IDENTIFYING EFFECTIVE GEOMETRIC AND TRAFFIC FACTORS TO PREDICT CRASHES AT HORIZONTAL CURVE SECTIONS

by

HOJR MOMENI

B.S., University of Kurdistan, 2003
M.S., International University of Ghazvin, 2008

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Civil Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Abstract

Driver workload increases on horizontal curves due to more complicated navigation compared to navigation on straight roadway sections. Although only a small portion of roadways are horizontal curve sections, approximately 25% of all fatal highway crashes occur at horizontal curve sections. According to the Fatality Analysis Reporting System (FARS) database, fatalities associated with horizontal curves were more than 25% during last years from 2008 to 2014, reinforcing that investigation of horizontal curve crashes and corresponding safety improvements are crucial study topics within the field of transportation safety. Improved safety of horizontal curve sections of rural transportation networks can contribute to reduced crash severities and frequencies. Statistical methods can be utilized to develop crash prediction models in order to estimate crashes at horizontal curves and identify contributing factors to crash occurrences, thereby correlating to the primary objectives of this research project.

Primary data analysis for 221 randomly selected horizontal curves on undivided two-lane highways with Poisson regression method revealed that annual average daily traffic (AADT), heavy vehicle percentage, degree of curvature, and difference between posted and advisory speeds affect crash occurrence at horizontal curves. The data, however, were relatively overdispersed, so the negative binomial (NB) regression method was utilized. Results indicated that AADT, heavy vehicle percentage, degree of curvature, and long tangent length significantly affect crash occurrence at horizontal curve sections. A new dataset consisted of geometric and traffic data of 5,334 horizontal curves on the entire state transportation network including undivided and divided highways provided by Kansas Department of Transportation (KDOT) Traffic Safety Section as well as crash data from the Kansas Crash and Analysis Reporting System (KCARS) database were used to analyze the single vehicle (SV) crashes. An R software package was used to write a code and combine required information from aforementioned databases and create the dataset for 5,334 horizontal curves on the entire state transportation network. Eighty percent of crashes including 4,267 horizontal curves were randomly selected for data analysis and remaining 20% horizontal curves (1,067 curves) were used for data validation. Since the results of the Poisson regression model showed overdispersion of crash data and many horizontal curves had zero crashes during the study period from 2010 to 2014, NB, zero-inflated Poisson (ZIP), and zero-inflated negative binomial (ZINB) methods were used for data analysis.

Total number of crashes and severe crashes were analyzed with the selected methods. Results of data analysis revealed that AADT, heavy vehicle percentage, curve length, degree of curvature, posted speed, difference between posted and advisory speed, and international roughness index influenced single vehicle crashes at 4,267 randomly selected horizontal curves for data analysis. Also, AADT, degree of curvature, heavy vehicle percentage, posted speed, being a divided roadway, difference between posted and advisory speeds, and shoulder width significantly influenced severe crash occurrence at selected horizontal curves. The goodness-of-fit criteria showed that the ZINB model more accurately predicted crash numbers for all crash groups at the selected horizontal curve sections. A total of 1,067 horizontal curves were used for data validation, and the observed and predicted crashes were compared for all crash groups and data analysis methods. Results of data validation showed that ZINB models for total crashes and severe crashes more accurately predicted crashes at horizontal curves.

This study also investigated the effect of speed limit change on horizontal curve crashes on K-5 highway in Leavenworth County, Kansas. A statistical t-test proved that crash data from years 2006 to 2012 showed only significant reduction in equivalent property damage only (EPDO) crash rate for adverse weather condition at 5% significance level due to speed limit reduction in June 2009. However, the changes in vehicles speeds after speed limit change and other information such as changes in surface pavement condition were not available.

According to the results of data analysis for 221 selected horizontal curves on undivided two-lane highways, tangent section length significantly influenced total number of crashes. Therefore, providing more information about upcoming changes in horizontal alignment of the roadway via doubling up warning sings, using bigger sings, using materials with higher retroreflectivity, or flashing beacons were recommended for horizontal curves with long tangent section lengths and high number of crashes. Also, presence of rumble strips and wider shoulders significantly and negatively influenced severe SV crashes at horizontal curve sections; therefore, implementing rumble strips and widening shoulders for horizontal curves with high number of severe SV crashes were recommended.

IDENTIFYING EFFECTIVE GEOMETRIC AND TRAFFIC FACTORS TO PREDICT CRASHES AT HORIZONTAL CURVE SECTIONS

by

HOJR MOMENI

B.S., University of Kurdistan, 2003
M.S., International University of Ghazvin, 2008

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Civil Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Co-Major Professor
Dr. Sunanda Dissanayake

Approved by:

Co-Major Professor
Dr. Malgorzata J. Rys

Copyright

Abstract

Driver workload increases on horizontal curves due to more complicated navigation compared to navigation on straight roadway sections. Although only a small portion of roadways are horizontal curve sections, approximately 25% of all fatal highway crashes occur at horizontal curve sections. According to the Fatality Analysis Reporting System (FARS) database, fatalities associated with horizontal curves were more than 25% during last years from 2008 to 2014, reinforcing that investigation of horizontal curve crashes and corresponding safety improvements are crucial study topics within the field of transportation safety. Improved safety of horizontal curve sections of rural transportation networks can contribute to reduced crash severities and frequencies. Statistical methods can be utilized to develop crash prediction models in order to estimate crashes at horizontal curves and identify contributing factors to crash occurrences, thereby correlating to the primary objectives of this research project.

Primary data analysis for 221 randomly selected horizontal curves on undivided two-lane two-way highways with Poisson regression method revealed that annual average daily traffic (AADT), heavy vehicle percentage, degree of curvature, and difference between posted and advisory speeds affect crash occurrence at horizontal curves. The data, however, were relatively overdispersed, so the negative binomial (NB) regression method was utilized. Results indicated that AADT, heavy vehicle percentage, degree of curvature, and long tangent length significantly affect crash occurrence at horizontal curve sections. A new dataset consisted of geometric and traffic data of 5,334 horizontal curves on the entire state transportation network including undivided and divided highways provided by Kansas Department of Transportation (KDOT) Traffic Safety Section as well as crash data from the Kansas Crash and Analysis Reporting System (KCARS) database were used to analyze the single vehicle (SV) crashes. An R software package was used to write a code and combine required information from aforementioned databases and create the dataset for 5,334 horizontal curves on the entire state transportation network. Eighty percent of crashes including 4,267 horizontal curves were randomly selected for data analysis and remaining 20% horizontal curves (1,067 curves) were used for data validation. Since the results of the Poisson regression model showed overdispersion of crash data and many horizontal curves had zero crashes during the study period from 2010 to 2014, NB, zero-inflated Poisson (ZIP), and zero-inflated negative binomial (ZINB) methods were used for data analysis.

Total number of crashes and severe crashes were analyzed with the selected methods. Results of data analysis revealed that AADT, heavy vehicle percentage, curve length, degree of curvature, posted speed, difference between posted and advisory speed, and international roughness index influenced single vehicle crashes at 4,267 randomly selected horizontal curves for data analysis. Also, AADT, degree of curvature, heavy vehicle percentage, posted speed, being a divided roadway, difference between posted and advisory speeds, and shoulder width significantly influenced severe crash occurrence at selected horizontal curves. The goodness-of-fit criteria showed that the ZINB model more accurately predicted crash numbers for all crash groups at the selected horizontal curve sections. A total of 1,067 horizontal curves were used for data validation, and the observed and predicted crashes were compared for all crash groups and data analysis methods. Results of data validation showed that ZINB models for total crashes and severe crashes more accurately predicted crashes at horizontal curves.

This study also investigated the effect of speed limit change on horizontal curve crashes on K-5 highway in Leavenworth County, Kansas. A statistical t-test proved that crash data from years 2006 to 2012 showed only significant reduction in equivalent property damage only (EPDO) crash rate for adverse weather condition at 5% significance level due to speed limit reduction in June 2009. However, the changes in vehicles speeds after speed limit change and other information such as changes in surface pavement condition were not available.

According to the results of data analysis for 221 selected horizontal curves on undivided two-lane highways, tangent section length significantly influenced total number of crashes. Therefore, providing more information about upcoming changes in horizontal alignment of the roadway via doubling up warning sings, using bigger sings, using materials with higher retroreflectivity, or flashing beacons were recommended for horizontal curves with long tangent section lengths and high number of crashes. Also, presence of rumble strips and wider shoulders significantly and negatively influenced severe SV crashes at horizontal curve sections; therefore, implementing rumble strips and widening shoulders for horizontal curves with high number of severe SV crashes were recommended.

Table of Contents

List of Figures xi

List of Tables xiii

Acknowledgements xv

Dedication xvi

Chapter 1. Introduction 1

1.1. Importance of Horizontal Curves 1

1.2. Problem Statement 8

1.3. Objectives 8

Chapter 2. Literature Review 10

2.1. Horizontal Curve Countermeasures 10

2.1.1. Low-Cost Countermeasures 10

2.1.1.1. Centerline and Edge Line 11

2.1.1.2. Horizontal Alignment Signs and Advisory Speed Signs 12

2.1.1.2.1. Larger Signs and Doublingup Signs 14

2.1.1.2.2. High Retro-reflective Intensity and Fluorescent Yellow Sheeting 15

2.1.1.3. Combination Horizontal Alignment Signs and Advisory Speed Signs 16

2.1.1.4. Chevrons and One-Direction Arrow Signs 16

2.1.1.5. Delineators 19

2.1.1.6. Profile Thermoplastic Markings and Raised Pavement Markings 20

2.1.1.7. Reflective Barrier Delineation 22

2.1.1.8. Speed Limit Advisory Marking 24

2.1.1.9. Optical Speed Bars 26

2.1.1.10. Rumble Strips 27

2.1.1.11. Centerline Rumble Strips 28

2.1.1.12. Edge Line or Shoulder Rumble Strips/Stripes 29

2.1.1.13. Roadway Rumble Strips 30

2.1.2. Intermediate and High-Cost Treatments 31

2.1.2.1. Flashing Beacons 31

2.1.2.2. Dynamic Curve Warning System 33

2.1.2.3. Paved Shoulders and Widening Shoulders ... 34
2.1.2.4. Shoulder Drop-Off Mitigation (Safety Edge) ... 35
2.1.2.5. Installation or Improvement of Lighting ... 36
2.1.2.6. Skid Resistant Pavement Surface ... 37
2.2. Crash Prediction Models ... 38
Chapter 3. Methodology and Data ... 45
3.1. Horizontal Curves for Preliminary Analysis ... 45
3.2. Poisson Regression Model ... 47
3.2.1. Overdispersion ... 48
3.3. Negative Binomial Regression ... 49
3.3.1. Maximum Likelihood ... 49
3.4. Zero-inflated Model ... 50
3.5. Goodness of Fit ... 51
3.6. Data Collection ... 52
3.6.1. Initial Analysis ... 52
3.6.1.1. Crashes at 221 Horizontal Curves ... 52
3.6.2. New KDOT Horizontal Curve Inventory ... 53
3.6.2.1. Crashes at Horizontal Curves in KDOT Inventory ... 55
3.6.2.2. Selecting Data for Data Analysis and Data Validation ... 57
Chapter 4. Effect of Speed Limit Change on Horizontal Curve Crashes on K-5 Highway in Leavenworth County, Kansas ... 64
4.1. Countermeasures on Horizontal Curves on K-5 ... 64
4.2. Roadside Hazard Rating ... 66
4.3. Superelevation of Horizontal Curves ... 67
4.4. K-5 Highway Curve-Related Crash Analysis ... 68
4.5. Speed Limit Reduction on K-5 Highway ... 68
4.6. T-test Analysis for Speed Limit Reduction ... 74
Chapter 5. Results and Discussion ... 75
5.1. Data Analysis and Regression Models for Datasets ... 75
5.1.1. First Dataset ... 75
5.1.1.1. Poisson Regression Model ... 76

5.1.1.2. Negative Binomial Regression Model 79
5.1.2. KDOT's Horizontal Curve Inventory 81
5.1.2.1. Data Analysis for KDOT New Horizontal Curves Dataset 82
5.1.2.1.1. Negative Binomial Regression for the KDOT Horizontal Curve Inventory Dataset 83
5.1.2.1.2. Zero-inflated Poisson Regression for the KDOT Horizontal Curve Inventory Dataset 91
5.1.2.1.3. Zero-inflated Negative Binomial Regression for the KDOT Horizontal Curve Inventory Dataset 97
5.1.2.1.4. Comparison of Developed Models 105
5.2. Data Validation for the New KDOT Horizontal Curve Inventory 108
Chapter 6. Summary, Conclusions and Recommendations 111
6.1. Summary 111
6.2. Conclusions 113
6.3. Countermeasures to Improve Safety of Horizontal Curves Based on the Developed Prediction Models 116
6.4. Recommendations for Future Studies 118
References 120
Appendix A. Summary of Prediction Model Studies 126
Appendix B. R Codes for Creating a Dataset of KDOT Horizontal Curves Inventory 131
Appendix C. SAS Codes for Data Analysis 138
Appendix D. Observed and Predicted Crashes for Randomly Selected Horizontal Curves for Data Validation 147

List of Figures

Figure 1.1 Percentage of the US fatal crashes occurred at horizontal curves (2008-2014)............ 3
Figure 1.2 Percentage of ROR crashes from horizontal curve fatal crashes (2008-2014) 4
Figure 1.3 Curve-related PDO crashes in Kansas (2008-2014)... 5
Figure 1.4 Curve-related injury crashes in Kansas (2008-2014) 5
Figure 1.5 Curve-related fatal crashes in Kansas (2008-2014).. 6
Figure 1.6 Curve-related crashes in Kansas (2008-2014).. 6
Figure 2.1 Advanced warning signs for horizontal curves 12
Figure 2.2 Combination curve alignment sign and advisory speed sign 16
Figure 2.3 Curve delineation signs 17
Figure 2.4 Post-mounted delineators 19
Figure 2.5 Profile thermoplastic markings.. 21
Figure 2.6 Raised pavement markings... 22
Figure 2.7 Reflective panels and mounted reflectors.. 23
Figure 2.8 Reflective tape on object close to the road 24
Figure 2.9 Speed limit advisory pavement marking 25
Figure 2.10 PennDOT curve advance marking.. 25
Figure 2.11 Optical speed bars... 26
Figure 2.12 Application of on-pavement chevrons.. 27
Figure 2.13 CLRS pattern 28
Figure 2.14 Example of CLRS... 29
Figure 2.15 Example of SRS 30
Figure 2.16 Example of roadway rumble strip 30
Figure 2.17 Use of flashing beacons with a sign 31
Figure 2.18 Chevron enhanced with LEDs.. 32
Figure 2.19 Use of in-pavement LED markers.. 32
Figure 2.20 Dynamic curve warning system in Texas.. 33
Figure 2.21 DCWS... 33
Figure 2.22 Chevrons enhanced with LEDs 34
Figure 2.23 Inside shoulder widening.. 35
Figure 2.24 Safety edges.. 36
Figure 2.25 Lighting to improve visibility on a curve 37
Figure 2.26 Application of skid resistant material on a curve 38
Figure 3.1 Measuring radius and length of curve using Google Maps and AutoCAD................. 46
Figure 3.2 Measuring lengths of tangent sections for a curve using Google Maps and AutoCAD 47
Figure 3.3 Total crash frequencies for 221 horizontal curves on two-lane highways 53
Figure 3.4 Crashes at horizontal curve sections based on severity.. 55
Figure 3.5 Crashes based on crash type and crash severity 57

Figure 3.6 Total crash frequencies for 5,334 horizontal curves in Kansas 59
Figure 3.7 Total crash frequencies for 4,267 horizontal curves selected for data analysis 60
Figure 3.8 Total crash frequencies for 1,067 horizontal curves selected for data validation 60
Figure 3.9 Severe (KA) crash frequencies for 5,334 horizontal curves in Kansas 61
Figure 3.10 Severe crash (KA) frequencies for 4,267 horizontal curves selected for data analysis 61
Figure 3.11 Severe crash (KA) frequencies for 1,067 horizontal curves selected for data validation.. 62
Figure 3.12 Severe (KAB) crash frequencies for 5,334 horizontal curves in Kansas 62
Figure 3.13 Severe crash (KAB) frequencies for 4,267 horizontal curves selected for data analysis.. 63
Figure 3.14 Severe crash (KAB) frequencies for 1,067 horizontal curves selected for data validation.. 63
Figure 4.1 Segment of K-5 highway with speed limit change... 64
Figure 4.2 Crash percentage before and after speed limit reduction 71
Figure 4.3 Crashes based on severity before and after speed limit change 71
Figure 4.4 EPDO crash changes before and after speed limit change for various conditions...... 73

List of Tables

Table 1.1 PDO, injury, fatal, and total number of crashes for the entire Kansas roadway network and curve sections from 2008 to 2014 7
Table 2.1 Guidelines for advance placement of curve warning signs 14
Table 2.2 Sizes of warning signs in inches 15
Table 2.3 Typical spacing of chevron alignment signs on horizontal curves 18
Table 2.4 Recommended spacing for delineators at horizontal curves 20
Table 2.5 Estimated cost of pavement markings in Texas 22
Table 2.6 Variables used in horizontal curve crashes prediction models 43
Table 2.7 Variables used in non-horizontal curve crashes prediction models 44
Table 3.1 Geometric and traffic variable descriptions for initial data collection 52
Table 3.2 Geometric and traffic variable descriptions for KDOT horizontal curve inventory 54
Table 3.3 Number of total, KA, and KAB crashes at horizontal curves based on crash type 56
Table 3.4 Comparison of statistical characteristics of randomly selected curves for data analysis and data validation 58
Table 4.1 Number of curves with specific treatment 65
Table 4.2 RHR for curves on K-5 highway, Leavenworth County, Kansas 66
Table 4.3 Superelevations of studied horizontal curves 67
Table 4.4 Number of crashes for each crash group before and after speed limit reduction 70
Table 4.5 Results of t-test for crash frequencies and crash rate at K-5 horizontal curves 74
Table 5.1 Poisson regression model variables for the first dataset with 221 horizontal curves 77
Table 5.2 Goodness of fit for Poisson regression model for the first dataset with 221 horizontal curves 79
Table 5.3 NB regression model variables for the first dataset with 221 horizontal curves 80
Table 5.4 Goodness of fit for the NB regression model for the first dataset with 221 horizontal curves 81
Table 5.5 NB regression model variables for total single-vehicle crashes of KDOT horizontal curve inventory 84
Table 5.6 Goodness of fit for the NB regression model of KDOT horizontal curve inventory 85
Table 5.7 NB regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory 86
Table 5.8 Goodness of fit for the NB regression model of KDOT horizontal curve inventory 88
Table 5.9 NB regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory 89
Table 5.10 Goodness of fit for the NB regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory 91
Table 5.11 Goodness of fit for the ZIP regression model of total crashes of KDOT horizontal curve inventory 92

Table 5.12 ZIP regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory ... 93
Table 5.13 Goodness of fit for the ZIP regression model of severe crashes (K and A levels) of KDOT horizontal curve inventory ... 94
Table 5.14 ZIP regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory ... 95
Table 5.15 Goodness of fit for the ZIP regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory ... 97
Table 5.16 ZINB regression model variables for total single-vehicle crashes of KDOT horizontal curve inventory ... 98
Table 5.17 Goodness of fit for the ZINB regression model of KDOT horizontal curve inventory ... 100
Table 5.18 ZINB regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory ... 101
Table 5.19 Goodness of fit for the ZINB regression model of KDOT horizontal curve inventory ... 102
Table 5.20 ZINB regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory ... 103
Table 5.21 Goodness of fit for the ZINB regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory ... 105
Table 5.22 Variables used to predict total crashes in the selected models ... 106
Table 5.23 Variables used to predict severe crashes (K and A levels) in the selected models... 107
Table 5.24 Variables used to predict severe crashes (K, A, and B levels) in the selected models ... 108
Table 5.25 Comparison of observed and predicted crashes at selected horizontal curve sections for data validation ... 109
Table 6.1 Recommended countermeasures for the studied horizontal curves based on the crash types ... 118
Table A.1 Studies related to crash prediction models at horizontal curve sections ... 126
Table A.2 Factors used in crash prediction models at rural highway sections ... 128
Table D.1 Sample of Comparison of 1,067 observed and predicted crashes for various crash groups and models ... 147

Acknowledgements

Foremost, I would like to express my immeasurable appreciation and sincere gratitude to Dr. Sunanda Dissanayake, my advisor, for the continuous support of my PhD study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this dissertation. I am also deeply grateful to Dr. Malgorzata J. Rys, my co-advisor, for her insightful comments and constructive advices. I would like to acknowledge my committee members, Dr. Robert W. Stokes, Dr. Nathan Albin, and Dr. Eugene R. Russell. I am also thankful to Dr. Russell, my first supervisor at Kansas State University, for his continuous encouragement and guidance.

I would like to thank the Kansas Department of Transportation (KDOT) for funding the major part of this research project. Also, I am indebted to Mr. Steven Buckley, Ms. Tina Cramer, and Mr. Kyle Gonterwitz from KDOT for providing me with the KDOT horizontal curve inventory. I would like to thank the members of the Department of Civil Engineering at KSU, including faculty, staff, and students, especially my officemates in the offices 0076, 2151, and 2136 of Fiedler Hall.

I would like to express my heartfelt gratitude to my parents for their support, encouragement, and love. And finally, I warmly appreciate my wife for her love, patience, and support.

Dedication

To my beloved wife Zohreh and my precious daughter Roshana

Chapter 1. Introduction

This chapter emphasizes the importance of safety at horizontal curve sections, including a discussion of curve-related crash history and statistics in the United States and in Kansas. This chapter also defines the problem statement and dissertation objectives.

1.1. Importance of Horizontal Curves

Horizontal curves on roadways are located at points of horizontal alignment alterations or changes in road direction. Horizontal curves affect vehicle movement by producing centrifugal force and causing altered driving conditions. According to the second edition of the National Cooperative Highway Research Program (NCHRP) 600 report, design aspects of a curve, such as lane width, degree of curvature, radius of curve, design superelevation, and design consistency, impact driver workload [**1**]. Appropriate horizontal curve design lessens driver workload related to curve geometry. Spirals and appropriate superelevation changes are suggested treatments for horizontal curves in order to inform and prepare drivers for changes in roadway horizontal alignment. Driver visibility changes on curved segments of a road, causing most drivers to focus on tangent points; on road segments including smooth curves, however, drivers focus on the horizon [**1**].

According to the second edition of NCHRP 600, curve navigation within a safe speed is the most influential factor affecting crash rates on curves [**1**]. Although drivers select their driving speed based on expectations of a curve (affected by design consistency) and road signage (including advisory speed), driver expectations typically outweigh the effects of road signage. A combination of horizontal curve and sag vertical curve, in which the curve radius or apparent radius as viewed from the driver's perspective is greater than the actual curve radius, negatively influences driver selection of safe driving speed to negotiate a curve. In addition, observed speeds may exceed the advisory speed due to poor judgment of drivers. This adverse result is also expected for a combination of a horizontal curve and a crest vertical curve because the apparent radius is less than the actual radius and the curve appears to be sharper than it really is, causing drivers to often beneficially reduce their driving speed [**1**].

A two-level process model describes steering control as "an open-loop anticipatory component (far view)" and "a closed-loop compensatory component (near view)" [**1**]. Drivers use "far view" to predict curvature and steering angle and "near view" to correct a deviation from the desired path. However, all path-decision behaviors, such as curve-cutting, are not completely explained by this model. Although steering action should be dependent upon direct visual feedback, drivers often rely on their estimation of vehicle characteristics and their previous experience navigating curves [**2**]. A driver chooses curve entry speed based on personal perception of curvature influenced by geometric alignment and delineation features of the curve segment. Drivers often enter a curve at an improper speed due to curvature misperception, and in order to correct driving mistakes, drivers often take compensatory control ("near view") actions, especially in sharp curves [**2**,**3**]. Consequently, instead of following the ideal radius or the radius at the center of the lane, drivers often follow a trajectory with a larger radius.

Based on Fatality Analysis Reporting System (FARS) data from 2002, 38,309 fatal crashes caused 42,815 deaths on US highways, and nearly 25% of fatal crashes occurred on curve segments. More than 75% of those crashes were single vehicle, run-off-road (ROR), and approximately 10% of them were head-on crashes [**1**]. In recent years, the number of fatal crashes on US highways has decreased, but the percentage of curve-related fatalities has remained constant. In 2008, horizontal curve-related fatal crashes accounted for 28% of all nationwide fatal crashes; approximately 74% of those crashes were road departure crashes [**4**].

Recent data from the FARS database verifies a similar trend for fatal crashes and fatalities on horizontal curves in 2014, indicating that 25% of fatal crashes occurred on horizontal curves of which approximately 72% were ROR crashes [**5**]. Figures 1.1 and 1.2 show the percentage of horizontal curve fatal crashes of the US fatal crashes and percentage of ROR crashes of the horizontal curve fatal crashes, respectively for last seven years from 2008 to 2014. An appropriate design that includes curve segment consistency with other roadway segments (especially close segments before and after the curve), proper curve radius, a suitable spiral, superelevation, and lane width on the curve can improve curve safety. Other treatments can increase curve segment safety, but adequate treatment selection must be conducted according to expert judgment and/or empirical data [**1**]. Countermeasures for improved curve safety are classified as low cost, intermediate cost, and high-cost treatments [**6**].

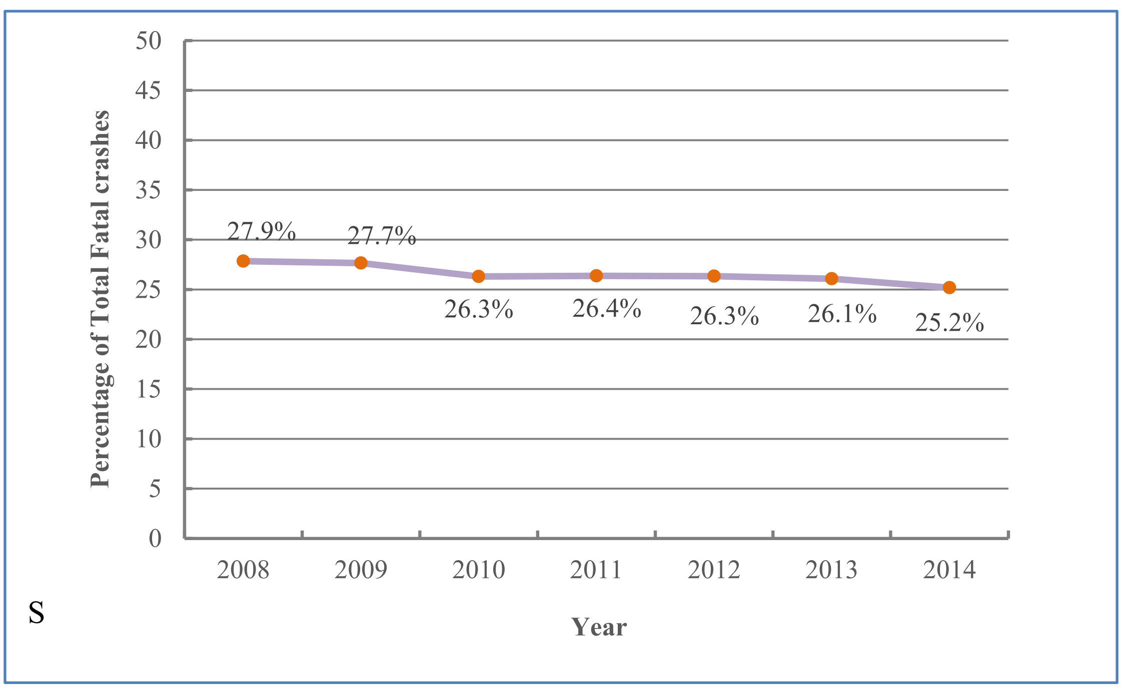


Figure 1.1 Percentage of the US fatal crashes occurred at horizontal curves (2008-2014)

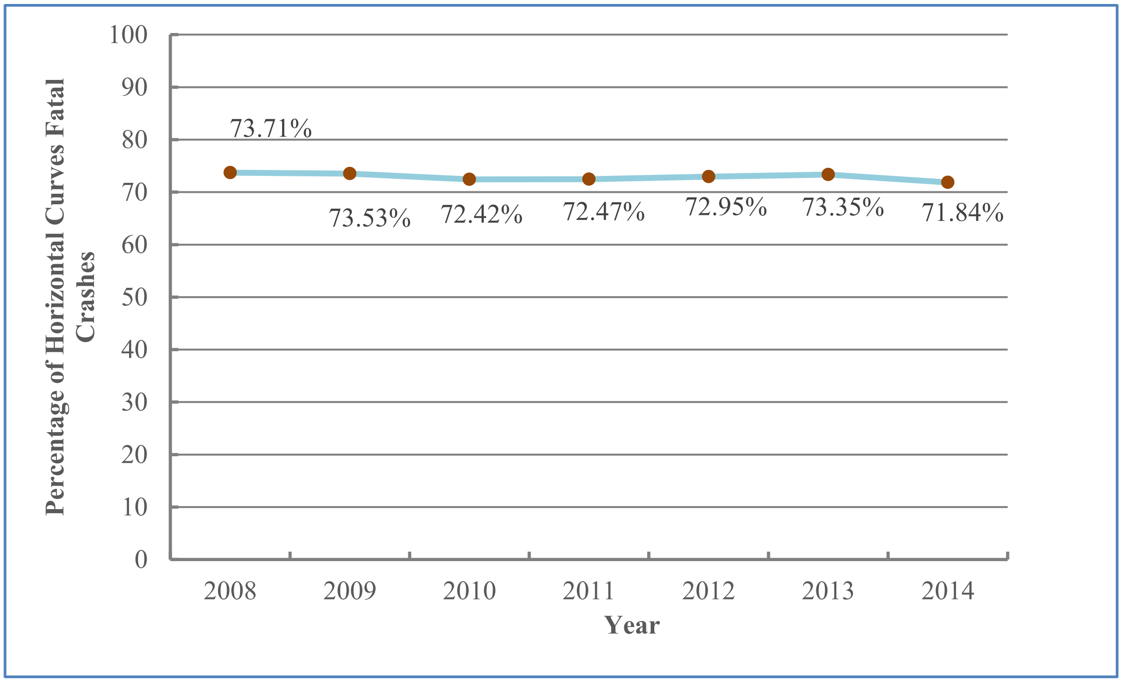


Figure 1.2 Percentage of ROR crashes from horizontal curve fatal crashes (2008-2014)

Seven years of crash data from the Kansas Department of Transportation (KDOT) showed a similar trend in increasing fatalities associated with horizontal curves: Approximately 29% of Property Damage Only (PDO) crashes, 24% of injury crashes, 19% of fatal crashes, and 27% of total crashes occurred at horizontal curves from 2008 to 2014. Figure 1.3–Figure 1.6 show the fatal, PDO, injury, and total number of crashes for horizontal curve sections, respectively.

