A Pilot Data System and Analytical Framework for Tribal and Rural Community Traffic Safety Equity Assessments

Christopher Gottsacker

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Civil Engineering

University of Washington

2019

Committee:

Yinhai Wang, Chair

Xuegang Ban

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

A Pilot Data System and Analytical Framework for Tribal and Rural Community Traffic Safety Equity Assessments

Christopher Gottsacker

Chair of the Supervisory Committee:
Professor Yinhai Wang
Department of Civil and Environmental Engineering

There is a systemic inequity in terms of traffic safety between rural and urban areas throughout the nation. Rural, isolated, tribal, and indigenous (RITI) communities often need more funding to address transportation safety concerns. However, recent focus on technological improvements in urban areas risks widening this divide, leaving RITI communities further behind. This study aims at investigating solutions to address the problem. In particular, a framework is proposed to address the issue of traffic safety equity. Research efforts are made to follow this framework from outreach to data analysis and visualization. The outreach activities have resulted in data sharing agreements with one tribe in Washington State and paved the road for signing to agreement with another tribe. Descriptive analyses are conducted to gain basic understanding of the data, and crash frequency prediction models are also used to find significant contributing

factors in crash occurrence. The results of this model are used to create an index of Crash Reduction Potential, and an online crash mapping application is developed to visualize the index.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis could not have come to fruition without the tremendous support of so many people. First, I would like to thank my research adviser, Dr. Yinhai Wang from the University of Washington Department of Civil and Environmental Engineering. Dr. Wang consistently pushed me to challenge myself and to persevere. He always allowed me to make this research project my own, even encouraged it, while still being available when I needed guidance or direction.

I would also like to thank the previous work of other researchers, practitioners, and experts who helped me understand this field and gave me insight to frame this work, while igniting a passion for traffic safety. The foundation of previous work is solid and it is an honor to contribute. I must thank some specific people who were quite instrumental in the success of this project. Dr. Ziqiang Zeng and Dr. Ying Jiang provided support, advice, and guidance while they completed the post-doctoral studies at the University of Washington. Thank you to John Ash, Kris Henrickson, and Mayuree Binjolkar for being excellent course instructors, mentors, and a sounding board of great advice throughout my program. This project allowed me to meet several people in the American Indian and Alaska Native community, and while I cannot name them all, these are some of the people whose input and participation contributed substantially to the success of this project: Kwis Logan, Nicole Ahlem, and Angelena Campobasso of the Confederated Tribes of the Colville Reservation; Dezerae Hayes of the Muckleshoot Indian Tribe; and Iisaaksiichaa Ross Braine from the University of Washington Intellectual House and the Apsáalooke Nation.

I would also like to thank Dr. Xuegang Ban of the University of Washington Department of Civil and Environmental Engineering for being the second reader of my thesis and defense committee member. Dr. Ban provided great comments and questions that helped me develop my thesis, and I will always be grateful for this.

Finally, I must thank my parents, siblings, and girlfriend for their unwavering support and encouragement – without their continuous belief in me even when I did not believe in myself, I would not have had the courage to chase my dreams. Throughout my years of education, my parents have always been my rock, the place I can go when I need to vent, unwind, focus, or just get advice. Many thanks for all they have done for me, as this would not have been possible without them. Thank you, thank you, thank you.

Christopher Gottsacker

Chapter 1. INTRODUCTION

1.1 RESEARCH BACKGROUND

Rural, isolated, tribal, and indigenous (RITI) communities are often overlooked for transportation improvements, which causes a certain disadvantage especially in terms of traffic safety. This is a major equity and public health concern, especially as technology continues to develop at impressive rates while implementation is primarily focused in urban or metropolitan areas, creating inequity in terms of transportation funding allocation. The technological nature of many urban transportation improvements creates an even greater divide between improvements in rural areas, where often relatively simple infrastructure improvements and proven countermeasures can have a significant impact. However, without the deserved attention, these RITI communities face their transportation issues, safety related and otherwise, with limited funding and other resource impediments such as personnel and expertise shortages. The Center for Safety Equity in Transportation states that “if you have a right to get there, you have a right to get there safely,” and the systemic disadvantage RITI communities face is directly at odds with this mentality as well as Vision Zero and Target Zero initiatives (Center for Safety Equity in Transportation, 2019). The cultural and environmental diversity found in RITI communities are two reasons why these areas deserve greater attention to traffic safety analysis.

The Vision Zero program adopted in the United States and Washington State aims to eliminate traffic fatalities by 2030, yet there is a discrepancy between traffic safety improvements in RITI and urban communities. In 2000, Washington State became the first state in the United States to adopt formal policy aimed at reducing roadway fatalities to zero and was influenced by Sweden’s Vision Zero program, started in 1997. The program has been a significant point in the conversation regarding traffic safety, with 42 cities adopting their own plans. This clearly suggests that the nation, as a whole, adequately recognizes the public health and economic challenge posed by traffic collisions. However, while there has undoubtedly been significant research and practical improvements in the realm of traffic safety to support the goals of Vision Zero, it is also clear that there is lacking attention to RITI community needs.

Figure 1 depicts the trend of fatalities from 2005 - 2014 throughout Washington State (Washington Traffic Safety Commission, 2016). It is important to note that the overall trend is promising, but recent years have seen an increase in fatalities. Because of this increase, there exists a performance gap between reaching the goal of Target Zero and the current trend that is seen. While this is from the most recent Strategic Highway Safety Plan, more recent state fatality data suggests that the performance gap has continued to increase since 2014.

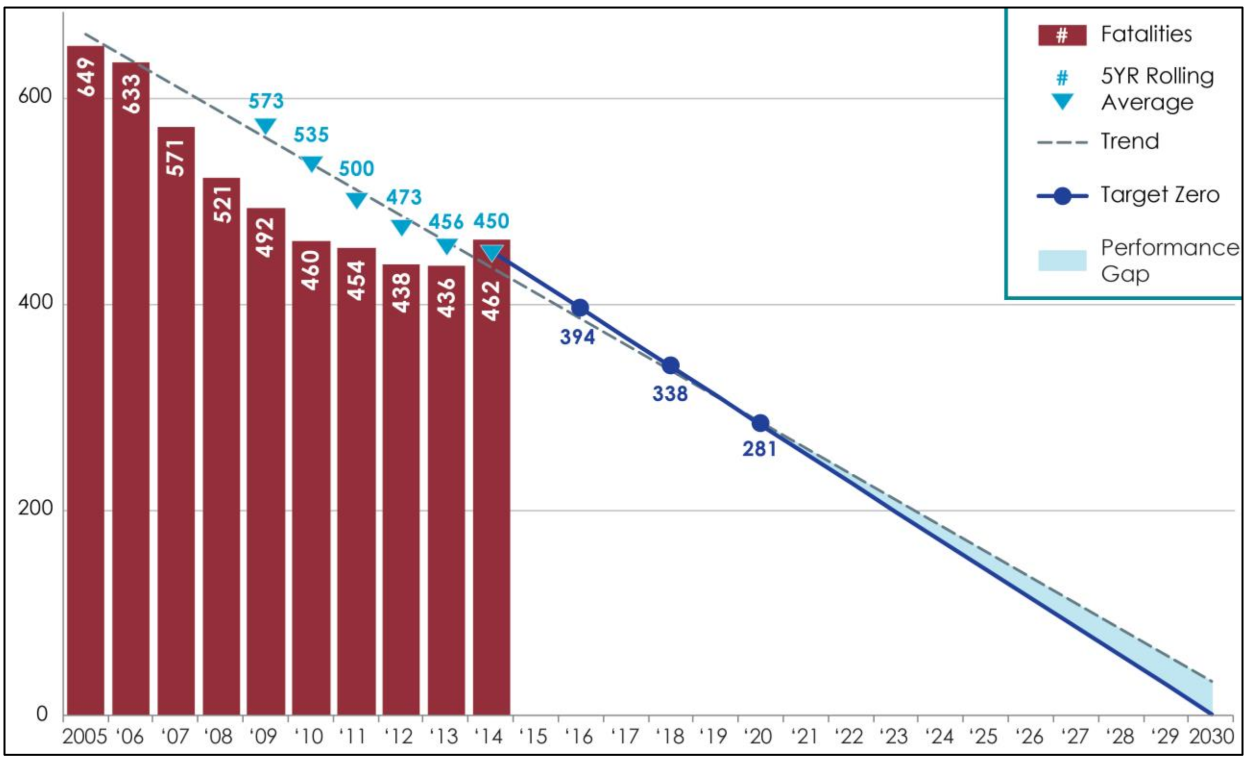

Figure 1 Trend of Fatalities in Washington State

While the Strategic Highway Safety Plan does include notes on the need to focus on rural areas and includes statistics for American Indian and Alaska Native (AIAN) communities, only cursory statistics are provided to indicate this performance and there is limited discussion of solutions to mitigate crashes in rural areas specifically. Figure 2 depicts the difference in rural and urban crashes throughout the entire nation from 2007 – 2016. This figure shows a promising, primarily downward trend in fatality counts for both urban and rural crashes, with a faster decrease in rural fatalities. However, Figure 3 depicts the fatality rate per 100 million vehicle miles traveled (VMT) for urban and rural areas throughout the nation for the same time period (National Highway Traffic Safety Administration, 2018). Fatality rate per VMT is a more useful and intuitive measure and trend investigation as it provides an extra layer of context. While fatality counts may look very similar, the fact remains that in rural areas, roadway crashes result in a fatality more than twice as frequently as urban area crashes. The overall trend is still promising, but the stark difference of rural fatalities occurring at a rate of about 2.5 times that of urban fatalities has persisted from 2007 through 2016. The evidence reinforces the disproportionate need for rural areas to receive attention and support to better address traffic safety.

Figure 2 Fatality Counts in Rural and Urban Areas

Figure 3 Fatality Rates in Rural and Urban Areas

1.2 Research Objective

This research, funded by the Center for Safety Equity in Transportation, serves to address some of the disparities between rural and urban transportation safety. The research does so by applying a framework of outreach, data collection, data management, and traffic safety analysis that is replicable. The outreach activities prioritize forming relationships with the underserved RITI populations and to better understand local challenges. The outreach was also meant to expand the data available to be used in the study, especially to include tribal traffic data to better quantify challenges on reservation roads as state data on AIAN crashes is known to be incomplete. Merging different sources of publicly available data with the tribal data is a crucial component of the framework. Even merging the state data can present a challenge, with crash data not often linked to the roadway characteristic data. Additionally, this study used different definitions of rurality which are needed to link to the crash data. Once merged, the goal was to analyze the crash frequency on rural state highways, which is one of the proposed applications of the framework. In summary, the research objective is a proof of concept of the proposed RITI traffic safety analysis framework from outreach to analysis for safety performance. A part of this included investigating the data and results from different definitions of rural.

1.3 PROBLEM STATEMENT

The disadvantage facing RITI communities in terms of traffic safety is even more clear when looking at crash data in recent years. Again, to be more specific, the data especially highlight the challenge facing AIAN communities – according to the national Fatality Analysis Reporting System (FARS), the fatal crash rate among the AIAN community is higher than any other race in Washington State (National Center for Statistics and Analysis, 2015). This holds true when accounting for deaths both on and off tribal lands. Figure 4 shows the fatality rate per 100,000 population for AIAN and non-AIAN races in Washington State using data from 2002 – 2011. Note that the fatality rate for AIANs is 4 times that of non-AIANs. A similar case is found when comparing pedestrian fatality rates per 100,000 population among AIANs and non-AIANs. Figure 5 shows this comparison - note that the AIAN pedestrian fatality rate is nearly 5 times that of the overall non-AIAN pedestrian fatality rate. While this data is from 2002 – 2011, the rates are still very similar; the Washington Traffic Safety Committee reports that as of 2016, the AIAN pedestrian crash rate was 5 times that of the non-AIAN pedestrian crash rate. They also reported that as of 2016, the overall fatality rate for AIANs was 4.2 times that of the non-AIAN fatality rate (Washington Traffic Safety Commission, June 2018). While the national traffic fatality rate is lower than it was in 2002, the AIAN has seen virtually no change in the overall fatality rate; greater attention must be paid to RITI traffic safety issues.

