

2011

Systemic highway safety assessment: A general analysis of funding allocation and a specific study of the horizontal curve crash problem

Corey Douglas Bogenreif
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Civil and Environmental Engineering Commons](#)

Recommended Citation

Bogenreif, Corey Douglas, "Systemic highway safety assessment: A general analysis of funding allocation and a specific study of the horizontal curve crash problem" (2011). *Graduate Theses and Dissertations*. 11976.
<https://lib.dr.iastate.edu/etd/11976>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Systemic highway safety assessment: A general analysis of funding allocation and a specific study of the horizontal curve crash problem

by

Corey Douglas Bogenreif

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee

Reginald Souleyrette, Major Professor

Thomas Stout

Kelly Strong

Iowa State University

Ames, Iowa

2011

Copyright © Corey Douglas Bogenreif, 2011. All rights reserved.

TABLE OF CONTENTS

LIST OF FIGURES	v
LIST OF TABLES.....	viii
ACKNOWLEDGMENTS.....	ix
ABSTRACT	x
CHAPTER 1. GENERAL INTRODUCTION.....	1
1.1 INTRODUCTION	1
1.2 THESIS ORGANIZATION	2
1.3 REVIEW OF LITERATURE.....	2
CHAPTER 2. OPTIMIZING SAFETY FUND ALLOCATION.....	6
2.1 INTRODUCTION	6
2.2 REVIEW OF LITERATURE.....	8
2.3 DATA	10
2.3.1 Crash and roadway data.....	10
2.3.2 Data related to statewide highway safety projects	12
2.4 METHODOLOGY	12
2.4.1 Crash data classification	12
2.4.2 Safety project classification.....	13
2.4.3 Combined crash and safety project classifications.....	13
2.5 ANALYSIS AND RESULTS	14
2.5.1 Statewide crash data classification	14
2.5.2 Statewide allocation of safety funds.....	16
2.5.3 Combined crash data and safety project classification.....	19
2.5.4 Safety funding allocation relative to crash density.....	25
2.6 CONCLUSIONS AND RECOMMENDATIONS.....	29
CHAPTER 3. SYSTEMWIDE IDENTIFICATION OF HORIZONTAL CURVES AND GEOMETRY PARAMETERS	31

3.1 INTRODUCTION	31
3.2 REVIEW OF LITERATURE	33
3.3 DATA	33
3.3.1 Roadway data	33
3.3.2 Calculated curve data	34
3.3.3 As-built curve data	34
3.4 METHODOLOGY	35
3.4.1 Curve identification	35
3.4.2 Curve identification validation process	37
3.5 ANALYSIS	38
3.5.1 Horizontal curve length estimation	39
3.5.2 Horizontal curve radius estimation using the circular regression method.....	42
3.5.3 Horizontal curve radius estimation using the long chord method	46
3.5.3 Horizontal curve radius estimation comparison	48
3.5.5 Sources of error	50
3.5.6 Sensitivity of errors	53
3.6 CONCLUSIONS AND RECOMMENDATIONS	56
CHAPTER 4. HORIZONTAL CURVE CRASH PREDICTION MODEL	58

4.1 INTRODUCTION	58
4.2 REVIEW OF LITERATURE	58
4.3 DESCRIPTIVE STATISTICS	62
4.3.1 All rural, paved, two-lane roadway horizontal curves	62
4.3.2 Primary rural, paved, two-lane roadway horizontal curves	70
4.3.3 Secondary rural, paved, two-lane roadway horizontal curves	74
4.4 METHODOLOGY	78
4.4.1 Data collection and preparation	78
4.4.2 Negative binomial regression	78
4.4.3 Variables	79
4.5 ANALYSIS	81
4.5.1 All crashes with $R_{\text{regression}}$	81
4.5.2 Serious crashes with $R_{\text{regression}}$	82
4.5.3 All crashes with R_{chord}	82
4.5.4 Serious crashes with R_{chord}	83
4.5.5 Goodness-of-fit comparison	84
4.5.6 Empirical Bayes usefulness comparison	85
4.5.6 Interpretation of models	86
4.6 CONCLUSIONS AND RECOMMENDATIONS	88
CHAPTER 5. GENERAL CONCLUSIONS	90
5.1 GENERAL DISCUSSION	90
5.2 RECOMMENDATIONS and CONCLUSIONS	90
5.3 FUTURE RESEARCH	91
5.3.1 Funding allocation	91
5.3.2 Expanded horizontal curve identification	91
5.3.3 Additional variables	91
5.4 REFERENCES	92

LIST OF FIGURES

Figure 2-1. Fatal crash trend on Iowa roadways by facility type, 1970-2009.	7
Figure 2-2. Fatal crash rate trend on Iowa roadways by facility type, 1970-2009.	7
Figure 2-3. The effect of cartography changes.	11
Figure 2-4. Crash and safety project data timeline	12
Figure 2-5. State system – fatal and serious injury crash data classification.....	15
Figure 2-6. Local system – fatal and serious injury crash classification.	15
Figure 2-7. State system – allocation of funding by project type and location.	17
Figure 2-8. Local system – allocation of funding by project type and location.	18
Figure 2-9. "Relative difference" example calculation.	19
Figure 2-10. "Relative difference" color scale.	20
Figure 2-11. Matched crash data and safety funding data classification for rural roadway facilities.....	21
Figure 2-12. Matched crash data and safety funding data classification for urban state system facilities.....	22
Figure 2-13. Matched crash data and safety funding data classification for urban local system facilities.....	23
Figure 2-14. Relative safety investment for Iowa roadway classifications (crash densities show in parentheses).	25
Figure 3-1. Two-lane horizontal curve distribution with paved two-lane, rural roads shown in gray.	32
Figure 3-2. Plan set curve data locations.	35
Figure 3-3. Curve identification process with GPS traces and simplified polyline vertices.	36
Figure 3-4. Actual curve length vs. estimated curve length.	40
Figure 3-5. Curve length versus percent error of estimated length value.	41
Figure 3-6. Curve length histogram comparison for sample curves and primary roadway curves.	42
Figure 3-7. Curve length histogram comparison for sample curves and all curves in database....	42

Figure 3-8. Actual radius vs. circular regression method estimated radius.	43
Figure 3-9. Curve radius versus percent error of estimated radius value, $R_{\text{regression}}$	44
Figure 3-10. Curve radius ($R_{\text{regression}}$) histogram comparison for sample curves and primary roadway curves.	45
Figure 3-11. Curve radius ($R_{\text{regression}}$) histogram comparison for sample curves and all curves in database.....	45
Figure 3-12. Actual curve radius vs. long chord method estimated curve radius.....	46
Figure 3-13. Curve radius versus percent error of estimated radius value, R_{chord}	47
Figure 3-14. Curve radius (R_{chord}) histogram comparison for sample curves and primary roadway curves.	48
Figure 3-15. Curve radius (R_{chord}) histogram comparison for sample curves and all curves in database.....	48
Figure 3-16. Percent RMSE comparison by curve radius category.....	49
Figure 3-17. Effect of lateral shift on travel path radius.....	50
Figure 3-18. Effect of incorrect GPS trace “location” data.	51
Figure 3-19. Long chord method underestimation example.....	52
Figure 3-20. Sensitivity of predicted crash frequency to radius percent difference.	54
Figure 3-21. Sensitivity of predicted crash frequency to $R_{\text{regression}}$ percent difference.	55
Figure 3-22. Sensitivity of predicted crash frequency to R_{chord} percent difference.	56
Figure 4-1. Curve crash rate as a function of radius.	59
Figure 4-2. Number of horizontal curves by number of all crashes.	63
Figure 4-3. All crashes by AADT on all rural, paved, two-lane roadway horizontal curves.	64
Figure 4-4. Serious crashes (K+A) by AADT on all rural, paved, two-lane roadway horizontal curves.	64
Figure 4-5. Number of crashes by curve radius category for all rural, paved, two-lane roadway curves.	65
Figure 4-6. Crash severity ratio by curve radius category for all rural, paved, two-lane roadway curves.	66
Figure 4-7. Crash rate per HMVMT for all crash severities by curve radius category.	67
Figure 4-8. Fatal crash rate per HMVMT by curve radius category for all curves.	68

Figure 4-9. Crash frequency for all horizontal curves by crash severity and lane width.	68
Figure 4-10. Crash frequency for all horizontal curves by crash severity and terrain adjacent to the roadway.....	69
Figure 4-11. Crash frequency for all horizontal curves by crash severity and shoulder type.....	70
Figure 4-12. Crash frequency for all horizontal curves by crash severity and shoulder width.	70
Figure 4-13. All crash frequency by AADT on primary, rural, paved, two-lane roadway horizontal curves.....	71
Figure 4-14. Serious crash (K+A) frequency by AADT on primary, rural, paved, two-lane roadway horizontal curves.....	72
Figure 4-15. Crash rate per HMVMT on primary roadway curves for all crash severities by curve radius category.....	72
Figure 4-16. All crash frequency by AADT on secondary, rural, paved, two-lane roadway horizontal curves.....	74
Figure 4-17. Serious crash frequency by AADT on secondary, rural, paved, two-lane roadway horizontal curves.....	75
Figure 4-18. Crash rate per HMVMT on secondary roadway curves for all crash severities by curve radius category.....	76
Figure 4-19. Fatal crash rate per HMVMT on secondary and primary roadway curves comparison.....	76
Figure 4-20. Expected all-crash frequency vs. curve radius on all horizontal curves.....	86
Figure 4-21. Expected serious crash frequency vs. curve radius on all horizontal curves.....	86
Figure 4-22. Effect of radius on curve speed.....	87

LIST OF TABLES

Table 1-1. KABCO scale for crash severity.	3
Table 2-1. Comparison of rural expressway sideslope flattening project and rural secondary two-lane rumble strip projects.	27
Table 2-2. SICL list crash comparison.	28
Table 2-3. Iowa 5 percent crash comparison for SVROR crashes.	29
Table 3-1. Iowa statewide crash comparison for horizontal curves (2001-2009).....	31
Table 3-2. Identified curves by system type.	39
Table 3-3. Percent RMSE comparison by curve radius category.	49
Table 3-4. Long chord method calculation example.	53
Table 4-1. Horizontal curve crashes by severity, lane width, and terrain for primary roads.....	73
Table 4-2. Horizontal curve crashes by severity, shoulder type, and shoulder width for primary roads.	74
Table 4-3. Horizontal curve crashes by severity, lane width, and terrain for secondary roads.	77
Table 4-4. Horizontal curve crashes by severity, shoulder type, and shoulder width for secondary roads.....	77
Table 4-5. All crash model using $R_{\text{regression}}$	81
Table 4-6. Serious crash model using $R_{\text{regression}}$	82
Table 4-7. All crash model using R_{chord}	83
Table 4-8. Serious crash model using R_{chord}	84
Table 4-9. McFadden's ρ^2 goodness-of-fit comparison.	84
Table 4-10. Calculated average weights for comparing a model's usefulness in the empirical Bayes process.....	85

ACKNOWLEDGMENTS

I would first like to thank Dr. Souleyrette for his guidance, encouragement and support throughout this research process. I would also like to thank Zach Hans for providing me with the GIS expertise and knowledge needed to complete this research. I would like to thank Yu Qui for the statistical help and the numerous undergraduate assistants who helped create, compile, and analyze the vast amount of data required for this research. Additionally, I want to thank my program of study committee for their support and assistance. Finally, I would like to thank my soon to be wife, Mallory, for always supporting me throughout my studies and work.

ABSTRACT

It is well documented that motor vehicle crashes are a public safety concern. However, traditional approaches do not always lend themselves to addressing the complete extent of this “safety problem”. Identifying the extent of the “safety problem” is an important step in optimizing safety fund allocation and analyzing horizontal curve safety. This study investigates the allocation of safety expenditures in Iowa, relative to crash data. The matching of crash data with safety expenditures suggests the shift of funds from the high crash density, state system to facilities on the low density, local system. However, the redistribution of funding should also consider factors such as crash density and benefit cost. Furthermore, because some crashes are too widely distributed to be identified using traditional high crash location methodology; a balance of systematic and high crash location methods should be considered. Ultimately, the optimum balance of safety resources should reduce the most possible fatal and serious injury crashes. This study also investigated a systematic method for identifying and estimating geometric parameters on horizontal curves. A validation of this method showed that as horizontal curve radius decrease, sensitivity to errors in the estimated curve radius increase. Although some large errors associated with the estimated curve radius were found, predicted crash frequency for all curves was found to be no more than twenty percent different than the actual predicted crash frequency. Lastly, safety performance functions created for the horizontal curve database did not yield a concrete correlation between curve radius and crash frequency. Because of the random nature of fatal and major injury crashes, care is advised when creating crash models for these crashes.

CHAPTER 1. GENERAL INTRODUCTION

1.1 INTRODUCTION

Highway safety in the United States is a national epidemic. Nationwide, in 2009, there were 33,808 traffic related fatalities. Furthermore, motor vehicle traffic crashes are the leading cause of death for people ages four to thirty-five, and the ninth leading cause of death for all age groups (NHTSA, 2006). Although there has been a recent downward trend, the number of traffic fatalities nationwide has remained largely constant since the mid 1980s. Subsequently, the approach for assessing highway safety has begun to transition from action based on “experience, intuition, judgment, and tradition, to action based on empirical evidence, science, and technology” (HSM Practitioner’s Guide, 2011). In Iowa, the implementation of these evolving methods is especially important for optimizing safety funding and analyzing horizontal curve safety performance.

In order to optimize highway safety expenditures, funds must not only be invested in actions that can most effectively mitigate the “safety problem”, but must also match the extent of the “safety problem”. However, traditional methods do not always lend themselves to evenly addressing these safety needs. Theoretically, if all “safety problem” types can be equally mitigated, funding should match the extent of the problem. Therefore, identifying the extent of the “safety problem” and determining the allocation of funds is an important first step in the optimization of safety expenditures.

Identifying the extent of the “safety problem” is also important for horizontal curve safety. In Iowa, rural horizontal curves comprise of only 1.2 percent of the total statewide roadway system, yet, 10.5 percent of the state’s fatal crashes occur on curves. In order to effectively address the safety performance of these horizontal curves, curve locations, characteristics and geometric parameters must be known. However, little is known about horizontal curves in Iowa. Furthermore, systematic curve identification and parameter estimation is difficult on a large system.

This thesis addresses key issues related to these two topics. First, a data-driven analysis method of balancing statewide safety funding relative to crash data is developed and critiqued. Secondly, a systemic horizontal curve identification and geometric parameter

estimation method is explored and validated. Lastly, crash prediction models are developed for estimating the number of expected crashes on horizontal curves.

1.2 THESIS ORGANIZATION

This thesis is divided into five chapters. Chapter 1 (this chapter) provides an introduction of the thesis and a review of basic highway safety literature related to high crash locations and systematic analysis. Chapter 2 investigates the highway safety funding allocation in Iowa relative to the eight year statewide crash data. It also discusses considerations for balancing high crash location and systematic analysis in the safety funding allocation process.

Chapter 3 provides an overview of a systemic horizontal curve identification and parameter estimation process as well as a validation of the method. Chapter 4 consists of the development of safety performance functions for predicting horizontal curve crashes on rural, paved, two-lane highways. Finally, Chapter 5 includes the general conclusions of the previous three chapters and provides final recommendations for funding allocation and horizontal curve identification, parameter estimation, and safety performance analysis.

1.3 REVIEW OF LITERATURE

There are two main methods of evaluating roadway safety performance: high crash location, or “black spot” analysis and systematic, or “mass action” analysis. Historically black spot analysis has been the most common method to identify candidate locations for safety improvements (Preston, et al., 2010). Black spot analysis finds intersections, horizontal curves, or even short roadway corridors that “exhibit unusually high crash frequencies or crash rates” (p. 3 Preston, e. al., 2010). Locations are then analyzed, ranked and prioritized. Black spot analysis typically use all crashes as a performance measure due to fatal and serious injury crashes being too widely dispersed and random to yield statistically significant locations.

Mass action analysis is a fairly new method deployed by state DOTs. Mass action is a proactive method that targets low density and random crashes by employing a system-wide improvement. The objective of the mass action method is to “identify candidates for a wide

deployment of lower-cost safety measures over many miles of roadway segments, corridors, or even over the entire roadway system” (p. 4, Preston et al., 2010).

Road departure and cross center line crashes are two examples of crashes that occur randomly and commonly on high speed rural roadways. These crashes are distributed widely across many miles of roadway and therefore are not identified using a “black spot” analysis. Systematic improvements such as shoulder or center line rumble strips are two low cost countermeasures that can be deployed to mitigate these widely dispersed crashes.

The safety of an entity, or roadway, cannot be measured solely by the count of accidents because of the random fluctuation of those accidents. If the safety of an entity were to be measured by only number of accidents in a year, a drop in crashes from one year to the next would mean that the safety of the roadway improved, when the roadway itself remained unchanged. One way to define safety is as “the number of accidents (crashes), or accident consequences, by kind and severity, expected to occur on an entity during a specific period” (p. 24, Hauer, 1997).

Crash severity is commonly measured on the KABCO scale. The KABCO was established by the American National Standards Institute, and is used by law enforcement officers in coding crash details at a crash scene (Sinha, 2007). The state of Iowa uses this scale to distinguish crash severity. Crashes are classified by the most severely injured person involved in a crash. Table 1-1 shows the KABCO scale for crash severity, including a description of each coding.

Table 1-1. KABCO scale for crash severity.

Code	Crash Severity	Definition
K	Fatal	One or more deaths
A	Serious Injury	incapacitating injury preventing victim from functioning normally (e.g., paralysis, broken/distorted limbs, etc.)
B	Minor Injury	non-incapacitating but visible injury (e.g., abrasions, bruising, swelling, limping, etc.)
C	Possible Injury/ Unknown	probable but not visible injury (e.g., sore/stiff neck)
O	Property Damage Only (PDO)	property-damage only

The expected number of crashes on a horizontal curve is estimated by applying crash modification factors (CMF) to base conditions. The base condition safety performance function (SPF) for a rural roadway segment is shown in Equation 1-1 (HSM Practitioner's Guide, 2011).

Equation 1-1:

$$N_{spf-rs} = (AADT_n) (L) (365) (10^{-6}) e^{-0.312}$$

$AADT_n$ = AADT of horizontal curve segment

L = horizontal curve length

CMFs are developed for different roadway attributes to assess the relative safety performance of an entity. The CMF for horizontal curves was developed to represent how the crash experience of tangent and horizontal curve segments differ. Equation 1-2 shows the CMF for the safety effect of horizontal curves. This CMF, along with several other CMFs related to the roadway, are then applied to the SPF for the base prediction model as shown in Equation 1-3 to determine the total safety effect of individual geometric features (HSM Practitioner's Guide, 2011).

Equation 1-2:

$$CMF_{curves} = \frac{1.55L_c + \left(\frac{80.2}{R}\right) - 0.012(S)}{1.55L_c}$$

L_c = horizontal curve length (miles)

R = horizontal curve radius (feet)

S = presence of spiral transition. One if yes, zero if no.

Equation 1-3:

$$N_{predicted-rs} = N_{spf-rs} * \left(\prod CMF_i\right) * C_r$$

$N_{predicted-rs}$ = predicted number of crashes for a rural horizontal curve

CMF_i = Crash modification factor for roadway attribute_i

$C_r = 1.0$ for base condition

CHAPTER 2. OPTIMIZING SAFETY FUND ALLOCATION

2.1 INTRODUCTION

In 2005, the Safe, Accountable, Flexible, Efficient Transportation Equality Act: A Legacy for Users (SAFETEA-LU) was signed into law. This new transportation bill built on its predecessors and became the largest surface transportation investment in U.S. history. One key component of this bill is highway safety. A separately funded Highway Safety Improvement Program (HSIP) was established to help finance projects that will aid in reducing highway fatalities (Federal Highway Administration, 2006).

The HSIP requires each state to develop a Strategic Highway Safety Plan (SHSP). “An SHSP is a statewide-coordinated safety plan that provides a comprehensive framework for reducing highway fatalities and serious injuries on all public roads” (Federal Highway Administration, 2006). Using and integrating the four E’s – engineering, education, enforcement, and emergency medical services (EMS), the SHSP establishes statewide safety goals, objectives and key emphasis areas. Moreover, the SHSP requires that safety investment decisions be data-driven.

1998’s Transportation Equality Act for the 21st Century (TEA-21) pushed for “safety conscious planning” with the goal to prevent “human and economic losses that result from motor vehicle and non-motorized traveler-related crashes” (NCHRP, 2010). Human and economic loss implies crashes of all severity, from fatal crashes to property damage only crashes. As previously mentioned the purpose of the SHSP is to reduce highway fatalities and serious injuries on *all public roads*. This is a change from the previous legislation which aimed to prevent all crash severities.

The need to address *all public roads* is apparent in Figure 2-1 and Figure 2-2. Figure 2-1 shows the fatal crash trend on Iowa roadways by facility type from 1970-2009. Figure 2-2 shows the fatal crash rate trend on Iowa roadways by facility type for the same time period. Fatal crashes on Iowa roadways have decreased since the 1970’s but not equally on all roadway facilities. Rural secondary roadways have actually seen an increase over the past ten years. Fatal crash rates have also decreased until about ten years, where they have stayed fairly constant since 2000.

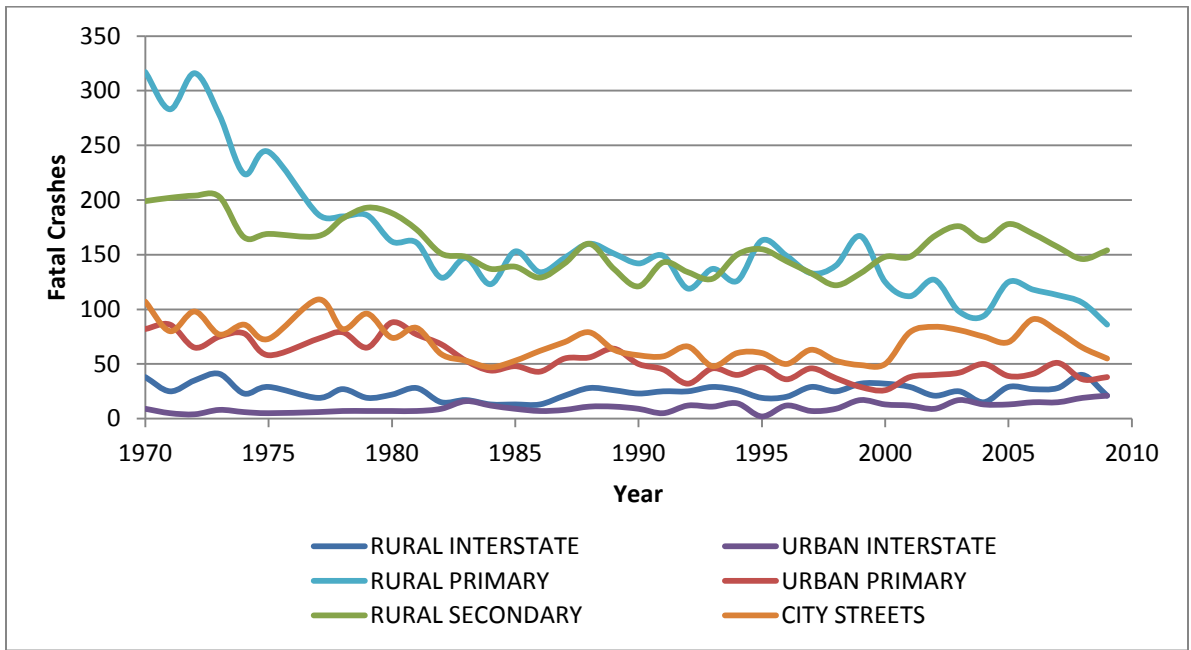


Figure 2-1. Fatal crash trend on Iowa roadways by facility type, 1970-2009.

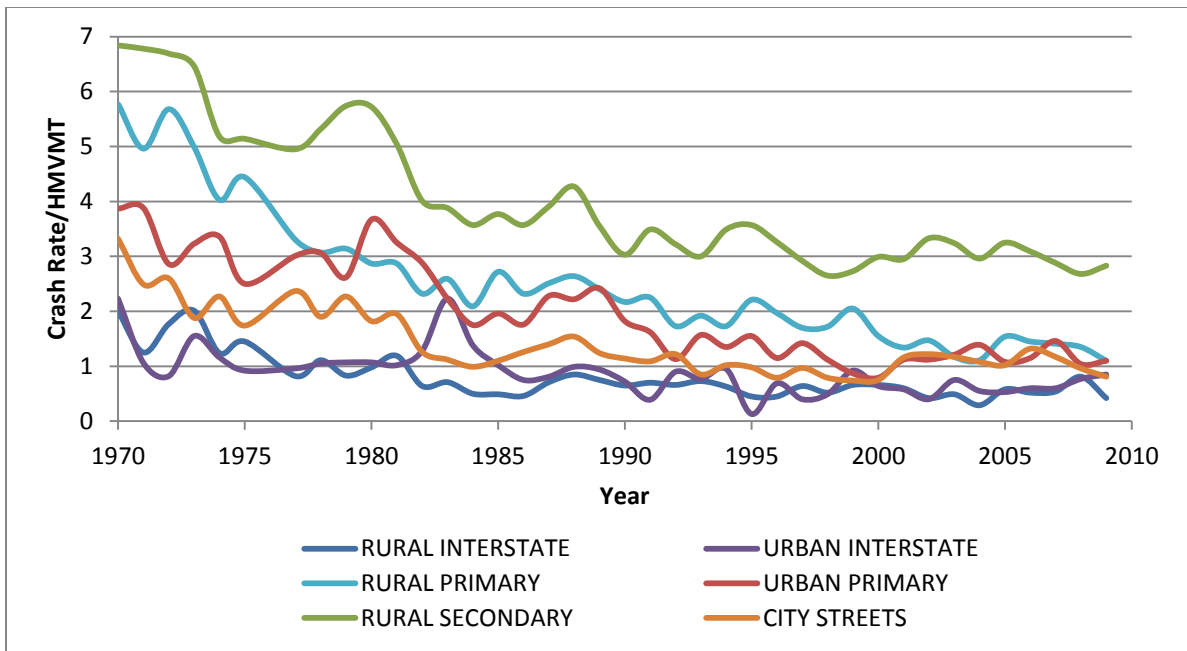


Figure 2-2. Fatal crash rate trend on Iowa roadways by facility type, 1970-2009.

This study was conducted to first complete a wholly data-driven analysis of crash data in Iowa to recognize potential facilities or crash types in need of safety mitigations. Secondly, safety funding allocations were matched with crash data for different categories of roadways to identify the funding investment relative to the crash data. Lastly, funding allocation considerations, including the need for finding a balance between high crash locations and mass-action analysis, are discussed.

2.2 REVIEW OF LITERATURE

In 2006 the Iowa Department of Transportation (DOT) published the first Iowa Comprehensive Highway Safety Plan (CHSP) as mandated by SAFETEA-LU and the HSIP. The Iowa CHSP was a joint effort by safety stakeholders throughout the state of Iowa. It aimed to reduce the annual average of traffic fatalities in Iowa to 400 by 2015. From this effort five policy and eight program strategies were recommended to aid in the reduction of traffic fatalities and injuries (CHSP, 2006). The top five safety policy areas recommended were:

- Young drivers
- Occupant protection
- Motorcycle safety
- Traffic safety enforcement
- Traffic safety improvement program

The top eight safety program areas recommended were:

- Lane departure
- Safety corridors
- Intersections
- Local roads
- State traffic records
- Senior mobility
- Safety training and education
- Unpaved rural roads

The “Iowa Five Percent Report” identifies areas of safety needs based on an analysis of fatal and major injury crashes. This analysis identified Iowa’s most severe safety needs are crashes associated with single vehicle running off the road (SVROR), vehicles crossing the centerline on two-lane highways, vehicles crossing the medians on freeways, horizontal curves, intersections, unbelted drivers and passengers, impaired drivers, and speeding (Iowa Five Percent Most Severe Safety Needs Report, 2010). Sites for these eight safety needs are then prioritized, separately, based on annual fatal and serious injury crash densities. After prioritization, safety mitigations are suggested for each site.

Iowa intersections are also prioritized in the “Safety Improvement Candidate Location (SICL) List”. The SICL list identifies the “200 highest ranked intersections relative to crash history” (p. 2 Iowa Five Percent Most Severe Safety Needs Report, 2010). The top five percent of these intersections are identified as the most severe intersections with safety needs in the “Iowa Five Percent Report”.

Safety improvement projects in Iowa are funded, primarily, by two sources. The first source of funding is federal funding from the HSIP. “The actual allocation is subjective based on need, the specific strategy selected, and the five percent process. Projects are prioritized by benefit-cost analysis consistent with requirements for reporting project evaluations to FHWA” (p. 14, Preston et. al., 2010). Of the available HSIP funding, approximately 90 percent is spent on rural roads. All HSIP funding is used on roadways not under the state jurisdiction.

Iowa also provides state safety funding through its Traffic Safety Improvement Program (TSIP). These funds are available to all local jurisdictions (cities and counties) as well as the Iowa DOT. Funding is available for three categories of projects; site-specific, traffic control devices, and research, studies, and public information. Overall, Iowa “directs approximately 18 percent of safety funds towards projects on local roads” (p. 13 Preston, et. al, 2010).

Preston, et al. (2010) also identified possible safety funding allocation inequalities based on state survey data and overall crash data. HSIP funding is available for the local system, however the federal reporting requirements are often cumbersome and few local agencies take advantage of the opportunity because there is a separate, less labor intensive

safety program (TSIP). Consequently, about 82 percent of available safety funds are allocated to the state system even though nearly 50 percent of fatal crashes occur on the local system (Preston, et al., 2010).

2.3 DATA

2.3.1 Crash and roadway data

Crash data were assembled from the Iowa SAVER crash database for 2001-2008. Fatal and major injury crashes were queried from the database of all crashes. Statewide, from 2001-2008, there were 3,018 fatal crashes and 13,370 injury crashes. Statewide roadway data were obtained from the Iowa GIMS roadway database for 2005. Crash data were assigned to GIMS roadway segments using a spatial join in ArcGIS.

Using a spatial join to assign crash data to the GIMS network can be problematic because of cartographic issues associated with GIMS data from year to year. Each year the cartography of the GIMS network improves and the GIMS data get closer and closer to their actual locations. As a result, the alignment of the GIMS database can shift slightly from year to year. This becomes a problem because crash locations are digitized based on the existing cartography. Therefore, crashes that are not digitized using a given year's GIMS database could be wrongly located on that year's GIMS network. Figure 2-3 shows a "location where the cartography changed and the new intersection location is fifty meters from the previous intersection location" (p. 18 Jackson, 2006).

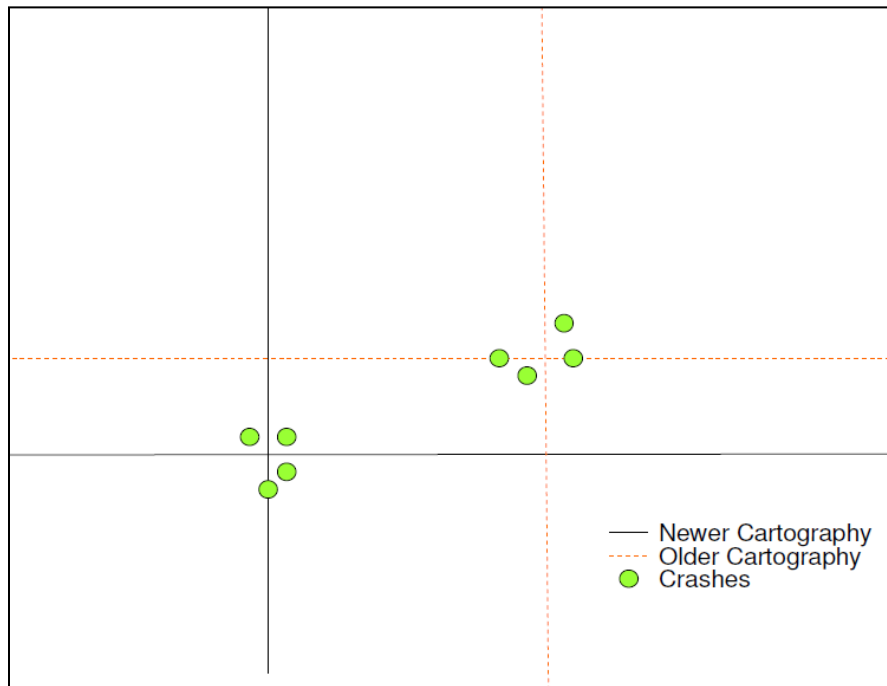


Figure 2-3. The effect of cartography changes.

