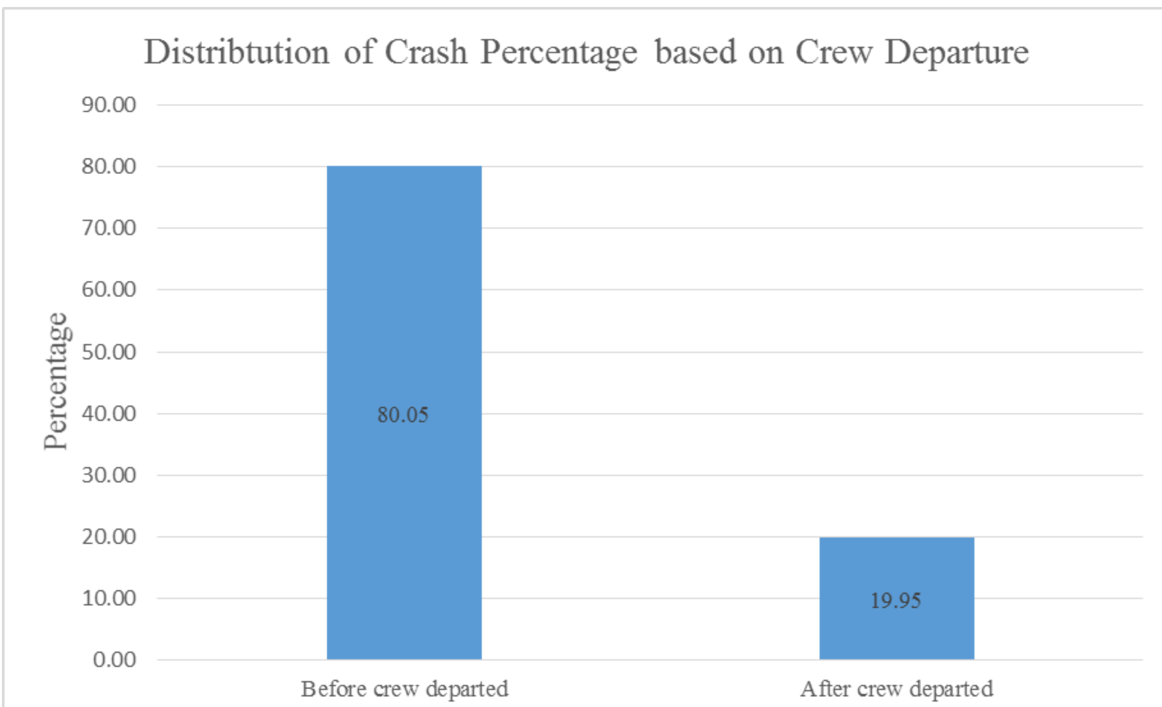


Figure 3.3 Distribution of crash percentage based on crew departure

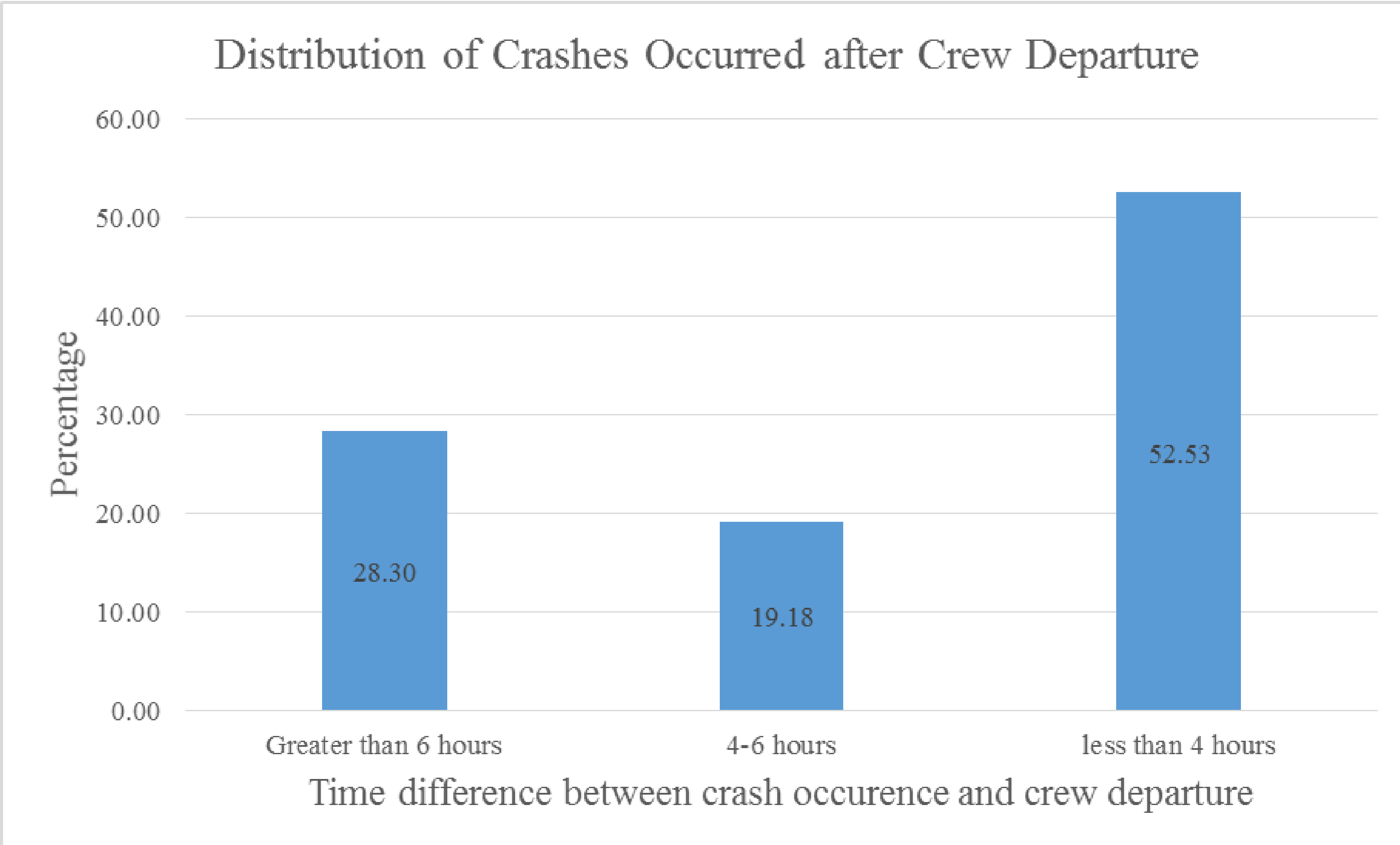


Figure 3.4 Distribution of crashes occurring after crew departure

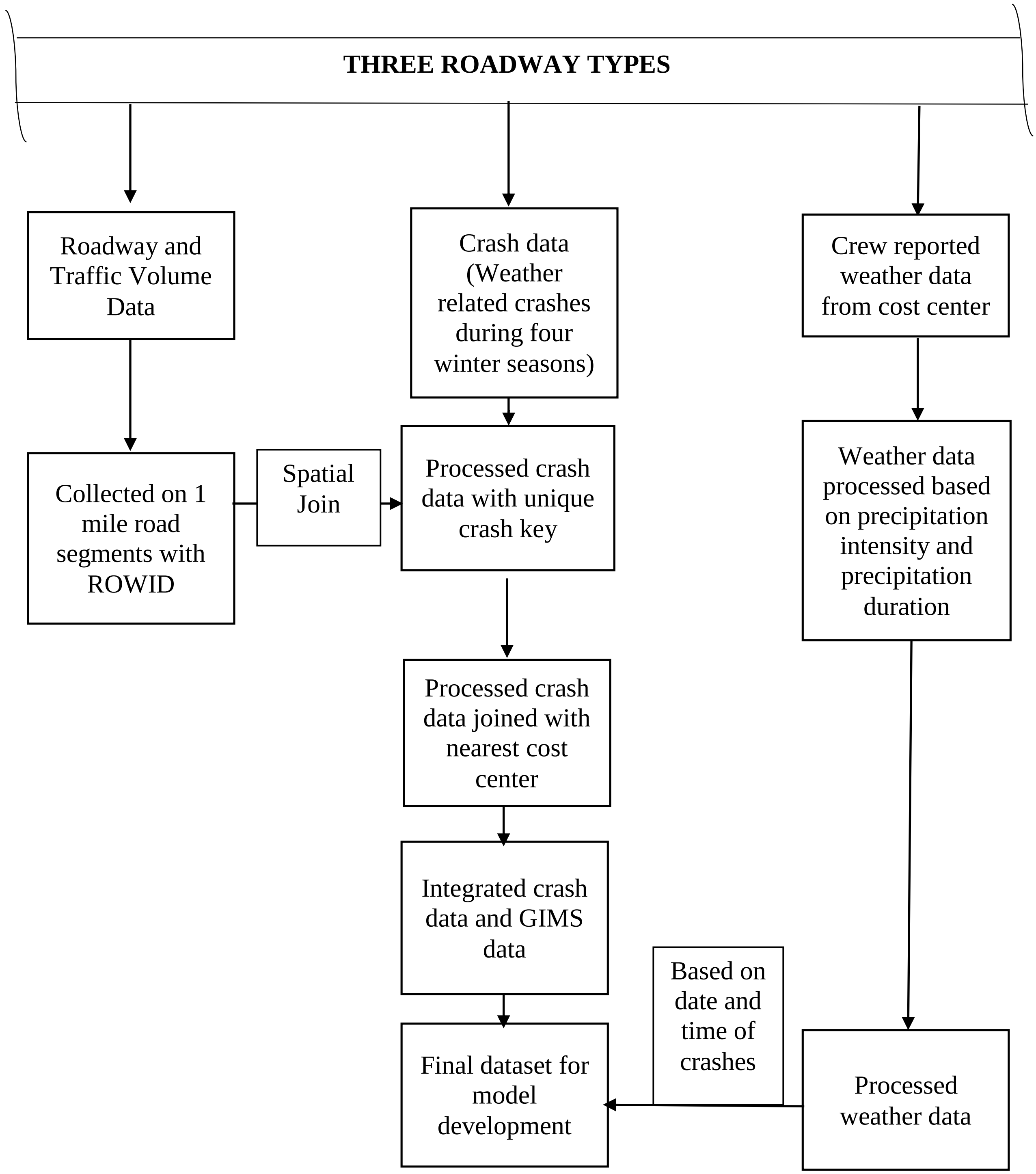


Figure 3.5 Crash data, weather data, and roadway data integration steps

3.3 Description of the Data Used to Develop Safety Performance Functions

A total of 13,859 weather related crashes during the winter seasons from 2008-2009 to 2011-2012 were considered for developing SPFs for different roadway types. Of the total 13,859 weather-related crashes during the winter seasons, 6,210 crashes occurred on interstate/freeway facilities, while 3,898 crashes occurred on multilane/ divided/undivided roadways and 3,751 crashes occurred on two-lane roadways. Descriptive statistics for the weather related factors considered in the development process of SPFs for each class of roadway are presented in the following tables.

Table 3.2 Descriptive statistics of weather attributes for crashes on Interstate/freeway

Total number of crashes	**N = 6,210**			
Variable	**Interstate/Freeway**			
	Min	Max	Mean	SD
Air temperature (F)	-25	51	20.46	10.4
Pavement temperature (F)	-21	54	22.46	8.9
Wind Velocity (mph)	0	117	14.96	8.7
Visibility (in mile)	1	5	3.65	1.31
Snow Amount (in inch)	0	12.5	2.01	2.13

Table 3.3 Descriptive statistics of weather attributes for crashes on multilane divided/undivided roadways

Total number of crashes	**N = 3,898**			
Variable	**Multilane divided/undivided**			
	Min	Max	Mean	SD
Air temperature (F)	-29	42	19.6	10.9
Pavement temperature (F)	-31	54	21.32	10.05
Wind Velocity (mph)	0	117	14.13	8.6
Visibility (in mile)	1	5	3.77	1.31
Snow Amount (in inch)	0	16	1.94	2.2

Table 3.4 Descriptive statistics of weather attributes for crashes on two-lane roadways

Total number of crashes	**N = 3,751**			
Variable	**Two-lane roadway**			
	Min	Max	Mean	SD
Air temperature (F)	-29	50	22.01	9.86
Pavement temperature (F)	-26	56	23	9.24
Wind Velocity (mph)	0	57	15.24	9.07
Visibility (in mile)	1	5	3.79	1.3
Snow Amount (in inch)	0	16	1.83	2.1

Once all the crashes and factors to be considered for modeling was ready, average number of crashes was computed for each one mile road segment for the three classes of roadway for the four winter seasons. Average values for each segment for the weather related variables were considered for developing the safety performance functions. The total numbers of one-mile road segments generated along which at least one weather-related crash occurred during the four winter seasons in Iowa for interstate/freeway, multilane divided/undivided, and two-lane roadways were 995, 887, and 2,325 respectively. Tables 3.4-3.6 show the descriptive statistics for the average values of the factors considered for final analysis. It is to be noted that there were some outliers in the data representing roadway geometry characteristics and traffic volumes. For example, the minimum posted speed limit for Interstate/freeway facilities was found to 35 miles per hour. This might be a segment located along ramps.

Table 3.5 Descriptive statistics of factors for SPF development (Interstate/freeway)

Number of roadway segment	**N = 995**			
Variable	**Interstate/Freeway**			
	Min	Max	Mean	SD
Segmental Crash frequency	1	54	6.24	6.6
Segmental Air temperature (F)	-25	37	21.67	6.54
Segmental Pavement temperature (F)	-9	38.8	22.28	7.21
Segmental Wind Velocity (mph)	0	37.57	14.03	5.9
Segmental Visibility (in mile)	0	5	3.35	1.36
Segmental Snow Amount (in inch)	0	10	1.97	1.36
Segmental AADT	90	113600	23958	17056
Segmental Surface Width	16	90	30	10
Segmental Posted Speed Limit	35	70	67	4.68

Table 3.6 Descriptive statistics of factors for SPF development (Multilane divided/undivided)

Number of roadway segment	**N = 887**			
Variable	**Multilane divided/undivided**			
	Min	Max	Mean	SD
Segmental Crash frequency	1	27	4.4	4.36
Segmental Air temperature (F)	-18	38	20.5	7.43
Segmental Pavement temperature (F)	-8	54	21.28	7.82
Segmental Wind Velocity (mph)	0	50	13.64	6.67
Segmental Visibility (in mile)	0	5	3.22	1.63
Segmental Snow Amount (in inch)	0	12.5	1.9	1.52
Segmental AADT	50	34225	11023	5891.44
Segmental Surface Width	12	72	32	11
Segmental Posted Speed Limit	20	65	52	13

Table 3.7 Descriptive statistics of factors for SPF development (Two-lane roadway)

Number of roadway segment	**N = 2,325**			
Variable	**Two-lane roadway**			
	Min	Max	Mean	SD
Segmental Crash frequency	1	12	1.61	1.26
Segmental Air temperature (F)	-15	50	21.85	9.43
Segmental Pavement temperature (F)	-12	56	22.23	9.54
Segmental Wind Velocity (mph)	0	50	14.18	8.81
Segmental Visibility (in mile)	0	5	3.1	1.81
Segmental Snow Amount (in inch)	0	14.5	1.84	1.92
Segmental AADT	50	52700	3345	2652
Segmental Surface Width	14	76	26.3	6.56
Segmental Posted Speed Limit	20	70	52.71	6.75

3.4 Additional Data Collection Effort for Crash Injury Severity Analysis

Data used in the analysis of injury severities of occupants involved in crashes during winter weather were obtained from the Iowa DOT maintained crash database. For the purpose of this research, a data set containing all crashes occurring over four winter seasons (2008 – 2012 with each winter season covering six months from October 15th to April 15th). Crash level, vehicle level, and occupant level attributes were obtained from the crash database for the four winter seasons. It was important to select a study corridor for analyzing injury severities of the crashes occurring during these winter seasons as it was not feasible to retrieve weather information for all crashes (both weather and non-weather related) occurred statewide along every type of roadways in Iowa. Considering different types of roadways, interstate roadways were ranked as the most important with respect to mobility and safety. As such, an interstate corridor was selected as the study corridor as described in the following section.

3.4.1 Study Corridor

The study corridor considered for the crash injury severity analysis in this study included the entire length of Interstate 80 (I-80) crossing the state of Iowa from the Missouri River in the Mississippi River. The total length of the corridor is 318 miles. 5,242 crashes occurred on this corridor during the four winter seasons from 2008 to 2012. Key factors considered while selecting this corridor were the availability of the Road Weather Information System (RWIS) stations located in near proximity to the study corridors, frequency and severity of crashes occurring on the roadway. Figure 3.6 shows the study corridor with green dots.

3.4.2 Data Processing for Crash Injury Severity Models

As mentioned in the previous section, all the crash level, vehicle level, and occupant level information were extracted from the Iowa DOT maintained crash database for the I-80 corridor for the specified period. The crash data set contained both weather and non-weather related crashes with 7,829 injured occupants. This data set is referred to as All Crash Data (ACD) in this dissertation. The ACD dataset contained 2,493 weather related crashes with 3,717 injured occupants. As this corridor-based analysis involved both weather and non-weather related crashes (ACD), RWIS data was considered as the preferable source of weather information rather than the cost center data considered in developing the frequency models previously. Weather related information for both weather and non-weather related crashes were obtained from the RWIS stations located nearest to the crash sites.

Visibility information was extracted from the Automated Weather Observing System (AWOS) consisting of weather stations located in airports. AWOS weather stations report visibility information at a ten minute interval. Weather data from these two sources were merged with the crash data based on the date, time, and location for 5,242 crashes that occurred along the I-80 corridor during the four winter seasons. There were instances with missing weather information from the nearest RWIS/AWOS station for crashes. In these cases, weather information from the second nearest RWIS/AWOS station from crash sites was obtained. Table 3.8 and 3.9 lists the RWIS and AWOS stations considered for extracting weather related information and the corresponding second nearest RWIS and AWOS stations used to fill in the missing data. This data set was then disaggregated at the vehicle level and then at the occupant level resulting in three level of aggregation and three data sets:

i) crash based data set - one level including details on crashes but aggregated information about vehicles and occupants;

ii) vehicle based data set – two levels including details on both crashes and vehicles involved but aggregate information about occupants, and

iii) occupant based data set – including details on crashes, vehicles, and occupants.

These data sets were used to develop crash severity models. The advantage of using this disaggregate data set is in making full use of the information available in the crash data while at the same time accounting for possible correlations in the severity levels of occupants (similar to this research) or vehicles involved in a given crash. Figure 3.7 shows the locations of the RWIS stations and airports with AWOS stations along the study corridor in the I-80 study corridor. The green circles indicate RWIS stations and the red planes indicate airports with AWOS stations.

Figure 3.6 The study corridor considered for crash injury severity models

Table 3.8 RWIS stations and next nearest RWIS stations considered for weather data

RWIS Stations	Second Nearest RWIS stations
Adair	Avoca
Altoona	Des Moines I-235
Avoca	Adair
Colfax	Altoona
Council Bluffs	Avoca
Davenport	Quad Cities
De Soto	De Moines I-235
Des Moines I-235	Altoona/De Soto
Grinnell	Colfax
Iowa City US 218	Iowa City I-80/Tipton
Iowa City I-80	Iowa City US 218/Tipton
Quad Cities	Davenport
Tipton	Iowa City US 218
Williamsburg	Iowa City US 218

Table 3.9 Airports with AWOS stations and next nearest airports for visibility data

Airports with AWOS stations	Second nearest Airport with AWOS stations
Ankeny	Newton
Atlantic	Harlan/Audobon
Audobon	Harlan
Council Bluffs	Harlan/Audobon
Grinnell	Newton/Perry
Harlan	Atlantic
Muscatine	Washington
Newton	Grinnell
Perry	Ankeny/Audobon
Washington	Muscatine

Figure 3.7 Locations of the RWIS stations and AWOS stations near to the I-80 study corridor

3.4.3 Exploratory Data Analysis

Crashes were categorized into six injury severity levels according to Iowa DOT as fatal, incapacitating, non-incapacitating, possible, uninjured, and unknown. The percentage of fatal, incapacitating, non-incapacitating, possible, and unknown injuries resulting from the crashes considered is merely 20% of the total. A large number of factors can influence the severity of crashes under winter conditions. Therefore, it was important to consider all the factors available from the prepared crash data sets for the multi-level analysis of occupant injury severity. Most of the variables considered to be used in the severity analysis were categorical. Table 3.10 provides a list of the variables considered in the crash severity analysis. It is to be noted that not all the variables were included in the severity models developed in this study. Some variables were derived from the variables presented in Table 3.10 for developing the models. Details about these variables are presented in table 3.11 – 3.13 for the three models developed using weather-related crashes, all crashes, and non-weather related crashes for the winter seasons from 2008-09 to 2011-12. Figure 3.8 shows the distribution of injury severity for these three types of crashes during these winter seasons.

Distribution of injury severity by weather related variables such as air temperature, pavement temperature, and visibility are provided in Appendix H.

Table 3.10 Variables considered for developing the multilevel crash severity models

Variable	**Definition**
Gender of the occupant	Male = 1, Female = 0
Seating position of the occupant	Driver/Motorcycle driver = 1, otherwise = 0
Occupant protection used	None = 1, safety protection used = 2, unknown/not reported = 3
Ejection status	Not ejected = 1, ejected = 2, N/A, unknown, not reported = 3
Air bag deployment	Deployed = 1, not deployed = 2, unknown, not reported = 3
Trap status	Not trapped = 1, trapped = 2, unknown, not reported = 3
First harmful event	Non-collision event = 1, collision w/ other vehicle = 2, collision with other objects (non-vehicle) = 3
Manner of crash/collision	non-collision = 1, rear end = 2, sideswipe = 3, others (head-on, angle, broadside, unknown, not reported) = 4
Contributing circumstances - Environment	Environment as a contributing circumstances = 1, otherwise = 0
Weather conditions	Clear = 1, cloudy/partly cloudy = 2, others (including rain/sleet/hail, freezing rain)=3
Light conditions	Daylight+dawn = 1, dusk+dark = 2
Surface conditions	Dry = 1, wet/ice, snow/slush = 2, other, not reported = 3
Time of day	8 am to 6 pm = 1, midnight to 8 am = 2, 6pm to 12 am = 3
Location of first harmful event	On roadway = 1, shoulder = 2, median = 3, roadside and others = 4
Location of the crash (urban)	In urban area = 1, otherwise = 0
Location of crash (ramp)	On ramp = 1, otherwise = 0
Contributing circumstances - Roadway	Road surface condition as the contributing circumstances = 1, otherwise = 0
Type of roadway junction	Intersection = 1, otherwise = 0
Total number of vehicles in a crash	More than one = 1, otherwise = 0
Total number of occupants in a crash	Two or more = 1, otherwise = 0
Air temperature	In degree Fahrenheit
Pavement temperature	In degree Fahrenheit
Visibility	In mile

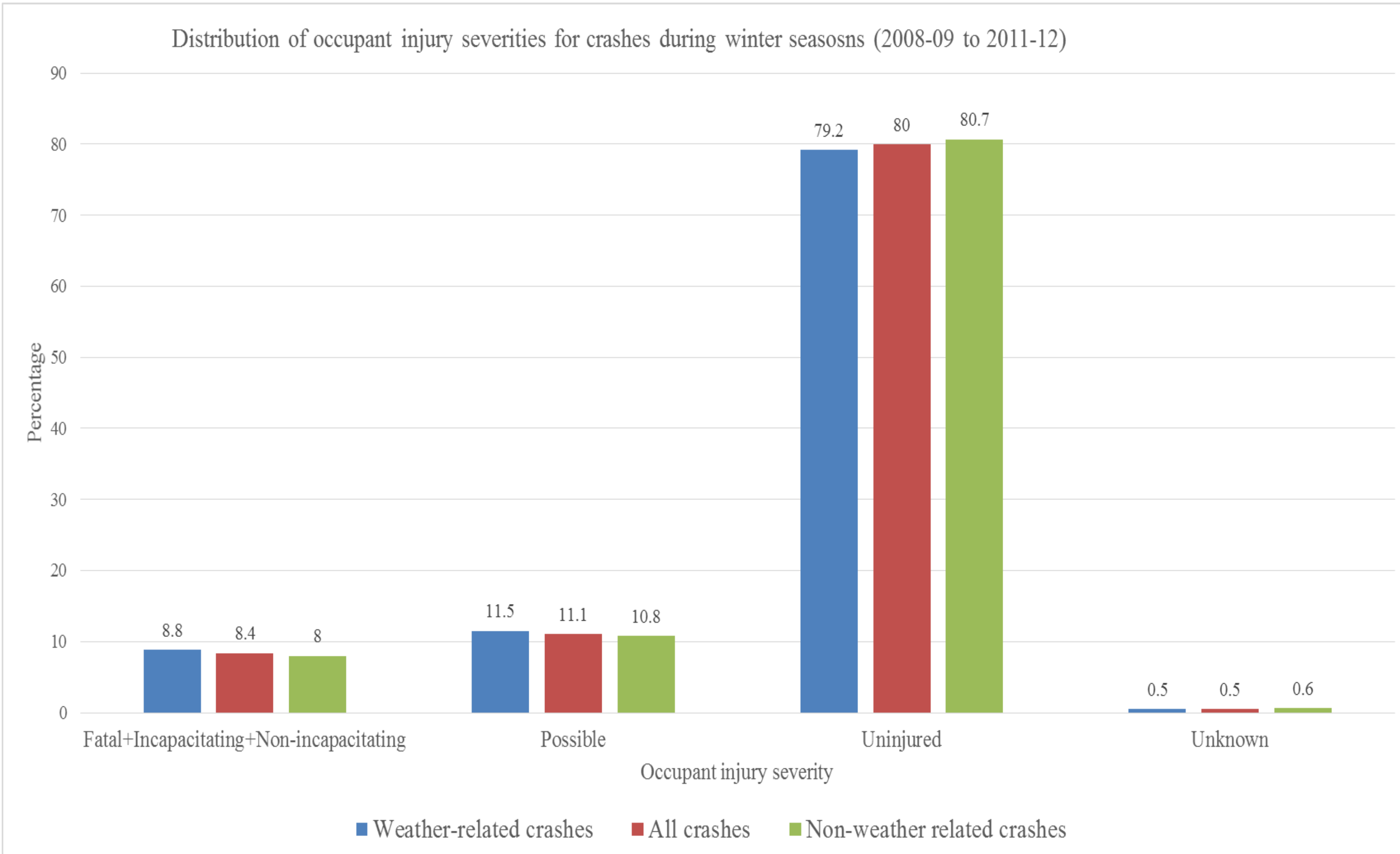


Figure 3.8 Distribution of occupant injury severities for crashes during 2008/09 to 2011/12 winters

Table 3.11 Descriptive statistics of the variables used in the model with weather-related crashes

Parameters	Descriptive Statistics	
	Mean	SD
Gender (if male = 1, otherwise = 0)	0.661	0.473
Seating position (if driver = 1, otherwise =0)	0.926	0.26
Occupant Protection		
None Used (If none used = 1, otherwise = 0)	0.022	0.148
Used (if used = 1, otherwise = 0)	0.875	0.331
unknown/not reported (if unknown or not reported = 1, otherwise = 0)	0.103	0.304
Airbag deployment		
Airbag deployed (if yes = 1, otherwise = 0)	0.1	0.3
Not deployed (If yes = 1, otherwise = 0)	0.734	0.441
Unknown/not reported (if yes = 1, otherwise = 0)	0.165	0.371
First harmful event		
Non-collision including overturn, rollover, jackknife) (If yes = 1, otherwise = 0)	0.232	0.422
Collision with vehicles	0.512	0.499
Collision with non-vehicles	0.255	0.436
Roadway condition as contributing circumstances (if yes = 1, otherwise = 0)	0.746	0.434
Trapped (if an occupant is trapped = 1, otherwise = 0)	0.036	0.188
Ejection status (if an occupant is ejected = 1, otherwise = 1)	0.05	0.073
Age of the occupant (if age of the injured is > 24 = 1, otherwise = 0)	0.77	0.42
Road type (if intersection = 1, otherwise = 0)	0.12	0.325
Road surface condition and air temperature (if surface icy and temperature below zero = 1, otherwise = 0)	0.426	0.494

Table 3.12 Descriptive statistics of the variables used in the model with all crashes

Parameters	Descriptive Statistics	
	Mean	SD
Gender (if male = 1, otherwise = 0)	0.645	0.478
Seating position (if driver = 1, otherwise =0)	0.935	0.246
Visibility		
Visibility 1 to 3 mile (If yes = 1, otherwise = 0)	0.136	0.343
Visibility 3 to 6 mile (If yes = 1, otherwise = 0)	0.122	0.327
Visibility 6 mile and above (If yes = 1, otherwise = 0)	0.74	0.438
Occupant Protection		
None Used (If none used = 1, otherwise = 0)	0.022	0.146
Used (if used = 1, otherwise = 0)	0.863	0.343
Unknown/not reported (if unknown or not reported = 1, otherwise = 0)	0.114	0.318
Airbag deployment		
Airbag deployed (if yes = 1, otherwise = 0)	0.102	0.302
Not deployed (If yes = 1, otherwise = 0)	0.749	0.433
Unknown/not reported (if yes = 1, otherwise = 0)	0.148	0.356
First harmful event		
Non-collision including overturn, rollover, jackknife) (If yes = 1, otherwise = 0)	0.162	0.368
Collision with vehicles (If yes = 1, otherwise = 0)	0.641	0.479
Collision with non-vehicles (If yes = 1, otherwise = 0)	0.196	0.397
Surface Condition		
Surface condition dry (If yes = 1, otherwise = 0)	0.346	0.475
Surface condition icy, wet, snowy, or slushy (If yes = 1, otherwise = 0)	0.576	0.494
Surface condition others (water, debris, sand, dirt) and not reported (If yes = 1, otherwise = 0)	0.077	0.267
Trap Status		
Not trapped (If yes = 1, otherwise = 0)	0.93	0.251
Trapped (If yes = 1, otherwise = 0)	0.033	0.179
unknown/not reported (if yes = 1, otherwise = 0)	0.034	0.182
Ejection status (If an occupant is ejected = 1, otherwise = 1)	0.04	0.063
Air temperature (If below zero = 1, otherwise = 0)	0.577	0.494

Table 3.13 Descriptive statistics of the variables used in the model with non-weather crashes

Parameters	Descriptive Statistics	
	Mean	SD
Gender (if male = 1, otherwise = 0)	0.63	0.482
Seating position (if driver = 1, otherwise =0)	0.943	0.231
Occupant Protection		
Occupant protection not used, unknown or not reported (If yes = 1, otherwise = 0)	0.146	0.353
Occupant protection used (If yes = 1, otherwise = 0)	0.853	0.353
Airbag deployment		
Airbag not deployed (If yes = 1, otherwise = 0)	0.762	0.425
Airbag deployed (if yes = 1, otherwise = 0)	0.103	0.304
First harmful event		
Non-collision (overturn, rollover, jackknife) (reference)	0.098	0.298
Collision with vehicles	0.757	0.428
Collision with non-vehicles	0.144	0.351
Surface Condition		
If surface has water (moving or standing) (If yes = 1, otherwise = 0)	0.145	0.352
Trap Status		
Not trapped (If yes = 1, otherwise = 0)	0.929	0.256
Trapped (If yes = 1, otherwise = 0)	0.03	0.171
Unknown/not reported (If yes = 1, otherwise = 0)	0.04	0.196
Ejection status		
Not ejected (If yes = 1, otherwise = 0)	0.956	0.204
Ejected (If yes = 1, otherwise = 0)	0.002	0.053
Unknown/not reported (If yes = 1, otherwise = 0)	0.041	0.198
Age of the occupant		
Occupants aged up to 24 years (If yes = 1, otherwise = 0)	0.201	0.401
Occupants aged 24 years or higher (If yes = 1, otherwise = 0)	0.798	0.401
Major cause (if run-off-road = 1, otherwise = 0)	0.119	0.323

CHAPTER 4

METHODS

As described in Chapter 1, the primary objective of this research is to develop quantitative models to understand the factors affecting winter weather crash frequency and crash injury severity. This chapter presents the methods that were developed to accomplish the research objective. In specific, count data models were used to estimate crash frequency relating to winter weather related factors to understand those affecting crash frequency and apply Empirical-Bayes technique to prioritize road segments with potential for safety improvement during winter weather. Crash injury severity models were also developed to account for the hierarchical nature in the crash data using multilevel modeling techniques in a Bayesian framework.

4.1 Prediction Models for Crash Frequency

A Poisson distribution is normally assumed for modeling the probability of crash frequency on road segments. However, when the crash counts are overdispersed (with variance greater than mean of crashes), the assumption that crash counts are Poisson distributed is no longer valid. Thus, negative binomial distribution is used to represent the distribution of crash counts. As both Poisson and negative binomial distribution was used to model winter weather related crash frequency for three different types of roadways in this research, the next section discusses the functional formulation of Poisson and negative binomial regression model.

4.1.1 Poisson Regression Model

Generalized linear models, also known as GLM techniques, are the most commonly employed models for predicting collision/crash frequency. GLM could be applied to model both continuous and discrete dependent variables. For the purpose of this research, it is assumed that crashes over a period of time follow a count process such as Poisson distribution. Mathematically, if the number of crashes (Y) is assumed to follow a Poisson distribution, the probability of crash frequency can be expressed as shown in the equation below

$$P(Y = K) = \frac{e^{-\mu}\mu^{k}}{k!}, K = 0, 1, 2, 3 \dots\dots \qquad (4.1)$$

where P (Y = *K*) = probability of having *K* crashes over a period of time

Y = number of crashes over a period of time

μ = expected number of crashes over a period of time, known as the Poisson parameter.

Poisson regression models are estimated by specifying the Poisson parameter as a function of explanatory variables (geometric conditions of roadways, traffic exposure, pavement conditions, visibility, etc.) potentially having significant impact on the occurrence of crashes over a period. The model parameter μ in Equation 4.1 is commonly assumed to be a function of different factors using a non-linear link function *g (.)*, as shown in the following equation

$$g(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \qquad (4.2)$$

where, β_0 = intercept,

β_k = coefficient of explanatory variable X_k,

X_k = k^{th} explanatory variable which could be related to road, traffic, or weather characteristics.

The most commonly used non-linear link function in road safety modeling is the log link function ensuring positive estimates for the mean. It can be expressed mathematically as follows

$$\ln(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad (4.3)$$

which can also be expressed as

$$\mu = \exp(\beta_0 + \textstyle\sum_{k=1}^{n} \beta_k X_k) \quad (4.4)$$

Coefficients of the explanatory variables can be estimated using the maximum likelihood method (ML) by using the following equation

$$LL(\beta) = \textstyle\sum_{i=1}^{n} [y_i \ln(\mu_i) - \mu_i - \ln(y_i!)] \quad (4.5)$$

where *LL(β)* is the log of the likelihood function. Exposure is one of the most important factors affecting crash frequency, which can be represented by traffic volume, segment length or the cross product of them. The exposure can be included in a crash frequency model either as a variable or as an offset. For the latter case, equation (4.5) can be written as

$$\ln(\mu) = \beta_0 + \textstyle\sum_{k=1}^{n} \beta_k X_k + \gamma \ln(EXP) \quad (4.6)$$

where, *EXP* is the exposure and γ is the exponent of the exposure.

4.1.2 Negative Binomial Regression Model

One limitation of the Poisson model is that the mean of the crash frequency is assumed to be equal to the variance. However, in practice, the variance of crash frequency is normally greater than its mean, which is known as the overdispersion problem. Overdispersion affects the standard error estimates of the parameters (Cameron and Trivedi, 1998) making some insignificant variables significant and drawing incorrect inferences from the model estimation. The Negative binomial distribution can address the phenomenon of overdispersion. The negative binomial model can be derived from the Poisson model by adding a Gamma distributed error term to the Equation (4.3).

$$\ln(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \qquad (4.7)$$

where exp (ε) is assumed to follow a Gamma distribution with both of its parameters equal to φ. The resulting crash frequency (Y) should have a variance which is a function of the mean and φ as given by the following equation

$$Var(Y) = \mu + \varphi\mu^2 = \mu + \frac{\mu^2}{\alpha} \qquad (4.8)$$

where α = 1/φ is known as the over dispersion factor

4.1.3 Empirical Bayes Method

Hot spot identification or prioritizing sites for safety improvement of road networks is an essential task for engineers in state agencies in order to ensure efficient allocation of limited resources for mitigating the safety problems in the identified sites or spots. There are various methods mostly relying on historic traffic crash records to obtain an estimate for safety for various traffic entities. The majority of these traditional methods use raw crash data such as the crash frequency method, the crash rate method, the rate quality control method, the crash severity method, and the safety index method. The most prominent problem associated with these naïve statistical methods to identify hot spots for safety improvements is the Regression to the Mean (RTM) problem. Analysts or engineers must take into account this phenomenon when identifying potential safety issues for a single site/spot or a group of sites/spots. RTM reflects the tendency of the observed crashes to regress or return to the mean in the year following an unusually high or low crash counts. The effect of RTM can arise when sites with high-short-term crash counts are selected as candidate sites for safety improvements or treatments. In this case, the counts of the crashes at these sites would decrease due to the RTM and regress towards their long-term mean irrespective of the implementation of the treatment. So, one can overestimate the safety effectiveness of the implemented treatment if the RTM is not taken into account. Because of the random variation in crash occurrences, the sites with highest number of crashes in one period are very likely to experience lower crash frequencies in the next period, and vice versa. So, relying solely on crash records and using one of these traditional methods does not warrant to account for the RTM and evaluate the effectiveness of a particular treatment aimed at improving safety at particular sites. So despite their simplicities, naïve statistical methods using raw crash records have serious limitations for screening road networks for safety improvement or evaluating the effectiveness of a treatment to trigger safety improvement at particular sites.

In recent years, techniques for screening road networks to identify crash locations have become more sophisticated and require more data as inputs. SPFs are frequently used in the network screening and evaluation process and can be used to reduce the effects of RTM. SPFs can used to estimate the expected safety of a roadway segment or location based on similar facilities. Typical SPFs have been developed to estimate crash frequency using site or roadway characteristics such as lane width and traffic exposure expressed as AADT. These typical SPFs normally do not incorporate weather related variables, as this would be more complex and labor intensive. This study develops SPFs for three function classes of roadways in Iowa to predict winter weather related crashes as a function of several factors related to winter weather conditions such as visibility, pavement temperature, air temperature, and wind speed. The Empirical-Bayes approach is used to combine the predicted number of crashes from the SPFs with the observed crash counts at a location to produce an improved estimate of the expected number of crashes. As crashes are random in nature, the Empirical-Bayes method takes into account the phenomenon of RTM. Extensive research has also shown that the Empirical-Bayes approach is the most consistent and reliable method for identifying sites with potential for safety improvement (Cheng and Washington, 2008).