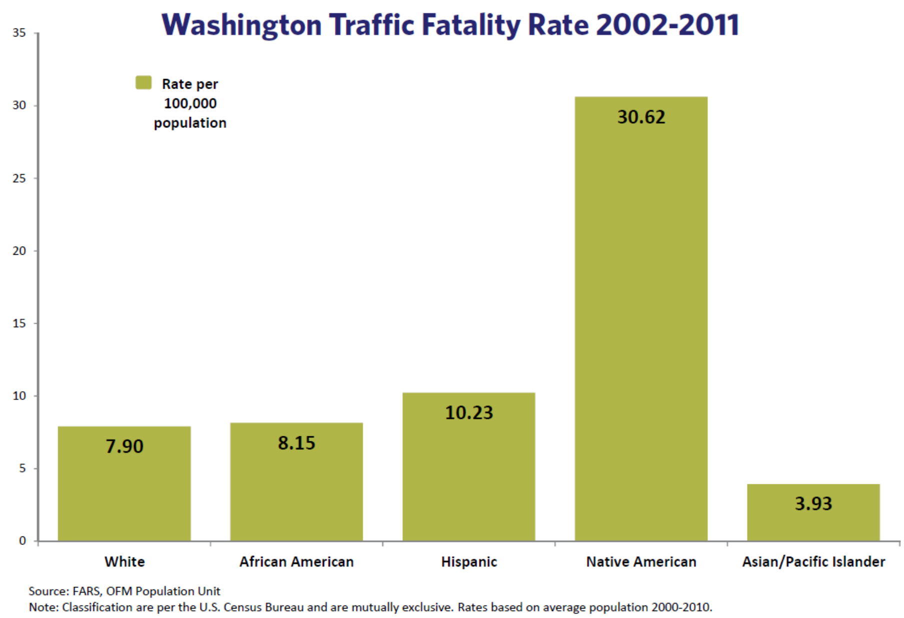

Figure 4 Comparison of Fatality Rate for AIAN and non-AIAN in Washington State

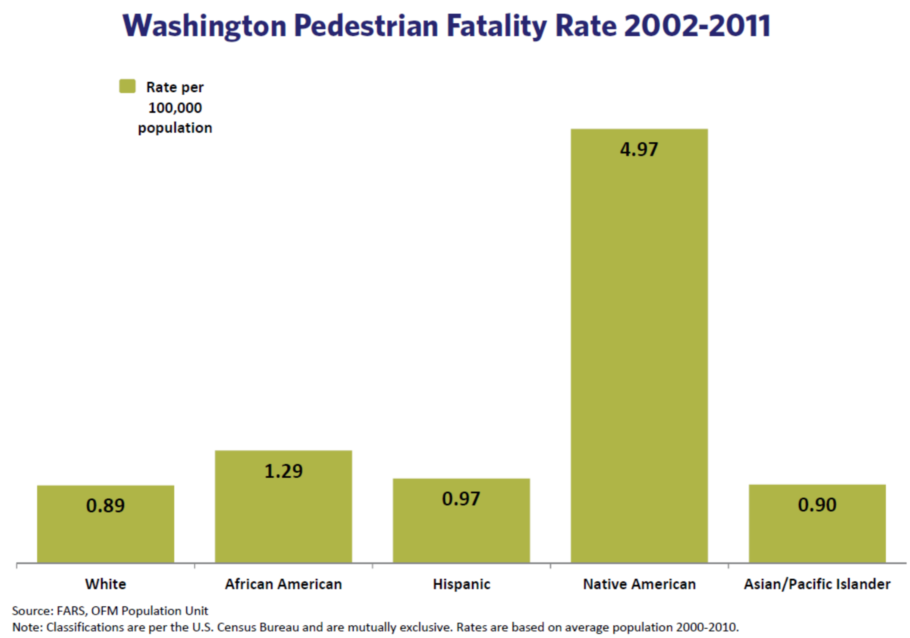

Figure 5 Comparison of Pedestrian Fatality rate for AIAN and non-AIAN in Washington State

A particular concern of AIAN tribes in Washington are behavioral-related fatalities, such as crashes where occupants are unrestrained, traveling at high speeds or above the speed limit, or under the influence of drugs and/or alcohol. Pointedly, the fatality rate per 100,000 population of unrestrained occupants in AIAN crashes is 9 times that of all non-AIAN crashes combined for years 2007-2016 in Washington (Washington Traffic Safety Commission, June 2018). It is also recognized that RITI communities may not have the resources to collect quality traffic safety data. AIAN communities are separate sovereign entities that are not required to collect or share all data with United States citizens or state departments, though the Centennial Accord in Washington State aimed to help facilitate more transparent communication. The baseline level of data available and the quality of these data varies by tribe and rural community and accessing these data to form a more uniform database is a multi-faceted problem that requires consistent and fulfilling outreach activities. The current lack of uniform data collection and quality control approaches impedes efforts to conduct thorough analysis and hinders ability to secure federal and state funding for transportation improvements. Any solution must be culturally sensitive and collaborative, allowing more efficient knowledge transfer and continuous management of a database system. The outreach activities conducted with AIAN communities and progress made can be modeled to repeat with other tribes as well as other rural communities throughout Washington State.

1.4 Scope of Study

This study is meant to serve as a pilot for a new rural and tribal traffic safety data management, analysis, and visualization framework. Data was obtained for all of Washington State and for one tribe – the Confederated Tribes of the Colville Nation – which will be used in the pilot program. As equity is a major concern, the pilot focuses on rural and tribal traffic safety, though does include some characteristics for statewide crashes. The framework, outlined later, begins with outreach activities and ends with analysis and visualization results. The primary contributions of this study include the comparison of rural data definitions and analysis results, the tribal outreach activities culminating in quantitative tribal traffic safety analysis, and the presentation of the framework for future implementations and broader analysis. The analysis for this pilot focuses on overall descriptive analyses and crash frequency modeling on state routes, as well as initial efforts to combine and compare tribal and state crash data where available. The outreach efforts served to understand local traditions, form relationships to build future work, and to understand how traffic collision data collection and management practices are utilized. The data acquisition and combination served to present a method which merges crash data for each year with other available data. The crash frequency regression analysis served to determine the significant contributing factors in rural traffic crashes in Washington State. The proposed visualization tools serve to help communicate the results in a digestible manner by mapping the crashes and obtaining a measure of Crash Reduction Potential (CRP) for each segment in the dataset.

Chapter 2. LITERATURE REVIEW

2.1 TRAFFIC CRASH ANALYSIS METHODS

While RITI communities are indeed disadvantaged from a transportation safety perspective, there has been growing amounts of research addressing some of the primary concerns. There are some unique challenges to overcome when performing rural crash data analyses, such as geospatial randomness, missing data or poor data quality, and in general lower data volume. There are also challenges in working with RITI communities in terms of effective communication, finding appropriate strategies to address concerns, and understanding local needs. Context and culturally sensitive solutions are the only acceptable solutions for any type of work but are especially important when working with RITI communities.

Traditional traffic safety analysis is conducted using generalized regression models. The primary goal is often a crash frequency prediction or an injury severity prediction relying on police report data. A Poisson model has historically been employed as the fundamental generalized regression model for crash prediction. However, there are several known downfalls to this modeling technique, such as being unable to handle overdispersion and underdispersion, as well as not accounting for spatial or temporal characteristics. Additionally, the reliance on police reports alone presents a challenge despite the availability and difficulty in using other data sources – for instance, it has been found that police report crash datasets will often underreport the occurrence of non-injury or property damage only crashes (Yamamoto, Hashiji, & Shankar, 2008). There have also been noted inaccuracies in the true injury severity of a traffic crash, such as incidents of no reported injuries eventually requiring a hospital visit, or apparent serious injuries being discharged from the hospital early (McDonald, Davie, & Langley, 2009). Given these and other challenges faced by traditional traffic safety analysis methods, many other models have been developed or applied for crash frequency prediction, crash severity prediction, and crash clearance time prediction. There has been some advancement in crash data sources used for analysis, but these are largely still at the cutting-edge and not always available, such as the linkage of police reports to hospital reports. Despite the challenges of relying on police report data, it is likely that data used in later RITI-focused projects will come from police reports given the access to at least some form of police report data in most RITI communities. This said, there have been a multitude of studies relying solely on police report data that have helped the industry better understand traffic crashes and their underlying causes, and with the deployment of increasingly accurate methods it is possible to better account for or even estimate errors from police reports.

As stated, crash frequency modeling has historically been conducted using basic Poisson models in practice; this is often not suitable for more advanced analysis given that Poisson models rely on the basis that a sample mean is equal to its variance and thus cannot handle overdispersion, underdispersion, or a large amount of zero count data which is typical in crash data. The negative binomial model emerged as an option to handle overdispersed data, or data whose variance is greater than its mean. And while this modeling method is consistently used in practice, it has its own weaknesses such as the inability to handle underdispersion, which occurs when the crash count mean is greater than variance, and inaccuracies occurring when a low sample volume and mean exist (Lord & Mannering, 2010). These are characteristics that could be reasonable to expect when dealing with crash data, emphasizing the need for more advanced modeling techniques to be used in later projects related to RITI communities and data analysis. Overdispersion is handled by the negative binomial model and other common models, and is still relatively easy to interpret and execute, leading to its prevalence. One way to handle underdispersion of crash data has been the development of Conway-Maxwell-Poisson models (Lord, Geedipally, & Guikema, 2010). Other advanced modeling methods that have been used for traffic crash analysis in order to address various challenges of prior models or otherwise improve accuracy include random-effects models (Shankar, Albin, Milton, & Mannering, 1998), spatial and temporal correlation models (Aguero-Valverde & Jovanis, 2006; Wang & Abdel-Aty, 2006), random parameters count models (Wu, Sharma, Mannering, & Wang, 2013; Castro, Paleti, & Bhat, 2012), and several different neural networks (Zeng, et al., 2017; Abdelwahab & Abdel-Aty, 2001). In this pilot, relatively lightweight modeling approaches are used. This is to maintain ease of interpretability and to ensure that online access and processing can be handled by local computers that may not always have advanced capabilities. This is very intentional given the equity component of this research, which relies on the framework to be readily available and operable with minimal training. More detail regarding the methods used in this pilot are discussed in Chapter 3.

2.2 Rural and Tribal Related Crashes

When comparing rural crashes to urban crashes, it is clear that the fatality rate on rural roads is higher than on urban roads even when controlling for crash severity (Muelleman, Wadman, Tran, Ullrich, & Anderson, 2007). This suggests that the distance to medical attention could be a factor diminishing the survivability of rural crashes. The scarcity and distance to medical resources is related to the spatial characteristics of rural areas. Given the many miles of rural roads with relatively low vehicle miles traveled (VMT), it is often necessary to consider spatial correlation when investigating rural crashes and their severity (Aguero-Valverde & Jovanis, 2008). Besides spatial characteristics, rural fatal crash rates are more heavily influenced by behavioral instances such as alcohol impairment, speeding, and overtake maneuvers (Kloeden, et al., 2001; Wu, et al., 2014; Wu, et al., 2016). The National Highway Traffic Safety Administration (NHTSA) found seatbelt use was low in AIAN communities and funded programs to help address this, though it continues to be a high contributing factor in serious injury and fatal crashes (Leaf & Solomon, 2005). These behavioral instances are more difficult to directly address with traffic engineering solutions, though proper data management can lead to identifying locations suitable for crash modification factors determined to be suitable for use. Outreach programs can also be created to help educate communities about the unique issues they face, but this will likely be more successful if the instruction comes from members of the community itself which is potentially why the NHTSA program has seen a relatively small impact in AIAN communities (Hill & Myers, 2016). Local engagement and collaboration is a necessity and should incorporate local traditions and culture to be effective. Effective tribal crash reporting can help identify areas needing greater attention and safety improvements. The data can then also be used to more successfully obtain funding to implement the improvements identified through data analysis (National Academies of of Sciences, Engineering, and Medicine, 2014).

Tribes do report some fatal crashes to the national Fatality Accident Reporting System (FARS) through agreements with the states and federal government, though this data is found to be significantly under the actual count of fatal crashes that occur on tribal roads (Ragland, Bigham, Oum, Chen, & Felschundneff, 2014). The underreporting could be due to several reasons, such as jurisdiction issues or available resources and training. Regardless of the cause, relying on FARS data is typically not sufficient enough for complete local analysis and is not sufficient for tribal use when applying for safety improvement funding. This project will fill this gap in the research by creating lasting relationships with Washington State tribes in order to achieve a higher level of local tribal traffic safety analysis. Besides the known high-risk factors for tribal crashes and recent push for programs supporting data-driven decision making, only some research has been completed to help realize the improved local data collection, management, and analysis. More research is needed in order to gain a better understanding of the crashes that occur on tribal roads and to improve the overall traffic safety in tribal communities.

Some efforts to improve local crash data collection and analysis have been completed in recent years. For instance, University of California – Berkeley researchers in their SafeTREC lab created a Tribal Traffic Safety Data tool. This tool uses their statewide crash data and overlays shapefiles from tribal lands and allows tribes to register and upload their own data. Importantly, this tool is only available to tribal members that have been verified, and the tool was created in collaboration with National Indian Justice Center (National Indian Justice Center, 2019). It was also emphasized that simply analyzing the current data is not sufficient enough, and that building connections and collaborating on the work is a crucial step towards action in improving traffic safety (Ragland, Bigham, Oum, Chen, & Felschundneff, 2014).

