DEVELOPING A DATA-DRIVEN SAFETY ASSESSMENT FRAMEWORK FOR RITI COMMUNITIES IN WASHINGTON STATE

FINAL PROJECT REPORT

by

Yinhai Wang, Wei Sun, Sam Ricord, Cesar Maia de Souza, Shuyi Yin, and Meng-Ju Tsai
University of Washington

for

Center for Safety Equity in Transportation (CSET)
USDOT Tier 1 University Transportation Center
University of Alaska Fairbanks
ELIF Suite 240, 1764 Tanana Drive
Fairbanks, AK 99775-5910

In cooperation with U.S. Department of Transportation, Research and Innovative Technology Administration (RITA)

rural · isolated · tribal · indigenous

DISCLAIMER

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Developing a Data-Driven Safety Assessment Framework for RITI Communities in Washington State		**5. Report Date** Sep 10, 2021 **6. Performing Organization Code**
7. Author(s) and Affiliations Yinhai Wang, Wei Sun, Sam Ricord, Cesar Maia de Souza, Shuyi Yin, and Meng-Ju Tsai University of Washington		**8. Performing Organization Report No.** INE/CSET 21.09
9. Performing Organization Name and Address Center for Safety Equity in Transportation ELIF Building Room 240, 1760 Tanana Drive Fairbanks, AK 99775-5910		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** Grant # 69A3551747129
12. Sponsoring Organization Name and Address United States Department of Transportation Research and Innovative Technology Administration 1200 New Jersey Avenue, SE Washington, DC 20590		**13. Type of Report and Period Covered** Final report, Sep 2019 – Sep 2021 **14. Sponsoring Agency Code**
15. Supplementary Notes Report uploaded to:		

16. Abstract

The roadway safety of the Rural, Isolated, Tribal, or Indigenous (RITI) communities has become an important social issue in the United States. Official data from the Federal Highway Administration (FHWA) shows that, in 2012, 54 percent of all fatalities occurred on rural roads while only 19 percent of the US population lived in rural communities. Under the serious circumstances, this research aims to help the RITI communities to improve their roadway safety through the development of a roadway safety management system. Generally, a roadway safety management system includes two critical components, the baseline data platform and safety assessment framework. In our Year 1 and Year 2 CSET projects, a baseline data platform was developed by integrating the safety related data collected from the RITI communities in Washington State. This platform is capable of visualizing the accident records on the map. The Year 3 project further developed the safety data platform by developing crash data analysis and visualization functions. In addition, various roadway safety assessment methods had been developed to provide safety performance estimation, including historical accident data averages, predictions based on statistical and machine learning (ML) models, etc. Beside roadway safety assessment methods, this project investigated the safety countermeasures selection and recommendation methods for RITI communities. Specifically, the research team has reached out to RITI communities and established a formal research partnership with the Yakama Nation. The research team has conducted research on safety countermeasures analysis and recommendation for RITI communities.

17. Key Words Safety Data Tool; Roadway Safety Assessment; Statistical and Machine Learning Modeling; RITI communities		**18. Distribution Statement**	
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 39	**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 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 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 (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the 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

EXECUTIVE SUMMARY

RITI communities often do not have the capability and resources to sufficiently solve roadway safety problems. In this case, several challenges are often encountered when addressing transportation safety issues in RITI communities, including: (1) crashes are often randomly distributed on local and rural roads in RITI areas; (2) there is a critical need for data-driven safety assessment methods for RITI communities; (3) RITI communities often lack safety data tools for data management and visualization to support decision making. A survey conducted by the National Association of Counties (NACo) in 2009 revealed that only 42 percent of counties surveyed maintained a database that tracks the number and types of crashes on their rural roads and less than half of the respondents had conducted a road safety audit. Existing databases are still incomplete for most of the RITI communities. It is necessary to develop safety data platform and assessment methods specifically for RITI communities for traffic safety data management and analysis.

In our Year 1 and Year 2 CSET projects, a baseline data platform, i.e., Safety Net, was developed by integrating the safety related data collected from the RITI communities in Washington State (Wang et al., 2019). This platform is capable of visualizing the accidents records on the map. The Year 3 project aimed to further develop the safety data platform by developing crash data analysis and visualization functions.

In addition, this project developed various roadway safety assessment methods to provide safety performance estimation, including historical accident data averages, predictions based on statistical and machine learning (ML) models, etc. This project investigated the potential influential factors, such as roadway geometric characteristics, environmental conditions, human behaviors, and traffic conditions on the injury severity of crashes occurred on rural roads. Four models, including ordered probit (OP), multinomial logit (MNL), artificial neural network (ANN) and random forest (RF), were trained, tested, and validated using five years of Washington State crash records from 2013 to 2017. It was found that the two Machine Learning models (ANN and RF) performed better than the two statistical models (OP and MNL), and the RF model had the best performance in predicting crash injury severities. The results also showed that variables such as grade percentage, degrees of curvature, shoulder width, driver's gender, roadway width, head on crash, pedestrian/cyclist involved, young driver, truck involved, etc. have significant impact on the crash injury severity on low-volume rural roads.

Beside roadway safety assessment methods, this project investigated the safety countermeasures selection and recommendation methods for RITI communities. Safety countermeasures are developed and implemented aiming at reducing crash frequency and accidents severity on road systems. Concerning the RITI communities, this is even more critical. Specifically, the research team reached out to RITI communities and established a formal research partnership with Yakama Nation. By working with the traffic engineers and planners Yakama Nation DNR Engineering Department, the research team conducted research on safety countermeasures analysis and recommendation for RITI communities.

CHAPTER 1. INTRODUCTION

1.1. Project Background

Road traffic crashes often cause property damage, injuries, and even fatalities. They also account for 25% of congestion in road networks (Cambridge Systematics Inc., 2005). According to the Federal Highway Administration (FHWA), while only 19% of the country's population lives in rural areas, about 54% of the traffic crashes occurred on rural roads (Federal Highway Administration, 2012). This indicates a clear disparity regarding transportation safety in rural areas of the country. According to the National Highway Traffic Safety Administration (NHTSA), from 2007 to 2016, the fatality rate in rural roads was more than two times higher than in urban areas (NHTSA, 2018). This discrepancy in fatalities reveals the urgency to improve the roadway safety conditions in the rural areas in order to achieve transportation equity.

To meet the transportation safety needs of RITI communities, Washington State also faced a lot of challenges. Twenty-two percent of the state's major rural locally and state-maintained roads are in poor condition. An additional 52 percent of rural roads are in mediocre or fair condition. The fatality rate on Washington's rural non-Interstate roads was 1.76 fatalities per 100 million vehicle miles of travel in 2013, nearly three and a half times higher than the 0.52 fatality rate on all other roads and highways in the state. According to the data from Washington State Strategic Highway Safety Plan 2016, more than half of impairment-involved fatalities occurred in rural areas during 2012-2014, and unrestrained occupants are also more likely to die in rural road crashes. It is obvious that rural roadway safety has become an important social issue influencing the sustainable development of RITI communities in Washington State.

1.2. Problem Statement

RITI communities often do not have the capability and resources to sufficiently solve roadway safety problems. In this case, several challenges are often encountered when addressing transportation safety issues in RITI communities, including: (1) crashes are often randomly distributed on local and rural roads in RITI areas; (2) there is a critical need for data-driven safety assessment methods for RITI communities; (3) RITI communities often lack safety data tools for data management and visualization to support decision making. A survey conducted by the National Association of Counties (NACo) in 2009 revealed that only 42 percent of counties surveyed maintained a database that tracks the number and types of crashes on their rural roads and less than half of the respondents had conducted a road safety audit. Existing databases are still incomplete for most of the RITI communities. It is necessary to develop safety data platform and assessment methods specifically for RITI communities for traffic safety data management and analysis.

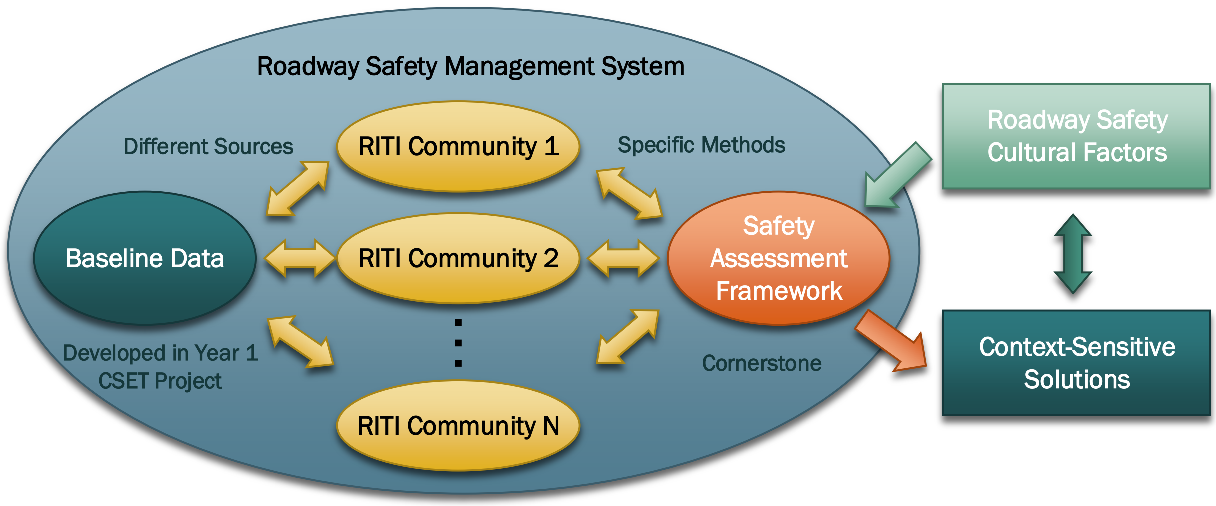


Figure 1.1 Structure of the roadway safety management system for RITI communities

Due to the importance of reducing the social and economic costs associated with traffic crashes, most transportation agencies apply some type of roadway safety management system, designed to improve the roadway safety performance. To build up the roadway safety management system, two critical components, i.e., the baseline data and safety assessment framework, are needed (shown in Figure 1.1). In the Year 1 CSET project, a baseline data platform was developed by integrating the safety related data collected from the RITI communities in Washington State. The safety assessment framework is the cornerstone of the roadway safety management system. Due to different RITI communities having different safety data sources, a general safety assessment method may not be adapted to all the RITI communities. To provide context-sensitive solutions, the roadway safety cultural factors, such as local driving habits and training level, will be considered. All the safety assessment methods form a data-driven safety assessment framework which can enable effective roadway safety management systems at all levels in RITI communities, and aid in roadway design and implementation appropriate countermeasures to mitigate rural crash severities and risks.

1.3. Research Objective

In our Year 1 and Year 2 CSET projects, a baseline data platform, i.e., Safety Net, was developed by integrating the safety related data collected from the RITI communities in Washington State (Wang et al., 2019). This platform is capable of visualizing the accidents records on the map. The Year 3 project aimed to further develop the safety data platform by developing crash data analysis and visualization functions. In addition, this project developed various roadway safety assessment methods to provide safety performance estimation, including historical accident data averages, predictions based on statistical and machine learning (ML) models, etc. Beside roadway safety assessment methods, this project investigated the safety countermeasures selection and recommendation methods for RITI communities.

CHAPTER 2. LITERATURE REVIEW

2.1. Roadway Safety Assessment Methods

Over the years, there have been many studies about crash injury severity modeling and analysis, with the objective to better understand the risk factors that influence the crash injury severities and help transportation agencies make decisions to improve the roadway safety conditions. Crash injury severities are usually classified into several categories by law enforcement. According to the KABCO crash injury scale developed by the National Safety Council, the crash injury severities include five categories: fatal (K), incapacitating injury (A), non-incapacitating injury (B), possible injury (C), and property damage only (PDO). Researchers have developed models with crash injury severities as the dependent variable and various contributing factors as the explanatory variables. Such contributing factors usually include roadway geometric characteristics (e.g., number of lanes and lane width, shoulder width and type, curve rate, grade, road surface type, median type, shoulder type), human behavior factors with respect to drivers, occupants, and pedestrians (e.g., driver and occupant characteristics such as gender and age, DUI, speeding, seat belt use, distractions), environmental conditions (e.g., adverse weather conditions such as fog, snow, ice, heavy rain, and lighting conditions), traffic conditions, and vehicle characteristics.

