Contributing Factors for Focus Crash and Facility Types

PUBLICATION NO. FHWA-HRT-20-052 NOVEMBER 2020

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

The research documented in this report was conducted as part of the Federal Highway Administration's (FHWA's) Evaluation of Low-Cost Safety Improvements Pooled Fund Study (ELCSI-PFS). FHWA established this PFS in 2005 to conduct research on the effectiveness of safety improvements identified by the National Cooperative Highway Research Program Report 500 Guides as part of the implementation of the American Association of State Highway and Transportation Officials' Strategic Highway Safety Plan. ELCSI-PFS studies provide a crash modification factor and benefit–cost economic analysis for each targeted safety strategy identified as a priority by member States of the PFS.

This project identified focus crash and facility types and associated contributing factors to better inform applications of systemic safety improvements. Three data sources were used to conduct contributing-factor analyses: crash and roadway inventory from the Highway Information System, climate data from the National Oceanic and Atmospheric Administration, and socioeconomic census data from the U.S. Census Bureau (FHWA 2018c; NOAA 2018; U.S. Census Bureau 2018). The contributing-factor analysis on road segments used data from Ohio and Washington State, and the analysis on intersections used data from California and Ohio. For the analysis, the research team used the random-forest method to identify the most predictive variables and then created plots of random forest–predicted crash frequencies as a function of the predictor variables to observe the general trends in the relationships.

A *Quick Reference Guide* (FHWA-HRT-20-053) was also developed from this project and aims to assist safety practitioners in selecting countermeasures to address focus crash types (Porter et al. 2020). This report and the *Quick Reference Guide* will benefit safety engineers and safety planners by providing greater insight into increased highway safety.

Brian P. Cronin, P.E.
Director, Office of Safety and Operations
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-20-052	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Contributing Factors for Focus Crash and Facility Types		5. Report Date November 2020
		6. Performing Organization Code
7. Author(s) Taha Saleem (ORCID: 0000-0002-2878-0980), Richard Porter (ORCID: 0000-0001-8535-3451), Raghavan Srinivasan (ORCID: 0000-0002-3097-5154), Daniel Carter (ORCID: 0000-0001-6572-6548), Scott Himes (ORCID: 0000-0003-2932-7691), and Thanh Le (ORCID: 0000-0001-5809-2751)		8. Performing Organization Report No.
9. Performing Organization Name and Address VHB 940 Main Campus Drive, Suite 500 Raleigh, NC 27606		10. Work Unit No.
		11. Contract or Grant No. DTFH61-13-D-000001
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Final Report; March 2017–February 2018
		14. Sponsoring Agency Code HRDS-20

15. Supplementary Notes
The Federal Highway Administration (FHWA) Office of Safety Research and Development managed this project under the Development of Crash Modification Factors program. The FHWA Office of Safety Research and Development Program Manager and Contracting Officer's Representative was Roya Amjadi (HRDS-20; ORCID: 0000-0001-7672-8485), and the Technical Manager was Karen Scurry (HSSP; ORCID: 0000-0002-3499-5515).

16. Abstract
The main goal of this project was to identify focus crash and facility types (FCFTs) and associated contributing factors to better inform applications of systemic safety improvements. The selection of FCFTs was guided by the use of the Fatality Analysis Reporting System and the Highway Safety Information System (HSIS) databases (NHTSA 2018a; FHWA 2018c). The method used to select potential FCFTs was based on the number of fatal crashes and number of fatal-plus-incapacitating-injury crashes associated with various crash-related variables. A total of 17 FCFTs (8 intersection FCFTs and 9 nonintersection FCFTs) were identified for analysis to identify contributing factors. Three data sources were used to conduct the contributing-factor analysis: crash and roadway inventory from HSIS, climate data from the National Oceanic and Atmospheric Administration, and socioeconomic census data from the U.S. Census Bureau (FHWA 2018c; NOAA 2018; U.S. Census Bureau 2018). For the analysis, the research team used the random-forest method to identify the most predictive variables and then created plots of random forest–predicted crash frequencies as a function of the predictor variables to observe the general trends in the relationships. The research laid out a process for identifying and selecting countermeasures for focus crash types based on contributing factors. The countermeasure guidance focused on traffic and roadway findings. Findings linked to socioeconomic- and weather-related factors showed promise, but there is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results of this effort.

17. Key Words Systemic safety, focus crash types, contributing factors, random forests, crash prediction, proven countermeasures		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 308	22. Price N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1,000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2,000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	2.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

A	incapacitating injury
AADP	annual average daily pedestrian
AADT	annual average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
ADT	average daily traffic
ANG	angle
ANG-D	angle daytime
ANG-KABCO-D	angle fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only daytime
ANG-KABCO-N	angle fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only nighttime
ANG-KAB-D	angle fatality, incapacitating injury, or nonincapacitating injury daytime
ANG-KAB-N	angle fatality, incapacitating injury, or nonincapacitating injury nighttime
ANG-N	angle nighttime crashes
B	nonincapacitating injury
B/C	benefit–cost
C	possible injury
CART	classification and regression trees
CMF	crash modification factor
CRSP	Country Road Safety Plan
DCMF	Development of Crash Modification Factors
DOT	Department of Transportation
FARS	Fatality Analysis Reporting System
FCFT	focus crash and facility type
FHWA	Federal Highway Administration
GLM	generalized linear model
HEO	head on
HEO-D	head-on daytime
HEO-KABCO-D	head-on fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only daytime
HEO-KABCO-N	head on fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only nighttime
HEO-KAB-D	head on fatality, incapacitating injury, or nonincapacitating injury daytime
HEO-KAB-N	head on fatality, incapacitating injury, or nonincapacitating injury nighttime
HEO-N	head-on nighttime
HSIS	Highway Safety Information System
HSM	*Highway Safety Manual*
K	fatal injury
LNDP	lane departure
LNDP-D	lane departure daytime
LNDP-KABCO-D	lane-departure fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only daytime

LNDP-KABCO-N	lane-departure fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only nighttime
LNDP-KAB-D	lane-departure fatality, incapacitating injury, or nonincapacitating injury daytime
LNDP-KAB-N	lane-departure fatality, incapacitating injury, or nonincapacitating injury nighttime
LNDP-N	lane departure nighttime
MnDOT	Minnesota Department of Transportation
MSE	mean squared error
NB	negative binomial
NCHRP	National Cooperative Highway Research Program
NOAA	National Oceanic and Atmospheric Administration
O	no injury
ODOT	Ohio Department of Transportation
PDO	property damage only
ROLL	rollover/overturn
ROLL-D	rollover/overturn daytime
ROLL-KABCO-D	rollover/overturn fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only daytime
ROLL-KABCO-N	rollover/overturn fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only nighttime
ROLL-KAB-D	rollover/overturn fatality, incapacitating injury, or nonincapacitating injury daytime
ROLL-KAB-N	rollover/overturn fatality, incapacitating injury, or nonincapacitating injury nighttime
ROLL-N	rollover/overturn nighttime
ROR	run off road
ROR-D	run off road daytime
ROR-KABCO-D	run-off-road fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only daytime
ROR-KABCO-N	run-off-road fatality, incapacitating injury, nonincapacitating injury, possible injury, or property-damage-only nighttime
ROR-KAB-D	run-off-road fatality, incapacitating injury, or nonincapacitating injury daytime
ROR-KAB-N	run-off-road fatality, incapacitating injury, or nonincapacitating injury nighttime
ROR-N	run off road nighttime
RRSI	rural road safety index
SAS	statistical analysis software
SHRP2	second Strategic Highway Research Program
SHSP	strategic highway safety plan
TTE	time-to-event
usRAP	United States Road Assessment Program

EXECUTIVE SUMMARY

This report describes the efforts to identify focus crash and facility types (FCFTs) and associated crash-contributing factors (herein referred to as "contributing factors") to better inform applications of systemic safety improvements. Systemic safety improvements—when selected and targeted appropriately—are a tremendous opportunity to proactively reduce crashes and their resulting harm. These improvements are particularly useful for crash types that are prevalent but somewhat disperse in their occurrence, or in other words, high numbers of certain crash types scattered across the road system at low densities of resultant fatalities. An FHWA publication, *Using Risk to Drive Safety Investments*, notes that "fatal and other life-threatening crashes often are distributed widely across State and local highway systems, in both urban and rural environments, with few individual locations experiencing a high number or sustained occurrence of severe crashes" (Preston et al. 2013b). These types of prevalent but scattered systemic safety issues do not lend themselves to the site-specific detection and diagnosis that characterize a more traditional approach to road-safety management.

Within the context of systemic safety analysis and management, the main objectives of this project were to select reliable and applicable data resources, statistical methodologies, analysis procedures, and tools; conduct data analysis to identify and validate FCFTs and associated contributing factors; and identify potential low-cost safety strategies that may effectively be used as systemic safety improvements.

FCFTS

The research team analyzed the Fatality Analysis Reporting System (FARS) and the Highway Safety Information System (HSIS) databases to select FCFTs (NHTSA 2018a; FHWA 2018c). The FARS database contains information on all fatal crashes involving a motor vehicle traveling on a public trafficway in all 50 States, the District of Columbia, and Puerto Rico. The HSIS database contains information on crashes of all severities (i.e., fatal crashes, injury crashes, and property-damage-only crashes) occurring on State-operated and -maintained roads for the participating States. The analysis incorporated data from four States that are part of HSIS: California, Minnesota, Ohio, and Washington State. The research team defined potential FCFTs by using combinations of variables in FARS and HSIS. To rank the list of FCFTs, the research team used the number of fatal crashes and, for the HSIS States, the number of fatal-plus-incapacitating-injury crashes during the observation period that corresponded to each potential FCFT. Based on the rankings, the research team selected 15 FCFTs for the first contributing-factors analysis:

- Intersection FCFTs:
  1. Angle (ANG) crashes on rural two-lane roads at four-leg minor-road stop-controlled intersections (daytime and nighttime).
  2. ANG crashes on urban two-lane roads at four-leg minor-road stop-controlled intersections (daytime).
  3. ANG crashes on rural two-lane roads at three-leg minor-road stop-controlled intersections (daytime).

4. ANG crashes on urban multilane divided roads at four-leg signalized intersections (daytime).
5. ANG crashes on urban multilane undivided roads at four-leg signalized intersections (daytime).
6. ANG crashes on rural multilane divided roads at four-leg minor-road stop-controlled intersections (daytime).

- Nonintersection FCFTs:
  1. Run-off-road (ROR) crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
  2. ROR crashes on rural two-lane roads on tangent segments (daytime and nighttime).
  3. Lane-departure (LNDP) crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
  4. LNDP crashes on rural two-lane roads on tangent segments (daytime and nighttime).
  5. Head-on (HEO) crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
  6. HEO crashes on rural two-lane roads on tangent segments (daytime and nighttime).
  7. ANG crashes on rural two-lane roads on tangent segments (daytime).
  8. Rollover/overturn (ROLL) crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
  9. ROLL crashes on rural two-lane roads on tangent segments (daytime and nighttime).

Pedestrian crashes represented a significant number of fatal crashes; however, identifying contributing factors for pedestrian crashes is a challenging task without quality exposure data (i.e., the total number of pedestrians crossing at segments and intersections or walking parallel to segments). Given the significant number of pedestrian crashes, this report incorporates the results of Thomas et al. (2017) to identify contributing factors for two types of pedestrian crashes:

- All types of pedestrian crashes at intersections.
- Pedestrian crashes at intersections involving a crossing pedestrian and a vehicle going straight.

ANALYSIS OF CONTRIBUTING FACTORS

The research team analyzed contributing factors using data from three different sources: crash and roadway inventory from HSIS, climate data from the National Oceanic and Atmospheric Administration, and socioeconomic census data from the U.S. Census Bureau (FHWA 2018c; NOAA 2018; U.S. Census Bureau 2018). The research team completed all linkages of road segments to climate and census data in the spatial environment. To simplify the linking process, roadway segments were represented as point features according to the midpoint of the segment. For climate data, the source data were in a point file with each weather station shown as a point on the map. The research team linked each roadway segment to the closest weather station by a simple straight-line distance measurement. For census data, the source data were in a polygon file with each census-block group shown as a shape on the map. Each roadway segment was linked to the census-block group that contained the midpoint of the segment.

The research team used the random-forest method to identify contributing factors that corresponded to the FCFTs. Random forests help identify predictors that may not appear in the output of a single classification or regression tree but are nevertheless highly related to the target variable. Researchers commonly use random forests to display the percentage increase in mean squared error with the removal of a variable from the random-forest model. Random forests do not directly indicate if variables correspond to contributing factors that increase or decrease predicted crash frequencies. However, plots of random forest–predicted crash frequencies as a function of the variables of interest provided the information needed to identify contributing factors, estimate the direction of the relationship, and inform countermeasure identification to mitigate the presence of factors that increase predicted crash frequencies. The research team used data from California and Ohio to analyze contributing factors for intersection FCFTs and data from Ohio and Washington to analyze contributing factors for nonintersection (i.e., segment) FCFTs.

Roadway factors uncovered by the analyses as influencing the frequencies of the different crash types were generally consistent with what was expected based on previous research and existing practice. Factors associated with increasing crash frequencies include the following:

- Larger average daily traffic volumes.
- Steeper vertical grades.
- Sharper horizontal curve radii.
- Narrower lane and shoulder widths.
- Unpaved shoulders or no shoulders.
- Mountainous terrains.
- Higher speed limits.
- Wider crossing distances at intersections (captured by lane and median widths on approaches).
- Absence of left- and right-turn channelization at intersections.

Agencies with sufficient data and analysis capabilities can refer to the FHWA *Systemic Safety Project Selection Tool* for discussion of how to analyze data to identify contributing factors given a specific FCFT (Preston et al. 2013a). Agencies without sufficient data and analysis capabilities can reference the factors developed in this research to help identify countermeasures and prioritize sites for systemic safety improvements.

Findings connected to socioeconomic- and weather-related factors showed promise, but there is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results of this effort. Findings related to socioeconomic variables are likely representing differences in travel behavior, driving behavior, and driving capabilities that seem key for safety analyses but are generally not incorporated into segment- and intersection-specific analyses that also include traffic and roadway factors. Weather-related findings are likely representing differences in visibility, road conditions, and driver experience and behavior. Both sets of factors bring significant potential to the process of making more informed decisions about sites that have higher levels of crash potential. (Appendix H includes an example analysis of model performance with and without consideration of socioeconomic variables.) A multiyear study focused on testing various alternatives and developing safety-analysis guidance on collecting, merging, and analyzing crash, traffic, roadway, census, and weather data is needed.

The most challenging aspect of implementing the random-forest method within the context of this project was the limited ability to interpret the direction and form of the relationship between a factor of interest and expected crash frequency. This research implemented a relatively simple approach, searching for primarily linear trends between each factor and the corresponding random forest–predicted crash frequency. Future efforts should explore more effective ways to uncover and interpret the relationships contained in the random forests.

While the research team was able to compare the random-forests results to the findings of previous research on related crash types, it was not able to incorporate prior information or knowledge into the analyses. Bayesian approaches hold significant promise, and future efforts should explore these approaches to identify contributing factors for FCFTs. Causal Bayesian networks, such as those applied by Karwa et al. (2011), have significant potential but will require a considerable amount of time to fully explore within the context of this project.

COUNTERMEASURE-SELECTION PROCESS

Following the selection of FCFTs and analysis of contributing factors, this report presents a six-step countermeasure-selection process, as follows:

1. Identify a focus crash type.
2. Identify contributing factors for the focus crash type.
3. Assemble a list of potential countermeasures that address the focus crash type.
4. Identify countermeasures that address roadway factors associated with the focus crash type.
5. Identify countermeasures with crash modification factors (CMFs).
6. Select a countermeasure.

To assist practitioners with implementing this process, this report also provides various examples demonstrating the process for selecting countermeasures to address focus crash types. Additionally, this report provides proven safety countermeasures to mitigate the presence of common contributing factors. In developing practitioner guidance, the research team observed that most CMFs have not been developed for crash types at the level of disaggregation that may be needed for systemic safety applications. Similarly, expected changes in crash frequencies reflected by CMFs typically represent evaluation results of site-specific applications of a treatment. Additional work on estimating network-wide safety impacts resulting from systemic applications of countermeasures is needed.

CHAPTER 1. INTRODUCTION

BACKGROUND

The Federal Highway Administration (FHWA) established the Development of Crash Modification Factors (DCMF) program in 2012 to address highway safety research needs for evaluating new and innovative safety-improvement strategies and developing reliable quantitative estimates of their effectiveness in reducing crash frequency and severity. The ultimate goal of the DCMF program is to save lives by identifying new strategies that effectively reduce crash frequency and severity and promoting those strategies for nationwide implementation by providing measures of their safety effectiveness and benefit–cost (B/C) ratios gleaned through research. State transportation departments and other transportation agencies should have objective measures of safety effectiveness before investing in broad applications of safety countermeasures. Forty State transportation departments provide technical input to the DCMF program and implement new safety improvements to facilitate evaluations. These States are members of the Evaluation of Low-Cost Safety Improvements Pooled Fund Study, which functions under the DCMF program.

Another goal of the DCMF program is to advance highway safety and related research by establishing a sound foundation for developing highway transportation-specific statistical methodologies in cooperation with the American Statistical Association and other statistician communities. Several efforts have been conducted or are underway in pursuit of that goal. One such effort, *Highway Safety Statistical Paper Synthesis*, included a review and critical synthesis of recent papers that explored refinements to current research methods (including study design and statistical analysis) and proposed new methods to assess the safety performance of highways and streets (Persaud et al. 2001). This effort included methods to predict expected crash frequencies and severities, assess underreporting in crash frequency models, use nontraditional datasets to analyze safety, and estimate crash modification factors (CMFs) using alternative approaches to study design and analysis. The intention of the critical synthesis was to serve as a resource to researchers and others looking to advance the science of highway safety.

As part of the *Highway Safety Statistical Paper Synthesis*, researchers demonstrated classification and regression trees (CART) and random forests within the context of conducting statistical road-safety analyses. They concluded that tree-based models hold strong potential for road safety analyses. These models are particularly effective in making predictions of expected crash frequency, which has applications in multiple contexts (e.g., network screening, alternatives assessment, predicting "what would have been" in before–after studies). Tree-based methods also have the potential to inform specifications that are part of more traditional modeling approaches through identifying the "most predictive" right-hand-side variables and uncovering informative relationships between left-hand-side and right-hand-side variables.

The DCMF effort applied CART and random forests to make predictions of expected crash frequency as a function of traffic, geometric design, and operational features along directional freeway segments that have a right-hand-side entrance ramp followed by a right-hand-side exit ramp. This current project situates the exploration of potential applications of tree-based methods within the context of quantitative approaches to systemic safety analysis.

SYSTEMIC APPROACH TO ROAD-SAFETY MANAGEMENT

Common approaches to road-safety management focus on identifying and diagnosing high-crash locations and implementing projects to address the predominant safety concerns at those locations. Such approaches to road-safety management generally consist of six steps:

1. Network screening.
2. Site diagnoses.
3. Countermeasure selections.
4. Economic appraisals.
5. Project prioritization.
6. Safety-effectiveness evaluations.

Researchers use network screening to identify "sites with promise" (i.e., sites that show potential for safety improvements), usually characterized as sites with substantially higher than expected crash frequencies given traffic volumes and other site characteristics. Network screening methods use "total" crashes (i.e., all crash types and severities) or crashes of a specific type and/or severity to inform site ranking. Once researchers select specific sites for investigation, the subsequent steps occur at the individual site level. Safety practitioners attempt to diagnose underlying crash-contributing factors (herein referred to as "contributing factors") occurring at a specific roadway segment or intersection and develop targeted countermeasures. Researchers estimate economic and safety benefits to ensure a project is cost effective and to prioritize investments. Prioritization can occur at two levels: the project level to prioritize proposed countermeasures within a project and the program level to determine an optimal combination of projects within a fixed budget to maximize the benefits of the entire program of projects. The final step is to evaluate the impacts of completed projects, including safety, for individual projects and for overall program effectiveness.

Significant progress has been made in the development of methodologies (e.g., empirical Bayes estimation) and tools (e.g., safety-performance functions, CMFs) to support data-driven, quantitative approaches to the safety-management process. The current states of related knowledge, tools, and practices are captured in the *Highway Safety Manual* (HSM), AASHTOWare® Safety Analyst™, CMF Clearinghouse, *Interactive Highway Safety Design Model*, and a number of other tools and resources listed in the Roadway Safety Data and Analysis Toolbox (AASHTO 2010; AASHTO n.d.; FHWA 2018a; FHWA 2018d; FHWA 2018e).

While identifying and treating sites with promise can result in significant safety benefits, some focus crash and facility types (FCFTs) cannot be fully addressed by this approach to road-safety management. Rather than a high number of crashes at specific locations, these FCFTs are characterized by higher numbers of crashes scattered across the road system at low densities. Examples of crash types include those involving nonmotorized-vehicle users (typically motor vehicles hitting pedestrians or bicycles), roadway-departure crashes on low-volume roads, and fatal and incapacitating-injury crashes (i.e., KA crashes) on almost any facility type. An FHWA publication, *Using Risk to Drive Safety Investments*, notes that "fatal and other life-threatening crashes often are distributed widely across State and local highway systems, in both urban and rural environments, with few individual locations experiencing a high number or sustained occurrence of severe crashes" (Preston et al. 2013b). These types of systemic safety issues do not lend themselves to site-specific detection and diagnosis. Alternative approaches that identify and diagnose potential for safety improvements at a system level are needed.

The systemic approach to road-safety management is a method of safety management that typically involves widely implementing lower (unit) cost safety improvements based on characteristics that indicate greater potential for future crashes. While the outcome of a site-based approach to identifying safety projects is a range of appropriate countermeasures for varying safety issues and crash types at individual locations, the outcome of a systemic approach is projects that implement one or more proven, cost-effective countermeasures across a number of sites to effectively manage the potential for future crashes of a particular focus crash type. This key distinction results in sites being treated not necessarily because of crash history but because of estimated crash potential, a strategy identified in FHWA's Highway Safety Improvement Program for reducing fatalities and serious injuries in the United States. The systemic approach is also different than a systematic approach, in which practitioners implement countermeasures at all sites regardless of the potential for future crashes.

Agency experiences with systemic approaches to safety management continue to increase, but a review of practices shows that, to date, systemic safety analysis often includes subjective approaches to identifying contributing factors and characterizing crash potentials. Agencies need additional guidance to help select and target their systemic safety improvements to make the most efficient investment decisions, including detailed and data-driven information on crash types that can be effectively addressed with systemic approaches to safety management as well as the situations (characterized by contributing factors and facility types) where these crashes are more likely to occur.

OBJECTIVE

The goal of this project was to identify FCFTs and associated contributing factors to inform applications of systemic safety improvements. Within the context of conducting systemic safety analysis and management, the objectives of this project were as follows:

- Select reliable and applicable data resources, statistical methodologies, analysis procedures, and tools.
- Conduct data analysis to identify and validate FCFTs and associated contributing factors.
- Identify potential low-cost safety strategies that may effectively be used as systemic safety improvements.
- Develop a technical report and *Quick Reference Guide* for identifying focus crash types and contributing factors (Porter et al. 2020).
- Develop a solicitation and evaluation criteria to identify volunteer agencies to implement systemic safety improvements.

The *Quick Reference Guide* is a stand-alone document separate from this technical report (Porter et al. 2020). The solicitation and evaluation-criteria documents are not available for distribution.

ORGANIZATION OF THE REPORT

This technical report is organized into seven chapters:

- Chapter 1 provides background information and project objectives.
- Chapter 2 includes a literature and database review, with a focus on promising data sources that support the objectives of this project.
- Chapter 3 outlines the framework for defining crash types, facility types, potential for future crashes, and contributing factors.
- Chapter 4 focuses on identifying FCFTs. It contains a discussion of databases used and implemented methodology for FCFT identification. A list of the 15 FCFTs identified for further analysis is also contained in this chapter.
- Chapter 5 describes the analysis of contributing factors for the 15 FCFTs. An overview of databases used for the analysis is followed by identifying the most influential predictor variables for each FCFT and determining whether the influential predictor variables are associated with increases or decreases in expected crash frequencies.
- Chapter 6 provides a process for identifying and selecting countermeasures for focus crash types, multiple examples of applying the countermeasure-selection process, and a concise overview of countermeasures that address the FCFTs and contributing factors.
- Chapter 7 concludes this report by providing a summary of accomplishments, conclusions, and recommendations for future work.

CHAPTER 2. LITERATURE REVIEW

The objectives of this literature review were to draw key information from past studies to identify common contributing factors for various FCFTs, including methodologies used to select these factors, and identify and review emerging research and practice on systemic safety analysis.

OVERVIEW OF CONTRIBUTING FACTORS AND METHODS

This section provides a review of relevant past studies, identifying common contributing factors for various FCFTs. Additionally, this section provides a review of methodologies commonly used to identify these contributing factors. Together, this information provided the basis for selecting user types, crash types, facility types, contributing factors, and methodologies for this research. This section also defines each of the aforementioned items and presents a discussion of literature findings for each. Last, contributing factors by applicable facility types, which the research team determined by the presence of statistical associations in cross-sectional or before–after models in the literature findings, are summarized.

The literature review is concentrated on studies, guides, and reports with a focus on the broad consideration of contributing factors. Since the motivation was to identify potential crash types, facility types, contributing factors, datasets, and statistical methodologies, and there are innumerable studies that have examined the impact of an isolated contributing factor, it was not cost- or time-effective to focus on studies looking only at individual contributing factors.

The following sections identify further considerations for this research with respect to user types, crash types, facility types, contributing roadway factors, datasets, and methodologies.

Contributing Factors

User Types

User types refer to the users of the roadway. The American Association of State Highway and Transportation Officials' (AASHTO's) *A Policy on Geometric Design of Highways and Streets* (2011) considers the following roadway users:

- Pedestrians.
- Vehicles:
  - Passenger cars.
  - Buses.
  - Trucks.
  - Recreational vehicles.
  - Bicycles.

Pedestrian crashes and bicycle crashes almost universally involve a different vehicle type (e.g., passenger car) and are typically considered as a separate crash type (e.g., bicycle–vehicle collisions). In crash analyses, buses and recreational vehicles are often not considered independently as vehicle types but instead are lumped into "all vehicles" or "large vehicles". Noting how these vehicle types are treated when identifying crash contributing factors is a key detail of the analysis; if considered separately, crashes involving these vehicle types may have differing crash contributing factors than crashes that do not distinguish between these vehicle types.

Crash Severities and Types

Crashes are generally defined by two attributes: type and level of severity. Type refers to the number of vehicles, crash location, and manner of collision (e.g., single-vehicle crash, run-off-road [ROR] crash), while severity refers to the level of injury resulting from a crash. Police officers code type and severity on crash-reporting forms, and codes for both attributes are determined subjectively based on physical evidence and statements made by those involved in the crash.

Most agencies use the KABCO crash-severity scale—which consists of five levels of injury—for reporting the most severe injury in a given crash as well as injuries of drivers and occupants involved in a crash. KABCO severities, as defined in the 5th edition of *Model Minimum Uniform Crash Criteria,* include the following (NHTSA 2017):

- Fatal Injury (K)—A fatal injury is any injury that results in death within 30 days after the motor vehicle crash in which the injury occurred. If the person did not die at the scene but died within 30 days of the motor vehicle crash in which the injury occurred, the injury classification is changed from the attribute previously assigned to the attribute “Fatal Injury.”

- Suspected Serious Injury (A)—A suspected serious injury is any injury other than fatal that results in one or more of the following:
  - Severe laceration resulting in exposure of underlying tissues/muscles/organs or significant loss of blood.
  - Broken or distorted extremity (arm or leg).
  - Crush injury.
  - Suspected skull, chest, or abdominal injury other than a bruise or minor laceration.
  - Significant burn (second- and third-degree burn over 10 percent or more of the body).
  - Unconsciousness when taken from the crash scene.
  - Paralysis.

- Suspected Minor Injury (B)—A minor injury is any injury that is evident at the scene of the crash other than fatal or serious injuries. Examples of minor injuries include lumps on the head, abrasions, bruises, and minor lacerations (i.e., cuts on the skin surface with minimal bleeding and no exposure of deeper tissue/muscle).

- Possible Injury (C)—A possible injury is any injury reported or claimed that is not a fatal, suspected serious, or suspected minor injury. Examples of possible injuries include a momentary loss of consciousness, claim of injury, limping, or complaint of pain or nausea. There are two ways to identify possible injuries: self-reported by the person involved in the crash or indicated by the person involved in the crash but no wounds or injuries are readily evident.

- No Apparent Injury (O)—No apparent injury is a situation where there is no reason to believe that the person involved in the crash received any bodily harm. There is no physical evidence of injury and the person does not report any change in normal function.

Crash types often fall into one of the broader categories of crash types characterized by the number of vehicles involved: single-vehicle crashes, multiple-vehicle crashes, or total crashes (i.e., single- and multiple-vehicle crashes combined). Since systemic safety focuses on crash types scattered across the roadway network at low densities, attention is most typically paid to specific crash types that result in fatalities or severe injuries. The literature review found that models use many variations of crash-type definitions to identify contributing factors. Research also often distinguishes crash types by severity (e.g., all angle [ANG] crashes versus all KABC ANG crashes); therefore, crash-severity levels can be considered part of crash-type definitions. Example publications using different crash-type definitions under the three broadly defined crash types (i.e., single-vehicle crashes, multiple-vehicle crashes, and total crashes) are shown in table 1 through table 3.

Table 1. Example publications addressing subsets of total crashes.

Crash Type	Reviewed Topical Literature
All crashes	AASHTO 2010; National Academy of Sciences 2003e, 2004a, 2004c, 2004f, 2008b, 2009; FHWA 2004, 2008a, 2008b, 2009; Fink and Krammes 1995; Fitzpatrick et al. 2005; Harkey et al. 2008; Karlaftis and Golias 2002; Lord et al. 2011; Milton and Mannering 1996, 1998; Orner and Drakopoulos 2007; Shankar et al. 1995; Wang et al. 1998
Severe crashes (KAB)	AASHTO 2010; Preston et al. 2013a; National Academy of Sciences 2004d
KABC crashes	AASHTO 2010; AAA Foundation for Traffic Safety 2012; FHWA 2008b; Harkey et al. 2008; MnDOT 2011
Nonfatal-injury crashes (ABC)	FHWA 2008b; Harkey et al. 2008
Young-driver crashes	FHWA 2009; Goodwin et al. 2013; Preston et al. 2013a
Older-driver crashes	FHWA 2004, 2009; Goodwin et al. 2013; National Academy of Sciences 2004c; Preston et al. 2013a
Weather condition–related crashes	Lord et al. 2011
Weather condition–related crashes: wet-road crashes	FHWA 2008a, 2009; Harkey et al. 2008; National Academy of Sciences 2008b
Weather condition–related crashes: dry-road crashes	Harkey et al. 2008
Weather condition–related crashes: snow-related crashes	FHWA 2008a
Light condition–related crashes	Lord et al. 2011
Light condition–related crashes: nighttime crashes	FHWA 2008a, 2009; Harkey et al. 2008; National Academy of Sciences 2004a
Light condition–related crashes: nighttime-wet crashes	FHWA 2009
Light condition–related crashes: daytime crashes	FHWA 2009
Light condition–related crashes: daytime/nighttime ratio	National Academy of Sciences 2004a
Aggressive driving–related crashes	Goodwin et al. 2013; Preston et al. 2013a
Drug- and alcohol-related crashes	Goodwin et al. 2013; Preston et al. 2013a
Inattentive, distracted, asleep crashes	Goodwin et al. 2013; National Academy of Sciences 2004g; Preston et al. 2013a

Crash Type	Reviewed Topical Literature
Motorcycle crashes	Goodwin et al. 2013; National Academy of Sciences 2008d; Preston et al. 2013a
Heavy-vehicle crashes	FHWA 2008a; National Academy of Sciences 2004g; Miaou 1994; Preston et al. 2013a
Intersection-related crashes	AAA Foundation for Traffic Safety 2012; Fitzpatrick et al. 2005; Poch and Mannering 1996; Preston et al. 2013a
Nonintersection-related crashes	Fitzpatrick et al. 2005, 2008
Work zone–related crashes	Preston et al. 2013a
Surface width–influence crashes	Fitzpatrick et al. 2005
Cross-median crashes	Harkey et al. 2008
Cross-median crashes: cross-median HEO crashes	National Academy of Sciences 2008b
Speed-related crashes	FHWA 2008a, 2009; National Academy of Sciences 2009
Speed-related crashes: speed-related daytime crashes	FHWA 2009
Speed-related crashes: speed-related nighttime crashes	FHWA 2009
Speed-related crashes: speed-related dry crashes	FHWA 2009
Speed-related crashes: speed-related wet crashes	FHWA 2009
Emergency-vehicle crashes	National Academy of Sciences 2004f; FHWA 2009
Red-light-running crashes	National Academy of Sciences 2004f; FHWA 2004
Lane-change crashes	National Academy of Sciences 2004f
Wrong-way crashes	National Academy of Sciences 2008b

HEO = head on.

Table 2. Example publications addressing subsets of single-vehicle crashes.

Crash Type	Reviewed Topical Literature
All single-vehicle crashes	AASHTO 2010; Harkey et al. 2008; Zegeer et al. 1987
Roadway-departure/ROR crashes	AASHTO 2010; AAA Foundation for Traffic Safety 2012; Cato et al. 2013; FHWA 2008a, 2009; Hallmark et al. 2006; Knapp et al. 2014; National Academy of Sciences 2003e, 2003f, 2004b, 2008b, 2009; Lee and Nam 2003; Liu and Ye 2011; Lord et al. 2011; MnDOT 2011; Orner and Drakopoulos 2007; Patel et al. 2007; Preston et al. 2013a
Fatal ROR crashes	AASHTO 2010; Liu and Subramanian 2009; Patel et al. 2007
Collision-with-animal crashes	AASHTO 2010; FHWA 2008a
Overturn crashes	AASHTO 2010; FHWA 2008a, 2009; Shankar et al. 1995
Collision-with-parked-vehicle crashes	AASHTO 2010; National Academy of Sciences 2004f; FHWA 2008a; Shankar et al. 1995
Collision-with-fixed-object crashes	AASHTO 2010; FHWA 2008a, 2009; Shankar et al. 1995
Collision-with-fixed-object crashes: collision-with-tree crashes	National Academy of Sciences 2003c
Collision-with-other-object crashes	AASHTO 2010
Noncollision	AASHTO 2010
Wet-road crashes	Harkey et al. 2008

Table 3. Example publications addressing subsets of multiple-vehicle crashes.

Total Multiple-Vehicle Crashes	Reviewed Topical Literature
All multivehicle crashes	AASHTO 2010; National Academy of Sciences 2004f; FHWA 2009
ANG crashes	AASHTO 2010; National Academy of Sciences 2003e, 2004f, 2009; Poch and Mannering 1996
ANG crashes: ANG-N crashes	Harkey et al. 2008
Right-ANG crashes	National Academy of Sciences 2004f; FHWA 2004, 2008a, 2009; Harkey et al. 2008; Preston et al. 2013a
Right-ANG crashes: wet-road right-ANG crashes	Harkey et al. 2008
HEO crashes	AASHTO 2010; AAA Foundation for Traffic Safety 2012; FHWA 2008a, 2009; Preston et al. 2013a
Rear-end crashes	AASHTO 2010; National Academy of Sciences 2003e, 2004f; FHWA 2004, 2008a; Harkey et al. 2008; Poch and Mannering 1996; Preston et al. 2013a; Shankar et al. 1995
Rear-end crashes: wet-road rear-end crashes	Harkey et al. 2008
Turning crashes	FHWA 2008a, 2009; National Academy of Sciences 2003e, 2009; Poch and Mannering 1996; Preston et al. 2013a
Turning crashes: left-turn crashes	National Academy of Sciences 2004f; FHWA 2004, 2008a, 2009; Harkey et al. 2008
Turning crashes: right-turn crashes	FHWA 2008a, 2009
Sideswipe crashes	AASHTO 2010; National Academy of Sciences 2004f; FHWA 2004, 2008a, 2009; National Acadmey of Science 2003e; Preston et al. 2013a; Shankar et al. 1995
Sideswipe crashes: same-direction sideswipe crashes	Zegeer et al. 1987
Sideswipe crashes: opposite-direction sideswipe crashes	Zegeer et al. 1987
Nondriveway Crashes	AASHTO 2010
Driveway crashes	AASHTO 2010; Fitzpatrick et al. 2005
Vehicle–pedestrian crashes	AASHTO 2010; AAA Foundation for Traffic Safety 2012; National Academy of Sciences 2004d, 2004f; FHWA 2004, 2008a, 2008b, 2009; Goodwin et al. 2013; Preston et al. 2013a
Vehicle–pedestrian crashes: crossing-roadway crashes	AAA Foundation for Traffic Safety 2014; National Academy of Sciences 2004d
Vehicle–pedestrian crashes: walking-along-roadway crashes	AAA Foundation for Traffic Safety 2012; FHWA 2008a, 2008b; National Academy of Sciences 2004d
Vehicle–bicycle crashes	AASHTO 2010; AAA Foundation for Traffic Safety 2012; National Academy of Sciences 2004f, 2008a; FHWA 2009; Goodwin et al. 2013; Preston et al. 2013a
Vehicle–bicycle crashes: crossing-roadway crashes	AAA Foundation for Traffic Safety 2012
Vehicle–bicycle crashes: biking-along-roadway crashes	AAA Foundation for Traffic Safety 2012
Vehicle–train crashes	Preston et al. 2013a

HEO = head on.

Researchers more commonly analyze KA crashes separately from minor injuries or property-damage-only (PDO) crashes because the reliability of the data is greatest for higher-severity crashes (KA or KAB). As crash severity decreases, the reliability of reporting decreases.

Underreporting increases as crash severity decreases, and it is generally accepted that fewer than 5 percent of K crashes are unreported while more than 50 percent of PDO crashes are unreported. Underreported data will produce biased estimates in crash-frequency and crash-severity models. Additionally, this distribution of underreporting leads to overrepresentation of crashes with higher severity and underrepresentation of crashes with lower severity. This fact, along with other factors like the higher economic and social costs of KA crashes and the purpose of the Highway Safety Improvement Program (i.e., to reduce K and A crashes), leads researchers and practitioners to focus on reducing KA crashes.

Conversely, considering only K, KA, and KB crashes for specific crash types leads to very small sample sizes. Disaggregation of analysis to this level may lead to analyses that do not yield contributing factors associated with specific crash types and severities. Inclusion of all severities or combining crash types may affect the underlying relationships between contributing factors and severe crash outcomes but help increase the confidence in statistical associations between contributing factors and crash outcomes. Based on the strengths and limitations associated with different levels of disaggregation, this research focuses on crash types well represented within the three databases.

Facility Types

Researchers present facility types in several different formats. Most commonly, researchers use the HSM definitions for facility types, which include the following (AASHTO 2010; AAA Foundation for Traffic Safety 2012; FHWA 2004, 2008a, 2008b, 2009, Fink and Krammes 1995; Fitzpatrick et al. 2005, 2008; Goodwin et al. 2013; Hallmark et al. 2006; Harkey et al. 2008; Karlaftis and Golias 2002; Lord et al. 2011; Miaou 1994; National Academy of Sciences 2003c, 2003e, 2003f, 2004a, 2004b, 2004c, 2004d, 2004f, 2004g, 2008a, 2008b, 2008d, 2009; Patel et al. 2007; Shankar et al. 1995; Wang et al. 1998; Zegeer et al. 1987):

- Rural two-lane two-way roads:
  - Segments.
  - Unsignalized intersections.
  - Signalized intersections.
- Rural multilane highways:
  - Segments.
  - Unsignalized intersections.
  - Signalized intersections.
- Urban and suburban arterials:
  - Segments.
  - Unsignalized intersections.
  - Signalized intersections.
- Rural and urban freeways:
  - Segments.
  - Ramp/collector–distributor segments.
  - Crossroad ramp terminals.

Basic facility types can be further subdivided by number of lanes and median type for all but rural two-lane two-way roads and by number of approach legs for intersections. Often, literature notes that contributing factors are applicable to specific facility types, such as rural and urban unsignalized intersections, but does not differentiate for which specific facility types each contributing factor is appropriate. If researchers consider the contributing factors of specific facility types but report them in a more generalized manner, then this creates potential for users to consider inappropriate contributing factors or apply incorrect countermeasures.

Several resources do not discuss when the authors use cross-sectional data for segments and intersections across different facility types (Lee and Nam 2003; Liu and Subramanian 2009; Liu and Ye 2011; Orner and Drakopoulos 2007; Poch and Mannering 1996; Preston et al. 2013a). These results were generalized for all facility types, excluding ramps and freeways unless specifically identified as being included. In these cases, contributing factors were reported for all facility types included in the data, with facility type–level indicators included in the model (i.e., number of lanes, area type, median presence and width, intersection approaches, and traffic control type). When larger amounts of data are available, identifying contributing factors at this level can be effectively accomplished by considering interactions between contributing factors and facility type–level information (e.g., an interaction between shoulder width and number of lanes will identify facilities where shoulder width has the greatest impact on safety by number of lanes).

Two resources identified facility types as paved or unpaved roads (Calvert et al. 1999; Knapp et al. 2014). These resources identified contributing factors, with specific factors identified for unpaved roadways versus paved roadways. In all cases, the roadways were rural two-lane two-way roads. Researchers can consider unpaved roadways as an additional facility type or a subset of rural two-lane roads.

Alternatively, two resources disaggregated the appropriate facilities at levels further than the HSM (Cato et al. 2013; MnDOT 2011). These authors considered contributing factors for horizontal curves and tangent segments separately and identified contributing factors in relation to bridge location. At this level, contributing factors can be considered more fully; however, crash data become sparser, resulting in more zero-crash sites in the database.

Milton and Mannering (1996, 1998) considered facility types by functional classification in cross-sectional regression models. In their research, roadway-segment models considered contributing factors for principal arterials, minor arterials, and all collectors. Other research by Shankar et al. (1995) considered only interstate segments, which are also a level of functional classification. Shankar et al.'s definition of facility type only considered segments and excluded intersections and interchanges. Since this research used cross-sectional models, HSM-defining characteristics were included as predictors in the models (e.g., number of lanes). However, this definition could become more cumbersome when considering intersections. In many cases, the intersecting roadways would consist of roadways with differing functional classifications, requiring contributing factors to be considered for several more levels of intersection types.

While differing facility-type definitions were identified in the literature, it is clear that the HSM definitions are most commonly used. As agencies are already aligning their data by these definitions for network screening, the HSM definitions are the most practical for use in this current project, possibly with an additional level of disaggregation of segments into horizontal curves and tangent segments, as outlined in Cato et al. (2013) and MnDOT (2011). The HSM definitions provide the most practical level of disaggregation for identifying appropriate contributing factors.

Contributing Roadway Factors

This section summarizes contributing factors and safety countermeasures identified in the literature. Contributing factors are characteristics of roadway segments and intersections that are associated with an increase or decrease in crash frequency or severity. Safety countermeasures or treatments are applied by agencies to roadway segments or intersections with the specific intention of reducing crash frequency or severity.

Table 4 summarizes roadway factors and table 5 summarizes safety countermeasures that published studies have identified as influencing crash frequencies and severities on the noted facility types for segments. With the intent of providing a high-level summary of possible influential factors, the tables do not distinguish findings by crash type or direction of safety effects (i.e., increases or decreases in crashes). The "Prevalence" column in each table indicates the prevalence of the roadway factors or countermeasures in the literature as being low, medium, or high. Based on the resources examined, these categories are defined as follows:

- Low = appeared in one to three resources.
- Medium = appeared in four to six resources.
- High = appeared in seven or more resources.

Roadway factors or countermeasures identified as having a higher prevalence are those that are known to be related to crash frequency or severity. These factors should be prioritized for inclusion in crash-frequency or -severity models. For instance, annual average daily traffic (AADT) was found to be the most commonly considered roadway factor across all resources, leading AADT to be labeled as high prevalence. Other factors were only found in one resource (e.g., sign-support density) due to the lack of readily available data and difficulty collecting data for a large number of sites, leading these factors to be labeled as low prevalence.

Table 4. Prevalence of roadway factors by facility type for segments.

Factor	Prevalence	Rural 2-Lane	Rural Multilane	Urban Street	Freeway	Ramp
AADT	High	●	●	●	●	●
Peak hour	Low	●	●	●	●	—
Area type	High	—	—	—	●	●
Number of lanes	High	—	●	●	●	●
Lane width	High	●	●	●	●	●
Shoulder width	High	●	●	●	●	●
Shoulder type	High	●	●	●	●	—
Shoulder location (e.g., inside)	Medium	—	—	—	●	●
Presence of surfaced shoulder	Medium	●	—	—	—	—
Presence of combination-surface/stabilized shoulder	Low	●	—	—	—	—
Horizontal curve density	High	●	●	●	●	●
Driveway type	Low	●	●	●	●	—

Factor	Prevalence	Rural 2-Lane	Rural Multilane	Urban Street	Freeway	Ramp
Driveway density	High	●	●	●	●	—
Presence of horizontal curve	High	●	●	●	●	●
Degree of curvature/radius	High	●	●	●	●	●
Length of horizontal curve	High	●	●	●	●	●
Horizontal curve-approach tangent length	Low	●	●	●	—	—
Superelevation rate	Medium	●	●	●	●	—
Presence of spiral transition	Low	●	—	—	—	—
Posted speed limit	High	●	●	●	●	—
Lateral-clearance distance	High	●		●	●	—
Side-slope rating	High	●	●	●		—
Pavement-surface condition (friction)	High	●	●	●	●	—
Combination horizontal and vertical alignment (visual trap)	Low	●	●	●	●	—
Pavement-surface condition (weather)	Low	●	●	●	●	—
Presence of intersection	Low	●	●	●		—
Proportion of commercial vehicles in the traffic stream	Medium	●	●	●	●	—
Topography	Low	●	●	●	●	—
Roadway-edge quality	Medium	●	●	●	●	—
Speed differential between horizontal curve and tangent	Low	●	●	●	●	—
Roadway gradient	High	●	●	●	●	—
Median width	High	—	●	●	●	—
Median type	High	—	●	●	●	—
Hazard rating of roadsides	Medium	●	●	●	●	—
Adjacent land use	Low	●	●	●	●	—
Location and presence of bus stops	Low	●	●	●	●	—
Presence of on-street parking	Low	—	—	●	—	—
Type of on-street parking	Low	—	—	●	—	—
Roadside fixed-object density	Low	—	—	●	—	—
Distance to shoulder barrier	Low	●	●	●	●	●
Speed-change lane presence	Low	—	—	—	●	●
Speed-change lane type	Low	—	—	—	●	—
Distance to inside barrier	Low	—	—	—	●	●
Length of ramp entrance	Low	—	—	—	●	—
Length of ramp exit	Low	—	—	—	●	—
Hours exceeding a threshold (1,000 hr for freeways; 2,500 hr for others)	Low	●	●	●	●	—
Ramp side (e.g., right side)	Low	—	—	—	●	—
Presence of type B weaving section	Low	—	—	—	●	●
Length of type B weaving section	Low	—	—	—	●	●
Ramp type	Low	—	—	—	—	●
Mean travel speed	Medium	●	●	●	●	—
Bridge width	Low	●	●	●	●	—
Presence of vertical curve	Medium	●	●	●	●	●
Vertical-curve rate	Low	●	●	●	●	●

Factor	Prevalence	Rural 2-Lane	Rural Multilane	Urban Street	Freeway	Ramp
Length of vertical curve	Low	●	●	●	—	—
Median-barrier type/condition	Low	—	●	●	●	●
Vertical-curve density	Low	●	●	●	—	—
Available sight distance	Medium	●	●	●	●	●
Edgeline-marking width	Low	●	●	●	●	—
Ditch design	Low	●	—	—	—	—
Outside-barrier type/condition	Medium	●	●	●	●	●
Intersection density	Medium	●	●	—	—	—
Functional classification	Low	●	●	—	—	—
Pavement width	Low	●	●	—	—	—
Season	Low	●	●	●	—	—
Year	Low	●	●	●	●	—
Catch-basin density	Low	●	●	●	—	—
Culvert density	Low	●	●	●	—	—
Presence of fence/wall	Low	●	●	●	—	—
Sign-support density	Low	●	●	●	—	—
Utility-pole density	Low	●	●	●	—	—
Pavement condition	High	●	●	●	●	●
Pavement type	Low	●	—	—	—	—
Monthly precipitation (i.e., rain/snow)	Low	—	—	—	●	—
Letter height of roadway signs	Low	●	●	●	—	—
Presence of variable message signs	Low	—	—	—	●	—

●Factor that influenced crash frequency and severity on the segment (as identified in a published study).
—No data available.

Table 5. Prevalence of countermeasures by facility type for segments.

Factor	Prevalence	Rural 2-Lane	Rural Multilane	Urban Street	Freeway	Ramp
Centerline markings	Low	●	●	●	●	—
Edge-line markings	Medium	●	●	●	●	—
Advisory signs	Medium	●	●	●	●	—
Chevrons	Medium	●	●	●	●	—
Post-mounted delineators	High	●	●	●	●	—
Flashing beacon	Low	●	—	—	—	—
Reflective-barrier delineation	Low	●	—	—	—	—
Profile thermoplastic markings	Low	●	—	—	—	—
Dynamic-curve warning system	Low	●	—	●	—	—
Speed-limit-advisory marking lane	Low	●	—		—	—
Lighting	High	●	●	●	●	—
SafetyEdge℠	Low	●	—	—	—	—
Curve-advance marking	Low	●	—	—	—	—
Optical speed bars	Low	●	—	—	—	—
Centerline rumble strips	High	●	●	●	●	—
Shoulder rumble strips	High	●	●	●	●	—
Raised pavement markings	Low	●	●	●	●	—
Outside barrier	High	●	—	—	●	—
Presence of median	High	—	●	●	●	—
Automated speed enforcement	Medium	●	●	●	●	—

Factor	Prevalence	Rural 2-Lane	Rural Multilane	Urban Street	Freeway	Ramp
Passing lane	Medium	●	—	—	—	—
Short four-lane section	Low	●	—	—	—	—
Two-way left-turn lane	Low	●	—	●	—	—
Median barrier	Low	—	—	—	●	—
Lane drop/add	Low	—	—	—	—	●
Raised median at crosswalk	Low	—	—	●	—	—
Presence of sidewalk	Low	●	●	—	—	—
Side friction	Low	●	●	—	—	—
Crosswalk	Low	●	●	—	—	—
Bicycle lane	Low	●	●	—	—	—
Impact attenuators	Low	●	●	●	●	—
Presence of bridge	Low	●	●	●	●	—
Truck escape ramp	Low	●	●	●	●	—
Access control	Low	—	●	—	—	—
Object delineation in clear zone	Medium	●	●	●	—	—
Truck prohibition	Low	●	—	●	—	—
Anti-icing system	Low	●	—	●	—	—
Break-away devices	Low	●	—	●	—	—
Advance guide signs and street names	Low	●	●	●	—	—
Pedestrian-crossing warning	Low	●	●	●	—	—
Traffic calming	Low	●	●	●	—	—
Interactive truck-rollover signing	Low	●	●	●	●	●
Truck-related posted speed limit	Low	—	—	—	●	—
Variable speed limit	Low	●	●	●	●	—
Active speed warning	Low	●	●	●	●	—
Reduced-traction warning sign	Low	●	●	●	—	—

●Factor that influenced crash frequency and severity on the segment (as identified in a published study).
—No data available.

Table 6 summarizes roadway factors and table 7 summarizes countermeasures that published studies have identified as influencing crash frequencies and severities on the noted facility types at signalized and unsignalized intersections. With the intent of providing a high-level summary of possible influential factors, the tables do not distinguish findings by crash type or direction of safety effects. The prevalence of the roadway factors or countermeasures in the literature is used in the same way as the segment tables.

Table 6. Prevalence of roadway factors by facility type for intersections.

Factor	Prevalence	R2U	R2S	RMU	RMS	UU	US
Major AADT	High	●	●	●	●	●	●
Minor AADT	High	●	●	●	●	●	●
Left-turn volume	Low	—	—	—	—	●	●
Right-turn volume	Low	—	—	—	—	●	●
Total opposing approach volume	Low	—	—	—	—	●	●
AADT ratio	Low	●	●	—	—	—	—
Number of approaches	High	●	●	●	●	●	●
Number of approach lanes	Medium	—	—	●	●	●	●
Number of signal heads	Low	—	●	—	●	—	●
Signal head–mount location	Low	—	●	—	●	—	●
Signal timing/phases	Medium	—	●	—	●	—	●

Factor	Prevalence	R2U	R2S	RMU	RMS	UU	US
Signal-head size	Low	—	—	—	—	—	●
Yellow interval length	Low	—	●	—	●	—	●
Intersection sight distance	High	●	●	●	●	●	●
Pedestrian-crossing distance	Low	●	●	●	●	●	●
Pedestrian signal–head type	Low	—	●	—	●	—	●
Skew ANG	High	●	●	●	●	●	●
Proximity to horizontal and vertical curves	Low	●	●	●	●	●	●
Proximity to at-grade railroad crossing	Low	●	●	●	●	●	●
Traffic control type (e.g., all-way stop)	High	●	—	●	—	●	—
Presence of commercial development	Low	●	●	●	●	●	●
Maximum number of lanes crossed by pedestrian in maneuver	Low	—	—	—	—	●	●
Number of bus stops within 1,000 ft	Low	—	—	—	—	—	●
Number of schools within 1,000 ft	Low	—	—	—	—	—	●
Number of alcohol-sales establishments with 1,000 ft	Low	—	—	—	—	—	●
Left turn–lane length	Low	●	●	●	●	●	●
Right turn–lane length	Low	●	●	●	●	●	●
Shoulder width	Low	●	●	●	●	—	—
Pavement-surface condition (friction)	Medium	●	●	●	●	●	●
Median type	Low	—	—	●	●	●	●
Presence of parking near intersection	Low	●	●	●	●	●	●
Minor route access density	Low	●	●	●	●	—	—
Functional classification	Low	—	—	—	—	●	●
Approach posted speed limit	Low	—	—	—	—	●	●
Opposing approach posted speed limit	Low	—	—	—	—	●	●
Intersection approach gradient	Low	—	—	—	—	●	●
Roadway drainage	Low	—	●	—	●	—	●
Signal hardware located in clear zone	Low	—	●	—	●	—	●

●Factor that influenced crash frequency and severity on the segment (as identified in a published study).
—No data available.
R2U = rural two-lane unsignalized intersection; R2S = rural two-lane signalized intersection; RMU = rural multilane unsignalized intersection; RMS = rural multilane signalized intersection; UU = urban unsignalized intersection; US = urban signalized intersection.

Table 7. Prevalence of countermeasures by facility type for intersections.

Factor	Prevalence	R2U	R2S	RMU	RMS	UU	US
Advance intersection warning	Low	●	●	●	●	●	●
Back plates	Low	—	●	—	●	—	●
Retroreflective sheeting on back plate	Low	—	●	—	●	—	●
Left-turn lane	High	●	●	●	●	●	●
Right-turn lane	High	●	●	●	●	●	●
Presence of crosswalk	Low	●	●	●	●	●	●
Intersection lighting	High	●	●	●	●	●	●
Right turn on red prohibition	Medium	—	—	—	—	—	●
Red-light camera	Medium	—	—	—	—	—	●
Automated speed enforcement	Low	—	—	—	—	—	●
Protected left-turn phase	High	—	●	—	●	—	●
Protected/permissive left-turn phase	Low	—	●	—	●	—	●
Right-turn channelization	Low	—	—	●	●	●	●
Left-turn channelization	Low	—	—	●	●	●	●
Exclusive pedestrian phasing	Low	—	●	—	●	—	●

Factor	Prevalence	R2U	R2S	RMU	RMS	UU	US
Emergency pre-emption	Low	—	●	—	●	—	●
Lead-pedestrian interval	Low	—	●	—	●	—	●
Lead-bicycle interval	Low	—	●	—	●	—	●
Actuated signal	Low	—	●	—	●	—	●
Signal coordination	Low	—	●	—	●	—	●
Advanced dilemma-zone detection	Low	—	●	—	●	—	—
Flash mode	Low	—	●	—	●	—	●
Stop-sign size	Low	●	—	●	—	●	—
Pedestrian signing	Low	●	●	●	●	●	●
Overhead lane-use signs	Low	●	●	●	●	●	●
Raised pavement marker	Low	●	●	●	●	●	●
Stop bars	Low	●	—	●	—	●	—
Directional median opening	Low	—	●	—	●	—	●
Double left-turn lane	Low	—	●	—	●	—	●
Left turn–painted separation	Low	—	●	—	●	—	●
Presence of median	Low	—	—	●	●	●	●
Raised median crosswalk	Low	●	—	●	—	●	—
Presence of raised island	Low	●	—	●	—	●	—
Splitter island on minor approach	Low	●	—	●	—	●	—
Bypass lane	Low	●	●	●	●	●	●
Double stop sign	Low	●	—	●	—	●	—
Flashing beacons	Low	●	—	●	—	●	—
Stop-ahead pavement markings	Low	●	—	●	—	●	—
Transverse rumble strips	Low	●	●	●	●	●	●
Bicycle lane	Low	●	●	●	●	●	●
Contraflow bicycle lane	Low	●	●	●	●	●	●
No-left-turn sign	Medium	●	●	●	●	●	●
No-U-turn sign	Low	●	●	●	●	●	●
Pedestrian overpass/underpass	Low	—	—	—	—	●	●
Presence of sidewalk	Medium	●	●	●	●	●	●
Offset left-turn lanes	Low	—	●	—	●	●	●
Extended edgelines	Low	●	—	●	—	●	—
Driveway closure	Low	●	—	●	—	●	—
Acceleration lane	Low	●	—	●	—	●	—
Offset T versus four-leg intersection	Low	●	—	●	—	●	—
ICWS	Low	●	—	●	—	●	—
Roadside markers	Low	●	—	●	—	●	—
Clearance interval	Medium	—	●	—	●	—	●
Pedestrian signals	Low	●	●	●	●	●	●
Pedestrian crossing warning	Low	●	●	●	●	●	●
Restricted turning maneuvers	Low	—	●	—	●	—	●
Access management near intersection	Low	—	●	—	●	—	●
Bicycle-detecting sensor	Low	—	●	—	●	—	●

●Factor that influenced crash frequency and severity on the segment (as identified in a published study).
—No data available.
R2U = rural two-lane unsignalized intersection; R2S = rural two-lane signalized intersection; RMU = rural multilane unsignalized intersection; RMS = rural multilane signalized intersection; UU = urban unsignalized intersection; US = urban signalized intersection; ICWS = intersection curve warning sign.

Datasets

Nearly all resources identified in this literature review used databases that were developed specifically for their research, including field data collections or State crash, traffic, and roadway inventory data. The databases were developed to suit the needs of the research at hand, and a limited set of predictors was considered in the developed models. As such, these databases were not usable for this study. However, as part of the DCMF program, States supplied information on their data sources and availability via a DCMF Needs Assessment. In addition to these data, the research team identified the following databases for further consideration for identifying focus crash types, facility types, and contributing roadway factors:

- Crash Injury Research and Engineering Network (NHTSA 2018d).
- Crashworthiness Data System (NHTSA 2018c).
- Fatality Analysis Reporting System (FARS) (NHTSA 2018a).
- Federal Transit Administration National Transit Database (FTA 2018).
- General Estimates System (NHTSA 2018b).
- Highway Safety Information System (HSIS) (FHWA 2018c).
- Motor Carriers Management Information System (FMCSA 2016).
- National EMS Information System (NEMSIS Technical Assistance Center 2018).
- National Motor Vehicle Crash Causation Study (NHTSA 2008).
- National Park Service Service-Wide Traffic Accident Reporting System.[1]
- Second Strategic Highway Research Program (SHRP2) Naturalistic Driving and Roadway Databases (Virginia Tech Transportation Institute 2018).

Chapter 4 and chapter 5 describe the ultimate selection of data sources for this project.

Methodologies

The previous sections of this chapter summarized roadway factors and countermeasures expected to be associated with crash frequency and severity for various facility types. The previous sections also identified commonly used data resources upon which these findings were based. This section summarizes relevant statistical methods that could be employed to the aforementioned data resources (or others) to identify contributing roadway factors for FCFTs. That is, this section provides an overview of statistical methods that can be used to estimate and quantify the uncertainty about factors that either increase or decrease the frequency and/or severity of various crash types.

In determining which statistical method to use to identify contributing roadway factors for FCFTs, it is imperative that the selected method be appropriate for the type of data under consideration. For example, researchers may use generalized linear modeling to identify roadway factors for data on the number of crashes that occur in a certain section of a highway (e.g., count data), but this method may not be applicable when trying to analyze the time between crashes on the same section of the highway. The chosen statistical method must match the selected data to identify the contributing roadway factor.

[1]This unpublished dataset is available through the National Park Service.

The following sections detail the selected methods many researchers have used to identify factors that increase or decrease the potential for future crashes of various crash types. Each section includes a high-level overview of each method, along with the type of data to which the method would apply, and a list of the most popular textbooks for using the proposed method.

Normal Linear Regression Models

Normal linear regression models (commonly referred to as "linear models") seek to quantify the linear relationship between or among an outcome (often called a response, dependent variable, or output) and one or more other measurements (often called predictors, independent variables, or inputs). In other words, the goal of a normal regression model is to estimate the change in an output that results from changing one or more of the input variables. By way of example, a normal regression model would estimate the change in friction on a road surface due to a change in surface temperature. Relevant results from a normal regression analysis include but are not limited to a measure, which represents the probability of a Type I error (e.g., rejecting a null hypothesis when it is actually true, also known as a false positive), of whether each input variable has a statistically significant effect on the output variable in the form of *p*-values and a summary of the estimated effect of each input variable on the output variable.

Certain assumptions that are used to estimate the effect of each input variable on the output variable restrict the applicability of linear models. As the name implies, linear models are only able to estimate linear relationships between inputs (predictors) and the output (the response) and any transformations of the input and output variables. Admittedly, many relationships are more complicated (i.e., nonlinear), and linear models could fail to recognize such relationships.

Normal linear models are only applicable to continuous outputs. A continuous variable is a variable that can take any value (e.g., rational and irrational numbers) on the real number line. Linear regression would not apply, for example, to count data because a count has to be an integer (e.g., 0, 1, 2…).

Linear regression is a well-established method and taught as a core course in many statistics department. The most popular introductory texts to linear regression include the following:

- *Applied Linear Regression* by S. Weisberg (2014).
- *Regression Analysis by Example* by S. Chatterjee and A. S. Hadi (2012).
- *Linear Models with R* by J. J. Faraway (2015).
- *Introduction to Linear Regression Analysis* by D. C. Montgomery, E. A. Peck, and G. G. Vining (2012).

The Faraway (2015) and Montgomery et al. (2012) books are particularly popular due to their use of R programming language and statistical analysis software (SAS).

Generalized Linear Models

As previously mentioned in the discussion about normal linear regression models, one of the drawbacks of such models is their reliance on continuous outputs. For many applications, particularly for traffic-safety data, the outcome of interest is not continuous. Generalized linear models (GLMs) are an extension of linear models by use of noncontinuous data. GLMs are similar to normal linear models in that GLM results include a measure of whether each input variable has a significant effect on the output variable in the form of p-values and a summary of the estimated effect of each input variable on the output variable. However, GLMs explicitly take into account the noncontinuous nature of the outcome variable to estimate these effects.

A few specific GLMs that are of particular applicability for traffic-safety data are Poisson, negative binomial (NB) regression, and logistic models. Poisson and NB regression models are specifically used to estimate effects of a particular input variable on a count output variable. For example, Poisson or NB regression models can be used to estimate effects of including a high occupancy–vehicle lane on the number of crashes along a particular stretch of highway. Estimated effects from Poisson or NB regression models are interpreted as the effect of a change in the input variable on the expected number of outcomes.

Binary logistic regression models are used to estimate the effect of input variables on a 0/1 outcome variable. A 0/1 outcome variable is any variable that either occurred (typically denoted as "1") or did not occur (typically denoted as "0"). For example, a binary logistic regression model could be used to estimate the effect of a 5-ft-wide shoulder on the occurrence of one or more ROR crashes. Estimated effects from a binary logistic regression model are interpreted as the effect of an input variable on the probability of the outcome variable occurring (e.g., the probability of one or more ROR crashes). Multinomial logistic regression models extend this concept to situations where the response variable has more than two outcomes. Similar to normal linear regression models, GLMs assume that the relationship between each input variable and the output variable is linear. While the assumption of linearity is often a good approximation, nonlinear relationships are often more realistic (but also more difficult to model and interpret).

GLMs inherently respect the noncontinuous nature of an outcome variable, but interpreting the estimated effects from a GLM can be more complicated than estimating those from a normal linear regression model. For example, in a logistic regression model, the effects displayed by SAS are typically on a log–odds ratio scale; that is, the displayed effects can be interpreted as the effect of an input variable on the log–odds ratio of an outcome variable. Therefore, special care must be given when considering the estimated effects in a GLM framework. If necessary, the scale of the estimated effects can be transformed back to the original scale of the data, but this back-transformation is typically not the default for most SAS packages.

Approachable, high-level textbooks for GLMs include the following:

- *Extending the Linear Model with R* by J. J. Faraway (2006).
- *An Introduction to Generalized Linear Models* by A. J. Dobson and A. G. Barnett (2008).
- Chapter 12 of *Applied Linear Regression* by S. Weisberg (2014).
- Chapter 4 of *Introduction to Statistical Learning and Data Mining* by G. James, D. Witten, T. Hastie, and R. Tibshirani (2013).

Chapter 4 of *The Elements of Statistical Learning and Data Mining* by T. Hastie, R. Tibshirani, and J. Friedman (2009) and *Generalized Linear Models* by P. McCullagh and J. A. Nelder (1994) are more theoretically based texts.

Longitudinal Analyses

Longitudinal data are repeated measurements collected on a single entity over time. For example, data on the number of crashes along a given road over time would constitute longitudinal data for that specific road. The goal of a longitudinal analysis is to relate input variables to changes in the output over time. For example, a longitudinal analysis could infer the effect of increasing the width of a rural road lane from 9 to 12 ft on the number of head-on (HEO) crashes along the road.

The primary advantage of a longitudinal analysis is that it allows unit-specific inference. For instance, increasing the lane width from 9 to 12 ft might have no effect on the number of HEO crashes for some rural roads but might significantly decrease the number of HEO crashes for others. Hence, measuring the same unit repeatedly over time allows for inference specific to the unit.

Standard textbooks for longitudinal analysis include the following:

- *Applied Longitudinal Analysis* by G. M. Fitzmaurice, N. M. Laird, and J. H. Ware (2011). This text includes examples of how to use SAS for longitudinal analyses.
- *Longitudinal Data Analysis* by D. Hedeker and R. D. Gibbons (2006).

Additionally, the following books have discussion on the subject:

- *Extending the Linear Model with R* by J. J. Faraway (2006).
- An Introduction to Generalized Linear Models by A. J. Dobson and A. G. Barnett (2008).

Survival Analysis

Survival analysis, also referred to as hazard analysis, is a branch of statistics primarily concerned with time-to-event (TTE) data. TTE data are collected on the amount of time passed until a certain event occurs (e.g., the time between consecutive crashes on the same section of highway). The focus of many survival analyses is the relation of input variables to the probability of an event occurring given that the event has not yet occurred. For example, a survival analysis would seek to relate the presence of a median to the probability of a crash occurring between 5 and 6 p.m. given that a crash has not previously occurred that day.

Reasons why TTE data are not amenable to traditional statistical analysis (e.g., normal or GLMs) include the following:

- TTE data are strictly positive.
- TTE data are typically described by nonsymmetric and highly skewed distributions.
- TTE data can often be censored.

Censored data occur when an event is not observed in the timeframe of the study and, hence, the TTE is unknown. Much of survival analysis is concerned with appropriately accounting for such data characteristics in a statistical analysis.

Common textbooks for introductory courses in survival analysis include the following:

- *Modelling Survival Data in Medical Research* by D. Collett (2003).
- Survival Analysis Using SAS: A Practice Guide by P. D. Allison (2010).

The text by P. D. Allison (2010) is more applied in nature, but both texts detail how to use SAS to perform analyses. Another approachable text is *Survival Analysis: A Self-Learning Text* by D. G. Kleinbaum and M. Klein (2012). Additionally, other books (e.g., *An Introduction to Generalized Linear Models* by A. J. Dobson and A. G. Barnett [2008]) have individual chapters dedicated to survival analysis.

Spatial Modeling

Spatial modeling, or spatial statistics, is a branch of statistics that analyzes data collected regarding a specific space and produces information about the spatial dimensions of the data (known as spatial data). Some examples of spatial data include the following:

- Weather measurements taken at different stations across the United States.
- Counts of flu cases observed in each county within a State.
- The number of crashes observed on adjacent sections of a highway.

Researchers collect each of these example datasets at various spatial locations (e.g., a weather station or section of highway) and use spatial methods for analysis.

The primary concern of spatial statistics is accounting for and exploiting correlations with data collections at nearby locations. For example, the weather in Raleigh, NC, is expected to be similar to (or correlated with) weather in Durham, NC, due to their proximity of approximately 25 mi. Similarly, the effect of a random, unmeasured event that occurs in Raleigh could “spillover” to nearby Durham. Accounting for and exploiting the spatial correlation in such data leads to more appropriate and correct statistical inferences and the ability to predict at unobserved locations.

Common tools to analyze spatial data include Gaussian process regression, autoregressive models, and point processes. Because these tools are most commonly used within the previously discussed methods (e.g., as part of a normal regression model or GLM), they are considered an extension of these methods when used with data collected over space.

Approachable textbooks in spatial statistics include the following:

- *Applied Spatial Statistics for Public Health Data* by L. A. Waller and C. A. Gotway (2004).
- Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology by A. B. Lawson (2013).

More theoretical, low-level texts include the following:

- *Statistics for Spatial Data* by N. A. C. Cressie (1993).
- *Statistics for Spatio-temporal Data* by N. A. C. Cressie and C. K. Wikle (2011).
- *Hierarchical Modeling and Analysis of Spatial Data* by S. Banerjee, B. P. Carlin, and A. E. Gelfand (2015).

Bayesian Modeling

Bayesian modeling is not a single method, per se, but more of a general philosophy toward any statistical analysis. Unlike traditional statistics, Bayesian analysis uses laws of probability to quantify the uncertainty associated with statistical inferences. That is, the Bayesian framework (or philosophy) performs statistical estimation using laws of probability to combine a priori information (prior distribution) with data (data distribution) into a composite summary of unknown parameters (posterior distribution).

Since the mid-1990s, researchers have widely adopted Bayesian philosophy to perform complex analyses of datasets. Advantages of Bayesian philosophy include the following:

- A unified, single approach to quantifying uncertainty regarding statistical conclusions.
- Grounding inference only in observed data (rather than relying on complex asymptotic mathematics).
- The ability to combine prior knowledge into the analysis.

Despite its popularity and advantages, the Bayesian paradigm is more complex to implement. Employing Bayesian estimation requires advanced computation skills and often requires writing problem-specific software (code). The analyst must also assign appropriate prior distributions to parameters they are estimating and be able to interpret the posterior distributions that are a result of the modeling. A single set of results may take hours of computing time to obtain. For these reasons, researchers who lack significant experience with or training in advanced statistical analyses often do not adopt the Bayesian paradigm for statistical inference.

Standard texts for Bayesian statistics include the following:

- *Bayesian Data Analysis* by A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013).
- *Bayesian Methods for Data Analysis* by B. P. Carlin and T. A. Louis (2008).
- *A First Course in Bayesian Statistical Methods* by P. D. Hoff (2009).
- *Doing Bayesian Data Analysis* by J. K. Kruschke (2015).

The texts by Hoff (2009) and Kruschke (2015) are perhaps the most approachable but still require a significant understanding of mathematical statistics.

Other Methods

While the previously discussed methods cover a breadth of statistical methodology that can be used to identify contributing factors from various types of data, the list is by no means comprehensive. Other methods worth mentioning are data-mining techniques, such as treed regression or nonlinear regression, and multivariate statistical methods, such as clustering, multivariate regression, and factor analysis. *The Elements of Statistical Learning and Data Mining* by T. Hastie, R. Tibshirani, and J. Friedman (2009) and *Introduction to Statistical Learning and Data Mining* by G. James, D. Witten, T. Hastie, and R. Tibshirani (2013) discuss data-mining techniques, while *Methods of Multivariate Analysis* by A. C. Rencher and W. F. Christensen (2012) discusses multivariate methods.

APPROACHES TO SYSTEMIC SAFETY ANALYSIS

This section of the literature review summarizes additional resources that address crash types, facility types, and contributing roadway factors, primarily within the context of systemic approaches to road-safety management. Table 8 provides brief summaries of the resources the research team identified and reviewed. The table also provides a more detailed summary of a selected number of these resources.

Table 8. Summary of selected literature on systemic safety analysis.

Publication	Summary of Publication
NCHRP Synthesis 128: "Methods for Identifying Hazardous Highway Elements" (Zegeer 1986)	Synthesizes methods used by 39 State agencies and 17 local agencies to identify and treat hazardous highway elements and determine which methods have been most successful.
"Variable Safety Improvements for Unpaved Roads" (Caldwell and Wilson 1996)	Provides a list of 12 roadway-design elements that should be considered in unpaved-road safety audits based on a survey of a steering committee.
NCHRP 500 Series (National Academy of Sciences 2003a, 2003b, 2003c, 2003d, 2003e, 2003f, 2004a, 2004b, 2004c, 2004d, 2004e, 2004f, 2004g, 2005a, 2005b, 2005c, 2006, 2007, 2008a, 2008b, 2008c, 2008d, 2009)	Presents 22 emphasis areas that affect overall highway safety, as well as strategies for reducing crashes corresponding to these emphasis areas and an outline of what is needed to implement each strategy.
HSM (AASHTO 2010)	Includes a framework for identifying issues, developing countermeasures, and evaluating treatment effectiveness. Provides safety performance functions and CMFs for various facility/treatment types. Details quantitative approaches that can help to understand relative safety impacts of various elements.
NCHRP Research Results Digest 345: Alternate Strategies for Safety Improvement Investments (National Academy of Sciences 2010)	Synthesizes practices and compares two approaches used by States to allocate safety funds: "black spot" analysis and "systematic" methods.
"Mn/DOT County Road Safety Plans" (Preston and Gute 2010)	Provides safety plans for 20 counties based on analysis of State-specific crash data. Identifies contributing factors that increase crash potential and focus crash types for various facility types, including roadway segments, horizontal curves, and rural stop-controlled intersections.
Minnesota Department of Transportation Traffic Safety Analysis Software State of the Art (Brown et al. 2011)	Identifies and assesses safety-analysis practices and existing supporting software tools currently used by State transportation agencies, including systemic approaches to safety analysis.
Quantitative Assessment of Local Rural Road Safety – Case Study (Mahgoub et al. 2011)	Proposes a rural road safety index for local rural roads to rank the road network according to safety features and identify deficiencies in road sections.

Publication	Summary of Publication
Benefit/Cost Evaluation of MODOT's Total Striping and Delineation Program (Potts et al. 2011)	Documents the methodology and results of a before–after safety evaluation and B/C analysis of a major Missouri Department of Transportation program, known as the Smooth Roads Initiative, to improve both rideability and visibility of over 2,300 mi of major roadways in Missouri.
Centerline Rumble Strips on Secondary Highways – A Systemic Crash Analysis (Wilder 2011)	Summarizes a systemic approach to evaluate potential locations for and recommend the best way to implement centerline rumble strips in New York State on nonfreeway highway segments.
"A Systemic Safety Project Identification Process – Minnesota's County Road Safety Plans" (Preston 2012)	Identifies contributing factors that increase crash potential on rural paved roadway segments and rural unsignalized intersections.
Development of a Systemic Road Safety Analysis Tool – Roadway Departure Crashes at Bridges in Salem County, New Jersey (Cato et al. 2013)	Presents a systemic road-safety analysis tool to examine roadway-departure crashes at bridges. Identifies numerous characteristics related to a high potential of roadway-departure crashes at bridges.
Comparison of Countermeasure Selection Methods for Use in Road Safety Management (Harwood et al. 2013)	Compares three methods for selecting highway-infrastructure countermeasures to reduce crash frequency and severity: the usRAP Tools software (usRAP 2020), the FHWA Systemic Safety Project Selection tool, and road-safety audits (Preston et al. 2013a).
Systemic Safety Project Selection Tool (Preston et al. 2013a)	Describes a systemic safety-planning process, which includes identifying focus crash types and contributing factors, screening and prioritizing candidate locations, selecting countermeasures, and prioritizing projects.
Using Risk to Drive Safety Investments (Preston et al. 2013b)	Characterizes the systemic approach to safety management, identifies context (e.g., crashes, road types) suited for systemic safety approaches, and summarizes the cyclical three-element process in the *Systemic Safety Project Selection Tool* (Preston et al. 2013a).
Local and Rural Road Safety Briefing Sheets: Applying the Systemic Safety Approach on Local Roads (FHWA 2014)	Identifies contributing factors that increase crash potential for roadway departures along horizontal curves in Thurston County, WA, for screening and prioritizing candidate locations for systemic improvements.
Systemic Safety Improvement Risk Factor Evaluation and Countermeasure Summary: Final Report (Knapp et al. 2014)	Investigates and compares two systemic safety tools/methodologies, an approach used to produce Minnesota county road-safety plans and usRAP, when applied to two counties in Iowa (usRAP 2020).
A Systematic Approach to Identifying Traffic Safety Needs and Intervention Programs for Indiana: Volume I—Research Report (Tarko et al. 2014)	Presents a method and example application for systemically identifying thresholds of specific characteristics that note horizontal curves with a high potential for roadway-departure crashes.
Risk Factors Associated with High Potential for Serious Crashes (Al-Kaisy et al. 2015)	Presents an approach that uses extensive data collection and analysis for a large sample of Oregon's low-volume roads to develop an index that expresses the crash potential for different road geometries and roadside features, as well as crash history and traffic exposure.
Reliability of Safety Management Methods: Systemic Safety Programs (Gross et al. 2016)	Describes a process for program managers, project managers, and data analysts to develop and enact a comprehensive safety-management program based on contributing factors for FCFTs.
A Systemic Safety Analysis of Pedestrian Crashes: Lessons Learned (Wang et al. 2016)	Compares candidate site-ranking methods that use contributing factors for systemic approaches to analyzing pedestrian crashes at intersections in Austin, TX. Estimates contributing factors that increase crash potential for pedestrian crashes at signalized and stop-controlled intersections and compares the effects of the ranking and weighting methods.

NCHRP = National Cooperative Highway Research Program; usRAP = United States Road Assessment Program.

The systemic approach to safety involves identifying sites based on site-specific geometric and operational attributes rather than observed crashes (Gross et al. 2016). FHWA developed a *Systemic Safety Project Selection Tool*, which identifies a systemic safety-planning process that includes the following steps (Preston et al. 2013a):

1. Identify focus crash types and contributing factors.
2. Screen and prioritize candidate locations.
3. Select countermeasures.
4. Prioritize projects.

The first step relates most closely to this research, and its objective is to identify contributing factors commonly associated with each focus crash type experienced across the road system (Preston et al 2013a). These are the crash types that have the greatest potential for reducing fatal and severe injuries. The FHWA *Systemic Safety Project Selection Tool* notes that the State's Strategic Highway Safety Plan (SHSP) is a good starting point for identifying such contributing factors, particularly the State's documented emphasis areas, which are typically identified through a data-driven approach, similar to those recommended by the FHWA tool. An example data-driven approach for identifying focus crash types is to disaggregate jurisdiction-wide crash data by SHSP emphasis area. The focus crash types may be selected as the most frequently occurring statewide, the most frequently occurring crash type by jurisdiction or area type (e.g., most frequent crash type for rural areas, counties, and/or urban areas), or the most overrepresented when comparing proportions of crashes in a specific jurisdiction or area type to the State numbers.

The second step is to identify where crashes are occurring (i.e., the focus facilities). The FHWA *Systemic Safety Project Selection Tool* recommends using a crash tree diagram for this purpose. For a focus crash type, the crash tree begins with total KA crashes and uses available data to subdivide each subsequent level (e.g., area type, ownership, segments/intersections, surface type, number of lanes, posted speed limit). Each branch of the tree contains nodes that identify the number of KA crashes for those conditions. The nodes with the highest counts yield the focus facility types (i.e., facility type where the focus crash type most frequently occurs).

Once analysts identify FCFTs, they can identify and evaluate contributing factors. This can be done as an extension of identifying focus facility type (using branches and nodes on crash tree diagrams), further focusing on curvature, posted speed limit, traffic volume as data are available, or by listing known contributing roadway factors for which data are available. The FHWA *Systemic Safety Project Selection Tool* provides the following list of example factors whose presence, absence, and/or characteristics may influence crash potential (Preston et al. 2013a, p. 18):

- Number of lanes.
- Lane width.
- Shoulder surface width and type.
- Median width and type.
- Horizontal curvature, superelevation, delineation, or advance warning devices.
- Horizontal curve density.
- Horizontal curve and tangent speed differential.

- Presence of a visual trap at a curve or combinations of vertical grade and horizontal curvature.
- Roadway gradient.
- Pavement condition and friction.
- Roadside or edge hazard rating (potentially including sideslope design).
- Driveway presence, design, and density.
- Presence of shoulder or centerline rumble strips.
- Presence of lighting.
- Presence of on-street parking.
- Intersection skew angle.
- Intersection traffic control device.
- Number of signal heads versus number of lanes.
- Presence of backplates.
- Presence of advanced warning signs.
- Intersection located in or near horizontal curve.
- Presence of left-turn or right-turn lanes.
- Left-turn phasing.
- Allowance of right-turn-on-red.
- Overhead versus pedestal-mounted signal heads.
- Pedestrian crosswalk presence, crossing distance, signal head type.
- Average daily traffic volumes.
- Average daily entering vehicles.
- Proportion of commercial vehicles in traffic stream.
- Posted speed limit or operating speed.
- Presence of nearby railroad crossing.
- Presence of automated enforcement.
- Adjacent land use type (e.g., schools, commercial, or alcohol-sales establishments).
- Location and presence of bus stops.

Further, the FHWA *Systemic Safety Project Selection Tool* suggests using the predictive method from the HSM as a resource for identifying potential contributing roadway factors. The HSM predictive method includes segment and intersection safety-performance functions and adjustment factors for facility types listed in the Facility Types section of the HSM. The HSM and HSM supplement contain an extensive list of adjustment factors that have been related to crash frequency (AASHTO 2014). However, the HSM predictive method generally focuses on all crashes or all crash severities for specific crash types (e.g., all multivehicle crashes). Adjustment factors are associated with total crashes and not necessarily severe injury crashes of a highly specific type.

Alternatively, analysts can use contributing roadway factors identified in other research or through analysis of their own data (including development of crash-prediction models). The following discussion focuses on research that has looked at the process for site selection, contributing roadway-factor identification, prioritization, and methods for implementing systemic safety approaches.

National Cooperative Highway Research Program's (NCHRP's) *Research Results Digest 345: Alternate Strategies for Safety Improvement Investments* reviewed the "black spot" and "systematic" methods for allocating safety resources (National Academy of Sciences 2010). The authors distributed a survey to all 50 State DOTs and received responses from 25 State agencies. Additionally, the researchers conducted in-depth case studies with four States (Iowa, Minnesota, Missouri, and North Carolina). The authors found that most agencies target safety funds at high-crash locations; however, there was a trend of increasing the proportion of safety funding toward systemwide improvements. The authors also noted that the characteristics associated with KA crashes have caused programs to be more focused on rural areas and include more projects involving widely deploying proactive, low-cost strategies across systems. The States identified two challenges associated with the safety planning process: the analytical process for identifying candidate locations for investment in rural areas is not well developed, and while States have increased engagement with local road authorities, concerns remain about their lack of safety-planning experience.

Gross et al. (2016) developed *Reliability of Safety Management Methods: Systemic Safety Programs* to help program managers, project managers, and data analysts develop comprehensive safety-management programs. The objectives of the guide were to raise awareness of the systemic approach to safety management, characterize typical projects identified and implemented through a comprehensive safety-management program, demonstrate the value of integrating systemic approaches as part of a comprehensive safety-management program, and provide information on allocating funding to systemic projects within a comprehensive safety-management program. The guide described the state of the practice and the latest tools to support systemic safety analyses. Additionally, the guide described the systemic approach as having the following steps:

1. Identify FCFTs.
2. Determine contributing factors.
3. Select countermeasures.
4. Screen network for suitable locations.
5. Evaluate safety effects.

The authors suggested that contributing factors should be identified using statistical modeling or cross-tabulations, identifying the association of specific roadway data with each crash type. If the agency does not have data or expertise to determine contributing factors for a focus crash type, the guide suggests using the NCHRP Report 500 series (National Academy of Sciences 2003a, 2003b, 2003c, 2003d, 2003e, 2003f, 2004a, 2004b, 2004c, 2004d, 2004e, 2004f, 2004g, 2005a, 2005b, 2005c, 2006, 2007, 2008a, 2008b, 2008c, 2008d, 2009) to identify contributing factors related to specific crash types. The authors suggested the following common segment and intersection features to define contributing factors:

- *Segment features:* number of lanes, lane width, shoulder type and width, median type and width, road edge features and quality, number and type of access points, radius and superelevation of horizontal curves, speed limit, speed differential between horizontal curves and tangent segments, roadside hazards, and pavement condition and friction.

- *Intersection features:* number of approaches, number of approach lanes, traffic control devices, skew, proximity to horizontal and vertical curves, signal timing, proximity to railroad crossings, presence of street lighting, proximity to nearby access points, and presence of commercial developments.

Some transportation agencies have been utilizing their State-specific datasets to conduct systemic analyses. For example, the Minnesota Department of Transportation (MnDOT) has been working with counties to address fatalities on local roads, where more than 50 percent of Minnesota traffic fatalities occur (Preston and Gute 2010). MnDOT recognized that local agencies have less experience in conducting systemwide crash analysis or linking crash causes with mitigation strategies at specific locations on their system. Therefore, MnDOT began working on an initiative to develop a roadway-safety plan for each county within the State, concentrating on information specific to individual counties and identifying opportunities to reduce KA crashes. MnDOT funds the planning process, which includes the following:

- Performing crash analyses and conducting a systemwide assessment of crash potential.
- Identifying unique, low-cost infrastructure-based safety projects that can be deployed across the county.
- Developing unique safety plans for each county.

The research team completed its review of safety plans for 20 counties with the focus on roadway-departure crashes on local roads and ANG crashes at rural stop-controlled intersections in 2010. The State DOT identified contributing factors for roadway segments, horizontal curves, and rural stop-controlled intersections, along with the following contributing factors relevant to crashes involving curves (Preston and Gute 2010):

- Average daily traffic (ADT) volume.
- Curve radius (most crashes occur on curves with radii ranging between 500 and 1,500 ft).
- History of KA crashes on curves.
- Presence of an intersection or visual trap on a curve.

Additionally, the State DOT identified the following contributing factors for rural paved segments (Preston 2012):

- Density of roadway-departure crashes.
- Traffic volume.
- Curve density.
- Access density.
- Pavement edges.

For rural unsignalized intersections, the State DOT identified the following contributing factors (Preston 2012):

- Skewed minor leg approach.
- Intersection on/near horizontal curve.
- Minor and major ADT ratios.
- Proximity to previous stop sign.

- Proximity to railroad crossing.
- Intersection-related crashes.
- Commercial development in quadrants.