Researchers at the University of Wyoming developed a methodology to work with tribes to address the rural nature of their crashes and lack of crash data. The primary goal of the method was to identify collision hot spots, and a secondary goal was to address gaps in crash data collection. A case study with the Wind River Indian Reservation was also included to showcase the success of the methodology. Notably, part of the implementation plan was listed as "communication, coordination, and cooperation," though this seems to take prominence after the methodology is developed and implemented (Shinstine & Ksaibati, 2013). While the Wyoming project was successful, outreach activities should be prominent from the beginning in order to facilitate collaboration.

Work to conduct analysis and improve crash reporting has also been completed. Tribes recognize the numerous issues impacting their traffic safety, and ways to improve crash reporting have been developed. Of the ideas to improve crash reporting, the most difficult to overcome involves the political relationship between tribes and the state and national governments (Bailey & Huft, 2008). Other work has also showed that currently available data and processes can only provide a broad perspective on tribal traffic safety conditions, and that to obtain any substantial progress it is important to work with tribal communities at a much more localized level and in particular with tribes that have implemented programs to achieve the goal of reduced fatalities (Vichika, Carlson, & Schertz, 2015). Additionally, recent research has found geospatial information systems (GIS) to be particularly useful for tribal safety analysis, as a tool to both analyze and visualize crash data. Researchers from the University of Minnesota have successfully created tools based on GIS for hot spot identification, pedestrian crash analysis, and overall crash mapping (Horan, Hilton, Robertson, & Mbugua, 2018).

The body of research has found several notable conclusions regarding tribal traffic safety, primarily focused on how crash data is lacking and working at the community level tends to yield more promising results. However, work to form relationships with tribes prior to conducting the safety research or creating the tools and methodologies has been lacking. In order to truly understand the needs of the tribal communities, it is important to form these connections first. The outreach activities conducted for this research will help establish connections and foster understanding for safety work that follows. This framework relies on the strong relationships in order to be implemented successfully and repeatedly, as opposed to a one-time analysis.

Chapter 3. TRAFFIC SAFETY EQUITY METHODS

Previously, generalized regression models were introduced with the Poisson method being described in slightly more detail than other methods. This method is explored as part of this framework but not ultimately implemented due to violation of its assumptions. Other methods implemented as part of the framework in this study include the negative binomial model and the Empirical Bayes model to incorporate the results of the negative binomial in to more familiar, user-friendly terms. A Google Maps-based platform is used to visualize fatal and injury crashes in Washington State.

3.1 POISSON REGRESSION

The first method explored is a simple Poisson model. This model is classically used for count data, of which traffic collisions are a subset. This model was chosen given its prevalence. A Poisson model does have its weaknesses, such as typically being unable to handle overdispersion and underdispersion. Overdispersion occurs when the sample variance is greater than the sample mean, whereas underdispersion occurs when the sample variance is less than the sample mean. Both would violate a key assumption of the model, that the mean and variance are equal. The Poisson process outlines a discrete probability distribution for which the mean is constant in time. This is used in crash modeling because as a count model, it ensures that values cannot be negative. At a basic level, crashes can be assumed to occur at a relatively fixed time interval, which would make this model attractive. Equation 1 shows the typical probability mass function for a Poisson model.

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!} \quad \text{......} \quad (1)$$

Here, λ is the distribution mean which must be equal to the distribution variance. However, as described in the following sections, the data collected for this pilot violates this condition, making the Poisson model inappropriate for use.

3.2 NEGATIVE BINOMIAL REGRESSION

The second method used is the negative binomial distribution model. While traffic collision data tends to be suitable for count models, Poisson models are not often suitable due to the assumption that the mean and variance are equal. The advantage of the negative binomial model is that it is able to handle overdispersion of the data, or when the variance is greater than the mean. As will be shown, this is the case with the data obtained for this pilot. The negative binomial is less suitable than other models when the sample size is low, but given that 5 years of data are collected, the sample size should be adequate. The negative binomial has some different characteristics than the Poisson model. Equation 2 outlines how the negative binomial distribution variance is computed.

$$Var(Y) = E\big(Var(Y|\lambda)\big) + Var\big(E(Y|\lambda)\big) = E(\lambda) + Var(\lambda) \quad \ldots\ldots\ldots\ldots\ldots.. \quad (2)$$
$$= \alpha\beta + \alpha\beta^2 = \mu + k\mu^2;$$

Where k is a dispersion parameter equal to $1/\alpha$ and μ is the mean. In RStudio, where the models were implemented, the dispersion parameter k is known as theta. The negative binomial model is implemented twice: the first with the crash data and characteristics from the initial data merging process, and the second including an extra predictor in the form of degree of rurality. This is described in greater detail in Chapter 4.

3.3 EMPIRICAL BAYES

The final step in the crash frequency prediction module is the implementation of the Empirical Bayes method to obtain expected crash counts and Crash Reduction Potential based on the results of the negative binomial model. The Bayesian approach specifies that there is a prior distribution describing model parameters. The prior distribution of model parameters are deemed the hyperparameters, and they are estimated from the observed data. In order to use the Empirical Bayes method to determine road segments with greater Crash Reduction Potential (CRP), the estimated crash count is considered along with the actual crash count at each location. The estimated crash count comes from a safety performance function (SPF), which the negative binomial serves as for this framework. Prior to estimating the CRP, the expected safety condition must be calculated. Equation 3 shows this process.

$$\pi_i = w_i * SPF_i + (1 - w_i)K_i \quad \text{.....} \quad (3)$$

Where π is the expected safety condition of site i, w_i is a weighting factor between 0 and 1, the SPF is the result from the negative binomial model, and K_i is the observed crash count for segment i. The weighting factor, w_i, comes from the dispersion factor, which represents the variance of the SPF estimate and can be calculated with Equation 4.

$$w_i = \frac{1}{1 + SPF/kL^{\lambda}} \quad \text{.....} \quad (4)$$

Where k is the dispersion parameter from the negative binomial model, or theta when working in RStudio's MASS package, L is the length of the segment, and λ is a constant typically taken to be 0 in practice. From the expected safety values, it is possible to rank locations to determine areas that may need greater attention. Another option, and one preferred in this pilot, is to calculate the Crash Reduction Potential, which can be used to determine road segments which are underperforming in terms of safety as a hot spot identification tool. The Crash Reduction Potential is calculated with Equation 5.

$$CRP = (1 - w_i)(K_i - SPF_i) \quad \text{.....} \quad (5)$$

Note that if the K_i, the observed crash counts, is much greater than the SPF, the estimated crash counts, the CRP will be larger. Additionally, if the weighting factor is closer to 1, the variance in the SPF estimation is higher, which results in a lower CRP.

3.4 FRAMEWORK DESIGN

The framework for RITI traffic crash analysis involves 6 broad steps, which can loosely be grouped into 3 categories. Figure 6 shows the 6 steps grouped into their 3 categories for implementation stages. The steps include outreach, data collection, data merging, descriptive analyses, prediction modeling, and results visualization.

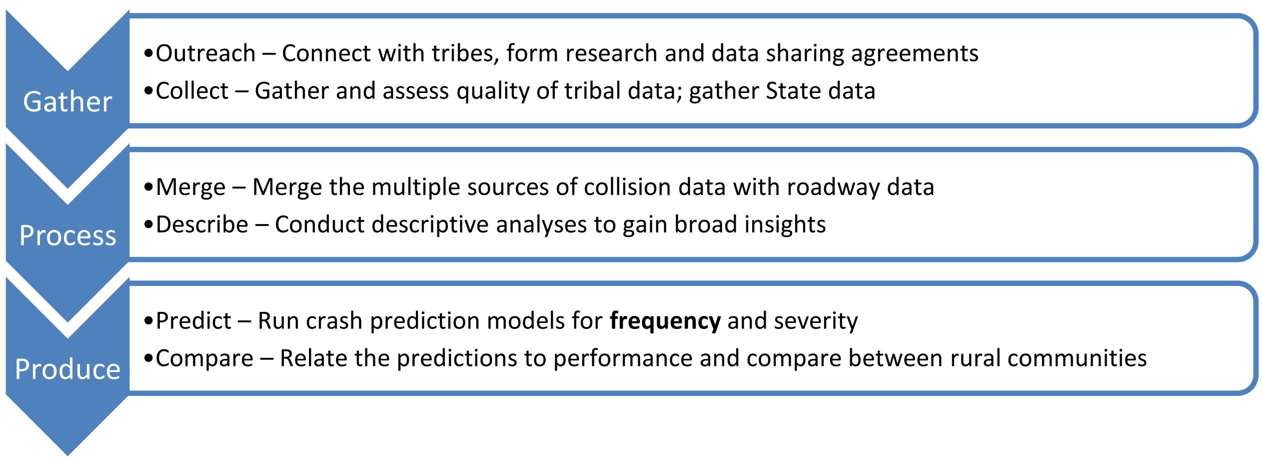

Figure 6 Framework Process Outline

The first step, outreach, serves as the foundation for the entire process and is arguably the most important. It is important that the outreach occur in order to form collaborative relationships and it is important to maintain the relationships. The outreach step is one that is never truly complete, as it is necessary to maintain direct communication with the RITI communities. Once the connections have been made, data sharing agreements should be formed and data obtained. Part of this should entail in depth discussion about what the current state of data collection and management is for the community, and if there are any plans to alter these. This can also entail some knowledge transfer for best practices. Additionally, if a baseline data format or organization has been established, this needs to be communicated. Once the data has been gathered, is must be merged with the state data. The state data should be merged prior to merging local community data to ensure that some degree of uniformity is maintained. For Washington State, crash and road characteristic data is available in many packages from the Washington State Department of Transportation (WSDOT), or in some pre-merged data from HSIS which aims to provide quality traffic safety data. The roadway segment data can be merged to the county level using ESRI ArcGIS. There are 4 tables from HSIS which can be merged for crash frequency prediction: roadway characteristics, crash records, roadway grade, and roadway curvature. These can be merged using a variety of tools, such as Python and the pandas and sqlite3 libraries. Descriptive analyses can be conducted to gain baseline understanding and insights from the merged data. Finally, the data can be used in the prediction models and also used in visualization tools such as maps and hot spot identification. Figure 7 shows another more detailed version of the framework that includes coordination with state agencies for improved data sharing and transportation improvement funding applications, which is a further goal of the project. The functions in the safety analysis module are proposed to be broken into two methods: crash severity prediction and crash frequency prediction. These are further defined as to relate to either local roads or to highways. The results are processed into a comprehensible and visual format and then distributed to the relevant parties. Currently, data does not come into the system in real time, but could in the future if outreach activities continue and relationships are developed further. It is also not yet stored on server supporting the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) developed by the Smart Transportation Applications and Research Laboratory

(STAR Lab), but could be in the future and it is proposed that tribal communities would maintain ownership of their data and control access to it.

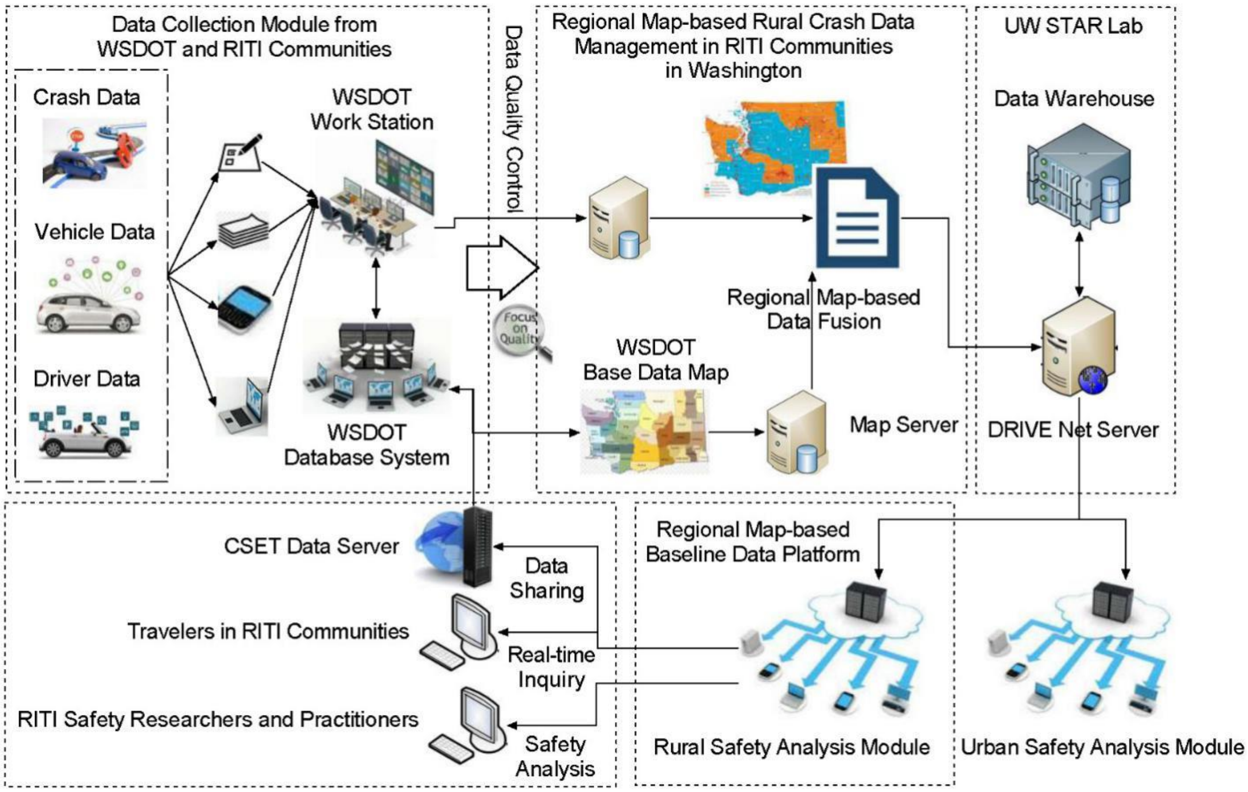

Figure 7 Detailed Framework for RITI Traffic Safety

Chapter 4. OUTREACH AND DATA COLLECTION EFFORTS

4.1 OUTREACH EFFORTS

To achieve the goal of this research, it was necessary to conduct outreach activities and to connect with Washington tribes and their leaders. In order to do so effectively, respectfully, and sustainably a meeting with tribal leaders at the University of Washington was organized. This served to build a profile of the AIAN communities in Washington State and identify key members to contact, in addition to learning more about the AIAN culture and current issues.

