Southern Plains Transportation Center

Crash Severity Formulation and Analysis under Extreme Weather Conditions

Rafiqul A. Tarefder, Ph.D. &
Guohui Zhang, Ph.D.

SPTC14.1-86

Southern Plains Transportation Center
201 Stephenson Parkway, Suite 4200
The University of Oklahoma
Norman, Oklahoma 73019

DISCLAIMER

Technical Report Documentation Sheet

1. REPORT NO. SPTC 14.1-86	2. GOVERNMENT ACCESSION NO.	3. RECIPIENTS CATALOG NO.
4. TITLE AND SUBTITLE Crash Severity Formulation and Analysis under Extreme Weather Conditions		5. REPORT DATE October 31, 2018
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Rafiqul A. Tarefder and Guohui Zhang		8. PERFORMING ORGANIZATION REPORT University of New Mexico
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Civil Engineering University of New Mexico 210 University Blvd NE MSC01 1070, Albuquerque, NM 87131-0001		10. WORK UNIT NO.
		11. CONTRACT OR GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Southern Plains Transportation Center 201 Stephenson Pkwy, Suite 4200 The University of Oklahoma Norman, OK 73019		13. TYPE OF REPORT AND PERIOD COVERED Final July 2016 – July 2018
		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES Matching fund provided by New Mexico Department of Transportation (NMDOT)		

ABSTRACT

This study focuses on vehicle crush with respect to extreme weather condition in the state of New Mexico. In the southwest region, unique weather extremes, including enormously high temperature, strong wind, flash flood, fog, dust, snow, etc., have resulted in more frequent and serious traffic crashes. For example, 501 traffic crashes were reported under extremely windy conditions in New Mexico in 2011. Although they only accounted for 1.2% of total crashes but result in about 3.4% of total fatalities. Under foggy conditions, fatal crashes account for 2.5% of total crashes, which is much higher than, 0.7%, the proportion of fatal crashes to total crashes under regular weather conditions in New Mexico in 2011. These statistics dramatically illustrate that traffic crashes impacted by weather extremes are more serious than those under regular weather conditions. Currents study attempts to develop a database of vehicle crash with respect to extreme weather condition. Statistical models are developed and analyzed to find a correlation of vehicle crash with extreme weather conditions.

17. KEY WORDS Concrete, Mechanistic-Empirical, Rigid pavement, Thermal expansion		18. DISTRIBUTION STATEMENT No restrictions. This publication is available at www.sptc.org and from the NTIS.	
19. SECURITY CLASSIF. (OF THIS REPORT) Unclassified	20. SECURITY CLASSIF. (OF THIS PAGE) Unclassified	21. NO. OF PAGES 28	22. PRICE

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic	m^3
	meters NOTE: volumes greater than 1000 L shall be			
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Table of Contents

List of Figures

List of Tables

EXECUTIVE SUMMARY

Weather conditions have tremendous impacts on traffic crash frequencies and severities. More than 1,312,000 crashes occurred under adverse weather conditions accounting for 23% of total crashes in the U.S. in 2012. On average, weather-related traffic crashes result in 6,250 fatalities and over 480,000 injuries nation-wide each year. Correspondingly, the annual economic loss of weather-related crashes is about $22 billion, which is a conservative estimate without considering unreported crashes. In the southwest region, unique weather extremes, including enormously high temperature, strong wind, flash flood, fog, dust, snow, etc., have resulted in more frequent and serious traffic crashes. For example, 501 traffic crashes were reported under extremely windy conditions in New Mexico in 2011. Although they only accounted for 1.2% of total crashes but result in about 3.4% of total fatalities. Under foggy conditions, fatal crashes account for 2.5% of total crashes, which is much higher than, 0.7%, the proportion of fatal crashes to total crashes under regular weather conditions in New Mexico in 2011. These statistics dramatically illustrate that traffic crashes impacted by weather extremes are more serious than those under regular weather conditions. Substantial research efforts are needed to better understand significant causal factors and their impacts on crash severities under extreme, adverse weather conditions in order to develop effective countermeasures and proper policies to minimize weather-related risks to traffic safety. This research project is proposed to develop a new multinomial Logit model-Bayesian network hybrid approach to discover the underlying patterns behind crash data and identify significant contributing attributes related to severe crashes impacted by weather extremes in the southwest region. A comprehensive database has been designed and developed with region-wide crash data, roadway geometric data, weather condition data, and traffic data. The research findings are helpful for transportation agencies to develop cost-effective countermeasures to mitigate crash severities under extreme weather conditions and minimize the weather-related risks to traffic safety in the southwest region.

CHAPTER 1
INTRODUCTION

1.1 PROBLEM STATEMENT

Weather conditions have tremendous impacts on traffic crash frequencies and severities. More than 1,312,000 crashes occurred under adverse weather conditions accounting for 23% of total crashes in the U.S. in 2012 (Federal Highway Administration, 2013). On average, weather-related traffic crashes result in 6,250 fatalities and over 480,000 injuries nation-wide each year (Federal Highway Administration, 2013). Correspondingly, the annual economic loss of weather-related crashes is about $22 billion, which is a conservative estimate without considering unreported crashes (Pisano et al., 2008). In the southwest region, unique weather extremes, including enormously high temperature, strong wind, flash flood, fog, dust, snow, etc., have resulted in more frequent and serious traffic crashes. For example, 501 traffic crashes were reported under extremely windy conditions in New Mexico in 2011. Although they only account for 1.2% of total crashes but result in about 3.4% of total fatalities (NMDOT, 2011). Under foggy conditions, fatal crashes account for 2.5% of total crashes, which is much higher than, 0.7%, the proportion of fatal crashes to total crashes under regular weather conditions in New Mexico in 2011 (NMDOT, 2011). These statistics dramatically illustrate that traffic crashes impacted by weather extremes are more serious than those under regular weather conditions. Substantial research efforts are needed to better understand significant causal factors and their impacts on crash severities under extreme, adverse weather conditions in order to develop effective countermeasures and proper policies to minimize weather-related risks to traffic safety. This study is proposed to quantify the impacts of region-wide weather extremes, driver behavior, demographic features, and environmental characteristics on crash severities. The developed crash record database and research findings are helpful for transportation agencies to develop cost-effective countermeasures to mitigate crash severities under extreme weather conditions in the southwest region.

1.2 OBJECTIVES

The objectives of this study are:

- Design and build a relational database that stores all the crash data, roadway geometric data, weather data, and traffic data;
- Develop a new hybrid approach to integrate multinomial Logt model with Bayesian network to discover the underlying patterns behind crash data and investigate the impacts of significant contributing attributes on crash severities impacted by weather extremes; and
- Identify high crash risk locations on the region-wide selected highways and better understand crash causes under extreme weather conditions; and
- Recommend cost-effective countermeasures for reducing crash severities impacted by weather extremes.

