IDENTIFICATION, INVESTIGATION, AND SPATIAL ANALYSIS OF VARIOUS CONTRIBUTING FACTORS TO CRASH AND INJURY SEVERITY IN DIFFERENT CRASH TYPES

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Ihsan Khan

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Transportation, Logistics, and Finance

November 2023

Fargo, North Dakota

North Dakota State University

Graduate School

Title

IDENTIFICATION, INVESTIGATION, AND SPATIAL ANALYSIS OF VARIOUS CONTRIBUTING FACTORS TO CRASH AND INJURY SEVERITY IN DIFFERENT CRASH TYPES

By

Ihsan Khan

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Diomo Motuba
Chair

Dr. Kimberly Vachal
Co-chair

Dr. Joseph Szmerekovsky

Dr. Ranjit Prasad Godavarthy

Dr. Megan Orr

Approved:

11/20/2023
Date

Dr. Tim Peterson
Department Chair

ABSTRACT

This dissertation had three main objectives related to improving road safety by investigating factors that contribute to injury severity in different types of single-vehicle crashes. The first objective was to develop a generalized ordered logit model to examine factors affecting injury severity of occupants in single-vehicle rollover crashes using 5 years of U.S. crash data from 2012-2016. Results showed likelihood of serious/fatal injuries increased in rollovers with occupant ejection, speeding, higher speed limits, roadside/median rollovers, undulating terrain, blacktop surfaces, rural roads, evenings, weekdays, older drivers, lack of occupant protection, previous driver crashes, distracted/aggressive driving, and passenger cars. Airbag deployment reduced serious/fatal injury risk. Regional variations also impacted injury severity.

The second objective identified high-risk areas for lane departure crashes on rural North Dakota roads using techniques like Global/Local Moran's I, network kernel density estimation (NetKDE), and emerging hotspot analysis. While Global Moran's I indicated clustering, Local Moran's I revealed specific hot/cold spots. NetKDE quantified and prioritized crash clusters by density along roadways. Emerging hotspot analysis evaluated temporal patterns of hot/cold spots. This approach can guide deployments of education, enforcement, and infrastructure countermeasures.

The third objective used a mixed logit model to analyze factors contributing to injury severity in single-vehicle run-off-road (ROR) crashes for passenger cars, SUVs, and pickups. Common factors increasing injury risk were older driver age, impaired driving, no seatbelt, no airbag, high speeds, and older vehicles. However, driver age impacts were most pronounced for pickups. Seatbelts substantially mitigated injury severity across all vehicle classes. Passenger

cars had a higher injury risk than SUVs/pickups, especially over 75 mph. Future research should examine additional factors stratified by vehicle class using larger datasets.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Diomo Motuba, and co-advisor Dr. Kimberly Vachal, whose guidance, expertise, and unwavering support have been instrumental in shaping this research journey. Their invaluable mentorship and encouragement have inspired me to push the boundaries of knowledge and strive for excellence.

I am profoundly grateful to the members of my dissertation committee, Dr. Joseph Szmerekovsky, Dr. Ranjit Prasad Godavarthy, and Dr. Megan Orr for their insightful feedback, constructive criticism, and dedication to ensuring the rigor and quality of this work. Your expertise and diverse perspectives have enriched this study immeasurably.

My heartfelt thanks go to my family and friends for their unwavering love, encouragement, and understanding throughout this demanding endeavor. Your support has been my rock, sustaining me through the highs and lows of doctoral study.

I am grateful to the Upper Great Plains Transportation Institute for providing the resources, facilities, and intellectual environment conducive to scholarly inquiry. Stimulating discussions, collaborative spirit, and exchange of ideas within the academic community have been integral to my growth as a researcher.

Lastly, I extend my appreciation to all those who have crossed my path and contributed in ways both big and small to the completion of this dissertation. Your support, encouragement, and belief in my capabilities have been invaluable, and for that, I am deeply thankful.

DEDICATION

I dedicate this dissertation to my younger brother, Captain Arif Ullah Khan (Martyred on June 7th, 2019), whose memory continues to inspire me with courage and resilience.

To my parents, Shah Daraz Khan and Mumtaz Begum, whose unwavering love and support have been my guiding light throughout this academic journey.

To my siblings: Inam Ullah Khan, Irshad Bibi, Shamshad Begum, Yasmin Bibi, Irfan Ullah Khan, and Laiba Khan.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGMENTS v

DEDICATION vi

LIST OF TABLES x

LIST OF FIGURES xi

CHAPTER 1: INTRODUCTION 1

1.1. Overview 1

1.2. Problem Statement 2

1.3. Research Objectives 3

1.3.1. Paper 1: Factors Affecting Injury Severity of Single-Vehicle Rollover Crashes in The United States 3

1.3.2. Paper 2: Hotspot Analysis of Single-Vehicle Lane Departure Crashes in North Dakota 4

1.3.3. Paper 3: Investigating Factors Affecting Injury Severity of Single-Vehicle Run-off-Road Crashes 5

CHAPTER 2: FACTORS AFFECTING INJURY SEVERITY OF SINGLE-VEHICLE ROLLOVER CRASHES IN THE UNITED STATES 6

2.1. Introduction 6

2.2. Methodology 9

2.2.1. Data Source and Preparation 9

2.3. Model Development 13

2.4. Results 14

2.5. Discussion 18

2.6. Acknowledgments 24

2.7. References 24

CHAPTER 3: HOTSPOT ANALYSIS OF SINGLE-VEHICLE LANE DEPARTURE CRASHES IN NORTH DAKOTA ... 27

3.1. Introduction ... 27

3.2. Literature Review ... 28

3.3. Scope and Data ... 31

3.4. Methodology ... 32

3.4.1. Global Spatial Autocorrelation ... 33

3.4.2. Local Spatial Autocorrelation ... 33

3.4.3. Network Kernel Density Estimation ... 36

3.4.4. Space-Time Cube Analysis ... 38

3.5. Results and Discussion ... 40

3.6. Conclusions ... 48

3.7. References ... 49

CHAPTER 4: INVESTIGATING FACTORS AFFECTING INJURY SEVERITY OF SINGLE VEHICLE RUN-OFF-ROAD CRASHES ... 54

4.1. Introduction ... 54

4.2. Literature Review ... 57

4.3. Data and Methodology ... 60

4.4. Elasticity Analysis ... 67

4.5. Results and Discussion ... 68

4.5.1. Continuous Variables ... 72

4.5.2. Driver Characteristics ... 72

4.5.3. Vehicle Characteristics ... 73

4.5.4. Environmental Information ... 74

4.5.5. Crash Information ... 74

4.6. Conclusions ... 81

4.7 References 82
CHAPTER 5: CONCLUSIONS 87
5.1. Driver Behavior and Occupant Protection 87
5.2. Vehicle Characteristics and Technology 88
5.3. Roadway Design and Infrastructure Improvements 88
5.4. Integrated Approach and Data-Driven Decision-Making 89
5.5. Future Research and Policy Implications 89

LIST OF TABLES

Table Page

1. Descriptive statistics of variables included in the model development 10

2. Results of generalized ordered logit model for serious injury in single-vehicle rollover crashes 15

3. Results of generalized ordered logit model for fatal injury in single-vehicle rollover crashes 16

4. Crash characteristics of the data 32

5. Names and definitions of space-time cube patterns... 40

6. The z-scores and p-values for a series of distance 42

7. Descriptive statistics of explanatory variables... 61

8. Mixed logit model results for PC... 69

9. Average Pseudo elasticities of significant variables for PC 71

10. Mixed logit model results for SUV.. 76

11. Average Pseudo elasticities of significant variables for SUV 78

12. Mixed logit model results for pickup/passenger van 80

13. Average Pseudo elasticities of significant variables for pickup/passenger van................ 81

LIST OF FIGURES

Figure Page

1. Statistics of different crash types from 2007 to 2016 in the United States 6

2. Predicted probabilities for injury severity of occupant's ejection versus speed limit 17

3. Predicted probabilities for injury severity of occupant's ejection versus seatbelt use 18

4. Single-vehicle rollover crashes in the United States from 2012 to 2016.......................... 22

5. Crash statistics of single-vehicle lane departure crashes in ND from 2016 – 2020 28

6. Kernel density estimation procedure on a line segment: (a) molehills or kernels (b) a big mountain or kernel density estimate.. 36

7. (Top) Structure of STC in 3D (Bottom) generated bins in 2D by running emerging hotspot analysis.. 38

8. Global Moran's I value for single-vehicle lane departure crashes 41

9. Hotspot classification using local Moran's I method 43

10. Hotspot analysis using local spatial autocorrelation (Getis-Ord Gi*) approach............... 44

11. Network Kernel density estimation for crashes in Williams County 45

12. Emerging hotspot analysis of single-vehicle lane departure crashes in ND..................... 46

13. Comparison of results from different hotspot methods – (top left) local Moran's I (top right) Getis Ord Gi* (bottom left) network kernel density estimation (bottom right) emerging hotspot analysis.. 47

14. Fatalities in single and multiple vehicle ROR crashes in the United States (2017-2021) 55

15. Bivariate scatter plot of random uniform draws 66

16. Bivariate scatter plot of Halton (9) and Halton (11) 66

CHAPTER 1: INTRODUCTION

1.1. Overview

In today's world, where transportation is an integral aspect of daily life, road safety stands as an imperative concern. The significance of road safety extends far beyond mere statistics; it resonates deeply with every individual who steps onto a road or behind the wheel of a vehicle. From preventing crashes to preserving lives, the importance of road safety cannot be overstated. Road traffic crashes continue to be a major public health concern worldwide, resulting in significant loss of life, injuries, and economic burdens. Despite ongoing efforts to enhance road safety measures, the number of crashes and their associated severity remain alarmingly high. The consequences of traffic crashes extend far beyond the immediate impact, affecting individuals, families, communities, and the broader society. Severe crashes not only result in fatalities and life-altering injuries but also impose substantial economic costs on healthcare systems, emergency services, and infrastructure maintenance.

Understanding the intricate interplay of factors that contribute to injury severity is crucial for developing effective countermeasures and implementing targeted interventions to mitigate these adverse outcomes. This dissertation endeavors to provide a comprehensive analysis of the multifaceted elements that influence injury severity, including environmental factors, road characteristics, vehicle dynamics, driver behavior, and other relevant variables. The findings of this research have the potential to inform policymakers, transportation authorities, and stakeholders in developing evidence-based strategies for enhancing road safety measures, infrastructure planning, and public awareness campaigns. Ultimately, this dissertation seeks to contribute to the broader goal of reducing the burden of traffic crashes and improving the overall safety and well-being of road users. The breakdown of each chapter is discussed below.

1.2. Problem Statement

The discussion in paper 1 illuminates a critical problem in the current understanding of rollover crashes. While numerous studies have concentrated on risk factors contributing to rollovers at the state level, a critical gap exists as no study has provided insight into the nationwide perspective. The lack of a national analysis hampers our ability to grasp the diverse variables influencing injury severity in rollovers, resulting in a fragmented understanding. This knowledge gap is significant as it may overlook crucial factors related to the nationwide geographic and infrastructural differences. Consequently, the development of effective, nationwide strategies and interventions to mitigate the impact of rollover crashes is impeded. To address this issue, this chapter aims to bridge the gap by employing an ordered logit model to comprehensively analyze injury severity factors in single-vehicle rollover crashes across the entire United States, thus contributing essential insights for the formulation of informed and effective national strategies.

In paper 2, the discussion is centered on single-vehicle lane departure crashes in North Dakota. North Dakota faces a pressing road safety challenge with 77% of lane departure crashes (Vision Zero, NDDOT) involving single vehicles, compounded by the predominantly rural landscape and limited resources in financial, personnel, and infrastructure domains. Despite the prevalence of these crashes, a comprehensive understanding of their spatial clustering is lacking, hindering the development of effective intervention strategies. The current analyses of crash hotspots fall short by not adequately considering the road network's configuration, potentially leading to the misidentification of high-risk areas. Furthermore, the neglect of the temporal dimension in existing analyses obscures dynamic patterns over time, preventing the formulation of responsive strategies. The critical problem at hand is the imperative need for a targeted identification approach to discern areas with elevated risks of lane departure crashes and implement interventions that address this specific challenge in North Dakota's unique road safety landscape.

In paper 3, the study delves into a critical problem within the literature concerning the severity of Run-Off-Road (ROR) crashes and their contributing factors. While numerous efforts have been made to unravel this complexity, the existing literature tends to consolidate vehicle types into a single variable, primarily utilizing binary or dummy variables in analyses. This oversimplified approach hinders a nuanced understanding of the factors contributing to injury severities in ROR crashes for distinct vehicle types, including large trucks, SUVs, and passenger cars. The homogenization of vehicle types may obscure unique risk profiles and contributory factors inherent to each category. To address this limitation, the present study seeks to disaggregate these vehicle types and conduct a detailed analysis to discern if and how factors influencing injury severity vary across different vehicle categories on a global scale. This research aims to provide a more refined understanding of the complexities associated with ROR crashes for distinct vehicle types, thereby contributing to enhanced road safety strategies tailored to the unique characteristics of each category.

1.3. Research Objectives

1.3.1. Paper 1: Factors Affecting Injury Severity of Single-Vehicle Rollover Crashes in The United States

The primary inquiry centers on identifying the most influential factors in predicting injury severity nationwide. Additionally, paper 1 aims to uncover regional variations in the impact of these factors across different regions within the United States, recognizing the diverse geographic landscape. An important goal is to discern not only the primary predictors but also how their effects may differ based on regional contexts. Beyond understanding these variations, the research seeks to draw implications for national traffic safety standards and vehicular design regulations. By addressing these questions, the study aspires to contribute valuable insights that can inform more effective and regionally nuanced approaches to enhance road safety standards across the United States.

1.3.2. Paper 2: Hotspot Analysis of Single-Vehicle Lane Departure Crashes in North Dakota

Firstly, paper 2 aims to identify discernible global hotspots that may necessitate targeted intervention strategies, establishing a baseline understanding of crash distribution across North Dakota. Building on this, the research aims to refine these findings by investigating more nuanced and localized crash hotspots within North Dakota’s road network, considering the intricacies of the transportation infrastructure. Additionally, the study strives to enhance the analysis by integrating temporal data through the Space-Time Cube method, exploring how the evolution of crash hotspots over time can inform the development of dynamic, time-sensitive safety interventions. These research questions collectively form a multifaceted approach to addressing the complexities of single-vehicle lane departure crashes, ranging from global patterns to localized nuances and temporal dynamics.

The primary objective of this paper is to comprehensively analyze the patterns of single-vehicle lane departure crashes in North Dakota, aiming to identify hotspot areas at higher risk and discern temporal trends for effective safety interventions. Specific objectives include conducting a global analysis to determine the presence and location of statistically significant clusters of lane departure crashes across the state, and prioritizing areas for safety improvements. Additionally, the study seeks to perform a local analysis to identify concentrated crash areas within broader hotspots, offering a detailed local perspective for targeted safety measures. Furthermore, the research aims to conduct a spatio-temporal analysis, exploring the temporal variability of crash hotspots and recognizing their dynamic nature, with the ultimate goal of proposing guidelines for time-specific interventions. This multifaceted approach ensures a comprehensive understanding of single-vehicle lane departure crashes, incorporating global, local, and temporal perspectives for informed and targeted safety interventions.

1.3.3. Paper 3: Investigating Factors Affecting Injury Severity of Single-Vehicle Run-off-Road Crashes

This paper endeavors to address critical questions regarding the injury severity of single-vehicle ROR crashes across various vehicle classes. The research questions include an exploration of the factors contributing to injury severity in these incidents and an investigation into how these factors differ among passenger cars, SUVs, pickups, and large trucks. The main objectives encompass the development of class-specific injury severity models for each vehicle type, the identification of key factors associated with increased injury severity within each class, and a comparative examination of similarities and differences in contributing factors across all four vehicle classes. Ultimately, the study aims to provide comprehensive insights into both class-specific and common factors influencing injury outcomes in single-vehicle ROR crashes, contributing to the development of targeted road safety strategies tailored to the unique characteristics of each vehicle class.

CHAPTER 2: FACTORS AFFECTING INJURY SEVERITY OF SINGLE-VEHICLE ROLLOVER CRASHES IN THE UNITED STATES

2.1. Introduction

Motor vehicle (MV) crashes are a major global concern because of their socioeconomic impacts in terms of human loss, productivity loss, and property damage. MV crashes result in injuries that are sudden and traumatic. One of the most severe occurrences reported in MV crash event sequences is a rollover. Rollover refers to a crash in which vehicle rotation of at least one-quarter turn (greater than or equal to 90 degrees) occurs about the vehicle's longitudinal or lateral axis. Once the rolling stops, the vehicle may land on a side, upside down, or upright (Conroy et al. 2006). Rollover crashes were found to cause more injuries and fatalities than other crash types (Jehle, Kuebler, and Auinger 2007). Figure 1 shows different types of crash statistics.

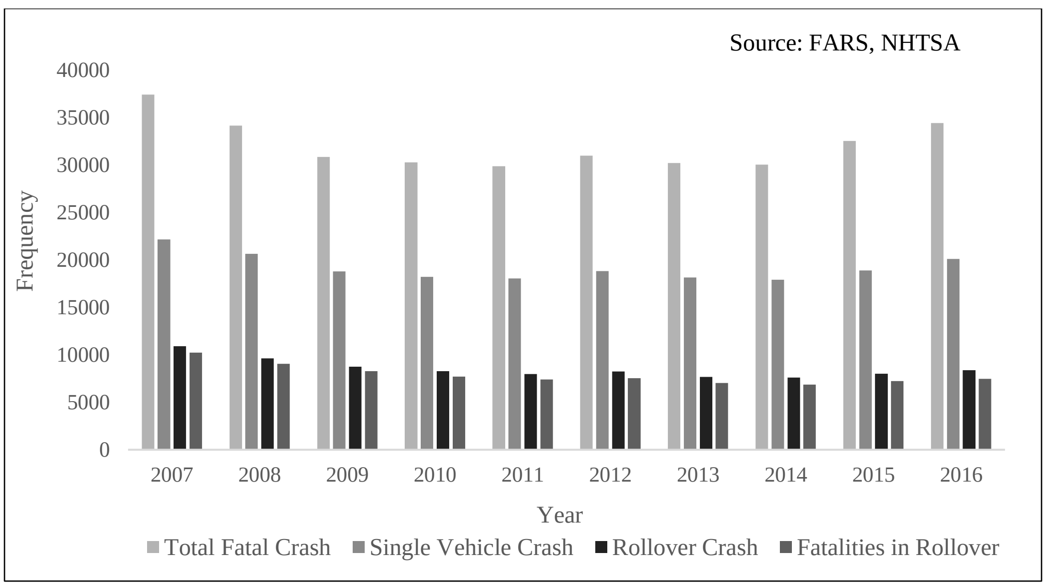


Figure 1. Statistics of different crash types from 2007 to 2016 in the United States

U.S. rollover crash trends show some evidence of improvement in terms of fatality reduction (see Figure 1). According to the National Highway Traffic Safety Administration (NHTSA) statistics, rollovers made up only 1.7% of traffic crashes in the United States but accounted for 33% of all motor vehicle occupant fatalities. In addition, single-vehicle crashes attributed nearly 60% of the total fatal traffic crash events between 2014 and 2016. A study in New Mexico showed that rollovers accounted for only 5% of total MV crashes, but were responsible for 35% of total fatal MV crashes and 36% of MV occupant fatalities (Chen et al. 2016).

Rollover events are complex as specific characteristics and confounding factors often influence injury outcomes. With regard to the rollover sequence, factors including pre-roll speed (Malliaris and DeBlois 1993), number of ground-to-roof impacts (Digges and Eigen 2004; Parenteau, Gopal, and Viano 2001), number of quarter turns (Digges and Eigen 2004), extent of roof crush (Cohen, Digges, and Nichols 1989), and the primary area of damage (Cohen, Digges, and Nichols 1989) are considered to significantly contribute to rollover crash severity. The number of quarter turns, affected by vehicle shape and deformation during the rollover, generally increase the ejection risk for occupants (Digges and Eigen 2004). According to Richardson et al. (2003), an increase in the number of quarter turns makes the occupant less effectively restrained. Some researchers believe that restraint effectiveness might be affected by the direction of the roll in relation to the occupant's seating position. While some of these factors are granular with regard to the rollover event, they further support the notion of a complex issue.

Discrete choice models have become common in studies of crash injuries. Traditional statistical techniques used to model crash injury severity include the ordered probit model, binary logit model, multinomial logit model, mixed logit model, heteroscedastic ordered logit model, hierarchical logit model, hierarchical ordered logit model, and nested logit model. To analyze the crash injury severity, the random-parameter logit model has been heavily utilized because it relaxed restrictive assumptions associated with traditional discrete outcome models. For example, Ye and Lord (2014) estimated ordered probit, mixed logit, and multinomial logit models to highlight the benefits and limitations of discrete choice models. The authors found that the mixed logit and multinomial logit models outperform the ordered probit model in interpretation power. However, despite accommodating unobserved heterogeneity in the data, a mixed logit model does not take into account the ordinal nature of injury severity outcomes. In light of this, different extensions of ordered probit or logit models with the capability to relax this limitation have been presented. These relatively new techniques, include the generalized ordered logit model, generalized ordered probit model, mixed generalized ordered logit model, spatial generalized ordered response model, random parameter ordered probit or logit model, and random-effects generalized ordered probit model. The literature also showed some promising statistical techniques such as the latent class model and Markov switching model that have been used recently in the analysis of vehicle crash injury severity.

Understanding MV rollover crash risk and injuries is a complex endeavor. Influential factors have been identified among the areas of environment, vehicle, roadway, driver, and event sequence. Numerous studies have focused on the risk factors contributing to rollovers in individual states. Some researchers have developed discrete choice models to identify the factors that may affect injury severity while others have investigated injury patterns in rollover crashes. The prevalence of rollover events in serious crashes prompts further investigation. This study offers insight into the rollover event nomenclature at the national level. Findings may be especially useful in promoting policies, practices, technologies, and programs to prevent future MV rollover crash injuries from a national perspective.

2.2. Methodology

2.2.1. Data Source and Preparation

Single-vehicle rollover crash data was extracted from the Fatality Analysis Reporting System (FARS), which is a nationwide census of yearly data regarding fatal injuries suffered in road crashes that occurred on public roads. Crash data was parsed to include rollover events on interstates and state highways from 2012 to 2016. The rollover crash events include both vehicle-induced and those tripped by objects. A unique identification number was assigned to each crash and associated occupant records were retained for the vehicle in the rollover event. The final data set was scrutinized to eliminate missing and incomplete records with unknown values. Each observation in the final dataset is at the occupant level. It contains information about the characteristics of the driver, environment, and roadway. Table 1 presents the list of variables that were included in the model development. A total of 20,735 valid single-vehicle rollover crash occupant records from 12,082 single-vehicle crashes with rollover events comprise the final data set. Among these records, 7,457 were single occupants, drivers only, while the remaining had one or more passengers. The injury severity variable was used to classify occupants' injury outcomes into four categories: no injury (985 or 4.75%), minor injury (4,160 or 20.06%), serious injury (2513 or 12.12%), and fatality (13077 or 63.07%). It may be noted that these injury counts are given when a fatal rollover crash has occurred. For the analysis purpose of the data, no injury crash outcomes and minor injury crash outcomes were collapsed into a minor or no injury (MNI) class. This new class increased the number of observations at that level so that variability could be reduced in developing the discrete choice models (Haleem and Abdel-Aty 2010).