Thurston County, WA, Public Works selected roadway departures along horizontal curves as a target crash type based on a review of severe crash data (FHWA 2014). After reviewing crash data linked with roadway characteristics, they selected the following contributing factors with high crash potential to screen and prioritize candidate locations for systemic improvements:

- Roadway class of major rural collector.
- Presence of an intersection.
- Traffic volume of 3,000 to 7,500 AADT.
- Edge clearance rating of 3 (on a scale of 1 [widest] to 3 [narrowest]).
- Paved shoulders ≥4 ft in width.
- Presence of a vertical curve.
- Consecutive horizontal curves (windy roads).
- Speed differential between posted approach speed and curve advisory speed of 0, 5, and 10 mph.
- Presence of a visual trap (a minor road on the tangent extended).

Salem County, NJ, developed a systematic road-safety analysis tool to examine roadway-departure crashes at bridges (Cato et al. 2013). The authors found the following roadway characteristics to be related to a higher potential of roadway-departure crashes at bridges:

- Pavement width is <22 ft.
- Shoulder width is <1.5 ft.
- Lane width is <10.5 ft.
- Pavement condition is fair or poor.
- Superelevation is minimal or nonexistent for a horizontal curve.
- Friction is fair or poor.
- Striping is fair or poor.
- Advance warning signs are minimal or nonexistent.
- Object markers are in poor condition or nonexistent.
- Abutment condition is fair or poor.
- Vertical curve exists.
- Horizontal curve exists.

Preston et al. (2013b) characterized the systemic approach to safety management and summarized the cyclical, three-element process in the *Systemic Safety Project Selection Tool* (Preston et al. 2013a). To demonstrate the contexts that are suited for systemic safety approaches, the researchers conducted a pilot test in three States: Kentucky, New York, and Washington. Table 9 provides a summary of how each State implemented the systematic approach, including the facility type, crash type, and contributing factors used for network screening.

Table 9. Systemic approach summary.

Method for Network Screening	Kentucky	New York	Washington
Facility type	County roadways	Rural undivided roadways	County arterial and collector roads
Crash type	Roadway-departure crashes on horizontal curves	Lane-departure crashes	Roadway-departure crashes
Contributing factors	Traffic volume, access density, curve density for critical radius curves, presence of advance signing, intersections on the curves, and visual traps	Number of lanes, speed, traffic volume, shoulder width, lighting conditions, and curve radius	Speed differential, visual trap, intersections, presence of advance warning signs, and edge assessment

Wang et al. (2017) compared the effects of the ranking and weighting methods using a dataset of intersection (signalized and stop controlled) pedestrian crashes from Austin, TX. They obtained crash data from the Crash Records Inventory System of the Texas DOT[2] and collected intersection characteristics data using Google® Street View™ (Google 2018). The researchers used the following contributing factors for the evaluation:

- Land-use entropy.
- Number of bus stops within 0.2 mi.
- Sidewalk presence/absence.
- Lighting presence/absence.
- Pedestrian crosswalk presence/absence.
- Percentage of one-way streets.
- Bike lane presence/absence.
- Number of approaches.
- Total number of lanes.
- Percentage of painted/raised medians.
- Average speed limit.
- Truck composition (percent single truck + percent combo truck).
- Average of approach ADTs.
- Pedestrian miles traveled.

Wang et al. (2017) used NB regression models to relate pedestrian crashes to traffic volume, intersection attributes, and contextual factors. The NB regression models revealed that land-use mix, number of approaches, percentage of one-way streets, and number of bus stops within 0.2 mi are positively correlated to pedestrian-crash potential at stop-controlled intersections. The total number of lanes and total number of bus stops within 0.2 mi of intersections are positively correlated to pedestrian-crash potential at signalized intersections. The percentage of one-way streets and percentage of painted/raised medians is negatively correlated with pedestrian crash potential at signalized intersections.

Al-Kaisy et al. (2015) developed an index that estimates crash potential to be used in proactively identifying locations or segments for potential safety improvements on low-volume roads in

[2]The Crash Records Inventory System of the Texas DOT is not available to the public.

Oregon. The index considers three major factors: geometric features, crash history, and traffic exposure. The researchers performed cross-tabulations and multivariate linear regression and correlation analyses to identify the relationship between road characteristics and observed crashes as well as their significance. Identified contributing factors include the following:

- Degree of curvature.
- Length of vertical curve.
- Lane width.
- Grade.
- Shoulder width.
- Driveway density.
- Side-slope rating.
- Fixed-object rating.
- AADT.
- Crash history.

The final crash index is a function of the roadway geometry and expressed as a value from 0.15 to 1.0, where higher values correspond to a greater crash potential. The weights of individual geometric features are based on the data analysis and not recommended for alteration. However, the authors suggested that the overall weights that define the proportion of the crash index due to geometric features versus crash history or traffic exposure may benefit from agency input.

Caldwell and Wilson (1996) described the need for a different approach to safety-improvement projects on unpaved roadways. Their approach used a survey of a steering committee to collect information about safety on unpaved roadways. The steering committee identified 12 roadway elements that should be the focus of unpaved road-safety audits. These elements included surface width, consistency, sight distance, signage, surface condition, prudent speed, horizontal curves, Percentage of trucks on the roadway, bridges, railroad crossings, vertical curves, foreslope, clear zone, and pedestrian/bicyclist presence.

Mahgoub et al. (2011) developed the rural road safety index (RRSI) using a dataset from Brookings County, SD. The RRSI ranks roadway network locations based on their safety features to identify deficiencies. The authors proposed the RRSI to support roadway-safety reviews. The study found the following factors that contributed to crash potential on the local roads of Brookings County:

- Intersection sight distance.
- ANG of approach.
- Signage.
- Road alignment.
- Vertical and horizontal curves.
- Culverts.
- Table drains.
- Location of signage.
- Visibility and legibility of signs.

Recently, two studies compared the FHWA *Systemic Safety Project Selection Tool* to other methods for selecting systemic safety improvements (Preston et al. 2013a; Knapp et al. 2014). Knapp et al. (2014) compared five selection tools/methodologies, including the following, through a comprehensive literature review:

- Minnesota Country Road Safety Plan (CRSP) (MnDOT 2020).
- FHWA *Systemic Safety Project Selection Tool* (Preston et al. 2013a).
- United States Road Assessment Program (usRAP) (usRap 2020).
- New Jersey Systemic Road Safety Tool.
- Safety Analyst (AASHTO n.d.).

The researchers compared general availability, required input data, ease of use, basis of prioritization, and potential for sensitivity analysis. The tools varied widely in terms of data-collection requirements and cost, but all tools used a rating (e.g., stars) of number of contributing factors and/or B/C ratios for location prioritization. The authors recommended the Minnesota CRSP and usRAP approaches for application on Iowa's paved secondary rural roadways.

Knapp et al. (2014) implemented the Minnesota CRSP approach by collecting five safety contributing factors for horizontal curves, seven for stop-controlled intersections, and five for segments, as outlined by Preston and Gute (2010) and Preston (2012). To apply usRAP, they collected the following data variables:

- Carriageway/roadway: number of divided and undivided highways and direction of travel.
- Distance: distance in kilometers from the start of the road segment.
- Length: length of roadway sections generally in kilometers (typically 100 m or 328 ft).
- Latitude and longitude: coordinates in decimal degrees for the starting point of each 328-ft roadway section.
- Landmark: presence of key landmarks (e.g., town/village sign, major bridge, toll booth).
- Traffic flow: number of vehicles recorded on each section of the road (e.g., AADT).
- Motorcycle percentage: percentage of motorcycles in the traffic flow.
- Observed bicycle flow: number of bicycles in the traffic flow.
- Pedestrian flow—crossing road: number of pedestrians observed crossing the road.
- Pedestrian flow—along road: number of pedestrians observed walking along the road.
- Area type: level of roadside development (e.g., urban, semiurban).
- Number of lanes for use by through traffic: total number of lanes in one direction of travel.
- One-way/two-way flow: traffic-flow operation.
- Speed: actual posted speed limit in miles per hour.
- Lane width for lanes serving through traffic: distance from the outside edge of the traveled way or the center of the edgeline marking to the center of the adjacent lane or centerline marking.
- Paved shoulder width: paved section of the roadway outside the edgeline that is safe and drivable usually measured from the center of the shoulder marking to the outside edge of the paved roadway or from the outside edge of the traveled way to the outside edge of the paved roadway.

- Unpaved shoulder width: space available for pedestrians to walk along the side of the road.
- Shoulder rumble strips: presence of rumble strips.
- Curvature: horizontal alignment (e.g., very sharp, sharp) of the roadway that may be based on advisory posted speed limits, if available.
- Quality of curve: measure of drivers' judgment to determine the sharpness of the curve and select a speed to traverse the curve.
- Overtaking demand: developed in the preprocessing stage rather than being coded and represents the frequency of or possibility that vehicles would undertake passing tactics by using the lane in the opposing direction of travel.
- Delineation: quality of the traffic-control devices.
- Vertical alignment variation: change in the roadway gradient along its length.
- Road condition: roadway-surface condition with respect to skid resistance.
- Sidewalk provision—right: presence of a sidewalk on the right side of the road.
- Sidewalk Provision—left: presence of a sidewalk on the left side of the road.
- Land use—right and left: measure of the possibility of generating pedestrian activity along the roadside.
- Side friction: extent of interaction between the activities along the roadside and traffic on the roadway.
- Pedestrian-crossing facilities: number of pedestrian-crossing facilities.
- Quality of crossing: measure of road-crossing visibility for drivers and the presence of warning signs.
- Bicycle facilities: presence of facilities (i.e., bike paths or shared lanes) for bicyclists.
- Roadside severity—separated bicycle path: qualitative measure of the severity of the roadside present on segregated bicycle facilities based on the presence, type, and location of roadside objects, crashworthy barrier, and roadside slopes.
- Motorcycle facilities: number of facilities for motorcyclists or other motorized two-wheel vehicles.
- Roadside severity—separated motorcycle path: qualitative measure of the severity of the roadside present on segregated motorcycle facilities based on the presence, type, and location of roadside objects, crashworthy barrier, and roadside slopes.
- Speed limit—separated motorcycle path: posted speed limit for segregated motorcycle paths.
- Median type—separated motorcycle path: presence of roadway infrastructure that separates opposing traffic flows for a segregated motorcycle path.
- Minor access—point density: number of driveways within each 327-ft roadway segment for both urban and semiurban areas.
- Roadside severity—right: distance to the nearest object likely to be struck by an errant vehicle on the right side of the road that could result in serious or fatal injury to vehicle occupants.
- Roadside severity—left: distance to the nearest object likely to be struck by an errant vehicle on the left side of the road that could result in serious or fatal injury to vehicle occupants.
- Intersection type: type of intersection (e.g., signalized four-leg, unsignalized three-leg).

- Intersection quality: measure of the quality (i.e., condition and location) of intersection design features (e.g., advance warnings, signs, markings) and sight distance to the intersection for approaching vehicles.
- Intersecting road volume: approach traffic volume (a rating based on volume, if known).
- Median type: presence of infrastructure separating opposing traffic flow.
- Major upgrade cost impact: measure that takes into consideration the influence of surrounding land use, environment, and topography on the cost of major developmental-upgrade projects.
- Comments: (optional) roadway safety and obstacles faced during the coding process.
- Roadwork (work zones): presence of any major roadway construction or work zones.

When applying the two approaches, Knapp et al. (2014) changed one or more input variables in each tool to adjust the coefficients (or weights) of contributing factors in the Minnesota CRSP approach. The researchers recommended that contributing factors should have the objective of identifying locations with characteristics known to impact rural roadway safety and locations should be differentiated by relatively unique contributing factors or combinations of contributing factors.

Harwood et al. (2013) compared three methods for selecting highway-infrastructure countermeasures to reduce crash frequency and severity using data from six Kentucky counties. The three methods were usRAP, the FHWA *Systemic Safety Project Selection Tool*, and road-safety audits (usRAP 2020; Preston et al. 2013a). The authors applied the FHWA tool to a 217-mi road network with a focus on ROR crashes on horizontal curves. The contributing factors selected as the basis for the network screening and indicating higher crash potential included the following:

- Horizontal curve density greater than mean density for critical curves with radii between 500 and 1,200 ft.
- Lane width <10.5 ft.
- Shoulder type not paved.
- Shoulder width <6 ft.
- Speed limit >30 mph.
- Traffic speed (higher than the speed limit or 85th-percentile speed).
- Number of lanes.
- Lane width.
- Shoulder width.
- Presence of shoulder rumble strips.
- Delineation.
- Passing demand.
- Median type.
- Curvature.
- Quality of curve signing and marking.
- Road-surface condition.
- Roadside severity.
- Number and type of intersections.

- Quality of intersection signing and marking.
- Intersecting road volume.
- Driveway density.

The results of the correlation analysis suggested that traffic volume should be a mandatory or, at least, a strongly considered contributing factor. The study also suggested that traffic volume should have a greater weight than other variables in a prioritization scheme. The researchers concluded that the FHWA tool is more flexible than usRAP when adapting to specific roadway networks and data-availability situations since it requires fewer data variables than usRAP. However, the authors noted that the FHWA tool's greater flexibility may be offset by its inability to weigh relevant contributing factors. Road-safety audits were most effective for identifying missing safety-related features or those in poor condition.

SUMMARY

Historically, researchers have attempted to identify contributing roadway factors related to various crash types, mainly through datasets developed to answer a specific question or focus on a specific set of characteristics (e.g., horizontal curve–related elements). Most research has focused on passenger cars or all user types combined. Researchers have used a variety of statistical methodologies to examine the relationship, but GLMs are most commonly applied to associate roadway characteristics with expected crash frequency. A relatively low number of studies focused on identifying contributing roadway factors for systemic safety analysis; however, the factors most commonly identified in the studies reviewed in this chapter should be prioritized for inclusion in future systemic safety analyses.

Literature that focused on systemic analyses consistently concluded that the analytical process for identifying candidate locations for investment in rural areas was not well developed and varied based on weights assigned to contributing factors. Multiple proposed strategies for identifying sites have strengths and limitations, with some being more data intensive than others. The less data-intensive methods may lack the ability to conduct more robust analyses. To this point, the ability to identify sites requires robust data of good quality, which may be more difficult to obtain for lower-volume roads, roads owned by local jurisdictions, or unpaved roads. Finally, researchers identified that local agencies sometimes lack experience in highway-safety planning, and staff may not be familiar with methods to conduct systemic safety analyses. Local agencies will benefit from the help of State agencies and research that can preidentify contributing factors associated with FCFTs. This research will provide a targeted list of data for agencies to collect to screen and prioritize locations, as well as identify related countermeasures.

CHAPTER 3. TERMINOLOGY

Drawing on information from the literature review in chapter 2, this chapter includes definitions for crash types, facility types, crash potential, and contributing factors. These definitions will guide the use of such terms throughout the remainder of this report.

Crash types are defined by the unique combinations of codes for two variables in crash reports: First Harmful Event and Manner of Collision. The First Harmful Event is the first injury- or damage-producing event of a crash. First Harmful Events generally fall into one of five categories, with more specific event codes under each category:

1. Noncollision Harmful Events (e.g., rollover/overturn [ROLL], fire/explosion, jackknife).
2. Collision with Motor Vehicle in Transport.
3. Collision with Object Not Fixed (e.g., pedestrian, pedalcyclist, railway vehicle, animal).
4. Collision with Fixed Object (e.g., tree, utility pole, embankment).
5. Not Reported or Unknown.

The Manner of Collision variable identifies the orientation of two motor vehicles in transport when they were involved in the First Harmful Event of a crash (e.g., front to rear, front to front, ANG), providing information on the crash type. Typically, a significant number of options are available on crash reports for defining the Manner of Collision. For example, the corresponding codes may indicate if the crash type is a particular type of ANG, rear-end, or sideswipe crash. If the First Harmful Event is not a collision between two motor vehicles, the Manner of Collision is classified as such (e.g., “Not a Collision with Motor Vehicle in Transport”).

Facility types are defined by surrounding area type/land use, number of lanes, level of access control, and median presence, resulting in the following example facility types:

- Rural four-lane freeway.
- Urban four-lane freeway.
- Urban six-lane freeway.
- Urban eight-lane freeway.
- Rural four-lane highway (nonfreeway):
  - Divided.
  - Undivided.
- Rural two-lane highway.
- Urban six-lane street:
  - Divided.
  - Undivided.
- Urban four-lane street:
  - Divided.
  - Undivided.
- Urban two-lane street.

Facility types are further broken down into site types, such as the following:

- Urban six-lane freeway:
  - Freeway segment.
  - Freeway-ramp terminal.
  - Ramp.
  - Ramp-crossroad terminal.
- Rural two-lane highway:
  - Tangent segment.
  - Horizontal curve segment.
  - Intersection.

The definition of crash potential captures both the probability of an event occurring and its negative impact. The event of interest is the occurrence of a specific crash type. Event occurrence can be defined in several ways, including the following:

- Expected frequency.
- Probability mass (probability of 0, 1, 2, 3, or more events).
- Cumulative probability (probability of 1 or more, 2 or more, 3 or more events).

The negative impact of an event refers to its negative impact with respect to cost, health, human life, or other personal or societal impacts. This negative impact by the severity of injuries resulting from a crash event is defined per the following scale:

- K = fatal injury.
- A = incapacitating injury.
- B = nonincapacitating injury.
- C = possible injury.
- O = no injury.

Contributing factors are defined as factors whose presence is associated with increases or decreases in expected frequencies of crashes or injury severities resulting from crashes. Chapter 5 will distinguish contributing factors associated with increases in crash frequency or crash severity and contributing factors associated with decreases in crash frequency or crash severity.

CHAPTER 4. FCFTS

For the purpose of this project, the research team conducted an analysis of the FARS and HSIS databases to inform the selection of FCFTs. FARS contains information on all K crashes that involved a motor vehicle traveling on a public trafficway in the 50 States, the District of Columbia, and Puerto Rico. HSIS contains information on crashes of all severities (fatal, injury, and PDO) that occurred on State-operated and -maintained roads for the participating States. This analysis incorporated data from four States that are part of HSIS (California, Minnesota, Ohio, and Washington). This chapter provides additional detail on the data, methods, and results regarding the selection of FCFTs.

FARS DATABASE AND ELEMENTS

A K crash is a crash in which the death of at least one person involved in the crash occurred within 30 days of the crash. FARS consists of several different files that contain information pertaining to K crashes. The research team analyzed three data files—Accident, Vehicle, and Person[1]—to inform the selection of FCFTs. The following are brief descriptions of the information available in these files:

- The Accident data file contains information about crash characteristics and environmental conditions at the time of the crash. This file contains one record per crash.
- The Vehicle data file contains information describing the in-transport motor vehicles involved in the crash. This file contains one record per in-transport motor vehicle involved in the crash.
- The Person data file contains information describing all persons involved in the crash, including both motorists and nonmotorists. This file contains one record per person involved in the crash.

These files can be merged using common variables, such as State, case number, vehicle number, and person number. The research team planned to use 6 years of FARS data ranging from 2009 to 2014; however, due to limitations in the 2009 data with respect to intersection information[2], the analysis incorporated FARS data from 2010 to 2014.

The research team narrowed down a list of key variables for identifying FCFTs for each file. This list and a brief description of each variable are provided in appendix A.

[1]Some variables in the Vehicle file used to be in the Accident file. For example, VTRAFCON (a traffic-control device) was part of the Accident file but is now in the Vehicle file.

[2]In FARS, data on intersection type were not collected prior to 2009.

HSIS DATABASE AND ELEMENTS

HSIS is operated by FHWA and consists of several different files that contain information pertaining to crashes of all severities (fatal, injury, and PDO) occurring on State-operated and -maintained roads for participating States. The four HSIS data files used for the purpose of informing the FCFT selection were the Accident, Roadway, Intersection[3], and Curve[4] files, which were used to collect data on California, Minnesota, Ohio, and Washington. The research team was able to efficiently identify intersections using Intersection files available for California, Minnesota, and Ohio and horizontal curves using Curve files available for Ohio and Washington.

The following are brief descriptions of the information available in these four file types:

- The Accident data file contains information about crash characteristics and environmental conditions at the time of the crash. This file contains one record per crash occurring on State-owned or -operated roads.
- The Roadway data file contains information on homogenous sections of roadway (i.e., stretches of road that are consistent in terms of certain characteristics, with new sections defined anytime a characteristic changes). Each record represents a homogenous roadway segment and contains characteristics of the roadway segment, such as width of the traveled way, number of lanes, width of paved shoulder and total shoulder, median type, and other variables.
- The Intersection data file contains information on both mainline routes and crossing routes of at-grade intersections. The information includes items such as intersection type, traffic control type, lighting, channelization, and AADT.
- The Curve data file contains information on sections of roadway that are part of horizontal curves. Each record includes variables related to deflection ANG, curve direction, degree of curvature (and radius), curve length, speed limit, and whether or not the horizontal curve overlaps with a preceding curve.

These HSIS files can be merged using common variables, such as route number, system number, milepost (for crashes and intersections), and beginning and end mileposts (for roadway segments). The analysis incorporated 6 years of HSIS data ranging from 2009 to 2014[5].

The research team narrowed down a list of key variables for identifying FCFTs for each file. This list and a brief description of each variable are provided in appendix A.

KEY VARIABLES FOR DEFINING POTENTIAL FCFTS

To define potential FCFTs, the research team identified the following variables (from the variable list in appendix A) as the most relevant:

[3]The Intersection file was only used for California during the FCFT analysis. For the other three States, relevant intersection information for identifying FCFTs was obtained from the Accident file.

[4]The Curve file was only used for Washington during the FCFT analysis. For the other three States, relevant road-alignment information for identifying FCFTs was obtained from the Accident file.

[5]The Intersection file for California was not available for 2013. As such, data from 2009 to 2012 and 2014 from Intersection files were used. For the other three States, data from 2009 to 2014 were used.

- Crash type.
- Area type (rural versus urban).
- Roadway type.
- Location (intersection versus nonintersection)[6].
- Intersection type.
- Type of traffic control.
- Lighting (day versus night).
- Road alignment (horizontal curve versus straight/tangent).

The research team explored the possibility of including other variables, such as weather, driver demographics, and driver behavior (e.g., speeding), but concluded that these variables would be analyzed as potential contributing factors rather than as factors for identifying and defining FCFTs. This conclusion is generally consistent with how systemic safety analyses in literature and practice have defined FCFTs and potential contributing factors, as well as with the state of knowledge, linking infrastructure-related countermeasures to crash types. For example, specific infrastructure-related countermeasures exist for single-vehicle ROR crashes at night, but existing practices do not distinguish different infrastructure-related countermeasures specific to single-vehicle ROR crashes at night for 16- to 19-year-old drivers versus other age groups. However, knowing that the number of 16- to 19-year-old drivers that use a segment is associated with the expected frequency of single-vehicle ROR crashes at night (i.e., a contributing factor) could help agencies identify which segments to treat systemically with proven countermeasures for ROR crashes.

The following sections include detailed descriptions of how the research team derived the key variables used to define FCFTs, along with information on how these key variables were extracted from FARS and HSIS.

Crash Type

The research team defined crash type based on different variables identifying the manner of collision and first harmful event. The information used from each dataset to extract the crash type was as follows:

- FARS: first harmful event and manner of collision.
- HSIS, California: accident type.
- HSIS, Minnesota: accident type and diagram of accident type.
- HSIS, Ohio: accident type.
- HSIS, Washington: accident type and first collision type.

Area Type

The research team defined area type as one of two categories: rural or urban. This information was extracted using both the area-type identifier present within the dataset and the roadway functional-class variable.

[6]In FARS, location type was not available in 2009, so the research team used data from 2010 to 2014.

Roadway Type

FARS and HSIS define roadway type differently. To define roadway type, FARS uses the following variables: roadway functional type, number of lanes, and median type. These variables allow for defining roadway type as interstates/freeways/expressways or two-lane/multilane divided or undivided roadways. HSIS uses the roadway classification variable, which disaggregates road type into freeways or two-lane/multilane divided or undivided roadways.

Work zone–related crashes were excluded from this analysis. Work zones have specific contributing factors linked to temporary traffic-control characteristics and the type of construction or maintenance work. This type of information was not readily available for this analysis as it is typically pulled from State DOT records and construction diaries.

Location Type

The research team further identified crashes as being intersection or nonintersection crashes using the intersection variable in FARS and the location type variable in HSIS. For this analysis, railroad-related crashes, as identified in the Accident files of FARS and HSIS, were excluded from all analyses. Contributing factors for at-grade railroad crossings often include the frequency with which trains pass through the location; however, this type of information was not readily available for this analysis.

Intersection Type

The research team classified crashes according to the type of intersection[7] (i.e., four-way intersection, T-intersection, Y-intersection) using the intersection type variable in FARS and location type in HSIS.

Traffic Control Type

The research team used the traffic control variable in FARS and HSIS[8] to define the type of traffic control (i.e., stop controlled, traffic signal, yield) at the location of the crash.

Light Condition

The research team used the light condition variable in FARS and HSIS to define the light condition at the time of the crash. Daytime crashes were defined as any crash that occurred during the daytime. To define nighttime crashes, the following categories were combined: dark–lighted, dark–not lighted, dark–no lights present, dawn, and dusk.

Road Alignment Type (Nonintersection Crashes)

The research team defined road alignment as either a straight segment (i.e., horizontal tangent) or horizontal curve. This information was extracted using the road alignment variable in FARS and

[7]HSIS data regarding intersection type were not available in the Accident files for Ohio and Washington; thus, crashes were identified as intersection- or nonintersection-related crashes for the purpose of informing FCFT selection.

[8]The traffic control variable was not available in the HSIS Accident files for Ohio and Washington.

road characteristic variable in HSIS. This information was only extracted for nonintersection crashes. The road alignment variable was not available for California. In Washington, the Curve file was used to determine whether the crash occurred on a curve or straight segment.

Fatal and Incapacitating Injury Crash

In the HSIS files, KA crashes were identified using the accident severity variable.

POTENTIAL FCFTS

After categorizing the data using the variables identified in the section Key Variables for Defining Potential FCFTs, the research team developed a list of potential FCFTs. First, they developed separate lists of potential FCFTs using data from FARS and each HSIS State to observe and address consistencies and differences across the databases before selecting FCFTs that would become the focus of the remainder of this research.

The research team defined potential FCFTs using a combination of the following variables:

- Area type (rural versus urban).
- Crash type.
- Intersection type.
- Lighting type (day versus night).
- Location type (intersection versus nonintersection).
- Road alignment type (horizontal curve versus straight segment).
- Roadway type.
- Traffic control type.

For example, HEO-D crashes on rural two-lane roads on straight segments represented one potential FCFT.

Potential FCFTs were then ranked using the number of K crashes and, for HSIS States, the number of KA crashes during the observation period corresponding to each potential FCFT. The analysis developed separate ranked lists based on FARS and each HSIS State to observe whether certain FCFTs consistently corresponded to higher or lower numbers of K crashes and KA crashes or whether the rankings varied considerably by database.

This analysis approach is slightly different from the systemic safety-planning process described in FHWA's *Systemic Safety Project Selection Tool* (discussion of this publication is in chapter 2) (Preston et al. 2013a). The systemic safety-planning process described by Preston et al. (2013a), as it applies to identifying FCFTs, is sequential and involves two main tasks.

The first task is the selection of focus crash types:

> This task involves conducting a system wide analysis of crash types to select those representing the greatest potential to reduce fatalities and severe injuries. This effort typically results in identifying the crash types that represent the greatest number of severe crashes across the system being analyzed. (Preston et al. 2013a, p. 11)

The second task is the selection of focus facilities for the focus crash types identified in the first task:

> After selecting the focus crash types, Task 2 answers the question *where are the crashes occurring*? A 'crash tree' diagram is the recommended approach to answer this question. The crash tree can have a number of different formats, depending on agency capabilities and data availability. One such example is to begin the crash tree with the total number of severe crashes at the highest level. Each subsequent level separates the severe crashes by facility type. (Preston et al. 2013a, p. 16)

The research team sought to determine whether the process detailed in this chapter for analyzing FCFTs yielded the same or different results as the process outlined in FHWA's *Systemic Safety Project Selection Tool* (Preston et al. 2013a). To accomplish this goal, the research team applied FHWA's approach to data from FARS on nonintersection crashes. The FCFTs selected from implementing this approach were consistent to those selected using the process detailed in this chapter (the results of this comparison are in appendix B).

Tables detailing potential FCFTs are available in appendix C. Table 93 and table 94 show the numbers of K crashes from FARS categorized by intersection and nonintersection crashes and sorted by the frequency of K crashes associated with particular combinations of the variables identified in the section Key Variables for Defining Potential FCFTs. Table 95 through table 108 show the numbers of K and KA crashes from the HSIS States categorized by intersection and nonintersection crashes and sorted by the frequency of K and KA crashes associated with particular combinations of the variables identified in the section Key Variables for Defining Potential FCFTs.

Looking at the results from FARS and three HSIS States (Minnesota, Ohio, and Washington) it was clear that two-lane roads are the facility type with the highest number of K crashes. More specifically, lane-departure (LNDP) crashes on segments (including ROR, ROLL, HEO, and ANG crashes) and ANG crashes at intersections seem to have the highest numbers of fatal or KA crashes. HSIS only includes information from State-maintained roads, and FARS includes K crashes from all public roads; however, this did not impact the type of consistency observed between the FARS and HSIS State datasets for these potential FCFTs. Freeway crashes were prevalent in the HSIS California data but not so much in the data from FARS and the other HSIS States. The large numbers of higher AADT on urban freeways in the HSIS California data partially explains this finding.

SELECTING FCFTS

The overall work plan for this project was based on an in-depth analysis of approximately 15 focus crash type–facility type combinations to identify contributing crash factors. Observations made during the FARS and HSIS State-specific analyses described earlier in this chapter served as the basis for selecting the 15 combinations for further analysis with respect to contributing crash factors. The remainder of this section provides additional information on how the research team addressed two crash types: pedestrian and ROR. Both of these crash types represent a significant number of K crashes and also have unique data-related challenges. The concluding sections of this chapter provide the nonintersection and intersection FCFTs.

Based on FARS data from 2010 to 2014, pedestrian crashes represented a significant number of K crashes (NHTSA 2018a):

- 5,286 (out of 11,507) K crashes were coded as intersection crashes.
- 14,817 (out of 77,664) K crashes were coded as nonintersection crashes.

Identifying contributing factors for pedestrian crashes is a challenging task without quality exposure data (i.e., the number of pedestrians crossing at segments and intersections or walking parallel to segments). Methods for analyzing pedestrian crashes were very limited in the first edition of the HSM for this reason, and filling this gap in the second edition is a high priority for the research community and AASHTO.

At the time of this project, two other efforts related to identifying factors contributing to pedestrian crash frequency and severity were ongoing:

- NCHRP 17-73, *Systemic Pedestrian Safety Analysis*, was funded a total of $300,000 for 2.5 years (Thomas et al. 2018).
- NCHRP 17-84, Pedestrian and Bicycle Safety Performance Functions for the *Highway Safety Manual*, was funded a total of $500,000 for 2 years (National Academy of Sciences n.d.).

The research team explored an additional original analysis using available data but was not confident that a quality analysis of contributing factors could be conducted with the existing data, which did not include pedestrian exposure and had some questionable crash-location coding (e.g., 350 K crashes and an additional 192 A crashes on California freeways at night). However, given California's significant numbers, chapter 5 includes a discussion of contributing factors for pedestrian crashes, which is based on information from related published literature.

For nonintersections, collisions with roadside fixed objects corresponded to the highest number of K crashes, but the research team decided not to specifically address fixed-object crashes for the following reasons:

- The HSIS Roadway files have very little information about roadsides, which are characterized by a number of factors that likely influence ROR crashes (e.g., roadside slopes, types and offsets to fixed objects, types and offsets to barriers).
- The Low-Cost Safety Improvements Pooled Fund Study, administered through the FHWA Office of Safety Research and Development, is sponsoring an effort to quantify safety effects of the presence, removal, or shielding of roadside fixed objects.

This analysis includes ROR and LNDP crashes since they represent a significant number of K crashes in the country. ROR crashes are a subset of LNDP crashes. LNDP crashes are crashes in which a vehicle departs the travel lane, resulting in a ROR, ROLL, HEO, or ANG crash. This definition of LNDP is intended to be synonymous with FHWA's definition of a roadway-departure crash: a crash that occurs after a vehicle crosses an edgeline or centerline or otherwise leaves the traveled way (FHWA 2017b). ROR crashes are any crashes in which a vehicle strikes a roadside object or otherwise leaves the travel lane and enters the roadside.

Intersection FCFTs

The following is a list of FCFTs for which an original analysis of contributing factors was executed as part of this project:

- ANG crashes on rural two-lane roads at four-leg stop-controlled intersections (daytime and nighttime).
- ANG crashes on urban two-lane roads at four-leg stop-controlled intersections (daytime).
- ANG crashes on urban multilane divided roads at four-leg signalized intersections (daytime).
- ANG crashes on urban multilane undivided roads at four-leg signalized intersections (daytime).
- ANG crashes on rural two-lane roads at three-leg stop-controlled intersections (daytime).
- ANG crashes on rural multilane divided roads at four-leg minor-road stop-controlled intersections (daytime).

Nonintersection FCFTs

The following is a list of FCFTs for which an original analysis of contributing factors was executed as part of this project:

- ROR crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
- ROR crashes on rural two-lane roads on straight segments (daytime and nighttime).
- LNDP crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
- LNDP crashes on rural two-lane roads on straight segments (daytime and nighttime).
- HEO crashes on rural two-lane roads on straight segments (daytime and nighttime).
- ANG crashes on rural two-lane roads on straight segments (daytime).
- HEO crashes on rural two-lane roads on horizontal curves (daytime and nighttime).
- ROLL crashes on rural two-lane roads on straight segments (daytime and nighttime).
- ROLL crashes on rural two-lane roads on horizontal curves (daytime and nighttime).

This project also addresses contributing factors for two types of pedestrian crashes using information from related published literature:

- All types of pedestrian crashes at intersections.
- Pedestrian crashes at intersections involving a crossing pedestrian and vehicle going straight.

CHAPTER 5. CONTRIBUTING FACTORS

This chapter includes descriptions of the data, methodology, and results for identifying and evaluating contributing factors corresponding to the selected FCFTs noted in chapter 4. According to *Systemic Safety Project Selection Tool*, "Tasks 1 and 2 [Select Focus Crash Types and Select Focus Facilities] relied on data typically in the crash record system. Task 3 [Identify and Evaluate Contributing Factors] is the first point where road and intersection inventories are likely needed to provide additional levels of detail to support the data analysis" (Preston et al. 2013a, p. 18).

DATA

Three different data sources were used to conduct the contributing-factor analysis:

- HSIS (FHWA 2018c).
- Databases from the National Oceanic and Atmospheric Administration's (NOAA's) National Centers for Environmental Information (2018).
- Databases from the U.S. Census Bureau (2018).

The research team initially planned on using HSIS data from four States—California, Minnesota, Ohio, and Washington—for the contributing-factor analysis. The plan was to use California and Minnesota data for the intersection contributing-factor analysis and Ohio and Washington data for the nonintersection contributing-factor analysis. However, the research team encountered some challenges using the Minnesota and Ohio datasets for this purpose. In the Minnesota data, major and minor roads at intersections were not always clearly defined and a framework to efficiently link climate and socioeconomic factors to the Minnesota HSIS Roadway file was not available. In the Ohio data, the main issue was that the HSIS Curve file had not been updated since 2009. During the analysis, issues with the Ohio data were resolved when the Ohio DOT (ODOT) provided their latest Curve Inventory, along with an intersection database. The research team resolved the issues with the Minnesota data, but the timing of these issues did not allow the research team to incorporate the data into the analysis. Thus, the analysis proceeded using Ohio and Washington data for the intersection contributing-factor analysis and California and Ohio data for the nonintersection contributing-factor analysis.

The research team assembled data on climate and socioeconomic factors for the sites in California, Ohio, and Washington and linked these data to the study locations for use in the contributing-factor analysis. This process involved the following steps:

- *Acquire data:* The research team obtained climate data from NOAA (2018) and socioeconomic data from the U.S. Census Bureau (2018). Both the climate and census datasets were available in spatial format.

- *Plot study sites:* For each State (California, Ohio, and Washington), the research team developed a list of locations for the analysis, either intersections, curve segments, or tangent segments based on HSIS data. They then plotted each location spatially using a spatial roadway network obtained from each State.

- *Associate data:* The research team joined the climate and census data to the study location in the spatial environment. Climate data was joined to the study-site location based on the nearest weather station in a straight-line distance. For census data, the join was based on the census block that contained the study site. For segments, the census block that housed the midpoint of the segment served as the basis for the join.

- *Assemble data fields required for analysis:* The research team assembled a working database that consisted of selected fields from the climate and census data. In the census data, the research team developed certain variables based on the raw data, such as "percentage of households with two or more vehicles."

POTENTIAL CONTRIBUTING FACTORS

Potential contributing factors were defined for three different categories: roadway, climate, and census. Among potential contributing factors, it is possible that some can be direct contributing factors while others may be surrogates for other characteristics that were not available (e.g., data were not available for roadsides, driveways, how well roads were maintained, or travel patterns or behavior other than AADT). In all cases in this report, findings regarding contributing factors are interpreted as statistical associations with expected crash frequencies and severities. The report does not spend time covering the philosophy of causation or causal inference. That said, the following characteristics raise confidence that a specific finding or set of findings are stable and transferable:

- Consistency across subsets of related FCFTs.
- Consistency across multiple States.
- Consistency with previous findings in the literature.

Roadway

The analysis explored potential contributing factors associated with roadways. The following list contains the contributing factors and their associated variable names:

- Average AADT (avg_aadt).
- Mainline AADT or major AADT (ML_AADT or majorAADT).
- Cross street AADT or minor AADT (XST_AADT or minorAADT).
- Segment length[1].
- Curve radius (curv_rad).
- Percent grade (pct_grad)[2].
- Grade type (grad_typ).
- Shoulder type (shl_typ).
- Shoulder width (shldwid).

---

[1]Instead of using segment length as a direct potential contributing factor in the analysis, the research team used it to define crash rate per mile.

[2]In the Ohio data, grade information was also available for segments and intersection approaches. The percent grade was divided into three categories (level = <3 percent, moderate = 3–6 percent, and steep = >6 percent) and used as a categorical variable for the intersection contributing-factor analysis. The segment analysis used the absolute value of grade as a continuous variable.

- Terrain (terrain).
- Speed limit (spd_limt or DESG_SPD).
- Lane width (lanewidth).
- Percentage of trucks on the roadway (trkpcts).
- Median width (MEDWID).
- Surface width (surfwid).
- Mainline left-turn channelization (ML_LEFT).
- Mainline right-turn channelization (ML_RIGHT).
- Cross street left-turn channelization (xstrtlft).
- Cross street right-turn channelization (XSTRTRGH).
- Number of approaches with left-turn lanes (leftLanes).
- Number of approaches with right-turn lanes (rightLanes).

Climate

The analysis explored the following potential contributing factors associated with climate:

- Average annual snowfall totals (snowavgyear).
- Average annual rainfall totals (rainavgyear).
- Average annual maximum temperatures (tempmaxavg).
- Average annual minimum temperatures (tempminavg).
- Average annual winter minimum temperatures (tempwintermin).
- Average annual number of days with a minimum temperature of ≤32°F (temp32fdays).

Census

The analysis explored the following potential contributing factors associated with socioeconomic characteristics:[3]

- Percentage of population ages 16+ unemployed (unempl16plus).
- Percentage of population ages 16–24 working full time (workftage16to24).
- Percentage of population ages 16–24 working part time (workptage16to24).
- Percentage of population ages 16–24 unemployed (noworkage16to24).
- Percentage of population ages 25+ without a high school diploma (noedctn25plus).
- Percentage of population ages 25+ with a high school diploma (diploma25plus).
- Percentage of population ages 25+ with a university degree (univ25plus).
- Percentage of households with income <$50,000 (income50k).
- Percentage of households with income between $50,000 and $100,000 (income50to100k).
- Percentage of households with income >$100,000 (income100kplus).
- Percentage of households with no vehicles (noveh).
- Percentage of households with one vehicle (X1veh).
- Percentage of households with two or more vehicles (X2vehplus).
- Percentage of population ages 15–19 (age15to19).

[3]These percentages represent people or households within the census block that contains the segment or intersection.

- Percentage of population ages 20–44 (age20to44).
- Percentage of population ages 45–64 (age45to64).
- Percentage of population ages 6–74 (age65to74).
- Percentage of population ages 75+ (age75plus).

SUMMARY OF DATA

Table 10 through table 25 provide summary statistics (minimum, maximum, average, and standard deviation) for roadway, climate, and socioeconomic variables for California, Ohio, and Washington datasets.

Roadway variables specified as categorical variables in the analysis include lane width, shoulder type, surface type, terrain, and curve radius (for segment and intersection analysis using Ohio data). The following sections detail the categories used for each variable.

Lane Width

The research team divided lane width into three categories:

- <11 ft.
- 11–12 ft.
- >12 ft.

Shoulder Type

The research team divided shoulder type into three categories with information about whether left and right shoulders were paved or unpaved:

- Paved–paved.
- Paved–unpaved.
- Unpaved–unpaved.

Surface Type

The research team divided surface type into six categories for Ohio and five categories for Washington. The six categories used for Ohio included the following:

- Combination surface.
- Brick.
- Reinforced concrete.
- Plain concrete.
- Dense-graded asphaltic concrete.
- Open-graded road mix.

The five categories used for Washington included the following:

- Asphalt.
- Bituminous.
- Gravel.
- Portland concrete.
- Other.

Terrain Type

The research team divided terrain into three categories:

- Level.
- Rolling.
- Mountainous.

Curve Radius

For the analysis of Ohio data only, the research team divided horizontal curve radius data into three categories for intersection analysis and four for nonintersection analysis. The three categories used for intersection analysis included the following:

- <500 ft.
- 500–1,000 ft.
- >1,000 ft (including intersections on tangent segments).

The four categories used for nonintersection analysis included the following:

- <500 ft.
- 500–1,000 ft.
- 1,000–1,500 ft.
- >1,500 ft (including tangent segments).

Data Quality Checks

The extent of the analysis across a broad range of FCFTs did not allow the research team to conduct a site-by-site data check and enhancement using tools like aerial photography and video logs. The research team utilized descriptive statistics, including mean, standard deviation, minimum, and maximum values, as well as scatter plots and histograms, to identify sites with variable values outside of reason. After this initial data-screening process, maximum measurements for mainline and cross street AADTs at intersections and shoulder widths at nonintersection locations still stood out as "borderline" to the research team. As such, the research team talked with ODOT to confirm maximum mainline and cross street AADTs for at-grade intersections across the State. They compared ranges of mainline and cross street AADTs for at-grade intersections in California to AADT ranges covered by the HSM predictive methods and found the California data to be within reason. Finally, the research team explored some of the wider shoulder-width segments (e.g., 16-ft shoulder widths). They found these measurements to be representative of a shoulder and also found that these segments make up a small portion of the overall samples. Therefore, these segments were kept in the analysis.

Table 10. Summary statistics for intersections (all types) in California: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Mainline AADT (ML_AADT) (veh/day)	52.00	111,600.00	13,927.96	14,184.35
Cross street AADT (XST_AADT) (veh/day)	1.00	77,000.00	1,309.18	3,593.86
Mainline number of lanes (ml_lanes)	1.00	8.00	2.82	1.22
Cross street number of lanes (xstlanes)	1.00	8.00	2.04	0.48
Median width (MEDWID) (ft)	0.00	99.00	7.56	17.09
Design speed (DESG_SPD) (mph)	25.00	70.00	51.67	11.19

Table 11. Summary statistics for intersections (all types) in California: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	0.00	202.60	11.06	32.94
Average annual rainfall totals (rainavgyear) (inches)	2.36	90.73	23.50	15.37
Average annual maximum temperatures (tempmaxavg) (°F)	56.10	91.40	71.56	6.79
Average annual minimum temperatures (tempminavg) (°F)	18.50	64.90	46.31	6.99
Average annual winter minimum temperatures (tempwintermin) (°F)	5.60	51.20	36.79	7.55
Average annual number of days with a minimum temperature of ≤32°F (temp32fdays)	0.00	307.50	43.43	60.30

Table 12. Summary statistics for intersections (all types) in California: census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	72.34	12.27	8.78
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	13.60	16.47
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	40.79	23.45
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	45.61	24.68
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	0.00	100.00	54.23	15.27
Percentage of population ages 25+ with a high school diploma (diploma25plus)	0.00	70.27	20.63	9.14
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	98.16	25.14	18.15
Percentage of households with income <$50,000 (income50k)	0.00	100.00	48.63	20.29
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	76.34	27.74	11.01
Percentage of households with income >$100,000 (income100kplus)	0.00	91.93	22.79	17.40
Percentage of households with 0 vehicles (noveh)	0.00	86.29	6.44	8.22
Percentage of households with 1 vehicle (X1veh)	0.00	100.00	31.45	13.28
Percentage of households with ≥2 vehicles (X2vehplus)	0.00	100.00	61.06	18.65
Percentage of population ages 15–19 (age15to19)	0.00	68.01	6.12	4.32
Percentage of population ages 20–44 (age20to44)	0.00	88.26	30.10	11.31
Percentage of population ages 45–64 (age45to64)	0.27	100.00	30.05	9.73
Percentage of population ages 65–74 (age65to74)	0.00	55.77	9.89	6.60
Percentage of population ages 75+ (age75plus)	0.00	74.59	6.85	5.39

Table 13. Summary statistics for curved segments in Washington: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Segment length (seg_lngth) (mi)	0.01	0.67	0.061	0.058
Curve radius (curv_rad) (ft)	100.00	12,000.00	2,240.91	3,022.83
Average AADT (avg_aadt) (veh/day)	109.00	25,599.00	2,638.68	2,794.56
Percent grade (pct_grad)	0.00	11.91	2.23	2.07
Shoulder width (shldwid) (ft)	0.00	20.00	3.96	2.31
Speed limit (spd_limt) (mph)	25.00	60.00	49.20	11.60
Percentage of trucks on the roadway (trkpcts)	0.00	61.52	14.50	8.61

Table 14. Summary statistics for curved segments in Washington: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	0.00	139.90	19.64	27.85
Average annual rainfall totals (rainavgyear) (inches)	0.00	119.72	43.25	33.90
Average annual maximum temperatures (tempmaxavg) (°F)	46.70	66.20	59.61	2.57
Average annual minimum temperatures (tempminavg) (°F)	30.60	45.70	38.76	3.57
Average annual winter minimum temperatures (tempwintermin) (°F)	15.90	38.30	27.58	5.88
Average annual number of days with a minimum temperature of ≤32°F (temp32fdays)	13.30	212.20	103.31	48.44

Table 15. Summary statistics for curved segments in Washington: census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	81.08	11.44	8.96
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	11.51	14.17
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	47.91	23.79
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	40.58	23.66
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	15.55	91.66	54.12	9.85
Percentage of population ages 25+ with a high school diploma (diploma25plus)	3.10	49.07	24.92	8.44
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	77.34	20.96	10.32
Percentage of households with income <$50,000 (income50k)	0.00	91.04	50.48	14.71
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	100.00	32.19	10.37
Percentage of households with income >$100,000 (income100kplus)	0.00	87.46	16.38	10.59
Percentage of households with 0 vehicles (noveh)	0.00	33.07	3.40	4.26
Percentage of households with 1 vehicle (X1veh)	0.00	62.10	23.21	10.30
Percentage of households with ≥2 vehicles (X2vehplus)	0.00	100.00	72.69	13.70
Percentage of population ages 15–19 (age15to19)	0.00	26.42	5.82	3.39
Percentage of population ages 20–44 (age20to44)	0.00	89.36	24.18	9.62
Percentage of population ages 45–64 (age45to64)	0.00	68.00	34.17	10.52
Percentage of population ages 65–74 (age65to74)	0.00	56.66	12.11	6.53
Percentage of population ages 75+ (age75plus)	0.00	34.64	6.68	4.14

Table 16. Summary statistics for tangent segments in Washington: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Segment length (seg_lngth) (mi)	0.01	4.82	0.095	0.137
Average AADT (avg_aadt) (veh/day)	95.00	25,599.00	2,920.16	3,016.06
Percent grade (pct_grad)	0.00	16.13.00	1.85	1.90
Shoulder width (shldwid) (ft)	0.00	20.00	4.02	2.49
Speed limit (spd_limt) (mph)	25.00	60.00	48.65	11.98
Percentage of trucks on the roadway (trkpcts)	0.00	59.62	14.58	9.04

Table 17. Summary statistics for tangent segments in Washington: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	0.00	139.90	22.74	28.06
Average annual rainfall totals (rainavgyear) (inches)	7.42	119.72	40.80	32.37
Average annual maximum temperatures (tempmaxavg) (°F)	46.70	66.20	59.83	2.66
Average annual minimum temperatures (tempminavg) (°F)	30.60	45.70	38.88	3.57
Average annual winter minimum temperatures (tempwintermin) (°F)	15.90	38.30	27.59	5.79
Average annual number of days with a minimum temperature of ≤32°F (temp32fdays)	13.30	212.20	103.42	47.98

Table 18. Summary statistics for tangent segments in Washington: census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	81.08	11.05	8.81
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	12.62	15.17
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	46.94	22.93
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	40.44	22.66
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	15.55	88.80	54.77	9.72
Percentage of population ages 25+ with a high school diploma (diploma25plus)	3.10	49.07	25.32	8.03
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	77.34	19.91	9.89
Percentage of households with income <$50,000 (income50k)	0.00	91.04	50.56	14.87
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	100.00	32.38	10.86
Percentage of households with income >$100,000 (income100kplus)	0.00	87.46	16.48	10.53
Percentage of households with 0 vehicles (noveh)	0.00	33.07	3.59	4.20
Percentage of households with 1 vehicle (X1veh)	0.00	59.91	22.89	10.30
Percentage of households with ≥2 vehicles (X2vehplus)	0.00	100.00	73.09	12.93
Percentage of population ages 15–19 (age15to19)	0.00	26.42	6.02	3.26
Percentage of population ages 20–44 (age20to44)	0.00	89.36	24.87	10.04
Percentage of population ages 45–64 (age45to64)	0.00	68.00	33.53	10.02
Percentage of population ages 65–74 (age65to74)	0.00	56.66	11.83	6.48
Percentage of population ages 75+ (age75plus)	0.00	38.30	6.73	4.04

Table 19. Summary statistics for curved segments in Ohio: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Segment length (seg_lngth) (mi)	0.02	0.33	0.058	0.031
Curve radius (curv_rad) (ft)	100.00	1,468.00	491.27	224.11
Percent grade (pct_grad)	0.00	20.00	3.63	4.74
Average AADT (avg_aadt) (veh/day)	93.00	12,210.00	1,133.51	1,134.64
Speed limit (spd_limt) (mph)	25.00	55.00	52.45	5.86
Shoulder width (shldwid) (ft)	0.00	11.00	2.33	1.56

Table 20. Summary statistics for curved segments in Ohio: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	4.70	109.90	20.16	9.46
Average annual rainfall totals (rainavgyear) (inches)	33.16	48.74	41.10	2.28
Average annual maximum temperatures (tempmaxavg) (°F)	57.00	69.00	62.43	1.99
Average annual minimum temperatures (tempminavg) (°F)	37.00	46.00	40.84	1.61
Average annual winter minimum temperatures (tempwintermin) (°F)	17.00	27.00	21.97	1.80
Average annual number of days with a minimum temperature of ≤32f (temp32fdays)	88.00	148.00	122.60	11.48

Table 21. Summary statistics for curved segments in Ohio: census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	54.58	9.04	6.42
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	16.99	17.65
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	43.67	22.50
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	39.34	21.44
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	10.85	86.20	44.79	8.94
Percentage of population ages 25+ with a high school diploma (diploma25plus)	5.00	76.36	43.06	8.69
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	65.41	12.16	7.35
Percentage of households with income <$50,000 (income50k)	6.02	95.77	55.00	12.33
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	74.37	32.24	9.83
Percentage of households with income >$100,000 (income100kplus)	0.00	69.78	12.76	7.94
Percentage of households with 0 vehicles (noveh)	0.00	88.11	5.62	6.72
Percentage of households with 1 vehicle (X1veh)	0.00	69.94	23.82	10.11
Percentage of households with ≥2 vehicles (X2vehplus)	5.41	96.38	70.56	11.34
Percentage of population ages 15–19 (age15to19)	0.00	32.10	6.42	3.31
Percentage of population ages 20–44 (age20to44)	1.28	79.23	27.38	6.18
Percentage of population ages 45–64 (age45to64)	2.97	59.86	31.92	7.97
Percentage of population ages 65–74 (age65to74)	0.00	33.27	9.65	4.17
Percentage of population ages 75+ (age75plus)	0.00	58.01	6.68	4.01

Table 22. Summary statistics for tangent segments in Ohio: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Segment length (seg_lngth) (mi)	0.02	10.82	0.275	0.581
Percent grade (pct_grad)	0.00	20.00	2.95	3.99
Average AADT (avg_aadt) (veh/day)	90.00	16,623.00	2,267.58	2,112.64
Speed limit (spd_limt) (mph)	20.00	65.00	51.34	7.35
Shoulder width (shldwid) (ft)	0.00	16.00	2.96	2.16

Table 23. Summary statistics for segments in Ohio: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	4.70	109.90	25.20	15.61
Average annual rainfall totals (rainavgyear) (inches)	31.48	48.74	40.39	2.37
Average annual maximum temperatures (tempmaxavg) (°F)	57.00	69.00	61.51	2.16
Average annual minimum temperatures (tempminavg) (°F)	37.00	46.00	40.58	1.77
Average annual winter minimum temperatures (tempwintermin) (°F)	17.00	27.00	21.38	1.99
Average annual number of days with a minimum temperature of ≤32°F (temp32fdays)	88.00	148.00	123.84	12.26

Table 24. Summary statistics for segments in Ohio: census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	54.66	8.66	6.19
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	16.65	15.78
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	45.84	21.09
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	37.51	20.30
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	10.85	86.20	44.50	9.13
Percentage of population ages 25+ with a high school diploma (diploma25plus)	2.75	76.81	41.42	9.03
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	82.82	14.09	8.64
Percentage of households with income <$50,000 (income50k)	3.68	100.00	50.77	13.69
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	74.37	34.14	9.95
Percentage of households with income >$100,000 (income100kplus)	0.00	78.45	15.08	9.41
Percentage of households with 0 vehicles (noveh)	0.00	88.11	5.51	7.38
Percentage of households with 1 vehicle (X1veh)	0.00	78.43	23.38	9.87
Percentage of households with ≥2 vehicles (X2vehplus)	5.41	98.16	71.11	11.85
Percentage of population ages 15–19 (age15to19)	0.00	51.54	6.66	3.28
Percentage of population ages 20–44 (age20to44)	0.56	91.47	27.68	6.09
Percentage of population ages 45–64 (age45to64)	1.68	59.86	30.83	7.42
Percentage of population ages 65–74 (age65to74)	0.00	33.27	9.49	4.08
Percentage of population ages 75+ (age75plus)	0.00	89.47	6.65	3.86

Table 25. Summary statistics for intersection (all types) in Ohio: roadway inventory data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Major AADT (MajorAADT) (veh/day)	89.00	96,948.00	5,055.14	6,773.42
Minor AADT (minorAADT) (veh/day)	100.00	84,905.00	1,221.91	2,603.86
Median width (medianWidth) (ft)	0.00	201.00	1.57	9.23
Speed limit (postedSpeed) (mph)	20.00	70.00	47.57	10.23
Number of approaches with left-turn lanes (leftLanes)	0.00	4.00	0.18	0.72
Number of approaches with right-turn lanes (rightLanes)	0.00	4.00	0.05	0.32

Table 26. Summary statistics for intersection (all types) in Ohio: climate data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Average annual snowfall totals (snowavgyear) (inches)	4.70	109.90	26.19	15.67
Average annual rainfall totals (rainavgyear) (inches)	31.48	48.74	39.60	2.66
Average annual maximum temperatures (tempmaxavg) (°F)	57.00	69.00	61.10	2.15
Average annual minimum temperatures (tempminavg) (°F)	37.00	46.00	40.91	1.88
Average annual winter minimum temperatures (tempwintermin) (°F)	17.00	27.00	21.41	2.03
Average annual number of days with a minimum temperature of ≤32°F (temp32fdays)	88.00	148.00	121.70	12.74

Table 27. Summary statistics for intersections (all types) in Ohio: Census data.

Data Element (Variable Name)	Minimum	Maximum	Average	Standard Deviation
Percentage of population ages 16+ unemployed (unempl16plus)	0.00	83.87	8.73	6.69
Percentage of population ages 16–24 working full time (workftage16to24)	0.00	100.00	17.27	16.05
Percentage of population ages 16–24 working part time (workptage16to24)	0.00	100.00	46.98	21.01
Percentage of population ages 16–24 unemployed (noworkage16to24)	0.00	100.00	35.75	20.27
Percentage of population ages 25+ without a high school diploma (noedctn25plus)	0.00	100.00	44.14	9.71
Percentage of population ages 25+ with a high school diploma (diploma25plus)	0.00	76.81	39.67	10.27
Percentage of population ages 25+ with a university degree (univ25plus)	0.00	100.00	16.19	11.31
Percentage of households with income <$50,000 (income50k)	0.00	100.00	50.49	15.53
Percentage of households with income between $50,000 and $100,000 (income50to100k)	0.00	100.00	33.64	10.59
Percentage of households with income >$100,000 (income100kplus)	0.00	96.66	15.87	10.84
Percentage of households with 0 vehicles (noveh)	0.00	93.25	6.05	8.48
Percentage of households with 1 vehicle (X1veh)	0.00	100.00	25.82	11.82
Percentage of households with ≥2 vehicles (X2vehplus)	0.00	100.00	68.12	15.48
Percentage of population ages 15–19 (age15to19)	0.00	83.50	6.65	3.59
Percentage of population ages 20–44 (age20to44)	1.28	94.01	28.50	7.11
Percentage of population ages 45–64 (age45to64)	0.00	68.23	30.10	7.33
Percentage of population ages 65–74 (age65to74)	0.00	33.27	9.10	4.18
Percentage of population ages 75+ (age75plus)	0.00	58.01	6.87	4.25

ANALYSIS METHODOLOGY

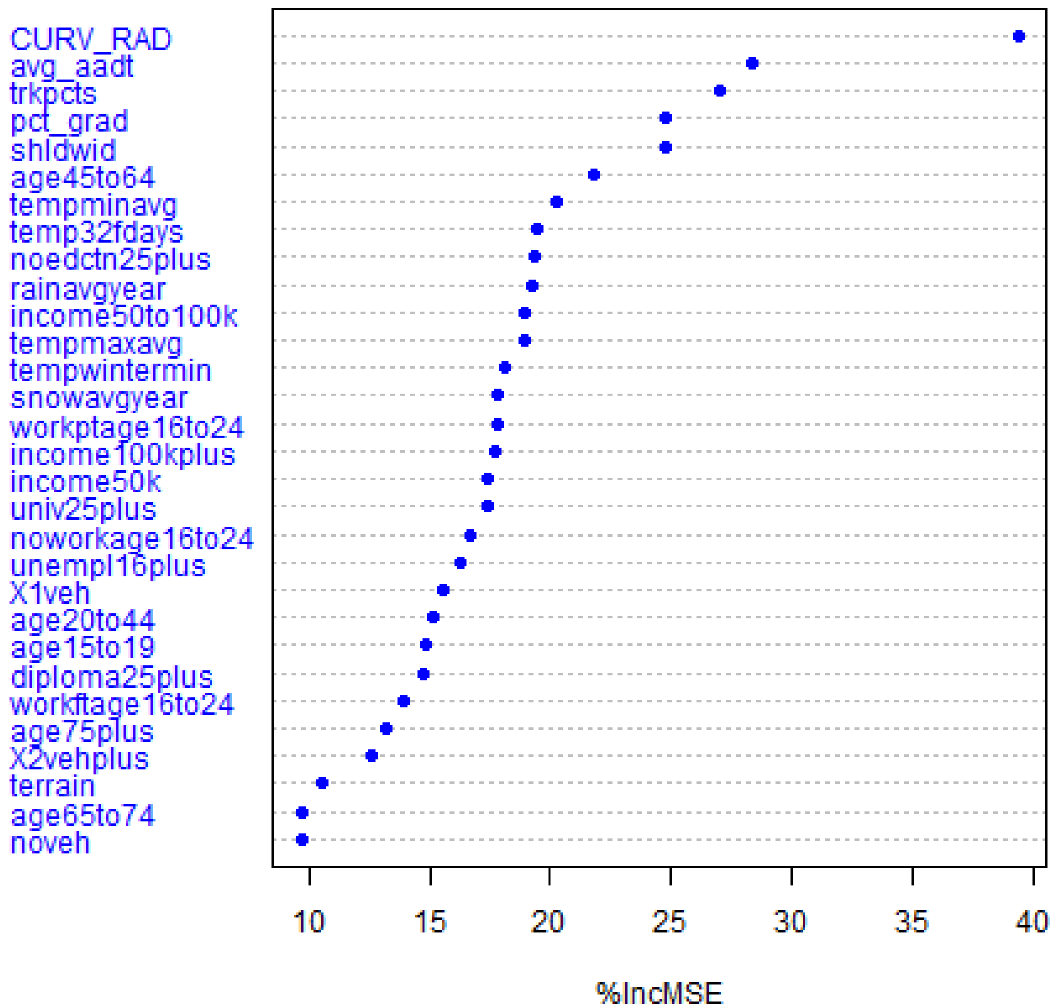
The research team used random forests to identify contributing factors corresponding to selected FCFTs. As noted in chapter 1, the *Highway Safety Statistical Paper Synthesis* demonstrated CART and random forests within the context of conducting statistical road-safety analyses (Persaud et al. 2001). The study concluded that tree-based models hold strong potential for road-safety analyses. This current project extends the exploration of potential applications of tree-based methods within the context of identifying contributing factors for systemic safety analysis.

Tree-based methods are a set of machine-learning and data-mining procedures. They use the form of a binary tree and act as predictive models that map values of a dependent variable or response variable (e.g., crash frequency) as a function of key explanatory variables (e.g., roadway, weather, sociodemographic characteristics). There are two types of tree analyses: a classification tree where the dependent variable is categorical and a regression tree where the dependent variable is continuous. The output of these analyses is a tree that shows the most predictive variable at the top that branches off into combinations of variables that best predict the outcome variable. A tree can be useful to determine contributing factors associated with different crash types and provide insights into interactions between contributing factors.

Breiman and Cutler (2013) developed the random-forest algorithm, which works within the framework of CART. With random forests, instead of having one tree, multiple trees are produced using a resampling method, and the aggregate results are then combined. Breiman and Cutler believed that a single decision tree may not reveal all variables that contribute to the dependent/target variable and that the contributions of some predictive independent variables can be masked by other independent variables. Random forests can help identify predictors that may not appear in the output of a single classification or regression tree but, nevertheless, are highly related to the target variable. The percentage increase in mean squared error (MSE) with the removal of a variable from the random-forest model are commonly displayed using random forests. The random forests do not directly indicate if the presence of variables correspond to increases or decreases in expected crash frequency. However, plots of random forest–predicted crash frequencies as a function of variables provide the information needed to identify contributing factors and inform countermeasure identification.

Figure 1 is an example of a random forest developed for crash frequency per mile of ROR-D crashes on Washington rural two-lane roads on horizontal curves. The vertical axis displays variables, and the horizontal axis shows percent increase in MSE (%IncMSE). Higher values of %IncMSE imply that a variable is a stronger predictor of crash frequency per mile. Typically, the strongest predictors are shown at the top of the plot. For example, figure 1 illustrates that curve radius is the strongest predictor of crash frequency per mile of ROR-D crashes on Washington rural two-lane roads on horizontal curves. Taking curve radius out of the analysis can potentially increase the MSE of predictions by approximately 40 percent.

A sample random-forest code used in R software is provided in appendix D.

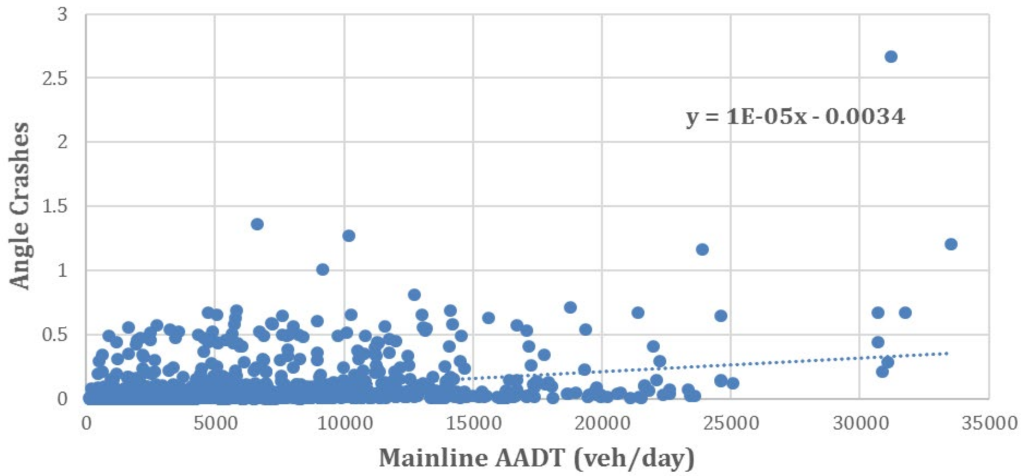


Source: FHWA.

Figure 1. Graph. Crash frequency per mile of ROR-D crashes on rural two-lane roads on horizontal curves: Washington.

After identifying the most predictive variables from the random forest, the research team identified the trend for each variable and created plots of random forest–predicted crash frequencies as a function of the predictor variables.

Figure 2 shows an example of such a plot for the random forest–predicted ANG crash frequency as a function of mainline AADT for three-leg stop-controlled intersections on rural two-lane roads in California. In this case, the plot shows mainline AADT as a contributing factor associated with increases in predicted crash frequency (i.e., the positive slope of the linear best-fit line indicating that an increase in mainline AADT is associated with more ANG crashes at three-leg stop-controlled intersections on rural two-lane roads). Similarly, figure 3 shows an example of such a plot for the random forest–predicted ANG crash frequency at four-leg stop-controlled intersections on urban two-lane roads in California as a function of percentage of households with income <$50,000 within the census block containing the intersection. In this case, the plot shows the percentage of households as a factor that decreases predicted crash frequency (i.e., the negative slope of the linear best-fit line indicating that an increase in the percentage of households with income <$50,000 in the census block containing the intersection is associated with fewer ANG crashes).



Source: FHWA.

Figure 2. Graph. ANG crash frequency versus mainline AADT for three-leg stop-controlled intersections on rural two-lane roads: California.

Source: FHWA.

Figure 3. Graph. ANG crash frequency versus percentage of households with income <$50,000 for four-leg stop-controlled intersections on urban two-lane roads: California.

RESULTS

The following sections of this chapter include tables that list the top 10 to 12 most influential predictor variables for each intersection and nonintersection FCFT according to the generated random forests. These top 10 to 12 variables for each FCFT serve as the focus of discussion of contributing factors for the respective FCFT. The tables also identify whether each variable contributes to an increase or decrease in the expected number of crashes. Appendix E provides the random-forest outputs corresponding to the analyses of all FCFTs. Appendix F provides plots of random forest–predicted crash frequencies as a function of selected predictor variables.

There are some counterintuitive results reported in the contributing-factor tables. For example, a roadway variable, in some cases, shows up as a contributing factor to an increase in crashes, whereas prior knowledge has indicated the variable to contribute to a decrease in crashes (e.g., wider shoulder width shows up as contributing to an increase in LNDP crashes on tangent segments in Washington). Such occurrences of counterintuitive findings for roadway variables have been denoted with two asterisks and are discussed in more detail in the following sections of this chapter. In general, the research team felt it useful to report the counterintuitive findings for the following reasons:

- There is a much broader discussion occurring at the national level about how to address counterintuitive findings, particularly when so little is known about the adaptive behavior of drivers (e.g., countermeasures that are thought to increase crashes could induce more careful driving and countermeasures that are thought to reduce crashes could induce higher-speed driving).

- There are relatively few examples of crash-data analysis at the level of crash-type disaggregation in this report, and therefore, there is not an abundance of existing knowledge on which to base a determination of intuitive versus counterintuitive findings for certain crash types. It is possible that some of the counterintuitive findings from this analysis could set the stage for future investigation, possibly using naturalistic-driving data collected as part of SHRP2.

In a limited number of cases, roadway variables presented in the following tables did not show up among the top 10 most influential predictor variables for that particular FCFT but were worth reporting in the tables for two reasons. The first reason the variables were included was because it made sense from an engineering perspective and the findings (e.g., smaller radius of curve increases the expected frequency of ROR crashes) are likely of interest to State and local agencies that focus on infrastructure improvements. The second reason the variables were included was because their removal would still lead to an increase in MSE of model predictions as derived from random-forest outputs even if the variable fell outside the top 10 to 12 predictor variables (i.e., the variable still appeared predictive even though it fell outside of the top 10 to 12 most predictive); when this occurs, the variables are included in the bottom rows of the tables below the 10 rows that identify the top most influential predictor variables.

For the nonintersection FCFT contributing-factor analysis, Washington data were separated into curved and tangent segments for the analysis and Ohio data for curved and tangent segments were analyzed together[4] in a single file.

The contributing-factor analyses of California and Washington data were based on the expected number of KAB crashes. The contributing-factor analysis of Ohio data was based on both KAB crashes and the expected number of KABCO crashes. This latter approach for Ohio allowed the research team to assess whether inclusion of only certain levels of severity (i.e., KAB versus KABCO) influenced contributing-factor results.

[4]The main reason behind combining the Ohio data was because ODOT defines any curve with a radius >1,800 ft as tangent sections, which limited the curve data compared to Washington, where radii ranged from 100 to 12,000 ft.

ROR CRASHES ON RURAL TWO-LANE HIGHWAY SEGMENTS

Ohio

Horizontal Curves and Highway Tangent Segments—Daytime

Table 28 and table 29 summarize the most influential predictor variables for the expected number of ROR-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROR-KABCO-D) according to random forests generated using Ohio data.

Table 28. Contributing factors for ROR-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Percent grade	Increases
Surface width	Decreases
Average shoulder width	Decreases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 20–44	Increases
Percentage of households with 0 vehicles	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 75+	Decreases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 29. Contributing factors for ROR-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Surface width	Decreases
Average shoulder width	Decreases
Percent grade	Increases
Speed limit	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 75+	Decreases
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves and Highway Tangent Segments—Nighttime

Table 30 and table 31 summarize the most influential predictor variables for the expected number of ROR-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROR-KABCO-N) according to random forests generated using Ohio data.

Table 30. Contributing factors for ROR-KAB-N crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Average shoulder width	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 45–64	Decreases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 75+	Decreases
Percentage of population ages 15–19	Increases
Percentage of population ages 20–44	Increases
Surface width	Decreases
Percent grade	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 31. Contributing factors for ROR-KABCO-N crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Average shoulder width	Decreases
Curve radius	Increases/decreases*
Speed limit	Increases
Percentage of population ages 16–24 working full time	Decreases
Surface width	Decreases
Percentage of population ages 75+	Decreases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Washington

Highway Tangent Segments—Daytime

Table 32 summarizes the most influential predictor variables for the expected number of ROR-D crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-D) according to random forests generated using Washington data. The results in table 32 offer an opportunity to revisit key considerations for interpreting the results in this table, as well as subsequent tables:

- Contributing factors are factors whose presence is associated with either increases or decreases in the expected frequencies of crashes according to the generated random forests.
- Findings regarding contributing factors are interpreted in this project as predictive relationships or statistical associations with expected crash frequencies. The report does not cover the philosophy of causation or causal inference. That said, the following characteristics raise confidence that a specific finding or set of findings are stable and transferable:
    - Consistency across subsets of related FCFTs.
    - Consistency across multiple States.
    - Consistency with previous findings in the literature.

Knowledge related to the safety impacts of traffic and roadway variables has grown substantially over the last two decades and offers a basis to interpret the results of this effort. There is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results of this effort. Findings related to socioeconomic variables are likely representative of differences in travel behavior, driving behavior, and driving capabilities that seem key for safety analyses but are generally not incorporated into analyses that also include traffic and roadway factors. Weather-related findings are likely representative of differences in visibility, road conditions, and driver experience and behavior. All findings are reported and discussed throughout the remainder of this report. In most cases, findings related to socioeconomic and weather variables set the stage for future analyses, possibly focused solely on these variables.

Table 32. Contributing factors for ROR-KAB-D crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Average AADT	Increases
Percentage of trucks on the roadway	Decreases
Average shoulder width	Decreases
Average annual rainfall total	Increases
Percentage of population ages 20–44	Increases
Percentage of population ages 16–24 unemployed	Increases
Annual average maximum temperature	Decreases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 25+ without a high school diploma	Decreases

Highway Tangent Segments—Nighttime

Table 33 summarizes the most influential predictor variables for the expected number of ROR-N crashes on rural two-lane tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-N) according to random forests generated using Washington data.

Table 33. Contributing factors for ROR-KAB-N crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Average shoulder width	Decreases
Percentage of trucks on the roadway	Decreases
Percent grade	Decreases**
Average AADT	Increases
Unpaved shoulders	Increases
Percentage of population ages 16+ unemployed	Increases
Annual average winter minimum temperature	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 65–74	Decreases
Annual average maximum temperature	Decreases

**Counterintuitive finding.

Horizontal Curves—Daytime

Table 34 summarizes the most influential predictor variables for the expected number of ROR-D crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-D) according to random forests generated using Washington data.

Table 34. Contributing factors for ROR-KAB-D crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Increases
Percentage of trucks on the roadway	Decreases
Percent grade	Increases
Average shoulder width	Decreases
Percentage of population ages 45–64	Increases
Annual average minimum temperature	Decreases
Annual Average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 25+ without a high school diploma	Decreases
Average annual rainfall total	Increases

*Increases crash frequency when comparing curves to tangents/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves—Nighttime

Table 35 summarizes the most influential predictor variables for the expected number of ROR-N crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-N) according to random forests generated using Washington data.

Table 35. Contributing factors for ROR-KAB-N crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Increases
Average shoulder width	Decreases
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 16–24 working part time	Decreases
Percentage of households with more than 2 vehicles	Increases
Percent grade	Decreases**
Percentage of households with 1 vehicle	Decreases
Percentage of trucks on the roadway	Decreases
Percentage of population ages 25+ without a high school diploma	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ROR crashes on rural two-lane horizontal curves and tangent segments:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of ROR-D and ROR-N crashes on curves and tangent segments.
- *Average shoulder width*: an increase in shoulder width was consistently associated with a decrease in the frequency of ROR-D and ROR-N crashes (and therefore a decrease in shoulder width was consistently associated with an increase in the frequency of ROR crashes).
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of ROR-D and ROR-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of ROR crashes).

There were several additional roadway variables that showed relationships with ROR crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Percent grade*: an increase in percent grade was associated with an increase in the frequency of ROR-D crashes in Washington and ROR-D and ROR-N crashes in Ohio; the relationship was the opposite for Washington during the nighttime.
- *Surface width*: an increase in surface width was associated with a decrease in the frequency of ROR-D and ROR-N crashes on curves and tangent segments (and therefore

a decrease in surface width was consistently associated with an increase in the frequency of ROR crashes); this variable was only available for Ohio.

- *Unpaved shoulders or no shoulders*: the presence of one or more unpaved shoulders or no shoulders showed up as a contributing factor that increases crash frequency when compared to paved shoulders but only for ROR-KAB-N crashes on tangent segments in Washington.
- *Speed limit*: an increase in speed limit was associated with an increase in the frequency of ROR-D and ROR-N crashes in Ohio.

With respect to sociodemographic characteristics, three variables appeared as contributing factors in both Ohio and Washington:

- The percentage of the population ages 16+ that are unemployed appeared to increase the frequency of ROR-KAB-D crashes on curves and tangent segments in Ohio and ROR-KAB-N crashes on tangent segments in Washington.
- The percentage of the population ages 20–44 appeared to increase ROR-KAB-D and ROR-KAB-N crashes on curves and tangent segments in Ohio and ROR-KAB-D crashes on tangent segments in Washington.
- The percentage of the population ages 25+ with a high school diploma but no university degree appeared to decrease ROR-KAB-D crashes for all crash types in Ohio and on tangent segments in Washington.

Other socioeconomic characteristics of note included the following:

- The percentage of the population ages 15–19 appeared to increase ROR-KAB-N crashes on curves and tangent segments in Ohio.
- The percentage of the population ages 75+ appeared to decrease ROR-D and ROR-N crashes on curves and tangent segments in Ohio.
- The percentage of households with an annual income >$100,000 also appeared to increase crashes for two (i.e., ROR-KABCO-D and ROR-KAB-N) of the four crash types in Ohio, possibly serving as a surrogate for travel amount or exposure.

With respect to weather characteristics, average annual rainfall total appeared to increase ROR-KAB-D crashes on curves and tangent segments in Washington, but data limitations prevented the validation of this finding in Ohio.

LNDP CRASHES ON RURAL TWO-LANE HIGHWAY SEGMENTS

Ohio

Horizontal Curves and Highway Tangent Segments—Daytime

Table 36 and table 37 summarize the most influential predictor variables for the expected number of LNDP-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (LNDP-KABCO-D) according to random forests generated using Ohio data.

Table 36. Contributing factors for LNDP-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Percent grade	Increases
Surface width	Decreases
Average shoulder width	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of households with income <$50,000	Decreases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of population ages 16+ unemployed	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 37. Contributing factors for LNDP-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Increases
Average shoulder width	Decreases
Surface width	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 16+ unemployed	Increases
Speed limit	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 75+	Decreases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves and Highway Tangent Segments—Nighttime

Table 38 and table 39 summarize the most influential predictor variables for the expected number of LNDP-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (LNDP-KABCO-N) according to random forests generated using Ohio data.

Table 38. Contributing factors for LNDP-KAB-N crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Average shoulder width	Decreases
Percentage of households with income >$100,000	Increases
Percentage of households with income <$50,000	Decreases
Percentage of population ages 16–24 working full time	Decreases
Percentage of population ages 45–64	Decreases
Percent grade	Increases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 20–44	Increases
Surface width	Decreases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 39. Contributing factors for LNDP-KABCO-N crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Surface width	Decreases
Curve radius	Increases/decreases*
Speed limit	Increases
Average shoulder width	Decreases
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 75+	Decreases
Percentage of population ages 65–74	Decreases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Washington

Highway Tangent Segments—Daytime

Table 40 summarizes the most influential predictor variables for the expected number of LNDP-D crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-D) according to random forests generated using Washington data.

Table 40. Contributing factors for LNDP-KAB-D crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Average shoulder width	Increases**
Percentage of trucks on the roadway	Decreases
Average annual rainfall total	Increases
Percentage of population ages 16–24 unemployed	Increases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 25+ with high school diploma but no university degree	Decreases
Percentage of households with 0 vehicles	Decreases
Average annual snowfall total	Increases
Lane width	Increases**

**Counterintuitive finding.

Highway Tangent Segments—Nighttime

Table 41 summarizes the most influential predictor variables for the expected number of LNDP-N crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-N) according to random forests generated using Washington data.

Table 41. Contributing factors for LNDP-KAB-N crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Decreases**
Average shoulder width	Increases**
Percentage of trucks on the roadway	Decreases
Average AADT	Increases
Unpaved shoulders	Increases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 65–74	Decreases
Annual average winter minimum temperature	Decreases
Percentage of population ages 25+ without a high school diploma	Increases
Annual average maximum temperature	Decreases

**Counterintuitive finding.

Horizontal Curves—Daytime

Table 42 summarizes the most influential predictor variables for the expected number of LNDP-D crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-D) according to random forests generated using Washington data.

Table 42. Contributing factors for LNDP-KAB-D crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Increases
Percent grade	Increases
Percentage of trucks on the roadway	Decreases
Average shoulder width	Decreases
Average annual rainfall total	Increases
Percentage of population ages 45–64	Increases
Annual average maximum temperature	Decreases
Percentage of households with income <$50,000	Decreases
Percentage of population ages 25+ with high school diploma but no university degree	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves—Nighttime

Table 43 summarizes the most influential predictor variables for the expected number of LNDP-N crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-N) according to random forests generated using Washington data.

Table 43. Contributing factors for LNDP-KAB-N crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Increases
Percentage of trucks on the roadway	Decreases
Percentage of households with 1 vehicle	Decreases
Annual average maximum temperature	Increases
Percentage of population ages 16+ unemployed	Decreases
Average shoulder width	Increases**
Percentage of households with ≥2 vehicles	Increases
Percentage of households with 0 vehicles	Decreases
Percentage of population ages 65–74	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with LNDP crashes on rural two-lane horizontal curves and tangent segments:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of LNDP-D and LNDP-N crashes on curves and tangent segments.
- *Percent grade*: an increase in percent grade was associated with an increase in the frequency of LNDP-D and LNDP-N crashes in Ohio and LNDP-D crashes in Washington.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of LNDP-D and LNDP-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of LNDP crashes).

There were several additional roadway variables that showed relationships with LNDP crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Surface width*: an increase in surface width was associated with a decrease in the frequency of LNDP-D and LNDP-N crashes on curves and tangent segments (and therefore a decrease in surface width was consistently associated with an increase in the frequency of LNDP crashes); this variable was only available for Ohio.
- *Average shoulder width*: an increase in shoulder width was associated with a decrease in the frequency of LNDP-D and LNDP-N crashes on horizontal curves in Ohio and LNDP-D crashes on horizontal curves in Washington (and therefore a decrease in shoulder width was associated with an increase in the frequency of LNDP crashes); the relationship was opposite for LNDP-N crashes on horizontal and LNDP-D and LNDP-N crashes on tangent segments in Washington.
- *Unpaved shoulders or no shoulders*: the presence of one or more unpaved shoulders or no shoulders was associated with an increase in crashes when compared to paved shoulders but only for LNDP-KAB-N crashes on tangent segments in Washington.
- *Speed limit*: an increase in speed limit was associated with an increase in the frequency of LNDP-D and LNDP-N crashes in Ohio.

With respect to sociodemographic characteristics, the percentage of the population ages 16+ that is unemployed in Ohio and Washington appeared to increase LNDP-KAB-D crashes (OH curves and tangent segments and WA tangent segments), LNDP-KAB-N crashes (OH curves and tangent segments and WA tangent segments), and LNDP-KABCO-D crashes (OH curves and tangent segments). As an exception to this trend, the percentage of the population ages 16+ that is unemployed appeared to reduce LNDP-KABCO-N crashes on horizontal curves and tangent segments in Ohio and LNDP-KAB-N crashes on horizontal curves in Washington. The percentage of the population ages 25+ with a high school diploma but no university degree appeared to decrease LNDP-KABCO-D crashes on curves and tangent segments in Ohio and LNDP-KAB-D crashes on curves and tangent segments in Washington.

The following also appeared with respect to LNDP crashes in Ohio and Washington:

- The percentage of households with an annual income <$50,000 appeared to decrease LNDP-KAB-D and LNDP-KAB-N crashes on curves and tangent segments in Ohio and LNDP-KAB-D crashes on horizontal curves in Washington.
- The percentage of households with an annual income >$100,000 appeared to increase LNDP crashes in Ohio.

With respect to weather characteristics, average annual rainfall total appeared to increase LNDP-KAB-D crashes on horizontal curves and tangent segments in Washington.

HEO CRASHES ON RURAL TWO-LANE HIGHWAY SEGMENTS

Ohio

Horizontal Curves and Highway Tangent Segments—Daytime

Table 44 and table 45 summarize the most influential predictor variables for the expected number of HEO-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (HEO-KABCO-D) according to random forests generated using Ohio data.

Table 44. Contributing factors for HEO-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Average AADT	Increases
Curve radius	Increases/decreases*
Percentage of households with income >$100,000	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of households with 1 vehicle	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 20–44	Increases
Percentage of population ages 25+ without a high school diploma	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 45. Contributing factors for HEO-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Curve radius	Increases/decreases*
Percentage of households with income >$100,000	Increases
Percentage of households with income <$50,000	Decreases
Percentage of households with 1 vehicle	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 16–24 unemployed	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves and Highway Tangent Segments—Nighttime

Table 46 and table 47 summarize the most influential predictor variables for the expected number of HEO-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (HEO-KABCO-N) according to random forests generated using Ohio data.

Table 46. Contributing factors for HEO-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Average AADT	Increases
Surface width	Increases
Percentage of population ages 15–19	Increases
Curve radius	Increases/decreases*
Percentage of population ages 75+	Increases
Percentage of population ages 65–74	Decreases
Percentage of population ages 25+ without high school diploma	Decreases
Percentage of population ages 16–24 working part time	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 47. Contributing factors for HEO-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Surface width	Increases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of population ages 16–24 working full time	Decreases
Percentage of population ages 15–19	Increases
Curve radius	Increases/decreases*
Percentage of population ages 75+	Decreases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 16–24 unemployed	Increases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Washington

Highway Tangent Segments—Daytime

Table 44 summarizes the most influential predictor variables for the expected number of HEO-D crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-D) according to random forests generated using Washington data.

Table 48. Contributing factors for HEO-KAB-D crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Percentage of trucks on the roadway	Decreases
Annual average maximum temperature	Decreases
Average annual rainfall total	Increases
Average AADT	Increases
Percent grade	Decreases**
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Unpaved shoulders	Increases
Percentage of households with 1 vehicle	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of population ages 16–24 working part time	Decreases

**Counterintuitive finding.

Highway Tangent Segments—Nighttime

Table 49 summarizes the most influential predictor variables for the expected number of HEO-N crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-N) according to random forests generated using Washington data.

Table 49. Contributing factors for HEO-KAB-N crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Decreases**
Annual average maximum temperature	Increases
Average AADT	Increases
Percentage of households with income <$50,000	Decreases
Annual average winter minimum temperature	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of population ages 20–44	Increases
Annual average minimum temperature	Decreases
Percentage of households with income between $50,000 and $100,000	Increases

**Counterintuitive finding.

Horizontal Curves—Daytime

Table 50 summarizes the most influential predictor variables for the expected number of HEO-D crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-D) according to random forests generated using Washington data.

Table 50. Contributing factors for HEO-KAB-D crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Percentage of trucks on the roadway	Decreases
Percent grade	Decreases**
Annual average minimum temperature	Decreases
Percentage of population ages 16–24 working part time	Decreases
Percentage of population ages 15–19	Increases
Percentage of households with income >$100,000	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Horizontal Curves—Nighttime

Table 51 summarizes the most influential predictor variables for the expected number of HEO-N crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-N) according to random forests generated using Washington data.

Table 51. Contributing factors for HEO-KAB-N crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Percentage of population ages 15–19	Increases
Percentage of trucks on the roadway	Decreases
Average AADT	Increases
Percentage of population ages 25+ with a high school diploma but not university degree	Decreases
Average shoulder width	Increases**
Percentage of population ages 16–24 working part time	Decreases
Percentage of households with 1 vehicle	Decreases
Percentage of households with income >$100,000	Increases
Annual average maximum temperature	Decreases
Percentage of population ages 65–74	Decreases

**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with HEO crashes on rural two-lane horizontal curves and highway tangent segments:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of HEO-D and HEO-N crashes on curves and highway tangent segments.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of HEO-D and HEO-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of HEO crashes).

