Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-22/0-7051-R1		
4. Title and Subtitle Develop a Real-Time Decision Support Tool for Rural Roadway Safety Improvements		5. Report Date December 2021 6. Performing Organization Code
7. Author(s) Subasish Das, Kay Fitzpatrick, Zihang Wei, Srinivas Geedipally, Eun Sug Park, Lingtao Wu, Stephanie Paal, and Ioannis Tsapakis		8. Performing Organization Report No. Report 0-7051-R1
9. Performing Organization Name and Address Texas A&M Transportation Institute The Texas A&M University System College Station, Texas 77843-3135		10. Work Unit No. (TRAIS) 11. Contract or Grant No. Project 0-7051
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office 125 E. 11th Street Austin, Texas 78701-2483		13. Type of Report and Period Covered Technical Report: September 2019–January 2022 14. Sponsoring Agency Code
15. Supplementary Notes Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration. Project Title: Develop a Real-Time Decision Support Tool for Rural Roadway Safety Improvements URL: http://tti.tamu.edu/documents/0-7051-R1.pdf		

16. Abstract

Speeding-related crashes are one of the dominant types of crashes on rural roadways. As a result, the Texas Strategic Highway Safety Plan has identified speeding as one of the seven research emphasis areas for 2017–2022. Additionally, inclement weather plays a role in crash occurrences. This study used crash data with the inclusion of roadway geometry, traffic volume, 5-minute interval operating speed data, and 5-minute interval weather data to develop safety performance functions for the major rural roadway facility types: rural two-lane roadways, rural multilane divided roadways, rural multilane undivided roadways, and rural freeways. The models were developed based on annual and daily levels. The results of the annual-level analysis show that increased variability in operational speed is associated with increased crashes for all four roadway facilities. Average precipitation shows a positive association for only rural multilane undivided and rural two-lane roadways. The results in the daily-level analysis show that increased variability in daily operational speed (the standard deviation of the daily average of 5-minute interval operation speeds) is associated with increased day-level crashes for all four roadway facilities. As daily average precipitation increases, the likelihood of a day-level crash occurrence increases for all facility types. This project also developed an interactive web-based decision support tool that assists practitioners in understanding safety scoring of roadway segments of interest with respect to not only roadway or traffic characteristics but also operating speed and weather conditions.

17. Key Words	18. Distribution Statement
Rural Roadways, Safety Evaluation, Safety Performance Functions, Decision Support Tool	No restrictions. This document is available to the public through NTIS: National Technical Information Service Alexandria, Virginia http://www.ntis.gov

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	133	

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

Develop a Real-Time Decision Support Tool for Rural Roadway Safety Improvements

by

Subasish Das, Ph.D.
Associate Research Scientist

Kay Fitzpatrick, Ph.D., PE, PMP
Senior Research Engineer

Zihang Wei
Graduate Research Assistant

Srinivas Geedipally, Ph.D., PE
Research Engineer

Eun Sug Park, Ph.D.
Senior Research Scientist

Lingtao Wu, Ph.D., PE
Associate Research Engineer

Stephanie Paal, Ph.D.
Assistant Professor

and

Ioannis Tsapakis, Ph.D.
Research Scientist

Report 0-7051-R1
Project 0-7051
Project Title: Develop a Real-Time Decision Support Tool
for Rural Roadway Safety Improvements

Performed in cooperation with the
Texas Department of Transportation
and the
Federal Highway Administration

January 2022

TEXAS A&M TRANSPORTATION INSTITUTE
College Station, Texas 77843-3135

DISCLAIMER

This research was performed in cooperation with the Texas Department of Transportation (TxDOT) and the Federal Highway Administration (FHWA). The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of FHWA or TxDOT. This report does not constitute a standard, specification, or regulation.

This report is not intended for construction, bidding, or permit purposes. The principal investigator of the project was Subasish Das, and Kay Fitzpatrick served as the co-principal investigator.

The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

ACKNOWLEDGMENTS

This research was conducted in cooperation with TxDOT and FHWA. The Texas A&M Transportation Institute (TTI) would like to thank TxDOT staff Heather Lott and Christopher Weber for providing valuable help, information, data, and advice throughout this project. TTI also gratefully acknowledges the support and assistance provided by TxDOT project manager Shelley Pridgen.

Project team members met with numerous other individuals at TxDOT to gather and/or complement data and information needed for the analysis. TTI would also like to thank Valerie Vierkant and Magdalena Theel for gathering and processing various data sets that were analyzed in this study.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ACRONYMS, ABBREVIATIONS, AND TERMS

AADT	Annual average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
ASOS	Automated Surface Observing System
CBR	Cost-benefit ratio
CMF	Crash modification factor
CMV	Commercial motor vehicle
CRIS	Crash Records Information System
DFO	Distance from origin
DOT	Department of transportation
GEE	Generalized estimating equation
Green Book	*A Policy on Geometric Design of Highways and Streets*
FFS	Free-flow speed
FHWA	Federal Highway Administration
GIS	Geographic information system
GPS	Global positioning system
HSIS	Highway Safety Information System
HSM	*Highway Safety Manual*
IRI	International Roughness Index
KABC/KABCO	Injury scale for fatal (K), incapacitating injury (A), non-incapacitating injury (B), possible injury (B), possible injury (C), and no injury (O) crashes
MCSSD	Mean of cross-sectional speed standard deviation
mph	Miles per hour
MUTCD	*Manual on Uniform Traffic Control Devices*
NB	Negative binomial
NCHRP	National Cooperative Highway Research Program
NHS	National Highway System
NMSL	National Maximum Speed Limit
NOAA	National Oceanic and Atmospheric Administration

NPMRDS	National Performance Management Research Data Set
NPV	Net present value
PDO	Property damage only
RHiNO	Roadway-Highway Inventory Network Offload
RID	Roadway Information Database
RWIS	Road weather information system
SDCSM	Standard deviation of the cross-sectional speed mean
SEM	Structural equation modeling
SHRP-2	Strategic Highway Research Program 2
SHSP	Strategic Highway Safety Plan
SMS	Space mean speed
SPF	Safety performance function
Texas WB	*Texas Roadway Safety Design Workbook*
TMAS	Travel Monitoring Analysis System
TMC	Traffic message channel
TTI	Texas A&M Transportation Institute
TxDOT	Texas Department of Transportation
VMT	Vehicle miles traveled
VOR	Value of research
vpd	Vehicles per day
VSL	Variable speed limit

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

Speeding is considered as one of the key contributing factors on rural roadway related traffic crashes. To address this issue, the Texas Strategic Highway Safety Plan (SHSP) has considered speeding as one of the seven research emphasis areas for 2017–2022. Additionally, inclement weather plays a role in crash occurrences. Traditional crash risk analysis typically excludes real-time speed, real-time volume, and weather or precipitation information. To alleviate this critical research gap, operational variables are needed to be considered in the modeling framework. The following three national databases can be considered in mitigating this research gap:

- National Performance Management Research Data Set (NPMRDS) with passenger car and truck speed data sets for the National Highway System (NHS) and other roadways.
- Travel Monitoring Analysis System (TMAS) data with traffic volume data through both temporary traffic counting and continuous traffic counting programs.
- National Oceanic and Atmospheric Administration (NOAA) with five-minute interval weather information such as precipitation.

Traffic crashes on the rural roadways are disproportionate in comparison to the crashes occurring on urban roadways. In 2017, 53 percent of traffic crash fatalities in Texas occurred on different facilities of rural roadways. There is a need for further investigation to improve safety on rural facility types, particularly focusing on the SHSP areas of emphasis (including speeding and inclement-weather-related crashes). In this study, the Texas A&M Transportation Institute (TTI) developed safety evaluation models for rural highways in Texas. Models were developed for both annual-level data and daily-level data. Moreover, an interactive decision support tool was developed to help visually illustrate the risk of roadway networks. The findings of this research can be integrated into TxDOT's future vision plans.

1.2 PROJECT GOAL AND RESEARCH TASKS

TTI outlined two goals for this study, which are summarized as follows:

- **Goal 1:** Develop a conflated database for analysis by incorporating roadway inventory, real-time traffic volume, operating speed, weather, and crash data while establishing a framework for the conflated database that will permit updating with an influx of new data. Develop facility-specific rural roadway safety performance functions (SPFs) with the inclusion of roadway, operating speed, traffic volume, and weather data.
- **Goal 2:** Develop a decision support tool with interactive mapping options.

In order to achieve the project goals, TTI conducted three major tasks, summarized as follows:

- **Task 1. Data Conflation Work:** TTI conflated the data from NPMRDS, the Crash Records Information System (CRIS), TMAS, and NOAA. The conflated data set has the crashes from CRIS snapped onto the base unit of analysis, and each unit of analysis has a definitive number of crashes (total and by severity level) associated with it.
- **Task 2. Safety Evaluation of Rural Roadway Networks:** TTI developed a conceptual data framework to organize collected data. Based on the data prepared, TTI developed two safety evaluation models based on two data aggregation intervals: an annual-level model and a daily-level model.
- **Task 3. Framework and Decision Support Tool:** Based on the modeling and exploratory results in Task 2, TTI developed a geographic information system (GIS)–based prototype decision support tool that can estimate and visually illustrate the risk on the roadway network. TTI used the open-source software platform Shiny[1] to develop the decision support tool.

1.3 REPORT ORGANIZATION

The remaining chapters of this report include the following:

- **Chapter 2: Literature Review:** an overview of the methods that can be used to perform highway safety evaluations.
- **Chapter 3: Data Preparation:** a brief overview of the data and the data conflation framework.
- **Chapter 4: Model Development:** description on developed models.
- **Chapter 5: Decision Support Tool:** the decision support tool the research group developed for this study.
- **Chapter 6: Conclusions and Recommendations:** the key research findings and recommendations.
- **Appendices:** the software code for the decision support tool and the value of research.

[1] More information on Shiny is available at https://shiny.rstudio.com/.

CHAPTER 2:
LITERATURE REVIEW

2.1 INTRODUCTION

This chapter provides a synthesis of relevant studies.

2.2 SPEED MEASURES IN THE *HIGHWAY SAFETY MANUAL*

Despite the necessity to include speed measures in highway safety evaluations, speed measures are not directly included in the SPFs for the first edition of the *Highway Safety Manual* (HSM) (American Association of State Highway and Transportation Officials [AASHTO], 2010) and are not proposed for the upcoming second edition of the HSM. Appendix 3E ("Speed and Safety") in the first edition of the HSM provides some context about speed measures and their effect on overall safety. The HSM also incorporates crash modification factors (CMFs) for the average operating speed change (for the before and after periods of the crash occurrence); however, these measures need to be reexamined due to the availability of newer data sources such as the NPMRDS.

2.3 RELATIONSHIP BETWEEN SPEED AND SAFETY

Research has shown that a vehicle's impact speed (but not prevailing speeds during travel) affects the injury severity of the roadway users, and that the speed differential between drivers affects crash frequencies. Rosén and Sander (2009) showed that fatality risks are highly associated with impact speed. However, there is minimal consistent evidence for the direct association between operational speed measures and crash counts although intuitively speed plays a major role in safety. It is anticipated that speed influences crash likelihood, but this association is more intricate and not as well understood.

National Cooperative Highway Research Program (NCHRP) Report 504 lists the speed percentiles for rural highways with various posted speed limits (Fitzpatrick et al., 2003). In addition, some researchers argue that speed variance should be considered in crash data analysis (Montella and Imbriani, 2015). Additional insights can be drawn from the point of view of speed-design consistency where speed is studied in relation to road alignment and safety (Banks et al., 2014). More detailed information about the speed-safety relation is summarized in the following subsections within three broader speed-related topics:

- Speed limit.
- Design speed.
- Operating speed.

2.3.1 Speed Limit

Speed limits form the basis for corridor-level speed enforcement strategies. A suitable speed limit can provide a safe, steady, and realistic speed to ensure the safety of all road users (including vehicles, cyclists, and pedestrians) along the roadway. Because drivers are not always capable of selecting suitable speeds, speed regulation measures such as speed limits are crucial as they provide specific assistance for proper speed choices while driving (Elvik, 2010). Figure 1 illustrates different forms of speed limits.

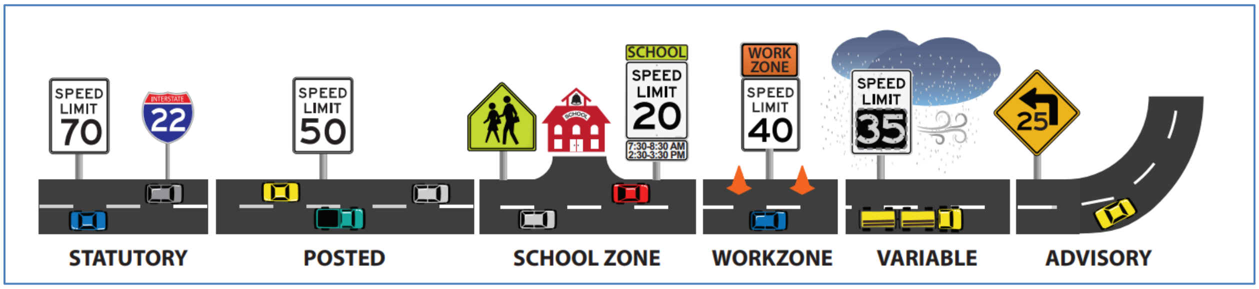

Figure 1. Types of Speed Limits (Source: Federal Highway Administration [FHWA], 2020).

2.3.1.1. Statutory Speed Limit

Statutory speed limits for roadway facility types are established by state legislatures and can vary by state. According to FHWA's *Speed Limit Basics*, "a statutory speed limit is enforceable by law and applicable even if the speed limit is not posted. Examples include 25 mph in residential or school districts, 55 mph on rural highways, and 70 mph on rural freeways" (FHWA, 2020).

2.3.1.2. Posted Speed Limit

According to FHWA's *Speed Limit Basics*, "posted speed limits (sometimes called regulatory speed limits) are those that are sign-posted along the road and are enforceable by law. A posted speed limit could be the same as the statutory speed set by the State legislature, or it could be established by a city, county, or state transportation agency as an adjustment to the statutory speed limit" (FHWA, 2020).

2.3.1.3. Advisory Speed Limit

Advisory speeds are designed to improve the safety at roadways alignments such as horizontal and vertical curve locations. These speed limits are usually set using an engineering speed study by following the guidance from the *Manual on Uniform Traffic Control Devices* (MUTCD) (FHWA, 2020). Bonneson et al. (2007) recommended that the advisory speed limit be based on the mean operating speed of truck drivers.

2.3.1.4. Variable Speed Limits

The variable speed limit (VSL) approach is an integral part of intelligent transportation systems. These systems can modify the speed limit based on real-time traffic conditions or predefined speed control algorithms on different road segments. Currently, many states in the United States (i.e., Arizona, New Mexico, Oregon, and Washington) have adopted VSL.

VSL practices have been observed to reduce overall speeds (Garber and Srinivasan, 1998; Ullman and Rose, 2005). VSLs are assumed to be an effective countermeasure for preventing speed-related crashes as well as helping to control congestion, especially in work zones (Ullman and Rose, 2005; Levin et al., 2019). De Pauw et al. (2018) showed that the number of injury crashes decreased significantly (an 18 percent reduction in total crashes and a 6 percent reduction in fatal and severe crashes) with the implementation of VSL. Saha et al. (2015) examined the interaction between road, weather, and crashes on rural mountainous freeway corridors for VSL. The results showed that interaction between horizontal curvatures and vertical grades with inclusion of weather-related variables had a substantial impact on crash occurrence.

2.3.1.5. Studies on Speed Limit and Safety

In the U.S., the states assign speed limits for different facility types. Until 1974, the speed limit was set as high as 75 miles per hour (mph). In 1974, Congress established the National Maximum Speed Limit (NMSL), which set the maximum speed limit at 55 mph. New legislation in 1987 allowed some rural and urban freeways to increase the speed limit to 65 mph. In November 1995, the NHS Designation Act of 1995 gave the states the power to set posted speed limits for their roadway facilities. Many researchers examined the consequences of speed limit change for different facility types.

In two earlier studies (Forester et al., 1984; Copulos, 1986), researchers examined the relationship between fuel usage (i.e., consumption, cost, and overall economic cost) and the lowering of posted speed limits. The findings indicated that the lowering of speed limits resulted in savings in fuel consumption. Later, the research focus was shifted toward safety. Some studies showed that the NMSL implementation was associated with fatality changes (Burritt et al., 1976; Dart, 1977).

Lave (1985) argued that crash severity might increase because of variance in the speed rather than from higher speed since slow drivers can potentially be as dangerous as drivers who select higher speeds. By performing a before-after observational study, Baum et al. (1989) studied the fatality rates of the speed limit change in 1987 on rural freeways in 38 states. The findings indicated that fatalities on rural freeways with higher speed limits were 15 percent more than the expected value. Farmer et al. (1997) also analyzed the impact of a speed limit increase from 70 to 75 mph (on freeway segments) for 12 states and compared safety performance to the list of 18 states that had maintained almost the same speed limit. The analysis showed an increase in fatalities of 16 percent for the study group and 4 percent for the comparison group. Parker (1997) investigated the safety outcome of posted speed limit change for non-limited-access roadways in 22 states. This study carefully selected the sites to represent different groups for which the safety effects of speed limit changes could be evaluated. The results generated several quantitative safety measures due to the changes in posted speed limit.

Vernon et al. (2004) studied the impact of raised speed limits on crash rate, fatality rate, and injury crash rates for Utah highways. The findings showed that the crash rate for urban freeways and the fatality rate for non-freeway highways increased. Kweon and Kockelman (2005) studied the safety implications of the raised speed limits using a large data set along seven freeways and 143 state highways in Washington State. The results of this study suggested that a hypothetical 5-mph speed limit increase on a road with average characteristics would not have a statistically significant impact on the count of fatal crashes.

Using crash data from Washington State, the results of Kockelman et al. (2006) demonstrated that around a 1 percent increase in the total crash numbers and a 13 percent increase in the fatal crash count was associated with a speed limit increase from 65 to 75 mph. Malyshkina and Mannering (2008) evaluated the impact of Indiana's speed limit increase in 2005. Their findings showed that the increased speed limits on rural freeways did not significantly affect crash severities. The authors said that a decrease in speed variability resulted from the increased speed limit. In contrast, on some multilane highways, the authors found that the increase in speed limits significantly increased the severity of crashes.

Islam et al. (2014) used Connecticut freeway data to develop SPFs for single-vehicle and multivehicle crashes. The models are developed for total and fatal/injury (FI) crashes. The results showed that consideration of the interaction between the speed limit and geometric variables (i.e., width of shoulder, number of lanes, shoulder width, and type of median) improves the accuracies of the model estimates.

Gayah et al. (2018) examined the operational and safety impact of setting speed limits lower than engineering recommendations. The authors performed an empirical Bayes before-after analysis using data from 41 miles of rural roads in Montana where the posted speed limit was decreased (the comparison group consisted of 131 miles of roadway). This study developed several CMFs for posted speed limit changes. Monsere et al. (2018) investigated the association between operating speed and traffic crashes for 1,400 miles of Oregon roadways where speed limits increased in 2016. The results showed that average operating speeds increased with the increase of posted speed limit.

In 2010, the Virginia Department of Transportation (VDOT) increased the posted speed limit from 65 to 70 mph for a portion of rural freeways. Himes et al. (2018) performed an empirical Bayes before-after study to examine the safety and operational effects of this change. No increase in any of the focus crash types at an aggregate level was observed. The disaggregate analysis showed some impacts: segment type influenced safety, and interchange segments experienced statistically significant increases in total, roadway departure, and truck involved traffic crashes.

Indiana rural freeways have several speed limits, including passenger cars (70 mph) and vehicles with a gross weight of 26,000 pounds or more (65 mph). For most urban freeways, the speed limit is 55 mph but varies at certain locations (from 50 mph on some downtown sections to 65 mph on some suburban sections). Tarko et al. (2019) assessed the impact of traffic mobility and safety for alternative speed limit scenarios on Indiana freeways. Posted speed limit is associated with safety and mobility mostly in non-congested traffic conditions. The findings showed no impact on congested conditions.

Several international studies also focused on posted speed limits and associated safety impacts. Jaarsma et al. (2011) explored the safety outcome of posted speed limit reduction from 80 to 60 km/h (50 to 37 mph) on rural roads in the Netherlands. The findings of the before-after with comparison group analysis revealed a statistically significant 24 and 27 percent overall reduction in fatal and fatal plus injury crashes, respectively. De Pauw et al. (2014) explored the safety impacts of decreasing speed limits from 90 to 70 km/h (56 to 44 mph) on roads in the Flemish Region of Belgium. Considering 61 of the treated sites with a total length of 116 km (72 mi), this study performed a meta-analysis by using the effectiveness of each section. The meta-analysis demonstrated a non-significant 5 and 6 percent decrease in injury and severe injury crashes, respectively.

Sayed and Sacchi (2016) evaluated the safety effects of raising speed limits on a rural highway in British Columbia, Canada, after a speed limit change in 2013. The authors found the speed limit increases to be associated with a statistically significant 11 percent increase in fatal and injury crashes. Imprialou et al. (2016) examined the association between crash, speed, and road geometry on United Kingdom motorways to estimate the impact of a potential speed limit increase (from 70 to 80 mph). The findings demonstrated that single-vehicle crashes of all severities and fatal or severe injury crashes involving multiple vehicles increased due to this change.

2.3.1.6. Studies on Speed Limit and Operating Speed

With a survey result of 128 speed limit intervals, Fitzpatrick et al. (2003) observed that the speeds of the 85th percentiles were higher than the posted speed limit and that 50th percentile speeds were close to the speed limit. Experts generally suggest that the 85th percentile speed only be used as the standard for the speed limit. Burritt et al. (1976) found that as the speed limit decreases, the average operating speed also reduces. Some researchers (Ossiander et al., 2002; Upchurch, 1989) found that with a speed limit increase from 55 to 65 mph, the mean operating speed increases by 2 to 7 mph.

Based on this summary, the speed limit or a metric closely linked to the speed limit appears to be an important predictor for the driver's speed choice. Fitzpatrick et al. (2003) examined the relationships between the posted speed limits, operating speeds, and design speeds as documented in NCHRP Report 504 (see Table 1). The authors acknowledged that the 85th percentile operating speed often exceeds the posted speed limits (Harkey et al., 1990). Parker (1997) showed that 64 percent of the vehicles adopted the posted speed limit, 86 percent met or exceeded the posted speed limit by 5 mph, and 97 percent met or exceeded the posted speed limit by 10 mph.

Table 1. Percentile Speeds Given the Posted Speed Limit.

Location	Posted Speed Limit (mph)	Percentile of Operating Speeds			Number of Sites
		Speed Limit	Plus 5 mph	Plus 10 mph	
Rural	50	81	99	100	12
	55	61	85	96	151
	60	91	95	98	8
	65	59	89	98	2
	70 (Fitzpatrick et al., 2003)	64	91	98	7
	50–70 (Parker, 1997)	64	86	97	126

McCarthy (2001) suggested that speed limit changes impact not only the average operating speeds but also the distribution and dispersion of drivers' speed profiles. Some studies showed that an increase in the posted speed limits appears to trigger higher operating speeds (Freedman and Esterlitz, 1990; Freedman and Williams, 1992; Retting and Cheung, 2008; Hu, 2017). Garber and Gadiraju (1989) found that the higher speed variances were in fact associated with relatively lower average traffic speeds on highways with posted speed limits of 55 mph. Parker (1997) also revealed that increasing and decreasing posted speed limits did not have a significant impact on the changes in the speed variances. Based on the documented studies, although posted speed limit can have a significant impact on a driver's speed choice and the average operating speed percentiles, the impact of the vehicle speed variance and speed distributions requires in-depth exploration.

2.3.2 Design Speed

Berry and Belmont (1951) defined design speed as "the highest continuous speed at which individual vehicles can travel with safety upon a highway when weather conditions are favorable, traffic density is low, and the highway design features are the governing conditions." Different versions of *A Policy on Geometric Design of Highways and Streets* (the Green Book) reveal different definitions of design speed:

- **Before 1954:** "the maximum appropriately uniform speed that probably will be adopted by the faster group of drivers but not, necessarily, by the small percentage of reckless ones" (AASHTO, 1940).

- **1954–2001:** “the maximum safe speed that can be maintained over a specified section of highway when conditions are so favorable that the design features of the highway govern” (AASHTO, 1984).
- **2001–2010:** “a selected speed used to determine the various geometric design features of the roadway” (AASHTO, 2004).

The design speed of a roadway is the speed used to determine minimal values for various design elements. Example geometric elements that may be influenced by design speed include the horizontal curvature radius, braking distance, horizontal sightline offset, length of vertical curvature, maximum super elevation of the roadway, maximum side friction factor, stopping sight distance, lane widths, shoulder inside widths, and shoulder outside widths. The design speed perception does not warrant design consistency with posted speed limits and operating speeds since designers are encouraged to use conservative design controls that are often higher than the minimum values identified in the Green Book. The use of design speed also varies. Many departments of transportation (DOTs) exclusively base their minimum criteria on the design speed, while other DOTs and transportation agencies may use a variety of geometric metrics.

2.3.3 Operating Speed

2.3.3.1. Operating Speed Measures

Data on the speed and travel times of passenger vehicles and trucks are essential for traffic engineers responsible for the design and operation of streets and highways. Operating speed is the driver-selected speed that can be observed during prevailing conditions. Operating speeds tend to be normally distributed and, as such, can be characterized by mean speed and standard deviation (Donnell et al., 2009; Donnell et al., 2018). Some of the key operating speed measures are described as follows:

- **Spot speed:** the instantaneous speed of a vehicle passing at a specified location.
- **Time mean speed:** the arithmetic average speed of all vehicles for a specified period. It is also known as mean speed or average speed. It is the simple average of spot speed and is associated with a point over time.
- **Space mean speed (SMS):** the average speed of all vehicles measured at an instant of time while traveling a given length of the roadway. It is a harmonic mean. SMS is always less than time mean speed. SMS also counts the mean measures of spot speed. However, the weightage is on spatial and not on temporal.
- **Standard deviation (SD):** a statistical measure to determine the centrality of the data. SD of speed measures is the square root of variance (the average of the squared differences from the average) of speed measures.

- **Percentile speeds:** the speed at or equal to which the indicated percentage of vehicle groups is traveling, such as 15th (slow-speed group), 50th (mid-range speed group), and 85th (high-speed group). The 85th percentile has had a strong historic relationship to speed limit, set in the belief that most drivers select a rational speed and desire to minimize their risk.
- **Free-flow speed (FFS):** the speed when there are no constraints placed on a driver by other vehicles, geometric constraints (e.g., horizontal curve or vertical curve), or traffic control devices (e.g., traffic signals) on the road.
- **10-mph pace speed:** This range contains the largest percentage of vehicles in a distribution of spot speeds at a location.
- **Speed dispersion:** Unlike the key operating speed measures (i.e., percentile speed and space mean speed), research on the characteristics of speed dispersion is very sparse. Vehicle speed dispersion is the main topic of speed research. Solomon (1964) first proposed the concept of vehicle speed dispersion. This study introduced a U-shaped curve (see Figure 2). Speed dispersion is defined as the difference in vehicle speeds. Researchers have used various vehicle speed distribution characteristic indicators to define vehicle speed dispersion based on the different purposes, methods, and data limitations of their research. For example, Wang et al. (2018) proposed two speed dispersion measures: the SD of the individual speed and the average speed difference of two neighboring vehicles.
- **Additional speed measures:** Two recent studies (Hutton et al., 2020; Das et al., 2020) used a wide range of operating speed measures. These speed measures are discussed in the next subsection.

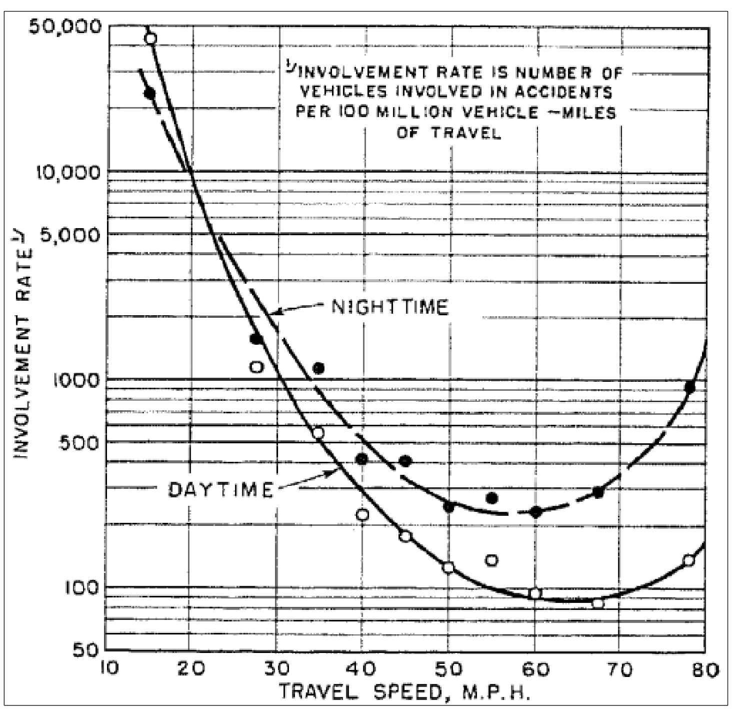

Figure 2. Solomon's Curve (Source: Solomon, 1964).

Findings from the current studies show that reduction of speed variation on a roadway facility is associated with the reduction of crash likelihood, and crash likelihood can be made minimal by keeping all vehicles operating with same speed in an ideal infrastructure. Note that Solomon (1964) study is the first study which took a deeper look at that the likelihood of traffic crash occurrence and the difference between an individual driver's speed and the average speed on the road, especially on rural roadways. The interpretations of the study's outcomes have been passed down through generations of traffic engineers. It is important to note that this study had methodological limitations with inclusion of data-related precision. Additionally, the data were collected sixty years ago. Since then, substantial changes have occurred in driver demographics, driving behavior, roadway characteristics, transportation systems, cyber-physical infrastructure, and vehicular safety features and performance.

In addition, it is helpful to understand to what degree a motor vehicle exceeds the posted or statutory speed limit. Conventional crash databases use the term *speeding* or *exceeding safe speed limit*. These terms require further explanation. A motorcycle speeding in a residential zone can be contextualized differently when compared to a large truck speeding on a freeway. Also, every state has a basic speed law that entails drivers to operate vehicles safely by following the safe speed thresholds.

2.3.3.2. Studies on Operating Speed and Safety

There is a substantial amount of research on the crash-speed association. Hauer (2009) described the relationship between operating speed and safety as a "causal two-link chain" whose main components are human actions (in terms of speed choice), the evolution of speed, and safety outcome. Established human activities such as setting and enforcing speed limits affect the speed evolution and drivers' speed choice, which in turn affect roadway safety. The resulting safety implications of this speed evolution are then used in future decision-making for human activities. Pre-event probabilities determine the number and frequency of crashes, while the "at the time of event" probabilities determine the severity of the crashes (Haddon, 1972). Therefore, the driver's speed choice affects not only the potential severity based on the operating speed choice, but also the probability that a crash will occur.

Researchers used various types of operating speed features to estimate the relationship between the speed choice and its safety impact. The common speed-related measures are the individual speeds, average road section speeds, and speed variance (Aarts and Van Schagen, 2006). By examining the operating speeds of various vehicle types and conducting surveys or questionnaires with individual drivers, it is argued that crash liability is associated with the increase of individual driving speeds (Maycock et al., 1998; Haglund and Åberg, 2000).

Some earlier studies showed a direct association between operating speed and traffic crashes, especially for fatal and injury crashes (Hauer, 1971; Nilsson, 2004; Elvik et al., 2004). Abdel-Aty and Radwan (2000) examined the magnitude of speeding relative to the speed limit. The results showed that speeding is associated with male and young drivers in most cases. Taylor et al. (2000) conducted a study in the UK which demonstrated that the mean speed measure was negatively associated with crash counts at the aggregate level. For different homogenous groups, the results showed that in each set of road and traffic conditions, the crash count increases with the operating speed.

Pei et al. (2012) executed a study investigating the association between operating speed and crash risk. The authors observed several descriptive factors of the relationship, including road design, weather conditions, and temporal distribution. The study concluded that crash risk decreases as speed increases. In their study, Yu et al. (2013) examined crash data from one year on I-70 in Colorado. The authors used a Bayesian inference model that included real-time weather, traffic, and road geometry variables. The results demonstrated that weather condition has a considerable relationship with crash likelihood. As in earlier studies, this study showed that crash segments with lower speeds and higher occupancy at the upstream segment 5–10 minutes before the crash time had a high likelihood of crash occurrence. This connection between speed and crash risk could be a result of traffic bottleneck and could also be impacted by inclement weather conditions.

Gargoum and El-Basyouny (2016) investigated the connection between average speed and crash counts on urban two-lane streets; the authors also investigated the potential effects of confounding factors. The study showed that the standard deviation of speed was negatively related to crash frequencies, so lower speeds were related to higher crash frequencies. However, this association was only statistically significant at the 10 percent significance level (p-value of 0.088). The results from a segment-based study conducted by Imprialou et al. (2016) had similar findings. The study established that the relationship between operation speed and crash occurrence was negative regardless of crash severity.

Yu et al. (2018) performed a study using advanced traffic-sensing data of urban expressway systems in Shanghai, China, to examine aggregation approaches and their influence on relationship analyses. In the study, the authors first conducted crash count analyses with both segment and scenario specific methods. The authors then developed crash risk analyses at the individual crash level. Based on the segment-based crash count modeling, the study found a negative relationship between speed and crash likelihood. The results demonstrated that an increase in operating speeds during congested traffic is associated with a reduced likelihood of crash occurrences. In another study, Wang et al. (2018) used taxi-based high-frequency global positioning system (GPS) data from 234 one-way urban arterial segments in Shanghai to investigate the average speed and speed variation. In the study, the authors developed a hierarchical Poisson log-normal model with random effects to analyze these data. The findings indicated that a 1 percent increase in average speed was associated with a 0.70 percent increase in total crashes.

Using both probe and point detector data, Dutta and Fontaine (2019) examined speed-crash association on rural and urban freeway segments in Virginia. The study found that a lower average speed was associated with a higher crash frequency on rural freeways, while an increase in the standard deviation of average speed was associated with higher crash frequencies on urban freeways.

Banihashemi et al. (2019) performed a study and observed that an increase in the severity of crashes (a ratio of fatal and injury crashes to total crashes) was associated with an increase in speed differential. Xu et al. (2019) introduced a unique filtering process to discern taxi GPS data points on elevated expressways from the data points on the surface roads underneath the expressways. In this study, the researchers derived speed measures such as the standard deviation of the cross-sectional speed mean (SDCSM) and mean of the cross-sectional speed standard deviation (MCSSD) to identify the spatial and temporal speed variations, respectively. The researchers also developed hierarchical and nonhierarchical Poisson-gamma measurement error models to determine the crash occurrences of the expressways. The findings of the models showed that the hierarchical model had a better performance out of the two, and the SDCSM and MCSSD were both positively associated with crash likelihood.