Table 1. Descriptive statistics of variables included in the model development

Variable name	**Variable description**	**Mean**	**Std. Dev.**
Occupant injury severity			
MNI	If true = 1, otherwise = 0	0.248	0.432
Serious injury	If true = 1, otherwise = 0	0.121	0.326
Fatality	If true = 1, otherwise = 0	0.631	0.483
Crash characteristics			
Ejection	If driver or occupants were partially or fully ejected during rollover = 1, otherwise = 0	0.425	0.494
Speeding related	If speed of vehicle was related to the crash as indicated by law enforcement = 1, otherwise = 0	0.489	0.499
Speed limit	Posted speed limit (kph)	74.076	17.238
Air bag	If air bag was available and deployed for this person = 1, otherwise = 0	0.428	0.495
Location of rollover			
On roadway	If location of the trip point or start of this vehicle's roll was on roadway/carriageway = 1, otherwise = 0	0.124	0.311
On shoulder	If location of the trip point or start of this vehicle's roll was on shoulder = 1, otherwise = 0	0.014	0.119
On median/separator	If location of the trip point or start of this vehicle's roll was on median or separator = 1, otherwise = 0	0.021	0.128
On roadside	If location of the trip point or start of this vehicle's roll was on roadside = 1, otherwise = 0	0.841	0.375
Region			
West	If rollover crash was occurred in the West region = 1, otherwise = 0	0.176	0.381
Southwest	If rollover crash was occurred in the Southwest region = 1, otherwise = 0	0.190	0.393
Midwest	If rollover crash was occurred in the Midwest region = 1, otherwise = 0	0.266	0.442
Northeast	If rollover crash was occurred in the Northeast region = 1, otherwise = 0	0.068	0.253
Southeast	If rollover crash was occurred in the Southeast region = 1, otherwise = 0	0.299	0.458

Table 1. Descriptive statistics of variables included in the model development (continued)

Variable name	**Variable description**	**Mean**	**Std. Dev.**
	Roadway information		
Roadway grade	If roadway grade was level = 1, otherwise = 0	0.620	0.485
Roadway alignment	If roadway alignment was straight = 1, otherwise = 0	0.552	0.497
Pavement type	If roadway surface type was blacktop or bituminous or asphalt = 1, otherwise = 0	0.634	0.482
Number of lanes	If number of travel lanes in each direction was two or greater = 1, otherwise = 0	0.980	0.137
Traffic way	If two opposing traffic flows were separated by median barrier (1 = Present, 0 = Not present)	0.063	0.244
Roadway surface condition	If roadway surface was wet at the time of crash = 1, otherwise = 0	0.123	0.329
Rural/urban	1 if rural, otherwise 0	0.735	0.452
	Environmental information		
Day of week	If crash occurred on weekend = 1, otherwise = 0	0.419	0.493
Weather	If it was raining at the time of crash = 1, otherwise = 0	0.054	0.225
Time of crash			
Daytime	If crash occurred during daytime (0900 hr-1600hr) = 1, otherwise = 0	0.347	0.475
Evening	If crash occurred during evening (1700 hr-0000hr) = 1, otherwise = 0	0.234	0.416
Midnight	If crash occurred during midnight/early morning (0100 hr-0800hr) = 1, otherwise = 0	0.418	0.491
	Lighting conditions		
Daylight	If lighting at the time of crash occurrence was daylight = 1, otherwise = 0	0.398	0.488
Dawn/dusk	If lighting at the time of crash occurrence was dawn or dusk = 1, otherwise = 0	0.043	0.202
Dark, lighted	If lighting at the time of crash occurrence was dark but lighted = 1, otherwise = 0	0.108	0.310
Dark, not lighted	If lighting at the time of crash occurrence was dark but not lighted = 1, otherwise = 0	0.446	0.497
	Driver information		
Driver age	Age of driver	36.312	16.867

Table 1. Descriptive statistics of variables included in the model development (continued)

Variable name	**Variable description**	**Mean**	**Std. Dev.**
Driver gender	Gender of driver (1 if male, otherwise 0)	0.789	0.408
Alcohol/drugs	If driver was under the influence of drug or alcohol = 1, otherwise = 0	0.583	0.493
Seatbelt	If seatbelt was used by driver or occupants = 1, otherwise = 0	0.286	0.452
Previous speed violation	If driver has any speeding convictions within three years of the crash date = 1, otherwise = 0	0.186	0.389
Previous crash	If driver has any crash record within three years of the crash date = 1, otherwise = 0	0.131	0.338
Previous DWI	If driver has any DWI (driving while intoxicated or impaired) convictions within three years of the crash date = 1, otherwise = 0	0.065	0.246
Driver distraction/error before the crash			
No distraction/error	If driver had no distraction or error = 1, otherwise = 0	0.594	0.491
Careless/inattentive	If driver was careless or inattentive = 1, otherwise = 0	0.054	0.225
Cellphone	If driver was using cellphone = 1, otherwise = 0	0.035	0.121
Eating/drinking	If driver was eating or drinking = 1, otherwise = 0	0.001	0.036
Outside event/object/person	If driver was distracted by any outside event or object or person etc. = 1, otherwise = 0	0.003	0.057
Other occupants/moving object	If driver was distracted by other occupants in vehicle or by a moving object in vehicle = 1, otherwise = 0	0.008	0.091
Aggressive driving	If driver was aggressively driving or expressing road rage = 1, otherwise = 0	0.028	0.134
Overloading	If vehicle was overloaded = 1, otherwise = 0	0.003	0.053
Vehicle information			
Vehicle age	If vehicle’s model year is after 2000 = 1, otherwise = 0	0.667	0.471
Towing	If vehicle was towed or pushed improperly = 1, otherwise = 0	0.001	0.025
Tire defect	If vehicle had a tire defect = 1, otherwise = 0	0.002	0.047

Table 1. Descriptive statistics of variables included in the model development (continued)

Variable name	**Variable description**	**Mean**	**Std. Dev.**
Vehicle type			
Passenger car	If vehicle is a passenger car = 1, otherwise = 0	0.377	0.484
SUV	If vehicle is a sport utility vehicle = 1, otherwise = 0	0.312	0.429
Pickup	If vehicle is a pickup or van= 1, otherwise = 0	0.283	0.450
Medium truck	If vehicle is a medium truck (Gross Vehicle Weight Rating < 4535.924 kg) = 1, otherwise = 0	0.013	0.112
Heavy truck	If vehicle is a heavy truck or bus (Gross Vehicle Weight Rating > 4535.924 kg) = 1, otherwise = 0	0.014	0.117

2.3. Model Development

Crash severity is typically measured by the highest level of occupant injury. These measurements represent the "level of severity" on an ordinal scale so ordered discrete choice models are usually applied. Ordered discrete outcome models, however, possess an important basic assumption of proportional odds (Quddus, Wang, and Ison 2009; Long 1997). Parameter estimates remain consistent across all severity levels under this assumption. The Brant test is used to test the violation of this assumption. In case of violation, a simple ordered probit or logit model produces spurious results so a different model should be used. One alternative is the generalized ordered logit model which relaxes the parallel regression assumption while accounting for the ordinal nature of injury severity. This model diverges from the traditional ordered logit model in that the coefficients of the variable are free to vary across severity levels. The generalized ordered logit model is given as follows:

$$P(y > i) = \frac{\exp(\beta_i X_n - \mu_i)}{[1+\exp(\beta_i X_n - \mu_i)]} \qquad \text{(Equation 1)}$$

where β_i is a p×1 vector of regression coefficients for p explanatory variables, X_n is a p×1 vector of explanatory variables, and μ_i and μ_{i-1} are the upper and lower bounds for injury category i. Sometimes, only a subset of variables violates the proportional odds assumption by varying across outcomes. In such a case, the model is referred to as the partial proportional odds model that uses the maximum likelihood estimation technique to estimate the parameters. It can be written as follows:

$$P(y > i) = \frac{\exp(\beta_{1i}X_1 + \beta_2 X_2 - \mu_i)}{[1 + \exp(\beta_{1i}X_1 + \beta_2 X_2 - \mu_i)]} \quad \text{(Equation 2)}$$

where β_1 can vary across different injury severity levels and β_2 remains consistent across each injury severity level i. It has been widely used in the area of road safety.

Various factors regarding crash characteristics, roadway attributes, environment, driver, and vehicle were considered to describe the variation in crash injury severity. The software, SAS© 9.4, was used to estimate the model parameters. The software uses the Akaike information criterion (AIC) and Schwarz criterion (SC) to test the model fit. Fitting was done to produce a robust and parsimonious model. In both methods, a smaller criterion statistic indicates a better model fit.

2.4. Results

The generalized ordered logit model quantifies the effects of various factors on injury severity outcomes for occupants in single-vehicle rollover crashes. Results of the model for serious and fatal injury are presented in Table 2 and Table 3, respectively development. The AIC and SC statistics for the intercept only and fitted model show that predictor variables contribute significantly to understanding rollover event injury outcomes.

Table 2. Results of generalized ordered logit model for serious injury in single-vehicle rollover crashes

Variables	**Coefficient**	**Odds ratio**	**Variables**	**Coefficient**	**Odds ratio**
Ejection	1.92*	3.81 (3.03, 4.90)	No seatbelt (Seatbelt use)	0.27**	1.33 (1.09, 1.72)
On median (On roadway)	0.76***	2.14 (1.12, 4.10)	On roadside (On roadway)	0.47*	1.59 (1.14, 2.28)
Roadway grade (Undulating)	0.18***	1.20 (0.98, 1.47)	Driver age	0.02*	1.02 (1.01, 1.03)
Speeding related	0.30***	1.36 (1.10, 1.68)	Outside event (No error)	1.04***	2.84 (1.18, 4.98)
Aggressive driving (No error)	0.46***	1.58 (1.08, 2.72)	Air bag (No air bag)	-0.56*	1.75 (1.07, 3.28)
SUV (Passenger car)	-0.32*	0.73 (0.57, 0.93)	Pickup (Passenger car)	-0.59*	0.55 (0.43, 0.71)
Southwest (Northeast)	0.46***	1.58 (1.01, 2.50)	Midwest (Northeast)	0.43**	1.54 (1.06, 2.63)
Ejection × Seatbelt	0.83***				

Table 3. Results of generalized ordered logit model for fatal injury in single-vehicle rollover crashes

Variables	**Coefficient**	**Odds ratio**	**Variables**	**Coefficient**	**Odds ratio**
Ejection	2.79*	9.92 (9.12, 11.34)	No seatbelt (Seatbelt use)	0.62*	1.92 (1.62, 2.17)
On roadside (On roadway)	0.44*	1.49 (1.12, 1.96)	Roadway grade (Undulating)	-0.25*	1.28 (1.12, 1.48)
Pavement type (Rigid pavement)	0.20**	1.22 (1.09, 1.44)	Rural/urban (Urban)	0.14**	0.87 (0.73, 0.98)
Day of week (Weekday)	-0.15**	0.86 (0.74, 0.97)	Evening (Daytime)	0.23**	1.44 (1.27, 1.59)
Midnight (Daytime)	-0.23***	0.79 (0.67, 0.93)	Dawn/dusk (Daylight)	0.19***	1.39 (1.06, 1.96)
Driver age	0.05*	1.056 (1.05, 1.061)	Speeding related	-0.02**	1.10 (1.01, 1.19)
Speed limit	0.02**	1.01 (1.001, 1.02)	Previous crash (No previous crash record)	0.30***	1.36 (1.10, 1.68)
Careless/inattentive (No error)	0.32*	0.72 (0.59, 0.87)	Other occupants (No error)	2.62*	0.77 (0.51, 0.98)
Vehicle age (Vehicle model before 2000)	-0.14***	1.39 (1.01, 2.13)	Air bag (No air bag)	-0.45*	1.57 (1.02, 3.06)
SUV (Passenger car)	-0.41*	0.66 (0.55, 0.78)	Pickup (Passenger car)	-0.49*	0.61 (0.52, 0.72)
Southwest (Northeast)	0.39**	1.47 (1.08, 2.02)	Midwest (Northeast)	0.38**	1.38 (1.10, 1.96)
Ejection × Speed limit	0.01***				

Note: Coefficient of Intercept for Serious injury = -0.985** and Coefficient of Intercept for Fatality = -0.897**

Number of observations = 20,735

AIC (Intercept only) = 12200.74 and AIC (Intercept and covariates) = 10183.35

SC (Intercept only) = 12215.26 and SC (Intercept and covariates) = 10894.62

Parenthesis contains the "reference category" for variables and the "95% confidence interval" for odds ratio. The reference category for injury severity is "MNI".

*Significant at 0.01 significance level; **Significant at 0.05 significance level; ***Significant at 0.10 significance level.

The model results also showed a significant two-way interaction term between occupant's ejection and speed limit. Figure 2 shows predicted probabilities for injury severity of occupant's ejection against speed limit while considering the reference categories for remaining predictors.

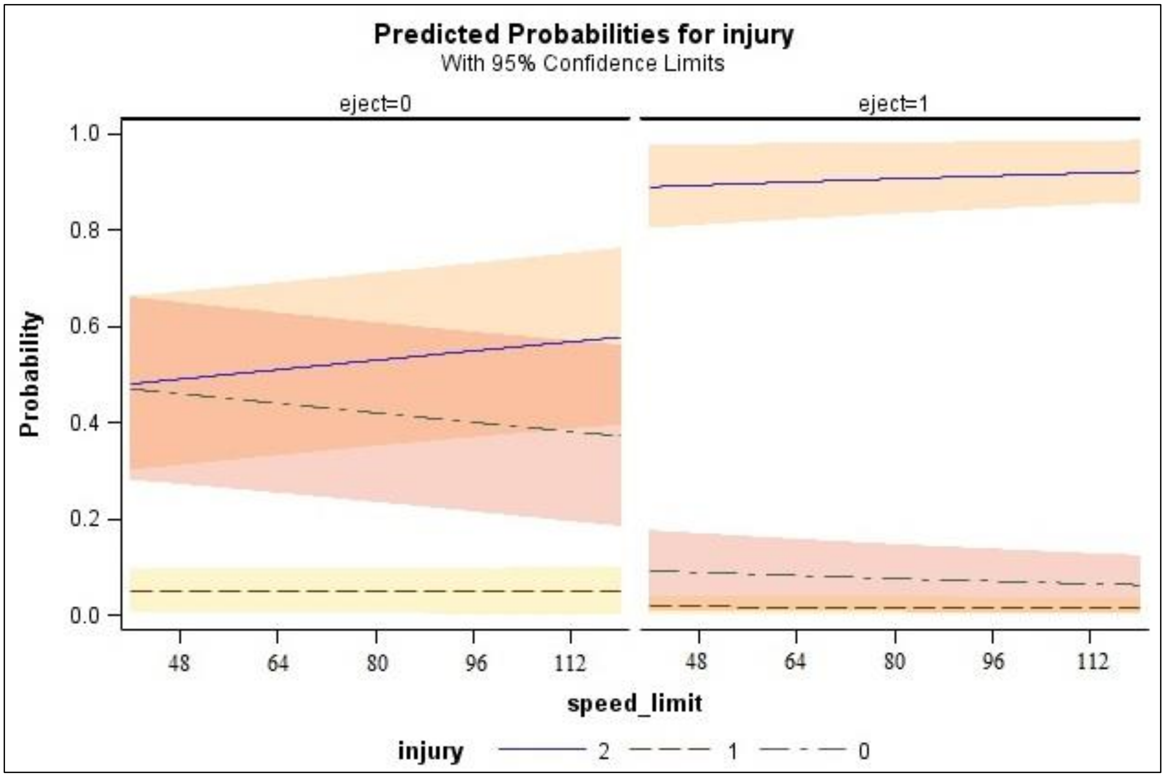


Figure 2. Predicted probabilities for injury severity of occupant's ejection versus speed limit
Note: 0 = MNI, 1 = Serious injury and 2 = Fatality

The other significant two-way interaction term was between occupant's ejection and seatbelt use. Predicted probabilities for injury severity of occupant's ejection against seatbelt use are shown in Figure 3.

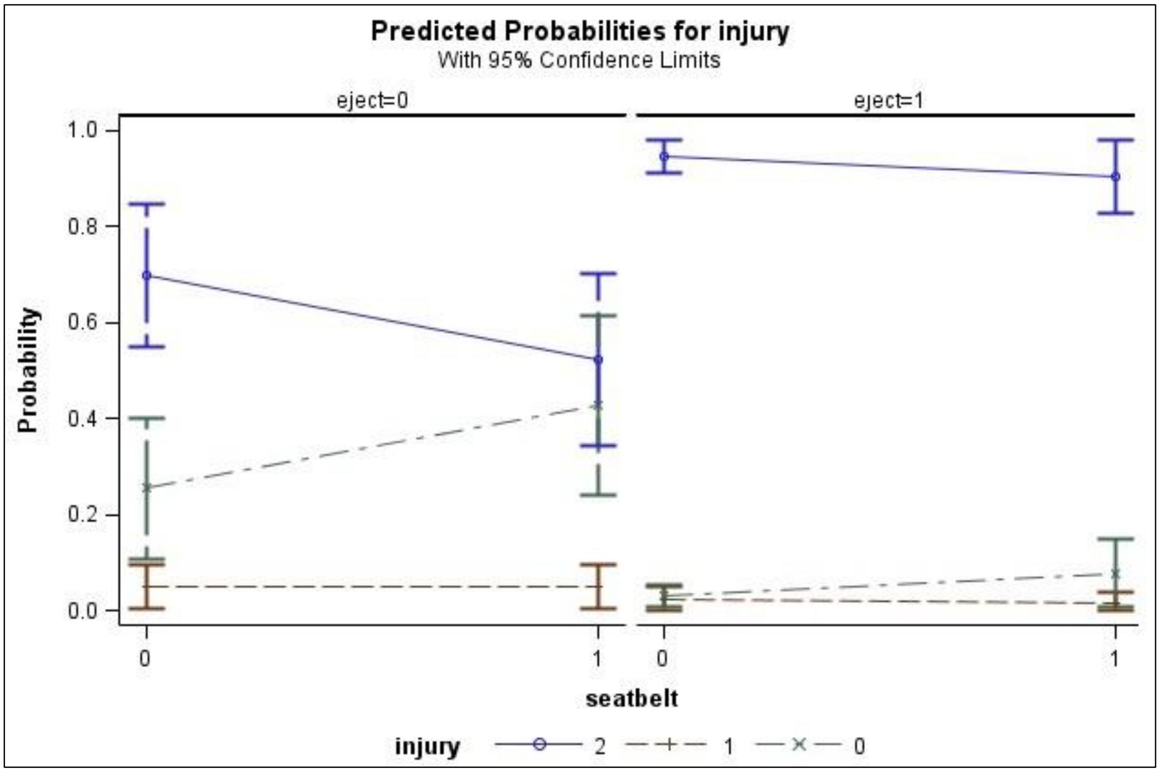


Figure 3. Predicted probabilities for injury severity of occupant's ejection versus seatbelt use
Note: 0 = MNI, 1 = Serious injury and 2 = Fatality

2.5. Discussion

The model results showed that the ejection (both partial and complete) of an occupant during a rollover was more likely to result in serious and fatal injuries while considering the use of a seatbelt. The odds of sustaining a fatal injury in a rollover crash for a partially or fully ejected occupant were almost 10 times as large as that for an unejected occupant. This finding is intuitive because the impact of partial and complete ejection on the risk of serious and fatal injury overshadows the influence of all other factors. The ejected occupant is at high risk of serious injury from striking the ground or being crushed during a rollover. This finding is consistent with previous research conducted by Funk et al. (2012) which states that completely ejected occupants are 20 times and 91 times more likely to be seriously injured or die, respectively, than those not ejected.

Seatbelt use was found to have a significant influence on the injury outcomes in rollover crashes. Occupants not using a seatbelt were twice as likely to suffer serious and fatal injuries. This result is intuitive and consistent with previous studies (Yasmin, Eluru, and Pinjari 2015). Since there is a correlation between seatbelt use and ejection, the risk of partial and complete ejection during a rollover is reduced due to seatbelt use. Hence, the possibility of sustaining serious and fatal injuries is also reduced.

Regarding the injury severity related to the initial location in a rollover, serious injuries were likely to be observed in crashes where the vehicle rollover began at the median or separator. Similarly, rollovers that began on the roadside also were likely to result in serious and fatal injuries. It is common for rollover crashes to occur off the roadway where the vehicle is more likely to strike barriers or fixed objects and result in serious injuries. The crash could be a tripped rollover, which occurs when a vehicle slides sideways upon leaving the roadway and its tires dig into soft soil or strike an object such as a guardrail or curb. NHTSA statistics show that 95% of single-vehicle rollovers are the tripped type. Also, the downward steep slope of the roadside will often move the vehicle's center of gravity outboards resulting in a rollover.

The flat or level terrain indicator was less likely to be associated with fatal injury than undulating terrain. These results are similar to findings in previous studies (Malliaris and DeBlois 1993). The type of pavement surface was also associated with rollover crash likelihood. A fatal injury was more likely on a blacktop or bituminous road but exposure may explain it. Blacktop is the most common surface for U.S. highways. With regard to land use, urban areas were less likely to experience fatal rollovers than rural areas. Higher posted speed limits, less traffic control, longer emergency response times, less prevalent enforcement, and less forgiving road design are typically associated with rural settings compared to urban settings. All of these characteristics are likely to be associated with a greater risk of fatal rollover crashes.

Fatal rollover crashes were more likely on weekdays, which may be associated with higher traffic volume during the work week. Time of the day was also associated with rollover injury severity. A greater likelihood of fatal rollover crashes was during the evening. Fatigued driving during evening hours and less late night traffic may explain these findings. Lighting conditions also affected the injury severity of rollovers. Crashes occurring at dawn or dusk were more likely to result in fatal injuries. This finding corresponds to the previous result. This increased risk may be attributed to poor lighting conditions at sunset and sunrise or to low-angle sunlight shining directly into drivers' eyes. In addition, less contrast in colors, reduction in natural light, and less effectiveness of vehicle headlights during sunset and sunrise may also affect drivers' vision.

The relationship between vehicle speed and rollover crashes was also significant as higher speeds are more likely to be associated with serious injuries. Speeding behavior, as indicated by law enforcement, includes racing, exceeding the speed limit, and speeding too fast for conditions. This finding could explain the greater number of quarter turns after rollover initiation and the likely greater vehicle rotation rate. Both turn and rotation factors have been associated with greater injury severity (Conroy et al. 2006). This finding supports education and enforcement programs that encourage drivers to obey speed limits. Similarly, the speed limit has a significant impact on injury severity. Higher posted speed limits tend to be associated with more serious and fatal injury outcomes compared to lower speed limits. In case of rollover occurrence at higher speed limits, the likelihood of occupant's partial or complete ejection increases. A similar finding can be found in a study conducted by Keall and Newstead (2009).

They showed that the risk of rollover crashes in areas with higher posted speed limits is higher than in other areas.

Previous crash involvement for the driver was associated with an increased likelihood of a fatal injury outcome. The use of this explanatory variable is relatively new in rollover injury severity modeling because it was recently added to the national dataset. This finding requires further investigation because of the key role of drivers in single-vehicle rollover crashes. The possible classification of previous crash types for the driver might help in understanding which behavior provides an opportunity for preemptive intervention.

The odds of sustaining a serious injury in a rollover crash for a driver distracted by an outside event or object were almost three times greater than the odds for an attentive driver. Similarly, driver interaction with passengers was more likely to result in fatal injury. Both findings related to driver distraction are consistent with previous work (Braitman and Braitman 2017). Talking with passengers is among the most commonly reported distractions. Aggressive driving also was significantly related to serious injury outcomes, a piece of additional evidence that driver behavior plays a significant role in injury outcomes. However, the nature of FARS data might be biased toward serious crashes because it is difficult for the driver to explain the surrounding circumstances.

Vehicle age significantly influenced injury severity in rollover crashes. Vehicle models manufactured before 2000 were more likely to result in fatal injuries compared to newer vehicles. Implementation of Federal Motor Vehicle Safety Standards and safety technologies including electronic stability control, roof strength, and advanced air bags have played a significant role in improving the overall crashworthiness of vehicles. Federal legislation made air bags mandatory for all cars and light vehicle trucks in 1998. The rule by FMVSS in 2000 to improve the design of air bags to reduce the risk of airbag-induced injuries has been effective in reducing crash injuries.

SUVs and pickups were less likely to be associated with serious and fatal rollover injuries. Although light trucks such as pickups and SUVs are more likely to roll over because they have a higher center of gravity and are less stable than passenger cars, they also protect occupants during rollovers because of their greater mass and crashworthiness (Khattak and Rocha 2003). Khattak and Rocha (2003) stated that SUVs reduce injury severity in rollover crashes because the protection they offer dwarfs any injurious effects. A region variable was used to capture geographic differences in rollover injury outcomes. Compared to the Northeast, rollover crashes in the Southwest and Midwest were more likely to result in serious and fatal injuries. The West and Southeast did not differ significantly from the Northeast.