These cartographic issues are compounded at intersection locations. Because of this, only crashes coded as occurring at an intersection were classified as intersection crashes. Other crashes that were visibly located near an intersection, but were not coded as occurring at an intersection, were classified as non-intersection crashes.

The US Bureau of the Census 2000 Urbanized Area Boundary Map was used to code the GIMS database as either urban or rural. “Urban” areas are classified by all territories, population, and housing units located within an urbanized area or urban cluster” (Bureau of the Census, 2000). “Rural” areas are anything outside of an urbanized area or cluster. Roadway segments were then classified as either “urban” or “rural” based on the land directly adjacent to the roadway segment. This designation was used in lieu of MPO and incorporated cities’ boundaries. Since MPO and incorporated cities’ boundaries contain areas of undeveloped land, some roadways within these boundaries are coded as urban when they are, in nature, rural roadways.

2.3.2 Data related to statewide highway safety projects

Data related to HSIP and TSIP safety projects were provided by the Iowa DOT Office of Traffic Safety. This data included project descriptions and funding information as well as a GIS “shapefile” with the location of each project. Data related to HSIP funded safety projects were available for the 2001-2009 fiscal years. TSIP funded projects data for fiscal years 2004-2011 were also provided. Figure 2-4 shows a timeline of the overlap in crash data and safety project data.

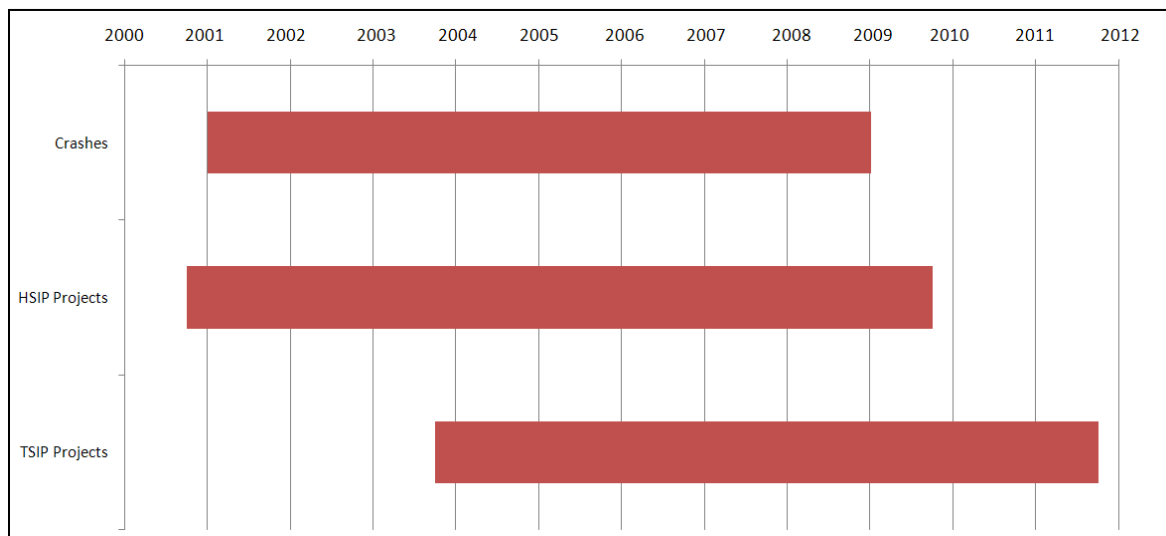


Figure 2-4. Crash and safety project data timeline

2.4 METHODOLOGY

2.4.1 Crash data classification

As previously mentioned, crash data were assigned to the GIMS network using a spatial join in ArcGIS. Roadway attributes including access control, number of lanes, median type, and jurisdiction responsible for the roadway, were used to code the facility type of each roadway segment. Secondary and municipal roadways were combined and coded as the local system. Crash types, such as single-vehicle-run-of-road (SVROR), head-on, right angle, rear end, ran stop sign, and ran signal, were coded using attributes available in the crash data. Queries in ArcGIS were then performed to categorize each crash by crash type

and crash location (e.g. roadway type, intersection/non-intersection, rural/urban).

2.4.2 Safety project classification

In order to categorize each safety project, the safety project description and funding data were joined to the project location data in ArcGIS. To distinguish between intersection and non-intersection, each project was coded manually using its project description data. Urban and rural coding was accomplished using the US census Urbanized Area Boundary Map information in ArcGIS.

To determine the facility type of each highway safety project, a spatial join was performed between the project location data and the GIMS network. Roadway attributes including access control, number of lanes, median type, and jurisdiction responsible for the roadway, were then used to code the facility type of each safety project. Safety project data were then categorized using the same criteria as per the crash classification. This process was performed for both the HSIP funded projects and TSIP funded projects.

2.4.3 Combined crash and safety project classifications

After the crash data classification and safety project data classifications were complete they were matched and combined. This was done to match safety funding to different categories of crash types and locations. For example, crashes classified as road departures occurring at non-intersection locations on rural freeways were matched with the safety funding allocated to non-intersection, rural freeway improvements mitigating road departures (e.g. rumble strips, shoulder improvements).

A “relative difference” for each category was then calculated. The “relative difference” of a crash location/type category shows the difference between the safety investment and the number of crashes for that category, relative to all other categories. Equation 2-1 shows equation used to calculate “relative difference”.

Equation 2-1:

$$\text{"Relative Difference"} = \frac{(\% \text{ funding})_i - (\% K + A)_i}{\Delta_{max}}$$

$\% \text{ funding} = (\text{funding allocated to category } i / \text{total statewide safety funding}) * 100$

$\% \text{ K+A} = (\text{number of K+A crashes for category } i / \text{total statewide K+A crashes}) * 100$

$\Delta_{max} = \text{maximum difference between } (\% \text{ funding})_i \text{ and } (\% \text{ K + A})_i$

The “relative difference” yields a number between -1 and +1. The closer a roadway category’s “relative difference” is to -1, the more crashes there are relative to the safety dollars invested in that category. The closer a roadway category’s “relative difference” is to +1, the more funding is invested relative to the number of crashes in that category. The closer to zero a category’s “relative difference”, the more balanced the funding is relative to the number of crashes in that category.

To avoid confusion, the term “classification” will be used to describe the process of separating crash locations and crash types. The term roadway “category” will be used to describe the specific crash location and crash type (e.g. rural, state expressway or rural, secondary, two-lane paved single vehicle run-off-road).

2.5 ANALYSIS AND RESULTS

To illustrate the balance between crash location and type and the allocation of safety funding, three classifications were completed. The first classified crash data only by location and type, the second classified the allocation of funding by location and type, and the final classification combined and matched the first two. Lastly, funding allocation relative to crash density is addressed and the need for balancing the black spot and systematic methods is discussed.

2.5.1 Statewide crash data classification

Crash data were classified first by system type (state and local) and then by urban and rural distinction. The local system in this analysis includes both secondary and municipal roadways. Figure 2-5 shows the classification of the primary system fatal and serious injury crashes. Figure 2-6 shows the classification of fatal and serious injury crashes on the local roadway system.

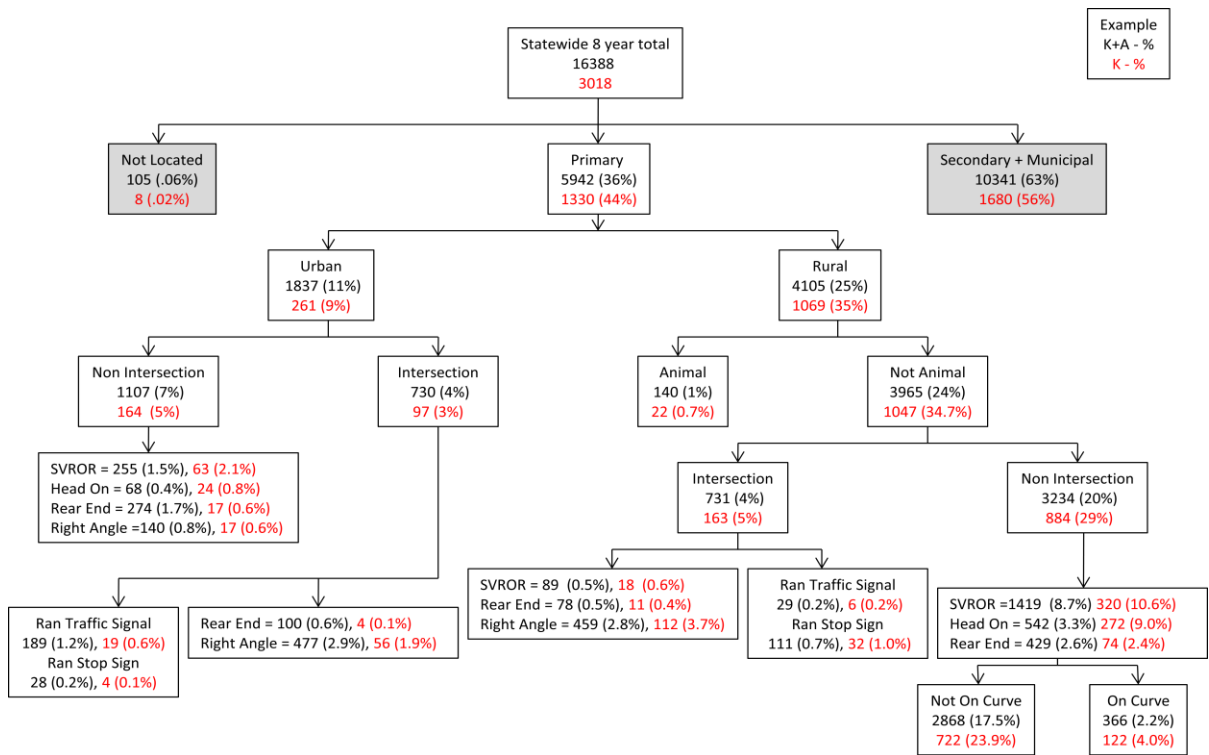


Figure 2-5. State system – fatal and serious injury crash data classification.

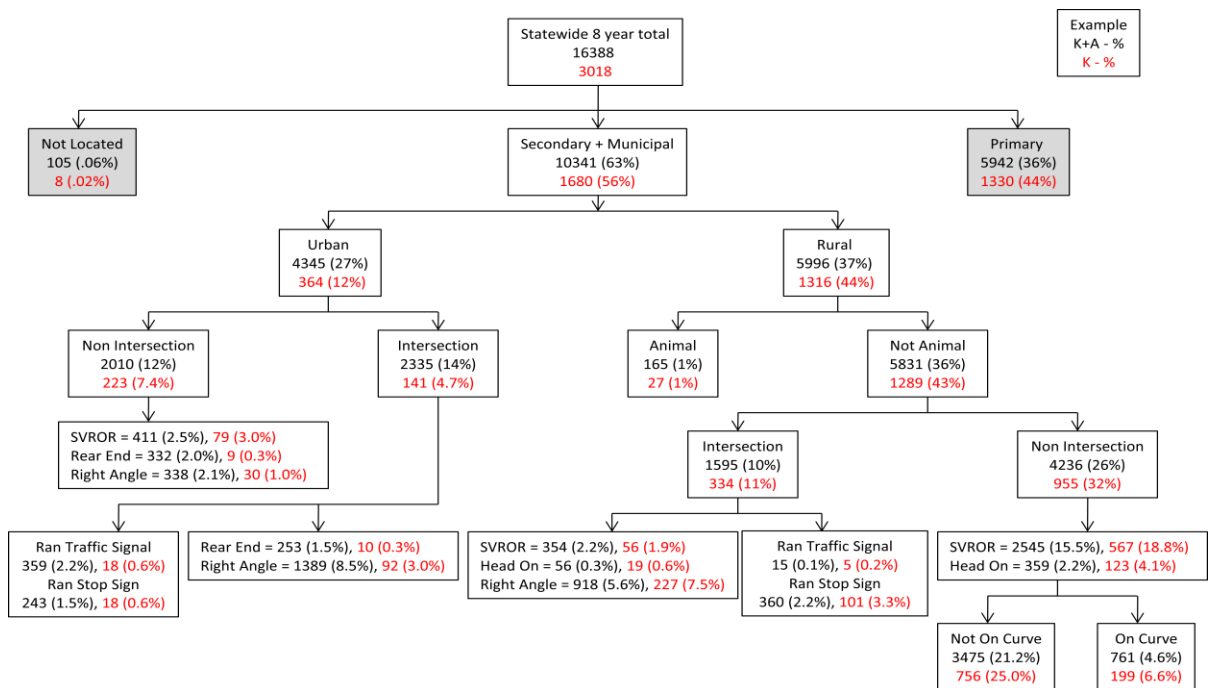


Figure 2-6. Local system – fatal and serious injury crash classification.

Of the 16,388 fatal and major injury crashes statewide during the 8 year analysis period, nearly two-thirds occurred on local roadways. For both the primary and local system there were about twice as many rural crashes as urban crashes. Furthermore there is a greater portion of fatal crashes on rural roadways than on urban roadways. Forty-four percent of all fatalities occur on the rural local system while 35 percent occur on rural primary roadways. Therefore over three in every four fatalities occurs on a rural roadway.

This over dispersion on the rural system can be attributed to a few factors. First, rural roadways tend to operate at higher speeds therefore the risk for a fatal crash is increased. Second, emergency response times for rural roadways are much higher than on urban roadways. According to a 1999 study for the International Symposium on Transportation Recorders the average elapsed time from the moment of the crash until the victim arrives at the hospital for rural fatal crashes is 17 minutes more than for urban fatal crashes (Champion. et. al, 1999). This additional time for emergency response on rural roadways could contribute to a greater number of fatalities.

At intersections, right angle crashes account for the majority of fatal and serious injury crashes. Right angle crashes by nature tend to be more severe. Nearly 9 percent of the statewide fatal and serious injury crashes are right angle crashes occurring at urban intersections on the local system.

Single vehicle run-off-road crashes are the most common non-intersection crash. Single vehicle run-off-road crashes tend to be primarily a rural roadway issue. Over 15 percent of all statewide K+A crashes and nearly 20 percent of all fatal crashes are local system run-off crashes. On the primary system, single vehicle run-off-road crashes account for 9 percent of the statewide K+A crashes and 11 percent of all fatal crashes.

2.5.2 Statewide allocation of safety funds

The second classification categorized safety project funds based on roadway on which the improvement was completed as well as by improvement type. Safety project data were first classified by system type (state and local) and then by urban and rural distinction. The local system in this analysis includes both secondary and municipal roadways. Projects

funded with HSIP funds and projects funded with TSIP funds were assessed separately and in total. Figure 2-7 illustrates the allocation of safety funding by project type and location on the state system. Figure 2-8 presents the same information for the local system.

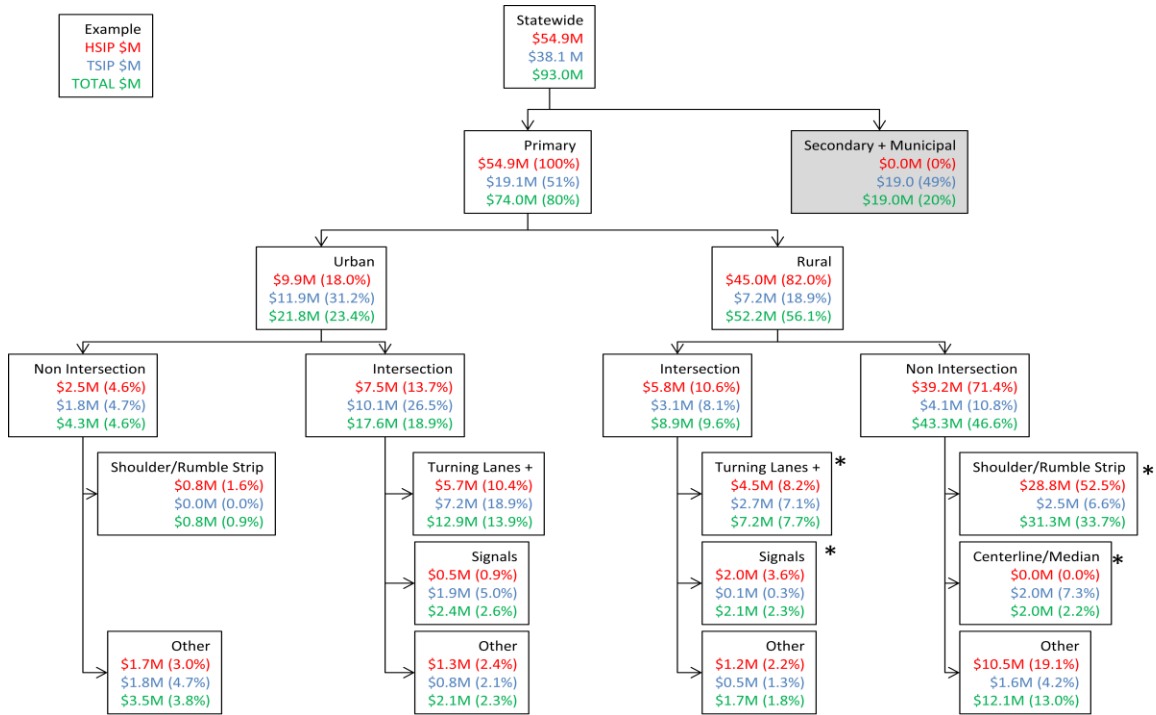


Figure 2-7. State system – allocation of funding by project type and location.

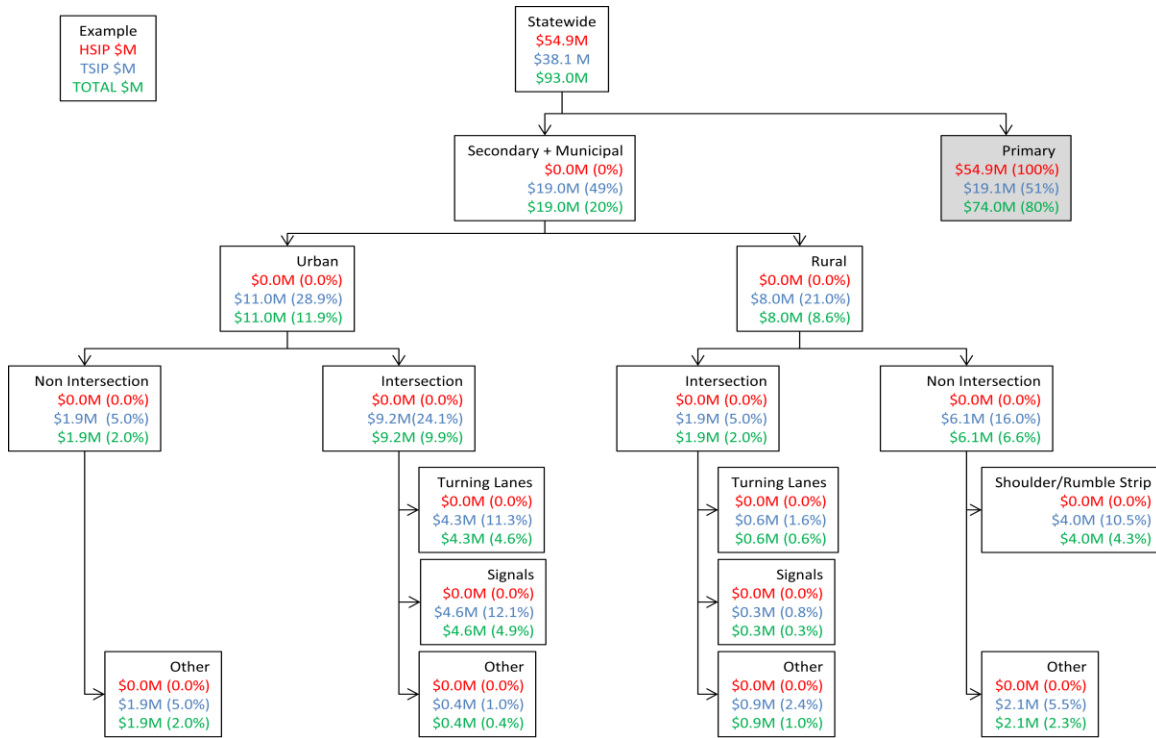


Figure 2-8. Local system – allocation of funding by project type and location.

Of the \$93 million invested in both the HSIP and TSIP projects, 80 percent of the total funding was allocated to the state system. All \$55 million of the HSIP program funds from FY 2001-2009 were invested on the state system. Funding from the TSIP program allocated from FY 2004-20011 was balanced, more or less evenly, between the state and local system.

Approximately 65 percent of the combined HSIP and TSIP funding was invested on the rural state system. Furthermore, approximately a third of all combined HSIP and TSIP funding was allocated to shoulder and edge-line rumble strip projects. The ease of implementation and relatively small capital cost of edge-line rumble strips make these types of projects very attractive. According to the Iowa DOT, “low-cost safety improvements (such as edge-line rumble strips, cable median barrier, and bigger and brighter curve and chevron signs) have proven to be very effective when data is systematically used to identify and address locations with high crash rates” (p.16, Iowa CHSP, 2006).

For intersection improvements, turning lanes are the most represented safety project.

Adding turning lanes can reduce rear-end crashes but are not a right-angle preventative mitigation (CMF Clearinghouse, 2010). Urban intersections on the state system are allocated about 19 percent of all statewide safety funds. Urban intersections are often considered black spots because of high volumes of traffic and a large number of conflict points. “Common black spot locations are intersections, particularly signalized intersections along multi-lane urban arterial roadways” (p. 3 Preston et. al., 2010).

2.5.3 Combined crash data and safety project classification

After the crash data classification and safety project data classification were completed, corresponding categories were matched. For example, crashes classified as road departures occurring at non-intersection locations on rural freeways were matched with the safety funding allocated to non-intersection, rural freeway improvements mitigating road departures (e.g. rumble strips, shoulder improvements). Once all categories of crash location and types were matched with their corresponding funding, the “relative difference” for each category was calculated. Figure 2-9 provides an example of how the “relative difference” was calculated using Equation 2-1.

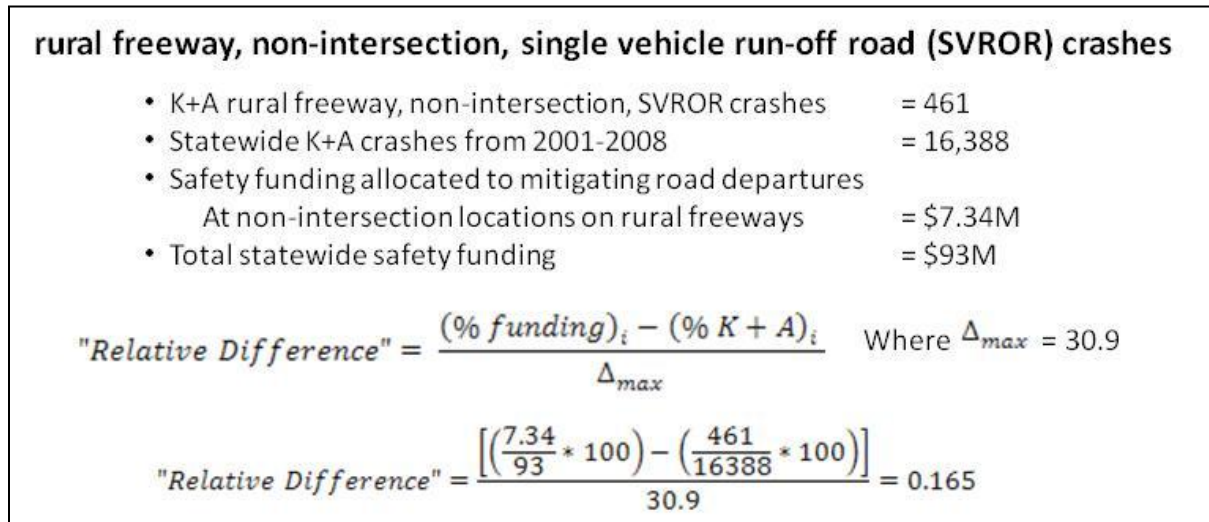


Figure 2-9. "Relative difference" example calculation.

The “relative difference” for each crash location/type category allows categories with different numbers of crashes and different amounts of funding to be compared relative to each other. To better depict how each category compares to one another, a “coloring scale” was created. Figure 2-10 shows the “color scale” used to visually compare different crash location and crash type categories.

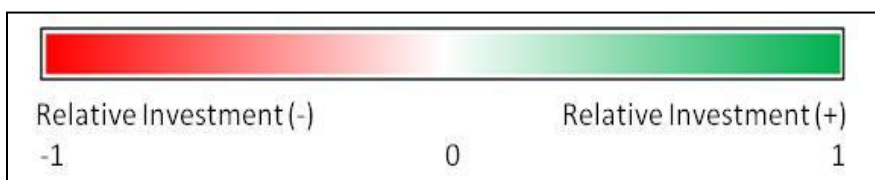


Figure 2-10. “Relative difference” color scale.

The maximum difference between $(\text{percent funding})_i$ and $(\text{percent } K + A)_i$ was found to be 30.9. Thirty point nine is therefore used as Δ_{max} in all “relative difference” calculations. The complete classification, matching crash data and safety funding, is shown in Figure 2-11, Figure 2-12, and Figure 2-13.

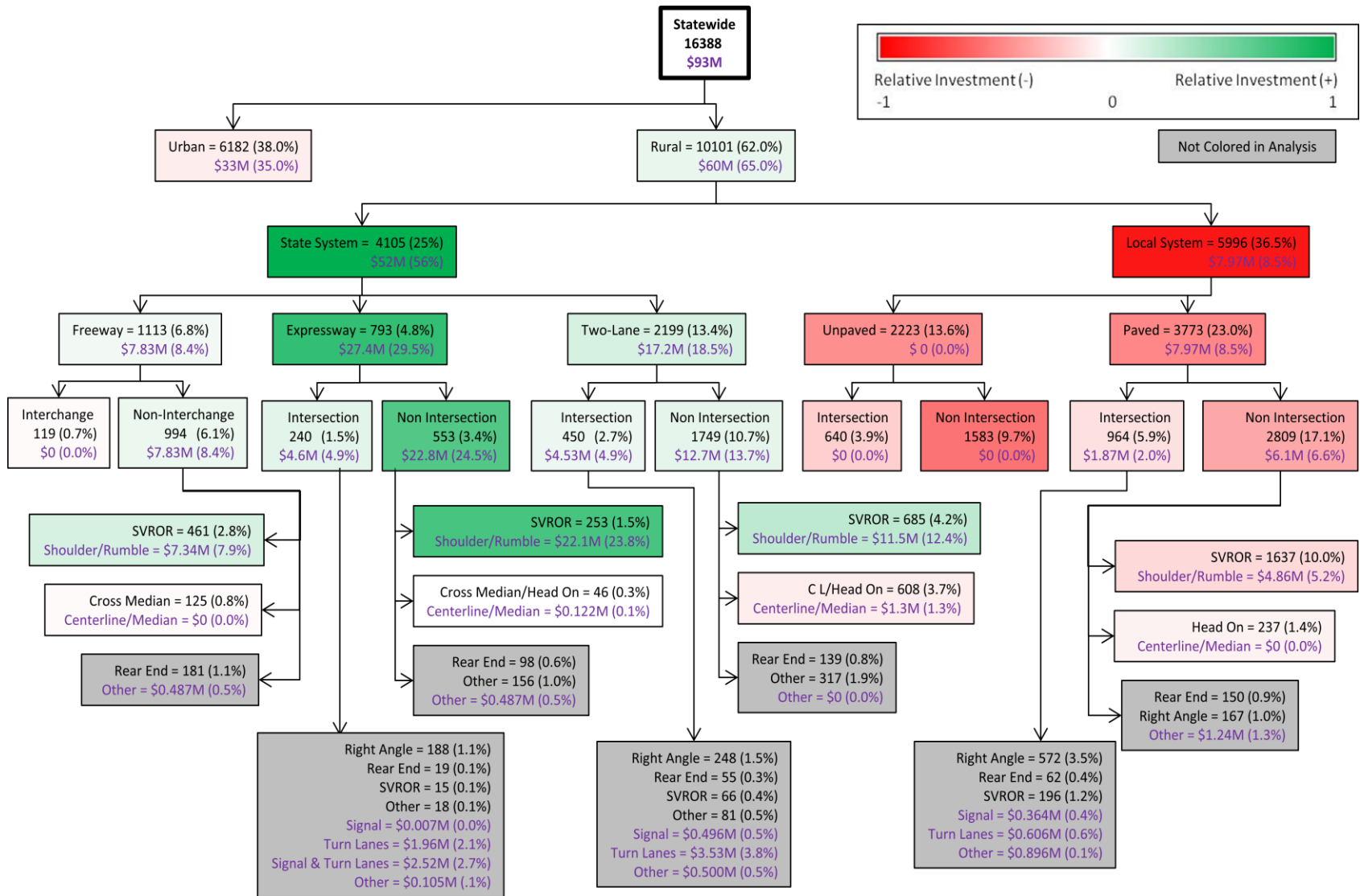


Figure 2-11. Matched crash data and safety funding data classification for rural roadway facilities.

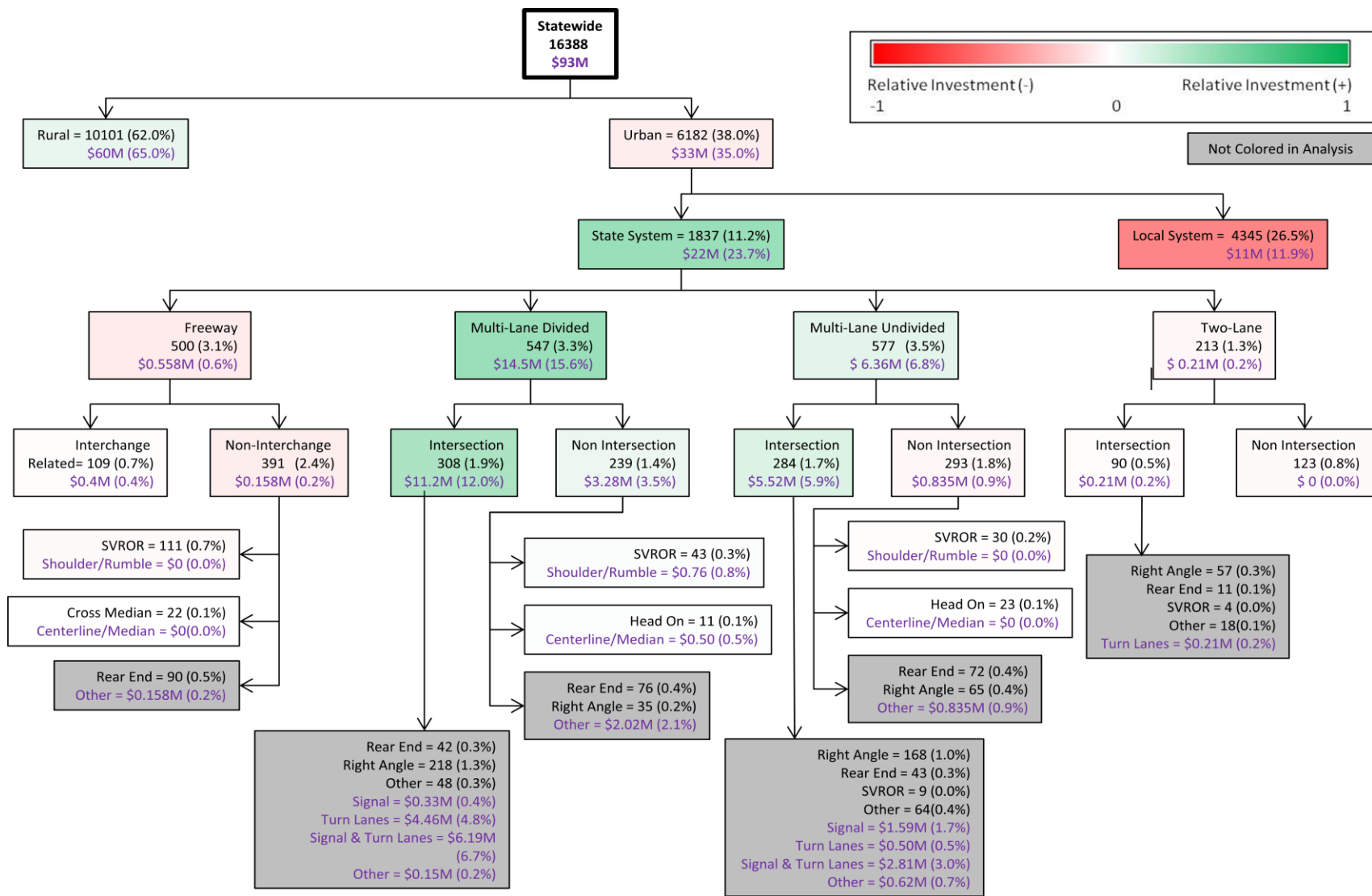


Figure 2-12. Matched crash data and safety funding data classification for urban state system facilities.

