Horizontal Curve Safety Performance Evaluation Based on the Naturalistic Driving Study Lane Position Data

PUBLICATION NO. FHWA-HRT-24-093

MAY 2024

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 authors conducted this research under Transportation Pooled Fund (TPF) study TPF-5(361), SHRP 2 [Second Strategic Highway Research Program] Naturalistic Driving Study [NDS] Pooled Fund: Advancing Implementable Solutions (Federal Highway Administration (FHWA) 2023a). The TPF study aimed to develop novel, multidisciplinary solutions based on the recorded natural behavior of vehicle operators interacting with infrastructure and other vehicles. Performance-based analysis is an approach that specifically addresses the purpose and needs of projects by providing design flexibility (FHWA 2017). Resources to evaluate roadway design based on performance are not currently available for transportation practitioners. Observed crash data alone does not provide sufficient information to quantify the effect of projects that were implemented based on PBD. This research project used the Roadway Information Database and the SHRP2 NDS data to estimate the safety effect of elements that influence driving behavior on rural undivided two-lane horizontal curves (Iowa State University 2024; Virginia Tech Transportation Institute 2023). The researchers evaluated and associated safety surrogates such as kinematics, speed, and lane position with observed crashes.

The researchers used the findings of this research study to develop an analytical tool for practitioners to use when considering curve design (University of Wisconsin–Madison 2024). The tool estimates crashes and associated crash costs to perform an economic assessment of curve design implementation. The economic assessment provides a quantitative measure for practitioners to evaluate implementing PBD alternatives by assessing the tradeoffs between safety and implementation costs. This research will interest roadway designers, safety professionals, and others interested in mitigating disruptions to vehicle flow through complex freeway interchanges.

Carl Andersen
Acting Director, Office of Safety and Operations
Research and Development

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation (USDOT) in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

Non-Binding Contents

Except for the statutes and regulations cited, the contents of this document do not have the force and effect of law and are not meant to bind the States or the public in any way. This document is intended only to provide information regarding existing requirements under the law or agency policies.

Quality Assurance Statement

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

Disclaimer for Product Names and Manufacturers

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this document only because they are considered essential to the objective of the document. They are included for informational purposes only and are not intended to reflect a preference, approval, or endorsement of any one product or entity.

TECHNICAL REPORT DOCUMENT PAGE

1. Report No. FHWA-HRT-24-093	2. Government Accession No.	3. Recipient’s Catalog No.

4. Title and Subtitle Horizontal Curve Safety Performance Evaluation Based on the Naturalistic Driving Study Lane Position Data	5. Report Date May 2024 6. Performing Organization Code
7. Authors Boris Claros (ORCID: 0000-0002-4787-1749), Madhav Chitturi (ORCID: 0000-0003-0580-3454), Glenn Vorhes (ORCID: 0000-0001-5256-9942) Andrea Bill (ORCID:0000-0001-9107-6745), and David A. Noyce (ORCID: 0000-0001-5887-8391)	8. Performing Organization Report No.
9. Performing Organization Name and Address Traffic Operations and Safety Laboratory University of Wisconsin–Madison 1415 Engineering Drive Madison, WI 53706	10. Purchase Order No. 11. Contract or Grant No. DTFH6116D00038
12. Sponsoring Agency Name and Address Office of Safety and Operations Research and Development Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296	13. Type of Report and Period Covered Final Report; April 2021–November 2023. 14. Sponsoring Agency Code HRSO-20

15. Supplementary Notes
The contracting officer’s representative was Charles Fay (HRSO-20; ORCID: 0000-0002-2498-9795).

16. Abstract

Geometric design based on performance is a renewed approach to project decisionmaking that specifically addresses the purpose and needs of projects by providing design flexibility. Transportation practitioners need resources to evaluate roadway design based on safety performance. This research project used the Roadway Information Database and the Second Strategic Highway Research Program Naturalistic Driving Study data to estimate the safety effect of elements that influence driving behavior on rural, undivided, two-lane horizontal curves (Iowa State University 2024, Virginia Tech Transportation Institute 2023). Available data included 3,292 curves and 150,233 traversals, which required a significant data processing effort to conduct data cleaning and quality assessment. From the safety surrogates evaluated and methods implemented, lane position provided the most consistent and statistically significant results. The researchers modeled centerline and edge-line encroachment events with the negative binomial regression modeling approach, using curve geometry as the predictor component. Encroachment estimates were associated with observed crashes from State data to convert encroachments to crashes. Predictor variables, such as curve radius, showed a decreasing trend in predicted crashes as the curve radius increased. Similarly, as shoulder or lane width increased, predicted crashes decreased. The researchers used crash estimates derived from safety surrogates to develop an analytical tool aimed at practitioners designing curves (University of Wisconsin–Madison 2024). Data input in the tool includes curve radius, shoulder width, lane width, length, annual average daily traffic, construction cost, and expected service life. The economic assessment provides a quantitative measure for practitioners to evaluate implementing alternative curve designs by assessing the tradeoffs between safety and implementation cost.

17. Key Words Performance-based design, geometric design, performance-based analysis, naturalistic driving, lane position, horizontal curve	18. Distribution Statement No restriction. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. https://www.ntis.gov

19. Security Classification (this report) Unclassified	20. Security Classification (this page) Unclassified	21. No. of pages 37	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

CHAPTER 1. INTRODUCTION **1**

CHAPTER 2. LITERATURE REVIEW **3**

CHAPTER 3. METHODOLOGY **5**

- **Curve and Traversal Data** **5**
- **Data Processing** **6**
  - Data Variables 6
  - Curve Geometry Review 7
  - Presence of Lead Vehicle 8
  - Presence of Rumble Strips 8
  - Exploratory Data Analysis 8
- **Data Analysis** **10**
  - Data Preparation 10
  - Safety Surrogate Measures 15
- **Modeling** **15**
  - Negative Binomial Regression Modeling 18
  - Association of Lane Departures and Observed Crashes 18
- **Tool Development** **18**
  - Data Input 19
  - Safety Analysis 19
  - Economic Assessment 20

CHAPTER 4. RESULTS **21**

- **Encroachment Models** **21**
- **Encroachments and Observed Crashes Association** **22**
- **Analytical Tool** **25**
  - Instructions 25
  - Automated Calculations 26

CHAPTER 5. CONCLUSIONS **29**

REFERENCES **31**

LIST OF FIGURES

Figure 1. Illustration. Curve and traversal data available. 6
Figure 2. Illustration. Direction of curves. 11
Figure 3. Illustration. NDS lane position (VTTI 2023). 16
Figure 4. Illustration. Encroachment thresholds. 17
Figure 5. Illustration. Encroachment and crash association. 24
Figure 6. Illustration. Observed crashes on existing curves. 26
Figure 7. Illustration. Diagram of simple curve geometric elements. 26
Figure 8. Illustration. Analytical section of the tool with safety and economic estimates. 27

LIST OF TABLES

Table 1. Summary of data processing. 6
Table 2. Count of all curves by curve radius and shoulder width. 10
Table 3. Count of all traversals by curve radius and shoulder width. 11
Table 4. Number of curves by curve radius, shoulder width, and speed limit/advisory speed for left-turning curves (−1). 12
Table 5. Number of traversals by curve radius, shoulder width, and speed limit/advisory speed for left-turning curves (−1). 12
Table 6. Number of curves by curve radius, shoulder width, and speed limit/advisory speed for right-turning curves (+1). 13
Table 7. Number of traversals by curve radius, shoulder width, and speed limit/advisory speed for right-turning curves (+1). 14
Table 8. HSM severity distribution of rural, two-lane roadways (AASHTO 2014). 19
Table 9. National KABCO crash unit costs in 2010 dollars (Blincoe et al. 2015; Harmon, Bahar, and Gross 2018). 20
Table 10. Encroachment model coefficients for curves with an advisory speed or speed limit <50 mph. 22
Table 11. Encroachment model coefficients for curves with an advisory speed or speed limit ≥50 mph. 22
Table 12. RID curves with AADT and State crash data available. 23
Table 13. Results of encroachments and observed crashes association. 25

LIST OF ABBREVIATIONS AND SYMBOLS

$1/k$	dispersion term
A	incapacitating injury
AADT	annual average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
a–e	model coefficients
B	nonincapacitating injury
B/C	benefit-cost ratio
C	possible injury
D	encroachment rate
E	encroachments
HSM	*Highway Safety Manual*
K	fatal
KABCO	injury classification scale
LEN	curve length
LNW	lane width
m	slope
N	observed crash rate
NDS	Naturalistic Driving Study
O	property damage only
PC	point of curve
PI	point of intersection
PT	point of tangency
QALY	quality-adjusted life years
R^2	coefficient off determination
RAD	curve radius
RID	Roadway Information Database
S_1	threshold between centerline and longitudinal axis of vehicle
S_2	threshold between the edge line and longitudinal axis of vehicle
SHRP2	Second Strategic Highway Research Program
SHW	shoulder width
SV	single vehicle
TRAV	number of traversals
TPF	Transportation Pooled Fund
VTTI	Virginia Tech Transportation Institute
YR	period of analysis (years)
Δ	curve central angle

CHAPTER 1. INTRODUCTION

Geometric design based on performance is a renewed approach to project decisionmaking that specifically addresses the purpose and needs of projects by providing design flexibility. Designs may be implemented for specific long- and short-term performance project goals, a whole corridor, or the overall system. This approach is not limited to geometric design elements but spans from project planning decisionmaking to identifying the type of facility, selecting design volumes and speed, and even selecting pavement materials and thickness. However, no project should experience safety or operation performance lower than acceptable performance measures nor affect mandates for people with disabilities or environmental requirements.

Performance-based analysis has mainly focused on crash data, expected crashes, or microsimulation. However, some limitations regarding the findings of these results exist despite the use of rigorous statistical methods and software. Observed crash data alone do not provide sufficient information to quantify the effect of projects that were implemented based on a performance metric. Conventional safety analysis may not confidently estimate the effect of the design approach because of limited data.