Mr. Iisaaksiichaa Ross Braine, director of the Intellectual House, an AIAN community center on the UW-Seattle campus, was the primary stakeholder at the meeting. He shared instruction, advice, and strategies for effectively communicating the project goals and purpose to other stakeholders that may be interested. He indicated that he wanted the project to succeed, was excited that this project was in place, and acknowledged traffic safety is an important public health issue that tribes are facing. Because of this, he was instrumental in connecting the researchers with tribal leaders in the transportation industry. Without this initial meeting, the researchers may not have been able to connect with the most interested parties and tribes, and future meetings would likely not have been as successful.

Advice given included emphasizing collaboration and commitment for continuous work with the tribes, as well as clearly communicating flexibility and willingness to follow tribes' wishes surrounding data privacy and ownership. Additionally, this meeting was crucial to the project development because it granted insight to ethical, cultural, and legal considerations that may be unique to tribal research, and thus require special attention. For instance, necessary steps to pursue a research agreement and what to expect from such an agreement was discussed. Researchers should expect to form the relationships with stakeholders and present their case prior to applying for a research agreement in order to facilitate and expedite the process. Some tribes may not be willing to enter into such an agreement for a variety of reasons, though with the help of the Director of the Intellectual House, some potentially interested tribes were identified. In fact, one of the key outcomes of this preparatory meeting was progress made to pinpoint which tribes might have the greatest initial interest in collaborating, and which stakeholders are already doing similar work or have implemented tribal safety plans.

Figure 8 shows the 29 federally recognized tribes in Washington State, each with its own characteristics, including varying sizes, organization, and resources. Each orange zone represents a reservation or tribal area. Clearly, their size varies quite substantially, and this is often reflected in their organization and available resources, which could impact their willingness to enter into a research agreement and share crash data. Table 1 lists these tribes, their size, whether they have a Tribal Transportation Plan in place, and the most recent year this plan was submitted. The Federal Highway Administration (FHWA) has a tribal traffic safety committee that has supported the adoption of Tribal Transportation Plans to help tribes successfully apply for grants. They have a publicly available list of tribes throughout the nation that have adopted plans. However, it is important to note that even if a tribe has not submitted a Tribal Transportation Plan to the FHWA, they can still apply for grants and may even have well-structured traffic safety committees or plans independently. If a tribe is known to have their own plans but has not submitted a Tribal Transportation Plan, this will be noted with a double-asterisk (**). This effort was conducted independently from the preparatory meeting and outreach activities but did reinforce what was learned from that meeting.

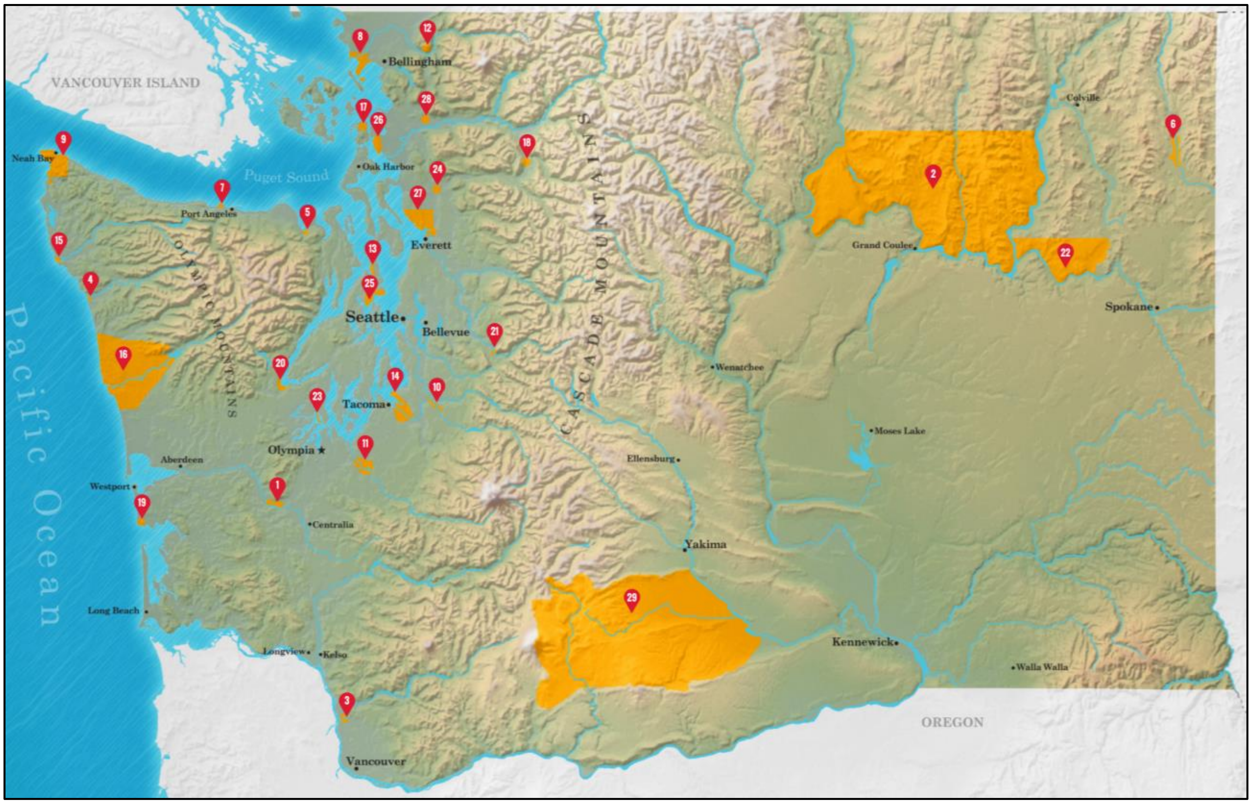

Figure 8 Washington Tribes Map

Table 1 List of Washington Tribes

Tribe Name	Enrolled Members	Size (acres)	Tribal Transportation Plan	Year Enacted
The Confederated Tribes of the Chehalis Reservation	833	4,438	Yes	2016
The Confederated Tribes of the Colville Reservation	9,365	1.4 million	Yes**	-
Cowlitz Indian Tribe	4,149	152	No	-
Hoh Indian Tribe	102	443	No	-
Jamestown S'Klallam Tribe	548	13.5 (1000+ owned outside reservation)	Yes	2016
Kalispel Indian Community of the Kalispel Reservation	470	292	Yes	2016
Lower Elwha Tribal Community	776	1,000	No	-
Lummi Tribe of the Lummi Reservation	4,483	13,000	Yes**	2015
Makah Indian Tribe of the Makah Indian Reservation	1,500	30,000	Yes	2011
Muckleshoot Indian Tribe	3,606	3,920	Yes	2016
Nisqually Indian Tribe	650	1,000	Yes	2011
Nooksack Indian Tribe	2,000	2,720	No	-
Port Gamble S'Klallam Tribe	1,234	1,303	No	-
Puyallup Tribe of the Puyallup Reservation	4,000	18,270	Yes	2015
Quileute Tribe of the Quileute Reservation	2,000	1,000	No	-
Quinault Indian Nation	2,453	208,150	No	-
Samish Indian Nation	1,200	200	No	-
Sauk-Suiattle Indian Tribe	200	23	No	-
Shoalwater Bay Indian Tribe of the Shoalwater Bay Indian Reservation	237	665	Yes	2017
Skokomish Indian Tribe	796	5,000	Yes	2014
Snoqualmie Indian Tribe	650	N/A	No	-
Spokane Tribe of the Spokane Reservation	2,153	154,000	Yes**	2015
Squaxin Island Tribe of the Squaxin Island Reservation	650	1,715	Yes	2016
Stillaguamish Tribe of Indians of Washington	237	64	Yes	2015
Suquamish Indian Tribe of the Port Madison Reservation	890	7,657	No	-
Swinomish Indian Tribal Community	778	8,155	Yes	2017
Tulalip Tribes of Washington	4,800	22,567	No	-
Upper Skagit Indian Tribe	504	84	Yes	2014
Confederated Tribes and Bands of the Yakama Nation	8,870	1,371,918	Yes**	2017

To turn the preparatory procedures into action, the researchers contacted the stakeholders that were recommended by the Director of the Intellectual House. This included emailing the people of interest to introduce the project and express interest in collaborating for a successful solution. There was a response from some but not all those who were contacted. This was expected based on the information given by the Director of the Intellectual House. Despite this, many did remember the initial email if the researchers were eventually able to meet and speak with them in person. It was important to have face-to-face meetings whenever possible in order to have more success and form stronger relationships. In person meetings are inherently more personal and fluid, which is invaluable when forming new relationships. From the initial preparatory meeting at the Intellectual House, the researchers were invited to attend the Tribal Leadership Summit, during which it would be possible to speak to some tribal leaders and gain connections to tribal transportation planners. Attending the event led to being invited to another conference, and from there several other connections and conferences were introduced. This was one way to mark progress of the project, as participation in conferences would increase from attendance, to sponsorship, to presentation. The conferences that were participated in included the aforementioned Tribal Leadership Summit, the Affiliated Tribes of Northwest Indians Conference, the Bureau of Indian Affairs Northwest Tribal Transportation Symposium, site visits. These events are discussed in detail in the following sections, including the level of participation and specific outcomes of each.

4.1.1 *Tribal Leadership Summit – 2018*

The researchers were invited to attend and present a poster during the annual Tribal Leadership Summit on May 11, 2018, where leaders from throughout Washington would convene to discuss tribal issues, projects, and future strategies with University leaders. The Tribal Leadership Summit took place at the Intellectual House on the University of Washington campus, which is a longhouse-style building that serves as a learning and gathering place for AIAN members of the University of Washington. The event was hosted by the Director of the Intellectual House, who knew specific stakeholders that would be interested in the project. While there was no transportation-centric discussion at the Summit, the researchers were able to speak with many tribal leaders and were subsequently invited to another, larger conference that would have a transportation section. The communications strategies learned from the Director were implemented in the presentation poster and proved vital. The Director made sure to introduce the researchers to key attendees and included an overview of the research in the Summit agenda and pamphlets. This was accomplished from the initial meeting with the Director of the Intellectual House, proving the importance of such a connection. In discussions with attendees, the researchers made sure to emphasize collaboration, tribal data ownership, and wishing to help the cultural heritage survive – focusing on traffic safety as a public health issue helped to root the research in some other relevant topics that were discussed at the Summit. Figures 9 shows the format of the Summit with several of the event organizers and tribal board members.

Figure 9 Tribal Leadership Summit 2018

4.1.2 *Affiliated Tribes of Northwest Indians Conference - 2018*

The second conference attended was the quarterly Affiliated Tribes of Northwest Indians (ATNI) Midyear Conference from May 21-24, 2018 in Yakama Nation at the Legends Casino Hotel. In addition to attending this conference, part of the CSET funding was used to become a silver sponsor of the event, allowing the logo to be displayed and project announced to the entire conference. This generated greater interest in the project with more people stopping by the poster after the announcement. Sponsorship was deemed crucial because it showed commitment to American Indian goals, community, and growth in the Northwest. There were many attendees at this conference, with several well-attended presentations on several key tribal issues. One of the breakout topics was concerned with traffic transportation planning and safety, during which the

Tribal Transportation Planning Organization (TTPO) held its quarterly meeting. Prior to, during, and following this meeting, the researchers met and interviewed tribal transportation leaders from different tribes in Washington State. There were leaders from 4 tribes that were especially interested, and one expressed interest in collaborating immediately. This began the process of applying for access to the crash data this tribe had collected. The tribe had received funding from the Washington Traffic Safety Commission to help support their own initiative of mapping fatal and serious injury crashes in the last decade. The result of this process was similar to the ultimate goal of CSET projects, and deemed a great place to start, but the process itself was not ideal because it was not sustainable and replicable. Figure 3 shows the researchers in attendance at the 2018 ATNI Conference. Following this conference, the interested tribal transportation leaders scheduled a phone conference to further discuss the project and invite the researchers to the Bureau of Indian Affairs Northwest Tribal Transportation Symposium, during which TTPO would have another meeting. The phone conference served to discuss the research plan in more detail and how it aligned with their own goals. Ultimately, the tribal planners were interested in continuing the discussion with a focus on basic data management, analysis, and visualization. Some interest was also expressed for DRIVE Net, a visualization platform created by the University of Washington STAR Lab with the support of WSDOT and Pacific Northwest Transportation Consortium (PacTrans), USDOT University Transportation Center for Federal Region 10. The link to WSDOT was seen as potentially useful given that WSDOT could possibly grant some funding for safety improvements. Figure 10 shows the researchers Dr. Ziqiang Zeng, Christopher Gottsacker, and Kris Henrickson with Tulalip Tribes board member and Vice President of ATNI Theresa Sheldon.