Hutton et al. (2020) performed a study to investigate the speeds of individual drivers along 100 study segments. The study used data from the second Strategic Highway Research Program (SHRP-2) Naturalistic Driving Study data and the Roadway Information Database (RID) to estimate speed variations of the drivers on the same roadway segment as well as variation measure within individual trips. The study also gathered roadway and roadside characteristic data from Google Street View and aerial imagery to explore the relationship between roadway characteristics and speed. This study used different speed measures to perform the analysis (Hutton et al., 2020):

- **Mean space mean speed (SMS):** the average of all individual trip SMSs for each site.
- **85th percentile of trip SMSs:** the 85th percentile of all individual trip SMSs for each site.
- **Mean FFS:** the average of all individual trip FFSs for each site (this measure was similar to the mean SMS).
- **Standard deviation of SMS:** the standard deviation of the distribution of all individual trip SMSs for each site. This measures speed variability between trips.
- **Standard deviation of FFS:** the standard deviation of the distribution of individual trips' FFS for each site.
- **85th percentile of trip SMS minus 15th percentile of trip SMS:** the 85th percentile of the distribution of individual trip SMS minus the 15th percentile of the SMS distribution for each site.
- **Mean of difference between trip maximum and minimum speed:** the average of all individual trips' arithmetic difference between the trip's maximum spot speed and minimum spot speed. This provides a measure of speed variability within trips.

- **The standard deviation of difference between trip maximum and minimum speed:** the standard deviation of the distribution of an individual trip's arithmetic difference between the trip's maximum speed and minimum speed.
- **Difference between posted speed limit and 85th percentile of trip SMS:** the site's 85th percentile of trip SMS minus the site's posted speed limit.

The findings from this study suggest that several geometric variables are related to speed and crash occurrence and that a variance of speed between trips was related to a higher crash frequency, particularly for multivehicle crashes. Most of the other evaluated speed measures either statistically insignificant or negatively correlated. The study did not conclude that adding a speed CMF in the SPF development can significantly improve prediction accuracies for urban and suburban arterials.

Das et al. (2020) recently completed a U.S. Department of Transportation Safety Data Initiative rural speed safety pilot project by achieving three research products:

- Conflated databases for Ohio and Washington by using the Highway Safety Information System (HSIS) and the NPMRDS.
- Interactive data visualization tools to demonstrate the speed-crash association.
- SPFs at daily and annual levels with inclusion of operating speed measures.

Das et al. (2020) used the following speed measures to define the relationship between crash and operating speed:

- Average hourly speed (SpdAvg).
- Average hourly speed during non-peak and non-event (1 hour before and 1 hour after a crash occurrence) periods (SpdNPNE).
- SD of hourly operating speeds (SDHrSpd).
- SD of monthly operating speeds (SDMonSpd).
- Differences in the operating speeds during weekdays and weekends (SpdW_W).

The overall finding of this project was that inclusion of speed measures provide more insights on highway safety evaluation. This project also built an interactive decision support tool (https://ruralspeedsafety.shinyapps.io/rss_sdi/) to demonstrate the annual risk score using Washington and Ohio data that provide the expected total crashes on the roadway segments (see Figure 3). For detailed overview of multi-source data fusion, and decision support tool development, readers can consult these relevant studies (Das et al., 2021; Fitzpatrick et al., 2021a; Fitzpatrick et al., 2021b; McCourt et al., 2019; Das and White, 2020; Das and Geedipally, 2020).

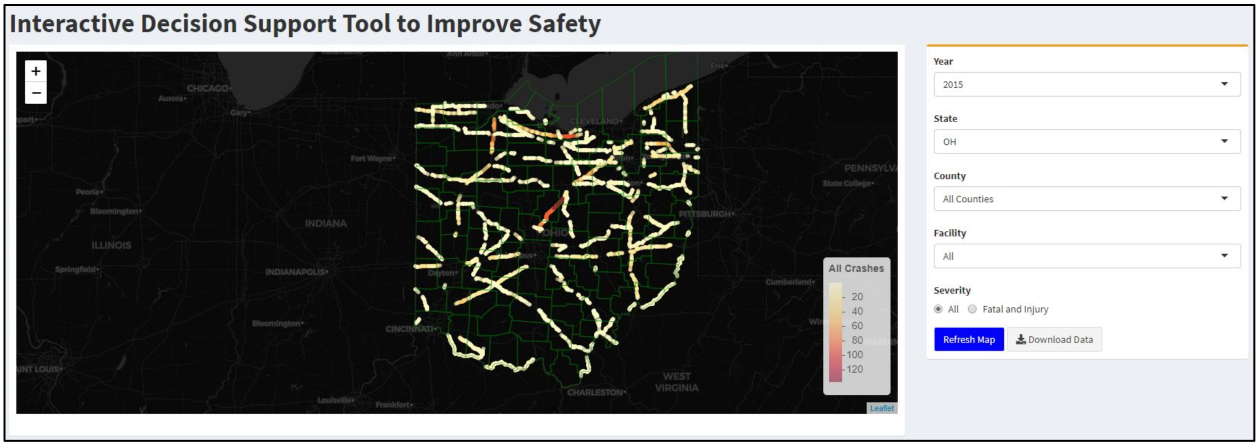

Figure 3. Interactive Decision Support Tool for Rural Speed Safety (Source: Das et al., 2020).

2.3.4 Summary

Table 2 lists the studies discussed in Section 2.3 and some key information.

Table 2. Studies on Operating Speed and Safety.

Study	Analysis Level	Roadway/ Location	Speed Measures	Operating Speed Data Source	Key Findings on Speed-Crash Relationship
Abdel-Aty and Radwan (2000)	Segment	Principal arterial, Florida	Speeding relative to posted speed limits	Crash data	Male and young drivers are mostly associated with exceeding posted speed limit
Taylor et al. (2000)	Segment	Different roadways, United Kingdom	Average speed	Road tubes	Speed is associated with high likelihood of crashes at disaggregate level
Pei et al. (2012)	Segment	Both urban and rural, Hong Kong	SD of average speed	Annual traffic census	With the increase of speed measures, crashes decrease
Yu et al. (2013)	Segment	Freeways, Colorado	Speed information prior to crash occurrence	Radars	Negative relationships between speed and crash occurrence
Gargoum and El-Basyouny (2016)	Segment	Urban two-lane roadways, Canada	SD of speed	Speed survey operations	SD of speed has negative likelihood of crash occurrences
Imprialou et al. (2016)	Traffic operation scenarios	Strategic road network, United Kingdom	Grouped average speed prior to crash occurrence	Inductive loop detectors	**Condition based approach:** speed measures are associated with crashes **Segment-based approach:** speed is negatively associated with crash likelihood
Yu et al. (2018)	Segment	Urban expressway, China	Average speed	Using algorithm	Speed is negatively associated with crash likelihood

Study	Analysis Level	Roadway/ Location	Speed Measures	Operating Speed Data Source	Key Findings on Speed-Crash Relationship
Wang et al. (2018)	Segment	234 one-way road segments from eight urban arterials, Shanghai, China	Mean speed	Taxi-based high-frequency GPS data	100% increase in average operating speed was associated with a 70% increase in total crashes
Banihashemi et al. (2019)	Segment	Urban freeway, Washington	Operating and posted speed differential	NPMRDS	Non-PDO/total crashes ratio is associated with the increase of speed differential
Dutta and Fontaine (2019)	Segment	Rural and urban freeways, Virginia	Speed variations	Physical sensors; INRIX®	Rural and urban roadways show different speed-crash association
Xu et al. (2019)	Segment	Expressways	Operating speed measures	Taxi GPS data points	Focused on model performance and precision
Hutton et al. (2020)	Segment	SHRP-2 states	Operating speed measures	SHRP-2 RID	Majority of the operational speed measures either had a negative correlation or were not found to be related to crash frequency
Das et al. (2020)	Segment	Rural facilities in Washington and Ohio	Operating speed measures	NPMRDS, HSIS	Speed-related operational information can be used to better understand safety outcomes

2.4 IMPACT OF WEATHER ON SAFETY

Crashes that occur during inclement weather (e.g., rain, sleet, snow, fog, snowy/slushy pavement, icy pavement, and wet pavement) are referred to as adverse- or inclement-weather-related crashes. Weather can include precipitation, wind gust, high temperature, and low visibility, all of which can influence pavement friction, driver performance, and vehicle condition. These effects can increase crash likelihood. Many studies examined driver behavior and crash occurrences under extreme weather condition. This section gives a brief summary of relevant studies.

2.4.1 Synthesis Studies

Strong et al. (2010) incorporated findings from several studies in the field of weather and surface transportation. The authors detailed the avenues of future research to address some of the existing research gaps. Considering the increasing availability of real-time traffic data and stimulation of the importance of proactive safety management, Theofilatos and Yannis (2014) reviewed the effect of weather-traffic impacts on roadway safety, identified the gaps, and discussed further research needs. Using a systematic approach, Lee et al. (2018) analyzed the extent to which the level of water depth and rainfall is responsible for traffic crashes, using Seoul City, Korea, as a case study. Over a 9-year period (2007–2015), the rainfall and traffic crash data for Seoul were analyzed using structural equation modeling (SEM) to determine the associations among variables by handling exogenous and endogenous variables simultaneously.

2.4.2 Crash-Weather Association

Using data from rural highways in Oregon, Monsere et al. (2008) performed an empirical study on weather related crashes. To rank the sites, both count and rate related computations were performed. Bijleveld et al. (2009) conducted a similar study to examine weather, crash and crash related injuries using data from the Netherlands. Using data from past two decades, Andrey (2010) examined the likelihood of crash related risks during inclement weather. The results suggest that the overall crash risk during snowfall shows no significant change over time.

El-Basyouny et al. (2014a) developed multivariate safety models by using 11 years of daily weather and crash data from Edmonton, Alberta, Canada. Results indicated that property damage only (PDO) crashes show the likelihood of 4.5 to 45 percent during adverse weather conditions. It is found that PDO crashes were affected more by inclement weather states than by severe (fatal and injury) crashes. Sudden weather changes show statistically significant and positive association with crash likelihood. El-Basyouny et al. (2014b) investigated how extreme inclement weather changes influence crash type. The results indicated that snowfall and temperature were statistically significant with the association results (crash likelihood increases as snowfall intensity increases; crash likelihood decreases with temperature increase) for all crash types. Using traffic and crash data from California freeways, Xu et al. (2018) examined the combined effects of real-time traffic conditions and environmental factors on crash risks. The model estimation results demonstrated that inclement weather conditions are associated with freeway crash likelihood. The interaction between upstream occupancy and light rainfall showed statistically significant association.

2.4.3 Spatial, Temporal, and Spatiotemporal

Jackson and Sharif (2016) examined the temporal and spatial distribution of rain-related fatal crashes from 1982 to 2011 in Texas. Rainfall was found statistically significant for few counties. Introducing time interdependencies in weather related crash count model, Brijs et al. (2008) introduced an integer autoregressive model. The daily level crash count models were developed using meteorological and traffic exposure data from the Netherlands. In their study, Bergel-Hayat et al. (2013) explored the relationship between short duration crash risk and weather conditions at an aggregate level. The primary results indicated associations between weather variables and monthly injury crash counts; however, impact of inclement weather measures vary by roadway facility types (rural roads, motorways, or urban roads). Malin et al. (2019) investigated the relative crash risk of various road weather conditions and combinations of weather conditions. Nearby road weather stations provided the hour-level weather and road condition data of each segment. The relative crash risks increased during poor road weather conditions; however, the highest relative crash risks occurred for icy rain and slippery and very slippery road conditions. When assessing the relative crash risk based on roadway facility types, the results demonstrated a higher risk in inclement weather and road conditions on motorways in comparison to two-lane and multiple-lane roads even though the overall risk on motorways was lower.

Using 8 years (2006–2013) of crash data from Pennsylvania, Kelarestaghi et al. (2017) analyzed the effect of inclement weather on crash injury types by performing a macroscopic analysis. The results found that factors such as inclement weather conditions and young drivers reduce crash severity, while the involvement of unbelted passengers, motorcycles, heavy trucks, and pedestrians are associated with the increase of the likelihood of having a severe injury crash. Wen et al. (2019) proposed the Bayesian spatiotemporal model to measure the association between crash count and contributing factors, such as weather conditions, traffic exposure, curve or slope, and their interactions. The results of the parameter estimation indicate that the interactions between precipitation and curve, between wind speed and slope, and between visibility and slope significantly correspond to the increase of freeway crash likelihood, while the interaction between precipitation and slope significantly corresponds to the reduction of freeway crash risk.

Several studies (Yu and Abdel-Aty, 2014; Yu et al., 2013; Yu et al., 2014; Yu et al., 2015) examined real-time crash prediction using weather, crash, and other key contributing factors. Theofilatos (2019) used negative binomial (NB) models and cusp catastrophe theory to predict crashes using real-time traffic and weather data collected in Athens, Greece, from urban motorways. The study found that rainfall intensity strongly influences crashes.

2.4.4 Pavement Condition

Using crash data on rural curved roadways in New York, Lamm et al. (1990) found that drivers failed to reduce their speeds sufficiently on curves during wet-pavement conditions. Mayora and Pina (2009) analyzed 10 years of crash data (including roadway geometry and skid resistance data) from two-lane rural roads on the Spanish National Road System to estimate a skid threshold. The results indicated that wet-pavement crash rates, averaging around 68 percent, were substantially reduced due to the improvement of pavement friction. The importance of sustaining acceptable levels of pavement friction was confirmed to improve safety. Buddhavarapu et al. (2013) investigated the association between crash severities on pavement surface condition and horizontal curves indices. This study used two Texas Department of Transportation (TxDOT) databases: 1) Pavement Management Information System data, 2) CRIS data. The results indicated a poor correlation between the skid number and crash injury severity on two-lane horizontal curves, and the Distress Index and International Roughness Index (IRI) had a statistically significant effect on crash severity types. Using New Jersey crash and pavement condition data, Najafi et al. (2015) developed regression models to examine how friction impacts the rate of surface condition based (wet vs. dry) traffic crashes for various urban facility types. The results concluded that friction is related to the rate of wet-condition vehicle crashes and also affects the rate of dry-condition-related traffic crashes.

2.4.5 Specific Issue and Condition

Using Colorado freeway data, Ahmed et al. (2012) examined the effect of the interaction between real-time weather, traffic data, and roadway geometric features on the occurrence of crashes. The results showed that the likelihood of a crash could double during the snowy weather condition due to the interaction between the steep grades of the mountainous freeways and pavement conditions. Saha et al. (2015) examined the relationship between adverse geometric roadway characteristics and weather conditions and their impact on crash occurrence on rural freeway corridors in mountainous terrain. This study concluded that the relationship between horizontal curves and grades with weather variables had a significant impact on the likelihood of traffic crashes.

Das et al. (2017) comprehended the impact of visibility on safety. The findings identified key associating factors such as curved roadways, younger and older drivers, high posted speed limit, low-friction roadways, undivided roadways, and no lighting at dark. Commercial motor vehicle (CMV) operations and driver safety are influenced by adverse weather conditions such as heavy precipitation, high winds, slit, low visibility, and other factors. Rossetti and Johnsen (2011) addressed this critical issue, which was less studied.

As part of a Wisconsin road weather safety initiative, Jung et al. (2012) applied a sequential logistic regression approach to examine safety impact on high-speed roadways during rainfall. The backward sequential logistic regression model produced the most preferable results, in which wind speed, vehicle types, and at-fault driver-related actions were found as statistically significant variables. To determine strategic locations for the deployment of road weather information system (RWIS) stations in a large regional transportation network, Jin et al. (2014) proposed a spatial optimization method. Weather-related crash data were transcribed into to a safety concern index.

2.4.6 Summary

Table 3 lists the studies discussed Section 2.4 and some key information.

Table 3. Studies on Weather and Safety.

Studies	Modeling Approach	Key Findings
Lamm et al. (1990)	Regression model	• Drivers do not control speed adequately on wet paved curve sections
Brijs et al. (2008)	Integer autoregressive model	• A serial temporal correlation can account for bias reduction
Monsere et al. (2008)	Empirical Bayes	• Performed an empirical analysis of screening and ranking for weather-related crashes on rural roadways in Oregon
Bijleveld et al. (2009)	Aggregate-level analysis	• Performed an analysis of the aggregate effect of weather conditions on crashes
Mayora and Pina (2009)	Cross-sectional analysis	• Maintenance of adequate levels of pavement friction is associated with safety improvement

Studies	Modeling Approach	Key Findings
Andrey (2010)	Matched-pair framework	• Analysis on two decades of wet-weather-related traveling patterns • Impact of snowfall on crash likelihood is not statistically significant
Strong et al. (2010)	Severity index	• Synthesized the findings from some of the major efforts in weather-crash association
Rossetti and Johnsen (2011)	Exploratory data analysis	• Examined the safety impact of weather on CMVs
Ahmed et al. (2012)	Bayesian logistic regression technique	• Investigated the interaction between roadway geometry and real-time weather and traffic data on traffic crashes on a mountainous freeway
Jung et al. (2012)	Sequential logistic regression	• Examined multivehicle-involved injury types on high-speed roadways during adverse weather • Used a sequential logistic regression approach to perform analysis
Buddhavarapu et al. (2013)	Ordered probit response model	• Skid number shows poor correlation with crash injury severity on two-lane horizontal curves • Distress Index and IRU show statistically significant effect on crash injury severity
Bergel-Hayat et al. (2013)	Time series model	• Examined the temporal level association between weather conditions and crash risk at an aggregate level
El-Basyouny et al. (2014a)	Full Bayesian context via a Markov chain Monte Carlo simulation	• Examined the aggregated effect of different weather conditions on crash occurrence
El-Basyouny et al. (2014b)	Multivariate Poisson lognormal	• Examined the impact of weather elements and sudden extreme snow or rain weather changes on crash type
Jin et al. (2014)	Spatial optimization method	• Examined the right deployment of RWISs
Theofilatos and Yannis (2014)	Linear regression	• Reviewed the effect of traffic and weather characteristics on road safety
Najafi et al. (2015)	Regression analysis	• Friction is not only associated with the rate of wet-condition vehicle crash likelihood but also influences the rate of dry-condition vehicle crashes
Saha et al. (2015)	NB	• Interaction between grades and horizontal curves with weather variables had a significant impact on crash occurrence
Jackson and Sharif (2016)	Spatial analysis	• Rain is found as a contributor factor to traffic crash likelihood in a few counties in Texas
Das et al. (2017)	Parametric model (ordinal logistic regression), non-parametric analysis (multiple correspondence analysis), and topic model development	• Examined the implications of adverse weather on safety from a perspective of visibility and other key issues
Kelarestaghi et al. (2017)	Spearman correlation test	• Macroscopic analysis showed that factors such as inclement weather conditions and young drivers reduce the severity of a crash • Motorcyclists, pedestrians, unbelted passengers, and heavy trucks are associated with the likelihood of having a crash with severe injury

Studies	Modeling Approach	Key Findings
Lee et al. (2018)	SEM	• Conducted a systematic approach to analyze the weather crash relation using data from Seoul, Korea
Xu et al. (2018)	Logistic regression models	• Presence of environmental infomation improves both the goodness of fit and prediction performance of the crash risk prediction model
Malin et al. (2019)	Concept of random point process	• Examined the relative crash risk of different scenarios of road weather conditions and combinations of conditions
Theofilatos (2019)	Cusp catastrophe theory and NB model	• Precipitation is linearly associated with crash likelihood
Wen et al. (2019)	Bayesian spatiotemporal model	• Measured the relationship between crash and factors such as curve and slope, traffic composition, weather conditions, and their interactions
Yu et al. (2013); Yu and Abdel-Aty (2014); Yu et al. (2014); Yu et al. (2015); Yu et al. (2018)	Different modeling techniques	• Examined the impact of real-time traffic and weather on crashes

Roadway geometry, traffic volume, and human factors are the key contributors of traffic crashes. Three review papers (Lord and Mannering, 2010; Savolainen et al., 2011; Mannering and Bhat, 2014) provide an in-depth understanding of the state-of-the-art methodological frameworks of crash data analysis (both the count data model and crash severity model). Interested readers can consult these papers for a comprehensive overview on crash data modeling techniques.

2.5 MODELING FRAMEWORKS

Researchers have used a wide variety of speed measures and different modeling techniques to explore the association between operating speed, posted speed, roadway geometry, traffic exposures, and safety. This section provides a brief overview of state-of-the-art modeling techniques.

2.5.1 Power Model and Meta-analysis

Researchers used different functional forms of operational speed and posted speed measures. Power and exponential models gained a wide audience. Nilsson (2004) studied the safety impact of a reduction of the speed limit on Swedish rural roads from 68 to 56 mph and vice versa. The comparison group was a group of 56 mph without speed limit changes. This study introduced the power model as a group of six different power functions (with different exponent measures based on the injury severity type) that estimates the expected number of crashes by severity type after a certain change in mean speeds on a road network. He assumed that crash frequency is always proportional to operating speed. Equations 1 through 3 illustrate this assumption.

$$Total\ crashes_{after} = Total\ crashes_{before}\left(\frac{Operating\ Speed_{after}}{Operating\ Speed_{before}}\right)^2 \tag{1}$$

$$Fatal\ and\ serious\ crashes_{after} = Fatal\ and\ serious\ crashes_{before}$$

$$\left(\frac{Operating\ Speed_{after}}{Operating\ Speed_{before}}\right)^3 \quad (2)$$

$$Fatal\ crashes_{after} = Fatal\ crashes_{before}\left(\frac{Operating\ Speed_{after}}{Operating\ Speed_{before}}\right)^4 \quad (3)$$

Elvik et al. (2004) and Elvik (2009) performed a meta-analysis of existing literature to develop a power model that would measure the changes in the safety measures as a result of the changes in average speed (see Table 4 for the exponent measures). Equation 4 shows this measure.

$$Crashes_{after} = Crashes_{before}\left(\frac{Operating\ Speed_{after}}{Operating\ Speed_{before}}\right)^p \quad (4)$$

Where:

p = the exponent to which the function is raised.

Table 4. Exponents Used in Elvik (2009).

Crash or Injury Severity	Rural Road/Freeways		Urban/Residential Roads		All Roads	
	Best Estimate	95% Confidence Interval	Best Estimate	95% Confidence Interval	Best Estimate	95% Confidence Interval
Fatal crashes	4.1	(2.9, 5.3)	2.6	(0.3, 4.9)	3.5	(2.4, 4.60)
Fatalities	4.6	(4.0, 5.2)	3.0	(–0.5, 6.5)	4.3	(3.7, 4.9)
Serious injury crashes	2.6	(–2.7, 7.9)	1.5	(0.9, 2.1)	2.0	(1.4, 2.6)
Seriously injured road user	3.5	(0.5, 5.5)	2.0	(0.8, 3.2)	3.0	(2.0, 4.0)
Slight injury crashes	1.1	(0.0, 2.2)	1.0	(0.6, 1.4)	1.0	(0.7, 1.3)
Slightly injured road user	1.4	(0.5, 2.3)	1.1	(0.9. 1.3)	1.3	(1.1, 1.5)
Injury crashes—all	1.6	(0.9, 2.3)	1.2	(0.7, 1.7)	1.5	(1.2, 1.8)
Injured road user—all	2.2	(1.8, 2.6)	1.4	(0.4, 2.4)	2.0	(1.6, 2.4)
PDO crashes	1.5	(0.1, 2.9)	0.8	(0.1, 1.5)	1.0	(0.5, 1.5)

Hauer and Bonneson (2006) explained that "the use of power functions is straightforward and transferable, but the exponents provided (as in Elvik et al., 2004) are independent of the baseline speed and could lead to inaccurate estimations." The authors proved the dependence of crash count on the baseline speed and examined whether there is a need to explore on the differential between speed measures. Hauer and Bonneson (2006) and Hauer (2009) also developed new exponential prediction models that incorporated maneuver time and distance for collision avoidance. The developed CMF was expressed as in Equation 5.

$$CMF\ (for\ speed\ change\ from\ v_o to\ v_1) = e^{\alpha[v_o - v_1 + \frac{\beta}{2}(v_0^2 - v_2^2)]} \quad (5)$$

Elvik (2013) conducted an updated analysis of the power model to assess speed-crash association while accounting for both speed change and initial speed (Equation 6). He found that the exponential functions can provide best fit models. Differences between the power model and the exponential model were not found significantly different.

$$Total\ crashes_{after} = Total\ crashes_{before}(exp)^{\text{Operating Speed}_{\text{after}} - \text{Operating Speed}_{\text{before}}} \quad (6)$$

2.5.2 Other Modeling Frameworks

2.5.2.1. Develop Speed-Based Crash Modification Factors

The predictive method in Part C of the HSM is an 18-step iterative procedure to estimate the average expected crash frequency at a segment or an intersection (AASHTO, 2010). The method uses three major components to predict the average expected crash frequency at a site:

- The base model estimates the predicted crashes using an SPF.
- The CMFs are used to adjust the estimate for additional site-specific conditions that differ from the base conditions.
- A calibration factor is used in improving model accuracy for a specific state or local area.

These components are combined in the general form shown in Equation 7.

$$N_{predicted} = N_{spf} \times \left(CMF_{1x} \times CMF_{2x} CMF_{1x} \ldots \times CMF_{yz}\right) \times C_x \quad (7)$$

Where:

$N_{predicted}$ = the predicted average yearly crash frequency for site x.
N_{spf} = the predicted average crash yearly frequency determined for base conditions of the SPF developed for site x.
CMF_{nx} = CMFs specific to the SPF for site type x.
C_x = the calibration factor to adjust the SPF for local conditions for site x.

Appendix 3E ("Speed and Safety") in the first edition of the HSM provides some contexts of speed measures and their impact on highway safety. To use operating speed measures in safety evaluation, the potential approach would be the identification of the locations with a range of the of mean speeds (e.g., 55 to 60 mph) and the development of a prediction model for these sites. The potential crash prediction model can serve as the base-condition SPF, where the base condition would be a range of operational speed measures.

One example of this approach is the CMFs developed by Hauer and Bonneson (2006) that involved conducting a meta-analysis of results from multiple studies where a treatment resulted in changes in average operating speed. These values were included in Table 3E-2 of the first edition of the HSM, as shown here in Table 5 and Table 6 (AASHTO, 2010).

Table 5. CMFs for Injury Crashes Based on Changes in Average Operating Speed.

Injury Crashes	$\overline{v_0}$ (mph)					
$\overline{v_1} - \overline{v_0}$ (mph)	30	40	50	60	70	80
–5	0.57	0.66	0.71	0.75	0.78	0.81
–4	0.64	0.72	0.77	0.80	0.83	0.85
–3	0.73	0.79	0.83	0.85	0.87	0.88
–2	0.81	0.86	0.88	0.90	0.91	0.92
–1	0.90	0.93	0.94	0.95	0.96	0.96
0	1.00	1.00	1.00	1.00	1.00	1.00
1	1.10	1.07	1.06	1.05	1.04	1.04
2	1.20	1.15	1.12	1.10	1.09	1.08
3	1.31	1.22	1.18	1.15	1.13	1.12
4	1.43	1.30	1.24	1.20	1.18	1.16
5	1.54	1.38	1.30	1.26	1.22	1.20

Note: Although data used to develop these CMFs are international, the results apply to North American conditions.
Source: AASHTO (2010), Table 3E-2 p. 3-57.

Table 6. CMFs for Fatal Crashes Based on Changes in Average Operating Speed.

Fatal Crashes	$\overline{v_0}$ (mph)					
$\overline{v_1} - \overline{v_0}$ (mph)	30	40	50	60	70	80
–5	0.22	0.36	0.48	0.58	0.67	0.75
–4	0.36	0.48	0.58	0.66	0.73	0.8
–3	0.51	0.61	0.68	0.74	0.8	0.85
–2	0.66	0.73	0.79	0.83	0.86	0.9
–1	0.83	0.86	0.89	0.91	0.93	0.95
0	1	1	1	1	1	1
1	1.18	1.14	1.11	1.09	1.07	1.05
2	1.38	1.28	1.22	1.18	1.14	1.1
3	1.59	1.43	1.34	1.27	1.21	1.16
4	1.81	1.59	1.46	1.36	1.28	1.21
5	2.04	1.75	1.58	1.46	1.36	1.27

Note: Although data used to develop these CMFs are international, the results apply to North American conditions.
Source: AASHTO (2010), Table 3E-2 p. 3-57.

2.5.2.2. Consideration of Speed Measures as a Covariate (Regression Model)

Another approach for incorporating speed in the prediction methodology is to estimate a fully specified model with speed measure as one of the covariates. However, this approach is not extensively different from the modeling approach discussed in the earlier section. These models can be used as estimation models where the coefficient for speed can be used as the measure of the speed-safety relationship. Most of the studies discussed in Section 2.3.3.2. ("Studies on Operating Speed and Safety") used this approach.

2.5.2.3. Consideration of Speed Measures as a Covariate (Rules-Based Model)

Regression models usually examine the mean effects of the influential factors and ignore sub-group effects in the data. As a result, the interpretations are mostly associated with the overall segment-level analysis without giving any consideration of the scenarios for different sub-groups. Rules-based modeling (i.e., decision tree and Cubist) can take sub-group effects into account and can identify the relation between influential factors without imposing any prior assumptions on the sub-group or the overall data.

2.5.2.4. Consideration of Speed Measures as Mediator or Latent Factor

An intuitive way to illustrate the interrelationship between speed, roadway characteristics, and safety is to express it in the form of multiple equations or performance functions. One way to consider both these equations in one modeling framework is using an SEM or path analysis. Both methods examine relationships between and among one or more dependent variables and two or more predictor or independent variables. The main difference between path analysis and SEM is that path analysis considers all variables to be measured without error and SEM uses latent variables to account for inaccuracy.

Gargoum and El-Basyouny (2016) explored the connotation between operating speed and safety using a path analysis model by considering speed as a mediator variable. The parameters in the crash model were estimated using the maximum likelihood estimation method, and those in the speed model were obtained from the ordinary least square method. The authors found that average speed was significantly and positively associated with traffic crash likelihood.

Cheng et al. (2013) used simultaneous equation models to develop the speed-crash relationship. At first, single-equation models were developed involving crash counts and speed limits. Then, a simultaneous equation modeling framework was developed for these same variables. The results indicated that the speed-crash relationship was not statistically significant in the locations considered in this study; however, the presence of endogenous variables was validated. It is suggested that in the future endogeneity needs to be considered in transportation models involving traffic crash histories and posted speed limits.

2.6 SUMMARY

This literature review presents some of the key findings on the association between crashes, speed, and weather. Results are not entirely consistent between studies. The following highlights indicate the inherent complexity of the speed-safety relationship and some potential methodological limitations:

- An increased posted speed limit is associated with average speeds and increases the likelihood of crash and injury severity.
- A direct relationship between operating speed and crash is still not consistent.
- In most cases, speed variability is associated with a high likelihood of crash occurrences.
- Inclement weather is usually associated with the increase of crashes. However, some studies show a low number of injury crashes due to the occurrence of lower operating speed during the inclement weather.
- In most studies, inclusion of speed in the modeling framework increases the model performance.

CHAPTER 3:
DATA PREPARATION

3.1 INTRODUCTION

This chapter provides a brief overview of the data sets and the data conflation framework.

3.2 DATA SOURCES

TTI identified the following four major data sources to perform the analysis:

- Traffic speed data from the NPMRDS.
- Traffic volume data from the TMAS.
- Traffic crash data from the CRIS.
- Roadway inventory data from the Road-Highway Inventory Network Offload (RHiNO).
- Weather data from NOAA.

3.2.1 Speed Data: NPMRDS

Three readily accessible options exist for capturing speed information on Texas roadways:

- The NPMRDS.[2]
- The recently released Performance Network[3] from FHWA.
- The INRIX XD network.[4]

TTI currently has a contract with INRIX to obtain travel time data on its XD network, which is conflated onto the RHiNO network for work with TTI's *100 Most Congested Roadways in Texas* report[5] and is conducted for TxDOT annually. The NPMRDS, procured by FHWA, is free to state DOTs and metropolitan planning organizations for research. The NPMRDS contains a road network shapefile. This GIS shapefile contains static roadway information to relate the travel time information to each traffic message channel (TMC) segment. The TMC file contains TMC segment geometry information. The data collection period for the NPMRDS is 2017 through 2020.

3.2.2 Volume Data: TMAS

State highway and transportation agencies maintain the TMAS, which is a system of traffic count stations that monitor traffic volume on different roadways by vehicle class, and vehicle weight

---

[2] The NPMRDS database is available at https://npmrds.ritis.org/analytics/help/#npmrds.
[3] The Performance Network database is available at https://www.fhwa.dot.gov/policyinformation/tables/performancenetwork/.
[4] More information about the XD network is available at https://inrix.com/press-releases/2664/.
[5] The report is available at https://mobility.tamu.edu/texas-most-congested-roadways/.

information. Note that that the TMAS data do not provide systemwide short-duration traffic counts for all roadway networks. Additionally, it is important to know that short-duration data may not represent the yearly traffic volume measures.

3.2.3 Crash Data: CRIS (2015–2020)

TTI collected 6 years (2015–2020) of crash data from TxDOT's CRIS. CRIS data elements are divided into three major groups:

- Crash event characteristics.
- Primary person characteristics.
- Vehicle (unit) characteristics.

3.2.4 Roadway Inventory Data: RHiNO

TTI acquired roadway inventory data from two different sources:

- 2018 RHiNO.[6]
- 2020 TxDOT Roadway Inventory.[7]

An examination of these data sets showed that they are the same. Since RHiNO provides a detailed data dictionary and additional supporting GIS files, TTI used the 2018 RHiNO data as the main layer on which the other data layers were conflated.

3.2.5 Weather Data: NOAA

The Automated Surface Observing System (ASOS) serves as a primary climatological observing network in the United States. TTI collected 5-minute data from NOAA. The data collection was conducted for 6 years (2015–2020) for all Texas ASOS sites.

3.3 DATA CONFLATION STEPS

This section provides a brief overview of the data conflation steps, grouped into three processes. Figure 4 shows the overall framework of the data conflation work performed in this study.

[6] More information on road inventory data is available at https://www.txdot.gov/inside-txdot/division/transportation-planning/roadway-inventory.html.
[7] The TxDOT Roadway Inventory is available at http://gis-txdot.opendata.arcgis.com/datasets/txdot-roadway-inventory.