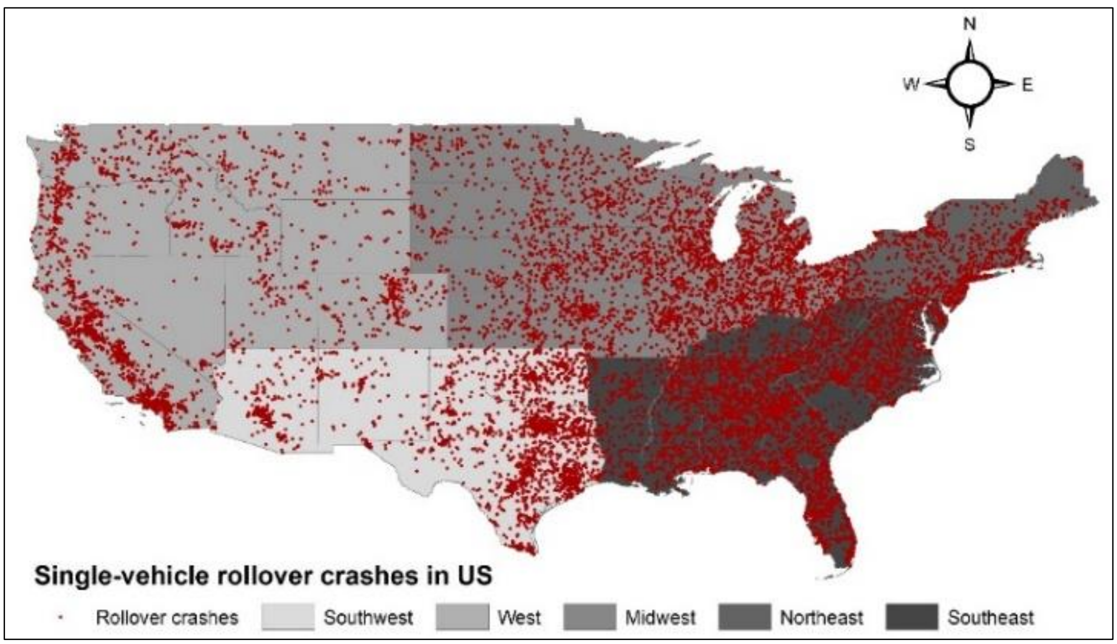


Figure 4. Single-vehicle rollover crashes in the United States from 2012 to 2016

This finding can be attributed to several reasons. The average posted speed limits on both rural and urban interstates are 116 kph, 110 kph, and 102 kph in Southwest, Midwest, and Northeast, respectively. The regional variations in posted speed limits might affect injury severity because the speed limit is associated with rollover crash injury severity as discussed earlier. The other differences among regions that affect rollover crash injury severity might include driving characteristics, topography, and driving laws. Also, inconsistencies among states for reporting rollover crash data by law enforcement cannot be overlooked. To this end, further investigation of variations among states is warranted. Figure 4 shows the spatial dispersion of single-vehicle rollover crashes in different regions of the US from 2012 to 2016.

The model results also showed a significant two-way interaction term between occupant's ejection and speed limit. For an unejected occupant, the likelihood of fatal injury increases with a higher speed limit. On the other hand, at all speed limits, the probability of fatal injury for ejected occupant is much higher than serious injury. The other significant two-way interaction term was between occupant's ejection and seatbelt use. The results indicate that seatbelt use decreased the likelihood of fatal injury for an unejected occupant.

To summarize the discussion, a generalized ordered logit model was developed to study the injury severity of occupants in single-vehicle rollover crashes. Increased likelihood of serious and fatal injuries in rollovers is associated with occupant ejection, speeding, higher posted speed limits, roadside and median rollovers, undulating terrain, rural roads, evening, not using seat belts, previous driver crash, and careless driving. Air bags reduce the likelihood of severe injury. In addition, the study showed regions vary with regard to injury severity risk. Findings suggest countermeasures that may reduce the injury severity of single-vehicle rollover crashes include seatbelt use, lower speed limits and heightened speed enforcement in high-risk areas, flattening roadside embankments, and in-vehicle stability-enhancement systems.

Efforts to promote the use of seatbelts as the first line of defense against rollover injuries need to be continued. Crashworthiness vehicle fleet will improve safety as older vehicles are replaced. The behavioral characteristics of drivers, such as those with a history of unsafe driving, could be targeted in education campaigns, policy decisions, and community awareness activities. Additional studies into the effect of driver education, behavior, and experience related to rollover crash events may also be useful in developing and deploying effective countermeasures.

A limitation of this study is that the fatal crash event scope may bias the results in terms of generalizing findings for other injury levels. Also, the scope of national and state highways may limit the transferability of findings to highly rural areas. Moreover, rollover crashes that include multiple occupants would include a range of injury levels, yet all of the crash, vehicle, and environmental factors would be identical. The study also did not account for multiple occupant effects.

2.6. Acknowledgments

I would like to thank the North Dakota Department of Transportation for sponsoring this research. The work presented in this paper was also supported by North Dakota State University and the Mountain-Plains Consortium, a University Transportation Center funded by the U.S. Department of Transportation. Khan thanks the National University of Sciences and Technology (NUST, Pakistan) for sponsoring his PhD program.

2.7. References

Braitman, Keli A, and Abby L Braitman. 2017. "Patterns of Distracted Driving Behaviors among Young Adult Drivers: Exploring Relationships with Personality Variables." *Transportation Research Part F: Traffic Psychology and Behaviour* 46: 169–76. https://doi.org/10.1016/j.trf.2017.01.015

Chen, Cong, Guohui Zhang, Zhen Qian, Rafiqul A Tarefder, and Zong Tian. 2016. “Investigating Driver Injury Severity Patterns in Rollover Crashes Using Support Vector Machine Models.” *Accident Analysis & Prevention* 90: 128–39. https://doi.org/10.1016/j.aap.2016.02.011

Cohen, Daniel, Kennerly Digges, and R. Hunter Nichols. 1989. “Rollover Crashworthiness Classification and Severity Indices.” http://wbldb.lievers.net/10125397.html.

Conroy, Carol, David B. Hoyt, A. Brent Eastman, Steve Erwin, Sharon Pacyna, Troy Lisa Holbrook, Teresa Vaughan, Michael Sise, Frank Kennedy, and Tom Velky. 2006. “Rollover Crashes: Predicting Serious Injury Based on Occupant, Vehicle, and Crash Characteristics.” *Accident Analysis & Prevention* 38 (5): 835–42. https://doi.org/10.1016/J.AAP.2006.02.002.

Digges, Kennerly H, and Ana Maria Eigen. 2004. “Crash Attributes That Influence the Severity of Rollover Crashes.” https://pdfs.semanticscholar.org/76e0/0a32d1fce93565c00afd36b5315409ac7999.pdf.

Fatality Analysis Reporting System. Accessed June 11, 2018. https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars.

Funk, James R, Joseph M Cormier, Charles E Bain, Jeffrey L Wirth, Enrique B Bonugli, and Richard A Watson. 2012. “Factors Affecting Ejection Risk in Rollover Crashes.” *Annals of Advances in Automotive Medicine. Association for the Advancement of Automotive Medicine. Annual Scientific Conference* 56: 203–11. http://www.ncbi.nlm.nih.gov/pubmed/23169130.

Haleem, Kirolos, and Mohamed Abdel-Aty. 2010. “Examining Traffic Crash Injury Severity at Unsignalized Intersections.” *Journal of Safety Research* 41 (4): 347–57. https://doi.org/10.1016/j.jsr.2010.04.006

Jehle, Dietrich, Joseph Kuebler, and Peggy Auinger. 2007. “Risk of Injury and Fatality in Single Vehicle Rollover Crashes: Danger for the Front Seat Occupant in the ‘Outside Arc.’” *Academic Emergency Medicine* 14 (10): 899–902. https://doi.org/10.1197/j.aem.2007.06.029.

Keall, Michael, and Stuart Newstead. 2009. “Induced Exposure Estimates of Rollover Risk for Different Types of Passenger Vehicles.” *Traffic Injury Prevention* 10 (1): 30–36. https://doi.org/10.1080/15389580802383117

Khattak, Asad, and Marta Rocha. 2003. “Are SUVs" Supremely Unsafe Vehicles"?: Analysis of Rollovers and Injuries with Sport Utility Vehicles.” *Transportation Research Record: Journal of the Transportation Research Board*, no. 1840: 167–77. https://doi.org/10.3141/1840-19

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Sage Publications. https://books.google.com/books/about/Regression_Models_for_Categorical_and_Li.html?id=CHvSWpAyhdIC.

Malliaris, A C, and J H DeBlois. 1993. "Pivotal Characterization of Car Rollovers," July. https://trid.trb.org/view/409125.

Parenteau, Chantal, Madana Gopal, and David C. Viano. 2001. "Near and Far-Side Adult Front Passenger Kinematics in a Vehicle Rollover." In . https://doi.org/10.4271/2001-01-0176.

Quddus, Mohammed A, Chao Wang, and Stephen G Ison. 2009. "Road Traffic Congestion and Crash Severity: Econometric Analysis Using Ordered Response Models." *Journal of Transportation Engineering* 136 (5): 424–35. https://doi.org/10.1061/(ASCE)TE.1943-5436.000004

Richardson, S, R H Grzebieta, and G Rechnitzer. 2003. "Proposal for a Dynamic Rollover Protective System Test." *International Journal of Crashworthiness* 8 (2): 133–41. https://doi.org/10.1533/ijcr.2003.0215.

Vision Zero NDDOT. Accessed September 27, 2021. https://visionzero.nd.gov/strategies/LaneDeparture/.

Yasmin, Shamsunnahar, Naveen Eluru, and Abdul R Pinjari. 2015. "Analyzing the Continuum of Fatal Crashes: A Generalized Ordered Approach." *Analytic Methods in Accident Research* 7: 1–15. https://doi.org/10.1016/j.amar.2015.03.001

Ye, Fan, and Dominique Lord. 2014. "Comparing Three Commonly Used Crash Severity Models on Sample Size Requirements: Multinomial Logit, Ordered Probit and Mixed Logit Models." *Analytic Methods in Accident Research* 1 (January): 72–85. https://doi.org/10.1016/J.AMAR.2013.03.001.

CHAPTER 3: HOTSPOT ANALYSIS OF SINGLE-VEHICLE LANE DEPARTURE CRASHES IN NORTH DAKOTA

3.1. Introduction

Lane departure crashes are a primary concern in the United States. The Federal Highway Administration (FHWA, 2021), defines a lane departure crash a one that occurs after a motor vehicle crosses a center line or an edge line, or otherwise leaves a roadway. From the driver's perspective, four components factor into these complex crashes: human error, vehicle component failure, roadway condition, and collision avoidance. About 51% of all U.S. traffic fatalities were attributed to lane departure crashes from 2016 to 2018 (FHWA, 2021). These crashes tend to be grievous, mostly occurring at relatively high speeds.

The rural lane departure crash is the most common type of fatal traffic crash in North Dakota. Nearly 90% of fatal lane departure crashes between 2015 and 2019 occurred on rural roads (Vision Zero, NDDOT). In addition, 77% of these crashes involved a single-vehicle. The two most common harmful events in single-vehicle fatal lane departure crashes were rollover/overturning (77%) and collision with fixed objects (11%). Figure 5 shows that single-vehicle lane departure crashes in North Dakota remain prevalent. The extensive rural road network and episodic crash event nature isan impediment to substantial improvement for rural roads. Moreover, these roads are often managed by local agencies that have limited resources and inadequate technical skills for safety analysis. This lack of analysis presents a challenge in determining where to apply lane departure countermeasures to effectively reduce fatalities and injuries.

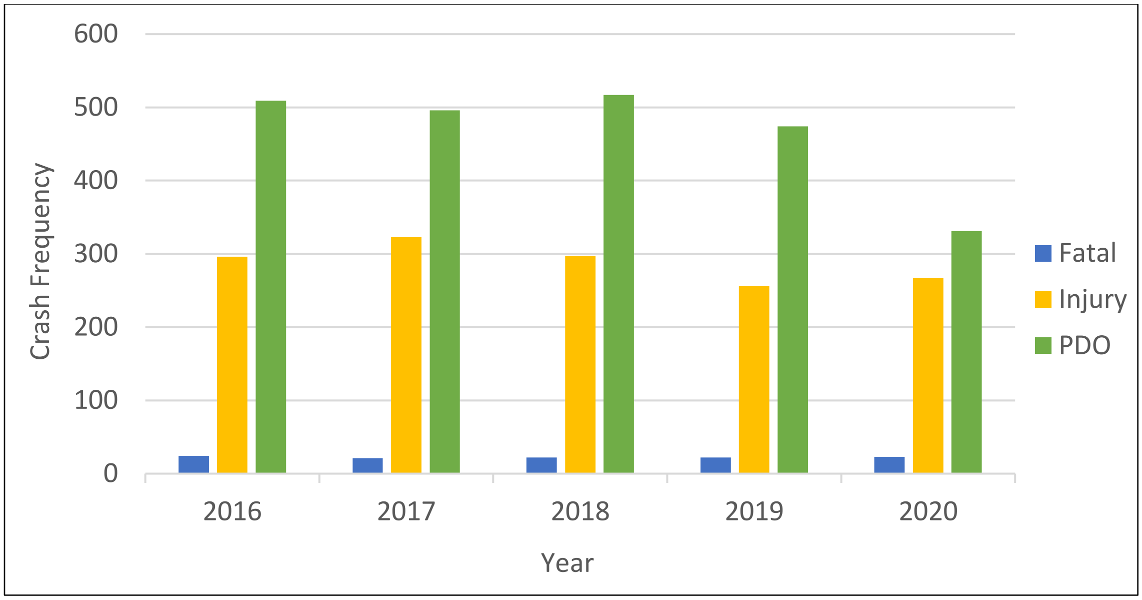


Figure 5. Crash statistics of single-vehicle lane departure crashes in ND from 2016 – 2020

Thus, tools to detect high-risk locations and ultimately deploy appropriate strategies are essential in moving traffic crash trends downward. Visual identification tools are evolving as a crucial technique. Hotspots, or crash-prone locations, are areas having clusters of relatively high-crash concentrations. Detecting these hotspots provides an opportunity to be more granular in studying and addressing crash causes. Refined knowledge is used to allocate limited resources more efficiently by prioritizing high-risk locations. The objective here was to unveil single-vehicle lane departure hotspot areas that are at a higher-risk for crashes across the state's rural road system.

3.2. Literature Review

It is important to examine crash locations and time in dispersal. One can then prioritize locations based on attribute clusters. Inherent geographical qualities increase the chance of a crash occurrence at certain locations. In other words, traffic crashes do not take place randomly. The tendency of crashes to be concentrated at given locations can be explained by factors such as land use, socio-economic parameters, and geometric design. Hotspots or black spots refer to high-concentration crash location clusters. The systematic strategy of crash hotspots is commonly used to identify crash-prone locations, establish ranking, reveal causation, and determine countermeasures on identified locations (Moons, Brijs, and Wets 2009; Dereli and Erdogan 2017). The primary purpose of hotspot analysis is to give decisionmakers network insight into the select and deployment of safety measure(s) for crash prevention (L. Li, Zhu, and Sui 2007; Anderson 2009; Vemulapalli et al. 2017; Xie and Yan 2013).

One way to conduct hotspot analysis is to identify locations by simply observing crash location maps. However, the results are subjective, providing weak support for decisions. Two empirical techniques have emerged in robust hotspot analysis for crash prevention. The first uses traditional statistical methods such as regression models (Vogt and Bared 1998; Zhang and Ivan 2005), empirical Bayesian (Elvik 2008; W. Cheng and Washington 2005), and full Bayesian (Sacchi, Sayed, and El-Basyouny 2015; Huang, Chin, and Haque 2009) for model-based analysis. The second technique is based on geostatistical analysis such as the spatial autocorrelation method (Getis and Ord 1992; Ord and Getis 1995; Moran 1950; Anselin 1995) or density estimation method (Sabel et al. 2005; Erdogan et al. 2008; Kaygisiz et al. 2015; Kuo, Zeng, and Lord 2012).

Generally, hotspot analysis based on a geostatistical approach is preferred. This approach requires less complex data and uses straightforward computation, unlike traditional statistical methods (Thakali, Kwon, and Fu 2015). In addition, the geostatistical method provides a visual representation of results, incorporating spatial factors to detect location-specific influences on crash occurrence (Mitra 2008). The Density estimation method is most commonly used in crash pattern detection, but it analyzes the crash location with no consideration for attributes. On the other hand, advanced approaches, like the spatial autocorrelation method, account for location as well as crash attributes. Hence, the underlying concept of spatial autocorrelation is associated with interdependence of a particular attribute over space.

Two categories of spatial autocorrelation are 1) global spatial analysis and 2) local spatial analysis. Global spatial analysis measures and tests the overall spatial phenomenon and identifies if the feature pattern is clustered, dispersed, or randomly distributed in space with respect to their attribute values (Getis and Ord 1992; Moran 1950). Local spatial analysis measures the level of spatial association at the local scale (Ord and Getis 1995; Anselin 1995; Moons, Brijs, and Wets 2009). It is preferred over global spatial analysis for crash hotspot identification because local spatial patterns are undetectable when generalized for a large area (Songchitruksa and Zeng 2010). Hence, local spatial analysis is used mostly for analyzing crash hotspots to capture regional heterogeneity across the study area.

The two most common approaches in local spatial analysis are local Moran's I (Anselin 1995; Dezman et al. 2016; Moons, Brijs, and Wets 2009) and Getis-Ord Gi* (Ord and Getis 1995; Zubaidi et al. 2021; Hazaymeh et al. 2022). The local Moran's I has the ability to identify local clusters but cannot differentiate the high-valued from low-valued clusters. This limitation was overcome by the Gi* spatial statistic approach that distinguishes low- and high-valued clusters (Ord and Getis 1995; Getis and Ord 1992; Chance Scott, Sen Roy, and Prasad 2016). Another approach for hotspot analysis is the space-time cube (STC) analysis (Z. Cheng, Zu, and Lu 2018; Kveladze, Kraak, and van Elzakker 2013). STC is a 3D visual representation of a geographical phenomenon such that the horizontal plane of the cube (x and y) represents space and the vertical axis represents time (Kveladze, Kraak, and van Elzakker 2013). This approach can show spatial patterns, spatial relationships, and changes over time. Various studies have used 3D visualizations for analysis of a wide array of spatial datasets such as atmospheric pollution, crime, crashes, and diseases (Demšar and Virrantaus 2010; Fang and Lu 2011; Nakaya and Yano 2010; Sadler et al. 2017; Yoon and Lee 2021; Kang, Cho, and Son 2018; Purwanto et al. 2021).

According to Kveladze et al. (Kveladze et al. 2018), understanding the underlying meaning of complex spatiotemporal data is difficult for users. However, spatial aggregation techniques, especially point-based data, can provide a solution. The integration of hotspots with STC is referred to as emerging hotspot analysis. Both techniques are integral in deploying crash cluster decision tools. Ample literature on the use of all these methods for spatial analysis with crash detection is available but few studies have considered crash severity when identifying crash hotspots. The objective here was to apply hotspot techniques to distinguish crash clusters across the road network based on injury severity. Stakeholders can use results in data-driven decisions to prioritize locations and strategies for preventing lane departure crashes.

3.3. Scope and Data

A total of 19,162 records for motor vehicle crashes that occurred on rural roads from 2016 to 2020 were collected from the state of ND. At least three years of crash data is needed to effectively conduct spatial analysis (W. Cheng and Washington 2005). Information regarding spatial location, crash, driver, vehicle, and geometrical characteristics was collected. Records were parsed to capture single-vehicle lane departure crashes on rural roads. Subsequently, rural road functional classes such as minor arterial, major collector, minor collector, and local roads were retained. This data deduction process resulted in a total of 3,878 single-vehicle lane departure crashes that occurred on local roads. Table 4 shows the crash characteristics of the data.

Table 4. Crash characteristics of the data

Variable name	Description	Crash severity		
		PDO	Minor injury	Serious and fatal injury
Road functional classification	Minor arterial	374	220	18
	Major collector	679	533	48
	Local	1274	686	46
Seatbelt use	Yes	1539	773	16
	No	129	423	87
	Unknown	659	243	9
Driver gender	Male	1587	1079	89
	Female	555	354	23
	Unknown	185	6	0
Speeding/too fast for conditions	Yes	846	644	38
	No	1481	795	74
Driver average age		34	35	42
Total crashes	3878			

Four categories were differentiated, property damage only, minor injury, serious injury, and fatal injury, for a spectrum of less- to more severe crash outcomes, respectively. The crash frequency with serious and fatal injury categories was limited, so these categories were merged. The transformed data contained three severity levels: 2,418 property damage crashes, 1,075 minor injury crashes, and 385 serious or fatal injury crashes. A roadway network layer was created based on the geographic coordinate system and was projected in metric linear units. All the crashes with latitude and longitude were projected on the roadway network. SAS software 9.4 was used to clean and prepare the data. ArcMap 10.4, SANET, and ArcScene 10.4 were used for analysis and visualization

3.4. Methodology

Spatial statistical methods were applied to analyze single-vehicle lane departure crashes to reveal black zone areas. Both global and local spatial analyses were used to determine significant clusters. For each crash location and the surrounding crashes, the spatial autocorrelation was performed using the linear spatial weight matrix. Approaches are described in the following sub-sections.

3.4.1. Global Spatial Autocorrelation

The spatial autocorrelation known as Global Moran's I detects general spatial patterns by considering both the location and feature values. Accounting for a set of features and related attributes, this approach examines if the pattern is clustered, random, or dispersed. Mathematically, Global Moran's I can be expressed as follows:

$$\mathrm{I} = \frac{\mathrm{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} W_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\mathrm{S_o}\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{x_i}-\bar{\mathrm{x}})} \quad \forall_i = 1,\dots,n \wedge \forall_j = 1,\dots,n \quad \text{(Equation 3)}$$

Where x_i is the value of the feature on location i, $\bar{x}$ is the feature mean, n is equal to the total number of locations, w_{ij} is the spatial weight representing connectivity relationships between features i and j, and S_o is the sum of all spatial weights.

$$S_{\mathrm{o}} = \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij} \qquad \text{(Equation 4)}$$

The results from the Global Moran's I statistics are interpreted with respect to its null hypothesis which states that among the features, an attribute is randomly distributed in the study area. The statistical significance of this test is calculated from the Z score assuming normal distribution with mean and variance equal to zero and one respectively. A positive Z score implies that the feature is surrounded by similar values. A negative Z score shows that the neighboring features have different values (Ord and Getis 1995; Getis and Ord 1992).

3.4.2. Local Spatial Autocorrelation

The local Moran's I index is used to identify local clusters and local spatial outliers of crashes in the study area. It can be written as follows:

$$I_i = \frac{Z_i - \bar{Z}}{\sigma^2} \sum_{j=1, j \neq 1}^{n} [W_{ij}(Z_i - \bar{Z})] \qquad \text{(Equation 5)}$$

Where I_i is the local Moran's I coefficient, z_i is the value of the feature on location i, $\bar{z}$ is the feature mean, z_j is the value at all other locations such that $j \neq i$, and σ^2 is the variance of z. A high positive value of the local Moran's I index indicates that the location under study is surrounded by features with similar values i.e. presence of a spatial cluster. On the other hand, a high negative value of local Moran's Index shows the presence of spatial outliers i.e. the site under study is surrounded by features with different values. Local Moran's index generates four types of results: 1) high-high clusters – high values surrounded by a high-value neighbored, 2) low-low clusters - low values surrounded by a low-value neighbored, 3) high-low outlier – a high value surrounded by low-value neighbored, and 4) low-high outlier – a low value surrounded by high-value neighbored. Note that high-high clusters are considered to be crash hotspots while high-low outliers are considered to be isolated individual hotspots.