The implementation of the Empirical-Bayes method is connected with the results from the modeling performed during the development of SPFs. Using the overdispersion parameter found during modeling (crashes fitting negative binomial model), a weight can be determined as follows:

$$w = \frac{1}{1+\alpha(n * E(\mu))} \tag{4.9}$$

where α is the overdispersion parameter derived from the SPFs modeled with negative binomial distribution and $E(\mu)$ is the predicted number of crashes for a given roadway with n being the number of years for crash observation.

The weight factor is then applied to the predicted number of crashes (calculated from SPFs) and actual observed number of crashes to determine the estimated number of crashes as follows:

$$\lambda = w.E(\mu) + (1-w)k \tag{4.10}$$

where λ is the improved estimated number of crashes and k is the total number of crashes observed in n years.

The difference between the EB adjusted crash frequency and the predicted crash frequency from a SPF is referred to as the Potential for Safety Improvement of PSI. The higher the PSI value for a road segment, the higher the potential for improving safety along that road segment. Considering the PSI, the roadway segments are ranked or prioritized for investing resources at those locations so that highest possible safety improvement can be achieved. The following figure represents the graphical definition of the PSI.

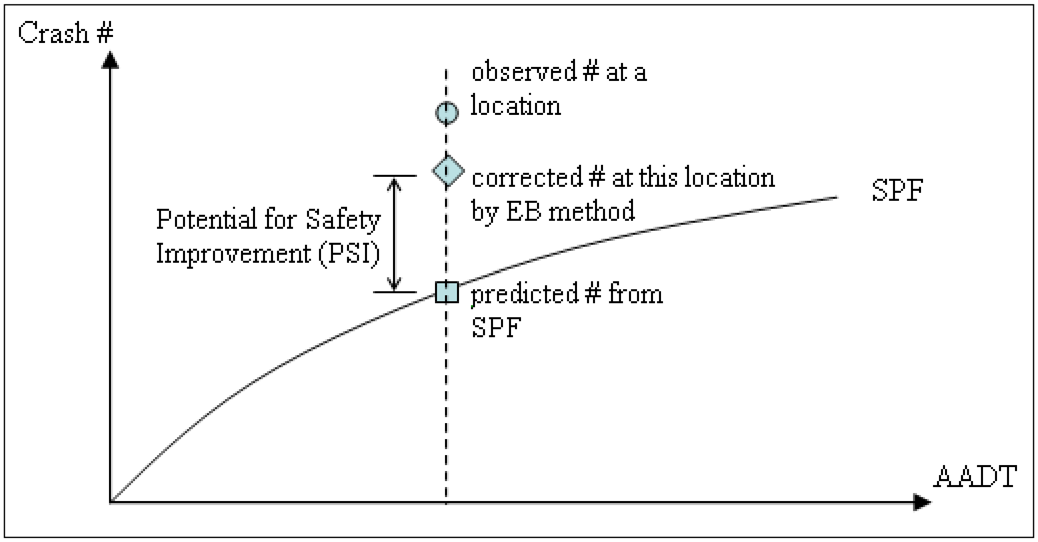


Figure 4.1 Calculation of PSI using Empirical-Bayes method (Source: Illinois DOT)

4.2 Multilevel Modeling Approach for Severity Analysis

Most of the crash data used for road safety research are of hierarchical nature and belong to structures with several hierarchically ordered levels. These hierarchical structures could be attributed to the spatial (and temporal) spread of the data or the hierarchical nature of the crash data where individuals or occupants are nested within vehicles and vehicles are nested within crashes. Information regarding vehicles, drivers or occupants is clustered within the crashes as each vehicle, driver, or occupant observation pertains to one crash only. These two types of hierarchies are associated with aggregated and disaggregated crash data respectively and can be distinguished as geographical and crash hierarchies. The analysis of aggregate crash data mainly focuses on the geographical part of the hierarchy, which can be characterized by accounting for the spatial dependence through spatial analyses. The analysis of disaggregate crash data focuses mainly on the individual occupants or vehicles involved in a crash requiring the crash hierarchy to be taken into consideration. The following figure illustrates the two hierarchies which are actually complementary and have been incorporated into a single framework by Huang and Abdel-Aty (2010) and Dupont et al. (2013) to represent the prevailing data structure in road safety. The current research focuses on addressing the crash hierarchy of disaggregated crash data by nesting occupants within crashes.

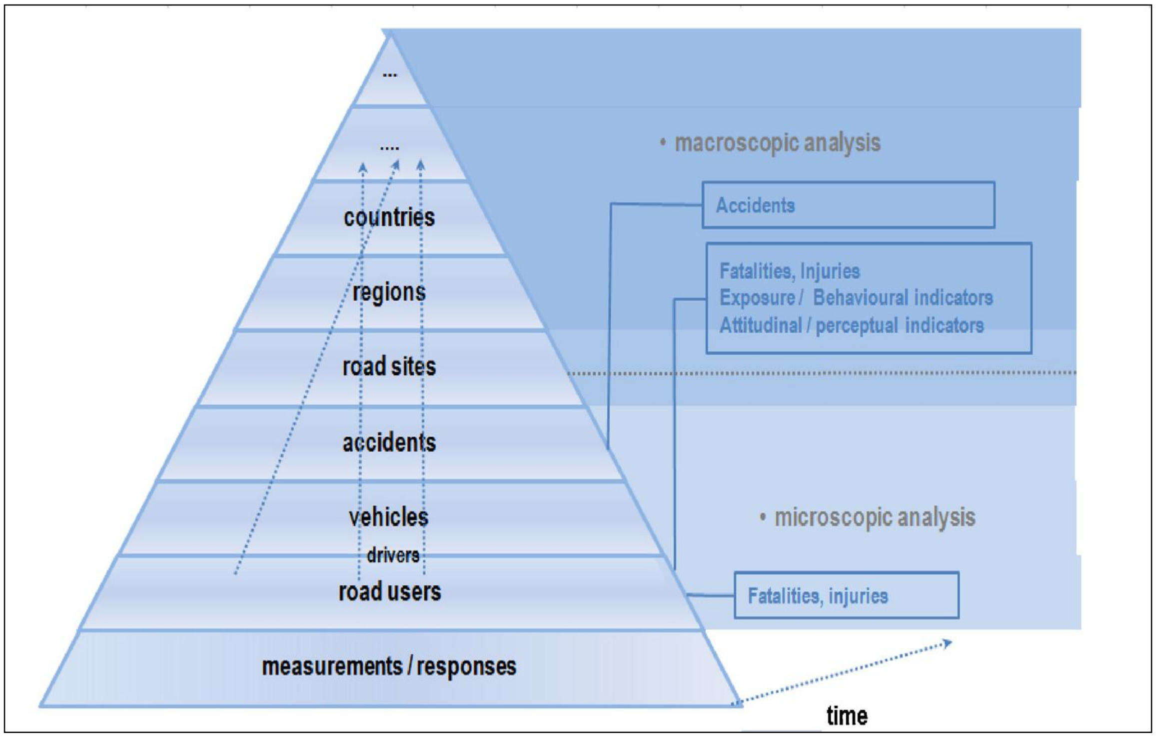


Figure 4.2 Multilevel structure of crash data (Adapted from Huang and Abdel-Aty, 2010 and Dupont et al., 2013)

4.2.1 Multilevel/hierarchical Models

Multilevel/hierarchical (ML/HL) models are regression models (linear or generalized linear models) with parameters that have been assigned a probability distribution. Hyperparameter is a term used in the ML/HL models to describe the parameters of the probability models. In this context, hierarchical models can be described in a Bayesian paradigm. For most of the multilevel models, conventional estimation methods can be used such as maximum likelihood or quasi-likelihood, which are based on Generalized Least Square (GLS) estimation (Browne et al., 2001; Dupont & Martensen, 2007). However, an important problem associated with these methods is that the likelihood estimation ratio is very approximate and cannot be used for assessing model fit. Applying these methods to complex data structure often result in numerical and convergence difficulties. Thus hierarchical/multilevel models can be grounded on Bayesian paradigm. Bayesian inference is the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations. Model parameters are assigned a probability distribution representing the knowledge the researcher possesses about each parameter prior to making any observation of data. These prior distributions may be informative or non-informative. Informative prior distributions about the parameters may be derived from the assumptions made by researchers from existing knowledge. To compensate for the lack of knowledge about the prior distributions of parameters, it is common to assign a typical distribution to these parameters with relatively large variance. Specifically, in Bayesian models, given model assumptions and parameters, the likelihood of the observed data is used to modify the prior knowledge of the unknowns, resulting in the updated knowledge summarized in posterior densities. Inference about the parameters is based on the posterior distribution, which is a combination of the prior information with information derived from the observations.

Although a wide range of literature on clustered/hierarchical data analysis with full and empirical Bayes application exists in health service research, econometrics and other fields, the application of full Bayesian treatment is still limited in the transportation research for addressing clustered/hierarchical nature of the crash data. With the significant computational advances enabling the estimation of formerly complex models, the Bayesian framework combined with Markov Chain Monte Carlo estimation enable the estimation of unordered discrete choice models. As this estimation method is based on interval estimates, it allows for the calculation of accurate likelihood values (Dupont & Martensen, 2007). This simulation-based Bayesian method is more powerful in dealing with complex hierarchical datasets along with missing data or few data. This method also incorporates all sources of uncertainty in estimating the random effects accounting for the variation in the hierarchical data.

4.2.2 General Model Formulation

To define the principles of multilevel models, a simplified two-level model can be used. Occupants nested within crashes are considered as the hierarchical form in this model formulation. For example, let the response variable be the probability of injury for occupant i involved in crash j. The response y_{ij} can only takes one of the two values: 1 in case of injury while 0 in case of no injury. The probability of $y_{ij} = 1$ is denoted by $\pi_{ij} = \Pr(y_{ij} = 1)$ following a binomial distribution.

The expected average probability for each crash can be defined as a non-linear function of the predictor or combination of predictors in linear form

$$\pi_{ij} = f(X\beta_{ij}) \tag{4.11}$$

If *f(x)* represents a link function, the logit link chosen in this case can be represented by the following equation

$$Logit(\pi_{ij}) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hij} + u_{0j} \tag{4.12}$$

The expected probability for driver i to be injured in crash j is now defined as being a logit function of the linear combination of an average value holding for the crash population (γ_0), of the effect of driver level and crash level predictors ($\sum_{h=1}^{r} \gamma_h x_{hij}$), and of crash-related random effect u_{oj} which is assumed to be normally distributed with mean 0 and variance $\sigma^2_{u_{0j}}$.

The model defined in equation 4.12 allows the intercept to vary across the crashes. Thus, this allows the expected probability of occupants' injury to be higher for some crashes than others. This random variation is intended to capture the unobserved characteristics of the crash-level units on the observations made of the of individual driver-level units.

Intra-class correlation (ICC) coefficients can be employed to calculate the proportion of variance in the outcome probability that is associated with each of the two levels (crash and driver) considered in this model formulation. This coefficient establishes the ratio of crash-level variance to the total variance in the outcome

$$\rho = \frac{\sigma^2_{u_0}}{\sigma^2_{u_0} + \sigma^2} \tag{4.13}$$

The logistic distribution for the driver-level residual implies a variance of $\pi^2/3 = 3.29$. So, for a two-level logistic random intercept model with an intercept variance of $\sigma^2_{u_0}$, the ICC for between-crash residuals is

$$\rho = \frac{\sigma^2_{u_0}}{\sigma^2_{u_0} + \frac{\pi^2}{3}} \tag{4.14}$$

The ρ is an indicator of the magnitude of the within-crash correlation. A ρ value close to 0 indicates a lack of variation among the crash observations indicating that multilevel modeling is not warranted. On the other hand, a relatively large value of ρ implies an inclination for estimating the multilevel or hierarchical model to fit to the data having hierarchical structure. Literature shows (Huang et al., 2008; Dupont et al., 2013) that ICC values ranging from 0.25 and above are considered high in terms of explaining the variance at the higher level (crash level considered herein).

4.2.2.1 Hierarchical Binary Logit Model in Bayesian Framework

A hierarchical binary logit model with two-level specification was considered for the current research. In equation 4.11, if the response variable y_{ij} only takes one of two values: y_{ij} = 1 in case of the occupant i sustaining injury in crash j and y_{ij} = 0 in case of not sustaining injury, then equation 4.13 can be re-written with π_{ij} following a binomial distribution

$$Logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hij} + u_{0j} \quad (4.15)$$

Full Bayesian inference was employed in this research. A Bayesian framework requires a researcher to think about prior information available on the parameters being estimated and to formally include that information in the model. If no prior information is available for the parameters of interest, one must specify an uninformative prior. The posterior distribution for a parameter θ given that observed data is y is subjected to the following rule according to Bayesian statistics

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (4.16)$$

where p (θ|y) is the posterior distribution for θ given observed y, p (y|θ) is the likelihood of observing y given θ, and p (θ) is the probability distribution arising from some statement of prior belief for the parameter.

The key hierarchical part of the model mentioned in equation 4.16 is the random effect u_{0j} which can be specified at the individual crash level by allowing its variance (τ_0^2) to follow a certain distribution varying across the crashes. For the current model specification, the crash level random effect was assumed to have a normal distribution with the variance of the normally distributed random effect having an inverse Gamma distribution (0.001, 0.001) as shown below

u_{0j} ~ N (0, τ_0^2) and τ_0^2 ~ Inverse-Gamma (0.001, 0.001)

Based on the full Bayesian inference, the joint prior distribution for the parameters (θ) and the random effects (represented by φ) is

$$p(\varphi,\theta) = p(\varphi)p(\theta|\varphi) \tag{4.17}$$

and the joint posterior distribution can be defined as

$$p(\varphi,\theta|y) \propto p(y|\varphi,\theta)p(\varphi,\theta) = p(\varphi,\theta)p(y|\theta) \tag{4.18}$$

In the absence of strong prior information for the model unknowns, uninformative priors were assumed for all regression coefficients (γ_0,γ_h) with normal distributions (0, 1000). Based on the above formulation, the model was computed via Metropolis-Hastings sampler, a Markov Chain Monte Carlo (MCMC) technique, which was implemented by using MLwiN software package (Rasbah et al., 2000). In MLwiN, the user does not have to choose between Gibbs sampling and Metropolis Hastings sampling directly, the software has the capability to choose the default and the most appropriate technique for the given model. In case of normal response models, Gibbs sampling is used for all parameters. For non-normal responses, MLwiN does not allow Gibbs sampling. For logistic regression models, Metroplis-Hastings algorithm is usually quicker than the Gibbs sampler approach usually used in the software WinBUGS. 5,000 iterations were used as the burn-in step for each model. This is the number of initial iterations that were not allowed to be used to describe the final parameter distributions. The monitoring chain length was decided from the Raftery-Lewis diagnostic. The Raftery-Lewis diagnostic (Raftery & Lewis, 1992) is a diagnostic based on a particular quantile of the distribution of a parameter. This diagnostic is used to estimate the length of the Markov chain required to estimate a particular quantile to a given accuracy. In MLwiN, the diagnostic is calculated for the two quantiles with the defaults being the 2.5% and 97.5% quantiles. More iterations would be required if the quantile values are greater than the number of iterations used for a parameter. A thinning to retain every tenth sample was used to reduce the autocorrelation. Thinning is the frequency with which successive values in the Markov chain are stored. Convergences of the models can be checked by monitoring the Markov Chain Monte Carlo (MCMC) dynamic plot traces for all the parameters considered in the models. The conclusion on the model convergence can be made once all the values of the parameters lies within a zone without strong periodicities. Plots of autocorrelation known as Auto Correlation Function (ACF) for all the parameters were also observed to make sure that the chain for each parameter was adequately close to independently and identically distributed (IID) data. A diagnostic known as the Deviance Information Criteria (DIC) is used to measure how well a model fits the data. It is derived by using the deviance with MCMC sampling. The DIC diagnostic can be used to compare models as it consists of the measure of fit and complexity of a particular model.

CHAPTER 5

DEVELOPMENT OF SITE PRIORITIZATION TECHNIQUES FOR WINTER WEATHER CRASHES IN IOWA

This chapter describes the development of a comprehensive site prioritization technique for identifying road segments with winter weather related crash problems using traditional naïve statistical methods. This chapter also discusses the development and results of SPFs used to develop a site prioritization technique using the Empirical-Bayes method to overcome the serious limitations possessed by the traditional methods as discussed before.

5.1 Development of Site Prioritization Techniques Using Naïve Statistical Methods

This section describes the development of a comprehensive site prioritization technique for identifying sites with winter weather related crash problems using crash data from 2002 to 2009. System-wide screening approaches were developed to identify and prioritize sites for further in-depth winter safety analysis. Combined metric analysis, standard deviation based analysis, and moving average analysis were employed to identify and analyze sites for winter-safety analysis.

5.1.1 Combined Metric Analysis

Winter weather-related crash density, crash proportion (the proportion of winter-weather related crashes to all winter crashes), and personal-level injury severity (injuries on each roadway segment by frequency and severity) were considered in the evaluation of one-mile roadway segments with respect to winter weather safety. Crash density was represented by crash frequency on a particular road segment by summing up the total winter-weather related crashes occurring during the winter seasons of 2002 to 2009 and dividing the total by the length of the road segment and the number of years in the analysis period. A Crash proportion metric was computed by summing the total winter-weather related crashes and dividing by the total winter crashes over a road segment for the analysis period. A Severity metric was computed from a total score assigned to each road segment based on the total frequency of injury severities experienced in winter weather-related crashes. The following equations show the three different metrics used for deriving the combined metric.

$$Crash\ density = (crash\ frequency\ on\ a\ road\ segment) \div (length\ of\ the\ road\ segment \times number\ of\ years)$$

$$Crash\ proportion = winter\ weather\ related\ crashes \div total\ crashes\ during\ winter\ season$$

$$Crash\ severity = \sum_i Crash_i \times Point_i$$

where *i* = fatal, major, minor, and possible injury crash.

The combined metric was created by computing the total frequency of injuries by severity over an analysis period and assigning each injury severity a certain number of points, based on the standard Iowa DOT scale shown in Table 5.1. It is to be noted that all three metrics for a specific one-mile road segment were normalized in order to index the metrics against a maximum value of 1.0. In order to do that, the crash density and crash severity metrics were divided by the maximum value of the metrics for a common type of roadway and an analysis period from 2002 to 2009. The crash proportion metric was already normalized against a maximum value of 1.0 as it was expressed as a percentage of total winter crashes during the analysis period. The resulting values were aggregated into a combined score

Table 5.1 Standard Iowa DOT scale for assigning points by injury severity

Injury severity	Points each occurrence
Fatality	200, with the first fatality at a site treated as a major injury
Major	100
Minor	10
Possible/Unknown	1

for each roadway segment. The combined score for each road segment was computed by putting equal weight to the normalized score of each of the three metrics. In this process, a one-third (1/3) weighting was assigned to each of the three metric for computing the combined score. While equal weight was put on the three metrics in this research to consider all three metrics equally important, it may be preferred to weight the three metrics differently. Combined road segment scores based on alternate weighting can be recomputed in the future at the discretion of concerned officials from Iowa DOT. The following equation shows the calculation of the combined score by putting equal weight on the three metrics.

$$Combined\ sc = \left(\frac{1}{3}\right) \times Crash\ density + \left(\frac{1}{3}\right) \times Crashpropportion + \left(\frac{1}{3}\right) \times Crash\ severity \qquad (5.1)$$

This combined score was categorized based on their relative magnitude within the appropriate road type and analysis period. The total mileage of a common road type (i.e. interstate/freeway and tow-lane roads) was computed and categories were created based on the combined score being among a certain percentage of the system mileage. Following the United States Road Assessment Program (usRAP) risk-mapping protocol, the five categories included five derived ranges for the combined metric. Table 5.2 shows the categories with the subsequent percentage of system mileage. Category 1 indicates the values of the combined score within the lowest 40 percent of the total mileage of a road type while category 5 indicates the values within highest 5 percent of the total system mileage. Road segments which were assigned to category 5 indicated maximum safety problems in term of winter weather safety when crash density, crash proportion and crash severity are considered while road segments belonging to category 1 were considered to be least prone to winter weather crashes. Figure 5.1 and 5.2 shows the map of interstate and tow-lane roadway systems categories based on the combined score considering crash density, crash proportion, and crash severity.

Table 5.2 Mileage category ranges (for each road type) by relative magnitude

Category	**Metric value is among percentage of system mileage**
1	Lowest 40 percent
2	Next 25 percent
3	Next 20 percent
4	Next 10 percent
5	Highest 5 percent

Figure 5.1 shows that Interstate 35 north has quite a few segments with higher categories compared to the Interstate 35 south. In fact, there is a continuous stretch of roadway belonging to categories equal or greater than 4. On the other hand, Interstate 80 has road segments with high category values scattered from east to west with the east portion having quite a few black and red colored segments. Interstate 80 has less clustered high category road segments unlike Interstate 35. The map for two-lane roadways also shows the segments or the cluster of segments to identify sites prone to winter weather crashes. Although it is difficult to derive any pattern from the two-lane roadway map, segments belonging to higher categories are more prominent in the upper region of the map compared to the lower region.

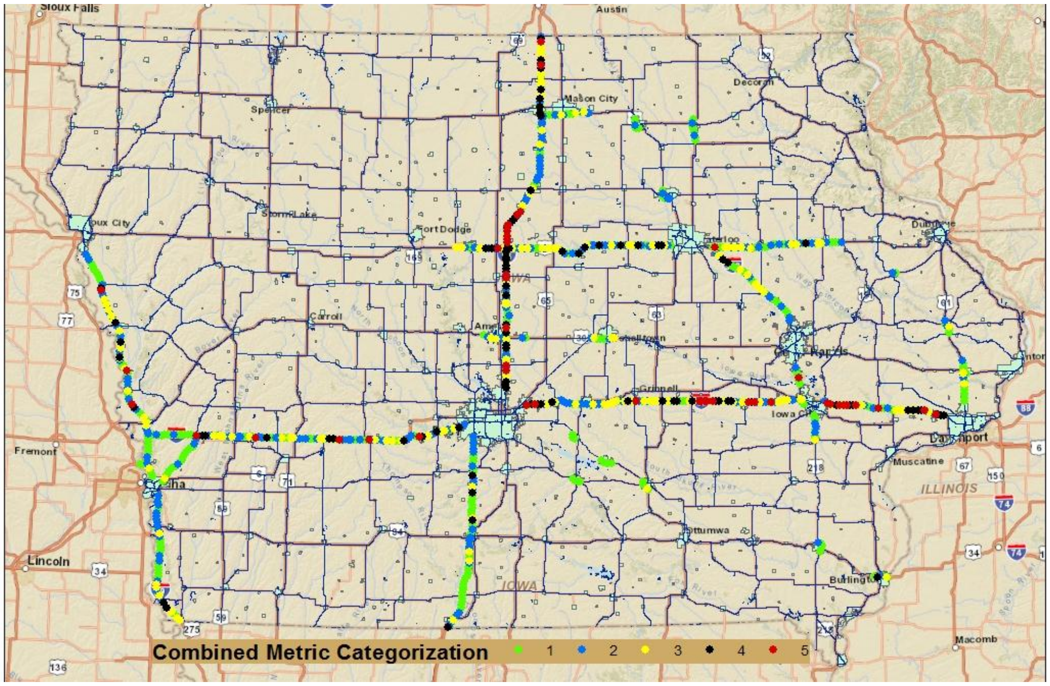


Figure 5.1 Combined metric categorization for Interstate/freeway road segments

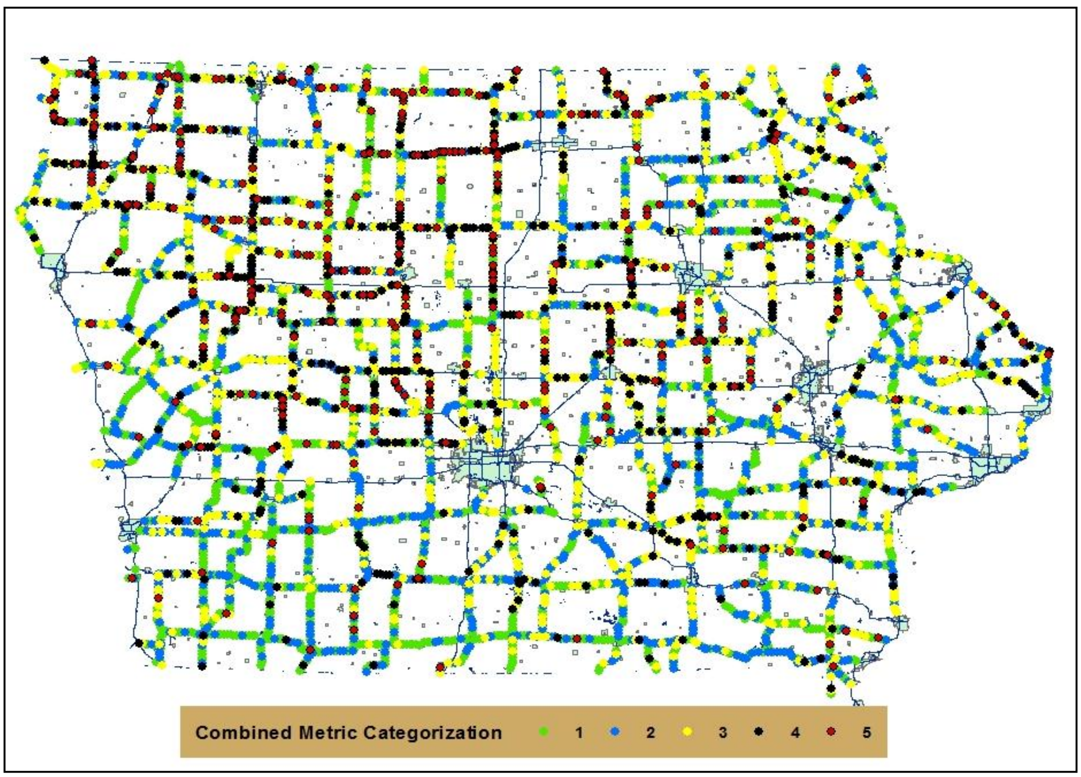


Figure 5.2 Combined metric categorization for two-lane road segments

5.1.2 Standard Deviation Based Analysis

A standard deviation based analysis on one-mile road segments for different functional types of roadway was conducted using the winter weather-related crashes from 2002 to 2009. The combined score (putting equal weight on crash frequency, crash proportion, and crash severity) was used to examine if the metric on one-mile road segments was within a certain range of the standard deviation for the crashes. This accounted for the wide variation in the combination of the three metrics on the road segments for a common road type. The metrics for crash density, crash proportion, and crash severity were also considered for the standard deviation based analysis. For this purpose, the metrics of all winter weather-related crashes (from 2002 to 2009) for each one-mile road segment were divided by the standard deviation of each metric for each type of roadway. The road segments were ranked based on this value with roads having higher values ranked higher to prioritize the segments for a common type of roadway. The following figures (Figure 5.3-5.5) show the standard deviation based analysis for the three individual metrics for interstate roadways along with the combined metric for the interstate and two-lane roadways (Figure 5.6-5.10)

Notable differences can be observed for prioritized road segments using the standard deviation based analysis of the three metrics. The analysis based on crash density shows a considerable number of roadway segments having yellow, black and red colors along the Interstate 80 corridor while a region with the potential for safety improvements (sites with promises) is shown in the north portion of Interstate 35. Similar results were obtained for the standard deviation based analysis of crash proportion metric. Interstate 35 shows road segments having the potential to be considered for safety improvements along the whole length from north to south with north portion having significant road segments with yellow, black, and red colors and south portion having mostly yellow colored segments with a very few red colored segments.

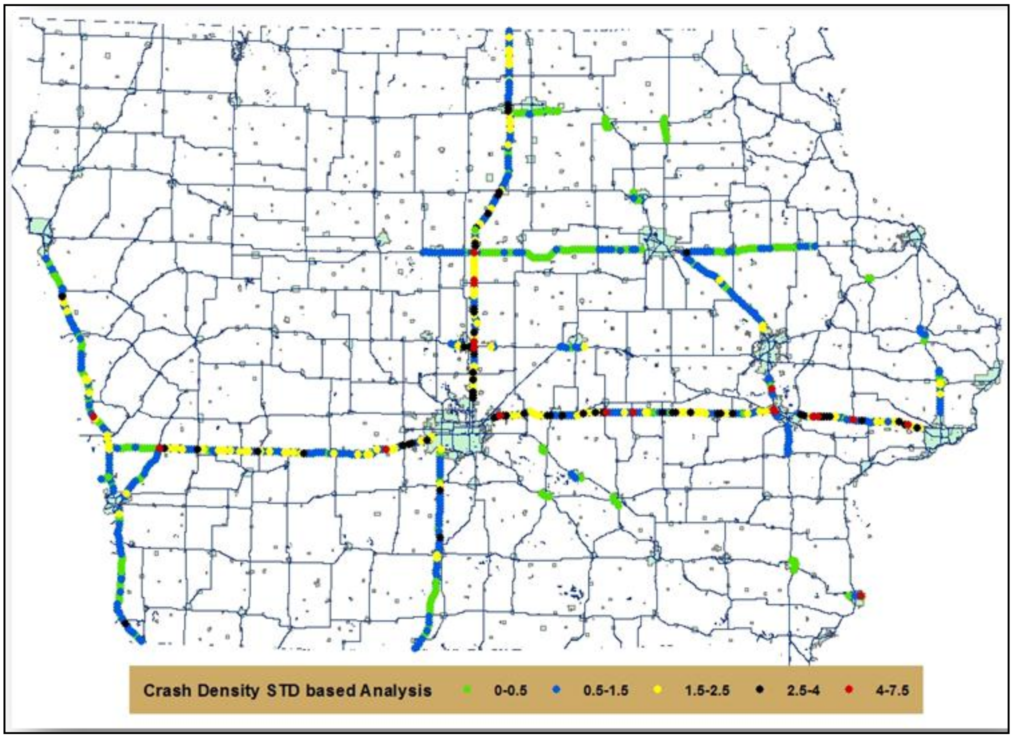


Figure 5.3 Standard deviation based analysis of crash density for Interstate/freeway

It is to note that Iowa DOT can utilize the three metrics (crash density, crash proportion, and crash severity metric) instead of the standard deviation based crash density, crash proportion, and crash severity metrics. The standard deviation based analysis was conducted to observe the sensitivity of the three different metrics with respect to the standard deviation. Caution also needs to be exercised for using any of these three metrics independently as using a single metric without considering the others might result in roadway segment to be wrongly prioritized. This is one reason to devise a combined metric by considering all three metrics.

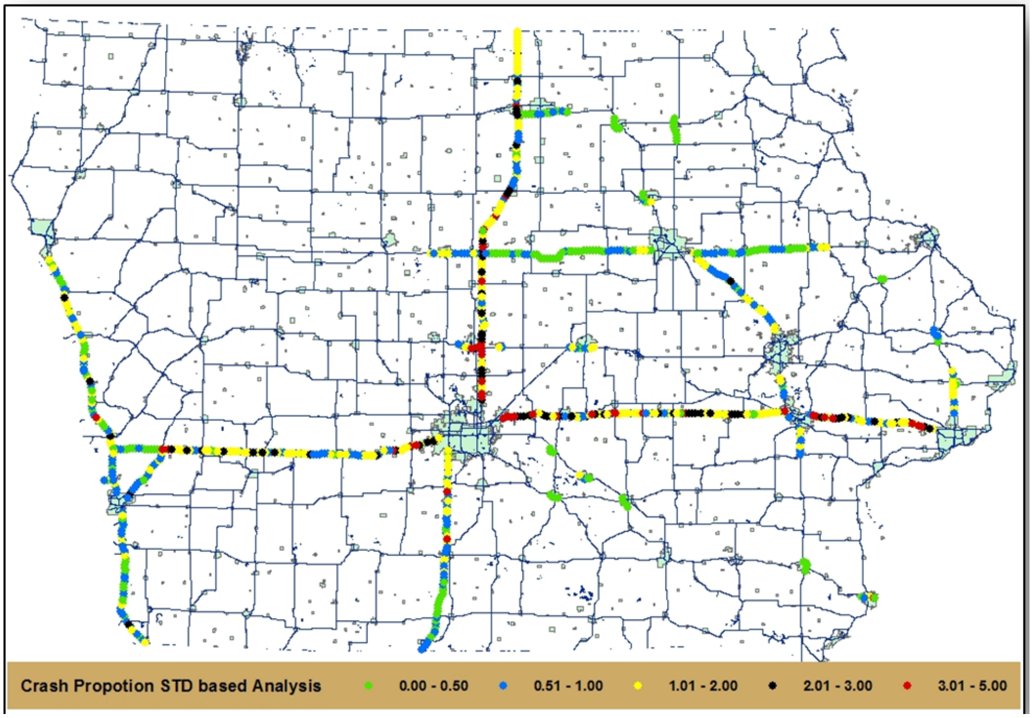


Figure 5.4 Standard deviation based analysis of crash proportion for Interstate/freeway

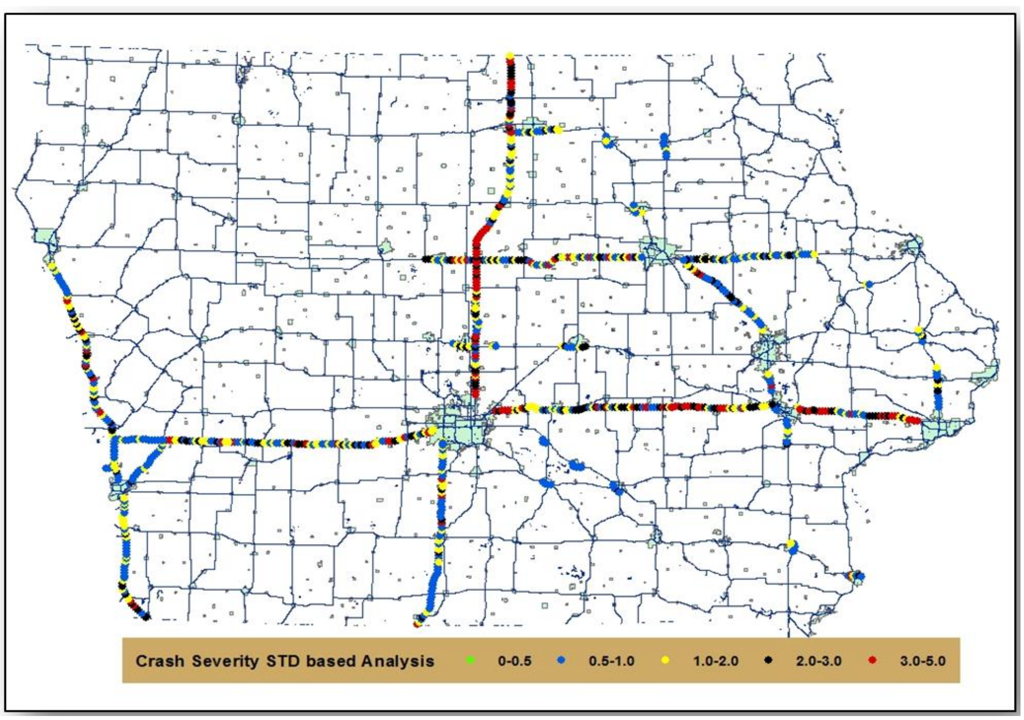


Figure 5.5 Standard deviation based analysis of crash severity for Interstate/freeway

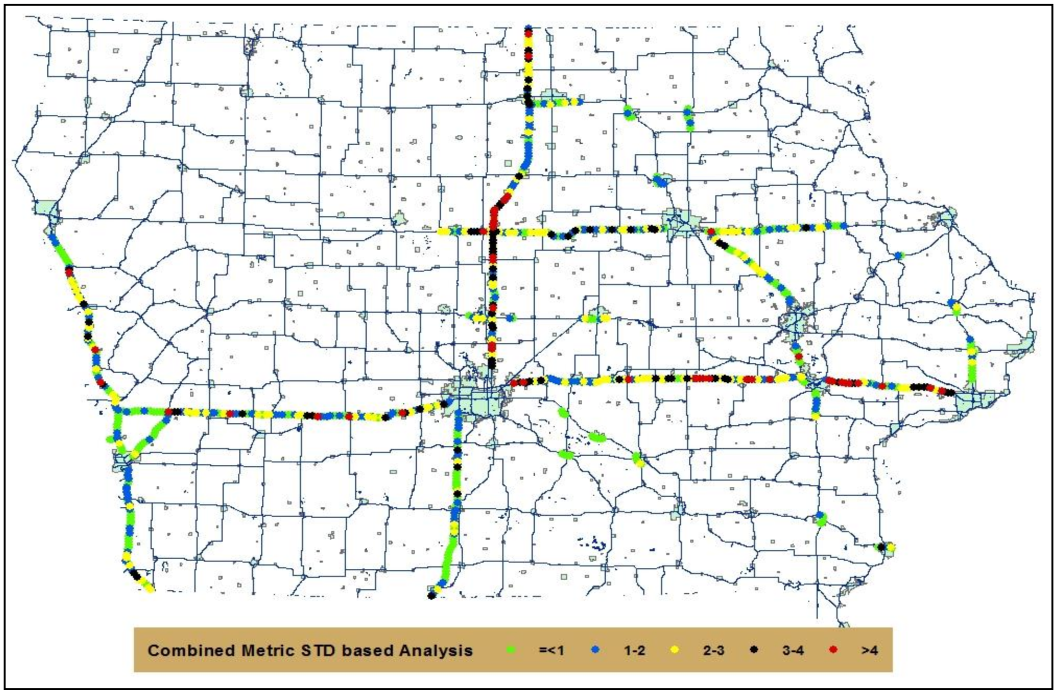


Figure 5.6 Standard deviation based analysis of combined metric for Interstate/freeway

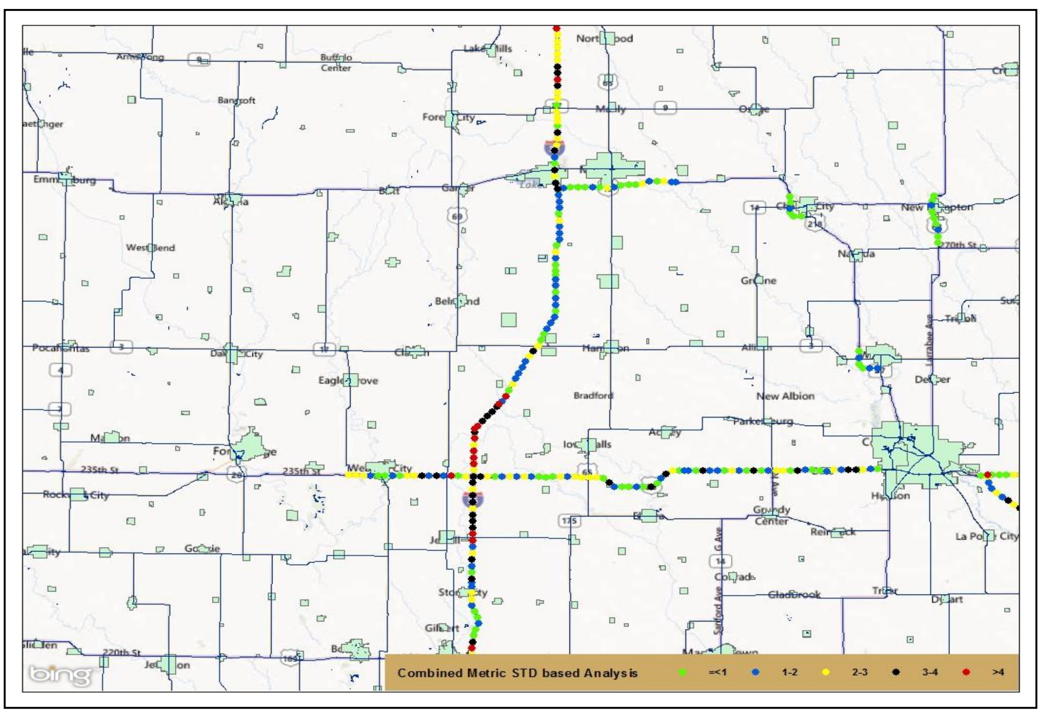


Figure 5.7 Standard deviation based analysis of combined metric for I-35 N

Figure 5.5 shows a continuous stretch along Interstate 35 north having red segments. This color indicates occurrence of severe winter weather-related crashes along this stretch. Similar but shorter stretches were also noticeable along the east portion of Interstate 80.