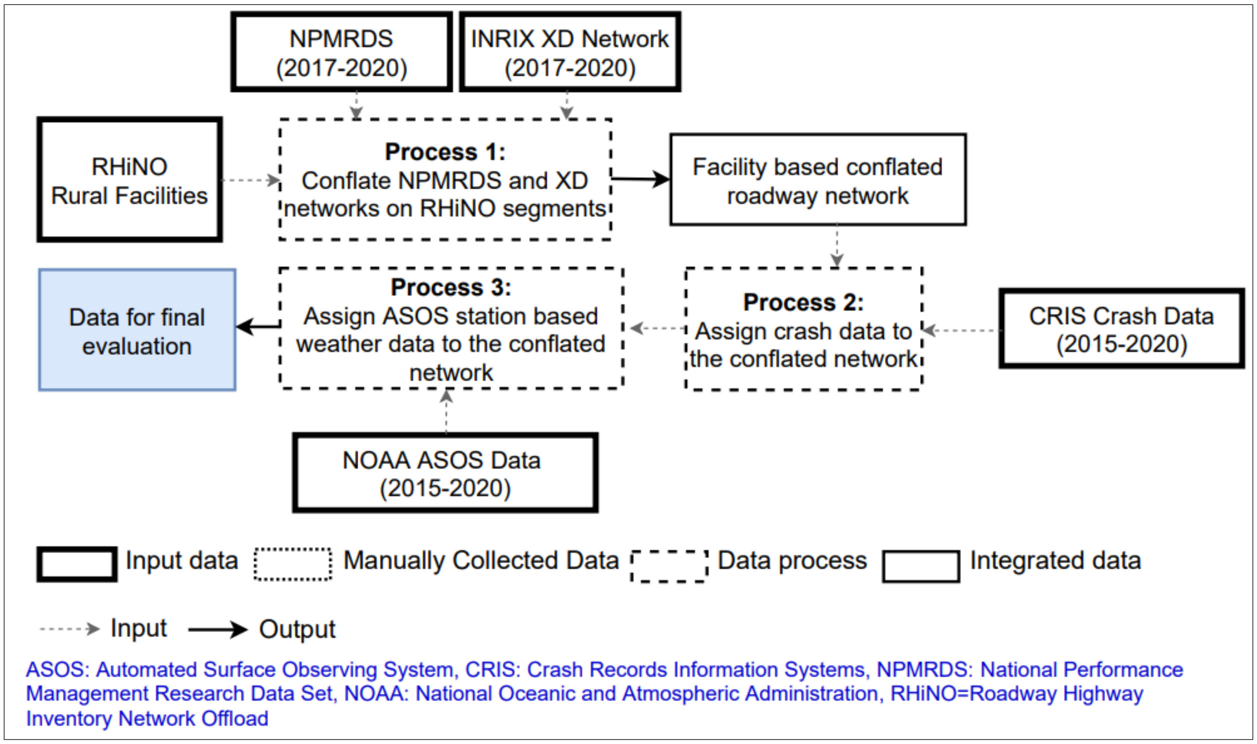

Figure 4. Flowchart of the Data Conflation Work.

3.3.1 Process 1: Conflation of RHiNO and NPMRDS/INRIX XD Networks

In this process, TTI integrated two linear systems (NPMRDS and RHiNO line files) to enable RHiNO segment-based analysis with speed measures. This study uses two formats of RHiNO files:

- Roadway linework without roadway characteristics (TxDOT_Roadway_Linework).
- Roadway linework with roadway characteristics (TxDOT_Roadway_Linework_wAssets).

In the former file, each route is a continuous polyline in the GIS database, and the file includes only the basic information for the route (e.g., route name and begin and end mileposts). In the latter file, routes are split into various numbers of homogeneous segments with roadway assets or features (e.g., traffic volume, posted speed limit, number of lanes, shoulder, and functional class).

The conflation work considered the 2018 NPMRDS file for Texas and the 2018 TxDOT RHiNO file. TTI used two software packages (ArcGIS and R) to conflate these databases. The following steps were taken in this process.

3.3.1.1. Step 1: Refine RHiNO Files

In this step, TTI removed non-centerline entries (i.e., frontage roadways, grade-separated connectors, and other nonessential roadways) in both RHiNO files. Definitions of centerline and non-centerline roadways can be found in the TxDOT RHiNO specification document.

3.3.1.2. Step 2: Divide the NPMRDS File by Direction

In this step, TTI divided the NPMRDS file into two files: positive and negative. The direction of the NPMRDS segments is determined by the TMC name: "+" or "P" indicates positive, and "–" or "N" indicates negative (Figure 5).

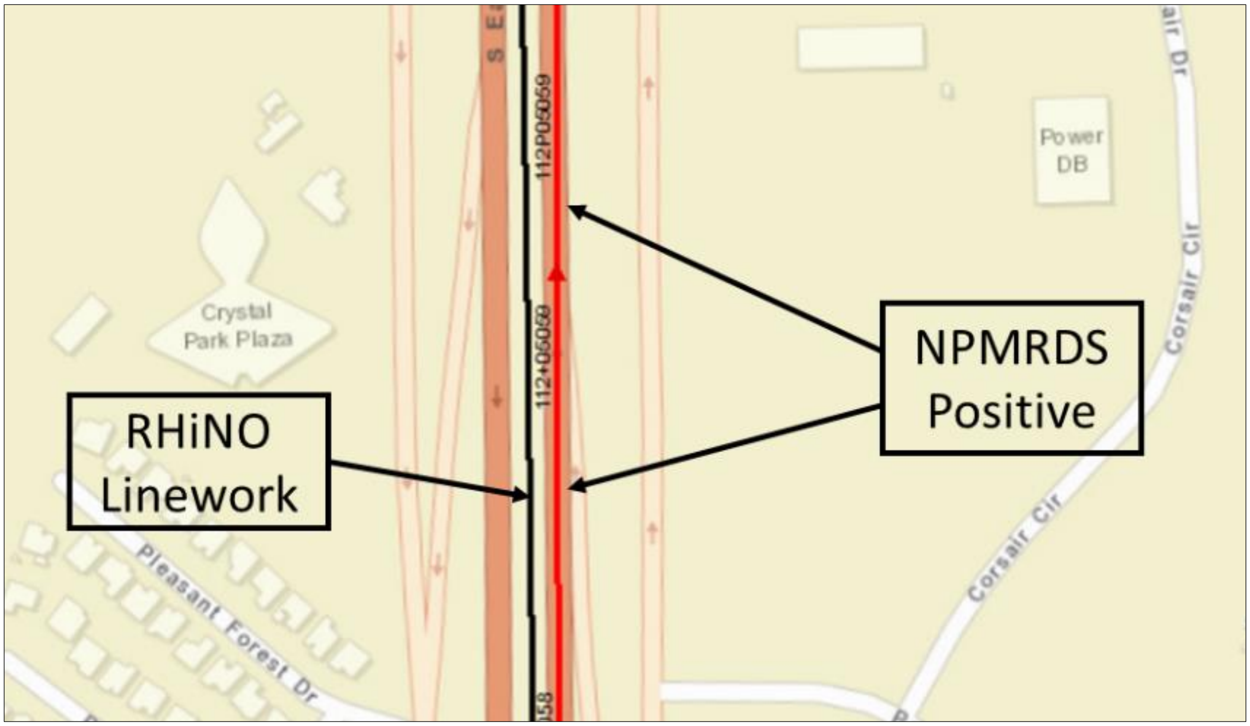

Figure 5. Step 1 and Step 2 (RHiNO and NPMRDS Linework).

3.3.1.3. Step 3: Locate NPMRDS Segments along with RHiNO Routes

In this step, TTI used one direction of NPMRDS files as the input feature and RHiNO linework without roadway characteristics as the input route feature, and then located the NPMRDS segments on the RHiNO routes. The event table generated from the locating process was exported as a csv file. Each direction of NPMRDS was located separately (see Figure 6).

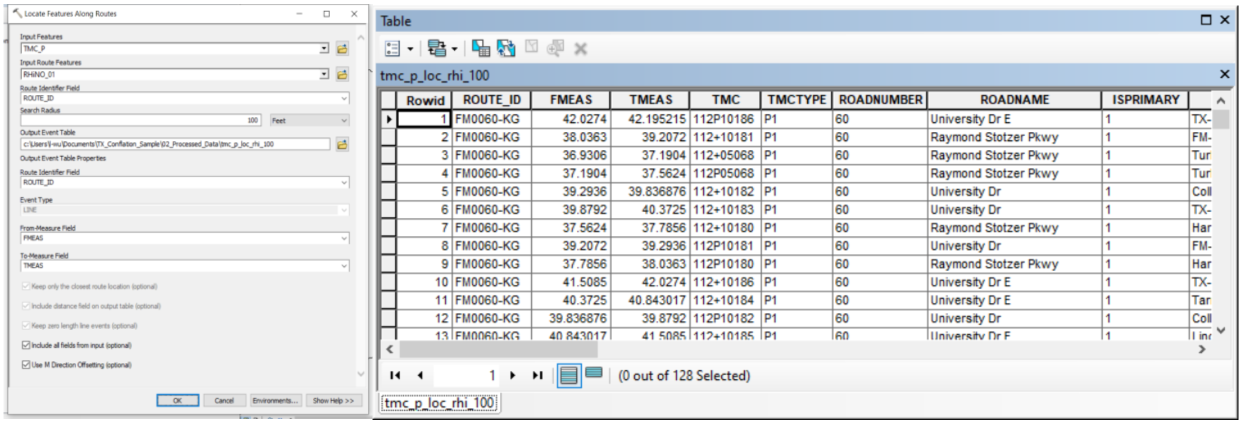

Figure 6. Step 3 (Locating NPMRDS Segments on RHiNO Routes).

3.3.1.4. Step 4: Refine the Event Table

In Step 3, the two files (i.e., NPMRDS and RHiNO) were located based on spatial relationships. A few segments were mismatched in the process. Step 4 eliminated the mismatched events based on the roadway names. In the refined event table, each NPMRDS segment had a route name and beginning and end mileposts relative to the RHiNO linework (see Table 7).

Table 7. Step 4 (Refined Event Table).

NPMRDS Segment				RHiNO Linework		
TMC	**TMC TYPE**	**ROAD NUMBER**	**ROAD NAME**	**ROUTE_ID**	**Beg MP**	**End MP**
112P10186	P1	60	University Dr E	FM0060-KG	42.027	42.195
112+10181	P1	60	Raymond Pkwy	FM0060-KG	38.036	39.207
112+05068	P1	60	Raymond Pkwy	FM0060-KG	36.931	37.190
112P05068	P1	60	Raymond Pkwy	FM0060-KG	37.190	37.562
112+10182	P1	60	University Dr	FM0060-KG	39.294	39.837
112+10183	P1	60	University Dr	FM0060-KG	39.879	40.372
112+10180	P1	60	Raymond Pkwy	FM0060-KG	37.562	37.786
112P10181	P1	60	University Dr	FM0060-KG	39.207	39.294
112P10180	P1	60	Raymond Pkwy	FM0060-KG	37.786	38.036
112+10186	P1	60	University Dr E	FM0060-KG	41.508	42.027
112+10184	P1	60	University Dr E	FM0060-KG	40.372	40.843

3.3.1.5. Step 5: Create the Final Table

TTI refined the event table and RHiNO linework with roadway characteristics. In this step, the association between NPMRDS segments and RHiNO segments with roadway characteristics was created based on the route name and mileposts. For each RHiNO segment, the data from NPMRDS segments were collected. The information included the TMC name and the effective length ratio of the TMC matching with the RHiNO segment (see Table 8).

Table 8. Step 5 (RHiNO Segments with NPMRDS Information).

RIA_RTE_ID	Unique ID	Beg MP	End MP	TMC	Effective Ratio
IH0010-KG	8	61.621	61.774	115+04763	0.024
IH0010-KG	8	61.621	61.774	115+04764	0.983
IH0010-KG	9	510.224	510.72	112+05215	0.044
FM3363-KG	10	0.444	0.487	NA	NA
IH0010-KG	29	316.695	317.161	112+05415	0.097
FM3181-KG	30	0	0.993	NA	NA
IH0010-KG	31	878.531	878.558	112+05594	0.016

Similar steps were taken for the conflation of the RHiNO and XD network. The current data conflation work is limited to the conflation of these roadway networks. Appropriate speed measures were developed in Process 3. TTI also assigned TMAS locations to the conflated RHiNO network. However, due to the missing values issues, TMAS values were not used in final analysis.

3.3.2 Process 2: Assignment of Crashes to the Conflated Network

TTI collected 6 years (2015–2020) of CRIS data. The crash data were assigned to the conflated RHiNO segments using the "near" function of ArcGIS. The data assignment required several data quality checks. These checks included removing intersection crashes from segment crashes and matching the crashes with the name of the roadway where they occurred in both the RHiNO and CRIS databases. Three different severity types were considered for the conflation work:

- All crashes.
- Fatal and injury crashes.
- Fatal and severe injury crashes.

3.3.3 Process 3: Assignment of ASOS Weather Station Data to the Conflated Network

TTI collected 5-minute interval weather data (i.e., temperature, wind speed, precipitation, and visibility) from ASOS for 5 years (2015–2019). For each of the weather stations, a 10-mile buffer was developed to assign conflated roadway segments to the weather station buffer.

CHAPTER 4: MODEL DEVELOPMENT

4.1 DATA PREPARATION

The following four subsections discuss the steps considered for the data preparation. The first subsection discusses the data sources. The second and third subsections provide discussions on data conflation and speed measure calculation, respectively. The fourth subsection presents database preparation on a daily level.

4.1.1 Data Sources

TTI identified the following five major data sources to perform the analysis:

- Traffic speed data from the NPMRDS.
- Traffic volume data from the TMAS.
- Traffic crash data from the CRIS.
- Roadway inventory data from the RHiNO.
- NOAA weather data from ASOS stations.

4.1.1.1. Speed Data: NPMRDS

TTI collected 4 years (2017–2020) of 5-minute interval operating speed data from the NPMRDS.

4.1.1.2. Volume Data: TMAS

After exploring the TMAS data, TTI decided to use RHiNO traffic volume data for the final analysis.

4.1.1.3. Crash Data: CRIS (2015–2020)

TTI collected 6 years (2015–2020) of crash data from TxDOT's CRIS.

4.1.1.4. Roadway Inventory Data: RHiNO

TTI used the 2018 RHiNO as the main layer on which the other data layers were conflated.

4.1.1.5. Weather Data: NOAA

TTI collected 5-minute data from NOAA. The data collection was conducted for 6 years (2015–2020) for all Texas ASOS sites.

The previous chapter provides details on the data conflation steps.

4.1.2 Annual-Level Databases

After the data conflation process was finished, TTI calculated several speed measurement variables based on the 5-minute interval speed data collected from the NPMRDS of each TMC section. For annual data, Table 9 lists the speed measurement variables. For all these speed measurement variables, the speed data were aggregated based on four years of data from 2017 to 2020.

Table 9. Speed Measures Used for Annual-Level Analysis.

Attribute Name	Definition
SpdAve	Average of all 5-minute intervals using all vehicle speeds for 4 years (2017–2020)
SpdStd	SD of all 5-minute intervals using all vehicle speeds for 4 years (2017–2020)
Spd85	85th percentile speed of all 5-minute intervals using all vehicle speeds for 4 years (2017–2020)
RefSpd	Average reference speed of all 5-minute intervals using all vehicle reference speeds for 4 years (2017–2020)
SpdAveDay	Average of 5-minute intervals using all vehicle speeds occurring in the daytime (5 < hour < 18) for 4 years (2017–2020)
SpdStdDay	SD of 5-minute intervals using all vehicle speeds occurring in the daytime (5 < hour < 18) for 4 years (2017–2020)
SpdAveNight	Average of 5-minute interval all vehicle speeds for daytime (17 < hour < 24 and –1 < hour < 6) for 4 years (2017–2020)
SpdStdNight	SD of 5-minute interval all vehicle speeds for daytime (17 < hour < 24 and –1 < hour < 6) for 4 years (2017–2020)
SpdAveMTWT	Average of 5-minute interval all vehicle speeds for Mon, Tues, Wed, and Thurs for 4 years (2017–2020)
SpdStdMTWT	SD of 5-minute interval all vehicle speeds for Mon, Tues, Wed, and Thurs for 4 years (2017–2020)
SpdAveFSS	Average of 5-minute interval all vehicle speeds for Fri, Sat, and Sun for 4 years (2017–2020)
SpdStdFSS	SD of 5-minute interval all vehicle speeds for Fri, Sat, and Sun for 4 years (2017–2020)
SpdFFAve	Average of all 5-minute interval all vehicle speeds for 4 years (2017–2020) when 5-min speed is > RefSpd
SpdFF85	Average of all 5-minute interval all vehicle speeds for 4 years (2017–2020) when 5-min speed is > RefSpd
Reference Speed	Provided value by the NPMRDS. An approximation of FFS for the segment. This value is calculated using the 95th percentile of the speeds between 10 p.m. and 5 a.m.

In the previous step, when NPMRDS segments were matched to RHiNO segments, most of the RHiNO segments had at least two TMC sections because the NPMRDS data set was separated into two directions (positive and negative). Therefore, these segments ended up with at least two sets of speed measurement variables, while only one set of speed measurement variables is needed. To address this problem, TTI used the following method to aggregate speed measurement variables.

Equation 8 was applied to the following speed measurement variables: SpdStd, SpdStdDay, SpdStdNight, SpdStdMTWT, and SpdStdFSS.

$$A = \sqrt{\sum_i^n (w_i s_i)^2} \tag{8}$$

Where:

A = the aggregated speed measurement variable.
n = the number of unique TMC names of a RHiNO segment.
w_i = the normalized effective ratio of the *i*th TMC name of a RHiNO segment.
s_i = the speed measurement value of the *i*th TMC name of a RHiNO segment.

Equation 9 was applied to the following speed measurement variables: SpdAve, Spd85, RefSpd, SpdAveDay, SpdAveNight, SpdAveMTWT, SpdAveFSS, SpdFFAve, and SpdFF85.

$$A = \sum_i^n w_i s_i \tag{9}$$

Where:

A = the aggregated speed measurement variable.
n = the number of unique TMC names of a RHiNO segment.
w_i = the normalized effective ratio of the *i*th TMC name of a RHiNO segment.
s_i = the speed measurement value of the *i*th TMC name of a RHiNO segment.

4.1.3 Daily-Level Databases

Considering the time frame and resources available for this project, TTI decided to develop daily-level models by using a case-control study design. The data preparation method for the daily model is similar to the annual model, except for the following steps:

- Speed measurement data and crash data were aggregated into daily intervals (see Table 10). For example, a segment has been repeated 1,860 times (365 days multiplied by 4 years) with all daily-level speed measures.
- The average of precipitation values was calculated by day.
- The RHiNO segment-level information was repeated for each day-level row.

Table 10. Speed Measures Used for Daily-Level Analysis.

Attribute Name	Definition
SpdAve	Average of all 5-minute interval all vehicle speeds by date
SpdStd	SD of all 5-minute interval all vehicle speeds by date
Spd85	85th percentile speed of all 5-minute interval all vehicle speeds by date
RefSpd	Average of all 5-minute interval all vehicle reference speeds by date
SpdAveDay	Average speed of 5-minute interval all vehicle speeds for daytime (5 < hour < 18) by date
SpdStdDay	SD of 5-minute interval all vehicle speeds for daytime (5 < hour < 18) by date
SpdAveNight	Average speed of 5-minute interval all vehicle speeds for daytime (17 < hour < 24 and –1 < hour < 6) by date
SpdStdNight	SD of 5-minute interval all vehicle speeds for daytime (17 < hour < 24 and –1 < hour < 6) by date
SpdFFAve	Average speed all 5-minute interval all vehicle speeds by date when 5-minute speed is > RefSpd
SpdFF85	Average speed all 5-minute interval all vehicle speeds by date when 5-minute speed is > RefSpd

When crash data are aggregated into the daily level, the daily crash frequency usually generates a massive number of zero values (over 99 percent of the daily data contain zero crash counts). TTI chose to add two variables: Daily_FI_Y_N (whether a daily fatal and injury crash occurred or not) and Daily_TOTAL_Y_N (whether a daily total crash occurred or not) to reflect crash occurrence. Then, TTI applied the following steps to develop the final daily-level data sets (see Figure 7):

- For each observation with Daily_TOTAL_Y_N as 1, the observation is repeated *N* times, where *N* equals the value of Daily_TOTAL, which is the daily total crash frequency. In a daily model, unique_id and Date define a unique observation.
- Then, for the variable Daily_FI_Y_N, it is replaced with *M* 1 and *N-M* 0 where *M* equals the value of Daily_FI, which is the daily fatal and injury crash frequency.
- Finally, for each observation with Daily_TOTAL_Y_N as 1, the same number of observations of the same RHiNO roadway segment (with the same unique_id) on the same day of the week (with the same DOW) with no daily total crash occurrence is matched into the data set.

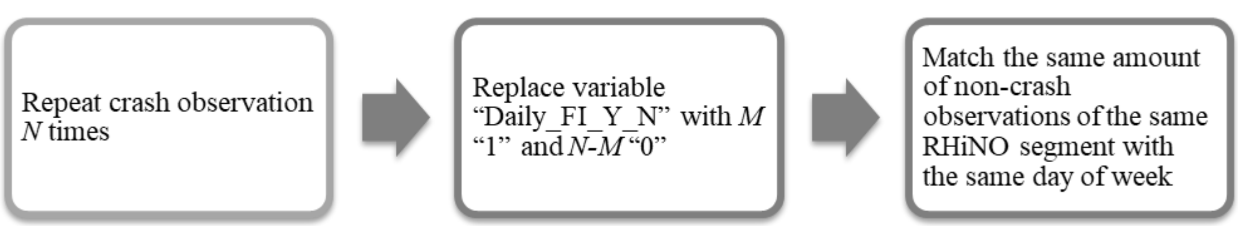

Figure 7. Case-Control Study Design Dataset Preparation Process.

Table 11 is an example of the control-case data set.

Table 11. Example of the Case-Control Study Design.

unique_id	Date	DOW	Daily_FI	Daily_ TOTAL	Daily_FI_Y_N	Daily_TOTAL_Y_N
1	2017-1-1	Mon	1	3	1	1
1	2017-1-1	Mon	1	3	0	1
1	2017-1-1	Mon	1	3	0	1
1	2018-5-2	Mon	0	0	0	0
1	2019-3-4	Mon	0	0	0	0
1	2017-12-4	Mon	0	0	0	0

4.2 DESCRIPTIVE STATISTICS

4.2.1 Annual-Level Databases

This section presents the descriptive statistics of the key variables used for the annual-level analysis. Table 12 through Table 15 list descriptive statistics of the key variables for annual-level analysis.

TTI populated a list of speed measures that quantify speed with respect to different aspects. Given the long list, it is imperative to select a measure that is appropriate and meaningful to include in the SPF development. The final variables were selected by performing co-relation analysis. The speed measures were highly correlated, so researchers decided to use 85th percentile FFS (SpdFF85) and SD in the speed (SpdStd) in the SPFs, given their wide range use.

Table 12. Descriptive Statistics of Rural Four-Lane Freeways (Annual-Level Data).

Variables	Code	Mean	SD	Min.	Max.
Annual average daily traffic (AADT)	ADT_CUR	29,588.537	17,325.103	4500	112,248
Truck proportion	TRK_AADT_P	33.736	11.710	5.9	78.8
Lane width	LANE_WIDTH	11.935	0.460	8	20
Inside shoulder width	S_WID_I	10.615	3.621	0	32
Outside shoulder width	S_WID_O	19.489	2.691	0	48
Median width	MED_WID	55.060	35.881	1	455
85th percentile FFS	SpdFF85	78.138	3.240	69.513	87.549
SD in speed	SpdStd	3.075	0.925	1.311	10.000
K-factor	K_FAC	9.498	1.591	5.3	19.8

Table 13. Descriptive Statistics of Rural Four-Lane Divided Roadways (Annual-Level Data).

Variables	Code	Mean	SD	Min.	Max.
AADT	ADT_CUR	12,636.625	7563.621	243	53,382
Truck proportion	TRK_AADT_P	19.663	9.810	1.2	90.4
Lane width	LANE_WIDTH	12.226	1.120	6	25
Inside shoulder width	S_WID_I	9.740	4.126	0	30
Outside shoulder width	S_WID_O	17.400	4.899	0	38
Median width	MED_WID	54.949	34.581	1	450
85th percentile FFS	SpdFF85	77.328	4.502	54.715	93.001
SD in speed	SpdStd	5.131	1.736	1.987	18.279
K-factor	K_FAC	9.648	1.322	7	20

Table 14. Descriptive Statistics of Rural Four-Lane Undivided Roadways (Annual-Level Data).

Variables	Code	Mean	SD	Min.	Max.
AADT	ADT_CUR	4695.331	3158.593	134	50191
Truck proportion	TRK_AADT_P	20.374	10.488	1.2	66.6
Lane width	LANE_WIDTH	12.540	1.625	6	32
Inside shoulder width	S_WID_I	8.143	2.586	0	25
Outside shoulder width	S_WID_O	8.360	2.563	0	25
85th percentile FFS	SpdFF85	76.216	4.936	36.647	85.227
SD in speed	SpdStd	6.217	1.953	2.219	17.420
K-factor	K_FAC	9.785	1.530	5.9	19.1

Table 15. Descriptive Statistics of Rural Two-Lane Roadways (Annual-Level Data).

Variables	Code	Mean	SD	Min.	Max.
AADT	ADT_CUR	4695.331	3158.593	134	50191
Truck proportion	TRK_AADT_P	20.374	10.488	1.2	66.6
Lane width	LANE_WIDTH	12.540	1.625	6	32
Inside shoulder width	S_WID_I	8.143	2.586	0	25
Outside shoulder width	S_WID_O	8.360	2.563	0	25
85th percentile FFS	SpdFF85	76.216	4.936	36.647	85.227
SD in speed	SpdStd	6.217	1.953	2.219	17.420
K-factor	K_FAC	9.785	1.530	5.9	19.1

4.2.2 Daily-Level Databases

This section presents the descriptive statistics of the key variables used for the daily-level analysis. Table 16 through Table 19 provide descriptive statistics of the key variables for daily-level analysis.

Table 16. Descriptive Statistics of Rural Four-Lane Interstate Freeways (Daily-Level Data).

Variables	Code	Mean	SD	Min.	Max.
Average operating speed	SpdAve	66.510	3.048	17.5	75.435
SD of operating speed	SpdStd	3.191	1.923	0.778	19.467
Daily precipitation	daily_precip	0.118	0.639	0	49.2
Truck AADT percentage	TRK_AADT_P	32.202	10.188	5.9	78.8
Segment length	Length	1.064	0.565	0.001	2
Shoulder width inside	SW_I	10.085	2.921	0	24
Shoulder width outside	SW_O	19.674	2.760	0	48
Average shoulder width	SW_Avg	14.879	1.725	0	28
Lane width	LANE_WIDTH	11.923	0.329	11	14
Median width	MED_WID	51.696	34.428	1	455

Table 17. Descriptive Statistics of Rural Four-Lane Divided Roadways (Daily-Level Data).

Variables	Code	Mean	SD	Min.	Max.
Average operating speed	SpdAve	61.806	8.722	11.5	76.711
SD of operating speed	SpdStd	4.986	2.043	1.532	28.991
Daily precipitation	daily_precip	0.102	0.459	0	41.2
Log of AADT	LnAADT	9.490	0.592	5.737	10.885
Truck AADT percentage	TRK_AADT_P	18.422	10.040	1.8	69.5
Segment length	Length	0.918	0.600	0.001	1.997
Shoulder width inside	SW_I	9.368	4.083	0	40
Shoulder width outside	SW_O	17.549	4.827	0	40
Average shoulder width	SW_Avg	13.459	3.623	0	40
Lane width	LANE_WIDTH	12.015	0.258	9	14
Median width	MED_WID	54.729	34.902	1	450

Table 18. Descriptive Statistics of Rural Four-Lane Undivided Roadways (Daily-Level Data).

Variables	Code	Mean	SD	Min.	Max.
Average operating speed	SpdAve	52.763	12.889	9.835	73.241
SD of operating speed	SpdStd	5.770	1.908	1.834	16.263
Daily precipitation	daily_precip	0.098	0.407	0	24.510
Log of AADT	LnAADT	9.122	0.656	4.522	1539
Truck AADT percentage	TRK_AADT_P	14.157	8.936	0.6	68
Segment length	Length	0.604	0.556	0.001	1.995
Shoulder width inside	SW_I	4.973	4.035	0	27
Shoulder width outside	SW_O	5.512	4.318	0	24
Average shoulder width	SW_Avg	5.243	3.971	0	25.5
Lane width	LANE_WIDTH	11.841	0.677	9	14

Table 19. Descriptive Statistic of Rural Two-Lane Roadways (Daily-Level Data).

Variables	Code	Mean	SD	Min.	Max.
Average operating speed	SpdAve	56.962	9.105	14.833	76
SD of operating speed	SpdStd	5.988	2.140	1.002	2726
Daily precipitation	daily_precip	0.100	0.526	0	32
Log of AADT	LnAADT	8.605	0.689	2.197	10.824
Truck AADT percentage	TRK_AADT_P	18.487	10.503	1.2	66.6
Segment length	Length	0.824	0.592	0.001	1.999
Shoulder width inside	SW_I	8.014	2.604	0	25
Shoulder width outside	SW_O	8.306	2.412	0	25
Average shoulder width	SW_Avg	8.160	2.326	0	25
Lane width	LANE_WIDTH	12.214	0.528	9	14

4.3 SAFETY EVALUATION

4.3.1 Annual-Level Models

For annual-level analysis, SPFs have been developed for four facility types (rural two-lane roadways, rural four-lane divided roadways, rural four-lane undivided roadways, and rural freeways) for two severity groups (fatal and severe injury crashes and PDO crashes).

4.3.2 SPFs for Freeways

4.3.2.1. Predicted Crash Frequency

For the SPF development, the team only considered the four-lane freeways, given the large sample size in that category. Crash severities are usually divided into five major categories:

- Fatal (K).
- Incapacitating injury (A).
- Non-incapacitating injury (B).
- Possible injury (C).
- No injury or PDO (O).

Since the PDO crashes are usually underreported, it was decided to develop separate models for FI (or KABC) and PDO crashes. TTI first examined different functional forms with various combinations of variables while modeling the FI crashes. It is assumed that the FI crash model provides a true relationship between crashes and independent variables. The form presented reflects the findings from several preliminary regression analyses. The same form is also used for modeling the PDO crashes, even if some variables are strongly insignificant or counterintuitive. The predicted crash frequency is calculated in Equation 10.

$$N = L \times y \times e^{b_0 + b_{aadt}\ln(AADT)} \times CMF_{tk} \times CMF_{lw} \times CMF_{isw} \times CMF_{osw} \times CMF_{mw} \times CMF_{spd} \times CMF_{std} \times CMF_{kf}; \quad (10)$$

With:

$$CMF_{tk} = e^{b_{tk}(tk\_perc/100)}$$
$$CMF_{lw} = e^{b_{lw}(lw-12)}$$
$$CMF_{isw} = e^{b_{isw}(isw-4)}$$
$$CMF_{osw} = e^{b_{osw}(osw-10)}$$
$$CMF_{mw} = e^{b_{mw}(mw-48)}$$
$$CMF_{spd} = \begin{cases} e^{b_{spd}(SpdFF85-65)}, if\ RefSpd < 70mph \\ e^{b_{spd}(SpdFF85-70)}, if\ 70mph \le RefSpd < 75mph \\ e^{b_{spd}(SpdFF85-75)}, \quad if\ RefSpd \ge 75mph \end{cases}$$
$$CMF_{std} = e^{b_{std}(SpdStd-3)}$$
$$CMF_{kf} = e^{b_{kf}(kf-10)}$$

Where:

N = the predicted annual average crash frequency.
L = the segment length, miles.
y = the number of years of crash data.
$AADT$ = the average annual daily traffic, vehicles per day (vpd).
CMF_{tk} = the CMF for the truck proportion in the traffic mix.
CMF_{lw} = the CMF for lane width.
CMF_{isw} = the CMF for inside shoulder width.
CMF_{osw} = the CMF for outside shoulder width.
CMF_{mw} = the CMF for median width.
CMF_{spd} = the CMF for 85th FFS.
CMF_{std} = the CMF for SD in speed.
CMF_{kf} = the CMF for K-factor.
tk_perc = the percent of trucks in the traffic mix, percent.
lw = the lane width, feet.
isw = the inside shoulder width, feet.
osw = the outside shoulder width, feet.
mw = the median width, feet.
$RefSpd$ = the reference speed, mph.
kf = K-factor.
b_j = the calibrated coefficients.

Table 20 and Table 21 provide calibrated coefficients for FI crashes and PDO crashes, respectively. A significance level of 5 percent is used to include the variables in the model. However, the coefficient is also considered if it is at a 20 percent level, is intuitive, and is within logical boundaries. This study used SAS software platform's NLMIXED method to evaluate the model coefficients. The reason behind using this method is nonlinear and discontinuous nature of the prediction model. To determine the best-fit coefficients, the log-likelihood function for the NB distribution was applied.

Table 20. Calibrated Coefficients for Fatal and Injury Crashes on Freeways.

Coefficient		Variable	Value	SD	t-statistic	p-value
b_0		Intercept	–12.2194	0.461	–26.52	<0.0001
b_{aadt}		AADT	1.2256	0.039	31.34	<0.0001
b_{tk}		Truck proportion	0.3074	0.174	1.76	0.08
b_{lw}		Lane width	–0.1032	0.038	–2.75	0.01
b_{isw}		Inside shoulder width	–0.0144	0.006	–2.41	0.02
b_{osw}		Outside shoulder width	–0.0212	0.007	–3.17	0.00
b_{mw}		Median width	–0.0014	0.001	–2.47	0.01
b_s	$RefSpd < 70mph$	85th percentile FFS	0.0148	0.011	1.36	0.17
	$70mph \leq RefSpd < 75mph$		0.0343	0.008	4.52	<0.0001
	$RefSpd \geq 75mph$		0.0457	0.011	4.25	<0.0001
b_{std}		SD in speed	0.1881	0.020	9.27	<0.0001
b_{kf}		K-factor	0.0417	0.009	4.84	<0.0001
k		Inverse dispersion parameter	2.0095	0.062	32.28	<0.0001

Table 21. Calibrated Coefficients for Property Damage Only Crashes on Freeways.