In the early stage of crash injury severity studies, researchers mostly used statistical models such as Logistic regression to investigate the risk factors related to crash injury severities. With the massive and complicated crash data being collected nowadays, it is difficult for statistical models to accurately capture the impacts of various risk factors to injury severity. Recently, the advancements in computing technology in the fields of artificial intelligence (AI), especially machine learning (ML), have allowed for more efficient and effective extraction of information from extensive traffic safety datasets. Consequently, researchers have been very active in applying ML technologies towards crash injury severity prediction. With an abundance of data available, the ML approaches could capture relationships among contributing factors and crash injury levels that traditional statistical models are not able to, and thus improve the accuracy of the prediction results.

2.2. Crash Injury Analysis on Rural Roads

Many researchers have applied the aforementioned statistical and machine learning methods for injury severity analysis of crashes occurred on rural roads. Vogt and Bared (1998) built negative binomial models to investigate relationships between crash injury and explainable variables including traffic, horizontal and vertical alignments, lane and shoulder widths, and number of driveways, etc. Karlaftis and Golias (2002) used tree-based regression and found that roadway geometric design and pavement conditions are two most important factors for crashes that occurred on rural two-lane and multilane roads. Chen et al. (2016) used the ordered logit model for crash injury severities analysis on rural non-interstate roadway, and the results showed that factors such as seatbelt use, driver age and gender, DUI, wet road surface, crash location, collision type, vehicle type, number of vehicles and maximum vehicle damage, have significant impacts on driver injury severity. A study by Ye and Lord (2014) compared MNL, OP, and mixed logit models on crash injury severity analysis on rural two-way highways. It was found that OP model has the least requirement on sample size comparing to the other two methods. Lin et al. (2020) developed RF models for the driver crash injury severity prediction, and it was found that road class, speed limit, and the first harmful event are the most important factors on the injury severity of teen driver involved crashes occurred on rural roads in West Texas.

Only a few studies were found in the literature that specifically analyzed traffic crashes on low-volume rural roads. Souleyrette et al. (2010) used the OP model to study the influential factors for crashes on rural roads with 400 Annual Average Daily Traffic (AADT) or less in Iowa. Several factors were identified to increase the severity of crashes on low-volume rural roads, such as paved surfaces, spring/summer months, weekends, impaired driving, speeding, younger or older driver involvement, etc. Prato et al. (2014) studied crashes that occurred on low-volume facilities in rural area in Denmark between 2007 and 2011. The authors applied the generalized ordered logit model for crash injury severity analysis. The results indicated that factors including alcohol, seatbelt usage, involvement of vulnerable road users such as pedestrian and cyclists, speed limits, etc. to be significantly associated with crash injury severity on low-volume rural roads.

2.3. Pedestrian Safety Assessment Methods

There have been many different studies related to pedestrian safety. This has been a subject of study for many decades. M Snyder conducted a study in 1971 to identify the causes and countermeasures of pedestrian collisions in Maryland. For this study over 2000 pedestrian collisions were analyzed, mostly focusing on pedestrian behavior. It was found that over 50% of crashes were caused by some form of pedestrians entering the roadway inappropriately (Snyder and Knoblauch, 1971). In 1983, Hall conducted a study to measure rural pedestrian safety in New Mexico. The study described the discrepancy between rural pedestrian fatalities and urban fatalities, with the results for each region being 49% and 34% respectively (Hall, 1983). These studies highlight the importance of pedestrian safety studies and some of the initial methodologies aimed at understanding them.

More recently, an NCHRP study was conducted for the National Academies of Sciences, Engineering, and Medicine that investigated the correlation between site-specific characteristics and pedestrian collisions. They found that site-specific characteristics increased the likelihood of pedestrian collisions (National Academy of Sciences, 2008). A similar study was conducted in Bangladesh. This study also found that specific characteristics precipitate pedestrian collisions in the city of Dhaka (Bhuiyan, 2019). Additionally, several studies were conducted utilizing different modeling techniques to assess pedestrian safety. Zajac created an ordered probit model that evaluated roadway features that are prevalent in pedestrian collisions (Zajac and Ivan, 2002). This model showed different characteristics that influence pedestrian collisions that are more appropriate for a rural setting than the other above studies. Chen conducted a similar study in 2019, which used the alternative method of a mixed logit model to predict rural pedestrian collisions (Chen and Fan, 2019). Baireddy conducted a study in rural Illinois that identified several factors that increase the pedestrian collision likelihood using multiple correspondence analysis (Baireddy et al., 2018). These previous studies hold several implications for this study. Firstly, they show that using roadway characteristics as a method to predict pedestrian collisions is a valid and well-documented methodology. Secondly, it highlights the research gaps filled by this paper where the severity of pedestrian collisions is not considered in any of these previous studies.

It can be noted that most of the literature relies on traditional statistical modeling approaches to address pedestrian safety issues. Nevertheless, with the recent advent of Machine Learning, some researchers have started applying these latest approaches to this type of problem. In 2018, Ding developed a study to examine built environmental effects on the frequency of crashes involving automobiles and pedestrians by applying Multiple Additive Poisson Regression Trees (MAPRT), a Machine Learning approach based on decision trees. Using data from Seattle, Washington, the study helped to detect non-linear relationships between the built environment and pedestrian collisions frequency, confronting the linearity assumption frequently used in studies that use statistical models (Ding et al., 2018). Das applied in 2020 distinct Machine Learning techniques to classify pedestrian collision types (intended vs. untended, pedestrian at fault vs. motorist at fault) using pedestrian crashing data from two locations in Texas (Das et al., 2020). These reference studies were essential for the development of our methodology applied specifically to fatal collisions, an unprecedented approach so far.

CHAPTER 3. SAFETY DATA MANAGEMENT AND VISUALIZATION PLATFORM

In our Year 1 and Year 2 CSET projects, a baseline data platform, Safety Net, was developed by integrating the safety related data collected from the RITI communities in Washington State. This platform is capable of visualizing the accidents records on the map. This project further developed the safety data platform by developing crash data analysis and visualization functions. The safety platform includes various visualization functions to support decision making, such as visualization of crash records in roadway network map based on filtering of crash attributes and roadway features, visualization of roadway segment safety performance based on the calculated safety performance indices, rural roadway speeding map, hotspot identification for pedestrian and vehicles, graphs and tables of crash statistics that support crash reporting, etc. Specifically, the following visualization functions had been developed for the crash data and crash modeling results.