CHAPTER 2
REVIEW OF CURRENT PRACTICES

2.1 BACKGROUND

The impacts of weather extremes on traffic safety have been investigated during the last decades. By analyzing crash severities under extreme weather conditions, Qiu and Nixon (2008) concluded that excess precipitation is likely to lead to an increase in crash injury rates, and specifically, the impact of snowfall is greater than that of rainfall. Brodsky and Hakkert (1988) found a significant impact of excess precipitation on fatal crashes. Fridstrøm et al. (1995) clarified that similar impacts were found in Denmark but not in either Finland or Norway. Eisenberg (2004) analyzed the mixed effects of excess precipitation on traffic crashes and concluded that the relationship between monthly precipitation and monthly fatal crash is negative, while the one at daily levels is positive. On the other hand, strong winds have remarkable effects on high-profile vehicles, such as busses, trucks, delivery vans, to have them involved in rollover crashes, which are very likely to associate with fatal or incapacitating injuries (Baker and Reynolds, 1992). Objects carried by strong winds, such as fallen trees, could also affect traffic safety (SWOV, 2012). Additionally, winds can considerably magnify the impacts of adverse weather conditions, such as snowstorms (Usman et al., 2012). Foggy conditions and their influence on crash severities have been investigated (Cai et al., 2013; Edwards, 1998). As can be expected, crash severities dramatically increase under foggy conditions due to reduced visibility (Perry and Symons, 1991). The researchers (Andersson and Chapman, 2011; Koetse and Rietveld, 2009) also found that weather extremes, such as intense precipitation, wind, and snowstorm, are more likely to happen due to climate changes currently we are facing. Such large-scale long-term weather extremes will result in more serious crashes and cause significant loss in human life and property. However, research findings from the existing studies are local data-driven and location-specified and they are not applicable to address traffic safety issues impacted by unique weather extremes in the southwest region. It is of practical importance to investigate and identify significant contributing factors for weather-related crashes considering region-wide weather extremes, driver behavior, demographic features, and environmental characteristics, and further develop cost-effective countermeasures for traffic safety performance improvements.

2.2 STATE OF ART

Considerable modeling techniques and analysis methods have been used to investigate crash severities. Ulfarsson and Mannering (2004) estimated two separated multinomial Logit models for male and female drivers, respectively, in order to explore differences in injury severity outcomes between those two groups. Abdel-Aty and Keller (2005) applied ordered Probit models, binary Probit models, and nested Logit models, to conduct data analyses for 2043 intersection-related crashes occurred in Florida from 2003 to 2006. Zhu et al. (2005) applied a logistic regression model to investigate the relationship between seatbelt usages and crash occurrences. A binary Logit model and a multinomial Logit model were developed to explore main factors resulting in severe injury outcomes for two types of crashes: cross-median crashes and rollover crashes (Hu and Donnell, 2011). Savolainen and Mannering (2007) investigated a wide-range of factors and formulated the probabilities of having more severe injuries of motorcyclists using nested Logit and multinomial Logit models. The bus crash severities were investigated by Kaplan and Prato (2012) through estimating a generalized Logit model. Shankar (1996) employed a multinomial Logit model to evaluate the determinants of motorcycle accidents classified by five severity levels. Malyshkina and Mannering (2009) proposed two-state Markov switching multinomial Logit models to analyze crash severities. These studies could be divided into two types of model, ordered models (e.g. ordered Logit and Probit) and unordered models (e.g. multinomial Logit and nested Logit). Recently, Bayesian method has been gaining popularity in traffic crash modeling research. For instance, Eksler (2010) applied a Full Bayes (FB) model to analyze spatial-temporal traffic data and investigate roadway safety performance at local levels. Elvik (2013) studied the influence of speed limit on traffic crashes based on Empirical Bayes (EB) methods. Ma et al. (2008) developed a Bayesian approach for parameter estimation to calibrate a multivariate Poisson-lognormal model for crash severity prediction. MacNab (2003, 2004) used Bayesian hierarchical and spatial-ecological models for traffic crash analysis and prevention. Abdalla (2005) developed a hierarchical Bayesian method to evaluate protection effects of seatbelt in traffic crashes.

2.3 RESEARCH CONTRIBUTION

These studies provide a comprehensive and insightful understanding of discrete choice models and Bayesian approaches in traffic crash modeling and analysis. However, very few previous studies were conducted to analyze crash injury severities using Bayesian network methods. This study aims to significantly contribute to the state of the art and the practice by developing a multinomial Logit model-Bayesian network hybrid approach to discover the underlying patterns behind crash data and investigate the impacts of significant contributing attributes on crash severities impacted by weather extremes in the southwest region. The proposed hybrid approach will outperform the existing methods due to its flexibility to capture cause-effect relationships between contributing attributes and crash severity outcomes to better interpret their heterogeneous impacts on crash severity outcomes from the attribute changes in terms of region-wide weather extremes, driver behavior, demographic features, and environmental characteristics. This research project is closely relevant to the research topic identified in the SPTC Strategic Plan: minimizing weather/climate-related risks to transportation safety. Guided by the USDOT's priorities to promote the safe, efficient, and environmentally sound movement of goods and people, this project will formulate, analyze, and mitigate crash severities impacted by unique weather extremes, such as enormously high temperature, strong wind, flash flood, fog, dust, snow, etc., in the southwest region. A multinomial Logt model- Bayesian network hybrid approach will be developed to discover the underlying patterns behind crash data and investigate the impacts of significant contributing attributes on crash severities in terms of region-wide weather extremes, driver behavior, demographic features, and environmental characteristics. Effective countermeasures will be developed and recommended to minimize weather-related risks to traffic safety. The developed crash record database and research findings are helpful for transportation agencies to develop cost-effective solutions to reduce crash severities and improve region-wide traffic safety performance under extreme weather conditions.