Coefficient		Variable	Value	SD	t-statistic	p-value
b_0		Intercept	–5.0457	0.529	–9.54	<0.0001
b_{aadt}		AADT	0.6926	0.045	15.49	<0.0001
b_{tk}		Truck proportion	–0.1090	0.189	–0.58	0.56
b_{lw}		Lane width	–0.1796	0.040	–4.44	<0.0001
b_{isw}		Inside shoulder width	–0.0283	0.006	–4.44	<0.0001
b_{osw}		Outside shoulder width	–0.0155	0.007	–2.21	0.03
b_{mw}		Median width	–0.0029	0.001	–5.03	<0.0001
b_{spd}	$RefSpd < 70mph$	85th percentile FFS	–0.0288	0.012	–2.35	0.02
	$70mph \leq RefSpd < 75mph$		–0.0335	0.008	–3.97	<0.0001
	$RefSpd \geq 75mph$		–0.1303	0.012	–10.54	<0.0001
b_{std}		SD in speed	0.3583	0.024	14.95	<.0001
b_{kf}		K-factor	0.0203	0.009	2.16	0.03
k		Inverse dispersion parameter	1.2125	0.044	27.51	<.0001

Figure 8 and Figure 9 show a comparison of different calibrated freeway SPFs for FI crashes and total crashes (total crashes were obtained by adding crash estimates obtained from FI and PDO models), respectively. The SPFs developed in this project were compared with the HSM SPFs

(AASHTO, 2010) and the *Texas Roadway Safety Design Workbook* (referred to as the Texas WB) (Bonneson and Pratt, 2008). The equations were plotted for the case of all CMFs equal to 1.0 (representing base conditions). The SPFs do not include the same set of base conditions, and so they are not directly comparable to each other. In addition, the HSM SPFs are not calibrated to Texas conditions, and Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since the Texas WB includes SPFs for FI crashes only, the comparison of 0-7051 total crashes is made just with HSM SPFs for total crashes.

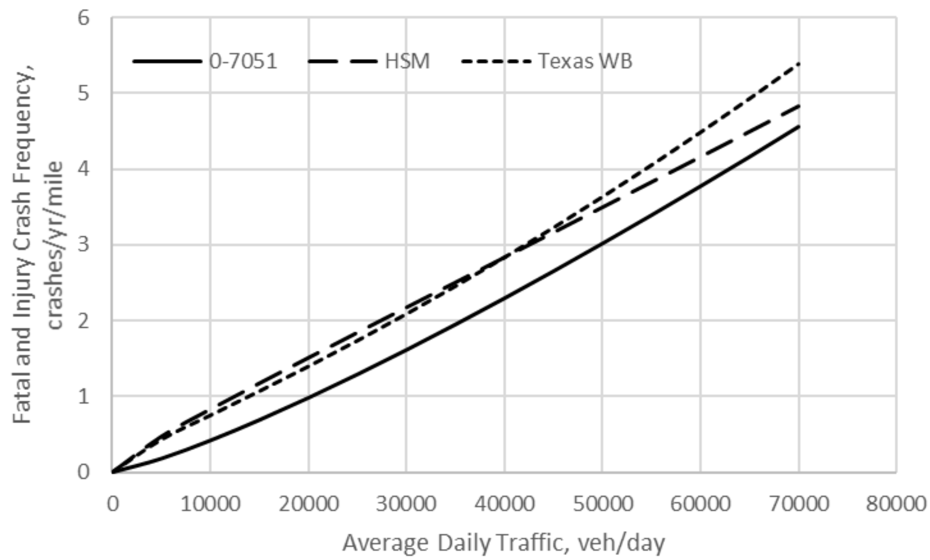

Figure 8. SPF Comparison for Fatal and Injury Crashes on Freeways.

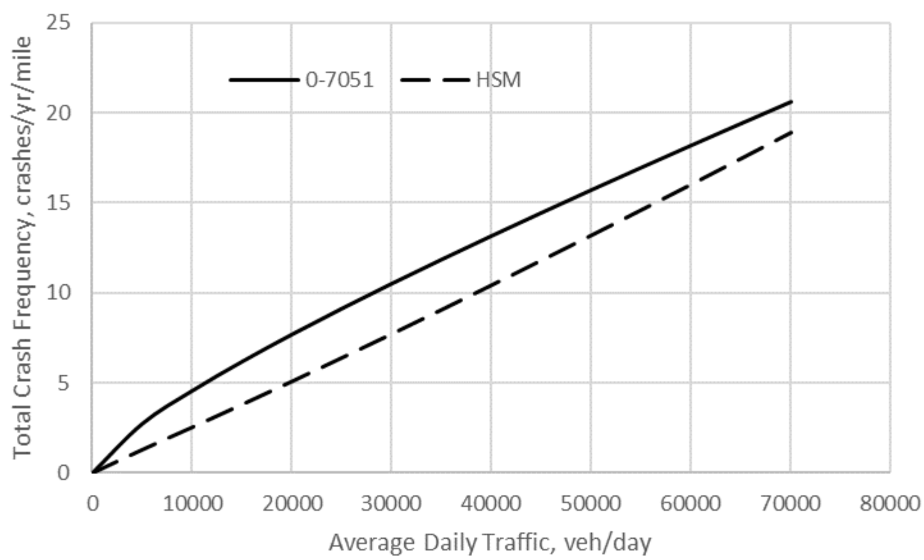

Figure 9. SPF Comparison for Total Crashes on Freeways.

4.3.2.2. Crash Modification Factors

This study calibrated several CMFs in aggregation with the SPFs. All of them were calibrated using the FI crash data. Collectively, they explain the association between various operational and geometric factors and traffic crash count. These CMFs are described in this section and, where possible, compared with the findings from the HSM and the Texas WB as a means of model validation.

4.3.2.3. Truck Proportion CMF

The truck proportion CMF is described using Equation 11.

$$CMF_{tk} = e^{0.3074(tk\_perc/100)} \tag{11}$$

The base condition for this CMF is no trucks on the highway. Figure 10 shows the truck proportion CMF. The CMF shows that there is an estimated 5 percent increase in crashes for every 15 percent increase in the truck proportion.

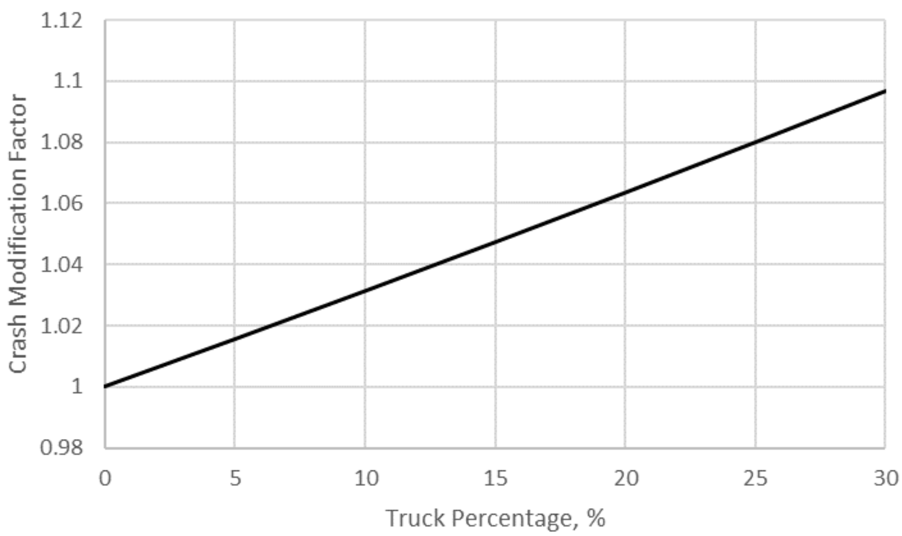

Figure 10. CMF for Truck Proportion on Freeways.

4.3.2.4. Lane Width CMF

The lane width CMF is described using Equation 12.

$$CMF_{lw} = e^{-0.1032(lw-12)} \tag{12}$$

The base condition for this CMF is a lane width of 12 feet. The lane width used in this CMF is an average for all through lanes on the segment. Figure 11 shows the lane width CMF developed in this study. The lane widths used to calibrate this CMF range from 10 to 13 feet. Figure 11 also shows CMFs presented in the HSM and the Texas WB. The proposed CMF is more sensitive to the lane width than the CMFs in the HSM and the Texas WB (the separate CMFs are for different numbers of lanes in the Texas WB; the comparison is based on the CMF for four-lane freeways).

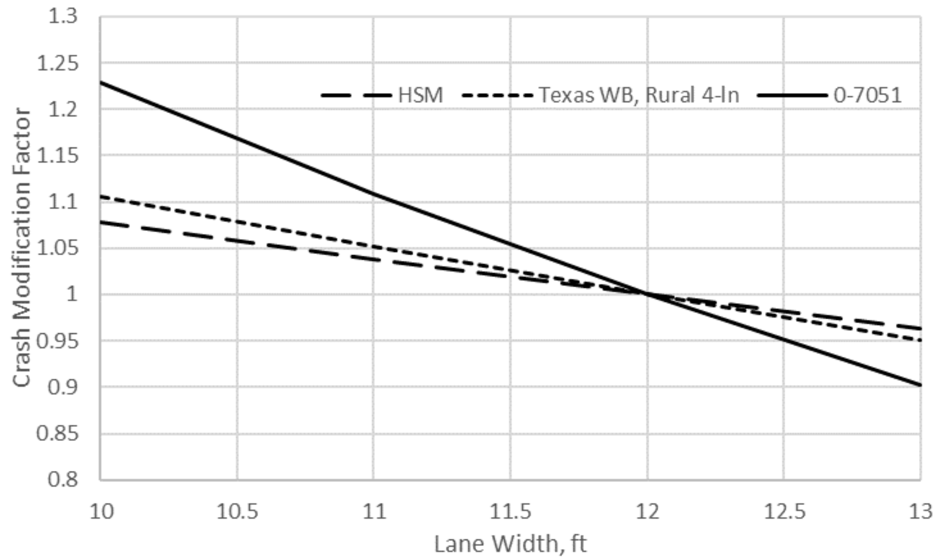

Figure 11. CMF for Lane Width on Freeways.

4.3.2.5. Inside Shoulder Width CMF

The inside shoulder width CMF is described using Equation 13.

$$CMF_{isw} = e^{-0.0144(isw-4)} \tag{13}$$

The base condition for this CMF is an inside shoulder width of 4 feet. The width used in this CMF is an average for inside shoulders in both directions. Figure 12 shows the inside shoulder width CMF developed in this study. The inside shoulder widths used to calibrate this CMF ranged from 2 to 10 feet. Figure 12 also shows CMFs presented in the HSM and the Texas WB. The proposed CMF closely tracks the CMFs presented in the HSM. The CMF presented in the Texas WB is more sensitive to the inside shoulder width than the proposed CMF or the one in the HSM.

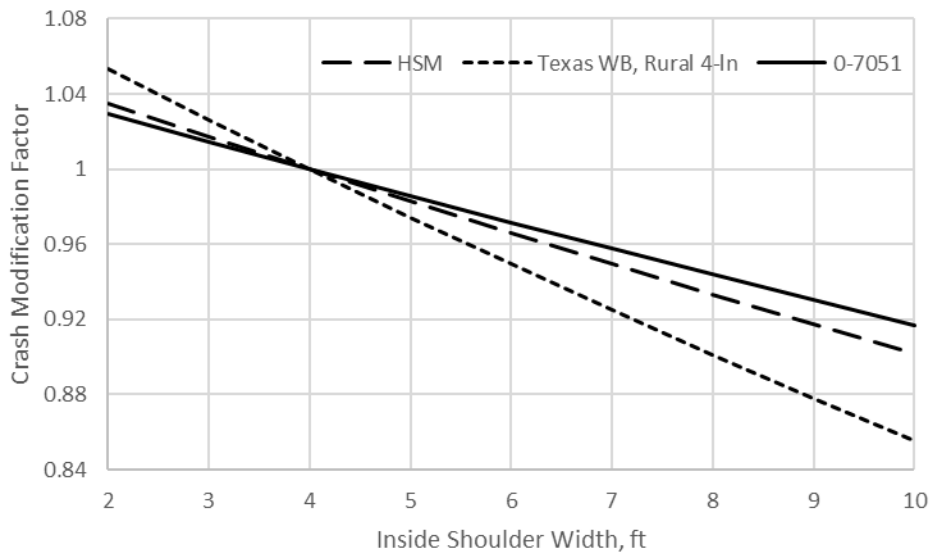

Figure 12. CMF for Inside Shoulder Width on Freeways.

4.3.2.6. Outside Shoulder Width CMF

The outside shoulder width CMF is described using Equation 14.

$$CMF_{osw} = e^{-0.0212(osw-10)} \tag{14}$$

The base condition for this CMF is an outside shoulder width of 10 feet. The width used in this CMF is an average for outside shoulders in both directions. Figure 13 shows the outside shoulder width CMF developed in this study. The outside shoulder widths used to calibrate this CMF range from 6 to 12 feet. Figure 13 also shows CMFs presented in the HSM and the Texas WB. The proposed CMF closely tracks the CMFs presented in the Texas WB. The CMF presented in the HSM is more sensitive to the outside shoulder width than the proposed CMF or the one in the Texas WB.

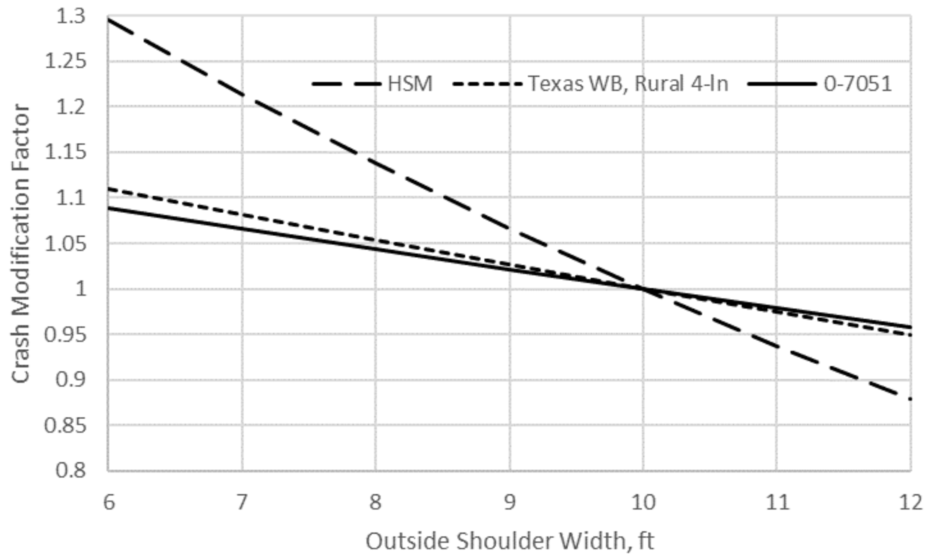

Figure 13. CMF for Outside Shoulder Width on Freeways.

4.3.2.7. Median Width CMF

The median width CMF is described using Equation 15.

$$CMF_{mw} = e^{-0.0057(W_m - 15)} \tag{15}$$

The base condition for this CMF is a median width of 48 feet. Figure 14 shows the median width CMF. This CMF applies only to traversable medians without any kind of traffic barriers. Figure 14 compares the CMF proposed in this research to the CMF in the HSM and the Texas WB. The proposed CMF closely tracks the CMFs presented in the HSM. The CMF presented in the Texas WB is more sensitive to the inside shoulder width than the proposed CMF or the one in the HSM.

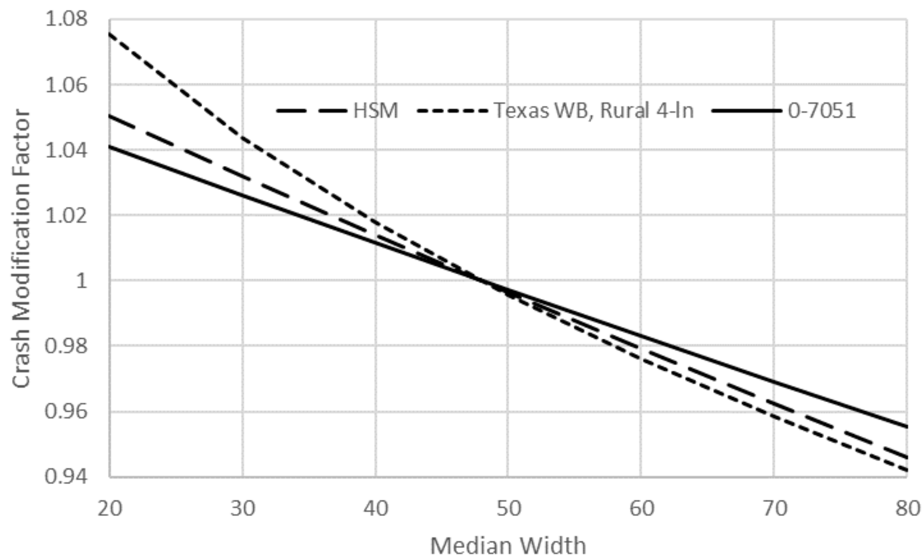

Figure 14. CMF for Median Width on Freeways.

4.3.2.8. Operating Speed CMF

The operating speed CMF is described using Equation 16.

$$CMF_{spd} = \begin{cases} e^{0.0148(SpdFF85-65)}, \ if\ RefSpd < 70mph \\ e^{0.0343(SpdFF85-70)}, if\ 70mph \leq \ RefSpd < 75mph \\ e^{0.0457(SpdFF85-75)}, \ if\ RefSpd \geq 75mph \end{cases} \tag{16}$$

The base condition for this CMF varies according to the reference speed. The reference speed (see definition in Table 9) acts as a surrogate for the posted speed limit. Since the operating speed CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 15 shows the operating speed CMFs by reference speed. The CMF shows that with the increase in operating speeds, the relative increase in crashes is more on the higher-reference-speed freeways than on the freeways with lower reference speeds.

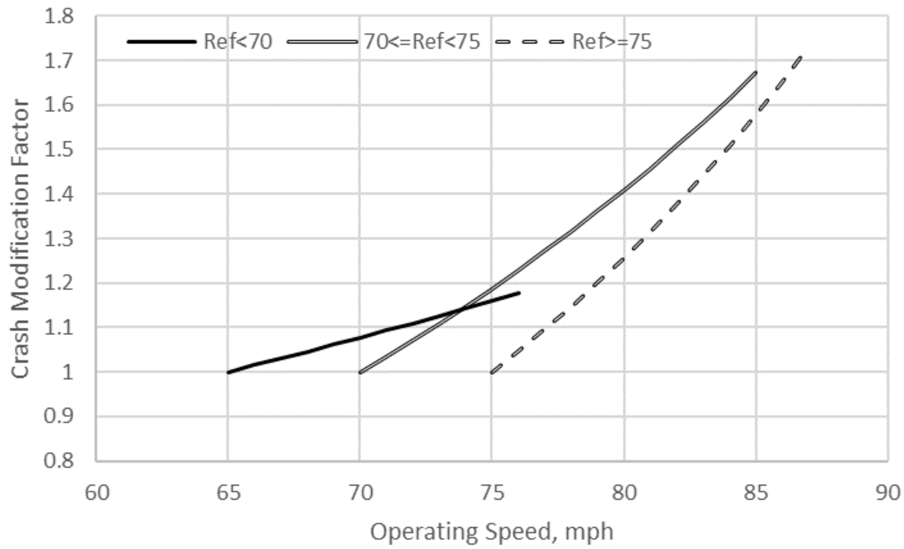

Figure 15. CMF for Median Width on Freeways.

4.3.2.9. Speed Variation CMF

The speed variation CMF is described using Equation 17.

$$CMF_{std} = e^{0.1881(SpdStd-3)} \tag{17}$$

Since the speed variation CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 16 shows the speed variation CMF. The CMF shows that the speed variation has a major influence on the occurrence of crashes.

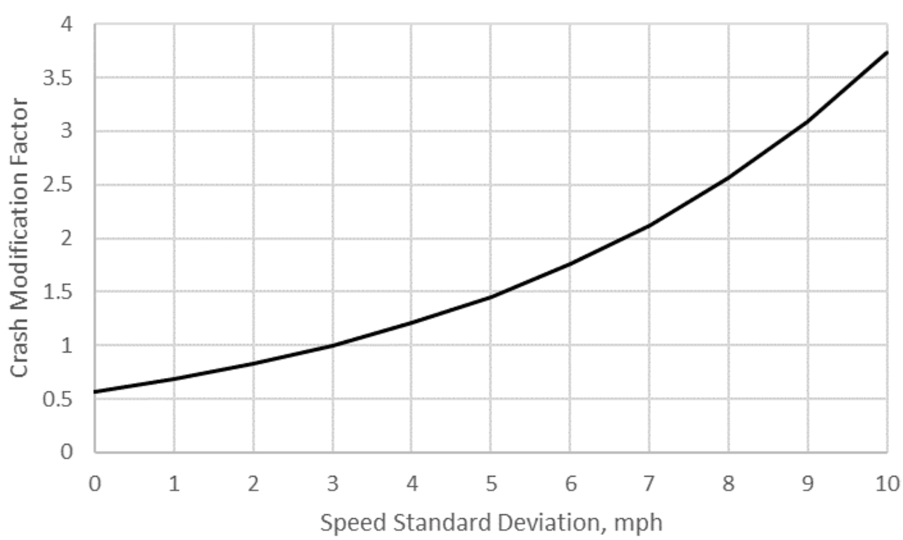

Figure 16. CMF for Speed Variation on Freeways.

4.3.2.10. K-Factor CMF

K-factor, also known as the design hour factor, illustrates the proportion of AADT expected to occur in the design hour. K-factors or bias factors are zone-to-zone factors, other than travel time, that are specific to an urban area and affect travel patterns. The K-factor CMF is described using Equation 18.

$$CMF_{kf} = e^{0.0417(kf-10)} \tag{18}$$

The base condition for this CMF is a k-factor of 10. Since the k-factor CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 17 shows the k-factor CMF. The CMF shows that the crashes increase with an increase in the k-factor.

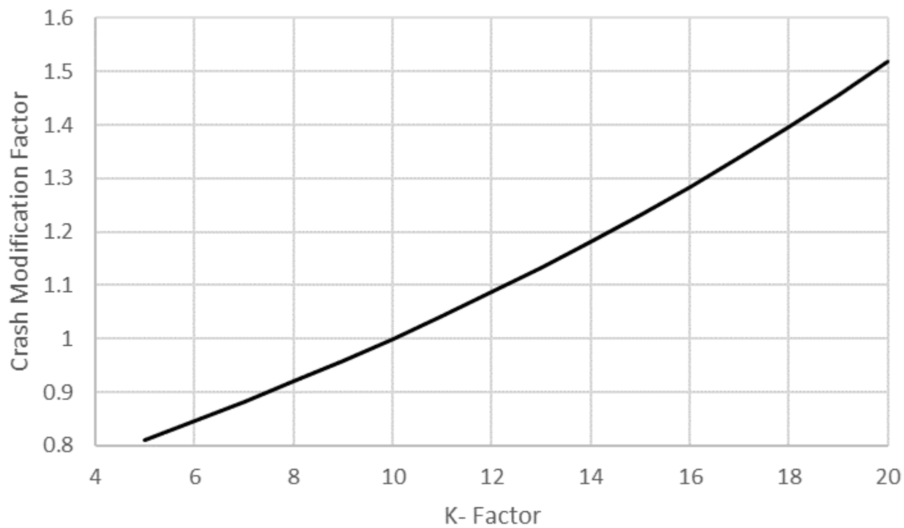

Figure 17. CMF for K-Factor on Freeways.

4.3.3 SPFs for Multilane Divided Highways

4.3.3.1. Predicted Crash Frequency

For the SPF development, the research team just considered the four-lane highways, given the large sample size in that category. The predicted crash frequency was calculated using Equation 19.

$$N = L \times y \times e^{b_0 + b_{aadt} \ln(AADT)} \times CMF_{tk} \times CMF_{isw} \times CMF_{osw} \times CMF_{spd} \times CMF_{std} \times CMF_{kf} \tag{19}$$

With:

$$CMF_{tk} = e^{b_{tk}(I_{tk})}$$
$$CMF_{osw} = e^{b_{osw}(osw-8)}$$

$$CMF_{spd} = \begin{cases} e^{b_{spd}(SpdFF85-60)}, \; if \; RefSpd < 65mph \\ e^{b_{spd}(SpdFF85-65)}, if \; 65mph \; \leq \; RefSpd \; < 70mph \\ e^{b_{spd}(SpdFF85-70)}, \; if \; 70mph \leq RefSpd < 75mph \\ e^{b_{spd}(SpdFF85-75)}, \; if \; RefSpd \geq 75mph \end{cases}$$

Where:

CMF_{tk} = the CMF for higher truck proportion in the traffic mix.
I_{tk} = the indicator variable for higher truck proportions (1 if ≥16 percent; 0 otherwise).

Table 22 and Table 23 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 22. Calibrated Coefficients for Fatal and Injury Crashes on Multilane Divided Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
b_0		Intercept	–9.7508	0.3577	–27.26	<0.0001
b_{aadt}		AADT	0.9839	0.035	28.28	<0.0001
b_{tk}		Higher truck proportion	0.0447	0.037	1.21	0.23
b_{isw}		Inside shoulder width	–0.0082	0.005	–1.75	0.08
b_{osw}		Outside shoulder width	–0.0161	0.006	–2.66	0.01
b_{spd}	$RefSpd < 65mph$	85th percentile FFS	0.0159	0.012	1.29	0.20
	$65mph \leq RefSpd < 70mph$		0.0133	0.009	1.55	0.12
	$70mph \leq RefSpd < 75mph$		0.0279	0.009	3.16	0.00
	$RefSpd \geq 75mph$		0.0287	0.011	2.62	0.01
b_{std}		SD in speed	0.0371	0.012	3.18	0.00
b_{kf}		K-factor	0.0292	0.015	1.99	0.05
k		Inverse dispersion parameter	1.7841	0.075	23.79	<0.0001

Table 23. Calibrated Coefficients for Property Damage Only Crashes on Multilane Divided Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
b_0		Intercept	–8.6983	0.295	–29.50	<0.0001
b_{aadt}		AADT	0.9368	0.028	32.94	<0.0001
b_{tk}		Higher truck proportion	0.0583	0.032	1.84	0.07
b_{isw}		Inside shoulder width	0.0054	0.004	1.38	0.17
b_{osw}		Outside shoulder width	–0.0047	0.005	–0.90	0.37
b_{spd}	$RefSpd < 65mph$	85th percentile FFS	0.0343	0.010	3.30	0.00
	$65mph \leq RefSpd < 70mph$		0.0147	0.007	2.05	0.04
	$70mph \leq RefSpd < 75mph$		0.0257	0.007	3.54	0.00
	$RefSpd \geq 75mph$		0.0160	0.009	1.79	0.07
b_{std}		SD in speed	0.0229	0.010	2.33	0.02
b_{kf}		K-factor	0.0095	0.012	0.78	0.44
k		Inverse dispersion parameter	1.6055	0.044	36.77	<0.0001

Figure 18 and Figure 19 compare different calibrated multilane divided highways for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the HSM SPFs and the Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). The SPFs do not include the same set of base conditions, so they are not directly comparable to each other. In addition, the HSM SPFs are not calibrated to Texas conditions, and the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since the Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

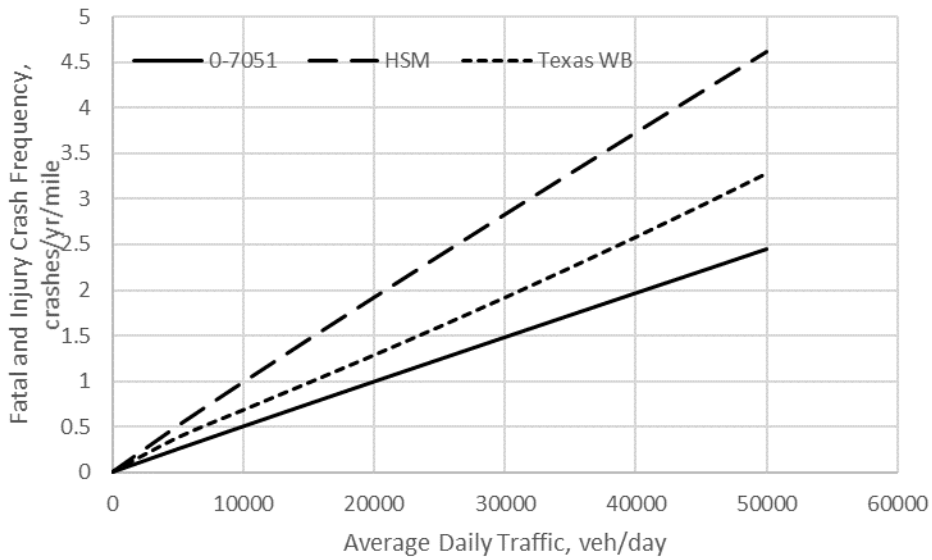

Figure 18. SPF Comparison for Fatal and Injury Crashes on Multilane Divided Highways.

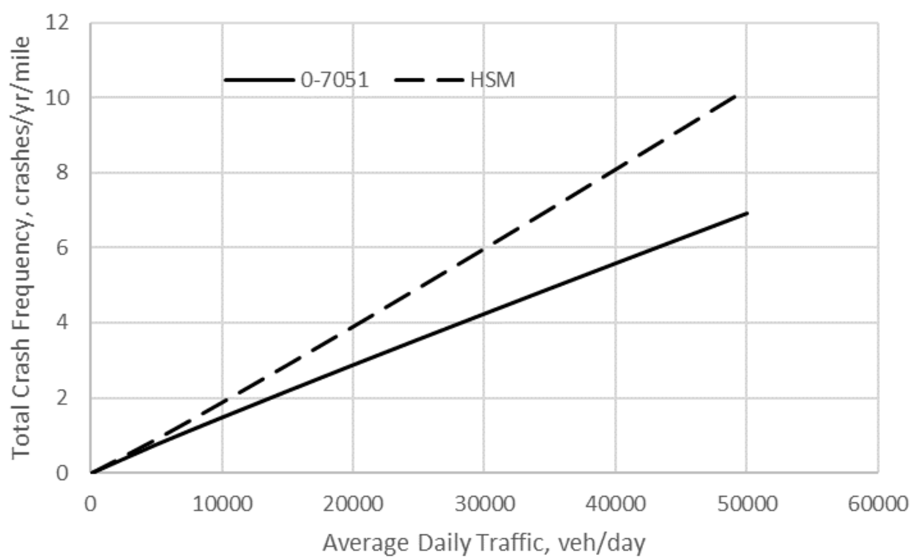

Figure 19. SPF Comparison for Total Crashes on Multilane Divided Highways.

4.3.3.2. Crash Modification Factors

Several CMFs were calibrated in conjunction with the SPFs. All of them were calibrated using the FI crash data. These CMFs are described in this section and, where possible, compared with the findings from the HSM and the Texas WB as a means of model validation.

4.3.3.3. Higher Truck Proportion CMF

The higher truck proportion CMF is described using Equation 20.

$$CMF_{tk} = e^{0.0447(I_{tk})} \tag{20}$$

The base condition for this CMF is traffic mix with less than 16 percent trucks. Figure 20 shows the higher truck proportion CMF. The CMF shows that highways with greater than 16 percent trucks experience an average of 4.5 percent more crashes than highways with less than 16 percent trucks.

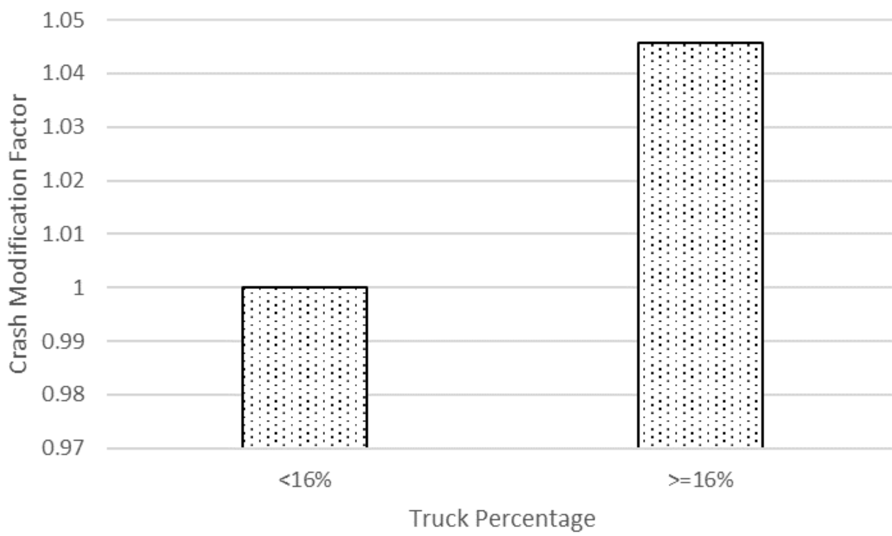

Figure 20. CMF for Higher Truck Proportion on Multilane Divided Highways.

4.3.3.4. Inside Shoulder Width CMF

The inside shoulder width CMF is described using Equation 21.

$$CMF_{isw} = e^{-0.0082(isw-4)} \tag{21}$$

The base condition for this CMF is an inside shoulder width of 4 feet. The width used in this CMF is an average for inside shoulders in both directions. Figure 21 shows the inside shoulder width CMF developed in this study. The inside shoulder widths used to calibrate this CMF range from 2 to 10 feet. Figure 21 also shows the CMF presented in the HSM. The proposed CMF is less sensitive to the inside shoulder width than the one in the HSM.

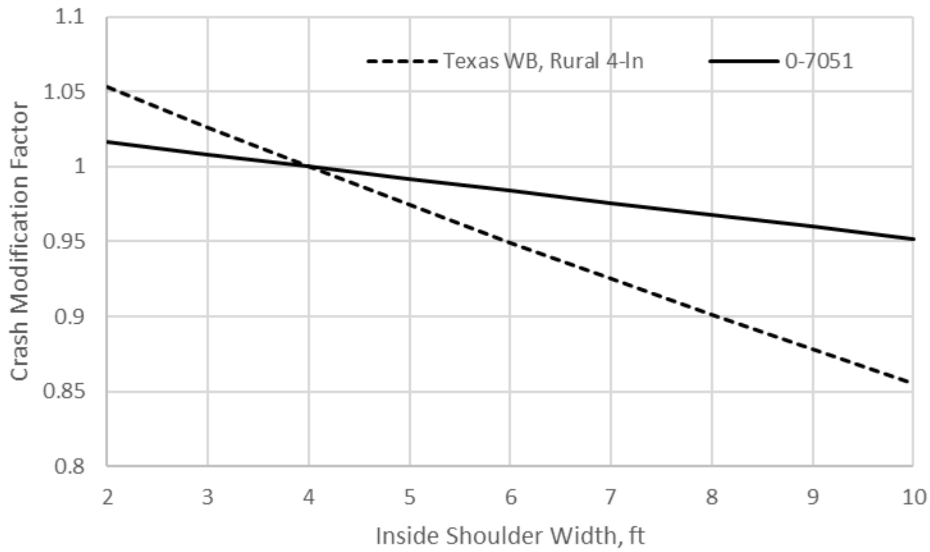

Figure 21. CMF for Inside Shoulder Width on Multilane Divided Highways.