- Point-based crash visualization

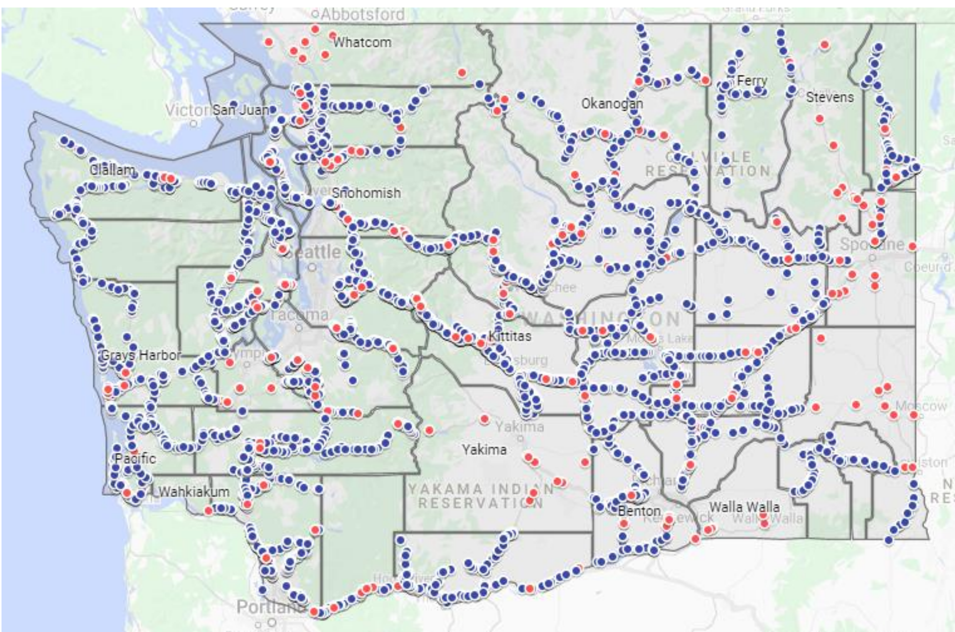


Figure 3.1. Point-based crash visualization

- Segment-based safety index visualization

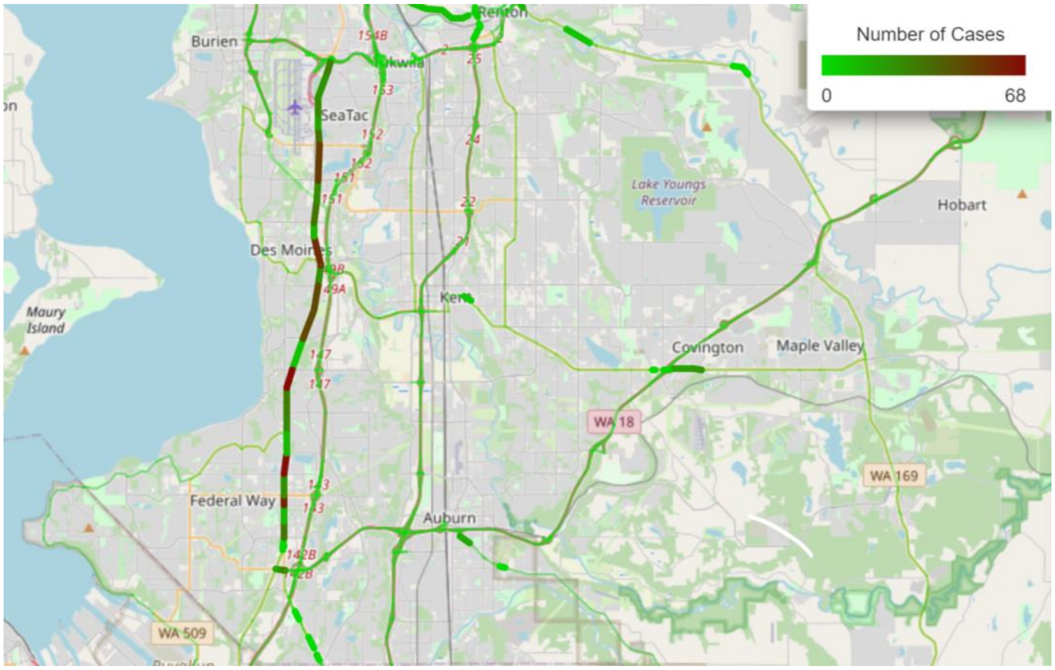


Figure 3.2. Segment-based safety index visualization

- Area-based safety index visualization

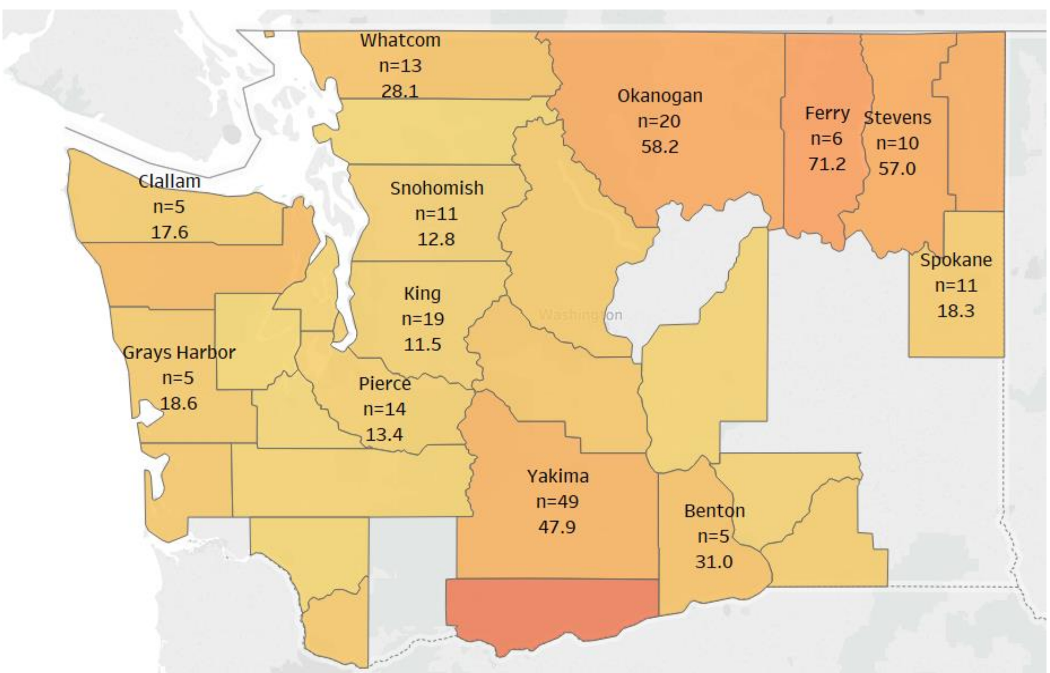


Figure 3.3. Area-based safety index visualization

- Crash heatmap

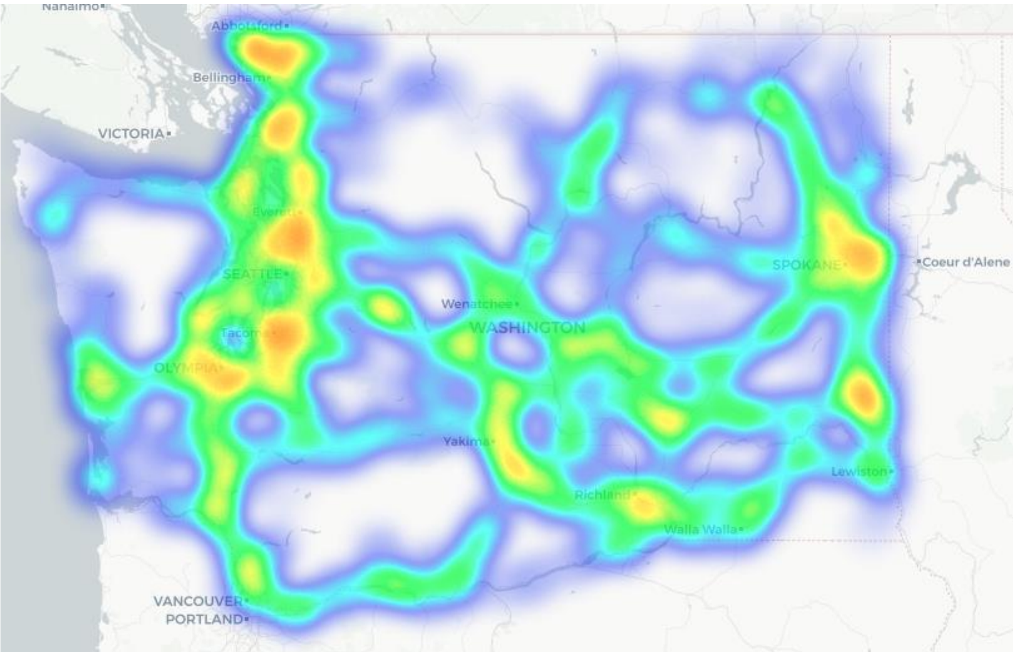


Figure 3.4. Crash heatmap

- Safety report

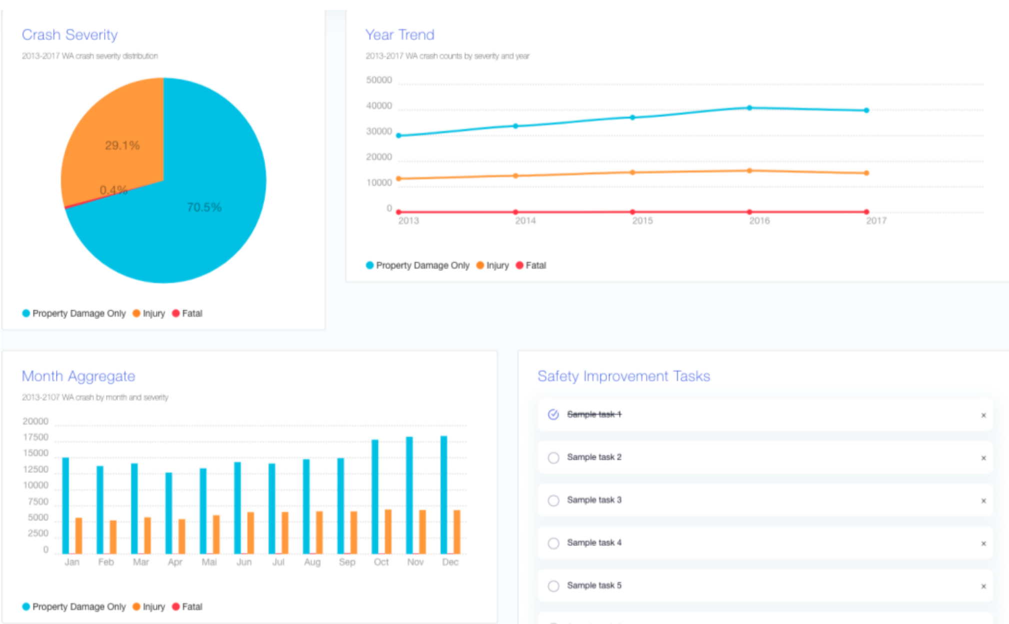


Figure 3.5. Safety report

The safety assessment methods coupled with powerful visualization could assist the decision makers by transferring data analysis results into actionable insights. With the help of the safety platform, RITI communities could obtain funding based on the analytical results of the safety tool, make effective use of the resources utilizing the safety tool to prioritize the high-risk roadway segments, and apply countermeasures based on the analysis of risk factors of the selected roadway sections. As many state and local agencies, especially in RITI communities, are experiencing similar safety issues, the analytical methods and tool developed in this project can be modified to support these agencies as well.

CHAPTER 4. ROADWAY SAFETY ASSESSMENT METHODS FOR RITI COMMUNITIES

It is well known that some of the differences of roadway geometric characteristics between roads in rural and urban areas, such as lane width, shoulder width and type, number of curves and curve rate, grade, etc. contribute to the higher fatality rate in rural areas, especially in the RITI communities. Human behavior factors, such as driving under the influence (DUI), speeding, unrestrained driver/occupant, and distracted driver, etc. could lead to serious injuries and fatalities. In addition, adverse weather conditions, such as fog, snow, ice, and severe rain on rural roads will lead to dangerous driving conditions and increase the risk of having crashes. While most rural road traffic safety related studies have focused on major facilities, few efforts have been made to investigate the crash injuries on low-volume roads. These less-travelled roads do not expect to serve high volumes of vehicles, and are often built with lower geometric standards. Nonetheless, drivers or pedestrians still have the same right to travel on these roads safely. In order to help decision makers better understand the influential factors on rural road safety and implement effective safety countermeasures accordingly, this project aims to investigate the potential influential factors, such as roadway geometric characteristics, environmental conditions, human behaviors, traffic conditions, etc., on the injury severity of crashes that occurred on low-volume rural roads.

4.1. Crash Injury Analysis

This project investigated the potential influential factors, such as roadway geometric characteristics, environmental conditions, human behaviors, and traffic conditions as shown in Table 4-1, on the injury severity of crashes occurred on low-volume rural roads. Four models, including ordered probit (OP), multinomial logit (MNL), artificial neural network (ANN) and random forest (RF), were trained, tested, and validated using five years (2013to 2017) of Washington State crash records on low-volume rural roads. It was found that the two Machine Learning models (ANN and RF) performed better than the two statistical models (OP and MNL), and the RF model had the best performance in predicting crash injury severities. The results also showed that variables such as grade percentage, degrees of curvature, shoulder width, driver's gender, roadway width, head on crash, pedestrian/cyclist involved, young driver, and truck involved, have significant impact on the crash injury severities on low-volume rural roads.

All the models applied in this project were developed using Python and libraries such as Scikit-Learn (Pedregosa et al., 2011) and TensorFlow (Abadi et al., 2016). The performance of the four models were measured using the accuracy scored calculated in **Equation 1:**

$$Accuracy = \frac{T_p+T_n}{T_p+T_n+F_p+F_n} \qquad (1)$$

Where T_p is true positive, T_n is true negative, F_p is false positive, and F_n is false negative.

Table 4-2 shows the accuracy score of the four models. Despite the fact that the OP model can capture the ordinal nature of the crash injury severities, the MNL method in this study outperformed the OP model in prediction accuracy. This is most likely because the crash injury severity variables for the OP model are the same while the MNL method had different variables to predict each crash injury severity category. The accuracy scores also indicated that the ML methods had better performance than the statistical method, and the RF model had the overall best performance in predicting crash injury severity on low-volume rural roads in this research.

Table 4-1. Selected variables

Variable		**Definition**	**Range/Categories**
Crash Injury Severity		Crash injury severities	Fatal, Injury, PDO
Roadway Geometrics	Road Surface Material	Surface material type	Asphalt, Bituminous, Gravel, Portland Concrete Cem, Soil, Other
	Lane Width	Calculate lane width: calculated by dividing the total roadway width by the total number of lanes	Continuous, in ft
	Roadway Width	Total roadway width for the roadway segment	Continuous, in ft
	Degree of Curvature	Degree of curvature for the curve: calculated from curve radius	Continuous, in ft
	Left Shoulder Width	The width of the inside (left) shoulder of road in feet in the increasing direction of the roadway.	Continuous, in ft
	Right Shoulder Width	The width of the outside (right) shoulder road in feet in the increasing direction of the roadway.	Continuous, in ft
	Grade Percentage	Percent grade for this roadway segment	Continuous, in %
Vehicle Information	Truck	If the involved vehicle is truck	Yes, No
	Old Car	If the involved vehicle was more than 15 years old at the time of crash	Yes, No
Traffic Characteristics	AADT	Calculated Annual average daily traffic (AADT)	Integer
	Truck percentage	Truck percentage for the roadway segment	Continuous, in %
	MVMT	Million vehicle miles traveled on road segment	Continuous, in veh-mile

Table 4-2. Model performance

Model	**Accuracy**
OP	0.47
MNL	0.67
ANN	0.72
RF	0.79

To identify the variables that have the most predictive power, variable importance is calculated and ranked for each variable in the final RF model. The variable importance is computed as the impurity decrease weighted by the probability of reaching that node. In this study, the Scikit-Learn library was used to calculate the impurity-based feature importance. As shown in Figure 1.1, the variable importance rankings indicate that variables such as grade percentage, degrees of curvature, shoulder width, driver's gender, roadway width, head on crash, pedestrian/cyclist involved, young driver, truck involved, all have a significant impact on the crash injury severities on low-volume rural roads.

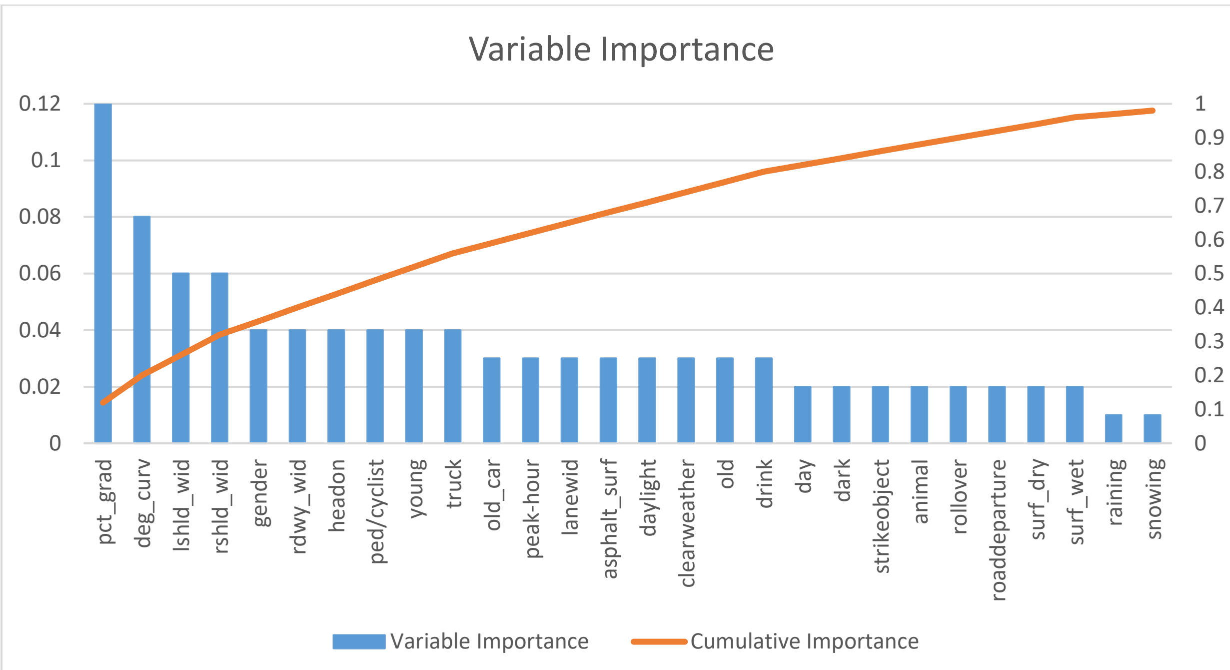


Figure 4.1. Variable importance ranking for significant factors

Roadway Geometrics

According to the analysis, roadway geometric factors such as grade percentage, degrees of curvature, left and right shoulder width, roadway width, lane width, and road surface material, were among the most significant factors for crash injury severity. Given their functional class and geometric design standards, many of the rural roads with low traffic volume are narrow and without shoulders. Some of these roads tend to have sharp curves and steep hills. Many such roadways in the rural area have rough/no pavement, which could be dangerous to travel, especially under adverse weather conditions. With the random nature of crash events and the low traffic volume of these rural facilities, it is difficult to identify the crash hotspots compared to the higher volume roads. In this case, it is critical to provide geometric design guidelines based on the unique characteristics of the low-volume rural roads, as well as using signs and markings to improve the safety conditions.