There were several additional roadway variables that showed relationships with HEO crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Surface width*: an increase in surface width was associated with a decrease in the frequency of HEO-N crashes on curves and highway tangent segments (and therefore a decrease in surface width was consistently associated with an increase in the frequency of HEO-N crashes); this variable was only available for Ohio.
- *Unpaved shoulders*: unpaved shoulders appeared to increase HEO-KAB-D crashes on highway tangent segments in Washington.
- *Percent grade*: an increase in percent grade was associated with an increase in the frequency of HEO-D and HEO-N crashes on curves and highway tangent segments in Ohio (counterintuitively, an increase in percent grade was associated with a decrease in HEO crashes in Washington).
- *Speed limit*: an increase in speed limit was associated with an increase in HEO-KAB-N crashes on curves and highway tangent segments in Ohio.

With respect to sociodemographic characteristics, several factors showed a consistent relationship to HEO crashes in Ohio and Washington. The percentage of the population ages 15–19 appeared to increase HEO-KAB-N and HEO-KABCO-N crashes on curves and highway tangent segments in Ohio and HEO-KAB-D and HEO-KAB-N crashes on horizontal curves in Washington. The percentage of the population ages 20–24 appeared to increase HEO-KAB-D crashes on curves and highway tangent segments in Ohio and HEO-KAB-N crashes on highway tangent segments in Washington.

The following also appeared with respect to HEO crashes in Ohio and Washington:

- Percentage of the population ages 25+ with a high school diploma but no university degree appeared to decrease HEO crashes (consistent with findings for ROR and LNDP crashes).
- Percentage of households with income <$50,000 appeared to decrease HEO crashes (consistent with findings for LNDP crashes).
- Percentage of households with income between $50,000 and $100,000 appeared to increase HEO crashes.
- Percentage of households with income >$100,000 appeared to increase HEO crashes (consistent with findings for ROR and LNDP crashes).

The percentage of the population ages 25+ with a university degree appeared to increase HEO-D crashes in Ohio. The percentage of households with two or more vehicles appeared to decrease HEO-D crashes on curves and highway tangent segments in Ohio and highway tangent segments in Washington.

With respect to weather characteristics, the average annual rainfall appeared to increase HEO-KAB-D crashes on highway tangent segments in Washington.

ROLL CRASHES ON RURAL TWO-LANE HIGHWAY SEGMENTS

Ohio

Horizontal Curves and Highway Tangent Segments—Daytime

Table 52 and table 53 summarize the most influential predictor variables for the expected number of ROLL-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROLL-KABCO-D) according to random forests generated using Ohio data.

Table 52. Contributing factors for ROLL-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Decreases**
Surface width	Decreases
Percentage of population ages 45–64	Increases
Average shoulder width	Decreases
Percentage of population ages 16+ unemployed	Increases
Percentage of population ages 16–24 working full time	Decreases
Percentage of population ages 20–44	Increases
Percentage of population ages 65–74	Decreases
Percentage of households with income >$100,000	Decreases
Speed limit	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Table 53. Contributing factors for ROLL-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average shoulder width	Decreases
Percentage of households with 0 vehicles	Decreases
Percentage of population ages 20–44	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 45–64	Increases
Surface width	Decreases
Average AADT	Decreases
Percentage of population ages 16+ unemployed	Increases
Percentage of households with 1 vehicle	Increases
Speed limit	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Horizontal Curves and Highway Tangent Segments—Nighttime

Table 54 and table 55 summarize the most influential predictor variables for the expected number of ROLL-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROLL-KABCO-N) according to random forests generated using Ohio data.

Table 54. Contributing factors for ROLL-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Percentage of population ages 65–74	Decreases
Average AADT	Increases
Curve radius	Increases/decreases*
Percentage of population ages 75+	Decreases
Percentage of population ages 15–19	Increases
Percentage of population ages 25+ without a high school diploma	Decreases
Percentage of population ages 45–64	Decreases
Percentage of population ages 16–24 working full time	Decreases
Percentage of population ages 25+ with a high school diploma but no university degree	Increases
Surface width	Increases
Shoulder width	Increases**

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Table 55. Contributing factors for ROLL-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Surface width	Decreases
Percentage of population ages 75+	Increases
Percentage of population ages 45–64	Decreases
Curve radius	Increases/decreases*
Percentage of population ages 65–74	Decreases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of population ages 15–19	Increases
Percentage of households with income <$50,000	Decreases
Speed limit	Increases
Average shoulder width	Decreases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Washington

Highway Tangent Segments—Daytime

Table 56 summarizes the most influential predictor variables for the expected number of ROLL-D crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-D) according to random forests generated using Washington data.

Table 56. Contributing factors for ROLL-KAB-D crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Average AADT	Decreases**
Average shoulder width	Increases**
Percentage of trucks on the roadway	Decreases
Percentage of population ages 16+ unemployed	Increases
Annual average minimum temperature	Decreases
Terrain	Increases
Average annual rainfall total	Increases
Percentage of households with income >$100,000	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases

**Counterintuitive finding.

Highway Tangent Segments—Nighttime

Table 57 summarizes the most influential predictor variables for the expected number of ROLL-N crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-N) according to random forests generated using Washington data.

Table 57. Contributing factors for ROLL-KAB-N crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percentage of trucks on the roadway	Increases
Annual average winter minimum temperature	Increases
Percentage of population ages 16–24 working full time	Increases
Annual average number of days with a minimum temperature of ≤32°F	Increases
Percentage of population ages 16–24 working part time	Decreases
Percentage of households with income between $50,000 and $100,000	Decreases
Percentage of households with 0 vehicles	Decreases
Percentage of population ages 25+ with a university degree	Decreases
Percent grade	Increases

Horizontal Curves—Daytime

Table 58 summarizes the most influential predictor variables for the expected number of ROLL-D crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-D) according to random forests generated using Washington data.

Table 58. Contributing factors for ROLL-KAB-D crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Percentage of population ages 16–24 working part time	Increases
Percentage of population ages 65–74	Increases
Percentage of trucks on the roadway	Decreases
Curve radius	Increases/decreases*
Percentage of population ages 75+	Decreases
Average AADT	Increases
Percent grade	Increases
Percentage of households with income <$50,000	Decreases
Annual average maximum temperature	Decreases
Percentage of population ages 16–24 unemployed	Decreases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Horizontal Curves—Nighttime

Table 59 summarizes the most influential predictor variables for the expected number of ROLL-N crashes on rural two-lane horizontal curves that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-N) according to random forests generated using Washington data.

Table 59. Contributing factors for ROLL-KAB-N crashes on rural two-lane horizontal curves: Washington.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average shoulder width	Increases**
Average AADT	Increases
Truck percentage	Decreases
Percent grade	Increases
Annual average maximum temperature	Increases
Percentage of population ages 25+ with a university degree	Decreases
Average annual rainfall total	Decreases
Percentage of households with income >$100,000	Increases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of population ages 65–74	Increases
Percentage of population ages 75+	Decreases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ROLL crashes on rural two-lane horizontal curves and highway tangent segments:

- *Average AADT*: an increase in AADT was associated with an increase in the frequency of ROLL-N crashes on curves and highway tangent segments in Ohio and Washington; interestingly, AADT appeared to ROLL-D crashes in both Ohio (on curves and highway tangent segments) and Washington (on highway tangent segments).
- *Percent grade*: an increase in percent grade was consistently associated with an increase in the frequency of ROLL-D and ROLL-N crashes on curves and highway tangent segments.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of ROLL-D and ROLL-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of ROLL crashes).

There were several additional roadway variables that showed relationships with ROLL crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Surface width*: an increase in surface width was associated with a decrease in the frequency of ROLL-KAB-D crashes and ROLL-KABCO-D and ROLL-KABCO-N crashes on curves and highway tangent segments (and therefore a decrease in surface width was associated with an increase in the frequency of those types of ROLL crashes); this variable was only available for Ohio.
- *Average shoulder width*: an increase in shoulder width was associated with a decrease in the frequency of ROLL-KAB-D crashes and ROLL-KABCO-D and ROLL-KABCO-N crashes on curves and highway tangent segments (and therefore a decrease in surface width was associated with an increase in the frequency of ROLL crashes); the relationship was the opposite for ROLL-KAB-D crashes on highway tangent segments and ROLL-KAB-N crashes on horizontal curves in Washington.
- *Terrain*: mountainous terrain appeared to increase ROLL-KAB-D crashes on highway tangent segments in Washington.
- *Speed limit*: an increase in speed limit was associated with an increase in the frequency of ROLL-D and ROLL-N crashes in Ohio.

With respect to sociodemographic characteristics, there were several predictors of ROLL crashes in Ohio and Washington:

- The percentage of households with incomes between $50,000 and $100,000 appeared to increase ROLL-KABCO-N crashes on curves and highway tangent segments in Ohio and ROLL-KAB-N crashes on horizontal curves in Washington. It appeared as a factor that decreases the crash frequency of ROLL-KAB-N crashes on highway tangent segments in Ohio.
- The percentage of households with an annual income <$50,000 appeared to decrease ROLL-KABCO-N crashes on curves and highway tangent segments in Ohio and

ROLL-KAB-D crashes on horizontal curves in Washington and (consistent with findings for LNDP and HEO crashes).

- The percentage of households with an annual income >$100,000 appeared to increase ROLL-KAB-D crashes on highway tangent segments and ROLL-KAB-N crashes on horizontal curves in Washington (consistent with findings for ROR, LNDP, and HEO crashes).

With respect to weather characteristics, average annual rainfall total appeared to increase ROLL-KAB-D crashes on highway tangent segments in Washington.

ANG CRASHES ON RURAL TWO-LANE HIGHWAY SEGMENTS

Ohio

Horizontal Curves and Highway Tangent Segments—Daytime

Table 60 and table 61 summarize the most influential predictor variables for the expected number of ANG-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 60. Contributing factors for ANG-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Decreases**
Average AADT	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 25+ without a high school diploma	Increases
Percentage of population ages 45–64	Decreases
Surface width	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 20–44	Increases
Percentage of population ages 75+	Decreases

**Counterintuitive finding.

Table 61. Contributing factors for ANG-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Average shoulder width	Decreases
Percent grade	Decreases**
Percentage of population ages 16+ unemployed	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of households with 0 vehicles	Decreases
Percentage of households with 1 vehicle	Decreases
Percentage of households with ≥2 vehicles	Increases
Surface width	Increases
Speed limit	Decreases**
Curve radius	Increases/decreases*

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Washington

Highway Tangent Segments—Daytime

Table 62 summarizes the most influential predictor variables for the expected number of ANG-D crashes on rural two-lane highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using Washington data.

Table 62. Contributing factors for ANG-KAB-D crashes on rural two-lane highway tangent segments: Washington.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Lane width	Increases
Average shoulder width	Increases**
Percentage of population ages 45–64	Increases
Percentage of population ages 20–44	Decreases
Average annual snowfall total	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Annual average maximum temperature	Increases
Annual average winter minimum temperature	Decreases
Percentage of households with income <$50,000	Decreases

**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes on rural two-lane horizontal curves and highway tangent segments:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of ANG-D crashes on curves and highway tangent segments.

There were several additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Percent grade*: an increase in percent grade was consistently associated with a decrease in the frequency of ANG-D crashes on curves and highway tangent segments in Ohio (and therefore a decrease in percent grade was consistently associated with an increase in the frequency of ANG crashes).
- *Surface width*: an increase in surface width was associated with an increase in the frequency of ANG-D crashes on curves and highway tangent segments (and therefore a decrease in surface width was consistently associated with a decrease in the frequency of ANG-D crashes); this variable was only available for Ohio.
- *Average shoulder width*: an increase in shoulder width was associated with a decrease in the frequency of ANG-KABCO-D crashes on curves and highway tangent segments in Ohio (counterintuitively, an increase in shoulder width was associated with an increase in the frequency of ANG-KAB-D crashes on highway tangent segments in Washington).
- *Curve radius*: an increase in curve radius was associated with a decrease in the frequency of ANG-KABCO-D crashes on curves in Ohio (and therefore a decrease in curve radius was associated with an increase in the frequency of ANG-KABCO-D crashes).

With respect to sociodemographic characteristics, none of the factors showed a consistent relationship to ANG crashes in Ohio or Washington:

- The percentage of the population ages 25+ with a university degree appears to increase ANG-D crashes on curves and highway tangent segments in Ohio.
- The percentage of households with income <$50,000 appears to decrease ANG-D crashes on highway tangent segments in Washington.
- The percentage of households with income >$100,000 appears to increase ANG-D crashes on curves and highway tangent segments in Ohio.
- The percentage of households with no vehicles appears to decrease ANG-D crashes on curves and highway tangent segments in Ohio.
- The percentage of households with one vehicle appears to decrease ANG-D crashes on curves and highway tangent segments in Ohio.
- The percentage of households with two or more vehicles appears to increase ANG-D crashes on curves and highway tangent segments in Ohio.

With respect to weather characteristics, none of the factors showed a consistent relation to ANG crashes in Ohio or Washington.

ANG CRASHES AT FOUR-LEG STOP-CONTROLLED INTERSECTIONS ON RURAL TWO-LANE ROADS

California

Daytime

Table 63 summarizes the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 63. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural two-lane roads: California.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Mainline AADT	Increases
Mainline left channelization	Decreases
Lane width	Increases
Annual average maximum temperature	Decreases
Percentage of population ages 25+ without a high school diploma	Decreases
Average annual rainfall total	Increases
Percentage of households with income >$100,000	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 45–64	Decreases

Nighttime

Table 64 summarizes the most influential predictor variables for the expected number of ANG-N crashes at four-leg stop-controlled intersections (with stop control on the minor road) on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-N) according to random forests generated using California data.

Table 64. Contributing factors for ANG-KAB-N crashes at four-leg stop-controlled intersections on rural two-lane roads: California.

Variable	Impact on Crash-Frequency Predictions
Design speed	Increases
Average annual rainfall total	Increases
Annual average maximum temperature	Decreases
Mainline AADT	Increases
Percentage of population ages 25+ with a university degree	Increases
Percentage of households with income <$50,000	Decreases
Cross street AADT	Increases
Average annual snowfall total	Decreases
Percentage of population ages 45–64	Increases
Annual average minimum temperature	Increases
Mainline left channelization	Decreases

Ohio

Daytime

Table 65 and table 66 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 65. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross Street AADT	Increases
Speed limit	Increases
Mainline AADT	Increases
Percentage of households with income <$50,000	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 16–24 working full time	Increases
Percentage of households with ≥2 vehicles	Decreases
Annual average maximum temperature	Decreases
Percentage of households with income between $50,000 and $100,000	Decreases
Lane width	Increases

Table 66. Contributing factors for ANG-KABCO-D crashes at four-leg stop-controlled intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Speed limit	Increases
Percentage of households with income <$50,000	Decreases
Mainline AADT	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of population ages 16–24 working part time	Increases
Average annual snowfall total	Increases
Percentage of households with ≥2 vehicles	Decreases
Lane width	Increases

Nighttime

Table 67 and table 68 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-N) according to random forests generated using Ohio data.

Table 67. Contributing factors for ANG-KAB-N crashes at four-leg stop-controlled intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Speed limit	Increases
Percentage of households with income <$50,000	Decreases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of households with 1 vehicle	Decreases
Mainline AADT	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of population ages 16–24 working full time	Increases
Percentage of population ages 45–64	Increases

Table 68. Contributing factors for ANG-KABCO-N crashes at four-leg stop-controlled intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Percentage of households with income <$50,000	Decreases
Speed limit	Increases
Mainline AADT	Increases
Percentage of population ages 16–24 working part time	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of households with 1 vehicle	Decreases
Percentage of households with 0 vehicles	Increases
Percentage of population ages 16–24 unemployed	Decreases
Average annual snowfall total	Increases

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes at four-leg stop-controlled intersections (with stop control on the minor road) on rural two-lane roads:

- *Mainline AADT*: an increase in mainline AADT was consistently associated with an increase in the frequency of ANG-D and ANG-N crashes at four-leg stop-controlled intersections on rural two-lane roads.
- *Cross street AADT*: an increase in cross street AADT was consistently associated with an increase in the frequency of ANG-D and ANG-N crashes at four-leg stop-controlled intersections on rural two-lane roads.
- *Speed*: an increase in speed was consistently associated with an increase in the frequency of ANG-D and ANG-N crashes at four-leg stop-controlled intersections on rural two-lane roads; design speed appeared to increase ANG-KAB-N crashes in California, and speed limit appeared to increase ANG-KAB-D, ANG-KABCO-D, ANG-KAB-N, and ANG-KABCO-N crashes in Ohio.

- *Lane width*: an increase in lane width was associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on rural two-lane roads in California and Ohio.

There were additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Mainline left-turn channelization*: presence of channelized left-turn lanes on the mainline was associated with a decrease in the frequency of ANG crashes at four-leg stop-controlled intersections on rural two-lane roads in California.

With respect to sociodemographic characteristics, two factors showed a consistent relationship to ANG crashes in California and Ohio:

- The percentage of population ages 45–64 appeared to increase ANG-KAB-N crashes at four-leg stop-controlled intersections on rural two-lane roads in California and Ohio (the relationship was opposite for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural two-lane roads in California).
- The percentage of households with income >$100,000 appeared to increase ANG-D crashes at four-leg stop-controlled intersections on rural two-lane roads in California and Ohio and ANG-KAB-N crashes at four-leg stop-controlled intersections on rural two-lane roads in Ohio.
- The percentage of households with income <$50,000 appeared to decrease ANG-D crashes at four-leg stop-controlled intersections on rural two-lane roads in California and Ohio; the contributing factor appeared to increase ANG-KAB-N crashes in California and all groups of ANG crashes in Ohio.

The following also appeared with respect to ANG crashes in California and Ohio:

- The percentage of the population ages 16–24 that is unemployed appeared to decrease ANG crashes at four-leg stop-controlled intersections on rural two-lane roads in Ohio.
- The percentage of households with two or more vehicles appeared to decrease ANG-D crashes at four-leg stop-controlled intersections on rural two-lane roads in Ohio.

With respect to weather characteristics, two factors showed a consistent relationship to ANG crashes:

- Average annual rainfall total appeared to increase ANG crashes at four-leg stop-controlled intersections on rural two-lane roads in California.
- Annual average maximum temperature appeared to increase ANG-KAB-D and ANG-KAB-N crashes in California and ANG-KAB-D in Ohio.

ANG CRASHES AT FOUR-LEG STOP-CONTROLLED INTERSECTIONS ON URBAN TWO-LANE ROADS

California

Daytime

Table 69 summarizes the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on urban two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 69. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads: California.

Variable	Impact on Crash-Frequency Predictions
Percentage of households with income <$50,000	Decreases
Design speed	Increases
Percentage of households with ≥2 vehicles	Increases
Percentage of households with income >$100,000	Increases
Annual average maximum temperature	Decreases
Annual average minimum temperature	Increases
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 20–44	Decreases
Annual average winter minimum temperature	Decreases
Percentage of households with 1 vehicle	Decreases

Ohio

Daytime

Table 70 and table 71 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on urban two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 70. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Speed limit	Increases
Percentage of households with income >$100,000	Increases
Percentage of households with income <$50,000	Decreases
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 25+ without a high school diploma	Decreases
Cross street AADT	Increases
Percentage of population ages 75+	Decreases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of households with ≥2 vehicles	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Mainline AADT	Increases

Table 71. Contributing factors for ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Percentage of households with ≥2 vehicles	Increases
Percentage of households with income >$100,000	Increases
Speed limit	Increases
Percentage of population ages 25+ without a high school diploma	Increases
Mainline AADT	Increases
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of households with 1 vehicle	Decreases
Percentage of population ages 25+ with a university degree	Increases

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes at four-leg stop-controlled intersections (with stop control on the minor road) on urban two-lane roads:

- *Mainline AADT*: an increase in mainline AADT was associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads in Ohio.
- *Cross street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads in Ohio.
- *Speed*: an increase in speed was consistently associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads; design speed appeared to increase ANG-KAB-D crashes in California, and speed limit appeared to increase ANG-KAB-D and ANG-KABCO-D crashes in Ohio.

With respect to sociodemographic characteristics, two factors showed a consistent relationship to ANG crashes in California and Ohio:

- The percentage of households with two or more vehicles consistently appeared to increase ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads in California and Ohio.
- The percentage of households with income >$100,000 consistently appeared to increase ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads in California and Ohio.

The following also appeared with respect to ANG crashes in California and Ohio:

- The percentage of the population ages 16–24 that is unemployed appeared to decrease ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads in Ohio. The percentage of the population ages 16+ appeared to decrease ANG-KAB-D crashes in California and ANG-KABCO-D crashes in Ohio.
- The percentage of households with income <$50,000 appeared to decrease ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads in California and Ohio.

With respect to weather characteristics, two factors showed a consistent relationship to ANG crashes:

- Average annual minimum temperature appeared to increase ANG crashes at four-leg stop-controlled intersections on urban two-lane roads in California.
- Average annual maximum temperature appeared to decrease ANG crashes at four-leg stop-controlled intersections on urban two-lane roads in California.

ANG CRASHES AT FOUR-LEG SIGNALIZED INTERSECTIONS ON URBAN MULTILANE DIVIDED ROADS

California

Daytime

Table 72 summarizes the most influential predictor variables for the expected number of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 72. Contributing factors for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane divided roads: California.

Variable	Impact on Crash-Frequency Predictions
Average annual rainfall total	Decreases
Percentage of households with income <$50,000	Increases
Annual average winter minimum temperature	Decreases
Annual average minimum temperature	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Design speed	Increases
Lane width	Increases
Percentage of population ages 75+	Increases
Mainline AADT	Increases
Median width	Increases

Ohio

Daytime

Table 73 and table 74 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 73. Contributing factors for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane divided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percentage of population ages 16+ unemployed	Decreases
Percentage of population ages 45–64	Decreases
Total AADT (mainline AADT + cross street AADT)	Increases
Speed limit	Increases
Percentage of population ages 16–24 working part time	Decreases
Percentage of population ages 25+ without a high school diploma	Increases
Percentage of households with income >$100,000	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 25+ with a high school diploma	Decreases
Percentage of population ages 16–24 unemployed	Increases

Table 74. Contributing factors for ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane divided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Mainline AADT	Increases
Percentage of population ages 20–44	Increases
Percentage of population ages 25+ without a high school diploma	Increases
Cross street AADT	Increases
Percentage of population ages 16–24 unemployed	Decreases
Percentage of population ages 45–64	Decreases
Percentage of population ages 25+ with a high school diploma	Decreases
Percentage of households with ≥2 vehicles	Decreases
Percentage of population ages 16–24 working full time	Increases
Median width	Decreases
Speed limit	Increases

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes at four-leg signalized intersections on urban multilane divided roads:

- *Mainline AADT*: an increase in mainline AADT was consistently associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads.
- *Cross street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads in Ohio.
- *Speed*: an increase in speed was consistently associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads; design speed appeared to increase ANG-KAB-D crashes in California, and speed limit appeared to increase ANG-KAB-D and ANG-KABCO-D crashes in Ohio.

There were additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Lane width*: an increase in lane width was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane divided roads in California.
- *Median width*: an increase in median width was associated with an increase in the frequency of ANG-KAB-D crashes at four-leg signalized intersections on urban multilane divided roads in California but a decrease in the frequency of ANG-KABCO-D crashes at four-leg signalized intersections on urban multi-lane divided roads in Ohio.

With respect to sociodemographic characteristics, the percentage of households with income <$50,000 appeared to increase ANG-KAB-D crashes at four-leg signalized intersections on urban multilane divided roads in California. Interestingly, this factor has generally shown decreases in crashes for other FCFTs but appeared to increase crashes for this FCFT. The percentage of the population between the ages of 45 and 64 appeared to decrease ANG-KAB-D and ANG-KABCO-D crashes in Ohio. The percentage of the population ages 25+ without a high school diploma appeared to increase ANG-KAB-D and ANG-KABCO-D crashes in Ohio.

With respect to weather characteristics, the average annual number of days with a minimum temperature of ≤32°F appeared to decrease ANG-KAB-D crashes in both California and Ohio.

ANG CRASHES AT FOUR-LEG SIGNALIZED INTERSECTIONS ON URBAN MULTILANE UNDIVIDED ROADS

California

Daytime

Table 75 summarizes the most influential predictor variables for the expected number of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 75. Contributing factors for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads: California.

Variable	Impact on Crash-Frequency Predictions
Percentage of population ages 25+ with a university degree	Increases
Percentage of population ages 45–64	Decreases
Percentage of population ages 65–74	Increases
Percentage of population ages 25+ without a high school diploma	Decreases
Percentage of households with income >$100,000	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of population ages 20–44	Decreases
Lane width	Decreases
Cross street right-turn channelization	Decreases
Cross street AADT	Decreases**

**Counterintuitive finding.

Ohio

Daytime

Table 76 and table 77 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 76. Contributing factors for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Mainline AADT	Increases
Cross street AADT	Increases
Speed limit	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Increases
Percentage of households with income <$50,000	Increases
Percentage of households with ≥2 vehicles	Increases
Percentage of households with 1 vehicle	Decreases
Percentage of population ages 25+ with a university degree	Decreases
Percentage of population ages 25+ without a high school diploma	Increases
Percentage of households with 0 vehicles	Increases

Table 77. Contributing factors for ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Mainline AADT	Increases
Percentage of households with ≥2 vehicles	Decreases
Annual average winter minimum temperature	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Percentage of population ages 45–64	Decreases
Percentage of households with income <$50,000	Decreases
Percentage of households with 0 vehicles	Increases
Percentage of households with 1 vehicle	Decreases
Percentage of population ages 16–24 working full time	Increases
Speed limit	Increases
Number of channelized left-turn lanes	Increases

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes at four-leg signalized intersections on urban multilane undivided roads:

- *Mainline AADT*: an increase in mainline AADT was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio.
- *Speed*: an increase in speed limit was consistently associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio.

There were additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Cross street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio (the relationship was opposite for ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads in California).
- *Lane width*: an increase in lane width was associated with a decrease in the frequency of ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads in California.
- *Cross street right-turn channelization*: presence of right-turn channelization on cross street was associated with a decrease in the frequency of ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads in California.

With respect to sociodemographic characteristics, two factors showed a relationship to ANG crashes in California and Ohio:

- The percentage of the population ages 45–64 appeared to decrease ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads in California and ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio.
- The percentage of households with no vehicles appeared to increase ANG-KAB-D and ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio. The percentage of households with one vehicle appeared to decrease those same crash types.

With respect to weather characteristics, the research did not produce any contributing factors.

ANG CRASHES AT THREE-LEG STOP-CONTROLLED INTERSECTIONS ON RURAL TWO-LANE ROADS

California

Daytime

Table 78 summarizes the most influential predictor variables for the expected number of ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 78. Contributing factors for ANG-KAB-D crashes at three-leg stop-controlled intersections on rural two-lane roads: California.

Variable	Impact on Crash-Frequency Predictions
Mainline AADT	Increases
Percentage of households with 1 vehicle	Decreases
Annual average maximum temperature	Increases
Percentage of population ages 25+ with a high school diploma but no university degree	Decreases
Percentage of households with income >$100,000	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases
Annual average minimum temperature	Decreases
Percentage of population ages 25+ with a university degree	Decreases
Annual average winter minimum temperature	Decreases
Percentage of households with ≥2 vehicles	Increases

Ohio

Daytime

Table 79 and table 80 summarize the most influential predictor variables for the expected number of ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 79. Contributing factors for ANG-KAB-D crashes at three-leg stop-controlled intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of population ages 16–24 working part time	Increases
Curve radius	Increases/decreases*
Percentage of households with income <$50,000	Decreases
Percentage of population ages 25+ without a high school diploma	Decreases
Percentage of households with ≥2 vehicles	Increases
Percentage of population ages 16–24 working full time	Increases
Percentage of population ages 16–24 unemployed	Decreases
Mainline AADT	Increases
Percentage of population ages 65–74	Decreases
Lane width	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).

Table 80. Contributing factors for ANG-KABCO-D crashes at three-leg signalized intersections on rural two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Mainline AADT	Increases
Percentage of households with ≥2 vehicles	Decreases
Percentage of households with income between $50,000 and $100,000	Increases
Percentage of households with income <$50,000	Decreases
Lane width	Increases
Percentage of households with 0 vehicles	Increases
Average annual rainfall total	Decreases
Percentage of population ages 25+ without a high school diploma	Increases
Percentage of population ages 20–44	Increases
Annual average number of days with a minimum temperature of ≤32°F	Increases
Cross street AADT	Increases
Speed limit	Decreases**
Percent grade	Decreases**
Curve radius	Increases/decreases*

*Increases crash frequency when comparing curves to highway tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends the following contributing roadway factors associated with ANG crashes at three-leg stop-controlled intersections on rural two-lane roads:

- *Mainline AADT*: an increase in mainline AADT was consistently associated with an increase in the frequency of ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads.

There were additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Cross Street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-KABCO-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of ANG-D crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of ANG crashes) at three-leg stop-controlled intersections on rural two-lane roads in Ohio.
- *Lane width*: an increase in lane width was associated with an increase in the frequency of ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio.
- *Speed*: an increase in speed limit was associated with an increase in the frequency of ANG-KAB-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio (the relationship was opposite for ANG-KABCO-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio).

With respect to sociodemographic characteristics, the percentage of households with two or more vehicles appeared to increase ANG-KAB-D crashes at three-leg signalized intersections on rural two-lane roads in California and Ohio.

The following also appeared with respect to ANG crashes in California and Ohio:

- The percentage of the population ages 16–24 that is unemployed appeared to decrease ANG-KAB-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio.
- The percentage of households with income between $50,000 and $100,000 appeared to increase ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio. The percentage of households with income <$50,000 appeared to decrease those same crash types.

With respect to weather characteristics and ANG cashes, average annual rainfall totals appeared to decrease ANG-KABCO-D crashes at three-leg stop-controlled intersections on rural two-lane roads in Ohio.

ANG CRASHES AT FOUR-LEG STOP-CONTROLLED INTERSECTIONS ON RURAL MULTILANE DIVIDED ROADS

California

Daytime

Table 81 summarizes the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections on rural multilane divided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) according to random forests generated using California data.

Table 81. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads: California.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Percentage of population ages 16+ unemployed	Decreases
Percentage of households with income >$100,000	Increases
Percentage of population ages 20–44	Decreases
Median width	Increases
Average annual snowfall total	Decreases
Percentage of households with income <$50,000	Decreases
Percentage of population ages 16–24 unemployed	Decreases
Annual average winter minimum temperature	Increases
Percentage of population ages 15–19	Increases

Ohio

Daytime

Table 82 and table 83 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections on rural multilane undivided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data.

Table 82. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Speed limit	Decreases**
Mainline AADT	Increases
Cross street AADT	Decreases**
Average annual snowfall total	Increases
Annual average maximum temperature	Decreases
Percentage of households with income between $50,000 and $100,000	Decreases
Average annual rainfall total	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 16–24 unemployed	Increases
Percentage of population ages 25+ with a university degree	Decreases

**Counterintuitive finding.

Table 83. Contributing factors for ANG-KABCO-D crashes at four-leg stop-controlled intersections on rural multilane divided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Speed limit	Decreases**
Mainline AADT	Increases
Percentage of households with income >$100,000	Increases
Percentage of population ages 16+ unemployed	Increases
Number of channelized left-turn lanes	Increases
Average annual snowfall total	Increases
Percentage of population ages 25+ with a university degree	Increases
Annual average maximum temperature	Increases
Percentage of population ages 16–24 working full time	Decreases

**Counterintuitive finding.

Discussion

Based on the analysis, the research team recommends only one contributing roadway factor associated with ANG crashes at four-leg stop-controlled intersections on rural multilane divided roads:

- *Mainline AADT*: an increase in mainline AADT was consistently associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in Ohio.

There were additional roadway variables that showed relationships with ANG crash frequency, but results were inconsistent or could not be validated across the two States due to data limitations:

- *Cross street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in California and ANG-KABCO-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in Ohio (the relationship was opposite for ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in Ohio).
- *Median width*: an increase in median width was associated with an increase in the frequency of ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in California.
- *Speed*: an increase in speed limit was associated with a decrease in the frequency of ANG crashes at four-leg stop-controlled intersections on rural multilane divided roads in Ohio. This finding was opposite of reported speed effects on other intersection FCFTs.

With respect to sociodemographic characteristics, the percentage of households with income >$100,000 appeared to increase ANG-D crashes at four-leg stop-controlled intersections on rural multilane divided roads in California and Ohio.

With respect to weather characteristics, the research did not produce any contributing factors.

PEDESTRIAN CRASHES

This section includes an overview based on information from related published literature of contributing factors for pedestrian crashes.

Specifically, a study by Thomas et al. (2017) looked at identifying and screening intersection locations with the potential for future pedestrian crashes and injuries to help Seattle, WA, broaden treatment priorities beyond high crash locations. The methodology used was similar to this current project, as they used random forests to identify potential contributing factors for pedestrian crashes.

Thomas et al. (2017) evaluated the following pedestrian crash types:

- All pedestrian crashes at intersections.
- Pedestrians crossing at intersections and struck by motorists going straight.

The categorical roadway variables used by Thomas et al. (2017) for inputs to the random-forest analysis included the following:

- Highest arterial class for intersection approaches.
- Type of traffic control at intersection.
- Presence of a mini-roundabout at the intersection.
- Total number of motor vehicle lanes on the largest approach leg.
- Number of through motor-vehicle lanes of all types present on the largest leg of the intersection.
- Number of legs present at the intersection.
- Number of legs at the intersection that are local streets.
- Total number of motor-vehicle travel lanes for all legs at the intersection.
- Sum of through motor-vehicle travel lanes on all legs at the intersection.
- Presence of parking on any legs of the intersection.
- Presence of bus-only lanes on any approach.
- Presence of painted or raised medians on any approach.
- Presence of right-turn lanes on any approach.
- Presence of left-turn lanes on any approach.
- Presence of two-way left-turn lanes on any approach.
- Presence of bike lanes on any approach.
- Presence of any type of two-way bicycle track.
- Presence of bicycle shared-lane markings on any approach.
- Presence of a shared-use path crossing at the intersection.

Different measures of exposure and scale factors used by Thomas et al. (2017) for inputs to the random-forest analysis included the following:

- Estimated annual average daily bicycle volume at intersection.
- Estimated annual average daily pedestrian (AADP) volume at intersection.
- Presence of one of more K–12 schools within 0.25 mi of the intersection.
- Presence of university campus within 0.25 mi of the intersection.
- Mean income within 150 ft of the intersection.
- Total population within 0.10 mi of the intersection.
- Total employment within 0.25 mi of the intersection.
- Number of commercial properties within 0.10 mi of the intersection.
- Proportion of population ages 65+ within 0.25 mi of the intersection.
- Proportion of population ages 18 and younger within 0.25 mi of the intersection.
- Network distance (in miles) to the nearest university campus.
- Maximum percent slope on any approach.
- Average slope of terrain within 0.50 mi surrounding the intersection.
- Number of buses/trains stopping within 150 ft of the intersection on a typical weekday.
- All building volume (height × area) within 0.10 mi of the intersection.
- Commercial building volume (height × area) within 0.10 mi of the intersection.

Table 84 and table 85 provide the most influential predictor variables and their direction of effect as identified by Thomas et al.'s framework (2017).

Table 84. Contributing factors for all pedestrian crashes at intersections.

Variable	Impact on Crash-Frequency Predictions
Number of commercial properties within 0.10 mi of the intersection	Increases
Number of buses/trains stopping within 150 ft of the intersection on a typical weekday	Increases
All building volume (height × area) within 0.10 mi of the intersection	Increases
AADP volume	Increases
Commercial building volume (height × area) within 0.10 mi of the intersection	Decreases
Number of legs at the intersection that are local streets	Decreases
Mean income within 150 ft of the intersection	Decreases
Average slope of terrain within 0.50 mi surrounding the intersection	Decreases
Total population within 0.10 mi of the intersection	Increases
Presence of traffic signal	Increases
Number of legs present at the intersection	Increases
Total number of motor-vehicle travel lanes for all legs at the intersection	Increases
Total number of motor-vehicle lanes on the largest approach leg	Increases
Highest arterial class entering the intersection	Increases
Presence of parking on any legs of the intersection	Increases

Table 85. Contributing factors for pedestrians crossing at intersections and struck by motorists going straight.

Variable	Impact on Crash-Frequency Predictions
Number of commercial properties within 0.10 mi of the intersection	Increases
Number of buses/trains stopping within 150 ft of the intersection on a typical weekday	Increases
All building volume (height × area) within 0.10 mi of the intersection	Increases
AADP volume	Increases
Commercial building volume (height × area) within 0.10 mi of the intersection	Decreases
Number of legs at the intersection that are local streets	Decreases
Mean income within 150 ft of the intersection	Decreases
Presence of traffic signal	Increases
Number of legs present at the intersection	Increases
Total number of motor-vehicle lanes on the largest approach leg	Increases
Highest arterial class entering the intersection	Increases

As can be seen from table 84 and table 85, the following variables were identified as factors contributing to potential increases in both types of pedestrian crashes:

- Number of commercial properties within 0.10 mi of the intersection.
- Number of buses/trains stopping within 150 ft of the intersection on a typical weekday.
- All building volume (height × area) within 0.10 mi of the intersection.
- Estimated AADP volume at intersection.
- Presence of traffic signal at the intersection.
- Number of legs present at the intersection.
- Total number of motor-vehicle travel lanes for all legs at the intersection.

The analyses results presented throughout this chapter, along with the review of a previously conducted pedestrian crash analysis by Thomas et al. (2017), provide insights to key contributing factors for FCFTs. This information served as the basis for developing a countermeasure-selection process, which is presented in the next chapter. While the analyses uncovered contributing factors related to traffic, roadway, socioeconomic, and weather characteristics, the countermeasure-selection process is focused on the traffic and roadway findings. Findings linked to socioeconomic- and weather-related factors show promise, but there is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results.

CHAPTER 6. COUNTERMEASURE-SELECTION PROCESS

In addition to identifying contributing factors for FCFTs (presented in chapter 5), another objective of this project was to use the contributing-factor findings to assist safety practitioners in making informed choices regarding countermeasures to address the focus crash types. This chapter describes a process other practitioners can use to identify potential countermeasures for each set of focus crash types. The process is accompanied by several examples to demonstrate the steps for various crash types and provide a list of countermeasures identified for those crash types. The chapter concludes with the countermeasures that address the FCFTs and contributing factors covered in chapter 4 and chapter 5. The overview includes multiple potential countermeasures for intersection and nonintersection FCFTs, respectively, for consideration by State and local transportation agencies as part of an approach to systemic road-safety management.

This process recognizes that, although there are certain contributing factors associated with focus crash types, selecting countermeasures must be broad and encompass many options for addressing the crash type. The usefulness of the contributing-factor identification is to allow a practitioner to identify which countermeasures specifically address those factors as a part of their safety effectiveness and also to prioritize locations for countermeasures. This identification and prioritization of countermeasures may in turn raise the priority for selecting and implementing those countermeasures. In addition, the contributing factors will inform practitioners' decisions about where to implement the selected countermeasures as part of their systemic safety-management efforts. The following six steps detail the countermeasure-selection process:

1. Identify a focus crash type.
2. Identify contributing factors for the focus crash type.
3. Assemble a list of potential countermeasures that address the focus crash type.
4. Identify countermeasures that address the contributing factors associated with the focus crash type.
5. Identify countermeasures with CMFs.
6. Select a countermeasure.

BACKGROUND RELEVANT TO COUNTERMEASURE SELECTION

Chapter 4 and chapter 5 provided tables of the most influential predictor variables and identified them as factors contributing to increases or decreases in potential future crashes. A brief discussion on similarities between the results across States and recommended contributing factors for systemic safety analysis based on these results accompanied the tables.

The research team conducted a contributing-factor analysis for the expected crash frequencies of KABCO crashes using Ohio data and compared results to those for KAB crashes in Ohio. Results from the contributing-factor analysis were quite similar when using either the KAB or KABCO severities. In almost all cases, the most influential predictor variables stayed consistent between the KAB and KABCO groupings, except for some changes in the percent increase of predictions of MSEs associated with removal of the variables.

The following sections summarize prevalent factors from each main category of factors explored in this analysis: roadway, socioeconomic, and climate.

Roadway Variables

Recommended roadway variables for contributing-factor analysis from intersection and nonintersection FCFT categories are as follows:

- Intersection:
    - Mainline AADT.
    - Cross street AADT.
    - Speed.
    - Lane width.
    - Curve radius.

- Nonintersection:
    - Average AADT.
    - Percent grade.
    - Curve radius.
    - Shoulder width.

A secondary set of roadway factors that also generally showed consistent results and could be presented as contributing factors include the following:

- Surface width (wider surface widths decreased predicted crash frequency).
- Unpaved shoulders or no shoulders (the presence of one or more unpaved shoulders or no shoulders increased predicted crash frequency compared to the presence of two paved shoulders).
- Terrain (rolling and mountainous terrain increased predicted crash frequency).
- Left- and right-turn channelization (the presence of left- and right-turn channelization at intersections decreased predicted crash frequency).

There were some counterintuitive results where a variable was expected to increase crash frequency and instead appeared to decrease crash frequency or vice versa. The majority of these counterintuitive results were related to percent grade, shoulder width, speed, and cross street AADT. The percentage of trucks on the roadway was another variable that was expected to increase crashes but consistently came up as a factor expected to decrease crashes. This finding was not identified as counterintuitive due to the fact that there have been studies conducted in the past by Milton and Mannering (1998), Lord et al. (2005), and Milton et al. (2008) that support this finding. Milton and Mannering (1998) looked at the relationship among highway geometrics, traffic-related elements, and motor vehicle–accident frequencies using Washington data. They concluded that an increase in the percentage of both single-unit trucks and all trucks was associated with lower accident frequency. Lord et al. (2005) concluded that an increase in the percentage of trucks was associated with a reduction in the number of rear-end and ROR crashes. Milton et al. (2008) concluded that an increase in percentages of trucks tends to lower the possible amount of injury crashes.

In most cases, counterintuitive findings would show up in the expected direction (i.e., increasing or decreasing crashes) when using data from a different State. For example, average shoulder width came up as a factor contributing to increases in LNDP-N crashes on rural two-lane roads when using Washington data. However, the same factor showed as contributing to a decrease in crash frequency (as expected) for LNDP-N crashes on rural two-lane roads when using Ohio data.

Counterintuitive findings could happen for various reasons. One possible reason could be the application of some other treatments for which data were not available for the analysis. Another possible reason could be the completeness of data. For example, the research team had the information for shoulder widths for Washington; however, information regarding paved and unpaved shoulder widths was not available.

Socioeconomic Variables

There were some interesting findings regarding socioeconomic variables. The following variables showed as most influential for intersection and nonintersection FCFT categories:

- Percentage of population ages 16–24 working full time (generally appeared to decrease segment crashes with one occurrence of increasing intersection crashes).
- Percentage of population ages 16–24 working part time (generally appeared to decrease segment crashes).
- Percentage of households with income <$50,000 (consistently appeared to decrease all crash types, except for crashes at urban signalized intersections).
- Percentage of households with income >$100,000 (consistently appeared to increase crashes).
- Percentage of population ages 16–24 that is unemployed (consistently appeared as a predictive factor, but whether it was expected to increase or decrease crashes varied by FCFT).
- Percentage of population ages 15–19 (consistently expected to increase ROR, HEO, and ROLL crashes, particularly on horizontal curves).

The research team did not recommend any of the socioeconomic variables for systemic safety analysis due to limitations in the data-linkage approach and scarcity of prior research linking these variables to crashes at segment and intersection levels. However, these variables showing up among the most influential predictor variables does provide a basis for future research to explore these as potential contributing factors for systemic analysis efforts. Likely, the variable with the strongest theoretical basis would be the percentage of the population ages 15–19 as a factor contributing to increased ROR, HEO, and ROLL crashes on rural roads.

Climate Variables

As with the socioeconomic variables, there were some interesting findings regarding climate variables. The following variables showed as most influential for intersection and nonintersection FCFT categories:

- Average annual snowfall totals (generally appeared to increase crashes).
- Average annual rainfall totals (generally appeared to increase crashes).
- Average annual number of days with a minimum temperature of ≤32°F (generally to decrease crashes).

The research team did not recommend any of the climate variables for systemic safety analysis due to scarcity of prior research linking these variables to crashes and some of the inconsistencies in the directions (i.e., increasing or decreasing crashes) of the effects. However, these variables showing up among the most influential predictor variables in some cases does provide a basis for future research to explore them as potential contributing factors.

The following sections describe the steps of a countermeasure-selection process for FCFTs based on traffic- and roadway-related contributing factors. The traffic- and roadway-related contributing factors identified in this process can inform the selection and implementation of countermeasures as a part of a systemic safety analysis. The steps include (1) identifying a focus crash type, (2) identifying contributing factors for the focus crash type, (3) assembling a list of potential countermeasures that address the focus crash type, (4) identifying countermeasures that address the contributing factors associated with the focus crash type, (5) identifying countermeasures with CMFs, and (6) selecting a countermeasure.

STEP 1. IDENTIFY A FOCUS CRASH TYPE

Approach

Identify the focus crash type of interest. Using the FCFT definitions from chapter 5, the focus crash type is defined by the type of maneuver (e.g., ROR), time of day (daytime versus nighttime), and type of segment (curve versus tangent) or type of intersection (e.g., four-leg signalized). Selecting a focus crash type should be done according to the priorities of the transportation agency (e.g., as outlined in a State or regional SHSP) and based on an analysis of available data. These methods may include the following:

- Jurisdiction-wide analysis of crash data to identify overrepresented crash types.
- Site-level analysis to determine overrepresented crash types or crash types in excess of the expected performance at a specific site (i.e., through crash-predictive methodologies).
- Systemic safety analysis of a network to identify crash types that represent the greatest number of KA crashes across the system being analyzed.

Example

Based on data analysis, an agency selected the following focus crash type: ROR-D crashes on curves. The agency conducted an analysis using the approach detailed in chapter 4 and found KA crashes of this type to be particularly prevalent on rural two-lane roads.

STEP 2. IDENTIFY CONTRIBUTING FACTORS FOR THE FOCUS CRASH TYPE

Approach

Using the results of this project (chapter 7), identify contributing factors associated with the focus crash type. Some factors may be expected to increase a particular crash type, such as the presence of an unpaved shoulder, whereas others may decrease certain crash types, such as increasing (i.e., flattening) the radius of a horizontal curve.

Example

Given the selected focus crash type of ROR-D crashes on curves, the results of the agency's analysis showed that contributing factors for this crash type (ROR-D crashes on curves) were as follows:

1. Percent grade (as grade increases, crash potential increases).
2. Average shoulder width (as shoulder width increases, crash potential decreases).

STEP 3. ASSEMBLE A LIST OF POTENTIAL COUNTERMEASURES THAT ADDRESS THE CRASH TYPE

Approach

Use countermeasure resources to assemble a list of potential countermeasures that may address the focus crash type. Such resources may include the following:

- NCHRP Report 500 Series (National Academy of Sciences 2003a, 2003b, 2003c, 2003d, 2003e, 2003f, 2004a, 2004b, 2004c, 2004d, 2004e, 2004f, 2004g, 2005a, 2005b, 2005c, 2006, 2007, 2008a, 2008b, 2008c, 2008d, 2009).
- HSM (AASHTO 2010).
- CMF Clearinghouse (FHWA 2018a).
- FHWA Proven Safety Countermeasures (FHWA 2017a).
- *Countermeasures That Work* (Goodwin et al. 2013).
- PEDSAFE (FHWA 2018f).
- BIKESAFE (FHWA 2018g).
- State-generated list of common countermeasures within a State (i.e., a toolbox).

This assembled list should contain a wide pool of potential countermeasures. Eligible countermeasures should address any part of the focus crash type. For example, if the focus crash type is ROR-N crashes on curves, a countermeasure that addresses ROR crashes should be considered eligible even if it does not specifically address curve locations or nighttime conditions.

Example

Given the selected focus crash type (ROR-D crashes on curves) and an examination of the previously listed countermeasure resources, the agency identified the following countermeasures as eligible for consideration:

- Install shoulder rumble strips.
- Install centerline rumble strips.
- Widen lanes.
- Enhance curve delineation (e.g., add chevrons, large arrows, or delineators on guardrails).
- Install or improve curve warning signs (e.g., add flashing beacons).
- Install warning arrows on the pavement prior to the curve.
- Install pavement markings to decrease speed prior to the curve.
- Flatten horizontal curves (i.e., increase the curve's radius).
- Flatten vertical curves (i.e., decrease the curve's grade).
- Enhance pavement markings (e.g., add edgelines).
- Install skid-resistant pavement or improve pavement friction.
- Widen paved shoulders.
- Pave shoulders.
- Install SafetyEdge.
- Design safer slopes and ditches to prevent ROLLs.
- Remove/relocate objects in hazardous locations.
- Delineate roadside objects.
- Improve the design of roadside hardware (e.g., bridge rails).
- Improve the design and application of barrier and attenuation systems.

STEP 4. IDENTIFY COUNTERMEASURES THAT EXPLICITLY ADDRESS CONTRIBUTING FACTORS ASSOCIATED WITH THE FOCUS CRASH TYPE

Approach

Compare the information known about each countermeasure on the list from step 3 to the contributing factors identified in step 2. Identify countermeasures that specifically address one or more contributing factors. The degree to which countermeasures address contributing factors is subjective; a logical link should be drawn between many countermeasures and contributing factors. For the purpose of this process, only clear, explicit relations to contributing factors should be indicated.

Example

The agency used table 86 to summarize which countermeasures address the identified contributing factors for the focus crash type (ROR-D crashes on curves).

Table 86. Example selection process: countermeasure summary with contributing factors.

Countermeasure	Factor 1: Percent Grade	Factor 2: Average Shoulder Width
Install shoulder rumble strips	—	—
Install centerline rumble strips	—	—
Widen lanes	—	—
Enhance curve delineation (e.g., add chevrons, large arrows, or delineators on guardrails)	—	—
Install or improve curve warning signs (e.g., add flashing beacons)	—	—
Install warning arrows on the pavement prior to the curve	—	—
Install pavement markings to decrease speed prior to the curve	—	—
Flatten horizontal curves (i.e., increase radius)	—	—
Flatten vertical curves (i.e., decrease grade)	Y	—
Enhance pavement markings (e.g., add edgelines)	—	—
Install skid-resistant pavement or improve pavement friction	—	—
Widen paved shoulder	—	Y
Pave shoulder	—	Y
Install SafetyEdge	—	—
Design safer slopes and ditches to prevent ROLLs	—	—
Remove/relocate objects in hazardous locations	—	—
Delineate roadside objects	—	—
Improve the design of roadside hardware (e.g., bridge rails)	—	—
Improve the design and application of barrier and attenuation systems	—	—

—No data available.
Y = contributing factor is addressed by the countermeasure.

STEP 5. IDENTIFY COUNTERMEASURES WITH CMFS

Approach

Use CMF resources to determine which countermeasures on the list from step 3 have established CMFs that quantify the safety effect. This is a critical part of the process as the systemic approach consists of widely implementing low-cost, proven countermeasures to achieve safety benefits across a large portion of the system. A countermeasure with a known effect can also be compared to other countermeasures in a prioritized selection and used to generate a B/C analysis of proposed alternatives. It is possible that countermeasures for which a robust set of CMFs are available may have CMFs that specifically address the focus crash type. If this is the case, extra consideration should be given to these countermeasures in the final selection.

The most comprehensive and accessible resource for CMF information is the CMF Clearinghouse (FHWA 2018a). Other resources for CMF information may include State-specific CMF lists (FHWA 2018b).

Example

The agency used table 87, which expands on table 86, to show which countermeasures have CMFs available for ROR crashes.

Table 87. Example selection process: countermeasure summary with contributing factors and CMF indicators.

Countermeasure	Factor 1: Percent Grade	Factor 2: Average Shoulder Width	CMF Available
Install shoulder rumble strips	—	—	Y
Install centerline rumble strips	—	—	Y
Widen lanes	—	—	Y
Enhance curve delineation (e.g., add chevrons, large arrows, or delineators on guardrails)	—	—	Y
Install or improve curve warning signs (e.g., add flashing beacons)	—	—	Y
Install warning arrows on the pavement prior to the curve	—	—	—
Install pavement markings to decrease speed prior to the curve	—	—	—
Flatten horizontal curves (i.e., increase radius)	—	—	Y
Flatten vertical curves (i.e., decrease grade)	Y	—	Y
Enhance pavement markings (e.g., add edgelines)	—	—	Y
Install skid-resistant pavement or improve pavement friction	—	—	Y
Widen paved shoulder	—	Y	Y
Pave shoulder	—	Y	Y
Install SafetyEdge	—	—	Y
Design safer slopes and ditches to prevent ROLLs	—	—	Y
Remove/relocate objects in hazardous locations	—	—	Y
Delineate roadside objects	—	—	—
Improve the design of roadside hardware (e.g., bridge rails)	—	—	—
Improve the design and application of barrier and attenuation systems	—	—	Y

—No data available.
Y = contributing factor is addressed by the countermeasure or a CMF is available for the countermeasure.

STEP 6. SELECT A COUNTERMEASURE

Approach

Select a countermeasure to address the focus crash type. This selection should use the pool of eligible countermeasures generated in step 3, and how well each countermeasure addresses the specific contributing factors for the focus crash type (as determined in step 4) and the extent to which CMFs are available for each countermeasure (as determined in step 5) should be considered.

Example

The agency referenced CMF resources and found that many of the countermeasures on the list from step 3 have CMFs that quantify the safety effect. Among these countermeasures, the agency determined that those related to paving the shoulder or widening the paved shoulder explicitly address the contributing factor of average shoulder width. Additionally, the agency concluded that the countermeasure of flattening a vertical curve to decrease the grade was also related to the contributing factor of percent grade. The agency first considered vertical curve flattening as it could be more effective and address an underlying design issue. However, the agency determined that flatting vertical curves would be expensive and decided that it would not be feasible given the agency's existing budget constraints.

The agency decided that the most appropriate countermeasure to address the focus crash type (ROR-D crashes on curves) based explicitly on the contributing factors was paving the shoulder or widening the paved shoulder. Alternatively, the agency recognized that it could select a countermeasure in table 87 and then use the contributing factors for ROR crashes to identify candidate locations at which to install the countermeasure using a systemic approach.

ADDITIONAL EXAMPLES

This section includes two additional hypothetical examples that demonstrate the process of selecting countermeasures to address focus crash types.

Example 1: HEO Crashes

Complete step 1: identify the focus crash type. An agency determined a focus crash type of HEO-D crashes on horizontal curves. The agency used a crash-tree diagram, as suggested in FHWA's *Systemic Safety Project Selection Tool*, to help identify facility types associated with the focus crash type. The agency determined that severe HEO-D crashes are particularly prevalent on rural two-lane roads.

Complete step 2: identify contributing factors for the focus crash type. Given the selected focus crash type and using the results of this project (chapter 7), the agency determined that the contributing factor for this focus crash type is curve radius (as curve radius increases, predicted crash frequency decreases).

Complete step 3: assemble a list of potential countermeasures that address the focus crash type. Given the selected focus crash type and an examination of countermeasure resources, the agency arrived at the list of potential countermeasures in table 88.

Complete step 4: identify countermeasures that address the contributing factors associated with the focus crash type. The agency identified countermeasures that address the contributing factors associated with the focus crash type by comparing the information known about each countermeasure on the list from step 3 to the contributing factor identified in step 2. The agency used the information in table 88, which summarizes which countermeasures address the identified contributing factors for the focus crash type. The agency determined that the only countermeasure that will address the contributing factor of curve radius is flatten horizontal curve.

Complete step 5: identify countermeasures with CMFs. The agency identified which countermeasures have CMFs using CMF resources, such as the CMF Clearinghouse or State-specific CMF lists (FHWA 2018a). The agency used table 88 to show which countermeasures have CMFs available for HEO crashes.

Complete step 6: select a countermeasure. The agency found that many of the countermeasures in table 88 have CMFs that quantify the safety effect of the countermeasure. From this pool of potential countermeasures, the agency determined in step 4 that only one countermeasure (flatten curve radius) explicitly addressed the contributing factor of curve radius. Thus, this countermeasure could be one possible selection to address the focus crash type. However, if the agency determined it could not afford the high cost of realigning horizontal curves, it could instead implement one or more of the other countermeasures identified in table 88, particularly those with known CMFs, such as pavement friction or curve delineation. In such cases, the agency could select other countermeasures and then use the contributing factors for HEO crashes to identify candidate locations at which to install the countermeasures using a systemic approach.

Table 88. Countermeasure summary for example 1: HEO crashes.

Countermeasure	Factor 1: Curve Radius	CMF Available
Install centerline rumble strips	—	Y
Install profiled thermoplastic stripes for the centerline	—	
Widen lanes	—	Y
Install center two-way left-turn lanes	—	Y
Reallocate lane and shoulder width to include a narrow buffer median	—	—
Install passing relief lanes, alternating passing lanes or four-lane sections at key locations	—	Y
Enhance curve delineation (e.g., add chevrons, large arrows, or delineators on guardrails)	—	Y
Install or improve curve warning signs (e.g., add flashing beacons)	—	Y
Install warning arrows on the pavement prior to the curve	—	—
Install pavement markings to decrease speed prior to the curve	—	—
Flatten horizontal curve (i.e., increase radius)	Y	Y
Install skid-resistant pavement or improve pavement friction	—	Y
Widen paved shoulder	—	Y
Pave shoulder	—	Y
Install raised median	—	Y
Install cable median barrier	—	Y
Install lighting	—	Y

—No data available.
Y = contributing factor is addressed by the countermeasure or a CMF is available for the countermeasure.

Example 2: ANG Crashes

Complete step 1: identify the focus crash type. An agency determined a focus crash type of ANG-N crashes at stop-controlled intersections. The agency used FHWA's *Systemic Safety Project Selection Tool*, which outlines a process for States or local agencies to identify facility types associated with a focus crash type using a crash-tree diagram, to create a crash-tree diagram (Preston et al. 2013a). Using the diagram, the agency determined that KA crashes of this type are particularly prevalent on rural two-lane roads in their jurisdiction.

Complete step 2: identify contributing factors for the focus crash type. Given the selected focus crash type and using the results of this project (chapter 7), the agency determined the following contributing factors for ANG-N crashes:

1. Design speed (as the design speed increases, crash potential increases).
2. Speed limit (as the speed limit increases, crash potential increases).
3. Mainline left-turn channelization (when present, crash potential decreases).

Complete step 3: assemble a list of potential countermeasures that address the focus crash type. Given the selected focus crash type and an examination of countermeasure resources, the agency arrived at the list of potential countermeasures in table 89.

Complete step 4: identify countermeasures that address the contributing factors associated with the focus crash type. The agency identified countermeasures that address the contributing factors associated with the focus crash type by comparing the information known about each countermeasure on the list from step 3 to the contributing factors identified in step 2. The agency used the information in table 89, which summarizes which countermeasures address the identified contributing factors for the focus crash type.

Complete step 5: identify countermeasures with CMFs. The agency identified which countermeasures have CMFs using CMF resources, such as the CMF Clearinghouse or State-specific CMF lists (FHWA 2018a). The agency used table 89 to show which countermeasures have CMFs available for ANG crashes.

Complete step 6: select a countermeasure. The agency found that many of the countermeasures in table 88 have CMFs that quantify the safety effect of the countermeasure. From this pool of potential countermeasures, the agency determined that the countermeasures of converting the intersection to a roundabout, installing or lengthening exclusive left-turn lanes, and providing offset left-turn lanes had known CMFs. The agency noted that these three countermeasures explicitly address the focus crash type's contributing factors and identified them as one group of priority countermeasures. However, the agency also noted that it could implement one or more of the other countermeasures in table 89, particularly those with known CMFs associated with the focus crash type. In such cases, the agency could select other countermeasures and then use the contributing factors for ANG crashes to identify candidate locations at which to install the countermeasures using a systemic approach.

Table 89. Countermeasure summary for example 2: ANG crashes.

Countermeasure	Factor 1: Design Speed	Factor 2: Speed Limit	Factor 3: Mainline Left-Turn Channelization	CMF Available
Replace direct left-turn design with a right turn/U-turn	—	—	—	Y
Convert the intersection to a roundabout	Y	—	—	Y
Install a traffic signal	—	—	—	Y
Install or lengthen an exclusive left-turn lane	—	—	Y	Y
Provide an offset left-turn lane	—	—	Y	Y
Close, relocate, or restrict turns at driveways near intersections	—	—	—	—
Provide an offset right-turn lane at an intersection	—	—	—	Y
Provide a full-width paved shoulder in intersection areas	—	—	—	Y
Restrict or eliminate turning maneuvers with signing	—	—	—	—
Restrict or eliminate turning maneuvers by providing channelization or closing median openings	—	—	—	—
Convert a four-leg intersection to two T-intersections (on a low-volume cross street)	—	—	—	Y
Convert offset T-intersections to four-leg intersection (on a high-volume cross street)	—	—	—	—
Reduce or eliminate intersection skew	—	—	—	Y

Countermeasure	Factor 1: Design Speed	Factor 2: Speed Limit	Factor 3: Mainline Left-Turn Channelization	CMF Available
Improve the sight distance at an intersection	—	—	—	Y
Install a dynamic advance intersection warning system	—	—	—	Y
Provide an automated real-time system to inform drivers of the suitability of available gaps for making turning and crossing maneuvers	—	—	—	Y
Provide intersection lighting	—	—	—	Y
Install a splitter island on the minor-road approach to an intersection	—	—	—	—
Provide a stop bar (or provide a wider stop bar) on the minor-road approach	—	—	—	—
Install a larger regulatory and warning sign at the intersection	—	—	—	—
Install transverse rumble strips on the intersection approach	—	—	—	Y
Provide dashed markings (i.e., extended left edgelines) for major-road continuity across the median opening at a divided highway intersection	—	—	—	—
Provide a supplementary stop sign mounted over the roadway	—	—	—	—
Install a double stop sign	—	—	—	Y
Provide a pavement marking with a supplementary message, such as stop ahead	—	—	—	Y
Install a flashing beacon at stop-controlled intersection	—	—	—	Y
Replace a standard stop sign with a flashing LED stop sign	—	—	—	Y
Convert a two-way intersection to an all-way stop-controlled intersection	—	—	—	Y
Provide traffic calming on an intersection approach through a combination of geometric and traffic-control devices	—	—	—	—
Lower speed on an intersection approach	Y	Y	—	—
Install a median acceleration lane	—	—	—	Y
Provide skid resistance at an intersection	—	—	—	Y
Narrow a lane through rumble strips and a painted median	—	—	—	Y

—No data available.
Y = contributing factor is addressed by the countermeasure or a CMF is available for the countermeasure.

TARGETED COUNTERMEASURES FOR CONTRIBUTING FACTORS

Many countermeasures have been researched and evaluated for improving highway safety. FHWA's *Proven Safety Countermeasures* identifies a small set of these countermeasures, which have proven repeatedly to reduce serious injuries and fatalities (FHWA 2017a). As a result, FHWA encourages their widespread implementation. Other proven and recommended countermeasures, along with additional information on the *Proven Safety Countermeasures*, are found in FHWA's *Low-Cost Treatments for Horizontal Curve Safety* (Albin et al. 2016), the Institute of Transportation Engineers' *Unsignalized Intersection Improvement Guide* (ITE 2015), FHWA's Intersection Safety website (FHWA 2015a), FHWA's *Roadway Departure Safety* (FHWA 2017b), and FHWA's *Intersection Safety Strategies Brochure* (FHWA 2015b). Table 90 and table 91 show how the *Proven Safety Countermeasures* applies to various intersection and nonintersection contributing factors identified previously in this report. This section is not meant to provide complete lists of countermeasures and countermeasure resources. Instead, it is meant to be a quick reference and starting point for agencies interested in applying countermeasures using the systemic approach to common FCFTs.

Table 90. Systemic countermeasures for intersection contributing factors.

Countermeasure	Smaller Curve Radius (Intersection on Curve)	Wider Mainline Lane Width	Wider Mainline Median Width	Absence of Mainline Left-Turn Channelization	Absence of Minor-Street Right-Turn Channelization (Signalized Intersections)	Design Speed/Higher Speed Limit
Add left- and right-turn lanes	—	—	—	●	●	●
Increase yellow change intervals	—	—	—	—	—	●
Add backplates with retroreflective borders	—	●	●	—	—	●
Apply multiple low-cost countermeasures	●	●	●	●	●	●
Install advance signs	●	●	●	—	—	●

—Countermeasure does not address contributing factor.
●Countermeasure addresses contributing factor.

Table 91. Systemic countermeasures for nonintersection contributing factors.

Countermeasure	Larger Percent Grade	Narrower Lane Width	Narrower Paved Surface Width	Narrower Shoulder Width	Unpaved Shoulder	Design Speed/Higher Speed Limit	Mountainous Terrain	Smaller Curve Radius
Install SafetyEdge	●	●	●	●	●	●	—	●
Install rumble strips/stripes	—	●	●	●	●	●	●	●
Enhance/improve friction for horizontal curves	●	—	●	—	—	●	—	●
Enhance delineation for horizontal curves	—	—	●	—	—	●	—	●
Implement roadside design improvements at curves	—	—	—	—	—	●	—	●
Install advance markings for curves	—	●	●	●	●	●	—	●
Advance signs	—	●	●	●	—	●[a]	●	●

—Countermeasure does not address contributing factor.
●Countermeasure addresses contributing factor.
[a]Refer to the MUTCD for advance-warning-sign requirements based on speed differentials (FHWA 2012).

The following sections provide brief descriptions of the countermeasures for the intersection contributing factors in table 90. For additional information on these countermeasures, refer to the *Quick Reference Guide* that accompanies this report (Porter et al. 2020). The CMF Clearinghouse provides additional intersection safety countermeasures, as well as other sources (FHWA 2018a).

Auxiliary Turn Lanes (FHWA 2017a)

Providing auxiliary turn lanes at intersection approaches separates slow-moving turning traffic from through traffic and provides vehicles space for deceleration prior to turning. With these improvements, auxiliary turn lanes can reduce the potential for turning vehicles to be involved in ANG and rear-end crashes. This treatment is recommended when there is an absence of mainline left-turn channelization, minor-street right-turn channelization, and a high design speed on the approaching roadway(s). Additionally, this treatment should be considered for major road approaches to intersections with minor stop control where there are operationally warranted turning-vehicle volumes and a history of crashes involving turning vehicles. When considering auxiliary turn lanes, the impact on pedestrian and bicycle safety and convenience should be considered.

Yellow Change Intervals (FHWA 2017a)

Evaluating and adjusting yellow change intervals can lead to reduced red-light-running crashes (which can arise from yellow change intervals that are too long) and rear-end crashes (which can arise from yellow change intervals that are too short). Agencies should regularly reevaluate yellow change intervals at all signalized intersections within their system, especially those with histories of rear-end crashes and red-light-running crashes. This treatment is recommended for intersections with high design speeds and/or speed limits on at least one approach.

Backplates with Retroreflective Borders (FHWA 2017a)

Adding backplates with retroreflective borders to signal heads can improve signal visibility by providing a contrasting background. These visibility improvements are evident in both daytime and nighttime conditions and can improve visibility for all drivers, especially older or color-blind drivers. This treatment can reduce crashes caused by signal-visibility issues, such as red-light-running and rear-end crashes. Contributing factors for which this treatment is recommended include intersections with wide mainline roadways and/or median widths, as well as those with high design speeds and/or speed limits on at least one approach.

Application of Multiple Low-Cost Countermeasures (FHWA 2017a)

Crashes can be reduced system wide by deploying a suite of low-cost countermeasures at multiple stop-controlled intersections. These countermeasures include improving pavement markings and signage. By improving the visibility of signage and pavement markings, driver awareness is increased and the potential for crashes due to driver inattention or confusion is reduced. Systemically, these improvements should be considered for all stop-controlled intersections and can be effective at intersections with any contributing factors.

Advance Signs (ITE 2015)

Installing advanced intersection warning signs, especially for unsignalized intersections, can reduce intersection crashes by warning approaching drivers of intersections or intersection traffic-control devices for which there is limited or inadequate visibility. These signs have the potential to reduce right-ANG and rear-end crashes on both major and minor road approaches. Specific signage for various intersection configurations and traffic-control devices can be found in the MUTCD (FHWA 2012). Contributing factors for which these signs can be effective include intersections on curve radii, intersections with wide mainline roadways and/or median widths, and high design speeds and/or speed limits on at least one approach.

SafetyEdge (FHWA 2017a)

Sharp pavement drop-offs at the pavement edge can exacerbate roadway departures. Adding a SafetyEdge—a gradual sloped drop from the edge of the paved surface to the roadside 30 degrees from the cross-slope—when paving can help drivers maintain control of their vehicles when departing the roadway. This treatment only has minimal impact on the cost of paving or resurfacing projects while providing a proven reduction in roadside crash frequency and severity. Contributing factors for which this treatment should be considered include tight curve radii, narrow traveled ways and paved surfaces, narrow and/or unpaved shoulders, and high design speeds and/or speed limits.

Rumble Strips or Stripes (FHWA 2017a)

Driver inattention, whether through distraction, drowsiness, or another reason, can lead to lane and roadway departures, both of which can lead to crashes, often with severe outcomes arising from a collision with oncoming traffic or roadside fixed objects. Installing rumble strips or stripes along the centerline or edgeline of a roadway can provide an audible alert to drivers straying from the traveled way. This treatment is a low-cost and proven countermeasure that can provide significant benefits if installed system wide. However, agencies need to consider bicyclists when installing shoulder rumble strips; regular gaps to allow for bicyclists to transition between the traveled way and shoulder should be provided where there is regular bicycle traffic. Contributing factors for which this treatment should be considered include tight horizontal radii; narrow traveled ways, shoulder widths, and paved surfaces; high speed limits; and mountainous terrain (e.g., steep grades and significant vertical curvature).

Enhanced Friction for Horizontal Curves (FHWA 2017a)

Insufficient friction on horizontal curves can lead to crashes. If the existing pavement is unable to provide the required friction for cornering, vehicles may lose traction and slide, sometimes off the roadway or into opposing traffic. High-friction surface treatments and pavement grooving (for concrete pavements) have both been found to improve friction on horizontal curves and reduce crash frequency by modifying the existing surface of the pavement. These treatments can be especially effective on curves with tight radii, poor cross-slope design, poor drainage, low pavement friction, and a history of driving too fast for the design conditions of the curve. Of that list, tight curve radii and high speeds are both contributing factors that can be identified systemically.

Enhanced Delineation for Horizontal Curves (FHWA 2017a)

Horizontal curves present a challenge to drivers in that they are a deviation from the existing vehicle path. Drivers who are inattentive or unable to see the change in alignment may fail to navigate the curve, possibly leading to KA crashes, such as ROR, fixed-object, and HEO crashes. Improving the delineation of the curve assists drivers with seeing and preparing for the change in alignment. Using one or a combination of pavement markings, post-mounded delineators, larger signs, more retroreflective signs, dynamic curve warning signs, and sequential curve signs can lead to a reduction in crash severity on a curve. Systemically, curves with small curve radii or high speed limits should be considered for this treatment.

Roadside-Design Improvements at Curves (FHWA 2017a)

Roadway departures are almost unavoidable on horizontal curves. Given the inevitability, making the roadside more forgiving to vehicles departing the roadway can reduce the severity of these events. Improvements to the roadside that have proven reliably effective include widening the clear zone, flattening side slopes, and paving and widening shoulders. Additionally, if the clear-zone width is unable to be obtained, the use of cable barriers, guardrails, or concrete barriers (as appropriate) can also reduce the severity of outcomes for departing vehicles. Contributing factors for which this treatment should be considered include tight curve radii, narrow shoulder widths, high speed limits, and mountainous terrain (e.g., steep grades and significant vertical curvature).

Advance Markings for Curves (Albin et al. 2016)

Various pavement markings can be used to supplement horizontal-curve treatments provided in FHWA's *Proven Safety Countermeasures* (FHWA 2017). A speed-advisory marking can be painted within the travel lane to provide additional guidance in drivers' line of sight. Optical speed bars spaced at gradually decreasing distances on the approach to the curve can manipulate drivers' speed perception and influence them to slow down. Advisory pavement markings are most appropriate for curves with advisory speeds much lower than the regular posted speed limit on the roadway; optical speed bars are most appropriate on isolated or unexpected horizontal curves. Contributing factors for which this treatment should be considered include tight radii and high design speeds.

Advance Signs (Albin et al. 2016)

A low-cost treatment to improve safety on horizontal curves is the installation of static advance curve warning signs. Warning signs at horizontal curves include curve warning signs and advisory speed-limit signs. These signs make drivers aware of the impending change in alignment and suggested speed at which the curve should be navigated. The MUTCD provides guidance for the installation of these signs (FHWA 2012). Contributing factors for which static advance signage should be considered include steep grades and tight curve radii.

The treatments described above and summarized in table 90 and table 91 represent a set of countermeasures that have been proven to successfully reduce crash frequency or severity at a relatively low cost, leading them to be recommended by FHWA and ITE. For more details on these treatments, refer to the *Quick Reference Guide* accompanying this report (Porter et al.

2020). For additional countermeasures, refer to resources such as the CMF Clearinghouse (FHWA 2018a).

CHAPTER 7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The main goal of this project was to identify focus crash types, focus facility types, and associated contributing factors to inform applications of systemic safety improvements. To achieve this goal, the research team executed three main technical tasks: identifying FCFTs, identifying contributing factors associated with the FCFTs, and developing a process to assist safety practitioners in making informed choices regarding countermeasures to address the focus crash types.

IDENTIFYING FCFTS

Selecting FCFTs was guided by the use of FARS and HSIS (NHTSA 2018a; FHWA 2018c). The method used by the research team to select potential FCFTs was based on the number of both K and KA crashes associated with the combination of various crash-related variables, such as location, crash type, area type, roadway type, intersection type, traffic control type, lighting, and road alignment. A case study showed that the method employed was analogous to the steps for selecting FCFTs outlined in FHWA's *Systemic Safety Project Selection Tool* (Preston et al. 2013a). Its application by the research team resulted in a total of 17 FCFTs (8 intersection FCFTs and 9 nonintersection FCFTs) for analysis to identify contributing factors for each FCFT:

- ROR crashes on rural two-lane roads on horizontal curves.
- ROR crashes on rural two-lane roads on straight segments.
- LNDP crashes on rural two-lane roads on horizontal curves.
- LNDP crashes on rural two-lane roads on straight segments.
- HEO crashes on rural two-lane roads on straight segments.
- ANG crashes on rural two-lane roads on straight segments.
- HEO crashes on rural two-lane roads on horizontal curves.
- ROLL crashes on rural two-lane roads on straight segments.
- ROLL crashes on rural two-lane roads on horizontal curves.
- ANG crashes on rural two-lane roads at four-leg stop-controlled intersections.
- ANG crashes on urban two-lane roads at four-leg stop-controlled intersections.
- ANG crashes on urban multilane divided roads at four-leg signalized intersections.
- ANG crashes on urban multilane undivided roads at four-leg signalized intersections.
- ANG crashes on rural two-lane roads at three-leg stop-controlled intersections.
- ANG crashes on rural multilane divided roads at four-leg stop-controlled intersections.
- All pedestrian crashes at intersections.
- Crashes involving a pedestrian crossing at an intersection struck by a motorist going straight.

Original analyses conducted as part of this research was the basis for identifying contributing factors for crash types involving only motor vehicles. Identifying contributing factors for pedestrian crashes was challenging without quality exposure data. The research team explored additional, original analyses of pedestrian crashes using available data, but were not confident that a quality analysis of contributing factors could be conducted with existing data, which did not include pedestrian exposure and included some questionable crash-location coding. Instead, a previous study by Thomas et al. (2017), which looked at identifying and screening intersection locations that have the potential for future pedestrian crashes, served as the basis for identifying contributing factors for pedestrian crashes.

IDENTIFYING CONTRIBUTING FACTORS ASSOCIATED WITH FCFTS

The research team defined contributing factors as factors whose presence are associated with increases or decreases in expected frequencies of crashes or injury severities resulting from crashes. The research team used random forests to identify contributing factors corresponding to selected FCFTs. As noted in chapter 1, a previous research effort, *Highway Safety Statistical Paper Synthesis*, demonstrated the use of CART and random forests within the context of conducting statistical road-safety analyses (Persaud et al. 2001). The study concluded that tree-based models have strong potential for effective use in road-safety analyses. This current project extended the exploration of potential applications of tree-based methods within the context of identifying contributing factors for systemic safety analysis.

The research team used three data sources to conduct the contributing-factor analysis: crash and roadway inventory from HSIS, climate data from the NOAA, and socioeconomic census data from the U.S. Census Bureau (FHWA 2018c; NOAA 2018; U.S. Census Bureau 2018). All linkage of road segments to climate and census data was done in the spatial environment. To simplify the joining process, roadway segments were represented as point features according to the midpoint of the segment. For climate data, the source data were in a point file, with each weather station shown as a point on the map. Each roadway segment was linked to the closest weather station by a simple straight-line distance measurement. For census data, the source data were in a polygon file, with each census-block group shown as a shape on the map. Each roadway segment was linked to the census-block group that contained the midpoint of the segment.

The research team used HSIS data from California and Ohio for the intersection contributing-factor analysis and HSIS data from Ohio and Washington for the nonintersection contributing-factor analysis. The research team conducted a detailed analysis of crash-severity groupings and found that the results of the contributing-factor analyses were quite similar when using either the frequencies of KAB or KABCO crashes. In almost all cases, the most influential predictor variables were consistent across the two severity levels, except for the changes in the increase of predictions of MSE associated with the variables.

Table 92 lists the most prevalent factors that were found to influence expected crash frequencies from each of the three main categories: roadway, socioeconomic, and climate. Findings regarding these factors are interpreted in this project as predictive relationships or statistical associations with expected crash frequencies. The following characteristics raise confidence that a specific finding or set of findings are stable and transferable:

- Consistency across subsets of related FCFTs.
- Consistency across multiple States.
- Consistency with previous findings in the literature.

Knowledge related to the safety impacts of traffic and roadway variables has grown substantially over the last two decades and offers a basis to interpret the results of this effort. In most cases, the traffic- and roadway-related findings were consistent with published literature and practice for related crash types.

There is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results of this effort. Findings related to socioeconomic variables likely represent differences in travel behavior, driving behavior, and driving capabilities that seem key for safety analyses but are generally not incorporated into analyses that also include traffic and roadway factors. Weather-related findings likely represent differences in visibility, road conditions, and driver experience and behavior. In most cases, findings related to socioeconomic and weather variables set the stage for future analyses, possibly focused solely on these variables.

Table 92. Most influential factors found to influence expected crash frequencies.

Variable Category	Factor
Roadway: intersection	Mainline AADT
Roadway: intersection	Cross street AADT
Roadway: intersection	Design speed/speed limit
Roadway: intersection	Lane width
Roadway: intersection	Median width
Roadway: intersection	Left-turn channelization
Roadway: intersection	Curve radius
Roadway: nonintersection	Average AADT
Roadway: nonintersection	Percent grade
Roadway: nonintersection	Curve radius
Roadway: nonintersection	Surface width
Roadway: nonintersection	Shoulder type
Roadway: nonintersection	Terrain
Socioeconomic	Percentage of population ages 16–24 working full time
Socioeconomic	Percentage of population ages 16–24 working part time
Socioeconomic	Percentage of population ages 25+ without a high school diploma
Socioeconomic	Percentage of population ages 25+ with a high school diploma but no university degree
Socioeconomic	Percentage of households with income <$50,000
Socioeconomic	Percentage of households with 0 vehicles
Socioeconomic	Percentage of households with 1 vehicle
Socioeconomic	Percentage of households with ≥2 vehicles
Socioeconomic	Percentage of population ages 15–19
Socioeconomic	Percentage of population ages 45–64
Climate	Average annual snowfall totals
Climate	Average annual rainfall totals
Climate	Average annual maximum temperatures
Climate	Average annual number of days with a minimum temperature of ≤32°F

There were some counterintuitive results where a variable was expected to increase crash frequency but instead was found to decrease crash frequency or vice versa. In most cases, counterintuitive findings would show up in only a single State. For example, increasing the average shoulder width appeared to increase the frequency of LNDP-N crashes on rural two-lane roads in Washington. However, the same increase in the average shoulder width decreased (as expected) the frequency of LNDP-N crashes on rural two-lane roads in Ohio.

Counterintuitive findings can happen for different reasons:

- Data related to the presence and type of safety countermeasures were not available for the analysis. For example, an agency may apply chevrons on certain types of curves based on curve characteristics (e.g., crash history of a specific curve, radius of a curve). Similarly, an agency may have installed centerline or shoulder rumble strips on some corridors. That type of countermeasure information was not available to the research team for this analysis.

- Details on specifics of some roadway features were not available. The research team had information on shoulder widths and types (i.e., paved or unpaved) for Washington. However, for paved shoulders, it was not clear if the complete shoulder was paved or only part of the shoulder was paved. Data on the nature of the roadside were also not available. The road safety research and practitioner communities generally believe that roadside design features and hazards are significantly associated with the frequency and severity of several of the focus crash types on rural two-lane roads. However, limitations in roadside data have made the ability to quantify these relationships elusive.

- In linking census data to roadway data, there was an implicit assumption that the characteristics of the census block represents the characteristics of the drivers at the sites investigated in this project. There was no additional analysis to determine whether this assumption is reasonable. The findings related to the socioeconomic variables are likely representing some complex interaction of travel behavior, driving behavior, and driving capabilities. This is an area that needs additional work.

- There is a possibility of an unexpected interaction of driver behavior with certain roadway factors or combinations of roadway factors that this analysis does not fully consider or explain. There are multiple examples of these unexpected interactions in road safety research, including those highlighted by Smiley (2008).

Information about these factors could have helped further explore the counterintuitive findings. Without full knowledge of the reason behind such counterintuitive findings, the research team did not consider these findings in providing guidance on the identification of contributing factors and selection of countermeasures. The guidance developed for practitioners focused on traffic and roadway findings that were generally consistent across crash types and States, as well as with related literature.

PROCESS FOR SELECTING COUNTERMEASURES

The research team laid out a process for identifying and selecting countermeasures for focus crash types based on identified contributing factors. This process recognizes that, although there are certain contributing factors associated with focus crash types, selecting countermeasures must be broad and encompass many options for addressing the crash type. The usefulness of the contributing-factor identification is to allow a practitioner to identify which countermeasures specifically address those factors as part of their assessment of potential safety effectiveness.

This practitioner identification may in turn raise the priority for selecting and implementing those countermeasures. The process is laid out in the following six steps:

1. Identify a focus crash type.
2. Identify contributing factors for the focus crash type.
3. Assemble a list of potential countermeasures that address the focus crash type.
4. Identify countermeasures that address the contributing factors associated with the focus crash type.
5. Identify countermeasures with CMFs.
6. Select a countermeasure.

Chapter 6 provides several examples to assist practitioners with implementing this process and descriptions of proven safety countermeasures for common crash types, facility types, and contributing factors.

CONCLUSIONS AND RECOMMENDATIONS

The following sections provide conclusions of this research and recommendations for future work with respect to identifying FCFTs, identifying contributing factors, and implementing and evaluating countermeasures applied as part of a systemic safety-management approach.

Identifying FCFTs

The methodology developed and applied by the research team to select potential FCFTs based on the number of K crashes across the United States, as well as the number of KA crashes in four selected States, resulted in a useful list of FCFTs. Agencies with sufficient data and analysis capabilities can refer to FHWA's *Systemic Safety Project Selection Tool* for discussion about how to analyze data to identify jurisdiction-specific FCFTs (Preston et al. 2013a). Agencies with limited data or analysis capabilities can refer to State, regional, or local SHSPs to identify focus crash types based on emphasis areas. Agencies with limited data or analysis capabilities can also refer to the list of FCFTs developed for this research, which contribute to a significant number of fatalities and serious injuries across the United States and represent common priorities.

During initial stages of this research, the research team disaggregated selected FCFTs by time of day (e.g., ROR-D crashes on rural two-lane roads on horizontal curves; ROR-N crashes on rural two-lane roads on horizontal curves). This approach was to allow the analysis to uncover the possibility of different factors influencing the frequency of the same crash type during the daytime or nighttime. This disaggregation did not lead to any key insights, likely due to the unavailability of traffic information by time of day. Future work should continue to explore this direction, as it seems logical that at least some contributing factors would differ by time of day. Acquiring traffic exposure by time of day will be critical to this analysis. Musunuru et al. (2017) provided one option for estimating day and night traffic volumes on rural two-lane roads in the absence of data.

Identifying Contributing Factors

A review of systemic safety practices shows that, to date, systemic safety analyses often include subjective approaches to identifying contributing factors and characterizing crash potential. Agencies need additional information and guidance to help select and target their systemic safety improvements to make the most of limited funds, including detailed and data-driven information on situations characterized by contributing factors and facility types where focus crash types are more likely to occur. The task is challenging, as the nature of specific crash types that are the target of systemic safety management is not necessarily compatible with statistical analyses of contributing factors that require significant sample sizes. The following conclusions and recommendations pertain to the methodologies and findings with respect to roadway-, socioeconomic-, and weather-related factors.

Methodologies

The research team explored the use of random forests to identify contributing factors for FCFTs and noted both strengths and limitations. The strengths of the technique include the following:

- Unlike traditional regression methods, the random-forest method does not require any predefined underlying relationship between the target (dependent) variable and predictors (independent variables).
- Unlike traditional regression methods, the random-forest method does not require any formal distributional assumptions.
- Strobl et al. (2009) noted that high-ranked variables on a random forest may appear on that list as a result of a complex interaction that cannot be captured in a traditional regression model.
- Random forests provided estimates of what independent variables were most predictive in their relationship with the dependent variable. Rossi et al. (2005) concluded that random-forest variable rankings with respect to predictive capabilities are more stable than those produced by stepwise logistic regressions.
- In addition to predictive rankings, random forests also provided the percentage increase in MSE that would result from removing a particular variable from the analysis.

The weaknesses of the technique include the following:

- When random forests are used for regressions, they cannot predict beyond the range in the training data.
- Random forests have been found to overfit datasets that are particularly noisy (Segal 2004).
- For data including categorical variables with different numbers of levels, random forests are biased in favor of attributes with more levels. Therefore, variable rankings with respect to predictive capabilities from random forests are not always reliable for this type of data.
- Random forests are primarily a predictive modeling tool, not a descriptive tool, even though they can be used to learn about underlying mechanisms.
- For large datasets, the model size for random forests can be quite large, using hundreds of megabytes of memory.

- Random forests can be black boxes that are difficult to interpret.
- In our application of random forests for segments, the dependent variable was the number of crashes per mile, which was calculated as the number of crashes in a segment divided by the length of the segment. The results could be unexpectedly and unduly influenced by short segments.

The most challenging aspect of implementing random forests within this context was in the limited ability to interpret the direction and form of the relationship between a factor of interest and expected crash frequency. This research implemented a relatively simple approach, searching for primarily linear trends between each factor and the random forest–predicted crash frequency. Future efforts should explore more effective ways to uncover and interpret relationships contained in random forests.

While the research team compared random-forest results to findings of previous research on related crash types, it was not able to incorporate prior information or knowledge into the analyses. Bayesian approaches hold significant promise, and future efforts should explore contributing factors for FCFTs. Causal Bayesian networks, such as those applied by Karwa et al. (2011), have significant potential but will need to be further explored.