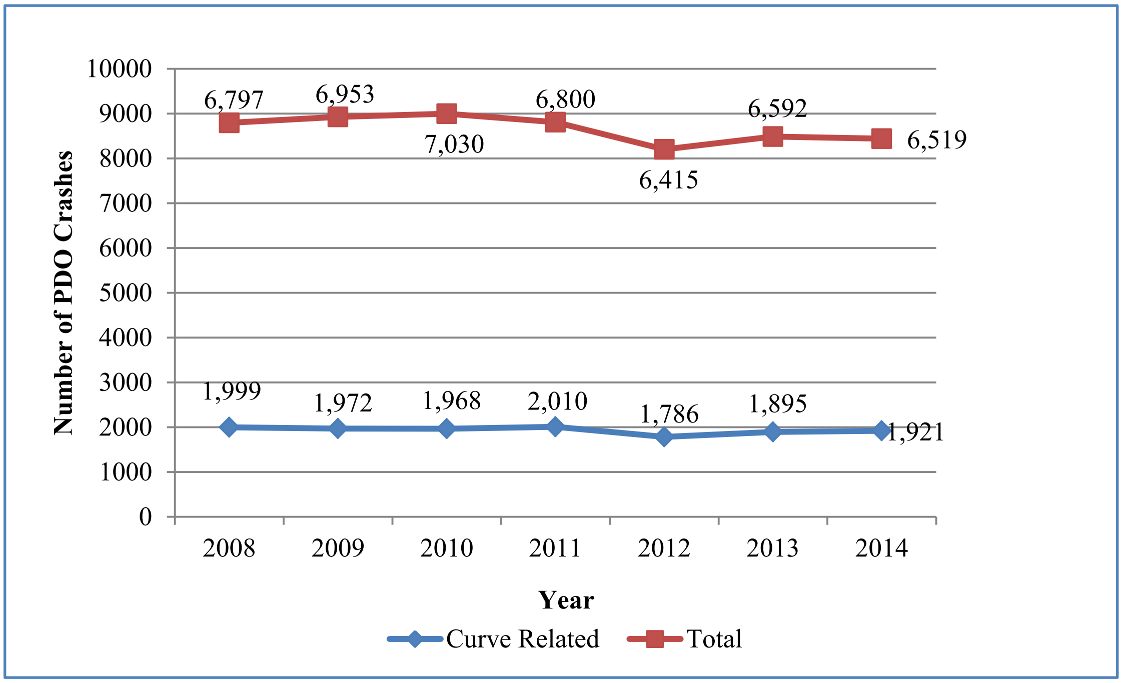


Figure 1.3 Curve-related PDO crashes in Kansas (2008-2014)

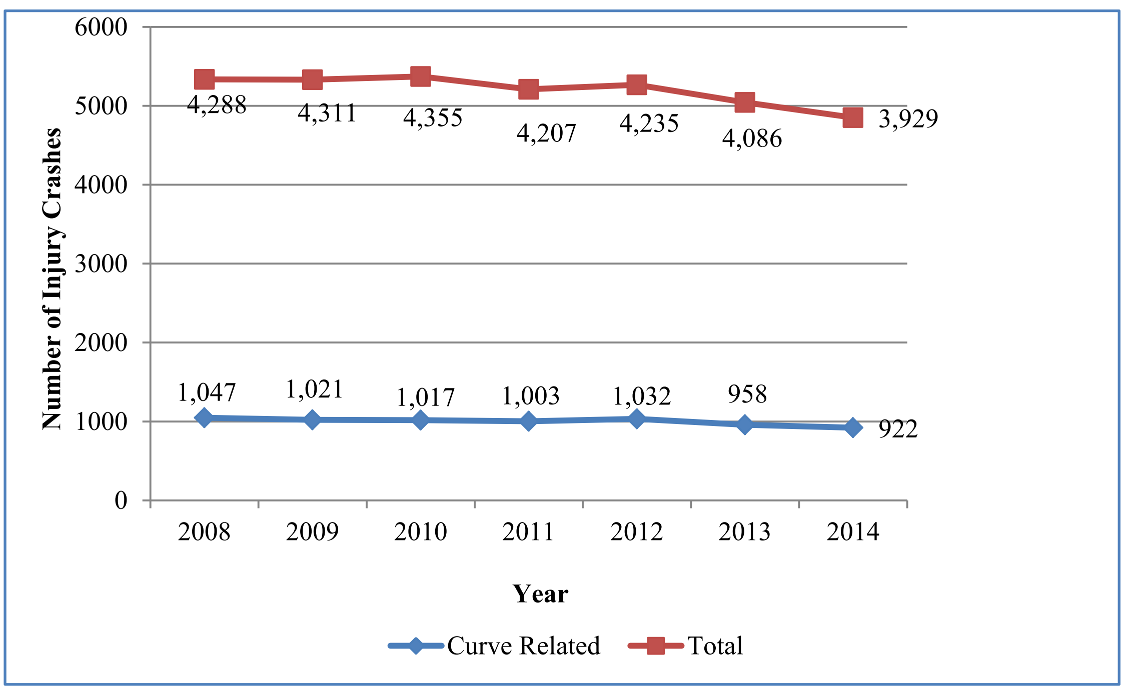


Figure 1.4 Curve-related injury crashes in Kansas (2008-2014)

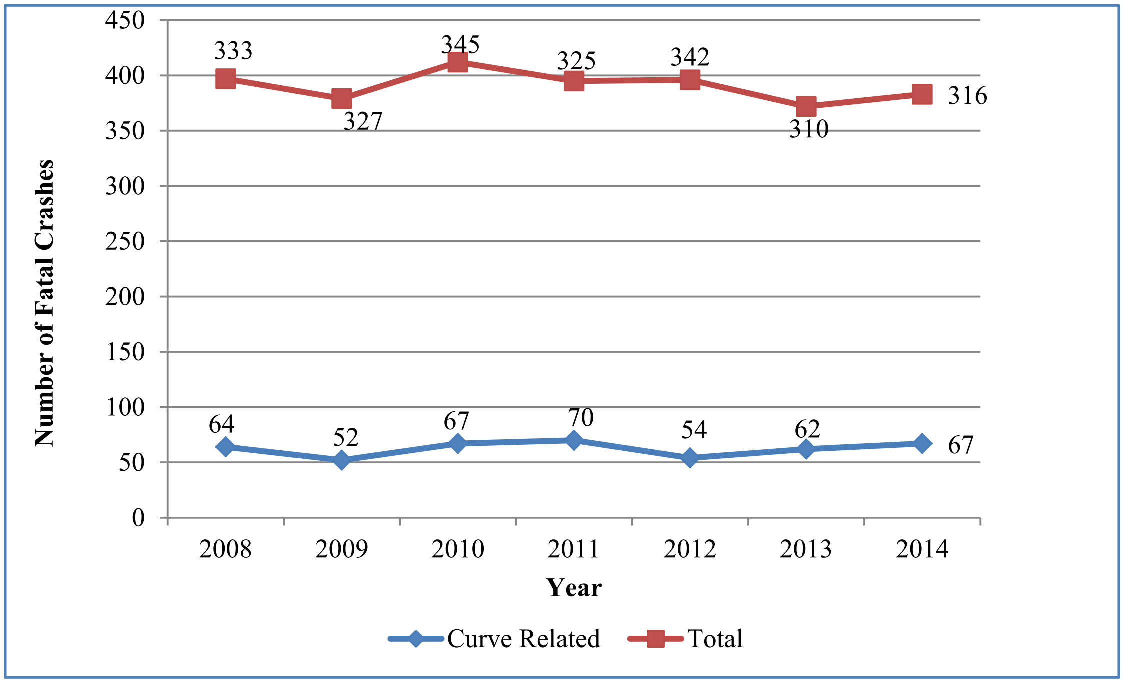


Figure 1.5 Curve-related fatal crashes in Kansas (2008-2014)

e

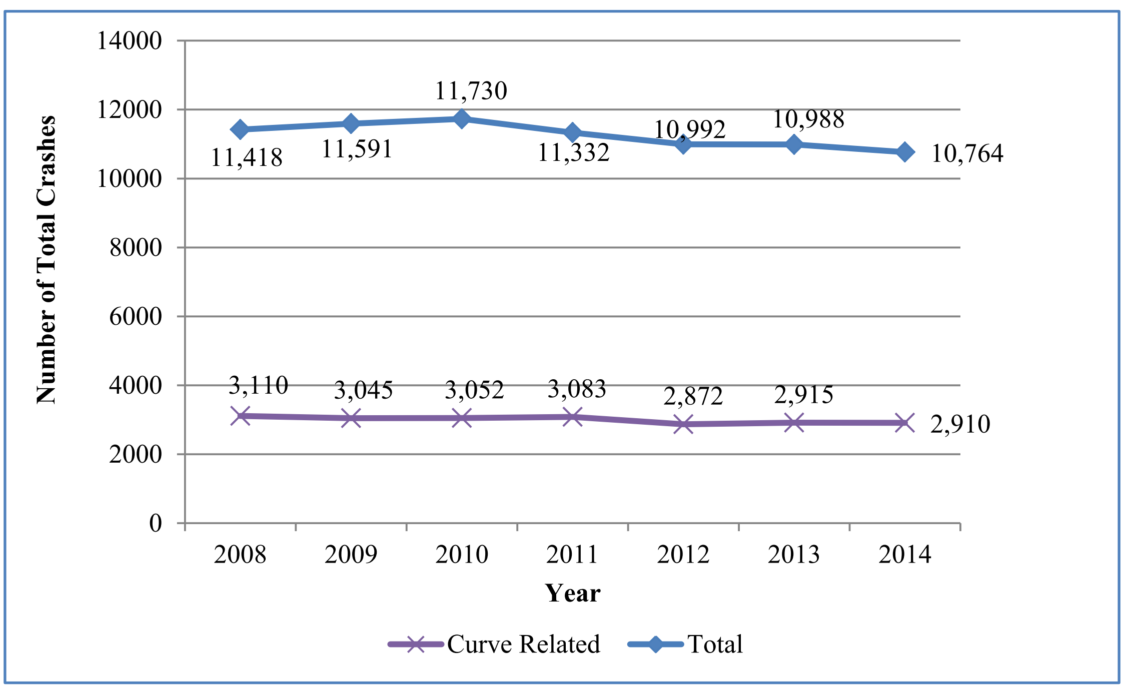


Figure 1.6 Curve-related crashes in Kansas (2008-2014)

Table 1.1 shows the number of PDO, injury, fatal, and total crashes for the roadway network and curve sections in Kansas from the years 2008 to 2014, as well as the percentages of crashes at horizontal curve sections.

Table 1.1 PDO, injury, fatal, and total number of crashes for the entire Kansas roadway network and curve sections from 2008 to 2014

Year	2008	2009	2010	2011	2012	2013	2014
Number of PDO Crashes							
Entire Network	6,797	6,953	7,030	6,800	6,415	6,592	6,519
Curve Sections	1,999	1,972	1,968	2,010	1,786	1,895	1,921
Percentage	29.4%	28.4%	28.0%	29.6%	27.8%	28.7%	29.5%
Number of Injury Crashes							
Entire Network	4,288	4,311	4,355	4,207	4,235	4,086	3,929
Curve Sections	1,047	1,021	1,017	1,003	1,032	958	922
Percentage	24.4%	23.7%	23.4%	23.8%	24.4%	23.4%	23.5%
Number of Fatal Crashes							
Entire Network	333	327	345	325	342	310	316
Curve Sections	64	52	67	70	54	62	67
Percentage	19.2%	15.9%	19.4%	21.5%	15.8%	20.0%	21.2%
Total Number of Crashes							
Entire Network	11,418	11,591	11,730	11,332	10,992	10,988	10,764
Curve Sections	3,110	3,045	3,052	3,083	2,872	2,915	2,910
Percentage	27.2%	26.3%	26.0%	27.2%	26.1%	26.5%	27.0%

1.2. Problem Statement

According to the FARS database, the number of fatalities associated with horizontal curves was more than 25% during last years from 2008 to 2014, with a majority of ROR horizontal curve-related crashes. Therefore, most adequate, cost-effective countermeasures must be implemented in order to decrease the number of vehicle crashes at horizontal curves. KDOT crash data from 2008 to 2014 showed that the proportion of fatal crashes at horizontal curve sections fluctuated between 16% and 21%, with an average of 19%, and ROR crashes constituted 63%–78% of fatal crashes, with an average of 71%. The proportion of curve-related fatal crashes was less than the national average but considerable, demonstrating maximum values of 20% and 21% for 2013 and 2014, respectively. This increasing trend reaffirms the importance of efficient safety improvements on horizontal curve sections by identifying factors that increase or decrease crash frequencies. Therefore, this research study utilized traffic and geometric data, such as annual average daily traffic (AADT), posted speed, advisory speed, heavy vehicle traffic, curve length, curve radius, pavement surface width, and shoulder type and width, as well as statistical methods and tools to devise recommended countermeasures for identified factors that affect crash occurrence at horizontal curves.

1.3. Objectives

The main objective of this dissertation is to identify geometric and traffic factors that affect crashes at horizontal curve sections. Two datasets were used for this objective: a dataset of 221 randomly selected horizontal curves on undivided two-lane two-way highways from approximately 750 horizontal curves from identified curves from KDOT horizontal curve crash database from 2004–2012 and KDOT curve inventory of curves with zero crashes from 2004-2012, and a new KDOT horizontal curve inventory that became available in 2015 and includes approximately 5,300 horizontal curves. Geometric data such as curve length, curve radius, tangent section length, pavement surface width, and paved and unpaved shoulder type and width were obtained from KDOT databases or measured. Also traffic data including traffic volume and heavy vehicle percentage were provided in KDOT databases or obtained directly from KDOT traffic maps. Suitable countermeasures are also identified in order to match geometric and traffic factors that positively or negatively affect crash occurrences at horizontal curve sections.

Chapter 2. Literature Review

This chapter reviews studies that have investigated basic and low-cost countermeasures recommended by the Federal Highway Administration (FHWA) in the Manual on Traffic Control Devices (MUTCD) and supplementary and innovative countermeasures and intermediate and high-cost countermeasures, including evaluation of countermeasure effectiveness. This chapter also describes previous studies that explored factors that positively or negatively affect crash occurrences or severe crash occurrences at horizontal curve sections or other roadway segments.

2.1. Horizontal Curve Countermeasures

Basic and low-cost treatments are discussed in MUTCD, and In addition to the MUTCD basic treatments, other supplementary and innovative countermeasures and also intermediate and high cost countermeasures have been introduced in other study sources. Studies have also been conducted to evaluate the effectiveness of horizontal curve countermeasures. Various utilized countermeasures and results of their evaluation are discussed in the following sections.

2.1.1. Low-Cost Countermeasures

McGee and Hanscom [7] studied the nine basic countermeasures introduced by MUTCD including centerline and edge line, horizontal alignment signs, advisory speed plaque, one-direction large arrow sign, combination horizontal alignment and advisory speed sign, curve speed sign, chevron alignment sign, and delineators. Following sections discuss basic low-cost countermeasures briefly.

2.1.1.1. Centerline and Edge Line

A centerline is the minimum treatment for a horizontal curve. Based on the MUTCD, use of a centerline for roadways with travel widths less than 16 ft. requires engineering judgment, but roadways with lane widths of 20 ft. or more with minimum average daily traffic (ADT) of 6000 vehicles per day (vpd) require edge lines [**8**,**9**]. When a curve does not provide adequate sight distance on two-lane roadways, a solid yellow line is necessary for one or both directions; edge lines are solid white lines along the right side of the road. The primary purpose of centerlines and edge lines is to provide a visual cue for drivers to follow the curve in order to impede encroachment into the opposite lane or edge line and prevent probable ROR incidents or crashes. NCHRP 600 states that pavement surface markings provide the strongest curvature guide [**1**].

Pavement markings utilize various materials, including common thermoplastic marking, which lasts longer than other materials, thereby increasing its cost-effectiveness [**7**]. Retro-reflective pavement materials (RPMs) and retro-reflective raised pavement materials (RRPMs) are also applicable for pavement markings depending on roadway conditions, but the FHWA prohibits the use of raised pavement markings for edge lines [14]. Studies have suggested that the combination of centerlines or edge lines with rumble strips improve curve safety [**10**,**11**].

Although conventional width for a centerline or edge line is 4–6 in., some states use widths of 8–12 in. [**12**]. Edge lines with widths of 8 in. were found to be appropriate alternatives for roadways with 12 ft. wide lanes, unpaved shoulders, and ADT of 2000–5000 vpd [**13**,**14**]. Hallmark et al. summarized the positive benefits, drivers' feedback, and improvements, including increased visibility (especially at night for older drivers), peripheral vision stimulation, lane keeping, comfort of drivers, and aesthetics [**15**].

Material used for centerline stripes significantly impacts cost, which varies by state. Lord et al. determined that average costs for type I, solid white edge lines are $0.30 per linear foot for 4-inch markings, $0.66 per linear foot for 6-inch markings, and $0.94 per linear foot for 8-inch markings [**12**]. The costs for type II, solid white edge lines for 4-inch, 6-inch, and 8-inch markings were estimated to be $0.12, $0.25, and $0.35 per linear foot, respectively [**12**].

2.1.1.2. Horizontal Alignment Signs and Advisory Speed Signs

A variety of signs presented in MUTCD are used in advance of a curve or a turn to warn drivers of an upcoming horizontal curve [**9**]. For a single curve, a turn sign (W1-1), a curve sign (W1-2), a hairpin curve sign (W1-11) for 135-degree change in alignment, or a 270-degree loop sign (W1-15) are applicable, as depicted in Figure 2.1 Advanced warning signs for horizontal curves. Similarly, two signs are used for two sequential curves or turns: reverse turns (W1-3) and reverse curves (W1-4). For several sequential curves, the winding road sign (W1-5) is appropriate [**9**].

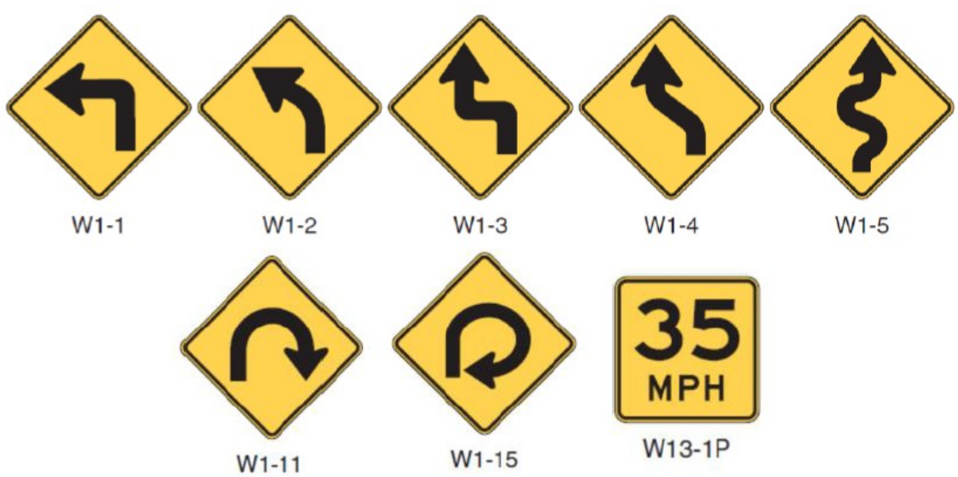


Figure 2.1 Advanced warning signs for horizontal curves
Source: [**9**]

The KDOT Handbook of Traffic Control Practices for Low Volume Rural Roads suggests placement of a turn sign when the advisory speed is equal to or less than 30 miles per hour (mph) and a curve sign for speeds greater than 30 mph [**16**]. An advisory speed plaque (W1-13) can also be added to curve-related signs. The advisory sign speed should be placed below the horizontal alignment sign [**1**]. McGee and Hanscom emphasized that advisory speed is not the legal speed limit but an advised speed to drivers [**7**]. The NCHRP 600 states that, although researchers agree about the use of warning signs in advance of a curve, disagreement exists regarding the use of symbols or text messages [**1**].

Placement of a highway curve sign is related to the curve's advisory speed and posted speed or 85th percentile speed of the tangent section of road prior to the curve [**9**]. McGee and Hanscom provided guidelines for warning sign placement in advance of highway curves in accordance with approach speed, as shown in Table 2.1. They asserted that all signs must be comprised of retro-reflective sheeting for increased night visibility and low-light conditions. The lower edge of the sign must be at least 5 ft. above the pavement surface, and the closest edge of the sign to the road must be at least 6 ft. from the outer edge of the shoulder [**7**].

Table 2.1 Guidelines for advance placement of curve warning signs
Source: [9]

Posted or 85th percentile speed (mph)	**Advance placement distance (ft.) for advisory speed of the curve (mph)**						
	10	**20**	**30**	**40**	**50**	**60**	**70**
20	n/a[1]	-	-	-	-	-	-
25	n/a[1]	n/a[1]	-	-	-	-	-
30	n/a[1]	n/a[1]	-	-	-	-	-
35	n/a[1]	n/a[1]	n/a[1]	-	-	-	-
40	100[2]	100[2]	n/a[1]	-	-	-	-
45	125	100[2]	100[2]	n/a[1]	-	-	-
50	200	175	125	100[2]	-	-	-
55	275	225	200	125	n/a[1]	-	-
60	350	325	275	200	100[2]	-	-
65	450	400	350	275	200	100[2]	-
70	525	500	450	375	275	150	-
75	625	600	550	475	375	250	100[2]

[1] No suggested distances are provided for these speeds since the placement location depends on site conditions and other signage in order to provide adequate advance warning for drivers.
[2] Minimum advance placement distance is listed as 100 ft. to provide adequate spacing between signs.

Amjadi studied the effectiveness of delineation improvements of curve-related signs, such as chevrons and one-arrow direction signs, on horizontal curves on two-lane rural roads in Connecticut [**17**]. Results indicated a reduction of 18% in all crashes, a 25% reduction in injury and fatal non-intersection crashes, and a 35% reduction in crashes during dark conditions. A few studies achieved varying results when relating the effectiveness of horizontal curve signs. Studies showed a reduction in crash occurrence due to advance curve warning signs; however, crash reduction percentages varied from 10% to 30% reduction in all crashes [**12**].

2.1.1.2.1. Larger Signs and Doublingup Signs

McGee and Hanscom affirmed that the MUTCD permits an increase in the size of horizontal curve signs when the volume, speed, or other conditions of a roadway require emphasis on sign readability. Table 2.2 lists sign sizes on various roadways [**9**].

Table 2.2 Sizes of warning signs in inches
Source: [**9**]

Description		Conventional Road		Expressway	Freeway	Minimum	Oversized
Shape	Sing Series	Single	Multi-				
Diamond	W1-1, 2, 3, 4,	30×30	36×36	36×36	36×36	–	48×48
	W1-1a, W1-	36×36	36×36	48×48	48×48	–	48×48
	W1-11, W1-	30×30	30×30	36×36	48×48	–	48×48
Rectangular	W1-6	48×24	48×24	60×30	60×30	–	60×30
	W1-8	18×24	18×24	30×36	36×48	–	24×30
	W13-1P	18×18	18×18	24×24	30×30	–	30×30

Doubling up sign is another method to help roadway users see warning signs and increase horizontal curve safety. A second similar sign is typically placed on the left side of the roadway.

2.1.1.2.2. High Retro-reflective Intensity and Fluorescent Yellow Sheeting

The MUTCD lists various types of retro-reflective materials used in roadway sign construction [**9**]. Increasing the retro-reflectivity of signs (measured in cd/lx/m2) has been shown to increase sign visibility [**1**]. For example, high intensity grade (Type III) and micro-prismatic sheeting (Type V) increase sign visibility compared to engineering grade (Type I) sheeting. A study in 2006 estimated a 2.4% cost increase for upgrading retro-reflective material from Type III to Type V [**7**]. Increased visibility allows timely driver responses to changes in roadway alignment, thereby increasing road safety [**7**].

2.1.1.3. Combination Horizontal Alignment Signs and Advisory Speed Signs

Combination horizontal alignment and advisory speed signs, referred to as supplementary signs, are used on curves with high numbers of crashes and significant differences between posted speeds and advisory speeds of curves. As shown in Figure 2.2, the W1-1a sign is a combination of a turn sign and advisory speed sign, and the W1-2a sign is a combination of a curve sign and an advisory speed sign [14]. These signs motivate drivers to reduce driving speed at the beginning point of a curve. However, W1-1a or W1-2a signs should not be used when the distance between the alignment sign and the beginning point of a curve is less than 200 ft. Campbell et al. suggested that increased visual demands for drivers on curves cause "conspicuous non-verbal information," such as chevrons, to be more effective than advisory speed signs [**1**].

Figure 2.2 Combination curve alignment sign and advisory speed sign
Source: [**9**]

2.1.1.4. Chevrons and One-Direction Arrow Signs

The W1-6 (one direction arrow) sign and the W1-8 (chevron) sign communicate an alteration in horizontal alignment. Both signs are also used with horizontal alignment signs, specifically when a sharp curve is present in the road. These signs must be placed on the outside of the curve at an approximate right angle with approaching traffic; one W1-6 sign is sufficient for each direction. If additional delineation is required, chevrons are appropriate alternatives, in which case at least two of them should always be in the driver's sight when navigating the curve [**12**]. Campbell et al. asserted that chevrons are the "strongest guidance cues for long-range guidance (anticipatory control)" [**1**]. Amjadi classified W1-6 and W1-8 signs as curve delineation signs [**17**].

Figure 2.3 Curve delineation signs
Source: [9]

The KDOT Highway Sign Manual recommends chevrons when the difference between posted speed and curve advisory speed is 15 mph or more [**18**]. According to FHWA, chevrons must be installed at least 4 ft. above the travel way [**9**]. McGee and Hanscom recommend that chevrons be posted 5 ft. above the surface of the roadway in rural areas, increasing to 7 ft. in urban areas [**7**]. Because MUTCD [**8**] did not specify spacing for chevrons, McGee and Hanscom use two states' spacing guidelines [**7**]. The latest edition of MUTCD [**9**], however, determines a guide for chevron spacing, as shown in Table 2.3.

Table 2.3 Typical spacing of chevron alignment signs on horizontal curves
Source: [**9**]

Advisory Speed	Curve Radius	Sign Spacing
15 mph or less	Less than 200 feet	40 feet
20 to 30 mph	200 to 400 feet	80 feet
35 to 45 mph	401 to 700 feet	120 feet
50 to 60 mph	701 to 1,250 feet	160 feet
More than 60 mph	More than 1,250 feet	200 feet

Note: The relationship between the curve radius and the advisory speed shown in this table should not be used to determine the advisory speed.

Iowa's traffic safety analysis manual recommends chevrons for curves with degree of curvature greater than or equal to six degrees and PMDs for curves with degree of curvature less than six degrees. The manual also recommends occasionally using chevrons for curves with degree of curvature less than 6 degrees if sight distance is reduced due to vegetation or a combination of horizontal and vertical curves, or whenever crash history indicates delineation improvement is needed [**19**]. McGee and Hanscom found that when the degree of curvature was more than seven degrees, chevrons significantly reduced centerline encroachment [**7**].

Amjadi reported a 20% reduction in crashes during dark conditions and a 20% reduction in departure crashes during dark conditions because of chevrons installed on horizontal curves in Washington state [**17**]; however, a variety of reduction percentages of crash occurrences were reported in other studies [**12**].

McGee and Hanscom approximated $500 for the installation of 10 chevrons [**7**], and Amjadi estimated $100 for the installation of each chevron on two-lane rural roads in

Washington state [**17**]. A recent study reported an average cost of $433 for the installation of one chevron in Texas [**12**].

2.1.1.5. Delineators

Delineators are retro-reflective devices mounted above the roadway surface, parallel to the roadway segment, in order to guide drivers through alignment changes at horizontal curves [**17**]. Although PMDs are not warning signs, they provide guidance information, as shown in Figure 2.4 [**8**]. Chevron alignment signs or PMDs are selected based on two criteria. The KDOT Highway Sign Manual recommends using delineators when the difference between posted speed and advisory speed is 10 mph or less [**18**]. McGee and Hanscom recommend PMDs for curves smaller than or equal to 7 degrees [**7**]. Hallmark et al. reported that PMDs and chevrons with retro-reflectorized posts more effectively help drivers recognize curve sharpness than standard PMDs and chevrons [**15**].

Figure 2.4 Post-mounted delineators
Source: [7**]**

The FHWA states that delineators must be posted approximately 4 ft. above the road surface, and they must be placed 2–8 ft. from the outer edge of the shoulder. The delineator color should be identical to the adjacent edge line. The FHWA also determines suitable spacing for delineators in accordance with the radius of the curve, as shown in Table 2.4 [**9**].

Table 2.4 Recommended spacing for delineators at horizontal curves
Source: [**9**]

Radius (R) of Curve	Approximate Spacing (S) on Curve
50 feet	20 feet
115 feet	25 feet
180 feet	35 feet
250 feet	40 feet
300 feet	50 feet
400 feet	55 feet
500 feet	65 feet
600 feet	70 feet
700 feet	75 feet
800 feet	80 feet
900 feet	85 feet
1,000 feet	90 feet

Notes:
1. Spacing for specific radii may be interpolated from table.
2. The minimum spacing should be 20 feet.
3. The spacing on curves should not exceed 300 feet.
4. In advance of or beyond a curve, and proceeding away from the end of the curve, the spacing of the first delineator is 2S, the second 3S, and the third 6S, but not to exceed 300 feet.
5. S refers to the delineator spacing for specific radii computed from the formula $S=3\sqrt{R-50}$.
6. The distances for S shown in the table above were rounded to the nearest 5 feet.

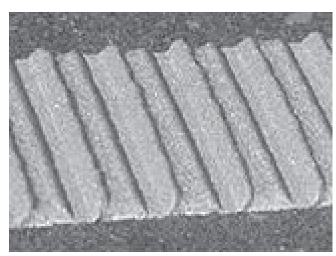
A study of installed PMDs on horizontal curves revealed a 25% reduction in all types of crashes at horizontal curves [**20**]; however, unique reduction percentage is not anticipated for crashes on horizontal curves [**12**]. Other low-cost treatments are available, but they are not classified as basic treatments [**14**]. These low-cost countermeasures are discussed in the following sections.

2.1.1.6. Profile Thermoplastic Markings and Raised Pavement Markings

Profile thermoplastic markings and RPMs produce a rumble effect and auditory warning to drivers in order to potentially increase safety. Profile thermoplastic markings and RRPMs increase visibility as compared to RPMs which have non-retro-reflective features. Although these two marking types typically are not effective for snowy regions since snow plows often damage them, design changes may make them compatible with snow plowing. According to FHWA, RRPMs are suitable for smooth curves (less than 3.5 degrees) and relatively high traffic volume (more than 5000 vpd) since they may create an unrealistic feeling of safety for drivers on sharp curves and consequently cause drivers to more quickly negotiate a sharp curve due to increased curve visibility [**7**].

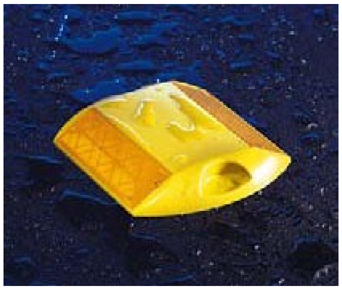
Campbell et al. suggested that a combination of RRPMs with centerlines and edge lines increases curve safety [**1**]. For very sharp curves (more than 12 degrees), the report recommends the use of RRPM pairs on the outside edges of the centerline, placed 800.5 ft. (244 meters) in advance of every curve with spacing intervals of 131.2 ft. (40 meters) for sharp curves and 262.5 feet (80 meters) for smooth curves [**1**]. Another study recommends the utilization of snow-plowable RPMs for curves that cause ROR crashes [**19**].

Raised profile thermoplastic



Inverted profile thermoplastic

Figure 2.5 Profile thermoplastic markings
Source: [7]



Standard RPM- for   Snow-plowable

Figure 2.6 Raised pavement markings
Source: [7]

Lord et al. estimated an average cost of $0.93/ft^2 for profile thermoplastic markings in Texas. The study also compared costs of pavement markings for various materials, as shown in Table 2.5 [**12**].

Table 2.5 Estimated cost of pavement markings in Texas
Source: [**27**]

Pavement Marking Material	Cost ($/mile)
Paint	1,056
Thermoplastic	1,584
Tape	3,960
Buttons	2,233

2.1.1.7. Reflective Barrier Delineation

Reflective sheeting panels installed on concrete barriers or guardrails improve visibility of horizontal curves, particularly at night [**7**]. Reflectors should be mounted on guardrails perpendicular to approaching headlights, as shown in Figure 2.7. The color of reflective sheeting or mounted reflectors must match the adjacent edge line color. Panels and reflectors typically have 18–36-inch spacing [**9**].

Figure 2.7 Reflective panels and mounted reflectors
Source: [7]

McGee and Hanscom estimated that each reflector costs $3 to install and almost $2.33 per linear foot of 4-inch wide reflective panels [**7**]. Lord et al. estimated that each reflector costs $3.42 and each linear foot of reflective panels with 4-inch, 6-inch, and 8-inch widths costs $0.30, $0.66, and $0.94, respectively [**12**].

As described in McGee and Hanscom, reflective sheeting can be implemented on obstructions near the edge of roadways and in the clear zone, but it is potentially hazardous for ROR crashes. Six-inch wide reflective tape is typically applied to the object, as illustrated in Figure 2.8. When the distance between the object and the shoulder is 8 ft. or less, the marker should be placed at least 4 ft. above the pavement surface, otherwise the 4-foot height should be measured from the ground. Yellow reflective materials are commonly used unless aesthetic consideration requires brown materials [**9**].

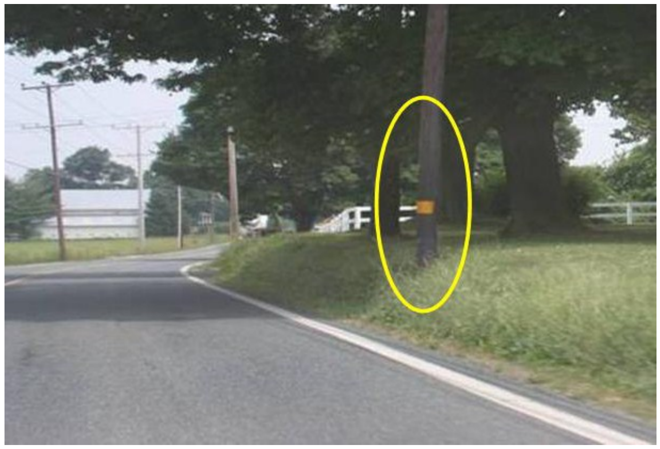


Figure 2.8 Reflective tape on object close to the road
Source: [7]

2.1.1.8. Speed Limit Advisory Marking

Speed limit markings can be used as a supplemental warning in advance of a curve with a common horizontal alignment sign. Use of an arrow sign and “SLOW” text on the pavement has the similar effect [**7**,**12**]. Campbell et al. suggested that an arrow sign and text should be placed 230 ft. (70 meters) in advance of the curve in high hazard areas or at sharp curves [**1**]. However, McGee and Hanscom stated that sign placement distance depends on approach speed and curve design speed [**7**]. Several studies [**7**,**12**,**1**] indicated expected speed reduction when this kind of treatment is used. Lord et al. reported that speed limit advisory pavement markings cost an average of $116 [**12**].

Figure 2.9 Speed limit advisory pavement marking
Source: [7]

Figure 2.10 PennDOT curve advance marking
Source: [7]

One study recommended placing on-pavement curve signs in locations where advisory sign placements are recommended [**7**]. In a study conducted by Charlton in 2007 in New Zealand, on-pavement curve signs were determined to be more effective than chevrons at low speeds (28 mph (45 km/hr)) [**21**].