As part of the Transportation Pooled Fund (TPF) study TPF-5(361), SHRP 2 [Second Strategic Highway Research Program] Naturalistic Driving Study [NDS] Pooled Fund: Advancing Implementable Solutions, the research team used the Roadway Information Database (RID) and the SHRP2 NDS data to estimate the safety effect of roadway geometric elements that directly influence driving behavior related to kinematics, especially on a curved roadway segment (Federal Highway Administration (FHWA) 2023a; Iowa State University 2024; Virginia Tech Transportation Institute (VTTI) 2023). In the absence of limited crash data, safety surrogates are an alternative approach with promising results, and the RID and SHRP2 NDS data are ideal sources of information for safety surrogates since RID and SHRP2 NDS provide alternative and realistic data that can be associated with safety measures. The research team used safety surrogates obtained from the SHRP2 NDS Trips and Event data, curves identified in the RID, and State crash data from the RID to model the safety impacts of curve geometric elements on rural, undivided, two-lane roadways and developed a tool to perform safety and economic assessments for different design alternatives University of Wisconsin–Madison 2024).

CHAPTER 2. LITERATURE REVIEW

The literature review focused on past studies that used naturalistic driving data to assess safety at curves on rural, two-lane highways. Using SHRP2 NDS data, Wang et al. (2018) evaluated operational speed on rural two-lane curves with nearly 10,000 vehicle traversals from 202 drivers on 219 horizontal curves. A log-linear relationship was found between curve radius and mean vehicle speeds. Speeds were relatively stable on curves with radii of 900–1,000 ft or more, decreasing more rapidly as the radius decreased below this range. Drivers reduced speeds when curve speed advisories were present, but the magnitude of reductions was much less than suggested by the advisory signs. Speeds were significantly lower when a W1-6 one-directional curve arrow sign was present (Wang et al. 2018). W1-6 signs may be used in place of or to supplement delineators or chevron alignment signs to indicate a change in horizontal alignment (FHWA 2023b).

Hallmark et al. (2015) evaluated driving behavior on rural two-lane curves using environmental and traffic characteristics reduced from SHRP2 NDS forward roadway video. Driver glance location and distraction data were obtained from the driver and over-the-shoulder videos. Hallmark et al. conducted an assessment on the probability (odds) of a given type of encroachment based on driver, roadway, and environmental characteristics. The results showed that the probability of a right-side encroachment increases as drivers spend less time glancing at the forward roadway. Also, the probability of 5 mph or more over the posted or advisory speed was higher with higher average speed upstream, younger drivers, and when edge-line markings were obscured or not present.

In terms of safety, crashes and near crashes available from SHRP2 NDS were evaluated (Wang et al. 2017, Wu et al. 2017). Some of the data elements evaluated included driver demographic characteristics, traffic environment, roadway, geometry, and driving behavior such as distractions. Wang et al. (2017) used logistic regression modeling to analyze factors affecting the likelihood of a driver being involved in a crash. Results of the evaluations showed that crash risk was three times higher when drivers were visually distracted. Wu et al. (2019) associated curve safety and operations on rural, two-lane roads. The researchers in that study evaluated the severity and rate of crashes, integrating 8 yr of crash records and associated geometric features for analysis. The results indicated that higher curve severity categories (related to side friction demand and tolerance) are positively correlated with crash rate.

CHAPTER 3. METHODOLOGY

Quantifying the effect of design practices on safety is a complex task. Identifying locations with homogeneous treatments and similar roadway characteristics where the performance-based approach was implemented is difficult, so limited data are available to conduct before and after observational studies. Implementing cross-sectional modeling with observed crashes is challenging because the number of observed crashes on rural roadways is small, and facilities have a wide variety of roadway characteristics (heterogeneity).

Implementing safety surrogate analysis has become a viable alternative when crash occurrence is limited or capturing the effect of a particular roadway feature is not possible. Tracking vehicle trajectories and driver operations provides valuable information to analyze events in which vehicle operations, proximity, and location measures help to identify unsafe events that may result in crashes.

After the exhaustive effort of data processing (described in the following sections), the researchers analyzed naturalistic traversal data, such as vehicle speed, deceleration, or steering; however, measures associated with lane position had the most consistent and statistically significant results in terms of curve geometry. Negative binomial regression was implemented to model centerline and edge-line encroachment frequency as a function of curve geometry. The researchers compared safety surrogate estimates of encroachments combined with traffic volumes to actual crashes (from State data) to find the degree of association and estimate potential crashes. A tool was developed using estimates of potential crashes based on encroachments to compare the performance of different curve design configurations in terms of crash benefit-cost (University of Wisconsin–Madison 2024).

CURVE AND TRAVERSAL DATA

From the data available in the RID, the researchers selected curves for rural, two-lane roads for evaluation (Iowa State University 2024). Figure 1 provides a breakdown of available curves and the different parameters that were considered. Overall, 43,296 curves were identified in the RID. Based on location, 20,945 curves were in urban areas, and 22,351 curves were in rural areas. In rural areas, 17,230 curves were undivided, 18,644 curves had two lanes, and 10,195 curves had grades between −1 and +1 percent. As a result, the researchers identified 6,269 curves on rural, undivided, two-lane highways with grades between −1 and +1 percent. Traversal data were requested for the curves of interest, and 3,292 curves (with a radius between 250 and 3,000 ft) with 150,233 traversals were available.

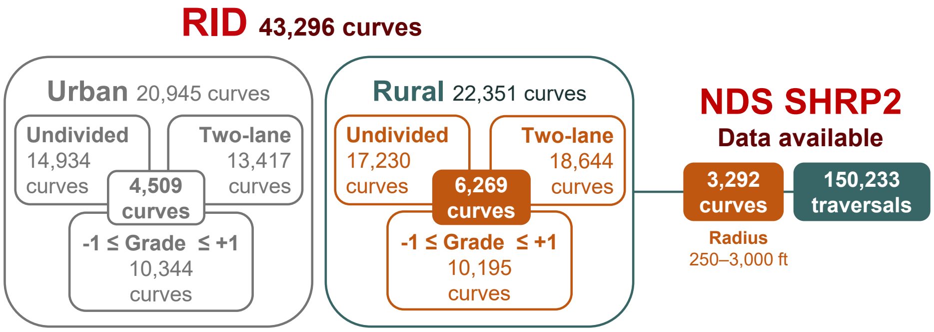


Figure 1. Illustration. Curve and traversal data available.

DATA PROCESSING

Data processing consisted of data review and quality assessment. The researchers took several steps to guarantee the quality and validity of the data. All curves were manually reviewed to confirm that curves were on undivided, two-lane segments. Traversals were processed to check for lead vehicles, rumble strips, missing data, duplicate data, low probability data quality, outliers, and overrepresentation of observations. Additionally, the researchers reviewed 455 videos to validate observations for specific traversals of interest in which the data indicated lane departure. Table 1 outlines the data cleaning process and how the number of curves and traversals available for modeling and analysis were reduced at each step of data processing. The following sections will describe each step of data processing in more detail.

Table 1. Summary of data processing.

Description	Removed		Remain	
	Curves	Traversals	Curves	Traversals
Data request	0	0	3,292	150,233
Review and validation of curves	355	16,371	2,937	133,862
Presence of lead vehicle	42	28,652	2,895	105,210
Presence of rumble strips	429	7,892	2,466	97,318
Exploratory data analysis	436	34,904	**2,030**	**62,414**

Data Variables

The RID data included geometric, operational, and crash information (Iowa State University 2024). Traversal data were obtained from 600 ft upstream of the point of curvature, through the curve, and 600 ft downstream of the point of tangency. Some differences in data availability and format existed since the RID contains roadway information from multiple States: Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington. The following variables were available for the curves in this study:

CurveID
Tangent
Radius
CurveDirectionID
SuperElevation
Grade
CurveLength
OppositeCurve
CurveWarning
Chevrons
CurveArrow
AdvisorySpeed

SpeedLimit
BiDirectionalLanes
AverageLaneWidth
TotalWidth
ThroughLanes
LeftTurnLane
RightTurnLane
ShoulderType
ShoulderWidth
Rumble
MedianType
BarrierType
Intersection
Lighting
Direction
RouteID
FrMeasure
ToMeasure
Length
FrX
FrY
ToX
ToY
CountTotal
CountCurve
CountIntersection
Count2006
Count2007
Count2008
Count2009
Count2010
Count2011
Count2012
Count2013
MaxSeverity
TotalInjury
NumFatalaties
NumSevInjury
NumMinorInjury
NumPossInjury
Shape_Length
AADT (annual average daily traffic)

The researchers requested traversals on identified curves in the RID for the study with kinematic information and lane position. The following information was available for the traversals of interest (Iowa State University 2024):

speed_network
steering_wheel_position
light_level
pedal_gas
accel_x
decel_x
accel_y
decel_y
accel_z
decel_z
gyro_x
gyro_y
gyro_z
lane_distance_off_center
lane_width
left_line_right
right_line_left
left_marker
right_marker
pedal_brake
lead_vehicle_headway

Curve Geometry Review

The researchers identified 3,292 curves from the RID that were classified as rural two-lane, undivided horizontal curves (Iowa State University 2024). All curves were manually reviewed to validate location and geometric data. During the review process, 355 curves were removed from the dataset, which resulted in removing 16,371 traversals. The curves were removed for several reasons:

- Presence of intersection approach with negative left-turn offset.
- Divided multilane roadway.
- Divided two-lane roadway.
- Undivided multilane roadway.
- Length of curves (<150 or >2,000 ft).
- Radius of curves (<250 or >3,000 ft).
- Vertical grade (<−1 or >+1 percent).
- Paved shoulder width (>8 ft).
- Roundabout.

- Curves on ramps.
- Curves not located in rural settings.
- Presence of right-side parking or sidewalk.
- Presence of all-way stop or traffic signal control intersection along curve.
- Inconsistent geometry along curve (change of number of lanes, median, or taper).