CHAPTER 3
Region-wide Data Collection and Database Development

3.1 SQL Database Design and Development

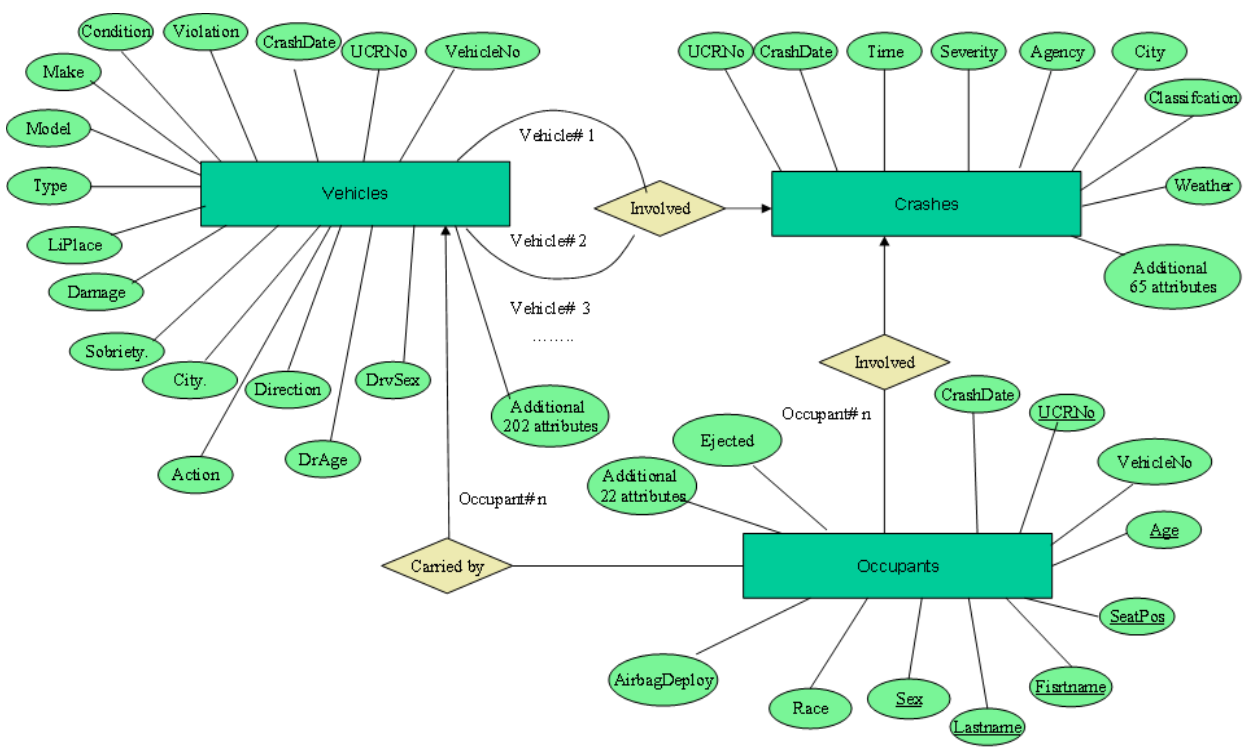
The research team has collected region-wide multiple-year crash data from state DOTs and MPOs in New Mexico, Arkansas, Louisiana, Oklahoma, and Texas to fully cover the representative roadway segments in the southwest region. Roadway geometric data, traffic data, and weather data will be collected from the corresponding agencies as well. The quality of these data will be verified. Table 3.1 shows the vehicle crash data due to weather condition.

Table 3.1. Crashes and Crash Fatality by Weather Condition, 2016

Weather	Crashes		Fatality	
	Count	Percent	Count	Percent
Clear	40,800	90.5%	363	89.6%
Inclement	3,035	6.7%	29	7.2%
Raining	1,683	3.7%	12	3.0%
Snowing	723	1.6%	5	1.2%
Wind	256	0.6%	4	1.0%
Other	221	0.5%	4	1.0%
Sleet or Hail	75	0.2%	3	0.7%
Fog	71	0.2%	1	0.2%
Dust	6	0.0%	0	0.0%
Missing Data	1,236	2.7%	13	3.2%
Total	45,071	100%	405	100%

Through the past several years, the research team has collected ten-year (2003-2012) traffic crash data in New Mexico from the New Mexico Department of Transportation (NMDOT) Traffic Safety Division and the Division of Government Research at UNM. Three major datasets are included: crash data, vehicle data, and driver data, which detail all the information regarding crash types, locations, severities, occurrence times, weather conditions, roadway geometric characteristics, vehicle characteristics, and driver demographic and behavior information. Based on this existing effort, more crash data and traffic data have been collected from the corresponding state Departments of Transportation (DOTs) and Metropolitan Planning Organization (MPOs) in Arkansas, Louisiana, Oklahoma, and Texas to fully cover the representative roadway segments in the southwest region. Roadway geometric characteristics, such as curvature data, gradient data, alignment data, etc., were obtained from the Highway Safety Information Systems (HSIS) at the U.S. Department of Transportation (USDOT). Additionally, detailed weather condition data will be collected from the Nation Weather Service. These datasets provide an excellent base for statistical analysis and modeling on traffic crashes impacted by weather extremes.

The database design and implementation have been concentrated on during this project period of three months. We use Microsoft SQL server 2012 to implement the database and manage the data. We use the E/R diagram method for database design. Relationships between database tables is set up through foreign keys and other constraints. Data collected are typically in Excel format. Microsoft SQL Server 2012 can import data directly from Excel files. The relationships between these imported files must be configured to make sure that they are connected correctly. The relational database was designed, and database architecture was proposed. Figure 3.1 shows the draft database design schema for crash, vehicle, and occupant tables. On a study of crash data analysis with respect to weather data analysis, Ahmed et al. (2012) found a strong correlation between month and number of crashes (Figure 3.2). The crash frequencies during the months of the snowy season were found to be more than double the frequencies during the months of the dry season.