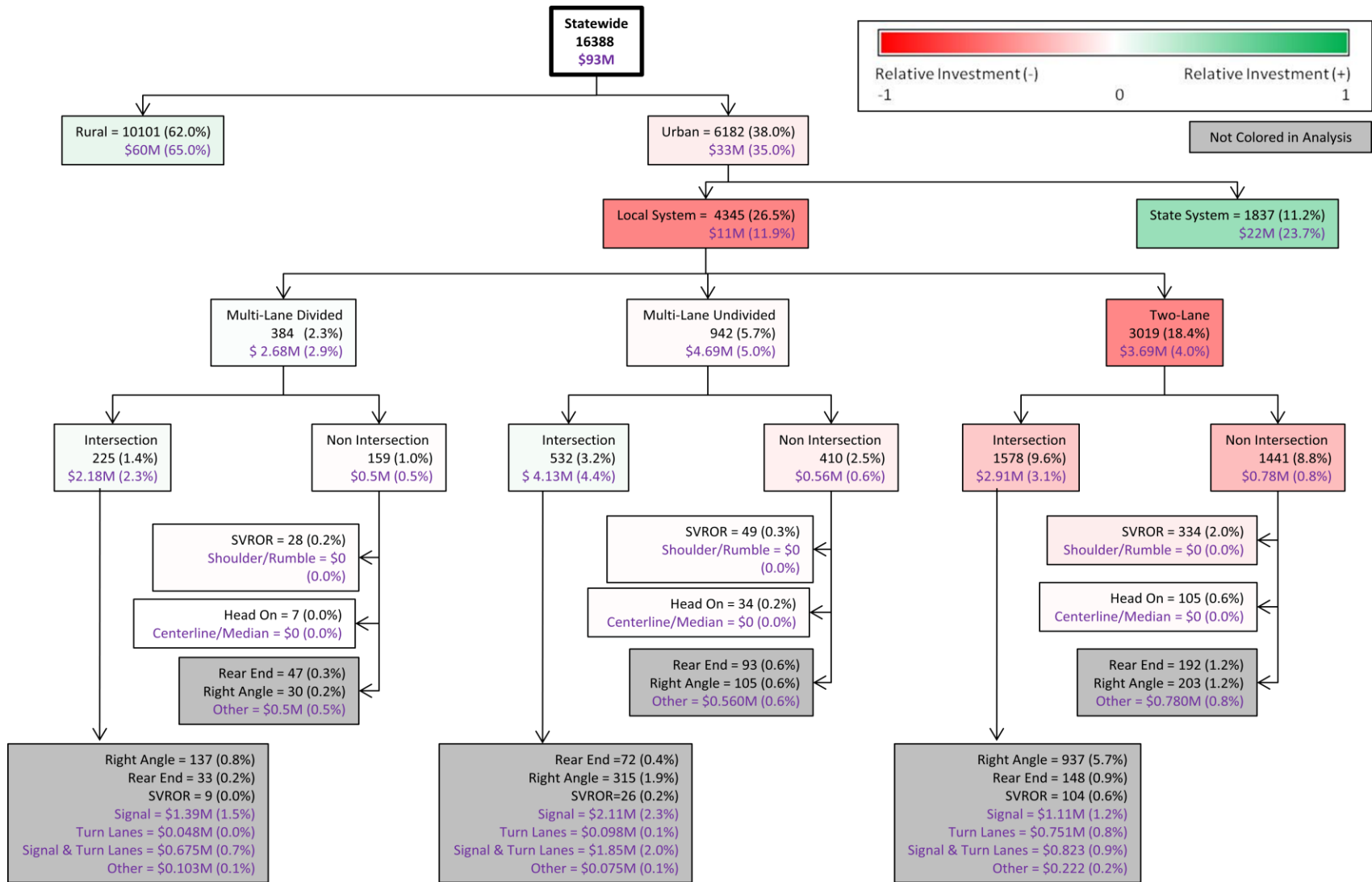


Figure 2-13. Matched crash data and safety funding data classification for urban local system facilities.

Some safety projects were not matched with crash data. Projects such as turn lanes and signals mitigate more than one type of crash; some projects included multiple project types, while some descriptions did not include enough adequate details to connect them with a specific crash type. For these reasons, some projects and crash types were combined and excluded from the “relative difference” analysis. These categories are identified with a gray color box.

In Iowa it appears that rural and urban funding is nearly balanced with the amount of statewide crashes occurring in each respective area. However, a quick examination of the classification shows an apparent disconnect of funding between the state and local system. About 25 percent of the statewide fatal and serious injury crashes occurred on the rural state system yet 56 percent of the statewide funding was allocated to these roadways. In contrast, approximately 36 percent of the statewide fatal and serious injury crashes occurred on the rural local system and only 9 percent of statewide funding was allocated to these roadways. There are many reasons for this apparent disconnect as discussed in the following section.

Furthermore, rural state expressways, where only about 5 percent of statewide crashes occur, were allocated more than one fourth of all statewide expenditures. Most of the projects on rural state expressways were shoulder and/or edge line rumble strip projects. On the rural local system, unpaved roadways received no funding, while paved roadways received about 9 percent of the statewide funding as compared to the 23 percent of K+A crashes that occur on these roadways.

The urban state system received almost a fourth of the statewide funding, yet only 11 percent of the statewide fatal and serious injury crashes during the study period occurred on these roadways. Multi-lane divided intersections on the state system received more funding relative to the number of crashes that occurred on these roadways. The urban local system, which had about one fourth of all statewide K+A crashes, received only 12 percent of all statewide funds.

2.5.4 Safety funding allocation relative to crash density

Upon first inspection there appears to be a significant disconnect between the statewide crash incidence and the safety funding allocation. According to current state policy, funding through the HSIP program is only available for state system projects; local jurisdictions do not have access to that funding. Moreover, this funding accounts for 60 percent of the statewide funding (\$54.9 million).

Like many states, Iowa invests safety dollars on more densely traveled roadways. Roadways such as rural, state expressways and multi-lane, urban roadways have received more funding because their crash densities are greater than other systems. The roadway systems with higher crash densities tend to receive more funding relative to number of crashes occurring on those systems. Figure 2-14 illustrates this investment in higher density roadways. Figure 2-14 shows the safety investment on Iowa roadways relative to the number of crashes on each system, with the average number of crashes per mile, per year from 2001-2008.

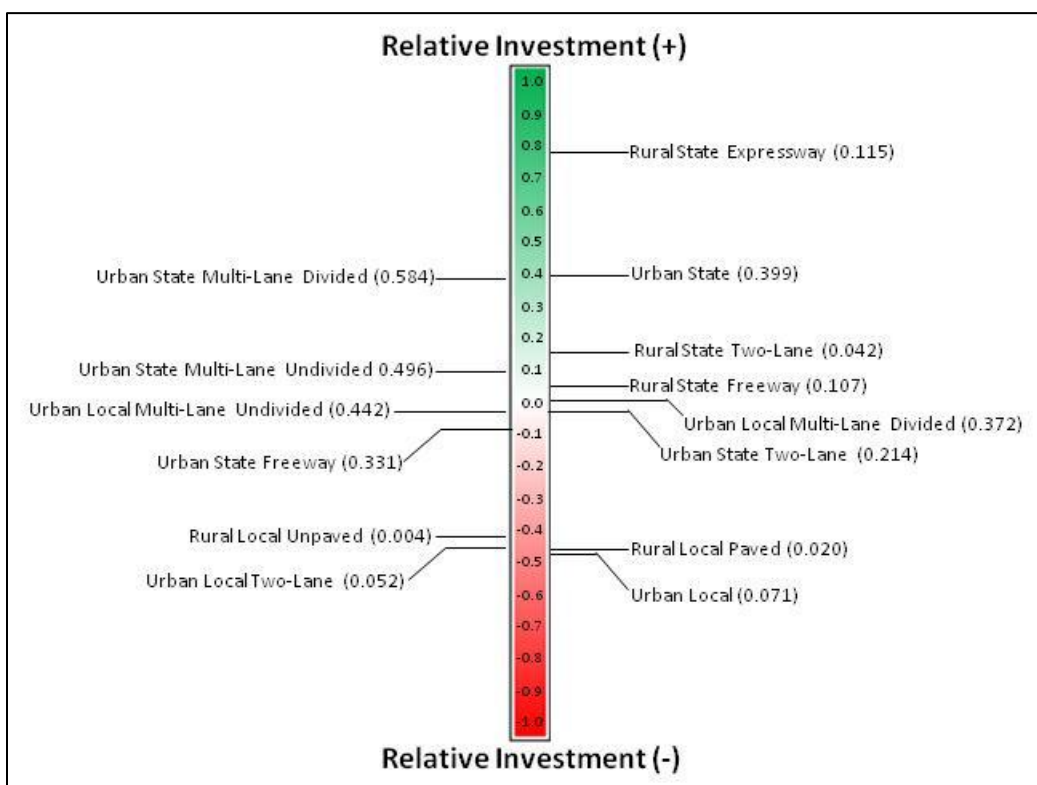


Figure 2-14. Relative safety investment for Iowa roadway classifications (crash densities show in parentheses).

This begs the question, if safety investments are made based on crash density, would it ever be more practical to invest safety dollars on roadway systems with many crashes spread over many miles of roadway like the classification analysis suggests? Consider this hypothetical example:

US Highway 218 directly south of Interstate 80 in Johnson County is a four-lane divided expressway. This 15.5 mile section of roadway has previously received funding from the HSIP program for paved shoulders and rumble strips. Assume that there is still a crash problem and the recommended mitigation is to flatten the sideslopes of the roadway from 1:4 to 1:6. Also consider a proposal to add milled-in rumble strips on 8 corridors, totaling 84.5 miles all on secondary, two-lane paved roadways. All of these corridors are from the 2009 Iowa 5 percent report and labeled as corridors with the highest fatal and serious injury crash density for single vehicle run-of-road crashes. The expressway project has a crash density of 4.7 crashes/mile/year compared to an aggregate crash density for all eight secondary roadways of 1.0 crashes/mile/year. Therefore if considering crash density, it would be recommended that the expressway project be completed over the secondary roadway projects.

The following assumptions were made for the sideslope flattening project:

- Assume other mitigations have been done and flattening the sideslope is the preferred option
- Assume a constant and typical slope throughout the 15.5 mile segment
- Assume the unit cost of fill/flattening is \$3.08 CY (Iowa DOT Bid Express)
- Assume flattening applies to roadsides and median
- Assume each slope is 24', therefore on average, need 3.00 yd² to flatten slope from 1:4 to 1:6
- Assume total CY of fill/flattening need is 327,360 CY
- Assume estimated cost of crash by severity as per National Safety Council, 2009

The following assumptions were made for the rumble strip projects:

- Assume roadways meet the minimum requirements for edge-line rumble strip projects in Iowa
- Assume the unit cost of milled-in rumble strips is \$658/mi/shoulder (Iowa DOT Bid Express)
- Assume a constant and typical slope throughout the 15.5 mile segment

Table 2-1 shows a comparison of these two proposed projects. The rumble strip projects on the rural, secondary, two-lane corridors yields a benefit/cost ratio (42.4) much higher than that of the rural expressway sideslope project (2.6). This is a good example of a project on a roadway with a lower crash density that should be implemented in lieu of a project on a high crash density roadway. Investing safety funds where it can make a difference is most prudent.

Table 2-1. Comparison of rural expressway sideslope flattening project and rural secondary two-lane rumble strip projects.

Project Type	Mileage	Average Crashes Per Year (2001-2008)						CMF ^a	Average Crashes Mitigated Per Year						Estimated Benefit ^b	Estimated Cost ^c	Estimated B/C
		K	A	B	C	O	Total		K	A	B	C	O	Total			
Rural State Expressway - Slope Flattening Project	15.5	0.5	2.125	8	8	54.13	72.75	0.76	0.38	1.62	6.08	6.08	41.1	55.29	\$2,578,984	\$1,008,269	2.6
Rural Secondary Two Lane -Rumble Strip Projects	84.5	1.13*	4.25*	6.00*	9.63*	12.75*	33.75*	0.74	0.83	3.15	4.44	7.12	9.44	24.975	\$4,717,084	\$111,202	42.4

*SVROR crashes only

^aSource: CMF Clearinghouse

^bSource: National Safety Council (estimated cost of crash)

^cSource: <http://www.bidx.com>

2.5.5 Black spot analysis vs. mass action analysis

Another consideration needed in the funding allocation process is the method utilized to identify and prioritize sites, corridors, or even systems for potential safety investment. The data identify the most common fatal and serious injury, intersection crash type as right angle. The most common fatal and serious injury, non-intersection crash type is single

vehicle run-off-road. Right angle crashes account for approximately 25 percent of all fatal and serious injury crashes from 2001 to 2008, while SVROR crashes account for 33 percent all fatal and serious injury crashes over the same period.

Historically, as identified by the literature, black spot analysis is the “most common method to identify candidate locations for safety investment” (p. 3, Preston et al., 2010). The Iowa 5 Percent Report and SICL List are two examples of the use of black spot analysis to identify and prioritize intersections and corridors for safety investment. These two processes are important and integral in addressing highway safety in Iowa but upon further inspection they only address a fraction of statewide crashes.

Table 2-2 compares the intersection crashes at the top 200 intersections, as prioritized by the SICL list for 2003-2007, to all statewide and intersection crashes for the same period. In Iowa there are approximately 160,000 intersections, therefore 17 percent of all fatal and serious injury crashes occur at the top 0.125 percent of all intersections. This is a very substantial number but it also means that 83 percent of all intersection crashes are spread across the other 150,800 intersections.

Table 2-2. SICL list crash comparison.

Top 200 Intersection Crashes (2003-2007)			
Number of crashes occurring at the top 200 intersections	8264	2.8%	of all crashes
		8.3%	of all intersection crashes
Number of K+A crashes occurring at the top 200 intersections	484	4.8%	of all K+A crashes
		17.0%	of all K+A intersection crashes

Table 2-3 compares the SVROR fatal and serious crashes occurring on the top 33 corridors, as prioritized by the 2010 Iowa 5 percent report, to all SVROR fatal and serious injury crashes occurring between 2001 and 2008. The total mileage of these corridors is approximately 345 miles. The state of Iowa has approximately 115,000 miles of roadway statewide, with about 105,300 miles of which is rural. Therefore, approximately 4.1 percent of all SVROR crashes occur on only 0.3 percent of the total roadway network. This also means that over 4,527 fatal and major injury SVROR crashes are distributed over about 105,000 miles of rural roadways.

Table 2-3. Iowa 5 percent crash comparison for SVROR crashes.

Top 33 SVROR Corridors by K+A Crash Density			
Number of K+A crashes occurring at highest 33 SVROR crash density corridors	195	1.2%	of all K+A crashes
		4.1%	of all K+A SVROR crashes

Using a high crash location approach to mitigate these widely distributed crashes is not effective because it will not yield many locations that exhibit unusually high crash frequencies or crash rates. The only way to address these widely distributed crashes is to use a systematic approach. This does not suggest, however, that black spot analysis not be utilized; rather it suggests that there needs to be a balance between the two methods.

2.6 CONCLUSIONS AND RECOMMENDATIONS

Several questions were addressed through the classification of statewide crash data. It appears that not all of Iowa's roadway system elements are equally at risk. For example, some facility types, such as state and local two-lane rural roadways are more at risk for single vehicle run-off-road crashes. The results of the matching of crash data with safety project funding data suggest the shifting of funds from the high crash density state system to facilities on the low density local system. However, it is clear that the redistribution of funds, from one system to another, includes many other factors such as crash density, benefit-cost, and other political issues.

The allocation of funding should also identify what mitigations have been implemented and what additional options are available to maximize safety spending. It is very possible for a safety project on a roadway with a lower crash density to be more effective than a project on a roadway with a very high crash density, depending on the projects and their benefit/cost ratios. Crash reduction factors and benefit cost analyses are integral in aiding the safety funding decision making process as well. Ultimately, the optimum allocation of resources would reduce the most possible fatal and serious injury crashes.

Some crashes are too widely distributed over many miles of roadway to be identified as possible sites in need of safety mitigation. It was recommended that the highway safety

process include both reactive (black spot) approaches as well as proactive (mass action) approaches. There should be a balance among these two methods. This optimum balance between black spot and mass action is yet to be determined. It is recommended that this balance of black spot and mass action be addressed in future research.

CHAPTER 3. SYSTEMWIDE IDENTIFICATION OF HORIZONTAL CURVES AND GEOMETRY PARAMETERS

3.1 INTRODUCTION

In order to analyze the safety performance of horizontal curves and mitigate associated crash problems, curve locations and characteristics must be known. However, curve identification is difficult on a large system. In Iowa, rural horizontal curves comprise of only 1.2 percent of the state's total of 115,335 miles. Yet 10.5 percent of the state's fatal crashes occur on these roadways.

Table 3-1 shows this over representation of fatal crashes on horizontal curves in Iowa. Nationwide more than 25 percent of fatal crashes are associated with horizontal curves (FHWA). Crash rates for horizontal curves are typically 1.5 to 4 times higher than the crash rates of tangent highway sections (Zegeer, et al., 1992). Figure 3-1 shows the distribution of horizontal curves on paved, two-lane, rural highways in Iowa.

Table 3-1. Iowa statewide crash comparison for horizontal curves (2001-2009).

	Crashes on Curves	Rural Crashes	Statewide Crashes	Curve Crashes/Rural Crashes	Curve Crashes/Statewide Crashes
Fatal	353	2437	3355	14.5%	10.5%
Injury	5384	50584	153362	10.6%	3.5%
PDO	5861	108443	370195	5.4%	1.6%
All	11598	161464	526912	7.2%	2.2%

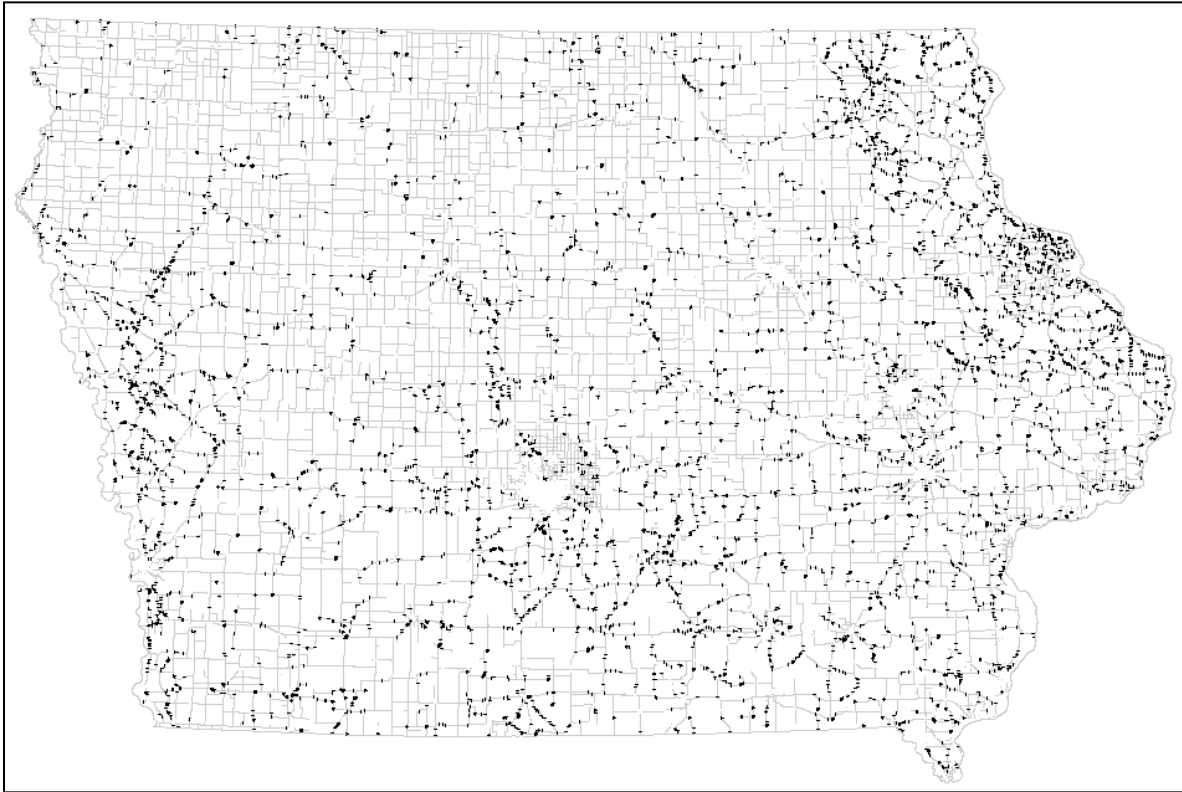


Figure 3-1. Two-lane horizontal curve distribution with paved two-lane, rural roads shown in gray.

Recently the Iowa DOT has addressed safety concerns on horizontal curves. In 2010, the Iowa 5 percent report identified and prioritized horizontal curves based on crash frequency. However past studies show that prioritization of horizontal curves needs to be based on more than just crash frequency (Preston et al., 2009). Other factors such as curve radii, traffic volume, the presence of visual traps, intersections and proximity to other high priority curves should also be considered (Preston et al., 2009).

The purpose of this chapter was to first create a statewide curve database by systematically identifying horizontal curves on two-lane rural roads in Iowa. Secondly, a validation of the curve identification methods was completed using a sample of curves with as-built geometric data. Lastly, the safety performance of horizontal curves in Iowa was explored using crash prediction models. Chapter 4 of this thesis presents the crash prediction model based on these data.

3.2 REVIEW OF LITERATURE

The most widely used method of curve geometry data collection includes the use of an in vehicle GPS receiver and some form of post-processing. Patterson, D., et al. (2006) used GPS and GIS applications to collect and analyze horizontal curve geometry data. The process included collecting field data at 0.1-second intervals using differential GPS surveying in a vehicle. Results demonstrated that GPS could quickly, accurately, and inexpensively produce horizontal alignment data.

Pratt, et al. (2009) used a similar method to collect curve geometry while driving through a curve. A GPS receiver was used to collect curve radius and deflection angle data. An electronic ball-bank indicator was used to gather superelevation data. These two instruments were directly connected with a laptop, and, while driving, curve data were simultaneously compiled into an in house software package called the Texas Roadway Analysis and Measurement Software (TRAMS). Data were collected at 25 foot increments, and curve radii were calculated. These data were then post-processed to calculate a recommended advisory speed. Results showed that this method provided an accurate and precise measurement of curve radius. However, these methods are impractical for collection of statewide curve data.

Sanders (2007) provided a methodology for statewide data collection of horizontal curves using GPS centerlines. GPS data for over 79,000 centerline miles of roadway were collected in Kentucky. An automated process was developed using GIS to extract curve data and determine roadway geometry. Results showed this GPS/GIS method provided much more accurate curve data than previous field-collected processes.

3.3 DATA

3.3.1 Roadway data

In this study, Iowa statewide road data were obtained from the 2007 GIMS roadway database. The complete GIMS database was reduced because the focus of this study is rural two-lane facilities. Rural, two-lane facilities with a speed limit of 45 mph or greater were extracted from the complete roadway dataset.

3.3.2 Calculated curve data

Calculated curve data were obtained through a manual identification process discussed in the methodology section of this chapter. The GIMS 2007 database was used for roadway attributes. The Iowa Pavement Management Program (IPMP) provided GPS traces of the state's roadways, originally obtained from in-vehicle GPS data collectors. The data consist of points at ten meter increments along all paved routes throughout the state.

3.3.3 As-built curve data

As-built data were required for the validation of the calculated curve data geometry. As-built curve data were identified using the Iowa DOT's Electronic Records Management System (ERMS). ERMS contains historic roadway plans for all primary road projects in Iowa. Secondary (county) road data were not available in the ERMS and were not included in the study.

Curve data available in the historic roadway plans were manually extracted for a specific set of counties. Data for 435 horizontal curves were identified in 15 counties throughout Iowa. Figure 3-2 shows the counties in which horizontal curve data were identified. Horizontal curve data were collected on paved two-lane rural roadways with a speed limit of 45 mph or greater to match the roadways used to identify curves in the created horizontal curve database. Curve data were extracted by county because roadway plans in ERMS are stored by county. Counties were chosen to yield a sample that is topographically diverse and geographically dispersed throughout the state.

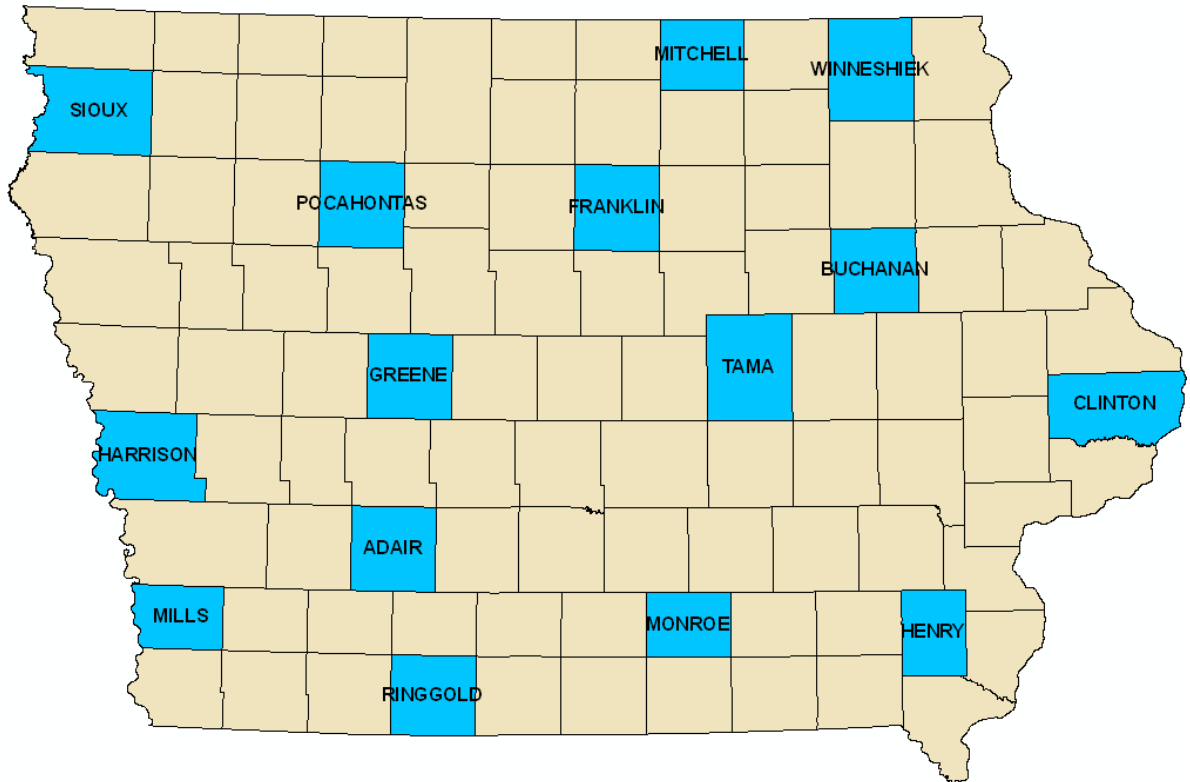


Figure 3-2. Plan set curve data locations.

3.4 METHODOLOGY

3.4.1 Curve identification

Horizontal curves were identified with the use of GIS tools. To limit the extent of required visual inspection and to more systematically identify possible locations of horizontal curvature, polylines were first created, and later simplified, from the available IPMP GPS traces. The remaining vertices in the simplified polylines primarily represented the locations of route termini and, more importantly, possible curvature. GPS traces from IPMP, the GIMS roadway network and the simplified polyline vertices were then added into an ArcGIS workspace for visual inspection. Aerial imagery was also used as a reference, where necessary.

Once a reviewer identified a section of roadway as a horizontal curve, the GPS traces were selected and manually coded as being a part of a curve. Post-processing was then performed on the GPS traces to extract only the “curve” records, combine the points for each

curve, attach a unique identifier, and calculate the curve geometry. After all curves were identified, they were compiled and reintroduced to the GIS environment for analysis. Figure 3-3 shows a roadway (black line), the location of the GPS traces (small green dots) and the location of the simplified polyline vertices (large green dots). The red line shows the final approximate location of the curve relative to the tangents.

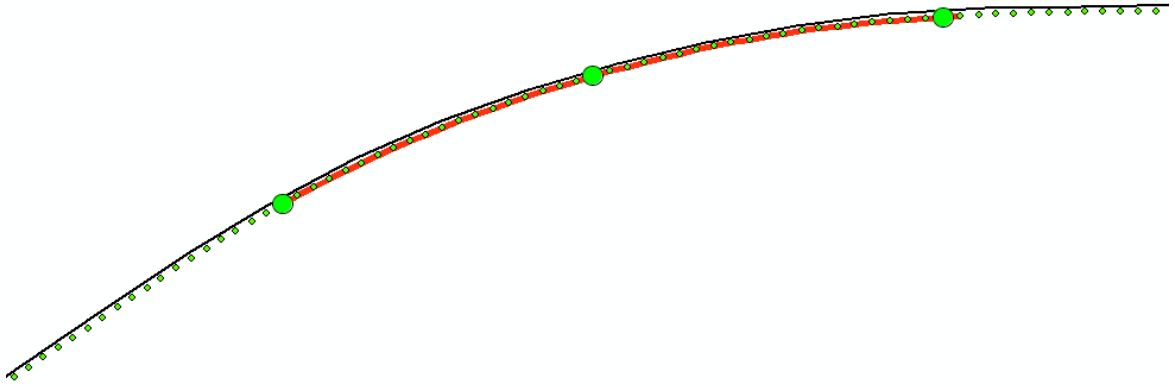


Figure 3-3. Curve identification process with GPS traces and simplified polyline vertices.

Spiral transition data were not collected. The reviewer identified possible locations of roadway curvature only. Every curve was assumed to be a circular curve in order for the radius value to be estimated using circular regression during post-processing. Therefore, if a curve was located it was assumed to be a circular curve.

Estimated curve geometry included length, degree of curvature, and two different radius values. The first curve radius value was calculated using circular regression. Once a curve was identified, a circle was fitted through each GPS trace of the curve to find a best fit. The radius of that circle was the estimated radius of the curve. This radius value is referred to as $R_{\text{regression}}$. In order for a curve to be fit, a curve was required to have at least five GPS traces associated with it. If a curve did not have five traces associated with it, the post-processing would not work properly and the $R_{\text{regression}}$ value would be estimated as zero.