Another approach for local spatial autocorrelation is based on Getis-Ord Gi* statistics. This method distinguishes between the areas of high and low value concentration on the local scale. It determines high or low corresponding values by comparing each feature with its neighboring features. For a location to be considered as a hotspot, both the feature itself as well as the features surrounding it should have high attribute values. In ArcMap, the hotspot analysis tool uses this statistic to determine a significant red/blue spot based on neighbors' feature values. The following equation shows the general form of G_i^*,

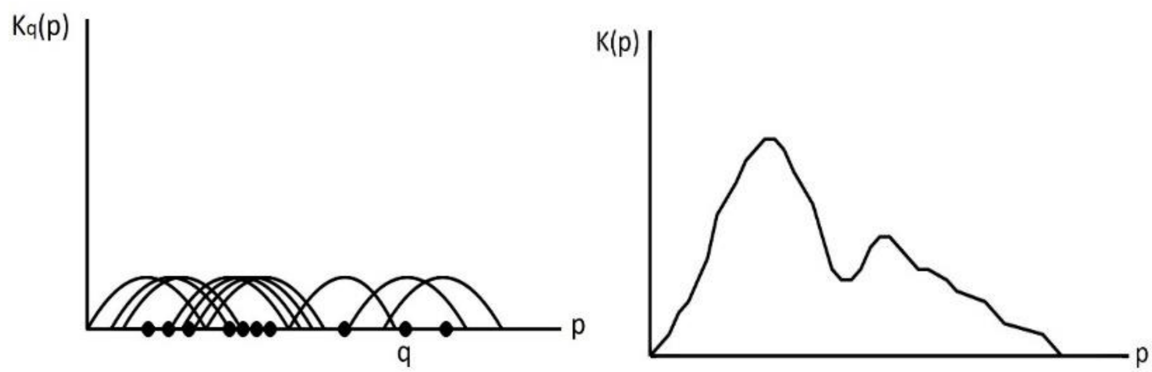
$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij} x_i - \left(\frac{\sum_{j=1}^{n} x_j}{n}\right) \times \sum_{j=1}^{n} w_{ij}}{\sqrt{\frac{\sum_{j=1}^{n} x_j^2 - \sum_{j=1}^{n} x_j}{n}} \times \sqrt{\frac{n \sum_{j=1}^{n} w_{ij}^2 - (\sum_{j=1}^{n} w_{ij})^2}{n-1}}} \qquad \text{(Equation 6)}$$

Where G_i^* is a statistic which describes the spatial dependency of feature i and x_j is an attribute value of j feature in the neighborhood. The G_i^* statistic represents the value of a target feature in the form of a z-score. Features with a significant positive z-score indicate that neighboring values are similar while those with negative z-scores imply that nearby values are dissimilar. The size of the z-score in either direction of association shows the magnitude of clustering. In other words, a higher positive or a smaller negative z-score represents a cluster concentration for each corresponding feature. A suitable calculation of the spatial relationship between features is integral for local spatial autocorrelation analysis (O'Sullivan and Unwin 2010). Some different approaches available for spatial relationships include fixed distance band, inverse distance, K nearest neighbors, and space-time window. Previous research shows that the fixed distance band approach is preferred for analyzing point datasets (Mitchell 2021). This study used a fixed distance band because the study was analyzing crashes.

The degree of spatial autocorrelation changes according to distance bands because of unique groups of neighbors. Knowing the precise nature of spatial crash patterns at a specific location is a complex phenomenon. Alternatively, previous research suggests using the optimum distance value that establishes the maximized spatial autocorrelation (Mitchell 2021; Maingi et al. 2012). Clustering patterns of crashes will be more evident when the distance used in calculation provides an enlarged degree of spatial autocorrelation across the area of interest. For this purpose, a tool known as incremental spatial autocorrelation in ArcMap measures the variation in spatial autocorrelation as the distance bandwidths change. For a set of increasing distances, it produces several z-scores related to corresponding Moran's I. At a critical distance bandwidth, a peaked z-score shows spatial autocorrelation which forms the most prominent clusters (Mitchell 2021; Maingi et al. 2012). But before using incremental spatial autocorrelation, the question of what starting distance should be used arises. A good starting distance would be the distance at which any given point has at least one neighbor. For this, a tool in ArcMap known as “calculate distance band from neighbor count” was used. Crashes were used as input features with neighbors equal to one in order to calculate the average distance.

3.4.3. Network Kernel Density Estimation

There is another non-parametric approach for hotspot analysis known as kernel density estimation (KDE). This method provides a high-quality visual representation of density estimates from observed data. It is known as KDE because a circular area (the kernel) of defined bandwidth is created around each point at which the indicator is observed. Figure 6 shows a 2D visualization of the kernel density estimation method. For each point, an individual kernel or molehill is created in the form of a smooth and continuous-density surface. Then, for a given cell, all overlapping density surfaces are added together to obtain a big mountain or kernel density estimate. In Figure 6, q represents individual points (crashes), $K_q(p)$ is a network kernel density function at q, and $K(p)$ is a network kernel density estimator.



(a) (b)

Figure 6. Kernel density estimation procedure on a line segment: (a) molehills or kernels (b) a big mountain or kernel density estimate (source: (Okabe and Sugihara 2012)

Conceptually, a road crash does not occur at an exact point location but occupies and affects some length along the roadway. The use of points ignores the impact on the nearby roadway by simplifying the dimension. The uncertainty created by the crash logging process is not considered. On the contrary, characterizing a crash as a spatial event that possesses a spread of risk is more appropriate. The spread of risk is defined as the neighboring area around a crash location where the likelihood for a crash to occur differs depending on how close areas are to the crash (Anderson 2009). Previous research shows that the surface-based density estimation may not be suitable for characterizing crash locations occurring on the roadway network (Okabe and Sugihara 2012; Steenberghen, Aerts, and Thomas 2010; Xie and Yan 2013). Network kernel density estimation (NetKDE) is an improved approach to avoid under- or over-estimation issues of the planar KDE technique for the road crash data. The NetKDE can be expressed as follows:

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{r} k\left(\frac{d_{is}}{r}\right) \qquad \text{(Equation 7)}$$

Where $\lambda(s)$ is the NetKDE at location s, r is the search radius of KDE, and $k\left(\frac{d_{is}}{r}\right)$ is the weight of point i at location d_{is} to location s. The kernel function, k provides information on the distance decay effect between two points as they get farther from each other. As the distance of a point increases from location s, the effect of the point will also decrease for overall density estimation (Xie and Yan 2008). There are various types of kernel functions but their selection has little effect on both local and global density formations (Xie and Yan 2008; O'Sullivan and Unwin 2010; Silverman 1998). However, the search radius r has a major impact on $\lambda(s)$ estimation because the points outside the search radius on the given road network are not considered in the estimation. Although the process of deciding the bandwidth is somewhat subjective, past studies suggest that 50 to 300 m and 1000 m bandwidth should be used for urban and rural areas, respectively (Steenberghen, Aerts, and Thomas 2010; Xie and Yan 2013). This study used a bandwidth of 300m for density estimation. We used a software called SANET Standalone (Ver. 4.1) for analyzing crashes that occur on the road network.

3.4.4. Space-Time Cube Analysis

Hotspot analysis and kernel density analysis show patterns of lane departure crash density but lack the temporal characteristics of these crashes. Hence, space-time cube (STC) analysis was used for spatiotemporal analysis of lane departure crashes. STC analysis refers to a 3D geovisualization approach that maps spatiotemporal data in the form of a cube. The x and y dimensions of a 3D space-time cube represents space while t dimension represents time (Figure 7). In this study, the value of grid cells was set at 8000m × 8000m with a time duration of 6 months. This generated a total of 31,680 grid cells for the study area.

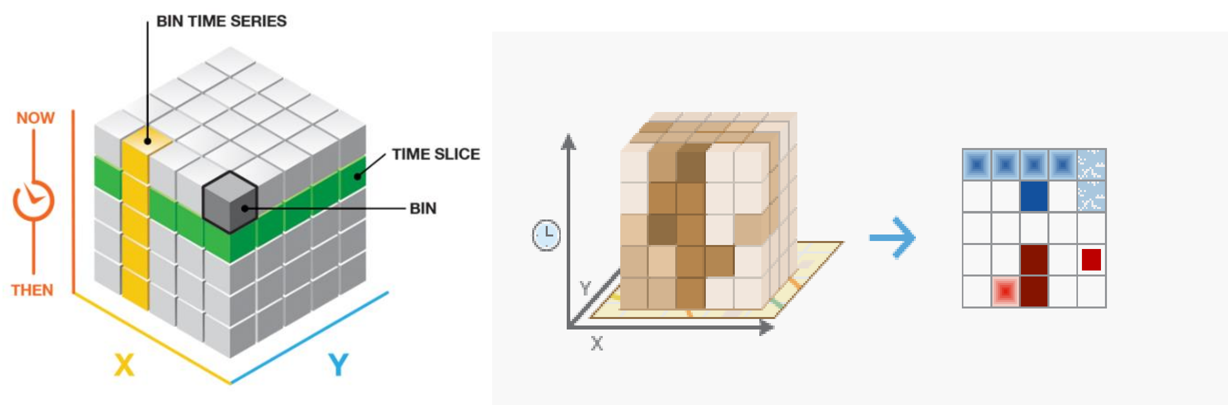


Figure 7. (Top) Structure of STC in 3D (Bottom) generated bins in 2D by running emerging hotspot analysis (Source: (Emerging Hot Spot Analysis (Space Time Pattern Mining)—ArcGIS Pro | Documentation))

This study used a time approach to analyze the trend of lane departure crashes. The input feature points (lane departure crashes, in our case) were aggregated into NetCDF (Network Common Data Form) data structures in the form of space-time bins. The points were counted within each bin and specific attributes were aggregated. The calculated value of each bin represents the crash frequency at a specific location during a given time interval. The STCs were then analyzed using emerging hotspot analysis (EHSA). The goal of EHSA is to evaluate changes in hot and cold spots with respect to time. The NetCDF data structure was analyzed using a tool known as Create Space Time Cube by Aggregating Points. It then calculates the Getis-Ord Gi* statistic for each bin and time period. After the space-time hotspot analysis completion, each bin is provided with a specific z-score, p-value, and hotspot bin classification. In the next step, the Mann-Kendall trend test (Henry B. Mann 1945; Maurice G. Kendall 1975) is used to evaluate all the hot and cold spot trends. To help better understand the change in locations over time, 17 unique categories (Table 5) are created comparing the Getis-Ord Gi* and Mann-Kendall trend test.

Table 5. Names and definitions of space-time cube patterns

Patterns	Pattern Name	Definition
	New Hot/Cold Spot	A location that is a statistically significant hot/cold spot for the final time step and has never been a previously statistically significant hot/cold spot.
	Consecutive Hot/Cold Spot	A location with a single uninterrupted run of statistically significant hot/cold spot bins in the final time-step intervals. The location has never been a statistically significant hot/cold spot before the final hot/cold spot run and fewer than 90% of all bins are statistically significant hot/cold spots.
	Intensifying Hot/Cold Spot	A location in which the intensity of clustering of high counts in each time step is increasing overall and for which the increase is statistically significant (90%).
	Persistent Hot/Cold Spot	A location that has been a statistically significant hot/cold spot for 90% of the time-step intervals with no discernible trend indicating an increase or decrease in the intensity of clustering over time.
	Diminishing Hot/Cold Spot	A location that has been a statistically significant hot/cold spot for 90% of the time-step intervals, including the final time step. Additionally, the intensity of clustering in each time step is decreasing overall and that decrease is statistically significant.
	Sporadic Hot/Cold Spot	A location that is an on-again off-again hot/cold spot. Fewer than 90% of the time-step intervals have been statistically significant hot/cold spots and none of the time-step intervals have been statistically significant cold/hot spots.
	Oscillating Hot/Cold Spot	A statistically significant hot/cold spot for the final time-step interval that has a history of being a statistically significant cold/hot spot during a prior time step. Fewer than 90% percent of the time-step intervals have been statistically significant hot/cold spots.
	Historical Hot/Cold Spot	The most recent time period is not hot/cold, but at least 90% of the time-step intervals have been statistically significant hot/cold spots.
	No Pattern Detected	Does not fall into any of the defined hot or cold spot patterns.

Source: (ESRI. How Emerging Hot Spot Analysis Works—ArcGIS Pro | Documentation)

3.5. Results and Discussion

This section provides findings for the various hotspot techniques used in analyzing single-vehicle lane departure crashes. To measure the tendency of crash events to cluster and estimate the overall degree of spatial autocorrelation, we used the Global Moran’s I. As shown in

Figure 8, the value of the Moran's I index was positive, which shows significant clustering of crashes in the study area. The z-score for crashes was 6.96 and the p-value was 0.00. The assumption of random distribution for crashes was rejected because the z-score of 6.96 was significantly greater than the threshold for rejection. It means there is less than 1% likelihood that this clustered pattern could be the result of random chance. The high value of absolute z-score indicates significant spatial autocorrelation of lane departure crashes in the study area.

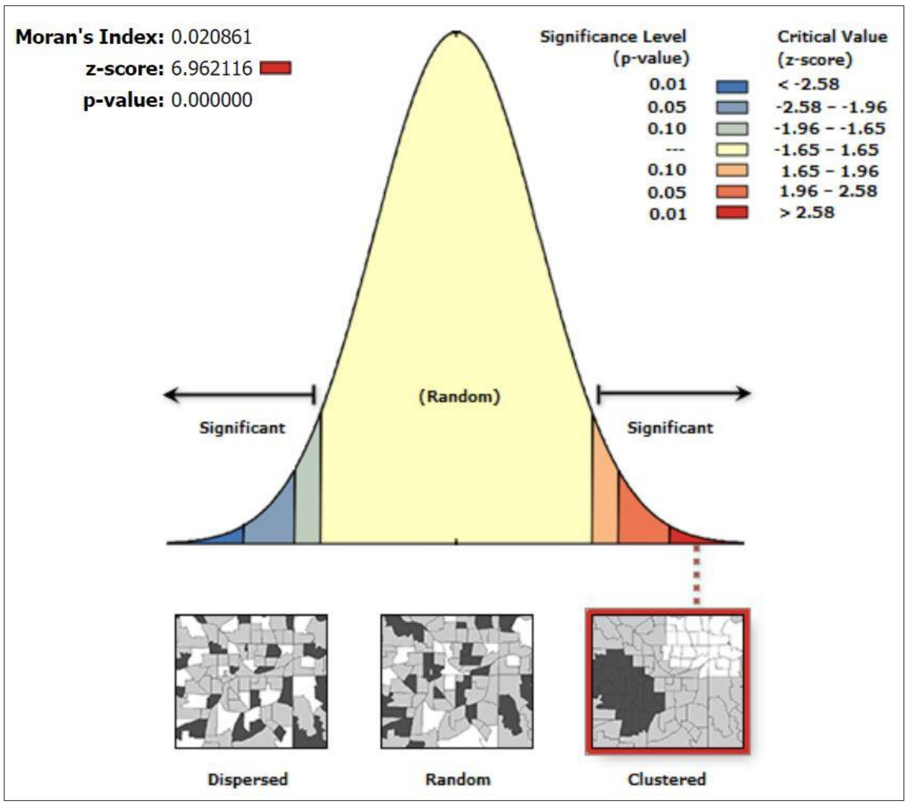


Figure 8. Global Moran's I value for single-vehicle lane departure crashes

Local spatial autocorrelation techniques were applied to assess clustering at the local level. First of all, the average distance at which any given point that has at least one neighbor was calculated to be 2,180 meters (=2.18 kilometers or 1.35 miles). This distance was set as a starting distance in the incremental spatial autocorrelation technique. Several trials to minimize the distance range of peaked z-scores were conducted. We extended the starting distance by increments of 500 to 2000 meters and observed a decreasing trend of the z-score. The concentration of spatial clustering is reflected by z-scores. The intensity of spatial clustering of crashes for the z-scores at each distance for a set number of segments is shown in Table 6. The first peak was at 3,500 m with a z-score of 4.27. The table also shows that, after reaching the peak, the intensity of spatial clustering decreases with the increase in distance. Where multiple peaks are present, the use of the first peak is recommended because it best describes the spatial variation of the analysis (Mitchell 2021). In our case, the first and maximum peak was at 3,500 m which would reflect more variations at the local level for crash hotspots.

Table 6. The z-scores and p-values for a series of distance

Distance (m)	**Moran's Index**	**Expected Index**	**Variance**	**z-score**	**p-value**
2000	0.05	-0.0004	0.0002	3.20	0.001
2500	0.05	-0.0004	0.0002	3.65	0.000
3000	0.05	-0.0003	0.0001	3.99	0.000
3500	0.04	-0.0003	0.0001	4.27	0.000
4000	0.04	-0.0003	0.0001	4.09	0.000
4500	0.03	-0.0003	0.0001	3.43	0.000
5000	0.03	-0.0003	0.0001	3.39	0.000
5500	0.02	-0.0003	0.0001	3.03	0.002
6000	0.02	-0.0003	0.0000	3.32	0.000
6500	0.02	-0.0003	0.0000	3.31	0.000

The local Moran’s I approach was applied to reveal the spatial pattern of local differences. This technique evaluated the significance of the spatial difference in crashes by comparing each cell to its surrounding cells. Figure 9 presents the results obtained for local spatial clustering through the local Moran’s I index. The figure shows high and low cluster areas together with high and low spatial outliers at the 95 percent confidence level. The red dots represent locations with high clusters, which means that serious injury crashes were surrounded by crashes with serious injuries. Similarly, locations with low clusters show the areas where

PDO crashes were mostly surrounded by similar crashes i.e. PDO crashes. The high spatial outliers show locations where the serious injury crashes were surrounded by crashes with dissimilar severity, i.e. PDO and injury crashes, and vice versa for low spatial outliers. The majority of the crash clusters with higher values, i.e. serious injury crashes, were located in Barnes, McKenzie, and Williams counties.

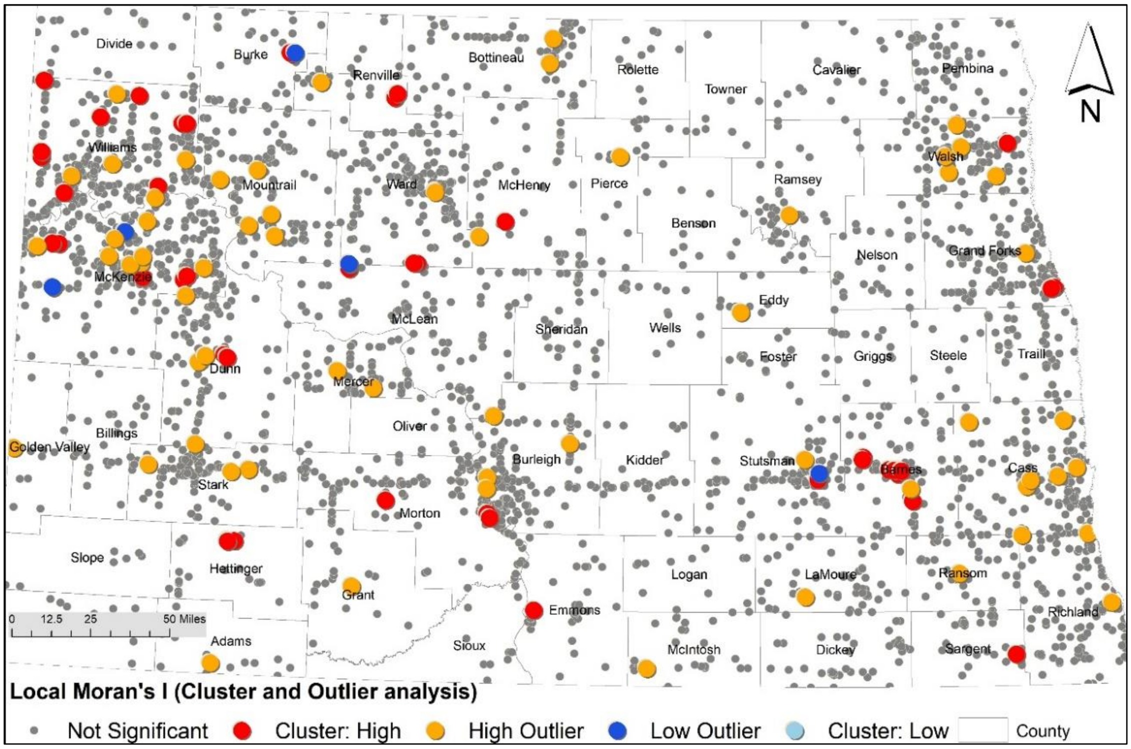


Figure 9. Hotspot classification using local Moran's I method

The local spatial autocorrelation analysis based on Getis-Ord Gi* statistics was also conducted to reveal crash severity patterns. The z-score and p-value for each crash point determine statistically significant red, blue, or gray spots while accounting for severity values of neighboring crashes. Red spots indicate statistically significant serious injury crashes surrounded by similar values, i.e. serious crashes. Blue spots indicate statistically significant minor injury crashes surrounded by minor injury crashes, and gray spots indicate points that are not statistically significant and surrounded by randomly distributed severity values. Three different shade for red and blue hotspots indicate statistical confidence levels of 90%, 95%, and 99%. The darkness and intensity of hotspots color represent the statistical significance level. Figure 10 shows serious and minor injury crash clusters in the study area.

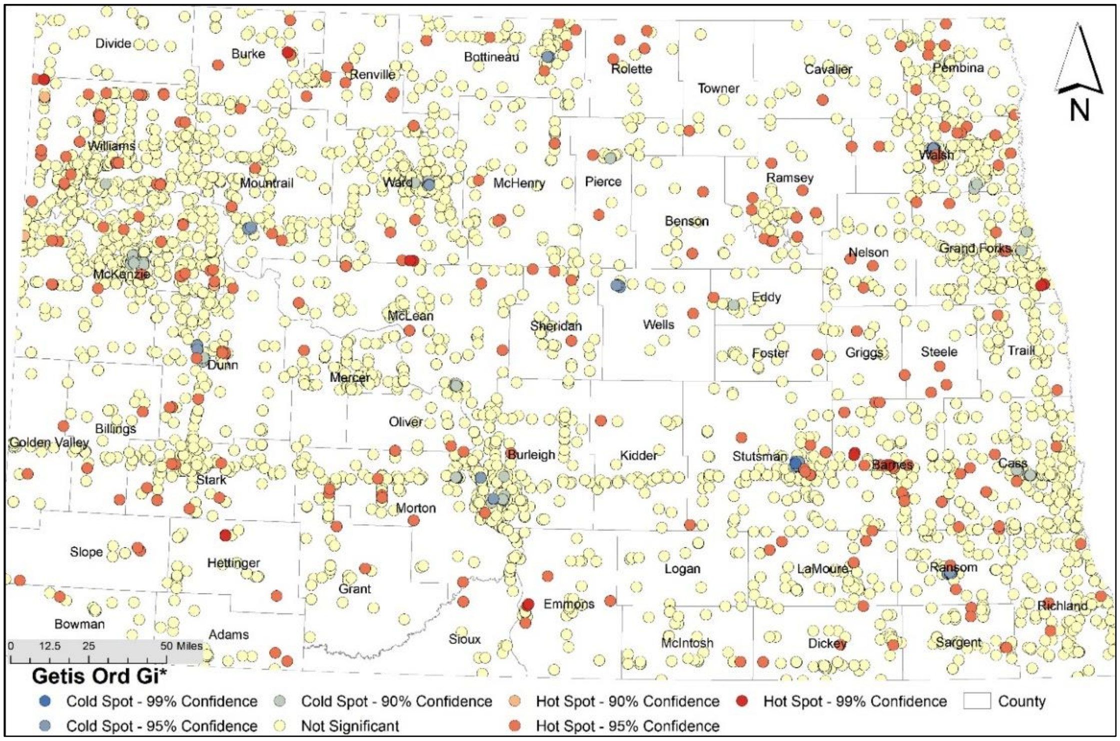


Figure 10. Hotspot analysis using local spatial autocorrelation (Getis-Ord Gi*) approach

The clusters in Figure 10 were distinguished from general crash-prone locations because their composition was based on points with similar crash severity values. Most serious injury crash clusters can be seen in Barnes, Emmons, Grand Forks, and Walsh counties. The south boundary of Ward County also showed some clustering of serious injury crashes. Similarly, the clustering of crashes with low severity, i.e. minor injury, was present throughout the study area. However, most of the minor injury crash clusters were present in Barnes, Cass, McKenzie, Stark, Walsh, and Williams counties. The frequency of minor injury crash clusters in Williams, McKenzie, and Barnes counties was 32, 27, and 19 respectively.