Figure 10 Researchers at ATNI 2018 with Theresa Sheldon, Vice President of ATNI

4.1.3 *Bureau of Indian Affairs Northwest Tribal Transportation Symposium – 2019*

The Bureau of Indian Affairs (BIA) held a tribal transportation conference in Spokane, WA in February 2019. The event consisted of networking, keynote addresses, and presentations by various tribal leaders and national traffic safety experts that work closely with tribes. The TTPO held another meeting as a part of this conference, which included a presentation by the tribe that had previously expressed interest in collaborating immediately (and with whom a research agreement was in the process of being written). The researchers gave their own presentation immediately following. The succession of one tribal safety project after another generated increased interest in the capabilities of the STAR Lab research team and in the DRIVE Net platform to analyze and visualize results. Subsequently, a site visit was organized for the tribal leaders to have a greater understanding of the technology used in the Lab, and how analysis models can be implemented and scaled to fit their own needs. The meeting site visit also helped the researchers understand the challenges facing the tribe and why they had to manually map serious injury and fatal crashes. It seems there is some resistance and challenges regarding the link between different departments, so that updates and information for a crash may not be transferred to each dataset. This presents a notable problem to solve and is one that the research team may be able to approach later. While this outreach project has been completed, it is crucial to note that outreach activities should never cease in order to maintain a trusting and mutually beneficial relationship with the tribes in Washington State. This conference presentation was primarily intended to showcase progress made in other related CSET projects, but new connections were made, and current relationships strengthened, thus continuing the impacts of this outreach project.

4.1.4 *Tribal Leadership Summit – 2019*

After the success of the Tribal Leadership Summit in 2018 the organizers invited the research team to attend in 2019 as well. However, at this year's Summit, tribal traffic safety was set to be a primary topic. Because of the success of the 2018 Summit, the organizers asked the team to present to the roundtable regarding the project and progress that had been made, and what was needed to further the research. This was one of the first examples of the outreach activities becoming successful, as tribal leaders were beginning to reach out as the reputation grew. This continued after the 2019 Tribal Leadership Summit as members of the Tulalip Tribes reached out following the event to inquire more about the project and set up a site visit to their reservation and meet with their planning department head, the chief of police, and other interested parties. This was another major point of progress for the project and illustrates the importance of continued outreach activities.

4.2 Data Collection

4.2.1 *Tribal Traffic Safety Data*

Using all these opportunities, communications, and follow ups, our research team was able to connect with leaders from 23 federally recognized tribes in Washington and ultimately established positive connections with twelve tribes, including the Confederated Tribes of the Colville Reservation (Colville), Spokane Tribe of Indians (Spokane), Muckleshoot Tribe (Muckleshoot), Swinomish Indian Tribal Community (Swinomish), the Confederated Bands and Tribes of the Yakama Nation (Yakama), Makah Tribe (Makah), Quinault Indian Nation (Quinault), Skokomish Indian Tribe (Skokomish), Puyallup Tribe (Puyallup), Lummi Nation (Lummi), Tulalip Tribes (Tulalip), and Sauk-Suiattle Indian Tribe (Sauk-Suiattle). Five of these tribes, i.e., Colville, Tulalip, Spokane, Muckleshoot, and Swinomish, have established strong connections with our research team. Ultimately, a formal research agreement with Colville has been signed, setting up an example of success for us to work with other tribes for safety data collection and analysis. There have also been promising meetings with Tulalip, with the hope that the Center for Safety Equity in Transportation will also enter research agreements with this tribe.

Two traffic collision datasets were obtained from Colville – Dataset A and Dataset B. These were similar, though Dataset A followed the standard police report more closely than Dataset B. However, Dataset B included GPS coordinates for collisions. In both datasets, only serious injury and fatal crashes were included. The outreach activities and relationship with Colville traffic safety leaders allowed to know more about the datasets. For instance, Dataset A was an output from the police report database with personal identifying information removed. Dataset B was manually created by the tribal traffic safety leaders, notably Nicole Ahlem, the Traffic Safety Coordinator in the Colville Tribal Public Safety department. Ms. Ahlem manually went to each location of serious and fatal collisions in recent years and recorded the GPS location with a smartphone app, along with some of the relevant information from the reports.

Dataset A contains 175 serious injury or fatality records for 2007 – 2018, with 144 traffic incidents. Dataset A includes fields such as incident severity; date; time; weather if known; location description; roadway jurisdiction (state route, county, residential) if known; whether the record is for a driver, passenger, or pedestrian; age; gender; race; seatbelt use; impairment; speeding; crash description; crash type; vehicle type; injury description; and number of people in the vehicle. Dataset B includes 170 serious injury or fatality records for 2007 – 2018, with 150 traffic incidents. Dataset B includes fields such as incident severity, date, time, gender, impairment, speeding, seatbelt use, collision comments, latitude, and longitude. The police records notably contain more descriptive and open-ended input fields than would typically be found in a state crash report, at least compared to the number of other fields. These datasets could be combined to a certain extent, however some of the records in each were unable to be matched and thus cannot be included. After merging the data, 138 complete unique records were included in the dataset. Of the merged data, 10 were from 2018 and were removed for comparative analyses, leaving 128 complete records for years 2007 – 2017. 54 incidents occurred on state routes, which is the analysis roadway for this research. 20 of these are also included in the state dataset, but notably none of the fatal incidents reported by Colville are contained in the WSDOT data. The Colville Reservation is located within Ferry and Okanogan counties. Roadway characteristic data was not able to be obtained from the Confederated Tribes of the Colville Reservation, but through outreach it was learned that GIS shapefiles for roadway inventory are in the process of being created. This is one reason that state routes were chosen as the roadway type for this research, but if the GIS data becomes available it could potentially be more detailed than what is publicly available through the Washington State local roads database. Also, because the data available does not include non-injury collisions, it is not as straightforward to model crash frequency or crash severity, as the models would be skewed. Data for years 20013 – 2017 was used from the state dataset which is discussed in the next section; there were 24 crashes on state routes recorded by the Confederated Tribes of the Colville Reservation for this same time period. Four of these were included in the statewide database, but again none of these were fatal crashes. Unfortunately, due to the low sample count of state route crashes that occurred, it was not possible to use the negative binomial model on the tribal data alone. Despite the small data volume, it is still eye-opening to have more information regarding the degree of underreporting. The 20 state route serious injury and fatal crashes over a 5 year period are just a small piece of information yet it changes some of the statistics for the state and region. Of the 24 crashes recorded, 8 involved a victim not wearing a seatbelt; 12 involved an impaired actor; and 8 involved high speed. There are more crashes that have these fields as unknown.

4.2.2 *Washington State Data*

Two different sources of data were used for Washington State crashes and roadway inventories. The first source of data is WSDOT. Geographic information system shapefiles were obtained through the WSDOT open data portal in order to determine the roadway segments of state routes in each county. Roadway characteristic data was also available in separate shapefiles from the open data portal, but these were not utilized due to the disconnected nature of the reference segments between each file. Instead, data was also requested from the Highway System Information System (HSIS) for designated state routes in Washington State from 2007 to 2017.

The HSIS data included collision records, road characteristic data, roadway curvature data, and roadway grade data for each segment and each year from 2007 – 2017. The road characteristic data includes many fields that would otherwise be difficult to merge from the WSDOT open data portal, such as lane width, speed limit, number of lanes, shoulder types, shoulder widths, and surface type. In order to merge the yearly HSIS roadway data to the crash data, the pandas and sqlite3 libraries in Python were used. Database tables were created for each year first, and these were merged into a final table. The table was exported to a .csv format which was then imported into an R script for cleaning and analysis. The HSIS data are not without their own limitations, however. A slightly more specific view was available due to the four data tables being available for each year under investigation, but for any number of reasons they do not necessarily share the same segment definitions year-to-year. One reason for this could be changes to route designation or if a route was altered. Regardless, this results in blank rows for some years for AADT and total crash count per segment. It was found that while using the HSIS data, merging data for years 2013 – 2017 was more reliable, with much fewer blank rows than the entire dataset due to these changes in segment definition. The entire range from 2007 – 2017 could not be used as crash records from WSDOT showed collisions occurring during years and at mileposts that were returned as NA in the merged HSIS database. Upon inspection of the database with merged data from 2007 – 2017, it was clear that there was a major disconnect beginning in year 2013. This also works in favor of using the Empirical Bayes method, as various temporal and spatial effects can come into play when a long range of years is used, and Empirical Bayes cannot easily handle these effects. Thus, the data merging program was rerun for years 2013 – 2017 with much more stable and reliable results. In sum, there were 48,157 collisions on 18,700 road segments designated as rural by WSDOT for the years 2013 – 2017. This data does not include severity codes so it was not possible to directly compare the results with the tribal data that was collected. For analysis, only one year of AADT and road segment definition is able to be utilized. Because the most recent information available via the open data portal from WSDOT is from 2014, this is the year of the AADT data used.

4.2.3 *Defining Rural*

One contribution of this study is the comparison of traffic crash analysis results under different definitions of what constitutes a rural roadway or area. The data from HSIS include a field to denote if a segment is rural or urban, and many of the segments are blank as well. WSDOT has a similar shapefile in their open data portal, which is the same data applied to the HSIS dataset. However, the methodology for determining what makes a segment rural or urban is difficult to uncover. The HSIS Washington guidebook lists population ranges under a rural-urban classification chart, shown below in Table 2 (Nujjetty, Mohamedshah, & Council, 2014).

Table 2 HSIS Rural-Urban Classification Table

WSDOT Urban/Rural Classification System	
Population Group	
0	Unknown
1	250,000+
2	100,000 - 249,999
3	50,000 - 99,999
4	25,000 - 49,999
5	10,000 - 24,999
6	5,000 - 9,999
7	2,500 - 4,999
8	< 2,500
9	Other rural areas

A major disadvantage to this is the lack of indication at what population level an area is designated as rural. Indeed, population alone would not be a great method to determine rurality. It might be assumed that the rural and urban designation comes from WSDOT and their functional classification system, which has been phasing out the process of using jurisdiction boundaries to label roadways as rural or urban and instead to label based on mobility and access (Washington Department of Transportation, 2013). However, this shift is not complete and so the WSDOT classification of urban and rural may not be consistent, making it potentially unreliable or at the very least not representative of communities necessarily. The classification is more useful at the microscopic level, in terms of classifying some roadways in a rural town as arterials despite potentially having significantly less traffic and conditions than an urban arterial. The determination of urban and rural by WSDOT is not necessarily meant to capture the context of an area; even if a rural town has arterials, the local population could be considered rural if, for instance, there is no emergency medical facility for a very large distance. Additionally, binary classification as urban or rural alone could strip away local characteristics; this thesis contends that there is a degree of urbanity and rurality that is worth an attempt to include in traffic safety analysis. There are numerous factors that could be included in the determination of what constitutes a rural area. Still, this classification will be considered in analysis described in the next chapter to serve as a baseline for what is easily available data.

As alluded to previously, some classifications rely on jurisdictional boundaries for urban and rural classifications. The most common boundary to use is perhaps the county boundary. This is due to the accessibility of data at the county level. The Washington State Office of Financial Management has created one such list of rural counties, with classification based on population density and county size (Washington State Office of Financial Management, 2019). Table 3 displays the counties in Washington State that are classified as rural under this definition. Crash data for these counties was obtained from WSDOT.