E/R Diagram and Database Schema

Figure 3. 1. Database E/R Diagram and Database Scheme

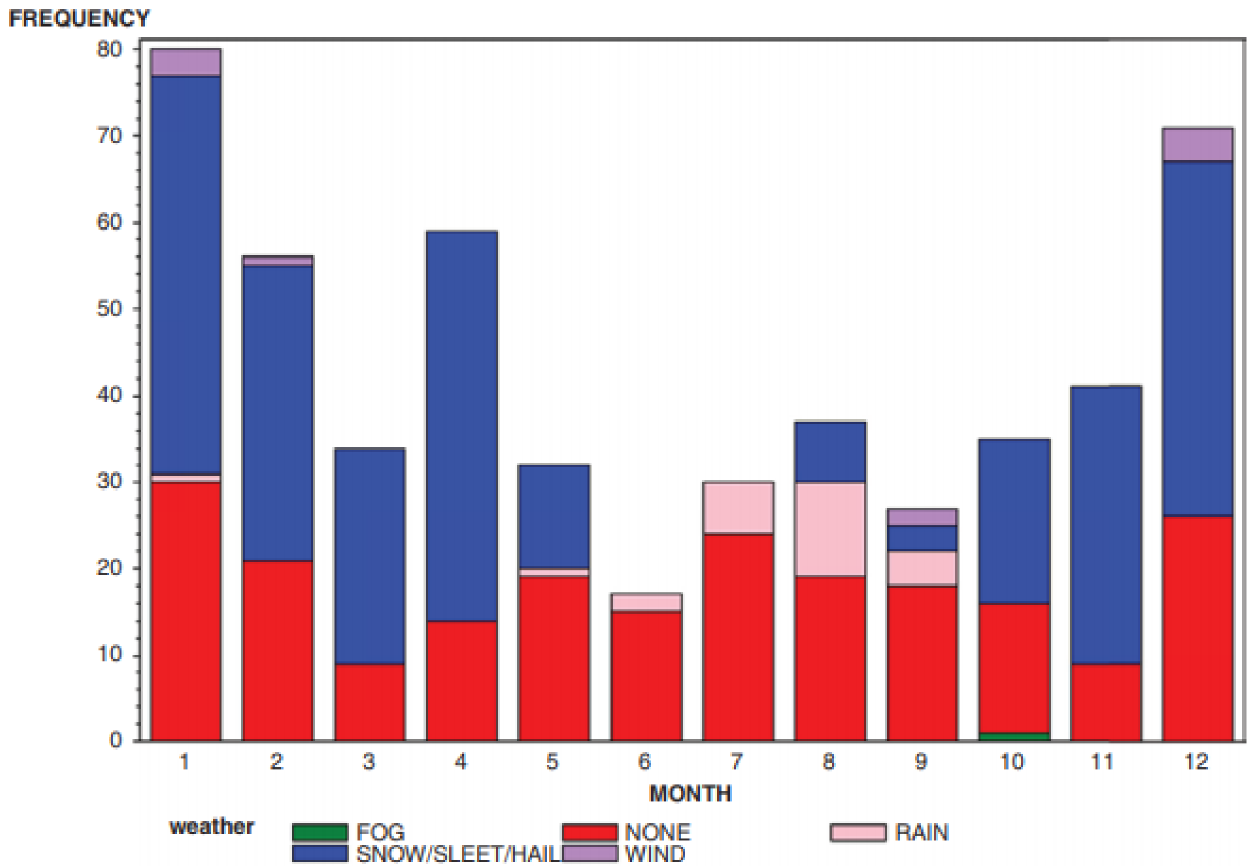


Figure 3.2. Crash Frequency by Month. Ahmed et al. (2012)

3.2 Fundamental Statistical Analysis

Statistical analysis and high crash risk location identification will help us better understand region-wide crash attributes impacted by weather extremes. Non-parametric analysis and *t* statistics is used for such analyses. High crash risk locations will be identified through statistical analyses. ArcGIS provides a great platform for visualizing data. To help identify crash-prone locations and understand crash distributions impacted by weather extremes, we convert crash and roadway data into GIS format. The analyses to be conducted include:

- Descriptive statistics
- Crash distributions and high crash risk location identification
- Conversion to GIS shape files

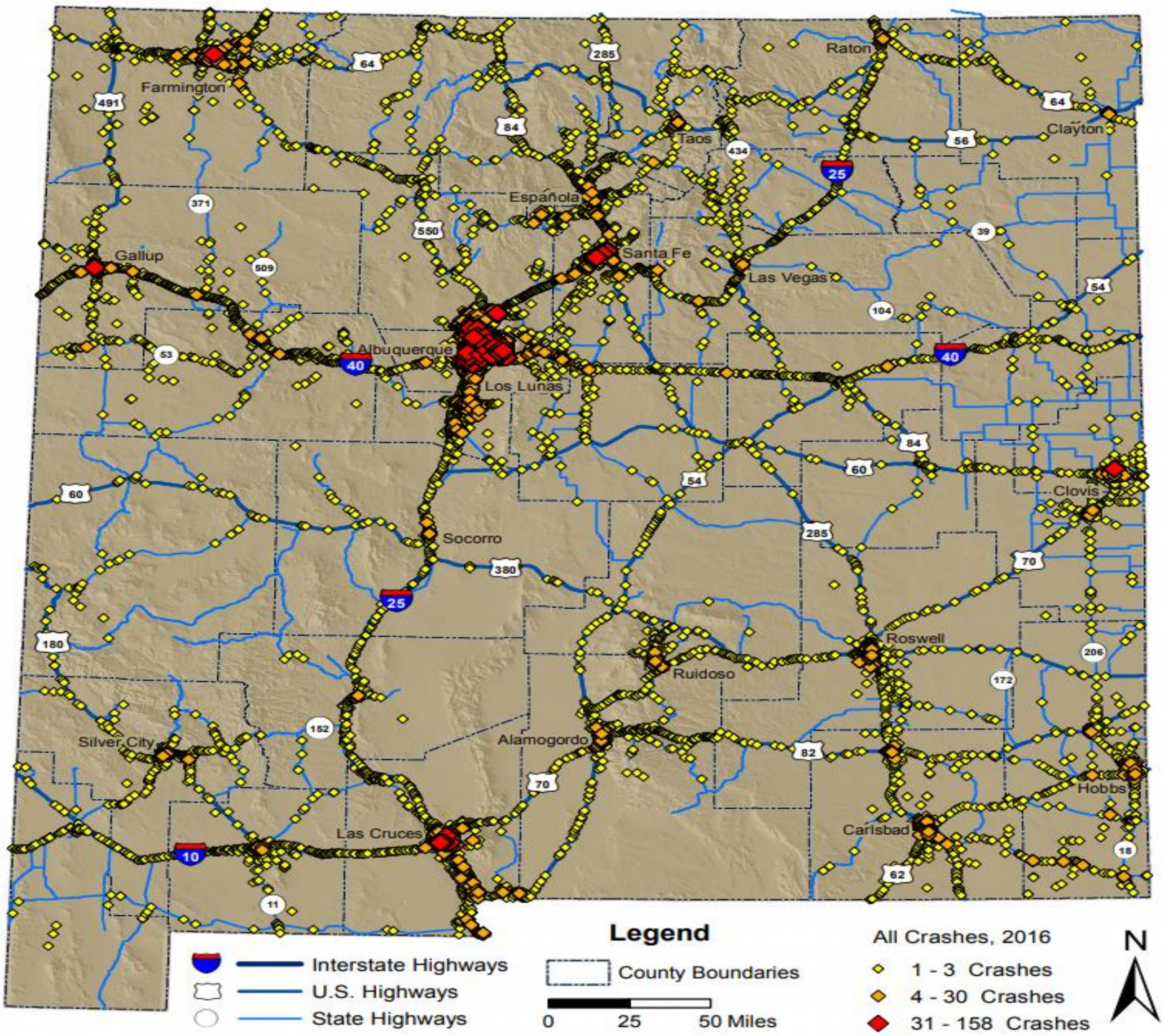


Figure 3.3. ArcGIS Crash Data for the State of New Mexico Year 2016

Figure 3.3 shows all crashes for the years of 2016 in the state of New Mexico. The GIS data analysis shows a strong correlation between crashes and weather condition. The analysis technique considered in this study to examine the effect of the interaction between geometric features, weather, and traffic data on crash occurrence. Although all geometric factors included in the models were significant during the dry and snowy seasons, the coefficient estimates indicate that the likelihood of a crash could be doubled during the snowy season because of the interaction between the snowy, icy, or slushy pavement conditions and the steep grades.

CHAPTER 4
MODEL DEVELOPMENT AND ANALYSIS

4.1 Methodology Design

Due to the difference in occurrence mechanisms, traffic crashes impacted by weather extremes will be modeled and investigated based on a new hybrid approach to integrate multinomial Logt model with Bayesian network to discover the underlying patterns behind crash data. In this study, a multinomial Logt model-Bayesian network hybrid approach will be developed to discover the underlying patterns behind data and investigate the impacts of significant contributing attributes on crash severity outcomes impacted by weather extremes in the southwest region. Bayesian networks have emerged as a powerful framework to extract expert knowledge and patterns hidden behind data through combining graph theory and probability theory. Graphical representations of Bayesian networks visualize complicated relationships and interactions among independent and dependent variables for constructing probabilistic inference and diagnosis. Therefore, Bayesian networks are capable of modeling inter-correlated independent variables to better interpret heterogeneous influence on weather-related crash injury severities from attribute changes. However, Bayesian network structure optimization in the global space is extremely computation-intensive considering a large amount of independent variables. The search space increases as a super-exponential function of the number of variables.