It was observed that blue spots (minor injury crash clusters) tend to be less spread out in comparison to red spots (serious injury crash clusters). The randomly distributed crash spots (gray-colored points) made the two different cluster types (i.e. red and blue) distinguishable. Moreover, the two different cluster types were not adjacent to each other. This shows the presence of an obvious spatial relationship between geographic locations and the occurrence of crash severity clusters. One limitation of this approach is that it is difficult to quantify clusters along a road segment and distinguish its boundaries because they are created by points. One way to overcome this issue is to use density estimation which shows the concentration of points along the road network.

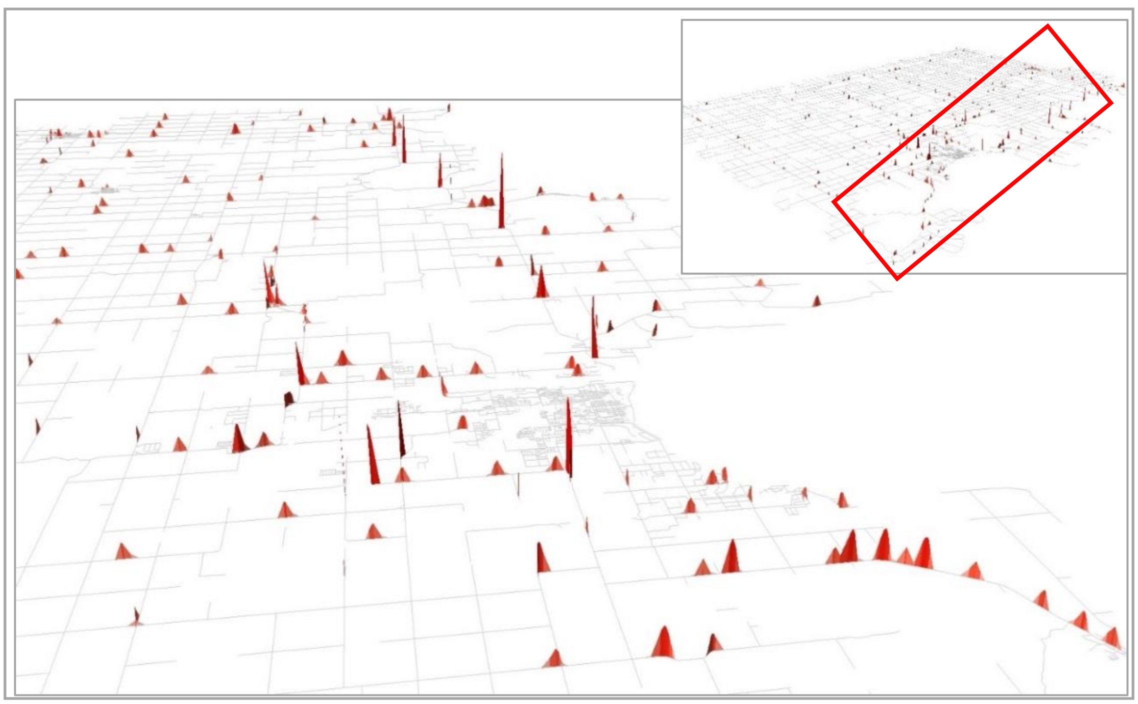


Figure 11. Network Kernel density estimation for crashes in Williams County

Hence, the third technique was applied. NetKDE is a successful spatial clustering method for investigating crashes on roadways. This method has a higher processing demand, so a county-level approach was needed. Williams was selected as the test county because it has the highest number of crash clusters. Figure 11 shows a 3-D density map of crash density patterns in Williams County with a subsection magnified for better visualization. The NetKDE results delivered comprehensive clustering patterns with clear boundaries. It can be observed that the majority of the crash clusters are located on curves, junctions, and intersections. Persistent clusters on horizontal curves can be seen throughout the county. The clusters emerging along the straight road segments might be related to speed. Investigation beyond this study's scope is needed to discern value with a granular understanding of this resource-intensive approach.

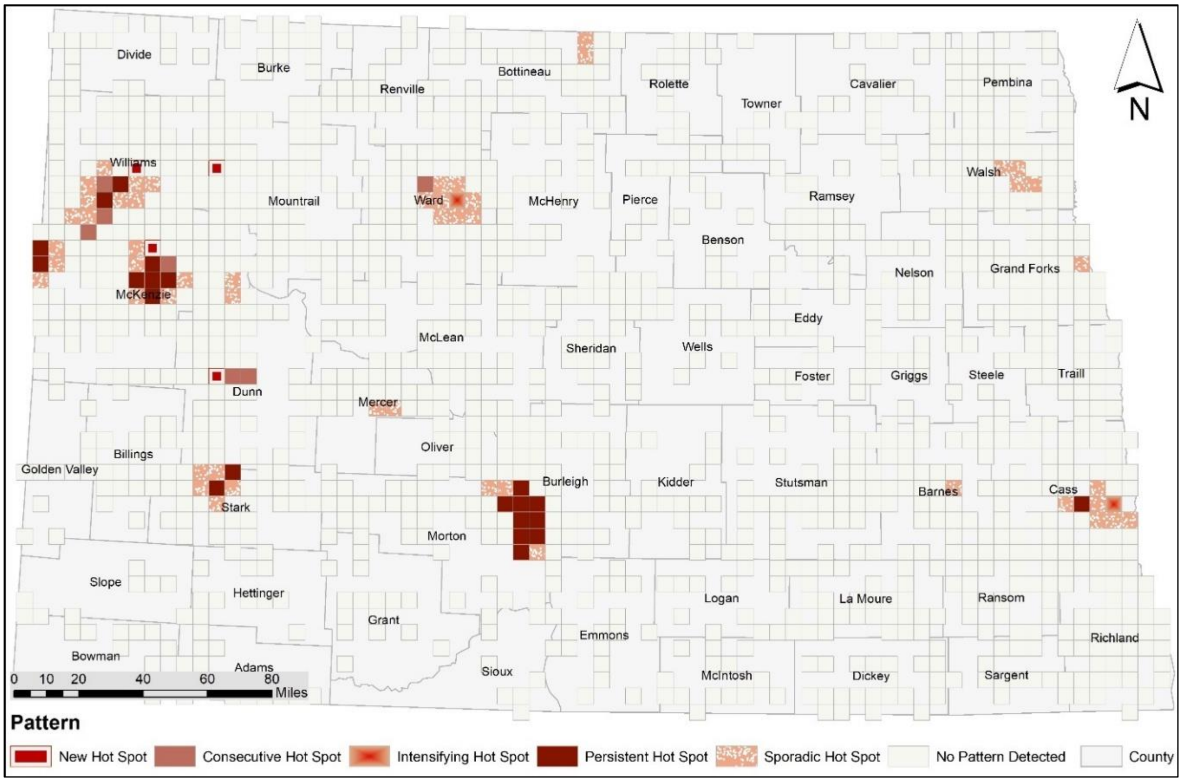


Figure 12. Emerging hotspot analysis of single-vehicle lane departure crashes in ND

This study conducted a space-time cube analysis of 3878 single-vehicle lane departure crashes. The results of this analysis included six types of patterns as can be seen in Figure 12. Persistent hotspots were distributed as spatial clusters in Burleigh, McKenzie, and Williams counties. This means that these locations have been statistically significant hotspots for 90% of the time-step intervals. There has been no visible trend with regard to a decrease or increase in the intensity of clusters over the time duration of the study. Two intensifying hotspots can also be seen in Cass and Ward counties. This refers to locations with statistically significant hotspots for 90% of the time-step intervals. Also, there has been a statistically significant increase in the intensity of clustering of high counts. Moreover, widely distributed new hotspots were generated in the western part of the state. These hotspots appeared with the passage of time. Finally, sporadic hotspots that repeatedly appeared and disappeared can be seen in both the East and West part of the state.

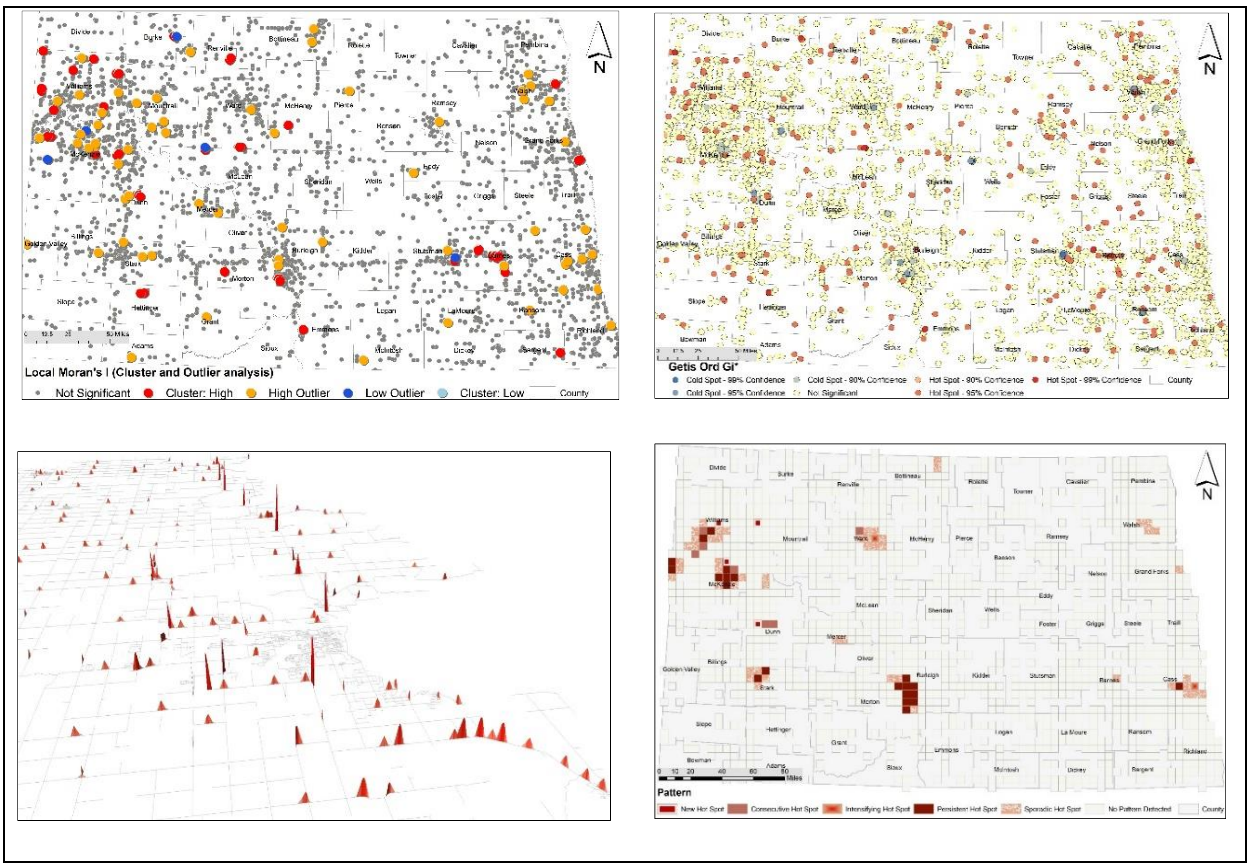


Figure 13. Comparison of results from different hotspot methods – (top left) local Moran's I (top right) Getis Ord Gi* (bottom left) network kernel density estimation (bottom right) emerging hotspot analysis

The results of different hotspot analysis methods are summarized in Figure 13. The high and low crash incidence locations are represented by hotspots and cold spots respectively. Using all the methods together to identify and locate crash hotspots is a promising approach. This study is focused more on the spatial analysis whereby we are trying to locate the spatial clusters for the crash severity. Relevant countermeasures can be implemented according to the specific spatial and temporal trends of these hotspots. The findings of the study indicate that various phenomena in certain areas of Burleigh, McKenzie, and Williams counties warrant the attention of traffic safety agencies. The findings of the study can be used to direct the limited resources and to propose preventive strategies targeting crash hotspots.

3.6. Conclusions

The benefit of classifying crash hotspots cannot be over-emphasized as an important method of discerning geospatial themes in road safety. Hence, this study employed hotspot analysis techniques to reveal spatial patterns of single-vehicle lane departure crashes. While the Global Moran's I index indicated the existence of crash clustering, the use of the local Moran's I statistic enabled the identification of hot and cold spots within the study area. Furthermore, data-driven enhancements were placed side by side with locations in terms of statistically significant high and low crash concentrations. A more granular crash cluster quantification, using the NetKDE technique, was presented for Williams County. A 3D density map created in ArcScene also defined boundaries for each cluster in terms of density values embedded in the roadway. Finally, the space-time cube analysis enabled us to identify different patterns in the spatial data with respect to time.

As effective as this method was, it does have shortcomings. Particularly, it fails to give an accurate statistical significance of the resulting crash clusters, which therefore suggests the need for further investigation. This article focused on lane departure crashes that involved only single vehicles. Another limitation of the study was limiting events to a specific road type in the state network. Future studies can overcome these limitations by broadening the crash type and locations. Results are valuable as a sample for road safety stakeholders to consider for future endeavors aimed at data-driven tools to reveal crash hotspots and to distinguish patterns across the state. Scrutiny of characteristics for crashes in the areas with a high frequency of clusters will further enhance the understanding of associated risk factors and appropriate prevention strategies such as heightened enforcement, public education, and infrastructure enhancements including rumble strip installations, safety edges, post-mounted delineators, oversized chevrons, and advisory speed marking in lanes on curves. The tools approach used in this study can be adopted by other jurisdictions seeking to empirically visualize hotspots and more effectively deploy traffic safety strategies.

3.7. References

Anderson, Tessa K. 2009. "Kernel Density Estimation and K-Means Clustering to Profile Road Accident Hotspots." *Accident Analysis & Prevention* 41 (3): 359–64. https://doi.org/10.1016/J.AAP.2008.12.014.

Anselin, Luc. 1995. "Local Indicators of Spatial Association—LISA." *Geographical Analysis* 27 (2): 93–115. https://doi.org/10.1111/J.1538-4632.1995.TB00338.X.

Chance Scott, Mackenzie, Shouraseni Sen Roy, and Shivangi Prasad. 2016. "Spatial Patterns of Off-the-System Traffic Crashes in Miami-Dade County, Florida, during 2005-2010." *Traffic Injury Prevention* 17 (7): 729–35. https://doi.org/10.1080/15389588.2016.1144878.

Cheng, Wen, and Simon P. Washington. 2005. "Experimental Evaluation of Hotspot Identification Methods." *Accident Analysis & Prevention* 37 (5): 870–81. https://doi.org/10.1016/J.AAP.2005.04.015.

Cheng, Zeyang, Zhenshan Zu, and Jian Lu. 2018. "Traffic Crash Evolution Characteristic Analysis and Spatiotemporal Hotspot Identification of Urban Road Intersections." *Sustainability 2019, Vol. 11, Page 160* 11 (1): 160. https://doi.org/10.3390/SU11010160.

Demšar, Urška, and Kirsi Virrantaus. 2010. "Space–Time Density of Trajectories: Exploring Spatio-Temporal Patterns in Movement Data." *Https://Doi.Org/10.1080/13658816.2010.511223* 24 (10): 1527–42.

https://doi.org/10.1080/13658816.2010.511223.

Dereli, Mehmet Ali, and Saffet Erdogan. 2017. “A New Model for Determining the Traffic Accident Black Spots Using GIS-Aided Spatial Statistical Methods.” *Transportation Research Part A: Policy and Practice* 103 (September): 106–17. https://doi.org/10.1016/J.TRA.2017.05.031.

Dezman, Zachary, Luciano de Andrade, Joao Ricardo Vissoci, Deena El-Gabri, Abree Johnson, Jon Mark Hirshon, and Catherine A. Staton. 2016. “Hotspots and Causes of Motor Vehicle Crashes in Baltimore, Maryland: A Geospatial Analysis of Five Years of Police Crash and Census Data.” *Injury* 47 (11): 2450–58. https://doi.org/10.1016/J.INJURY.2016.09.002.

Elvik, Rune. 2008. “Comparative Analysis of Techniques for Identifying Locations of Hazardous Roads:” *Https://Doi.Org/10.3141/2083-08*, no. 2083 (January): 72–75. https://doi.org/10.3141/2083-08.

Emerging Hot Spot Analysis (Space Time Pattern Mining)—ArcGIS Pro | Documentation. Accessed September 10, 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/space-time-pattern-mining/emerginghotspots.htm.

Erdogan, Saffet, Ibrahim Yilmaz, Tamer Baybura, and Mevlut Gullu. 2008. “Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar.” *Accident Analysis and Prevention* 40 (1): 174–81. https://doi.org/10.1016/j.aap.2007.05.004.

ESRI. How Emerging Hot Spot Analysis Works—ArcGIS Pro | Documentation. Accessed September 10, 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/space-time-pattern-mining/learnmoreemerging.htm.

Fang, Tianfang B., and Yongmei Lu. 2011. “Constructing a Near Real-Time Space-Time Cube to Depict Urban Ambient Air Pollution Scenario.” *Transactions in GIS* 15 (5): 635–49. https://doi.org/10.1111/J.1467-9671.2011.01283.X.

Roadway Departure Safety, Federal Highway Administration. https://safety.fhwa.dot.gov/roadway_dept/. Accessed September 27, 2021.

Getis, Arthur, and J. K. Ord. 1992. “The Analysis of Spatial Association by Use of Distance Statistics.” *Geographical Analysis* 24 (3): 189–206. https://doi.org/10.1111/J.1538-4632.1992.TB00261.X.

Hazaymeh, K ;, A ; Almagbile, A H Alomari, Khaled Hazaymeh, Ali Almagbile, and Ahmad H Alomari. 2022. “Spatiotemporal Analysis of Traffic Accidents Hotspots Based on Geospatial Techniques.” *ISPRS International Journal of Geo-Information 2022, Vol. 11, Page 260* 11 (4): 260. https://doi.org/10.3390/IJGI11040260.

Henry B. Mann. 1945. “Nonparametric Tests Against Trend.” *Econometrica* 13 (3): 245–59. https://www.jstor.org/stable/1907187.

Huang, H.L., H.C. Chin, and M.M. Haque. 2009. “Hotspot Identification: A Full Bayesian Hierarchical Modeling Approach.” *Transportation and Traffic Theory 2009: Golden Jubilee*, 441–62. https://doi.org/10.1007/978-1-4419-0820-9_22.

Kang, Youngok, Nahye Cho, and Serin Son. 2018. "Spatiotemporal Characteristics of Elderly Population's Traffic Accidents in Seoul Using Space-Time Cube and Space-Time Kernel Density Estimation." *PLOS ONE* 13 (5): e0196845. https://doi.org/10.1371/JOURNAL.PONE.0196845.

Kaygisiz, Ömür, Şebnem Düzgün, Ahmet Yildiz, and Metin Senbil. 2015. "Spatio-Temporal Accident Analysis for Accident Prevention in Relation to Behavioral Factors in Driving: The Case of South Anatolian Motorway." *Transportation Research Part F: Traffic Psychology and Behaviour* 33 (August): 128–40. https://doi.org/10.1016/J.TRF.2015.07.002.

Kuo, Pei-Fen, Xiaosi Zeng, and Dominique Lord. 2012. "Guidelines For Choosing Hot-Spot Analysis Tools Based on Data Characteristics, Network Restrictions, and Time Distributions." In *91st Annual Meeting of the Transportation Research Board, January 22-26,*. Washington, DC. https://trid.trb.org/view/1130485

Kveladze, Irma, Menno-Jan Kraak, & Corné, and P J M Van Elzakker. 2018. "Cartographic Design and the Space-Time Cube." *The Cartographic Journal* 56 (1): 73–90. https://doi.org/10.1080/00087041.2018.1495898.

Kveladze, Irma, Menno Jan Kraak, and Corné P.J.M. van Elzakker. 2013. "A Methodological Framework for Researching the Usability of the Space-Time Cube." *Http://Dx.Doi.Org/10.1179/1743277413Y.0000000061* 50 (3): 201–10. https://doi.org/10.1179/1743277413Y.0000000061.

Li, Linhua, Li Zhu, and Daniel Z. Sui. 2007. "A GIS-Based Bayesian Approach for Analyzing Spatial–Temporal Patterns of Intra-City Motor Vehicle Crashes." *Journal of Transport Geography* 15 (4): 274–85. https://doi.org/10.1016/J.JTRANGEO.2006.08.005.

Maingi, John K, Joseph M Mukeka, Daniel M Kyale, and Robert M Muasya. 2012. "Spatiotemporal Patterns of Elephant Poaching in South-Eastern Kenya." *Wildlife Research* 39 (3): 234–49. https://doi.org/10.1071/WR11017.

Maurice G. Kendall. 1975. *Rank Correlation Methods*. 4th ed. Griffin, London. https://www.worldcat.org/title/rank-correlation-methods/oclc/3827024.

Mitchell, Andy. 2021. *The Esri Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics, Second Edition*. *Esri Press*. Second. Esri press. https://www.esri.com/en-us/esri-press/browse/the-esri-guide-to-gis-analysis-volume-2-spatial-measurements-and-statistics-second-edition.

Mitra, Sudeshna. 2008. "Enhancing Road Traffic Safety: A GIS Based Methodology to Identify Potential Areas of Improvement." *Project Report Submitted to William and Barbara Leonard University Transportation Center*, 1–23. https://rosap.ntl.bts.gov/view/dot/18096

Moons, Elke, Tom Brijs, and Geert Wets. 2009. "Improving Moran's Index to Identify Hot Spots in Traffic Safety." *Studies in Computational Intelligence* 176: 117–32. https://doi.org/10.1007/978-3-540-89930-3_7.

Moran, P. A. P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37 (1/2): 17–23. https://doi.org/10.2307/2332142.

Nakaya, Tomoki, and Keiji Yano. 2010. “Visualising Crime Clusters in a Space-Time Cube: An Exploratory Data-Analysis Approach Using Space-Time Kernel Density Estimation and Scan Statistics.” *Transactions in GIS* 14 (3): 223–39. https://doi.org/10.1111/J.1467-9671.2010.01194.X.

O’Sullivan, David, and David J. Unwin. 2010. “Geographic Information Analysis: Second Edition.” *Geographic Information Analysis: Second Edition*, March. https://doi.org/10.1002/9780470549094.

Okabe, Atsuyuki, and Kōkichi Sugihara. 2012. “Spatial Analysis along Networks : Statistical and Computational Methods,” 288. https://www.wiley.com/en-us/Spatial+Analysis+Along+Networks%3A+Statistical+and+Computational+Methods-p-9780470770818.

Ord, J. K., and Arthur Getis. 1995. “Local Spatial Autocorrelation Statistics: Distributional Issues and an Application.” *Geographical Analysis* 27 (4): 286–306. https://doi.org/10.1111/J.1538-4632.1995.TB00912.X.

Purwanto, Purwanto, Sugeng Utaya, Budi Handoyo, Syamsul Bachri, Ike Sari Astuti, Kresno Sastro, Bangun Utomo, and Yulius Eka Aldianto. 2021. “Spatiotemporal Analysis of COVID-19 Spread with Emerging Hotspot Analysis and Space–Time Cube Models in East Java, Indonesia.” *ISPRS International Journal of Geo-Information 2021, Vol. 10, Page 133* 10 (3): 133. https://doi.org/10.3390/IJGI10030133.

Sabel, C., S. Kingham, A. Nicholson, and P. Bartie. 2005. “Road Traffic Accident Simulation Modelling - A Kernel Estimation Approach.” *Proc., 17th Annual Colloquium of the Spatial Information Research Centre University of Otago, Dunedin, New Zealand*, 67–75. https://www.researchgate.net/publication/252156589_Road_Traffic_Accident_Simulation_Modelling_-_A_Kernel_Estimation_Approach

Sacchi, Emanuele, Tarek Sayed, and Karim El-Basyouny. 2015. “Multivariate Full Bayesian Hot Spot Identification and Ranking: New Technique.” *Https://Doi.Org/10.3141/2515-01* 2515 (January): 1–9. https://doi.org/10.3141/2515-01.