Table 3 Office of Financial Management Designated Rural Counties in Washington State

Counties with population density < 100 persons per square mile or counties smaller than two hundred twenty-five square miles as of April 1, 2018		
Adams (10.40)	Grant (36.33)	Pacific (22.97)
Asotin (35.24)	Grays Harbor (38.70)	Pend Oreille (9.67)
Chelan (26.64)	Island (402.30)	San Juan (96.66)
Clallam (43.22)	Jefferson (17.51)	Skagit (73.08)
Columbia (4.78)	Kittitas (19.85)	Skamania (7.18)
Cowlitz (92.12)	Klickitat (11.75)	Stevens (18.17)
Douglas (23.15)	Lewis (32.62)	Wahkiakum (15.57)
Ferry (3.53)	Lincoln (4.68)	Walla Walla (48.66)
Franklin (74.50)	Mason (66.73)	Whitman (22.79)
Garfield (3.11)	Okanogan (8.07)	Yakima (59.25)

The Washington State Department of Health (DOH) has outlined other definitions in use, such as the Rural-Urban Continuity Codes (RUCCs), which are used in this research, and the Rural-Urban Commuting Area Codes (RUCAs). RUCCs were first developed by the United States Department of Agriculture (USDA) and include 9 levels of rurality; 3 of these are grouped as metropolitan areas while 6 are grouped into non-metropolitan categories. RUCCs provide degree of rurality at the county level while RUCAs provide rural and urban status and relationships at the zip code level and census tract level (Hailu & Wasserman, 2016). However, inspection found RUCA codes to be less ideal due to local anomalies – areas known to be urban in nature and even within 1 mile of emergency hospitals can be classified as rural given the reliance on the inflow and egress travel volumes. This could be due to RUCAs being relatively newly developed and could be useful in future analysis. RUCCs still accomplish the goal of considering degrees of rurality by recognizing some urban areas are more urban than others, while some rural areas are more rural than others, and it is not necessarily a binary distinction. Being at the county level also makes it easier to merge this data with the roadway and crash data that were obtained. The use of RUCCs was also influenced by their promotion by the DOH, because traffic safety is ultimately a public health issue, especially when considering rural-urban disparity. The RUCC levels are outlined in Table 4 and the RUCC classification for each of the rural counties identified by the Office of Financial Management (OFM) are shown in Table 5. Note that not all these counties are characterized as rural and that there are counties classified as urban in both of these definitions that are included in the HSIS rural dataset. In total, rural traffic crash frequency prediction will be conducted incorporating definitions from WSDOT/HSIS and USDA.

Table 4 USDA Rural Urban Continuum Codes Definitions

Metropolitan Counties	
Code	Description
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of fewer than 250,000 population
Nonmetropolitan Counties	
4	Urban population of 20,000 or more, adjacent to a metro area
5	Urban population of 20,000 or more, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to a metro area
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area

Table 5 Rural Urban Continuum Codes for Office of Financial Management Designated Rural Counties

County Name (RUCC)		
Adams (6)	Grant (5)	Pacific (7)
Asotin (3)	Grays Harbor (4)	Pend Oreille (2)
Chelan (3)	Island (4)	San Juan (9)
Clallam (5)	Jefferson (6)	Skagit (3)
Columbia (3)	Kittitas (4)	Skamania (1)
Cowlitz (3)	Klickitat (6)	Stevens (2)
Douglas (3)	Lewis (4)	Wahkiakum (8)
Ferry (9)	Lincoln (8)	Walla Walla (3)
Franklin (2)	Mason (4)	Whitman (4)
Garfield (8)	Okanogan (6)	Yakima (3)

Chapter 5. RITI TRAFFIC CRASH DATA ANALYSIS

5.1 DESCRIPTIVE ANALYSES

5.1.1 *Tribal Data*

While the tribal data was not able to be included in the crash frequency prediction modeling, it is still valuable to investigate descriptive statistics and relevant information. For instance, on all roads, including both state route and local roads, from 2013 – 2017, there were 67 crashes recorded. Information about these crashes can be found in Table 6.

Table 6 Description of Tribal Crash Data, All Roads 2013 – 2017

Total Crashes	Serious Injury	# Injured	Fatal	# Fatalities	Impaired	No Seatbelt	Speeding
67	44	59	23	26	37	33	32

It is importation to note that these data cannot be interpreted to be constant for all tribal areas, and that there are roadway characteristics associated with each crash as well that could be contributing factors. These data measures can also not be compared to the state data as the state data did not include severity measures or factors such as impairment, seatbelt use, or speeding. These are more often included in crash severity prediction and thus these data are only summarized here; another application of the framework is crash severity prediction which would involve the same information to be requested for the state data as well.

5.1.2 *State Data*

The data obtained from HSIS can be investigated for means and variances as opposed to severity measures. Table 7 shows some of these relevant statistics about the crash data that was investigated in this pilot.

Table 7 HSIS Data Description

HSIS Data				
Number of Collisions	Number of Segments	Average Crashes per Segment	Variance	Max Crashes per Segment
48157	18700	2.575	26.944	171

From this information, it is clear that the variance of the data is much higher than the mean, exhibiting overdispersion and evidence that using the Poisson model for crash frequency prediction would be inappropriate. Results for one Poisson model are still shown in the next section to further examine the extent to which a negative binomial model is a better fit for the data. In addition, Figure 11 shows the distribution of the number of crashes per segment in the dataset. As one would expect, there are many zero values and otherwise low count values. However, there still are a significant amount of crashes on the segments, and some with relatively high counts, meaning that models such as a zero-inflated negative binomial might not be appropriate. Thus, the negative binomial model was chosen for other modeling as well.

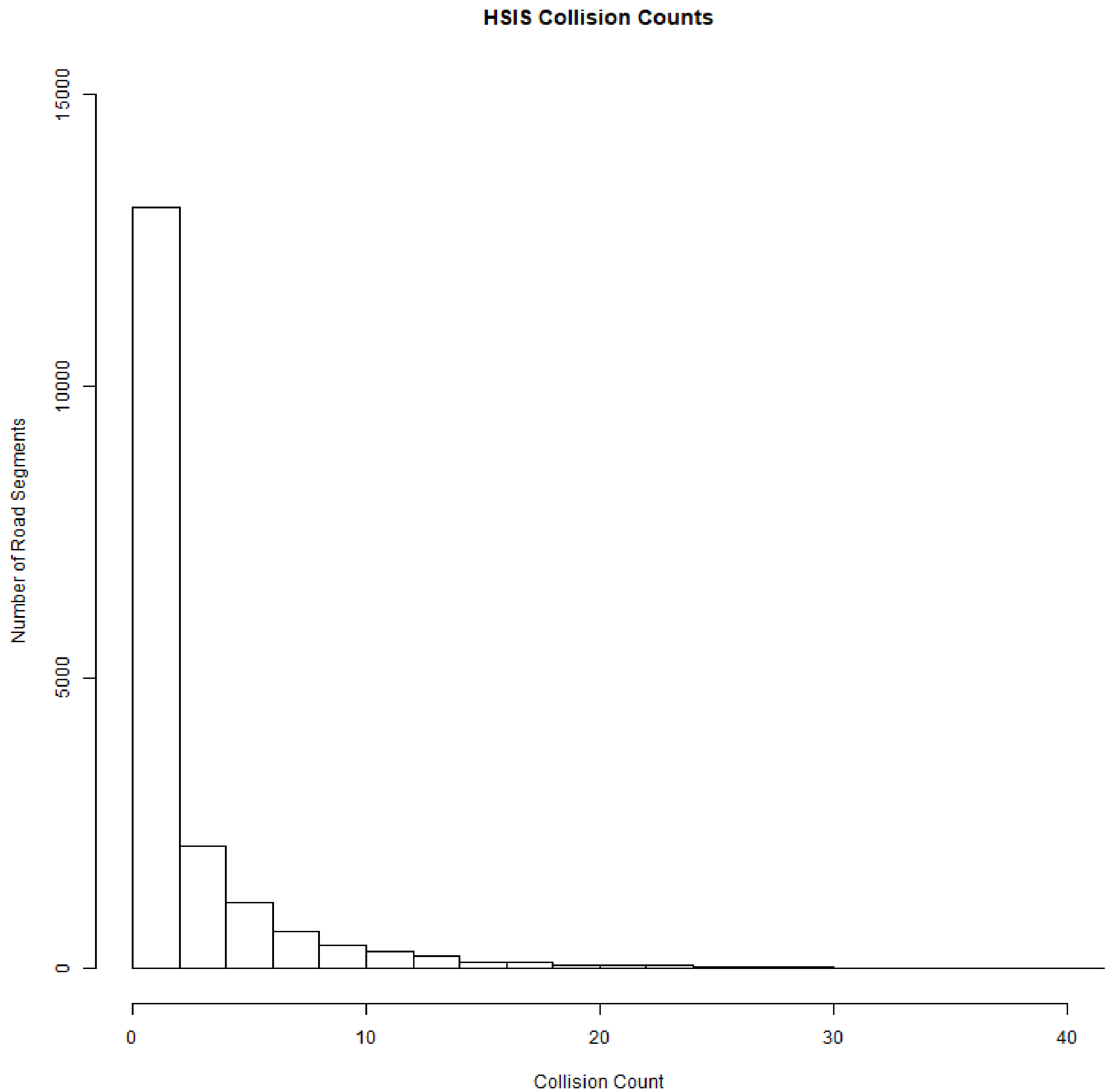

Figure 11 Distribution of HSIS Data

5.2 Traffic Crash Modeling Analysis

First, results for the Poisson and Negative Binomial models are presented for the HSIS data with state route segments labeled as rural under the WSDOT functional classification system. Second, results were compiled for the rural-designated segments but with RUCCs included as an independent variable. This was done to determine if this was a significant contributing factor that would result in a better fit. Tribal crashes that were not already included in the state dataset were added, though it is important to recall that these were only serious injury and fatal collisions. As more tribal data is obtained it could be possible to run a negative binomial model on solely tribal data to determine if there are a separate set of significant contributing factors. No other counties were included because it is impossible to determine which data from tribal crashes are recorded in the HSIS dataset for other counties. This is a step that will need to occur whenever new tribal data is obtained.

5.2.1 *Results for Highway Safety Information System Rural-Designated Segments*

HSIS was used to gather a baseline understanding of what could be the most readily available statewide data, with the built-in designation for each segment as rural or urban. Despite the descriptive statistics above, a Poisson model was run to show how results can vary under the assumption of equal mean and variance. The Poisson model results are not used to determine expected safety and crash reduction potential due to the violation of this condition. Nevertheless, Table 8 shows the Poisson results for the HSIS rural segment designation with all factors considered.

Table 8 Poisson Model Results

term	estimate	std.error	statistic	p.value
(Intercept)	-10.4921	63.5162	-0.1652	0.8688
lshl_typConcrete	***1.0169***	***0.0486***	***20.9218***	***0.0000***
lshl_typOther	-0.0902	0.1663	-0.5425	0.5874
lshl_typStructure	0.0394	0.0607	0.6493	0.5161
lshl_typUnknown	***-0.8410***	***0.1175***	***-7.1578***	***0.0000***
med_typeConcrete	-0.0577	0.0686	-0.8414	0.4001
med_typeOther	***-0.3033***	***0.0265***	***-11.4484***	***0.0000***
med_typeUnknown	***0.2982***	***0.0406***	***7.3379***	***0.0000***
rshl_typOther	0.1004	0.1668	0.6016	0.5474
rshl_typStructure	***-0.1729***	***0.0614***	***-2.8181***	***0.0048***
surf_typAsphalt	4.7717	63.5162	0.0751	0.9401
surf_typConcrete	4.8308	63.5162	0.0761	0.9394
surf_typOther	4.2833	63.5168	0.0674	0.9462
spd_limt	***0.0026***	***0.0006***	***4.6113***	***0.0000***
lanewid	***0.0115***	***0.0034***	***3.3883***	***0.0007***
no_lanes	***-0.0889***	***0.0080***	***-11.0966***	***0.0000***
lshldwid	***-0.0184***	***0.0029***	***-6.2850***	***0.0000***
rshldwid	***-0.0072***	***0.0028***	***-2.5906***	***0.0096***
medwid	***0.0003***	***0.0001***	***6.8287***	***0.0000***
avg_grad	***-0.0132***	***0.0028***	***-4.6725***	***0.0000***
curv_count	***0.0283***	***0.0022***	***12.9806***	***0.0000***
max_deg_curv	***0.0060***	***0.0009***	***6.5228***	***0.0000***
log_length	***0.5896***	***0.0040***	***147.5929***	***0.0000***
log_aadt	***0.8816***	***0.0068***	***130.2826***	***0.0000***
AIC	***81744***			

The significant predictors are in bold and italics in the above table. There are many that do make intuitive sense, but perhaps most striking is the intercept not being significant, which is something that should be expected. The AIC is also included for reference when completing the negative binomial model. Table 9 shows the results for the negative binomial model when initially run with all predictors included except for RUCC. The negative binomial models were created and run using the MASS package in RStudio, using the natural log link function.