The second radius value was calculated using the long chord of the curve. A straight line was fitted through the reviewer's estimated point of curvature (PC) and point of tangency (PT) for each curve. The radius value was then estimated using the long chord and the angle between the long chord and curve. This radius value is referred to as R_{chord} . For

simplicity and mass production the processes used to estimate both radius values were performed using a macro program in Microsoft Excel.

3.4.2 Curve identification validation process

In order to validate the curve identification process, as-built horizontal curve data needed to be compared to the estimated curve data for a sample of curves. As previously mentioned, as-built curve data were extracted from historic roadway plans using ERMS and compiled. A unique, as-built identifier, different from the unique identifier for the curve identification process, was given to each curve. The location of the horizontal curves was then found using Google Map tools. The unique, as-built identifier was then matched and attached to its corresponding horizontal curve data using GIS tools. Once this process was completed, as-built curve data could be compared with the estimated curve data and validated for precision.

Percent RMSE was used as a measure of precision to validate the curve identification process. RMSE measures the deviation between the actual geometric feature value (e.g. length, radius), and the estimated geometric feature value. A large percent RMSE indicates a large deviation between the actual and estimated values.

Equation 3-1:

$$\% RMSE = \frac{100 * \sqrt{\frac{\sum_j (Estimated_j - Actual_j)^2}{(Count - 1)}}}{\left(\frac{\sum_j Actual_j}{Count}\right)}$$

Estimated_j = calculated geometric feature value (e.g. $L_{predicted}$, R_{reg} , R_{chord})

Actual_j = as-built geometric feature value (e.g. L_{actual} , $L_{adjusted}$, R_{actual})

Count = number of horizontal curves

Further analysis was completed to investigate the relationship between percent error and geometric features as well as how well the sample represents the entire database. In order to explore the representation of roadway curves, the coincidence ratio was compared between the histogram of the sample of curves with as-built data and the histogram of all

primary roadway curves, as well as the histogram of entire population of curves (primary and secondary curves). Histograms were compared for curve length and both radius estimation methods. The estimated curve length values were divided into “bins” at 100 foot increments while both estimated radius values were divided in increments of 250 feet.

The coincidence ratio compares two distributions and measures the percentage of total area in common between the two distributions. It should be noted that the coincidence variable is only used to check whether or not the sample is representative of the entire population and is not a measure of precision. Equation 3-2 shows the formula for calculating the coincidence ratio.

Equation 3-2:

$$\text{Coincidence Ratio} = \frac{\sum \min \left\{ \frac{f^e(j)}{F^e}, \frac{f^a(j)}{F^a} \right\}}{\sum \max \left\{ \frac{f^e(j)}{F^e}, \frac{f^a(j)}{F^a} \right\}}$$

$f^e(j)$ = frequency of geometric values (e.g. $R_{\text{regression}}$) with radius value j from sample of curves

$f^a(j)$ = frequency of geometric values (e.g. $R_{\text{regression}}$) with radius value j from population of curves (or primary roadway curves)

F^e = total number of geometric values (e.g. $R_{\text{regression}}$) in sample of curves (329)

F^a = total number of geometric values (e.g. $R_{\text{regression}}$) in population of curves (11,279) or primary roadway curves (2,349)

3.5 ANALYSIS

11,882 curves were identified during the curve identification process. If a curve did not have at least five GPS traces assigned to it the $R_{\text{regression}}$ value would be estimated as zero. Of the 11,882, 603 had less than five GPS traces and therefore an $R_{\text{regression}}$ value of zero. Because of the zero value for the $R_{\text{regression}}$, these curves were removed from the analysis. Therefore a total of 11,279 curves were included in the statewide horizontal curve database.

Table 3-2 gives the number of horizontal curves identified and the average curve geometry for the primary and secondary systems. As expected, curves on secondary

roadway curves, tend to be shorter in length and have a sharper radius.

Table 3-2. Identified curves by system type.

Roadway System	Number of Curves	Average Curve Geometry		
		L (ft)	R _{regression} (ft)	R _{chord} (ft)
Primary	2349	870	2162	2078
Secondary	8930	576	1158	1136
All	11279	637	1367	1332

Using the historic roadway plans, as-built data were identified for 435 curves on the primary system. Of these, 329 were matched with curves from the statewide curve database. The curves that were not matched were mostly large radius curves that were so large they were not identified as curves in the manual identification process. A few curves were matched, but upon further inspection were found to have incorrect data. These curves were omitted.

3.5.1 Horizontal curve length estimation

Figure 3-4 shows the actual horizontal curve length plotted against the estimated horizontal curve length. The curve length was estimated during the curve identification process by manually estimating the PC and PT locations. The percent RMSE of the curve length is 29.03 percent. This magnitude error was expected for the curve length because of the way curve lengths were estimated. It is difficult to identify exact location of where the curve ends and begins. Furthermore there appears to be a systematic bias to underestimate the curve length. These sources of error are discussed further in the “Sources of error” section of this chapter.

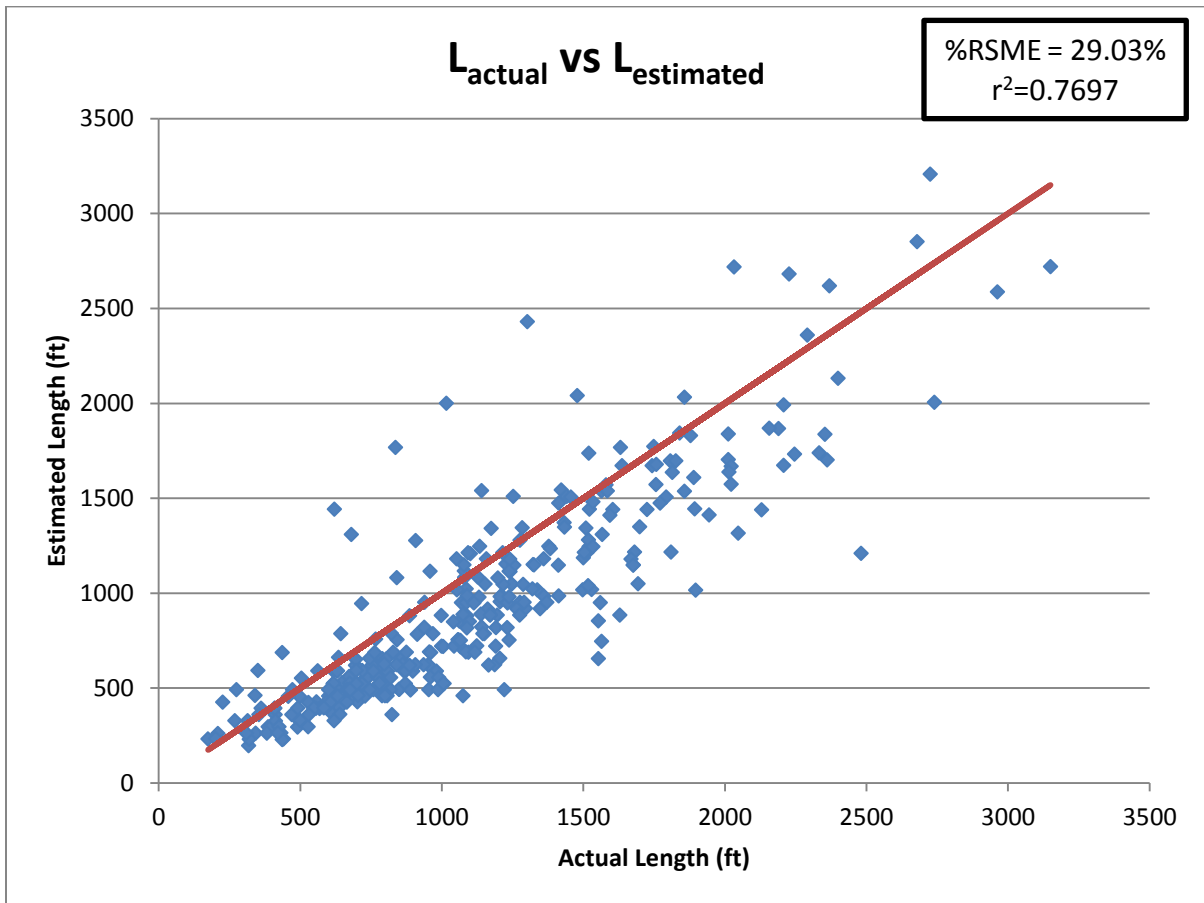


Figure 3-4. Actual curve length vs. estimated curve length.

The absolute, percent error between the estimated curve length and as-built curve length versus the as-built curve length is plotted in Figure 3-5. One would expect as the curve length increases, the percent error would decrease. However, the percent errors do not necessarily follow this form although all large errors (>60 percent) are on curves shorter than 1,500 feet in length. Eighty-seven percent of all estimated curve length data have a percent error of less than 40 percent.

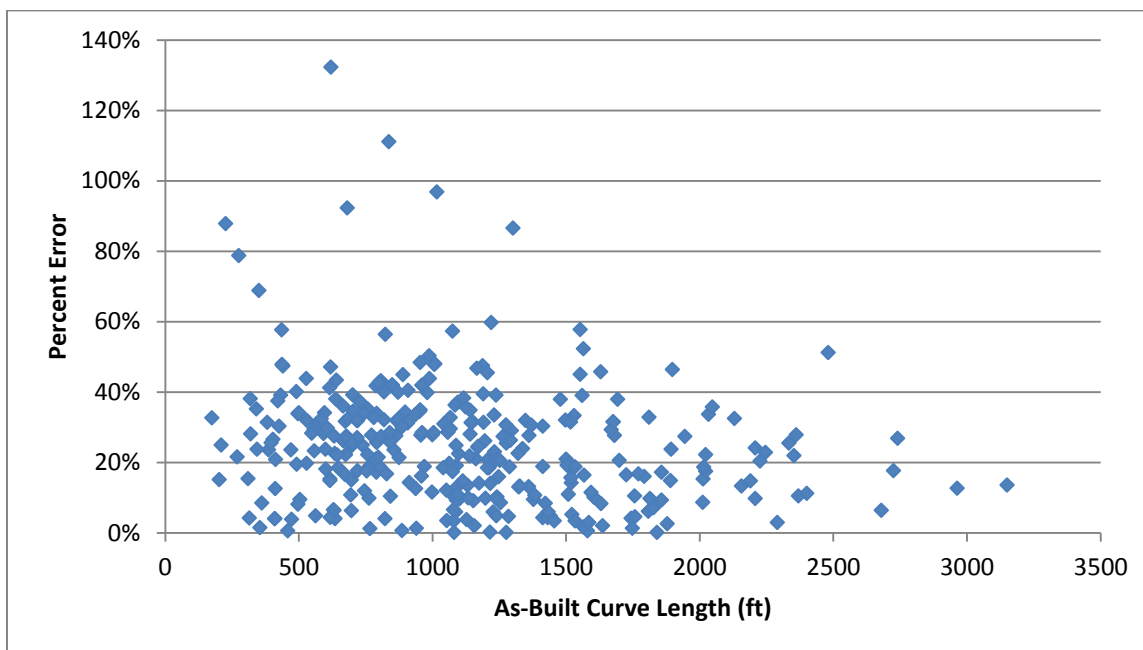


Figure 3-5. Curve length versus percent error of estimated length value.

Figure 3-6 shows both the curve length histogram for the sample of curves with as-built data and the curve length histogram for all curves identified on primary roadways. Comparing the two histograms yields a coincidence ratio of 0.826, indicating that 82 percent of the primary roadway data length values are described by the sample of curves with as-built data.

When comparing the curve length histogram for the sample of curves to the curve length histogram of the entire population of curves (primary and secondary roadways) the coincidence ratio decreases to 0.580. Figure 3-7 displays both the curve length histogram for the sample of curves data and the curve length histogram for the entire population of identified curves.

The reason for this decrease is due to the inclusion of secondary roadway curves. Since the sample of curves data contain only data from primary roadway curves, the sample of curves is a better representative of the primary roadway curve data. Secondary roadways, on average have shorter length curves and therefore the population of curves histogram is skewed towards shorter length curves.

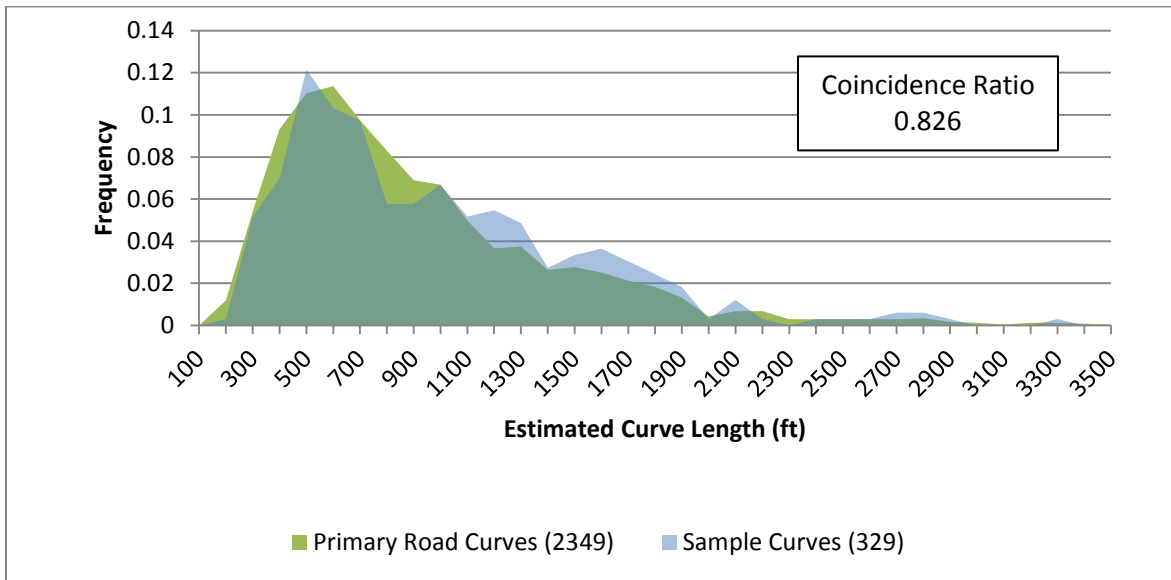


Figure 3-6. Curve length histogram comparison for sample curves and primary roadway curves.

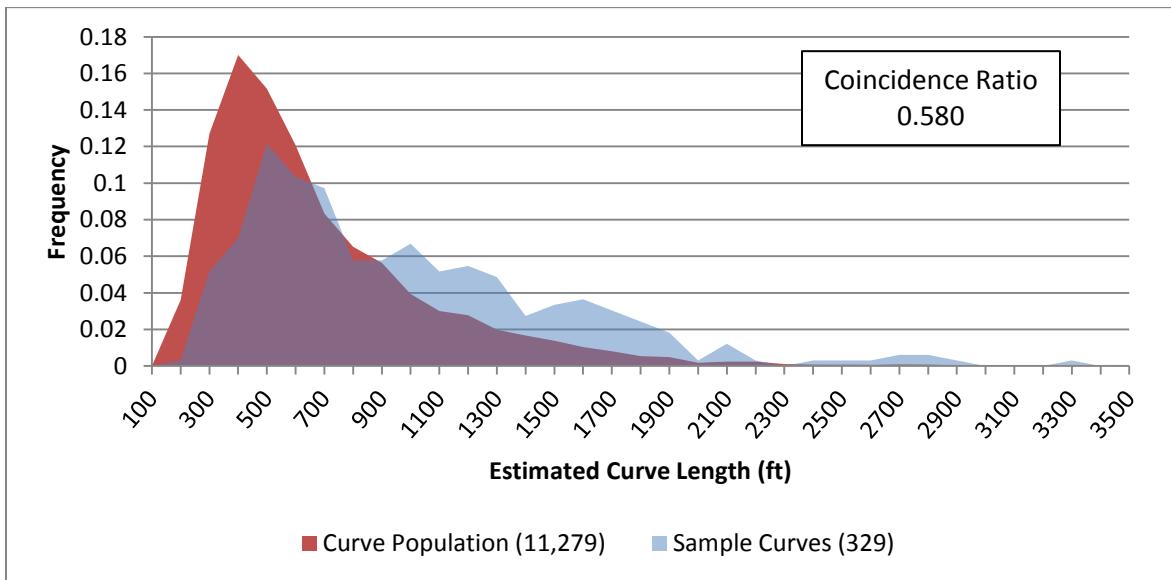


Figure 3-7. Curve length histogram comparison for sample curves and all curves in database.

3.5.2 Horizontal curve radius estimation using the circular regression method

Figure 3-8 shows the actual, as-built curve radius plotted against the calculated curve radius using the circular regression method, $R_{\text{regression}}$. The circular regression method is

relatively precise, with a percent RMSE of only 16.27 percent and a coefficient of determination of 0.9257.

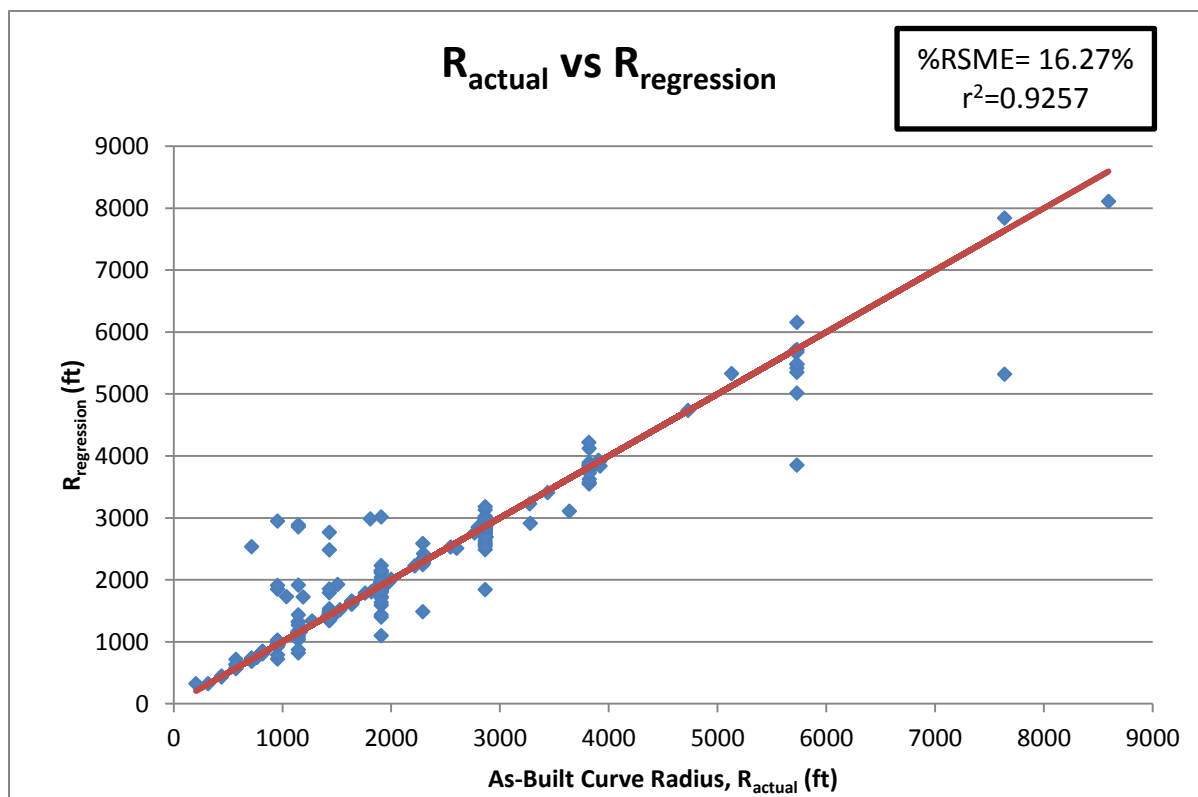


Figure 3-8. Actual radius vs. circular regression method estimated radius.

The percent error between $R_{\text{regression}}$ and the as-built curve radius is plotted against the as-built curve radius in Figure 3-9. All large errors are associated with curves with radii less than 1,500. It appears that as curve radius increases, percent error decreases. On average, $R_{\text{regression}}$ is very precise with 94 percent of all curves with as-built data have an $R_{\text{regression}}$ percent error less than 30 percent. Moreover, over 75 percent of all curves with as-built data have an $R_{\text{regression}}$ percent error equal to or less than 5 percent.

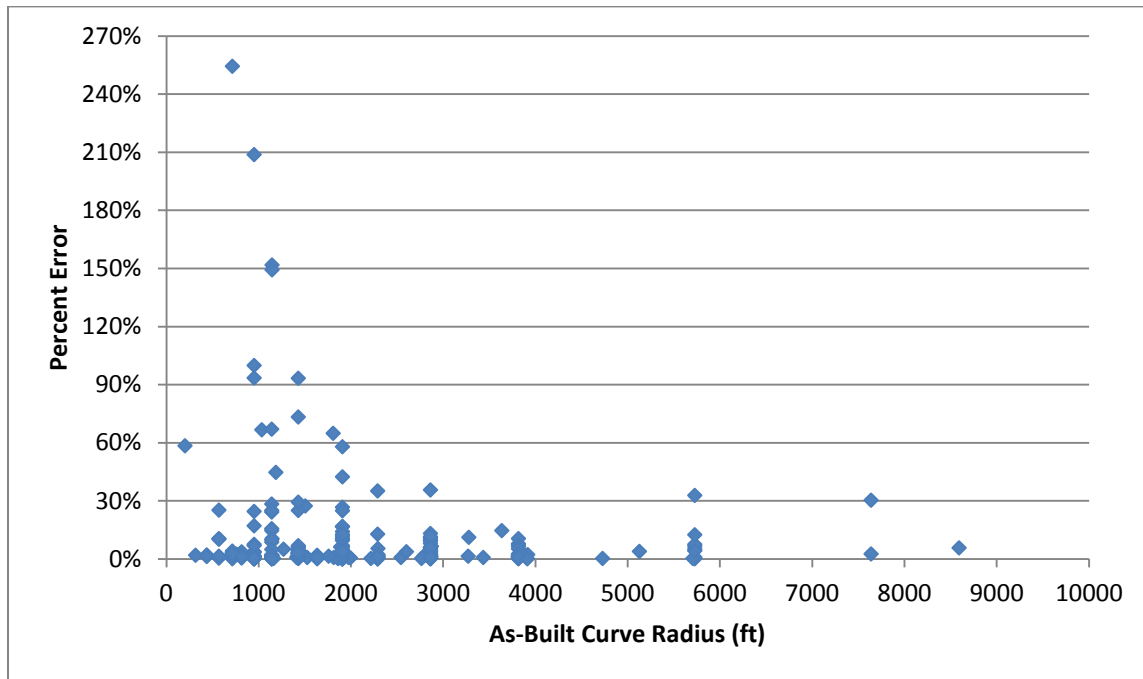


Figure 3-9. Curve radius versus percent error of estimated radius value, $R_{\text{regression}}$.

Figure 3-10 displays the curve radius ($R_{\text{regression}}$) histogram for both the sample of curves with as-built data and the histogram for all curves identified on primary roadways. The distribution of the sample of curves data coincides with 76.5 percent of primary roadway curves distribution. The $R_{\text{regression}}$ histogram for the sample of curves coincides with only 46 percent of the $R_{\text{regression}}$ histogram for the entire population of curves. Figure 3-11 shows the comparison of these histograms. An explanation for this decrease in coincidence ratio is similar to that of the curve length. The inclusion of secondary roadway curves skews the histogram towards smaller radius curves because, on average, secondary roadway curves tend to be sharper.

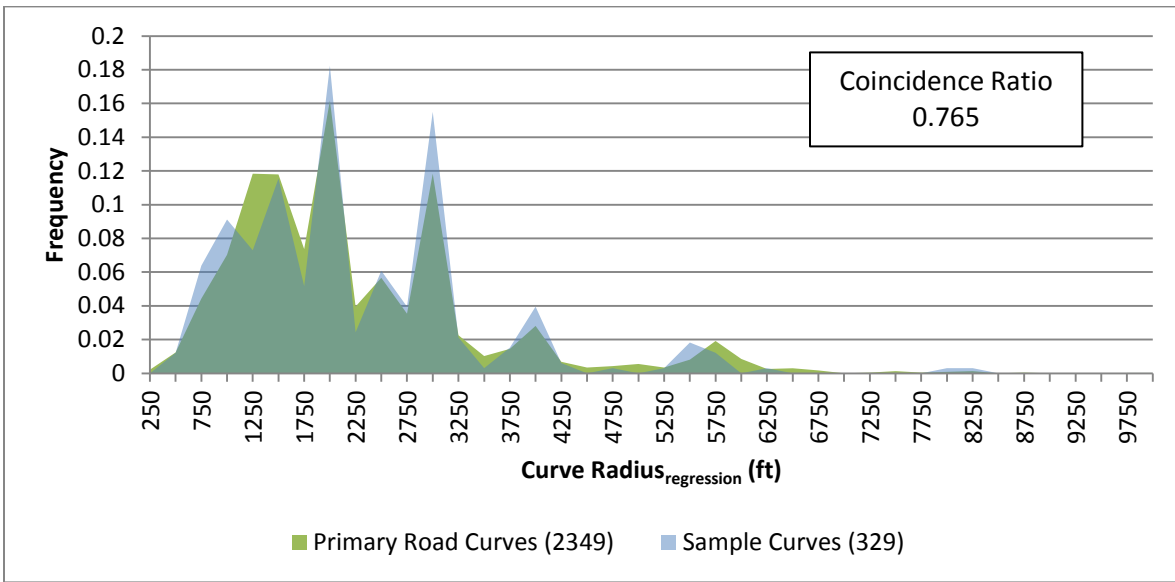


Figure 3-10. Curve radius ($R_{\text{regression}}$) histogram comparison for sample curves and primary roadway curves.

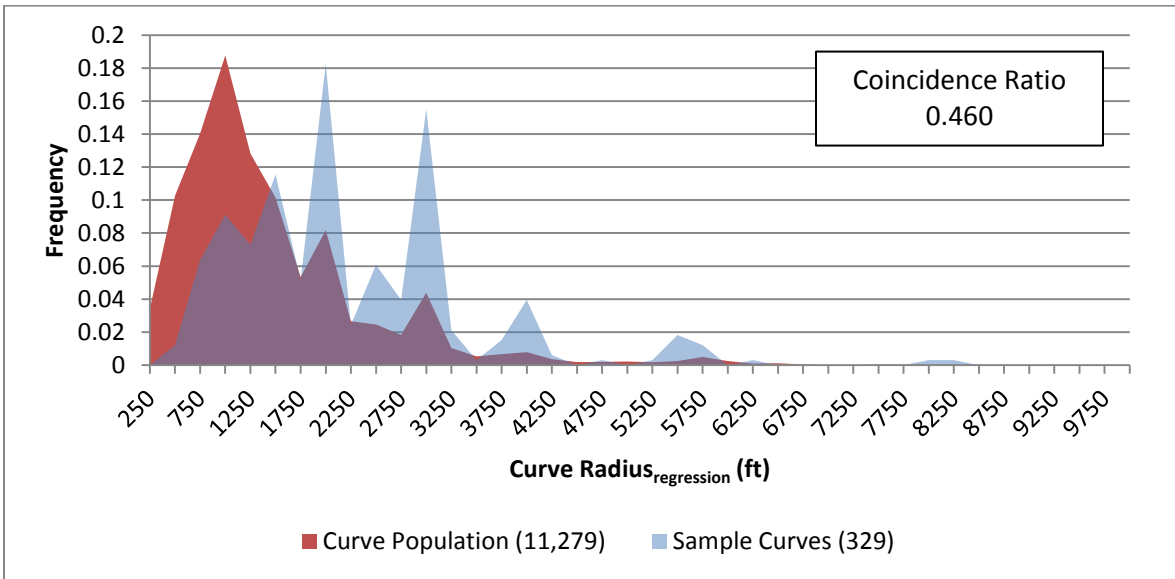


Figure 3-11. Curve radius ($R_{\text{regression}}$) histogram comparison for sample curves and all curves in database.

3.5.3 Horizontal curve radius estimation using the long chord method

Figure 3-12 shows the actual, as-built curve radius plotted against the calculated curve radius using the long chord method. The long chord method is just slightly less precise than the circular regression method, with a percent RMSE of 19.45 percent and coefficient of determination of 0.9016.

There appears to be a systematic bias to underestimate the radius value in larger radius curves (>5000). Possible explanations for this underestimation are described in the forthcoming sources of error section of this chapter.

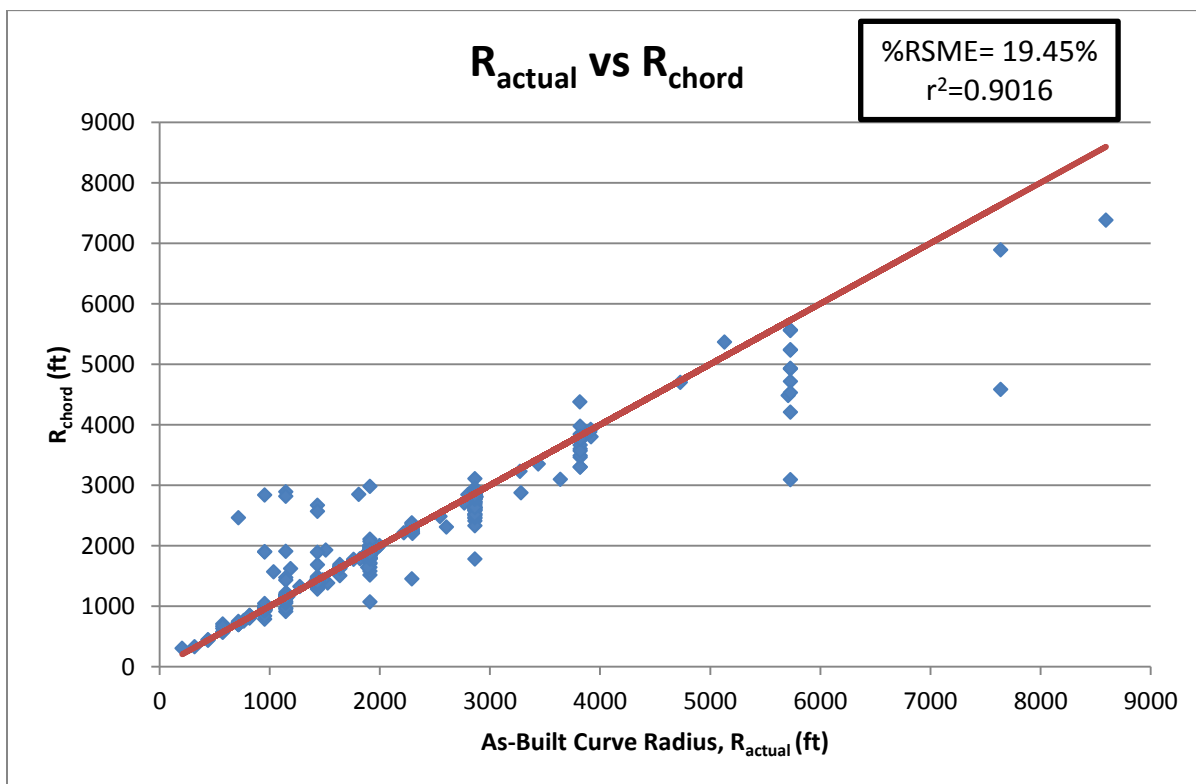


Figure 3-12. Actual curve radius vs. long chord method estimated curve radius.

Figure 3-13 displays the percent error between R_{chord} and the as-built curve radius plotted against the as-built curve radius. All percent errors greater than 50 percent are associated with curves with radii less than 2,000. As with the $R_{\text{regression}}$ percent error curve, it appears that as curve radius, R_{chord} increases, percent error decreases. Ninety-five percent of

all estimated R_{chord} values have a percent error less than 30 percent with 68 percent of the R_{chord} values having a percent error less than 5 percent.