To achieve feasible and efficient network structure estimation, the significant variables must be selected to reduce the search space. Variable selection is very important to find a set of significant contributing variables and screen out variables that do not influence model performance. Many different variable selection criteria and methods have been used, such as the most commonly used correlation-based variable selection. During this process, a set of variables are selected due to their strong correlations with the output outcomes but low inter-correlations with each other (Mujalli and de Oña, 2011). Such correlation-based variable examination may not produce statistically stable results regarding variable predictive ability and redundancy reduction among variables. In this study, a multinomial Logit model is used to select significant variables to comprehensively screen out unnecessary and redundant attributes and increase optimal Bayesian network structure search efficiency.

4.2 Multinomial Logit Model-based Variable Selection

The severity of crashes in this study will be normally classified into three discrete categories, which describe the injury outcomes in weather-related crashes, including no injury, injury and fatality. Multinomial Logit models are developed to estimate the probability of three driver injury outcomes in weather-related crashes. It is assumed that for any attribute changes, the marginal costs for each severity outcome (no injury, injury, and fatality) are different. P_{is}, the probability of driver, *s*, being involved in a weather-related crash with severity level, *i*, is determined by the utility function U_{is}:

$$P_{is} = P\left(U_{is} \geq U_{js}, \forall i, j \in C, i \neq j\right) = P(u_{is} + \varepsilon_{is} \geq u_{js} + \varepsilon_{js}, \forall i, j \in C, i \neq j) \quad (1)$$

where u_{is} is the deterministic component that is only modeled by significant variables describing the instance; ε_{is} is the random component representing the hidden effect on driver injury severity; *C* is the choice set of possible driver injury severity outcomes. u_{is} is defined as a linear function for driver *s*,

$$u_{is} = \boldsymbol{\beta_i} \times \boldsymbol{V_{is}} + \alpha_{is} \quad (2)$$

where V_{is} is the exogenous variable vector influencing injury severity, *i*, for driver, *s*, and β_i is a coefficient vector to be estimated for measuring the influence of V_{is} on driver injury severity, *i*; α_{is} is the constant term. ε_{is} is normally assumed to follow a Generalized Extreme Value (GEV) distribution, and a multinomial Logit model can be derived as

$$P_{is} = \frac{e^{u_{is}}}{\sum_{j \in C} e^{u_{js}}} = \frac{e^{\beta_i \times V_{is} + \alpha_{is}}}{\sum_{j \in C} e^{\beta_j \times V_{js} + \alpha_{js}}} \quad (3)$$

where, P_{is} is the probability of driver, *s,* suffering injury outcome, *i,* in a crash. The coefficients $\boldsymbol{\beta_i}$ and α_{is} are estimated via maximum likelihood estimation methods. All the variables are used for multinomial Logit model development and significant ones are identified based on their T-ratios and P-values at the confidence level of p=0.01. These identified significant variables will be used for Bayesian network structure establishment and probabilistic parameter learning to explicitly formulate cause-effect relationships between injury severity outcomes and explanatory attributes.

In the study, model was developed in the premise of roll over crush and different variety of factors including severe weather was used in the analysis. The data used in this study majorly concentrate on all single vehicle rollover crashes occurring in the State of New Mexico from 2010 to 2012, which were obtained from the New Mexico Department of Transportation (NMDOT), the Traffic Safety Division (TSD) and the Division of Government Research (DGR) at the University of New Mexico. In this database, the code, CLASS, is used to classify the different types of crashes. Itis a rollover crash when the value of this code is equal to "01". Based on the variable, NVEH, which indicates the total number of vehicle involved in crashes, rollover crashes which only one vehicle get involved in are selected into this study. All types of vehicles including Passenger Car, Pickup, Truck, Van and other Four Wheel, were explored. Rollover crashes happened on the left/right side of the road and on the roads are all included in this study. After screening out observations with missing data in variables, a total of 4022 single-vehicle rollover crashes were used for model estimation in this study. Each record is at the individual level and contains all information about characteristics of crashes and drivers. Table 4.1 presents the distributions of the single vehicle rollover crashes across driver injuries for each variable considered in the study. The variable, Severity, is used to classify driver injury outcomes into five categories: No Injury (46.0%), Possible Injury (19.0%), Non-incapacitating Injury (23.5%), Incapacitating Injury (9.4%) and Fatality (2.1%). As suggested by previous studies (Chen et al., 2016b, 2015a), due to the small size of the last two categories and their similarities of injury severity attributes, they were combined as one category, which was named Incapacitating injury and fatality.

Table 4.1. Variable Definition and Description

Variable Description	**N[b]**		**P[C]**		**I[d]**		**II/F[e]**		**All**	
Crash Characteristics Severity	1851	46.0%	763	19%	944	23.5%	464	11.5%	4022	100%
Light Condition										
Daylight	1077	46.6%	547	23.7%	446	19.3%	242	10.5%	2312	57.5%
Dawn/Dusk	113	51.6%	47	21.5%	37	16.9%	22	10.0%	219	5.4%
Dark	661	44.3%	350	23.5%	280	18.8%	200	13.4%	1491	37.1%
Road Surface Condition										
Dry	1117	38.3%	581	19.9%	805	27.6%	410	14.1%	2913	72.4%
Wet	154	58.1%	42	15.8%	46	17.4%	23	8.7%	265	6.6%
Snow	197	71.6%	48	17.5%	22	8.0%	8	2.9%	275	6.8%
Ice	306	72.6%	65	15.3%	37	8.7%	17	4.0%	425	10.6%

4.3 Bayesian Network Definition

Bayesian network will be employed as a classifier to analyze driver injury severity outcomes in weather-related crashes based on the significant factors identified in the multinomial Logit model in this study. Bayesian network is capable of quantifying conditional probability relationships among variables via graphic presentation, known as a Directed Acyclic Graph (DAG) (Bouckaert, 2008). A BN can be represented by a network structure B_s over a set of variables, $V = \{x_1, x_2, \ldots, x_v\}$, $v > 1$. The DAG topology is portrayed to show cause-effect relationships among variables. A set of probability tables $B_p = \{p(x_i|pr(x_i)), x_i \in V\}$ are provided to quantitatively interpret these cause-effect relationships depicted by the graphical structure, B_s, where $pr(x_i)$ is the set of parent variables of x_i in B_s and i=1,2,…,v. Technically speaking, A BN over a set of variables, V, represents joint probability distributions, $P(V)=\prod_{x_i \in V} p(x_i|pr(x_i))$ for i=1,2,…, v. Using BN to analyze crash injury severities is to classify a potential driver injury outcome, $y=y_0$ (e.g. no injury, injury, fatality), given a set of significant attribute variables identified, $X = \{x_1, x_2, \ldots, x_k\}$, $k = v - 1$. The driver injury outcome, y, and the attribute variables, X, constitute the overall variable set $V=(X, y)$. The classifier is a function mapping a case of X to an outcome of y, which could be trained from a given dataset *D* that contains sample instances of (*X*, *y*). To use BN as a classifier, we need to calculate $argmax_y P(y|X)$, the value of *y* that maximizes $P(y|X)$, using the distribution *P(V)*, where

$$P(y|X) = \frac{P(X,y)}{P(X)} = \frac{P(V)}{P(X)} \propto P(V) = \prod_{x_i \in V} p(x_i|pr(x_i)) \qquad (4)$$