Table 9 Initial Negative Binomial Model Results

term	estimate	std.error	statistic	p.value
(Intercept)	-26.0166	114739.8	-0.00023	0.999819
lshl_typConcrete	***0.8873***	***0.15495***	***5.72649***	***1.03E-08***
lshl_typOther	-0.14457	0.221364	-0.65307	0.513711
lshl_typStructure	-0.01537	0.09698	-0.15844	0.874106
lshl_typUnknown	***-0.86035***	***0.16245***	***-5.29605***	***1.18E-07***
med_typeConcrete	***-0.46471***	***0.19505***	***-2.38253***	***0.01719***
med_typeOther	***-0.54456***	***0.06274***	***-8.68012***	***3.95E-18***
med_typeUnknown	***0.34595***	***0.0841***	***4.11353***	***3.90E-05***
rshl_typOther	0.105544	0.222293	0.474799	0.63493
rshl_typStructure	***-0.20136***	***0.09789***	***-2.05686***	***0.0397***
surf_typAsphalt	19.70612	114739.8	0.000172	0.999863
surf_typConcrete	19.62159	114739.8	0.000171	0.999864
surf_typOther	19.47519	114739.8	0.00017	0.999865
spd_limt	***0.00403***	***0.00101***	***4.00922***	***6.09E-05***
lanewid	***0.01539***	***0.00526***	***2.92678***	***0.00342***
no_lanes	-0.01321	0.019663	-0.67162	0.501826
lshldwid	***-0.01481***	***0.00557***	***-2.66016***	***0.00781***
rshldwid	0.002559	0.005228	0.489389	0.624566
medwid	***0.00027***	***0.00015***	***1.79538***	***0.07259***
avg_grad	***-0.01727***	***0.00585***	***-2.95405***	***0.00314***
curv_count	***0.05674***	***0.00604***	***9.38861***	***6.08E-21***
max_deg_curv	***0.00687***	***0.00169***	***4.06612***	***4.78E-05***
log_length	***0.49816***	***0.00778***	***64.0037***	***0***
log_aadt	***0.88671***	***0.01291***	***68.6574***	***0***
AIC	***65320***			

Again, the intercept is not found to be significant, but the AIC is much lower than the Poisson model. This is a measure of the comparative goodness of fit between the two models, with a lower AIC signifying a better fit. This is one other reason why the negative binomial was chosen. Table 10 shows the results of the negative binomial model when it is run with only the significant predictors from the previous model definition. For simplicity, only the final iteration is shown here. The negative binomial model is specified in Equation 6.

$$\ln(\mu) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad \ldots\ldots\ldots\ldots\ldots\; (6)$$

Table 10 Final Negative Binomial Results for Rural Segments

term	estimate	std.error	statistic	p.value
(Intercept)	-4.2160	0.0731	57.7049	0.0000
spd_limt	-0.0059	0.0009	-6.5992	0.0000
avg_grad	-0.0129	0.0060	-2.1646	0.0304
curv_count	0.0576	0.0062	9.3314	0.0000
max_deg_curv	0.0054	0.0017	3.1762	0.0015
log_length	0.4966	0.0076	65.4050	0.0000
log_aadt	0.7339	0.0094	77.8293	0.0000
AIC	66006			

The dispersion parameter for this model is 1.0362 and the dispersion parameter standard error is 0.0200. A t-statistic can be found by dividing the dispersion parameter by its standard error, resulting in 51.8849. With 5 degrees of freedom, this is significant at the 0.05 level and thus adequately accounts for the overdispersion in the data and is an appropriate model. The final predictors that were found to be significant include the speed limit of a segment, the average grade, the amount of curves in a segment, the maximum degree of curvature if there is a curve, the log of the length of the segment, and the log of the annual average daily traffic. The log was taken of these last two predictors in order to reduce the range and skew the results. The AIC is slightly higher than the initial model, but the results do make better intuitive sense and the AIC values are fairly similar. The speed limit was found to have a decreasing impact. This implies that a higher speed limit reduces the overall number of crashes on a segment, which may seem counterintuitive. A likely reason is that segments with higher speed limits are likely associated with better visibility and curvature conditions. Rural highways with higher speeds could also have greater safety features incorporated into their design. It is also possible that there were many crashes on segments with lower speed limit when previous segments had high speed limits; this change could result in high speed crashes being recorded in lower speed limit segments. The average grade was found to have a decreasing impact on crash frequency as well, meaning that with increased grade the crash frequency is slightly reduced. This could be due to a large amount of crashes occurring on flat or zero grade segments, of which there are many more than high grade segments. It is also possible that crashes are more likely to occur immediately prior to or following a change in grade, or similar to the speed limit could be the result of graded segments having some safety features in place already. These are two contributing factors that deserve greater investigation to determine their true impact. The remaining significant factors all had increasing impacts, meaning that crash frequency increases as their value increases. As expected, length and average annual daily traffic have the largest coefficients, as these are most directly associated with the risk and exposure associated with any one segment.

5.2.2 *Results When Considering Rural-Urban Continuum Codes*

Using the same set of data from the HSIS, the county of the collision could be retrieved from each crash record. Thus, it was also possible to link each segment to its Rural-Urban Continuum Code. The negative binomial model was run with the new set of predictors, including the RUCC. The results of the final model with only significant predictors chosen to be in the model definition are shown in Table 11. This model had a dispersion parameter of 1.1051 and a dispersion parameter standard error of 0.0218. A t-statistic can be found by dividing the dispersion parameter by its standard error, resulting in 50.6975. With 7 degrees of freedom, this is well beyond a significance at the 0.05 level, meaning that this model appropriately accounts for the overdispersion present and is suitable.

Table 11 Negative Binomial Results When Including RUCC

term	estimate	std.error	statistic	p.value
(Intercept)	-4.3987	0.0965	45.5671	0.0000
rucc	-0.0887	0.0051	17.4436	0.0000
lanewid	0.0097	0.0037	2.6464	0.0081
no_lanes	-0.1891	0.0116	16.2333	0.0000
avg_grad	-0.0188	0.0059	-3.1913	0.0014
curv_count	0.0619	0.0060	10.2493	0.0000
max_deg_curv	0.0055	0.0017	3.2930	0.0010
log_length	0.4997	0.0076	65.7805	0.0000
log_aadt	0.7979	0.0120	66.2400	0.0000
AIC	65450			

These results show that the AIC is lower than when RUCC is considered, meaning it is a better fit for the data. The number of lanes and lane width are also found to be significant when including RUCC as a predictor. These results suggest that adding a degree of rurality to analysis can help improve the results of a model when predicting traffic crash frequency on rural roadways. Speed limit was not found to be a significant predictor when RUCC was considered, which perhaps can signify the previous results regarding speed limit vary depending on model specification – more research is needed to determine the effect of speed limit on crash frequency on rural highways. Lane width has an increasing impact on crash frequency suggests increasing lane width is related to an increase in crash frequency; this is counterintuitive to urban areas as often times narrow lane width is considered less desirable in practice. These results could come from a very high amount of roadway segments with one lane width (12 feet) and some segments being recorded with increasingly wider lane widths and very few narrow lanes. The number of lanes was found to be significant with a negative impact, meaning that crash frequency decreases as more lanes are added. This could be intuitively true but could also be influenced by the number of segments that included only 2 lanes, with some segments having more lanes but very few having only 1 lane. The grade, curve count, maximum degree of curvature, length, and average annual daily traffic are all reported as having the same sign and relatively similar coefficient values compared to the negative binomial model without considerations for RUCC. The RUCC predictor has a decreasing impact, suggesting that crash frequency decreases with an increase in rurality. This could be due to more rural areas having lower average annual daily traffic, reducing the exposure of a segment. These results do not investigate the severity of crashes with different levels or rurality.

5.3 EMPIRICAL BAYES METHOD

With the negative binomial model completed with only significant predictors, it is possible to use the Empirical Bayes method to convert the results to more actionable items. One such item is an index of expected safety, and another is an index of Crash Reduction Potential (CRP). Both of these could be used to help determine where more focus should be applied and perhaps more funding allocated to address concerns. Table 12 shows the summary of the expected safety as determined from Equation 3 compared to the summary of the actual data. It can be noted that the expected safety is mostly accurate, though tends to underestimate the number of crashes when the number of crashes exceeds the mean.

Table 12 Summary of Results for Expected Safety and Actual Safety

Expected Safety					
Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
0.024	0.495	1.010	2.575	2.692	158.635
Actual Safety					
Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
0	0	1	2.575	3	171

The CRP could also be calculated using the results of the negative binomial model, following Equation 5. These results can be plotted to show the index of crashes and the Crash Reduction Potential. Figure 12 shows such a plot with red indicating a CRP greater than 3 and Blue indicating a CRP greater than 1. This could be used to identify crash hot spots that potentially have the greatest ability to have increased safety based on the negative binomial and Empirical Bayes results. There are a total of 2,569 road segments with a CRP greater than 1.0, with 1,191 road segments having a CRP greater than 3.0. A CRP of 1.0 represents approximately the 86th percentile, while a CRP of 3.0 represents approximately the 94th percentile. These are both the upper end of percentiles, and often times in transportation engineering the 85th percentile is used as a starting point of estimates – thus, the use of 1.0 as a cutoff in this example. It would be relatively straightforward to alter the results of this analysis framework component to fit the needs of the user.

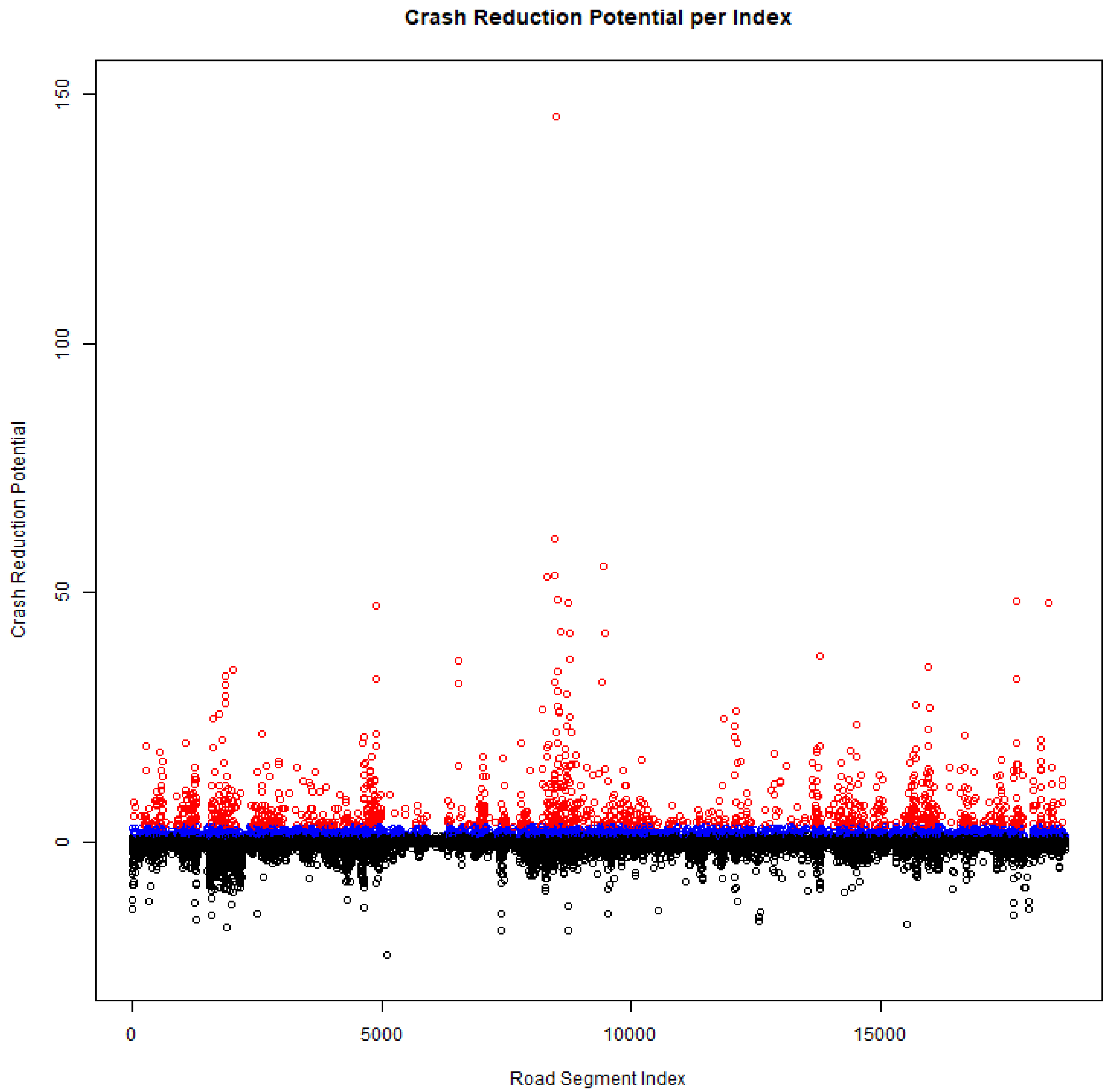

Figure 12 Crash Reduction Potential per Road Segment Index

5.4 Traffic Crash Visualization

Other tools are developed as a part of this framework to help communities gain a better understanding of their crashes. One such tool developed has been a crash mapping tool based on the familiar Google Maps interface that allows for data to be uploaded if it has coordinates. The model proposed here includes serious injury and fatal collisions throughout the state on the rural segments. The interface allows users to filter to years and severity, and allows a user to focus on one selected area to gain better insights. The user can also select specific crashes or groups of crashes to understand the factors that were recorded with each crash and proportions of fatal to serious injury if selecting a group of crashes. Figure 13 shows a screenshot of the interface when viewing a larger portion of the mapping application. Figure 14 shows a screenshot of the application when selecting a specific crash to view its characteristics. Tribal data was not included in this demo.