2.1.1.9. Optical Speed Bars

Thermoplastic painted stripes, or transverse stripes, are implemented perpendicular to roadway alignment in advance of the curve and on the curve [**1**], as shown in Figure 2.11. The primary objective of this treatment is to give drivers the illusion of increased speed by decreasing the distance between stripes, thus causing drivers to slow their driving speed. Based on McGee and Hanscom, these white stripes are typically 18 in. long and 12 in. wide; their effectiveness is attributed to decreasing spaces between the stripes in relation to the curve [**7**]. However, no clear conclusion confirms the effectiveness of this treatment since various studies have indicated occasional reduction in speed, no reduction in speed, or a slight increase in speed [**1**]. Campbell et al. suggested that a combination of rumble strips and transverse stripes would be more effective [2]. McGee and Hanscom estimated a $2000 cost for the implementation of optical bars on two directions of a curve in Virginia in 2006 [**7**].

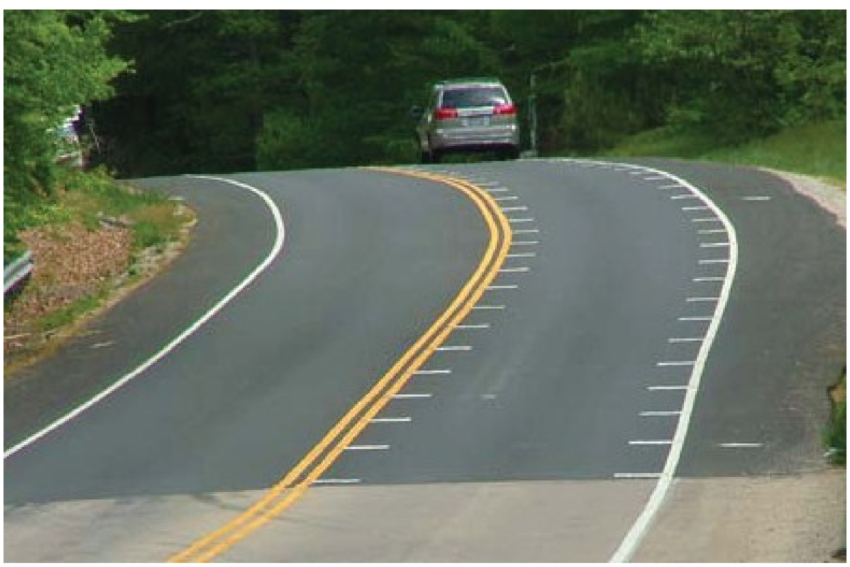


Figure 2.11 Optical speed bars
Source: [**7**]

Hallmark et al. categorized optical speed bars as "transverse pavement markings," "on-pavement chevrons," and "herringbone." On-pavement chevrons are shown in Figure 2.12.

According to their study, on-pavement chevrons are typically applied on freeway ramps, in advance of curves, and as entrance treatment to rural communities. Anticipated results of on-pavement chevron applications include a decrease in mean speed or 85th percentile speed at horizontal curves [**15**].

Figure 2.12 Application of on-pavement chevrons
Source: [**15**]

2.1.1.10. Rumble Strips

Rumble strips, which are grooved or raised elements installed on the pavement surface, can be milled, rolled, formed, and raised rumble strips. Bogenreif reported that one-third of Iowa safety funding was allocated for the installation of shoulder and edge line rumble strips [**10**]. Rumble strips placed near or on curve sections of roads cause noise and vibration to alert drivers of their lateral placement on the curve. Rumble strips on a horizontal curve can be utilized as centerline rumble strips (CLRS) to prevent drivers from encroaching into the opposite lane, edge line or shoulder rumble strips to warn drivers of ROR crashes, and transverse rumble strips to encourage drivers to reduce their driving speed. KDOT practice only allows transverse rumble strips in advance of a stop condition.

2.1.1.11. Centerline Rumble Strips

In general, CLRSs are milled rumble strips that are installed at or near the centerline. Factors such as operating condition, cross section characteristics, and potential road users affect optimum dimensions for milled centerline rumble strips [**12**]. McGee and Hanscom designated common dimensions of CLRSs to be 12–16-inch length (vertical to centerline), 7-inch width, and 0.5-inch depth (or height), as shown in Figure 2.13 [**7**]. Russell and Rys recommended milled rumble strips with 12–16-inch length (perpendicular to the centerline), 7-inch width (along the centerline), 0.5-inch depth, with a 12-inch continuous apart or alternating pairs 12-inches apart with the pairs 24-inches apart [**22**]. In a recent study by Karkle, Rys, and Russell, typical CLRSs are milled strips with 16-inch length, 7-inch width, 0.5-inch depth, and 12-inch continuous spacing [**23**,**24**].

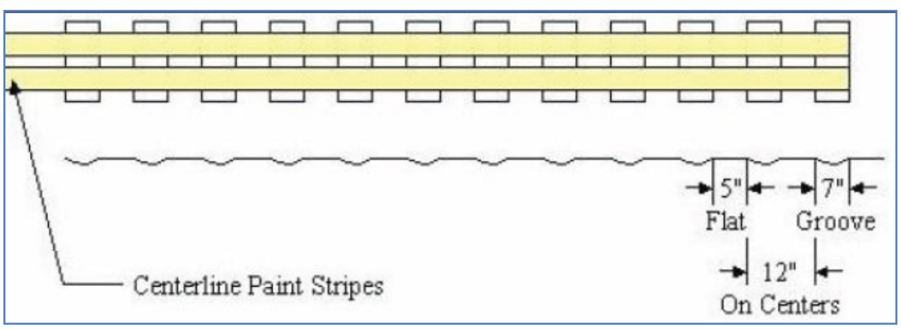


Figure 2.13 CLRS pattern
Source: [7]

McGee and Hanscom estimated the cost of CLRSs to be approximately $0.40 per linear foot [**7**], and Lord et al. reported CLRSs to cost approximately $8.63 per linear foot in Texas [**12**].

Figure 2.14 Example of CLRS
Source: [**68**]

2.1.1.12. Edge Line or Shoulder Rumble Strips/Stripes

Shoulder rumble strips (SRSs) require sufficient width of roadway shoulder; an SRS can be applied on the edge line (commonly called rumble stripes) or shoulder, depending on shoulder width, as shown in Figure 2.15. Lord et al. suggested 4–12 in. of offset distance from the edge line when an 8-foot clear shoulder width is available after installation [**12**]. McGee and Hanscom recommend a 7.1-inch longitudinal width and 15.8-inch transverse width with a repeating pattern of approximately 5.1 inches [**7**]. Implementation cost has been estimated at $8.63 per linear foot in Texas [**7**]. Hallmark et al. introduced edge line rumble stripes, an innovative combination of rumble strips and edge line markings, in order to improve visibility during wet conditions [**15**]. McDonald also recommended rumble stripes or milled-in wet-weather visibility pavement markings on curves with an ROR crash history [**19**].

Figure 2.15 Example of SRS
Source: [7]

2.1.1.13. Roadway Rumble Strips

Transverse rumble strips, which are grooved or raised stripes across the road pavement, remind drivers to reduce speed or increase caution when negotiating a curve section [**9**]. The maximum height or depth should not exceed 0.5 in., and a warning “RUMBLE STRIPS AHEAD” sign is recommended in advance of this treatment to warn motorists, bicyclists, and motorcyclists [**7**]. However, maintenance concerns should be considered, especially when raised rumble strips are implemented in snowy regions. To prevent motorists from using the opposite lane when they encounter transverse rumble strips, a discontinuous pattern design, such as gaps in the bars or grooves across the pavement, is recommended.

Figure 2.16 Example of roadway rumble strip
Source: [7]

2.1.2. Intermediate and High-Cost Treatments

2.1.2.1. Flashing Beacons

Flashing beacons are used as a supplementary treatment when conventional safety improvement countermeasures have not remedied a safety problem [**7**]. A typical circular yellow section from a standard traffic signal is used for flashing beacons, thereby attracting driver attention to existing signs. Flashing beacons are commonly installed above the signs and at least 12 in. from the nearest edge of the signs. Lord et al. estimated the cost of a traditional unit to be $2300 and a solar-powered unit to be $4900 in Texas [**12**].

Figure 2.17 Use of flashing beacons with a sign
Source: [**28**]

Similar to flashing beacons, LEDs are used in traffic signs such chevrons, as shown in Figure 2.18. This treatment advantageously directs driver attention to the sign [**6**].

Figure 2.18 Chevron enhanced with LEDs
Source: [**6**]

As shown in Figure 2.19, LEDs can also be used in pavement markers to enhance delineation, particularly during low visibility conditions, or in RPMs or markers (solar or hardwired) on the pavement since they are snowplow-safe and bike-safe [**9**]. The cost of a photocell-powered LED RPM is approximately $50 including material and installation [**25**]. The cost of 20 embedded LED markers connected by wire and installed with 20-foot spacing on a curve section is estimated to be $48,000 for a 110 volts of alternative current (VAC) power system and $55,000 for a solar-powered source [**26**].

Figure 2.19 Use of in-pavement LED markers
Source: [**69**]

2.1.2.2. Dynamic Curve Warning System

The DCWS detects an approaching vehicle, measures its speed, and activates a warning variable message and/or beacons whenever the vehicle navigates the curve faster than a safe speed. The DCWS consists of loop detectors or radar. Various designs have been suggested for this system, but the simplest design includes a constant message sign enhanced with flashing beacons on the corner of the sign. When a vehicle with excessive speed approaches the curve, the detection system activates the beacons to warn the driver. Because of their high cost, however, DCWSs are suggested for curves in which common treatments have not improved curve safety. A wide diversity of design options influences DCWS implementation cost. McGee and Hanscom reported a $61,000 system installation cost in California [**7**], while in Texas a system was estimated to cost $18,000 [**12**]. However, Caltrans reported a 44% reduction of total crashes in the first year after DCWS installation [**7**].

Figure 2.20 Dynamic curve warning system in Texas
Source: [**12**]

Figure 2.21 DCWS
Source: [**71**]

Traffic & Parking Control Company, Inc. (TAPCO) recently introduced a Sequential Dynamic Curve Warning System (SDCWS) consisting of chevrons enhanced with LEDs and solar-powered sources with a radar detector and activator or controller. When the radar detector senses a speeding vehicle, it triggers the controller, and the controller wirelessly activates LEDs on the chevrons to flash sequentially or synchronously at a desired rate. According to TAPCO, a wireless, vehicle-activated, 5-sign system with solar-powered 30-inch by 36-inch signs would cost $14000. The SDCWS was studied in Colorado, Missouri, Texas, Washington, and Wisconsin. Study results have not yet been published. Use of chevrons enhanced with LEDs on a horizontal curve section is shown in Figure 2.22.

Figure 2.22 Chevrons enhanced with LEDs
Source: [**70**]

2.1.2.3. Paved Shoulders and Widening Shoulders

Paved shoulders provide extra space for drivers to escape if another vehicle is approaching head-on, thus increasing the safety of roadway curve segments. Removing material from the old shoulder, recompacting the shoulder, and replacing it with new appropriate asphalt constitutes construction activity. Various aggregates and colors can be used to distinguish the shoulder and travel lane for drivers. Widening paved or stabilized shoulders provides additional space for drivers and increases curve safety.

Figure 2.23 Inside shoulder widening
Source: [7]

McGee and Hanscom estimated a \$1 /yd^2 cost for seal-coating a gravel shoulder [**7**]. Another study reported that asphalt costs approximately \$3.80 per gallon and aggregates cost \$72 per cubic yard [**27**].

2.1.2.4. Shoulder Drop-Off Mitigation (Safety Edge)

Horizontal curves often contain drop-offs because unstabilized pavement edges erode, resulting in a height difference and causing drop-offs [**7**,**12**]. McGee and Hanscom recommended a 45-degree angle fillet of asphalt on each side of the roadway [**7**], and Lord et al. recommended a formed slope with a 30-degree angle [**12**]. Fillet or slope-formed shoulders enable drivers who leave the travel lane to return their vehicles to the roadway with less hazard or risk. FHWA states that the treatment is cost-effective because it requires less than 1% of the asphalt required for a new surfacing project [**28**].

Figure 2.24 Safety edges
Source: [7]

New safety edge guidelines at KDOT, effective since January 18, 2013, recommends 1.7H: 1V for asphalt and concrete shoulders with 0–3 ft. of width, 1/4H: 1V for asphalt shoulders wider than 3 ft., and vertical edges for concrete shoulders wider than 3 ft.

2.1.2.5. Installation or Improvement of Lighting

Installation of new lighting or improvement of old lighting can increase the visibility of a curve section of roadway. Lighting is particularly beneficial at nighttime and in adverse weather conditions. However, the installation of new lighting is expensive and should be considered only if economically justifiable. Lord et al. reported the average cost of lighting to be $2,336 for each unit installed lighting in Texas [**12**].

Figure 2.25 Lighting to improve visibility on a curve
Source: [**12**]

2.1.2.6. Skid Resistant Pavement Surface

The use of aggregates such as calcined bauxite can increase the friction coefficient of a pavement surface. During resurfacing this treatment can be applied on horizontal curves to increase curve safety, especially when surfaces are wet. This treatment can be obtained by overlaying existing asphalt with appropriate asphalt or applying grooves on the pavement surface. Moreover, the pavement surface must be well drained to meet the purpose. In order to obtain a proper asphalt overlay, voids should be present on the surface to help drainage and improve skid resistance. Voids can be formed using aggregate that lacks particular particle gradations. Longitudinal or transversal grooving provides drainage and increases friction [**12**].

Investigating the safety effectiveness of a skid resistive overlay in New York showed a 50% reduction in wet condition crashes and a 20% reduction in total crashes [**7**]. Results indicated that grooved pavement performed better in wet weather conditions than the other. McGee and Hanscom reported a 72% reduction in wet condition crashes on some studied horizontal curves in California but only a 7% reduction in dry pavement crashes in those curves. They also reported the cost of a 2-mile overlay in California to be $200,000 in the year 1996 [**7**].

Figure 2.26 Application of skid resistant material on a curve
Source: [7]

2.2. Crash Prediction Models

Many studies have developed prediction models for various crash types on rural highways. Zegeer et al. studied 10,900 curves in the state of Washington in order to determine effective features of horizontal curves regarding crash occurrence on two-lane rural highways, identify curve geometric improvements related to curve safety, and develop a crash reduction factor for horizontal curves. They utilized a linear regression model to analyze geometric and traffic data. Their findings showed that sharp curves with narrow width, deficient superelevation, and no spiral transitions increased the number of crashes. In similar conditions, the probability of crash occurrence increased with increased traffic volume and curve length [**29**]. Miaou and Lu evaluated the effect of geometric characteristics of horizontal curves in Utah in relation to truck crash rates using data from the Highway Safety Information System (HSIS). They developed a Poisson regression model to identify factors that affect crash occurrence and estimate how improvement of those factors would affect crash rate at horizontal curve sections. They found that truck involvement in crashes increased when AADT, horizontal curvature, curve length, vertical grade, and length of vertical grade increased and paved inside shoulder width and truck percentage in similar AADTs decreased [**30**]. In an extended study, Dissanayake and Amarasingha investigated truck crashes in rural networks in Kansas. They utilized negative binomial (NB) regression when the results of Poisson regression method showed overdispersed data, and they compared large truck crashes on limited access highways. Their findings revealed that horizontal curvature, vertical grade, lane width, and shoulder width affect truck crash occurrence on limited access highways [**31**].

Poch and Mannering studied crash frequencies at urban and suburban intersections in Bellevue, Washington. They developed statistical models for annual total number of crashes, rear-end crashes, angle crashes, and approach-turn crashes. Geometric and traffic factors considered in the model development included approach volumes, number of approach lanes, speed limits, highway grades, signal control characteristics, the presence of horizontal curves, and sight distance restrictions. An NB regression method was used, and prediction models were developed for four crash groups [**32**].

Chang and Chen compared use of the Classification and Regression Tree (CART) method and the NB regression method. They developed a prediction model for a national freeway in Taiwan, concluding that CART could be used as an alternative for other prediction models. According to their study, variables that affect crash occurrence include traffic volume, precipitation in terms of number of days and amount, and roadway alignment in terms of grade and degree of curvature [**33**].

Hallmark et al. conducted a before and after study to investigate the effect of dynamic curve signs on speed reduction and crash occurrence at horizontal curve sections using a Bayesian or generalized linear regression model. Results were published in early 2015 [**34**,**35**]. Although a wide variety of variables were considered in the model, they used AADT, curve length, curve radius, sign type, posted speed, curve advisory speed, and the difference between posted speed and advisory speed. They also used a full Bayes model to develop a crash modification factor (CMF) using crash frequency as the response variable [**35**].

Schneider et al. investigated the impact of rural two-lane horizontal curves on non-intersection truck-related crashes. They used an NB regression model with full Bayes method to develop a prediction model that considered 15,390 truck crashes from 2002 to 2006. Their study showed that traffic volume and horizontal curvature influence truck crash occurrence at horizontal curve sections [**36**]. In another study, Schneider et al. (2010) explored the impact of roadway geometry at horizontal curve sections on single-vehicle crash frequency by implementing a Bayesian technique to improve the frequency estimation. According to the findings of their study, curve length and radius, shoulder width, and annual daily traffic (ADT) significantly influenced single-vehicle motorcycle crash frequency at horizontal curve sections [**37**].

Easa and You studied 3600 miles of rural two-lane highways in the state of Washington. The Highway Safety Information System (HSIS) was used to obtain required geometric and traffic data, and crash data from 2002 to 2005 were extracted from police departments reports in Washington. They selected five groups of combinations of horizontal curves and vertical alignment conditions, including crest vertical curve, sag vertical curve, multiple vertical curve, grade with absolute value of less than 5%, and grade with absolute value equal to or greater than 5%. Poisson, NB, ZIP, and ZINB regression methods were used to develop the most appropriate prediction model. Among the five studied groups of roadway sections, the ZIP method was found to be ideal for one group and the ZINB model was best for the remaining four groups. Variables influencing crash occurrence at the studied sections were AADT, length of curve, degree of curvature, roadway width, access density, and grade [**38**].

Khan et al. used a comprehensive database of horizontal curves in Wisconsin, provided by the Wisconsin Department of Transportation (WisDOT). They studied crashes from 2006 to 2010, and they utilized quasi Poisson regression and NB regression methods to model crashes at horizontal curve sections. In order to consider crash severity, they developed a prediction model for all crashes and for severe crashes, including K, A, and B crashes from the KABCO scale [**39**]. Bogenreif explained each term of KABCO that determines a level of crash severity [**10**]:

K: Fatal
A: Incapacitating Injury
B: Non Incapacitating Injury
C: Possible Injury
O: Property Damage Only.

Khan et al. found that approximately similar variables were included in the prediction models for all crashes and severe crashes; however, the coefficient of the variables varied for each model. Variables that affected crash occurrence at horizontal curve sections consisted of curve radius, curve length, AADT, posted speed, average international roughness index (IRI) for the pavement surface, shoulder width, and tangent section length [**39**]. Table 2.66 and Table 2.77

summarize explanatory and dependent variables used in transportation safety studies to develop prediction models for curve-related crashes or other types of crashes.

Table 2.6 Variables used in horizontal curve crashes prediction models

Source	Explanatory Variables	Dependent Variables
Khan et al. (2013)	Curve radius (R), curve length (L), log AADT, posted speed, left and right shoulder width and type, average IRI, pavement surface age and type, upstream tangent (0–600 ft., 601–1200 ft., and 1201–2600 ft.), truck percentage, travel way width, difference between posted speed and advisory speed, presence of curve-related signs	Number of crashes at horizontal curves, number of KAB crashes At horizontal curve sections
Hallmark et al. (2007)	Number of lanes, lane width, shoulder width and type, speed limit, pavement type and condition, presence and location of street lighting, grade, horizontal curve radius, degree of curvature, superelevation, sight distance, presence and characteristics of spirals, density of curves upstream in terms of number of curves per mile, length of connection tangent section, location and type of signage before and within the curve (e.g., location of speed reduction zones, chevrons, etc.), speed, volume, any other feature that may influence driver expectation and curve approach speed	Speed change, crash frequency
Schneider et al. (2009)	Shoulder width, horizontal curve radius, curve length, passenger vehicle ADT, truck ADT, degree of curvature	Truck crashes
Schneider et al. (2010)	Lane width, overall surface width, posted speed limit, additional land use categories (e.g., population density), ADT, segment length, curve radius, shoulder width	Single-vehicle motorcycle crashes
Hallmark et al. (2015)	AADT, section length, season, sign type, posted speed limit, curve advisory speed, differences between speed limit and advisory speed, radius	Total crashes for both directions, total crashes for the direction of the sign, total single vehicle crashes, single vehicle crashes in the direction of the sign
Zegeer et al. (1993)	ADT, curve length, degree of curvature, total surface width, presence of spiral transition, superelevation, roadside hazard rate, roadside recovery distance	Total number of crashes

Table 2.7 Variables used in non-horizontal curve crashes prediction models

Source	Explanatory Variables	Dependent Variables
Chang and Chen (2005)	Highway geometric design information including number of lanes, horizontal curvature, VG, and shoulder width, traffic information including ADT of various vehicle types, peak hour factors, traffic distribution over lanes, weather information from the annual report of climatological data for cities and towns along National Freeway 1 (including pressure, temperature, humidity, precipitation, wind speed, and cloudiness)	Crash frequency from 2001 to 2002
Dissanayake and Amarasingha (2012)	Section length, speed limit, median width, functional class, AADT, AADT of heavy vehicles, right rumble strips, inside rumble strips, right shoulder width, inside shoulder width, horizontal curve, vertical grade, number of lanes	Truck crashes from 2005 to 2010
Miaou and Lum (1992)	Section length, truck miles or truck exposure, dummy intercept, dummy variables from the years 1986 to 1989, AADT per lane, horizontal curvature (HC), length of original horizontal curve (LHC), vertical grade, length of vertical grade, deviation of paved inside shoulder width, percent trucks, interaction between HC and LHC, interaction between VG and LVG	Number of trucks involved in accidents
Hosseinpour et al. (2014)	Posted speed, shoulder width, horizontal curvature, terrain type, heavy vehicle traffic, land use, side friction factor, presence of median, access points	Number of head-on crashes, crash severity

Appendix A provides additional details regarding variables in prediction models in previous studies

Chapter 3. Methodology and Data

In order to identify geometric and traffic factors that affect horizontal curve crashes, crash data from the years 2004 to 2012 were obtained from KDOT and crashes were verified to ensure that they occurred at horizontal curve sections. In addition, data of curves with zero number of crashes were obtained from KDOT's horizontal curve inventory developed by the KDOT Traffic Safety Unit. Because the KDOT database did not initially contain sufficient required geometric and traffic data, 221 horizontal curves on rural two-lane highways were selected. Geometric data were measured, including curve length, curve radius, and tangent sections lengths in advance of horizontal curves and paved and unpaved shoulder widths. Traffic data such as AADT, heavy vehicle percentage, posted speed, advisory speed, and presence of rumble strips were obtained from KDOT traffic maps or images from Google Maps. Recently, however, the Geometric and Accident Data (GAD) Unit of the Traffic Safety section of KDOT provided updated horizontal curve inventory with complete data variables of horizontal curves in Kansas and additional geometric and traffic data that were used in data analysis for this study. Data analyses of 221 horizontal curves on undivided rural two-lane highways and 5,334 horizontal curves on the entire state transportation network in KDOT inventory are represented in the following sections of this chapter and in Chapter 5.

3.1. Horizontal Curves for Preliminary Analysis

Required geometric and traffic data were measured or obtained from KDOT databases or KDOT traffic maps for 221 randomly selected horizontal curves on undivided two-lane two-way highways. The collected data variables included AADT from 2004 to 2012, heavy vehicle percentage, radius of curve, degree of curvature, curve length, tangent sections length, posted speed, advisory speed of the curve, crashes (fatal, injury, and PDO), presence of centerline or edge line rumble strips at curves, paved or unpaved shoulder, width of paved shoulder (≤3ft. and >3ft.), and width of unpaved shoulder (≤3ft., >3ft. and ≤7ft., and >7). The 221 randomly selected curves undivided two-lane two-way highways were initially identified on Google Maps using longitude and latitude of at least one point on the horizontal curve, and then an AutoCAD 2013 software package was used to measure related horizontal curve characteristics. Figures 3.1 and 3.2 illustrate the measuring of curve radius, curve length, and curve tangent sections length using AutoCAD 2013 on extracted photos from Google Maps. The complete inventory of horizontal curves with approximately 5300 curves and related geometric and traffic data were later obtained from the KDOT database in April 2015.

Figure 3.1 Measuring radius and length of curve using Google Maps and AutoCAD

Figure 3.2 Measuring lengths of tangent sections for a curve using Google Maps and AutoCAD

3.2. Poisson Regression Model

Crashes occurring at transportation sections and segments have nonnegative integer values, thereby requiring utilization of count data models such as Poisson and NB regression methods; other methods, such as applying standard least squares regression method for continuous data should be avoided [**40**]. Because crashes are discrete response variables with integer values as possible outcomes and random variables, generalized linear models (GLMs) can be implemented to analyze crashes. Crashes as Poisson variables take nonnegative integer values. Poisson distribution has one mean parameter, μ, that is also a variance of response variable [**41**]. The general form of mean and variance of Poisson distribution is defined according to Equation (1).

$$E(Y) = Var\,(Y) = \mu, \qquad \sigma(Y) = \sqrt{\mu} \tag{1}$$

The mean of Poisson distribution always has a positive value, and the log of the mean is commonly modeled. The general form of the Poisson regression model is

$$\log(\mu) = \beta_0 + \beta_i X_i \quad \textbf{(2)}$$

which can be written as

$$\mu = e^{(\beta_0 + \beta_i X_i)} \quad \textbf{(3)}$$

where

μ: the mean of Y
β_0: the constant value
β_i: coefficient for ith independent variable
X_i: he ith independent variable

The probability distribution or mass function can be expressed as the following equation [**42**]:

$$f(y;\mu) = \frac{e^{-\mu_i} . {\mu_i}^{y_i}}{y_i!} \quad \textbf{(4)}$$

where y is the actual number of response variables, which represents crash frequency.

3.2.1. Overdispersion

When the response variance is greater than the mean, overdispersion occurs in the Poisson model due to data clustering, unaccounted temporal correlation, or model misspecification; however, according to Lord and Miranda-Moreno, overdispersion for crashes occurs primarily due to a Bernoulli trial with unequal probability of events, also known as Poisson trials [**43**]. Positive correlation between responses and excess variation between response probabilities or counts cause overdispersion, causing standard errors of estimates to be deflated or underestimated. In other words, some variables can be identified as significant variables although they are not significant variables. In order to control data overdispersion, the value of the Pearson Chi Square (χ^2) or Pearson statistic divided by the degree of freedom (df) is greater than 1.0. A correction is required when the value of the Pearson statistic divided by the df is greater than 1.25 and 1.05 for moderately sized and large models, respectively [**42**]. Although overdispersion does not occur for normally distributed data that follow ordinary regression models, it should be considered in data analysis with Poisson distribution pattern [**41**].

3.3. Negative Binomial Regression

The NB regression model contains a nonnegative parameter, called dispersion parameter, that enables the model to consider data overdispersion [**41**]. Equation (5) shows the difference between Poisson and NB regression assumptions:

$$E(Y) = \mu, \qquad Var(Y) = \mu + D\mu^2 \tag{5}$$

The basic structure of the NB regression model is described as

$$y_i = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \tag{6}$$

where

y_i: response variable for i^{th} observation (e.g., number of crashes for i^{th} section)
β_0: constant
$\beta_1, \ldots, \beta_n$: estimated parameters in vector form
$x_1, \ldots, x_n$: explanatory variables for i^{th} observation

3.3.1. Maximum Likelihood

The maximum likelihood parameter (L) can be utilized to estimate parameters of the NB regression model [**37**]. The maximum likelihood function can be obtained from Equation (7).

$$L(\lambda_i) = \prod_{i=1}^{N} \frac{\Gamma(y_i + \theta)}{y_i \Gamma(\theta)} \left[\frac{\theta}{\theta + \lambda_i}\right]^{\theta} \left[\frac{\lambda_i}{\theta + \lambda_i}\right]^{\alpha} \quad \textbf{(7)}$$

where

$L(\lambda_i)$: maximum likelihood estimator of λ_i
N: total number of observation groups (e.g., number of horizontal curve sections)
$\Gamma(\ )$: Gamma function
θ: dispersion parameter
α: inverse dispersion parameter (1/θ)

3.4. Zero-inflated Model

Some studies recommend use of ZI models in which many zeros are observed and in this case, two states are assumed for the studied transportation segments. Some of the segments are safe or virtually safe with zero crashes, known as true-zero, and the rest of the segments are unsafe or are crash-prone locations with zero or non-zero crashes [**44**]. However, the assumption that excess zeros are due to virtual safe segments is questionable; the preponderance zeros could be the result of inappropriate selection of time period or/and section length [**27**]. Lord et al. recommended avoiding the use if ZI models when the study time period is not long enough; however, they did not clarify the appropriate time scale to apply ZI models [**27**]. Washington et al. stated that a zero state for roadway segments or sections can be true when some sections have insufficiently small probabilities of crash occurrence, thereby prompting application of ZI models that consider a dual-state system (normal-count and zero-count states) [**40**]. ZI models assume that events $Y = (y_1, y_2, \dots y_n)$ are independent [**40**]. The ZIP model is defined according to Equation (9) [**40**].

$$y_i = 0 \; with \; probability \; p_i + (1 + p_i)EXP(-\lambda_i)$$
$$y_i = y \; with \; probability \; \frac{(1 - p_i)EXP(-\lambda_i)\lambda_i^y}{y!} \tag{8}$$

where y is the number of events, such as crashes per study period. The ZINB model also is defined as

$$y_i = 0 \; with \; probability \; p_i + (1 + p_i)\left[\frac{\frac{1}{\alpha}}{\left(\frac{1}{\alpha}\right) + \lambda_i}\right]^{1/\alpha}$$
$$y_i = y \; with \; probability \; (1 - p_i)\left[\frac{\Gamma\left(\left(\frac{1}{\alpha}\right) + y\right)u_i^{1/\alpha}(1 - u_i)^y}{\Gamma\left(\frac{1}{\alpha}\right)y!}\right], y = 1, 2, 3 \ldots \tag{9}$$

where $u_i = \left(\frac{1}{\alpha}\right)\left[\left(\frac{1}{\alpha}\right) + \lambda_i\right]$.

3.5. Goodness of Fit

Various criterion have been recommended to verify goodness of fit for Poisson, NB, and ZI models. A majority of studies have utilized Akaike information criterion (AIC) [**39**,**45**]. According to Khan et al., "The AIC is a measure of the relative goodness of fit of a statistical model which loosely describes the tradeoff between the accuracy and complexity of the model" [**39**]. The general form of the AIC equation is

$$AIC = 2k - 2lin(L) \tag{10}$$

where

k: the number of parameters in the statistical model
L: the maximum value of the likelihood function for the estimated model

Although no specific threshold exists for the value of AIC, small values are desired for studied models: A model with the smallest AIC value is the best model.

3.6. Data Collection

3.6.1. Initial Analysis

As mentioned, only 221 horizontal curves on undivided two-lane two-way highways were initially randomly selected for this study due to lack of horizontal curve geometric and traffic data. Table 3.1 provides descriptive statistics of geometric and traffic variables related to the selected curves data set.

Table 3.1 Geometric and traffic variable descriptions for initial data collection

Variable	Minimum	Maximum	Mean	Standard Deviation
Curve length (ft.)	93	5,621	1,144	970.7
Curve radius (ft.)	115	18,833	1,967	2,161.6
AADT (vpd)	205	10,498	2,214	1,679.6
Heavy vehicle percentage	1	24	10	4.4
Tangent length (ft.)	25	39,449	2,888	4,962.5
Posted speed (mph)	35	65	57	7.4
Difference between posted speed and advisory speed	0	40	7	9.1

3.6.1.1. Crashes at 221 Horizontal Curves

Four hundred sixty-nine crashes occurred at 221 randomly selected horizontal curves on two-lane highways during the years 2004–2012. Out of those crashes, 283, 170, and 14 crashes were PDO, injury, and fatal crashes, respectively. Figure 3.3 shows total crash frequencies for the 221 horizontal curves on undivided two-lane two-way highways during the study period.

According to the figure, 120 horizontal curves, approximately 54.3% of the total studied curves, experienced 0 or 1 crash during the nine-year study period. Only four curves had more than 1 crash per year, approximately 1.8% of the studied horizontal curves. Less than 10% of the studied horizontal curves, amounting to 20 curves, had more than one crash per every other year.