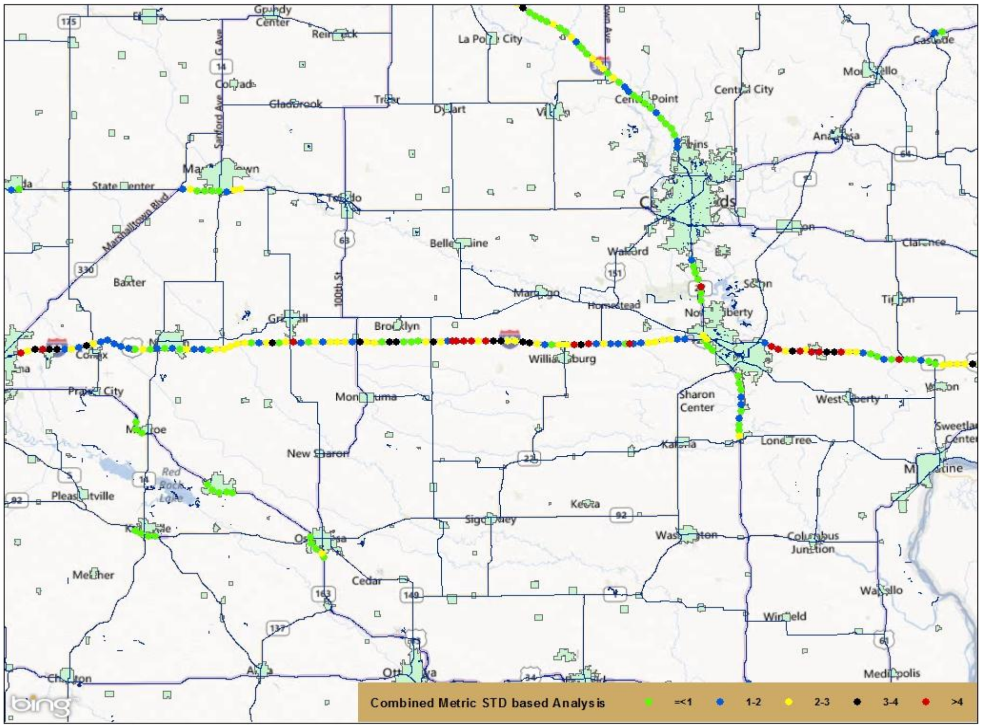


Figure 5.8 Standard deviation based analysis of combined metric for I-80 E

The standard deviation based analysis of the combined metric shows highly prioritized road segments along Interstate 80 east and Interstate 35 north with some hot spots along US 20. Figure 5.7 and 5.8 show high-resolution images of the I-35 north, US 20, and I-80 east. Results show that the curved portion of the I-35 north has road segments with consistent black and red colors. The I-80 east section also has some hot spot clusters near Williamsburg and Iowa City with the road segments near Iowa City having a curved section. The standard deviation based analysis for two-lane road segments can be used to identify crash hotspots with segments having higher values indicated by black and red colors. Figure 5.10 shows only the highly prioritized road segments based on the standard deviation based analysis of the combined metric.

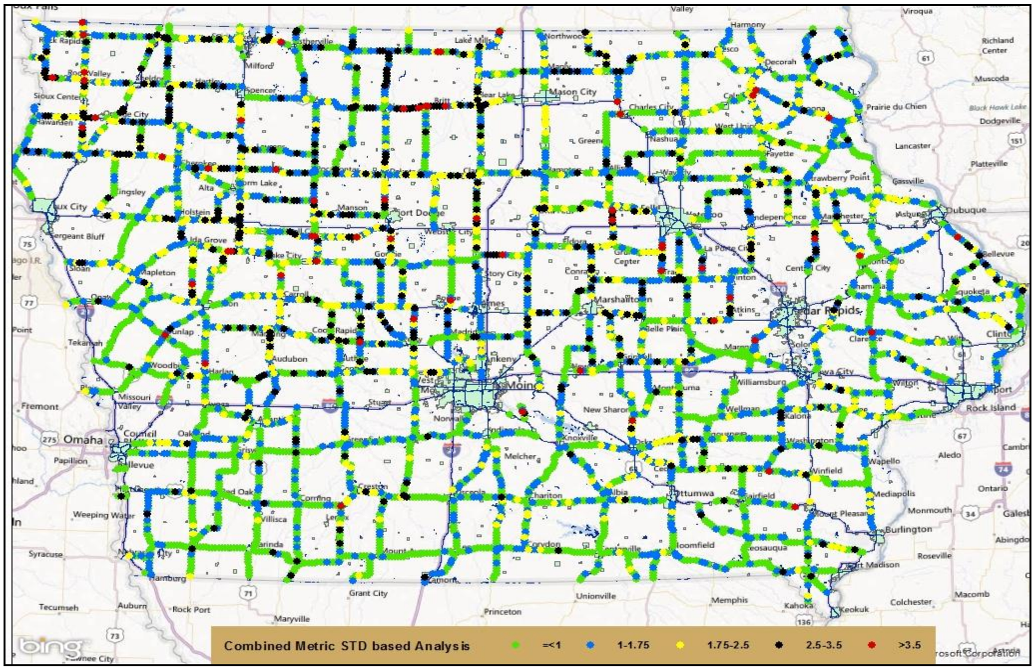


Figure 5.9 Standard deviation based analysis of combined metric for two-lane roadways

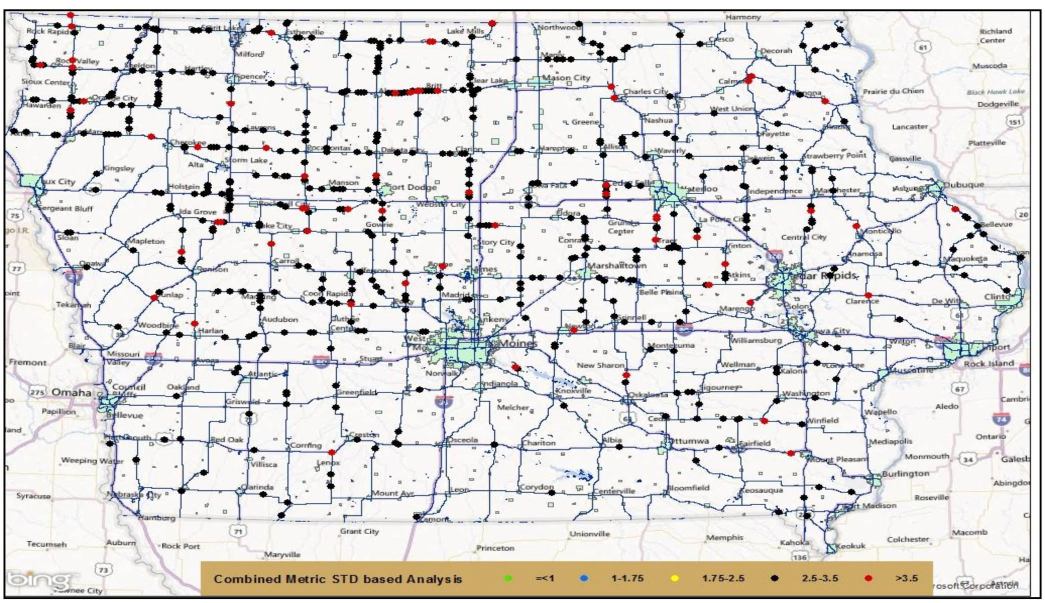


Figure 5.10 Crash hotspots for two-lane roadways (STD based analysis of combined metric)

5.1.3 Moving Average Analysis

To incorporate the spatial proximity, a moving average analysis was performed on all the Interstate/Freeway roads in Iowa to identify potential sites for safety improvements during winter weather. Three-mile road segments were considered for this purpose. The combined metric was used for each road segment to calculate the moving averages for 3-mile road segments. This moving average analysis was performed for I-29, I-35, I-80, I-680, I-380 and US-20 roadways. Figure 5.11 shows the road segments with moving average values of the combined metric and screened road segments having higher moving average values for considering those as high priority roadway section for improving winter weather safety. Results show two sections along I-35 north and three sections along I-80 east having a considerable number of road segments with black and red color.

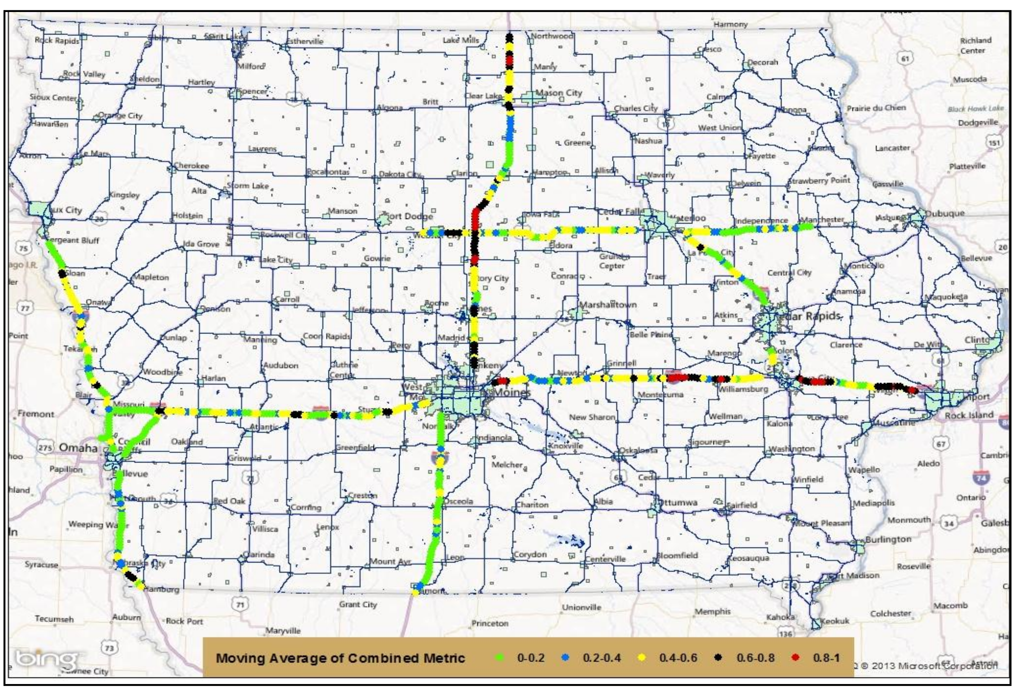


Figure 5.11 Moving average analysis of combined metric for interstates

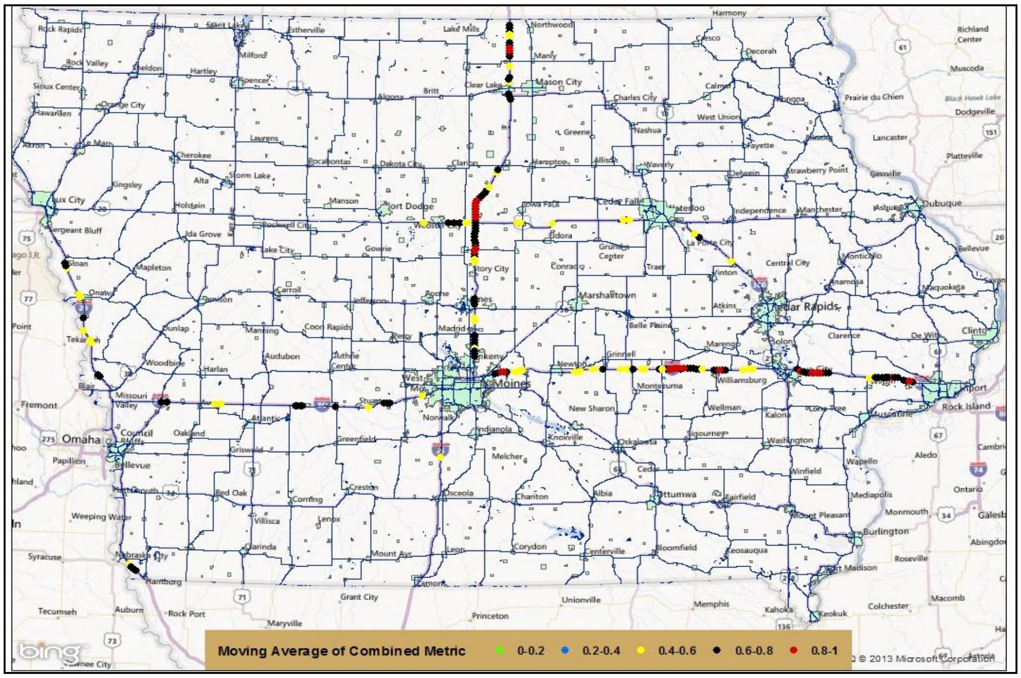


Figure 5.12 Crash hotspots based on moving average values of the combined metric

The ranking results of these methods (standard deviation based analysis of crash density, crash proportion, crash severity, combined metric, and moving average analysis) for 25 road segments are presented in the Table 5.3 for interstate/freeway roadway system.

A visual inspection of the Table 5.3 reveals that these rankings based on different methods identify different road segments. There are no segments identified by all the methods. Rankings based on moving average analysis and standard deviation based combined metric analysis have the most numbers of common road segments with eleven segments. Ranking based on standard deviation analysis of crash proportion has nine road segments in common with the combined metric based rankings, while rankings produced using crash density and crash severity based analysis have only two and one road segments respectively in common with the ranked segments based on combined metric. One reason for this discrepancy might be attributed to

Table 5.3 Results of Interstate/freeway network screening and ranking

	Crash density	Crash proportion	Crash severity	Combined metric	Moving Average
Ranking	Segment ID	Segment ID	Segment ID	Segment ID	Segment ID
1	4150	6792	6504	2676	4162
2	6779	3281	4119	2746	6713
3	9354	4151	6517	3281	6897
4	6664	6818	6665	3293	4150
5	6676	4122	6541	3342	4166
6	9347	4162	6731	4122	4167
7	6792	4163	6605	4150	4168
8	6820	4271	6714	4151	6664
9	3847	4130	6746	4162	6729
10	6884	6729	4157	4163	4151
11	3281	3252	6723	4164	4161
12	4215	3342	6581	4166	4163
13	6850	4164	6713	4168	4164
14	4213	4217	4045	4202	4259
15	4160	6731	4194	4204	4260
16	6693	4167	4128	4217	4271
17	4151	4273	6542	4260	4274
18	4068	4259	6761	4268	6665
19	6699	6639	3270	4271	6666
20	6713	3318	4150	4276	6714
21	4202	3331	6614	6618	6715
22	4198	4166	6572	6628	6717
23	6751	6688	4019	6664	6788
24	4197	2664	6538	6665	6792
25	4240	3321	9088	6668	6793

the non-consideration of traffic volume in developing these two metrics. Roads with high volumes typically experience more crashes and therefore, have a higher crash density. The severity metric also does not account for the changes in total injuries resulting from different traffic volumes. Rate based metrics considering traffic volume were not derived for this study as availability of traffic volume data during winter weather conditions was limited.

5.2 Development of Safety Performance Functions for Empirical-Bayes Analysis

This section discusses the development and results of the crash frequency models, SPFs to associate crashes with a variety of variables related to weather factors, traffic, and roadway related factors for three types of roadways. These SPFs are used to calculate the potential for safety improvement along road segments during winter seasons. The modeling techniques used for developing the SPFs are based on the methodology describe in the Chapter 4. Development of SPFs is based on the integrated dataset already described in the Chapter 3. After the integration of the crash data for the 2008 to 2012 winter periods, crash frequency was modeled as a function of geometric and traffic characteristics of different roadway classes, and weather related variables derived from crew reported weather information. Frequency models were developed for interstate/freeway, multilane divided/undivided, and two lane roadways. Overdispersion was present in the crash frequencies for interstate/freeway and multilane divided/undivided roadway segments. As such, a negative binomial modeling (with variance greater than mean of crashes) approach was taken to estimate the frequency of crashes on these two types of roadway segments. Although the value of overdispersion parameter was not very high, it was statistically significant for both classes of roadways. For two-lane roadway segments, a Poisson regression model was developed to estimate the probability of the number of crashes on this type of roadway as the variance and mean of crash frequencies along the road segments were same showing no overdispersion. The following section presents the model results for different types of roadways followed by a section describing the model results

5.2.1 Model Results

Tables 5.4 to 5.5 show the negative binomial regression estimation results for Interstate/freeway and multilane divided/undivided roadway segments. Table 5.6 shows the Poisson regression model estimation results along two-lane roadway segments. It is to be noted that the “snow amount” variable was not found to be significant in any of the SPFs developed for the three roadway types. Measures of model fit are represented by pseudo R^2 (1- residual deviance/null deviance) proposed by Cameron and Trivedi (1998) for negative binomial and Poisson models.

Table 5.4 Negative binomial model for Interstate/freeway road segments

Variables	**Estimates**	**Std. Errors**	**Pr**
Intercept	-4.72	0.48	<0.001
Log of AADT	0.65	0.038	<0.001
Air temperature (Fahrenheit)	-0.02	0.005	<0.001
Pavement temperature	0.017	0.005	<0.001
Surface Width (in feet)	0.02	0.003	<0.001
Visibility (in mile)	0.03	0.016	0.053
Posted Speed Limit (in mph)	-0.01	0.005	0.091
Null deviance	1915.33		
Residual Deviance	961.79		
Overdispersion factor	0.2343		
Pseudo R^2	0.5		

Table 5.5 Negative binomial model for multilane road segments

Variables	Estimates	Std. Errors	Pr
Intercept	-5.42	0.41	<0.001
Log of AADT	0.73	0.041	<0.001
Visibility (in mile)	0.04	0.015	0.004
Air temperature (Fahrenheit)	-0.006	0.003	0.0853
Posted Speed Limit (in mph)	-0.011	0.002	<0.001
Surface Width (in feet)	0.02	0.002	<0.001
Null deviance	1622.93		
Residual Deviance	797.63		
Overdispersion factor	0.16		
Pseudo R^2	0.51		

Table 5.6 Poisson model for two-lane road segments

Variables	Estimates	Std. Errors	Pr
Intercept	-3.04	0.22	<0.001
Log of AADT	0.37	0.03	<0.001
Wind Speed (in mph)	0.005	0.002	0.0156
Visibility (in mile)	0.023	0.009	0.0109
Surface Width (in feet)	0.0162	0.002	<0.001
Null deviance	1493.6		
Residual Deviance	1143		
Pseudo R^2	0.24		

5.2.2 Model Interpretation

The Poisson and negative-binomial models are of an exponential functional form, a measure of sensitivity of crash frequency to the corresponding variable can be attributed to the exponent in the model. Thus, elasticity estimates for the variables in the models were computed to measure the sensitivity of crash frequency to the corresponding variables. Elasticity is defined as the percentage change in the dependent variable resulting from a 1% change in an explanatory variable. Table 5.7 shows the elasticity values for the variables considered in the models for the three different types of roadways. The following specific observations can be made from the modeling outcomes for the three types of roadways. The results for the Interstate/freeway class and multilane divided/undivided classes are discussed together as the findings from the models were similar.

Table 5.7 Estimated elasticities

Variable	**Elasticties (for Interstate/freeway)**	**Elasticties (Multilane)**	**Elasticities (two-lane)**
Log of AADT	6.4	6.74	2.93
Air temperature	-0.43	-0.12	
Pavement temperature	0.38		
Road surface width	0.58	0.71	0.43
Visibility	0.06	0.13	0.07
Posted speed limit	0.66	0.53	
Wind speed			0.07

5.2.2.1 Interstate/Freeway and Multilane Divided/Undivided Roadways

Traffic volume

As expected, traffic volume, AADT for each specific road segment, was found to be significant with a positive sign, suggesting that an increase in traffic volume would result in an increase in the mean number of weather-related crashes expected to occur on the road segment during the winter season. The value of the coefficient associated with the traffic volume is 0.65, which is less than one and suggests the moderating effect of traffic volume is non-linear with a decreasing trend. A Similar value was found for road segments belonging to multilane divided/undivided roadways. Previous literature also reports similar effects of traffic volume on speed and weather-related crash frequency (Monsere et al., 2008). Traffic volume represented by AADT in the current study has a considerable impact on safety, as an increase in traffic volume (ranging from 90 to 113,600 during the winter seasons) by 1% would cause the mean number of crashes to increase by 6.4% on interstate/freeways. The elasticity value for volume reveals that a 1% increase in AADT (ranging from 50 to 34,225 during the winter seasons) will result in 6.74% increase in the mean crash frequency on multilane divided/undivided roadway segments.

Air temperature

Air temperature was found to be significant with a negative sign suggesting that the mean number of weather-related crashes will increase with the decrease in the air temperature. The elasticity value for the air temperature reveals that a 1% increase in air temperature during the winter season would decrease the mean number of crashes by 0.43% on interstate/freeway roads for air temperature ranging from -25 to 37 degree Fahrenheit. This result is in agreement with some of the previous findings (Fu et al. 2006). The elasticity value for multilane divided/undivided roadway indicates a 0.12% decrease in the mean number of crashes with a 1% increase in air temperature (from -18 to 38 degree Fahrenheit) during the winter season.

Pavement temperature

Modeling results reveal a significant relationship between average pavement temperature and mean number of crashes during the winter season on interstate/freeway road segments. The elasticity value for the pavement temperature shows that a 1% increase in pavement temperature would cause the mean number of weather-related crashes to increase by 0.38% when pavement temperature ranges from -9 to 38.8 degree Fahrenheit. Though the finding seem counterintuitive, it is possible that the increase in pavement temperature might result in different level of variation in road surface condition affecting crash frequency. Pavement temperature was not found statistically significant with crash frequency occurring on multilane divided/undivided roadway segments.

Road surface width (in feet)

Road surface width (measured from edge line to edge line) was found to be statistically significant and the positive sign indicates that roadway segments with wider surface were associated with a higher number of weather-related crashes during the winter seasons. Results revealed that a 1% increase in the roadway surface (from 16 to 90 feet) width would result in 0.58% increase in the mean number of crashes on interstate/freeway and 0.71% on multilane divided/undivided road segments (from 12 to 72 feet) during the winter seasons. On interstate/freeway roadways, wider roadways might make drivers feel safer and the number of crashes might increase due to the risk compensating behaviors of drivers during winter seasons. It might also suggest that drivers slow down during winter weather on narrower roadways. Drivers are also prone to changing lanes on a multilane roadway segment and this might have increased potential for greater number of crashes on road segments having larger surface width. Previous studies have also found similar results while developing crash frequency models for speed and winter-weather related crashes (Monsere et al. 2006).

Visibility (in mile)

Visibility was found to be significant with a positive relationship suggesting that the mean number of weather-related crashes during winter seasons will increase with better visibility. Although, this finding might seem counterintuitive, the increase in the frequency of crashes during winter seasons might be attributed to the risk compensating behaviors of the drivers due to increased visibility. Previous research findings showed a decrease in average vehicle speeds during winter weather with a decrease in visibility below 0.4 kilometer (Knapp and Smithson, 2001). The elasticity value for visibility indicates that a 1% increase in visibility (ranging from 0 to 5 mile) during winter seasons will increase the mean number of weater-related crashes expected along interstate/freeway road segments by 0.06%. The elasticity value for visibility (ranging from 0 to 5 mile) was found to be 0.13 for multilane divided/undivided roadway segments. Note that these results are similar with that of a past study Hermans et al. (2006a) using data from 37 sites. However, our results are different from the study conducted by Usman (2011) that found a negative relationship between visibility and crash frequency during a storm event. Note that the models developed in this research are not winter-storm event based models but rather consider all the weather-related crashes that occurred during the winter seasons. Large aggregation of data at the temporal level may have masked the effect of visibility in the current model.

Posted speed limit

Posted speed limit was found to be significant with a negative sign suggesting that the mean number of weather-related crashes would increase with a decrease in posted speed limits along roadway segment during winter weather seasons. This finding is in agreement with a previous study (Monsere et al., 2006). The elasticity value for the posted speed limit variable reveals that a 1% increase in posted speed limit will (35-70 mph for interstate and 20-65 mph for multilane roadways) result in a 0.66% and 0.53% decrease in the mean number of winter weather-related crashes on interstate/freeway road segments and multilane divided/undivided roadways segments, respectively. The greater variability of vehicle speeds during winter weather conditions compared to non-winter conditions (Knapp and Smithson, 2001) might contribute to this finding. Literature also shows evidence of decrease in the average winter weather speed compared to the typical average speed at the same location during non-winter weather conditions (Knapp and Smithson, 2001).

5.2.2.2 Two-lane roadways

Traffic volume

Traffic volume represented by AADT was found to have a statistically significant relation with the mean number of crashes. The sign of the value of the coefficient was positive suggesting an increase in the mean weather-related crash frequency with the increase in the traffic volume. The elasticity value for the AADT suggests that a 1% increase in the traffic volume will result in a 2.93% increase in mean crash frequency for AADT ranging from 50 to 52,700 on two-lane roadways.

Road surface width (in feet)

Road surface width was also found to have a significant effect with a positive sign on winter weather crash frequency for two-lane roadway segment. A similar effect was also found for weather-related crashes occurring on interstate/freeway and multilane divided/undivided roadway segment during the winter season. A 1% increase in road surface width (ranging from 14 to 76 feet) would result in a 0.43% increase in weather related mean crash frequency.

Visibility (in mile)

Visibility was found to have an effect on mean winter weather crash frequency for two-lane roadway segments similar to that of interstate/freeway and multilane divided/undivided roadway segments. Results reveal that a 1% increase in visibility (0-5 mile) would result in a 0.07% increase in the mean number of weather-related crashes on two-lane of roadway segments.

Wind speed

While wind speed was not found to be significant for the frequency models developed for interstate/freeway and multilane divided/undivided roadway segments, it was found to be statistically significant for two-lane roadways. The positive sign indicates that higher wind speeds were associated with a higher number of crashes. The elasticity value for wind speed (0-50 mph) shows that a 1% increase in wind speed would result in a 0.07% increase in mean number of weather-related crash frequency along two-lane roadway segments. The results make intuitive sense as higher wind speeds could cause blowing snow effects, which might impair driver performance during winter seasons. This result is in agreement with previous research findings (Knapp et al., 2000; Usman, 2011).

5.2.3 Ranking Results of Roadway Segments Using Empirical-Bayes

The results showing the ranking of the top 25 road segments based on the value of PSI are presented in Table 5.8 and 5.9 for Interstate/freeway and multilane divided/undivided roadways respectively. Sample ranking of roadway segments for these two classes of roadways are presented in Appendix C.

Table 5.8 and 5.9 show only two road segments in common for the ranking produced using the PSI and the previously discussed combined metric for intestate/freeway roadway (Table 5.3). The EB adjusted frequency and observed frequency produced similar ranking. This is expected as crash counts are included in the EB adjustment. Four years of crash counts were incorporated into the EB adjustment and as a result more weight was put on observed crash counts compared to the expected number of crash counts predicted from the SPFs. A similar result was also reported in the literature (Monsere et al., 2008). If less years of crash data would have been considered, more weight would be assigned to the predicted number of crash frequency from the SPFs. The tables show that the ranking based on the PSI is different from the ranking produced by EB adjusted crash frequency or the observed crash frequency. For example, segment 4186 of Interstate/freeway system had thirty eight observed crashes during the four years compared to twenty four crashes experienced by the segment 3342. But segment 3342 ranked ahead of segment 4186 when PSI was considered for the ranking purpose. The predicted crash frequency from the SPF was low for the segment 3342 compared to the observed number of crashes while the predicted crashes for the segment 4186 were close to the observed number of crashes. More weight was put on the predicted crashes for the segment 3342 to in comparison with the weight put on predicted crashes for the segment 4286.

Table 5.8 Top 25 roadway segments for potential for safety improvements (Interstate/freeway)

Interstate/freeway road segments					
Segment ID	Observed	Predicted	Weight	Adjusted	PSI
3393	50	8.65	0.03	48.93	40.28
9078	35	10.93	0.02	34.50	23.57
4187	48	26.56	0.01	47.82	21.26
9364	36	16.30	0.01	35.73	19.42
3342	24	3.21	0.07	22.61	19.40
3219	29	9.74	0.02	28.56	18.82
4198	26	8.93	0.03	25.57	16.65
6818	24	7.11	0.03	23.47	16.36
3394	30	13.98	0.02	29.74	15.76
4163	20	3.84	0.06	19.09	15.24
4154	21	5.11	0.04	20.31	15.21
6429	27	11.57	0.02	26.70	15.13
6714	22	6.45	0.03	21.46	15.02
4199	23	7.69	0.03	22.56	14.86
6430	30	15.12	0.01	29.78	14.66
4157	21	5.87	0.04	20.43	14.56
6792	22	7.13	0.03	21.53	14.41
6915	24	9.62	0.02	23.66	14.05
9161	50	36.19	0.01	49.91	13.73
9084	25	11.01	0.02	24.71	13.70
4150	20	5.98	0.04	19.48	13.50
6875	20	6.70	0.03	19.56	12.86
4178	44	31.26	0.01	43.91	12.65
3387	17	4.24	0.05	16.34	12.10
4186	38	26.81	0.01	37.90	11.10

Table 5.9 Top 25 roadway segments for potential for safety improvements (multilane divided/undivided)

Multilane divided/undivided roadway segments					
Segment ID	Observed	Predicted	Weight	Adjusted	PSI
1092	27	9.78	0.02	26.73	16.94
6039	26	11.76	0.01	25.81	14.05
6074	26	12.25	0.01	25.82	13.58
2633	27	13.68	0.01	26.85	13.17
1096	23	10.60	0.01	22.82	12.22
1091	24	11.76	0.01	23.84	12.08
6075	25	13.08	0.01	24.86	11.78
9077	26	14.47	0.01	25.88	11.41
9330	17	5.41	0.03	16.67	11.26
9481	19	7.55	0.02	18.77	11.21
2293	18	6.67	0.02	17.74	11.07
8209	21	9.76	0.02	20.82	11.06
1960	21	10.05	0.02	20.83	10.78
5267	17	6.83	0.02	16.77	9.94
9472	14	3.79	0.04	13.59	9.81
5659	16	5.98	0.03	15.74	9.77
6051	19	9.09	0.02	18.83	9.75
1053	19	9.70	0.02	18.85	9.15
3230	19	9.72	0.02	18.85	9.13
5455	13	3.72	0.04	12.62	8.91
9095	14	4.82	0.03	13.71	8.88
646	16	6.98	0.02	15.80	8.82
5316	16	7.01	0.02	15.80	8.79
6042	16	7.56	0.02	15.83	8.27
1089	16	7.73	0.02	15.83	8.11

5.3 Summary

Site prioritization techniques for identifying roadway segments with the potential for safety improvements related to winter weather crashes were developed through traditional naïve statistical methods by using raw crash data. Crash frequency models were developed using integrated crash data for four winter seasons (2008 to 2012) with the objective to identify factors affecting crash frequency during winter seasons and screen roadway segments using the Empirical Bayes (EB) technique. EB accounts for RTM phenomenon by overcoming the limitations caused by traditional methods. Weather factors such as visibility, wind speed, air temperature were found to have statistically significant effects on crash frequency along different types of roadways. The ranking of roadway segments for PSI also differed from the ranking produced by simple crash frequency which does not take into account the RTM. This type of ranking produced by employing Empirical-Bayes technique can be useful to identify roadway segments for consideration of potential safety improvement and allocate agency resources in an effective manner to mitigate winter weather-related crashes. SPFs developed in this research can be used to produce ranking based on PSI by using crash observations made over a specific number of years for winter weather crashes.

CHAPTER 6

MULTILEVEL MODELS FOR OCCUPANT INJURY SEVERITY ANALYSIS

Chapter 5 describes the identifying factors linked to winter weather crash frequency. This chapter describes the factors identified by models that predict occupant injury severity for weather-related crashes, all crashes (weather and non-weather-related), and non-weather related crashes during the winter seasons along the whole length of I-80 in Iowa. To account for the hierarchical nature of the crash data, models with a two-level specification (occupants nested within crashes) were developed to estimate the effects of several covariates on occupant injury risk.

6.1 Estimation Results

A binomial logistic distribution was used in a Bayesian framework for this modeling purpose. The response variable takes one of two values: 1 if the occupant was injured or 0 if the occupant was not injured. In the absence of strong prior information for the predictors, uninformative priors were assumed for all regression coefficients with normal distributions (0,1000) and the variance of the normally distributed random effect with inverse Gamma distribution (0.001. 0.001). The models were computed via the Metropolis Hastings sampler, a Markov Chain Monte Carlo (MCMC) technique which was implemented using MLwiN software (Rasbah et al., 2012). A ninety five percent Bayesian Credible Interval (BCI) was used to examine the significance of the covariates. Deviation Information Criteria (DIC) was used as the model diagnostic. Details about the modeling method were described in Chapter 4.

6.1.1 Modeling Results of Weather-Related Crashes

To develop this model, 2,493 weather-related crashes during the four winter seasons (2008-09 to 2011-12) were used with 3,717 vehicle occupants. This resulted in an average involvement rate of 1.49 individuals per crash. Table 6.1 shows the estimation results of the model predicting occupant injury severity for weather-related crashes.

6.1.1.1 Gender of the Occupant

The estimation results showed that the likelihood of sustaining an injury by a male occupant is higher compared to female occupants. Male occupants have 63% lower odds to be injured than female occupants when involved in crashes. Several past studies (Dissanayake, 2004; Duncan et al., 1998; Jung et al., 2010; Khattak et al., 1998) suggested that the level of injury severity in a crash was lower for male drivers compared to female drivers.

6.1.1.2 Seating Position of the Occupant

The model results revealed a decrease in the odds of an occupant being injured if he/she is in the driver seating position. The driver has 99% less likelihood of being injured in coparison to the other occupants involved in a crash. This might be a reflection of the improved protection offered to car occupant in the case of a frontal impact compared to side or rare impacts. Previous research (Lenguerrand et al., 2006) shows increased likelihood of occupant injury with crash impact from the side compared to front.

Table 6.1 Posterior summaries of parameter estimates for model with weather-related crashes

Parameters	Effect Estimate		Odds ratio	95% Bayesian Credible Interval (BCI)	
	Mean	SD		2.50%	97.50%
Gender (if male = 1, otherwise = 0)	-1.017	0.154	0.362	-1.322	-0.725
Seating position (if driver = 1, otherwise =0)	-5.306	0.417	0.005	-6.235	-4.551
Occupant Protection					
None used (reference)	0	0	1.000	0	0
Used	-1.674	0.486	0.187	-2.661	-0.78
Unknown/not reported	-1.745	0.537	0.175	-2.861	-0.725
Airbag deployment					
Airbag deployed (reference)	0	0	1.000	0	0
Not deployed	-2.117	0.011	0.120	-2.565	-1.726
Unknown/Not reported	-2.234	0.285	0.107	-2.825	-1.693
First harmful event					
Non-collision (overturn, rollover, jackknife) (reference)	0	0	1.000	0	0
Collision with vehicles	-1.598	0.197	0.202	-1.998	-1.231
Collision with non-vehicles (animal, debris, work zone equipment, etc.)	-1.432	0.214	0.239	-1.869	-1.041
Roadway condition as contributing circumstance (if yes = 1, otherwise = 0)	-0.387	0.179	0.679	-0.732	-0.037
Trapped (if an occupant is trapped = 1, otherwise = 0)	1.143	0.35	3.136	0.467	1.829
Ejection status (if an occupant is ejected = 1, otherwise = 1)	2.105	1.159	8.207	-0.096	4.412
Age of the occupant (if age of the occupant is greater than 24 years old = 1, otherwise = 0)	0.373	0.179	1.452	0.028	0.73
Road type (if intersection = 1, otherwise = 0)	-0.682	0.263	0.506	-1.2	-0.181
Road surface condition and air temperature (if surface icy and temperature below zero = 1, otherwise = 0)	0.366	0.157	1.442	0.066	0.679
Random effects					
Between crash variance	3.708	0.851		2.322	5.732
Within-crash variance	3.29				
Intra- Class correlation	0.53				

6.1.1.3 Occupant Protection

The types of occupant protection used by occupants at the time of a crash occurrence included shoulder and lap belt, helmet, and child safety seat. The effect of occupant protection was investigated by means of three variables. These variables were available for each occupant involved in the crashes and were coded to include an "unknown/not reported" value for the cases without any available information. With respect to no occupant protection used, it was found that occupants using some type of protection have 81% less likelihood of being injured. This finding confirms previous results from several studies (Khattak et al., 2002; Dissanayake, 2004; Hermans et al., 2006b) reported in the literature. One interesting finding was that the occupants are less likely to be injured when no information about the type of protection used by them is available. These occupants are 83% less likely to be injured compared to occupants using no protection. This could be due to the fact that these occupants may have used some type of occupant protection, even though no such information was reported in the crash database.