Roadway, Socioeconomic, and Weather Factors

Roadway factors uncovered by the analyses as influencing frequencies of crash types were generally consistent with the research team's expectations based on previous research and existing practice. Factors associated with higher crash frequencies included the following:

- Larger ADT volumes.
- Steeper vertical grades.
- Sharper curve radii.
- Narrower lane and shoulder widths.
- Unpaved shoulders or no shoulders.
- Mountainous terrain.
- Higher speed limits.
- Wider crossing distances at intersections (captured by lane and median widths on approaches).
- Absence of left- and right-turn channelization at intersections.

Agencies with sufficient data and analysis capabilities can refer to the FHWA *Systemic Safety Project Selection Tool* for discussion of how to analyze data to identify contributing factors given a specific FCFT (Preston et al. 2013a). Agencies without sufficient data and analysis capabilities can reference factors developed in this research to help identify countermeasures and prioritize sites for systemic safety improvements.

Findings related to socioeconomic- and weather-related factors showed promise, but there is not yet a significant amount of theory to support or refute the socioeconomic- and weather-related results of this effort. Findings related to socioeconomic variables are likely representing differences in travel behavior, driving behavior, and driving capabilities that seem key for safety analyses but are generally not incorporated into segment- and intersection-specific analyses that also include traffic and roadway factors. Weather-related findings are likely representing differences in visibility, road conditions, and driver experience and behavior. Both sets of factors bring significant potential to the process of making more informed decisions about sites that have higher levels of crash potential. (Appendix H provides examples, including an analysis of model performance with and without consideration of socioeconomic variables.) A multiyear study focused on testing various alternatives and developing safety-analysis guidance on collecting, merging, and analyzing crash, traffic, roadway, census, and weather data is needed.

As with many road-safety analyses, the dataset for this project consisted of correlated variables, particularly various socioeconomic characteristics. The research team conducted an additional analysis that used factor analysis to reduce dimensions of socioeconomic variables. As part of an exploratory effort, the research team developed and interpreted factors (e.g., a low socioeconomic status factor that weighs low income, no vehicle ownership, unemployment, and so on) and then included the factors in the random-forest analysis in place of individual variables. This approach was based on a study from transit ridership literature that reduced the dimensionality of a set of possibly relevant predictor variables from the same category (e.g., land use) prior to implementing a regression-tree approach. The approach and findings were promising but did not significantly change the overall conclusions when compared to the original analysis. Appendix G provides details of this analysis. Future efforts should explore the contributing factor–regression tree approach for correlated roadway, socioeconomic, and weather factors.

Implementing and Evaluating Countermeasures

The research team laid out a process for identifying and selecting countermeasures for focus crash types based on contributing factors and identified and proven safety countermeasures for common contributing factors. In doing so, the research team observed that most CMFs have not been developed for crash types at the level of disaggregation that may be needed for systemic safety applications. Similarly, expected changes in crash frequencies reflected by CMFs typically represent results of evaluations of site-specific applications of a treatment. Additional work on estimating network-wide safety improvements resulting from systemic applications of countermeasures is needed.

APPENDIX A. LIST OF VARIABLES USED FROM FARS AND HSIS FOR FCFT SELECTION

The following sections detail variables the research team used from FARS and HSIS for California, Ohio, and Washington for FCFT selection (NHTSA 2018a; FHWA 2018c).

LIST OF FARS VARIABLES USED FOR FCFT SELECTION

Common Variables

- State number: State in which the crash occurred.
- Consecutive number: Unique case number assigned to each crash.
- Vehicle number: Number assigned to each vehicle in the case.
- Person number: Number assigned to each person in the case.

Accident Data File

- Number of vehicle forms: Number of contact motor vehicles involved in the crash.
- Number of forms submitted for persons in motor vehicles: Count of the number of persons in motor vehicles that are applicable to the case.
- Number of forms submitted for persons not in motor vehicles: Count of the number of persons not in motor vehicles that are applicable to the case.
- Month of crash: Month in which the crash occurred.
- Day of crash: Month on which the crash occurred.
- Day of week: Day of the week on which the crash occurred.
- Year of crash: Year in which the crash occurred.
- Hour of crash: Hour in which the crash occurred.
- Minute of crash: Minutes after the hour at which the crash occurred.
- First harmful event: First injury or damage-producing event of the crash.
- Manner of collision: Orientation of two motor vehicles in transport when they were involved in the crash.
- Relation to junction—specific location: Location of the crash with respect to its presence in or proximity to components in a junction or interchange area.
- Type of intersection: Various intersection type in which the crash occurred.
- Relation to trafficway: Identifies the location of the crash as it relates to its position on the trafficway.
- Work zone: If the crash occurred within the boundaries of a work zone.
- Light conditions: Type and level of light at the time of the crash.
- Atmospheric conditions: Prevailing atmospheric conditions that existed at the time of the crash.
- Drunk drivers: Number of drunk drivers involved in the crash.
- Fatalities: Number of fatally injured persons in the crash.
- Roadway function class: Functional classification of the trafficway on which the crash occurred.

Vehicle Data File

- Initial contact point: Area on the vehicle that produced the first instance of injury to nonmotorists or damage to other property or to the vehicle itself.
- Most harmful event: Event that resulted in the most severe injury.
- Trafficway description: Trafficway flow just prior to the vehicle's critical precrash event.
- Total lanes in roadway: Number of travel lanes prior to the vehicle's critical precrash event.
- Roadway Alignment: Roadway alignment prior to the vehicle's critical precrash event.
- Roadway grade: Roadway grade prior to the vehicle's critical precrash event.
- Roadway-surface condition: Roadway-surface condition prior to the vehicle's critical precrash event.
- Traffic control device: Traffic controls in the vehicle's environment prior to the vehicle's critical precrash event.

Person Data File

- Age: Person's age at the time of the crash in years.
- Sex: Sex of the person involved in the crash.
- Person type: Role of the person involved in the crash.
- Injury severity: Severity of injury (using the KABCO scale) to the person in the crash.
- Police-reported alcohol involvement: If alcohol was involved for the person.
- Police-reported drug involvement: If drugs were involved for the person.

LIST OF HSIS WASHINGTON VARIABLES USED FOR FCFT SELECTION

Common Variables

- Accident number: Case number of the accident.
- Roadway inventory: Crash location information (i.e., county, route, and milepost) used in linkage to other files.
- Accident reference point: Reference point where the crash occurred.
- Beginning milepost: Beginning milepost of the road segment on which the crash occurred.
- Ending milepost: Ending milepost of the road segment on which the crash occurred.

Accident Data Files

- Accident type: Type of accident that occurred.
- Light conditions: Type and level of light at the time of the crash.
- Relation to intersection: Location of the crash in relation to the intersection at which the crash occurred.
- Collision type: Types of first and second collisions in the crash.
- Roadway surface: Condition of the road surface where the crash occurred.
- Number of pedestrians/cyclists: Number of pedestrians/cyclists involved in the crash.
- Crash severity: Most severe injury that resulted from the crash.

- Number of persons injured: Total number of persons injured in the crash.
- Total killed: Total number of persons killed in the crash.
- Weather conditions: Weather conditions at the time the crash occurred.
- Work zone status: Work zone details if the crash occurred in a work zone.
- Object struck: Fixed object struck in the crash.
- Driver's age: Age of the driver of the vehicle involved in the crash.
- Driver's injury: Extent of injury to the driver of the vehicle involved in the crash.
- Driver's sex: Sex of the driver of the vehicle involved in the crash.

Roadway and Curve Data Files

- Calculated average AADT: Calculated average AADT of the location where the crash occurred.
- Control of access: Access control in place at the location where the crash occurred.
- Functional class: Functional class of roadway segment on which the crash occurred.
- Median barrier type: Type of median barrier on the roadway segment on which the crash occurred.
- Total number of lanes: Total number of lanes (in both directions) of the roadway segment on which the crash occurred.
- Roadway classification: Roadway classification of the roadway segment on which the crash occurred.
- Rural/urban identification: Rural or urban classification of the roadway where the crash occurred.
- Curve beginning milepost: Beginning milepost of the curved segment on which the crash occurred.
- Curve ending milepost: Ending milepost of the curved segment on which the crash occurred.
- Grade beginning milepost: Beginning milepost of a grade segment on which the crash occurred.
- Grade ending milepost: Ending milepost of a grade segment on which the crash occurred.
- Percent grade: Percent grade for the roadway segment on which the crash occurred.

LIST OF HSIS OHIO VARIABLES USED FOR FCFT SELECTION

Common Variables

- Accident number: Case number of accident.
- County route: Crash location information used in linkage to other files.
- Accident reference point: Reference point where the crash occurred.
- Beginning milepost: Beginning milepost of the road segment on which the crash occurred.
- Ending milepost: Ending milepost of the road segment on which the crash occurred.

Accident Data Files

- Type of crash (first harmful event): First harmful event in the crash sequence.
- Light conditions: Type and level of light at the time of the crash.
- Relation to intersection: Location of crash in relation to the intersection in which the crash occurred.
- Number of vehicles involved: Number of vehicles involved in the crash.
- Road characteristics: Characteristics of the road on which the crash occurred.
- Number of pedestrians: Number of pedestrians involved in the crash.
- Crash severity: Most severe injury that resulted from the crash.
- Serious visible injury: Total number of A injuries in the crash.
- Minor visible injury: Total number of B injuries in the crash.
- No visible injury: Total number of C injuries in the crash.
- Total killed: Total number of persons killed in the crash.
- Weather conditions: Weather conditions at the time the crash occurred.
- Population: Rural/urban population where the crash occurred.
- Driver's age: Age of the driver of the vehicle involved in the crash.
- Driver's injury: Extent of injury to the driver of the vehicle involved in the crash.
- Driver's sex: Sex of the driver of the vehicle involved in the crash.

Roadway Data Files

- Calculated average AADT: Calculated average AADT of the crash location.
- Control of access: Access control in place at the location of crash.
- Functional class: Functional class of the roadway segment on which crash occurred.
- Median type: Type of median on the roadway segment on which the crash occurred.
- Total number of lanes: Total number of lanes (in both directions) of the roadway segment on which the crash occurred.
- Roadway classification: Roadway classification of the roadway segment on which the crash occurred.

LIST OF HSIS CALIFORNIA VARIABLES USED FOR FCFT SELECTION

Common Variables

- Accident number: Case number of accident.
- County route: Crash location information used in linkage to other files.
- Accident reference point: Reference point where the crash occurred.
- Beginning milepost: Beginning milepost of the road segment on which the crash occurred.
- Ending milepost: Ending milepost of the road segment on which the crash occurred.

Accident Data Files

- Type of collision: Type of accident that occurred.
- Light conditions: Type and level of light at the time of the crash.
- Relation to intersection: Location of the crash in relation to the intersection at which the crash occurred.
- Number of vehicles: Number of vehicles involved in the crash.
- Roadway surface: Condition of the road surface on which the crash occurred.
- Pedestrian involvement: Whether or not a pedestrian was involved in the crash.
- Crash severity: Most severe injury that resulted from the crash.
- Number of persons injured: Total number of persons injured in the crash.
- Total killed: Total number of persons killed in the crash.
- Weather conditions: Weather conditions at the time the crash occurred.
- Primary collision factor: Primary collision factor of the crash.
- Motor vehicles involved: Vehicles or nonvehicles that were involved in the crash.
- Driver's age: Age of the driver of the vehicle involved in the crash.
- Driver's sex: Sex of the driver of the vehicle involved in the crash.

Roadway and Intersection Data Files

- Calculated average AADT: Calculated average AADT of the location where the crash occurred.
- Control of access: Access control in place at the location where the crash occurred.
- Functional class: Functional class of the roadway segment on which the crash occurred.
- Median barrier type: Type of median on the roadway segment on which the crash occurred.
- Total number of lanes: Total number of lanes (in both directions) of the roadway segment on which the crash occurred.
- Roadway classification: Roadway classification of the roadway segment on which the crash occurred.
- Rural/urban identification: Rural or urban classification of the roadway where the crash occurred.
- Traffic control type: Traffic control type at the intersection where the crash occurred.

LIST OF HSIS MINNESOTA VARIABLES USED FOR FCFT SELECTION

Common Variables

- Accident number: Case number of accident.
- Combined route system/route number: Combined route system and route number where the crash occurred.
- Accident reference point: Reference point where the crash occurred.
- Beginning milepost: Beginning milepost of the road segment on which the crash occurred.
- Ending milepost: Ending milepost of the road segment on which the crash occurred.

Accident Data Files

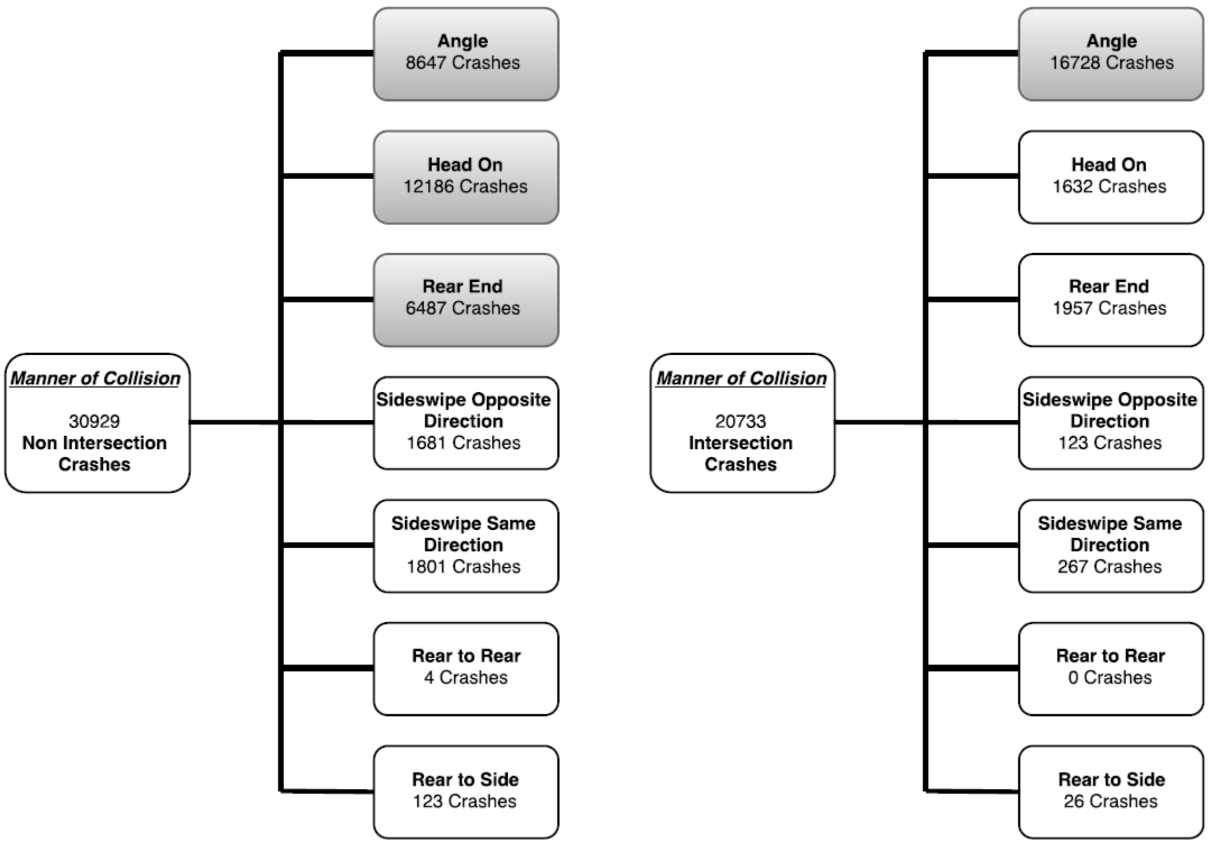
- Diagram of accident code: Accident configuration describing the direction and maneuvers of the vehicles involved in the crash.
- Type of accident: Type of accident that occurred.
- Road design: Design of the roadway where the crash occurred.
- Light conditions: Type and level of light at the time of the crash.
- Relation to intersection: Location of the crash in relation to the intersection.
- Number of vehicles involved: Number of vehicles involved in the crash.
- Road characteristics: Characteristics of the road on which the crash occurred.
- Road surface conditions: Condition of the road surface on which the crash occurred.
- Accident severity: Most severe injury in the crash.
- Posted speed limit: Posted speed limit where the crash occurred.
- Number of persons injured: Total number of persons injured in the crash.
- Number of persons killed: Total number of persons killed in the crash.
- Traffic-control devices: Traffic-control devices where the crash occurred.
- Weather conditions: Weather conditions at the time the crash occurred.
- Work zone marked: Type of work zone where the crash occurred.
- Driver's age: Age of the driver of the vehicle involved in the crash.
- Driver's injury: Extent of injury to the driver of the vehicle involved in the crash.
- Driver's sex: Sex of the driver of the vehicle involved in the crash.

Roadway Data Files

- Calculated average AADT: Calculated average AADT of the crash location.
- Control of access: Access control in place at the location of crash.
- Functional class: Functional class of the roadway segment on which the crash occurred.
- Median type: Type of median on the roadway segment on which the crash occurred.
- Total number of lanes: Total number of lanes (in both directions) of the roadway segment on which the crash occurred.
- Roadway classification: Roadway classification of the roadway segment on which the crash occurred.

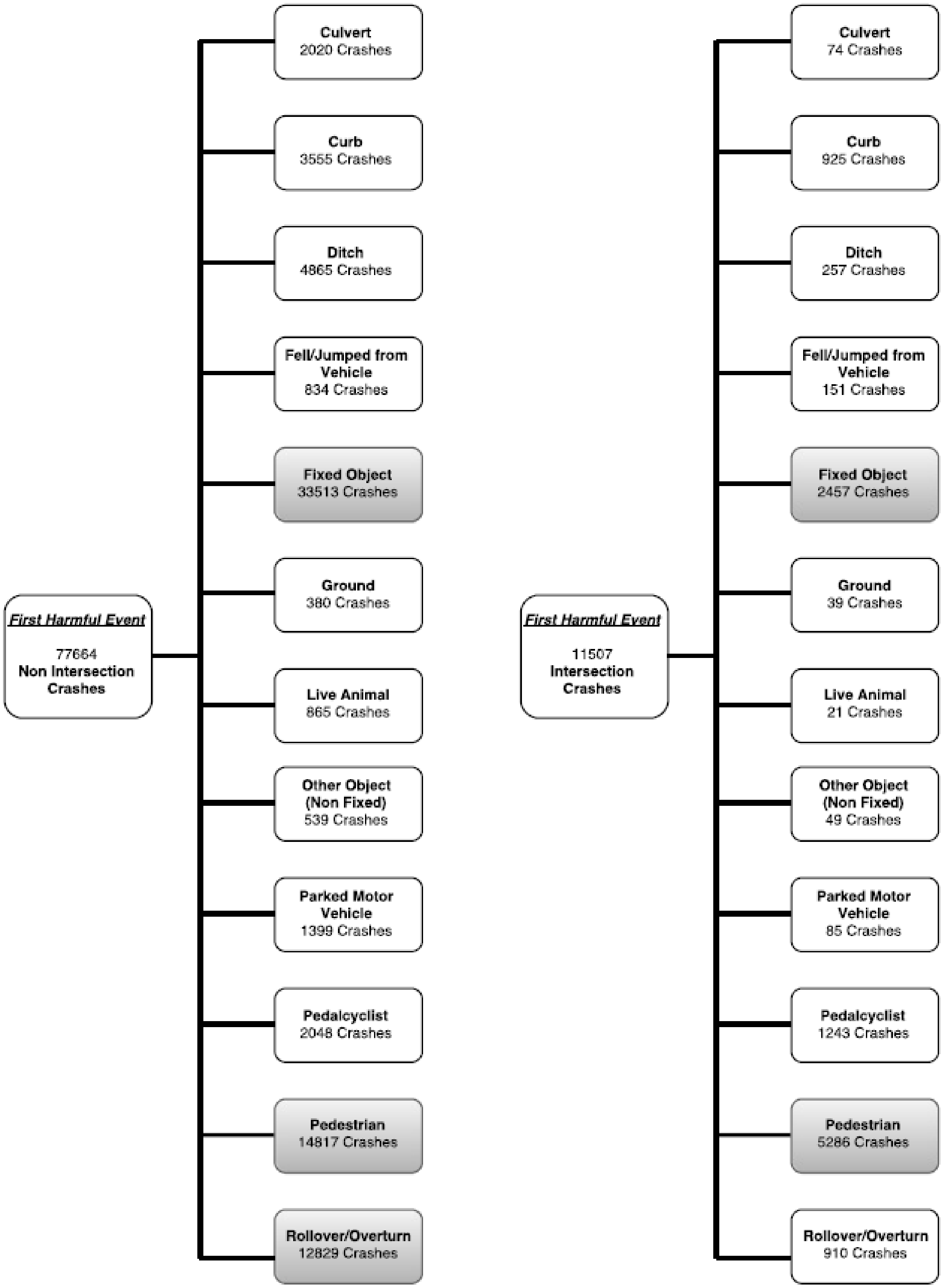
APPENDIX B. SYSTEMIC SAFETY-PLANNING PROCESS

Figure 4 and figure 5 show the selection of focus crash types (i.e., task 1 of the systemic safety-planning process). Each box shows a crash or facility type and the number of corresponding crashes. The shaded boxes are the selected focus crash types. The crash trees for selecting focus facility types for nonintersection crashes (i.e., task 2 of the systemic safety-planning process) are shown in figure 5 through figure 13.



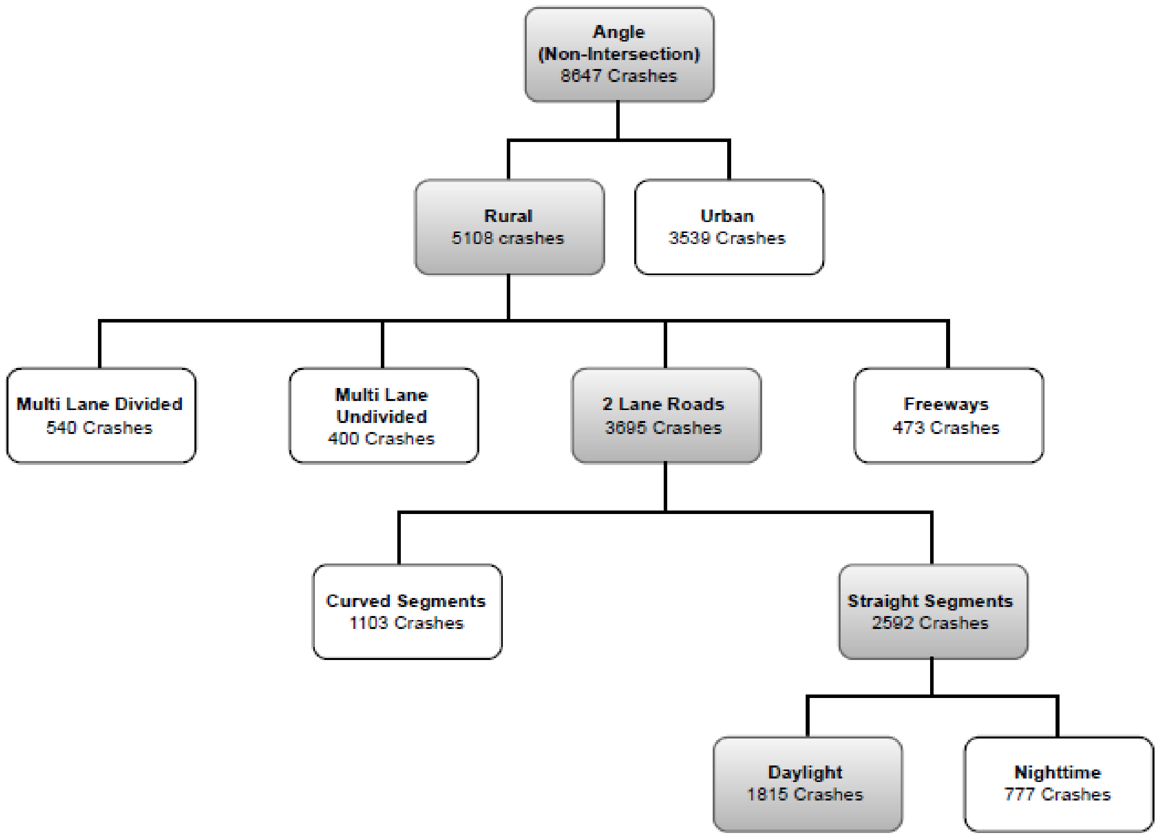
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 4. Chart. Crash types defined by Manner of Collision.



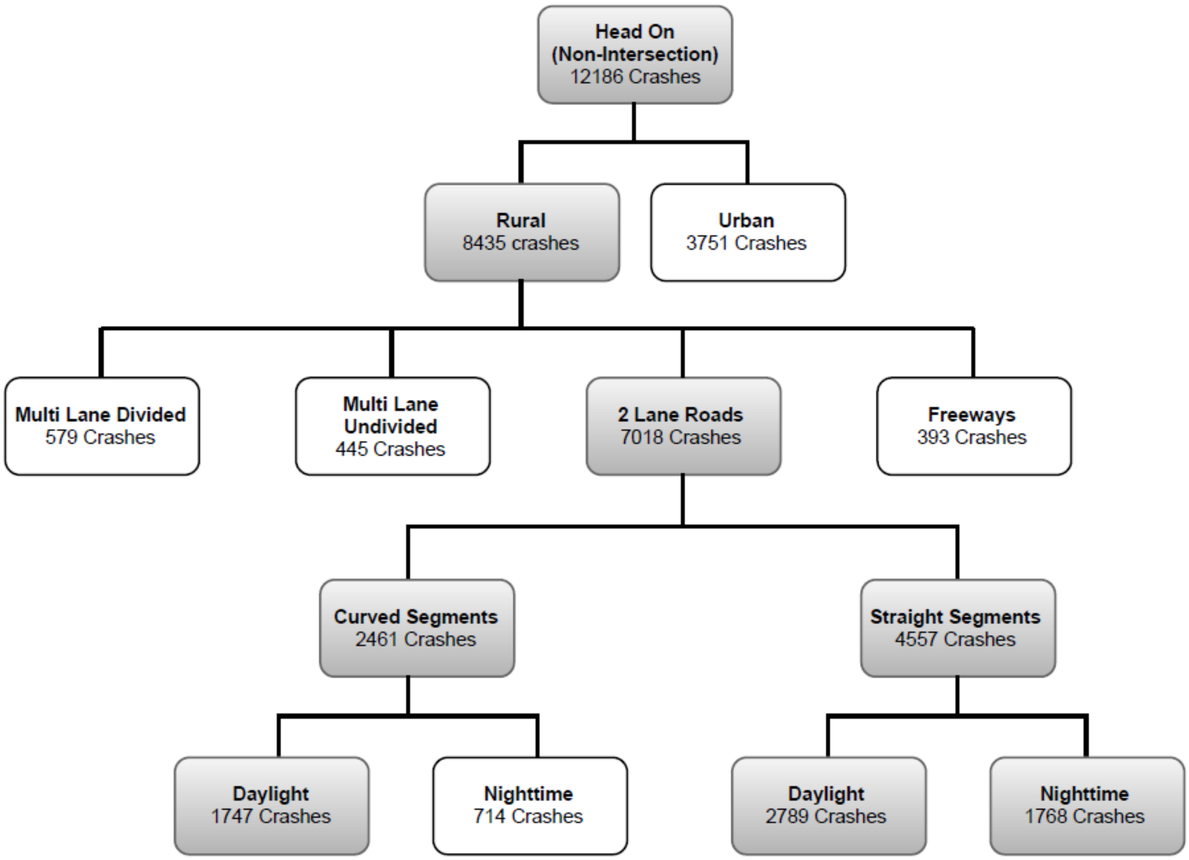
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 5. Chart. Crash types defined by First Harmful Event.



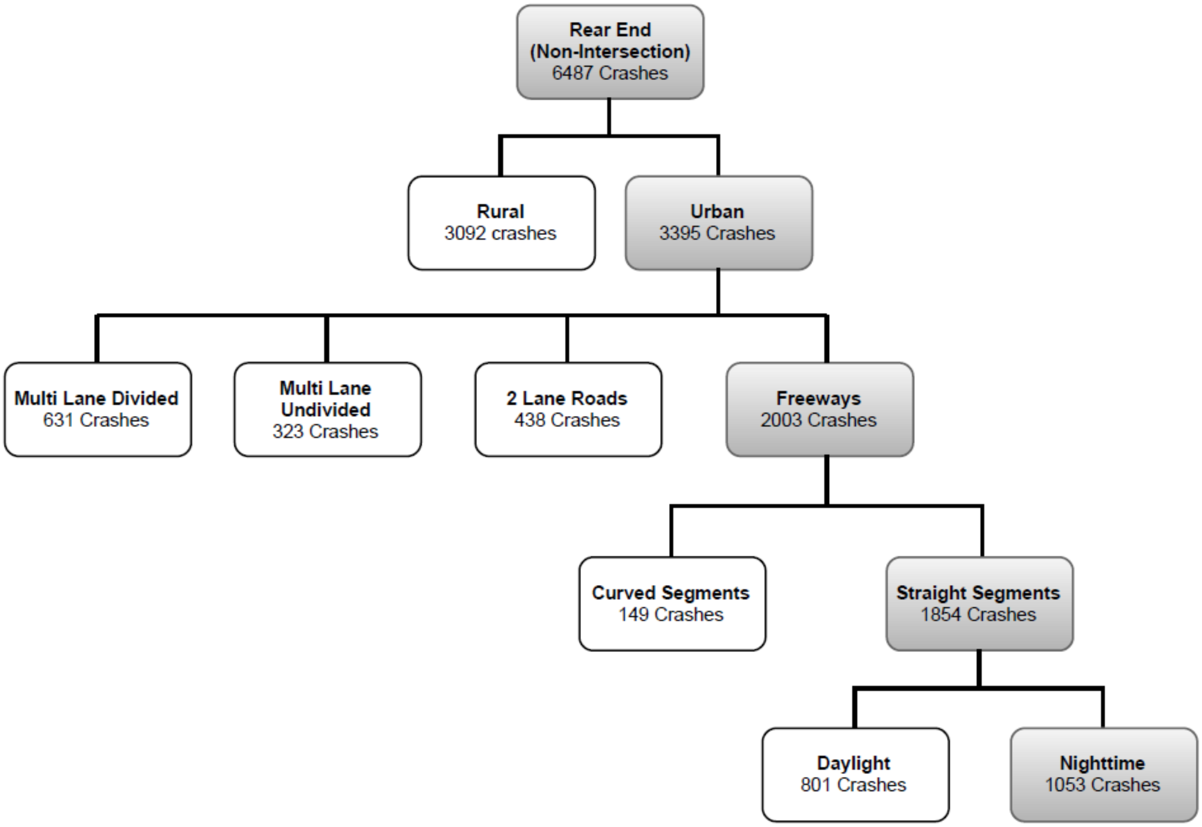
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 6. Chart. Crash tree for ANG crashes at nonintersection locations.



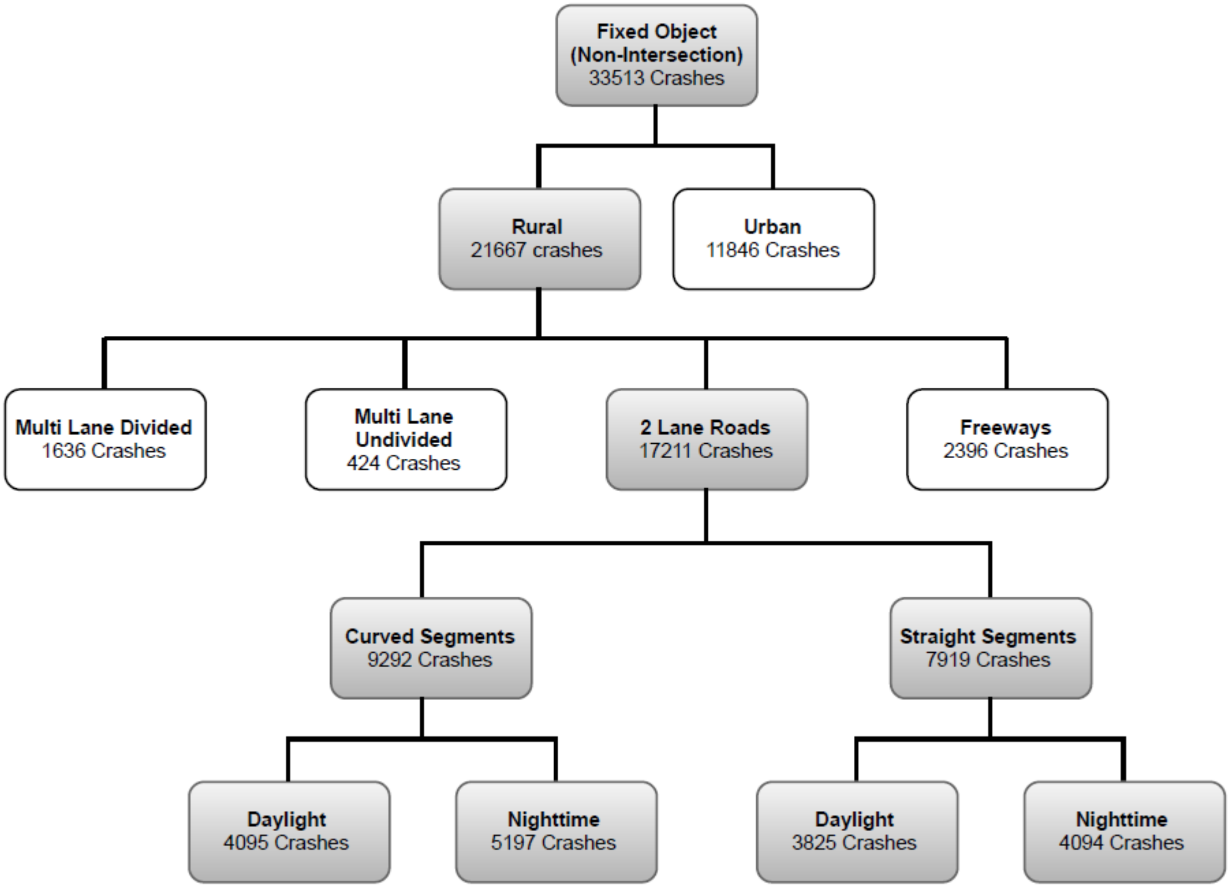
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 7. Chart. Crash tree for HEO crashes at nonintersection locations.



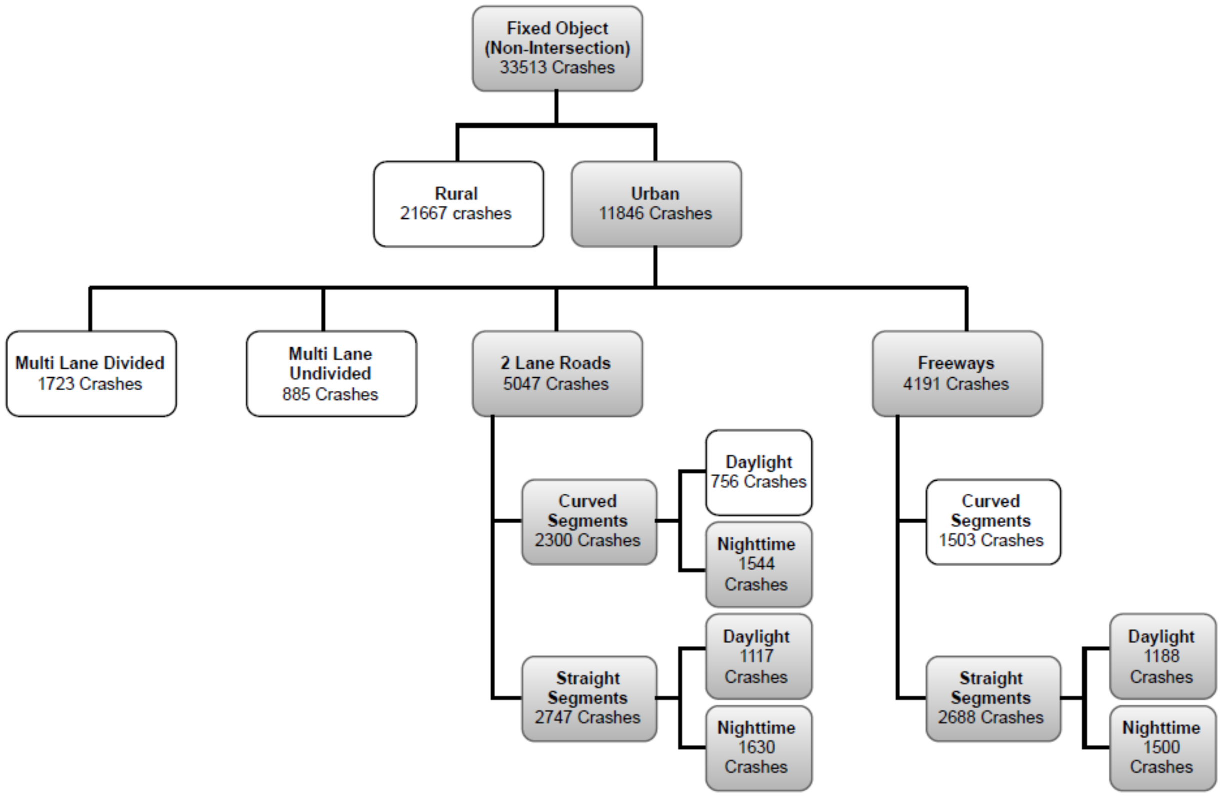
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 8. Chart. Crash tree for rear-end crashes at nonintersection locations.



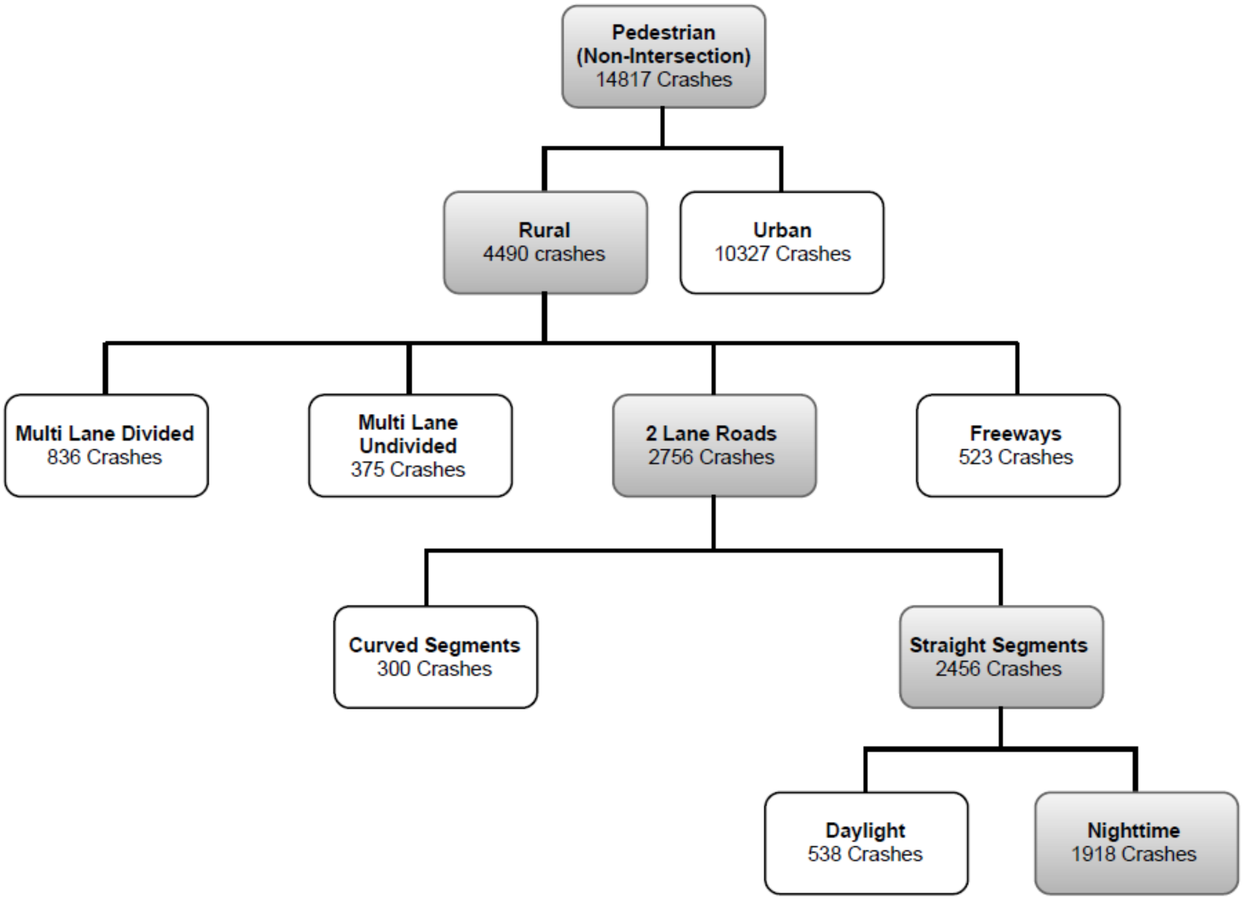
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 9. Chart. Crash tree for fixed-object crashes in rural areas at nonintersection locations.



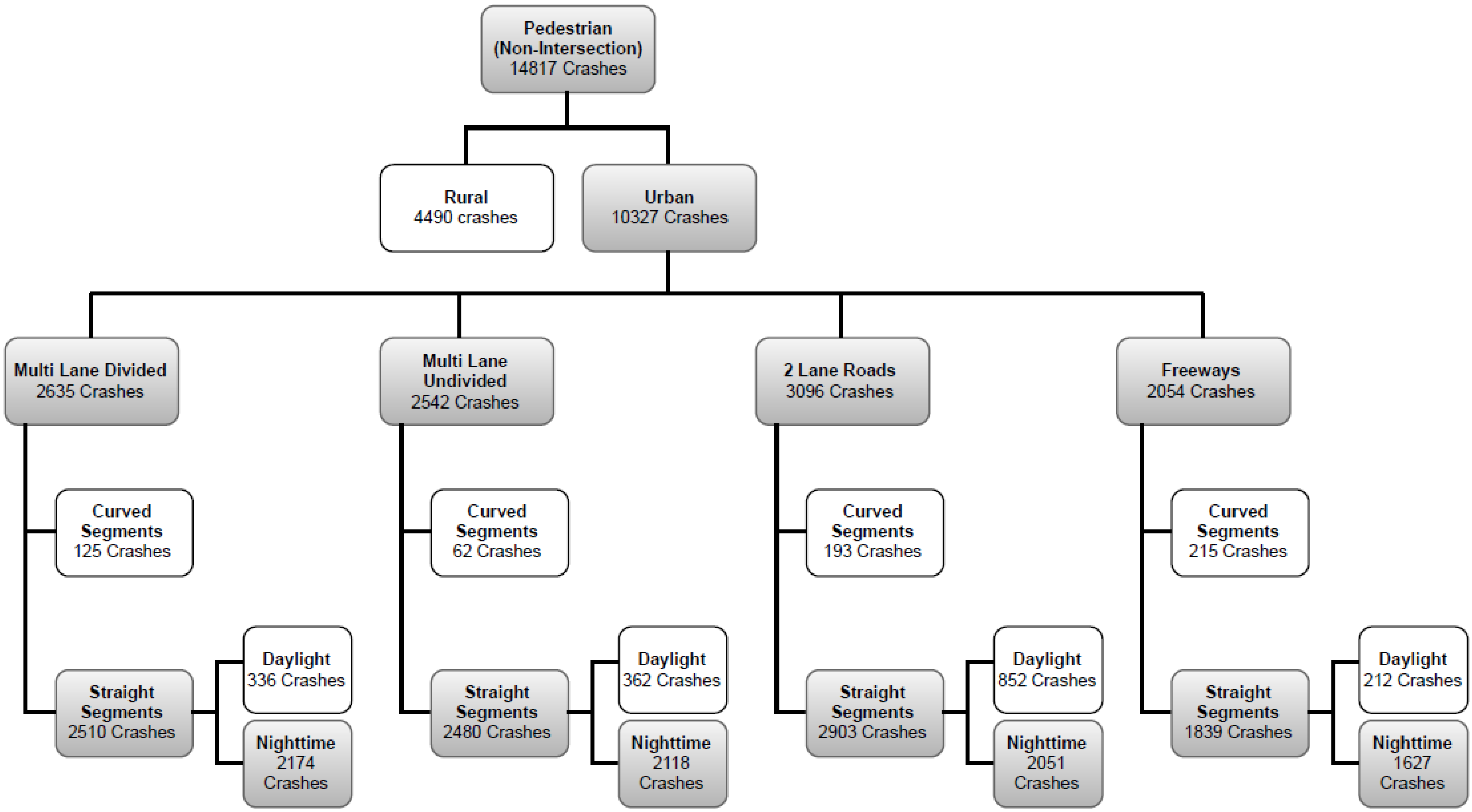
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 10. Chart. Crash tree for fixed-object crashes in urban areas at nonintersection locations.



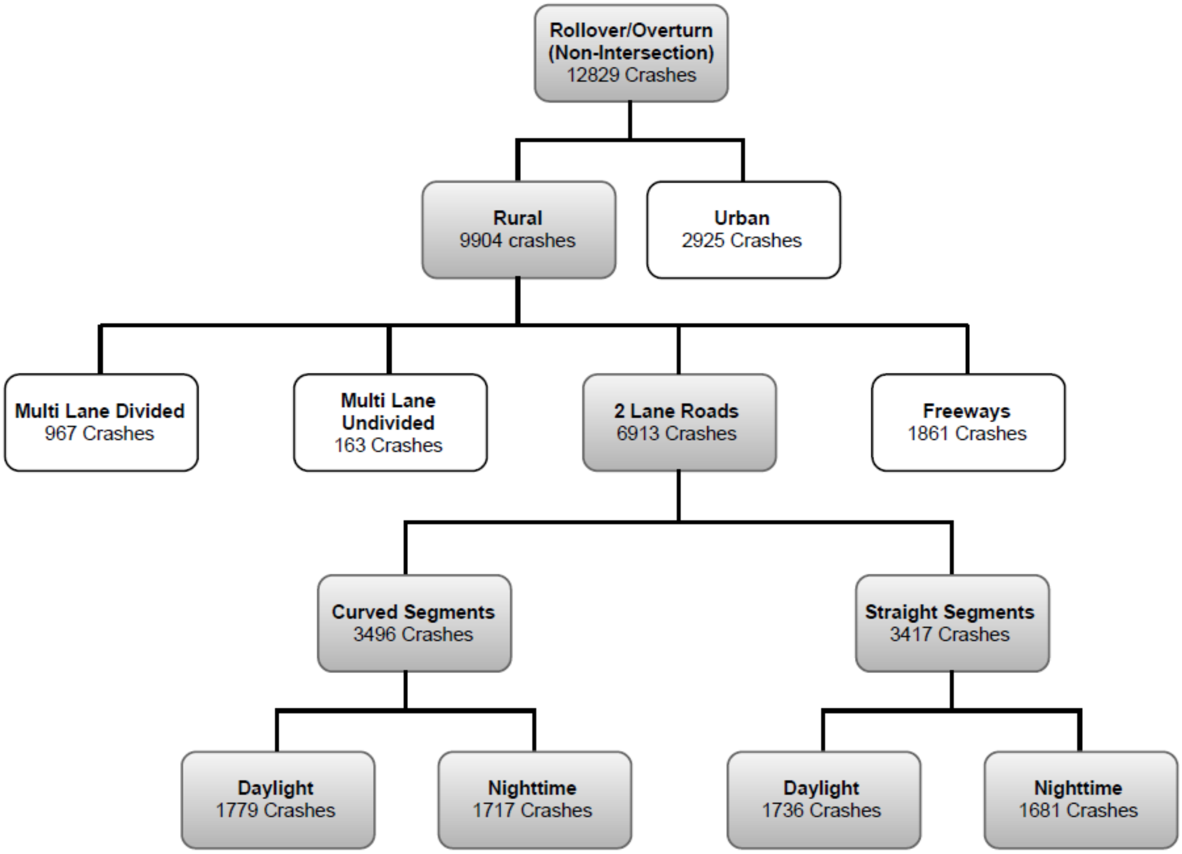
Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 11. Chart. Crash tree for pedestrian crashes in rural areas at nonintersection locations.



Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 12. Chart. Crash tree for pedestrian crashes in urban areas at nonintersection locations.



Source: FHWA.
Note: Shaded boxes are selected focus crash types.

Figure 13. Chart. Crash tree for ROLL crashes at nonintersection locations.

APPENDIX C. POTENTIAL INTERSECTION AND NONINTERSECTION FCFTS FROM FARS AND HSIS

Table 93 through table 110 list potential intersection and nonintersection FCFTs as derived from FARS and HSIS (NHTSA 2018a; FHWA 2018c). Potential FCFTs from HSIS datasets are listed in order based on both K and KA crashes.

Table 93. Potential nonintersection FCFTs from FARS.

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes
1	Fixed object	Rural	2-lane	Nighttime	Curve	5,197
2	Fixed object	Rural	2-lane	Daytime	Curve	4,095
3	Fixed object	Rural	2-lane	Nighttime	Straight	4,094
4	Fixed object	Rural	2-lane	Daytime	Straight	3,825
5	HEO	Rural	2-lane	Daytime	Straight	2,789
6	Pedestrian	Urban	Multilane divided	Nighttime	Straight	2,174
7	Pedestrian	Urban	Multilane undivided	Nighttime	Straight	2,118
8	Pedestrian	Urban	2-lane	Nighttime	Straight	2,051
9	Pedestrian	Rural	2-lane	Nighttime	Straight	1,918
10	ANG	Rural	2-lane	Daytime	Straight	1,815
11	Rollover/ overturn	Rural	2-lane	Daytime	Curve	1,779
12	HEO	Rural	2-lane	Nighttime	Straight	1,768
13	HEO	Rural	2-lane	Daytime	Curve	1,747
14	Rollover/ overturn	Rural	2-lane	Daytime	Straight	1,736
15	Rollover/ overturn	Rural	2-lane	Nighttime	Curve	1,717
16	Rollover/ overturn	Rural	2-lane	Nighttime	Straight	1,681
17	Fixed object	Urban	2-lane	Nighttime	Straight	1,630
18	Pedestrian	Urban	Interstates/ freeways/ expressways	Nighttime	Straight	1,627
19	Fixed object	Urban	2-lane	Nighttime	Curve	1,544
20	Fixed object	Urban	Freeways	Nighttime	Straight	1,500
21	Fixed object	Urban	Freeways	Daytime	Straight	1,188
22	Fixed object	Urban	2-lane	Daytime	Straight	1,177
23	Rear end	Urban	Interstates/ freeways/ expressways	Nighttime	Straight	1,053
24	Ditch	Rural	2-lane	Nighttime	Straight	1,026
25	Ditch	Rural	2-lane	Nighttime	Curve	993
26	Fixed object	Rural	Freeways	Daytime	Straight	987
27	Fixed object	Urban	Freeways	Nighttime	Curve	924
28	Rollover/ overturn	Rural	Interstates/ freeways/ expressways	Daytime	Straight	886
29	Ditch	Rural	2-lane	Daytime	Straight	868
30	Pedestrian	Urban	2-lane	Daytime	Straight	852
31	ANG	Rural	2-lane	Daytime	Curve	813

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes
32	Rear end	Urban	Interstates/ freeways/ expressways	Daytime	Straight	801
33	Fixed object	Rural	Freeways	Nighttime	Straight	785
34	ANG	Rural	2-lane	Nighttime	Straight	777
35	Fixed object	Urban	2-lane	Daytime	Curve	756
36	Ditch	Rural	2-lane	Daytime	Curve	755
37	HEO	Urban	2-lane	Daytime	Straight	736
38	HEO	Rural	2-lane	Nighttime	Curve	714
39	Rear end	Rural	2-lane	Daytime	Straight	714
40	Fixed object	Urban	Multilane divided	Nighttime	Straight	686
41	Pedestrian	Rural	Multilane divided	Nighttime	Straight	680
42	ANG	Urban	2-lane	Nighttime	Straight	625
43	Rollover/ overturn	Rural	Interstates/ freeways/ expressways	Nighttime	Straight	600
44	Fixed object	Urban	Freeways	Daytime	Curve	579
45	ANG	Urban	Multilane undivided	Daytime	Straight	549
46	Fixed object	Rural	Multilane divided	Daytime	Straight	539
47	Pedestrian	Rural	2-lane	Daytime	Straight	538
48	Rear end	Rural	Interstates/ freeways/ expressways	Nighttime	Straight	526
49	Fixed object	Rural	Multilane divided	Nighttime	Straight	519

Table 94. Potential intersection FCFTs from FARS.

No.	Crash Type	Area Type	Roadway Type	Location Type	Traffic Control	Light Conditions	K Crashes
1	ANG	Rural	2-lane	4-leg intersection	Stop controlled	Daytime	2,424
2	ANG	Urban	2-lane	4-leg intersection	Stop controlled	Daytime	1,126
3	ANG	Urban	Multilane divided	4-leg intersection	Traffic signal	Daytime	977
4	ANG	Urban	Multilane undivided	4-leg intersection	Traffic signal	Daytime	864
5	ANG	Rural	2-lane	4-leg intersection	Stop controlled	Nighttime	750
6	ANG	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	684
7	ANG	Rural	2-lane	T intersection	Stop controlled	Daytime	592
8	Pedestrian	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	589
9	ANG	Urban	Multilane undivided	4-leg intersection	Traffic signal	Nighttime	555
10	ANG	Urban	2-lane	T intersection	Stop controlled	Daytime	536
11	Pedestrian	Urban	Multilane undivided	4-leg intersection	Traffic signal	Nighttime	432

Table 95. Potential nonintersection FCFTs from HSIS: Minnesota (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	HEO	Rural	2-lane	Daytime	Straight	94	175
2	ROR	Rural	2-lane	Nighttime	Curve	75	104
3	ROR	Rural	2-lane	Daytime	Straight	71	226
4	ROR	Rural	2-lane	Daytime	Curve	68	222
5	ROR	Rural	2-lane	Nighttime	Straight	66	262
6	HEO	Rural	2-lane	Nighttime	Straight	51	231
7	ROLL	Rural	2-lane	Daytime	Straight	44	78
8	HEO	Rural	2-lane	Daytime	Curve	42	71
9	ROLL	Rural	2-lane	Nighttime	Curve	42	84
10	ROLL	Rural	2-lane	Nighttime	Straight	38	103
11	HEO	Urban	2-lane	Nighttime	Straight	33	140
12	ROLL	Rural	2-lane	Daytime	Curve	29	143
13	Pedestrian	Urban	2-lane	Nighttime	Straight	19	140
14	HEO	Urban	2-lane	Daytime	Straight	17	65

Table 96. Potential nonintersection FCFTs from HSIS: Minnesota (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	ROR	Rural	2-lane	Nighttime	Straight	66	262
2	HEO	Rural	2-lane	Nighttime	Straight	51	231
3	ROR	Rural	2-lane	Daytime	Straight	71	226
4	ROR	Rural	2-lane	Daytime	Curve	68	222
5	HEO	Rural	2-lane	Daytime	Straight	94	175
6	ROLL	Rural	2-lane	Daytime	Curve	29	143
7	HEO	Urban	2-lane	Nighttime	Straight	33	140
8	Pedestrian	Urban	2-lane	Nighttime	Straight	19	140
9	ROR	Rural	2-lane	Nighttime	Curve	75	104
10	ROLL	Rural	2-lane	Nighttime	Straight	38	103
11	ROLL	Rural	2-lane	Nighttime	Curve	42	84
12	ROLL	Rural	2-lane	Daytime	Straight	44	78
13	HEO	Rural	2-lane	Daytime	Curve	42	71
14	HEO	Urban	2-lane	Daytime	Straight	17	65

Table 97. Potential intersection FCFTs from HSIS: Minnesota (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Traffic Control	Light Conditions	K Crashes	KA Crashes
1	ANG	Rural	2-lane	4-leg intersection	Minor road stop control	Daytime	97	203
2	ANG	Rural	Multilane divided	4-leg intersection	Minor road stop control	Daytime	24	42
3	ANG	Urban	Multilane divided	4-leg intersection	Traffic signal	Daytime	22	80
4	ANG	Urban	2-lane	4-leg intersection	Minor road stop control	Daytime	15	116
5	Pedestrian	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	13	42

Table 98. Potential intersection FCFTs from HSIS: Minnesota (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Traffic Control	Light Conditions	K Crashes	KA Crashes
1	ANG	Rural	2-lane	4-leg intersection	Minor road stop control	Daytime	97	203
2	ANG	Urban	Multilane divided	4-leg intersection	Traffic signal	Daytime	22	116
3	ANG	Urban	2-lane	4-leg intersection	Minor road stop control	Daytime	15	80
4	ANG	Rural	Multilane divided	4-leg intersection	Minor road stop control	Daytime	24	42
5	Pedestrian	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	13	42

Table 99. Potential nonintersection FCFTs from HSIS: Ohio (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	Fixed object	Rural	2-lane	Daytime	Curve	123	773
2	Fixed object	Rural	2-lane	Daytime	Straight	111	1027
3	Fixed object	Rural	2-lane	Nighttime	Curve	110	569
4	Fixed object	Rural	2-lane	Nighttime	Straight	101	792
5	Sideswipe—meeting	Rural	2-lane	Daytime	Straight	85	391
6	HEO	Rural	2-lane	Daytime	Straight	78	181
7	Fixed object	Urban	Freeways	Nighttime	Straight	52	522
8	Fixed object	Urban	Freeways	Nighttime	Curve	49	291
9	Fixed object	Urban	Freeways	Daytime	Straight	48	590
10	Sideswipe—meeting	Rural	2-lane	Daytime	Curve	48	196
11	Pedestrian	Urban	Multilane undivided	Nighttime	Straight	45	169
12	Sideswipe—meeting	Rural	2-lane	Nighttime	Straight	42	199
13	Rear end	Urban	Freeways	Nighttime	Straight	35	238
14	Fixed object	Urban	Freeways	Daytime	Curve	30	297
15	Rear end	Urban	Freeways	Daytime	Straight	28	565
16	Fixed object	Urban	2-lane	Nighttime	Straight	25	191
17	ROLL	Rural	2-lane	Daytime	Curve	22	225
18	Fixed object	Rural	Freeways	Daytime	Straight	21	155
19	Fixed object	Urban	2-lane	Daytime	Straight	21	202
20	Sideswipe—passing	Urban	Freeways	Daytime	Straight	21	348
21	Fixed object	Urban	Multilane undivided	Nighttime	Straight	18	158
22	ROLL	Rural	2-lane	Daytime	Curve	18	166
23	Rear end	Rural	2-lane	Daytime	Straight	16	409
24	Sideswipe—passing	Urban	Freeways	Daytime	Straight	14	169

Table 100. Potential nonintersection FCFTs from HSIS: Ohio (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	Fixed object	Rural	2-lane	Daytime	Straight	111	1,027
2	Fixed object	Rural	2-lane	Nighttime	Straight	101	792
3	Fixed object	Rural	2-lane	Daytime	Curve	123	773
4	Fixed object	Urban	Freeways	Daytime	Straight	48	590
5	Fixed object	Rural	2-lane	Nighttime	Curve	110	569
6	Rear end	Urban	Freeways	Daytime	Straight	28	565
7	Fixed object	Urban	Freeways	Nighttime	Straight	52	522
8	Rear end	Rural	2-lane	Daytime	Straight	16	409
9	Sideswipe—meeting	Rural	2-lane	Daytime	Straight	85	391
10	Sideswipe—passing	Urban	Freeways	Daytime	Straight	21	348
11	Rear end	Urban	Multilane undivided	Daytime	Straight	6	318
12	Fixed object	Urban	Freeways	Daytime	Curve	30	297
13	Fixed object	Urban	Freeways	Nighttime	Curve	49	291
14	Rear end	Urban	2-lane	Daytime	Straight	6	270
15	Rear end	Urban	Freeways	Nighttime	Straight	35	238
16	ROLL	Rural	2-lane	Daytime	Curve	22	225
17	ANG	Urban	Multilane undivided	Daytime	Straight	4	213
18	Fixed object	Urban	2-lane	Daytime	Straight	21	202
19	Sideswipe—meeting	Rural	2-lane	Nighttime	Straight	42	199
20	Sideswipe—meeting	Rural	2-lane	Daytime	Curve	48	196
21	Fixed object	Urban	2-lane	Nighttime	Straight	25	191
22	HEO	Rural	2-lane	Daytime	Straight	78	181
23	Pedestrian	Urban	Multilane undivided	Nighttime	Straight	45	169
24	Sideswipe—passing	Urban	Freeways	Daytime	Straight	14	169

Table 101. Potential intersection FCFTs from HSIS: Ohio (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Light Conditions	K Crashes	KA Crashes
1	ANG	Rural	2-lane	At or related to intersection	Daytime	153	1,152
2	ANG	Rural	2-lane	At or related to intersection	Nighttime	48	296
3	ANG	Urban	Multilane divided	At or related to intersection	Daytime	33	511
4	ANG	Urban	Multilane undivided	At or related to intersection	Daytime	29	657
5	ANG	Urban	Multilane divided	At or related to intersection	Daytime	27	243
6	ANG	Rural	2-lane	At or related to intersection	Daytime	25	213
7	Pedestrian	Urban	Multilane undivided	At or related to intersection	Nighttime	20	163
8	ANG	Urban	2-lane	At or related to intersection	Nighttime	17	169
9	Rear end	Rural	2-lane	At or related to intersection	Daytime	13	205
10	Pedestrian	Urban	Multilane undivided	At or related to intersection	Daytime	11	184
11	ANG	Urban	Multilane undivided	At or related to intersection	Nighttime	10	242
12	Rear end	Urban	Multilane undivided	At or related to intersection	Daytime	10	338
13	Rear end	Urban	2-lane	At or related to intersection	Daytime	5	243

Table 102. Potential intersection FCFTs from HSIS: Ohio (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Light Conditions	K Crashes	KA Crashes
1	ANG	Rural	2-lane	At or related to intersection	Daytime	153	1,152
2	ANG	Urban	Multilane undivided	At or related to intersection	Daytime	29	657
3	ANG	Urban	Multilane divided	At or related to intersection	Daytime	33	511
4	Rear end	Urban	Multilane undivided	At or related to intersection	Daytime	10	338
5	ANG	Rural	2-lane	At or related to intersection	Nighttime	48	296
6	ANG	Urban	Multilane divided	At or related to intersection	Daytime	27	243
7	Rear end	Urban	2-lane	At or related to intersection	Daytime	5	243
8	ANG	Urban	Multilane undivided	At or related to intersection	Nighttime	10	242
9	ANG	Rural	2-lane	At or related to intersection	Daytime	25	213
10	Rear end	Rural	2-lane	At or related to intersection	Daytime	13	205
11	Pedestrian	Urban	Multilane undivided	At or related to intersection	Daytime	11	184
12	ANG	Urban	2-lane	At or related to intersection	Nighttime	17	169
13	Pedestrian	Urban	Multilane undivided	At or related to intersection	Nighttime	20	163

Table 103. Potential nonintersection FCFTs from HSIS: California (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	K Crashes	KA Crashes
1	Hit object	Urban	Freeways	Nighttime	530	1,983
2	Pedestrian	Urban	Freeways	Nighttime	350	542
3	Hit object	Urban	Freeways	Daytime	279	1,182
4	Rear end	Urban	Freeways	Nighttime	276	1,172
5	HEO	Rural	2-lane	Daytime	223	530
6	Hit object	Rural	2-lane	Daytime	214	867
7	Hit object	Rural	2-lane	Nighttime	173	547
8	Rear end	Urban	Freeways	Daytime	162	1,182
9	Hit object	Rural	Freeways	Nighttime	137	326
10	ROLL	Rural	Freeways	Daytime	135	418
11	HEO	Rural	2-lane	Nighttime	129	281
12	Hit object	Rural	Freeways	Daytime	121	394
13	Sideswipe	Urban	Freeways	Daytime	120	738
14	ROLL	Rural	2-lane	Daytime	107	716
15	Pedestrian	Urban	Multilane divided	Nighttime	106	252
16	ROLL	Urban	Freeways	Nighttime	106	504
17	ROLL	Urban	Freeways	Daytime	96	623
18	Sideswipe	Urban	Freeways	Nighttime	88	436
19	ROLL	Rural	Freeways	Nighttime	84	235
20	Broadside	Urban	Freeways	Nighttime	72	260
21	Rear end	Rural	Freeways	Nighttime	71	215
22	ROLL	Rural	2-lane	Nighttime	54	221
23	Sideswipe	Rural	2-lane	Daytime	40	201

Table 104. Potential nonintersection FCFTs from HSIS: California (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	K Crashes	KA Crashes
1	Hit object	Urban	Freeways	Nighttime	530	1,983
2	Hit object	Urban	Freeways	Daytime	279	1,182
3	Rear end	Urban	Freeways	Daytime	162	1,182
4	Rear end	Urban	Freeways	Nighttime	276	1,172
5	Hit object	Rural	2-lane	Daytime	214	867
6	Sideswipe	Urban	Freeways	Daytime	120	738
7	ROLL	Rural	2-lane	Daytime	107	716
8	ROLL	Urban	Freeways	Daytime	96	623
9	Hit object	Rural	2-lane	Nighttime	173	547
10	Pedestrian	Urban	Freeways	Nighttime	350	542
11	HEO	Rural	2-lane	Daytime	223	530
12	ROLL	Urban	Freeways	Nighttime	106	504
13	Sideswipe	Urban	Freeways	Nighttime	88	436
14	ROLL	Rural	Freeways	Daytime	135	418
15	Hit object	Rural	Freeways	Daytime	121	394
16	Hit object	Rural	Freeways	Nighttime	137	326
17	HEO	Rural	2-lane	Nighttime	129	281
18	Broadside	Urban	Freeways	Nighttime	72	260
19	Pedestrian	Urban	Multilane divided	Nighttime	106	252
20	ROLL	Rural	Freeways	Nighttime	84	235
21	ROLL	Rural	2-lane	Nighttime	54	221
22	Rear end	Rural	Freeways	Nighttime	71	215
23	Sideswipe	Rural	2-lane	Daytime	40	201

Table 105. Potential intersection FCFTs from HSIS: California (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Traffic Control	Light Conditions	K Crashes	KA Crashes
1	Broadside	Rural	2-lane	4-leg intersection	Minor road stop control	Daytime	30	23
2	Broadside	Urban	Multilane divided	4-leg intersection	Traffic signal	Daytime	13	71
3	Broadside	Rural	2-lane	4-leg intersection	Minor road stop control	Nighttime	12	33
4	Broadside	Rural	Multilane divided	4-leg intersection	Minor road stop control	Daytime	11	35
5	Broadside	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	10	48
6	Pedestrian	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	9	37
7	Hit object	Rural	2-lane	T-intersection	Minor road stop control	Nighttime	4	15

Table 106. Potential intersection FCFTs from HSIS: California (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Traffic Control	Light Conditions	K Crashes	KA Crashes
1	Broadside	Rural	2-lane	4-leg intersection	Minor road stop control	Daytime	30	23
2	Broadside	Urban	Multilane divided	4-leg intersection	Traffic signal	Daytime	13	71
3	Broadside	Rural	2-lane	4-leg intersection	Minor road stop control	Nighttime	12	33
4	Broadside	Rural	Multilane divided	4-leg intersection	Minor road stop control	Daytime	11	35
5	Broadside	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	10	48
6	Pedestrian	Urban	Multilane divided	4-leg intersection	Traffic signal	Nighttime	9	37
7	Hit object	Rural	2-lane	T-intersection	Minor road stop control	Nighttime	4	15

Table 107. Potential nonintersection FCFTs from HSIS: Washington (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	Hit object	Rural	2-lane	Daytime	Straight	35	118
2	Hit object	Rural	2-lane	Nighttime	Straight	33	111
3	HEO	Rural	2-lane	Daytime	Straight	33	59
4	Hit object	Rural	2-lane	Nighttime	Curve	24	64
5	Hit object	Rural	2-lane	Daytime	Curve	23	92
6	Hit object	Urban	Freeways	Nighttime	Straight	22	66
7	Pedestrian	Urban	Multilane undivided	Nighttime	Straight	18	57
8	ANG	Rural	2-lane	Daytime	Straight	16	59
9	Rear end	Urban	Freeways	Nighttime	Straight	15	63
10	Hit object	Urban	Freeways	Daytime	Straight	11	60
11	Rear end	Urban	Freeways	Daytime	Straight	8	110
12	Rear end	Rural	2-lane	Daytime	Straight	8	57

Table 108. Potential nonintersection FCFTs from HSIS: Washington (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Light Conditions	Road Alignment	K Crashes	KA Crashes
1	Hit object	Rural	2-lane	Daytime	Straight	35	118
2	Hit object	Rural	2-lane	Nighttime	Straight	33	111
3	Rear end	Urban	Freeways	Daytime	Straight	8	110
4	Hit object	Rural	2-lane	Daytime	Curve	23	92
5	Hit object	Urban	Freeways	Nighttime	Straight	22	66
6	Hit object	Rural	2-lane	Nighttime	Curve	24	64
7	Rear end	Urban	Freeways	Nighttime	Straight	15	63
8	Hit object	Urban	Freeways	Daytime	Straight	11	60
9	ANG	Rural	2-lane	Daytime	Straight	16	59
10	HEO	Rural	2-lane	Daytime	Straight	33	59
11	Pedestrian	Urban	Multilane undivided	Nighttime	Straight	18	57
12	Rear end	Rural	2-lane	Daytime	Straight	8	57

Table 109. Potential intersection FCFTs from HSIS: Washington (by K crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Light Conditions	K Crashes	KA Crashes
1	ANG	Rural	2-lane	At or related to intersection	Daytime	28	64
2	ANG	Urban	Multilane undivided	At or related to intersection	Daytime	13	110
3	Pedestrian	Urban	Multilane undivided	At or related to intersection	Nighttime	11	129
4	ANG	Rural	2-lane	At or related to intersection	Nighttime	10	95
5	ANG	Urban	2-lane	At or related to intersection	Daytime	7	42
6	Rear end	Urban	Multilane undivided	At or related to intersection	Daytime	5	49
7	ANG	Urban	Multilane undivided	At or related to intersection	Nighttime	5	48
8	ANG	Urban	2-lane	At or related to intersection	Daytime	5	40
9	Rear end	Rural	2-lane	At or related to intersection	Daytime	3	49
10	Pedestrian	Urban	Multilane undivided	At or related to intersection	Daytime	3	40

Table 110. Potential intersection FCFTs from HSIS: Washington (by KA crashes).

No.	Crash Type	Area Type	Roadway Type	Location Type	Light Conditions	K Crashes	KA Crashes
1	Pedestrian	Urban	Multilane undivided	At or related to intersection	Nighttime	11	129
2	ANG	Urban	Multilane undivided	At or related to intersection	Daytime	13	110
3	ANG	Rural	2-lane	At or related to intersection	Nighttime	10	95
4	ANG	Rural	2-lane	At or related to intersection	Daytime	28	64
5	Rear end	Urban	Multilane undivided	At or related to intersection	Daytime	5	49
6	Rear end	Rural	2-lane	At or related to intersection	Daytime	3	49
7	ANG	Urban	Multilane undivided	At or related to intersection	Nighttime	5	48
8	ANG	Urban	2-lane	At or related to intersection	Daytime	7	42
9	ANG	Urban	2-lane	At or related to intersection	Daytime	5	40
10	Pedestrian	Urban	Multilane undivided	At or related to intersection	Daytime	3	40

APPENDIX D. RANDOM-FOREST R CODE

Figure 14 provides a sample random-forest code that was used in R software for the contributing-factor analysis. In this particular example, the code was used to analyze ANG crashes at horizontal curves on rural two-lane highway segments. The code can be modified to analyze other focus crash types.

```
######## Set and Call Working Directory #########
setwd ("C:/Rdatasets/WA_CurveDay")
##################################################

getwd()

library(randomForest)

# import dataset
y <- read.delim("C:/Rdatasets/WA_CurveDay/WaCurvDayforR_small.txt")
View(y)
summary(y)

# convert some variables into factors
y$grade <- as.factor (y$grade)
y$GRAD_TYP <- as.factor (y$GRAD_TYP)
y$access <- as.factor (y$access)
y$shl_typ <- as.factor (y$shl_typ)
y$lanewidth <- as.factor (y$lanewidth)
y$terrain <- as.factor (y$terrain)


# Change ANGrate with crash type being analyzed
set.seed(123)
rf <- randomForest      (ANGrate ~ CURV_RAD + avg_aadt + pct_grad + GRAD_TYP + shl_typ + terrain +
                        shldwid + lanewidth + trkpcts + snowavgyear + rainavgyear + tempmaxavg +
                        tempminavg + temp32fdays + tempwintermin + unempl16plus + diploma25plus
                        + univ25plus + noedctn25plus + income50k + income50to100k + income100kplus
                        + noveh + X1veh + X2vehplus + age15to19 + age20to44 + age45to64 +
                        age65to74 + age75plus + workftage16to24 + workptage16to24 +
                        noworkage16to24,      data=y,importance=T,ntree=2000,na.action=na.exclude)


# How many trees are needed to reach the minimum error estimate?
which.min(rf$mse)
```

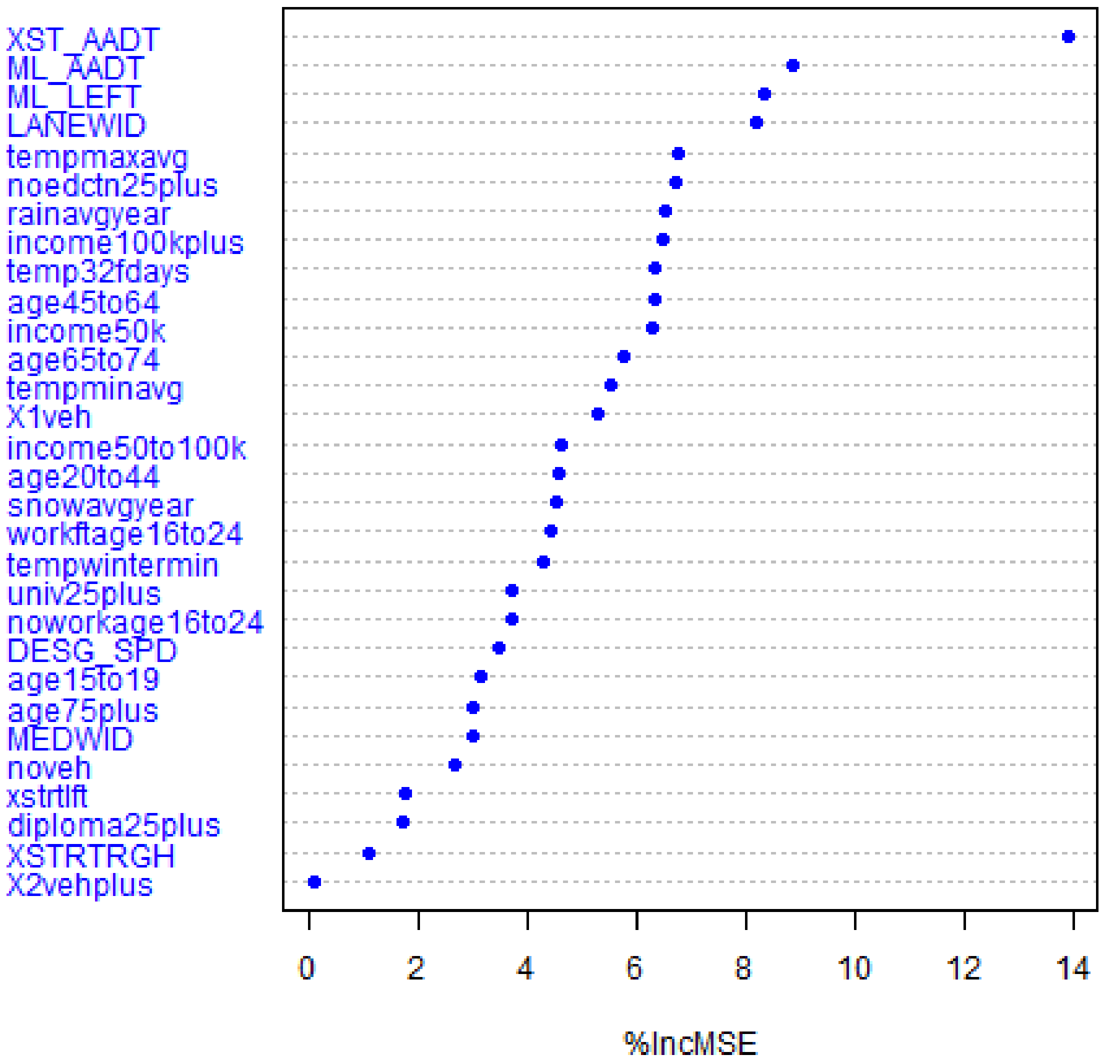
Source: FHWA.

Figure 14. Screenshot. Sample random-forest code for R software.

APPENDIX E. RANDOM-FOREST OUTPUTS

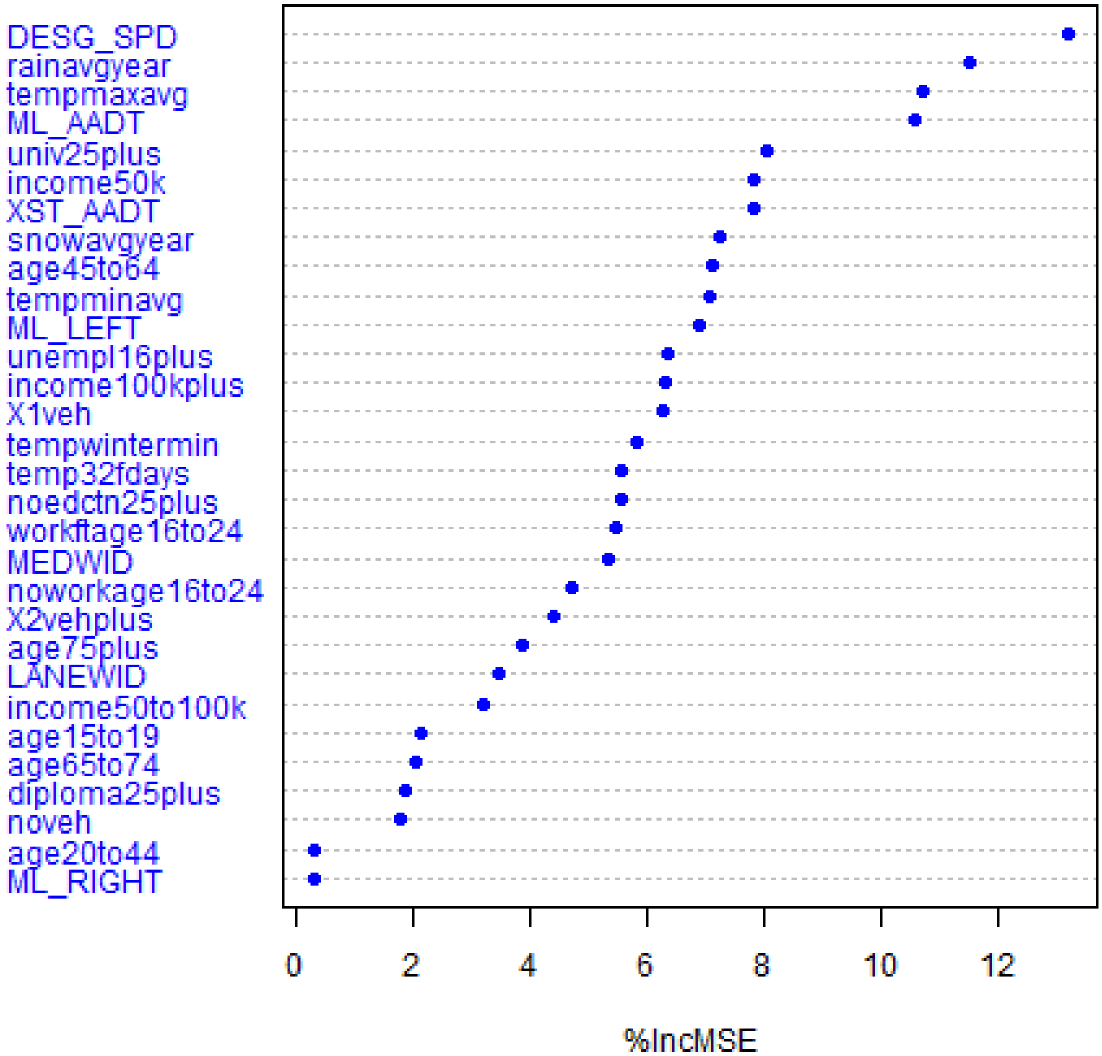
Figure 15 through figure 70 show random-forest outputs of the contributing-factor analysis for California, Ohio, and Washington data.

RANDOM FORESTS OF CALIFORNIA DATA



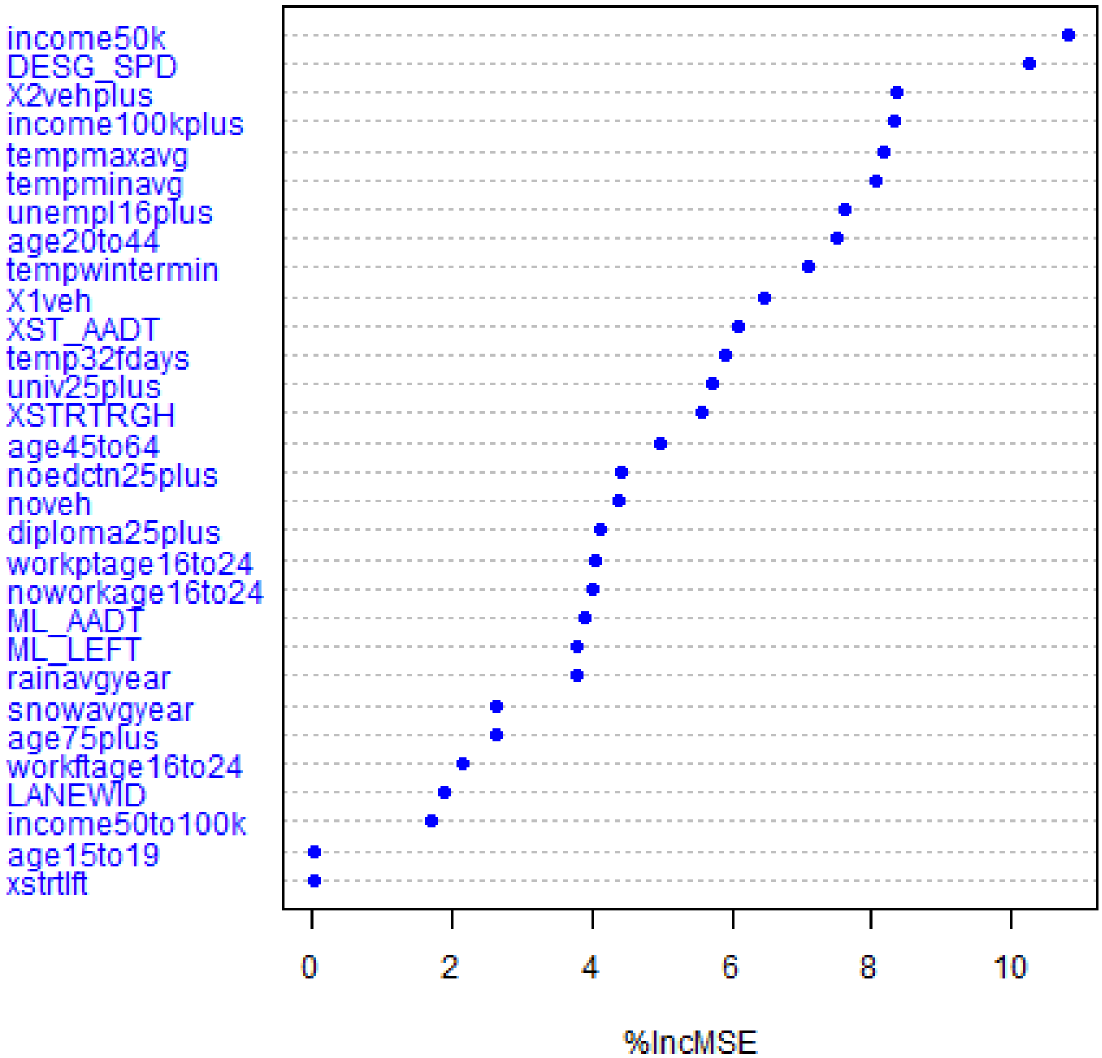
Source: FHWA.

Figure 15. Graph. ANG-D crashes at four-leg stop-controlled intersections on rural two-lane roads.



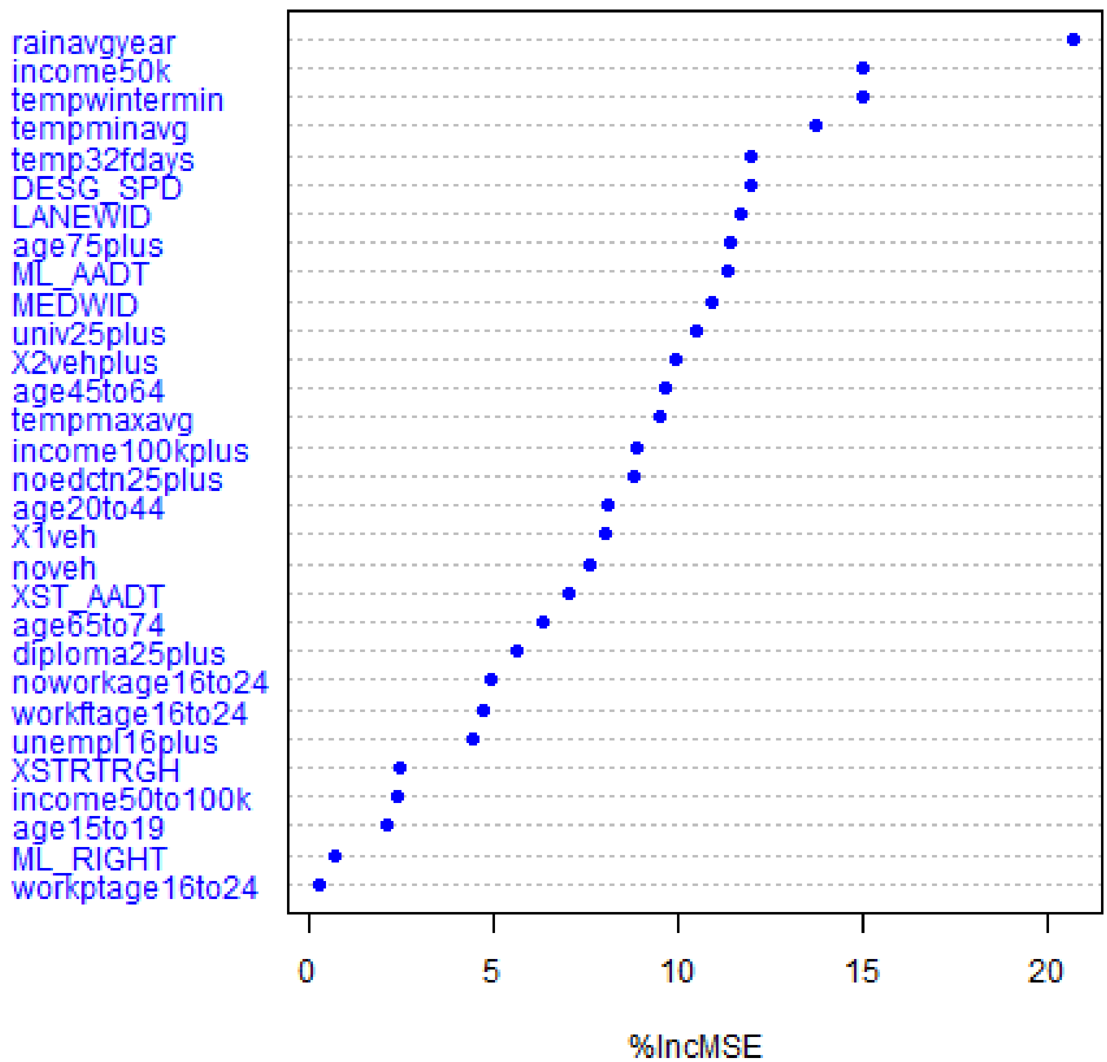
Source: FHWA.