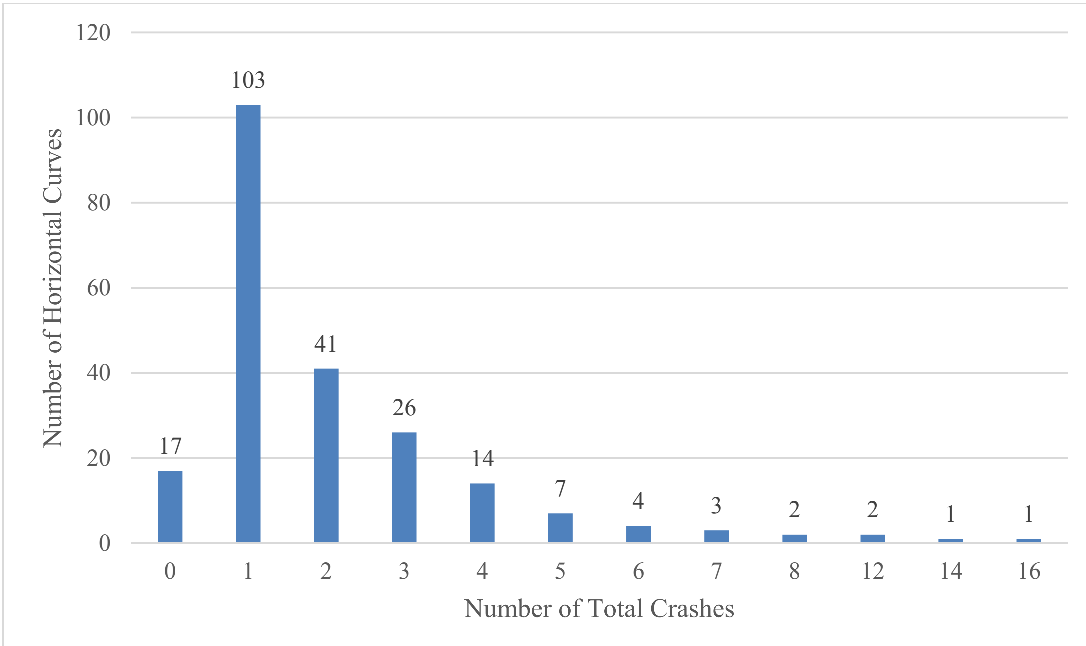


Figure 3.3 Total crash frequencies for 221 horizontal curves on two-lane highways

3.6.2. New KDOT Horizontal Curve Inventory

A complete horizontal curve inventory and related existing geometric and traffic data and crashes at each horizontal curve section in 2014 were obtained from KDOT. The inventory consisted of approximately 5,300 horizontal curves and recorded data that included the following:

1. Curve length (ft.), assuming all curves were circular curves
2. Curve radius (ft.), assuming all curves were circular
3. Speed limit (mph) in the new KDOT database for horizontal curves

4. Advisory speed (mph),in the new KDOT database for horizontal curves
5. IRI (in./mi. or m/km) based on the simulated response of a generic motor vehicle to roughness in a single wheel path of the road surface
6. AADT (vpd) in the new KDOT database for horizontal curves
7. Surface width (ft.) for the roadway surface at horizontal curve sections measured by surface width between edge lines
8. Shoulder width (ft.), paved and unpaved, in the KDOT database

Horizontal curve data were in different Microsoft Excel files, and the Kansas Crashes and Analysis Reporting System (KCARS) database was in Microsoft Access format. Therefore, an R software package was used to combine the required data of horizontal curves in one file and create a dataset for data analysis. The written codes are provided in Appendix B. Table 3.2 presents descriptive statistics of horizontal curves in KDOT's curve inventory.

Table 3.2 Geometric and traffic variable descriptions for KDOT horizontal curve inventory

Variable	Minimum	Maximum	Mean	Standard Deviation
Curve Radius (ft.)	43	70,317	5,918	5,984.1
Curve length (ft.)	31	9,513	1,328	926.0
AADT (vpd)	79	76,000	14,789	14,379.9
Heavy vehicle percentage (%)	1.63	72.65	14.97	9.68
Surface Pavement width (ft.)	16	80	25.5	9.9
Outside Shoulder Width (ft.)	0	12	8.3	3.5
Inside Shoulder Width (ft.)	0	12	2.3	3.4
IRI (in./mile)	23	245	69.2	26.28
Posted Speed (mph)	20	75	63	10.4
Advisory Speed (mph)	15	75	62.5	12.2

3.6.2.1. Crashes at Horizontal Curves in KDOT Inventory

The KCARS database was used to find crashes at horizontal curve sections of KDOT inventory; crashes from the beginning of the year 2010 to the end of 2014 were identified for each of the 5,334 horizontal curves on the entire state transportation network. Crashes at each horizontal curve section were determined using KDOT's horizontal curve inventory and common parameters such as roadway code, start and end mileposts of the horizontal curves, and lane class. From January 1, 2010, to December 31, 2014, 13,874 crashes occurred at 5,334 horizontal curve sections on the entire state transportation network. A majority of those crashes were PDO crashes. Figure 3.4 shows crashes at horizontal curve sections based on severity.

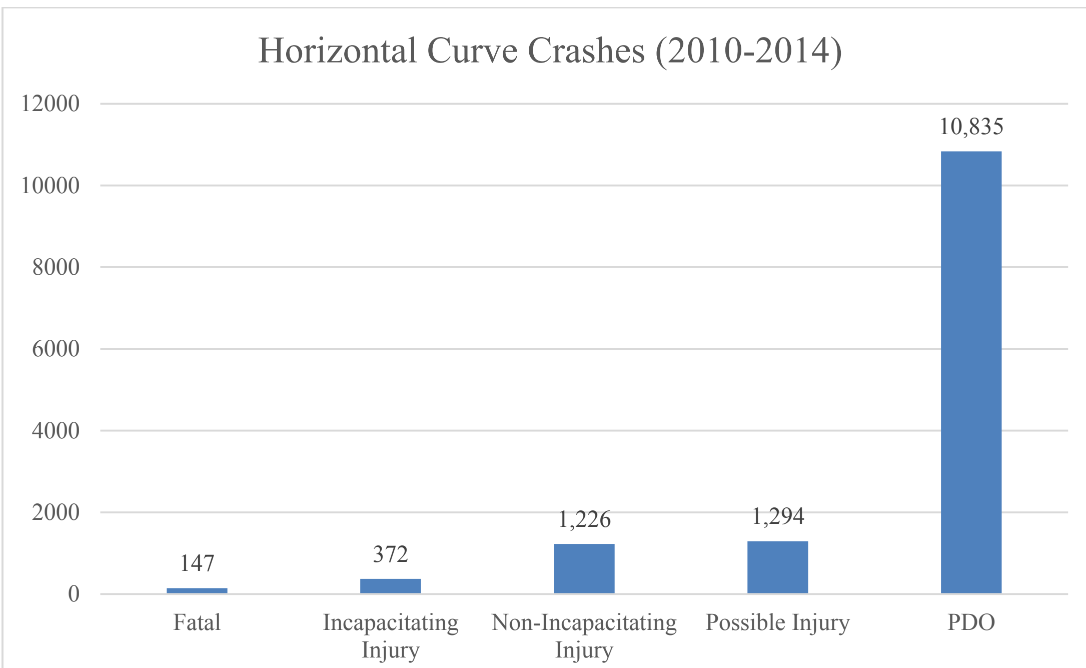


Figure 3.4 Crashes at horizontal curve sections based on severity

This research study considered severe crashes in addition to the total number of crashes. Using the five levels of crash severity (KABCO scale), some studies [**46**,**47**,**48**,**49**] consider fatal and incapacitating injuries (K and A levels) to be severe crashes, while the other studies [**39**,**50**] consider fatal and injury crashes, whether incapacitating or non-incapacitating, (K, A, and B levels) to be severe crashes. Therefore, this research study considered severe crashes with both K and A levels and K, A, and B levels since the number of fatal and incapacitating injury crashes may be very low on the studied horizontal curve sections. Crashes were considered based on crash type. Six crash types were identified for crashes at horizontal curve sections: single-vehicle crashes, multi-vehicle crashes, crashes with animals, crashes with pedal-cyclists, crashes with pedestrians, and crashes with railway trains. Table 3.3 and Figure 3.5 show the number of total crashes, severe crashes (K,A and K,A,B levels) for the six crash categories at 5,334 horizontal curve sections on the entire state transportation network.

Table 3.3 Number of total, KA, and KAB crashes at horizontal curves based on crash type

Crash Types	Single-vehicle	Multi-vehicle	With animal	With pedal-cycle	With pedestrian	With railway train	Total
Severe Crashes (KA)	289	216	14	0	0	0	**519**
Severe Crashes (KAB)	1,008	658	77	1	1	0	**1,745**
Total Crashes	5,095	4,930	3,813	11	24	1	**13,874**

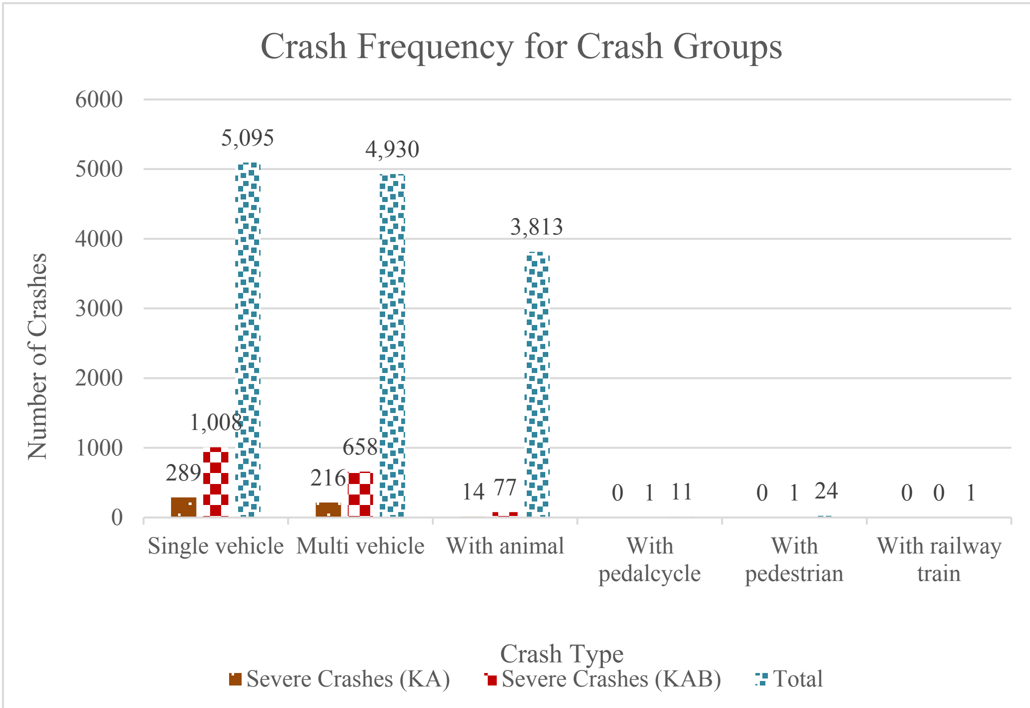


Figure 3.5 Crashes based on crash type and crash severity

3.6.2.2. Selecting Data for Data Analysis and Data Validation

Data from the KDOT horizontal curve inventory included geometric characteristics of 5,334 horizontal curve sections on the entire state transportation network. The data were randomly divided into two parts in order to determine if the developed model predicted crashes with acceptable accuracy. The first part of the data, consisting of 80% (4,267) of horizontal curve sections, was used for data analysis, and the second part, including 1,067 horizontal curve sections, was used for data validation. Table 3.4 compares statistical characteristics of the horizontal curves randomly selected for data analysis and data validation. According to the table, datasets for data analysis and validation reasonably followed statistical characteristics of the entire database.

Table 3.4 Comparison of statistical characteristics of randomly selected curves for data analysis and data validation

	Statistical Characteristics	Total Crashes	Severe Crashes	AADT*1000	Heavy Vehicle (%)	Curve Length (ft.)	Degree of Curvature
All of the Curves	**Min.**	0	0	0.079	1.63	30.54	0.075
	Max.	58	11	76	72.65	9,512.60	122.13
	Ave.	0.95	0.19	14.79	14.97	1,327.68	2.63
Selected Curves for Data Analysis	**Min.**	0	0	0.079	1.63	30.54	0.075
	Max.	27	11	76	72.65	9,512.60	122.13
	Ave.	0.96	0.19	14.80	14.87	1,328.64	2.71
Selected Curves for Data Validation	**Min.**	0	0	0.116	1.95	47.67	0.111
	Max.	58	6	76	64.66	7,377.74	88.13
	Ave.	0.92	0.18	14.71	15.21	1,323.82	2.35

During the study period from 2010 to 2014, 5,090 single-vehicle crashes occurred at 5,334 Kansas horizontal curves, and among the entire single-vehicle crashes 4,108 and 982 of them occurred at selected 4,267 and 1,067 horizontal curves for data analysis and data validation, respectively. From the 288 severe crashes (K and A levels) on all Kansas horizontal curves, 232 and 56 severe crashes occurred at selected curves for data analysis and data validation, respectively. In addition, 819 and 187 severe crashes (K, A, and B levels) occurred at the randomly selected horizontal curves for data analysis and validation, respectively, altogether constituting 1,006 severe crashes at 5,334 horizontal curve sections on the entire state transportation network. Figure 3.6–Figure 3.14 show the number of horizontal curves for each crash group and all horizontal curve sections, as well as horizontal curves selected for data analysis and data validation.

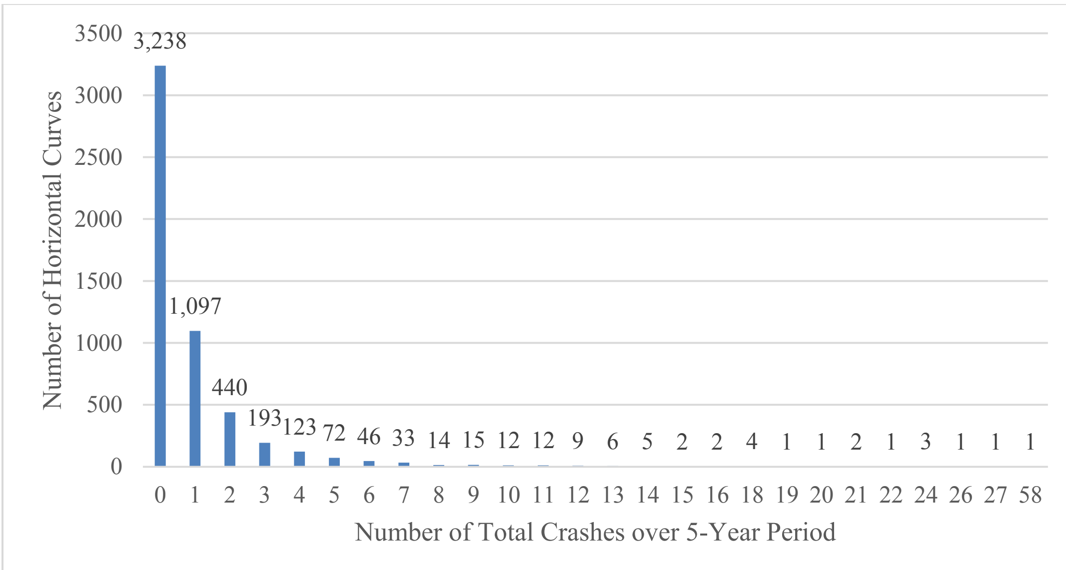


Figure 3.6 Total crash frequencies for 5,334 horizontal curves in Kansas

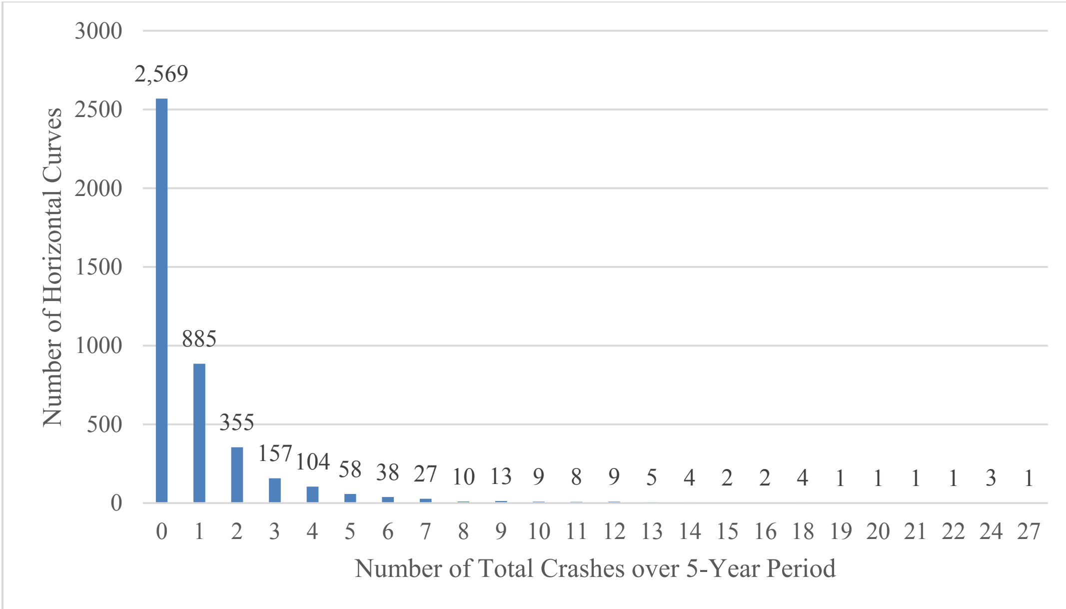


Figure 3.7 Total crash frequencies for 4,267 horizontal curves selected for data analysis

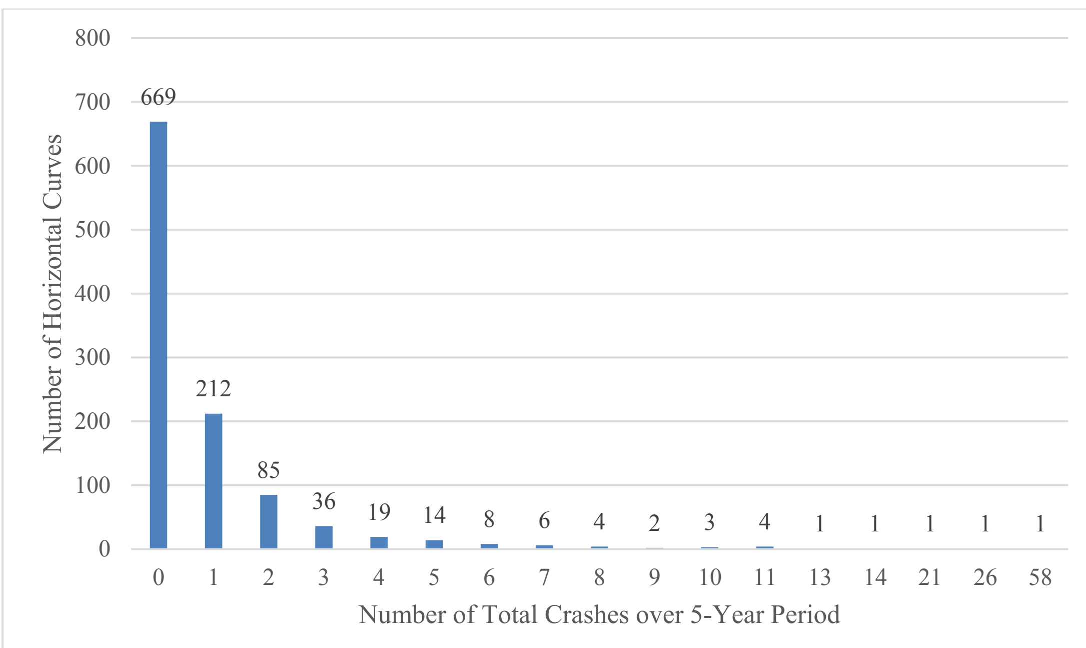


Figure 3.8 Total crash frequencies for 1,067 horizontal curves selected for data validation

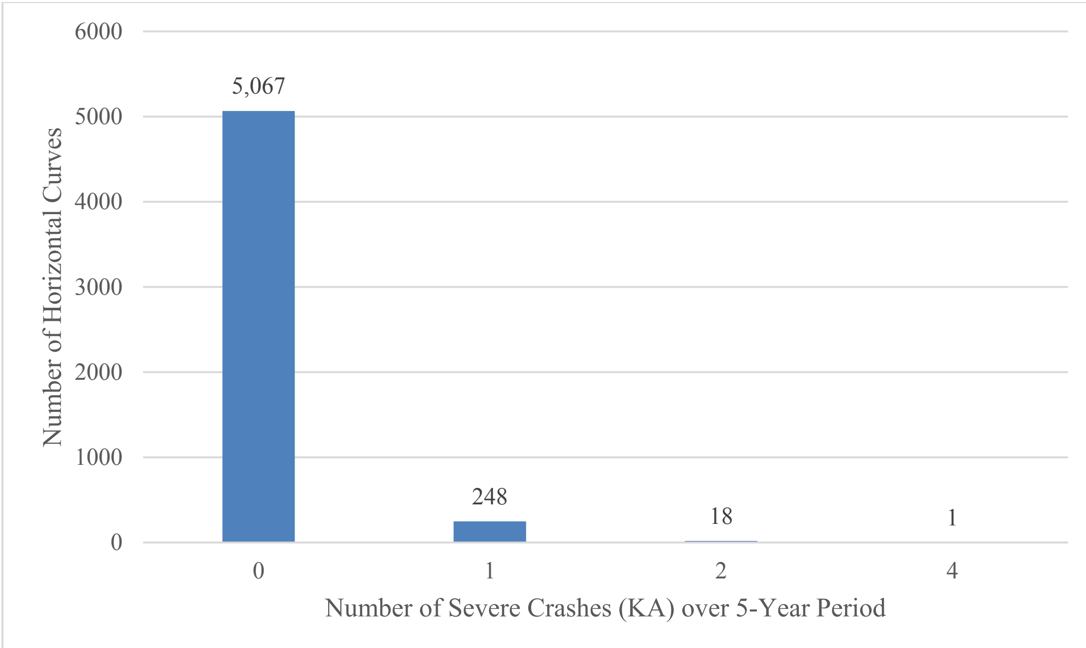


Figure 3.9 Severe (KA) crash frequencies for 5,334 horizontal curves in Kansas

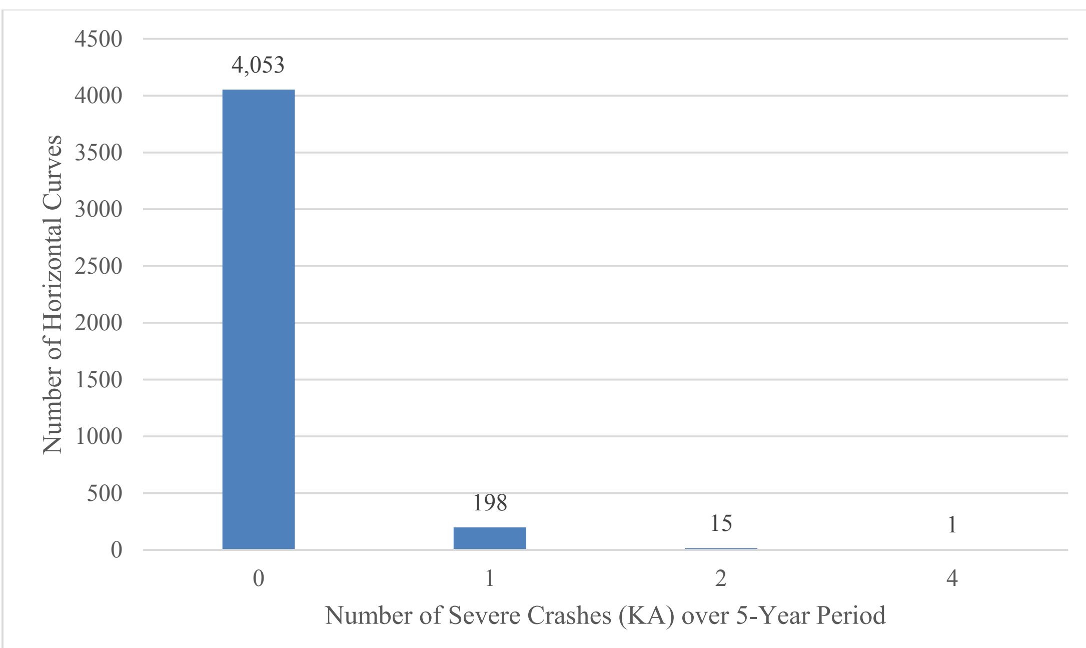


Figure 3.10 Severe crash (KA) frequencies for 4,267 horizontal curves selected for data analysis

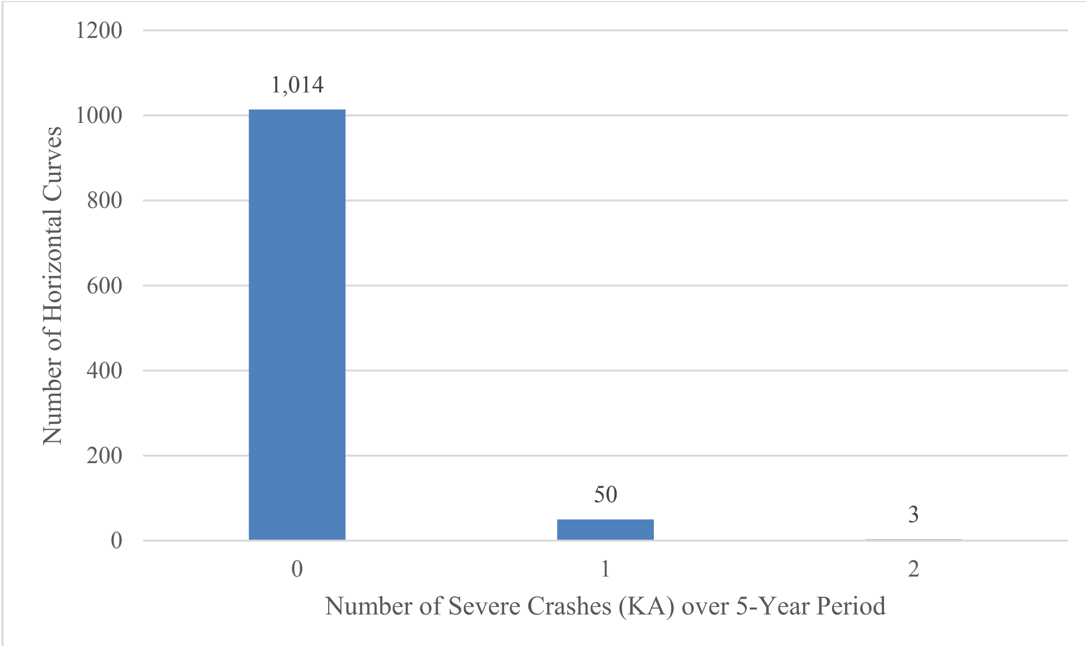


Figure 3.11 Severe crash (KA) frequencies for 1,067 horizontal curves selected for data validation

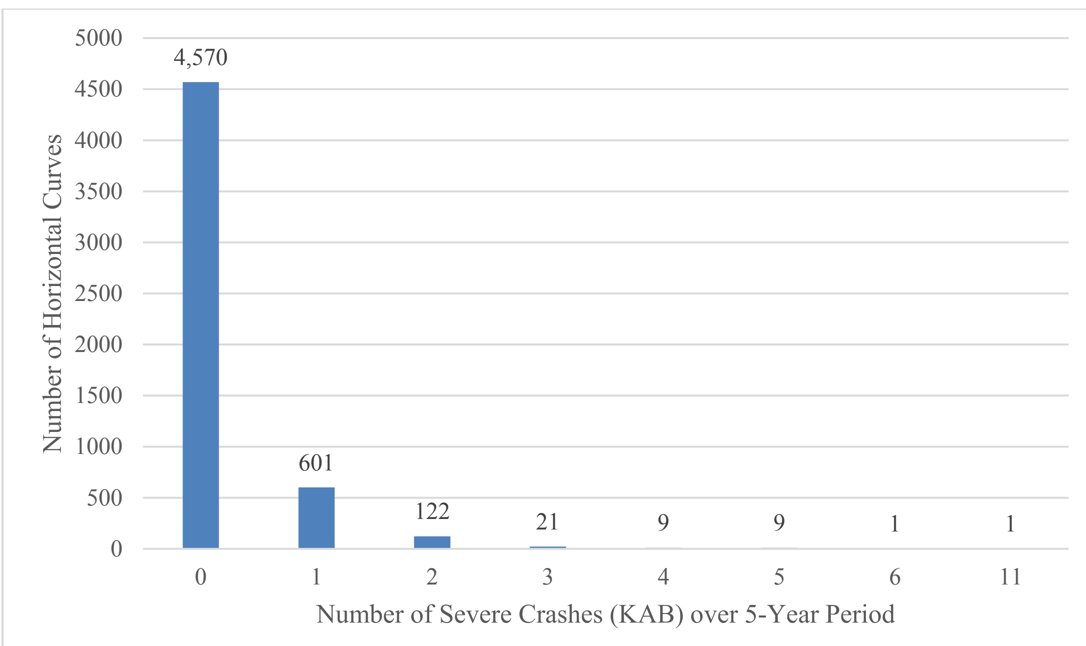


Figure 3.12 Severe (KAB) crash frequencies for 5,334 horizontal curves in Kansas

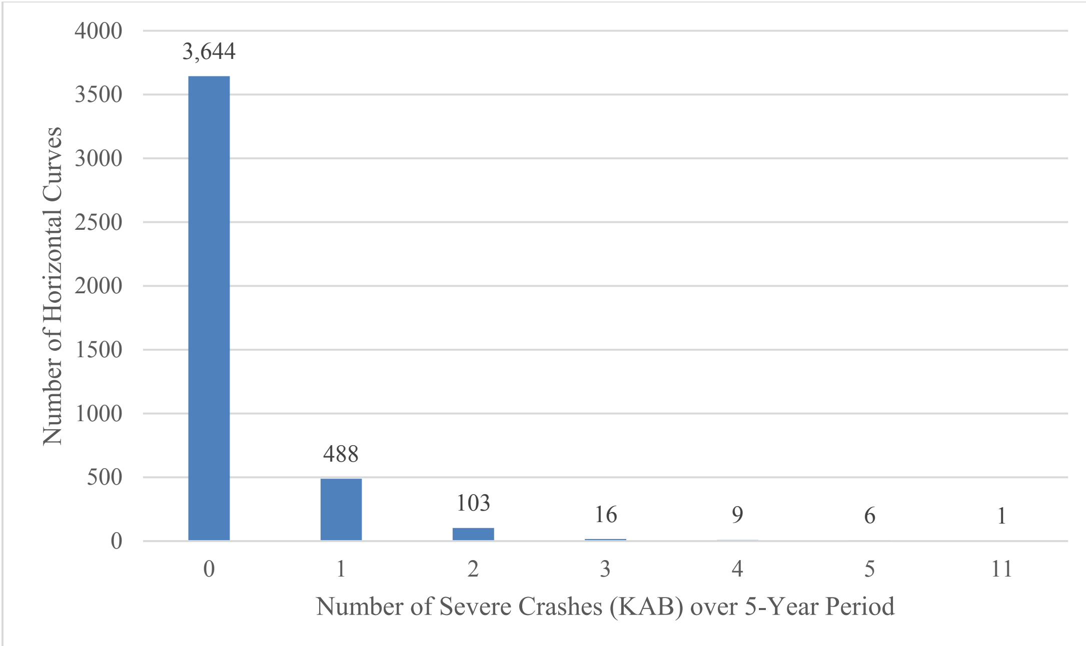


Figure 3.13 Severe crash (KAB) frequencies for 4,267 horizontal curves selected for data analysis

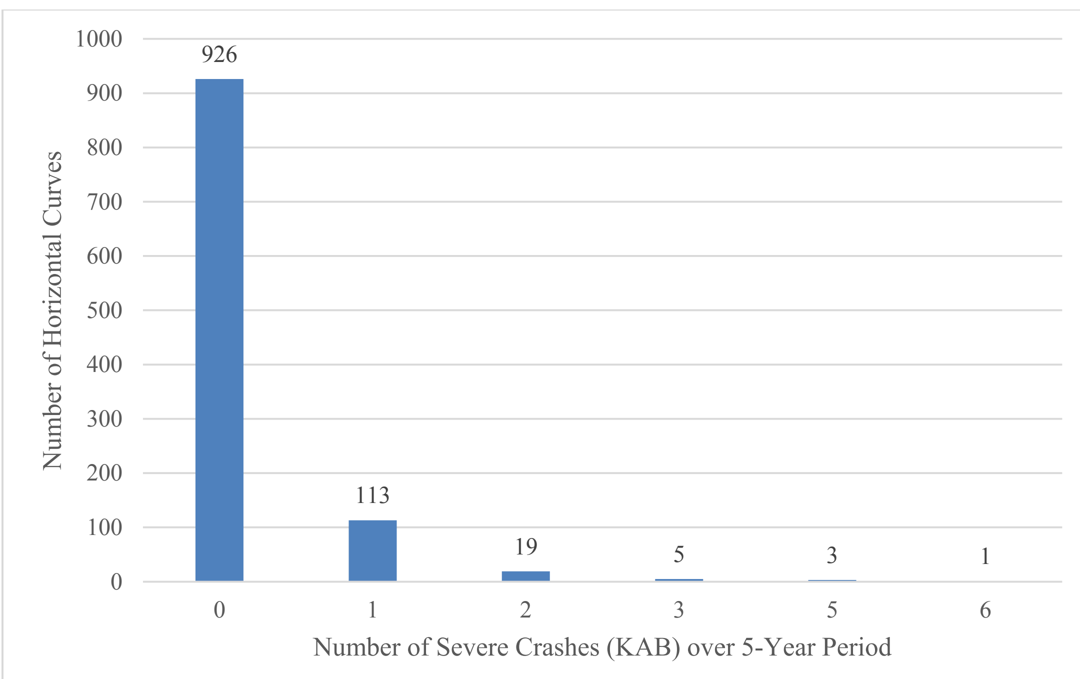


Figure 3.14 Severe crash (KAB) frequencies for 1,067 horizontal curves selected for data validation

Chapter 4. Effect of Speed Limit Change on Horizontal Curve Crashes on K-5 Highway in Leavenworth County, Kansas

Various studies determined speed of vehicles as the most important factor causing crashes at horizontal curve sections [**1**,**51**,**12**,**7**,**21**]; therefore, investigating the relationship between speed management and crash occurrence is an interesting topic. Due to a considerable number of recent crashes, in June 2009 the posted speed on K-5 highway in Leavenworth County, Kansas, was reduced to 50 mph. The effect of speed limit reduction as a policy countermeasure was investigated, and other roadway changes such as implemented countermeasures and roadside conditions were considered and examined. Figure 4.1 shows the segment of highway on which the speed limit change was applied.