Driver Characteristics

Driver's characteristics including gender, age (young and older drivers), and behavior (DUI) were also found to have significant impact on crash injury severity on low-volume rural roads. Other researchers, such as Souleyrette et al. (2010) also recognized that impaired driving, younger or older driver to be increasing the severity of crashes on low-volume rural roads. This is evident as not only the human behavior factors are always significant towards traffic safety, but also the relatively challenging geometrics conditions of the low-volume rural facilities pose more requirements on the driver's ability to travel safely.

Crash Type and Vehicle Information

Crash types, especially head on crashes and pedestrian/cyclist involved crashes were found to be important in the crash injury severities prediction models. Head on crashes and pedestrian/cyclist involved crashes have a much higher rate of injury and fatality compared to other crash types. In addition, if the involved vehicle is a truck and if the vehicle is older than 15 years were recognized as important contributing factors as well. While the majority of pedestrian related injuries and fatalities occurred in the urban areas because of higher population rate, pedestrians and cyclists are exposed to many risks in the rural areas, especially when travelling on the low-volume facilities with poor infrastructure and lack of regulations towards pedestrian safety. As trucks require higher levels of visibility conditions and road geometrics to ensure driving safety, the sharp curves, steep grades, and narrow road width in low-volume rural facilities cause blind spots to truck drivers. The implementation of signs or sensors at the risky locations to warn the truck drivers and pedestrians could help prevent crashes from happening. Several other crash types, such as rollover, animal strikes and other objects and road departure, were also found to have significant impact.

Environmental Conditions

The model results indicated that the light conditions have significant impact on crush injury severities. A study by Abdel-Aty (2003) also found that dark lighting conditions contribute to higher probability of roadway injuries. Plainis et al. found that in the United Kingdom, the crash fatality rate is higher during nighttime than daytime (Plainis et al., 2006). In this case, it is important for decision makers to enhance the lighting conditions on rural roads, especially during the nighttime. In addition, weather conditions such as rain and snow were also found to be significant as well as road surface conditions (dry/wet). Wet road surface could be caused by weather conditions such as snow, ice and rain.

4.2. Pedestrian Safety Analysis

Four classification techniques were applied to assess how roadway features mainly correlate to pedestrian fatal crashes: Logistic Regression, Nearest Neighbor Classification, Decision Tree, and Random Forest Classifier. Each of the four modeling approaches was implemented using K-fold cross-validation, a process that allows choosing the best parameters for the model. Their results were evaluated and then compared in terms of accuracy score and confusion matrices for the testing data set. It was found that the Decision tree had consistent results and the best performance among all models, showing how the distinct predictors relate to each other to predict fatal pedestrian collisions. This project focuses on specific roadway factors that precipitate fatal pedestrian collisions. Specifically, we aimed to predict the severity of a pedestrian collision based upon the existing roadway characteristics at the location of the collision. This allows practitioners to pinpoint the locations where the highest severity of pedestrian collision is likely to occur so that they can prioritize locations and treatments for pedestrian safety countermeasures.

For this study, four different Machine Learning classification techniques were used to assess how roadway features and some other factors correlate to pedestrian fatal collisions. An overview of each of these modeling techniques is presented below.

Logistic Regression

Logistic regression is a classical method for binary classification whose parameters are estimated by maximizing the likelihood estimation through the following equation:

$$log \frac{Pr\ (y=1)}{1-Pr(y=1)} = \beta_0 + \sum_{i=1}^{p} \beta_i x_i \qquad (2)$$

Where: Pr represents the probability of a sample belonging to class 1, β_0 = intercept and β_i = coefficients. The expression $\frac{Pr\ (y=1)}{1-Pr(y=1)}$ corresponds to the odds, where $log \frac{Pr\ (y=1)}{1-Pr(y=1)}$ represents the log odds. The goal is to predict the log-odds, which is converted to probability through the logistic function.

Logistic regressions can also have a penalty term related to the model complexity. It is represented for the hyperparameter C that controls the inverse of model complexity (smaller values imply stronger regularization). The hyperparameter selection was made using the k-fold cross-validation method, where the training data is split into k groups to select the best value of C. The value of k usually varies from 5 to 10, and due to the data size used in this study, k = 5 was selected. Cross-validation is also an appropriate technique to avoid overfitting issues, a Machine Learning sign of poor performance occurs when the model fits perfectly the training data set, including its noise or outliers.

We built an initial logit regression with all the variables having Fatal as the response variable. Results were then analyzed to verify the significance of each of the variables' coefficients. As a common practice, a level of significance of 0.05 was established, so that each variable with a statistical p-value less than 0.05 was considered significant. A second logistic regression model was built using only the significant variables identified in the initial model.

The following table presents the coefficients for the final logistic model using k-fold cross-validation, where all the variables are significant to a level of 0.05, except for the "rural" variable (p-value = 0.1272). However, since we are equally interested in understanding whether rural areas may have a distinct impact on pedestrian fatal collisions compared to urban areas, we decided to keep this variable in the model.Table 4-3. Results for the Logistic regression are the final set of variables used in all the other models that will be presented below.

Table 4-3. Results for the Logistic regression

Variable	**Coefficient**
Intercept	-0.0026
SPD_LIMT	-0.0124
LANEWID	-0.0273
RSHLDWID	0.1999

Variable	Coefficient
AADT	0.0000
TRKPCTS	0.0706
RURAL	-0.0313
LIGHT_DAYLIGHT	-0.3674
LIGHT_DARKLIGHTSON	-0.2014
LIGHT_DARKNOLIGHT	0.0002
WEATHER_CLEAR	-0.2106
FREEWAY	0.0142
2-LANEROAD	-0.0580
MULTILANE_NON-FREWAY	-0.1868
Penalty value: 5.0	
Accuracy on training data set: **0.759**	
Accuracy on testing data set: **0.826**	

The results show noteworthy insights about the relationship between each variable and their impact on the occurrence of fatal crashes involving pedestrians. SPD_LIMIT and LANEWID have surprisingly negative impacts on the outcome, suggesting that roadways with higher speed limits and lane width tend to be related to fewer fatal collisions. This outcome should be interpreted with caution though, particularly because this analysis does not include features that may be linked to these variables' effects. For example, demographic predictors are not present in this model.

On the other hand, RSHLDWID and TRKPCTS all have positive coefficients, which may indicate that roads with wider right shoulders and a higher percentage of trucks are more likely to have pedestrian fatal collisions. Rural roads seem to be less likely to have fatal crashes compared to urban roads, whereas daylight periods have the most negative impact on the response variable. However, dark periods when street lights are off or absent are more likely to generate fatal crashes, which intuitively makes sense. As expected, days with clear or partly cloudy weather are less likely to have lethal collisions, and freeways are more likely to be associated with this type of crash, which may be explained by their general higher volume compared to other roads (although AADT has a positive coefficient, its value is practically equal to zero).

Regarding model accuracy, we observe that the testing data set has a higher accuracy when compared to the training data set (0.826 against 0.759, respectively). Since the model is fitted using the training data and then used to predict testing samples, having a higher accuracy on the testing data set does not seem to be conceivable. This may indicate that logistic regression is not a suitable model type for this study.

Nearest Neighbor Classification

Nearest Neighbor Classification is a method that uses class labels of the K nearest neighbors to determine the class label of an unknown record using proximity metrics to calculate distance/similarity between records. The number of nearest neighbors (K) is a hyperparameter that must be provided, along with the distance metric (Minkowski distances are usually used). Choosing the values of K can be difficult, since a too-small K may lead to neighborhoods that are sensitive to noise points, whereas a too-large K may make a neighborhood include points from other classes. K-fold cross-validation is an effective method to handle this issue, where distinct values of K can be analyzed using the training data set, as well as distinct values for parameters such as the exponent factor "p" for the Minkowski distance and the weights associated with the distances. A final model with the best parameters is then fitted and can be applied to the testing data set. Therefore, a third model using Nearest Neighbor Classification with k-fold cross-validation was built using the same set of significant variables applied to the second logistic regression.

The following figure is a dashboard with the results for the Nearest Neighbor Classification. It shows the best value of parameters selected after cross-validation using Minkowski as the distance metric: number of neighbors, the exponent factor "p" for the Minkowski distance, and the criteria of weights (two possible options: uniform and distance). The table also presents the accuracy scores for the training and testing data sets, in addition to the confusion matrix (right column) showing the right and wrong predictions for the testing data set.

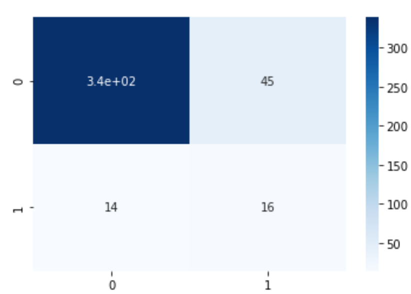
Best parameters N. neighbors: 3 p: 1 Metric: Minkowski Weights: Distance	
Accuracy on training data set: **0.997** Accuracy on testing data set: **0.857**	

Figure 4.2. Dashboard with the results for the Nearest Neighbor Classification

We observe from Figure 4.2 that the accuracy of the training data set is higher than the testing data set, which is reasonable. However, the accuracy for the training data is practically equal to 1, almost a "perfect" fit to the training data. As previously mentioned, this is a sign of overfitting, which suggests that this model, even after cross-validation, may not represent a good fit for the studied data set. Like the logistic regression, the confusion matrix for testing predicted values shows that the model is biased predicting 0 values (non-fatal collisions).

Decision Tree

Decision trees are another effective tool to handle classification problems. The goal is to classify data (leaf nodes of the tree) from the characteristics of the predictor variables (decision nodes). If D_t is the set of training data points reaching a node t, two options exist: if D_t contains data points that belong to the same class y_t, then t is a leaf node labeled as y_t; if D_t contains records that belong to more than one class, we should use an attribute test to split the data into smaller subsets and recursively apply the procedure to each subset. However, early terminations are often applied to stop the splitting procedure to avoid overfitting issues.

To define the best split, we follow the Greedy approach which establishes that nodes with “purer” class distribution are preferred, i.e. nodes with samples mainly distributed towards one of the classes. We applied the Gini Index as a measure of node impurity, computed as follows:

$$GINI\ (t) = 1 - \sum_j\ [p(j|t)]^2 \qquad (2)$$

Where $p(j|t)$ is the relative frequency of the class j at node t. GINI is maximized when the points are equally distributed among all classes, showing the least interesting information. The minimum (0) occurs when all records belong to one class, indicating the most interesting information for that node.

To avoid overfitting and to select an appropriate number of parameters such as minimum samples per leaf nodes and the maximum depth, we applied k-fold cross-validation as was done for the other models to choose the optimal hyper-parameters (always with k=5). A new model was thus fitted using the same set of final variables that were applied for the previous two models.

The results of fitting a decision tree using cross-validation are summarized in Figure 4.3. We can observe that two parameters were tested during the cross-validation process in order to get the best modeling performance: the maximum depth of the three (i.e. the number of horizontal levels of a top-down tree from its root) and the minimum samples in leaf nodes. Additionally, there are the accuracy scores for the training and testing data sets, as well as the confusion matrix with predictions for the testing data.

Best parameters

Maximum depth: 5

Minimum samples in leaf node: 5

Accuracy on training data set: **0.885**

Accuracy on testing data set: **0.874**

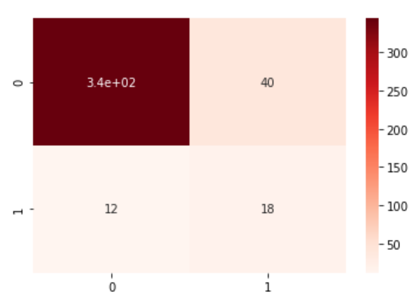


Figure 4.3. Dashboard with the results for the Decision Tree

The results of Figure 4.3 show that the best tree that fits the training data has a maximum depth of 5 levels, with a minimum of 5 samples for the leaf nodes. These parameters help to avoid overfitting issues, which can be seen for the accuracy score of the training data set: 0.885 (a high value not too close to 1). The accuracy score for the testing data (0.874) is also high and slightly less than the training’s score, which is a good performance indicator for this model. However, the confusion matrix once again shows that the model is biased to predicting 0 values (non-fatal collisions).