Figure 16. Graph. ANG-N crashes at four-leg stop-controlled intersections on rural two-lane roads.



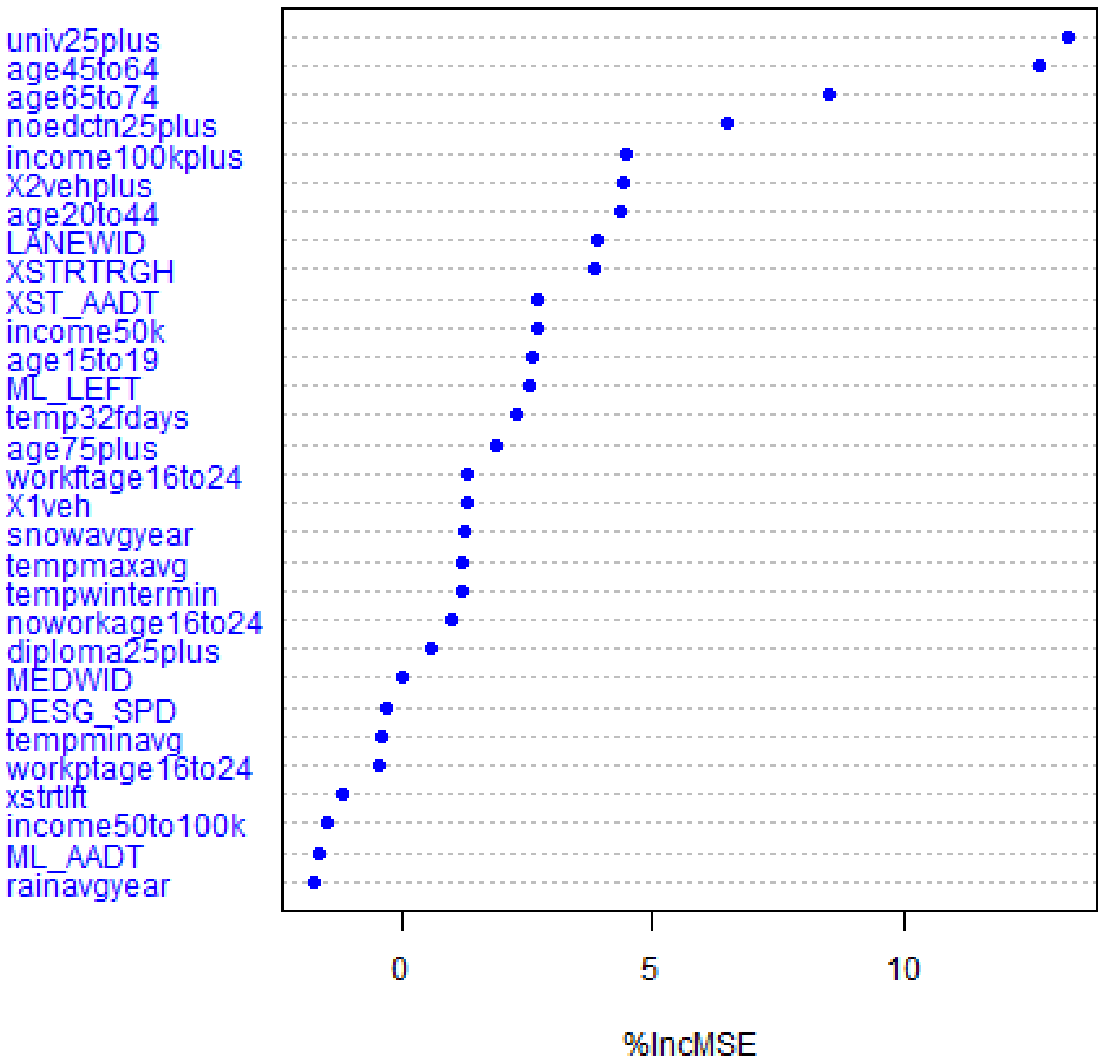
Source: FHWA.

Figure 17. Graph. ANG-D crashes at four-leg stop-controlled intersection on urban two-lane roads.



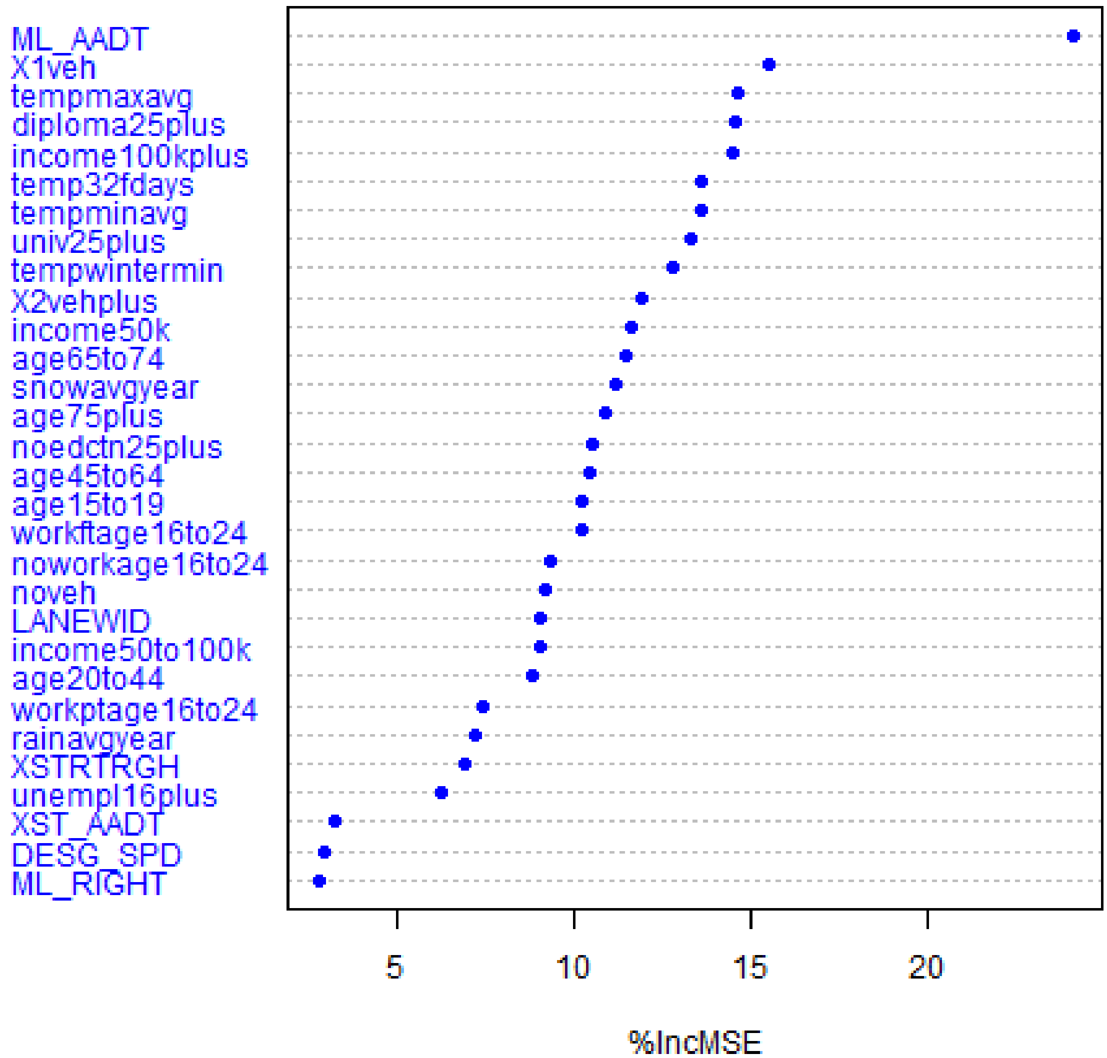
Source: FHWA.

Figure 18. Graph. ANG-D crashes at four-leg signalized intersections on urban multilane divided roads.



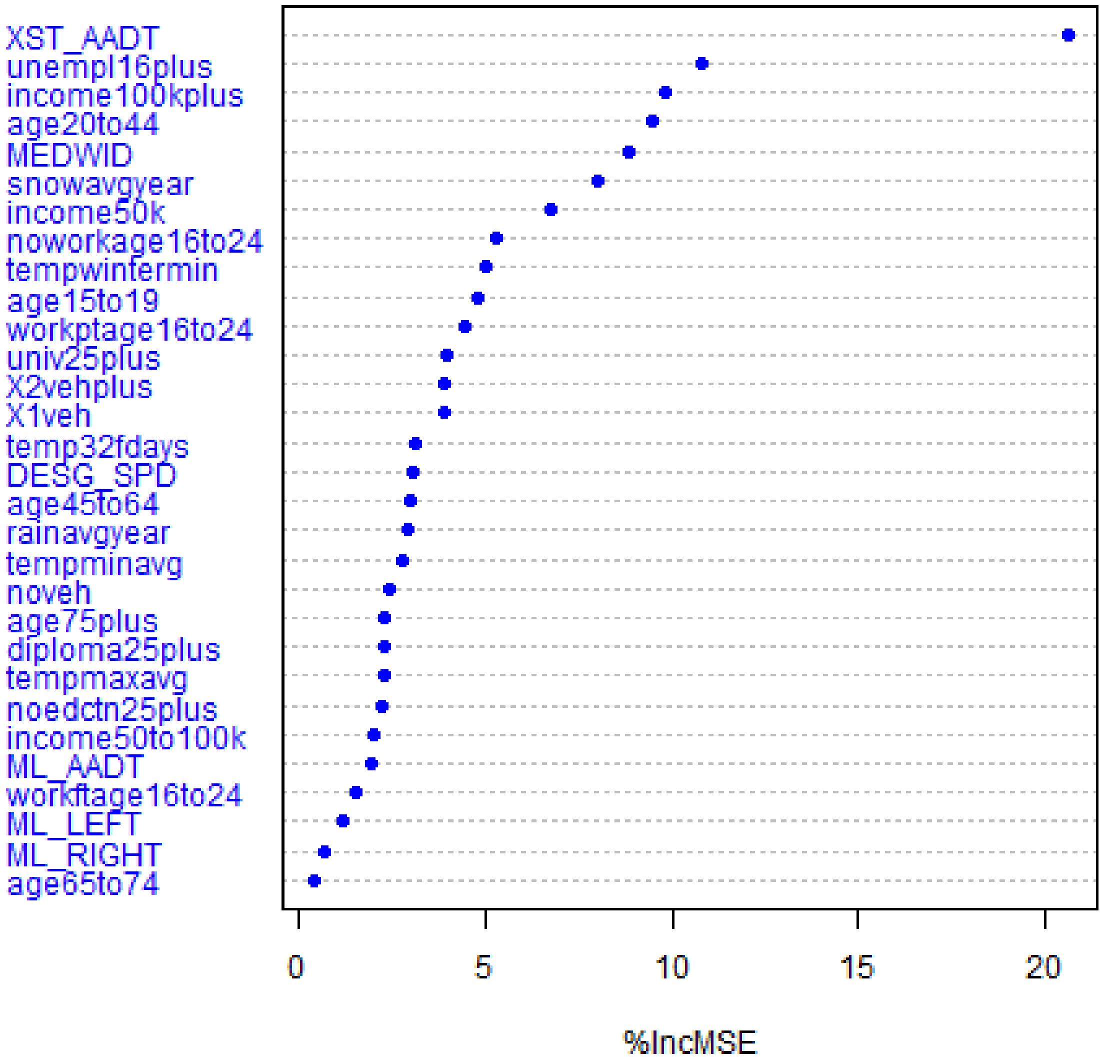
Source: FHWA.

Figure 19. Graph. ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads.



Source: FHWA.

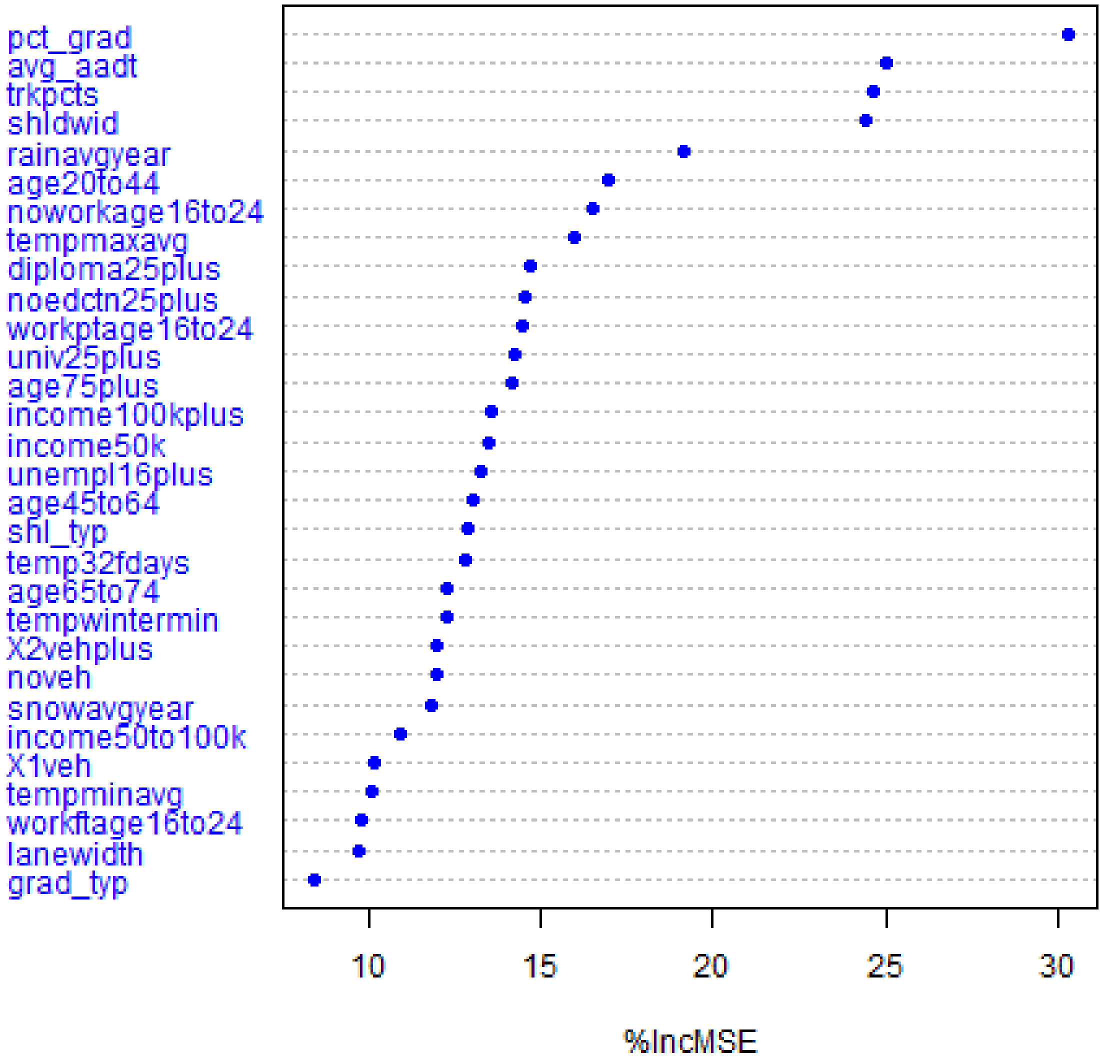
Figure 20. Graph. ANG-D crashes at three-leg stop-controlled intersections on rural two-lane roads.



Source: FHWA.

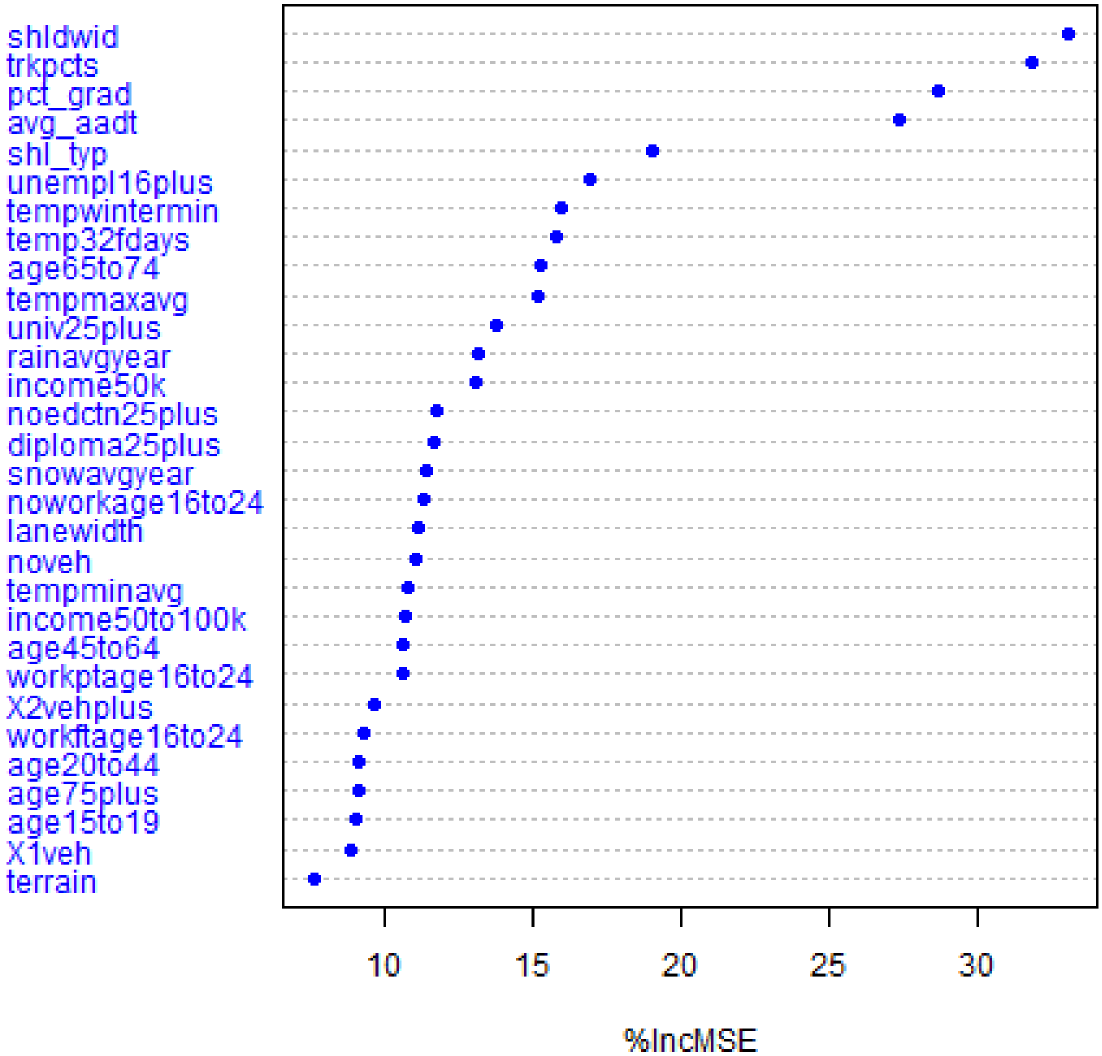
Figure 21. Graph. ANG-D crashes at four-leg stop-controlled intersections on rural multilane divided roads.

RANDOM FORESTS OF WASHINGTON DATA



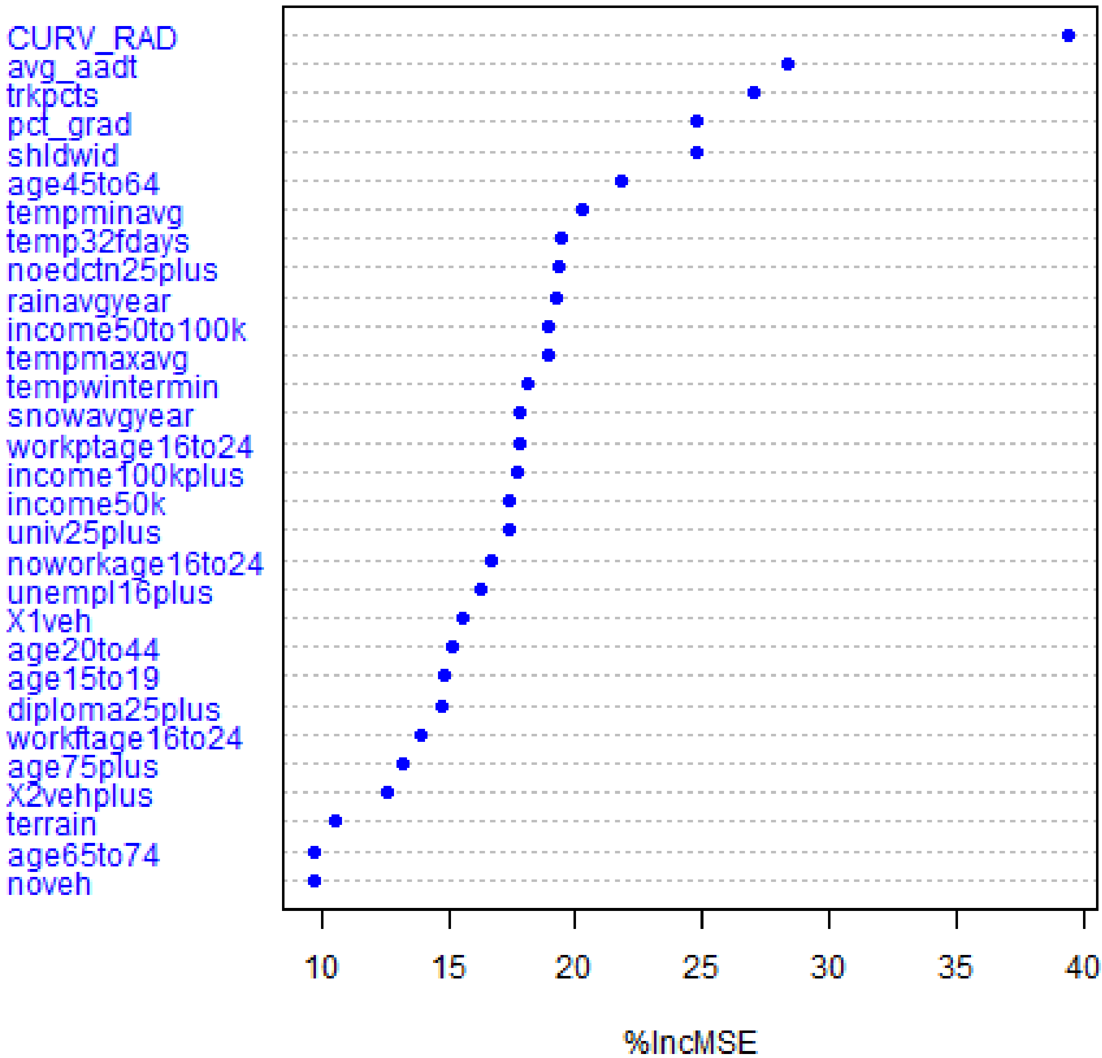
Source: FHWA.

Figure 22. Graph. ROR-D crashes at tangent segments on rural two-lane highway segments.



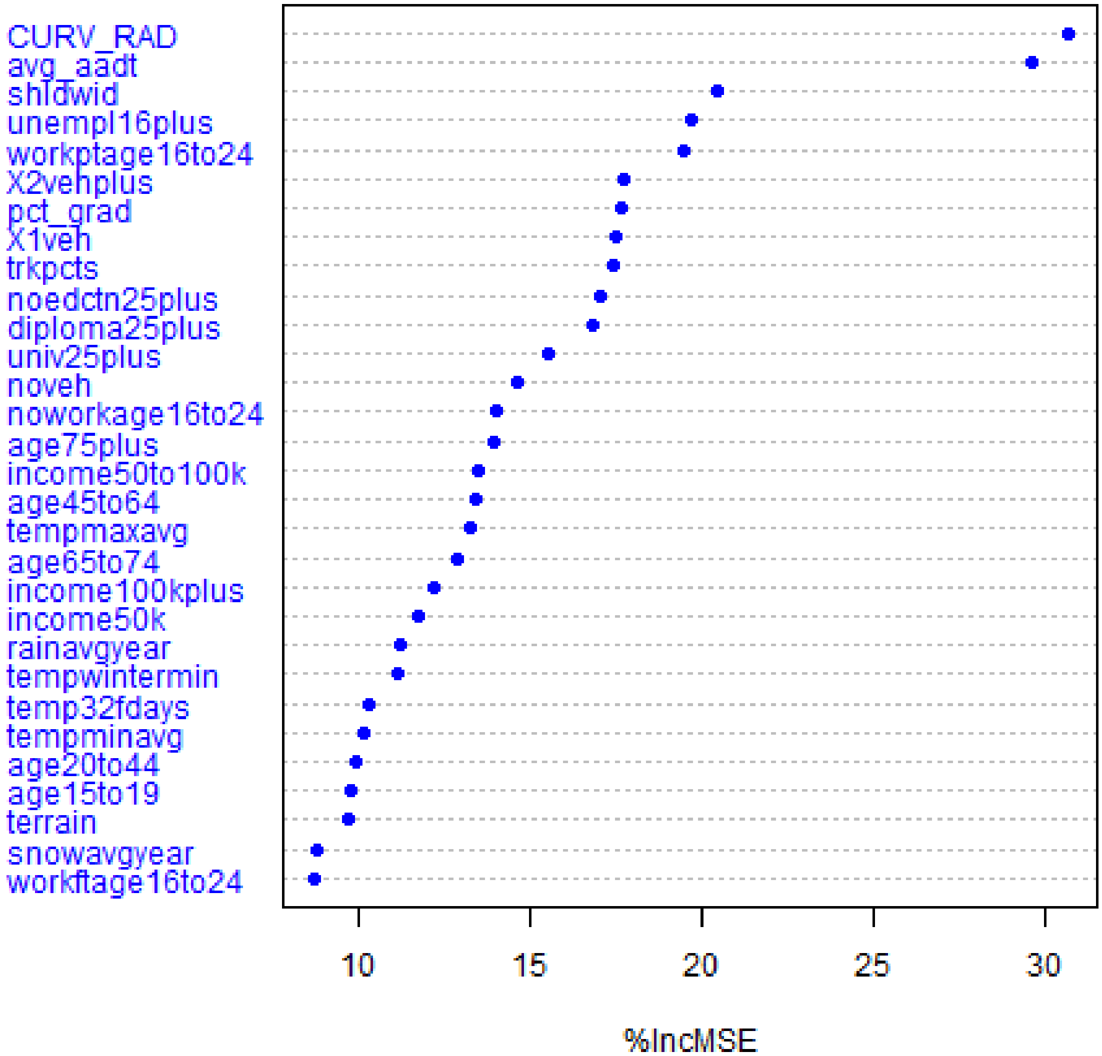
Source: FHWA.

Figure 23. Graph. ROR-N crashes at tangent segments on rural two-lane highway segments.



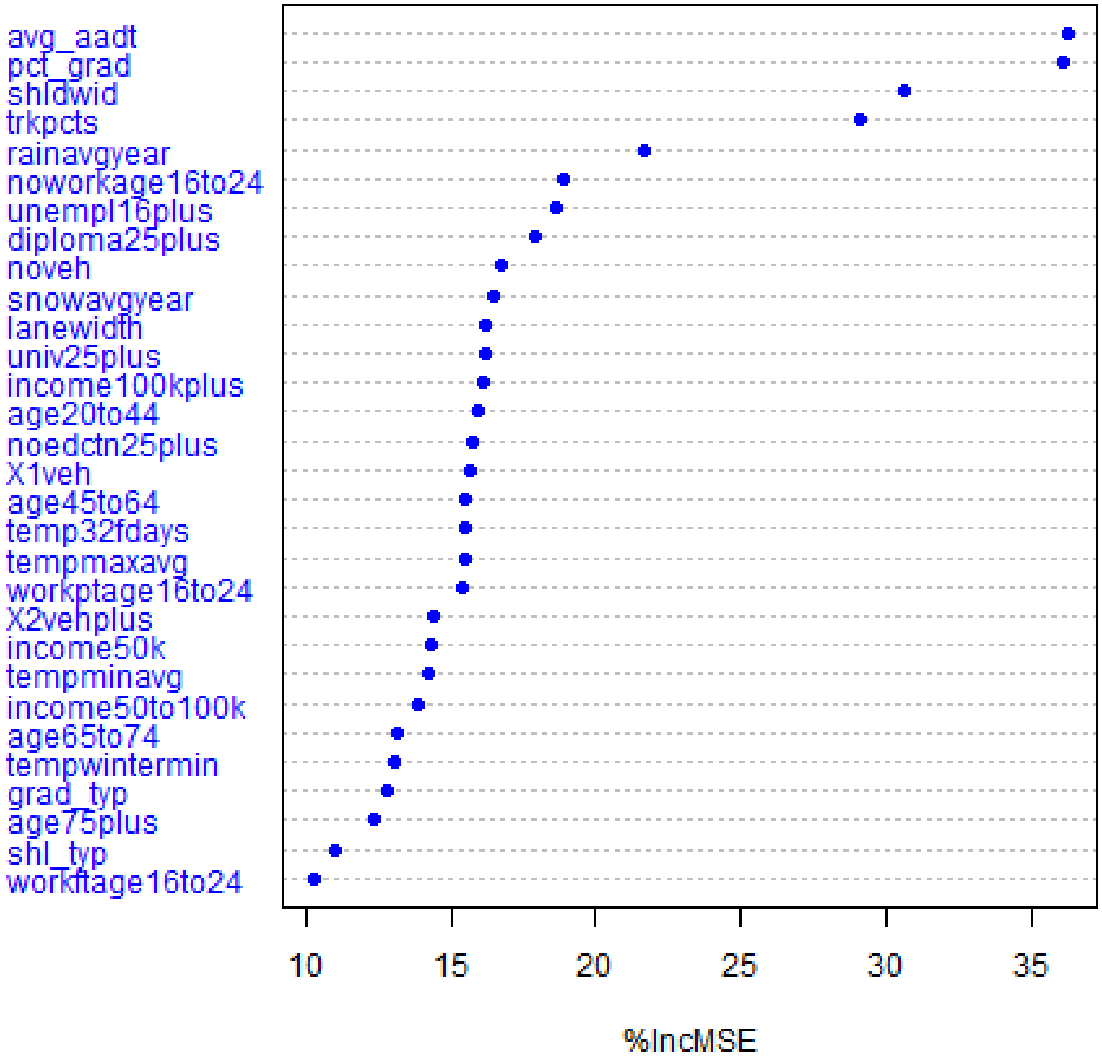
Source: FHWA.

Figure 24. Graph. ROR-D crashes at horizontal curves on rural two-lane highway segments.



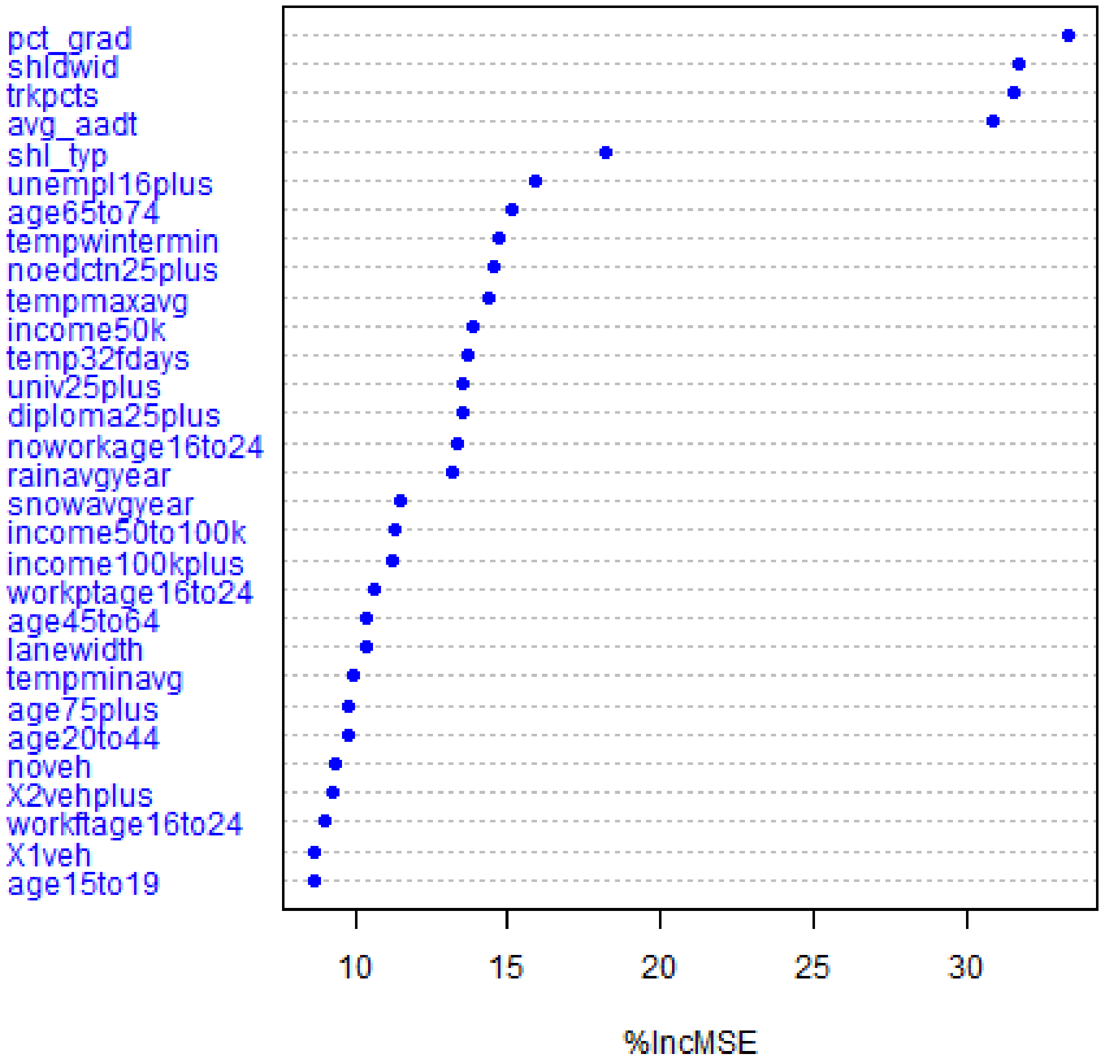
Source: FHWA.

Figure 25. Graph. ROR-N crashes at horizontal curves on rural two-lane highway segments.



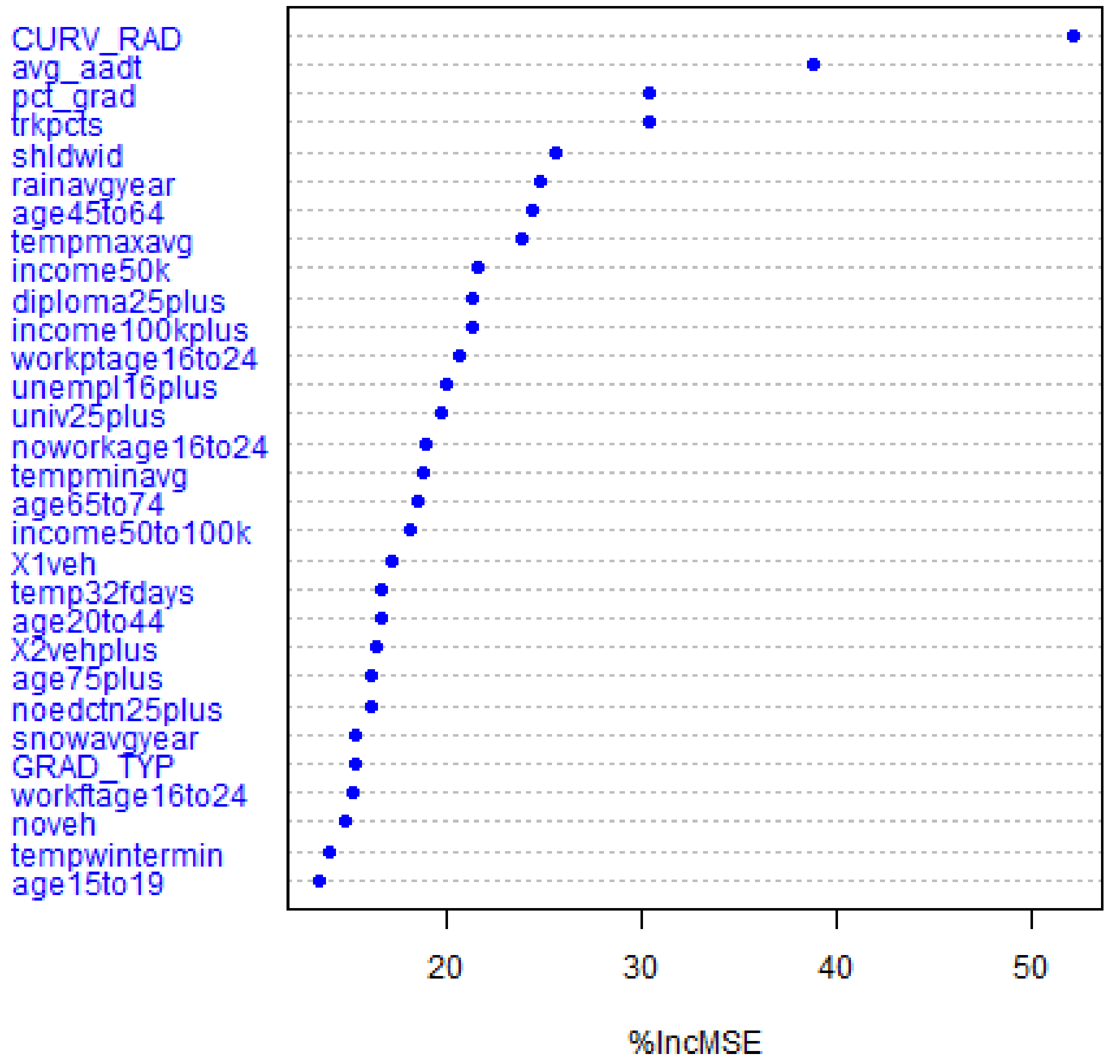
Source: FHWA.

Figure 26. Graph. LNDP-D crashes at tangent segments on rural two-lane highway segments.



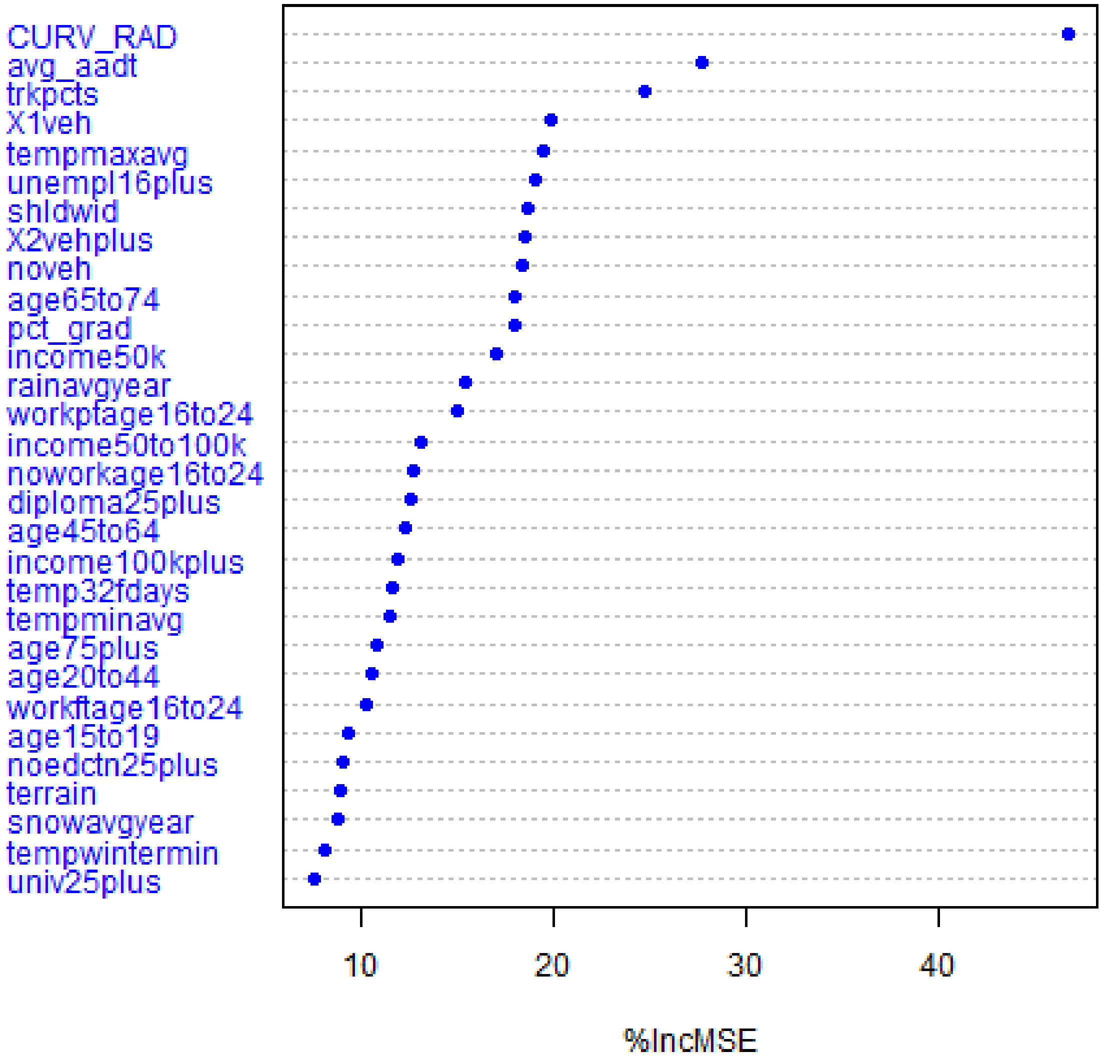
Source: FHWA.

Figure 27. Graph. LNDP-N crashes at tangent segments on rural two-lane highway segments.



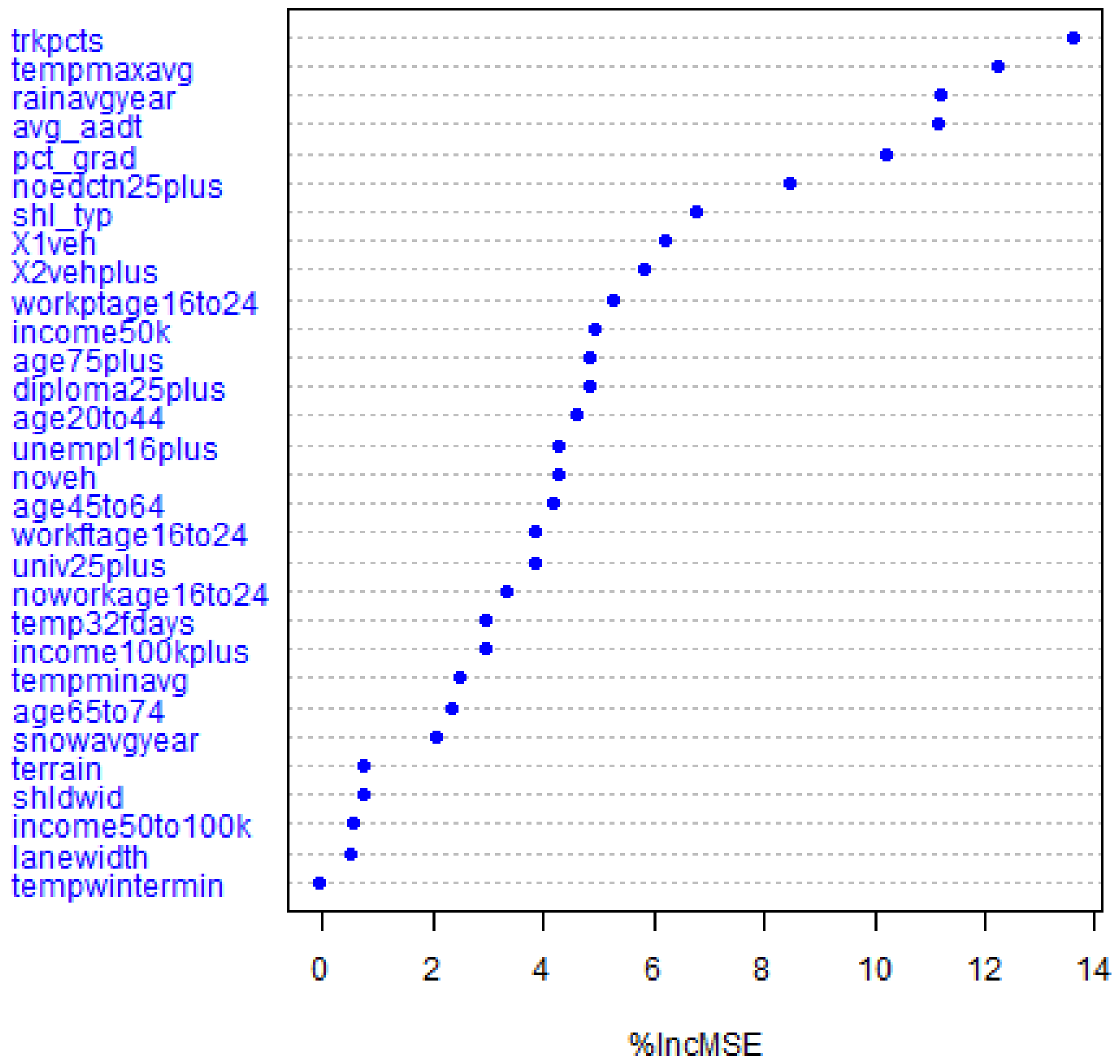
Source: FHWA.

Figure 28. Graph. LNDP-D crashes at horizontal curves on rural two-lane highway segments.



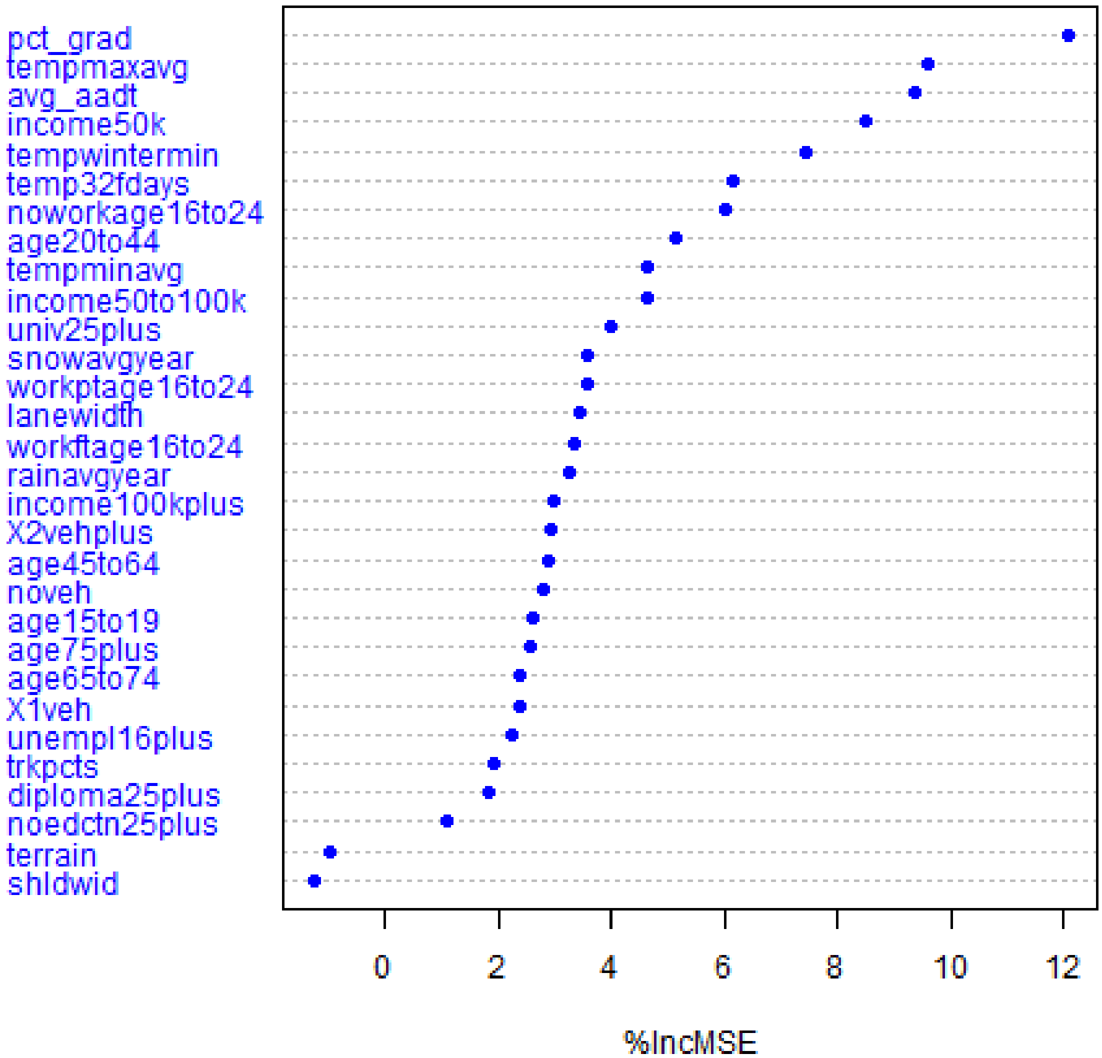
Source: FHWA.

Figure 29. Graph. LNDP-N crashes at horizontal curves on rural two-lane highway segments.



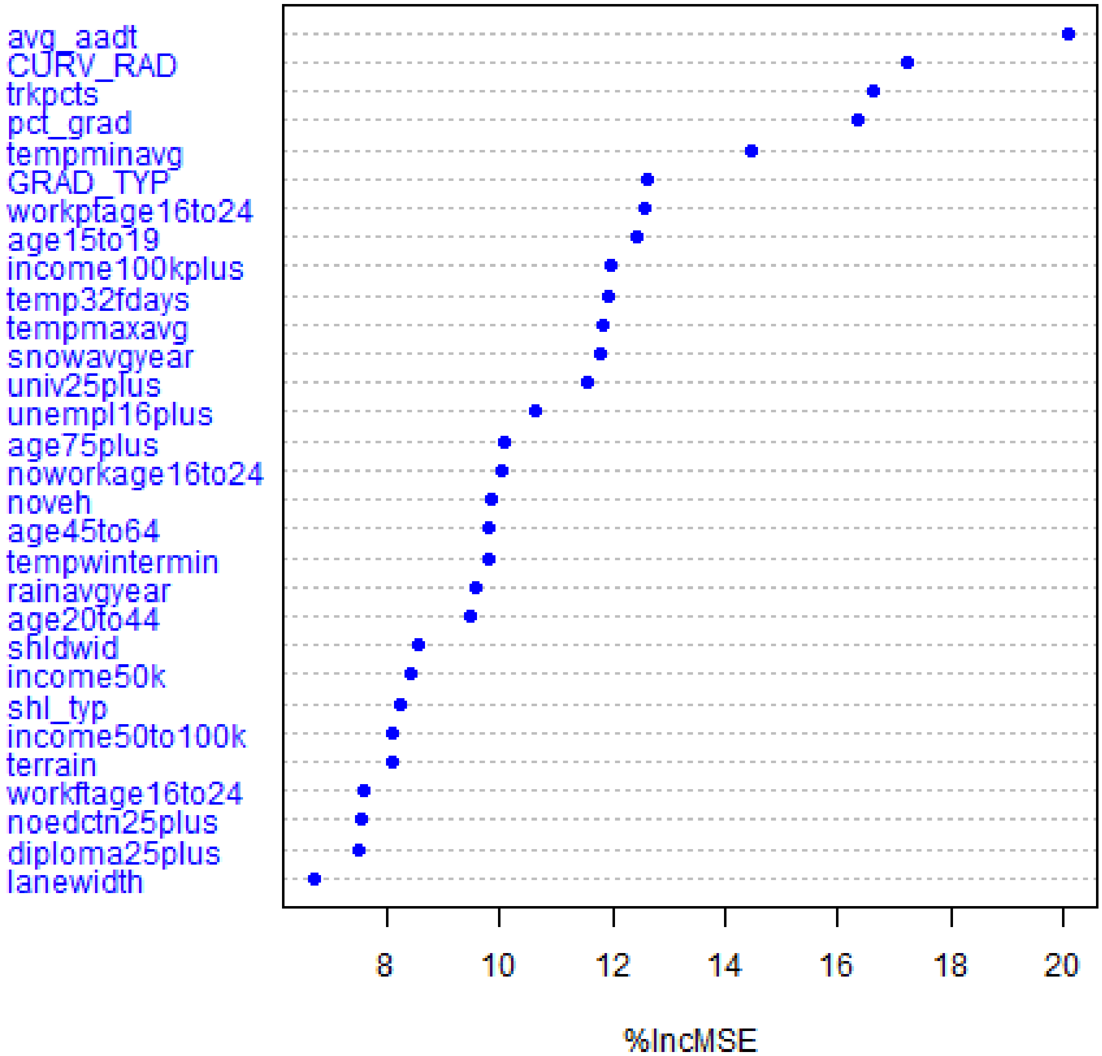
Source: FHWA.

Figure 30. Graph. HEO-D crashes at tangent segments on rural two-lane highway segments.



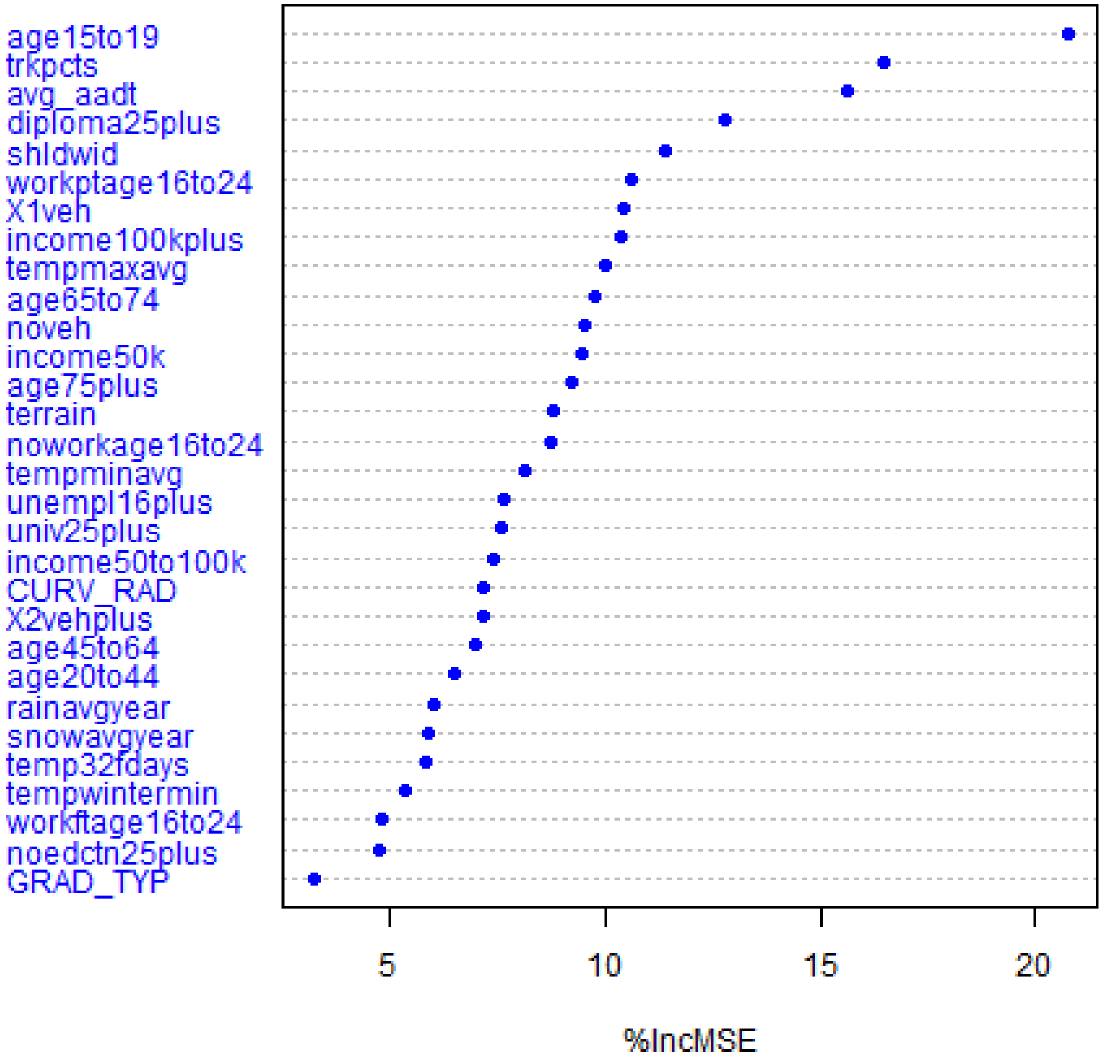
Source: FHWA.

Figure 31. Graph. HEO-N crashes at tangent segments on rural two-lane highway segments.



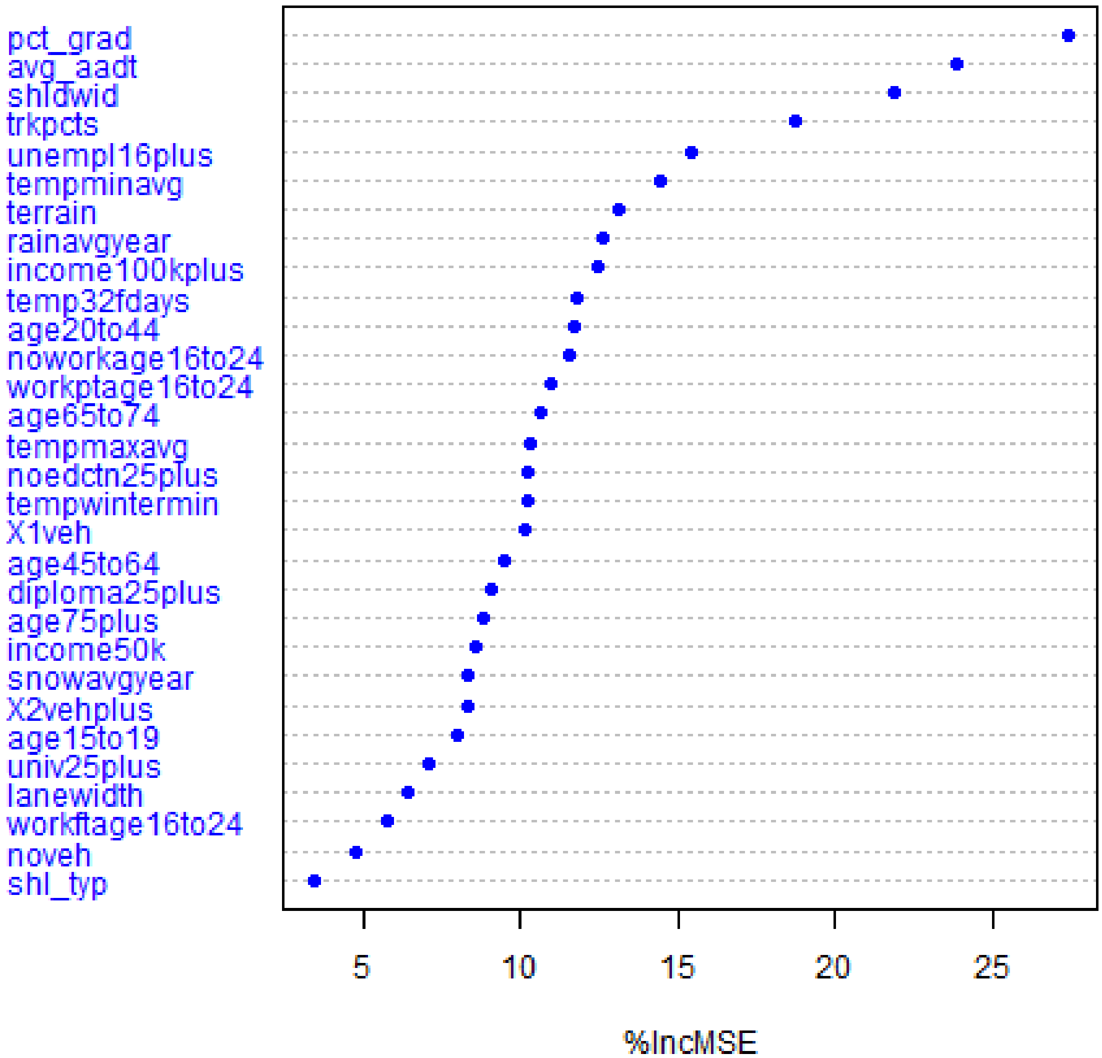
Source: FHWA.

Figure 32. Graph. HEO-D crashes at horizontal curves on rural two-lane highway segments.



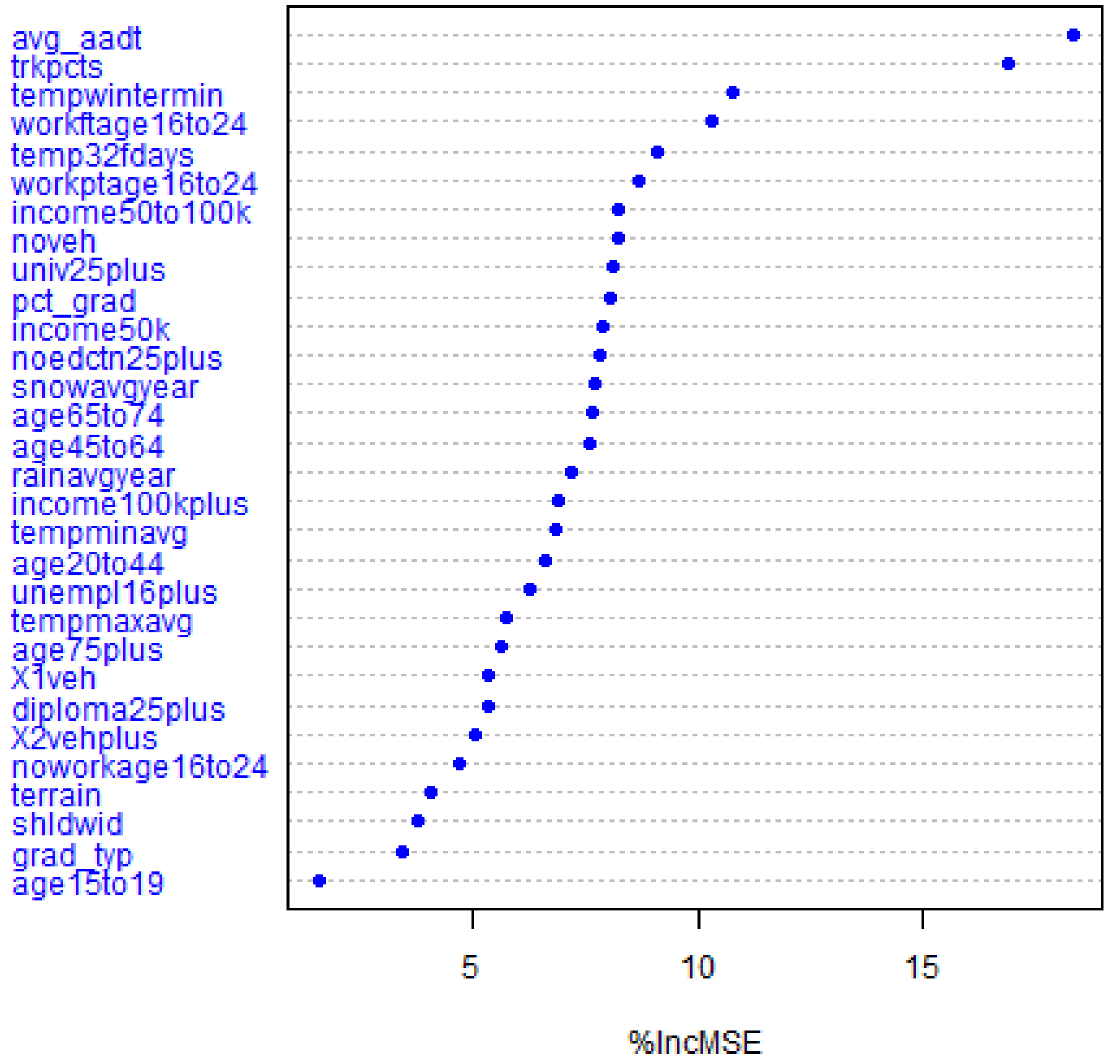
Source: FHWA.

Figure 33. Graph. HEO-N crashes at horizontal curves on rural two-lane highway segments.



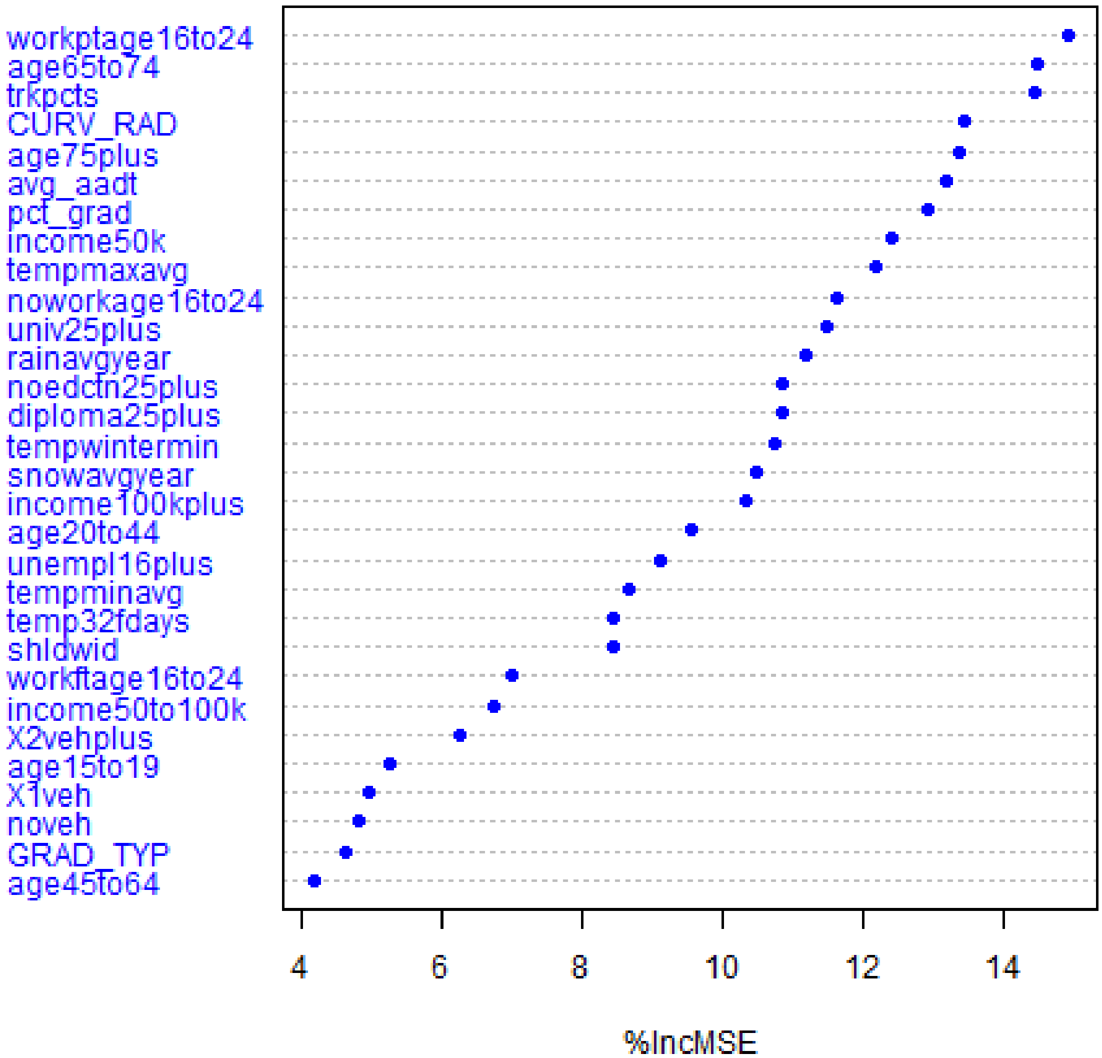
Source: FHWA.

Figure 34. Graph. ROLL-D crashes at tangent segments on rural two-lane highway segments.



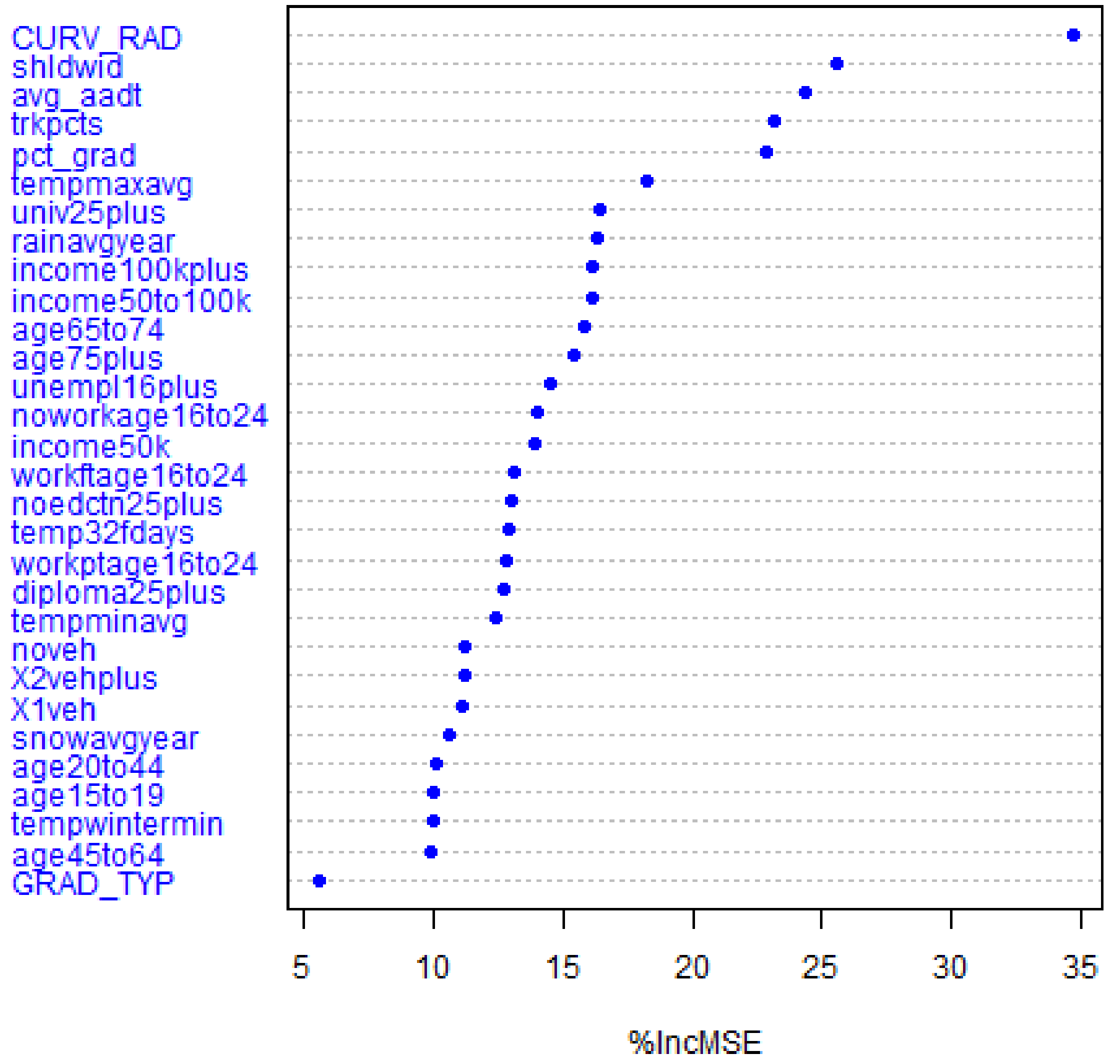
Source: FHWA.

Figure 35. Graph. ROLL-N crashes at tangent segments on rural two-lane highway segments.



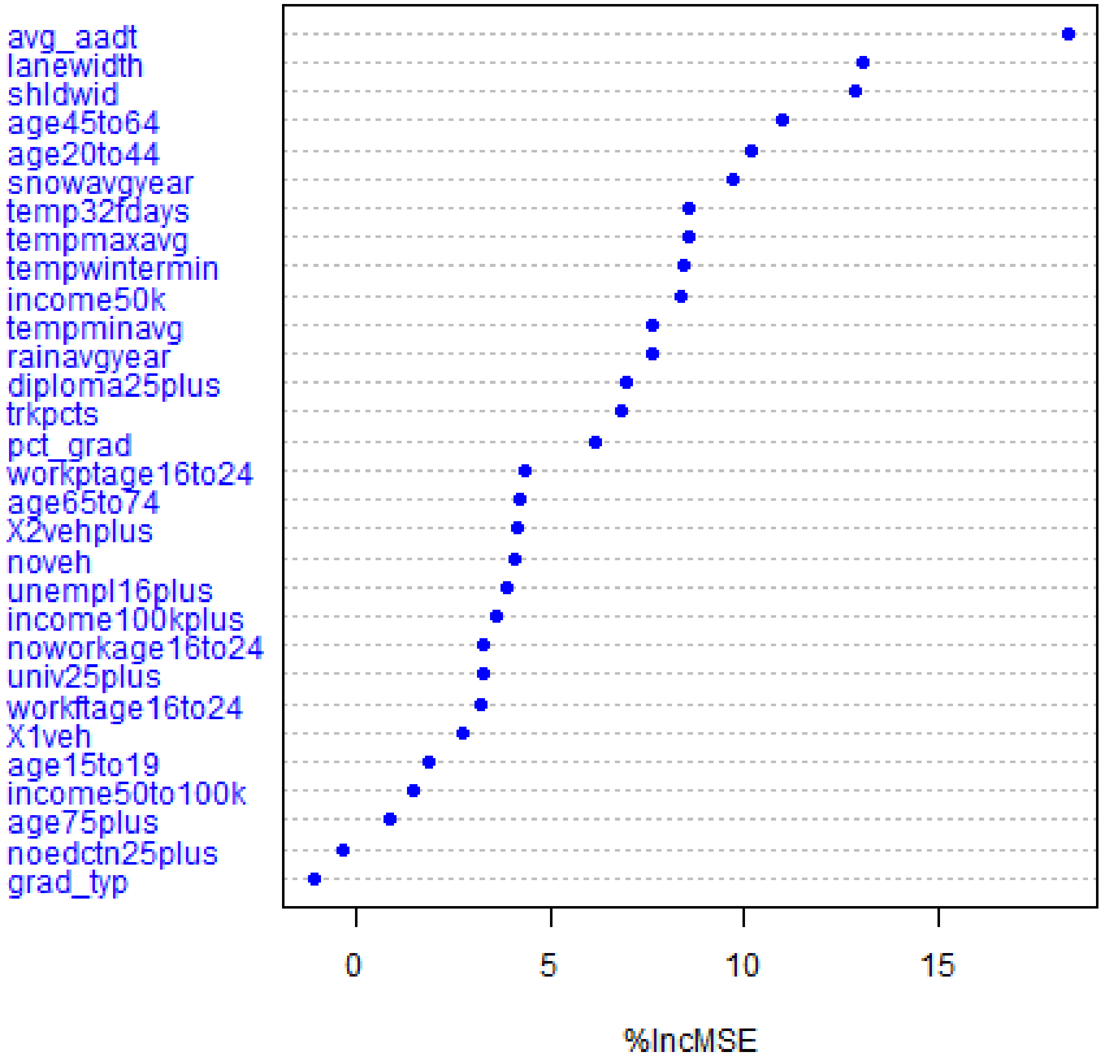
Source: FHWA.

Figure 36. Graph. ROLL-D crashes at horizontal curves on rural two-lane highway segments.



Source: FHWA.

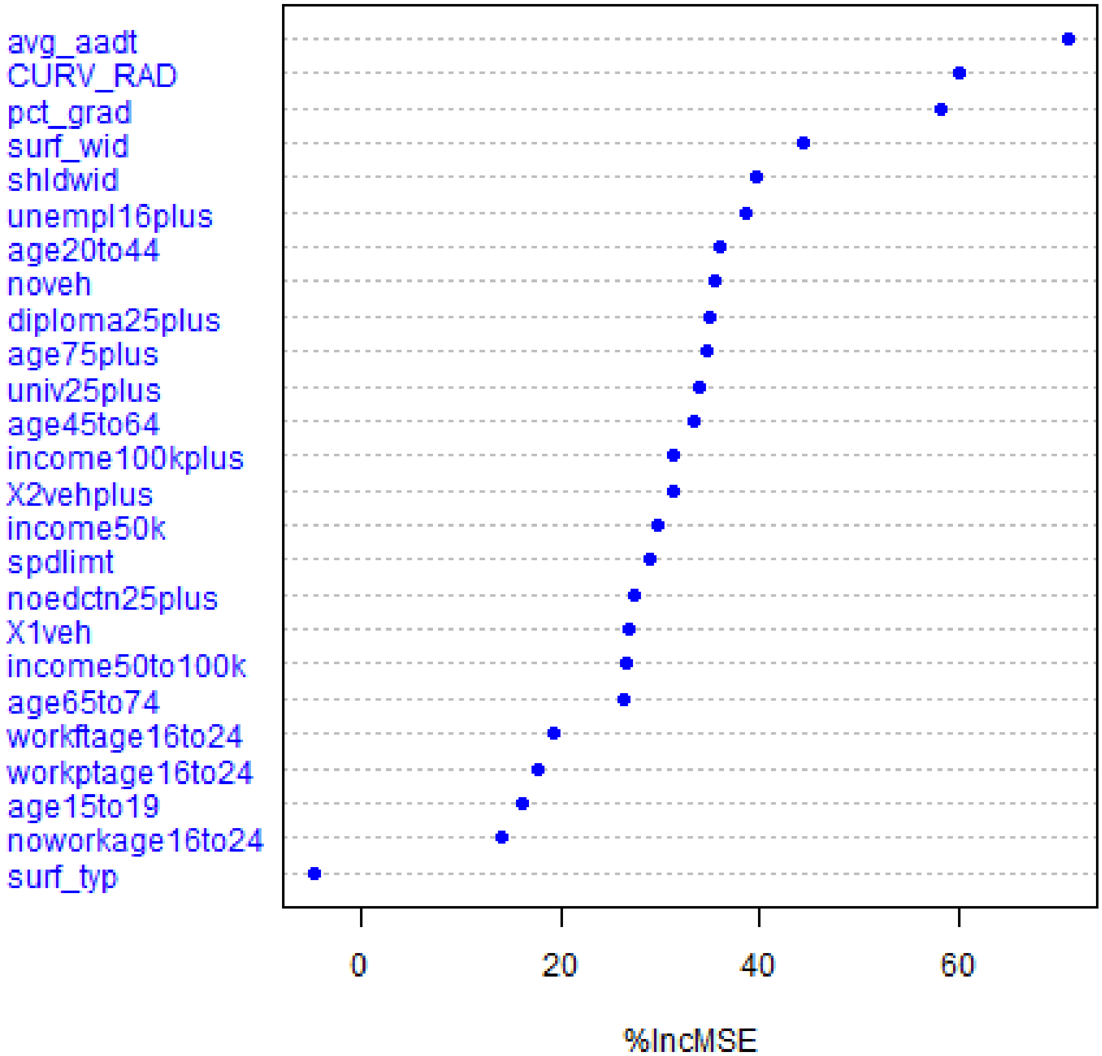
Figure 37. Graph. ROLL-N crashes at horizontal curves on rural two-lane highway segments.



Source: FHWA.

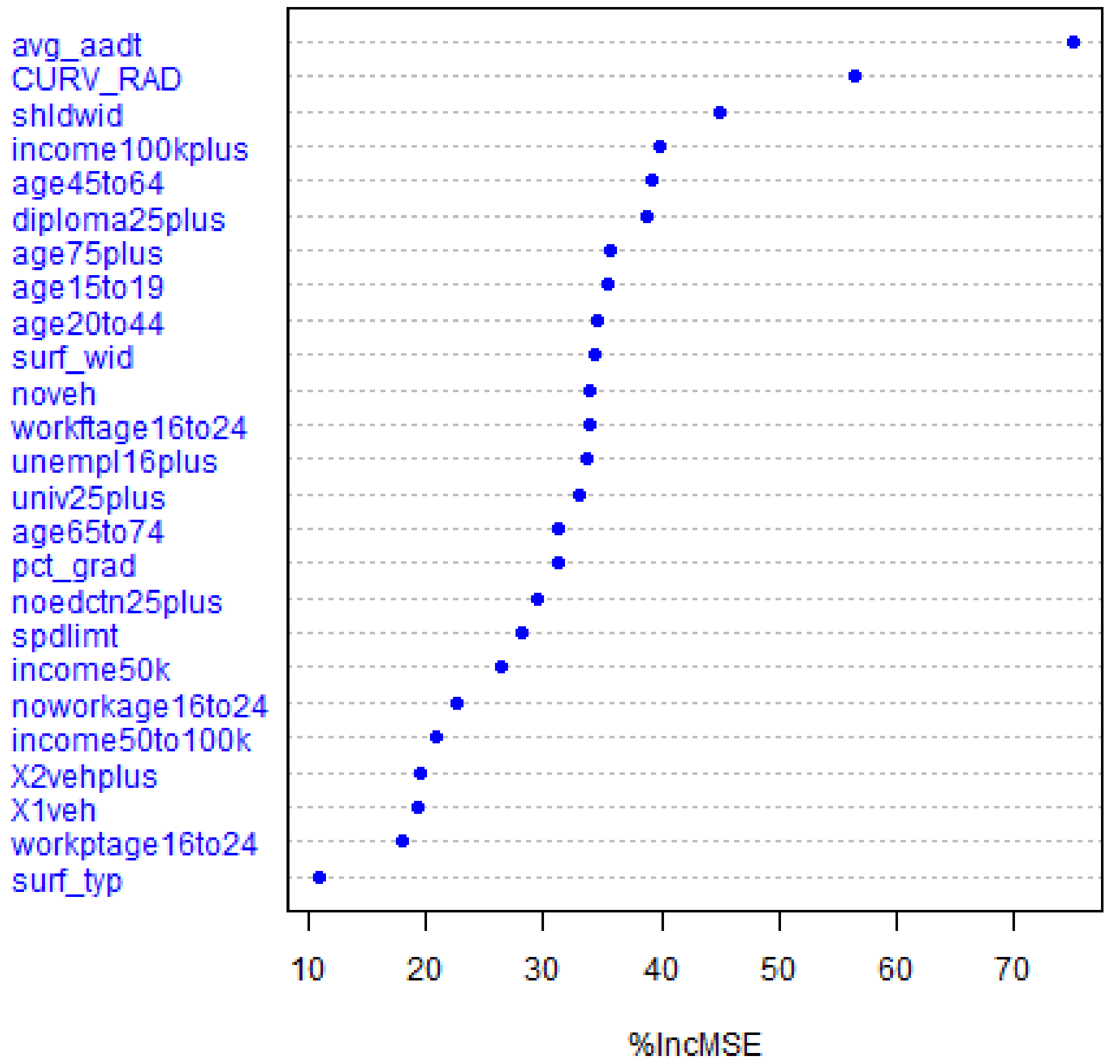
Figure 38. Graph. ANG-D crashes at tangent segments on rural two-lane highway segments.

RANDOM FORESTS OF OHIO DATA



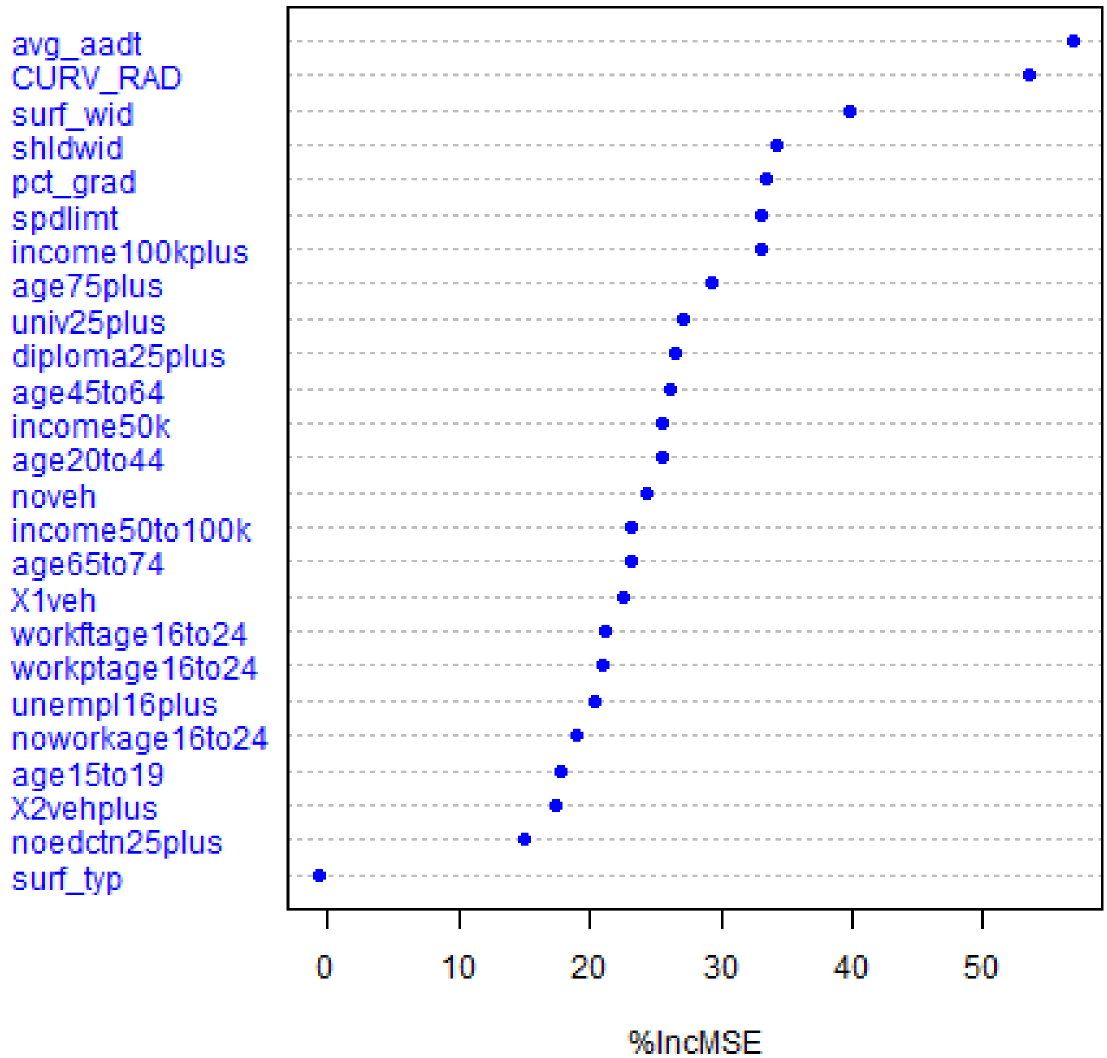
Source: FHWA.