4.3.3.5. Outside Shoulder Width CMF

The outside shoulder width CMF is described using Equation 22.

$$CMF_{osw} = e^{-0.0161(osw-8)} \tag{22}$$

The base condition for this CMF is an outside shoulder width of 8 feet. The width used in this CMF is an average for outside shoulders in both directions. Figure 22 shows the outside shoulder width CMF developed in this study. The outside shoulder widths used to calibrate this CMF range from 6 to 12 ft. Figure 22 also shows CMFs presented in the HSM and the Texas WB. The proposed CMF is less sensitive to the outside shoulder width than the HSM and the Texas WB CMFs. The HSM states that no additional benefits can be achieved for shoulders greater than 8 feet. However, both this project and the Texas WB found an increase in benefits for outside shoulder widths between 8 and 12 feet.

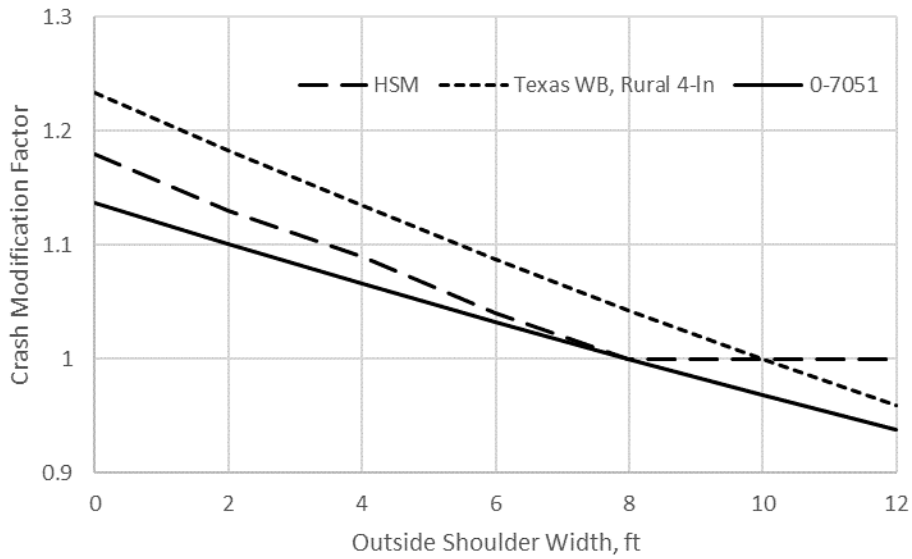

Figure 22. CMF for Outside Shoulder Width on Multilane Divided Highways.

4.3.3.6. Operating Speed CMF

The operating speed CMF is described using Equation 23.

$$CMF_{spd} = \begin{cases} e^{0.0159(SpdFF85-60)}, if\ RefSpd < 65mph \\ e^{0.0133(SpdFF85-65)}, if\ 65mph \le RefSpd < 70mph \\ e^{0.0279(SpdFF85-70)}, if\ 70mph\ \le\ RefSpd\ <\ 75mph \\ e^{0.0287(SpdFF85-75)}, if\ RefSpd \ge 75mph \end{cases} \quad (23)$$

The base condition for this CMF varies according to the reference speed. Since the operating speed CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 23 shows the operating speed CMF for highways with different reference speeds. The CMF shows that with the increase in operating speeds, the relative increase in crashes is more on higher reference speed highways than on highways with lower reference speeds.

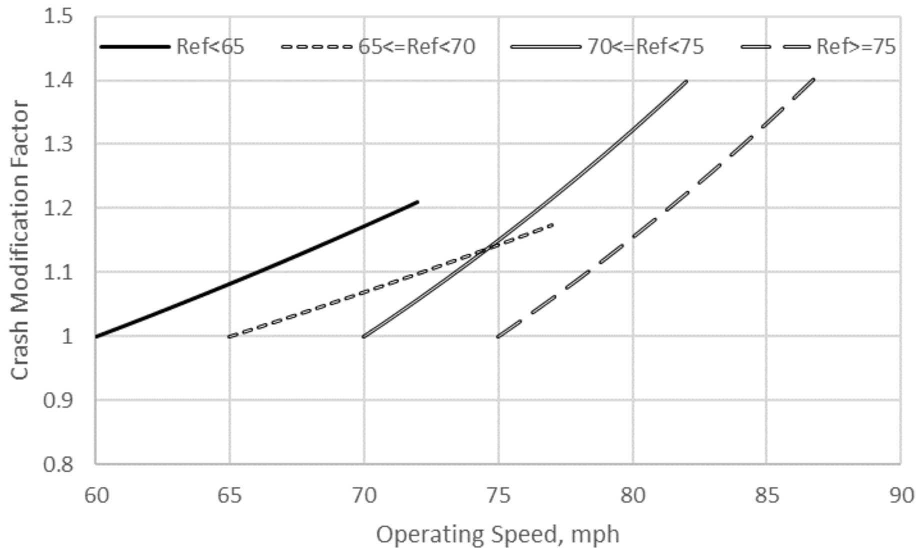

Figure 23. CMF for Operating Speeds (85th Percentile FFS) on Multilane Divided Highways.

4.3.3.7. Speed Variation CMF

The speed variation CMF is described using Equation 24.

$$CMF_{std} = e^{0.0371(SpdStd-3)} \tag{24}$$

The base condition for this CMF is an SD of 3 mph in speed. Since the speed variation CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 24 shows the speed variation CMF. The CMF shows that the speed variation has an effect on the occurrence of rural multilane highway crashes but has relatively less influence when compared to freeways (see Figure 24).

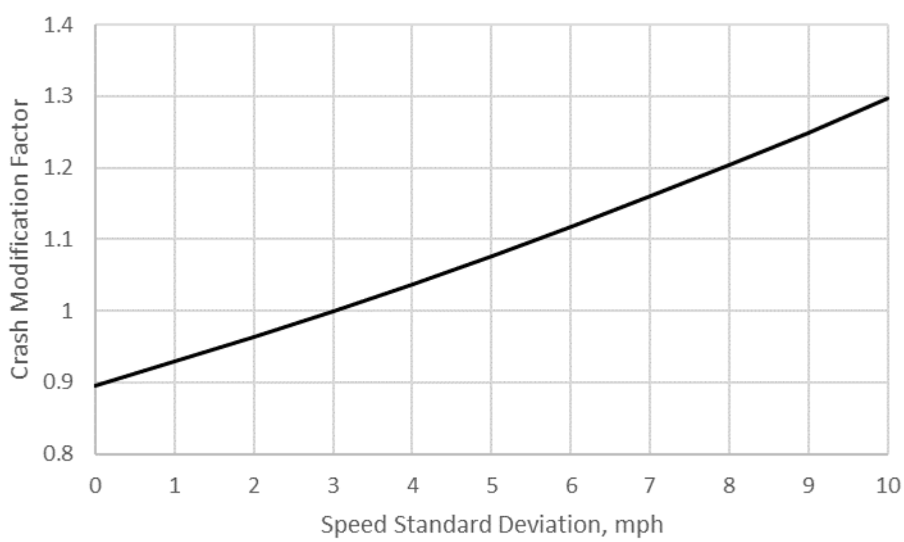

Figure 24. CMF for Speed Variation on Multilane Divided Highways.

4.3.3.8. K-Factor CMF

The K-factor CMF is described using Equation 25.

$$CMF_{kf} = e^{0.0292(kf-10)} \tag{25}$$

The base condition for this CMF is a k-factor of 10. Since the k-factor CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 25 shows the k-factor CMF. The CMF shows that crashes increase with the increase in k-factor although the effect is slightly lower when compared to freeways (which is shown in Figure 25).

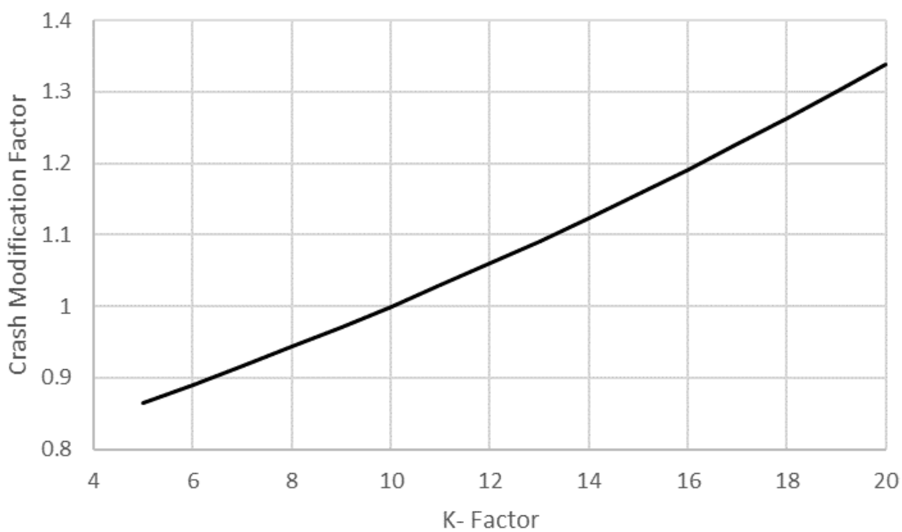

Figure 25. CMF for K-Factor on Multilane Divided Highways.

4.3.4 SPFs for Multilane Undivided Highways

4.3.4.1. Predicted Crash Frequency

For the SPF development, the team just considered the four-lane highways, given the large sample size in that category. The predicted crash frequency was calculated using Equation 26.

$$N = L \times y \times e^{b_0 + b_{aadt}\ln(AADT)} \times CMF_{sw} \times CMF_{spd} \times CMF_{pre} \tag{26}$$

With:

$$CMF_{sw} = e^{b_{sw}(sw-8)}$$

$$CMF_{spd} = \begin{cases} e^{b_{spd}(SpdFF85-60)}, \ if \ RefSpd < 65mph \\ e^{b_{spd}(SpdFF85-65)}, if \ 65mph \ \leq \ RefSpd < 70mph \\ e^{b_{spd}(SpdFF85-70)}, \ if \ 70mph \leq RefSpd < 75mph \\ e^{b_{spd}(SpdFF85-75)}, \ if \ RefSpd \geq 75mph \end{cases}$$

$$CMF_{pre} = e^{b_{pre}(pre-0.003)\times 100}$$

Where:

CMF_{sw} = the CMF for average shoulder width.
CMF_{pre} = the CMF for average daily precipitation.
sw = the average shoulder width, feet.
pre = the average daily precipitation, inches.

Table 24 and Table 25 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 24. Calibrated Coefficients for Fatal and Injury Crashes on Multilane Undivided Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
$\boldsymbol{b_0}$		Intercept	–6.7131	0.511	–13.14	<0.0001
$\boldsymbol{b_{aadt}}$		AADT	0.6528	0.055	11.87	<0.0001
$\boldsymbol{b_{sw}}$		Shoulder width	–0.0275	0.009	–3.13	0.00
$\boldsymbol{b_{spd}}$	$RefSpd < 65mph$	85th percentile FFS	—	—	—	—
	$65mph \leq RefSpd < 70mph$		—	—	—	—
	$70mph \leq RefSpd < 75mph$		0.0123	0.009	1.39	0.16
	$RefSpd \geq 75mph$		0.0123	0.009	1.39	0.16
$\boldsymbol{b_{pre}}$		Precipitation	0.2187	0.133	1.64	0.10
$\boldsymbol{k}$		Inverse dispersion parameter	1.4374	0.127	11.34	<0.0001

Table 25. Calibrated Coefficients for Property Damage Only Crashes on Multilane Undivided Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
$\boldsymbol{b_0}$		Intercept	–6.4512	0.415	–15.54	<0.0001
$\boldsymbol{b_{aadt}}$		AADT	0.7156	0.045	15.96	<0.0001
$\boldsymbol{b_{sw}}$		Shoulder width	–0.0188	0.007	–2.68	0.01
$\boldsymbol{b}$	$RefSpd < 65mph$	85th percentile FFS	—	—	—	—
	$65mph \leq RefSpd < 70mph$		—	—	—	—
	$70mph \leq RefSpd < 75mph$		–0.0133	0.007	–1.85	0.06
	$RefSpd \geq 75mph$		–0.0133	0.007	–1.85	0.06
$\boldsymbol{b_{pre}}$		Precipitation	0.5838	0.104	5.62	<0.0001
$\boldsymbol{k}$		Inverse dispersion parameter	1.5707	0.080	19.68	<0.0001

Figure 26 and Figure 27 compare different calibrated multilane undivided SPFs for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the HSM SPFs and the Texas WB. The equations are plotted for the case of all CMFs equal to 1.0

(representing base conditions). The SPFs do not include the same set of base conditions, so they are not directly comparable to each other. In addition, the HSM SPFs are not calibrated to Texas conditions, and the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since the Texas WB include SPFs for FI crashes only, the comparison is made with just HSM SPFs for the total crashes.

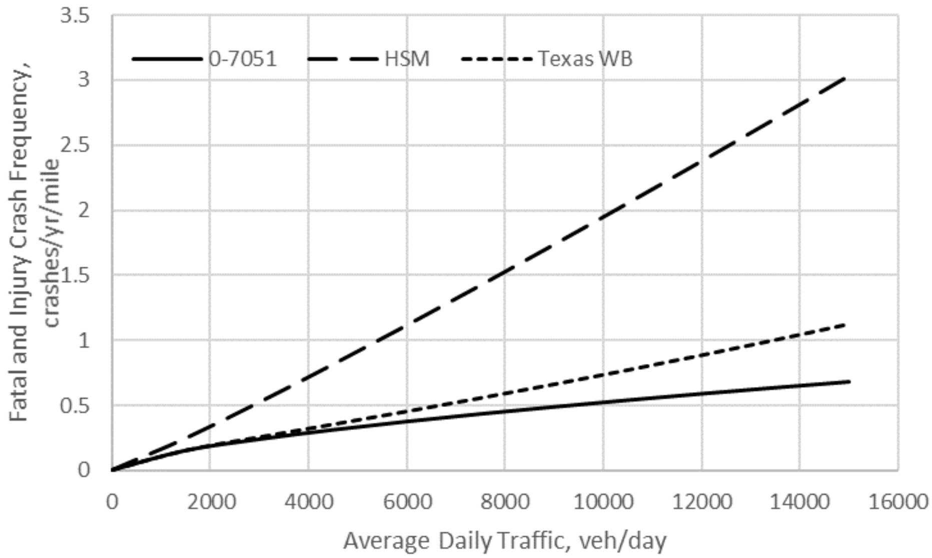

Figure 26. SPF Comparison for Fatal and Injury Crashes on Multilane Undivided Highways.

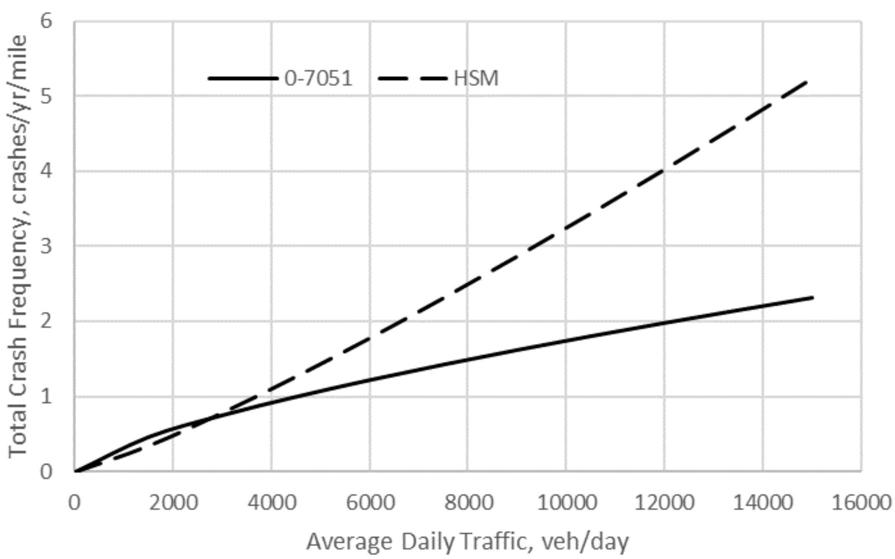

Figure 27. SPF Comparison for Total Crashes on Multilane Undivided Highways.

4.3.4.2. Crash Modification Factors

Several CMFs were calibrated in conjunction with the SPFs. All of them were calibrated using the FI crash data. This section describes these CMFs and, where possible, compares them with the findings from the HSM and the Texas WB as a means of model validation.

4.3.4.3. Shoulder Width CMF

The shoulder width CMF is described using Equation 27.

$$CMF_{sw} = e^{-0.0275(sw-6)} \tag{27}$$

The base condition for this CMF is a shoulder width of 6 feet. The value used in this CMF is an average width for shoulders in both directions. Figure 28 shows the shoulder width CMF developed in this study. Figure 28 also shows the CMF presented in the HSM. The proposed CMF is less sensitive to the shoulder width than the HSM CMF for narrow widths but closely tracks for wider shoulders.

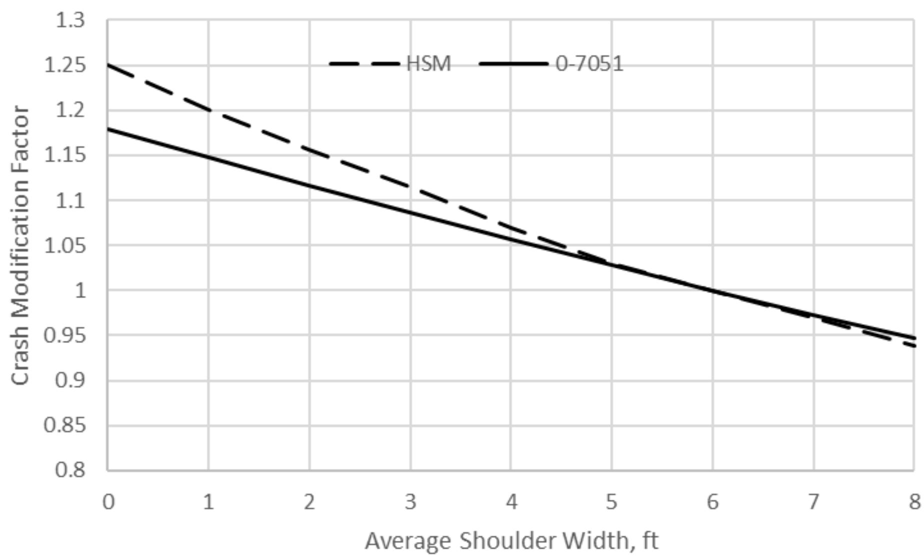

Figure 28. CMF for Shoulder Width on Multilane Undivided Highways.

4.3.4.4. Operating Speed CMF

The operating speed CMF is described using Equation 28.

$$CMF_{spd} = \begin{cases} 1.0, if\ RefSpd < 65mph \\ 1.0, if\ 65mph \le RefSpd < 70mph \\ e^{0.0123(SpdFF85-70)}, if\ 70mph \le RefSpd < 75mph \\ e^{0.0123(SpdFF85-75)} if\ RefSpd \ge 75mph \end{cases} \tag{28}$$

The base condition for this CMF varies according to the reference speed. Since the operating speed CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 29 shows the operating speed CMFs for highways with different reference speeds. A relationship between operating speeds and crashes could not be established for highways with reference speed less than 70 mph. In addition, the relationship was much more similar for highways with reference speed greater than 70 mph and 75 mph, so a single coefficient was used.

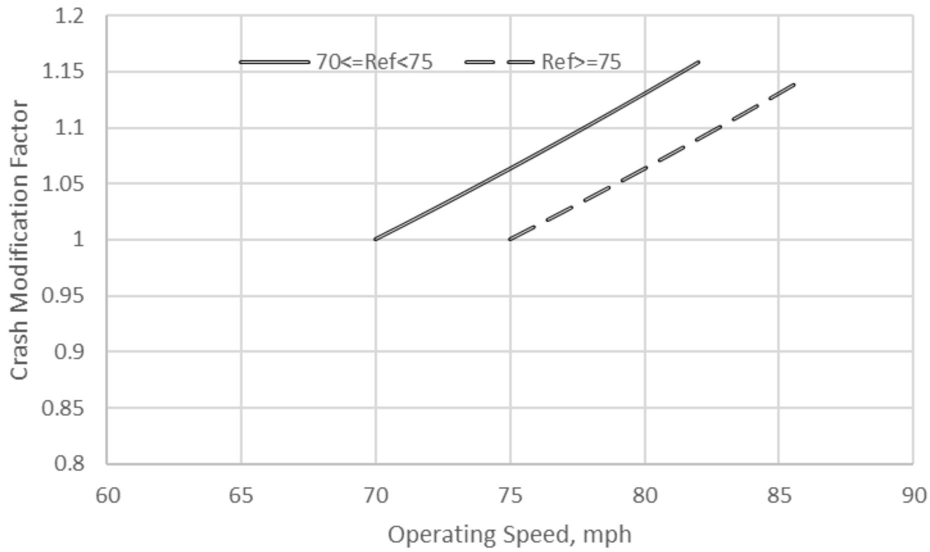

Figure 29. CMF for Operating Speeds (85th Percentile FFS) on Multilane Undivided Highways.

4.3.4.5. Precipitation CMF

The precipitation variable was not found significant in rural freeways and rural multilane divided roadways. The precipitation CMF is described using Equation 29.

$$CMF_{pre} = e^{0.2187(pre-0.003)\times 100} \tag{29}$$

The base condition for this CMF is an average annual precipitation of 0.003 inches per day (this is the average precipitation value in the data considered). Since the precipitation CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 30 shows the precipitation CMF. The CMF shows that there will be an increase in crashes by about 50 percent with an increase in precipitation of 0.01 inches per day. The effect seems to be overestimated, and it is possible that this variable is capturing the effects of additional variables, such as a higher number of horizontal and vertical curves that are more common in East Texas than in West Texas. The precipitation in East Texas is much higher than in West Texas, on average.

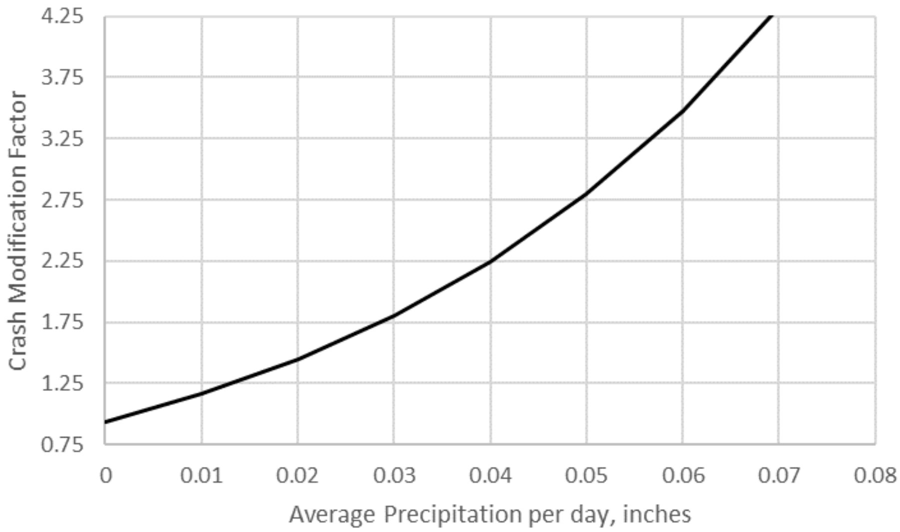

Figure 30. CMF for Precipitation on Multilane Undivided Highways.

4.3.5 SPFs for Two-Lane Highways

4.3.5.1. Predicted Crash Frequency

The predicted crash frequency is calculated using Equation 30.

$$N = L \times y \times e^{b_0 + b_{aadt}\ln(AADT)} \times CMF_{lw} \times CMF_{sw} \times CMF_{spd} \times CMF_{pre} \times CMF_{kf} \qquad (30)$$

With:

$$CMF_{lw} = e^{b_{lw}(lw-12)}$$
$$CMF_{sw} = e^{b_{sw}(sw-8)}$$
$$CMF_{spd} = \begin{cases} e^{b_{spd}(SpdFF85-60)}, if\ RefSpd < 65mph \\ e^{b_{spd}(SpdFF85-65)}, if\ 65mph \le RefSpd < 70mph \\ e^{b_{spd}(SpdFF85-70)},\ if\ 70mph \le RefSpd < 75mph \\ e^{b_{spd}(SpdFF85-75)},\ if\ RefSpd \ge 75mph \end{cases}$$
$$CMF_{pre} = e^{b_{pre}(pre-0.003)\times 100}$$
$$CMF_{kf} = e^{b_{kf}(kf-10)}$$

Where:

CMF_{lw} = the CMF for lane width.
CMF_{kf} = the CMF for K-factor.
lw = the average lane width, feet.
kf = K-factor.

Table 26 and Table 27 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 26. Calibrated Coefficients for Fatal and Injury Crashes on Two-Lane Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
$\boldsymbol{b_0}$		Intercept	–8.2367	0.275	–29.94	<0.0001
$\boldsymbol{b_{aadt}}$		AADT	0.8353	0.033	25.44	<0.0001
$\boldsymbol{b_{lw}}$		Lane width	–0.0408	0.029	–1.42	0.16
$\boldsymbol{b_{sw}}$		Shoulder width	–0.0460	0.011	–4.30	<0.0001
$\boldsymbol{b_s}$	$RefSpd < 65mph$	85th percentile FFS	0.0191	0.008	2.47	0.01
	$65mph \leq RefSpd < 70mph$		—	—	—	—
	$70mph \leq RefSpd < 75mph$		—	—	—	—
	$RefSpd \geq 75mph$		—	—	—	—
$\boldsymbol{b_{pre}}$		Precipitation	0.2106	0.059	3.55	0.00
$\boldsymbol{b_{kf}}$		K-factor	0.0433	0.013	3.21	0.00
$\boldsymbol{k}$		Inverse dispersion parameter	1.6606	0.109	15.19	<0.0001

Table 27. Calibrated Coefficients for Property Damage Only Crashes on Two-Lane Highways.

Coefficient		Variable	Value	SD	t-statistic	p-value
$\boldsymbol{b_0}$		Intercept	–7.6079	0.208	–36.59	<0.0001
$\boldsymbol{b_{aadt}}$		AADT	0.8473	0.025	34.03	<0.0001
$\boldsymbol{b_{lw}}$		Lane width	–0.0642	0.022	–2.89	0.00
$\boldsymbol{b_{sw}}$		Shoulder width	–0.0677	0.008	–8.20	<0.0001
$\boldsymbol{b}$	$RefSpd < 65mph$	85th percentile FFS	0.0142	0.006	2.33	0.02
	$65mph \leq RefSpd < 70mph$		—	—	—	—
	$70mph \leq RefSpd < 75mph$		—	—	—	—
	$RefSpd \geq 75mph$		—	—	—	—
$\boldsymbol{b_{pre}}$		Precipitation	0.1997	0.047	4.26	<0.0001
$\boldsymbol{b_{kf}}$		K-factor	0.0800	0.011	7.62	<0.0001
$\boldsymbol{k}$		Inverse dispersion parameter	1.7530	0.066	26.61	<0.0001

Figure 31 and Figure 32 compare different calibrated two-lane SPFs for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the HSM and Texas WB SPFs. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). The SPFs do not include the same set of base conditions, so they are not directly comparable to each other. In addition, the HSM SPFs are not calibrated to Texas conditions, and the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since the Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

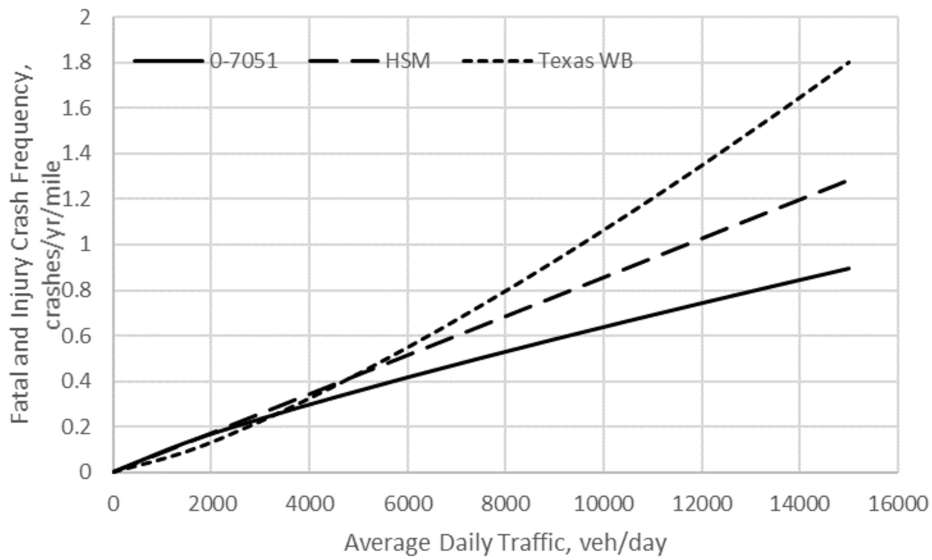

Figure 31. SPF Comparison for Fatal and Injury Crashes on Two-Lane Highways.

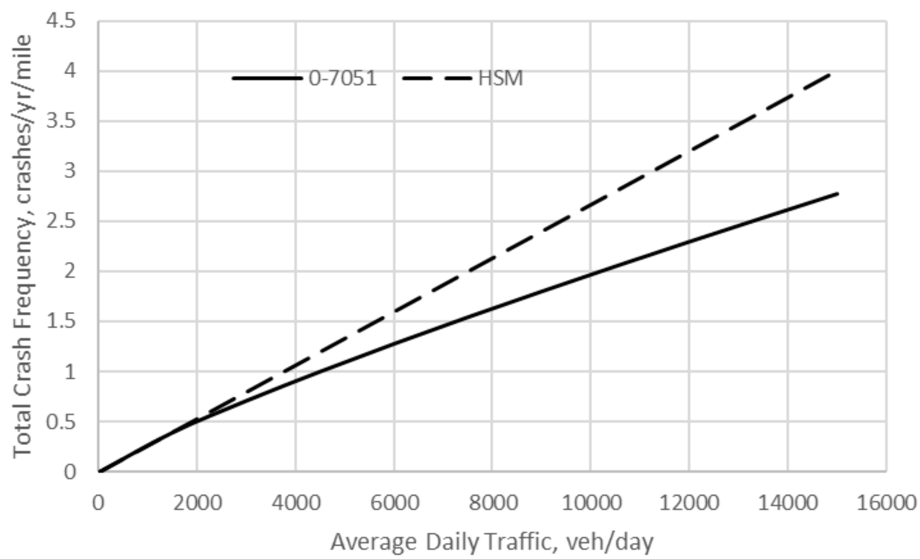

Figure 32. SPF Comparison for Total Crashes on Two-Lane Highways.

4.3.5.2. Crash Modification Factors

Several CMFs were calibrated in conjunction with the SPFs. All of them were calibrated using the FI crash data. This section describes these CMFs and, where possible, compares them with the findings from the HSM and the Texas WB as a means of model validation.

4.3.5.3. Lane Width CMF

The lane width CMF is described using Equation 31.

$$CMF_{lw} = e^{-0.0408(lw-12)} \tag{31}$$

The base condition for this CMF is a lane width of 12 ft. The lane width used in this CMF is an average for lanes in both directions on the segment. Figure 33 shows the lane width CMF developed in this study. The lane widths used to calibrate this CMF range from 9 to 12 feet Figure 33 also shows CMFs presented in the HSM and the Texas WB (considering 8 ft shld or shoulder). The proposed CMF is less sensitive to lane width than the CMFs in the HSM and the Texas WB.

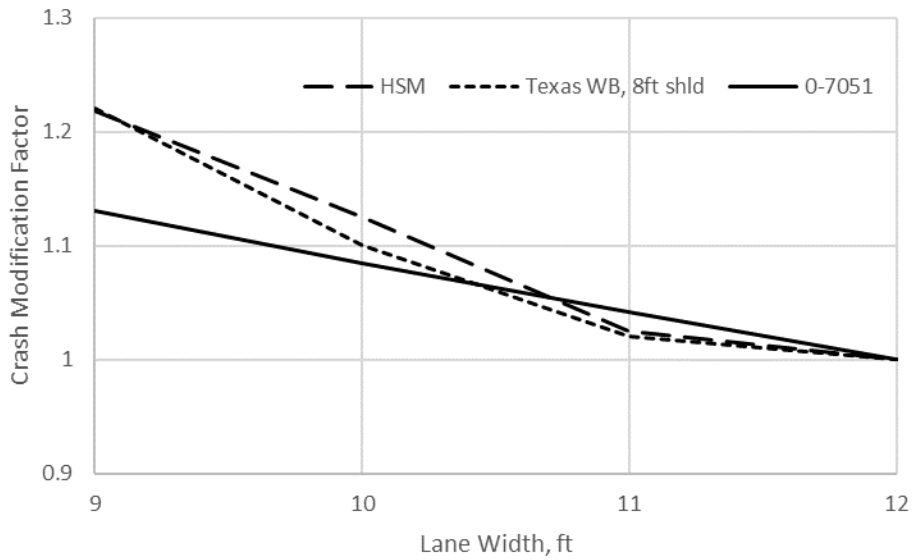

Figure 33. CMF for Lane Width on Two-Lane Highways.

4.3.5.4. Shoulder Width CMF

The shoulder width CMF is described using Equation 32.

$$CMF_{mw} = e^{-0.0057(W_m-15)} \quad (32)$$

The base condition for this CMF is a shoulder width of 6 feet. The value used in this CMF is an average width for shoulders in both directions. Figure 34 shows the shoulder width CMF developed in this study. Figure 34 also shows the CMF presented in the HSM. The proposed CMF is more sensitive to the shoulder width than the HSM CMF.