Sadler, Richard Casey, Jesenia Pizarro, Brandon Turchan, Stephen P. Gasteyer, and Edmund F. McGarrell. 2017. “Exploring the Spatial-Temporal Relationships between a Community Greening Program and Neighborhood Rates of Crime.” *Applied Geography* 83 (June): 13–26. https://doi.org/10.1016/J.APGEOG.2017.03.017.

Silverman, B. W. 1998. “Density Estimation for Statistics and Data Analysis” 175. https://ned.ipac.caltech.edu/level5/March02/Silverman/paper.pdf

Songchitruksa, Praprut, and Xiaosi Zeng. 2010. “Getis–Ord Spatial Statistics to Identify Hot Spots by Using Incident Management Data:” *Https://Doi.Org/10.3141/2165-05*, no. 2165 (January): 42–51. https://doi.org/10.3141/2165-05.

Steenberghen, Thérèse, Koen Aerts, and Isabelle Thomas. 2010. “Spatial Clustering of Events on a Network.” *Journal of Transport Geography* 18 (3): 411–18. https://doi.org/10.1016/J.JTRANGEO.2009.08.005.

Thakali, Lalita, Tae J. Kwon, and Liping Fu. 2015. "Identification of Crash Hotspots Using Kernel Density Estimation and Kriging Methods: A Comparison." *Journal of Modern Transportation* 23 (2): 93–106. https://doi.org/10.1007/S40534-015-0068-0/FIGURES/10.

Vemulapalli, Sai Saylesh, Mehmet Baran Ulak, Eren Erman Ozguven, Thobias Sando, Mark W. Horner, Yassir Abdelrazig, and Ren Moses. 2017. "GIS-Based Spatial and Temporal Analysis of Aging-Involved Accidents: A Case Study of Three Counties in Florida." *Applied Spatial Analysis and Policy* 10 (4): 537–63. https://doi.org/10.1007/S12061-016-9192-4/FIGURES/9.

Vision Zero NDDOT. Accessed September 27, 2021. https://visionzero.nd.gov/strategies/LaneDeparture/.

Vogt, Andrew, and Joe Bared. 1998. "Accident Models for Two-Lane Rural Segments and Intersections:" *Transportation Research Record: Journal of the Transportation Research Board*, no. 1635 (January): 18–29. https://doi.org/10.3141/1635-03.

Xie, Zhixiao, and Jun Yan. 2008. "Kernel Density Estimation of Traffic Accidents in a Network Space." *Computers, Environment and Urban Systems* 32 (5): 396–406. https://doi.org/10.1016/J.COMPENVURBSYS.2008.05.001.

Xie, Zhixiao, and Jun Yan. 2013. "Detecting Traffic Accident Clusters with Network Kernel Density Estimation and Local Spatial Statistics: An Integrated Approach." *Journal of Transport Geography* 31 (July): 64–71. https://doi.org/10.1016/J.JTRANGEO.2013.05.009.

Yoon, Junho, and Sugie Lee. 2021. "Spatio-Temporal Patterns in Pedestrian Crashes and Their Determining Factors: Application of a Space-Time Cube Analysis Model." *Accident Analysis & Prevention* 161 (October): 106291. https://doi.org/10.1016/J.AAP.2021.106291.

Zhang, Chen, and John N. Ivan. 2005. "Effects of Geometric Characteristics on Head-On Crash Incidence on Two-Lane Roads in Connecticut:" *Transportation Research Record: Journal of the Transportation Research Board* 1908 (1): 159–64. https://doi.org/10.1177/0361198105190800119.

Zubaidi, H., I. Obaid, H. A. Mohammed, S. Das, and N. S.S. Al-Bdairi. 2021. "Hot Spot Analysis of the Crash Locations at the Roundabouts through the Application of GIS." *Journal of Physics: Conference Series* 1895 (1): 012032. https://doi.org/10.1088/1742-6596/1895/1/012032.

CHAPTER 4: INVESTIGATING FACTORS AFFECTING INJURY SEVERITY OF SINGLE VEHICLE RUN-OFF-ROAD CRASHES

4.1. Introduction

A run-off-road (ROR) crash, also known as a roadway departure crash, can be defined as an incident in which a motor vehicle crosses a center line or an edge line, encroaching onto the shoulder, roadside, median, or otherwise leaving the roadway (Federal Highway Administration (FHWA), 2020). ROR crashes typically involve a single vehicle and account for a significant portion of serious and fatal traffic accidents. According to the Fatality and Injury Reporting System Tool (FIRST), a query tool developed by the National Highway Traffic Safety Administration (NHTSA), a total of 67,811 individuals lost their lives in single-vehicle ROR crashes between 2017 and 2021 (FIRST). Out of these fatalities, 14,821 occurred in 2021. Figure 14 illustrates the fatalities in single and multiple-vehicle ROR crashes in the United States from 2017 to 2021. The likelihood of ROR crashes occurring in rural environments is higher than in urban areas. This may be attributed to inherent differences in driving behaviors between urban and rural settings. Factors such as higher speed limits, lower traffic volumes, and fewer roadway safety measures contribute to the increased risk of ROR crashes on rural roads. For instance, in an analysis of FARS data, Liu and Subramanian (2009) found that ROR crashes accounted for nearly 81% and 56% of all accidents on rural and urban roads, respectively.

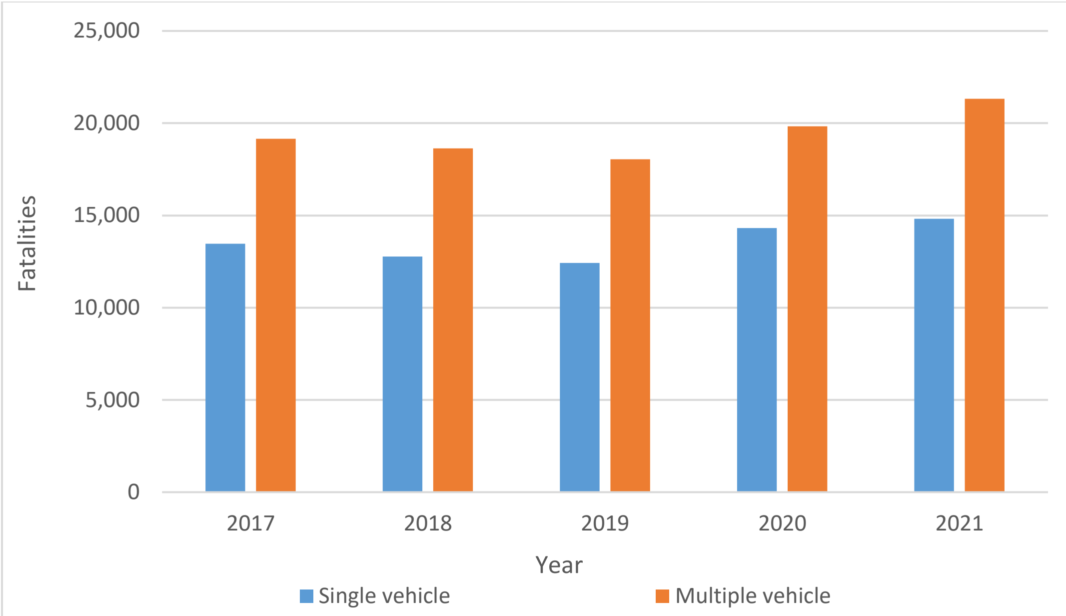


Figure 14. Fatalities in single and multiple vehicle ROR crashes in the United States (2017-2021)

Understanding the factors involved in run-off-road (ROR) crashes is crucial for enhancing ROR crash safety. Consequently, numerous researchers have investigated the impact of contributory factors on the occurrence and severity of these incidents. These factors can be broadly categorized into three distinct groups: human-related factors, environmental factors, and roadside factors. For instance, in a study conducted by Davis, Davuluri, and Pei (2006), data collected from Minnesota and Australia were analyzed using Bayesian relative risk regression. They discovered a positive relationship between higher speed limits and serious and fatal ROR crashes. In other words, an increase in speed was associated with a higher risk of serious and fatal ROR crashes. Similarly, Gong and Fan (2017) explored factors contributing to driver injury severity in ROR crashes involving single vehicles. They classified drivers into three age groups: young drivers (16-24 years), middle-aged drivers (25-65 years), and older drivers (over 65 years). Among these age groups, they observed variations in the relationship between contributing factors and injury severity in single-vehicle ROR crashes.

Given the discrete nature of crash severity data, which includes categories like fatal injury, severe injury, minor injury, and property damage only (PDO), various discrete choice models have been employed to explore the factors that influence the severity of run-off-road (ROR) crashes. Numerous studies have delved into the causal factors impacting the occurrence and injury severity outcomes of ROR crashes (Roque and Cardoso 2014; Peng and Boyle 2012; S. A. Khan, Yasmin, and Haque 2023; C. Liu and Ye 2011). This study primarily focuses on identifying the factors contributing to single-vehicle (SV) ROR crashes and their influence on injury severity. For example, Liu and Ye (2011) utilized a binary logit model to estimate the impact of driver and vehicle-related factors on injury severity in ROR crashes. Similarly, Palamara et al. (2013) employed a multinomial logit (MNL) model to identify the attributes contributing to the occurrence of ROR crashes in Western Australia. As previously mentioned, numerous factors, including roadway features, traffic-related aspects, vehicle characteristics, and environmental conditions, play a critical role in determining the outcomes of ROR crashes. Typically, data used for traditional crash severity models are derived from police reports, which include only a subset of explanatory variables.

However, there may be unobserved factors that influence the injury outcome in a crash. These unobserved factors result in heterogeneity across injury levels, and failing to account for them can lead to inaccurate parameter estimates (Washington, Karlaftis, and Mannering 2010). To address the effects of unobserved heterogeneity, random parameters models have been developed, which consider the variation of estimated parameters across observations. Nevertheless, introducing a mixing distribution makes model development more complex. The benefit of accounting for unobserved heterogeneity is that it provides valuable insights into the relationships among variables within the study dataset.

Numerous efforts have been undertaken to untangle and comprehend the complex relationship between the severity of ROR crashes and the potential contributing factors. Some studies have compared ROR crashes to non-ROR crashes to highlight differences in the underlying causal factors (Roy and Dissanayake 2011). Similarly, other efforts include studies that focus on ROR crashes involving large trucks and passenger vehicles (Peng and Boyle 2012; Roque and Cardoso 2014; Gong and Fan 2017; Al-Bdairi and Hernandez 2017; Hossain et al. 2023). These studies investigated ROR crashes primarily involving specific types or classes of vehicles. However, the physical and operational characteristics of vehicles in one class can differ significantly from those in another class. In light of these differences, the current study aims to analyze and compare the contributing factors affecting injury severity in single vehicle ROR crashes involving various vehicle classes, including passenger cars (PC), sport utility vehicles (SUVs), passenger vans/pickups, and large trucks. This goal will be achieved by developing mixed logit models separately for each vehicle class. The consistent occurrence of fatal ROR crashes in the United States underscores the ongoing need for further research. A deeper understanding of the commonalities and differences among the factors influencing injury severity in single vehicle ROR crashes across various vehicle classes is essential for the more efficient implementation of countermeasures.

4.2. Literature Review

Reducing crashes and minimizing crash severity on roads has always been of paramount importance to transportation engineers and other safety stakeholders involved in planning, designing, constructing, and maintaining highways. Ensuring a safe driving environment is not only a responsibility but also the primary focus of highway projects. Numerous efforts have been made to understand the nature and outcomes of different types of crashes. One such crash type is run-off-road (ROR) crashes, which are notorious for their severe consequences, particularly when a vehicle departs the roadway and collides with fixed objects on the roadside or when a vehicle crosses the centerline and is involved in a head-on collision. Given the limited information available from collected crash data, significant efforts are required to enhance our understanding of the factors contributing to ROR crash occurrence and severity.

The literature highlights two types of countermeasures with the potential to mitigate injury severity in ROR crashes: infrastructure-based countermeasures and vehicle-based countermeasures. Infrastructure-based countermeasures include cable barriers, guardrails, and concrete barriers, which aim to protect vehicles from roadside hazards by redirecting them back onto the roadway. On the other hand, vehicle-based countermeasures for addressing road departures include systems like lane departure warning and lane departure prevention (Riexinger, Sherony, and Gabler 2019). The distinction between these systems is that the former requires driver input for corrective maneuvers, while the latter can automatically steer the vehicle back into the lane without driver intervention. However, despite the advancements in vehicle-based safety measures, it's important to acknowledge that not all ROR crashes can be prevented. Therefore, infrastructure-based countermeasures will remain an indispensable element in ensuring roadside safety.

Turning to the methodological efforts that have been carried out to support informed decision-making in ROR crash safety. These efforts include the development of various discrete choice models to explain injury severity in ROR (Peng and Boyle 2012; Lord et al. 2011; C. Liu and Ye 2011; Palamara et al. 2013; Gong and Fan 2017; Hu and Donnell 2011; Ye and Lord 2014). A comprehensive review of discrete choice models reveals that they can be classified into two categories: ordinal or nominal (Savolainen et al. 2011).

Within the realm of ordinal models, the most common ones include ordered logit models, ordered probit models, and ordered mixed logit models. On the other hand, nominal models can be subdivided into three groups: multinomial logit models, nested logit models, and mixed logit models. Interested readers can refer to the following sources for a more in-depth review of crash injury severity models: Savolainen et al. (2011) and Mannering and Bhat (2013). Few researchers have directly compared various crash severity models, despite each model type having its own set of distinctive advantages and limitations. So far, there is no prevailing consensus on which model stands out as the best choice, as the choice of model selection is typically influenced by data availability and specific attributes of the data. Some researchers tend to prefer nominal models over ordinal models due to the restrictions that come with the influence of variables on probabilities of ordered discrete outcomes, which involve using the same coefficient for a variable across various levels of crash severity. On the other hand, some researchers favor ordinal models because of their simplicity and effective performance, particularly when working with less detailed data.

Regarding unordered discrete outcome structures, latent-class, and mixed logit models have been proposed as viable methods for modeling different levels of crash injury severity. The purpose is to address the potential limitations associated with the Independence from Irrelevant Alternatives (IIA) property found in the MNL model. The IIA property means that the probability of choosing a particular choice is independent of the presence or absence of other alternatives. The presence of unobserved effects correlation among certain injury severity levels will violate the model's IIA assumption. Previous studies have indicated the existence of shared unobserved effects at lower levels of crash severity (Hu and Donnell 2010; Lee and Mannering 2002). The mixed logit model is an appropriate substitution to overcome this limitation.

However, when estimating a mixed logit model for crash severity, a relatively larger sample size is required if more random parameters are incorporated into the model (Ye and Lord 2014).

4.3. Data and Methodology

The data used for analysis in this study was obtained from the Crash Report Sampling System (CRSS, NHTSA), which is one of NHTSA's many crash data collection programs. CRSS comprises a dataset that includes police-reported crashes involving various types of motor vehicles, cyclists, and pedestrians, ranging from minor property damage to fatal events. These crashes are collected from 60 different selected areas across the United States. For this study, five years of crash data (2016 to 2020) were obtained from CRSS. The data was acquired in separate Statistical Analysis System (SAS) files, including accident (data file containing information about crash characteristics and environmental conditions at the time of the crash), vehicle (data file containing information describing the in-transport motor vehicles and the drivers of in-transport motor vehicles who are involved in the crash), person (data file contains information describing all persons involved in the crash including motorists and non-motorists), parkwork (data file containing information about parked and working vehicle involved in crashes), pbtype (data file containing information about bicycle and pedestrian crashes), safetyeq (data file containing information about safety equipment used by people who are not occupants of motor vehicles), and cevent (data file containing information for all of the qualifying events both harmful and non-harmful). Details about the available data files in CRSS can be found in the CRSS Analytical User’s Manual (Center for Statistics and Highway Traffic Safety Administration, 2023). All the data files have a common unique case number associated with a specific crash event. The data files were merged using the case number. The ROR crash type was identified based on the first harmful event and the pre-crash circumstance configured as a right roadside departure or left roadside departure. There were a total of 13,479 single vehicle ROR crashes that occurred from 2017 to 2020. After carefully examining the dataset, variables with missing and unknown values were filtered out, resulting in 8,618 remaining observations. Initially, there were five levels of injury, including no injury (3,850), possible injury (1,490), minor injury (1,430), severe injury (1,620), and fatal injury (228). For modeling purposes, three levels of injury severity were used by merging possible injury with minor injury and severe injury with fatal injury. Table 7 shows the descriptive statistics of all the variables.

Table 7. Descriptive statistics of explanatory variables

Variable	**Categories**		**Frequency**
	Value	**Description**	
PDO	0	No injury	3850
MINJ	1	Minor injury	2920
SFINJ	2	Severe/Fatal injury	1848
Driver Information			
Age in years	35.65		8615
Gender	1	Male	5526
	0	Female	3092
Seatbelt use	1	Used	7518
	0	Not used	1100
Alcohol/Drug usage	1	Yes	1639
	0	No	6979
Contributing factor	0	Speeding	1972
	1	Careless	1109
	2	Overcorrection	585
	3	Aggressive	121
	4	Overloading	5
	5	Towing	7
Vehicle Information			
Vehicle type	0	PC	5019
	1	SUV	1667
	2	Passenger van/Pickup	1612
	3	Large truck	269

Table 7. Descriptive statistics of explanatory variables (continued)

Variable	Categories		Frequency
	Value	Description	
Vehicle model year	0	Before 2000	844
	1	Between 2000 and 2010	4047
	2	After 2010	3727
Roadway Information			
Road location	0	Rural	2891
	1	Urban	5727
Surface condition	0	Wet/Slush/Snow/Ice	2833
	1	Dry	5785
Roadway alignment	0	Straight	6176
	1	Curve	2442
Roadway grade	0	Level	6700
	1	Hilly	1918
Numbers of lane	0	One lane	3275
	1	Two or more lanes	5343
Trafficway	0	One-way	5870
	1	Two-way	2748
Speed limit	0	<= 45 mph	4613
	1	46-60 mph	2554
	2	>= 61 mph	1451
Environmental Information			
Weather	0	Rain/Hail/Snow/Fog	2132
	1	Clear	6486
Light conditions	0	Daylight	4467
	1	Dawn/Dusk	379
	2	Dark - lighted	1492
	3	Dark – not lighted	2279
Region	0	Northeast	358
	1	Midwest	1504
	2	South	6064
	3	West	692
Crash Information			
Time	0	Day	2612
	1	Evening	2753
	2	Night	3253

Table 7. Descriptive statistics of explanatory variables (continued)

Variable	Categories		Frequency
	Value	Description	
Day of week	0	Weekend	2738
	1	Weekday	5880
Harmful event	0	Tree/pole	3119
	1	Concrete barrier	665
	2	Guardrail/fence	1149
	3	Ditch	1793
	4	Other fixed object	1356
	5	Rollover	536
Rollover location	0	On roadway	82
	1	On shoulder	16
	2	On median	162
	3	On roadside	1765
Vehicle travel speed	0	<= 45 mph	2888
	1	46 – 60 mph	3217
	2	61 – 75 mph	2207
	3	>= 76 mph	306
Airbag deployed	0	No	4697
	1	Yes	3921
Ejection	0	Fully/partially ejected	8345
	1	Not ejected	273

The gaps and subpar quality of crash data can create challenges when addressing road safety issues. Considering the absence of detailed and high-quality crash data, the use of adequate methodology becomes more important to avoid difficulties in making statistical inferences. This study developed a mixed logit model (random parameters multinomial logit model) to account for unobserved heterogeneity arising due to randomness of unobserved factors. The flexibility in model definition has given this technique an important place in traffic safety research. The mixed logit model is an extension of the multinomial logit model by allowing the parameter β_i to vary across individual observations, giving the option for the constant specific to injury outcomes and each element of the parameter vector β_i to be either fixed or randomly distributed with fixed means. Within the mixed logit framework, a linear utility function is defined to determine the particular injury severity level *j* for observation *i* as follows (Washington, Karlaftis, and Mannering 2010):

$$U_{ij} = \boldsymbol{X}_{ij}\beta_j + \varepsilon_{ij} \qquad \text{(Equation 8)}$$

where U_{ij} is the specific injury severity function that determines the probability of injury severity level *j* in crash *i*, $\boldsymbol{X}_{ij}$ is a vector of explanatory variables for injury severity level *j* in crash *i*, β_j is a vector of estimable parameters for injury-severity level *j,* and accounts for the unobserved heterogeneity, and ε_{ij} is the disturbance term assumed to be independent and identically distributed. The conditional probability for alternative *j* and crash *i* can be written as:

$$P_{ij}/\beta_j = \frac{\exp(X_{ij}\beta_j)}{\sum_{j=1}^{J}\exp(X_{ij}\beta_j)} \qquad \text{(Equation 9)}$$

Similarly, the unconditional probability is expressed as:

$$P_{ij} = \int (P_{ij}/\beta_j) f(\beta/\varphi) d\beta \qquad \text{(Equation 10)}$$

where $f\,(\beta/\phi)$ describes the probability density function (PDF) of the random vector β and ϕ is the parameter (mean and variance) vector with known density function. It is worth stating that some parameters of the vector β may be randomly distributed and some may be fixed. It has been observed while estimating mixed logit models with the introduction of random parameters for certain variables, the mean effects may not significantly differ from zero but their standard deviation may. This phenomenon points out the presence of unobserved heterogeneity in the dataset across all of the observations. To address this situation, a likelihood ratio (LR) test is conducted, comparing the model with random coefficients to the model with fixed coefficients. The model with the random coefficient is utilized if the hypothesis that the coefficient is not random is rejected; if not, the coefficient is fixed.

In this study, we tried three distributional forms including normal, uniform, and lognormal for the identified random parameters. The log-normal distribution was not a good fit as it requires the estimated parameter impact to be strictly positive or negative (Train 2009). However, in crash data as is the case in this dissertation, a random effect can cause positive and negative impacts. As far as uniform and normal distribution are concerned, it is noted noted that the log-likelihood values at convergence with uniform distribution were lower than those with normal distribution. Therefore, normal distribution was used in this study as it provided the best estimation results. This is found to be consistent with past studies which conclude that the normal distribution provides the best estimation results for crash injury severity data (Moore et al. 2011; Behnood and Mannering 2016; Wu et al. 2014; Gong and Fan 2017).

Typically, in mixed logit model development, a simulation-based maximum likelihood estimation is used for coefficient estimation. There is a direct dependence of the required number of simulations on model complexity. With the increase in the number of randomized variables, the number of simulation points also increases. Past literature shows that the standard Halton draws provide more efficient and accurate parameter estimates than the random draws (Train 2009; Bhat 2003; Zeng 2016). This study used 500 Halton draws for simulation. Considering the number of Halton draws, it is sufficient to produce precise maximum likelihood estimates. Figure 15 and Figure 16 show the bivariate scatter plot of random uniform draws and Halton draws respectively.