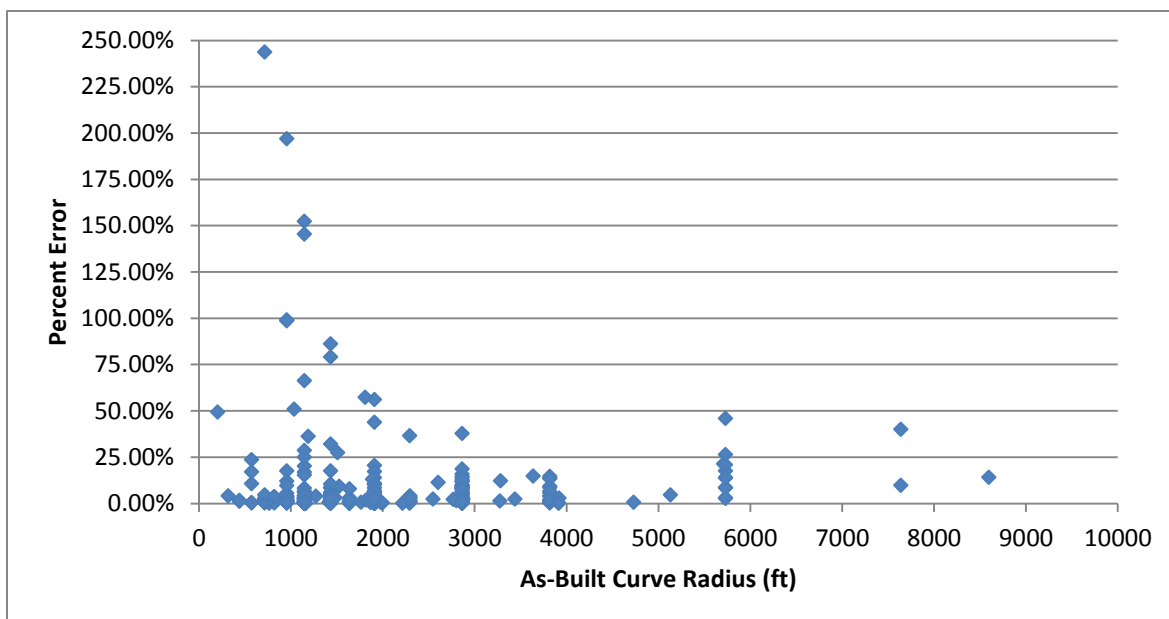


Figure 3-13. Curve radius versus percent error of estimated radius value, R_{chord} .

The distribution of R_{chord} data for the sample of curves coincides with 77.3 percent of the distribution of R_{chord} data for all primary roadway curves. However, when the sample of curves distribution for R_{chord} is compared to the entire population distribution, the coincidence ratio drops to 47.7 percent. Figure 3-14 and Figure 3-15 show the R_{chord} distribution comparison between the sample of curves and the primary roadway curves and the entire population of curves, respectfully. Again the explanation for this decrease in coincidence ratio is similar to that of the curve length and $R_{\text{regression}}$. Again, the inclusion of secondary roadway curves skews the histogram towards smaller radius curves because, on average, secondary roadway curves tend to be sharper.

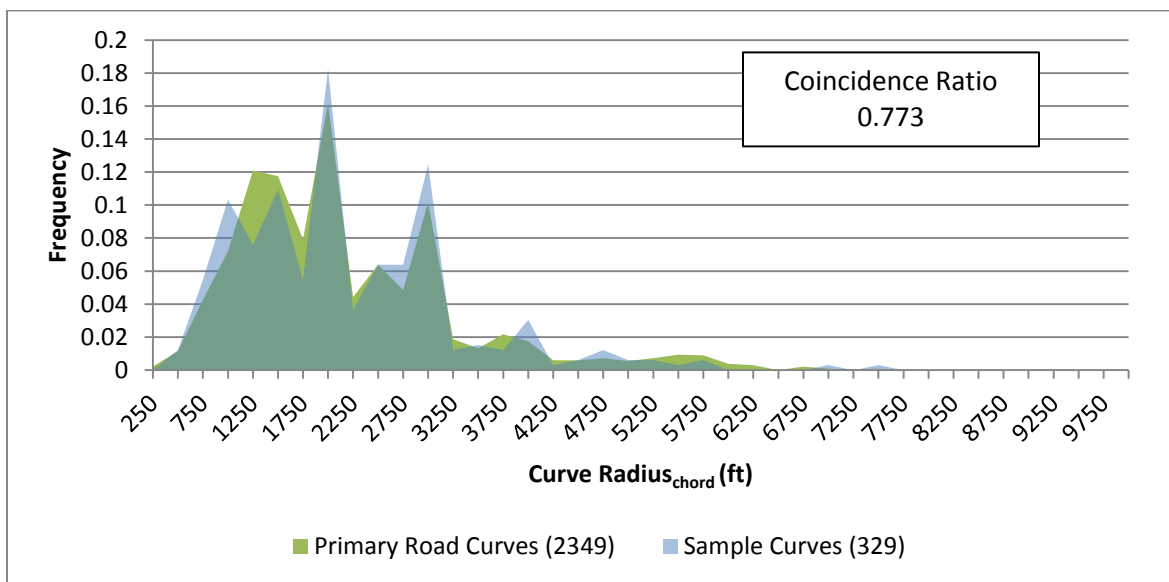


Figure 3-14. Curve radius (R_{chord}) histogram comparison for sample curves and primary roadway curves.

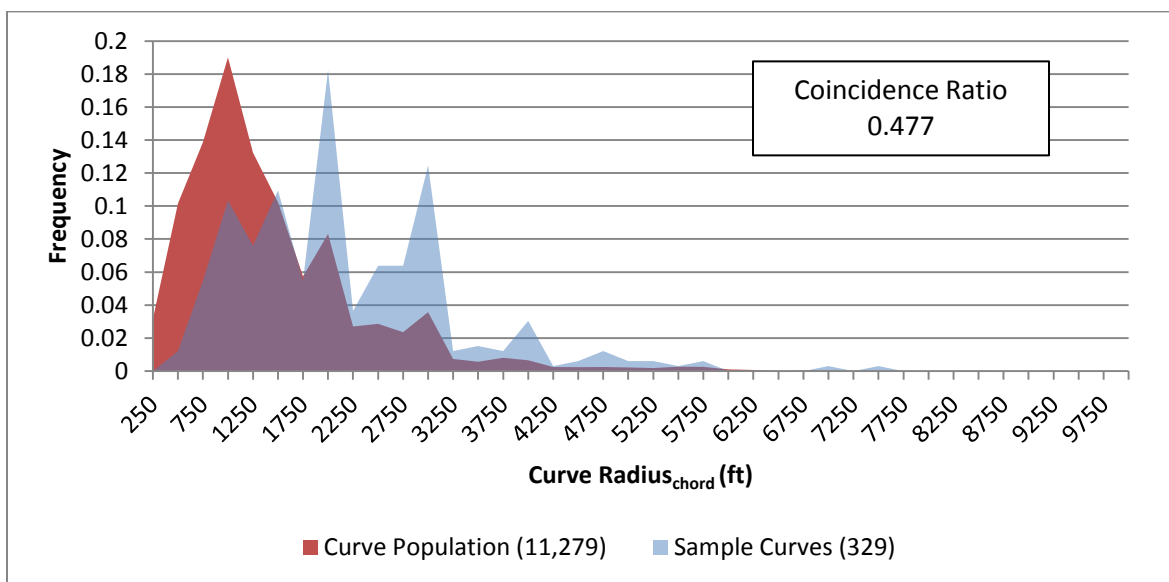


Figure 3-15. Curve radius (R_{chord}) histogram comparison for sample curves and all curves in database.

3.5.3 Horizontal curve radius estimation comparison

This section compares the precision of the two estimated radius values. The data sets for the two radius values were compared using a difference of means test. The difference of means test yielded no significant difference between the two sample sets. The total percent

RMSE for both radius values were very similar, with the $R_{\text{regression}}$ only slightly more precise when compared to the as-built curve radius.

Table 3-3 shows the percent RMSE for both estimated curve radius values for different radius ranges. Figure 3-16 shows these same data only in graphic form. R_{chord} is slightly more precise with lower radius curves (< 2,000 feet). However, it is difficult to distinguish which estimated radius measure is more precision.

Table 3-3. Percent RMSE comparison by curve radius category.

Radius Range (ft)	# of Curves	%RMSE	
		R_{reg}	R_{chord}
< 500	4	19.80%	15.51%
500- 1000	56	48.45%	41.72%
1000- 1500	64	30.77%	28.07%
1500- 2000	82	13.92%	12.45%
2000- 2500	22	8.32%	8.30%
>=2500	101	10.09%	15.81%
All Curves	329	16.27%	19.45%

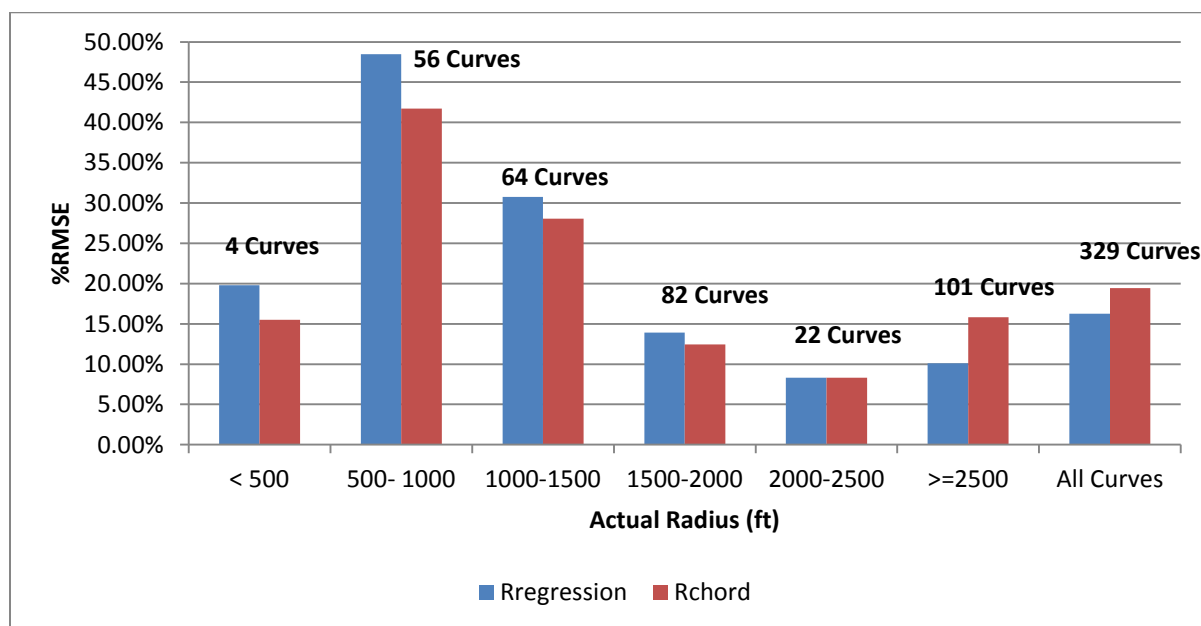


Figure 3-16. Percent RMSE comparison by curve radius category.

3.5.5 Sources of error

There are multiple sources of error that may explain the variation between the calculated radius and length values and the as-built radius and length values. The first source of error is associated with off-tracking of the vehicle during the collection of the GPS traces. Bonneson, et al. (2007) observed that when travelling through a curve, drivers, in order to limit the speed reduction needed to negotiate the curve, laterally shifted in their lane. This lateral shift resulted in a slightly flattened curve radius. Figure 3-17 shows the difference between the curve radius and the vehicle path radius. This same behavior during the collection of GPS data could account for errors between the calculated geometric values and the as-built geometric values.

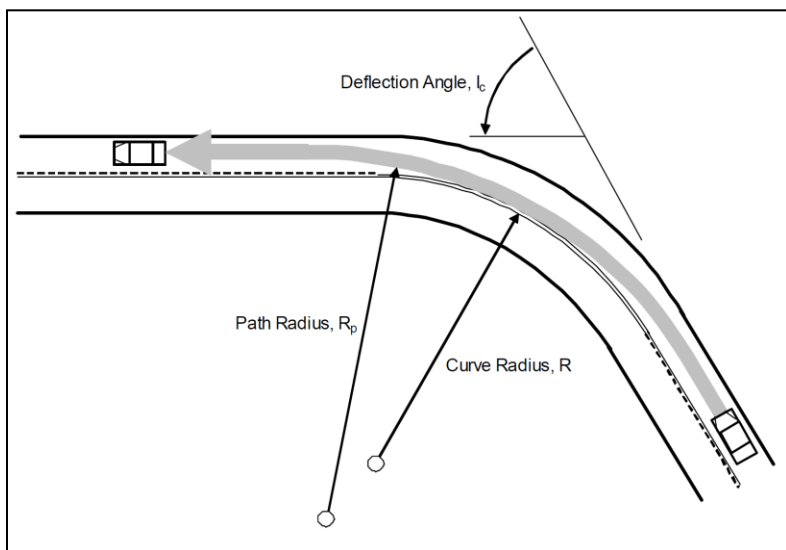


Figure 3-17. Effect of lateral shift on travel path radius.

A second source of error resulting in the flattening of curve radii was also observed. Some horizontal curve PC and PT locations were manually located outside of the curve segment along the tangent sections. Because of this, portions of the tangent sections were used in the curve geometry estimation resulting in estimated curve radii larger than as-built radii. This error was observed in approximately five percent of curves with as built data radii mostly ranging from 1,000 feet to 2,000 feet.

The manual identification of the PC and PT location inside the curve segment was also detected in a large portion of horizontal curves. It was observed that the reviewer was

more apt to estimate the PC and PT within the curve rather than farther out on the tangent. The curve length was underestimated in 82 percent of the 329 curves with as built data. This underestimation of the horizontal curve length explains the larger errors associated with the prediction of the curve length.

Another source of error can be attributed to the post processing of the GPS traces. After the GPS traces are coded as being a part of a curve, they are combined for each individual curve. Each trace has a beginning and ending “location” associated with it. If these points were not correctly ordered, the calculation of the long chord was incorrectly estimated. Figure 3-18 shows the effect of incorrect GPS trace “location” data. This error affected the estimated length of the curve as well as the estimated radius value, R_{chord} .

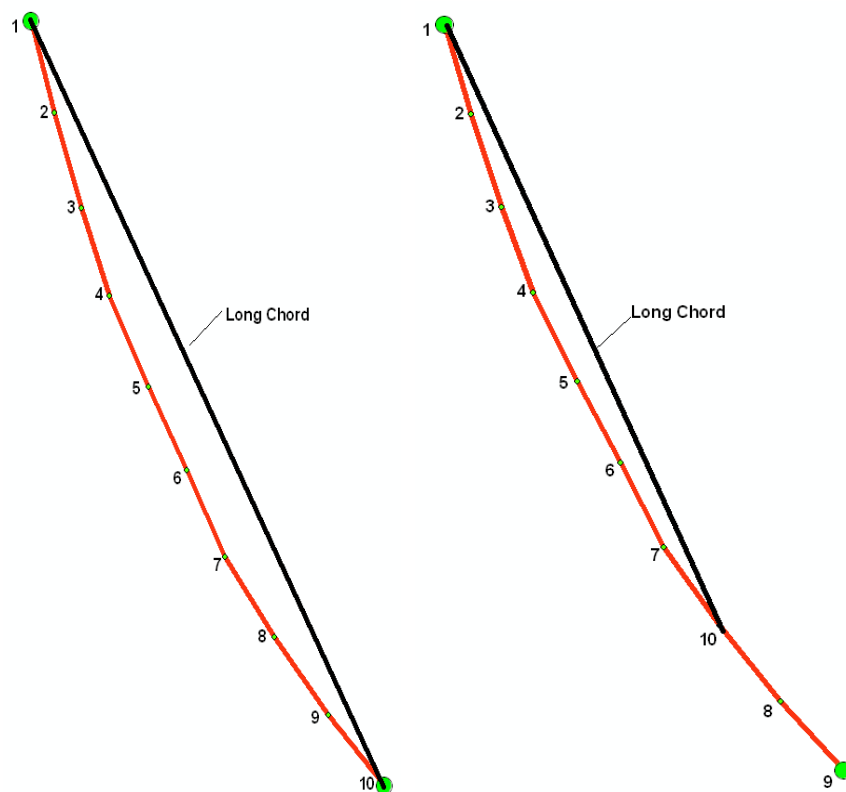


Figure 3-18. Effect of incorrect GPS trace “location” data.

As mentioned previously, in comparing the as-built curve radius and R_{chord} (Figure

3-12) there appears to be a bias to underestimate the radius value. An investigation of this underestimation points to a possible error in the post-processing estimation of the radius value using the long chord method. The long chord method uses the curve length and long chord between the estimated PC and PT to estimate the deflection angle, θ of the curve. When a curve is post-processed, the deflection angle is estimated using a series of iterations. If the deflection angle is underestimated the radius value, R_{chord} is underestimated.

Figure 3-19 and Table 3-4 illustrate an example of this error. Using ArcGIS and the as-built curve plan set, as shown in Figure 3-19, the actual and calculated curves could be scaled and compared. Then using the post-processing spreadsheet, the as-built curve data could be back-calculated to show the error, as shown in Table 3-4. The variables, c and d are the deviation from the actual PC and PT, respectively. Other instances of this error were investigated, and yielded similar results.

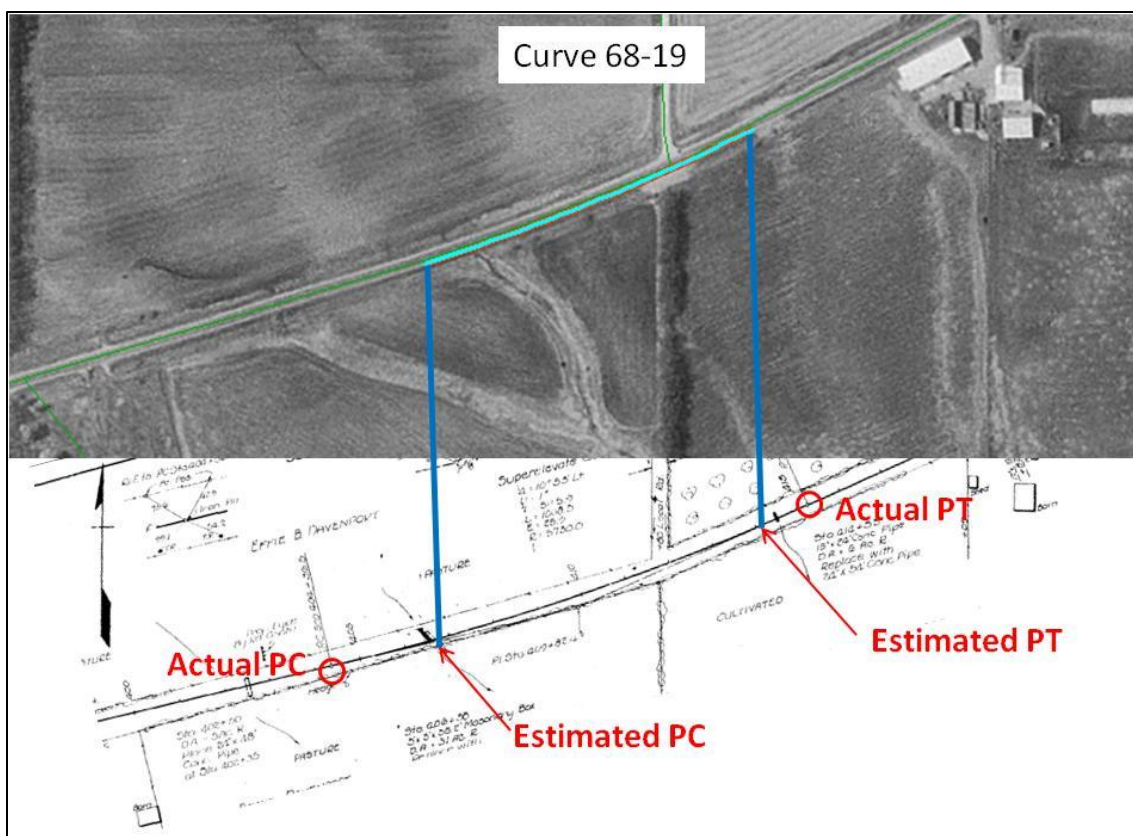


Figure 3-19. Long chord method underestimation example.

Table 3-4. Long chord method calculation example.

	Inputs				7th Iteration θ_7		Outputs			
	L (ft)	LC (ft)	c (ft)	d(ft)			R (ft)	D (deg)	Δ (deg)	T (ft)
Calculated	817.38	816.55	171	100	0.078	Calculated	5240.97	1.09	8.94	409.52
Actual	1088.38	1086.74	-	-	0.095	Actual	5726.60	1.00	10.89	545.83

Lastly, conventional human error could also be a source of error. The majority of the as-built curve data were extracted from roadway plans more than forty years old. These plans are handwritten and there is a large possibility of human error in their creation, as well as the possibility of human error in the extraction of geometric data from these plans.

3.5.6 Sensitivity of errors

Because horizontal curve geometry is estimated, the sensitivity of safety performance to estimation errors is important to identify. Assuming all other roadway attributes are constant and a spiral transition is not present, the expected safety performance of horizontal curves with different radii values could be estimated using Equations 1-1 through 1-3. The safety performance of these curves could then be compared and effects of error could be measured. Figure 3-20 shows the sensitivity of predicted crash frequency to errors in radius estimation process.

The percent change in predicted crash frequency is relatively linear for percent differences of -25 percent and greater. However, the percent change in predicted crash frequency increases exponentially from -25 percent differences in radius value and lower. Furthermore, the percent change in predicted crash frequency decreases as the actual radius increases. In other words, as the actual radius value decreases, the more sensitive expected safety performance is to errors in the estimated radius value. Furthermore, the underestimation of curve radii has a much larger effect on safety performance than overestimation.

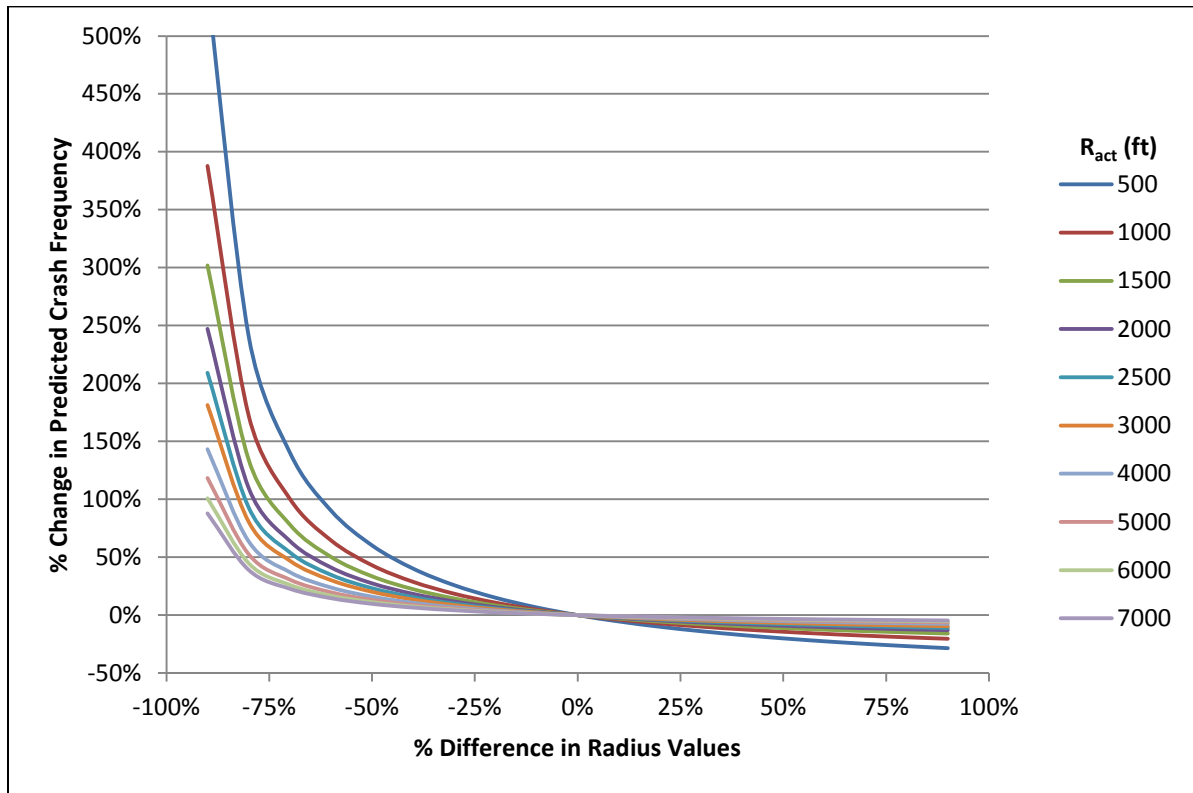


Figure 3-20. Sensitivity of predicted crash frequency to radius percent difference.

Using the circular regression method, approximately 54 percent of the curves with as-built data had an underestimated radius value. However, only 20 radius values were underestimated by greater than ten percent using the regression method. Figure 3-21 displays the sensitivity of predicted crash frequency for 20 horizontal curves with a percent difference in radius values greater than ten percent. Only two underestimated radius values had a deviation from the actual predicted crash frequency of greater than 15 percent. Furthermore, only four radius values had a deviation of greater than ten percent.

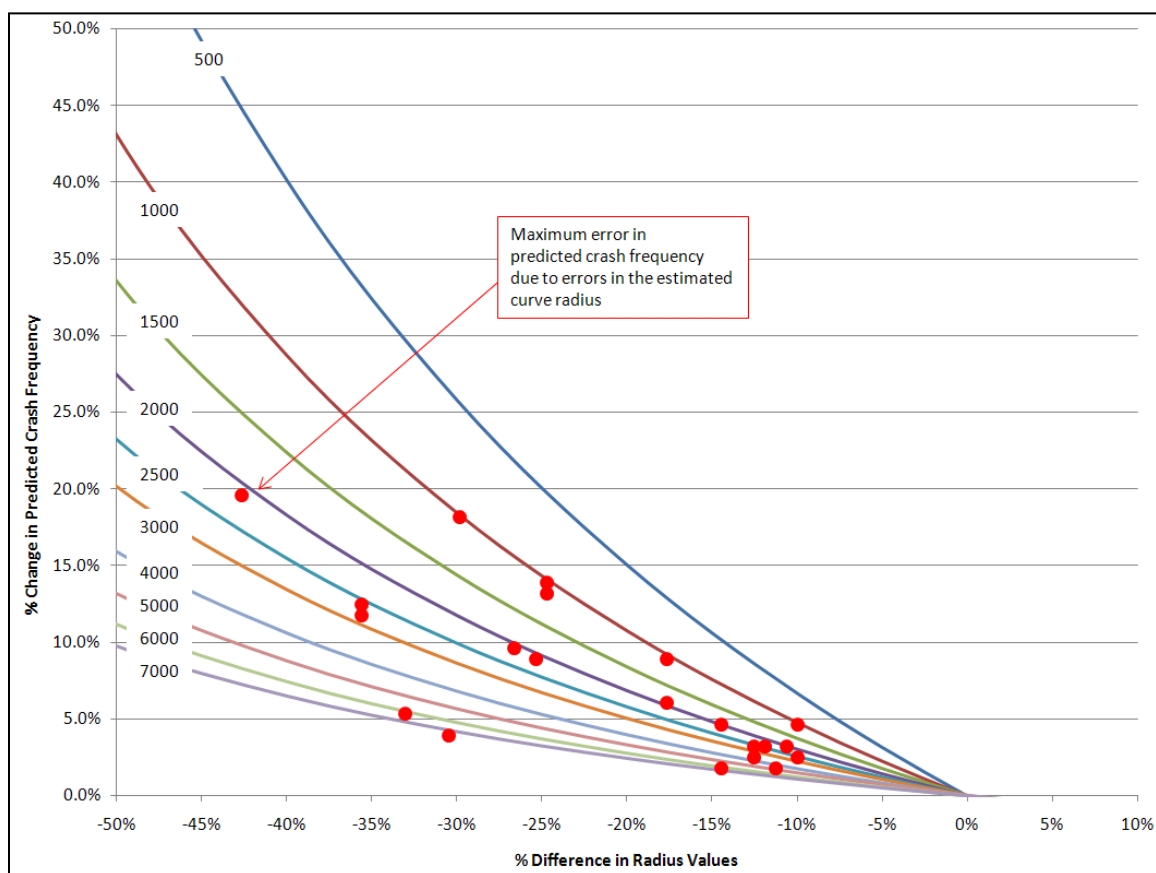


Figure 3-21. Sensitivity of predicted crash frequency to $R_{\text{regression}}$ percent difference.

The long chord method produced an underestimated radius value in 66 percent of the curves with as-built data. Thirty-seven (11 percent) horizontal curves had a percent difference in radius values greater than ten percent. Figure 3-22 shows the sensitivity of predicted crash frequency for the 37 horizontal curves with a percent difference in radius values greater than ten percent. Only one horizontal curve had a radius value with a percent change in predicted crash frequency greater than 15 percent.

As discussed previously, there appears to be a large bias to underestimate the radius value using the long chord method for curves with radii greater than 5,000 feet. However, the predicted crash frequency for large radius curves are less sensitive to errors in the estimated radius value. All but one of these large radius curves could expect less than a five percent change in predicted crash frequency as a result of underestimating the curve radius.

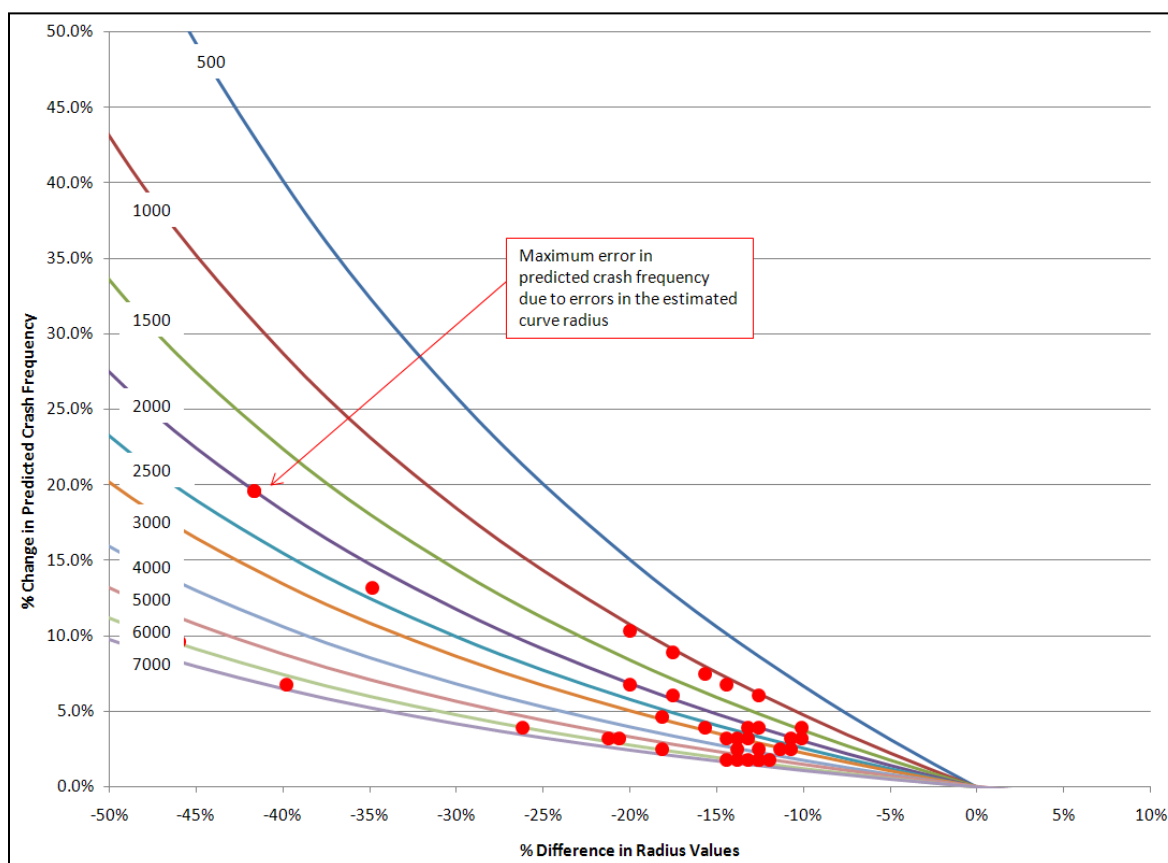


Figure 3-22. Sensitivity of predicted crash frequency to R_{chord} percent difference.