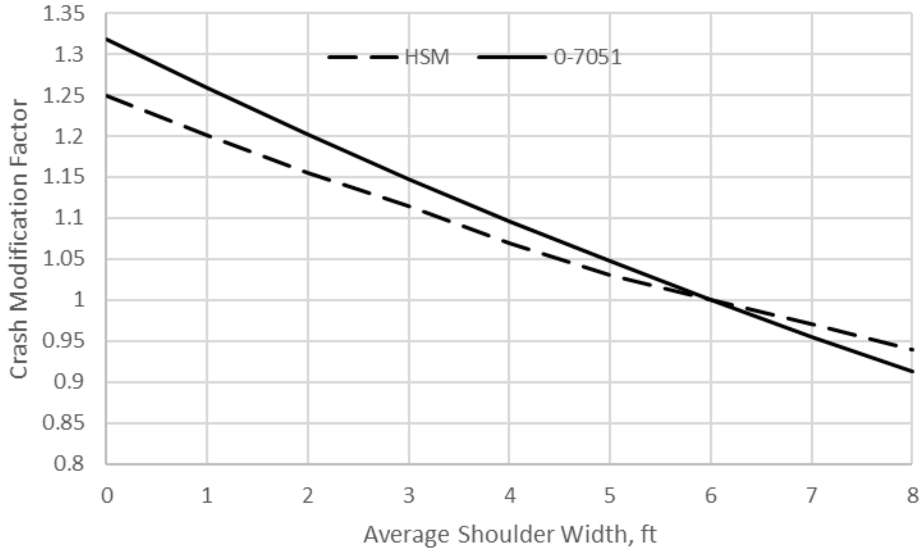

Figure 34. CMF for Shoulder Width on Two-lane Highways.

4.3.5.5. Operating Speed CMF

The operating speed CMF is described using Equation 33.

$$CMF_{spd} = \begin{cases} e^{0.0191(SpdFF85-60)}, if\ RefSpd < 65mph \\ 1.0, if\ 65mph \leq RefSpd < 70mph \\ 1, if\ 70mph \leq RefSpd < 75mph \\ 1, if\ RefSpd \geq 75mph \end{cases} \tag{33}$$

The base condition for this CMF varies according to the reference speed. Since the operating speed CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 35 shows the operating speed CMF for highways with a reference speed less than 65 mph. A relationship between operating speed and crashes could not be established for highways with other reference speeds.

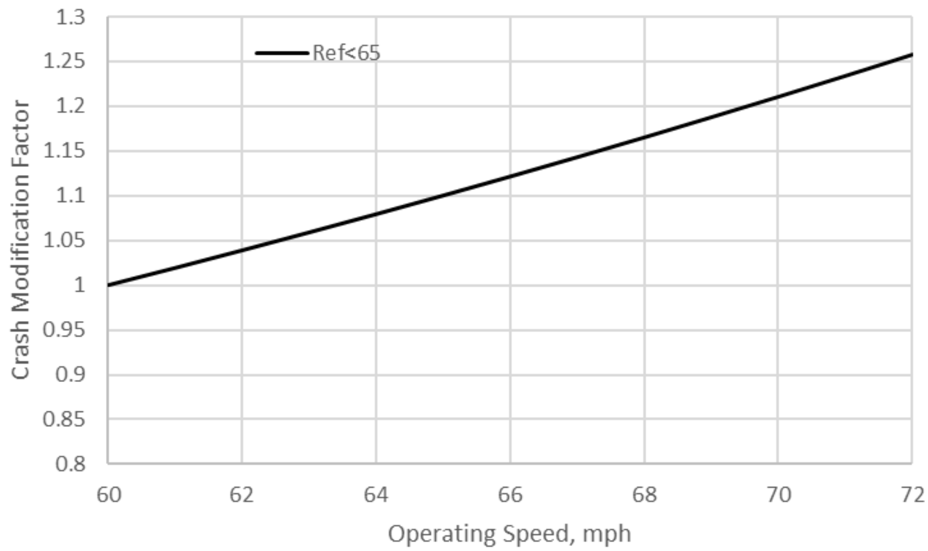

Figure 35. CMF for Operating Speeds (85th Percentile FFS) on Rural Two-Lane Highways.

4.3.5.6. Precipitation CMF

The precipitation CMF is described using Equation 34.

$$CMF_{pre} = e^{0.2106(pre-0.003)\times 100} \tag{34}$$

The base condition for this CMF is an annual average precipitation of 0.003 inches per day. Since the precipitation CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 36 shows the precipitation CMF. The CMF shows that there will be an increase in crashes by about 50 percent with an increase in precipitation of 0.01 inches per day. The effect seems to be overestimated, and it is possible that this variable is capturing the effects of additional variables such as a higher number of horizontal and vertical curves that are more common in East Texas than in West Texas. The precipitation in East Texas is much higher than in West Texas, on average.

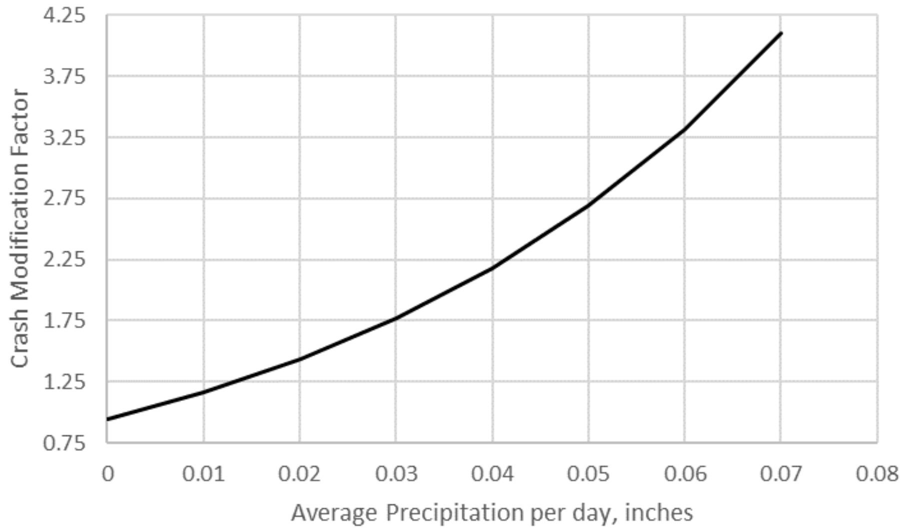

Figure 36. CMF for Speed Variation on Two-Lane Highways.

4.3.5.7. K-Factor CMF

The K-factor CMF is described using Equation 35.

$$CMF_{kf} = e^{0.0433(kf-10)} \tag{35}$$

The base condition for this CMF is a k-factor of 10. Since the k-factor CMF does not exist in the HSM or the Texas WB, a comparison could not be made. Figure 37 shows the k-factor CMF. The CMF shows that crashes increase with an increase in k-factor.

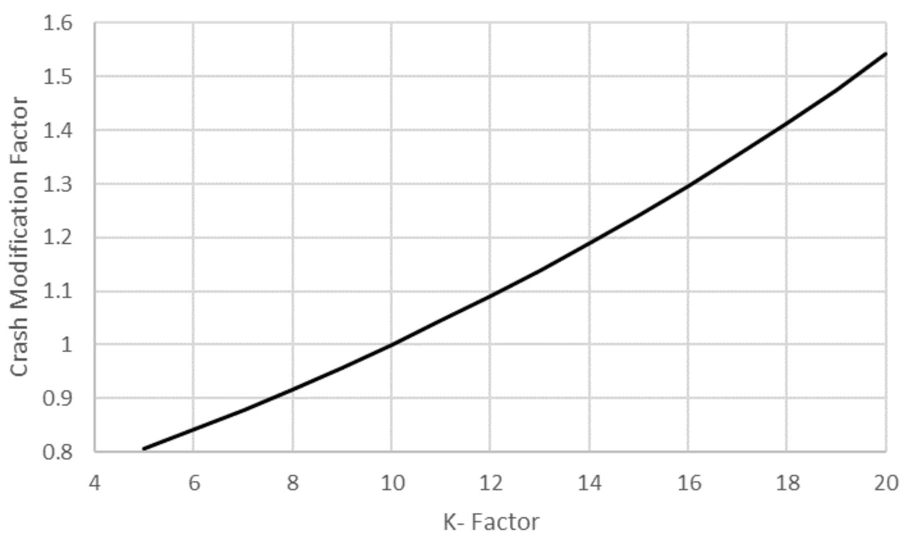

Figure 37. CMF for K-Factor on Two-Lane Highways.

4.3.6 Other Considerations

The truck proportion variable is not included in the rural multilane undivided or two-lane highway SPFs. This variable was statistically significant but counterintuitive. The results suggested that crashes decrease with an increase in truck proportion in the traffic mix. It is widely known that the highways on which trucks travel are usually wider. Figure 38 presents the relationship between surface width and truck proportion on two-lane highways. Highways with higher truck proportions generally contain wider surface width. Thus, the truck proportion is a confounding effect for the surface width variable. The width of a highway is generally positively correlated to traffic safety.

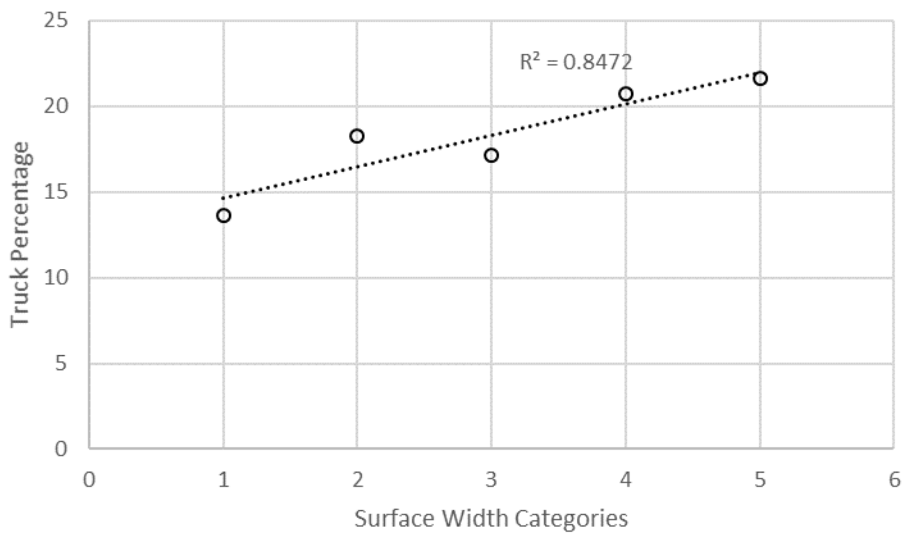

Figure 38. Relationship of Surface Width and Truck Proportion.

4.3.7 Daily-Level Model

Unlike annual crash frequency, the daily crash frequencies per segment are mostly zeros or ones. Thus, for daily crashes, a crash occurrence (which has a value 1 if a crash occurred and 0 if no crash occurred, rather than a crash count) at each segment was modeled. A logistic regression model was adopted for the daily crash data, which express the log-odds of the probability of a crash occurrence as a function of speed variables and other roadway characteristic variables, given in Equation 36.

$$g(\mathbf{x}) = \ln\left[\frac{P(Y_{it}=1|\mathbf{x})}{1-P(Y_{it}=1|\mathbf{x})}\right] = \beta_0 + \beta_1 x_{1,it} + \cdots + \beta_K x_{K,it} \qquad (36)$$

Where:

$g(x)$ = the logit (log-odds).

Y_{it} = a crash occurrence on segment *i* ($i = 1, \cdots, I$) and day *t*.

$\mathbf{x}$ = a value of predictor variables $X_{1it}, \cdots, X_{K,it}$.

$\beta_0, \beta_1, \cdots, \beta_K$ = regression coefficients for daily crashes.

The underlying assumption for this model is that the relationship between the logit, $g(x)$, and predictor variables is linear. The intercept β_0 represents the baseline level of the logit, and β_k represents the change in the logit that occurs with a unit change in X_k. The conditional probability that a crash occurs at site *i* in time *t* can be expressed as Equation 37.

$$P\left(Y_{it} = 1 | \mathbf{x}\right) = \frac{e^{g(x)}}{1 + e^{g(x)}} = \frac{e^{\beta_0 + \beta_1 x_{1,it} + \cdots + \beta_K x_{K,it}}}{1 + e^{\beta_0 + \beta_1 x_{1,it} + \cdots + \beta_K x_{K,it}}} \tag{37}$$

TTI constructed the daily crash database for logistic regression modeling using a case-control study design, which matched every observation with a total crash occurrence (case) with a no-crash observation (control) from the same segment or a segment with similar roadway characteristics (1-1 matching). The resulting data sets for rural two-lane roadways, rural freeways, rural multilane divided roadways, and rural multilane undivided roadways consist of 23,450 observations of rural two-lane roadways obtained from 3,557 segments (corresponding to 2355.0 miles), 61,554 observations of rural freeways obtained from 2,480 segments (corresponding to 1992.2 miles), 47,954 observations of rural multilane divided roadways obtained from 4,147 segments (corresponding to 2811.4 miles), and 25,768 observations of rural multilane undivided roadways obtained from 2,853 segments (corresponding to 1236.5 miles), respectively, for 2017–2019.

TTI analyzed both total and FI daily crashes. The following variables were included as predictors in the logistic regression model for the Texas daily crash data:

- Average operating speed (SpdAve).
- SD of operating speed (SpdStd).
- Daily precipitation (daily_precip).
- Log of AADT (LnAADT).
- Truck AADT percentage (TRK_AADT_P).
- Segment length (Length).
- Shoulder width inside (SW_I).
- Shoulder width outside (SW_O).
- Average shoulder width (SW_Avg).
- Lane Width.
- Median Width.

For daily crash data, some variables (e.g., SpdAve, SpdStd, and daily_precip) have daily variability at each segment, some (e.g., AADT and TRK_AADT_P) have only yearly variability, and some (e.g., segment length, shoulder widths, and lane width) do not have any temporal variability and have only spatial variability (i.e., segment-to-segment variability). To account for potential correlation in the outcomes obtained for multiple days from the same segment of daily crash data in the estimation, a generalized estimating equation (GEE) approach was used to estimate model parameters in Equation 36.

TTI explored several logistic regression models for daily crash occurrence with including SpdAve, SpdStd, daily_precip, and other roadway characteristic variables such as AADT, TRK_AADT_P, Length, SW_I, SW_O, or SW_Avg, Lane Width, and/or Median Width as predictor variables for each of four roadway types (rural two-lane roadways, rural freeways, rural multilane undivided roadways, and rural multilane divided roadways). Only four-lane roadways are included in the data sets for rural freeways, rural multilane undivided roadways, and rural multilane divided roadways, so the number of lane variables is not included as a predictor. Predictor variables are retained in the final model as long as p-values for the estimated coefficients are less than 0.2 and the signs of coefficients are physically meaningful.

The estimated regression coefficients for final selected models and the corresponding odds ratio (OR) estimates are given in Table 28 for rural freeways, Table 29 for rural multilane divided highways, Table 30 for rural multilane undivided highways, and Table 31 for rural two-lane highways. The OR for each predictor variable can be estimated by exponentiating the corresponding regression coefficient. The OR is expressed as the expected increase or decrease in the crash risk due to a unit increase in the predictor variable. The OR can serve as a direct estimate of the CMF. An OR greater than 1.0 suggests that an increase in the predictor variable increases crash risk, while a value less than 1.0 suggests a decrease in crash risk.

For rural freeway highways in Table 28, daily_precip, SpdAve, SpdStd, TRK_AADT_P, and Length have statistically significant (at $\alpha = 0.05$) positive effects on the odds of total crash occurrence. SpdAve, SpdStd, TRK_AADT_P, and Length have statistically significant positive effects on the odds of FI crash occurrence. The OR is large for daily_precip (1.6153) for total crashes (i.e., a total crash is 1.6153 times as likely to occur as daily precipitation increases by 1 inch) and for SpdStd (1.4113 for total crashes and 1.2752 for FI crashes).

Table 28. Modeling Results of Daily-Level Analysis for Rural Freeways.

Variables	Total Crashes		FI Crashes	
	Estimates	Standard Error	Estimates	Standard Error
Outcome Variable: Crash Occurrence (1: Crash, 0: No Crash)				
Intercept	0.2311	0.4595	**–8.4173**	0.5334
daily_precip	**0.4795**	0.0716	*0.0283*	0.0149
SpdAve	**–0.0275**	0.0069	**0.0361**	0.0063
SpdStd	**0.3445**	0.0111	**0.2431**	0.0098
TRK_AADT_P	**0.0107**	0.0011	**0.0086**	0.0016
Length	**0.1306**	0.0220	—	—
S_WID_I	—	—	**–0.0105**	0.0047
RefSpd	—	—	**0.0431**	0.0081
Odds Ratios: Crash Counts				
daily_precip	**1.6153**	0.1157	*1.0287*	0.0153
SpdAve	**0.9728**	0.0067	**1.0368**	0.0065
SpdStd	**1.4113**	0.0157	**1.2752**	0.0125
TRK_AADT_P	**1.0107**	0.0011	**1.0087**	0.0016
Length	**1.1395**	0.0250	—	—
S_WID_I	—	—	**0.9896**	0.0047
RefSpd	—	—	**1.0441**	0.0085
Goodness of Fit Statistics				
QIC	66,453.7866		39,119.0298	
QICu	66,431.2078		39,114.6848	

Notes:
1. The GEE approach was used as an estimation method to account for possible correlation in the observations obtained from the same segment.
2. The coefficient "0" denotes that the corresponding variable was excluded from the model.
3. LnAADT = Log(AADT).
4. OR estimates are obtained by Exp(β_k) where β_k represents the estimated coefficient of the *k*th variable.
5. Statistically significant results at **95%** (*90%*) confidence level are shown in **bold** (*italic*).

Table 29 shows that for rural multilane divided roadways, daily_precip, SpdAve, SpdStd, LnAADT, TRK_AADT_P, and Length have statistically significant (at $\alpha = 0.05$) positive effects on the odds of total crash occurrence, and SpdAve, SpdStd, LnAADT, and Length have statistically significant positive effects on the odds of KABC crash occurrence, while S_WID_I has a statistically significant negative effect on the odds of FI crash occurrence. The OR is large for daily_precip for total crashes, SpdStd for total crashes, and SpdStd for FI crashes, which are 1.3421, 1.1459, and 1.1814, respectively.

Table 29. Modeling Results of Daily-Level Analysis for Rural Multilane Divided Roadways.

Variables	Total Crashes		FI Crashes	
	Estimates	Standard Error	Estimates	Standard Error
Outcome Variable: Crash Occurrence (1: Crash, 0: No Crash)				
Intercept	**–2.4032**	0.2992	**–4.1043**	0.5769
daily_precip	**0.2942**	0.0432	—	—
SpdAve	**0.0094**	0.0019	**0.0210**	0.0055
SpdStd	**0.1362**	0.0115	**0.1675**	0.0159
LnAADT	**0.1004**	0.0153	**0.0458**	0.0286
TRK_AADT_P	**0.0086**	0.0009	—	—
Length	**0.0376**	0.0085	**0.0611**	0.0253
S_WID_I	–0.0038	0.0023	**–0.0154**	0.0048
Odds Ratios: Crash Counts				
daily_precip	**1.3421**	0.0579	—	—
SpdAve	**1.0095**	0.0020	**1.0212**	0.0056
SpdStd	**1.1459**	0.0132	**1.1824**	0.0188
LnAADT	**1.1057**	0.0169	*1.0469*	0.0300
TRK_AADT_P	**1.0086**	0.0009	—	—
Length	**1.0384**	0.0088	**1.0630**	0.0269
S_WID_I	0.9962	0.0023	**0.9847**	0.0048
Goodness of Fit Statistics				
QIC	45,866.8577		33,652.0673	
QICu	45,868.4580		33,634.7837	

Notes:
1. The GEE approach was used as an estimation method to account for possible correlation in the observations obtained from the same segment.
2. The coefficient "0" denotes that the corresponding variable was excluded from the model.
3. LnAADT = Log(AADT).
4. OR estimates are obtained by Exp(β_k) where β_k represents the estimated coefficient of the *k*th variable.
5. Statistically significant results at **95%** (*90%*) confidence level are shown in **bold** (*italic*).

For rural multilane undivided roadways in Table 30, daily_precip, SpdStd, TRK_AADT_P, and Length have statistically significant (at α = 0.05) positive effects on the odds of total crash occurrence, while SpdAve, SpdStd, TRK_AADT_P, and Length have statistically significant positive effects on the odds of FI crash occurrence. The OR is 1.2441 for daily_precip for total crashes, 1.1188 for SpdStd for total crashes, and 1.1770 for SpdStd for FI crashes.

Table 30. Modeling Results of Daily-Level Analysis for Rural Multilane Undivided Roadways.

Variables	Total Crashes		FI Crashes	
	Estimates	Standard Error	Estimates	Standard Error
Outcome Variable: Crash Occurrence (1: Crash, 0: No Crash)				
Intercept	**–0.8021**	0.0612	**–3.4289**	0.2207
daily_precip	**0.2184**	0.0690	—	—
SpdAve	—	—	**0.0110**	0.0031
SpdStd	**0.1122**	0.0095	**0.1630**	0.0169
TRK_AADT_P	**0.0052**	0.0008	**0.0069**	0.0032
Length	**0.0816**	0.0118	**0.2307**	0.0465
Odds Ratios: Crash Counts				
daily_precip	**1.2441**	0.0859	—	—
SpdAve	—	—	**1.0110**	0.0031
SpdStd	**1.1188**	0.0106	**1.1770**	0.0199
TRK_AADT_P	**1.0052**	0.0008	**1.0069**	0.0032
Length	**1.0851**	0.0128	**1.2595**	0.0585
Performance Measures				
QIC	18,667.7756		13,311.0957	
QICu	*18,670.7172*		*13,305.3697*	

Notes:
1. The GEE approach was used as an estimation method to account for possible correlation in the observations obtained from the same segment.
2. The coefficient "0" denotes that the corresponding variable was excluded from the model.
3. LnAADT = Log(AADT).
4. OR estimates are obtained by Exp(β_k) where β_k represents the estimated coefficient of the *k*th variable.
5. Statistically significant results at **95%** (*90%*) confidence level are shown in **bold** (*italic*).

Table 31 shows that for rural two-lane roadways, SpdStd, LnAADT, TRK_AADT_P, and Length have statistically significant (at $\alpha = 0.05$) positive effects on the odds of total crash occurrence, while SpdAve, SpdStd, LnAADT, and Length have statistically significant positive effects on the odds of FI crash occurrence. The OR for SpdStd is 1.1780 for FI crashes (i.e., a FI crash is 1.1780 times as likely to occur as SpdStd increases by 1 mph) and is 1.1149 for total crashes, which indicates a larger effect size of SpdStd for FI crashes (as expected) compared to total crashes.

Table 31. Modeling Results of Daily-Level Analysis for Rural Two-Lane Roadways.

Variables	Total Crashes		FI Crashes	
	Estimates	Standard Error	Estimates	Standard Error
Outcome Variable: Crash Occurrence (1: Crash, 0: No Crash)				
Intercept	**–1.1344**	0.1591	**–4.7582**	0.4587
daily_precip	0.1273	0.0916	—	—
SpdAve	0.0014	0.0010	**0.0171**	0.0030
SpdStd	**0.1087**	0.0071	**0.1638**	0.0111
LnAADT	**0.0337**	0.0105	**0.1269**	0.0365
TRK_AADT_P	**0.0037**	0.0005	—	—
Length	**0.0341**	0.0089	**0.1137**	0.0328
Odds Ratios: Crash Counts				
daily_precip	1.1358	0.1041	—	—
SpdAve	1.0014	0.0010	**1.0172**	0.0030
SpdStd	**1.1149**	0.0079	**1.1780**	0.0131
LnAADT	**1.0342**	0.0108	**1.1353**	0.0414
TRK_AADT_P	**1.0037**	0.0005	—	—
Length	**1.0347**	0.0092	**1.1204**	0.0367
Goodness of Fit Statistics				
QIC	27,380.2441		20,807.2303	
QICu	27,379.6606		20,804.3426	

Notes:
1. The GEE approach was used as an estimation method to account for possible correlation in the observations obtained from the same segment.
2. The coefficient "0" denotes that the corresponding variable was excluded from the model.
3. LnAADT = Log(AADT).
4. OR estimates are obtained by Exp(β_k) where β_k represents the estimated coefficient of the *k*th variable.
5. Statistically significant results at **95%** (*90%*) confidence level are shown in **bold** (*italic*).

4.4 KEY FINDINGS

4.4.1 Findings from Annual-Level Analysis

TTI developed SPFs using aggregated annual data for FI crashes and PDO crashes separately, as well as for both states together. The general findings are:

- Geometric variables such as lane width, shoulder width, and median width show an association with crash counts, and the association meets the conventional expectations.
- Truck proportion shows a positive association with FI crashes on rural freeways and rural multilane divided roadways. This variable is not statistically significant for rural multilane undivided roadways and rural two-lane roadways. One possible explanation is that the truck proportion is usually less in the latter two roadway facility types.
- K-factor shows a positive association with crashes on rural facilities. For rural multilane undivided roadways, this factor is not statistically significant.
- Increased variability in operational speed is associated with increased crashes for all four roadway facilities.

- In the absence of a posted speed limit, this study used reference speed (see Table 1) as a surrogate. For rural freeways and rural multilane roadways, with the increase in operating speeds, the relative increase in crashes is greater on highways with higher reference speeds than highways with lower reference speeds.
- Average precipitation shows a positive association for only rural multilane undivided and rural two-lane roadways. This variable is not statistically significant for rural freeways and rural multilane divided roadways.

4.4.2 Findings from Daily-Level Analysis

Crash prediction models using annual level or multi-year average data have limitations such as these models do not take into account the effects of variables such as operating speeds, operating speed variance, or seasonal differences, which can fluctuate over the temporal duration. TxDOT requires the need in assessing seasonal or daily changes of operational variables on safety impact. Thus, daily-level models were developed in this study by considering different roadway facilities and crash injury types. The general findings from the modeling that used daily data are:

- Geometric variables such as length and shoulder width show an association with the potential of day-level crash occurrences, and the association meets the conventional expectations.
- Truck proportion shows a positive association with the potential of day-level crash occurrences for all facilities.
- Increased variability in daily operational speed (the SD of the daily average of 5-minute interval operation speeds) is associated with increased day-level crashes for all four roadway facilities.
- Average operating speed (the average of the daily average of 5-minute interval operation speeds) increases are associated with the potential of day-level crash occurrences. However, average operating speed is negatively associated with the potential of total crash occurrences on rural freeways. This finding for rural freeways could be because of the high design standards for the freeways. This variable is not statistically significant for rural multilane undivided roadways.
- As daily average precipitation increases, the potential of a day-level crash occurrence increases for all facility types.

CHAPTER 5:
DECISION SUPPORT TOOL

5.1 INTRODUCTION

This chapter presents a brief overview of the developed decision support tool. TTI developed a GIS-based prototype decision support tool that can estimate and visually illustrate the expected number of annual crashes on the roadway network. Segments with a high number of expected crashes have the highest potential for improvement.

5.2 DECISION SUPPORT TOOL

TTI used the open-source software platform Shiny[8] to develop the decision support tool. A dedicated weblink (https://ruralspeedsafety.shinyapps.io/0_7051Tool_v2/) has been developed so that users can use the tool in any browser with internet access.

This chapter is a software manual that provides guidance on the use of the interactive tool developed for this project.

5.2.1 Interface

Figure 39 shows the interface of the opening page for the decision support tool. This page includes a brief introduction to this project, the components of the tool, and the basic steps of using the tool.

The web interface has two tabs:

- Introduction (the interface shown in Figure 39).
- 0-7051 Tool (users can go to this page by clicking on the tab).

[8] More information about Shiny is available at https://shiny.rstudio.com/.

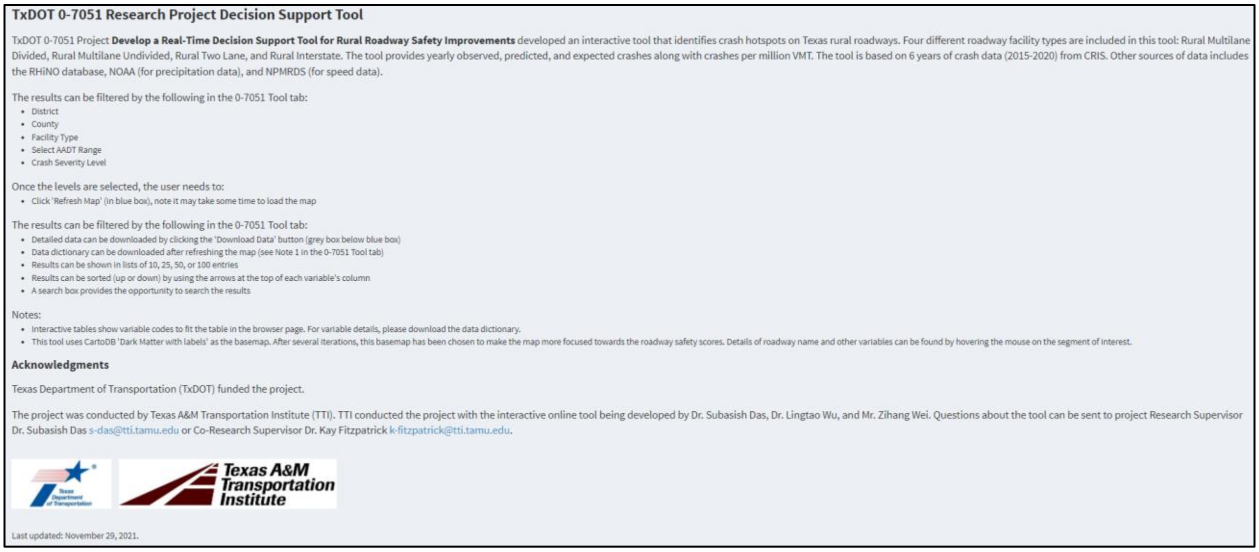

Figure 39. Interface of the 0-7051 Decision Support Tool.

5.2.2 Tool Description

Figure 40 shows the interface of the decision support tool. This page contains two components in the top panel: the map (on the left side) and the drop-down selection panel (on the right side). After the user selects the filters and clicks the Refresh Map button, an interactive table will appear below the top panel.

The top panel has the following features:

- **Filtering option selection:** several drop-down panels (District, County, Facility, AADT Ranges, and Crash Severity).
- **Plot:** Refresh Map button under the drop-down panels.
- **Data download:** Download Data button to download data after the filters are selected.
- **Removal of very short segments**: check box for "Remove very short segments (segment length less than 0.1 mi)." By checking this box, users can remove very short segments in the map visualization.
- **Note:** two notes providing instruction on the data dictionary and interactive table.
- **Zoom in/out of the map:** plus/minus button on the top left side of the map.
- **Popup information in the map:** hovering on a segment to see the information of the segment (this will show up after selection of the filters and map refreshing).

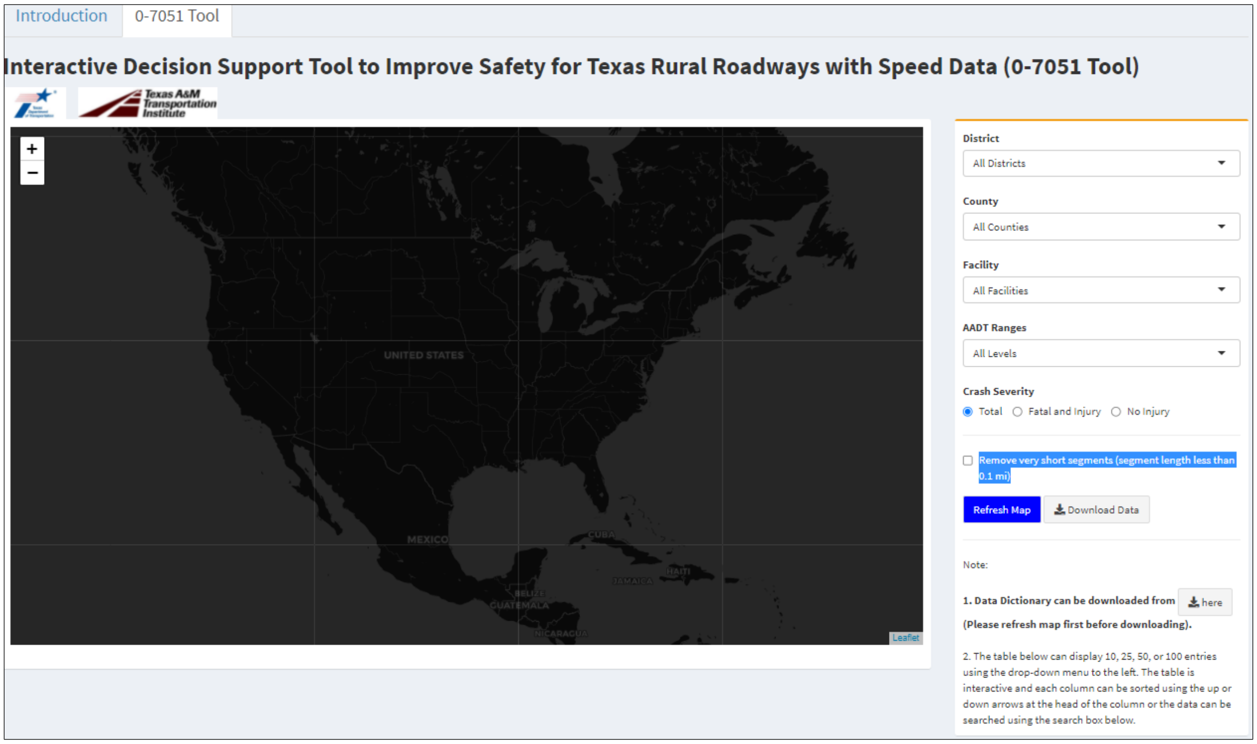

Figure 40. Decision Support Tool.

The Introduction tab has two notes:

- "Interactive tables show variable codes to fit the table in the browser page. For variable details, please download the data dictionary."
- "This tool uses CartoDB 'Dark Matter with labels' as the base map. After several iterations, this base map has been chosen to make the map more focused towards the roadway safety scores. Details of roadway name and other variables can be found by hovering the mouse on the segment of interest."

5.2.3 Map Generation Steps

Four different roadway facility types are included in this tool: rural multilane divided roadways, rural multilane undivided roadways, rural two-lane roadways, and rural freeways. The tool provides yearly observed, predicted, and expected crashes along with crashes per million vehicle miles traveled (VMT). The tool is based on 6 years of crash data (2015–2020) from the CRIS. Other sources of data include roadway inventory data from the RHiNO database, weather data from ASOS stations from NOAA, and operating speed data from the NPMRDS.

The results can be filtered by the following in the 0-7051 Tool tab:

- District (TxDOT districts).
- County (counties in Texas).

- Facility Type (all facilities, rural two-lane roadways, rural multilane undivided roadways, rural multilane divided roadways, and rural freeways).
- AADT Ranges (less than 2000 vpd, 2001–10,000 vpd, and greater than 10,000 vpd).
- Crash Severity Level (total, fatal and injury, and no injury).

Once the levels are selected, the user needs to click Refresh Map (in the blue box; it may take some time to load the map).

The results can be filtered by the following in the 0-7051 Tool tab:

- Detailed data can be downloaded by clicking the Download Data button (the gray box below the blue box).
- The data dictionary (see Table 32) can be downloaded after refreshing the map (see Note 1 in the 0-7051 Tool tab).
- Results can be shown in lists of 10, 25, 50, or 100 entries.
- Results can be sorted (up or down) by using the arrows at the top of each variable's column.
- A search box provides the opportunity to search the results.