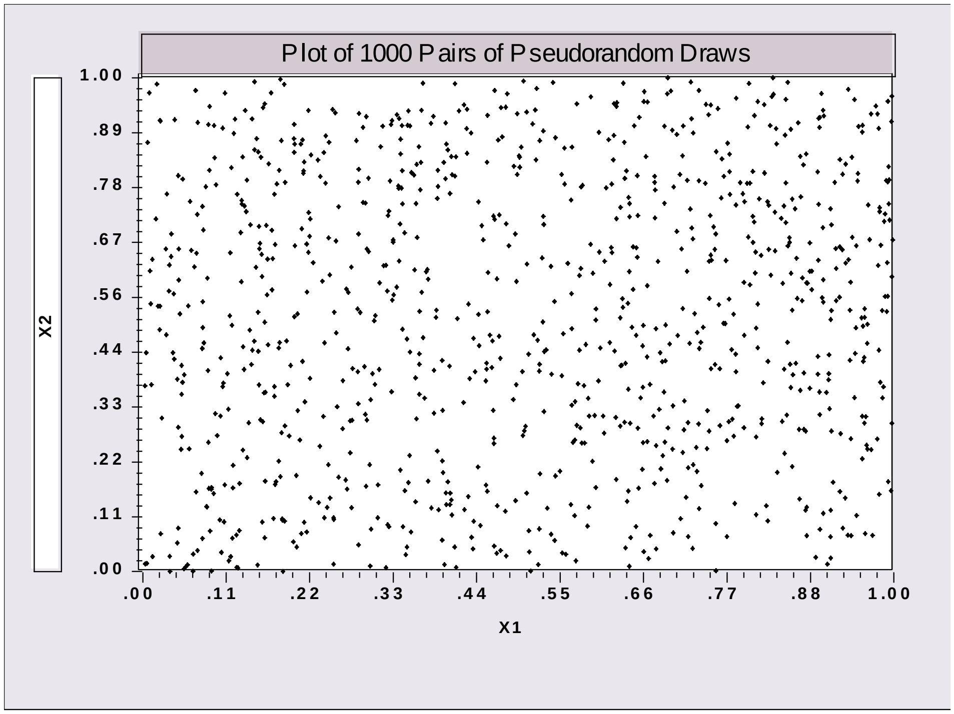


Figure 15. Bivariate scatter plot of random uniform draws

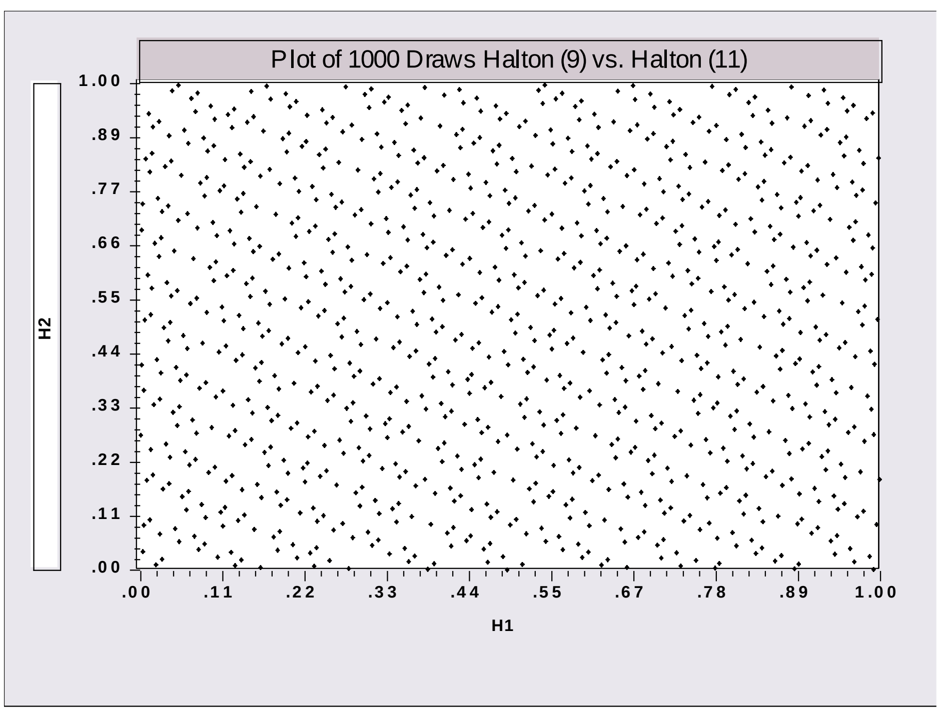


Figure 16. Bivariate scatter plot of Halton (9) and Halton (11)

4.4. Elasticity Analysis

The estimated model coefficients by the mixed logit model are insufficient to explore how changes in the explanatory factors affect the probabilities of outcomes. This limitation arises from the fact that the marginal effect of a variable is dependent on all the coefficients within the model. Therefore, determining the true net effect is not possible only from the individual value or sign of a single coefficient (Khorashadi et al. 2005). To make a comparative evaluation of each model parameter, elasticities were calculated to quantify the degree of impact that individual parameters have on the probabilities of the three levels of injury severity. Since there was a continuous variable (i.e. driver age) in the data, Equation 11 was used to compute the elasticity corresponding to it.

$$E_{X_{in}}^{P_n(i)} = [1 - \textstyle\sum_{I=I_n} P(i)]\beta_i \boldsymbol{X}_{in} \qquad \text{(Equation 11)}$$

where I_n represents the subgroup of injury severity levels that incorporates the variable $\mathbf{X}_{in}$ in the severity function, $P_n(i)$ is the probability of a person (i), and β_i denotes the estimated coefficient linked to $\mathbf{X}_{in}$. As outlined in Equation 11, a 1 percent elastic change corresponds to an approximate 1 percent change in the probability of the injury severity outcome. As far as the quantification of binary variables (those with a value of 1 or 0) is concerned, Equation 12 was used to calculate pseudo-elasticities for those variables.

$$E_{X_{in}}^{P_n(i)} = \frac{EXP[\Delta(\beta_i \boldsymbol{X}_{in})] \sum_{\forall I} EXP(\beta_i \boldsymbol{X}_{in})]}{EXP[\Delta(\beta_i \boldsymbol{X}_{in})] \sum_{I=I_n} EXP(\beta_i \boldsymbol{X}_{in}) + \sum_{I \neq I_n} EXP(\beta_i \boldsymbol{X}_{in})} - 1 \qquad \text{(Equation 12)}$$

where $E_{X_{in}}^{P_n(i)}$ is the pseudo-elasticity of the nth variable in the vector $\mathbf{X}_i$. Given that using the average value of dummy variables wouldn't be appropriate, the elasticities were computed by averaging them across the entire sample. The approximate interpretation of the elasticity value for a specific variable $\mathbf{X}_{in}$ is the percentage effect that a 1% change in $\mathbf{X}_{in}$ has on the probability of the injury severity outcome $P_n(i)$. The term "pseudo-elasticity" refers to the percentage change in the probability of a specific injury severity category when a binary variable changes from zero to one within the context of an ROR injury severity category. For example, the results of this study indicate that in the case of a ROR crash involving a passenger car, when the seatbelt is used (i.e., when seatbelt usage changes from 0 to 1), the probability of a serious or fatal injury decreases by an average of 11.54%. The literature reflects a widespread use of average pseudo-elasticity calculation due to its computational efficiency. (Moore et al. 2011; Kim et al. 2013; Wu et al. 2014; Gong and Fan 2017; Roque, Jalayer, and Hasan 2021).

4.5. Results and Discussion

As mentioned in the introduction section, the aim of this study was to analyze and compare the factors influencing injury severity in single-vehicle run-off-road (ROR) crashes involving various vehicle classes. Initially, four vehicle classes were chosen for the analysis: passenger cars (PC), SUVs, passenger vans/pickups, and large trucks. However, as indicated in Table 7, the dataset for large trucks comprised only 269 observations. This sample size is significantly smaller than the generally recommended size for estimating a mixed logit model (Ye and Lord 2014). Despite the larger sample size requirement for mixed logit modeling, we attempted to develop a model for large trucks. Unfortunately, the model results did not reveal any significant explanatory variables. Consequently, the analysis of large trucks was excluded from this study.

Table 8. Mixed logit model results for PC

Variable	**Coefficient**	**S.E.**	**p-Value**
Constant (PDO)	1.0300*	0.2356	0.0810
Age (PDO)	-0.02230***	0.00720	0.0020
Gender (PDO)	0.49348***	0.18587	0.0079
Alcohol (MINJ)	0.62559**	0.28300	0.0271
Alcohol (SFINJ)	1.14647***	0.32592	0.0007
Vehicle model after 2010 (PDO)	-0.52699**	0.22830	0.0210
Vehicle model after 2010 (MINJ)	-0.43154**	0.20485	0.0351
Northeast (PDO)	0.80843*	0.45528	0.0758
Midwest (PDO)	1.25438***	0.42131	0.0029
South (PDO)	0.58757**	0.26299	0.0255
Weather (PDO)	0.59782***	0.22693	0.0084
Harmful event-Guardrail/Fence (PDO)	0.41981*	0.23293	0.0715
Rollover on shoulder (PDO)	5.77467*	3.30541	0.0806
Travel speed above 75 mph (PDO)	-1.53736**	0.72644	0.0343
Airbag (MINJ)	-0.59771***	0.16325	0.0003
Eject (SFINJ)	4.04742***	0.99380	0.0001
Heterogeneity in means			
Seatbelt (SFINJ)	-3.78095***	1.05889	0.0006
Seatbelt (MINJ)	-1.75022***	0.57476	0.0037
Airbag (PDO)	-3.39067***	0.84516	0.0001
Travel speed above 75 mph (MINJ)	-1.88051*	1.13213	0.0967
Heterogeneity in variance			
Seatbelt (MINJ) (SFINJ)	2.51111**	0.98708	0.0178
Airbag (PDO)	2.91109***	1.01521	0.0047
Travel speed above 75 mph (MINJ)	4.96018**	2.10130	0.0358
Number of observations		5019	
Log-likelihood		-2774.27526	

Table 8. Mixed logit model results for PC (continued)

Variable	Coefficient	S.E.	p-Value
Restricted log-likelihood	-3416.68422		
Chi squared	1284.81793		
Note: ***, **, * indicates significance at 1%, 5%, 10% level, respectively. (PDO) defined for no injury; (MINJ) defined for minor injury; (SFINJ) defined for serious or fatal injury.			

A mixed logit model with heterogeneity in means and variance was created for single-vehicle ROR crashes involving PCs. Heterogeneity in means refers that the parameters (e.g., coefficients) of the model are not constant but vary across individuals or groups. Similarly, heterogeneity in variance refers to the variability around the mean (i.e., the variance) is not constant but varies across individuals or groups. We used NLOGIT (version 5) for all models' development. The analysis focused on three levels of injury severity: property damage only/no injury, minor injury, and severe or fatal injury, utilizing a total dataset of 5,019 observations. The model was constructed using a variety of crash-related attributes, including driver characteristics, roadway features, environmental conditions, temporal features, vehicle information, and crash details. The parameter estimation results for the PC model are presented in Table 8. The log-likelihood and restricted log-likelihood values for the developed model are -2774.27 and -3416.68, respectively. A significantly higher (i.e., less negative) log-likelihood value compared to the restricted log-likelihood value suggests that our model offers a better fit to the data. Some of the robustness tests for the mixed logit model, to ensure the model reliability and validity of the estimated parameters and model performance, included the multicollinearity test, random parameter distribution, and model comparison with the multinomial logit model.

Table 9. Average Pseudo elasticities of significant variables for PC

Variable	**Elasticity effect**		
	PDO	MINJ	SFINJ
Age (PDO)	-0.0798	-	-
Gender (PDO)	0.0310	-	-
Alcohol (MINJ)	-	0.0134	-
Alcohol (SFINJ)	-	-	0.0201
Seatbelt (SFINJ)	-	-	-0.1154
Seatbelt (MINJ)	-	-0.0927	-
Vehicle model after 2010 (PDO)	-0.0243	-	-
Vehicle model after 2010 (MINJ)	-	-0.0235	-
Northeast (PDO)	0.0032	-	-
Midwest (PDO)	0.0214	-	-
South (PDO)	0.0457	-	-
Weather (PDO)	0.0151	-	-
Harmful event-Guardrail/Fence (PDO)	0.0065	-	-
Rollover on shoulder (PDO)	0.0005	-	-
Travel speed above 75 mph (PDO)	-0.0035	-	-
Travel speed above 75 mph (MINJ)	-	-0.0006	-
Airbag (PDO)	-0.0858	-	-
Airbag (MINJ)	-	-0.0332	-
Eject (SFINJ)	-	-	0.0936

As previously discussed, it's important to exercise caution when interpreting parameter coefficient estimates, as a positive coefficient doesn't necessarily imply a higher probability of that specific injury severity level. To offer a more precise assessment of the estimated parameter coefficients, we calculated parameter-specific elasticities (for continuous variables) and pseudo-elasticities (for categorical variables), as presented in Table 9. This approach assists in quantifying the impact of individual parameters on the likelihood of the three injury severity levels.

4.5.1. Continuous Variables

The continuous variable "driver age" was found to be statistically significant for the severity level PDO, with an estimated coefficient of -0.022. The negative coefficient value indicates that as the driver's age increases, the probability of experiencing "no injury" decreases. The results from Table 9 offer additional insights, demonstrating that the probability of no injury decreases by 7.98% for each 1% increase in driver age.

4.5.2. Driver Characteristics

The driver gender was found to be statistically significant for the severity level of “PDO” in single vehicle ROR crashes involving PC. The model estimate for driver gender is significant with a p-value of $p < 0.0079$ for PDO. Since male and female drivers exhibit variations in skills and decision-making abilities in complex situations, gender is associated with an influence on crash severity. The pseudo-elasticity analysis in Table 9 reveals that male drivers have a lower likelihood of sustaining injuries compared to female drivers. The presence of male drivers reduced the likelihood of PDO injury by 3.10% in a single vehicle ROR crash involving a PC. This is in line with prior research, which has consistently demonstrated that male drivers exhibit superior driving performance and contribute to enhanced safety in complex scenarios compared to their female counterparts (Yasmin et al. 2014).

The presence of alcohol or drug involvement significantly elevates the risk of minor and serious or fatal injuries. More specifically, when alcohol or drug involvement is identified as a contributing factor in single vehicle ROR crashes involving a PC, the probability of minor and serious or fatal injuries increases by 1.34% and 2.01%, respectively. These findings align with prior research in the field of traffic safety (Wu et al. 2014; Gong and Fan 2017; Hasan et al. 2022; P. Liu and Fan 2020). The observed changes in fatal injury probabilities may be attributed to impaired judgment or aggressive driving behaviors associated with alcohol or drug use. Consequently, the implementation of educational and enforcement initiatives aimed at deterring drunk driving could serve to mitigate the severity of single-vehicle ROR crashes.

The mixed logit model revealed a normally distributed random parameter with heterogeneity in mean and variance for seatbelt use in relation to MINJ and SFINJ severity levels. This suggests that there are unknown uncertainties associated with seatbelt use, and these uncertainties vary when modeling injury severity for MINJ and SFINJ. Consequently, the impact of seatbelt use on the probabilities of MINJ and SFINJ severity levels is significant but may vary in magnitude. The results of this study indicate that seatbelt use was associated with a decreased likelihood of minor injury (MINJ) and serious or fatal injury (SFINJ) resulting from a single-vehicle run-off-road (ROR) crash involving a passenger car (PC). These findings align with previous research (Abdel-Aty 2003; Z. Li et al. 2019; Kim et al. 2013; Sawtelle et al. 2023) and are intuitive, as the use of seatbelts is a major contributing factor to reducing injury severity (I. U. Khan and Vachal 2020). The estimated variation in standard mean and standard deviation for this random parameter is -3.78 and 2.51, respectively, for the SFINJ severity level. As shown in Table 9, seatbelt use reduced the probability of MINJ and SFINJ by 9.27% and 11.54%, respectively, in a single-vehicle ROR crash involving a PC.

4.5.3. Vehicle Characteristics

The vehicle model year after 2010 was found to be statistically significant for PDO and MINJ severity levels. The likelihood of a PC occupant sustaining PDO or MINJ injuries in a single-vehicle ROR crash was reduced by 2.43% and 2.34%, respectively, if the vehicle's model year was after 2010. This result is consistent with the advancements in vehicle safety systems, such as lane departure warning and lane-keeping assist, that have been made over the past decade. Previous studies have demonstrated the effectiveness of these safety technologies in enhancing safety in ROR crashes (Cicchino 2018; Hickman et al. 2015; Sternlund et al. 2017).

4.5.4. Environmental Information

Two variables including region and weather were found to be significant. The likelihood of PDO increases by 0.32%, 2.14%, and 4.57% if the single vehicle ROR crash involving PC occurred in the Northeast, Midwest, and South, respectively. Similarly, the likelihood of PDO increases by 1.51% if the single vehicle ROR crash involving a PC occurred during bad weather (i.e., rain, hail, snow, or fog). It could be attributed to the fact that inclement weather serves as an alert to drivers to exercise caution and drive safely.

4.5.5. Crash Information

The initial object struck in an ROR crash plays a crucial role in determining injury severity. In the case of a single-vehicle ROR crash involving a PC and hitting a guardrail or fence, there was a positive association with PDO. This means that the probability of PDO increased by 0.56%. This finding is consistent with previous research and is intuitive, as guardrails or fences serve as safety barriers designed to protect occupants from serious or fatal injuries (Roque, Jalayer, and Hasan 2021). It prevents vehicles from colliding with more hazardous objects and acts as a shock absorber, which in turn increases the probability of PDO. Similarly, in ROR crashes involving a PC, the initiation of a rollover on the shoulder was found to be significant for PDO. Prior research has demonstrated that the initial location of a rollover has an impact on the outcome of injury severity (I. U. Khan and Vachal 2020).

The travel speed of over 75 mph was significantly associated with two injury severity outcomes, namely PDO and MINJ. However, a normally distributed random parameter with heterogeneity in mean and variance was found to be significant for speeds above 75 mph for the minor injury category. This indicates that a speed of over 75 mph has an impact of varying magnitude on minor injuries. The pseudo-elasticity analysis in Table 9 reveals a decrease in the likelihood of PDO and MINJ severity levels for speeds above 75 mph. This could be attributed to the fact that higher speeds are linked to greater injury severity, including fatal injuries, as indicated in previous studies (Malyshkina and Mannering 2008; Alnawmasi and Mannering 2022). Higher speeds may contribute to increased injury severity in single-vehicle ROR crashes.

The airbag was found to be statistically significant. The model estimates for the airbag are significant with p-values of less than 0.0003 for MINJ, and less than 0.0001 and 0.0047 for mean heterogeneity and variance heterogeneity for PDO. The pseudo-elasticity analysis reveals that the airbag decreases the likelihood of PDO and MINJ severity by 8.58% and 3.32%, respectively. It should be noted that due to the randomness of this variable, its impact varies in magnitude on PDO severity. Past studies also highlight the safety significance of airbag presence (Roque, Jalayer, and Hasan 2021; Gabauer and Gabler 2010).

The last significant variable for an ROR crash involving a PC is the ejection of the occupant, which is highly intuitive. The model estimates the coefficient for ejection as 4.04 with a p-value of 0.0001 for serious or fatal injury. To put this into perspective, the probability of an occupant sustaining a serious or fatal injury in a single-vehicle ROR crash involving a passenger car increases by 9.36% if the occupant is partially or fully ejected. This finding aligns with past studies that demonstrate a significant increase in the probability of serious injury when an occupant is ejected (I. U. Khan and Vachal 2020; Viano, Parenteau, and Edwards 2007; Viano and Parenteau 2018).

Table 10. Mixed logit model results for SUV

Variable	**Coefficient**	**S.E.**	**p-Value**
Constant (MINJ)	0.07300*	0.15135	0.0902
Age (PDO)	-0.01317**	0.00551	0.0168
Alcohol (MINJ)	1.04952***	0.30718	0.0006
Alcohol (SFINJ)	1.17039***	0.34365	0.0007
Seatbelt (MINJ)	-1.72756***	0.29409	0.0000
Seatbelt (SFINJ)	-2.57183***	0.31419	0.0000
Vehicle model after 2010 (PDO)	-0.82367***	0.30859	0.0076
Vehicle model after 2010 (MINJ)	-0.89927***	0.29462	0.0023
Daylight (PDO)	-0.40467*	0.20688	0.0505
Travel speed b/w 61 and 75 mph (SFINJ)	0.52305**	0.24943	0.0360
Harmful event - Rollover (PDO)	-1.71703***	0.25553	0.0000
Harmful event–Concrete Barrier (PDO)	1.30911*	0.67187	0.0514
Harmful event–Concrete Barrier (MINJ)	1.52995**	0.65495	0.0195
Weekend (PDO)	0.40492*	0.20977	0.0536
Eject (SFINJ)	3.65703***	0.86062	0.0000
Heterogeneity in means			
Airbag (PDO)	-2.40659***	0.40833	0.0000
Airbag (MINJ)	-1.10676***	0.31430	0.0004
Weather (PDO)	1.22862***	0.36864	0.0009
Heterogeneity in variance			
Airbag (PDO) (MINJ)	3.14902***	0.73367	0.0000
Weather (PDO)	1.53542*	0.82195	0.0618
Number of observations	1667		

Table 10. Mixed logit model results for SUV (continued)

Variable	Coefficient	S.E.	p-Value
Log-likelihood		-920.26489	
Restricted log-likelihood		-1124.97898	
Chi squared		409.42818	
Note: ***, **, * indicates significance at 1%, 5%, 10% level, respectively. (PDO) defined for no injury; (MINJ) defined for minor injury; (SFINJ) defined for serious or fatal injury.			

Similarly, a mixed logit model was developed with heterogeneity in mean and variance for single-vehicle ROR crashes involving sport utility vehicles. A total of 1,667 observations were used in the model development. The log-likelihood and restricted log-likelihood values for the developed model are -920.265 and -1124.979, respectively. The mixed logit model estimation results for SUV can be seen in Table 10, and average pseudo-elasticity values in Table 11. The majority of the significant variables for SUVs are the same as for PCs, with a few exceptions, including daylight, concrete barriers, weather, and weekends. The probability of sustaining no injury in a ROR crash involving an SUV decreases by 3.1% if there is daylight. The presence or absence of daylight significantly affects driver visibility. Previous studies have also shown that daylight decreases the probability of high-severity crashes (Kim et al. 2013; Wu et al. 2014). The parameter for weather was found to be normally distributed with heterogeneity in mean and variance of 1.228 and 1.535, respectively, for no injury. This indicates that for a single-vehicle ROR crash involving an SUV, clear weather increases the probability of no injury by 2.53% (i.e., it decreases the probability of minor and serious or fatal injury). Concrete barriers were also significant, with a coefficient estimate of 1.309 and 1.529 for PDO and MINJ severity categories. This means that concrete barriers are associated with increased occurrences of no injury and minor injury. To further reduce the severity of injuries, the installation of energy-absorbing devices on the concrete barrier surface is recommended.

Table 11. Average Pseudo elasticities of significant variables for SUV

Variable	Elasticity effect		
	PDO	MINJ	SFINJ
Age (PDO)	-0.0658	-	-
Alcohol (MINJ)	-	0.0300	-
Alcohol (SFINJ)	-	-	0.0291
Seatbelt (MINJ)	-	-0.2204	-
Seatbelt (SFINJ)	-	-	-0.2281
Vehicle model after 2010 (PDO)	-0.0976	-	-
Vehicle model after 2010 (MINJ)	-	-0.1346	-
Weather (PDO)	0.0253	-	-
Daylight (PDO)	-0.0310	-	-
Travel speed b/w 61 and 75 mph (SFINJ)	-	-	0.0141
Harmful event - Rollover (PDO)	-0.0607	-	-
Harmful event–Concrete Barrier (PDO)	0.0125	-	-
Harmful event–Concrete Barrier (MINJ)	-	0.0150	-
Weekend (PDO)	0.0181	-	-
Eject (SFINJ)	-	-	0.0709
Airbag (PDO)	-0.0422	-	-
Airbag (MINJ)	-	-0.0183	-

The final mixed logit model was developed for a pickup/passenger van. There was a total of eleven significant variables, including two random parameters with a normal distribution. The model development included a total of 1612 observations. Careless driving was found to be significant for pickup drivers involved in single-vehicle ROR crashes. The probability of a pickup occupant sustaining no injury during an ROR crash increases by 1.10% due to careless driving. The estimated parameter of road surface condition was positively associated with no injury and minor injury. It is surprising to observe that wet/icy road conditions increase the probability of no injury and minor injury. However, one reasonable explanation for this could be that drivers become more alert in response to poor road surface conditions and drive carefully by reducing their speed. In contrast, drivers are more likely to drive at high speeds on normal road surface conditions, which may result in serious injury in the event of an ROR crash, as demonstrated in past research (Hou et al. 2019). Traveling at speeds between 61 and 75 mph was significant for pickups, with an estimated parameter of 1.160 for the serious or fatal injury category. The elasticity value shows that the probability of serious or fatal injury for pickup occupants involved in ROR crashes increases by 2.74% when traveling at speeds between 61 and 75 mph. This finding is intuitive, as the likelihood of an increase in injury severity is higher at higher speeds.