Presence of Lead Vehicle

The researchers evaluated safety surrogates along horizontal curves under free-flow conditions, so traversals with leading vehicles were removed. The NDS traversal data have the variable of time headway with respect to a leading vehicle (VTTI 2023). Since the traversal data are in a time-series format, several time headway observations were made along the curve if a lead vehicle was present. Thus, the researchers computed the average time headway of all the observations along the curves and the number of observations with a time headway measurement. The average time headway indicated the average proximity of the leading vehicle and the number of observations with a time headway measurement, which helped identify the proportion of observations along the curve where a lead vehicle was present. If the average time headway was less than 4.0 s and the number of observations with a recorded time headway measurement along the curve was greater than half of all observations (>50 percent of all observations had the presence of a leading vehicle), the traversal was considered to have a leading vehicle. As a result, the researchers removed 28,652 traversals and 42 curves.

Presence of Rumble Strips

The researchers also removed curves with rumble strips from the dataset for modeling and analysis because not enough representative data existed to capture the effect of rumble strips. Since the rumble strip information was integrated in the RID database, 429 curves with rumble strips either on the centerline, the shoulder, or both were removed, which resulted in the removal of 7,892 traversals.

Exploratory Data Analysis

After removing several curves and corresponding traversals, the researchers explored different variables in the traversal data for modeling. The exploratory data analysis consisted of evaluating the standard data metrics, such as the number of observations along the curve, mean, standard deviation, minimum, and maximum. Based on the results of the exploratory analysis, the researchers conducted several checks to identify traversals with a significant number of missing or unique observations in the time series, extreme observations, and the number of traversals for each curve. A total of 436 curves and 34,904 associated traversals were removed from the dataset.

Probability of Data Accuracy

Lane position available in the NDS data was obtained through machine-vision evaluation and processing. The accuracy of each lane position measurement is indicated by an associated probability with a range of 1–1,024 (VTTI 2023). Based on the recommendations of VTTI and previous research, the researchers deemed unreliable measurements with an associated probability lower than 512. Thus, traversals with lane position measurements with probability less than 512 were removed.

Missing Data

Traversal data are in a time series format, meaning they contain measurements of location, operations, and lane position in short intervals and recorded with timestamps when the vehicles were traversing curves. During the exploratory data analysis, the researchers observed that, although several traversals had many intervals where timestamps were recorded, no operations or lane position data were available for some of the timestamps. For instance, when network speed measurements were evaluated, some timestamps were missing data. Missing data reduced the number of observations and accuracy of aggregate measurements through the curve. Missing data may be attributed to data collection and processing of information or loss of connectivity among tracking or measurement devices. The researchers counted timestamps with missing measurements and compared them to the overall number of observations to address traversals with a significant number of missing observations. If the number of missing observations was greater than half (>50 percent of all observations), the researchers dropped the traversals and associated curves. Thresholds were selected based on the distribution of the proportion of missing data.

Unique Observations

The researchers used a process similar to that used for missing data to evaluate timestamps with measurements with the same repeated data. If timestamps provided the same measurements when traversing a curve, data recorded were not considered reliable, and repeated observations may be attributed to data collection and processing of information or loss of connectivity among tracking or measurement devices. If timestamps with repeated measurements were greater than 25 percent of all observations, traversals and associated curves were removed to address repeated observations.

Outliers

As part of the exploratory data analysis, the researchers evaluated observations that significantly differed from the overall distribution of observations. Based on the characteristics of the data, thresholds were selected and observations beyond those thresholds were further reviewed and validated with performance metrics or video data. For instance, in the case of speed measurements, the speed limit and curve radius were used as a reference to estimate the highest speed that could possibly be reached through a curve. For lane position measurements, the researchers requested videos for extreme situations in which the data indicated lane departure. The researchers reviewed a total of 455 videos and validated or removed lane position measurements if the video contradicted the lane position data measurement.

Overrepresentation of Traversals

The number of traversals available for each curve varied widely. The number of traversals available was an important aspect for modeling and analysis in this project, so the researchers assessed the effect of curves with a significant number of traversals. The researchers partitioned data by bins according to speed limit and advisory speed, curve radius, and paved shoulder width. For each bin, empirical probability density functions and histograms were plotted to visually identify the distribution of predictor variables such as speed and lane position. If probability density functions and histograms displayed potential issues with skewness and inconsistent distribution (mixed distribution), data in each bin were reviewed at the curve level to identify curves that could possibly be skewing the data or introducing a predominant distribution into the overall bin data. If the overrepresented traversals of a curve in a bin were identified to skew and bias the distribution of the bin data, the overrepresented curve data were completely removed, or a sample of overrepresented traversals was randomly selected and kept in the bin. The decision to remove or keep randomly selected traversals was based on the degree of the effect of overrepresented samples and the number of curves and traversals in each bin.

DATA ANALYSIS

Through exploratory data analysis, the researchers narrowed down the data of interest to 2,030 curves and 62,414 traversals. Data analysis consisted of evaluating predictor variables that could provide safety surrogate measures in relation to curve geometry and vehicle operations. Data analysis and preparation for modeling consisted of establishing bins by speed limit and advisory speed, curve direction, curve radius, and paved shoulder width.

Data Preparation

Based on data availability for the number of curves and traversals, the researchers prepared data for probability distribution modeling, which required data with specific attributes that would contribute to capturing the effect of curve geometric traits, such as curve radius and paved shoulder width. Table 2 summarizes all 2,030 curves according to curve radius and shoulder width available for the study. Similarly, table 3 provides a summary of the corresponding traversals available for all curves according to curve radius and shoulder width.

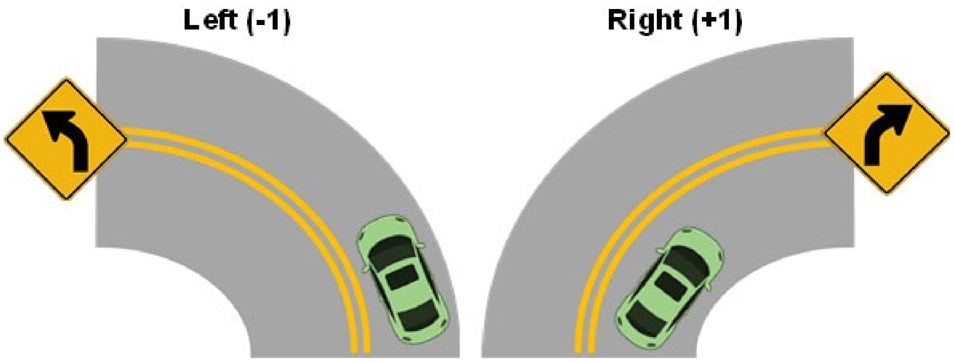
Table 2. Count of all curves by curve radius and shoulder width.

Curve Radius (ft)	Paved Shoulder Width (ft)									
	0	1	2	3	4	5	6	7	8	All
250–499	22	45	23	11	9	3	5	3	2	**123**
500–749	61	86	63	28	7	15	11	1	2	**274**
750–999	30	93	101	36	17	14	4	5	3	**303**
1,000–1,249	28	106	87	21	9	12	11	9	4	**287**
1,250–1,499	22	73	65	32	19	9	15	10	13	**258**
1,500–1,749	12	49	31	23	10	7	4	7	6	**149**
1,750–1,999	17	44	47	37	18	12	9	6	6	**196**
2,000–2,249	10	44	30	10	10	6	7	4	6	**127**
2,250–2,499	7	17	33	18	9	7	10	5	4	**110**
2,500–2,749	9	23	18	6	9	6	5	6	2	**84**
2,750–3,000	12	32	29	12	9	9	3	9	4	**119**
All	**230**	**612**	**527**	**234**	**126**	**100**	**84**	**65**	**52**	**2,030**

Table 3. Count of all traversals by curve radius and shoulder width.

Curve Radius (ft)	Paved Shoulder Width (ft)									
	0	1	2	3	4	5	6	7	8	All
250–499	471	1,541	523	148	139	182	106	61	22	**3,193**
500–749	1,124	4,540	2,372	831	157	425	511	17	9	**9,986**
750–999	288	2,472	3,018	725	595	534	85	334	35	**8,086**
1,000–1,249	914	4,365	3,505	476	85	224	115	715	298	**10,697**
1,250–1,499	1,048	2,463	2,293	823	370	247	359	436	1,096	**9,135**
1,500–1,749	171	1,684	2,825	418	420	75	15	200	109	**5,917**
1,750–1,999	550	939	818	638	279	175	118	127	67	**3,711**
2,000–2,249	337	1,576	641	149	257	319	198	66	683	**4,226**
2,250–2,499	349	932	514	211	183	62	748	35	168	**3,202**
2,500–2,749	35	1,297	370	78	75	131	111	24	40	**2,161**
2,750–3,000	185	669	374	43	82	456	34	230	27	**2,100**
All	**5,472**	**22,478**	**17,253**	**4,540**	**2,642**	**2,830**	**2,400**	**2,245**	**2,554**	**62,414**

Using the available data, the researchers introduced additional classifications based on the direction of curves and the speed limit or advisory speed. Traversals were defined based on the direction of travel at curves since vehicle operations are different according to the curve direction. Two categories were defined for curve direction: left (−1) and right (+1). The left direction (−1) indicated that the vehicle traversing on the right side of the roadway had a left-turning curve direction. The right direction (+1) indicated that the vehicle traversing on the right side of the roadway had a right-turning curve direction. Figure 2 illustrates curve directions.



© 2023 University of Wisconsin–Madison.

Figure 2. Illustration. Direction of curves.

An additional classification was introduced for speed limit/advisory speed. If an advisory speed was available, the curve was designated with that speed; otherwise, the speed limit was assigned. Based on the speed limit or advisory speed, the researchers classified curves into two groups: <50 mph and ≥50 mph. According to the breakdown of curves and traversals by the direction of travel and speed limit/advisory speed, table 4 provides the number of curves for left-turning curves (−1), and table 5 provides the number of corresponding traversals.