Random Forest Classifier

Random Forest classifiers are part of the so-called Ensemble Methods, ML classification techniques aiming at building a set of base classifiers from the training data set and predicting the class label of test records by combining the predictions made by all base classifiers (through majority vote). Ensemble methods also aim to reduce the variance of complex models by aggregating responses of multiple base classifiers.

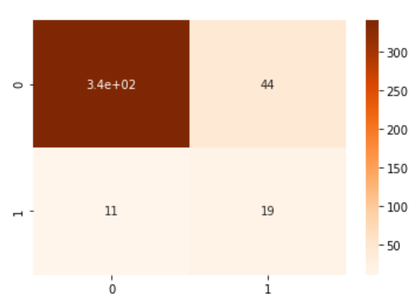
Ensemble methods generally need independent base classifiers, and Random Forest techniques are well aligned with this. They fit a full decision tree by randomizing which predictors would be available for a given node, which alleviates the split on similar predictors for bagged trees.

We developed a final model by applying a Random Forest Classifier for the same final set of significant variables used in the previous models. Likewise, we used k-fold cross-validations to select parameters needed for this method, such as maximum depth for the trees and the number of trees in the forest ("number of estimators").

All the models in this study were evaluated and compared in terms of the accuracy score for the training and testing data sets. Accuracy scores vary from 0 to 1, and values close to one denote effective predictions. However, models with accuracy scores approximately equal to 1 may indicate overfitting issues and are not suitable. Additionally, we applied confusion matrices to the testing data of each of the models to evaluate their level of predictions. A confusion matrix is a representation of the classification rate for a classifier method and has 4 quadrants indicating the number of right and wrong predictions for the class.

We compared each model to verify which one performed better when predicting the occurrence of fatal collisions involving pedestrians. The following table summarizes the best parameters related to the Random Forest Classifier after cross-validation (maximum depth and number of estimators), as well as the accuracy score for the training and testing data sets and the confusion matrix for the predicted tested values.

Table 4-4. Dashboard with the results for the Random Forest Classifier

<table>
<tr><td>Best parameters
Maximum depth: 5
Number of estimators: 100</td><td rowspan="2">
</td></tr>
<tr><td>Accuracy on training data set: 0.883
Accuracy on testing data set: 0.867</td></tr>
</table>

The two classification techniques that best fit the data and have consistent and high accuracy scores for the training and testing data sets are the Decision Tree and the Random Forest Classifier. Since both models are based on building classifier trees and have similar parameters, we would expect them to have related performances. Nevertheless, the Decision Tree has a slightly higher accuracy score on the testing data (0.874) when compared to the same metric for the Random Forest Classifier (0.867), thus this is the model with the best performance among all. Furthermore, we noted from the results that the confusion matrices for all the models were alike: even though they have substantially more correct than incorrect predictions, they are all biased to predict 0 values (non-fatal collisions). Since this is happening with all the models, we believe that is derived from the dataset itself and its errors rather than specificities related to any of the modeling approaches used in this study.

With the best performance among all models, the Decision Tree represents how the combination of predictor variables leads to fatal or non-fatal accidents involving pedestrians. However, not all the predictor variables are used as decision nodes for the final tree, and this is due to the selected parameters after the cross-validation process, particularly for the maximum depth. Indeed, since we have 12 final predictor variables, a tree having almost all of them used as decision nodes would require a higher depth, but large-sized trees are generally not easy to interpret and may likely lead to overfitting. In this case, the Machine Learning algorithm searches for the set of variables that mostly impact the outcome for the selected tree depth.

CHAPTER 5. ROADWAY SAFETY COUNTERMEASURES ANALYSIS AND RECOMMENDATION

5.1. Introduction

In addition to developing the safety assessment methods, this project also investigated the safety countermeasures for RITI communities. Safety countermeasures are developed and implemented to reduce crash frequency and accident severity on road systems. Concerning the rural, isolated, tribal and indigenous (RITI) communities, this is even more critical. Studies indicated that crashes involving pedestrians in RITI communities often lead to severe injuries or fatalities (Marshall and Ferenchak, 2017; Baireddy et al., 2018; Chen and Fan, 2019). The lack of accommodation, such as sidewalks, marked crosswalks, lighting conditions, and traffic control, makes pedestrians and bicyclists at a high disadvantage in these RITI communities. Over seventy percent of pedestrian fatalities on tribal lands occurred in the rural areas (Awwad-Rafferty et al., 2019), where approximately seventy-five percent of them happened at night. Additionally, aggravating cultural and human behaviors such as speeding, driving under the influence, and pedestrian behavior make the problem even worse.

Specifically, the research team has reached out to RITI communities and established a formal research partnership with the Yakama Nation. By working with the traffic engineers and planners in the Yakama Nation DNR Engineering Department, the research team has conducted research on safety countermeasures analysis and recommendation for RITI communities.

5.2. Crash Modification Factors (CMFs)

Use of Crash Modification Factors (CMFs) is a common way of assessing the efficacy of safety countermeasures. According to the CMF Clearinghouse (Crash Modification Factors Clearinghouse, accessed 03/23/2021), a Crash Modification Factor is “a multiplicative factor that indicates the proportion of crashes that would be expected after implementing a countermeasure” (installing a traffic signal or a median barrier, increasing the width of edgelines, etc). CMFs with values less than 1.0 indicate expected decreases in crashes, whereas values greater than 1.0 indicate a likely increase in crashes. Harkey et al. (2008), as cited by CMF Clearinghouse, gives a practical example for using the CMF: a specific stop-controlled intersection is expected to have 5.2 total crashes per year, so a traffic signal is planned to be installed. The CMF for installing the traffic signal is estimated at 0.56 for the total of crashes, thus the expected total crashes after installing the signal would be 5.2 x 0.56 = 2.9 (total crashes per year).

The Crash Reduction Factor (CRF) is a distinct way of evaluating the effect of countermeasures related to the expected decrease in crashes and can be calculated as:

$$CRF = 100\ (1 - CMF) \tag{3}$$

Although some transportation agencies in the United States have been utilizing CRFs, the use of CMF has been encouraged in recent years for the safety field. This is due to interpretative confusions that may be raised since CRFs can present negative reduction values for CMFs larger than 1, which actually indicates expected crashes increasing.

Figure 5.1 shows how distinct countermeasures can be visualized and compared on the CMF Clearinghouse page for distinct categories.

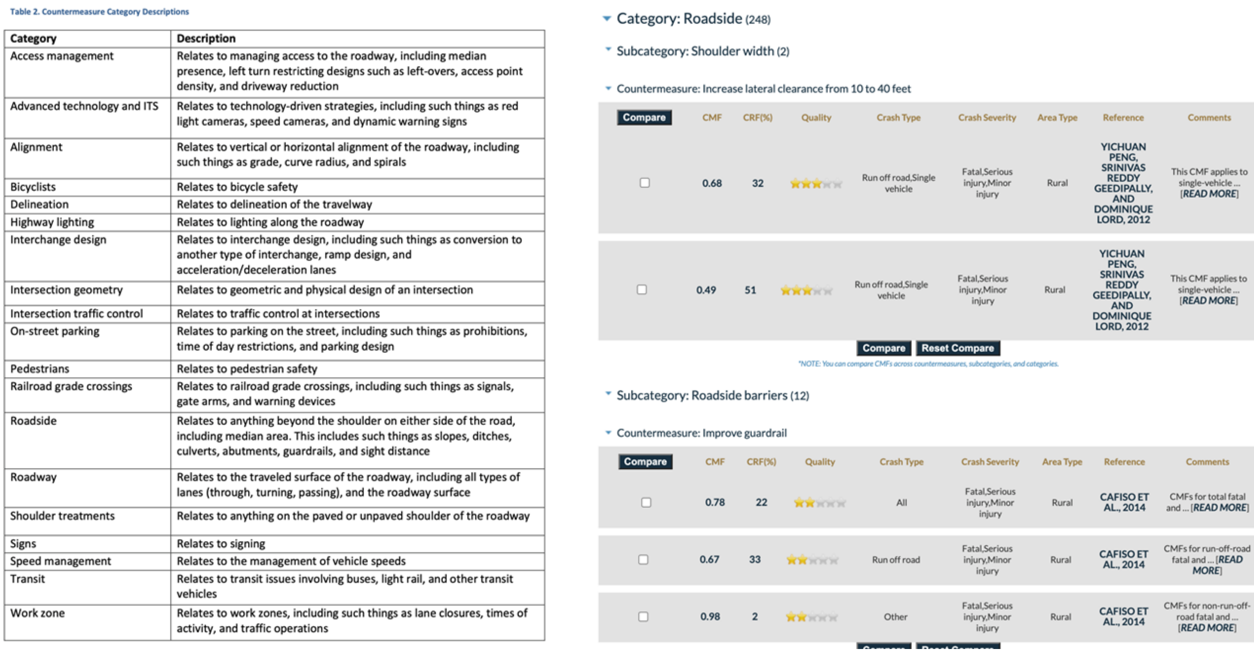


Table 2. Countermeasure Category Descriptions

Category	Description
Access management	Relates to managing access to the roadway, including median presence, left turn restricting designs such as left-overs, access point density, and driveway reduction
Advanced technology and ITS	Relates to technology-driven strategies, including such things as red light cameras, speed cameras, and dynamic warning signs
Alignment	Relates to vertical or horizontal alignment of the roadway, including such things as grade, curve radius, and spirals
Bicyclists	Relates to bicycle safety
Delineation	Relates to delineation of the travelway
Highway lighting	Relates to lighting along the roadway
Interchange design	Relates to interchange design, including such things as conversion to another type of interchange, ramp design, and acceleration/deceleration lanes
Intersection geometry	Relates to geometric and physical design of an intersection
Intersection traffic control	Relates to traffic control at intersections
On-street parking	Relates to parking on the street, including such things as prohibitions, time of day restrictions, and parking design
Pedestrians	Relates to pedestrian safety
Railroad grade crossings	Relates to railroad grade crossings, including such things as signals, gate arms, and warning devices
Roadside	Relates to anything beyond the shoulder on either side of the road, including median area. This includes such things as slopes, ditches, culverts, abutments, guardrails, and sight distance
Roadway	Relates to the traveled surface of the roadway, including all types of lanes (through, turning, passing), and the roadway surface
Shoulder treatments	Relates to anything on the paved or unpaved shoulder of the roadway
Signs	Relates to signing
Speed management	Relates to the management of vehicle speeds
Transit	Relates to transit issues involving buses, light rail, and other transit vehicles
Work zone	Relates to work zones, including such things as lane closures, times of activity, and traffic operations

Category: Roadside (248)

Subcategory: Shoulder width (2)

Countermeasure: Increase lateral clearance from 10 to 40 feet

Compare	CMF	CRF(%)	Quality	Crash Type	Crash Severity	Area Type	Reference	Comments
	0.68	32		Run off road,Single vehicle	Fatal,Serious injury,Minor injury	Rural	YICHUAN PENG, SRINIVAS REDDY GEEDIPALLY, AND DOMINIQUE LORD, 2012	This CMF applies to single-vehicle ... [READ MORE]
	0.49	51		Run off road,Single vehicle	Fatal,Serious injury,Minor injury	Rural	YICHUAN PENG, SRINIVAS REDDY GEEDIPALLY, AND DOMINIQUE LORD, 2012	This CMF applies to single-vehicle ... [READ MORE]

Compare Reset Compare

*NOTE: You can compare CMFs across countermeasures, subcategories, and categories.