The Bayesian network structure graphically represents various intersections among variables. The variables are denoted as nodes and their interactions are represented by directional arcs and edges between two nodes. Unconnected nodes signify direct independence between the variables represented by the corresponding nodes. The optimal network structure, DAG, can be examined based on the prior knowledge constraints and predefined scoring metrics. In this study, prior knowledge and network structure score will be combined to achieve an efficient network structure estimation. The structure scoring metrics, Minimum Description Length (MDL) and Akaike information criterion (AIC), will be used as structure quality measurements.

4.4 Bayesian Learning and Model Specification

Supervised Bayesian network learning will be conducted to find an appropriate network structure and estimate the corresponding parameters based on the scoring metrics and prior expert knowledge to identify the best Bayesian network structure given a data set. Through training processes for both network structures and parameters, a Bayesian network is able to interpret observed crash injury severity data based on their probabilistic relationships and predicting unobserved crash injury outcomes based on attribute variables. In this study, an improved *K2* searching algorithm will be applied. A *K2* algorithm is a type of greedy hill climbing search algorithm, and based on this staring point, all the neighboring DAGs are established by adding, removing, and reversing an existing arc of the initial DAG. The scoring metrics are used to evaluate each DAG performance. A DAG with the highest score is the optimal network structure. The identified optimal BN structure presents the dependence relationships among the variables in the model based on the data set. The conditional probability tables can be estimated by maximum likelihood estimation methods during the parameter estimation process.

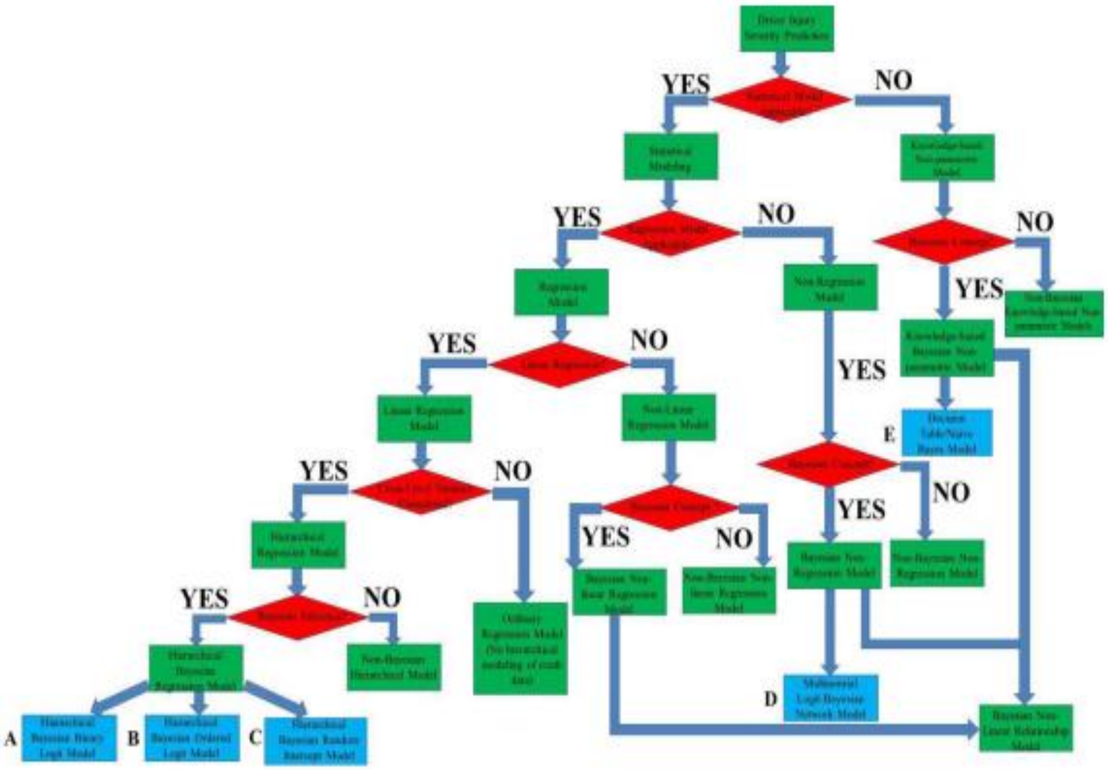


Figure 4.1. Conceptual Framework for Appropriate Selection of Bayesian Models for Driver Injury Severity Analysis.

A two-level hierarchical Bayesian logit model with binary response (indicated as Box A in Figure 4.1) was developed to estimate the effects of crash-level variables and vehicle/driver-level variables on driver injury severities, with the consideration of within crash correlations. In the lower level (vehicle/driver level), the injury severity of driver i in crash j, denoted as Sij, is a binary variable with Sij = 0 indicating no injury or slight injury, and Sij = 1 representing incapable injury or death. The probability of Sij = 1, denoted as Pij = Pr (Sij = 1).

For modeling result interpretation, the odds ratio rather than the estimated mean was utilized to explain the influence of the identified variables on driver injury severity. The odds ratio is the exponential output of the estimated mean for γ, exp(γ). An odds ratio equal to 1 means no effect for the studied variable on driver injury severity, which is corresponding to γ = 0; an odds ratio larger than 1.0 indicates that an increase of one unit on the studied variable would increase the odds of drivers being incapably injured or killed in a rural interstate crash by 100(exp(γ) − 1)% compared with the base case. An odds ratio less than 1.0 implies that an increase of one unit on the studied variable would decrease the odds of drivers being incapably injured or killed in a rural interstate crash by 100(1 − exp(γ)%. The 95% Bayesian Credible Interval (BCI) is provided to indicate the significance of the variables (Gelman et al., 2013), and 90% BCI is also calculated as an additional reference. A variable is considered significant in affecting driver injury severity if the 95% BCI of its odds ratio does not cover 1 and is not significant if otherwise. Experience and consensus on traffic safety analyses are also referred to for result reasonableness examination.

According to model analysis, driver fatalities are more likely to occur in a comfortable traffic environment, such as clear weather, level road grade, and paved road surface, whereas drivers would be more aware of potential risk under adverse driving conditions. Driver fatal injuries are most likely to happen on rural roads, especially on rural two-lane highways. Maximum vehicle damage in rear-end crashes is positively associated with driver injury severities, and drivers are most likely to suffer fatal injuries when vehicles involved in rear-end crashes are disabled. The Bayesian network reference analyses indicate that the factors, such as truck-involvement, inferior lighting conditions, windy weather conditions, the number of vehicles involved, etc. could significantly increase driver injury severities in rear-end crashes.