Table 4. Number of curves by curve radius, shoulder width, and speed limit/advisory speed for left-turning curves (−1).

Advisory or Speed Limit (mph)	Curve Radius (ft)	Paved Shoulder Width (ft)									
		0	1	2	3	4	5	6	7	8	All
<50	250–499	9	22	12	1	4	2	4	1	2	**57**
	500–749	29	45	31	16	3	7	2	1	1	**135**
	750–999	14	39	38	18	4	7	3	2	1	**126**
	1,000–1,249	9	39	31	7	2	4	4	3	1	**100**
	1,250–1,499	13	23	15	11	2	3	6	5	2	**80**
	1,500–1,749	4	16	4	5	3	2	1	0	1	**36**
	1,750–1,999	10	12	12	9	4	1	3	2	3	**56**
	2,000–2,249	5	10	6	3	2	0	3	2	2	**33**
	2,250–2,499	3	6	12	1	2	1	4	0	0	**29**
	2,500–2,749	5	8	8	2	2	2	2	1	1	**31**
	2,750–3,000	7	8	1	1	3	0	1	1	1	**23**
	All with <50 mph	**108**	**228**	**170**	**74**	**31**	**29**	**33**	**18**	**15**	**706**
≥50	250–499	0	0	0	0	0	0	0	0	0	**0**
	500–749	0	0	0	0	0	0	0	0	0	**0**
	750–999	0	7	10	2	0	1	0	0	0	**20**
	1,000–1,249	0	9	12	2	1	2	2	1	1	**30**
	1,250–1,499	0	14	11	6	6	2	0	0	5	**44**
	1,500–1,749	0	6	6	7	2	0	1	3	4	**29**
	1,750–1,999	0	8	13	11	3	2	3	2	1	**43**
	2,000–2,249	0	8	7	1	3	1	0	0	1	**21**
	2,250–2,499	0	3	4	7	2	2	1	1	0	**20**
	2,500–2,749	0	3	4	2	3	2	1	3	0	**18**
	2,750–3,000	0	4	15	4	3	2	0	3	1	**32**
	All with ≥50 mph	**0**	**62**	**82**	**42**	**23**	**14**	**8**	**13**	**13**	**257**
All	**All**	**108**	**290**	**252**	**116**	**54**	**43**	**41**	**31**	**28**	**963**

Table 5. Number of traversals by curve radius, shoulder width, and speed limit/advisory speed for left-turning curves (−1).

Advisory or Speed Limit (mph)	Curve Radius (ft)	Paved Shoulder Width (ft)									
		0	1	2	3	4	5	6	7	8	All
<50	250–499	75	861	297	2	26	67	72	42	22	**1,464**
	500–749	624	2,405	980	392	22	190	52	17	1	**4,683**
	750–999	130	839	1,132	512	116	142	43	39	7	**2,960**
	1,000–1,249	313	1,034	1,077	50	17	52	22	417	26	**3,008**
	1,250–1,499	735	468	471	573	48	72	76	306	85	**2,834**
	1,500–1,749	25	519	243	67	71	9	3	0	6	**943**
	1,750–1,999	298	130	303	174	15	71	91	34	42	**1,158**
	2,000–2,249	54	212	62	52	52	0	57	8	58	**555**
	2,250–2,499	296	462	73	4	6	1	355	0	0	**1,197**
	2,500–2,749	9	616	211	26	12	106	10	9	29	**1,028**

Advisory or Speed Limit (mph)	**Curve Radius (ft)**	**Paved Shoulder Width (ft)**									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**All**
	2,750–3,000	134	71	41	8	13	0	8	16	9	**300**
	All with <50 mph	**2,693**	**7,617**	**4,890**	**1,860**	**398**	**710**	**789**	**888**	**285**	**20,130**
≥50	250–499	0	0	0	0	0	0	0	0	0	**0**
	500–749	0	0	0	0	0	0	0	0	0	**0**
	750–999	0	379	135	6	0	114	0	0	0	**634**
	1,000–1,249	0	665	560	136	4	7	81	4	20	**1,477**
	1,250–1,499	0	599	602	83	91	7	0	0	51	**1,433**
	1,500–1,749	0	297	740	147	1	0	6	91	92	**1,374**
	1,750–1,999	0	212	288	191	67	26	19	2	1	**806**
	2,000–2,249	0	158	243	1	102	15	0	0	1	**520**
	2,250–2,499	0	6	74	96	107	3	6	1	0	**293**
	2,500–2,749	0	14	30	35	22	21	7	10	0	**139**
	2,750–3,000	0	11	233	11	12	35	0	105	1	**408**
	All with ≥50 mph	**0**	**2,341**	**2,905**	**706**	**406**	**228**	**119**	**213**	**166**	**7,084**
All	**All**	**2,693**	**9,958**	**7,795**	**2,566**	**804**	**938**	**908**	**1,101**	**451**	**27,214**

Table 6 provides the number of curves for right-turning curves (+1), and table 7 provides the number of corresponding traversals.

Table 6. Number of curves by curve radius, shoulder width, and speed limit/advisory speed for right-turning curves (+1).

Advisory or Speed Limit (mph)	**Curve Radius (ft)**	**Paved Shoulder Width (ft)**									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**All**
<50	250–499	13	23	11	10	5	1	1	2	0	**66**
	500–749	32	41	32	12	4	8	9	0	1	**139**
	750–999	16	41	42	13	11	4	1	3	2	**133**
	1,000–1,249	19	43	34	10	5	5	4	2	2	**124**
	1,250–1,499	9	30	24	6	7	3	7	4	3	**93**
	1,500–1,749	8	25	13	8	5	1	2	2	1	**65**
	1,750–1,999	7	15	10	8	4	6	2	1	1	**54**
	2,000–2,249	5	17	7	2	3	3	2	1	3	**43**
	2,250–2,499	4	4	8	4	1	2	4	2	2	**31**
	2,500–2,749	4	8	5	1	3	0	2	0	1	**24**
	2,750–3,000	5	8	5	2	2	5	1	2	0	**30**
	All with <50 mph	**122**	**255**	**191**	**76**	**50**	**38**	**35**	**19**	**16**	**802**
≥50	250–499	0	0	0	0	0	0	0	0	0	**0**
	500–749	0	0	0	0	0	0	0	0	0	**0**
	750–999	0	6	11	3	2	2	0	0	0	**24**

Advisory or Speed Limit (mph)	Curve Radius (ft)	Paved Shoulder Width (ft)									
		0	1	2	3	4	5	6	7	8	All
	1,000–1,249	0	15	10	2	1	1	1	3	0	**33**
	1,250–1,499	0	6	15	9	4	1	2	1	3	**41**
	1,500–1,749	0	2	8	3	0	4	0	2	0	**19**
	1,750–1,999	0	9	12	9	7	3	1	1	1	**43**
	2,000–2,249	0	9	10	4	2	2	2	1	0	**30**
	2,250–2,499	0	4	9	6	4	2	1	2	2	**30**
	2,500–2,749	0	4	1	1	1	2	0	2	0	**11**
	2,750–3,000	0	12	8	5	1	2	1	3	2	**34**
	All with ≥50 mph	**0**	**67**	**84**	**42**	**22**	**19**	**8**	**15**	**8**	**265**
All	**All**	**122**	**322**	**275**	**118**	**72**	**57**	**43**	**34**	**24**	**1,067**

Table 7. Number of traversals by curve radius, shoulder width, and speed limit/advisory speed for right-turning curves (+1).

Advisory or Speed Limit (mph)	Curve Radius (ft)	Paved Shoulder Width (ft)									
		0	1	2	3	4	5	6	7	8	All
<50	250–499	396	680	226	146	113	115	34	19	0	**1,729**
	500–749	500	2,135	1,392	439	135	235	459	0	8	**5,303**
	750–999	158	1,041	1,496	181	467	108	42	295	28	**3,816**
	1,000–1,249	601	2,139	1,761	268	63	164	12	16	252	**5,276**
	1,250–1,499	313	1,008	1,073	100	156	167	238	129	716	**3,900**
	1,500–1,749	146	776	702	172	348	9	6	24	11	**2,194**
	1,750–1,999	252	273	121	243	52	73	6	80	9	**1,109**
	2,000–2,249	283	837	182	36	90	78	133	51	624	**2,314**
	2,250–2,499	53	457	279	38	1	37	364	16	157	**1,402**
	2,500–2,749	26	615	83	17	40	0	94	0	11	**886**
	2,750–3,000	51	303	73	4	55	358	3	30	0	**877**
	All with <50 mph	**2,779**	**10,264**	**7,388**	**1,644**	**1,520**	**1,344**	**1,391**	**660**	**1,816**	**28,806**

Advisory or Speed Limit (mph)	**Curve Radius (ft)**	**Paved Shoulder Width (ft)**									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**All**
≥50	250–499	0	0	0	0	0	0	0	0	0	**0**
	500–749	0	0	0	0	0	0	0	0	0	**0**
	750–999	0	213	255	26	12	170	0	0	0	**676**
	1,000–1,249	0	527	107	22	1	1	0	278	0	**936**
	1,250–1,499	0	388	147	67	75	1	45	1	244	**968**
	1,500–1,749	0	92	1,140	32	0	57	0	85	0	**1,406**
	1,750–1,999	0	324	106	30	145	5	2	11	15	**638**
	2,000–2,249	0	369	154	60	13	226	8	7	0	**837**
	2,250–2,499	0	7	88	73	69	21	23	18	11	**310**
	2,500–2,749	0	52	46	0	1	4	0	5	0	**108**
	2,750–3,000	0	284	27	20	2	63	23	79	17	**515**
	All with ≥50 mph	0	**2,256**	**2,070**	**330**	**318**	**548**	**101**	**484**	**287**	**6,394**
All	**All**	**2,779**	**12,520**	**9,458**	**1,974**	**1,838**	**1,892**	**1,492**	**1,144**	**2,103**	**35,200**