3.6 CONCLUSIONS AND RECOMMENDATIONS

Systemically identifying and precisely estimating curve geometry is an important step in understanding safety performance of horizontal curves. This study investigated the creation of a statewide curve database and attempted to validate the precision of the estimated geometric features of each curve. The horizontal curve identification method was found to be an accurate and complete method for identifying possible locations of curvature on the road network.

The validation results show that the curve identification method, as outlined previously, yielded a relatively precise method for estimating circular curve length and radius. Some large errors in the estimation of curve length were observed, however, this was expected because of the manual identification of the PC and PT and its effect on curve

length. In regards to estimating curve radius, the circular curve method is slightly more precise than the long chord method. However, with only a 3 percent difference in percent RMSE and very similar percent error, this difference is not significant.

It was found that the safety performance of smaller radius curves is more sensitive to errors in the estimated curve radius value. Although some horizontal curves were found to have large errors associated with the estimated curve radius, the maximum expected change in the predicted crash frequency was found to be less than twenty percent of the actual predicted crash frequency. For the use of safety performance evaluation, the majority of the horizontal curves in the database appear to have a predicted crash frequency within ten percent of the actual predicted crash frequency.

One limitation to this study was that no as-built curve data were available for secondary roadway horizontal curves. This, however, does not mean that the estimated geometric features of these secondary roadway curves are not estimated to the same precision of the primary roadway curves. It only means that validation was not performed over this range of curves. This study also ignored the presence of spiral transitions. The literature shows that the presence of a spiral transition could affect the safety performance of curve. It is recommended that the presence of spiral transitions and the possibility of estimating its value be investigated further.

It is also recommended that further studies, including a larger sample of data, be performed to solidify the validation of this curve identification method. It is further recommended that facilities with greater than two-lanes of travel be included in this investigation.

CHAPTER 4. HORIZONTAL CURVE CRASH PREDICTION MODEL

4.1 INTRODUCTION

To reduce problems associated with small sample sizes and regression to the mean, it is important to consider both crash history as well as the expected safety performance of similar sites when identifying the safety performance of a highway segment. Crash frequency, rate and cost are metrics commonly used to identify high crash locations (black spots). The use of such metrics using data from study sites alone is known as naïve analysis. The empirical Bayes (EB) method accounts for both crash history as well as the safety performance of similar sites (Hauer, 2001). The EB method utilizes a crash prediction model (safety performance function) to determine if a site is experiencing an unusually high frequency, rate, or severity of crashes. Safety performance functions have been developed for a range of roadway attributes (HSM Practitioner's Guide, 2011). The safety performance function for horizontal curves was developed using a regression model developed by Zegeer et al. (1992). However, for more accurate analysis, SPFs should be developed or at least calibrated for conditions specific to a study area. The purpose of this chapter is to develop safety performance functions for the horizontal curve database validated in Chapter 3 of this thesis.

A review was conducted to identify and summarize literature related to the safety performance of horizontal curves relative to their geometric and operational features as represented in previous research producing curve crash prediction models. In this study, crash models were developed for both serious crashes (fatal and major injury) and all crashes for curve radius values estimated by using the circular regression and long chord methods in Chapter 3. Lastly, the reliability of the safety performance functions for all four models are compared.

4.2 REVIEW OF LITERATURE

Bonneson, et al. (2007) developed a relationship between injury and fatal crash frequency and curve design using data from 1,757 curves in Texas. Included in the analysis

was the development of the relationship between curve radius and crash rate, shown in Figure 4-1. This curve indicates that crash rate increases sharply for curves with radii less than 1000 feet and that crashes on longer curves are less likely to result in an injury or fatality.

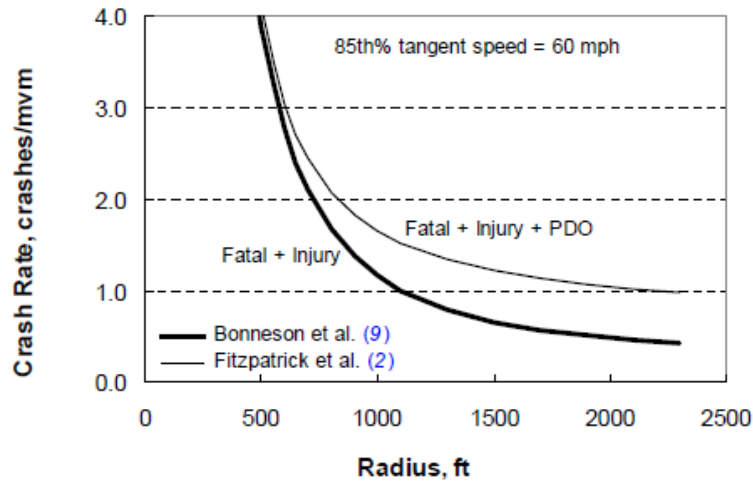


Figure 4-1. Curve crash rate as a function of radius.

Preston et al. (2009) suggests similar findings between the crash rate and curve radii. Compared to 2000 foot radius curves, crash rates for 1500 foot curves are twice as high; crash rates for 1000 foot curves are five times as high, and crash rates for 500 foot curves are eleven times as high.

The identification of promising horizontal curves (curves where improvements may result in significant reduction in crashes) should be based on more than just crashes. Curve radii, traffic volume, presence of visual traps, intersections and proximity to other high priority curves should be considered. Crash severity should also be included.

A Safety Performance Function (SPF) is a curve relating expected crash frequency to traffic level for a roadway segment or intersection, over some fixed period of time (usually one year). While SPFs could be developed for roads with specific features (e.g., lane width, shoulder type, etc.) they are typically developed for traffic (AADT) only. The impact of specific features on safety performance is now usually accounted for by the application of crash modification factors or functions (CMFs).

As crashes are statistically random, positive count events with variance difference

than mean values, the negative binomial distribution is assumed to model their distribution (Hauer, 2001). For models with a large number of zero event observations a zero-inflated negative binomial regression model could yield a better fit. To compare and select the best fit between a negative binomial model and zero inflated negative binomial model a Vuong statistic is used (Washington, 2011).

Tarko (2007) developed SPFs for multiple facility types, including two-lane rural roadways. Significant variables were found to be lane width, shoulder width, roadside hazard rating, driveway density, average grade for vertical curves and average degree of curvature in the segment (Tarko, 2007). Equation 4-1 shows the general form used for the SPFs developed in the report.

Equation 4-1:

$$A = \exp(k)LQ^{\beta}\exp(\sum \gamma_i X_i)$$

A = number of crashes in a year

L = length of the section in miles

Q = AADT of the section

X = explanatory variables

k, β, γ = constants

In *Accident Models for Two-Lane Rural Segments and Intersections* (Vogt, 1998) crash prediction models were developed for both segments and intersections in Minnesota and Washington. Poisson, negative binomial, extended binomial, and logistic techniques were tested and results showed that all but the logistic model yielded consistent values for regression variables. Additionally, overdispersion was found to be present, thus the negative binomial models were preferred. Exposure, lane and shoulder width, a roadside hazard rating, driveway density, degree of curvature, horizontal, and vertical alignment variables were all found to be significant in the segment models.

Effective Safety Factors on Horizontal Curves of Two-lane Highways (Aram, 2010) developed crash prediction models for horizontal curves on two-lane rural highways. The variables found to be significant were degree of curvature, curve segment length,

superelevation, length of spiral curve, shoulder width, and AADT (as an offset variable). Equation 4-2 shows the general form of the crash prediction model for this study.

Equation 4-2:

$$CR = ADT * L * 365 * 10^{-6} e^{(-10.5606 + 0.1087D_c + 0.000840L_{ct} + 0.0963E_c - 0.5842L_{sp} - 0.1970S_w)}$$

CR = number of horizontal curve-related crashes

ADT = average daily traffic (veh day⁻¹)

L = roadway section length (m)

D_c = Degree of curvature (18000 / πR)

E_c = Superelevation horizontal curve (%)

L_{sp} = Length spiral curve (m)

S_w = Shoulder Width (m)

L_{ct} = Total length segment of horizontal curve (m), ($L + 2L_{sp}$)

Bonneson, et al. (2006) calibrated the accident modification factor (AMF) for horizontal curve radii for 1,757 curves in Texas. This study included the calibration of a negative binomial regression crash prediction model. Variables for the model included, AADT, curve segment length, degree of curvature, and a categorical region variable. Separate calibrations of the AMF for lane width and shoulder width were also computed.

To compare models several methods are available. The Akaike Information Criterion (AIC) is a measure that is used to compare models with different error distributions using the same set of data. The AIC is a relative measure of the information lost when a model is created. The lower the AIC the better the model (Hu, 2007). The AIC, however is not a goodness-of-fit measure. For goodness-of-fit of a regression model, the McFadden ρ^2 statistic is a common measure. The McFadden ρ^2 statistic yields a value between zero and one and a “statistic close to one suggests that the model is predicting the outcomes with near certainty” (p.322, Washington 2011). The McFadden ρ^2 increases with the inclusion of additional parameters; to account for this a corrected ρ^2 is estimated as shown in Equation 4-3. The McFadden ρ^2 tends to be small with a value better 0.2 and 0.4 to be considered satisfactory (Ainsworth, 2010).

Equation 4-3:

$$\rho^2 = 1 - \frac{LL(\beta) - K}{LL(0)}$$

$LL(\beta)$ = log likelihood at convergence with parameter vector β

$LL(0)$ = initial log likelihood (with all parameters set to zero)

K = number of parameters in the vector β

4.3 DESCRIPTIVE STATISTICS

It is important to investigate general data relationships before a crash prediction model is created. In order to gain a better understanding of horizontal curve safety, trends related to attributes of horizontal curves are explored. This section provides an overview of statistics related to attributes of horizontal curves for all Iowa rural, paved, two-lane roadways, including both primary and secondary roadways.

4.3.1 All rural, paved, two-lane roadway horizontal curves

Figure 4-2 shows the number of horizontal curves by number of all crashes regardless of severity. Fifty-three percent of horizontal curves statewide had no crashes of any severity from 2001-2009. Eighty-eight percent of all horizontal curves have no more than two crashes of any severity during that same period. Only 22 horizontal curves had greater than nine crashes (or an average of one crash of any severity per year) over the study period.

When limiting the analysis to only serious crashes (fatal + serious injury), 90 percent of all horizontal curves have zero crashes recorded over the nine year study period. Only one horizontal curve experienced over three serious crashes during the study period. The many zeros in the database can present special problems for regression analysis and this topic will be discussed later.

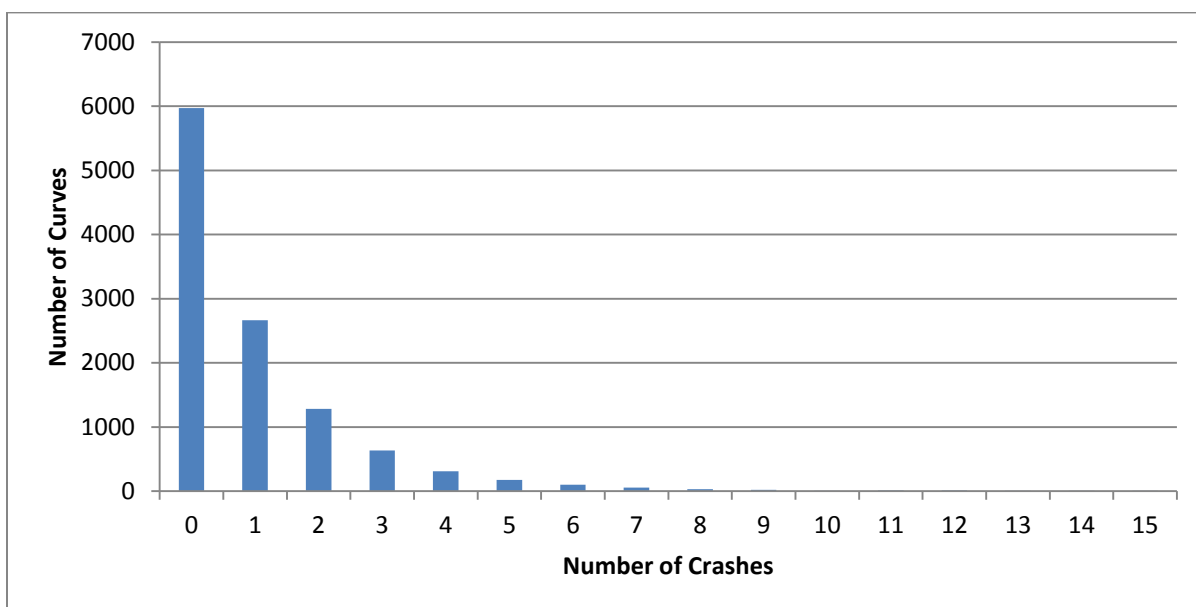


Figure 4-2. Number of horizontal curves by number of all crashes.

Figure 4-3 displays the total number of all crashes, regardless of severity, for all horizontal curves by AADT. AADT for all horizontal curves in this study, range from under 100 to approximately 10,000. Forty-one percent of all crashes occur on roadways with an AADT of less than 1,000, while 39 percent of all crashes occur on roadways with an AADT between 1,000 and 3,000.

Figure 4-4 charts the total number of serious crashes for all horizontal curves against AADT. Nearly half of all serious crashes on horizontal curves occurred on roadways with less than 1,000 AADT. All crashes and serious crashes over the nine year study period appear to be distributed in a similar fashion.

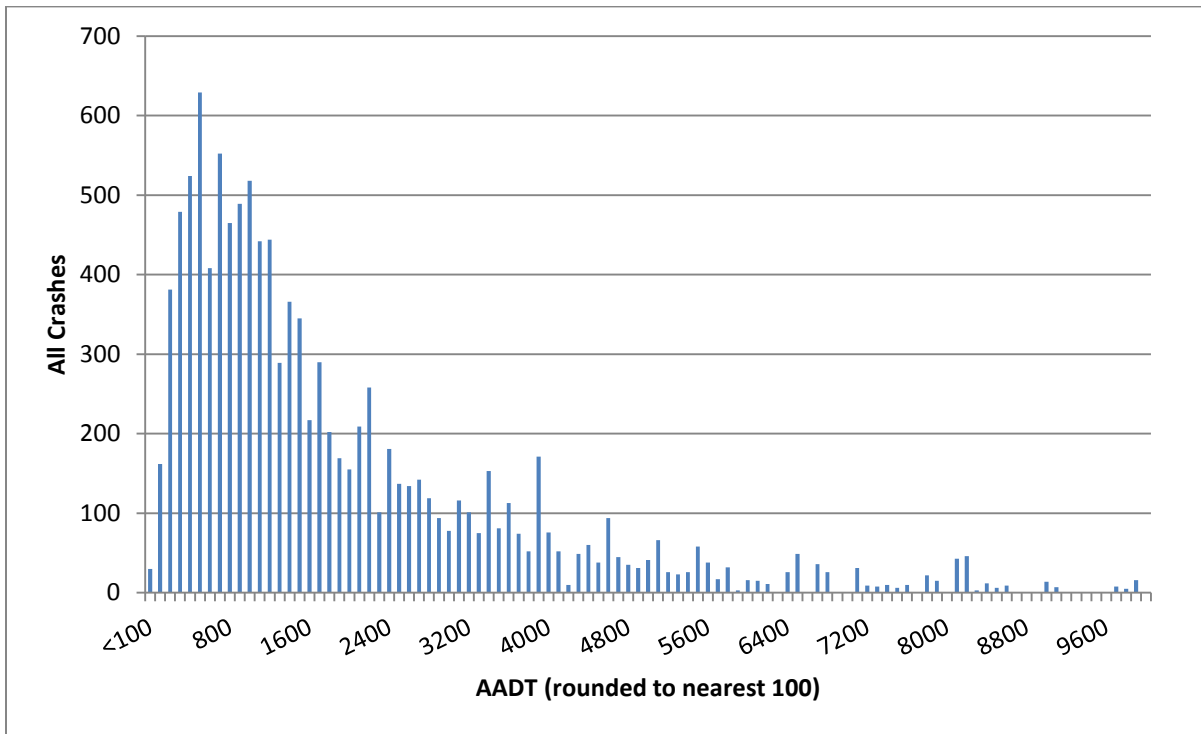


Figure 4-3. All crashes by AADT on all rural, paved, two-lane roadway horizontal curves.

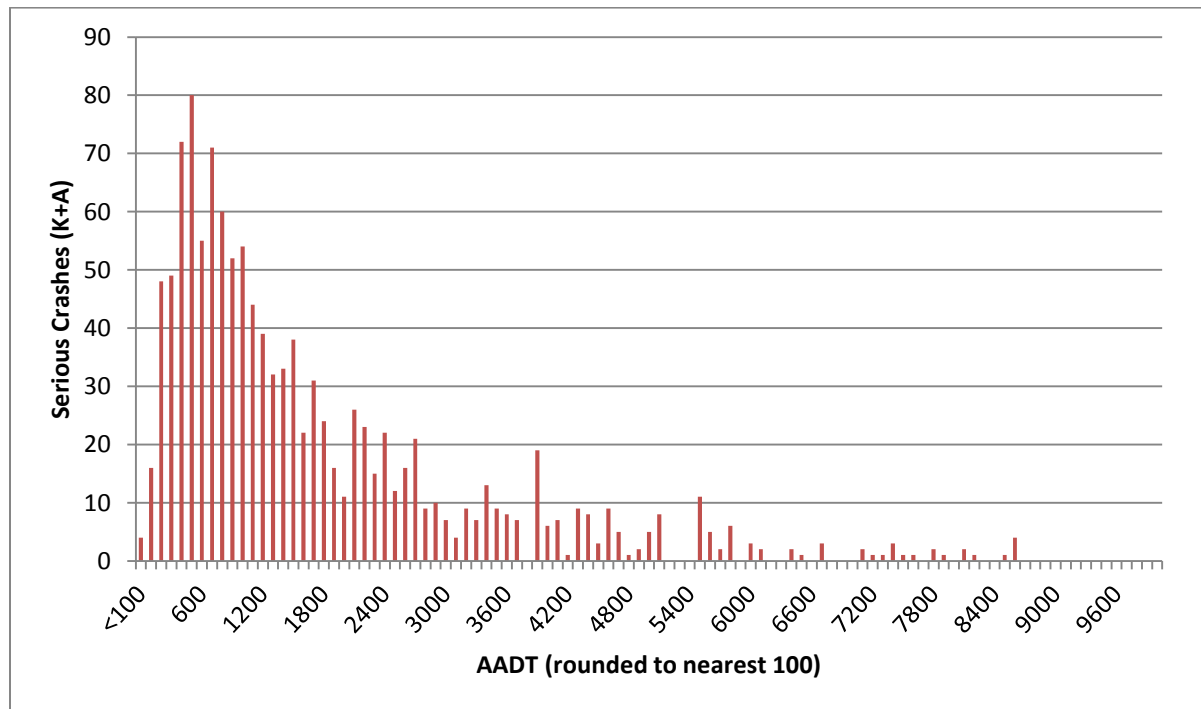


Figure 4-4. Serious crashes (K+A) by AADT on all rural, paved, two-lane roadway horizontal curves.

Figure 4-5 shows the crash frequency for six different curve radius categories, as well as for all curves, by crash type. It appears that more crashes occur on horizontal curves with a radius between 500 and 1,500 feet, but when compared to the number of curves in each category, all categories are relatively similar. Figure 4-6 shows the crash severity for each of these categories as a percentage. Curves with radii between 500 and 1,500 appear to experience slightly more serious crashes (K+A) relative to other radius value ranges.

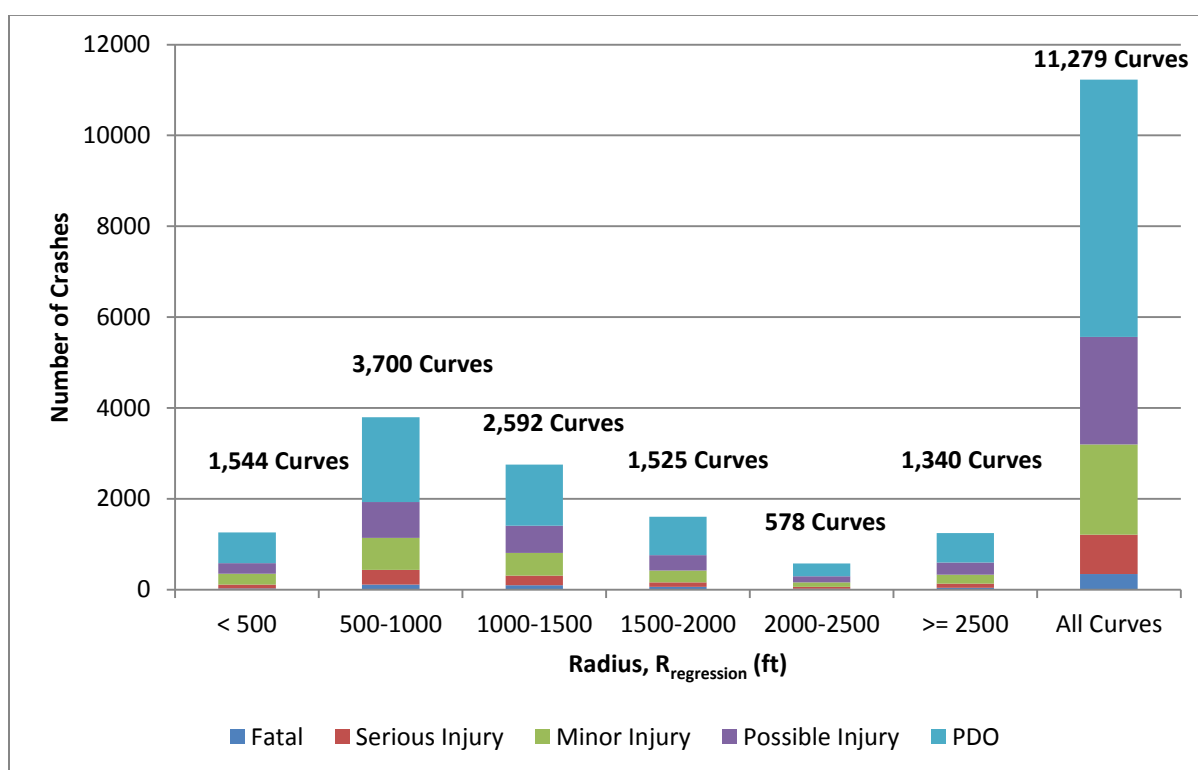


Figure 4-5. Number of crashes by curve radius category for all rural, paved, two-lane roadway curves.

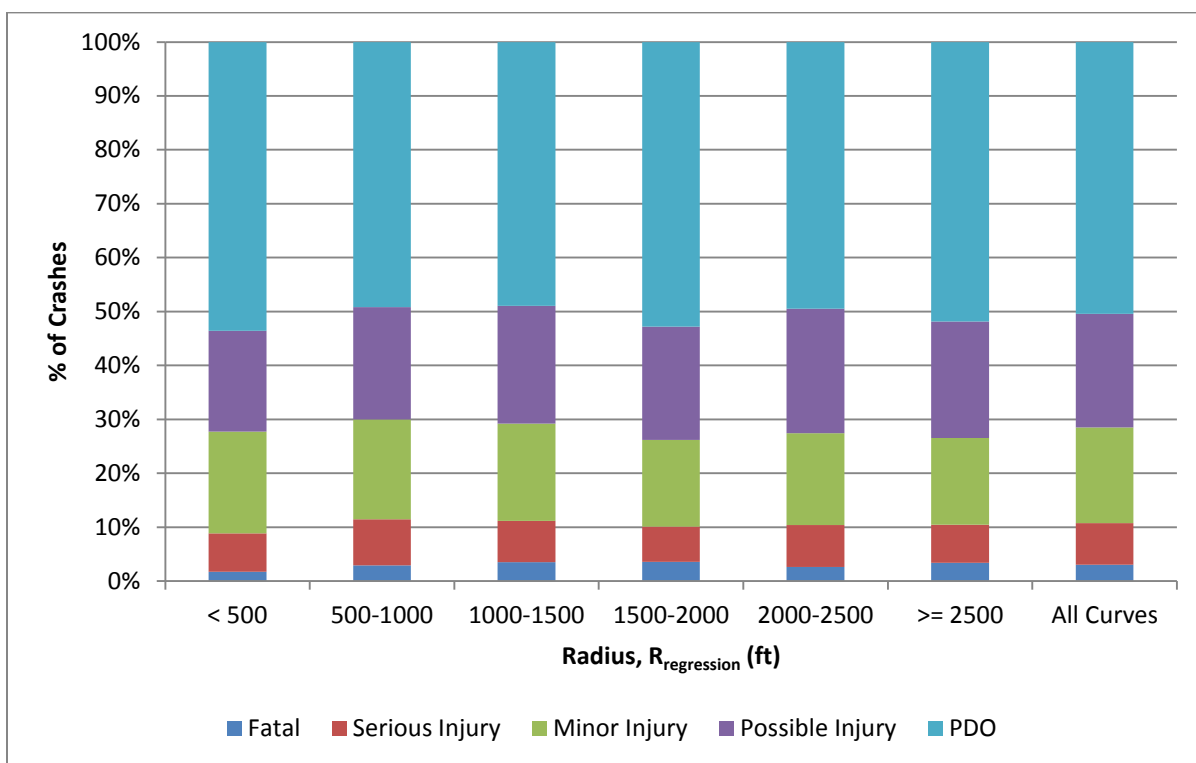


Figure 4-6. Crash severity ratio by curve radius category for all rural, paved, two-lane roadway curves.

From the data it would appear that smaller radius horizontal curves have higher crash rates. Using Equation 4-4 (Iowa DOT, 1989) to compute rate, Figure 4-7 shows the crash rate per HMVMT by crash severity for six curve radius categories. For each successively smaller curve radius category, the all-crash rate appears to roughly double. Severity appears to be inversely related to the all-crash rate.

Equation 4-4:

$$\text{Crash Rate/HMVMT} = \frac{(\sum_j \text{Crashes}_j)(100,000,000)}{(\sum_j \text{AADT}_j)(365)(\sum_j L_j)(9 \text{ years})}$$

Crashes = number of crashes linked to each Horizontal Curve

AADT = average annual daily traffic for horizontal curve roadway segment

L = length of curve (miles)

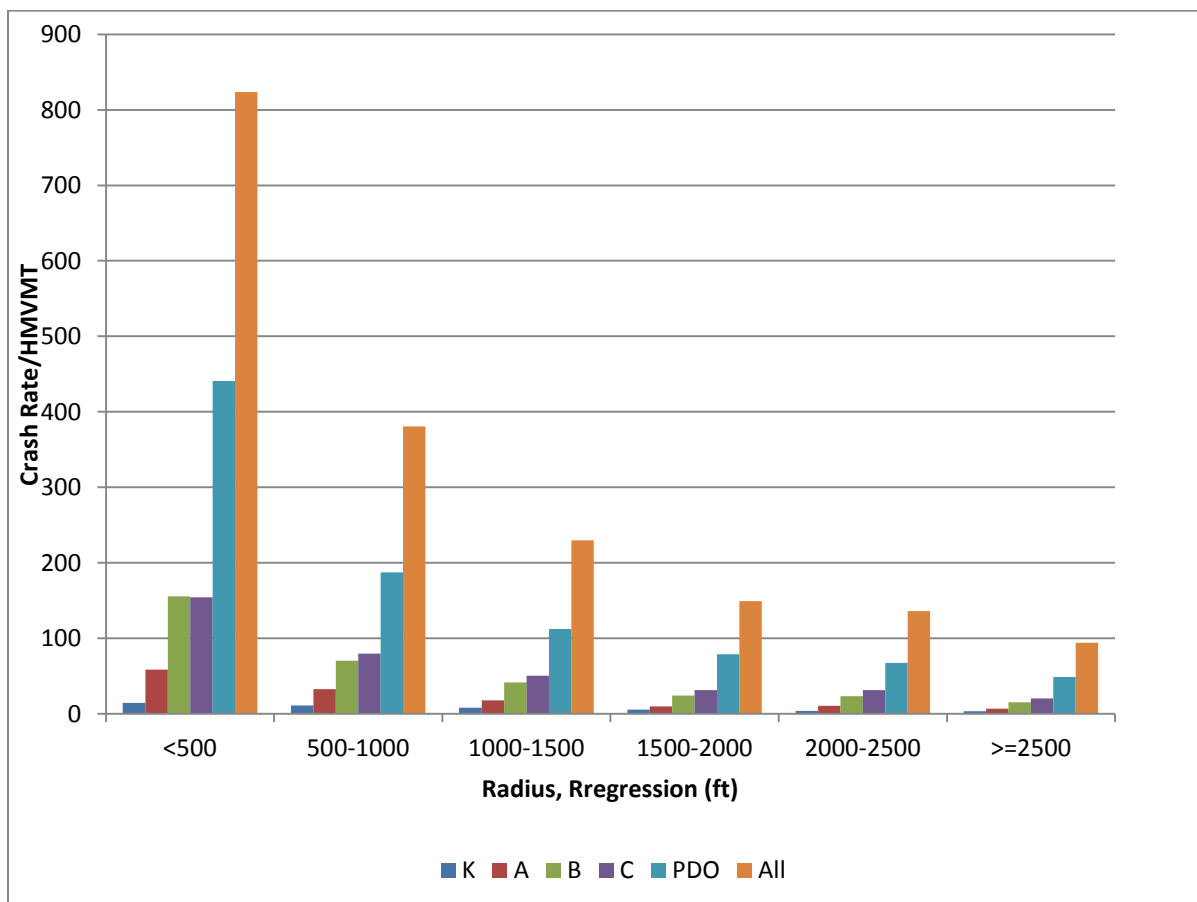


Figure 4-7. Crash rate per HMVMT for all crash severities by curve radius category.

Figure 4-8 examines the fatal crash rate per HMVMT for each curve radius category. The trend for fatal crashes is similar to the trend for all crashes: the smaller the curve radii, the higher the crash rate. The fatal crash rate for all horizontal curves in this study is 6.64/HMVMT, nearly five times higher than Iowa's rural roadway fatal crash rate of 1.42/HMVMT.

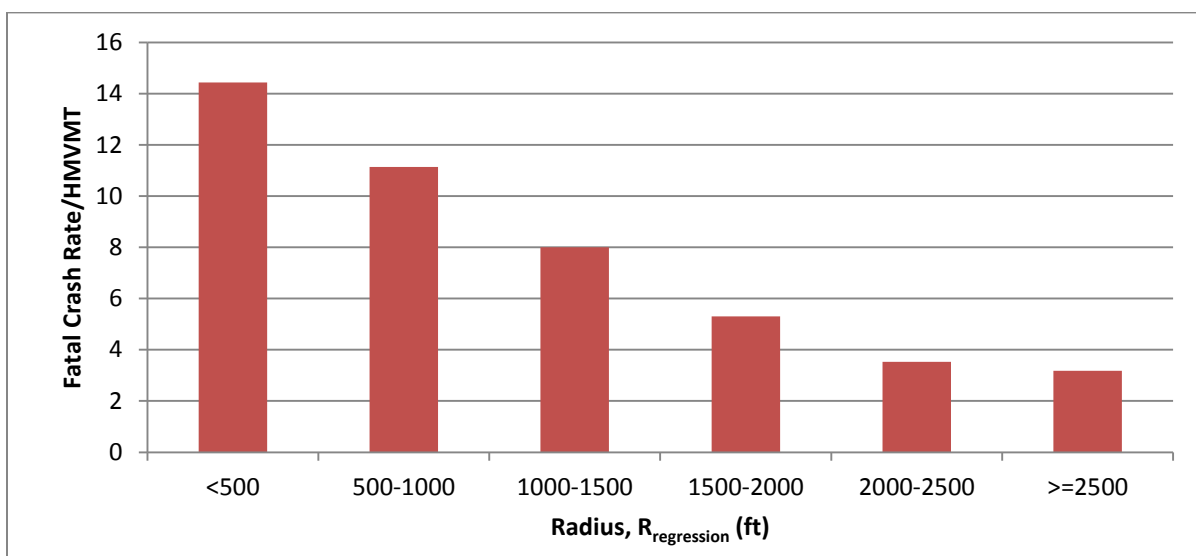


Figure 4-8. Fatal crash rate per HMVMT by curve radius category for all curves.