Figure 39. Graph. ROR-KAB-D crashes on rural two-lane highway segments.



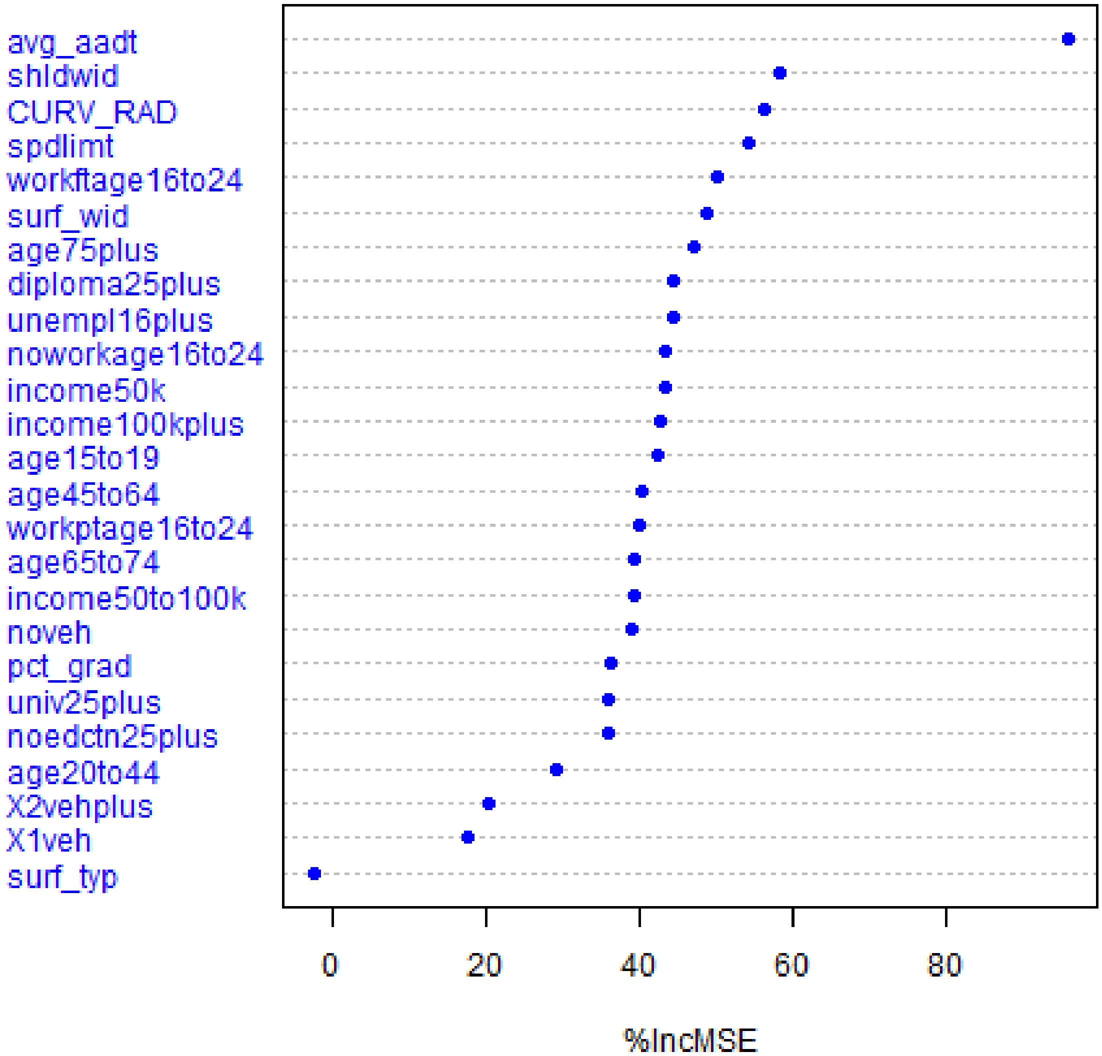
Source: FHWA.

Figure 40. Graph. ROR-KAB-N crashes on rural two-lane highway segments.



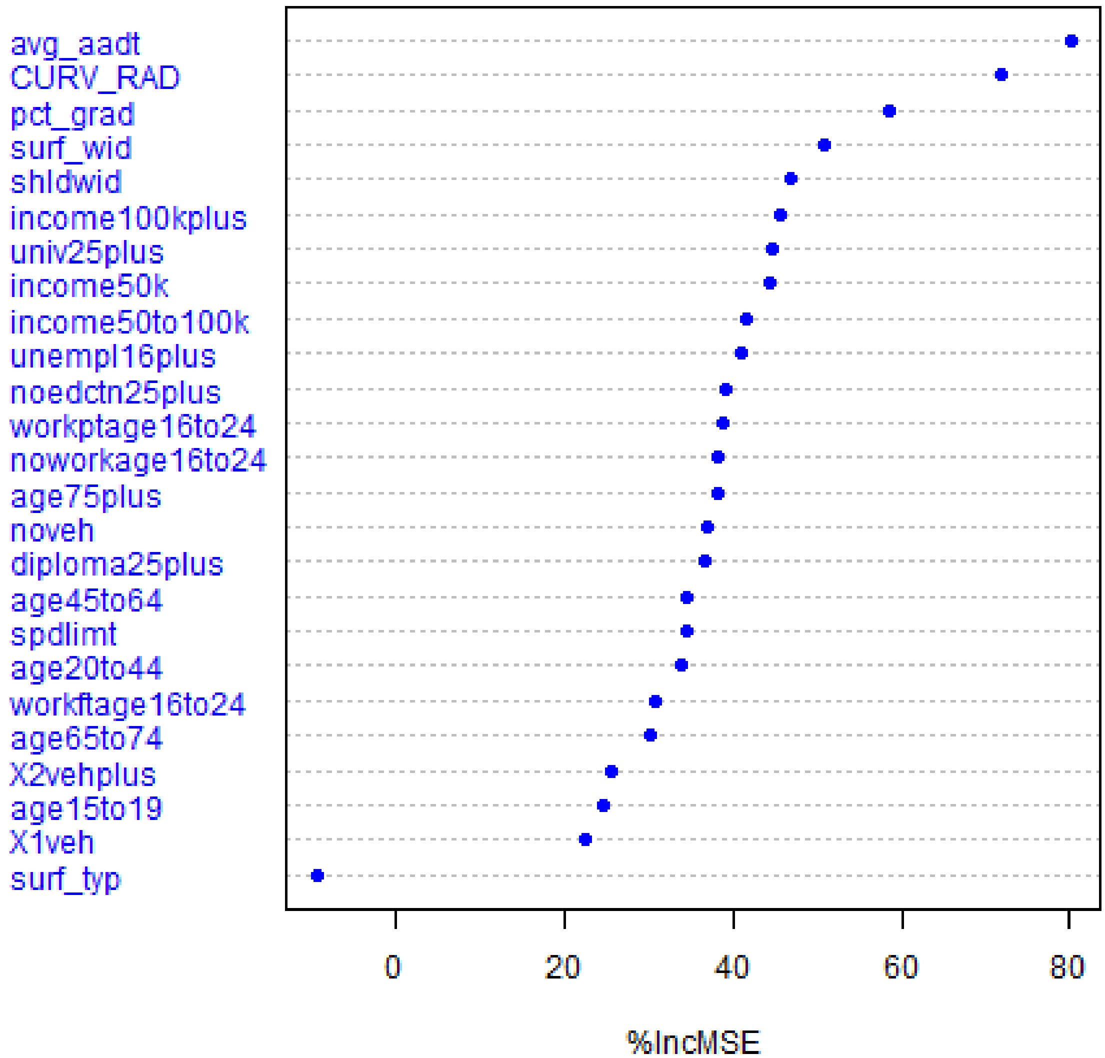
Source: FHWA.

Figure 41. Graph. ROR-KABCO-D crashes on rural two-lane highway segments.



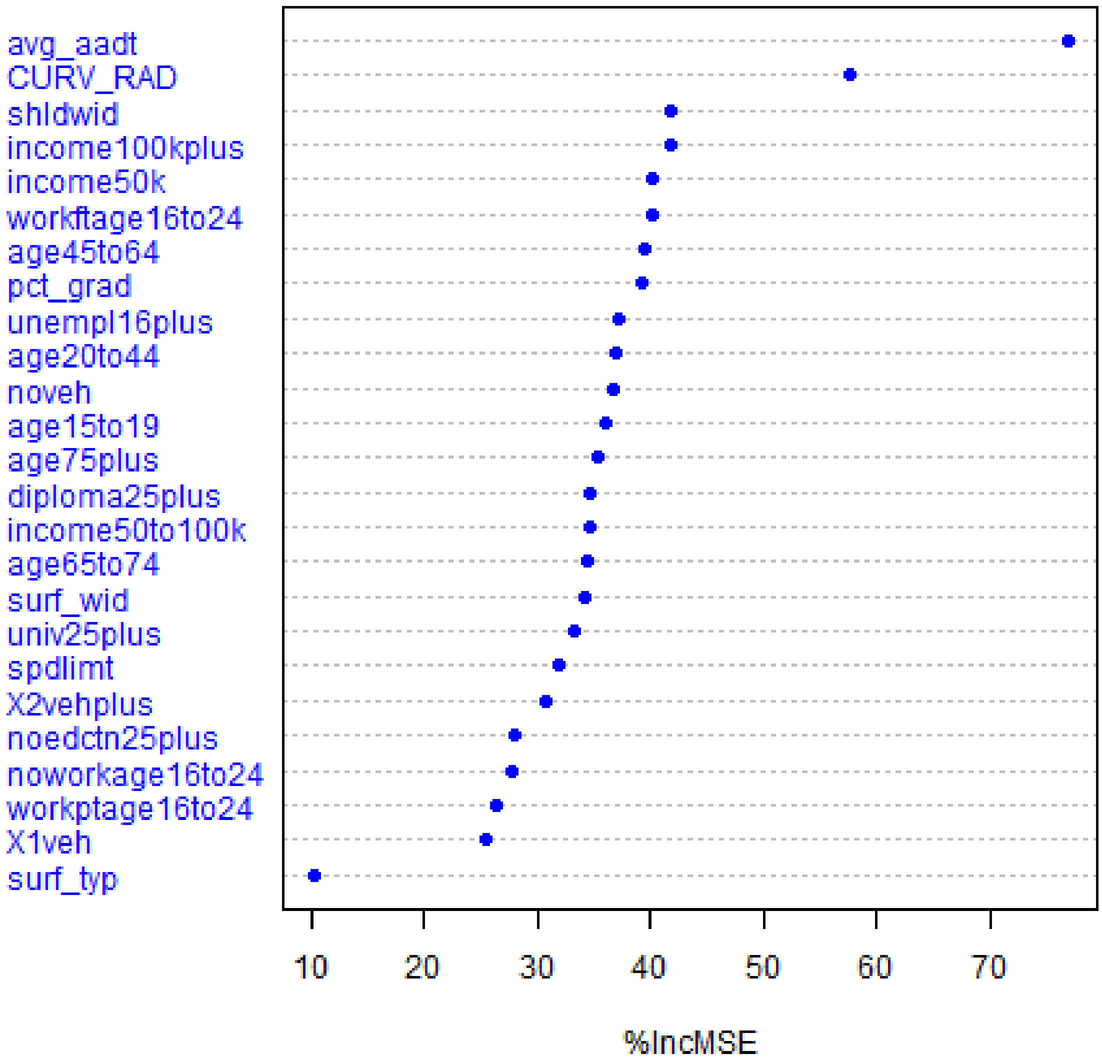
Source: FHWA.

Figure 42. Graph. ROR-KABCO-N crashes on rural two-lane highway segments.



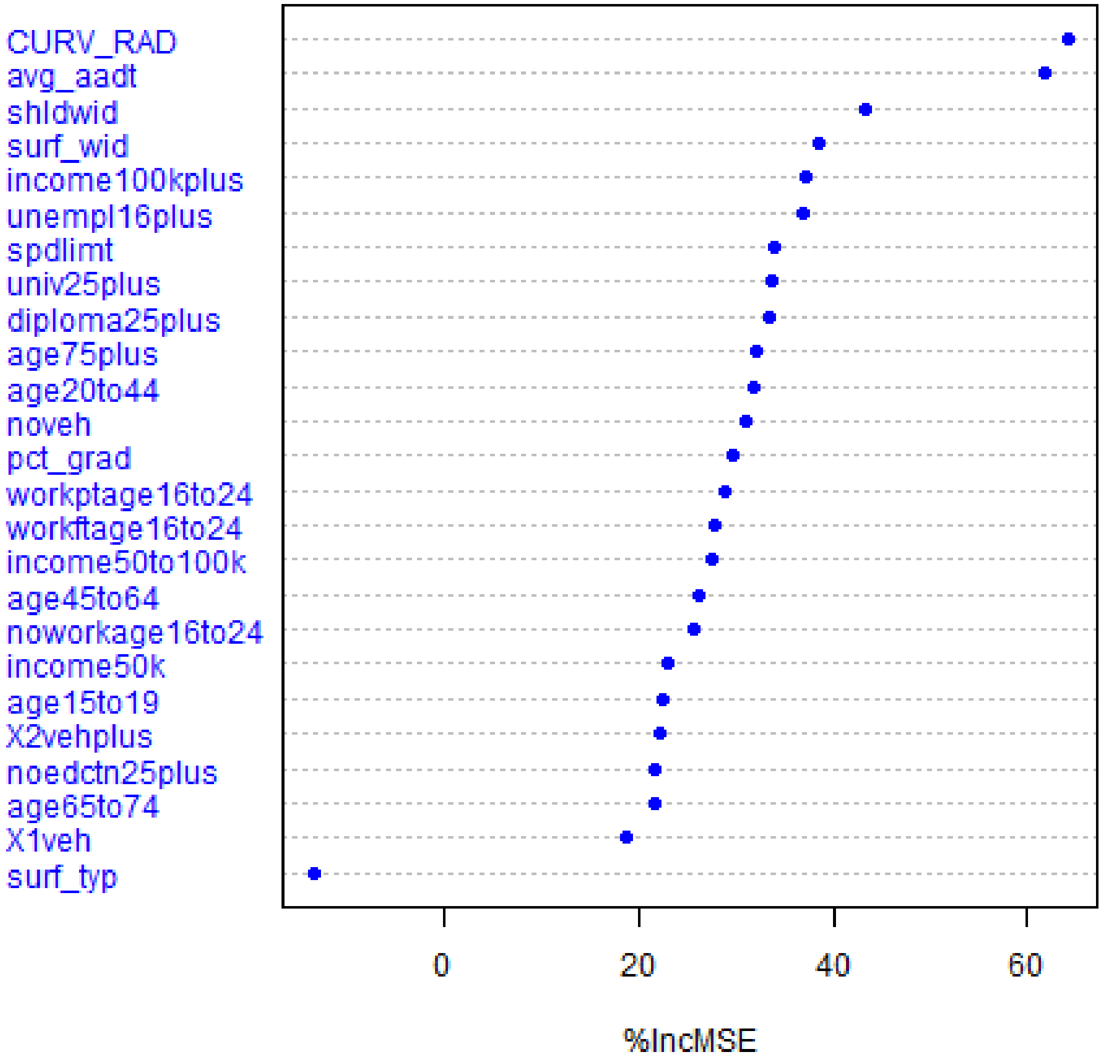
Source: FHWA.

Figure 43. Graph. LNDP-KAB-D crashes on rural two-lane highway segments.



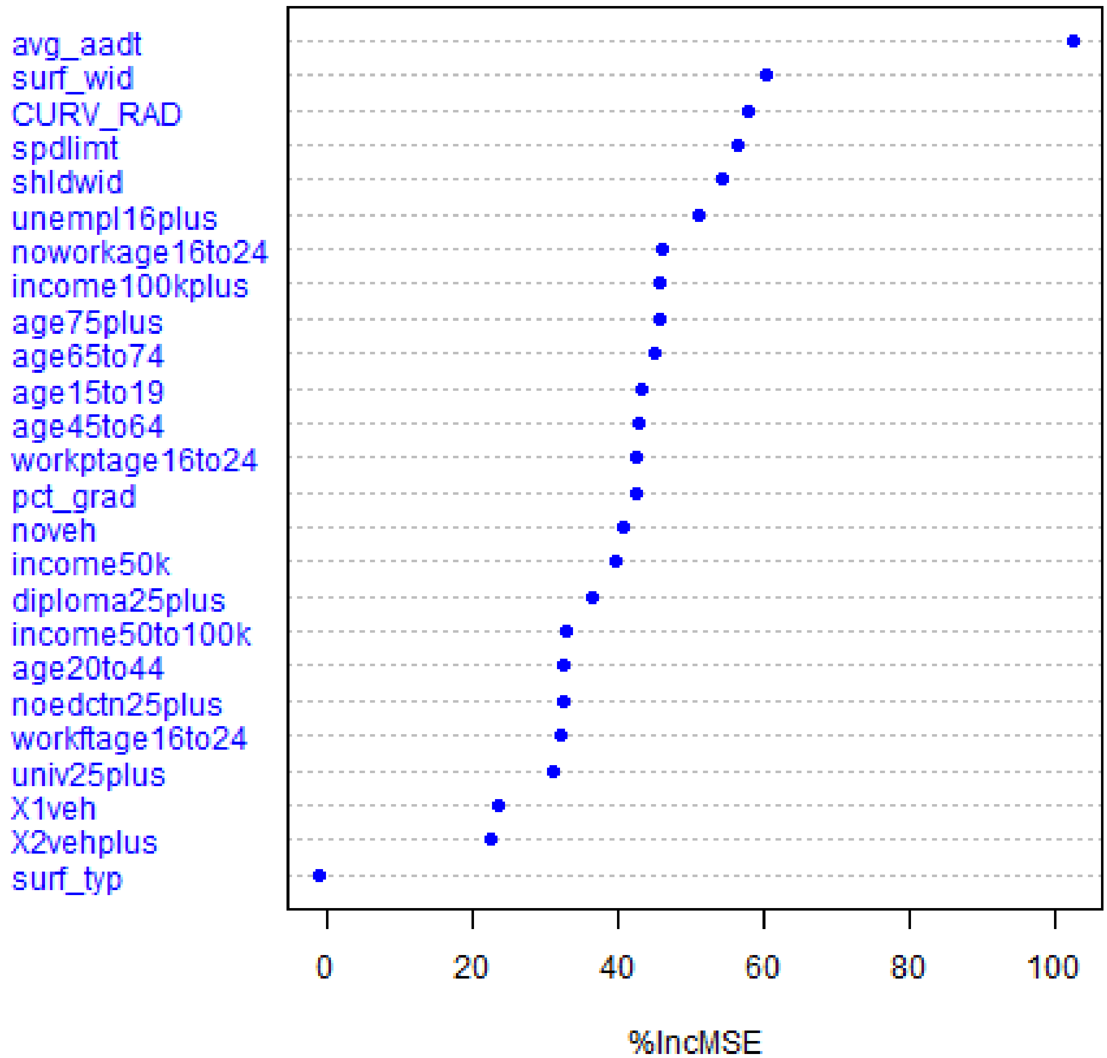
Source: FHWA.

Figure 44. Graph. LNDP-KAB-N crashes on rural two-lane highway segments.



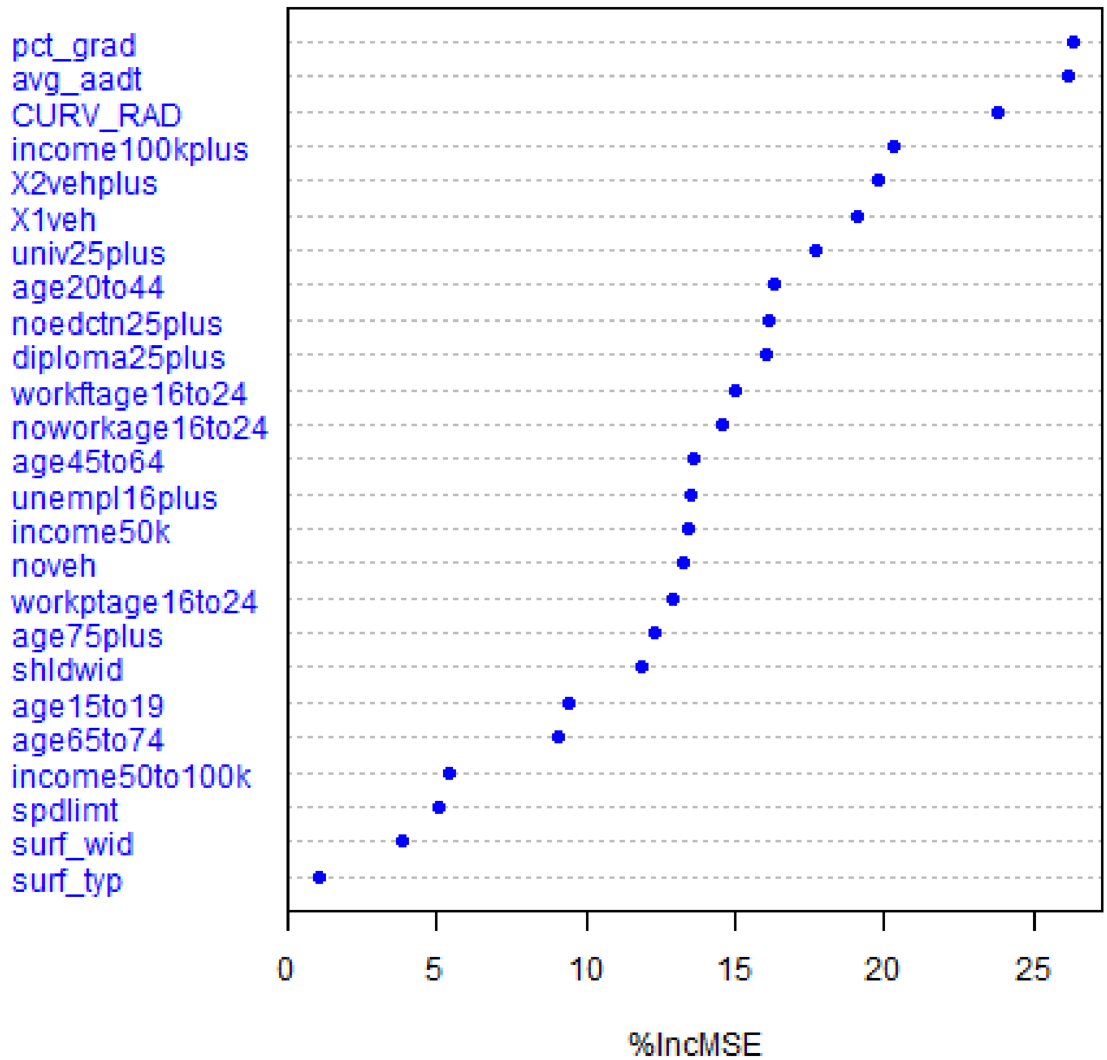
Source: FHWA.

Figure 45. Graph. LNDP-KABCO-D crashes on rural two-lane highway segments.



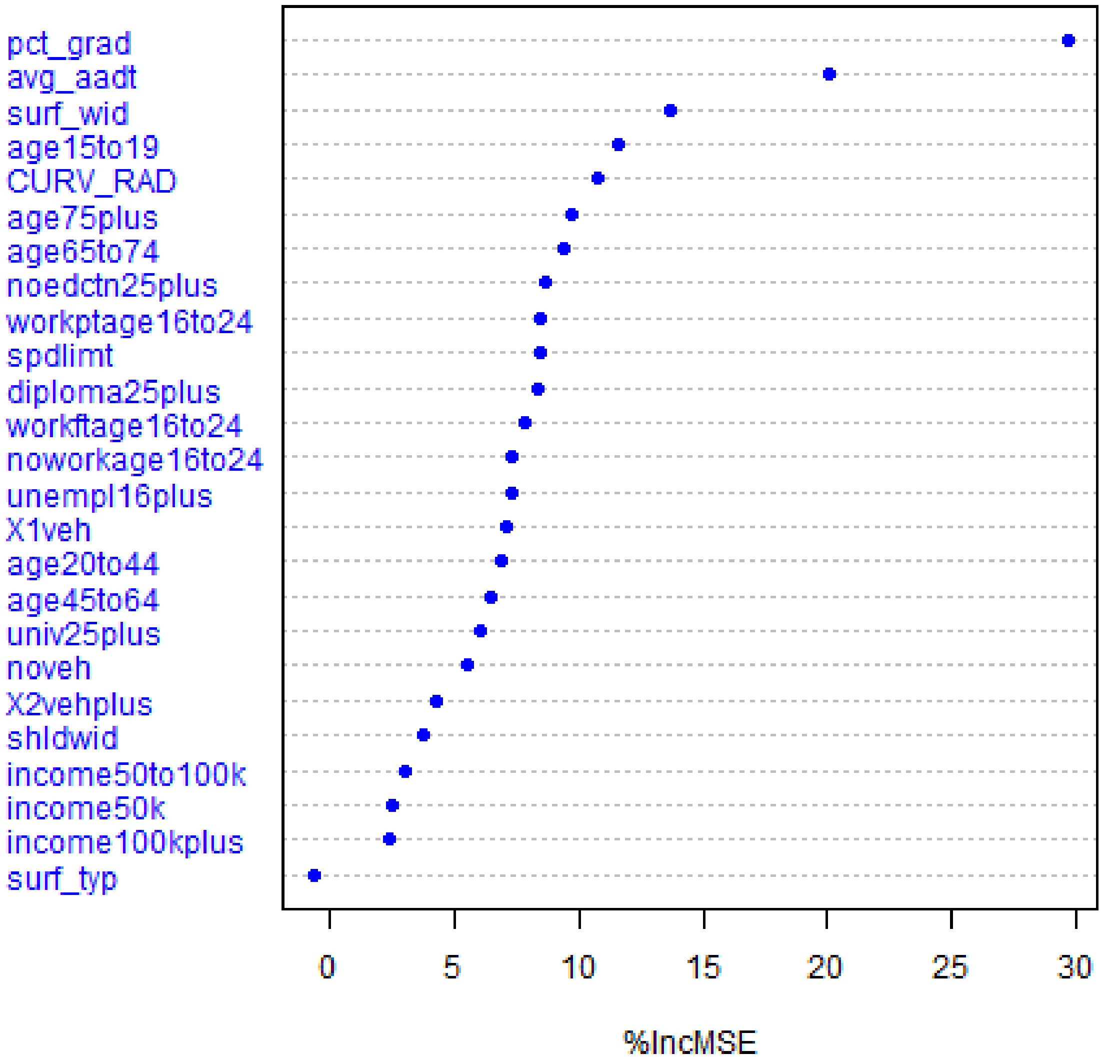
Source: FHWA.

Figure 46. Graph. LNDP-KABCO-N crashes on rural two-lane highway segments.



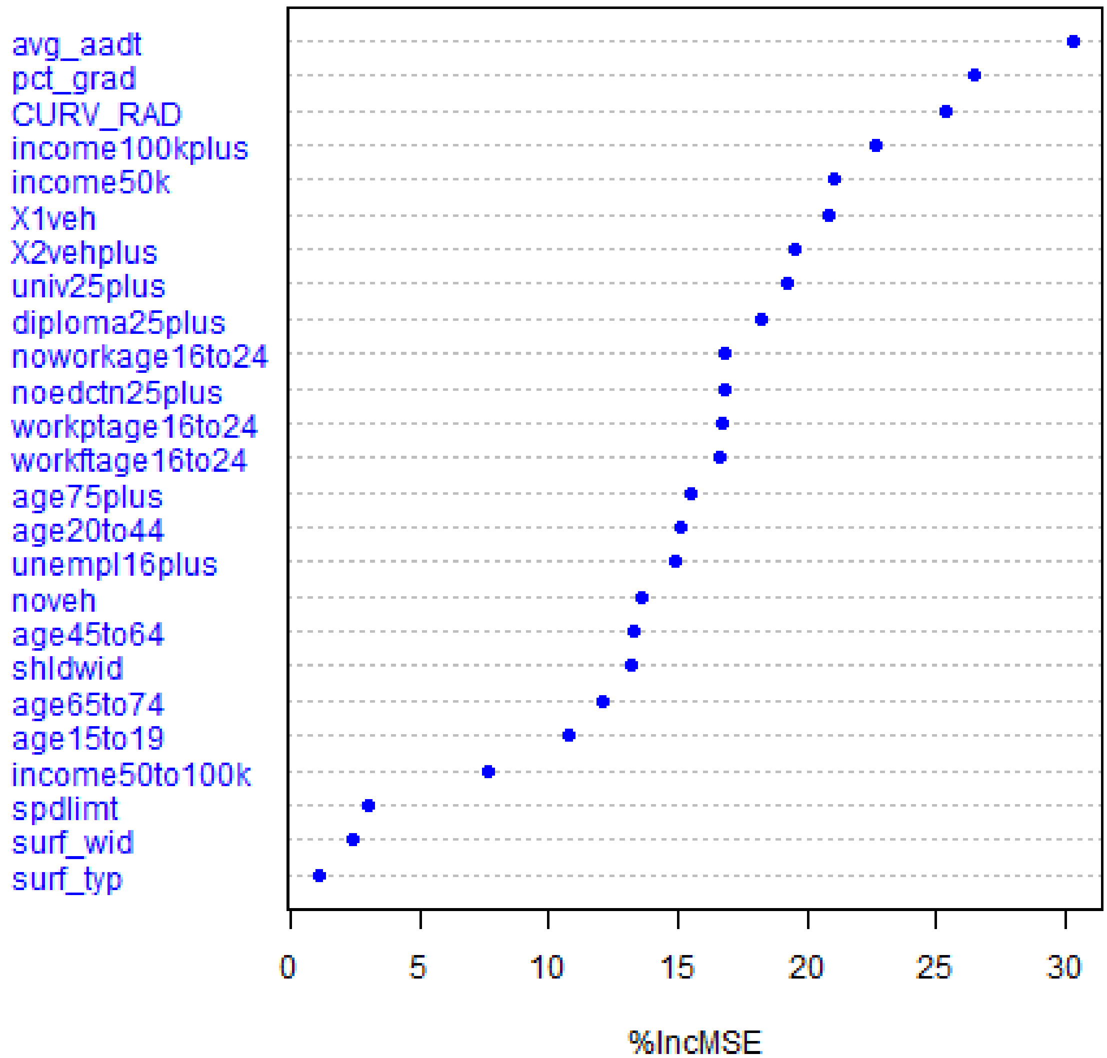
Source: FHWA.

Figure 47. Graph. HEO-KAB-D crashes on rural two-lane highway segments.



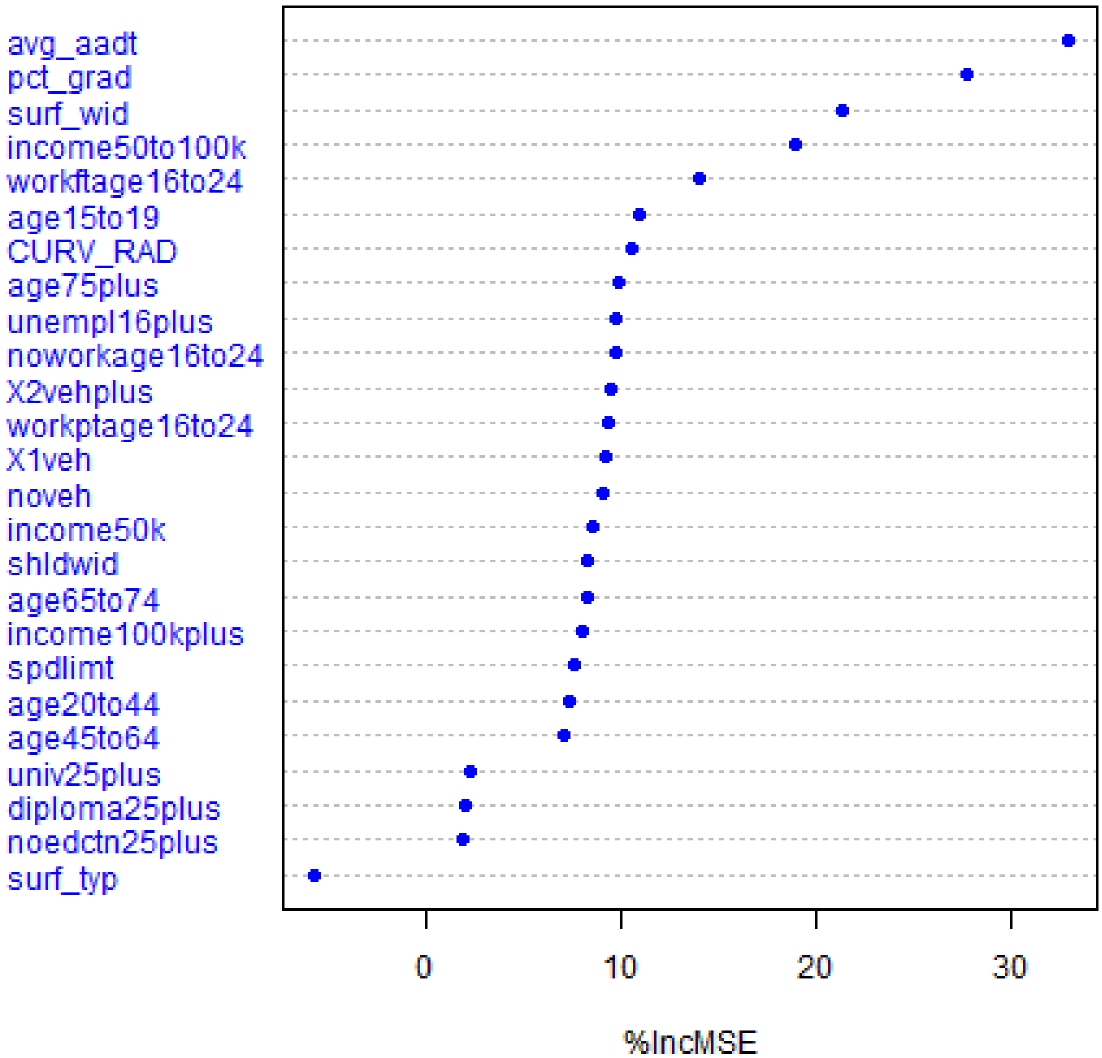
Source: FHWA.

Figure 48. Graph. HEO-KAB-N crashes on rural two-lane highway segments.



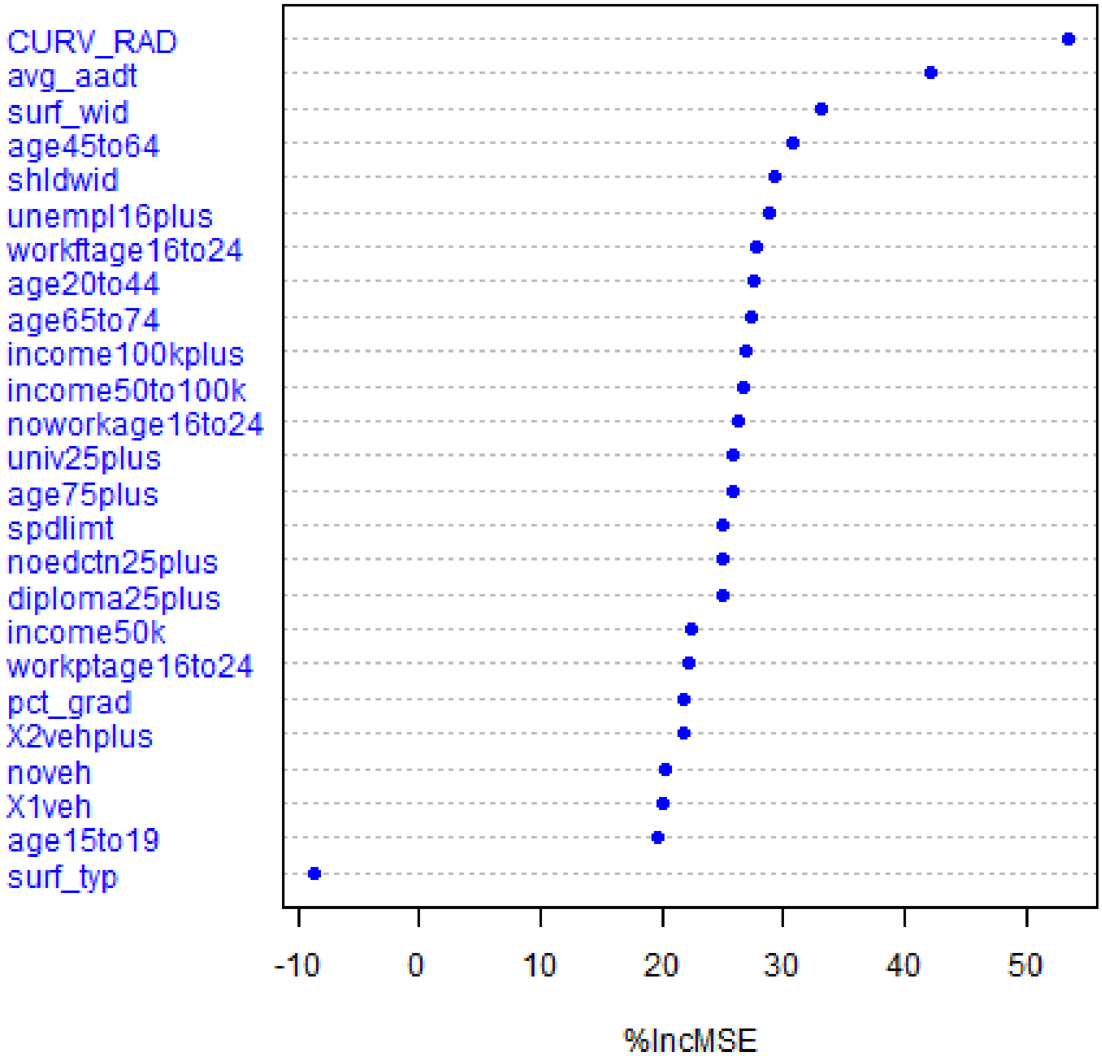
Source: FHWA.

Figure 49. Graph. HEO-KABCO-D crashes on rural two-lane highway segments.



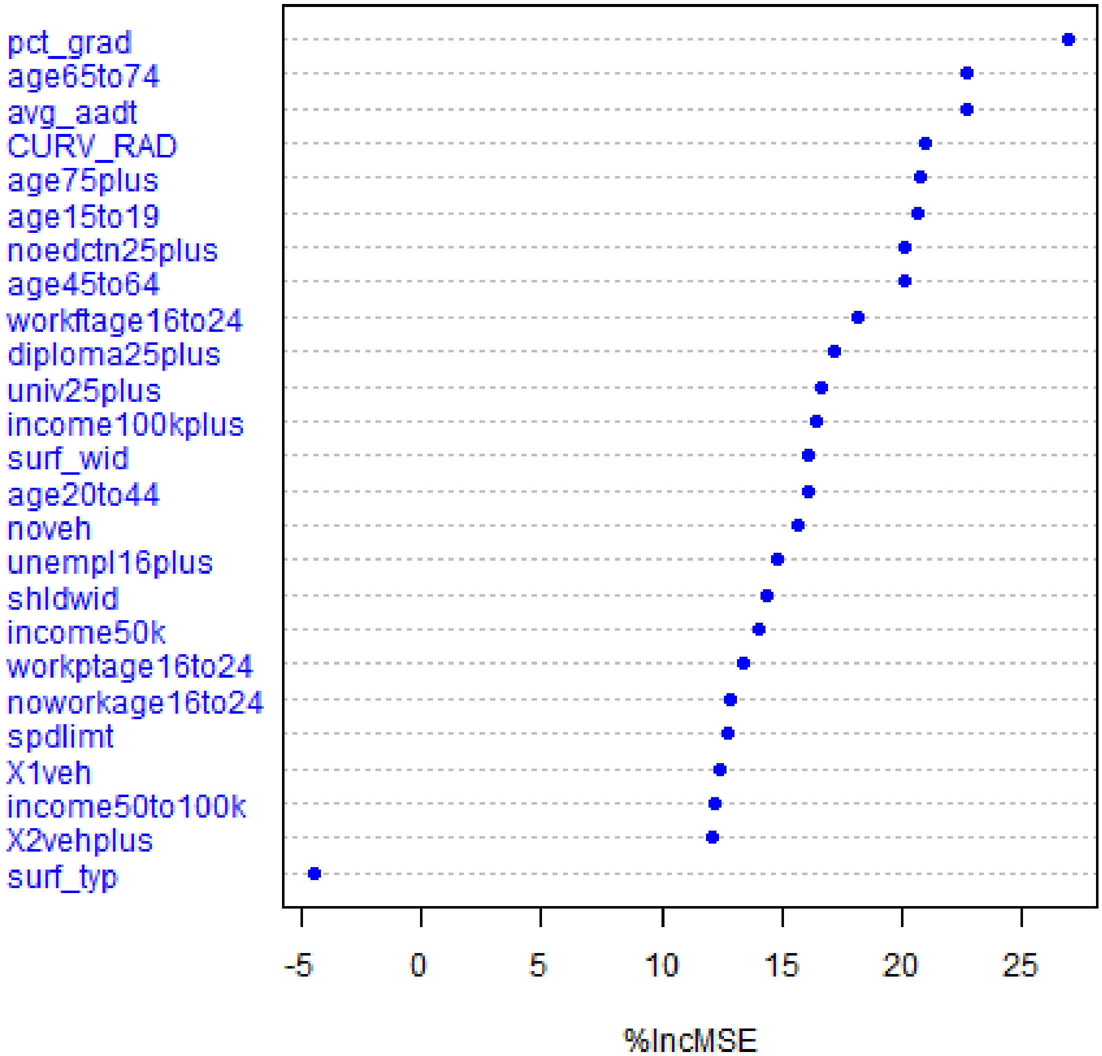
Source: FHWA.

Figure 50. Graph. HEO-KABCO-N crashes on rural two-lane highway segments.



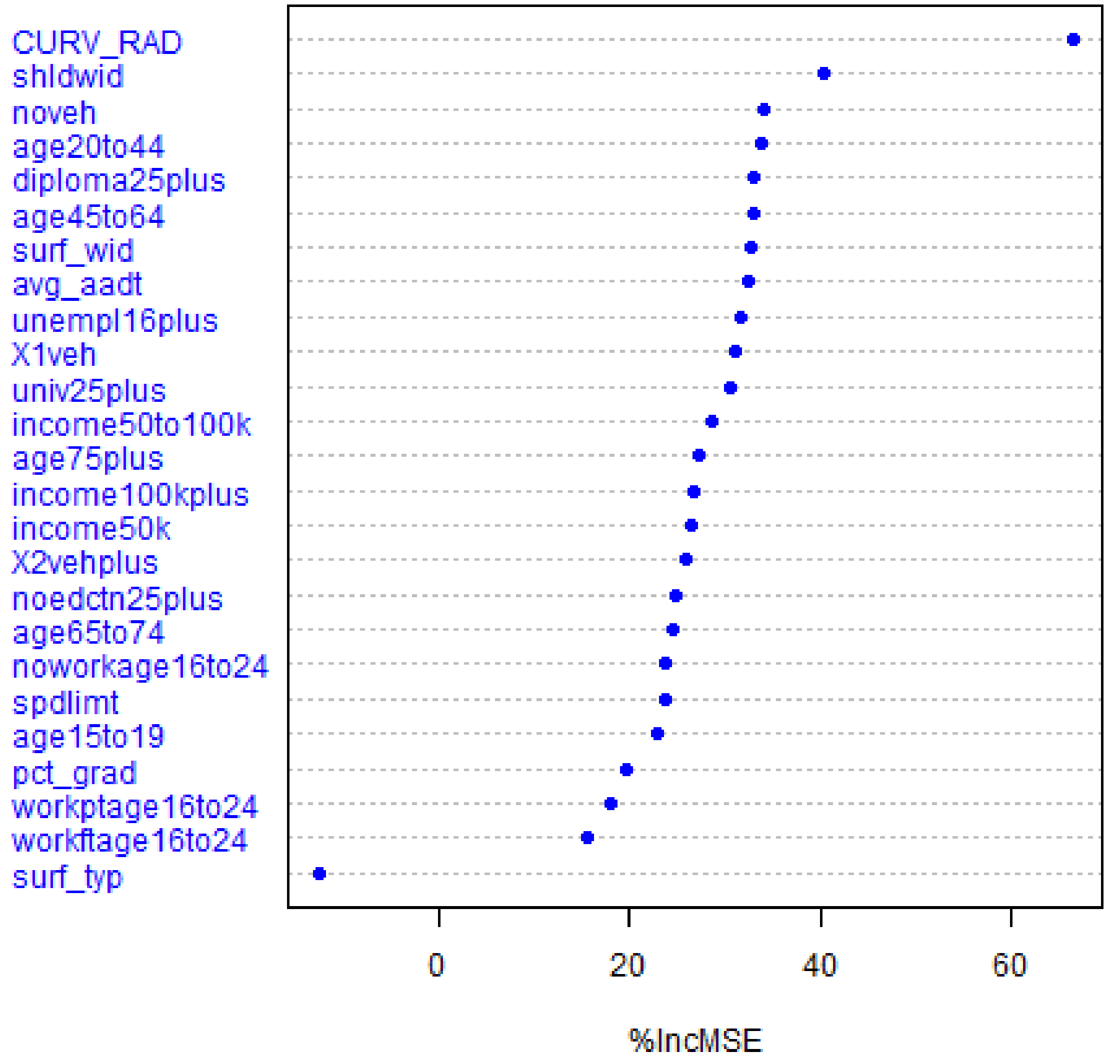
Source: FHWA.

Figure 51. Graph. ROLL-KAB-D crashes on rural two-lane highway segments.



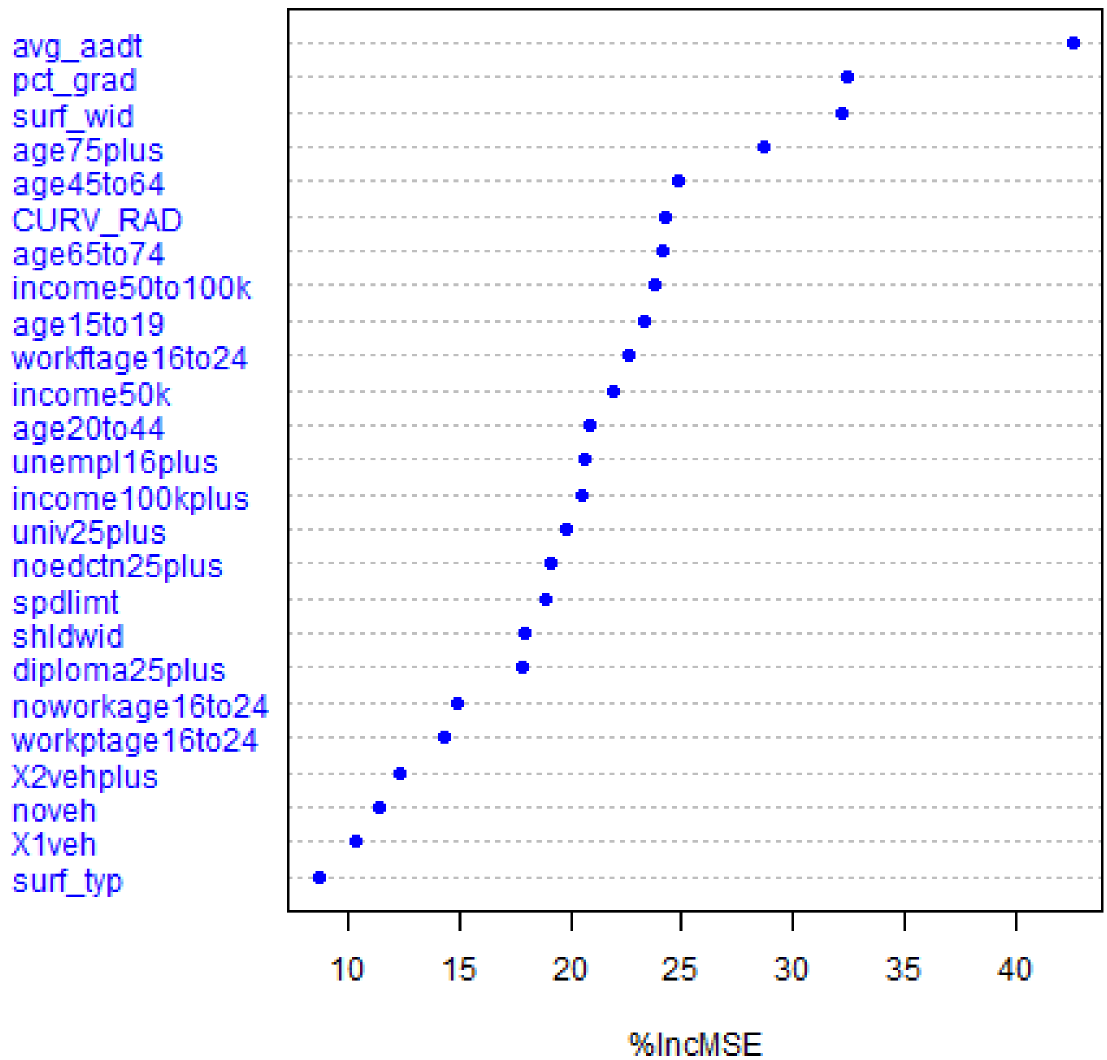
Source: FHWA.

Figure 52. Graph. ROLL-KAB-N crashes on rural two-lane highway segments.



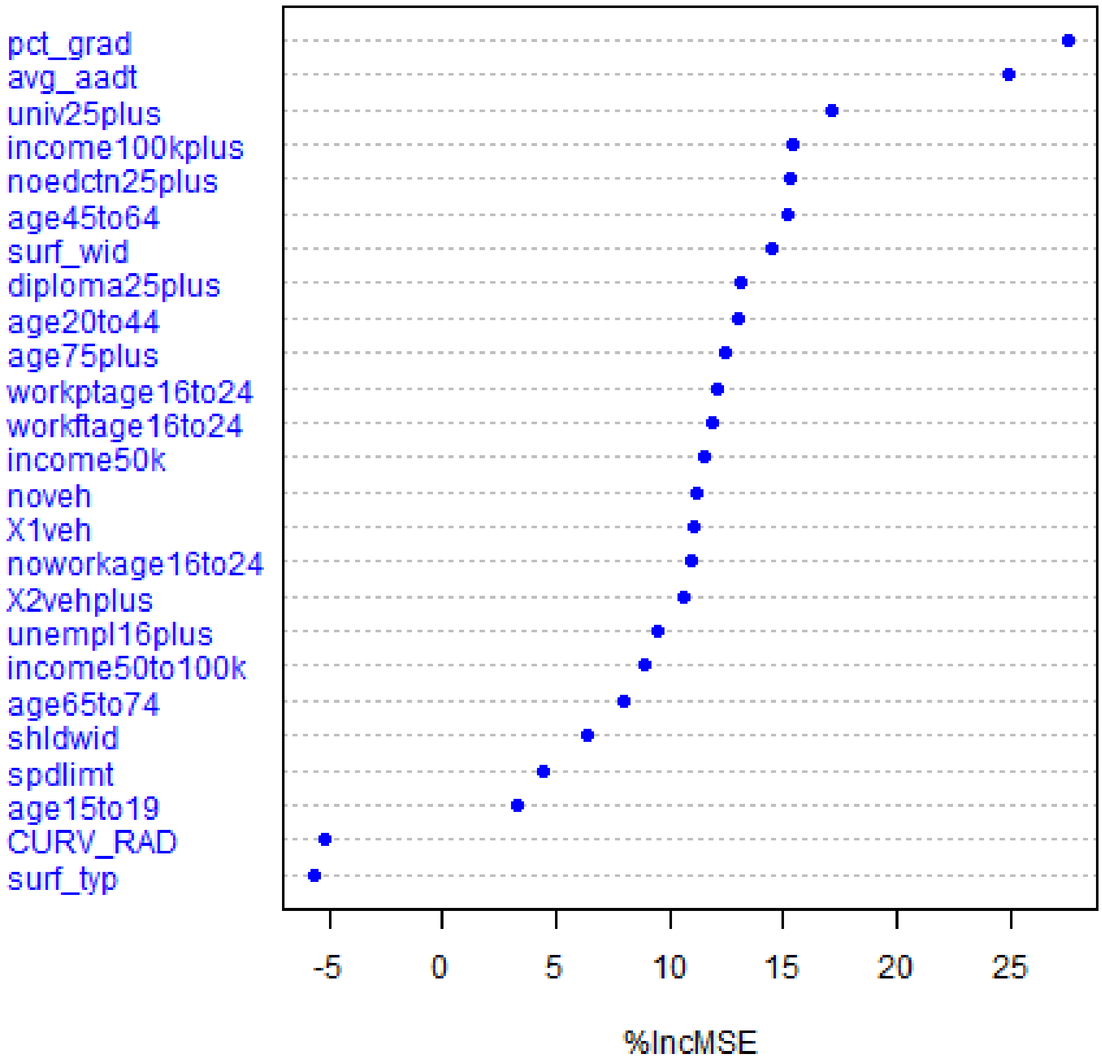
Source: FHWA.

Figure 53. Graph. ROLL-KABCO-D crashes on rural two-lane highway segments.



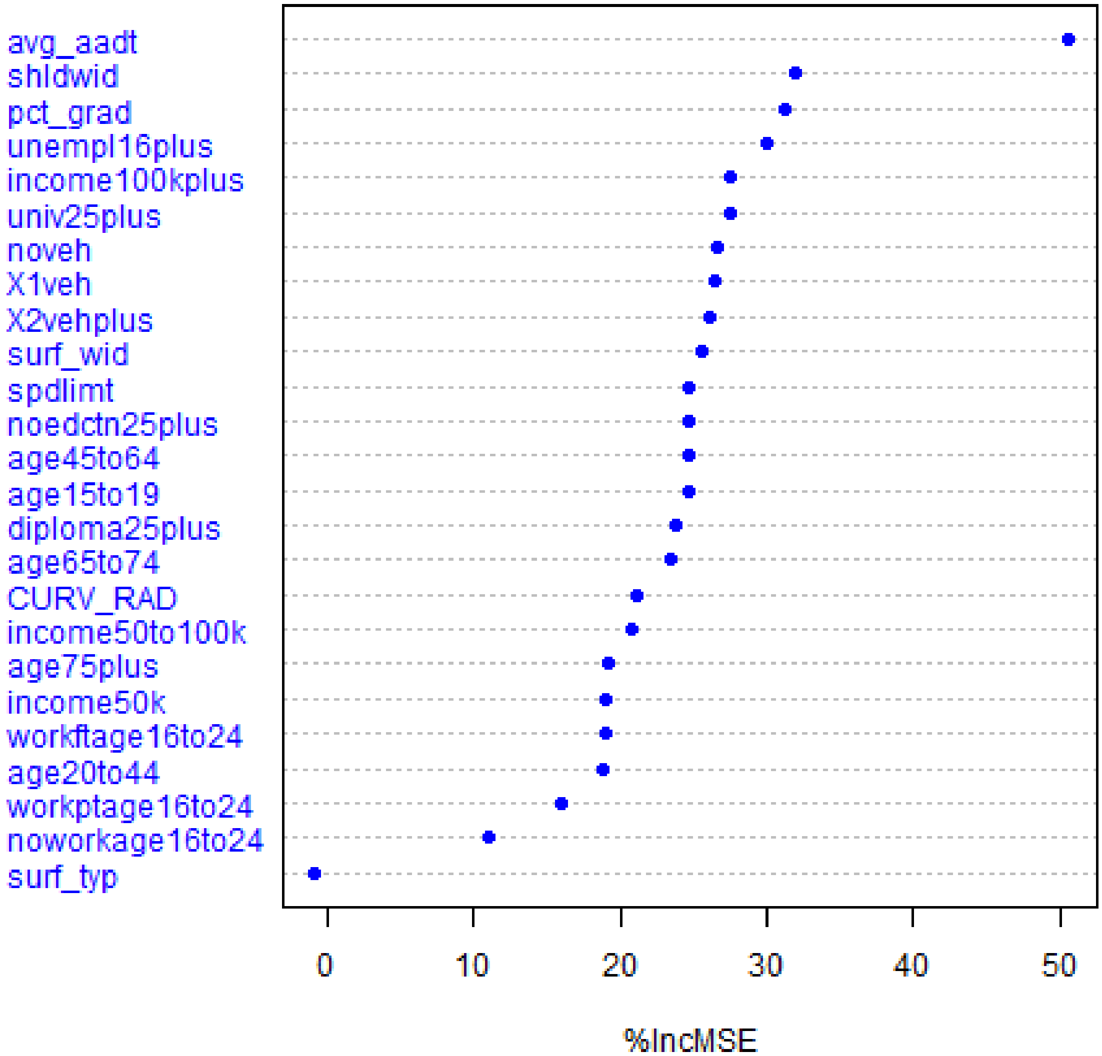
Source: FHWA.

Figure 54. Graph. ROLL-KABCO-N crashes on rural two-lane highway segments.



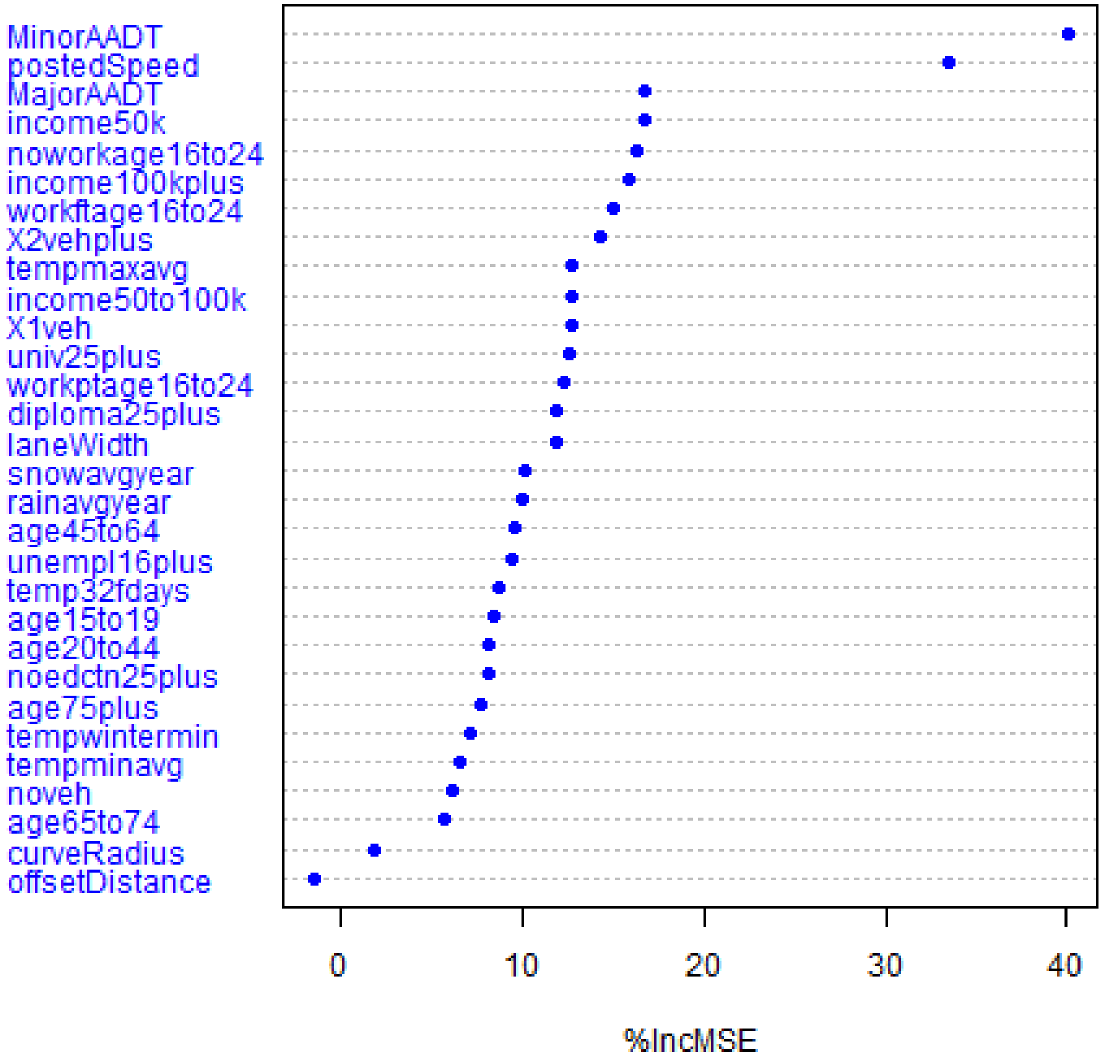
Source: FHWA.

Figure 55. Graph. ANG-KAB-D crashes on rural two-lane highway segments.



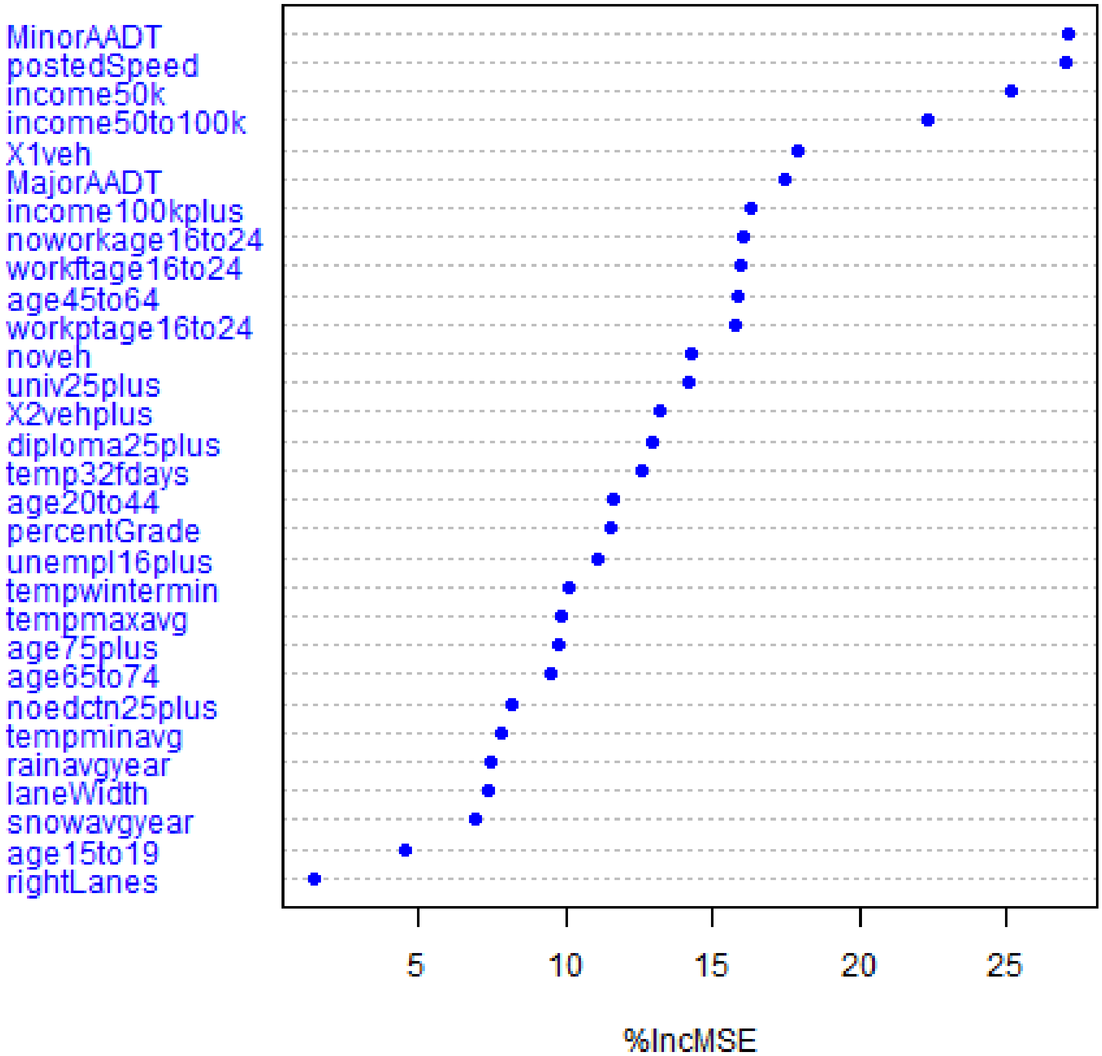
Source: FHWA.

Figure 56. Graph. ANG-KABCO-D crashes on rural two-lane highway segments.



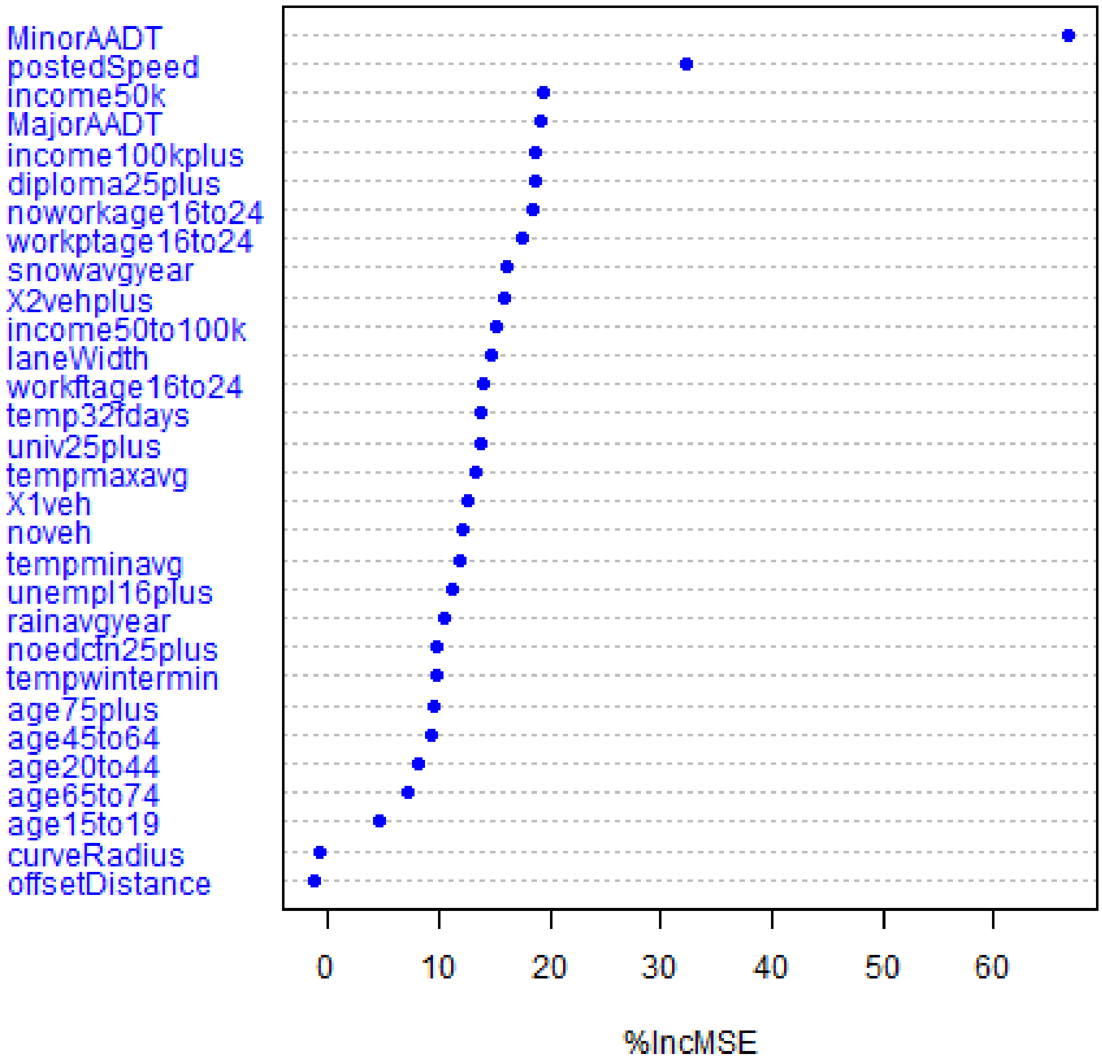
Source: FHWA.

Figure 57. Graph. ANG-KAB-D crashes at four-leg stop-controlled intersections on rural two-lane roads.



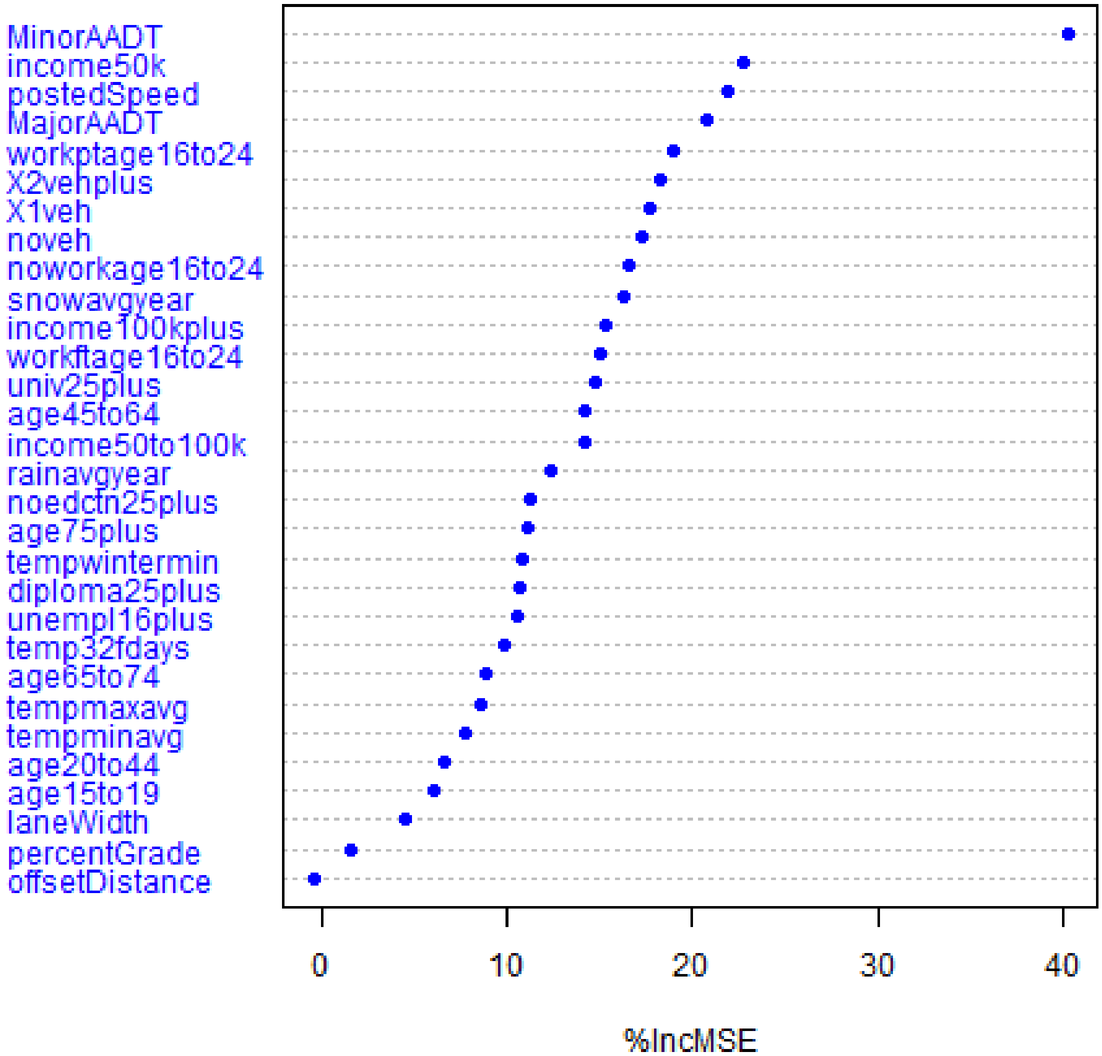
Source: FHWA.

Figure 58. Graph. ANG-KAB-N crashes at four-leg stop-controlled intersections on rural two-lane roads.



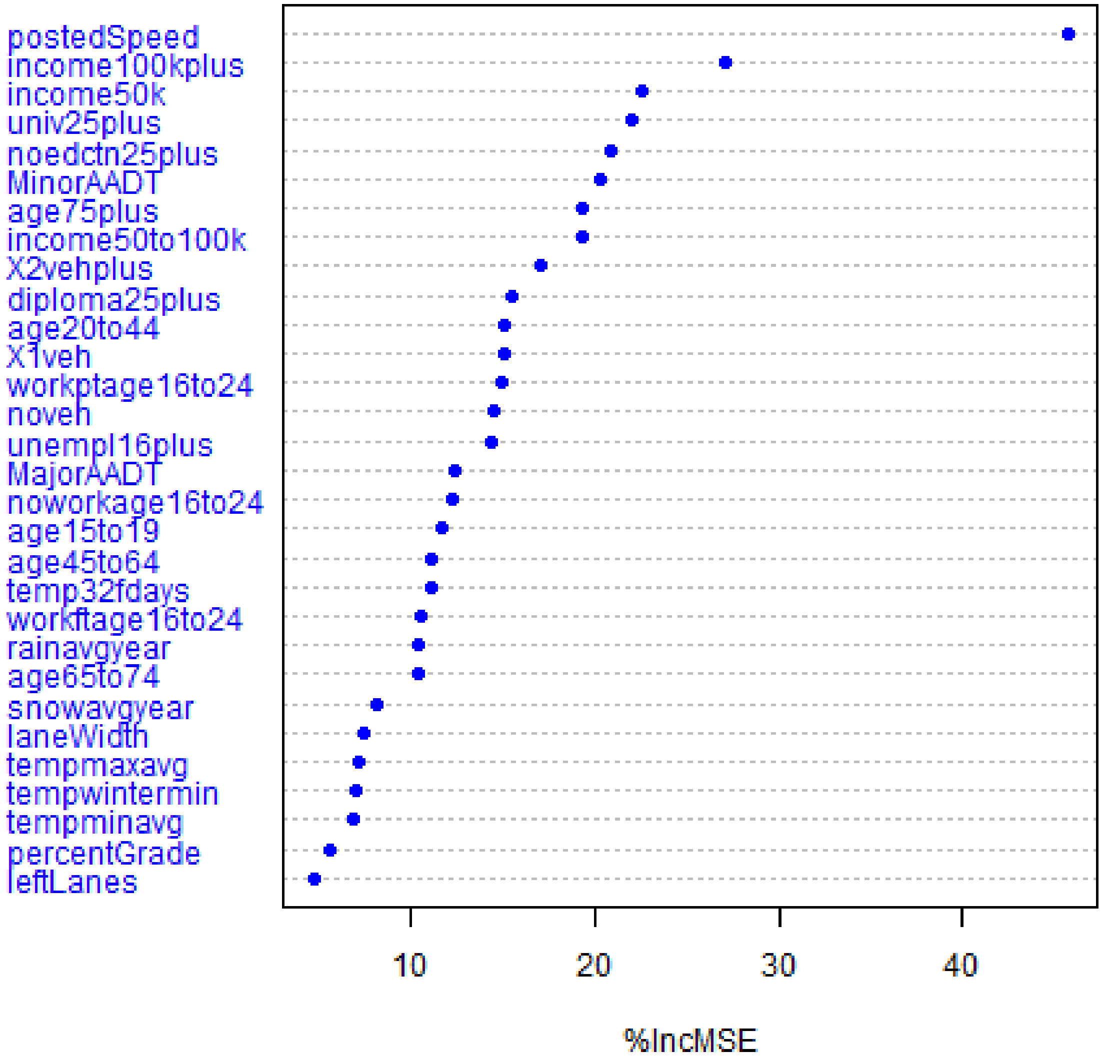
Source: FHWA.

Figure 59. Graph. ANG-KABCO-D crashes at four-leg stop-controlled intersections on rural two-lane roads.



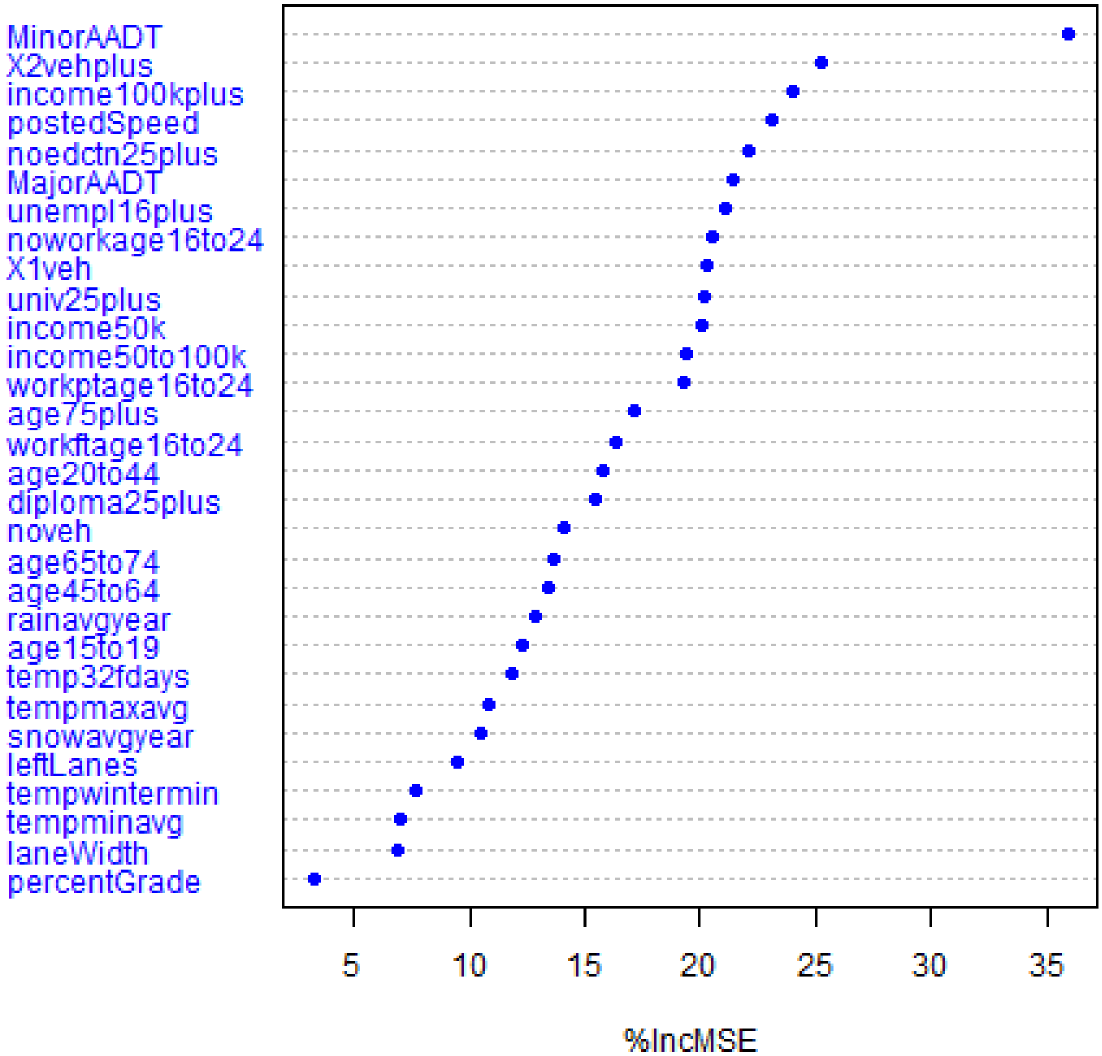
Source: FHWA.

Figure 60. Graph. ANG-KABCO-N crashes at four-leg stop-controlled intersections on rural two-lane roads.



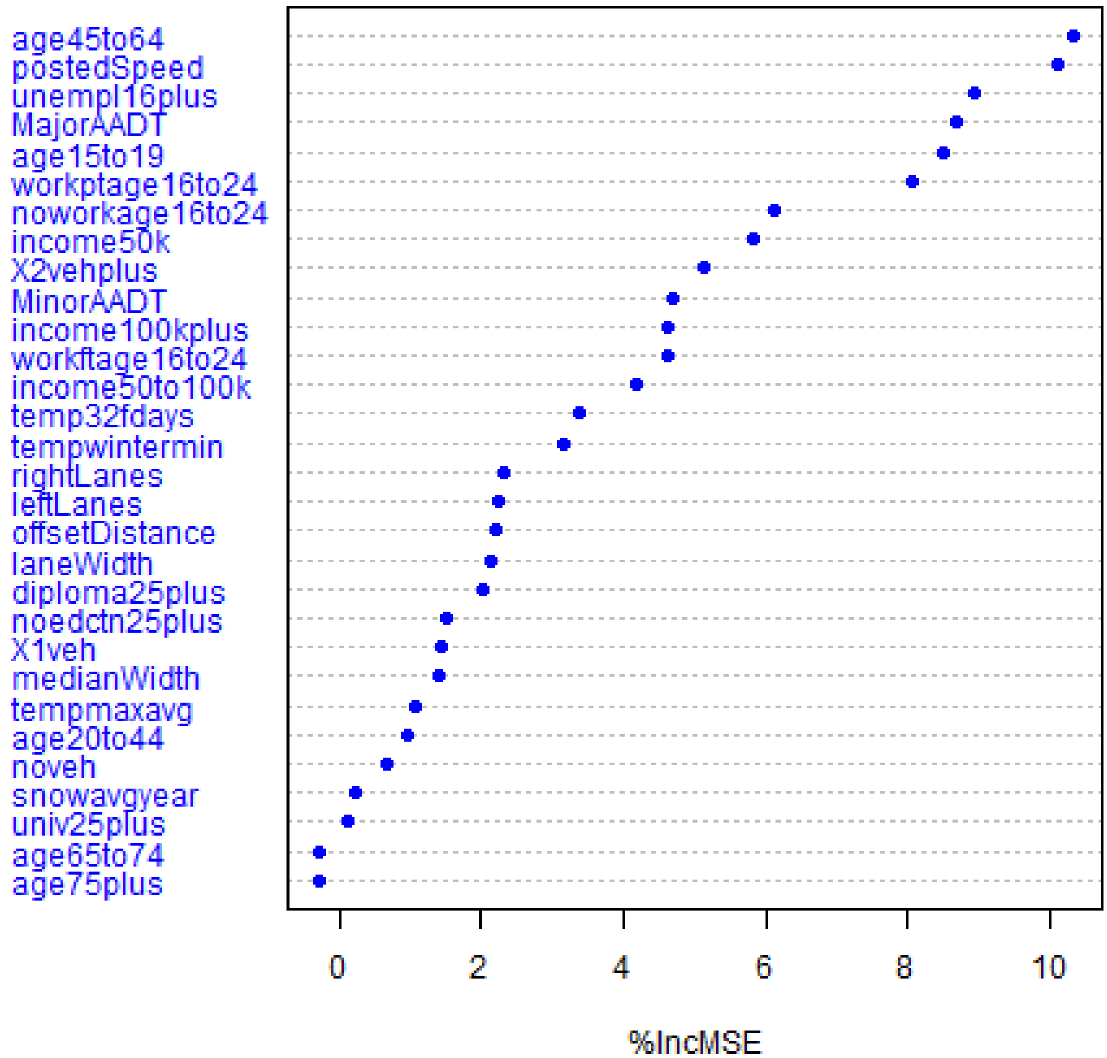
Source: FHWA.

Figure 61. Graph. ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads.



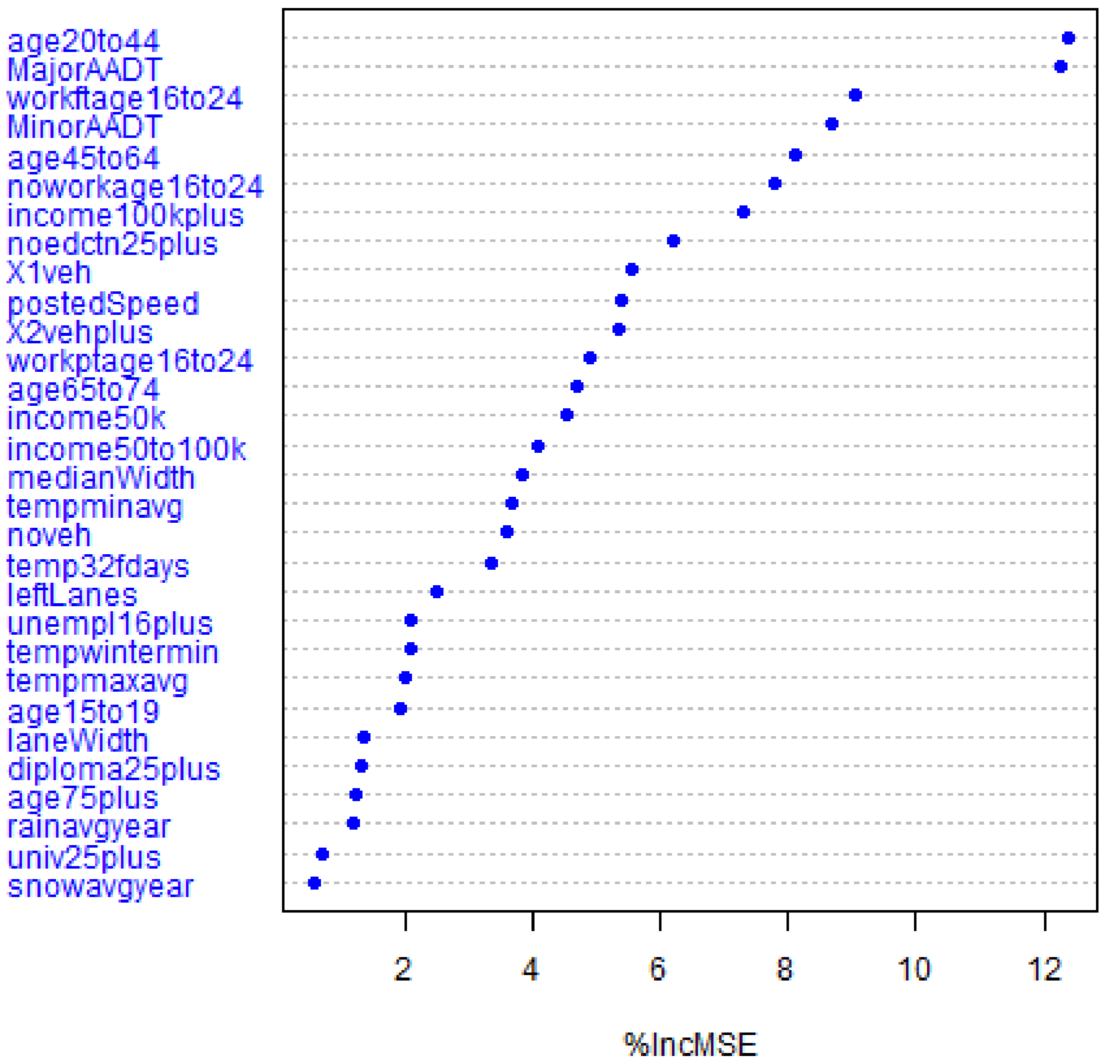
Source: FHWA.

Figure 62. Graph. ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads.