6.1.1.4 Airbag Deployment

The variable indicating the deployment of airbag was classified as a categorical variable in this study with three categories: airbag deployed (reference category), not deployed, and unknown/not reported. It was found that occupants have 87% lower likelihood of being injured when airbags were not deployed during the crashes compared to the instances when airbag was deployed. It might be possible that deployment of airbag increases the likelihood of injuries because of the impact sustained by the occupants but reduces the likelihood of occupants being fatally or seriously injured. A similar result was obtained for occupants involved in weather-related crashes with no information available related to the deployment of airbag in the crash database. It might be possible that airbag was deployed in those crashes involving the occupants yielding similar results.

6.1.1.5 First Harmful Event

This variable indicates the type of first harmful event in relation to the chain of crash events. The “first harmful event” variable was categorized in three categories: non-collision (overturn, rollover, jack knife), collision with vehicles, and collision with non-vehicle objects (e.g. fixed object, animal, debris, work zone equipment). It was found that occupant injury risk decreases by 80% and 77% respectively for crashes with first harmful event of “collision with vehicles” and “collision with non-vehicle objects” compared to crashes with first harmful event of “non-collision” events. As the first harmful event of “non-collision” is not an actual collision but a loss of vehicle control or other inappropriate maneuver recorded, it is possible that following events resulted in more harmful events (including one or more collisions) with higher probability of sustaining injury by occupants involved in the crashes

6.1.1.6 Roadway Surface Condition as the Contributing Circumstance

The variable indicating whether roadway surface condition was a contributing circumstance for weather-related crashes was found to be a significant factor affecting the likelihood of occupant injury risk. The odds ratio value indicates a 33% decrease in the risk of occupant injury for crashes with the roadway surface condition reported as the contributing circumstance. This might likely be an effect of reduced speed levels of the vehicles on poor road surface conditions in adverse weather conditions during the winter seasons. However, it is difficult to represent the roadway condition at the time of the crash by taking into account the pavement friction. Findings reported in past studies (Cheng and Mannering, 1999; Khattak et al., 1998; Quddus et al., 2002; Quddus et al., 2010) indicated that poor road surface conditions were associated with reduced level of injury severity.

6.1.1.7 Trap Status

It was found that trapped occupants are more likely to be injured compared to those who were not trapped as a result of a crash. Results indicated that trapped occupants have 213% increased odds of being injured when involved in crashes. Trapped occupants in vehicles might remain in critical position before they are rescued. If the rescue efforts take time, the likelihood of an occupant being injured might increase. On the other hand, occupants able to get out of the vehicle after the crash occurrence can be attended quickly reducing the likelihood of injury or severe injury.

6.1.1.8 Ejection Status

The indicator variable for ejection status was found to significantly affect the probability of injury for the occupants. Ejected occupants are eight times more likely to be injured compared to those not ejected. This finding is intuitive as occupants are more likely to be injured once they are ejected from the vehicles and experience an impact. No findings related to ejection status and occupant injury was found in previous studies reported in the literature.

6.1.1.9 Age of the Occupant

The demographic variable indicating the age of the occupant was found to significantly affect the occupant injury risk. Occupants older than 24 years were identified to be more likely to be injured. Results revealed a 45% increase in the odds of injury for occupants older than 24 years. Aged occupants included in this group might have relatively weak risk detecting and reacting abilities along with deteriorating muscle strength and visual power. Thus, they are more susceptible to sustain injuries compared to younger occupants when involved in a crash.

6.1.1.10 Type of Roadway Junction

The indicator variable for type of roadway junction revealed that the probability of sustaining injury by occupants is low when crashes occurred at intersections in comparison to crashes occurred at non-intersections locations. This might be attributed to the caution exercised by the drivers when reached at intersections. This finding is in agreement with a previous study (Lenguerrand et al., 2006) considering the hierarchical nature of the crash data.

6.1.1.11 Road Surface Condition and Air Temperature

This indicator variable was used to represent whether the road surface was icy and the temperature was below zero degrees Fahrenheit during the crashes. Results indicated a 44% increase in the odds of occupant injury when both conditions related to road surface and air temperature existed. Icy road surfaces with extremely low temperature might result in poor road surface condition with reduced pavement friction. This might lead to difficulty for the drivers to maneuver safely and could result in crashes. Similar findings were reported in past studies (Donnell and Mason, 2004; Deng et al., 2006; Mergia, 2010).

6.1.2 Modeling Results of Non-Weather Related Crashes

Table 6.2 shows the estimation results for the model considering non-weather crashes only. 2,749 non-weather related crashes occurring along Interstate 80 during the four winter seasons spanning from 2008-09 to 2011-12 were considered. There were 4,112 occupants involved in these crashes resulting in an average involvement of 1.5 individuals per crash.

6.1.2.1 Gender of the Occupant

This indicator variable was found to an effect similar to the model results for weather-related crashes. Male occupants have 50% less odds of being injured compared to the female occupants when involved in a crash not related to weather during winter seasons.

6.1.2.2 Seating Position of the Occupant

It was found that occupants seated in the driver position are 99% less likely to be injured compared to occupants seated in other positions. This result is similar to those of the models predicting occupant injury risk for weather-related crashes and all crashes. Both models showed 99% lower probability for driver occupants to be injured compared to occupants in other seating positions when involved in a crash.

6.1.2.3 Occupant Protection

Occupants using some type of protection were found to have 38% less odds of being injured when compared to occupants not using any protection. This finding makes intuitive sense and is also in agreement with findings from previous studies reported in the literature (Khattak et al., 2002; Dissanayake, 2004; Hermans et al., 2006b).

6.1.2.4 Airbag Deployment

Deployment of airbag affects occupant injury risk significantly in this model also. Occupants have 243% increased risk of being injured when an airbag was deployed in comparison to occupants not experiencing airbag deployment during crash occurrences. As already discussed in the models results for weather-related crashes, airbags might have prevented the occupants from being fatally or seriously injured but might have increased the probability of occupants being injured because of the airbag impact.

6.1.2.5 First Harmful Event

This categorical variable was found to have identical effect on occupant injury risk when compared to the effect for all crashes during the four winter seasons. Occupants have 82% less odds of being injured in crashes with a first harmful event of “collision with vehicles” compared to crashes with a first harmful event of “non-collision”. It was also found that occupants have 60% less odds of sustaining injury when involved in crashes with a first harmful event of “collision with non-vehicle object”. This reduction in likelihood odds for crashes having the first harmful event as “collision with non-vehicle objects” might be due to the fact that fixed objects (such as tree, poles, guardrail, sign post) are less likely to absorb energy when experiencing a collision and result in increased occupant injury risk.

Table 6.2 Posterior summaries of parameter estimates for model with non-weather crashes

Parameters	Effect Estimate		Odds ratio	95% Bayesian Credible Interval (BCI)	
	Mean	SD		2.50%	97.50%
Gender (if male = 1, otherwise = 0)	-0.702	0.14	0.496	-0.984	-0.439
Seating position (if driver = 1, otherwise =0)	-5.575	0.461	0.004	-6.614	-4.819
Occupant Protection					
Occupant protection not used (reference)	0	0	1.000	1	1
Occupant protection used	-0.476	0.213	0.621	-0.901	-0.069
Airbag deployment					
Airbag not deployed (reference)	0	0	1.000	1	1
Airbag deployed	1.235	0.157	3.438	0.936	1.547
First harmful event					
Non-collision (overturn, rollover, jackknife) (reference)	0	0	1.000	1	1
Collision with vehicles	-1.695	0.232	0.184	-2.187	-1.267
Collision with non-vehicles (animal, debris, work zone equipment, etc.)	-0.915	0.244	0.401	-1.418	-0.469
Surface Condition					
If surface has water (moving or standing) = 1, otherwise = 0	-2.213	0.479	0.109	-3.209	-1.309
Trap Status					
Occupant not trapped (reference)	0	0	1.000	1	1
Occupant trapped	4.583	0.481	97.807	3.706	5.567
Unknown/not reported	-3.088	0.925	0.046	-4.992	-1.355
Ejection status					
Occupant not ejected (reference)	0	0	1.000	1	1
Occupant ejected	4.869	1.738	130.191	1.923	8.902
Unknown/not reported	1.101	0.862	3.007	-0.505	2.91
Age of the occupant					
Occupants aged up to 24 years (reference)	0	0	1.000	1	1
Occupants aged 24 years or higher	0.371	0.17	1.449	0.051	0.714
Weather condition (if rain, mist, snow, fog, wind = 1, otherwise = 0)	-0.488	0.218	0.614	-0.923	-0.067
Major cause (if run-off-road = 1, otherwise = 0)	0.858	0.206	2.358	0.454	1.27
Random effects					
Between crash variance	2.951	0.656		1.926	4.431
Within-crash variance	3.29				
Intra- Class correlation	0.47				

6.1.2.6 Trap Status

Trap status of the occupants was defined as a categorical variable in this model with three categories: "not trapped", "trapped", and "unknown/not reported". Compared to occupants not trapped, trapped occupants are almost hundred times more likely to be injured when involved in a non-weather-related crash. On the other hand, occupants with unavailable information on their trap status have 96% less odds of being injured with respect to the reference category.

6.1.2.7 Ejection Status

Ejection status was defined as a categorical variable for this model with the categories: "not ejected", "ejected", and "unknown/not reported". It was found that ejected occupants are hundred and thirty times more likely to be injured compared to the occupants not ejected. With respect to the occupants not ejected, occupants having "unknown/not reported" status related to their ejection status are three times more likely to be injured. Occupants falling into "unknown/not reported" category might have been ejected and sustained injury because of the ejection. However, "unknown/not reported" category was not statistically significant in the current model

6.1.2.8 Age of the Occupant

This variable was defined as categorical in this model. The two categories are "occupants aged up to 24 years" and "occupants aged 24 years and higher". Considering "occupants up to 24 years" as the reference category, it was found that occupants aged 24 years or higher have 45% increased risk of being injured when involved in a non-weather related crash.

This finding is similar to the results of the weather-related crash model except a reference category was used in the current model.

6.1.2.9 Surface Condition

Surface condition affects the risk of injury severity significantly in the current model. Occupants involved in crashes occurred on a surface with water (standing or moving) have 89% less odds to be injured. This might be attributed to the capability of porous asphalt to drain water away quickly. Risk of hydroplaning reduces with the drainage of water and visibility of road marking might become better reducing the risk of an injury. Also, drivers might be more cautious while driving on wet road surface reducing the risk of injury.

6.1.2.10 Weather Condition

This variable indicates bad weather conditions (fog/smoke, mist, rain, sleet, hail, freezing rain, snow, severe winds, blowing snow) during the crash. Bad weather conditions result in 39% less odds of occupants being injured. It might be possible that driver experience enable the involved drivers to adjust to the type of bad weather conditions occurring frequently during winter seasons in Iowa and lower the probability of injury risk for the occupants. It is reasonable to speculate that drivers adapt to adverse weather conditions by adjusting their speeds, and driving behaviors that could result in a lower probability of a crash.

6.1.2.11 Major Cause: Run-Off-Road

This indicator variable was found significant in this model unlike the models considering weather-related crashes during the winter seasons. Occupants involved in crashes with the major cause being “run-off-road” have increased odds of being injured by 135%. Run-off-road crashes might follow events (such as collision with oncoming vehicles, collision with fixed objects) that are responsible for the occupants being injured.

6.1.3 Modeling Results of All (Weather and non-weather related) Crashes

Table 6.3 shows the estimation results for the model considering all the crashes (weather and non-weather-related) that occurred during the winter seasons from 2008/09 to 2011/12. 5,242 crashes (both weather and non-weather) were considered in the development of this model. 7,829 occupants resulted in an average involvement rate of 1.5 per crash.

6.1.3.1 Gender of the Occupant

According to the estimation results, male occupants are 49% less likely to be injured compared to the female occupants involved in crashes occurred during the seasons. A previous study (Morgan and Mannering, 2011) confirmed that drivers’ adaptation to weather-induced changes is a complex process that might potentially be influenced by gender.

6.1.3.2 Seating Position of the Occupant

The model results revealed that there is 99% less odds of being injured when an occupant is in the driver position. This finding is similar to that of the model results that consider weather-related and non-weather related crashes. Considering all crashes (weather- and non-weather-related) did not impact type of effect this indicator variable had on the likelihood of occupant injury.

6.1.3.3 Visibility

The visibility variable was categorized as “1 to 3 miles”, “3 to 6 miles”, and “6 miles and above” when the model was developed for all crashes during the winter seasons. It was found that visibility affects the risk of occupant injury significantly. Compared with visibility of “1 to 3 miles”, visibility ranging from “3 to 6 miles” and “6 miles and beyond” increase the odds of occupant injury by 60% and 36% respectively. The reason for increased likelihood of occupant injury with increased visibility might be attributed to the increase in vehicle speed with improved visibility. Previous research (Knapp and Smithson, 2001) showed that vehicle speed decreased with decrease in visibility below 0.4 km. Risk compensating behavior of drivers with increased visibility might also be responsible for increased possibility of crash occurrence and injury for occupants. It is to be noted that this variable was not found to be significant in the model that considered weather-related crashes only. Incorporating non-weather related crashes in the current model might have induced additional variability in visibility with respect to occupant injury.

6.1.3.4 Occupant Protection

This variable was classified in a manner similar to the model that considered weather-related crashes only. With respect to the occupants not using any occupant protections, it was found that occupants using some type of protection are at 82% less odds of being injured. Occupants have 79% less likelihood of being injured when no information is available for the type of occupant protection used by them in the crash database. The findings are similar to that of the model results considering weather-related crashes only. No information is available on the use of occupant protection by some occupants. It might be possible that the use of some of type of protection by these occupants might have decreased the likelihood of them being injured compared to those who used no occupant protection.

6.1.3.5 Airbag Deployment

According to the model results, occupants have 85% lower odds of being injured when an airbag was not deployed in their vehicles compared to crash occurrences when it was deployed. Occupants in the category “unknown/not reported” have 82% lower odds of being injured with respect to occupants experiencing airbag deployment. These findings are also similar to those reported from the model results that considered weather-related crashes only.

6.1.3.6 First Harmful Event

The variable indicating first harmful event is a significant factor affecting occupant injury in this model. This variable was classified according to the same categorization used in the model for weather-related crashes only. Results revealed an 80% reduction in the likelihood of occupant injury for crashes with first harmful event of “collision with vehicles” compared to crashes with first harmful event of “non-collision”. However, it was found that risk of occupant injury was reduced by 62% for crashes with a first harmful event of “collision with non-vehicle objects” compared to crashes with a first harmful event of “non-collision”. Non-vehicle objects include fixed objects, debris, animals, and work zone equipment. The same effect of this variable was discussed in the model results for non-weather-related crashes.

Table 6.3 Posterior summaries of parameter estimates for model with all (weather-related and non-weather) crashes

Parameters	Effect Estimate		Odds ratio	95% Bayesian Credible Interval (BCI)	
	Mean	SD		2.50%	97.50%
Gender (if male = 1, otherwise = 0)	-0.661	0.096	0.516	0.096	-0.654
Seating position (if driver = 1, otherwise =0)	-4.886	0.052	0.008	-5.387	-4.429
Visibility					
Visibility 1 to 3 mile (reference)	0	0	1.000	0	0
Visibility (3 to 6 mile)	0.468	0.172	1.597	0.132	0.796
Visibility (6 mile and above)	0.306	0.132	1.358	0.051	0.556
Occupant Protection					
None used (reference)	0	0	1.000	0	0
Used	-1.708	0.298	0.181	-2.344	-1.203
Unknown/not reported	-1.562	0.321	0.210	-2.218	-0.92
Airbag deployment					
Airbag deployed (reference)	0	0	1.000	0	0
Not deployed	-1.903	0.143	0.149	-2.207	-1.634
Unknown/not reported	-1.758	0.198	0.172	-2.147	-1.394
First harmful event					
Non-collision (overturn, rollover, jackknife) (reference)	0	0	1.000	0	0
Collision with vehicles	-1.619	0.012	0.198	-1.905	-1.361
Collision with non-vehicles (animal, debris, work zone equipment, etc.)	-0.976	0.008	0.377	-1.263	-0.707
Surface Condition					
Surface condition dry (reference)	0	0	1.000	0	0
Surface condition icy, wet, snow or slush	-0.266	0.101	0.766	-0.46	-0.067
Surface condition (others and not reported)	-2.27	0.358	0.103	-3.03	-1.646
Trap Status					
Occupant not trapped (reference)	0	0	1.000	0	0
Occupant trapped	3.336	0.275	28.106	2.84	3.907
Unknown/not reported	-0.808	0.332	0.446	-1.464	-0.155
Ejection status (if an occupant is ejected = 1, otherwise = 1)	1.755	0.667	5.783	0.441	3.037
Air temperature (if below zero = 1, otherwise = 0)	-0.186	0.095	0.830	-0.369	-0.002
Random effects					
Between crash variance	2.391	0.457		1.706	3.349
Within-crash variance	3.29				
Intra- Class correlation	0.42				

6.1.3.7 Surface Condition

Surface condition was classified as a categorical variable with three categories: “dry”, “icy, wet, snow or slush”, and “others and not reported”. The category “others and not reported” includes road surface with sand, mud, dirt, oil, gravel, and water. With respect to the reference category of dry surface, it was found that risk of occupant injury reduced by 23% for crashes occurring on icy, wet, snowy or slushy road conditions. This might be attributed to the reduced speed of the vehicles on road surface with these conditions. As all crashes are considered in this model, effect of icy or wet road conditions on occupant injury risk might have been superseded by the effect of dry road condition on occupant injury risk. Compared to “dry” road surface condition, occupants have 90% less likelihood to be injured when involved in crashes occurring on road surface with “others and not reported” condition. It is possible that drivers use extra caution while driving on this type of road condition and reduce the probability of injury. Again, reduced speed of the vehicles might also have played a role in reducing the occupant injury risk for crashes on this type of road surface condition.

6.1.3.8 Trap Status

Trap status of the occupants was defined as a categorical variable in this analysis with three categories: “not trapped”, “trapped”, and “unknown/not reported”. Results from the model showed that trapped occupants have twenty eight times higher odds of being injured compared to occupants not trapped. Similar results were found for the model that considered non- weather-related crashes but with smaller effects. With respect to non-trapped occupants, occupants with unavailable information (unknown/not reported) on their trap status have 55% lower odds of being injured. Occupants belonging to the "unknown/not reported" category might not be trapped reducing the likelihood of being injured.

6.1.3.9 Ejection Status

Ejected occupants are almost eight times more likely to be injured compared to those not ejected involved in crashes during the winter seasons. This finding is similar to that of the model that considered weather-related crashes only. The higher injury risk for the ejected occupants might be attributed to the impact experienced by the occupants due to the ejection from vehicles.

6.1.3.10 Air Temperature

Air temperature was found to be a significant factor affecting the risk of occupant injury in the model that considered all crashes. An indicator variable was used to represent whether air temperature was below zero degree Fahrenheit during the crash. Occupants involved in crashes with air temperature below zero degrees Fahrenheit were found to have 17% less odds of being injured. Information about the physical effect of air temperature on injury risk is sparse in the literature. However, according to a German study (DVR, 2000), emotions rise with temperature and as a result, people get tired and lose concentration with an increase in reaction time. This might be attributed to the decreased injury risk with decrease in air temperature.

6.2 Measure of Between-Crash Variance and Model Diagnostic

As shown in Tables 6.1 to 6.3, the variance of random effect indicating the magnitude of the between-crash variance or within-crash correlation for the three models are 3.708, 2.391, and 2.95, respectively. The Intra-Class Correlation (ICC) is calculated according to Equation 4.15.

Results indicate ICC values of 0.53, 0.42, and 0.47 for the models considering weather-related crashes, all crashes, and non-weather-related crashes. These values mean that 53%, 42%, and 47% of unexplained variations in occupant injury risk resulted from between-crash variance suggesting the usefulness of this model that considers the hierarchical structure of the crash data.

A diagnostic known as Deviance Information Criteria (DIC) (described in Chapter 4) is used to assess the models with a DIC. A small DIC value indicates for a better model. The comparison results of different models on the basis of DIC values are presented in the following table. Ordinary logistic models were developed without considering the hierarchical structure of the crash data. The model diagnostics show that the models considering hierarchical crash data have higher DIC value even after penalizing the mean deviance (Dbar) by the effective number of parameters (pD). This further strengthens the notion of considering the hierarchical structure of crash data to develop crash inury severity prediction models.

Table 6.4 Model diagnostic for the models

	Dbar	D(thetabar)	pD	DIC
Model with weather-related crashes				
Ordinary logistic model	2831.06	2816.05	15.01	2846.07
Hierarchical logistic model	1918.3	1265.5	652.81	2571.1
Model with non-weather-related crashes				
Ordinary logistic model	2033.3	1502.69	530.63	2563.9
Hierarchical logistic model	2765.14	2750.19	14.95	2780.09
Model with all crashes				
Ordinary logistic model	5626.25	5611.17	15.08	5641.32
Hierarchical logistic model	4231.9	3203.36	1028.6	5260.5

6.3 Summary

Three models of occupant injury crash severity with binomial logit formulation were developed considering the hierarchical structure of the crash data in a Bayesian framework. These models were developed using disaggregate crash data with occupants nested within crashes. The vehicle level was ignored as majority of the vehicles in the sample included only a single occupant. This did not allow a differentiation between vehicle level and occupant level. Based on the modeling results, it was found that the model developed with weather-related crashes had the highest ICC value with crash-level variance accounting for 53% of the occupant injury risk. It was found that factors related to occupants (gender, seating position, trap status, ejection status, airbag deployment, safety equipment used) had statistically significant effects on occupant injury risk for all the models. Weather-related variables such as visibility and air temperature were found significant predictors of all crashes occurring during the winter seasons.

CHAPTER 7

CONCLUSIONS, LIMITATIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH

The primary objectives of this dissertation were to develop a systematic prioritization technique for screening road segments while incorporating winter weather related factors, and gain insights on the factors affecting injury crash severity during winter seasons. This chapter summarizes the major contributions and conclusions of this study followed by the limitations of this study and recommendations for future research.

7.1 Major Contributions and Conclusions

7.1.1 Data Processing

Integrating weather data with crash data is a major concern for conducting winter weather road safety research. This research effort has resulted in a comprehensive crash database covering four winter seasons (2008-09 to 2011-12) in Iowa for three types of roadways as Interstate/freeway, multilane divided/undivided, and two-lane roadways. The compiled dataset contains variables representing roadway and traffic information, and most importantly weather related information. The novelty of this research lies in integrating the weather related information reported by Iowa DOT maintenance and operational crew members during winter seasons with the crass data. It would be a cumbersome effort to collect weather data from different sources (RWIS and AWOS) for all the crashes considered during the four winter periods representing weather conditions during the time of crash occurrences. Instead, this research devoted effort to collect weather information for each crash from the nearest maintenance garage where maintenance crew members report important weather related information (such as air temperature, pavement temperature, wind speed, visibility, and precipitation type) while performing maintenance activities. This research is the first to create a process for extracting weather data from crew reported information. A comprehensive code in the C programming platform was developed to integrate the crew reported weather data with the crash information based on time and date of crash occurrences. This code can be used for extracting weather data in an efficient and comparably quicker way than extracting weather information from RWIS and AWOS stations for winter crash data analysis at network level. Extracting weather data from RWIS and AWOS stations at network level is a time consuming and resource intensive process at network level. Agencies can modify this code to integrate weather data and crash data for conducting winter crash data analysis at network level. The resulting database in this research includes 13,859 winter weather-related crashes with assigned roadway geometry, traffic volume, and weather information for the three types of roadways, which represents a rich dataset for developing crash frequency models.

Furthermore, as part of this research effort, a comprehensive database for all the crashes that occurred on the full stretch of Interstate 80 covering the state of Iowa spanning the winter seasons from 2008-09 to 2011-12 was developed. As both weather-related and non-weather related crashes during the winter seasons were considered for developing crash-injury severity models, weather data were collected and integrated from RWIS and AWOS stations nearest to the crash sites. An efficient code was developed in C programming platform to integrate weather data with crash data based on the time and date of crash occurrence by considering nearest RWIS and AWOS stations assigned to each crash. The resulting database contains detailed crash level, vehicle level, and occupant level information along with weather information for 5,242 weather and non-weather related crashes.

7.1.2 Development of a Systematic Prioritization Technique for Winter Crashes

As there is no systematic method to identify potentially problematic winter weather-related crash locations in Iowa, initially this research effort identified candidate locations for safety improvements during winter weather conditions on the basis of traditional metrics such as crash proportion, crash density, crash severity and a combined metric considering these three metrics. Using historic raw crash data from 2002 to 2009 winter seasons, combined metric analysis, standard deviation based analysis, and moving average analysis were employed to identify and analyze sites for winter-safety analysis. The rankings of the road segments produced from these different analyses are useful for identifying high crash locations during winter weather conditions. However, these naïve statistical methods suffer from serious limitations including the regression-to-mean problem.

To overcome this problem, this study employed a systematic prioritization technique in Empirical-Bayes framework for screening road segments with high potential for safety improvements during the winter seasons. The crash frequency models or PSFs developed for Interstate/freeway and multilane divided/undivided roadways accounted for the overdispersion of the crash data after incorporating the weather related factors in the models. Typical SPFs developed to estimate crash frequency use roadway geometry characteristics such as lane width, surface width and traffic volume expressed as average annual daily traffic (AADT). These SPFs normally do not consider weather variables as weather data collection and incorporation of weather information in SPF are complex processes. This research is the first to develop safety performance functions for winter weather-related crashes incorporating weather information extracted from crew reported data at network level. However, the distribution of crash frequency for the two-lane roadways did not exhibit overdispersion as most of the two-lane road segments experience a single crash during the study period. Quite a few weather related factors as air temperature, pavement temperature, wind speed, and velocity were statistically significant in the developed models. Goodness-of-fit of the developed models showed promising results in terms of explaining the variance in winter weather crash frequency after incorporating the weather related factors. Models for Interstate/freeway and multilane divide/undivided road facilities performed well with 50% and 51% of the variance in crash frequency explained by these two models respectively. The ranking of road segments based on Potential for Safety Improvement (PSI) derived from Empirical-Bayes method using the developed SPFs differed from the ranking based on crash frequency which ignores the regression-to-mean phenomenon. The crash frequency based models developed in this research can be used to predict the occurrence of winter weather crashes on road segments based on information from crew reported winter weather information and characteristics of roadway segments of different roadway types to evaluate the crash risk. Iowa DOT can use the developed SPFs to apply the Empirical Bayes technique for screening road networks for improving winter weather safety. Greater weight can be put on the predicted crash frequency from the developed SPFs along similar roadway segments if crash observations are made over a small period and vice versa when observations are made over a considerable number of years. However, the safety performance functions developed in this research should be applied to different types of roadway networks in different regions during winter seasons to test the robustness and reliability of the developed SPFs for using outside Iowa. Availability of detail geometric features of the roadway segments and weather information during winter seasons in different regions should improve the transferability of the models substantially. State transportation agencies can also screen the road networks to identify potential high risk sites or locations for safety improvement and introduce countermeasures in an effective way. The developed crash frequency models or the SPFs can be used to make data driven decision to prioritize road segments for improving winter weather safety in Iowa.

7.1.3 Development of Multilevel Models for Crash-injury Severity during Winter Seasons

This dissertation developed three statistical models to predict occupant crash injury risk for weather, non-weather, and all crashes during the four winter seasons (2008/09-2011/12) considering the hierarchical nature of the crash data. The comparison of the model estimation results helps to identify the differences in the injury risk of occupants involved in weather-related crashes versus all crashes. Binary logit models in a Bayesian framework were developed to predict the injury risk of occupants relating to several covariates at the occupant and the crash level. These models were developed using crash data at disaggregate levels with occupants nested within crashes. Most previous studies on crash-injury severity were conducted at the crash level ignoring the potential correlation of the severity levels for the vehicles involved in the same crashes or individuals involved in the same vehicles. This research applied a multilevel approach to take into account the hierarchical nature of the crash data by nesting the occupants within crashes. Results from the developed models revealed significant within-crash correlation in the dataset analyzed for the study. More than 40% of unexplained variation in individual occupant injury risk resulted from between-crash variance for all three models developed in this research. Most of the person level attributes were found to have statistically significant relationship with the occupant injury risks. Interestingly, no weather-related factors were directly related to occupant injury risk in the model considering weather-related crashes only. This might be attributed to the lack of variability in the weather-related factors during the winter conditions. However, two weather-related factors were found to be statistically significant in the model considering all crashes (weather and non-weather related). It was found that icy road surface condition has a positive effect on increasing the occupant injury risk for weather-related crashes.

The developed crash-injury risk models revealed several factors affecting occupant injury risk during winter weather conditions. This would be helpful to understand the factors contributing to increased injury risk for the occupants involved in crashes during winter weather seasons. Findings from the model results might have practical implications when considered from the perspective of the “4E’s” (e.g. engineering, enforcement, education, emergency response) of road safety. Results from the models can be helpful for educating road users to use some of type of occupant protection at all times to reduce the risk of sustaining injuries during winter seasons. Model results also revealed increased injury risk for aged occupants involved in weather-related crashes. Thus, the research finding can be useful to increase the awareness among aged populations while travelling during winter weather and decrease the potential injury risk. Interesting findings from the model results revealed increased injury risk for occupants involved in weather-related crashes on icy surface conditions. This finding strengthens the need to incorporate road surface condition by considering pavement friction or a friction surrogate accounting for road surface conditions in severity modeling to understand the effect of road surface condition on crash-injury risk. This finding can also have significant implications in highway and pavement design and operations. Occupant injury risk was found higher for all types of crashes (weather, all, non-weather related crashes) when the occupants are involved in non-collision crashes resulting overturn, rollover, and jackknife. Innovative and effective engineering countermeasures can be introduced for reducing injury risk during winter weather in this regard. For example, variable speed sign with dynamic display on roadway with icy conditions can be introduced during winter weather conditions. As the trapped occupants are found to be more susceptible to sustain injury risk, effort can be invested to make the emergency response more efficient to reach the crash site as soon as possible.

Finally, the findings of this modeling effort underscore the importance of accounting for the hierarchical structure of the crash data when developing crash severity models. The development of the multilevel logit models in a Bayesian framework revealed the unexplained variations in the individual occupant injury risk. Such information can be used to justify the use of multilevel model specification for addressing the hierarchy of the crash data over traditional econometric models that do not consider the within-crash correlation. The development of models ignoring the hierarchy of the crash data might result in biased parameter estimates and variables to be found incorrectly significant. Thus, the strong Intra-Class Correlations derived from the hierarchical models in this research reinforce the idea of considering the hierarchical nature of the crash data.

7.2 Limitations

There are some limitations in viewing the results of this dissertation, mainly pertaining to the data and methodology applied:

- A small percentage of weather-related crashes considered for developing SPFs occurred after the crew departed the sites. As crew reported weather information was used for developing crash frequency models at network level in this research, weather conditions during crash occurrences were assumed to be similar to that reported by the crew members in these cases.

- The crash frequency models did not consider the one-mile road segments that did not experience crashes during the four-year study period (zero-crash segments). SPFs are normally developed by considering the zero crash segments. Caution needs to be exercised for using the crash frequency models developed in this study. The frequency models developed in this research are capable to predict only winter weather-related crashes. The winter weather crash frequency is modeled as a function of weather related factors, roadway geometry characteristics, and traffic volume for similar type of roadway facilities.
- While use of event based or hourly based data at disaggregate level has been argued as more representative for weather related variables, aggregated weather data over four years for road segments were used to develop the crash frequency models. This level of aggregation was selected as frequency models were developed at network level for one-mile road segments rather than considering sites or group of sites.
- Road surface condition could not be included in the crash frequency models due to the unavailability of quantitative information on surface condition (e.g. pavement friction, road surface index).
- Real time traffic volume collected ATR could have been used in the crash frequency models instead of AADT. Estimate about the AADT during the winter season can be derived by factoring AADT. The factored AADT can be used in the models to represent traffic volume during the winter seasons. Real time traffic volumes were not used and the AADT was not corrected for the seasonality.

- As the multilevel severity models were developed by nesting occupants within crashes, a few vehicle-related variables (such as vehicle type, vehicle action) could be included in the models. Vehicle types might have significant effect on crash injury severity during the winter seasons.

7.3 Recommendations for Future Research

While this dissertation offered valuable insights on incorporating weather information to develop safety performance functions along with considering the importance of hierarchical structure in crash dataset for developing severity models, the following areas for future research efforts are recommended:

- Assess the stability of the models developed in this study over time via additional data collection (additional winter seasons).
- Develop hierarchical crash frequency models by considering spatial and temporal effects. Spatial models are particularly useful for quantifying the unmeasured effect of weather varying over space. Hierarchical frequency models by nesting the crashes within region or county can be useful in that regard.
- Examine the estimation of random effects of the slopes with respect to varying regions for the crashes considered instead of random effects over the constant terms only (considered herein).
- Compile broader datasets with large number of observations at crash level, vehicle level, and person level in a bid to consider the crash-vehicle-person hierarchy when developing multilevel severity models. However, considering large datasets do not

guarantee that number of vehicles nested within crashes would be large enough to consider the all three levels of hierarchy in the crash data.

- Incorporate winter maintenance information (such as salt consumption, manpower, cost of labor and equipment) in quantitative models for predicting crash frequency and severity. This would be beneficial from a policy perspective.
- Investigate the impact of salt and deicing materials on roadway infrastructure during winter seasons.

REFERENCES

Adams, T. M., Juni, E., Sproul, M., Xu, L., 2006. Regression tree models to predict winter storm costs. Transportation Research Record: Journal of the Transportation research Board 1948 (13), 117-124.

Andreescu, M. P., Frost. D. B., 1998. Weather and traffic accidents in Montreal, Canada. Climate Research 9, 225-230.

Andrey, J., 2010. Long-term trends in weather-related crash risks. Journal of Transport Geography 18, 247–258.

Andrey, J., Knapper, C., 2003. Weather as a risk factor in road transport. What are the most significant weather-related changes to the physical operating environment? In: Weather and Transportation in Canada, Department of Geography publication series 55.

Andrey, J., Mills. B., Suggett, J., Leahy. M., 2003. Weather as a chronic hazard for road transportation in Canadian cities‖. Natural Hazards 28, 319-343.

Blomqvist, G., Gustafsson, M., Eram, M., Unver, K., 2011. Prediction of salt on road surface: Tool to minimize salt use. Transportation Research Record: Journal of the Transportation research Board 2258 (16), 131-138.

Browne, W., Goldstein, H., Rasbah, J., 2001. Multiple membership multiple classification (MMMC) models. Statistical Modeling 1, 103-124.

Cameron, C. A., Trivedi, P. K., 1998. Regression Analysis of Count Data, Econometric Society Monograph No. 30, Cambridge University Press, Cambridge, U. K.

Chang, L. Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accident Analysis and Prevention 31 (5), 579–592.

Cheng, W., Washington, S. P., 2005. Experimental evaluation of hotspot identification methods. Accident Analysis & Prevention 37 (5), 870–881.

Cheng, W., Washington, S. P., 2008. New criteria for evaluating methods of identifying hot spots. Transportation Research Record: Journal of the Transportation research Board 2083 (9), 76-85.

Deng, Z., Ivan, J. N., Gårder, P., 2006. Analysis of factors affecting the severity of head-on crashes in two-lane rural highways in Connecticut. Transportation Research Record: Journal of the Transportation Research Board 1953 (16), 137–146.

Dissanayake, S., 2004. Comparison of severity affecting factors between young and older drivers involved in single vehicle crashes. IATSS Research 28 (2), 48-54.

Donnell, E. T., and Mason Jr, J. M., 2004. Predicting the severity of median-related crashes in Pennsylvania by using logistic regression. Transportation Research Record: Journal of the Transportation Research Board 1897 (8), 55–63.

Duncan, C. S., Khattak, A. J., Council, F. M., 1998. Applying the Ordered Probit Model to Injury Severity in Truck–Passenger Car Rear-End Collisions. Transportation Research Record: Journal of the Transportation research Board 1635 (9), 63-71.

Dupont, E., Martensen, H. (Eds.), 2007. Multilevel modelling and time series analysis in traffic research – methodology. Deliverable D7.4 of the EU FP6 project SafetyNet., Available on-line at: http://erso.swov.nl/safetynet/fixed/WP7/D7 4/ D7.4.pdf.

Dupont, E., Martensen, H., Papadimitriou, E., Yannis, G., 2010. Risk and protection factors in fatal accidents. Accident Analysis and Prevention 42 (2), 645–653

Dupont, E., Papadimitiou, H., Martensen, H., Yannis, G., 2013.Multilevel analysis in road safety research. Accident Analysis and Prevention 60, 402-411.

Edwards, J.B., 1998. The Relationship between road accident severity and recorded weather. Journal of Safety Research 29 (4), 249–26.

Eisenberg, D., 2004. The mixed effects of precipitation on traffic crashes. Accident Analysis and Prevention 36, 637–647.

Eisenberg, D., Warner, K. E., 2005. Effects of snowfalls on motor vehicle collisions, injuries, and fatalities. American Journal of Public Health 95 (1), 120-124.

Environment Canada, 2002. Winter Road Maintenance Activities and the Use of Road Salts in Canada: A Compendium of Costs and Benefits Indicators.

Federal Highway Administration (FHWA), 2010. How Do Weather Events Impact Roads? http://ops.fhwa.dot.gov/weather/q1_roadimpact.htm Accessed January 19, 2014.

Feng, C., 2001. Synthesis of studies on speed and safety. Transportation Research Record: Journal of the Transportation Research Record 1779 (12), 86-92.

Fu, L., Perchanok, M. S., Miranda-Moreno, L. F., Shah, Q. A., 2006. Effects of winter weather and maintenance treatments on highway safety. In: Transportation Research Board 85th Annual Meeting Compendium of Papers CD-ROM, Transportation Research Board of the National Academies, Washington, DC.

Hauer, E., 1996. Identification of sites with promise. Transportation Research Record: Journal of the Transportation Research Board 1542 (9), 54-60.

Hauer, E., Kononov, J., Allery, B., Griffith, M. S., 2002. Screening the road network for sites with promise. Transportation Research Record: Journal of the Transportation Research Board 1784 (4), 27-32.

Helai, H., Chor, C.H., Haque, M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: a Bayesian hierarchical analysis. Accident Analysis and Prevention 40, 45–54.