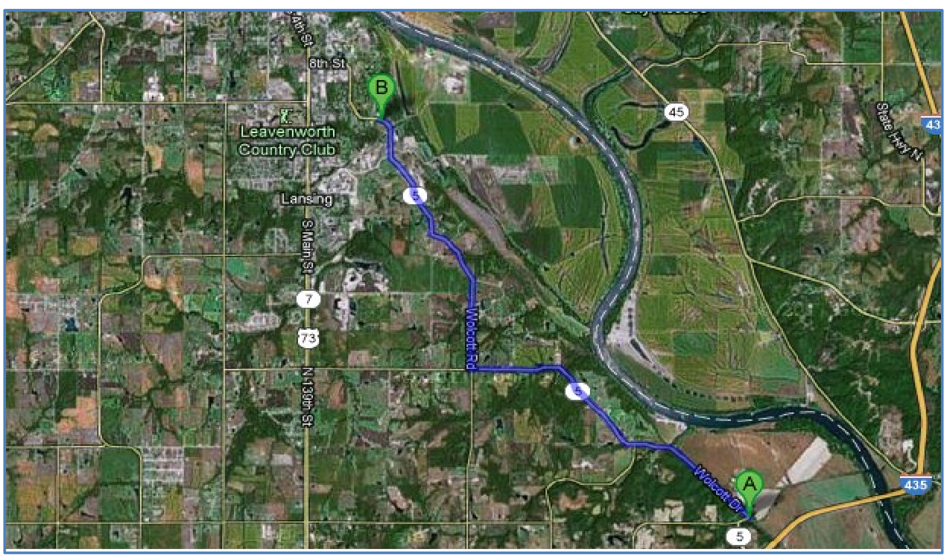


Figure 4.1 Segment of K-5 highway with speed limit change

4.1. Countermeasures on Horizontal Curves on K-5

In order to study countermeasure effectiveness, installation dates of the studied measures had to be known. Because the KDOT database does not contain this information, available video logs from the years 2004, 2007, and 2010 were studied in order to list applied measures at each horizontal curve section. Investigation of those logs pertaining to 25 horizontal curves on K-5 highway showed that the only alteration of applied countermeasures was the change of a "Winding Road" sign on one curve to a "Reverse Turn" sign. The first sign existed in video logs from 2004 and 2007, while the second sign was observed in the video log from 2010. Utilized countermeasures on the 25 studied curves consisted of centerline, edge line, horizontal alignment signs, advisory speed plaque, a one-direction large arrow sign, PMDs, chevrons, and a no-passing zone sign. During the study period, crashes occurred on 10 of the curves, and the remaining 15 curves had no crashes. With the exception of centerline and edge line markings on all curves, the use of other treatments is shown in Table 4.1.

Table 4.1 Number of curves with specific treatment

Treatment	Number of Curves	
	With Crashes	**Without Crashes**
Horizontal Alignment Signs	8	7
Advisory Speed Plaque	7	6
On-Direction Large Arrow	2	1
PMDs	0	6
Chevrons	8	1
No-Passing Zone Sign	0	2

Approximately 80% of curves with crashes had signs and supplemental treatments, such as chevrons, while less than half of curves without crashes had applied countermeasures that were identical to countermeasures of curves with crashes.

4.2. Roadside Hazard Rating

Individual roadside characteristics for each curve were investigated from the video logs. Roadside hazard rating (RHR) was determined for each curve using the first edition of the Highway Safety Manual (HSM) [**52**]. According to the video logs, few changes were evident in roadside characteristics. Table 4.2 shows RHRs for each roadside of the studied curves.

Table 4.2 RHR for curves on K-5 highway, Leavenworth County, Kansas

Year	2004		2007		2010	
Roadside	North	South	North	South	North	South
Curve No.	RHR	RHR	RHR	RHR	RHR	RHR
C-160	6	6	-	-	6	6
B-160	5	5	-	-	5	5
A-160	5	4	-	-	5	4
160	5	5	5	5	5	5
160-A	4	5	4	5	4	5
160-B	4	5	4	5	4	5
160-C	4	4	4	4	3	3
160-D	5	5	5	5	4	5
160-E	6	6	6	6	6	6
160-F	4	5	4	5	4	5
160-G	5	5	5	5	5	5
160-H	4	4	4	4	4	4
160-I	5	4	5	4	5	4
161	4	4	4	4	4	4
161-A	4	4	4	4	4	4
162	5	5	5	5	5	5
162-A	4	4	4	4	4	4
163	3	5	3	5	3	5
164	5	5	5	5	5	5
164-A	5	5	5	5	5	5
165	3	5	3	5	3	5
166	4	4	4	4	4	4
167	3	4	3	4	3	4
168	5	5	5	5	5	5
169	4	4	4	4	4	3

Images for the first three curves of the list did not exist in the video logs from 2007. Only four differences in RHRs of the three curves (160-C, 160-D, and 169) were observed among the 25 studied curves, indicating that no significant changes in roadside characteristics occurred during the study period. Similarly, no change of roadside characteristics of curves with crashes was observed.

4.3. Superelevation of Horizontal Curves

Information regarding superelevations of horizontal curve sections of the K-5 highway in Leavenworth County and other roadways was not available from KDOT; therefore, superelevations of horizontal curves of the studied highway were measured in the field. However, measurement of all superelevations of all horizontal curves was not possible due to lack of required safety measures, such as inadequate walking space along curve sections and the absence of traffic control equipment. Maximum superelevations of the studied horizontal curve sections are shown in Table 4.3.

Table 4.3 Superelevations of studied horizontal curves

Curve No.	Max. Superelevation (%)	Curve No.	Max. Superelevation (%)	Curve No.	Max. Superelevation (%)	Curve No.	Max. Superelevation (%)
C-160	n/a	160-D	n/a	161-A	n/a	166	6.1
B-160	n/a	160-E	n/a	162	4.7	167	4.4
A-160	n/a	160-F	n/a	162-A	n/a	168	4.9
160	4.4	160-G	n/a	163	4.7	169	5.2
160-A	n/a	160-H	n/a	164	n/a		
160-B	n/a	160-I	n/a	164-A	n/a		
160-C	n/a	161	10.5	165	n/a		

n/a: Not Available

According to roadway design principles, maximum superelevation is required at the one-third past the point of curvature (PC) and before the point of tangent (PT). Also, because the cross slope changes from normal slope (typically 1.6%) via a superelevation runoff length, the amount of superelevation should be less at the beginning of a curve compared to the center of a curve [**53**]. However, superelevations of horizontal curves on K-5 did not follow this principle.

4.4. K-5 Highway Curve-Related Crash Analysis

No changes in geometric characteristics and implemented conventional countermeasures were identified in the KDOT video logs; the only change applied to K-5 was speed limit reduction. In order to study the effectiveness of an applied policy countermeasure, a statistical t-test approach was used for crash frequencies and crash rates. This statistical method is recommended to analyze small sample sizes. For this study, a paired t-test was applied using a Statistical Analysis System (SAS 9.2) software package (SAS Institute, Cary, North Carolina). Crash frequencies and crash rates before and after the speed limit change for all 25 horizontal curves on K-5 highway were used for each group of crashes.

4.5. Speed Limit Reduction on K-5 Highway

According to the KDOT database, in 2009 the speed limit of K-5 highway was reduced from 55 mph to 50 mph. Data for 3.5 years before and 3.5 years after the speed limit change from 2006 to 2012 were used. During the study period, 45 crashes occurred at 10 horizontal curves out of the 25 horizontal curves on K-5 highway. Among those crashes, 29 occurred before the speed limit reduction and 16 occurred after the speed limit reduction. Thirty-six PDO crashes were noted, 24 of which occurred before the speed limit reduction and 12 which occurred after the speed limit reduction. Nine injury crashes occurred during the study time period: five occurred before the speed limit change and four occurred after the speed limit change. No fatal crashes were recorded at the horizontal curve sections during the study period. To provide a better perception of crash severity, equivalent property damage only (EPDO) crashes were used according to Equation (11).

$$EPDO = PDO + 15 \times (I + F) \tag{11}$$

Where:

EPDO: number of equivalent property damage only crashes,

PDO: number of property damage only crashes,

15: coefficient representing equivalent PDO crashes for injury and fatal crashes for Kansas,

I: number of injury crashes, and

F: number of fatal crashes.

Overall, 171 EPDO crashes occurred during the study period, with 99 EPDO crashes occurring before the speed limit reduction and 72 EPDO crashes occurring after the speed limit reduction.

Weather, light, and road surface conditions were considered for the EPDO crashes. Two characteristics were defined for weather conditions: no adverse weather conditions and adverse weather conditions. For no-adverse weather conditions, 76 EPDO crashes and 69 EPDO crashes occurred before and after the speed limit reduction, respectively. For adverse weather conditions, 26 EPDO crashes occurred: 23 before the speed limit change occurred and only three after the speed limit change. Light conditions were categorized as daylight conditions or dark-time conditions. For daylight condition, 44 and 53 EPDO crashes occurred before and after the speed limit reduction, respectively. For dark-time condition, 55 and 19 EPDO crashes occurred before and after the speed limit change, respectively. Road surface conditions were defined as dry conditions or wet conditions. For dry road surface conditions, 59 and 53 EPDO crashes occurred before and after the speed limit change, respectively. For wet road surface conditions, the dispersion before and after the speed limit change was 40 and 19 EPDO crashes, respectively. The numbers of crashes for each group and each time period are summarized in Table 4.4.

According to traffic count maps for the study period from 2006 to 2012, AADTs were 1636 and 2088 vpd before and after the speed limit change, respectively, indicating a 27% increase in AADT for the studied sections. KDOT video logs did not show specific geometric change on the studied sections.

Table 4.4 Number of crashes for each crash group before and after speed limit reduction

Crash group	Number of crashes		Percent Difference (%)
	Before speed limit change	After speed limit change	
Overall crashes	29	16	-44.83
PDO crashes	24	12	-50.00
Injury crashes	5	4	- 20.00
EPDO crashes	99	72	-27.27
EPDO crashes no-adverse weather condition	76	69	-9.21
EPDO crashes- adverse weather condition	23	3	-86.96
EPDO crashes- day light condition	44	53	20.45
EPDO crashes- dark time condition	55	19	-65.45
EPDO crashes- dry on road surface condition	59	53	-10.17
EPDO crashes wet on road surface condition	40	19	-52.50

Figure 4.2 and 4.3 show percentages of crash occurrence and crash severity, respectively, before and after the speed limit change at horizontal curve sections on K-5 highway.

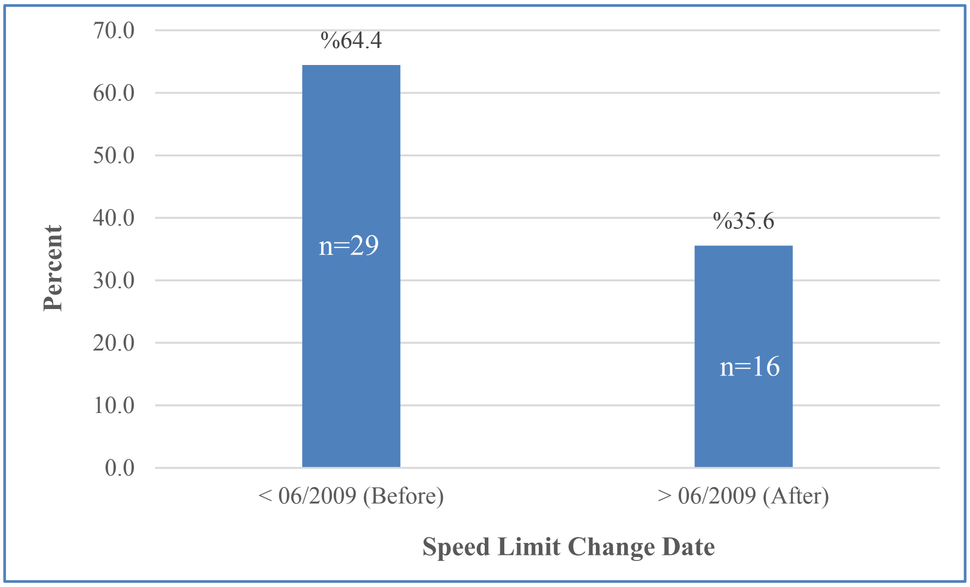


Figure 4.2 Crash percentage before and after speed limit reduction

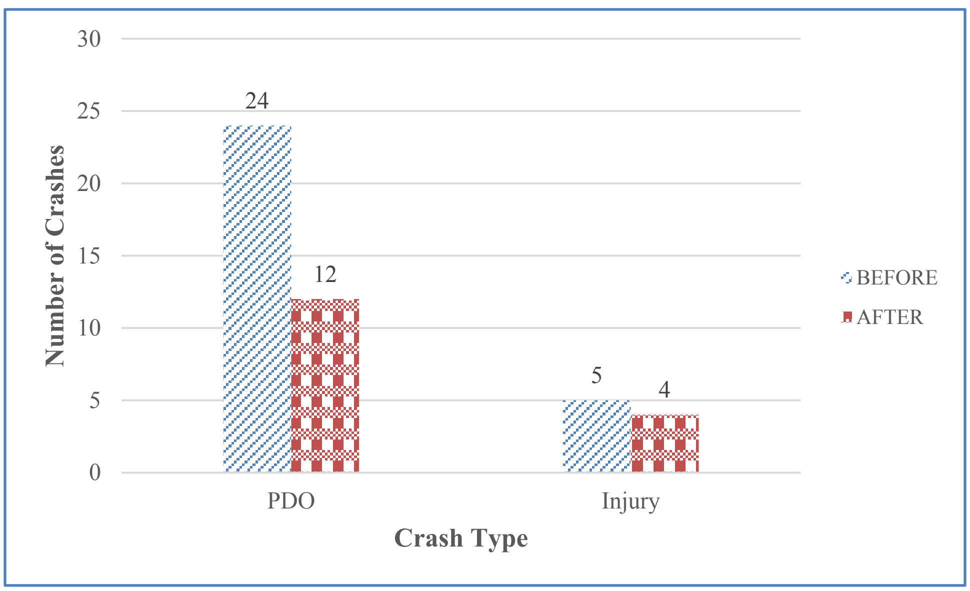
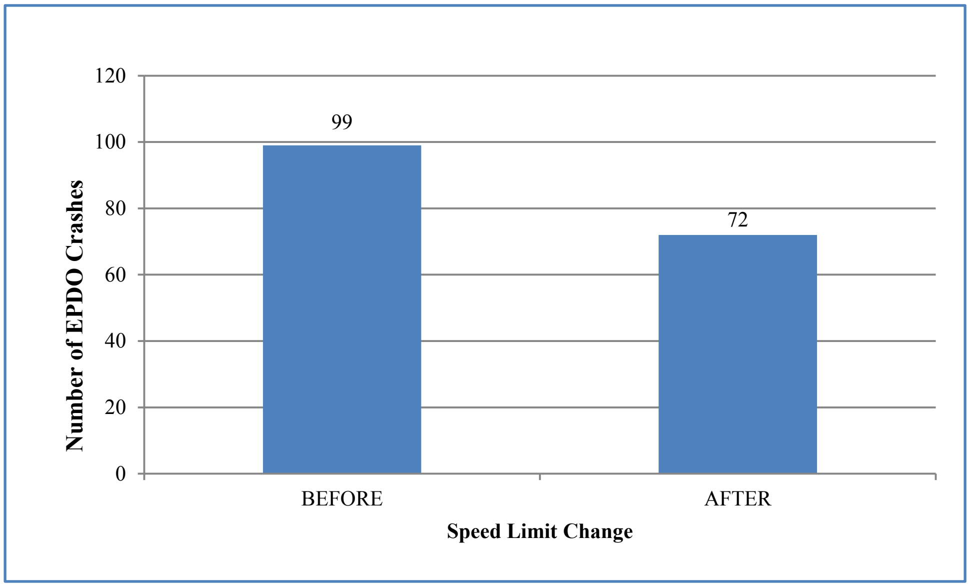
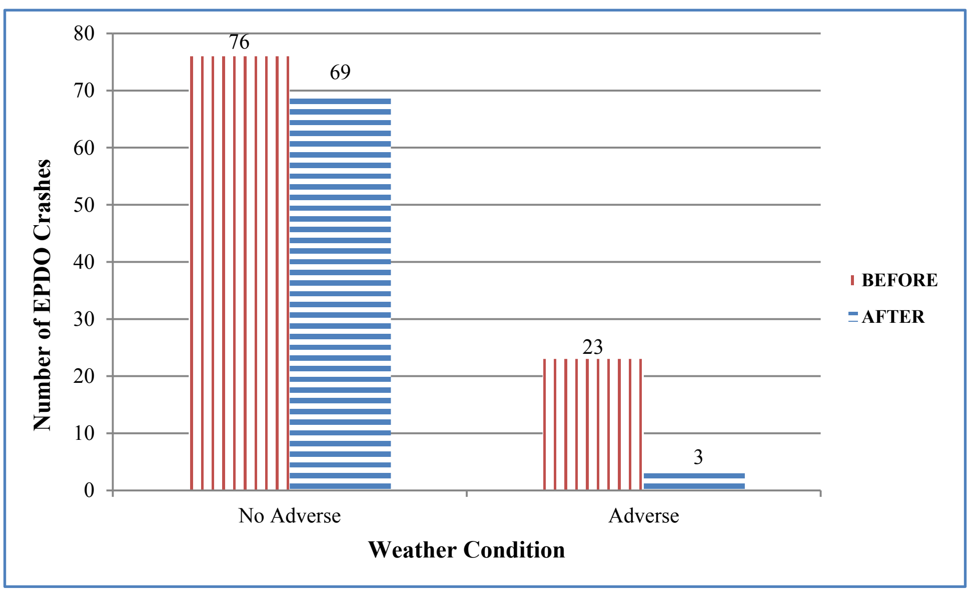


Figure 4.3 Crashes based on severity before and after speed limit change

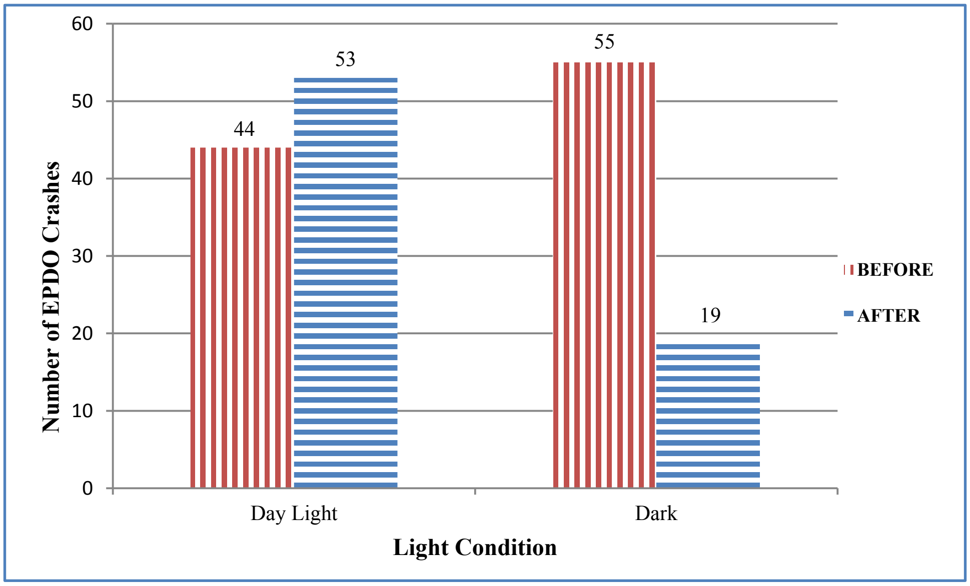
Figure 4.4(a)–(d) show changes in EPDO crashes and EPDO crashes for various weather, light, and road surface conditions.



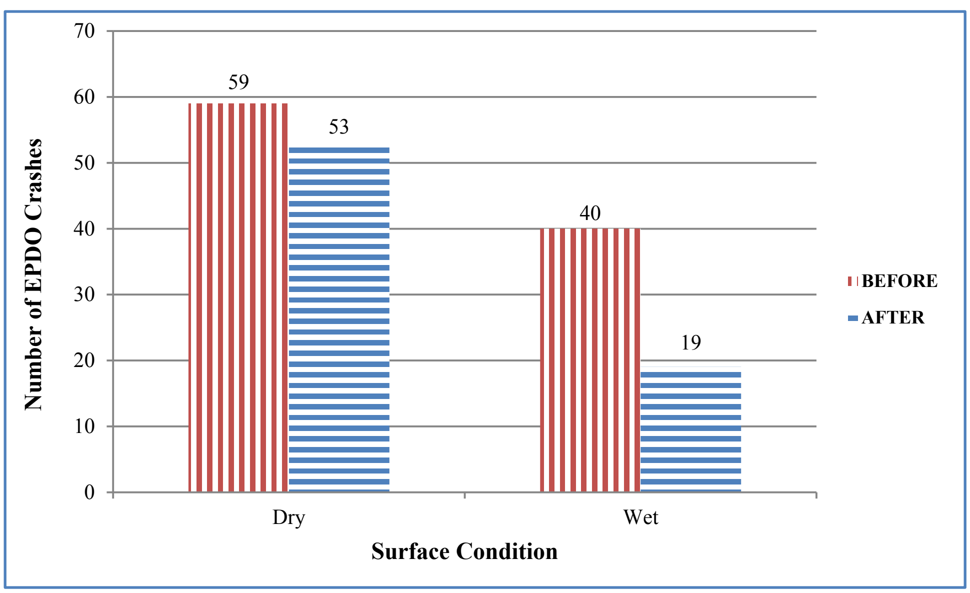
(a) EPDO crashes before and after speed limit change



(b) EPDO crashes before and after speed limit change for weather conditions



(c) EPDO crashes before and after speed limit change for light conditions



(d) EPDO crashes before and after speed limit change for surface conditions

Figure 4.4 EPDO crash changes before and after speed limit change for various conditions

4.6. T-test Analysis for Speed Limit Reduction

In order to determine if speed limit reduction influenced crash occurrences at horizontal curve sections on K-5 highway, a statistical SAS software calculated the t-value and p-value for each group of crashes and conditions. A comparison of p-values to the significance level of 5% indicated whether or not the speed limit reduction significantly influenced the particular crash group. Results of the applied method for crash frequency and crash rate are shown in Table 4.5.

Table 4.5 Results of t-test for crash frequencies and crash rate at K-5 horizontal curves

Crash group	Crash Frequency		Crash Rate	
	t-value	p-value	t-value	p-value
Overall crashes	1.74	0.115	1.84	0.099
PDO crashes	1.96	0.081	2.04	0.072
Injury crashes	0.43	0.678	0.71	0.497
EPDO crashes	0.71	0.497	0.995	0.346
EPDO crashes no-adverse weather condition	0.21	0.831	1.02	0.333
EPDO crashes- adverse weather condition	1.36	0.206	2.33	0.045
EPDO crashes- day light condition	-0.21	0.838	-0.15	0.887
EPDO crashes- dark time condition	1.54	0.157	1.75	0.114
EPDO crashes- dry road surface condition	0.23	0.820	0.73	0.426
EPDO crashes not dray road surface condition	0.77	0.463	0.75	0.474

Results of the t-test for crash frequencies, assuming a 5% significance level (equal to 95% confidence level), indicated that no crash group changes were statistically significant due to speed limit reduction. The EPDO crash rate in adverse weather conditions, however, was statistically significant at the 95% confidence level after the speed limit change. At the 10% significance level, the only statistically significant change was PDO crashes for crash frequency. Because p-values of overall crash rate, PDO crash rate, and EPDO crash rate for adverse weather conditions were less than 0.10, crash rates were significantly reduced at the 90% confidence level for PDO and EPDO.

Chapter 5. Results and Discussion

5.1. Data Analysis and Regression Models for Datasets

As mentioned, two datasets were used in data collection and data analysis procedures due to lack of information regarding horizontal curve sections in KDOT databases. Although the results of data analysis for the two datasets are presented in this chapter, each dataset contained different variables. For example, KDOT's horizontal curve inventory did not include information related to tangent sections prior to the horizontal curve sections. KDOT's GAD unit in the Traffic Safety Section in April 2015 prepared comprehensive data of horizontal curve sections on Kansas highways and distributed the data for use in this study.

5.1.1. First Dataset

At the beginning of data analysis, Statistical Package for the Social Science software (SPSS 18.0) was used to model Poisson regression and NB regression for all crashes on selected 221 horizontal curves on undivided two-lane two-way highways. A set variables including AADT from 2004 to 2012, heavy vehicle percentage, radius, degree of curvature, short and long tangent lengths, curve length, posted speed, advisory speed, difference between posted and advisory speeds, presence of rumble strips, shoulder type, and shoulder width was used to develop a crash prediction model. The results of data analysis with SPSS for the data of 221 horizontal curves on undivided two-lane two-way highways without considering tangent length sections showed neither Poisson nor NB methods, models crashes perfectly. Tangent length sections were then added to the dataset, and the data were analyzed by an SAS 9.3 software package, yielding a much better result since the tangent length variable was a significance variable in the new model. Results of Poisson and NB regression models for the complete data of

221 horizontal curves on undivided two-lane two-way highways are given in the following sections.

5.1.1.1. Poisson Regression Model

A Poisson regression model with all variables included was developed to predict crashes at randomly selected 221 horizontal curves on undivided two-lane two-way highways. The variable with the highest p-value (greater than significance level 5%) was eliminated from the model, and then the model was run for the remaining variables. This process was repeated until all p-values were less than the significance level, typically set to 0.05 or 5% in the study. Table 5.1 shows explanatory variables for the Poisson regression model to predict crashes at horizontal curve sections. Considering acceptable variables from Table 5.1 and regarding the general form of a Poisson regression model, the number of crashes (y_i) depended on the natural logarithm of AADT (logAADT), heavy vehicle percentage (HVPct), degree of curvature (DoC), and difference between posted speed and advisory speed (Diff_PS_AS), according to Equation (12).

Table 5.1 Poisson regression model variables for the first dataset with 221 horizontal curves

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3529	0.5805	-2.4907	-0.2151	5.43	0.0198
Ln AADT	1	0.2818	0.0699	0.1448	0.4188	16.24	<.0001
Heavy vehicle percentage	1	-0.0362	0.0121	-0.0598	-0.0125	8.97	0.0027
Degree of curvature	1	0.0337	0.0066	0.0209	0.0466	26.42	<.0001
Difference between posted speed and advisory speed	1	0.0130	0.0063	0.0007	0.0253	4.26	0.0391
Scale	0	1.0000	0.0000	1.0000	1.0000		

$$y_i = \exp(-1.353 + 0.282\, lnAADT - 0.036\, HVPct + 0.034\, DoC + 0.013\, Diff\_PS\_AS) \qquad \textbf{(12)}$$

As shown in Table 5.1, the Wald Chi-Square column, which is a conservative chi-square, is a squared *t*, where *t* is the value of the slope in the logistic regression divided by its standard error. Table 5.2 provides criterion for goodness of fit for the Poisson regression model in which the Pearson statistic value divided by the degree of freedom is higher than 1.0, indicating that the data were overdispersed. The AIC value was also high, meaning that the proposed model in Equation (12) may not accurately predict the number of crashes at horizontal curve sections. According to the Institute for Digital Research and Education (IDRE) at the University of California, Los Angeles (UCLA), for terms presented in the goodness-of-fit table, deviance is the log-likelihood of the regression model, Poisson or NB, multiplied by (-2). The deviance is calculated according to Equation (13) [**54**].

$$\sum_{i=1}^{n} 2\left(y_i log\frac{y_i}{\hat{y}_i} - (y_i - \hat{y}_i)\right) \quad \textbf{(13)}$$

where $\hat{y}_i$ is the predicted value of y_i.

The Pearson chi-square is a goodness-of-fit measure that compares outcome values with actual values according to Equation (14).

$$\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{\hat{y}_i} \quad \textbf{(14)}$$

where $\hat{y}_i$ is the predicted value of y_i.

The Bayesian information criterion (BIC) is another goodness-of-fit measure calculated using Equation (15).

$$\frac{-2\ln L + k\ln(n)}{n} \quad \textbf{(15)}$$

Similar to AIC, low values of BIC are preferred; the model with the lowest BIC is considered to be the best model.

Table 5.2 Goodness of fit for Poisson regression model for the first dataset with 221 horizontal curves

Criterion	DF	Value	Value/DF
Deviance	216	282.682	1.309
Scaled Deviance	216	282.682	1.309
Pearson Chi-Square	216	301.684	1.397
Scaled Pearson X2	216	301.684	1.397
Log-Likelihood		-75.765	
Full Log-Likelihood		-397.714	
AIC (smaller is better)		805.428	
AICC (smaller is better)		805.707	
BIC (smaller is better)		822.419	

5.1.1.2. Negative Binomial Regression Model

Table 5.3 shows explanatory variables for the NB regression model. Compared to the Poisson regression model, the same number of variables influenced crash occurrences at horizontal curve sections; however, posted speed in the Poisson regression method and the natural logarithm of the long tangent length in the NB regression method were essential variables. The values of Pearson chi-square divided by the degree of freedom and goodness-of-fit criteria showed that the NB regression method was a more accurate prediction model.

Table 5.3 NB regression model variables for the first dataset with 221 horizontal curves

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.853	0.711	-3.246	-0.460	6.80	0.009
Ln AADT	1	0.217	0.078	0.065	0.369	7.86	0.005
Heavy vehicle percentage	1	-0.029	0.014	-0.056	-0.002	4.32	0.038
Degree of curvature	1	0.047	0.008	0.031	0.062	34.98	0.000
Ln Long tangent length	1	0.120	0.044	0.034	0.208	7.49	0.006
Dispersion	1	0.160	0.051	0.085	0.300		

Equation (16) can be used to calculate the number of crashes at horizontal curve sections (y_i) according to the natural logarithm of AADT (logAADT), heavy vehicle percentage (HVPct), degree of curvature (DoC), and the natural logarithm of the length of the long tangent of the horizontal curve in feet (Ln_Long_Lt).

$$y_i = \exp(-1.853 + 0.217 \cdot logAADT - 0.029 \cdot HVPct + 0.047 \cdot DoC + 0.012 \cdot Ln\_Long\_Lt) \qquad \textbf{(16)}$$

Considering the Pearson chi-square divided by the degree of freedom in the goodness-of-fit table (Table 5.4), the conclusion was made that the NB regression method models crashes best at the horizontal curve sections because Pearson χ^2 divided by the degree of freedom was close enough to 1.0 and the AIC was less than its value for the Poisson method, as shown in Table 5.4.

Table 5.4 Goodness of fit for the NB regression model for the first dataset with 221 horizontal curves

Criterion	DF	Value	Value/DF
Deviance	216	197.4684	0.9142
Scaled Deviance	216	197.4684	0.9142
Pearson Chi-Square	216	217.3163	1.0061
Scaled Pearson X2	216	217.3163	1.0061
Log-Likelihood		-63.3780	
Full Log-Likelihood		-385.3271	
AIC (smaller is better)		782.6543	
AICC (smaller is better)		783.0468	
BIC (smaller is better)		803.0432	

Data analysis on the first dataset with 221 randomly selected horizontal curves on undivided two-lane two-way highways showed that the Poisson regression method was not an appropriate method to predict crashes at horizontal curve sections since the data were overdispersed, but the NB regression method, according to the goodness-of-fit criteria (particularly the value of Pearson chi-square divided by the degree of freedom), provided an acceptable estimate of crash numbers at horizontal curve sections. Because no other curve data was available to validate the results, prediction accuracy of the accepted NB model for crashes at horizontal curve sections could not be determined.