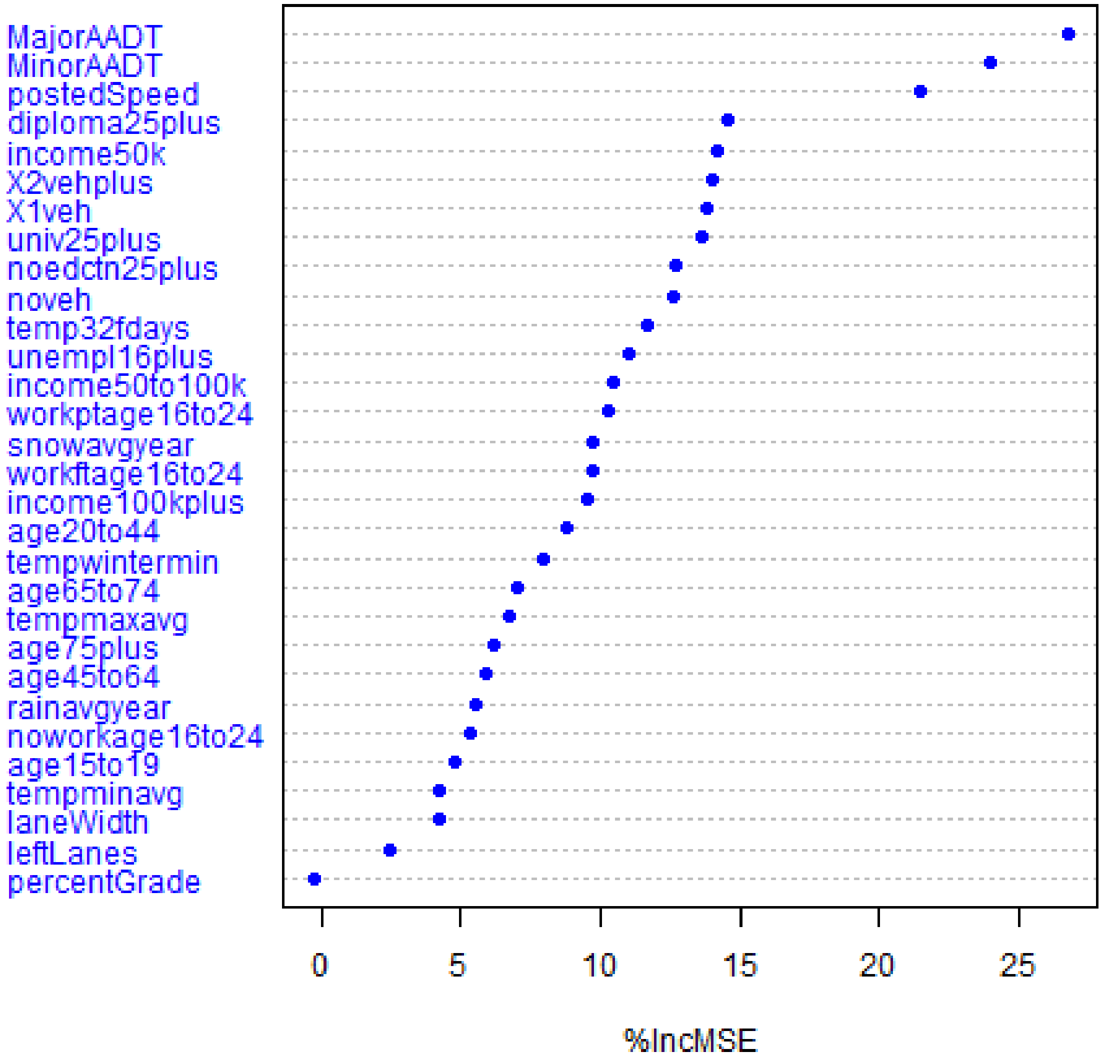
Source: FHWA.

Figure 63. Graph. ANG-KAB-D crashes at four-leg signalized intersections on urban multilane divided roads.



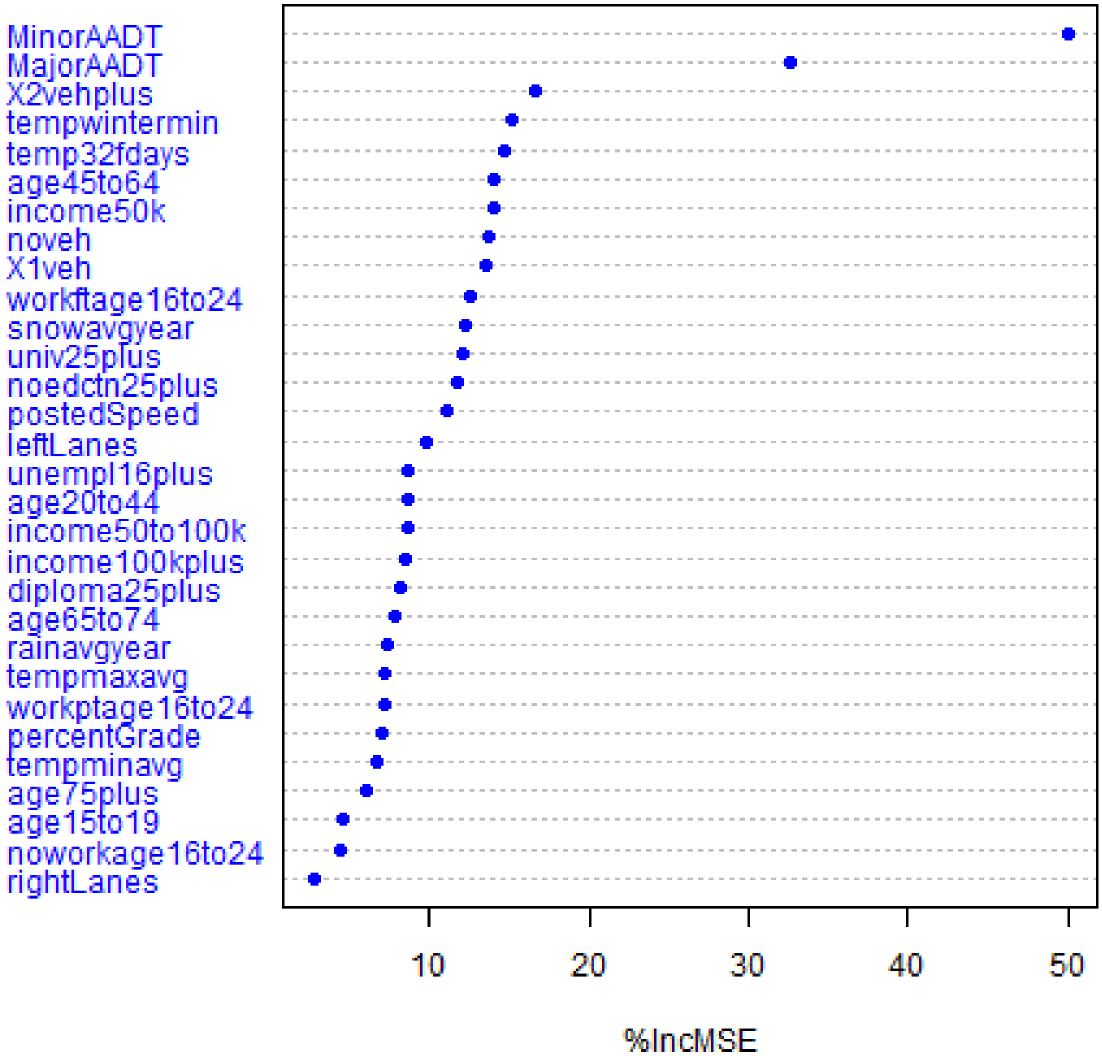
Source: FHWA.

Figure 64. Graph. ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane divided roads.



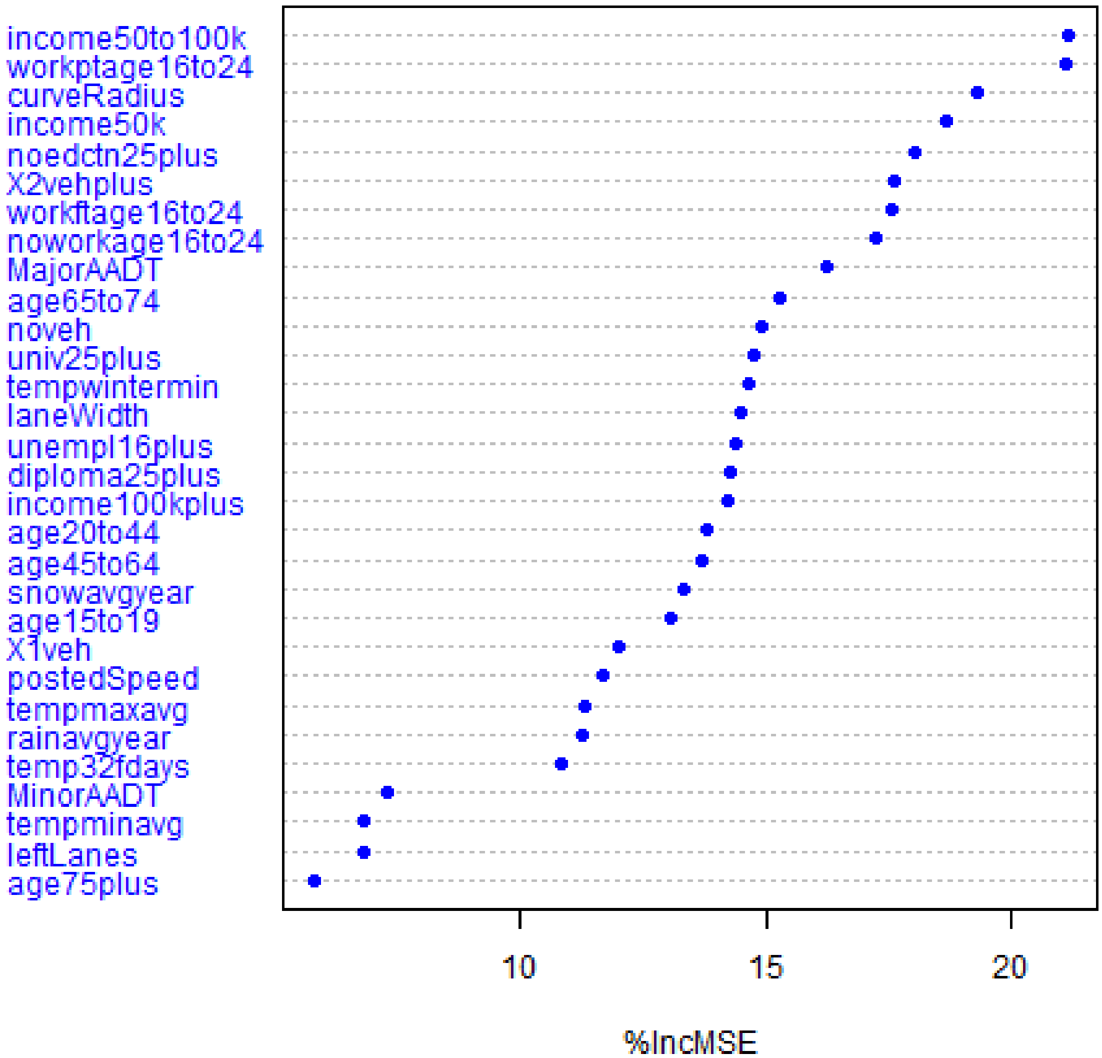
Source: FHWA.

Figure 65. Graph. ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads.



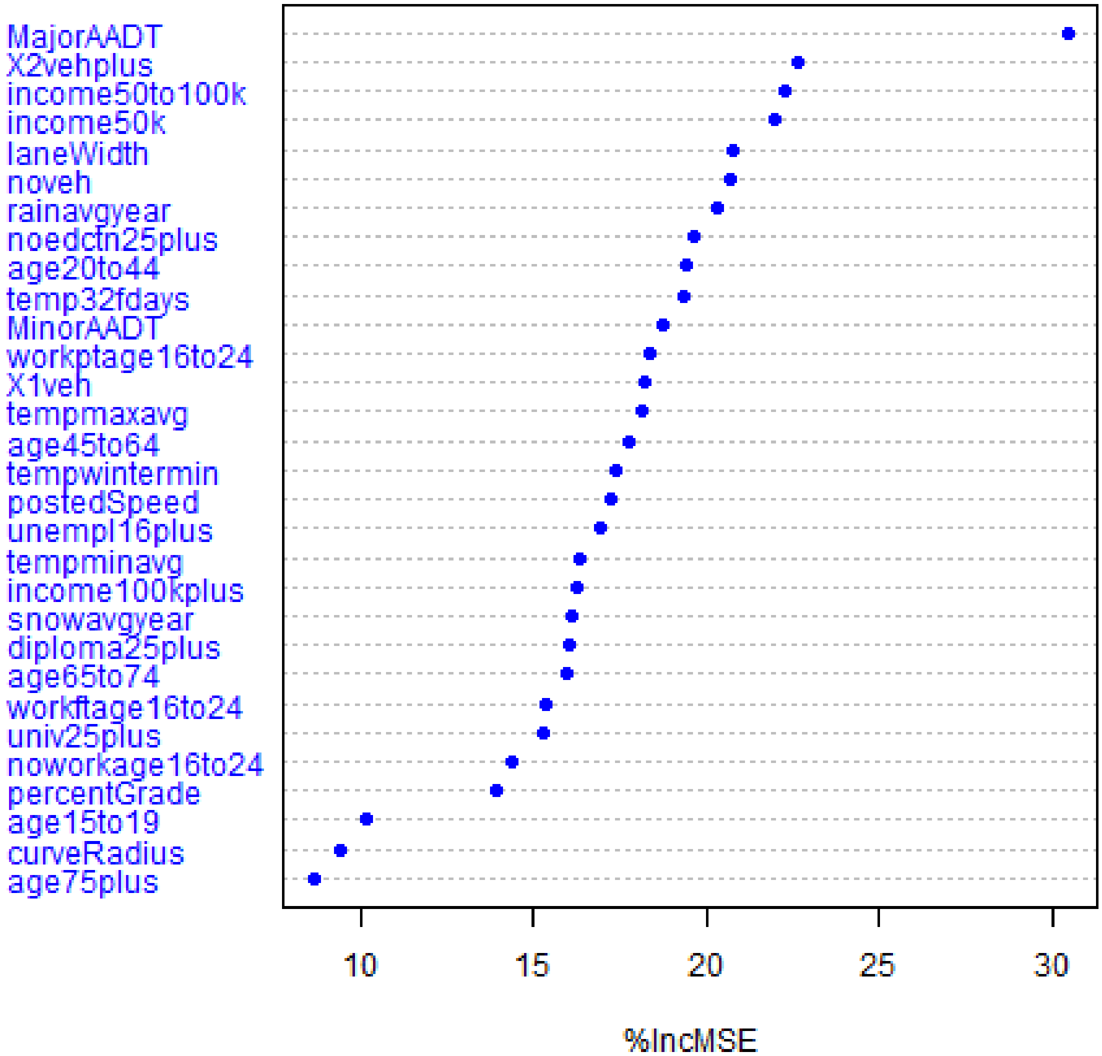
Source: FHWA.

Figure 66. Graph. ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads.



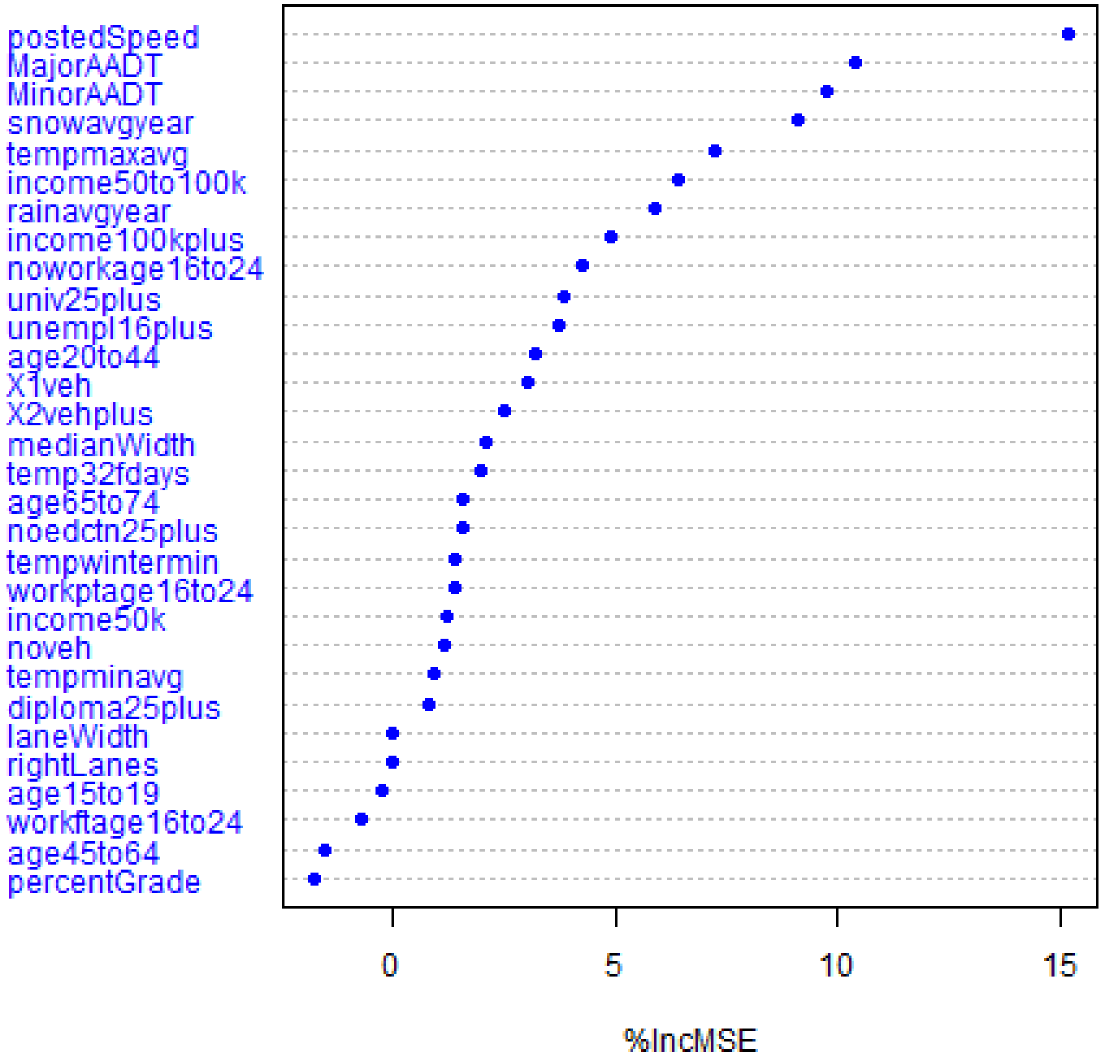
Source: FHWA.

Figure 67. Graph. ANG-KAB-D crashes at three-leg stop-controlled intersections on rural two-lane roads.



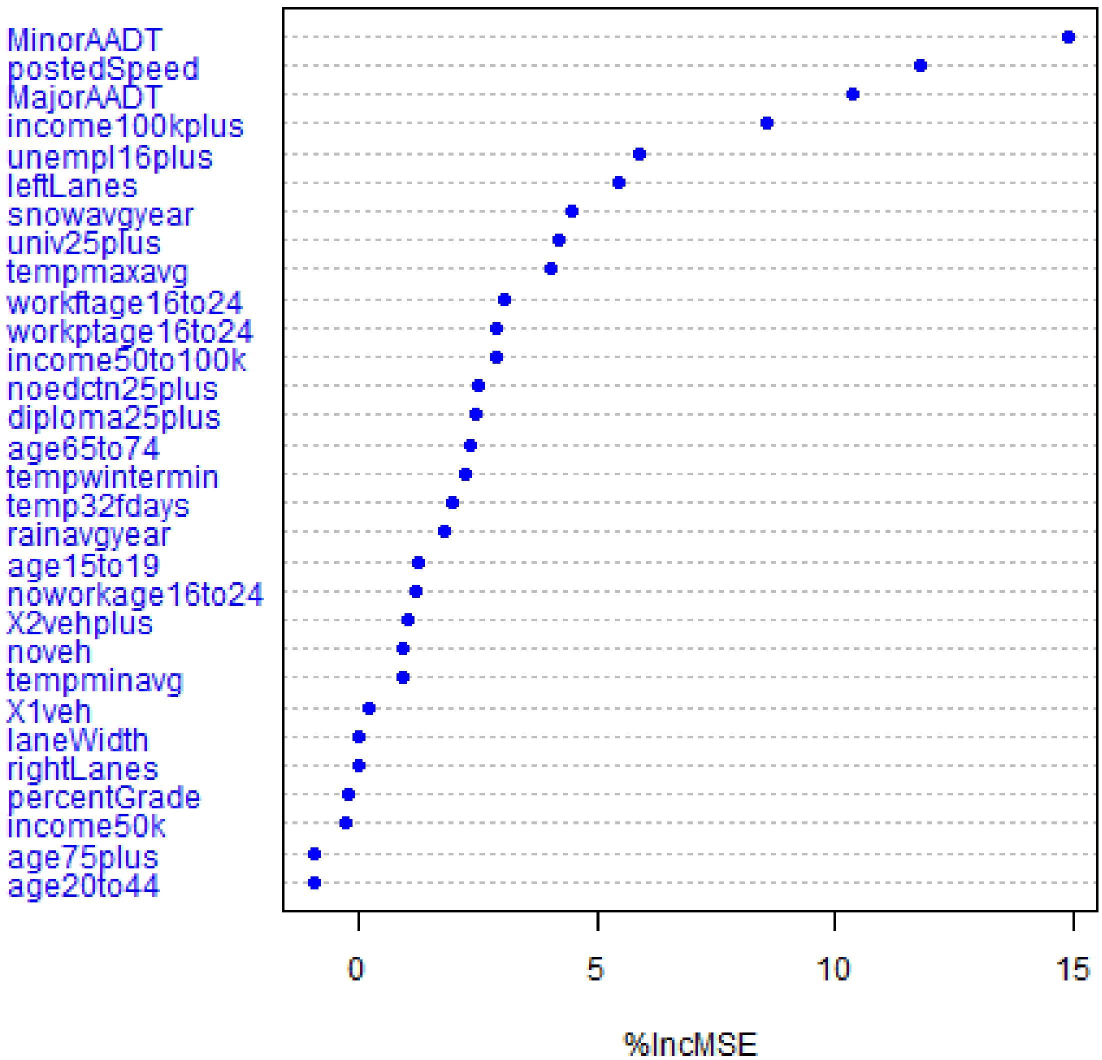
Source: FHWA.

Figure 68. Graph. ANG-KABCO-D crashes at three-leg stop-controlled intersections on rural two-lane roads.



Source: FHWA.

Figure 69. Graph. ANG-KAB-D crashes at four-leg stop-controlled intersections on rural multilane divided roads.



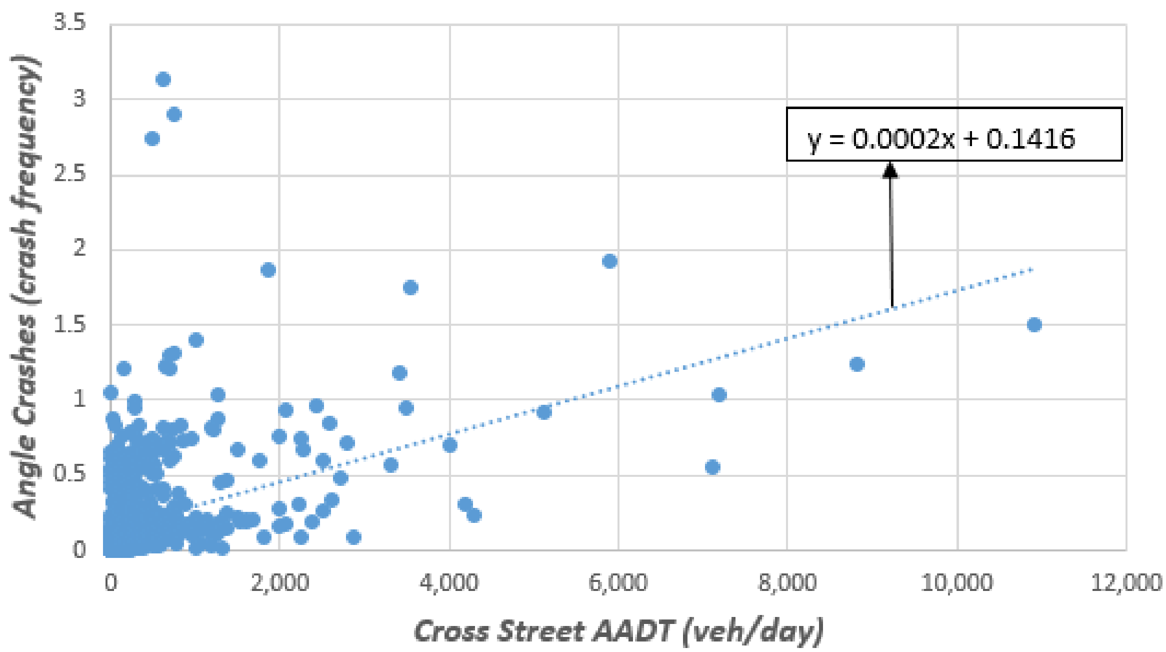
Source: FHWA.

Figure 70. Graph. ANG-KABCO-D crashes at four-leg stop-controlled intersections on rural multilane divided roads.

APPENDIX F. SAMPLE PLOTS OF RANDOM FOREST–PREDICTED CRASH FREQUENCIES VERSUS PREDICTOR VARIABLES

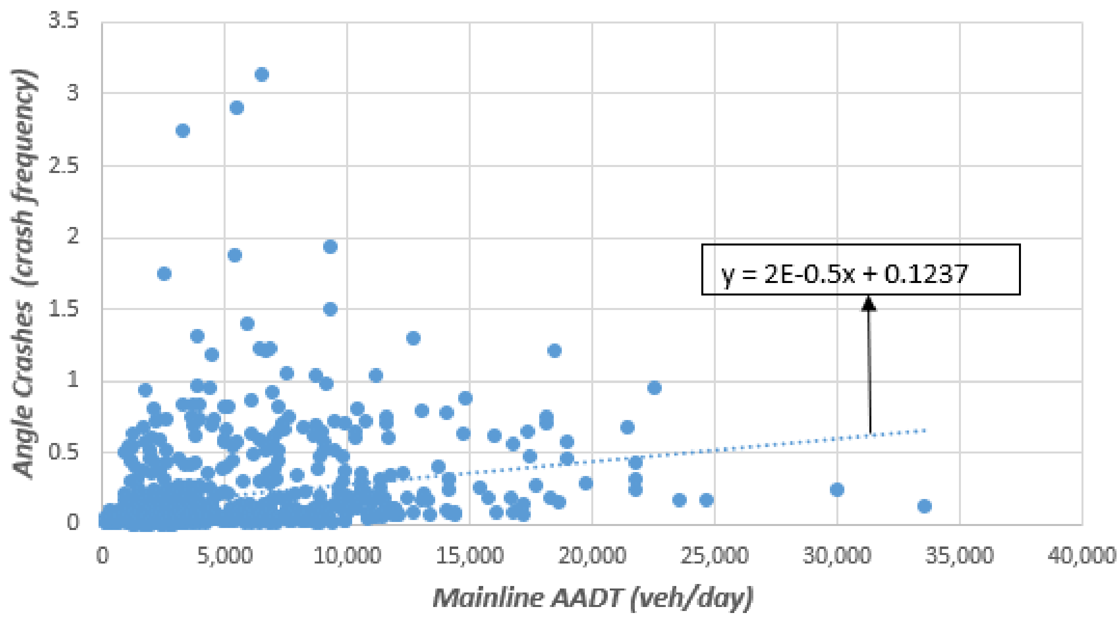
Figure 71 through figure 89 show sample plots of random forest–predicted crash frequencies versus predictor variables for ANG crashes at four-leg stop-controlled intersections on rural two-lane roads and ROR crashes at horizontal curves on rural two-lane highway segments.

ANG-D CRASHES AT FOUR-LEG STOP-CONTROLLED INTERSECTIONS ON RURAL TWO-LANE ROADS



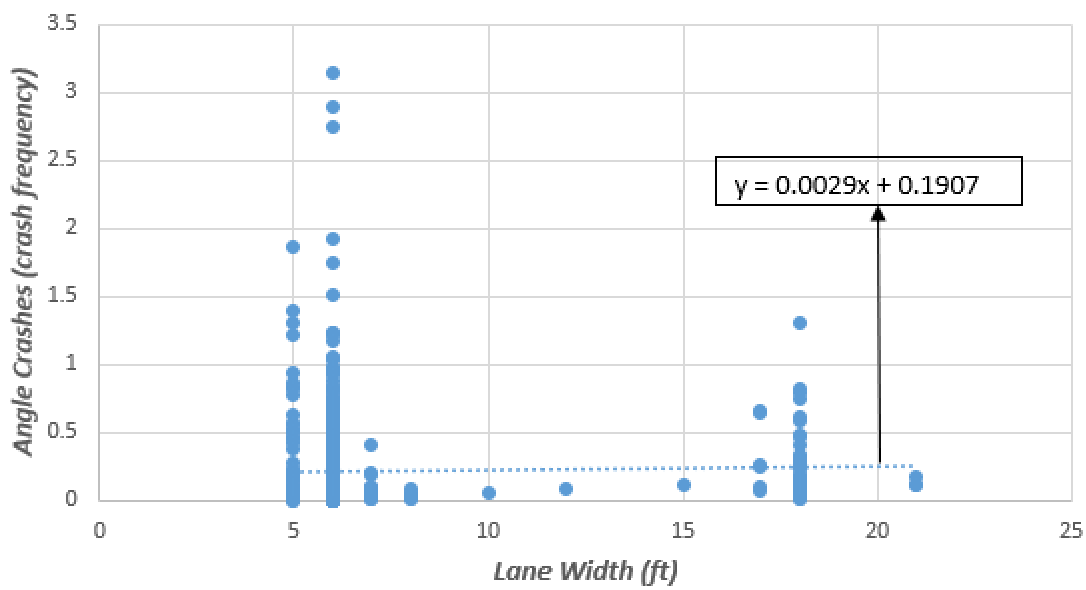
Source: FHWA.

Figure 71. Graph. Predicted ANG crash frequency versus cross street AADT.



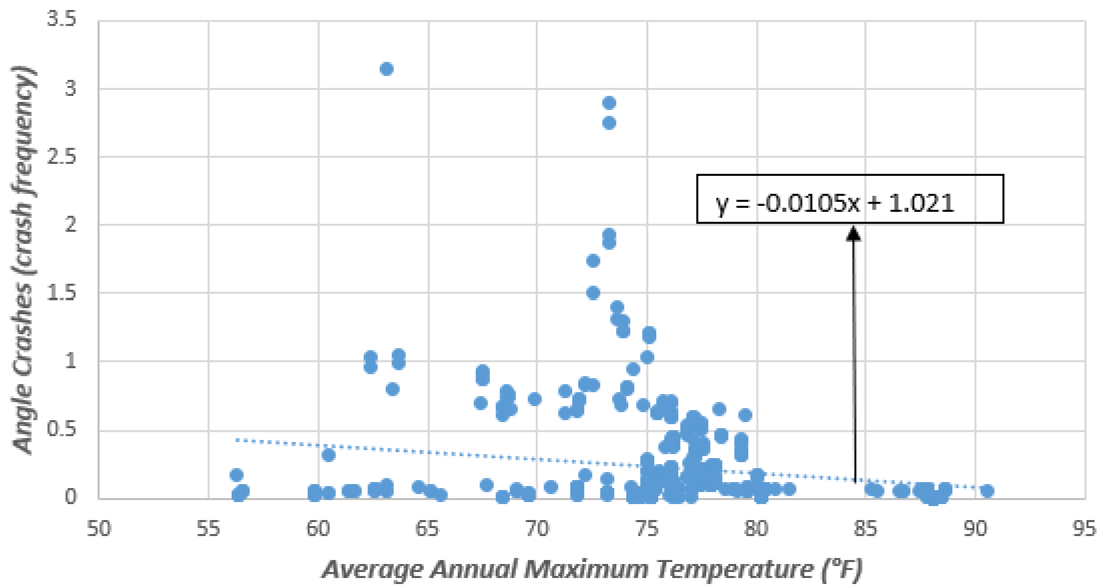
Source: FHWA.

Figure 72. Graph. Predicted ANG crash frequency versus mainline AADT.



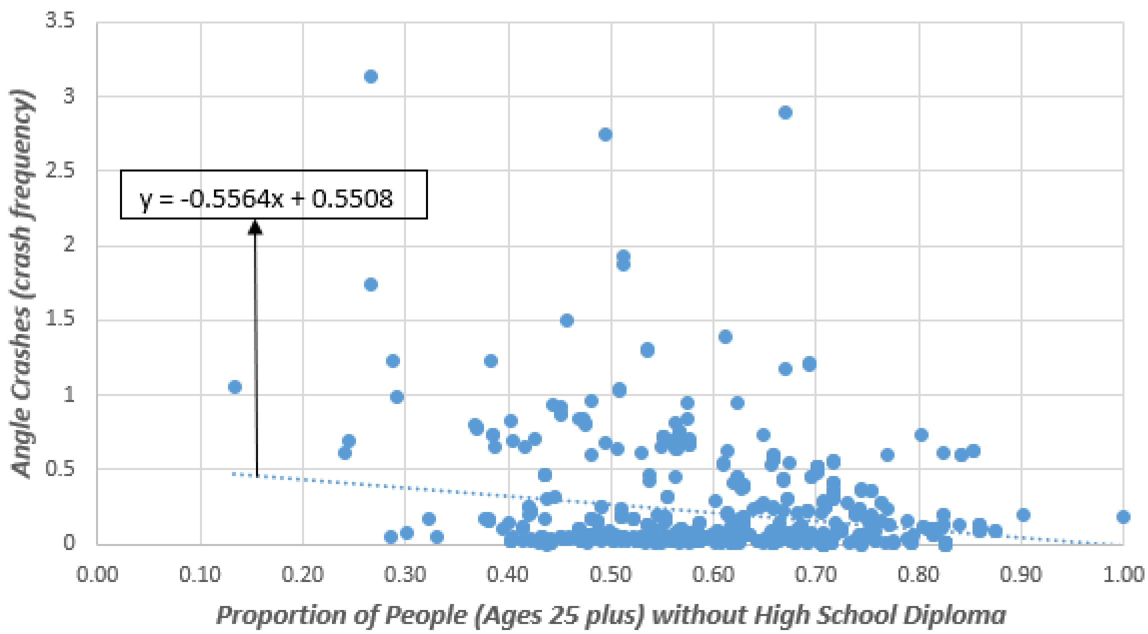
Source: FHWA.

Figure 73. Graph. Predicted ANG crash frequency versus lane width.



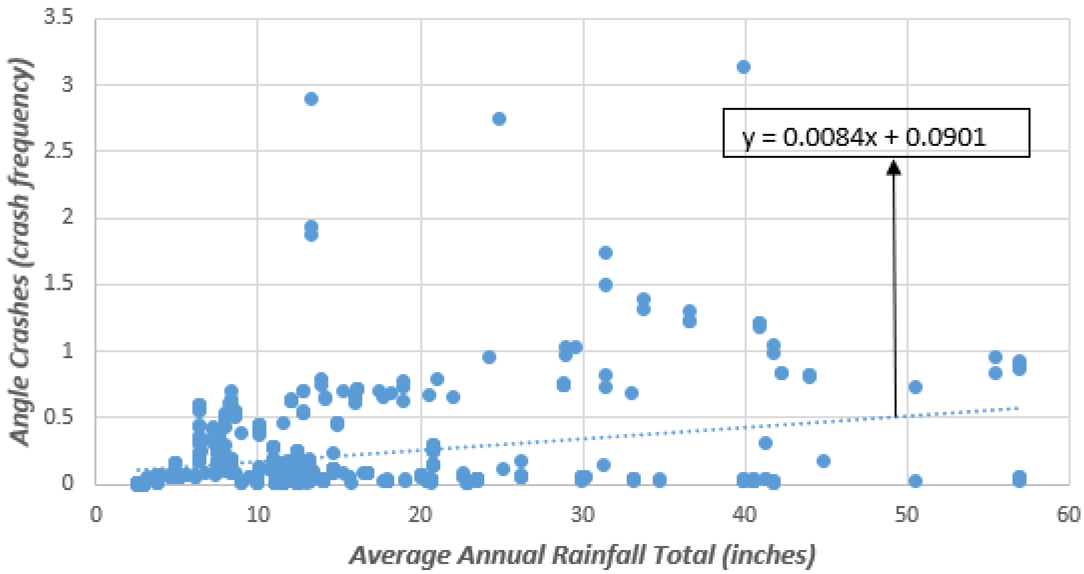
Source: FHWA.

Figure 74. Graph. Predicted ANG crash frequency versus average annual maximum temperature.



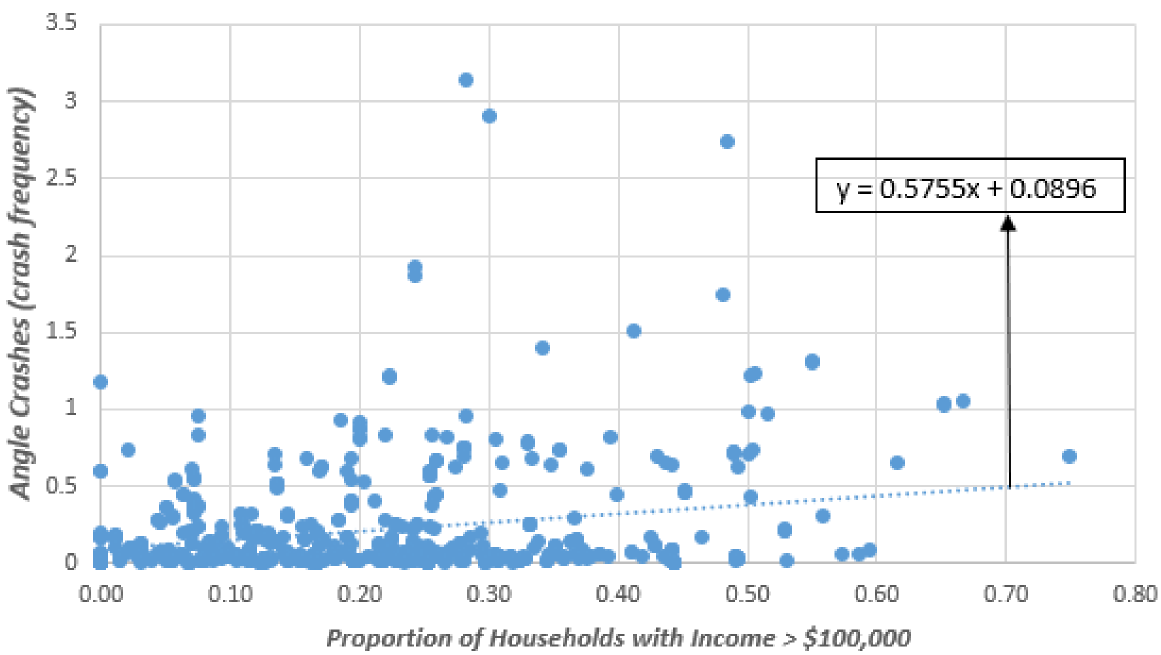
Source: FHWA.

Figure 75. Graph. Predicted ANG crash frequency versus proportion of people (ages 25+) without a high school diploma.



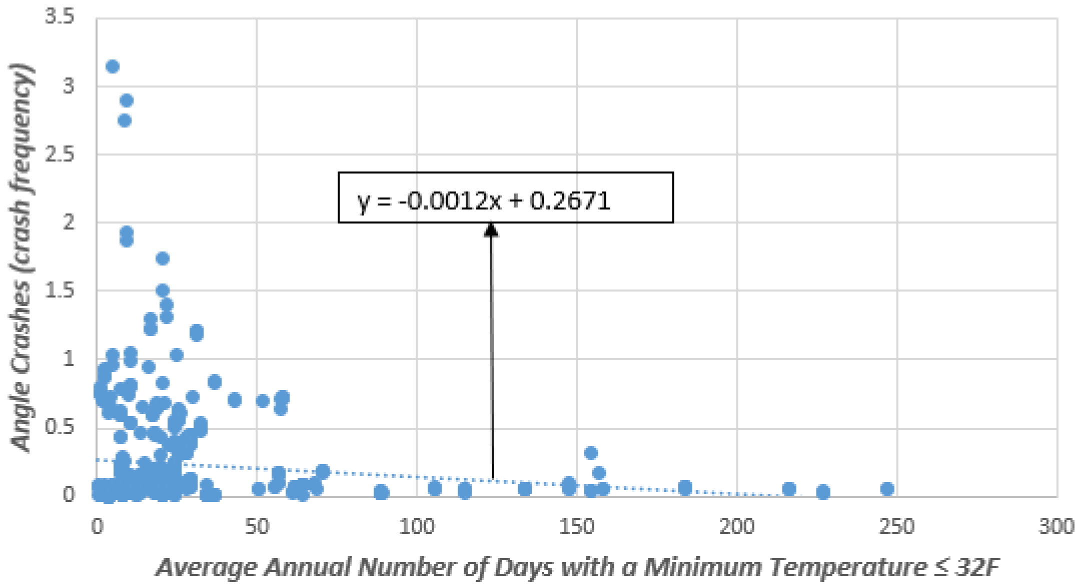
Source: FHWA.

Figure 76. Graph. Predicted ANG crash frequency versus average annual rainfall.



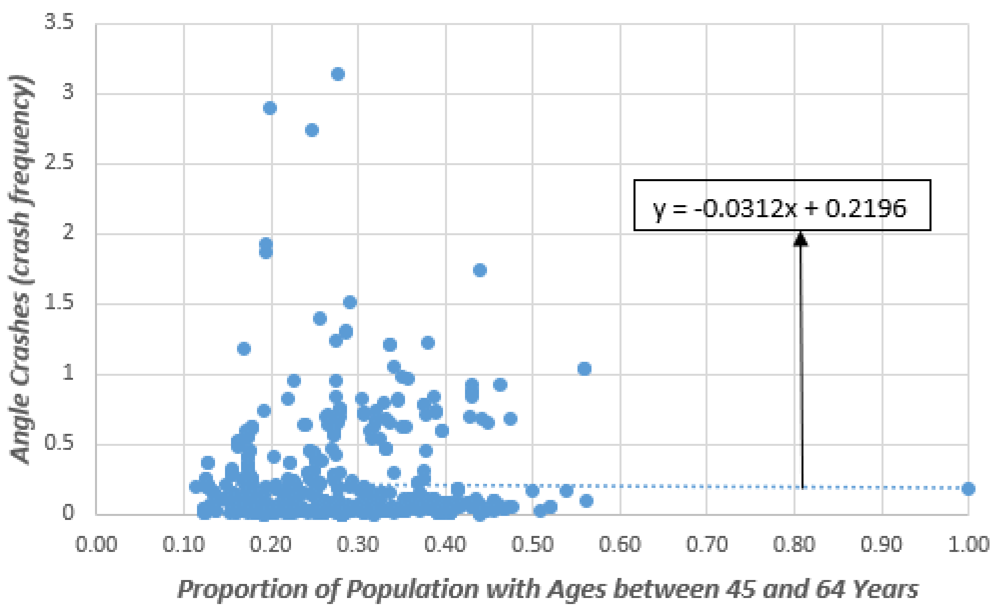
Source: FHWA.

Figure 77. Graph. Predicted ANG crash frequency versus proportion of households with income >$100,000.



Source: FHWA.

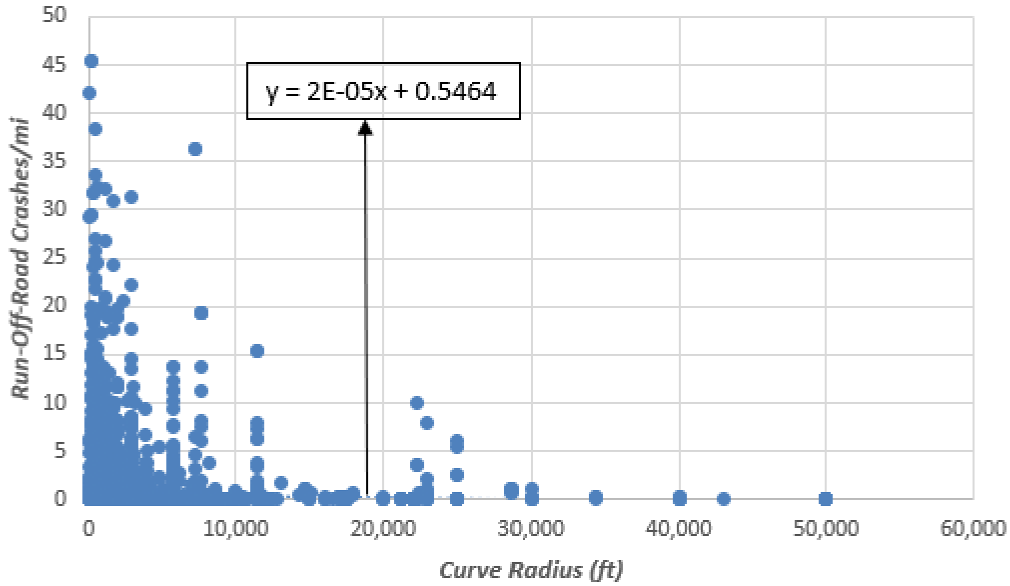
Figure 78. Graph. Predicted ANG crash frequency versus average annual number of days with a minimum temperature of ≤32°F.



Source: FHWA.

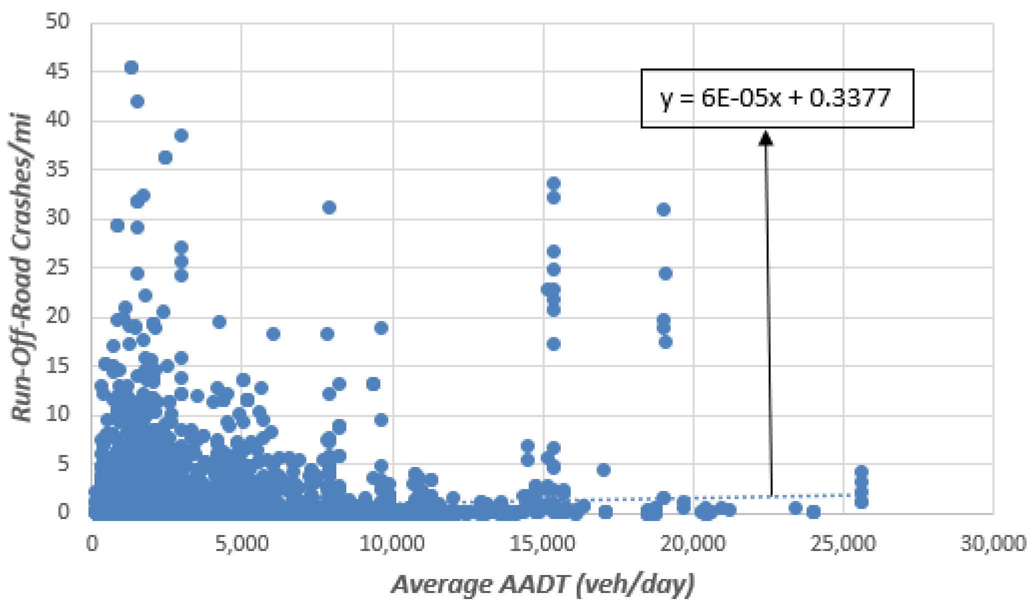
Figure 79. Graph. Predicted ANG crash frequency versus proportion of population ages 45–64.

ROR-D CRASHES AT HORIZONTAL CURVES ON RURAL TWO-LANE HIGHWAY SEGMENTS



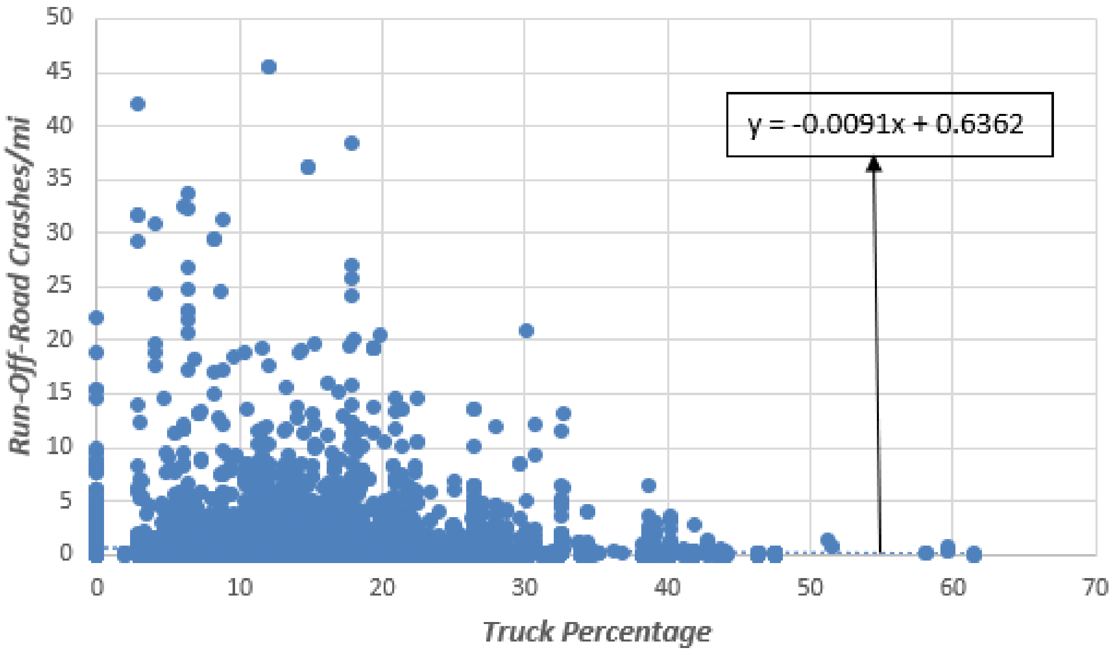
Source: FHWA.

Figure 80. Graph. Predicted ROR crash frequency per mile versus curve radius.



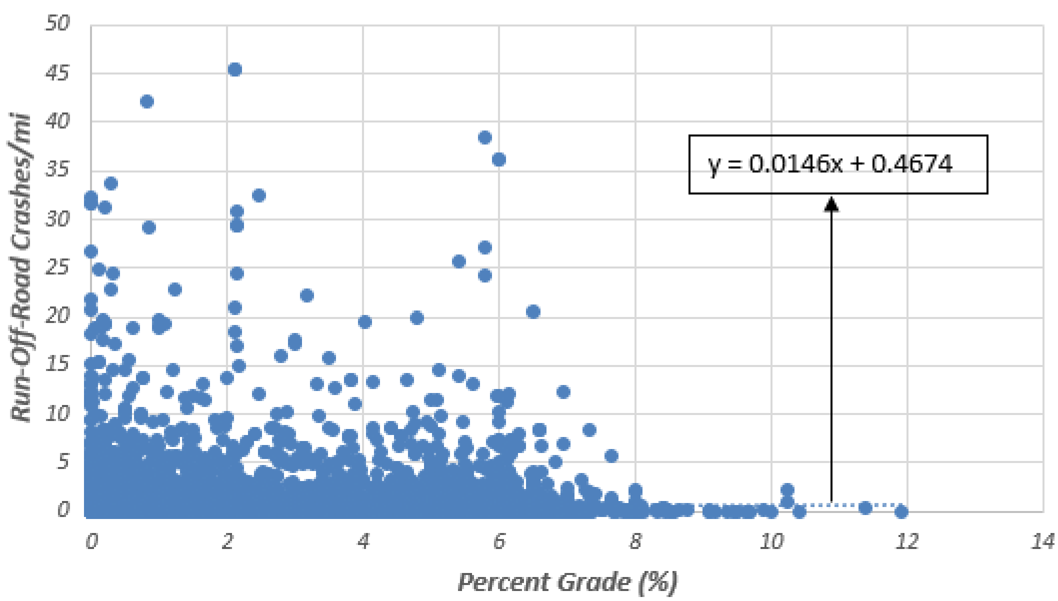
Source: FHWA.

Figure 81. Graph. Predicted ROR crash frequency per mile versus AADT.



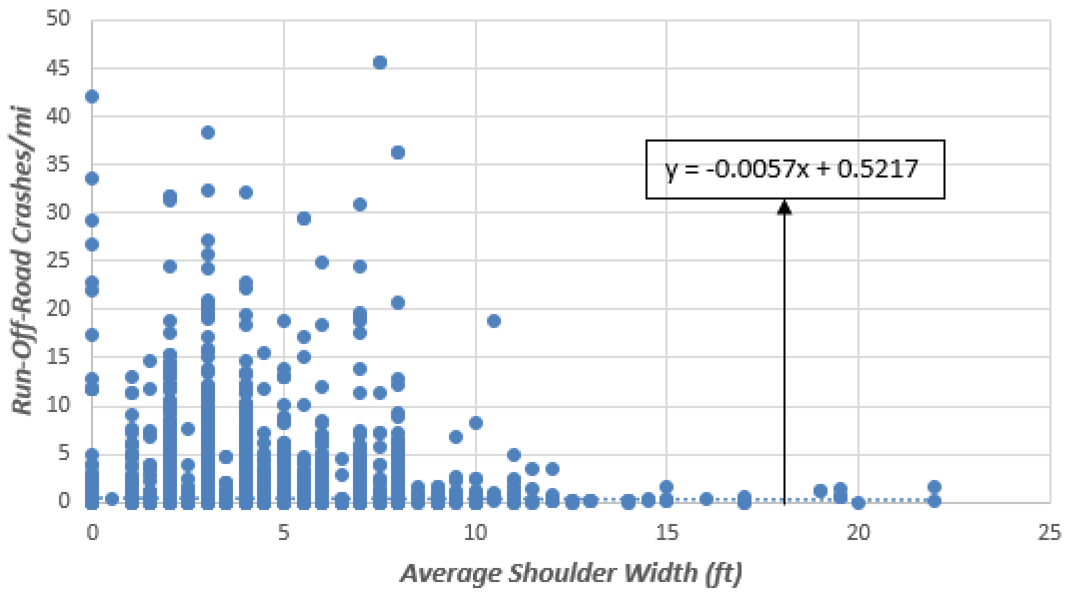
Source: FHWA.

Figure 82. Graph. Predicted ROR crash frequency per mile versus percentage of trucks on the roadway.



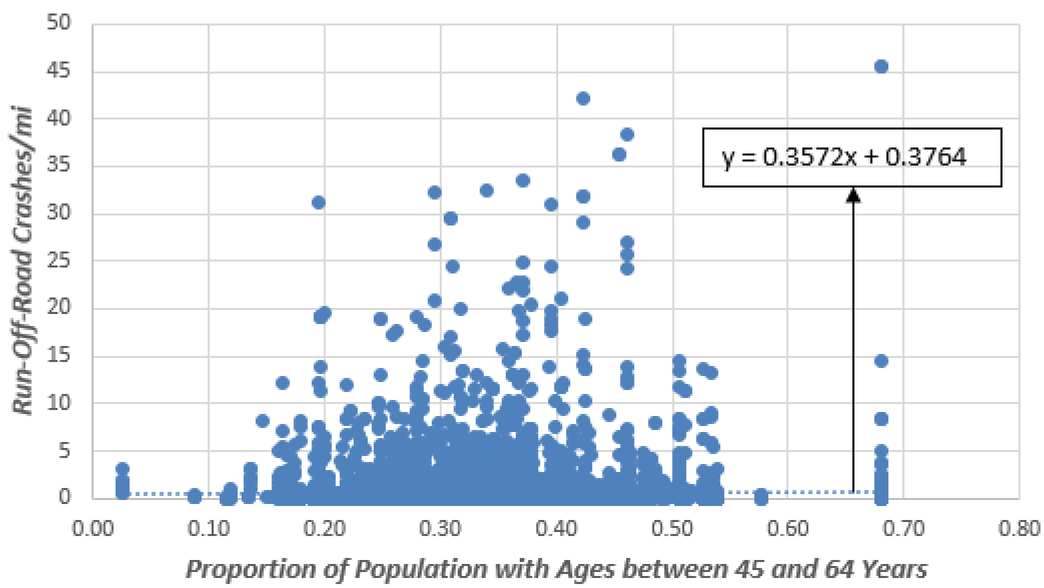
Source: FHWA.

Figure 83. Graph. Predicted ROR crash frequency per mile versus percent grade.



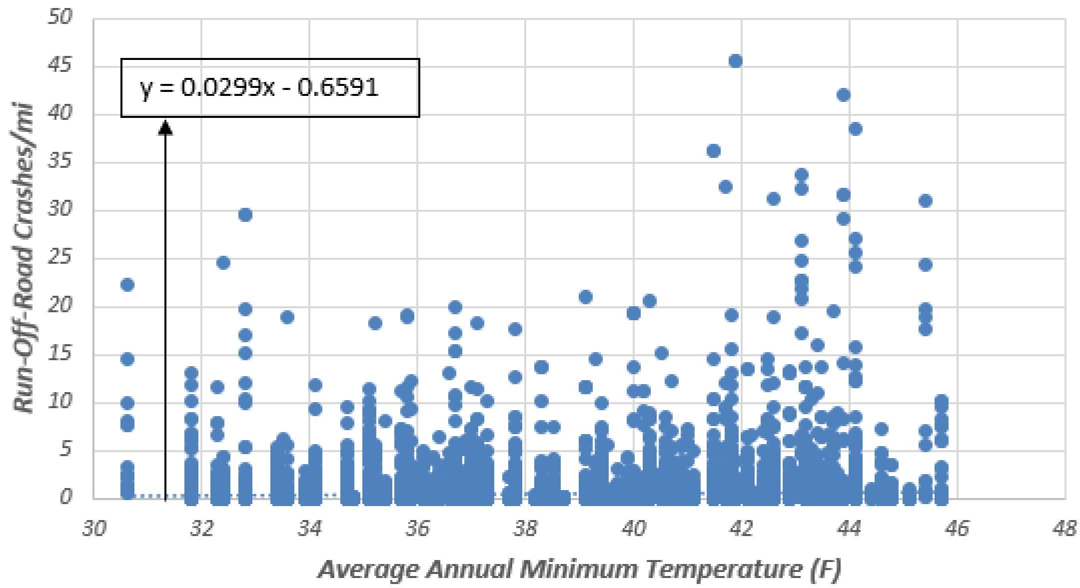
Source: FHWA.

Figure 84. Graph. Predicted ROR crash frequency per mile versus average shoulder width.



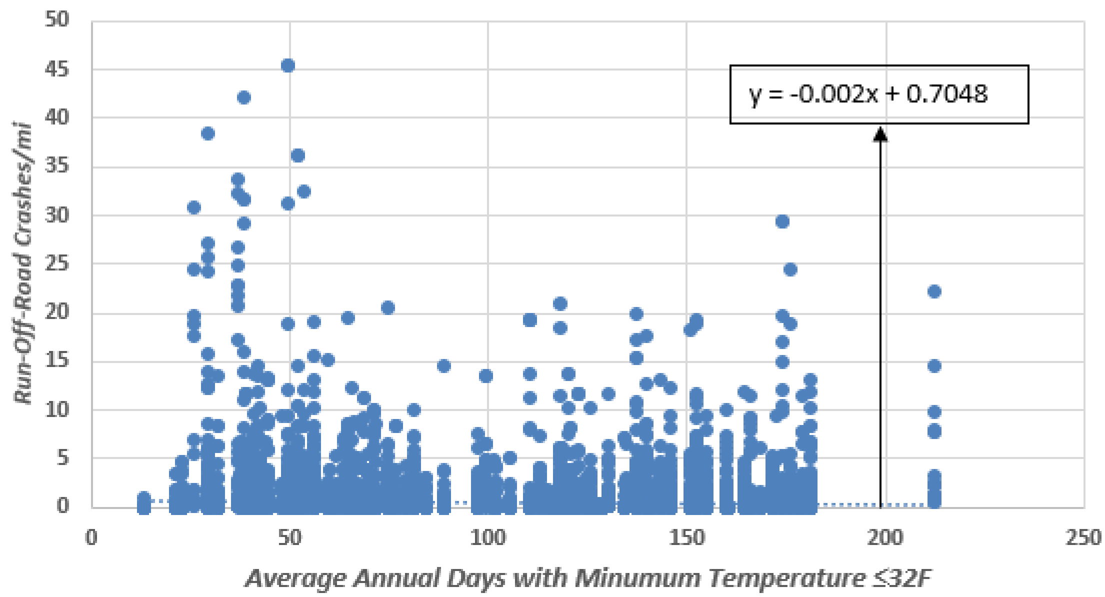
Source: FHWA.

Figure 85. Graph. Predicted ROR crash frequency per mile versus proportion of population ages 45–64.



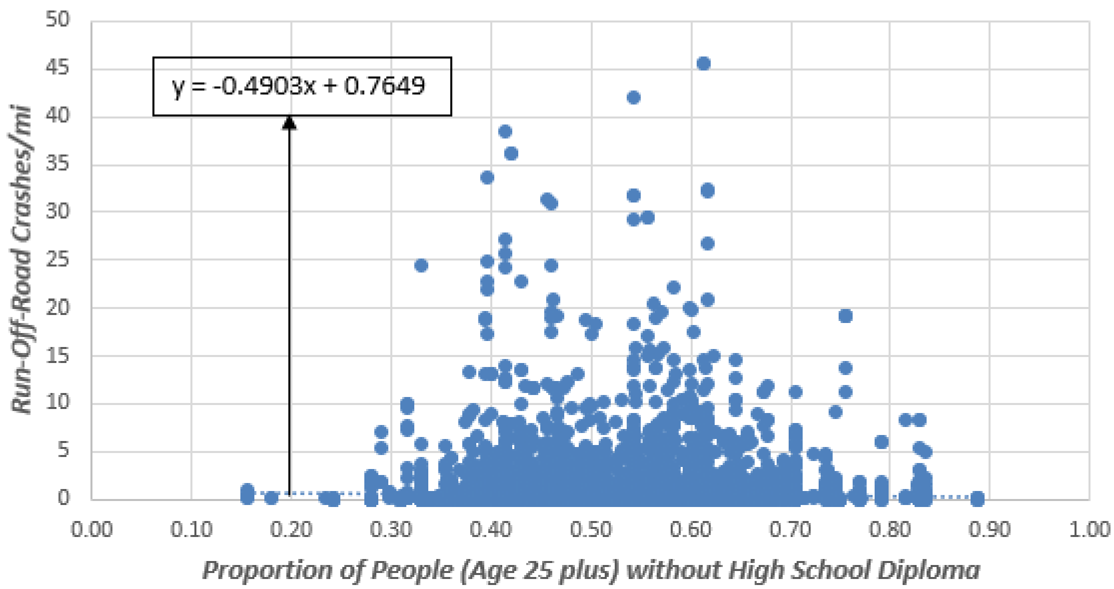
Source: FHWA.

Figure 86. Graph. Predicted ROR crash frequency per mile versus average annual minimum temperature.



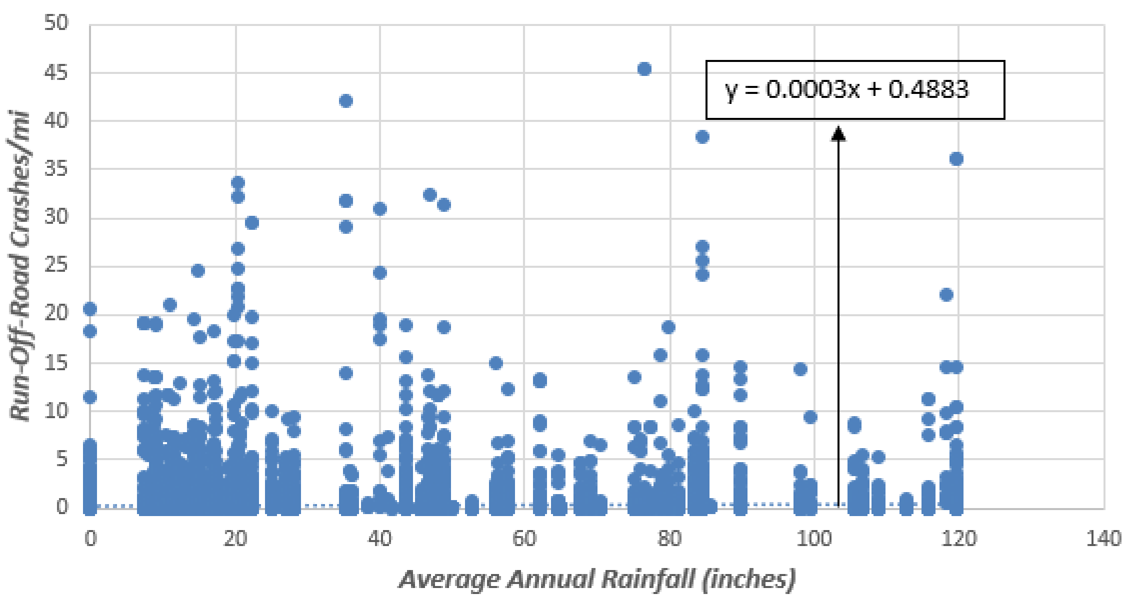
Source: FHWA.

Figure 87. Graph. Predicted ROR crash frequency per mile versus average annual number of days with a minimum temperature of ≤32°F.



Source: FHWA.

Figure 88. Graph. Predicted ROR crash frequency per mile versus proportion of people (ages 25+) without a high school diploma.



Source: FHWA.

Figure 89. Graph. Predicted ROR crash frequency per mile versus average annual rainfall.

APPENDIX G. EXPLORATION OF FACTOR ANALYSIS

The research team explored use of a factor analysis to address the correlation between socioeconomic characteristics used in the contributing-factor analysis. A study by Li et al. (2016) presented a decision tree–based model to forecast rail-transit ridership at the station level according to surrounding land-use patterns. To avoid the impact of irrelevant land-use variables, Li et al. employed a canonical correlation-analysis method to investigate the relationship between all land-use variables and the demand variables in a multivariate framework. As a result, they were able to define different factors that were a combination of different land-use variables for use in their decision tree.

The research team applied the technique used by Li et al. (2016) to derive factors from a combination of various socioeconomic variables for use in the contributing-factor analysis. For the purpose of this exploration, the research team used the data from Ohio for the following factors:

- Curves and tangent segments on rural two-lane roads.
- Four-leg stop-controlled intersections on urban two-lane roads.
- Four-leg signalized intersections on urban multilane undivided roads.

FACTOR ANALYSIS

The following sections provide results of the factor analysis for Ohio segments and intersections.

Curves and Tangent Segments on Rural Two-Lane Highway Roads

The factor analysis for curves and tangent segments on rural two-lane highway roads in Ohio resulted in three factors for the contributing-factor analysis:

- Factor 1 shows more weight on population/households with less education, low income, and less vehicle ownership.
- Factor 2 shows more weight on population/households with less education, high income, and less vehicle ownership.
- Factor 3 shows more weight on population/households with more education, high income, and more vehicle ownership.

Table 111 shows the factor-analysis results for curves and tangent segments on rural two-lane highway roads in Ohio.

Table 111. Factor analysis for curves and tangent segments.

Socioeconomic Variable	Factor 1	Factor 2	Factor 3
Percentage of population ages 16+ unemployed	0.32420	−0.20406	0.14908
Percentage of population ages 25+ without a high school diploma	0.51173	0.64328	−0.34034
Percentage of population ages 25+ with a high school diploma but no university degree	0.10912	−0.89584	−0.19936
Percentage of population ages 25+ with a university degree	−0.66791	0.26139	0.57936
Percentage of households with income <$50,000	0.87527	−0.24687	0.14978
Percentage of households with income between $50,000 and $100,000	−0.52770	0.09029	−0.53644
Percentage of households with income >$100,000	−0.71865	0.26610	0.39014
Percentage of households with 0 vehicles	0.40680	0.54877	−0.30267
Percentage of households with 1 vehicle	0.62705	−0.01826	0.50984
Percentage of households with ≥2 vehicles	−0.78055	−0.32344	−0.24351

Four-Leg Stop-Controlled Intersections on Urban Two-Lane Roads

The factor analysis for four-leg stop-controlled intersections on urban two-lane roads in Ohio resulted in two factors for the contributing-factor analysis:

- Factor 1 is associated with the prevalence of population/households with no high school education and low income.
- Factor 2 is associated with the prevalence of population/households with a high school education and medium to high income.

The number of vehicles in a household were not included in the factor analysis due to their presence leading to some noninterpretable results. Thus, the research team decided to introduce two more factors dealing with household vehicle ownership to the analysis:

- Factor 3 is associated with the percentage of households with no vehicles.
- Factor 4 is associated with the percentage of households with any number of vehicles.

Table 112 shows the factor analysis results for four-leg stop-controlled intersections on urban two-lane roads in Ohio.

Table 112. Factor analysis for four-leg stop-controlled intersections.

Socioeconomic Variable	Factor 1	Factor 2
Percentage of people ages 16+ unemployed	0.58068	−0.30712
Percentage of population ages 25+ without a high school diploma	0.71821	−0.26190
Percentage of population ages 25+ with a high school diploma but no university degree	0.48666	0.83686
Percentage of population ages 25+ with a university degree	−0.89247	−0.33733
Percentage of households with income <$50,000	0.91095	−0.23481
Percentage of households with income between $50,000 and $100,000	−0.41974	0.72009
Percentage of households with income >$100,000	−0.88462	−0.25201

Four-Leg Signalized Intersections on Urban Multilane Undivided Roads

The factor analysis for four-leg signalized intersections on urban multilane undivided roads in Ohio resulted in two factors for the contributing-factor analysis:

- Factor 1 is associated with the prevalence of population/households with no high school education and low income.
- Factor 2 is associated with the prevalence of population/households with a high school education and medium to high income.

The number of vehicles in a household were not included in the factor analysis due to their presence leading to some noninterpretable results.

Table 113 shows the factor analysis results for four-leg signalized intersections on urban multilane undivided roads in Ohio.

Table 113. Factor analysis for four-leg signalized intersections.

Socioeconomic Variable	Factor 1	Factor 2
Percentage of people ages 16+ unemployed	0.65462	−0.35922
Percentage of population ages 25+ without a high school diploma	0.79210	−0.10944
Percentage of population ages 25+ with a high school diploma but no university degree	0.54037	0.65992
Percentage of population ages 25+ with a university degree	−0.90269	−0.34545
Percentage of households with income <$50,000	0.92500	−0.21541
Percentage of households with income between $50,000 and $100,000	−0.50650	0.69851
Percentage of households with income >$100,000	−0.86864	−0.23768

RESULTS FROM THE CONTRIBUTING-FACTOR ANALYSIS

The following sections provide results of the updated contributing-factor analysis for the intersection and nonintersection FCFTs based on KAB and KABCO crashes using Ohio data.

ROR Crashes on Rural Two-Lane Highway Roads in Ohio

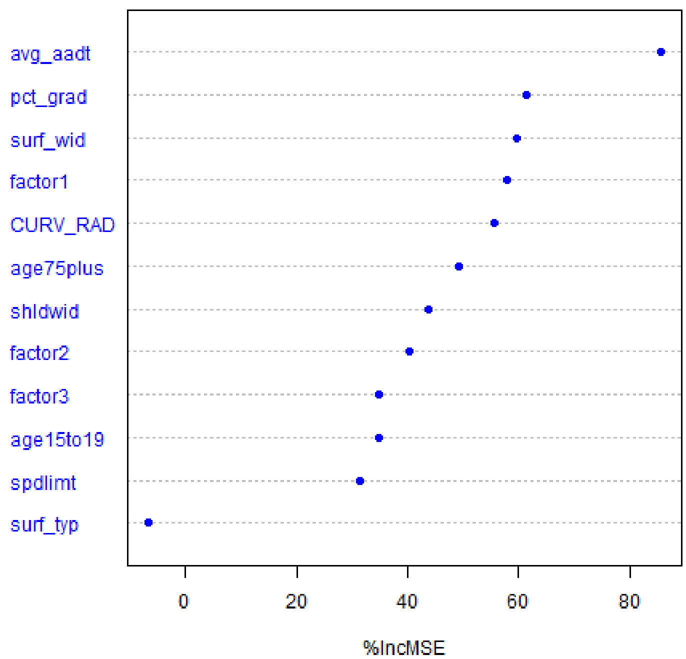
Horizontal Curves and Highway Tangent Segments—Daytime

Table 114 and table 115 summarize the most influential predictor variables for the expected number of ROR-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROR-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 90 and figure 91.

Table 114. Contributing factors for ROR-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Surface width	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Curve radius	Increases/decreases*
Percentage of population ages 75+	Decreases
Average shoulder width	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 3 (more education, high income, more vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



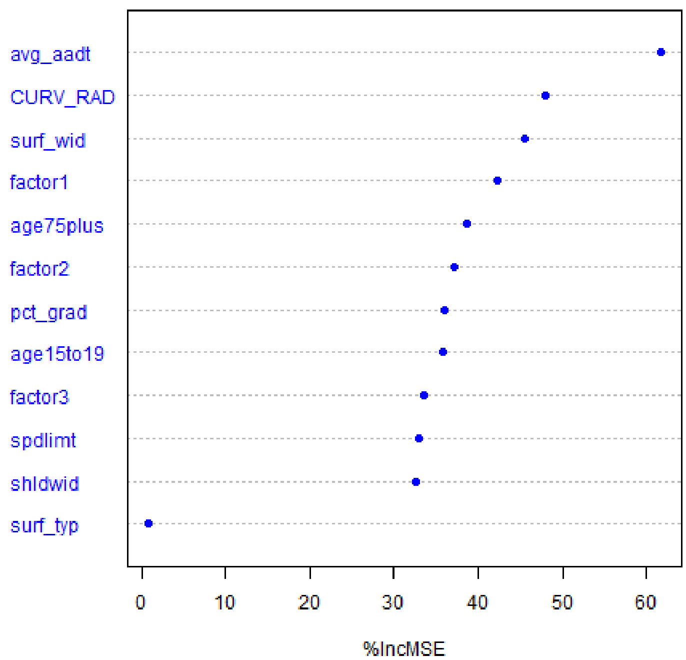
Source: FHWA.

Figure 90. Graph. Output for ROR-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 115. Contributing factors for ROR-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Surface width	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percentage of population ages 75+	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Percent grade	Increases
Percentage of population ages 15–19	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Speed limit	Increases
Average shoulder width	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



Source: FHWA.

Figure 91. Graph. Output for ROR-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

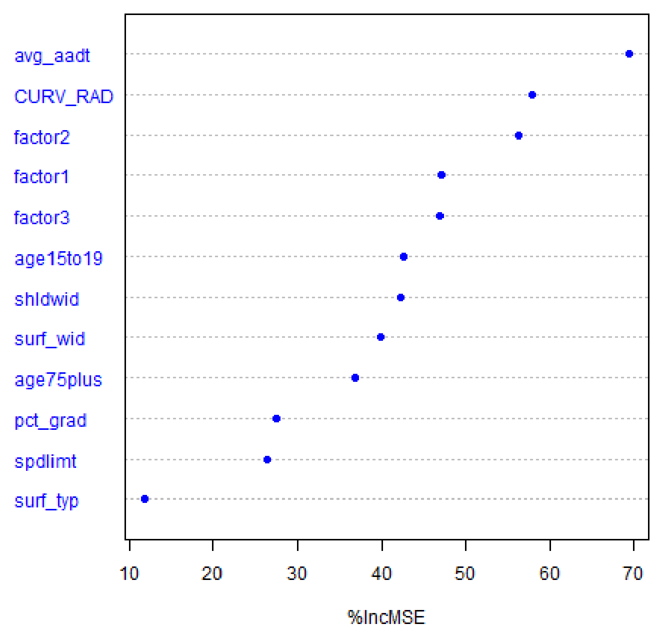
Horizontal Curves and Highway Tangent Segments—Nighttime

Table 116 and table 117 summarize the most influential predictor variables for the expected number of ROR-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROR-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROR-KABCO-N) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 92 and figure 93.

Table 116. Contributing factors for ROR-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Average shoulder width	Decreases
Surface width	Decreases
Percentage of population ages 75+	Decreases
Percent grade	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



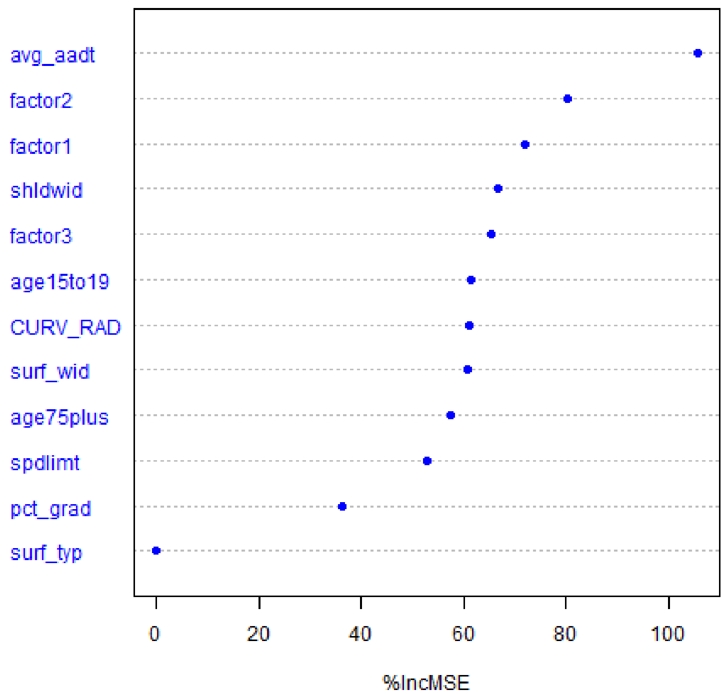
Source: FHWA.

Figure 92. Graph. Output for ROR-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 117. Contributing factors for ROR-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Average shoulder width	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Curve radius	Increases/decreases*
Surface width	Decreases
Percentage of population ages 75+	Decreases
Speed limit	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



Source: FHWA.

Figure 93. Graph. Output for ROR-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ROR crashes on rural two-lane horizontal curves and highway tangent segments in Ohio:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of ROR-D and ROR-N crashes.
- *Percent grade*: an increase in percent grade was consistently associated with an increase in the frequency of ROR-D crashes.
- *Average shoulder width*: an increase in shoulder width was consistently associated with a decrease in the frequency of ROR-D and ROR-N crashes (and therefore a decrease in shoulder width was consistently associated with an increase in the frequency of ROR crashes).
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of ROR-D and ROR-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of ROR crashes).
- *Surface width*: an increase in surface width was associated with a decrease in the frequency of ROR-D and ROR-N crashes on curves and highway tangent segments (and therefore a decrease in surface width was consistently associated with an increase in the frequency of LNDP crashes).
- *Speed*: an increase in speed limit was consistently associated with an increase in the frequency of ROR-D and ROR-N crashes.

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 consistently appeared as a factor that increases predicted daytime and nighttime crash frequency.
- The percentage of population ages 75+ consistently appeared as a factor that decreases daytime and nighttime predicted crash frequency.
- Factor 1 (less education, low income, and less vehicle ownership) consistently appeared as a factor that decreases daytime and nighttime predicted crash frequency.
- Factor 2 (less education, high income, and less vehicle ownership) consistently appeared as a factor that increase daytime and nighttime predicted crash frequency.
- Factor 3 (more education, high income, and more vehicle ownership) consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.

LNDP Crashes on Rural Two-Lane Highway Segments in Ohio

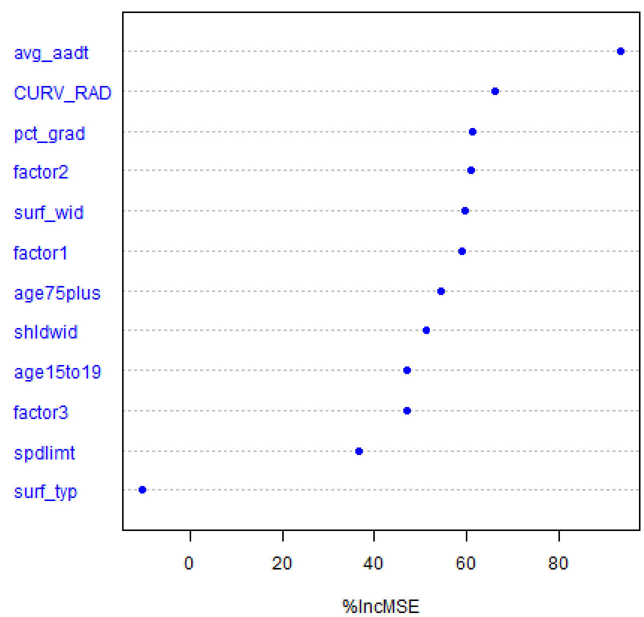
Horizontal Curves and Highway Tangent Segments—Daytime

Table 118 and table 119 summarize the most influential predictor variables for the expected number of LNDP-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (LNDP-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 94 and figure 95.

Table 118. Contributing factors for LNDP-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Percent grade	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Surface width	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percentage of population ages 75+	Decreases
Average shoulder width	Decreases
Percentage of population ages 15–19	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



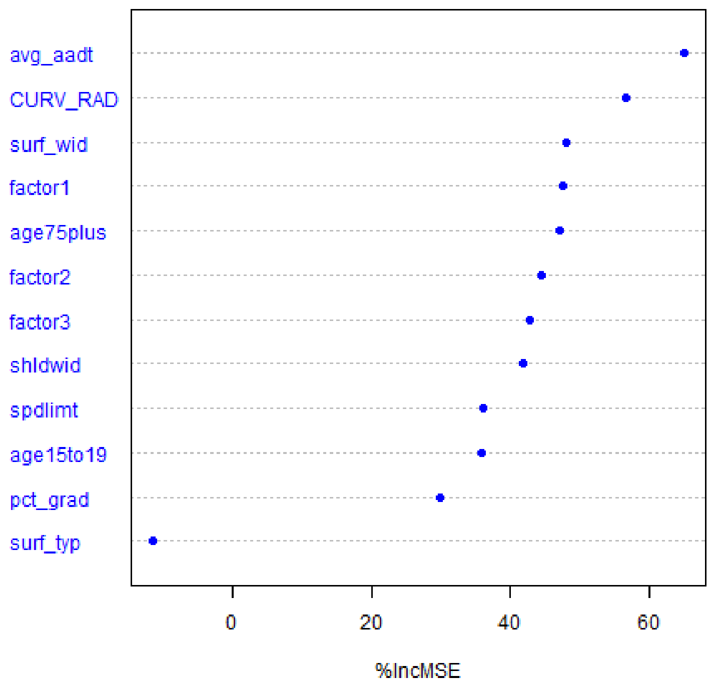
Source: FHWA.

Figure 94. Graph. Output for LNDP-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 119. Contributing factors for LNDP-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Surface width	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percentage of population ages 75+	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Average shoulder width	Decreases
Speed limit	Increases
Percentage of population ages 15–19	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



Source: FHWA.

Figure 95. Graph. Output for LNDP-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

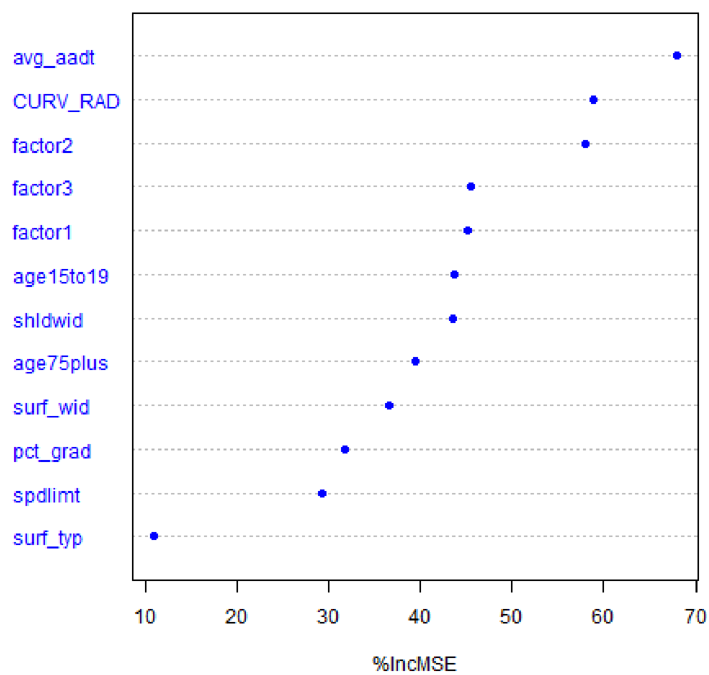
Horizontal Curves and Highway Tangent Segments—Nighttime

Table 120 and table 121 summarize the most influential predictor variables for the expected number of LNDP-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (LNDP-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (LNDP-KABCO-N) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 96 and figure 97.

Table 120. Contributing factors for LNDP-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Curve radius	Increases/decreases*
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percentage of population ages 15–19	Increases
Average shoulder width	Decreases
Percentage of population ages 75+	Increases
Surface width	Decreases
Percent grade	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



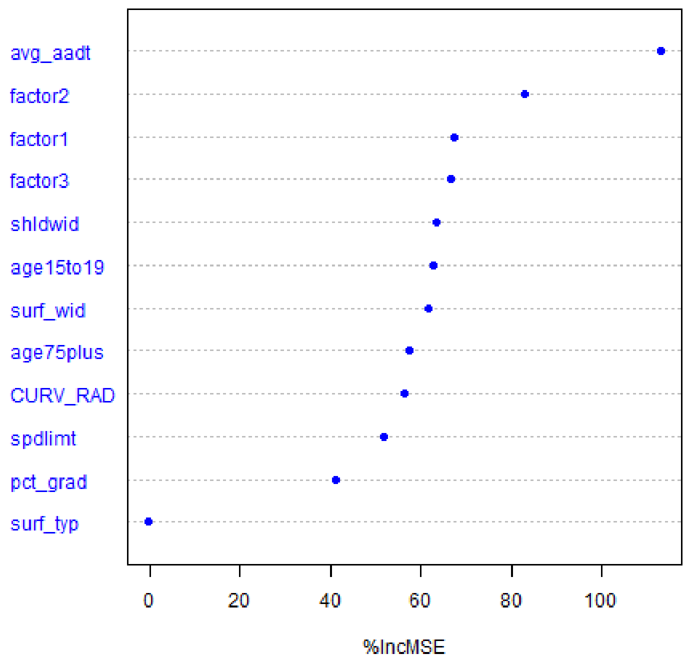
Source: FHWA.

Figure 96. Graph. Output for LNDP-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 121. Contributing factors for LNDP-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Average shoulder width	Decreases
Percentage of population ages 15–19	Increases
Surface width	Decreases
Percentage of population ages 75+	Decreases
Curve radius	Increases/decreases*
Speed limit	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



Source: FHWA.

Figure 97. Graph. Output for LNDP-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with LNDP crashes on rural two-lane horizontal curves and highway tangent segments in Ohio:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of LNDP-D and LNDP-N crashes.
- *Percent grade*: an increase in percent grade was consistently associated with an increase in the frequency of LNDP-D and LNDP-N crashes.
- *Average shoulder width*: an increase in shoulder width was consistently associated with a decrease in the frequency of LNDP-D and LNDP-N crashes (and therefore a decrease in shoulder width was consistently associated with an increase in the frequency of LNDP crashes).
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of LNDP-D and LNDP-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of LNDP crashes).
- *Surface width*: an increase in surface width was associated with a decrease in the frequency of LNDP-D and LNDP-N crashes on curves and tangent segments (and therefore a decrease in surface width was consistently associated with an increase in the frequency of LNDP crashes).
- *Speed*: an increase in speed limit was consistently associated with an increase in the frequency of LNDP-D and LNDP-N crashes.