Figure 4-9 shows the crash frequency for all horizontal curves in this study by crash severity and lane width. It is difficult to recognize any trends as lane width categories are not represented equally, and 11 and 12 foot lanes are most prevalent. These two categories have similar distributions of curve crash severity.

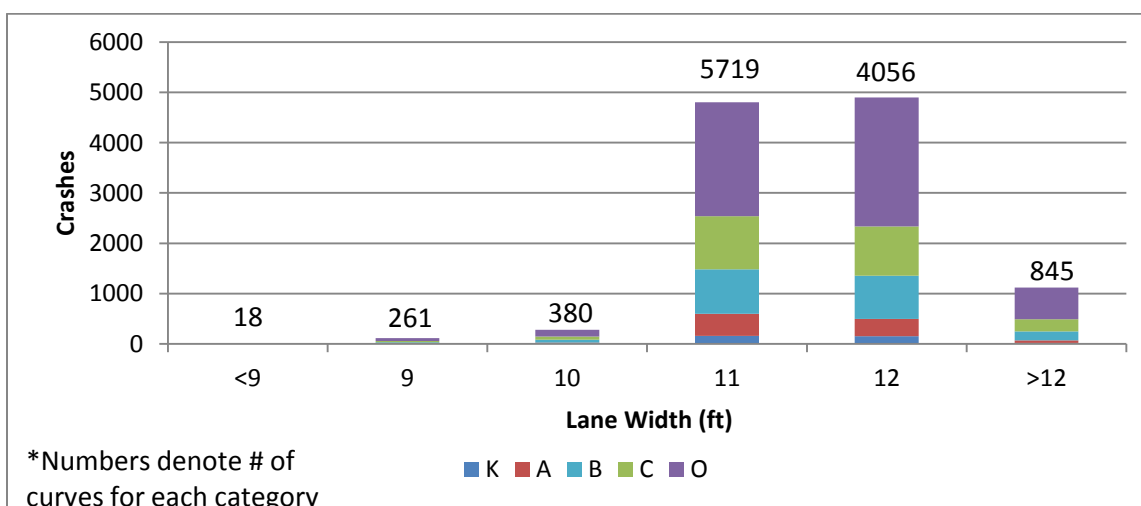


Figure 4-9. Crash frequency for all horizontal curves by crash severity and lane width.

Figure 4-10 shows the crash frequency for all horizontal curves by crash severity and terrain. Terrain is a categorical attribute that refers to the lane adjacent to the roadway. Categories for terrain include: not applicable (N/A), flat, rolling, and hilly. Horizontal curves appear to be more prevalent in rolling terrain. It appears that each terrain has a similar crash frequency relative to the number of curves in each category, but again it is difficult to distinguish any definite trends in the data because each category is not equally represented.

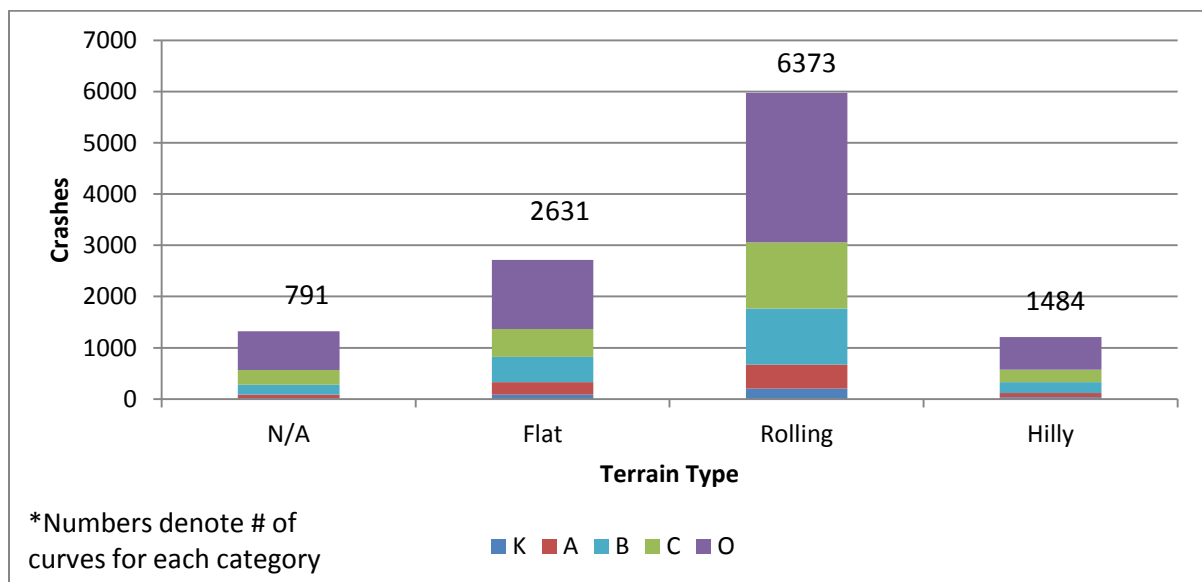


Figure 4-10. Crash frequency for all horizontal curves by crash severity and terrain adjacent to the roadway.

Figure 4-11 displays the crash frequency for all horizontal curves by crash severity and shoulder type. Shoulder types are divided into four separate categories: no shoulder, earth, gravel, and paved. Earth and gravel shoulders are the most prevalent shoulder types for Iowa roadways. Figure 4-12 shows the crash frequency for all horizontal curves by crash severity and shoulder width (in feet). Curve crash severity of all shoulder width categories appear to have similar distributions. It is difficult to distinguish any trends, in either figure, as all categories are not equally represented.

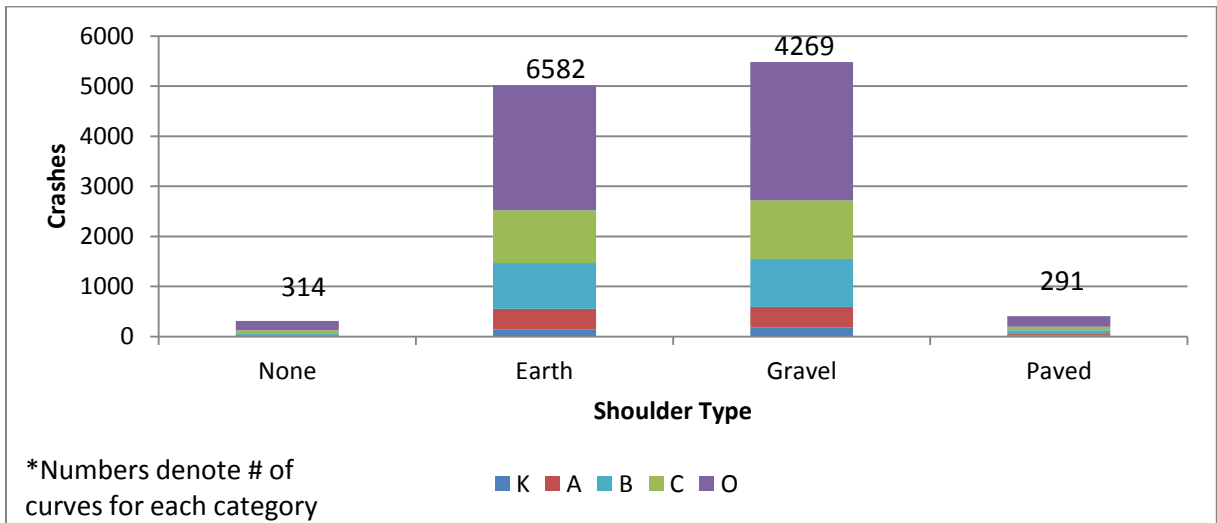


Figure 4-11. Crash frequency for all horizontal curves by crash severity and shoulder type.

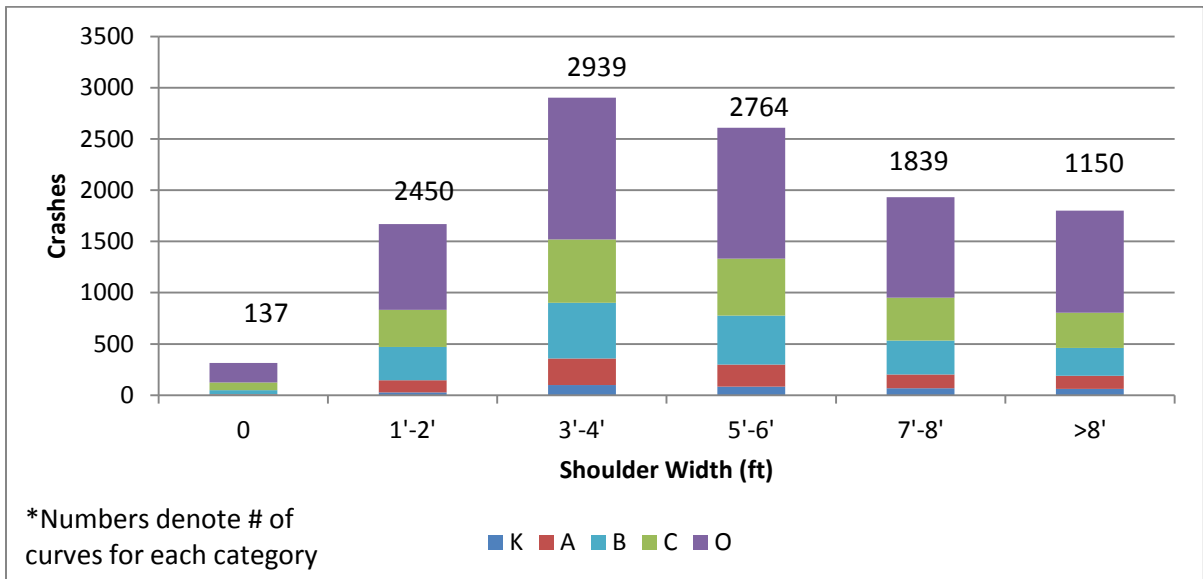


Figure 4-12. Crash frequency for all horizontal curves by crash severity and shoulder width.

4.3.2 Primary rural, paved, two-lane roadway horizontal curves

Figure 4-13 charts the crash frequency of all crashes by AADT on primary road horizontal curves. Primary roadways experience more traffic than secondary roadways, with no primary roadway having an AADT less than 400. Crashes appear to be more evenly

distributed between 1,000-4,000 AADT than the AADT versus crash frequency figure for all roadways (Figure 4-3). Over 70 percent of all crashes occurring on primary roadway curves occur with traffic volumes between 1,000 and 4,000.

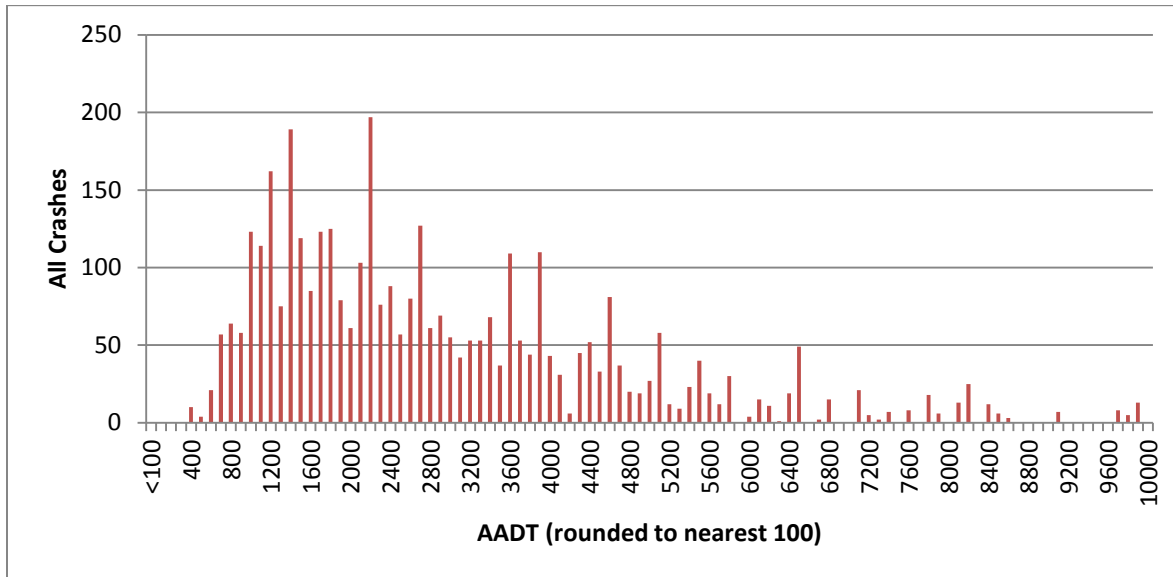


Figure 4-13. All crash frequency by AADT on primary, rural, paved, two-lane roadway horizontal curves.

Figure 4-14 shows the serious crash frequency by AADT on primary, rural, paved, two-lane roadway horizontal curves. Serious crashes on these roadway curves appear to be very variable making it difficult to identify any trend related to AADT for serious crashes on these roadways.

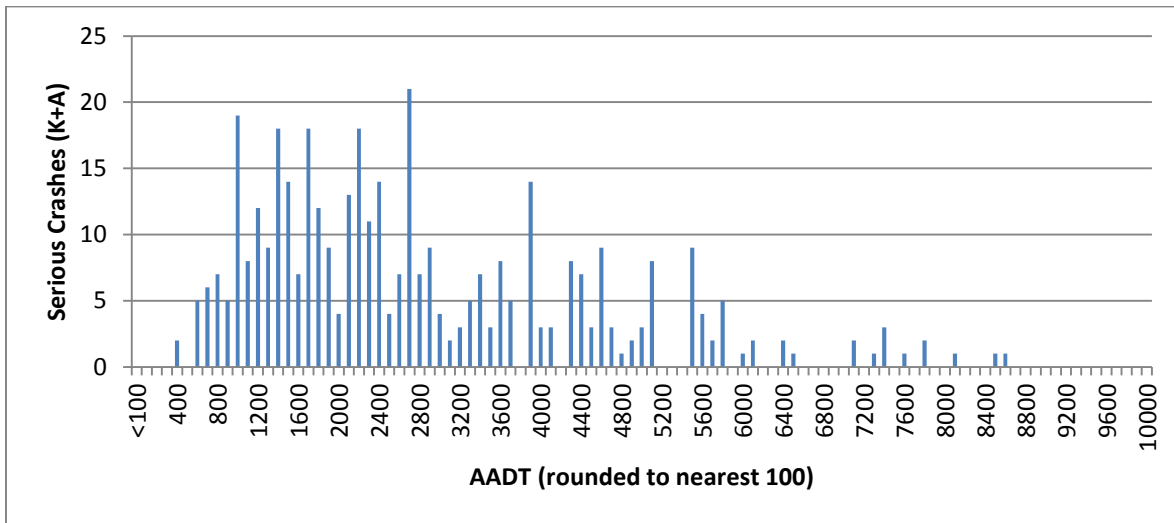


Figure 4-14. Serious crash (K+A) frequency by AADT on primary, rural, paved, two-lane roadway horizontal curves.

Crash rate trends for primary road horizontal curves are similar to that of all roadways as shown by Figure 4-15. Crash rates decrease as curve radii, $R_{\text{regression}}$ increases. Crashes rates also increase with decreasing severity in each curve radius category.

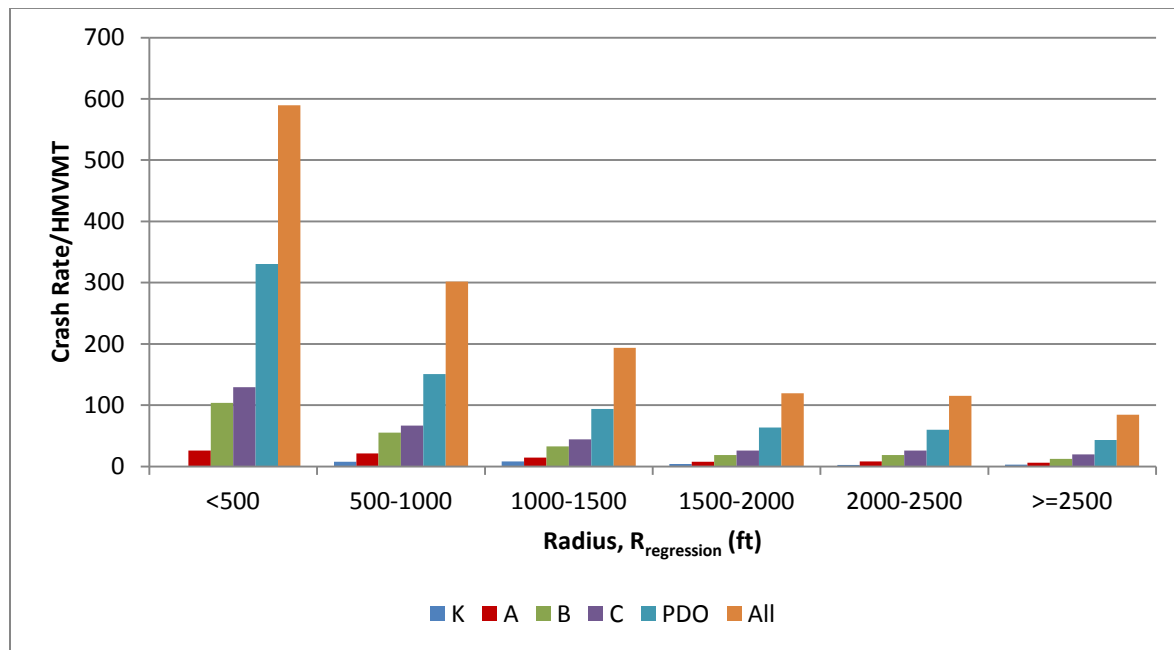


Figure 4-15. Crash rate per HMVMT on primary roadway curves for all crash severities by curve radius category.

Horizontal curve crashes by severity, lane width, and terrain for primary roads is shown in Table 4-1. Twelve foot lanes are the most common lane width on paved, two-lane primary roadways in Iowa. It is difficult to detect trends in the data related to lane width and terrain because each category is not equally represented.

Table 4-1. Horizontal curve crashes by severity, lane width, and terrain for primary roads.

PRIMARY ROADWAY HORIZONTAL CURVES							
Attribute		Crash Severity					ALL CRASHES
		K	A	B	C	O	
Lane Width (ft)	<9	0	0	0	1	2	3
	9	2	4	5	8	10	29
	10	3	5	15	12	45	80
	11	23	44	85	147	238	537
	12	95	193	441	581	1349	2659
	>12	7	22	77	113	298	517
Terrain	N/A	9	34	91	148	370	652
	Flat	39	86	161	200	480	966
	Hilly	74	126	310	448	932	1890
	Rolling	8	22	61	66	160	317
Total Crashes		130	268	623	862	1942	3825

Table 4-2 displays horizontal curve crashes by severity, shoulder type, and shoulder width for primary roads in Iowa. Again, it is difficult to distinguish crash severity trends associated with these attributes because they are not equally represented.

Table 4-2. Horizontal curve crashes by severity, shoulder type, and shoulder width for primary roads.

PRIMARY ROADWAY HORIZONTAL CURVES							
Attribute		Crash Severity					ALL CRASHES
		K	A	B	C	O	
Shoulder Type	None	0	4	26	48	122	200
	Earth	38	85	175	235	494	1027
	Gravel	82	166	392	556	1254	2450
	Paved	10	13	30	23	72	148
Shoulder Width (ft)	0	0	4	26	48	122	200
	1'-2'	4	6	16	22	33	81
	3'-4'	30	49	120	191	308	698
	5'-6'	14	44	103	118	278	557
	7'-8'	36	76	182	232	504	1030
	>8'	46	89	176	251	697	1259
Total Crashes		130	268	623	862	1942	3825

4.3.3 Secondary rural, paved, two-lane roadway horizontal curves

Figure 4-16 charts the all-crash frequency by AADT on secondary roadway curves. Since most secondary roadways experience low traffic volumes, crashes are concentrated on segments with an AADT less than 1,500. Over 90 percent of all secondary roadway curve have an AADT less than 1,500. Data for serious injury crashes on secondary roadways have a similar trend as shown in Figure 4-17.

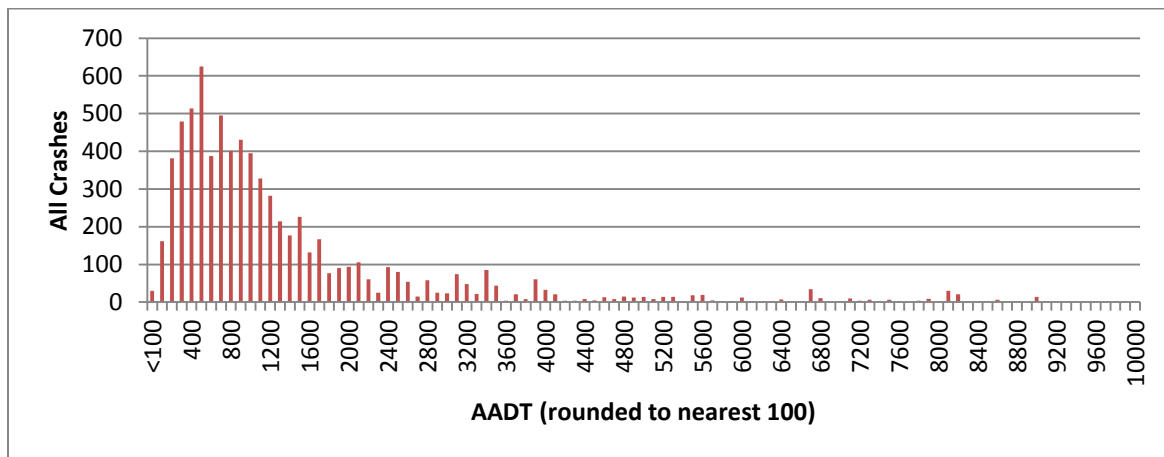


Figure 4-16. All crash frequency by AADT on secondary, rural, paved, two-lane roadway horizontal curves.

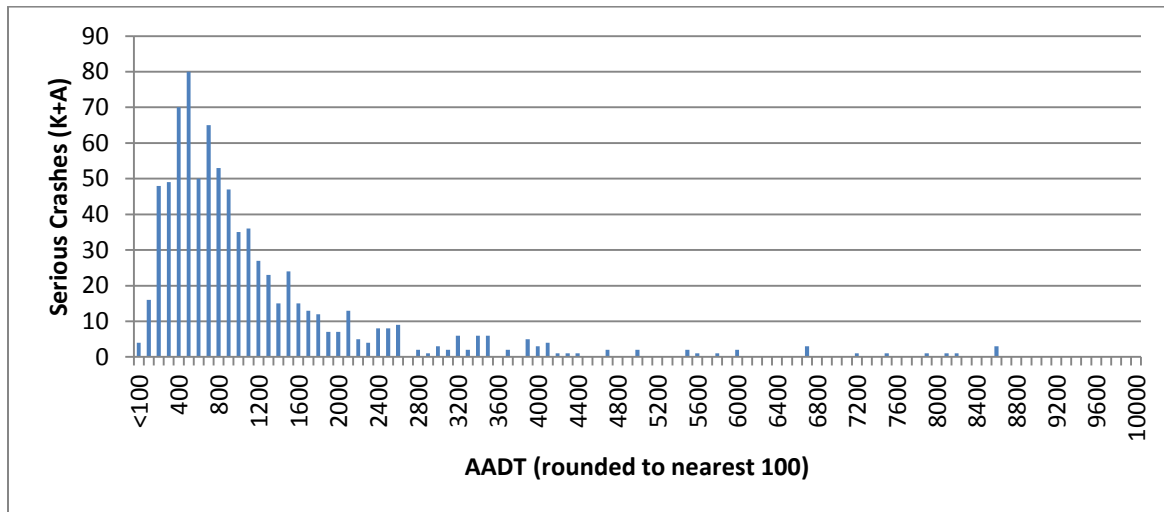


Figure 4-17. Serious crash frequency by AADT on secondary, rural, paved, two-lane roadway horizontal curves.

The data suggest that smaller radius horizontal curves on secondary roads have higher crash rates as displayed in Figure 4-18. The all crash rate trend on secondary road curves is similar to that of primary roadway curves however secondary roadways tend to have higher all-crash rates. The same can be said for fatal crashes on secondary roadway horizontal curves. This trend is illustrated in Figure 4-19. It should be noted that from curve radii ranging from 1,000 feet to 1,500 feet, primary road horizontal curves have a slightly higher crash rate than secondary road curves. This is the only curve radii range where primary roads experience a higher crash rate than secondary roads.

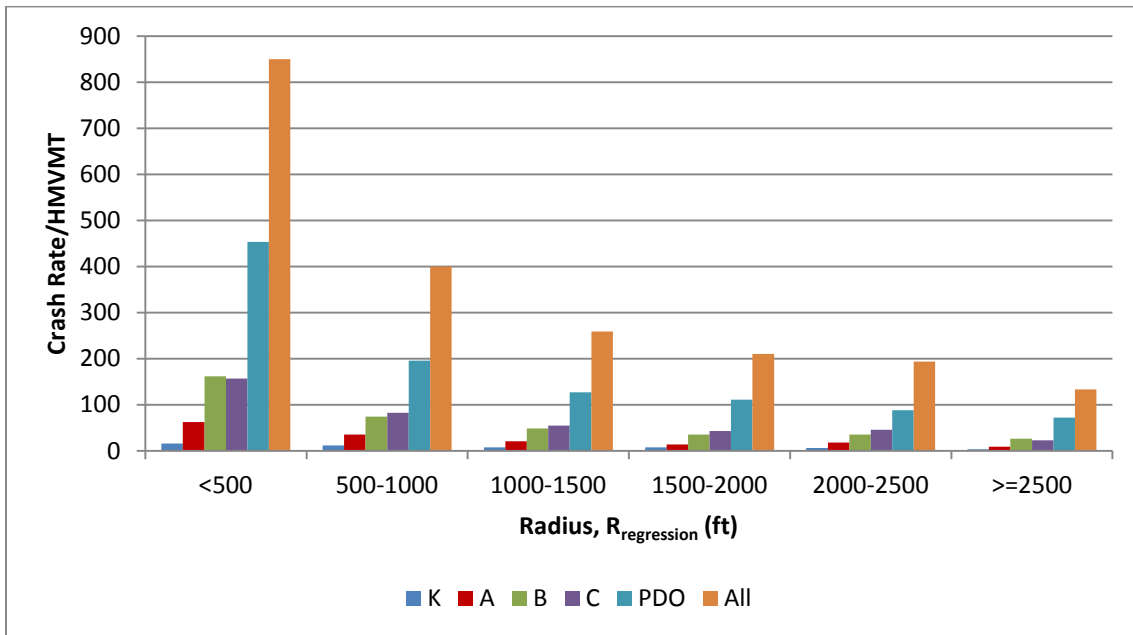


Figure 4-18. Crash rate per HMVMT on secondary roadway curves for all crash severities by curve radius category.

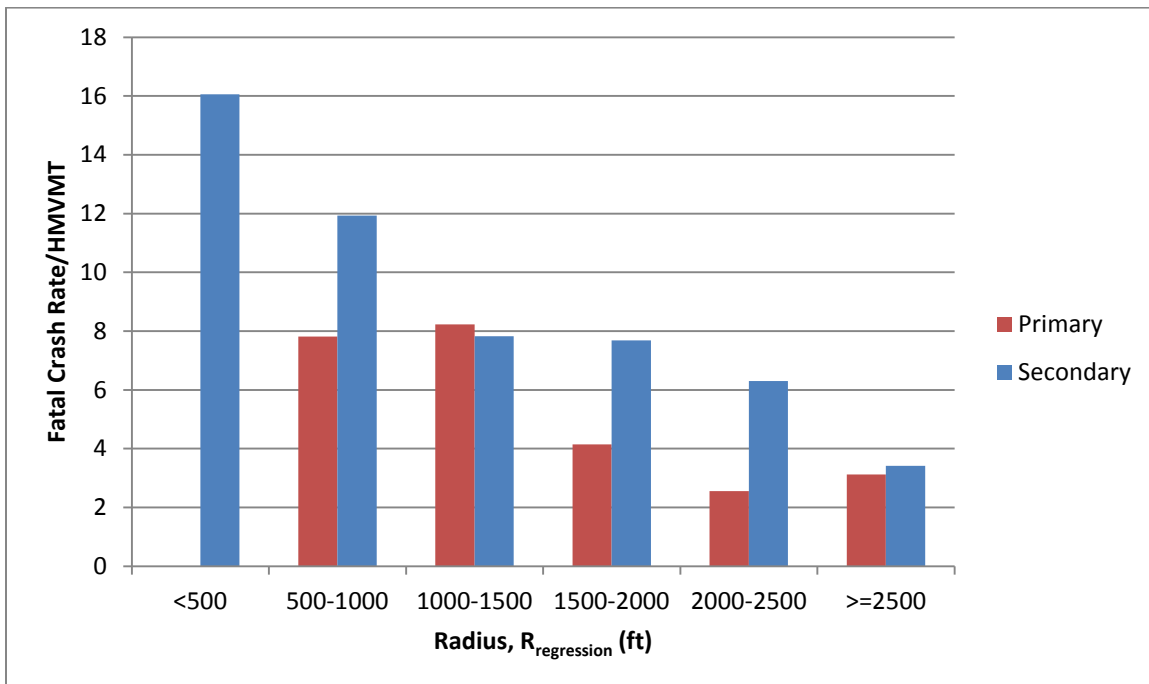


Figure 4-19. Fatal crash rate per HMVMT on secondary and primary roadway curves comparison.

Table 4-3 shows horizontal curve crashes by severity, lane width, and terrain for secondary roads. Table 4-4 displays the same data except with roadway attributes, shoulder type and shoulder width. Because the classes of each attribute are not equally represented, it is difficult to identify trends in the data.

Table 4-3. Horizontal curve crashes by severity, lane width, and terrain for secondary roads.

SECONDARY ROADWAY HORIZONTAL CURVES							
Attribute		Crash Severity					ALL CRASHES
		K	A	B	C	O	
Lane Width (ft)	<9	1	2	2	1	4	10
	9	4	4	10	18	48	84
	10	3	16	38	51	92	200
	11	138	392	802	907	2028	4267
	12	57	151	416	400	1212	2236
	>12	10	31	99	129	335	604
Terrain	N/A	8	35	106	135	386	670
	Flat	48	155	333	345	869	1750
	Hilly	131	342	780	849	1984	4086
	Rolling	26	64	148	177	480	895
Total Crashes		130	268	623	862	1942	3825

Table 4-4. Horizontal curve crashes by severity, shoulder type, and shoulder width for secondary roads.