Subcategory: Roadside barriers (12)

Countermeasure: Improve guardrail

Compare	CMF	CRF(%)	Quality	Crash Type	Crash Severity	Area Type	Reference	Comments
	0.78	22		All	Fatal,Serious injury,Minor injury	Rural	CAFISO ET AL., 2014	CMFs for total fatal and ... [READ MORE]
	0.67	33		Run off road	Fatal,Serious injury,Minor injury	Rural	CAFISO ET AL., 2014	CMFs for run-off-road fatal and ... [READ MORE]
	0.98	2		Other	Fatal,Serious injury,Minor injury	Rural	CAFISO ET AL., 2014	CMFs for non-run-off-road fatal and ... [READ MORE]

Figure 5.1. Safety countermeasures examples from CMF Clearinghouse. Source: CMF Clearinghouse (03/01/2021)

5.3. Safety Improvements on Rural Roads

5.3.1. Avoiding Vehicle-Vehicle Crashes

As previously mentioned, safety countermeasures can be classified according to distinct categories. Each one has its specificities and parameters, such as CMF, CRF, and Benefit-Cost ratios. Examples for some categories are listed below (National Center for Rural Road Safety, 2016):

- **Road alignment**
  - Horizontal alignment signs
  - Flashing beacons
  - Chevrons
  - Post mounted delineators
  - Raised pavement markers
- **Cross Section**
  - Widen lanes
  - Widen shoulders
  - Adding shoulders
  - Stabilizing shoulders
  - High friction surface treatments

- **Roadside Features**
    - Flatten side slopes
    - Install safety edge
    - Object markers
    - Relocate objects
    - Remove objects
- **Miscellaneous**
    - Shoulder rumble strips/stripes
    - Centerline rumble stripes
    - Edge-line markings
    - Centerline markings
    - Widen edge-line markings
    - Widen centerline markings

Additionally, rural collectors and rural local roads are within the scope of High-Risk Rural Roads by the Manual for Selecting Safety Improvements on High Risk Rural Roads (Atkinson et. al., 2014)**.** This manual presents information about cost and benefits as well as the CMF of safety treatments on high-risk rural roads. The manual is organized by roadway feature type, such as horizontal curves, intersections, roadside, signing, etc., providing a treatment matrix for each treatment presented. This matrix presents an overview of benefits and costs related to each safety countermeasure and can be used to compare them. An example of a treatment matrix for pavement and shoulder resurfacing is shown in Figure 5.2.

SAFETY TREATMENT	For more information, visit page	COST		Frequency of Maintenance (years)	SAFETY BENEFIT		BENEFIT-COST RATIO[30]			
		Initial Implementation	Ongoing Maintenance		NCHRP 500 Performance Rating	Crash Modification Factor (CMF)	Lower Volume*, Optimal Conditions***	Higher Volume**, Optimal Conditions***	Lower Volume*, Narrower Conditions****	Higher Volume**, Narrower Conditions****
Install a Safety Edge	86	$	–	20	P	0.85-0.92	33.4	267.2	40.9	403.2
Install Center Line Rumble Strips	87	$	–	10	P	0.75-0.85	21.3	170.6	26.1	257.5
Install Edge Line or Shoulder Rumble Strips	88	$	–	10	P	0.78-0.90	58.6	469.0	71.8	707.7
Install Transverse Rumble Strips	89	$			P	0.76-0.91				
Regrade or Recondition Gravel Lanes	90	$-$$			T					
Install Targeted Longitudinal Rumble Strips at Key Locations (Such as on the Outside of Horizontal Curves Only)	91	$$-$$$			T	0.85				
Install or Maintain a Graded Shoulder	92	$$-$$$			P	0.52				
Provide Turnout Areas	93	$$-$$$			T					
Improve Pavement Friction/Increase Skid Resistance	94	$$$$	–	10	P	0.25-0.60	3.3	26.7	4.1	40.3
Add Paved Shoulder	95	$$$$$	$$	2	P	0.86	n/a	n/a	0.5	4.5
Widen Existing Travel Lanes by Two Feet or Less per Lane	96	$$$$$	$$$	10	P	0.95	n/a	n/a	0.3	2.8
Install Passing or Climbing Lanes	97	$$$$$	$$$	10	P		0.3	2.3	0.4	3.5
Increase Shoulder Width	98	$$$$$			P	0.90-0.97				
Improve Superelevation at Horizontal Curve Locations	99	$$$$$			P					

Cost:
$ = $0 to $5,000
$$ = $5,001 to $20,000
$$$ = $20,001 to $50,000
$$$$ = $50,001 to $100,000
$$$$$ = $100,001 and up

NCHRP 500 Performance Rating [31]
P – Proven
T – Tried
E – Experimental
U – Unknown

*Lower Volume ≤1000 vpd
**Higher Volume = Between 1,001 and 8000 vpd
***Optimal Conditions = 12-foot lanes, 6-foot paved shoulders
****Narrower Conditions = 10-foot lanes and no shoulders

Figure 5.2. Treatment matrix for pavement and shoulder resurfacing (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Some examples of countermeasures with high benefit costs ratios related to pavement and shoulder resurfacing are presented as follows.

Edge line (or Shoulder rumble strips) and Center Line Rumble Strips

Rumble strips provide both an audible warning and a physical vibration that alerts drivers they are leaving the driving lane. The treatment matrices for these types of safety countermeasure are presented as follows[1].

---

[1] Definition from the FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014:
Lower Volume ≤1000 vpd
Higher Volume = Between 1,001 and 8000 vpd
Optimal Conditions = 12-foot lanes, 6-foot paved shoulders
Narrower Conditions = 10-foot lanes and no shoulders

NCHRP 500 Performance. Proven: The safety effect for other similar applications has shown a proven benefit. Tried: The treatment has indications that it can be expected to reduce crashes, but has some conflicting reports as to its associated safety effects or has been deployed and observed to be effective. Experimental: New treatments that still need to be tested and for which the safety effect is unknown. Unknown: Not enough is known about an associated safety performance.

Install Edge Line or Shoulder Rumble Strips - Initial Investment: $3,000 - Cost of Maintenance: n/a - Frequency of Maintenance: 10 years (2 applications)	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume Optimal Conditions	58.6	Proven	0.78–0.90
Higher Volume Optimal Conditions	469.0	Proven	0.78–0.90
Lower Volume Narrower Conditions	71.8	Proven	0.78–0.90
Higher Volume Narrower Conditions	707.7	Proven	0.78–0.90

Figure 5.3. Treatment matrix for installing Edge line or Shoulder rumble strips (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Install Center Line Rumble Strips - Initial Investment: $5,000 - Cost of Maintenance: n/a - Frequency of Maintenance: 10 years (2 applications)	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume Optimal Conditions	21.3	Proven	0.75–0.85
Higher Volume Optimal Conditions	170.6	Proven	0.75–0.85
Lower Volume Narrower Conditions	26.1	Proven	0.75–0.85
Higher Volume Narrower Conditions	257.5	Proven	0.75–0.85

Figure 5.4. Treatment matrix for installing Center Line Rumble Strips (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

- Edge line: On roads with a history of departure crashes. Center Line Rumble Strips: Any roads, especially those with a history of head-on crashes.
- “For all rumble strips, pavement conditions should be sufficient to accept milled rumble strips”. (Atkinson et. al., 2014)
- “Rumble strips should be provided on all new rural freeways and on all new rural two-lane highways with travel speeds of 50 mph or greater”. (Atkinson et. al., 2014)

<u>Safety Edge</u>

This treatment aims at minimizing drop-off related crashes by sloping the pavement edge at an angle, so the driver can safely reenter the road after driving onto the shoulder.

Install a Safety Edge - Initial Investment: $2,145 - Cost of Maintenance: n/a - Frequency of Maintenance: 20 years	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume Optimal Conditions	33.4	Proven	0.85–0.92
Higher Volume Optimal Conditions	267.2	Proven	0.85–0.92
Lower Volume Narrower Conditions	40.9	Proven	0.85–0.92
Higher Volume Narrower Conditions	403.2	Proven	0.85–0.92

Figure 5.5. Treatment matrix for installing Safety Edge (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

- "Each State should implement policies and procedures to incorporate the Safety Edge where pavement and non-pavement surfaces interface on all paving and resurfacing projects with surface differentials of 2.5 inches or more". (Atkinson et. al., 2014)
- The Safety Edge is properly used at spots where pavement edge drop-offs occur through everyday use, which is the case of rural roads with unpaved shoulders.

The National Center for Rural Road Safety, founded by the Federal Highway Administration in 2014, is also focused on improving safety on rural roads by "supporting local, state, and tribal road owners and their stakeholders" (National Center for Rural Road Safety web page, accessed 03/23/2021). They publish several analyses regarding rural road safety on their web portal, where the presentation "Crash Risk Factors for Low-Volume Roads: an ODOT Case Study" (2016) shows an evaluation of low-cost safety countermeasures that can be implemented on low-volume road roads in terms of benefit-cost (B/C) ratios. The following figure summarizes the overall B/C ranges for several safety countermeasures on rural roads:

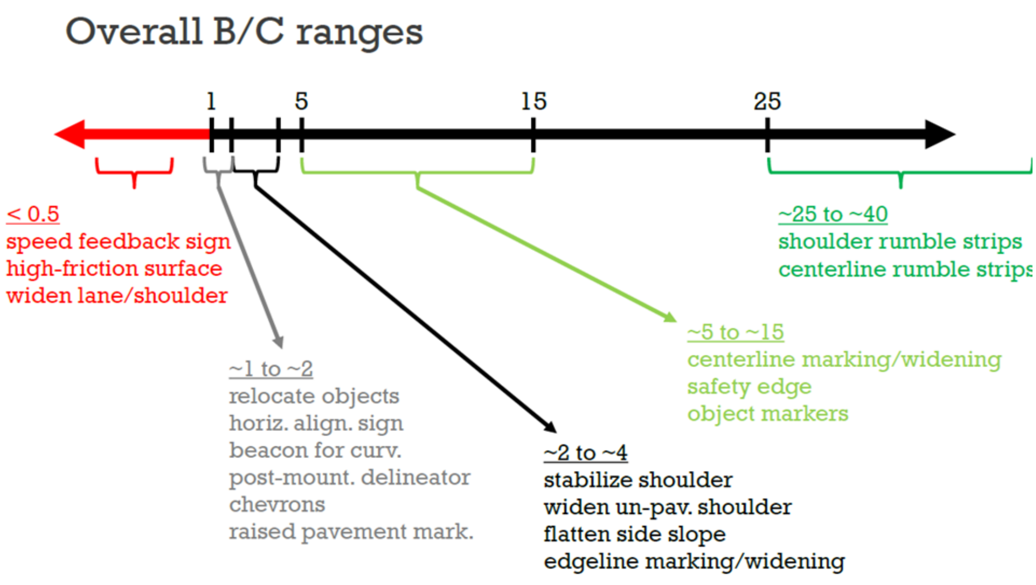


Figure 5.6. Benefit-Cost Diagram for distinct safety countermeasures used on rural roads (Crash Risk Factors for Low-Volume Roads: an ODOT Case Study, 2016. National Center for Rural Road Safety)

The analysis of the diagram shows that countermeasures related to pavement and shoulder resurfaces as well as to cross-sections and roadside features have higher B/C rations when compared to those from road alignment. Nonetheless, according to the study, alignment countermeasures are highly used for low-volume rural roads particularly for their low-cost implementation, as well as for their effectiveness concerning crash reduction. For instance, the study results show that chevrons and arrow signs reduced injury and fatal crashes by 18% and night-time crashes by 27.5%[2], while raised pavement marks reduced total crashes from 224 to 33, fatalities from 7 to 0, and injuries from 152 to 10[3]. Some of these safety countermeasures often used on low-volume rural roads are described below based on the FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014.

<u>Horizontal Alignment Signs</u>

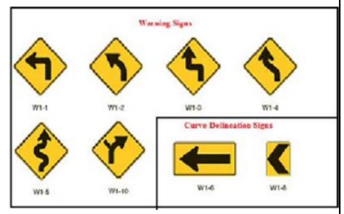
Horizontal alignment signs can be used to alert drivers about changes related to the road geometry, providing them with some information about the type of curve they are approaching. The table below shows CMFs equal to 0.7, which represents a significant Crash Reduction Factor (CRF) of 30%.

[2] Empirical Bayes before and after (average of 5.6 years before data and 5.4 years of after data) for 89 rural two-lane curves in Connecticut and 139 rural two-lane curves in Washington State.

[3] Simple before and after (4 years data before and 4 years data after) for 10 rural roadways (tangents and curves) in Mobile County, AL, with documented high run-off-road crashes.

Table 5-1. Treatment matrix for installing Horizontal Alignment Signs (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014).

Install Curve Warning Signs - Initial Investment: $2,400 - Cost of Maintenance: $1,280 - Frequency of Maintenance: 5 years	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume Optimal Conditions	33.8	Proven	0.70
Higher Volume Optimal Conditions	270.1	Proven	0.70
Lower Volume Narrower Conditions	43.5	Proven	0.70
Higher Volume Narrower Conditions	428.4	Proven	0.70



Use:

- Horizontal alignment signs can be applied to any curve or turn with a history of roadway departure crashes and to those with similar geometry or traffic volume that have not experienced crashes yet.
- "Warning signs are required on curves or turns where the advisory speed is 10 mph less than the posted speed". [2009 Manual on Uniform Traffic Control Devices (MUTCD) cited by Atkinson et. al., 2014].
- "Studies have shown that reductions in crashes due to the installation of curve warning signs are more prominent at locations with expressive traffic volumes, sharper curves, or hazardous roadsides". (Atkinson et. al., 2014)

Flashing Beacons

Flashing beacons are generally introduced to show the presence of an intersection, improving safety particularly at spots with night visibility issues, such as the case of Yakima Nation.