Figure 41 shows a screenshot of the interface after selecting "All" from the four drop-down panels. The map shows entire rural roadway networks on the NHS. The color of the segments is based on the number of expected crashes on the individual segment. Below the interactive map, an interactive table produces the result of the final selection. The column names and associated descriptions can be downloaded by clicking the Data Dictionary button in Note 1. The interactive table can display 10, 25, 50, or 100 entries using the drop-down menu to the left. Each column in the table can be sorted using the up or down arrows at the head of the column, or the data can be searched using the search box.

Table 32. Data Dictionary.

Column Name	Source	Original Source Name	Definition	Unit	Data Years
ID	TTI generated	—	—	—	—
District_Code	RHiNO	DI	Code of TxDOT district	—	2018
District	RHiNO	DISTRICT	District name	—	2018
County_Code	RHiNO	CO	Code of county	—	2018
County	RHiNO	COUNTY	County name	—	2018
FaciTy	TTI generated (Source: CRIS)	—	Facility type	—	2018
Length	RHiNO	LEN_SEC	Segment length (re-segmented when segment length > 2 miles)	mi	2018

Column Name	Source	Original Source Name	Definition	Unit	Data Years
ADT	RHiNO	ADT_CUR	Current year AADT	vehicle per day (vpd)	2018
TrkPer	RHiNO	TRK_AADT_PCT	Percent of trucks in AADT	percentage	2018
ADTYear	RHiNO	ADT_YEAR	Year of current year AADT	—	2018
ControlS	RHiNO	C_SEC	Control + section with hyphen	—	2018
FrmDFO	RHiNO	FRM_DFO	From distance from origin (DFO)	—	2018
ToDFO	RHiNO	TO_DFO	To DFO	—	2018
Hwy	RHiNO	HWY	Highway system + highway number + highway suffix	—	2018
KFac	RHiNO	K_FAC	Peak factor	—	2018
MedType	RHiNO	MED_TYPE	0 = No median 2 = Unprotected 3 = Curbed 4 = Positive barrier, unspecified 5 = Positive barrier, flexible 6 = Positive barrier, semi-rigid 7 = Positive barrier, rigid 99 = Unknown	—	2018
MedWid	RHiNO	MED_WID	Median width (does not include inside shoulder widths)	ft	2018
NumLane	RHiNO	NUM_LANES	Number of through lanes (does not include turning, climbing, or auxiliary lanes but does include Super 2 and exclusive high-occupancy vehicle/high-occupancy toll lanes)	—	2018
LaneWid	RHiNO	LANE_WIDTH	Lane width	ft	2018
CLMBLN	RHiNO	CLMB_PS_LANE	Climbing/passing center-turning lane (1 = continuous two-way left-turn lane 2 = Super 2 lane 3 = climbing/passing lane)	—	2018

Column Name	Source	Original Source Name	Definition	Unit	Data Years
SHDWidI	RHiNO	S_WID_I	Shoulder width inside	ft	2018
SHDWidO	RHiNO	S_WID_O	Shoulder width outside	ft	2018
MaxPSL	RHiNO	SPD_MAX	Speed limit maximum	mph	2018
SpdAve	TTI generated (source: NPMRDS)	—	Average operating speed for all vehicles from 5-minute interval NPMRDS data	mph	2017–2020
SpdStd	TTI generated (source: NPMRDS)	—	SD of operating speed for all vehicles from 5-minute interval NPMRDS data	mph	2017–2020
RefSpd	TTI generated (source: NPMRDS)	—	Average reference speed[3] for all vehicles from 5-minute interval NPMRDS data	mph	2017–2020
Spd85	TTI generated (source: NPMRDS)	—	85th percentile operating speed for all vehicles from 5-minute interval NPMRDS data	mph	2017–2020
Precip	TTI generated (source: NOAA ASOS)	—	Average daily precipitation from NOAA ASOS 1-hour precipitation data	in	2017–2020
FICrO	TTI generated (source: CRIS)	—	Yearly observed fatal and injury crashes	crashes/ 1 yr	2015–2020
PDOCrO	TTI generated (source: CRIS)	—	Yearly observed PDO (no injury) crashes	crashes/ 1 yr	2015–2020
TotalCrO	TTI generated (source: CRIS)	—	Yearly observed total crashes	crashes/ 1 yr	2015–2020
FIPre	TTI generated (source: CRIS)	—	Yearly predicted fatal and injury crashes	Crashes/ 1 yr	2015–2020
PDOPre	TTI generated (source: CRIS)	—	Yearly predicted PDO (no injury) crashes	crashes/ 1 yr	2015–2020
TotalPre	TTI generated (source: CRIS)	—	Yearly predicted total crashes	crashes/ 1 yr	2015–2020
FICrE	TTI generated (source: CRIS)	—	Yearly expected fatal and injury crashes	crashes/ 1 yr	2015–2020
PDOCrE	TTI generated (source: CRIS)	—	Yearly expected PDO (no injury) crashes	crashes/ 1 yr	2015–2020

Column Name	Source	Original Source Name	Definition	Unit	Data Years
TotalCrE	TTI generated (source: CRIS)	—	Yearly expected total crashes	crashes/ 1 yr	2015–2020
FICrR	TTI generated (source: CRIS, RHiNO)	—	Yearly fatal and injury yearly crash rate	crashes per million VMT/1 yr	2015–2020
PDOCrR	TTI generated (source: CRIS, RHiNO)	—	Yearly PDO (no injury) crash rate	crashes per million VMT/1 yr	2015–2020
TotalCrR	TTI generated (source: CRIS, RHiNO)	—	Yearly total crash rate	crashes per million VMT/1 yr	2015–2020
TMC_RD	TTI generated (source: CRIS)	—	TMC presence	1	2015–2020

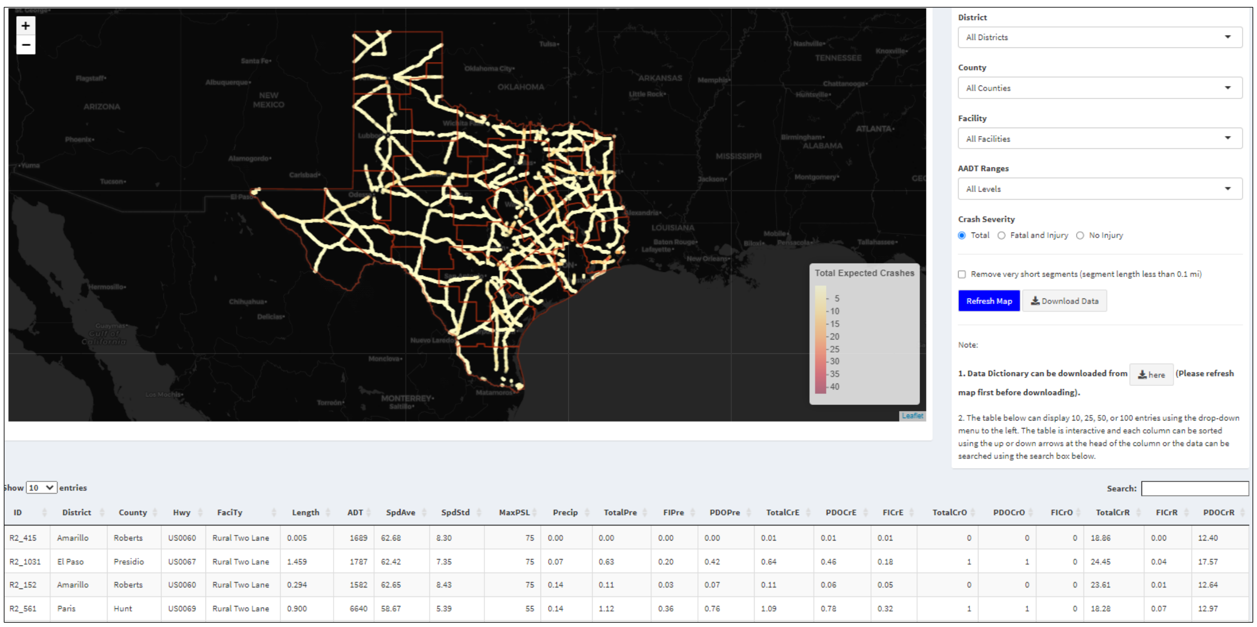

Figure 41. Interface of Tool after Selecting "All" from the Four Drop-Down Panels.

After the generation of the map, the user can hover on a segment to see the segment-specific information (see Figure 42).

Figure 42. Screenshot Showing Hovering Option.

5.2.4 Tool Usage Example

A safety engineer from the Wichita Falls District wants to explore the tool to understand the safety condition of low-volume rural two-lane roadways.

The safety engineer needs to select the following options:

- **District:** Wichita Falls.
- **County:** All Counties.
- **Facility:** Rural Two-Lane.
- **AADT Ranges:** Less than 2000.

The user can select the Crash Severity option as needed. For example, if Total is selected from Crash Severity, a map will be generated after clicking Refresh Map. Figure 43 displays the generated map after selecting the options. The red boundary indicates the boundary of the district, and the green boundaries indicate the boundaries of the counties. The segments are color-coded based on the total number of expected crashes. The lighter yellow color indicates a lower number of expected crashes, and a darker red indicates a higher number of expected crashes.

Figure 43. Screenshot Showing District-Specific Map.

The engineer also has the option to explore a specific county. For example, if the engineer selects Baylor County, the map will display only low-volume rural two-lane roadways in Baylor County (see Figure 44). The engineer can also download the data after finalizing the selection by clicking the Download Data button. To get more details on the segment, the engineer can zoom in or out of the map by clicking the plus or minus buttons on the top left of the map. As mentioned earlier, the map is interactive and has a hovering option to get more details on a particular segment (see Figure 42).

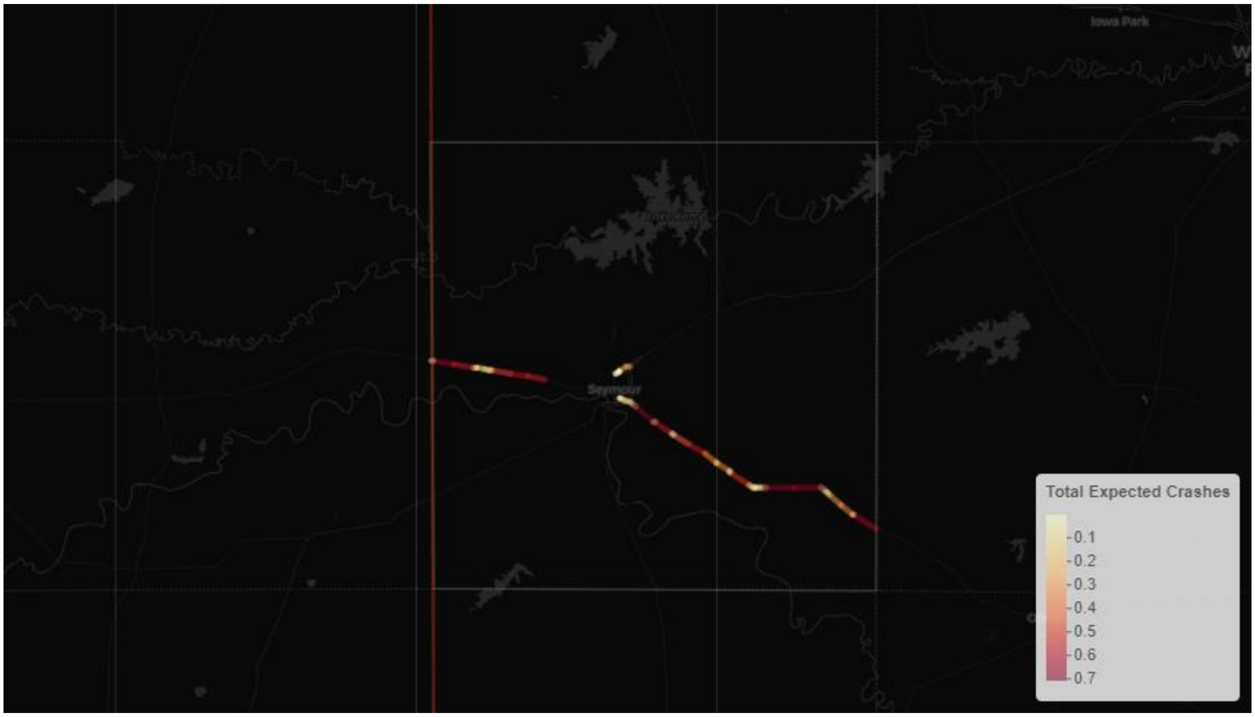

Figure 44. Screenshot Showing County-Specific Map.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

The SHSP has identified speeding-related crashes as one of the seven research emphasis areas for 2017–2022. The conventional crash risk analysis method typically omits real-time speed, real-time volume, and weather data. This can significantly limit its safety predictive performances. To address this gap, TTI used data from five sources in this study:

- NPMRDS (real-time speed data).
- TMAS (traffic volume data).
- CRIS (crash data).
- RHiNO (roadway inventory data).
- NOAA (weather data).

For annual-level data analysis, TTI developed SPFs for four facility types (rural two-lane roadways, rural four-lane divided roadways, rural four-lane undivided roadways, and rural freeways) for two severity groups: fatal and severe injury crashes and PDO crashes. Moreover, since the annual-level safety prediction model can limit the SPFs' performance to reflect the effects of time-sensitive variables such as operating speeds, operating speed variance, and weather condition factors, TTI applied logistic regression to develop a daily-level model as well. In the final step, TTI developed an interactive decision support tool using the open-source software platform Shiny.

6.2 FINDINGS AND CONCLUSIONS

TTI developed the SPFs using the crash data with annual aggregation intervals. The SPFs are developed for FI and PDO crashes and both types together. The findings from the annual-level model are as follows:

- Geometric variables such as lane width, shoulder width, and median width show an association with the crash counts, and the association meets the conventional expectations.
- Truck proportion shows a positive association with FI crashes in rural freeways and rural multilane divided roadways. This variable is not statistically significant for rural multilane undivided roadways and rural two-lane roadways. One possible explanation is that the truck proportion is usually less in the latter two roadway facility types.
- K-factor shows a positive association with crashes on rural facilities. For rural multilane undivided roadways, this factor is not statistically significant.
- Increased variability in operational speed is associated with increased crashes for all four roadway facilities.

- In the absence of a posted speed limit, this study used reference speed (see Table 1) as a surrogate. For rural freeways and rural multilane roadways, with an increase in operating speeds, the relative increase in crashes is greater on highways with higher reference speeds than highways with lower reference speeds.
- Average precipitation shows a positive association for only rural multilane undivided roadways and rural two-lane roadways. This variable is not statistically significant for rural freeways and rural multilane divided roadways.

TTI also developed a daily model with daily-level crash data using a logistic regression method. The findings are as follows:

- Geometric variables such as length and shoulder width show an association with the potential of day-level crash occurrences, and the association meets the conventional expectations.
- Truck proportion shows a positive association with the potential of day-level crash occurrences for all facilities.
- Increased variability in daily operational speed (the SD of the daily average of 5-minute interval operation speeds) is associated with increased day-level crashes for all four roadway facilities.
- Average operating speed (the average of the daily average of 5-minute interval operation speeds) increases are associated with the potential of day-level crash occurrences. However, average operating speed is negatively associated with the potential of total crash occurrences on rural freeways. This finding for the rural freeway could be because of the high design standards for the freeways. This variable is not statistically significant for rural multilane undivided roadways.
- As daily average precipitation increases, the potential of a day-level crash occurrence increases for all facility types.

The developed interactive web-based decision support tool assists practitioners in understanding safety scoring of roadway segments of interest with respect to not only roadway or traffic characteristics but also operating speed and weather conditions. TTI recommends using the tool to take specific measures to mitigate significant safety issues.

6.3 RECOMMENDATIONS

The following recommendations are developed from this study:

- Update the rural roadway SPFs for Texas with the inclusion of operating speed and weather information.
- Update guidelines to provide emphasis on the impact of operating speed on traffic crashes.
- Use the decision support tool to take specific measures to mitigate significant safety issues on rural roadways.

The following are additional recommendations for future work:

- Update the decision support tool with the applicability of data uploading.
- Update the decision support tool with the functionality of short-duration (e.g., daily) safety evaluation.
- Update the daily-level models with short-duration (e.g., daily) traffic counts.
- Conduct a project on urban roadway safety evaluation with the inclusion of operating speed and weather information. Develop an interactive web-based decision support tool for urban roadway safety evaluation.

REFERENCES

Aarts, L., and Van Schagen, I., 2006. Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention*, 38(2): 215-224.

Abdel-Aty, M. A., and Radwan, A. E., 2000. Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5): 633-642.

Ahmed, M. M., Abdel-Aty, M., and Yu, R., 2012. Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data. *Transportation Research Record: Journal of the Transportation Research Board*, 2280: 51-59.

American Association of State Highway and Transportation Officials (AASHTO), 1940. *A Policy on Geometric Design of Highways and Streets*. Washington, DC.

American Association of State Highway and Transportation Officials (AASHTO), 1984. *A Policy on Geometric Design of Highways and Streets*. Washington, DC.

American Association of State Highway and Transportation Officials (AASHTO), 2004. *A Policy on Geometric Design of Highways and Streets*. Washington, DC.

American Association of State Highway and Transportation Officials (AASHTO), 2010. *Highway Safety Manual*, first edition. Washington, DC.

Andrey, J., 2010. Long-term trends in weather-related crash risks. *Journal of Transport Geography*, 18(2): 247-258.

Banihashemi, M., Dimaiuta, M., Zineddin, A., Spear, B., Smadi, O., and Hans, Z., 2019. Using linked SHRP2 RID and NPMRDS data to study speed-safety relationships on urban interstates and major arterials. *Proceeding of the 98th TRB Annual Meeting*, Washington, DC.

Banks, D., Persaud, B., Lynn, C., Eccles, K. A., and Himes, S., 2014. *Enhancing Statistical Methodologies for Highway Safety Research: Impetus from FHWA*. No. FHWA-HRT-14-081| HRDS-20/11-14 (500) E. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Baum, H. M., Lund, A. K., and Wells, J. K., 1989. The mortality consequences of raising the speed limit to 65 mph on rural interstates. *American Journal of Public Health*, 79(10): 1392-1395.

Bergel-Hayat, R., Debbarh, M., Antoniou, C., and Yannis, G., 2013. Explaining the road accident risk: Weather effects. *Accident Analysis & Prevention*, 60: 456-465.

Berry, D. S., and Belmont, D. M., 1951. Distribution of vehicle speeds and travel times. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Regents of the University of California, Oakland, CA.

Bijleveld, F., and Churchill, T., 2009. *The Influence of Weather Conditions on Road Safety: An Assessment of Precipitation and Temperature*. Institute for Road Safety Research, SWOV, The Hague, Netherlands.

Bonneson, J., Lord, D., Zimmerman, K., Fitzpatrick, K., and Pratt, M., 2007. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions*. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Bonneson, J., and M. Pratt, 2008. *Roadway Safety Design Workbook*. Report 0-4703-P2. Texas Transportation Institute, 2008. https://static.tti.tamu.edu/tti.tamu.edu/documents/0-4703-P2.pdf.

Brijs, T., Karlis, D., and Wets, G., 2008. Studying the effect of weather conditions on daily crash counts using a discrete time-series model. *Accident Analysis & Prevention*, 40(3): 1180-1190.

Buddhavarapu, P., Banerjee, B., and Prozzi, J., 2013. Influence of pavement condition on horizontal curve safety. *Accident Analysis & Prevention*, 52: 9-18.

Burritt, B. E., Moghrabi, A., and Matthias, J. S., 1976. Analysis of the relation of accidents and the 88-km/h (55-mph) speed limit on Arizona highways. *Transportation Research Record: Journal of the Transportation Research Board*, 609: 34-35.

Cheng, W., Wang, J., Bryden, G., Ye, X., and Jia, X., 2013. An examination of the endogeneity of speed limits and accident counts in crash models. *Journal of Transportation Safety & Security*, 5: 314-326.

Copulos, M. R., 1986. *The High Cost of the 55 MPH Speed Limit.* Heritage Foundation, Washington, DC.

Dart, O. K., 1977. *The Effects of the 88.5 KPH (55 MPH) Speed Limit and Its Enforcement on Traffic Speeds and Accidents*. Department of Civil Engineering, Louisiana State University, Baton Rouge, LA.

Das, S., Brimley, B. K., Lindheimer, T., and Pant, A., 2017. *Safety Impacts of Reduced Visibility in Inclement Weather.* Texas A&M Transportation Institute, Center for Advancing Transportation Leadership and Safety, and Office of the Assistant Secretary for Research and Technology, Washington, DC.

Das, S., Geedipally, S., Avelar, R., Wu, L., Fitzpatrick, K., Banihashemi, M., and Lord, D., 2020. *Rural Speed Safety Project Final Report*. https://rosap.ntl.bts.gov/view/dot/48840

Das, S., Geedipally, S., 2020. Rural Speed Safety Project for USDOT Safety Data Initiative: Findings and Outcome. ITE J. 90, pp 38-42.

Das, S., Geedipally, S.R., Fitzpatrick, K., 2021. Inclusion of speed and weather measures in safety performance functions for rural roadways. IATSS Res. 45, pp 60-69.

Das, S., White, L.D., 2020. RuralSpeedSafetyX: Interactive decision support tool to improve safety. SoftwareX 11, 100493. https://doi.org/10.1016/j.softx.2020.100493.

De Pauw, E., Daniels, S., Franckx, L., and Mayeres, I., 2018. Safety effects of dynamic speed limits on motorways. *Accident Analysis & Prevention*, 114: 83-89.

De Pauw, E., Daniels, S., Thierie, M., and Brijs, T., 2014. Safety effects of reducing the speed limit from 90 km/h to 70 km/h. *Accident Analysis & Prevention*, 62: 426-431.

Donnell, E. T., Hines, S. C., Mahoney, K. M., Porter, R. J., and McGee, H., 2009. *Speed Concepts: Informational Guide.* No. FHWA-SA-10-001. Office of Safety, Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Donnell, E. T., Kersavage, K., and Tierney, L. F., 2018. *Self-Enforcing Roadways: A Guidance Report.* No. FHWA-HRT-17-098. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Dutta, N., and Fontaine, M. D., 2019. Improving freeway segment crash prediction models by including disaggregate speed data from different sources. *Accident Analysis & Prevention*, 132: 105253.

El-Basyouny, K., Barua, S., and Islam, M. T., 2014a. Investigation of time and weather effects on crash types using full Bayesian multivariate Poisson lognormal models. *Accident Analysis & Prevention*, 73: 91-99.

El-Basyouny, K., Barua, S., Islam, M. T., and Li, R., 2014b. Assessing the effect of weather states on crash severity and type by use of full Bayesian multivariate safety models. *Transportation Research Record: Journal of the Transportation Research Board*, 2432: 65-73.

Elvik, R., 2009. *The Power Model of the Relationship between Speed and Road Safety: Update and New Estimates.* Report 1034. Institute of Transport Economics, Oslo, Norway.

Elvik, R., 2010. A restatement of the case for speed limits. *Transport Policy*, 17(3): 196-204.

Elvik, R., 2013. A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims. *Accident Analysis & Prevention*, 50: 854-860.

Elvik, R., Christensen, P., and Amundsen, A. H., 2004. *Speed and road accidents: An evaluation of the Power Model.* TØI Report 740. Institute of Transport Economics, Oslo, Norway.

Farmer, C. M., Retting, R. A., and Lund, A. K., 1997. *Effect of 1996 Speed Limit Changes on Motor Vehicle Occupant Fatalities.*

Federal Highway Administration (FHWA), 2020. *Speed Limit Basics*. Report FHWA-SA-16-076. Washington, DC. https://safety.fhwa.dot.gov/speedmgt/ref_mats/fhwasa16076/fhwasa16076.pdf.

Fitzpatrick, K., Carlson, P., Brewer, M., Wooldrige, M., and Miaou, S., 2003. *Design Speed, Operating Speed and Posted Speed Practices.* No. 504. National Cooperative Highway Research Program Report, Washington, DC.

Fitzpatrick, K., Das, S., Pratt, M.P., Dixon, K., Gates, T., 2021a. Posted Speed Limit Setting Procedure and Tool: User Guide. National Cooperative Highway Research Program Report, Washington, DC.

Fitzpatrick, K., Das, S., Pratt, M.P., Dixon, K., Gates, T., 2021b. Development of a Posted Speed Limit Setting Procedure and Tool, National Cooperative Highway Research Program Report, Washington, DC.

Fitzpatrick, K., Das, S., 2019. Vehicle Operating Speed on Urban Arterial Roadways. Safe D Report. Washington, DC.

Forester, T. H., McNown, R. F., and Singell, L. D., 1984. A cost-benefit analysis of the 55 mph speed limit. *Southern Economic Journal*, 50(3): 631-641.

Freedman, M., and Esterlitz, J. R., 1990. *Effect of the 65-mph Speed Limit on Speeds in Three States*. No. 1281. Transportation Research Board, Washington, DC.

Freedman, M., and Williams, A. F., 1992. Speeds associated with 55 mph and 65 mph speed limits in northeastern states. *ITE Journal*, 62(2): 17-22.

Garber, N. J., and Gadiraju, R., 1989. Factors affecting speed variance and its influence on accidents. *Transportation Research Record: Journal of the Transportation Research Board*, 1213: 64-71.

Garber, N. J., and Srinivasan, S., 1998. Influence of exposure duration on the effectiveness of changeable-message signs in controlling vehicle speeds at work zones. *Transportation Research Record: Journal of the Transportation Research Board*, 1650(1): 62-70.

Gargoum, S. A., and El-Basyouny, K., 2016. Exploring the association between speed and safety: A path analysis approach. *Accident Analysis & Prevention*, 93: 32-40.

Gayah, V. V., Donnell, E. T., Yu, Z., and Li, L., 2018. Safety and operational impacts of setting speed limits below engineering recommendations. *Accident Analysis & Prevention*, 121: 43-52.

Haddon, Jr, W., 1972. A logical framework for categorizing highway safety phenomena and activity. *Journal of Trauma and Acute Care Surgery*, 12(3): 193-207.

Haglund, M., and Åberg, L., 2000. Speed choice in relation to speed limit and influences from other drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 3(1): 39-51.

Harkey, D. L., Robertson, H. D., and Davis, S. E., 1990. *Assessment of Current Speed Zoning Criteria*. No. 1281. Transportation Research Board, Washington, DC.

Hauer, E., 1971. Accidents, overtaking and speed control. *Accident Analysis & Prevention*, 3(1): 1-13.

Hauer, E., 2009. Speed and safety. *Transportation Research Record: Journal of the Transportation Research Board*, 2103(1): 10-17.

Hauer, E., and Bonneson, J., 2006. An Empirical Examination of the Relationship between Speed and Road Accidents Based on Data by Elvik, Christensen and Amundsen. Unpublished manuscript data, March 5, 2006. Prepared for the Highway Safety Manual Task Force.

Himes, S., Gross, F., Nichols, M., and Lockwood, M., 2018. Safety evaluation of change in posted speed limit from 65 to 70 mph on rural Virginia interstate system. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(38): 35-45.

Hu, W., 2017. Raising the speed limit from 75 to 80 mph on Utah rural interstates: Effects on vehicle speeds and speed variance. *Journal of Safety Research*, 61: 83-92.

Hutton, J. M., Cook, D. J., Grotheer, J., and Conn, M., 2020. *Research Utilizing SHRP2 Data to Improve Highway Safety: Development of Speed—Safety Relationships*. No. FHWA-HRT-20-035. Office of Safety Research and Development, Federal Highway Administration, U.S. Department of Transportation, McLean, VA.

Imprialou, M. I. M., Quddus, M., and Pitfield, D. E., 2016. Predicting the safety impact of a speed limit increase using condition-based multivariate Poisson lognormal regression. *Transportation Planning & Technology*, 39(1): 3-23.

Islam, M. S., Ivan, J. N., Lownes, N. E., Ammar, R. A., and Rajasekaran, S., 2014. Developing safety performance function for freeways by considering interactions between speed limit and geometric variables. *Transportation Research Record: Journal of the Transportation Research Board*, 2435(1): 72-81.

Jaarsma, R., Louwerse, R., Dijkstra, A., de Vries, J., and Spaas, J. P., 2011. Making minor rural road networks safer: The effects of 60 km/h-zones. *Accident Analysis & Prevention*, 43(4): 1508-1515.

Jackson, T. L., and Sharif, H. O., 2016. Rainfall impacts on traffic safety: Rain-related fatal crashes in Texas. *Geomatics, Natural Hazards & Risk*, 7(2): 843-860.

Jin, P. J., Walker, A., Cebelak, M., and Walton, C. M., 2014. Determining strategic locations for environmental sensor stations with weather-related crash data. *Transportation Research Record: Journal of the Transportation Research Board*, 2440: 34-42.

Jung, S., Qin, X., and Noyce, D. A., 2012. Injury severity of multivehicle crash in rainy weather. *Journal of Transportation Engineering*, 138(1): 50-59.

Kelarestaghi, K. B., Zhang, W., Wang, Y., Xiao, L., Hancock, K., and Heaslip, K. P., 2017. Impacts to crash severity outcome due to adverse weather and other causation factors. *Advances in Transportation Studies*, 43: 31-42.

Kockelman, K., Bottom, J., Kweon, Y. J., Ma, J., and Wang, X., 2006. *Safety Impacts and Other Implications of Raised Speed Limits on High-Speed Roads*. Vol. 90. Transportation Research Board, Washington, DC.

Kweon, Y. J., and Kockelman, K. M., 2005. Safety effects of speed limit changes: Use of panel models, including speed, use, and design variables. *Transportation Research Record: Journal of the Transportation Research Board*, 1908(1): 148-158.

Lamm, R., Choueiri, E., and Mailaender, T., 1990. Comparison of operating speeds on dry and wet pavements of two-lane rural highways. *Transportation Research Record: Journal of the Transportation Research Board*, 1280: 199-207.

Lave, C. A., 1985. Speeding, coordination, and the 55mph limit. *The American Economic Review*, 75(5): 1159-1164.

Lee, J., Chae, J., Yoon, T., and Yang, H., 2018. Traffic accident severity analysis with rain-related factors using structural equation modeling—A case study of Seoul City. *Accident Analysis & Prevention*, 112: 1-10.

Levin, M., Chen, R. C., Liao, C. F. L., and Zhang, T., 2019. *Improving Intersection Safety through Variable Speed Limits for Connected Vehicles*. Center for Transportation Studies, University of Minnesota, Minneapolis, MN.

Lord, D., and 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.

Malin, F., Norros, I., and Innamaa, S., 2019. Accident risk of road and weather conditions on different road types. *Accident Analysis & Prevention*, 122: 181-188.

Malyshkina, N. V., and Mannering, F., 2008. Effect of increases in speed limits on severities of injuries in accidents. *Transportation Research Record: Journal of the Transportation Research Board*, 2083(1): 122-127.

Mannering, F., and Bhat, C., 2014. Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research*, 1: 1-22.

Maycock, G., Brocklebank, P. J., and Hall, R. D., 1998. *Road Layout Design Standards and Driver Behaviour*. TRL Report 332. Transportation Research Laboratory, Berkshire, England.

Mayora, J., and Pina, R., 2009. An assessment of the skid resistance effect on traffic safety under wet-pavement conditions. *Accident Analysis & Prevention*, 41(4): 881-886.

McCarthy, P., 2001. Effect of speed limits on speed distributions and highway safety: A survey of recent literature. *Transport Reviews*, 21(1): 31-50.

McCourt, R., Fitzpatrick, K., Koonce, P., Das, S., 2019. Speed Limits: Leading to Change. ITE J. 89, pp 38-43.

Monsere, C. M., Bosa, P. G., and Bertini, R. L., 2008. Combining climate, crash, and highway data for improved ranking of speed and winter-weather related crash locations in Oregon. *Journal of Transportation Engineering*, 134(7): 287-296.

Monsere, C., Kothuri, S., and Anderson, J., 2018. *Preliminary Analysis of Speed Limit Changes in Eastern Oregon*. Oregon Department of Transportation, Salem, OR.

Montella, A., and Imbriani, L. L., 2015. Safety performance functions incorporating design consistency variables. *Accident Analysis & Prevention*, 74: 133-144.

Najafi, S., Flintsch, G., and Medina, A., 2015. Linking roadway crashes and tire-pavement friction: A case study. *International Journal of Pavement Engineering*, 18(2), 119-127.

Nilsson, G., 2004. *Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety*. Department of Technology and Society, Lund Institute of Technology, Lund, Sweden

Ossiander, E. M., and Cummings, P., 2002. Freeway speed limits and traffic fatalities in Washington State. *Accident Analysis & Prevention*, 34(1): 13-18.

Parker, Jr., M. R., 1997. *Effects of Raising and Lowering Speed Limits on Selected Roadway Sections*. No. FHWA-RD-97-084. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Pei, X., Wong, S. C., and Sze, N. N., 2012. The roles of exposure and speed in road safety analysis. *Accident Analysis & Prevention*, 48: 464-471.

Retting, R., and Cheung, I., 2008. Traffic speeds associated with implementation of 80 mph speed limits on West Texas rural interstates. *Journal of Safety Research*, 39(5): 529-534.

Rosén, E., and Sander, U., 2009. Pedestrian fatality risk as a function of car impact speed. *Accident Analysis & Prevention*, 41(3): 536-542.

Rossetti, M. A., and Johnsen, M., 2011. *Weather and Climate Impacts on Commercial Motor Vehicle Safety*. Research and Innovative Technology Administration and Federal Motor Carrier Safety Administration, Washington DC.

Saha, P., Ahmed, M. M., and Young, R. K., 2015. Safety effectiveness of variable speed limit system in adverse weather conditions on challenging roadway geometry. *Transportation Research Record: Journal of the Transportation Research Board*, 2521(1): 45-53.

Savolainen, P., Mannering, F., Lord, D., and Quddus, M., 2011. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention*, 43(5): 1666-1676.

Sayed, T., and Sacchi, E., 2016. Evaluating the safety impact of increased speed limits on rural highways in British Columbia. *Accident Analysis & Prevention*, 95: 172-177.

Solomon, D. H., 1964. *Accidents on Main Rural Highways: Related to Speed, Driver, and Vehicle*. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Strong, C. K., Ye, Z., and Shi, X., 2010. Safety effects of winter weather: The state of knowledge and remaining challenges. *Transport Reviews*, 30(6): 677-699.

Tarko, A. P., Pineda-Mendez, R., and Guo, Q., 2019. *Predicting the Impact of Changing Speed Limits on Traffic Safety and Mobility on Indiana Freeways.* Indiana Department of Transportation, Indianapolis, IN.