Table 12. Mixed logit model results for pickup/passenger van

Variable	**Coefficient**	**S.E.**	**p-Value**
Constant (PDO)	1.35121**	0.58949	0.0219
Age (PDO)	-0.03998***	0.01122	0.0004
Careless (PDO)	0.67899*	0.38217	0.0756
Vehicle model after 2010 (PDO)	0.77900**	0.35995	0.0305
Vehicle model after 2010 (MINJ)	0.83265**	0.33671	0.0134
Surface condition (PDO)	1.38978***	0.46724	0.0029
Surface condition (MINJ)	0.88208**	0.41895	0.0353
Travel speed b/w 61 and 75 mph (SFINJ)	1.16003***	0.35142	0.0010
Northeast (PDO)	1.65410**	0.82899	0.0460
Harmful event - Rollover (PDO)	-1.62051***	0.41761	0.0001
Airbag (PDO)	-1.81220***	0.44215	0.0000
Eject (SFINJ)	4.57945***	1.36518	0.0008
Heterogeneity in means			
Seatbelt (MINJ)	-1.77162***	0.46824	0.0002
Seatbelt (SFINJ)	-2.95515***	0.73713	0.0001
Alcohol (MINJ)	0.82151*	0.43924	0.0614
Alcohol (SFINJ)	0.95061*	0.49060	0.0527
Heterogeneity in variance			
Seatbelt (MINJ) (SFINJ)	2.43698***	0.83101	0.0034
Alcohol (MINJ) (SFINJ)	2.81807**	1.11604	0.0116
Number of observations		1612	
Log-likelihood		-915.26533	
Restricted log-likelihood		-1096.41506	
Chi squared		362.29946	

Note: ***, **, * indicates significance at 1%, 5%, 10% level, respectively.
(PDO) defined for no injury; (MINJ) defined for minor injury; (SFINJ) defined for serious or fatal injury.

Table 13. Average Pseudo elasticities of significant variables for pickup/passenger van

Variable	Elasticity effect		
	PDO	MINJ	SFINJ
Age (PDO)	-0.1759	-	-
Careless (PDO)	0.0110	-	-
Alcohol (MINJ)	-	0.0236	-
Alcohol (SFINJ)	-	-	0.0250
Seatbelt (SFINJ)	-	-	-0.0921
Seatbelt (MINJ)	-	-0.0973	-
Vehicle model after 2010 (PDO)	-0.0267	-	-
Vehicle model after 2010 (MINJ)	-	-0.0901	-
Surface condition (PDO)	0.0566	-	-
Surface condition (MINJ)	-	0.0315	-
Travel speed b/w 61 and 75 mph (SFINJ)	-	-	0.0274
Northeast (PDO)	0.0052	-	-
Harmful event - Rollover (PDO)	-0.0480	-	-
Airbag (PDO)	-0.0684	-	-
Eject (SFINJ)	-	-	0.0409

4.6. Conclusions

To recap, the primary objective of this study was to assess and compare the various factors influencing injury severity in single-vehicle run-off-road crashes across different vehicle categories, such as passenger cars, sport utility vehicles, passenger vans, and large trucks. For each of these vehicle types, three distinct mixed logit models were developed, incorporating variations in means and standard deviations. The majority of the significant variables were consistent across all three vehicle classes, including driver age, alcohol or drug usage, seatbelt utilization, airbag deployment, higher travel speeds, and vehicle model years post-2010. Notably, it was observed that as driver age increased, the likelihood of changes in injury severity outcomes was notably higher for pickup drivers when compared to passenger car and SUV drivers. Across all three vehicle classes, the utilization of seatbelts proved highly effective in mitigating injury severity during run-off-road crashes. Passenger cars exhibited a connection with heightened injury severity at relatively higher travel speeds (above 75 mph), especially when compared to SUVs and pickups traveling between 61 and 75 mph. For future research, it is advisable to separately examine each vehicle class to discern the factors contributing to injury severity in run-off-road crashes, using more extensive crash datasets than those employed in this study. This study serves to provide valuable insights into both the commonalities and distinctions among the factors impacting injury severity in run-off-road crashes across various vehicle classes. Furthermore, it offers practical guidance for decision-making among practitioners and safety stakeholders, aimed at enhancing safety measures in run-off-road incidents.

4.7 References

Abdel-Aty, Mohamed. 2003. "Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models." *Journal of Safety Research* 34 (5): 597–603. https://doi.org/10.1016/J.JSR.2003.05.009.

Al-Bdairi, Nabeel Saleem Saad, and Salvador Hernandez. 2017. "An Empirical Analysis of Run-off-Road Injury Severity Crashes Involving Large Trucks." *Accident Analysis and Prevention* 102 (May): 93–100. https://doi.org/10.1016/j.aap.2017.02.024.

Alnawmasi, Nawaf, and Fred Mannering. 2022. "The Impact of Higher Speed Limits on the Frequency and Severity of Freeway Crashes: Accounting for Temporal Shifts and Unobserved Heterogeneity." *Analytic Methods in Accident Research* 34 (June): 100205. https://doi.org/10.1016/J.AMAR.2021.100205.

Behnood, Ali, and Fred L. Mannering. 2016. "An Empirical Assessment of the Effects of Economic Recessions on Pedestrian-Injury Crashes Using Mixed and Latent-Class Models." *Analytic Methods in Accident Research* 12 (December): 1–17. https://doi.org/10.1016/J.AMAR.2016.07.002.

Bhat, Chandra R. 2003. "Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences." *Transportation Research Part B: Methodological* 37: 837–55. https://doi.org/10.1016/S0191-2615(02)00090-5.

Center for Statistics, National, and National Highway Traffic Safety Administration. 2023. “Crash Report Sampling System Analytical User’s Manual, 2016-2021.” https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813436

Cicchino, Jessica B. 2018. “Effects of Lane Departure Warning on Police-Reported Crash Rates.” *Journal of Safety Research* 66 (September): 61–70. https://doi.org/10.1016/J.JSR.2018.05.006.

Crash Report Sampling System | NHTSA. Accessed October 21, 2023. https://www.nhtsa.gov/crash-data-systems/crash-report-sampling-system.

Davis, Gary A., Sujay Davuluri, and Jianping Pei. 2006. “Speed as a Risk Factor in Serious Run-off-Road Crashes: Bayesian Case-Control Analysis with Case Speed Uncertainty.” *Journal of Transportation and Statistics* 9 (1): 17–28. https://www.bts.gov/archive/publications/journal_of_transportation_and_statistics/volume_09_number_01/paper_02/index

Fatality and Injury Reporting System Tool (FIRST). Accessed October 15, 2023. https://cdan.dot.gov/query.

Federal Highway Administration (FHWA), Roadway Departure Safety. Accessed October 15, 2023. https://safety.fhwa.dot.gov/roadway_dept/.

Gabauer, Douglas J., and Hampton C. Gabler. 2010. “The Effects of Airbags and Seatbelts on Occupant Injury in Longitudinal Barrier Crashes.” *Journal of Safety Research* 41 (1): 9–15. https://doi.org/10.1016/J.JSR.2009.10.006.

Gong, Linfeng, and Wei (David) Fan. 2017. “Modeling Single-Vehicle Run-off-Road Crash Severity in Rural Areas: Accounting for Unobserved Heterogeneity and Age Difference.” *Accident Analysis and Prevention* 101 (April): 124–34. https://doi.org/10.1016/j.aap.2017.02.014.

Hasan, Ahmed Sajid, Muntahith Mehadil Orvin, Mohammad Jalayer, Eric Heitmann, and Joseph Weiss. 2022. “Analysis of Distracted Driving Crashes in New Jersey Using Mixed Logit Model.” *Journal of Safety Research* 81: 166–74. https://doi.org/10.1016/j.jsr.2022.02.008.

Hickman, Jeffrey S., Feng Guo, Matthew C. Camden, Richard J. Hanowski, Alejandra Medina, and J. Erin Mabry. 2015. “Efficacy of Roll Stability Control and Lane Departure Warning Systems Using Carrier-Collected Data.” *Journal of Safety Research* 52 (February): 59–63. https://doi.org/10.1016/J.JSR.2014.12.004.

Hossain, Ahmed, Xiaoduan Sun, Shahrin Islam, Shah Alam, and Md Mahmud Hossain. 2023. “Identifying Roadway Departure Crash Patterns on Rural Two-Lane Highways under Different Lighting Conditions: Association Knowledge Using Data Mining Approach.” *Journal of Safety Research* 85: 52–65. https://doi.org/10.1016/j.jsr.2023.01.006.

Hou, Qinzhong, Xiaoyan Huo, Junqiang Leng, and Yuxing Cheng. 2019. “Examination of Driver Injury Severity in Freeway Single-Vehicle Crashes Using a Mixed Logit Model with Heterogeneity-in-Means.” *Physica A* 531: 121760. https://doi.org/10.1016/j.physa.2019.121760.

Hu, Wen, and Eric T. Donnell. 2011. “Severity Models of Cross-Median and Rollover Crashes on Rural Divided Highways in Pennsylvania.” *Journal of Safety Research* 42 (5): 375–82. https://doi.org/10.1016/j.jsr.2011.07.004.

Hu, Wen, and Eric T Donnell. 2010. “Median Barrier Crash Severity: Some New Insights.” *Accident Analysis and Prevention* 42: 1697–1704. https://doi.org/10.1016/j.aap.2010.04.009.

Khan, Ihsan Ullah, and Kimberly Vachal. 2020. “Factors Affecting Injury Severity of Single-Vehicle Rollover Crashes in the United States.” *Traffic Injury Prevention* 21 (1): 66–71. https://doi.org/10.1080/15389588.2019.1696962.

Khan, Shinthia Azmeri, Shamsunnahar Yasmin, and Mazharul Haque. 2023. “Effects of Design Consistency Measures and Roadside Hazard Types on Run-off-Road Crash Severity: Application of Random Parameters Hierarchical Ordered Probit Model.” https://doi.org/10.1016/j.amar.2023.100300.

Khorashadi, Ahmad, Debbie Niemeier, Venky Shankar, and Fred Mannering. 2005. “Differences in Rural and Urban Driver-Injury Severities in Accidents Involving Large-Trucks: An Exploratory Analysis.” *Accident Analysis and Prevention* 37 (5): 910–21. https://doi.org/10.1016/j.aap.2005.04.009.

Kim, Joon-Ki, Gudmundur F Ulfarsson, Sungyop Kim, and Venkataraman N Shankar. 2013. “Driver-Injury Severity in Single-Vehicle Crashes in California: A Mixed Logit Analysis of Heterogeneity Due to Age and Gender.” *Accident Analysis and Prevention* 50: 1073–81. https://doi.org/10.1016/j.aap.2012.08.011.

Lee, Jinsun, and Fred Mannering. 2002. “Impact of Roadside Features on the Frequency and Severity of Run-off-Roadway Accidents: An Empirical Analysis.” *Accident Analysis and Prevention* 34 (2): 149–61. https://doi.org/10.1016/S0001-4575(01)00009-4.

Li, Zhenning, Yusheng Ci, Cong Chen, Guohui Zhang, Qiong Wu, Zhen (Sean) Qian, Panos D. Prevedouros, and David T. Ma. 2019. “Investigation of Driver Injury Severities in Rural Single-Vehicle Crashes under Rain Conditions Using Mixed Logit and Latent Class Models.” *Accident Analysis & Prevention* 124 (March): 219–29. https://doi.org/10.1016/J.AAP.2018.12.020.

Liu, Cejun, and Rajesh Subramanian. 2009. “Factors Related to Fatal Single-Vehicle Run-Off-Road Crashes.” https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811232

Liu, Cejun, and Tony Jianqiang Ye. 2011. “Run-Off-Road Crashes: An On-Scene Perspective.” https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811500

Liu, Pengfei, and Wei Fan. 2020. “Exploring Injury Severity in Head-on Crashes Using Latent Class Clustering Analysis and Mixed Logit Model: A Case Study of North Carolina.” *Accident Analysis & Prevention* 135 (February): 105388. https://doi.org/10.1016/J.AAP.2019.105388.

Lord, Dominique, Marcus A Brewer, Kay Fitzpatrick, Srinivas R Geedipally, and Yichuan Peng. 2011. “Analysis of Roadway Departure Crashes on Two-Lane Rural Roads in Texas.” http://tti.tamu.edu/documents/0-6031-1.pdf.

Malyshkina, Nataliya V., and Fred Mannering. 2008. “Effect of Increases in Speed Limits on Severities of Injuries in Accidents.” *Transportation Research Record*, no. 2083 (January): 122–27. https://doi.org/10.3141/2083-14.

Mannering, Fred L, and Chandra R Bhat. 2013. “Analytic Methods in Accident Research: Methodological Frontier and Future Directions.” https://doi.org/10.1016/j.amar.2013.09.001.

Moore, Darren N., William H. Schneider IV, Peter T. Savolainen, and Mohamadreza Farzaneh. 2011. “Mixed Logit Analysis of Bicyclist Injury Severity Resulting from Motor Vehicle Crashes at Intersection and Non-Intersection Locations.” *Accident Analysis & Prevention* 43 (3): 621–30. https://doi.org/10.1016/J.AAP.2010.09.015.

Palamara, Peter, Michelle Broughton, Michelle Fraser, P, Palamara, M, Broughton, and M Fraser. 2013. “An Investigation of Urban Area Run-Off-Road Crashes in Western Australia 2005–2009. Report No. RR 10-005.” https://www.wagov.pipeline.preproduction.digital.wa.gov.au/system/files/2021-08/Urban-area-crash-report.pdf

Peng, Yiyun, and Linda Ng Boyle. 2012. “Commercial Driver Factors in Run-off-Road Crashes.” *Transportation Research Record: Journal of the Transportation Research Board* 2281 (1): 128–32. https://doi.org/10.3141/2281-16.

Riexinger, Luke E., Rini Sherony, and Hampton C. Gabler. 2019. “Residual Road Departure Crashes after Full Deployment of LDW and LDP Systems.” *Traffic Injury Prevention* 20 (sup1): S177–81. https://doi.org/10.1080/15389588.2019.1603375.

Roque, Carlos, and João Lourenço Cardoso. 2014. “Investigating the Relationship between Run-off-the-Road Crash Frequency and Traffic Flow through Different Functional Forms.” *Accident Analysis and Prevention* 63 (February): 121–32. https://doi.org/10.1016/j.aap.2013.10.034.

Roque, Carlos, Mohammad Jalayer, and Ahmed Sajid Hasan. 2021. “Investigation of Injury Severities in Single-Vehicle Crashes in North Carolina Using Mixed Logit Models.” https://doi.org/10.1016/j.jsr.2021.02.013.

Roy, Uttara, and Sunanda Dissanayake. 2011. “Comparison of Factors Associated with Run-Off-Road and Non-Run-Off-Road Crashes in Kansas.” *Journal of the Transportation Research Forum* 50 (2): 69–86. https://doi.org/10.22004/AG.ECON.207298.

Savolainen, Peter T, Fred L Mannering, Dominique Lord, and Mohammed A Quddus. 2011. “The Statistical Analysis of Highway Crash-Injury Severities: A Review and Assessment of Methodological Alternatives.” *Accident Analysis and Prevention* 43: 1666–76. https://doi.org/10.1016/j.aap.2011.03.025.

Sawtelle, Alainie, Mohammadali Shirazi, Per Erik Garder, and Jonathan Rubin. 2023. “Driver, Roadway, and Weather Factors on Severity of Lane Departure Crashes in Maine.” *Journal of Safety Research* 84 (February): 306–15. https://doi.org/10.1016/J.JSR.2022.11.006.

Sternlund, Simon, Johan Strandroth, Matteo Rizzi, Anders Lie, and Claes Tingvall. 2017. “The Effectiveness of Lane Departure Warning Systems—A Reduction in Real-World Passenger

Car Injury Crashes.” *Traffic Injury Prevention* 18 (2): 225–29. https://doi.org/10.1080/15389588.2016.1230672.

Train, Kenneth. 2009. “Discrete Choice Methods with Simulation, Cambridge University Press.” 2009. https://eml.berkeley.edu/books/choice2.html.

Viano, David C., and Chantal S. Parenteau. 2018. “Rollover Injury in Vehicles with High-Strength-to-Weight Ratio (SWR) Roofs, Curtain and Side Airbags, and Other Safety Improvements.” *Traffic Injury Prevention* 19 (7): 734–40. https://doi.org/10.1080/15389588.2018.1482489.

Viano, David C., Chantal S. Parenteau, and Mark L. Edwards. 2007. “Rollover Injury: Effects of Near- and Far-Seating Position, Belt Use, and Number of Quarter Rolls.” *Traffic Injury Prevention* 8 (4): 382–92. https://doi.org/10.1080/15389580701583379.

Washington, Simon P., Matthew G. Karlaftis, and Fred L. Mannering. 2010. *Statistical and Econometric Methods for Transportation Data Analysis, Second Edition*. *Statistical and Econometric Methods for Transportation Data Analysis, Second Edition*. CRC Press. https://doi.org/10.1198/tech.2004.s238.

Wu, Qiong, Feng Chen, Guohui Zhang, Xiaoyue Cathy Liu, Hua Wang, and Susan M Bogus. 2014. “Mixed Logit Model-Based Driver Injury Severity Investigations in Single-and Multi-Vehicle Crashes on Rural Two-Lane Highways.” *Accident Analysis and Prevention* 72: 105–15. https://doi.org/10.1016/j.aap.2014.06.014.

Yasmin, Shamsunnahar, Naveen Eluru, Chandra R. Bhat, and Richard Tay. 2014. “A Latent Segmentation Based Generalized Ordered Logit Model to Examine Factors Influencing Driver Injury Severity.” *Analytic Methods in Accident Research* 1 (January): 23–38. https://doi.org/10.1016/J.AMAR.2013.10.002.

Ye, Fan, and Dominique Lord. 2014. “Comparing Three Commonly Used Crash Severity Models on Sample Size Requirements: Multinomial Logit, Ordered Probit and Mixed Logit Models.” *Analytic Methods in Accident Research* 1 (January): 72–85. https://doi.org/10.1016/J.AMAR.2013.03.001.

Zeng, Tong. 2016. “Using Halton Sequences in Random Parameters Logit Models.” *Journal of Statistical and Econometric Methods* 5 (1): 1792–6939. http://ssrn.com/abstract=2741934https://ssrn.com/abstract=2741934Electroniccopyavailable at:http://ssrn.com/abstract=2741934.

CHAPTER 5: CONCLUSIONS

This dissertation investigated factors contributing to injury severity in three distinct types of single-vehicle crashes: rollover crashes, rural lane departure crashes, and run-off-road (ROR) crashes. By employing advanced statistical modeling techniques, spatial analysis methods, and a comprehensive examination of crash data, this research offers invaluable insights into the underlying causes and potential countermeasures for mitigating the consequences of these high-risk crash scenarios.

Integrating the findings from the three chapters, several overarching themes and implications emerge, highlighting the critical role of driver behavior, vehicle characteristics, roadway design, and targeted interventions in improving road safety.

5.1. Driver Behavior and Occupant Protection

Across all three types of crashes examined, driver behavior and occupant protection measures emerged as pivotal factors influencing injury severity. Speeding, impaired driving (alcohol/drug use), and lack of seatbelt usage consistently contributed to an increased likelihood of severe or fatal injuries. These findings underscore the importance of promoting responsible driving practices, enhancing enforcement efforts, and implementing educational campaigns to raise awareness about the life-saving benefits of seatbelt usage.

Notably, the analysis of rollover crashes (Chapter 2) and ROR crashes (Chapter 4) identified seatbelt usage as a highly effective countermeasure in mitigating injury severity. This finding aligns with well-established research on the protective capabilities of seatbelts and highlights the need for continued efforts to promote their widespread adoption.

Furthermore, the study on ROR crashes (Chapter 4) revealed that certain driver demographics, such as older age, particularly among pickup drivers, were associated with a higher risk of severe injuries. This insight suggests the need for targeted educational programs and age-specific interventions to address the unique challenges faced by different driver groups.

5.2. Vehicle Characteristics and Technology

The research findings also shed light on the role of vehicle characteristics and technological advancements in enhancing occupant safety. For instance, the deployment of airbags was found to be beneficial in reducing the severity of injuries in both rollover crashes (Chapter 2) and ROR crashes (Chapter 4). This underscores the importance of promoting the adoption of advanced safety features and encouraging the integration of new technologies into vehicle design.

Additionally, the analysis of ROR crashes (Chapter 4) revealed that newer vehicle models (post-2010) were associated with a lower risk of severe injuries, potentially due to improved crashworthiness and safety features. This finding highlights the continuous progress in vehicle safety engineering and the potential benefits of accelerating the turnover of aging vehicle fleets.

5.3. Roadway Design and Infrastructure Improvements

The research findings also emphasize the crucial role of roadway design and infrastructure improvements in mitigating the consequences of single-vehicle crashes. For rollover crashes (Chapter 2), factors such as higher posted speed limits, undulating terrain, and roadside/median rollovers were identified as contributing to increased injury severity. These insights suggest the need for targeted interventions, such as speed limit reductions in high-risk areas, roadside embankment flattening, and the installation of barriers or guardrails to prevent roadside departures.

Furthermore, the spatial analysis of rural lane departure crashes (Chapter 3) identified hotspots and high-risk areas within the study region. This information can guide the prioritization of infrastructure improvements, such as rumble strip installations, enhanced delineation, and advisory speed markings in curves, effectively addressing localized risk factors.

5.4. Integrated Approach and Data-Driven Decision-Making

One of the overarching implications of this research is the need for an integrated approach to road safety that combines engineering, enforcement, and educational strategies. By leveraging the insights gained from comprehensive crash data analysis, spatial modeling, and vehicle-specific factors, transportation agencies and safety stakeholders can develop targeted and effective countermeasures tailored to specific crash types, road conditions, and driver demographics.

Moreover, the methodologies employed in this research, including advanced statistical modeling and spatial analysis techniques, demonstrate the value of data-driven decision-making in the field of traffic safety. By integrating diverse data sources, such as crash records, citation data, and geographic information systems (GIS), this research provides a framework for identifying high-risk areas, quantifying crash clusters, and evaluating the effectiveness of interventions over time.

5.5. Future Research and Policy Implications

While this dissertation offers significant contributions to the understanding of single-vehicle crash dynamics and injury severity factors, it also highlights several avenues for future research and policy implications.

First, the study on ROR crashes (Chapter 4) suggests the need for vehicle-specific analyses to gain deeper insights into the unique factors contributing to injury severity for different vehicle classes. By leveraging more extensive crash datasets and focusing on individual vehicle types, future research can inform targeted safety initiatives and design standards tailored to the specific characteristics of passenger cars, SUVs, pickups, and other vehicle categories.

Second, the findings related to driver behavior and occupant protection underscore the importance of ongoing public education campaigns, enforcement strategies, and policy interventions. Initiatives such as graduated licensing programs, stricter impaired driving laws, and incentives for the adoption of advanced vehicle safety features can play a crucial role in fostering responsible driving practices and enhancing occupant protection.

Third, the spatial analysis techniques employed in this research (Chapter 3) highlight the potential for integrating geographic information systems (GIS) and spatial modeling into road safety decision-making processes. By identifying high-risk areas and crash hotspots, transportation agencies can prioritize infrastructure improvements, allocate resources more effectively, and collaborate with local communities to address localized safety concerns.

Finally, the findings related to roadway design and infrastructure improvements (Chapters 2 and 3) emphasize the need for continued investment in road safety engineering, including the implementation of proven countermeasures such as rumble strips, safety edges, and enhanced delineation. Furthermore, the adoption of innovative technologies, such as connected and automated vehicle systems, could potentially mitigate the risks associated with single-vehicle crashes by enhancing vehicle stability, collision avoidance, and occupant protection.

In conclusion, this dissertation provides a comprehensive examination of the factors influencing injury severity in single-vehicle crashes, offering valuable insights and implications for transportation agencies, policymakers, vehicle manufacturers, and safety stakeholders. By integrating the findings from multiple perspectives, this research highlights the importance of a collaborative, data-driven approach to road safety, emphasizing the roles of driver behavior, vehicle technology, roadway design, and targeted interventions in reducing the devastating consequences of these high-risk crash scenarios.