Provide Flashing Beacons at Intersection Approaches - Initial Investment: $25,000 - Cost of Maintenance: $1,000 - Frequency of Maintenance: 2 years	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume 4-Way Intersections	16.3	Proven	0.85
Higher Volume 4-Way Intersections	56.8	Proven	0.85
Lower Volume 3-Leg Intersections	6.8	Proven	0.85
Higher Volume 3-Leg Intersections	35.8	Proven	0.85

Figure 5.7. Treatment matrix for Flashing Beacons at intersections approaches (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

- Installed at intersections with no signaling characterized by "patterns of right-angle crashes related to lack of driver awareness of the intersection on an uncontrolled approach and lack of driver awareness of the Stop sign on a stop-controlled approach". (Atkinson et. al., 2014)

- They can be implemented either atop Stop signs or Advance Intersection Warning Signs.

<u>Chevrons</u>

Chevrons (also known as curve delineation signs) show the road's alignment when drivers are within the actual horizontal alignment of a curve. The signs show the shape and degree of curvature, acting as a guide for drivers through the entire curves/turns.

Install Chevron Signs - Initial Investment: $7,200 - Cost of Maintenance: $3,600 - Frequency of Maintenance: 5 years	Benefit-Cost Ratio	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Lower Volume Optimal Conditions	10.6	Proven	0.75
Higher Volume Optimal Conditions	84.7	Proven	0.75
Lower Volume Narrower Conditions	13.0	Proven	0.75
Higher Volume Narrower Conditions	127.7	Proven	0.75

Figure 5.8. Treatment matrix for Chevrons (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

- Installed at any curve/turn with a history of roadway departure crashes and at those with similar geometry or traffic volume that have not experienced crashes yet.
- "Alignment delineation (or a one direction large arrow) is required on curves or turns where the advisory speed is 15 mph less than the posted speed limit". [2009 Manual on Uniform Traffic Control Devices (MUTCD) cited by Atkinson et. al., 2014].

<u>Raised Pavement Markers</u>

Raised pavement markers increase the visual alignment provided by pavement markers, making them more salient for drivers, especially during adverse weather conditions.

Safety Treatment	Initial Implementation Cost	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Install Raised Pavement Markers	$0 to $20,000	Tried	≤ 0.76

Figure 5.9. Treatment matrix for Raised Pavement Markers (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

- On roads with adequate pavement quality to hold the devices in place.
- The type of the marker depends on regional climate (e.g., in areas subject to snowfall, snow plowable devices should be used)

5.3.2. Improving Pedestrian and Bicyclist's Safety

The safety countermeasures already mentioned can also contribute to the reduction of crashes involving pedestrians and bicyclists. For example, edge lines and chevrons can alert drivers when they are leaving from the driving lane, protecting pedestrians and cyclists who circulate along the shoulder in areas without sidewalks or exclusive bike lanes. Likewise, flashing beacons can improve pedestrian and bicyclist crossing at intersections, particularly during night periods. However, there are specific treatments whose main goal is to improve safety for these non-motorized users. Some of the pedestrian/bicyclist safety-oriented countermeasures recommended for Yakima Nation are presented below (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014).

Crosswalks

Providing crosswalks at target locations is an effective treatment to define spots for pedestrian crossings and to draw drivers' attention. Indeed, their NCHRP 500 Performance Rating is classified as Proven and Tried[4].

Safety Treatment	Initial Implementation Cost	NCHRP 500 Performance Rating
Provide Crosswalks at Targeted Locations	$0 to $5,000	Proven & Tried

Figure 5.10. Treatment matrix for implementing crosswalks at target locations (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

Crosswalks can be implemented at:

- Locations with stop signs or traffic signals to indicate crossing sites to pedestrians and to prevent vehicular traffic from blocking pedestrian paths.
- Non-signalized street crossing sites in specific school zones.
- Non-signalized locations where engineering judgment shows that the number of motor vehicle lanes, pedestrian exposure, average daily traffic (ADT), posted speed limit, and site's geometry would make the use of crosswalks desirable for safety.

However, marked crosswalks alone (i.e., without traffic/pedestrian signals or other expressive crossing improvements) are not sufficient and should be not used:

- "Where the speed limit exceeds 40 mph (64.4 km/h)
- On roads with four or more lanes without a raised median or crossing island that has (or will soon have) an ADT of 12,000 or greater.

[4] No CMF information provided

- On roads with four or more lanes with a raised median or crossing island that has (or soon will have) an ADT of 15,000 or greater". (Atkinson et. al., 2014)

Sidewalks

Sidewalks provide a refuge for pedestrians and enhance road operations. Their implementation can considerably improve safety for both drivers and pedestrians, especially at locations with heavy pedestrian volumes.

Safety Treatment	Initial Implementation Cost	NCHRP 500 Performance Rating
Build Sidewalks	$5,001 to $50,000	Proven

Figure 5.11. Treatment matrix for building sidewalks (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

Building sidewalks should be considered for heavy pedestrian volumes existing along a corridor or specific location.

Pedestrian Hybrid Beacons or High Intensity Activated Crosswalk (HAWK)

The High Intensity Activated Crosswalk (HAWK) is "a pedestrian-activated beacon located on the roadside and on mast arms over major approaches to an intersection" (Atkinson et. al., 2014). It consists of two red lenses over a single yellow lens, displaying a red indication to drivers when activated. The device is illuminated only by pedestrian activation, changing to yellow and then to red to make drivers stop. It also shows a walking person symbol to pedestrians at the beginning and an upraised hand symbol with a countdown display at the conclusion of the walk phase. The estimated CMF is 0.712, indicating a significant safety improvement for pedestrians and drivers.

Safety Treatment	Initial Implementation Cost	NCHRP 500 Performance Rating	Crash Modification Factor (CMF)
Install Pedestrian Hybrid Beacons or High Intensity Activated Crosswalk (HAWK)	$20,001 to $100,000	Proven	0.712

Figure 5.12. Treatment matrix for Hybrid Beacons or High Intensity Activated Crosswalk (HAWK), (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

This countermeasure may be used at locations with a significant number of pedestrian crashes where additional visibility is needed. In Yakima Nation, for example, installing HAWKs could be appropriate at locations with important historical numbers of crashes involving pedestrians and bicyclists, such as the W 1st Ave and S Elm St/Buena Way in Toppenish and the W 1st St and Donald Wapato Rd in Wapato.

Shared-Use Paved Shoulders for Horse & Buggy Road Users or Bicyclists

Shared-use paved shoulders provide a paved shoulder next to the roadway with sufficient width to allow movements for other modes of transportation such as horses, buggies, and bicycles. The implementation of this treatment helps separate slower moving traffic from the main traffic lane, reducing crashes' incidences[5].

Safety Treatment	Initial Implementation Cost	NCHRP 500 Performance Rating
Construct Shared-Use Paved Shoulders for Horse & Buggy Road Users or Bicyclists	$5,001 to $50,000	Tried

Figure 5.13. Treatment matrix for Shared-Use Paved Shoulders for Horse & Buggy Road Users or Bicyclists, (FHWA Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014)

Use:

This treatment may be used at locations with frequent slower moving traffic, such as bicycle routes and sites with horse and buggy users.

5.4. Summary of Safety Countermeasures for Yakima Nation

As a RITI community, Yakima Nation faces several road safety problems related to the ongoing road conditions. Generally, there is a lack of pedestrian facilities, and most roads do not have a shoulder, but an embankment or a drainage ditch instead. This forces pedestrians to walk essentially on the fog line or in the live traffic lane along most of these roads. Additionally, most of the population do not have access to private vehicles, relying on either public transportation (which has limited routes and schedule) or walking. Several intersections have only stop signs with poor visibility, which can be more hazardous during winter months when fog regularly limits drivers' visibility.

Therefore, the countermeasures previously described can be effective treatments for improving safety in the Yakama Nation road network. However, the choice between them will depend on the characteristics of each road, which will define the most appropriate treatments to be implemented.

Figure 5.14 shows some road profile examples taken from distinct roads of Yakima Nation.

[5] No CMF information provided


a. no shoulder


b. no signs/markings


c. lack of signs


d. faded pavement markings

Figure 5.14. Set of road photos taken from Yakima Nation. Source: Yakama Nation Tribal Traffic Safety

Most of the profiles are formed by low-volume roads, with exceptions for some routes within more urban centers like Toppenish. From the above photos, we also note a lack of pedestrian and bicycle facilities, which may be related to the significant number of crashes on certain roads. The following image shows the spots for crashes involving pedestrians from 2010 to current times, where the majority are located in the area of Toppenish along W 1st Ave and S Elm St/Buena Way.

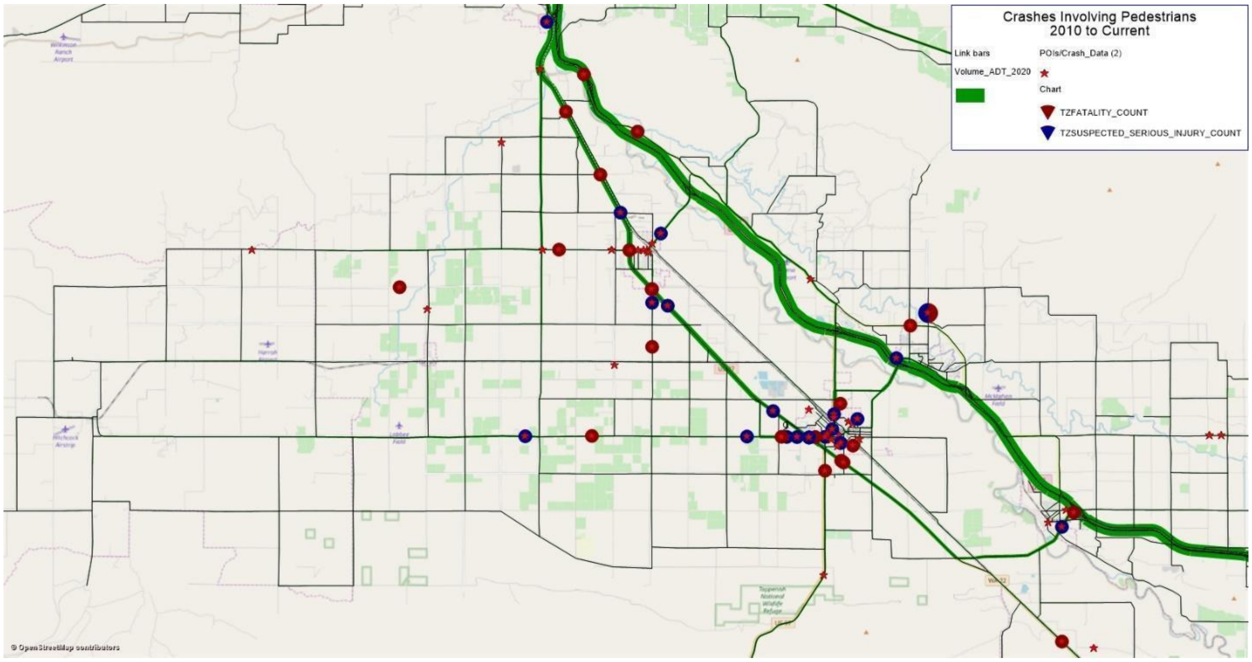


Figure 5.15. Locations for crashes involving pedestrians in Yakima Nation (2010-current). Source: Yakama Nation Tribal Traffic Safety

The following table shows the mechanism of motor vehicle-related deaths of Yakima County residents by race between 1999 and 2016. We observe a higher proportion of AI/AN (American Indian/Alaska Native) motor vehicle-related deaths for pedestrians when compared to other races (roughly 23% for AI/AN against 11% for Non-Hispanic Whites - NHW - and 10% for other races). This illustrates the alarming pedestrian crashes rates among the AI/AN community.

Table 5-2. Mechanism of motor vehicle-related deaths of Yakima County residents by race, 1999-2016. Source: Yakama Nation Tribal Traffic Safety.

	AI/AN	**Percentage**	**NHW**	**Percentage**	**Other races**	**Percentage**	**Total**
Occupant	90	**62.1**	169	**66.5**	217	**74.6**	476
Motorcyclist	0	**0.0**	27	**10.6**	9	**3.1**	36
Pedal cyclist	1	**0.7**	3	**1.2**	4	**1.4**	8
Pedestrian	33	**22.8**	29	**11.4**	29	**10.0**	91
Unspecified	21	**14.5**	26	**10.2**	32	**11.0**	79
Total	145	**100.0**	254	**100.0**	291	**100.0**	690

The table below shows the Yakima County resident pedestrian deaths by county of death occurrence. 83 out of the total of 91 Yakima County residents who were killed by motor vehicles as pedestrians were in Yakima County.