Safety Surrogate Measures

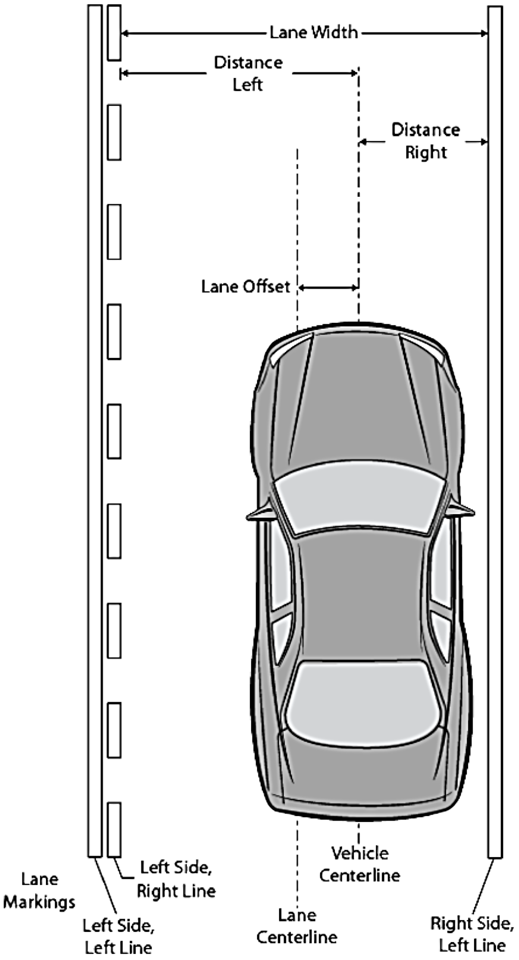
Safety surrogate measures are observations derived from proximity, operation, or location measurements. Based on a specified threshold, a safety surrogate would be deemed unsafe and may be associated with crash occurrence. In this study, the researchers evaluated kinematic measures, driver operation (braking or steering), and lane position.

The time series data were available at various time intervals during the traversal through the curve. Since several observations were available for each curve, the mean estimate and corresponding descriptive statistics were computed for each of the traversals and predictor variables. For example, for vehicle speed, the average speed through the curve was computed with corresponding descriptive statistics (minimum, maximum, and standard deviation). Safety surrogates were evaluated for acceleration and deceleration, speed, steering, braking, and lane position data. The researchers evaluated the distribution of the data and specified thresholds based on observed distributions, speed limits (speeding), or lane position markers (centerline or edge-line marking encroachment).

MODELING

Through exploratory data analysis, several traversal variables were evaluated to serve as safety surrogates. The researchers evaluated variables at univariate and multivariable levels, considering potential combinations of variables. Variables that showed promise were related to lane position.

Lane position refers to the vehicle’s lateral position during the traversal through the curve. Several lateral position measurements were available in the NDS data, which were extracted through a machine vision-based lane tracker. Figure 3 illustrates the variable names and dimensions of the lane position data. In this study, the lane position variables that provided the most promise were lateral distance from the vehicle’s longitudinal axis to the centerline (distance to the left) and the edge line (distance to the right).

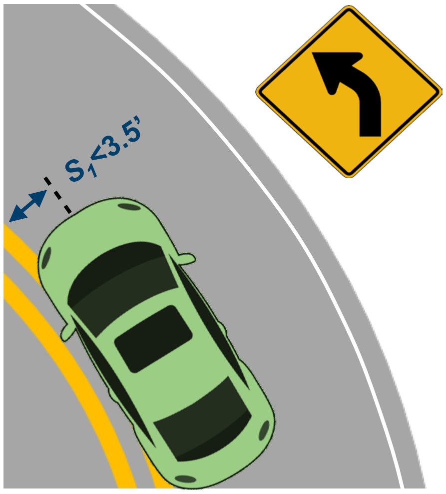


© 2023 VTTI.

Figure 3. Illustration. NDS lane position (VTTI 2023).

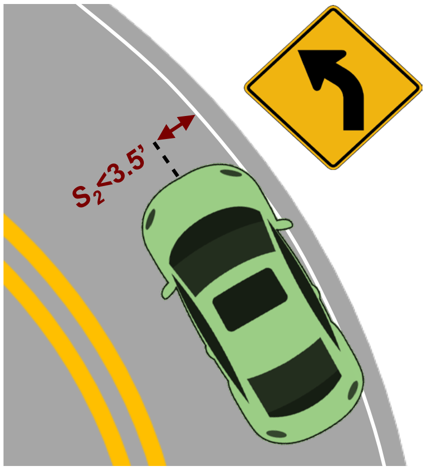
The safety surrogate associated with lane position was defined as centerline and edge-line encroachment. Since vehicle width was not available, for consistency, a fixed threshold was selected to define encroachment. Based on the passenger car vehicle design in the American Association of State Highway and Transportation Officials (AASHTO) *A Policy on Geometric Design of Highways and Streets* (also known as the Green Book), a vehicle width of 7.0 ft was assumed (AASHTO 2011). Since the lane position data are referenced to the vehicle’s longitudinal axis, the researchers defined encroachment events with a threshold of 3.5 ft from the roadway centerline or edge-line marking. For instance, in figure 4-A, assuming a vehicle width of 7.0 ft and a left-turning curve, a roadway centerline encroachment was considered when the vehicle’s longitudinal axis had a lateral distance from the roadway centerline smaller than 3.5 ft.

Following the same criteria in figure 4-B, an edge-line encroachment was considered when the vehicle’s longitudinal axis had a lateral distance from the edge line smaller than 3.5 ft. The same threshold applies when evaluating the opposite direction of travel (right-turning curves). The researchers implemented several methodological approaches, including extreme value theory (Songchitruksa and Tarko 2006; Tarko 2012); however, cross-sectional modeling of safety surrogates with the negative binomial provided the most consistent and statistically significant results.



© 2023 University of Wisconsin–Madison.
S_1 = threshold between the vehicle’s centerline and longitudinal axis.

A. Centerline encroachment.



© 2023 University of WisconsinMadison.
S_2 = threshold between the vehicle’s edge line and longitudinal axis.

B. Edge-line encroachment.

Figure 4. Illustration. Encroachment thresholds.

Negative Binomial Regression Modeling

The negative binomial regression is a generalization of Poisson regression in which the variance is greater than the mean. The probability distribution is a mixture of Poisson-Gamma distribution in which the Gamma distribution provides additional shape and scale parameters that introduce additional flexibility to account for dispersion. This modeling approach is conventionally used to estimate the number of outcomes given a particular number of events. For instance, crash frequency is commonly modeled as the number of crashes over a specified period. Variables modeled include a measure of exposure, such as traffic volume (reference number of vehicles going through the roadway facility), and roadway factors, such as geometry. This study modeled encroachment events as a function of the number of traversals at each curve (measure of exposure) and curve geometry traits (radius, lane width, and paved shoulder width). The researchers developed models for two curve categories with advisory speed or speed limits <50 mph and ≥50 mph.

Association of Lane Departures and Observed Crashes

The methodological approach consisted of obtaining curves with available AADT and observed crashes from State data. Although the reduced dataset included 2,030 curves, the sample size was significantly reduced because some of the curves did not have AADT or State crash data available. Several efforts—including collecting additional crash and traffic volume data from State department of transportation databases and additional queries—were implemented to increase the number of curves. Thus, 629 curves were available with AADT and crash data. The data provided between one and three AADT estimates for each curve in different years, and crash data were available for eight consecutive years between 2006 and 2013.

The researchers evaluated the correlation between encroachment estimates and observed crashes. A correlation statistical measure was used to determine the size and direction of the association, including the statistical significance. Also, linear regression modeling was implemented with frequency of lane departure as the predictor variable and frequency of observed crashes as the dependent variable. The model coefficient, statistical significance, and coefficient of determination (R^2) were evaluated to assess goodness of fit.

TOOL DEVELOPMENT

One of the project’s main objectives was implementing the findings through an analytical tool for practitioners to consider in curve design. The results of modeling in this research project provide an association of encroachment with observed crashes (from State data) to quantify safety and costs among different geometric design alternatives explored in the design of horizontal curves on rural, two-lane roads. Using these estimates, the research team developed a tool to analyze and compare different curve design alternatives. Methods implemented in the tool calculations include the Empirical Bayes to estimate expected crashes, safety effect, severity, crash type distribution factors, and crash benefit-cost.

Data Input

Required data input includes the proposed curve radius, length, AADT, estimated cost to construct the curve, and expected service life. Data for the baseline curve and one or more proposed alternative curves will be required.

Safety Analysis

With the availability of encroachment estimates using geometric and operational features of curves, equivalent crash estimates can be obtained. In the tool, crashes are distributed according to the *Highway Safety Manual* (HSM) severity distribution of rural, two-lane roadways,[1] which was obtained using Highway Safety Information System data for the years 2002–2006 (AASHTO 2014). However, the user can introduce local severity distribution factors if desired. Table 8 provides the severity distribution factors based on crash severity as defined by the KABCO injury classification scale.

Table 8. HSM severity distribution of rural, two-lane roadways (AASHTO 2014).

Crash Severity	Percentage
Fatal (K)	1.3
Incapacitating injury (A)	5.4
Nonincapacitating injury (B)	10.9
Possible injury (C)	14.5
Property damage only (O)	67.9
Total	**100**

Based on the State crash data available for the curves in this study, the researchers estimated the percentage of single-vehicle crashes. For rural, two-lane roadways with an advisory speed or speed limit <50 mph, the percentage of single-vehicle crashes was 51.4 percent, and for curves with an advisory speed or speed limit ≥50 mph, the percentage of single-vehicle crashes was 67.2 percent. With the tool, the user can introduce local estimates of the percentage of single-vehicle crashes, if desired.