5.1.2. KDOT's Horizontal Curve Inventory

KDOT's horizontal curve inventory, which was completed in April 2015, is comprised of geometric data of 5,334 horizontal curves in the entire Kansas roadways, including curve length, curve radius, surface width, IRI, roadway type (divided or undivided), number of lanes, shoulder type and width, speed limit, advisory speed, presence of rumble strips, and roadway grade in percentage. In addition, traffic data such as AADT and heavy vehicle percentage were obtained from KDOT's GAD unit and added to the horizontal curves inventory data. Crashes and crash severities at horizontal curve sections were obtained from the KCARS database and added to the main dataset of the GAD unit of KDOT's Traffic Safety Section. Geometric crash characteristics from the KCARS database for crashes from 2010 to 2014 were compared to geometric characteristics of horizontal curve sections obtained from GAD unit database. In order to analyze data and validate data analysis, 80% (4,267 curves) of horizontal curve sections were selected for data analysis, and 20% (1,067 curves) of horizontal curve sections were selected for data validation. A SAS 9.4 software package (SAS Institute, Cary, North Carolina) was used for data analysis.

5.1.2.1. Data Analysis for KDOT New Horizontal Curves Dataset

In addition to Poisson and NB regression methods, ZIP and ZINB methods were utilized to analyze data and develop crash prediction models for total crashes and severe crashes due to the considerable number of horizontal curves with zero crashes. KABCO severity levels were used to consider crash severity, and two methods were utilized for severe crashes. The first approach considered K and A levels (fatal and incapacitating injury crashes) and the second approach included K, A, and B levels (fatal, incapacitating, and non-incapacitating injury) crashes. From the crash data extracted from the KCARS database, six types of crashes occurred at horizontal curve sections, including single-vehicle crashes, multi-vehicle crashes, crashes with pedestrians, crashes with pedal-cyclists, crashes with trains, and crashes with animals. For this study, single-vehicle crashes were selected for data analysis and validation because fewer external factors influence crash occurrence and consideration of one type of crash decreases complexity. Since the results of Poisson regression showed overdispersion of the crash data, NB, ZIP, and ZINB methods were used. Results are presented in the following sections.

5.1.2.1.1. Negative Binomial Regression for the KDOT Horizontal Curve Inventory Dataset

Because the results of the Poisson regression method showed overdispersion of crashes, the NB method was used to model total crashes, KA crashes, and KAB crashes. Backward and stepwise methods were utilized in order to eliminate explanatory variables that were not significant according to their p-values and to determine the best model for each crash group. Final remaining variables with accepted p-values were selected, and models were developed for each crash group.

Total Crashes

Total single-vehicle crashes for 4,267 horizontal curves were used for data analysis in SAS version 9.4 (SAS Institute, Cary, North Carolina). Explanatory variables were examined via backward and stepwise approaches and by adding and eliminating from the model to obtain the best NB regression model with significant variables. Table 5.5 shows significant variables for total single-vehicle crashes at horizontal curves sections randomly selected from KDOT's horizontal curves inventory.

Table 5.5 NB regression model variables for total single-vehicle crashes of KDOT horizontal curve inventory

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.1248	0.1811	-2.4798	-1.7698	137.62	<.0001
AADT (1000 vpd)	1	0.0303	0.0015	0.0274	0.0331	422.19	<.0001
Heavy vehicle percentage	1	-0.0378	0.0034	-0.0445	-0.0311	121.78	<.0001
Curve length	1	0.0005	0.0000	0.0004	0.0005	332.92	<.0001
Degree of curvature	1	0.0141	0.0044	0.0055	0.0226	10.30	0.0013
Posted speed	1	0.0193	0.0028	0.0138	0.0248	47.10	<.0001
Presence of rumble strips	1	0.0211	0.0062	0.0091	0.0332	11.79	0.0006
Dispersion	1	0.8813	0.0526	0.7840	0.9907		

According to the NB regression model, AADT (1000 vpd), curve length (ft.), degree of curvature, posted speed (mph), and difference between posted and advisory speeds (mph) positively influence crash occurrences at horizontal curve sections, and heavy vehicle percentage negatively influences crash occurrences at horizontal curve sections. The total number of crashes can be calculated according to Equation (17), in which coefficients were obtained from Table 5.5.

$$ttcrsh_i = \exp(-2.1248 + 0.0303 \cdot AADT\_th - 0.0378 \cdot HVPct + 0.0005 \cdot Curve\_Length + 0.0141 \cdot D\_o\_C + 0.0193 \cdot PS + 0.0193 \cdot DiffPSAS) \quad \textbf{(17)}$$

where

$ttcrsh_i$: total number of single-vehicle crashes at horizontal curve section *i*

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

DiffPSAS: difference between posted and advisory speeds (mph)

Table 5.6 shows goodness-of-fit criterion for the NB regression model for total single-vehicle crashes at horizontal curve sections in which the value of Pearson chi-square divided by the degree of freedom is approximately 1.05, which is acceptable.

Table 5.6 Goodness of fit for the NB regression model of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance	4,260	3,600.1709	0.8451
Scaled Deviance	4,260	3,600.1709	0.8451
Pearson Chi-Square	4,260	4,488.4948	1.0536
Scaled Pearson X2	4,260	4,488.4948	1.0536
Log-Likelihood		-1,622.5514	
Full Log-Likelihood		-4,847.3091	
AIC (smaller is better)		9,710.6182	
AICC (smaller is better)		9,710.6520	
BIC (smaller is better)		9,761.4875	

Severe Crashes (K and A levels)

As mentioned, two approaches were considered in order to analyze severe crashes in this study since fatal and incapacitating injury crashes are typically considered to be severe crashes in most studies. K and A levels were considered for the randomly selected horizontal curves from the KDOT inventory. Table 5.7 shows the best model of the NB regression method for severe crashes (K and A levels) for the selected horizontal curve sections.

Table 5.7 NB regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.5470	0.5778	-5.6795	-3.4145	61.93	<.0001
AADT (1000 vpd)		1	0.0178	0.0026	0.0128	0.0228	48.23	<.0001
Heavy vehicle percentage		1	-0.0211	0.0092	-0.0391	-0.0030	5.25	0.0220
Curve length		1	0.0003	0.0001	0.0002	0.0005	29.17	<.0001
Degree of curvature		1	0.0223	0.0085	0.0057	0.0389	6.93	0.0085
posted speed		1	0.0198	0.0095	0.0013	0.0384	4.40	0.0359
Presence of rumble strips	1	1	-0.3783	0.1588	-0.6897	-0.0670	5.67	0.0172
Dispersion		1	0.8943	0.4673	0.3212	2.4902		

Using the NB regression method for severe crashes (K and A levels), AADT, percentage, curve length, and posted speed positively influence severe crash occurrences and the number of severe crashes at horizontal curve sections; heavy vehicle and presence of rumble strips negatively influence severe crash occurrences and the number of severe crashes at those sections. Equation (18) presents the prediction model from the NB regression model for severe crashes (K and A levels).

$$KAcrsh_i = \exp(-4.5470 + 0.0178 \cdot AADT\_th - 0.0211 \cdot HVPct + 0.0003 \cdot Curve\_Length + 0.0223 \cdot D\_o\_C + 0.0198 \cdot PS - 0.3783 \cdot RS) \quad \textbf{(18)}$$

where

$KAcrsh_i$: the number of severe crashes (K and A levels) at horizontal curve section *i*

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section in feet,

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

RS: presence of rumble strips at horizontal curve section

Table 5.8 shows goodness-of-fit criterion for the NB regression model for severe crashes (K and A levels) at horizontal curve sections in which the value of Pearson chi-square divided by the degree of freedom is approximately 1.02, which is acceptable.

Table 5.8 Goodness of fit for the NB regression model of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance	4,260	1,127.0997	0.2646
Scaled Deviance	4,260	1,127.0997	0.2646
Pearson Chi-Square	4,260	4,328.3983	1.0161
Scaled Pearson X2	4,260	4,328.3983	1.0161
Log-Likelihood		-848.3476	
Full Log-Likelihood		-861.9229	
AIC (smaller is better)		1,739.8457	
AICC (smaller is better)		1,739.8795	
BIC (smaller is better)		1,790.7150	

Severe Crashes (K, A, and B levels)

Some studies define severe crashes as fatal and injury crashes (incapacitating and non-incapacitating injury crashes), K, A, and B levels from KABCO crash severity scale [**39**,**50**]. Since many horizontal curve sections in this study had zero K and A severe crashes, a severe crash dataset containing K, A, and B levels was determined, and the NB regression method was utilized to develop a prediction model for a new severe crashes dataset. Table 5.9 shows the best model of the NB regression method for severe crashes (K, A, and B levels) for the selected horizontal curve sections.

Table 5.9 NB regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.8088	0.3649	-4.5240	-3.0937	108.98	<.0001
AADT (1000 vpd)		1	0.0216	0.0020	0.0177	0.0255	116.98	<.0001
Heavy vehicle percentage		1	-0.0297	0.0059	-0.0414	-0.0181	25.07	<.0001
Curve length		1	0.0005	0.0000	0.0004	0.0006	151.86	<.0001
Degree of curvature		1	0.0168	0.0071	0.0029	0.0308	5.58	0.0182
Posted speed		1	0.0259	0.0062	0.0137	0.0381	17.31	<.0001
Difference between posted speed and advisory speed		1	0.0225	0.0100	0.0029	0.0422	5.08	0.0242
Presence of rumble strips	1	1	-0.2046	0.1114	-0.4230	0.0139	3.37	0.0664
Shoulder width	2	1	-0.1826	0.1621	-0.5003	0.1351	1.27	0.2600
Shoulder width	3	1	-0.2813	0.1367	-0.5492	-0.0134	4.23	0.0396
Shoulder width	4	1	-0.3733	1.3236	-2.9675	2.2209	0.08	0.7779
Dispersion		1	0.9212	0.1497	0.6700	1.2666		

For severe crashes (K, A, and B levels), the results of the NB regression method showed that AADT, curve length, degree of curvature, posted speed, and difference between posted and advisory speeds positively influence severe crash (K, A, and B levels) occurrences and the number of severe crashes (K, A, and B levels) at horizontal curve sections; heavy vehicle percentage, presence of rumble strips, and shoulder width negatively influence severe crash (K, A, and B levels) occurrences and the number of severe crashes (K, A, and B levels) at those sections. Equation (19) shows the resultant prediction model from the NB regression model for severe crashes (K, A, and B levels).

$$KABcrsh_i = \exp(-3.8088 + 0.0216 \cdot AADT\_th - 0.0297 \cdot HVPct + 0.0005 \cdot Curve\_Length + 0.0168 \cdot D\_o\_C + 0.0259 \cdot PS + 0.0225 \cdot DiffPSAS - 0.2046 \cdot RS - 0.2813 \cdot RLSH\_W_3) \quad \textbf{(19)}$$

where

$KABcrsh_i$: the number of severe crashes (K, A, and B levels) at horizontal curve section *i*

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

DiffPSAS: difference between posted and advisory speed on the roadway section (mph)

RS: presence of rumble strips at horizontal curve section

$RLSH\_W_3$: shoulder width between 3 and 7 ft.

Table 5.10 shows goodness-of-fit criterion for the NB regression model for severe crashes (K, A, and B levels) at horizontal curve sections in which the value of Pearson chi-square divided by the degree of freedom is approximately 1.02, which is acceptable.

Table 5.10 Goodness of fit for the NB regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance	4,256	2,120.7650	0.4983
Scaled Deviance	4,256	2,120.7650	0.4983
Pearson Chi-Square	4,256	4,328.5662	1.0171
Scaled Pearson X2	4,256	4,328.5662	1.0171
Log-Likelihood		-1,820.3338	
Full Log-Likelihood		-1,995.2258	
AIC (smaller is better)		4,014.4517	
AICC (smaller is better)		4,014.5250	
BIC (smaller is better)		4,090.7557	

5.1.2.1.2. Zero-inflated Poisson Regression for the KDOT Horizontal Curve Inventory Dataset

In addition to the overdispersion of crash data, many horizontal curves had zero crashes, thereby justifying use of ZI models that consider overdispersion and excessive zeros. Therefore, ZI models were developed for total crashes and severe crashes. Results of the ZIP method are explained in the following sections.

Total Crashes

Backward and stepwise approaches were used to determine the best ZIP model for total crashes. Because the Pearson chi-square divided by the degree of freedom of the model was very high, 1.29, as shown in Table 5.11, the method was not used to develop the prediction model for total crashes at horizontal curve sections from KDOT's inventory.

Table 5.11 Goodness of fit for the ZIP regression model of total crashes of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		10,077.9552	
Scaled Deviance		10,077.9552	
Pearson Chi-Square	4,236	5,477.4541	1.2931
Scaled Pearson X2	4,236	5,477.4541	1.2931
Log-Likelihood		-1,814.2199	
Full Log-Likelihood		-5,038.9776	
AIC (smaller is better)		10,139.9552	
AICC (smaller is better)		10,140.4236	
BIC (smaller is better)		10,337.0738	

Severe Crashes (K and A levels)

The ZIP method was utilized for severe crashes (K and A levels), and backward and stepwise approaches were used to select explanatory variables with significant p-values. The final ZI model with significant variables is provided in Table 5.12.

Table 5.12 ZIP regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.1961	0.6689	-5.5072	-2.8849	39.35	<.0001
AADT (1000 vpd)		1	0.0161	0.0032	0.0098	0.0225	24.80	<.0001
Curve length		1	0.0003	0.0001	0.0002	0.0005	25.88	<.0001
Degree of curvature		1	0.0518	0.0164	0.0198	0.0839	10.03	0.0015
Posted speed		1	0.0334	0.0111	0.0117	0.0551	9.11	0.0025
Divided	1	1	-1.1471	0.2707	-1.6777	-0.6165	17.95	<.0001
Presence of rumble strips	1	1	-0.5342	0.1799	-0.8867	-0.1816	8.82	0.0030
Scale		0	1.0000	0.0000	1.0000	1.0000		

For severe crashes (K and A levels), results of the ZIP regression method showed that AADT, curve length, degree of curvature, and posted speed positively influence severe crash (K and A levels) occurrences and the number of severe crashes (K and A levels) at horizontal curve sections; divided situation and presence of rumble strips negatively influence severe crash (K and A levels) occurrences and the number of severe crashes (K and A levels) at those sections. Equation (20) presents the prediction model for severe crashes (K and A levels).

$$KAcrsh_i = \exp(-4.1961 + 0.0161 \cdot AADT\_th + 0.0003 \cdot Curve\_Length + 0.0518 \cdot D\_o\_C + 0.0334 \cdot PS - 1.1471 \cdot Divided - 0.5342 \cdot RS) \quad \textbf{(20)}$$

where

$KAcrsh_i$: the number of severe crashes (K and A levels) at horizontal curve section i

$AADT\_th$: average annual daily traffic in thousands in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

Divided: binary variable for undivided roadways is 0 and 1 for divided roadways

RS: presence of rumble strips at horizontal curve section

Table 5.13 shows goodness-of-fit criterion for the ZIP regression model for severe crashes (K and A levels) at horizontal curve sections in which the value of Pearson chi-square divided by the degree of freedom is approximately 0.96, which is acceptable. However, according to results of the SAS program, the Hessian convergence criterion equaled 0.0028, which is greater than the set value of SAS (0.0001), thereby making the ZIP model results questionable.

Table 5.13 Goodness of fit for the ZIP regression model of severe crashes (K and A levels) of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		1,704.8333	
Scaled Deviance		1,704.8333	
Pearson Chi-Square	4,258	4,071.9362	0.9563
Scaled Pearson X2	4,258	4,071.9362	0.9563
Log-Likelihood		-838.8414	
Full Log-Likelihood		-852.4166	
AIC (smaller is better)		1,722.8333	
AICC (smaller is better)		1,722.8756	
BIC (smaller is better)		1,780.0613	

Severe Crashes (K, A, and B levels)

The ZIP regression method was also used to develop a prediction model for severe crashes (K, A, and B levels). Table 5.14 shows the best model of the ZIP method for severe crashes (K, A, and B levels) after applying backward and stepwise approaches.

Table 5.14 ZIP regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1081	0.1528	-1.4076	-0.8086	52.58	<.0001
AADT (1000 vpd)		1	0.0086	0.0015	0.0056	0.0116	31.22	<.0001
Heavy vehicle percentage		1	-0.0267	0.0063	-0.0390	-0.0144	18.04	<.0001
Curve length		1	0.0003	0.0000	0.0002	0.0004	49.51	<.0001
Degree of curvature		1	0.0322	0.0094	0.0137	0.0507	11.66	0.0006
Difference between posted speed and advisory speed		1	0.0203	0.0105	-0.0003	0.0408	3.73	0.0534
Presence of rumble strips	1	1	-0.1639	0.0894	-0.3391	0.0112	3.37	0.0666
Scale		0	1.0000	0.0000	1.0000	1.0000		

According to the ZIP regression method, independent variables that influence K, A, and B levels of crashes are AADT, heavy vehicle percentage, curve length, degree of curvature, difference between posted and advisory speeds, and presence of rumble strips. AADT, curve length, and difference between posted and advisory speeds positively impact severe crash occurrences at horizontal curve sections; heavy vehicle percentage and presence of rumble strips negatively impact severe crash occurrences at horizontal curve sections. A significance level of

5% was assumed in this research study; the p-values of difference between posted and advisory speeds and the presence of rumble strips were relatively close to 5%, so they were included in the model. Equation (21) shows the resultant prediction model from the ZIP regression model for severe crashes (K, A, and B levels).

$$KABcrsh_i = \exp(-1.1081 + 0.0086 \cdot AADT\_th - 0.0267 \cdot HVPct + 0.0003 \cdot Curve\_Length + 0.0322 \cdot D\_o\_C + 0.0203 \cdot DiffPSAS - 0.1639 \cdot RS) \quad \textbf{(21)}$$

where

KABcrsh$_i$: the number of severe crashes (K, A, and B levels) at horizontal curve section *i*

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

DiffPSAS: difference between posted and advisory speed on the roadway section (mph)

RS: presence of rumble strips at horizontal curve section

Table 5.15 shows goodness-of-fit criterion for the ZIP regression model for severe crashes (K, A, and B levels) at horizontal curve sections. Results of SAS version 9.4 (SAS Institute, Cary, North Carolina) showed that the value of Pearson chi-square divided by the degree of freedom was in an acceptable range, as shown in Table 5.15.

Table 5.15 Goodness of fit for the ZIP regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		3,944.1577	
Scaled Deviance		3,944.1577	
Pearson Chi-Square	4,248	4,067.0491	0.9574
Scaled Pearson X2	4,248	4,067.0491	0.9574
Log-Likelihood		-1,797.1868	
Full Log-Likelihood		-1,972.0789	
AIC (smaller is better)		3,982.1577	
AICC (smaller is better)		3,982.3367	
BIC (smaller is better)		4,102.9724	

5.1.2.1.3. Zero-inflated Negative Binomial Regression for the KDOT Horizontal Curve Inventory Dataset

The zero-ZINB regression method also considers overdispersion and excessive zeros. Results of the ZINB method for the studied crash groups are discussed in the following sections.

Total Crashes

This study utilized a ZINB method to develop a prediction model of all the crashes that occurred at the randomly selected horizontal curves during the study period from 2010 to 2014. Backward and stepwise approaches were implemented to find the best model with significant variables at 5% significance level. Table 5.16 lists the significant explanatory variables of the ZINB model that influenced the occurrence of total crashes at the selected horizontal curves.

Table 5.16 ZINB regression model variables for total single-vehicle crashes of KDOT horizontal curve inventory

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.2397	0.2395	-2.7091	-1.7704	87.48	<.0001
AADT (1000vpd)	1	0.0237	0.0014	0.0209	0.0265	279.74	<.0001
Heavy vehicle percentage	1	-0.0335	0.0036	-0.0406	-0.0264	86.21	<.0001
Curve length	1	0.0003	0.0000	0.0003	0.0004	162.42	<.0001
Degree of curvature	1	0.0236	0.0060	0.0117	0.0354	15.23	<.0001
Posted speed	1	0.0243	0.0032	0.0181	0.0306	58.54	<.0001
Difference between posted speed and advisory speed	1	0.0292	0.0072	0.0150	0.0434	16.25	<.0001
International roughness index (IRI)	1	0.0031	0.0010	0.0011	0.0052	9.07	0.0026
Dispersion	1	0.6440	0.0479	0.5566	0.7451		

According to the ZINB regression model, AADT (1000 vpd), curve length (ft.), degree of curvature, posted speed (mph), difference between posted and advisory speeds (mph), and IRI positively influence total crash occurrences at horizontal curve sections; heavy vehicle percentage negatively influences total crash occurrences at those sections. Based on the ZINB prediction model, the total number of crashes was calculated according to Equation (22).

$$ttcrsh_i = \exp(-2.2397 + 0.0237 \cdot AADT\_th - 0.0335 \cdot HVPct + 0.0003 \cdot Curve\_Length + 0.0236 \cdot D\_o\_C + 0.0243 \cdot PS + 0.0292 \cdot DiffPSAS + 0.0031 \cdot IRI) \quad \textbf{(22)}$$

where

ttcrsh$_i$: the total number of single-vehicle crashes at horizontal curve section *i*

AADT_th: average annual daily traffic on the section (1000 vpd)

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

DiffPSAS: difference between posted and advisory speeds (mph)

IRI: international roughness index

Table 5.17 shows goodness-of-fit criterion for the NB regression model for total single-vehicle crashes at horizontal curve sections in which the value of Pearson chi-square over the degree of freedom is close to one and the developed model can accurately predict the total number of crashes.

Table 5.17 Goodness of fit for the ZINB regression model of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		9,512.5679	
Scaled Deviance		9,512.5679	
Pearson Chi-Square	4,249	4,018.9449	0.9459
Scaled Pearson X2	4,249	4,018.9449	0.9459
Log-Likelihood		-4,756.2840	
Full Log-Likelihood		-4,756.2840	
AIC (smaller is better)		9,550.5679	
AICC (smaller is better)		9,550.7469	
BIC (smaller is better)		9,671.3826	

Severe Crashes (K and A levels)

This study used ZINB regression method, and sets of explanatory variables were examined via backward and stepwise approaches in order to determine with the best set of independent variables for the prediction model. Table 5.18 shows the predictor variables and their coefficients for the best model developed based on the ZINB regression method for severe crashes (K and A levels) for the selected horizontal curve sections.

Table 5.18 ZINB regression model variables for severe crashes (K and A levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-6.8915	0.7845	-8.4291	-5.3539	77.17	<.0001
AADT(1000 vpd)		1	0.0184	0.0029	0.0127	0.0240	40.70	<.0001
Heavy vehicle percentage		1	-0.0227	0.0115	-0.0453	-0.0001	3.88	0.0488
Curve length		1	0.0002	0.0001	0.0000	0.0003	3.91	0.0479
Degree of curvature		1	0.0949	0.0154	0.0648	0.1251	38.07	<.0001
Posted speed		1	0.0712	0.0139	0.0440	0.0984	26.37	<.0001
Divided	1	1	-0.8036	0.3038	-1.3990	-0.2082	7.00	0.0082
Presence of rumble strips	1	1	-0.4817	0.1767	-0.8281	-0.1353	7.43	0.0064
Dispersion		1	0.3513	0.3408	0.0525	2.3524		

According to the ZINB model, variables AADT, curve length, degree of curvature, and posted speed positively influence severe crash occurrences and the number of severe crashes (K and A levels) at horizontal curve sections; heavy vehicle percentage, divided roadways, and presence of rumble strips negatively influence severe crash occurrences and the number of severe crashes (K and A levels) at those sections. Equation (23) shows the best ZINB prediction model for severe crashes (K and A levels) at the studied horizontal curve sections.

$$KAcrsh_i = \exp(-6.8915 + 0.0184 \cdot AADT\_th - 0.0227 \cdot HVPct + 0.0002 \cdot Curve\_Length + 0.0949 \cdot D\_o\_C + 0.0712 \cdot PS - 0.8036 \cdot Divided - 0.4817 \cdot RS) \qquad \textbf{(23)}$$

where

$KAcrsh_i$: the number of severe crashes (K and A levels) at horizontal curve section *i*

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

Divided: binary variable for undivided roadways is 0 and 1 for divided roadways

RS: presence of rumble strips at horizontal curve section

Table 5.19 shows that the Pearson chi-square divided by the degree of freedom value was acceptable. The table also shows other goodness-of-fit criterion of the ZINB model for severe crashes (K and A levels).

Table 5.19 Goodness of fit for the ZINB regression model of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		1,683.2516	
Scaled Deviance		1,683.2516	
Pearson Chi-Square	4,254	4,119.6591	0.9684
Scaled Pearson X2	4,254	4,119.6591	0.9684
Log-Likelihood		-841.6258	
Full Log-Likelihood		-841.6258	
AIC (smaller is better)		1,711.2516	
AICC (smaller is better)		1,711.3503	
BIC (smaller is better)		1,800.2729	

Severe Crashes (K, A, and B levels)

The ZINB regression method was used to analyze K, A, and B level severe crashes. Table 5.20 shows the independent variables, their coefficients, and other statistical characteristics of the best model obtained from the ZINB regression method.

Table 5.20 ZINB regression model variables for severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.3404	0.4152	-4.1542	-2.5267	64.73	<.0001
AADT(1000 vpd)		1	0.0121	0.0019	0.0085	0.0158	43.03	<.0001
Heavy vehicle percentage		1	-0.0285	0.0064	-0.0410	-0.0160	20.04	<.0001
Curve length		1	0.0003	0.0000	0.0002	0.0004	43.85	<.0001
Degree of curvature		1	0.0571	0.0138	0.0300	0.0842	17.09	<.0001
Posted speed		1	0.0318	0.0062	0.0196	0.0441	26.01	<.0001
Different between posted and advisory speed		1	0.0247	0.0113	0.0025	0.0468	4.75	0.0294
Presence of rumble strips	1	1	-0.4279	0.1146	-0.6525	-0.2033	13.94	0.0002
Dispersion		1	0.3692	0.1086	0.2074	0.6572		

According to the ZINB regression method for severe crashes (K, A, and B levels), AADT, curve length, degree of curvature, posted speed, and difference between posted and advisory speeds positively influence the number of severe crashes (K, A, and B levels) at horizontal curve sections; heavy vehicle percentage and presence of rumble strips negatively influence the number of severe crashes (K, A, and B levels) at those sections. The prediction model for severe crashes (K, A, and B levels) based on the ZINB method is presented in Equation (24).

$$\begin{aligned} KABcrsh_i = \exp(&-3.3404 + 0.0121 \cdot AADT\_th - 0.0285 \cdot HVPct + 0.0003 \\ &\cdot Curve\_Length + 0.0571 \cdot D\_o\_C + 0.0318 \cdot PS + 0.0247 \\ &\cdot DiffPSAS - 0.4279 \cdot RS) \end{aligned} \quad \textbf{(24)}$$

where

KABcrsh$_i$: the number of severe crashes (K, A, and B levels) at horizontal curve section i

AADT_th: average annual daily traffic in thousands in the section

HVPct: heavy vehicle percentage in the section

Curve_Length: length of the horizontal curve section (ft.)

D_o_C: degree of curvature of the section

PS: posted speed on the roadway section (mph)

DiffPSAS: difference between posted and advisory speed on the roadway section (mph)

RS: presence of rumble strips at horizontal curve section

Table 5.21 shows goodness-of-fit criterion for the ZINB regression model for severe crashes (K, A, and B levels) at horizontal curve sections.

Table 5.21 Goodness of fit for the ZINB regression model of severe crashes (K, A, and B levels) of KDOT horizontal curve inventory

Criterion	DF	Value	Value/DF
Deviance		3,908.5624	
Scaled Deviance		3,908.5624	
Pearson Chi-Square	4,251	4,021.4718	0.9460
Scaled Pearson X2	4,251	4,021.4718	0.9460
Log-Likelihood		-1,954.2812	
Full Log-Likelihood		-1,954.2812	
AIC (smaller is better)		3,942.5624	
AICC (smaller is better)		3,942.7064	
BIC (smaller is better)		4,050.6597	

Appendix C contains written codes in SAS for data analysis in order to develop prediction models in this research study.

5.1.2.1.4. Comparison of Developed Models

This study utilized a total of four methods for data analysis for crashes at horizontal curve sections, including, Poisson, NB ZIP, and ZINB methods. The Poisson regression method was used first, but due to overdispersion, the NB, ZIP, and ZINB methods were selected for data analyses. This section includes the results of these models for the studied crash groups.

Total Crashes

Because the value of Pearson chi-square divided by the degree of freedom was very high for total crashes in the ZIP model, the NB and ZINB models were compared. As shown in Table 5.5 and Table 5.17 the values of goodness-of-fit AIC criterion of NB and ZINB were 9,710.6 and 9,550.6, respectively. AADT, heavy vehicle percentage, curve length, degree of curvature, posted speed, and difference between posted and advisory speeds variables commonly influenced crash occurrence in the NB and ZINB models. IRI also influenced crash occurrence in the ZINB model. Table 5.22 summarizes the significant variables for total crashes.

Table 5.22 Variables used to predict total crashes in the selected models

Variables	NB Model	ZINB Model
AADT_th (AADT (1000 vpd))	✓	✓
HVPct (Heavy vehicle percentage)	✓	✓
Curve_Length	✓	✓
D_o_C (Degree of curvature)	✓	✓
PS (Posted speed)	✓	✓
DiffPSAS (Difference between posted and advisory speeds)	✓	✓
IRI (international roughness index)		✓

Severe Crashes (K and A levels)

According to the results of data analysis from SAS 9.4 (SAS Institute, Cary, North Carolina) for severe crashes, the NB, ZIP, and ZINB models can be used to develop the prediction model. According to Table 5.7, Table 5.13, and Table 5.17, the ZINB model had the lowest AIC criteria value and the NB model had the highest AIC criteria value, with 1739.8, 1722.9, and 1711.3 for NB, ZIP, and ZINB models, respectively. Variables of AADT, curve length, degree of curvature, posted speed, and rumble strips were included in all selected NB, ZIP, and ZINB models. The divided roadway variable had to be added for ZIP and ZINB models, and heavy vehicle percentage had to be included in the NB and ZINB models. Table 5.23 summarizes the significant variables for severe crashes (K and A levels).

Table 5.23 Variables used to predict severe crashes (K and A levels) in the selected models

Variables	NB Model	ZIP Model	ZINB Model
AADT_th (AADT (1000 vpd))	✓	✓	✓
HVPct (Heavy vehicle percentage)	✓		✓
Curve_Length	✓	✓	✓
D_o_C (Degree of curvature)	✓	✓	✓
PS (Posted speed)	✓	✓	✓
RS (Rumble strips)	✓	✓	✓
Divided		✓	✓

Severe Crashes (K, A, and B levels)

Results of data analysis for severe crashes (K, A, and B levels) from SAS 9.4 (SAS Institute, Cary, North Carolina) showed that NB, ZIP, and ZINB methods could be used to develop the prediction model for this crash group. The values of AIC criterion for NB, ZIP, and ZINB were 4014.5, 3982.2, and 3942.6, respectively, as obtained from Table 5.10, Table 5.15, and Table 5.21. Considering K, A, and B crash occurrences at horizontal curve sections, the results of data analysis revealed that AADT, heavy vehicle percentage, curve length, degree of curvature, difference between posted and advisory speed, and presence of rumble strips influence severe crashes (K, A, and B levels) using the NB, ZIP, or ZINB models. Posted speed is another influencing parameter that must be included in NB and ZINB models. In addition, in NB model, the shoulder width is another parameter which influence crash occurrence. A summary of influencing parameters on K, A, and B crashes is provided in Table 5.24.

Table 5.24 Variables used to predict severe crashes (K, A, and B levels) in the selected models

Variables	NB Model	ZIP Model	ZINB Model
AADT_th (AADT (1000 vpd))	✓	✓	✓
HVPct (Heavy vehicle percentage)	✓	✓	✓
Curve_Length	✓	✓	✓
D_o_C (Degree of curvature)	✓	✓	✓
PS (Posted speed)	✓		✓
DiffPSAS (Difference between posted and advisory speeds)	✓	✓	✓
RLSH_W_3 (Shoulder width between 3 ft. and 7 ft.)	✓		
RS (Rumble strips)	✓	✓	✓

5.2. Data Validation for the New KDOT Horizontal Curve Inventory

Although goodness-of-fit criteria showed that the selected models can predict crash occurrence at horizontal curves, the accuracy of the prediction models was not clear. Therefore, KDOT's horizontal curve inventory was divided into two sets: data analysis and data validation. Out of the 5,334 horizontal curves on the entire state transportation network in the KDOT inventory, 80% (4,267) of the curves were randomly selected for data analysis. The remaining 20% (1,067) curve sections were used to conduct data validation.

For data validation, total crashes and severe crashes (K and A levels and K, A, and B levels) were calculated from applicable NB, ZIP, and ZINB models for each randomly selected horizontal curves. Then, crash numbers of horizontal curves for each crash group and each model were added up. The sum of crashes of each method was compared to the sum of observed crashes at the randomly selected horizontal curve sections. Table 5.25 shows data validation results and compares predicted crashes of each crash group to observed crashes at the randomly selected horizontal curve sections for NB, ZIP, and ZINB models.