Taylor, M. C., Lynam, D. A., and Baruya, A., 2000. *The Effects of Drivers' Speed on the Frequency of Road Accidents*. Crowthorne, Transport Research Laboratory, Berkshire, England.

Theofilatos, A., 2019. Utilizing real-time traffic and weather data to explore crash frequency on urban motorways: A cusp catastrophe approach. *Transportation Research Procedia*, 41: 471-479.

Theofilatos, A., and Yannis, G., 2014. A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis & Prevention*, 72: 244-256.

Ullman, G. L., and Rose, E. R., 2005. Evaluation of dynamic speed display signs. *Transportation Research Record: Journal of the Transportation Research Board*, 1918(1): 92-97.

Upchurch, J., 1989. Arizona's experience with the 65-mph speed limit. *Transportation Research Record: Journal of the Transportation Research Board*, 1244: 1-6.

Vernon, D. D., Cook, L. J., Peterson, K. J., and Dean, J. M., 2004. Effect of repeal of the national maximum speed limit law on occurrence of crashes, injury crashes, and fatal crashes on Utah highways. *Accident Analysis & Prevention*, 36(2): 223-229.

Wang, X., Zhou, Q., Quddus, M., and Fan, T., 2018. Speed, speed variation and crash relationships for urban arterials. *Accident Analysis & Prevention*, 113: 236-243.

Wen, H., Zhang, X., Zeng, Q., and Sze, N. N., 2019. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accident Analysis & Prevention*, 132: 105249

Xu, C., Wang, C., and Liu, P., 2018. Evaluating the Combined effects of weather and real-time traffic conditions on freeway crash risks. *Weather, Climate, and Society*, 10(4): 837-850.

Xu, C., Wang, X., Yang, H., Xie, K., and Chen, X., 2019. Exploring the impacts of speed variances on safety performance of urban elevated expressways using GPS data. *Accident Analysis & Prevention*, 123: 29-38.

Yu, R., and Abdel-Aty, M., 2014. Analyzing Crash Injury Severity for a Mountainous Freeway Incorporating Real-Time Traffic and Weather Data. *Safety Science*, 63: 50-56.

Yu, R., Abdel-Aty, M., and Ahmed, M., 2013. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. *Accident Analysis & Prevention*, 50: 371-376.

Yu, R., Abdel-Aty, M., Ahmed, M. M., and Wang, X., 2014. Utilizing microscopic traffic and weather data to analyze real-time crash patterns in the context of active traffic management. *IEEE Transactions on Intelligent Transportation Systems*, 15(1): 205-213.

Yu, R., Quddus, M., Wang, X., and Yang, K., 2018. Impact of data aggregation approaches on the relationships between operating speed and traffic safety. *Accident Analysis & Prevention*, 120: 304-310.

Yu, R., Xiong, Y., and Abdel-Aty, M., 2015. A correlated random parameter approach to investigate the effects of weather conditions on crash risk for a mountainous freeway. *Transportation Research Part C: Emerging Technologies*, 50: 68-77.

APPENDIX A:
SOFTWARE CODES OF THE DECISION SUPPORT TOOL

```
### The codes developed for this tool were prepared by Subasish Das, Lingtao Wu, and
### Zihang Wei.

# Load Libraries

library(shiny)
library(shinydashboard)
library(shinyjs)
library(sf)
library(leaflet)
library(leaflet.extras)
library(dplyr)
library(DT)
library(htmltools)
library(shinyWidgets)
library(readxl)
library(writexl)

# Read State/Counties CSV File

StateCountyData = read.csv("County_District_List/CountyList.csv")

# Create State and Initial County List
StateCountyData$District <- as.character(StateCountyData$District)
StateCountyData$County <- as.character(StateCountyData$County)


DistrictList <- unique(StateCountyData$District)

CountyList <- StateCountyData$County

DistrictNameList <- data.frame(
  unique(StateCountyData[c("District","DistrictID")])
)


# Read TMC_Tes Texas SHP file

TX_shp = st_transform(st_read("ShapeFiles/TMC_Yes_Polyline/TMC.shp"), 4326)
TX_shp$FICrO <- round(TX_shp$FICrO/5,0)
TX_shp$PDOCrO <- round(TX_shp$PDOCrO/5,0)
TX_shp$TotalCrO <- round(TX_shp$TotalCrO/5,0)
TX_shp$FICrE <- TX_shp$FICrE/5
```

```
TX_shp$PDOCrE <- TX_shp$PDOCrE/5
TX_shp$TotalCrE <- TX_shp$TotalCrE/5
TX_shp$FIPre <- TX_shp$FIPre/5
TX_shp$PDOPre <- TX_shp$PDOPre/5
TX_shp$TotalPre <- TX_shp$TotalPre/5
TX_shp$FICrR <- TX_shp$FIPre/5
TX_shp$PDOCrR <- TX_shp$PDOCrR/5
TX_shp$TotalCrR <- TX_shp$TotalCrR/5


# Read District SHP file

TXDistricts_shp = st_transform(st_read("ShapeFiles/Texas_Districts/District_Poly.shp"), 4326)

TXDistricts_shp = st_as_sf(select(as.data.frame(TXDistricts_shp),c('DIST_NBR','geometry')))

# Read Counties SHP file
TXcounties_shp = st_transform(st_read("ShapeFiles/Texas_Counties/County.shp"), 4326)

TXcounties_shp =
    st_as_sf(select(as.data.frame(TXcounties_shp),c('CNTY_NBR','DIST_NBR','geometry'))
    )

# vardefini <- read_excel("ShapeFiles/Data_Dictionary.xlsx")

# Start Creating Dashboard layout

body <- dashboardBody(useShinyjs(),
            setBackgroundColor("black"),
            tabsetPanel(
              tabPanel(HTML(paste(tags$span(style="font-size: 22px", "Introduction"))),
                    tags$br(),
                    tags$h2(tags$b("TxDOT 0-7051 Research Project Decision Support
    Tool")),
                    h2(),
                    div(style = "font-size: 22px;", HTML("TxDOT 0-7051 Project <b>Develop
    a Real-Time Decision Support Tool for Rural Roadway Safety Improvements</b>
    developed an
                                        interactive tool that identifies crash hotspots on
    Texas rural roadways. Four different roadway facility types are included
                                        in this tool: Rural Multilane Divided, Rural
    Multilane Undivided, Rural Two Lane, and Rural Interstate. The tool provides
                                        yearly observed, predicted, and expected crashes
    along with crashes per million VMT. The tool is based on 6 years of crash
                                        data (2015-2020) from CRIS. Other sources of data
    includes the RHiNO database,  NOAA (for precipitation data), and
```

```
                                    NPMRDS (for speed data).")),
          h2(),
          div(style = "font-size: 22px;", HTML("The results can be filtered by the
following in the 0-7051 Tool tab:",
                                    "<br>"
          )),
          tags$span(style="font-size: 18px", tags$ul(
            tags$li("District"),
            tags$li("County"),
            tags$li("Facility Type"),
            tags$li("Select AADT Range"),
            tags$li("Crash Severity Level")
          )),
          h2(),
          div(style = "font-size: 22px;", HTML("Once the levels are selected, the user
needs to:",
                                    "<br>"
          )),
          tags$span(style="font-size: 18px", tags$ul(
            tags$li("Click 'Refresh Map' (in blue box), note it may take some time to
load the map")
          )),

          h2(),
          div(style = "font-size: 22px;", HTML("The results can be filtered by the
following in the 0-7051 Tool tab:",
                                    "<br>"
          )),
          tags$span(style="font-size: 18px", tags$ul(
            tags$li("Detailed data can be downloaded by clicking the 'Download Data'
button (grey box below blue box)"),
            tags$li("Data dictionary can be downloaded after refreshing the map (see
Note 1 in the 0-7051 Tool tab)"),
            tags$li("Results can be shown in lists of 10, 25, 50, or 100 entries"),
            tags$li("Results can be sorted (up or down) by using the arrows at the top
of each variable's column"),
            tags$li("A search box provides the opportunity to search the results")
          )),
          h2(),
          div(style = "font-size: 22px;", HTML("Notes:",
                                    "<br>"
          )),
          tags$span(style="font-size: 18px", tags$ul(
            tags$li("Interactive tables show variable codes to fit the table in the
browser page. For variable details, please download the data dictionary."),
```

```
                  tags$li("This tool uses CartoDB 'Dark Matter with labels' as the basemap.
After several iterations, this basemap has been chosen to make the map more focused
towards the roadway safety scores. Details of roadway name and other variables can be
found by hovering the mouse on the segment of interest. ")
                )),

                tags$h3(tags$b("Acknowledgments")),
                h2(),
                div(style = "font-size: 22px;", HTML("Texas Department of Transportation
(TxDOT) funded the project.")),

                h2(),
                div(style = "font-size: 22px;", HTML("The project was conducted by Texas
A&M Transportation Institute (TTI).",
                                        "TTI conducted the project with the interactive
online tool being developed by Dr. Subasish Das, Dr. Lingtao Wu, and Mr. Zihang Wei. ",
                                        "Questions about the tool can be sent to project
Research Supervisor Dr. Subasish Das <a href='mailto:s-das@tti.tamu.edu'>s-
das@tti.tamu.edu</a> or",
                                        "Co-Research Supervisor  Dr. Kay Fitzpatrick <a
href='mailto:k-fitzpatrick@tti.tamu.edu'>k-fitzpatrick@tti.tamu.edu</a>."
                )),

                hr(),
                tags$img(src='txdot_Logo.png', height=100), HTML("   "),
                tags$img(src='TTI_Logo.png', height=100),
                hr(),
                div(style = "font-size: 18px;", HTML("Last updated: November 29,
2021."))),

            tabPanel(HTML(paste(tags$span(style="font-size: 22px", "0-7051 Tool"))),
id="RuralSpeedTool",
                tags$h2(tags$b("Interactive Decision Support Tool to Improve Safety for
Texas Rural Roadways with Speed Data (0-7051 Tool)")),
                tags$img(src='txdot_Logo.png', height=40), HTML("   "),
                tags$img(src='TTI_Logo.png', height=40),
                fluidRow(
                  column(width = 9,
                      box(width = NULL, solidHeader = TRUE,
                        leafletOutput("MapOut", height = 650),
                        h2()                              )
                  ),
                  column(width = 3,
                      box(width = NULL, status = "warning",
                        selectInput("DistrictInput","District",choices = c("All
Districts",sort(DistrictList)),selected = "All Districts"),
```

```
                              selectInput("CountyInput","County",choices = c("All Counties",sort(CountyList)),selected = "All Counties"),
                              selectInput("FacilityInput","Facility",choices = c("All Facilities","Rural Mulitlane Divided","Rural Mulitlane Undivided","Rural Two Lane","Rural Interstate"), selected = "All Facilities"),
                              selectInput("AADTInput", "AADT Ranges", choices = c("All Levels", "Less than 2000", "2001 to 10000", "Greater than 10000"), selected = "All"),
                              radioButtons("Severity", label = "Crash Severity", choices = list("Total", "Fatal and Injury","No Injury"), inline=TRUE),
                              hr(),
                              checkboxInput("showlongseg","Remove very short segments (segment length less than 0.1 mi)", F),
                              actionButton(inputId = "RefreshMap", label = "Refresh Map", class = "butt"),
                              tags$head(tags$style(".butt{background-color:#0000FF;}
.butt{color: white;}")), # background color and font color
                              downloadButton("downloadData",label ="Download Data"),
                              hr(),
                              HTML("Note:"),
                              h2(),
                              HTML("<b>1. Data Dictionary can be downloaded from</b>"),
                              downloadButton("downloadDefination",label ="here"),
                              HTML("<b>(Please refresh map first before downloading).</b>"),
                              h2(),
                              HTML("2. The table below can display 10, 25, 50, or 100 entries using the drop-down menu to the left. The table is interactive and each column can be sorted using the up or down arrows at the head of the column or the data can be searched using the search box below.")
                            )
                         )

                  ),
                  DT::DTOutput('outputDT'),
                  h2(),tags$br(),
                  h2(),tags$br()


            )
        ))

# Put them together into a dashboardPage
ui <- dashboardPage(
   #header,
   dashboardHeader(disable = TRUE),
   dashboardSidebar(disable = TRUE),
   body
```

```
)

server <- function(input, output, session) {


  observeEvent(input$DistrictInput,{
    if(input$DistrictInput != "All Districts"){
      updateSelectInput(session, "CountyInput","County",choices = c("All
      Counties",subset(StateCountyData$County, StateCountyData$District ==
      input$DistrictInput)))
    }
    else{
      updateSelectInput(session, "CountyInput","County",choices = c("All Counties"))
    }
  }

  )

  output$MapOut <- renderLeaflet({
    leaflet() %>%
      addTiles(urlTemplate = "//cartodb-basemaps-
      {s}.global.ssl.fastly.net/dark_all/{z}/{x}/{y}{r}.png", layerId = 'Carto DB Dark Matter')
      %>%
      setView(lng = -95.7129, lat = 37.0902, zoom = 4)
  })

  observeEvent(input$RefreshMap, {

    Districtin <- input$DistrictInput
    if (Districtin != "All Districts"){
      Districtin_Code <- unique(select(filter(StateCountyData, District == Districtin),
      DistrictID))[1,1]
    } else {
      Districtin_Code <- 0
    }


    COUNTYin = input$CountyInput
    if (COUNTYin != "All Counties"){
      Countyin_Code <- select(filter(filter(StateCountyData, DistrictID == Districtin_Code),
      (County == COUNTYin)), CountyID)[1,1]
    } else {
      Countyin_Code <- 0
    }

    Facilityin_Code <- switch(input$FacilityInput,
```

```
                    "All Facilities" = 0,
                    "Rural Mulitlane Divided" = "RMD",
                    "Rural Mulitlane Undivided" = "RMU",
                    "Rural Two Lane" = "R2",
                    "Rural Interstate" = "RI"
)

AADTin_Code <- switch(input$AADTInput,
                 "All Levels" = 0,
                 "Less than 2000" = 1,
                 "2001 to 10000" = 2,
                 "Greater than 10000" = 3
)

Longsegment <- input$showlongseg

### Will need to change MapOutputData switch to include other states

if (Longsegment == 1){
   MapOutputData <- st_as_sf(filter(as.data.frame(TX_shp), Length >= 0.1))
}
else{
   MapOutputData <- TX_shp
}



if (Districtin_Code == 0){
   MapOutputDataTempDistrict <- MapOutputData
   TXdistricts_shp_selected <- st_as_sf(as.data.frame(TXDistricts_shp))

}
else{
   MapOutputDataTempDistrict <- st_as_sf(filter(as.data.frame(MapOutputData), District
   == Districtin_Code))
   TXdistricts_shp_selected <- st_as_sf(filter(as.data.frame(TXDistricts_shp), DIST_NBR
   == Districtin_Code))

}

if (Countyin_Code == 0){
   MapOutputDataTempCounty <- MapOutputDataTempDistrict

   if (Districtin_Code != 0){
      TXcounties_shp_selected <- st_as_sf(filter(as.data.frame(TXcounties_shp),
   DIST_NBR == Districtin_Code))
```

```
      corr <- as.data.frame(st_coordinates(st_centroid(TXdistricts_shp_selected)))

      LATzoom <- corr$Y
      LONzoom <- corr$X

      zoomLevel <- 8
    }
    else{
      TXcounties_shp_selected         <-          st_as_sf(filter(as.data.frame(TXcounties_shp),
    DIST_NBR == 0))
      LATzoom <- 31.9686
      LONzoom <- -99.9018

      zoomLevel <- 6

    }


  } else {
    MapOutputDataTempCounty                                                              <-
    st_as_sf(filter(as.data.frame(MapOutputDataTempDistrict), County == Countyin_Code))
    TXcounties_shp_selected <- st_as_sf(filter(as.data.frame(TXcounties_shp), CNTY_NBR
    == Countyin_Code))
    corr <- as.data.frame(st_coordinates(st_centroid(TXcounties_shp_selected)))

    LATzoom <- corr$Y
    LONzoom <- corr$X

    zoomLevel <- 9

  }

  if(Facilityin_Code == 0){
    MapOutputDataTempFacility <- MapOutputDataTempCounty
  }
  else{
    MapOutputDataTempFacility                                                            <-
    st_as_sf(filter(as.data.frame(MapOutputDataTempCounty), FaciTy ==  Facilityin_Code))
  }


  if(AADTin_Code == 0){
    MapOutputDataTempAADT <-  MapOutputDataTempFacility
  }
  else if(AADTin_Code == 1){
```

```
  MapOutputDataTempAADT <-
  st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT <= 2000))
}
else if(AADTin_Code == 2){
  MapOutputDataTempAADT <-
  st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT > 2000 & ADT <=
  10000))
}
else {
  MapOutputDataTempAADT <-
  st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT > 10000))
}


MapOutputDataFinal <- MapOutputDataTempAADT

displaydata <- as.data.frame(MapOutputDataFinal)
displaydata <- within(displaydata, FaciTy[FaciTy == 'RI'] <- 'Rural Interstate')
displaydata <- within(displaydata, FaciTy[FaciTy == 'R2'] <- 'Rural Two Lane')
displaydata <- within(displaydata, FaciTy[FaciTy == 'RMU'] <- 'Rural Multilane Undivided')
displaydata <- within(displaydata, FaciTy[FaciTy == 'RMD'] <- 'Rural Multilane Divided')

displaydata <- left_join(displaydata,StateCountyData, by = c("District" =
  "DistrictID","County" = "CountyID"))

displaydata <- within(displaydata, TotalCrE[TotalCrE < 0.01] <- 0.01 )
displaydata <- within(displaydata, PDOCrE[PDOCrE < 0.01] <- 0.01 )
displaydata <- within(displaydata, FICrE[FICrE < 0.01] <- 0.01 )

names(displaydata)[names(displaydata) == "District"] <- "District_Code"
names(displaydata)[names(displaydata) == "County"] <- "County_Code"
names(displaydata)[names(displaydata) == "District.y"] <- "District"
names(displaydata)[names(displaydata) == "County.y"] <- "County"

DataForPal <- switch(input$Severity,
                "Total" = MapOutputDataFinal$TotalCrE,
                "Fatal and Injury" = MapOutputDataFinal$FICrE,
                "No Injury" = MapOutputDataFinal$PDOCrE
)


pal_Total <- colorNumeric("YlOrRd", DataForPal)

labelOut <- as.list(paste0('<b><font size="2">Segment Level
  Information</font></b>','</p><font size="2">',
                    'ID: ', displaydata$ID, "<br>",
```

```
                'Highway: ', displaydata$Hwy, "<br>",
                'Facility Type: ', displaydata$FaciTy, "<br>",
                'Length (mi): ', format(round(displaydata$Length,3),nsmall = 3), "<br>",
                'ADT (Year 2018): ', displaydata$ADT, "<br>",
                'Control Section: ', displaydata$ControlS, "<br>",
                'From DFO: ', displaydata$FrmDFO, "<br>",
                'To DFO: ', displaydata$ToDFO, "<br>",
                'Avg. Speed (mph): ', format(round(displaydata$SpdAve,2),nsmall = 2),
   "<br>",
                'Std. of Speed (mph): ', format(round(displaydata$SpdStd,2),nsmall = 2),
   "<br>",
                'Reference Speed (mph): ', format(round(displaydata$RefSpd,2),nsmall =
   2), "<br>",
                '85th Perc. Speed (mph): ', format(round(displaydata$Spd85,2),nsmall = 2),
   "<br>",
                'Avg. Precipitation (in): ', format(round(displaydata$Precip,2),nsmall = 2),
   "<br>",
                'Lane Wid. (ft): ', displaydata$LaneWid, "<br>",
                'Inside Shoulder Wid. (ft): ', displaydata$SHDWidI, "<br>",
                'Outside Shoulder Wid. (ft): ', displaydata$SHDWidO, "<br>",
                'Number of Lanes: ', displaydata$NumLane, "<br>",
                'Truck Perc.: ', displaydata$TrkPer, "<br>",
                'Max. PSL (mph): ', displaydata$MaxPSL, "<br>",
                'Obs. Total Crash (1 yr): ', displaydata$TotalCrO, "<br>",
                'Obs. FI Crash (1 yr): ', displaydata$FICrO, "<br>",
                'Obs. PDO Crash (1 yr): ', displaydata$PDOCrO, "<br>",
                'Exp. Total Crash (1 yr): ', format(round(displaydata$TotalCrE,2),nsmall =
   2), "<br>",
                'Exp. FI Crash (1 yr): ', format(round(displaydata$FICrE,2),nsmall = 2),
   "<br>",
                'Exp. PDO Crash (1 yr): ', format(round(displaydata$PDOCrE,2),nsmall =
   2), "<br>"
))


popupOut <- paste0('ID: ', displaydata$ID, "<br>",
            'Highway: ', displaydata$Hwy, "<br>",
            'Facility Type: ', displaydata$FaciTy, "<br>",
            'Length (mi): ', format(round(displaydata$Length,3),nsmall = 3), "<br>",
            'ADT (Year 2018): ', displaydata$ADT, "<br>",
            'Control Section: ', displaydata$ControlS, "<br>",
            'From DFO: ', displaydata$FrmDFO, "<br>",
            'To DFO: ', displaydata$ToDFO, "<br>",
            'Avg. Speed (mph): ', format(round(displaydata$SpdAve,2),nsmall = 2), "<br>",
            'Std. of Speed (mph): ', format(round(displaydata$SpdStd,2),nsmall = 2), "<br>",
```

```
            'Reference Speed (mph): ', format(round(displaydata$RefSpd,2),nsmall = 2),
   "<br>",
            '85th Perc. Speed (mph): ', format(round(displaydata$Spd85,2),nsmall = 2),
   "<br>",
            'Avg. Precipitation (in): ', format(round(displaydata$Precip,2),nsmall = 2),
   "<br>",
            'Lane Wid. (ft): ', displaydata$LaneWid, "<br>",
            'Inside Shoulder Wid. (ft): ', displaydata$SHDWidI, "<br>",
            'Outside Shoulder Wid. (ft): ', displaydata$SHDWidO, "<br>",
            'Number of Lanes: ', displaydata$NumLane, "<br>",
            'Truck Perc.: ', displaydata$TrkPer, "<br>",
            'Max. PSL (mph): ', displaydata$MaxPSL, "<br>",
            'Obs. Total Crash (1 yr): ', displaydata$TotalCrO, "<br>",
            'Obs. FI Crash (1 yr): ', displaydata$FICrO, "<br>",
            'Obs. PDO Crash (1 yr): ', displaydata$PDOCrO, "<br>",
            'Exp. Total Crash (1 yr): ', format(round(displaydata$TotalCrE,2),nsmall = 2),
   "<br>",
            'Exp. FI Crash (1 yr): ', format(round(displaydata$FICrE,2),nsmall = 2), "<br>",
            'Exp. PDO Crash (1 yr): ', format(round(displaydata$PDOCrE,2),nsmall = 2),
   "<br>"
)

leafletProxy("MapOut") %>% clearPopups() %>% clearGroup("Total/Fata/Injury") %>%
   clearGroup("CountiesSHP") %>% clearGroup("DistrictsSHP") %>% clearControls()
   %>%
   setView(lng = LONzoom, lat = LATzoom, zoom = zoomLevel) %>%
   addPolylines(data=MapOutputDataFinal,
            color=~pal_Total(
               switch(input$Severity,
                    "Total" = MapOutputDataFinal$TotalCrE,
                    "Fatal and Injury" = MapOutputDataFinal$FICrE,
                    "No Injury" = MapOutputDataFinal$PDOCrE)
            ),
            group="Total/Fata/Injury",
            popup = popupOut,
            label = lapply(labelOut, HTML)) %>%
   addPolylines(data=TXcounties_shp_selected,
            color='#81A88D',
            group="CountiesSHP",
            weight = 1) %>%
   addPolylines(data=TXdistricts_shp_selected,
            color='#C93312',
            group="DistrictsSHP",
            weight =2) %>%
   addLegend("bottomright", pal = pal_Total,
          values = switch(input$Severity,
```

```
                        "Total" = MapOutputDataFinal$TotalCrE,
                        "Fatal and Injury" = MapOutputDataFinal$FICrE,
                        "No Injury" = MapOutputDataFinal$PDOCrE),
            title = paste0(input$Severity, " Expected Crashes")
    )




if (nrow(displaydata) > 0){
    MapOutputDataFinalDTtemp <- cbind(select(displaydata,
                                        c('ID',
                                          'District',
                                          'County',
                                          'Hwy',
                                          'FaciTy',
                                          'Length',
                                          'ADT',
                                          'SpdAve',
                                          'SpdStd',
                                          'MaxPSL',
                                          'Precip',
                                          'TotalPre',
                                          'FIPre',
                                          'PDOPre',
                                          'TotalCrE',
                                          'PDOCrE',
                                          'FICrE',
                                          'TotalCrO',
                                          'PDOCrO',
                                          'FICrO',
                                          'TotalCrR',
                                          'FICrR',
                                          'PDOCrR'
                                        )
    )
    )

    MapOutputDataFinalDTtemp <-  distinct(MapOutputDataFinalDTtemp, ID,  .keep_all =
    TRUE)
    MapOutputDataFinalDTtemp$SpdAve                                                   <-
    format(round(MapOutputDataFinalDTtemp$SpdAve,2),nsmall = 2)
    MapOutputDataFinalDTtemp$SpdStd                                                   <-
    format(round(MapOutputDataFinalDTtemp$SpdStd,2),nsmall = 2)
```

```
    MapOutputDataFinalDTtemp$Precip                                                   <-
    format(round(MapOutputDataFinalDTtemp$Precip,2),nsmall = 2)
    MapOutputDataFinalDTtemp$TotalCrE                                                 <-
    format(round(MapOutputDataFinalDTtemp$TotalCrE,2),nsmall = 2)
    MapOutputDataFinalDTtemp$PDOCrE                                                   <-
    format(round(MapOutputDataFinalDTtemp$PDOCrE,2), nsmall = 2)
    MapOutputDataFinalDTtemp$FICrE                                                    <-
    format(round(MapOutputDataFinalDTtemp$FICrE,2),nsmall = 2)
    MapOutputDataFinalDTtemp$TotalPre                                                 <-
    format(round(MapOutputDataFinalDTtemp$TotalPre,2),nsmall = 2)
    MapOutputDataFinalDTtemp$PDOPre                                                   <-
    format(round(MapOutputDataFinalDTtemp$PDOPre,2), nsmall = 2)
    MapOutputDataFinalDTtemp$FIPre                                                    <-
    format(round(MapOutputDataFinalDTtemp$FIPre,2),nsmall = 2)
    MapOutputDataFinalDTtemp$Length                                                   <-
    format(round(MapOutputDataFinalDTtemp$Length,3),nsmall = 3)
    MapOutputDataFinalDTtemp$TotalCrR                                                 <-
    format(round(MapOutputDataFinalDTtemp$TotalCrR,2), nsmall = 2)
    MapOutputDataFinalDTtemp$PDOCrR                                                   <-
    format(round(MapOutputDataFinalDTtemp$PDOCrR,2),nsmall = 2)
    MapOutputDataFinalDTtemp$FICrR                                                    <-
    format(round(MapOutputDataFinalDTtemp$FICrR,2),nsmall = 2)
  } else {
    MapOutputDataFinalDTtemp <- cbind('None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None',
                               'None'
```

```
    )

  }

  MapOutputDataFinalDT <- datatable(MapOutputDataFinalDTtemp,
                          class = 'cell-border stripe',rownames = FALSE,options =
    list(searchHighlight = TRUE)
  )

  output$outputDT = DT::renderDT(MapOutputDataFinalDT, options = list(lengthChange =
    FALSE))


  if (nrow(displaydata) > 0){
    outputDTdowload <- select(displaydata,
                      -c('geometry')
    )
    outputDTdowload <- distinct(outputDTdowload, ID, .keep_all = TRUE)
    outputDTdowload <- outputDTdowload %>% relocate(District, .after = District_Code)
    %>% relocate(County, .after = County_Code)
  }else{
    outputDTdowload <- cbind('None')

  }

  output$downloadData <- downloadHandler(
    filename = function() {gsub(" ","",paste(input$DistrictInput,"_",
                                  input$CountyInput,"_",
                                  switch(input$FacilityInput,
                                         "All Facilities"= "All",
                                         "Rural Mulitlane Divided" = "RMD",
                                         "Rural Mulitlane Undivided" = "RMU",
                                         "Rural Two Lane" = "R2",
                                         "Rural Interstate" = "RI"
                                  ),
                                  input$YearInput,".csv"))},
    content = function(file) {
      write.csv(outputDTdowload,file, row.names=FALSE)
    })
  output$downloadDefination <- downloadHandler(
    filename = function() {
      paste0("Data_Dictionary", ".xlsx")
    },
    content = function(file) {
      file.copy("data_dictionary.xlsx",file)
    }
```

```
    )
  })

}
shinyApp(ui, server)
```

APPENDIX B: VALUE OF RESEARCH

OVERVIEW

The research team conducted a value of research (VOR) analysis of TxDOT Research Project 0-7051 to produce an estimate of the benefit that the project will likely yield for TxDOT. The temporal scope for this analysis is an 11-year period (labeled as years 1–11), starting with the beginning of the 2-year project. The value of the project is described in terms of net present value (NPV) and cost-benefit ratio (CBR), which are computed using economic discounting formulas.

The primary objective of TxDOT Research Project 0-7051 is to improve safety evaluation of rural roadways by incorporating operating speed and precipitation in the modeling framework. The project aims at quantifying the safety (in terms of the precision of crash frequency) benefits that can be obtained by considering updated SPFs with inclusion of operational measures such as operating speed and precipitation. The research team focused the VOR analysis on the safety benefits of the conversions and the resulting cost savings that can be obtained by improving this knowledge.

METHODOLOGY

The research team used a VOR template provided by TxDOT to compute the NPV and CBR measures. The template requires the following items:

- Project budget: $300,000 ($129,947 in year 1 + $153,199 in year 2 + $16,854 in year 3).
- Project duration: 2.17 years.
- Expected value duration: 11 years (convention chosen by TxDOT).
- Discount rate: 3 percent (default value assumed by TxDOT).
- Expected value per year: $325,750.

The project's expected value per year is estimated based on savings obtained from reduced crashes. The analysis method is described in the following sections.

Concept

To conduct the VOR analysis, the following steps were taken:

1. Compare the precision of crash frequencies by Texas WB SPFs and 0-7051 SPFs.
2. Determine the reduced crash frequency by severity by comparing the expected crash outcomes from Texas WB SPFs and 0-7051 SPFs.
3. Apply the procedure to estimate the expected value of the research.

Input Data

The VOR analysis was conducted by randomly selecting 1,000 miles of rural two-lane roadways. The SPFs developed for 0-7051 provide better precision than Texas WB SPFs. The calculated benefit for 8 years is $2,606,000.

Crash Cost

The research team derived crash severity distribution proportions from the sample considered in TxDOT Research Project 0-7051. These proportions are as follows:

- K: 3.14 percent.
- A: 6.92 percent.
- B: 14.15 percent.
- C: 16.35 percent.
- PDO: 59.43 percent.

To estimate the costs of crashes on rural four-lane undivided highways, the research team chose two sources:

  - First, the research team used crash costs from TxDOT's Highway Safety Improvement Program guideline.[9] The crash value is $3.7 million for K (fatal) and A (incapacitating injury) crashes. The B (non-incapacitating injury) crash value is $520,000.
- Second, the National Safety Council values of $155,000 and $51,000 for C (possible injury) and O (no injury) crashes, respectively, were used.

Cost

The research team used an annual maintenance cost of $0 for analysis based on the assumption that TxDOT would provide the same amount of periodic maintenance and monitoring for the reconfigured sites as in the existing condition.

RESULTS

The research team conducted the VOR analysis using the SRPW program and obtained an annual VOR estimate of $325,750. This value represents the benefit that can be obtained if the results of the research project are used to analyze 1000 miles of rural two-lane roadways.

Figure 45 summarizes the VOR calculations. The payback period for Research Project 0-7051 is 0.92 years, and the CBR is 6.02.

---

[9] Texas Department of Department, 2020. *Highway Safety Improvement Guidelines.* https://ftp.txdot.gov/pub/txdot-info/trf/hsip/hsip-guidance-june-2020.pdf.

The findings shown in Figure 45 are as follows:

- The benefits included in the VOR calculations include only those incurred by TxDOT. In reality, other agencies (e.g., local and county agencies within Texas and other state DOTs) will be able to implement and benefit from the published findings from the project.
- The estimated benefits included only crash reduction, which will occur when the safety prediction model is applied to evaluate the precision of safety evaluation. TxDOT will likely incur additional benefits that are more difficult to quantify.
- The VOR analysis focused on rural two-lane roadways. In reality, other rural roadway facilities may also realize similar benefits from the application of these project results. The estimated VOR, NPV, and CBR would increase if these sites were included in the analysis.

Project #	0-7051			
Project Name:	Develop a Real-Time Decision Support Tool for Rural Roadway Safety Improvements			
Agency:	TTI	**Project Budget**	$	300,000
Project Duration (Yrs)	2.17	**Exp. Value (per Yr)**	$	325,750
Expected Value Duration (Yrs)	10	**Discount Rate**		3%
Economic Value				
Total Savings:	$2,306,000	**Net Present Value (NPV):**	$	1,806,631
Payback Period (Yrs):	0.92	**Cost Benefit Ratio (CBR):**		6.02

Years	Expected Value
1	-$129,947
2	-$153,199
3	-$16,854
4	$325,750
5	$325,750
6	$325,750
7	$325,750
8	$325,750
9	$325,750
10	$325,750
11	$325,750

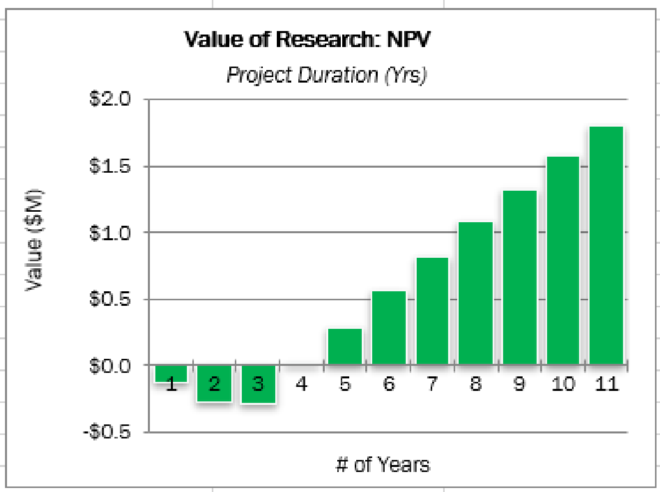

Figure 45. VOR Analysis Results.