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.
- The percentage of the population ages 75+ consistently appeared as a factor that decreases daytime and nighttime predicted crash frequency.
- Factor 1 (less education, low income, and less vehicle ownership) consistently appeared as a factor that decreases daytime and nighttime predicted crash frequency.
- Factor 2 (less education, high income, and less vehicle ownership) consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.
- Factor 3 (more education, high income, and more vehicle ownership) consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.

HEO Crashes on Rural Two-Lane Highway Segments in Ohio

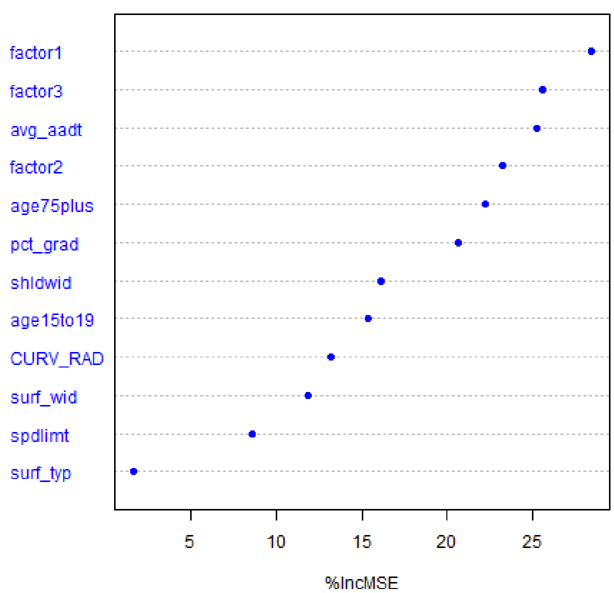
Horizontal Curves and Highway Tangent Segments—Daytime

Table 122 and table 123 summarize the most influential predictor variables for the expected number of HEO-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (HEO-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 98 and figure 99.

Table 122. Contributing factors for HEO-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Average AADT	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Percentage of population ages 75+	Increases
Percent grade	Increases
Average shoulder width	Increases**
Percentage of population ages 15–19	Decreases
Curve radius	Increases/decreases*
Surface width	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



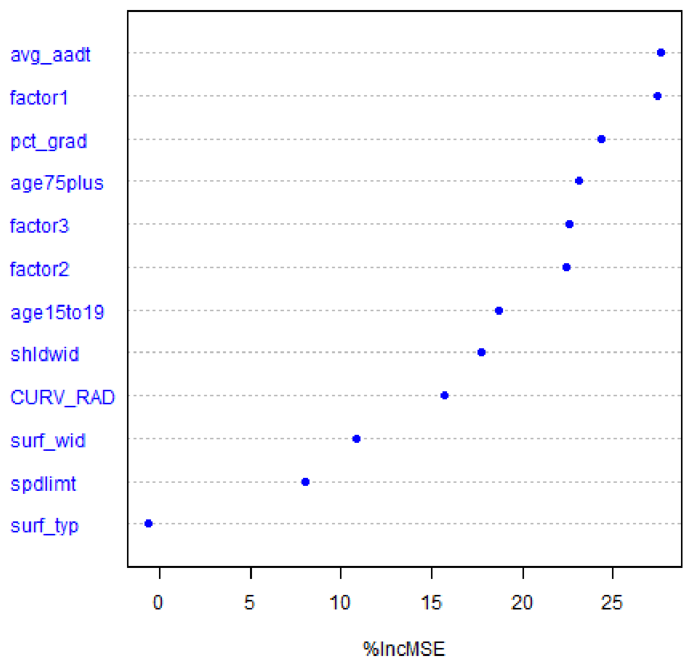
Source: FHWA.

Figure 98. Graph. Output for HEO-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 123. Contributing factors for HEO-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percent grade	Increases
Percentage of population ages 75+	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Percentage of population ages 15–19	Decreases
Average shoulder width	Increases**
Curve radius	Increases/decreases*
Surface width	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



Source: FHWA.

Figure 99. Graph. Output for HEO-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

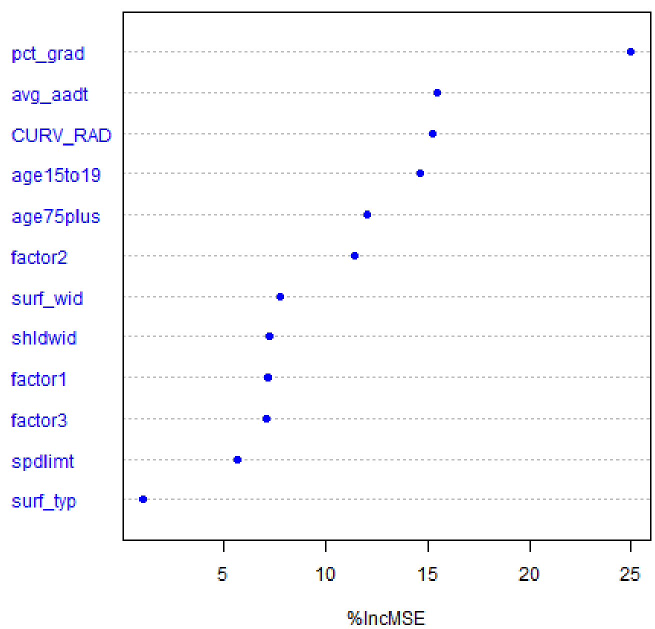
Horizontal Curves and Highway Tangent Segments—Nighttime

Table 124 and table 125 summarize the most influential predictor variables for the expected number of HEO-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (HEO-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (HEO-KABCO-N) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 100 and figure 101.

Table 124. Contributing factors for HEO-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Increases
Average AADT	Increases
Curve radius	Increases/decreases*
Percentage of population ages 15–19	Increases
Percentage of population ages 75+	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Surface width	Increases
Average shoulder width	Increases**
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Speed limit	Decreases**

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



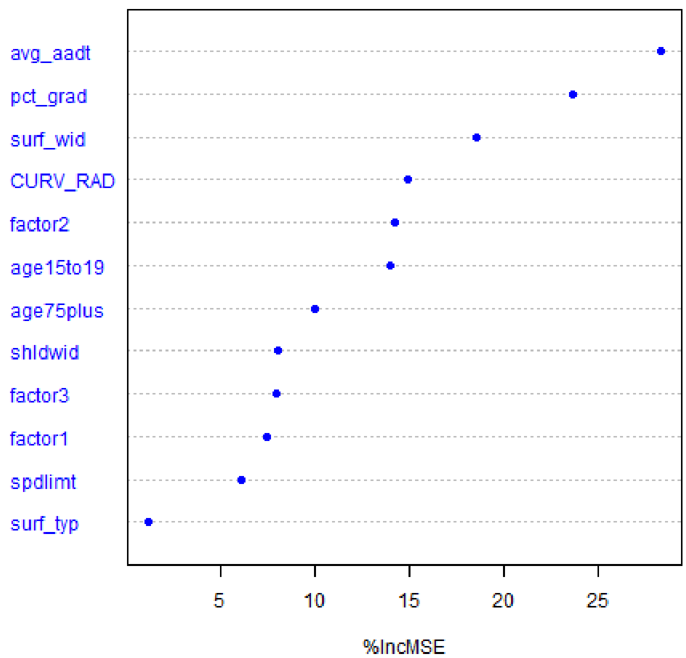
Source: FHWA.

Figure 100. Graph. Output for HEO-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 125. Contributing factors for HEO-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Percent grade	Increases
Surface width	Increases
Curve radius	Increases/decreases*
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Percentage of population ages 75+	Decreases
Average shoulder width	Increases**
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Speed limit	Decreases**

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



Source: FHWA.

Figure 101. Graph. Output for HEO-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with HEO crashes on rural two-lane horizontal curves and highway tangent segments in Ohio:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of HEO-D and HEO-N crashes.
- *Percent grade*: an increase in percent grade was consistently associated with an increase in the frequency of HEO-D and HEO-N crashes.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of HEO-D and HEO-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of LNDP crashes).
- *Surface width*: an increase in surface width was associated with an increase in the frequency of HEO-D and HEO-N crashes.

There were additional roadway variables that showed relationships with HEO crash frequency, but they were not recommended as contributing factors due inconsistent results:

- *Average shoulder width*: an increase in shoulder width was consistently associated with an increase in the frequency of HEO-D and HEO-N crashes (this is a counterintuitive finding, as an increase in shoulder width is generally associated with a decrease in crashes).
- *Speed*: an increase in speed limit was associated with an increase in the frequency of HEO-D crashes (during the nighttime, an increase in speed limit was associated with a decrease in the frequency of HEO crashes).

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 appeared as a factor that increases the predicted crash frequency of nighttime crashes and a factor that decreases the predicted crash frequency of daytime crashes.
- The percentage of the population ages 75+ appeared as a factor that increases the predicted crash frequency of HEO-KAB-D, HEO-KAB-N, and HEO-KABCO-D crashes. It also appeared as a factor that decreases the predicted crash frequency of HEO-KABCO-N crashes.
- Factor 1 (less education, low income, and less vehicle ownership) consistently appeared as a factor that decreases daytime and nighttime predicted crash frequency.
- Factor 2 (less education, high income, and less vehicle ownership) consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.
- Factor 3 (more education, high income, and more vehicle ownership) consistently appeared as a factor that increases daytime and nighttime predicted crash frequency.

ROLL Crashes on Rural Two-Lane Highway Segments in Ohio

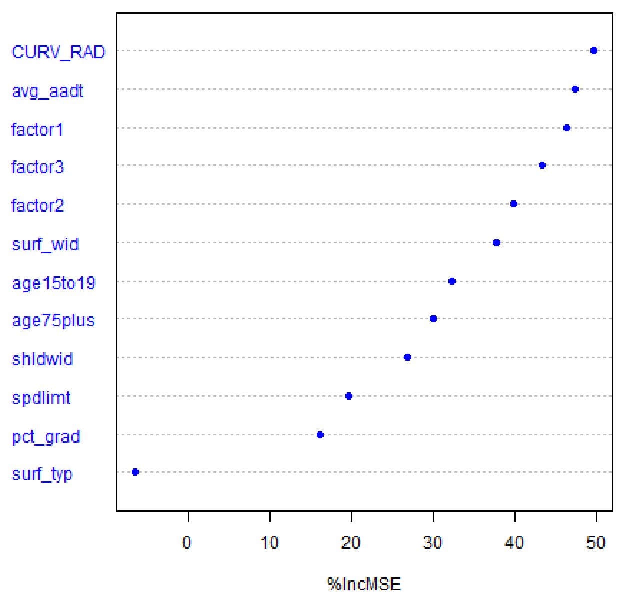
Horizontal Curves and Highway Tangent Segments—Daytime

Table 126 and table 127 summarize the most influential predictor variables for the expected number of ROLL-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROLL-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 102 and figure 103.

Table 126. Contributing factors for ROLL-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Average AADT	Decreases**
Factor 1 (less education, low income, and less vehicle ownership)	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Surface width	Decreases
Percentage of population ages 15–19	Increases
Percentage of population ages 75+	Decreases
Average shoulder width	Increases**
Speed limit	Increases
Percent grade	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



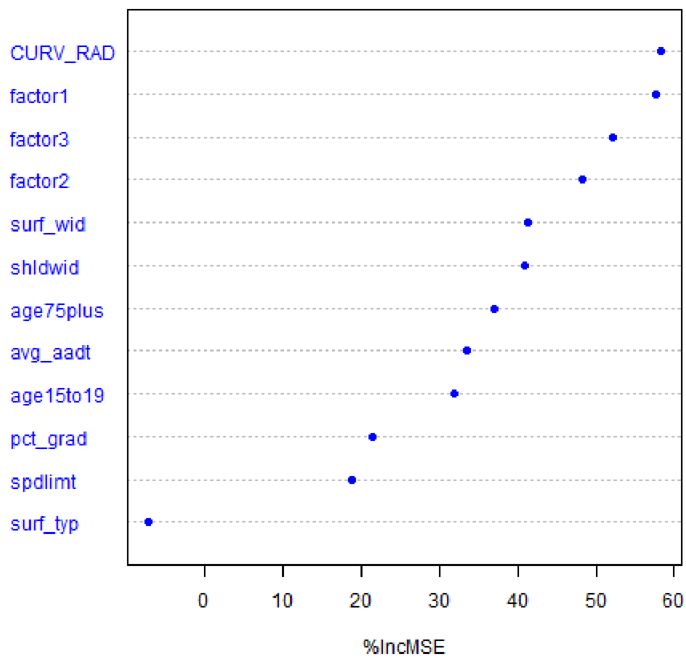
Source: FHWA.

Figure 102. Graph. Output for ROLL-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 127. Contributing factors for ROLL-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Surface width	Decreases
Average shoulder width	Decreases
Percentage of population ages 75+	Decreases
Average AADT	Decreases**
Percentage of population ages 15–19	Increases
Percent grade	Increases
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



Source: FHWA.

Figure 103. Graph. Output for ROLL-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

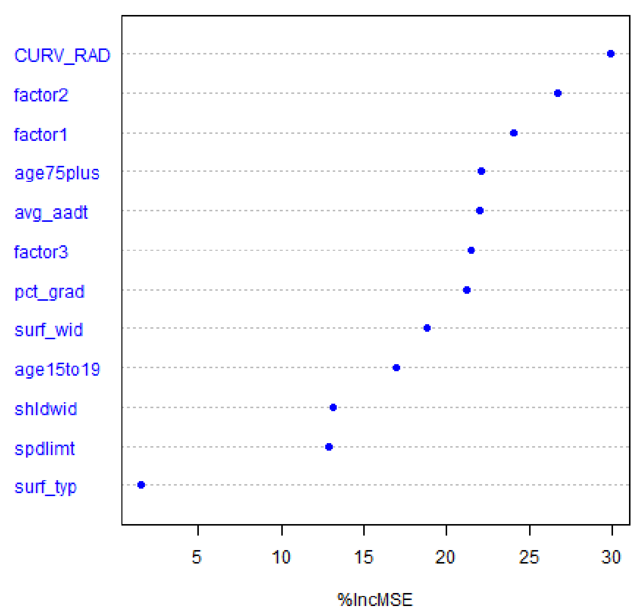
Horizontal Curves and Highway Tangent Segments—Nighttime

Table 128 and table 129 summarize the most influential predictor variables for the expected number of ROLL-N crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ROLL-KAB-N) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ROLL-KABCO-N) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 104 and figure 105.

Table 128. Contributing factors for ROLL-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Curve radius	Increases/decreases*
Factor 2 (less education, high income, and less vehicle ownership)	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percentage of population ages 75+	Decreases
Average AADT	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Percent grade	Increases
Surface width	Increases
Percentage of population ages 15–19	Increases
Average shoulder width	Increases**
Speed limit	Increases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



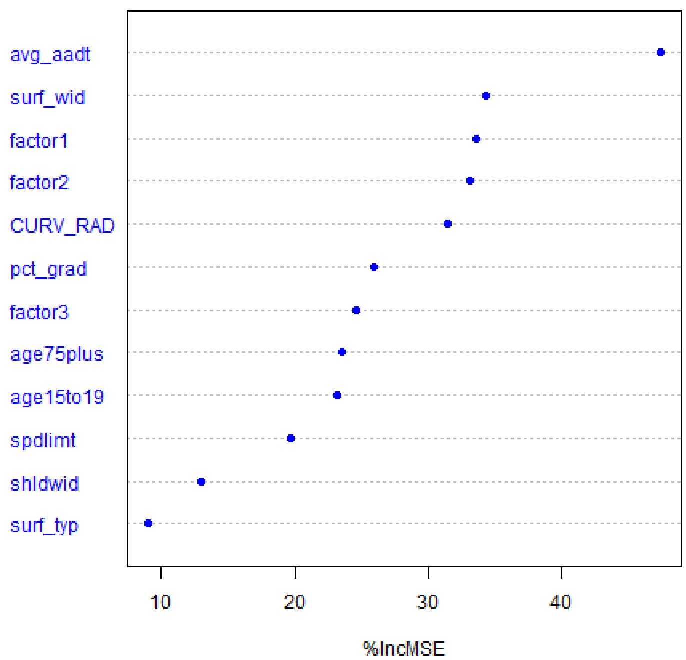
Source: FHWA.

Figure 104. Graph. Output for ROLL-KAB-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 129. Contributing factors for ROLL-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Average AADT	Increases
Surface width	Decreases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Curve radius	Increases/decreases*
Percent grade	Increases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Percentage of population ages 75+	Increases
Percentage of population ages 15–19	Decreases
Speed limit	Increases
Average shoulder width	Decreases

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).



Source: FHWA.

Figure 105. Graph. Output for ROLL-KABCO-N crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ROLL crashes on rural two-lane horizontal curves and highway tangent segments in Ohio:

- *Percent grade*: an increase in percent grade was consistently associated with an increase in the frequency of ROLL-D and ROLL-N crashes.
- *Curve radius*: an increase in curve radius was consistently associated with a decrease in the frequency of ROLL-D and ROLL-N crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of LNDP crashes).
- *Speed*: an increase in speed limit was associated with an increase in the frequency of ROLL-D and ROLL-N crashes.

There were additional roadway variables that showed relationships with HEO crash frequency, but they were not recommended as contributing factors due to inconsistent results:

- *Average AADT*: an increase in AADT was associated with an increase in the frequency of ROLL-D and ROLL-N crashes (during the daytime, an increase in AADT was associated with a decrease in the frequency of ROLL crashes).
- *Surface width*: an increase in surface width was associated with a decrease in the frequency of ROLL-KAB-D, ROLL-KACO-D, and ROLL-KABCO-N (and therefore a decrease in surface width was associated with an increase in the frequency of these crashes).
- *Average shoulder width*: an increase in shoulder width was consistently associated with an increase in the frequency of ROLL-KAB-D and ROLL-KAB-N crashes (this is a counterintuitive finding, as an increase in shoulder width is generally associated with a decrease in crashes).

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 appeared as a factor that increases the predicted crash frequency of ROLL-KAB-D, ROLL-KABCO-D, and ROLL-KAB-N crashes. It also appeared as a factor that decreases the predicted crash frequency of HEO-KABCO-N crashes.
- The percentage of the population ages 75+ appeared as a factor that decreases the predicted crash frequency of ROLL-KAB-D, ROLL-KAB-N, and ROLL-KABCO-D crashes. It also appeared as a factor that decreases the predicted crash frequency of ROLL-KABCO-N crashes.
- Factor 1 (less education, low income, and less vehicle ownership) consistently appeared as a factor that decreases the predicted crash frequency of ROLL-KABCO-D, ROLL-KAB-N, and ROLL-KABCO-N crashes. It also appeared as a factor that increases the predicted crash frequency of ROLL-KAB-D crashes.
- Factor 2 (less education, high income, and less vehicle ownership) appeared as a factor that increases the predicted crash frequency of ROLL-KAB-D, ROLL-KABCO-D, and

ROLL-KABCO-N crashes. It also appeared as a factor that decreases the predicted crash frequency of ROLL-KAB-N crashes.

- Factor 3 (more education, high income, and more vehicle ownership) appeared as a factor that increases the predicted crash frequency of ROLL-KABCO-D, ROLL-KAB-N, and ROLL-KABCO-N crashes. It also appeared as a factor that decreases the predicted crash frequency of ROLL-KAB-D crashes.

ANG Crashes on Rural Two-Lane Highway Segments in Ohio

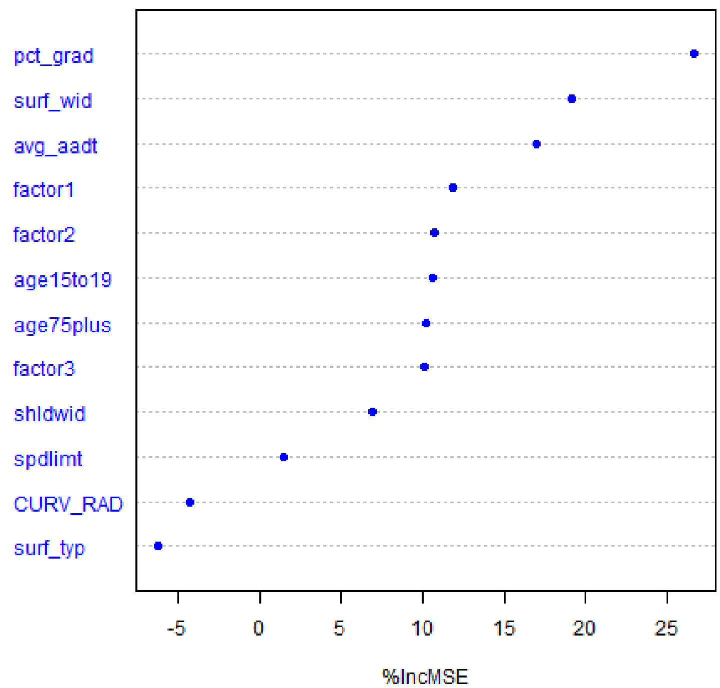
Horizontal Curves and Highway Tangent Segments—Daytime

Table 130 and table 131 summarize the most influential predictor variables for the expected number of ANG-D crashes on rural two-lane horizontal curves and highway tangent segments that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 106 and figure 107.

Table 130. Contributing factors for ANG-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Percent grade	Decreases**
Surface width	Increases
Average AADT	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Percentage of population ages 75+	Decreases
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Average shoulder width	Increases**
Speed limit	Decreases**

**Counterintuitive finding.



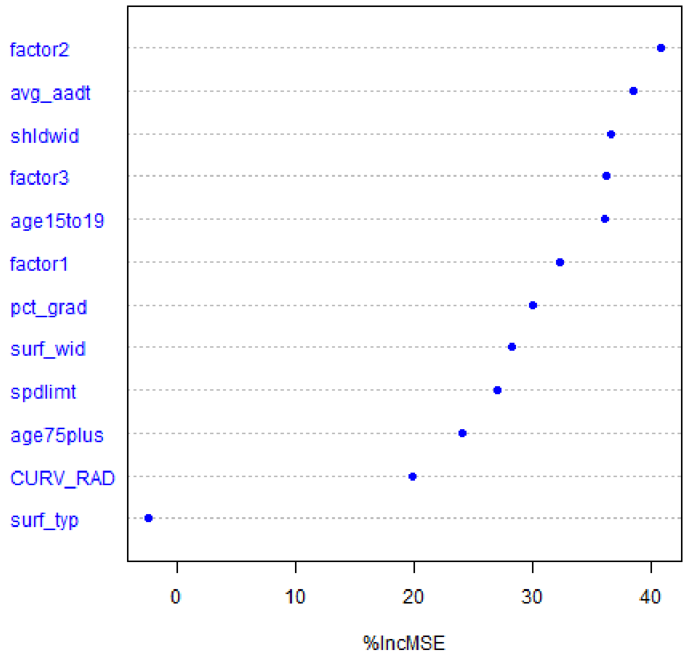
Source: FHWA.

Figure 106. Graph. Output for ANG-KAB-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Table 131. Contributing factors for ANG-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Variable	Impact on Crash-Frequency Predictions
Factor 2 (less education, high income, and less vehicle ownership)	Increases
Average AADT	Increases
Average shoulder width	Increases**
Factor 3 (more education, high income, and more vehicle ownership)	Increases
Percentage of population ages 15–19	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Percent grade	Decreases**
Surface width	Increases
Percentage of population ages 75+	Increases
Curve radius	Increases/decreases*

*Increases crash frequency when comparing curves to tangent segments/decreases crash frequency when comparing curves (i.e., a larger radius is associated with fewer crashes).
**Counterintuitive finding.



Source: FHWA.

Figure 107. Graph. Output for ANG-KABCO-D crashes on rural two-lane horizontal curves and highway tangent segments: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ANG crashes on rural two-lane horizontal curves and highway tangent segments in Ohio:

- *Average AADT*: an increase in AADT was consistently associated with an increase in the frequency of ANG-D crashes.
- *Curve radius*: an increase in curve radius was associated with a decrease in the frequency of ANG-KABCO-D crashes (and therefore a decrease in curve radius was consistently associated with an increase in the frequency of ANG-KABCO-D crashes).
- *Surface width*: an increase in surface width was consistently associated with an increase in the frequency of ANG-D crashes.

There were additional roadway variables that showed relationships with HEO crash frequency, but they were not recommended as contributing factors due to inconsistent results:

- *Percent grade*: an increase in percent grade was consistently associated with a decrease in the frequency of ANG-D crashes (this is a counterintuitive finding, as an increase in percent grade is generally associated with an increase in crashes).
- *Average shoulder width*: an increase in shoulder width was consistently associated with an increase in the frequency of ANG-D crashes (this is a counterintuitive finding, as an increase in shoulder width is generally associated with a decrease in crashes).
- *Speed*: an increase in speed limit was consistently associated with a decrease in the frequency of ANG-D crashes.

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 consistently appeared as a factor that increases predicted crash frequency.
- The percentage of the population ages 75+ appeared as a factor that decreases the predicted crash frequency of ANG-KAB-D crashes and a factor that increases the predicted crash frequency of ANG-KABCO-D crashes.
- Factor 1 (less education, low income, and less vehicle ownership) consistently appeared as a factor that decreases predicted crash frequency.
- Factor 2 (less education, high income, and less vehicle ownership) consistently appeared as a factor that increases predicted crash frequency.
- Factor 3 (more education, high income, and more vehicle ownership) consistently appeared as a factor that increases predicted crash frequency.

ANG Crashes at Four-Leg Stop-Controlled Intersections on Urban Two-Lane Roads in Ohio

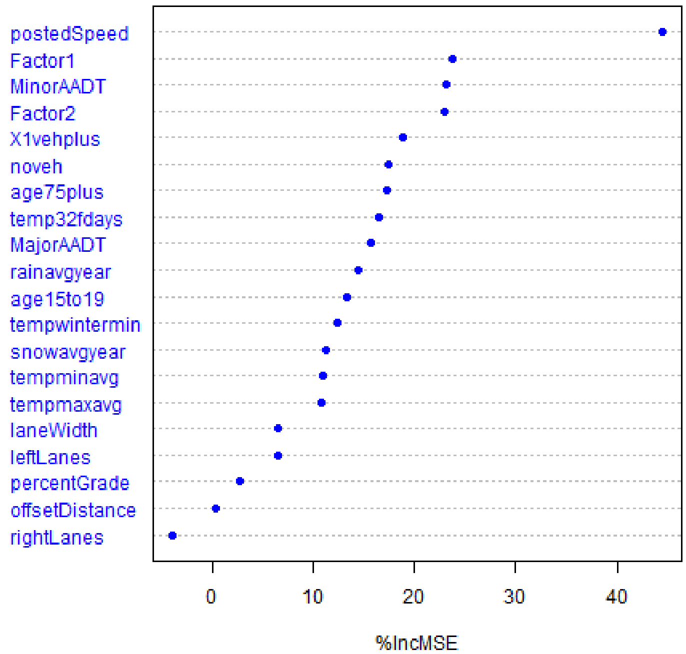
Daytime

Table 132 and table 133 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg stop-controlled intersections (with stop control on the minor road) on urban, two-lane roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 108 and figure 109.

Table 132. Contributing factors for ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Speed limit	Increases
Factor 1 (less education, low income, and less vehicle ownership)	Decreases
Cross street AADT	Increases
Factor 2 (less education, high income, and less vehicle ownership)	Decreases
Percentage of households with any number of vehicles	Increases
Percentage of households with 0 vehicles	Decreases
Percentage of population ages 75+	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Increases
Mainline AADT	Decreases**
Average annual rainfall totals	Decreases
Percentage of population ages 15–19	Decreases

**Counterintuitive finding.



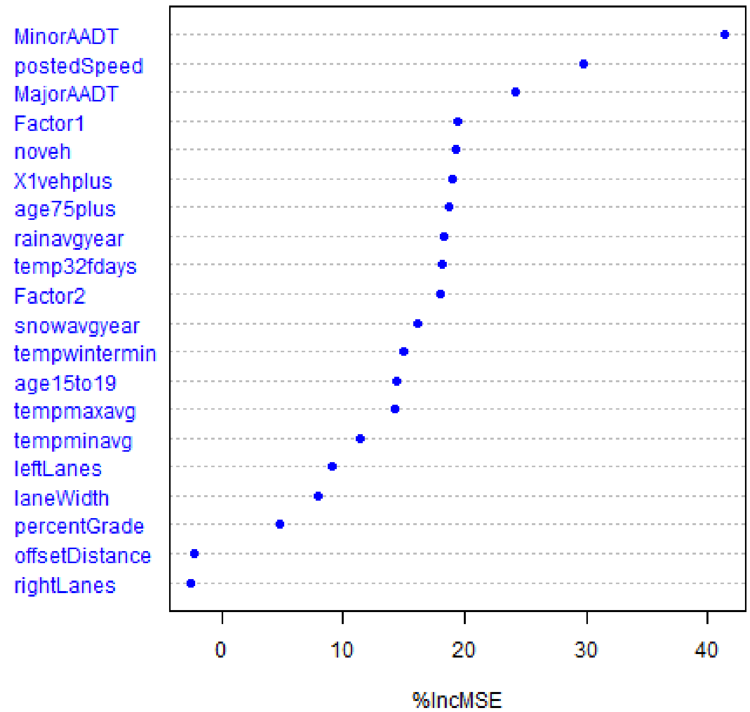
Source: FHWA.

Figure 108. Graph. Output for ANG-KAB-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Table 133. Contributing factors for ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Speed limit	Increases
Mainline AADT	Increases
Factor 1 (no education and low income)	Decreases
Percentage of households with 0 vehicles	Increases**
Percentage of households with any number of vehicles	Decreases**
Percentage of population ages 75+	Decreases
Average annual rainfall totals	Decreases
Annual average number of days with a minimum temperature of ≤32°F	Increases
Factor 2 (diploma education and medium income)	Decreases
Average annual snowfall totals	Increases
Average annual winter minimum temperature	Decreases**
Percentage of population ages 15–19	Decreases

**Counterintuitive finding.



Source: FHWA.

Figure 109. Graph. Output for ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ANG crashes at four-leg stop-controlled intersections (with stop control on the minor road) on urban two-lane roads in Ohio:

- *Cross Street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads.
- *Speed*: an increase in speed was consistently associated with an increase in the frequency of ANG-D crashes at four-leg stop-controlled intersections on urban two-lane roads.

There were additional roadway variables that showed relationships with HEO crash frequency, but they were not recommended as contributing factors due to inconsistent results:

- *Mainline AADT*: an increase in mainline AADT was associated with an increase in the frequency of ANG-KABCO-D crashes at four-leg stop-controlled intersections on urban two-lane roads.

With respect to weather and sociodemographic characteristics, the following factors were observed:

- The percentage of the population ages 15–19 consistently appeared as a factor that decreases predicted crash frequency.
- The percentage of the population ages 75+ consistently appeared as a factor that decreases predicted crash frequency.
- Factor 1 (no education and low income) consistently appeared as a factor that decreases predicted crash frequency.
- Factor 2 (diploma education and medium income) consistently appeared as a factor that decreases predicted crash frequency.
- The percentage of households with no vehicles appeared as a factor that increases the predicted crash frequency of ANG-KAB-D crashes and as a factor that decreases the predicted crash frequency of ANG-KABCO-D crashes.
- The percentage of households with any number of vehicles appeared as a factor that increases the predicted crash frequency of ANG-KAB-D crashes and as a factor that decreases the predicted crash frequency of ANG-KABCO-D crashes.
- Average annual rainfall totals consistently appeared as a factor that decreases predicted crash frequency.
- Average annual snowfall totals appeared as a factor that increases the predicted crash frequency of ANG-KABCO-D crashes.
- Average annual number of days with a minimum temperature of ≤32°F appeared consistently as a factor that increases predicted crash frequency.

ANG Crashes at Four-Leg Stop-Controlled Intersections on Urban Two-Lane Roads in Ohio

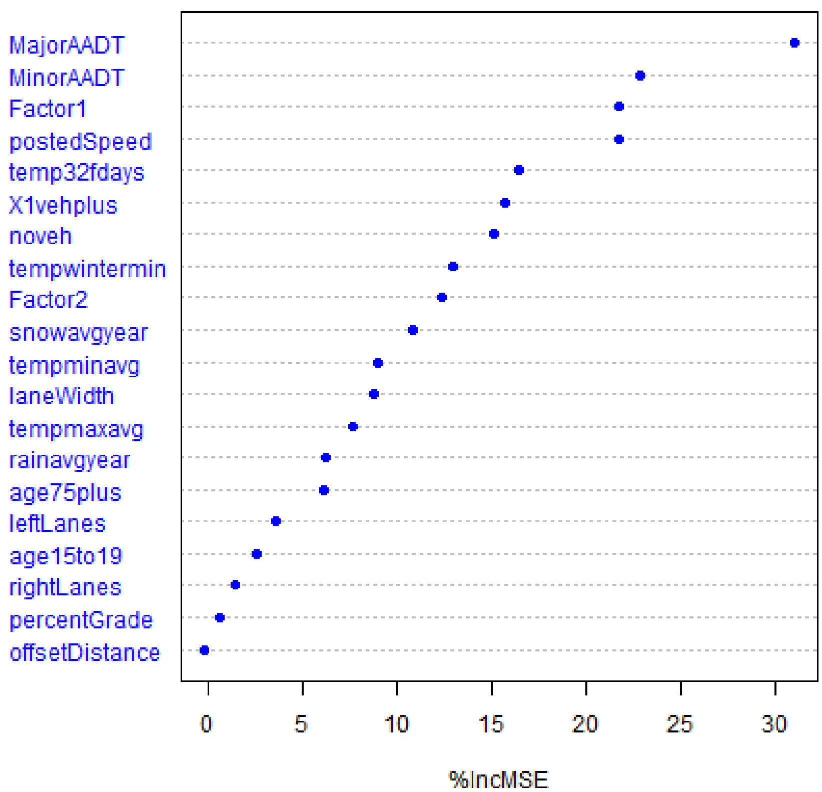
Daytime

Table 134 and table 135 summarize the most influential predictor variables for the expected number of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads that result in fatality, incapacitating injury, or nonincapacitating injury (ANG-KAB-D) or fatality, incapacitating injury, nonincapacitating injury, possible injury, or PDO (ANG-KABCO-D) according to random forests generated using Ohio data. The random-forest outputs for both crash severities are shown in figure 110 and figure 111.

Table 134. Contributing factors for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Mainline AADT	Increases
Cross street AADT	Increases
Factor 1 (no education and low income)	Increases**
Speed limit	Increases
Annual average number of days with a minimum temperature of ≤32°F	Increases
Percentage of households with any number of vehicles	Decreases**
Percentage of households with 0 vehicles	Increases**
Average annual winter minimum temperature	Decreases**
Factor 2 (diploma education and medium income)	Decreases**
Average annual snowfall totals	Decreases**
Average annual minimum temperature	Increases
Lane width	Decreases

**Counterintuitive finding.



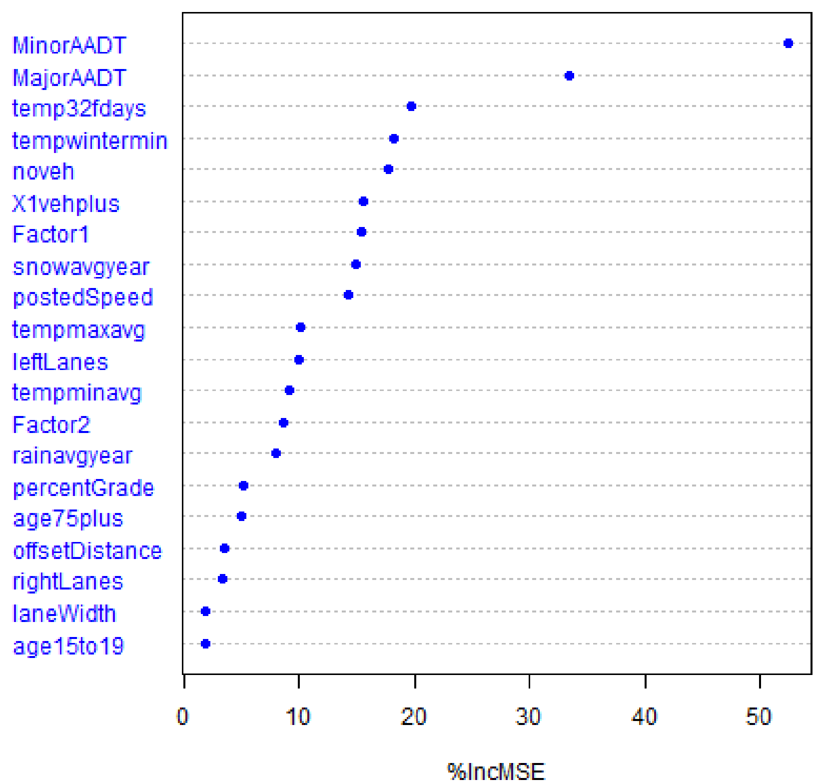
Source: FHWA.

Figure 110. Graph. Output for ANG-KAB-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Table 135. Contributing factors for ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Variable	Impact on Crash-Frequency Predictions
Cross street AADT	Increases
Mainline AADT	Increases
Annual average number of days with a minimum temperature of ≤32°F	Decreases**
Annual average winter minimum temperature	Increases
Percentage of households with 0 vehicles	Increases**
Percentage of households with any number of vehicles	Decreases**
Factor 1 (no education and low income)	Decreases
Average annual snowfall totals	Decreases**
Speed limit	Increases
Average annual maximum temperature	Increases
Number of channelized left-turn lanes	Increases
Average annual minimum temperature	Increases
Factor 2 (diploma education and medium income)	Decreases**

**Counterintuitive finding.



Source: FHWA.

Figure 111. Graph. Output for ANG-KABCO-D crashes at four-leg signalized intersections on urban multilane undivided roads: Ohio.

Discussion

Based on the analysis, the research team recommends the following roadway contributing factors associated with ANG crashes at four-leg signalized intersections on urban multilane undivided roads in Ohio:

- *Mainline AADT*: an increase in mainline AADT was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads.
- *Cross Street AADT*: an increase in cross street AADT was associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads.
- *Speed*: an increase in speed was consistently associated with an increase in the frequency of ANG-D crashes at four-leg signalized intersections on urban multilane undivided roads.

There were additional roadway variables that showed relationships with HEO crash frequency, but they were not recommended as contributing factors due to inconsistent results:

- *Lane width*: an increase in lane width was associated with a decrease in the frequency of ANG-KAB-D crashes at four-leg stop signalized intersections on urban multilane undivided roads.
- *Number of approaches with left turn lanes*: an increase in the number of approaches with left-turn lanes at an intersection was associated with an increase in the frequency of ANG-KABCO-D crashes at four-leg stop signalized intersections on urban multilane undivided roads.

With respect to weather and sociodemographic characteristics, the following factors were observed:

- Factor 1 (no education and low income) appeared as a factor that decreases the predicted crash frequency of ANG-KABCO-D crashes and a factor that increases the predicted crash frequency of ANG-KAB-D crashes.
- Factor 2 (diploma education and medium income) consistently appeared as a factor that decreases predicted crash frequency.
- The percentage of households with no vehicles consistently appeared as a factor that increases predicted crash frequency.
- The percentage of households with any number of vehicles consistently appeared as a factor that decreases predicted crash frequency.
- Average annual snowfall totals consistently appeared as a factor that decreases predicted crash frequency.
- Average annual minimum temperature consistently appeared as a factor that increases predicted crash frequency.

SUMMARY OF RESULTS FROM THE CONTRIBUTING-FACTOR ANALYSIS

The previous sections provided the updated analysis for Ohio data using factors to reduce the number of socioeconomic variables and to account for high correlation between them. The updated analysis showed good results with respect to nonintersection FCFTs, but the results for the intersections FCFTs were not as expected.

Intersection FCFTs

The contributing-factor analysis for urban intersections produced consistent results with AADT and speed limit showing as roadway contributing factors. The analysis, however, did not show consistent results with respect to socioeconomic variables. The percentage of households with no vehicles came up as a factor that increases predicted crash frequency, whereas the percentage of households with any number of vehicles came up as factor that decreases predicted crash frequency. Similarly, Factor 1 (no education and low income) showed up as a factor that increases predicted crash frequency at urban signalized intersections. Weather variables also showed some counterintuitive results; for example, the average annual winter minimum temperature and average annual snowfall totals showed up as factors that decrease crash frequencies for urban signalized intersections when, for the same site type, the average annual minimum temperature showed up as a factor that increases predicted crash frequency.

Nonintersection FCFTs

The contributing-factor analysis for rural highway segments produced consistent results with AADT, percent grade, curve radius, and average shoulder width showing as roadway contributing factors. In almost all cases, the percentage of the population ages 15–19 showed up as a factor that increases predicted crash frequency, while the percentage of the population ages 75+ showed up as factor that decreases predicted crash frequency. For socioeconomic factors, in almost all cases, Factor 1 (less education, low income, and less vehicle ownership) showed up as a factor that decreases predictive crash frequency, whereas Factor 2 (less education, high income, and less vehicle ownership) and Factor 3 (more education, high income, and more vehicle ownership) showed up as factors that increase predicted crash frequency.

APPENDIX H. OBSERVING PREDICTION MSEs TO EXPLORE CONTRIBUTIONS OF SOCIOECONOMIC VARIABLES

To further explore the contributions of socioeconomic variables to the random-forest analysis, the research team calculated and compared prediction MSEs using random-forest predictions from analyses using data with and without socioeconomic variables. For these comparisons, the research team used daytime crashes on horizontal curves and tangent segments on rural two-lane roads in Ohio. Random forests were estimated using only roadway variables and compared to the models reported in appendix G (the roadway-only random-forest outputs are provided in figure 112 and figure 113).

MSE values for predictions from the roadway-only random forests were compared to MSE values for predictions from random forests presented in appendix G, which included socioeconomic variables as factors to account for the high correlation between the different socioeconomic variables.

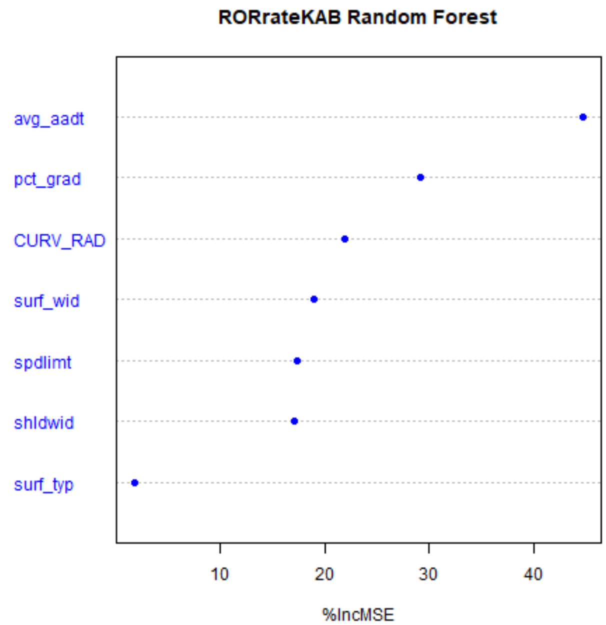
Table 136 provides a comparison of the MSEs for the two analyses.

Table 136. Comparison of MSEs for analyses using data with and without socioeconomic variables.

Crash-Severity Type	Predictions with Socioeconomic Variables	Predictions without Socioeconomic Variables
ROR-KAB-D	7.19	9.39
ROR-KABCO-D	30.99	41.65
LNDP-KAB-D	9.11	12.18
LNDP-KABCO-D	35.05	47.80
ROLL-KAB-D	1.47	1.85
ROLL-KABCO-D	2.36	3.04
HEO-KAB-D	0.23	0.29
HEO-KABCO-D	0.29	0.37
ANG-KAB-D	0.27	0.31
ANG-KABCO-D	1.37	1.77

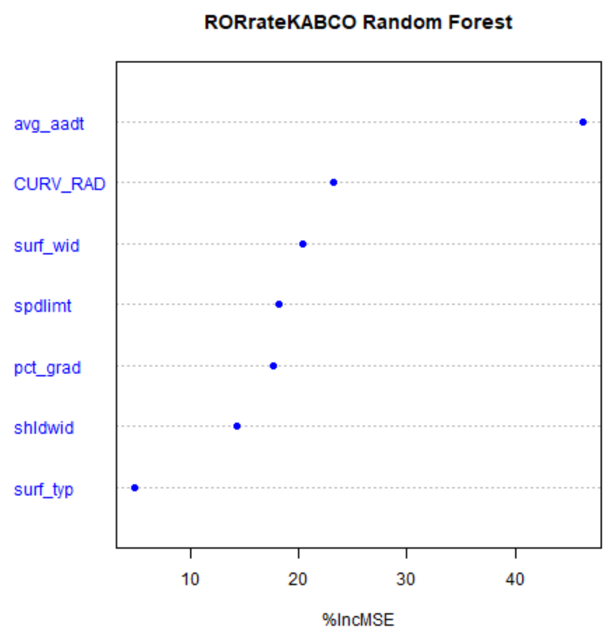
When comparing MSEs, lower values are better as they suggest that predicted values were closer to observed values. It can be seen from table 136 that MSEs for predictions using data with socioeconomic variables are consistently 20- to 30-percent lower than MSEs for predictions using data without socioeconomic variables.

Based on this analysis, inclusion of socioeconomic variables in analyses leads to improved crash predictions.



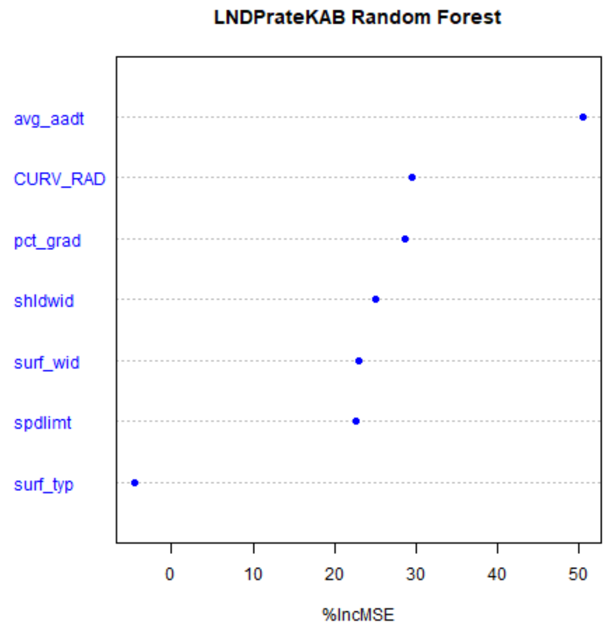
Source: FHWA.

Figure 112. Graph. Output for ROR-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



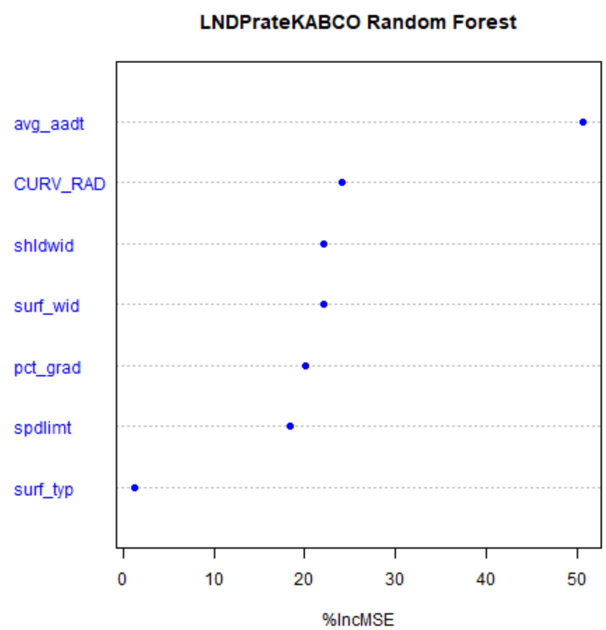
Source: FHWA.

Figure 113. Graph. Output for ROR-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



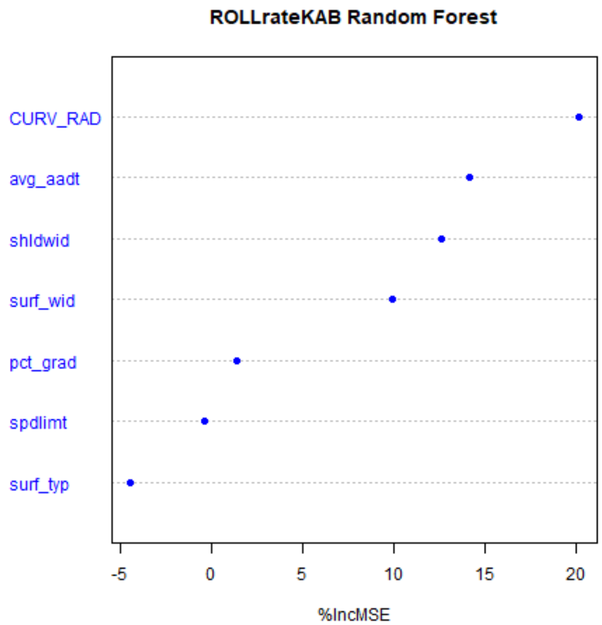
Source: FHWA.

Figure 114. Graph. Output for LNDP-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



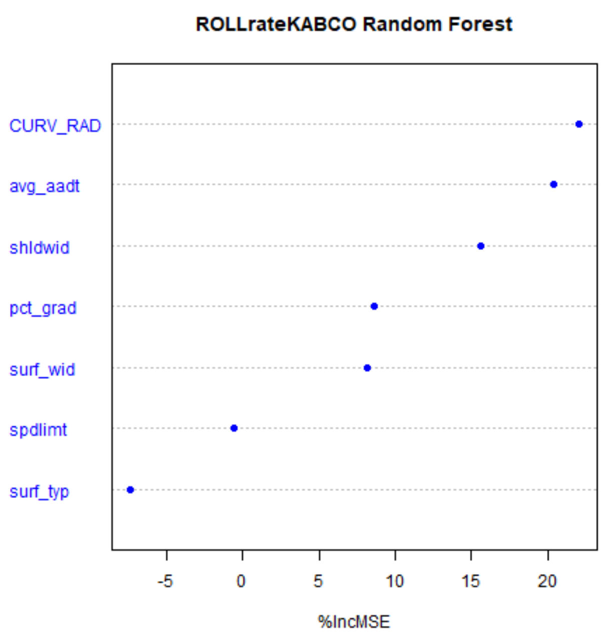
Source: FHWA.

Figure 115. Graph. Output for LNDP-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



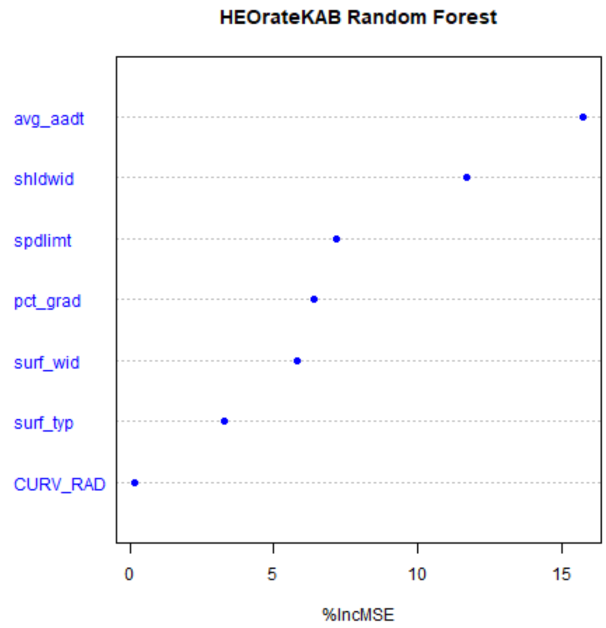
Source: FHWA.

Figure 116. Graph. Output for ROLL-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



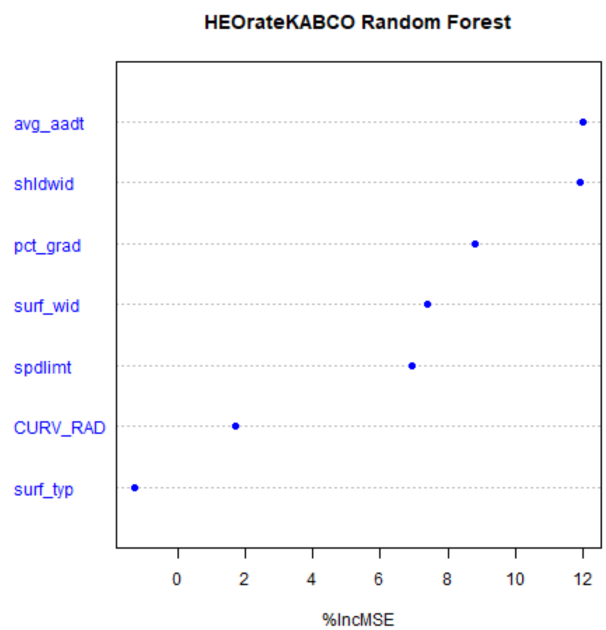
Source: FHWA.

Figure 117. Graph. Output for ROLL-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



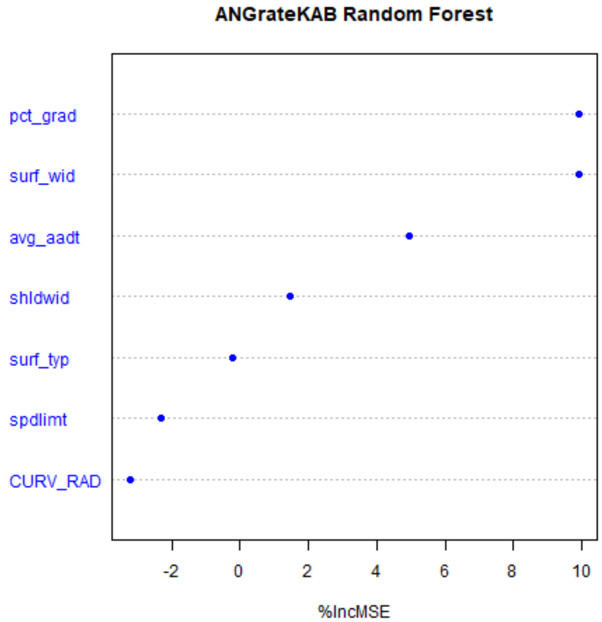
Source: FHWA.

Figure 118. Graph. Output for HEO-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



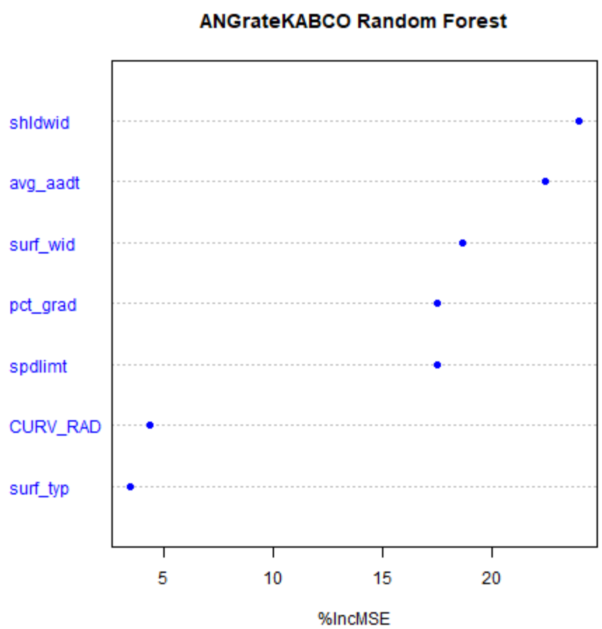
Source: FHWA.

Figure 119. Graph. Output for HEO-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



Source: FHWA.

Figure 120. Graph. Output for ANG-KAB-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.



Source: FHWA.

Figure 121. Graph. Output for ANG-KABCO-D crashes on rural two-lane horizontal curves and tangent segments: Ohio.

REFERENCES

AAA Foundation for Traffic Safety. 2012. *usRAP Coding Manual for Star Ratings and Safer Roads Investment Plans*. Washington, DC: American Automobile Association Foundation for Traffic Safety.

AASHTO. 2010. *Highway Safety Manual*. Washington, DC: American Association of State Highway and Transportation Officials. Accessed May 15, 2020. http://www.highwaysafetymanual.org/Pages/default.aspx.

AASHTO. 2014. *Highway Safety Manual 1st Edition Supplement.* Washington, DC: American Association of State Highway and Transportation Officials. Accessed May 15, 2020. https://store.transportation.org.

AASHTO. n.d. "AAHSHTOWare Safety Analyst." Accessed May 12, 2020. http://www.safetyanalyst.org.

Albin, R., V. Brinkly, J. Cheung, F. Julian, C. Satterfield, W. Stein, E. Donnell, et al. 2016. *Low-Cost Treatments for Horizontal Curve Safety 2016*. Report number FHWA-SA-15-084. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/roadway_dept/horicurves/fhwasa15084.

Al-Kaisy, A., L. Ewam, D. Venaziano, and F. Hossain. 2015. *Risk factors associated with high potential for serious crashes.* Report No. FHWA-OR-RD-16-05, Washington, DC: Federal Highway Administration; Salem, OR: Oregon Department of Transportation. Accessed May 15, 2020. https://ruralsafetycenter.org/resources/list/risk-factors-associated-with-high-potential-for-serious-crashes.

Breiman, L., and A. Cutler. 2013. "Random Forests" (website). Accessed January 19, 2018. http://www.stat.berkeley.edu/~breiman/RandomForests.

Brown, J. V., M. Martello, and R. R. Souleyrette. 2011. *Minnesota Department of Transportation Traffic Safety Analysis Software State of the Art.* Report No. MN/RC 2011-10. Saint Paul, MN: Minnesota Department of Transportation. Accessed May 15, 2020. http://www.dot.state.mn.us/research/TS/2011/2011-10.pdf.

Caldwell, R. C., and E. M. Wilson. 1996. "Variable Safety Improvements for Unpaved Roads." Presented at the *1996 Institute of Transportation Engineers International Conference, Dana Point, CA, March 1996*.

Calvert, E., D. Haiar, and E. Wilson. 1999. "Development of a Field Evaluation Guide for Unpaved Rural Roads." *Transportation Research Record* 1652. Accepted May 15, 2020. https://journals.sagepub.com/doi/pdf/10.3141/1652-11.

Carlin, B. P. and T. A. Louis. 2008. *Bayesian Methods for Data Analysis*. Boca Raton, FL: Chapman and Hall.

Cato. J., A. Ghafoori, A. Kaplan, J. Marandino, P. Ott, and E. Gonzales. 2013. *Development of a Systemic Road Safety Analysis Tool – Roadway Departure Crashes at Bridges in Salem County, New Jersey*. Piscataway, NJ: Rutgers University.

Chatterjee, S., and A. S. Hadi. 2012. *Regression Analysis by Example*. 5th ed. Hoboken, FL: Wiley.

Collett, D. 2003. *Modelling Survival Data in Medical Research*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC. Accessed May 15, 2020. https://www.worldcat.org/title/modelling-survival-data-in-medical-research/oclc/64401278.

Cressie, N. A. C. 1993. *Statistics for Spatial Data*. New York, NY: Wiley.

Cressie, N. A. C., and C. K. Wikle. 2011. *Statistics for Spatio-temporal Data*. Hoboken, NJ: Wiley. Accessed May 15, 2020. https://www.wiley.com/en-us/Statistics+for+Spatio+Temporal+Data-p-9780471692744.

Dobson, A. J., and A. G. Barnett. 2008. *An Introduction to Generalized Linear Models*. 3rd ed. Boca Raton, FL: Chapman & Hall.

Faraway, J. J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, FL: Chapman & Hall/CRC.

Faraway, J. J. 2015. *Linear Models with R*. 2nd ed. Boca Raton, FL: Taylor & Francis. Accessed May 15, 2020. http://www.utstat.toronto.edu/~brunner/books/LinearModelsWithR.pdf.

FHWA. 2004. *Toolbox of Countermeasures and Their Potential Effectiveness to Make Intersections Safer*. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://www.ite.org/pub/?id=e26c7e9c-2354-d714-5181-4cc79fba5459.

FHWA. 2008a. *Toolbox of Countermeasures and Their Potential Effectiveness for Roadway Departure Crashes*. Report No. FHWA-SA-07-013. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/tools/crf/resources/briefs/rdwydepartissue.cfm.

FHWA. 2008b. *Toolbox of Countermeasures and Their Potential Effectiveness for Pedestrian Crashes*. Report No. FHWA-SA-014. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. http://www.pedbikeinfo.org/cms/downloads/pedToolboxofCountermeasures2013.pdf.

FHWA. 2009. *Toolbox of Countermeasures and Their Potential Effectiveness for Intersection Crashes*. Report No. FHWA-SA-10-005. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://cdn.ymaws.com/www.azace.org/resource/resmgr/imported/CrashReductionIntersectionIssueBrief.pdf.

FHWA. 2012. *Manual on Uniform Traffic Control Devices, 2009 Edition with Revision Numbers 1 and 2*. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://mutcd.fhwa.dot.gov/pdfs/2009r1r2/pdf_index.htm.

FHWA. 2014. *Local and Rural Road Safety Briefing Sheets: Applying the Systemic Safety Approach on Local Roads*. Report No. FHWA-SA-14-081. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/local_rural/training/fhwasa14081.

FHWA. 2015a. "Intersection Safety" (website). Accessed January 19, 2018. https://safety.fhwa.dot.gov/intersection.

FHWA. 2015b. *Intersection Safety Strategies Brochure*. 2nd ed. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/intersection/conventional/signalized.

FHWA. 2017a. "Proven Safety Countermeasures" (website). Accessed January 19, 2018. https://safety.fhwa.dot.gov/provencountermeasures.

FHWA. 2017b. "Roadway Departure Safety" (website). Accessed January 19, 2018. https://safety.fhwa.dot.gov/roadway_dept.

FHWA. 2018a. "CMF Clearinghouse. 2018a" (website). Accessed January 19, 2018. http://www.cmfclearinghouse.org.

FHWA. 2018b. "CMF Clearinghouse State Selected CMF Lists" (website). Accessed January 19, 2018. http://www.cmfclearinghouse.org/stateselectedlist.cfm.

FHWA. 2018c. "Highway Safety Information System (HSIS)" (website). Accessed January 19, 2018. https://www.hsisinfo.org.

FHWA. 2018d. "Interactive Highway Safety Design Model (IHSDM)" (website). Accessed January 19, 2018. https://highways.dot.gov/safety/interactive-highway-safety-design-model/interactive-highway-safety-design-model-ihsdm.

FHWA. 2018e. "Roadway Safety Data Program" (website). Accessed January 19, 2018. https://safety.fhwa.dot.gov/rsdp.

FHWA. 2018f. "PEDSAFE: Pedestrian Safety Guide and Countermeasure Selection System" (website). Accessed January 19, 2018. http://www.pedbikesafe.org/PEDSAFE.

FHWA. 2018g. "BIKESAFE: Bicycle Safety Guide and Countermeasure Selection System" (website). Accessed January 19, 2018. http://www.pedbikesafe.org/bikesafe.

Fink, K., and R. Krammes. 1995. "Tangent Length and Sight Distance Effects on Accident Rates at Horizontal Curves on Rural Two-Lane Highways." *Transportation Research Record* 1500: 162–168. Accessed May 15, 2020. http://onlinepubs.trb.org/Onlinepubs/trr/1995/1500/1500-020.pdf.

Fitzmaurice, G. M., N. M. Laird, and J. H. Ware. 2011. *Applied Longitudinal Analysis*. 2nd ed. Hoboken, NJ: Wiley. Accessed May 15, 2020. https://content.sph.harvard.edu/fitzmaur/ala2e.

Fitzpatrick, K., W. Schneider IV, and E. Park. 2005. *Comparisons of Crashes on Rural Two-Lane and Four-Lane Highways in Texas*. Report No. FHWA/TX-06/0-4618-1. Austin, TX: Texas Department of Transportation. Accessed May 15, 2020. https://d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/0-4618-1.pdf.

Fitzpatrick, K., E. Park, and W. Schneider IV. 2008. "Potential Accident Modification Factors for Driveway Density on Rural Highways from Texas Data." *Transportation Research Record* 2083: 49–61. Accessed May 15, 2020. https://journals.sagepub.com/doi/10.3141/2083-06?icid=int.sj-abstract.similar-articles.2.

FMCSA. 2016. "MCMIS Catalog and Documentation" (website). Accessed January 19, 2018. https://ask.fmcsa.dot.gov/app/mcmiscatalog/mcmishome.

FTA. 2018. "The National Transit Database (NTD)" (website). Accessed January 19, 2018. https://www.transit.dot.gov/ntd.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC. Accessed May 15, 2020. https://statisticalsupportandresearch.files.wordpress.com/2017/11/bayesian_data_analysis.pdf.

Goodwin, A., B. Kirley, L. Sandt, W. Hall, L. Thomas, N. O'Brien, and D. Summerlin. 2013. *Countermeasures That Work: A Highway Safety Countermeasure Guide for State Highway Safety Offices, Seventh Edition, 2013*. Report No. DOT-HS-811-727. Washington, DC: National Highway Traffic Safety Administration. Accessed May 15, 2020. http://www.nhtsa.gov/staticfiles/nti/pdf/811727.pdf.

Google. 2018. "Google® Street View™". Google (website). Accessed January 19, 2018. https://www.google.com/streetview.

Gross, F., T. Harmon, G. Bahar, and K. Peach. 2016. *Reliability of Safety Management Methods: Systemic Safety Programs*. Report No. FHWA-SA-16-041. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa16041.pdf.

Hallmark. S., D. Veneziano, T. McDonald, J. Graham, K. Bauer, R. Patel, and F. Council. 2006. *Safety Impacts of Pavement Edge Drop-Offs*. Washington, DC: Foundation for Traffic Safety. Accessed May 15, 2020. https://www.semanticscholar.org/paper/Safety-Impacts-of-Pavement-Edge-Drop-Offs-Hallmark-Veneziano/519e19e603a8c783d4bfb54a6c250d05af9ea60f.

Harkey. D., R. Srinivasan, J. Baek, F. Council, K. Eccles, N. Lefler, F. Gross, et al. 2008. *NCHRP Report 617: Accident Modification Factors for Traffic Engineering and ITS Improvements*. Washington, DC: Transportation Research Board. Accessed May 15, 2020. http://www.trb.org/Publications/Blurbs/156844.aspx.

Harwood, D. W., R. R. Souleyrette, M. A. Fields, and E. R. Green. 2013. "Comparison of Countermeasure Selection Methods for Use in Road Safety Management." Presented at the *92nd Transportation Research Board Annual Meeting, Washington, DC*. Accessed May 15, 2020. https://trid.trb.org/view.aspx?id=1287952.

Hastie, T., Tibshirani, R. and Friedman, J. H. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer. Accessed May 15, 2020. https://web.stanford.edu/~hastie/ElemStatLearn/.

Hedeker, D. R., and R. D. Gibbons. 2006. *Longitudinal Data Analysis*. Hoboken, NJ: Wiley-Interscience.

Hoff, P. 2009. *A First Course in Bayesian Statistical Methods*. New York: Springer-Verlag. Accessed May 15, 2020. https://link.springer.com/book/10.1007/978-0-387-92407-6.

ITE. 2015. "Unsignalized Intersection Improvement Guide" (website). Accessed January 19, 2018. http://www.ite.org/uiig.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning: With Applications in R*. New York, NY: Springer. Accessed May 15, 2020. https://link.springer.com/article/10.1007/s13253-014-0179-9.

Karlaftis, M. and I. Golias. 2002. "Effects of Road Geometry and Traffic Volumes on Rural Roadway Accident Rates." *Accident Analysis and Prevention* 34, 2: 357–365. Accessed May 15, 2020. https://www.sciencedirect.com/science/article/abs/pii/S0001457501000331.

Karwa, V., A. B. Slavković, and E. T. Donnell. 2011. "Causal inference in transportation safety studies: Comparison of potential outcomes and causal diagrams." *The Annals of Applied Statistics* 5, 2b: 1428–1455. Accessed May 15, 2020. https://projecteuclid.org/euclid.aoas/1310562728.

Kleinbaum, D. G. and M. Klein. 2012. *Survival Analysis: A Self-learning Text*. 2nd ed. New York, NY: Springer. Accessed May 15, 2020. https://www.springer.com/gp/book/9781441966452.

Knapp. K., S. Hallmark, and G. Bou-Saab. 2014. *Systemic Safety Improvement Risk Factor Evaluation and Countermeasure Summary: Final Report*. Report No. WBS 25-1121-0003-140. Lincoln, NE: Mid-America Transportation Center. Accessed May 15, 2020. https://intrans.iastate.edu/research/completed/systemic-safety-improvement-risk-factor-evaluation-and-countermeasure-summary.

Lawson, A. B. 2013. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. 2nd ed. Boca Raton, FL: Taylor & Francis. Accessed May 15, 2020. https://doi.org/10.1201/b14073.

Lee, J., and D. Nam. 2003. "Run-Off-Roadway Accident Frequency Models by Type of Accident Severity." *Journal of the Eastern Asia Society for Transportation Studies* 5.

Li, X., Y. Liu, A. M. Asce, Z. Gao, and D. Liu. 2016. "Decision Tree Based Station-Level Rail Transit Ridership Forecast." *Journal of Urban Planning and Development* 142, 4. Accessed May 15, 2020. https://ascelibrary.org/doi/10.1061/%28ASCE%29UP.1943-5444.0000331.

Liu, C. and R. Subramanian. 2009. *Factors Related to Fatal Single-Vehicle Run-Off-Road Crashes*. Report No. DOT-HS-811-232. Washington, DC: National Highway Traffic Safety Administration. Accessed May 15, 2020. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811232.

Liu, C. and T. Ye. 2011. *Run-Off-Road Crashes: An On-Scene Perspective*. Report No. DOT-HS-811-500. Washington, DC: National Highway Traffic Safety Administration. Accessed May 15, 2020. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811500.

Lord, D., D. Middleton, and J. Whitrace. 2005. "Does separating trucks from other traffic improve overall safety?" *Transportation Research Record* 1922: 156–166. Accessed May 15, 2020. https://journals.sagepub.com/doi/abs/10.1177/0361198105192200120.

Lord, D., M. Brewer, K. Fitzpatrick, S. Geedipally, and Y. Peng. 2011. *Analysis of Roadway Departure Crashes on Two-Lane Rural Roads in Texas*. Report No. FHWA/TX-11/0-6031-1. Austin, TX: Texas Department of Transportation. Accessed May 15, 2020. https://static.tti.tamu.edu/tti.tamu.edu/documents/0-6031-1.pdf.

Mahgoub, H., A. Selim, and K. C. Pramod. 2011. "Quantitative Assessment of Local Rural Road Safety – Case Study." Presented at *the 90th Transportation Research Board Annual Meeting, Washington, DC*. Accessed May 15, 2020. https://trid.trb.org/view/1093226.

McCullagh, P., J. A. Nelder. 1994. *Generalized Linear Models*. 2nd ed. London, UK: Chapman & Hall.

Miaou, S. 1994. "The Relationship between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions." *Accident Analysis and Prevention* 26, 4: 471–482. Accessed May 15, 2020. https://www.sciencedirect.com/science/article/abs/pii/0001457594900388.

Milton, J., and F. Mannering. 1996. *The Relationship between Highway Geometrics, Traffic Related Elements, and Motor Vehicle Accidents*. Report No. WA-RD 403.1. Olympia, WA: Washington State Department of Transportation. Accessed May 15, 2020. https://www.wsdot.wa.gov/research/reports/fullreports/403.1.pdf.

Milton, J., and F. Mannering. 1998. "The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies." *Transportation* 25, 4: 395–413. Accessed May 15, 2020. https://www.researchgate.net/publication/226003223_The_relationship_among_highway_geometrics_traffic-related_elements_and_motor-vehicle_accident_frequencies.

Milton, J., V. N. Shankar, and F. Mannering. 2008. "Highway accident severities and the mixed logit model: an exploratory empirical analysis." *Accident Analysis and Prevention* 40, 1: 260–266. Accessed May 15, 2020. https://www.researchgate.net/publication/5635292_Highway_accident_severities_and_the_mixed_logit_model_An_exploratory_empirical_analysis.

MnDOT. 2011. *Otter Tail County Safety Plan: Moving Toward Zero Deaths*. Saint Paul, MN: Minnesota Department of Transportation. Accessed May 15, 2020. https://www.dot.state.mn.us/stateaid/trafficsafety/county/ottertail-crsp-final-aug2011.pdf.

MnDOT. 2020. "County Roadway Safety Plans." Accessed May 12, 2020. https://www.dot.state.mn.us/stateaid/county-roadway-safety-plans.html.

Montgomery, D. C., E. A. Peck, and G. G. Vining. 2012. *Introduction to Linear Regression Analysis*. 5th ed. Hoboken, NJ: Wiley. Accessed May 15, 2020. https://www.wiley.com/en-us/Introduction+to+Linear+Regression+Analysis%2C+5th+Edition-p-9780470542811.

Musunuru, A., R. Wei, and R. J. Porter. 2017. "Predicting Day and Night Traffic Volumes on Rural Roads for Statistical Road Safety Modeling." *Transportation Research Record* 2659: 192–203. Accessed May 15, 2020. https://journals.sagepub.com/doi/abs/10.3141/2659-21.

National Academy of Sciences. 2003a. *NCHRP Report 500: Volume 1: A Guide for Addressing Aggressive-Driving Conditions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v1.pdf.

National Academy of Sciences. 2003b. *NCHRP Report 500: Volume 2: A Guide for Addressing Collissions Involving Unlicensed Drivers and Drivers with Suspended or Revoked Licenses.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v2.pdf.

National Academy of Sciences. 2003c. *NCHRP Report 500: Volume 3: A Guide for Addressing Collisions with Trees in Hazardous Locations.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v3.pdf.

National Academy of Sciences. 2003d. *NCHRP Report 500: Volume 4: A Guide for Addressing Head-On Collisions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v4.pdf.

National Academy of Sciences. 2003e. *NCHRP Report 500: Volume 5: A Guide for Addressing Unsignalized Intersection Collisions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v5.pdf.

National Academy of Sciences. 2003f. *NCHRP Report 500: Volume 6: A Guide for Addressing Run-Off-Road Collisions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v6.pdf.

National Academy of Sciences. 2004a. *NCHRP Report 500: Volume 7: A Guide for Reducing Collisions on Horizontal Curves.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_on_Horizontal_Curv_154782.aspx.

National Academy of Sciences. 2004b. *NCHRP Report 500: Volume 8: A Guide for Reducing Collisions Involving Utility Poles.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. https://www.nap.edu/catalog/23426/a-guide-for-reducing-collisions-involving-utility-poles.

National Academy of Sciences. 2004c. *NCHRP Report 500: Volume 9: A Guide for Reducing Collisions Involving Older Drivers.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_Involving_Older_Dr_154829.aspx.

National Academy of Sciences. 2004d. *NCHRP Report 500: Volume 10: A Guide for Reducing Collisions Involving Pedestrians.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_Involving_Pedestri_154863.aspx.

National Academy of Sciences. 2004e. *NCHRP Report 500: Volume 11: A Guide for Increasing Seatbelt Use.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Increasing_Seatbelt_Use_154846.aspx.

National Academy of Sciences. 2004f. *NCHRP Report 500: Volume 12: A Guide for Reducing Collisions at Signalized Intersections.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_at_Signalized_Inte_154972.aspx.

National Academy of Sciences. 2004g. *NCHRP Report 500: Volume 13: A Guide for Reducing Collisions Involving Heavy Trucks.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_Involving_Heavy_Tr_154900.aspx.

National Academy of Sciences. 2005a. *NCHRP Report 500: Volume 14: A Guide for Reducing Crashes Involving Drowsy and Distracted Drivers.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Crashes_Involving_Drowsy_and_156333.aspx.

National Academy of Sciences. 2005b. *NCHRP Report 500: Volume 15: A Guide for Enhancing Rural Emergency Medical Services.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Enhancing_Rural_Emergency_Medical_Serv_156339.aspx.

National Academy of Sciences. 2005c. *NCHRP Report 500: Volume 16: A Guide for Reducing Alcohol-Related Collisions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_AlcoholRelated_Collisions_156343.aspx.

National Academy of Sciences. 2006. *NCHRP Report 500: Volume 17: A Guide for Reducing Work Zone Collisions.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Work_Zone_Collisions_156711.aspx.

National Academy of Sciences. 2007. *NCHRP Report 500: Volume 19: A Guide for Reducing Collisions Involving Young Drivers.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_Involving_Young_Dr_159494.aspx.

National Academy of Sciences. 2008a. *NCHRP Report 500: Volume 18: A Guide for Reducing Colllisions Involving Bicycles.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_Collisions_Involving_Bicycles_156839.aspx.

National Academy of Sciences. 2008b. *NCHRP Report 500: Volume 20: A Guide for Reducing Head-On Crashes on Freeways.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_HeadOn_Crashes_on_Freeways_159999.aspx.

National Academy of Sciences. 2008c. *NCHRP Report 500: Volume 21: Safety Data Ana Analysis In Developing Emphasis Area Plans.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/Safety_Data_and_Analysis_in_Developing_Emphasis_Ar_160164.aspx.

National Academy of Sciences. 2008d. *NCHRP Report 500: Volume 22: A Guide for Addressing Collisions Involving Motorcycles.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Addressing_Collisions_Involving_Motorc_160626.aspx.

National Academy of Sciences. 2009. *NCHRP Report 500: Volume 23: A Guide for Reducing Speeding-Related Crashes.* Washington, DC: National Academy of Sciences. Accessed May 11, 2020. http://www.trb.org/Publications/Public/Blurbs/A_Guide_for_Reducing_SpeedingRelated_Crashes_160862.aspx.

National Academy of Sciences. 2010. *NCHRP Research Results Digest 345: Alternate Strategies for Safety Improvement Investments*. Washington, DC: National Academy of Sciences.

National Academy of Sciences. n.d. "NCHRP 17-84 [Active]: Pedestrian and Bicycle Safety Performance Functions for the Highway Safety Manual" (website). Accessed July 1, 2019. https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4203.

NEMSIS Technical Assistance Center. 2018. *National EMS Database NEMSIS Public Release Research Data Set V3.3.4 and V3.4.0: 2017 User Manual*. Washington, DC: National Highway Traffic Safety Administration. Accessed January 19, 2018. https://nemsis.org/wp-content/uploads/2019/06/NEMSIS-RDS-340-User-Manual_V4.pdf.

NHTSA. 2008. *National Motor Vehicle Crash Causation Survey Report to Congress*. Report No. DOT HS 811 059. Washington, DC: National Highway Traffic Safety Administration. Accessed January 19, 2018. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811059.

NHTSA. 2017. *MMUCC Guideline: Model Minimum Uniform Crash Criteria*. 5th ed. Washington DC: National Highway Traffic Safety Administration. Accessed May 15, 2020. https://crashstats.nhtsa.dot.gov/#/PublicationList/2.

NHTSA. 2018a. "Fatality Analysis Reporting System (FARS)" (website). Accessed January 19, 2018. https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars.

NHTSA. 2018b. "National Automotive Sampling System (NASS) General Estimates System" (website). Accessed January 19, 2018. https://www.nhtsa.gov/national-automotive-sampling-system-nass/nass-general-estimates-system.

NHTSA. 2018c. "National Automotive Sampling System (NASS) Crashworthiness Data System" (website). Accessed January 19, 2018. https://www.nhtsa.gov/national-automotive-sampling-system-nass/crashworthiness-data-system.

NHTSA. 2018d. "Crash Injury Research" (website). Accessed January 19, 2018. https://www.nhtsa.gov/research-data/crash-injury-research.

NOAA. 2018. “National Oceanic and Atmospheric Administration” (website). Accessed January 19, 2018. https://www.noaa.gov.

Orner, E., and A. Drakopoulos. 2007. *Analysis of Run-Off-Road Crashes in Relation to Roadway Features and Driver Behavior.* Proceedings from the 2007 Mid-Continent Transportation Research Symposium, Ames, Iowa. Accessed May 15, 2020. https://www.semanticscholar.org/paper/Analysis-of-Run-Off-Road-Crashes-in-Relation-to-and-Ornek-Drakopoulos/0a53c6390036db3a671c3132d26e185baba799b7.

Patel. R., F. Council, and M. Griffith. 2007. “Estimating Safety Benefits of Shoulder Rumble Strips on Two-Lane Rural Highways in Minnesota: Empirical Bayes Observational Before-and-After Study.” *Transportation Research Record* 2019: 205–211. Accessed May 15, 2020. https://trid.trb.org/view/801951.

Persaud, B., M. Griffith, C. Hayden, J. Kononov, C. Lewis II, R. Pain, P. M. Salzberg, et al. 2001. *NCHRP Synthesis 295: Statistical Methods in Highway Safety Analysis: A Synthesis of Highway Practice.* Washington, DC: National Cooperative Highway Safety Research Program. Accessed January 19, 2018. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_295.pdf.

Poch, M., and F. Mannering. 1996. “Negative Binomial Analysis of Intersection-Accident Frequencies.” *Journal of Transportation Engineering* 122, 2. Accessed May 15, 2020. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.4343&rep=rep1&type=pdf.

Porter, R., T. Le, F. Gross, D. Carter, T. Saleem, and R. Srinivasan. 2020. *Focus Crash Types and Contributing Factors: Quick Reference Guide*. Report No. FHWA-HRT-20-053. Washington, DC: Federal Highway Administration.

Potts, I. B., D. W. Harwood, D. Bokenkroger, and M. M. Knoshaug. 2011. *Benefit/Cost Evaluation of MODOT’s Total Striping and Delineation Program.* Project No. 110749. Kansas City, KS: MRI Global. Accessed May 15, 2020. https://trid.trb.org/view/1117526.

Preston, H. 2012. “A Systemic Safety Project Identification Process – Minnesota’s County Road Safety Plans.” Presented at the *South Dakota Transportation Safety Conference, April 2012*.

Preston, H., and M. Gute. 2010. “Mn/DOT County Road Safety Plans (CRSP).” *Inciter* (fall 2010): 12–13. Accessed January 19, 2018. http://www.dot.state.mn.us/stateaid/trafficsafety/county/inciter-2010-fall-crsp.pdf.

Preston, H., R. Storm, J. Bennett, and B. Wemple. 2013a. *Systemic Safety Project Selection Tool*. Report No. FHWA-SA-13-019. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://safety.fhwa.dot.gov/systemic/fhwasa13019.

Preston, H., R. Storm, K. Scurry, and B. Wemple. 2013b. “Using Risk to Drive Safety Investments.” *Public Roads* 76, 6. Accessed May 15, 2020. https://www.fhwa.dot.gov/publications/publicroads/13mayjun/03.cfm.

Rencher, A. C., and W. F. Christensen. 2012. *Methods of Multivariate Analysis*. 3rd ed. Hoboken, NJ: John Wiley & Sons. Accessed May 15, 2020. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118391686.

Rossi, A., Amaddeo, F., Sandri, M., and Tansella, M. 2005. "Determinants of once-only contact in a community-based psychiatric service." *Social Psychiatry and Psychiatric Epidemiology* 40, 1: 50–56. Accessed May 15, 2020. https://www.ncbi.nlm.nih.gov/pubmed/15624075.

Segal, M. R. 2004. *Machine Learning Benchmarks and Random Forest Regression*. San Francisco, CA: Center for Bioinformatics and Molecular Biostatistics.

Shankar, V., F. Mannering, and W. Barfield. 1995. "Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies." *Accident Analysis and Prevention* 27, 3: 371–389. Accessed May 15, 2020. https://www.sciencedirect.com/science/article/abs/pii/000145759400078Z.

Smiley, A. 2008. "Point of View: Driver Behavior: A Moving Target." *TR News* 254. Accessed May 15, 2020. https://trid.trb.org/view/850893.

Strobl, C., J. Malley, and G. Tutz. (2009). *An Introduction to recursive partitioning*. Report No. 55. Munich, Germany: University of Munich.

Tarko, A. P., L. Mingyang, M. Romero, and J. Thomaz. 2014. *A Systematic Approach to Identifying Traffic Safety Needs and Intervention Programs for Indiana: Volume I—Research Report*. Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2014/03. West Lafayette, IN: Purdue University. Accessed May 15, 2020. https://docs.lib.purdue.edu/jtrp/1556/.

Thomas, L., B. Lan, R. L. Sanders, A. Frackelton, S. Gardner, and M. Hintze. 2017. "Changing the future? Development and application of pedestrian safety performance functions to prioritize locations in Seattle, Washington." *Transportation Research Record* 2659: 212–223. Accessed May 15, 2020. https://journals.sagepub.com/doi/10.3141/2659-23.

Thomas, L., W. Kumfer, K. Lang, C. Zeeger, L. Sandt, B. Lan, K. Nordback, et al. 2018. *Systemic Pedestrian Safety Analysis: Contractor's Technical Report*. Project No. 17-27. Washington, DC: National Cooperative Highway Safety Research Program. Accessed July 1, 2019. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_893_Contractor.pdf.

usRAP. 2020. "What is usRAP?" (website). Accessed May 11, 2020. http://www.usrap.org/what-usrap.

U.S. Census Bureau. 2018. "Socioeconomic Census Data" (website). Accessed January 19, 2018. https://www.census.gov/data.html.

Virginia Tech Transportation Institute. 2018. "InSight Data Access Website SHRP2 Naturalistic Driving Study" (website). Accessed January 19, 2018. https://insight.shrp2nds.us.

Waller, L. A., and C. A. Gotway. 2004. *Applied Spatial Statistics for Public Health Data*. Hoboken, NJ: John Wiley & Sons. Accessed May 15, 2020. https://onlinelibrary.wiley.com/doi/book/10.1002/0471662682.

Wang, J., W. Hughes, and R. Stewart. 1998. *Safety Effects of Cross-Section Design on Rural Multi-Lane Highways*. Report No. FHWA-RD-98-071. Washington, DC: Federal Highway Administration.

Wang, Y., S. Sharda, and H. Wang. 2016. "A Systemic Safety Analysis of Pedestrian Crashes: Lessons Learned." Presented at *95th Annual Meeting of the Transportation Research Board. Washington, DC*. Accessed May 15, 2020. https://trid.trb.org/view/1393638.

Weisberg, S. 2014. *Applied Linear Regression*. 4th ed. Hoboken, NJ: Wiley. Accessed May 15, 2020. https://www.wiley.com/en-us/Applied+Linear+Regression%2C+4th+Edition-p-9781118386088.

Wilder, R. D. 2011. *Centerline Rumble Strips on Secondary Highways – A Systemic Crash Analysis.* New York, NY: New York State Department of Transportation. Accessed January 19, 2018. https://www.dot.ny.gov/programs/rumblestrips/repository/Centerline%20Rumble%20Strips%20on%20Secondary%20Highways.pdf.

Zegeer, C. V. 1986. "Methods for identifying hazardous highway elements." *NCHRP Synthesis of Highway Practice* 128. Washington, DC: Transportation Research Board. Accessed May 15, 2020. http://www.trb.org/Publications/Blurbs/154502.aspx.

Zegeer, C., J. Hummer, D. Reinfurt, L. Herf, and W. Hunter. 1987. *Safety Effects of Cross-Section Design for Two-Lane Roads – Volume I – Final Report*. Report No. FHWA-RD-87/008. Washington, DC: Federal Highway Administration. Accessed May 15, 2020. https://trid.trb.org/view/274947.

Recycled
Recyclable

HRDS-20/11-20(300)E