4.5 Countermeasure Identification and Evaluation

Based on the model specifications and research findings, effective countermeasures and proper policies are identified and recommended. The meetings were set up with state traffic engineers to discuss and evaluate the effectiveness of these countermeasures. Before-and-after studies may be conducted to evaluate these countermeasures depending on data availability. The goodness of fit of the model's prediction versus the actual results determines our confidence levels in using the proposed approach for traffic safety performance improvements under extreme weather conditions. This study will significantly contribute to the state of the art and the practice. The proposed hybrid approach is capable of capturing cause-effect relationships between contributing attributes and crash severities to better interpret their heterogeneous impacts on crash severities impacted by weather extremes. The methodology and crash record database developed in this study can be used to identify high crash risk locations and improve region-wide safety conditions. The research findings are helpful for transportation agencies to develop cost-effective solutions to reduce crash severities under extreme weather conditions and minimize the weather-related risks to traffic safety in the southwest region. Rumble stripes have been proposed and implemented as an effective solution to increase nighttime lane edge visibility to prevent roadway departure and overturn occurrences. A rumble strip becomes a rumble stripe when a retroreflective pavement marking is placed on it. Rumble stripes integrate the benefits of rumble strips and retroreflective lane markings. The contour of the rumble strip drains water, and the reflective rumble striping provides a back wall allowing the retroreflective markings to highlight the lane edge during nighttime and other low visibility weather conditions (Federal Highway Administration (FHWA), 2011b). It was also found that retroreflective markings under rumble strips have more reflectivity than the standard edge-line/centerline markings, and these rumble strips are more resilient and durable than standard markings, especially in heavy winter climates (Torbic et al., 2009).

Retroreflective materials have wide applications in transportation systems, including traffic signs and pavement markings. They also provide road users necessary information for safety and expedite trip activities. The Manual on Uniform Traffic Control Device (U.S. Department of Transportation, 2012) defines transportation-related retroreflective colors and the specific meaning for each color in traffic signs and pavement markings and also addresses the use of markings in combinations with longitudinal (shoulder, centerline) rumble strips and transverse rumble strips. Figure 4.2 shows the examples of longitudinal rumble strip markings. According to MUTCD, “if it is desirable to use a color other than the color of the pavement for a longitudinal rumble strip, the color of the rumble strip shall be the same color as the longitudinal line the rumble strip supplements”.

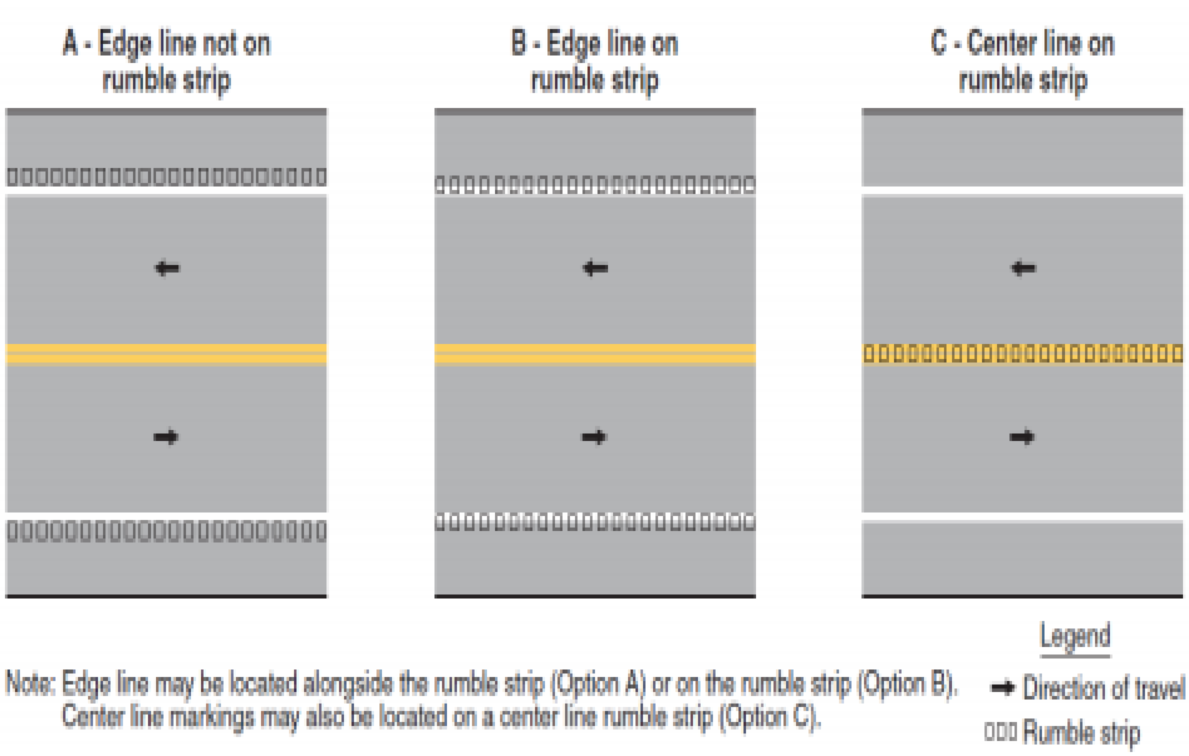


Figure 4.2. MUTCD Longitudinal Rumble Strip Markings.

Rumble stripes have been utilized in peer regions to increase lane edge visibility during nighttime and under low visibility weather conditions such as rain, fog, and snow. As discussed above, the MUTCD specifies the minimum maintained retro-reflectivity for pavement markings, which also applies to the retroreflective paintings on rumble strips

(FHWA, 2010). However, the rumble strips are deteriorated due to aging, traffic loading, and weather change, and become less visible during nighttime and adverse weather conditions.

Figure 4.3. Installed White Rumble Stripes (Near Outside Shoulder)

To reduce the potential for overturn crashes and injury severities, NMDOT initiated a project and applied retroreflective rumble striping with elements on existing rumble strips along U.S. 285 to increase their visibility. In this project, the rumble strip paintings were applied by using two applications of high-durable acrylic traffic paint installed at 22 to 25 mils wet film thickness per application as per Item Number 018 on existing rumble strips. Following the second application, Double Drop dry elements were placed as per Item Number 022. Both paintings and elements add to the visual representation of edge line location as well as the angles associated with a rumble strip magnifying the reflective capability of a painted stripe. These applications were implemented parallel to the existing shoulder stripe on the inside and outside rumble strips for multi-lane median divided sections and only on the outside rumble strips for two-lane sections. Their colors matched the adjacent should stripe color, Figure 4.3. The rumble strips were fully within the width of the existing milled rumble strip and did not overlap onto the shoulder pavement.