According to the KABCO severity distribution, crash estimates by severity would be combined with crash cost estimates by severity to obtain overall crash costs for a specific curve design. The researchers used crash costs in 2010 dollars by severity from the revised *Economic and Societal Impact of Motor Vehicle Crashes* report (Blincoe et al. 2015; Harmon, Bahar, and Gross 2018). Table 9 provides the crash cost estimates. The severity distribution and crash costs serve as a reference and can be used as a default in the tool, but the tool also allows the user to introduce local severity distribution factors, crash costs, and other costs. The tool allows the user to update the crash cost to any specific year with economic factors such as the Consumer Price Index and Usual Weekly Earnings of Wage and Salary Workers (U.S. Bureau of Labor Statistics 2024; U.S. Bureau of Labor Statistics 2023).

---

[1] HSM chapter 10, table 10-3, page 10–17.

Table 9. National KABCO crash unit costs in 2010 dollars (Blincoe et al. 2015; Harmon, Bahar, and Gross 2018).

Crash Severity	Economic Crash Unit Cost	QALY Crash Unit Cost	Comprehensive Crash Unit Cost
K	$1,565,406	$8,583,550	$10,148,956
A	$118,172	$470,675	$588,847
B	$48,789	$129,835	$178,624
C	$38,645	$74,450	$113,095
O	$10,817	—	$10,817

—No data.
QALY = quality-adjusted life years.

Economic Assessment

The economic assessment compares the difference in construction costs to the difference in crash costs when a proposed design is compared to a baseline design. The economic assessment provides a quantitative measure for practitioners to confidently evaluate the implementation of curve design alternatives by assessing the tradeoffs between safety and construction costs.

CHAPTER 4. RESULTS

Results of this study include negative binomial regression model coefficients and measures of goodness of fit (centerline and edge-line encroachment models), encroachment model estimates associated with observed crashes, and an analytical tool for rural two-lane horizontal curves.

ENCROACHMENT MODELS

Using the negative binomial model, the researchers estimated the number of encroachments to the centerline or edge line as a function of the number of observed traversals and curve geometric features. The dependent variable was the number of encroachments in both directions of travel of a curve. Encroachment is defined as an event with a distance between the longitudinal axis of the vehicle and the roadway centerline or edge line less than 3.5 ft. The predictor variables were the number of traversals through the curve, curve length, curve radius, lane width, and shoulder width. The researchers developed encroachment models for two categories of curves with an advisory speed or speed limit <50 mph or ≥50 mph. Equation 1 illustrates the encroachment model.

$$E = LEN \times e^{a} \times TRAV^{b} \times e^{c \times RAD + d \times SHW + e \times LNW} \tag{1}$$

Where:

E = encroachments.
LEN = curve length (feet).
TRAV = number of traversals (both directions of travel).
RAD = curve radius (feet).
SHW = shoulder width (feet).
LNW = lane width (feet).
a–e = model coefficients.

The model follows conventional functional forms used for predicting roadway segment crashes in which the length is specified as an offset or scaling variable, intercept or constant term, or power function for the measure of exposure, and the rest of the terms included as exponential functions. Table 10 provides the model coefficients for curves with an advisory speed or speed limit <50 mph, and table 11provides the model coefficients for curves with an advisory speed or speed limit ≥50 mph.

Table 10. Encroachment model coefficients for curves with an advisory speed or speed limit <50 mph.

Description	Estimate	Standard Error	P-Value
(Intercept, *a*)	−5.0710	0.5724	<0.001
Traversals (*TRAV*, *b*)	1.0080	0.0410	<0.001
Radius (*RAD*, *c*)	−0.0004	0.0001	<0.001
Shoulder width (*SHW*, *d*)	−0.0948	0.0286	<0.001
Lane width (*LNW*, *e*)	−0.3034	0.0569	<0.001
Dispersion term (*1/k*)	0.8975	0.0823	<0.001

Table 11. Encroachment model coefficients for curves with an advisory speed or speed limit ≥50 mph.

Description	Estimate	Standard Error	P-Value
(Intercept, *a*)	−6.4856	1.4688	<0.001
Traversals (*TRAV*, *b*)	1.1686	0.1047	<0.001
Radius (*RAD*, *c*)	−0.0004	0.0002	0.0992
Shoulder width (*SHW*, *d*)	−0.0893	0.0706	0.2059
Lane width (*LNW*, *e*)	−0.2671	0.1274	0.0360
Dispersion term (*1/k*)	0.4503	0.0700	<0.001

Model coefficients consistently indicate a positive effect (negative coefficient in an exponential function) as the curve radius, shoulder width, and lane width increase. Fewer encroachments are predicted with a larger curve radius, wider shoulder, or wider lane width. Regarding the measure of exposure, the predictor variable traversals (*TRAV*) have the expected positive coefficient in a power function. With an increasing number of traversals, an increasing number of encroachments are predicted. Encroachment models have an additional term for overdispersion, which can be used with the Empirical Bayes method.

ENCROACHMENTS AND OBSERVED CRASHES ASSOCIATION

The amount of data used for the association of encroachments and observed crashes was significantly reduced because of limited availability of AADT and State crash data. Since some of the roadways in the analysis were in remote areas, AADT measures were not available. Table 12 provides a breakdown of the data available for the association, which included 629 curves with available AADT and a total of 3,104 crashes. Between one and three AADT estimates were available for each curve in different years, and crash data were available for eight consecutive years between 2006 and 2013. The researchers reviewed available crash data for every State and further classified crashes by single-vehicle crashes. Distribution of single-vehicle crashes accounted for 51.4 percent (advisory speed or speed limit <50 mph) and 67.2 percent (advisory speed or speed limit of ≥50 mph) of all crashes.

Table 12. RID curves with AADT and State crash data available.

State	Curves			Single-Vehicle Crashes (8 yr)			All Crashes (8 yr)		
	<50 mph	≥50 mph	All	<50 mph	≥50 mph	All	<50 mph	≥50 mph	All
FL	0	6	**6**	0	15	**15**	0	39	**39**
IN	92	127	**219**	203	365	**568**	335	510	**845**
NC	130	48	**178**	405	208	**613**	844	343	**1,187**
NY	107	34	**141**	252	204	**456**	496	292	**788**
PA	56	23	**79**	96	30	**126**	151	40	**191**
WA	6	0	**6**	11	0	**11**	54	0	**54**
Total	**391**	**238**	**629**	**967**	**822**	**1,789**	**1,880**	**1,224**	**3,104**

To associate encroachments with crashes, the researchers adjusted the units to the same terms or dimensions. Using observed crashes from State data, the researchers obtained the crash rate in terms of crashes/year-mile. Equation 2 shows the formula to compute the crash rate. In the case of encroachments, estimates had to be normalized according to similar units of crash rate, as illustrated in equations 3 and 4. In equation 5, the number of traversals variable was replaced with the AADT multiplied by 365, which would be the overall number of traversals that would be expected in both directions of travel through the curve in a calendar year.

$$N = \frac{crashes}{8\ years \times \frac{LEN}{5{,}280}} \quad \left(\frac{crashes}{year - mi}\right) \tag{2}$$

$$D = \frac{E}{1{,}000{,}000 \times \frac{LEN}{5{,}280}} \quad \left(\frac{million\ encroachments}{year - mi}\right) \tag{3}$$

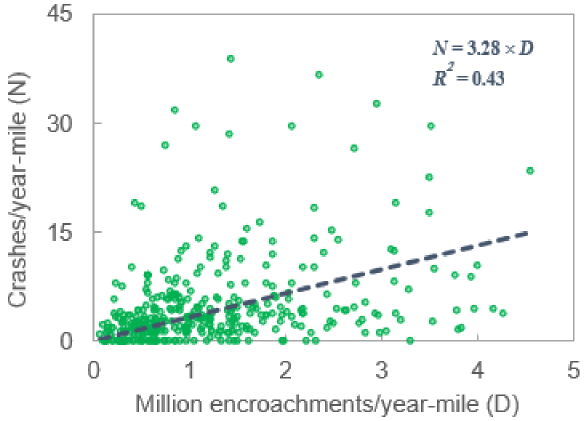
$$D = \frac{5{,}280 \times e^{a} \times TRAV^{b} \times e^{c \times RAD + d \times SHW + e \times LNW}}{1{,}000{,}000} \quad \left(\frac{million\ encroachments}{year - mi}\right) \tag{4}$$

$$D = \frac{5{,}280 \times e^{a} \times (AADT \times 365)^{b} \times e^{c \times RAD + d \times SHW + e \times LNW}}{1{,}000{,}000} \quad \left(\frac{million\ encroachments}{year - mi}\right) \tag{5}$$

Where:

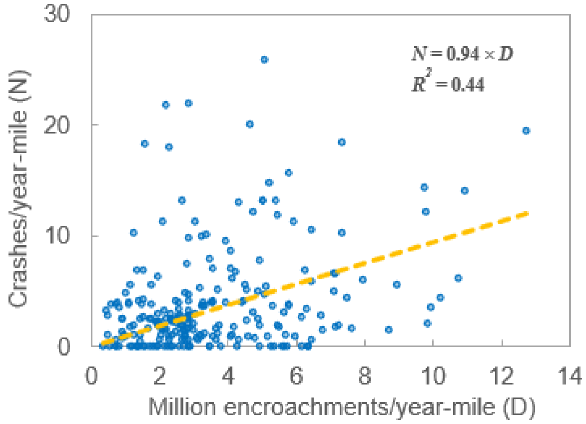
N = observed crash rate (crashes/year-mile).
D = encroachment rate (million encroachments/year-mile).
$AADT$ in vehicles per day (vpd).
RAD = curve radius (feet).