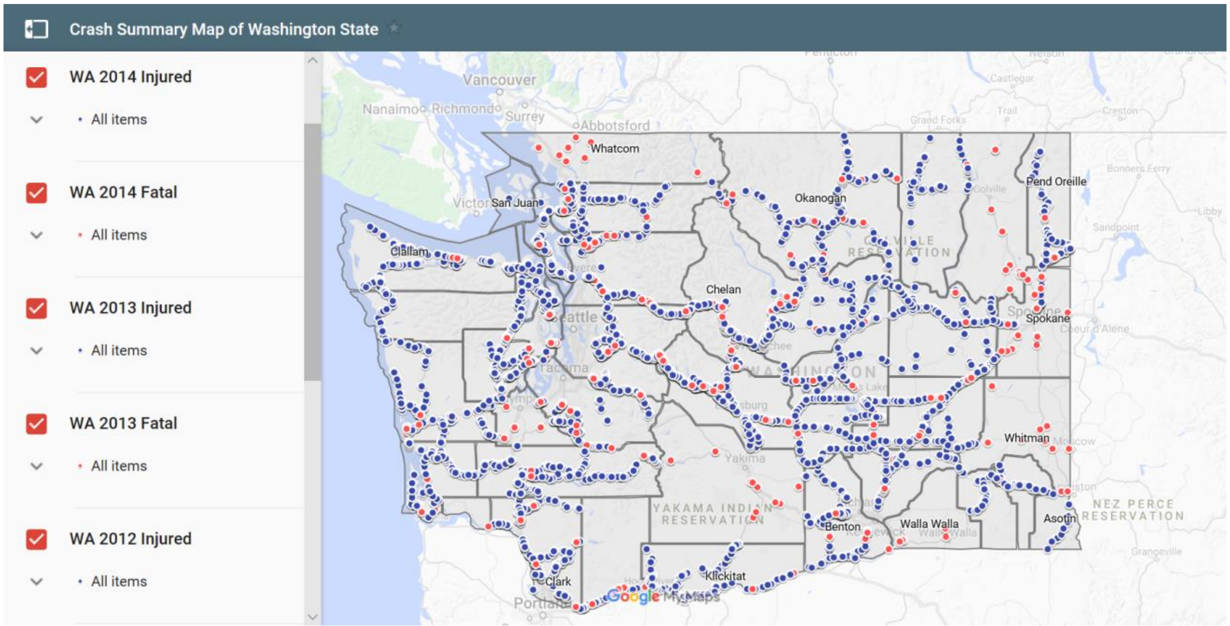

Figure 13 Year and Crash Severity Options in the Google Maps Crash Visualization Application

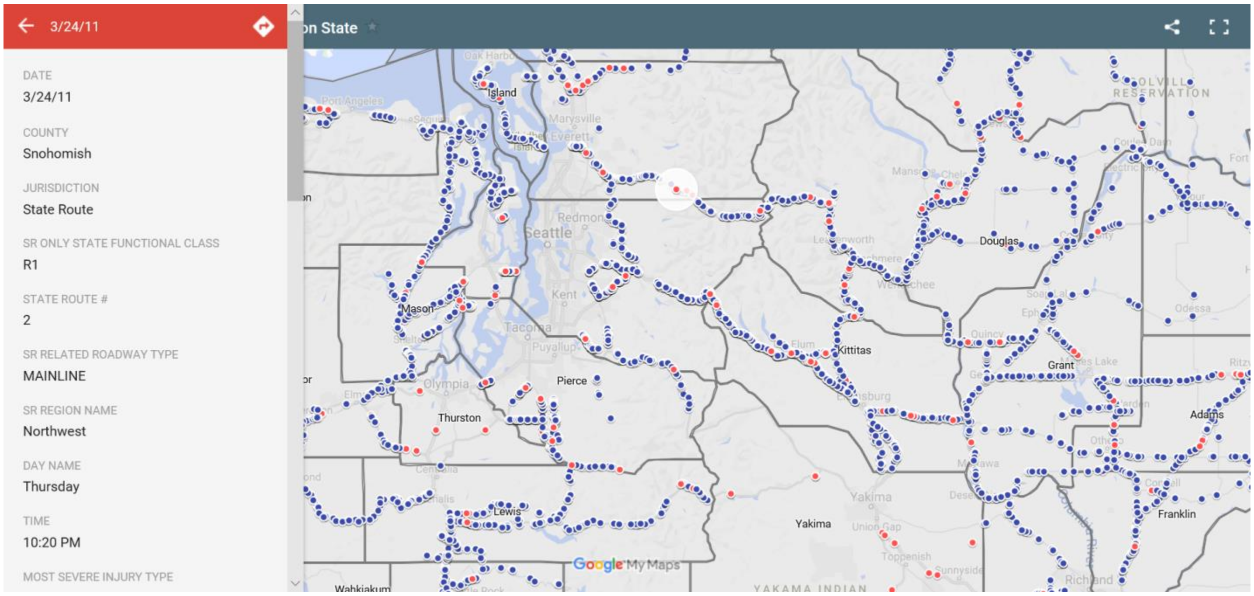

Figure 14 Crash Detail View in the Google Maps Crash Visualization Application

Chapter 6. CONCLUSION

This research is an important step towards better understanding rural, isolated, tribal, and indigenous community traffic safety issues by presenting a simple and lightweight framework. The proposed framework begins with outreach activities, and follows to data aggregation, data merging, data descriptive analyses, crash prediction analyses, and finally visualization of some pertinent results. While other research has focused on rural and tribal traffic safety, there is a lack of research initiatives aimed at forming lasting collaborative relationships with the RITI communities. The outreach portion of the framework is not merely a step to follow but a paradigm to maintain during the process. By connecting first with local tribal liaisons at the University of Washington Intellectual House, it was possible to gain a better understanding of the community and learn how to successfully communicate the project and form connections. This proved vital and traffic crash data was obtained in part from the Confederated Tribes of the Colville Reservation after entering a formal research agreement. The inclusion of this data and being able to determine the degree of underreporting from just one region of Washington State is significant, even if the data could not be included in crash frequency prediction models. Another aspect of this framework was the inclusion of a degree of rurality in the prediction models, which allowed for a greater context to be considered and resulted in a better fitting model. The results show that significant contributing factors to rural crashes include the degree of rurality, the lane width, the number of lanes, the average grade, the number of curves, the maximum degree of curvature, the log of length, and the log of annual average daily traffic. Being able to map the crashes was an important consideration learned from collaborating with tribal transportation leaders, and an application to rank the road segments according the expected safety and Crash Reduction Potential is also proposed as a method to aid in transportation improvement proposals. While the issue of traffic safety inequity cannot be solved overnight, it is one that can be addressed through thoughtful collaborations with RITI communities and efficient solutions.

BIBLIOGRAPHY

Abdelwahab, H. T., & Abdel-Aty, M. A. (2001). Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board, 1746*(1), 6-13.

Aguero-Valverde, J., & Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis & Prevention, 38*(3), 618-625.

Aguero-Valverde, J., & Jovanis, P. P. (2008). Analysis of Crash Frequency with Spatial Models. *Transportation Research Record: Journal of the Transportation Research Board, 2061*(1), 55-63.

Bailey, L., & Huft, D. (2008). Improving Crash Reporting: Study of Crash Reporting Practice on Nine Indian Reservations. *Transportation Research Record: Journal of the Transportation Research Board, 2078*(1), 72-79.

Castro, M., Paleti, R., & Bhat, C. R. (2012). A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transportation Research Part B: Methodological, 46*(1), 253-272.

Center for Safety Equity in Transportation. (2019). *CSET Mission*. Retrieved from http://cset.uaf.edu/mission/

Hailu, A., & Wasserman, C. (2016). *Guidelines for Using Rural-Urban Classification Systems for Community Health Assessment.* Washington State Department of Health.

Hill, M. L., & Myers, C. S. (2016). Creating a Culture of Traffic Safety on Reservation Roads: Tribal Law & Order Codes and Data-Driven Planning. *The Indigenous Peoples' Journal of Law, Culture & Resistance*, *3*(1), 43-80.

Horan, T., Hilton, B., Robertson, J., & Mbugua, J. (2018). *Using GIS to Improve Tribal Safety: Applications, Trends, and Implementation Dimensions.* University of Minnesota, Center for Transportation Studies. Roadway Safety Institute.

Leaf, W. A., & Solomon, M. G. (2005). *Safety belt use estimate for Native American tribal reservations.* National Highway Traffic Safety Administration.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291-305.

Lord, D., Geedipally, S. R., & Guikema, S. D. (2010). Extension of the Application of Conway-Maxwell-Poisson Models: Analyzing Traffic Crash Data Exhibiting Underdispersion. *Risk Analysis, 30*(8), 1268-1276.

McDonald, G., Davie, G., & Langley, J. (2009). Validity of Police-Reported Information on Injury Severity for Those Hospitalized from Motor Vehicle Traffic Crashes. *Traffic Injury Prevention, 10*(2), 184-190.

Muelleman, R., Wadman, M., Tran, T., Ullrich, F., & Anderson, J. (2007). Rural Motor Vehicle Crash Risk of Death is Higher After Controlling for Injury Severity. *The Journal of Trauma: Injury, Infection, and Critical Care, 62*(1), 221-226.

National Academies of of Sciences, Engineering, and Medicine. (2014). *Guide for Effective Tribal Crash Reporting.* Washington, DC: The National Academies Press.

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

National Indian Justice Center. (2019). *Tribal Traffic Safety Data Center*. Retrieved February 16, 2019, from https://www.nijc.org/online-tool-training-manual.html

Nujjetty, A. P., Mohamedshah, Y. M., & Council, F. M. (2014). *HSIS Guidebook for State Data Files: Washington.* Federal Highway Administration Office of Safety and Office of Safety Research and Development.

Ragland, D., Bigham, J., Oum, S., Chen, K., & Felschundneff, G. (2014). *Traffic Injury on Tribal Lands in California.* UC Berkeley: Safe Transportation Research and Education Center.

Shankar, V. N., Albin, R. B., Milton, J. C., & Mannering, F. (1998). Evaluating Median Crossover Likelihoods with Clustered Accident Counts: An Empirical Inquiry Using the Random Effects Negative Binomial Model. *Transportation Research Record: Journal of the Transportation Research Board, 1635*(1), 44-48.

Shinstine, D. S., & Ksaibati, K. (2013). *Indian reservation safety improvement program : a methodology and case study.* University of Wyoming, Department of Civil and Architectural Engineering. Wyoming Department of Transportation.

Vichika, I., Carlson, P., & Schertz, G. (2015). Review of Tribal Transportation Safety. *Transportation Research Recod: Journal of the Transportation Research Board, 2531*(1), 153-160.

Wang, X., & Abdel-Aty, M. (2006). Temporal and spatial analyses of rear-end crashes at signalized intersections. *Accident Analysis & Prevention, 38*(6), 1137-1150.

Washington Department of Transportation. (2013). *Guidelines for Amending Functional Classification in Washington State.* Transportation Data & GIS Office.

Washington State Office of Financial Management. (2019). *Population density and land area criteria used for rural area assistance and other programs*. Retrieved from Office of Financial Management: Washington Data and Research: https://www.ofm.wa.gov/washington-data-research/population-demographics/population-estimates/population-density/population-density-and-land-area-criteria-used-rural-area-assistance-and-other-programs

Washington Traffic Safety Commission. (2016). *Washington State Strategic Highway Safety Plan 2016: Target Zero.*

Washington Traffic Safety Commission. (June 2018). *Motor Vehicle Deaths in the American Indian/Alaska Native Community.*

Wu, Z., Sharma, A., Mannering, F. L., & Wang, S. (2013). Safety impacts of signal-warning flashers and speed control at high-speed signalized intersections. *Accident Analysis & Prevention, 54*, 90-98.

Yamamoto, T., Hashiji, J., & Shankar, V. N. (2008). Underreporting in traffic accident data, bias in parameters and the structure of injury severity models. *Accident Analysis & Prevention, 40*(4), 1320-1329.

Zeng, Z., Zhu, W., Ke, R., Ash, J., Wang, Y., Xu, J., & Xu, X. (2017). A generalized nonlinear model-based mixed multinomial logit approach for crash data analysis. *Accident Analysis & Prevention, 99*(Part A), 51-65.