Table 5.25 Comparison of observed and predicted crashes at selected horizontal curve sections for data validation

Crash Type		Number of Crashes			
		Observed	Predicted		
			NB Model	ZIP Model	ZINB Model
Total Crashes		982	1,208 (+23%)	-	1,123 (+14%)
Severe Crashes	K, A, and B Crashes	187	119 (-36%)	233 (+25%)	158 (-15%)
	K and A Crashes	56	1 (-98%)	14 (-75%)	33 (-41%)

As shown in Table 5.25, during the study period from 2010 to 2014, a total of 982 single-vehicle crashes occurred at horizontal curve sections selected for data validation. The developed NB model for total crashes predicted 1,208 crashes at the randomly selected 1,067 horizontal curves, while the ZINB model predicted 1,123 crashes. The numbers of crashes predicted by NB and ZINB models were 23% and 14%, respectively higher than the number of observed crashes. For severe crashes (K, A, and B levels), 187 crashes were observed at the 1,067 randomly selected curves, but NB, ZIP, and ZINB models predicted 119, 233, and 158 severe crashes, respectively. NB and ZINB predicted severe crashes (K, A, and B levels) 36% and 15% less than the observed crashes, respectively, while the ZIP model showed a 25% increase in predicted severe crashes compared to observed crashes. Predicted severe crashes (levels K and A) for NB, ZIP, and ZINB models were 1, 14, and 33, respectively. Comparison of predicted and observed severe crashes (K and A levels) showed that NB and ZIP results were much less than actual severe crashes: 98% and 75%, respectively. However, the result of ZINB model for severe crashes (K and A levels) was only 41% less than the observed severe crashes (K and A levels).

According to data validation results for all crash groups, total crashes, and severe crashes, ZINB model results were closer to the observed crashes at the randomly selected horizontal curve sections for data validation compared to the NB and ZIP methods. However, comparison of crash groups for each model showed that the difference between observed and predicted crashes was relatively greater when crash severity was considered, especially for fatal and incapacitating injury crashes. Comparisons of observed crashes and predicted crashes for individual curves are included in Appendix D.

Chapter 6. Summary, Conclusions and Recommendations

6.1. Summary

The number of fatalities due to vehicle crashes on rural and urban highways has decreased during the last decade, but the percentage of fatal crashes on horizontal curve sections was approximately constant and more than 25% during last years from 2008 to 2014 with the majority of ROR crashes [**5**]. Therefore, identification of factors that contribute to crash occurrences at horizontal curve sections would improve curve safety and reduce crash occurrence risk at these sections. This dissertation focused on geometric and traffic data of horizontal curve sections in Kansas. Because a majority of the required data was unavailable at the beginning of the research, only 221 horizontal curves on undivided two-lane two-way highways were randomly selected for this study. Geometric and traffic data were measured or obtained using various software tools or databases. However, a comprehensive curve inventory with extensive data was provided in April 2015 and is also included in this dissertation.

Data analysis for the first horizontal curves dataset with 221 horizontal curves on undivided two-lane two-way highways was conducted using SAS 9.3 (SAS Institute, Cary, North Carolina), and Poisson and NB regression models were developed for 221 selected horizontal curve sections on undivided two-lane two-way highways using collected AADT, heavy vehicle percentage, curve length, radius, degree of curvature, posted speed, advisory speed, difference between posted and advisory speeds, presence of rumble strips, shoulder type, and shoulder width. Results for the Poisson regression model showed overdispersion since Pearson χ^2 divided by the degree of freedom was much greater than 1. Therefore, the NB regression method was required to take into account the overdispersion of crash data. Variables that affect crash occurrence at horizontal curves for the NB regression method include AADT, heavy vehicle percentage, degree of curvature, and the length of long tangent section. The value of Pearson χ^2 divided by the degree of freedom (1.006) was very close to 1.0, proving the accuracy of the NB model. Moreover, comparison of AIC values for Poisson and NB regression methods showed that the NB regression method more accurately estimated crash numbers at horizontal curve sections due to its lower AIC value (742.6 for the NB model versus 805.4 for the Poisson model). However, previous data analysis without considering tangent section length did not lead to a reasonable prediction model for the selected 221 horizontal curves on undivided two-lane two-way highways with Poisson and NB regression methods using SPSS 18.0.

In addition to the 221 horizontal curves on undivided two-lane two-way highways, KDOT completed a Kansas horizontal curve inventory in April 2015, which contained 5,334 horizontal curves on the entire state transportation network. This study utilized the curve inventory dataset from the KCARS database and traffic data from KDOT to construct a dataset with all required and combined data for 5,334 horizontal curves on the entire state transportation network. In order to verify accuracy of the developed models, the horizontal curve data were divided into two groups: one group for data analysis and the other group for data validation. Eighty percent (4,247) of horizontal curves were randomly selected for data analysis, and the remaining 20% (1,067) of horizontal curves were used for data validation. SAS 9.4 (SAS Institute, Cary, North Carolina) was used for data analysis, and models were developed using Poisson, NB, ZIP, and ZINB methods. Results of the Poisson method showed overdispersed crash data and many curves with zero number of crashes. Therefore, NB and ZI methods were examined for data analysis. Total crashes, severe crashes (K and A levels), and severe crashes (K, A, and B levels) were analyzed. Two severe crash groups were utilized because some studies only consider fatal and incapacitating injury crashes to be severe crashes, while other studies also consider non-incapacitating injury crashes to be severe crashes. Consequently, this study considered two types of severe crashes and compared the results of data analysis for each group.

The KDOT crash database indicated that K-5 highway contained the highest numbers of crashes on horizontal curves; therefore, additional investigations were conducted on horizontal curve sections of that roadway. Video logs from years 2004, 2007, and 2010 did not show changes in applied countermeasures at horizontal curve sections or significant changes in roadside characteristics of the studied curves. The measured superelevation of some curves with crashes showed that these sections met the minimum requirement of superelevations; however, changes in superelevations along curve sections were not constant and did not meet current design guidance or criteria.

Existing data revealed that the speed limit of a roadway segment with approximately 25 horizontal curves was reduced in June 2009. Therefore, impact of the speed limit change on crash reduction at horizontal curve sections was studied. For data analysis, a time period from 2006 to 2012, including 3.5 years before and 3.5 years after the speed limit change, was selected. Although initial data showed a reduction in crash frequencies for crashes and EPDO crashes in various light, weather, and road surface conditions, a statistical t-test did not indicate numbers high enough to conclude that crash frequencies and crash rates showed statistically significant reduction due to speed limit change at the 95% confidence level (5% significance level). However, EPDO crash rates for adverse weather conditions significantly decreased at the 5% significance level.

6.2. Conclusions

The results of data analysis with NB regression method for randomly selected 221 horizontal curves on undivided two-lane two-way highways revealed that the important parameters influencing crash occurrence at Kansas two-lane highways were AADT, heavy vehicle percentage, and the length of long tangent section of the curve. For the KDOT horizontal curve inventory including 5,334 horizontal curves on the entire state transportation network, 80% of horizontal curve selected randomly for data analysis and NB, ZIP, and ZINB methods were used to develop prediction models. For total crashes, according to the best NB and ZINB developed models AADT, heavy vehicle percentage, curve length, degree of curvature, posted speed, and difference between posted and advisory speeds were significantly influence the total crash occurrences at the randomly selected 4,267 horizontal curves for data analysis. In addition, IRI that shows the rideability of the roadway's pavement is another parameter which influence total crash occurrence at the selected horizontal curves. For severe crashes (K and A levels), AADT, curve length, degree of curvature, posted speed and presence of rumble strips found to be the significant parameters influencing K and A crashes at horizontal curve sections for all NB, ZIP, and ZINB models. For NB and ZINB models, heavy vehicle percentage and for ZIP and ZINB models being divided or undivided roadways influenced KA crash occurrence at the selected horizontal curves. For severe crashes (K, A, and B levels), in all of the NB, ZIP, and ZINB models, AADT, heavy vehicle percentage, curve length, degree of curvature, difference between posted and advisory speeds, and presence of rumble strips significantly influenced K, A, and B crash occurrences at the selected horizontal curves. Additionally, for NB model shoulder width between 3ft to 7ft and for ZINB model posted speed were the other variables that influenced severe crash (K, A, and B levels) occurrences at the selected horizontal curve sections. Furthermore, the comparisons of models for various crash groups indicated that AADT, heavy vehicle percentage, curve length, degree of curvature, and speed (in posted speed or difference between posted and advisory formats) significantly influence crash occurrences for all crash groups, while presence of rumble strips is only significant when the crash severity is considered. For the most severe crashes, fatal and incapacitating crashes, a binary variable for divided or undivided roadways has significant impact on crash occurrences at the selected horizontal curves. However, for total crashes using ZINB method, IRI is a significant parameter influencing the crash occurrences, but it is not a significant variable for the severe crashes.

Considering the goodness-of-fit criteria for all of the NB, ZIP, and ZINB models for all of the crash groups showed that ZINB is the best model and can better predict number of total and severe crashes since the value of AIC criterion is less than the other prediction models. Moreover, data validation was conducted using the data of 1,067 horizontal curves, remaining 20% of 5,334 horizontal curves on the entire state transportation network, and the observed total and severe crashes were compared. The results of data validation showed that the ZINB models for all crash groups are closer to the observed crashes; however, the accuracy of models decreased when the severity of crashes was considered. For total number of crashes, ZINB model predicts crashes 14% less than observed total crashes, while for severe crashes (K, A, and B levels) the difference percentage is -15% and for severe crashes this value dramatically drops to -41%. Overall, the results of data analysis and data validation show the acceptable correlation between the prediction models numbers of crashes and the observed crashes for the studied crash groups.

The findings of this research study and the developed crash prediction models are in accordance with the results of other studies. Various studies concluded that crashes increase with increase in AADT [**30**,**39**,**55**,**45**,**37**,**33**,**30**,**31**], increase in degree of curvature or decrease in curve radius [**29**,**31**,**39**,**33**,**36**,**29**,**30**], increase in posted speed or difference between posted and advisory speeds [**31**,**39**,**29**,**56**], increase in curve length [**36**,**37**,**29**,**57**,**30**], and increase in IRI

[**39**]. However, there is no consistency in the effect of heavy vehicle percentage on crash occurrence among different studies. Some studies concluded that the increase in truck ADT increases the number of crashes [**37**,**45**,**33**]. Kapetanakis concluded that crashes decrease due to increase in heavy vehicle percentage [**55**]. Sharma et al. found that crashes increase with heavy vehicle percentage increase. In another study Kim et al. concluded truck percentage-mile-per-lane has a dual impact on crash occurrences [**58**]. Increase in heavy vehicle percentage increases the crashes that a vehicle becomes an obstacle; on the contrary, increase in heavy vehicle percentage reduces the crashes that occurred due to the following vehicle's reaction failure [**58**].

6.3. Countermeasures to Improve Safety of Horizontal Curves Based on the Developed Prediction Models

. Various results of this research study confirm the findings of previous researches and provide useful suggestions for improving the safety of horizontal curve sections.

According to the results of data analysis of the first dataset with 221 horizontal curves on undivided two-lane two-way highways, the length of the long tangent section positively influenced the number of total crashes at horizontal curve sections. Thus the increase in length of the long tangent section increased the possibility of crash occurrences at horizontal curve section. In other words, when drivers navigate a long tangent section, they are less prepared for a horizontal curve, especially a sharp curve [**1**]. Therefore, sufficient warning and information must be provided to drivers when the length of the tangent section is long and the drivers do not expect a change in roadway alignment. Increased visibility of warning signs due to better materials with higher retroreflectivity can improve the safety of horizontal curve sections, particularly in dark and adverse weather conditions [**58**,**59**].

For the KDOT horizontal curve inventory, results of the ZINB model showed that IRI significantly and positively influenced the total single-vehicle crashes at horizontal curve sections. In other word, higher values of IRI indicate an increased probability for single-vehicle crashes at horizontal curve sections. High IRI also indicates an uneven pavement surface, which negatively affects driving quality on the roadway and negatively impacts rideability of the roadway. Thus, drivers feel less comfortable when IRI has a high value. Therefore, a smooth pavement surface at horizontal curve sections can help improve the safety of horizontal curve sections.

Although the presence of rumble strips was not found a significant parameter for crash occurrences at horizontal curve sections in Kansas, results of data analysis showed that rumble strips mitigate crash severity at horizontal curve sections. Results of data analysis for severe crashes showed that the presence of rumble strips at a horizontal curve section can cause up to $e^{-0.53}$ reduction in the number of severe crashes. Various studies have been conducted to find the appropriate pattern of rumble strips in order to increase their effectiveness in Kansas [**60**,**61**,**62**]. According to the results of NB model for severe crashes, increasing shoulder width to 7 ft. influenced severe crash occurrences and reduced single-vehicle crash severity at horizontal curve sections. Therefore, use of rumble strips and increased shoulder width at horizontal curve sections are low-cost countermeasures to improve safety at these sections. Table 6.1 summarizes the recommended countermeasures to safety improvement of the studied horizontal curves based on the crash types.

Table 6.1 Recommended countermeasures for the studied horizontal curves based on the crash types

Countermeasure	Crash Type		
	Total Crashes	Single Vehicle Crashes	Severe Single Vehicle Crashes
Centerline and edge line	✓	✓	✓
Advance warning signs	✓	✓	✓
Supplemental treatments with advanced warning signs such as doubling up signs, higher retroreflective materials	✓	✓	✓
Speed management treatments such as chevrons enhanced with LEDs	✓	✓	✓
On pavement markings	✓	✓	✓
Flashing beacons	✓	✓	✓
Widening shoulder			✓
Rumble strips			✓
Dividing opposite lanes			✓

6.4. Recommendations for Future Studies

This research study made assumptions and simplifications in order to increase the feasibility of the data analysis. For example, only geometric and traffic characteristics were considered for data collection and data analysis. In reality, however, other parameters, such as weather conditions, roadway surface conditions, and driver circumstances, affect crash occurrences at horizontal curve sections. Future research studies should consider the effect of other environmental characteristics and human factors in conjunction with the current collected and analyzed data.

Data analysis of collected data for 221 selected horizontal curves on undivided two-lane two-way highways showed that long tangent section length significantly influences crash occurrence at horizontal curve sections. Therefore, recommendations were made for additional warning and guide signs in advance of horizontal curves with long tangent sections. However, the results did not show the critical tangent length. A future study should find the critical tangent section length as a threshold for providing additional information about changes in roadway alignment via additional warning and guide signs. Also, according to data analysis results, fewer severe crashes occur on divided highways potentially due to better geometric characteristics of divided highways. Future research should compare geometric characteristics of divided and undivided highways and determine effective parameters that decrease the single-vehicle crash severity on divided highways. Also, a future study should further investigate the effect of heavy vehicle percentage on crash occurrences at horizontal curve sections due to the inconsistency between the results of various studies conducted in this area.

References

[1] J. L. Campbell, M. G. Lichty, J. L. Brown, C. M. Ricahrd, J. S. Graving and J. Graham, "NCHRP Report 600 Second Edition:Human factors Guidlines for road sysytems," Transportation Research Board, 2012.

[2] H. Godthelp, "Vehicle Control During Curve Driving," *Human Factors,* vol. 28, no. 2, pp. 211-221, 1986.

[3] O. Simsek, A. C. Bittner, W. H. Levison and S. A. Garness, "Development of Prototype Driver Model for Highway Design," *Task B: Speed-Decision Simulator Experiment. Battelle Human Factors Transportation Center,* 2000.

[4] J. Cheung, "Horizontal Curve Safety," FHWA, U. S. Department of Transportation, 2010.

[5] "Fatality Analysis Reporting System (FARS)," National Highway Traffic Safety Administration, [Online]. Available: http://www.nhtsa.gov/FARS.

[6] R. Christiansen, "Ahead of the Curves," *Thinking Highways,* pp. 44-46, 2011.

[7] H. W. McGee and F. R. Hanscom, "Low-cost treatment for horizontal curve safety," Federal Highway Adminstration, 2006.

[8] FHWA, "Manual of Uniform Traffic Control Devices (MUTCD)," FHWA, U. S. Department of Transportation, 2003.

[9] FHWA, "Manual of Uniform Traffic Control Devices (MUTCD)," FHWA, U. S. Department of Transportation, 2009.

[10] C. D. Bogenreif, "Systemic Highway Safety Assessment: A General Analysis of Funding Allocation and a Specific Study of the Horizontal Curve Crash Problem," Iowa State University, 2011.

[11] D. Camacho, "An overview of FHWA's nine proven safety countermeasures," Institute of Transportation Engineers Purto Rico Section, San Juan, 2012.

[12] D. M. Lord, M. A. Brewer, K. Fitzpatrick, S. R. Geedipally and Y. Peng, "Analysis of roadway departure crashes on two-lane rural roads in Texas," Texas Department of Transportation, 2011.

[13] K. Fitzpatrick, K. Balke, D. W. Harwood and I. B. Anderson, "NCHRP Report 440: Accident Mitigation Guide for Congested Rural Two-Lane Highways," Transportation Research Board, Washington DC, 2000.

[14] T. R. Neuman, R. Pfefer, K. L. Slack, K. K. Hardy, F. Council, H. McGee, L. Prothe and K. Eccles, "NCHRP500: Guidance for Implementation of the AASHTO Strategic Highway Safety Plan. Volume 6: A Guide for Addressing Run-Off-Road Collisions," Transportation Research Board, Washington DC, 2003.

[15] S. Hallmark, N. Hawkins and O. Smadi, "Toolbox of countermeasures for rural two-lane curves," Midwestern Transportation Consortium, Ames, 2012.

[16] Handbook of Traffic Control Practices for Low Volume Rural Roads, Topeka: Kansas Department of Transportaion, 1991.

[17] R. Amjadi, "Safety evaluation of improved curve delineation," Federal Highway Adminstration, Washington D.C., 2009.

[18] "Highway Sign Manual," Kansas Department of Transportation, 2007.

[19] T. McDonald, "Iowa's Traffic Safety Manual," Iowa Department of Transportation, Ames, 2012.

[20] A. Gan, J. Shen and A. Rodriguez, "Update of Florida Crash Reduction Factors and Countermeasures to improve the Development of District Safety Improvement Projects," State of Florida Department of Transportation , Tallahassee, 2005.

[21] S. G. Charlton, "The role of attention in horizontal curves: A comparison of advance warning, delineaion, and road marking treatments," Accident analysis and prevention, 2007.

[22] E. R. Russell and M. J. Rys, "Centerline Rumble Strips," Transportation Research Board(TRB), Washington D.C., 2005.

[23] D. E. Karkle, M. J. Rys and E. R. Russell, "State-of-the-Art: Centerline Rumble Strips Usage in the United States," *Transportation Research Forum,* vol. 50, pp. 100-117, 2011.

[24] M. J. Rys, D. E. Karkle and E. R. Russell, "Study of KDOT Policy on Lane and Shoulder Minimum Width for Application Shoulder Minimum Width for Application," Kansas Department of Transportation(KDOT), Topeka, 2012.

[25] C. Ibarra, "LED Raised Pavement Markers," FHWA, U. S. Department of Transportation, 2009.

[26] B. Everard, "Complementary Information of Sequential Dynamic Curve Warning System (SDCWS) and In-Pavement Markings," Milwaukee, 2013.

[27] D. Lord, S. P. Washington and J. N. Ivan, "Poisson, Poisson-gamma and Zero-inflated Regression Models of Motor Vehicle Crashes: Balancing Statistical Fit and Theory," *Accident Analysis & Prevention,* vol. 1, no. 37, pp. 35-46, 2005.

[28] "Mitigation Strategies For Design Exceptions: FHWA," [Online]. Available: http://safety.fhwa.dot.gov/geometric/pubs/mitigationstrategies/chapter6/6_mitigameas.cfm. [Accessed 6 December 2013].

[29] C. V. Zegeer, R. J. Stewart, F. M. Council, D. W. Reinfurt and E. Hamilton, "Safety Effects of Geometric Improvements on Horizontal Curves," *Transportation Research Record,* vol. 1356, pp. 11-19, 1992.

[30] S.-P. Miaou and H. Lum, "A Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements," *Transportation Research Record,* vol. 1407, pp. 11-23, 1993.

[31] S. Dissanayake and N. Amarasingha, "Effects of Geometric Design Features on Truck Crashes on Limited-Access Highways," Mid-America Transportation Center, 2012.

[32] M. Poch and F. Mannering, "Negative Binomial Analysis of Intersection-accident Frequencies," *Journal of Transportation Engineering,* vol. 2, no. 122, pp. 105-113, 1996.

[33] L.-Y. Chang and W.-C. Chen, "Data Mining of Tree-based Models to Analyze Freeway Accident Frequency," *Journal of Safety Research ,* vol. 36, no. 4, pp. 365-375, 2005.

[34] S. Hallmark, N. Hawkins and O. Smadi, "Low Cost Strategies to Reduce Speed and Crashes on Curves," Iowa State University, Ames, 2007.

[35] S. Hallmark, Y. Qiu, N. Hawkins and O. Smadi, "Crash Modification Factors for Dynamic Speed Feedback Signs on Rural Curves," *Journal of Transportation Technologies,* vol. 5, pp. 9-23, 2015.

[36] W. H. Schneider, K. Zimmerman, D. Van Boxel and S. Vavilikolanu, "Bayesian analysis of the effect of horizontal curvature on truck crashes using training and validation data sets," *Transportation Research Record: Journal of the Transportation Research Board ,* vol. 2096, no. 1, pp. 41-46, 2009.

[37] W. H. Schneider, P. T. Savolainen and D. N. Moore, "Effects of horizontal curvature on single-vehicle motorcycle crashes along rural two-lane highways," *Transportation Research Record: Journal of the Transportation Research Board ,* vol. 2194 , no. 1, pp. 91-98, 2010.

[38] S. M. Easa and Q. C. You, "Collision Prediction Models for Three-dimensional Two-lane Highways: Horizontal Curves," *Journal of the Transportation Research Board,* no. 2092, pp. 48-56, 2009.

[39] G. Khan, A. R. Bill, M. V. Chitturi and D. A. Noyce, "Safety Evaluation of Horizontal Curves on Rural Undivided Roads," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2386, no. 1, pp. 147-157, 2013.

[40] S. P. Washington, M. G. Karlaftis and . F. L. Mannering, Statistical and Econometric Methods for Transportation Data Analysis, CRC press, 2010.

[41] A. Agresti, An Introduction to Categorical Data Analysis, Hoboken, New Jersey: Wiley Interscience, 2007.

[42] J. M. Hilbe, Negative Binomial Regression, New York: Cambridge University Press, 2011.

[43] D. Lord and L. F. Miranda-Moreno, "Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter of Poisson-gamma Models for Modeling Motor Vehicle Crashes: a Bayesian Perspective," *Safety Science,* vol. 5, no. 46, pp. 751-770, 2008.

[44] D. Lord, S. P. Washington and J. N. Ivan, "Further Notes on the Application of Zero-inflated Models in Highway Safety," *Accident Analysis & Prevention,* vol. 1, no. 39, pp. 53-57, 2007.

[45] M. Hosseinpour, A. S. Yahaya and A. F. Sadullah, "Exploring the effects of roadway characteristics on the frequency andseverity of head-on crashes: Case studies from Malaysian FederalRoads," *Accident Analysis and Prevention,* vol. 62, pp. 209-222, 2014.

[46] C. A. Plaxico, "Recommended Guidelines for Curb and Curb-Barrier Installations," Transportation Research Board, Washington D. C., 2005.

[47] H. F. Huang, R. . J. Schneider, C. V. Zegeer, A. J. Khattak, V. . J. Amerlynck and K. J. Lacy, "Identification of Severe Crash Factors and Countermeasures in North Carolina," North Carolina Department of Transportation, Raleigh, NC, 2001.

[48] R. Schneider and J. Stefanich, "Wisconsin Pedestrian and Bicycle Crash Analysis: 2011-2013," Wisconsin Department of Transportation (WisDOT), Milwaukee, 2015.

[49] K. K. Dixon and C. M. Monsere, "Highway Safety Investigation Manual for the Oregon Department of Transportation," Oregon Department of Transportation, Salem, 2011.

[50] A. Matteson and D. Blower , "Evaluation of California Crash Data Reported to MCMIS Crash File," Center for National Truck and Bus Statistics, Ann Arbor, MI, 2005.

[51] S. L. Hallmark, Y. Qiu, N. Hawkins and O. Smadi, "Crash Modification Factors for Dynamic Speed Feedback Signs on Rural Curves," *Journal of Transportation Technologies,* vol. 5, p. 9, 2015.

[52] Highway Safety Manual, Washington DC: AASHTO, 2010.

[53] "Bureau of Local Roads & Streets Manual," Illinois Department of Transportation, 2006.

[54] "Institute for Digital Research and Education," University of California, Los Angeles (UCLA), [Online]. Available: http://www.ats.ucla.edu/stat/spss/. [Accessed 29 April 2015].

[55] K. Kapetanakis, "Investigation of the Impact of Junctions on the Frequency of Road Crashes in Urban Areas," MS Thesis, Florida Atlantic University, Boca Raton, 2014.

[56] A. K. Sharma, V. S. Landge, and T. K. Rao, "Access Density and Standard Deviation of Speed as Contributing Factors for Accident of High Speed Cars," *International Journal of Computer Applications*, 2014.

[57] C. V. Zegeer, et al., "Cost-effective Geometric Improvements for Safety Upgrading of Horizontal Curves," Federal Highway Administration Report No. FHWA-RD-90-021, 1991.

[55] . M. S. Obeidat, M. J. Rys, E. R. Russell and A. Gund, "Determining Cost-Effective Policy for Visibility of Overhead Guide Signs on Highways," *Journal of the Transportation Research Forum,* vol. 53, pp. 45-58, 2014.

[56] M. S. Obeidat, "Modeling of Guide Sign Illumination and Retroreflectivity to Improve Driver's Visibility and Safety," Kansas State University, Manhattan, KS, 2015.

[57] E. R. Russell, M. J. Rys and T. S. Brin , "US Experience with Centerline Rumble Strips on Two-Lane Roads: Pattern Research and North American," in *Mid-Continent Transportation Research Symposium*, Ames, IA, 2003.

[58] L. W. Gardner, M. J. Rys and E. R. Russell, "Comparison of Football Shaped Rumble Strips versus Rectangular Rumble Strips," Kansas Department of Transportation, Topeka, KS, 2007.

[59] M. J. Rys, L. W. Gardner and E. R. Russell, "Evaluation of Football Shaped Rumble Strips Versus Rectangular Rumble Strips," *Journal of the Transportation Research Forum,* vol. 47, no. 2, pp. 41-54, 2008.

[60] V. A. W. J. Marchau, R. E. C. M. Van der Heijden and E. J. E. Molin, "Desirability of Advanced Driver Assistance from Road Safety Perspective: the Case of ISA," *Safety Science,* vol. 43, no. 1, pp. 11-27, 2005.

[61] I. M. Dias, "Work Zone Crash Analysis and Modeling to Identify Factors Associated with Crash Severity and Frequency," Kansas State University, Manhattan, KS, 2015.

[62] "AutoCAD Release 2013 (). Copyright© 2012-2014 by Autodesk Inc., San Rafael, CA.".

[63] "SAS Software Release 9.4 (). Copyright© 2010-2015 by SAS Institute Inc., Cary, NC.".

[64] "R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/".

[65] "Pages/CrashReductionFactors: AASHTO," [Online]. Available: http://scohts.transportation.org/Pages/CrashReductionFactors(CRF).aspx. [Accessed 5 December 2013].

[66] "In-Road Warning Lights: TAPCO," [Online]. Available: http://www.tapconet.com/solar-led-division/in-road-warning-lights. [Accessed 19 January 2013].

[67] O. Smadi, "Sequential Dynamic Curve Warning System (SDCWS): Product Safety Performance Evaluation," FHWA, 2011.

[68] J. Britton, "FHWA Select Dorman Varitext VATCS to Feature on National Traffic Safety Study of Dynamic Curve Warning Vehicle Activated Signs," *IMSA Journal,* pp. 46-50, 2009.

[72] J.-K. Kim, Y. Wang, and G. F. Ulfarsson, "Modeling the Probability of Freeway Rear-End Crash Occurrence," *Journal of transportation engineering*, vol. 133, no. 1, pp. 11-19, Jan. 2007.

Appendix A. Summary of Prediction Model Studies

Table A.1 Studies related to crash prediction models at horizontal curve sections

Source	Explanatory Variables	Dependent Variable	Sample Size	Statistical Method Used	Summary
Schneider et al. (2009)	Shoulder width, horizontal curve radius, curve length, passenger vehicle ADT, truck ADT, degree of curvature	Truck crashes	15390 crash records (2002–2006)	NB regression model with full Bayes methods for improving model performance	The objective of this paper was to develop an NB regression model to examine the impact of rural two-lane horizontal curves on non-intersection truck-related crashes. Traffic volume and horizontal curvature influenced truck crash occurrences at horizontal curve sections
Schneider et al. (2010)	Lane width, overall surface width, posted speed limit, additional land use categories (e.g., population density), ADT, segment length, curve radius, shoulder width	Single-vehicle motorcycle crashes	30379 roadway segments. Single-vehicle motorcycle crashes: 225 from 2002 to Spring 2008	NB model at first step to examine the impact of each variable. Second step includes implementation of a full Bayes methodology, resulting in an NB model with posterior distributions of parameters.	The objective of this study was to investigate the impact of roadway geometry at horizontal curve sections on single-vehicle motorcycle crash frequency by implementing a Bayesian technique to improve the frequency estimation. Curve length and radius, shoulder width, and ADT significantly influenced single-vehicle motorcycle crash frequency at horizontal curve sections
Hallmark et al. (2015)	AADT, section length, season, sign type, posted speed limit, curve advisory speed, differences between speed limit and advisory speed, radius	Total crashes for both directions, total crashes for direction of the sign, total single-vehicle crashes, single-vehicle crashes in the direction of the sign	624 observations for control sites and 492 observations for treatment sites	Full Bayes modeling methodology was utilized to develop crash modification factor (CMF).	This study investigated the effectiveness of "Dynamic Speed Feedback Sign Systems" (DSFS) at rural horizontal curves.

Table A.1 Studies related to crash prediction models at horizontal curve sections (Cont'd)

Source	Explanatory Variables	Dependent Variable	Sample Size	Statistical Method used	Summary
Khan et al. (2013)	Curve radius (R), Curve length (L), log AADT, posted speed, left and right shoulder width and type, average IRI, pavement surface age and type, upstream tangent (0–600 ft., 601–1200 ft., and 1201–2600 ft.), truck percentage, travel way width, difference between posted speed and advisory speed, presence of curve-related signs	Number of crashes at horizontal curves Number of KAB crashes	20842 Horizontal curves on undivided roadways	Quasi-Poisson NB regression	This study developed a crash prediction model for crashes, considering crashes separately for each direction on undivided roadways. They found effective variables were for each type of crash, but most considered variables were not effective. They concluded that curves with radii greater than 2500 ft. had fewer crashes than other curves, so they excluded curves with radii greater than 2500 ft.
Hallmark (2007)	Number of lanes, lane width, shoulder width and type, speed limit, pavement type and condition, presence and location of street lighting, grade, horizontal curve radius, degree of curvature, superelevation, sight distance, presence and characteristics of spirals, density of curves upstream for number of curves per mile, length of connection tangent section, any feature that may influence driver expectations and curve approach speed, location and type of signage before and within the curve (e.g., location of speed reduction zones, chevrons, etc.), speed, volume,	Speed change Crash frequency		Before and after study Bayesian or generalized linear regression models	The objective of the study was to investigate the effect of dynamic curve signs on speed reduction and crash occurrence at horizontal curve sections.

Table A.2 Factors used in crash prediction models at rural highway sections

Source	Explanatory Variables	Dependent Variables	Sample Size	Statistical Methods	Effective Explanatory Variables	Summary
Chang and Chen (2005)	Highway geometric design information includes number of lanes, HC, vertical grade (VG), and shoulder width, traffic information includes ADT of various vehicle types, peak hour factors, and traffic distribution over lanes. Weather information was obtained from the annual report of climatological data, which records detailed weather information of cities and towns along National Freeway 1, including pressure, temperature, humidity, precipitation, wind speed, and cloudiness.	Crash frequency from 2001 to 2002	1072 fatal and injury crashes and 1484 highway sections	CART NB regression	Degree of curvature, ADT, heavy vehicle ADT, VG, annual precipitation, precipitation day	The objective of this study was to investigate whether CART can be used to analyze the relationship between risk factors and crashes. CART was used to predict crash frequency for roadway sections. Development of CART consisted of three steps: tree growing, creation of a sequence of simpler trees by cutting off increasingly important nodes (i.e., pruning), and selection of the right tree from the pruned trees. Selected sections had lengths of 1 km. In order to compare predictions between the CART model and the statistical model, collected data were also randomly divided into two subsets: one for training and the other for testing. The number of cases used for model training and testing was 1,113 (75% of total observations) and 371, respectively. A chi-squared test showed that accident frequency distributions between the two sub-samples were not significantly different.
Dissanayake and Amarasingha (2012)	Section length, speed limit, median width, functional class, AADT, AADT of heavy vehicles, right rumble strips, inside rumble strips, right shoulder width, inside shoulder width, HC, VG, number of lanes	Truck crashes from 2005 to 2010	7273 segments	Poisson regression model NB regression model	Length of section, number of lanes, HC, VG, AADT, truck percentage, and inside shoulder width	The objective of the study was to find the relationship between large truck crash probability and traffic and geometric characteristics.