Table 5-3. Yakima County resident pedestrian deaths by county of death occurrence. Source: Yakama Nation Tribal Traffic Safety

County	**AI/AN**	**NHW**	**Other**	**Total**
Non-Washington State or Unknown	2	1	0	3
King County	1	0	0	1
Klickitat County	3	0	0	3
Walla Walla County	0	0	1	1
Yakima County	27	28	28	83
Total	33	29	29	91

Additionally, 50% of the AI/AN pedestrian deaths occurred at night (9 PM – 5 AM), against 24% for NHW and 21% for other races, as seen in the following table. In fact, different from the AI/AN community, the highest percentage of non-AI/AN pedestrian deaths occurred from 4 PM to 8 PM. These numbers highlight the predominance of accidents involving pedestrians in RITI communities during night periods.

Table 5-4. Pedestrian deaths by hour of death that occurred in Yakima County 1999-2016. Source: Yakama Nation Tribal Traffic Safety

	AI/AN		**NHW**		**Other/Unknown**		
Hour of injury	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**	**Total**
Day time (6AM-3PM)	1	**3.57**	9	**31.03**	3	**10.71**	**13**
Night time (9PM-5AM)	14	**50.00**	7	**24.14**	6	**21.43**	**27**
Rush hour (4-8PM)	9	**32.14**	10	**34.48**	10	**35.71**	**29**
Missing	4	**14.29**	3	**10.34**	9	**32.14**	**16**
Total	28	**100.00**	29	**100.00**	28	**100.00**	**85**

Countermeasures for both low-volume and more urban roads can effectively be implemented to improve road safety within the area. Some locations may be prioritized, such as spots with historical numbers of crashes as well as those with actual or expected important demands of pedestrians or cyclists. Indeed, Yakima Nation is planning to implement a trail for pedestrians and cyclists connecting its major city (Toppenish) with other distant locations, so that installing safety countermeasures on roads that are close to this future trail will be an efficient effort to protect users and drivers. Figure 5.16 depicts the location of this trail plan with some images of nearby roads[6].

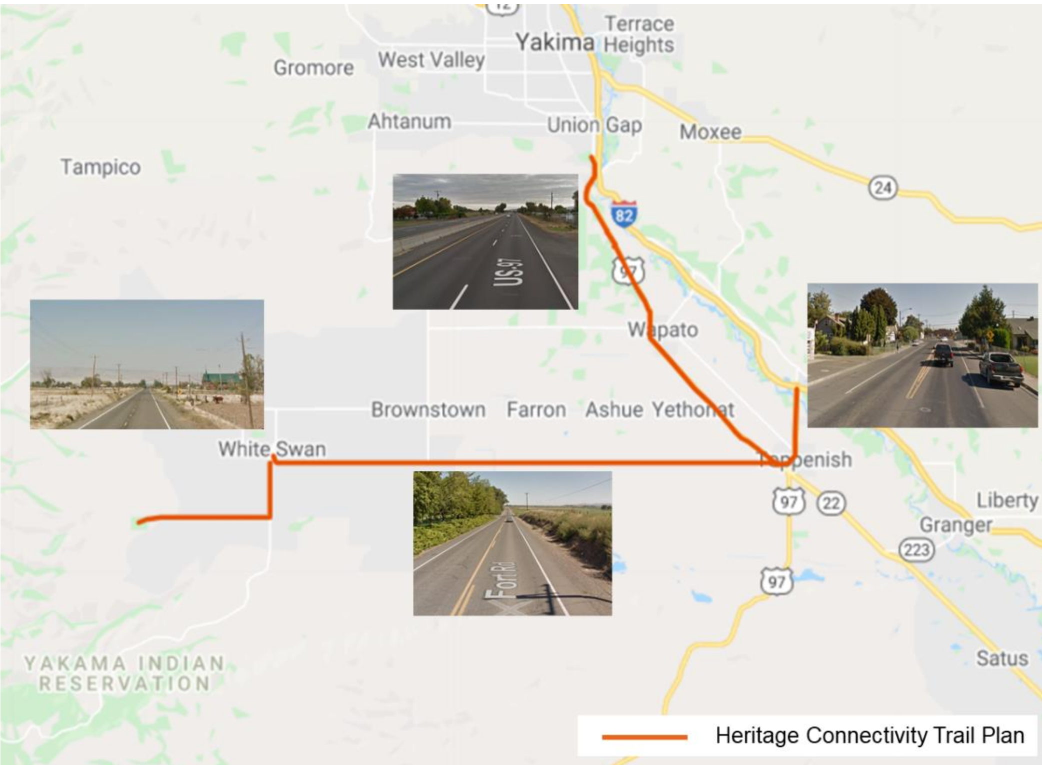


Figure 5.16. Trail plan for Yakima Nation. Source: Own elaboration based on information provided by Yakama Nation Tribal Traffic Safety

The table below summarizes our safety countermeasures recommendations for Yakima Nation according to their roads' profile.

Table 5-5. Summary of safety countermeasures recommendations for Yakima Nation

Safety countermeasure	**Where to implement**
Crosswalks	Target locations with a high volume of pedestrian circulation (schools, hospitals, public spaces) and spots with historical pedestrian crashes.
Sidewalks	Locations with urban characteristics and those with significant pedestrian circulation but without pedestrian facilities.
High Intensity Activated Crosswalks (HAWK)	Target locations with a significant number of pedestrian crashes where additional visibility is needed.

[6] Two major highways cross Yakima Nation territories: Interstate 82 (I-82) and State Route 22 (SR 22). Although some of the described countermeasures for high volume roads can be used in specific cases, these roads are not addressed in the Manual for Selecting Safety Improvements on High Risk Rural Roads, 2014

Shared-Use Paved Shoulders	Locations with frequent slower moving traffic
Horizontal Alignment Signs	Hazardous curves or turns.
Flashing Beacons	Intersections with no signaling, either atop Stop signs or Advance Intersection Warning Signs, especially on spots with night visibility issues.
Chevrons	Any curve/turn with a history of roadway departure crashes or with similar geometry or traffic volume yet to experience crashes.
Raised Pavement Markers	Any route with sufficient pavement quality to hold the devices in place.
Edge Lines & Center Line Rumble Strips	Roads with a history of road departure and head-on crashes.
Safety Edge	Locations where pavement edge drop-offs occur through everyday use, particularly on rural roads with unpaved shoulders. Policies and procedures subject to each State.

CHAPTER 6. REFERENCES

Abdel-Aty, M. 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of safety research*, *34*(5), pp.597-603.

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*. https://www.tensorflow.org/

Atkinson, J.E., Chandler, B.E., Betkey, V., Weiss, K., Dixon, K, Giragosian, A., Donoughe, K. and O'Donnell, C., 2014. Manual for Selecting Safety Improvements on High Risk Rural Roads. U.S. Department of Transportation – Federal Highway Administration. Report No. FHWA-SA-14-075.

Awwad-Rafferty, R., Chang, K., and Brown, H., 2019. Reaching Out to Tribal Communities: Lessons Learned and Approaches to Consider.

Baireddy, R., Zhou, H. and Jalayer, M., 2018. Multiple correspondence analysis of pedestrian crashes in rural Illinois. Transportation research record, 2672(38), pp.116-127.

Bhuiyan, N. 2019. Enhancing Pedestrian Safety in Bangladesh. *CONE 2019 Discipline Specific Proposal (Transportation Engineering)*. https://www.researchgate.net/publication/332057524

Cambridge Systematics Inc., and Texas Transportation Institute, 2005. "Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation." FHWA Office of Operations, US Department of Transportation." *Washington, DC: Federal Highway Administration*.

Chen, C., Zhang, G., Huang, H., Wang, J. and Tarefder, R.A. 2016. Examining driver injury severity outcomes in rural non-interstate roadway crashes using a hierarchical ordered logit model. *Accident Analysis & Prevention*, *96*, pp.79-87.

Chen, Z., and Fan, W., 2019. Modeling pedestrian injury severity in pedestrian-vehicle crashes in rural and urban areas: mixed logit model approach. Transportation research record, 2673(4), pp.1023-1034.

Crash Modification Factors Clearinghouse <http://www.cmfclearinghouse.org/> (Accessed 03.23.2021)

Ding, C., Chen, P., & Jiao, J. 2018. Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: A machine learning approach. *Accident Analysis and Prevention, Vol. 112*, 116-126. https://doi.org/10.1016/j.aap.2017.12.026

Das, S., Le, M, & Dai, B. 2020. Application of machine learning tools in classifying pedestrian crash types: A case study. *Transportation Safety and Environment*, *Vol. 2*, *No. 2*, 106–119. https://doi.org/10.1093/tse/tdaa010

Federal Highway Administration, 2012. *Traffic Safety Facts "Rural and Urban Comparison".* U.S. Department of Transportation.

Hall, J. 1983. Pedestrian Accidents on Rural Highways. *Transportation Research Record*, *904*, 46–50

Harkey, D.L., Srinivasan, R., Baek, J., Council, F., Eccles, K., Lefler, N., Gross, F., Persaud, B., Lyon, C., Hauer, E., and Bonneson, J., 2008. National Cooperative Highway Research Report 617: Accident Modification Factors for Traffic Engineering and ITS Improvements, NCHRP, Transportation Research Board, Washington, DC, 2008.

Karlaftis, M.G. and Golias, I. 2002. Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis & Prevention*, *34*(3), pp.357-365.

Lin, C., Wu, D., Liu, H., Xia, X. and Bhattarai, N. 2020. Factor identification and prediction for teen driver crash severity using machine learning: a case study. *Applied Sciences*, *10*(5), p.1675.

Marshall, W.E. and Ferenchak, N.N., 2017. Assessing equity and urban/rural road safety disparities in the US. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 10(4), pp.422-441.

National Highway Traffic Safety Administration. 2018. "Traffic Safety Facts: 2016 Data-Rural/Urban Comparison of Traffic Fatalities." *United States Department of Transportation*.

National Academy of Sciences. 2008. *Pedestrian Safety Prediction Methodology*. National Academy of Sciences, Engineering and Medicine.

National Center for Rural Road Safety < https://ruralsafetycenter.org/> (Accessed 03.23.2021)

National Center for Rural Road Safety, 2016. Crash Risk Factors for Low-Volume Roads: an ODOT Case Study. Presentation. < https://ruralsafetycenter.org/wp-content/uploads/2016/04/Low-Volume-Risk-Factors_Handout.pdf> (Accessed 03.23.2021)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, pp.2825-2830. https://scikit-learn.org/stable/

Plainis, S., Murray, I.J. and Pallikaris, I.G. 2006. Road traffic casualties: understanding the night-time death toll. *Injury Prevention*, *12*(2), pp.125-138.

Prato, C.G., Rasmussen, T.K. and Kaplan, S. 2014. Risk factors associated with crash severity on low-volume rural roads in Denmark. *Journal of Transportation Safety & Security*, *6*(1), pp.1-20.

Snyder, M., & Knoblauch, R. 1971. *Pedestrian Safety The Identification of Precipitating Factors and Possible Countermeasures*. US Department of Transportation.

Souleyrette, R.R., Caputeu, M., McDonald, T.J., Sperry, R.B., Hans, Z.N. and Cook, D. 2010. *Safety Analysis of Low-Volume Rural Roads in Iowa* (No. InTrans Project 07-309). Ames, IA: Institute for Transportation, Iowa State University.

Souleyrette, R.R., Caputeu, M., McDonald, T.J., Sperry, R.B., Hans, Z.N. and Cook, D. 2010. *Safety Analysis of Low-Volume Rural Roads in Iowa* (No. InTrans Project 07-309). Ames, IA: Institute for Transportation, Iowa State University.

Vogt, A. and Bared, J. 1998. Accident models for two-lane rural segments and intersections. *Transportation Research Record*, *1635*(1), pp.18-29.

Wang, Y., Zeng, Z., Gottsacker, C. and Yang, H.F., 2019. *Develop a Regional Multi-Source Database System for Safety Data Management and Analysis in RITI Communities in Washington State*.

Wang, Y., Sun, W., Yang, H., Gottsacker, C., Ricord, S. and Yin, S., 2019. *Developing a Data-Driven Safety Assessment Framework for RITI Communities in Washington State*.

Ye, F. and Lord, D. 2014. Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic methods in accident research*, *1*, pp.72-85.

Zajac, S., & Ivan, J. 2002. Factors Influencing Injury Severity of Motor-Vehicle Crossing Pedestrian Crashes in Rural Connecticut. *Accident Analysis and Prevention*, *35*, 369–379. https://doi.org/10.1016/S0001-4575(02)00013-1