The correlation analysis between encroachment estimates and observed crashes from State data showed a positive linear correlation of 0.3317 (advisory speed or speed limit <50 mph) and 0.3032 (advisory speed or speed limit ≥50 mph). Also, linear regression modeling of encroachments as the predictor variable and observed crashes as the dependent variable resulted in a linear model with statistically significant slopes of 3.2810 (p-value <0.001, advisory speed or speed limit <50 mph) and 0.9397 (p-value <0.001, advisory speed or speed limit ≥50 mph), and model R^2 equal to 0.4286 and 0.4386, respectively (figure 5). Table 13 provides the results of the encroachment-crash association. Thus, crashes may be estimated using encroachment estimates (*D*) obtained using equation 5 and multiplied by a conversion factor (*m*).



D = encroachment rate; N = observed crash rate.

A. Advisory or speed limit <50 mph.



B. Advisory or speed limit ≥50 mph.

Figure 5. Illustration. Encroachment and crash association.

Table 13. Results of encroachments and observed crashes association.

Advisory or speed limit (mph)	Description	Estimate	P-Value
<50	Slope (*m*)	3.2810	<0.001
	Correlation	0.3317	<0.001
	R^2	0.4286	NA
≥50	Slope (*m*)	0.9394	<0.001
	Correlation	0.3032	<0.001
	R^2	0.4386	NA

NA = not applicable.

Based on the association analysis, crashes for a given period of analysis and curve geometry can be estimated using equation 6.

$$P = m \times \frac{YR \times LEN \times e^{a} \times (AADT \times 365)^{b} \times e^{c \times RAD + d \times SHW + e \times LNW}}{1{,}000{,}000} \quad \left(\frac{crashes}{period}\right) \tag{6}$$

Where:

P = predicted crashes (crashes/period).

m = slope (encroachment to crashes conversion factor).

YR = period of analysis (years).

ANALYTICAL TOOL

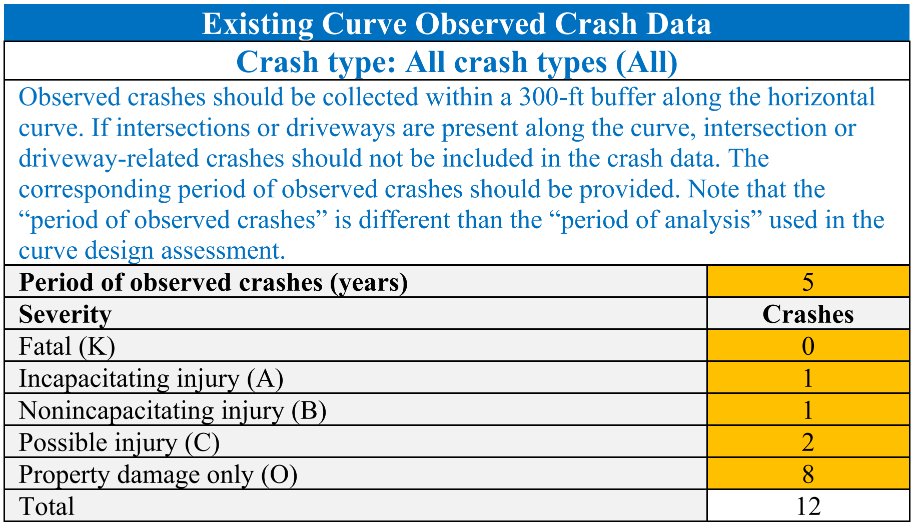
The researchers incorporated the findings of this research study into an analytical tool (University of Wisconsin–Madison 2024) for practitioners to compare multiple horizontal curve designs on two-lane, undivided roadways. The tool estimates crashes and associated crash costs to perform an economic assessment of a curve design. The results are provided as part of the economic assessment, which compares the construction costs to the difference in crash costs of proposed designs to a baseline design. The economic assessment provides a quantitative measure for practitioners to evaluate different design alternatives by assessing the tradeoffs between safety and implementation costs. The HSM severity distribution, single-vehicle distribution, and crash costs are used for reference in the tool, but the tool allows users to introduce preferred or local severity distribution factors, crash costs, or economic measures to adjust crash costs to the current year (AASHTO 2014). The tool has five tabs, including instructions, automated calculations, model coefficients and equations, distribution factors, and crash costs.

Instructions

The instructions tab includes the project information and guidance to navigate the tool. An overview of the project is provided, and descriptions of the elements and color coding of the tool are explained. A brief description of every tab of the tool is provided. In the instructions section is contact information for support, and a section for keeping a record of updates to the tool is also included.

Automated Calculations

This section of the tool provides elements for data inputs used in automated calculations. Required data include observed crash data, the curves' geometric design elements, costs, and analysis period. Figure 6 shows the section of the tool for inputting observed crashes at an existing curve.



Existing Curve Observed Crash Data	
Crash type: All crash types (All)	
Observed crashes should be collected within a 300-ft buffer along the horizontal curve. If intersections or driveways are present along the curve, intersection or driveway-related crashes should not be included in the crash data. The corresponding period of observed crashes should be provided. Note that the "period of observed crashes" is different than the "period of analysis" used in the curve design assessment.	
Period of observed crashes (years)	5
Severity	**Crashes**
Fatal (K)	0
Incapacitating injury (A)	1
Nonincapacitating injury (B)	1
Possible injury (C)	2
Property damage only (O)	8
Total	12

Figure 6. Illustration. Observed crashes on existing curves.

Figure 7 describes some of the basic elements of a simple horizontal curve used for reference in the tool.

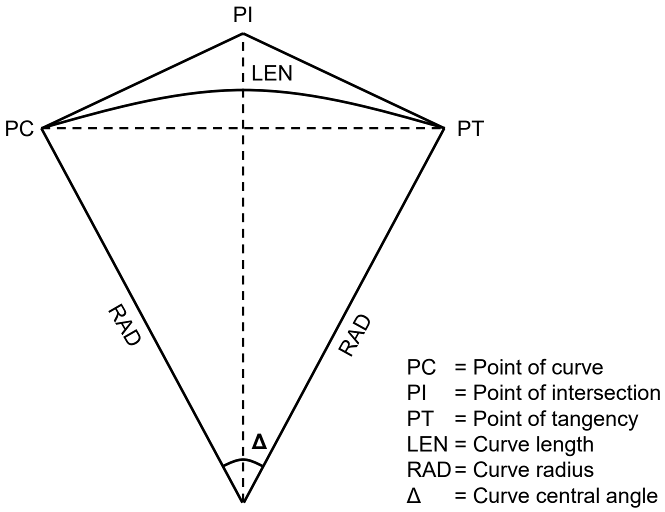
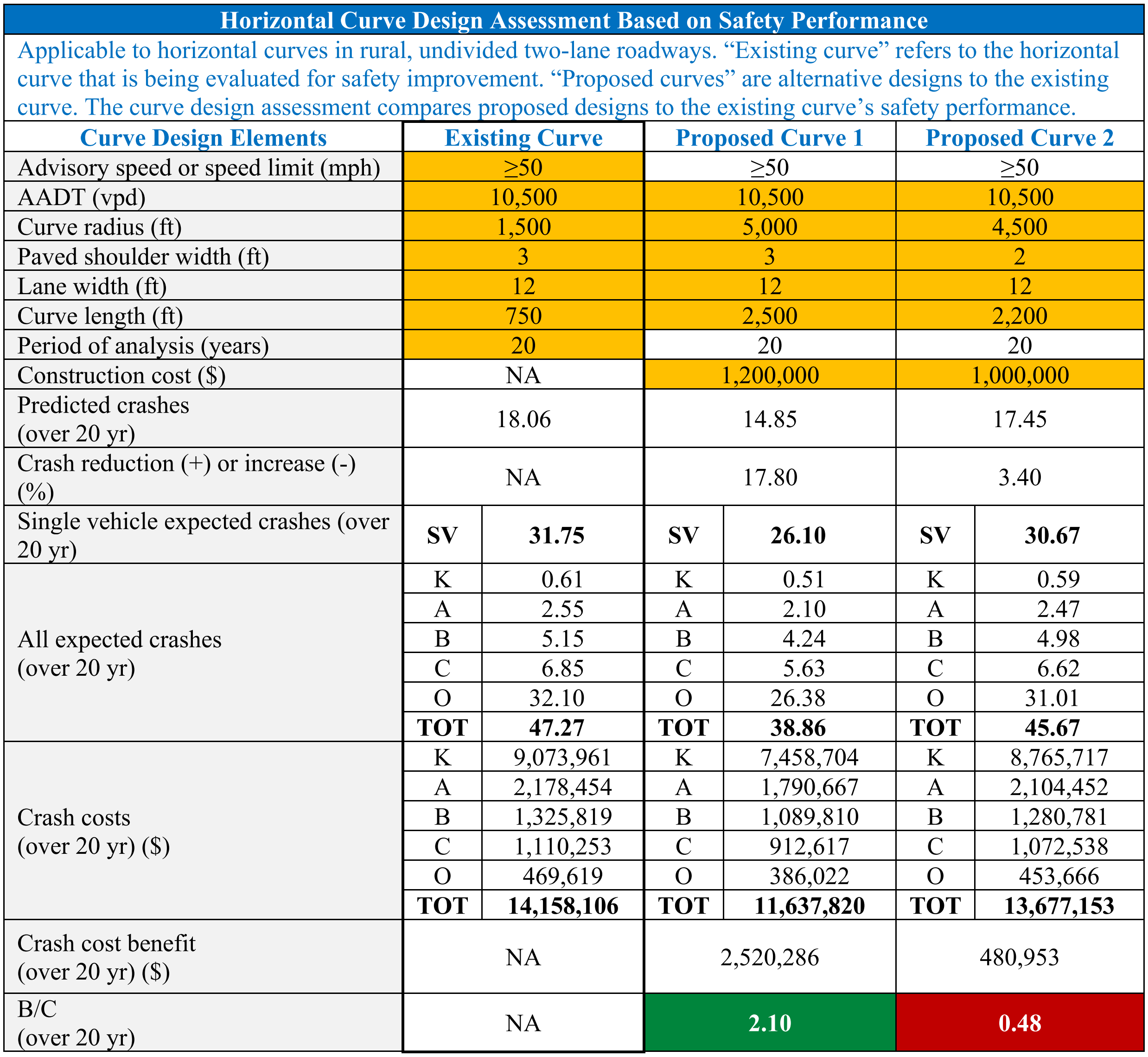


Figure 7. Illustration. Diagram of simple curve geometric elements.