SECONDARY ROADWAY HORIZONTAL CURVES							
Attribute		Crash Severity					ALL CRASHES
		K	A	B	C	O	
Shoulder Type	None	2	4	13	27	68	114
	Earth	109	323	739	820	2003	3994
	Gravel	93	253	566	610	1510	3032
	Paved	9	16	49	49	138	261
Shoulder Width (ft)	0	2	4	13	27	68	114
	1'-2'	23	114	309	339	803	1588
	3'-4'	69	209	426	425	1074	2203
	5'-6'	70	172	375	435	1001	2053
	7'-8'	32	58	149	187	476	902
	>8'	17	39	95	93	297	541
Total Crashes		130	268	623	862	1942	3825

4.4 METHODOLOGY

4.4.1 Data collection and preparation

Crash data were acquired from the Iowa SAVER crash database for the period 2001-2009. Data for all crashes, regardless of severity, were collected for this analysis. Statewide road data were collected from the Iowa GIMS roadway database for 2007. Rural paved two-lane roadways with a speed limit equal to or greater than 45 mph were then selected from the complete roadway database. Since this study is focused on horizontal curve safety and identification, crashes within 100 meters of an identified curve were queried from the initial complete set of 2001-2009 crashes. To account for the topographic errors associated with each different year of GIMS file, multiple GIMS layers were overlaid to identify horizontal curve crashes geocoded to other years of GIMS data.

Crashes located within 100 meters of each curve were identified as being possible curve crashes. Multiple years of GIMS data were overlaid to account for cartographic changes from year to year. Since crashes within 100 meters of each curve were included, some crashes on the tangent section were included as being a crash on a horizontal curve.

Crashes occurring on horizontal curves caused by other factors not related to curve geometry were omitted from the safety performance analysis. Crashes where the first harmful event or major cause was an animal collision were excluded. Crashes identified as occurring at an intersection or crashes suspected of occurring at an intersection were also omitted. Examples of these intersection crashes are those caused by running stop sign, running traffic signal, or failing to yield right-of-way at an intersection/traffic control device. After obviously unrelated crashes were removed, crash data were joined with their respective horizontal curves for analysis.

4.4.2 Negative binomial regression

Crash data for two radius estimation methods, $R_{\text{regression}}$ and R_{chord} , were fitted to a generalized linear model using negative binomial regression in SAS. Equation 4-5 shows the general form of the safety performance function.

Equation 4-5:

$$\mu = LENG * AADT^\alpha * e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}$$

μ = expected number of crashes (per unit time)

$LENG$ = length of curve segment in feet

$AADT$ = AADT of curve segment

X_k = model covariates (roadway attributes)

α, β_k = model coefficients

$LENG$ is an offset variable and is considered directly proportional to the expected number of crashes. Equation 4-5 was derived from Equation 4-6. For modeling purposes, the natural log of the AADT and length for each horizontal curve was used in the regression procedure in order to utilize the model form in Equation 4-5:

Equation 4-6:

$$\mu = e^{\ln(LENG) + \alpha(AADT) + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}$$

4.4.3 Variables

The following variables were used in the regression procedure:

LENG: The curve length in feet was considered an offset variable in the model

Log(AADT): The natural log of the horizontal curve segment's AADT.

R_{regression}: Curve radius in feet calculated using the circular regression method.

R_{chord}: Curve radius in feet calculated using the long chord method

LANEWID: The width in feet of the roadway lane (half the surface width)

SHDWIDTH: The width in feet of the shoulder.

SHDTYPE 0: Equals 1 if there is no shoulder. Zero if not.

SHDTYPE 1: Equals 1 if the shoulder type is earth. Zero if not.

SHDTYPE 2: Equals 1 if the shoulder type is gravel. Zero if not.

SHDTYPE 6: Equals 1 if the shoulder type is paved. Zero if not.

TERRAIN 0: Equals 1 if terrain is not applicable. Zero if not.

TERRAIN 1: Equals 1 if terrain is flat. Zero if not.

TERRAIN 2: Equals 1 if terrain is rolling. Zero if not.

TERRAIN 3: Equals 1 if terrain is hilly. Zero if not.

LIMITMPH: Speed limit of the roadway in miles per hour.

Other variables were considered for the safety performance functions but were not included because the data were not easily obtained or Iowa does not maintain it. As vertical alignment is not maintained by Iowa, a terrain variable was used to represent the general vertical alignment of the roadway. Spiral transition geometry, superelevation, driveway density, and proximity to other high priority curves were also considered for the crash prediction model, however, this data were not readily available.

The shoulder type (SHDTYPE) and terrain (TERRAIN) variables are categorical variables. For modeling purposes, dummy variables were created for each category of each variable.

4.5 ANALYSIS

Four different safety performance functions were created using the negative binomial regression model. Two crash models, one for all crashes and one for serious crashes (fatal + serious injury) crashes, were formed for both horizontal curve radius estimation methods, for a total of four models. Variables were included if their p-value indicated the variable's coefficient was statistically significant.

Equation 4-5 shows the general form for the crash prediction models. Each safety performance function estimates the predicted number of horizontal curve crashes over a 9 year period. In each model, the dispersion factor was found to be statistically significant.

4.5.1 All crashes with $R_{\text{regression}}$

Table 4-5 shows the negative binomial regression output for predicting all crashes using $R_{\text{regression}}$. The estimated coefficient for each parameter and their respective p-values are included in the output. The p-value is used to test a variable's significance in the model. Also included are the dispersion parameter, ϕ , model comparison statistic, AIC, and both log likelihood statistics. All parameters except, lane width, speed limit, and shoulder type were found to significant. Equation 4-7 shows the form of the safety performance function for all crashes using $R_{\text{regression}}$.

Table 4-5. All crash model using $R_{\text{regression}}$.

Parameter	Estimate	SD	P-Value
Intercept	-11.2556	0.1056	<.0001
$R_{\text{regression}}$	-0.0004	<0.0001	<.0001
SHDWIDTH	-0.083	0.0055	<.0001
TERRAIN 0	0.0838	0.0589	0.1551
TERRAIN 1	0.1094	0.047	0.0199
TERRAIN 2	0.1128	0.0423	0.0077
Log(AADT)	0.8522	0.0163	<.0001
ϕ (dispersion)	0.6132	0.0248	
AIC	28031.59		
Log Likelihood	-7410.2984		
Full Log Likelihood	-14007.7941		

Equation 4-7:

$$\mu = LENG * AADT^{0.8522} * e^{-11.2556 - 0.0004R_{\text{regression}} - 0.0830SHDWIDTH + 0.0838TERRAIN0 + 0.1094TERRAIN1 + 0.1128TERRAIN2}$$

4.5.2 Serious crashes with $R_{\text{regression}}$

Table 4-6 shows the negative binomial regression output for predicting fatal and serious injury crashes using $R_{\text{regression}}$. All parameters except speed limit and shoulder type were found to significant. Lane width was found to be significant in predicting fatal and serious injury crashes but not all crashes when using $R_{\text{regression}}$ in a crash prediction model.

Equation 4-8 shows the form of the safety performance function for all crashes using $R_{\text{regression}}$.

Table 4-6. Serious crash model using $R_{\text{regression}}$.

Parameter	Estimate	SD	P-Value
Intercept	-11.7813	0.3408	<.0001
$R_{\text{regression}}$	-0.0004	<0.0001	<.0001
LANEWID	-0.1198	0.0264	<.0001
SHDWIDTH	-0.054	0.0128	<.0001
TERRAIN 0	-0.1906	0.152	0.21
TERRAIN 1	0.2593	0.1122	0.0208
TERRAIN 2	0.22	0.1037	0.0338
Log(AADT)	0.7735	0.0387	<.0001
φ (dispersion)	0.5215	0.1357	
AIC	7428.649		
Log Likelihood	-3611.5007		
Full Log Likelihood	-3705.3247		

Equation 4-8:

$$\mu = LENG * AADT^{0.7735} * e^{-11.7813 - 0.0004R_{\text{regression}} - 0.1198LANEWID - 0.0540SHDWIDTH - 0.1906TERRAIN0 + 0.2593TERRAIN1 + 0.2200TERRAIN2}$$

4.5.3 All crashes with R_{chord}

Table 4-7 shows the negative binomial regression output for predicting all crashes using R_{chord} . Similar to the all crashes model with $R_{\text{regression}}$, all parameters except, lane

width, speed limit, and shoulder type were found to significant. The AIC for predicting all crashes using R_{chord} is slightly lower than when using $R_{\text{regression}}$. Equation 4-9 shows the form of the safety performance function for all crashes using R_{chord} .

Table 4-7. All crash model using R_{chord} .

Parameter	Estimate	SD	P-Value
Intercept	-11.26	0.1054	<.0001
R_{chord}	-0.0005	<0.0001	<.0001
SHDWIDTH	-0.0803	0.0054	<.0001
TERRAIN 0	0.0814	0.0588	0.1659
TERRAIN 1	0.1093	0.0469	0.0198
TERRAIN 2	0.114	0.0422	0.0069
Log(AADT)	0.8591	0.0163	<.0001
ϕ (dispersion)	0.6059	0.0246	
AIC	27971.53		
Log Likelihood	-7380.2715		
Full Log Likelihood	-13977.7671		

Equation 4-9:

$$\mu = LENG * AADT^{0.8591} * e^{-11.2600 - 0.0005R_{\text{chord}} - 0.0803SHDWIDTH + 0.0814TERRAIN0 + 0.1093TERRAIN1 + 0.1140TERRAIN2}$$

4.5.4 Serious crashes with R_{chord}

Table 4-8 shows the negative binomial regression output for predicting fatal and serious injury crashes using R_{chord} . All parameters except speed limit and shoulder type were found to significant. Lane width was found to be significant in predicting fatal and serious injury crashes but not all crashes when using R_{chord} in a crash prediction model. The AIC for predicting fatal and serious injury crashes using R_{chord} is slightly lower than when using $R_{\text{regression}}$. Equation 4-10 shows the form of the safety performance function for all crashes using R_{chord} .

Table 4-8. Serious crash model using R_{chord} .

Parameter	Estimate	SD	P-Value
Intercept	-11.7941	0.3403	<.0001
R_{chord}	-0.0005	<0.0001	<.0001
LANEWID	-0.1187	0.0264	<.0001
SHDWIDTH	-0.0514	0.0129	<.0001
TERRAIN 0	-0.1986	0.1519	0.1911
TERRAIN 1	0.2597	0.1121	0.0205
TERRAIN 2	0.2206	0.1036	0.0332
Log(AADT)	0.779	0.0387	<.0001
ϕ (dispersion)	0.512	0.1348	
AIC	7419.531		
Log Likelihood	-3606.9417		
Full Log Likelihood	-3700.7657		

Equation 4-10:

$$\mu = LENG * AADT^{0.7790} * e^{-11.7941 - 0.0005R_{\text{chord}} - 0.1187LANEWID - 0.0514SHDWIDTH - 0.1986TERRAIN0 + 0.2597TERRAIN1 + 0.2206TERRAIN2}$$

4.5.5 Goodness-of-fit comparison

The goodness-of-fit was computed for each of the four models using the McFadden's ρ^2 statistic as shown in Equation 4-3. The all crash models for $R_{\text{regression}}$ and R_{chord} are very comparable. Both models explain approximately 47 percent of the actual crash frequency. The serious injury crash models are also comparable but explain only two percent of the data associated with the actual serious crash frequency. Fatal and serious injury crashes are quite random and therefore it is difficult to model these crashes with certainty.

Table 4-9. McFadden's ρ^2 goodness-of-fit comparison.

	All Crashes	K+A Crashes
$R_{\text{regression}}$	0.47042	0.02289
R_{chord}	0.47143	0.02292

4.5.6 Empirical Bayes usefulness comparison

In order to estimate expected crashes at a site using the empirical Bayes process, a weighted average of the safety performance function must be calculated. This weighted average determines how much the safety performance function contributes to the expected number of crashes at a site (Hauer, 2001). The weight can range from a value of zero to one. The closer the weight is to one, the more reliable the safety performance function estimates the expected number of crashes at that site.

Equation 4-11 shows the weight calculation.

Equation 4-11:

$$Weight = \frac{1}{1 + \frac{(\mu * Y)}{\varphi}}$$

μ = model predicted number of crashes

Y = number of years of crash data

φ = dispersion parameter

The average weight for each of the four models was calculated. The all crashes models for $R_{\text{regression}}$ and R_{chord} were compared as was the fatal and serious injury crashes for $R_{\text{regression}}$ and R_{chord} . These weights were compared to see which model contributes more to the empirical Bayes process. The model with the higher average weight is said to be represented more in the EB crash prediction. Table 4-10 shows the average weight for each of the four crash prediction models.

Table 4-10. Calculated average weights for comparing a model's usefulness in the empirical Bayes process.

	All Crashes	K+A Crashes
$R_{\text{regression}}$	0.4733	0.8389
R_{chord}	0.4844	0.8457

4.5.6 Interpretation of models

It would be expected that as curve radius decreases, crash frequency increases, but this is not necessarily the case. Figure 4-20 charts the expected all crash frequency verses the estimated curve radius, $R_{\text{regression}}$. Figure 4-21 shows the serious crash frequency verses the estimated curve radius, $R_{\text{regression}}$. The same figures using R_{chord} yielded nearly identical results.

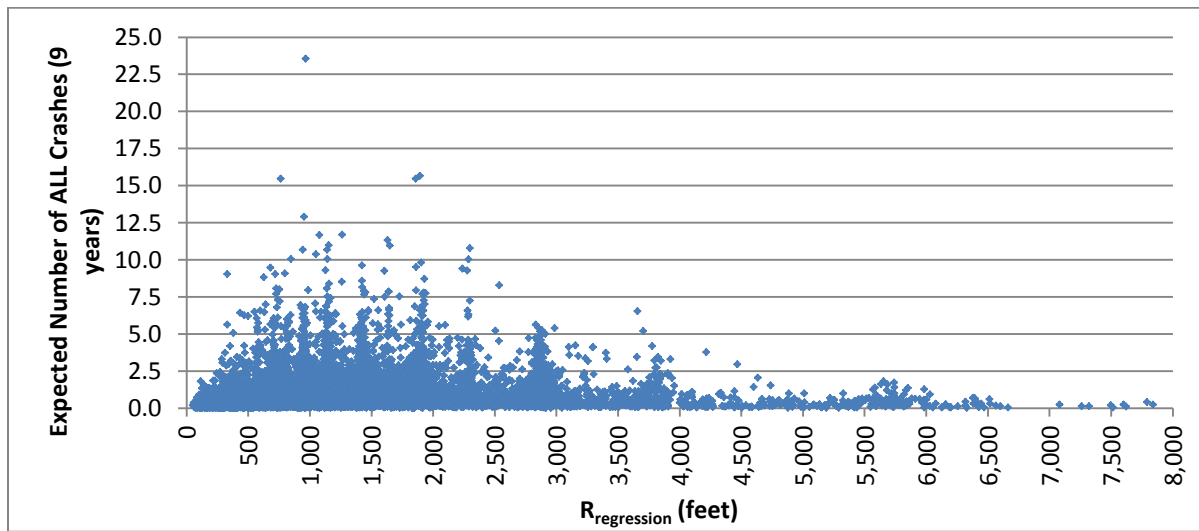


Figure 4-20. Expected all-crash frequency vs. curve radius on all horizontal curves.

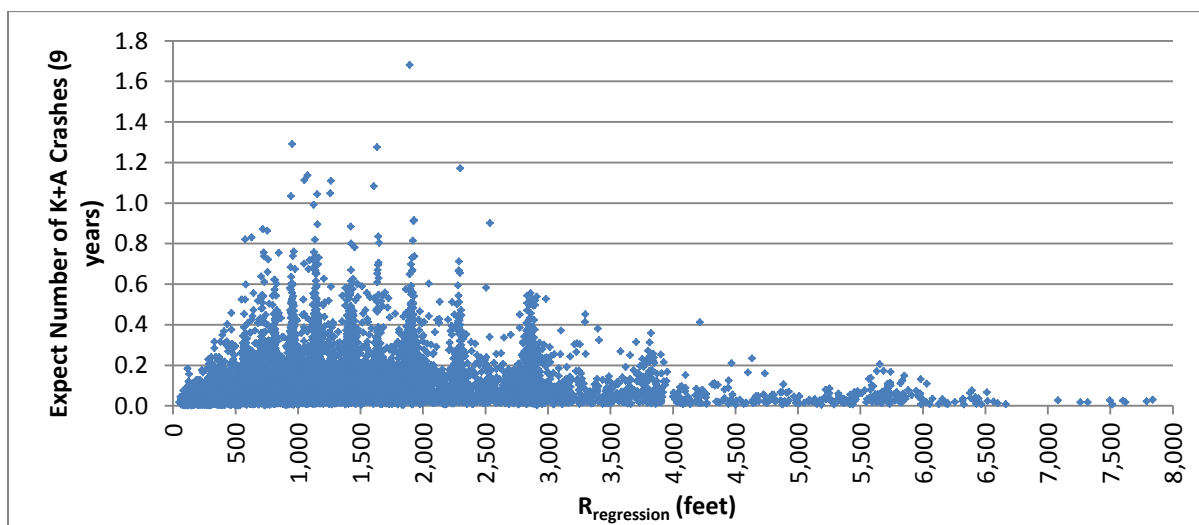


Figure 4-21. Expected serious crash frequency vs. curve radius on all horizontal curves.

The expected crash frequency for all and serious crashes appears to be even from curve radii of 500 feet to 3000 feet. This is due to the effect of other roadway attributes such as lane width, shoulder type and width, AADT, superelevation, and curve length on the safety performance of horizontal curves. Changes in these attributes can make a larger radius curve perform similarly to a smaller radius curve with superior roadway characteristics. For this reason, curve radius, by itself is not related to the severity of crashes on horizontal curves.

It was unexpected that crash frequency would decrease from a curve radius of 500 to a curve radius of zero. One possible explanation for this is shown in Figure 4-22 (Bonneson et al., 2007).

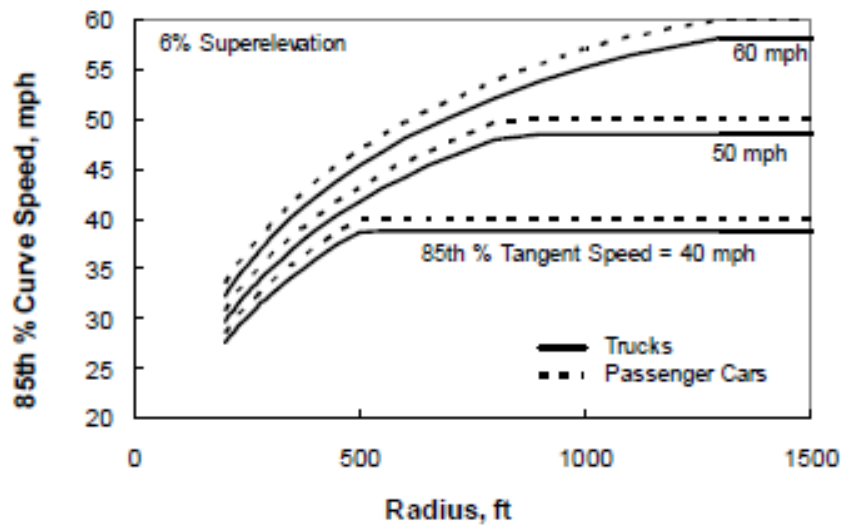


Figure 4-22. Effect of radius on curve speed.

For sharper curves, drivers tend to reduce their speed as they transition from a tangent segment to a curve segment. As curve radius decreases, the amount of speed reduction increases. For curves radii less than 500 feet, this reduction is especially significant. This resultant speed reduction could explain the reduction in crash frequency for horizontal curves with radii less than 500 feet.

4.6 CONCLUSIONS AND RECOMMENDATIONS

Crash modeling is an important step in understanding and improving the safety performance of horizontal curves. Descriptive statistics indicated a strong inverse relation between curve radius and crash rate. However, using descriptive statistics alone, it is difficult to find a relationship between curve radius and crash frequency.

Because such a large number of curves experienced zero fatal and serious injury crashes during the nine year study period, a zero-inflated negative binomial regression model was considered. However, the Vuong statistical test, comparing the zero-inflated negative binomial and negative binomial regression models for fatal and serious injury crashes was inconclusive. Therefore a zero-inflated was not used.

Crash models were created for all crashes and serious crashes (fatal + serious injury) using both radius estimation methods, $R_{\text{regression}}$ and R_{chord} . A goodness-of-fit statistic, ρ^2 , showed that both estimated curve radii models predicted crash expectancy with similar certainty. Both models explain approximately 47 percent of the expected crash frequency. The serious injury crash models, however, explain only two percent of the data associated with the expected crash frequency. Fatal and serious injury crashes are quite random and therefore it is difficult to model these crashes with certainty. Other attributes related to crash severity and consequence, such as sideslope and clear zone data, could improve the fit of the serious crash models.

Crash models were also compared using the weighted average from the empirical Bayes process. The crash model for all crashes and serious crashes using R_{chord} was identified as contributing more to the empirical Bayes crash estimation than the crash models using $R_{\text{regression}}$. However, the difference between the models using R_{chord} and the models using $R_{\text{regression}}$ is so small it can be considered negligible.

Interpretation of the crash models demonstrated that curve radius by itself is not related to the severity of horizontal curve crashes. The presence of other roadway attributes coupled with curve radius, however, is related. It is recommended that additional variables such as sideslope data, clear zone data, and the presence of a spiral transition be included in future research.

It is also recommended that future work include speed limit as a categorical variable

because speed limit as a continuous variable was not found to be significant in any models. An explanation for this is that all curves are located on paved, two-lane, rural roadways with a speed limit of at least 45 mph. The maximum speed limit on two-lane Iowa facilities is 55 mph, and speed limits are commonly set at 5 mph increments. Therefore, there are only three possible speed limits for the curves in this study, 45, 50, and 55 mph. Defining LIMITMPH as a categorical variable, with these three speed limit categories, could yield different results in regards to the significance of speed limit in a crash prediction model.

CHAPTER 5. GENERAL CONCLUSIONS

5.1 GENERAL DISCUSSION

Identifying where safety funding should be allocated continues to be a challenge. Knowing the type of roadway facilities most at-risk and the types of crashes occurring on those roadways is but one piece of the puzzle. Blindly shifting funding from one facility to another, based solely on what the crash data are suggesting, is not advised without having an estimate of potential consequences. Safety funding should be allocated based on a combination of where the crash data suggest and where the funding has the most benefit in reducing fatalities and serious injuries. In doing so, finding an optimum balance between black spot analysis and mass-action could be more closely achieved.

Understanding safety performance on horizontal curves also continues to be a challenge. When creating and analyzing a statewide horizontal curve database for safety performance it is important to have a reliable and precise estimation method. This thesis validated the use of a systematic horizontal curve geometry estimation method; however care should be taken when relying on these estimated values in understanding the safety performance of horizontal curves.

5.2 RECOMMENDATIONS AND CONCLUSIONS

When attempting to understand how to allocate safety funding it is important to consider not only where the data suggests funding should be spent but also where the greatest benefit for the cost can be achieved. It is recommended that the analysis performed in Chapter 2 be combined with a benefit cost analysis of potential safety projects, site specific and mass-action, to determine how best to allocate safety funding.

Chapter 3 provided a validation of a systematic identification method for horizontal curves in Iowa. The method for identifying horizontal curves was found to be an acceptable method for finding and estimating curve geometry. The circular curve method for estimating curve radius was found to be slightly more precise than the long chord method: however, the difference between the two datasets was not found to be statistically significant.

A sensitivity analysis showed that the safety performance of smaller radius curves is

more sensitive to errors in the estimated curve radius value. Although some horizontal curves were found to have large errors associated with the estimated curve radius, the maximum expected change in the predicted crash frequency was found to be less than twenty percent of the actual predicted crash frequency. For the use of safety performance evaluation, the majority of the horizontal curves in the database appear to have a predicted crash frequency within ten percent of the actual predicted crash frequency.

Crash prediction models for all crashes and serious crashes were developed using both estimated curve radii values, $R_{\text{regression}}$ and R_{chord} . The McFadden's ρ^2 statistic and average weight computed during the empirical Bayes process were used to compare the safety performance functions. It was found that the models for both radius values are, relatively speaking, equally good.

5.3 FUTURE RESEARCH

5.3.1 Funding allocation

Future work should be focused on determining which “problem” areas can be mitigated in the most cost effective manner. Furthermore, future research should include a benefit cost analysis of both black-spot locations and mass-action projects. This work should also include the utilization the Highway Safety Manual and methods involving crash reduction factors.

5.3.2 Expanded horizontal curve identification

This study focused only on identifying horizontal curves on rural, paved, two-lane highways. Further research should include facilities with more than two-lanes. Validation could also be expanded if as-built data included some horizontal curves from secondary roadways.

5.3.3 Additional variables

The inclusion of additional variables should be considered for the development of crash prediction models for horizontal curves in future research. Identifying the presence of spiral transitions and specifying speed limit as a categorical variable should be considered in

future research. Other variables related to crash severity and consequence, such as sideslope and clear zone data should also be considered.

5.4 REFERENCES

- Ainsworth, Andrew. Logistic Regression. California State University – Northridge. 2010.
- Aram, Ali. "Effective Safety Factors on Horizontal Curves of Two-lane Highways." *Journal of Applied Science* 10.22 (2010): 2814-822.
<<http://docsdrive.com/pdfs/ansinet/jas/2010/2814-2822.pdf>>. Accessed February 7, 2011.
- Bonneson, J., et al. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions*. FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.
- Bonneson, J., et al. *Development of Guidelines for Establishing Effective Curve Advisory Speeds*, Report No. FHWA/TX-07/0-5439-1., October 2007.
- Bonneson, J., et al. *Horizontal Curve Signing Handbook*, Report No. FHWA/TX-07/0-5439-P1., October 2007.
- Bureau of the Census. *2000 Urbanized Area Boundary Map*.
http://www.bts.gov/publications/national_transportation_atlas_database/2007. Accessed February 19, 2010.
- Champion, H., et al. *Reducing Highway Deaths and Disabilities with Automatic Wireless Transmission of Serious Injury Probability Ratings from Crash Recorders to Emergency Medical Service Providers*. International Symposium on Transportation Recorders, May, 1999. <http://www.nhtsa.gov/cars/problems/studies/acns/champion.htm>. Accessed January 10, 2011.

Cook, Dan. "Secondary low volume rural road safety: Segmentation, crash prediction, and identification of high crash locations." Thesis. Iowa State University, 2010.

Crash Modification Factors (CMF) Clearinghouse. 2010.
<<http://www.cmfclearinghouse.org/>>. Accessed March 3, 2011.

Federal Highway Administration. *Strategic Highway Safety Plan (SHSP) – FHWA Safety Program*. <http://safety.fhwa.dot.gov/hsip/shsp/>. 2006. Accessed January 3, 2011.

Hauer, Ezra, et al. *Estimating Safety by the Empirical Bayes Method: A Tutorial*. 2001.

Hauer, Ezra. *Observational Before-After Studies in Road Safety*. Toronto: Pergamon, 1997.

HSM Practitioner's Guide. *Predicting Highway Safety for Curves on Two-Lane Rural Highways*. 2011.

"Horizontal Curve Safety - FHWA Safety Program." *FHWA Safety Programs*. Accessed February 10, 2011. <http://safety.fhwa.dot.gov/roadway_dept/horicurves/>.

Hummer, J., et al. *An Examination of Horizontal Curve Collisions Characteristics and Corresponding Countermeasures*, October 2009.

Hu, Shuhua. *Akaike Information Criterion*. North Carolina State University. 2007.

Iowa Comprehensive Highway Safety Plan (CHSP). Ames, IA: Iowa Department of Transportation, 2006.

Iowa Department of Transportation. *Accident and Related Data for Rural and Municipal Intersections in Iowa*. Ames, Iowa: Iowa Department of Transportation, 1989.

Iowa Department of Transportation. *Traffic Safety Improvement Program (TSIP)*. <<http://www.iowadot.gov/tsip.htm>>. Accessed January 3, 2011.

"Iowa DOT Bid Express." *Bid Express Secure Internet Bidding*. Feb. 2011. Web. 07 Mar. 2011. <<https://www.bidx.com/index.html>>.

"Iowa DOT Top 200 Safety Improvement Candidate Locations." Iowa Department of Transportation. Ames, Iowa: Iowa Department of Transportation, 2010. <<http://www.iowadot.gov/crashanalysis/top200.htm>>. Accessed March 7, 2011.

Iowa's Five Percent Most Severe Safety Needs Report. Ames, IA: Iowa Department of Transportation, 2010.

Jackson, Justin. "Effect of Spatial Data Aggregation on Highway Safety Analysis." Thesis. Iowa State University, 2006.

Lincoln MPO Travel Demand Model Report. Lincoln, NE: Lincoln Metropolitan Planning Organization, 2006. <<http://lincoln.ne.gov/city/plan/mpo/travel/pdf/chap9.pdf>>. Accessed February 22, 2011.

National Highway Traffic Safety Administration (NHTSA). *FARS: Fatality Analysis Reporting System Encyclopedia*. <<http://www-fars.nhtsa.dot.gov/Main/index.aspx>>. Accessed January 3, 2011.

National Highway Traffic Safety Administration (NHTSA). *Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States*, 2006. <<http://www-nrd.nhtsa.dot.gov/Pubs/811226.pdf>>. Accessed March 21, 2011.

National Highway Traffic Safety Administration (NHTSA). *Motor Vehicle Traffic Fatalities and Fatality Rate: 1899-2003*. <<http://www.saferoads.org/federal/2004/TrafficFatalities1899->

2003.pdf>. Accessed January 3, 2011.

National Safety Council. *Estimating the Costs of Unintentional Injuries*.

<http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/EstimatingtheCostsofUnintentionalInjuries.aspx>. Accessed March 3, 2011.

NCHRP. *Model Curriculum for Highway Safety Core Competencies: Model 2-1*. Report 667. Washington, D.C.: Transportation Research Board, 2010.

Patterson, D., et al. "GPS–GIS-Based Procedure for Tracking Vehicle Path on Horizontal Alignments." *Computer-Aided Civil and Infrastructure Engineering* 21.5 (2006): 383-94.

Pawlovich, Michael. "Office of Traffic and Safety: Data." *Iowa Department of Transportation*. 2007. <<http://www.iowadot.gov/crashanalysis/data.htm>>. Accessed February 10, 2010.

Preston, H., et al. *Benefit: Cost Analysis of In-Vehicle Technologies and Infrastructure Modifications as a Means to Prevent Crashes Along Curves and Shoulders*. Report No. MN/RC 2009-39., December 2009.

Preston, Howard and Richard Storm. *Minnesota Comprehensive Highway Safety Plan (CHSP)*. St. Paul, MN: Minnesota Department of Transportation, 2005.

Preston, H., et al. *Alternative Strategies For Safety Improvement Investments*. Report Number 345. Washington, DC: Transportation Research Board, April 2010. NCHRP Research Results Digest.

Pratt, M., et al. *Workshops on Using the GPS Method to Determine Curve Advisory Speeds*. Report No. FHWA/TX-10/5-5439-01-1., December 2009.

Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users - A Summary of Highway Provisions. Washington, DC: Federal Highway Administration. <<http://www.fhwa.dot.gov/safetealu/summary.htm>>. Accessed January 3, 2011.

Sanders, Brandon. *Updating Horizontal Curve Data Using GPS Centerlines*. Kentucky Transportation Cabinet. 2007.

Sinha, Kumares, and Samuel Labi. *Transportation Decision Making: Principles of Project Evaluation and Programming*. Hoboken, NJ: John Wiley & Sons, 2007.

Souleyrette, Reginald, et al. Systematic Identification of High Crash Locations. Publication TR-442. Iowa Department of Transportation, 2001.

Tarko, Andrew, et al. *Safety Concious Planning in Indiana: Predicting Safety Benefits in Corridor Studies, Volume 1*. 2007.

Washington, Simon, et al. *Statistical, Econometric Methods: Transportation Data Analysis*. Second Edition. Boca Raton, FL: Taylor & Francis Group, 2011.

Zegeer, C. V., J. R. Stewart, F. M. Council, D. W. Reinfurt, and E. Hamilton. Safety Effects of Geometric Improvements on Horizontal Curves. *Transportation Research Record 1356*, Transportation Research Board, National Research Council, Washington, D.C., 1992, pp. 11–19.