CHAPTER 5
CONCLUSIONS

This project aims to develop comprehensive, cost-effective, and implementable solutions to region-wide infrastructure-related issues in support of the nation's transportation systems through regionally-based, nationally-recognized research, education, and outreach activities. This study significantly contributes to the state of the art and the practice minimizing weather/climate-related risks to transportation safety. Guided by the USDOT's priorities to promote the safe, efficient and environmentally sound movement of goods and people, this project will formulate, analyze, and mitigate crash severities impacted by unique weather extremes in the southwest region. Effective countermeasures will be developed to minimize weather-related risks to traffic safety. The developed crash record database and research findings are helpful for transportation agencies to develop cost-effective solutions to reduce crash severities and improve region- and nation-wide traffic safety performance under extreme weather conditions.

The region-wide crash record database developed in this study can also be used for future safety studies in the southwest region. Results from the statistical analyses will help state DOTs to better understand crash severities impacted by weather extremes and identify locations with higher crash risks, particularly in New Mexico, Arkansas, Louisiana, Oklahoma, and Texas. The hybrid multinomial Logt model-Bayesian network modeling approach can also be used for evaluating the effectiveness of safety improvement plans. Based on the model specifications and research findings, effective countermeasures will be identified and recommended. Before-and-after studies may be conducted to evaluate these countermeasures depending on data availability. The goodness of fit of the model's prediction versus the actual results determines our confidence levels in using the proposed approach for traffic safety performance improvements under extreme weather conditions. The research findings are helpful for transportation agencies to develop cost-effective countermeasures to mitigate crash severities under extreme weather conditions and minimize the weather-related risks to traffic safety in the southwest region.

5.1 Research Outcome and Impact

The expected outcomes of this study include:

- A relational database with all the crash data, roadway geometric data, weather data, and traffic data.
- A region-wide GIS map with high crash risk locations on the selected roadways impacted by weather extremes;
- A new hybrid approach to integrate multinomial Logt model with Bayesian network to formuate crash severities under weather extremes, including

enormously high temperature, strong wind, flash flood, heavy dust, fog, and snow.

- A better understanding of significant contributing attributes and their impacts on crash severities in terms of weather extremes, driver behavior, demographic features, and environmental characteristics.

This study will significantly contribute to the state of the art and the practice. The proposed hybrid approach is capable of capturing cause-effect relationships between contributing attributes and crash severities to better interpret their heterogeneous impacts on crash severities impacted by weather extremes. The methodology and crash record database developed in this study can be used to identify high crash risk locations and improve region-wide safety conditions. The research findings are helpful for transportation agencies to develop cost-effective solutions to reduce crash severities under extreme weather conditions and minimize the weather-related risks to traffic safety in the southwest region.

References

Abdel-Aty, M.A., Keller, J., 2005. Exploring the overall and specific crash severity levels at signalized intersections. Accid. Anal. Prev. 37, 417–25.

Andersson, A.K., Chapman, L., 2011. The impact of climate change on winter road maintenance and traffic accidents in West Midlands, UK. Accid. Anal. Prev. 43, 284–9.

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

Andrey, J., 2010. Long-term trends in weather-related crash risks. J. Transp. Geogr. 18, 247–258.

Baker, C.J., 1993. The behaviour of road vehicles in unsteady cross winds. J. Wind Eng. Ind. Aerodyn. 49, 439–448.

Baker, C.J., Reynolds, S., 1992. Wind-induced accidents of road vehicles. Accid. Anal. Prev. 24, 559–575.

Bilionis, D.V., 2013. Interaction effects of prevailing weather conditions and crash characteristics on crash severity : A case study on two corridors in Iowa. Grad. Theses Diss. Paper 13182.

Brodsky, H., Hakkert, A.S., 1988. Risk of a road accident in rainy weather. Accid. Anal. Prev. 20, 161–176.

Cai, X., Lu, J.J., Xing, Y., Jiang, C., Lu, W., 2013. Analyzing Driving Risks of Roadway Traffic under Adverse Weather Conditions: In Case of Rain Day. Procedia - Soc. Behav. Sci. 96, 2563–2571.

Edwards, J.B., 1998. The Relationship Between Road Accident Severity and Recorded Weather. J. Safety Res. 29, 249–262.

Eisenberg, D., 2004. The mixed effects of precipitation on traffic crashes. Accid. Anal. Prev. 36, 637–47.

Federal Highway Administration, 2013. How do weather events impact roads? [WWW Document]. URL http://www.ops.fhwa.dot.gov/weather/q1_roadimpact.htm

Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., Thomsen, L.K., 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Accid. Anal. Prev. 27, 1–20.

Hu, W., Donnell, E.T., 2011. Severity models of cross-median and rollover crashes on rural divided highways in Pennsylvania. J. Safety Res. 42, 375–82.

Kaplan, S., Prato, C.G., 2012. Risk factors associated with bus accident severity in the United States: a generalized ordered logit model. J. Safety Res. 43, 171–80.

Kim, J.-K., Kim, S., Ulfarsson, G.F., Porrello, L. a, 2007. Bicyclist injury severities in bicycle-motor vehicle accidents. Accid. Anal. Prev. 39, 238–51.

Koetse, M.J., Rietveld, P., 2009. The impact of climate change and weather on transport: An overview of empirical findings. Transp. Res. Part D Transp. Environ. 14, 205–221.

Malyshkina, N. V, Mannering, F.L., 2009. Markov switching multinomial logit model: An application to accident-injury severities. Accid. Anal. Prev. 41, 829–38.

McFadden, D., 1981. Econometric models of probabilistic choice, in: A Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge.

New Mexico Department of Transportation (NMDOT), 2011. New Mexico Traffic Crash Annual Report.

Perry, A.H., Symons, L., 1991. Highway Meteorology. Taylor & Francis, New York.

Pisano, P.A., Goodwin, L.C., Rossetti, M.A., 2008. U.S. highway crashes in adverse road weather conditions, in: 24th Conference on IIPS. pp. 1–15.

Qiu, L., Nixon, W. a., 2008. Effects of Adverse Weather on Traffic Crashes: Systematic Review and Meta-Analysis. Transp. Res. Rec. J. Transp. Res. Board 2055, 139–146.

Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accid. Anal. Prev. 39, 955–63.

Shankar, V., Mannering, F., 1996. An exploratory multionmial logit analysis of single-vehicle motorcycle accident severity. J. Safety Res. 27, 183–194.

SWOV Institute for Road Safety Research, 2012. The influence of weather on road safety.

Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36, 135–47.

Usman, T., Fu, L., Miranda-Moreno, L.F., 2012. A disaggregate model for quantifying the safety effects of winter road maintenance activities at an operational level. Accid. Anal. Prev. 48, 368–78.

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2003. Statistical and econometric methods for transportation data analysis. Chapman & Hall/CRC, Washington, D.C.

Zhu, M., Hardman, S.B., Cook, L.J., 2005. Backseat safety belt use and crash outcome. J. Safety Res. 36, 505–7.