Table A.2 Factors used in crash prediction models at rural highway sections (Cont'd)

Source	Explanatory Variables	Dependent Variables	Sample Size	Statistical Methods	Effective Explanatory Variables	Summary
Miaou and Lum (1992)	Section length, truck miles or truck exposure, dummy intercept, dummy variables for years 1986–1989, AADT per lane, HC, LHC, VG, LVG, deviation of paved inside shoulder width, percent trucks, interaction between HC and LHC, and interaction between VG and LVG	Number of trucks involved in accidents	8263 road sections, 1643 crashes in which trucks were involved during five years	Poisson regression model	HC, LHC, VG, LVG, and paved inside shoulder width	This study used data of Highway Safety Information System (HSIS) of the state of Utah to develop crash prediction models for truck accident involvements. They developed models with different variables and compared models by comparing the results of developed models.
Zegeer et al. (1993)	ADT, curve length, degree of curvature, total surface width, presence of spiral transition, superelevation, RHR, roadside recovery distance	Total number of crashes	1039 horizontal curves	Linear regression model	Degree of curvature, shoulder width, presence of spirals, superelevation	The objective of this research was to evaluate the relationship between horizontal curve features and safety levels. The authors also quantified the effects of curve flattening, curve widening, the addition of a spiral, improved deficient superelevation, and roadside clearing for crash occurrence.

Table A.2 Factors used in crash prediction models at rural highway sections (Cont'd)

Source	Explanatory Variables	Dependent Variables	Sample Size	Statistical Methods	Effective Explanatory Variables	Summary
Hosseinpour et al. (2014)	Posted speed, shoulder width, HC, terrain type, heavy vehicle traffic, land use, side friction factor, presence of median, and access points, segment length, number of lanes	Number of head-on crashes Crash severity	448 segments	For crash frequency: Poisson, standard NB, random-effect NB, hurdle Poisson, hurdle NB, ZIP, and ZINB models were used. For crash severity, a random-effect generalized ordered probit model (REGOPM) was used.	HC, terrain type, heavy vehicle, heavy vehicle, access points, posted speed, shoulder width	This study used head-on crashes data for 4 years (from 2007 to 2011) on 448 segments of five federal roadways in Malaysia. They modeled crash frequency and crash severity. To model crash frequency, they found random-effect NB more accurately fit the head-on crashes compared to the other regression methods.

Appendix B. R Codes for Creating a Dataset of KDOT Horizontal Curves Inventory

Adding AADTs to Dataset

```
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("file1.txt",header=TRUE,sep="\t")
f2=read.csv("file2.txt",header=TRUE,sep="\t")
f3=read.csv("file3.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
Mf2=as.matrix(f2)
Mf3=as.matrix(f3)
p=mat.or.vec(5343,12)
for (i in 1:5343) {
                for (j in 1:8308) {
                                if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))     for (k
in 1:12) {p[i,k]=Mf1[j,45+k] }
                                        }
                        }
        for (x in 1:5343) {
                        for (y in 1:12290) {
                                        if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]))
        for (k in 1:12) {p[x,k]=Mf2[y,45+k] }
                                        }
                        }
        for (x in 1:5343) {
                        for (y in 1:12358) {
                                        if ((Mmain[x,2]==Mf3[y,1]) & (Mmain[x,3]==Mf3[y,2]) &
(Mmain[x,4]==Mf3[y,3]) & (Mmain[x,7]<=Mf3[y,6]) & (Mmain[x,8]>=Mf3[y,6]) &
(Mmain[x,10]==Mf3[y,9]) & (Mmain[x,12]==Mf3[y,31]) & (Mmain[x,13]==Mf3[y,32]))
        for (k in 1:12) {p[x,k]=Mf3[y,45+k] }
                                        }
                        }
        write.csv(p, file="p.csv")
        main=read.csv("main.txt",header=TRUE,sep="\t")
        Mmain=as.matrix(main)
        f1=read.csv("file1.txt",header=TRUE,sep="\t")
        f2=read.csv("file2.txt",header=TRUE,sep="\t")
        f3=read.csv("file3.txt",header=TRUE,sep="\t")
        Mf1=as.matrix(f1)
        Mf2=as.matrix(f2)
```

```
	Mf3=as.matrix(f3)
	p=mat.or.vec(5343,3)
	for (i in 1:5343) {
			for (j in 1:8308) {
					if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))	for (k
in 1:3) {p[i,k]=Mf1[j,10+k] }
					}
			}
	for (x in 1:5343) {
			for (y in 1:12290) {
					if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]))
	for (k in 1:3) {p[x,k]=Mf2[y,10+k] }
					}
			}
	for (x in 1:5343) {
			for (y in 1:12358) {
					if ((Mmain[x,2]==Mf3[y,1]) & (Mmain[x,3]==Mf3[y,2]) &
(Mmain[x,4]==Mf3[y,3]) & (Mmain[x,7]<=Mf3[y,6]) & (Mmain[x,8]>=Mf3[y,6]) &
(Mmain[x,10]==Mf3[y,9]) & (Mmain[x,12]==Mf3[y,31]) & (Mmain[x,13]==Mf3[y,32]))
	for (k in 1:3) {p[x,k]=Mf3[y,10+k] }
					}
			}
write.csv(p, file="p.csv")
```

Adding Other Variables to the Dataset

```
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("f1.txt",header=TRUE,sep="\t")
f2=read.csv("f2.txt",header=TRUE,sep="\t")
f3=read.csv("f3.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
Mf2=as.matrix(f2)
Mf3=as.matrix(f3)
p=mat.or.vec(5343,22)
for (i in 1:5343) {
		for (j in 1:23997) {
				if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))	for (k
in 1:22) {p[i,k]=Mf1[j,10+k] }
					}
```

```
                    }
          for (x in 1:5343) {
                    for (y in 1:31007) {
                              if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]))
          for (k in 1:22) {p[x,k]=Mf2[y,10+k] }
                              }
                    }
          for (x in 1:5343) {
                    for (y in 1:29452) {
                              if ((Mmain[x,2]==Mf3[y,1]) & (Mmain[x,3]==Mf3[y,2]) &
(Mmain[x,4]==Mf3[y,3]) & (Mmain[x,7]<=Mf3[y,6]) & (Mmain[x,8]>=Mf3[y,6]) &
(Mmain[x,10]==Mf3[y,9]) & (Mmain[x,12]==Mf3[y,31]) & (Mmain[x,13]==Mf3[y,32]))
          for (k in 1:22) {p[x,k]=Mf3[y,10+k] }
                              }
                    }
write.csv(p, file="p.csv")
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("file1.txt",header=TRUE,sep="\t")
f2=read.csv("file2.txt",header=TRUE,sep="\t")
f3=read.csv("file3.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
Mf2=as.matrix(f2)
Mf3=as.matrix(f3)
p=mat.or.vec(5343,19)
for (i in 1:5343) {
               for (j in 1:8308) {
                         if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))     for (k
in 1:19) {p[i,k]=Mf1[j,13+k] }
                              }
                    }
          for (x in 1:5343) {
                    for (y in 1:12290) {
                              if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]))
          for (k in 1:19) {p[x,k]=Mf2[y,13+k] }
                              }
                    }
          for (x in 1:5343) {
                    for (y in 1:12358) {
```

```
                                        if ((Mmain[x,2]==Mf3[y,1]) & (Mmain[x,3]==Mf3[y,2]) &
(Mmain[x,4]==Mf3[y,3]) & (Mmain[x,7]<=Mf3[y,6]) & (Mmain[x,8]>=Mf3[y,6]) &
(Mmain[x,10]==Mf3[y,9]) & (Mmain[x,12]==Mf3[y,31]) & (Mmain[x,13]==Mf3[y,32]))
        for (k in 1:19) {p[x,k]=Mf3[y,13+k] }
                                        }
                        }
write.csv(p, file="p.csv")
```

Adding Crashes Based on their Severities to the Dataset

```
F11=read.csv("F11.txt",header=TRUE,sep="\t") ##ino  bara tarfie file neveshtam
F11=as.matrix(F11)
MF1=read.csv("MF1.txt",header=TRUE,sep="\t")
MF1=as.matrix(MF1)
cm=mat.or.vec(14418,2)
for (i in 1:10) {
                                        cm[i,1]=F11[i,1]
                for (j in 1:10) {
                                if ((F11[i,1]=MF1[j,1])
                                        if (cm[i,2]=0 cm[i,2]=MF1[j,2])
                }                       ifelse (if (MF1[j,2]<cm[i,2] cm[i,2]=MF1[j,2]))

write.csv(cm, file="cm.csv")
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("file1.txt",header=TRUE,sep="\t")
f2=read.csv("file2.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
Mf2=as.matrix(f2)
p=mat.or.vec(5343,5)
## For first file
for (i in 1:5343) {
                for (j in 1:15992) {
                                if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))     for (k
in 1:5) { if (Mf1[j,35]==k) p[i,k]=p[i,k]+1 }
                                        }
                        }
## For next File
for (x in 1:5343) {
                for (y in 1:16330) {
                                if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]))
        for (z in 1:5) { if (Mf2[y,35]==z) p[x,z]=p[x,z]+1 }
```

```
                                        }
                                }
write.csv(p, file="p.csv",sep="\t")
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("f14.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
ttcrsh=mat.or.vec(5343,5)
for (i in 1:5343) {
                for (j in 1:13874) {
                                if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]))     for (k
in 1:5) { if (Mf1[j,45]==k) ttcrsh [i,k]= ttcrsh [i,k]+1 }
                                }
                }
write.csv(ttcrsh, file=" ttcrsh.csv")
```

Selecting Single-Vehicle Crashes Based on Their Severities

```
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("file1.txt",header=TRUE,sep="\t")
f2=read.csv("file2.txt",header=TRUE,sep="\t")
f3=read.csv("file3.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
Mf2=as.matrix(f2)
Mf3=as.matrix(f3)
p=mat.or.vec(5343,5)
## For first file
for (i in 1:5343) {
        for (j in 1:8308) {
                                        if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]) &
((as.POSIXct(Mf1[j,34], format="%m/%d/%Y"))>= (as.POSIXct("01/01/2010",
format="%m/%d/%Y"))) & ((as.POSIXct(Mf1[j,34], format="%m/%d/%Y"))<
(as.POSIXct("01/01/2015", format="%m/%d/%Y"))) & (as.numeric(Mf1[j,58])!=2) &
(as.numeric(Mf1[j,58])!=3) & (as.numeric(Mf1[j,58])!=6) & (as.numeric(Mf1[j,58])!=7))
        for (k in 1:5) { if (Mf1[j,35]==k) p[i,k]=p[i,k]+1 }
                                        }
                                }
## For next File
for (x in 1:5343) {
                for (y in 1:12290) {
                                if ((Mmain[x,2]==Mf2[y,1]) & (Mmain[x,3]==Mf2[y,2]) &
(Mmain[x,4]==Mf2[y,3]) & (Mmain[x,7]<=Mf2[y,6]) & (Mmain[x,8]>=Mf2[y,6]) &
```

```
(Mmain[x,10]==Mf2[y,9]) & (Mmain[x,12]==Mf2[y,31]) & (Mmain[x,13]==Mf2[y,32]) &
((as.POSIXct(Mf2[y,34], format="%m/%d/%Y"))>= (as.POSIXct("01/01/2010",
format="%m/%d/%Y"))) & ((as.POSIXct(Mf2[y,34], format="%m/%d/%Y"))<
(as.POSIXct("01/01/2015", format="%m/%d/%Y"))) & (as.numeric(Mf2[y,58])!=2) &
(as.numeric(Mf2[y,58])!=3) & (as.numeric(Mf2[y,58])!=6) & (as.numeric(Mf2[y,58])!=7))
        for (z in 1:5) { if (Mf2[y,35]==z) p[x,z]=p[x,z]+1 }
                                        }
                        }
## For next File
for (x in 1:5343) {
                for (y in 1:12358) {
                                if ((Mmain[x,2]==Mf3[y,1]) & (Mmain[x,3]==Mf3[y,2]) &
(Mmain[x,4]==Mf3[y,3]) & (Mmain[x,7]<=Mf3[y,6]) & (Mmain[x,8]>=Mf3[y,6]) &
(Mmain[x,10]==Mf3[y,9]) & (Mmain[x,12]==Mf3[y,31]) & (Mmain[x,13]==Mf3[y,32]) &
((as.POSIXct(Mf3[y,34], format="%m/%d/%Y"))>= (as.POSIXct("01/01/2010",
format="%m/%d/%Y"))) & ((as.POSIXct(Mf3[y,34], format="%m/%d/%Y"))<
(as.POSIXct("01/01/2015", format="%m/%d/%Y"))) & (as.numeric(Mf3[y,58])!=2) &
(as.numeric(Mf3[y,58])!=3) & (as.numeric(Mf3[y,58])!=6) & (as.numeric(Mf3[y,58])!=7))
        for (z in 1:5) { if (Mf2[y,35]==z) p[x,z]=p[x,z]+1 }
                                        }
                        }
write.csv(p, file="p.csv")
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f1=read.csv("SVRTCRCT.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
svcrsh=mat.or.vec(5343,5)
## For first file
for (i in 1:5343) {
                for (j in 1:14418) {
                                if ((Mmain[i,2]==Mf1[j,1]) & (Mmain[i,3]==Mf1[j,2]) &
(Mmain[i,4]==Mf1[j,3]) & (Mmain[i,7]<=Mf1[j,6]) & (Mmain[i,8]>=Mf1[j,6]) &
(Mmain[i,10]==Mf1[j,9]) & (Mmain[i,12]==Mf1[j,31]) & (Mmain[i,13]==Mf1[j,32]) &
((as.POSIXct(Mf1[j,34], format="%m/%d/%Y"))>= (as.POSIXct("01/01/2010",
format="%m/%d/%Y"))) & ((as.POSIXct(Mf1[j,34], format="%m/%d/%Y"))<
(as.POSIXct("01/01/2015", format="%m/%d/%Y"))) & (as.numeric(Mf1[j,46])!=2) &
(as.numeric(Mf1[j,46])!=3) & (as.numeric(Mf1[j,46])!=6) & (as.numeric(Mf1[j,46])!=7))
        for (k in 1:5) { if (Mf1[j,45]==k) svcrsh [i,k]= svcrsh [i,k]+1 }
                                }
                }
write.csv(svcrsh, file=" svcrsh.csv")
```

Adding Severe Crashes (K and A levels) to the Randomly Selected Horizontal Curves for Data Analysis and Validation

```
main=read.csv("main.txt",header=TRUE,sep="\t")
```

```
Mmain=as.matrix(main)
f1=read.csv("fda.txt",header=TRUE,sep="\t")
Mf1=as.matrix(f1)
KAan=mat.or.vec(4267,2)
for (i in 1:4267) {
                for (j in 1:5343) {if (Mmain[j,1]==Mf1[i,1]) {p[i,1]=Mf1[i,1] ; KAan
[i,2]=Mmain[j,4]} }
                }
write.csv(KAan, file=" KAan.csv")
main=read.csv("main.txt",header=TRUE,sep="\t")
Mmain=as.matrix(main)
f2=read.csv("fdv.txt",header=TRUE,sep="\t")
Mf2=as.matrix(f2)
KAva=mat.or.vec(1067,2)
for (i in 1:1067) {
                for (j in 1:5343) {if (Mmain[j,1]==Mf2[i,1]) {q[i,1]=Mf2[i,1] ; KAva
[i,2]=Mmain[j,4]} }
                }
write.csv(KAva, file=" KAva.csv")
```

Appendix C. SAS Codes for Data Analysis

Codes for Analyzing Crashes at 221 Horizontal Curves on Rural Two-lane Highways

```
data Hcurves;
	infile '/folders/myfolders/sasuser.v94/221_HCs.csv' dlm=',' firstobs=2;
	input curve_number $ road_name $ AADT logAADT HVPct Radius DoC Tangent_l_1
Tangent_l_2 Posted_Spd Adv_spd Diff_PS_AS F_crash I_crash PDO_crash Total_Crash
Rumble_S PavedSh UnpavedSh PSH_t1 PSH_t2 UPSH_t1 UPSH_t2 UPSH_t3 Lt_less_60L
t_bw600_1200 Lt_bw1200_2600 Lt_greater_2600 Long_Lt Short_Lt;
Run;

proc means data=Hcurves;
	var AADT logAADT HVPct Radius DoC Tangent_l_1 Tangent_l_2 Posted_Spd
Adv_spd Diff_PS_AS F_crash I_crash PDO_crash Total_Crash Long_Lt Short_Lt;
run;
data Hcurves;
	infile '/folders/myfolders/sasuser.v94/221_HCs.csv' dlm=',' firstobs=2;
	input curve_number $ road_name $ AADT logAADT HVPct Radius DoC Tangent_l_1
Tangent_l_2 Curve_length Posted_Spd Adv_spd Diff_PS_AS F_crash I_crash PDO_crash
Total_Crash Rumble_S PavedSh UnpavedSh PSH_t1 PSH_t2 UPSH_t1 UPSH_t2 UPSH_t3
Lt_less_600 Lt_greater_2600 Lt_bw600_1200 Lt_bw1200_2600  Long_Lt Short_Lt;
Run;
proc means data=Hcurves;
	var AADT logAADT HVPct Radius DoC Tangent_l_1 Tangent_l_2 Curve_length
Posted_Spd Adv_spd Diff_PS_AS F_crash I_crash PDO_crash Total_Crash Long_Lt Short_Lt;
run;
proc means data=Hcurves;
	var AADT;
run;

* Frequenies;
Proc freq data=Hcurves;
	tables F_crash I_crash PDO_crash Total_crash;
run;
*Poisson;
Proc genmod data=Hcurves;
	model Total_crash= logAADT HVPct DoC Adv_Spd Diff_PS_AS/ dist=poisson
link=log;
run;

*Poisson2;
Proc genmod data=Hcurves;
	model Total_crash= logAADT HVPct DoC Adv_Spd Diff_PS_AS/ dist=poisson
link=log obstats;
run;
```

```
*Poisson3;
Proc genmod data=Hcurves;
	class Rumble_S;
	class PavedSh;
	class UnpavedSh;
	class PSH_t1;
	class PSH_t2;
	class UPSH_t1;
	class UPSH_t2;
	class UPSH_t3;
	class Lt_less_600;
	class Lt_greater_2600;
	class Lt_bw600_1200;
	class Lt_bw1200_2600;
	model Total_crash= logAADT HVPct DoC Posted_Spd Adv_Spd Diff_PS_AS
Curve_length Long_Lt Short_Lt Rumble_S PavedSh UnpavedSh PSH_t1 PSH_t2 UPSH_t1
UPSH_t2 UPSH_t3 Lt_less_600 Lt_greater_2600 Lt_bw600_1200 Lt_bw1200_2600/
dist=poisson link=log;
run;
*Poisson4;
Proc genmod data=Hcurves;
	model Total_crash= logAADT HVPct DoC Diff_PS_AS/ dist=poisson link=log;
run;
data Hcurves;
	infile '/folders/myfolders/sasuser.v94/221_HCs.csv' dlm=',' firstobs=2;
	input curve_number $ road_name $ AADT logAADT HVPct Radius DoC Tangent_l_1
Tangent_l_2 Curve_length Posted_Spd Adv_spd Diff_PS_AS F_crash I_crash PDO_crash
Total_Crash Rumble_S PavedSh UnpavedSh PSH_t1 PSH_t2 UPSH_t1 UPSH_t2 UPSH_t3
Lt_less_600 Lt_greater_2600 Lt_bw600_1200 Lt_bw1200_2600  Long_Lt Short_Lt
Ln_Long_Lt Ln_Short_Lt;
Run;
proc means data=Hcurves;
	var AADT logAADT HVPct Radius DoC Tangent_l_1 Tangent_l_2 Curve_length
Posted_Spd Adv_spd Diff_PS_AS F_crash I_crash PDO_crash Total_Crash Long_Lt Short_Lt
Ln_Long_Lt Ln_Short_Lt;
run;
proc means data=Hcurves;
	var AADT;
run;

* Frequenies;
Proc freq data=Hcurves;
	tables F_crash I_crash PDO_crash Total_crash;
run;
*Poisson5;
Proc genmod data=Hcurves;
```

```
	model Total_crash= logAADT HVPct DoC Diff_PS_AS Ln_Long_Lt/ dist=poisson
link=log;
run;

*NB1;
Proc genmod data=Hcurves;
	model Total_crash= logAADT HVPct DoC Ln_Long_Lt/ dist=negbin link=log;
run;

*Poisson6;
Proc genmod data=Hcurves;
	class Lt_greater_2600;
	model Total_crash= logAADT HVPct DoC Diff_PS_AS Ln_Long_Lt Lt_greater_2600/
dist=poisson link=log;
run;
```

Codes for Analyzing Randomly Selected 4,267 Horizontal Curves of KDOT Curve Inventory

```
proc import datafile="C:\Users\hmomeni.ENGG\Desktop\Data_Validation\frdv021516.csv" out=mydata dbms=csv
replace;
  getnames=yes;
run;
Proc corr data=mydata;
	var AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI Sur_w RS Grade RLSH ISH LSH
RSH_W ISH_W LSH_W;
run;
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\nb_pvalue_consrd_rslts.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref
ref="1") LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref
ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RLSH ISH
RSH_W/ type3 dist= negbin link=log;
run;
ODS RTF CLOSE;


ODS RTF FILE='c:\desktop\Hojr\CategoricalVarChisq.doc';

Proc genmod data=mydata;

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI Sur_w RS
RLSH ISH RSH_W ISH_W/ type3 dist= negbin link=log;
run;
ODS RTF CLOSE;
/*I removed the zeromodel vars when they were insig*/
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\zim_pvlu_cnsrd_rslts.doc';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");
```

```
  model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI RS RLSH/dist=zinb;
  zeromodel AADT_th Curve_Length AS IRI RSH_W/link = logit;
run;
ODS RTF CLOSE;

/*I removed the zeromodel vars when they were insig*/
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\zip_rslts.doc';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

  model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI RS RLSH ISH RSH_W
ISH_W LSH_W/dist=zip;
  zeromodel AADT_th Curve_Length IRI RLSH RSH_W/link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_nb_most_vars_cnsrd_rslts.doc';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

  model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RS RLSH RSH_W
ISH_W/ type3 dist= negbin link=log;
 run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_nb_10%sig_vars_cnsrd_rslts.doc';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

  model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RS RLSH RSH_W/ type3
dist= negbin link=log;
 run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_nb_5%sig_vars_cnsrd_rslts.doc';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

  model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RLSH RSH_W/ type3
dist= negbin link=log;
 run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_zi_10%sig_vars_cnsrd_rslts.doc';
'(the sign of the Surf_W is not good at all)';
proc genmod data=mydata;
  class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

  model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided Sur_w RLSH
RSH_W/dist=zinb;
  zeromodel AADT_th Curve_Length IRI RLSH RSH_W/link = logit;
```

```
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_zi_5%sig_vars_cnsrd_rslts.doc';
'the sign of the Surf_W is not good at all as well as RLSH vars';
proc genmod data=mydata;
 class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

 model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Sur_w RLSH RSH_W/dist=zinb;
 zeromodel AADT_th Curve_Length IRI RLSH RSH_W/link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_zip_all important vars_rslts.doc';

proc genmod data=mydata;
 class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

 model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided Sur_w RS
RLSH/dist=zip;
 zeromodel AADT_th Curve_Length IRI Sur_w RS RLSH RSH_W/link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svcr_zip_10%sig_Vars consrd_rslts.doc';

proc genmod data=mydata;
 class Divided(param=ref ref="0") RS(param=ref ref="0") RLSH(param=ref ref="1") ISH(param=ref ref="1")
LSH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1") LSH_W(param=ref ref="1");

 model Severe_Crashes = HVPct Curve_Length D_o_C PS DiffPSAS Divided Sur_w RS RLSH/dist=zip;
 zeromodel AADT_th Curve_Length IRI Sur_w RS RLSH RSH_W/link = logit;
run;
ODS RTF CLOSE;

proc import datafile="C:\Users\hmomeni.ENGG\Desktop\Data_Validation\03032016\frdv2.csv" out=mydata
dbms=csv replace;
  getnames=yes;
run;
Proc corr data=mydata;
	var AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI Sur_w RS Grade RLSH ISH
RSH_W ISH_W LSH_W RL_SH;
run;
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\KAcrsh_NB_020262016_sigvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model KA_Crashes = AADT_th HVPct Curve_Length D_o_C PS RS/ type3 dist= negbin link=log;
run;
ODS RTF CLOSE;
```

```
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\KAcrsh_ZIP_03032016_sigvars.doc';
Proc genmod data=mydata;
        class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

        model KA_Crashes = AADT_th Curve_Length D_o_C PS Divided RS/dist=zip;
  zeromodel AADT_th/link = logit;
run;
ODS RTF CLOSE;

Proc genmod data=mydata;
        class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

        model KA_Crashes = AADT_th HVPct Curve_Length D_o_C PS Divided RS RL_SH ISH/dist=zinb;
  zeromodel AADT_th Curve_Length D_o_C Sur_w RL_SH ISH/link = logit;
run;



Proc genmod data=mydata;
        class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

        model KA_Crashes = AADT_th HVPct Curve_Length D_o_C PS Divided RS ISH/dist=zinb;
  zeromodel AADT_th Curve_Length D_o_C Sur_w RL_SH ISH/link = logit;
run;



Proc genmod data=mydata;
        class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

        model KA_Crashes = AADT_th HVPct Curve_Length D_o_C PS Divided RS/dist=zinb;
  zeromodel AADT_th  D_O_C/ link = logit;
run;
proc import datafile="C:\Users\hmomeni.ENGG\Desktop\Data_Validation\frdv022216.csv" out=mydata dbms=csv
replace;
   getnames=yes;
run;
Proc corr data=mydata;
        var AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI Sur_w RS Grade RLSH ISH
RSH_W ISH_W LSH_W RL_SH;
run;
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_NB_020262016_sigvars.doc';
Proc genmod data=mydata;
        class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

        model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RL_SH ISH
RSH_W/ type3 dist= negbin link=log;
run;
ODS RTF CLOSE;
```

```
ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_NB_020282016_sigvars_accepted signs.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS/ type3 dist= negbin link=log;
run;
ODS RTF CLOSE;



ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_ZIP_020272016_sigvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI RS RL_SH ISH RSH_W ISH_W/dist=zip;
 zeromodel AADT_th Curve_Length IRI RL_SH ISH RSH_W/link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_ZINB_020282016_SignVarsPLus_srfw_n_RS_accepted signs.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS sur_w IRI RS/dist=zinb;
 zeromodel AADT_th Curve_Length PS IRI RL_SH ISH /link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_ZINB_020282016_SignVars_acceptedSigns.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS sur_w IRI RS/dist=zinb;
 zeromodel AADT_th Curve_Length PS IRI RL_SH ISH /link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_ZINB_020282016_SignVars_acceptedSigns_corrected.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS IRI/dist=zinb;
```

```
  zeromodel AADT_th Curve_Length PS IRI RL_SH ISH /link = logit;
run;
ODS RTF CLOSE;


ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\ttlcrsh_ZINB_020272016_SignVars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Total_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided IRI RS RL_SH ISH
RSH_W ISH_W/dist=zinb;
  zeromodel AADT_th Curve_Length PS IRI RL_SH ISH RSH_W/link = logit;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_NB_020262016_sigvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS RS RSH_W/ type3 dist=
negbin link=log;
run;
ODS RTF CLOSE;

ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_NB_020262016_impvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RS RSH_W/
type3 dist= negbin link=log;
run;
ODS RTF CLOSE;



ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_ZIP_020272016_imp_vars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C DiffPSAS Divided RS RSH_W/dist=zip;
  zeromodel AADT_th Curve_Length PS RS RL_SH ISH RSH_W/link = logit;
run;
ODS RTF CLOSE;



ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_ZIP_020272016_sigvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C DiffPSAS RS/dist=zip;
```

```
  zeromodel AADT_th Curve_Length PS RL_SH ISH RSH_W/link = logit;
run;
ODS RTF CLOSE;


ODS RTF
FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_ZINB_020272016_Important_vars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS Divided RS ISH
RSH_W/dist=zinb;
  zeromodel AADT_th Curve_Length Divided RS/link = logit;
run;
ODS RTF CLOSE;



ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\svrcrsh_ZINB_020272016_sigvars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model Severe_Crashes = AADT_th HVPct Curve_Length D_o_C PS DiffPSAS RS/dist=zinb;
  zeromodel AADT_th Curve_Length Divided RS RSH_W/link = logit;
run;
ODS RTF CLOSE;


ODS RTF FILE='C:\Users\hmomeni.ENGG\Desktop\Data_Validation\KAcrsh_ZINB_03032016_SignVars.doc';
Proc genmod data=mydata;
	class Divided(param=ref ref="0") RS(param=ref ref="0") RL_SH(param=ref ref="1") ISH(param=ref
ref="1") RSH_W(param=ref ref="1") ISH_W(param=ref ref="1");

	model KA_Crashes = AADT_th HVPct Curve_Length D_o_C PS Divided RS/dist=zinb;
  zeromodel AADT_th D_O_C Curve_Length PS / link = logit;
run;
ODS RTF CLOSE;
```

Appendix D. Observed and Predicted Crashes for Randomly Selected Horizontal Curves for Data Validation

Table D.1 Sample of Comparison of 1,067 observed and predicted crashes for various crash groups and models

Number	Curve number	Total Crashes			Severe (KAB) Crashes				Severe (KA) Crashes			
		Observed	Predicted		Observed	Predicted			Observed	Predicted		
			NB	ZINB		NB	ZIP	ZINB		NB	ZIP	ZINB
1	1700	2	2	2	0	0	0	0	0	0	0	0
2	3301	2	6	3	0	1	1	1	0	0	0	0
3	3335	0	1	1	0	0	0	0	0	0	0	0
4	4861	0	0	0	0	0	0	0	0	0	0	0
5	2254	0	0	0	0	0	0	0	0	0	0	0
6	696	1	1	1	0	0	0	0	0	0	0	0
7	4179	0	1	1	0	0	0	0	0	0	0	0
8	3455	0	0	1	0	0	0	0	0	0	0	0
9	4424	2	3	3	0	0	1	0	0	0	0	0
10	4459	1	0	0	1	0	0	0	1	0	0	0
11	4806	1	1	1	0	0	0	0	0	0	0	0
12	3291	1	1	1	1	0	0	0	0	0	0	0
13	4368	1	0	0	0	0	0	0	0	0	0	0
14	3643	0	0	1	0	0	0	0	0	0	0	0
15	1254	0	0	0	0	0	0	0	0	0	0	0
16	2366	58	64	33	6	5	2	2	0	1	1	1
17	4775	1	1	1	0	0	0	0	0	0	0	0
18	1058	5	1	1	1	0	0	0	0	0	0	0
19	531	1	1	1	0	0	0	0	0	0	0	0
20	4786	0	0	0	0	0	0	0	0	0	0	0
21	552	0	0	0	0	0	0	0	0	0	0	0
22	5111	4	3	2	1	1	1	1	0	0	0	0
23	2738	2	1	1	0	0	0	0	0	0	0	0
24	3014	0	0	1	0	0	0	0	0	0	0	0
25	4602	0	1	1	0	0	0	0	0	0	0	0
26	499	0	0	0	0	0	0	0	0	0	0	0
27	3893	0	0	1	0	0	0	0	0	0	0	0
28	615	1	1	1	0	0	0	0	0	0	0	0
29	1053	0	0	0	0	0	0	0	0	0	0	0

. . . . . .

. . . .												
1041	2176	0	0	0	0	0	0	0	0	0	0	0
1042	231	0	0	0	0	0	0	0	0	0	0	0
1043	2143	1	3	1	0	1	1	0	0	0	0	0
1044	3545	0	1	1	0	0	0	0	0	0	0	0
1045	2283	2	12	7	0	2	1	2	0	0	0	0
1046	3353	0	0	0	0	0	0	0	0	0	0	0
1047	1692	0	0	0	0	0	0	0	0	0	0	0
1048	2455	0	0	0	0	0	0	0	0	0	0	0
1049	1563	1	0	0	1	0	0	0	1	0	0	0
1050	3977	2	0	1	0	0	0	0	0	0	0	0
1051	3234	0	0	0	0	0	0	0	0	0	0	0
1052	4440	1	0	0	0	0	0	0	0	0	0	0
1053	3323	1	0	1	1	0	0	0	0	0	0	0
1054	4987	1	0	0	1	0	0	0	0	0	0	0
1055	4268	3	1	1	0	0	0	0	0	0	0	0
1056	4818	0	0	0	0	0	0	0	0	0	0	0
1057	4557	0	1	1	0	0	0	0	0	0	0	0
1058	3691	0	0	0	0	0	0	0	0	0	0	0
1059	1055	1	3	3	0	1	1	1	0	0	0	0
1060	1231	0	0	1	0	0	0	0	0	0	0	0
1061	1723	0	0	1	0	0	0	0	0	0	0	0
1062	2491	0	1	1	0	0	0	0	0	0	0	0
1063	2033	0	1	1	0	0	0	0	0	0	0	0
1064	2189	0	4	3	0	1	1	1	0	0	0	0
1065	3420	0	1	1	0	0	1	0	0	0	0	0
1066	3443	1	0	1	0	0	0	0	0	0	0	0
1067	239	0	0	0	0	0	0	0	0	0	0	0
Sum		**982**	**1208**	**1123**	**187**	**119**	**233**	**158**	**56**	**1**	**14**	**33**