Hermans, E., Brijs, T., Stiers, T., Offermans, C., 2006a .The Impact of weather conditions on road safety investigated on an hourly basis. In: Transportation Research Board 85th Annual Meeting Compendium of Papers CD-ROM, Transportation Research Board of the National Academies, Washington, DC.

Hermans, E., Wets; G., Van-Den Bossche, F., 2006b. Frequency and severity of Belgian road traffic accidents studied by state-space methods. Journal of Transportation and Statistics 9 (6).

Hox, J., 2002. Multilevel Analysis, Techniques and Applications, Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey.

Huang, H., Abdel-Aty, M., 2010.Multilevel data and Bayesian analysis in traffic safety. Accident Analysis and Prevention 42 (6), 1555-1565.

Jones, A.P., Jørgensen, S.H., 2003. The use of multilevel models for the prediction of road accident outcomes. Accident Analysis and Prevention 35, 59–70.

Jung, S, Qin, X., Noyce, D. A., 2010. Rainfall effect on single-vehicle crash severities using polychotomous response models. Accident Analysis and Prevention 42 (1), 213–224.

Khattak, A. J., Pawlovich. M. D., Souleyrette, R. R., Hallmark, S. L., 2002. Factors related to more severe older driver traffic crash injuries. Journal of Transportation Engineering 128(3), 243-249.

Khattak, A., Kantor, P., Council, F. M., 1998. Role of adverse weather in key crash types on limited access: Roadways implications for advanced weather systems. Transportation Research Record: Journal of the Transportation Research Board 1621 (2), 10-19.

Khattak, A.J., Knapp, K. K., 2001. Interstate Highway Crash Injuries during Winter Snow and non-snow Events. Transportation Research Record: journal of the Transportation Research Board 1746 (5), 30-36.

Kim, D. G., Lee, Y., Washington, S., Choi, K., 2007. Modeling crash outcome probabilities at rural intersections: application of hierarchical binomial logistic models. Accident Analysis and Prevention 39, 125–134.

Knapp, K. K., Smithson, L. D., 2000. Winter storm event volume impact analysis using multiple-source archived monitoring data. Transportation Research Record: Journal of the Transportation research Board 1700 (3), 10-16.

Knapp, K. K., Smithson, L. D., 2001. Use of mobile video data collection equipment to investigate winter weather vehicle speeds. Transportation Research Record: Journal of the Transportation research Board 1745 (7), 53-60.

Knapp, K.K., Smithson. D. L., Khattak, A. J., 2000. The mobility and safety impacts of winter storm events in a freeway environment. In: Mid-Continent Transportation Symposium, May 15-16, Iowa State University, Ames, Iowa.

Lee, C., Loh, W. Y., Qin. X., Sproul, M., 2008. Development of new performance measure for winter maintenance by using vehicle speed data. Transportation Research Record: Journal of the Transportation research Board 2055 (11), 89-98.

Lee, C., Ran, L., 2004. Pilot study of speed-recovery duration as a measure of winter maintenance performance. Transportation Research Record: Journal of the Transportation research Board 1877 (15), 137-143.

Lenguerrand, E., Martin, J.L., Laumon, B., 2006. Modelling the hierarchical structure of road crash data – application to severity analysis. Accident Analysis and Prevention 38, 43–53.

Mergia, W. Y., 2010. Exploring Factors Contributing to Injury Severity at Freeway Merging and Diverging areas. In: MS Thesis, University of Dayton.

Monsere, C. M., Bosa, P. G., Bertini, R. L., 2008. Combining climate, crash, and highway data for improved ranking of speed and winter-weather related crash locations in Oregon. Journal of Transportation Engineering 134 (7), 287-296.

Morgan, A., Mannering, F., 2011. The effect of road-surface conditions, age, and gender on driver-injury severity. Accident Analysis and Prevention 43 (5), 1852-1863.

Nilsson, G., Obrenovic, A., 1998. Road accident risk twice as high in the winter, Report No. 435. Swedish Road and Traffic Research Institute (VTI).

Nixon, W. A., Qiu, L., 2005. Developing a storm severity index. Transportation Research Record: Journal of the Transportation research Board 1911 (14), 143-148.

Padget, E. D., Knapp. K. K., Thomas, G. D., 2001. Investigation of winter-weather speed variability in sports utility vehicles, pickup trucks, and passenger cars. Transportation Research Record: Journal of the Transportation research Board 1779 (16), 116-124.

Qin, X., Noyce, D. A., Lee, C., Kinar, J. R., 2006. Snowstorm event-based crash analysis. Transportation Research Record: Journal of the Transportation research Board 1948 (15), 135-141.

Qiu, L., Nixon, W. A., 2008. Effects of adverse weather on traffic crashes: systematic review and meta-analysis. Transportation Research Record: Journal of the Transportation research Board 2055 (16), 139–146.

Quddus, M. A., Noland, R. B., Chin, H. C., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. Journal of Safety Research 33(4), 445– 462.

Quddus, M. A., Wang, C., Ison, S. G., 2010. Road traffic congestion and crash severity: Econometric analysis using ordered response models. Journal of Transportation Engineering 136 (5), 424–435.

Raftery, A.E. & Lewis, S.M., 1992. How many iterations in the Gibbs sampler? In J.M. Bernado, J.O. Berfer, A.P. Dawid & A.F.M. Smith, eds., Bayesian Statistics 4, 765-776, Oxford University Press.

Rasbash, J., Goldstein, H., Browne, W.J., Yang, M., Woodhouse, G., 2000. The MLwiN Command Interface Version 1.3. Institute of Education, University of London, London.

Russ, A., Mitchell, G. F., Richardson, W., 2008. Decision tree for pretreatments for winter maintenance. Transportation Research Record: Journal of the Transportation research Board 2055 (13), 106-115.

Sherif, A., 2005. Impact of road surface temperature and condition on the risk of winter vehicle collisions. In: PhD thesis, Department of Civil and Environmental Engineering, Carleton University, Ottawa, Ontario, Canada.

Strong, C.K., Ye, Z., Shi, X., 2010. Safety effects of winter weather: The state of knowledge and remaining challenges. Transport Reviews 6, 677-699.

US Department of Commerce, 2002. Weather information for surface transportation, national needs assessment report. Report No. FCM-R18-2002, Washington, DC.

Usman, T., Fu, L., Miranda-Moreno, L. F., 2010. Quantifying safety benefit of winter road maintenance. Accident Analysis and Prevention 42 (6), 1878-1887.

Usman, T., Fu, L., Miranda-Moreno, L. F., 2011. Accident prediction models for winter road safety. Transportation Research Record: Journal of the Transportation research Board 2237 (16), 144-151.

Usman, T., Fu, L., Miranda-Moreno, L. F., 2012. A disaggregate model for quantifying the safety effects of winter road maintenance at an operational level. Accident Analysis and Prevention 48, 368-378.

Usman, T., Miranda-Moreno, L. F., Fu, L., 2012 .Analysis of injury severity outcomes of highway winter crashes: A multi-level modeling approach. In: Transportation Research Board 91st Annual Meeting Compendium of Papers CD-ROM, Transportation Research Board of the National Academies, Washington, DC.

Vanlaar, W., 2005. Multilevel modelling in traffic safety research: two empirical examples illustrating the consequences of ignoring hierarchies. Traffic Injury Prevention 6 (4), 311–316.

Yannis, G., Papadimitriou, E., Dupont, E., Martensen, H., 2010. Estimation of fatality and injury risk by means of in-depth fatal accident investigation data. Traffic Injury Prevention 11 (5), 492–502.

Ye, Z., Shi, X., Strong, C. K., Greenfield, T. H., 2009. Evaluation of effects of weather information on winter maintenance costs. Transportation Research Record: Journal of the Transportation research Board 2107 (11), 104-110.

APPENDIX A

COLLECTED DATA SAMPLES

A.1 RWIS DATA SAMPLE

SiteID	DATE	TIME	Air temperature	Pave temp (sensor 0)	Pave temp (sensor 1) (Sensor	Pave temp (sensor 2)	Pave temp (sensor 3)
RDAI4	10/15/2008	0:01	47.7	49.28	49.1	51.08	49.28
RDAI4	10/15/2008	0:11	47.7	49.28	M	50.9	49.28
RDAI4	10/15/2008	0:21	47.3	49.28	M	50.72	49.28
RDAI4	10/15/2008	0:31	47.1	49.1	M	50.54	49.1
RDAI4	10/15/2008	0:41	47.1	48.92	48.74	50.54	48.92
RDAI4	10/15/2008	0:51	47.1	48.92	48.56	50.36	48.92
RDAI4	10/15/2008	1:01	46.9	48.74	48.56	50.36	48.74
RDAI4	10/15/2008	1:11	46.9	48.74	48.38	50.36	48.74
RDAI4	10/15/2008	1:21	47.1	48.74	48.56	50.36	48.74
RDAI4	10/15/2008	1:31	47.1	48.56	M	50.18	48.56
RDAI4	10/15/2008	1:41	46.6	48.38	48.2	50.18	48.38
RDAI4	10/15/2008	1:51	46.2	48.2	48.2	50	48.2
RDAI4	10/15/2008	2:02	45.7	47.84	47.84	49.64	47.84
RDAI4	10/15/2008	2:11	45.3	47.66	47.48	49.46	47.66
RDAI4	10/15/2008	2:21	45.3	47.48	47.48	49.28	47.48
RDAI4	10/15/2008	2:31	45.5	47.3	M	49.1	47.3
RDAI4	10/15/2008	2:42	45.7	47.3	M	49.28	47.3
RDAI4	10/15/2008	2:51	45.9	47.48	47.48	49.1	47.48

A.2 AWOS DATA SAMPLE

FAACODE	Airport	DATE	TIME	Air temperature	Wind speed	Wind gust	Visibility
IKV	ANKENY	10/15/2008	0:00	53	0	0	10
IKV	ANKENY	10/15/2008	0:10	52	0	0	10
IKV	ANKENY	10/15/2008	0:20	52	0	0	10
IKV	ANKENY	10/15/2008	0:30	52	0	0	10
IKV	ANKENY	10/15/2008	0:40	52	0	0	10
IKV	ANKENY	10/15/2008	0:50	53	0	0	10
IKV	ANKENY	10/15/2008	1:00	53	0	0	10
IKV	ANKENY	10/15/2008	1:10	52	0	0	10
IKV	ANKENY	10/15/2008	1:20	53	3	0	10
IKV	ANKENY	10/15/2008	1:30	53	0	0	10
IKV	ANKENY	10/15/2008	1:40	53	3	0	10
IKV	ANKENY	10/15/2008	1:50	53	0	0	10
IKV	ANKENY	10/15/2008	2:00	53	0	0	10
IKV	ANKENY	10/15/2008	2:10	53	0	0	10
IKV	ANKENY	10/15/2008	2:20	53	3	0	10
IKV	ANKENY	10/15/2008	2:30	52	0	0	10
IKV	ANKENY	10/15/2008	2:40	52	0	0	10
IKV	ANKENY	10/15/2008	2:50	52	0	0	10

A.3 COST CENTER DATA SAMPLE

COST CENTER	EVENT_DATE	PRECIP_START_TIME	PRECIP_END_TIME	PRECIP_TYPE	AIR_TEMP	PAVEMENT_TEMP	WIND_DIRECTION	VELOCITY	VISIBILITY
551602	15-Oct-06	15-Oct-06	15-Oct-06	2	48	48	4	2	5
551602	16-Oct-06	16-Oct-06	16-Oct-06	2	51	51	2	12	5
551602	21-Oct-06	21-Oct-06	21-Oct-06	2	39	38	4	10	5
551602	10-Nov-06	10-Nov-06	10-Nov-06	2	37	37	4	15	5
551602	10-Nov-06	10-Nov-06	10-Nov-06	6	32	32	4	15	4
551602	28-Nov-06	28-Nov-06	28-Nov-06	2	40	40	6	10	5
551602	29-Nov-06	29-Nov-06	29-Nov-06	2	34	35	4	12	5
551602	06-Dec-06	06-Dec-06	06-Dec-06	6	32	30	4	15	5
551602	20-Dec-06	20-Dec-06	20-Dec-06	2	36	36	7	3	5
551602	21-Dec-06	21-Dec-06	21-Dec-06	2	36	36	7	3	4
551602	22-Dec-06	22-Dec-06	22-Dec-06	2	37	37	7	5	5
551602	22-Dec-06	22-Dec-06	22-Dec-06	2	35	36	8	9	5
551602	30-Dec-06	30-Dec-06	30-Dec-06	2	50	50	6		5
551602	31-Dec-06	31-Dec-06	31-Dec-06	5	36	40	4	18	2
551602	31-Dec-06	31-Dec-06	31-Dec-06	6	31	30	4	22	1
551602	01-Jan-07	01-Jan-07	01-Jan-07	8	25	24	4	15	5
551602	05-Jan-07	05-Jan-07	05-Jan-07	2	35	35	4	5	4
551602	05-Jan-07	05-Jan-07	05-Jan-07	6	35	35	4	5	3

A.4 CREW ACTIVITY INFORMATION SAMPLE

COST CENTER	CREW_ON_ROAD	CREW_OFF_ROAD	PRECIP_TYPE	SERVICE_LEVEL	OPERATION
551602	10-Nov-06	10-Nov-06	2	1	2
551602	10-Nov-06	10-Nov-06	6	1	2
551602	10-Nov-06	10-Nov-06	2	2	2
551602	10-Nov-06	10-Nov-06	6	2	2
551602	10-Nov-06	10-Nov-06	2	3	2
551602	10-Nov-06	10-Nov-06	6	3	2
551602	29-Nov-06	29-Nov-06	2	1	1
551602	29-Nov-06	29-Nov-06	2	2	1
551602	29-Nov-06	29-Nov-06	2	3	1
551602	22-Dec-06	22-Dec-06	2	2	1
551602	22-Dec-06	22-Dec-06	2	2	1
551602	22-Dec-06	22-Dec-06	2	3	1
551602	22-Dec-06	22-Dec-06	2	3	1
551602	31-Dec-06	31-Dec-06	5	1	2
551602	31-Dec-06	31-Dec-06	6	1	2
551602	31-Dec-06	31-Dec-06	5	2	2
551602	31-Dec-06	31-Dec-06	6	2	2

APPENDIX B

PREPARED DATA SAMPLES FOR DEVELOPING SPFs

B.1 SAMPLE PREPARED DATA FOR DEVELOPING SPF (INTERSTATE/FREEWAY)

ROWID/SEGMENT	CRASH FREQUENCY	AIR TEMPERATURE	PAVEMENT TEMPERATURE	WIND SPEED	VISIBILITY	SNOW AMOUNT	SURFACE WIDTH	SPEED LIMIT	AADT
3393	50	15.17	18.12	10.63	2.56	1.23	34	65	36400
9078	35	15.57	15.17	12.55	4.03	0.94	56	55	23400
4187	48	20.5	23.58	22.77	3.1	2.63	64	65	79900
9364	36	14.25	17.89	11.19	0	4	49	60	62600
3342	24	17.29	15.96	29.4	1.57	0.88	24	70	13700
3219	29	11.52	13.93	11.43	3.85	1.67	51	55	21100
4198	26	24.46	25.68	18.23	3.08	3	34	70	42600
6818	24	15.96	20.59	15.63	4.46	1.58	24	70	34400
3394	30	14.04	14.31	13.46	3	2.08	51	55	41300
4163	20	25	26	18.14	4.21	0.75	24	70	14900
4154	21	18.52	24.19	15.9	4.29	1.63	24	70	20000
6429	27	23.78	29.17	13.3	3.74	2.01	32	55	45800
6714	22	26.05	25.27	21	3.23	1.64	34	70	27100
4199	23	25.39	28.39	15.39	3.61	3.52	24	70	42600
6430	30	16.57	23.78	13.3	3.4	2.21	40	55	52000
4157	21	22.76	25.1	12.86	4.52	1.17	34	70	20000
6792	22	24.32	27.09	14.45	4.5	2.45	24	70	36500
6915	24	23.21	25.28	10.79	2.54	3.04	34	65	44200

B.2 SAMPLE PREPARED DATA FOR DEVELOPING SPF (MULTILANE DIVIDED/UNDIVIDED)

ROWID/SEGMENT	CRASH FREQUENCY	AIR TEMPERATURE	PAVEMENT TEMPERATURE	WIND SPEED	VISIBILITY	SNOW AMOUNT	SURFACE WIDTH	SPEED LIMIT	AADT
1092	27	16.74	22.32	18.68	3.74	1.68	48	35	22000
6039	26	21.31	21.5	13.31	4.12	2.23	48	40	30900
6074	26	18.96	19.92	14.42	3.92	1.98	64	45	25200
2633	27	9.12	13.26	9.07	3.77	2.24	43	35	36400
1096	23	14.26	18.73	23.23	3.87	1.72	60	35	18700
1091	24	10.92	17.89	19.39	3.79	1.38	58	35	22000
6075	25	18.32	19.8	12.72	3.84	2.01	63	35	24200
9077	26	15.5	16.92	15.08	3.58	1.49	68	55	33700
9330	17	16.42	18.12	12.91	3.86	2.32	30	65	22000
9481	19	17.79	17.58	15.11	3.53	1.95	48	45	18300
2293	18	14.06	16.83	17.61	2.94	2.33	52	35	12300
8209	21	23.38	26.24	17.71	3.24	2.62	42	35	26800
1960	21	20.71	21.45	11.1	4.88	1.39	62	35	16500
5267	17	26.31	26.56	9.92	4.77	1.97	48	55	18400
9472	14	24.71	24.29	7.43	0	1.43	24	45	15000
5659	16	13.51	15.69	15.81	3.63	2.49	48	35	11000
6051	19	20.26	22.05	21.95	3.74	2.01	48	35	20400
1053	19	19.95	23.06	10.74	4.37	2.13	42	40	26200
3230	19	19.42	23.51	16.67	3.42	2.45	40	40	28800

B.3 SAMPLE PREPARED DATA FOR DEVELOPING SPF (TWO-LANE)

ROWID/SEGMENT	CRASH FREQUENCY	AIR TEMPERATURE	PAVEMENT TEMPERATURE	WIND SPEED	VISIBILITY	SNOW AMOUNT	SURFACE WIDTH	SPEED LIMIT	AADT
5338	1	31	30	16	3	3.5	24	55	240
5551	1	0	0	0	0	0	24	55	240
6002	2	9.5	12	20	2	3.25	16	65	320
8375	1	0	37	0	5	0.75	24	55	360
7366	1	2	9	10	3	4	16	45	380
1465	1	26	31	14	4	3	16	55	390
8843	1	34	36	6	4	1.5	22	55	400
8881	1	22	40	25	5	0	24	55	420
8882	1	22	22	15	5	0	24	55	420
2046	1	0	0	20	0	2	22	55	450
4314	1	18	20	13	4	0	22	55	480
8822	1	31	30	9	4	4	22	55	480
1983	1	25	0	40	0	2.5	24	55	530
7444	1	-7	-1	7	5	0	22	55	530
7448	1	31	32.5	13	3	1	22	55	530
1614	1	24	27	5	5	0.5	24	55	540
4781	1	30	31	6	4	1.75	22	55	560
4783	1	30	29	10	4	2	22	55	560
7371	1	23	19	15	2	2	24	55	560

APPENDIX C

SAMPLE RANKING OF ROADWAY SEGMENTS

C.1 SAMPLE RANKING OF ROADWAY SEGMENTS (INTERSTATE/FREEWAY) BASED ON PSI

Segment ID	Observed	Predicted Crash Frequency	Weight	Empirical-Bayes Adjusted Crash Frequecny	PSI
9164	54	43.41924071	0.005269278	53.94424703	10.525006
6783	25	14.32131115	0.015806136	24.83121119	10.5099
2705	14	2.634370856	0.080296865	13.08737561	10.453005
4271	15	3.958977459	0.054906001	14.39378161	10.434804
3388	16	5.243966574	0.042017063	15.54806306	10.304096
6665	18	7.602777982	0.029363784	17.69469822	10.09192
9354	19	8.645464001	0.025914138	18.73167112	10.086207
9082	16	5.533481353	0.039906436	15.58231854	10.048837
3392	17	6.64780945	0.03344088	16.65381364	10.006004
4156	16	5.814367639	0.038051954	15.61241678	9.7980491
6852	15	4.781309356	0.045896189	14.53100105	9.7496917
6693	16	5.872159675	0.037691574	15.61826576	9.7461061
6850	16	5.982355776	0.037022992	15.62911684	9.6467611
6811	24	14.22013301	0.015916809	23.84433573	9.6242027
4162	14	3.904518823	0.055629206	13.4383964	9.5338776
3689	16	6.241350487	0.035541268	15.65316523	9.4118147
6721	15	5.260209569	0.041892754	14.59197335	9.3317638
4266	14	4.156078421	0.052438643	13.48379811	9.3277197
3281	13	3.352231995	0.06420578	12.38055753	9.0283255
6820	16	6.717523634	0.033105321	15.69270064	8.975177
9171	23	13.89407483	0.016284252	22.85171682	8.957642
4160	15	5.687848366	0.038865477	14.63807878	8.9502304
6874	16	6.915943819	0.032186091	15.70761974	8.7916759
9079	23	14.0911036	0.016060215	22.85692121	8.7658176
4196	18	9.545272322	0.023528756	17.80107077	8.2557985
4215	15	6.847142093	0.032498994	14.73504032	7.8878982
4213	17	8.930320963	0.025108291	16.79738415	7.8670632
4130	12	3.72917032	0.058092979	11.51952286	7.7903525
6853	13	4.890482656	0.044917641	12.63573961	7.745257
4153	13	4.921591893	0.044646394	12.63932821	7.7177363
5762	15	7.072392328	0.031496527	14.75030789	7.6779156
4124	12	3.912664165	0.055519828	11.55099251	7.6383283

C.1 SAMPLE RANKING OF ROADWAY SEGMENTS (INTERSTATE/FREEWAY) BASED ON PSI (CONTIN'D)

Segment ID	Observed	Predicted Crash Frequency	Weight	Empirical-Bayes Adjusted Crash Frequecny	PSI
2680	11	2.871275563	0.074163032	10.39714915	7.5258736
6922	14	6.378560351	0.034803344	13.73474841	7.3561881
6884	14	6.408258183	0.034647643	13.73696404	7.3287059
4217	13	5.403797676	0.040825037	12.68988476	7.2860871
4122	11	3.802727279	0.057033363	10.58951533	6.7867881
9081	16	9.06080292	0.024755664	15.82821557	6.7674126
6821	15	8.132103147	0.027505042	14.81109821	6.6789951
4152	12	5.037107626	0.04366723	11.69594978	6.6588422
4067	12	5.100235883	0.04315006	11.70227476	6.6020389
6662	10	2.889929287	0.07371962	9.475848292	6.585919
3396	13	6.212027886	0.035703043	12.75764874	6.5456209
6796	21	14.34953561	0.015775537	20.89508535	6.5455497
6784	19	12.34101132	0.018296062	18.87816673	6.5371554
4066	11	4.268312004	0.051130291	10.65580684	6.3874948
6719	13	6.41602632	0.034607146	12.77214746	6.3561211
2810	14	7.602245378	0.029365781	13.81212494	6.2098796
6871	11	4.511525324	0.048507597	10.68525968	6.1737344
6791	13	6.900840566	0.032254262	12.80327611	5.9024355
6684	12	5.866606648	0.037725904	11.76861219	5.9020055
6738	13	7.000192358	0.031811049	12.80913983	5.8089475
4148	11	4.931957189	0.044556743	10.72962778	5.7976706
3648	9	2.77039557	0.076656559	8.52245996	5.7520644
9165	35	29.2669362	0.00779742	34.95529689	5.6883607
6761	13	7.300150275	0.030543879	12.82590448	5.5257542
6431	22	16.40469124	0.013826527	21.92263631	5.5179451
6715	11	5.279328823	0.041747372	10.76117701	5.4818482
9223	12	6.321365884	0.035107183	11.80063915	5.4792733
2245	9	3.145722342	0.068133566	8.601127188	5.4554048
3691	13	7.37925841	0.030226336	12.83010558	5.4508472
4170	17	11.44088762	0.019707156	16.89044571	5.4495581
4184	22	16.49289542	0.013753599	21.92425749	5.4313621
6917	15	9.518594902	0.023593144	14.87067642	5.3520815
6701	11	5.4413705	0.040554571	10.77457216	5.3332017
6816	18	12.62708646	0.017888967	17.90388413	5.2767977

C.2 SAMPLE RANKING OF ROADWAY SEGMENTS (MULTILANE DIVIDE/UNDIVIDED) BASED ON PSI

Segment ID	Observed	Predicted Crash Frequency	Weight	Empirical-Bayes Adjuested Crash Frequency	PSI
1709	15	6.868870355	0.022482	14.81719562	7.948325
1959	16	8.007330401	0.019347	15.84536223	7.838032
1050	26	18.13298094	0.008637	25.93205304	7.799072
3227	19	11.18284012	0.01393	18.89110676	7.708267
4681	13	5.265561984	0.029128	12.77470984	7.509148
5214	10	1.984371417	0.07374	9.408923636	7.424552
6038	19	11.71180571	0.013309	18.90299962	7.191194
5220	10	2.963702845	0.050607	9.643916394	6.680214
1097	14	7.182695547	0.021521	13.85328549	6.67059
1086	17	10.24666073	0.015183	16.89746129	6.650801
5660	12	5.212291848	0.029417	11.80032516	6.588033
5905	15	8.37332221	0.018517	14.87729086	6.503969
3216	14	7.460418064	0.020736	13.8643928	6.403975
7359	13	6.442560225	0.023934	12.84305364	6.400493
9480	14	7.891861099	0.019625	13.88012793	5.988267
5295	13	7.06525288	0.021871	12.87020229	5.804949
7358	14	8.184246666	0.018937	13.88986626	5.70562
3487	10	4.073206577	0.037337	9.778713907	5.705507
1958	15	9.397478628	0.016533	14.90737497	5.509896
9489	13	7.417923211	0.020853	12.88359871	5.465675
5218	8	2.203300054	0.066904	7.612180179	5.40888
7820	11	5.443083819	0.028205	10.84326724	5.400183
1054	13	7.591025795	0.020387	12.88972798	5.298702
1772	8	2.438870842	0.060834	7.661691725	5.222821
4040	8	2.459061269	0.060365	7.665520567	5.206459
6077	13	7.766631711	0.019935	12.89567228	5.129041
3217	11	5.75804181	0.026703	10.86002186	5.10198
2373	9	3.691821239	0.041035	8.782176987	5.090356
9488	13	7.811379073	0.019823	12.89714511	5.085766
9471	9	3.72146053	0.040722	8.785048143	5.063588
177	8	2.854050708	0.052449	7.730100113	4.876049
2796	7	1.670843053	0.086382	6.539654803	4.868812
7650	8	2.880667732	0.05199	7.73384811	4.85318
3474	8	2.89171478	0.051801	7.735384455	4.84367
6052	11	6.096309907	0.025259	10.87613704	4.779827
5219	8	2.981270189	0.050323	7.747440058	4.76617
5289	10	5.090462088	0.0301	9.852223059	4.761761

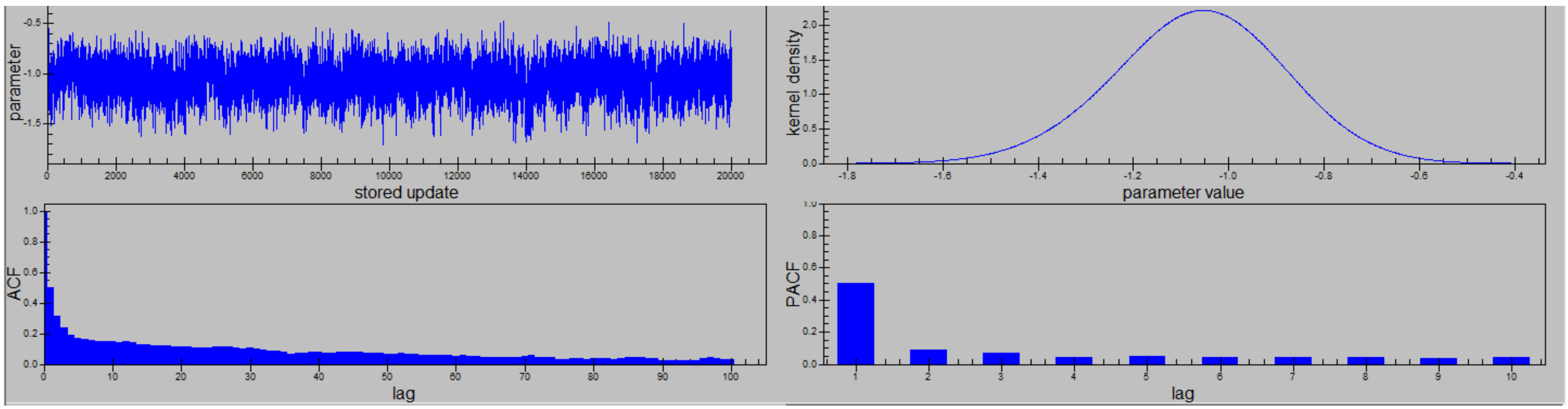
C.2 SAMPLE RANKING OF ROADWAY SEGMENTS (MULTILANE DIVIDE/UNDIVIDED) BASED ON PSI (CONTIN'D)

Segment ID	Observed	Predicted Crash Frequency	Weight	Empirical-Bayes Adjuested Crash Frequency	PSI
9338	8	3.172656241	0.047432	7.771030525	4.598374
6045	13	8.314257723	0.018647	12.91262712	4.598369
1961	13	8.31664087	0.018641	12.91269611	4.596055
8210	16	11.38017822	0.013692	15.93674642	4.556568
9337	10	5.36212118	0.028619	9.867270085	4.505149
2449	9	4.333803549	0.03517	8.835887866	4.502084
2625	9	4.44628628	0.034311	8.843756564	4.39747
8170	8	3.42474421	0.044094	7.798256966	4.373513
6040	14	9.636621331	0.016129	13.92962271	4.293001
960	9	4.858161196	0.031494	8.86955726	4.011396
67	7	2.775214012	0.053859	6.772458547	3.997245
1099	11	6.912077912	0.022345	10.90865683	3.996579
2617	7	2.936925642	0.051045	6.792602296	3.855677
2485	11	7.061761292	0.021881	10.91382589	3.852065
2615	7	2.99301528	0.050136	6.799106208	3.806091
1782	6	1.885434991	0.077311	5.681899689	3.796465
8140	7	3.016518724	0.049765	6.801763234	3.785245
5885	9	5.163990602	0.029684	8.886131486	3.722141
8452	7	3.169826282	0.047472	6.818173575	3.648347
6459	7	3.170824341	0.047458	6.818275456	3.647451
5818	14	10.35885736	0.015021	13.94530484	3.586447
3639	6	2.180139679	0.067566	5.741906285	3.561767
1052	18	14.429197	0.01083	17.9613285	3.532131
2822	6	2.261556217	0.065293	5.75590696	3.494351
2384	10	6.438778097	0.023948	9.914716521	3.475938
6464	7	3.363542752	0.044861	6.83686598	3.473323
9476	17	13.5166279	0.011553	16.95975783	3.44313
1708	10	6.52492689	0.023639	9.917852384	3.392925
5124	6	2.470141151	0.060111	5.787817971	3.317677
2363	6	2.48489366	0.059775	5.789884204	3.304991
3779	7	3.559919263	0.042491	6.853826867	3.293908
3680	7	3.580341917	0.042259	6.855488462	3.275147
5535	5	1.390872616	0.101997	4.63188035	3.241008
2376	6	2.572653604	0.057854	5.801714391	3.229061
1322	6	2.586672127	0.057558	5.803534035	3.216862
3477	6	2.590315481	0.057482	5.804003914	3.213688

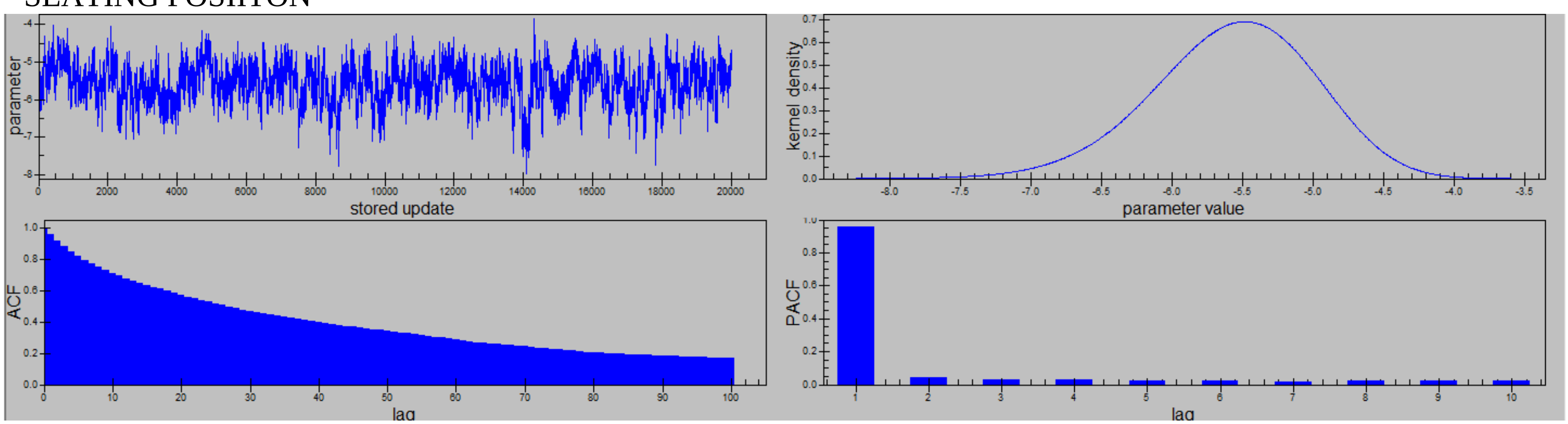
APPENDIX D

PLOTS OF DYNAMIC TRACE, KERNEL DENSITY AND ACF FOR THE PARAMETERS IN SEVERITY MODEL (WEATHER-RELATED CRASHES ONLY)

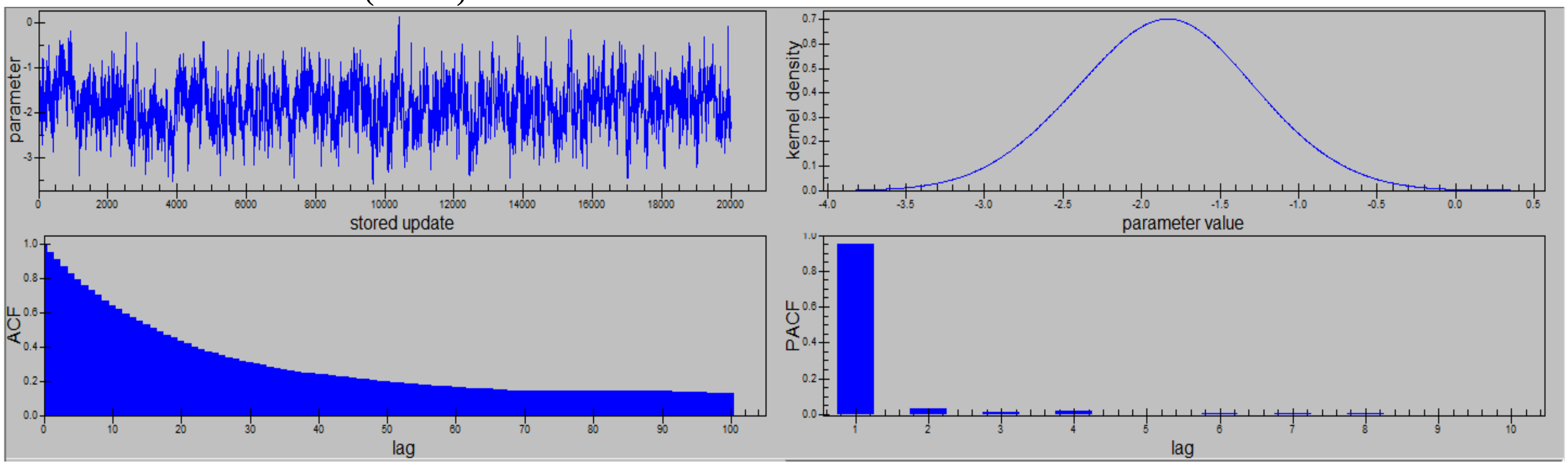
GENDER



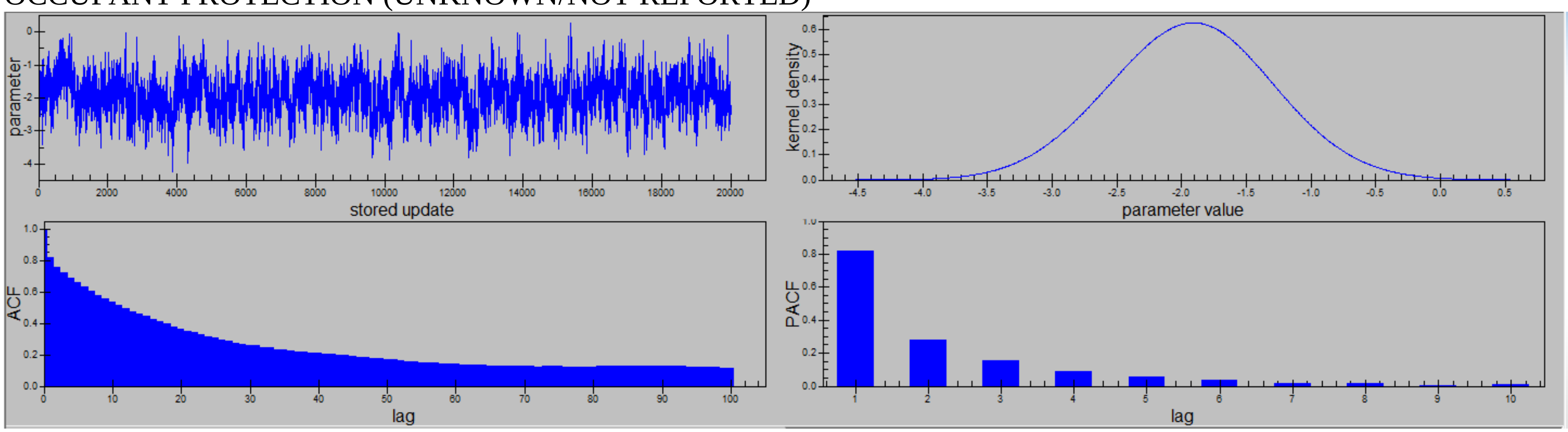
SEATING POSIITON



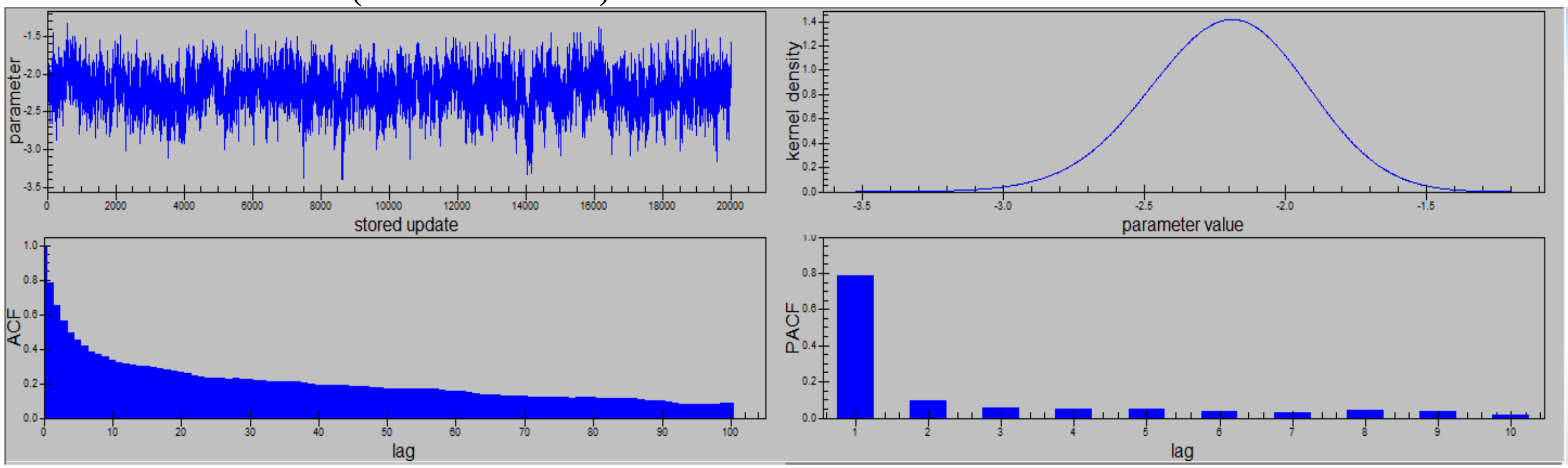
OCCUPANT PRTECTION (USED)