The analytical tool requires data for an existing curve and proposed curve designs. Data required include the advisory speed or speed limit, AADT, curve radius, paved shoulder width, lane width, curve length, period of analysis, and construction cost. Based on data inputs, the tool estimates predicted crashes to quantify the proposed curves' effect on crashes compared to the existing curve design. Using the Empirical Bayes method—along with the safety effects of the proposed curves, predicted crashes, and observed crashes—the tool can estimate expected overall and single-vehicle crashes. Expected crashes can be further evaluated by severity based on distribution factors. Based on expected crashes by severity, crash costs can be obtained to evaluate the crash cost benefit of the proposed designs compared to the existing curve. Crash cost benefits are then compared to the construction cost of the proposed designs to estimate the benefit-cost ratios. Figure 8 illustrates the analytical section of the tool.



Horizontal Curve Design Assessment Based on Safety Performance

Applicable to horizontal curves in rural, undivided two-lane roadways. "Existing curve" refers to the horizontal curve that is being evaluated for safety improvement. "Proposed curves" are alternative designs to the existing curve. The curve design assessment compares proposed designs to the existing curve's safety performance.

Curve Design Elements	**Existing Curve**		**Proposed Curve 1**		**Proposed Curve 2**	
Advisory speed or speed limit (mph)	≥50		≥50		≥50	
AADT (vpd)	10,500		10,500		10,500	
Curve radius (ft)	1,500		5,000		4,500	
Paved shoulder width (ft)	3		3		2	
Lane width (ft)	12		12		12	
Curve length (ft)	750		2,500		2,200	
Period of analysis (years)	20		20		20	
Construction cost ($)	NA		1,200,000		1,000,000	
Predicted crashes (over 20 yr)	18.06		14.85		17.45	
Crash reduction (+) or increase (-) (%)	NA		17.80		3.40	
Single vehicle expected crashes (over 20 yr)	**SV**	**31.75**	**SV**	**26.10**	**SV**	**30.67**
All expected crashes (over 20 yr)	K	0.61	K	0.51	K	0.59
	A	2.55	A	2.10	A	2.47
	B	5.15	B	4.24	B	4.98
	C	6.85	C	5.63	C	6.62
	O	32.10	O	26.38	O	31.01
	TOT	**47.27**	**TOT**	**38.86**	**TOT**	**45.67**
Crash costs (over 20 yr) ($)	K	9,073,961	K	7,458,704	K	8,765,717
	A	2,178,454	A	1,790,667	A	2,104,452
	B	1,325,819	B	1,089,810	B	1,280,781
	C	1,110,253	C	912,617	C	1,072,538
	O	469,619	O	386,022	O	453,666
	TOT	**14,158,106**	**TOT**	**11,637,820**	**TOT**	**13,677,153**
Crash cost benefit (over 20 yr) ($)	NA		2,520,286		480,953	
B/C (over 20 yr)	NA		**2.10**		**0.48**	

© 2023 University of Wisconsin–Madison.
SV = single vehicle; TOT = total.

Figure 8. Illustration. Analytical section of the tool with safety and economic estimates.

Model Coefficients and Equations, Distribution Factors, and Crash Costs.

All the equations, coefficients, factors, and costs are provided in separate sections of the tool. For severity and single-vehicle crash distribution factors, the tool uses factors from the HSM and this project's estimates as default values for calculation; however, the user may introduce local factors, if desired (AASHTO 2014). In terms of crash costs, economic factors may be introduced to adjust default crash costs to the current year, if desired.

CHAPTER 5. CONCLUSIONS

Using performance measures for project design decisionmaking provides design flexibility by specifically addressing the purpose and needs of projects. Transportation engineers need resources to evaluate roadway design based on safety and economic performance. In this project, the researchers used the RID database along with SHRP2 NDS data to estimate the safety effect of geometric elements that influence driving behavior on rural undivided two-lane horizontal curves (Iowa State University 2024; VTTI 2023). Lane position was used as a safety surrogate to define encroachment. The researchers modeled encroachment events with the negative binomial using curve geometry as the predictor component. Then the researchers associated encroachment estimates with observed crashes from State data to convert encroachments to crashes. As curve radius, shoulder, or lane width increased, predicted crashes decreased. The findings of this research study were used to develop an analytical tool for practitioners to use when considering curve designs. The tool estimates crashes and associated crash costs to perform an economic assessment of different curve designs. The economic assessment provides a quantitative measure for practitioners to evaluate different curve designs by assessing the tradeoffs between safety and costs of implementation.

Following are some of the lessons learned through the process of providing safety estimates using naturalistic data:

- Significant effort was required to process, clean, and validate the data.
- Many data points were present in a time series format for each traversal along curves.
- A wide range of traversals by curve showed that some curves had a significant overrepresentation of traversals, which may bias data distribution.
- Curve attributes need to be reviewed and confirmed. Curves in the RID may have mixed geometry, the presence of intersections, or attributes that may not be completely accurate.
- Diverse regions, curve attributes (including length, radius, shoulder, speed limit, and shoulder and lane width), vehicles, and drivers introduce variability in the data.
- Time series data cleaning included lack of accuracy for lane position variable (less than 512), repeated values, missing data, integer values, and outliers.
- Traversal lane position data indicating lane departure should be verified with video data.
- Data aggregation may be required to account for the variability of observations in a time series format.

REFERENCES

AASHTO. 2014. *Highway Safety Manual*. Washington, DC: American Association of State Highway and Transportation Officials. https://www.highwaysafetymanual.org/pages/default.aspx, last accessed December 28, 2023.

AASHTO. 2011. *A Policy on Geometric Design of Highways and Streets*, 6th edition. Washington, DC: American Association of State Highway and Transportation Officials.

Blincoe, L., T. R. Miller, E. Zaloshnja, and B. A. Lawrence. 2015. *The Economic and Societal Impact of Motor Vehicle Crashes, 2010 (Revised)*. Report No. DOT HS 812 013. Washington, DC: National Highway Traffic Safety Administration.

FHWA. 2023a. "SHRP 2 Naturalistic Driving Study Pooled Fund: Advancing Implementable Solutions." Study No. TPF-5(361). Federal Highway Administration. https://www.pooledfund.org/Details/Study/613, last accessed January 29, 2024.

FHWA. 2023b. *Manual on Uniform Traffic Control Devices for Streets and Highways*, 11th Edition. Washington, DC: Federal Highway Administration.

FHWA. 2017. "Brief: Overview of Performance-Based Practical Design (PBPD)" (web page). https://www.fhwa.dot.gov/design/pbpd/documents/pbpd_brief.cfm, last accessed January 29, 2024.

Hallmark, S. L., S. Tyner, N. Oneyear, C. Carney, and D. McGehee. 2015. "Evaluation of Driving Behavior on Rural 2-Lane Curves Using the SHRP 2 Naturalistic Driving Study Data." *Journal of Safety Research* 54: 17–27.

Harmon, T., G. B. Bahar, and F. B. Gross. 2018. *Crash Costs for Highway Safety Analysis*. Report No. FHWA-SA-17-071. Washington, DC: Federal Highway Administration.

Iowa State University Center for Transportation Research and Education (CTRE). 2024. "Roadway Information Database (RID)" (web page) . https://ctre.iastate.edu/roadway-information-database-rid/, latest accessed January 29, 2024.

Songchitruksa, P., and A. P. Tarko. 2006. "The Extreme Value Theory Approach to Safety Estimation." *Accident Analysis & Prevention* 38, no. 4: 811–822.

Tarko, A. P. 2012. "Use of Crash Surrogates and Exceedance Statistics To Estimate Road Safety." *Accident Analysis & Prevention* 45: 230–240.

University of Wisconsin–Madison. 2024. *Horizontal Curve Safety Performance and Economic Evaluation Analytical Tool* (software). https://topslab.wisc.edu/wp-content/uploads/2024/01/Curve_Tool_July2023_FINAL.xlsx, last accessed February 21, 2024.

U.S. Bureau of Labor Statistics. 2024. "Consumer Price Index" (web page). https://www.bls.gov/cpi/tables/supplemental-files/, last accessed February 14, 2024.

U.S. Bureau of Labor Statistics. 2023. "Usual Weekly Earnings of Wage and Salary Workers Archived News Releases" (web page). https://www.bls.gov/bls/news-release/wkyeng.htm#2010, last accessed February 14, 2024.

VTTI. 2023. *Second Strategic Highway Research Program Naturalistic Driving Study Dataset*. Available from the SHRP2 NDS InSight Data Dissemination website. https://insight.shrp2nds.us, last accessed December 27, 2023.

Wang, B., S. Hallmark, P. Savolainen, and J. Dong. 2017. "Crashes and Near-Crashes on Horizontal Curves Along Rural Two-Lane Highways: Analysis of Naturalistic Driving Data." *Journal of Safety Research* 63: 163–169.

Wang, B., S. Hallmark, P. Savolainen, and J. Dong. 2018. "Examining Vehicle Operating Speeds on Rural Two-Lane Curves Using Naturalistic Driving Data." *Accident Analysis & Prevention* 118: 236–243.

Wu, L., B. Dadashova, S. Geedipally, M. P. Pratt, and M. Shirazi. 2019. "Using Naturalistic Driving Study Data To Explore the Association Between Horizontal Curve Safety and Operation on Rural Two-Lane Highways." *Journal of Transportation Safety & Security* 13, no. 8: 1–18.

Recycled
Recyclable

Recommended citation: Federal Highway Administration,
*Horizontal Curve Safety Performance Evaluation Based
on the Naturalistic Driving Study Lane Position Data*
(Washington, DC: 2024) https://doi.org/10.21949/1521769

HRSO-20/05-24(WEB)E