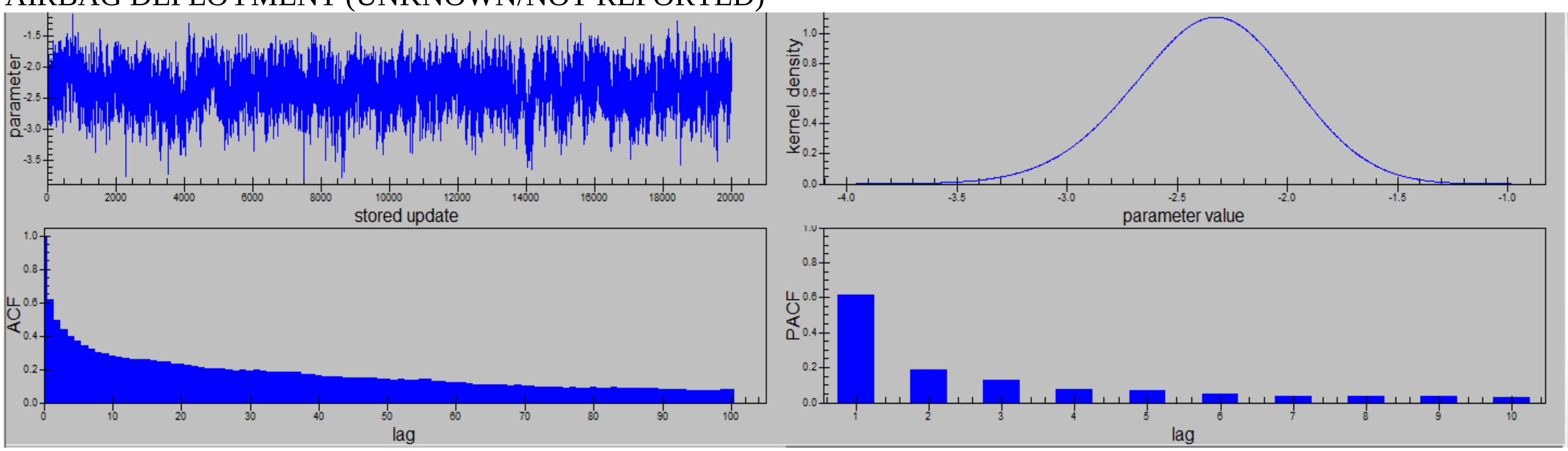
OCCUPANT PROTECTION (UNKNOWN/NOT REPORTED)



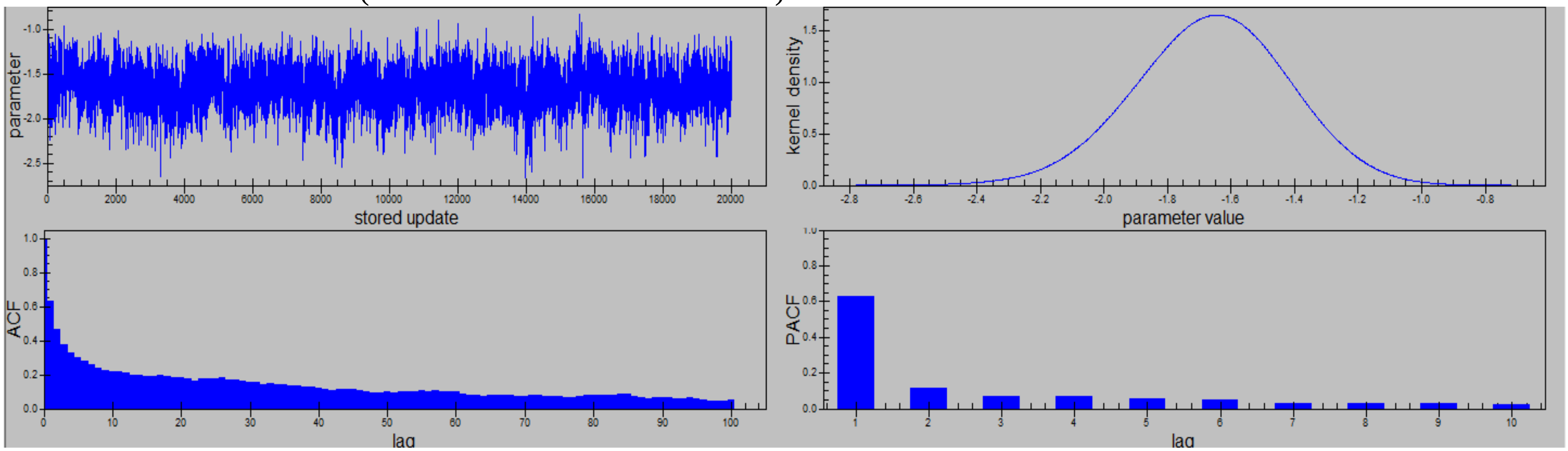
AIRBAG DEPLOYMENT (NOT DEPLOYED)



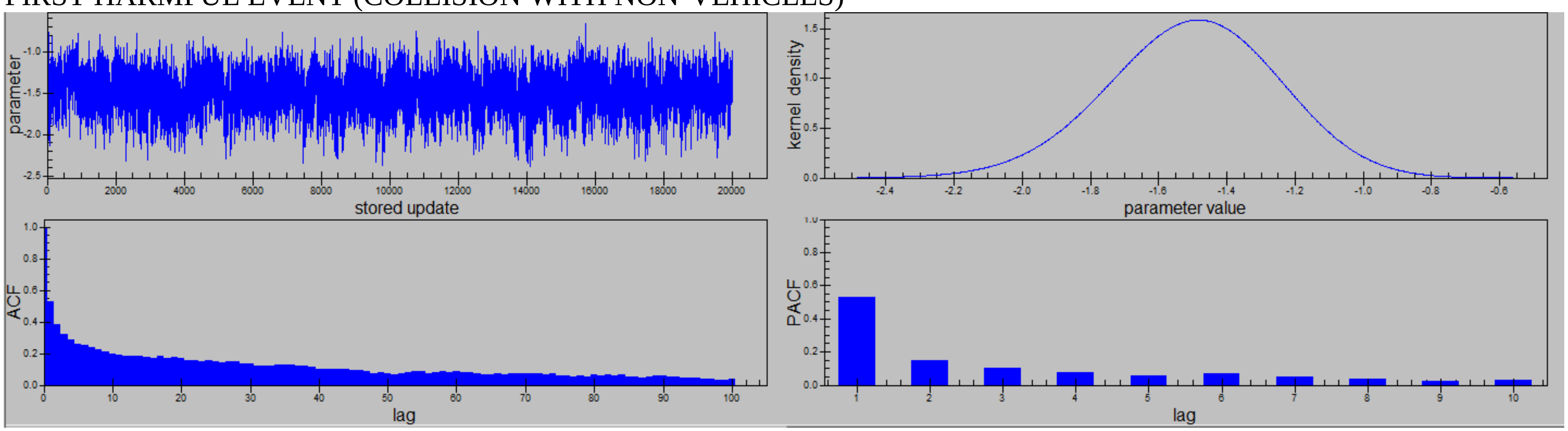
AIRBAG DEPLOYMENT (UNKNOWN/NOT REPORTED)



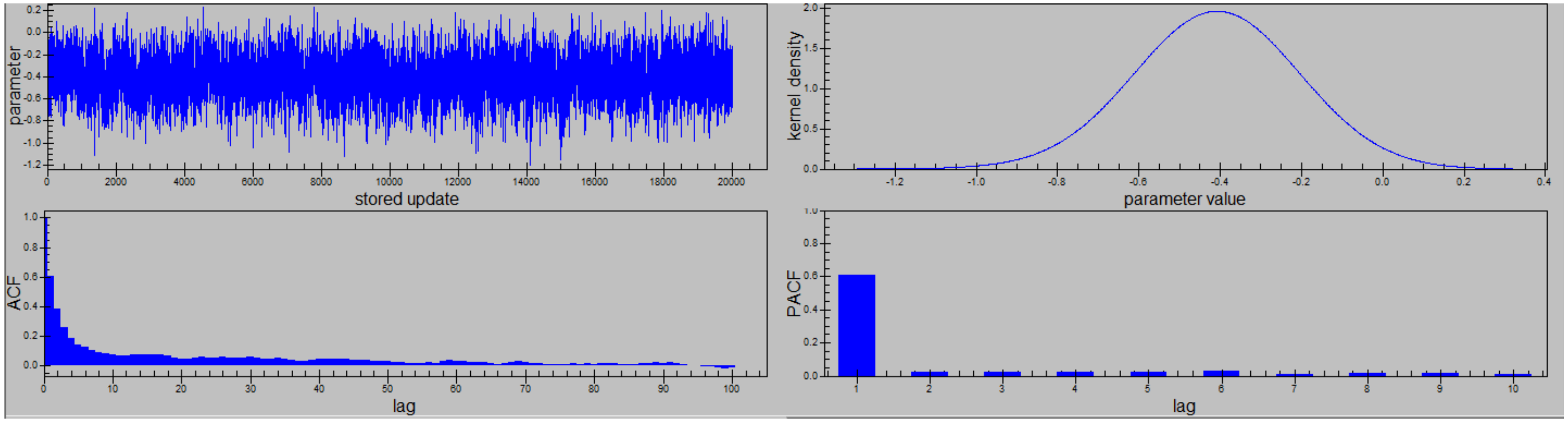
FIRST HARMFUL EVENT (COLLISION WITH VEHICLES)



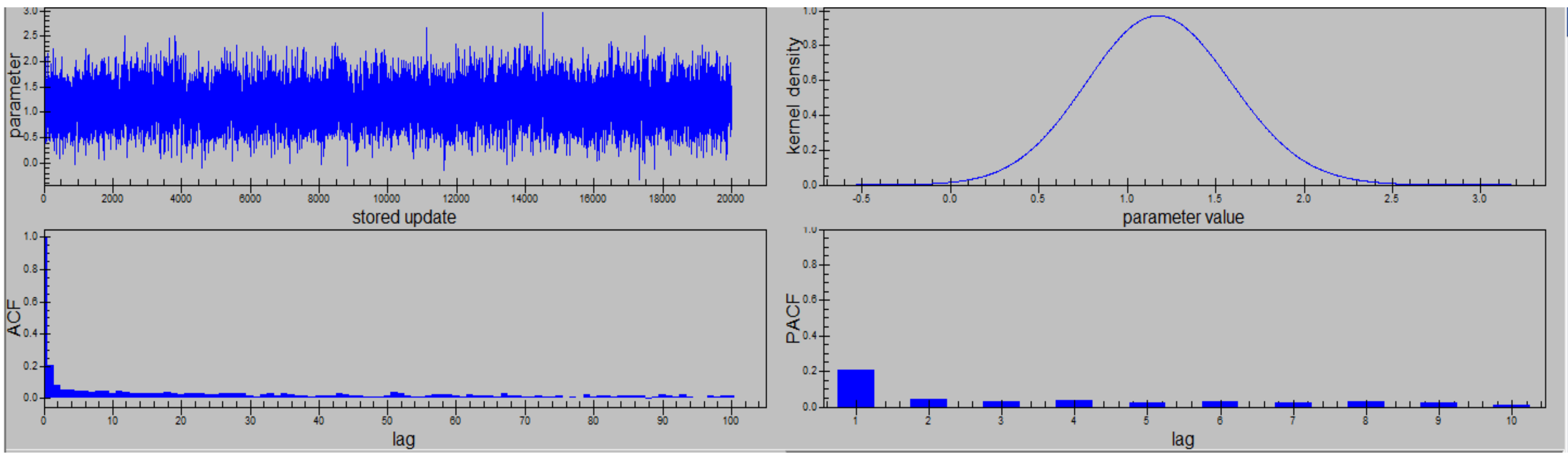
FIRST HARMFUL EVENT (COLLISION WITH NON-VEHICLES)



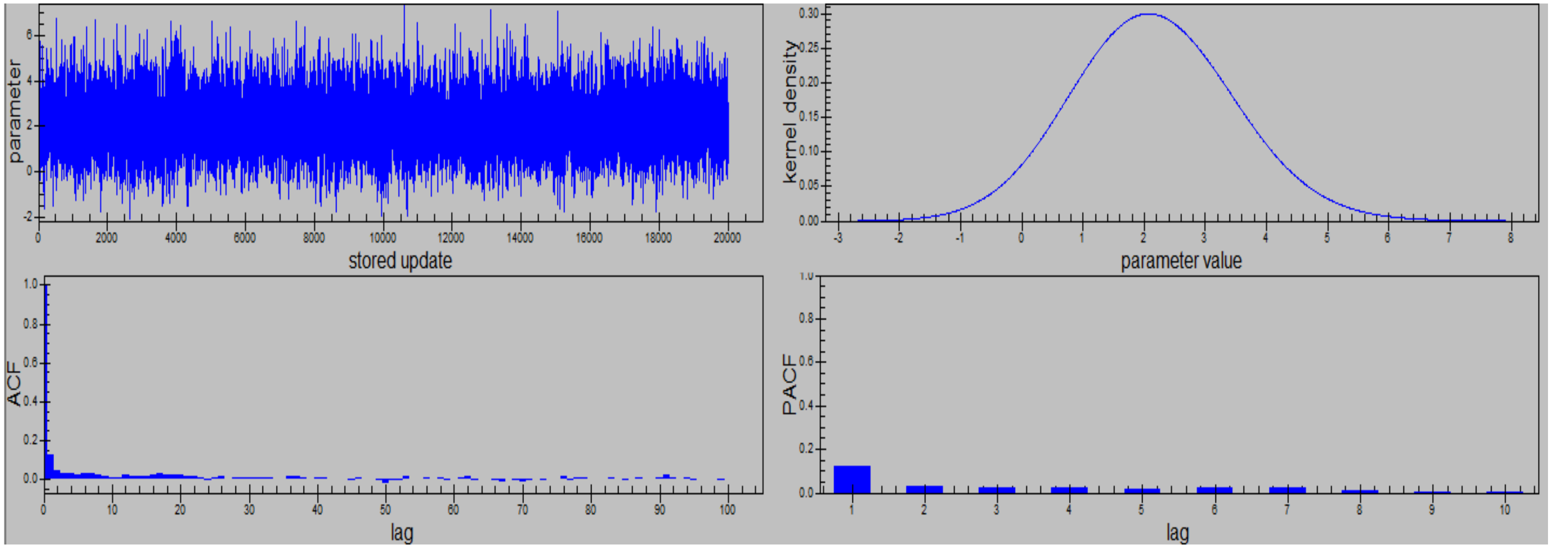
ROADWAY CONDITION AS THE CONTRIBUTIING CIRCURNTANCE



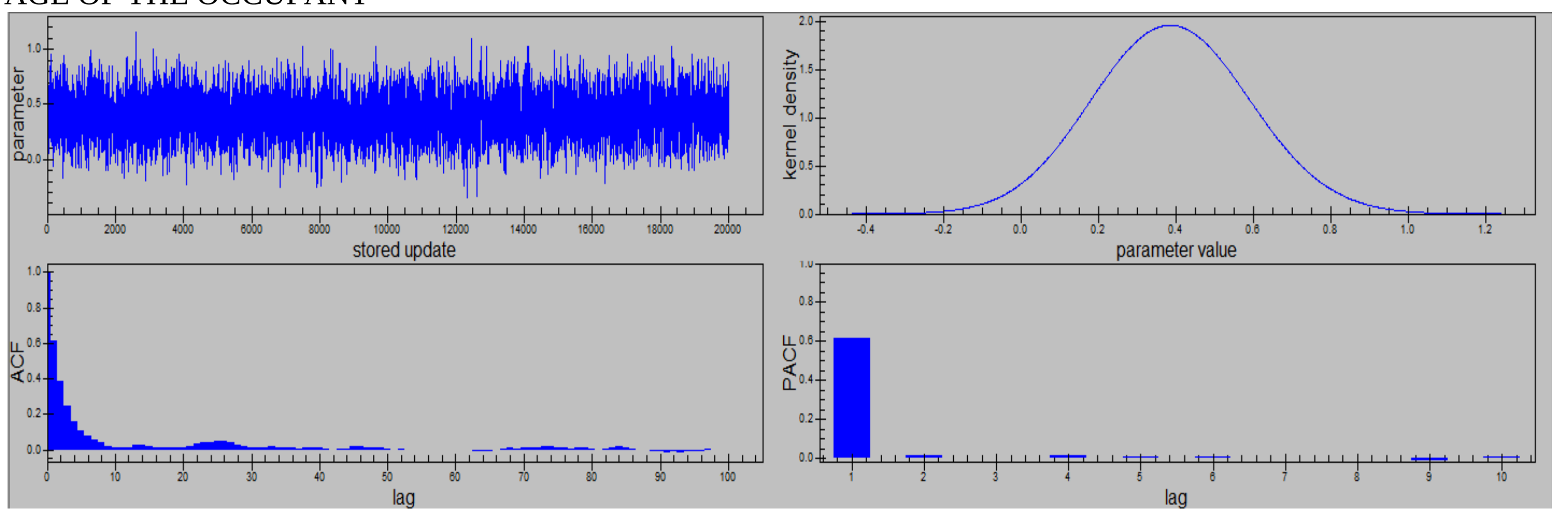
TRAPPED



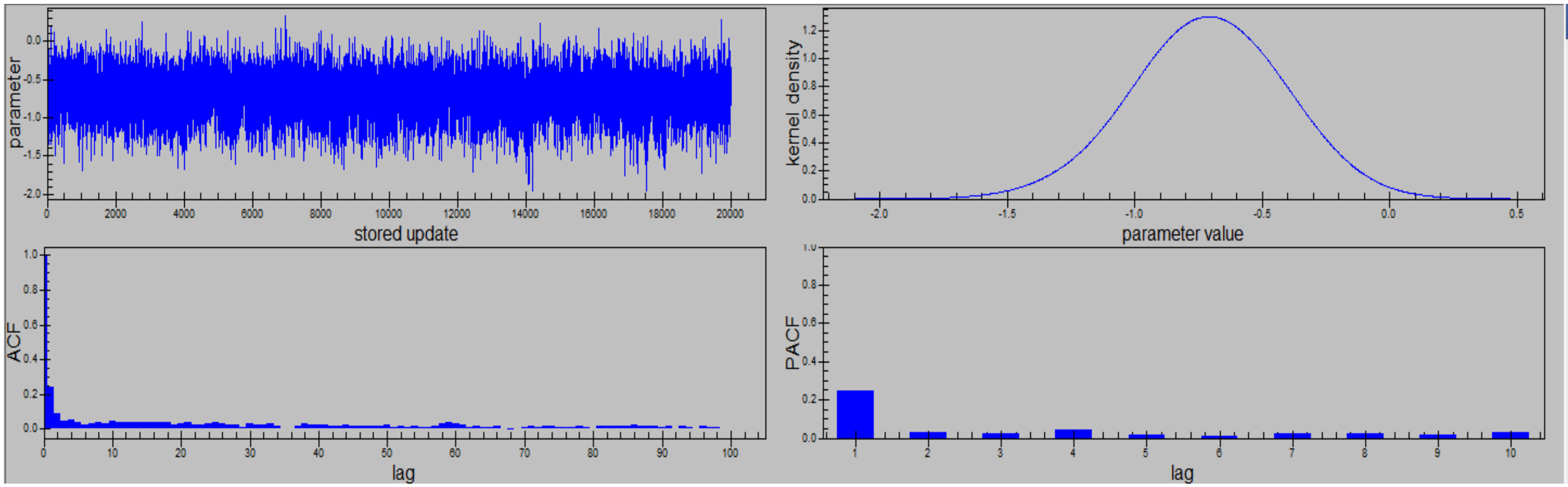
EJECTION



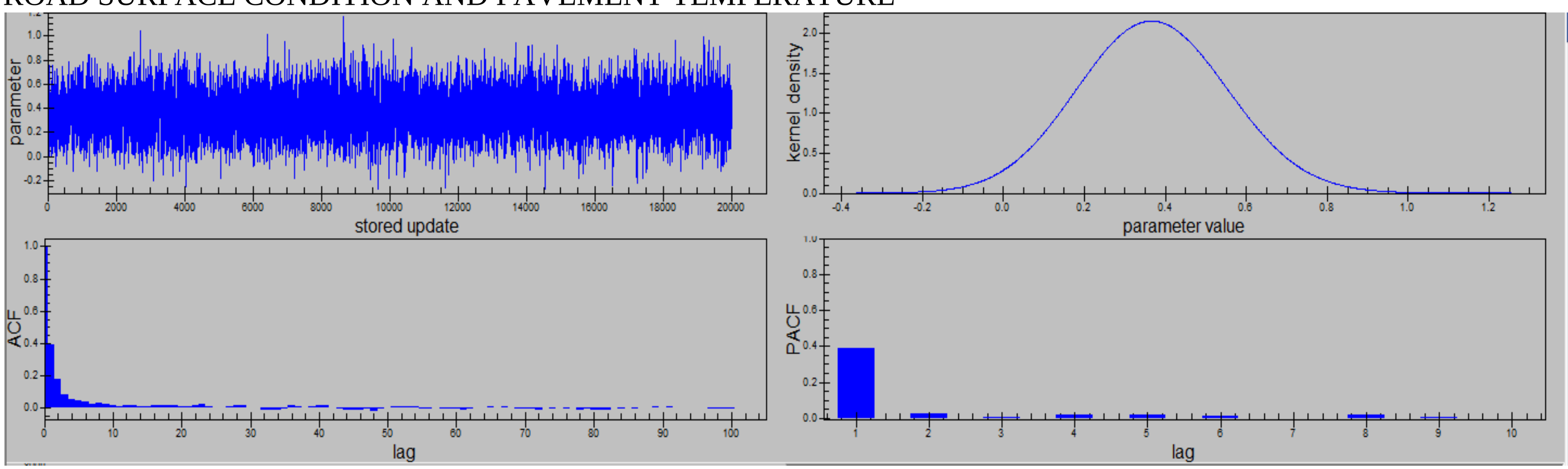
AGE OF THE OCCUPANT



ROADTYPE



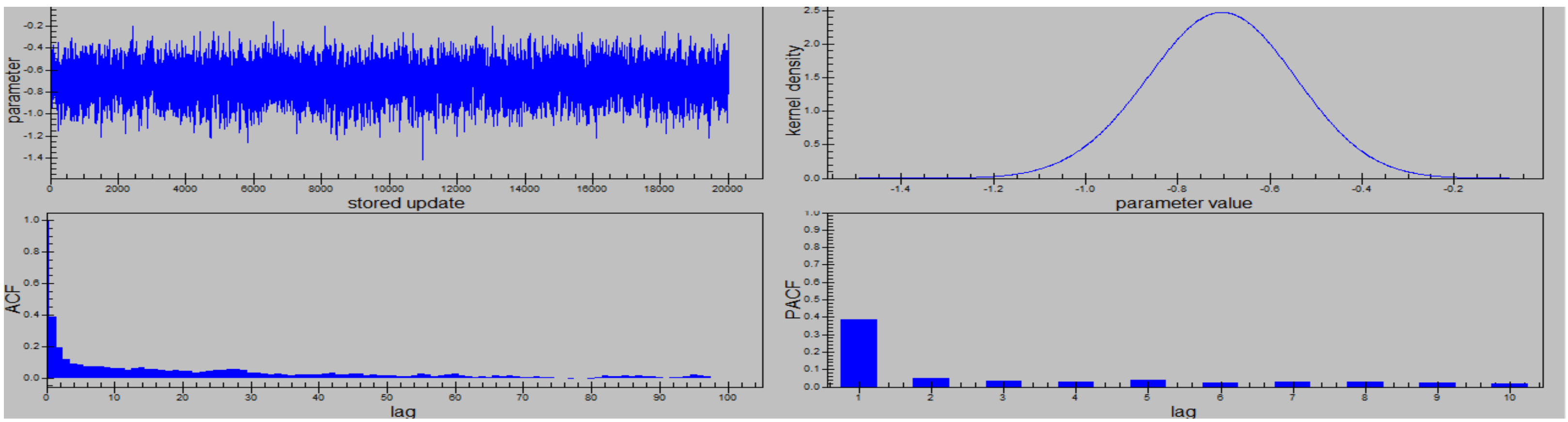
ROAD SURFACE CONDITION AND PAVEMENT TEMPERATURE



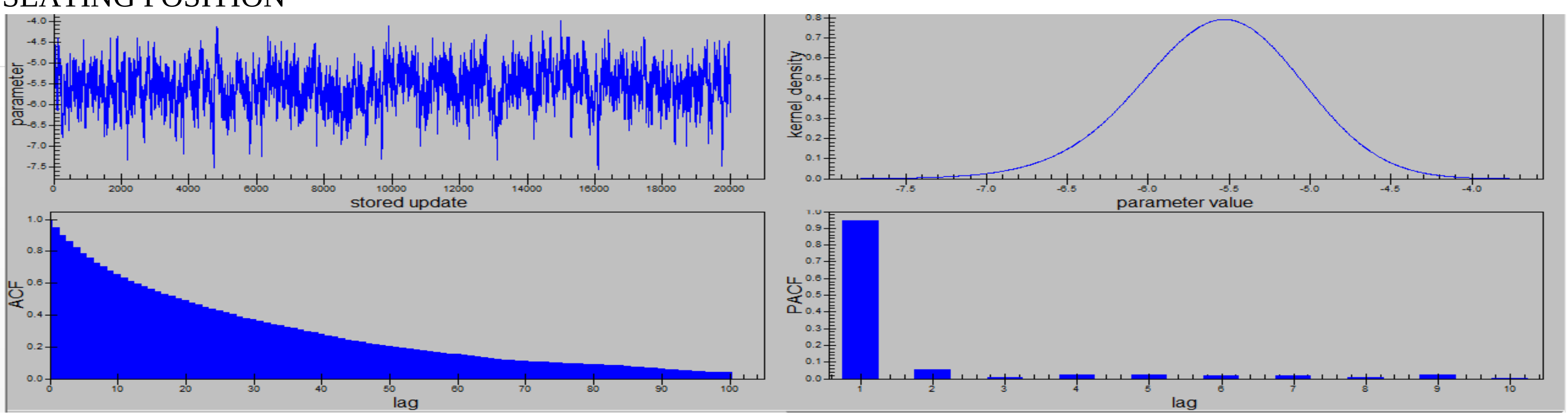
APPENDIX E

PLOTS OF DYNAMIC TRACE, KERNEL DENSITY AND ACF FOR THE PARAMETERS IN SEVERITY MODEL (NON-WEATHER-RELATED CRASHES ONLY)

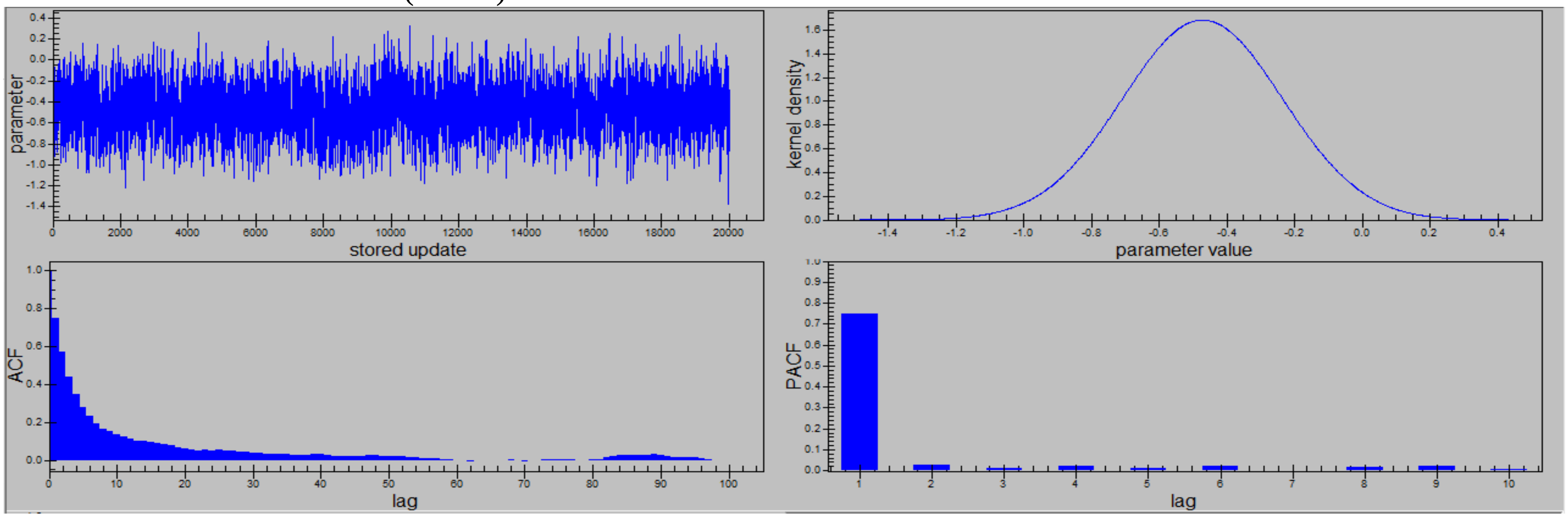
GENDER



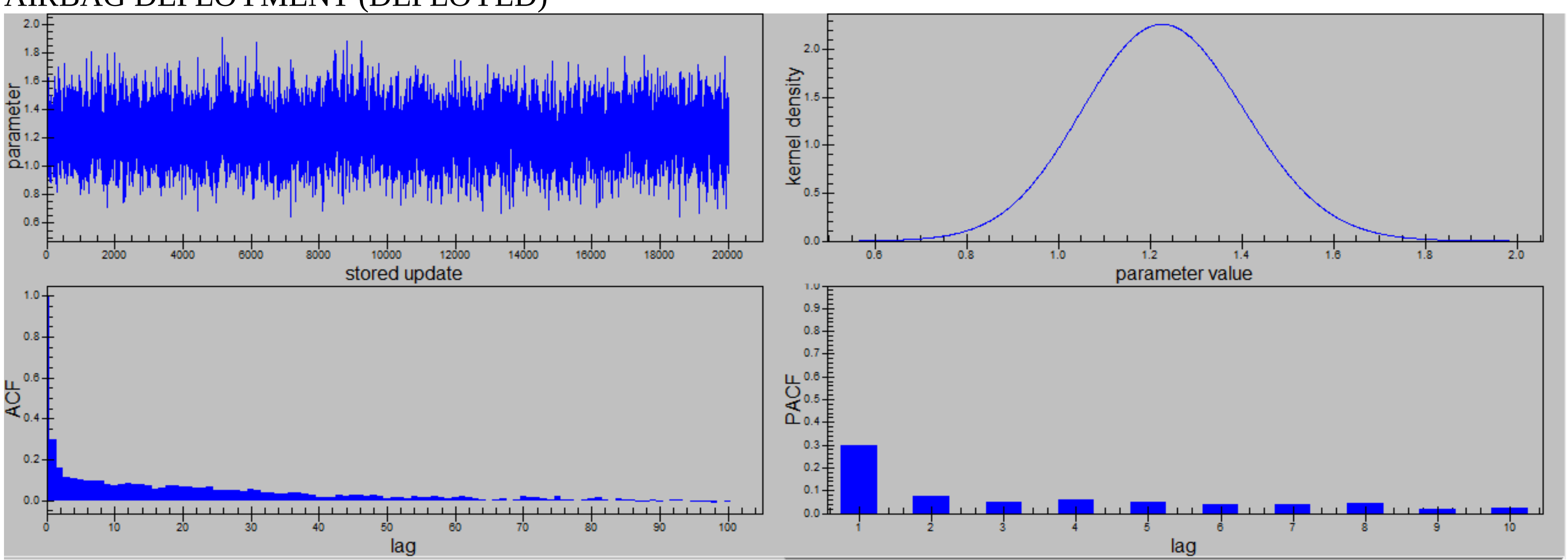
SEATING POSITION



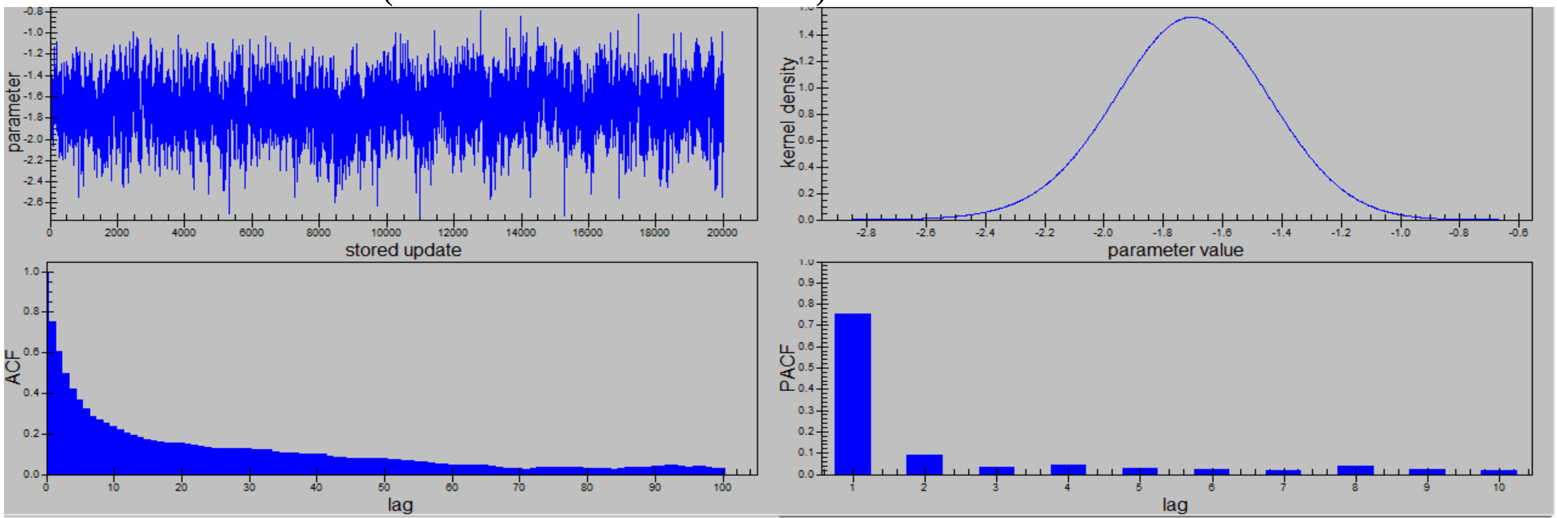
OCCUPANT PROTECTION (USED)



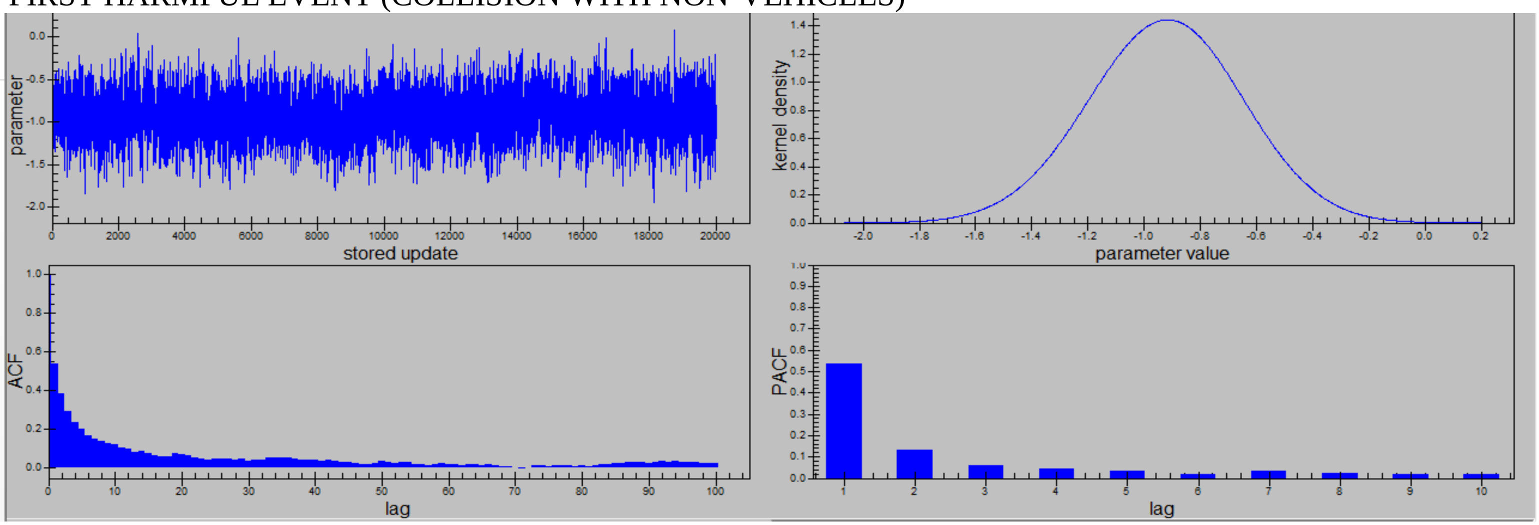
AIRBAG DEPLOYMENT (DEPLOYED)



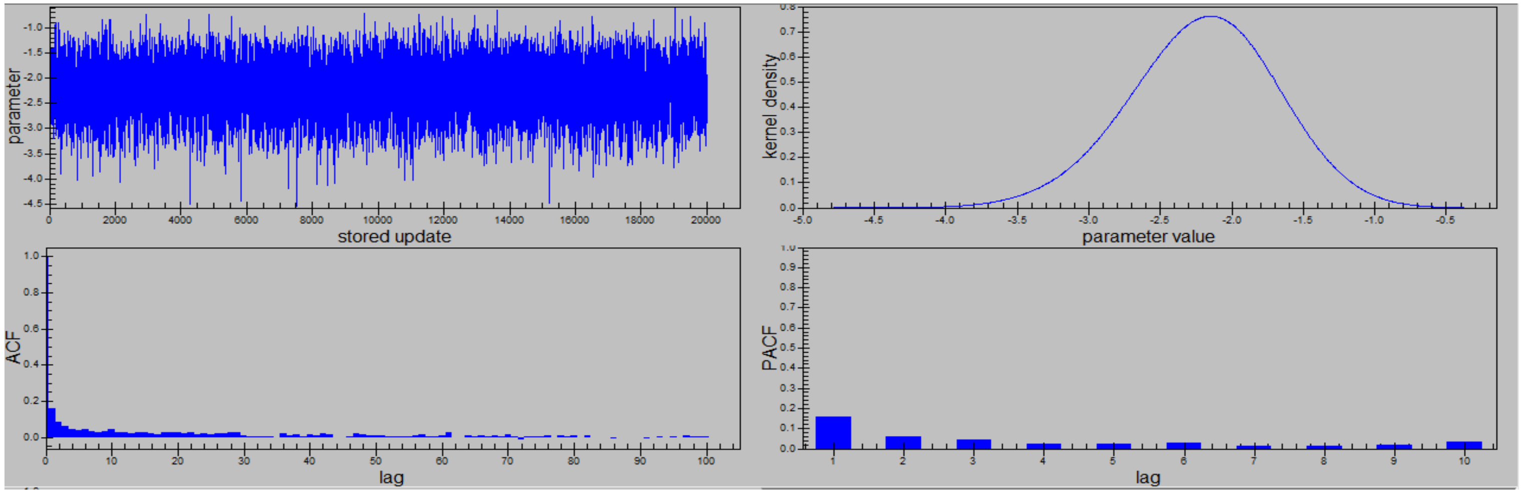
FIRST HARMFUL EVENT (COLLISION WITH VEHICLES)



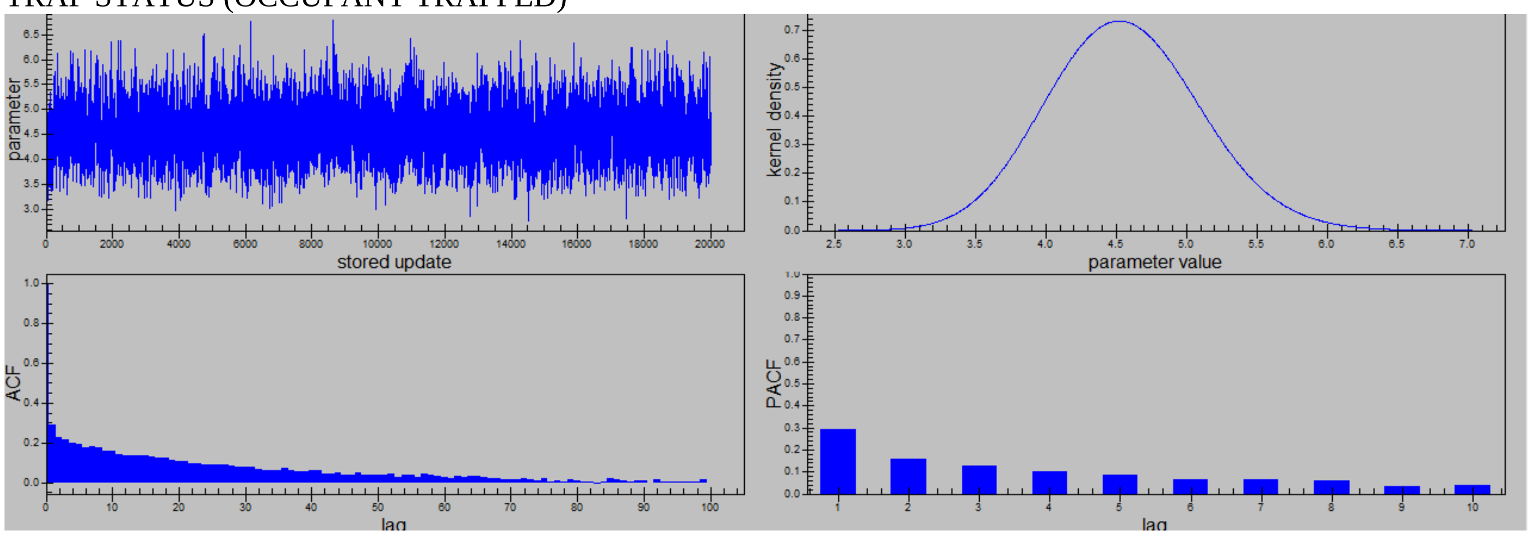
FIRST HARMFUL EVENT (COLLISION WITH NON-VEHICLES)



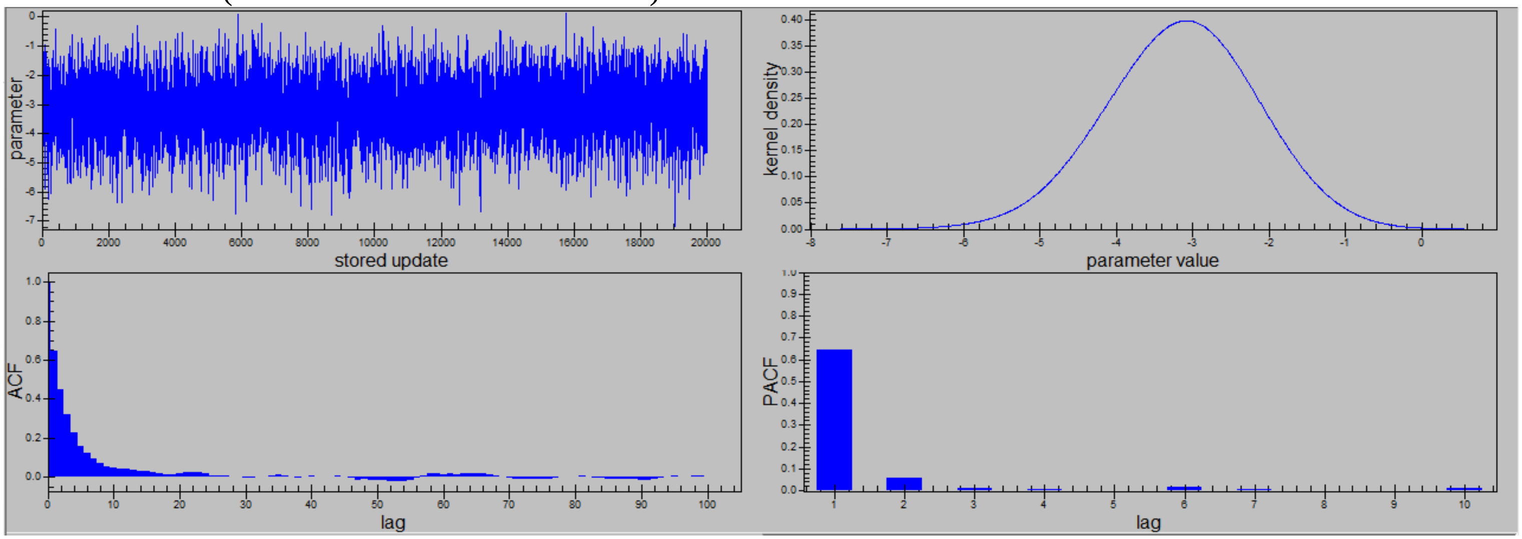
SURFACE CONDITION



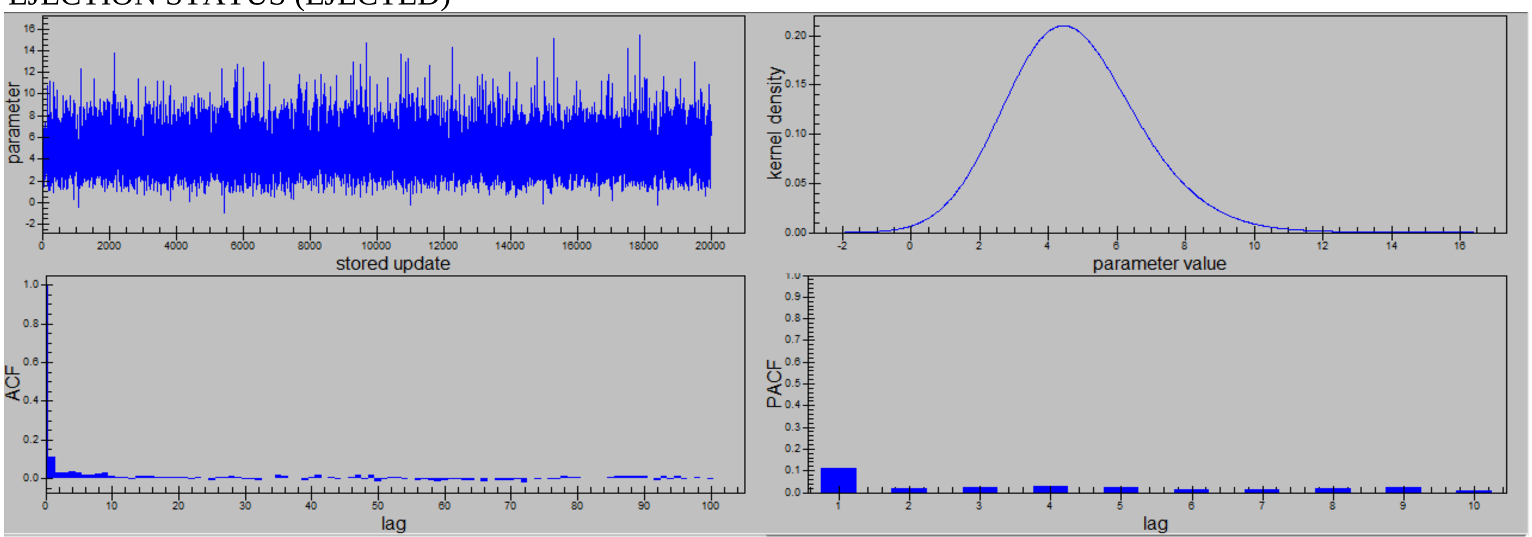
TRAP STATUS (OCCUPANT TRAPPED)



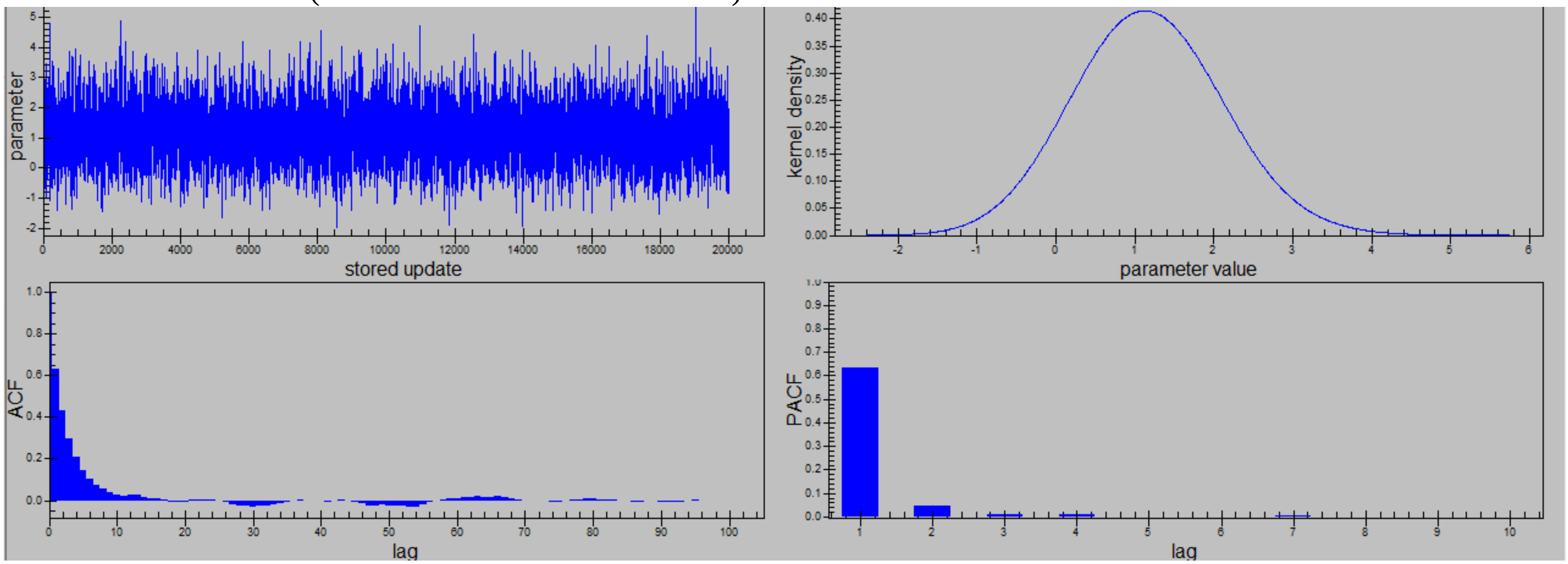
TRAP STATUS (NOT REPORTED/UNKNOWN)



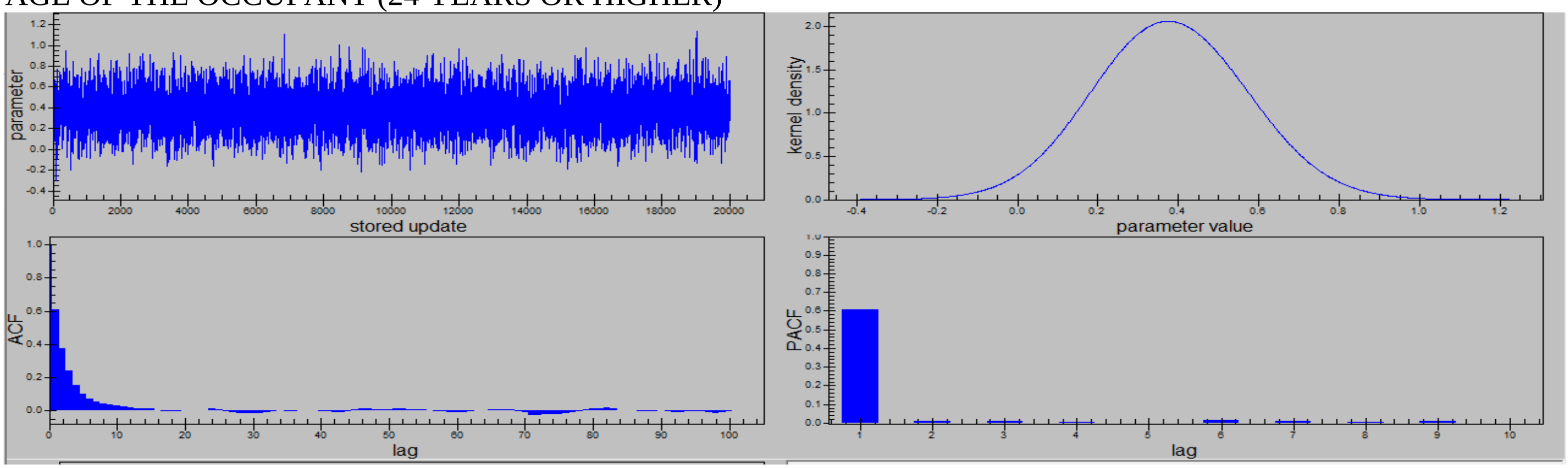
EJECTION STATUS (EJECTED)



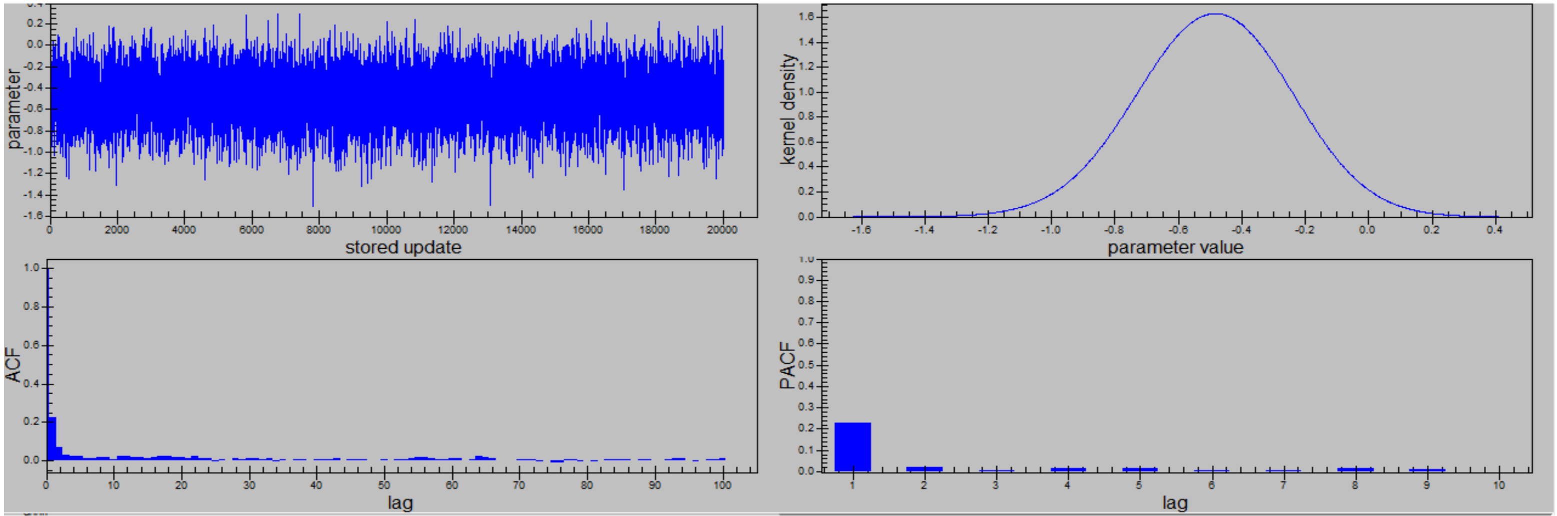
EJECTION STATUS (UNKNOWN/NOT REPORTED)



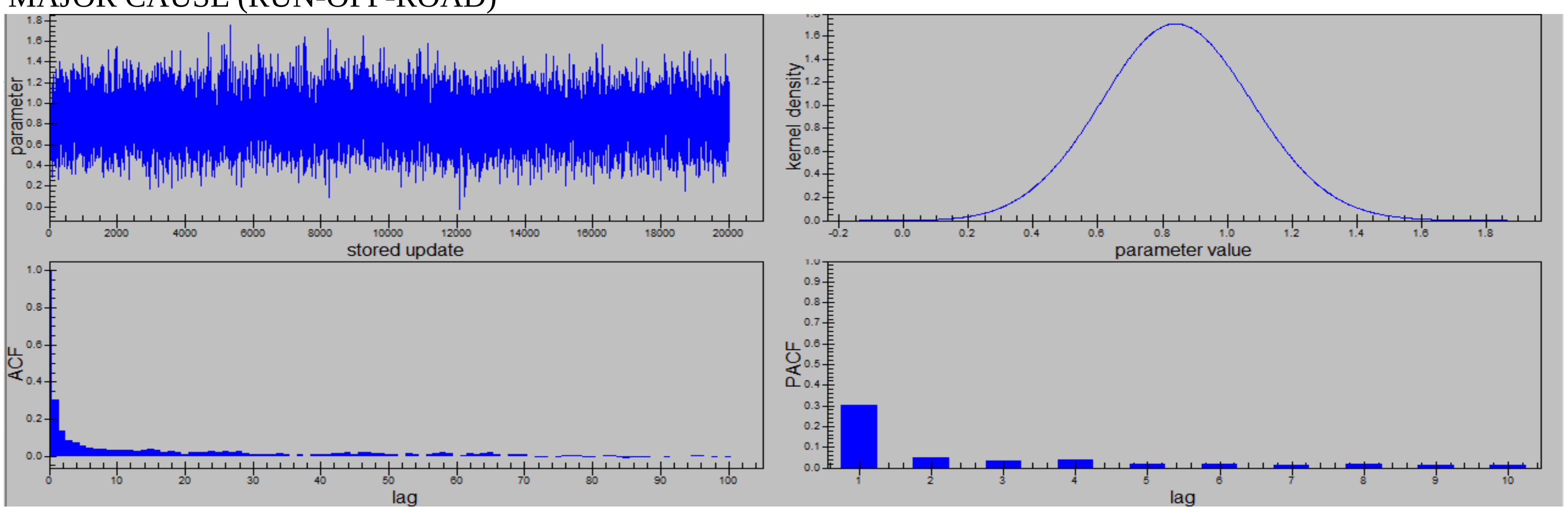
AGE OF THE OCCUPANT (24 YEARS OR HIGHER)



WEATHER CONDITION



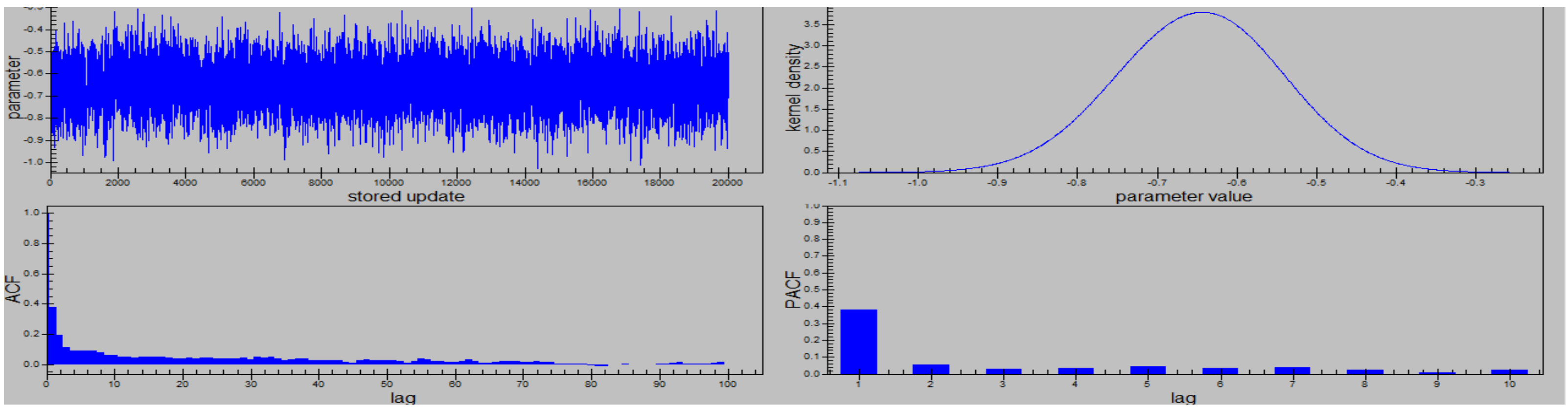
MAJOR CAUSE (RUN-OFF-ROAD)



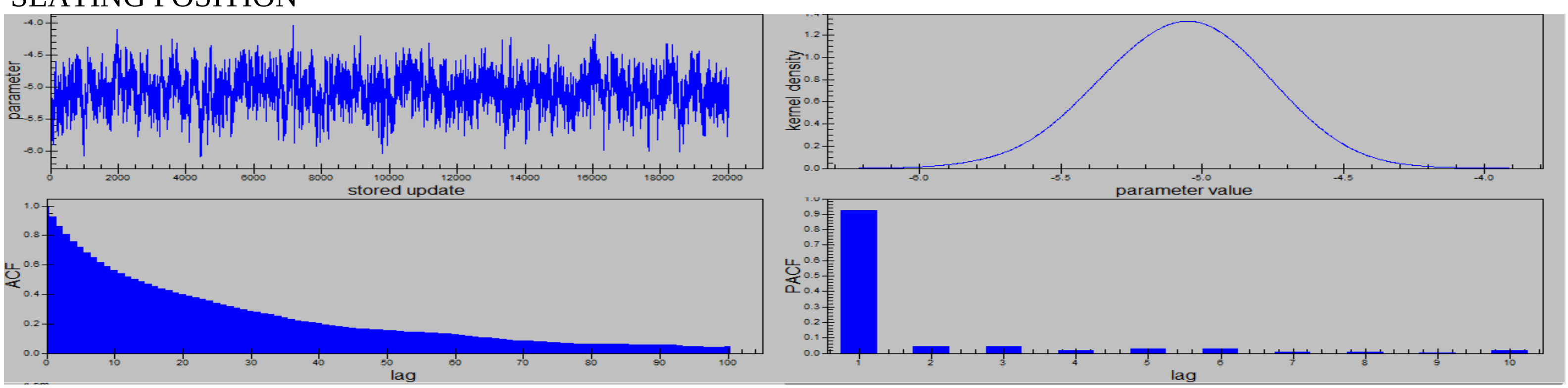
APPENDIX F

PLOTS OF DYNAMIC TRACE, KERNEL DENSITY AND ACF FOR THE PARAMETERS IN SEVERITY MODEL (ALL CRASHES)

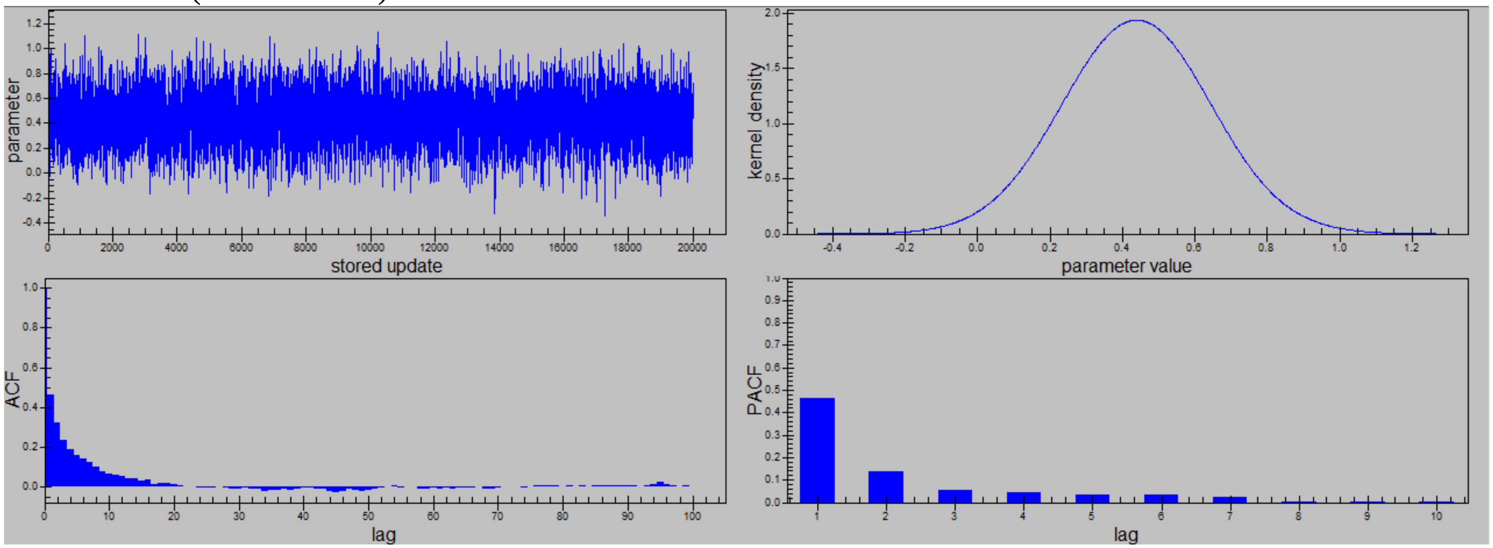
GENDER



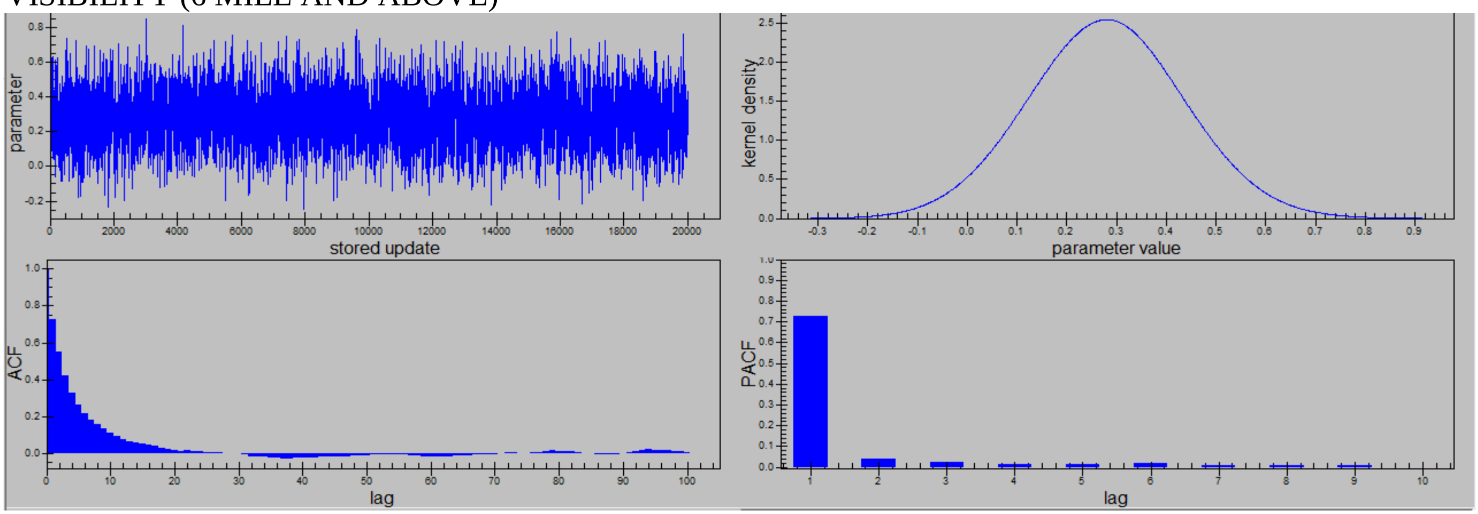
SEATING POSITION



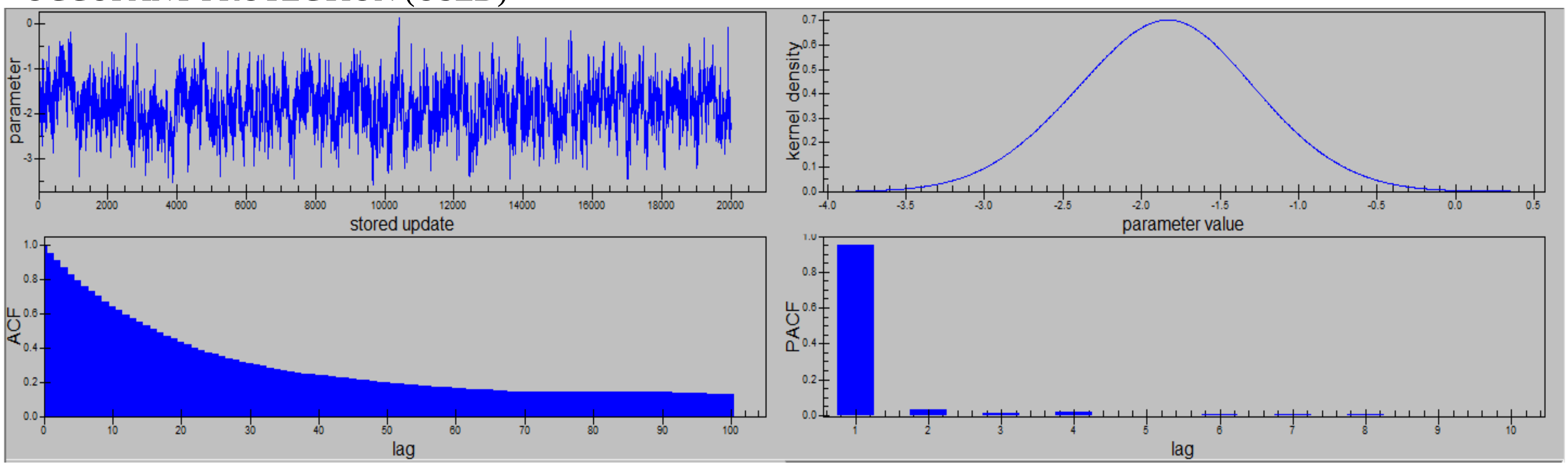
VISIBILITY (3 TO 6 MILE)



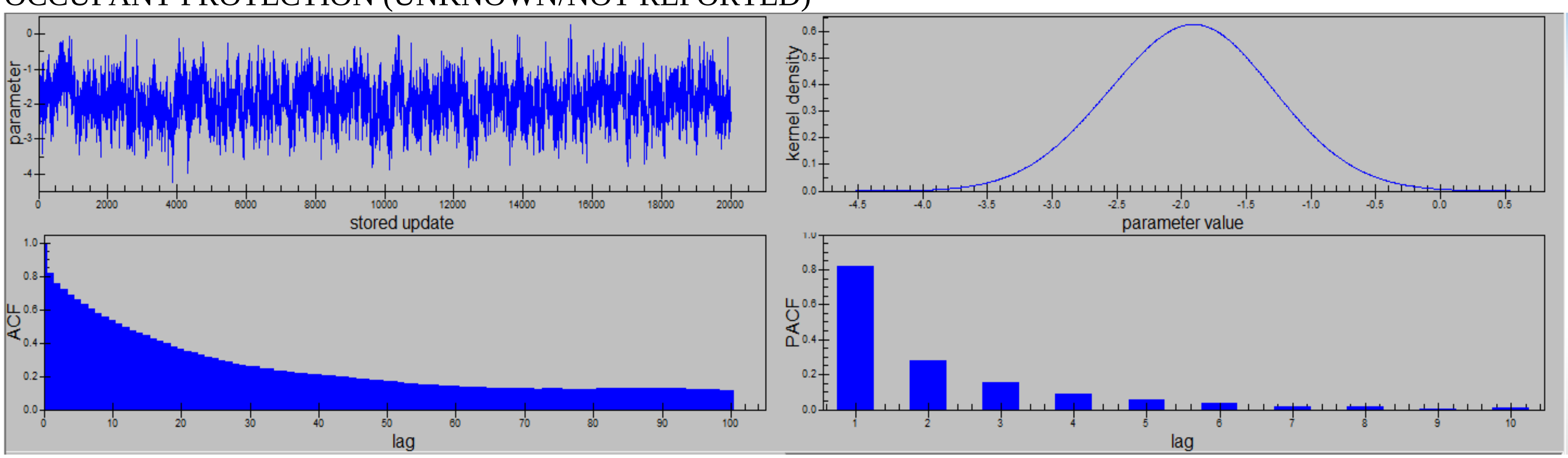
VISIBILITY (6 MILE AND ABOVE)



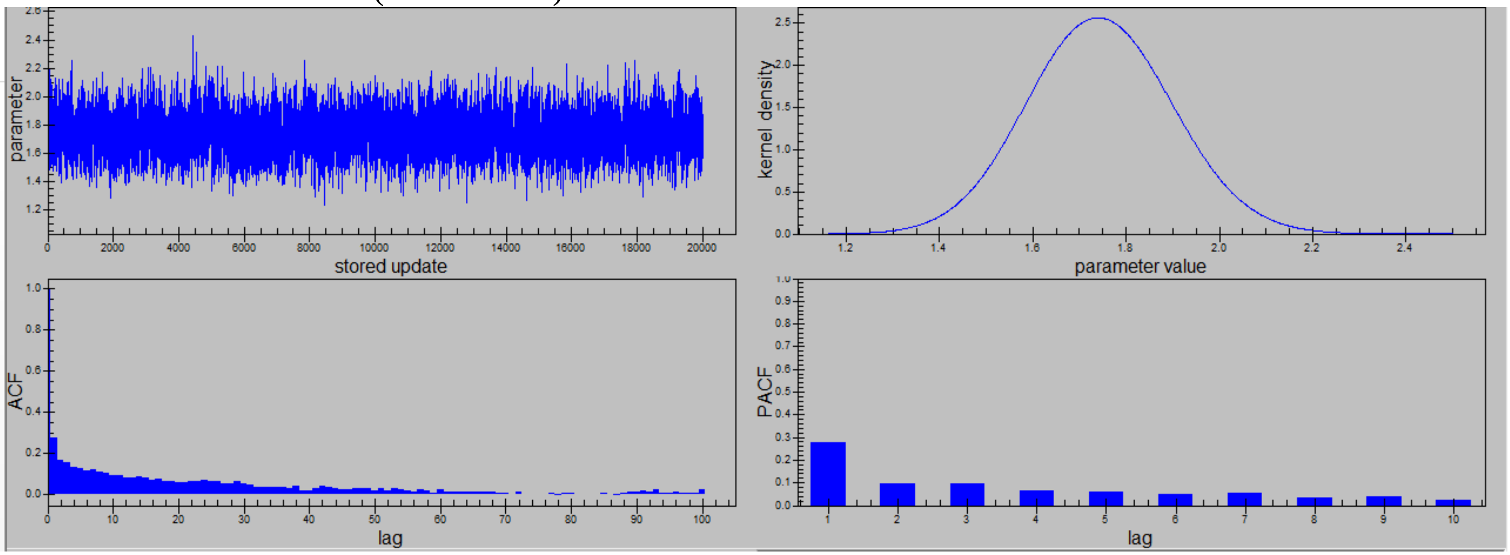
OCCUPANT PROTECTION (USED)



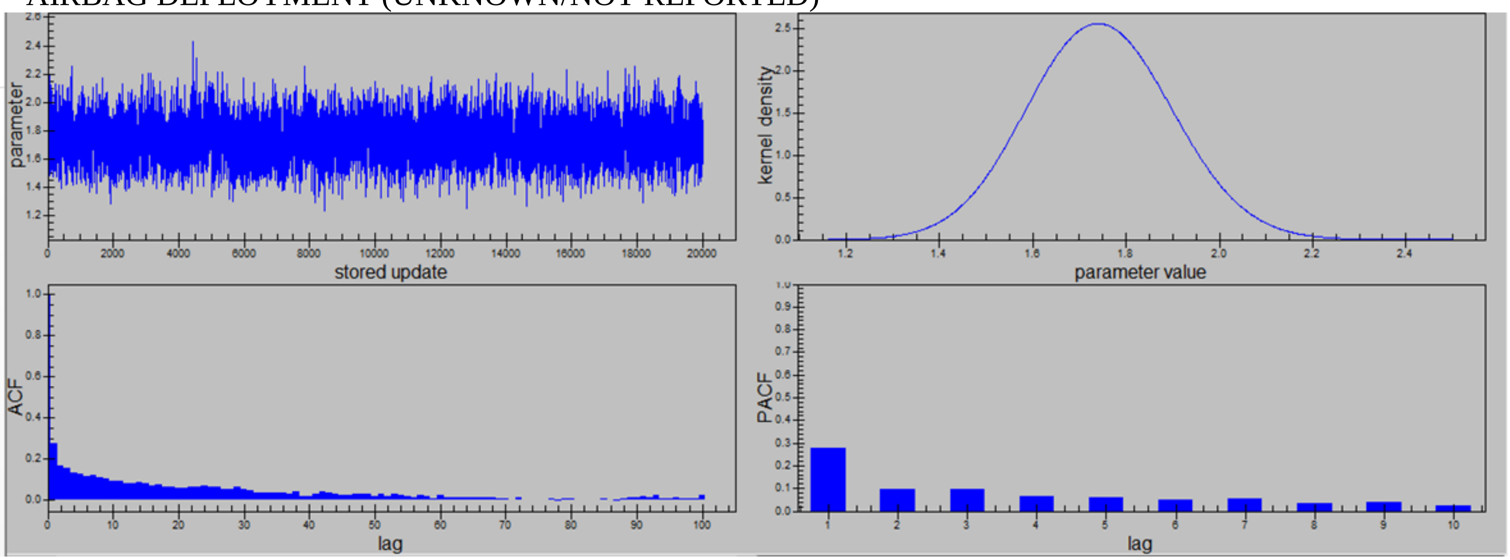
OCCUPANT PROTECTION (UNKNOWN/NOT REPORTED)



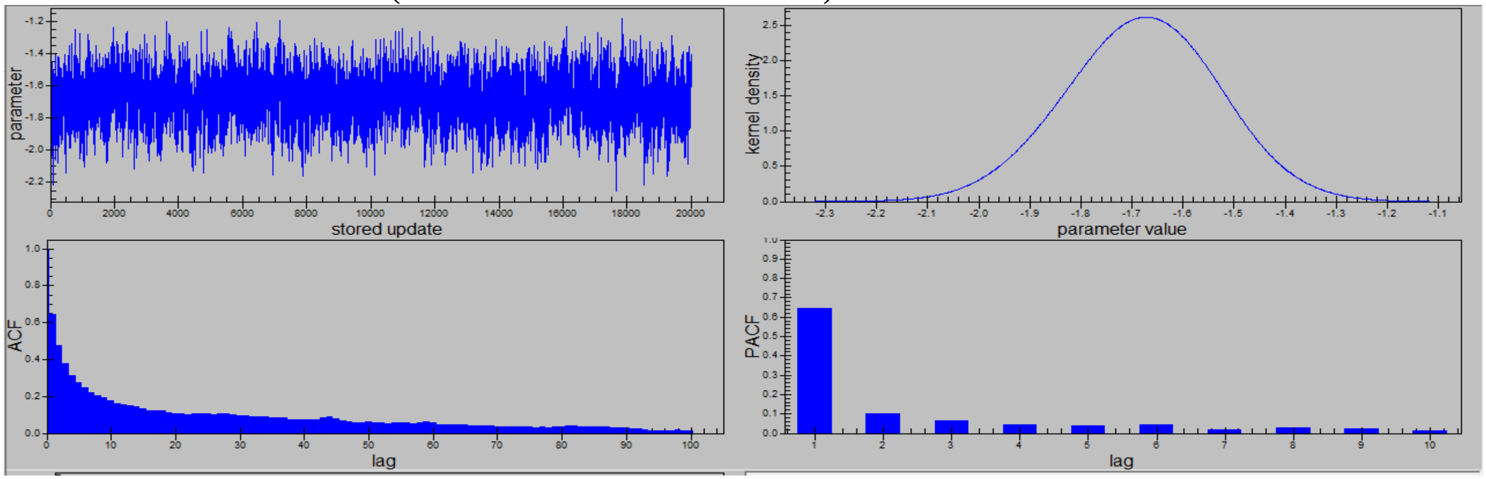
AIRBAG DEPLOYMENT (DEPLOYED)



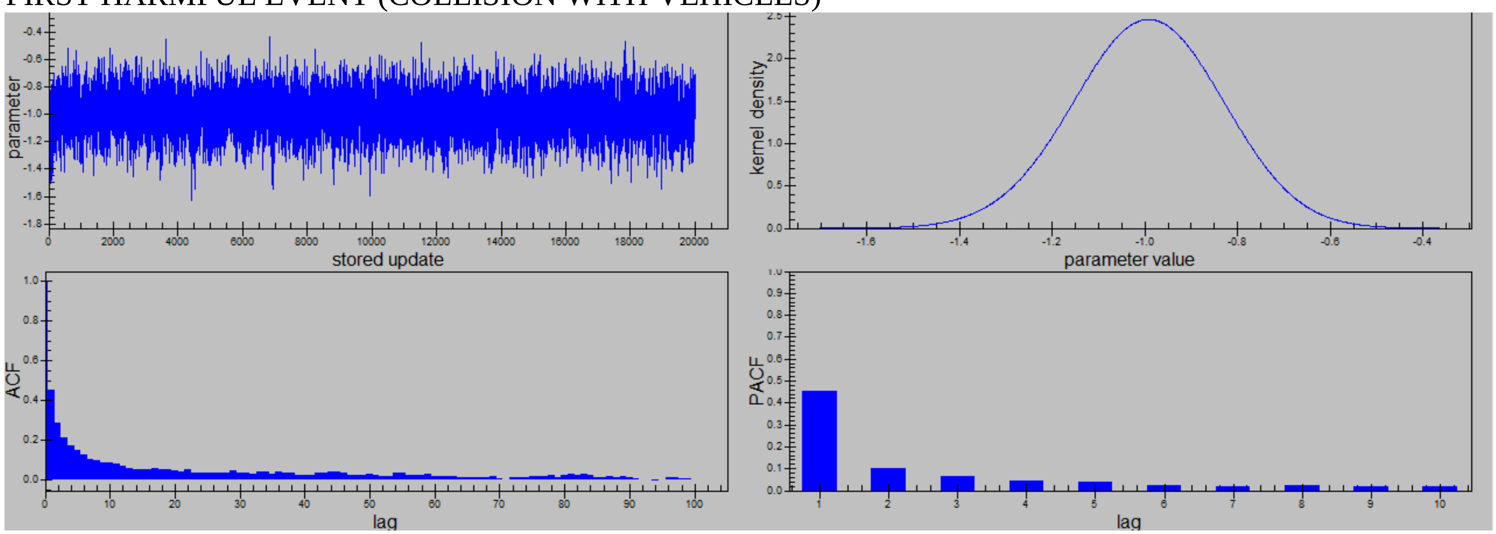
AIRBAG DEPLOYMENT (UNKNOWN/NOT REPORTED)



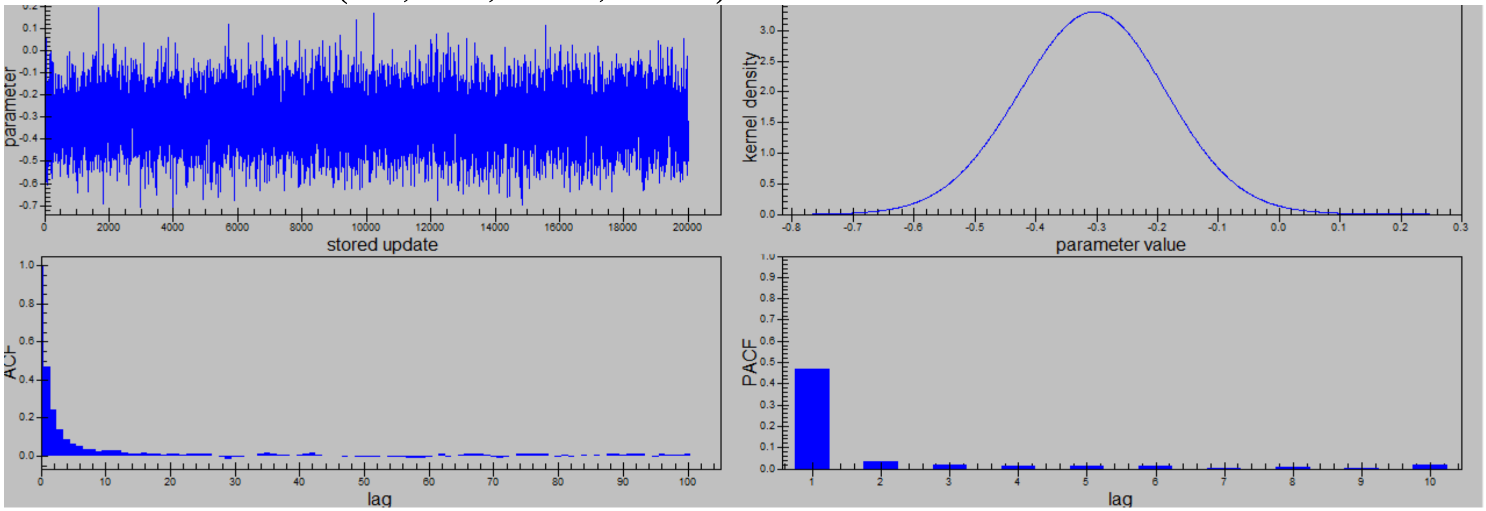
FIRST HARMFUL EVENT (COLLISION WITH VEHICLES)



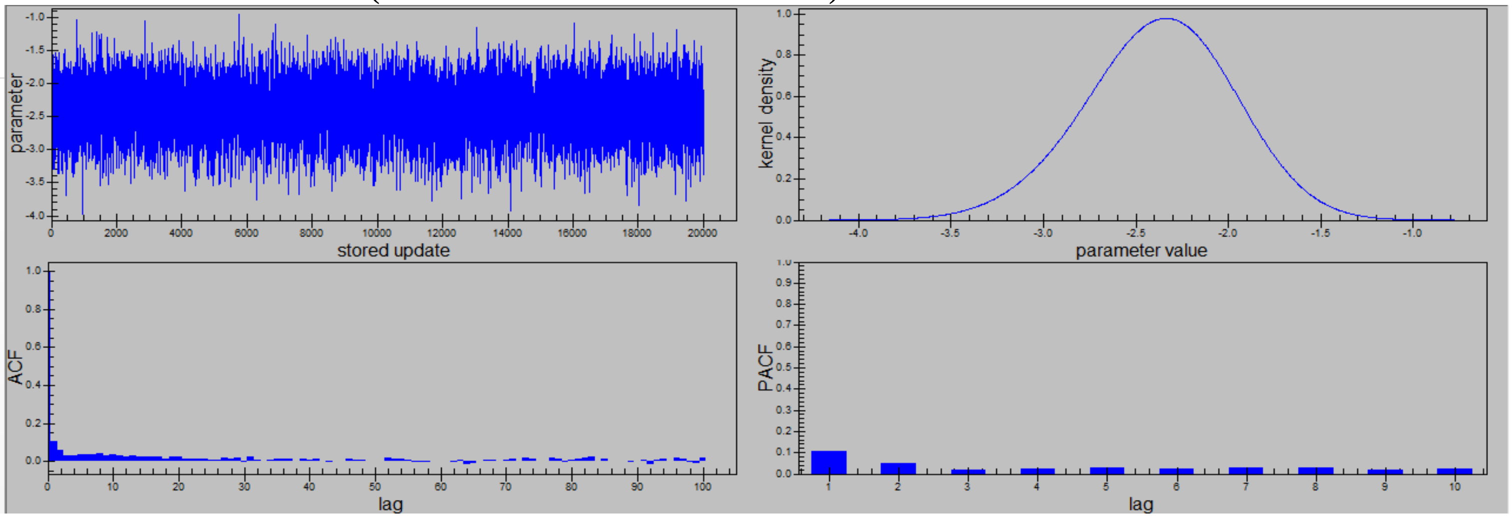
FIRST HARMFUL EVENT (COLLISION WITH VEHICLES)



SURFACE CONDTION (ICY, WET, SNOW, SLUSH)

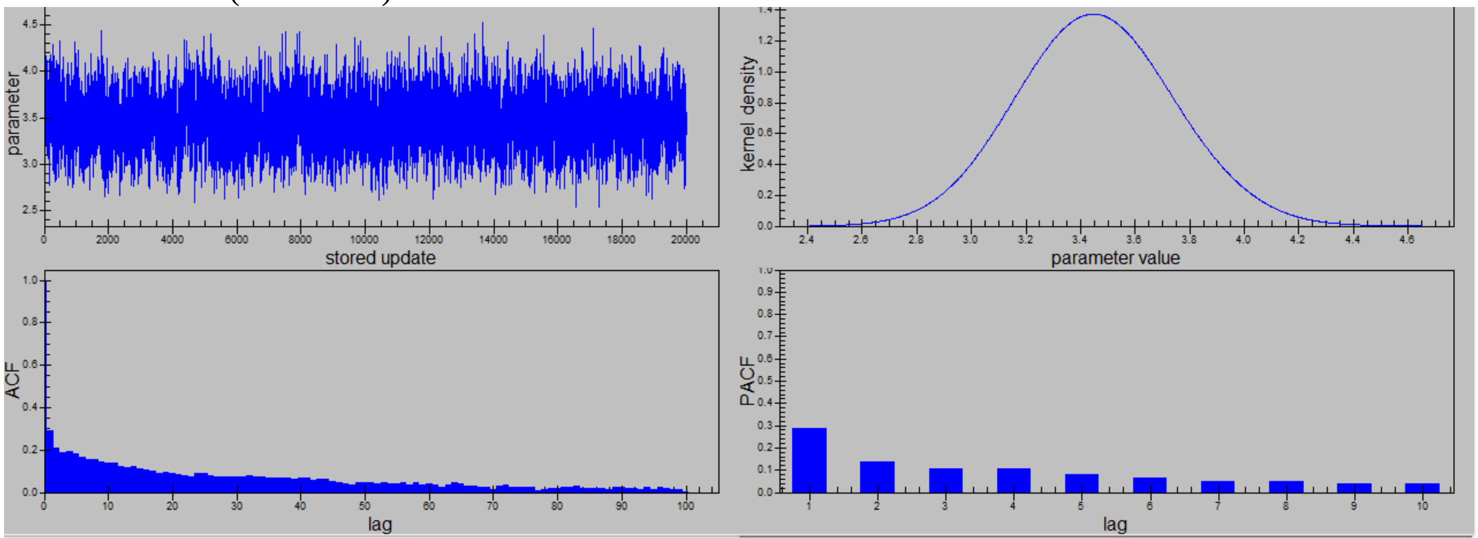


SURAFACE CONDITION (OTHERS AND NOT REPORTED)

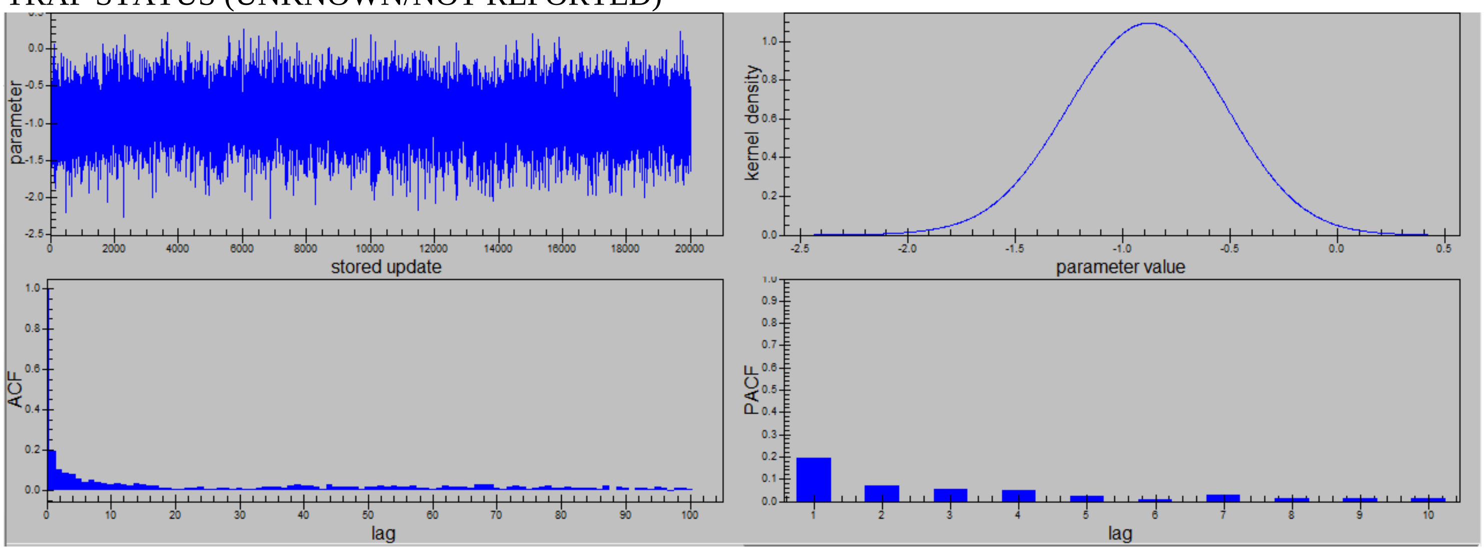


161

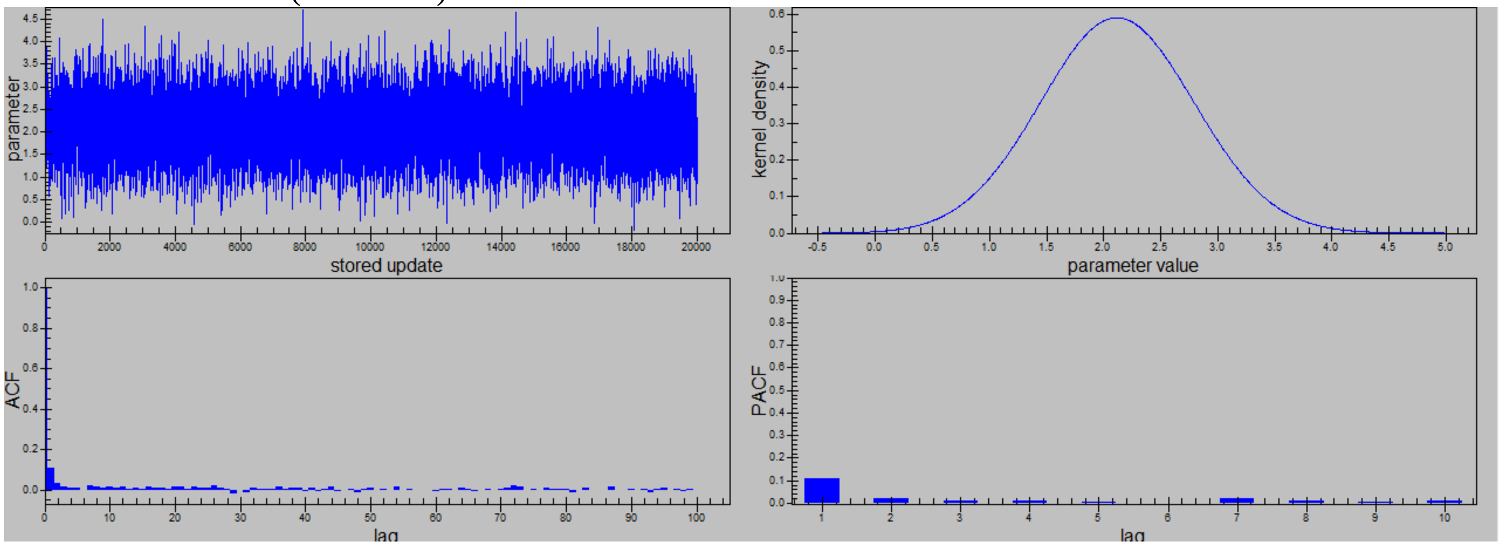
TRAP STATUS (TRAPPED)



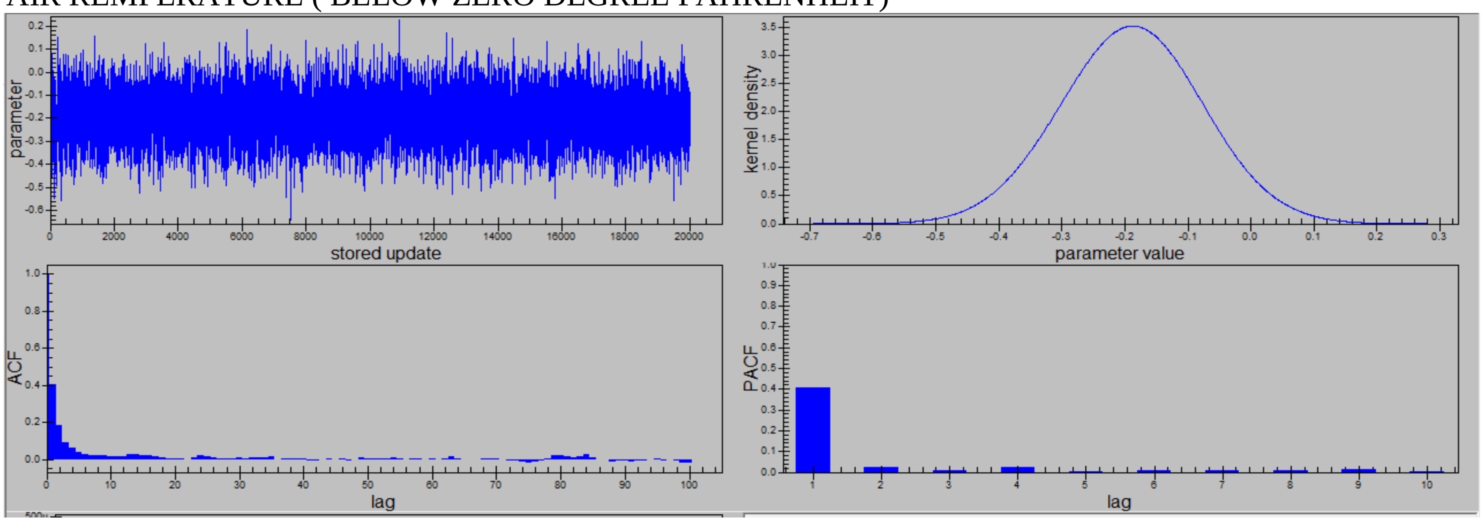
TRAP STATUS (UNKNOWN/NOT REPORTED)



EJECTION STATUS (EJECTED)



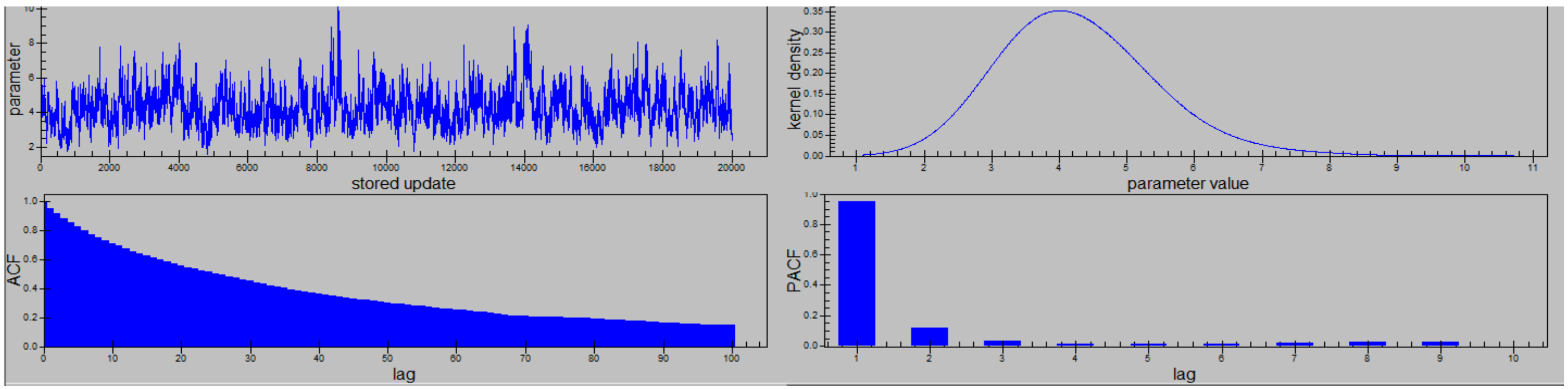
AIR REMPERATURE (BELOW ZERO DEGREE FAHRENHEIT)



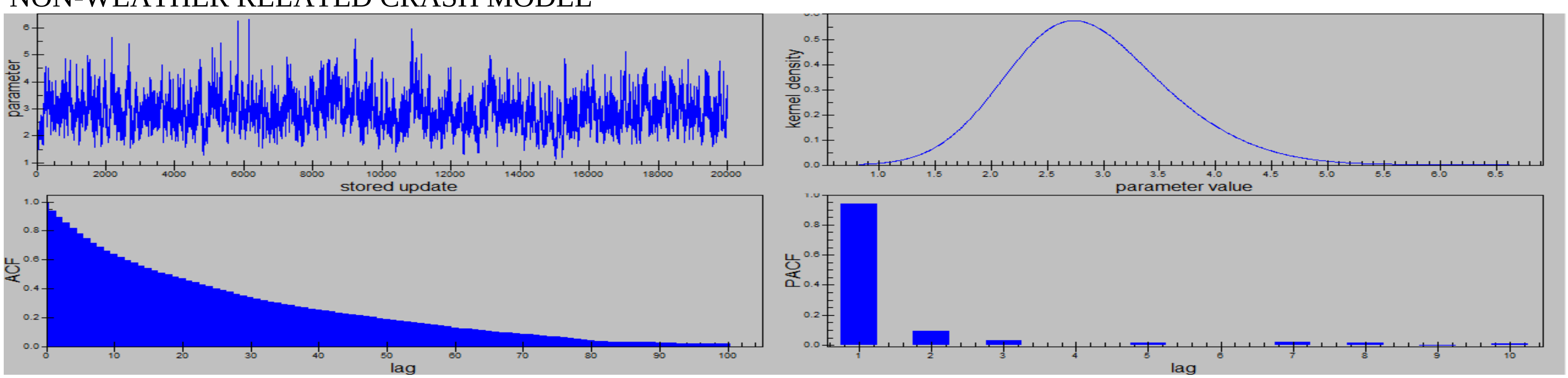
APPENDIX G

PLOTS OF DYNAMIC TRACE, KERNEL DENSITY AND ACF FOR THE RANDOM EFFECTS IN THE SEVERITY MODELS

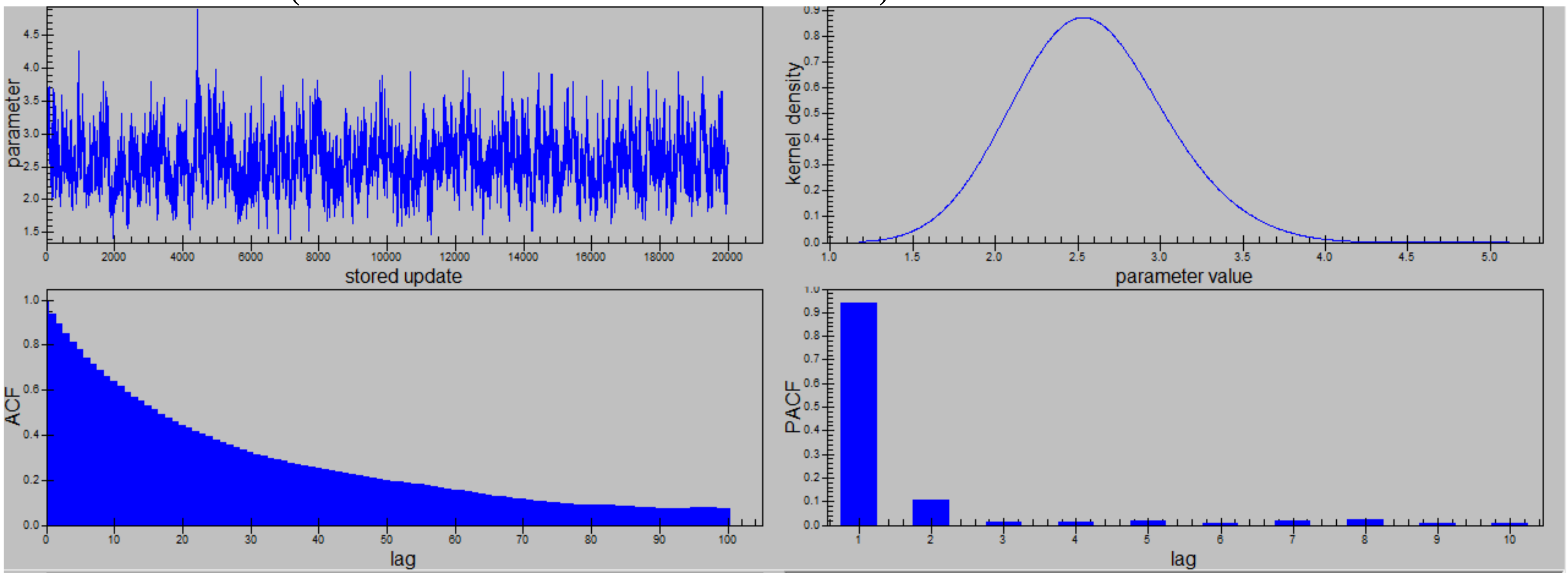
WEATHER-RLEATED CRASH MODEL



NON-WEATHER RELATED CRASH MODEL

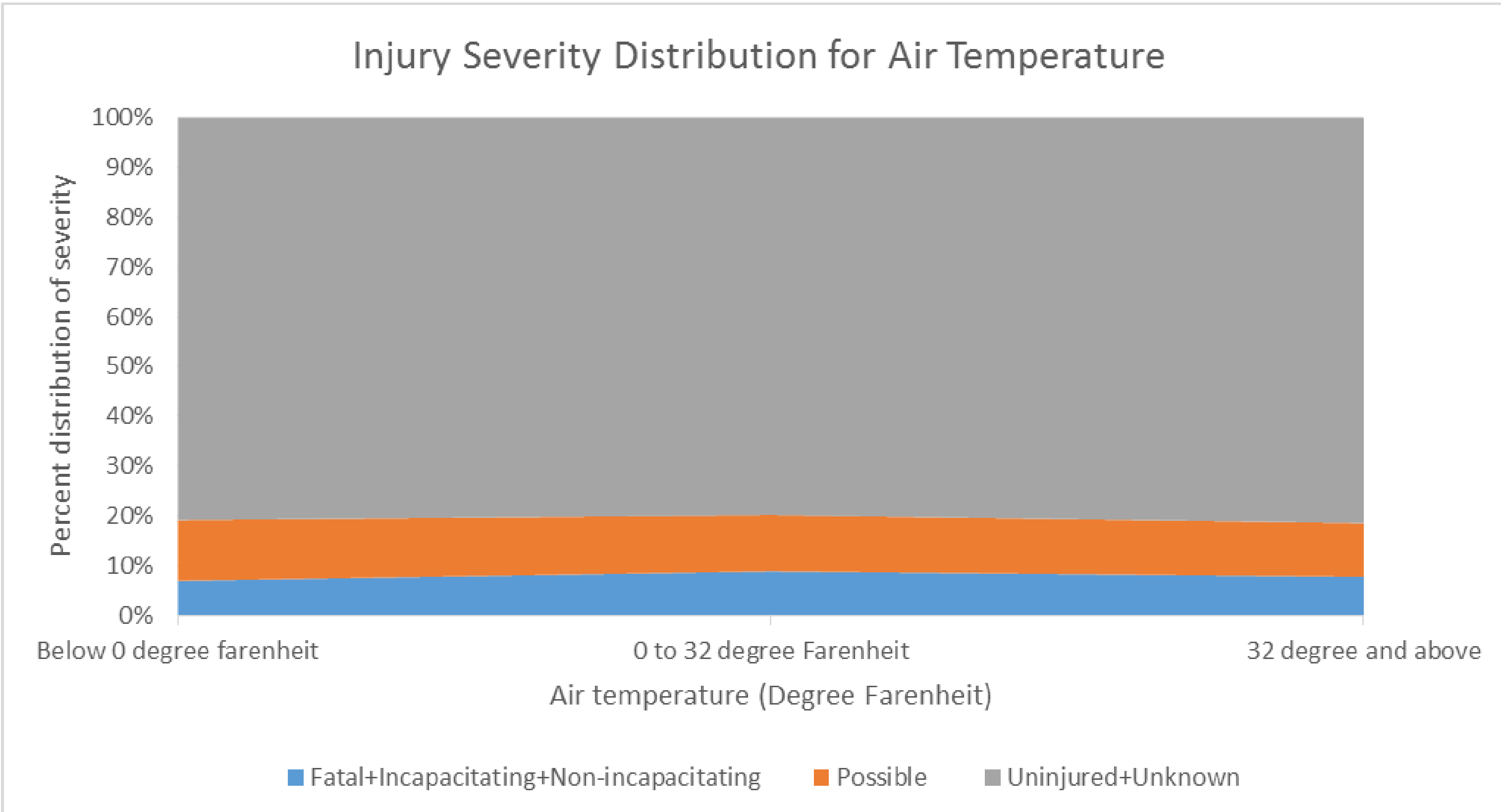


ALL CRASH MODEL (WEATHER AND NON-WEATHER RELATED)

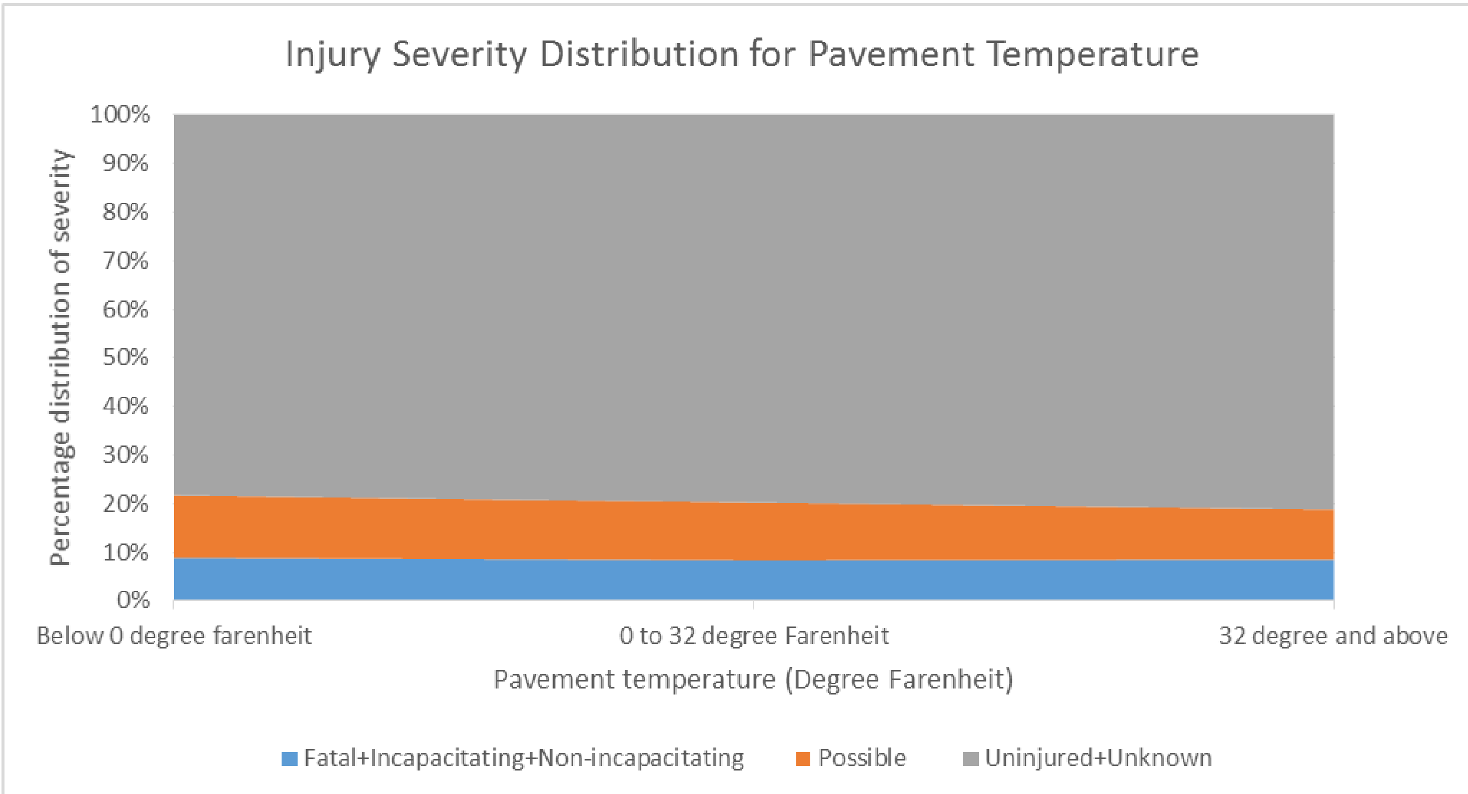


APPENDIX H

DISTRIBUTION OF INJURY SEVERITY BY WEATHER RELATED VARIABLES FOR ALL CRASHES



DISTRIBUTION OF INJURY SEVERITY BY WEATHER RELATED VARIABLES FOR ALL CRASHES (CONTIN'D)



DISTRIBUTION OF INJURY SEVERITY BY WEATHER RELATED VARIABLES FOR ALL CRASHES (CONTIN'D)